Hybrid Control Laws From Convex Dynamic Programming

Hedlund, Sven; Rantzer, Anders

Published in:
Proceedings of the 39th IEEE Conference on Decision and Control, 2000.

DOI:
10.1109/CDC.2000.912810

2000

Proceedings of the 39th IEEE
Conference on Decision and Control
Sydney, Australia • December, 2000

Hybrid Control Laws From Convex Dynamic Programming

Sven Hedlund and Anders Rantzer

Department of Automatic Control
Lund Institute of Technology, Box 118, 221 00 Lund, Sweden.
Phone:(+46)-46-222 42 87, Fax:(+46)-46-138118, Email:sven@control.lth.se

Abstract

In a previous paper, we showed how classical ideas for dynamic programming in discrete networks can be adapted to hybrid systems. The approach is based on discretization of the continuous Bellman inequality which gives a lower bound on the optimal cost. The lower bound is maximized by linear programming to get an approximation of the optimal solution.

In this paper, we apply ideas from infinite-dimensional convex analysis to get an inequality which is dual to the well known Bellman inequality. The result is a linear programming problem that gives an estimate of the approximation error in the previous numerical approaches.

Keywords: optimal control, duality, convex dynamic programming, hybrid systems.

1. Introduction

One of the most important aspects of the current research activity in the field of hybrid systems is the exchange of ideas between the research fields of discrete and continuous dynamics. This paper can be viewed as an attempt to approach optimal continuous and hybrid systems using a classical linear programming perspective for discrete transportation and flow problems.

The transportation problem was formulated by Hitchcock [6] and a classic reference for network flow theory is [4]. A continuous analog is the "Monge-Kantorivich mass transfer problem" dating back to Monge in 1781 and nicely surveyed in [3].

A primal/dual formulation of a continous optimal control problem for a continuous system is presented in [9] based on the Bellman inequality. A central concept is L.C. Youngs notion of generalized flow [10].

Discretization of the Bellman inequality for numerical computations can be done in several ways [1, 8, 7, 5].

This paper is devoted to an inequality which is dual to the "Hybrid Bellman inequality" which served as basis for the computations in [5]. The dual gives valuable information about the conservatism introduced by the discretization.

In Section 2, a discrete transportation problem is discussed as a preparation for the hybrid problem of Section 3.

2. Discrete Problem Formulation

Define a discrete dynamic system as

$$q(k+1) = \nu(q(k), \mu(k)) \tag{1}$$

where $q \in Q = \{1, 2, \ldots, N\}$ is the discrete state, $\mu \in \Omega_\mu$ is the input signal of the system, and $\nu : Q \times \Omega_\mu \mapsto Q$ is a function telling what state transitions are possible.

Let $\Gamma \subset Q$ be the set of final states and consider the optimal control problem of bringing the system from an initial state, $q_0 \in Q$, to a final state, $q_f \in \Gamma$, while minimizing

$$V_\mu(q_0) = \sum_{k=1}^{k_f} s(q(k-1), q(k)) \tag{2}$$

Here $s(q,r) > 0$, $(q,r) \in S$ is the cost for switching from state q to r. The set S contains all pairs (q,r) such that a transition from mode q to mode r is possible. The time when Γ is reached is represented by the variable k_f.

The function V is commonly referred to as the value function or "cost-to-go" function of the system.

EXAMPLE 1—THE TRANSPORTATION PROBLEM
A simple discrete dynamic system is shown in Fig. 1. Here the final state is $\Gamma = \{4\}$ and the goal is to find the cheapest path from the initial state $q_0 = 1$. □

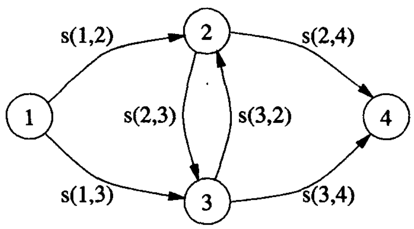

Figure 1: A simple discrete dynamic system.

Define the optimal value function $V^*(q_0) = \min_{\mu\in\Omega_\mu} V_{\mu(\cdot)}(q_0)$. A lower bound on V^* is then given by any function $V : \boldsymbol{Q} \mapsto \mathbf{R}^+$ that satisfies

$$0 \leq s(q,r) + V(r) - V(q) \qquad (q,r) \in S \quad (3)$$
$$0 = V(q) \qquad q \in \Gamma \quad (4)$$

Moreover, a bound on V^* can be found for all possible initial states simultaneously by solving *one* LP, maximizing a sum of V for those states, i.e.

$$\max_{V(q)} \sum_{q\in \boldsymbol{Q}\setminus\Gamma} \psi(q)V(q), \qquad (5)$$

where a suitable choice of ψ would be $\psi(q) = 1$.

EXAMPLE 2—THE LP APPROACH TO THE TRANSPORTATION PROBLEM

The transportation problem of Fig. 1 can be viewed as an LP problem according to (3)–(5), i.e.

$$\begin{aligned} \text{maximize} \quad & V(1) + V(2) + V(3) \\ \text{subject to} \quad & V(1) - V(2) \leq s(1,2) \\ & V(1) - V(3) \leq s(1,3) \\ & \vdots \\ & V(3) - V(4) \leq s(3,4) \\ & V(4) = 0 \end{aligned}$$

A common way to solve an LP of this structure is Dijkstra's algorithm [2]. □

2.1 Upper Bound on the Value Function via the Dual Problem

Knowing that every solution to the above problem gives a lower bound on the value function, it would be interesting to compute an upper bound as well. If the gap between the lower and upper bound is small, then the bounds are close to the optimal function. Fortunately, there is a dual LP problem that gives an upper bound on the value function.

EXAMPLE 3—DUAL LP OF THE TRANSPORTATION PROBLEM

The dual LP of the transportation problem is to

$$\begin{aligned} \text{minimize} \quad & s(1,2)\lambda_{12} + s(1,3)\lambda_{13} + s(2,3)\lambda_{23} + \\ & + s(3,2)\lambda_{32} + s(2,4)\lambda_{24} + s(3,4)\lambda_{34} \\ \text{subject to} \quad & -\lambda_{12} + \lambda_{23} - \lambda_{32} + \lambda_{24} \geq 1 \\ & -\lambda_{13} - \lambda_{23} + \lambda_{32} + \lambda_{34} \geq 1 \\ & \lambda_{12} + \lambda_{13} \geq 1 \end{aligned}$$

where $\lambda_{qr} \geq 0$ are the decision variables.

□

In general, every constraint in a primal problem appears as a variable in the dual problem. For the transportation problem, every possible transition gives rise to a constraint via the switching cost, s, and the corresponding dual variable when switching from node q to r is denoted λ_{qr}. (Conversely, every variable in the primal problem gives rise to a constraint in the dual problem, so $V(1)$, $V(2)$, and $V(3)$ correspond to the λ inequalities above.)

An interpretation of the dual problem can be given in terms of mass flow instead of the single mass unit transportation of the primal problem. The variable λ_{qr} is the flow from node q to r. There is a unit mass production in the starting states, and mass consumption in the end states. The dual problem is then to minimize the cost of the overall flow for this system. Conservation of mass implies that the production in a single node must not exceed the net flow out from the node. This corresponds to the inequality constraints for λ.

The general formulation of the dual problem to the maximization of (5) subject to (3) and (4) is thus

$$\min_{\lambda_{qr}} \sum_{(q,r)\in S} \lambda_{qr}s(q,r) \qquad (6)$$

$$\text{subject to} \sum_{r|(q,r)\in S} \lambda_{qr} - \sum_{r|(r,q)\in S} \lambda_{rq} \geq \psi(q), \ \forall q \in \boldsymbol{Q}\setminus\Gamma \qquad (7)$$

The trajectories that solve the original optimal control problem of minimizing (2) subject to the dynamics in (1) are easily found in the solutions to the primal and dual problem above. In the solution to the primal problem, the constraint (3) is active (equality holds) only for transitions (q,r) along the optimal trajectory. The same information is available in the variables of the dual problem: if there is a unique solution to the problem, then λ_{qr} is greater than zero for transitions (q,r) along the optimal trajectory and zero elsewhere.

3. Hybrid Problem Formulation

The idea of how to obtain a lower bound on the value function of the discrete dynamic system was the primal problem of maximizing the sum of the value function in a number of points (where $\psi(q) = 1$ above) subject to the constraint that the value function in two neighboring states must not differ more than the cost of switching between those states.

The dual problem to find an upper bound was interpreted as minimizing the entire mass flow in the graph subject to the constraint that the mass production in each state of the system ($\psi(q)$ above) must not exceed the net flow out from that state.

A similar primal/dual problem can also be set up for a continuous system based on this reasoning. The discrete and continuous problems can then be combined to the hybrid version presented below (including the continuous problem as a special case).

Define a hybrid system as

$$\begin{cases} \dot{x}(t) = \mathbf{f}_{q(t)}(x(t), u(t)) \\ q(t) = \nu(x(t), q(t^-), \mu(t)) \end{cases} \tag{8}$$

where $x(t) \in X \subset \mathbf{R}^n$ is the state vector, $u(t) \in \Omega_u = Co\{u^1, u^2, \dots, u^K\} \subset \mathbf{R}^m$ is a continuous input signal of the system. There is also a discrete input, $\mu(t) \in \Omega_\mu$, which allows for the selection between N different system modes, $q(t) \in \boldsymbol{Q} = \{1, 2, \dots, N\}$. The notation $q(t^-)$ is used for the left-hand limit of q at t. $S_{q,r}$ is a set (parameterized by q and r) such that switching from mode q to r is possible when $x \in S_{q,r} \subseteq X$. The continuous state, x, is constrained to a hyperrectangle $X = \{x | \underline{c}_i \le x_i \le \bar{c}_i,\ \underline{c}_i \in \mathbf{R},\ \bar{c}_i \in \mathbf{R},\ i = 1, \dots, n\}$.

The optimal control problem is to minimize the cost function

$$V_{u,\mu}(x_0, q_0) = \int_{t_0}^{t_f} \mathbf{l}_q(x, u)dt + \sum_{k=1}^{k_f} s(x(t_k), q(t_k^-), q(t_k^+)) \tag{9}$$

subject to (8) while bringing the system from an initial state (x_0, q_0) at time t_0, to a final state $(x_f, q_f)|x_f \in \Gamma_{q_f}$ at time t_f, where the end time, t_f, is free. Here, k_f is an arbitrary finite number of switches occurring at times $t_0 < t_1 < t_2 < \dots < t_{k_f} < t_f$ and $s(x, q, r) > 0$ is an associated cost for switching from discrete state q to r, the continuous part being x just before the switch. Note that $s(\cdot) > 0$ prevents the problem of infinitely many jumps in a finite interval. The final set is represented such that $\Gamma_q \subset X$ contains those x that are final in mode q. (If finishing in mode q is not allowed, then $\Gamma_q = \emptyset$.)

Sufficient conditions for a lower bound on the value function were given in [5]:

Let $V_q : X \mapsto R,\ q \in \boldsymbol{Q}$ be a set of continuous, piecewise C^1 functions that satisfy

$$0 \le \nabla V_q(x) \cdot \mathbf{f}_q(x, u) + \mathbf{l}_q(x, u) \qquad \forall x \in X,\ u \in \Omega_u,\ q \in \boldsymbol{Q} \tag{10}$$

$$0 \le V_r(x) - V_q(x) + s(x, q, r) \qquad \forall x \in S_{q,r}\ \ q, r \in \boldsymbol{Q} : q \neq r \tag{11}$$

$$0 = V_q(x) \qquad \forall (x, q)|x \in \Gamma_q \tag{12}$$

where $\mathbf{f}_q(x, u)$ gives the dynamics of a hybrid system according to (8), $\mathbf{l}_q(x, u)$ and $s(x, q, r)$ define a cost function for the system according to (9). Then, for every (x_0, q_0), $V_{q_0}(x_0)$ gives a lower bound on the cost for optimally bringing the system from (x_0, q_0) to $(x_f, q_f)|x_f \in \Gamma_{q_f}$.

3.1 Upper Bound on the Hybrid Value Function

One way of discretizing (10)–(12) to numerically obtain a lower bound of V_q was shown in [5]. The original inequalities were stated to give a lower bound on the optimal value function and the discretization was chosen to preserve this property, i.e. the solution to the discretized problem is in turn a lower bound on a function V_q that satisfies (10)–(12).

It is desirable to estimate the approximation error for the discretized problem and to grasp the importance of various discretization parameters, e.g. the grid size. The dual problem formulation, that renders an upper bound on V_q, can give such insight.

To state the dual hybrid formulation, the following assumptions are made:

$$\mathbf{f}_q(x, u) = f_q(x) + g_q(x)u \tag{13}$$

$$\mathbf{l}_q(x, u) = l_q(x) + m_q(x)u \tag{14}$$

Note that these assumptions are not as restrictive as they might look at first glance. They allow any functions $\mathbf{f}_q(x, u)$ and $\mathbf{l}_q(x, u)$ to be approximated arbitrarily well.

EXAMPLE 4—APPROXIMATION OF A QUADRATIC COST FUNCTION

Consider the cost function $\mathbf{l}_q(x, u') = u'^2$ where $u' \in [-1, 1]$. This function can be approximated by $m_q(x)u$, $u \in Co\{e_1, e_2, \dots, e_K\}$, where $e_i \in \mathbf{R}^K$ is a unit vector in the direction of the i:th coordinate axis, $m_q(x) = [u_1^2, u_2^2, \dots, u_K^2]$ is a row vector where $u_i = (2i - K - 1)/(K - 1)$.

The accuracy of the approximation increases with K. An example of $K = 5$ is shown in Fig. 2. □

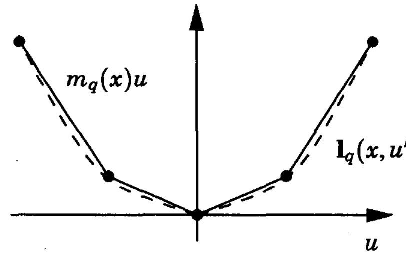

Figure 2: The solid line shows one possible approximation of u^2 (the dashed line) when $K = 5$.

THEOREM 1—UPPER BOUND ON THE INTEGRAL OF THE VALUE FUNCTION
Assume that $\rho_q^k : X \mapsto \mathbf{R}^+$, $q \in Q$, $k = 1, 2, \ldots, K$ is piecewise C^1 and $\lambda_{q,r} : X \mapsto \mathbf{R}^+$, $(q,r) \in Q \times Q$, $q \neq r$ such that

$$0 = \rho_q^k(x) \qquad x \in \partial X,\ q \in Q,\ k = 1, 2, \ldots, K$$

and

$$\psi_q(x) \leq \sum_{k=1}^{K} \nabla \cdot (\rho_q^k(x)(f_q(x) + g_q(x)u^k)) + \sum_{r|x \in S_{q,r}} \lambda_{q,r}(x) - \sum_{r|x\in S_{r,q}} \lambda_{r,q}(x), \qquad (15)$$

for all $(x,q) \in (X \backslash \Gamma_q) \times Q$.

Then the following inequality holds for every $V_q : X \mapsto R$, $q \in Q$ satisfying (10)–(12) with $\mathbf{f}_q$ and $\mathbf{l}_q$ given by (13) and (14).

$$\begin{aligned}
&\sum_q \int_{X\backslash\Gamma_q} \psi_q(x) V_q(x) dx \\
&\leq \sum_q \int_{X\backslash\Gamma_q} \left(\sum_{k=1}^{K} \rho_q^k(x)(l_q(x) + m_q(x)u^k) + \right. \\
&\qquad \left. + \sum_{r|x\in S_{q,r}} \lambda_{q,r}(x) s(x,q,r) \right) dx
\end{aligned} \qquad (16)$$

□

Remark 1. This theorem can be interpreted the same way as was done for the purely discrete case. Noting that the continuous control signal can be written as $u(x,q) = \sum_k \rho_q^k(x)u^k / \sum_k \rho_q^k(x)$, the inequality (15) corresponds to the mass production in state (x,q), $\psi_q(x)$ not exceeding the outflow (represented by the flow to other continuous states within the same discrete mode, ρ_q^k, and the flow to other modes, $\lambda_{q,r}$). The inequality (16) shows that a summation of the value function is bounded from above by the cost of the overall flow in the dual setting.

Proof. Let $f_q^k(x) \equiv f_q(x) + g_q(x)u^k$ and $l_q^k(x) \equiv l_q(x) + m_q(x)u^k$.

$$\begin{aligned}
&-\sum_{q\in Q} \int_{X\backslash\Gamma_q} \left(\sum_{r|x\in S_{q,r}} \lambda_{q,r}(x) s(x,q,r) \right. \\
&\qquad \left. + \sum_{k=1}^{K} \rho_q^k(x) l_q^k(x) \right) dx \\
&\qquad + \sum_{q\in Q} \int_{X\backslash\Gamma_q} V_q(x)\psi_q(x) dx \\
&\leq \sum_{q\in Q} \int_{X\backslash\Gamma_q} \left(\sum_{r|x\in S_{q,r}} \lambda_{q,r}(V_r(x) - V_q(x)) \right) dx \\
&\qquad + \sum_{q\in Q} \int_{X\backslash\Gamma_q} \nabla V_q(x) \cdot \left(\sum_{k=1}^{K} \rho_q^k(x) f_q^k(x) \right) dx \\
&\qquad + \sum_{q\in Q} \int_{X\backslash\Gamma_q} V_q(x) \left(\sum_{k=1}^{K} \nabla \cdot (\rho_q^k(x) f_q^k(x)) \right. \\
&\qquad \left. + \sum_{r|x\in S_{q,r}} \lambda_{q,r}(x) - \sum_{r|x\in S_{r,q}} \lambda_{r,q}(x) \right) dx \\
&= \sum_{q\in Q} \int_{X\backslash\Gamma_q} \nabla \cdot \left(V_q(x) \sum_{k=1}^{K} \rho_q^k(x) f_q^k(x) \right) dx \\
&\qquad + \sum_{q\in Q} \sum_{r\in Q} \int_{(\Gamma_r\backslash\Gamma_q)\cap S_{q,r}} \lambda_{q,r}(x) V_r(x) dx \\
&\qquad - \sum_{q\in Q} \sum_{r\in Q} \int_{(\Gamma_q\backslash\Gamma_r)\cap S_{q,r}} \lambda_{q,r}(x) V_r(x) dx \\
&= \sum_{q\in Q} \int_{\partial(X\backslash\Gamma_q)} \left(V_q(x) \sum_{k=1}^{K} \rho_q^k(x) f_q^k(x) \right) \cdot \mathbf{n} dS = 0
\end{aligned}$$

where the inequality above makes use of (10), (11), and (15).

Gauss' theorem is applied to the first equality on the last row ($\mathbf{n}$ is a unit vector that is orthogonal to $\partial(X \backslash \Gamma_q)$, pointing outwards from $X \backslash \Gamma_q$). Note that

$$\int_{(\Gamma_r\backslash\Gamma_q)\cap S_{q,r}} \lambda_{q,r}(x) V_r(x) dx = 0,$$

since $V_r(x) = 0$, $x \in \Gamma_r$ and that

$$\int_{(\Gamma_q\backslash\Gamma_r)\cap S_{q,r}} \lambda_{q,r}(x) V_r(x) dx = 0,$$

since $\Gamma_q \cap S_{q,r} = \varnothing$ (switching is not allowed from a final state). □

4. Discretization

Utilizing a computer to solve (15) for a specific control problem, a straight forward approach is to grid the state space to require the inequality to be met at a set of uniformly distributed points in X. This approximation will, however, not guarantee an upper bound on the integral of the value function, unless the nature of $f_q(x)$ *between* the grid points is taken into consideration. This can be dealt with using a method similar to the one in [5].

For readability, discretization of a purely continuous system in a two-dimensional state space is presented below ($Q = \{1\}$, $n = 2$), i.e. the discrete mode subscript q and the mode switching terms containing λ vanish. Knowing how to handle the discretization of the continuous states, the discretization is easily extended to the hybrid case.

Introduce the notation

$$\begin{aligned}
x^{jp} &= x_f + jhe_1 + phe_2 \\
X^{jp} &= \{x^{jp} + \theta_1 he_1 + \theta_2 he_2 : \ 0 \le \theta_i \le 1\} \\
\hat{X}^{jp} &= \{x^{jp} + \theta_1 he_1 + \theta_2 he_2 : \ -1 \le \theta_i \le 1\} \\
\rho^k(j,p) &= \rho^k(x^{jp}) \\
\Delta_i \rho^k(j,p) &= (\rho^k(x^{jp} + he_i) - \rho^k(x^{jp}))/h \\
\Delta_{-i} \rho^k(j,p) &= (\rho^k(x^{jp}) - \rho^k(x^{jp} - he_i))/h
\end{aligned}$$

where e_1 and e_2 are unit vectors along the coordinate axes, and h is the grid size. Define the $\min_{jp}$ operator such that

$$\min{}_{jp} f = \min_{x \in \hat{X}^{jp}} f(x)$$

and the $\max_{jp}$ operator analogously.

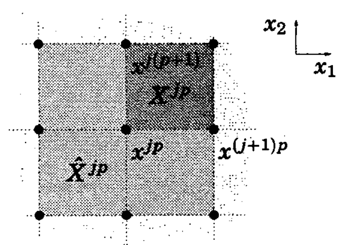

Figure 3: Illustration of X^{jp} and $\hat{X}^{jp}$.

Also introduce new variables, $\alpha_i^k(j,p) \in \mathbf{R}$ and $\beta^k(j,p) \in \mathbf{R}$ for $k = 1,2,\ldots,K$, $i = 1,2$, and (j,p) such that $x^{jp} \in X\backslash\Gamma$. The inequality (15) can then be replaced by the following combination of backward and forward difference approximations that should hold for all $k = 1,2,\ldots,K$, $i = -2,-1,1,2$, and (j,p) such that $x^{jp} \in X\backslash\Gamma$:

$$\max{}_{jp}\psi \le \sum_{k=1}^{K}\left(\sum_{i=1}^{2}\alpha_i^k(j,p) + \beta^k(j,p)\right) \tag{17}$$

$$\beta^k(j,p) \le \min{}_{jp}\{\nabla\cdot(f + gu^k)\}\rho^k(j,p) \tag{18}$$

$$\beta^k(j,p) \le \max{}_{jp}\{\nabla\cdot(f + gu^k)\}\rho^k(j,p) \tag{19}$$

$$\alpha_i^k(j,p) \le \min{}_{jp}\{f_{|i|} + g_{|i|}u^k\}\Delta_i\rho^k(j,p) \tag{20}$$

$$\alpha_i^k(j,p) \le \max{}_{jp}\{f_{|i|} + g_{|i|}u^k\}\Delta_i\rho^k(j,p) \tag{21}$$

For $x = x^{jp} + \theta_1 he_1 + \theta_2 he_2 \in X^{jp}$, $k = 1,2,\ldots,K$, define the interpolating functions

$$\begin{aligned}
\rho^k(x) =& (1-\theta_1)(1-\theta_2)\rho^k(j,p) \\
&+ \theta_1(1-\theta_2)\rho^k(j+1,p) \\
&+ (1-\theta_1)\theta_2\rho^k(j,p+1) \\
&+ \theta_1\theta_2\rho^k(j+1,p+1)
\end{aligned} \tag{22}$$

The following result applies.

THEOREM 2—DISCRETIZATION IN $\mathbf{R}^2$
If $Q = \{1\}$ and $\rho^k(j,p)$ satisfy (17)–(21) for all grid points $x^{jp} \in X \subset \mathbf{R}^2$ such that X^{jp} intersects X, then the interpolating functions $\rho^k(x)$ defined by (22) satisfies (15) and an upper bound of $\int_{X\backslash\Gamma} V(x)\psi(x)dx$ is given by (16). $\square$

Applying the above discretization scheme to a couple of examples, the resulting problem often seem to be ill conditioned. The reason for this is likely the inequalities (17)–(21) being to conservative. Various other discretization schemes will be tried.

5. Summary

We have derived an inequality which is dual to the "Hybrid Bellman inequality" presented in an earlier paper. The dual optimization problem has a simple physical interpretation in terms of particle flows. For a given control law one should envision particles flowing along the system trajectories everywhere in the state space. In a steady state situation, with particle production everywhere, the concentration of particles must be infinite near the equilibrium. The dual linear programming problem is stated in terms of the particle concentrations. It can be viewed as a generalization of the classical flow problems in discrete optimization to the case of hybrid systems.

The dual gives an upper bound on the optimal cost and thus contains valuable information about the conservatism introduced in the discretization of the primal problem. A discretization scheme that preserves the upper bound property has been proposed. Numerical problems call for further research on alternate discretization.

6. References

[1] M. S. Branicky and S. K. Mitter. "Algorithms for optimal hybrid control." In *Proceedings of the 34th Conference on Decision & Control*, New Orleans, 1995.

[2] T. H. Cormen, C. E. Leiserson, and R. L. Rivest. *Introduction to Algorithms.* MIT Press, 1989.

[3] L. C. Evans. "Partial differential equations and Monge-Kantorovich mass transfer.". Department of Mathematics, University of California, Berkeley, 1997.

[4] L. K. Ford and D. K. Fulkerson. *Flows in Networks.* Princeton University Press, Princeton, New Jersey, 1962.

[5] S. Hedlund and A. Rantzer. "Optimal control of hybrid systems." In *IEEE Conference on Decision and Control*, Phoenix, 1999.

[6] F. L. Hitchcock. "The distribution of a product from several sources to numerous localities." *J. Math. Phys.*, **20**, pp. 224–230, 1941.

[7] A. Rantzer. "Dynamic programming via convex optimization." In *Proceedings of the IFAC World Congress*, Beijing, 1999.

[8] A. Rantzer and M. Johansson. "Piecewise linear quadratic optimal control." In *Proceedings of American Control Conference*, Albuquerque, 1997.

[9] R. Vinter. "Convex duality and nonlinear optimal control." *SIAM J. Control and optimization*, **31:2**, pp. 518–538, March 1993.

[10] L. C. Young. *Lectures on the Calculus of Variations and Optimal Control Theory.* W. B. Saunders Company, Philadelphia, Pa, 1969.