A Game Theoretic Model for Aggregate Lane Change Behavior of Vehicles at Traffic Diverges

Negar Mehr, Ruolin Li, and Roberto Horowitz[1]

Abstract— **Lane changes are known to negatively affect traffic delays. However, due to the complexities of this phenomena, accurate and yet simple models of lane change maneuvers are hard to develop. In this work, we present a macroscopic model for predicting the number of vehicles that change lanes at a traffic diverge. We take into account the selfishness of vehicles in selecting their lanes; every vehicle selects lanes such that its own cost is minimized. We discuss how we model such costs experienced by the vehicles. Then, taking into account the selfish behavior of the vehicles, we model lane choice of vehicles at a traffic diverge as a Wardrop equilibrium. We state and prove the properties of our Wardrop equilibrium. We show that there always exists an equilibrium for our model. Moreover, we prove that the equilibrium is unique under mild assumptions. We discuss how our model can be easily calibrated by running a simple optimization problem. Using our calibrated model, we validate our model through simulation studies and demonstrate that our model successfully predicts the aggregate lane change maneuvers at a diverge. We further discuss how our model can be employed to obtain the optimal lane choice behavior of the vehicles, where the social or total cost of the vehicles is minimized.**

I. Introduction

Due to the huge delays and costs that are incurred by traffic congestion, the task of modeling traffic behavior is of paramount importance since such models can be used to analyze traffic networks to get an insight about how the traffic conditions can be improved in urban and freeway networks. However, as traffic networks normally exhibit very complex behaviors, developing models that are both accurate and simple enough for analysis and control purposes is non-trivial. Among the different phenomena that needs to be modeled in traffic networks, lane change behavior of the vehicles has been one of the most challenging behaviors to model. This is due to the fact lane change maneuvers are very complex by nature, and their negative effects on traffic streams are hard to quantify. The existing literature on modeling lane change behaviors is mainly divided into two categories: 1) modeling the microscopic lane change decisions of vehicles, and 2) investigating and quantifying the macro effect of aggregate lane change maneuvers on the traffic conditions.

Related to the first category of the existing research work, the lane change behavior of vehicles was first systematically studied in [7]. In [7], a set of rule and conditions were developed under which a single vehicle was assumed to change lane. The set of derived conditions were assumed to depend on microscopic parameters of the vehicle and road segment such as vehicle velocity and available distance in the neighboring links. Aligned with this, several car following models were proposed by researchers that model the micro behavior of vehicles such as acceleration and lane change maneuvers. Examples of such works can be found in [19], [2], [10]. Due to complexities of such models, researchers studied lane change behavior of vehicles using car following models in simulation studies [8]. Recently, a game theoretic approach was used in [20] and [17] to model the lane change behavior of a single vehicle, where the lane change decision was assumed to be taken by a vehicle for increasing its speed. In [12], similar approach was taken to mimic the behavior of the drivers at traffic merges.

With the recent advances in autonomous vehicles technology, a large body of literature has been devoted on how to design and control autonomous vehicles such that they manifest lane change behaviors that are both safe and similar to the lane change decisions made by humans. In [6], the intention of the drivers when merging to freeway lanes was estimated. In [21], a decision making approach for performing lane changes while driving fully automated in urban environments was presented and evaluated. In [11], the requirements associated with an optimal lane change behavior were described where minimizing fuel consumption and travel time were considered as objectives. In [16], computer vision techniques were utilized to infer lane change intents.

In the second category of the present research work investigating macro effects, there has been a focus on how to quantify the negative effects of lane change maneuvers on upstream traffic congestion. In [15], lane changing vehicles were modeled as particles endowed with mechanical properties. In this work, freeway sections that are away from diverges were considered, and the freeway was modeled as interacting streams which could be linked by these particles. The implications and applications of this model were discussed in [14]. In [4], it was demonstrated via case studies that lane change maneuvers could lead to reductions in freeway capacity. In [9], the impacts of lane change behaviors was modeled via introduction of lane–changing intensity variables and modified fundamental diagrams. In [23], a stochastic lane change model was developed for capturing the system–level lane changing characteristics.

In this paper, we study the *aggregate* lane change maneuvers taken by the vehicles at traffic diverges. Despite the majority of the existing literature on lane change modeling,

[1]Negar Mehr, Ruolin Li, and Roberto Horowitz are with the Mechanical Engineering Department, University of California, Berkeley, CA, USA `negar.mehr@berkeley.edu`, `ruolin_li@berkeley.edu`, `horowitz@berkeley.edu`

we study it at the *macro* scale, where we predict the number of vehicles who will change their lane in order to take an appropriate exit that corresponds to their route. In particular, given the number of vehicles who wish to take a certain exit, we develop a novel model which can predict how many vehicles will change their lanes close to the diverge in order to take an exit. We assume that vehicles act selfishly, i.e. every vehicle decides on its route and lane choice such that its own cost is minimized. We describe how the costs incurred on the vehicles can be modeled. Since our focus is on developing a macroscopic fluid–like model for the behavior of the vehicles, we model the equilibrium that results from selfishness of the vehicles as a Wardropian equilibrium. We prove that our model always has an equilibrium, and further, its equilibrium is unique under mild assumptions. We describe how our model can easily be calibrated by solving a mixed-integer linear program, and show through simulation studies that our model shows promising results, it can successfully predict the aggregate lane change behavior of the vehicles at a fork.

Our framework, albeit simple, provides a powerful tool for quantifying the inefficiencies that arise from the selfish lane change behavior of vehicles at traffic diverges. Moreover, our model can be used to determine the optimal aggregate lane change maneuvers of vehicles. It has been shown via multiple simulation studies that lane change can reduce the efficiency of the road, but our framework not only predicts the lane change behavior of the vehicles but it can also be used to quantitatively study and analyze this effect. Our model is particularly beneficial in scenarios when a central authority can route a fraction of vehicles. For instance, in networks with mixed autonomy, autonomous vehicles might be routed by a central planner. In such scenarios, our model can be used for deciding on the optimal lane change behavior of vehicles where the resulting equilibrium has the minimum social cost. To the best of our knowledge, there is no such work in the literature.

The organization of this paper is as follows. In Section II, we describe our modeling framework. In Section III, we state and prove the properties of our model. Simulation studies including model calibration and validation are described in Section IV. In Section V, we describe how our model can be used for choosing the optimal lane change pattern at traffic diverges. We conclude the paper and discuss our future directions in Section VI.

II. THE MODEL

We consider a traffic diverge where a link is bifurcated into two links, which is a common scenario for freeway and arterial forks. We wish to study the route choice behavior of vehicles in such diverges, where certain lanes correspond to a certain route or exit link. Normally, in these scenarios, among the vehicles with the same target exit link, a fraction of vehicles choose the appropriate lanes that correspond to their exit, far upstream of the diverge, while the remaining fraction of the vehicles choose their lane and route very close to the diverge. We wish to obtain a model that given

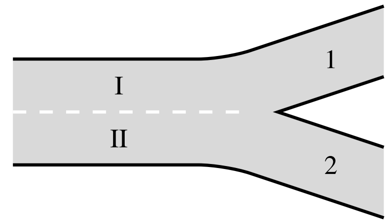


Fig. 1: Example of a traffic diverge with two destination links.

the demand of vehicles for each possible exit link it can predict the fraction of vehicles which perform either of these two behaviors in their route choice at the diverge. Note that we wish to capture the aggregate lane choice behavior of vehicles in a macro scale rather than deriving the conditions under which a single vehicle decides to change lane.

Let I be the index set of exit links at a diverge. Consider a traffic diverge with two exit links $I = \{1, 2\}$. Let d_1 and d_2 be the demand of vehicles who wish to take the link 1 and 2 respectively. We use $f_1 = \frac{d_1}{d_1+d_2}$ and $f_2 = \frac{d_2}{d_1+d_2}$ to represent the fraction of vehicles whose destinations are links 1 and 2. We describe our model in terms of these normalized flows rather than the actual flows since it simplifies our analysis. Let $F = \{f_1, f_2\}$ be the flow configuration which is the set of normalized demands for the diverge destinations. For each link $i \in I$, let x_i^s denote the fraction of "steadfast" vehicles among f_i, which are the vehicles that take the lanes that correspond to their destination link i far upstream of the diverge, whilst x_i^a denotes the fraction of "altering" vehicles who take the lanes that correspond to their exit link i at or at the vicinity of the diverge. We assume that vehicles change lane only once, i.e. if a vehicle is in its target lane, it remains there. We let $\mathbf{x} = (x_i^s, x_i^a : i \in I)$ be the vector of such flows for the two possible destinations of a fork. A flow vector $\mathbf{x}$ is feasible if satisfies:

$$f_i = x_i^s + x_i^a, \quad \forall i \in I, \tag{1}$$
$$\textstyle\sum_{i \in I} f_i = 1, \tag{2}$$
$$f_i \leq 1, \forall i \in I, \tag{3}$$
$$x_i^s \geq 0,\ x_i^a \geq 0, \quad \forall i \in I. \tag{4}$$

Example 1. Consider the diverge shown in Figure 1. In this example, there are two freeway lanes I and II which bifurcate to exit links 1 and 2. For this diverge, x_1^s is the fraction of vehicles that remain on lane I and take the exit link 1, whereas x_1^a is the fraction of vehicles that move along lane II and change their lane from II to I at vicinity of the diverge.

For each $i \in I$, we assume that all steadfast vehicles constituting x_i^s experience the same cost. Likewise, all altering vehicles taking an exit link i, experience the same cost. For each destination link $i \in I$, we let J_i^s and J_i^a be the cost incurred on the vehicles forming x_i^s and x_i^a respectively. It is important to note that for each $i \neq j \in I$, J_i^s or J_i^a depends not only on x_i^s and x_i^a but it can also depend on x_j^s and x_j^a. For each $i \neq j \in I$, we model the cost of the steadfast vehicles by

$$J_i^s(\mathbf{x}) = C_i^t \left(x_i^s + x_j^a\right) + C_i^c x_i^a \left(x_i^s + x_j^a\right), \tag{5}$$

where C_i^t and C_i^c are positive constants. The constant C_i^t is the cost of traversing the lanes that correspond to the exit i. Since $(x_i^s + x_j^a)$ is the total fraction of vehicles that traverse the link that leads to exit i, $(x_i^s + x_j^a)$ is multiplied by C_i^t (e.g. $x_1^s + x_2^a$ traverses link I in Figure 1 and the cost is C_1^t). It indicates that the more occupied the lanes that correspond to an exit are, the more expensive their traversal is. On the other hand, the constant C_i^c is used to reflect the negative cross effects caused by the lane change behavior of altering vehicles x_i^a. This term is used to mimic the fact that as the vehicles in x_i^a change their lanes to take the exit i, they use the roads (resources) that join the exit i; thus, they will create delays for the vehicles that are already in their target lanes. Note that since x_i^s and x_j^a both share the target link of x_i^a up to the vicinity of the diverge, the total fraction of the vehicles present in the target lanes of x_i^a is $(x_i^s + x_j^a)$. Hence, C_i^c is multiplied by $(x_i^s + x_j^a)$ and x_i^a. This multiplication implies that the higher the number of vehicles that change lanes (x_i^a) is, or, the more occupied the lanes that joins the exist i is, the larger the incurred cost is.

Now, we describe how we model the costs that the altering vehicles experience. For each $i \neq j \in I$, we model J_i^a via

$$J_i^a(\mathbf{x}) = C_j^t \left(x_j^s + \gamma_i x_i^a\right) + C_j^c x_j^a \left(x_j^s + x_i^a\right) \tag{6}$$

where γ_i is a constant assumed to satisfy $\gamma_i \geq 1$, and C_j^t and C_j^c are as previously defined. If $\gamma_i = 1$, the first and second terms in (6) are the costs that are incurred due to traversing the links that join the exit j defined by (5). But, if $\gamma_i > 1$, the additional cost that the altering vehicles must pay due to traversing a longer path for joining their appropriate exit, is modeled. In fact, $\gamma_i > 1$ can model the cost incurred on altering vehicles due to the additional distance they need to traverse as well as the discomfort cost they will face for changing lanes.

Example 2. Consider the diverge shown in Figure 1. For this example, x_1^s is the fraction of the vehicles that remain on lane I and take exit 1, whereas x_2^a is the fraction of the vehicles that use lane I and leave lane I close to the diverge to take the exit 2. In this case, $J_1^s = C_1^t\left(x_1^s + x_2^a\right) + C_1^c x_1^a \left(x_1^s + x_2^a\right)$. Note that $C_1^t\left(x_1^s + x_2^a\right)$ is the cost of traversing lane I, where $(x_1^s + x_2^a)$ is the total fraction of vehicles present on lane I. Now, consider J_1^a which is the cost of the vehicles who aim to take the exit 1 but via lane change maneuvers from II to I. These vehicles remain on lane II until close to the diverge and then change their lane to take the exit I. For this type of vehicles, $J_1^a = C_2^t\left(x_2^s + \gamma_1 x_1^a\right) + C_2^c x_2^a \left(x_2^s + x_1^a\right)$. In this case, $C_2^t\left(x_2^s + x_1^a\right) + C_2^c x_2^a \left(x_2^s + x_1^a\right)$ is the cost of traversing lane II (up to vicinity of the diverge); whilst, the additional $C_2^t((\gamma_1 - 1)x_1^a)$ is due to traversing the extra distance required for leaving lane II and finally joining exit I as well as the discomfort cost the vehicles have to pay for changing their lane.

We let $\mathbf{C} = (C_i^t, C_i^c, \gamma_i : i \in I)$ be the vector of the cost coefficients in our model. Before proceeding, we need the following definition.

Definition 1. A function $h(.) : \mathbb{R}^n \longrightarrow \mathbb{R}$ is called monotone if and only if for every $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^n$ such that $\mathbf{x} \leq \mathbf{x}'$, where inequalities are interpreted elementwise, we have

$$h(\mathbf{x}) \leq h(\mathbf{x}').$$

Using Equations (5) and (6), the following remark is evident.

Remark 1. For each $i \in I$, the cost functions J_i^s and J_i^a, are monotone in the sense of Definition 1.

We will later use Remark 1 to guarantee certain properties of our model.

A reasonable and realistic assumption is that vehicles act *selfishly*, i.e. each vehicle acts such that its own cost is minimized. Each vehicle has two options: to choose its appropriate lane upstream of the diverge or to take its target exit close to the diverge via lane change. Therefore, at equilibrium, if for $i \in I$, both x_i^s and x_i^a are nonzero, we must have $J_i^s(\mathbf{x}) = J_i^a(\mathbf{x})$; otherwise, vehicles will move to the lanes with lower cost. If either x_i^s or x_i^a is zero, then, its corresponding cost must be already larger than the one with nonzero flow. These conditions are called Wardrop conditions [22] in the transportation literature. In order to describe the formal definition of Wardrop conditions, let $G = (F, \mathbf{C})$ be a tuple enclosing F and $\mathbf{C}$ which are respectively the configuration of demand fractions and the vector of cost coefficients.

In our setting, an equilibrium is defined via the following.

Definition 2. For a given $G = (F, \mathbf{C})$, a flow vector $\mathbf{x}$ is an equilibrium if and only if for every $i \neq j \in I$, we have:

$$x_i^s(J_i^s(\mathbf{x}) - J_i^a(\mathbf{x})) \leq 0, \tag{7a}$$

$$x_i^a(J_i^a(\mathbf{x}) - J_i^s(\mathbf{x})) \leq 0. \tag{7b}$$

Note that Equations (7) imply that for an exit link $i \in I$, if $x_i^s \neq 0$ and $x_i^a \neq 0$ at equilibrium, then, we must have that $J_i^s(\mathbf{x}) = J_i^a(\mathbf{x})$. Alternatively, if $x_i^s = 0$ ($x_i^a = 0$) at equilibrium, we have $J_i^s(\mathbf{x}) \geq J_i^a(\mathbf{x})$ ($J_i^a(\mathbf{x}) \geq J_i^s(\mathbf{x})$). Note that adoption of a Wardrop assumption implies that vehicles can be treated infinitesimally, i.e. the change caused by the unilateral lane change of a single vehicle is negligible. This is in accordance with our goal of modeling the macroscopic behavior of vehicles at diverges

III. Equilibrium Properties

In this section, we state the equilibrium properties of our model including existence and uniqueness.

A. Equilibrium Existence

Using the existence Theorem in [3] for the setting of our problem, we can conclude that there always exists at least one equilibrium for a given $G = (F, \mathbf{C})$ if the following holds.

Proposition 1. *Given a $G = (F, \mathbf{C})$ for a diverge, if the cost functions $J_i^s(\mathbf{x}), J_i^a(\mathbf{x}), i \in I$ are continuous and monotone*

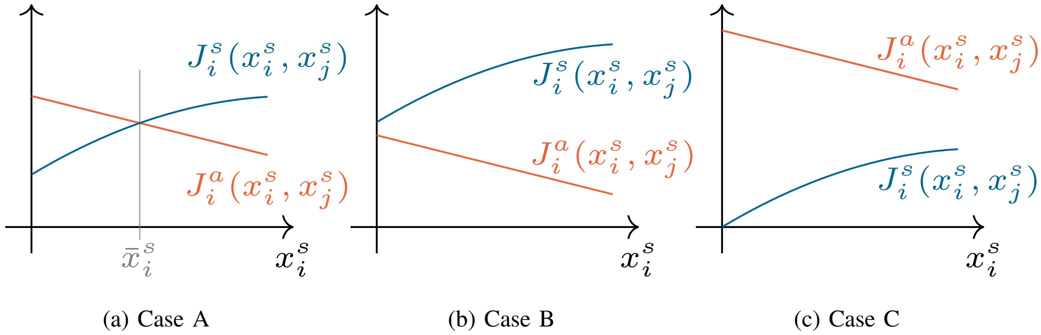


Fig. 2: Three possible configurations of $J_i^s(.)$ and $J_i^a(.)$.

in $\mathbf{x}$ *(in the sense of Definition 1), then, there exists at least one Wardrop equilibrium for* G.

Remark 2. For a diverge with two exit links, using Remark 1 and continuity of $J_i^s(.)$ and $J_i^a(.)$, we can conclude that there always exists at least one equilibrium for every $G = (F, \mathbf{C})$.

B. Equilibrium Uniqueness

As Equations (5) and (6) show, $J_i^s(.)$ depends not only on x_i^s but also on x_i^a and x_j^a. In the transportation literature, this dependence is interpreted as cost functions being *nonseparable* [5]. This nonseparability is further asymmetric, meaning that the incurred costs are not the same across J_i^s and J_i^a for $i \in I$. In addition to asymmetric nonseparability of (5) and (6), the cost functions of (5) and (6) are nonlinear. It is known that in such settings, the equilibrium exhibits very complicated behavior such as nonuniqueness. Equilibrium uniqueness is only achieved under very strong conditions which do not hold in the majority of applications [1]. Despite this complication, and the fact that none of the existing results on sufficient conditions for uniqueness of equilibrium in nonseparable asymmetric nonlinear settings can be applied to our problem, we state the conditions under which a given $G = (F, \mathbf{C})$ is guaranteed to have a unique equilibrium.

To prove uniqueness, we define an auxiliary game such that there exists a connection between the Wardrop equilibrium of our setting with the Nash equilibrium of the auxiliary game. For any given $G = (D, \mathbf{C})$, we define a two–player game $\tilde{G} = \langle P, A, (\tilde{J}_p : p \in P)\rangle$, where $P = \{1, 2\}$ is the set of players. Since both I and P are the set $\{1, 2\}$, we assume that there is a correspondence between every $p \in P$ and $i \in I$. In fact, $p = 1$ ($p = 2$) implies that $i = 1$ ($i = 2$) and vice versa. $A = A_1 \times A_2$ is the action space, and $A_p = [0, f_p]$ is the action set of player p. Morever, $\tilde{J}_p$ is the cost associated with each player $p \in P$. We let $\mathbf{y} = (y_p, p \in P)$ be the vector of the actions taken by the two players of the game $\tilde{G}$. We define $\mathbf{y}$ to be

$$\mathbf{y} = (x_i^s, i \in I). \tag{8}$$

Then, $\forall p \neq p' \in P$, we define $\tilde{J}_p(\mathbf{y})$ to be

$$\tilde{J}_p(\mathbf{y}) = \left(J_i^s(\mathbf{x}) - J_i^a(\mathbf{x})\right)^2. \tag{9}$$

For the auxiliary game $\tilde{G}$, $\mathbf{y} = (y_p : p \in P)$ is a pure Nash equilibrium if and only if

$$\forall p, p' \in P,\ y_p = B_p(y_{p'}) \tag{10}$$
$$= \text{argmin}_{y_p \in [0, f_p]} \tilde{J}_p(y_p, y_{p'}) \tag{11}$$

where B_p is the best response function of the player p. Note that since $\tilde{J}(\mathbf{y})$ is a continuous function on any closed interval, thus, the minimum is achieved. Equation (10) implies that if $y_{p'}$ is fixed, player p takes its best possible action which is minimizing its own cost $\tilde{J}_p(\mathbf{y})$. The following proposition establishes the connection between the Wardrop equilibrium of G and Nash equilibrium of $\tilde{G}$.

Proposition 2. *A flow vector* $\mathbf{x} = (x_i^s, f_i - x_i^s : i \in I)$ *is a Wardrop equilibrium for* $G = (F, \mathbf{C})$ *if and only if* $\mathbf{y} = (x_i^s, i \in I)$ *is a pure Nash equilibrium for* $\tilde{G}$ *provided that*

$$C_i^t \geq C_i^c,\ \forall i \in I. \tag{12}$$

Proof. First, note that given the demand fractions $F = \{f_1, f_2\}$, flow conservation requires that $x_i^a = f_i - x_i^s, \forall i \in I$. Thus, with a little abuse of notation, $J_i^s(\mathbf{x})$ and $J_i^a(\mathbf{x})$ can be written as $J_i^s(x_i^s, x_j^s)$ and $J_i^a(x_i^s, x_j^s)$ for every $i \neq j \in I$. We show that for every $i \neq j \in I$, (12) is a sufficient condition for $J_i^s(x_i^s, x_j^s)$ to be increasing in x_i^s, and $J_i^a(x_i^s, x_j^s)$ to be a decreasing function of x_i^s for any given x_j^s. To see this, note that for a given F, f_i's are fixed , and for every $i \neq j \in I$, we have:

$$\frac{\partial J_i^s}{\partial x_i^s} = -2C_i^c x_i^s + C_i^t + C_i^c(x_i^s) - C_i^c(f_j - x_j^s). \tag{13}$$

Equation (13) is linear in x_i^s. Moreover, for each $i \in I$, x_i^s is allowed to only take values in the interval $[0, f_i]$. Therefore, in order to obtain sufficient conditions for the positivity of (13), it is sufficient to guarantee that for every $i \neq j \in I$, $\frac{\partial J_i^s}{\partial x_i^s}$ is positive at all possible extreme points (x_i^s, x_j^s) which are $\{(0,0), (f_1, 0), (0, f_2), (f_1, f_2)\}$. Using Equation (3), it is easy to verify that the smallest possible value of (13) might be attained in $(f_1, 0)$ where $f_1 = 1$. At the point $(1, 0)$, we have $\frac{\partial J_i^s}{\partial x_i^s}(1, 0) = C_i^t - C_i^c$. Therefore, (12) is a sufficient condition for J_i^s to be increasing in x_i^s. Similarly, we can compute $\frac{\partial J_i^a}{\partial x_i^s}$. For every $i \neq j \in I$

$$\frac{\partial J_i^a}{\partial x_i^s} = -C_j^t \gamma_j - C_j^c (f_j - x_j^s). \tag{14}$$

Since $(f_j - x_j^s)$ is always greater than or equal to zero, clearly, for every $i \neq j \in I$, $J_i^a(x_i^s, x_j^s)$ is always decreasing in x_i^s for any given x_j^s.

Now, consider the best response function $B_p(y_{p'})$ in (10). In order to minimize $\tilde{J}_p(y_p, y_{p'})$ over y_p as the best response for a given $y_{p'}$, with $y_p = x_i^s$ and $y_{p'} = x_j^s$, since $\frac{\partial J_i^s}{\partial x_i^s}$ is increasing, and $\frac{\partial J_i^a}{\partial x_i^s}$ is decreasing in x_i^s under (12), the following scenarios might occur for a given x_j^s, where $i \neq j \in I$ (see Figure 2).

- Case A: $J_i^s(x_i^s, x_j^s)$ and $J_i^a(x_i^s, x_j^s)$ have an intersection on the interval $[0, f_i]$. In this case, there exists a point $\bar{x}_i^s(x_j^s) \in [0, f_i]$ such that $J_i^s(\bar{x}_i^s, x_j^s) = J_i^a(\bar{x}_i^s, x_j^s)$. Using (9), it is easy to see that in this case, $y_p = \bar{x}_i^s$ is the best response for a given $y_{p'} = x_j^s$ in the game $\tilde{G}$. If this is the case, Equations (7) are also satisfied by $\bar{x}_i^s$ for a given x_j^s. It is easy to see that the reverse is also true. Indeed, if $\bar{x}_i^s$ satisfies (7) for the given x_j^s, then, $\bar{x}_i^s$ must be the intersection of $J_i^s(x_i^s, x_j^s)$ and $J_i^a(x_i^s, x_j^s)$ on the interval $[0, f_i]$. Therefore, $y_p = x_i^s$ is the best response of $y_{p'} = x_j^s$.
- Case B: $J_i^s(x_i^s, x_j^s)$ and $J_i^a(x_i^s, x_j^s)$ do not intersect on the interval $[0, f_i]$, and $J_i^s(0, x_j^s) \geq J_i^a(0, x_j^s)$ for a given x_j^s. In this case, if $y_{p'} = x_j^s$, then $y_p = B_p(y_{p'}) = 0$ (See Figure 2, case B). It is easy to see that, $x_i^s = 0$ satisfies (7) for a given x_j^s since if $x_i^s = 0$, $x_i^a = f_i$ while $J_i^s \geq J_i^a$. The reverse is also true, if $x_i^s = 0$ satisfies (7) for a given x_j^s, $y_p = x_p^s = 0$ is the best response of $y_{p'} = x_j^s$.
- Case C: $J_i^s(x_i^s, x_j^s)$ and $J_i^a(x_i^s, x_j^s)$, do not intersect on the interval $[0, f_p]$, and $J_i^s(0, x_j^s) \leq J_i^s(0, x_j^s)$. In this case, if $y_{p'} = x_j^s$, $y_p = B_p(y_{p'}) = 1$. Similar to the case B, one can conclude that if $y_{p'} = x_j^s$, $y_p = x_i^s = 1$ is equal to $B_p(y_{p'})$ if and only if $x_i^s = 1$ satisfies (7) for a given x_j^s.

So far, we have shown that for every $p \neq p' \in P$, given $y_{p'}$, $y_p = B_p(y_{p'})$ if and only if $\mathbf{x} = (y_p, f_i - y_p : i \in I)$ satisfies (7). Therefore, $\mathbf{y} = (x_i^s, i \in I)$ is a Nash equilibrium of $\tilde{G}$ if and only if $\mathbf{x} = (x_i^s, f_i - x_i^s)_{i \in I}$ is a Wardrop equilibrium of G. $\square$

Remark 3. Notice that using the three cases described in the proof of Proposition 2, for a given $y_{p'} = x_j^s$, $B_p(y_{p'})$ can be found by first intersecting $J_i^s(x_i^s, x_j^s)$ and $J_i^a(x_i^s, x_j^s)$ and then projecting their intersection $\bar{x}_i^s(x_j^s)$ onto the interval $[0, f_p]$. We will use this fact in the remainder to prove equilibrium uniqueness.

Having Proposition 2 in mind, we are ready to state and prove the following.

Theorem 1. *For a given $G = (F, \mathbf{C})$, the Wardrop equilibrium flow vector $\mathbf{x}$ is unique if*

$$C_i^t \geq C_i^c, \ \forall i \in I, \tag{15}$$

$$(\gamma_i - 1) C_j^t \geq C_i^c, \ \forall i \in I. \tag{16}$$

Proof. Construct the auxiliary game $\tilde{G} = \langle P, A, (\tilde{J}_p, p \in P) \rangle$ described above from G. Using Proposition 2, we know that if (15) holds, $\mathbf{x}$ is a Wardrop equilibrium for G if and only if $(y_p : p \in P) = (x_i^s : i \in I)$ is a Nash equilibrium for $\tilde{G}$. We now prove that under (16), $\tilde{G}$ has a unique equilibrium; thus, G must also have a unique equilibrium if (15) and (16) hold. To see this, note that using (10), $\mathbf{y} = (x_i^s : i \in I)$ is a Nash equilibrium for $\tilde{G}$ if and only if for every $p \neq p'$, $y_p = B_p(y_{p'})$, and $y_{p'} = B_{p'}(y_p)$. These conditions can be rewritten as

$$y_p = B_p(B_{p'}(y_p)), \tag{17a}$$

$$y_{p'} = B_{p'}(B_p(y_{p'})). \tag{17b}$$

Equations (17) indicate that $\mathbf{y}$ is an equilibrium if and only if for every $p \neq p' \in P$, y_p is a fixed point for $B_p\,(B_{p'}(.))$. Thereby, $(y_p, y_{p'})$ is an equilibrium for $\tilde{G}$ if and only if $B_p(B_{p'}(.))$ intersects the line going through the origin with slope 1, at y_p. In the remainder, we prove that under (16), the slope of $B_p(B_{p'}(.))$ is always positive and smaller than 1 for every $p \neq p' \in P$. Therefore, $B_p(B_{p'}(.))$ can intersect the identity line at most once. Thus, we can then conclude that $\tilde{G}$ and therefore G will always have a unique equilibrium if (15) and (16) hold. To prove this, it suffices to prove that $0 \leq \frac{dB_p}{dy_{p'}} \leq 1$, for every $p \neq p' \in P$. To see this, let x_j^s be such that $J_i^s(x_i^s, x_j^s)$ and $J_i^a(x_i^s, x_j^s)$ intersect each other at $\bar{x}_i^s(x_j^s) \in [0, f_i]$. Using (5), and the fact that $x_i^a = f_i - x_i^s, \forall i \in I$, we have that

$$\frac{\partial J_i^s}{\partial x_j^s} = -C_i^t - C_i^c (f_i - x_i).$$

Since $(f_i - x_i) \geq 0$, we can conclude that $\frac{\partial J_i^s}{\partial x_j^s} \leq 0$. Similarly, we can compute

$$\frac{\partial J_i^a}{\partial x_j^s} = C_j^t + C_j^c (f_j - x_j^s) - C_j^c (x_j^s + f_i - x_i^s).$$

Since $(f_j - x_j^s) \geq 0$, it is easy to see that if (15) holds, $\frac{\partial J_i^a}{\partial x_j^s}$ is always positive. This implies that as x_j^s increases, J_i^s decreases while J_i^a increases. Therefore, as x_j^s increases, $\bar{x}_i^s(x_j^s)$ can only increase. However, Remark 3 implies that if $\bar{x}_i^s(x_j^s)$ lies outside the interval $[0, f_i]$, it is projected on this interval. Thus, the interval $[0, f_i]$ can be divide into three intervals $[0, f_i] = [0, m_i] \cup [m_i, n_i] \cup [n_i, f_i]$, such that $\bar{x}_i^s(x_j^s)$ is always 0 on $[0, m_i]$, and always 1 on $[n_i, f_i]$. Note that either of the intervals $[0, m_i]$, $[m_i, n_i]$ and $[n_i, f_i]$ can possibly be empty. Hence, in order to show that the slope of the best response function $B_p(.)$ is always smaller than 1 it suffices to show it for the interval $[m_i, n_i]$ where $J_i^s(x_i^s, x_j^s)$ and $J_i^a(x_i^s, x_j^s)$ do intersect.

On the interval $[m_i, n_i]$, for a given x_j^s, $\bar{x}_i^s(x_j^s)$ must satisfy:

$$J_i^s(\bar{x}_i^s, x_j^s) - J_i^a(\bar{x}_i^s, x_j^s) = 0.$$

Therefore, using implicit differentiation, $\frac{d\bar{x}_i(x_j^s)}{dx_j}$ can be computed via

$$\begin{aligned}\frac{\partial}{\partial x_i^s}\left(J_i^s(\bar{x}_i^s,x_j^s)-J_i^a(\bar{x}_i^s,x_j^s)\right)\frac{d\bar{x}_i^s(x_j^s)}{dx_j^s}+\\ \frac{\partial}{\partial x_j^s}\left(J_i^s(\bar{x}_i^s,x_j^s)-J_i^a(\bar{x}_i^s,x_j^s)\right)=0.\end{aligned}\tag{18}$$

Using (5) and (6) and flow conservation $x_i^a = f_i - x_i^s$ for all $i \in I$, we have

$$\begin{aligned}\frac{\partial}{\partial x_j^s}\left(J_i^s(\bar{x}_i^s,x_j^s)-J_i^a(\bar{x}_i^s,x_j^s)\right)=-C_i^t-C_i^c(f_i-x_i^s)-\\ C_i^c+C_j^c(x_j^s+f_i-x_i^s)-C_j^c(f_j-x_j^s).\end{aligned}\tag{19}$$

Since (19) is linear in x_i^s and x_j^s, its maximum and minimum are attained in its extreme points. It is easy to check that the maximum possible value for (19) is $-C_i^t - C_j^t + C_j^c$. If (15) holds, $-C_i^t - C_j^t + C_j^c \leq 0$. Therefore $\frac{\partial}{\partial x_j^s}\left(J_i^s(\bar{x}_i^s,x_j^s)-J_i^a(\bar{x}_i^s,x_j^s)\right) \leq 0$ under (15). Using the same trick of computing the function at its extreme points, it can be easily seen that under (15), it always the case that for every $i \in I$,

$$\frac{\partial}{\partial x_i^s}\left(J_i^s(\bar{x}_i^s,x_j^s)-J_i^a(\bar{x}_i^s,x_j^s)\right)\geq 0.$$

Hence, using (18), under (15),

$$\frac{d\bar{x}_i^s(x_j^s)}{dx_j^s}\geq 0,\quad \forall i\neq j\in I.$$

Now that we have shown that the slope of the best response function is always positive, it only remains to prove that $\frac{d\bar{x}_i^s(x_j^s)}{dx_j^s} \leq 1$. To prove this, it suffices to show that

$$\begin{aligned}\frac{\partial}{\partial x_i^s}(J_i^s(\bar{x}_i^s,x_j^s)-J_i^a(\bar{x}_i,x_j^s))\geq\\ -\left(\frac{\partial}{\partial x_j^s}\left(J_i^s(\bar{x}_i^s,x_i^s)-J_i^a(\bar{x}_i^s,x_j^s)\right)\right).\end{aligned}\tag{20}$$

Substituting (5), (6), and (19) in (20) and computing the linear function at its extreme points, we observe that (16) is a sufficient condition for (20). □

IV. Simulation Results

Up to now, we have described our modeling framework and the properties of our model. In this section, we describe how our simulation results indicate that our model can successfully predict the observed behaviors. A key element of our model which affects its functionality is the vector $\mathbf{C}$. In order to study the performance of the model, first, the model needs to be calibrated, i.e. the $\mathbf{C}$ that best fits a diverge must be found.

A. Model Calibration

Consider a diverge with two exit links $I = \{1, 2\}$. Fix the total flow of vehicles $d = d_1 + d_2$ that enter the diverge. For the fixed d, consider different configurations of demand fractions, $F^k = \{f_1^k, f_2^k\}, 1 \leq k \leq K$, where K is the total number of possible demand fraction configurations available from the data or simulation. For each value of f_1^k and $f_2^k = 1 - f_1^k$, record $(x_i^s)^k$ and $(x_i^a)^k$ which are the fraction of steadfast and altering vehicles for each destination $i \in I$ when the k^{th} demand pattern is used. We let $\mathbf{x}^k$ represent the vector $\mathbf{x}$ for the k^{th} demand configuration. Using our model, the vector of cost coefficients $\mathbf{C}$ must be found such that (7) is satisfied by $(x_i^s)^k$ and $(x_i^v)^k$ for every $k \leq K$. But since (7) contains nonlinear inequalities, finding such a $\mathbf{C}$ is nontrivial. We propose the following for calibrating $\mathbf{C}$.

For every $k \leq K$ and $i \in I$, define the integer variables $(z_i^s)^k \in \{0, 1\}$, and $(z_i^a)^k \in \{0, 1\}$ such that:

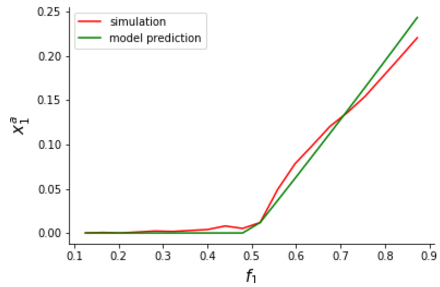
$$(x_i^s)^k(J_i^s(\mathbf{x}^k)-J_i^a(\mathbf{x}^k))\leq 0 \Longleftrightarrow (z_i^s)^k=0 \tag{21a}$$
$$(x_i^s)^k(J_i^s(\mathbf{x}^k)-J_i^a(\mathbf{x}^k))> 0 \Longleftrightarrow (z_i^s)^k=1 \tag{21b}$$
$$(x_i^a)^k(J_i^a(\mathbf{x}^k)-J_i^s(\mathbf{x}^k))\leq 0 \Longleftrightarrow (z_i^a)^k=0 \tag{21c}$$
$$(x_i^a)^k(J_i^a(\mathbf{x}^k)-J_i^s(\mathbf{x}^k))> 0 \Longleftrightarrow (z_i^a)^k=1 \tag{21d}$$

Then, we propose to solve the following optimization problem for calibrating $\mathbf{C}$.

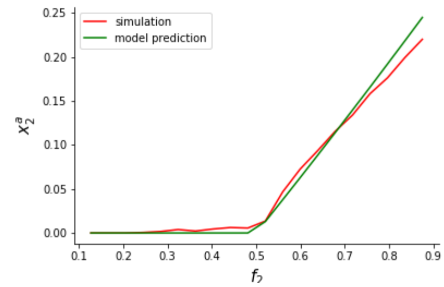
$$\begin{aligned}\underset{\mathbf{C}}{\text{minimize}}\quad & \sum_{k\in K}\sum_{i\in I}\left((z_i^s)^k+(z_i^a)^k\right)\\ \text{subject to}\quad & \text{Equations (21)}\\ & \mathbf{C}_j\geq 1,\end{aligned}\tag{22}$$

where $\mathbf{C}_j$ is the j_{th} element of $\mathbf{C}$. We use the constraint $\mathbf{C}_j \geq 1$ to avoid setting all the elements of $\mathbf{C}$ to be zero. It is important to note that since in (5) and (6), every term is multiplied by one and only one element of $\mathbf{C}$, and also, multiplying all the cost functions by the same constant does not change the Wardrop conditions, scaling $\mathbf{C}$ by a single number will not affect the model. Therefore, this constraint does not affect the model. Note that for every inequality constraint that is violated in (22), the cost is increased by 1. Thus, (22) penalizes for not satisfying (7) which are the equilibrium conditions. But, how can the problem optimization (22) be solved where the constraints are of the form (21)? To answer this, we use the procedure introduced in [18]. Let M be a large positive number, and ϵ be a small positive number close to zero. For every k, the following is equivalent to (21).

$$(x_i^s)^k(J_i^s(\mathbf{x}^k)-J_i^a(\mathbf{x}^k))\leq M(z_i^s)^k-\epsilon, \tag{23a}$$
$$-(x_i^s)^k(J_i^s(\mathbf{x}^k)-J_i^a(\mathbf{x}^k))\leq M(1-z_i^s)-\epsilon, \tag{23b}$$
$$(x_i^a)^k(J_i^a(\mathbf{x}^k)-J_i^s(\mathbf{x}^k))\leq M(z_i^a)^k-\epsilon, \tag{23c}$$
$$-(x_i^a)^k(J_i^a(\mathbf{x}^k)-J_i^s(\mathbf{x}^k))\leq M(1-z_i^a)-\epsilon. \tag{23d}$$



(a) x_1^a



(b) x_2^a

Fig. 3: The fraction of altering vehicles, x_i^a, predicted by our model and the values measured from simulation versus f_i.

Therefore, our model can be calibrated by solving the following optimization problem:

$$\begin{aligned} \underset{\mathbf{C}}{\text{minimize}} \quad & \sum_k \sum_{i \in I} \left((z_i^s)^k + (z_i^a)^k \right) \\ \text{subject to} \quad & \text{Equations (23)} \\ & \mathbf{C}_j \geq 1. \end{aligned} \tag{24}$$

Note that (24) is now a mixed–integer linear program that can be easily solved using optimization packages. Since (24) is solved offline, and, further, the number of required integer variables is small, the computational complexities for solving (24) are not overtaxing in our model.

B. Model Validation

Consider the diverge shown in Figure 1. We used the microscopic traffic simulator SUMO [13] to simulate the traffic behavior at the diverge of Figure 1 for different demand configurations. A total flow of $d_1 + d_2 = 3000\frac{\text{veh}}{\text{hour}}$ enters the diverge. The capacity of every lane is $930\frac{\text{veh}}{\text{hour}}$. At every simulation, a fraction of vehicles f_1 is assumed to take the exit link 1 while the remaining fraction of vehicles $f_2 = 1 - f_1$ is assumed to take the exit link 2. For different values of f_1, x_1^s, x_1^a, x_2^s, and x_2^a are measured. Then, this data set is used to calibrate the model, i.e. finding the $\mathbf{C}$ that best fits the data. By solving the optimization problem (24), we found the values of $\mathbf{C}$. Since our road geometry is symmetric, we introduced the additional constraints that $C_1^t = C_2^t$, $C_1^c = C_2^c$, and $\gamma_1 = \gamma_2$ in (24), and obtained the following values for $\mathbf{C}$.

$$C_1^t = C_2^t = 1,\ C_1^c = C_2^c = 1,\ \gamma_1 = \gamma_2 = 2.7.$$

Note that the obtained values of $\mathbf{C}$ satisfy (15) and (16); thus, in every scenario, Theorem 1 implies that there exits only one equilibrium. The objective function of (24) was 4 when fitting $\mathbf{C}$, meaning that only 4 inequalities were unsatisfied among our data set.

With the calibrated $\mathbf{C}$, we used our model to predict x_1^s, x_1^a, x_2^s, and x_2^a for the scenarios where the total flow entering the diverge is different from the one we used in calibration. Figure 3 demonstrates such a study for the case where the total flow of the diverge is $2500\frac{\text{veh}}{\text{hour}}$. Figure 3 shows both our simulation results and our model prediction for different configuration of the demands of exit links. As Figure 3 shows, our model can successfully predict the fraction of altering vehicles for each destination. Note that when the demand for exit 1 is low $f_1 \leq 0.5$, none of the vehicles who aim to take the exit 1 would take the more crowded lane II; therefore, $x_1^a \simeq 0$. But, with the increase of f_1, vehicles will take lane II since it will reduce their cost. Our simulation results indicate that our model is capable of predicting with great accuracy the behavior of the vehicles. We obtained similar results when the total flow that enter the diverge was varied.

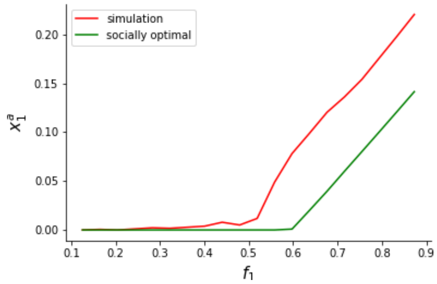
V. Socially Optimal Lane Change Behavior

Having shown that our model can lead to promising results, we can deploy it for further analysis. Intuitively, one might argue that if vehicles were less selfish, and would have chosen their destination lane far upstream of the diverge, it would have reduced the total cost of the vehicles. But how can we quantify this? Our model provides a powerful framework for analytically studying this conjecture. Assume that there is a central authority which can dictate vehicle lanes such that the total cost of the vehicles is minimum, or equivalently, that the social optimum is achieved. How would such authority pick lanes for the vehicles? To answer this question, using our model, we can solve the following optimization problem

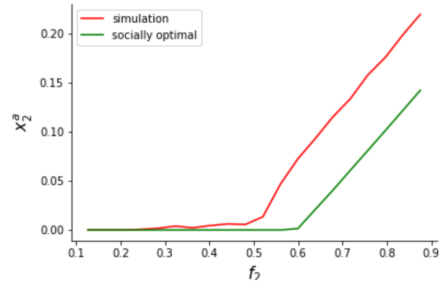
$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \sum_{i \in I} \left(x_i^s J_i^s + x_i^a J_i^a \right) \\ \text{subject to} \quad & x_i^s + x_i^a = f_i, \quad \forall i \in I, \\ & x_i^s \geq 0,\ x_i^a \geq 0, \quad \forall i \in I. \end{aligned} \tag{25}$$

Optimization (25) can be solved to find the optimal lane change behavior. Note that in (25), the decision variables are $x_i^s, x_i^a, \forall i \in I$; thus, (25) is a quadratic program which can be easily solved. This simplicity is in contrast to the existing models, where strategies for finding better lane choices are heuristically proposed through simulation.

Using the $\mathbf{C}$ that was obtained from our model calibration, we solved (25) for the case when the total flow entering the diverge is $3000\frac{\text{veh}}{\text{hour}}$. Figure 4 demonstrates the optimal lane choice of the vehicles. As Figure 4 shows, since people choose their lanes selfishly, at equilibrium, the number of



(a) x_1^a



(b) x_2^a

Fig. 4: The fraction of altering vehicles, x_i^a, from simulation and the values of altering vehicles required for social optimality as a function of f_i.

altering vehicles is larger than the optimal one. Moreover, as Figure 4 suggests, a key observation is that the optimal lane choice is *not* preventing all vehicles from changing lanes. Therefore, the optimal lane choice is in between the equilibrium and zero lane change. Our model can be used for quantitatively obtaining this trade-off, which has not been captured in previous studies.

VI. Conclusion and Future Work

We provided a game theoretic framework for macroscopically modeling the aggregate lane change maneuver of vehicles at traffic diverges. We modeled the fraction of vehicles who change lanes to take an exit, where vehicles were assumed to be selfish. We modeled the resulting equilibrium as a Wardrop equilibrium and proved the existence and uniqueness of this equilibrium. We described how our model can be easily calibrated and demonstrated via simulation studies that our model yielded promising results. For future steps, we are excited about the applications this model can have. For instance, when an authority might have control over a fraction of vehicles, we can study how to enforce lanes such that resulting equilibrium has lower social cost.

Acknowledgments

This work is supported by the National Science Foundation under Grants CPS 1446145 and CPS 1545116.

References

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

[2] M. Aycin and R. Benekohal. Comparison of car-following models for simulation. *Transportation Research Record: Journal of the Transportation Research Board*, (1678):116–127, 1999.

[3] D. Braess and G. Koch. On the existence of equilibria in asymmetrical multiclass-user transportation networks. *Transportation Science*, 13(1):56–63, 1979.

[4] B. Coifman, S. Krishnamurthy, and X. Wang. Lane-change maneuvers consuming freeway capacity. In *Traffic and Granular Flow03*, pages 3–14. Springer, 2005.

[5] J. R. Correa, A. S. Schulz, and N. E. Stier-Moses. A geometric approach to the price of anarchy in nonatomic congestion games. *Games and Economic Behavior*, 64(2):457–469, 2008.

[6] C. Dong, J. M. Dolan, and B. Litkouhi. Intention estimation for ramp merging control in autonomous driving. In *Intelligent Vehicles Symposium (IV), 2017 IEEE*, pages 1584–1589. IEEE, 2017.

[7] P. G. Gipps. A model for the structure of lane-changing decisions. *Transportation Research Part B: Methodological*, 20(5):403–414, 1986.

[8] P. Hidas. Modelling lane changing and merging in microscopic traffic simulation. *Transportation Research Part C: Emerging Technologies*, 10(5-6):351–371, 2002.

[9] W.-L. Jin. A kinematic wave theory of lane-changing traffic flow. *Transportation research part B: methodological*, 44(8-9):1001–1021, 2010.

[10] A. Kesting, M. Treiber, and D. Helbing. General lane-changing model mobil for car-following models. *Transportation Research Record: Journal of the Transportation Research Board*, (1999):86–94, 2007.

[11] U. Khan, P. Basaras, L. Schmidt-Thieme, A. Nanopoulos, and D. Katsaros. Analyzing cooperative lane change models for connected vehicles. In *Connected Vehicles and Expo (ICCVE), 2014 International Conference on*, pages 565–570. IEEE, 2014.

[12] H. Kita. A merging–giveway interaction model of cars in a merging section: a game theoretic analysis. *Transportation Research Part A: Policy and Practice*, 33(3-4):305–312, 1999.

[13] D. Krajzewicz, J. Erdmann, M. Behrisch, and L. Bieker. Recent development and applications of SUMO - Simulation of Urban MObility. *International Journal On Advances in Systems and Measurements*, 5(3&4):128–138, December 2012.

[14] J. Laval, M. Cassidy, and C. Daganzo. Impacts of lane changes at merge bottlenecks: a theory and strategies to maximize capacity. In *Traffic and Granular Flow05*, pages 577–586. Springer, 2007.

[15] J. A. Laval and C. F. Daganzo. Lane-changing in traffic streams. *Transportation Research Part B: Methodological*, 40(3):251–264, 2006.

[16] J. C. McCall, D. P. Wipf, M. M. Trivedi, and B. D. Rao. Lane change intent analysis using robust operators and sparse bayesian learning. *IEEE Transactions on Intelligent Transportation Systems*, 8(3):431–440, 2007.

[17] F. Meng, J. Su, C. Liu, and W.-H. Chen. Dynamic decision making in lane change: game theory with receding horizon. In *Control (CONTROL), 2016 UKACC 11th International Conference on*, pages 1–6. IEEE, 2016.

[18] V. Raman, A. Donzé, M. Maasoumy, R. M. Murray, A. Sangiovanni-Vincentelli, and S. A. Seshia. Model predictive control with signal temporal logic specifications. In *2014 IEEE 53rd Annual Conference on Decision and Control (CDC)*, pages 81–87. IEEE, 2014.

[19] R. W. Rothery. Car following models. *Trac Flow Theory*, 1992.

[20] A. Talebpour, H. S. Mahmassani, and S. H. Hamdar. Modeling lane-changing behavior in a connected environment: A game theory approach. *Transportation Research Part C: Emerging Technologies*, 59:216–232, 2015.

[21] S. Ulbrich and M. Maurer. Probabilistic online pomdp decision making for lane changes in fully automated driving. In *Intelligent Transportation Systems-(ITSC), 2013 16th International IEEE Conference on*, pages 2063–2067. IEEE, 2013.

[22] J. G. Wardrop. Some theoretical aspects of road traffic research. In *Inst Civil Engineers Proc London/UK/*, 1952.

[23] C. Wu, E. Vinitsky, A. Kreidieh, and A. Bayen. Multi-lane reduction: A stochastic single-lane model for lane changing. In *Intelligent Transportation Systems (ITSC), 2017 IEEE 20th International Conference on*, pages 1–8. IEEE, 2017.