The combined distribution and stochastic assignment problem*

Jan T. Lundgren[1] and Michael Patriksson[2]

[1] *Department of Mathematics, Linköping University, S-581 83 Linköping, Sweden* E-mail: jalun@mai.liu.se

[2] *Department of Mathematics, Chalmers University of Technology, S-412 96 Göteborg, Sweden* E-mail: mipat@math.chalmers.se

The combined distribution and assignment problem is the problem of the simultaneous determination of the distribution of trips between origins and destinations in a transportation network and the assignment of trips to routes in each origin–destination pair. In the model most widely used the distribution is assumed to follow a gravity model with negative exponential deterrence function and the assignment is made according to the deterministic user equilibrium principle. In this paper we describe an extension of this model in which the allocation of trips to routes is made according to the principle of stochastic user equilibrium.

We discuss the behavioural foundations of trip assignment and combined models assuming deterministic or stochastic route choice. In particular, we describe how they can be derived using the efficiency principle from discrete choice theory; the combined distribution and stochastic assignment model is obtained as the continuous approximation of the discrete problem of finding the most probable flow pattern under the assumption of efficient trip making behaviour.

We outline an algorithm for the solution of the model which is based on route (column) generation, disaggregate simplicial decomposition, and partial linearization.

Keywords: Combined transportation planning models, stochastic user equilibrium, efficiency principle, column generation algorithms, partial linearization

1 Introduction

In the design and management of urban transportation systems there is a need for efficient planning tools for analyzing future scenarios and predicting the use of the transportation system. Such tools rely on the use of mathematical models of the transportation system, the users of this system and their aggregate behaviour under different traffic conditions. The area under study is divided into a number of zones, and the transportation system is represented by a network which describes the transportation possibilities (i.e., roads,

*This work was financed by the Swedish Transport and Communications Research Board (KFB), grant 96-206-33.

transit lines, etc.) between the zones. The models are usually applied sequentially, following the well-known four-stage planning process, where output from one model is used as input in the next.

In the first, *trip generation*, stage, the number of trips originating and terminating in different zones are determined, and the demand is specified for different traveller groups, trip purposes and times of day. In the *trip distribution* stage the number of trips between each pair of zones is determined, producing an origin–destination (OD) trip matrix. The portion of the demand that uses a specific mode of transport is determined next; this is the *modal split* stage. The final stage in the sequential planning process is the *trip assignment* stage; at this, final, stage the number of trips allocated to specific routes in the network is determined, giving flows on the individual links and corresponding route and link costs.

In order to obtain a consistent output from the planning process, the stages have to be repeated, since, for example, the travel costs used in the trip distribution stage is a function of the allocation of trips made at the trip assignment stage. An alternative to this sequential treatment of the planning stages is to consider some of them simultaneously in *combined traffic models*. In this paper we focus on *combined distribution and assignment models*; extensions to modal split are however briefly mentioned. This work concerns the behavioural foundations of the traffic models most often used, and what can be derived from basic assumptions of rational traveller behaviour.

Traditional traffic models describe steady-state (static) conditions resulting from travellers' dynamical decision-making processes and some behavioural principle. In the context of trip assignment (TA) the standard behavioural principle was coined by Wardrop [24], and states that given the origin and destination of a trip, the route is chosen such that the individual travel cost (or, time) is minimized. In the steady-state describing the trip assignment model, therefore, no individual traveller can decrease her own travel cost by revising her route-choice decision; further, it means that in any given OD pair, all routes used will have the same travel cost. The situation thus described is therefore often referred to as a (deterministic) *user equilibrium* (UE).

Characteristic features of the situation described by the UE principle are that all travellers have perfect information about travel costs, and are uniform, in the sense that they have the same travel cost perception. If travellers have different preferences or perceive travel costs differently, it is natural to assume that traffic flows do not satisfy the user equilibrium conditions. To model the variations in the travellers' perception of travel cost we can include random components in the travel cost functions. Depending on the probability distributions chosen for the random components, different models are obtained (see, e.g., Sheffi [20]); the probability distribution most widely used implies the multinomial *logit* route choice model. The *stochastic user equilibrium* (SUE) principle states that in a steady-state, the *perceived costs* are equal among the routes used in each OD pair.

In a combined distribution and assignment (CDA) problem, the assignment of trips to routes and the distribution of trips between origin–destination pairs are determined simultaneously. The behavioural principle most widely used for trip distribution is given by a gravity model with negative exponential deterrence function, while trip assignment is made according to the deterministic Wardrop (user equilibrium) principle (e.g., Evans [8]).

In this paper we describe a combined distribution and assignment model, in which the allocation of trips to routes is made according to the principle of *stochastic* user equilibrium, and the distribution follows a gravity model with respect to the composite cost of the routes in each OD pair. (This model was first presented by Erlander [7].) We discuss the behavioural foundations of the model, and outline an algorithm for its solution.

The models described in this paper presume that the travel cost functions are separable, that is, that the travel cost on a link depends only on the total flow on that link. (Examples where more general cost functions are warranted are multimodal networks with interacting modes and networks with interacting flows in unsignalized junctions.) Under this separability assumption, the user equilibrium conditions and gravity model structures described above can be derived as the optimality conditions of *convex optimization problems*. Nothing prevents us, however, from using more general cost structures in the models; the equilibrium conditions would then be formulated as *variational inequality problems* ([18]). The algorithms to be presented can easily be extended to apply to this more general case. The chief motivation for making the separability assumption on the cost functions is that the resulting optimization problems then can be given a direct behavioural interpretation using the *efficiency principle* (Smith [21, 22]); this is not the case for more general cost structures, as pointed out in Bernstein and Smith [1].

The efficiency principle expresses the probability of a trip pattern (resulting from the simultaneous choices of all trip makers) and a flow pattern (consisting of the resulting flows on all routes in the network). The basic assumption concerns efficient trip making behaviour, which is defined as an overall tendency of trip makers to prefer lower travel costs (or, times) to higher ones. From this assumption the form of the probability distribution function can be derived, and hence also the most probable trip and flow patterns. In the context of the trip assignment problem, these patterns correspond to the UE and SUE route choice models, respectively. The combined distribution and stochastic assignment model is derived as the continuous approximation of the discrete problem of finding the most probable flow pattern in a combined choice situation. The outline of how the above mentioned transportation planning models can be given a behavioural motivation using the efficiency principle is the first contribution of this work.

The second contribution is the outline of a solution algorithm for the combined distribution and stochastic assignment model. The algorithm has its foundations in the disaggregate simplicial decomposition (DSD) algorithm of Larsson and Patriksson [12] for the TA-UE model, and in the extensions to this algorithm for the solutions of the TA-SUE model (Damberg *et al.* [3]) and the CDA-UE model (Lundgren and Patriksson [14]); as such, it is a combination of a route (column) generation algorithm and a partial linearization algorithm.

The rest of the paper is thus organized: Section 2 contains formulations and short discussions on the four traffic models described above; Section 3 provides a derivation of the models using the efficiency principle; Section 4, finally, outlines the new solution algorithm.

2 Model formulations

Let us first introduce some necessary notation. Let $\mathcal{G} = (\mathcal{N}, \mathcal{A})$ denote a network of nodes (representing the OD zones and intersections) and directed links (representing road sections). A subset $\mathcal{C}$ of $\mathcal{N} \times \mathcal{N}$ defines a set of OD pairs. The demand of transportation in a given pair (p, q) is denoted by d_{pq}. Denoting by h_{pqr} the flow on the (loop-free) route r between the nodes p and q, and by $\mathcal{R}_{pq}$ the set of all such routes in the OD pair, the total flow on a link $a \in \mathcal{A}$ is

$$f_a = \sum_{(p,q)\in\mathcal{C}} \sum_{r\in\mathcal{R}_{pq}} \delta_{pqra} h_{pqr},$$

where

$$\delta_{pqra} = \begin{cases} 1, & \text{route } r \text{ passes through link } a, \\ 0, & \text{otherwise.} \end{cases}$$

Associated with each link $a \in \mathcal{A}$ is a positive and strictly increasing travel cost function c_a of the total flow f on the link, representing the disutility of using the link.

2.1 *The trip assignment problem*

The deterministic model

The trip (or, traffic) assignment (TA) problem predicts the allocation of trips onto routes in the network given a demand matrix and travel cost functions. The Wardrop conditions can be stated as follows: for any given OD pair $(p, q) \in \mathcal{C}$ and route $r \in \mathcal{R}_{pq}$,

$$\begin{aligned} h_{pqr} > 0 &\implies c_{pqr}(h) = \pi_{pq}, \\ h_{pqr} = 0 &\implies c_{pqr}(h) \geq \pi_{pq}, \end{aligned}$$

where $c_{pqr}(h)$ is the cost on the route given the flows $h = (h_{pqr})$, and π_{pq} is the minimal travel cost in the OD pair. A solution satisfying these conditions is found by solving (e.g., [20, 18])

[TA-UE]

$$\text{minimize } T(f) = \sum_{a\in\mathcal{A}} \int_0^{f_a} c_a(s)\, ds, \tag{2a}$$

subject to

$$\sum_{r\in\mathcal{R}_{pq}} h_{pqr} = d_{pq}, \qquad \forall (p,q) \in \mathcal{C}, \tag{2b}$$

$$\sum_{(p,q)\in\mathcal{C}} \sum_{r\in\mathcal{R}_{pq}} \delta_{pqra} h_{pqr} = f_a, \qquad \forall a \in \mathcal{A}, \tag{2c}$$

$$h_{pqr} \geq 0, \qquad \forall (p,q) \in \mathcal{C}. \tag{2d}$$

The most common method applied to this problem is that of Frank and Wolfe ([10]), a simple descent algorithm for differentiable programs with linear constraints which has some attractive features in this application. Given a feasible link flow f^k, a search direction is found by solving a linear approximation of [TA-UE], in which $T(f)$ is replaced by $T(f^k) + \nabla T(f^k)^{\mathrm{T}}(f - f^k)$; this problem is equivalent to the solution of a set of shortest route problems over the network with link costs $c_a(f_a^k)$. The solution to this problem defines a lower bound on the optimal value of [TA-UE] (for use in termination criteria), and a feasible descent direction in which a line search is made to define the next iterate, f^{t+1}.

More efficient approaches for [TA-UE] are of course available; consult [18] for a thorough overview. A recent algorithm, which is the basis for the algorithm presented later in the paper, is the *disaggregate simplicial decomposition* (DSD) algorithm of Larsson and Patriksson [12]. In this algorithm, the same set of shortest route problems as in the Frank–Wolfe algorithm is solved, but where that method uses a line search, DSD stores the routes generated, maintaining a growing subset $\hat{\mathcal{R}}_{pq}$ of the (in general unknown) complete set $\mathcal{R}_{pq}$ of routes in each OD pair, and (approximately) solves the restriction of [TA-UE] to that (much smaller) subset of routes to define the next iterate.

The stochastic model

To account for the variations in the travellers' perception of travel cost we can extend the deterministic trip assignment model by including random components in the travel cost functions. The perceived travel cost on route r in OD pair (p,q) then is described as $C_{pqr} = c_{pqr} + \varepsilon_{pqr}$, where c_{pqr} is the deterministic travel cost at a given flow and ε_{pqr} is the random component. Depending on the probability distributions chosen for the random components, different models are obtained (see, e.g., Sheffi [20]). If we assume that the random terms ε_{pqr} are independently and identically distributed Gumbel variates, we obtain the widely used multinomial *logit* model. Given a fixed travel cost, the route flow solution can then be given in closed form as

$$h_{pqr} = d_{pq} \frac{\exp(-\Theta c_{pqr})}{\sum_{\ell \in \mathcal{R}_{pq}} \exp(-\Theta c_{pq\ell})}, \qquad \forall r \in \mathcal{R}_{pq},\ \forall (p,q) \in \mathcal{C}, \tag{3}$$

where Θ is a positive scaling factor related to the variances in the perceived travel costs.

The stochastic user equilibrium (SUE) principle states that in a steady-state, the perceived costs are equal among the routes used within each OD pair. In the case of the logit model, this condition is equivalent to (3), with $c_{pqr} = c_{pqr}(h)$; the stochastic user equilibrium conditions then are the optimality conditions for the following convex optimization problem (Fisk [9]):

[TA-SUE]

$$\text{minimize } Z(h) = \sum_{a \in \mathcal{A}} \int_0^{f_a} c_a(s)\, ds + \frac{1}{\Theta} \sum_{(p,q) \in \mathcal{C}} \sum_{r \in \mathcal{R}_{pq}} h_{pqr} \log h_{pqr},$$

subject to

$$\begin{aligned} \sum_{r \in \mathcal{R}_{pq}} h_{pqr} &= d_{pq}, && \forall (p,q) \in \mathcal{C}, \\ \sum_{(p,q) \in \mathcal{C}} \sum_{r \in \mathcal{R}_{pq}} \delta_{pqra} h_{pqr} &= f_a, && \forall a \in \mathcal{A}, \\ h_{pqr} &\geq 0, && \forall (p,q) \in \mathcal{C}. \end{aligned}$$

For all finite values of Θ, the solution to [TA-SUE] will yield a nonzero flow for each route in the sets $\mathcal{R}_{pq}$. Since the number of routes in the network is extremely large in general (and will in particular include many unrealistic routes with respect to their cost or their topology), in any practical application the sets $\mathcal{R}_{pq}$ must be restricted to some reasonable size. Section 4.2 describes some methods for achieving this.

Most algorithms proposed for this problem have mainly been based on generating the optimal *link* flows instead of the optimal *route* flows. In the method of successive averages (MSA) (e.g., [20, Sec 12.2]), simulations of the link costs are performed iteratively, and the resulting shortest route flows are weighted together with the current flow solution using predetermined weighting factors.

Damberg *et al.* [3] combine the concept of disaggregate simplicial decomposition, taken from [12] for the solution of [TA-UE], with a partial linearization algorithm for the solution of restrictions to [TA-SUE] to route subsets $\hat{\mathcal{R}}_{pq} \subset \mathcal{R}_{pq}$ that are generated algorithmically; the term *partial linearization* refers to the fact that the underestimating function used to define the search direction is constructed by linearizing only the first term of Z. It can be shown that the solution to this problem is equivalent to applying the formula (3) with $c_{pqr} = c_{pqr}(h^k)$, h^k being the current iterate.

2.2 The combined distribution and assignment problem

The deterministic model

The combined distribution and assignment problem is the problem of the simultaneous determination of the distribution of trips between origins and destinations in a transportation network and the assignment of trips to routes in each OD pair. This is a natural extension of the user equilibrium model [TA-UE] in that it also takes into account how the demand for transportation varies with the travel cost.

We introduce subsets $\mathcal{P}$ and $\mathcal{Q}$ of $\mathcal{N}$ to denote the origin and destination zones, respectively; for notational simplicity we will assume that *all* pairs $(p, q) \in \mathcal{P} \times \mathcal{Q}$ define potential OD pairs. The total flow emanating from an origin node $p \in \mathcal{P}$ is known, and equals O_p; correspondingly, the total flow terminating at a node $q \in \mathcal{Q}$ is D_q.

We consider a model for such a problem, first formulated by Evans [8], where the trip distribution follows a gravity model with negative exponential deterrence function, according to

$$d_{pq} = a_p O_p b_q D_q \exp(-\gamma \pi_{pq}), \qquad p \in \mathcal{P}, \quad q \in \mathcal{Q}, \tag{5}$$

where γ is the dispersion parameter, and a_p and b_q are balancing factors determined so that the flows given by (5) are consistent with O_p and D_q; further, the trip assignment is made according to the deterministic user equilibrium principle.

The combined distribution and assignment problem is

[CDA-UE]

$$\text{minimize } V(d, f) = \sum_{a \in \mathcal{A}} \int_0^{f_a} c_a(s)\, ds + \frac{1}{\gamma} \sum_{p \in \mathcal{P}} \sum_{q \in \mathcal{Q}} d_{pq} \log d_{pq} \tag{6a}$$

subject to

$$\sum_{r \in \mathcal{R}_{pq}} h_{pqr} = d_{pq}, \qquad \forall p \in \mathcal{P}, \quad \forall q \in \mathcal{Q}, \tag{6b}$$

$$\sum_{q \in \mathcal{Q}} d_{pq} = O_p, \qquad \forall p \in \mathcal{P}, \tag{6c}$$

$$\sum_{p \in \mathcal{P}} d_{pq} = D_q, \qquad \forall q \in \mathcal{Q}, \tag{6d}$$

$$\sum_{p \in \mathcal{P}} \sum_{q \in \mathcal{Q}} \sum_{r \in \mathcal{R}_{pq}} \delta_{pqra} h_{pqr} = f_a, \qquad \forall a \in \mathcal{A}, \tag{6e}$$

$$h_{pqr} \geq 0, \qquad \forall r \in \mathcal{R}_{pq}, \quad \forall p \in \mathcal{P}, \quad \forall q \in \mathcal{Q}. \tag{6f}$$

The most commonly used algorithm for this problem is the partial linearization method of Evans [8]. Given an iterate (d^k, f^k), the shortest routes given link costs $c(f^k)$ are calculated. Based on the shortest route costs π_{pq}^k, the doubly constrained entropy maximization problem of minimizing $T(d) = \sum_{p \in \mathcal{P}} \sum_{q \in \mathcal{Q}} (\pi_{pq}^k d_{pq} + \frac{1}{\gamma} d_{pq} \log d_{pq})$ subject to (6c)–(6d) and $d \geq 0$ is solved. The resulting auxiliary demand is assigned onto the shortest routes; this determines a descent direction in the space of demands and link flows, and a line search defines the new iterate.

The algorithm of [14] combines Evans' algorithm with the DSD algorithm, thus essentially replacing the simple line search in Evans' algorithm with a multi-dimensional one.

The stochastic model

We next consider a model for the combined distribution and assignment problem assuming a stochastic user equilibrium principle of route-choice. This problem can be formulated as ([7])

[CDA-SUE]

$$\text{minimize } W(h) = \sum_{a \in \mathcal{A}} \int_0^{f_a} c_a(s)\, ds + \frac{1}{\Theta} \sum_{p \in \mathcal{P}} \sum_{q \in \mathcal{Q}} \sum_{r \in \mathcal{R}_{pq}} h_{pqr} \log h_{pqr} \tag{7a}$$

subject to

$$\sum_{r \in \mathcal{R}_{pq}} h_{pqr} = d_{pq}, \qquad \forall p \in \mathcal{P}, \quad \forall q \in \mathcal{Q}, \tag{7b}$$

$$\sum_{q \in \mathcal{Q}} d_{pq} = O_p, \qquad \forall p \in \mathcal{P}, \tag{7c}$$

$$\sum_{p \in \mathcal{P}} d_{pq} = D_q, \qquad \forall q \in \mathcal{Q}, \tag{7d}$$

$$\sum_{p \in \mathcal{P}} \sum_{q \in \mathcal{Q}} \sum_{r \in \mathcal{R}_{pq}} \delta_{pqra} h_{pqr} = f_a, \qquad \forall a \in \mathcal{A}, \tag{7e}$$

$$h_{pqr} \geq 0, \qquad \forall r \in \mathcal{R}_{pq}, \quad \forall p \in \mathcal{P}, \quad \forall q \in \mathcal{Q}. \tag{7f}$$

An interesting feature of this model is that a solution strategy may utilize a projection of the feasible set onto either a set described in terms of route flow variables only or, alternatively, a set described in terms of demand variables only. We next utilize a projection of [CDA-SUE] into a problem described by route flow variables in order to derive its optimality conditions. The variables d_{pq} are substituted through (7b), and the constraints (7c)–(7d) thus reduce to

$$\sum_{q\in\mathcal{Q}}\sum_{r\in\mathcal{R}_{pq}} h_{pqr} = O_p, \qquad \forall p \in \mathcal{P}, \tag{8a}$$

$$\sum_{p\in\mathcal{P}}\sum_{r\in\mathcal{R}_{pq}} h_{pqr} = D_q, \qquad \forall q \in \mathcal{Q}. \tag{8b}$$

Next, using the Lagrangean function

$$\mathcal{L}(h,\alpha,\beta) = W(h) + \sum_{p\in\mathcal{P}}\alpha_p\left(\sum_{q\in\mathcal{Q}}\sum_{r\in\mathcal{R}_{pq}} h_{pqr} - O_p\right) + \sum_{q\in\mathcal{Q}}\beta_q\left(\sum_{p\in\mathcal{P}}\sum_{r\in\mathcal{R}_{pq}} h_{pqr} - D_q\right),$$

where α_p, $p \in \mathcal{P}$, and β_q, $q \in \mathcal{Q}$, are the Lagrange multipliers for the constraints (8a) and (8b), respectively, the optimality conditions can be expressed as

$$\frac{\partial\mathcal{L}(h,\alpha,\beta)}{\partial h_{pqr}} = c_{pqr} + \frac{1}{\Theta}(\log h_{pqr} + 1) + \alpha_p + \beta_q = 0,$$

where $c_{pqr} = c_{pqr}(h)$, and equivalently written as

$$h_{pqr} = \exp(-1 - \Theta\alpha_p)\exp(-\Theta\beta_q)\exp(-\Theta c_{pqr}). \tag{9}$$

Using (7b), we obtain

$$d_{pq} = \exp(-1 - \Theta\alpha_p)\exp(-\Theta\beta_q)\sum_{r\in\mathcal{R}_{pq}}\exp(-\Theta c_{pqr}),$$

which can be used together with (9) to express the logit model (3). Also, using (9) and identifying

$$a_p O_p = \exp(-1 - \Theta\alpha_p); \qquad b_q D_q = \exp(-\Theta\beta_q); \qquad \pi_{pq} = -\frac{1}{\Theta}\log\sum_{r\in\mathcal{R}}\exp(-\Theta c_{pqr}),$$

we recognize the gravity model (5) with $\gamma = \Theta$.

Thus, the optimal solution is a gravity model with respect to the composite route cost π_{pq} in each OD pair. Note that the dispersion parameter γ at the distribution level is in this model equal to the dispersion parameter Θ at the assignment level.

The relationships between the models of this section are summarized in Table 1.

3 Derivation from the efficiency principle

In this section we show how the models presented in the previous section can be derived from an assumption of efficient trip making behaviour, defined according to the *efficiency*

		Route-choice principle	
		UE	SUE
Model	TA	[TA-UE]	[TA-SUE]
	CDA	[CDA-UE]	[CDA-SUE]

Table 1: Model relationships.

principle (Smith [21, 22]). The basic assumption of efficiency in the context of trip making is that there is an overall tendency of trip makers to prefer lower travel costs (or, time) to higher ones.

3.1 Basic concepts and notation

Consider a set of T trip makers, where each trip maker t chooses a route r_t in the transportation network. (The choice of route may also include the choice of origin and destination of the trip.) The vector of routes chosen by all T trip makers, $u = (r_1, \dots, r_T)$, defines a *trip pattern*. For each trip pattern u we can compute the corresponding *flow pattern* $h(u) = [h_{pqr}(u)]$ consisting of the number of trip makers on each of the routes in the network.

Using the assumption of efficient trip making behaviour according to the efficiency principle we can derive an expression for the probability distribution of a trip pattern, and formulate the problems of finding the most probable trip pattern and the most probable flow pattern. These problems are defined in the discrete space, so the models derived will be discrete models. The models presented in Section 2 are continuous approximations of these models. Thus, by the use of the efficiency principle, the form of the models can be given a behavioural interpretation.

For a given trip pattern u we define

$$\begin{aligned} d_{pq}(u) &= \text{number of trips in } u \text{ originating in } p \text{ and terminating in } q, \\ O_p(u) &= \text{number of trips in } u \text{ originating in } p, \\ D_q(u) &= \text{number of trips in } u \text{ terminating in } q, \\ f_a(u) &= \text{number of trips in } u \text{ using link } a, \end{aligned}$$

and, clearly,

$$\begin{aligned} d_{pq}(u) &= \sum_{r \in \mathcal{R}_{pq}} h_{pqr}(u), \\ O_p(u) &= \sum_{q \in \mathcal{Q}} d_{pq}(u), \\ D_q(u) &= \sum_{p \in \mathcal{P}} d_{pq}(u), \\ f_a(u) &= \sum_{p \in \mathcal{P}} \sum_{q \in \mathcal{Q}} \sum_{r \in \mathcal{R}_{pq}} \delta_{pqra} h_{pqr}(u). \end{aligned}$$

Depending on the choice situation we may be interested only in trip patterns where the number of trip makers in each OD pair is given (i.e., $d_{pq}(u) = d_{pq}$), corresponding to a trip assignment problem, or in trip patterns satisfying constraints on the total flow from each origin and to each destination (i.e., $O_p(u) = O_p$ and $D_q(u) = D_q$), corresponding to a combined distribution and assignment problem.

We wish to derive an expression for the probability of a given trip pattern. Let the trip pattern u be a random vector taking values in the discrete sample space U consisting of all possible trip patterns, and let $u_k \in U$, $k = 1, \ldots, K$ denote all the possible trip patterns. The number of elements in U is very large but finite, since the number of origins and destinations is finite, as is the number of routes in each origin–destination pair. We want to characterize the form of the probability distribution $p = (p_1, \ldots, p_K)$ where $p_k = P[u = u_k]$, and $\sum_{k=1}^{K} p_k = 1$.

Important questions are how to define the cost for a trip pattern capturing the route costs faced by the individual trip makers, and how this cost influences the form of the probability distribution. The cost function used is the *cumulative cost function* ([22]) which is defined as

$$c(u) = \sum_{t=1}^{T} \sum_{a \in r_t} c_a(f_a^t), \tag{10}$$

where f_a^t denotes the flow on link a when the first t trip makers in u have chosen a route in the network, and $c_a(f_a^t)$ is the corresponding cost. The cumulative cost function can be interpreted as the summation of route costs facing the trip makers at the moment they sequentially enter the network. The cumulative cost function can equivalently be expressed as ([22])

$$c(u) = \sum_{a \in A} \sum_{n=1}^{f_a(u)} c_a(n), \tag{11}$$

which shows that the function takes the same value irrespective of the order of the trips in the trip pattern. Thus, every trip maker can be considered as the last one to enter, choosing a route r_T which minimizes the cost $\sum_{a \in r_T} c_a(f_a^T) = \sum_{a \in r_T} c_a(f_a(u))$. Therefore, minimizing $c(u)$ implies a cost minimizing behaviour among the trip makers, and the problem ([22])

$$\min_{u \in U} \quad c(u) \tag{12a}$$

$$\text{s.t.} \quad d_{pq}(u) = d_{pq}, \qquad \forall p \in \mathcal{P}, \quad \forall q \in \mathcal{Q} \tag{12b}$$

yields a solution that satisfies the discrete user equilibrium conditions for the trip assignment model. The reader should note that the reformulation (11) of the cumulative cost (10) is equivalent only under the presumption that the cost functions c_a are separable.

The objective function (11) is identical to the objective function used by Rosenthal [19] to define equilibrium conditions for integer valued network flows.

The form of the probability distribution p will be derived using samples taken from the distribution. Let $(u^1, \ldots, u^S)$ be an independent sample of size S of trip patterns and let z_k denote the number of observations in the sample that is identical to trip pattern

u_k. Note that each observation in the sample includes a specification of the routes taken by all T trip makers. The probability of the sample can, because of the independence, be written as

$$P(u^1, \ldots, u^S) = \prod_{k=1}^{K} p_k^{z_k}.$$

We want to compare samples, but only compare samples which are defined to be *activity equivalent.* Depending on the problem we are interested in, trip assignment or combined distribution and assignment, activity equivalence is defined differently (see below).

A natural hypothesis about individual trip making behaviour is that, other things being equal, a trip maker tends to make less costly trips rather than more costly. In the context of the efficiency principle this hypothesis can be translated to an aggregate level by saying that we assume that trip makers as a group behave efficiently.

Definition 1 (Efficiency) [21, 22]
The probability distribution p is efficient *if and only if for every S and every pair of activity equivalent samples, $(u^1, \ldots, u^S)$ and $(w^1, \ldots, w^S)$, we have*

$$\sum_{s=1}^{S} c(u^s) \leq \sum_{s=1}^{S} c(w^s) \quad \Longrightarrow \quad P(u^1, \ldots, u^S) \geq P(w^1, \ldots, w^S).$$

Hence, when comparing activity equivalent samples from the probability distribution, samples with a lower cost are at least as probable as samples with a higher cost.

3.2 The trip assignment problem

Let us first consider the trip assignment problem. Activity equivalence is then defined in terms of the number of trips in each OD pair.

Definition 2 (Activity equivalence) [22]
Two independent trip pattern samples $(u^1, \ldots, u^S)$ and $(w^1, \ldots, w^S)$ of size S are activity equivalent *if and only if*

$$\sum_{s=1}^{S} d_{pq}(u^s) = \sum_{s=1}^{S} d_{pq}(w^s), \qquad \forall p \in P, \quad \forall q \in Q.$$

Given the definition of activity equivalence and the definition of efficiency, the following theorem can be formulated, which gives the form of the probability distribution.

Theorem 3 (Representation theorem) [22, 4]
Assume that activity equivalence is defined as in Definition 2. Then, the strictly positive probability distribution p is efficient if and only if it is of log-linear type, that is, for some

real numbers τ_{pq}, $p \in \mathcal{P}$, $q \in \mathcal{Q}$, *and* $\gamma \geq 0$,

$$p_k = \exp\left(\sum_{p\in P}\sum_{q\in Q}\tau_{pq}d_{pq}(u_k) - \gamma c(u_k)\right), \qquad k = 1,\dots,K. \tag{13}$$

The most probable trip pattern can now be derived by maximizing p_k over all u_k, $k = 1,\dots,K$, given a specified level of the OD demand d_{pq}. This optimization problem can be formulated as

$$\max_{u\in U} \quad P(u) \tag{14a}$$
$$\text{s.t.} \quad d_{pq}(u) = d_{pq}, \qquad \forall p \in \mathcal{P}, \quad \forall q \in \mathcal{Q}. \tag{14b}$$

From the form of the probability distribution (13) it follows that the first part of p_k is constant for all trip patterns satisfying the constraints in (12) or (14), and the optimal solution can therefore equivalently be found by minimizing $c(u)$. From the problem (12) we know that this solution is a discrete user equilibrium and the conclusion is that *the most probable trip patterns are discrete user equilibria.* Thus, assuming efficient population behaviour in a network gives a certain form of the probability distribution for a trip pattern, and the trip pattern obtained when maximizing this probability distribution satisfies the discrete user equilibrium conditions. The continuous approximation of the problem of finding the most probable trip pattern, which is constructed by relaxing the integrality condition and replacing the second sum in (11) with a Riemann integral [hence obtaining the objective function (2a)], can be recognized as the [TA-UE] problem.

Next, we want to find an expression for the most probable flow pattern under the efficiency assumption. There are many trip patterns corresponding to one specific flow pattern, and in order to obtain the probability of a flow pattern the probability for all trip patterns producing the same flow pattern have to be computed. Let U_h be the set of all trip patterns u that corresponds to the flow pattern h, that is,

$$U_h = \{\, u \mid h(u) = h \,\}.$$

The number of elements in U_h is

$$|U_h| = \frac{T!}{\prod_{p\in P}\prod_{q\in Q}\prod_{r\in R_{pq}} h_{pqr}!}.$$

Theorem 4 [23, 6, 13]
Let the strictly positive probability distribution $p = (p_k)$ be efficient and define activity equivalence according to Definition 2. The probability of the specific flow pattern $h = (h_{pqr})$ is given, for $\gamma \geq 0$, and for any $u \in U_h$, by

$$P(h) = |U_h|\ P(u) = |U_h|\ \exp\left(\sum_{p\in\mathcal{P}}\sum_{q\in\mathcal{Q}}\tau_{pq}d_{pq}(u) - \gamma c(u)\right).$$

The theorem can be proved by noticing that $d_{pq}(u)$ is constant for all $u \in U_h$. Also, $f_a(u)$ is constant for all $u \in U_h$ and it follows from (11) that $c(u)$ is constant for all

$u \in U_h$. Hence, $P(u)$ is constant for all $u \in U_h$ and $P(h)$ is obtained by multiplying $P(u)$ by the number of elements in U_h.

The most probable flow pattern satisfying efficient population behaviour can be obtained by maximizing the probability $P(h)$. Smith [23] showed that for large trip activities this flow pattern is the "overwhelmingly" most probable. Computing the most probable flow pattern, given a certain level of activity, $d_{pq}(u) = d_{pq}$, in the discrete space is however a computationally very difficult problem. A continuous approximation of this problem can be obtained in the following way: minimize the logarithm of $-P(h)$ and relax the integer constraints on h_{pqr}; next, use Stirling's formula ($\log n! = n \log n - n$) to approximate the factorials; then, neglect the constant terms and replace $c(u)$ by its continuous approximation. The resulting problem is recognized as the stochastic user equilibrium model [TA-SUE]. Thus, the latter model can be interpreted as a continuous approximation of a problem of finding the most probable flow pattern under efficient trip making behaviour.

3.3 The combined distribution and assignment problem

In order to derive an expression for the probability for a trip pattern and a flow pattern in a combined trip distribution and assignment problem, samples with the same total flow emanating from and terminating in each origin and destination have to be compared. Therefore, activity equivalence is defined as follows.

Definition 5 (Activity equivalence) [22, 5]
Two independent trip pattern samples of size S, $(u^1, \dots, u^S)$ and $(w^1, \dots, w^S)$, are activity equivalent *if and only if*

$$\sum_{s=1}^{S} O_p(u^s) = \sum_{s=1}^{S} O_p(w^s), \qquad \forall p \in P,$$

and

$$\sum_{s=1}^{S} D_q(u^s) = \sum_{s=1}^{S} D_q(w^s), \qquad \forall q \in Q.$$

The form of the probability distribution is given in the following theorem.

Theorem 6 (Representation theorem) [22, 4]
Assume that activity equivalence is defined as in Definition 5. Then, the strictly positive probability distribution p is efficient if and only if it is of log-linear type, that is, for some real numbers α_p, $p \in \mathcal{P}$, β_q, $q \in \mathcal{Q}$, and $\gamma \geq 0$,

$$p_k = \exp\left(\sum_{p \in P} \alpha_p O_p(u_k) + \sum_{q \in Q} \beta_q D_q(u_k) - \gamma c(u_k) \right), \qquad k = 1, \dots, K.$$

The most probable trip patterns are user equilibria also in this case. For given values of the marginal totals, these trip patterns can be found by minimizing $c(u)$ resulting in a discrete user equilibrium. However, there are many trip patterns with the same maximal probability, and they result in different OD matrices.

The probability for a given OD matrix can be calculated as the number of trip patterns producing this matrix multiplied by the corresponding probability (equal for all those trip patterns). Let $U_d = \{ u \mid d(u) = d \}$ be the set of trip patterns u corresponding to trip matrix d. The number of elements in U_d is

$$|U_d| = \frac{T!}{\prod_{p \in P} \prod_{q \in Q} d_{pq}!} n(d)$$

where $n(d)$ denotes the number of trip patterns that produce trip matrix d and differ only in how the trip makers select their route. Erlander [5] discusses in detail the problem of finding the most probable trip matrix, and concludes that $|U_d|$ can not be computed unless $n(d)$ is constant, for example, if only one route is used in each OD pair. This assumption is however rather unrealistic in practice. In this special case, the [CDA-UE] model can be interpreted as the continuous approximation of the problem of finding the most probable OD matrix.

Instead of searching for the most probable OD matrix, the most probable flow pattern under efficient trip making behaviour can be computed. The OD matrix corresponding to this flow pattern is then easily computed.

The probability for a flow pattern can be derived in a similar way as in the trip assignment case (see Theorem 4).

Theorem 7 [7]
Let the strictly positive probability distribution $p = (p_k)$ be efficient and define activity equivalence according to 5. The probability of the specific flow pattern h is given, for $\gamma \geq 0$, and for any $u \in U_h$, by

$$P(h) = |U_h| \; P(u) = |U_h| \; \exp\left(\sum_{p \in \mathcal{P}} \alpha_p O_p(u_k) + \sum_{q \in \mathcal{Q}} \beta_q D_q(u) - \gamma c(u) \right). \quad (15)$$

The most probable flow pattern satisfying efficient population behaviour is found by maximizing (15). The continuous approximation of this problem is obtained in the same way as in the trip assignment case, whence we obtain the combined distribution and assignment model with stochastic user equilibrium route choice [CDA-SUE].

4 An algorithm for [CDA-SUE]

The algorithm which we will outline for the solution of [CDA-SUE] has its foundations in the disaggregate simplicial decomposition (DSD) algorithm of Larsson and Patriksson [12] for the solution of [TA-UE], and the extensions of this method to the solution of [TA-SUE] (Damberg *et al.* [3]) and [CDA-UE] (Lundgren and Patriksson [14]). The algorithm alternates between two phases: (1) given known subsets $\hat{\mathcal{R}}_{pq} \subseteq \mathcal{R}_{pq}$, $p \in \mathcal{P}$, $q \in \mathcal{Q}$, of the total sets of routes in the network, the corresponding restriction of [CDA-SUE] (the *restricted master problem*) is solved (approximately) using a partial linearization (descent) algorithm; and (2) at the (approximate) solution to this restricted master problem, the subsets $\hat{\mathcal{R}}_{pq}$ are augmented by the generation of new routes, for example through the solution of a set of shortest route problems, given appropriately chosen link costs (the *column generation problem*).

4.1 The restricted master problem phase

Consider the restriction of [CDA-SUE], where subsets $\hat{\mathcal{R}}_{pq}$ of $\mathcal{R}_{pq}$, $p \in \mathcal{P}$, $q \in \mathcal{Q}$, are known. This problem will henceforth be referred to as [CDA-SUE-R]. This restriction is solved by the use of a descent algorithm, which is briefly outlined below.

We use the projection of [CDA-SUE] onto the set of feasible route flows. Given a feasible iterate h^k at some iteration k, we (roughly only) solve an approximation of [CDA-SUE-R] in order to define an auxiliary feasible solution and a search direction. The approximate problem is constructed as follows: we linearize the first term of the objective function W of [CDA-SUE-R]; the effect of this linearization is that the link costs are fixed at their levels given the current flow h^k. This linearization is analogous to those made in the Frank–Wolfe algorithm for [TA-UE] (e.g., [18, Sec. 4.1]), in Evans' [8] algorithm and its extension presented in [14] for [CDA], and in the algorithm of [3] for [SUE]. The corresponding route costs are calculated as

$$c_{pqr}^k = c_{pqr}(h^k) = \sum_{a \in \mathcal{A}} \delta_{pqra} c_a(f_a^k), \qquad r \in \hat{\mathcal{R}}_{pq}, \quad p \in \mathcal{P}, \quad q \in \mathcal{Q},$$

where f^k is the link flow corresponding to the route flow h^k.

The partially linearized problem becomes

[CDA-SUE-R-PL]

$$\text{minimize} \sum_{p \in \mathcal{P}} \sum_{q \in \mathcal{Q}} \sum_{r \in \hat{\mathcal{R}}_{pq}} \left(c_{pqr}^k h_{pqr} + \frac{1}{\Theta} h_{pqr} \log h_{pqr} \right)$$

subject to

$$\begin{aligned}
\sum_{q \in \mathcal{Q}} \sum_{r \in \hat{\mathcal{R}}_{pq}} h_{pqr} &= O_p, \qquad \forall p \in \mathcal{P}, \\
\sum_{p \in \mathcal{P}} \sum_{r \in \hat{\mathcal{R}}_{pq}} h_{pqr} &= D_q, \qquad \forall q \in \mathcal{Q}, \\
h_{pqr} &\geq 0, \qquad \forall r \in \hat{\mathcal{R}}_{pq}, \quad \forall p \in \mathcal{P}, \quad \forall q \in \mathcal{Q}.
\end{aligned}$$

An optimal solution to this problem is obtained by first solving the entropy maximization problem

$$\text{minimize} \sum_{p \in \mathcal{P}} \sum_{q \in \mathcal{Q}} \left(c_{pq}^k d_{pq} + \frac{1}{\Theta} d_{pq} \log d_{pq} \right) \tag{17a}$$

subject to

$$\sum_{q \in \mathcal{Q}} d_{pq} = O_p, \qquad \forall p \in \mathcal{P}, \tag{17b}$$

$$\sum_{p \in \mathcal{P}} d_{pq} = D_q, \qquad \forall q \in \mathcal{Q}, \tag{17c}$$

$$d_{pq} \geq 0, \qquad \forall p \in \mathcal{P}, \quad \forall q \in \mathcal{Q}, \tag{17d}$$

where

$$c_{pq}^k = -\frac{1}{\Theta} \log \sum_{r \in \hat{\mathcal{R}}_{pq}} \exp(-\Theta c_{pqr}^k), \tag{17e}$$

for example by using a balancing procedure (see, e.g., Lamond and Stewart [11]). The result is an auxiliary demand $\underline{d}^k$; the solution to [CDA-SUE-R-PL] is then obtained by applying the logit formula [cf. (3)]

$$\underline{h}_{pqr}^k = \underline{d}_{pq}^k \frac{\exp(-\Theta c_{pqr}^k)}{\sum_{\ell \in \hat{\mathcal{R}}_{pq}} \exp(-\Theta c_{pq\ell}^k)}, \qquad r \in \hat{\mathcal{R}}_{pq}, \quad p \in \mathcal{P}, \quad q \in \mathcal{Q}. \tag{18}$$

This procedure for solving [CDA-SUE-R-PL] can be motivated by viewing its projection onto the demand space. Consider the following equivalent formulation of the problem:

$$\text{minimize } U(d) \tag{19a}$$

subject to

$$\sum_{q \in \mathcal{Q}} d_{pq} = O_p, \qquad \forall p \in \mathcal{P}, \tag{19b}$$

$$\sum_{p \in \mathcal{P}} d_{pq} = D_q, \qquad \forall q \in \mathcal{Q}, \tag{19c}$$

$$d_{pq} \geq 0, \qquad \forall p \in \mathcal{P}, \quad \forall q \in \mathcal{Q}. \tag{19d}$$

where

$$U(d) = \min_h \sum_{p \in \mathcal{P}} \sum_{q \in \mathcal{Q}} \sum_{r \in \hat{\mathcal{R}}_{pq}} \left(c_{pqr}^k h_{pqr} + \frac{1}{\Theta} h_{pqr} \log h_{pqr} \right) \tag{19e}$$

$$\text{s.t.} \sum_{r \in \mathcal{R}_{pq}} h_{pqr} = d_{pq}, \qquad \forall p \in \mathcal{P}, \quad \forall q \in \mathcal{Q}, \tag{19f}$$

$$h_{pqr} \geq 0, \qquad \forall r \in \hat{\mathcal{R}}_{pq}, \quad \forall p \in \mathcal{P}, \quad \forall q \in \mathcal{Q}. \tag{19g}$$

This formulation utilizes the fact that the solution to (19e)–(19g) is easily obtained by the use of the logit formula [cf. (3)]

$$h_{pqr} = d_{pq} \frac{\exp(-\Theta c_{pqr}^k)}{\sum_{\ell \in \hat{\mathcal{R}}_{pq}} \exp(-\Theta c_{pq\ell}^k)}, \qquad r \in \hat{\mathcal{R}}_{pq}, \quad p \in \mathcal{P}, \quad q \in \mathcal{Q};$$

By performing the above substitution in (19e), it is an easy matter to prove that the implicit function U actually has the explicit form (17a). Hence, it is clear that [CDA-SUE-R-PL] is solved through the solution of the entropy maximization problem (17) followed by the application of the logit formula (18).

An (approximative) line search is then made with respect to W in the (feasible) direction of $\underline{h}^k - h^k$, resulting in the new solution h^{k+1} and the corresponding demand solution d^{k+1}, and the process is repeated with $k := k+1$, until some convergence criterion terminates the solution of [CDA-SUE-R].

The algorithm described briefly above is an instance of the partial linearization algorithm, which is analyzed in [15, 16, 17]. According to [16, Lem. 1] and [15, Thm. 2.1], $\underline{h}^k = h^k$ holds if and only if h^k solves [CDA-SUE-R]; otherwise, the direction of $\underline{h}^k - h^k$ is a feasible direction of descent with respect to the objective function W of [CDA-SUE-R]. Since W is strictly convex in h, the sequence $\{h^k\}$ of route flows converges to the unique solution of the restricted master problem from any feasible initial flow h^0; see [15, Thm. 2.2]; the same property holds of course for the sequence $\{d^k\}$ of demands.

We finally remark that the convergence analysis of the class of partial linearization methods allows for a large freedom-of-choice in the selection of the accuracy by which each step is carried out; this introduces a trade-off between, say, the number of main iterations of the partial linearization algorithm to be performed and the accuracy by which the entropy maximization problem is solved.

The restricted master problem is combined with a column (or, route) generation scheme, that is, a method for augmenting the sets $\hat{\mathcal{R}}_{pq}$. Strategies for generating routes are discussed a detail in [3]; we provide a short summary in the next section based on the numerical experience reported there.

4.2 The column generation phase

The descent approach described in the previous section solves the restricted master problem given subsets $\hat{\mathcal{R}}_{pq}$ of the complete sets of routes. How well this solution describes the real travel pattern depends on how well the subsets $\hat{\mathcal{R}}_{pq}$ approximate the sets of routes that are actually used. The choice of strategy by which the routes are generated is obviously of crucial importance. The column generation strategy should, ideally, be such that only the routes that are most utilized will be generated, and that the error made by assuming that the remaining routes have zero flow is small. There exist many ways of generating routes; in fact, in the proposed algorithm the definition and generation of the elements of the sets $\hat{\mathcal{R}}_{pq}$ are quite arbitrary.

The two route generation strategies suggested and evaluated in [3] are based on the calculation of shortest routes given the solution to a restricted master problem. The first approach is to generate routes through the solution of shortest route problems based on the *actual* (i.e., deterministic) travel costs, thereby temporarily ignoring the random components of the travel costs. This is the familiar approach used in column generation algorithms for [TA-UE] (e.g., Larsson and Patriksson [12]). The second approach is to generate routes based on *randomized* link costs, in this way incorporating the differences in the travelers' perceptions of travel costs. The latter strategy was found to lead to the generation of significantly more routes than the first one.

To comply with the independence assumption on the random cost terms of the logit model, each set $\hat{\mathcal{R}}_{pq}$ of routes should consist only of routes that have no links in common, that is, the routes should not overlap each other. The algorithm of [3] is supplied with a mechanism for identifying and deleting routes that overlap to any pre-specified degree. This mechanism can also be used to keep the number of routes generated to a realistic level.

In both approaches, initial sets of routes may be defined by a rough approximation of the solution to [TA-UE], for instance by a few steps of an *incremental assignment* algorithm (see [20, pp. 115–116]), or by an early truncation of a column generation algorithm for [TA-UE]. (The latter strategy is used in the algorithm of [3], which uses the DSD algorithm of [12] to create the starting flow.)

The algorithm terminates when no new routes are obtained from the column generation phase. In neither of the two above column generation approaches is the algorithm guaranteed to converge to the unique optimal solution of [CDA-SUE], since it, in general, terminates before all routes are generated. It is, however, guaranteed to solve the restriction of [CDA-SUE] to any set of routes generated.

4.3 Extensions of the model

The parameter γ in a gravity model and the parameter Θ in a logit model may be interpreted as measures of the travellers' cost sensitivity at the route choice and origin–destination choice levels, respectively, and to what extent the trips are dispersed among the alternatives. In the [CDA-SUE] model the parameters have identical values, indicating that the travellers are equally sensitive to cost at both choice levels; this fact stems from the derivation of the model and the assumption of efficient trip making behaviour according to the efficiency approach.

It is possible to formulate combined distribution and assignment models having different values of the parameters at the route choice and origin–destination choice levels ([2]). This implies a nested choice structure. It is an open question whether the efficiency approach can be used to derive such a nested model. If we use the objective function

$$\text{minimize} \sum_{a \in \mathcal{A}} \int_0^{f_a} c_a(s)\, ds + \frac{1}{\Theta} \sum_{p \in \mathcal{P}} \sum_{q \in \mathcal{Q}} \sum_{r \in \mathcal{R}_{pq}} \left(h_{pqr} \log \frac{h_{pqr}}{d_{pq}} \right) + \frac{1}{\gamma} \sum_{p \in \mathcal{P}} \sum_{q \in \mathcal{Q}} \left(d_{pq} \log d_{pq} \right),$$

then the optimal solution can be shown to satisfy the gravity model (5) at the origin–destination level (defining π_{pq} as the composite route choice cost), and a logit model (3) at the route choice level. (Note that this function is strictly convex whenever $\Theta > \gamma$, that is, whenever the sensitivity to cost at the route choice level is greater than that at the origin–destination choice level.)

The proposed algorithm can easily be adapted to cover also this general nested case; the parameter γ is then used in the balancing procedure to compute the auxiliary demand $\underline{d}^k$, while the parameter Θ is used in the logit formula to compute the auxiliary route flow $\underline{h}^k$.

We can also extend the model to include a mode choice level, in this way building a three level nested choice structure. The objective function will then contain a component describing the dispersion over the modes and a corresponding third parameter. The proposed algorithm can be adapted also to this case.

References

[1] D. Bernstein and T. E. Smith, Programmability of discrete network equilibria, Statistics and Operations Research Technical Report SOR-94-19, Princeton University, Princeton, NJ, 1994.

[2] D. E. Boyce, K. S. Chon, Y. J. Lee, K. T. Lin, and L. J. LeBlanc, Implementation and computational issues for combined models of location, destination, mode and route choice, Environment and Planning, A15 (1983) 1219–1230.

[3] O. Damberg, J. T. Lundgren, and M. Patriksson, An algorithm for the stochastic user equilibrium problem, Transportation Research, 30B (1996) 115–131.

[4] S. Erlander, On the principle of monotone likelihood and log-linear models, Mathematical Programming Study, 25 (1985) 108–123.

[5] S. Erlander, On the relationship between the discrete and continuous models for combined distribution and assignment, Transportation Research, 22B (1988) 371–382.

[6] S. Erlander, Efficient population behavior in networks, dispersed equilibrium and the combined distribution and assignment model, Report LiTH-MAT-R-89-07, Department of Mathematics, Linköping Institute of Technology, Linköping, Sweden, 1989.

[7] S. Erlander, Efficient behavior and the simultaneous choices of origins, destinations and routes, Transportation Research, 24B (1990) 363–373.

[8] S. P. Evans, Derivation and analysis of some models for combining trip distribution and assignment, Transportation Research, 10 (1976) 37–57.

[9] C. Fisk, Some developments in equilibrium traffic assignment, Transportation Research, 14B (1980) 243–255.

[10] M. Frank and P. Wolfe, An algorithm for quadratic programming, Naval Research Logistics Quarterly, 3 (1956) 95–110.

[11] B. Lamond and N. F. Stewart, Bregman's balancing method, Transportation Research, 15B (1981) 239–248.

[12] T. Larsson and M. Patriksson, Simplicial decomposition with disaggregated representation for the traffic assignment problem, Transportation Science, 26 (1992) 4–17.

[13] J. T. Lundgren, *Optimization approaches to travel demand modelling*, Doctoral dissertation, Department of Mathematics, Linköping Institute of Technology, Linköping, Sweden, 1989.

[14] J. T. Lundgren and M. Patriksson, An algorithm for the combined distribution and assignment model, Report, Department of Mathematics, Linköping University, Linköping, Sweden.

[15] M. Patriksson, Partial linearization methods in nonlinear programming, Journal of Optimization Theory and Applications, 78 (1993) 227–246.

[16] M. Patriksson, A unified description of iterative algorithms for traffic equilibria, European Journal of Operational Research, 17 (1993) 154–176.

[17] M. Patriksson, *A unified framework of descent algorithms for nonlinear programs and variational inequalities*, Doctoral dissertation, Department of Mathematics, Linköping Institute of Technology, Linköping, Sweden, 1993.

[18] M. Patriksson, *The Traffic Assignment Problem—Models and Methods*, Topics in Transportation Series, VSP BV, Utrecht, The Netherlands, 1994.

[19] R. W. Rosenthal, The network equilibrium problem in integers, Networks, 3 (1973) 53–59.

[20] Y. Sheffi, *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*, Prentice-Hall, Englewood Cliffs, NJ, 1985.

[21] T. E. Smith, A cost-efficiency principle of spatial interaction behaviour, Regional Science and Urban Economics, 8 (1978) 313–337.

[22] T. E. Smith, A cost-efficiency approach to the analysis of congested spatial-interaction behavior, Environment and Planning, 15A (1983) 435–464.

[23] T. E. Smith, A cost-efficiency theory of dispersed network equilibria, Environment and

Planning, 20A (1988) 231–261.
[24] J. G. Wardrop, Some theoretical aspects of road traffic research, Proceedings of the Institute of Civil Engineers, Part II (1952) 325–378.