Adjoint-based optimization on a network of discretized scalar conservation law PDE's: application to coordinated ramp metering

Jack

May 3, 2013

Abstract

The adjoint method provides a computationally efficient means of calculating the gradient for applications in constrained optimization. In this article, we consider a network of scalar conservation law PDE's with general topology, whose behavior is modified by a set of control parameters in order to minimize some objective. After discretizing the PDE system via the Godunov scheme, we detail the computation of the gradient of the discretized system with respect to the control parameters and show that the complexity of its computation scales linearly for networks with small vertex degree. The method is applied to solve the problem of coordinated ramp metering on freeway networks and is shown to improve the performance and running time over existing methods.

1 Introduction

Control of PDE's with the adjoint method

- Applications
  - Optimization
  - Estimation
- Types
  - Continuous
  - discrete
  - numerical differentiation
- Tradeoffs
  - compactness of formulation
  - well-posedness

Applications

- Shape op
  - Alonso
  - Jameson
  - Pierce
- Optimization
  - Pirroneau[1]
- Estimation
  - Tomlin[10]
- Traffic (air, vehicle)
  - Bayen, Jacquet, Dengfeng
- Water
  - [4]Honnorat, Strub [13]
- EnKF adjoint
  - Bewley [5]

Network of PDE's

- work in discretized domain
  - Godunov
- Optimal control difficult for D-PDE
  - nonlinear, complex, piecewise constraints
- Alternative approaches
  - Relaxation of constraints
    * Gomes, Ziliaskoupoulos
  - decoupled systems
    * Frejo [6, 11]
  - explicit perturbation
    * something like SPSA

Traffic

- Problem and related Work
  - Roberto, Lyapunov
  - Papageourgiou
  - Jacquet
  - all in community

Contributions

- Complexity analysis for generic discretized network of PDE's
- Improvement to existing ramp metering work.

Outline

2 Preliminaries

2.1 Conservation law PDE's

A conservation law for a single space dimension can written in the form:

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

where $u : [0,\infty] \times \mathbb{R} \to \mathbb{R}$ is the scalar "conserved" quantity, and $f : \mathbb{R} \to \mathbb{R}$ is the flux function. We assume f to be smooth and rewrite Equation (1) as:

$$u_t + f(u)_x = 0 \tag{2}$$

We also assume the system is hyperbolic, which implies that f_x has real values everywhere. Since we are considering a scalar system, this is automatically satisfied.

For a particular initial condition $u_0 : \mathbb{R} \to \mathbb{R}$, we seek "weak" solutions of u for the Cauchy problem:

$$\begin{cases} u_t + f(u)_x & = 0 \\ u(0, x) & = u_0(x) \end{cases}$$

Please see [7] for more details on weak solutions to hyperbolic PDE's. For all systems of the form in Equation (2) that we consider, there exists a unique weak solution for all Cauchy problems.

Definition 1. Riemann Problem

A Riemann problem is a Cauchy problem with a piecewise-constant initial condition that has one point of discontinuity, $x = \bar{x}$. Without loss of generality, we take $\bar{x} = 0$.

2.2 Network of PDE's

A network of PDE's is defined as a set of links ($\mathcal{I}$) and junctions ($\mathcal{J}$). Each link $I \in \mathcal{I}$ has an associated space domain $]0, L_I[$ over which $u_I(t, x)$ is defined and obeys the PDE in Equation (2). Each junction $J \in \mathcal{J}$ is the union of two distinct subsets of links: $Inc(J)$ being the links entering junction J and $Out(J)$ being the links exiting.

While the behavior each link $u_I(t, x)$ is determined by the dynamics described in Equation (2), the behavior of the links still needs to be defined when they meet at junctions.

Definition 2. Riemann problem at junctions

Consider a junction J, with $m = |Inc(J)|$ incoming links and $n = |Out(J)|$ outgoing links. All links $I \in Inc(J) \cup Out(J)$ have a constant initial profile $u_I(0, x) = u_I^0$ (called the Riemann data) with incoming links having an x domain $=]-\infty, 0[$ and outgoing links having an x domain $=]0, \infty[$. Let $\mathbf{u}_J^0$, where $\left\{\mathbf{u}_J^0\right\}_I = u_I^0$[1], be the union of all Riemann data. The solution of $u_I(x, t)$ for all links $I \in Inc(J) \cup Out(J)$ and $t \geq 0$ we define as the solution to the Riemann problem at junction J with initial data $\left\{\mathbf{u}_J^0\right\}$.

We define a Riemann solver:

$$\begin{aligned} RS : \quad \mathbb{R}^{m+n} & \to \mathbb{R}^{m+n} \\ \mathbf{u}_J^0 & \mapsto RS\left(\mathbf{u}_J^0\right) = \hat{\mathbf{u}}_J \end{aligned}$$

where $\{\hat{\mathbf{u}}_J\}_I = \hat{u}_I$ provides the boundary condition for link I at the x position of the junction for all time $t \geq 0$. Specifically, the solution $u_I(t, x)$ for a link $I \in Inc(J)$ to the Riemann problem at junction J is given by the solution over the domain $x < 0$ to the following Riemann problem:

$$\begin{cases} (u_I)_t + f(u_I)_x & = 0 \\ u_I(0, x) & = \begin{cases} u_I^0 & x < 0 \\ \hat{u}_I & x \geq 0 \end{cases} \end{cases}$$

The Riemann problem for outgoing link is defined similarly, with the exception that $u_I(0, x > 0) = u_I^0$. Figure 1 gives a depiction of Riemann solution at a junction J.

The following conditions are required for a valid Riemann solver as part of the definition:

- The Riemann solver must produce self-similar solutions, i.e.

$$RS(\mathbf{u}_J) = RS(\hat{\mathbf{u}}_J) = \hat{\mathbf{u}}_J$$

[1] If $\mathbf{x}$ is a set, where each element $x_i \in \mathbf{x}$ has a corresponding index i which is obvious from context, then we use notation $\{\mathbf{x}\}_i$ to denote the element x_i.

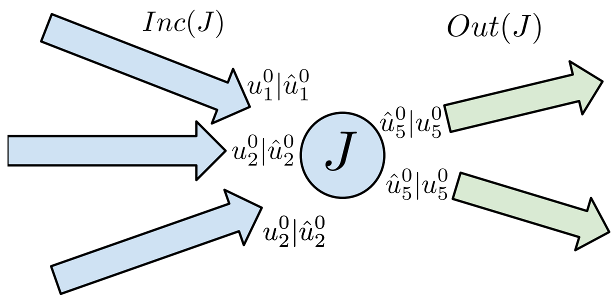

Figure 1: Solution of boundary conditions at junction. The boundary conditions $\hat{\mathbf{u}}^0$ are produced from the initial conditions Riemann solution $RS\left(\mathbf{u}^0\right)$.

- All waves produced from the solution to Riemann problems on all links, generated by the boundary conditions at a junction, must emanate out from the junction. For example, the solution to the Riemann problem on an incoming link must produce waves with negative speeds, while the solution on an outgoing link must produce waves with positive speed.
- The sum of all incoming fluxes must equal the sum of all outgoing fluxes, i.e. if $\hat{\mathbf{u}} = RS\left(\mathbf{u}\right)$, then

$$\sum_{i \in Inc(J)} f\left(\{\hat{\mathbf{u}}_J\}_i\right) = \sum_{j \in Out(J)} f\left(\{\hat{\mathbf{u}}_J\}_j\right)$$

The justification for these conditions can be found in [7].

2.3 Godunov Discretization

To approximate a hyperbolic PDE of the form in Equation 2 with a discretization in both space and time, we use Godunov's scheme [8], an upwind-downwind numerical method. Given an initial condition that is piecewise-constant, Godunov's scheme gives a procedure to advance the system one time-step forward. This process can be repeated *ad infinum* by applying the scheme to the new condition generated from the previous time-step.

Godunov scheme for a single link Consider a hyperbolic PDE with space domain $]0, L[$ that has a piecewise-constant initial condition $u\left(0, x\right)$. There are c discontinuities at locations $\{x_i : i \in [1, \ldots, c]\}$, taking the value $U_i \in \mathbb{R}$ over the range $[x_i, x_{i+1}[$ for $i \in [1, \ldots, c-1]$, U_0 over the range $]0, x_1[$ and U_c over the range $[x_c, L[$. The U_i values can be considered the discrete-valued approximations of u over the range $[x_i, x_{i+1}[$. Then, at time $\triangle t$ and over the range $[x_i, x_{i+1}[$, the *average* value of the system, $\bar{U}_i$, is given by:

$$\bar{U}_i = U_i - \frac{1}{x_{i+1} - x_i} \int_0^{\triangle t} \left(f\left(u\left(t, x_{i+1}\right)\right) - f\left(u\left(t, x_i\right)\right)\right) dt$$

Instead of evaluating the time integral term $\int_0^{\triangle t} f\left(u\left(t, x_i\right)\right) dt$, which requires a solution of the u term forward in time, the Godunov scheme approximates the term with a function g^G that depends on the constants U_i and U_{i+1} such that:

$$\int_0^{\triangle t} f\left(u\left(t, x_i\right)\right) dt \approx \triangle t g^G\left(U_i, U_{i+1}\right)$$

The g^G function can be seen as either an approximation or an exact solution of the flux across the discontinuity point of the Riemann problem with a left initial datum U_i and right initial datum U_{i+1}. Therefore, the Godunov update step for U_i is given by:

$$\bar{U}_i = U_i - \frac{\triangle t}{x_{i+1} - x_i}\left(g^G\left(U_i, U_{i+1}\right) - g^G\left(U_{i-1}, U_i\right)\right) \tag{3}$$

Again, this process can be repeated to approximate the average values of the system at the subsequent time-step by applying the Godunov scheme with initial data $\bar{U}_i$ rather than U_i.

CFL condition for discretized hyperbolic PDE's The Courant–Friedrichs–Lewy (CFL) condition is a condition on the time and space discretization size that guarantees that no two Riemann problems interact with one another over the course of a single time-step. This is a necessary condition for the convergence of the discrete solution to the continuous solution.

$$\lambda^{\max} \leq \frac{\triangle x}{2\triangle t}$$

where $\lambda^{\max} = \max_a \left|f'(a)\right|$ is the maximum absolute wave speed, and $\triangle x$ and $\triangle t$ are the space and time discretization steps, respectively. We assume for the rest of the paper that the time-step $\triangle t$ is fixed and that the discretization step-size $\triangle x_i = L_i = x_{i+1} - x_i$ is chosen to not violate the CFL condition[2].

Godunov scheme at junctions We can view each discontinuity point x_i above as a junction with one incoming link and one outgoing link, which gives us a Godunov scheme for the specific case of 1-to-1 junctions. Given this view, it becomes more convenient to assume that a link $I \in \mathcal{I}$ has a constant initial profile $u_I(0, x) = U_I$ over its entire domain $x \in \left]0, L_I\right[$ rather than a constant piecewise profile, since a constant piecewise profile can be mapped to a series of individual links with 1-to-1 junctions. We now generalize the approach to junctions with an arbitrary number of input links and output links.

Take a junction J with links $I \in Inc(J) \cup Out(J)$ with Riemann data $\{\mathbf{u}(0, x)\}_I = \{\mathbf{U}\}_I$. Applying the Riemann solver to the Riemann data gives the boundary solution $\left\{\hat{\mathbf{U}}\right\}_I$. Then for $I \in Inc(J)$, the downstream flux ($x = L_I$) is determined by the solution of a Riemann problem with initial conditions $\left(U_I, \hat{U}_I\right)$. Furthermore, since all waves emanate *outwards* from a junction, the flux at the downstream boundary is equal to $f\left(\hat{U}_I\right)$ for all t. This leads to a natural choice for the approximate flux function g^G at the junction for link $I \in Inc(J)$:

$$g^G\left(U_I, \hat{U}_I\right) = f\left(\hat{U}_I\right)$$

A similar argument says that the approximate flux function at the junction for outgoing links $I \in Out(J)$ is also the flux function f applied to the Riemann solution $\hat{U}_I$.

For junctions with no incoming or no outgoing links, a boundary condition is required in order to specify a Riemann problem at the upstream and downstream junction of every link. For the discretized problem, the boundary condition is taken to be constant over each discrete time step.

Composing the Riemann solver and flux function The equations are sometimes simplified if we allow ourselves to consider the composition of the Riemann solver RS and the approximate flux function g^G directly. Using the same notation as defined for the Riemann solver, we define the function $\hat{F}$ as just this composition:

$$\begin{aligned} \hat{F}: \quad \mathbb{R}^{m+n} &\to \mathbb{R}^{m+n} \\ \mathbf{U}_J^0 &\mapsto \left\{f\left(\left\{RS\left(\mathbf{U}_J^0\right)\right\}_I\right) : I \in Inc(J) \cup Out(J)\right\} \end{aligned}$$

[2] For the purposes of the Godunov scheme, we only require that the solution originating from one Riemann problem does not influence the flux across the boundary of another Riemann problem. This is equivalent to $\lambda^{\max} \leq \frac{\triangle x}{\triangle t}$

We illustrate the convenience with an example. Take the flux function to be linear $f(u) = \alpha u, \alpha > 0$. Then we can define the Riemann solver for a junction J as follows:

$$\{RS(\mathbf{U}_J)\}_I = \begin{cases} U_I & I \in Inc(J) \\ f^{-1}\left(\frac{\sum_{i \in Inc(J)} f(U_i)}{|Out(J)|}\right) & I \in Out(J) \end{cases}$$

Then to solve for the flux across the junction for an outgoing link $I \in Out(J)$, we again apply the flux function:

$$g^G\left(U_I, f^{-1}\left(\frac{\sum_{i \in Inc(J)} f(U_i)}{|Out(J)|}\right)\right) = f\left(f^{-1}\left(\frac{\sum_{i \in Inc(J)} f(U_i)}{|Out(J)|}\right)\right) = \frac{\sum_{i \in Inc(J)} f(U_i)}{|Out(J)|}$$

It is clear that the intermediate step of solving for the boundary states is unnecessary in this case for the Godunov scheme, since the scheme uses *flux* values at boundaries rather than *state* values, and it is often the case (particularly for the linear flux function considered above) that boundary states are determined by initially solving for the fluxes. Therefore, we can skip the intermediate step of applying the inverse flux function and reapplying the same flux function by considering the composed function $\hat{F}$:

$$\left\{\hat{F}(\mathbf{U}_J)\right\}_I = \begin{cases} f(U_I) & I \in Inc(J) \\ \frac{\sum_{i \in Inc(J)} f(U_i)}{|Out(J)|} & I \in Out(J) \end{cases}$$

Full discrete solution method Assume a discrete scalar hyperbolic network of PDE's with links $\mathcal{I}$, junctions $J \in \mathcal{J}$, and initial conditions $\{u_I(0,x) = U_I : x \in]0, L_I[, I \in \mathcal{I}\}$. The solution method for advancing the discrete system forward one time-step is given by:

1. Begin with initial condition $(t = 0)$ $\{U_I : I \in \mathcal{I}\}$.
2. For every junction $J \in \mathcal{J}$:
   (a) Apply the Riemann solver to Riemann data to obtain the boundary condition $\hat{\mathbf{U}}_J = RS(\mathbf{U}_J)$.
3. For every link $I \in \mathcal{I}$:
   (a) Letting $J_I^{\text{Up}} = J \in \mathcal{J} : I \in Out(J)$ and $J_I^{\text{Down}} = J \in \mathcal{J} : I \in In(J)$, the discrete value over link I at time $\triangle t$, $\bar{U}_I$, is given by:

$$\bar{U}_I = U_I - \frac{\triangle t}{L_I}\left(f\left(\left\{\hat{\mathbf{U}}_{J_I^{\text{Down}}}\right\}_I\right) - f\left(\left\{\hat{\mathbf{U}}_{J_I^{\text{Up}}}\right\}_I\right)\right)$$

If we instead use the composed flux junction solver $\hat{F}$, then the method simplifies to:

1. Begin with initial condition $(t = 0)$ $\{U_I : I \in \mathcal{I}\}$.
2. For every link $I \in \mathcal{I}$:
   (a) Letting $J_I^{\text{Up}} = J \in \mathcal{J} : I \in Out(J)$ and $J_I^{\text{Down}} = J \in \mathcal{J} : I \in In(J)$, the discrete value over link I at time $\triangle t$, $\bar{U}_I$, is given by:

$$\bar{U}_I = U_I - \frac{\triangle t}{L_I}\left(\hat{F}_I\left(\mathbf{U}_{J_I^{\text{Down}}}\right) - \hat{F}_I\left(\mathbf{U}_{J_I^{\text{Up}}}\right)\right) \tag{4}$$

3 State, control, and governing equations

Let's assume we have a discrete system of the form in Section 2.3, and that there is some arbitrary ordering of the links $\mathcal{I}$ such that we can uniquely associate an integer $i_I \in [1, \ldots, |\mathcal{I}|]$ to each link $I \in \mathcal{I}$. For simplicity, we assume $i_I = I$, i.e. I also represents its own unique integer. We wish to solve the system for T time-steps into the future, i.e. we wish to determine the discrete state values $U_{I,k}$ for all links $I \in \mathcal{I}$ and all time-steps $k \in [1, \ldots, T]$. We represent the entire state of the solved system with the vector $\mathbf{x} \in \mathbb{R}^{|\mathcal{I}|T}$, where for $I \in [1, \ldots, |\mathcal{I}|]$ and $k \in [1, \ldots, T]$, such that $x_{|\mathcal{I}|*(k-1)+I} = U_{I,k}$. Furthermore, we assume some "control" variable $\mathbf{v} \in \mathbb{R}^{N_v T}$ that influences the solution of the Riemann problems, where N_v is the number of controllable values at each time-step and each control may be updated at each time-step. Thus, it would be more accurate to represent the solution of the Riemann problem at a junction $J \in \mathcal{J}$ as a function of the current state of connecting links, $\mathbf{x}_{J,k} = \{x_{I,k} : I \in Inc\,(J) \cup Out\,(J)\}$ and the current control parameters, $\mathbf{v}_k = \{v_i : i \in [(k-1)N_v + 1, \ldots, (k)N_v]\}$. More specifically, we let $\mathbf{v}_{J,k} \subseteq \mathbf{v}_k$ be the set of control variables that the Riemann solver at junction J depends on at time k. Then using the same notation as before, we express the Riemann solver as:

$$\begin{aligned} RS: \quad \mathbb{R}^{m+n} \times \mathbb{R}^{|\mathbf{v}_{J,k}|} &\to \mathbb{R}^{m+n} \\ (\mathbf{x}_{J,k}, \mathbf{v}_{J,k}) &\mapsto RS\,(\mathbf{x}_{J,k}, \mathbf{v}_{J,k}) = \hat{\mathbf{U}}_{J,k} \end{aligned}$$

Additionally, for each variable, there is a corresponding update equation. More specifically, the update equations for link $I \in \mathcal{I}$ are:

$$h_{I,1}\,(\mathbf{x}, \mathbf{v}) = x_{I,1} - U_I = 0 \tag{5}$$

$$h_{I,k}\,(\mathbf{x}, \mathbf{v}) = x_{I,k} - x_{I,k-1} + \frac{\triangle t}{L_I}\left(\hat{F}_I\left(\mathbf{x}_{J_I^{\text{Down}},k-1}, \mathbf{v}_{J_I^{\text{Down}},k-1}\right) - \hat{F}_I\left(\mathbf{x}_{J_I^{\text{Up}},k-1}, \mathbf{v}_{J_I^{\text{Up}},k-1}\right)\right) = 0 \quad \forall k \in [2, \ldots, T] \tag{6}$$

for all links $I \in \mathcal{I}$, where U_I is the initial condition for link I, where again $\hat{F}_I\,(a, b) = f\,(\{RS\,(a, b)\}_I)$. Thus, we can construct the system of $|\mathcal{I}|T$ governing equations $H\,(\mathbf{x}, \mathbf{v}) = 0$, where the $h_{I,k}$ is the equation in H at index $|\mathcal{I}|(k-1) + I$, identical to the ordering of the state vectors.

4 Adjoint method

4.1 Optimization Problem

In addition to our governing equations $H\,(\mathbf{x}, \mathbf{v}) = 0$, we also introduce a cost function J:

$$\begin{aligned} J: \quad \mathbb{R}^{|\mathcal{I}|T} \times \mathbb{R}^{N_v T} &\to \mathbb{R} \\ (\mathbf{x}, \mathbf{v}) &\mapsto J\,(\mathbf{x}, \mathbf{v}) \end{aligned}$$

which returns a scalar that serves as a metric of performance of the state and control values of the system. We wish to minimize the quantity J over the set of control parameters $\mathbf{v}$, while constraining the state of the system to satisfy the governing equations $H\,(\mathbf{x}, \mathbf{v}) = 0$. We summarize this with the following optimization problem:

$$\begin{aligned} \min_{\mathbf{v}} \quad & J\,(\mathbf{x}, \mathbf{v}) \\ \text{subject to:} \quad & H\,(\mathbf{x}, \mathbf{v}) = 0 \end{aligned} \tag{7}$$

Both the cost function and governing equations may be non-convex in this problem.

4.2 Calculating the gradient

We wish to use gradient information in order to find control values $\mathbf{v}^*$ that give small costs $J^* = J\left(\mathbf{x}\left(\mathbf{v}^*\right), \mathbf{v}^*\right)$. Since there may exist many local minima for the optimization problem 7 due to the possible non-convexity of J and H, gradient methods do not guarantee global optimality of $\mathbf{v}^*$. Still, nonlinear optimization methods such as interior point optimization utilize gradient information to improve performance[14].

Let us assume we have an initial state and control vector $\mathbf{x}^0, \mathbf{v}^0$ that satisfies $H\left(\mathbf{x}^0, \mathbf{v}^0\right)=0$, and we wish to solve for the derivative of the cost function at the current state and control with respect to $\mathbf{v}$:

$$d_{\mathbf{v}} J\left(\mathbf{x}^0, \mathbf{v}^0\right) = \frac{\partial J\left(\mathbf{x}^0, \mathbf{v}^0\right)}{\partial \mathbf{x}} \frac{d\mathbf{x}}{d\mathbf{v}} + \frac{\partial J\left(\mathbf{x}^0, \mathbf{v}^0\right)}{\partial \mathbf{v}} \tag{8}$$

where J^i is the cost function expression local to points $\left(\mathbf{x}^i, \mathbf{v}^i\right)$. The current difficulty is computing the term $\frac{d\mathbf{x}}{d\mathbf{v}}$. Next we take advantage of the fact that the derivative of $H\left(\mathbf{x}, \mathbf{v}\right)$ with respect to $\mathbf{v}$ is equal to zero since its evaluation must always be zero:

$$d_{\mathbf{v}} H\left(\mathbf{x}^0, \mathbf{v}^0\right) = \frac{\partial H\left(\mathbf{x}^0, \mathbf{v}^0\right)}{\partial \mathbf{x}} \frac{d\mathbf{x}}{d\mathbf{v}} + \frac{\partial H\left(\mathbf{x}^0, \mathbf{v}^0\right)}{\partial \mathbf{v}} = 0 \tag{9}$$

The partial derivative terms, H_x, H_v, J_x, and J_v, can all be evaluated (more details provided in Section 4.3) and then treated as constant matrices. Thus, in order to evaluate $d_{\mathbf{v}} J\left(\mathbf{x}^0, \mathbf{v}^0\right)$, we must solve a coupled system of matrix equations.

Forward system If we solve for $\frac{d\mathbf{x}}{d\mathbf{v}}$ in Equation 9 (which we call the *forward system*):

$$H_x \frac{d\mathbf{x}}{d\mathbf{v}} = -H_v,$$

then we can substitute the solved value for $\frac{d\mathbf{x}}{d\mathbf{v}}$ into Equation 8 to obtain the full expression for the gradient. Section 4.3 gives details on the invertibility of H_x, guaranteeing a solution for $\frac{d\mathbf{x}}{d\mathbf{v}}$.

Adjoint system Instead of evaluating $\frac{d\mathbf{x}}{d\mathbf{v}}$ directly, the adjoint method instead solves the following system, called the *adjoint system*:

$$H_x^T \lambda = -J_x^T \tag{10}$$

Then the expression for the gradient becomes:

$$d_{\mathbf{v}} J\left(\mathbf{x}^0, \mathbf{v}^0\right) = \lambda^T H_v + J_v \tag{11}$$

We show in Section 4.4 how the complexity of computing the gradient is reduced from $O(D_x|\mathbf{x}||\mathbf{v}|)$ to $O(D_x|\mathbf{x}| + D_v|\mathbf{v}|)$ by considering the adjoint method over the forward method, where D_x is the maximum junction degree on the network

$$D_x = \max_{J \in \mathcal{J}} |Inc\left(J\right) \cup Out\left(J\right)| \tag{12}$$

and D_v is the maximum number of constraints that a single control variable appears in:

$$D_v = \max_{v_{i,k} \in \mathbf{v}} |\bigcup Inc\left(J\right) \cup Out\left(J\right) : v_{i,k} \in \mathbf{v}_{J,k}| \tag{13}$$

A graphical depiction of D_x and D_v are given in Figure 2. Since the D_x and D_v typically do not scale with $|\mathbf{x}|$ or $|\mathbf{v}|$ for freeway networks (which have low node degree), the complexity of evaluating the gradient for such networks can be considered linear for the adjoint method.

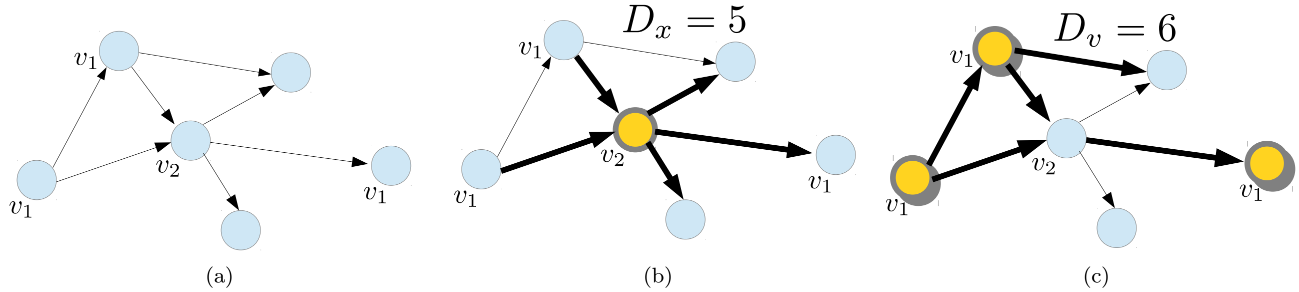

Figure 2: Depicting D_x and D_v for an arbitrary graph. Figure 2a shows the underlying graphical structure for an arbitrary PDE network. Figure 2b depicts the center junction having the largest number of connecting edges, thus giving $D_x = 5$. Figure 2c shows that control parameter v_1 influences three junctions with sum of junctions degrees equal to six, which is maximal over the other control parameter v_2. leading to the result $D_v = 6$.

4.3 Evaluating the partial derivatives

Computing the gradient relies on the existence of partial derivatives of both J and H. We assume that the partial derivatives of the cost function J with respect to both the state $\mathbf{x}$ and control $\mathbf{v}$ exist and have no special structure or sparsity. The networked-structure of the PDE system and the Godunov discretization scheme allows one to say much more about the structure and sparsity of H_x and H_v.

Partial derivative expressions Given that the governing equations require the evaluation of a Riemann solver at each step, we detail some of the necessary computational steps in evaluating the H_x and H_v matrices.

If we consider a particular governing equation $h_{I,k}$, then we may determine the partial term with respect to $x \in \mathbf{x}$ by applying the chain rule:

$$\frac{\partial h_{I,k}}{\partial x} = -\frac{\partial U_{I,k-1}}{\partial x} + \frac{\triangle t}{L_i}\left(\frac{\partial f}{\partial \left\{\hat{\mathbf{U}}_{J_I^{\text{Down}},k}\right\}_I}\frac{\partial \left\{\hat{\mathbf{U}}_{J_I^{\text{Down}},k}\right\}_I}{\partial x} - \frac{\partial f}{\partial \left\{\hat{\mathbf{U}}_{J_I^{\text{Up}},k}\right\}_I}\frac{\partial \left\{\hat{\mathbf{U}}_{J_I^{\text{Up}},k}\right\}_I}{\partial x}\right) \tag{14}$$

or if we consider the composed Riemann flux solver $\hat{F}$ in Section 2.3:

$$\frac{\partial h_{I,k}}{\partial x} = -\frac{\partial U_{I,k-1}}{\partial x} + \frac{\triangle t}{L_i}\left(\frac{\partial \left\{\hat{\mathbf{F}}_{J_I^{\text{Down}},k}\right\}_I}{\partial x} - \frac{\partial \left\{\hat{\mathbf{F}}_{J_I^{\text{Up}},k}\right\}_I}{\partial x}\right) \tag{15}$$

For links $i \neq I$, the $\frac{\partial U_{I,k-1}}{\partial x}$ term is 0. For links i connecting to J_I^{Down} (and assuming no parallel links for simplicity), the $\frac{\partial \left\{\hat{\mathbf{U}}_{J_I^{\text{Up}},k}\right\}_I}{\partial U_{i,l}}$ term will be zero, and for links i connecting to J_I^{Up}, $\frac{\partial \left\{\hat{\mathbf{U}}_{J_I^{\text{Down}},k}\right\}_I}{\partial U_{i,l}} = 0$. It is also clear that for links i non-adjacent to I, $\frac{\partial h_{I,k}}{\partial U_{i,k-1}} = 0$.

Also, for all i and $j \neq k-1$, $\frac{\partial h_{I,k}}{\partial U_{i,j}} = 0$, guaranteeing a lower triangular structure for the $\frac{\partial H}{\partial x}$ matrix. Coupling this result that with the fact that all diagonal terms, $\frac{\partial H_i}{\partial x_i}$, are equal to 1 guarantees the invertibility of $\frac{\partial H}{\partial x}$.

These results are for a generic scalar hyperbolic network of PDE's, and demonstrate that partial derivatives expressions for H_x are only needed for the flux function f and for the Riemann solver RS with respect to state variables adjacent to the junction of the Riemann solver for a given time.

Similarly for H_v, we take a control parameter $v \in \mathbf{v}$, and derive the expression:

$$\frac{\partial h_{I,k}}{\partial v} = \frac{\triangle t}{L_i}\left(\frac{\partial f}{\partial\left\{\hat{\mathbf{U}}_{J_I^{\mathrm{Down}},k}\right\}_I}\frac{\partial\left\{\hat{\mathbf{U}}_{J_I^{\mathrm{Down}},k}\right\}_I}{\partial v} - \frac{\partial f}{\partial\left\{\hat{\mathbf{U}}_{J_I^{\mathrm{Up}},k}\right\}_I}\frac{\partial\left\{\hat{\mathbf{U}}_{J_I^{\mathrm{Up}},k}\right\}_I}{\partial v}\right) \tag{16}$$

or for the composed Riemann flux solver $\hat{F}$:

$$\frac{\partial h_{I,k}}{\partial v} = \frac{\triangle t}{L_i}\left(\frac{\partial\left\{\hat{\mathbf{F}}_{J_I^{\mathrm{Down}},k}\right\}_I}{\partial v} - \frac{\partial\left\{\hat{\mathbf{F}}_{J_I^{\mathrm{Up}},k}\right\}_I}{\partial v}\right) \tag{17}$$

With the same argument, the $\frac{\partial\left\{\hat{\mathbf{U}}_{J,k}\right\}_I}{\partial v_{i,j}}$ terms are only non-zero when $j = k-1$ and $v_{i,j} \in \mathbf{v}_{J,k}$.

Example: linear flux function with split-ratio control For illustrative purposes, we will use the linear-flux function example in Section 2.3 as a concrete problem. We assume for simplicity that each junction J has two outgoing links $\left(I_J^a, I_J^b\right)$ and at each time-step k, J has a control parameter $v_{J,k} \in \mathbf{v}_k \subseteq \mathbf{v}$ that controls the "split ratio" of the flux in the outgoing links such that:

$$\frac{f\left(\hat{U}_{I_{J,k}^a}\right)}{f\left(\hat{U}_{I_{J,k}^a}\right) + f\left(\hat{U}_{I_{J,k}^b}\right)} = v_{J,k}$$

Thus, we have $\mathbf{v}_{J,k} = \{v_{J,k}\}$ and:

$$\{RS\left(\mathbf{U}_{J,k}, \mathbf{v}_{J,k}\right)\}_I = \begin{cases} U_I & I \in Inc\left(J\right) \\ v_{J,k}\sum_{i\in Inc(J)} U_i & I = I_{J,k}^a \\ \left(1 - v_{J,k}\right)\sum_{i\in Inc(J)} U_i & I = I_{J,k}^b \end{cases}$$

or:

$$\left\{\hat{F}\left(\mathbf{U}_{J,k}, \mathbf{v}_{J,k}\right)\right\}_I = \begin{cases} f\left(U_I\right) & I \in Inc\left(J\right) \\ v_{J,k}\alpha\sum_{i\in Inc(J)} U_i & I = I_{J,k}^a \\ \left(1 - v_{J,k}\right)\alpha\sum_{i\in Inc(J)} U_i & I = I_{J,k}^b \end{cases}$$

We can now evaluate the individual terms in Equations 14 and 16.

$$\begin{aligned}
\frac{\partial f}{\partial\left\{\hat{\mathbf{U}}_{J,}\right\}_I} &= \alpha \\
\frac{\partial\left\{\hat{\mathbf{U}}_{J_I^{\mathrm{Down}},k}\right\}_I}{\partial x_{i,j}} &= \frac{\partial\left\{RS\left(\mathbf{U}_{J_I^{\mathrm{Down}},k-1}, v_{J_I^{\mathrm{Down}},k-1}\right)\right\}_I}{\partial x_{i,j}} = \frac{\partial U_{I,k-1}}{\partial x_{i,j}} = \begin{cases} 1 & i = I \text{ and } j = k-1 \\ 0 & \text{otherwise} \end{cases} \\
\frac{\partial\left\{\hat{\mathbf{U}}_{J_I^{\mathrm{Up}},k}\right\}_I}{\partial x_{i,j}} &= \frac{\partial\left\{RS\left(\mathbf{U}_{J_I^{\mathrm{Up}},k-1}, v_{J_I^{\mathrm{Up}},k-1}\right)\right\}_I}{\partial x_{i,j}} = \frac{\partial\left(v_{J_I^{\mathrm{Up}},k}\sum_{i\in Inc(J)} U_i\right)}{\partial x_{i,j}} = 0 \\
\frac{\partial\left\{\hat{\mathbf{U}}_{J_I^{\mathrm{Up}},k}\right\}_I}{\partial v_{i,j}} &= \frac{\partial\left\{RS\left(\mathbf{U}_{J_I^{\mathrm{Down}},k-1}, v_{J_I^{\mathrm{Down}},k-1}\right)\right\}_I}{\partial v_{i,j}} = 0 \\
\frac{\partial\left\{\hat{\mathbf{U}}_{J_I^{\mathrm{Up}},k}\right\}_I}{\partial v_{i,j}} &= \frac{\partial\left\{RS\left(\mathbf{U}_{J_I^{\mathrm{Up}},k-1}, v_{J_I^{\mathrm{Up}},k-1}\right)\right\}_I}{\partial v_{i,j}} = \begin{cases} \sum_{i\in Inc(J)} U_i & j = k-1 \text{ and } i = J_I^{\mathrm{Up}} \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

where the 3rd equation is 0 since $I \in Out\left(J_I^{Up}\right)$ and only incoming link terms appear and the 4th equation is 0 since $I \in Inc\left(J_I^{\text{Down}}\right)$. These explicit solutions can be substituted back into Equations 14 and 16 to completely solve H_x and H_v.

Structure and sparsity We can show the lower-triangular structure of H_x by examining the governing equations in (5) and (6). For $k = 0$, we have $\frac{\partial h_{i,0}}{\partial x_{j,0}} = 1$ if $i = j$ and 0 otherwise, giving all 1's on the diagonal ($\frac{\partial H_{i,j}}{\partial x_{i,j}} = 1 \forall i, j$). For $k \in [2, \ldots, T]$, we have that $h_{I,k}$ is only a function of $x_{I,k}$ and of the state variables from the previous time-step $k-1$. Thus, based on our ordering scheme in Section 3, we know when $i \neq j$, $\frac{\partial h_{i,k}}{\partial x_{j,l}} = 0$ for all $l \geq k$, and $\frac{\partial h_{i,k}}{\partial x_{i,k}} = 1$. These two conditions demonstrate both that H_x is lower-triangular and is also invertible.

Additionally, if we take a single state variable $x_{I,k}$, then we can deduce from Equation (6) that all partial terms will be zero except for those terms involving constraints at time $k+1$ and links connecting to the downstream and upstream junctions J_I^{Down} and J_I^{Up} respectively. To summarize, H_x matrices for systems described in Section 3 will be invertible, lower-triangular and each column will have a maximum cardinality equal to Equation 12.

Using the same line of argument for maximum cardinality in H_x, we can bound the maximum cardinality of each column of H_v. Taking a single control variable $v_{i,k}$ at time-step k, the variable can only appear in constraints at time-step $k+1$ that correspond to a link that connects to a junction J such that $v_{i,k} \in \mathbf{v}_{J,k}$. These conditions give us the expression for D_v in Equation 13, or the maximum cardinality over all columns in H_v.

4.4 Complexity of solving gradient via forward method vs. adjoint method

If we only consider the lower triangular form of H_x, then the complexity of solving for the gradient using the forward system is $O(|\mathbf{x}|^2|\mathbf{v}|)$, where the dominating term comes from solving Equation 8, which requires the solution of $|\mathbf{v}|$ separate $|\mathbf{x}| \times |\mathbf{x}|$ lower-triangular system. The lower-triangular system allows for forward substitution, we can be solved in $O(|\mathbf{x}|^2)$ steps, giving the overall complexity $O(|\mathbf{x}|^2|\mathbf{v}|)$. The complexity of the adjoint method is $O(|\mathbf{x}|^2 + |\mathbf{x}||\mathbf{v}|)$, which is certainly more efficient than the forward-method, as long as $|\mathbf{v}| > 1$. The efficiency is gained by considering that Equation 10 only requires the solution of 1 $|\mathbf{x}| \times |\mathbf{x}|$ *upper*-triangular system (via backward-substitution), followed by the multiplication of $\lambda^T H_v$, a $|\mathbf{x}| \times |\mathbf{x}|$ and a $|\mathbf{x}| \times |\mathbf{v}|$ matrix in Equation 11, for a complexity of $O(|\mathbf{x}|^2 + |\mathbf{x}||\mathbf{v}|)$.

For the adjoint method, this complexity can be improved upon by considering the sparsity of the H_x and H_v matrices, as detailed in Section 4.3. For the backward-substitution step, each entry in the λ vector is solved by *at-most* D_x multiplications, and thus the complexity of solving Equation 10 is reduced to $O(D_x|\mathbf{x}|)$. Similarly, for the matrix multiplication of $\lambda^T H_v$, while λ is not necessarily sparse, we know that each entry in the resulting vector requires at most D_v multiplications, giving a complexity of $O(D_v|\mathbf{v}|)$. The total complexity for the adjoint method on a scalar hyperbolic network of PDE is then $O(D_x|\mathbf{x}| + D_v|\mathbf{v}|)$.

5 Applications to Optimal coordinated ramp metering on freeways

Model Consider a freeway section as a sequence of junctions $\mathcal{J} = [J_1, \ldots, J_n]$. At discrete time $t = k\triangle T, 1 \leq k \leq T$, Junction J_i has an incoming (resp. outgoing) mainline density $\rho_{i,k}^{\text{Inc}}$ (resp. $\rho_{i,k}^{\text{Out}}$) $\in \mathbb{R}$ (units of vehicles per unit length), onramp density $\rho_{i,k}^{\text{On}} \in \mathbb{R}$ (units of vehicles), and offramp split ratio $\beta_{i,k}$ representing the ratio of cars who stay on the freeway vs. total cars leaving the upstream mainline of junction J_i. Since J_1 is the source of the network, it has no upstream mainline or offramp, and similarly J_n has no downstream mainline or onramp. These values are depicted in Figure 3

As control input, an onramp at junction J_i and time k has a metering rate $u_{i,k}$ which limits the flux of vehicles leaving the onramp. Its mathematical model is expressed later.

The vehicle flow dynamics on all links (mainlines, onramps, and offramps) are modeled using the LWR equation:

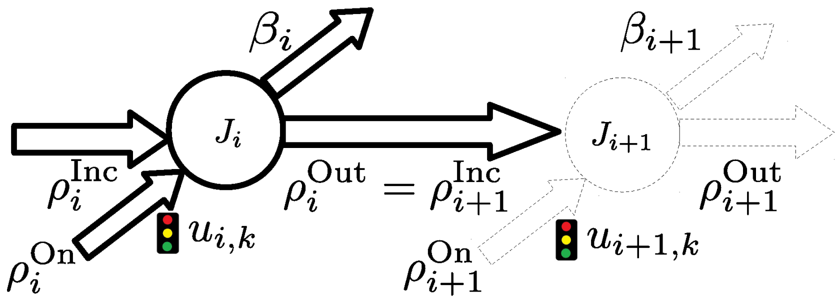

Figure 3: Freeway network junction J_i. Mainline densities represented by ρ, onramp queues by l, and offramp split ratios by β. Note that ρ_i^{Out} is equivalent to $\rho_{i+1}^{\mathrm{Inc}}$.

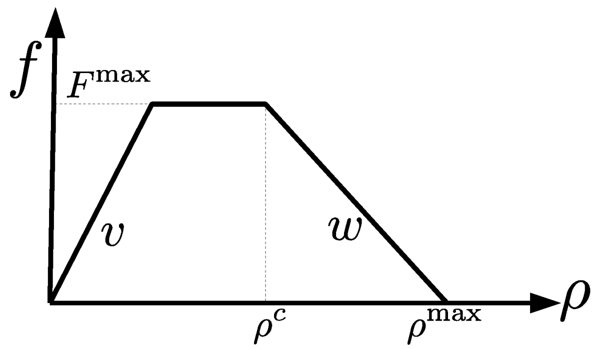

Figure 4: Fundamental diagram with free-flow speed v, congestion wave speed w, max flux $F^{\max}$, critical density ρ^c, and max density $\rho^{\max}$.

$$\rho_t + f'\rho_x = 0$$

where ρ is the density state, and f' is the derivative of the flux function (or fundamental diagram) $f(\rho)$. We assume the fundamental diagram has a trapezoidal form as depicted in Figure 4 .

For notational simplicity we consider the set of densities of links adjacent to a junction J_i at time-step k as $\boldsymbol{\rho}_{J_i,k} = \left\{\rho_{i,k}^{\mathrm{Inc}}, \rho_{i,k}^{\mathrm{Out}}, \rho_{i,k}^{\mathrm{On}}\right\}$. The offramp is considered to have infinite capacity, and thus has no bearing on the solution of junction problems. Taking a junction J_i at timestep k, we discretize the downstream mainline density $\rho_{i,k}^{\mathrm{Out}}$ using the Godunov scheme from Equation 6 to obtain the discrete update equation on a mainline:

$$h_{i,k}^{\mathrm{Out}}(\mathbf{x},\mathbf{v}) = \rho_{i,k}^{\mathrm{Out}} - \rho_{i,k-1}^{\mathrm{Out}} + \frac{\triangle t}{L_i}\left(\left\{\hat{F}\left(\boldsymbol{\rho}_{J_{i+1},k}, u_{i+1,k}\right)\right\}_{\mathrm{Inc}} - \left\{\hat{F}\left(\boldsymbol{\rho}_{J_i,k}, u_{i,k}\right)\right\}_{\mathrm{Out}}\right) = 0 \quad (18)$$

where L_i is the distance between junctions J_i and J_{i+1} and $\left\{\hat{F}\left(\boldsymbol{\rho}_{J,k}\right)\right\}_i$ is the flux from the boundary condition of density $\rho_{J,k}^i$ at junction J. We also make the assumption that onramps have infinite capacity and unit length and can then simplify the onramp update equation to be:

$$h_{i,k}^{\mathrm{On}}(\mathbf{x},\mathbf{v}) = \rho_{i,k}^{\mathrm{On}} - \rho_{i,k-1}^{\mathrm{On}} + \triangle t\left(\left\{\hat{F}\left(\boldsymbol{\rho}_{J_i,k}, u_{i+1,k}\right)\right\}_{\mathrm{On}} - D_{i,k}\right) = 0 \quad (19)$$

where $D_{i,k}$ is the onramp *flux* demand. This formulation results in "strong" boundary conditions at the onramps which guarantees all demand enters the network. Details on weak versus strong boundary conditions can be found in [9, 15, 12].

Riemann solver For the ramp metering problem, there are many potential Riemann solvers that satisfy the properties required in Section 2.2. Following the model of [15, 9], for each junction J_i, we add two modeling decisions: the flux solution maximizes the outgoing mainline flux $\left\{\hat{F}_i\right\}_{\text{Out}}$, and any remaining non-uniqueness in the incoming flux is resolved using a merging parameter p_i. This leads to the following system of equations that gives the flux solution of the Riemann solver at time-step $k > 1$ and junction J_i:

$$\begin{aligned}
\delta &= \min\left(v_i \rho_{i,k}^{\text{Inc}}, F_{i,\text{Inc}}^{\max}\right) \\
\sigma &= \min\left(w_i\left(\rho_i^{\max} - \rho_{i,k}^{\text{Out}}\right), F_{i,\text{Out}}^{\max}\right) \\
d &= u_{i,k} \min\left(\rho_{i,k}^{\text{On}}, F_{i,\text{On}}^{\max}\right) \\
\left\{\hat{F}_{J_i,k+1}\right\}_{\text{Out}} &= \min\left(\beta_{i,k}\delta + d, \sigma\right) \\
\left\{\hat{F}_{J_i,k+1}\right\}_{\text{Inc}} &= \begin{cases} \delta & \frac{\sigma p_i}{1+p_i} \geq \frac{\delta}{\beta_{i,k}} \\ \frac{\left\{\hat{F}_{J_i,k+1}\right\}_{\text{Out}} - d_{i,k}}{\beta_{i,k}} & \frac{\sigma}{1+p_i} \geq d \\ \frac{\sigma p_i}{(1+p_i)\beta_{i,k}} & \text{otherwise} \end{cases} \\
\left\{\hat{F}_{J_i,k+1}\right\}_{\text{On}} &= \left\{\hat{F}_{J_i,k+1}\right\}_{\text{Out}} - \beta_{i,k}\left\{\hat{F}_{J_i,k+1}\right\}_{\text{Inc}} \\
\left\{\hat{F}_{J_i,k+1}\right\}_{\text{Off}} &= \left(1 - \beta_{i,k}\right)\left\{\hat{F}_{J_i,k+1}\right\}_{\text{Inc}}
\end{aligned} \tag{20}$$

Note that the equations can be solved sequentially via forward substitution. Also, we only include the flux results for offramps explicitly here to demonstrate that indeed the flux solution satisfies the flux conservation property. We will ignore its explicit computation for the remainder of the article since its value has no bearing on further calculations.

Optimal coordinated ramp-metering Including the initial conditions as specified in Equation 5 with Equations 18 and 19 gives a complete description of the system $H(\mathbf{x}, \mathbf{v}) = 0$, where if $|\mathcal{I}| = 2 * (n-1)$ is the number of links:

$$\begin{aligned}
x_{|\mathcal{I}|(k-1)+2*i-1} &:= \rho_{i,k}^{\text{Out}} := \rho_{i+1,k}^{\text{Inc}} && 1 \leq i \leq n-1, 1 \leq k \leq T \\
x_{|\mathcal{I}|(k-1)+2*i} &:= \rho_{i,k}^{\text{On}} && 1 \leq i \leq n-1, 1 \leq k \leq T \\
v_{|\mathcal{I}|(k-1)+i} &:= u_{i,k} && 1 \leq i \leq n-1, 1 \leq k \leq T
\end{aligned}$$

The objective of the control is to minimize the *total travel time* of flux on the network. This is expressed with the cost function C:

$$C(\mathbf{x}, \mathbf{v}) = \triangle t \sum_{k=1}^{T} \sum_{i=1}^{n-1} L_i x_{|\mathcal{I}|(k-1)+2*i-1} + x_{|\mathcal{I}|(k-1)+2*i} = \triangle t \sum_{k=1}^{T} \sum_{i=1}^{n-1} L_i \rho_{i,k}^{\text{Out}} + \rho_{i,k}^{\text{On}}$$

Thus the optimal coordinated ramp-metering problem can be formulated as a network PDE-constrained optimization problem:

$$\begin{aligned}
\min_{\mathbf{v}} \quad & C(\mathbf{x}, \mathbf{v}) \\
\text{subject to:} \quad & H(\mathbf{x}, \mathbf{v}) = 0 \\
& 0 \leq v \leq 1 \quad \forall v \in \mathbf{v}
\end{aligned} \tag{21}$$

The inequalities constrain the metering control to neither impose negative flows nor send more vehicles than present in the queue. Since the adjoint method in Section 4 does not permit inequalities in the constraints, we add barrier penalties to the cost function [3, 2]:

$$\tilde{C}(\mathbf{x}, \mathbf{v}, R) = C(\mathbf{x}, \mathbf{v}) + R \sum_{v \in \mathbf{v}} \log((1 - v)(v - 0))$$

As R tends to zero, $\tilde{C}$ tends towards C. Thus we can solve Problem 21 by iteratively solving the augmented problem:

$$\begin{aligned} \min_{\mathbf{v}} \quad & \tilde{C}(\mathbf{x}, \mathbf{v}, R) \\ \text{subject to:} \quad & H(\mathbf{x}, \mathbf{v}) = 0 \end{aligned} \tag{22}$$

with decreasing values of R. As a result, $\tilde{C}$ will approach C as the number of iterations increases.

Applying the adjoint method To use the adjoint method as described in Section 4, we need to compute the partial matrices H_x, H_v, $\tilde{C}_x$ and $\tilde{C}_v$. Computing the partials with respect to the cost function is straight forward:

$$\begin{aligned} \frac{\partial \tilde{C}}{\partial \rho^{\text{Out}}_{i,k}} &= \triangle t L_i & 1 \le i \le n-1, 1 \le k \le T \\ \frac{\partial \tilde{C}}{\partial \rho^{\text{On}}_{i,k}} &= \triangle t & 1 \le i \le n-1, 1 \le k \le T \\ \frac{\partial \tilde{C}}{\partial u_{i,k}} &= R\left(\tfrac{-1}{1-u_{i,k}} + \tfrac{1}{u_{i,k}}\right) & 1 \le i \le n-1, 1 \le k \le T \end{aligned}$$

To compute the partials of H, we follow the procedure in Section 4.2. For a given junction J_i and time-step k, we only need to compute the partial derivatives of the flux solver $\hat{F}_{J_i,k}$ with respect to the adjacent state variables $\boldsymbol{\rho}_{J_i,k}$ and ramp metering control $u_{i,k}$. We calculate the partials of the variables in Equations 20 for a junction J_i at time-step $k > 1$ with respect to either a state or control variable s:

$$\begin{aligned} \frac{\partial \delta}{\partial s} &= \begin{cases} v_i & s = \rho^{\text{Inc}}_{i,k}, v_i \rho^{\text{Inc}}_{i,k} \le F^{\max}_{i,\text{Inc}} \\ 0 & \text{otherwise} \end{cases} \\ \frac{\partial \sigma}{\partial s} &= \begin{cases} -w_i & s = \rho^{\text{Out}}_{i,k}, w_i\left(\rho^{\max}_i - \rho^{\text{Out}}_{i,k}\right) \le F^{\max}_{i,\text{Out}} \\ 0 & \text{otherwise} \end{cases} \\ \frac{\partial d}{\partial s} &= \begin{cases} u_{i,k} & s = \rho^{\text{On}}_{i,k}, \rho^{\text{On}}_{i,k} \le F^{\max}_{i,\text{On}} \\ \min\left(\rho^{\text{On}}_{i,k}, F^{\max}_{i,\text{On}}\right) & s = u_{i,k} \\ 0 & \text{otherwise} \end{cases} \\ \frac{\partial \left\{\hat{F}_{J_i,k}\right\}_{\text{Out}}}{\partial s} &= \begin{cases} \beta_{i,k} \frac{\partial \delta}{\partial s} + \frac{\partial d}{\partial s} & \beta_{i,k}\delta + d \le \sigma \\ \frac{\partial \sigma}{\partial s} & \text{otherwise} \end{cases} \\ \frac{\partial \left\{\hat{F}_{J_i,k}\right\}_{\text{Inc}}}{\partial s} &= \begin{cases} \frac{\partial \delta}{\partial s} & \frac{\sigma p_i}{1+p_i} \le \frac{\delta}{\beta_{i,k}} \\ \beta^{-1}_{i,k}\left(\frac{\partial \left\{\hat{F}_{J_i,k}\right\}_{\text{Out}}}{\partial s} - \frac{\partial d}{\partial s}\right) & \frac{\sigma}{1+p_i} \le d \\ \frac{p_i}{\beta_{i,k}(+p_i)} \frac{\partial \sigma}{\partial s} & \text{otherwise} \end{cases} \\ \frac{\partial \left\{\hat{F}_{J_i,k}\right\}_{\text{On}}}{\partial s} &= \frac{\partial \left\{\hat{F}_{J_i,k}\right\}_{\text{Out}}}{\partial s} - \beta_{i,k} \frac{\partial \left\{\hat{F}_{J_i,k}\right\}_{\text{Inc}}}{\partial s} \end{aligned}$$

These expressions fully quantify the partial values needed in Equations 15 and 17. Thus we can construct the H_x and H_v matrices, which enables us to apply the adjoint method to the ramp metering problem.

Initial and boundary conditions

6 Numerical results for model predictive control simulations

7 Conclusions

References

[1] Claude Bardos and Olivier Pironneau. Derivatives and Control in the Presence of Shocks. pages 1–11, 2002.

[2] A.M. Bayen, R.L. Raffard, and C.J. Tomlin. Adjoint-based control of a new eulerian network model of air traffic flow. *IEEE Transactions on Control Systems Technology*, 14(5):804–818, September 2006.

[3] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*, volume 25. Cambridge University Press, 2010.

[4] W Castaings, D Dartus, M Honnorat, F Le Dimet, Y Loukili, and D Monnier. Automatic differentiation : a tool for variational data assimilation and adjoint sensitivity analysis for flood modeling. 50:249–262, 2006.

[5] C. H. Colburn, J. B. Cessna, and T. R. Bewley. State estimation in wall-bounded flow systems. Part 3. The ensemble Kalman filter. *Journal of Fluid Mechanics*, 682:289–303, August 2011.

[6] J R D Frejo and E F Camacho. Feasible Cooperation Based Model Predictive Control for Freeway Traffic Systems. (2):5965–5970, 2011.

[7] M. Garavello and B. Piccoli. *Traffic flow on networks*, volume 1. American institute of mathematical sciences Springfield, MA, USA, 2006.

[8] S K Godunov. A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics. *Matematicheskii Sbornik*, 89(3):271–306, 1959.

[9] ML. Maria Laura Modeling.

[10] RL Raffard. An adjoint-based parameter identification algorithm applied to planar cell polarity signaling. *Automatic Control, ...*, (January):109–121, 2008.

[11] José Ramón, Domínguez Frejo, and Eduardo Fernández Camacho. Global Versus Local MPC Algorithms in Freeway Traffic Control With Ramp Metering and Variable Speed Limits. 13(4):1556–1565, 2013.

[12] I S Strub and A M Bayen. Weak formulation of boundary conditions for scalar conservation laws: An application to highway traffic modelling. *International Journal of Robust and Nonlinear Control*, 16(16):733–748, 2006.

[13] Issam S. Strub, Julie Percelay, Mark T. Stacey, and Alexandre M. Bayen. Inverse estimation of open boundary conditions in tidal channels. *Ocean Modelling*, 29(1):85–93, January 2009.

[14] Andreas Wachter and Lorenz T Biegler. *On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming.* 2005.

[15] Walid. Walid ramp metering modeling.