2024 IEEE 63rd Conference on Decision and Control (CDC)
December 16-19, 2024. MiCo, Milan, Italy

Controlling Traffic Flow for Electric Fleets via Optimal Transport

Michele Mascherpa and Johan Karlsson

Abstract— **In this paper we consider the problem of optimally steering an ensemble of battery-powered agents over a network. This is an important problem in applications such as traffic flow control for electric vehicles, where both capacity constraints from the roads and the locations of charging stations need to be taken into account. We extend previous work where origin-destination problems have been formulated using optimal transport. By introducing a state representing the charge level, we can formulate the steering problem as a structured multi-marginal optimal transport problem. The computational method is based on a dual coordinate ascent algorithm applied to the entropy regularized problem, in which we can exploit the decomposable structure of the cost tensor for efficient computations. In this formulation the capacity constraints are represented in terms of certain linear operators, and we derive explicit expressions for the corresponding updates of blocks of the dual variables. Finally, the method is illustrated with a numerical example where vehicles having different charges are required to travel over a grid from origin to destination while minimizing the total energy consumed.**

I. Introduction

With technology advancing rapidly, electric vehicles (EV) and electric robots are becoming ubiquitous in many areas of the society. In the coming years, EVs will have a key role in smart cities [1], while fleets of electric robots could be employed in numerous sectors, such as warehouse operations [2], mining industry [3] and agricultural applications [4]. In all these contexts, the batteries and their relatively long recharging times still represent a major limitation. Furthermore, in many of these applications, vehicles and robots are often not indistinguishable as vehicles can have assigned origins and destinations, or the tasks that robots can performs might not be interchangeable, making the macroscopic steering problem more challenging. Due to the increasing interest in this area, there has recently been a large amount of research proposing strategies and optimization methods for routing and control of electric vehicles [5], [6], [7], [8], [9], [10], [11], [12]. Nevertheless, the results for effective solutions of the steering problem of electric fleets, including both origin-destination and charge constraints, is limited and leave space for additional contributions. An example of a routing problem for electric vehicles, incorporating such constraints is given in [13]. The problem is there formulated as a MILP, which is known to be NP-hard, and a metaheuristic bilevel solution is proposed.

A classical approach to the macroscopic traffic control problem is the dynamic network flow formulation. This method goes back to the work [14], and the procedure builds on transforming the discrete time-horizon to a static problem by time-expansion of the underlying network. However, by doing this, the time structure in the problem is not fully utilized, and thus limits current techniques in terms of number of commodities and problem sizes. Therefore, the focus has mainly been to include only few commodities representing, e.g., a couple of classes of vehicles [15]. An alternative approach is the optimal transport framework, which has recently been used to solve control, steering, and estimation problems [16], [17], [18], [19] and utilizes the underlying time structure. Indeed, by exploiting the graph structure of the constraints imposed on the marginals, we can efficiently solve problems such as the origin-destination problem [20], and multi-commodity flow problems [21].

In this paper we formulate the steering problem over a network of an electric fleet with origin-destination constraint, using a multi-marginal optimal transport formulation. We first introduce the charge state of the batteries, creating multiple copies of the network for different charge levels. Then we formulate a multi-marginal optimal transport problem on the charge-expanded graph, which we approximate with the addition of entropic regularization. We propose an algorithm for solving the problem, based on dual coordinate ascent. This generalizes the results in [20] as the new formulation can handle charge constraints. We also derive explicit expressions for the updates in the blocks dual ascent method, which leads to an efficient algorithm. Finally, we illustrate the algorithm on a numerical example where fleets of different agents are steered from origin to destination over a grid, while maintaining positive charge.

II. Background

A. Notation

By $\odot$, $\oslash$, $\exp(\cdot)$, $\log(\cdot)$ we denote element-wise multiplication, division, exponential and logarithm of matrices and vectors. The Kronecker product of two matrices is denoted by $\otimes$. With $\mathbf{1}_n$ we denote a $n \times 1$ column vector of ones (or just $\mathbf{1}$ if the dimension is clear from the context). Similarly $\mathbf{I}_n$ is the identity matrix of size n. The set of integers from 1 to n is indicated with $[n]$. The set of non-negative real numbers is denoted by $\mathbb{R}_+ = [0, +\infty)$. By $\langle \cdot, \cdot \rangle$ we indicate the standard Frobenius inner product between vectors, matrices or tensors.

B. Optimal Transport

The optimal transport problem is to find the most efficient way to transport resources from one distribution to another while minimizing the transportation costs [22]. In the discrete setting, which will be of interest for this paper, the

This work was supported by KTH Digital Futures and by the Swedish Research Council (VR) under grant 2020-03454

The authors are with the Department of Mathematics, KTH Royal Institute of Technology, Stockholm, Sweden micmas@kth.se, johan.karlsson@math.kth.se

979-8-3503-1632-2/24/$31.00 ©2024 IEEE

distributions can be represented by two non-negative vectors $\mu_0, \mu_1 \in \mathbb{R}_+^n$, and costs by a matrix $C \in \mathbb{R}_+^{n\times n}$, where C_{ij} is the cost of moving one unit of mass from location i to j. Analogously we consider the transport plan $M \in \mathbb{R}_+^{n\times n}$ between μ_0 and μ_1, where M_{ij} is the amount of mass moved from location i to j. The optimal transport problem can then be formulated as

$$\begin{aligned} \underset{M\in\mathbb{R}^{n\times n}}{\text{minimize}} \quad & \langle C, M\rangle \\ \text{subject to} \quad & M\mathbf{1}_n = \mu_0, \quad M^T\mathbf{1}_n = \mu_1. \end{aligned} \tag{1}$$

Problem (1) has been generalized to a multi-marginal setting [23], [24], in which $(\mathcal{T}+1)$-mode tensors $\mathbf{C}, \mathbf{M} \in \mathbb{R}_+^{n^{\mathcal{T}+1}}$ represent the cost and the transport plan between a set of $\mathcal{T}+1$ marginals $\mu_0, \ldots, \mu_{\mathcal{T}} \in \mathbb{R}_+^n$. Each entry of the tensors $\mathbf{C}, \mathbf{M}$ will be denoted by a tuple of indexes $\mathbf{i} = (i_0, \ldots, i_{\mathcal{T}}) \in [n]^{\mathcal{T}+1}$, so that $\mathbf{C_i}$ is the cost and $\mathbf{M_i}$ is the mass associated with the tuple $\mathbf{i} = (i_0, \ldots, i_{\mathcal{T}})$. The multi-marginal optimal transport problem can then be formulated as

$$\begin{aligned} \underset{\mathbf{M}\in\mathbb{R}_+^{n^{\mathcal{T}+1}}}{\text{minimize}} \quad & \langle \mathbf{C}, \mathbf{M}\rangle \\ \text{subject to} \quad & P_t(\mathbf{M}) = \mu_t, \quad t = 0, \ldots, \mathcal{T}, \end{aligned} \tag{2}$$

where $P_t(\mathbf{M}) \in \mathbb{R}_+^n$ is the projection operator on the t-th mode of the tensor $\mathbf{M}$, defined by

$$(P_t(\mathbf{M}))_{i_t} := \sum_{i_1,\ldots,i_{t-1},i_{t+1},\ldots,i_{\mathcal{T}}} \mathbf{M}_{i_0,\ldots,i_{t-1},i_t,i_{t+1},\ldots,i_{\mathcal{T}}}. \tag{3}$$

In recent years its application to large scale problems have become widespread, mainly thanks to entropy regularization [25]. This approach has been applied also to the multi-marginal setting [26], but the complexity of computing the projections scales exponentially with the number of marginals, making it in general computationally intractable. However, for cost tensors $\mathbf{C}$ corresponding to certain graph-structures, these structures can be utilized to compute the projections efficiently, e.g., for tree graphs or graphs with small treewidth [27], [28], [21], [29]. This is also the case for the model proposed in this paper. The multi-marginal entropy regularized optimal transport problem is

$$\begin{aligned} \underset{\mathbf{M}\in\mathbb{R}_+^{n^{\mathcal{T}+1}}}{\text{minimize}} \quad & \langle \mathbf{C}, \mathbf{M}\rangle + \epsilon D(\mathbf{M}) \\ \text{subject to} \quad & P_t(\mathbf{M}) = \mu_t, \quad t = 0, \ldots, \mathcal{T}, \end{aligned} \tag{4}$$

where $\epsilon > 0$ is the regularization parameter, and

$$D(\mathbf{M}) = \sum_{\mathbf{i}\in[n]^{\mathcal{T}+1}} \left(\mathbf{M_i}\log(\mathbf{M_i}) - \mathbf{M_i} + 1\right)$$

is an entropy term. The addition of D makes problem (4) strictly convex. It can be shown that the unique solution of (4) has the form $\mathbf{K}\odot\mathbf{U}$, where $\mathbf{K} = \exp(-\mathbf{C}/\epsilon)$, while $\mathbf{U} = \bigotimes_{t=0}^{\mathcal{T}} u_t$, in which $u_t = \exp(-\lambda_t/\epsilon)$ for $t = 0, 1 \ldots, \mathcal{T}$. Here $\lambda_t \in \mathbb{R}^n$ are the Lagrangian multipliers associated with the constraints $P_t(\mathbf{M}) = \mu_t$. The vectors u_t can be found as the fixed point of the Sinkhorn iterations

$$u_t \leftarrow (u_t \odot \mu_t) \oslash P_t(\mathbf{K}\odot\mathbf{U}), \tag{5}$$

which can also be seen as block coordinate ascent to the dual of (4), and therefore are guaranteed to converge [30], [31]. In this paper we generalize these concepts in order to apply the theory to traffic flow problems with charge constraints.

III. Problem formulation

In this section we formulate the problem of optimally steering an ensemble of electric vehicles over a network. First, we expand the network by introducing charge levels, and then we describe how to formulate the steering problem as a multi-marginal optimal transport problem, taking into account the temporal structure of the problem. Similar to [20], we also include origin-destination constraints, allowing agents with different tasks or destinations to be distinguished from one another.

A. Expanded network with charge levels

We consider an ensemble of agents, each of them equipped with a battery with a certain charge level. The ensemble operates on a directed network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{v_1, \ldots, v_m\}$ is the set of vertices and where $\mathcal{E} = \{e_1, \ldots, e_n\}$ is the set of directed edges, so that each edge is a tuple of vertices (v_i, v_j) for $v_i, v_j \in \mathcal{V}$. We further assume that a subset of nodes $\mathcal{C} \subseteq \mathcal{V}$ represents charging stations, i.e., nodes where agents can increase the charge of their battery. The charge level is modeled with $Q+1$ discrete levels of charge, where $q = 0$ corresponds to 0% and $q = Q$ corresponds to 100%. We create a charge-expanded network $\tilde{\mathcal{G}} = (\tilde{\mathcal{V}}, \tilde{\mathcal{E}})$, such that the charge-expanded node set $\tilde{\mathcal{V}}$ consists of $Q+1$ copies of the original node set $\mathcal{V}$:

$$\tilde{\mathcal{V}} := \{(v_i, q) : v_i \in \mathcal{V},\ q = 0, \ldots, Q\}.$$

The state space is then defined as $\tilde{\mathcal{E}} := \tilde{E} \cup \tilde{C}$. Here, the charge-expanded edge set

$$\tilde{E} = \{((v_i, q), (v_j, q-1)) : (v_i, v_j) \in \mathcal{E},\ q = 1, \ldots, Q)\}$$

contains Q copies of every original edge $(v_i, v_j) \in \mathcal{E}$, where the copies represent different charge levels of the same physical edge. The set $\tilde{\mathcal{C}}$ represents the states where vehicles are stationary and charging,

$$\tilde{\mathcal{C}} = \{((v_i, q-1), (v_i, q)) : v_i \in \mathcal{C},\ q = 1, \ldots, Q\},$$

which consists of Q copies each element of $\mathcal{C}$ representing different charge levels. This means that it is possible, for an agent who is stationary in a node in $\mathcal{C}$, to increase its charge by one level at a time. This is a simplified model where travelling along any road consumes exactly one level of charge and where recharging is also one level per time unit. However, it is straightforward to generalize this to settings where, e.g., recharge pace and loss of charge depends on properties of the road or the vehicle type (cf. [21], [17]).

The total number of states is $\tilde{n} := |\tilde{\mathcal{E}}| = (|\mathcal{E}|+|\mathcal{C}|)\times Q$, and we define an ordering among them for subsequent reference:

$$\tilde{e}_{(k-1)Q+q} = \begin{cases} ((v_i, q), (v_j, q-1)) \\ \quad \text{for } 1 \le q \le Q, (v_i, v_j) = e_k, 1 \le k \le n, \\ ((v_{i_{k-n}}, q-1), (v_{i_{k-n}}, q)) \\ \quad \text{for } 1 \le q \le Q, v_{i_{k-n}} \in \mathcal{C}, k \ge n+1. \end{cases}$$

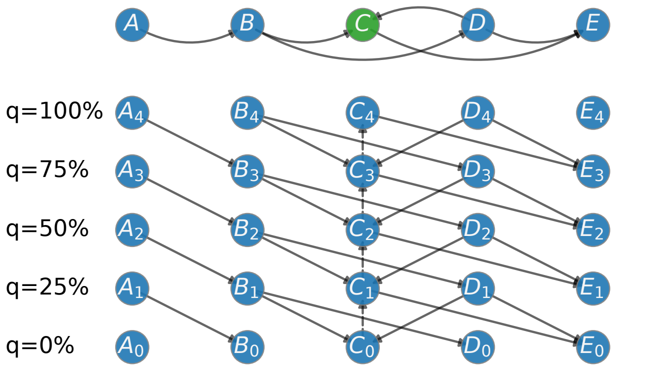

Fig. 1. The network $\mathcal{G}$ (above), with the charging station in green, and the corresponding charge expanded network $\tilde{G}$ for $Q = 4$ (below). The solid edges represent $\tilde{E}$ while dashed edges illustrate $\tilde{\mathcal{C}}$.

Example 1: We consider a network $\mathcal{G}$ with 5 nodes, 6 edges, and only 1 charging station (the node C). Both $\mathcal{G}$, and the corresponding charge-expanded network $\tilde{\mathcal{G}}$ for $Q = 4$ are illustrated in Figure 1.

Remark 1: This approach is similar to the time-expanded network used in the dynamic network flow formulation of the minimum cost network flow problem [14], [32]. However, rather than expanding the graph in both time and charge dimensions, we introduce the time dimension using an entropy regularized multi-marginal optimal transport approach, circumventing the need to solve a large linear program.

B. Steering of ensembles on the expanded network

In this section, we outline how to include the charge-expanded network $\tilde{G} = (\tilde{\mathcal{V}}, \tilde{\mathcal{E}})$, described in Section III-A, into an optimal transport framework. In a setting with $\mathcal{T}+1$ time steps, we introduce the cost tensor $\mathbf{C} \in \mathbb{R}_+^{\tilde{n}^{(\mathcal{T}+1)}}$. The key assumption is that $\mathbf{C}$ decouples as the sum over time of a cost matrix $C \in \mathbb{R}_+^{\tilde{n}\times\tilde{n}}$, in which the element C_{ij} represents the cost of going from state $\tilde{e}_i$ to state $\tilde{e}_j$ in $\tilde{\mathcal{E}}$. We obtain

$$\mathbf{C}_{i_0,\dots,i_\mathcal{T}} = \sum_{t=1}^{\mathcal{T}} C_{i_{t-1}i_t}. \tag{6}$$

The element $\mathbf{C}_{i_0,\dots,i_\mathcal{T}}$ thus represents the cost of taking the path formed by the edges $\tilde{e}_{i_0}, \dots, \tilde{e}_{i_\mathcal{T}}$. Here we use a constant cost matrix C for ease of notation, but our method easily extends to the time-varying scenario, allowing for networks that change over time.

We introduce the transport plan $\mathbf{M} \in \mathbb{R}_+^{\tilde{n}^{(\mathcal{T}+1)}}$: its component $\mathbf{M}_{i_0,\dots,i_\mathcal{T}}$ accounts for the amount of mass transported over the path formed by the edges $\tilde{e}_{i_0}, \dots, \tilde{e}_{i_\mathcal{T}}$. Our objective will then be to minimize $\langle \mathbf{C}, \mathbf{M} \rangle$, while satisfying origin-destination and capacity constraints. Origin-destination constraints are commonly encountered in traffic applications, where agents are required to have distinct identities. In such scenarios, the flow of agents from one location to another can be effectively represented using an Origin-Destination (OD) matrix [33], which encodes the number of agents traveling between specific origin-destination pairs. In our context, it is reasonable to assume that both initial position and charge are known, while the target is to reach a specified location, regardless of the battery level at the end. Final states with higher charge can be made more attractive by adjusting the cost of the paths reaching them. Exploiting the previously introduced ordering of the states, we define

$$B := \mathbf{I}_{|\mathcal{E}|+|\mathcal{C}|} \otimes \mathbf{1}_Q. \tag{7}$$

The OD matrix is then $R \in \mathbb{R}_+^{\tilde{n}\times(|\mathcal{E}|+|\mathcal{C}|)}$, and we impose the constraint as $P_{0,\mathcal{T}}(\mathbf{M})B = R$, where $P_{t_1,t_2} : \mathbb{R}^{\tilde{n}^{\mathcal{T}+1}} \to \mathbb{R}^{\tilde{n}\times\tilde{n}}$ is the bi-marginal projection operator, defined as

$$(P_{t_1,t_2}(\mathbf{M}))_{i_{t_1}i_{t_2}} := \sum_{i_0,\dots,i_\mathcal{T}\setminus\{i_{t_1},i_{t_2}\}} \mathbf{M}_{i_0,\dots,i_\mathcal{T}}, \tag{8}$$

while the matrix $B \in \{0,1\}^{\tilde{n}\times(|\mathcal{E}|+|\mathcal{C}|)}$ is utilized to aggregate the duplicates of the edges considered as destinations, by summing over all charge levels. We also assume that the edges in the original graph, as well as the charging stations, have a maximum capacity $d \in \mathbb{R}_+^{(|\mathcal{E}|+|\mathcal{C}|)}$. Similarly to the origin-destination constraint, we need to impose a maximum capacity on the aggregated Q duplicates in $\tilde{\mathcal{E}}$. The constraint can be written as $B^T P_t(\mathbf{M}) \le d$, for $t = 1, \dots, \mathcal{T}-1$.

Remark 2: In traffic applications, one often assumes nodes as origins and destination, instead of edges, as done in this work. The two approaches are equivalent. Indeed, given the sources and sinks nodes sets $\mathcal{S}^+, \mathcal{S}^- \subseteq \mathcal{V}$, one can augment the state space $\tilde{\mathcal{E}}$ with the sets $\tilde{\mathcal{S}}^+, \tilde{\mathcal{S}}^-$, defined as

$$\tilde{\mathcal{S}}^+ := \{((v_i, q), (v_i, q)) : v_i \in \mathcal{S}^+,\ q = 0, \dots, Q\},$$
$$\tilde{\mathcal{S}}^- := \{((v_i, q), (v_i, q)) : v_i \in \mathcal{S}^-,\ q = 0, \dots, Q\}.$$

Then B can be adjusted so that it performs also a summation over the copies of $\mathcal{S}^+, \mathcal{S}^-$ representing different charge levels of the same physical origin or destination, i.e.,

$$B := \begin{bmatrix} \mathbf{I}_{|\mathcal{E}|+|\mathcal{C}|} \otimes \mathbf{1}_Q & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{|\mathcal{S}^+|+|\mathcal{S}^-|} \otimes \mathbf{1}_{Q+1} \end{bmatrix}. \tag{9}$$

Even though the capacity constraints are here imposed on the edges, it would be possible to instead consider capacity constraints on the vertices by aggregating all the edges entering into each node.

The problem that we are solving is then (4) with modified constraints

$$\underset{\mathbf{M}\in\mathbb{R}_+^{\tilde{n}^{\mathcal{T}+1}}}{\text{minimize}} \quad \langle \mathbf{C}, \mathbf{M} \rangle + \epsilon D(\mathbf{M}) \tag{10a}$$
$$\text{subject to} \quad P_{0,\mathcal{T}}(\mathbf{M})B = R \tag{10b}$$
$$B^T P_t(\mathbf{M}) \le d, \quad \text{for } t = 1, \dots, \mathcal{T}-1. \tag{10c}$$

The graph-structure of the problem is illustrated in Figure 2, where the marginals of (10) are shown as circular nodes. Two nodes are connected with an edge if there is a cost or constraint on the corresponding bi-marginal. The marginals $P_t(\mathbf{M})$ and $P_{t+1}(\mathbf{M})$ are connected due to the cost (6), and the marginals $P_0(\mathbf{M})$ and $P_\mathcal{T}(\mathbf{M})$ are connected due to the constraint (10b) which creates a cyclic structure.

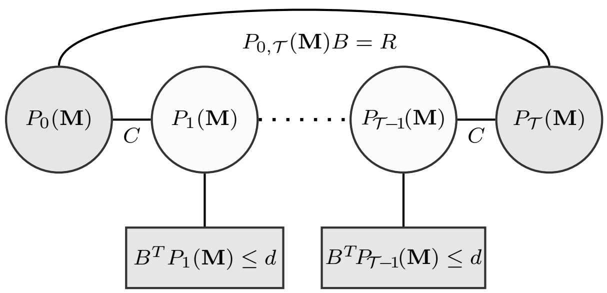

Fig. 2. Origin-destination problem with charge constraints. The grey nodes indicate that there is a constraint in the corresponding marginal.

IV. Main results

In this section we illustrate the computational aspects involved in solving problem (10). Proceeding as for the solution of (4), we start by finding the dual problem.

Theorem 1: Consider the optimization problem (10) and assume that $R \in \mathbb{R}_+^{\tilde{n}\times(|\mathcal{E}|+|\mathcal{C}|)}$ and $d \in \mathbb{R}_+^{|\mathcal{E}|+|\mathcal{C}|}$ are strictly positive with $\mathbf{1}^T R \mathbf{1} \leq \mathbf{1}^T d$. Then the Lagrangian dual of problem (10) is

$$\max_{\substack{\Lambda\in\mathbb{R}^{\tilde{n}\times(|\mathcal{E}|+|\mathcal{C}|)} \\ \lambda_1,\ldots,\lambda_{\mathcal{T}-1}\in\mathbb{R}_+^{|\mathcal{E}|+|\mathcal{C}|}}} -\epsilon\langle \mathbf{K},\mathbf{U}\rangle - \langle \Lambda, R\rangle - \sum_{t=1}^{\mathcal{T}-1}\langle \lambda_t, d\rangle, \quad (11)$$

where $\mathbf{U} \in \mathbb{R}^{\tilde{n}^{(\mathcal{T}+1)}}$ has the form

$$\mathbf{U}_{i_0,\ldots,i_{\mathcal{T}}} = (UB^T)_{i_0 i_{\mathcal{T}}} \cdot (Bu_1)_{i_1} \cdot \ldots \cdot (Bu_{\mathcal{T}-1})_{i_{\mathcal{T}-1}} \quad (12)$$

with $U := \exp(-\Lambda/\epsilon)$ and $u_t := \exp(-\lambda_t/\epsilon)$, and $\mathbf{K} = \exp(-\mathbf{C}/\epsilon)$. Furthermore, the optimal solution of (10) is given by $\mathbf{M} = \mathbf{K} \odot \mathbf{U}$ where $\mathbf{U}$ is on the form (12).

Proof: The Lagrangian of (10) is

$$\begin{aligned} L(\mathbf{M},\Lambda,\lambda_1,\ldots,\lambda_{\mathcal{T}-1}) &= \langle \mathbf{C},\mathbf{M}\rangle + \epsilon D(\mathbf{M}) + \\ &+ \langle \Lambda, P_{0,\mathcal{T}}(\mathbf{M})B - R\rangle + \sum_{t=1}^{\mathcal{T}-1}\langle \lambda_t, B^T P_t(\mathbf{M}) - d\rangle, \quad (13) \end{aligned}$$

where $\Lambda \in \mathbb{R}^{\tilde{n}\times(|\mathcal{E}|+|\mathcal{C}|)}$ represents the dual variable associated with the constraint $P_{0,\mathcal{T}}(\mathbf{M})B = R$, while $\lambda_t \in \mathbb{R}_+^{(|\mathcal{E}|+|\mathcal{C}|)}$ denotes the dual variables corresponding to $B^T P_t(\mathbf{M}) \leq d$, for $t = 1,\ldots,\mathcal{T}-1$. We minimize the Lagrangian with respect to $\mathbf{M_i}$, and noting that the derivative of the entropy term tends to minus infinity as the argument tends to zero, the optimal solution is always strictly positive and attained by setting the derivative equal to zero

$$\frac{\partial L}{\partial \mathbf{M_i}} = \mathbf{C_i} + \epsilon\log(\mathbf{M_i}) + (\Lambda B^T)_{i_0 i_{\mathcal{T}}} + \sum_{t=1}^{\mathcal{T}-1}(B\lambda_t)_{i_t} = 0,$$

which can be rewritten as

$$\mathbf{M_i} = \exp\left(\frac{-\mathbf{C_i} - (\Lambda B^T)_{i_0 i_{\mathcal{T}}}}{\epsilon} - \sum_{t=1}^{\mathcal{T}-1}\frac{(B\lambda_t)_{i_t}}{\epsilon}\right).$$

We observe that, since B is binary with one nonzero element in each row, it holds that $\exp\big(AB^T\big) = \exp(A)B^T$ and $\exp(BA) = B\exp(A)$ for any matrix A of appropriate dimension. Therefore

$$\mathbf{M_i} = \exp\left(-\frac{\mathbf{C_i}}{\epsilon}\right)(UB^T)_{i_0 i_{\mathcal{T}}}\prod_{t=1}^{\mathcal{T}-1}(Bu_t)_{i_t}, \quad (14)$$

which in matrix form corresponds to $\mathbf{M} = \mathbf{K} \odot \mathbf{U}$, with $\mathbf{K} = \exp(-\mathbf{C}/\epsilon)$ and $\mathbf{U}$ defined as in (12). Then by plugging expression (14) into the Lagrangian (13) we obtain (neglecting constant terms)

$$\begin{aligned} &L(\mathbf{K}\odot\mathbf{U},\Lambda,\lambda_1,\ldots,\lambda_{\mathcal{T}-1}) \\ &\qquad = -\epsilon\langle \mathbf{K},\mathbf{U}\rangle - \langle \Lambda, R\rangle - \sum_{t=1}^{\mathcal{T}-1}\langle \lambda_t, d\rangle. \end{aligned} \quad (15)$$

Since (15) is the minimum of the Lagrangian for given dual variables $\Lambda \in \mathbb{R}^{\tilde{n}\times(|\mathcal{E}|+|\mathcal{C}|)}$ and $\lambda_t \in \mathbb{R}_+^{|\mathcal{E}|+|\mathcal{C}|}$ for $t = 1,\ldots,\mathcal{T}-1$, this is also the objective function of the dual, and thus (11) is the dual problem. Finally, since R and d are positive with $\mathbf{1}^T R\mathbf{1} \leq \mathbf{1}^T d$, there exists a positive feasible solution, and by Slater's condition strong duality holds. By strict convexity of the problem a unique solution exists on the form $\mathbf{M} = \mathbf{K} \odot \mathbf{U}$. ∎

We proceed to describe the numerical solution of the dual problem (11) through a dual coordinate ascent algorithm. The method involves iteratively optimizing the dual objective function with respect to one of the dual variables at a time, while holding the remaining variables constant.

Theorem 2: Let the same assumptions of Theorem 1 hold. Then dual coordinate ascent is to iteratively update the components of $\mathbf{U}$ according to

$$U \leftarrow (U \odot R) \oslash (P_{0,\mathcal{T}}(\mathbf{K}\odot\mathbf{U})B), \quad (16a)$$

$$u_t \leftarrow \min(u_t \odot d \oslash (B^T P_t(\mathbf{K}\odot\mathbf{U})), \mathbf{1}) \quad \text{for } t = 1,\ldots,\mathcal{T}-1, \quad (16b)$$

where the minimum in (16b) is taken entry-wise. This converges to a limit point $\Lambda = -\epsilon\log(U)$ and $\lambda_t = -\epsilon\log(u_t)$ for $t = 1,\ldots,\mathcal{T}-1$ which is optimal for problem (11).

Proof: First note that, as shown in Theorem 1, strong duality holds and there exists an optimal solution to the dual problem. In dual coordinate ascent, the dual objective function is maximized with respect to one dual variable at a time while keeping the others fixed:

$$\Lambda \leftarrow \operatorname*{arg\,max}_{\Lambda\in\mathbb{R}^{\tilde{n}\times(|\mathcal{E}|+|\mathcal{C}|)}} -\epsilon\langle \mathbf{K},\mathbf{U}\rangle - \langle \Lambda, R\rangle, \quad (17a)$$

$$\lambda_t \leftarrow \operatorname*{arg\,max}_{\lambda_t\in\mathbb{R}_+^{|\mathcal{E}|+|\mathcal{C}|}} -\epsilon\langle \mathbf{K},\mathbf{U}\rangle - \langle \lambda_t, d\rangle \quad t = 1,\ldots,\mathcal{T}-1. \quad (17b)$$

The iterations are guaranteed to converge by [31]. First, we show that (17a) is equivalent to (16a). The optimization problems (17) are unconstrained and the objective functions are strictly concave with compact superlevel sets due to the positivity of K, R, d, and thus a necessary and sufficient condition for optimality is that the gradient vanishes. Recalling that $U = \exp(-\Lambda/\epsilon)$, we set the derivative of the objective function in (17a) with respect to Λ_{kl} equal to zero, obtaining

$$U_{k\ell}\sum_{i_\mathcal{T}} B^T_{\ell i_\mathcal{T}} \sum_{i_1,\dots,i_{\mathcal{T}-1}} \mathbf{K}_{k,\dots,i_\mathcal{T}} \prod_{t=1}^{\mathcal{T}-1}(Bu_t)_{i_t} = R_{k\ell}.$$

By observing that

$$(P_{0,\mathcal{T}}(\mathbf{M}))_{i_0 i_\mathcal{T}} = \sum_{i_1,\dots,i_{\mathcal{T}-1}} \mathbf{K}_{i_0,\dots,i_\mathcal{T}} (UB^T)_{i_0 i_\mathcal{T}} \prod_{t=1}^{\mathcal{T}-1}(Bu_t)_{i_t}$$

we get

$$\sum_{i_\mathcal{T}} B^T_{\ell i_\mathcal{T}} \sum_{i_1,\dots,i_{\mathcal{T}-1}} \mathbf{K}_{k,\dots,i_\mathcal{T}} \prod_{t=1}^{\mathcal{T}-1}(Bu_t)_{i_t} = \sum_{i_\mathcal{T}} \frac{(P_{0,\mathcal{T}}(\mathbf{K}\odot\mathbf{U})_{k i_\mathcal{T}}}{(UB^T)_{k i_\mathcal{T}}} B_{i_\mathcal{T}\ell}.$$

Using Lemma 1 in Appendix A with $V = P_{0,\mathcal{T}}(\mathbf{K}\odot\mathbf{U})$ we obtain

$$U_{k\ell} \leftarrow \frac{R_{k\ell}}{(P_{0\mathcal{T}}(\mathbf{K}\odot\mathbf{U})B)_{k\ell}/U_{k\ell}},$$

which in matrix form is (16a). Consider now the objective function in (17b), and take the derivative with respect to $(\lambda_t)_k$. The minimizer is obtained either where the gradient vanishes, if it does so for a positive value, or in 0 otherwise, because of the nonnegativity constraint on λ_t. By recalling that $u_t = \exp(-\lambda_t/\epsilon)$ we get

$$d_k = (u_t)_k \sum_{i_t} B_{i_t k} \sum_{\substack{i_0,\dots,i_{t-1}\\ i_{t+1},\dots,i_\mathcal{T}}} \mathbf{K}_\mathbf{i} (UB^T)_{i_0 i_\mathcal{T}} \prod_{\substack{1\le s\le \mathcal{T}-1\\ s\neq t}} (Bu_s)_{i_s}.$$

Proceeding as before:

$$\sum_{i_t} B_{i_t k} \sum_{\substack{i_0,\dots,i_{t-1}\\ i_{t+1},\dots,i_\mathcal{T}}} \mathbf{K}_\mathbf{i} (UB^T)_{i_0 i_\mathcal{T}} \prod_{\substack{1\le s\le \mathcal{T}-1\\ s\neq t}} (Bu_s)_{i_s}$$
$$= \sum_{i_t} B^T_{k i_t} \frac{P_t(\mathbf{K}\odot\mathbf{U})_{i_t}}{(Bu_t)_{i_t}} = \left(B^T(P_t(\mathbf{K}\odot\mathbf{U}) \oslash (Bu_t))\right)_k.$$

By transposing Lemma 1 we obtain $B^T(P_t(\mathbf{M})\oslash(Bu_t)) = (B^T P_t(\mathbf{M}))\oslash u_t$, which gives us, if the positivity constraint is respected,

$$(u_t)_k \leftarrow \frac{d_k (u_t)_k}{(B^T P_t(\mathbf{K}\odot\mathbf{U}))_k}.$$

The update (16b) is then the minimum between the expression above and $1 = \exp(0)$, resulting from $\lambda_t \in \mathbb{R}_+^{|\mathcal{E}|+|\mathcal{C}|}$. ■

The results can be extended to the case of a sparse network, allowing us to incorporate the network topology directly in the cost matrix, by setting $C_{ij} = +\infty$ whenever $\tilde{e}_i$ is not adjacent to $\tilde{e}_j$, resulting in infinite cost for infeasible paths. Indeed, under certain regularity conditions, as stated in Theorem 3 below, the previous results still hold. The theorem can be proved by removing the variables that are trivially constrained, with analogous arguments to those used in the proofs of Theorem 4 and Proposition 1 in [21].

Theorem 3: Let $\mathbf{C} \in \overline{\mathbb{R}}_+^{n^{\mathcal{T}+1}}$, $R \in \mathbb{R}_+^{\tilde{n}\times(|\mathcal{E}|+|\mathcal{C}|)}$ and $d \in \mathbb{R}_+^{|\mathcal{E}|+|\mathcal{C}|}$, where d is strictly positive. Assume that there is a feasible solution $\mathbf{M}$ of (10) for which $\mathbf{M}_{i_0\dots i_\mathcal{T}} > 0$ for all indices $i_0\dots i_\mathcal{T}$ where $\mathbf{C}_{i_0\dots i_\mathcal{T}} < \infty$ and $(RB^T)_{i_0 i_\mathcal{T}} > 0$. Then the optimal solution to (10) has the structure $\mathbf{M} = \mathbf{K}\odot\mathbf{U}$ where $\mathbf{K} = \exp(-\mathbf{C}/\epsilon)$ and $\mathbf{U}$ factorizes as in (12). Moreover, the iterative scheme (16) converges, and the limit point corresponds to a tensor $\mathbf{U}$ for which $\mathbf{M} = \mathbf{K}\odot\mathbf{U}$ is the optimal solution of (10).

The computationally expensive part of the iterations (16) is the evaluation of the projections $P_{0,\mathcal{T}}(\mathbf{K}\odot\mathbf{U})$ and $P_t(\mathbf{K}\odot\mathbf{U})$. Here we need to exploit the sequential decoupling of the cost tensor (6): after defining $K := \exp(-C/\epsilon)$, we obtain

$$\mathbf{K}_{i_0,\dots,i_\mathcal{T}} = \prod_{t=1}^{\mathcal{T}} K_{i_{t-1} i_t}. \tag{18}$$

Furthermore, the entries of $\mathbf{K} = \exp(-\mathbf{C}/\epsilon)$ corresponding to infeasible paths become zero, leading to a sparse tensor. The projections can then be computed as described by the following result.

Theorem 4: Consider $\mathbf{K} \in \mathbb{R}^{\tilde{n}^{\mathcal{T}+1}}$ decoupling as in (18), and $\mathbf{U} \in \mathbb{R}^{\tilde{n}^{\mathcal{T}+1}}$ decoupling as in (12). Then the projections $P_{0,\mathcal{T}}(\mathbf{M})$ and $P_t(\mathbf{M})$ of $\mathbf{M} = \mathbf{K}\odot\mathbf{U}$ can be computed as:

$$P_{0,\mathcal{T}}(\mathbf{M}) = \hat{\phi}_\mathcal{T} \odot (UB^T) = (UB^T)\odot\phi_0, \tag{19a}$$
$$P_t(\mathbf{M}) = Bu_t \odot (\hat{\phi}_t^T \odot (\phi_t (UB^T)^T))\mathbf{1} \tag{19b}$$
$$= Bu_t \odot (\hat{\phi}_t^T UB^T)\odot\phi_t)\mathbf{1}$$

for $t = 1,\dots,\mathcal{T}-1$, in which

$$\hat{\phi}_t = K\mathrm{diag}(Bu_1)K\dots K\mathrm{diag}(Bu_{t-1})K, \tag{20a}$$
$$\phi_t = K\mathrm{diag}(Bu_{t+1})K\dots K\mathrm{diag}(Bu_{\mathcal{T}-1})K. \tag{20b}$$

Proof: See Appendix B ■

We underline how the variables $\phi, \hat{\phi}$ can be computed recursively, allowing for faster computation. Indeed, after setting $\phi_{\mathcal{T}-1} = \hat{\phi}_1 = K$, from (20) it directly follows

$$\hat{\phi}_t = \hat{\phi}_{t-1}\mathrm{diag}(Bu_{t-1})K, \qquad 2\le t\le\mathcal{T}, \tag{21a}$$
$$\phi_t = K\mathrm{diag}(Bu_{t+1})\phi_{t+1}, \qquad 0\le t\le\mathcal{T}-2. \tag{21b}$$

Finally, in Algorithm 1 we recap the procedure for numerically solving (11), using the updates (16) while exploiting the recursions (21) for computing the marginal projections. The algorithm returns the marginal projections $P_t(\mathbf{M})$ for all t as well as the dual variables U and u_t for $t = 1,\dots,\mathcal{T}-1$, which can be used for reconstructing $\mathbf{M} = \mathbf{K}\odot\mathbf{U}$.

The computational complexity of one iteration (updating all the dual variables) in Algorithm 1 is $\mathcal{O}(\mathcal{T}\tilde{n}^3)$, because of the matrix-matrix multiplications in the updates of $\phi, \hat{\phi}$ (21). However, the actual computational complexity is typically considerably lower. First, due to the charge-expanded construction, K has a sparse structure even if the underlying network is complete. Furthermore, if the matrix R has only L non-zero entries (i.e., the problem has only L commodities), only LQ columns of ϕ and only L rows of $\hat{\phi}$ need to be computed. Both these factors significantly reduce the computational cost.

Algorithm 1 Dual coordinate ascent for problem (11)

$U \leftarrow \mathbf{1}\cdot\mathbf{1}^T$, $u_t \leftarrow \mathbf{1}$ for $t = 1, \dots \mathcal{T}-1$
$\hat{\phi}_1 \leftarrow K$, $\phi_{\mathcal{T}-1} \leftarrow K$
for $t = \mathcal{T}-2, \mathcal{T}-3, \dots, 0$ **do**
    $\phi_t \leftarrow K\text{diag}(Bu_{t+1})\phi_{t+1}$
end for
while constraint violation (10b)-(10c) $>$ tolerance **do**
    $P_{0,\mathcal{T}}(\mathbf{K}\odot\mathbf{U}) \leftarrow \phi_0 \odot UB^T$
    $U \leftarrow (U \odot R) \oslash (P_{0,\mathcal{T}}(\mathbf{K}\odot\mathbf{U})B)$
    for $t = 1, \dots, \mathcal{T}-1$ **do**
        $\hat{\phi}_t \leftarrow \hat{\phi}_{t-1}\text{diag}(Bu_{t-1})K$ **if** $t > 1$
        $P_t(\mathbf{K}\odot\mathbf{U}) \leftarrow Bu_t \odot ((\hat{\phi}_t^T UB^T)\odot \phi_t)\mathbf{1}$
        $u_t \leftarrow \min((u_t \odot d) \oslash (B^T P_t(\mathbf{K}\odot\mathbf{U})), \mathbf{1})$
    end for
    for $t = \mathcal{T}-2, \mathcal{T}-3, \dots, 0$ **do**
        $\phi_t \leftarrow K\text{diag}(Bu_{t+1})\phi_{t+1}$
    end for
end while

V. Numerical simulations

In this section we illustrate an application of the methodology described in Section III and Section IV, where we seek to transfer $L = 5$ groups of battery-powered agents from assigned origins to destinations, while respecting capacity and charge constraints. We consider a grid of dimension 30×30, with $Q = 60$ charge levels, and assume that the dynamics allow for agents to move from one square to a neighbouring one (up, down, left, or right) in one time step. Neighbouring squares are connected by two directed edges, allowing capacity constraints to differentiate the flow in opposite directions. Let the capacity be $d_i = 10$ for the roads and $d_i = 50$ for each of three charging stations. In order to allow for nodes to be origins and destinations, we augment the state space with sources and sinks, as described in Remark 2. The set of states is then $\tilde{\mathcal{E}} = \tilde{E} \cup \tilde{\mathcal{C}} \cup \tilde{\mathcal{S}}^+ \cup \tilde{\mathcal{S}}^-$. The details on the initial data are given in Table I, and the setup is illustrated in Figure 3. Note that units can have the same origin and destination but different initial charge, hence they need to be treated as separate commodities.

TABLE I
Data of the example

Origin (row,col)	Destination (row,col)	Initial charge q	# units
$(9, 1)$	$(15, 30)$	15	50
$(9, 1)$	$(15, 30)$	50	100
$(1, 11)$	$(30, 23)$	15	50
$(1, 11)$	$(30, 23)$	35	100
$(26, 1)$	$(6, 30)$	55	200

We assume that the cost is 1 for all roads. In this setting, we allow for agents to freely stay in sources and sinks (without cost and loss of charge), but not in other states. At the same time, it is possible to exit sources and to enter sinks, but not vice-versa. In this case, the cost is 0.5, as sources and sinks do not correspond to physical roads. We also assume that using a charging station costs 1 for each level of charge gained. As we are not imposing a final charge level, in order to favour arrivals with higher battery charge, we add the cost $(Q - q)$ for ending up in a sink with charge q. The cost matrix C is summarized in Table II, where we say that $\tilde{e}_i$ is adjacent to $\tilde{e}_j$ if $\text{tail}(\tilde{e}_i) = \text{head}(\tilde{e}_j)$. In other words, two edges are adjacent if it is possible to directly go from the first one to the second.

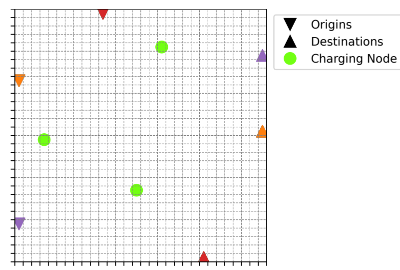

Fig. 3. Origin-destination pairs and charging stations on the grid. Each pair is assigned a different color

TABLE II
The cost matrix C

C	$\tilde{E}$	$\tilde{\mathcal{C}}$	$\tilde{\mathcal{S}}^+$	$\tilde{\mathcal{S}}^-$
$\tilde{E}$	1 if adj. ∞ else	1 if adj. ∞ else	∞	$\frac{1}{2} + (Q-q)$if adj. ∞ else
$\tilde{\mathcal{C}}$	1 if adj. ∞ else	1 if adj. ∞ else	∞	$\frac{1}{2} + (Q-q)$if adj. ∞ else
$\tilde{\mathcal{S}}^+$	$\frac{1}{2}$ if adj. ∞ else	$\frac{1}{2}$ if adj. ∞ else	0 if adj. ∞ else	$(Q-q)$if adj. ∞ else
$\tilde{\mathcal{S}}^-$	∞	∞	∞	0 if adj. ∞ else

We run Algorithm 1 for $\mathcal{T} = 80$, $\epsilon = 0.5$. Despite the large number of states ($|\tilde{\mathcal{E}}| \approx 2 \cdot 10^5$), the matrix K is highly sparse, making the matrix-matrix multiplication computationally feasible. The computation required 1390 iterations for the algorithm to converge within a tolerance of 10^{-3}, with a total runtime of 10 hours and 40 minutes. The simulation was performed on a machine equipped with an Intel Core i7-1165G7 CPU and 16 GB of RAM. The software environment utilized Python 3.9, with the SciPy library used for sparse matrix operations. The results of the simulation are shown in Figure 4. In the first row, we show the flow over time of the entire ensemble. In the following three rows we divide the agents in groups based on their remaining charge. These groups are referred to as *low* with $0 \leq q \leq 19$, *medium* with $20 \leq q \leq 39$, and *high* with $40 \leq q \leq 60$. One can observe how the groups of agents with same origin and destination but different initial charge behave differently. In both cases with agents starting in $(9, 1)$ and $(1, 11)$, the agents with *low* charge are immediately steered towards the nearest charging stations, where they spend several time steps to recharge, while the others continue towards their destination. We can also observe how the capacity constraints force agents to spread across different paths or to wait in their starting point. Also note that as time passes, more vehicles move from *high* to *medium* charge groups and from *medium* to *low*. These are nontrivial behaviour, but they match what one would intuitively expect for an optimal solution.

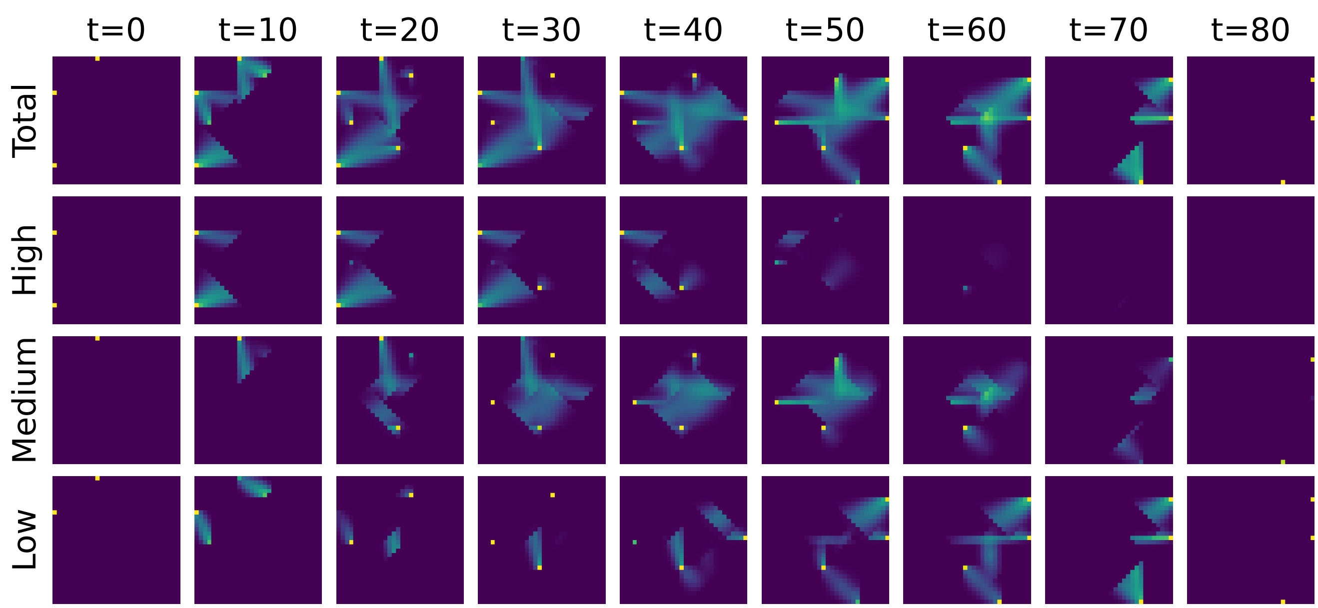

Fig. 4. Optimal traffic flow, agents divided in groups based on their charge levels.

VI. Conclusions

In this paper we introduced a method for the steering an ensemble of electric vehicles over a network with road-capacity and origin-destination constraints. We expanded the network in the charge-dimension, presenting a dual coordinate ascent algorithm for efficiently solving the problem. Finally, we illustrated an application of the algorithm on a grid network.

In the proposed approach very large optimization problems are considered. However, in order to handle realistic scenarios even larger problems need to be solved [34]. Further work therefore include the study of different graph structures to develop even more computationally efficient algorithms, while also including the impact of traffic congestion [34], [35] and limitations imposed by the electrical grid on charging infrastructure.

Appendix A
Lemma 1

Let $B \in \{0,1\}^{n\times m}$ be a binary matrix such that $B\mathbf{1}_m = \mathbf{1}_n$ and let $V \in \mathbb{R}^{p\times n}$, $U \in \mathbb{R}^{p\times m}$ be arbitrary matrices. Then

$$\big(V \oslash (UB^T)\big)B = (VB) \oslash U. \tag{22}$$

Proof: Consider the i,j-th element.

$$\big[\big(V \oslash (UB^T)\big)B\big]_{ij} = \sum_{k=1}^{n} \big(V_{ik} \oslash \big(\sum_{\ell=1}^{m} U_{i\ell}B_{\ell k}^T\big)\big)B_{kj}.$$

The property $B\mathbf{1}_m = \mathbf{1}_n$ implies that each row of B has exactly one nonzero element. Therefore, if $B_{kj} = 1$, we have $\sum_{\ell=1}^{m} U_{i\ell}B_{\ell k}^T = U_{ij}$. This gives us

$$\sum_{k=1}^{n} \Big(V_{ik} \oslash \Big(\sum_{\ell=1}^{m} U_{i\ell}B_{\ell k}^T\Big)\Big)B_{kj} = \sum_{k=1}^{n} \frac{V_{ik}B_{kj}}{U_{ij}} = \left[(VB)\oslash U\right]_{ij}.$$

which is the result of the lemma. As a remark, we notice that $U_{ij} = 0$ would result in infinity on both sides. ■

Appendix B
Proof of Theorem 4

Observe that

$$\begin{aligned}(\hat\phi_t)_{i_0 i_t} &= (K\text{diag}(Bu_1)K \dots K\text{diag}(Bu_{t-1})K)_{i_0 i_t}\\ &= \sum_{i_1,\dots,i_{t-1}} K_{i_0 i_1}(Bu_1)_{i_1}K_{i_1 i_2}\dots(Bu_{t-1})_{i_{t-1}}K_{i_{t-1}i_t}\\ &= \sum_{i_1,\dots,i_{t-1}} \left(\prod_{s=1}^{t} K_{i_{s-1}i_s}\prod_{s=1}^{t-1}(Bu_s)_{i_s}\right),\end{aligned}$$

and analogously

$$(\phi_t)_{i_t i_\mathcal{T}} = \sum_{i_{t+1},\dots,i_{\mathcal{T}-1}} \left(\prod_{s=t+1}^{\mathcal{T}} K_{i_{s-1}i_s}\prod_{s=t+1}^{\mathcal{T}-1}(Bu_s)_{i_s}\right).$$

By considering $(\phi_0)_{i_0 i_\mathcal{T}}$ and $(\hat\phi_\mathcal{T})_{i_0 i_\mathcal{T}}$ we get

$$\begin{aligned}&P_{0,\mathcal{T}}(\mathbf{M})_{i_0 i_\mathcal{T}} =\\ &= \sum_{i_1,\dots,i_{\mathcal{T}-1}} \left(\prod_{s=1}^{\mathcal{T}} K_{i_{s-1}i_s}\right)\left(\prod_{s=1}^{\mathcal{T}-1}(Bu_s)_{i_s}\right)(UB^T)_{i_0 i_\mathcal{T}}\\ &= (UB^T)_{i_0 i_\mathcal{T}}(K\text{diag}(Bu_1)K\dots K\text{diag}(Bu_{\mathcal{T}-1})K)_{i_0 i_\mathcal{T}}\\ &= (UB^T)_{i_0 i_\mathcal{T}}(\phi_0)_{i_0 i_\mathcal{T}} = (UB^T)_{i_0 i_\mathcal{T}}(\hat\phi_\mathcal{T})_{i_0 i_\mathcal{T}},\end{aligned}$$

which is (19a). Now consider the projections

$$\begin{aligned}&P_{0,t}(\mathbf{K}\odot\mathbf{U})_{i_0 i_t} =\\ &= \sum_{\substack{i_1,\dots,i_{t-1}\\ i_{t+1},\dots,i_\mathcal{T}}} \left(\prod_{s=1}^{\mathcal{T}} K_{i_{s-1}i_s}\right)\left(\prod_{s=1}^{\mathcal{T}-1}(Bu_s)_{i_s}\right)(UB^T)_{i_0 i_\mathcal{T}}.\end{aligned}$$

The product in the right hand side can be split up into two components, each accounting for the terms with indices smaller and larger than t, respectively. Then we identify (the sum of) each such term with ϕ_t and $\hat{\phi}_t$

$$
\begin{aligned}
&P_{0,t}(\mathbf{K}\odot\mathbf{U})_{i_0 i_t} = \\
&= (Bu_t)_{i_t} \sum_{i_1,\ldots,i_{t-1}} \left(\prod_{s=1}^{t} K_{i_{s-1}i_s} \prod_{s=1}^{t-1} (Bu_s)_{i_s} \right) \cdot \\
&\cdot \sum_{i_\mathcal{T}} \sum_{i_{t+1},\ldots,i_{\mathcal{T}-1}} \left(\prod_{s=t+1}^{\mathcal{T}} K_{i_{s-1}i_s} \prod_{s=t+1}^{\mathcal{T}-1} (Bu_s)_{i_s} \right) (UB^T)_{i_0 i_\mathcal{T}} \\
&= (Bu_t)_{i_t} (\hat{\phi}_t)_{i_0 i_t} \sum_{i_\mathcal{T}} (\phi_t)_{i_t i_\mathcal{T}} (UB^T)_{i_0 i_\mathcal{T}} \\
&= (Bu_t)_{i_t} (\hat{\phi}_t)_{i_0 i_t} ((UB^T)\phi_t^T)_{i_0 i_t}.
\end{aligned}
$$

Proceeding in the same way for $P_{t,\mathcal{T}}(\mathbf{K}\odot\mathbf{U})$ we obtain

$$P_{0,t}(\mathbf{K}\odot\mathbf{U}) = \big(\hat{\phi}_t \odot (UB^T\phi_t^T)\big)\mathrm{diag}(Bu_t), \tag{23a}$$

$$P_{t,\mathcal{T}}(\mathbf{K}\odot\mathbf{U}) = \mathrm{diag}(Bu_t)\big((\hat{\phi}_t^T UB^T) \odot \phi_t\big). \tag{23b}$$

Finally, observe that

$$P_t(\mathbf{K}\odot\mathbf{U}) = P_{0,t}(\mathbf{K}\odot\mathbf{U})^T\mathbf{1} = P_{t,\mathcal{T}}(\mathbf{K}\odot\mathbf{U})\mathbf{1}, \tag{24}$$

which leads to (19b). ■

References

[1] J. A. Sanguesa, V. Torres-Sanz, P. Garrido, F. J. Martinez, and J. M. Marquez-Barja, "A review on electric vehicles: Technologies and challenges," *Smart Cities*, vol. 4, no. 1, pp. 372–404, 2021.

[2] Í. R. da Costa Barros and T. P. Nascimento, "Robotic mobile fulfillment systems: A survey on recent developments and research opportunities," *Robotics and Autonomous Systems*, vol. 137, p. 103729, 2021.

[3] R. R. Murphy, J. Kravitz, S. L. Stover, and R. Shoureshi, "Mobile robots in mine rescue and recovery," *IEEE Robotics & Automation Magazine*, vol. 16, no. 2, pp. 91–103, 2009.

[4] R. R Shamshiri, C. Weltzien, I. A. Hameed, I. J Yule, T. E Grift, S. K. Balasundram, L. Pitonakova, D. Ahmad, and G. Chowdhary, "Research and development in agricultural robotics: A perspective of digital farming," *International Journal of Agricultural and Biological Engineering*, 2018.

[5] X. Bi and W. Tang, "Logistical planning for electric vehicles under time-dependent stochastic traffic," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 10, pp. 3771–3781, 2018.

[6] H. Wang, Q. Meng, J. Wang, and D. Zhao, "An electric-vehicle corridor model in a dense city with applications to charging location and traffic management," *Transportation Research Part B: Methodological*, vol. 149, pp. 79–99, 2021.

[7] J. Shi, T. Zeng, and S. Moura, "Electric fleet charging management considering battery degradation and nonlinear charging profile," *Energy*, vol. 283, p. 129094, 2023.

[8] J. Hu, F. Xiao, B. Mei, Z. Lin, and C. Fu, "Optimal energy efficient control of pure electric vehicle power system based on dynamic traffic information flow," *IEEE Transactions on Transportation Electrification*, vol. 8, no. 1, pp. 510–526, 2021.

[9] A. Le Rhun, F. Bonnans, G. De Nunzio, T. Leroy, and P. Martinon, "A bilevel energy management strategy for HEVs under probabilistic traffic conditions," *IEEE Transactions on Control Systems Technology*, vol. 30, no. 2, pp. 728–739, 2021.

[10] X. Shi, Y. Xu, Q. Guo, H. Sun, and W. Gu, "A distributed EV navigation strategy considering the interaction between power system and traffic network," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3545–3557, 2020.

[11] M. Neuman, H. Sandberg, B. Wahlberg, and A. Folkesson, "Modelling and control of series HEVs including resistive losses and varying engine efficiency," in *SAE World Congress 2009, Detroit, Michigan United States*. SAE International, 2009.

[12] N. Murgovski, L. Johannesson, X. Hu, B. Egardt, and J. Sjöberg, "Convex relaxations in the optimal control of electrified vehicles," in *2015 American control conference (ACC)*. IEEE, 2015, pp. 2292–2298.

[13] A. Montoya, C. Guéret, J. E. Mendoza, and J. G. Villegas, "The electric vehicle routing problem with nonlinear charging function," *Transportation Research Part B: Methodological*, vol. 103, pp. 87–110, 2017.

[14] L. Ford and D. Fulkerson, "Constructing maximal dynamic flows from static flows," *Operations research*, vol. 6, no. 3, pp. 419–433, 1958.

[15] C. Pasquale, S. Sacone, S. Siri, and A. Ferrara, "Traffic control for freeway networks with sustainability-related objectives: Review and future challenges," *Annual Reviews in Control*, 2019.

[16] Y. Chen, T. T. Georgiou, and M. Pavon, "Optimal transport in systems and control," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 4, pp. 89–113, 2021.

[17] A. Ringh, I. Haasler, Y. Chen, and J. Karlsson, "Mean field type control with species dependent dynamics via structured tensor optimization," *IEEE Control Systems Letters*, 2023.

[18] A. Terpin, N. Lanzetti, and F. Dörfler, "Dynamic programming in probability spaces via optimal transport," *SIAM Journal on Control and Optimization*, vol. 62, no. 2, pp. 1183–1206, 2024.

[19] M. Mascherpa, I. Haasler, B. Ahlgren, and J. Karlsson, "Estimating pollution spread in water networks as a Schrödinger bridge problem with partial information," *European Journal of Control*, p. 100846, 2023.

[20] I. Haasler, Y. Chen, and J. Karlsson, "Optimal steering of ensembles with origin-destination constraints," *IEEE Control Systems Letters*, vol. 5, no. 3, pp. 881–886, 2020.

[21] I. Haasler, A. Ringh, Y. Chen, and J. Karlsson, "Scalable computation of dynamic flow problems via multimarginal graph-structured optimal transport," *Mathematics of Operations Research*, 2023.

[22] C. Villani, *Optimal transport: old and new*. Springer, 2009, vol. 338.

[23] G. Carlier, "On a class of multidimensional optimal transportation problems," *Journal of convex analysis*, vol. 10, no. 2, pp. 517–530, 2003.

[24] B. Pass, "Multi-marginal optimal transport: theory and applications," *ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique*, vol. 49, no. 6, pp. 1771–1790, 2015.

[25] M. Cuturi, "Sinkhorn distances: Lightspeed computation of optimal transport," *Advances in neural information processing systems*, vol. 26, pp. 2292–2300, 2013.

[26] J.-D. Benamou, G. Carlier, M. Cuturi, L. Nenna, and G. Peyré, "Iterative Bregman projections for regularized transportation problems," *SIAM Journal on Scientific Computing*, vol. 37, no. 2, pp. A1111–A1138, 2015.

[27] J. M. Altschuler and E. Boix-Adsera, "Polynomial-time algorithms for multimarginal optimal transport problems with structure," *Mathematical Programming*, vol. 199, no. 1, pp. 1107–1178, 2023.

[28] I. Haasler, A. Ringh, Y. Chen, and J. Karlsson, "Multimarginal optimal transport with a tree-structured cost and the Schrödinger bridge problem," *SIAM Journal on Control and Optimization*, vol. 59, no. 4, pp. 2428–2453, 2021.

[29] L. Engström, S. Källblad, and J. Karlsson, "Computation of robust option prices via structured multi-marginal martingale optimal transport," *arXiv preprint arXiv:2406.09959*, 2024.

[30] A. Ringh, I. Haasler, Y. Chen, and J. Karlsson, "Graph-structured tensor optimization for nonlinear density control and mean field games," *SIAM Journal on Control and Optimization*, vol. 62, no. 4, pp. 2176–2202, 2024.

[31] P. Tseng, "Dual ascent methods for problems with strictly convex costs and linear constraints: A unified approach," *SIAM Journal on Control and Optimization*, vol. 28, no. 1, pp. 214–242, 1990.

[32] M. Skutella, "An introduction to network flows over time," *Research Trends in Combinatorial Optimization: Bonn 2008*, pp. 451–482, 2009.

[33] L. G. Willumsen, "Estimation of an od matrix from traffic counts–a review." *Working Paper 99*, 1978.

[34] A. Ferrara, S. Sacone, S. Siri *et al.*, *Freeway traffic modelling and control*. Springer, 2018, vol. 585.

[35] G. Como, "On resilient control of dynamical flow networks," *Annual Reviews in Control*, vol. 43, pp. 80–90, 2017.