Variable Speed Limit and Ramp Metering Control of Highway Networks Using Lax-Hopf Method: A Mixed Integer Linear Programming Approach

Suyash C. Vishnoi and Christian G. Claudel, *Member, IEEE*

***Abstract*— This paper presents a novel optimization formulation to solve the problem of variable speed limit control on road networks modeled by the Lighthill-Whitham-Richards (LWR) partial differential equation. It also presents some mathematical rules that allow for a reduction in the size and computational time of the optimization problem. Using the analytical solutions to the LWR model, an optimization problem is formulated for the variable speed limit and ramp metering control of traffic on highway networks using the Lax-Hopf algorithm. The resulting problem, which is non-linear in the decision variables, is transformed into a Mixed Integer Linear Program. An example is presented to show the effectiveness of the approach, including its application to a real-world highway network with multiple ramp connections. The possibility of linear relaxation of integer variables in the problem is also considered. Lastly, the method is compared to a classical Link Transmission Model formulation of the variable speed limit control problem.**

***Index Terms*— Lax-Hopf method, variable speed limits, network control, mixed integer linear programming.**

I. Introduction

OPTIMAL control of traffic on highway networks is one of the most efficient and inexpensive ways to mitigate congestion, and has been studied extensively by researchers in the past, for instance in [1]–[27] and references therein. The most popular forms of control include ramp metering (RM) and variable speed limits (VSL) [1]–[18]. A technique that is widely used to implement traffic control is Model Predictive Control (MPC) [28] such as in [15]–[20], [29] which allows network operators to obtain the optimal control inputs for a network based on its current state, over a certain time horizon using a realistic traffic model and further implement the obtained control input to the network. The various studies done in the past differ in the types of traffic models they use and the concerns that they aim to address by applying traffic control, for example safety, environmental, or congestion alleviation. In this work, we implement MPC similar to the mentioned approaches but with a different model than used in the above articles. Some popularly used models include discrete time models such as the Cell Transmission Model (CTM) [18], [30] and METANET [2], [31] and continuous time models such as the Link Transmission Model (LTM) [15], [32]. Broadly, these models differ in their theoretical accuracy and computational time. Discrete time models are more computationally intensive and require larger running times. This is because to avoid large discretization errors, they require breaking down the network into a large number of small cells and performing density computation for each cell at each time-step where the time-step value again needs to be small. On the other hand, continuous time approaches, such as the LTM as well as the one presented in this paper, can compute the evolution of density over the entire link space over longer acceptable duration at once. This makes continuous time approaches more attractive for real-time applications. Since control is generally applied in real-time, network operators desire a technique that is not only accurate but also robust and fast enough to complete the optimization before the next control input is due, which in case of online control is most often the very next time-step.

The CTM and LTM approaches are based on the Lighthill-Whitham-Richards (LWR) model [33], [34] of traffic flow which is a first order conservation law that describes the evolution of traffic density on highways through a partial differential equation (PDE). These approaches provide an approximate solution to the LWR model. While the CTM is prone to discretization errors, the LTM though exact in several scenarios, does not provide an exact solution in the presence of expansion waves in the traffic. It is possible to solve the LWR equation exactly using the Lax-Hopf formula [35], [36] after writing the LWR PDE in the form of a *Hamilton-Jacobi* equation. Obtaining an exact solution to the process model is important in network controls where minute errors can increase exponentially over time. It has been shown previously that the optimization problem associated with the Lax-Hopf formula could be solved semi-analytically [37]. Furthermore, in the case of a triangular fundamental diagram, we can reduce the complexity of the problem [38] by skipping the evaluation of some solution components. Finally, it can be shown that when the initial density on a highway section is uniform, the computational complexity of the *Fast* Lax-Hopf approach [38] is identical to that of the LTM [32] which is arguably the fastest of the solution schemes for the LWR model to date.

Highway state estimation and boundary control problems have been solved in the literature [27], [39] using efficient

Manuscript received 21 April 2020; revised 10 November 2020 and 11 March 2021; accepted 25 March 2021. Date of publication 7 April 2021; date of current version 8 July 2022. This work was supported by the National Science Foundation (NSF) CIS under Grant 1917056, NSF CPS under Grant 1739964, the United States Department of Transportation, and the Texas Department of Transportation. The Associate Editor for this article was I. Papamichail. *(Corresponding author: Suyash C. Vishnoi.)*

The authors are with the Department of Civil, Architectural and Environmental Engineering, The University of Texas at Austin, Austin, TX 78712 USA (e-mail: scvishnoi@utexas.edu).

Digital Object Identifier 10.1109/TITS.2021.3069971

optimization schemes based on the LWR model and Lax-Hopf method. However, these earlier studies assume that the speed limits on the links, which are parameters of the model, remain static over time. On the contrary, we assume in this paper that the speed limits are dynamic, and controlled and hence are also a variable in the formulation. One should note that in this paper, unlike many earlier work on variable speed limits such as [15], we assume that all vehicles of the highway section immediately observe the updated posted speed limit. In earlier work including [15], speed limits changes only affect the vehicles that enter the section after the switch. The former assumption is probably more realistic since typically variable speed limit signs are posted across entire highway sections, and any change in posted speed limit is propagated through an entire section [40].

To this end, we first develop analytical formulae similar to [37] for the Lax-Hopf approach under variable speed limits. These formulae however are non-linear in the decision variables which makes the optimization problem non-linear as well. This poses a problem for real-time control as non-linear problems are computationally harder and take longer to solve. In this article, we formulate the problem as a Mixed Integer Linear Program (MILP) consisting of continuous and binary variables and linear constraints using methods presented in [41], [42]. MILPs can generally be solved within a reasonable amount of time with the help of available toolboxes like CPLEX. While earlier papers [27], [39] also lead to a MILP, the formulation of this MILP is considerably different. Previously, the nonlinearities were only caused by junction conditions (the control problem on a single link was formulated as a Linear Program (LP)), while in the present case the nonlinearities arise both from junction conditions and model parameter switches (the control problem on a single link is a MILP in general). Similar techniques have been used in the past to transform the LTM variable speed limit control problem into a MILP as well [15]. Additionally, we also show that the principles that allow for the reduction of the number of steps in the computation, in case of static speed limits [38], do not apply to the variable speed limit case in general. Thereafter, we establish a new set of rules that allow for a reduction of computational complexity in the latter problem.

The paper is organized as follows: In Section II we introduce and discuss the theory behind the Lax-Hopf approach for solving the LWR model. In Section III we present the mathematical definitions of initial and boundary condition blocks and the analytical solutions for the variable speed limit case. In Section IV we present the formulation of the optimization problem. In that, we first discuss some properties that allow us to reduce the number of solution steps and the assumptions taken in formulating the problem. Then we discuss the model constraints, the objective functions and the decision variables along with the mixed-integer transformations. In Section V we discuss an example of implementation of the optimization problem on a real-world highway with multiple ramps. In Section VI we compare LTM based optimal control with the present approach, using a example. The paper is concluded with remarks, along with the scope of future work. The work in this paper derives from the first author's master's thesis 'Variable Speed Limit and Ramp Metering Control of Highway Networks using Lax-Hopf Method: A Mixed Integer Linear Programming Approach,' copyright (2020) by Suyash Chandra Vishnoi.

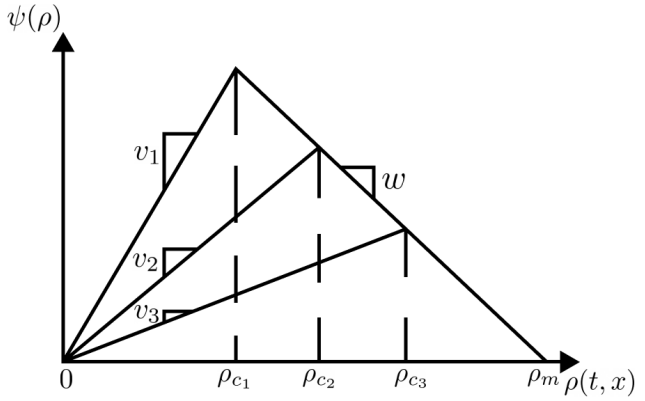

Fig. 1. Triangular fundamental diagram under variable speed limits. ρ_{c_1}, ρ_{c_2} and ρ_{c_3} represent the values of critical density corresponding to the three values of speed limits v_1, v_2 and v_3. The congestion wave speed w and maximum density ρ_m are taken to remain unchanged with the change in speed limit.

II. Background

A. The Lighthill-Whitham-Richards Traffic Flow Model

Traffic flow can be described by the *density* function, denoted as $\rho(\cdot,\cdot)$ which is equal to the number of vehicles per space unit. The LWR model [33], [34] that describes the evolution of $\rho(\cdot,\cdot)$ is written as follows

$$\frac{\partial \rho(t,x)}{\partial t} + \frac{\partial \psi(\rho(t,x))}{\partial x} = 0 \tag{1}$$

The function $\psi(\cdot)$ is called *flux function* for which several models have been proposed called *fundamental diagram*. In this paper we use the triangular fundamental diagram defined below which has been extensively used in the literature [30], [43], [44].

$$\psi(\rho) = \begin{cases} v\rho & \text{if } \rho \le \rho_c \\ w(\rho - \rho_m) & \text{otherwise} \end{cases} \tag{2}$$

Here, v denotes the free-flow speed which is assumed to be equal to the speed limit value, w denotes the congestion-wave speed, ρ_m denotes the maximum density, and ρ_c denotes the critical density. The part of the fundamental diagram which is below ρ_c is called the free-flow branch and above it is called the congestion branch.

Under variable speed limits, which means that v is variable, it is assumed that the change in v only affects the value of ρ_c while ρ_m and w remain unchanged. In that, an increase in v results in a smaller value of ρ_c while a decrease in v results in a larger value of ρ_c. The impact of variable speed limits on the fundamental diagram is depicted in Figure 1 which shows three possible values of speed limits v_1, v_2 and v_3, and the corresponding critical densities. Under static speed limit, the diagram would only consist of the free-flow branch corresponding to one of the speeds which would be the static speed limit in that case. The ensuing results and definitions in this section are independent of whether the speed limit is variable or static.

B. Hamilton-Jacobi Equation

The state of traffic can also be described by the *Moskowitz function* [45], [46] denoted by $\mathbf{M}(\cdot,\cdot)$ which can be interpreted as the cumulative number of vehicles at a point.

The density function $\rho(\cdot,\cdot)$ is related [46] to the spatial derivative of the Moskowitz function $\mathbf{M}(\cdot,\cdot)$ as follows:

$$\rho(t,x) = -\frac{\partial \mathbf{M}(t,x)}{\partial x} \tag{3}$$

If the density function is to be modeled by the LWR PDE (1), the Moskowitz function satisfies a *Hamilton-Jacobi* (HJ) PDE obtained [35], [47] by integration of the LWR PDE:

$$\frac{\partial \mathbf{M}(t,x)}{\partial t} - \psi\left(-\frac{\partial \mathbf{M}(t,x)}{\partial x}\right) = 0 \tag{4}$$

Several classes of weak solutions to equation (4) exist, such as viscosity solutions [48], [49] or Barron-Jensen/Frankowska (B-J/F) solutions [50], [51]. For the problem investigated in this article, these solutions are equivalent and can be computed implicitly using a Lax-Hopf formula.

C. Barron-Jensen/Frankowska Solutions to Hamilton Jacobi Equations

To characterize the B-J/F solutions, we first define the Legendre-Fenchel transform of the Hamiltonian $\psi(\cdot)$ as follows.

[Legendre-Fenchel transform] For an upper semicontinuous Hamiltonian $\psi(\cdot)$, the Legendre-Fenchel transform $\varphi^*(\cdot)$ is given by:

$$\varphi^*(u) := \sup_{p \in \mathrm{Dom}(\psi)} [p \cdot u + \psi(p)] \tag{5}$$

To solve the HJ PDE (4) we need to define *value conditions*. In this paper, we assume that for any highway section ξ represents the upstream boundary while χ represents the downstream boundary of the link space. Then a *value condition* can be defined as follows:

[Value condition] A value condition $\mathbf{c}(\cdot,\cdot)$ is a lower semicontinuous function defined on a subset of $[0, t_{\max}] \times [\xi, \chi]$.

Each of these functions is defined on a subset of $\mathbb{R}_+ \times [\xi, \chi]$. The value conditions are formally defined in Section III-A.

For each value condition $\mathbf{c}(\cdot,\cdot)$, we define the partial solution [37] to the HJ PDE (4) using the Lax-Hopf formula [35], [47].

[Lax-Hopf formula] Let $\psi(\cdot)$ be a concave Hamiltonian, and let $\varphi^*(\cdot)$ be its Legendre-Fenchel transform (5). Let $\mathbf{c}(\cdot,\cdot)$ be a lower semicontinuous value condition, as in Definition II-C. The B-J/F solution $\mathbf{M_c}(\cdot,\cdot)$ to (4) associated with $\mathbf{c}(\cdot,\cdot)$ can be algebraically represented [35], [47] by:

$$\mathbf{M_c}(t,x) = \inf_{(u,T) \in \mathrm{Dom}(\varphi^*) \times \mathbb{R}_+} \left(\mathbf{c}(t-T, x+Tu) + T\varphi^*(u)\right) \tag{6}$$

D. Inf-Morphism Property

The inf-morphism property can be formally derived through capture basins, such as in [47] and is given as follows

[Inf-morphism property] Let the value condition $\mathbf{c}(\cdot,\cdot)$ be minimum of a finite number of lower semicontinuous functions:

$$\forall (t,x) \in [0, t_{\max}] \times [\xi, \chi], \quad \mathbf{c}(t,x) := \min_{j \in J} \mathbf{c}_j(t,x) \tag{7}$$

The solution $\mathbf{M_c}(\cdot,\cdot)$ associated with the above value condition can be decomposed [35], [36], [47] as:

$$\forall (t,x) \in [0, t_{\max}] \times [\xi, \chi], \quad \mathbf{M_c}(t,x) = \min_{j \in J} \mathbf{M}_{\mathbf{c}_j}(t,x) \tag{8}$$

The inf-morphism property is essential for the derivation of the LWR PDE model constraints and allows us to instantiate these constraints as inequalities.

[Model compatibility constraints for block value conditions] Let $\mathbf{c}(\cdot,\cdot) = \min_{j \in J} \mathbf{c}_j(\cdot,\cdot)$ be given, and let $\mathbf{M_c}(\cdot,\cdot)$ be defined as in (6). The value condition $\mathbf{c}(\cdot,\cdot)$ satisfies $\forall (t,x) \in \mathrm{Dom}(\mathbf{c}), \mathbf{M_c}(t,x) = \mathbf{c}(t,x)$ if and only if the following inequality constraints are satisfied:

$$\mathbf{M}_{\mathbf{c}_j}(t,x) \geq \mathbf{c}_i(t,x) \quad \forall (t,x) \in \mathrm{Dom}(\mathbf{c}_i), \ \forall (i,j) \in J^2 \tag{9}$$

Readers are referred to [52] for various properties and proofs related to the above definitions and propositions. To solve the problem completely, we need to evaluate the functions $\mathbf{M}_{\mathbf{c}_j}(\cdot,\cdot)$ explicitly. In the following section we present the mathematical formulae for the value condition blocks along with the explicit solutions for the Moskowitz function under variable speed limits.

III. Explicit Solutions to Piecewise Affine Initial and Boundary Conditions

A. Definition of Piecewise Affine Initial, Upstream and Downstream Boundary Conditions

The initial, upstream and downstream boundary conditions associated with the HJ PDE (4) can be formally defined as follows:

[Piecewise affine initial, upstream and downstream boundary conditions] Let us define $\mathbb{K} = \{1, \ldots, k_{\max}\}$ and $\mathbb{N} = \{1, \ldots, n_{\max}\}$. For all $k \in \mathbb{K}$ and $n \in \mathbb{N}$, we define the following functions, respectively called initial, upstream and downstream conditions:

$$M_k(t,x) = \begin{cases} -\sum_{i=1}^{k-1} \rho_{\mathrm{ini}}(i)(x_{i+1} - x_i) \\ -\rho_{\mathrm{ini}}(k)(x - x_k) & \text{if } t = 0 \\ & \text{and } x \in [x_k, x_{k+1}] \\ +\infty & \text{otherwise} \end{cases} \tag{10}$$

$$\gamma_n(t,x) = \begin{cases} \sum_{i=1}^{n-1} q_{\mathrm{in}}(i)T \\ +q_{\mathrm{in}}(n)(t - (n-1)T) \\ \qquad \text{if } x = \xi \\ \qquad \text{and } t \in [(n-1)T, nT] \\ +\infty \quad \text{otherwise} \end{cases} \tag{11}$$

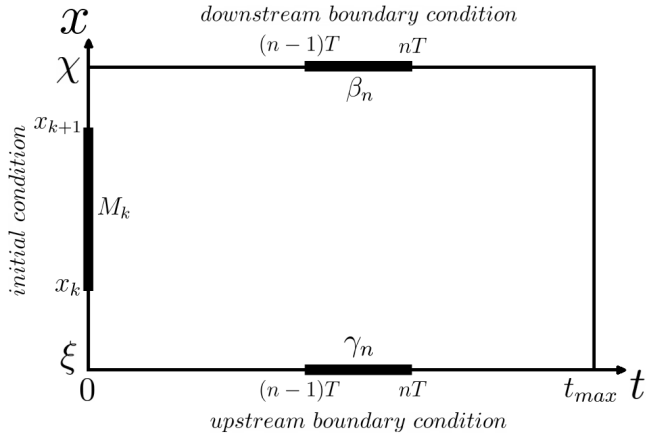

Fig. 2. Domains of the initial, upstream and downstream boundary conditions. The upstream and downstream boundary condition blocks respectively denoted by $\gamma_n(\cdot,\cdot)$ and $\beta_n(\cdot,\cdot)$ are defined on line segments corresponding to the upstream and downstream boundaries of the physical domain. In contrast, the block initial conditions $M_k(\cdot,\cdot)$ are defined on line segments corresponding to the initial time. Note that in the actual problem block initial conditions cover the entire physical domain $[\xi,\chi]$, while block boundary conditions cover the temporal domain $[0, t_{\max}]$.

$$\beta_n(t,x) = \begin{cases} \sum_{i=1}^{n-1} q_{\text{out}}(i)T \\ +q_{\text{out}}(n)(t-(n-1)T) \\ -\sum_{k=1}^{k_{max}} \rho_{ini}(k)(x_{k+1}-x_k) \\ \qquad \text{if } x=\chi \\ \qquad \text{and } t\in[(n-1)T, nT] \\ +\infty \quad \text{otherwise} \end{cases} \tag{12}$$

Here, $q_{in}(n) \in [0, \upsilon\rho_c]$ and $q_{out}(n) \in [0, \upsilon\rho_c], \forall n \in \mathbb{N}$ denote the inflow and outflow for the given highway section respectively and $\rho_{ini}(k) \in [0, \rho_m], \forall k \in \mathbb{K}$ denotes the initial density conditions. $T > 0$ is the duration of each time-step such that $t_{max} = n_{max}T$. Note that all the upstream and downstream boundary conditions have the same time duration that is T, while the initial conditions can have different lengths in space with their bounds denoted by $[x_k, x_{k+1}]\forall k \in \mathbb{K}$ such that $x_k \le x_{k+1}$, $x_1 = \xi$ and $x_{k_{max}+1} = \chi$. The domains of definitions of these functions are illustrated in Fig. 2. Since the formulation of the above definitions is independent of the free-flow speed, they are unaltered by the introduction of variable speed limits. One difference, however, will be in the upper bound for the flow variables that will be the maximum flow based on the different speed limit values and which will be obtained from the fundamental diagram.

B. Analytical Solutions to Piecewise Affine Initial, Upstream and Downstream Boundary Conditions Under Variable Speed Limits

This section presents a methodology to calculate $\mathbf{M}_{\mathbf{c}}(\cdot,\cdot)$ for any value condition block $\mathbf{c}(\cdot,\cdot)$ at any point in the time-space domain with the help of analytical solutions for the static speed limit case that can be found in [37], [52] and have been put in Appendix for reference. The idea is to define temporary density condition blocks (say $\mathbf{c}'(\cdot,\cdot)$) over the link space at the time of a switch in speed limit, and use the analytical solutions for the static speed limit case to calculate $\mathbf{M}_{\mathbf{c}'}(\cdot,\cdot)$ due to those temporary condition blocks at points from that switch time to the next. We define the following values to formulate the analytical solutions later in this section: N_{sec} is the number of sections in the space-time domain of the link separated in time by the time of switches in speed limit given by the set $\{T_{sw(j)} |\ \ j \in [1, N_{sec}-1]\}$. So each section is defined by the set $[\xi,\chi]\times[T_{sw(j)}, T_{sw(j+1)}]\forall j \in [1, N_{sec}-1]$ with the exception of the first and last section that are defined by $[\xi,\chi] \times [0, T_{sw(1)}]$ and $[\xi,\chi] \times [T_{sw(N_{sec}-1)}, t_{max}]$ respectively. The speed limit and the corresponding critical density value for a section defined by $[\xi,\chi]\times[T_{sw(j-1)}, T_{sw(j)}]\forall j \in [2, N_{sec}-1]$ is denoted by υ_j and $\rho_{c,j}$, and $\upsilon_1, \rho_{c,1}$ and $\upsilon_{N_{sec}}, \rho_{c,N_{sec}}$ for the first and last sections respectively. $M^c_{sw(j),i}$ denotes a temporary density condition block created at $T_{sw(j)}$ due to a condition block $\mathbf{c}(\cdot,\cdot)$. Here, subscript i denotes the index of the temporary condition block created due to block $\mathbf{c}(\cdot,\cdot)$ at that time starting from the most upstream block. The density of this temporary condition block is denoted by $\rho^c_{sw(j),i}$ and its bound in space is given by $[x^c_{sw(j),i}, x^c_{sw(j),i+1}]$ similar to (10). We now look at how the formulation is done for the initial density condition blocks.

1) Analytical Solutions for Initial Density Conditions Under Variable Speed Limits: Let an initial density condition block M_k be defined as in (10) for an arbitrary k. For this condition block we already know the analytical formulae (34) to obtain the value of $\mathbf{M}_{M_k}(\cdot,\cdot)$ at any point in the time-space domain when there are no switches in the speed limit. We can use those formulae to calculate $\mathbf{M}_{M_k}(\cdot,\cdot)$ at anytime before $T_{sw(1)}$, that is before a switch in speed limit occurs. To obtain $\mathbf{M}_{M_k}(\cdot,\cdot)$ at points after $T_{sw(1)}$, we follow the idea of defining temporary density condition blocks at the switch times which is implemented as follows.

First, we use the analytical formulae in (34) to calculate $\mathbf{M}_{M_k}(t,x)$ at $t = T_{sw(1)}$ for all x on the edges of the characteristic traffic waves emerging from the value condition block and lying within the link space. If any edges of the characteristic waves cross the link boundaries before $T_{sw(1)}$, we then calculate $\mathbf{M}_{M_k}(\cdot,\cdot)$ at the link boundaries, that is ξ or χ, instead of that edge. Further, using (3) with the values of $\mathbf{M}_{M_k}(t,x)$ obtained at the said x, we can calculate the density of traffic for the different traffic regions formed at $T_{sw(1)}$ due to the condition block M_k. As mentioned before, these densities are denoted by $\rho^{M_k}_{sw(1),i}$ and the density regions by $[x^{M_k}_{sw(1),i}, x^{M_k}_{sw(1),i+1}]$. From these density values, we define temporary density condition blocks denoted by $M^{M_k}_{sw(1),i}$ along the link space at time $T_{sw(1)}$. These are defined as in (13), as shown at the bottom of the next page. In defining these temporary condition blocks, we assume temporarily that $\mathbf{M}_{M_k}(t,x)$ at $(T_{sw(1)}, x^{M_k}_{sw(1),1})$ is 0. This does not affect the formulation ahead since this value would only be added as a constant in the defined formulae. This allows us to avoid writing this value in the ensuing formulae which otherwise would just make the formulae look even more complex. Instead, now it can just be added when calculating the final value of $\mathbf{M}_{M_k}(\cdot,\cdot)$ at any later point.

For these newly defined temporary density condition blocks, the analytical formulae to get the solution values denoted by $\mathbf{M}_{M^{M_k}_{sw(1),i}}(t,x)$ for the section $[\xi,\chi]\times[T_{sw(1)},T_{sw(2)}]$, taking $\mathbf{M}_{M_k}(T_{sw(1)},x^{M_k}_{sw(1),1})$ as 0, can be written as in (14), as shown at the bottom of the page. For this section, the speed limit is denoted by v_2 and the critical density is denoted by $\rho_{c,2}$.

Fig. 3 presents an example of how characteristic traffic waves move forward from an initial density condition block M_k. We can utilise the inf-morphism property to obtain the combined solution value due to this set of temporary density condition blocks at any point in $[\xi,\chi]\times[T_{sw(1)},T_{sw(2)}]$. Let this solution value be denoted by $\mathbf{M}_{M^{M_k}_{sw(1)}}(t,x)$. Note that this notation does not have the subscript i in it because it is the combined value from all temporary blocks at that switch and not just the value due to individual temporary blocks that are indexed using i. Now to obtain the value $\mathbf{M}_{M_k}(t,x)$ due to the original condition block M_k for $(t,x)\in[T_{sw(1)},T_{sw(2)}]\times[\xi,\chi]$, we can simply add the actual value of $\mathbf{M}_{M_k}(T_{sw(1)},x^{M_k}_{sw(1),1})$ to $\mathbf{M}_{M^{M_k}_{sw(1)}}(t,x)$. From here, we can calculate $\mathbf{M}_{M_k}(\cdot,\cdot)$ at the ends of traffic regions that will occur at $T_{sw(2)}$ and repeat the above procedure from defining temporary condition blocks $M^{M_k}_{sw(2),i}$ to obtaining the solution values $\mathbf{M}_{M_k}(t,x)$ for $(t,x)\in[T_{sw(2)},T_{sw(3)}]\times[\xi,\chi]$ and so on. Therefore, for any point (t,x) with $t\in[T_{sw(sec-1)},T_{sw(sec)}]$ where sec denotes the number of the section containing the point, such that $sec>1$, the formula for calculating $\mathbf{M}_{M_k}(t,x)$ can be written as in (15), as shown at the bottom of the page.

2) Analytical Solutions for Boundary Conditions Under Variable Speed Limits: We can use the same approach as the initial density condition blocks to calculate $\mathbf{M}_{\gamma_n}(\cdot,\cdot)$ and $\mathbf{M}_{\beta_n}(\cdot,\cdot)$ for any upstream and downstream boundary condition blocks γ_n and β_n respectively at all points throughout the time horizon. Let $secn$ denote the section containing

$$M^{M_k}_{sw(1),i}(t,x)=\begin{cases}-\sum_{j=1}^{i-1}\rho^{M_k}_{sw(1),j}(x^{M_k}_{sw(1),j+1}-x^{M_k}_{sw(1),j})\\ -\rho^{M_k}_{sw(1),i}(x-x^{M_k}_{sw(1),i}) & \text{if } t=T_{sw(1)}\\ & \text{and } x\in[x^{M_k}_{sw(1),i},x^{M_k}_{sw(1),i+1}]\\ +\infty & \text{otherwise}\end{cases}\tag{13}$$

$$\mathbf{M}_{M^{M_k}_{sw(1),i}}(t,x)=\begin{cases}+\infty & \text{if } x\le x^{M_k}_{sw(1),i}+(t-T_{sw(1)})w\\ & \text{or } x\ge x^{M_k}_{sw(1),i+1}+(t-T_{sw(1)})v_2\\ -\sum_{j=1}^{i-1}\rho^{M_k}_{sw(1),j}(x^{M_k}_{sw(1),j+1}-x^{M_k}_{sw(1),j})\\ +\rho^{M_k}_{sw(1),i}((t-T_{sw(1)})v_2+x^{M_k}_{sw(1),i}-x) & \text{if } x^{M_k}_{sw(1),i+1}+(t-T_{sw(1)})v_2\ge x\\ & \text{and } x^{M_k}_{sw(1),i}+(t-T_{sw(1)})v_2\le x\\ & \text{and } \rho^{M_k}_{sw(1),i}\le\rho_{c,2}\\ -\sum_{j=1}^{i-1}\rho^{M_k}_{sw(1),j}(x^{M_k}_{sw(1),j+1}-x^{M_k}_{sw(1),j})\\ +\rho_{c,2}((t-T_{sw(1)})v_2+x^{M_k}_{sw(1),i}-x) & \text{if } x^{M_k}_{sw(1),i}+(t-T_{sw(1)})v_2\ge x\\ & \text{and } x^{M_k}_{sw(1),i}+(t-T_{sw(1)})w\le x\\ & \text{and } \rho^{M_k}_{sw(1),i}\le\rho_{c,2}\\ -\sum_{j=1}^{i-1}\rho^{M_k}_{sw(1),j}(x^{M_k}_{sw(1),j+1}-x^{M_k}_{sw(1),j})\\ +\rho^{M_k}_{sw(1),i}((t-T_{sw(1)})w+x^{M_k}_{sw(1),i}-x)\\ -\rho_m(t-T_{sw(1)})w & \text{if } x^{M_k}_{sw(1),i}+(t-T_{sw(1)})w\le x\\ & \text{and } x^{M_k}_{sw(1),i+1}+(t-T_{sw(1)})w\ge x\\ & \text{and } \rho^{M_k}_{sw(1),i}\ge\rho_{c,2}\\ -\sum_{j=1}^{i}\rho^{M_k}_{sw(1),j}(x^{M_k}_{sw(1),j+1}-x^{M_k}_{sw(1),j})\\ +\rho_{c,2}((t-T_{sw(1)})w+x^{M_k}_{sw(1),i+1}-x)\\ -\rho_m(t-T_{sw(1)})w & \text{if } x^{M_k}_{sw(1),i+1}+(t-T_{sw(1)})v_2\ge x\\ & \text{and } x^{M_k}_{sw(1),i+1}+(t-T_{sw(1)})w\le x\\ & \text{and } \rho^{M_k}_{sw(1),i}\ge\rho_{c,2}\end{cases}\tag{14}$$

$$\mathbf{M}_{M_k}(t,x)=\mathbf{M}_{M_k}(T_{sw(1)},x^{M_k}_{sw(1),1})+\sum_{j=1}^{sec-2}\mathbf{M}_{M^{M_k}_{sw(j)}}(T_{sw(j+1)},x^{M_k}_{sw(j+1),1})+\mathbf{M}_{M^{M_k}_{sw(sec-1)}}(t,x)\tag{15}$$

$$\mathbf{M}_{\gamma_n}(t,x)=\mathbf{M}_{\gamma_n}(T_{sw(secn)},x^{\gamma_n}_{sw(secn),1})+\sum_{j=secn}^{sec-2}\mathbf{M}_{M^{\gamma_n}_{sw(j)}}(T_{sw(j+1)},x^{\gamma_n}_{sw(j+1),1})+\mathbf{M}_{M^{\gamma_n}_{sw(sec-1)}}(t,x)\tag{16}$$

$$\mathbf{M}_{\beta_n}(t,x)=\mathbf{M}_{\beta_n}(T_{sw(secn)},x^{\beta_n}_{sw(secn),1})+\sum_{j=secn}^{sec-2}\mathbf{M}_{M^{\beta_n}_{sw(j)}}(T_{sw(j+1)},x^{\beta_n}_{sw(j+1),1})+\mathbf{M}_{M^{\beta_n}_{sw(sec-1)}}(t,x)\tag{17}$$

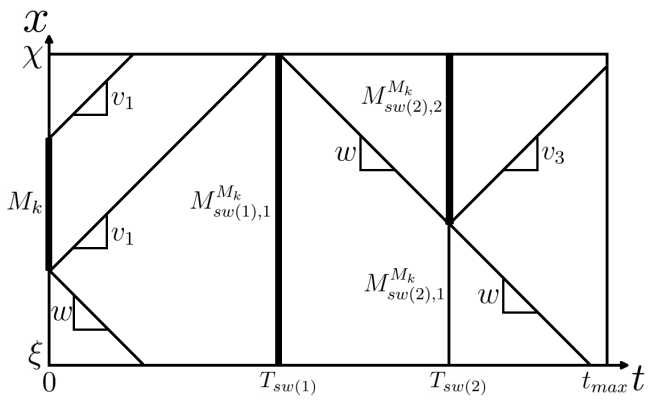

Fig. 3. Characteristic traffic wave propagation from an initial condition block M_k in time over the link space under variable speed limits. The speed limit for the link is assumed to change at times $T_{sw(1)}$ and $T_{sw(2)}$. It is also assumed here that $\rho_{ini}(k) \leq \rho_{c,1}$. The initial condition block M_k leads to the formation of one temporary density condition block, $M^{M_k}_{sw(1),1}$ at the first switch with density denoted by $\rho^{M_k}_{sw(1),1}$. From this block, assuming $\rho^{M_k}_{sw(1),1} \geq \rho_{c,2}$, the traffic waves propagate forward to form two temporary condition blocks, $M^{M_k}_{sw(2),1}$ and $M^{M_k}_{sw(2),2}$, at the second switch. After the second switch, the figure depicts the propagation of traffic waves from only $M^{M_k}_{sw(2),2}$ assuming $\rho^{M_k}_{sw(2),2} \leq \rho_{c,3}$.

the boundary condition block. Then the solutions at points within the section can be obtained using formulae from [37], [52]. The formulae to obtain $\mathbf{M}_{\gamma_n}(t,x)$ and $\mathbf{M}_{\beta_n}(t,x)$ for $(t,x) \in [T_{sw(sec-1)}, T_{sw(sec)}] \times [\xi, \chi]$, where $sec > secn$, can be written as (16) and (17), as shown at the bottom of the previous page respectively.

While these formulae can be used to calculate the solution values due to any value condition block at any point in the time-space domain, we will see that by the virtue of certain observations and assumptions presented in Section IV-C.1, we only need to calculate these values at points up to the next section in the time horizon. This eliminates the need to evaluate the summation portion of these formulae.

IV. Formulation of the Control Problem as a Mixed Integer Linear Program

In this section, we formulate the optimization problem for the VSL and RM control of networks and transform it into a MILP. We discuss how the problem can be applied in a Model Predictive Control (MPC) framework. Furthermore, we discuss the different components of the optimization problem namely the objective functions, constraints and decision variables and how to cast them as Mixed Integer Linear. In this paper, we assume that only some links have VSL control while other links operate at the normal speed limit of the highway which cannot be controlled. Let the set of all links be denoted by $\mathfrak{L}$, the set of input links be denoted by $\mathfrak{I}$, the set of links with VSL control be denoted by $\mathfrak{V}$ and the set of on-ramps with RM control be denoted $\mathfrak{M}$.

A. Model Predictive Control

Model predictive control (MPC) [28] is an advanced control technique that uses a prediction model to obtain optimal control inputs for a system over a prediction horizon by solving an optimization problem with the desired objective function. In this paper, we implement an online control scheme that works on a rolling horizon MPC approach. In that, we solve the optimization problem to obtain the control inputs for the entire prediction horizon but only apply the first step of the input to the actual system. We then record the state of the system at the end of this step and use it as the initial data for the next run of the optimization. In general, the prediction horizon should be larger than the time taken for vehicles to travel from the controlled link to the end of the highway which is satisfied in the examples considered in this paper. In this paper, the prediction model used is the Lax-Hopf scheme discussed in the previous sections. In the following sections, we will formulate the various parts of the optimization problem implemented in the MPC.

B. Objective Function

The objective function is chosen to fit the specific requirements for which the control is performed. Objective functions including total flow denoted by J_{TF}, total travel time denoted by J_{TTT}, and total average congestion denoted by J_{TAC}, can be used to deal with matters of traffic flow and congestion. These can be written as

$$J_{\mathrm{TF}} = -\sum_{\mathfrak{L}} \sum_{n=1}^{n_{max}} \big(q_{in}(n) + q_{out}(n)\big), \tag{18}$$

$$J_{\mathrm{TTT}} = T.\Bigg[\sum_{\mathfrak{L}} \sum_{n=1}^{n_{max}} \big(\mathbf{M}(nT,\xi) - \mathbf{M}(nT,\chi)\big) + T \cdot \sum_{\mathfrak{I}} \sum_{n=1}^{n_{max}} \big(Q_{in}(n) - q_{in}(n)\big)\Bigg], \tag{19}$$

$$J_{\mathrm{TAC}} = \sum_{\mathfrak{L}} \max\big(\mathbf{M}(t_{\max},\xi) - \mathbf{M}(t_{\max},\chi)\big)/L, k_{c,max}\big) \cdot L \tag{20}$$

where $\mathbf{M}(\cdot,\cdot)$ represents the final solution value due to all the condition blocks at the given point, $Q_{in}(n)$ is the demand for the input link during time-step n, $k_{c,max}$ is the maximum of the possible critical density values, and all the variables are specific to the link being summed over in the summations. We can also add the second term from (19) related to the input flows as a penalty term in other objectives to penalize queue formation at network entrances.

Also to ensure safety by avoiding large fluctuations in speed limits and ramp control, we introduce the following penalty terms:

$$\gamma \sum_{\mathfrak{V}} \sum_{n=1}^{n_{max}-1} \mid v_n - v_{n+1} \mid \tag{21}$$

$$\zeta \sum_{\mathfrak{M}} \sum_{n=1}^{n_{max}-1} \mid r_n - r_{n+1} \mid \tag{22}$$

where γ and ζ are the weights attached to the penalty terms and are set arbitrarily.

TABLE I
$\mathbf{M}(\cdot,\cdot)$ VALUES DUE TO DOWNSTREAM CONDITION BLOCKS AT UPSTREAM BOUNDARY POINTS

	Point 4	**Point 5**	**Point 6**	**Point 7**	**Point 8**	**Point 9**	**Point 10**	**Point11**
Block 1	269.0476	321.4286	383.9286	446.4286	508.9286	571.4286	633.9286	696.4286
Block 2	$+\infty$	321.4286	383.9286	446.4286	508.9286	571.4286	633.9286	696.4286
Block 3	$+\infty$	$+\infty$	384.5238	447.0238	509.5238	572.0238	634.5238	697.0238
Block 4	$+\infty$	$+\infty$	$+\infty$	448.8095	511.3095	573.8095	636.3095	698.8095
Block 5	$+\infty$	$+\infty$	$+\infty$	$+\infty$	513.0952	575.5952	638.0952	700.5952
Block 6	$+\infty$	$+\infty$	$+\infty$	$+\infty$	$+\infty$	573.4127	633.9286	696.4286
Block 7	$+\infty$	$+\infty$	$+\infty$	$+\infty$	$+\infty$	$+\infty$	633.9286	696.4286

C. Problem Constraints

The problem constraints comprise of the model compatibility constraints defined by (9), the demand and supply constraints and the junction constraints. Time-varying speed limits affect the compatibility constraints for controlled links while RM affects the demand constraints for controlled on-ramps. In this section, we will first look at some rules and assumptions to reduce the number of model constraints and then present a MILP formulation for various constraints.

1) Constraint Reduction Rules and Assumptions: In theory, the model compatibility constraints for the Lax-Hopf approach are defined for all possible combinations of value condition blocks (9). However, many of these constraints are redundant as the $\mathbf{M}(\cdot,\cdot)$ values on the left hand side in (9) from some condition blocks are larger than values from other condition blocks for the same points in time-space and thus for any condition block on the right hand side, the constraints resulting from the condition blocks with the smallest $\mathbf{M}(\cdot,\cdot)$ value on the left hand side is the only one that matters and the rest can be omitted from the formulation without loss of exactness. Removing such redundant constraints helps in improving the efficiency of the MILP as it reduces the number of required binary variables.

In [38], the authors established a set of rules for omitting such redundant constraints by comparing the $\mathbf{M}(\cdot,\cdot)$ values resulting from different condition blocks at the link boundaries with the help of mathematical analysis, but they did so in a static speed limit setting. According to the analysis in [38], under static speed limits, the condition blocks resulting in the smallest $\mathbf{M}(\cdot,\cdot)$ values at any boundary points on the link at a given time are the ones formed most recently whose traffic wave reaches that point. Thus, we can ignore all other condition blocks that are older than such a block even though waves originating from those blocks still reach the said point. In this section, we see through an example, that the rules defined under static speed limit cannot be directly applied to the dynamic time-varying case, that is, it is not sufficient to just take the effect of the latest condition blocks affecting a point.

Example: We assume an arbitrary highway section of length 500 m fitted with variable speed limit signs without on-ramps or off-ramps. We simulate the traffic in this section using the Lax-Hopf approach with time-steps of 30 s. The speed limit is switched once after the fifth time-step, that is, after 150 s when it is changed from 30 m/s to 25 m/s. Other parameters of the fundamental diagram are: $w = -5$ m/s, and $\rho_m = 0.5$ veh/m. The link is assumed to be empty at the start of the simulation. The demand and supply for the section are taken as the maximum capacity at a free-flow speed of 30 m/s that is 2.1428 veh/s. In this example, we focus on the $\mathbf{M}(\cdot,\cdot)$ values at upstream boundary points right at the end of time-steps 4 to 11 resulting from the first to seventh downstream boundary condition blocks, each block covering one time-step. The exact values are presented in Table I. Our objective is to verify if the aforementioned idea from [38] will still hold under dynamic speed limits.

Discussion: Note, that a congestion traffic wave originating from a downstream condition block at the downstream link boundary will take 100 s or about 3.3 time-steps to reach the upstream end of the link. Thus, if the speed limits were static, then by the analysis in [38] Block 2 would have resulted in the smallest $\mathbf{M}(\cdot,\cdot)$ value on just one of the considered points on the upstream boundary and that is Point 5. For Point 6 to 11, the minimum $\mathbf{M}(\cdot,\cdot)$ values would have resulted from Block 3 to 6 respectively and we could have left out constraints (9) due to Block 2, Block 3 and so on at those points respectively. However, in Table I one can see that the minimum value at all these points is not because of the said blocks but instead due to Block 1 and 2 only (or just Block 2 which is mathematically capable of giving lower values than Block 1), and hence constraints due to these condition blocks at these points are not redundant. So if we want the exact solution to the LWR model, we cannot leave out the constraints due to Block 2 at these points. If one derives the mathematical expressions for $\mathbf{M}(\cdot,\cdot)$ values at those points from the given blocks, one will be able to see that this irregularity in values occurs due to the switch in speed limit at Point 5 which separates Point 6 to 11 in time from Block 2 to 5. This shows that it is not sufficient to just consider the latest condition block affecting a point when setting the constraints as earlier discussed.

This requires us to redefine the set of condition blocks that we need to consider to formulate the compatibility constraints. In general, when considering constraints at points lying after a switch, in addition to the latest condition block affecting the point, we also need to consider the latest condition blocks that affect the $\mathbf{M}(\cdot,\cdot)$ value at any point throughout the link space at the switch. This means that in the above example, for instance, we also need to consider Block 2 to 5 in addition to Block 6 when setting the constraints associated with Point 9 due to downstream boundary conditions. We can

leave out Block 1 here since the $\mathbf{M}(\cdot,\cdot)$ values due to it over the link space at the switch will be redundant. This logic of leaving out Block 1 also tells us that we do not need to consider the effect of the additional condition blocks beyond a certain point in time after the switch. While the analysis under static speed limits given in [38] can still be applied when comparing conditions within a single section within which the speed is indeed static, we can write certain rules similar to ones presented in [38] to compare $\mathbf{M}(\cdot,\cdot)$ values from conditions on opposite sides of a switch. These rules can be used to limit the number of constraints in the variable speed limit case. The following rules accompanied by the first assumption below allow us to restrict the number of constraints.

Lemma 1: Let a set of upstream boundary conditions be defined as in (11). Let us assume that $\mathbf{M}_{\gamma_j}(t',\chi) \leq \mathbf{M}_{\gamma_i}(t',\chi)$ for some $i < j$ where γ_i belongs to section p and γ_j belongs to section $p+1$, for some $t' > t_j + \frac{\chi-\xi}{v_{p+1}}$. We have that $\forall t > t'$, $\mathbf{M}_{\gamma_j}(t,\chi) \leq \mathbf{M}_{\gamma_i}(t,\chi)$.

Lemma 2: Let a set of downstream boundary conditions be defined as in (12). Let us assume that $\mathbf{M}_{\beta_j}(t',\xi) \leq \mathbf{M}_{\beta_i}(t',\xi)$ for some $i < j$ where β_i belongs to section p and β_j belongs to the section $p+1$, for some $t' > t_j + \frac{\xi-\chi}{w}$. We have that $\forall t > t'$, $\mathbf{M}_{\beta_j}(t,\xi) \leq \mathbf{M}_{\beta_i}(t,\xi)$.

Lemma 3: Let a set of upstream and downstream boundary conditions be defined as in (11) and (12). Let us assume that $\mathbf{M}_{\gamma_j}(t',\chi) \leq \mathbf{M}_{\beta_i}(t',\chi)$ for some $i < j$ where β_i belongs to section p and γ_j belongs to the section $p+1$, for some $t' > t_j + \frac{\chi-\xi}{v_{p+1}}$. We have that $\forall t > t'$, $\mathbf{M}_{\gamma_j}(t,\chi) \leq \mathbf{M}_{\beta_i}(t,\chi)$.

Lemma 4: Let a set of upstream and downstream boundary conditions be defined as in (11) and (12). Let us assume that $\mathbf{M}_{\beta_j}(t',\xi) \leq \mathbf{M}_{\gamma_i}(t',\xi)$ for some $i < j$ where γ_i belongs to section p and β_j belongs to the section $p+1$, for some $t' > t_j + \frac{\xi-\chi}{w}$. We have that $\forall t > t'$, $\mathbf{M}_{\beta_j}(t,\xi) \leq \mathbf{M}_{\gamma_i}(t,\xi)$.

Next, we define certain assumptions taken to limit the number of constraints, simplify the MILP formulations and ensure computational feasibility.

1) For any link in the network, $T_{sw(j+1)} - T_{sw(j)} \geq T \cdot \lceil \max(L/v_j T, -L/wT) \rceil$, $j = 1, \ldots, N_{sec} - 2$. This assumption helps restrict the constraints as it ensures that the $\mathbf{M}(\cdot,\cdot)$ values from any value condition block are relevant at most up to points in the section following the one containing that block and not beyond that. In that too we only need to check up to a maximum of $T \cdot \lceil -L/wT \rceil$ upstream points and $T \cdot \lceil L/v_{j+1}T \rceil$ downstream points in the next section.
2) For any link in the network, $v_j \geq L/T, \forall j \in [1, N_{sec}]$. This assumption ensures that any free-flow traffic wave originating from any condition block either covers the entire link space at the switch or exits the link before encountering a switch. Since in our formulation the free-flow speed is unknown, this assumption removes the need for further complex formulations than presented in this paper.
3) The speed limit can take values only from a discrete set of values that lie within a realistic range. This assumption allows us to transform the non-linear constraints into a mixed-integer form.

The next section discusses how the non-linear constraints can be written as MILP constraints with the help of transformations suggested in [41], [42].

2) Transformation of Constraints: Non-linearity in the optimization problem is introduced from model compatibility constraints, demand and supply constraints, junction constraints and the objective function. We address them one by one in this section.

Non-linearity

a) Model compatibility constraints: in these constraints is introduced due to two types of terms- $\rho_{c,j} v_j$ and $q_{in}(i)/v_j$ where $i \in \mathbb{N}$ represents the time-step and $j \in [1, N_{sec}]$ represents the time section. While these terms, in theory, are the product of continuous variables that cannot be handled by linear programs, our assumption regarding the discreteness of speed limit values helps to convert them into the products of constant values and binary variables, and that of continuous variables and binary variables respectively. These combinations can be expressed as MILP constraints by techniques provided in [41], [42]. Let the set of speed limits be $\{VSL_1\ VSL_2\ \ldots\ VSL_S\}$ where S is the number of possible speed limit values. Since the parameters w and ρ_m of the fundamental diagram are assumed to be constant and known throughout the time horizon, we can calculate the critical density corresponding to each value of speed limit in the set beforehand using the continuity of the fundamental diagram at critical density. Let the set of calculated possible critical density values be written as $\{k_{c_1}\ k_{c_2}\ \ldots\ k_{c_S}\}$. Then the term $\rho_{c,j} v_j$ can be linearized as follows

$$\rho_{c,j} v_j = \sum_{s=1}^{S} \delta_{s,j} k_{c_s} VSL_s, \quad \forall j \tag{23a}$$

$$\sum_{s=1}^{S} \delta_{s,j} = 1, \quad \forall j \tag{23b}$$

where $\delta_{1,j}, \delta_{2,j} \ldots, \delta_{S,j}$ are binary variables corresponding to section j. For the second type of non-linear term we use a different approach where we define a new continuous variable $k_{in,i}$ at each time-step corresponding to each $q_{in}(i)$ to store the value $q_{in}(i)/v_j$ by writing a set of linear constraints. Let the possible maximum flow values corresponding to the different speed limit values be denoted by $\{Q_1\ Q_2\ \ldots\ Q_S\}$. We use the same binary variables defined in (23). In addition, we define S new auxiliary continuous variables named $k_{a_1,i}, k_{a_2,i}, \ldots, k_{a_S,i}$ to store the values of $q_{in}(i)/v_j$ at the S possible values of v_j. Also let Q_{max} be the maximum flow value in the set of flows. Then the aforementioned linear constraints are given as follows for $s = 1, \ldots, S$:

$$0 \leq k_{a_s,i} \leq \delta_{s,j} \frac{Q_s}{VSL_s}, \quad \forall i, j \tag{24a}$$

$$\frac{q_{in}(i)}{VSL_s} - (1 - \delta_{s,j}) \frac{Q_{max}}{VSL_s} \leq k_{a_s,i} \leq \frac{q_{in}(i)}{VSL_s}, \quad \forall i, j \tag{24b}$$

$$k_{in,i} = \sum_{s=1}^{S} k_{a_s,i}, \quad \forall i \tag{24c}$$

Thus we can replace the non-linear terms $q_{in}(i)/v_j$ everywhere with the continuous variable $k_{in,i}$ making the constraints

linear. Thus, we can write all the model constraints in a linear form using auxiliary binary and continuous variables and some additional linear constraints as above.

b) Demand and Supply constraints: The demand of a link is defined as the maximum traffic that can flow out of the link if it was connected to a sink of infinite capacity while the supply of a link is the maximum traffic that can enter the link given that it is connected to a source of infinite capacity. This implies that in calculating the demand and supply of the link we want to take the maximum cumulative flow permitted by the model constraints. Let us, for instance, denote the maximum cumulative number of vehicles exiting the link up to a certain time-step by L. Let three model constraints be constraining L as follows

$$M_{c_1}(\cdot,\cdot) \geq L \tag{25a}$$
$$M_{c_2}(\cdot,\cdot) \geq L \tag{25b}$$
$$M_{c_3}(\cdot,\cdot) \geq L \tag{25c}$$

We want that L should be equal to the smallest of the three values on the left-hand side. This is ensured by introducing two new binary variables, say δ_1 and δ_2, and writing additional constraints as follows

$$M_{c_1}(\cdot,\cdot) \leq L + C\delta_1 + C\delta_2 \tag{26a}$$
$$M_{c_2}(\cdot,\cdot) \leq L + C(1-\delta_1) + C\delta_2 \tag{26b}$$
$$M_{c_3}(\cdot,\cdot) \leq L + C\delta_1 + C(1-\delta_2) \tag{26c}$$

where C is a value larger than any $\mathbf{M}(\cdot,\cdot)$ value. We also put a constraint on the sum of the binary variables as follows

$$\delta_1 + \delta_2 \leq 1 \tag{27}$$

Now, if say $\mathbf{M}_{c_2}(\cdot,\cdot)$ is the smallest of the three, then L should be equal to $\mathbf{M}_{c_2}(\cdot,\cdot)$, which is ensured by setting $\delta_1 = 1$ and $\delta_2 = 0$. Any other combination of values of the binary variables would result in the violation of at least one of the constraints and thus will not be selected by the solver. In general, if there are N number of constraints then we need to define $\lceil \log_2(N) \rceil$ number of binary variables and write the constraints in a similar manner as above.

From L we can easily calculate the actual demand values in each time-step using the following expression. Let L_n be the cumulative traffic up to and including time-step n and L_0 be the value at the start of the horizon, then the demand $D(n)$ for time-step n can be calculated as follows

$$D_n = \frac{L_n - L_0}{nT} - \sum_{i=1}^{n-1} q_{out}(i) \tag{28}$$

Similarly, we can also calculate the supply $S(n)$ for any link over any time-step n using the variables for cumulative traffic entering the link and that for inflow.

Ramp Metering Constraints: Ramp metering allows us to pose an additional constraint on the demand of a controlled on-ramp by controlling the total vehicles that can leave the ramp at any time. Apart from the demand and supply constraints discussed above, a controlled on-ramp has an additional constraint on the maximum number of vehicles that can leave the on-ramp up to any time-step n (same as L defined above) and therefore, on the demand of the on-ramp during any time-step n. This additional ramp metering constraint is imposed with the help of a new ramp metering rate variable $r \in [0, 1]$ such that

$$T \cdot \sum_{i=1}^{n-1} q_{out}(i) + r \cdot (\rho_{c,j} v_j T) \geq L \tag{29}$$

where j is the section (based on speed limit switch times) to which the time-step n belongs. Here, the first term on the left side of the inequality is the total number of vehicles leaving the ramp up to time-step $n-1$. In the second term, r essentially represents the fraction of the maximum possible number of vehicles that are allowed to leave that on-ramp during time-step n. Note that in calculating the maximum possible vehicles for this time-step we use v_j and $\rho_{c,j}$ which are the speed and critical density corresponding to section j to which the time-step n belongs. In our formulation, a ramp metering rate variable is defined per time-step per controlled on-ramp. Again, we also provide a lower bound constraint (similar to (26)) corresponding to (29) for all n, to ensure that we take the minimum of the possible demands for the on-ramp.

c) Junction constraints: Junction models are based on the conservation of vehicles, that is, the number of vehicles entering a junction is the same as the number of vehicles leaving it. The allocation of incoming flow on to the exit links is done using an allocation matrix. Since junctions have no storage capacity the sum of all the allocation parameters from the incoming options among all the outgoing options must be equal to one. Also, the flows must obey the demand and supply constraints imposed at the boundary of the incoming and outgoing links respectively. Traditionally, we maximize the flow through the junctions while imposing the aforementioned constraints [53]. For instance, let there be a diverge junction with one incoming link having a demand D, and two outgoing links having supplies S_1 and S_2, and the allocation matrix is $\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$. For any time-step, if the in-flow from the incoming link is denoted by q_{in} and the out-flow into the outgoing links is denoted by q_{out1} and q_{out2}, then these flows obey the following constraints:

$$\begin{aligned} &q_{out1} = a_1 q_{in} \\ &q_{out2} = a_2 q_{in} \\ &q_{in} = \min(D, S_1/a_1, S_2/a_2) \\ &a_1 + a_2 = 1 \\ &a_1, a_2 \in [0, 1] \end{aligned} \tag{30}$$

Here the $\min(\cdot)$ function can be expressed using mixed integer linear inequalities similar to (25) and (26).

In the case of a merge junction, we also need to define priority rules to allocate the supply of the output links to different input links. For instance, if there are two incoming links having demands D_1 and D_2, and one outgoing link having supply S, then we can define two non-negative parameters α_1 and α_2 such that $\alpha_1 + \alpha_2 = 1$ and the supply corresponding to the two incoming links are $\alpha_1 S$ and $\alpha_2 S$ respectively. If the

inflow from the links is q_{in1} and q_{in2} and the outflow from the junction is q_{out} then these flows must obey the following constraints similar to [54]:

$$\begin{aligned} q_{out} &= q_{in1} + q_{in2} \\ q_{in1} &= \min(D_1, \alpha_1 S) \\ q_{in2} &= \min(D_2, \alpha_2 S) \\ q_{in1}, q_{in2} &\geq 0 \end{aligned} \tag{31}$$

Again, the $\min(\cdot)$ functions can be written in a mixed-integer linear form as mentioned before.

d) Objective function constraints: The objective function J_{TJD} contains the term $\max\left(\mathbf{M}(t_{\max}, \xi) - \mathbf{M}(t_{\max}, \chi)\right)/L, k_{c,max}\big)$ for each link which introduces the non-linearity. We can write it with the help of a new continuous variable, say t, and some more linear constraints as follows

$$\begin{aligned} &\min \ t \\ &\text{subject to } t \geq \left(\mathbf{M}(t_{\max}, \xi) - \mathbf{M}(t_{\max}, \chi)\right)/L \\ &\qquad\qquad t \geq k_{c,max} \\ &\qquad\qquad t \in \mathbb{R} \end{aligned} \tag{32}$$

For the minimization of the penalty terms, we can use a simple trick to write the absolute function in a linear form using a new continuous variable, say z, as follows

$$\begin{aligned} &\min \ \mathbf{1}^T z \\ &\text{subject to } -z_i \leq v_i - v_{i+1} \\ &\qquad\qquad -z_i \leq v_{i+1} - v_i \\ &\qquad\qquad i = 1, \ldots, N_{sec} - 1 \\ &\qquad\qquad z \in \mathbb{R}^{Nsec-1} \end{aligned} \tag{33}$$

which makes z_i essentially equal to $\mid v_i - v_{i+1} \mid$. The ramp metering penalty can also be incorporated in the same way.

The above will further require us to write v_i in terms of a combination of binary variables and the known discrete values, similar to (23). The next section summarizes the different variables used in the overall formulation.

D. Decision Variables

From the above discussion it is clear that the given optimization problem contains both continuous and binary variables.

1) Continuous Variables: These include the in-flow and out-flow variables, demand and supply variables, linear transformation variables for q_{in}/v (in model constraints), and the slack variables for the J_{TAC} objective and penalty terms.

2) Binary Variables: These include the auxiliary variables required for selecting the speed limit for each time section, variables for calculating the demand and supply for each link at each time-step based on model constraints and those required for setting the actual in-flow and out-flow equal to the minimum of the demand and supply based on the junction model.

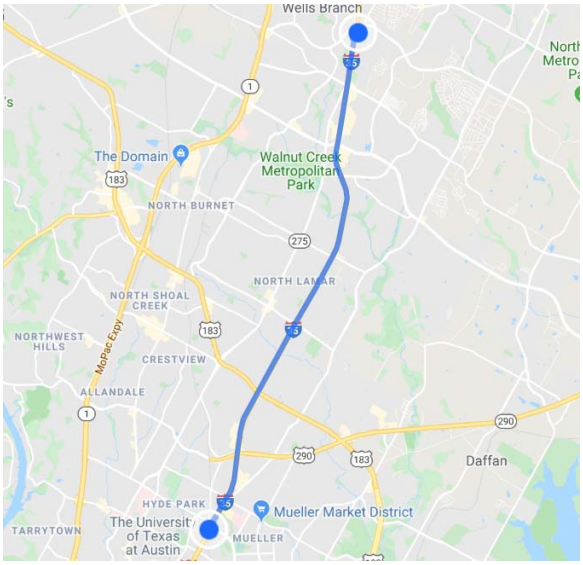

Fig. 4. Google maps illustration of the Interstate 35 highway in Austin, Texas. The 10.1 miles stretch considered for the implementation example in Section V has been highlighted in blue color.

V. Traffic Control on a Highway Network

In this section, we implement the optimization framework on a real highway network comprising of a stretch of a highway with on-ramps and off-ramps. The simulation model used in this case is the same model as used in the optimization that is the LWR model solved using the Lax-Hopf approach.

A. Example Setup

We consider a 10.1 miles (about 16.3 km) stretch of the Interstate-35 highway passing through Austin, Texas, highlighted in the Google Maps illustration in Fig. 4. This stretch has 8 on-ramps and 8 off-ramps. We divide the main highway into 18 links of different lengths, while each ramp is considered as one link, making a total of 34 links. The length of the highway links in kilometers, ordered from upstream to downstream end are 0.98, 0.48, 0.75, 0.37, 0.48, 0.82, 0.67, 1.30, 1.12, 1.28, 1.23, 1.88, 0.53, 1.12, 1.11, 0.78, 0.98, and 0.35 km. Let the links be numbered 1 to 18 in the specified order. The ramps are placed at the junction of the links, on-ramps located after link 2, 4, 6, 8, 10, 11, 13, and 16, and off-ramps located after link 1, 3, 5, 7, 9, 12, 15, and 17. Let the ramps be numbered 1 to 8 in the specified order for both on- and off-ramps. There are a total of 8 merge junctions, 8 diverge junctions and 1 one-to-one junction (between link 14 and 15). We assume that links 15 and 16 of the highway are equipped with VSL signs and on-ramps 6 and 7 are equipped with ramp meters. The allocation and priority matrix for junctions are set arbitrarily. The split ratios are taken as 10% for off-ramps 1, 2, 3, 5, 6 and 8, and 20% for off-ramps 4 and 7. The allocated flow from on-ramps 1 to 7 onto the highway is kept as 10%, while that from on-ramp 8 is kept as 20%. The demands for the various input links is chosen arbitrarily and is kept varying in a reasonable range of values to create a realistic scenario. The demands are presented in Table II. The supply for the

TABLE II
DEMANDS FOR HIGHWAY INPUT LINK AND ON-RAMPS FOR EXAMPLE 1 (VEH/S)

Timing (s)	Highway	Ramp 1 & 2	Ramp 3 & 4	Ramp 5	Ramp 6 & 7	Ramp 8
0-600	1.8	0.25	0.28	0.2	0.25	0.3
600-1200	1.9	0.28	0.3	0.24	0.26	0.28
1200-1800	2.0	0.28	0.28	0.25	0.25	0.26
1800-2400	2.1	0.3	0.3	0.25	0.28	0.28
2400-3000	2.0	0.29	0.28	0.24	0.26	0.3
3000-3600	1.9	0.25	0.26	0.22	0.28	0.3

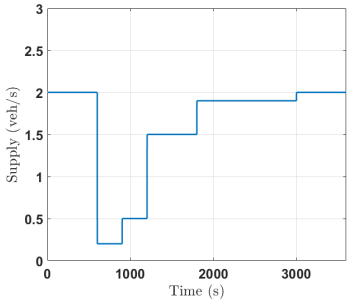

Fig. 5. Supply (veh/s) at downstream end of the highway.

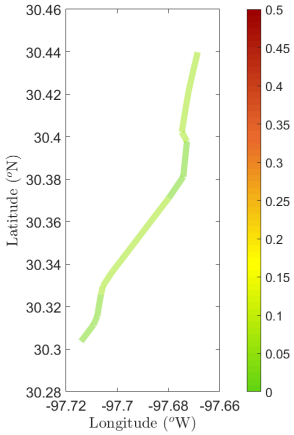

Fig. 6. Initial density condition (veh/m) for highway network example.

off-ramps is assumed to be unrestricted. The supply for the output section of the highway is depicted in Fig 5. We assume the occurrence of an incident, for instance an accident, ahead of the downstream end of the considered highway stretch at around 10 min from the start of the simulation which results in the drop of supply to 0.2 veh/s. The supply slowly recovers over time however the incident creates a jam wave that travels upstream from the downstream end of the highway. The initial density on the highway at the start of the optimization process is as shown in Fig. 6.

The fundamental diagram parameters values for highway links with no VSL are chosen as: $v = 33$ m/s, $w = -5$ m/s, $\rho_m = 0.5$ veh/m. The set of possible speed limits for the VSL links is {19 25 33} m/s. ρ_m for ramp links is 0.125 veh/m. The objective function used in this optimization is J_{TAC} (20) along with VSL (21) and RM (22) penalties with a weight

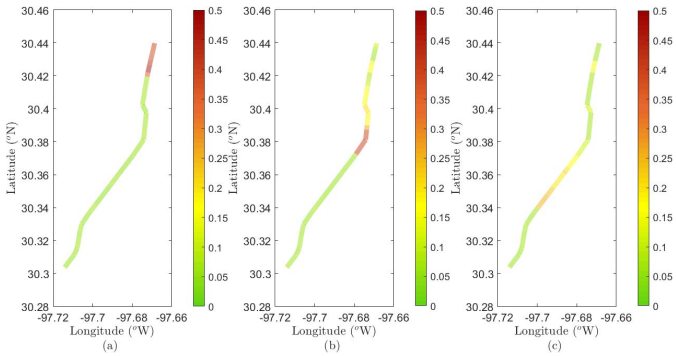

Fig. 7. Evolution of traffic density (veh/m) without control. The three figures depict conditions after (a) 20 minutes, (b) 40 minutes, and (c) 60 minutes.

of 10 for both. A penalty for unsatisfied demand is included in the objective similar to the second term in (19). The weight for the same is selected as the demand multiplied by a scaling coefficient, in this case 0.1. Thus, the penalty weight increases as more demand is unmet resulting in larger queues at the entrances of the input links and on-ramps. The time-step value T is taken as 60 s. The prediction horizon is set to 10-time-steps that is $10 \times T = 600$ s or 10 min. In the optimization setup, a change in speed limit is allowed once at 5 min from the start of the prediction horizon. The control horizon is kept as one time-step T. This means that the control signal for the first one time-step, or the duration of the control horizon, as obtained from the solver is applied to the system in real-time and the corresponding initial condition at the end of the control horizon, is again fed to the solver to perform the optimization to obtain optimal control signals for the next prediction horizon. Again, the control for the first time-step is applied to the system and so on.

B. Results and Discussion

Fig. 7 presents the uncontrolled scenario at three different time stamps during the simulation namely 20 min, 40 min and 60 min. It can be seen that a jam forms at the downstream end of the highway corresponding to the incident that took place ahead of the stretch end. As the supply at the downstream end restores, the jam starts to dissolve near that end as the vehicles at the downstream front of the jam are able to flow freely again. However, the congestion that has already formed results in a congestion wave that travels upstream on the highway and

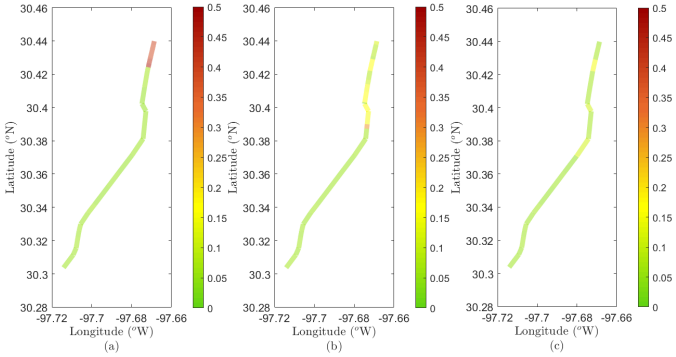

Fig. 8. Evolution of traffic density (veh/m) with VSL and RM control. The three figures depict conditions after (a) 20 minutes, (b) 40 minutes, and (c) 60 minutes.

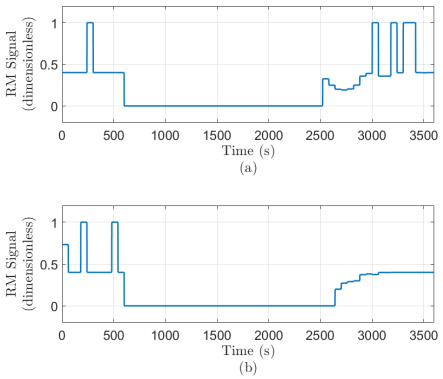

Fig. 10. Optimal ramp metering control inputs (dimensionless). (a) on-ramp 6 and (b) on-ramp 7.

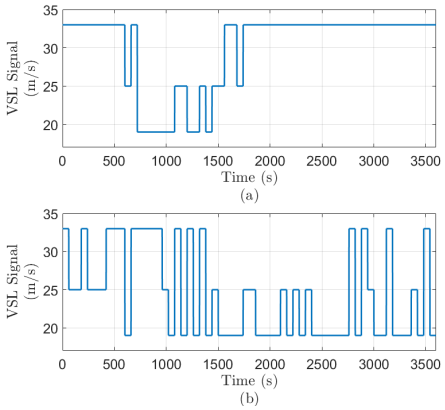

Fig. 9. Optimal variable speed limit control inputs (m/s). (a) link 15, and (b) link 16 of the highway.

continues to persist even after 50 minutes past the occurrence of the incident.

We apply our control formulation to the system under the same conditions of demands and supply as the uncontrolled case. Fig. 8 shows the controlled scenario again at the same time stamps as before. One can observe that in presence of control, the jam dissolves much quicker and almost disappears within around 50 min of its formation, as can be seen in Fig. 8(c). At the end of the 60 min period the value of J_{TAC} for the uncontrolled case is 2549.1 and that for the controlled case is 1945.4 which is an improvement of about 23.7% for the controlled case over the uncontrolled case.

The obtained control signals for speed limit and ramp meters are given in Fig. 9 and Fig. 10 respectively. Note that in this example a metering rate value of around 0.4 is equivalent to a value of 1, that is both values will give the same solution. This is because the metering rate is a fraction of the maximum flow based on fundamental diagram, but in the current scenario, given the percentage of highway supply allocated to the on-ramps, the possible flow from the ramps is already not higher than about 0.4 of the maximum flow based on fundamental diagram.

As expected from the controller, the speed limits are reduced corresponding to the formation of the jam (Fig. 9) in order to reduce the flow moving towards the jam to help the jam dissolve faster. Of course we expect this reduction to only be to the extent that another jam does not form due to the control on or behind the VSL controlled links. The ramp controls also restrict the flow of the controlled on-ramps corresponding to the formation of the jam. As observed in Fig. 10, the permitted on-ramp flows eventually restores to normal as the jam dissolves. The restricted flow from the on-ramps not only reduces the number of vehicles joining and thus worsening the jam on the highway, but it also provides more space for the out-flow of vehicles already in the jam thus helping the jam dissolve faster. This does result in a jammed state on the controlled on-ramps for a while but it helps to dissolve the jam on the highway which would otherwise have been bad for both the travellers coming from the on-ramps as well as those on the highway.

The problem consists of 1888 continuous variables, 1383 binary variables, and 9064 constraints, and took an average time of 6.47 s per run of the solver using CPLEX on a Windows computer with processor Intel Core i7-8750H CPU@2.2GHz. The relatively small running time of the optimization solver for a reasonably large 10.1 mile stretch of highway network with several ramps, and its effectiveness in reducing congestion on highways, as illustrated in the example, makes it attractive for use in real-world network control problems.

C. Relaxation Into a Linear Program

In this section we want to check if the relaxation of the MILP into a corresponding LP (by allowing all Boolean variables to continuously vary between 0 and 1) is tight. We test the effect of this relaxation on the example considered in this section. The optimal value of the relaxed problem differs from the originally computed optimal value, which shows that these integer constraints are necessary in general. The relaxed problem also results in theoretically incorrect (non-physical) values of other decision variables such as critical density variables, demand and supply variables and junction flow variables. Though an unrelaxed mixed-integer program generally demands more computational power than a relaxed linear program, we have seen that we are still able to get a decent performance in solving this problem for a reasonably large highway system with ordinary computational power.

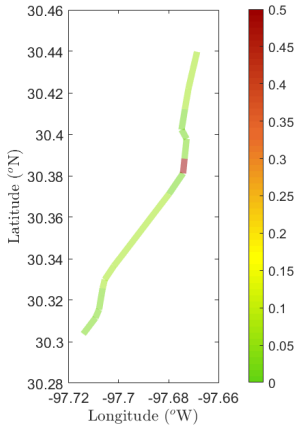

Fig. 11. Initial Density conditions (veh/m) for example comparing Lax-Hopf method with LTM.

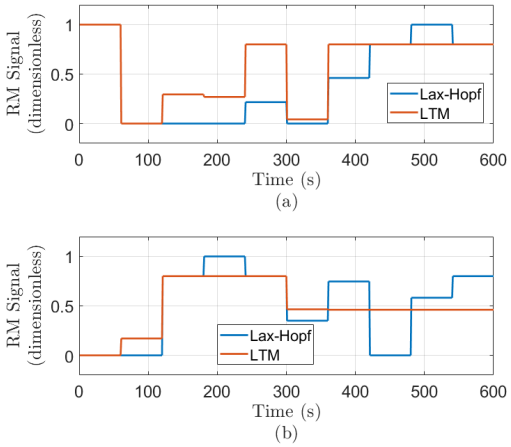

Fig. 12. Optimal ramp metering control inputs (dimensionless) for example comparing Lax-Hopf method with LTM. (a) on-ramp 6 and (b) on-ramp 7.

VI. Comparison With the Link Transmission Model

In this section, we discuss the difference between the LTM and the Lax-Hopf method in terms of implementation and compare the control input obtained from the optimization formulation using both methods for a specific type of initial density profile setting on the same network as in Section V-A. We also compare the general running time of the optimization obtained over several runs under different settings while applying both methods.

The major point of difference between the LTM [32] and the Lax-Hopf method [35], [36] is that the LTM does not account for expansion waves in traffic while the latter does. This difference between the two methods manifests in terms of different solutions to the LWR model whenever there is an expansion wave in the traffic stream, such as when a congested initial density block is followed by a free-flow density block. As a result, in the presence of expansion waves, the predicted traffic flows using both methods will be different. A difference in the predicted traffic evolution can result in different optimal control signals being found for the same initial condition and demand/supply settings. The same is illustrated in the ensuing example.

We use the same configuration of parameters as in Section V-A except for the allocated flow percentage for ramps 6 and 7 onto the highway which is set as 20%. The demand and supply profile are kept the same as in Section V. We put a different initial density profile on link 12 from that in Fig. 6, that is, we set an initial density of 0.5 veh/m for the first 800 m of the link and 0.02 veh/m from 800 m to the end of the link. The new initial density profile for the highway is depicted in Fig. 11. The objective function is J_{TAC} along with the input penalty term with the penalty weight as described in the Section V-A. We run the optimization with the given initial conditions for a prediction horizon of 600 s. The optimal solution from the two methods particularly differs in the ramp metering controls which are presented in Fig. 12. As can be observed, the optimal control signal from both methods is different. The value of the objective function as obtained by applying the control generated from both cases to the simulation is 1759.3 in the case of LTM and 1730.4 in the case of Lax-Hopf. This shows that the optimal control generated for this particular case using LTM is only sub-optimal in reality. Note that in this example, results differ because of the presence of an expansion wave in the initial conditions of link 12. Initial conditions for which the density increases over space (no expansion wave) lead to identical results for the proposed approach and the LTM approach. The running time for LTM based optimization problem was 1.46 s while that for Lax-Hopf based problem was 1.87 s.

In our formulation using Lax-Hopf method, the expansion waves are considered originating from the initial density condition blocks as well as forming in the middle of the time horizon due to switches in speed limits (Section III-B) as we assume the formation of temporary density condition blocks. Since LTM does not consider expansion waves in its formulation, it reduces the number of constraints at both these occasions. This does have an impact on the running time, however the said impact as observed from several runs of the solver under different conditions of initial density and demand/supply is not very large. The running time in general lies in the same order of magnitude. For instance, we ran both LTM and Lax-Hopf based optimizations under different conditions of initial density and demand/supply to get a mean running time of 2.34 s for LTM with standard deviation 0.43 s and that of 3.89 s for Lax-Hopf with standard deviation of 1.57 s.

VII. Conclusion

In this paper, an analytical formulation of the variable speed limit control problem applicable to the LWR PDE is developed. The analytical formulation is based on the classical Lax-Hopf formulas. Rules are developed for limiting the number of solution steps required in the Lax-Hopf scheme, therefore, reducing the computational time. The analytical scheme is further used to formulate an optimization problem to apply variable speed limit and ramp metering control of traffic on highway networks. Example numerical simulations are presented to validate the method on both a single highway section as well as an over 10 miles sub network of the Interstate-35 highway running through Austin, Texas. The

control inputs generated from the solver are shown to dissipate a traffic jam on the highway faster than the uncontrolled case. We also compare our method to another popular method, based on the Link Transmission Model (LTM), which provides a fast, approximate solution to the LWR PDE. In that, we see that in some cases (expansion wave at the initial time), the control input generated from the optimization problem using LTM are sub-optimal as compared to the results generated from the Lax-Hopf based optimization. In this paper, we also test the influence of relaxing the binary variables used in the problem. The tests show a significant optimality gap that can lead to inaccurate, non-physical results when doing so.

Future work will focus on robust traffic control by allowing for uncertainty in the demand and supply variables. The analytical framework can also be used to perform other types of control such as through route assignment. The proposed mathematical formulation can also serve as a stepping-stone to develop optimization schemes using second order models such as in [55] or LWR model incorporating bounded acceleration as in [56] which are more realistic than the LWR model.

APPENDIX
ANALYTICAL SOLUTIONS TO PIECEWISE AFFINE INITIAL, UPSTREAM AND DOWNSTREAM BOUNDARY CONDITIONS UNDER STATIC SPEED LIMITS

Given the piecewise affine initial, upstream and downstream boundary conditions, (10) to (12), defined in Section III-A, the corresponding solutions $\mathbf{M}_{M_k}(\cdot,\cdot)$, $\mathbf{M}_{\gamma_n}(\cdot,\cdot)$, $\mathbf{M}_{\beta_n}(\cdot,\cdot)$ at any point on the link space at a time before the first change in speed limit occurs from the time of start of the condition block is given [37], [52] by the following formulae:

$$\mathbf{M}_{M_k}(t,x) = \begin{cases} +\infty & \text{if } x \le x_k + wt \\ & \text{or } x \ge x_{k+1} + vt \\ -\sum_{i=1}^{k-1} \rho_{\text{ini}}(i)(x_{i+1}-x_i) & \\ +\rho_{\text{ini}}(k)(tv + x_k - x) & \text{if } x_k + tv \le x \\ & \text{and } x_{k+1} + tv \ge x \\ & \text{and } \rho_{\text{ini}}(k) \le \rho_c \\ -\sum_{i=1}^{k-1} \rho_{\text{ini}}(i)(x_{i+1}-x_i) & \\ +\rho_c(tv + x_k - x) & \text{if } x_k + tv \ge x \\ & \text{and } x_k + tw \le x \\ & \text{and } \rho_{\text{ini}}(k) \le \rho_c \\ -\sum_{i=1}^{k-1} \rho_{\text{ini}}(i)(x_{i+1}-x_i) & \\ +\rho_{\text{ini}}(k)(tw + x_k - x) & \\ -\rho_m tw & \text{if } x_k + tw \le x \\ & \text{and } x_{k+1} + tw \ge x \\ & \text{and } \rho_{\text{ini}}(k) \ge \rho_c \\ -\sum_{i=1}^{k} \rho_{\text{ini}}(i)(x_{i+1}-x_i) & \\ +\rho_c(tw + x_{k+1} - x) & \\ -\rho_m tw & \text{if } x_{k+1} + tv \ge x \\ & \text{and } x_{k+1} + tw \le x \\ & \text{and } \rho_{\text{ini}}(k) \ge \rho_c \end{cases} \tag{34}$$

$$\mathbf{M}_{\gamma_n}(t,x) = \begin{cases} +\infty & \text{if } t \le (n-1)T + \frac{x-\xi}{v} \\ \sum_{i=1}^{n-1} q_{\text{in}}(i)T & \\ +q_{\text{in}}(n)(t - \frac{x-\xi}{v} - (n-1)T) & \text{if } (n-1)T + \frac{x-\xi}{v} \le t \\ & \text{and } t \le nT + \frac{x-\xi}{v} \\ \sum_{i=1}^{n} q_{\text{in}}(i)T & \\ +\rho_c v(t - nT - \frac{x-\xi}{v}) & \text{otherwise} \end{cases} \tag{35}$$

$$\mathbf{M}_{\beta_n}(t,x) = \begin{cases} +\infty & \text{if } t \le nT + \frac{x-\chi}{w} \\ -\sum_{k=1}^{k_{max}} \rho_{\text{ini}}(k)(x_{k+1}-x_k) & \\ +\sum_{i=1}^{n-1} q_{\text{out}}(i)T & \\ +q_{\text{out}}(n)(t - \frac{x-\chi}{w} - (n-1)T) & \\ -\rho_m(x-\chi) & \text{if } (n-1)T + \frac{x-\chi}{w} \le t \\ & \text{and } t \le nT + \frac{x-\chi}{w} \\ -\sum_{k=1}^{k_{max}} \rho_{ini}(k)(x_{k+1}-x_k) & \\ +\sum_{i=1}^{n} q_{\text{out}}(i)T & \\ +\rho_c v(t - nT - \frac{x-\chi}{v}) & \text{otherwise} \end{cases} \tag{36}$$

Readers are referred to [39] for various properties associated with the affine condition blocks, their analytical solutions, and the model constraints.

REFERENCES

[1] M. Papageorgiou and A. Kotsialos, "Freeway ramp metering: An overview," *IEEE Trans. Intell. Transp. Syst.*, vol. 3, no. 4, pp. 271–281, Dec. 2002.

[2] A. Hegyi, B. DeSchutter, and J. Hellendoorn, "Optimal coordination of variable speed limits to suppress shock waves," *IEEE Trans. Intell. Transp. Syst.*, vol. 6, no. 1, pp. 102–112, Mar. 2005.

[3] G. Gomes and R. Horowitz, "Optimal freeway ramp metering using the asymmetric cell transmission model," *Transp. Res. C, Emerg. Technol.*, vol. 14, no. 4, pp. 244–262, Aug. 2006.

[4] M. Papageorgiou, E. Kosmatopoulos, and I. Papamichail, "Effects of variable speed limits on motorway traffic flow," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2047, no. 1, pp. 37–48, Jan. 2008.

[5] R. C. Carlson, I. Papamichail, M. Papageorgiou, and A. Messmer, "Optimal motorway traffic flow control involving variable speed limits and ramp metering," *Transp. Sci.*, vol. 44, no. 2, pp. 238–253, May 2010.

[6] R. C. Carlson, I. Papamichail, and M. Papageorgiou, "Integrated feedback ramp metering and mainstream traffic flow control on motorways using variable speed limits," *Transp. Res. C, Emerg. Technol.*, vol. 46, pp. 209–221, Sep. 2014.

[7] Y. Wang, E. B. Kosmatopoulos, M. Papageorgiou, and I. Papamichail, "Local ramp metering in the presence of a distant downstream bottleneck: Theoretical analysis and simulation study," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 5, pp. 2024–2039, Oct. 2014.

[8] P. Goatin, S. Göttlich, and O. Kolb, "Speed limit and ramp meter control for traffic flow networks," *Eng. Optim.*, vol. 48, no. 7, pp. 1121–1144, Jul. 2016.

[9] Y. Han, M. Ramezani, A. Hegyi, Y. Yuan, and S. Hoogendoorn, "Hierarchical ramp metering in freeways: An aggregated modeling and control approach," *Transp. Res. C, Emerg. Technol.*, vol. 110, pp. 1–19, Jan. 2020.

[10] I. Karafyllis and M. Papageorgiou, "Feedback control of scalar conservation laws with application to density control in freeways by means of variable speed limits," *Automatica*, vol. 105, pp. 228–236, Jul. 2019.

[11] X. Yang, Y. C. Lu, and Y. Lin, "Optimal variable speed limit control system for freeway work zone operations," *J. Comput. Civil Eng.*, vol. 31, no. 1, Jan. 2017, Art. no. 04016044.

[12] Y. Zhang and P. A. Ioannou, "Combined variable speed limit and lane change control for highway traffic," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 7, pp. 1812–1823, Jul. 2017.

[13] Y. Han, D. Chen, and S. Ahn, "Variable speed limit control at fixed freeway bottlenecks using connected vehicles," *Transp. Res. B, Methodol.*, vol. 98, pp. 113–134, Apr. 2017.

[14] Z. Li, P. Liu, C. Xu, H. Duan, and W. Wang, "Reinforcement learning-based variable speed limit control strategy to reduce traffic congestion at freeway recurrent bottlenecks," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 11, pp. 3204–3217, Nov. 2017.

[15] M. Hajiahmadi *et al.*, "Integrated predictive control of freeway networks using the extended link transmission model," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 1, pp. 65–78, Jan. 2016.

[16] I. Papamichail, A. Kotsialos, I. Margonis, and M. Papageorgiou, "Coordinated ramp metering for freeway networks—A model-predictive hierarchical control approach," *Transp. Res. C, Emerg. Technol.*, vol. 18, no. 3, pp. 311–331, Jun. 2010.

[17] J. R. D. Frejo and E. F. Camacho, "Global versus local MPC algorithms in freeway traffic control with ramp metering and variable speed limits," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 4, pp. 1556–1565, Dec. 2012.

[18] Y. Han, A. Hegyi, Y. Yuan, S. Hoogendoorn, M. Papageorgiou, and C. Roncoli, "Resolving freeway jam waves by discrete first-order model-based predictive control of variable speed limits," *Transp. Res. C, Emerg. Technol.*, vol. 77, pp. 405–420, Apr. 2017.

[19] G. S. van de Weg, A. Hegyi, S. P. Hoogendoorn, and B. De Schutter, "Efficient freeway MPC by parameterization of ALINEA and a speed-limited area," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 1, pp. 16–29, Jan. 2019.

[20] J. R. D. Frejo and E. F. Camacho, "Feasible cooperation based model predictive control for freeway traffic systems," in *Proc. 50th IEEE Conf. Decis. Control Eur. Control Conf.*, Dec. 2011, pp. 5965–5970.

[21] Y. Wang, M. Papageorgiou, G. Sarros, and W. J. Knibbe, "Feedback route guidance applied to a large-scale express ring road," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1965, no. 1, pp. 79–88, Jan. 2006.

[22] A. Hegyi, S. P. Hoogendoorn, M. Schreuder, H. Stoelhorst, and F. Viti, "SPECIALIST: A dynamic speed limit control algorithm based on shock wave theory," in *Proc. 11th Int. IEEE Conf. Intell. Transp. Syst.*, Oct. 2008, pp. 827–832.

[23] Y. Li, A. H. F. Chow, and D. L. Cassel, "Optimal control of motorways by ramp metering, variable speed limits, and hard-shoulder running," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2470, no. 1, pp. 122–130, Jan. 2014.

[24] A. Tympakianaki, A. Spiliopoulou, A. Kouvelas, I. Papamichail, M. Papageorgiou, and Y. Wang, "Real-time merging traffic control for throughput maximization at motorway work zones," *Transp. Res. C, Emerg. Technol.*, vol. 44, pp. 242–252, Jul. 2014.

[25] Y. Li, C. Tang, K. Li, X. He, S. Peeta, and Y. Wang, "Consensus-based cooperative control for multi-platoon under the connected vehicles environment," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 6, pp. 2220–2229, Jun. 2019.

[26] S. Jafari and K. Savla, "A decentralized optimal feedback flow control approach for transport networks," 2018, *arXiv:1805.11271*. [Online]. Available: http://arxiv.org/abs/1805.11271

[27] Y. Li *et al.*, "Efficient robust control of first order scalar conservation laws using semi-analytical solutions," *Discrete Continuous Dyn. Syst., S*, vol. 7, no. 3, pp. 525–542, 2014.

[28] J. M. Maciejowski, *Predictive Control: With Constraints*. London, U.K.: Pearson, 2002.

[29] A. Muralidharan and R. Horowitz, "Computationally efficient model predictive control of freeway networks," *Transp. Res. C, Emerg. Technol.*, vol. 58, pp. 532–553, Sep. 2015.

[30] C. F. Daganzo, "The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory," *Transp. Res. B, Methodol.*, vol. 28, no. 4, pp. 269–287, Aug. 1994.

[31] A. Kotsialos, M. Papageorgiou, C. Diakaki, Y. Pavlis, and F. Middelham, "Traffic flow modeling of large-scale motorway networks using the macroscopic modeling tool METANET," *IEEE Trans. Intell. Transp. Syst.*, vol. 3, no. 4, pp. 282–292, Dec. 2002.

[32] I. Yperman, "The link transmission model for dynamic network loading," Ph.D. dissertation, Dept. Civil Eng., KU Leuven, Leuven, Belgium, 2007.

[33] M. J. Lighthill and G. B. Whitham, "On kinematic waves II. A theory of traffic flow on long crowded roads," *Proc. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 229, no. 1178, pp. 317–345, 1955.

[34] P. I. Richards, "Shock waves on the highway," *Oper. Res.*, vol. 4, no. 1, pp. 42–51, Feb. 1956.

[35] C. G. Claudel and A. M. Bayen, "Lax–Hopf based incorporation of internal boundary conditions into Hamilton–Jacobi equation. Part I: Theory," *IEEE Trans. Autom. Control*, vol. 55, no. 5, pp. 1142–1157, May 2010.

[36] C. G. Claudel and A. M. Bayen, "Lax–Hopf based incorporation of internal boundary conditions into Hamilton–Jacobi equation. Part II: Computational methods," *IEEE Trans. Autom. Control*, vol. 55, no. 5, pp. 1158–1174, May 2010.

[37] P. E. Mazare, A. Dehwah, C. G. Claudel, and A. M. Bayen, "Analytical and grid-free solutions to the Lighthill–Whitham–Richards traffic flow model," *Transp. Res. B, Anal.*, vol. 45, no. 10, pp. 1727–1748, 2011.

[38] M. Simoni and C. Claudel, "A fast Lax–Hopf formula to solve the Lighthill-Whitham-Richards traffic flow model on networks," 2018, *arXiv:1802.05391*. [Online]. Available: http://arxiv.org/abs/1802.05391 and http://search.proquest.com/docview/2071560832/

[39] E. S. Canepa and C. G. Claudel, "Networked traffic state estimation involving mixed fixed-mobile sensor data using Hamilton–Jacobi equations," *Transp. Res. B, Methodol.*, vol. 104, pp. 686–709, Oct. 2017.

[40] B. Katz, J. Ma, H. Rigdon, K. Sykes, Z. Huang, and K. Raboy, "Synthesis of variable speed limit signs," Leidos ToXel, LLC FHWA, U.S. Dept. Transp., Washington, DC, USA, Tech. Rep. FHWA-HOP-17-003, 2017. [Online]. Available: https://ops.fhwa.dot.gov/publications/fhwahop17003/fhwahop17003.pdf

[41] F. Glover, "Improved linear integer programming formulations of nonlinear integer problems," *Manage. Sci.*, vol. 22, no. 4, pp. 455–460, Dec. 1975.

[42] A. Bemporad and M. Morari, "Control of systems integrating logic, dynamics, and constraints," *Automatica*, vol. 35, no. 3, pp. 407–427, Mar. 1999.

[43] C. F. Daganzo, "A variational formulation of kinematic waves: Basic theory and complex boundary conditions," *Transp. Res. B, Methodol.*, vol. 39, no. 2, pp. 187–196, Feb. 2005.

[44] C. F. Daganzo, "On the variational theory of traffic flow: Well-posedness, duality and applications," *Netw. Heterogeneous Media*, vol. 1, no. 4, pp. 601–619, 2006.

[45] K. Moskowitz, "Discussion of 'freeway level of service as influenced by volume and capacity characteristics' by D. R. Drew and C. J. Keese," *Highway Res. Rec.*, vol. 99, pp. 43–44, Jan. 1965.

[46] G. F. Newell, "A simplified theory of kinematic waves in highway traffic, Part (I), (II) and (III)," *Transp. Res. B, Methodol.*, vol. 27, no. 4, pp. 281–313, 1993.

[47] J.-P. Aubin, A. M. Bayen, and P. Saint-Pierre, "Dirichlet problems for some Hamilton–Jacobi equations with inequality constraints," *SIAM J. Control Optim.*, vol. 47, no. 5, pp. 2348–2380, Jan. 2008.

[48] M. G. Crandall and P.-L. Lions, "Viscosity solutions of Hamilton–Jacobi equations," *Trans. Amer. Math. Soc.*, vol. 277, no. 1, pp. 1–42, 1983.

[49] M. Bardi and I. Capuzzo-Dolcetta, *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*. Boston, MA, USA: Birkhäuser, 1997.

[50] E. N. Barron and R. Jensen, "Semicontinuous viscosity solutions for Hamilton–Jacobi equations with convex Hamiltonians," *Commun. Partial Differ. Equ.*, vol. 15, no. 12, pp. 1713–1742, 1990.

[51] H. Frankowska, "Lower semicontinuous solutions of Hamilton–Jacobi–Bellman equations," *SIAM J. Control Optim.*, vol. 31, no. 1, pp. 257–272, 1993.

[52] C. G. Claudel and A. M. Bayen, "Convex formulations of data assimilation problems for a class of Hamilton–Jacobi equations," *SIAM J. Control Optim.*, vol. 49, no. 2, pp. 383–402, Jan. 2011.

[53] G. M. Coclite, M. Garavello, and B. Piccoli, "Traffic flow on a road network," *SIAM J. Math. Anal.*, vol. 36, no. 6, pp. 1862–1886, 2005.
[54] J.-P. Lebacque, "The Godunov scheme and what it means for first order traffic flow models," in *Proc. 13th Int. Symp. Transp. Traffic Theory*, Lyon, France, 1996, pp. 647–677.
[55] G. Costeseque and J.-P. Lebacque, "A variational formulation for higher order macroscopic traffic flow models: Numerical investigation," *Transp. Res. B, Methodol.*, vol. 70, pp. 112–133, Dec. 2014.
[56] S. Qiu, M. Abdelaziz, F. Abdellatif, and C. G. Claudel, "Exact and grid-free solutions to the Lighthill–Whitham–Richards traffic flow model with bounded acceleration for a class of fundamental diagrams," *Transp. Res. B, Methodol.*, vol. 55, pp. 282–306, Sep. 2013.

Suyash C. Vishnoi received the B.Tech. degree in civil engineering from the Indian Institute of Technology at Roorkee in 2018, and the M.S.E. degree in civil engineering from UT Austin in 2020. He is currently pursuing the Ph.D. degree with the Department of Civil, Architectural and Environmental Engineering, The University of Texas at Austin. His research interests include state estimation and control of traffic on highway networks. He was a recipient of the Talbot S. Huff, Sr. Highway Engineering Graduate Fellowship from the Cockrell School of Engineering.

Christian G. Claudel (Member, IEEE) received the M.S. degree in plasma physics from the Ecole Normale Superieure de Lyon in 2004 and the Ph.D. degree in electrical engineering and computer sciences (EECS) from UC Berkeley in 2010. He is currently an Assistant Professor of Civil, Architectural and Environmental Engineering with The University of Texas, Austin. His research interests include control and estimation of distributed parameter systems, wireless sensor networks, and unmanned air vehicles. He received the Leon Chua Award from UC Berkeley in 2010 for his work on Mobile Millennium.