The stochastic multi-path traveling salesman problem with dependent random travel costs.

Submitted to *Transportation Science*
manuscript TS-2019-0022

The stochastic multi-path Traveling Salesman Problem with dependent random travel costs

Edoardo Fadda[1,2], Lohic Fotio Tiotsop[1], Daniele Manerba[2,3] *, Roberto Tadei[1]
[1]Department of Control and Computer Engineering, Politecnico di Torino - corso Duca degli Abruzzi 24, Torino (Italy)
[2]ICT for City Logistics and Enterprises Lab, Politecnico di Torino - corso Duca degli Abruzzi 24, Torino (Italy)
[3]Department of Information Engineering, University of Brescia - via Branze 38, Brescia (Italy)

The objective of the stochastic multi-path Traveling Salesman Problem is to determine the expected minimum-cost Hamiltonian tour in a network characterized by the presence of different paths between each pair of nodes, given that a random travel cost with an unknown probability distribution is associated with each of these paths. Previous works have proved that this problem can be deterministically approximated when the path travel costs are independent and identically distributed. Such an approximation has been demonstrated to be of acceptable quality in terms of the estimation of an optimal solution compared to consolidated approaches such as stochastic programming with recourse, completely overcoming the computational burden of solving enormous programs exacerbated by the number of scenarios considered. Nevertheless, the hypothesis regarding the independence among the path travel costs does not hold when considering real settings. It is well known, in fact, that traffic congestion influences travel costs and creates dependence among them. In this paper, we demonstrate that the independence assumption can be relaxed and a deterministic approximation of the stochastic multi-path Traveling Salesman Problem can be derived by assuming just asymptotically independent travel costs. We also demonstrate that this deterministic approximation has strong operational implications because it allows the consideration of realistic traffic models. Computational tests on extensive sets of random and realistic instances indicate the excellent efficiency and accuracy of the deterministic approximation.

Key words: TSP, stochastic travel costs, asymptotic independence, deterministic approximation.
History:

---

* Corresponding author: daniele.manerba@polito.it

Other e-mail addresses: edoardo.fadda@polito.it (E. Fadda); lohic.fotiotiotsop@polito.it (L. Fotio Tiotsop); roberto.tadei@polito.it (R. Tadei).

1. Introduction

Owing to its theoretical interest and wide applicability, the Traveling Salesman Problem (TSP) is undoubtedly one of the most studied problems in combinatorial optimization. Several logistics and routing problems, as well as other combinatorial problems (such as *job scheduling*), can be modeled as a TSP or contain the TSP as a critical sub-problem. In the classical TSP version, travel costs are deterministically known a priori and associated with arcs representing a unique path to travel from one node to another (in general, the shortest path between the two nodes).

In this paper, we address a TSP in real routing applications where uncertainty deriving from numerous factors including accidents, traffic congestion, and bad weather conditions can strongly influence the travel costs. Furthermore, a decision maker can frequently choose among different travel paths between the same pair of nodes, e.g., a different combination of streets or transportation modes (as commonly occurs in multi-modal transportation networks). The stochastic multi-path TSP ($smpTSP$) introduced in Tadei, Perboli, and Perfetti (2017) addresses these two features simultaneously. The authors demonstrate that, despite the increase in the complexity of the problem, the savings obtained by explicitly incorporating stochasticity into a multi-path TSP strongly justifies the effort. Unfortunately, the time required to solve the $smpTSP$ increases exponentially with the problem size and number of considered scenarios. Hence, a deterministic approximation has been developed to solve large instances. In Maggioni, Perboli, and Tadei (2014), the same approximation is applied to a real City Logistics application, where routing instances from the literature have been extended to incorporate real data collected from a sensor network (Fadda, Perboli, and Tadei 2018, Tadei et al. 2016).

The mentioned approximation, as with several other methods existing in the literature for similar problems (see Section 2), assumes that travel cost variations are independent and identically distributed (i.i.d.). However, it is well known that in real settings, network travel costs are strongly not independent. First, in traffic conditions, it is common that time delays on a link can be propagated through the preceding and consecutive links. Furthermore, under network congestion, users are likely to select the less crowded path between two nodes, thus creating dependence on the times influencing users traveling on different paths. For example, if a high-speed road is congested because of an accident, then users will choose and tend to congest a sufficiently close regional alternate route going in the same direction.

For this reason, in this paper, we investigate an $smpTSP$ variant called stochastic multi-path TSP with dependent travel costs ($smpTSP_{dc}$), where travel costs are assumed to be identically distributed, yet just asymptotically independent. In particular, we propose both a stochastic programming (SP) and asymptotic approximation approach for solving the $smpTSP_{dc}$. The relaxation of the travel cost independence assumption to the asymptotic independence poses challenges that must be overcome. First, there does not appear to be any asymptotic approximation approach in the literature that does not explicitly require cost independence to be proven; therefore, new theoretical results are required. Secondly, it is necessary to demonstrate how the new assumptions are consistent with realistic traffic networks.

The contribution of this work is as follows.

1. To the best of our knowledge, we address, for the first time, a TSP problem considering both multi-path networks and the dependence of travel cost variations. As previously mentioned, only the asymptotic independence is assumed for the random variations.

2. We formally prove that a deterministic approximation can be derived by using the asymptotic theory of extreme values (Galambos 1978). In fact, in this work, we could further relax another assumption on the shape of the probability distribution of the cost variations required for the approximation provided in Tadei, Perboli, and Perfetti (2017). This means that the new approximation holds for a greater number of theoretical distributions.

3. We demonstrate that the asymptotic independence assumption is not overly restrictive in real applications when addressing realistic traffic models based on the well-known *Wardrop's traffic equilibrium* principle (Wardrop 1952).

4. The proposed deterministic approximation provides a powerful and reasonably accurate decision support tool to address the $smpTSP_{dc}$. Its quality and efficiency are assessed through an extensive set of computational experiments, where both random networks and realistic traffic models are considered.

Finally, we stress the fact that our new theoretical results are not actually problem-dependent and, therefore, can be generalized and applied to solve other similar optimization problems under uncertainty.

The remainder of this paper is organized as follows. In Section 2, we briefly review the literature available on stochastic TSP and highlight several common assumptions. In Section 3, we present the mathematical model of the problem. In Section 4, we prove our main results enabling us to develop a deterministic approximation of the $smpTSP_{dc}$. In Section 5, we discuss the existing links between Wardrop's traffic equilibrium principle and the asymptotic independence of the travel costs of a network. In Section 6, we present the performance of the proposed approximation using several numerical examples. Section 7 provides the conclusions of our work and identifies possible future investigations.

2. Literature review

The TSP is one of the most well-known $\mathcal{NP}$-hard combinatorial optimization problems. It appears in numerous practical applications, either directly or as a sub-problem. Although many excellent books devoted to the TSP have been published in the past (Lawler et al. 1985, Reinelt 1994, Gutin and Punnen 2002), the problem continues to attract interest among researchers. Consequently, several different generalizations or versions of the problem have appeared in the literature. Vehicle Routing (Toth and Vigo 2014), Orienteering (Vansteenwegen, Souffriau, and Oudheusden 2011, Hanafi, Mansini, and Zanotti 2020), and Traveling Purchaser (Manerba, Mansini, and Riera-Ledesma 2017) problems all belong to this broad class. All of these routing problems have also been addressed in their stochastic counterpart (see, e.g., Kenyon and Morton 2003, Campbell, Gendreau, and Thomas 2011, Evers et al. 2014, Verbeeck, Vansteenwegen, and Aghezzaf 2016, Beraldi et al. 2017).

The explicit consideration of stochasticity in a problem tends to make it more realistic and, therefore, is a fundamental feature for those routing problems that embed the optimization of low-level details, such as the minimization of pollution (*green* routing). Several papers have appeared in this context (see, e.g., Ehmke, Campbell, and Thomas 2016, Huang et al. 2017, or Hwang and Ouyang 2015). Similar to our work, Ehmke, Campbell, and Thomas (2016) approximate the stochastic problem using a deterministic program. However, their approximation is based on the pre-calculation of complex expectations considering the load of the vehicles, whereas we propose an asymptotic approximation.

Of particular importance is the work by Kirkpatrick and Toulouse (1985), which introduces the stochastic version of the TSP. A TSP model is said to be stochastic if at least one of its components is assumed to be a random variable. The most traditional approach adopted to solve a stochastic TSP is to assume that all the random variables considered in the problem are i.i.d. with a given distribution. For example, in Carraway, Morin, and Moskowit (1989), Kao (1978), Sniedovich (1981), Huang et al. (2019), the authors study a TSP with independent and normally distributed arc costs. However, the restrictive assumptions of these problems are not sufficient to ensure that deterministic methods function correctly in the stochastic setting, as is the case for the Shortest Path Problem under an exponential probability distribution of the costs (see Eiger, Mirchandani, and Soroush 1985). In other papers (e.g., Wästlund 2010, Mezard and Parisi 1986), the authors approach stochastic TSP using statistical mechanics tools such as the *mean field* approximation or *replica and cavity* methods. In all previous works, the arc costs are considered to be i.i.d. uniformly or exponentially. Other important studies consider variants of the stochastic TSP problem. For example, in Campbell and Thomas (2008), the authors present two recourse problems and one chance-constrained model formalizing a stochastic TSP where there is a deadline associated with each node.

As can be expected, the results derived in all of the aforementioned papers are strongly related to the properties of the underlying distribution. However, in real applications, travel costs are determined by complex mechanisms and thus a precise derivation of the distribution that describes the variations of the arc costs is a difficult, or even impossible, task. Nevertheless, in the literature, there are papers where the authors overcome this problem. For example, in Toriello, William, and Poremba (2013), the authors present and study a dynamic stochastic model of the TSP where the realizations of the random cost vector connecting a single node to the others is known only when the salesman is about to leave that particular node. They demonstrate that, regardless of the distribution, if the costs are assumed to be independent with known expected values and supports, the problem can be formulated as a dynamic programming problem solvable by an approximation through a linear programming (LP) model. Another paper that considers a wide class of distributions is Tadei, Perboli, and Perfetti (2017). In this work, the authors prove that if the random costs are i.i.d. according to a probability distribution belonging to the Gumbel distribution domain of attraction, it is possible to derive an asymptotic approximation of the expected minimum Hamiltonian tour using the extreme value theory.

Despite the fact that the above papers propose approaches that allow the consideration of different types of distributions, the i.i.d. assumption on the random arc costs makes them inapplicable in many real situations. Recently, a considerable number of papers have studied both the spatial and temporal correlation among travel times in real-life road networks (Fan, Kalaba, and Moore 2005, Samaranayake, Blandin, and Bayen 2012, Chen et al. 2014). They found that real networks have strongly dependent arc costs. This occurs, for example, when vehicles are prone to delays due to rush hours, road works, accidents, or in general, when the traffic is congested. Unfortunately, the vehicle routing literature continues to lack consideration of both stochastic and dependent costs. In Letchford and Nasiri (2015), however, the authors do study a Steiner TSP with stochastic correlated costs and find a Pareto frontier through integer programming techniques.

In this work, we consider stochastic and dependent costs and demonstrate that an asymptotic independence among random travel costs is sufficient to derive an effective deterministic approximation and to justify the use of realistic traffic models. In particular, the mathematical description of the flow principles in real traffic networks is an active and demanding field of research. Because the topic is not central in our discussion, we just recall here one of the most well-known results in the field, i.e., *Wardrop's first traffic equilibrium principle* (Wardrop 1952, Wardrop and Whitehead 1952). This principle states that, at equilibrium, no single driver can unilaterally reduce his/her travel cost by shifting to another route. That is, the traffic tends to be distributed such that all alternative paths between two nodes indicate the same cost. Since its introduction in the context of road traffic research, transportation planners have developed Wardrop's equilibrium-based models to predict commuter decisions in real-life networks. Certain models have been and continue to be used today to evaluate alternative future scenarios and plan future actions on the networks. Other models describe how the traffic flow increases with respect to the traffic conditions, such as the *U.S. Bureau of Public Roads* (BPR) function (U.S. Department of Commerce, Bureau of Public Roads 1964). We consider this last model to demonstrate how random dependencies influence the traffic network.

3. Problem definition and mathematical formulation

Let us consider:

- $G = (\mathcal{I}, \mathcal{E})$: a directed complete graph where $\mathcal{I}$ is the set of nodes and $\mathcal{E} = \{(i,j) \mid i,j \in \mathcal{I}, i \neq j\}$ is the set of arcs;
- $\mathcal{P}_{ij}$: a set of paths for arc $(i,j) \in \mathcal{E}$;
- c_{ij}^p: a deterministic travel cost for arc $(i,j) \in \mathcal{E}$ on path $p \in \mathcal{P}_{ij}$;
- Θ_{ij}^p: a random variable, associated to the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, representing the stochastic variation of the deterministic travel cost c_{ij}^p for arc $(i,j) \in \mathcal{E}$ on path $p \in \mathcal{P}_{ij}$.

Therefore, for each arc $(i,j) \in \mathcal{E}$ on path $p \in \mathcal{P}_{ij}$, the total cost becomes

$$C_{ij}^p := c_{ij}^p + \Theta_{ij}^p. \tag{1}$$

The objective of the *smpTSP* is to determine the expected minimum-cost Hamiltonian tour in G and identify what path must be used to travel between each pair of nodes in that tour.

In the following, we propose a two-stage SP formulation with recourse for the *smpTSP*. The first stage composes the Hamiltonian tour and the second stage selects what path to use between each pair of nodes (e.g., what transport mode must be used) in the Hamiltonian tour determined in the first stage. Let us consider a first-stage binary variable y_{ij}, $(i,j) \in \mathcal{E}$, taking a value of one if node j is visited directly after node i, and zero otherwise. The *smpTSP* is defined as follows

$$\min_{y} \; \mathbb{E}_{\mathbf{\Theta}}[h(\mathbf{y}, \mathbf{\Theta})] \tag{2}$$

subject to

$$\sum_{(i,j)\in\mathcal{E}} y_{ij} = 1, \quad i \in \mathcal{I}, \tag{3}$$

$$\sum_{(j,i)\in\mathcal{E}} y_{ij} = 1, \quad j \in \mathcal{I}, \tag{4}$$

$$\sum_{i\in U} \sum_{j\notin U} y_{ij} \geq 1, \quad U \neq \emptyset, U \subset \mathcal{I}, \tag{5}$$

$$y_{ij} \in \{0,1\}, \quad (i,j) \in \mathcal{E}. \tag{6}$$

The term $h(\mathbf{y},\boldsymbol{\Theta})$, dependent on the solution vector $\mathbf{y}$ and, through the costs C_{ij}^p, on the random vector $\boldsymbol{\Theta}$ (whose components are the random variables Θ_{ij}^p), is the following problem

$$h(\mathbf{y},\boldsymbol{\Theta}) \;:=\; \min_x \sum_{(i,j)\in\mathcal{E}} \sum_{p\in\mathcal{P}_{ij}} C_{ij}^p x_{ij}^p \tag{7}$$

subject to

$$\sum_{p\in\mathcal{P}_{ij}} x_{ij}^p = y_{ij}, \quad (i,j)\in\mathcal{E}, \tag{8}$$

$$x_{ij}^p \in \{0,1\}, \quad (i,j)\in\mathcal{E}, p\in\mathcal{P}_{ij}, \tag{9}$$

where x_{ij}^p, $(i,j)\in\mathcal{E}$ and $p\in\mathcal{P}_{ij}$, is a binary variable taking value one if path p is selected for traveling across arc (i,j), and zero otherwise.

The objective function (2), strictly depending on (7), minimizes the expected total travel cost. Constraints (3) and (4) are the assignment constraints ensuring that each node is visited once and only once and connectivity constraint (5) prevents the formation of sub-tours in the solution. Constraints (8) link variables x_{ij}^p and y_{ij} to each other. In particular, when arc (i,j) is not selected by the first stage ($y_{ij}=0$), no path belonging to that arc (i,j) can be used. Conversely, when arc (i,j) is selected by the first stage ($y_{ij}=1$), then one and only one path must be selected for that arc. Finally, (6) and (9) are binary conditions on the variables.

Please note that the above formulation is somewhat different from the one proposed in Tadei, Perboli, and Perfetti (2017), which included a nonlinear objective function. This new formulation, rather, leads to an integer linear programming formulation of a deterministic equivalent problem (see Section 3.1).

3.1. Deterministic Equivalent Problem (DEP) formulation

The stochastic model (2)–(9) is practically impossible to solve because of the difficulty of calculating the expected value in the objective function as a multi-dimensional integral, which cannot be solved analytically. A common SP approach to overcome this problem (see, e.g., Wallace and Ziemba 2005) is to discretize the probability distribution of the random variables by creating a finite number of possible realizations (called *scenarios*), and then to approximate the stochastic model with a deterministic model, in our case called DEP.

Hence, in the following, we consider a set $\mathcal{S}$ of possible scenarios. Each scenario $s\in\mathcal{S}$, occurring with a probability π^s, is associated with a random cost variation Θ_{ij}^{ps} for each arc $(i,j)\in\mathcal{E}$ and path $p\in\mathcal{P}_{ij}$. Because π^s is a probability, we have $\sum_{s\in\mathcal{S}}\pi^s = 1$. The DEP of the *smpTSP* can be stated as follows

$$\min_{y,x} \; \sum_{s\in\mathcal{S}} \pi^s \sum_{(i,j)\in\mathcal{E}} \sum_{p\in\mathcal{P}_{ij}} x_{ij}^{ps}\; C_{ij}^{ps} \tag{10}$$

subject to

$$\sum_{(i,j)\in\mathcal{E}} y_{ij} = 1, \quad i \in \mathcal{I}, \tag{11}$$

$$\sum_{(j,i)\in\mathcal{E}} y_{ij} = 1, \quad j \in \mathcal{I}, \tag{12}$$

$$\sum_{i\in U}\sum_{j\notin U} y_{ij} \geq 1, \quad U \neq \emptyset,\ U \subset \mathcal{I}, \tag{13}$$

$$\sum_{p\in\mathcal{P}_{ij}} x_{ij}^{ps} = y_{ij}, \quad (i,j)\in\mathcal{E}, s\in\mathcal{S}, \tag{14}$$

$$x_{ij}^{ps} \in \{0,1\}, \quad (i,j)\in\mathcal{E}, p\in\mathcal{P}_{ij}, s\in\mathcal{S}, \tag{15}$$

$$y_{ij} \in \{0,1\}, \quad (i,j)\in\mathcal{E}, \tag{16}$$

where $C_{ij}^{ps} := c_{ij}^{p} + \Theta_{ij}^{ps}$ for each arc $(i,j)\in\mathcal{E}$, path $p\in\mathcal{P}_{ij}$, and scenario $s\in\mathcal{S}$, and x_{ij}^{ps} is a binary variable taking value the one if path $p\in\mathcal{P}_{ij}$ is selected for traveling across arc $(i,j)\in\mathcal{E}$ in scenario $s\in\mathcal{S}$, and zero otherwise.

It is worth noting that model (10)–(16), although deterministic, has severe drawbacks. First, its complexity increases dramatically with the size of $\mathcal{S}$ and, therefore, determining an optimal solution by considering a reasonable number of scenarios can be computationally difficult. Secondly, to create the scenario set, it is necessary to have precise knowledge regarding the distribution of all the random variables involved. The proposed approach overcomes both of these drawbacks. In fact, the complexity of the deterministic model resulting from our approximation presented in Section 4.2 is not influenced by the number of scenarios, and the knowledge of the random variables distribution is not necessary.

3.2. The *smpTSP* with dependent travel costs (*smpTSP*$_{dc}$)

As highlighted in Section 2, the *smpTSP* has always been studied assuming that the stochastic travel costs are i.i.d. In this paper, however, we specifically address a generalization of the *smpTSP* where the random variables Θ_{ij}^{p} representing travel costs have an unknown joint probability distribution and show inter-dependencies. We just assume them to be *asymptotically independent in their left tail*.

DEFINITION 1. Let X_1 and X_2 be two random variables. They are said to be *asymptotically independent in their left tail* if

$$\lim_{r\to-\infty}\big(\mathbb{P}[X_1 \leq r \mid X_2 \leq r] - \mathbb{P}[X_1 \leq r]\big) = 0. \tag{17}$$

That is, assuming two random variables to be asymptotically independent in their left tail means that the probability of having a large variation of one variable in its left tail is not influenced by the knowledge of a large variation in the left tail of the other variable.

In this particular case, we require asymptotic independence in the left tail of the random cost distributions because the approximation approach we develop follows the optimization perspective of the problem (10)–(16). Therefore, because our problem aims at minimizing travel cost, we are only interested in the behavior of the cost random variables in their left tail.

Relaxing the strong independence assumption allows us to address the traffic congestion effects in real networks, where the travel cost independence assumption is frequently unrealistic. We call this problem variant the $smpTSP$ with dependent travel costs ($smpTSP_{dc}$).

4. Deterministic approximation of the stochastic problem

To develop the deterministic approximation presented in this section, we consider the $smpTSP_{dc}$ as a discrete choice model where the decision maker selects the best alternative among a finite set of mutually exclusive ones, i.e., the best path to move from node i to node j. The approximation functions in two main steps; the first, where it is possible to derive how the costs of the best alternatives are asymptotically distributed, and the second, where an estimator for the travel cost variations can be analytically determined. This approach has been used previously in other application domains such as location, routing, loading, and packing problems (Perboli, Tadei, and Baldi 2012, Tadei et al. 2012, Perboli, Tadei, and Gobbato 2014). However, the independence of the stochastic variables has never previously been relaxed.

To set our approximation, we adopt an optimistic view (i.e., guided by the objective function of the $smpTSP_{dc}$) and relax the problem by assuming that we can choose among all scenarios the one that minimizes the random travel cost variations. More precisely, we define $\tilde{\Theta}^p_{ij}$ as the minimum random travel cost variations Θ^{ps}_{ij} among all scenarios $s \in \mathcal{S}$, i.e.,

$$\tilde{\Theta}^p_{ij} := \min_{s\in\mathcal{S}} \Theta^{ps}_{ij}, \quad (i,j)\in\mathcal{E},\ p\in\mathcal{P}_{ij}. \tag{18}$$

We also define $F^p_{ij}(x)$ as the survival function of $\tilde{\Theta}^p_{ij}$, i.e., $F^p_{ij}(x) = \mathbb{P}[\tilde{\Theta}^p_{ij} > x]$.

Remark 1. Because Θ^p_{ij} are asymptotically independent in their left tail, then $\tilde{\Theta}^p_{ij}$ are also asymptotically independent in their left tail.

Now, it is apparent that among all the alternative paths from node i to node j, the path indicating the least cost will be selected in the optimal solution of the $smpTSP_{dc}$. For simplicity and without loss of generality, we assume such path to be unique. We define $\tilde{C}_{ij}$ as the cost of such optimal path for traveling from node i to node j, i.e.,

$$\tilde{C}_{ij} := \min_{p\in\mathcal{P}_{ij}} \left(c^p_{ij} + \tilde{\Theta}^p_{ij}\right), \quad (i,j)\in\mathcal{E}. \tag{19}$$

Note that $\tilde{C}_{ij}$ remains a random variable because of its dependence on $\tilde{\Theta}^p_{ij}$. We call its survival function

$$G_{ij}(x) = \mathbb{P}[\tilde{C}_{ij} > x]. \tag{20}$$

Clearly, a variable x^p_{ij} will take the value one in an optimal solution of the $smpTSP_{dc}$ if and only if p is the optimal path from i to j and, therefore, variables x^p_{ij} can be surrogated by the already existing variables y_{ij}. Hence, because of the linearity of the expected value, problem (2)–(9) becomes

$$\min_y \sum_{(i,j)\in\mathcal{E}} \mathbb{E}_{\boldsymbol{\Theta}} \left[\tilde{C}_{ij}\right] y_{ij} \tag{21}$$

subject to constraints (3)–(6).

Unfortunately, the distribution of $\tilde{C}_{ij}$ is unknown because the distribution of $\tilde{\Theta}^p_{ij}$ is unknown. Thus, the expected value in (21) is not solvable. We provide in the following section an asymptotic approximation of the distribution of $\tilde{C}_{ij}$, or, equivalently, of its survival function $G_{ij}(x)$.

4.1. Asymptotic approximation of $G_{ij}(x)$

First, note that by subtracting a constant α from all random cost variations Θ^{ps}_{ij}, the optimal solution of problem (2)–(9) does not change. In fact, let us denote the original objective function by

$$f_0 := \mathbb{E}_{\boldsymbol{\Theta}} \left[\sum_{(i,j)\in\mathcal{E}} \sum_{p\in\mathcal{P}_{ij}} x^p_{ij}\left(c^p_{ij} + \Theta^p_{ij}\right)\right],$$

and the same objective function after the normalization of the cost variations by

$$f_\alpha := \mathbb{E}_{\boldsymbol{\Theta}} \left[\sum_{(i,j)\in\mathcal{E}} \sum_{p\in\mathcal{P}_{ij}} x^p_{ij}\left(c^p_{ij} + \Theta^p_{ij} - \alpha\right)\right].$$

Then, the following condition holds

$$\begin{aligned} f_\alpha &= f_0 - \alpha \sum_{(i,j)\in\mathcal{E}} \sum_{p\in\mathcal{P}_{ij}} x^p_{ij} = \\ &= f_0 - \alpha \sum_{(i,j)\in\mathcal{E}} y_{ij} = \\ &= f_0 - \alpha|\mathcal{E}|. \end{aligned}$$

Hence, we can restate (19) as

$$\tilde{C}_{ij} = \min_{p\in\mathcal{P}_{ij}} (c^p_{ij} + \min_{s\in\mathcal{S}}(\Theta^{ps}_{ij} - \alpha)), \quad (i,j)\in\mathcal{E}. \tag{22}$$

THEOREM 1. *Let us consider any arc $(i,j)\in\mathcal{E}$. If the random cost variations Θ^p_{ij} of each path $p\in\mathcal{P}_{ij}$ are asymptotically independent in their left tail, and if*

$$\lim_{|\mathcal{S}|\to+\infty} (F^p_{ij}(x+\alpha_{|\mathcal{S}|}))^{|\mathcal{S}|} = exp(-e^{\beta x}) \text{ for some real number } \beta > 0 \text{ and for some sequence } \alpha_{|\mathcal{S}|}, \tag{23}$$

then

$$\lim_{|\mathcal{S}|\to+\infty} G_{ij}(x) = \lim_{|\mathcal{S}|\to+\infty} \mathbb{P}[\tilde{C}_{ij} > x] = \lim_{|\mathcal{S}|\to+\infty} \mathbb{P}\left[\min_{p\in\mathcal{P}_{ij}} (c_{ij}^p + \min_{s\in\mathcal{S}} (\Theta_{ij}^{ps} - \alpha_{|\mathcal{S}|})) > x\right] = e^{-A_{ij}e^{\beta x}}, \quad (24)$$

where

$$A_{ij} = \sum_{p\in\mathcal{P}_{ij}} e^{-\beta c_{ij}^p}. \quad (25)$$

Proof. Let $P_{ij} = |\mathcal{P}_{ij}|$ and $F_{ij}(x_1, x_2, \ldots, x_{P_{ij}})$ be the unknown joint survival function of all the P_{ij} random variations associated to the paths connecting node i to node j under any given scenario $s \in \mathcal{S}$, i.e.,

$$F_{ij}(x_1, x_2, \ldots, x_{P_{ij}}) = \mathbb{P}[\bigcap_{p=1,2,\ldots,P_{ij}} \Theta_{ij}^{ps} > x_p]. \quad (26)$$

Using De Morgan's laws and the property of the probability over the union of a finite number of events, we have that

$$\begin{aligned} F_{ij}(x_1, x_2, \ldots, x_{P_{ij}}) &= \mathbb{P}[\bigcap_{p=1,2,\ldots,P_{ij}} \Theta_{ij}^{ps} > x_p] = \\ &= 1 - \mathbb{P}[\bigcup_{p=1,2,\ldots,P_{ij}} \Theta_{ij}^{ps} \le x_p] = \\ &= 1 - \sum_{k=1}^{P_{ij}} (-1)^{k+1} \sum_{\{p^1,p^2,\ldots,p^k\}\in 2^{\mathcal{P}_{ij}}} \mathbb{P}[\Theta_{ij}^{p^1 s} \le x_1, \Theta_{ij}^{p^2 s} \le x_2, \ldots, \Theta_{ij}^{p^k s} \le x_k] \end{aligned} \quad (27)$$

where $2^{\mathcal{P}_{ij}}$ is the power set of $\mathcal{P}_{ij}$, i.e., the set containing all subsets of $\mathcal{P}_{ij}$.

Without loss of generality, let us consider two paths $p^1, p^2 \in \mathcal{P}_{ij}$. From Def. 1, it can be seen that

$$\lim_{\substack{x_1\to-\infty,\\ x_2\to-\infty}} \mathbb{P}[\Theta_{ij}^{p^1 s} \le x_1 | \Theta_{ij}^{p^2 s} \le x_2] = 0, \quad s \in \mathcal{S}, \quad (28)$$

or, equivalently

$$\lim_{\substack{x_1\to-\infty,\\ x_2\to-\infty}} \frac{\mathbb{P}[\Theta_{ij}^{p^1 s} \le x_1, \Theta_{ij}^{p^2 s} \le x_2]}{\mathbb{P}[\Theta_{ij}^{p^2 s} \le x_2]} = 0, \quad s \in \mathcal{S}. \quad (29)$$

Note that (29) can be generalized as follows:

$$\lim_{\substack{x_1\to-\infty,\\ x_2\to-\infty}} \frac{\mathbb{P}[\Theta_{ij}^{p^1 s} \le x_1, \Theta_{ij}^{p^2 s} \le x_2]}{\mathbb{P}[\Theta_{ij}^{p^2 s} \le \min(x_1, x_2)]} = 0, \quad s \in \mathcal{S}. \quad (30)$$

In fact, when $x_1 \to -\infty$ and $x_2 \to -\infty$, it holds that

$$0 \le \frac{\mathbb{P}[\Theta_{ij}^{p^1 s} \le x_1, \Theta_{ij}^{p^2 s} \le x_2]}{\mathbb{P}[\Theta_{ij}^{p^2 s} \le \min(x_1, x_2)]} \le \frac{\mathbb{P}[\Theta_{ij}^{p^1 s} \le \min(x_1, x_2), \Theta_{ij}^{p^2 s} \le \min(x_1, x_2)]}{\mathbb{P}[\Theta_{ij}^{p^2 s} \le \min(x_1, x_2)]} \to 0. \quad (31)$$

The limit in (30) has the following important interpretation: when both x_1 and x_2 tend to $-\infty$, the implication is that $\mathbb{P}[\Theta_{ij}^{p^1 s} \le x_1, \Theta_{ij}^{p^2 s} \le x_2] \to 0$. Hence in (27), if $x_1 \to -\infty, x_2 \to -\infty, \ldots, x_{P_{ij}} \to$

$-\infty$, all probabilities including the intersection of two or more events are negligible. Thus, when $x_1 \to -\infty, x_2 \to -\infty, \dots, x_{P_{ij}} \to -\infty$,

$$F_{ij}(x_1, x_2, \dots, x_{P_{ij}}) \to 1 \quad - \sum_{p=1,2,\dots,P_{ij}} \mathbb{P}[\Theta_{ij}^{ps} \leq x_p], \tag{32}$$

under any scenario $s \in \mathcal{S}$.

Because of (22), $\mathbb{P}[\tilde{C}_{ij} > x]$ can be written as a function of $|S|$ as follows

$$\begin{aligned}
\mathbb{P}[\tilde{C}_{ij} > x] &= \mathbb{P}[\min_{p\in\mathcal{P}_{ij}} (c_{ij}^p + \min_{s\in\mathcal{S}}(\Theta_{ij}^{ps} - \alpha_{|\mathcal{S}|})) > x] = \\
&= \mathbb{P}[\bigcap_{p\in\mathcal{P}_{ij}} (c_{ij}^p + \min_{s\in\mathcal{S}}(\Theta_{ij}^{ps} - \alpha_{|\mathcal{S}|})) > x] = \\
&= \mathbb{P}[\bigcap_{p\in\mathcal{P}_{ij}} (\min_{s\in\mathcal{S}}(\Theta_{ij}^{ps} - \alpha_{|\mathcal{S}|})) > x - c_{ij}^p] = \\
&= \mathbb{P}[\bigcap_{p\in\mathcal{P}_{ij}} \bigcap_{s\in\mathcal{S}} (\Theta_{ij}^{ps} - \alpha_{|\mathcal{S}|}) > x - c_{ij}^p] = \\
&= \mathbb{P}[\bigcap_{p\in\mathcal{P}_{ij}} \bigcap_{s\in\mathcal{S}} \Theta_{ij}^{ps} > x - c_{ij}^p + \alpha_{|\mathcal{S}|}] = \\
&= \mathbb{P}[\bigcap_{s\in\mathcal{S}} \bigcap_{p\in\mathcal{P}_{ij}} \Theta_{ij}^{ps} > x - c_{ij}^p + \alpha_{|\mathcal{S}|}] = \\
&= \prod_{s\in\mathcal{S}} \mathbb{P}[\bigcap_{p\in\mathcal{P}_{ij}} \Theta_{ij}^{ps} > x - c_{ij}^p + \alpha_{|\mathcal{S}|}] = \\
&= [F_{ij}(x - c_{ij}^{p^1} + \alpha_{|\mathcal{S}|}, x - c_{ij}^{p^2} + \alpha_{|\mathcal{S}|}, \dots, x - c_{ij}^{p^{P_{ij}}} + \alpha_{|\mathcal{S}|})]^{|\mathcal{S}|}.
\end{aligned} \tag{33}$$

From the assumption of (23), it can be observed that, when $|\mathcal{S}| \to +\infty$,

$$F_{ij}^p(x + \alpha_{|\mathcal{S}|}) \to 1, \quad p \in \mathcal{P}_{ij}. \tag{34}$$

In fact, if $F_{ij}^p(x + \alpha_{|\mathcal{S}|})$ was bounded by any real number strictly less than one, then $(F_{ij}^p(x + \alpha_{|\mathcal{S}|}))^{|\mathcal{S}|}$ would tend to zero for any real number x and that would contradict (23).

Using (34), one has that $\lim_{|\mathcal{S}|\to\infty}(x + \alpha_{|\mathcal{S}|}) = -\infty, \quad x \in \mathbb{R}$. Thus, under any scenario $s \in \mathcal{S}$

$$\lim_{|\mathcal{S}|\to\infty} (x - c_{ij}^p) + \alpha_{|\mathcal{S}|} = -\infty, \quad p \in \mathcal{P}_{ij}, x \in \mathbb{R}. \tag{35}$$

Because of (32) and (35), when $|\mathcal{S}| \to +\infty$,

$$F_{ij}((x_1 - c_{ij}^{p^1}) + \alpha_{|\mathcal{S}|}, (x_2 - c_{ij}^{p^2}) + \alpha_{|\mathcal{S}|}, \dots, (x_{P_{ij}} - c_{ij}^{p^{P_{ij}}}) + \alpha_{|\mathcal{S}|}) \to 1 - \sum_{p\in\mathcal{P}_{ij}} \mathbb{P}[\Theta_{ij}^{ps} \leq (x_p - c_{ij}^p) + \alpha_{|\mathcal{S}|}]. \tag{36}$$

Hence, using (33) and (36), one obtains

$$\lim_{|\mathcal{S}|\to+\infty} \mathbb{P}[\tilde{C}_{ij} > x] = \lim_{|\mathcal{S}|\to+\infty} e^{|\mathcal{S}| \log(1 - \sum_{p\in\mathcal{P}_{ij}} \mathbb{P}[\Theta_{ij}^{ps} \leq (x - c_{ij}^p) + \alpha_{|\mathcal{S}|}])}. \tag{37}$$

Because (35) implies that $\lim_{|\mathcal{S}|\to+\infty}\mathbb{P}[\Theta_{ij}^{ps}\leq(x_p-c_{ij}^p)+\alpha_{|\mathcal{S}|}]=0, p\in\mathcal{P}_{ij}, s\in\mathcal{S}$, then (37) leads to

$$\begin{aligned}
\lim_{|\mathcal{S}|\to+\infty}\mathbb{P}[\tilde{C}_{ij}>x] &= \lim_{|\mathcal{S}|\to+\infty} e^{-|\mathcal{S}|(\sum_{p\in\mathbb{P}_{ij}}\mathbb{P}[\Theta_{ij}^{ps}\leq(x-c_{ij}^p)+\alpha_{|\mathcal{S}|}])}= \\
&= \lim_{|\mathcal{S}|\to+\infty}(e^{-(\sum_{p\in\mathbb{P}_{ij}}\mathbb{P}[\Theta_{ij}^{ps}\leq(x-c_{ij}^p)+\alpha_{|\mathcal{S}|}])})^{|\mathcal{S}|}= \\
&= \lim_{|\mathcal{S}|\to+\infty}\prod_{p\in\mathcal{P}_{ij}}(e^{-\mathbb{P}[\Theta_{ij}^{ps}\leq(x-c_{ij}^p)+\alpha_{|\mathcal{S}|}]})^{|\mathcal{S}|}= \\
&= \lim_{|\mathcal{S}|\to+\infty}\prod_{p\in\mathcal{P}_{ij}}(1-\mathbb{P}[\Theta_{ij}^{ps}\leq(x-c_{ij}^p)+\alpha_{|\mathcal{S}|}]))^{|\mathcal{S}|}= \\
&= \prod_{p\in\mathcal{P}_{ij}}\lim_{|\mathcal{S}|\to+\infty}(F_{ij}^p((x-c_{ij}^p)+\alpha_{|\mathcal{S}|}))^{|\mathcal{S}|}.
\end{aligned}\tag{38}$$

Now, owing to (23) and (38), it holds that

$$\lim_{|\mathcal{S}|\to\infty}\mathbb{P}[\tilde{C}_{ij}>x]=\prod_{p\in\mathcal{P}_{ij}}\exp(-e^{\beta(x-c_{ij}^p)})=e^{-A_{ij}e^{\beta x}}.\tag{39}$$

This proves the theorem. □

Basically, Theorem 1 states that if the unknown probability distribution of the stochastic cost variations satisfies assumption (23), then the costs asymptotically converge in the distribution to a Gumbel function (double exponential), even if the costs are asymptotically independent. Note that the expression A_{ij} in (25) represents the "*accessibility*" in the sense of Hansen (1959), which is a measure of "visibility" that the decision maker has for each arc (i,j) on the entire set of its alternative paths $\mathcal{P}_{ij}$. In turn, this accessibility depends on a parameter $\beta>0$ that must be calibrated (see Section 6.1) and represents the dispersion of the alternatives in the decision making process. A small value of β means to consider a large number of alternatives to choose from, while a large value means to concentrate the choice just on a small subset of alternatives (when $\beta\to+\infty$, the set of alternatives collapses to the most convenient one).

4.1.1. Applicability of Theorem 1. After having presented, proved, and commented on Theorem 1, we now present a brief discussion highlighting the vast applicability of the results obtained by demonstrating the mildness of assumption (23) made on the structure of the distribution of the random cost variations.

First, note that assumption (23) can be equivalently rewritten as

$$\lim_{|\mathcal{S}|\to+\infty}(F_{ij}^p(\frac{1}{\beta}x+\alpha_{|\mathcal{S}|}))^{|\mathcal{S}|}=\exp(-e^x)$$

and thus, for an accurate calibration of β, it just requires that the distribution belongs to the *domain of attraction* of the Gumbel distribution. This domain consists a large family of distributions including common distributions such as the Normal, Gumbel, Weibull, Logistic, Laplace, Lognormal, and numerous others (i.e., any distribution of the form $1 - e^{-P(x)}$, where $P(x)$ is a polynomial function). Hence, (23) can actually be considered an extremely mild assumption.

Secondly, it is important to notice that (23) is a more general assumption compared to the one used in Tadei, Perboli, and Perfetti (2017) and in similar approaches appearing in the literature, where a more restrictive behavior on the distribution tails is imposed. In particular, the following asymptotic exponential behavior for the left tail of the distribution $F_{ij}^p(x)$ is required

$$\lim_{y\to-\infty} \frac{1-F_{ij}^p(x+y)}{1-F_{ij}^p(y)} = e^{\beta x} \quad \text{for some real number } \beta > 0. \tag{40}$$

We prove in the following proposition that (23) is a more general assumption than (40).

Proposition 1 *Assumption in* (40) *implies assumption in* (23).

Proof. From (23), it holds that $\lim_{|\mathcal{S}|\to+\infty} \alpha_{|\mathcal{S}|} = -\infty$. From (40), we have that

$$\lim_{|\mathcal{S}|\to+\infty} \frac{1-F_{ij}^p(x+\alpha_{|\mathcal{S}|})}{1-F_{ij}^p(\alpha_{|\mathcal{S}|})} = e^{\beta x}. \tag{41}$$

Using (23), (41) becomes

$$\lim_{|\mathcal{S}|\to+\infty} \frac{1-F_{ij}^p(x+\alpha_{|\mathcal{S}|})}{\frac{1}{|\mathcal{S}|}} = e^{\beta x},$$

and thus,

$$\lim_{|\mathcal{S}|\to+\infty} F_{ij}^p(x+\alpha_{|\mathcal{S}|}) = \lim_{|\mathcal{S}|\to+\infty} \left(1 - \frac{e^{\beta x}}{|\mathcal{S}|}\right).$$

Hence,

$$\lim_{|\mathcal{S}|\to+\infty} (F_{ij}^p(x+\alpha_{|\mathcal{S}|}))^{|\mathcal{S}|} = \lim_{|\mathcal{S}|\to+\infty} \left(1 - \frac{e^{\beta x}}{|\mathcal{S}|}\right)^{|\mathcal{S}|} = \exp(-e^{\beta x}).$$

This proves the proposition. □

This means that any distribution satisfying (40) also satisfies our assumption. Conversely, it is easy to verify that several of the previously mentioned distributions satisfying (23), e.g., the Normal and Lognormal, do not follow the behavior expressed in (40).

4.2. Deterministic approximation of $smpTSP_{dc}$

It is worth noting that given the result of Theorem 1, it is also possible to derive for each arc, a Multinomial Logit model for the choice probability of its alternative paths (see Tadei, Perboli, and Manerba 2018). This could lead to a continuous assignment of paths to arcs, and to a possible feasible solution of the proposed $smpTSP_{dc}$ through rounding. However, given the hard feasibility constraints of our problem, we noticed in preliminary experiments that this rounding leads to considerably poor approximated solutions. Therefore, we decided instead to exploit the knowledge of the asymptotic distribution of the random cost variations to calculate their expected value, and thus to achieve an approximated model for the problem based only on deterministic parameters.

More precisely, if $|\mathcal{S}|$ is sufficiently large, the limit obtained in (24) can be used as the survival function of costs $\tilde{C}_{ij}$ and therefore, we can calculate their expected value as follows

$$\mathbb{E}_{\boldsymbol{\Theta}}[\tilde{C}_{ij}] = \int_{-\infty}^{+\infty} x \; d\mathbb{P}[\tilde{C}_{ij} \leq x] = -\int_{-\infty}^{+\infty} x \; d\mathbb{P}[\tilde{C}_{ij} > x] = A_{ij}\beta \int_{-\infty}^{+\infty} x e^{-A_{ij}e^{\beta x}} e^{\beta x} dx. \quad (42)$$

After manipulation (see details in Appendix A), the above expected value becomes

$$\mathbb{E}_{\boldsymbol{\Theta}}[\tilde{C}_{ij}] \approx -\frac{1}{\beta}(\ln(A_{ij}) + \gamma), \quad (43)$$

where $\gamma = -\int_0^{\infty} \ln(t) e^{-t} dt \approx 0.5772$ is the Euler constant.

Now, using (43) and disregarding the constant terms, the following deterministic approximation of the stochastic problem (2)–(9) is obtained

$$\min_{y} \quad -\frac{1}{\beta} \sum_{(i,j)\in\mathcal{E}} y_{ij} \ln(A_{ij}) \quad (44)$$

subject to constraints (3)–(5), and (6).

Note that the deterministic approximation developed allows a reduction of the combinatorial structure of the DEP of the proposed $smpTSP_{dc}$ in formulation (10)–(16) to a common TSP, overcoming the complexity resulting from the presence of both multiple paths and multiple scenarios.

5. Asymptotic independence of the travel costs for realistic networks

In this section, we discuss how requiring the random cost variations to be asymptotically independent in their left tail (see Def. 1) is not in any manner a restrictive condition to exhaustively model the stochastic behavior of a realistic traffic network. Our argument is performed by proving that, under common conditions for realistic networks, travel costs are actually highly correlated, yet remain asymptotically independent in their left tail.

Realistic networks can include peculiarities compared to theoretical networks. In the following, we list and discuss the realistic features of a network that we want to consider and that make reasonable use of our approximation approach.

1. In real routing applications, the support of random travel cost variations is such that the total cost of a path remains strictly positive. Hence, for each $(i,j) \in \mathcal{E}$, $p \in \mathcal{P}_{ij}$, the random cost variation Θ_{ij}^p is lower-bounded by $-c_{ij}^p$.

2. The travel cost on a given path increases when the flow of traffic increases on the same path. More precisely, we assume that there exists an increasing function H_{ij}^{ps} such that

$$C_{ij}^{ps} = H_{ij}^{ps}(Q_{ij}^{ps}), \quad (i,j) \in \mathcal{E}, \; p \in \mathcal{P}_{ij}^U, \; s \in \mathcal{S}, \quad (45)$$

where Q_{ij}^{ps} denotes the actual flow of traffic on path $p \in \mathcal{P}_{ij}^U$ under scenario $s \in \mathcal{S}$. Such an assumption is commonly made in transportation engineering studies (see, e.g., U.S. Department of Commerce, Bureau of Public Roads 1964, where the cost of each path is evaluated by the *Bureau of Public Road* function). Please note that the considered function H_{ij}^{ps} also depends on the particular conditions of a specific scenario s, such as bad weather or accidents.

3. The network dynamics are coherent with the well-known concept of *user equilibrium* based on Wardrop's first principle of route choice (Wardrop 1952, Wardrop and Whitehead 1952). We recall that Wardrop's first principle states: "*The traffic arranges itself in networks such that all used routes between an origin and destination pair have equal and minimum costs, whereas all unused routes demonstrate greater costs*". The principle derives basically from the classical game theory field, and in particular from Nash's equilibrium (see, e.g., Osborne and Rubinstein 1994). This means that the network traffic tends to equilibrium where each traveler cannot obtain savings in travel costs by choosing a different path. Let us denote by $\mathcal{P}_{ij}^U \subseteq \mathcal{P}_{ij}$, the subset of paths used at the equilibrium for arc $(i,j) \in \mathcal{E}$. From Wardrop's first principle, given a scenario s and an arc $(i,j) \in \mathcal{E}$, if $p^1, p^2 \in \mathcal{P}_{ij}^U$, the relative random costs $C_{ij}^{p^1 s}$ and $C_{ij}^{p^2 s}$ are equal, i.e.,

$$\Theta_{ij}^{p^1 s} - \Theta_{ij}^{p^2 s} = c_{ij}^{p^2} - c_{ij}^{p^1} \quad (i,j) \in \mathcal{E},\ p^1, p^2 \in\ \mathcal{P}_{ij}^U,\ s \in \mathcal{S}. \tag{46}$$

Equation (46) demonstrates the high correlation that exists among the random cost variations.

4. The *user equilibrium* state discussed in Wardrop's first principle is virtually never achieved, i.e., the network is essentially always in transition from an equilibrium to another one. This is reasonable as the theoretical equilibrium is continuously perturbed by new users entering or exiting the network (and, therefore, changing the amount of flow on the different paths). Moreover, perturbations also depend on the stochastic nature of the travel costs, which can vary because of accidents or other unforeseen events.

We can now state the following proposition.

Proposition 1. *Let us consider a realistic traffic network $G = (\mathcal{I}, \mathcal{E})$ (defined as in Section 3), i.e., a network where*

- $C_{ij}^{ps} > 0$*, for each arc* $(i,j) \in \mathcal{E}$*, each path* $p \in \mathcal{P}_{ij}$*, and each scenario* $s \in \mathcal{S}$*;*
- *the cost on a path increases as the flow on the same path increases, according to an increasing function* H_{ij}^{ps} *as in* (45)*;*
- *condition* (46) *holds at the user equilibrium (Wardrop's first principle).*

Then, when the network is not at equilibrium, for any pair of different paths $p^1, p^2 \in \mathcal{P}_{ij}$ *of any arc* $(i,j) \in \mathcal{E}$*, the following condition holds*

$$\lim_{r^1 \to -c_{ij}^{p^1}, r^2 \to -c_{ij}^{p^2}} \mathbb{P}[\Theta_{ij}^{p^1} \le r^1 \mid \Theta_{ij}^{p^2} \le r^2] = 0. \tag{47}$$

Proof. See Appendix B.

The above proposition basically states that the travel cost variations Θ_{ij}^p are asymptotically independent in the right-hand side neighborhood of the lower bound $(-c_{ij}^p)$ of their distribution support. Note that this is not strictly equivalent to the definition of asymptotic independence proposed in Definition 1, which considers an unbounded support. However, in our realistic case, it is reasonable to evaluate the asymptotic independence for the largest possible negative variations of the costs, i.e., when the costs tend to zero. Finally, note that the behavior described in Proposition 1 holds when the network is not at equilibrium. This is compatible with the previous observation where realistic traffic networks are virtually never at equilibrium because of perturbations. Thus, the assumption made in Section 3.2 holds and the use of the derived approximation approach (which is based on such an assumption) is suitable for realistic networks as well.

6. Computational experiments

To assess the performance of the proposed approach, we compared the results obtained by the Deterministic Approximation (DA) proposed in Section 4.2 with those of the DEP formulated in (10)–(16) on a large set of benchmark instances. The DA was implemented using *MATLAB* v9.4 and its internal integer solver; the DEP was solved using Cplex v12.7.1 and its C++ Concert Technology. In all experiments, we considered a discretization of the probability space in 100 scenarios ($|\mathcal{S}| = 100$). This choice enforces both *in-sample* and *out-of-sample* stability of the problem (see Kaut et al. 2007). We executed all experiments on an Intel Core I7 2.5 GHz workstation with 16 GB RAM, running the *Windows* 10 operating system.

In Section 6.1, we propose an empirical method to calibrate the parameter β, required to calculate our deterministic approximation. In Section 6.2, we discuss the generation of the instances. In Section 6.3, the computational results are given.

6.1. Calibration of parameter β

As mentioned previously, the DA depends on the parameter β, which must be calibrated. In all experiments, the calibration of β was accomplished as in Tadei, Perboli, and Perfetti (2017). More precisely, let us consider the standard Gumbel distribution $\exp(-e^{-x})$. If an approximation error of 2‰ is accepted, then $\exp(-e^{-x}) = 1 \iff x = 6.08$ and $\exp(-e^{-x}) = 0 \iff x = -1.76$. Hence, if the support of the unknown distribution of the cost variations is $[m,\ M]$, then

$$\beta(m - \zeta) = -1.76, \tag{48}$$

and

$$\beta(M - \zeta) = 6.08, \tag{49}$$

where ζ is the mode of the Gumbel with distribution $\exp(e^{-\beta(x-\zeta)})$. Then, by subtracting (48) from (49), we obtain

$$\beta = \frac{7.84}{M-m}. \tag{50}$$

In our experiments, m is the minimum arc cost of the considered instance and $M := 2\frac{|\mathcal{P}|_{max} f_{det}}{|\mathcal{I}|}$, where $|\mathcal{P}|_{max} = \max_{(i,j)\in\mathcal{E}} |\mathcal{P}_{ij}|$ is the number of paths considered between each pair of nodes. f_{det} is the value of a deterministic TSP obtained by choosing, for each arc, the minimum cost path. In this manner, we hold M to be proportional to the magnitude of the average cost variation in the final solution ($f_{det}/|\mathcal{I}|$), without considering paths with extreme costs.

6.2. Benchmark instances

To better assess the quality and efficiency of our approximation, we generate two different sets of instances. In the first set (presented in Section 6.2.1), we use well-known distributions for modeling the random cost variations; in the second set (presented in Section 6.2.2), we use a more realistic traffic model.

In both sets, nodes are randomly selected from a database providing the position of 16,862 Italian locations (`http://www.math.uwaterloo.ca/tsp/world/it16862.tsp`. Last access: Ovtober 13, 2019) in terms of Cartesian coordinates, and we assume to have the same number $|\mathcal{P}|$ of available paths for each arc of the network, i.e., $|\mathcal{P}| = |\mathcal{P}_{ij}|, (i,j) \in \mathcal{E}$.

6.2.1. Randomly generated instances In this set of random instances, the deterministic costs c_{ij}^p are computed as follows $c_{ij}^p := \tau_p d_{ij}$, where d_{ij} is the Euclidean distance between nodes i and j, and τ_p is randomly sampled in the interval $[1,3]$. We first create five different deterministic instances for each combination of number of nodes ($|\mathcal{I}| = \{50, 100\}$) and number of paths ($|\mathcal{P}| = \{3,4,5\}$), i.e., 30 deterministic instances in total. For each deterministic instance, the random cost variations $\Theta_{ij}^{ps},\ (i,j) \in \mathcal{E}, p \in \mathcal{P}_{ij}, s \in \mathcal{S}$ are generated according to five different marginal distributions (Gumbel, Normal, Logistic, Laplace, and Uniform). The set of randomly generated instances is therefore composed by 150 instances. It is important to note that the Uniform distribution does not satisfy condition (23), required to apply our deterministic approximation. Nevertheless, because in real settings it is not always possible to derive a precise knowledge in terms of the distribution of the observed scenarios, we want to test the approximation, in addition, for distributions that would not fulfill the assumptions of our theory.

To combine the aforementioned marginal distributions into a multivariate distribution, we use the Normal copula (see Nelsen 2006). This simulates the dependency structure of random variations and maintains the asymptotic independence property. In all cases, the support of the distribution of the random cost variation Θ_{ij}^{ps} has been truncated to $[-0.8c_{ij}^p,\ 0.8c_{ij}^p]$ to consider significant changes in costs.

6.2.2. Instances based on a realistic traffic model. In this set of realistic instances, the deterministic cost c_{ij}^p was assumed to be proportional to the time required to travel from node i to j on path p, thus $c_{ij}^p := c_0 t_{ij}^p$, where c_0 is a constant value representing the cost per unit of time and t_{ij}^p is the time required to travel from i to j in normal traffic condition (i.e., without congestion) on path p. For each path $p \in \mathcal{P}_{ij}$, let q_{ij}^p, v_{ij}^p, and l_{ij}^p denote the capacity (i.e., the maximum traffic flow that such path can support), the average speed, and the length of path p, respectively. To model a network containing both high capacity paths (main roads or highways) and low capacity paths (secondary roads), q_{ij}^p are randomly generated from a Uniform distribution with support in $\mathcal{Q}_1 = [70, 100]$ for half of the paths and in $\mathcal{Q}_2 = [20, 50]$ for the other half. The average speed v_{ij}^p on any path $p \in \mathcal{P}_{ij}$ is assigned according to the type of the path (p). More precisely, we set $v_{ij}^p = 100$ if $q_{ij}^p \in \mathcal{Q}_1$ and $v_{ij}^p = 40$ if $q_{ij}^p \in \mathcal{Q}_2$. The length l_{ij}^p is obtained by uniformly sampling a value in the interval $[d_{ij}, 3d_{ij}]$. The time t_{ij}^p is then computed as $t_{ij}^p = l_{ij}^p / v_{ij}^p$.

For each deterministic instance, the random cost variations Θ_{ij}^{ps} are obtained using the following formula

$$\Theta_{ij}^{ps} = c_{ij}^p \left(0.15 \left(\frac{Q_{ij}^{ps}(\lambda_{ij})}{q_{ij}^p} \right)^{3+\lambda_{ij}} + \delta_{ij}^{ps} \right), \tag{51}$$

where $Q_{ij}^{ps}(\lambda_{ij})$ is the actual traffic flow on path $p \in \mathcal{P}_{ij}$ under scenario $s \in \mathcal{S}$ and δ_{ij}^{ps} is an additive term generated from a standard Normal distribution truncated in $[-0.3,\ 0.3]$. δ_{ij}^{ps} models the effect of exogenous events (e.g., weather conditions and road works) that can influence the travel cost, other than the traffic flow. Note that equation (51) is based on the previously presented *Bureau of Public Road* (BPR) model (U.S. Department of Commerce, Bureau of Public Roads 1964), which is widely used and consolidated in transportation engineering. More precisely, the function has been mildly modified by including a positive real parameter λ_{ij} that enables us to modulate the traffic flow on a specific arc (i, j) and therefore, to simulate different traffic conditions.

The flows $Q_{ij}^{ps}(\lambda_{ij})$ are computed as follows. First, under each scenario $s \in \mathcal{S}$, the total flow of traffic Q_{ij}^s is randomly generated in the interval $[0.3 \sum_{p \in P_{ij}} q_{ij}^p,\ 0.7 \sum_{p \in P_{ij}} q_{ij}^p]$ for low-congested networks, and in $[0.7 \sum_{p \in P_{ij}} q_{ij}^p,\ \sum_{p \in P_{ij}} q_{ij}^p]$ for high-congested networks. Then, the flows $Q_{ij}^{ps}(\lambda_{ij})$ for each path $p \in \mathcal{P}_{ij}$ are computed as $Q_{ij}^{ps}(\lambda_{ij}) := Q_{ij}^s \pi_{ij}^p$, where π_{ij}^p denotes the probability of choosing path p among all paths available. This can be calculated according to the following Logit model

$$\pi_{ij}^p = \frac{\exp(-\lambda_{ij}(l_{ij}^p - l_{ij}^o))}{\sum_{p \in \mathcal{P}_{ij}} \exp(-\lambda_{ij}(l_{ij}^p - l_{ij}^o))}, \tag{52}$$

where l_{ij}^o denotes the length of the shortest path p_{ij}^o between node i and j. It is worth noting that the costs obtained on the paths by this simulation are necessarily dependent because $\sum_{p \in \mathcal{P}_{ij}} Q_{ij}^{ps}(\lambda_{ij}) = Q_{ij}^s$. The rationale behind the above formula is the following. Assume that a user must make a choice among all paths linking nodes i and j. It is clear that under normal traffic conditions (no congestion) the user would choose, with high probability, the shortest path p^o_{ij}. Furthermore, the longer a path, the less the probability of being selected. However, under traffic congestion, the shortest path is likely overused. Hence, users attempt to minimize their travel time by evaluating the possibility of using alternative paths and thus, the traffic tends to be redistributed uniformly among all paths (Wardrop principle). These aspects are captured by the Logit model in (52). In fact, when λ_{ij} is close to zero (high-congested network), the probability of choosing path $p \in \mathcal{P}_{ij}$ tends to $\frac{1}{|\mathcal{P}_{ij}|}$ for all paths. For large values of λ_{ij} (low-congested network), the probability tends to zero for all paths $p \in \mathcal{P}_{ij} \setminus p^o_{ij}$ and to one if $p = p^o_{ij}$.

Eventually, we generated 144 instances. More precisely, for each combination of $|\mathcal{I}| = \{50, 100\}$ number of nodes and $|\mathcal{P}| = \{3, 4, 5\}$ number of paths, we create

- ten instances (representing high-congested networks) where λ_{ij} was randomly selected in the interval $[0.1, 2]$, $(i, j) \in \mathcal{E}$;
- ten instances (representing low-congested networks) where λ_{ij} was randomly selected in the interval $[8, 20]$, $(i, j) \in \mathcal{E}$;
- four instances (representing a mixed situation indicating both congested and uncongested paths) where λ_{ij} was randomly selected in the interval $[0.1, 20]$, $(i, j) \in \mathcal{E}$.

6.3. Results and analysis

To quantify the performance of the proposed methodology, we performed the DA and DEP (here used as a benchmark) approaches on each generated instance and calculated the percentage gaps $f\%$ and $t\%$ in terms of objective function value of the returned solution and computational time, i.e.,

$$f\% := 100 \frac{f_{DA} - f_{DEP}}{f_{DEP}},$$

where f_{DA} and f_{DEP} are the values of the objective function of the solution computed using the proposed DA and by solving DEP in (10)–(16) using Cplex within a threshold time of 7200 seconds (2 hours), respectively.

It is important to note that the value of f_{DA} was not obtained directly from the objective function in (44), which only represents an approximation of the overall decision process cost. Rather, a more reasonable evaluation of the real objective function can be obtained, for each instance, through the following steps.

1. Optimally solve the model in (44) and derive, for each arc $(i, j) \in \mathcal{E}$, the values y^*_{ij} of the variables y_{ij} in the optimal solution, which represent the first-stage decisions;
2. For each scenario $s \in \mathcal{S}$, solve DEP (10)–(16) where the y_{ij} variables are fixed to the values y^*_{ij} determined in Step 1, calculating the relative objective function f^s_{DA}. Note that through this variable fixing, the optimization problem actually resorts to simply computing $f^s_{DA} := \sum_{(i,j)\in\mathcal{E}} y^*_{ij} \min_{p\in\mathcal{P}_{ij}} C^{ps}_{ij}$;

3. Finally, compute $f_{DA} := \sum_{s \in \mathcal{S}} \pi^s f^s_{DA}$.

Basically, f_{DA} represents the expected cost that can be obtained by implementing, at the first stage, the decisions suggested by the deterministic problem derived through our approximation. Consequently, t_{DA} represents, in our tests, the computational time to execute all three of the above steps, including the β calibration procedure and calculation of the logarithm of the accessibility for each arc in the objective function of (44).

Table 1 displays the percentage gaps $f\%$ obtained comparing the DA and DEP approaches on the 150 instances randomly generated from the theoretical distributions. More precisely, each entry reports the average and standard deviation (in square brackets) of the percentage gaps $f\%$ among the five random instances generated for each number of nodes, number of paths, and type of distribution. Columns and rows labelled with $Total$ report averages and standard deviations on the relative aggregation of instances.

Instance		Distribution					
$\|\mathcal{I}\|$	$\|\mathcal{P}\|$	Gumbel	Laplace	Logistic	Normal	Uniform	Total:
50	3	0.23 [0.17]	2.12 [2.44]	0.92 [1.08]	1.09 [0.93]	2.14 [2.85]	**1.30** [1.80]
	4	0.58 [0.32]	0.53 [0.57]	0.61 [0.64]	2.00 [2.29]	2.58 [0.93]	**1.26** [1.39]
	5	1.62 [3.19]	0.66 [0.45]	0.31 [0.68]	0.66 [0.75]	1.08 [0.76]	**0.87** [1.49]
	Total:	**0.81** [1.82]	**1.10** [1.55]	**0.62** [0.81]	**1.25** [1.50]	**1.93** [1.78]	**1.14** [1.56]
100	3	1.76 [2.71]	0.69 [0.67]	0.71 [0.37]	2.46 [2.25]	2.02 [1.52]	**1.53** [1.75]
	4	1.26 [0.82]	0.70 [0.49]	1.27 [0.83]	0.62 [0.65]	1.41 [1.16]	**1.05** [0.82]
	5	0.60 [0.69]	2.58 [1.51]	6.11 [9.25]	2.10 [2.22]	1.52 [0.61]	**2.58** [4.39]
	Total:	**1.21** [1.63]	**1.32** [1.30]	**2.70** [5.56]	**1.72** [1.91]	**1.65** [1.11]	**1.72** [2.81]

Table 1 Percentage gaps ($f\%$) obtained for random generated instances. Each entry indicates average gap and its standard deviation in square brackets over all random generated instances characterized by $|\mathcal{I}|$ number of nodes, $|\mathcal{P}|$ number of paths, and specific distribution (Gumbel, Laplace, Logistic, Normal, Uniform).

The results were excellent in terms of quality. The overall average gaps were 1.14% for all instances with 50 nodes and 1.72% for those with 100 nodes. The standard deviations confirm the acceptable stability of the approximation. All average gaps (and deviations) increased only marginally by increasing the size of the instances (in terms of nodes or paths) for virtually all types of distribution. The poorest behavior was demonstrated by the Logistic distribution, in particular concerning the largest instances ($|\mathcal{I}| = 100$ and $|\mathcal{P}| = 5$), for which an average gap of approximately 6% was observed. In all remaining cases, the average gaps did not exceed 2.6% and were less than 1% in approximately half of the cases.

Special attention should be given to the results concerning the Uniform distribution. As previously highlighted, although it does not satisfy assumption (23), we decided to include these experiments to investigate the behavior of the proposed method when the distribution of the variations was not known. In this case, the observed gaps were interesting and in line with the other distributions (i.e., 1.93% for $|\mathcal{I}| = 50$ and 1.65% for $|\mathcal{I}| = 100$). This indicates a broader applicability for the proposed deterministic approximation, as it can be expected to provide accurate results for a class of distribution even larger than the Gumbel domain of attraction (possibly indicating that assumption (23) is sufficient, yet not necessary for the derivation of our results).

The quality of the proposed approximation method was also tested considering a more realistic traffic model, which totally disregarded any assumption on the resulting empirical distribution of the costs. The percentage gaps $f\%$ obtained comparing the DA and DEP approaches on the 144 realistic instances are displayed in Table 2. Again, each entry reports the average and standard deviation (in square brackets) of the percentage gaps $f\%$ among the random instances generated for each number of nodes, number of paths, and type of traffic congestion. In this case, the results obtained remain acceptable, considering the complexity of the underlying problem. The approximation provides, on average, a solution differing from that of DEP by less than 4.6% for all type of networks (with reasonable standard deviations). Furthermore, the gaps do not appear to increase as the number of nodes or paths increase, demonstrating, again, acceptable stability and scalability of the approach. In fact, the approximation appears to function more effectively when the number of nodes and number of possible alternatives per arc (paths) are greater.

Instance		Type of traffic			
$\lvert\mathcal{I}\rvert$	$\lvert\mathcal{P}\rvert$	Low congestion	Mixed situation	High congestion	Total:
50	3	2.99 [1.61]	7.36 [11.02]	3.98 [2.81]	**4.78** [4.73]
	4	3.33 [1.78]	3.87 [1.16]	3.92 [2.03]	**3.71** [1.76]
	5	5.74 [2.13]	2.35 [2.67]	5.58 [2.79]	**4.56** [2.71]
	Total:	**4.02** [2.60]	**4.53** [6.34]	**4.49** [2.18]	**4.35** [3.31]
100	3	2.70 [1.41]	2.16 [0.80]	3.91 [1.91]	**2.92** [1.67]
	4	3.99 [1.43]	3.46 [1.32]	5.02 [1.18]	**4.16** [1.40]
	5	4.64 [2.52]	3.07 [1.42]	4.65 [2.40]	**4.12** [2.31]
	Total:	**3.78** [1.87]	**2.90** [1.23]	**4.53** [1.97]	**3.74** [1.89]

Table 2 Percentage gaps ($f\%$) obtained for realistic instances. Each entry indicates average gap and its standard deviation in square brackets over all realistic instances characterized by $|\mathcal{I}|$ number of nodes, $|\mathcal{P}|$ number of paths, and specific traffic condition (Low or High congestion, and Mixed situation).

To confirm that the performances presented actually depend on the estimator derived from our approximation, we also report the results obtained by approximating the problem simply using mean values for the stochastic costs. This mean-value approximation (MVA) also creates a deterministic problem where, for each traveled arc, the minimum-cost path is selected by default. More precisely, for each arc $(i,j) \in \mathcal{E}$, we consider a deterministic cost

$$c_{ij} = \arg\min_{p \in \mathcal{P}_{ij}} \left(\frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} c_{ij}^{ps} \right).$$

Similar to the previous tests, we compared MVA to DEP in terms of quality, i.e., we calculated

$$f_m\% := 100 \frac{f_{MVA} - f_{DEP}}{f_{DEP}},$$

where f_{MVA} is the value of the objective function of the optimal solution of the problem approximated through mean-values. The detailed results of $f_m\%$ for the random-generated and realistic instances are reported in Table 6 and 7, respectively (see Appendix C). For the random instances, we observed an average gap of 6%, whereas for the realistic instances, the average gap was higher (approximately 18%). Moreover, in specific instances, the theoretical instances also presented extremely high standard deviation, highlighting the unstable behavior of the mean-value estimator. These results were strongly outperformed by those obtained using our approximation. In fact, on average, DA provided solutions that were 5% and 14% better than those obtained by MVA on the theoretical and realistic instances, respectively.

In Table 3, we have summarized the computational times t_{DA} and t_{DEP} observed for the solution approaches of DA and DEP, respectively, on both random generated and realistic instances. For

Instance		Random instances		Realistic instances	
$\lvert\mathcal{I}\rvert$	$\lvert\mathcal{P}\rvert$	$t_{DA}(s)$	$t_{DEP}(s)$	$t_{DA}(s)$	$t_{DEP}(s)$
50	3	7.0	714.7	1.9	722.1
	4	5.6	1037.0	1.8	437.4
	5	6.0	655.8	2.1	830.0
	Avg:	**6.2**	**802.5**	**1.9**	**663.2**
100	3	97.7	6791.1	33.5	6805.1
	4	86.9	5930.6	29.2	6670.0
	5	104.7	7006.2	24.8	7162.9
	Avg:	**96.4**	**6576.0**	**29.1**	**6879.3**

Table 3 **Computational times in seconds of DA (t_{DA}) vs. DEP (t_{DEP}) approaches for both random-generated and realistic instances, grouped by number of nodes $|\mathcal{I}|$ and number of paths $|\mathcal{P}|$.**

both cases, we do not provide the detailed results per type of distribution or type of network because the differences were not meaningful, and depended primarily on the number of nodes. In all experiments, the time required to obtain the DA solution was negligible compared to the time required to perform the DEP solution. DEP required, on average, approximately 800 seconds to solve instances with 50 nodes and approached the threshold time of 2 hours for instances with 100 nodes. Conversely, DA required, on average, less than 7 seconds and less than 100 seconds to solve instances with 50 and 100 nodes, respectively. We can also observe that, on average, the realistic instances were solved in less time by DA compared to the random instances. Approximately 2 seconds and 30 seconds were required for solving $|\mathcal{I}| = 50$ and 100 instances, respectively.

Finally, in Tables 4 and 5, we summarize the value of the compromise offered by DA for the two main sets of instances in terms of efficiency and effectiveness. More precisely, we compared the loss in effectiveness $f\%$ and gain in efficiency $t\%$, calculated as

$$t\% := 100\frac{t_{DEP} - t_{DA}}{t_{DEP}},$$

when using the proposed DA compared to DEP. On average, by sacrificing from 1% to 4% of the solution quality, the approximation allowed a gain of two orders of magnitude in efficiency.

Distribution	$f\%$	$t\%$
Gumbel	1.01	97.93
Laplace	1.21	98.69
Logistic	1.66	96.98
Normal	1.49	98.10
Uniform	1.79	98.46
Avg:	**1.43**	**98.03**

Table 4 Loss in effectiveness ($f\%$) vs. gain in efficiency ($t\%$) of DA compared to DEP for random instances.

Traffic situation	$f\%$	$t\%$
Low congestion	4.51	99.60
Mixed situation	3.71	99.80
High Congestion	3.90	99.52
Avg:	**4.12**	**99.60**

Table 5 Loss in effectiveness ($f\%$) vs. gain in efficiency ($t\%$) of DA compared to DEP for realistic instances.

7. Conclusions

In this paper, we studied the $smpTSP_{dc}$. We demonstrated that, under a mild assumption on the distribution of the random cost variations and if such variations are just asymptotically independent, a deterministic approximation of the problem can be derived using the theory of extreme values. We also demonstrated that the asymptotic independence assumption on the travel costs is not overly restrictive in real network applications. In fact, it allows the addressing of realistic traffic models such as the well-known BPR function. Finally, we validated the behavior of the proposed methodology using extensive computational experiments on random generated and realistic instances with up to 100 nodes and five possible different paths per arc. The deterministic approximation is definitely capable of solving the problem with a very good compromise between the quality of the solution and overall efficiency compared to the standard equivalent SP approaches and state-of-the-art solvers. On average, the deterministic approximation can determine solutions with a 1–4% gap compared to the optimal ones in less than 100 seconds. Conversely, an exact solver requires hours to obtain an optimal solution. Please note that the proposed approach is expected to achieve similar (or superior) performance for instances with a greater number of paths and considered scenarios, being an asymptotic approximation.

Our future research can be outlined as follows. First, encouraged by the good results obtained even in those cases where the theoretical assumptions for the derivation of our approximation did not hold, we would like to further investigate the possible relaxations of such assumptions. Moreover, we would like to concentrate on determining a general method to calibrate the β parameter that better exploits the instance features. Finally, a time-dependent version of the problem can be studied and approximated through other recent developments on random utility choice models (Tadei, Perboli, and Manerba 2019, Roohnavazfar et al. 2019, 2020).

References

Beraldi P, Bruni ME, Manerba D, Mansini R, 2017 *A stochastic programming approach for the traveling purchaser problem. IMA Journal of Management Mathematics* 28(1):41–63.

Campbell AM, Gendreau M, Thomas BW, 2011 *The orienteering problem with stochastic travel and service times. Annals of Operations Research* 186(1):61–81.

Campbell AM, Thomas BW, 2008 *Probabilistic traveling salesman problem with deadlines. Transportation Science* 42(1):1–21.

Carraway R, Morin T, Moskowit H, 1989 *Generalized dynamic programming for stochastic combinatorial optimization. Operations Research* 5(37):819–829.

Chen BY, Lam W, Sumalee A, Li Q, Lam Tam M, 2014 *Reliable shortest path problems in stochastic time-dependent networks. Journal of Intelligent Transportation Systems* 18:177–189.

Ehmke JF, Campbell AM, Thomas BW, 2016 *Vehicle routing to minimize time-dependent emissions in urban areas. European Journal of Operational Research* 251(2):478–494.

Eiger A, Mirchandani PB, Soroush H, 1985 *Path preferences and optimal paths in probabilistic networks. Transportation Science* 19(1):75–84.

Evers L, Glorie K, van der Ster S, Barros AI, Monsuur H, 2014 *A two-stage approach to the orienteering problem with stochastic weights. Computers & Operations Research* 43:248–260.

Fadda E, Perboli G, Tadei R, 2018 *Customized multi-period stochastic assignment problem for social engagement and opportunistic iot. Computers & Operations Research* 93:41–50.

Fan Y, Kalaba R, Moore J, 2005 *Shortest paths in stochastic networks with correlated link costs. Computers & Mathematics with Applications* 49(9):1549–1564.

Galambos J, 1978 *The Asymptotic Theory of Extreme Order Statistics* (John Wiley, New York).

Gutin G, Punnen AP, eds., 2002 *The traveling salesman problem and its variations.* Combinatorial Optimization (Dordrecht, London: Kluwer Academic).

Hanafi S, Mansini R, Zanotti R, 2020 *The multi-visit team orienteering problem with precedence constraints. European Journal of Operational Research* 282(2):515–529.

Hansen W, 1959 *How accessibility shapes land use. Journal of the American Institute of Planners* 25:73–76.

Huang Y, Zhao L, Woensel TV, Gross JP, 2017 *Time-dependent vehicle routing problem with path flexibility. Transportation Research Part B: Methodological* 95:169–195.

Huang Z, Zheng QP, Pasiliao E, Boginski V, Zhang T, 2019 *A cutting plane method for risk-constrained traveling salesman problem with random arc costs. Journal of Global Optimization* 74(4):839–859.

Hwang T, Ouyang Y, 2015 *Urban freight truck routing under stochastic congestion and emission considerations. Sustainability* 7:6610–6625.

Kao E, 1978 *A preference order dynamic program for a stochastic traveling salesman problem. Operations Research* 6(26):1033–1045.

Kaut M, Vladimirou H, Wallace S, Zenios S, 2007 *Stability analysis of portfolio management with conditional value-at-risk. Quantitative Finance* 7:397–409.

Kenyon AS, Morton DP, 2003 *Stochastic vehicle routing with random travel times. Transportation Science* 37(1):69–82.

Kirkpatrick S, Toulouse G, 1985 *Configuration space analysis of traveling salesman problem. J. Phys. France* 46:1277–1292.

Lawler E, Lenstra JK, Rinnooy Kan A, Shmoys D, 1985 *Traveling salesman problem: a guided tour of combinatorial optimization* (Wiley, Chichester).

Letchford AN, Nasiri SD, 2015 *The steiner traveling salesman problem with correlated costs. European Journal of Operational Research* 245(1):62–69.

Maggioni F, Perboli G, Tadei R, 2014 *The multi-path traveling salesman problem with stochastic travel costs: Building realistic instances for city logistics applications. Transportation Research Procedia* 3:528–536, 17th Meeting of the EURO Working Group on Transportation, July 2-4, 2014. Sevilla (Spain).

Manerba D, Mansini R, Riera-Ledesma J, 2017 *The traveling purchaser problem and its variants. European Journal of Operational Research* 259(1):1–18.

Mezard M, Parisi G, 1986 *A replica analysis of the traveling salesman problem. J. Phys. France* 47:1285–1296.

Nelsen RB, 2006 *An Introduction to Copulas*. Springer series in Statistics (Springer-Verlag New York).

Osborne MJ, Rubinstein A, 1994 *A course in game theory* (MIT press).

Perboli G, Tadei R, Baldi M, 2012 *The stochastic generalized bin packing problem. Discrete Applied Mathematics* 160:1291–1297.

Perboli G, Tadei R, Gobbato L, 2014 *The multi-handler knapsack problem under uncertainty. European Journal of Operational Research* 236(3):1000–1007.

Reinelt G, 1994 *The traveling salesman: computational solutions for TSP applications*, volume 840 (Springer, Berlin).

Roohnavazfar M, Manerba D, De Martin JC, Tadei R, 2019 *Optimal paths in multi-stage stochastic decision networks. Operations Research Perspectives* 6:100124, URL `http://dx.doi.org/https://doi.org/10.1016/j.orp.2019.100124`.

Roohnavazfar M, Manerba D, Pasandideh SHR, Tadei R, 2020 *Single-machine job-scheduling problem as a multi-stage dynamic stochastic decision process (submitted)* .

Samaranayake S, Blandin S, Bayen A, 2012 *A tractable class of algorithms for reliable routing in stochastic networks. Transportation Research Part C: Emerging Technologies* 20(1):199–217.

Sniedovich M, 1981 *Analysis of a preference order traveling salesman problem. Operations Research* 6(29):1234–1237.

Tadei R, Fadda E, Gobbato L, Perboli G, Rosano M, 2016 *An ict-based reference model for e-grocery in smart cities. Proceedings of the First International Conference on Smart Cities - Volume 9704*, 22–31, Smart-CT 2016 (Berlin, Heidelberg: Springer-Verlag), ISBN 978-3-319-39594-4, URL `http://dx.doi.org/10.1007/978-3-319-39595-1_3`.

Tadei R, Perboli G, Manerba D, 2018 *A recent approach to derive the multinomial logit model for choice probability. P. Daniele and L. Scrimali (eds.), New Trends in Emerging Complex Real Life Problems, AIRO Springer Series 1 ODS2018, Sept 10-13, 2018. Taormina (Italy).*

Tadei R, Perboli G, Manerba D, 2019 *The multi-stage dynamic stochastic decision process with unknown distribution of the random utilities. Optimization Letters (in press),* URL `http://dx.doi.org/10.1007/s11590-019-01412-1`.

Tadei R, Perboli G, Perfetti F, 2017 *The multi-path traveling salesman problem with stochastic travel costs. EURO Journal on Transportation and Logistics* 6:3–23.

Tadei R, Perboli G, Ricciardi N, Baldi MM, 2012 *The capacitated transshipment location problem with stochastic handling costs at the facilities. International Transactions in Operational Research* 19(6):789–807.

Toriello A, William B, Poremba HM, 2013 *A dynamic traveling salesman problem with stochastic arc costs. Operations Research* 62(5):1107–1125.

Toth P, Vigo D, 2014 *Vehicle Routing: Problems, Methods, and Applications - Second Ed.* MOS-SIAM Series on Optimization (Philadelphia, PA: Society for Industrial and Applied Mathematics).

US Department of Commerce, Bureau of Public Roads, 1964 *Traffic assignment manual* (Washington, U.S. Dept. of Commerce, Bureau of Public Roads, Office of Planning, Urban Planning Division).

Vansteenwegen P, Souffriau W, Oudheusden DV, 2011 *The orienteering problem: A survey. European Journal of Operational Research* 209(1):1–10.

Verbeeck C, Vansteenwegen P, Aghezzaf EH, 2016 *Solving the stochastic time-dependent orienteering problem with time windows. European Journal of Operational Research* 255(3):699–718.

Wallace S, Ziemba W, 2005 *Applications of Stochastic Programming.* MPS-SIAM Series on Optimization (Society for Industrial and Applied Mathematics).

Wardrop JG, 1952 *Road paper. Some theoretical aspects of road traffic research. Proceedings of the Institution of Civil Engineers* 1(3):325–362, URL `http://dx.doi.org/10.1680/ipeds.1952.11259`.

Wardrop JG, Whitehead JI, 1952 *Correspondence. Some theoretical aspects of road traffic research. Proceedings of the Institution of Civil Engineers* 1(5):767–768, URL `http://dx.doi.org/10.1680/ipeds.1952.11362`.

Wästlund J, 2010 *The mean field traveling salesman and related problems. Acta Mathematica* 204(1):91–150.

Appendix A: Analytical derivation of $\mathbb{E}_{\mathbf{\Theta}}[\tilde{C}_{ij}]$ in (43).

From (42), we have

$$\mathbb{E}_{\mathbf{\Theta}}[\tilde{C}_{ij}] = \int_{-\infty}^{+\infty} x \; d\mathbb{P}[\tilde{C}_{ij} \leq x] = -\int_{-\infty}^{+\infty} x \; d\mathbb{P}[\tilde{C}_{ij} > x] = \int_{-\infty}^{+\infty} x\beta A_{ij} e^{\beta x} e^{-A_{ij}e^{\beta x}} dx.$$

By performing the substitution $t = A_{ij}e^{\beta x}$, i.e., $x = \frac{1}{\beta}\ln(t/A_{ij})$ and $dx = \frac{1}{\beta t}dt$, the integral in (43) can be analytically calculated as follows

$$\begin{aligned}
\mathbb{E}_{\mathbf{\Theta}}[\tilde{C}_{ij}] = & \int_{0}^{+\infty} \frac{1}{\beta}\ln(t/A_{ij})\beta t e^{-t}\frac{1}{\beta t}dt = \\
= & \frac{1}{\beta}\int_{0}^{+\infty} \ln(t/A_{ij})e^{-t}dt = \\
= & \frac{1}{\beta}\int_{0}^{+\infty} (\ln(t) - \ln(A_{ij}))e^{-t}dt = \\
= & \frac{1}{\beta}\left(\int_{0}^{+\infty} \ln(t)e^{-t}dt\right) - \frac{1}{\beta}\ln(A_{ij})\left(\int_{0}^{+\infty} e^{-t}dt\right) = \\
= & -\frac{1}{\beta}\gamma - \frac{1}{\beta}\ln(A_{ij}) = \\
= & -\frac{1}{\beta}(\gamma + \ln(A_{ij})),
\end{aligned}$$

where $\gamma = -\int_{0}^{+\infty}\ln(t)e^{-t}dt$ is the Euler constant.

Appendix B: Proof of Proposition 1

Proof. Let us consider a multi-agent system characterized by a set $\mathcal{N} = \{1, \ldots, N\}$ of rational agents solving our two stage stochastic problem (2)–(9). We assume that the agents arrive in the network at different times.

Without loss of generality, let us consider a given arc $(i,j) \in \mathcal{E}$ and the set of agents that have chosen to travel on that arc (i,j) during the first stage. When each agent approaches the actual travel on (i,j), he must complete the information regarding the costs of the different alternative paths (i.e., the stochasticity has been realized). Then, according to the second stage of our problem, each agent (being rational) will choose the least expensive path.

From Wardrop's first principle, at equilibrium, each path used by at least one agent is the minimum-cost path. Thus, it holds that

$$C_{ij}^{p^1 s} = C_{ij}^{p^2 s}, \quad \forall p^1 \neq p^2 \in \mathcal{P}_{ij}^U, s \in \mathcal{S},$$

where $\mathcal{P}_{ij}^U \subseteq \mathcal{P}_{ij}$ is the subset of paths used at equilibrium. Let us now assume that on a path $\hat{p} \in \mathcal{P}_{ij}$, a scenario $\hat{s}$ occurs such that $\Theta_{ij}^{\hat{p}\hat{s}} \leq \hat{r}$, with $\hat{r}$ close to $-c_{ij}^{\hat{p}}$.

This event perturbs the equilibrium and leads to the following new conditions

$$C_{ij}^{\hat{p}\hat{s}} \sim 0,$$

$$C_{ij}^{\hat{p}\hat{s}} < C_{ij}^{p\hat{s}}, \quad p \in \mathcal{P}_{ij}.$$

This means that from this point onward, all agents willing to travel on arc (i,j) will use path $\hat{p}$. Consequently, the traffic flow $Q_{ij}^{\hat{p}\hat{s}}$ will increase and the traffic flow $Q_{ij}^{p\hat{s}}$ for all $p \neq \hat{p}$ will decrease. Because the cost function $H_{ij}^{p\hat{s}}$ is increasing, the cost $C_{ij}^{\hat{p}\hat{s}}$ will increase, whereas the cost $C_{ij}^{p\hat{s}}$ for all $p \neq \hat{p}$ will decrease. Moreover, during all transition to the next equilibrium, it holds that

$$C_{ij}^{p\hat{s}} > C_{ij}^{\hat{p}\hat{s}} > 0, \quad p \in \mathcal{P}_{ij}.$$

Thus, for each $p \neq \hat{p} \in \mathcal{P}_{ij}$, the random cost variation $\Theta_{ij}^{p\hat{s}}$ is not expected to assume a value close to its lower bound $-c_{ij}^{p}$. This proves that the travel cost variations Θ_{ij}^{p} are asymptotically independent in the right-hand side neighborhood of the lower bound $(-c_{ij}^{p})$ of the support of their distribution. $\square$

Appendix C: Results of the mean-value approximation (MVA)

Tables 6 and 7 present the detailed results obtained by solving the mean-value approximation (MVA) for random and realistic instances, respectively.

Instance		Distribution					
$\|\mathcal{I}\|$	$\|\mathcal{P}\|$	Gumbel	Laplace	Logistic	Normal	Uniform	Total:
50	3	20.82 [33.87]	3.53 [2.63]	3.84 [2.32]	4.57 [1.70]	17.70 [23.45]	**10.09** [18.18]
	4	16.47 [25.79]	3.39 [1.74]	4.18 [1.83]	7.45 [2.35]	4.23 [09.34]	**7.15** [12.06]
	5	9.86 [04.27]	3.91 [1.88]	2.93 [0.93]	6.59 [1.46]	7.27 [01.36]	**6.11** [03.26]
	Total:	**15.72** [22.83]	**3.61** [1.93]	**3.65** [1.71]	**6.20** [2.11]	**9.74** [14.51]	**7.78** [12.63]
100	3	7.31 [5.06]	2.56 [1.19]	4.18 [1.29]	6.48 [3.86]	12.36 [15.30]	**6.58** [7.46]
	4	5.62 [1.86]	3.84 [1.57]	4.56 [1.44]	3.59 [1.74]	8.38 [07.14]	**5.20** [3.60]
	5	4.98 [2.03]	3.53 [1.96]	4.61 [1.89]	11.47 [7.92]	7.58 [01.71]	**6.43** [4.55]
	Total:	**5.97** [3.18]	**3.31** [1.56]	**4.45** [1.42]	**7.18** [5.79]	**9.44** [9.13]	**6.07** [5.40]

Table 6 Percentage gaps ($f_m\%$) obtained for random generated instances. Each entry displays average gap and its standard deviation in square brackets over all random generated instances characterized by $|\mathcal{I}|$ number of nodes, $|\mathcal{P}|$ number of paths, and specific distribution (Gumbel, Laplace, Logistic, Normal, Uniform).

Instance		Type of traffic			
$\|\mathcal{I}\|$	$\|\mathcal{P}\|$	Low congestion	Mixed situation	High congestion	Total:
50	3	15.11 [0.94]	14.62 [1.10]	14.58 [1.13]	**14.81** [1.02]
	4	18.29 [1.18]	18.31 [1.08]	19.18 [1.02]	**18.48** [1.11]
	5	22.23 [1.11]	21.76 [1.24]	21.54 [1.72]	**21.90** [1.25]
	Total:	**18.54** [3.15]	**18.23** [3.17]	**18.43** [3.25]	**18.40** [3.13]
100	3	14.54 [0.69]	14.10 [0.75]	14.26 [0.79]	**14.31** [0.72]
	4	17.99 [0.59]	17.96 [1.07]	19.00 [0.82]	**18.18** [0.91]
	5	21.06 [0.72]	20.85 [0.83]	21.89 [0.76]	**21.14** [0.83]
	Total:	**17.86** [2.79]	**17.63** [2.95]	**18.38** [3.36]	**17.88** [2.94]

Table 7 Percentage gaps ($f_m\%$) obtained for realistic instances. Each entry displays average gap and its standard deviation in square brackets over all realistic instances characterized by $|\mathcal{I}|$ number of nodes, $|\mathcal{P}|$ number of paths, and specific traffic condition (Low or High congestion, and Mixed situation).