# Robust traffic density estimation using Discontinuous Galerkin formulation of a macroscopic model

Tigran T. Tchrakian
IBM Research – Ireland
Dublin, Ireland
Email: tigran@ie.ibm.com

Sergiy Zhuk
IBM Research – Ireland
Dublin, Ireland
Email: sergiy.zhuk@ie.ibm.com

Alberto C. Nogueira Jr.
IBM Research – Brazil
São Paulo, Brazil
Email: albercn@br.ibm.com

***Abstract*—In this paper, we develop a data-assimilation algorithm for a macroscopic model of traffic flow. The algorithm is based on the Discontinuous Galerkin Method and Minimax Estimation, and is applied to a macroscopic model based on a scalar conservation law. We present numerical results which demonstrate the shock-capturing capability of the algorithm under high uncertainty in the initial traffic condition, using only sparse measurements, and under time-dependent boundary conditions. The latter makes it possible for estimation to be performed on merge/diverge sections, allowing the possibility of the deployment of the algorithm to road networks.**

## I. Introduction

Traffic state estimation has gained a lot of attention over the last few years, partly due to the increased availability of traffic data. Data from fixed-infrastructure sensors such as loop detectors and traffic cameras are widely used as inputs to traffic control systems, and can also be used to determine the state of traffic networks and provide predictions. In addition to those fixed data sources, mobile phones can be used to generate what is known as floating car data (FCD), whose pervasiveness has led to an abundance of real-time traffic data in cities.

Traffic data can be used in conjunction with a flow model to provide estimates of the traffic state. This can either be done in real-time or offline, depending on the objective. State estimation based on macroscopic models of traffic flow has been studied by numerous authors using techniques such as Extended or Ensemble Kalman Filtering [1], [2], [3], [4], [5], [6], [7], Particle Filtering [8] and Newtonian relaxation [9]. Recently, Tchrakian and Zhuk [10], [11] proposed using a minimax filter for the macroscopic traffic data-assimilation problem. For this, the traffic flow model was solved using a spectral method, namely the Fourier-Galerkin method. While the numerical convergence results looked promising, the Fourier Galerkin method required the use of periodic boundary conditions, which are unrealistic in practice. In this paper, we address this issue by developing a macroscopic traffic data-assimilation algorithm which employs a different spectral method in place of Fourier Galerkin, namely, the Discontinuous Galerkin method.

The paper is organized as follows. In Section II, we describe the traffic flow model used in the paper. In Section III, we discuss the proposed work in the context of previous works on traffic state estimation. In Section IV, we describe the Discontinuous Galerkin formulation of the macroscopic traffic flow model. In Section V, the data assimilation procedure is described, and in Section VI, we present numerical results, including the solution of an initial value problem for the macroscopic model using the Discontinuous Galerkin method, as well as the data assimilation results using synthetic observations. We conclude with a discussion of future work in Section VII.

## II. Macroscopic traffic flow model

In this paper, we employ the LWR [12], [13] model which is based on the following scalar conservation law

$$u(x,t)_t + f(u(x,t))_x = 0, \tag{1}$$

with initial data

$$u_0(x) = u(x,0), \tag{2}$$

and boundary conditions,

$$u(0_-,t) = g(t), \quad u(L_+,t) = h(t). \tag{3}$$

The scalar valued function, $u : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$ is the traffic density. The flux function is given by $f : \mathbb{R} \to \mathbb{R}$, and the independent variables, $x \in \mathbb{R}$ and $t \in \mathbb{R}_+$ denote space (along roadway) and time respectively. We use the quadratic flux function,

$$f(u) = uV_m\left(1 - \frac{u}{u_m}\right), \tag{4}$$

where $u_m$ and $V_m$ are the maximum density and maximum attainable speed respectively. Clearly, $u \in [0, u_m]$. In this work, we take $V_m = u_m = 1$.

## III. State-estimation for macroscopic models

In Section I, we cited several existing data assimilation works. Most of those works are based on the use of finite difference based methods. That is, the PDE describing the macroscopic model is first discretized using the finite difference method, and a data assimilation algorithm is subsequently built on the resulting discrete model. An example of this

is Sun et al. [3], [4] who used the CTM [14] and Mixture Kalman Filter for the estimation of highway traffic. Another example is Herrera and Bayen [9] who used the CTM to model the traffic flow, and then used a Newtonian Relaxation method as the data assimilation technique. An exception is Tchrakian and Zhuk [10], [11] who used a global solution method for the PDE. In that work, the Fourier-Galerkin method was employed for the solution of the flow model. This resulted in a model which was continuous in space. A minimax estimator was then applied to this model in order to perform the data assimilation. Although that algorithm showed good numerical convergence results even with sparse and noisy observations, it had limitations due to the choice of basis functions (Fourier modes) used in the Galerkin method. Specifically, the boundary conditions were restricted to be periodic. This corresponds physically to the flow of traffic on a ring-road: a scenario of limited appeal in practice. In order to address this limitation, we employ a method which does not enforce the same restriction on the boundaries. In Section IV, we describe the Discontinuous Galerkin method which, as we will see, allows the incorporation of more general boundary conditions, leading to a data-assimilation algorithm with the potential for network deployment.

## IV. Numerical Scheme

In this section, we provide a brief description of the Discontinuous Galerkin (DG) method, and describe its application to the macroscopic model, (1). In an effort to keep the following material as compact as possible, we keep the detail to a bare minimum. For further details on Spectral methods, see [15] and [16].

### A. Nodal expansions over elements

We wish to solve the LWR model, (1), for $t > 0$ on the space interval $x \in [0, L] = \Omega$. First, $\Omega$ is divided into $K$ non-overlapping elements, $x \in [x_l^k, x_r^k] = D^k$, i.e., $x_l^k$ and $x_r^k$ are the left and right boundaries of element $D^k$ respectively. We also define the space variable, $\xi \in [-1, 1] = I$, where $I$ is a reference interval whose significance will be explained later. The mapping from $I$ to the element $D^k$ is given by

$$x(\xi) = x_l^k + \frac{1+\xi}{2} h^k,\ h^k = x_r^k - x_l^k, \quad x \in D^k, \quad (5)$$

In the *nodal* DG method [16], the solution, $u(x,t)$, is approximated by $u_h(x,t)$ on each element. This is the $N+1$-term series,

$$\begin{aligned} u_h^k(x(\xi),t) &= \sum_{i=1}^{N+1} u^k(x_i,t)\ell_i^k(x) \\ &= \sum_{i=1}^{N+1} u^k(x(\xi_i),t)\ell_i(\xi), \end{aligned} \quad (6)$$

where the superscript $k$ refers to element $D^k$. The $\xi_i$ are specific grid-points on $I$, and the points $x_i$ are transformations of $\xi_i$ under (5), i.e., $x_i = x(\xi_i)$. The functions $\ell_i^k$ and $\ell_i$ are Lagrange interpolating polynomials defined on $D^k$ and $I$ respectively. The latter are given by

$$\ell_i(\xi) = \prod_{\substack{j=1 \\ j\neq i}}^{N+1} \frac{\xi - \xi_j}{\xi_i - \xi_j}, \quad (7)$$

which have the cardinal property,

$$\ell_i(\xi_j) = \delta_{ij}. \quad (8)$$

The Lagrange polynomials, $\ell_i^k(x)$, defined on elements, $D^k$, are simply transformations of $\ell_i(\xi)$ under (5). At this point, we can define the reference grid-points, $\xi_i$: these are quadrature points for the numerical integration of degree-$N$ polynomials on $I$, such as (7). Specifically, $\xi_i$ are the Legendre-Gauss-Lobatto (LGL) quadrature points, which are used for the integration of Legendre polynomials on $I$. The reason that quadrature for *Legendre* polynomials is used is because in addition to the nodal expansion, (6), there is a complementary *modal* expansion for $u_h^k(x,t)$ which uses Legendre polynomials, and the two expansions are linked through the use of the associated quadrature points (see [16] for more details). We will not go into any more detail about that here, since we work exclusively with nodal DG. Associated with the LGL points, $\xi_i$, are integration weights, $w_i$. A polynomial of up to degree-$N$ can be integrated exactly on $I$ (see [15, Chapter 5]) using LGL points and weights. The integral of a degree-$N$ polynomial, $P_N$ over $I$ is computed numerically by LGL quadrature as follows:

$$\int_I P_N d\xi = \sum_{i=1}^{N+1} P_N(\xi_i) w_i. \quad (9)$$

In short, the Lagrange polynomials are defined such that the coefficients of the nodal expansion of a function are values of the function itself at the LGL points, transformed from interval $I$ to the element interval $D^k$ using the transformation, (5).

### B. Nodal DG method for LWR model

We will now perform the nodal DG method on the LWR model. In addition to the nodal expansion for the density given by (6), the flux function, $f(\cdot)$, is also approximated by an expansion. On element $D^k$, the flux is approximated by

$$\begin{aligned} f_h^k(x(\xi),t) &= \sum_{i=1}^{N+1} f^k(u^k(x_i,t))\ell_i^k(x) \\ &= \sum_{i=1}^{N+1} f^k(u^k(x(\xi_i),t))\ell_i(\xi). \end{aligned} \quad (10)$$

Considering a single element, the nodal DG involves applying the classical Galerkin projection method [15] on $D^k$. That is, we form the residual, $R_h$ by replacing $u$ and $f$ in (1) with their series approximations, $u_h^k$ and $f_h^k$, from (6) and (10):

$$R_h^k(x,t) = u_h^k(x,t)_t + f_h^k(x,t)_x. \quad (11)$$

The residual is required to be orthogonal to each of the Lagrange polynomials, $\ell_i^k(x)$, $i = 1 \dots N+1$ on $D^k$, giving

$$\langle R_h^k, \ell_n^k(x) \rangle_{D^k} = \int_{D^k} R_h^k(x,t)\ell_n^k(x)dx = 0, \; n = 1 \dots N+1. \tag{12}$$

We have

$$\int_{D^k} u_h^k(x,t)_t \ell_n^k(x)dx + \int_{D^k} f_h^k(x,t)_x \ell_n^k(x)dx = 0, \quad n = 1 \dots N+1. \tag{13}$$

Using integration by parts on the second integral in (13) to move the $x$-derivative off $f_h^k$ and onto $\ell_n^k$, we get

$$\int_{D^k} \left( u_h^k(x,t)_t \ell_n^k(x) - f_h^k(x,t)\frac{d\ell_n^k}{dx} \right) dx = - \int_{\partial D^k} \hat{\boldsymbol{n}} \cdot f^* \ell_n^k dx, \quad n = 1 \dots N+1, \tag{14}$$

where $\partial D^k$ denotes the boundary of the element, $\hat{\boldsymbol{n}}$ is the outward facing unit normal on the boundary, and $f^*$ is the numerical flux at the boundary, which we will come back to. Changing the integration variable of the left-hand-side integral from $x$ to $\xi$ using the transformation, (5), we get

$$\frac{h^k}{2} \int_I u_h^k(x(\xi),t)_t \ell_n(\xi)d\xi - \int_I f_h^k(x(\xi),t)\frac{d\ell_n}{d\xi}d\xi = - \int_{\partial D^k} \hat{\boldsymbol{n}} \cdot f^* \ell_n^k dx, \quad n = 1 \dots N+1, \tag{15}$$

where we recall that $\ell_n(\xi)$ are the Lagrange polynomials defined over $I$ as opposed to $D^k$. Using the Gauss-Lobatto quadrature rule, (9), and invoking the cardinal property of Lagrange polynomials, (8), the weak form of DG on element $D^k$ is:

$$M^k \frac{d}{dt}\boldsymbol{u}_h^k - S^T \boldsymbol{f}_h^k = - \left[ \boldsymbol{\ell}^k(x) f^* \right]_{x_l^k}^{x_r^k}, \tag{16}$$

where

$$\begin{aligned} \boldsymbol{u}_h^k &= \left[ u^k(x_1,t), u^k(x_2,t), \dots u^k(x_{N+1},t) \right]^T, \\ \boldsymbol{f}_h^k &= \left[ f^k(x_1,t), f^k(x_2,t), \dots f^k(x_{N+1},t) \right]^T, \\ \boldsymbol{\ell}^k(x) &= \left[ \ell_1^k(x), \ell_2^k(x), \dots \ell_{N+1}^k(x) \right]^T, \end{aligned}$$

and

$$\begin{aligned} M_{ij}^k &= \frac{h^k}{2} \int_I \ell_i(\xi)\ell_j(\xi)d\xi, \\ S_{ij} &= \int_I \ell_i(\xi)\frac{d\ell_j(\xi)}{d\xi}d\xi, \end{aligned}$$

are known as the mass and stiffness matrices respectively.

Now, we return to the numerical flux, $f^*$. For this work, we choose the Godunov [17], [18] flux, since it has been a popular choice in the traffic context due to its intuitive interpretation as a supply/demand function [19], [20]. It is given by:

$$f^*(x) = \begin{cases} \min\limits_{u_h^k(x_-) \le u \le u_h^k(x_+)} f(u) & \text{if} \quad u_h^k(x_-) \le u_h^k(x_+), \\ \max\limits_{u_h^k(x_+) \le u \le u_h^k(x_-)} f(u) & \text{if} \quad u_h^k(x_-) > u_h^k(x_+). \end{cases} \tag{17}$$

The $x_-$ and $x_+$ refer to the left and right faces of the point $x$. Since $f^*(\cdot)$ takes either $x_l^k$ or $x_r^k$ as an argument, it returns the flux at either interface of an element. For an element, $D^k$, bordered by two other elements $D^{k-1}$ and $D^{k+1}$, the numerical flux is computed using the solution on the neighbouring elements, since $x_l^k = x_r^{k-1}$ and $x_r^k = x_l^{k+1}$. It is through these fluxes that the elements "communicate", allowing the global solution to be obtained. In the case that element $D^k$ is on the boundary of $\Omega$, then $x_l^k = 0$ or $x_r^k = L$ (recall that $x \in [0, L] = \Omega$), so the boundary conditions, (3), must be used, i.e. $u_h^k(x_-) = g(t)$ if $x_l^k = 0$ and $u_h^k(x_+) = h(t)$ if $x_r^k = L$

### C. Time-integration and post-processing

The DG system for the LWR model (16) can be solved forward in time from an initial condition, (2), and subject to boundary conditions, (3). This can be done numerically using either explicit or implicit time-integration methods. In Section V-B, we will see that the data-assimilation algorithm makes use of an implicit integration method.

The scalar conservation law, (1), can develop shock discontinuities even for smooth initial data, (2). In order to deal with discontinuities, we must identify steep gradients forming in the solution. To do this, a flux limiter is applied. The limiter we use is known as a *minmod* limiter [16], which identifies changes in sign in the slope of the solution at neighbouring LGL points within elements. Such a change in sign indicates a spurious oscillation in the solution, which the minmod filter then suppresses. As we will see in Section VI, the numerical results indicate that post-processing the DG solution with a minmod filter appears provide good tracking of shocks in the solution and in the estimate.

## V. Data assimilation

In this section, we describe the data-assimilation portion of the paper. First, we describe the state and observation equations which we generate using the DG discretization of the PDE, and then we describe the data-assimilation algorithm which we employ, namely, the minimax filter.

### A. State and observation equations for DG formulation of LWR model

For the purpose of data assimilation, we express the DG-discretized LWR model in the following form:

$$\frac{dU}{dt} = A(U)U + B(t). \tag{18}$$

This is the global system obtained by assembling the $K$ elemental systems given by (16). The state vector, $U$, is thus composed of the elemental state vectors, $\boldsymbol{u_h^k}$, and matrix, $A(U)$ is a block-diagonal matrix:

$$A = \begin{bmatrix} A^1 & 0 & \cdots & 0 \\ 0 & A^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A^K \end{bmatrix} \tag{19}$$

where

$$A^k = (M^k)^{-1} S^T \left( V_m - \frac{1}{u_m} \text{diag}\left(\boldsymbol{u_h^k}\right) \right), \tag{20}$$

which we get using the flux function, (4). The vector, $B(t)$, is formed using the inter-element and boundary flux terms on the right hand side of (16), i.e.,

$$B = \left[F^1, F^2, \dots F^K\right]^T, \tag{21}$$

where

$$F^k = -(M^k)^{-1} \left[\boldsymbol{\ell}^k(x) f^*\right]_{x_l^k}^{x_r^k}. \tag{22}$$

We note that this representation is not unique; for certain choices of numerical flux, it may be possible for the $A(U)U$ term to absorb vector $B(t)$. However, our choice of numerical flux, (17), would make this difficult to do. We also note that each of the elemental systems should be solved separately for the sake of computational efficiency. The reason we express the system in the form (18) is so that the description of the filter in Section V-B is kept simple, without the need for any emphasis on elements.

Next, we introduce the observation equation,

$$Y(t) = HU(t) + \eta(t), \tag{23}$$

where $Y : \mathbb{R} \to \mathbb{R}^M$ is an $M$-element vector containing observations at time, $t$, and $H$ is an observation operator. The vector, $\eta$, is a noise term which accounts for the measurement error. Since the state vector, $U$, comprises elemental grid functions, $\boldsymbol{u}_h^k$ over transformed LGL points, the operator, $H$, takes a particularly simple form. If observations are available at all grid-points, then $H$ is simply the identity matrix. In the case where the number of available observations, $M$, is less than the total number of grid-points, the matrix $H$ comprises only the corresponding $M$ rows of the full $H$ (identity matrix).

In Section V-B, we will use the system, (18) and (23), for data assimilation.

### B. Minimax filter

The data assimilation algorithm we present is based on the minimax state estimation framework. For this, we require a system of ODEs governing the evolution of a state vector, $U$, of the form:

$$\frac{dU}{dt} = A(p_1)U + B(p_2) + e^m(t), \quad U(0) = U_0 + e^b, \tag{24}$$

where $U_0$, is the initial condition, $e^m$ represents model error, $e^b$ describes the error in the initial condition, and $p_1$ and $p_2$ are parameters, possibly depending on the state vector $U$ and/or time, $t$. In our case, comparing with (18), we see that while both $A$ and $B$ are both time-dependant, there is an explicit dependence of $A$ on $U$. The initial state-, model- and measurement-error terms, $e^m$, $e^b$ and $\eta$ can be used to define the model and error covariance matrices, $t \mapsto Q(t)$ and $t \mapsto R(t)$. For details on this, see [11].

The minimax filter equations are given by:

$$\begin{aligned} \frac{dP}{dt} &= A(p_1)P + PA^T(p_1) + Q^{-1} - PH^T RHP, \\ & \qquad K(0) = Q_0^+, \\ \frac{d\widehat{U}}{dt} &= A(p_1)\widehat{U} + B(p_2) + P^{-1}H^T R(Y(t) - H\widehat{U}(t)), \\ & \qquad \widehat{U}(0) = U_0, \end{aligned} \tag{25}$$

where $t \mapsto P(t)$ is the minimax gain, $Q_0 = Q(0)$, and $Q_0^+$ denotes the pseudoinverse of the matrix $Q_0$.

Our implementation of the minimax filter is summarized as Algorithm 1, where we call it *implicitly-linearized minimax*. This term was introduced by two of the authors of this paper in [11] to describe the process of linearizing the filter equations, (25), about each new state estimate in a manner analogous to the Extended Kalman Filter. In practice, this approach would be useful for online estimation. The state and observation error covariance matrices, $Q(t)$ and $R(t)$ can be varied over time, and depend on the trust we place in the model and measurements respectively.

**Algorithm 1** Implicitly-linearized minimax

```
Set initial data, u_0 (known or estimated)
Set computational time-step, Δt
Set number of time-steps, N_t
Generate initial elemental grid functions, u_h^k,  k = 1...K
Combine grid functions, u_h^k, to form initial state vector, U_0
p_1 ← U_0
p_2 ← 0
for i = 1 to N_t do
    Take measurements, Y(iΔt), from real or simulated traffic flow
    Assemble A(p_1) and B(p_2)
    Advance system (25) forward in time using implicit midpoint method to obtain Û(iΔt)
    p_1 ← Û(iΔt)
    p_2 ← iΔt
end for
```

## VI. Numerical results

In this section, we present some numerical results. In Section VI-A, we test the DG discretization of LWR to show that it can capture the solution correctly under presence of time-dependent boundary conditions. In Section VI-B, we present data-assimilation results using the DG-based minimax algorithm using synthetic data.

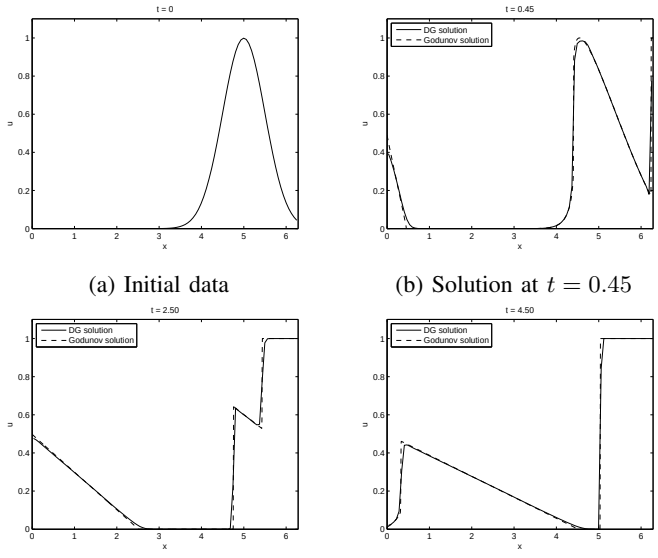
### A. Numerical results of DG-LWR

We test the DG-discretization of LWR with $x \in [0, 2\pi]$ using the following initial condition:

$$u_0(x) = u_m e^{-2(x-5)^2}, \tag{26}$$

and with the left- and right-boundary conditions:

$$g(t) = \frac{1}{2} u_m \left(1 + \sin(t)\right), \quad h(t) = u_m. \tag{27}$$




(a) Initial data

(b) Solution at $t = 0.45$

(c) Solution at $t = 2.50$

(d) Solution at $t = 4.50$

Fig. 1: DG and Godunov numerical solutions of LWR with initial and boundary conditions, (26) and (27)

The left boundary condition is clearly a sinusoidally-varying inflow, while the right boundary condition describes a red traffic light (i.e., zero-outflow is mandated, since $f(u_m) = 0$). We discretize the LWR model, (1), using the DG approach described in Section IV-B with $4$-th order elements (i.e., $N = 4$), to arrive at a system of elemental systems of the form (16). These are then solved forward in time using the implicit midpoint method (although an explicit method could be used instead). In Figure 1, we show the results of this initial-value problem at four time instants. Also shown is the solution to the same problem using a Godunov difference scheme. The purpose of this is to test the DG formulation of LWR against the commonly used Godunov scheme. In Figure 1b, we see two shocks forming: one close to $x = 4.5$, and the other at the right boundary, which indicates a queue beginning to form. In Figure 1c, we see that one shock is traveling forward, while the queue-shock is traveling backward as expected. In Figure 1d, we see that the two shocks from the previous time have coalesced, and a new shock has formed near the left boundary due to the inflow condition. Over all time, we observe that the DG solution remains close to the Godunov solution.

### *B. Data assimilation results*

The data assimilation is performed using Algorithm 1 together with synthetic data. These data are produced using the Godunov finite-difference solution to the LWR model. In other words, the Godunov solution to LWR, including initial and boundary conditions, is taken to mimic reality. From this simulated reality, a sparse set of observations, $Y(t)$, is taken at each time-step and assimilated into the DG model, (16), using the minimax filter, (25), implemented in the fashion described in Algorithm 1. To mimic a real-time estimation scenario, the DG/minimax-based filter and the Godunov scheme (reality) are run concurrently.

We initialize the truth/reality (Godunov) using the initial data, (26), and with left and right boundary conditions,

$$g(t) = \frac{1}{2}u_m\left(1 + \sin(t)\right), \quad h(t) = \frac{1}{2}u_m\left(1 + \cos(t)\right). \quad (28)$$

The data, $Y(t)$, generated by the Godunov scheme are made available to the minimax filter at every computational time-step. However, these data are spatially sparse; we take observations at $9$ evenly-spaced locations over the spatial interval, i.e., $M = 9$. The DG model for the filter is initialized using the data, (26), with an additive perturbation. This perturbation is taken to be large, meaning that we do not assume accurate knowledge of the initial condition of the truth. Rather than supplying the boundary conditions, (28), to the minimax algorithm, we instead infer them from the data, $Y(t)$.

The filter results are shown in Figure 2 alongside the truth and the observations sampled from the latter. In Figure 2a, we see the estimate close to the initial time, where we observe the large difference between the truth and the estimate. In Figure 2b, we see that the estimate has begun to converge to the truth; in particular, a shock that has formed in the truth has been captured. In Figure 2c, we see that the estimate appears to have converged to the truth, and is tracking shocks correctly.

## VII. Conclusion

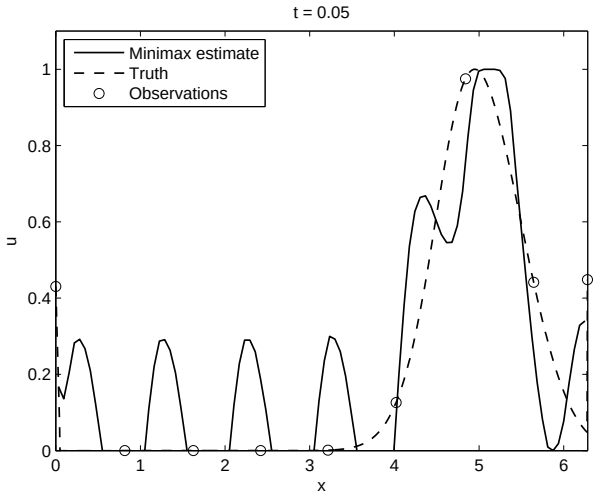
In this paper, we presented a macroscopic traffic data assimilation algorithm based on the Discontinuous Galerkin (DG) method and Minimax Estimation. The numerical results demonstrated the ability of the algorithm to track the traffic state (including the capturing of shocks) with imprecise knowledge of the initial traffic condition, sparse observations in space, and with the inclusion of general time-dependent boundary conditions. The last feature addresses a significant drawback of previous traffic state estimation work based on spectral methods, namely the restriction to periodic (ring-road) boundary conditions. With this shortcoming addressed, the way is paved for the estimation of signal controlled junctions, merge/diverge sections, and consequently networks of roads. These applications will be addressed in a subsequent paper.

The elemental nature DG may provide a considerable computational benefit: for certain regions of space where there are no observations available, it may be sufficient to run the DG model forward in time without filtering. In the DG framework, the filter may be run on elements which contain sensors. This concept of running "local filters" on a subset of elements could potentially reduce the computational cost of the algorithm. This idea will be developed in a future paper.

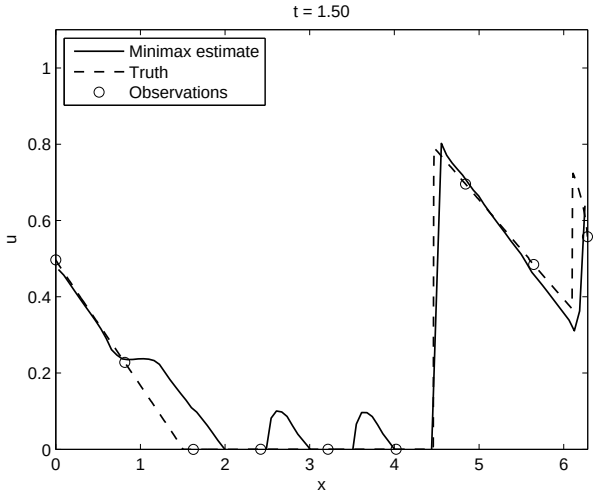
## References

[1] Y. Yuan, H. Van Lint, F. Van Wageningen-Kessels, and S. Hoogendoorn, "Network-wide traffic state estimation using loop detector and floating car data," *Journal of Intelligent Transportation Systems*, vol. 18, no. 1, pp. 41–50, 2014. [Online]. Available: http://dx.doi.org/10.1080/15472450.2013.773225

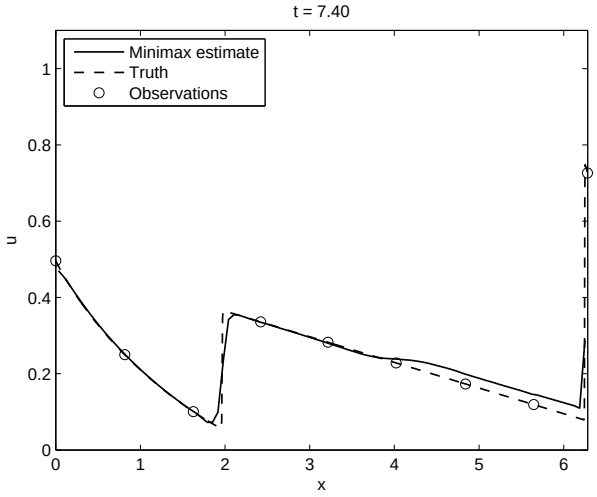
[2] D. B. Work, S. Blandin, O.-P. Tossavainen, B. Piccoli, and A. M. Bayen, "A traffic model for velocity data assimilation," *Applied Mathematics Research eXpress*, vol. 2010, no. 1, pp. 1–35, 2010.




(a) Estimate and truth at $t = 0.05$




(b) Estimate and truth at $t = 1.50$




(c) Estimate and truth at $t = 7.40$

Fig. 2: Results of Minimax Filtering


[3] X. Sun, L. Muñoz, and R. Horowitz, "Highway traffic state estimation using improved mixture Kalman filters for effective ramp metering control," in *Decision and Control, 2003. Proceedings. 42nd IEEE Conference on*, vol. 6, 2003, pp. 6333 – 6338 Vol.6.

[4] ——, "Mixture Kalman filter based highway congestion mode and vehicle density estimator and its application," in *American Control Conference, 2004. Proceedings of the 2004*, vol. 3, 2004, pp. 2098 – 2103 vol.3.

[5] L. Muñoz, X. Sun, R. Horowitz, and L. Alvarez, "Traffic density estimation with the Cell Transmission Model," in *American Control Conference, 2003. Proceedings of the 2003*, vol. 5, 2003, pp. 3750 – 3755 vol.5.

[6] Y. Wang and M. Papageorgiou, "Real-time freeway traffic state estimation based on extended Kalman filter: a general approach," *Transportation Research Part B: Methodological*, vol. 39, no. 2, pp. 141 – 167, 2005.

[7] C. Bucknell and J. C. Herrera, "A trade-off analysis between penetration rate and sampling frequency of mobile sensors in traffic state estimation," *Transportation Research Part C: Emerging Technologies*, vol. 46, no. 0, pp. 132 – 150, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0968090X1400120X

[8] D. Ngoduy, "Applicable filtering framework for online multiclass freeway network estimation," *Physica A: Statistical Mechanics and its Applications*, vol. 387, no. 2-3, pp. 599 – 616, 2008.

[9] J. C. Herrera and A. M. Bayen, "Incorporation of Lagrangian measurements in freeway traffic state estimation," *Transportation Research Part B: Methodological*, vol. 44, no. 4, pp. 460 – 481, 2010.

[10] T. T. Tchrakian and S. Zhuk, "A macroscopic traffic data assimilation framework based on fourier-galerkin method and minimax estimation," in *Intelligent Transportation Systems - (ITSC), 2013 16th International IEEE Conference on*, Oct 2013, pp. 1858–1863.

[11] ——, "A macroscopic traffic data-assimilation framework based on the fourier-galerkin method and minimax estimation," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 16, no. 1, pp. 452–464, Feb 2015.

[12] M. J. Lighthill and G. B. Whitham, "On kinematic waves II. A theory of traffic flow on long crowded roads," *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, vol. 229, no. 1178, pp. 317–345, 1955.

[13] P. I. Richards, "Shock waves on the highway," *Operations Research*, vol. 4, no. 1, pp. 42–51, 1956.

[14] C. F. Daganzo, "The Cell Transmission Model: A dynamic representation of highway traffic consistent with the hydrodynamic theory," *Transportation Research Part B: Methodological*, vol. 28, no. 4, pp. 269 – 287, 1994.

[15] J. Hesthaven, D. Gottlieb, and S. Gottlieb, *Spectral Methods for Time-Dependent Problems*. Cambridge: Cambridge Univ. Press, 2007.

[16] J. Hesthaven and T. Warburton, *Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications*, ser. Texts in Applied Mathematics. Springer, 2008. [Online]. Available: http://books.google.co.uk/books?id=APQkDOmwyksC

[17] R. J. LeVeque, *Numerical Methods for Conservation Laws*, ser. Lectures in Mathematics ETH Zürich, Department of Mathematics Research Institute of Mathematics. Springer, 1992. [Online]. Available: http://books.google.co.uk/books?id=3WhqLPcMdPsC

[18] S. K. Godunov, "A difference scheme for numerical solution of discontinuous solution of hydrodynamic equations," *Math. Sbornik*, vol. 47, pp. 271 – 306, 1959.

[19] C. F. Daganzo, "A behavioral theory of multi-lane traffic flow. part I: Long homogeneous freeway sections," *Transportation Research Part B: Methodological*, vol. 36, no. 2, pp. 131 – 158, 2002. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0191261500000424

[20] ——, "A behavioral theory of multi-lane traffic flow. part II: Merges and the onset of congestion," *Transportation Research Part B: Methodological*, vol. 36, no. 2, pp. 159 – 169, 2002. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0191261500000436