Technical report number 2001-04

Model Predictive Control with Active Steady-State Input Constraints: Existence and Computation

Gabriele Pannocchia*
Department of Chemical Engineering
University of Pisa
Via Diotisalvi, 2
56126 Pisa (ITALY)

Stephen J. Wright†
Computer Sciences Department, University of Wisconsin,
1210 W. Dayton Street, Madison, WI 53706

James B. Rawlings‡
Department of Chemical Engineering
University of Wisconsin-Madison
1415 Engineering Drive
Madison, WI 53706

Revised on February 6, 2003

*email: pannocchia@ing.unipi.it
†email: swright@cs.wisc.edu.
‡email: jbraw@che.wisc.edu
Author to whom all correspondence should be addressed.

Abstract

This paper addresses the existence and implementation of the infinite-horizon controller for the case of active steady-state input constraints. This case is important because, in many practical applications, controllers are required to operate at the boundary of the feasible region (for instance, in order to maximize global economic objectives). For this case, the usual finite horizon parameterizations with terminal cost cannot be applied, and optimal solutions are not generally available.

We propose here an iterative algorithm that generates two finite-horizon approximations to the true infinite-horizon problem, where the solution to one of the approximations yields an upper bound on the true optimum, while the other approximation yields a lower bound. We show convergence of both bounding approximations to the optimal solution, as the horizon length in the approximations is increased. We outline a procedure, based on this result, to provide a solution to the infinite-horizon problem that is exact to within any user-specified tolerance. Two examples with comparison between optimal and suboptimal controllers are presented.

Keywords

Model Predictive Control, Steady-state constraints, Optimal Control

1 Introduction

Model Predictive Control (MPC) is a technique in which a process model is used to forecast future process behavior, and the sequence of future control inputs is computed as the solution to an open-loop optimization problem. The first element of the optimal input sequence is used as the process input, and the remaining elements of the input sequence are discarded. The optimization procedure is repeated at each sampling time. Feedback from measurements is considered by correcting the model prediction, based on the error between the measurement and prediction. Many methods are available for this correction. Several recent reviews [10, 8, 6] summarize the theoretical formulations and industrial implementations of MPC.

In this paper, we consider the infinite horizon formulation of model predictive control (IHMPC) and address the case of constraints that are active at steady state. Process constraints arise both from physical limitations (for example, a valve can be at maximum fully open and at minimum totally closed) and from safety and performance specifications. Most papers on constrained infinite horizon MPC rely on the assumption that the origin is in the interior of the feasible region [5, 1, 15]. It is frequently the case, however, that in order to maximize performance objectives, the MPC controller operates at the boundary of the feasible region with respect to both input and output constraints. Moreover, when a nonzero disturbance enters the process, it is often the case that one or more manipulated inputs ride at their corresponding saturation values during a period of steady-state operation. These cases give rise to problem formulations in which the origin lies on the boundary of the feasible region. This situation was treated in [12], where a suboptimal solution for this problem is given. The main contribution of this paper is to provide an algorithm for finding upper and lower bounds on the optimum of the constrained infinite horizon optimization problem, and a procedure for identifying the optimal solution to any given level of accuracy.

Several authors have addressed the solution of convex optimization problems on an infinite dimensional space via finite-dimensional approximations; some, like us, specialize their analysis to quadratic programs from control. A lower-bound approximation is derived in [17] in a general setting. Convergence of finite approximation schemes for problems related to control are discussed in [13, 14], but these works make key assumptions on the properties of the functions and solutions that are not satisfied in our case, and the second presents a stopping rule that is difficult to implement.

The paper is organized as follows. In Section 2, we recall the formulation of the infinite horizon controller, presenting the common finite parameterization with terminal constraint and discussing feasibility limitations. In Section 3, we discuss the proposed algorithm and present results of existence and convergence to the optimal solution. Some implementation issues are addressed in Section 4 and we present applications of this method in Section 5. Finally, in Section 6, we summarize the main results of this work. Proofs of the existence and converge results are given in Appendix A.

2 Formulation of the problem

2.1 Infinite Horizon Model Predictive Control

In this paper we consider linear, time-invariant, discrete systems described by

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + B_d d_k, \\ d_{k+1} &= d_k, \\ y_k &= Cx_k + C_d d_k \ . \end{aligned} \tag{2.1}$$

in which $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the input, $y \in \mathbb{R}^p$ is the output, $d \in \mathbb{R}^p$ is the integrated disturbance added for offset-free purpose, and A, B, B_d, C, C_d are matrices of appropriate dimensions. It is assumed that the pair (A, B) is stabilizable, the pair (C, A) is detectable, and (B_d, C_d) are such that the augmented system (2.1) is detectable. We assume that, at each sampling time k, a state estimator designed for (2.1) (*e.g.* Kalman filter) provides estimate of the state $\hat{x}_{k|k}$ and of the disturbance $\hat{d}_{k|k}$.

Given the current disturbance estimate $\hat{d}_{k|k}$, we solve the following target calculation problem [12]:

$$\min_{x_{s,k}, u_{s,k}, y_{s,k}, \eta} \frac{1}{2} \left\{ \eta^T Q_s \eta + (u_{s,k} - \bar{u})^T R_s (u_{s,k} - \bar{u}) \right\} + q_s^T \eta, \tag{2.2}$$

subject to:

$$
\begin{gathered}
x_{s,k} = Ax_{s,k} + Bu_{s,k} + B_d\hat{d}_{k|k}, \\
y_{s,k} = Cx_{s,k} + C_d\hat{d}_{k|k}, \\
\bar{y} - \eta \leq y_{s,k} \leq \bar{y} + \eta, \qquad \eta \geq 0, \\
u_{\min} \leq u_{s,k} \leq u_{\max}, \qquad y_{\min} \leq y_{s,k} \leq y_{\max},
\end{gathered}
$$

in which Q_s and R_s are positive definite matrices and $\bar{y}$ and $\bar{u}$ are the setpoints for the output and input, respectively. The target optimization handles the case in which the setpoints $\bar{y}$ and $\bar{u}$ may come from a plant-wide optimization using a model and disturbance different from those in (2.2). The target problem provides a steady state for the model in (2.2) close to the setpoint, which serves as the origin for the deviation variables used subsequently in (2.4). We assume that the strict inequalities $y_{\min} < y_{\max}$, $u_{\min} < u_{\max}$ are satisfied. For appropriate choices of q_s, the linear penalty $q_s^T\eta$ guarantees that the output slack variable η is zero whenever it is possible to use this value without violating feasibility; that is, whenever the target value $\bar{y}$ is a feasible choice for $y_{s,k}$.

Next, we pose the following infinite horizon optimization problem to compute the control action:

$$
\min_{\{x_{k+j},u_{k+j},y_{k+j}\}_{j=0}^{\infty}} \frac{1}{2}\sum_{j=0}^{\infty}(y_{k+j} - y_{s,k})^T Q(y_{k+j} - y_{s,k}) + (u_{k+j} - u_{s,k})^T R(u_{k+j} - u_{s,k}), \tag{2.3a}
$$

subject to:

$$
\hat{x}_{k+j+1|k} = A\hat{x}_{k+j|k} + Bu_{k+j} + B_d\hat{d}_{k|k}, \quad j = 0, 1, 2, \ldots, \tag{2.3b}
$$

$$
y_{k+j} = C\hat{x}_{k+j|k} + C_d\hat{d}_{k|k}, \quad j = 0, 1, 2, \ldots, \tag{2.3c}
$$

$$
u_{\min} \leq u_{k+j} \leq u_{\max}, \qquad j = 0, 1, 2 \ldots, \tag{2.3d}
$$

$$
y_{\min} \leq y_{k+j} \leq y_{\max}, \qquad j = 0, 1, 2, \ldots \ . \tag{2.3e}
$$

We assume that Q and R are positive definite matrices of appropriate dimension.

This optimization problem can be rewritten in a more convenient way by considering the following new deviation variables.

$$
\begin{gathered}
w_j = \hat{x}_{k+j|k} - x_{s,k} \qquad v_j = u_{k+j} - u_{s,k}, \qquad Q \leftarrow C^T QC, \\
D = \begin{bmatrix} I \\ -I \end{bmatrix}, \quad d = \begin{bmatrix} u_{\max} - u_{s,k} \\ -u_{\min} + u_{s,k} \end{bmatrix}, \qquad E = \begin{bmatrix} C \\ -C \end{bmatrix}, \quad e = \begin{bmatrix} y_{\max} - y_{s,k} \\ -y_{\min} + y_{s,k} \end{bmatrix} .
\end{gathered}
$$

Notice that from (2.2), we have that $d \geq 0$ and $e \geq 0$. Thus (2.3) becomes:

$$
\mathcal{O}(N): \qquad \min_{\{w_j,v_j\}_{j=0}^{\infty}} \frac{1}{2}\sum_{j=0}^{\infty} w_j^T Q w_j + v_j^T R v_j, \tag{2.4a}
$$

subject to:

$$w_0 = w^{\text{init}}, \qquad w_{j+1} = Aw_j + Bv_j, \qquad j = 0, 1, 2, \ldots, \tag{2.4b}$$

$$Dv_j \leq d, \qquad j = 0, 1, 2, \ldots, \tag{2.4c}$$

$$Ew_j \leq e, \qquad j = 0, 1, 2, \ldots. \tag{2.4d}$$

Problem (2.4) may be infeasible due to the presence of input and state constraints. For example, a disturbance may enter the plant and cause the current state w^{init} to leave the set of admissible initial conditions. The problem of transient infeasibility for MPC caused by inconsistent state constraints has been addressed in several ways [3, 18, 16]. Any of these state constraint softening approaches can be incorporated into the methodology proposed here. Therefore, we can assume that for the given w^{init}, a sequence $\{v_j, w_j\}_{j=0}^{\infty}$ exists that is feasible with respect to constraints (2.4b), (2.4c), (2.4d) and gives a finite value of the objective function in (2.4). This assumption is commonly referred to as *constrained stabilizability.*

It is important to note that the pair $(Q^{1/2}, A)$ is detectable, because (C, A) is detectable and the "original" Q in (2.3) is positive definite. Because $(Q^{1/2}, A)$ is detectable, unstable modes cannot evolve without affecting the objective function. This condition and the fact that R is definite imply that for a feasible sequence $\{(w_j, v_j)\}_{j=0}^{\infty}$ in (2.4), we have $w_j, v_j \to 0$ as $j \to \infty$ [15].

2.2 Finite parameterization of the optimal control problem

When all components of d and e are strictly positive, the origin is in the interior of the feasible region, and algorithms for solving (2.4) are available [1, 15]. The key step in these analyses is to recognize that inequality constraints remain active only for a finite number of sampling times while the states and the inputs are approaching the origin.

The solution of the unconstrained infinite horizon problem is the well-known linear feedback control law:

$$v_j = Kw_j, \tag{2.5}$$

in which K is computed from the solution of the discrete-algebraic Riccati equation. For nonlinear systems, Michalska and Mayne [7] present the "dual-mode" controller in which the optimal linear control law in (2.5) is appended to the input sequence after a finite horizon. The same idea is used for linear systems [1, 15], where the following finite horizon objective function is used as replacement of (2.4):

$$\min_{\{w_j\}_{j=0}^{N}, \{v_j\}_{j=0}^{N-1}} \frac{1}{2} \sum_{j=0}^{N-1} \left\{ w_j^T Q w_j + v_j^T R v_j \right\} + \frac{1}{2} w_N^T \Pi w_N, \tag{2.6a}$$

subject to:

$$w_0 = w^{\text{init}}, \qquad w_{j+1} = Aw_j + Bv_j, \qquad j = 0, 1, 2, \ldots, N-1, \tag{2.6b}$$

$$Dv_j \leq d, \qquad j = 0, 1, 2, \ldots N-1, \tag{2.6c}$$

$$Ew_j \leq e, \qquad j = 0, 1, 2, \ldots N, \tag{2.6d}$$

in which the cost-to-go Π is the solution of the discrete-algebraic Riccati equation. In addition to constraints (2.6c), and (2.6d), the final state w_N is required to be in the following positive invariant convex set [4]:

$$\mathbb{O} = \left\{ w | H(A+BK)^i w \leq h \qquad \forall i \geq 0 \right\}, \tag{2.7}$$

in which

$$H = \begin{bmatrix} DK \\ E \end{bmatrix}; \qquad h = \begin{bmatrix} d \\ e \end{bmatrix} .$$

If the final state w_N is in the invariant set defined by (2.7), the optimal unconstrained control law (2.5) yields a solution that satisfies state and input constraints at all future sampling times. When the origin is in the relative interior of the feasible region, which is true if and only if d and e are strictly positive, the existence of a nontrivial invariant set is guaranteed.

When the origin lies inside the feasible region for (2.4), we may construct a solution for the infinite-horizon problem (2.4) from the solution of the finite-horizon problem (2.6) provided that the horizon index N is sufficiently large. Typically, one solves (2.6) for some N and then checks to see whether w_N lies in the output admissible set $\mathbb{O}$. If so, it can be shown that the optimal values of w_j, $j = 0, 1, 2, \ldots, N$ and v_j, $j = 0, 1, 2, \ldots, N-1$ are identical for (2.4) and (2.6). Otherwise, one increases the value of N in (2.6) and repeats the process.

When state constraints are active at steady state, arbitrarily small constant disturbances would render the hard constrained problem infeasible, which means that there is no feasible sequence that brings the system to the origin without persistently violating the active state constraints. This condition should register a process exception and possibly shut down the process. We assume, therefore, that state constraints are not active at steady state; that is, $e > 0$. Finally, we consider the possibility that input constraints may be active at steady state; that is, some components of d are zero (equivalently, some components of $u_{s,k}$ equal their lower or upper bound). This case is often encountered in practice, and is the one we consider in the remainder of this paper.

3 Optimal solution of the infinite horizon problem

In this section, we present a method for finding the solution of problem (2.4) to arbitrary accuracy. Our approach is to construct two finite-horizon problems that approximate (2.4) and for which solutions can be calculated. One of these approximations has an optimal objective value that is an upper bound for the optimal objective of (2.4), while the other approximation yields a lower bound. We show that the two bounds approach each other as the horizon length N approaches infinity, and use the difference between the bounds to estimate the difference between the solution of the approximating problems and the solution of (2.4). An added benefit of our analysis is that it proves existence of the optimal solution to problem (2.4) whenever a feasible sequence exists.

3.1 Upper bound on the optimal solution

An upper bound on the optimal objective Φ^* of (2.4) can be computed by using the method proposed in [12]. In this approach, a suboptimal solution to (2.4) is found by restricting the evolution of the input and state trajectories to the null space of the active steady-state constraints, after the finite horizon $N > 0$. This solution is found by minimizing the following infinite horizon objective function:

$$\mathcal{U}(N): \quad \min_{\{w_j, v_j\}_{j=0}^{\infty}} \frac{1}{2} \sum_{j=0}^{\infty} w_j^T Q w_j + v_j^T R v_j, \tag{3.1a}$$

$$\text{subject to: (2.4b), (2.4c), (2.4d)} \tag{3.1b}$$

and

$$\bar{D} v_j = 0, \qquad j = N, N+1, \ldots, \tag{3.1c}$$

in which $\bar{D}$ denotes the row sub-matrix of D corresponding to input inequality constraints active at steady state, that is the rows of D for which the corresponding elements of d are zero. Let $\overline{\Phi}_N$ be the optimal objective value for $\mathcal{U}(N)$ in (3.1). Since (3.1) has more constraints than (2.4), its feasible region is smaller, so we have:

$$\Phi^* \leq \overline{\Phi}_N \; . \tag{3.2}$$

We can reformulate the infinite horizon problem (3.1) as a finite-horizon problem that can be solved by practical means as follows:

$$\min_{\{w_j\}_{j=0}^{N}, \{v_j\}_{j=0}^{N-1}} \frac{1}{2} \sum_{j=0}^{N-1} \left\{ w_j^T Q w_j + v_j^T R v_j \right\} + \frac{1}{2} w_N^T \bar{\Pi} w_N, \tag{3.3a}$$

$$\text{subject to: (2.6b), (2.6c), (2.6d)}\ , \tag{3.3b}$$

in which the cost-to-go matrix $\bar{\Pi}$ is associated with the unconstrained control law:

$$v_j = \bar{K} w_j \; . \tag{3.4}$$

The off-line computation of $\bar{K}$ is described in [12]; we outline the procedure here. Let $\mathcal{N}_{\bar{D}}$ be an orthonormal basis for the null space of $\bar{D}$, so that vectors v_j that satisfy (3.1c) have the form $\mathcal{N}_{\bar{D}} p_j$, $j = N, N+1, \ldots$ for arbitrary p_j. We then solve the optimal unconstrained LQR problem for the system with characteristic matrices $(A, B\mathcal{N}_{\bar{D}})$ and with $(Q, \mathcal{N}_{\bar{D}}^T R \mathcal{N}_{\bar{D}})$ as state and input penalty, respectively, to obtain optimal gain and cost-to-go matrices $K_{\bar{D}}$ and $\bar{\Pi}$, respectively. It follows that $\bar{K} = \mathcal{N}_{\bar{D}} K_{\bar{D}}$. In order for such a linear control law to exist, the pair $(A, B\mathcal{N}_{\bar{D}})$ must be stabilizable. When this condition is not satisfied, we require the controller to zero, at the end of the finite horizon N, the unstable modes that are not controllable in this subspace. To this aim we have the following existence result.

Theorem 3.1. *If the optimization problem* (2.4) *is feasible, there exists a finite integer* N' *such that* (3.1) *is feasible for any* $N \geq N'$.

Proof. See Appendix A.1. □

The problems (3.1) and (3.3) are identical in the sense that the solution components $v_0, v_1, \ldots, v_{N-1}$ are the same for each. We compute the remaining components by using the unconstrained evolution of the system under the feedback gain $\bar{K}$:

$$v_j = \bar{K} w_j \qquad j = N, N+1, \ldots, \tag{3.5a}$$

$$w_{j+1} = A w_j + B v_j \qquad j = N, N+1, \ldots\ . \tag{3.5b}$$

If the final state w_N does not lie in the output admissible set for the subset of inequalities not active at steady state under the feedback gain $\bar{K}$ ([12], [4]), the horizon N must be increased in order for the solution components of (3.1) and (3.3) to be equal. That is, we need to ensure that the sequence $\{v_j, w_j\}_{j=N}^{\infty}$ generated by (2.4b) together with (3.4) satisfies the inactive constraints from (2.4c), (2.4d) at all subsequent stages $j = N, N+1, \ldots$.

3.2 Lower bound on the optimal solution

A lower bound on the optimal objective Φ^* of (2.4) can be found by minimizing the following infinite horizon objective function:

$$\mathcal{L}(N): \quad \min_{\{w_j, v_j\}_{j=0}^{\infty}} \frac{1}{2} \sum_{j=0}^{\infty} w_j^T Q w_j + v_j^T R v_j, \tag{3.6a}$$

$$\text{subject to: (2.4b), (2.6c), (2.6d).} \tag{3.6b}$$

Notice that constraints (2.6c), (2.6d) are enforced over a finite horizon N only. Therefore, if the optimal problem (2.4) (equivalently, the upper bounding problem (3.1)) is feasible, we have that (3.6) also is feasible.

Let $\underline{\Phi}_N$ be the optimal objective value for $\mathcal{L}(N)$. Since (3.6) has fewer constraints than (2.4), it is clear that

$$\underline{\Phi}_N \leq \Phi^*, \qquad \forall N > 0\ . \tag{3.7}$$

The infinite horizon problem (3.6) can be solved by using a finite parameterization, since after stage N the optimal control law is $v_j = K w_j$ in which K is the well-known unconstrained LQR gain computed from the corresponding Riccati equation. Thus, we solve the following problem:

$$\min_{\{w_j\}_{j=0}^{N}, \{v_j\}_{j=0}^{N-1}} \frac{1}{2} \sum_{j=0}^{N-1} \left\{ w_j^T Q w_j + v_j^T R v_j \right\} + \frac{1}{2} w_N^T \Pi w_N, \tag{3.8a}$$

$$\text{subject to: (2.6b), (2.6c), (2.6d),} \tag{3.8b}$$

in which Π is the steady-state solution of the Riccati equation. Note that since the constraints (2.6c), (2.6d) are imposed only over a finite horizon, the solution of (3.8) may violate input and state constraints (2.4c) and (2.4d) at some stages $j > N$.

The solution components $v_0, v_1, \ldots, v_{N-1}$ are the same for (3.6) and (3.8). We obtain $v_N, v_{N+1}, \ldots,$ from (3.8) by using the unconstrained evolution:

$$v_j = Kw_j \qquad j = N, N+1, \ldots, \tag{3.9a}$$

$$w_{j+1} = Aw_j + Bv_j \qquad j = N, N+1, \ldots . \tag{3.9b}$$

3.3 Convergence of the Optimal Sequences

The following results show the convergence properties of the lower and upper bound problems. See Appendix A.2 for the proofs of the following theorems.

Theorem 3.2. *Let z^* and $\bar{z}_N$ be the optimal infinite dimensional input sequences solution to the optimal problem* (2.4) *and to the upper bounding problem* (3.1)*, respectively. We have*

$$\lim_{N\to\infty} \bar{z}_N = z^*, \qquad \overline{\Phi}_N \downarrow \Phi^* . \tag{3.10}$$

where the last limit indicates that $\{\overline{\Phi}_N\}$ is nonincresing and converges to Φ^*.

Theorem 3.3. *Let z^* and $\underline{z}_N$ be the optimal infinite dimensional input sequences solution to the optimal problem* (2.4) *and to the lower bounding problem* (3.6)*, respectively.*

$$\lim_{N\to\infty} \underline{z}_N = z^*, \qquad \underline{\Phi}_N \uparrow \Phi^*, \tag{3.11}$$

where the last limit indicates that $\{\underline{\Phi}_N\}$ is nondecreasing and converges to Φ^*.

Theorem 3.4. *Let z_0^* and $\bar{z}_{N,0}$ be the first input component of the solution to the optimal problem* (2.4) *and to the upper bounding problem* (3.1)*, respectively. There exists a positive scalar α such that for any N the following inequality holds:*

$$\|\bar{v}_{N,0} - v_0^*\| \leq \|\bar{z}_N - z^*\| \leq \left[\frac{2}{\alpha}(\overline{\Phi}_N - \underline{\Phi}_N)\right]^{1/2} . \tag{3.12}$$

The scalar α depends on the problem data and in particular on the matrix R, as we show at the end of the next section.

4 Implementation issues

The results of the previous section suggest an iterative approach to determining an approximation to the v_0^* component of the solution of the infinite horizon problem (2.4). In this approach, we solve a series of quadratic programs for the upper and the lower bound problems (3.1) and (3.6). If the difference between the optimal objective values for these problems does not satisfy a chosen stopping criterion, the horizon is increased; otherwise the first input $\bar{v}_{N,0}$ of the computed sequence of the upper bound problem (3.1) is accepted as a good approximation to v_0, and is injected into the plant.

As stopping criterion we use a relative difference between the upper and the lower bound solution:

$$\frac{\overline{\Phi}_N - \underline{\Phi}_N}{1 + \underline{\Phi}_N} \leq \rho, \tag{4.1}$$

where ρ is a small positive number.

At each sampling time we apply the following algorithm, starting with a positive horizon $N > 0$.

Algorithm 4.1. *Start with a positive horizon $N > 0$.*

1. *Solve* (3.3)*. If the problem is infeasible, go to 5. Otherwise, let $\overline{\Phi}_N$ be the optimal value of its objective function.*
2. *If the final state w_N does not belong to the output admissible set for constraints inactive at steady state [12], go to 5.*
3. *Solve* (3.8)*. Let $\underline{\Phi}_N$ be the optimal value of its objective function.*
4. *Check* (4.1)*. If satisfied, go to 6.*
5. *Increase the horizon N and go to 1.*
6. *Set v_0 equal to the first solution component of* (3.3)*.*

The proposed algorithm always terminates because from (3.11) and (3.10) we have that

$$\lim_{N\to\infty} \overline{\Phi}_N - \underline{\Phi}_N = 0, \tag{4.2}$$

which implies that for any $\rho > 0$ there exists a N' such that for $N > N'$ the stopping criterion (4.1) is satisfied.

In (3.12) the monotonicity constant α appears in the denominator, so the bound is tight when α is large. It is straightforward to show that $\alpha \geq \lambda_{\min}(R) > 0$, in which we used the fact that R is positive definite. A better bound on α can be obtained numerically by computing the smallest eigenvalue of the finite dimensional Hessian $U(N)$, and then letting the horizon increase until we reach convergence on α.

5 Case studies

In this section we present two examples of systems with constraints active at steady state. The first example concerns a system in which the control action does not become permanently active or inactive at steady state. The second example is a heavy oil fractionator of the Shell Control Problem, which has input and output constraints.

5.1 Example # 1

The following system is considered:

$$x_{k+1} = \begin{bmatrix} 0.5477 & 0.8208 & 0 \\ -0.8208 & 0.5067 & 0 \\ 0 & 0 & 0.8 \end{bmatrix} x_k + \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} u_k, \tag{5.1}$$

$$y_k = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix} x_k. \tag{5.2}$$

The controller is required to drive the system state to the origin, from the initial state $x_0 = \begin{bmatrix} 3 & 3 & 0 \end{bmatrix}^T$. The control action is bounded as follows:

$$u_k \leq 0. \tag{5.3}$$

Tuning parameters for the controller are $Q = 1$ and $R = 1$.

For this system the target values for both input and state are zero, which is also a bound for the input. Hence, the origin lies on the boundary of the feasible region. The unconstrained control law $u_k = Kx_k$ for this system requires both positive and negative inputs. As shown in Figure 2, the corresponding constrained control action does not become permanently active or inactive. A logarithmic scale is used in the figure to emphasize this behavior.

The relative tolerance between the upper bound and the lower bound solutions in (4.1) is chosen equal to $\rho_1 = 10^{-6}$. In Figure 1, we plot the horizon length against the relative difference between the upper bound and the lower bound objective functions, obtained at time $k = 0$. This plot shows that a horizon of about 70 is needed to satisfy the chosen tolerance ρ_1. This relative tolerance guarantees a precise convergence of the entire input sequence $(u_0, u_1, \ldots)$ to the optimal value. However, even with a bigger tolerance (and therefore a shorter horizon), the injected input u_0 is still close to the optimal value. In fact, in Figure 2 the injected input is reported for the chosen tolerance ρ_1 and for the larger tolerance $\rho_2 = 10^{-3}$. These two input sequences are close to each other except for the last part of the simulation, where the magnitude of the input is small in any case. The output for the closed-loop system is reported in Figure 3, both for the tolerance ρ_1 and ρ_2. The two lines are essentially indistinguishable, showing that even a weaker convergence tolerance produced a response close to the optimal one.

5.2 Example # 2

As second example we consider the heavy oil fractionator of the Shell Control Problem [9]. A linear model of the system is:

$$G(s) = \begin{bmatrix} \frac{4.05e^{-27s}}{50s+1} & \frac{1.77e^{-28s}}{60s+1} & \frac{5.88e^{-27s}}{50s+1} \\ \frac{5.39e^{-18s}}{50s+1} & \frac{5.72e^{-14s}}{60s+1} & \frac{6.90e^{-15s}}{40s+1} \\ \frac{4.38e^{-20s}}{33s+1} & \frac{4.42e^{-22s}}{44s+1} & \frac{7.20}{19s+1} \end{bmatrix}. \tag{5.4}$$

where the three inputs are the top product draw rate (u_1), the side product draw rate (u_2) and the reflux heat duty for the bottom of the column (u_3). The three outputs are the top

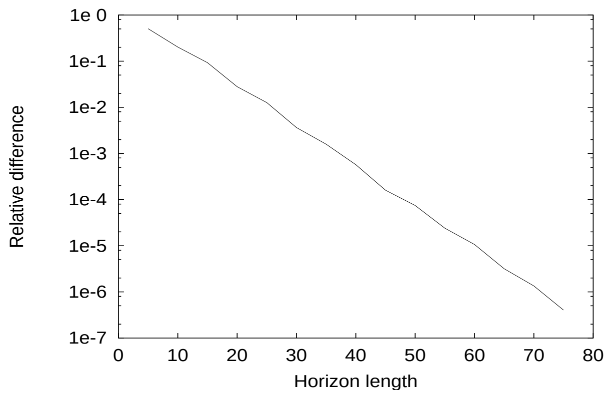

Figure 1: Horizon length *vs* relative difference of objective functions

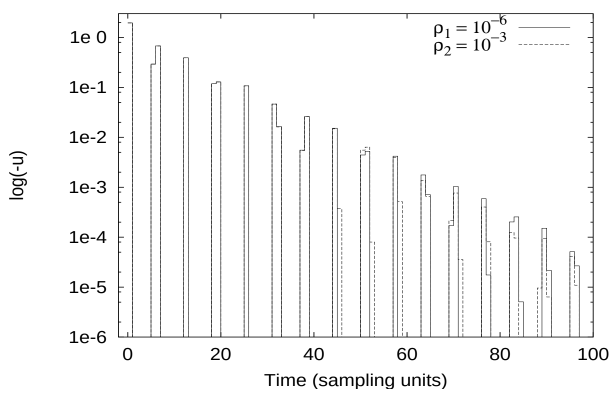

Figure 2: Closed-loop input for Example # 1

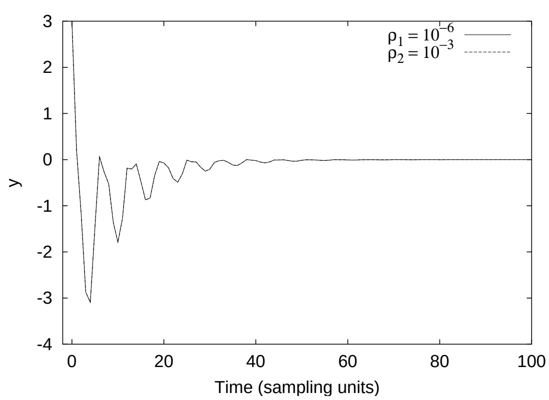

Figure 3: Closed-loop output for Example # 1

product composition (y_1), the side product composition (y_2) and the reflux temperature at the bottom of the column (y_3). The following input and output constraints are considered:

$$-0.5 \leq u_i \leq 0.5, \qquad i = 1, 2, 3, \qquad -0.5 \leq y_1 \leq 0.5, \qquad -0.5 \leq y_3 \ .$$

The controller tuning matrices are $Q = R = I$. We consider a set-point: $\bar{y} = \begin{bmatrix} 0.3 & 0.3 & -0.3 \end{bmatrix}^T$, which is not reachable due to input constraints. The target calculations returns (with $Q_s = R_s = I, q_s^T = 10^6 \cdot \mathbf{1}$) the following feasible targets for input and output:

$$u_{s,k} = \begin{bmatrix} 0.5 \\ -0.1026 \\ -0.2702 \end{bmatrix}, \qquad y_{s,k} = \begin{bmatrix} 0.2540 \\ 0.2436 \\ -0.2086 \end{bmatrix}, \qquad k = 0, 1, 2, \dots \ . \tag{5.5}$$

Because the first component of $u_{s,k}$ equals its upper bound, the origin lies on the boundary of the feasible region.

In Figure 4 we report inputs and outputs for three optimal regulators obtained with different relative tolerance ($\rho = 10^{-6}$, 10^{-3}, 10^{-1}, respectively). The regulators with relative tolerance of 10^{-6} and 10^{-3} show essentially the same closed-loop response. Also the regulator with relative tolerance of 10^{-1} guarantees a performance not too far from optimal. Clearly, the use of a larger tolerance has a direct impact on the computation time because the horizon length N required to satisfy the stopping criterion decreases. We show in Figure 5 the horizon required by each controller to meet the specified convergence tolerance, at each point in the simulation. It is interesting to notice that the required horizon is larger at the beginning of the horizon because the system state is far away from its steady state, and then it decreases.

For this particular example the sampling time was 4 minutes [9], while the average computation time required for the controller with $\rho = 10^{-6}$ has been of about 4 seconds on a 1.2 GHz Athlon computer and using a dense hessian approach, which scales cubically with the horizon length. In fact, interior-point structured solvers for the MPC problem that scale linearly with the horizon are available and can be directly applied to this algorithm.

6 Conclusions

In this paper the existence and the implementation of the infinite horizon controller for the case of active steady-state input constraints has been discussed. This case is important because, in practical applications, controllers are often required to operate at the boundary of the feasible region. Previously, only suboptimal solutions were available for this case, based on finite horizon formulations with terminal equality constraints or infinite horizon formulations with appropriate suboptimal finite parameterization. We presented here an iterative algorithm that determines the optimal solution of this problem within a user specified tolerance. Availability of a near-optimal solution makes the proposed controller simple to understand and tune, and improves its performance. Finally, when the computation time is a limiting factor we can still apply this algorithm with a larger tolerance (shorter horizon) and obtain a bound on the difference between the optimal and the computed suboptimal input.

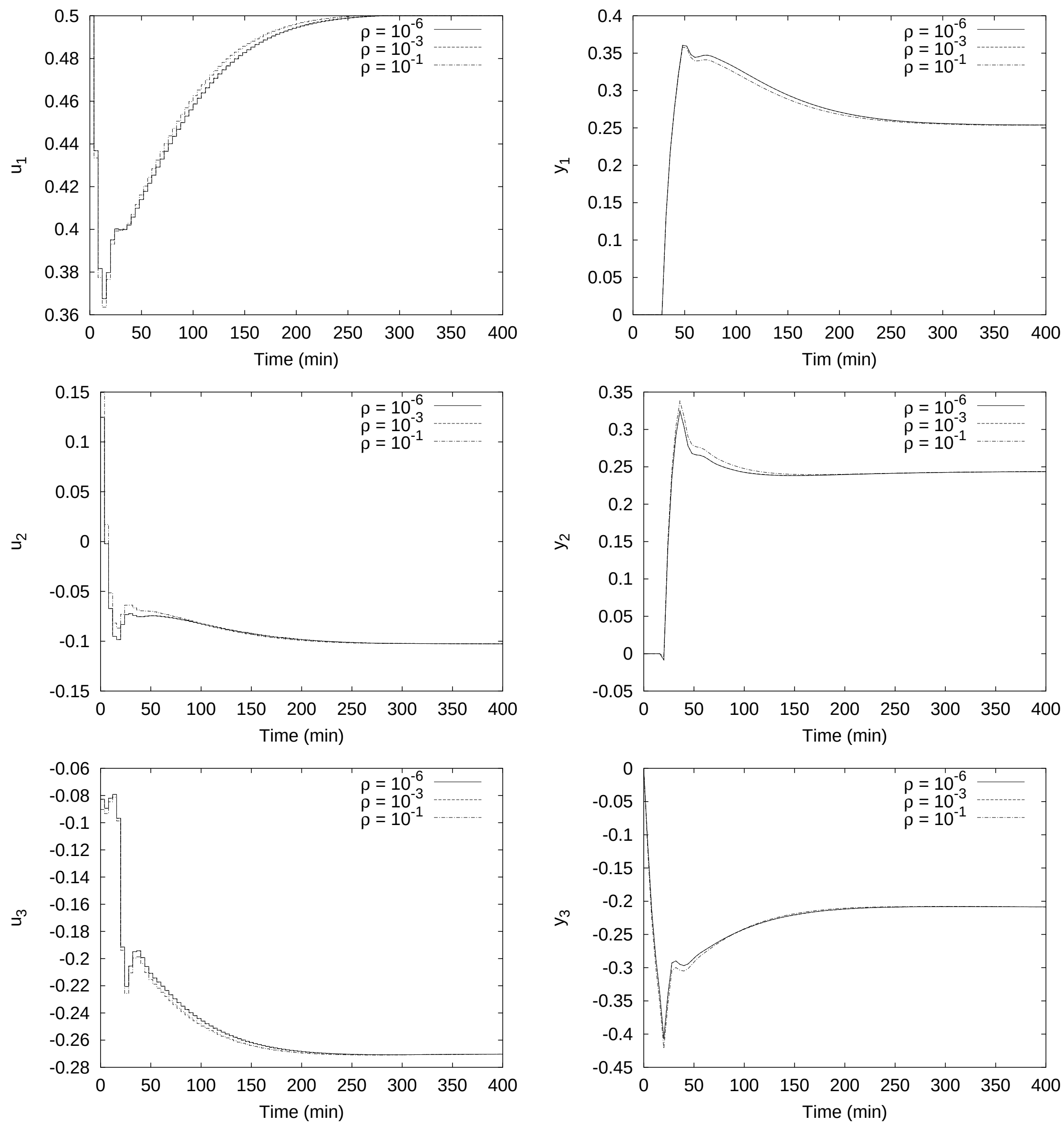

Figure 4: Closed-loop inputs (left side) and outputs (right side)

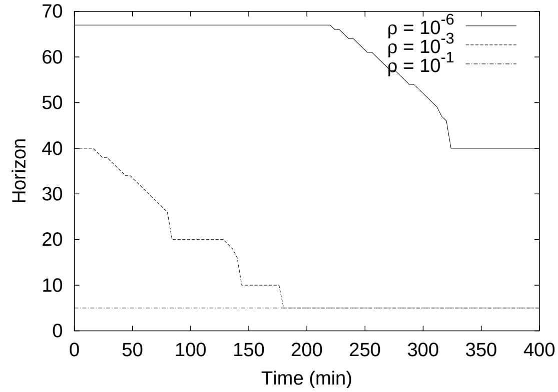

Figure 5: Closed-loop required horizon

References

[1] D. Chmielewski and V. Manousiouthakis. On constrained infinite-time linear quadratic optimal control. *Sys. Cont. Let.*, 29:121–129, 1996.

[2] F. H. Clarke. *Optimization and Nonsmooth Analysis*. John Wiley, New York, 1983.

[3] N. M. C. de Oliveira and L. T. Biegler. Constraint handling and stability properties of model-predictive control. *AIChE J.*, 40:1138–1155, 1994.

[4] E. G. Gilbert and K. T. Tan. Linear systems with state and control constraints: The theory and application of maximal output admissible sets. *IEEE Trans. Auto. Cont.*, 36:1008–1019, 1991.

[5] S. S. Keerthi and E. G. Gilbert. Optimal infinite-horizon feedback laws for a general class of constrained discrete-time-systems. stability and moving-horizon approximations. *J. Optim. Theory Appl.*, 57:265–293, 1988.

[6] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert. Constrained model predictive control: Stability and optimality. *Automatica*, 36:789–814, 2000.

[7] H. Michalska and D. Q. Mayne. Robust receding horizon control of constrained nonlinear systems. *IEEE Trans. Auto. Cont.*, 38:1623–1633, 1993.

[8] M. Morari and J. H. Lee. Model predictive control: Past, present and future. *Comput. Chem. Eng.*, 23:667–682, 1999.

[9] D. M. Prett and M. Morari. *The Shell Process Control Workshop*. Butterworth Publishers, 1986.

[10] S. J. Qin and T. A. Badgwell. An overview of industrial model predictive control technology. In Jeffrey C. Kantor, Carlos E. Garcia, and Brice Carnahan, editors, *Fifth*

International Conference on Chemical Process Control, AIChE Symposium Series 316, 93, pages 232–256. AIChE, 1997.

[11] J. R. Randolph. *Basic Real and Abstract Analysis*, chapter 3.10. Academic Press, 1968.

[12] C. V. Rao and J. B. Rawlings. Steady states and constraints in model predictive control. *AIChE J.*, 45:1266–1278, 1999.

[13] Irwin E. Schochetman and Robert L. Smith. Existence and discovery of average cost optimal solutions in deterministic infinite horizon optimization. *Math. Oper. Res.*, 23(2):416–432, 1998.

[14] Irwin E. Schochetman and Robert L. Smith. A finite algorithm for solving infinite dimensional optimization problems. *Ann. Oper. Res.*, 101:119–142, 2001.

[15] P. O. M. Scokaert and J. B. Rawlings. Constrained linear quadratic regulation. *IEEE Trans. Auto. Cont.*, 43:1163–1169, 1998.

[16] P. O. M. Scokaert and J. B. Rawlings. Feasibility issues in linear model predictive control. *AIChE J.*, 45:1649–1659, 1999.

[17] J. Semple. Infinite positive-definite quadratic programming in a Hilbert space. *J. Optim. Theory Appl.*, 88:743–749, 1996.

[18] A. Zheng and M. Morari. Stability of model predictive control with mixed constraints. *IEEE Trans. Auto. Cont.*, 40:1818–1823, 1995.

A Proofs

A.1 Proof of Theorem 3.1

We show that the existence of an infinite feasible sequence for the optimal problem (2.4) implies that the upper bounding optimization problem (3.3) is feasible for finite N. In the upper bounding optimization problem (3.3), the use of the linear control law $v_k = \bar{K} w_k$ requires the controller to zero in finite time the unstable modes that are not controllable in the null space of the active steady-state constraints. We show here the stronger result that the existence of an infinite feasible sequence for the optimal problem (2.4) implies we can zero all unstable modes in finite time.

Without loss of generality, we make two simplifying assumptions. First, consider a Schur decomposition of the system matrix partitioned into stable and unstable parts. We consider only the unstable part, *i.e.* we consider a purely unstable system, and show we can zero the entire state in finite time.

Second, we assume that the initial state w_0 is sufficiently close to the origin that all constraint boundaries that do not intersect the origin (i.e. constraints that are not active at steady state) appear arbitrarily far away and can be neglected. This assumption is valid since, for any sequence $\{w_k, v_k\}$ feasible for (2.4), we have that $\lim_{k\to\infty} w_k = 0$,

$\lim_{k\to\infty} v_k = 0$ and, therefore, the system state can be brought arbitrarily close to the origin in finite time.

Preliminary definitions

Definition A.1 (General problem). *We consider the following problem:*

$$w_0 \text{ given }, \qquad w_{k+1} = Aw_k + Bv_k \quad k = 0, 1, \ldots, \tag{A.1a}$$

$$\bar{D}v_k \leq 0 \quad k = 0, 1, \ldots, \tag{A.1b}$$

in which $A \in \mathbb{R}^{n\times n}$ *has all eigenvalues outside the open unit circle,* i.e. $|\lambda_i(A)| \geq 1$, $i = 1, \ldots, n$, $B \in \mathbb{R}^{n\times m}$ *and* $\bar{D} \in \mathbb{R}^{m_a \times m}$, *with* $m_a \leq m$, *and* $\text{rank}(\bar{D}) = m_a$. *We assume that a sequence* $\{w_k, v_k\}_{k=0}^{\infty}$ *exists such that* (A.1a) *and* (A.1b) *are satisfied and*

$$\sum_{k=0}^{\infty} w_k^T Q w_k + v_k^T R v_k < \infty \tag{A.2}$$

for Q, R *symmetric positive definite matrices.*

Without loss of generality, we assume that $\bar{D} = \begin{bmatrix} I_{m_a} & 0 \end{bmatrix}$. The transformation of the input u and, consequently, of the matrix B that leads to this form is

$$v \leftarrow Tv = \begin{bmatrix} \bar{D} \\ \bar{D}^c \end{bmatrix} v, \qquad B \leftarrow BT^{-1}, \qquad R \leftarrow T^{-T} R T^{-1}$$

in which $\bar{D}^c \in \mathbb{R}^{(m-m_a)\times m}$ is such that $T \in \mathbb{R}^{m\times m}$ is invertible.

Lemma A.1. *Given a nonsingular matrix* $A \in \mathbb{R}^{n\times n}$ *and* $B \in \mathbb{R}^{n\times n}$, (A, B) *is controllable iff* (A^{-1}, B) *is controllable.*

Proof. If (A, B) is controllable we have that:

$$\begin{aligned} n &= \text{rank} \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix} = \text{rank}\left(A^{n-1} \begin{bmatrix} A^{-n+1}B & A^{-n+2}B & \cdots & B \end{bmatrix} \right) \\ &= \text{rank} \begin{bmatrix} B & A^{-1}B & \cdots & A^{-(n-1)}B \end{bmatrix}, \end{aligned}$$

which implies that (A^{-1}, B) is controllable. Necessity is proven in an analogous way. □

Definition A.2. *We define the controllability matrix of order* k *for* (A^{-1}, B) *as*

$$\mathbb{C}_k = \begin{bmatrix} B & A^{-1}B & A^{-2}B & \cdots & A^{-(k-1)}B \end{bmatrix}. \tag{A.3}$$

We also define the infinite dimensional controllability matrix for (A^{-1}, B) *as*

$$\mathbb{C}_\infty = \lim_{k\to\infty} \mathbb{C}_k. \tag{A.4}$$

The matrix $\mathbb{C}_\infty$ has bounded elements since A has unstable eigenvalues.

Definition A.3. *We define the matrix* $\mathbb{D}_k \in \mathbb{R}^{m_a k \times mk}$ *as*

$$\mathbb{D}_k = \begin{bmatrix} \bar{D} & 0 & \cdots & 0 \\ 0 & \bar{D} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \bar{D} \end{bmatrix},$$

and the infinite dimensional matrix $\mathbb{D}_\infty$ *as:*

$$\mathbb{D}_\infty = \lim_{k \to \infty} \mathbb{D}_k.$$

Main proof

Let $\pi_\infty = (v_0, v_1, \cdots) \in \ell^2$ be an input vector that satisfies (A.1a), (A.1b) and (A.2). We can write

$$w_k - A^k w_0 = A^{k-1} B v_0 + A^{k-2} B v_1 + \cdots + B v_{k-1},$$

or alternatively

$$A^{-(k-1)} w_k - A w_0 = B v_0 + A^{-1} B v_1 + \cdots + A^{-(k-1)} B v_{k-1}.$$

As $k \to \infty$ we have that $w_k \to 0$ and, therefore, we have

$$\mathbb{C}_\infty \pi_\infty = -A w_0, \tag{A.5a}$$

$$\mathbb{D}_\infty \pi_\infty \leq 0. \tag{A.5b}$$

Theorem A.1. *Given an infinite vector* $\pi_\infty \in \ell^2$ *that satisfies* (A.5)*, there exists a finite vector* $\bar{\pi}_N$ *for some* N *such that*

$$\mathbb{C}_N \bar{\pi}_N = -A w_0, \tag{A.6a}$$

$$\mathbb{D}_N \bar{\pi}_N \leq 0. \tag{A.6b}$$

Proof. Some elements of the vector π_∞ may be zero. We can remove these elements from π_∞ and define a new vector $\tilde{\pi}_\infty$. Consequently, we remove the corresponding columns from $\mathbb{C}_\infty$ and $\mathbb{D}_\infty$ obtaining new matrices $\tilde{\mathbb{C}}_\infty$ and $\tilde{\mathbb{D}}_\infty$. If $\tilde{\pi}_\infty$ has a finite number of elements, the proof is complete because we can construct $\bar{\pi}_N$ from π_∞ by choosing N such that $v_j = 0$, $j > N$. If all elements of π_∞ are zero, the proof is also complete because $w_0 = 0$. We assume, therefore, that $\tilde{\pi}_\infty$ has an infinite number of elements. We can rewrite (A.5) as

$$\tilde{\mathbb{C}}_\infty \tilde{\pi}_\infty = -A w_0, \tag{A.7a}$$

$$\tilde{\mathbb{D}}_\infty \tilde{\pi}_\infty < 0. \tag{A.7b}$$

in which the strict inequality comes from the structure of $\mathbb{D}_\infty$ and from the fact that the zero elements of π_∞ have been removed in $\tilde{\pi}_\infty$.

Let $l \leq n$ be the rank of $\tilde{\mathbb{C}}_\infty$. We have that there exists a M such that $\text{rank}(\mathbb{C}_J) = l$ for any $J \geq M$. From (A.7a) we also have that $\text{rank}([\tilde{\mathbb{C}}_\infty | Aw_0]) = l$. Given a $N > M$, we partition the vector $\tilde{\pi}_\infty$ and the matrices $\tilde{\mathbb{C}}_\infty$ and $\tilde{\mathbb{D}}_\infty$ as follows

$$\tilde{\pi}_\infty = (\tilde{\pi}_N | \tilde{\pi}_{N|\infty}), \quad \tilde{\mathbb{C}}_\infty = [\tilde{\mathbb{C}}_N | \tilde{\mathbb{C}}_{N|\infty}], \quad \tilde{\mathbb{D}}_\infty = [\tilde{\mathbb{D}}_N | \tilde{\mathbb{D}}_{N|\infty}].$$

From (A.7a) we have

$$\tilde{\mathbb{C}}_N \tilde{\pi}_N + \tilde{\mathbb{C}}_{N|\infty} \tilde{\pi}_{N|\infty} = -Aw_0 \tag{A.8}$$

We wish to construct a finite dimensional vector π_N as $\pi_N = \tilde{\pi}_N + \rho_N$ in which

$$\rho_N = (v_0, v_1, \ldots, v_{M-1}, 0, \ldots, 0)$$

such that

$$\tilde{\mathbb{C}}_N \pi_N = -Aw_0, \tag{A.9a}$$

$$\tilde{\mathbb{D}}_N \pi_N < 0. \tag{A.9b}$$

Using (A.7a), we have that

$$\tilde{\mathbb{C}}_N (\tilde{\pi}_N + \rho_N) = -Aw_0 = \tilde{\mathbb{C}}_N \tilde{\pi}_N + \tilde{\mathbb{C}}_{N|\infty} \tilde{\pi}_{N|\infty},$$

from which

$$\tilde{\mathbb{C}}_N \rho_N = \tilde{\mathbb{C}}_{N|\infty} \tilde{\pi}_{N|\infty}. \tag{A.10}$$

Since the last $N - M$ terms of ρ_N are zero, we can rewrite (A.10) as

$$\tilde{\mathbb{C}}_M \rho_M = \tilde{\mathbb{C}}_{N|\infty} \tilde{\pi}_{N|\infty} \tag{A.11}$$

Using (A.9b) instead, we obtain

$$\tilde{\mathbb{D}}_N \rho_N < -\tilde{\mathbb{D}}_N \tilde{\pi}_N \tag{A.12}$$

Since $N > M$, we can take a sub-matrix of (A.12) and use the particular structure of $\tilde{\mathbb{D}}_N$ to obtain:

$$\tilde{\mathbb{D}}_M \rho_M < -\tilde{\mathbb{D}}_M \tilde{\pi}_M. \tag{A.13}$$

Since $N > M$, we have that $\tilde{\mathbb{C}}_{N|\infty} \tilde{\pi}_{N|\infty} \in \text{range}(\tilde{\mathbb{C}}_N) = \text{range}(\tilde{\mathbb{C}}_M)$ and, therefore, (A.11) admits solution. Let $\tilde{\mathbb{C}}_M^+$ be a left inverse of $\tilde{\mathbb{C}}_M$ (*i.e.* $\tilde{\mathbb{C}}_M^+ \tilde{\mathbb{C}}_M = I$). One solution of (A.11) is:

$$\rho_M = \tilde{\mathbb{C}}_M^+ \tilde{\mathbb{C}}_{N|\infty} \tilde{\pi}_{N|\infty}. \tag{A.14}$$

Since $\tilde{\mathbb{C}}_{N|\infty} \tilde{\pi}_{N|\infty} \to 0$ as $N \to \infty$, we have that

$$\lim_{N \to \infty} \|\rho_M\|_2 = 0 \tag{A.15}$$

Hence, since $-\tilde{\mathbb{D}}_M \tilde{\pi}_M > 0$ in (A.13), there exists N' such that (A.13) holds for all $N \geq N'$. Choosing any $N \geq N'$, we have found a vector π_N that satisfies (A.9a) and (A.9b). We can obtain the vector $\bar{\pi}_N$ that satisfies (A.6a) and (A.6b) by reinserting the zero elements that have been removed from π_∞ to obtain $\tilde{\pi}_\infty$. □

A.2 Quadratic Programs and Sequences of Sets

Convex Quadratic Programs in ℓ^2

We consider the space ℓ^2, which is the infinite-dimensional set of objects of the form:

$$z = (z_1, z_2, z_3, \ldots), \quad z_i \in \mathbb{R} \text{ for all } i = 1, 2, \ldots.$$

such that

$$\sum_{i=1}^{\infty} z_i^2 < \infty.$$

When equipped with the following inner product:

$$\langle a, b \rangle = \sum_{i=1}^{\infty} a_i b_i,$$

ℓ^2 is a separable Hilbert space [11]. We define the norm on this space in the obvious way:

$$\|z\| = \langle z, z \rangle^{1/2}.$$

Consider the following convex optimization problem over ℓ^2:

$$\min_z f(z) = \frac{1}{2}\langle z, Uz \rangle + \langle c, z \rangle, \qquad \text{subject to } z \in \mathcal{C}, \tag{A.16}$$

where

- $U : \ell^2 \to \ell^2$ is a linear, self-adjoint, strictly monotone operator; that is, there is $\alpha > 0$ such that
$$\langle z, Uz \rangle \geq \alpha \langle z, z \rangle, \qquad \text{for all } z \in \ell^2. \tag{A.17}$$
- $\mathcal{C} \subset \ell^2$ is convex, closed, and nonempty.

We define the normal cone to $\mathcal{C}$ at a point $\bar{z}$ as follows:

$$N_{\mathcal{C}}(\bar{z}) = \{v \mid \langle v, \bar{z} - z \rangle \geq 0, \text{ for all } z \in \mathcal{C}\}. \tag{A.18}$$

Theorem A.2. *If z^* solves* (A.16)*, then we have*

$$-(Uz^* + c) \in N_{\mathcal{C}}(z^*), \tag{A.19}$$

and in particular, we have

$$\langle Uz^* + c, z - z^* \rangle \geq 0, \qquad \text{for all } z \in \mathcal{C}. \tag{A.20}$$

Proof. We can apply the Corollary in [2, p.52]. At the solution z^*, we have:

$$0 \in \partial f(z^*) + N_{\mathcal{C}}(z^*), \tag{A.21}$$

where ∂f denotes the generalized gradient of f at z^*. Since f is quadratic and U has the properties described above, we have $\partial f(z^*) = \{Uz^* + c\}$, so that (A.19) follows immediately from (A.21). The second claim follows from (A.18). □

Theorem A.3. *The problem* (A.16) *has a unique minimizer* z^*.

Proof. Given any $z_0 \in \mathcal{C}$, the level set

$$\{z \,|\, f(z) \le f(z_0)\}$$

is closed and bounded, by the monotonicity property (A.17). Hence f attains a minimum on this set, say at z^*. Uniqueness of z^* is also a consequence of monotonicity. □

Theorem A.4. *Let* z^* *be the solution of* (A.16). *Then for all other* $z \in \mathcal{C}$, *we have that*

$$\|z - z^*\| \le \left[2\frac{f(z) - f(z^*)}{\alpha}\right]^{1/2}. \tag{A.22}$$

Proof. Using (A.20), we have $\langle c, z - z^* \rangle \ge -\langle Uz^*, z - z^* \rangle$. Hence, we have

$$\begin{aligned} f(z) - f(z^*) &= \frac{1}{2}\langle z, Uz \rangle - \frac{1}{2}\langle z^*, Uz^* \rangle + \langle c, z - z^* \rangle \\ &\ge \frac{1}{2}\langle z, Uz \rangle - \frac{1}{2}\langle z^*, Uz^* \rangle - \langle Uz^*, z - z^* \rangle = \frac{1}{2}\langle z - z^*, U(z - z^*) \rangle \\ &\ge \frac{1}{2}\alpha\|z - z^*\|^2, \end{aligned}$$

where the last inequality follows from (A.17). □

Increasing and Decreasing Sequences of Sets

We now consider monotonic sequences of subsets of ℓ^2, their limits, and the behavior of the sequence of points obtained by minimizing the function $f(z)$ defined in (A.16) over each of the sets in these sequences.

Consider first a decreasing sequence of sets $\{\underline{\mathcal{C}}_J\}_{J=1,2,\ldots}$ such that

$$\text{each } \underline{\mathcal{C}}_J \subset \ell^2 \text{ is closed, convex, and nonempty;} \tag{A.23}$$

$$\underline{\mathcal{C}}_1 \supset \underline{\mathcal{C}}_2 \supset \cdots. \tag{A.24}$$

This sequence has a limit $\underline{\mathcal{C}}$ defined by

$$\underline{\mathcal{C}} = \bigcap_{J=1,2,\ldots} \underline{\mathcal{C}}_J; \tag{A.25}$$

see [11, p. 66]. It is clear that $\underline{\mathcal{C}}$ too is closed, convex, and nonempty. $\underline{\mathcal{C}}$ is simply the set of points that belong to every set $\underline{\mathcal{C}}_J$ in the sequence. We also have the following characterization.

Lemma A.2.

$$\underline{\mathcal{C}} = \{z \,|\, z = \lim_{J\to\infty} z_J, \textit{ for any convergent sequence } \{z_J\} \textit{ with } z_J \in \underline{\mathcal{C}}_J \textit{ for all } J\}. \tag{A.26}$$

Proof. Assume first that $z \in \underline{\mathcal{C}}$. The trivial sequence defined by $z_J \equiv z$ suffices to show that z belongs to the set on the right-hand side of (A.26).

Now assume that $z \notin \underline{\mathcal{C}}$. We show that there can exist no sequence $\{z_J\}$ with $z_J \in \underline{\mathcal{C}}_J$, all J with the property that $\|z_J - z\| \to 0$.

Since $z \notin \underline{\mathcal{C}}$, we have that $z \notin \underline{\mathcal{C}}_K$ for some K and indeed, by closedness of $\underline{\mathcal{C}}_K$, we have $\text{dist}(z, \underline{\mathcal{C}}_K) > 0$. Therefore, by the monotonicity property of the sequence $\{\underline{\mathcal{C}}_J\}$, we have that $z \notin \underline{\mathcal{C}}_J$ for all $J \geq K$, and in fact that

$$\text{dist}(z, \underline{\mathcal{C}}_J) \geq \text{dist}(z, \underline{\mathcal{C}}_K) > 0, \quad \text{for all } J \geq K.$$

It follows that for *any* sequence $\{z_J\}$ with the property $z_J \in \underline{\mathcal{C}}_J$, we have that

$$\|z - z_J\| \geq \text{dist}(z, \underline{\mathcal{C}}_J) \geq \text{dist}(z, \underline{\mathcal{C}}_K) > 0, \quad \text{for all } J \geq K,$$

so we cannot have $\|z - z_J\| \to 0$. □

We consider the sequence of problems $\underline{P}(J)$ defined as follows

$$\underline{P}(J): \quad \min_z f(z) \text{ subject to } z \in \underline{\mathcal{C}}_J, \tag{A.27}$$

where $f(z)$ is defined as in (A.16). By applying Theorem A.2, we can identify points $\underline{z}_J$, $J = 1, 2, \ldots$ such that $\underline{z}_J$ is the unique solution of $\underline{P}(J)$ for each J. Similarly, we define $\underline{z}$ to be the unique minimizer of f over the limiting set $\underline{\mathcal{C}}$. By the decreasing property (A.24), and the definition (A.25), we have that

$$f(\underline{z}_1) \leq f(\underline{z}_2) \leq \cdots \leq f(\underline{z}). \tag{A.28}$$

In fact, we have the following result.

Theorem A.5. *For $\underline{z}_J$, $J = 1, 2, \ldots$ and $\underline{z}$ defined in the previous paragraph, we have*

$$\lim_{J \to \infty} \underline{z}_J = \underline{z}. \tag{A.29}$$

Proof. The sequence of real numbers

$$\{f(\underline{z}_J)\}_{J=1,2,\ldots} \tag{A.30}$$

is increasing and bounded above. Using the following argument, we can show that this sequence $\{\underline{z}_J\}$ is Cauchy. Given any indices J_1 and J_2 with $J_1 < J_2$, we have that $\underline{z}_{J_2}$ is feasible in $P(J_1)$. Hence, by applying Theorem A.4 to $P(J_1)$ with $z^* = \underline{z}_{J_1}$ and $z = \underline{z}_{J_2}$, we have that

$$\|\underline{z}_{J_1} - \underline{z}_{J_2}\| \leq \left[2 \frac{f(\underline{z}_{J_1}) - f(\underline{z}_{J_2})}{\alpha} \right]^{1/2}.$$

Hence, for any $\epsilon > 0$, there exists $J_{(\epsilon)}$ such that

$$\|\underline{z}_{J_1} - \underline{z}_{J_2}\| \leq \epsilon, \quad \text{for all } J_1,\ J_2 \text{ with } J_{(\epsilon)} \leq J_1 < J_2.$$

Because of this Cauchy property and the fact that ℓ^2 is a Hilbert space, the sequence $\{\underline{z}_J\}$ converges to a limit in ℓ^2, say z^*. In fact, because of the characterization (A.26), we have $z^* \in \underline{\mathcal{C}}$. Because $f(\underline{z}_J) \uparrow f(z^*)$, we have from (A.28) that $f(z^*) \leq f(\underline{z})$. But since $z^* \in \underline{\mathcal{C}}$ and $\underline{z}$ is the unique minimizer of f over $\underline{\mathcal{C}}$, we must have $\underline{z} = z^*$, completing the proof. □

We next consider an increasing sequence of sets. Let $\{\bar{\mathcal{C}}_J\}_{J=1,2,\ldots}$ be a sequence of sets such that

$$\text{each } \bar{\mathcal{C}}_J \subset \ell^2 \text{ is closed, convex, and nonempty;} \tag{A.31}$$
$$\bar{\mathcal{C}}_1 \subset \bar{\mathcal{C}}_2 \subset \cdots . \tag{A.32}$$

This sequence has a limit $\bar{\mathcal{C}}$ defined by

$$\bar{\mathcal{C}} = \bigcup_{J=1,2,\ldots} \bar{\mathcal{C}}_J; \tag{A.33}$$

see [11, p. 66]. The set $\bar{\mathcal{C}}$ is convex and nonempty but *not* necessarily closed. As an example, consider the sets defined by

$$\bar{\mathcal{C}}_J = \{w = (w_1, w_2, w_3, \ldots) \,|\, w \in \ell^2,\ w_i = 0 \text{ for all } i \geq J\},$$

which yield an increasing sequence whose limit is

$$\bar{\mathcal{C}} = \{w = (w_1, w_2, w_3, \ldots) \,|\, w \in \ell^2,\ w_i = 0 \text{ for all } i \text{ sufficiently large}\}.$$

Although the sets $\bar{\mathcal{C}}_J$ are closed for all J, the limit $\bar{\mathcal{C}}$ is open. The point $w = (1, 1/2, 1/4, \ldots)$ lies in the closure of $\bar{\mathcal{C}}$ though not in $\bar{\mathcal{C}}$ itself.

Since the limit $\bar{\mathcal{C}}$ may be an open set, the function $f(z)$ may not attain its minimizer on this set. We can still however show convergence of the sequence of minimizers of f over $\bar{\mathcal{C}}_J$ to a point $\bar{z}$ that minimizes f over some closed set containing $\bar{\mathcal{C}}$, which we denote by $\mathcal{C}^*$.

Similarly to (A.27), we consider the sequence of problems $\bar{P}(J)$ defined as follows

$$\bar{P}(J): \quad \min_z f(z) \ \text{ subject to } z \in \bar{\mathcal{C}}_J, \tag{A.34}$$

where $f(z)$ is defined as in (A.16). By applying Theorem A.2, we can identify points $\bar{z}_J$, $J = 1, 2, \ldots$ such that $\bar{z}_J$ is the unique solution of $\bar{P}(J)$ for each J. We also define a point $\bar{z}$ and a set $\mathcal{C}^*$ as follows:

$$\bar{z} = \arg\min_{z \in \mathcal{C}^*} f(z), \quad \text{where } \mathcal{C}^* \text{ is a set satisfying} \tag{A.35a}$$
$$\mathcal{C}^* \text{ is closed}, \ \ \bar{\mathcal{C}} \subset \mathcal{C}^*, \ \text{ and } \bar{z} \in \mathrm{cl}(\bar{\mathcal{C}}). \tag{A.35b}$$

(Note that since $\mathcal{C}^*$ is closed and $\bar{\mathcal{C}} \subset \mathcal{C}^*$, we certainly have $\mathrm{cl}(\bar{\mathcal{C}}) \subset \mathcal{C}^*$.) Clearly, we have that

$$f(\bar{z}_1) \geq f(\bar{z}_2) \geq \cdots \geq f(\bar{z}). \tag{A.36}$$

In fact, we have the following result.

Theorem A.6. *For $\bar{z}_J$, $J = 1, 2, \ldots$ and $\bar{z}$ defined in the previous paragraph, we have*

$$\lim_{J \to \infty} \bar{z}_J = \bar{z}. \tag{A.37}$$

Proof. We can show that the sequence $\{\bar{z}_J\}$ is Cauchy by using as similar argument as in the proof of Theorem A.5. Hence the sequence converges, say to a point z^*. Moreover, since $\bar{z}_J \in \bar{\mathcal{C}}$ for all J, we have that $z^* \in \mathrm{cl}(\bar{\mathcal{C}})$.

Because of (A.36), we have that $f(z^*) \geq f(\bar{z})$ Suppose for the moment that this inequality is strict. Since $\bar{z} \in \mathrm{cl}(\bar{\mathcal{C}})$, there is a sequence $\{y_K\}$ with $y_K \in \bar{\mathcal{C}}$ for all K, such that $y_K \to \bar{z}$. By the definition (A.33), we can choose indices K and J_K sufficiently large that the following properties hold:

$$f(y_K) < f(z^*), \tag{A.38}$$

$$y_K \in \bar{\mathcal{C}}_J, \;\; \text{for all } J \geq J_K. \tag{A.39}$$

In particular, we have that

$$f(y_K) < f(z^*) \leq f(\bar{z}_{J_K}),$$

which is a contradiction since $\bar{z}_{J_K}$ is the minimizer of f over $\bar{\mathcal{C}}_{J_K}$. Therefore, we must have $f(z^*) = f(\bar{z})$.

Since $\bar{z}$ is the unique minimizer of f over the set $\mathcal{C}^*$, it certainly is the minimizer of f over $\mathrm{cl}(\bar{\mathcal{C}})$. Since $z^* \in \mathrm{cl}(\bar{\mathcal{C}})$ and $f(z^*) = f(\bar{z})$, we have $z^* = \bar{z}$, completing the proof. □

Proofs of the convergence theorems

We now prove the results of Theorems 3.2, 3.3 and 3.4 by treating (2.4), (3.1), (3.6) as strictly convex quadratic programs in the variable

$$z = (v_0, v_1, \ldots) \in \ell^2,$$

where

$$\begin{aligned} \ell^2 &= \left\{ z = (z_1, z_2, \ldots) \,|\, z_i \in \mathbb{R},\; i = 1, 2, \ldots;\; \sum_{i=1}^{\infty} z_i^2 < \infty \right\} \\ &= \left\{ (v_0, v_1, v_2, \ldots) \,|\, v_j \in \mathbb{R}^m,\; j = 0, 1, \ldots;\; \sum_{j=0}^{\infty} \|v_j\|_2^2 < \infty \right\}. \end{aligned} \tag{A.40}$$

By using the state equation to eliminate w_j for $j = 1, 2, \ldots$, all three problems (2.4), (3.1), (3.6) have the following form:

$$\min_z f(z) = \frac{1}{2}\langle z, Uz\rangle + \langle c, z\rangle, \qquad \text{subject to } z \in \mathcal{C}, \tag{A.41}$$

where $\mathcal{C}$ is a closed, convex subset of ℓ^2.

Note that the restriction (A.40) does not hamper our ability to consider interesting points. An input sequence $\{v_j\}_{j=0}^{\infty}$ for which $\sum_{j=0}^{\infty} \|v_j\|_2^2 = \infty$ is such that the objective function in (2.4) would be infinite, since R is positive definite and Q is positive semidefinite.

Proof of Theorem 3.3. Let $\underline{\mathcal{C}}_N$ denote the feasible set in ℓ^2 for this problem; that is, the set of vectors $v = (v_0, v_1, v_2, \dots)$ for which there is a $w = (w_0, w_1, w_2, \dots)$ such that (v, w) satisfy (3.6b) and, in addition, $\sum_{j=0}^{\infty} \|v_j\|_2^2 < \infty$. We obtain (3.6) by setting $\mathcal{C} = \underline{\mathcal{C}}_N$ in (A.41). It is clear that $\{\underline{\mathcal{C}}_N\}_{N=1,2,3\dots}$ is a decreasing sequence of sets. Moreover, it is easy to see that the set defined by

$$\underline{\mathcal{C}} = \bigcap_{N=1,2,\dots} \underline{\mathcal{C}}_N$$

is simply the feasible set for (2.4); that is, the set of vectors $v = (v_0, v_1, v_2, \dots)$ for which there is an $w = (w_0, w_1, w_2, \dots)$ such that (v, w) satisfy (2.4b), (2.4c), (2.4d) and, in addition, $\sum_{j=0}^{\infty} \|v_j\|_2^2 < \infty$. We obtain (2.4) by setting $\mathcal{C} = \underline{\mathcal{C}}$ in (A.41).

We introduce the notation $\underline{z}_N = (\underline{v}_{N,0}, \underline{v}_{N,1}, \underline{v}_{N,2}, \dots)$ for the minimizer of (A.41) with $\mathcal{C} = \underline{\mathcal{C}}_N$ (equivalently, (3.6)), and $z^* = (v_0^*, v_1^*, v_2^*, \dots)$ for the minimizer of (A.41) with $\mathcal{C} = \underline{\mathcal{C}}$ (equivalently, (2.4)). By applying Theorem A.5 from Appendix A.2, we have that

$$\lim_{N \to \infty} \underline{z}_N = z^*, \qquad \Phi_N^l \uparrow \Phi^* \, . \tag{A.42}$$

□

Proof of Theorem 3.2. Let $\bar{\mathcal{C}}_N$ denote the feasible set in ℓ^2 for this problem; that is, the set of vectors $v = (v_0, v_1, v_2, \dots)$ for which there is an $w = (w_0, w_1, w_2, \dots)$ such that (v, w) satisfy (3.1b), (3.1c), and, in addition, $\sum_{j=0}^{\infty} \|v_j\|_2^2 < \infty$. We obtain (3.1) by setting $\mathcal{C} = \bar{\mathcal{C}}_N$ in (A.41). It is clear that $\{\bar{\mathcal{C}}_N\}_{N=1,2,3\dots}$ is an increasing sequence of sets. Its limit, defined by

$$\bar{\mathcal{C}} = \bigcup_{N=1,2,\dots} \bar{\mathcal{C}}_N$$

is the set of vectors $v \in \ell^2$ for which there is an w such that (v, w) satisfy (3.1b), as well as satisfying (3.1c) *for some value of* N. This is not the same set as the feasible set $\underline{\mathcal{C}}$ for (2.4), in which the restriction (3.1c) does not appear at all.

The remainder of our discussion below shows that we can apply Theorem A.6 in to this case, and arrive at the desired conclusion that the sequence of minimizers of the upper-bounding problem (3.1) converges to the minimizer of the optimal problem (2.4).

We can identify the feasible set $\underline{\mathcal{C}}$ for (2.4) with the set $\mathcal{C}^*$ in (A.35), and identify the minimizer z^* of (2.4) with $\bar{z}$ of (A.35). Note that $\underline{\mathcal{C}}$ is certainly closed, and that since $\bar{\mathcal{C}}_N \subset \underline{\mathcal{C}}$ for every N we certainly have $\bar{\mathcal{C}} \subset \underline{\mathcal{C}}$. It remains only to show that $z^* \in \text{cl}(\bar{\mathcal{C}})$; that is, the solution of (2.4) lies in the closure of the set formed by the union of the feasible sets for (3.1), over all N.

To show that z^* lies in the closure of $\bar{\mathcal{C}}$, we construct a sequence $\{z_N\}$ such that

$$z_N \in \bar{\mathcal{C}}_N \subset \bar{\mathcal{C}}, \;\; \text{for all } N \text{ sufficiently large, and } z_N \to z^*.$$

Writing

$$z^* = (v_0^*, v_1^*, v_2^*, \dots),$$

we have by the definition of $\underline{\mathcal{C}}$ that there is a vector $w^* = (w_0^*, w_1^*, w_2^*, \dots)$ such that (v^*, w^*) satisfies the conditions (2.4b), (2.4c), (2.4d). Since $z^* \in \ell^2$, we also have that

$$\lim_{N\to\infty} \sum_{j=N}^{\infty} \|v_i^*\|_2^2 = 0. \tag{A.43}$$

We now construct z_N by perturbing the optimal vector z^* in such a way that all the unstable modes of the system are zeroed at time N. Theorem A.1 shows that such a z_N exists and it is feasible with respect to all constraints when $N > N'$ for some positive N'. After stage N, the input is set to zero, and z_N is as follows

$$z_N = (v_{N,0}, v_{N,1}, \dots, v_{N,N-1}, 0, \dots).$$

Clearly $z_N \in \bar{\mathcal{C}}_N$. We can write

$$\|z_N - z^*\|^2 \le \sum_{j=0}^{N-1} \|v_{N,j} - v_j^*\|_2^2 + \sum_{j=N}^{\infty} \|v_j^*\|_2^2$$

From (A.14) and (A.15), we have that the first term goes to zero as $N \to \infty$ and, from (A.43), we have that also the second term goes to zero as $N \to \infty$. We conclude that $\|z_N - z^*\| \to 0$ as $N \to \infty$ and hence that $z^* \in \mathrm{cl}(\bar{\mathcal{C}})$, as claimed.

Having verified that the assumptions of Theorem A.6 in are satisfied, we can now apply this theorem to deduce that the sequence of minimizers $\bar{z}_N$ of the upper bounding problem (3.1) (alternatively, the problem obtained by replaying $\mathcal{C}$ by $\bar{\mathcal{C}}_N$ in (A.41)), approaches the solution of the optimal problem (2.4), and that the sequence of objective values converges monotonically to the optimal objective value. That is, we have

$$\lim_{N\to\infty} \bar{z}_N = z^*, \qquad \Phi_N^u \downarrow \Phi^*. \tag{A.44}$$

□

Proof of Theorem 3.4. We can apply the result of Theorem A.4 to obtain an estimate of the distance of the first component $\bar{v}_{N,0}$ from the solution $\bar{z}_N$ of $\mathcal{U}(N)$ to the corresponding optimal component v_0^*. In Theorem A.4, let the canonical problem (A.16) correspond to the optimal problem (2.4). We have from this result that

$$\frac{1}{2}\alpha\|\bar{z}_N - z^*\|^2 \le \overline{\Phi}_N - \Phi^* \le \overline{\Phi}_N - \underline{\Phi}_N,$$

where we use the fact that $\Phi^* \ge \underline{\Phi}$ in the second inequality. Thus we deduce that

$$\|\bar{v}_{N,0} - v_0^*\| \le \|\bar{z}_N - z^*\| \le \left[\frac{2}{\alpha}(\Phi_N^u - \Phi_N^l)\right]^{1/2}. \tag{A.45}$$

□