



Jack D. Reilly
Ph.D. Student
Tel. 916-768-1755
Email: jackdreilly@berkeley.edu

Civil and Environmental Engineering
652 Sutardja Dai Hall
Berkeley, California 94720-1710

June 26, 2014

Object: Revised version of the article *Adjoint-based optimization on a network of discretized scalar conservation law PDEs with applications to coordinated ramp metering* by Reilly, Krichene, Delle Monache, Samaranayake, Goatin and Bayen, submitted on November 24, 2013 for publication in Journal of Optimization Theory and Applications.

The authors are very grateful to the two referees for their reports sent on March 19, 2014, and their careful reading of the paper. The revised version takes into account *all* their recommendations and suggestions, which we feel greatly improve the article.

We now address point by point all comments of the reviewers (presented here in *italic*).

1 Reviewer 1

The derivation of the adjoint equation is fully discrete without convergence proof or other theoretical investigations. Numerical examples for a realistic setting are given. The paper cannot yet be accepted for publication. The authors should present a revised version taking into account the following points. There is extensive literature on derivatives of hyperbolic equations with respect to controls and convergence of suitable schemes which the authors are apparently not aware of. I suggest a more intense research in the literature. Some references are given also in the following answer and should be reviewed by the authors.

The reviewer is correct in pointing out that the submitted version of the article was lacking in a discussion of the existing literature and results on the convergence of discrete adjoint systems to their continuous counterparts. We have found the suggested references to be helpful in placing the article in a broader context. We have elaborated on the suggested literature in the edits within the article and again in the responses below.

In particular, the authors should be careful with the relation of the discrete KKT system and its continuous counterpart which in the case of hyperbolic equations are not necessarily related!

We agree with the reviewer that the relationship between discrete KKT systems (analogously described in the article as the discrete adjoint equations) and the continuous counterparts is missing from the article. We have revised the article to include a discussion of the recommended articles and where our assumptions diverge from those references. The specific references and edits are discussed in the edits in response to the previous comment.

It is not correct that the formal definition is more complicated (line 53,p1). In fact, the problem is much more subtle: the cost is not differentiable in the L^1 or any L^p space. It is required to introduce a new notion of derivatives introduced in the work by Bressan and Marson [1], Bressan, Crasta and Piccoli [4] and Bressan and Shen [5]. In this sense there exists a solution. The authors should elaborate this point and add in particular the important and missing references to the correct notion of derivatives and sensitivities.

It is correct that appropriate references are missing from the discussion on adjoint approaches on conservation laws, and we are grateful for the references provided by the reviewer. The differentiability of the cost function in the setting of conservation laws using generalized tangent vectors and shift-differentiability concepts have been included in the introduction in the first paragraph. We also note the presence of junctions in our general treatment does not allow us to take advantage of these techniques. We have added a discussion of our lack of results on differentiability and convergence of the discrete KKT system in the introduction in the third and fourth paragraphs, and have emphasized that the focus of the paper is rather on a general and practical framework for applying discrete adjoint methods to conservation law networks and the linear scaling capabilities of the approach, even in the absence of rigorous theoretical guarantees. We have also removed the incorrect wording that the continuous adjoint derivation is “complicated”. We are grateful for the clarification provided by the reviewer.

Also, the adjoint equation itself might be ill-posed leading to many solutions (see James et al on conservation laws with discontinuous coefficients). Therefore, the choice of the numerical scheme is not immediately clear and has to be investigated in order to ensure to obtain an approximation to the correct adjoint solution. In fact, it is not true that any discrete system will converge to the continuous and correct derivative (of the reduced cost). Additional conditions are required on the scheme to obtain this convergence. The authors formulation however suggests that one may equivalently use discrete adjoints or even automatic differentiation and obtain an approximation to the continuous adjoints. This is not true in general. Many studies on the convergence exists and also rigorous results are presented. I strongly suggest to at least review the papers by [10, 11], or [2] on this topic and include a more careful discussion of the differences.

We agree that this point was not duly highlighted in the introduction of the first submitted version. We have followed the reviewer’s suggestion and included discussion and references about rigorous convergence results. In particular, we mention that the numerical treatment of the adjoint equation imposes a careful choice of the discretization scheme due to the possible presence of discontinuities in the solution of the forward system (27). Rigorous convergence results have been provided for Lax-Friedrichs type schemes [11] and relaxation schemes [2]. However, these results do not apply directly to our more general setting, and rigorous convergence results are out of the scope of our paper.

Another point consists of the choice of discretization of the nonlinear flux. Here, a Godunov method is employed giving rise to a highly nonlinear optimization problem. Other approaches leading to more efficient schemes use a piecewise linear or mixed-integer reformulation. Since those approaches are comparable to the presented one the authors should at least comment on the differences. Examples of the mentioned approach are found for traffic in for example [8], and for production in [6].

We thank the reviewer for directing our attention to a number of approaches to optimal control on PDE networks we did not explicitly consider. The references provided aid the article immensely in giving a broader scope to the landscape of solution methods than the article previously contained. We have added their references and a discussion of their advantages and drawbacks to the introduction section. Our approach differs from the ones above in that there is no restriction in the directionality of the shockwaves that may arise, which is vital in modeling real-world traffic patterns where ramp-metering is an effective method. While the Godunov method does give rise to nonlinear constraints, we have chosen an approach which works within these limitations to allow for a more expressive model.

Therein, also the performance for large networks is studied which yield problems of storage of data. The authors should comment on how there method scales with respect to network size (when not considering MPC).

It is true we did not discuss how the required memory scales with the size of the network, as the reviewer accurately mentions. We have added some discussion of the linear scaling rate of memory usage with respect to the number of state and control variables to both the introduction and to the sparsity discussion in Section 3.4, and mention that the result is a consequence of the ability to use sparse data structures when construction partial derivative matrices.

Section 2: This almost exactly copied from the book of Piccoli and Garavello [9] and I think it is not necessary in order to understand the scheme to introduce the definition of weak solution, Riemann problem and Riemann solver at the junction.

It is true that the presentation follows that of *Traffic Flow on Networks* by Piccoli and Garavello [9]. While the original presentation was too long, we intended this work to be viewed as a somewhat self-contained procedure one could follow to transform a control problem on PDE networks into a discrete-adjoint formulation, and as such, required at least some depth of discussion on Riemann solvers, and the Godunov discretization method. Additionally, Reviewer 2 supported the in-depth presentation of preliminaries, so we tried to reconcile both points of view.

We appreciate the advice of the reviewer and agree there were places where discussions could be shortened. We have removed the explicit description of weak solutions, as it is unused in the rest of the article. Additionally, we have removed Figure 1 from the presentation for the sake of abbreviating the preliminaries section.

Instead the authors should add which additional condition they use to solve the Riemann problem at the junction. Different have been proposed in the literature and the actually used condition should be added here. They have discrete condition in section 4.1, but for readability reasons it would be interesting to have the continuous version of these conditions in Section 2.

We fully agree that the Riemann solver section (Section 2.2) would benefit from a discussion of the additional conditions added for our application to ramp metering. Thank you for giving attention to this issue. We have added a discussion of additional conditions in the Riemann solver for ramp metering applications to the Riemann solver section, while leaving the notational and mathematical details of the conditions to the ramp metering section (Section 4.1) to avoid introducing application-specific notation too early. We mention that the conditions expressed on the Riemann solver apply equally well to the continuous and discrete case formulations of the dynamics, as the densities are assumed to be constant at the interacting boundaries.

I do not see a point in writing explicitly Algorithm 1 and 2. This already clear from the text (p8) and the algorithms should be removed.

We have removed any reference to the algorithms. We are grateful for the suggestion and believe that the presentation of the preliminaries is now more readable.

Section 3.2 and Section 3.3. This is a classical computation find in many textbooks and other papers and should be removed from the presentation. It suffices to state the KKT System for equation (14). There should be a remark that the KKT is only necessary but not sufficient and that they did not check for any constraint qualification. Maybe they can also comment on that?

We appreciate the critical analysis of the reviewer and thank them for illuminating a point missing from the article, regarding the equivalence of the adjoint presentation in the article and that which can be derived from the first-order KKT conditions. Based on the review of Reviewer 2 to include the longer presentation style, we have left some of the derivations of the adjoint in the article, and have added some additional discussion, as follows.

We agree with the reviewer that the derivation using KKT conditions belongs in the article as well for completeness and have added a discussion just after the original adjoint derivation. We thank the reviewer for the accurate observations on the lack of sufficient conditions for optimality given by the first order KKT conditions for the assumptions on our system. We have added a discussion on the issue, and reiterate the fact that the gradient computations are useful in themselves for finding descent directions to decrease the cost function, and we do not require guarantees on global or local optimality for practical purposes.

The authors should comment which continuous equations the discretizations approximate and in view of the theoretical results why this approximation is not sufficient to have convergence.

The discrete adjoint system should give an approximation of the continuous adjoint equations computed from the original continuous system of equations and weak boundary conditions. Nevertheless, the general setting adopted in our paper does not fit established convergence results, which rely upon increasing fluxes [14] and modified Lax-Friedrichs schemes that ensure convergence also in the presence of shocks [10, 11]. We have

included a discussion of these issues in the derivation of the adjoint in Section 3.2. The reviewer is correct in noting that the above discussion provides better context of existing results in adjoint calculus.

The reformulation of (35) using a penalty term is completely unnecessary and can be simply avoided by projection! It is not true that the adjoint method only works for equality constraints (see for example book of D'Apice, Goettlich et. al on supply chains where a similar computation is presented). In fact, all control constraints can be taken into account by adding a projection formula to the derivative of the Lagrangian with respect to the control (section 3.3). The authors should modify this because it is clearly wrong. I also expect better results with the modified version being independent of epsilon.

We fully agree with the arguments of the reviewer, and thank them for helping to clarify the presentation on handling geometric constraints on the control variables. The wording was incorrect with regards to the adjoint method not being able to account for inequality constraints. We have removed such wording and the discussion of ϵ and have replaced it with a much shorter discussion on handling geometric constraints with either projection methods (while adding the reference from D'Apice and Goettlich the reviewer helpfully provided) or barrier methods as an alternative.

Reviewer 2

In this paper the problem of maximization of traffic flow by optimal ramp metering for a freeway network is considered. The system is modeled by conservation laws on the corresponding networks. For the numerical solution of the optimization problem, the conservation laws are discretized by the Godunov scheme. For the nodes of the networks, the corresponding Godunov flux solutions are defined. To obtain the gradients used in the numerical optimization of the system, the adjoint of the discretized system is considered. The sparsity patterns of the resulting matrices are studied in detail. This allows to show that the complexity of the adjoint method grows only linearly with the size of the network and the number of decision variables. The application to optimal coordinated ramp metering on freeways is presented in detail. The optimal control strategy is compared with the results of a feedback control. This comparison illustrates very clearly the merits of the optimal control approach versus the feedback approach. In the application, the proposed adjoint method generates significant improvements. Also the effect of noisy input data on the quality of the results of the two control strategies is presented in detail. The paper convinces by a very clear presentation. The system of balance laws is presented with the Krushkov entropy condition and the analytical framework of the solutions in $BV \cap L^1$. The Godunov discretization is described in detail. The paper contains only one theoretical result, which demonstrates the advantages of the adjoint approach in the numerical optimization of systems with balance laws. The paper is of high value as a tutorial survey of the application of optimization in the important application of traffic flow control. Moreover, the comparison of the optimal control results with the results of the feedback control is also a very interesting contribution, since it clearly shows that if the numerical optimization is sufficiently fast (real time) which is possible due to the adjoint approach, it produces results that are superior to a feedback approach. The consideration of noisy data is extremely important. The results show that the adjoint method is able to outperform the feedback control when the noise level is less than 80 percent. In my opinion, this paper is a valuable contribution for the readers of JOTA. I recommend it for publication subject to a minor revision that takes into account the following remarks. 1. At the end of the abstract, it should be mentioned that also the robustness to noise is studied.

We thank the reviewer for their positive assessment of our work, and the nice remarks on our contributions. We have followed their suggestion of mentioning the robustness result in the abstract, and have gladly made the addition at the end of the abstract to highlight the contribution.

2. In the introduction, the applications with gas pipeline flow and water channels are mentioned. However, these systems are modeled by balance laws and not conservation laws, since friction effects have to be taken into account. So the term balance laws should appear here. Concerning gas flow, a more recent reference than [15] is [13]. Concerning the applications with water flow in networks, a recent reference is [12]

The reviewer is correct in pointing out that gas flow and water flow systems are, indeed, balance laws. We have made the correction in the revised article. We thank the reviewer for the suggestion. We also have

updated the references regarding gas flow and water flow networks to account for the more recent work being done in the field and thank the reviewer for suggesting the given references.

3. Please state in section 2.1 that f should be a C^2 -function. Comment on the fact that the trapezoidal f from Figure 7 does not satisfy this assumption.

It is true there are inconsistencies in the stated assumptions in Equation (2) and the application to ramp metering given in the paper. After further consideration, we determined from a work by Bressan [3] that only Lipschitz continuity of the flux function is required for the well-posedness of the Cauchy problem (2). The rest of the manuscript is now consistent, given that trapezoidal fundamental diagrams are indeed Lipschitz. We have fixed the assumptions on the flux function in Section 2.1 to only required Lipschitz continuity.

4. p. 5: To make the meaning of WR on line 39 clearer, please give a reference to Definition 22 here.

We fully agree with the advice of the reviewer and have added the missing reference to Definition 22 at the spot mentioned.

5. The constraints are written in the general form $H(\rho, u) = 0$. Since the cost function C is assumed to be a C^2 -function, it would be reasonable to have the same regularity for H . However, a discussion of the regularity of H is missing. Please add comments here!

We thank the reviewer for pointing out this error and fully agree the presentation is confusing and incorrect. We have corrected the discussion of C and H to require them both to belong to C^1 , which ensures the existence of partial derivatives on the relevant domain. We discuss our departure from these assumptions in the response to the issue of regularity given subsequently.

Also in section (3.2), the regularity is an essential issue. Therefore, please state directly after (15), that (15) only holds if all the required gradients exist. Please include a remark about the existence of the derivatives of the cost functional here, which is linked to the regularity of the solutions. This comes later in Note 1, but this is too late. For the application with water channel networks within the framework a classical solutions, this issue has been discussed in detail in [14]

We fully agree with the reviewer that the regularity conditions needs to be discussed centrally in the derivation of the adjoint. We have taken the advice of the reviewer and moved the note on regularity and existence of derivatives for both the cost functional and constraints to just after Equation (15). We have also elaborated on when these conditions may hold or not hold. We also thank the reviewer for providing us with a very relevant reference on the subject and have included it in Note 1 in Section 3.2.

6. p. 15, (37): After the presentation of the logarithmic barrier, please give a reference for the convergence with ϵ , for example [7]

We appreciate the advice given by the reviewer with regards to providing a reference for barrier terms, and have included the suggested reference the modified discussion.

7. Please mention in the conclusions some open questions concerning the mathematical theory, for example the convergence of the generated optimal controls that solve the discretized optimal control problem as the step-size δx converges to zero. There has been some work in this direction in the group of Michael Herty.

We are grateful for the suggestion of the reviewer to include a discussion on the convergence of the optimal control given from the discrete adjoint approach to the continuous adjoint system. The general framework considered in the paper does not allow to draw any conclusion about convergence from the existing literature [14]. Moreover, general well-posedness results are still lacking for the evolution system on networks, a major focus of the current article, making any effort in this direction useless. The authors believe that this point has been sufficiently highlighted in the discussion in the introduction, including a reference on relaxation schemes in non-network settings from the group of Michael Herty [2].

Once again we would like to thank each of the reviewers for their thoughtful comments and helpful suggestions that we believe greatly improved the paper.

Sincerely,

Jack D. Reilly, *corresponding author*

References

- [1] Bressan Alberto, Marson Andrea, Alberto Bressan, and Andrea Marson. A variational calculus for discontinuous solutions of systems of conservation laws. *Communications in partial differential equations*, 20(9):1491–1552, January 1995.
- [2] Mapundi K Banda and Michael Herty. Adjoint IMEX-based schemes for control problems governed by hyperbolic conservation laws. *Computational Optimization and Applications*, 51(2):909–930, 2012.
- [3] Alberto Bressan. Hyperbolic systems of conservation laws. 2006.
- [4] Alberto Bressan, Graziano Crasta, and Benedetto Piccoli. Well-posedness of the cauchy problem for n systems of conservation laws. *Memoirs of the American Mathematical Society*, 146(694), 2000.
- [5] Alberto Bressan and Wen Shen. Optimality conditions for solutions to hyperbolic balance laws. *Contemporary Mathematics*, 426:129–152, 2007.
- [6] Ciro D’Apice, Simone Gottlich, Michael Herty, and Benedetto Piccoli. *Modeling, Simulation, and Optimization of Supply Chains: A Continuous Approach*. SIAM, 2010.
- [7] Anthony V Fiacco and Garth P McCormick. *Nonlinear programming: sequential unconstrained minimization techniques*, volume 4. Siam, 1990.
- [8] Armin Fügenschuh, Michael Herty, Axel Klar, and Alexander Martin. Combinatorial and continuous models for the optimization of traffic flows on networks. *SIAM Journal on Optimization*, 16(4):1155–1176, 2006.
- [9] M. Garavello and B. Piccoli. *Traffic flow on networks*, volume 1. American institute of mathematical sciences Springfield, MA, USA, 2006.
- [10] Mike Giles and Stefan Ulbrich. Convergence of linearized and adjoint approximations for discontinuous solutions of conservation laws. Part 1: Linearized approximations and linearized output functionals. *SIAM Journal on Numerical Analysis*, 48(3):882–904, 2010.
- [11] Mike Giles and Stefan Ulbrich. Convergence of linearized and adjoint approximations for discontinuous solutions of conservation laws. Part 2: Adjoint approximations and extensions. *SIAM Journal on Numerical Analysis*, 48(3):905–921, 2010.
- [12] Martin Gugat. Contamination source determination in water distribution networks. *SIAM Journal on Applied Mathematics*, 72(6):1772–1791, 2012.
- [13] Martin Gugat, Markus Dick, and Günter Leugering. Gas flow in fan-shaped networks: classical solutions and feedback stabilization. *SIAM Journal on Control and Optimization*, 49(5):2101–2117, 2011.
- [14] Martin Gugat, Michael Herty, Axel Klar, and Günter Leugering. Optimal Control for Traffic Flow Networks. *Journal of Optimization Theory and Applications*, 126(3):589–616, September 2005.
- [15] B Rothfarb, H Frank, D M Rosenbaum, K Steiglitz, and D J Kleitman. Optimal design of offshore natural-gas pipeline systems. *Operations Research*, 18(6):992–1020, 1970.