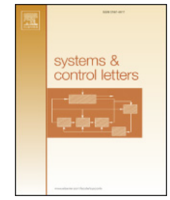
Systems & Control Letters 58 (2009) 178–182

Contents lists available at ScienceDirect

Systems & Control Letters

journal homepage: www.elsevier.com/locate/sysconle



Performance bounds for linear stochastic control

Yang Wang *, Stephen Boyd [1]

Room 243, Packard Electrical Engineering, 350 Serra Mall, Stanford, CA 94305-9505, United States

ARTICLE INFO

Article history:
Received 16 August 2008
Received in revised form
7 October 2008
Accepted 7 October 2008
Available online 22 November 2008

Keywords:
Stochastic control
Model predictive control
Linear matrix inequality
Convex optimization

ABSTRACT

We develop computational bounds on performance for causal state feedback stochastic control with linear dynamics, arbitrary noise distribution, and arbitrary input constraint set. This can be very useful as a comparison with the performance of suboptimal control policies, which we can evaluate using Monte Carlo simulation. Our method involves solving a semidefinite program (a linear optimization problem with linear matrix inequality constraints), a convex optimization problem which can be efficiently solved. Numerical experiments show that the lower bound obtained by our method is often close to the performance achieved by several widely-used suboptimal control policies, which shows that both are nearly optimal. As a by-product, our performance bound yields approximate value functions that can be used as control Lyapunov functions for suboptimal control policies.

1. Linear stochastic control

We consider a discrete-time linear time-invariant system (or plant), with dynamics

$$x(t+1) = Ax(t) + Bu(t) + w(t), \quad t = 0, 1, \ldots, \tag{1}$$

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, $A \in \mathbf{R}^{n\times n}$ is the dynamics matrix, and $B \in \mathbf{R}^{n\times m}$ is the input matrix. We assume that $w(t)$, for different values of t, are zero mean IID. We will also assume that $x(0)$ is random, and 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 $x(0), \ldots, x(t)$, *i.e.*,

$$u(t) = \phi_t(x(0), \ldots, x(t)), \quad t = 0, 1, \ldots,$$

where $\phi_t : \mathbf{R}^{(t+1)n} \to \mathbf{R}^m$. The collection of functions $\phi_0, \phi_1, \ldots$ is called the control policy. For the problem we will consider, it can be shown that there is an optimal policy that is time-invariant and depends only on the current state, *i.e.*, has the form

$$u(t) = \phi(x(t)), \quad t = 0, 1, \ldots, \tag{2}$$

where $\phi : \mathbf{R}^n \to \mathbf{R}^m$. We will refer to ϕ as the state feedback function, or the control policy. For fixed state feedback function ϕ, the Eqs. (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 function, the state and input trajectories become stochastic processes.

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

$$J = \limsup_{T\to\infty} \frac{1}{T} \mathbf{E} \sum_{t=0}^{T-1} \left(\ell_x(x(t)) + \ell_u(u(t))\right), \tag{3}$$

where $\ell_x : \mathbf{R}^n \to \mathbf{R}$ is the state stage cost function, and $\ell_u : \mathbf{R}^m \to \mathbf{R}$ is the input state cost function. (We assume that the expectations exist.) The objective J is the average stage cost. Finally, we impose the control input constraint

$$u(t) \in \mathcal{U} \quad \text{(a.s.)}\,, \; t = 0, 1, \ldots, \tag{4}$$

where $\mathcal{U} \subseteq \mathbf{R}^m$ is a nonempty constraint set with $0 \in \mathcal{U}$. The stage cost functions ℓ_x and ℓ_u, and the input constraint set $\mathcal{U}$, need not be convex.

We can now describe the stochastic control problem. The problem data are A, B, the distribution of $w(t)$, the stage cost functions ℓ_x and ℓ_u, and the input constraint set $\mathcal{U}$; the optimization variable is the state feedback function ϕ. The stochastic control problem is to find the state feedback function ϕ that minimizes the objective J, among those that satisfy the input constraint (4). We will let $J^\star$ denote the optimal value of J, and we let $\phi^\star$ denote an optimal state feedback function.

For more on the formulation of the linear stochastic control problem, including technical details (*e.g.*, finiteness of $J^\star$, existence and uniqueness of an optimal state feedback function), see, *e.g.*, [5,6,26,2,13,27].

* Corresponding author.
E-mail addresses: yw224@stanford.edu (Y. Wang), boyd@stanford.edu (S. Boyd).

1 This material is based upon work supported by the Precourt Institute on Energy Efficiency, by NSF award 0529426, by NASA award NNX07AEIIA, by AFOSR award FA9550-06-1-0514, and by AFOSR award FA9550-06-1-0312.

0167-6911/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.sysconle.2008.10.004

The stochastic control problem can be effectively solved in only a few special cases. The most famous example (described in Section 2 in more detail) is when $\mathcal{U} = \mathbf{R}^m$ (*i.e.*, there are no constraints on the input) and ℓ_x and ℓ_u are convex quadratic functions [14]. In this case the optimal state feedback function is linear, *i.e.*, $u(t) = Kx(t)$, where $K \in \mathbf{R}^{m\times n}$ can be effectively computed from the problem data.

1.1. Suboptimal control policies

Many methods can be used to find a suboptimal state feedback function, *i.e.*, one with (one hopes) a small value of J. We describe three methods in this section; many others can be found in the literature.

1.1.1. Projected linear state feedback

Perhaps the simplest form is a *projected linear state feedback*,

$$\phi_{\mathrm{plsf}}(z) = \mathcal{P}(K_{\mathrm{plsf}}z), \tag{5}$$

where $K_{\mathrm{plsf}} \in \mathbf{R}^{m\times n}$ is a gain matrix (to be chosen), and $\mathcal{P}$ is projection onto $\mathcal{U}$. When $\mathcal{U}$ is a box, *i.e.*,

$$\mathcal{U} = \{u \mid \|u\|_\infty \le U^{\max}\},$$

projection is the same as entry-wise saturation, so the projected linear state feedback policy has the form

$$\phi_{\mathrm{plsf}}(z) = U^{\max}\mathbf{sat}((1/U^{\max})K_{\mathrm{plsf}}z),$$

where the **sat** function is defined for scalar argument as

$$\mathbf{sat}(a) = \begin{cases} a & |a| \le 1 \\ 1 & a > 1 \\ -1 & a < -1, \end{cases}$$

and extended to vectors by acting entry-wise. (Projected linear state feedback is sometimes called saturated linear state feedback in this case.)

1.1.2. Control-Lyapunov feedback

A more sophisticated state feedback function is given by

$$\phi_{\mathrm{clf}}(z) = \operatorname*{argmin}_{v\in\mathcal{U}} \left(\ell_u(v) + \mathbf{E}V_{\mathrm{clf}}(Az + Bv + w(t))\right), \tag{6}$$

where $V_{\mathrm{clf}} : \mathbf{R}^n \to \mathbf{R}$ (which is to be chosen) is called a *control-Lyapunov function* [11,22,12,23]. (The optimal control has this form, for a particular choice of V_{clf}, called the value function or Bellman function for the problem.) When V_{clf} is quadratic, the control-Lyapunov policy (6) can be simplified to

$$\phi_{\mathrm{clf}}(z) = \operatorname*{argmin}_{v\in\mathcal{U}} \left(\ell_u(v) + V_{\mathrm{clf}}(Az + Bv)\right). \tag{7}$$

1.1.3. Certainty-equivalent model predictive control

An even more sophisticated feedback control function is given by *certainty-equivalent model predictive control* (MPC) [16,21,10,15,4,19], in which $\phi(z)$ is found by solving the optimization problem

$$\begin{array}{ll} \text{minimize} & V_{\mathrm{mpc}}(\tilde{x}(T)) + \sum_{\tau=0}^{T-1}\left(\ell_x(\tilde{x}(\tau)) + \ell_u(v(\tau))\right) \\ \text{subject to} & \tilde{x}(\tau+1) = A\tilde{x}(\tau) + Bv(\tau), \quad \tau = 0,\dots,T-1 \\ & v(\tau) \in \mathcal{U}, \quad \tau = 0,\dots,T-1 \\ & \tilde{x}(0) = z, \end{array} \tag{8}$$

with variables $v(0),\dots,v(T-1)$, $\tilde{x}(0),\dots,\tilde{x}(T)$. The function $V_{\mathrm{mpc}} : \mathbf{R}^n \to \mathbf{R}$ is the terminal cost (to be chosen), and T is the horizon (also to be chosen). Let $v^\star(0),\dots,v^\star(T-1)$, $\tilde{x}^\star(0),\dots,\tilde{x}^\star(T)$ be a solution of this problem. The MPC policy is $\phi_{\mathrm{mpc}}(z) = v^\star(0)$, which is a (complicated) function of z through the optimization problem (8). As the horizon T becomes larger, the choice of V_{mpc} becomes less and less important. When the horizon is $T = 1$, the MPC policy reduces to the control-Lyapunov policy (7), with $V_{\mathrm{mpc}} = V_{\mathrm{clf}}$.

1.1.4. Parameters in suboptimal control policies

The art in finding a good suboptimal control policy is in choosing good values for the parameters that appear in them. For projected state feedback, the gain matrix K_{plsf} must be chosen; for a control-Lyapunov policy, V_{clf} must be chosen; and for MPC, the terminal cost function V_{mpc} (and horizon T) must be chosen. A common choice for V_{clf} or V_{mpc} is the (quadratic) value function for a related linear stochastic control problem with no constraints and quadratic stage cost; K_{plsf} can be chosen as the associated optimal gain matrix.

These methods (as well as many others) can give suboptimal state feedback functions that give good performance, *i.e.*, a low value for J. (The objective J is typically evaluated by stochastic simulation, *e.g.*, Monte Carlo.) A natural question that arises is: how close to optimal are these suboptimal control policies? In other words, how much larger than the optimal performance $J^\star$ is J, the performance obtained with a suboptimal control policy? To answer this question, we need to compute a lower bound on $J^\star$, *i.e.*, a bound on achievable performance over all feasible state feedback control functions.

1.2. Performance bounds

In this paper we show how a numerical lower bound on $J^\star$ can be effectively computed, using convex optimization, from the problem data. Our bound is not a generic one, that depends only on the problem dimensions and general assumptions about ℓ_x, ℓ_u, and $\mathcal{U}$; instead, it is computed for each specific problem instance.

We cannot (at this time) guarantee that the bound will be close to $J^\star$. But in a large number of numerical simulations, we have found that our bound is often not too far from the performance achieved by a suboptimal control policy. It is very valuable knowledge in practice to know that a proposed suboptimal control policy attains a specific cost J (found by Monte Carlo simulation), and that the optimal value of the stochastic control problem must exceed a known lower bound J^{lb} (found by the method described in this paper). If the gap between the two is small, we can be certain that our suboptimal control policy is nearly optimal (and that our bound is nearly tight) for this problem instance. The gap can be large, of course, for two reasons: our suboptimal controller is substantially suboptimal, or, our lower bound is poor (for this problem instance).

2. Linear quadratic control

It is well known that the linear stochastic control problem can be effectively solved when $\mathcal{U} = \mathbf{R}^m$ (*i.e.*, there are no constraints on the input) and the stage cost functions have the form

$$\ell_x(z) = z^{\mathrm{T}}Qz, \qquad \ell_u(v) = v^{\mathrm{T}}Rv,$$

where $Q \succeq 0$, $R \succeq 0$ (meaning, they are symmetric positive semidefinite). In this section we (briefly) review these results, since our bound relies on them. For more detailed discussion of the linear quadratic stochastic control problem, see, *e.g.*, [5,6,26].

The optimal cost is

$$J^\star = \mathbf{Tr}(P^\star W),$$

and the optimal state feedback function is

$$\phi^\star(z) = K^\star z,$$

where

$$K^\star = -(R + B^{\mathrm{T}}P^\star B)^{-1}B^{\mathrm{T}}P^\star Az.$$

The matrix W is the covariance of $w(t)$, $W = \mathbf{E}w(t)w(t)^{\mathrm{T}}$. The symmetric matrix $P^\star$ can be effectively found by several methods.

The traditional characterization (and computation) of $P^\star$ is via the algebraic Riccati equation (ARE),

$$P^\star = Q + A^T P^\star A - A^T P^\star B(R + B^T P^\star B)^{-1} B^T P^\star A,$$

along with $P^\star \succeq 0$. (For a more detailed discussion of the ARE and its relation to the linear quadratic stochastic control problem, see, *e.g.*, [5, Section 4.1] [7]) The control policy $\phi^\star(z) = K^\star z$ is called the linear quadratic regulator (LQR) [5, Section 4.1] [8, Section 10.8] [14].

It will be more convenient for us to use a characterization of $P^\star$ and $J^\star$ in terms of a convex optimization problem. Consider the optimization problem

$$\begin{array}{ll} \text{maximize} & \mathbf{Tr}(PW) \\ \text{subject to} & \begin{bmatrix} R + B^T PB & B^T PA \\ A^T PB & Q + A^T PA - P \end{bmatrix} \succeq 0 \\ & P \succeq 0, \end{array} \tag{9}$$

with variable P. The optimal point is $P = P^\star$, and the optimal value of this problem is $J^\star$ [8,1,20,3,7,29].

For future use we make some important comments about this problem.

- The problem (9) is a convex optimization problem, more specifically, a semidefinite program, and can be effectively solved [24,9,18,17,28].
- The block matrix inequality appearing as a constraint is called a linear matrix inequality (LMI) [8].
- The solution $P^\star$ does not depend on W, as long as $W \succ 0$.
- Since the objective and constraints are also convex (in fact, affine) jointly in the variables (P, Q, R), it follows that the optimal value $J^\star$ is a concave function of (Q, R). (See [9, Section 5.6.1].)

In the sequel we will refer to the optimal cost $J^\star$ (feedback gain $K^\star$, value function matrix $P^\star$) for the linear quadratic stochastic control problem with the subscript 'lq' (for 'linear quadratic'), and explicitly show the dependence on the stage cost data Q and R:

$$J^\star_{lq}(Q, R).$$

(The value also depends on the other problem data, specifically, A, B, and W.) We have seen that $J^\star_{lq}$ is concave in (Q, R).

3. Performance bound

3.1. Basic bound

We return now to the general linear stochastic control problem, with input constraint set $\mathcal{U}$ and general stage cost functions ℓ_x and ℓ_u, with optimal value $J^\star$. Suppose $Q \succeq 0$, $R \succeq 0$, and s satisfy the condition

$$z^T Qz + v^T Rv + s \le \ell_x(z) + \ell_u(v), \quad \text{for all } z \in \mathbf{R}^n,\ v \in \mathcal{U}, \tag{10}$$

which can be expressed as

$$\sup_z \left(z^T Qz - \ell_x(z)\right) + \sup_{v\in\mathcal{U}} \left(v^T Rv - \ell_u(v)\right) + s \le 0.$$

Then we have the lower bound

$$J^\star_{lq}(Q, R) + s \le J^\star. \tag{11}$$

The lefthand side can be effectively computed, as described in Section 2. The challenge will be in verifying the condition (10).

We now justify the lower bound (11). Assume that (10) holds. Let $\phi^\star$ be an optimal state feedback function (for the general stochastic problem), and let x and u be the associated (stochastic) trajectories with control policy $\phi^\star$. We have $u(t) \in \mathcal{U}$ a.s., so by (10), we have

$$\begin{aligned} J_{lq} &= \lim_{T\to\infty} \frac{1}{T} \mathbf{E} \sum_{t=0}^{T-1} \left(x(t)^T Qx(t) + u(t)^T Ru(t) + s\right) \\ &\le \lim_{T\to\infty} \frac{1}{T} \mathbf{E} \sum_{t=0}^{T-1} (\ell_x(x(t)) + \ell_u(u(t))) \\ &= J^\star. \end{aligned}$$

Here J_{lq} is the objective of the optimal policy (for the general problem), evaluated with the quadratic objective. The bound (11) follows from $J^\star_{lq} \le J_{lq}$, which holds when the righthand side is evaluated with any feedback function.

We mention one simple case. When the stage costs are quadratic,

$$\ell_x(z) = z^T Q_0 z, \qquad \ell_u(v) = v^T R_0 v, \tag{12}$$

where $Q_0 \succeq 0$, $R_0 \succeq 0$, the simple choice $Q = Q_0$, $R = R_0$, and $s = 0$ satisfies (10). The corresponding lower bound on $J^\star$ is just $J^\star_{lq}(Q_0, R_0)$. This is obvious: the optimal average stage cost, with constraints on the input, is larger than the optimal average stage cost, with no constraints on the input (which we can effectively compute).

3.2. Optimizing the bound

We can optimize the lower bound (11), over the parameters Q, R, and s, by solving the optimization problem

$$\begin{array}{ll} \text{maximize} & J^\star_{lq}(Q, R) + s \\ \text{subject to } (10), & Q \succeq 0, \quad R \succeq 0, \end{array} \tag{13}$$

with variables Q, R, and s. This is a convex optimization problem, since the objective is concave, and the condition (10) is convex (since it is convex for each z and v). In the general case the constraint (10) is a semi-infinite constraint, since it is really a family of constraints, parametrized by the (infinite) set $z \in \mathbf{R}^n$, $u \in \mathcal{U}$.

The idea behind (13) is similar to the basic idea in Lagrangian duality. In Lagrangian duality, we ignore the constraints, but add to the objective an augmenting function that is nonpositive on the feasible set. Minimizing this composite function gives a lower bound on the optimal value of the original problem; optimizing over the parameters that parametrize the augmenting function we obtain the best lower bound obtainable using this technique. We use the same technique here, relying on our ability to solve the stochastic control problem in the special case when the stage costs are quadratic and there are no input constraints.

In a few special cases we can solve the problem (13) exactly. In other cases, we can replace the condition (10) with a conservative approximation, which still yields a lower bound on $J^\star$. We give more specific examples of each of these cases below.

3.3. Quadratic stage cost and finite input constraint set

We assume the stage costs are quadratic, with the form given in (12), and the input constraint set is finite, $\mathcal{U} = \{u_1, \ldots, u_K\}$. The condition (10) is then

$$z^T Qz + u_i^T Ru_i + s \le z^T Q_0 z + u_i^T R_0 u_i, \quad \text{for all } z \in \mathbf{R}^n,\ i = 1, \ldots, K,$$

which is the same as $Q \preceq Q_0$, and

$$u_i^T Ru_i + s \le u_i^T R_0 u_i, \quad i = 1, \ldots, K,$$

a set of linear inequalities on R and s. The problem (13) can be expressed as the SDP

$$\begin{array}{ll} \text{maximize} & \mathbf{Tr}(PW) + s \\ *\text{subject to} & \begin{bmatrix} R + B^T PB & B^T PA \\ A^T PB & Q + A^T PA - P \end{bmatrix} \succeq 0 \\ & P \succeq 0, \quad Q \succeq 0, \quad R \succeq 0, \quad Q \preceq Q_0 \\ & u_i^T R u_i + s \le u_i^T R_0 u_i, \quad i = 1, \dots, K, \end{array}$$

with variables $P, Q, R,$ and s. The LMI is (matrix) monotone in Q, so the optimal value of Q is $Q = Q_0$. So we can just as well drop the variable Q, and replace it with Q_0, to obtain the problem

$$\begin{array}{ll} \text{maximize} & \mathbf{Tr}(PW) + s \\ \text{subject to} & \begin{bmatrix} R + B^T PB & B^T PA \\ A^T PB & Q_0 + A^T PA - P \end{bmatrix} \succeq 0 \\ & P \succeq 0, \quad R \succeq 0 \\ & u_i^T R u_i + s \le u_i^T R_0 u_i, \quad i = 1, \dots, K, \end{array} \tag{14}$$

with variables P, R, and s.

3.4. S-procedure relaxation

We suppose again that the stage costs are quadratic, with the form (12). Suppose we can find $R_1, \dots, R_M$ and $s_1, \dots, s_M$ for which

$$\mathcal{U} \subseteq \tilde{\mathcal{U}} = \{v \mid v^T R_i v + s_i \le 0,\ i = 1, \dots, M\}.$$

A sufficient condition for (10) to hold is

$$z^T Q z + v^T R v + s \le z^T Q_0 z + v^T R_0 v, \quad \text{for all } z \in \mathbf{R}^n,\ v \in \tilde{\mathcal{U}},$$

which holds if and only if $Q \preceq Q_0$ and

$$v^T R_i v + s_i \le 0,\ i = 1, \dots, M \Longrightarrow v^T R v + s \le v^T R_0 v.$$

A sufficient condition for this to hold is (by the so-called S-procedure; see, *e.g.*, [8, Section 2.6.3]) the existence of nonnegative $\lambda_1, \dots, \lambda_M$ which satisfy

$$R - R_0 \preceq \sum_{i=1}^M \lambda_i R_i, \qquad s \le \sum_{i=1}^M \lambda_i s_i.$$

We can optimize over the lower bound by solving the SDP

$$\begin{array}{ll} \text{maximize} & \mathbf{Tr}(PW) + s_0 \\ *\text{subject to} & \begin{bmatrix} R + B^T PB & B^T PA \\ A^T PB & Q + A^T PA - P \end{bmatrix} \succeq 0 \\ & P \succeq 0, \quad Q \succeq 0, \quad R \succeq 0, \quad Q \preceq Q_0 \\ & R - R_0 \preceq \sum_{i=1}^M \lambda_i R_i, \qquad s_0 \le \sum_{i=1}^M \lambda_i s_i \\ & \lambda_i \ge 0, \quad i = 1, \dots, M, \end{array}$$

with variables $P, Q, R, \lambda_1, \dots, \lambda_M$, and $s_0, \dots, s_M$. As before, we can replace Q with Q_0; we can also take $s_0 = \sum_{i=1}^M \lambda_i s_i$. This yields the SDP

$$\begin{array}{ll} \text{maximize} & \mathbf{Tr}(PW) + s^T \lambda \\ \text{subject to} & \begin{bmatrix} R + B^T PB & B^T PA \\ A^T PB & Q_0 + A^T PA - P \end{bmatrix} \succeq 0 \\ & P \succeq 0, \quad R \succeq 0, \quad \lambda \succeq 0 \\ & R - R_0 \preceq \sum_{i=1}^M \lambda_i R_i, \end{array} \tag{15}$$

with variables P, R, and $\lambda \in \mathbf{R}^M$, where we use $\preceq$ between vectors to mean componentwise inequality. The optimal value of this SDP gives a lower bound on $J^\star$.

3.4.1. Box constraints

As a more specific example, consider the case where $\mathcal{U}$ is a box, $\{v \in \mathbf{R}^m \mid \|v\|_\infty \le U^{\max}\}$. Since we have $v_i^2 \le 1$, we have the quadratic inequalities

$$v^T(e_i e_i^T)v - (U^{\max})^2 \le 0, \quad i = 1, \dots, m,$$

for $v \in \mathcal{U}$, where e_i is the ith unit vector. If we use only these inequalities to define $\tilde{\mathcal{U}}$, we have $M = m$. The objective in the problem (15) becomes

$$\mathbf{Tr}(PW) - (U^{\max})^2 \mathbf{1}^T \lambda,$$

where $\mathbf{1}$ is the vector with all entries one, and the last inequality in the problem becomes

$$R - R_0 \preceq \mathbf{diag}(\lambda).$$

3.5. Suboptimal control policies

When we solve the problem (13), or a restriction of it obtained by replacing the condition (10) with some stronger set of tractable inequalities, we obtain $P_{\rm lb}$ and $R_{\rm lb}$ (where 'lb' stands for 'lower bound'). Very roughly speaking, we can interpret our method as finding an unconstrained quadratic problem that approximates our original problem. This suggests that

$$V_{\rm lb}(z) = z^T P z,$$

and

$$K_{\rm lb} = -(R_{\rm lb} + B^T P_{\rm lb} B)^{-1} B^T P_{\rm lb} A z$$

would be good candidates for use in synthesizing suboptimal control policies, as described in Section 1.1. Examples show that this is the case.

4. Numerical examples

We will illustrate our bound, and compare it with the performance of several suboptimal control policies, on three problem instances. (We have used the method on many other problem instances, with similar results. All the data and code required to replicate the results reported here are available online.)

The first instance is a small problem, with $n = 8$ states, $m = 2$ inputs. The data are generated randomly: the entries of the matrices A and B are drawn from a standard normal distribution, after which A is scaled so that its spectral radius is one, *i.e.*, so the open-loop system is marginally stable. The stage costs are quadratic with $R_0 = I$ and $Q_0 = I$. The process disturbance $w(t)$ has distribution $\mathcal{N}(0, 0.25I)$. The input constraint set is finite: $\mathcal{U} = \{-0.2, 0, 0.2\}^2$. At each time step, each of the two inputs can only have three possible values, so we refer to a controller for this system as trilevel.

The second instance is generated in the same way, but is larger, with $n = 30$ states, $m = 10$ inputs. The stage costs are quadratic, with $R_0 = I$ and $Q_0 = I$, and $w(t) \sim \mathcal{N}(0, 0.25I)$. For this example, the input constraint set $\mathcal{U}$ is a box with $U^{\max} = 0.1$. For the S-procedure, we use $R_i = e_i e_i^T$, where e_i is the ith unit vector, and $s_i = -(U^{\max})^2$, so the inequality $v^T R_i v - s_i \le 0$ is equivalent to $|v_i| \le U^{\max}$.

Our third problem instance is a discretized mechanical control system, consisting of 6 masses, connected by springs, with three input forces that can be applied between pairs of the masses. This is the same example as described in [25]. For this problem, we have $n = 12$, $m = 3$, quadratic stage costs with $Q_0 = I$, $R_0 = I$; each entry of $w(t)$ is uniformly distributed on the interval $[-0.5, 0.5]$. The input constraint set is a box with $U^{\max} = 0.1$. We use the same S-procedure relaxation as in the second example.

For each example, we evaluate the performance for several suboptimal controllers (via Monte Carlo simulation), as well as several lower bounds. The suboptimal controllers are as follows.

- Projected linear state feedback (PLSF) with $K_{\rm lq}^\star$ and with $K_{\rm lb}$, respectively.

Table 1
Performance of suboptimal control policies (top half) and lower bounds (bottom half) for three examples.

	Small trilevel	Large random	Masses
PLSF, $K^\star_{\rm lq}$	12.91	31.27	269.84
PLSF, $K_{\rm lb}$	11.15	27.13	70.91
CLF, $P^\star_{\rm lq}$	11.12	27.23	68.01
CLF, $P_{\rm lb}$	10.76	25.60	61.06
MPC	10.87	25.70	58.87
$J^{\rm lb}$	9.12	23.83	43.15
LQR	7.47	16.81	4.66
Prescient	6.67	17.30	11.78

- Control-Lyapunov function (CLF) with $V_{\rm clf}(z) = z^{\rm T}P^\star_{\rm lq}z$ and $V_{\rm clf}(z) = z^{\rm T}P_{\rm lb}z$, respectively.
- Model predictive control (MPC) with horizon $T = 30$.

We do not carry out exact MPC for the trilevel example, since this requires solving a mixed-integer quadratic program at each step. Instead we solve the convex relaxation, with $u(t) \in [-0.2, 0.2]$, and round the value obtained to $\{-0.2, 0, 0.2\}$.

These suboptimal performance values are compared with several lower bounds, described below.

- The prescient bound, which is the value of J obtained when the control $u(t)$ is computed with knowledge of all (past and future) disturbances. This is found by Monte Carlo: in each step, we generate a realization of w, and solve the resulting quadratic program over a long horizon. For the trilevel example, we do not solve prescient optimal control problem exactly, since it is a mixed-integer quadratic program. Instead we solve the relaxation, which gives a lower bound.
- The LQR cost, which gives a simple lower bound.
- The lower bound $J^{\rm lb}$ found by our method. For the first example, with finite $\mathcal{U}$, we use the optimal value of (14) as our lower bound; for the second and third examples, with box $\mathcal{U}$, we use the optimal value of (15) as our lower bound.

The results are shown in Table 1. We can see that, in general, projected linear state feedback with the gain matrix found as a by-product of our bound computation does better than the simple LQR gain matrix. Control-Lyapunov control policies do better still, again better with the control-Laypunov function that arises as a by-product of our bound computation. Finally, MPC does very well.

On the lower bound side, we can see that $J^{\rm lb}$ is often much better than the prescient bound, and the simple bound from LQR. In two of the examples, our lower bound is quite close to the performance achieved by the best suboptimal control policies. For the masses example, the gap is larger, but it is still interesting and useful to know that the MPC control policy is no more than $100(58.87 - 43.15)/43.15 = 36\%$ suboptimal. This is not at all obvious.

5. Conclusions

We have shown how to effectively compute performance bounds for linear stochastic control problems using convex optimization. Our bounds are often close to the performance attained by suboptimal controllers based on control-Lyapunov functions and MPC. In these cases, our method shows that these controllers are close to optimal.

Acknowledgement

Yang Wang is supported by a Rambus Corporation Stanford Graduate Fellowship.

References

[1] M. Ait-Rami, L. El Ghaoui, Solving non-standard Riccati equations using LMI optimization, in: Proceedings of the 34th Conference on Decision and Control, 1995.
[2] K.J. Åström, Introduction to Stochastic Control Theory, Academic Press, 1970.
[3] F.A. Badawi, On a quadratic matrix inequality and the corresponding algebraic Riccati equation, International Journal of Control 36 (1982) 313–322.
[4] A. Bemporad, M. Morari, V. Dua, E.N. Pistikopoulos, The explicit linear quadratic regulator for constrained systems, Automatica 38 (1) (2002) 3–20.
[5] D.P. Bertsekas, Dynamic Programming and Optimal Control, vol. 1, Athena Scientific, 2005.
[6] D.P. Bertsekas, Dynamic Programming and Optimal Control, vol. 2, Athena Scientific, 2007.
[7] S. Bittanti, A.J. Laub, J.C. Willems, The Riccati Equation, Springer-Verlag, 1991.
[8] S. Boyd, L. El Ghaoui, E. Feron, V. Balakrishnan, Linear Matrix Inequalities in Systems and Control Theory, SIAM Books, Philadelphia, 1994.
[9] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University Press, 2004.
[10] E.F. Camacho, C. Bordons, Model Predictive Control, Springer-Verlag, 2004.
[11] M. Corless, G. Leitmann, Controller design for uncertain system via Lyapunov functions, in: Proceedings of the American Control Conference, vol. 3, 1988.
[12] R.A. Freeman, J.A. Primbs, Control Lyapunov functions, new ideas from an old source, in: Proceedings of the 35th IEEE Conference on Decision and Control, vol. 4, 1996.
[13] O. Hernàndez-Lerma, J.B. Lassere, Discrete-Time Markov Control Processes: Basic Optimality Criteria, Springer, 1997.
[14] R.E. Kalman, When is a linear control system optimal? Journal of Basic Engineering 86 (1) (1964) 1–10.
[15] W.H. Kwon, S. Han, Receding Horizon Control, Springer-Verlag, 2005.
[16] J.M. Maciejowski, Predictive Control with Constraints, Prentice-Hall, 2002.
[17] J. Nocedal, S.J. Wright, Numerical Optimization, Springer, 1999.
[18] F.A. Potra, S.J. Wright, Interior-point methods, Journal of Computational and Applied Mathematics 124 (1–2) (2000) 281–302.
[19] S.J. Qin, T.A. Badgwell, A survey of industrial model predictive control technology, Control Engineering Practice 11 (7) (2003) 733–764.
[20] M.A. Rami, X.Y. Zhou, Linear matrix inequalities, Riccati equations and indefinite stochastic linear quadratic controls, IEEE Transactions on Automatic Control 45 (6) (2000) 1131–1143.
[21] R. Soeterboek, Predictive Control: A Unified Approach, Prentice Hall, 1992.
[22] E.D. Sontag, A Lyapunov-like characterization of asymptotic controllability, SIAM Journal on Control and Optimization 21 (3) (1983) 462–471.
[23] M. Sznaier, R. Suarez, 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) (2003) 247–259.
[24] L. Vandenberghe, S. Boyd, Semidefinite programming, SIAM Review 38 (1) (1996) 49–95.
[25] Y. Wang, S. Boyd, Fast model predictive control using online optimization, in: Proceedings of the 17th IFAC World Congress, 2008.
[26] P. Whittle, Optimization Over Time, John Wiley & Sons, Inc., New York, NY, USA, 1982.
[27] H.S. Witsenhausen, On performance bounds for uncertain systems, SIAM Journal on Control 8 (1) (1970) 55–89.
[28] S.J. Wright, Primal-Dual Interior-Point Methods, SIAM, 1997.
[29] K. Zhou, J.C. Doyle, K. Glover, Robust and Optimal Control, Prentice Hall, 1995.