Distributionally Robust Markovian Traffic Equilibrium [*]

Selin Damla Ahipaşaoğlu[†]   Uğur Arıkan[‡]   Karthik Natarajan[§]

May 2017

Abstract

Stochastic user equilibrium models are fundamental to the analysis of transportation systems. Such models are typically developed under the assumption of route based choice models for the users. A class of link based models under a Markovian assumption on the route choice behavior of the users has been proposed to deal with the drawbacks of route based choice models. However, the application of this model has been thus far mainly restricted to the multinomial logit model. Furthermore, the complete distribution of the random utilities in such a model is rarely known to a system planner. In this paper, we propose a distributionally robust Markovian traffic equilibrium model and a corresponding choice model under the assumption that the marginal distributions of the link utilities are known but the joint distribution is unknown. By using a distributionally robust approach, we develop a new convex optimization formulation and propose an efficient algorithm to compute equilibrium flows. In the special case of exponential marginals, our formulation reduces to the entropy formulation of the Markovian multinomial logit model. Importantly, our formulation is completely link based and relaxes the assumption of independence

[†]Engineering Systems and Design, Singapore University of Technology and Design, 8 Somapah Road, Singapore 487372, ahipasaoglu@sutd.edu.sg

[‡]Engineering Systems and Design, Singapore University of Technology and Design, 8 Somapah Road, Singapore 487372, ugur_arikan@sutd.edu.sg

[§]Engineering Systems and Design, Singapore University of Technology and Design, 8 Somapah Road, Singapore 487372, karthik_natarajan@sutd.edu.sg

and identical distributions in the link utilities. Our numerical experiments indicate that this provides modeling flexibility and computational tractability for system planners interested in calculating traffic equilibrium.

Keywords: Markovian traffic equilibrium, distributionally robust, convex optimization

1   Introduction

Traffic equilibrium models are fundamental to the analysis of transportation systems. The inputs to such models are the network topology, origin-destination (OD) pairs, demands of OD pairs and link performance functions. The goal is to estimate the equilibrium link flows. Daganzo and Sheffi (1977) relaxed the perfect information assumption of the user equilibrium (Wardrop, 1952) and proposed a stochastic user equilibrium (SUE) with a discrete choice model to capture randomness in user route choice behavior. Route based SUE models assign the demand probabilistically to the alternative routes. However, traffic networks are often cyclic and might include infinitely many routes. A common approach is to then first generate a small set of plausible routes in a choice set with respect to some criteria. Bekhor et al. (2006) provides an evaluation of such choice set generation algorithms. These algorithms fall in the class of deterministic methods, and ignore the effect of congestion. While there has been much focus on developing route based SUE models integrated with discrete choice models to accurately model user behavior, lesser attention has been paid to computational drawbacks of such route based models. One approach that has been proposed in this context is to assume a sequential Markovian decision-making process where the route choice of a user is determined by successive link choices that is made independently of how the users reached the current node (see Akamatsu (1996, 1997); Bell (1995); Baillon and Cominetti (2008)). Evaluating the steady state flows in this Markov chain then allows the system planner to develop models that evaluate all routes without explicitly generating them. Despite its obvious computational advantages, this idea has attracted lesser attention than the route based models, and thus far its implementation has been primarily limited to the multinomial logit (MNL) model.

1.1 Markovian Choice

Let $G = (N, A)$ denote a directed traffic network with a set of nodes N and a set of links A, where W is the set of OD pairs and $D \subseteq N$ is the set of destination nodes in the set of OD pairs. In a Markovian choice model, a node can be considered as the state of the user. A user at a particular node chooses one of the links emanating from that node to move to the next state with a goal to reach the destination, and proceeds in this manner until she reaches the destination. We define $G_d = (N_d \cup \{d\}, A_d)$ as the subnetwork that can be used by the users towards destination d. Here, A_d is the subset of links $(i, j) \in A$ for which i is reachable from an origin o such that $(o, d) \in W$ without visiting d, and destination d is reachable from j. The subset N_d contains all tail nodes of links in A_d, i.e., the transient states for destination d. Finally, we let $N_d^+(i)$ denote the out-neighbourhood of node i and $N_d^-(i)$ denote the in-neighbourhood of node i in G_d (see Figure 1). Without loss of generality, we assume that $A = \cup_{d \in D} A_d$. The choice at a node is independent of the previous states the user visits. The uncertainty in the choice behavior is introduced by modelling the link cost of arc (i, j) as a random variable which is equal to $t_{ij} - \widetilde{\epsilon}_{ij}^d$ where t_{ij} is the deterministic cost of travelling on the link and $\widetilde{\epsilon}_{ij}^d$ is a random utility error term of using the link towards destination d. The random term captures the components of the preferences of the users that is unobservable to the system planner. Let w_j^d denote the expected minimum cost of travelling from node j to destination d. A Markovian choice model for the users travelling towards destination d is then characterized as the solution to the following system of equations:

$$w_i^d = E_{\theta_{id}} \left[\min_{j \in N_d^+(i)} \left\{ t_{ij} - \widetilde{\epsilon}_{ij}^d + w_j^d \right\} \right], \qquad \forall i \in N_d, \qquad (1)$$

$$w_d^d = 0, \qquad (2)$$

where θ_{id} is the joint distribution of the error terms $\{\widetilde{\epsilon}_{ij}^d ; j \in N_d^+(i)\}$.

The error terms of the links emanating from two different nodes are assumed to be independent, while the error terms of the links exiting a particular node might be correlated. The corresponding conditional link choice probability of selecting link (i, j) at node i is defined as:

$$p_{ij}^d = P_{\theta_{id}} \left\{ j = \operatorname*{argmin}_{k \in N_d^+(i)} \left\{ t_{ik} - \widetilde{\epsilon}_{ik}^d + w_k^d \right\} \right\}, \qquad \forall (i, j) \in A_d.$$

Figure 1: Illustration of network terminology.
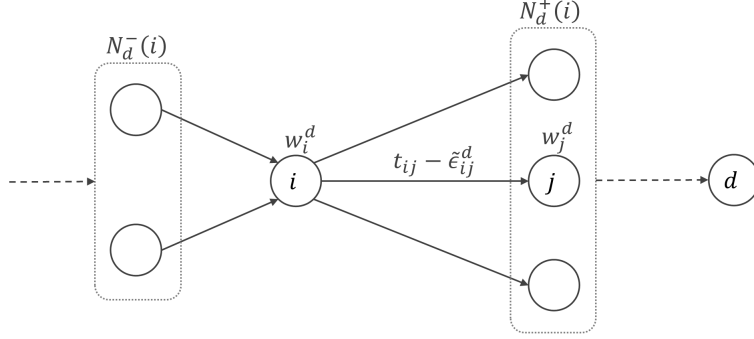
Let p_r be the probability of choosing route $r \in R_d$, where R_d is the set of routes with destination d. This set might be countably infinite to allow for the possibility of cycles. Let $\delta_{ij,r}$ be the number of times link (i,j) is used in a route r. Then the following relation exists between the link and route choice probabilities:

$$ p_r = \prod_{(i,j) \in A_d} \left(p_{ij}^d \right)^{\delta_{ij,r}}, \qquad \forall r \in R_d, d \in D. \qquad (3) $$

The flows towards a designated destination are obtained directly from the route choice probabilities, as follows:

$$ y_r = h_{o_r}^d \cdot p_r, \qquad \forall r \in R_d, $$
$$ x_{ij}^d = \sum_{r \in R_d : (i,j) \in r} y_r, \qquad \forall (i,j) \in A_d, $$

where y_r is the flow on route r, o_r is the origin of route r, h_o^d is the demand of the OD pair (o,d), and x_{ij}^d is the flow on link (i,j) towards destination d. Finally, the total link flow on the arc denoted by f_{ij} is obtained by summing the flows as follows:

$$ f_{ij} = \sum_{d \in D : (i,j) \in A_d} x_{ij}^d, \qquad \forall (i,j) \in A. $$

Building on the earlier work of Dial (1971) on link based models, Akamatsu (1996, 1997) and Bell (1995) proposed the Markovian logit model. In this model, the random utility error terms $\widetilde{\epsilon}_{ij}^d$ are assumed to be iid Gumbel random variables with a dispersion parameter β. This model is also sometimes referred to as the Recursive Logit (RL) model (see Fosgerau et al.

(2013)). Under this assumption, a closed form expression for the conditional link probability expression in the w_j^d variables is given as follows:

$$p_{ij}^d = \frac{\exp\left[-\beta\left(t_{ij} + w_j^d\right)\right]}{\displaystyle\sum_{k\in N_d^+(i)} \exp\left[-\beta\left(t_{ik} + w_k^d\right)\right]}, \qquad \forall(i,j) \in A_d.$$

Solving for the Markovian logit model then reduces to solving the following set of nonlinear equations in the w_j^d variables:

$$w_i^d = -\frac{1}{\beta}\ln\left(\sum_{j\in N_d^+(i)} \exp\left[-\beta\left(t_{ij} + w_j^d\right)\right]\right), \qquad \forall i \in N_d, \qquad (4)$$

$$w_d^d = 0. \qquad (5)$$

A straightforward approach to solve this system of nonlinear equations is to convert it to a system of linear equations through an exponential transformation of the w_j^d variables (see Fosgerau et al. (2013) for such an approach). Recently, Mai et al. (2015) extended this to a Markovian nested recursive logit model that allows for correlations among the error terms. However in this case, the choice model is much harder to solve as there is no straightforward transformation of the system of nonlinear equations to a system of linear equations.

For general distributions such as the multivariate normal distribution, however no closed form expression is known for the link choice probability. As a result, to the best of our knowledge, the Markovian choice model has not been explored in much detail for general discrete choice models. One notable exception is the work of Baillon and Cominetti (2008) who proposed a Markovian Traffic Equilibrium (MTE) model for the congested case with general distributions. Let $\tau_{ij}(f_{ij})$ be a strictly increasing continuous cost function which defines the cost of a link as a function of its flow f_{ij}. Then the MTE in Baillon and Cominetti (2008) is defined as the solution to a fixed point problem in the variables $(f_{ij}, t_{ij}, w_i^d, x_{ij}^d, n_i^d)$ where n_i^d is the number

of users at node i moving towards destination d as follows:

$$t_{ij} = \tau_{ij}\left(f_{ij}\right), \qquad\qquad\qquad \forall(i,j) \in A,$$

$$w_i^d = E_{\theta_{id}}\left[\min_{j \in N_d^+(i)}\left\{t_{ij} - \widetilde{\epsilon}_{ij}^d + w_j^d\right\}\right], \qquad\qquad \forall i \in N_d, d \in D,$$

$$w_d^d = 0, \qquad\qquad\qquad \forall d \in D,$$

$$n_i^d = h_i^d + \sum_{k \in N_d^-(i)} x_{ki}^d, \qquad\qquad\qquad \forall i \in N_d, d \in D,$$

$$x_{ij}^d = n_i^d \cdot \mathrm{P}_{\theta_{id}}\left\{j = \operatorname*{argmin}_{k \in N_d^+(i)}\left\{t_{ik} - \widetilde{\epsilon}_{ik}^d + w_k^d\right\}\right\}, \quad \forall(i,j) \in A_d, d \in D,$$

$$f_{ij} = \sum_{d \in D:(i,j) \in A_d} x_{ij}^d, \qquad\qquad\qquad \forall(i,j) \in A.$$

Under mild assumptions on the distribution of the random utilities which is denoted by $\boldsymbol{\theta}$, the authors showed that the MTE exists and is unique. Furthermore, the link costs at equilibrium are obtained as the solution to the following unconstrained convex optimization formulation:

$$Z\left(\boldsymbol{\theta}\right) = \max_{\mathbf{t}} \sum_{d \in D}\sum_{i \in N_d} h_i^d \cdot w_i^d(\mathbf{t}) - \sum_{(i,j) \in A} \int_0^{t_{ij}} \tau_{ij}^{-1}(\omega)d\omega, \qquad (6)$$

where $w_i^d(\mathbf{t})$ is a concave and smooth function of the link costs and is implicitly defined in (1)-(2). The equilibrium link flow is given as $f_{ij} = \tau_{ij}^{-1}(t_{ij}^*)$ where t_{ij}^* is the unique optimal solution in the convex optimization formulation (6). While the optimization problem (6) is convex, the application of this model has been primarily restricted to the logit model. The reason for this largely stems from the challenge in computing the choice probabilities for general discrete choice models. Furthermore in link based models, the system planner also needs to solve a set of nonlinear equations in the w_j^d variables unlike the route based models. Another important assumption that is implicitly made in the model but is itself subject to scrutiny is that the system planner has complete knowledge on the distribution of the random utilities. In this paper, we derive a new class of Markovian traffic equilibrium models which relaxes this assumption while at the same time preserving computational tractability.

6

1.2 Distributionally Robust Markovian Traffic Equilibrium

To provide a distributionally robust perspective on the MTE, we start by re-formulating (6) as a constrained convex optimization problem in the decision variables (t_{ij}, w_i^d) as follows:

$$Z\left(\boldsymbol{\theta}\right) = \max_{\mathbf{t},\mathbf{w}} \ \sum_{d \in D} \sum_{i \in N_d} h_i^d \cdot w_i^d - \sum_{(i,j) \in A} \int_0^{t_{ij}} \tau_{ij}^{-1}(\omega)d\omega \tag{7}$$

$$\text{s.t. } w_i^d \leq E_{\theta_{id}} \left[\min_{j \in N_d^+(i)} \left\{ t_{ij} - \widetilde{\epsilon}_{ij}^d + w_j^d \right\} \right], \quad \forall i \in N_d, d \in D, \tag{8}$$

$$w_d^d = 0, \qquad\qquad\qquad\qquad\qquad \forall d \in D. \tag{9}$$

Note that at optimality since $h_i^d > 0$, the inequalities in (8) will be tight. This is a convex optimization formulation since the objective function of the maximization problem is concave and the constraints are convex in the decision variables. However an important assumption in the model is that the joint distribution of the error terms is known or at the very least samples from this distribution are available to the system planner.

We now provide a distributionally robust formulation for the MTE which relaxes this assumption as follows. Suppose that the joint distribution θ_{id} of the error terms $\{\widetilde{\epsilon}_{ij}^d; j \in N_d^+(i)\}$, is not completely known. Rather it is only known to lie in a set of joint distributions Θ_{id}. Then the distributionally robust Markovian traffic equilibrium is formulated as follows:

$$\max_{\mathbf{t},\mathbf{w}} \ \sum_{d \in D} \sum_{i \in N_d} h_i^d \cdot w_i^d - \sum_{(i,j) \in A} \int_0^{t_{ij}} \tau_{ij}^{-1}(\omega)d\omega \tag{10}$$

$$\text{s.t. } w_i^d \leq E_{\theta_{id}} \left[\min_{j \in N_d^+(i)} \left\{ t_{ij} - \widetilde{\epsilon}_{ij}^d + w_j^d \right\} \right], \quad \forall \theta_{id} \in \Theta_{id}, \forall i \in N_d, d \in D,$$

$$\tag{11}$$

$$w_d^d = 0, \qquad\qquad\qquad\qquad\qquad \forall d \in D, \tag{12}$$

where (11) enforces that the constraint is valid for all joint distributions in the set Θ_{id}. Since the seminal work of Scarf (1958) on the minmax newsvendor problem, there has been significant progress in developing robust optimization techniques for network flow problems. The applications studied include network design (see Atamturk and Zhang (2007)), dynamic empty container repositioning (see Erera and Savelsbergh (2009)), vehicle routing (see Carlsson and Delage (2013)), path based traffic equilibrium models (see

Ahipaşaoğlu et al. (2015, 2016)) and the traveling salesperson problem (see Carlsson and Behroozi (2017)). In this paper, we study the distributionally robust Markovian traffic equilibrium model for link based choice behavior. The main contributions of the paper are as follows:

(a) We define a Markovian marginal distribution choice model (referred to as M-MDM) and propose a convex optimization formulation to compute the M-MDM link choice probabilities. We show that the choice probabilities are unique under mild assumptions on the marginal distributions, and propose an efficient algorithm to calculate the link choice probabilities. This extends the entropy formulation that was first developed in Akamatsu (1996, 1997) to general marginal distributions. The result also extends an earlier convex optimization formulation developed in Natarajan et al. (2009) for link based models in acyclic networks by explicitly incorporating the Markovian assumption into the choice process.

(b) We propose a convex optimization formulation to compute the flows for the distributionally robust Markovian traffic equilibrium with the marginal distribution choice model (referred to as MTE-MDM). We propose the method of successive averages as an algorithm that can be used to find the equilibrium flows in congested traffic networks in this case. The algorithm is integrated with an efficient subroutine for calculating the link choice probabilities at a given flow. Our experiments indicate that the proposed solution approach is applicable to large traffic networks.

(c) Unlike the logit based models, the proposed model relaxes the independent and identically distributed error assumption and allows for using different families of distributions. We showcase the flexibility of the model in capturing nonidentical distributions and in incorporating heterogeneity at the link level. Moreover, the model incorporates the shape of the marginal distribution thereby capturing a wider range of user choice behavior. Our results thus extends the recent work of Ahipaşaoğlu et al. (2016) on the marginal distribution model from a route based model to a link based model.

The outline of the paper is as follows. In Section 2, we define the characteristics of the choice model M-MDM and propose an equivalent optimization problem. In Section 3, we propose the equilibrium model MTE-MDM, and show that the equilibrium can be computed by solving a pair of primal and dual optimization problems, in flow and cost variables, respectively. In Section 4, we propose the solution algorithms. In Section 5, we report nu-

merical results and discuss the flexibility of the proposed model. Finally, we give concluding remarks in Section 6.

2 Markovian Marginal Distribution Model (M-MDM)

In this section, we introduce a new discrete choice model for a single destination motivated by formulation (10)-(12). We consider the uncongested case in this section with a fixed link cost and extend the model to the congested case in the next section. We make the following assumptions:

A1 Random utility error terms of the links that emanate from two different nodes are independent of each other.

A2 The joint distribution θ_{id} of the random error terms $\{\tilde{\epsilon}_{ij}^d; j \in N_d^+(i)\}$ that emanate from a node is incompletely specified and only known to lie in a set of distributions with given marginal distributions. This set of distributions is designated as Θ_{id}.

A3 The random utility error term of each link has a finite first moment, namely $E_{\theta_{id}}[|\tilde{\epsilon}_{ij}^d|] < \infty$ for all $\theta_{id} \in \Theta_{id}$. The support of the random error term is either the whole real line $(-\infty, \infty)$, or the semi-infinite interval $(\underline{\epsilon}, \infty)$, where $\underline{\epsilon}$ is the minimum value for the corresponding error term. The cumulative distribution function is strictly increasing on the support and is continuous with a probability density function $f(.) > 0$ on the support.

Assumption A3 on the marginals is satisfied by a wide range of distribution families, such as the exponential, normal, student-t, logistic, Gumbel and gamma distributions to name a few. Building on the results of Natarajan et al. (2009) on discrete optimization problems with random objective function coefficients, Mishra et al. (2014) proposed a discrete choice model, termed as the Marginal Distribution Model (MDM), where the utility of an alternative is defined as the sum of a deterministic component and a random error term. We next describe this model in our context.

Consider the node i where the link costs t_{ij} and the expected costs to go from nodes $j \in N_d^+(i)$ denoted by w_j^d are known. In MDM, we evaluate the minimum expected cost at node i over all joint distributions with the given marginals. The minimum expected cost and the corresponding choice

probabilities at node i are calculated as follows:

$$w_i^d = \min_{\theta_{id} \in \Theta_{id}} E_{\theta_{id}} \left[\min_{j \in N_d^+(i)} \left\{ t_{ij} - \tilde{\epsilon}_{ij}^d + w_j^d \right\} \right], \tag{13}$$

$$p_{ij}^d = P_{\theta_{id}^*} \left\{ j = \underset{k \in N_d^+(i)}{\operatorname{argmin}} \left\{ t_{ik} - \tilde{\epsilon}_{ik}^d + w_k^d \right\} \right\}, \qquad \forall j \in N_d^+(i), \tag{14}$$

where θ_{id}^* is the extremal distribution in (13). For fixed values of w_j^d, Mishra et al. (2014) provided an equivalent reformulation of problem (13) as a convex minimization problem over finite dimensional choice probabilities rather than infinite dimensional probability measures as follows:

$$w_i^d = \min_{\mathbf{p_i^d}} \left\{ \sum_{j \in N_d^+(i)} \left((t_{ij} + w_j^d) p_{ij}^d - \int_{1-p_{ij}^d}^1 F_{ij}^{d\,(-1)}(\omega) d\omega \right) : \mathbf{p_i^d} \in \Delta_i^d \right\}, \tag{15}$$

where Δ_i^d is a unit simplex defined as follows:

$$\Delta_i^d = \left\{ \mathbf{p_i^d} : \sum_{j \in N_d^+(i)} p_{ij}^d = 1, p_{ij}^d \geq 0, \forall j \in N_d^+(i) \right\}.$$

Under the assumptions on the marginal distributions, (15) is a strictly convex optimization problem over the unit simplex and the choice probabilities are strictly positive. The Karush-Kuhn-Tucker (KKT) optimality conditions of (15) is:

$$p_{ij}^d = 1 - F_{ij}^d \left(\lambda_i^d + t_{ij} + w_j^d \right), \qquad \forall j \in N_d^+(i), \tag{16}$$

$$1 = \sum_{j \in N_d^+(i)} \left[1 - F_{ij}^d \left(\lambda_i^d + t_{ij} + w_j^d \right) \right], \tag{17}$$

where λ_i^d is the dual variable corresponding to the equality constraint $\sum_j p_{ij}^d = 1$. The choice probabilities in this case are easily computed by an efficient line search method over the single variable λ_i^d using the condition (17). Alternatively, the minimum expected cost and the optimal value of the dual variable is obtained by solving the following concave maximization problem:

$$w_i^d = \max_{\lambda_i^d} \left\{ -\lambda_i^d - \sum_{j \in N_d^+(i)} \int_{\lambda_i^d + t_{ij} + w_j^d}^\infty \left[1 - F_{ij}^d(\omega) \right] d\omega \right\}. \tag{18}$$

It is easy to verify again that under the assumptions, there exists a unique λ_i^d satisfying the normalization condition (17), which is an optimality condition that arises from (18).

We now generalize the result to the Markovian Marginal Distribution Model (M-MDM) as follows.

Definition 1. *Given the link cost vector $\boldsymbol{t}$, the link choice probabilities of M-MDM, for the users moving towards destination d are characterized as the solution of the following system of equations in the variables w_i^d:*

$$w_i^d = \min_{\theta_{id} \in \Theta_{id}} E_{\theta_{id}} \left[\min_{j \in N_d^+(i)} \left\{ t_{ij} - \widetilde{\epsilon}_{ij}^d + w_j^d \right\} \right] \qquad \forall i \in N_d, \qquad (19)$$

$$w_d^d = 0, \qquad (20)$$

where:

$$p_{ij}^d = P_{\theta_{id}^*} \left\{ j = \underset{k \in N_d^+(i)}{\operatorname{argmin}} \left\{ t_{ik} - \widetilde{\epsilon}_{ik}^d + w_k^d \right\} \right\}, \qquad \forall (i,j) \in A_d, \qquad (21)$$

and θ_{id}^* is the extremal distribution in (19).

It is important to note that there exists at most one solution for the M-MDM system. To see this, note that the right-hand-side of the M-MDM recursion (19), and hence, w_i^d, is componentwise nondecreasing in $w_j^d, \forall j \in N_d^+(i)$. Under Assumption A3, it is componentwise increasing. Suppose that $\mathbf{w^{d(1)}}$ and $\mathbf{w^{d(2)}}$ are two solutions. Let $\delta_i^d = w_i^{d(1)} - w_i^{d(2)}$, and let $\delta = \max_{i \in N_d} \{\delta_i^d\} \geq 0$. Let $i \in N_d$ be a node where the maximum difference is attained, i.e., $w_i^{d(1)} - w_i^{d(2)} = \delta$. By definition of δ, we have $\delta_j^d \leq \delta, \forall j \in N_d^+(i)$. Strict monotonicity of w_i^d in w_j^d implies that $\delta_j^d = \delta, \forall j \in N_d^+(i)$. Moving forward in the same manner, we can show that $\delta_j^d = \delta$ for all j reachable from node i including the destination. Then, the boundary condition implies that $\delta = 0$. Thus we have $\delta_i^d = 0, \forall i \in N_d$, implying that there is at most one solution.

Next, we propose an equivalent convex optimization reformulation to solve the system of equations in M-MDM and compute the probabilities.

Proposition 1. *Given a link cost vector $\boldsymbol{t}$, the conditional link choice probabilities of the users moving to destination d in M-MDM is calculated as:*

$$p_{ij}^d = 1 - F_{ij}^d \left(\lambda_i^d + t_{ij} + w_j^d \right), \qquad \forall (i,j) \in A_d,$$

where the values of λ_i^d and w_i^d are obtained from the optimal solution to the following convex optimization problem:

$$Z_d^* (t) = \max_{w^d, \lambda^d} \sum_{i \in N_d} h_i^d \cdot w_i^d \qquad (22)$$

$$s.t. \ w_i^d \leq -\lambda_i^d - \sum_{j \in N_d^+(i)} \int_{\lambda_i^d + t_{ij} + w_j^d}^{\infty} \left[1 - F_{ij}^d(\omega)\right] d\omega, \quad \forall i \in N_d, \qquad (23)$$

$$w_d^d = 0. \qquad (24)$$

Proof. The Lagrangian function of the formulation (22)-(24) is given below:

$$L = \sum_{i \in N_d} h_i^d \cdot w_i^d - \sum_{i \in N_d} w_i^d \cdot n_i^d - \sum_{i \in N_d} \lambda_i^d \cdot n_i^d$$
$$- \sum_{(i,j) \in A_d} n_i^d \int_{\lambda_i^d + t_{ij} + w_j^d}^{\infty} \left[1 - F_{ij}^d(\omega)\right] d\omega,$$

where n_i^d is the dual variable corresponding to constraint (23). The KKT optimality conditions are as follows:

$$\frac{\partial L}{\partial \lambda_i^d} = 0 \implies n_i^d = n_i^d \sum_{j \in N_d^+(i)} \left[1 - F_{ij}^d\left(\lambda_i^d + t_{ij} + w_j^d\right)\right], \qquad \forall i \in N_d,$$

$$\frac{\partial L}{\partial w_i^d} = 0 \implies n_i^d = h_i^d + \sum_{k \in N_d^-(i)} n_k^d \left[1 - F_{ki}^d\left(\lambda_k^d + t_{ki} + w_i^d\right)\right], \quad \forall i \in N_d.$$

For any origin node o, i.e., $(o, d) \in W$, we have $h_o^d > 0$ and, therefore, the second KKT condition implies that $n_o^d > 0$. Then, the summation in the right-hand-side of the first KKT condition (for node o) must be equal to one. In other words, λ_o^d solves the MDM optimality condition (17). Since MDM choice probabilities are strictly positive, there will be a positive flow into each node that is adjacent to o. The same reasoning can be used to show that $n_i^d > 0, \forall i \in N_d$ and to conclude that all λ_i^d's solve the MDM optimality conditions (17) for all $\forall i \in N_d$. The right-hand-side of constraint (23) is componentwise increasing in w_j^d. Monotonicity and the fact that we are maximizing the objective function imply that the constraint holds at equality for the optimal solution. Then, the optimal values of w_i^d and λ_i^d solve the M-MDM system of equations (19)-(21). $\square$

The KKT optimality conditions of the M-MDM formulation are equivalent to the constraints of a network flow problem where h_i^d units enter the network at node $i \in N_d$, and $\sum_{i \in N_d} h_i^d$ units leave the network from node d. Let x_{ij}^d be the flow on link (i, j), and n_i^d be the outflow from node i. Using the substitution $x_{ij}^d = n_i^d \cdot p_{ij}^d$, the KKT optimality conditions can be rewritten as follows:

$$\sum_{j \in N_d^+(i)} x_{ij}^d - \sum_{k \in N_d^-(i)} x_{ki}^d = h_i^d, \qquad \forall (i, j) \in A_d.$$

The choice process can be considered as a Markov chain where nodes are the states and destination d is the single absorbing state. Let $\mathbf{Q^d}$ be an $|N_d| \times |N_d|$ matrix where $(i, j)^{th}$ entry of the matrix, $\mathbf{Q_{ij}^d}$, is equal to p_{ij}^d if $(i, j) \in A_d$ and zero otherwise. Since $p_{ij}^d > 0$, $\mathbf{Q^d}$ is sub-stochastic; and $\mathbf{I} - \mathbf{Q^d}$ is nonsingular (Kemeny et al., 1960), where $\mathbf{I}$ is the identity matrix of size $|N_d|$. The fundamental matrix of the Markov chain is $\mathbf{M^d} = \sum_{n=0}^{\infty} (\mathbf{Q^d})^n = (\mathbf{I} - \mathbf{Q^d})^{-1}$, where $\mathbf{M_{ij}^d}$ is the expected number of times a user entering the system at node i visits node j before reaching the destination. Then, n_i^d can be expressed as follows:

$$n_i^d = \sum_{o \in N_d} h_o^d \cdot \mathbf{M_{oi}^d}, \qquad \forall i \in N_d. \qquad (25)$$

Now consider the M-MDM recursion (19). Given optimal link choice probabilities p_{ij}^d, the expression can be rewritten using the primal MDM formulation (15) as follows:

$$\underline{w}_i^d = \sum_{j \in N_d^+(i)} \left(t_{ij} \cdot p_{ij}^d - \int_{1 - p_{ij}^d}^1 F_{ij}^{d^{(-1)}}(\omega) \, d\omega \right), \qquad \forall i \in N_d,$$

$$w_i^d = \underline{w}_i^d + \sum_{j \in N_d^+(i)} p_{ij}^d \cdot w_j^d, \qquad \forall i \in N_d.$$

Here, $\underline{w}_i^d$ corresponds to the instantaneous cost and the second term calculates the expected cost from the downstream nodes. In matrix form, the system can be expressed as $\mathbf{W^d} = \underline{\mathbf{W}}^\mathbf{d} + \mathbf{Q^d W^d}$ which is equivalent to the following expression:

$$w_i^d = \sum_{j \in N_d} \mathbf{M_{ij}^d} \cdot \underline{w}_j^d \qquad \forall i \in N_d. \qquad (26)$$

13

Using equations (25) and (26), the optimal objective function value of M-MDM given the link choice probabilities can be expressed as follows:

$$
\sum_{i \in N_d} h_i^d \cdot w_i^d = \sum_{i \in N_d} h_i^d \sum_{j \in N_d} \mathbf{M}_{ij}^{\mathbf{d}} \cdot \underline{w}_j^d
$$

$$
= \sum_{j \in N_d} \underline{w}_j^d \sum_{i \in N_d} h_i^d \cdot \mathbf{M}_{ij}^{\mathbf{d}}
$$

$$
= \sum_{j \in N_d} n_j^d \cdot \underline{w}_j^d
$$

$$
= \sum_{i \in N_d} n_i^d \cdot \underline{w}_i^d
$$

$$
= \sum_{i \in N_d} n_i^d \sum_{j \in N_d^+(i)} \left(t_{ij} \cdot p_{ij}^d - \int_{1-p_{ij}^d}^1 F_{ij}^{d\,(-1)}(\omega)\, d\omega \right)
$$

$$
= \sum_{i \in N_d} \sum_{j \in N_d^+(i)} \left(t_{ij} \cdot x_{ij}^d - n_i^d \int_{1-\frac{x_{ij}^d}{n_i^d}}^1 F_{ij}^{d\,(-1)}(\omega)\, d\omega \right).
$$

Using the KKT conditions and the alternative expression of the objective function, the M-MDM problem can be reformulated in flow variables as we show in the next proposition.

Proposition 2. *Given a link cost vector $\boldsymbol{t}$, the link choice probabilities of the users moving towards destination d in M-MDM is calculated by:*

$$
p_{ij}^d = \frac{x_{ij}^d}{n_i^d}, \qquad\qquad \forall (i,j) \in A_d,
$$

where the unique values of x_{ij}^d and n_i^d are obtained from the optimal solution to the following convex optimization problem:

$$
Z_d^*(\boldsymbol{t}) = \min_{\boldsymbol{x}^d, \boldsymbol{n}^d} \sum_{(i,j) \in A_d} \left(t_{ij} \cdot x_{ij}^d - n_i^d \int_{1-\frac{x_{ij}^d}{n_i^d}}^1 F_{ij}^{d\,(-1)}(\omega)\, d\omega \right) \qquad (27)
$$

$$
s.t. \sum_{j \in N_d^+(i)} x_{ij}^d = n_i^d, \qquad \forall i \in N_d, \qquad (28)
$$

$$
\sum_{k \in N_d^-(i)} x_{ki}^d + h_i^d = n_i^d, \qquad \forall i \in N_d, \qquad (29)
$$

$$
x_{ij}^d \geq 0, \qquad\qquad \forall (i,j) \in A_d. \qquad (30)
$$

Proof. Let $\varphi(\mathbf{x^d}, \mathbf{n^d}) = \sum_{(i,j) \in A_d} \varphi_{ij}(x_{ij}^d, n_i^d)$ represent the objective function where:

$$\varphi_{ij}(x_{ij}^d, n_i^d) = \left(t_{ij} \cdot x_{ij}^d - n_i^d \int_{1-\frac{x_{ij}^d}{n_i^d}}^{1} F_{ij}^{d\,(-1)}(\omega) d\omega \right).$$

This is the perspective of the function $\varphi_{ij}^0(x)$ defined as:

$$\varphi_{ij}^0(x) = \left(t_{ij} \cdot x - \int_{1-x}^{1} F_{ij}^{d\,(-1)}(\omega) d\omega \right),$$

such that $\varphi_{ij}(x_{ij}^d, n_i^d) = n_i^d \varphi_{ij}^0(x_{ij}^d / n_i^d)$. Since the function $\varphi_{ij}^0(x)$ is strictly convex under the assumptions on the marginal distribution, the perspective of the function denoted by $\varphi_{ij}(x_{ij}^d, n_i^d)$ is also strictly convex. This implies that $\varphi(\mathbf{x^d}, \mathbf{n^d})$ is strictly convex. Furthermore the constraints are convex and hence it is a convex optimization problem with a unique optimal solution. The KKT optimality conditions of the formulation are given below:

$$\frac{x_{ij}^d}{n_i^d} = 1 - F_{ij}^d \left(\lambda_i^d + t_{ij} + w_j^d \right), \qquad\qquad \forall (i,j) \in A_d,$$

$$w_i^d = -\lambda_i^d - \sum_{j \in N_d^+(i)} \left(\int_{1-\frac{x_{ij}^d}{n_i^d}}^{1} F_{ij}^{d\,(-1)}(\omega) d\omega - \frac{x_{ij}^d}{n_i^d} F_{ij}^{d\,(-1)} \left(1 - \frac{x_{ij}^d}{n_i^d} \right) \right)$$

$$= -\lambda_i^d - \sum_{j \in N_d^+(i)} \int_{\lambda_i^d + t_{ij} + w_j^d}^{\infty} \left[1 - F_{ij}^d (\omega) \right] d\omega, \qquad\qquad \forall i \in N_d,$$

where λ_i^d and w_i^d are the dual variables corresponding to constraints (28) and (29), respectively. By constraint (28), λ_i^d satisfies the normalization condition (17). Therefore, the KKT conditions are equivalent to the M-MDM system of equations.

$\square$

2.1 Special Case: Multinomial Logit and Extensions

In this section, we show that for special instances of the marginal distributions, the link flows in M-MDM reduces to results for the multinomial logit model and some of its extensions. We do this by showing the equivalence of a special instance of Proposition 2 with an entropy maximization problem that was first proposed in Akamatsu (1996) for the Markovian logit

model. Under the assumption of independent and identical Gumbel random variables, Akamatsu (1996) provided a convex optimization formulation to compute the link flows in the Markovian logit model. For a given destination d, their formulation is given as:

$$Z_d\left(\boldsymbol{\theta}_{\mathbf{d}}^{(G)}, \mathbf{t}\right) = \min_{\mathbf{x}} \sum_{(i,j)\in A_d} x_{ij}^d \cdot t_{ij} + \frac{1}{\beta} \sum_{(i,j)\in A_d} x_{ij}^d \cdot \ln x_{ij}^d$$

$$- \frac{1}{\beta} \sum_{i\in N_d} \left(\sum_{j\in N_d^+(i)} x_{ij}^d\right) \cdot \ln\left(\sum_{j\in N_d^+(i)} x_{ij}^d\right) \qquad (31)$$

$$\text{s.t.} \sum_{j\in N_d^+(i)} x_{ij}^d - \sum_{k\in N_d^-(i)} x_{ki}^d = h_i^d, \qquad \forall i \in N_d, \quad (32)$$

$$\sum_{i\in N_d^-(d)} x_{id}^d = \sum_{i\in N_d} h_i^d, \qquad (33)$$

$$x_{ij}^d \geq 0, \qquad \forall (i,j) \in A_d. \qquad (34)$$

We now provide an alternative derivation of this model from our result.

Proposition 3. *Let $\boldsymbol{\theta}_{\mathbf{d}}^{(G)}$ be the joint distribution of independent and identical Gumbel random error terms with dispersion parameter β; and $\boldsymbol{\Theta}_{\mathbf{d}}^{(E)}$ be the set of all joint distributions of the random error terms with exponential marginals having location parameter $-1/\beta$ and scale parameter $1/\beta$. The cumulative distribution function of the marginal distribution of the random error term $\widetilde{\epsilon}_{ij}^d$ in $\boldsymbol{\Theta}_{\mathbf{d}}^{(E)}$ is given as:*

$$F_{ij}^d(\omega) = 1 - \exp\left(-1 - \beta\omega\right), \qquad \forall \omega \geq -\frac{1}{\beta}, (i,j) \in A_d.$$

Then, the M-MDM formulation given in (27)-(30) reduces to the Markovian logit formulation given in (31)-(34) with

$$Z_d^{*(E)}(\boldsymbol{t}) = Z_d\left(\boldsymbol{\theta}_{\boldsymbol{d}}^{(G)}, \boldsymbol{t}\right),$$

where $Z_d^{*(E)}(\boldsymbol{t})$ is the optimal objective value in (27) with exponential marginals.

Proof. The exponential distribution satisfies the assumption on marginals, A3. Substituting the inverse cumulative distribution function $F_{ij}^{d(-1)}(\omega) =$

16

$-(1 + \ln(1 - \omega))/\beta$ in the objective function (27), the integral is calculated as follows:

$$\int_{1 - \frac{x_{ij}^d}{n_i^d}}^{1} F_{ij}^{d\,(-1)}(\omega)d\omega = \frac{1}{\beta n_i^d} \left(-x_{ij}^d \cdot \ln x_{ij}^d + x_{ij}^d \cdot \ln n_i^d \right),$$

and the objective function simplifies to the following:

$$\sum_{(i,j) \in A_d} \left(t_{ij} \cdot x_{ij}^d + \frac{x_{ij}^d \cdot \ln x_{ij}^d}{\beta} - \frac{x_{ij}^d \cdot \ln n_i^d}{\beta} \right).$$

Constraints (28) and (29) can be expressed as a single constraint by substituting n_i^d in the objective function. Using substitution $n_i^d = \sum_{j \in N_d^+(i)} x_{ij}^d$, the last term in the objective function can be expressed as follows:

$$-\frac{1}{\beta} \sum_{(i,j) \in A_d} x_{ij}^d \cdot \ln n_i^d = -\frac{1}{\beta} \sum_{i \in N_d} \sum_{j \in N_d^+(i)} x_{ij}^d \cdot \ln \left(\sum_{j \in N_d^+(i)} x_{ij}^d \right)$$

$$= -\frac{1}{\beta} \sum_{i \in N_d} \left(\sum_{j \in N_d^+(i)} x_{ij}^d \right) \cdot \ln \left(\sum_{j \in N_d^+(i)} x_{ij}^d \right).$$

Then, the M-MDM formulation reduces to the Markovian logit model given in (31)-(34). $\square$

Similarly, it is possible to show the equivalence of the model in the expected cost variables with the formulation for the Markovian logit model studied in Fosgerau et al. (2013) and given by the equations (4)-(5). Let λ_i^d be the optimal dual variable. Substituting the cumulative distribution function, we obtain the following choice probability expression for exponential marginals:

$$p_{ij}^d = \exp\left(-1 - \beta \lambda_i^d \right) \cdot \exp\left(-\beta \left(t_{ij} + w_j^d \right) \right), \qquad \forall (i,j) \in A_d.$$

The normalization condition (17) provides a closed form expression for λ_i^d:

$$\sum_{j \in N_d^+(i)} p_{ij}^d = 1 \implies \lambda_i^d = \frac{1}{\beta} \ln \left(\sum_{j \in N_d^+(i)} \exp\left(-\beta \left(t_{ij} + w_j^d \right) \right) \right) - \frac{1}{\beta}.$$

Substituting λ_i^d in the choice probability expression, we obtain:

$$p_{ij}^d = \frac{\exp\left(-\beta\left(t_{ij} + w_j^d\right)\right)}{\sum_{k \in N_d^+(i)} \exp\left(-\beta\left(t_{ik} + w_k^d\right)\right)}, \qquad \forall(i,j) \in A_d.$$

Finally, substituting the cumulative distribution function in the recursive equation, we obtain the following:

$$w_i^d = -\lambda_i^d - \frac{\exp\left(-1 - \beta\lambda_i^d\right)}{\beta} \sum_{j \in N_d^+(i)} \exp\left(-\beta\left(t_{ij} + w_j^d\right)\right)$$

$$= -\frac{1}{\beta} \ln\left(\sum_{j \in N_d^+(i)} \exp\left(-\beta\left(t_{ij} + w_j^d\right)\right)\right),$$

which is exactly the expected minimum cost expression of the Markovian logit model given in (4).

Corollary 1. *When the marginal distribution of the random error term $\widetilde{\epsilon}_{ij}^d$ is exponential with location parameter $-1/\beta_i$ and scale parameter $1/\beta_i$, the M-MDM system of equations given in (19)-(20) reduces to the system of equations in the Nested Recursive Logit model proposed by Mai et al. (2015).*

The proof follows directly from Proposition 3 by using node specific dispersion parameters β_i instead of identical β. The Nested Recursive Logit model aims to model the correlation among the links by allowing heterogeneity in the node level. M-MDM generalizes this model by relaxing the indentically distributed assumption and allows heterogeneity in the link level.

2.2 Special Case: Acyclic Network

In this section, we compare the Markovian model with an alternate link based model that was proposed in Natarajan et al. (2009) for acyclic networks. To discuss the model, we focus on a single origin-destination pair (o, d) in a directed acyclic network with supply of 1 at node o and a demand of 1 at node d. The link cost for arc (i, j) is a random variable and is equal to $t_{ij} - \widetilde{\epsilon}_{ij}$. We make the following assumptions as in Natarajan et al. (2009):

B1 Error terms of the links emanating from the different nodes are arbitrarily dependent.

B2 The joint distribution θ of the random error terms $\{\widetilde{\epsilon}_{ij}; (i,j) \in A_d\}$ is incompletely specified and only known to lie in a set of distributions with given marginal distributions. This set of distributions is designated as Θ.

In Natarajan et al. (2009), the authors study the problem of finding the minimum expected shortest path from node o to node d which is formulated as follows:

$$\min_{\theta \in \Theta} E_\theta \left[\min_{x \in X} \sum_{(i,j) \in A_d} (t_{ij} - \widetilde{\epsilon}_{ij}) \cdot x_{ij} \right], \tag{35}$$

where the set X is the flow polytope defined as:

$$X = \left\{ x : \sum_{j \in N_d^+(i)} x_{ij} - \sum_{k \in N_d^-(i)} x_{ki} = \left\{ \begin{array}{ll} 1, & i = o \\ -1, & i = d \\ 0, & \text{otherwise} \end{array} \right. , \ x_{ij} \geq 0, \forall (i,j) \in A_d \right\}.$$

Under assumptions B1, B2 and A3, Natarajan et al. (2009) derived a convex reformulation for this problem in the link flows as follows:

$$\min_{\mathbf{x}} \left\{ \sum_{(i,j) \in A_d} \left(t_{ij} \cdot x_{ij} - \int_{1-x_{ij}}^{1} F_{ij}^{(-1)}(\omega) d\omega \right) : x \in X \right\}. \tag{36}$$

On the other hand under the Markovian assumption, the choice probabilities for a acyclic network from Proposition 2 is obtained by solving for the link flows as:

$$\min_{\mathbf{x}} \left\{ \sum_{(i,j) \in A_d} \left(t_{ij} \cdot x_{ij} - \left(\sum_{j \in N_d^+(i)} x_{ij} \right) \int_{1 - \frac{x_{ij}}{\sum_{j \in N_d^+(i)} x_{ij}}}^{1} F_{ij}^{(-1)}(\omega) d\omega \right) : x \in X \right\}. \tag{37}$$

Let the objective function of the model (36) be $\sum_{(i,j) \in A_d} \varphi_{ij}^0(x_{ij})$ and the objective function of the model (37) be $\sum_{(i,j) \in A_d} \varphi_{ij}^*(x_{ij}, n_i)$ where $n_i = \sum_{j \in N_d^+(i)} x_{ij}$, respectively. These models are referred to as MDM and M-MDM respectively. Table 1 summarizes the differences in the cost functions for different marginal distributions. The abbreviations MEM, MNM and MLM are used for the MDM model with exponential, normal and logistic marginals, respectively. The $\Phi(.)$ is the standard normal cumulative distribution function and erf(.) is the error function. Since the network is acyclic, we have $n_i \leq 1$; and hence $\ln n_i \leq 0, \forall i \in N$.

Table 1: Cost terms of the MDM and M-MDM.

Marginal Distribution	Objective function term	
MEM / M-MEM	$\varphi_{ij}^{(0)} =$	$(t_{ij} - A_{ij} - B_{ij})\, x_{ij} +$
$F_{ij}(\omega) = 1 - \exp\left(\frac{A_{ij}-\omega}{B_{ij}}\right)$		$B_{ij} x_{ij} \ln x_{ij}$
A_{ij}: location parameter	$\varphi_{ij}^{*} =$	$(t_{ij} - A_{ij} - B_{ij})\, x_{ij} +$
B_{ij}: scale parameter		$B_{ij} x_{ij} \ln x_{ij} - B_{ij} x_{ij} \ln n_i$
MNM / M-MNM	$\varphi_{ij}^{(0)} =$	$(t_{ij} - \mu_{ij})\, x_{ij}$
$F_{ij}(\omega) = \Phi\left(\frac{\omega - \mu_{ij}}{\sigma_{ij}}\right)$		$-\frac{\sigma_{ij}}{\sqrt{2\pi}} \exp\left(-\left[\mathrm{erf}^{-1}\left(1 - 2x_{ij}\right)\right]^2\right)$
μ_{ij}: mean	$\varphi_{ij}^{*} =$	$(t_{ij} - \mu_{ij})\, x_{ij}$
σ_{ij}: standard deviation		$-\frac{n_i \sigma_{ij}}{\sqrt{2\pi}} \exp\left(-\left[\mathrm{erf}^{-1}\left(1 - \frac{2x_{ij}}{n_i}\right)\right]^2\right)$
MLM / M-MLM	$\varphi_{ij}^{(0)} =$	$(t_{ij} - \alpha_{ij})\, x_{ij} + \frac{x_{ij} \ln x_{ij}}{\eta_{ij}}$
$F_{ij}(\omega) = \frac{1}{1 + \exp[-\eta_{ij}(\omega - \alpha_{ij})]}$		$+\frac{(1 - x_{ij}) \ln(1 - x_{ij})}{\eta_{ij}}$
α_{ij}: location parameter	$\varphi_{ij}^{*} =$	$(t_{ij} - \alpha_{ij})\, x_{ij} + \frac{x_{ij} \ln x_{ij}}{\eta_{ij}}$
η_{ij}: rate parameter		$+\frac{(n_i - x_{ij}) \ln(n_i - x_{ij})}{\eta_{ij}} - \frac{n_i \ln n_i}{\eta_{ij}}$

3 Markovian Traffic Equilibrium with MDM (MTE-MDM)

In Section 2, we have described the choice model M-MDM and provided convex optimization formulations to compute link flows for a given link cost vector. In congested traffic networks however users' choices affect the link costs and hence, link choices towards different destinations are interdependent. The equilibrium flow is then obtained as the solution to a fixed-point problem defined as follows.

Definition 2. *Let $p_{ij}^d(\boldsymbol{t})$ be the link choice probabilities obtained by M-MDM for given link cost vector $\boldsymbol{t}$. The Markovian traffic equilibrium under the Marginal Distribution Model, termed as MTE-MDM, is characterized as the solution to the following fixed-point problem in the variables $(f_{ij}, t_{ij}, x_{ij}^d, n_i^d, p_{ij}^d)$:*

$$n_i^d = h_i^d + \sum_{k \in N_d^-(i)} x_{ki}^d, \qquad \forall i \in N_d, d \in D, \tag{38}$$

$$x_{ij}^d = n_i^d \cdot p_{ij}^d(\boldsymbol{t}), \qquad \forall (i,j) \in A_d, d \in D, \tag{39}$$

$$f_{ij} = \sum_{d \in D:(i,j) \in A_d} x_{ij}^d, \qquad \forall (i,j) \in A, \tag{40}$$

$$t_{ij} = \tau_{ij}(f_{ij}), \qquad \forall (i,j) \in A. \tag{41}$$

Conditions (38)-(39) are equivalent to the flow conservation constraints (28)-(29) in M-MDM. Therefore, the M-MDM formulation given in (27)-(30) can be naturally extended to handle congestion by using the link cost functions. We make the following additional assumption in this section for the congested network case.

A4 The link cost function $\tau_{ij}(.)$ is strictly increasing with $\tau_{ij}(f) > t_{ij}^0 \geq 0$, for any positive flow $f > 0$; where t_{ij}^0 is the free-flow travel cost of link (i,j).

One such function that is commonly used is the Bureau of Public Roads (BPR) cost function which accounts for congestion as follows:

$$\tau_{ij}(f_{ij}) = t_{ij}^0 \left[1 + \alpha_{ij} \left(\frac{f_{ij}}{q_{ij}} \right)^{\gamma_{ij}} \right], \quad \forall (i,j) \in A,$$

where q_{ij} is the capacity, and α_{ij} and γ_{ij} are constants. Therefore, f_{ij}/q_{ij} serves as a measure of congestion.

Proposition 4. *Under assumptions A1-A4, the equilibrium link flows for the MTE-MDM is obtained as the unique optimal solution to the following convex optimization problem:*

$$Z^* = \min_{\boldsymbol{x,n,f}} \sum_{(i,j)\in A} \int_0^{f_{ij}} \tau_{ij}(\omega)d\omega - \sum_{d\in D} \sum_{(i,j)\in A_d} n_i^d \int_{1-\frac{x_{ij}^d}{n_i^d}}^1 F_{ij}^{d\,(-1)}(\omega)d\omega \quad (42)$$

$$s.t. \sum_{j\in N_d^+(i)} x_{ij}^d = n_i^d, \qquad \forall i \in N_d, d \in D, \quad (43)$$

$$\sum_{k\in N_d^-(i)} x_{ki}^d + h_i^d = n_i^d, \quad \forall i \in N_d, d \in D, \quad (44)$$

$$\sum_{d\in D:(i,j)\in A_d} x_{ij}^d = f_{ij}, \quad \forall (i,j) \in A, \quad (45)$$

$$x_{ij}^d \geq 0, \qquad \forall (i,j) \in A_d, d \in D. \quad (46)$$

Furthermore, the link choice probabilities at equilibrium are obtained as follows:

$$p_{ij}^d = \frac{x_{ij}^d}{n_i^d}, \qquad \forall (i,j) \in A_d, d \in D.$$

Proof. Let λ_i^d, w_i^d and t_{ij} be the dual variables corresponding to constraints (43), (44) and (45), respectively. Then, the KKT optimality conditions of the convex optimization problem are as follows:

$$t_{ij} = \tau_{ij}(f_{ij}), \qquad\qquad\qquad \forall (i,j) \in A,$$

$$\frac{x_{ij}^d}{n_i^d} = 1 - F_{ij}^d \left(\lambda_i^d + t_{ij} + w_j^d \right), \qquad\qquad \forall (i,j) \in A_d, d \in D,$$

$$w_i^d = -\lambda_i^d - \sum_{j\in N_d^+(i)} \int_{\lambda_i^d+t_{ij}+w_j^d}^{\infty} \left[1 - F_{ij}^d(\omega) \right] d\omega, \quad \forall i \in N_d, d \in D.$$

The last two optimality conditions solve the M-MDM system of equations for a given link cost vector. Then, conditions (40) and (41) are satisfied since $x_{ij}^d/n_i^d = p_{ij}(\mathbf{t})$, where $t_{ij} = \tau_{ij}(f_{ij})$ by the first KKT condition. Conditions (38) and (39) are satisfied by the constraints of the model. Therefore, the optimal solution of the formulation (42)-(45) solves the MTE-MDM fixed-point problem. The constraints of the model are linear, and the objective function is strictly convex and hence, the solution of the model is unique. $\quad\square$

The MTE-MDM flows can alternatively be obtained by solving a convex optimization problem using the cost variables.

$$Z^* = \max_{\mathbf{w}, \boldsymbol{\lambda}, \mathbf{t}} \sum_{d \in D} \sum_{i \in N_d} h_i^d \cdot w_i^d - \sum_{(i,j) \in A} \int_0^{t_{ij}} \tau_{ij}^{-1}(\omega) \, d\omega \tag{47}$$

$$\text{s.t. } w_i^d \le -\lambda_i^d - \sum_{j \in N_d^+(i)} \int_{\lambda_i^d + t_{ij} + w_j^d}^{\infty} \left[1 - F_{ij}^d(\omega) \right] d\omega, \quad \forall i \in N_d, d \in D, \tag{48}$$

$$w_d^d = 0, \qquad\qquad\qquad\qquad\qquad \forall d \in D. \tag{49}$$

4   Solution Algorithm

In Sections 2 and 3, we proposed convex optimization models for the choice and equilibrium models. This implies that it is possible to use convex optimization solvers for this problem depending on the choice of the marginal distribution, or standard algorithms such as the Frank-Wolfe algorithm can be implemented to obtain the solutions. In this section, we show that a fixed-point iteration can be used to find the choice probabilities for a given link cost vector. These iterations include a step to solve the MDM problem using a line search method over the single dual variable. The fixed-point algorithm can be integrated as the stochastic network loading procedure within a method of successive averages (MSA) algorithm to find the equilibrium flow. The reason of our choice for this solution approach is threefold. First, the fixed-point algorithm provides an efficient sub-procedure for the choice model and is easy to extend to different marginal distributions. Second, MSA is one of the popular approaches used by system planners to solve traffic equilibrium problems. We show how the new choice model can be integrated into this approach. Third, the proposed approach allows to solve route based and Markovian models for any marginal setting under the same framework, allowing for a fair comparison of the convergence performance.

Consider the Markovian choice model for a given link cost vector and a particular destination d. For ease of exposition, we drop the desination index from the cost variables w_i^d. We define the function $\varphi_i(\mathbf{w})$ as follows:

$$\varphi_i(\mathbf{w}) = \min_{\theta_{id} \in \Theta_{id}} E_{\theta_{id}} \left[\min_{j \in N_d^+(i)} \left\{ t_{ij} - \widetilde{\epsilon}_{ij}^d + w_j \right\} \right], \qquad \forall i \in N_d,$$

where w_d is fixed to zero. Since the solution of the M-MDM system is unique, there exists a unique fixed-point $\boldsymbol{\varphi}(\mathbf{w^*}) = \mathbf{w^*}$. Next, consider the

23

sequence $\mathbf{w^n}$ defined as follows:

$$
\begin{aligned}
w_d^n &= 0, & \forall n \in \{0, 1, \dots\}, \\
w_i^0 &= s_i^d, & \forall i \in N_d, \\
w_i^n &= \varphi_i\left(\mathbf{w^{n-1}}\right), & \forall i \in N_d, n \in \{1, 2, \dots\},
\end{aligned}
$$

where s_i^d is the shortest path distance from node i to destination d with repsect to arc costs $t_{ij} - E[\tilde{\epsilon}_{ij}^d]$. The following lemma shows that the sequence monotonically converges to the fixed point.

Lemma 1. *Consider the unique fixed-point $\boldsymbol{w^*}$ and the sequence $\boldsymbol{w^n}$ defined above. The following inequalities hold:*

$$
\begin{aligned}
w_i^* &\leq s_i^d, & \forall i \in N_d, \\
w_i^n &\leq w_i^{n-1}, & \forall i \in N_d, n \in \{1, 2, \dots\}.
\end{aligned}
$$

Then, the sequence $\boldsymbol{w^n}$ converges to $\boldsymbol{w^*}$.

Proof. The M-MDM recursion (19) implies that $w_i \leq t_{ij} - E[\tilde{\epsilon}_{ij}^d] + w_j, \forall j \in N_d^+(i)$. This inequality recursively implies that s_i^d is an upper bound for w_i^*. We can make the induction assumption that $w_i^n \leq w_i^{n-1}, \forall n \in \{1, \dots, m-1\}, i \in N_d$, since $w_i^1 \leq w_i^0 = s_i^d, \forall i \in N_d$. By this assumption, we have $t_{ij} + w_j^{m-1} \leq t_{ij} + w_j^{m-2}, \forall (i,j) \in A_d$. The monotonicity of w_i in $w_j, \forall j \in N_d^+(i)$, then implies that $w_i^m \leq w_i^{m-1}, \forall i \in N_d$. By induction, the second inequality holds for all n. Thus, the sequence is monotonic nonincreasing, moreover bounded from below provided that the M-MDM solution exists. This proves that the sequence converges to the unique fixed-point $\mathbf{w^*}$. $\qquad\square$

The corresponding fixed-point algorithm is given in Algorithm 1. It solves the MDM for each node independently at each iteration. This is equivalent to finding the unique value of the dual variable $\lambda_i^{d(n)}$ at line 7 satisfying the normalization condition, which can be achieved by efficient line search algorithms. In our experiments, we use the following stopping condition with sufficiently small ϵ_1:

$$
\max_{i \in N_d} \left\{ \frac{w_i^{d(n-1)} - w_i^{d(n)}}{w_i^{d(n-1)}} \right\} \leq \epsilon_1.
$$

The loading algorithm M-MDM$(\mathbf{t}, d)$ finds the unique link choice probabilities for given $\mathbf{t}$ and destination d. Given the choice probabilities, link

---
Algorithm 1 M-MDM Loading Algorithm

Input: $\mathbf{t}; d$

 1: **procedure** M-MDM$(\mathbf{t}, d)$

 2:     $w_i^{d(0)} = s_i^d, \forall i \in N_d$.

 3:     $n = 0$.

 4:     **while** stopping condition is not satisfied **do**

 5:         $n = n + 1$.

 6:         **for** each $i \in N_d$ **do**

 7:            find $\lambda_i^{d(n)}$ s.t. $\sum_{j \in N_d^+(i)} \left[1 - F_{ij}^d \left(\lambda_i^{d(n)} + t_{ij} + w_j^{d(n-1)} \right) \right] = 1$.

 8:            $w_i^{d(n)} = -\lambda_i^{d(n)} - \sum_{j \in N_d^+(i)} \int_{\lambda_i^{d(n)} + t_{ij} + w_j^{d(n-1)}}^{\infty} \left[1 - F_{ij}^d(\omega) \right] d\omega$.

 9:         **end for**

10:     **end while**

11:     $p_{ij}^d = 1 - F_{ij}^d \left(\lambda_i^{d(n)} + t_{ij} + w_j^{d(n)} \right)$

12: **end procedure**

Output: $\mathbf{p_d} = \left(p_{ij}^d \right)_{(i,j) \in A_d}$.

---

flows can be obtained by Markovian operations. Then a standard MSA algorithm can be implemented, which is represented in Algorithm 2. We use the following stopping condition in our experiments:

$$rmse = \sum_{(i,j) \in A} \left(f_{ij}^{(n)} - f_{ij}^{(n-1)} \right)^2 \leq \epsilon_2.$$

5   Computational Study

In this section, we investigate the user choice behavior in M-MDM, and equilibrium flows in MTE-MDM for small and large traffic networks[1].

5.1   Choice Behavior under M-MDM

Consider the four-node network in Figure 2 taken from Akamatsu (1996) with a single destination 4, where the numbers next to the arcs represent the deterministic link costs. Since there is a single destination, we drop the destination index d throughout this section. Note that there are infinitely

---

[1]All experiments are carried out with the open-source software, seSue, provided in http://people.sutd.edu.sg/~ugur_arikan/seSue/.

Algorithm 2 MTE-MDM Algorithm
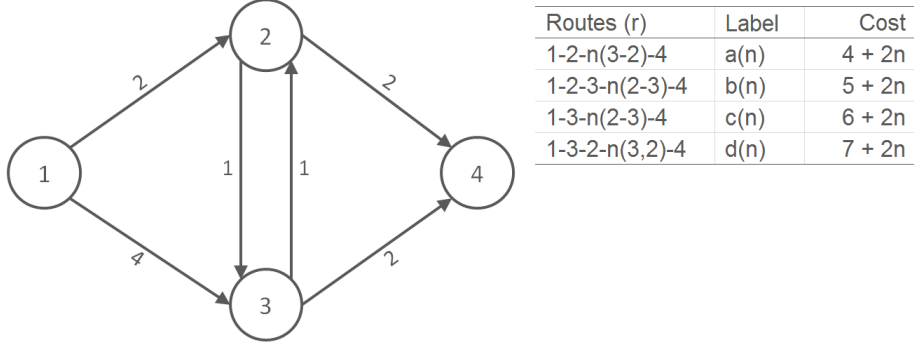
1: **procedure** MTE-MDM
2:     $f_{ij}^{(0)} = 0, \forall (i,j) \in A.$
3:     $n = 0.$
4:     **while** stopping condition is not satisfied **do**
5:         $n = n + 1.$
6:         $t_{ij} = \tau_{ij}\left(f_{ij}^{(n-1)}\right),$ $\forall (i,j) \in A.$
7:         **for each** $d \in D$ **do**
8:             $\mathbf{p_d} = \text{M-MDM}\left(\mathbf{t}, d\right).$
9:             $\mathbf{Q_{ij}^d} = \begin{cases} p_{ij}^d, & \text{if } (i,j) \in A_d \text{ and } j \neq d, \\ 0, & \text{otherwise.} \end{cases}$
10:             $\mathbf{M^d} = \left(\mathbf{I} - \mathbf{Q^d}\right)^{-1}$
11:             $\mathbf{n_d} = \left(\mathbf{M^d}\right)^{\mathsf{T}} \mathbf{h_d}$
12:             $x_{ij}^d = n_i^d \cdot p_{ij}^d,$ $\forall (i,j) \in A_d.$
13:         **end for**
14:         $f_{ij}^* = \sum_{d \in D: (i,j) \in A_d} x_{ij}^d,$ $\forall (i,j) \in A.$
15:         $f_{ij}^{(n)} = f_{ij}^{(n-1)} + \frac{1}{n}\left(f_{ij}^* - f_{ij}^{(n-1)}\right),$ $\forall (i,j) \in A.$
16:     **end while**
17: **end procedure**

Figure 2: Four-node network with single OD pair (1,4).

Routes (r)	Label	Cost
1-2-n(3-2)-4	a(n)	4 + 2n
1-2-3-n(2-3)-4	b(n)	5 + 2n
1-3-n(2-3)-4	c(n)	6 + 2n
1-3-2-n(3,2)-4	d(n)	7 + 2n

many routes in this network due to the presence of the cycle between nodes 2 and 3. All the possible routes can be represented in four groups as illustrated in the table next to the network. In order to illustrate the notation, the route $a(n) = $ 1-2-n(3-2)-4 corresponds to route 1-2-4 when $n = 0$, and corresponds to cyclic route 1-2-3-2-4 when $n = 1$. In a route based model considering all routes, we have:

$$\sum_{n=0}^{\infty} \left(p_{a(n)} + p_{b(n)} + p_{c(n)} + p_{d(n)} \right) = 1,$$

where p_r is the choice probability of route r.

5.1.1 Relaxing the Independence Assumption

The Recursive Logit (RL) model assumes a joint distribution $\boldsymbol{\theta}^G$ with iid Gumbel random variables. M-MDM does not assume independence; instead it calculates the choice probabilities with respect to an extremal joint distribution minimizing the expected cost over the set of all joint distributions satisfying the given marginals. Let $\boldsymbol{\Theta}^G$ be the set of all joint distributions with Gumbel marginals, i.e.,

$$F_{ij}(\omega) = \exp\left[-\exp\left(-\gamma - \beta\omega\right)\right],$$

where γ is the Euler's constant. The behavioral difference between the RL and M-MDM with Gumbel marginals can be interpreted analytically. Since $\boldsymbol{\theta}^G \in \boldsymbol{\Theta}^G$, the optimal expected cost of the M-MDM is smaller than that of the RL. The expected costs are bounded above by the deterministic

Figure 3: Effect of independence assumption on link choice probabilities.
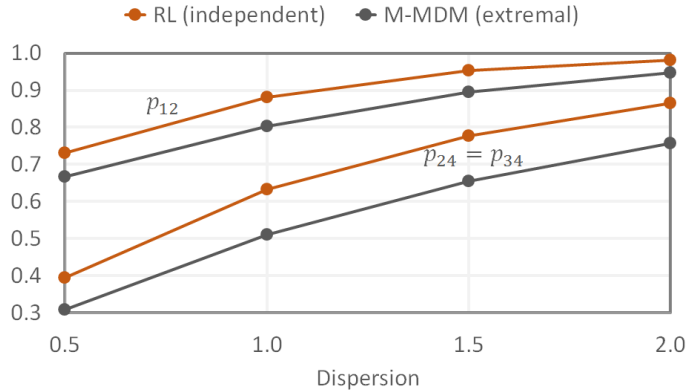
shortest path costs, s_i^d. Therefore, RL assignment is closer to deterministic assignment than the M-MDM, regardless of the dependence structure of the extremal distribution solving the M-MDM recursion. We calculate the choice probabilities of the network in Figure 2 for different values of the dispersion parameter, $\beta \in \{0.5, 1.0, 1.5, 2.0\}$. Figure 3 represents the choice probabilities of the links $(1, 2), (2, 4)$, and $(3, 4)$, which are on the shortest paths from nodes 1, 2 and 3 to the destination, respectively. As discussed above, the RL model assigns a greater portion of the flow to the deterministic shortest routes. It is also worth noting that the difference between the choice probabilities is not monotonic due to the fact that the extremal joint distribution solving the M-MDM does not necessarily have the same dependency structure for different values of the dispersion parameter.

5.1.2 Incorporating Scale Heterogeneity

Suppose that the variance of the error terms of all links is equal to one, except for the link $(1, 3)$. We assume that the marginal distributions are exponential random variables with possibly different parameters and refer to the corresponding model as the Markovian Marginal Exponential Model (M-MEM). The marginal distributions are given as:

$$F_{ij}(\omega) = 1 - \exp(-1 - \beta_{ij}\omega), \qquad \forall \omega \geq -\frac{1}{\beta_{ij}}, (i, j) \in A, \qquad (50)$$

where the rate parameter $\beta_{ij} = 1, \forall (i, j) \in A \setminus \{(1, 3)\}$. The error term $\widetilde{\epsilon}_{ij}$ under this model has a mean of zero and a variance of $(\beta_{ij})^{-2}$. We vary

28

the standard deviation of link $(1,3)$ in range $[0.1, 10.0]$. Figure 4 plots p_{13} and w_1 with respect to the standard deviation for link $(1,3)$. When the variance is very small, almost all users choose link $(1,2)$, $p_{12} = 1 - p_{13} \simeq 1$, which is on the shortest route 1-2-4. However, p_{13} increases as the standard deviation increases, and reaches around 0.20 when the standard deviation is 2. There is also a monotonic relation between the standard deviation and the expected minimum cost. Finally, note that the Markovian logit model corresponds to the dashed line where $\beta_{13} = 1$ due to identically distributed assumption; while the M-MEM models user behavior under nonidentical error terms with scale heterogeneity. For any value of β_{13}, the choice probabilities $p_{23}, p_{24}, p_{32}, p_{34}$, and expected costs w_2 and w_3 remain unaffected since there is no path from nodes 2 and 3 to 1. However, the change in p_{13} still makes a significant change in the link flows. Figure 5 plots the link flows x_{ij} with respect to the standard deviation for demand $h_1 = 100$.

Next, we experiment the model M-MEMs defined by exponential marginals where the error variances are scaled with respect to the free-flow link costs as follows:

$$F_{ij}\left(\omega\right) = 1 - \exp\left(-1 - \frac{\omega}{\nu_i \cdot t_{ij}^0}\right), \qquad \forall \omega \geq \nu_i \cdot t_{ij}^0, (i, j) \in A. \qquad (51)$$

The random error term $\widetilde{\epsilon}_{ij}$ has mean zero and standard deviation $\nu_i \cdot t_{ij}^0$. This variant scales the error variance of each link separately with respect to the link cost and hence, it is suitable under the assumption that longer links are subject to greater error variance. We use $\nu_i = \nu \in \{0.1, 0.2, 0.3, 0.4\}$, and compare the link choice probabilities in M-MEM with the Nested Recursive Logit model (NRL) where the variances are scaled in the same manner, except that error variance is equal for the links emanating from the same node. Figure 6 illustrates the choice probabilities p_{12}, p_{24} and p_{34} for different values of ν. For small values of ν, the majority of the users choose the shortest route; while the choice probabilities of the links on the shortest routes decrease as the coefficient of variation increases. Furthermore, we observe that the probabilities of these links are less in the M-MEMs when compared with the logit model. The reason for this is that the logit model scales all error variances at node i with respect to the link with smallest cost due to the identically distributed assumption. This underestimates the variance in the longer links, and hence favors the shortest link. On the other hand, M-MEMs assigns the variance of each link separately and provides more accurate variance scaling.

Figure 4: Effect of standard deviation on expected cost and link choice probability.
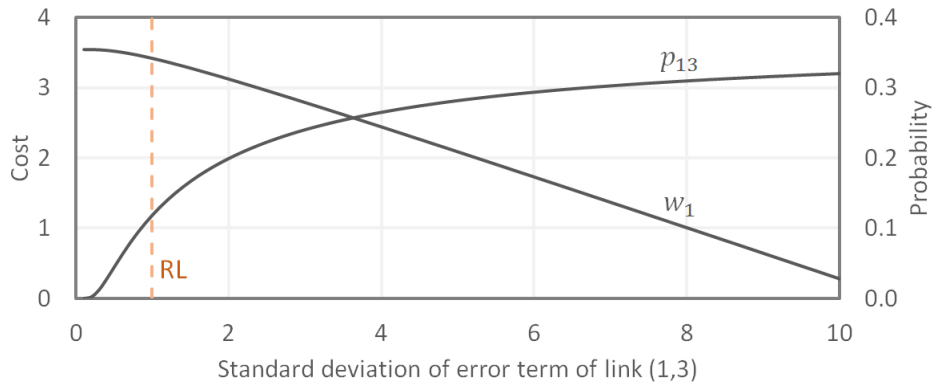

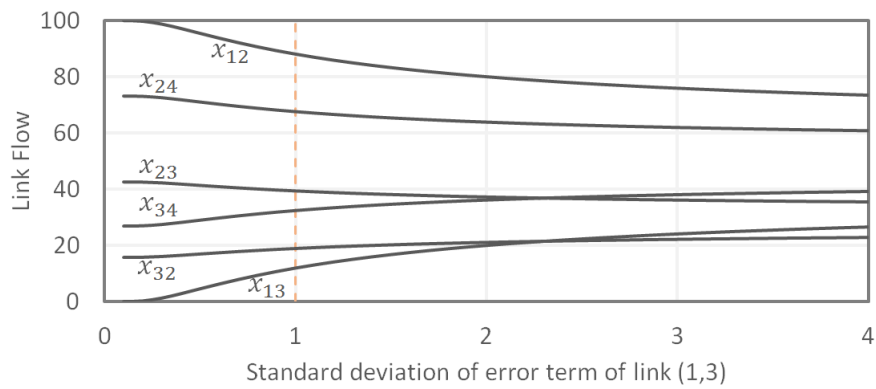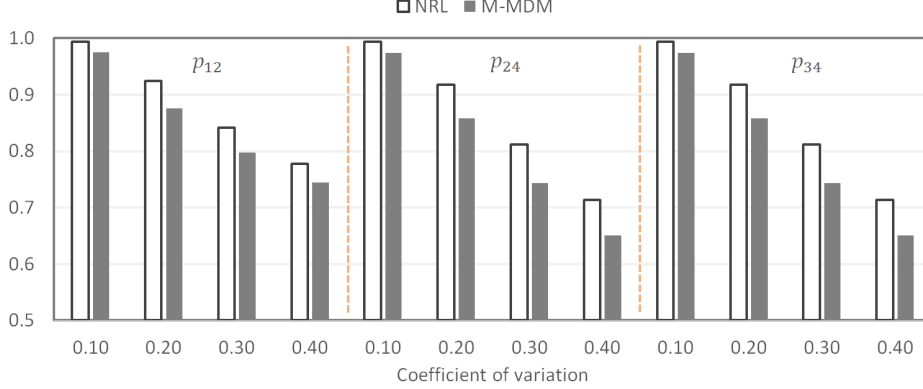Figure 5: Effect of standard deviation on link flows.

Figure 6: Effect of link specific standard deviation on link choice probabilities.

5.1.3   Comparison with Route Based Models

In this section, we compare the link based model with a route based model. Consider the network in Figure 2 with the link costs displayed next to the arcs, except that $t_{32} = \infty$. The network in this case is acyclic with three routes, $R = \{1\text{-}2\text{-}4, 1\text{-}2\text{-}3\text{-}4, 1\text{-}3\text{-}4\}$. In this example, we assume normal random error terms. Consider the M-MDM variant, termed as the M-MNM, with the following marginals:

$$F_{ij}\left(\omega\right) = \Phi\left(\frac{\omega - \mu_{ij}}{\sigma_{ij}}\right), \qquad\qquad \forall (i,j) \in A. \qquad (52)$$

In route based models, the random cost of route r is usually represented as $\sum_{(i,j)\in r} t_{ij} - \widetilde{\epsilon}_r$. Consider the route based model proposed by Ahipaşaoğlu et al. (2016) for the uncongested case. Assume that the marginal distribution of $\widetilde{\epsilon}_r$ is normal with mean μ_r and standard deviation σ_r, $N\left(\mu_r, \sigma_r\right)$, where:

$$\mu_r = \sum_{(i,j)\in r} \mu_{ij}, \qquad\qquad \forall r \in R,$$

$$\sigma_r^2 = \sum_{(i,j)\in r} \sigma_{ij}^2, \qquad\qquad \forall r \in R.$$

The route choice probabilities under this setting are calculated as:

$$p_r = 1 - \Phi\left(\frac{\lambda + \sum_{(i,j)\in r} t_{ij} - \mu_r}{\sigma_r}\right), \qquad\qquad \forall r \in R,$$

31

where λ satisfies the following normalization condition:

$$\sum_{r \in R} \left[1 - \Phi \left(\frac{\lambda + \sum_{(i,j) \in r} t_{ij} - \mu_r}{\sigma_r} \right) \right] = 1.$$

Note that the error terms of links emanating from different nodes are independent. Therefore, for acyclic graphs, the reproductive property holds, i.e., $\widetilde{\epsilon}_r = \sum_{(i,j) \in r} \widetilde{\epsilon}_{ij}$, where $\widetilde{\epsilon}_{ij} \sim N(\mu_{ij}, \sigma_{ij})$. In this case, the difference in user choice behavior under the Markovian and route based models stems from the Markovian assumption.

We set $\mu_{ij} = 0$ and set an identical standard deviation $\sigma_{ij} = \sigma$ for all links, varying in the interval $(0, 4]$. The route choice probabilities with the M-MNM are calculated using link choice probabilities in equation (3). Figure 7 illustrates the route choice probabilities under the link and route based models. We observe that there is a significant difference in choice probabilities. In particular, M-MNM assigns a greater choice probability to the shortest route 1-2-4 than the route based model. This behavior can be interpreted using the link choice probability expression (21). The choice probability p_{ij} depends directly on the error variance of $\widetilde{\epsilon}_{ij}$, while the effect of the error variances of the downstream links are encapsulated in w_j. This leads to a reduced error variance compared to the route based model; and hence, more users choose the shortest route.

Next, we compare the route based and the Markovian models but allow for cyclic routes by setting $t_{32} = 1$. For the route based model, we use the predetermined route choice set $R = \{a(n), b(n), c(n), d(n); n = 0, 1, \ldots, 5\}$ with the route labels given in Figure 2. Figure 8 represents the choice probabilities of the simple routes $b(0) = $ 1-2-3-4, $c(0) = $ 1-3-4, $d(0) = $ 1-3-2-4. Note that the deterministic costs of 1-3-4 and 1-3-2-4 are 6 and 7, respectively. However, we observe $p_{1324} > p_{134}$ in the route based model when the error variance is large. This behavior might be acceptable for these paths, however, for the cyclic routes it leads to a paradox which is explained next.

Consider the routes $a(0) = $ 1-2-4 and $a(1) = $ 1-2-3-2-4. Intuitively, we expect $p_{a(0)} \geq p_{a(1)}$. The MNL assignment always satisfies this inequality. On the other hand, it might be violated for sufficiently large error variance with the logit variants such as the C-logit (Cascetta et al., 1996) or the path-size logit (Ben-Akiva and Bierlaire, 1999); as well as for certain settings of the marginals with the MDM-SUE model. Therefore, the parameter space should carefully be constrained for route based models with nonidentical error terms in order to prevent this paradox of cyclic routes. On the other hand, we always have $\prod_{(i,j) \in a(0)} p_{ij} \geq \prod_{(i,j) \in a(1)} p_{ij}$. Therefore, the

Figure 7: Comparison of route based and Markovian choice models on acyclic network.
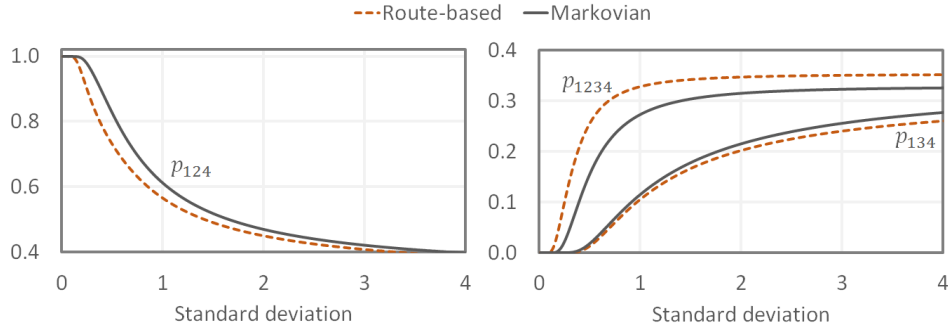
cycle paradox is never observed with the Markovian choice model. Figure 9 represents the choice probabilities for different values of the standard deviation. First, note that the choice probability of the shortest route 1-2-4 is greater in the Markovian model, as in the acyclic case. Second, we observe the paradox $p_{12324} > p_{124}$ for $\sigma \geq 2.3$ in the route based model.

Furthermore, choice probabilities of route based models depend on the choice set which is necessarily finite. Let $R^m = \{a(n), b(n), c(n), d(n); n = 0, 1, \ldots, m\}$ be the choice set for the route based MNL model with dispersion parameter $\beta = 0.1$. Figure 10 displays the MNL choice probabilities for the simple routes with respect to $m \in \{0, 1, \ldots, 10\}$. The choice probabilities are significantly affected by the route choice set. For $m = 20$, the route choice probabilities aproach to those obtained by the M-MEM with marginals (50) where $\beta_{ij} = 0.1$. These analyses indicate the better fit of the Markovian choice model in traffic assignment when users are allowed to follow cyclic routes and when the uncertainty is specified in the link level.

5.1.4 Comparison with the MDM

In this section we compare the M-MDM assignment with the alternate link based MDM summarized in Section 2.2. We again use the acyclic network with $t_{32} = \infty$, and assume that a single unit is transported from node 1 to node 4. We use the link specific scaling with exponential marginals given in (51). Figure 11 represents flows of the MDM next to the links and those of the M-MDM in parentheses, for $\nu = 0.4$. We observe that the M-MDM assigns a larger flow on the shortest path 1-2-4. Figure 12 displays the objective function values of the two models for $\nu \in (0, 1]$. As ν approaches

Figure 8: Comparison of route based and Markovian choice models on cyclic network.
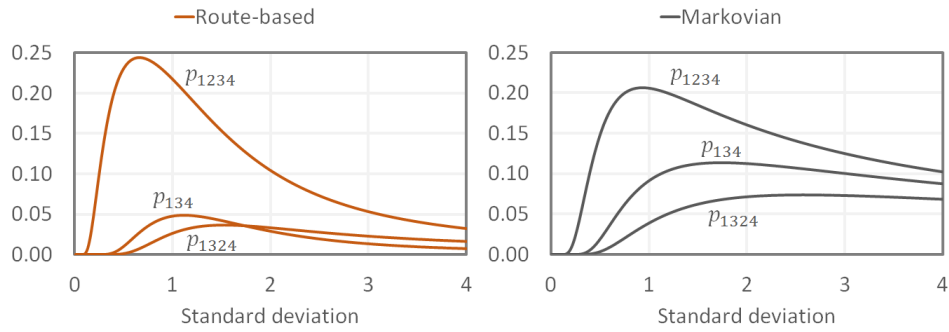

Figure 9: Comparison of route based and Markovian choice models on cyclic network.
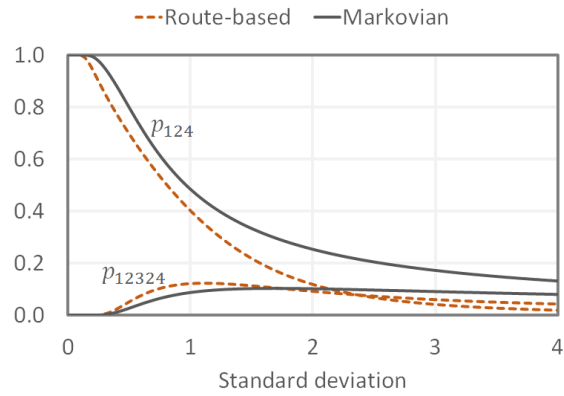

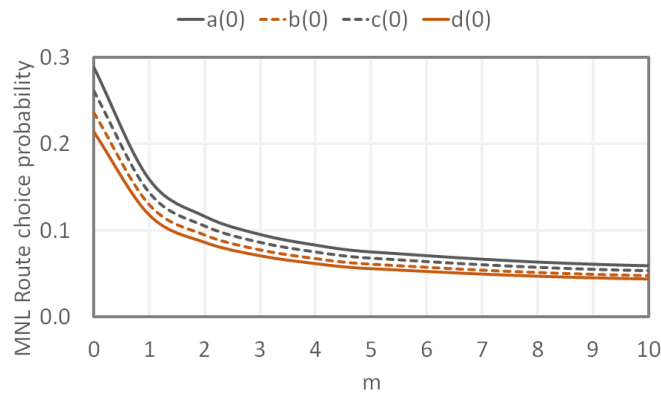Figure 10: Effect of choice set on the route based MNL choice probabilities.

Figure 11: Comparison of link based MDM and M-MDM flows on acyclic network.
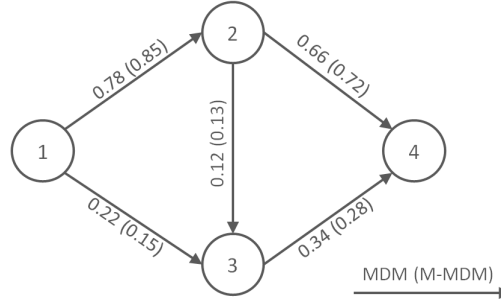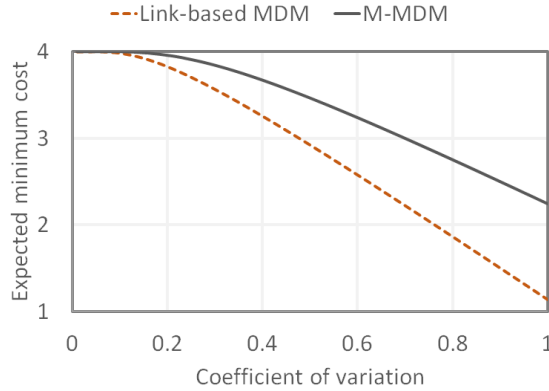

Figure 12: Comparison of link based MDM and M-MDM costs on acyclic network.

to zero, the optimal solutions of both models approach to the deterministic shortest path distance 4. As the error variances increase, the gap between the two models increases. The optimal M-MDM cost is always greater than that of the link based MDM due to the difference in the decision processes. In the link based MDM the user observes the entire network, while M-MDM models a dynamic decision making process where only the emanating links are observed at each node. This can also be observed in the additional penalty term in the M-MEM objective function given in Table 1.

5.1.5 Incorporating Other Tail Behavior

M-MDM allows utilizing different distribution families to model user choice behavior. In this section, we compare the effect of the family of the marginal

distributions on the traffic assignment. We use identical marginal distributions with mean zero and standard deviation σ. We consider three distributions. The first setting is $\beta_{ij} = \sigma^{-1}$ in the exponential distribution function (50); and the second setting is $\mu_{ij} = 0, \sigma_{ij} = \sigma$ in the normal distribution function (52). In addition, we consider a variant with the logistic marginal:

$$F_{ij}(\omega) = (1 + \exp\left[-\eta_{ij}(\omega - \alpha_{ij})\right])^{-1}, \qquad \forall \omega, (i,j) \in A,$$

where α_{ij} and η_{ij} are the location and rate parameters, respectively. The mean and variance are α_{ij} and $\pi^2/3(\eta_{ij})^2$. We set $\alpha_{ij} = 0, \eta_{ij} = \pi/\sqrt{3}\sigma$ in the logistic cumulative distribution function for all links. In this case, we obtain M-MEM, M-MNM and M-MLM error terms with mean zero and variance σ^2 for all links. We compare the user behavior under these models for $\sigma \in [0.5, 2.0]$.

Figure 13 illustrates the choice probabilities of links $(1,2)$, $(2,4)$ and $(3,4)$. Due to identical marginals, we have $p_{24} = p_{34}$. We first observe that user choice behavior changes significantly with different distribution families even when the error terms have equal means and variances. M-MEM assigns a greater choice probability to the links on the shortest route when compared with M-MNM and M-MLM choice probabilities are in between these two models. Note that exponential distribution has positive skewness while normal and logistic distributions are symmetric. M-MEM assigns a greater mass close to zero for the random error terms; and hence, resulting link choice probabilities are closer to the deterministic choice ($p_{12} = p_{24} = p_{34} = 1$). Normal and logistic distributions have similar shapes. The reason for the greater choice probabilities assigned to links on shortest routes under the M-MLM is that the logistic distribution has heavier tails than the normal distribution. Figure 14 represents the expected minimum costs from nodes 1, 2 and 3 to the destination 4 under these models. The models in increasing order of the difference of deterministic costs ($z_{14} = 4, z_{24} = z_{34} = 2$) from expected costs is M-MEM, M-MLM, and M-MNM, due to similar reasons.

5.2 Convergence Behavior of the Algorithm for MTE-MDM

In this section, we study the convergence performance of the solution algorithm, Algorithm 2, to calculate the MTE-MDM flows. Consider the M-MEM model where the marginal distribution of the error term $\widetilde{\epsilon}_{ij}$ is exponential with rate parameter $\beta_{ij} = 2/t_{ij}^0$ for the network given in Figure 2. M-MDM link choice probabilities for the free-flow travel costs are given

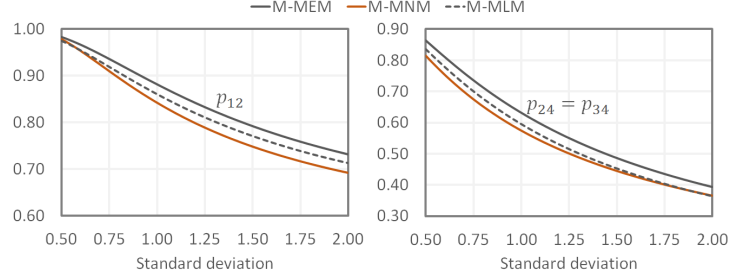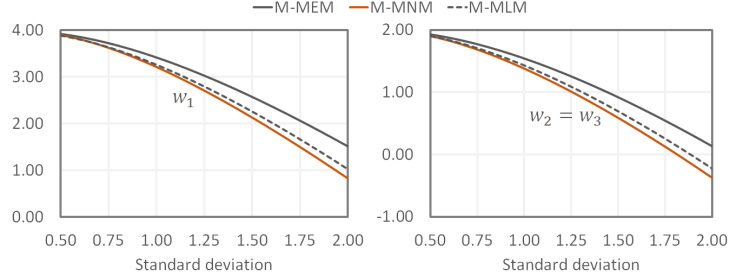Figure 13: Effect of distribution family on link choice probability.

Figure 14: Effect of distribution family on expected minimum cost.

below:

$$p_{12} = 0.80, \quad p_{13} = 0.20, \quad p_{23} = p_{32} = 0.28, \quad p_{24} = p_{34} = 0.72.$$

Suppose that there are two OD pairs $(1, 4)$ and $(2, 4)$ with demands 10 and 5 respectively. The Markovian steps of Algorithm 2 (lines 9-14) and the corresponding link flows f_{ij}^* are given below:

$$\mathbf{Q} = \begin{bmatrix} 0.00 & 0.80 & 0.20 \\ 0.00 & 0.00 & 0.28 \\ 0.00 & 0.28 & 0.00 \end{bmatrix}, \mathbf{M} = (\mathbf{I} - \mathbf{Q})^{-1} = \begin{bmatrix} 1.00 & 0.93 & 0.46 \\ 0.00 & 1.09 & 0.30 \\ 0.00 & 0.30 & 1.09 \end{bmatrix},$$

$$\mathbf{h} = \begin{bmatrix} 10 \\ 5 \\ 0 \end{bmatrix}, \mathbf{n} = \begin{bmatrix} 10.00 \\ 14.71 \\ 6.12 \end{bmatrix}, \quad \begin{array}{ll} x_{12} = f_{12}^* = 8.00, & x_{13} = f_{13}^* = 2.00, \\ x_{23} = f_{23}^* = 4.09, & x_{24} = f_{24}^* = 10.61, \\ x_{32} = f_{32}^* = 1.70, & x_{34} = f_{34}^* = 4.39. \end{array}$$

The descent direction is obtained using f_{ij}^* in line 15 of the algorithm. We use the link cost function $\tau_{ij}(f) = t_{ij}^0 (1 + 0.02f), \forall (i, j) \in A$. The equilibrium flows, costs, and link choice probabilities obtained by Algorithm 2 are given

37

Figure 15: Convergence behavior of Algorithm 2.
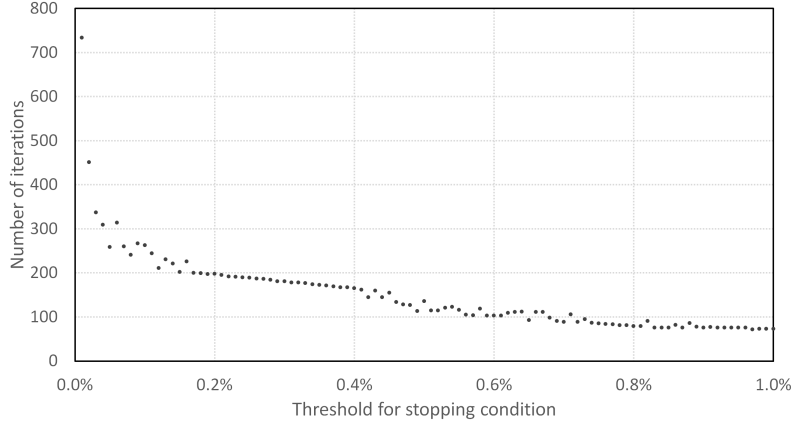
below:

$$f_{12} = 7.75, \quad \tau_{12}(f_{12}) = 2.31, \quad p_{12} = 0.77,$$
$$f_{13} = 2.25, \quad \tau_{13}(f_{13}) = 4.18, \quad p_{13} = 0.23,$$
$$f_{23} = 4.48, \quad \tau_{23}(f_{23}) = 1.09, \quad p_{23} = 0.30,$$
$$f_{24} = 10.31, \quad \tau_{24}(f_{24}) = 2.41, \quad p_{24} = 0.70,$$
$$f_{32} = 2.04, \quad \tau_{32}(f_{32}) = 1.04, \quad p_{32} = 0.30,$$
$$f_{34} = 4.69, \quad \tau_{34}(f_{34}) = 2.19, \quad p_{34} = 0.70.$$

Note that the Markovian steps using the equilibrium link choice probabilities return back the equilibrium link flows; and hence, solves the fixed point problem of MTE-MDM given in (38)-(41).

We use the Sioux Falls network to experiment the convergence behavior of the algorithm. Figure 15 plots the number of MSA iterations against the threshold ϵ_2. We observe that the algorithm shows a nearly linear convergence rate for $\epsilon_2 \geq 0.05\%$.

Finally, we experiment the effect of distribution parameters on convergence performance of Algorithm 2. Markovian models might be considerably more complex than the route based SUE models, due to the recursive relation of the expected minimum costs. In particular, the recursion might diverge when the error variance is sufficiently large. We observe in our experiments on the small network that the choice probabilities of cyclic routes increase as the expected minimum costs decrease. This not only increases the solution time required for the M-MDM algorithm but also for the matrix inverse operation. On the other hand, execution time of these steps is much faster when the expected costs are close to the deterministic shortest path

cost, or when majority of the users follow the deterministic shortest routes. Similar issues are discussed by several authors (see Akamatsu (1996); Wong (1999); Baillon and Cominetti (2008); Fosgerau and Bierlaire (2009); Si et al. (2010)). Therefore, understanding the effect of the distribution parameters on convergence rate is crucial. In our experimental setup, we first find the parameters that result in zero expected minimum cost at each node. This represents the farthest setting from the deterministic choice while we constraint the costs to be nonnegative. Then, we test settings that are closer to the deterministic choice to understand the sensitivity of the solution approach to the parameters. We use the M-MEM choice model within the equilibrium model. Let $\underline{\beta}_i$ be a lower bound on the dispersion parameter that guarantees nonnegative instantaneous expected minimum cost for the logit model:

$$\sum_{j \in N_d^+(i)} \exp\left(-\underline{\beta}_i \cdot t_{ij}^0\right) = 1, \qquad \forall i \in N_d.$$

We set the scale parameter of the exponential marginals to $B_{ij} = B_i = (\alpha_1 \cdot \underline{\beta}_i)^{-1}$, and vary α_1 in the range $\{1.25, 1.50, 1.75, 2.00\}$. Note that as α_1 increases, the error variances decrease. We allow non-zero means and calculate upper bounds on the location parameters, $\overline{A}_i$, for every B_i to guarantee nonnegative expected costs as follows:

$$\overline{A}_i = \operatorname*{argmax}_{a \in R} \left\{ \sum_{j \in N_d^+(i)} \exp\left(\frac{a - t_{ij}^0}{B_i}\right) \leq \exp\left(-1\right) \right\}, \qquad \forall i \in N_d.$$

We set the location parameters of the exponential marginals to $A_{ij} = A_i = -B_i + \alpha_2(\overline{A}_i + B_i)$, for $\alpha_2 \in \{-1.0, -0.5, 0.0, 0.5\}$. The model reduces to the traffic equilibrium with the Nested Recursive Logit model when $\alpha_2 = 0$. All combinations of α_1 and α_2 guarantee nonnegative costs. Finally, note that as α_2 increases, the expected costs decrease for given variance. Due to the discussion above, we expect the solution time to increase as α_1 decreases and α_2 increases. We use the same MSA algorithm for NRL model utilizing closed form choice probability and expected minimum cost expressions to compare the convergence performance of the MTE-MDM.

We carry out the experiment on the small network of Sioux Falls and the large network of Winnipeg. Winnipeg network is composed of 1040 nodes (138 of which are destinations), 2836 links and 4344 OD pairs. The topology, link characteristics, and OD demands of these networks are obtained

Table 2: Convergence behavior of Algorithm 2 with respect to distribution parameters.

α_1	α_2	Sioux Falls		Winnipeg		Winnipeg (NRL)
		Iter.	Time(s)	Iter.	Time(m)	Time(m)
2.00	-1.0	255	8.19	117	10.88	
	-0.5	530	20.21	154	14.63	
	0.0	430	19.78	188	16.37	12.07
	0.5	357	20.73	229	23.15	
1.75	-1.0	450	16.85	152	11.27	
	-0.5	480	20.25	175	15.16	
	0.0	363	17.42	194	16.96	11.35
	0.5	327	19.29	218	23.99	
1.50	-1.0	567	22.41	168	11.32	
	-0.5	356	16.27	174	15.23	
	0.0	319	16.99	206	17.03	12.50
	0.5	288	17.26	223	24.10	
1.25	-1.0	280	12.71	185	11.09	
	-0.5	308	15.71	194	14.91	
	0.0	273	14.87	204	16.68	13.71
	0.5	264	16.69	223	23.59	

from Bar-Gera's Traffic Assignment Test Problem website[2]. The results are summarized in Table 2. The columns with header 'Iter' represent the number of MSA iterations, (s) and (m) represent seconds and minutes, and the rightmost column represents the NRL solution time for the Winnipeg network. The effect of the distribution parameters on the convergence rate is not significant for the small network. On the other hand, we observe an increase in the number of MSA iterations and solution time as α_2 increases for the Winnipeg network. Interestingly, the effect of variance on the solution performance is less significant. This is due to the parameter settings that guarantee nonnegative costs which protects the M-MDM algorithm against divergence. Finally, the average solution time of the MTE-MDM for the large network is 16.65 minutes, while that of the NRL model is 12.41 minutes. Therefore, we conclude that the proposed equilibrium model provides modeling flexibility with a reasonable additional computational burden.

---

[2]See https://github.com/bstabler/TransportationNetworks

6   Conclusion

Markovian choice models use a universal choice set and avoid route set generation to overcome the drawbacks of the route based SUE models. Baillon and Cominetti (2008) proposed the Markovian Traffic Equilibrium model assuming complete knowledge on the distribution of the random utilities. Despite the convexity of the MTE formulation, the application of the model has been primarily restricted to the logit model due to the challenge in computing choice probabilities and expected minimum costs for general choice models. In this paper, we have developed a distributionally robust Markovian choice model (M-MDM) and Markovian traffic equilibrium model (MTE-MDM) under the assumption that the joint distribution of the error terms is not known. We develop an equivalent convex optimization reformulation and an efficient solution algorithm to obtain the M-MDM link choice probabilities, which is then integrated in a method of successive averages algorithm to obtain the MTE-MDM flows in the congested case. Proposed choice model extends an earlier MDM formulation developed in Natarajan et al. (2009) for link based models in acyclic networks by explicitly incorporating the Markovian assumption into the choice process. Furthermore, proposed equilibrium model extends the recent work of Ahipaşaoğlu et al. (2016) on the marginal distribution model from a route based model to a link based model.

M-MDM allows using different families of distributions, and provides a framework for a wide range of Markovian choice models to accurately model user choice behavior. We show that the Markovian logit model can be recreated by the M-MDM for a particular choice of the marginals. Moreover, it provides further flexibility by relaxing the iid assumption and incorporating different tail behavior. We present the results of a comprehensive experimentation to illustrate the modeling flexibility, as well as the convergence performance of the MTE-MDM model. The results indicate that the model provides computational tractability for calculating the distributionally robust traffic equilibrium in large traffic networks. We believe that the flexible nature of the model and the practicality of the solution approach makes MTE-MDM interesting for future research. A related question is to perform demand estimation from observed traffic flows using such a framework. Also, the Markovian assumption allows to adapt the model for dynamic networks where the state of the users are defined by a node-time pair. We leave these questions for future research.

Appendix

The main notations used throughout the paper is tabulated below.

G	:	Traffic network with set of nodes N and set of links A, i.e., $G = (N, A)$.
W	:	Set of OD pairs.
D	:	Set of destinations: $D = \{d \in N : (o, d) \in W\}$.
G_d	:	Subnetwork for users towards destination d with set of nodes $N_d \cup \{d\}$ and set of links A_d, i.e., $G_d = (N_d \cup \{d\}, A_d)$.
$N_d^+(i)$	:	Out-neighbourhood of node i in G_d: $N_d^+(i) = \{j \in N : (i, j) \in A_d\}$.
$N_d^-(i)$	:	In-neighbourhood of node i in G_d: $N_d^-(i) = \{k \in N : (k, i) \in A_d\}$.
h_i^d	:	Demand for the OD pair (i, d).
t_{ij}^0	:	Free-flow travel cost of link (i, j).
t_{ij}	:	Travel cost of link (i, j).
w_i^d	:	Expected minimum cost from node i to d.
$\widetilde{\epsilon}_{ij}^d$	:	Random utility error term of using the link toward destination d.
θ_{id}	:	Joint distribution of error terms of links emanating from node i, $\{\widetilde{\epsilon}_{ij}^d; j \in N_d^+(i)\}$, for destination d.
$F_{ij}^d(.)$	:	Cumulative distribution function of the random error term $\widetilde{\epsilon}_{ij}^d$.
$F_{ij}^{d\,(-1)}(.)$	:	Inverse of the cumulative distribution function $F_{ij}^d(.)$.
Θ_{id}	:	Set of joint distributions of random error terms $\{\widetilde{\epsilon}_{ij}^d; j \in N_d^+(i)\}$.
p_{ij}^d	:	Choice probability of link (i, j) towards d.
x_{ij}^d	:	Number of users using link (i, j) towards d.
n_i^d	:	Number of users leaving node i towards d.
f_{ij}	:	Number of users using link (i, j).
$\tau_{ij}(.)$	:	Cost function of link (i, j).

References

S. D. Ahipaşaoğlu, R. Meskarian, T. L. Magnanti, and K. Natarajan. Beyond normality: A cross moment-stochastic user equilibrium model. *Transportation Research Part B: Methodological*, 81(2):333–354, 2015.

S. D. Ahipaşaoğlu, U. Arıkan, and K. Natarajan. On the flexibility of using marginal distribution choice models in traffic equilibrium. *Transportation Research Part B: Methodological*, 91:130–158, 2016.

T. Akamatsu. Cyclic flows, Markov process and stochastic traffic assignment. *Transportation Research Part B: Methodological*, 30(5):369–386, 1996.

T. Akamatsu. Decomposition of path choice entropy in general transport networks. *Transportation Science*, 31(4):349–362, 1997.

A. Atamturk and M. Zhang. Two-stage robust network flow and design under demand uncertainty. *Operations Research*, 55(4):662–673, 2007.

J.-B. Baillon and R. Cominetti. Markovian traffic equilibrium. *Mathematical Programming*, 111(1-2):33–56, 2008.

S. Bekhor, M. E. Ben-Akiva, and M. S. Ramming. Evaluation of choice set generation algorithms for route choice models. *Annals of Operations Research*, 144(1):235–247, 2006.

M. G. Bell. Alternatives to Dial's logit assignment algorithm. *Transportation Research Part B: Methodological*, 29(4):287–295, 1995.

M. Ben-Akiva and M. Bierlaire. Discrete choice methods and their applications to short term travel decisions. *Hall, R.W. (Ed.), Handbook of Transportation Science, Kluwer Publishers*, pages 5–33, 1999.

J. G. Carlsson and M. Behroozi. Wasserstein distance and the distributionally robust TSP. *Working Paper*, 2017.

J. G. Carlsson and E. Delage. Robust partitioning for stochastic multivehicle routing. *Operations Research*, 61(3):727–744, 2013.

E. Cascetta, A. Nuzzolo, F. Russo, and A. Vitetta. A modified logit route choice model overcoming path overlapping problems: specification and some calibration results for interurban networks. *In: Proceedings of the 13th International Symposium on Transportation and Traffic Theory, Leon, France*, pages 697–711, 1996.

C. F. Daganzo and Y. Sheffi. On stochastic models of traffic assignment. *Transportation Science*, 11(3):253–274, 1977.

R. Dial. A probabilistic multipath traffic assignment model which obviates path enumeration. *Transportation Research*, 5(2):83–111, 1971.

M. J. C. Erera, A. L. and M. Savelsbergh. Robust optimization for empty repositioning problems. *Operations Research*, 57(2):468–483, 2009.

M. Fosgerau and M. Bierlaire. Discrete choice models with multiplicative error terms. *Transportation Research Part B: Methodological*, 43(5):494–505, 2009.

M. Fosgerau, E. Frejinger, and A. Karlstrom. A link based network route choice model with unrestricted choice set. *Transportation Research Part B: Methodological*, 56:70–80, 2013.

J. G. Kemeny, J. L. Snell, et al. *Finite Markov chains*, volume 356. van Nostrand Princeton, NJ, 1960.

T. Mai, M. Fosgerau, and E. Frejinger. A nested recursive logit model for route choice analysis. *Transportation Research Part B: Methodological*, 75:100–112, 2015.

V. K. Mishra, K. Natarajan, D. Padmanabhan, C.-P. Teo, and X. Li. On theoretical and empirical aspects of marginal distribution choice models. *Management Science*, 60(6):1511–1531, 2014.

K. Natarajan, M. Song, and C.-P. Teo. Persistency model and its applications in choice modeling. *Management Science*, 55(3):453–469, 2009.

H. Scarf. *A min-max solution of an inventory problem in Studies in the Mathematical Theory of Inventory and Production.* Stanford University Press, 1958.

B.-F. Si, M. Zhong, H.-Z. Zhang, and W.-L. Jin. An improved Dial's algorithm for logit-based traffic assignment within a directed acyclic network. *Transportation Planning and Technology*, 33(2):123–137, 2010.

J. Wardrop. Some theoretical aspects of road traffic research. *Proceedings, Institution of Civil Engineers*, II(1):325–378, 1952.

S. Wong. On the convergence of Bell's logit assignment formulation. *Transportation Research Part B: Methodological*, 33(8):609–616, 1999.