Preprints of the 18th IFAC World Congress
Milano (Italy) August 28 - September 2, 2011

Finite-Time State-Constrained Optimal Control for Input-Affine systems with Actuator Noise

Per Rutquist * Claes Breitholtz ** Torsten Wik **

* *Tomlab Optimization AB, Västerås, Sweden*
** *Department of Signals and Systems, Chalmers University of Technology, Göteborg, Sweden*

Abstract: We show that a linearizing transformation of the Hamilton-Jacobi-Bellman (HJB) equation can be applied to certain finite-time problem such that the time dependence can be separated and also has a simple analytical solution. The remaining state dependence is the solution to a linear eigenvalue problem that may have an analytical solution or is readily solved numerically. The efficiency of the method is illustrated by an inventory control problem.

Keywords: Stochastic optimal control; Dynamic programming; Hamilton-Jacobi-Bellman equation.

1. INTRODUCTION

Consider a nonlinear, state constrained, input-affine, time invariant, stochastic system where the disturbance enters the system in the same way as the control input:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{G}(\mathbf{x})\,(\mathbf{u} + \dot{\mathbf{w}})\,, \tag{1}$$

where $\mathbf{x} \in \Omega \subset \mathbb{R}^n$ is the system state, $\mathbf{u} \in \mathbb{R}^m$ is the control signal, $\mathbf{f} : \Omega \to \mathbb{R}^n$ and $\mathbf{G} : \Omega \to \mathbb{R}^{n\times m}$ are functions that describe the system dynamics, and $\dot{\mathbf{w}} \in \mathbb{R}^m$ is a Gaussian white noise having the covariance matrix $\mathbf{W} \in \mathbb{R}^{m\times m}$ and $\mathbf{w}$ is a Wiener process. The boundary $\partial\Omega$ of the state space Ω defines the state constraints, i.e. the system must be controlled such that $\mathbf{x}$ never leaves Ω.)

This type of systems is common in, for example, process control where stochastic feed variations are the major disturbances, and also in cases when the actuators themselves are the dominating source of noise.

Optimal control problems for (1) result in a nonlinear HJB-equation that is numerically problematic to solve by standard PDE solvers because of the nonlinearity and the infinite boundary conditions on $\partial\Omega$ (due to the state constraints). For the infinite horizon case though, Rutquist et al. [2008] have derived an efficient method based on a transformation of the cost functional. However, there are many situations where the non-stationary solution, i.e. the optimal control problem with a finite end-time, is of interest. For example finite time buffer problems, which are typically applicable in process control, both for batch processes and for continuous processes with stop time, as in one-shift operation.

Solution of state constrained deterministic finite time optimal control problems have been extensively treated in connection with discrete time Model Predictive Control (Jones et al. [2007], Borelli et al. [2003]) and for linear time-invariant systems explicit piecewise affine control laws can be determined (Bemporad et al. [2002]). Working in discrete time for systems with stochastic disturbances, however, imposes problems related to constraint violation. Unless the disturbance is either bounded or known beforehand, there may not exist any control that is guaranteed to keep the system within bounds in the next time step. Bemporad et al. [2002] therefore suggest that safety margins are introduced.

Here, we consider the continuous time case and derive an explicit control law that guarantees optimality and no constraint violation. By using the linearizing transformation introduced in (Rutquist et al. [2008]) the HJB-equation is transformed into a form that allows for solution by separation of variables. The time dependent factors have a simple analytical solution and the remaining state-dependent factors are the solution to a readily solved linear eigenvalue problem with zero Dirichlet boundary conditions.

2. PROBLEM FORMULATION

Now, consider the fixed final time control problem, where the goal is to find a feedback control policy $\mathbf{u}(t, \mathbf{x})$ that minimizes

$$V(x(t), t) = \mathrm{E}\left\{ \int_t^{t_f} \left(l(\mathbf{x}(\tau)) + \mathbf{u}(\tau)^T \mathbf{Q} \mathbf{u}(\tau) \right) \,\mathrm{d}\tau + V_f(\mathbf{x}(t_f)) | x(t) \right\} \tag{2}$$

where $t \in \mathbb{R}$ is the current time, $t_f > t$ is the final time, $l : \Omega \to \mathbb{R}$ describes the (time independent) cost (non-singular on Ω) associated with the state, and $V_f : \Omega \to \mathbb{R}$ is the final cost.

The positive definite symmetric matrix $\mathbf{Q} \in \mathbb{R}^{m\times m}$ defines the cost of the control signal. A natural choice is

$$\mathbf{Q} = \kappa \mathbf{W}^{-1}\,, \tag{3}$$

where $\kappa \in \mathbb{R} > 0$ is an arbitrary constant, since this corresponds to a lower cost of control for input nodes with large noise. The parameter κ can be used to tune the trade-off between performance and controller actuation.

3. LINEARIZATION OF THE HJB EQUATION

In the following all vectors are column vectors with the exception of ∇V, which is defined as a row vector.

The stochastic HJB equation for the minimization of (2) can be formulated as

$$-\frac{\partial V}{\partial t} = \min_{\mathbf{u}}\{l + (\nabla V)(\mathbf{f} + \mathbf{G}\mathbf{u}) + \mathbf{u}^T\mathbf{Q}\mathbf{u} + \frac{1}{2}\mathrm{tr}\big[(\nabla^T\nabla V)\mathbf{G}\mathbf{W}\mathbf{G}^T\big]\} \tag{4}$$

where $V : [t, t_f] \times \Omega \to \mathbb{R}$ is the so-called cost-to-go function (see for instance Dorato et al. [1995] for details on the derivation of this equation). Applying $\mathrm{tr}[AB] = \mathrm{tr}[BA] = \mathrm{tr}\big[(AB)^T\big]$ and the fact that the covariance matrix $\mathbf{W}$ is symmetric it is readily verified that this formulation equals that of Hanson [2007], for example. Minimization of this quadratic expression with respect to u gives the optimal control input

$$\mathbf{u} = -\frac{1}{2}\mathbf{Q}^{-1}\mathbf{G}^T(\nabla V)^T\,. \tag{5}$$

Insertion into (4) gives

$$-\frac{\partial V}{\partial t} = l - \frac{1}{4}(\nabla V)\mathbf{G}\mathbf{Q}^{-1}\mathbf{G}^T(\nabla V)^T + (\nabla V)\mathbf{f} + \frac{1}{2}\mathrm{tr}\big[(\nabla^T\nabla V)\mathbf{G}\mathbf{W}\mathbf{G}^T\big] \tag{6}$$

with infinite boundary conditions on $\partial\Omega$ due to the state constraints. The solution to this equation inserted into (5) gives the optimal control. However, this nonlinear partial differential equation has in general no analytic solution and is costly to solve numerically.

Rutquist et al. [2008] showed that the variable transformation

$$V = -2\kappa\log(Z), \tag{7}$$

where $Z \in ([t, t_f] \times \Omega \to \mathbb{R} > 0)$, linearizes the stationary HJB-equation. Analogously it is shown that it also transforms equation (6) into the linear partial differential equation

$$\frac{\partial Z}{\partial t} = \frac{l}{2\kappa}Z - (\nabla Z)\mathbf{f} - \frac{1}{2}\mathrm{tr}\big[(\nabla^T\nabla Z)\mathbf{G}\mathbf{W}\mathbf{G}^T\big] \tag{8}$$

with boundary conditions

$$Z = 0, \quad \mathbf{x} \in \partial\Omega \tag{9}$$

and

$$Z_f \triangleq Z(t_f, \mathbf{x}) = \exp\left(\frac{-V_f}{2\kappa}\right). \tag{10}$$

For some systems, the HJB equation (4) does not have a solution in the classical (continuously differentiable) sense, and in such cases neither does (8). Both equations may still have viscosity solutions (Crandall et al. [1992]) that can be used.

4. VARIABLE SEPARATION

In contrast to (6), Equation (8) is linear and also in a form that allows for solution by separation of variables (Folland [1992]). The solutions then have the form $Z(t, \mathbf{x}) = T(t)\phi(\mathbf{x})$, where $T : ([t, t_f] \to \mathbb{R})$ is independent of $\mathbf{x}$, and $\phi : \Omega \to \mathbb{R}$ is independent of time. This gives

$$\phi\frac{\mathrm{d}T}{\mathrm{d}t} = T\left(\frac{l}{2\kappa}\phi - (\nabla\phi)\mathbf{f} - \frac{1}{2}\mathrm{tr}\big[(\nabla^T\nabla\phi)\mathbf{G}\mathbf{W}\mathbf{G}^T\big]\right) \tag{11}$$

which can be separated into

$$\frac{1}{T}\frac{\mathrm{d}T}{\mathrm{d}t} = \lambda\,, \tag{12}$$

and

$$\lambda = \frac{l\phi - 2\kappa(\nabla\phi)\mathbf{f} - \kappa\mathrm{tr}\big[(\nabla^T\nabla\phi)\mathbf{G}\mathbf{W}\mathbf{G}^T\big]}{2\kappa\phi}\,. \tag{13}$$

The boundary condition (9) is translated to

$$\phi = 0, \quad \mathbf{x} \in \partial\Omega. \tag{14}$$

The time-dependent part (12) has an explicit analytical solution,

$$T(t) = \gamma e^{\lambda t}, \tag{15}$$

where γ is an arbitrary constant, while the state-dependent part (13) may need to be solved numerically. However, (13) is a linear eigenvalue problem that is readily solved compared to (6). Solving this eigenvalue problem one obtains a family of solutions (λ_n, T_n, ϕ_n). The solution to (8), (9) and (10) is then a linear combination of the solutions for different eigenvalues λ_n. We can write

$$Z(t, \mathbf{x}) = \sum_{n=1}^{\infty}\beta_n\exp(-\lambda_n(t_f - t))\phi_n(\mathbf{x})\,, \tag{16}$$

where the coefficients $\beta_n \in \mathbb{R}$ are given by the projection of the final condition (10) onto the space spanned by the eigenfunctions ϕ_n.

$$\langle\phi_j, Z_f\rangle = \sum_{n=1}^{N}\beta_n\langle\phi_j, \phi_n\rangle\ \forall j = 1, \ldots N, \tag{17}$$

where $\langle\phi_n, Z_f\rangle$ is the scalar product of ϕ_n and Z_f on Ω.

For systems where the eigenfunctions are orthogonal the projection results in

$$\beta_n = \frac{\langle\phi_n, Z_f\rangle}{\langle\phi_n, \phi_n\rangle}\,, \quad n \in \mathbb{N}. \tag{18}$$

For first order systems we may use this to get closer to an analytical solution since the eigenfunctions are always orthogonal then. To see this we begin by writing (13) as

$$\frac{1}{2}G^2W\phi'' + f\phi' - \alpha\phi = -\lambda\phi, \tag{19}$$

where $\alpha = l(x)/(2\kappa)$.

A problem on the form

$$D(p(x)D)\phi + q(x)\phi = \lambda r(x), \tag{20}$$

with $p \neq 0$, q and $r > 0$ real-valued, and p, p' and q continuous, is a Sturm-Liouville problem (Griffel [1992]). In the scalar case we may write

$$p\phi'' + p'\phi' + q\phi = \lambda r\phi. \tag{21}$$

Multiplication by $G^2W/(2p)$, $p \neq 0$, gives

$$\frac{G^2W}{2}\phi'' + \frac{G^2W}{2}\frac{p'}{p}\phi' + \frac{qG^2W}{2p}\phi = \lambda\frac{G^2W}{2p}r\phi. \tag{22}$$

Identifying the coefficients, noting that W is positive because it is a variance and assuming $G \neq 0$, gives

$$\begin{cases} p' = \dfrac{2f}{G^2W}p \\ q = -\dfrac{2\alpha}{G^2W}p \\ r = -\dfrac{2p}{G^2W} \end{cases} \Rightarrow \begin{cases} p = ke^{\Gamma(x)} \\ q = -\dfrac{l}{G^2W}ke^{\Gamma(x)} \end{cases} \quad (23)$$

where

$$\Gamma(x) = \int \frac{2f}{G^2W}dx, \quad (24)$$

and k is an arbitrary constant we may set to $-W/2$, for example. Hence, our eigenvalue problem is a Sturm-Liouville problem for which the eigenfunctions are always an orthogonal basis in the Hilbert space $L_2(0,1,r)$, with the scalar product (see e.g. Griffel [1992])

$$\langle v_1, v_2 \rangle \equiv \int_0^1 v_1 \bar{v}_2 r dx. \quad (25)$$

5. CONTROL POLICY

Inserting $\nabla V = -(2\kappa/Z)\nabla Z$ into (5) yields the optimal control

$$\mathbf{u} = \frac{1}{Z}\mathbf{W}\mathbf{G}^T(\nabla Z)^T. \quad (26)$$

If Z is an exact solution to (8), then $Z = 0$ corresponds to $V = \infty$ so the division by Z in (26) does not introduce any difficulty that was not in the original problem formulation. [1] For approximate solutions to (8), Z may change sign in some part of the state-space, and that should then be taken as a warning signal that the resulting control policy is very bad. Figure 4 shows an example of this.

If Z is a viscosity solution, then $\mathbf{u}$ may be discontinuous in some points. Equation (26) cannot be used to compute $\mathbf{u}$ in those points. Instead, $\mathbf{u}$ is computed using the right hand side of (4), taking ∇V in the direction of travel for each possible $\mathbf{u}$.

It can be noted that the policy (26) is unchanged if Z is multiplied by any function that is independent of $\mathbf{x}$. An implication of this is that the parameter γ in the analytical solution for $T(t)$ has no effect on the control. We may also multiply (16)

by $\beta_1^{-1}\exp(\lambda_1(t_f - t))$ and then use

$$Z_r(t,\mathbf{x}) = \phi_1(\mathbf{x}) + \sum_{n=2}^{\infty}\frac{\beta_n}{\beta_1}\exp(-(\lambda_n - \lambda_1)(t_f - t))\phi_n(\mathbf{x}). \quad (27)$$

This formulation is numerically more robust, as it avoids the risk of floating point underflow.

Because $(\lambda_n - \lambda_1) > 0, \forall n > 1$, all terms except ϕ_1 vanish as the "remaining time" $(t_f - t)$ approaches infinity. So for infinite-horizon optimal control, only the eigenfunction corresponding to the lowest eigenvalue needs to be computed, which is the stationary solution derived by Rutquist et al. [2008].

6. EXAMPLE

As an example, we consider the following problem: A chemical plant contains a buffer tank, which regulates the flow of a certain chemical. The inflow varies stochastically as a result of upstream events, while the outflow is controlled to be as steady as possible. Specifically, we want to minimize u^2 over time, where u is the deviation from desired outflow. The buffer state x is allowed to vary in the interval 0 (empty) to 1 (full).

The system is then described by

$$\dot{x} = u + \mu + \sigma\dot{w}, \quad x \in (0,1), \quad (28)$$

where $\dot{w} \in \mathbb{R}$ is a random disturbance, which we model as a white noise with zero mean and intensity 1, and μ is the average net filling rate from the buffer. [2]

The plant operates in batch mode, and at the end of the batch, at time t_f all the chemical remaining in the buffer will have to be discarded. The optimal control problem is therefore to minimize both variation in outflow during operation from current time t to t_f and the level in the buffer at the end. Mathematically the goal is to minimize

$$V(x(t),t) = \int_t^{t_f} u^2 \mathrm{d}\tau + ax(t_f), \quad (29)$$

where the constant a is the specific cost of the discarded chemical.

This problem can be expressed by equations (1), (2) and (3) if we let $\mathbf{f} = \mu$, $\mathbf{G} = 1$, $l = 0$, $\mathbf{Q} = 1$ and $\mathbf{W} = \sigma^2$. Equation (29) then implies $\kappa = \sigma^2$.

In order to compute an optimal control policy, we first solve the eigenvalue problem given by (13) and (14), i.e.

$$\lambda\phi = -\mu\frac{\mathrm{d}\phi}{\mathrm{d}x} - \frac{\sigma^2}{2}\frac{\mathrm{d}^2\phi}{\mathrm{d}x^2} \quad (30)$$

$$\phi(0) = \phi(1) = 0. \quad (31)$$

This can be reformulated as a Sturm-Liouville problem with the solutions

$$\lambda_n = \frac{\mu^2}{2\sigma^2} + \frac{1}{2}n^2\pi^2\sigma^2, \quad n = 1,2,3,\ldots \quad (32)$$

$$\phi_n = \exp\left(-\frac{\mu}{\sigma^2}x\right)\sin(n\pi x) \quad (33)$$

where the eigenfunctions form an orthogonal basis for the Hilbert space $L_2(0,1,\exp(2\mu x/\sigma^2))$.

Next, we compute the projection of the final condition Z_f onto the space spanned by the ϕ_n, using (18).

Since $V_f = \alpha x$, by (10) we have $Z_f = \exp(-\alpha x/2\kappa)$,

which gives

$$\beta_n = \frac{\int_0^1 \phi_n(x)\exp(-\frac{ax}{2\sigma^2} + \frac{2\mu x}{\sigma^2})\mathrm{d}x}{\int_0^1 \phi_n^2(x)\exp(\frac{2\mu x}{\sigma^2})\mathrm{d}x}. \quad (34)$$

and the optimal control

$$\mathbf{u} = \sigma^2\frac{\sum_{n=1}^{\infty}\beta_n\exp(-\lambda_n(t_f - t))\phi_n'(\mathbf{x})}{\sum_{n=1}^{\infty}\beta_n\exp(-\lambda_n(t_f - t))\phi_n(\mathbf{x})}. \quad (35)$$

Figures 1 – 3 show an example of what Z and u look like for this inventory control problem. In this case $\mu = -0.2$, $\sigma = 0.3$ and $a = 1$. The plots show the control policies for $t_f - t$ equal to 10, 1, and 0.1. It can be seen how the control

[1] Boundary conditions of $Z = 0$ is not a problem, because the system state will never reach the boundary, so the control need not be computed there.

[2] It is possible to eliminate μ from the calculations by a change of variables, but we will not do that in this example.

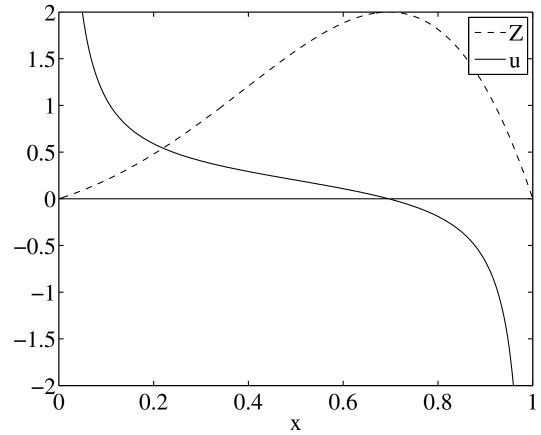

Figure 1. The log-transformed value function Z and the control policy u at $t - t_f = 10$. At this point, only the first term in the series is of any consequence. The long-term control policy has $u = -\mu$ when the buffer is half-full.

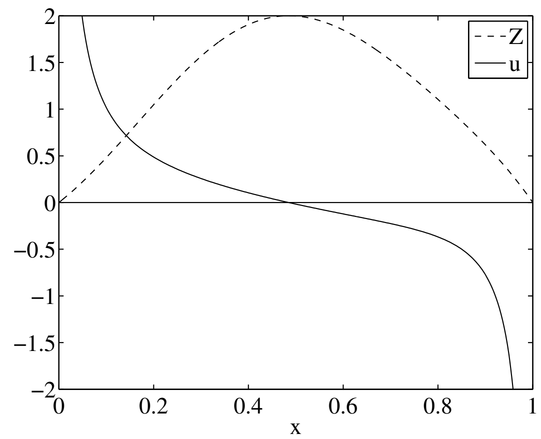

Figure 2. The result at $t_f - t = 1$. The control is starting to favor lower x.

policy moves towards smaller x as time passes, while the buffering capability is maintained.

The sums in (35) converge rapidly on most of the time interval because of the exponential factors. When $t_f - t$ approaches zero, however, a practical difficulty may arise as an increasing number of terms need to be included in the sum. Truncating too early can result in very bad control policies, as shown in Figure 4. The effect is very strong in this example, because the final condition (the buffer should be as empty as possible) is in conflict with the continuous operation condition (the buffer must never be empty). The resulting discontinuity in V (and hence Z) makes the Fourier series converge very slowly.

In practice, this problem can be circumvented by using more terms and/or slightly increasing t_f in the calculations and then interrupt in advance (at the proper t_f). It also helps to select the final cost V_f compatible with the buffer boundary constraints (that is: V_f approaching infinity at $x = 0$ and $x = 1$). With a continuous boundary condition, the Fourier series will converge much more rapidly.

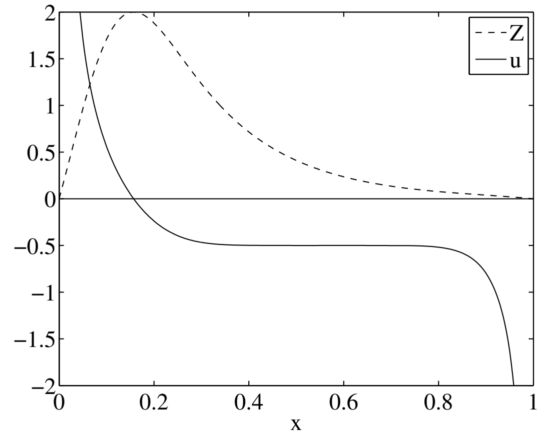

Figure 3. At $t_f - t = 0.1$, the control policy favors even lower x. Still the level is not allowed to drop too far, as buffering capability must be maintained until the end. (At this point, the Fourier series converges slowly. In this graph, 15 terms are used, which is sufficient.)

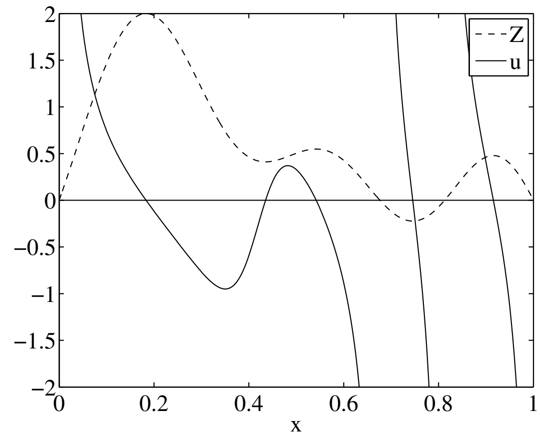

Figure 4. The result of truncating the series too soon. The value function and control computed for $t_f - t = 0.1$ using only 5 terms of the Fourier series. (Note that Z passes below zero near $x = 0.7$ causing u to become infinite, which is why there are gaps in the plot of u.)

We note that this method requires the least computational effort to compute when the final time is far away. The closer to the final time, the more terms must be included in the sums in order to obtain the optimal control. This makes the solution an excellent complement to strategies such as model predictive control, which requires more computing power the farther away the time horizon is.

7. CONCLUSION

A previously presented method for the solution of state constrained stationary optimal control problems has been extended to the finite-time case. The time dependence is handled by separation of variables which gives an analytical solution for the time varying factors. The state dependent factors are given by the solution to a linear eigenvalue problem, which can readily be solved analytically or numerically if the original optimal control problem is formulated properly.

REFERENCES

A. Bemporad, M. Morari, V. Vivek, and E. N. Pistikopoulos. The explicit linear quadratic regulator for constrained systems. *Automatica*, 38(1):3–20, January 2002. doi: 10.1016/S0005-1098(01)00174-1.

F. Borelli, A. Bemporad, and M. Morari. Geometric algorithm for multiparametric linear programming. *Journal of optimization theory and applications*, 118(3):152–170, 2003.

Michael G. Crandall, Hitoshi Ishii, and Pierre L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bulletin of the American Mathematical Society*, 27(1):1–67, 1992.

P. Dorato, T. C. Abdallah, and V. Cerone. *Linear Quadratic Control: An Introduction*. Prentice-Hall, 1995. ISBN 1575241560.

G. B. Folland. *Fourier Analysis and Its Applications (Wadsworth and Brooks/Cole Mathematics Series)*. Brooks Cole, January 1992. ISBN 0534170943.

D. H. Griffel. *Applied functional analysis*. Ellis Horwood Ltd., England, 1992. ISBN 0-85312226-1.

F. B. Hanson. *Applied Stochastic Processes and Control for Jump-Diffusion*. SIAM, Philadelphia, US, 2007. ISBN 978-0-898716-33-7.

C.N. Jones, M. Baric, and M. Morari. Multiparametric linear programming with application to control. *European Journal of Control*, 13:152–170, 2007.

P. Rutquist, C. Breitholtz, and T. Wik. On the infinite time solution to state-constrained stochastic optimal control problems. *Automatica*, 44:1800–1805, 2008. doi: 10.1016/j.automatica.2007.10.018.