

# Performance Bounds and Suboptimal Policies for Linear Stochastic Control via LMIs

Yang Wang <sup>\*</sup>      Stephen Boyd <sup>†</sup>

May 25, 2009

## Abstract

In a recent paper, the authors showed how to compute performance bounds for infinite horizon stochastic control problems with linear system dynamics and arbitrary constraints, objective, and noise distribution. In this paper we extend these results to the finite-horizon case, with asymmetric costs and constraint sets. The method is based on bounding the objective with a general quadratic function, and using linear matrix inequalities (LMIs) and semidefinite programming (SDP) to optimize the bound. The resulting LMIs are more complicated than in the previous paper (which only used quadratic forms) but this extension allows us to obtain good bounds for problems with substantial asymmetry, such as supply chain problems. The method also yields very good suboptimal control policies, using control-Lyapunov methods.

---

<sup>\*</sup>Yang Wang is supported by a Rambus Corporation Stanford Graduate Fellowship.

<sup>†</sup>This material is based upon work supported by the Precourt Institute on Energy Efficiency, by NSF award 0529426, by NASA award NNX07AEI1A, by AFOSR award FA9550-06-1-0514, and by AFOSR award FA9550-06-1-0312.

# Contents

- 1 Introduction** **3**
  - 1.1 Prior and related work . . . . . 4
  - 1.2 Outline . . . . . 5
  
- 2 Finite horizon** **5**
  - 2.1 Dynamic programming ‘solution’ . . . . . 6
  - 2.2 Basic bound . . . . . 7
  - 2.3 Bellman inequality as an LMI . . . . . 9
  - 2.4 Optimizing the bound . . . . . 9
  - 2.5 Finite input constraint set . . . . . 10
  - 2.6 S-procedure relaxation . . . . . 11
  - 2.7 Control-Lyapunov policy . . . . . 12
  
- 3 Infinite horizon** **13**
  - 3.1 Dynamic programming ‘solution’ . . . . . 14
  - 3.2 Basic bound . . . . . 16
  - 3.3 Bellman inequality as an LMI . . . . . 17
  - 3.4 Optimizing the bound . . . . . 18
  - 3.5 Finite input constraint set . . . . . 18
  - 3.6 S-procedure relaxation . . . . . 19
  - 3.7 Control-Lyapunov policy . . . . . 20
  
- 4 Examples** **20**
  - 4.1 Small example . . . . . 20
  - 4.2 Nonnegative control . . . . . 21
  - 4.3 Supply chain . . . . . 21
  
- 5 Conclusions** **22**

# 1 Introduction

In this paper we consider a stochastic control problem with linear dynamics, and arbitrary objective and constraint sets. This problem can be effectively solved in only a few cases. For example, when the objective is quadratic and there are no constraints, it is well known that the optimal control is linear state feedback [Kal64, Ber05, Ber07]. In other cases, when the problem cannot be solved analytically, many methods can be used to find suboptimal controllers, *i.e.*, one that achieves a small objective value. While this paper does not focus on suboptimal policies, one suboptimal control that we will discuss in more detail is called the control-Lyapunov policy (CLF), sometimes also known as approximate dynamic programming (ADP) [Son83, FP96, CL88, SSC03]. In CLF, the control policy is obtained by replacing the true value function for the stochastic control problem with a computationally tractable approximation. We will see later that our lower bound naturally yields an approximate value function for use in a control-Lyapunov policy; examples suggest that this control policy achieves surprisingly good performance. For more detailed discussion of suboptimal policies, see, *e.g.*, [Ber05, Ber07, Pow07, Åst70, Whi82, GPM89, BMDP02, Mac02, QB03, Soe92, CB04, KH05, WB08, Wri97a].

We present a method for computing a numerical lower bound on the optimal objective value for the linear stochastic control problem. Our bound is not generic, *i.e.*, it does not depend only on the problem dimensions and some basic assumptions about the objective and constraints. Instead, the bound is computed for each specific problem instance. We will see that for many practical control problems, the bound can be effectively computed by solving a convex optimization problem. Thus, the complexity of computing the bound does not grow significantly with the problem dimensions, and avoids the ‘curses of dimensionality’.

The bound we compute can be compared to the objective achieved by suboptimal control policies, which can be found through Monte Carlo simulation. If the gap between the bound and the objective achieved by a suboptimal policy is small, we can confidently conclude that our suboptimal policy is nearly optimal, and that our bound is nearly tight. On the other hand, if the gap between the two is large, then either our suboptimal controller is substantially suboptimal, or our lower bound is poor (for this problem instance). We cannot (at this time) guarantee that our bound will be close to the optimal objective value. However, in a large number of numerical simulations, we have found that the bound we compute is often close to the objective achieved by a suboptimal control policy.

In a previous paper [WB09b], we presented a special case of our lower bound, where the constraint sets and objective functions are symmetric. In this paper we present a more general method that produces better bounds for problems with substantial asymmetry. The main idea in our method is to bound the objective and value functions with general quadratic functions, and then optimize the bound over this family of quadratic functions. We will illustrate our bound with several numerical examples. In all cases, we find that the bound we compute is close to the objective achieved by a control-Lyapunov policy, which shows that both are nearly optimal.

## 1.1 Prior and related work

Previous work related to performance bounds can be found in several areas. In approximate dynamic programming, a common approach is to parameterize the approximate value function using a set of basis functions, and then to find a combination of these basis functions with a guarantee on the distance from the optimal solution. For example, in [LR06], the authors represent the approximate value function as a combination of simple (*i.e.*, linear, quadratic) basis functions. They use an iterative method for adding basis functions to this set, based on a modified value iteration with relaxed stage costs. This gives lower and upper bounds within a prespecified distance from the true value function, and can be effectively applied to many practical problems. Another work closely related to ours is [DV03], where the authors consider a stochastic control problem with a finite number of states and inputs. In this paper, the approximate value functions are represented as a weighted sum of pre-selected basis functions. The weights are then chosen to give a lower bound on the true value function by solving a linear program (LP). Here, the lower bound property is obtained by relaxing the Bellman equation to an inequality—a technique we will also use. The authors show that as long as the basis functions are ‘well chosen’, a maximum distance from the true value function can be guaranteed.

Another area in which performance bounds have been studied is in the context of Markov decision processes and in particular, queueing systems. In [CL06], the authors derive performance bounds for Markov decision processes by finding upper and lower bounds on the average cost incurred in each period. This method is applied to a multiclass queueing system [CL06], as well as event-based sampling [CLH07], and typically yields analytic bounds that apply to entire problem classes. Another example of performance bounds in this area is [Cas05]. Here, the authors consider the problem of controlling a sensor network to minimize estimation error, subject to a sensor resource constraint. To get a lower bound, the resource constraint is ‘dualized’ by adding the constraint function into the objective, weighted by a nonnegative Lagrange multiplier. The lower bound is then optimized over the dual variable. We will see that in special cases, our bound can also be interpreted as an application of Lagrange duality.

There are also many works on deriving upper bounds on the performance of a suboptimal control policy. A common approach here is to find a quadratic Lyapunov function to establish an upper bound on the objective. This is sometimes called guaranteed cost control, and has been studied extensively in the context of robust control. We will not consider the problem of upper bounding the performance of suboptimal control policies in this paper; interested readers are referred to [PM94, CP72, BH88, LB76, ZDG95].

There are many other related works we will not summarize, including more theoretical contributions [Wit70, RZ94, SS85], other application focussed papers [KK94, BGT01, VGN04, PSSV07], as well as books on approximate dynamic programming methods and stochastic control [BT96, Pow07, Alt99]. Many of the ideas we will use appear in these, and will be pointed out.

## 1.2 Outline

The structure of our paper is as follows. In section 2 we describe our bound for the finite horizon stochastic control problem. In section 2.1 we outline the dynamic programming ‘solution’, followed by our method for finding a bound in sections 2.2–2.4. Then, in sections 2.5–2.6 we describe two cases for which our bound can be effectively computed by solving a semidefinite program (SDP), and in section 2.7 we describe the control-Lyapunov suboptimal policy. In section 3 we repeat this for the infinite horizon, average cost-per-stage problem. Finally, in section 4 we illustrate our bound with three numerical examples.

## 2 Finite horizon

We consider a discrete-time linear system, over the time interval  $t = 0, \dots, N$ , with dynamics

$$x_{t+1} = A_t x_t + B_t u_t + w_t, \quad t = 0, 1, \dots, N-1, \quad (1)$$

where  $x_t \in \mathbf{R}^n$  is the state,  $u_t \in \mathbf{R}^m$  is the control input,  $w_t \in \mathbf{R}^n$  is the process noise (or exogenous input),  $A_t \in \mathbf{R}^{n \times n}$  is the dynamics matrix, and  $B_t \in \mathbf{R}^{n \times m}$  is the input matrix, at time  $t$ . We assume that  $w_t$ , for different values of  $t$ , are independent with mean  $\bar{w}_t = \mathbf{E} w_t$ , and covariance  $W_t = \mathbf{E}(w_t - \bar{w}_t)(w_t - \bar{w}_t)^T$ . We will also assume that  $x_0$  is random, with mean  $\bar{x}_0 = \mathbf{E} x_0$ , and covariance  $X_0 = \mathbf{E}(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T$ , and that  $x_0$  is independent of all  $w_t$ .

We consider causal state feedback control policies, where the current input  $u_t$  is determined from the current and previous states  $x_0, \dots, x_t$ . For the problem we will consider, it can be shown that there is an optimal policy that depends only on the current state, *i.e.*,

$$u_t = \psi_t(x_t), \quad t = 0, 1, \dots, N-1, \quad (2)$$

where  $\psi_t : \mathbf{R}^n \rightarrow \mathbf{R}^m$  is the state feedback function, or control policy, at time  $t$ . The equations (1) and (2) determine the state and control input trajectories as functions of  $x_0$  and the process noise trajectory. Thus, for fixed choice of state feedback functions  $\psi_0, \dots, \psi_{N-1}$ , the state and input trajectories become stochastic processes.

The objective function has the form

$$J = \mathbf{E} \left( \sum_{t=0}^{N-1} \ell_t(x_t, u_t) + \ell_N(x_N) \right),$$

where  $\ell_t : \mathbf{R}^n \times \mathbf{R}^m \rightarrow \mathbf{R}$ ,  $t = 0, \dots, N-1$ , is the stage cost function at time  $t$ , and  $\ell_N : \mathbf{R}^n \rightarrow \mathbf{R}$  is the stage cost function at time  $N$ , sometimes referred to as the terminal cost function. We will assume that the expectation above exists.

We also have state and control constraints

$$(x_t, u_t) \in \mathcal{C}_t \text{ (a.s.)}, \quad t = 0, 1, \dots, N-1, \quad x_N \in \mathcal{C}_N \text{ (a.s.)}, \quad (3)$$

where  $\mathcal{C}_0 \subseteq \mathbf{R}^n \times \mathbf{R}^m, \dots, \mathcal{C}_{N-1} \subseteq \mathbf{R}^n \times \mathbf{R}^m$  and  $\mathcal{C}_N \subseteq \mathbf{R}^n$  are nonempty constraint sets. The stage cost functions  $\ell_0, \dots, \ell_N$  and the constraint sets  $\mathcal{C}_0, \dots, \mathcal{C}_N$  need not be convex.

The stochastic control problem is to find the state feedback functions  $\psi_0, \dots, \psi_{N-1}$  that minimize the objective  $J$ , among those that satisfy the constraint (3). The problem data consists of  $A_0, \dots, A_{N-1}, B_0, \dots, B_{N-1}$ , the distribution of  $x_0$  and each  $w_t$ , the stage cost functions  $\ell_0, \dots, \ell_N$ , and the constraint sets  $\mathcal{C}_0, \dots, \mathcal{C}_N$ . We let  $J^*$  denote the optimal value of the stochastic control problem, *i.e.*, the minimum value of  $J$ .

For more on the formulation of the linear stochastic control problem, including technical details, see, *e.g.*, [Ber05, Ber07, BS96, Whi82, Åst70, BB91, Pow07, KV86, HLL97, Wit70].

## 2.1 Dynamic programming ‘solution’

In this section we give the standard dynamic programming solution of the stochastic control problem, for later use. We first define the extended value stage cost functions  $\bar{\ell}_t : \mathbf{R}^n \times \mathbf{R}^m \rightarrow \mathbf{R} \cup \{\infty\}$ ,  $t = 0, \dots, N - 1$ , as

$$\bar{\ell}_t(z, v) = \begin{cases} \ell_t(z, v) & (z, v) \in \mathcal{C}_t \\ \infty & \text{otherwise,} \end{cases} \quad t = 0, \dots, N - 1.$$

Similarly we define  $\bar{\ell}_N : \mathbf{R}^n \rightarrow \mathbf{R} \cup \{\infty\}$  as

$$\bar{\ell}_N(z) = \begin{cases} \ell_N(z) & z \in \mathcal{C}_N \\ \infty & \text{otherwise.} \end{cases}$$

Let  $V_t(z)$  denote the optimal value of the objective  $J$  starting from time  $t$ , at state  $x_t = z$ ,

$$V_t(z) = \min_{\psi_t, \dots, \psi_{N-1}} \mathbf{E} \left( \sum_{\tau=t}^{N-1} \bar{\ell}_\tau(x_\tau, u_\tau) + \bar{\ell}_N(x_N) \right),$$

subject to the dynamics (1). ( $V_t : \mathbf{R}^n \rightarrow \mathbf{R} \cup \{\infty\}$  is sometimes called the value function, or the optimal cost-to-go function, at time  $t$ .) We know that  $V_N(z) = \bar{\ell}_N(z)$  and  $J^* = \mathbf{E} V_0(x_0)$ , where the expectation is over  $x_0$ . The functions  $V_0, \dots, V_N$  satisfy the Bellman recursion,

$$V_t(z) = \min_v \{ \bar{\ell}_t(z, v) + \mathbf{E} V_{t+1}(A_t z + B_t v + w_t) \}, \quad t = N - 1, \dots, 0, \quad (4)$$

where the minimization is over the variable  $v$ , and the expectation is over  $w_t$ . We can write this in abstract form as

$$V_t = \mathcal{T}_t V_{t+1}, \quad t = N - 1, \dots, 0,$$

where  $\mathcal{T}_t$  is the Bellman operator at time  $t$ , defined as

$$(\mathcal{T}_t f)(z) = \min_v \{ \bar{\ell}_t(z, v) + \mathbf{E} f(A_t z + B_t v + w_t) \},$$

for any  $f : \mathbf{R}^n \rightarrow \mathbf{R} \cup \{\infty\}$ .

The optimal feedback functions are

$$\psi_t^*(z) = \underset{v}{\operatorname{argmin}} \{ \bar{\ell}_t(z, v) + \mathbf{E} V_{t+1}(A_t z + B_t v + w_t) \}, \quad t = 0, \dots, N-1. \quad (5)$$

The value functions and optimal feedback functions can be effectively computed in only a few special cases. The most famous example is when  $\mathcal{C}_0 = \dots = \mathcal{C}_{N-1} = \mathbf{R}^n \times \mathbf{R}^m$ ,  $\mathcal{C}_N = \mathbf{R}^n$  (there are no constraints on the input and state) and  $\ell_0, \dots, \ell_N$  are convex quadratic functions [Kal64]. In this case the optimal state feedback functions are affine, *i.e.*,  $u_t = K_t x_t + g_t$ ,  $t = 0, \dots, N-1$ , where  $K_t \in \mathbf{R}^{m \times n}$  and  $g_t \in \mathbf{R}^m$  are easily computed from the problem data. For more details, including proofs of these results, and other cases for which the optimal feedback function can be computed, see [Ber05, Ber07, BS96, Whi82].

## 2.2 Basic bound

Let  $\tilde{\ell}_t : \mathbf{R}^n \times \mathbf{R}^m \rightarrow \mathbf{R}$ ,  $t = 0, \dots, N-1$  be quadratic functions with the form

$$\tilde{\ell}_t(z, v) = \begin{bmatrix} z \\ v \end{bmatrix}^T \begin{bmatrix} Q_t & S_t \\ S_t^T & R_t \end{bmatrix} \begin{bmatrix} z \\ v \end{bmatrix} + 2q_t^T z + 2r_t^T v + s_t, \quad t = 0, \dots, N-1,$$

and let  $\tilde{\ell}_N : \mathbf{R}^n \rightarrow \mathbf{R}$  be quadratic with the form

$$\tilde{\ell}_N(z) = z^T Q_N z + 2q_N^T z + s_N.$$

We define operators  $\tilde{\mathcal{T}}_t$ ,  $t = 0, \dots, N-1$  as

$$(\tilde{\mathcal{T}}_t f)(z) = \min_v \{ \tilde{\ell}_t(z, v) + \mathbf{E} f(A_t z + B_t v + w_t) \}, \quad t = 0, \dots, N-1,$$

where  $f : \mathbf{R}^n \rightarrow \mathbf{R}$ . The operators  $\tilde{\mathcal{T}}_0, \dots, \tilde{\mathcal{T}}_{N-1}$  are Bellman operators with stage costs  $\tilde{\ell}_0, \dots, \tilde{\ell}_{N-1}$ , instead of  $\bar{\ell}_0, \dots, \bar{\ell}_{N-1}$ . Now suppose  $\tilde{\ell}_0, \dots, \tilde{\ell}_N$  satisfy

$$\tilde{\ell}_t \leq \bar{\ell}_t, \quad t = 0, \dots, N-1, \quad \tilde{\ell}_N \leq \bar{\ell}_N,$$

where the notation  $f \leq g$  for functions  $f$  and  $g$  means pointwise, *i.e.*,  $f(x) \leq g(x)$  for all  $x$ . This can be expressed as

$$\begin{aligned} \sup_{(z,v) \in \mathcal{C}_t} (\tilde{\ell}_t(z, v) - \bar{\ell}_t(z, v)) &\leq 0, \quad t = 0, \dots, N-1, \\ \sup_{z \in \mathcal{C}_N} (\tilde{\ell}_N(z) - \bar{\ell}_N(z)) &\leq 0. \end{aligned} \quad (6)$$

Then for any function  $f : \mathbf{R}^n \rightarrow \mathbf{R}$ , we have

$$\tilde{\ell}_t(z, v) + \mathbf{E} f(A_t z + B_t v + w_t) \leq \bar{\ell}_t(z, v) + \mathbf{E} f(A_t z + B_t v + w_t)$$

for all  $z \in \mathbf{R}^n$ ,  $v \in \mathbf{R}^m$ , which implies that

$$\tilde{\mathcal{T}}_t f \leq \mathcal{T}_t f, \quad t = 0, \dots, N-1. \quad (7)$$

Now let  $\tilde{V}_t : \mathbf{R}^n \rightarrow \mathbf{R}$ ,  $t = 0, \dots, N$  be quadratic functions with the form

$$\tilde{V}_t(z) = z^T P_t z + 2p_t^T z + c_t, \quad t = 0, \dots, N.$$

Suppose  $\tilde{V}_0, \dots, \tilde{V}_N$  satisfy the Bellman inequalities

$$\tilde{V}_t \leq \tilde{\mathcal{T}}_t \tilde{V}_{t+1}, \quad t = 0, \dots, N-1, \quad \tilde{V}_N = \tilde{\ell}_N. \quad (8)$$

Then we claim that

$$\tilde{V}_t \leq V_t, \quad t = 0, \dots, N, \quad (9)$$

which gives us the lower bound

$$\mathbf{E} \tilde{V}_0(x_0) \leq \mathbf{E} V_0(x_0) = J^*. \quad (10)$$

The lefthand side can be explicitly given as

$$\mathbf{E} \tilde{V}_0(x_0) = \mathbf{Tr}(P_0 X_0) + 2p_0^T \bar{x} + c_0.$$

(The lower bound depends only on the first and second moments of  $x_0$ , while the righthand side can depend on the particular distribution of  $x_0$ .)

We now establish our claim (9). We know  $\tilde{V}_N = \tilde{\ell}_N \leq \bar{\ell}_N = V_N$ , which implies that

$$\tilde{V}_{N-1} \leq \tilde{\mathcal{T}}_{N-1} \tilde{V}_N \leq \mathcal{T}_{N-1} \tilde{V}_N \leq \mathcal{T}_{N-1} V_N = V_{N-1}.$$

Here, the first and second inequalities follow from (8) and (7). The third inequality follows from monotonicity of the Bellman operator [RZ94, BS96, Ber07, Wit70], *i.e.*,  $f \leq g$  implies  $\mathcal{T}_{N-1} f \leq \mathcal{T}_{N-1} g$ , and the condition (6) Using the same argument we get

$$\tilde{V}_{N-2} \leq \tilde{\mathcal{T}}_{N-2} \tilde{V}_{N-1} \leq \mathcal{T}_{N-2} \tilde{V}_{N-1} \leq \mathcal{T}_{N-2} V_{N-1} = V_{N-2}.$$

Continuing this argument recursively we get  $\tilde{V}_t \leq V_t$ , for  $t = 0, \dots, N$ .

In other words, if we can find

$$\begin{aligned} Q_t, q_t, s_t, P_t, p_t, c_t, & \quad t = 0, \dots, N, \\ S_t, R_t, r_t, & \quad t = 0, \dots, N-1 \end{aligned}$$

for which (6) and (8) hold, then we have the lower bound on achievable performance

$$\mathbf{Tr}(P_0 X_0) + 2p_0^T \bar{x} + c_0 \leq J^*.$$

In the rest of this paper we will focus on cases where we can effectively compute this bound. In particular, we will see that we can optimize our bound over these variables by solving a convex optimization problem.



## 2.3 Bellman inequality as an LMI

We can express the Bellman inequalities (8) as

$$\begin{bmatrix} z \\ 1 \end{bmatrix}^T \begin{bmatrix} P_t & p_t \\ p_t^T & c_t \end{bmatrix} \begin{bmatrix} z \\ 1 \end{bmatrix} \leq \min_v \left\{ \begin{bmatrix} v \\ z \\ 1 \end{bmatrix}^T \begin{bmatrix} \tilde{R}_t & \tilde{S}_t^T & \tilde{r}_t \\ \tilde{S}_t & \tilde{Q}_t & \tilde{q}_t \\ \tilde{r}_t^T & \tilde{q}_t^T & \tilde{s}_t \end{bmatrix} \begin{bmatrix} v \\ z \\ 1 \end{bmatrix} \right\}, \quad t = 0, \dots, N-1,$$

for all  $z \in \mathbf{R}^n$ , where we define

$$\begin{aligned} \tilde{R}_t &= R_t + B_t^T P_{t+1} B_t, & \tilde{Q}_t &= Q_t + A_t^T P_{t+1} A_t, & \tilde{S}_t &= S_t + A_t^T P_{t+1} B_t, \\ \tilde{r}_t &= r_t + B_t^T P_{t+1} \bar{w}_t + B_t^T p_{t+1}, & \tilde{q}_t &= q_t + A_t^T P_{t+1} \bar{w}_t + A_t^T p_{t+1}, \\ \tilde{s}_t &= s_t + \mathbf{Tr}(P_{t+1}(W_t + \bar{w}_t \bar{w}_t^T)) + 2p_{t+1}^T \bar{w}_t + c_{t+1}. \end{aligned}$$

This is equivalent to the condition,

$$\begin{bmatrix} v \\ z \\ 1 \end{bmatrix}^T \begin{bmatrix} 0 & 0 & 0 \\ 0 & P_t & p_t \\ 0 & p_t^T & c_t \end{bmatrix} \begin{bmatrix} v \\ z \\ 1 \end{bmatrix} \leq \begin{bmatrix} v \\ z \\ 1 \end{bmatrix}^T \begin{bmatrix} \tilde{R}_t & \tilde{S}_t^T & \tilde{r}_t \\ \tilde{S}_t & \tilde{Q}_t & \tilde{q}_t \\ \tilde{r}_t^T & \tilde{q}_t^T & \tilde{s}_t \end{bmatrix} \begin{bmatrix} v \\ z \\ 1 \end{bmatrix}, \quad \text{for all } z \in \mathbf{R}^n, v \in \mathbf{R}^m,$$

for  $t = 0, \dots, N-1$ , which can be written as

$$\begin{bmatrix} \tilde{R}_t & \tilde{S}_t^T & \tilde{r}_t \\ \tilde{S}_t & \tilde{Q}_t - P_t & \tilde{q}_t - p_t \\ \tilde{r}_t^T & \tilde{q}_t^T - p_t^T & \tilde{s}_t - c_t \end{bmatrix} \succeq 0, \quad t = 0, \dots, N-1. \quad (11)$$

Each of the terms  $\tilde{R}_t, \tilde{S}_t, \tilde{r}_t, \tilde{Q}_t - P_t, \tilde{q}_t - p_t, \tilde{s}_t - c_t$  in the block matrix inequalities are linear functions of the variables  $Q_t, q_t, s_t, P_t, p_t, c_t, S_t, R_t$  and  $r_t$ . Thus, the inequalities (11) are linear matrix inequalities (LMIs) [BV04, BEFB94, VB96, RZ00, Bad82, ARE95, BLW91]. In particular, the set of matrices  $Q_t, q_t, s_t, P_t, p_t, c_t, S_t, R_t$ , and  $r_t$  that satisfy (11) is convex.

The terminal condition  $\tilde{V}_N = \ell_N$  can be written as

$$P_N = Q_N, \quad p_N = q_N, \quad c_N = s_N, \quad (12)$$

which is a set of linear equality constraints.

## 2.4 Optimizing the bound

We can optimize the lower bound (10) over the variables  $Q_t, q_t, s_t, P_t, p_t, c_t, t = 0, \dots, N$ , and  $S_t, R_t, r_t, t = 0, \dots, N-1$ , by solving the optimization problem

$$\begin{aligned} & \text{maximize} && \mathbf{E} \tilde{V}_0(x_0) \\ & \text{subject to} && (6), \quad (11), \quad (12). \end{aligned} \quad (13)$$

This is a convex optimization problem. The objective can be written as

$$\mathbf{E} \tilde{V}_0(x_0) = \mathbf{Tr}(P_0 X_0) + 2p_0^T \bar{x} + c_0,$$

which is a linear function of  $P_0, p_0$  and  $c_0$ . The LMIs (11) are convex constraint sets, (12) is convex and in addition, the condition (6) is convex. To see this, notice that the constraint

$$\tilde{\ell}_t(z, v) \leq \ell(z, v),$$

is linear in the variables  $Q_t, S_t, R_t, q_t, r_t, s_t$  for each  $z$  and  $v$ , and the supremum over a family of linear functions is convex. In the general case, the constraint (6) is a semi-infinite constraint, since it is really a family of constraints parameterized by the infinite sets  $\mathcal{C}_0, \dots, \mathcal{C}_N$  [BV04, NW99].

The idea behind our bound is to find functions  $\tilde{\ell}_t$  that are everywhere smaller than the stage cost functions  $\ell_t$ . Then, ignoring the constraints, the optimal value of new stochastic control problem with stage costs  $\tilde{\ell}_t$  is already a lower bound on  $J^*$ . If, in addition, the Bellman equations for the new stochastic control problem are relaxed to Bellman inequalities, then the functions  $\tilde{V}_t$  that satisfy these inequalities are certainly also lower bounds. Finally, we optimize the bound over the parameters by solving the optimization problem (13).

In some cases we can solve the problem (13) exactly. In other cases, we can replace the condition (6) with a conservative approximation, which still yields a lower bound on  $J^*$ . We give more specific examples of each of these cases below.

## 2.5 Finite input constraint set

Here is a case for which we can solve the optimization problem (13) exactly. We assume that the stage costs are quadratic with the form

$$\ell_t(z, v) = \begin{bmatrix} z \\ v \end{bmatrix}^T \begin{bmatrix} \bar{Q}_t & \bar{S}_t \\ \bar{S}_t^T & \bar{R}_t \end{bmatrix} \begin{bmatrix} z \\ v \end{bmatrix} + 2\bar{q}_t^T z + 2\bar{r}_t^T v + \bar{s}_t, \quad t = 0, \dots, N-1, \quad (14)$$

and terminal cost is quadratic with the form

$$\ell_N(z) = z^T \bar{Q}_N z + 2\bar{q}_N^T z + \bar{s}_N. \quad (15)$$

We also assume that there are no state constraints, and the input constraint sets are finite, *i.e.*,

$$\mathcal{C}_t = \mathbf{R}^n \times \mathcal{U}_t, \quad t = 0, \dots, N-1,$$

where  $\mathcal{U}_t = \{u_1^{(t)}, \dots, u_K^{(t)}\}$  and  $\mathcal{C}_N = \mathbf{R}^n$ . The condition (6) becomes

$$\begin{aligned} \begin{bmatrix} z \\ u_i^{(t)} \end{bmatrix}^T \begin{bmatrix} Q_t & S_t \\ S_t^T & R_t \end{bmatrix} \begin{bmatrix} z \\ u_i^{(t)} \end{bmatrix} + 2q_t^T z + 2r_t^T u_i^{(t)} + s_t \\ \leq \begin{bmatrix} z \\ u_i^{(t)} \end{bmatrix}^T \begin{bmatrix} \bar{Q}_t & \bar{S}_t \\ \bar{S}_t^T & \bar{R}_t \end{bmatrix} \begin{bmatrix} z \\ u_i^{(t)} \end{bmatrix} + 2\bar{q}_t^T z + 2\bar{r}_t^T u_i^{(t)} + \bar{s}_t, \end{aligned} \quad (16)$$

for all  $z \in \mathbf{R}^n$ ,  $i = 1, \dots, K$ ,  $t = 0, \dots, N-1$ , and

$$z^T Q_N z + 2q_N^T z + s_N \leq z^T \bar{Q}_N z + 2\bar{q}_N^T z + \bar{s}_N, \quad \text{for all } z \in \mathbf{R}^n. \quad (17)$$

We can write conditions (16) and (17) as LMIs,

$$\begin{bmatrix} \bar{Q}_t - Q_t & (\bar{S}_t - S_t)u_i^{(t)} + \bar{q}_t - q_t \\ ((\bar{S}_t - S_t)u_i^{(t)} + \bar{q}_t - q_t)^T & 2(\bar{r}_t - r_t)^T u_i^{(t)} + u_i^{(t)T}(\bar{R}_t - R_t)u_i^{(t)} + \bar{s}_t - s_t \end{bmatrix} \succeq 0 \quad (18)$$

for  $t = 0, \dots, N-1$ , and

$$\begin{bmatrix} \bar{Q}_N - Q_N & \bar{q}_N - q_N \\ (\bar{q}_N - q_N)^T & \bar{s}_N - s_N \end{bmatrix} \succeq 0. \quad (19)$$

Thus in this case, problem (13) can be expressed as the semidefinite program (SDP),

$$\begin{aligned} & \text{maximize} && \text{Tr}(P_0 X_0) + 2p_0^T \bar{x} + c_0 \\ & \text{subject to} && (18), \quad (19), \quad (11), \quad (12), \end{aligned} \quad (20)$$

with variables  $Q_t, q_t, s_t, P_t, p_t, c_t$ ,  $t = 0, \dots, N$ , and  $S_t, R_t, r_t$ ,  $t = 0, \dots, N-1$ . This can be effectively solved using interior-point methods (see, *e.g.*, [BV04, NW99, PW00, VB96, GBY06, Wri97b, NN94]).

## 2.6 S-procedure relaxation

We suppose again that the stage costs are quadratic, with the form in (14) and (15). Let  $f_t^{(i)} : \mathbf{R}^n \times \mathbf{R}^m \rightarrow \mathbf{R}$ ,  $i = 1, \dots, M_t$ ,  $t = 0, \dots, N-1$ , be quadratic (not necessarily convex) functions, with the form

$$f_t^{(i)}(z, v) = \begin{bmatrix} z \\ v \end{bmatrix}^T \begin{bmatrix} F_t^{(i)} & G_t^{(i)} \\ G_t^{(i)T} & H_t^{(i)} \end{bmatrix} \begin{bmatrix} z \\ v \end{bmatrix} + 2g_t^{(i)T} z + 2h_t^{(i)T} v + d_t^{(i)},$$

and let  $f_N^{(i)} : \mathbf{R}^N \rightarrow \mathbf{R}$ ,  $i = 1, \dots, M_N$  have the form

$$f_N^{(i)}(z) = z^T F_N^{(i)} z + 2g_N^{(i)T} z + d_N^{(i)}.$$

Now suppose we can find matrices  $F_t^{(i)}, G_t^{(i)}, H_t^{(i)}, g_t^{(i)}, h_t^{(i)}$  and  $d_t^{(i)}$  so that

$$\mathcal{C}_t \subseteq \tilde{\mathcal{C}}_t = \{(z, v) \mid f_t^{(i)}(z, v) \leq 0, \quad i = 1, \dots, M_t\}, \quad t = 0, \dots, N-1,$$

and

$$\mathcal{C}_N \subseteq \tilde{\mathcal{C}}_N = \{z \mid f_N^{(i)}(z) \leq 0, \quad i = 1, \dots, M_N\}.$$

A sufficient condition for (6) is

$$\sup_{(z,v) \in \tilde{\mathcal{C}}_t} (\tilde{\ell}_t(z, v) - \ell_t(z, v)) \leq 0, \quad t = 0, \dots, N-1, \quad \sup_{z \in \tilde{\mathcal{C}}_N} (\tilde{\ell}_N(z) - \ell_N(z)) \leq 0,$$

which is equivalent to

$$f_t^{(i)}(z, v) \leq 0, \quad i = 1, \dots, M_t \quad \implies \quad \tilde{\ell}_t(z, v) \leq \ell_t(z, v), \quad (21)$$

for  $t = 0, \dots, N - 1$ , and

$$f_N^{(i)}(z) \leq 0, \quad i = 1, \dots, M_N \quad \implies \quad \tilde{\ell}_N(z) \leq \ell_N(z). \quad (22)$$

A sufficient condition for (21) and (22) is (by the so-called  $\mathcal{S}$ -procedure [BEFB94, BV04]) the existence of nonnegative  $\lambda_t^{(1)}, \dots, \lambda_t^{(M_t)}$ ,  $t = 0, \dots, N$  such that

$$\tilde{\ell}_t(z, v) - \ell_t(z, v) - \sum_{i=1}^{M_t} \lambda_t^{(i)} f_t^{(i)}(z, v) \leq 0, \quad \text{for all } z \in \mathbf{R}^n, v \in \mathbf{R}^m,$$

and

$$\tilde{\ell}_N(z) - \ell_N(z) - \sum_{i=1}^{M_N} \lambda_N^{(i)} f_N^{(i)}(z) \leq 0, \quad \text{for all } z \in \mathbf{R}^n.$$

These can be written as LMIs

$$\begin{bmatrix} \bar{Q}_t - Q_t & \bar{S}_t - S_t & \bar{q}_t - q_t \\ (\bar{S}_t - S_t)^T & \bar{R}_t - R_t & \bar{r}_t - r_t \\ (\bar{q}_t - q_t)^T & (\bar{r}_t - r_t)^T & \bar{s}_t - s_t \end{bmatrix} + \sum_{i=1}^{M_t} \lambda_t^{(i)} \begin{bmatrix} F_t^{(i)} & G_t^{(i)} & g_t^{(i)} \\ G_t^{(i)T} & H_t^{(i)} & h_t^{(i)} \\ g_t^{(i)T} & h_t^{(i)T} & d_t^{(i)} \end{bmatrix} \succeq 0, \quad (23)$$

for  $t = 0, \dots, N - 1$  and

$$\begin{bmatrix} \bar{Q}_N - Q_N & \bar{q}_N - q_N \\ (\bar{q}_N - q_N)^T & \bar{s}_N - s_N \end{bmatrix} + \sum_{i=1}^{M_N} \lambda_N^{(i)} \begin{bmatrix} F_N^{(i)} & g_N^{(i)} \\ g_N^{(i)T} & d_N^{(i)} \end{bmatrix} \succeq 0. \quad (24)$$

Thus, to get a lower bound we relax the condition (6) to conditions (23) and (24). This gives us the SDP

$$\begin{aligned} & \text{maximize} \quad \mathbf{Tr}(P_0 X_0) + 2p_0^T \bar{x} + c_0 \\ & \text{subject to} \quad (23), \quad (24), \quad (11), \quad (12) \\ & \quad \quad \quad \lambda_t^{(i)} \geq 0, \quad t = 0, \dots, N, \quad i = 1, \dots, M. \end{aligned} \quad (25)$$

with variables  $Q_t, q_t, s_t, P_t, p_t, c_t$ ,  $t = 0, \dots, N$ ,  $S_t, R_t, r_t$ ,  $t = 0, \dots, N - 1$ , and  $\lambda_t^{(i)}$ ,  $i = 1, \dots, M$ ,  $t = 0, \dots, N$ . Again, this can be solved very efficiently [BV04, NW99, PW00, VB96, GBY06, Wri97b, NN94].

## 2.7 Control-Lyapunov policy

There are many methods for implementing suboptimal controllers. In this paper we will consider one of these methods, called the control-Lyapunov feedback policy (CLF).

In control-Lyapunov feedback, we modify the optimal feedback function (5) by replacing the optimal value function  $V_{t+1}$ , with an approximate value function  $V_{t+1}^{\text{clf}} : \mathbf{R}^n \rightarrow \mathbf{R}$ , which we call a *control-Lyapunov function* [Son83, FP96, CL88, SSC03]. The state feedback function at time  $t$  is given by

$$\psi_t^{\text{clf}}(z) = \underset{(z,v) \in \mathcal{C}_t}{\operatorname{argmin}} (\ell_t(z, v) + \mathbf{E} V_{t+1}^{\text{clf}}(A_t z + B_t v + w_t)). \quad (26)$$

The performance of this feedback policy clearly relies on good choices for  $V_1^{\text{clf}}, \dots, V_N^{\text{clf}}$ . Ideally, a control-Lyapunov function should be a good approximation for the optimal value function, but it should also allow the state feedback function to be effectively evaluated. For example, if the stage costs are convex and quadratic, then common choices for  $V_1^{\text{clf}}, \dots, V_N^{\text{clf}}$  would be the quadratic value functions for the associated linear stochastic control problem with no constraints.

In this paper, when we optimize our bound on  $J^*$  (either exactly, or by solving a conservative approximation of (13)), we obtain the lower bound functions  $V_0^{\text{lb}}, \dots, V_N^{\text{lb}}$ , where

$$V_t^{\text{lb}}(z) = z^T P_t^{\text{lb}} z + 2p_t^{\text{lb}T} z + c_t^{\text{lb}}, \quad t = 0, \dots, N.$$

Here,  $P_t^{\text{lb}}, p_t^{\text{lb}}, c_t^{\text{lb}}$  denote the  $P_t, p_t,$  and  $c_t$  matrices we obtain by optimizing our lower bound (13). Very roughly speaking, we can interpret  $V_0^{\text{lb}}, \dots, V_N^{\text{lb}}$ , as value functions for an unconstrained problem that approximates our original problem. Thus, we expect that  $V_1^{\text{lb}}, \dots, V_N^{\text{lb}}$  would be good choices for a control-Lyapunov policy. In this case, the state feedback function can be written as

$$\psi_t^{\text{clf}}(z) = \underset{(z,v) \in \mathcal{C}_t}{\operatorname{argmin}} \left( \ell_t(z, v) + \begin{bmatrix} v \\ z \\ 1 \end{bmatrix}^T \begin{bmatrix} \hat{R}_t & \hat{S}_t^T & \hat{r}_t \\ \hat{S}_t & \hat{Q}_t & \hat{q}_t \\ \hat{r}_t^T & \hat{q}_t^T & 0 \end{bmatrix} \begin{bmatrix} v \\ z \\ 1 \end{bmatrix} \right),$$

where

$$\begin{aligned} \hat{R}_t &= B_t^T P_{t+1}^{\text{lb}} B_t, & \hat{Q}_t &= A_t^T P_{t+1}^{\text{lb}} A_t, & \hat{S}_t &= A_t^T P_{t+1}^{\text{lb}} B_t, \\ \hat{r}_t &= B_t^T P_{t+1}^{\text{lb}} \bar{w}_t + B_t^T p_{t+1}^{\text{lb}}, & \hat{q}_t &= A_t^T P_{t+1}^{\text{lb}} \bar{w}_t + A_t^T p_{t+1}^{\text{lb}}. \end{aligned}$$

In particular, when the stage cost  $\ell_t(z, v)$  is convex and quadratic and the constraint set  $\mathcal{C}_t$  is polyhedral, we can evaluate this feedback function by solving a convex quadratic program (QP) with  $m$  variables. This can be done *very* efficiently: For instance, for a system with say 10 inputs, the QP can be solved in tens of *microseconds*, allowing control to be carried out at tens of kilohertz [WB09a, WB08, WB07, MB09]. Alternatively, we can also solve the QP explicitly offline, as a multiparametric quadratic program (parameterized by the state  $z$ ). Then, online evaluation of the control policy reduces to searching through a lookup table of pre-computed affine controllers. When the state and input dimensions are small, this method also yields extremely fast computation times [BMDP02, TJB01, ZJM08, CZJM07, TJ02, BF04, KGBM03, FBD08, JGR06].

In §4 we will show that for many examples, the gap between the objective achieved by the control-Lyapunov policy and our lower bound is small, which shows that these controllers are nearly optimal.

### 3 Infinite horizon

We now derive a lower bound for the infinite horizon, average cost-per-stage problem. Here, we consider a discrete time-invariant linear system with dynamics,

$$x_{t+1} = Ax_t + Bu_t + w_t, \quad t = 0, 1, \dots, \tag{27}$$

where  $x_t \in \mathbf{R}^n$  is the state,  $u_t \in \mathbf{R}^m$  is the control input,  $w_t \in \mathbf{R}^n$  is the disturbance at time  $t$ , and  $A \in \mathbf{R}^{n \times n}$  and  $B \in \mathbf{R}^{n \times m}$  are the dynamics and input matrices. We assume that  $w_t$  for different values of  $t$  are independent identically distributed (IID) with mean  $\bar{w} = \mathbf{E} w_t$ , and covariance  $W = \mathbf{E}(w_t - \bar{w})(w_t - \bar{w})^T$ . We will also assume that  $x_0$  is random, and independent of all  $w_t$ , but we'll see that the distribution of  $x_0$  will not matter in the problem we consider.

As with the finite horizon case, we consider causal state feedback control policies, where the current input  $u_t$  is determined from the current and previous states  $x_0, \dots, x_t$ . For the problem we'll consider, it is also possible to show that there is an optimal policy that is time invariant and depends only on the current state, *i.e.*,

$$u_t = \psi(x_t), \quad t = 0, 1, \dots, \quad (28)$$

where  $\psi : \mathbf{R}^n \rightarrow \mathbf{R}^m$  is called the state feedback function. For a fixed state feedback function (28) and system dynamics (27), the state and input trajectories are stochastic processes.

We now introduce the objective function, which we assume has the form

$$J = \limsup_{N \rightarrow \infty} \frac{1}{N} \mathbf{E} \sum_{t=0}^{N-1} \ell(x_t, u_t) \quad (29)$$

where  $\ell : \mathbf{R}^n \times \mathbf{R}^m \rightarrow \mathbf{R}$  is the stage cost function. (Here, we'll assume that the expectations exist.) The objective  $J$  is the average stage cost. We also impose constraints on the state and input

$$(x_t, u_t) \in \mathcal{C} \text{ (a.s.)}, \quad t = 0, 1, \dots, \quad (30)$$

where  $\mathcal{C} \subseteq \mathbf{R}^m$  is a nonempty constraint set. The stage cost  $\ell$  and the constraint set  $\mathcal{C}$  need not be convex.

The time-invariant infinite horizon stochastic control problem is to choose the state feedback function  $\psi$  that minimizes the objective  $J$  and satisfies the constraint (30). We let  $J^*$  denote the optimal value of  $J$  and we let  $\psi^*$  denote the optimal state feedback function. The problem data are  $A$ ,  $B$ , the distribution of  $w_t$ , the stage cost function  $\ell$ , and the constraint set  $\mathcal{C}$ .

For more on the formulation of the stochastic control problem, including technical details (*e.g.*, finiteness of  $J^*$ , existence and uniqueness of an optimal state feedback function), see [Ber05, Ber07, BS96, Whi82, Åst70, BB91, Pow07, KV86, HLL97, Wit70].

### 3.1 Dynamic programming ‘solution’

As with the finite horizon case, we first give the dynamic programming solution of the stochastic control problem. We'll use these results (and the notation) later. First we define the extended value stage cost function  $\bar{\ell} : \mathbf{R}^n \times \mathbf{R}^m \rightarrow \mathbf{R} \cup \{\infty\}$ , as

$$\bar{\ell}(z, v) = \begin{cases} \ell(z, v) & (z, v) \in \mathcal{C} \\ \infty & \text{otherwise.} \end{cases}$$

The Bellman equation for the average cost-per-stage problem can be written as

$$\alpha + V = \mathcal{T}V, \quad (31)$$

where  $V : \mathbf{R}^n \rightarrow \mathbf{R} \cup \{\infty\}$ ,  $\alpha \in \mathbf{R}$ . Here,  $\mathcal{T}$  is the steady-state Bellman operator, defined as

$$(\mathcal{T}f)(z) = \min_v \{\bar{\ell}(z, v) + \mathbf{E} f(Az + Bv + w_t)\},$$

for any  $f : \mathbf{R}^n \rightarrow \mathbf{R} \cup \{\infty\}$ , and  $\alpha + V$  is a function defined as

$$(\alpha + V)(x) = \alpha + V(x),$$

for all  $x$ . If we can find a function  $V$  and a constant  $\alpha$  that satisfies (31), then  $J^* = \alpha$ , and the optimal feedback functions are

$$\psi^*(z) = \operatorname{argmin}_v \{\bar{\ell}(z, v) + \mathbf{E} V(Az + Bv + w_t)\}. \quad (32)$$

Notice that if  $\alpha, V$  satisfy (31), then  $\alpha, V + \beta$  also satisfy (31), for any  $\beta \in \mathbf{R}$ . Thus, we can assume, without loss of generality, that  $V(0) = 0$ .

Here, several pathologies can occur. The stochastic control problem can be infeasible—there exists no causal state feedback policy that satisfies the constraints, and attains a finite average cost per stage  $J$ . The stochastic control can also be unbounded below, which means that we can find policies for which  $J = -\infty$ . Finally, the Bellman equation may not have any solutions, *i.e.*, there exist no  $\alpha, V$  that satisfy (31). In this paper, we will consider only the cases where the stochastic control problem is feasible, the optimal average cost per stage is finite, and a solution exists to the Bellman equation (31). For the technical details, including the conditions under which a solution to the Bellman equation exists, see, *e.g.*, [Ber05, Ber07, BS96, Whi82].

The value iteration method for the average cost problem can be written as

$$\hat{V}^{(k)} = \mathcal{T}V^{(k)}, \quad V^{(k+1)} = \hat{V}^{(k)} - \hat{V}^{(k)}(0), \quad \alpha^{(k)} = \hat{V}^{(k)}(0), \quad (33)$$

where  $V^{(k)} : \mathbf{R}^n \rightarrow \mathbf{R} \cup \{\infty\}$ ,  $\hat{V}^{(k)} : \mathbf{R}^n \rightarrow \mathbf{R} \cup \{\infty\}$ ,  $k = 1, 2, \dots$ , and  $V^{(0)} : \mathbf{R}^n \rightarrow \mathbf{R}$  is any real-valued function. As  $k \rightarrow \infty$ ,

$$V^{(k)} \rightarrow V \text{ (pointwise)}, \quad \alpha^{(k)} \rightarrow \alpha,$$

where  $V : \mathbf{R}^n \rightarrow \mathbf{R} \cup \{\infty\}$  and  $\alpha \in \mathbf{R}$  satisfy the Bellman equation (31). As with the finite horizon case, the function  $V$  and the constant  $\alpha$  can be computed in only a few special cases. One example is where  $\mathcal{C} = \mathbf{R}^n \times \mathbf{R}^m$  and the stage cost  $\ell$  is a convex quadratic function. In this case, the optimal state feedback function is affine, *i.e.*,  $u_t = Kx_t + g$  (and  $K, g$  are easily computed from the problem data).

### 3.2 Basic bound

Our development of the performance bound for the infinite horizon problem will be very similar compared with the finite horizon case in §2.2. Here we will make use of the value iteration described in §3.1 to show one of our inequalities.

Let  $\tilde{\ell} : \mathbf{R}^n \times \mathbf{R}^m \rightarrow \mathbf{R}$  be quadratic with the form

$$\tilde{\ell}(z, v) = \begin{bmatrix} z \\ v \end{bmatrix}^T \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \begin{bmatrix} z \\ v \end{bmatrix} + 2q^T z + 2r^T v + s.$$

We define an operator  $\tilde{\mathcal{T}}$  as

$$(\tilde{\mathcal{T}}f)(z) = \min_v \{ \tilde{\ell}(z, v) + \mathbf{E} f(Az + Bv + w_t) \},$$

where  $f : \mathbf{R}^n \rightarrow \mathbf{R}$ . The operator  $\tilde{\mathcal{T}}$  is a Bellman operator with stage cost  $\tilde{\ell}$ , instead of  $\bar{\ell}$ . Now suppose that  $\tilde{\ell}$  satisfies  $\tilde{\ell} \leq \bar{\ell}$ . As before, the notation  $f \leq g$  for functions  $f$  and  $g$  means pointwise inequality, *i.e.*,  $f(x) \leq g(x)$  for all  $x$ . This condition can be expressed as

$$\sup_{(z,v) \in \mathcal{C}} (\tilde{\ell}(z, v) - \bar{\ell}(z, v)) \leq 0. \quad (34)$$

Then, for any function  $f \in \mathbf{R}^n \rightarrow \mathbf{R}$  we have

$$\tilde{\ell}(z, v) + \mathbf{E} f(Az + Bv + w_t) \leq \bar{\ell}(z, v) + \mathbf{E} f(Az + Bv + w_t),$$

for all  $z \in \mathbf{R}^n$ ,  $v \in \mathbf{R}^m$ , so we get

$$\tilde{\mathcal{T}}f \leq \mathcal{T}f. \quad (35)$$

Now let  $\tilde{V} : \mathbf{R}^n \rightarrow \mathbf{R}$ , be a quadratic function with the form

$$\tilde{V}(z) = z^T P z + 2p^T z,$$

and suppose  $\tilde{\alpha} \in \mathbf{R}$  and  $\tilde{V}$  satisfy the Bellman inequality

$$\tilde{\alpha} + \tilde{V} \leq \tilde{\mathcal{T}}\tilde{V}. \quad (36)$$

We define  $\alpha$  and  $V : \mathbf{R}^n \rightarrow \mathbf{R}$  to be

$$\alpha = \lim_{k \rightarrow \infty} \alpha^{(k)}, \quad V(z) = \lim_{k \rightarrow \infty} V^{(k)}(z), \quad \text{for all } z \in \mathbf{R}^n,$$

where  $\alpha^{(k)}$  and  $V^{(k)}$  satisfy the value iteration (33), with  $V^{(0)} = \tilde{V}$ . In particular,  $\alpha$  and  $V$  satisfy the Bellman equation (31), *i.e.*,  $\alpha + V = \mathcal{T}V$ , and  $\alpha = J^*$ . Now we claim that

$$\tilde{\alpha} \leq \alpha = J^*, \quad (37)$$

so  $\tilde{\alpha}$  is our lower bound. To prove this, we can write  $V^{(k)}$  (the  $k$ th iterate of the value iteration), as

$$V^{(k)} = \mathcal{T}^k V^{(0)} - \sum_{i=0}^{k-1} \alpha^{(i)} = \mathcal{T}^k \tilde{V} - \sum_{i=0}^{k-1} \alpha^{(i)}.$$



This implies

$$\mathcal{T}V^{(k)} = \mathcal{T}^k \mathcal{T}\tilde{V} - \sum_{i=0}^{k-1} \alpha^{(i)} \geq \tilde{\alpha} + \mathcal{T}^k \tilde{V} - \sum_{i=0}^{k-1} \alpha^{(i)} = \tilde{\alpha} + V^{(k)}.$$

Here, the inequality follows from  $\tilde{\alpha} + \tilde{V} \leq \tilde{\mathcal{T}}\tilde{V} \leq \mathcal{T}\tilde{V}$  (*i.e.*, inequalities (36) and (35)). Taking the pointwise limit as  $k \rightarrow \infty$  we get

$$\alpha + V(z) = (\mathcal{T}V)(z) = \lim_{k \rightarrow \infty} (\mathcal{T}V^{(k)})(z) \geq \tilde{\alpha} + \lim_{k \rightarrow \infty} V^{(k)}(z) = \tilde{\alpha} + V(z), \text{ for all } z \in \mathbf{R}^n.$$

Thus we conclude  $\tilde{\alpha} \leq \alpha$ .

This means that if we can find

$$Q, S, R, q, s, r, P, p, \tilde{\alpha}$$

for which (34) and (36) hold, then  $\tilde{\alpha}$  is a lower bound on  $J^*$ . As with the finite horizon case, we will focus on the cases where we can effectively compute (and optimize) this bound.

### 3.3 Bellman inequality as an LMI

We can express the Bellman inequality (36) as

$$\begin{bmatrix} z \\ 1 \end{bmatrix}^T \begin{bmatrix} P & p \\ p^T & \tilde{\alpha} \end{bmatrix} \begin{bmatrix} z \\ 1 \end{bmatrix} \leq \min_v \left\{ \begin{bmatrix} v \\ z \\ 1 \end{bmatrix}^T \begin{bmatrix} \tilde{R} & \tilde{S}^T & \tilde{r} \\ \tilde{S} & \tilde{Q} & \tilde{q} \\ \tilde{r}^T & \tilde{q}^T & \tilde{s} \end{bmatrix} \begin{bmatrix} v \\ z \\ 1 \end{bmatrix} \right\},$$

for all  $z \in \mathbf{R}^n$ , where we define

$$\begin{aligned} \tilde{R} &= R + B^T P B, & \tilde{Q} &= Q + A^T P A, & \tilde{S} &= S + A^T P B, \\ \tilde{r} &= r + B^T P \bar{w} + B^T p, & \tilde{q} &= q + A^T P \bar{w} + A^T p, \\ \tilde{s} &= s + \mathbf{Tr}(P(W + \bar{w}\bar{w}^T)) + 2p^T \bar{w}. \end{aligned}$$

This is equivalent to the condition,

$$\begin{bmatrix} v \\ z \\ 1 \end{bmatrix}^T \begin{bmatrix} 0 & 0 & 0 \\ 0 & P & p \\ 0 & p^T & \tilde{\alpha} \end{bmatrix} \begin{bmatrix} v \\ z \\ 1 \end{bmatrix} \leq \begin{bmatrix} v \\ z \\ 1 \end{bmatrix}^T \begin{bmatrix} \tilde{R} & \tilde{S}^T & \tilde{r} \\ \tilde{S} & \tilde{Q} & \tilde{q} \\ \tilde{r}^T & \tilde{q}^T & \tilde{s} \end{bmatrix} \begin{bmatrix} v \\ z \\ 1 \end{bmatrix}, \text{ for all } z \in \mathbf{R}^n, v \in \mathbf{R}^m,$$

which can be written as

$$\begin{bmatrix} \tilde{R} & \tilde{S}^T & \tilde{r} \\ \tilde{S} & \tilde{Q} - P & \tilde{q} - p \\ \tilde{r}^T & \tilde{q}^T - p^T & \tilde{s} - \tilde{\alpha} \end{bmatrix} \succeq 0. \quad (38)$$

Each of the terms  $\tilde{R}$ ,  $\tilde{S}$ ,  $\tilde{r}$ ,  $\tilde{Q} - P$ ,  $\tilde{q} - p$ ,  $\tilde{s} - \tilde{\alpha}$  in the block matrix inequality is a linear function of the variables  $Q, S, R, q, s, r, P, p, \tilde{\alpha}$ . Thus, the inequality (38) is an LMI [BV04, BEFB94, VB96, RZ00, Bad82, ARE95, BLW91].

### 3.4 Optimizing the bound

As with the finite horizon case, we can optimize our lower bound  $\tilde{\alpha}$ , over the variables  $Q, S, R, q, s, r, P, p, \tilde{\alpha}$ , by solving the optimization problem

$$\begin{aligned} & \text{maximize} && \tilde{\alpha} \\ & \text{subject to} && (34), \quad (38). \end{aligned} \tag{39}$$

The condition (34) is convex, since the constraint

$$\tilde{\ell}(z, v) \leq \ell(z, v)$$

is linear in the variables  $Q, S, R, q, r, s$ , for each  $z$  and  $v$ , and the supremum over a family of linear functions is convex. In addition, the LMI (38) defines a convex constraint set, thus the optimization problem (39) is a convex optimization problem [BV04, NW99].

In the general case, the condition (34) is a semi-infinite constraint, since it is a family of constraints parametrized by the infinite set  $\mathcal{C}$ . In the next few sections, we discuss cases where we can handle the semi-infinite constraint exactly, and cases where we can replace (34) with a relaxation, which still yields a lower bound on  $J^*$ .

### 3.5 Finite input constraint set

We now describe a case for which we can solve the optimization problem (39) exactly. First we assume that the stage cost is quadratic with the form

$$\ell(z, v) = \begin{bmatrix} z \\ v \end{bmatrix}^T \begin{bmatrix} \bar{Q} & \bar{S} \\ \bar{S}^T & \bar{R} \end{bmatrix} \begin{bmatrix} z \\ v \end{bmatrix} + 2\bar{q}^T z + 2\bar{r}^T v + \bar{s}. \tag{40}$$

We also assume that there are no state constraints, and the input constraint set is finite, *i.e.*,

$$\mathcal{C} = \mathbf{R}^n \times \mathcal{U},$$

where  $\mathcal{U} = \{u^{(1)}, \dots, u^{(K)}\} \subseteq \mathbf{R}^m$ . The condition (34) becomes

$$\begin{aligned} & \begin{bmatrix} z \\ u^{(i)} \end{bmatrix}^T \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \begin{bmatrix} z \\ u^{(i)} \end{bmatrix} + 2q^T z + 2r^T u^{(i)} + s \\ & \leq \begin{bmatrix} z \\ u^{(i)} \end{bmatrix}^T \begin{bmatrix} \bar{Q} & \bar{S} \\ \bar{S}^T & \bar{R} \end{bmatrix} \begin{bmatrix} z \\ u^{(i)} \end{bmatrix} + 2\bar{q}^T z + 2\bar{r}^T u^{(i)} + \bar{s}, \end{aligned} \tag{41}$$

for all  $z \in \mathbf{R}^n, i = 1, \dots, K$ . We can write the conditions (41) as LMIs,

$$\begin{bmatrix} \bar{Q} - Q & (\bar{S} - S)u^{(i)} + \bar{q} - q \\ ((\bar{S} - S)u^{(i)} + \bar{q} - q)^T & 2(\bar{r} - r)^T u^{(i)} + u^{(i)T}(\bar{R} - R)u^{(i)} + \bar{s} - s \end{bmatrix} \succeq 0, \tag{42}$$

for  $i = 1, \dots, K$ . Thus, the problem (39) becomes the SDP,

$$\begin{aligned} & \text{maximize} && \tilde{\alpha} \\ & \text{subject to} && (42), \quad (38) \end{aligned} \tag{43}$$

with variables  $Q, S, R, q, r, s, P, p$ , and  $\tilde{\alpha}$  (which we can solve using interior-point methods [BV04, NW99, PW00, VB96, GBY06, Wri97b, NN94]).

### 3.6 S-procedure relaxation

Now we consider the case where we can relax the condition (34), and still obtain a lower bound. We suppose again that the stage cost is quadratic, with the form in (40). Let  $f^{(i)} : \mathbf{R}^n \times \mathbf{R}^m \rightarrow \mathbf{R}$ ,  $i = 1, \dots, M$ , be quadratic (not necessarily convex) functions, with the form

$$f^{(i)}(z, v) = \begin{bmatrix} z \\ v \end{bmatrix}^T \begin{bmatrix} F^{(i)} & G^{(i)} \\ G^{(i)T} & H^{(i)} \end{bmatrix} \begin{bmatrix} z \\ v \end{bmatrix} + 2g^{(i)T}z + 2h^{(i)T}v + d^{(i)}.$$

Now suppose we can find matrices  $F^{(i)}, G^{(i)}, H^{(i)}, g^{(i)}, h^{(i)}$  and  $d^{(i)}$  so that

$$\mathcal{C} \subseteq \tilde{\mathcal{C}} = \{(z, v) \mid f^{(i)}(z, v) \leq 0, i = 1, \dots, M\}.$$

A sufficient condition for (34) is

$$\sup_{(z,v) \in \tilde{\mathcal{C}}} (\tilde{\ell}(z, v) - \ell(z, v)) \leq 0,$$

which is equivalent to

$$f^{(i)}(z, v) \leq 0, i = 1, \dots, M \implies \tilde{\ell}(z, v) \leq \ell(z, v). \quad (44)$$

A sufficient condition for (44) is (by the  $\mathcal{S}$ -procedure) the existence of nonnegative  $\lambda_1, \dots, \lambda_M$  such that

$$\tilde{\ell}(z, v) - \ell(z, v) - \sum_{i=1}^M \lambda_i f^{(i)}(z, v) \leq 0, \text{ for all } z \in \mathbf{R}^n, v \in \mathbf{R}^m.$$

This can be written as the LMI

$$\begin{bmatrix} \bar{Q} - Q & \bar{S} - S & \bar{q} - q \\ (\bar{S} - S)^T & \bar{R} - R & \bar{r} - r \\ (\bar{q} - q)^T & (\bar{r} - r)^T & \bar{s} - s \end{bmatrix} + \sum_{i=1}^M \lambda_i \begin{bmatrix} F^{(i)} & G^{(i)} & g^{(i)} \\ G^{(i)T} & H^{(i)} & h^{(i)} \\ g^{(i)T} & h^{(i)T} & d^{(i)} \end{bmatrix} \succeq 0. \quad (45)$$

Thus, to get a lower bound we relax the condition (34) to the above LMI (45). Then, to optimize the bound we solve the SDP

$$\begin{aligned} & \text{maximize} && \tilde{\alpha} \\ & \text{subject to} && (45), \quad (38) \\ & && \lambda_i \geq 0, \quad i = 1, \dots, M, \end{aligned} \quad (46)$$

with variables  $Q, S, R, q, r, s, P, p, \tilde{\alpha}$ , and  $\lambda_1, \dots, \lambda_M$ . Again, this can be effectively solved (see [BV04, NW99, PW00, VB96, GBY06, Wri97b, NN94]).

### 3.7 Control-Lyapunov policy

As with the finite horizon case, to get the control-Lyapunov feedback policy for the infinite horizon problem we modify the optimal feedback function (32) by replacing the optimal  $V$ , with an approximation  $V^{\text{clf}} : \mathbf{R}^n \rightarrow \mathbf{R}$ . We call  $V^{\text{clf}}$  the steady-state control-Lyapunov function [Son83, FP96, CL88, SSC03]. In this case, the state feedback function is given by

$$\psi^{\text{clf}}(z) = \underset{(z,v) \in \mathcal{C}}{\operatorname{argmin}} (\ell(z, v) + \mathbf{E} V^{\text{clf}}(Az + Bv + w_t)). \quad (47)$$

As before, one natural choice for  $V^{\text{clf}}$  that arises from our lower bound is

$$V^{\text{lb}}(z) = z^T P^{\text{lb}} z + 2p^{\text{lb}T} z,$$

where  $P^{\text{lb}}, p^{\text{lb}}$  denote the  $P, p$  matrices we obtain by optimizing our bound (39). Again, for this choice of  $V^{\text{clf}}$  we can write the feedback function as

$$\psi^{\text{clf}}(z) = \underset{(z,v) \in \mathcal{C}}{\operatorname{argmin}} \left( \ell(z, v) + \begin{bmatrix} v \\ z \\ 1 \end{bmatrix}^T \begin{bmatrix} \hat{R} & \hat{S}^T & \hat{r} \\ \hat{S} & \hat{Q} & \hat{q} \\ \hat{r}^T & \hat{q}^T & 0 \end{bmatrix} \begin{bmatrix} v \\ z \\ 1 \end{bmatrix} \right),$$

where

$$\begin{aligned} \hat{R} &= B^T P^{\text{lb}} B, & \hat{Q} &= A^T P^{\text{lb}} A, & \hat{S} &= A^T P^{\text{lb}} B, \\ \hat{r} &= B^T P^{\text{lb}} \bar{w} + B^T p^{\text{lb}}, & \hat{q} &= A^T P^{\text{lb}} \bar{w} + A^T p^{\text{lb}}. \end{aligned}$$

When the stage cost  $\ell(z, v)$  is convex and quadratic and the constraint set  $\mathcal{C}$  is polyhedral, we can evaluate this feedback function by solving a QP with  $m$  variables (which we can do *very* efficiently [WB09a, WB08, WB07, MB09, BMDP02, TJB01, ZJM08, CZJM07, TJ02, BF04, KGBM03, FBD08, JGR06]). Examples in the next section show that  $V^{\text{lb}}$  is often a good choice for  $V^{\text{clf}}$ .

## 4 Examples

In this section we compute our bound for several example problems, and compare it to the performance achieved by the control-Lyapunov suboptimal policy (which we evaluate via Monte Carlo simulation). To keep our problem descriptions short, we will illustrate our bound for only the infinite horizon problem. Very similar results are observed for the finite horizon case. Table 1 summarizes our results for all three examples. Here,  $J^{\text{lb}}$  is the lower bound found by our method,  $J^{\text{clf}}$  is the objective achieved by control-Lyapunov policy, and  $\eta = (J^{\text{clf}} - J^{\text{lb}})/J^{\text{lb}}$ .

### 4.1 Small example

The first example is a small problem with  $n = 6$  states and  $m = 2$  inputs. The  $A, B$  matrices are generated randomly: The entries of each matrix is drawn from a standard

	finite input	positive input	supply chain
$J^{\text{clf}}$	160.0	51.0	44.6
$J^{\text{lb}}$	157.0	47.4	39.5
$\eta$	1.9%	7.6%	12.9%

**Table 1:** Performance of control-Lyapunov policy and lower bounds for three examples.

normal distribution, and then  $A$  is scaled so that its spectral radius is less than one (which ensures that the open-loop is stable). This is not needed to compute the bound, but we find that the performance of suboptimal policies can often be poor for highly unstable systems. The stage costs are quadratic with the form in (40), where  $\bar{R} = I$ ,  $\bar{Q} = I$ ,  $\bar{S} = 0$ ,  $\bar{q} = \mathbf{1}$ ,  $\bar{r} = \mathbf{1}$  and  $\bar{s} = 0$ . The disturbance  $w_t$  has distribution  $\mathcal{N}(\mathbf{1}, 0.25I)$ . There are no state constraints, and the input constraint set is finite with  $K = 15$  points, *i.e.*,  $\mathcal{C} = \mathbf{R}^n \times \mathcal{U}$ , where  $\mathcal{U} = \{u^{(1)}, \dots, u^{(K)}\}$ . Each entry of  $u^{(i)}$  is randomly drawn from a standard normal distribution.

**Results:** For this small problem, the average objective value achieved by the control-Lyapunov policy is 160.0. (This is averaged over 1000 time steps in statistical steady state.) The lower bound we compute for this problem is 157.0. Thus, for this problem instance we conclude that the control-Lyapunov policy, as well as our lower bound, are both within 2% of  $J^*$ .

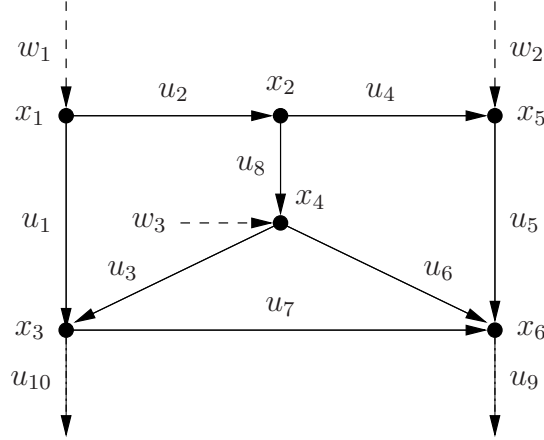
## 4.2 Nonnegative control

The second example is generated in the same way as the first example, except that it is larger, with  $n = 30$  states and  $m = 5$  inputs. The stage costs are quadratic with  $\bar{R} = I$ ,  $\bar{Q} = I$ ,  $\bar{S} = 0$ ,  $\bar{q} = \mathbf{1}$ ,  $\bar{r} = \mathbf{1}$  and  $\bar{s} = 0$ . The disturbance  $w_t$  is Gaussian with mean  $\bar{w} = 0$  and covariance  $W = I$ . Again, there are no state constraints; the input constraint set is  $\mathcal{U} = \{v \mid v \geq 0\}$ .

**Results:** For the nonnegative control problem, the average objective value achieved by the control-Lyapunov policy is 51.0. (As before, this is averaged over 1000 time steps in statistical steady state.) Our method gives the lower bound 47.4. Thus we conclude that both the suboptimal policy, as well as our bound, are within 10% of  $J^*$ .

## 4.3 Supply chain

Our third problem instance is a single commodity supply chain with  $n = 6$  nodes, that represent warehouses (or buffers), and 13 uni-directional links, over which the commodity can be transported from one node to another (this is the same example as [WB08]). This is shown



**Figure 1:** Supply chain model. Dots represent nodes or warehouses. Arrows represent links or commodity flow. Dashed arrows are inflows and dash-dot arrows are outflows.

in figure 1. Three of these links, represented by dashed arrows, are inflows, which represent random arrivals of the commodity at each warehouse (these cannot be controlled). We denote the vector of inflows at time  $t$  by  $w_t$ . We assume that  $w_t$  is exponentially distributed with  $\bar{w} = \mathbf{1}$  (hence  $W = I$ ). The remaining  $m = 10$  links are the controls. At time  $t$  we denote the vector of commodity transported along these links by  $u_t$ . Each component of  $u_t$  is constrained to lie in the interval  $[0, 2.5]$ . The system state  $x_t$  denotes the amount of commodity present at each node, and is constrained to be nonnegative, *i.e.*,  $x_t \geq 0$ . The final constraint is that the total flow out of any node, at any time, cannot exceed the amount of commodity available at the node (which is a linear inequality constraint involving  $x_t$  and  $u_t$ ). The objective is also quadratic with  $\bar{Q} = I$ ,  $\bar{R} = 0$ ,  $\bar{S} = 0$ ,  $\bar{q} = \mathbf{1}$ ,  $\bar{r} = \mathbf{1}$ . This means that there is a storage cost at each node, with value  $(x_t)_i + (x_t)_i^2$ , and a charge for transporting the commodity along each edge.

**Results:** For the supply chain problem, the average objective value achieved by the control-Lyapunov policy is 44.6 (averaged over 1000 time steps in statistical steady state). Our lower bound is 39.5. This shows that the control-Lyapunov policy, as well as our lower bound, are both within around 10% of  $J^*$ .

## 5 Conclusions

In this paper we have described a method for computing lower bounds on the optimal objective value of linear stochastic control problems. Our method naturally yields an approximate value function that can be used with a suboptimal control method, such as the control-Lyapunov policy. In many examples we find that the gap between the objective

achieved by the control-Lyapunov policy and our lower bound is small (say, less than 10%), which shows that both are close to  $J^*$ , the optimal value of the control problem. In other words, the controller is nearly optimal, in practical terms.

## References

- [Alt99] E. Altman. *Constrained Markov Decision Processes*. Chapman & Hall, 1999.
- [ARE95] M. Ait-Rami and L. El Ghaoui. Solving non-standard Riccati equations using LMI optimization. In *Proceedings of the 34th Conference on Decision and Control*, pages 3855–3856, 1995.
- [Åst70] K. J. Åström. *Introduction to stochastic control theory*. Academic Press, 1970.
- [Bad82] F. Badawi. On a quadratic matrix inequality and the corresponding algebraic Riccati equation. *International Journal of Control*, 36:313–322, 1982.
- [BB91] S. Boyd and C. Barratt. *Linear Controller Design: Limits of Performance*. Prentice-Hall, Inc., 1991.
- [BEFB94] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in Systems and Control Theory*. SIAM books, Philadelphia, 1994.
- [Ber05] D. Bertsekas. *Dynamic Programming and Optimal Control: Volume 1*. Athena Scientific, 2005.
- [Ber07] D. Bertsekas. *Dynamic Programming and Optimal Control: Volume 2*. Athena Scientific, 2007.
- [BF04] A. Bemporad and C. Filippi. Suboptimal explicit receding horizon control via approximate multiparametric quadratic programming. *Journal of Optimization Theory and Applications*, 117(1):9–38, November 2004.
- [BGT01] D. Bertsimas, D. Gamarnik, and J. Tsitsiklis. Performance of multiclass markovian queueing networks via piecewise linear Lyapunov functions. *Annals of Applied Probability*, 11(4):1384–1428, 2001.
- [BH88] D. Bernstein and W. Haddad. The optimal projection equations with petersen-hollot bounds: Robust stability and performance via fixed order dynamic compensation for systems with structured real valued parameter uncertainty. *IEEE Transactions on Automatic Control*, 33(8):578–582, 1988.
- [BLW91] S. Bittanti, A. Laub, and J. Willems. *The Riccati equation*. Springer-Verlag, 1991.

- [BMDP02] A. Bemporad, M. Morari, V. Dua, and E. Pistikopoulos. The explicit linear quadratic regulator for constrained systems. *Automatica*, 38(1):3–20, January 2002.
- [BS96] D. Bertsekas and S. Shreve. *Stochastic optimal control: The discrete-time case*. Athena Scientific, 1996.
- [BT96] D. Bertsekas and J. Tsitsiklis. *Neuro-Dynamic Programming*. Athena Scientific, 1996.
- [BV04] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [Cas05] D. Castañón. Stochastic control bounds on sensor network performance. In *Proceedings of the 44th IEEE Conference on Decision and Control*, pages 4939–4944, December 2005.
- [CB04] E. Camacho and C. Bordons. *Model Predictive Control*. Springer-Verlag, 2004.
- [CL88] M. Corless and G. Leitmann. Controller design for uncertain systems via Lyapunov functions. In *Proceedings of the American Control Conference*, volume 3, pages 2019–2025, 1988.
- [CL06] R. Cogill and S. Lall. Suboptimality bounds in stochastic control: A queueing example. In *Proceedings of the 2006 American Control Conference*, pages 1642–1647, 2006.
- [CLH07] R. Cogill, S. Lall, and J. Hespanha. A constant factor approximation algorithm for event-based sampling. In *Proceedings of the 2007 American Control Conference*, pages 305–311, 2007.
- [CP72] S. Chang and T. Peng. Adaptive guaranteed cost control of systems with uncertain parameters. *IEEE Transactions on Automatic Control*, 17(4):474–483, 1972.
- [CZJM07] C. Christophersen, M. Zeilinger, C. Jones, and M. Morari. Controller complexity reduction for piecewise affine systems through safe region elimination. In *IEEE Conference on Decision and Control*, pages 4773–4778, December 2007.
- [DV03] D. De Farias and B. Van Roy. The linear programming approach to approximate dynamic programming. *Operations Research*, 51(6):850–865, 2003.
- [FBD08] H. J. Ferreau, H. G. Bock, and M. Diehl. An online active set strategy to overcome the limitations of explicit MPC. *International Journal of Robust and Nonlinear Control: Special Issue on Nonlinear Model Predictive Control of Fast Systems*, 18(8):816–830, 2008.



- [FP96] R. Freeman and J. Primbs. Control Lyapunov functions, new ideas from an old source. In *Proceedings of the 35th IEEE Conference on Decision and Control*, volume 4, pages 3926–3931, 1996.
- [GBY06] M. Grant, S. Boyd, and Y. Ye. CVX: Matlab software for disciplined convex programming, 2006. Available at <http://www.stanford.edu/~boyd/cvx>.
- [GPM89] C. Garcia, D. Prett, and M. Morari. Model predictive control: theory and practice. *Automatica*, 25(3):335–348, 1989.
- [HLL97] O. Hernández-Lerma and J. Lassere. *Discrete-Time Markov Control Processes: Basic Optimality Criteria*. Springer, 1997.
- [JGR06] C. Jones, P. Grieder, and S. Rakovic. A logarithmic-time solution to the point location problem. *Automatica*, 42(12):2215–2218, December 2006.
- [Kal64] R. Kalman. When is a linear control system optimal? *Journal of Basic Engineering*, 86(1):1–10, 1964.
- [KGBM03] M. Kvasnica, P. Grieder, M. Baotić, and M. Morari. Multi-parametric toolbox (MPT). Technical report, December 2003.
- [KH05] W. Kwon and S. Han. *Receding Horizon Control*. Springer-Verlag, 2005.
- [KK94] S. Kumar and P. Kumar. Performance bounds for queueing networks and scheduling policies. *IEEE Transactions on Automatic Control*, 39(8):1600–1611, 1994.
- [KV86] P. Kumar and P. Varaiya. *Stochastic systems*. Prentice-Hall, 1986.
- [LB76] A. Laub and F. Bailey. Suboptimality bounds and stability in the control of non-linear dynamic systems. *IEEE Transactions on Automatic Control*, 21(3):396–399, 1976.
- [LR06] B. Lincoln and A. Rantzer. Relaxing dynamic programming. *IEEE Transactions on Automatic Control*, 51(8):1249–2006, 2006.
- [Mac02] J. Maciejowski. *Predictive Control with Constraints*. Prentice-Hall, 2002.
- [MB09] J. Mattingley and S. Boyd. Automatic code generation for real-time convex optimization. In *Convex optimization in signal processing and communications*, 2009. To appear.
- [NN94] Y. Nesterov and A. Nemirovsky. *Interior-Point Polynomial Methods in Convex Programming*. SIAM, 1994.
- [NW99] J. Nocedal and S. Wright. *Numerical Optimization*. Springer, 1999.

- [PM94] I. Petersen and D. McFarlane. Optimal guaranteed cost control and filtering for uncertain linear systems. *IEEE Transactions on Automatic Control*, 39(9):1971–1977, 1994.
- [Pow07] W. Powell. *Approximate dynamic programming: solving the curses of dimensionality*. John Wiley & Sons, Inc., 2007.
- [PSSV07] A. Peters, M. Salgado, and E. Silva-Vera. Performance bounds in MIMO linear control with pole location constraints. In *Proceedings of the 2007 Mediterranean Conference on Control and Automation*, pages 1–6, 2007.
- [PW00] F. Potra and S. Wright. Interior-point methods. *Journal of Computational and Applied Mathematics*, 124(1-2):281–302, 2000.
- [QB03] S. Qin and T. Badgwell. A survey of industrial model predictive control technology. *Control Engineering Practice*, 11(7):733–764, 2003.
- [RZ94] U. Rieder and R. Zagst. Monotonicity and bounds for convex stochastic control models. *Mathematical Methods of Operations Research*, 39(2):1432–5217, 1994.
- [RZ00] M. Rami and X. Zhou. Linear matrix inequalities, Riccati equations and indefinite stochastic linear quadratic controls. *IEEE Transactions on Automatic Control*, 45(6):1131–1143, June 2000.
- [Soe92] R. Soeterboek. *Predictive Control: A Unified Approach*. Prentice Hall, 1992.
- [Son83] E. Sontag. A Lyapunov-like characterization of asymptotic controllability. *SIAM Journal on Control and Optimization*, 21(3):462–471, 1983.
- [SS85] P. Schweitzer and A. Seidmann. Generalized polynomial approximations in Markovian decision processes. *Journal of Mathematical Analysis and Applications*, 110(2):568–582, 1985.
- [SSC03] M. Sznaier, R. Suarez, and J. Cloutier. Suboptimal control of constrained nonlinear systems via receding horizon constrained control Lyapunov functions. *International Journal on Robust and Nonlinear Control*, 13(3-4):247–259, 2003.
- [TJ02] P. Tøndel and T. A. Johansen. Complexity reduction in explicit linear model predictive control. In *15th IFAC World Congress on Automatic Control*, July 2002.
- [TJB01] P. Tøndel, T. A. Johansen, and A. Bemporad. An algorithm for multi-parametric quadratic programming and explicit MPC solutions. In *IEEE Conference on Decision and Control*, pages 1199–1204, 2001.
- [VB96] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38(1):49–95, 1996.

- [VGN04] P. Vuthandam, H. Genceli, and M. Nikolaou. Performance bounds for robust quadratic dynamic matrix control with end condition. *AIChE Journal*, 41(9):2083–2097, 2004.
- [WB07] B. Wegbreit and S. Boyd. Fast computation of optimal contact forces. *IEEE Transactions on Robotics*, 23(6):1117–1132, December 2007.
- [WB08] Y. Wang and S. Boyd. Fast model predictive control using online optimization. In *Proceedings of the 17th IFAC world congress*, pages 6974–6997, 2008.
- [WB09a] Y. Wang and S. Boyd. Fast evaluation of control-Lyapunov policy. 2009. In preparation.
- [WB09b] Y. Wang and S. Boyd. Performance bounds for linear stochastic control. *System and Control Letters*, 58(3):178–182, 2009.
- [Whi82] P. Whittle. *Optimization over Time*. John Wiley & Sons, Inc., 1982.
- [Wit70] H. Witsenhausen. On performance bounds for uncertain systems. *SIAM Journal on Control*, 8(1):55–89, 1970.
- [Wri97a] S. Wright. Applying new optimization algorithms to model predictive control. *Chemical Process Control-V*, 93(316):147–155, 1997. AIChE Symposium Series.
- [Wri97b] S. Wright. *Primal-Dual Interior-Point Methods*. SIAM, 1997.
- [ZDG95] K. Zhou, J. Doyle, and K. Glover. *Robust and optimal control*. Prentice Hall, 1995.
- [ZJM08] M. Zeilinger, C. Jones, and M. Morari. Real-time suboptimal model predictive control using a combination of explicit MPC and online computation. In *IEEE Conference on Decision and Control*, pages 4718–4723, December 2008.