INTERNATIONAL JOURNAL OF ROBUST AND NONLINEAR CONTROL
Int. J. Robust. Nonlinear Control 2011; **21**:1710–1728
Published online 3 December 2010 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/rnc.1665

Performance bounds and suboptimal policies for linear stochastic control via LMIs

Yang Wang*,† and Stephen Boyd

Packard Electrical Engineering, 350 Serra Mall, Stanford, CA 94305, U.S.A.

SUMMARY

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. In addition, we derive our bounds using a new method, where we relax the Bellman equation to an inequality. 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. Copyright © 2010 John Wiley & Sons, Ltd.

Received 17 September 2009; Revised 7 August 2010; Accepted 14 September 2010

KEY WORDS: dynamic programming; stochastic control; convex optimization

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 [1–3]. 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) [4–7]. 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., [2, 3, 8–13].

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 see that for many practical control problems, the bound can be effectively computed by solving a convex optimization problem. Thus,

*Correspondence to: Yang Wang, Packard 243, 350 Serra Mall, Stanford, CA 94305, U.S.A.
†E-mail: yw224@stanford.edu

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 with 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 [14], 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. Our method is based on relaxing the Bellman equation to an inequality, and looking for quadratics that bound the stage cost and value functions. We can then optimize over this family of bounds by solving an optimization problem. This new method is conceptually more elegant, and it will also allow us to extend our bounds to more general settings, such as problems with polynomial dynamics, constraints, and objective functions. (However, in this paper we focus exclusively on the linear quadratic case.) 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 [15], 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 [16], 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 [17], 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 [17], as well as event-based sampling [18], and typically yields analytic bounds that apply to entire problem classes. Another example of performance bounds in this area is [19]. 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 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

DOI: 10.1002/rnc

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 [20–24].

There are many other related works we will not summarize, including more theoretical contributions [25–27], other application focussed papers [28–33], as well as books on approximate dynamic programming methods and stochastic control [8, 34, 35]. 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 $t=0,\ldots,N$, with dynamics

$$x_{t+1}=A_t x_t + B_t u_t + w_t, \quad t=0,1,\ldots,N-1, \tag{1}$$

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 $W_t = \mathbf{E}(w_t-\bar{w}_t)(w_t-\bar{w}_t)^{\mathrm{T}}$. We will also assume that x_0 is random, with mean $\bar{x}_0=\mathbf{E}x_0$, and covariance $X_0=\mathbf{E}(x_0-\bar{x}_0)(x_0-\bar{x}_0)^{\mathrm{T}}$, 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 $x_0,\ldots,x_t$. 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.

$$u_t=\psi_t(x_t), \quad t=0,1,\ldots,N-1, \tag{2}$$

where $\psi_t : \mathbf{R}^n \to \mathbf{R}^m$ is the state feedback function, or control policy, at time t. 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 $\psi_0,\ldots,\psi_{N-1}$, 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 $\ell_t : \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}$, $t=0,\ldots,N-1$ is the stage cost function at time t, and $\ell_N : \mathbf{R}^n \to \mathbf{R}$ is the stage cost function at time N, sometimes referred to as the terminal cost function. We will assume that the above expectation exists.

We also have state and control constraints

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

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

The stochastic control problem is to find the state feedback functions $\psi_0,\ldots,\psi_{N-1}$ that minimize the objective J, among those that satisfy constraint (3). The problem data consists of $A_0,\ldots,A_{N-1}$,

$B_0, \ldots, B_{N-1}$, the distribution of x_0 and each w_t, the stage cost functions $\ell_0, \ldots, \ell_N$, and the constraint sets $\mathcal{C}_0, \ldots, \mathcal{C}_N$. We let $J^\star$ 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., [2, 3, 8–10, 25, 36–39].

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 $\bar{\ell}_t : \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R} \cup \{\infty\}$, $t = 0, \ldots, N-1$, as

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

Similarly, we define $\bar{\ell}_N : \mathbf{R}^n \to \mathbf{R} \cup \{\infty\}$ 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 $x_t = z$,

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

subject to the dynamics (1). ($V_t : \mathbf{R}^n \to \mathbf{R} \cup \{\infty\}$ 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 $J^\star = \mathbf{E} V_0(x_0)$, where the expectation is over x_0. The functions $V_0, \ldots, V_N$ satisfy the Bellman recursion,

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

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

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

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 $f : \mathbf{R}^n \to \mathbf{R} \cup \{\infty\}$.

The optimal feedback functions are

$$\psi_t^\star(z) = \arg\min_v \{\bar{\ell}_t(z, v) + \mathbf{E} V_{t+1}(A_t z + B_t v + w_t)\}, \quad t = 0, \ldots, N-1. \tag{5}$$

The value functions and optimal feedback functions can be effectively computed in only a few special cases. The most famous example is when $\mathcal{C}_0 = \cdots = \mathcal{C}_{N-1} = \mathbf{R}^n \times \mathbf{R}^m$, $\mathcal{C}_N = \mathbf{R}^n$ (there are no constraints on the input and state) and $\ell_0, \ldots, \ell_N$ are convex quadratic functions [1]. In this case the optimal state feedback functions are affine, i.e., $u_t = K_t x_t + g_t$, $t = 0, \ldots, N-1$, 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 [2, 3, 10, 36].

DOI: 10.1002/rnc

2.2. Basic bound

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

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

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

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

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

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

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

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

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

$$\begin{aligned} &\sup_{(z,v)\in\mathcal{C}_t} \left(\tilde{\ell}_t(z, v) - \ell_t(z, v)\right) \leqslant 0, \quad t = 0, \ldots, N-1, \\ &\sup_{z\in\mathcal{C}_N} \left(\tilde{\ell}_N(z) - \ell_N(z)\right) \leqslant 0. \end{aligned} \tag{6}$$

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

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

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

$$\tilde{\mathcal{T}}_t f \leqslant \mathcal{T}_t f, \quad t = 0, \ldots, N-1. \tag{7}$$

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

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

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

$$\tilde{V}_t \leqslant \tilde{\mathcal{T}}_t \tilde{V}_{t+1}, \quad t = 0, \ldots, N-1, \quad \tilde{V}_N = \tilde{\ell}_N. \tag{8}$$

Then we claim that

$$\tilde{V}_t \leqslant V_t, \quad t = 0, \ldots, N, \tag{9}$$

which gives us the lower bound

$$\mathbf{E}\tilde{V}_0(x_0) \leqslant \mathbf{E} V_0(x_0) = J^\star. \tag{10}$$

The left-hand side can be explicitly given as

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

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

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

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

DOI: 10.1002/rnc

Here, the first and second inequalities follow from (8) and (7). The third inequality follows from monotonicity of the Bellman operator [3, 25, 26, 36], i.e. $f \leqslant g$ implies $\mathcal{T}_{N-1} f \leqslant \mathcal{T}_{N-1} g$, and the condition (6). Using the same argument we get

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

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

In other words, if we can find

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

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

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

In the remainder 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

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

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

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

This is equivalent to the condition,

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

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

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

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, inequalities (11) are linear matrix inequalities (LMIs) [40–46]. 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 $\tilde{V}_N = \tilde{\ell}_N$ can be written as

$$P_N = Q_N, \quad p_N = q_N, \quad c_N = s_N, \tag{12}$$

which is a set of linear equality constraints.

DOI: 10.1002/rnc

2.4. Optimizing the bound

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

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

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

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

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

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

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, constraint (6) is a semi-infinite constraint, since it is really a family of constraints parameterized by the infinite sets $\mathcal{C}_0, \ldots, \mathcal{C}_N$ [40].

The idea behind our bound is to find functions $\tilde{\ell}_t$ that are everywhere smaller than the stage cost functions ℓ_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 $J^\star$. 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 problem (13) exactly. In other cases, we can replace the condition (6) with a conservative approximation, which still yields a lower bound on $J^\star$. 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 optimization problem (13) exactly. We assume that the stage costs are quadratic with the form

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

and terminal cost is quadratic with the form

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

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

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

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

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

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

$$z^{\mathrm{T}} Q_N z + 2q_N^{\mathrm{T}} z + s_N \leqslant z^{\mathrm{T}} \bar{Q}_N z + 2\bar{q}_N^{\mathrm{T}} z + \bar{s}_N \quad \text{for all } z \in \mathbf{R}^n. \tag{17}$$

DOI: 10.1002/rnc

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

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

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

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

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

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

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. [40, 42, 47–49]).

2.6. S-procedure relaxation

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

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

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

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

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

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

and

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

A sufficient condition for (6) is

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

which is equivalent to

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

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

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

A sufficient condition for (21) and (22) is (by the so-called $\mathcal{S}$-procedure [40, 41]) the existence of nonnegative $\lambda_t^{(1)}, \dots, \lambda_t^{(M_t)}$, $t=0,\dots,N$ such that

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

and

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

These can be written as LMIs

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

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

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

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

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

with variables $Q_t, q_t, s_t, P_t, p_t, c_t$, $t = 0, \ldots, N$, S_t, R_t, r_t, $t = 0, \ldots, N-1$, and $\lambda_t^{(i)}$, $i = 1, \ldots, M$, $t = 0, \ldots, N$. Again, this can be solved very efficiently [40, 42, 47–49].

2.7. *Control-Lyapunov policy*

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

In control-Lyapunov feedback, we modify optimal feedback function (5) by replacing the optimal value function V_{t+1}, with an approximate value function $V_{t+1}^{\mathrm{clf}} : \mathbf{R}^n \to \mathbf{R}$, which we call a *control-Lyapunov function* [4–7]. The state feedback function at time t is given by

$$\psi_t^{\mathrm{clf}}(z) = \arg \min_{(z,v) \in \mathcal{C}_t} (\ell_t(z, v) + \mathbf{E} V_{t+1}^{\mathrm{clf}}(A_t z + B_t v + w_t)). \tag{26}$$

The performance of this feedback policy clearly relies on good choices for $V_1^{\mathrm{clf}}, \ldots, V_N^{\mathrm{clf}}$. 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 $V_1^{\mathrm{clf}}, \ldots, V_N^{\mathrm{clf}}$ 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 $J^\star$ (either exactly, or by solving a conservative approximation of (13)), we obtain the lower bound functions $V_0^{\mathrm{lb}}, \ldots, V_N^{\mathrm{lb}}$, where

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

Here, P_t^{lb}, p_t^{lb}, c_t^{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 $V_0^{\mathrm{lb}}, \ldots, V_N^{\mathrm{lb}}$, as value functions for an unconstrained problem that approximates our original problem. Thus, we expect that $V_1^{\mathrm{lb}}, \ldots, V_N^{\mathrm{lb}}$ would be good choices for a control-Lyapunov policy. In this case, the state feedback function can be written as

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

DOI: 10.1002/rnc

where

$$\hat{R}_t = B_t^{\mathrm{T}} P_{t+1}^{\mathrm{lb}} B_t, \quad \hat{Q}_t = A_t^{\mathrm{T}} P_{t+1}^{\mathrm{lb}} A_t, \quad \hat{S}_t = A_t^{\mathrm{T}} P_{t+1}^{\mathrm{lb}} B_t,$$
$$\hat{r}_t = B_t^{\mathrm{T}} P_{t+1}^{\mathrm{lb}} \bar{w}_t + B_t^{\mathrm{T}} p_{t+1}^{\mathrm{lb}}, \quad \hat{q}_t = A_t^{\mathrm{T}} P_{t+1}^{\mathrm{lb}} \bar{w}_t + A_t^{\mathrm{T}} p_{t+1}^{\mathrm{lb}}.$$

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 [13, 50–52]. 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 [11, 53–59].

In Section 4 we 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,

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

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 $W = \mathbf{E}(w_t - \bar{w})(w_t - \bar{w})^{\mathrm{T}}$. We also assume that x_0 is random, and independent of all w_t, but we 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 $x_0, \ldots, x_t$. For the problem, we will consider, it is also possible to show that the there is an optimal policy that is time invariant and depends only on the current state, i.e.

$$u_t = \psi(x_t), \quad t = 0, 1, \ldots, \tag{28}$$

where $\psi : \mathbf{R}^n \to \mathbf{R}^m$ 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

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

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

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

where $\mathcal{C} \subseteq \mathbf{R}^m$ is a nonempty constraint set. The stage cost ℓ 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 ψ that minimizes the objective J and satisfies constraint (30). We let $J^\star$ denote the optimal value of J and we let $\psi^\star$ denote the optimal state feedback function. The problem data are A, B, the distribution of w_t, the stage cost function ℓ, and the constraint set $\mathcal{C}$.

For more on the formulation of the stochastic control problem, including technical details (e.g. finiteness of $J^\star$, existence, and uniqueness of an optimal state feedback function), see [2, 3, 8–10, 36].

DOI: 10.1002/rnc

3.1. Dynamic programming 'solution'

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

$$\alpha + V = \mathcal{T}V, \tag{31}$$

where $V:\mathbf{R}^n \to \mathbf{R}\cup\{\infty\}$, $\alpha\in\mathbf{R}$. 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 $f:\mathbf{R}^n \to \mathbf{R}\cup\{\infty\}$, 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 α that satisfies (31), then $J^\star = \alpha$, and the optimal feedback functions are

$$\psi^\star(z) = \arg\min_v \{\bar{\ell}(z,v) + \mathbf{E} V(Az+Bv+w_t)\}. \tag{32}$$

Notice that if α, V satisfy (31), then α, $V+\beta$ 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 α, V that satisfy (31). In this paper, we consider only the cases where the stochastic control problem is feasible, the optimal average cost per stage is finite, and a solution exists to Bellman equation (31). For the technical details, including the conditions under which a solution to the Bellman equation exists, see, e.g., [2, 3, 10, 36].

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

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

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

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

where $V:\mathbf{R}^n \to \mathbf{R}\cup\{\infty\}$ and $\alpha\in\mathbf{R}$ satisfy Bellman equation (31). As with the finite-horizon case, the function V and the constant α can be computed only in a few special cases. One example is where $\mathcal{C}=\mathbf{R}^n\times\mathbf{R}^m$ and the stage cost ℓ 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 Section 2.2. Here, we make use of the value iteration described in Section 3.1 to show one of our inequalities.

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

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

DOI: 10.1002/rnc

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

$$\sup_{(z,v)\in\mathcal{C}} \left(\tilde{\ell}(z, v) - \ell(z, v)\right) \leqslant 0. \tag{34}$$

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

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

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

$$\tilde{\mathcal{T}} f \leqslant \mathcal{T} f. \tag{35}$$

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

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

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

$$\tilde{\alpha} + \tilde{V} \leqslant \tilde{\mathcal{T}} \tilde{V}. \tag{36}$$

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

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

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

$$\tilde{\alpha} \leqslant \alpha = J^\star \tag{37}$$

so $\tilde{\alpha}$ is our lower bound. To prove this, we can write $V^{(k)}$ (the kth 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

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

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

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

Thus, we conclude $\tilde{\alpha} \leqslant \alpha$.

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 $J^\star$. As with the finite-horizon case, we will focus on the cases where we can effectively compute (and optimize) this bound.

DOI: 10.1002/rnc

3.3. Bellman inequality as an LMI

We can express the Bellman inequality (36) as

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

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

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

This is equivalent to the condition,

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

which can be written as

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

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, inequality (38) is an LMI [40–42, 44, 45].

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

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

Condition (34) is convex, since the constraint

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

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, LMI (38) defines a convex constraint set, thus the optimization problem (39) is a convex optimization problem [40].

In the general case, condition (34) is a semi-infinite constraint, since it is a family of constraints parametrized by the infinite set $\mathcal{C}$. In the following 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 $J^\star$.

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

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

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 $\mathcal{U}=\{u^{(1)},\ldots,u^{(K)}\}\subseteq \mathbf{R}^m$. Condition (34) becomes

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

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

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

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

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

with variables Q, S, R, q, r, s, P, p, and $\tilde{\alpha}$ (which we can solve using interior-point methods [40, 42, 47–49]).

3.6. S-procedure relaxation

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

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

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

$$\mathcal{C}\subseteq\tilde{\mathcal{C}}=\{(z,v)\,|\,f^{(i)}(z,v)\leqslant 0,\ i=1,\ldots,M\}.$$

A sufficient condition for (34) is

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

which is equivalent to

$$f^{(i)}(z,v)\leqslant 0,\ i=1,\ldots,M \Longrightarrow \tilde{\ell}(z,v)\leqslant\ell(z,v). \tag{44}$$

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

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

DOI: 10.1002/rnc

This can be written as the LMI

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

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

$$\begin{array}{ll} \text{maximize} & \tilde{\alpha} \\ \text{subject to} & (45), (38) \\ & \lambda_i \geqslant 0, \quad i=1,\ldots,M \end{array} \tag{46}$$

with variables Q, S, R, q, r, s, P, p, $\tilde{\alpha}$, and $\lambda_1,\ldots,\lambda_M$. Again, this can be effectively solved (see [40, 42, 47–49]). Note that when $\bar{S}=0$, $\bar{q}=0$, $\bar{r}=0$, and $F^{(i)}=0$, $G^{(i)}=0$, $g^{(i)}=0$, and $h^{(i)}=0$, $i=1,\ldots,M$, (46) simplifies to the SDP obtained in our previous paper [14], and gives the same bound.

3.7. Control-Lyapunov policy

As with the finite-horizon case, to get the CLF for the infinite-horizon problem we modify the optimal feedback function (32) by replacing the optimal V, with an approximation $V^{\mathrm{clf}}:\mathbf{R}^n \to \mathbf{R}$. We call V^{clf} the steady-state control-Lyapunov function [4–7]. In this case, the state feedback function is given by

$$\psi^{\mathrm{clf}}(z) = \arg\min_{(z,v)\in\mathcal{C}} (\ell(z,v) + \mathbf{E} V^{\mathrm{clf}}(Az+Bv+w_t)). \tag{47}$$

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

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

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

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

where

$$\hat{R} = B^{\mathrm{T}} P^{\mathrm{lb}} B, \quad \hat{Q} = A^{\mathrm{T}} P^{\mathrm{lb}} A, \quad \hat{S} = A^{\mathrm{T}} P^{\mathrm{lb}} B,$$
$$\hat{r} = B^{\mathrm{T}} P^{\mathrm{lb}} \bar{w} + B^{\mathrm{T}} p^{\mathrm{lb}}, \quad \hat{q} = A^{\mathrm{T}} P^{\mathrm{lb}} \bar{w} + A^{\mathrm{T}} p^{\mathrm{lb}}.$$

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 [11, 13, 50–60]). Examples in the following section show that V^{lb} is often a good choice for V^{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). The first three problems are all infinite-horizon problems, while problem 4 is

DOI: 10.1002/rnc

Table I. Performance of control-Lyapunov policy and lower bounds for three examples.

	Finite input	Positive input	Supply chain	Finite horizon
J^{clf}	160.0	51.0	44.6	245.9
J^{lb}	157.0	47.4	39.5	211.0
η	1.9%	7.6%	12.9%	16.5%

a finite-horizon problem. Table I summarizes our results for all three examples. Here, J^{lb} is the lower bound found by our method, J^{clf} is the objective achieved by the 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. A, B matrices are generated randomly: The entries of each matrix are drawn from a standard 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 $\mathcal{U}=\{u^{(1)}, \ldots, u^{(K)}\}$. 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 $J^\star$.

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 $\mathcal{U}=\{v \mid v \geqslant 0\}$.

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 $J^\star$.

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 [13]). This is shown 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., $x_t \geqslant 0$. 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.

DOI: 10.1002/rnc

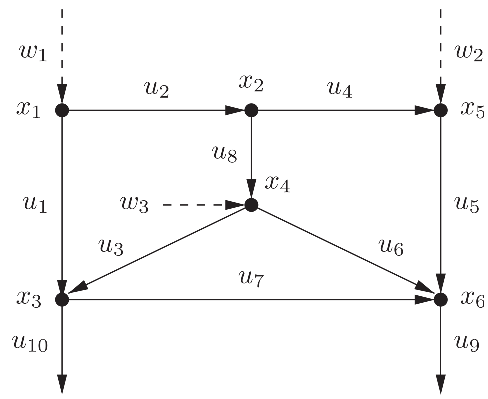

Figure 1. Supply chain model. Dots represent nodes or warehouses. Arrows represent links or commodity flow. Dashed arrows are inflows and dash-dot arrows are outflows.

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 $J^\star$.

4.4. Finite horizon

Our last example is a finite-horizon nonnegative control example. The problem instance is generated in the same way as examples 1 and 2, with $n=8$ states, $m=3$ inputs, and horizon $N=15$. The stage costs are all quadratic with $\bar{R}_t=I$, $\bar{Q}_t=I$, $\bar{S}_t=0$, $\bar{q}_t=\mathbf{1}$, $\bar{r}_t=\mathbf{1}$, $\bar{s}_t=0$, $t=0,\ldots,N-1$, and $\bar{Q}_N=I$, $\bar{q}_N=\mathbf{1}$, $\bar{s}_N=0$. The disturbance w_t is Gaussian with mean $\bar{w}=0$ and covariance $W=I$. The initial state is also Gaussian with mean $\bar{x}_0=0$ and covariance $X_0=I$. There are no state constraints; the input constraint set is $\mathcal{U}_t=\{v|v\geqslant 0\}$.

Results: For this problem instance, the average objective value achieved by the finite-horizon control-Lyapunov policy is 245.9 (averaged over 1000 runs, where each run consists of $N=15$ steps). The bound we get is 211.0, so both the J^{clf} and J^{lb} are within around 15% of $J^\star$.

5. CONCLUSIONS AND EXTENSIONS

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 $J^\star$, the optimal value of the control problem. In other words, the controller is nearly optimal, in practical terms.

Our method directly extends to the case where the dynamics, constraints and objective functions are polynomials. In this case, we look for polynomial lower bounds on the stage cost and value functions. The derivation of the bounds is exactly the same as for the quadratic case, except that to get a sufficient condition for the lower bound we use the sum-of-squares procedure instead of the $\mathcal{S}$-procedure. The resulting set of inequalities is still convex, with a tractable number of variables and constraints as long as the degree of the polynomials as well as the state and input dimensions are small.

The same methods can also be used to obtain piecewise quadratic bounds. For example, for the finite-horizon case, the lower bound condition defines a family of quadratic lower bounds on the value function. A simple observation is that the supremum over this family of lower bounds is also a lower bound (in many cases a much better bound than the ones shown here). We will examine these extensions in more detail in forthcoming publications.

DOI: 10.1002/rnc

REFERENCES

1. Khalman R. When is a linear control system optimal? *Journal of Basic Engineering* 1964; **86**(1):1–10.
2. Bertsekas D. *Dynamic Programming and Optimal Control*, vol. 1. Athena Scientific: Nashua, NH, U.S.A., 2005.
3. Bertsekas D. *Dynamic Programming and Optimal Control*, vol. 2. Athena Scientific: Nashua, NH, U.S.A., 2007.
4. Sontag E. A Lyapunov-like characterization of asymptotic controllability. *SIAM Journal on Control and Optimization* 1983; **21**(3):462–471.
5. Freeman R, Primbs J. Control Lyapunov functions, new ideas from an old source. *Proceedings of the 35th IEEE Conference on Decision and Control*, Kobe, Japan, vol. 4, 1996; 3926–3931.
6. Corless M, Leitmann G. Controller design for uncertain systems via Lyapunov functions. *Proceedings of the American Control Conference*, Atlanta, U.S.A., vol. 3, 1988; 2019–2025.
7. Sznaier M, Suarez R, Cloutier J. Suboptimal control of constrained nonlinear systems via receding horizon constrained control Lyapunov functions. *International Journal on Robust and Nonlinear Control* 2003; **13**(3–4):247–259.
8. Powell W. *Approximate Dynamic Programming*: *Solving the Curses of Dimensionality*. Wiley: New York, 2007.
9. Åström KJ. *Introduction to Stochastic Control Theory*. Academic Press: New York, 1970.
10. Whittle P. *Optimization Over Time*. Wiley: New York, 1982.
11. Bemporad A, Morari M, Dua V, Pistikopoulos E. The explicit linear quadratic regulator for constrained systems. *Automatica* 2002; **38**(1):3–20.
12. Maciejowski J. *Predictive Control with Constraints*. Prentice-Hall: Englewood Cliffs, NJ, 2002.
13. Wang Y, Boyd S. Fast model predictive control using online optimization. *Proceedings of the 17th IFAC World Congress*, Seoul, South Korea, 2008; 6974–6997.
14. Wang Y, Boyd S. Performance bounds for linear stochastic control. *System and Control Letters* 2009; **58**(3):178–182.
15. Lincoln B, Rantzer A. Relaxing dynamic programming. *IEEE Transactions on Automatic Control* 2006; **51**(8):1249–2006.
16. De Farias D, Van Roy B. The linear programming approach to approximate dynamic programming. *Operations Research* 2003; **51**(6):850–865.
17. Cogill R, Lall S. Suboptimality bounds in stochastic control: a queueing example. *Proceedings of the 2006*; *American Control Conference*, Minneapolis, MN, U.S.A., 2006; 1642–1647.
18. Cogill R, Lall S, Hespanha J. A constant factor approximation algorithm for event-based sampling. *Proceedings of the 2007 American Control Conference*, New York City, U.S.A., 2007; 305–311.
19. Castañón D. Stochastic control bounds on sensor network performance. *Proceedings of the 44th IEEE Conference on Decision and Control*, Seville, Spain, 2005; 4939–4944.
20. Petersen I, McFarlane D. Optimal guaranteed cost control and filtering for uncertain linear systems. *IEEE Transactions on Automatic Control* 1994; **39**(9):1971–1977.
21. Chang S, Peng T. Adaptive guaranteed cost control of systems with uncertain parameters. *IEEE Transactions on Automatic Control* 1972; **17**(4):474–483.
22. Bernstein D, Haddad W. 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* 1988; **33**(8):578–582.
23. Laub A, Bailey F. Suboptimality bounds and stability in the control of nonlinear dynamic systems. *IEEE Transactions on Automatic Control* 1976; **21**(3):396–399.
24. Zhou K, Doyle J, Glover K. *Robust and Optimal Control*. Prentice-Hall: Englewood Cliffs, NJ, 1995.
25. Witsenhausen H. On performance bounds for uncertain systems. *SIAM Journal on Control* 1970; **8**(1):55–89.
26. Rieder U, Zagst R. Monotonicity and bounds for convex stochastic control models. *Mathematical Methods of Operations Research* 1994; **39**(2):1432–5217.
27. Schweitzer P, Seidmann A. Generalized polynomial approximations in Markovian decision processes. *Journal of Mathematical Analysis and Applications* 1985; **110**(2):568–582.
28. Kumar S, Kumar P. Performance bounds for queueing networks and scheduling policies. *IEEE Transactions on Automatic Control* 1994; **39**(8):1600–1611.
29. Bertsimas D, Gamarnik D, Tsitsiklis J. Performance of multiclass Markovian queueing networks via piecewise linear Lyapunov functions. *Annals of Applied Probability* 2001; **11**(4):1384–1428.
30. Vuthandam P, Genceli H, Nikolaou M. Performance bounds for robust quadratic dynamic matrix control with end condition. *AIChE Journal* 2004; **41**(9):2083–2097.
31. Peters A, Salgado M, Silva-Vera E. Performance bounds in MIMO linear control with pole location constraints. *Proceedings of the 2007 Mediterranean Conference on Control and Automation*, Athens, Greece, 2007; 1–6.
32. Shen B, Wang Z, Shu H, Wei G. On nonlinear H_∞ filtering for discrete-time stochastic systems with missing measurements. *IEEE Transactions on Automatic Control* 2008; **53**(9):2170–2180.
33. Shen B, Wang Z, Shu H, Wei G. H_∞ filtering for nonlinear discrete-time stochastic systems with randomly varying sensor delays. *Automatica* 2009; **45**(4):1032–1037.
34. Bertsekas D, Tsitsiklis J. *Neuro-Dynamic Programming*. Athena Scientific: Nashua, NH, U.S.A., 1996.
35. Altman E. *Constrained Markov Decision Processes*. Chapman & Hall: London, 1999.
36. Bertsekas D, Shreve S. *Stochastic Optimal Control*: *The Discrete-time Case*. Athena Scientific: Nashua, NH, U.S.A., 1996.

DOI: 10.1002/rnc

37. Boyd S, Barratt C. *Linear Controller Design*: *Limits of Performance*. Prentice-Hall Inc: Englewood Cliffs, NJ, 1991.
38. Kumar P, Varaiya P. *Stochastic Systems*. Prentice-Hall: Englewood Cliffs, NJ, 1986.
39. Hernàndez-Lerma O, Lassere J. *Discrete-Time Markov Control Processes*: *Basic Optimality Criteria*. Springer: Berlin, 1997.
40. Boyd S, Vandenberghe L. *Convex Optimization*. Cambridge University Press: Cambridge, 2004.
41. Boyd S, El Ghaoui L, Feron E, Balakrishnan V. *Linear Matrix Inequalities in Systems and Control Theory*. SIAM Books: Philadelphia, 1994.
42. Vandenberghe L, Boyd S. Semidefinite programming. *SIAM Review* 1996; **38**(1):49–95.
43. Rami M, Zhou X. Linear matrix inequalities, Riccati equations and indefinite stochastic linear quadratic controls. *IEEE Transactions on Automatic Control* 2000; **45**(6):1131–1143.
44. Badawi F. On a quadratic matrix inequality and the corresponding algebraic Riccati equation. *International Journal of Control* 1982; **36**:313–322.
45. Ait-Rami M, El Ghaoui L. Solving non-standard Riccati equations using LMI optimization. *Proceedings of the 34th Conference on Decision and Control*, New Orleans, LA, U.S.A., 1995; 3855–3856.
46. Bittanti S, Laub A, Willems J. *The Riccati Equation*. Springer: Berlin, 1991.
47. Grant M, Boyd S, Ye Y. CVX: Matlab software for disciplined convex programming 2006. Available from: http://www.stanford.edu/~boyd/cvx.
48. Wright S. *Primal-Dual Interior-Point Methods*. SIAM: Philadelphia, PA, 1997.
49. Nesterov Y, Nemirovsky A. *Interior-Point Polynomial Methods in Convex Programming*. SIAM: Philadelphia, PA, 1994.
50. Wang Y, Boyd S. Fast evaluation of control-Lyapunov policy. *IEEE Transactions on Control Systems Technology* 2010; Available from: http://www.stanford.edu/~boyd/papers/fast_clf.html, to appear.
51. Wegbreit B, Boyd S. Fast computation of optimal contact forces. *IEEE Transactions on Robotics* 2007; **23**(6):1117–1132.
52. Mattingley J, Boyd S. Automatic code generation for real-time convex optimization. In *Convex Optimization in Signal Processing and Communications*, Palomar DP, Eldar YC (eds). Cambridge University Press, 2010; 1–41.
53. Tøndel P, Johansen TA, Bemporad A. An algorithm for multi-parametric quadratic programming and explicit MPC solutions. *IEEE Conference on Decision and Control*, Orlando, FL, U.S.A., 2001; 1199–1204.
54. Zeilinger M, Jones C, Morari M. Real-time suboptimal model predictive control using a combination of explicit MPC and online computation. *IEEE Conference on Decision and Control*, Cancun, Mexico, 2008; 4718–4723.
55. Christophersen C, Zeilinger M, Jones C, Morari M. Controller complexity reduction for piecewise affine systems through safe region elimination. *IEEE Conference on Decision and Control*, New Orleans, LA, U.S.A., 2007; 4773–4778.
56. Tøndel P, Johansen TA. Complexity reduction in explicit linear model predictive control. *15th IFAC World Congress on Automatic Control*, Barcelona, Spain, 2002.
57. Bemporad A, Filippi C. Suboptimal explicit receding horizon control via approximate multiparametric quadratic programming. *Journal of Optimization Theory and Applications* 2004; **117**(1):9–38.
58. Kvasnica M, Grieder P, Baotić M, Morari M. Multi-parametric toolbox (MPT). *Technical Report*, Eidgenössische Technische Hochschule Zürich, December 2003.
59. Jones C, Grieder P, Rakovic S. A logarithmic-time solution to the point location problem. *Automatica* 2006; **42**(12):2215–2218.
60. Ferreau HJ, Bock HG, Diehl M. 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* 2008; **18**(8):816–830.

DOI: 10.1002/rnc