Optimal Control of Linear Systems with State Equality Constraints ⋆

Sangho Ko and Robert R. Bitmead

Department of Mechanical and Aerospace Engineering, University of California, San Diego, 9500 Gilman Drive, La Jolla, California, 92093-0411, USA.

Abstract

This paper deals with the optimal control problem of systems with state linear equality constraints. For deterministic linear systems, first we find various existence conditions for constraining state feedback control and determine all constraining feedback gains, from which the optimal feedback gain is derived by using the result of singular optimal control. For systems with stochastic process noises, it is shown that the same gain used for constraining the deterministic system also optimally constrains the expectation of states inside the constraint subspace and minimizes the expectation of the squared constraint error. We compare and discuss performance differences between unconstrained (using penalty method), projected, and constrained controllers for both deterministic and stochastic systems. Finally, numerical examples are used to demonstrate the performance difference of the three controllers.

Key words: Linear optimal control; Singular control; Constraints; Kalman filters; Projection

1 Introduction

Dealing with constraints on the state and/or input variable is one of the fundamental tasks in control synthesis problems and, hence, has drawn much attention of the dynamics and control community, since it is closely connected with system performance and, thus, fulfillment of given system specifications. Recently, a number of modern model-based control design methods seek to deal with system constraints directly rather than through their implicit incorporation via penalty or barrier functions. Such is the case for Model Predictive Control, where part of the attraction of the approach is the introduction of constraints into the formulation without compromising the scalar control objective function (Maciejowski 2002).

From the viewpoint of its origin, a state/input constraint can be a *physical constraint*, physically imposed upon the system state and/or input, or a *design constraint*, deliberately imposed to avoid undesirable states by using corrective control action. Of the various kinds of constraints, this paper focuses on a special case: equality state constraints (which are also known as algebraic equation state constraints.) This topic has been extensively researched under the notion of *invariant sets*. Specially, the so-called *controlled invariant set* [1] has been a fundamental subject of many

⋆ Corresponding author Robert R. Bitmead. Tel. +1-858-822-3477. Fax +1-858-822-3107.
Email addresses: sko@ucsd.edu (Sangho Ko), rbitmead@ucsd.edu (Robert R. Bitmead).

[1] A set is *controlled invariant* if, for all initial conditions chosen among its elements, we can keep the trajectory inside the set by means of a proper control action.

researchers, because of its direct relation to finding a set of feasible initial states whose future trajectories meet design specifications and also characterizing such constraining control laws (Blanchini 1999).

In the case of physical state equality constraints, it is always possible to reduce the system parametrization to fit in a lower dimensional state space. For robotic systems with (hard) holonomic constraints, McClamroch & Wang (1988) derived stable controllers by decomposing the constrained system into a reduced order dynamic system and a static system. But, sometimes keeping the non-reduced state space has also good reasons as described in Hemami & Wyman (1979), where a general dynamic model for biped locomotion was derived in a non-reduced state equation form and a pole-assignment algorithm was devised for a linearized state equation, which, in general, does not satisfy the hard constraint without control input. Hence, their design methodology can be applied for the system having design constraints. That is, for a system represented by

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u},$$

they found, so called, $\mathcal{X}_c$-constrained linear state feedback gain $\mathbf{K}$ such that

$$(\mathbf{A} - \mathbf{B}\mathbf{K})(\mathcal{X}_c) \subset \mathcal{X}_c,$$

where $\mathcal{X}_c$ is a given constraint subspace.

This pole-assignment problem was also studied in Yu & Müller (1994) in which a method of designing pole-assignment controllers was developed to constrain systems, firstly by finding a suitable form of feedback gain and then by designing a specific pole-assignment controller. Existence conditions for the pole-assignment controller were also studied. Similarly to the above-mentioned equality (or algebraic) constraint approach, a method of asymptotic constraint satisfaction was considered in the name of *subspace stabilization* (Johnson 1973, Johnson 2000) or *stabilized constraint* (Hahn 1992, Yu, Lin & Müller 1996). Here, instead of considering algebraic constraint relations, the stabilized constraints having stable stationary solutions were used, where the limiting solution is identical to the algebraic constraints. Recently, this method was applied to flight control problems as in Tournes & Johnson (2001) and Tournes & Landrum (2003). Hahn (1992) studied a pole-assignment controller and Yu et al. (1996) designed a linear quadratic (LQ) regulator for systems with stabilized constraints.

In the case of LQ regulator problems, we expect that the state equality constraints cause reduction of the allowable input space which, in turn, produces a performance degradation in terms of optimal performance index, compared to that of the unconstrained LQ regulator problem. This paper studies this and its related problems. Mathematically, this problem is very similar to that considered in Oloomi & Shafai (1997) for avoiding the transient mismatch phenomenon in singularly perturbed systems, where the authors found a way for determining the gains taking a certain structure without finding existence conditions. Therefore, the result of this paper can be extended for finding the desired gain for that problem in a more organized manner.

The contributions of this paper are fourfold, which are

- Identification of the existence conditions for $\mathcal{X}_c$-constrained feedback inputs and finding of all $\mathcal{X}_c$-constrained feedback gains by solving a certain linear matrix equation (Section 2.2)
- finding of the constrained LQ optimal controller among all the $\mathcal{X}_c$-constrained feedback controllers through the result of singular optimal control (Section 3)
- Comparison of the performance of three controllers that can be used for constraining the system: the unconstrained LQ optimal controller using penalty method, the projected LQ optimal controller obtained by projecting the unconstrained LQ optimal feedback gain onto the $\mathcal{X}_c$-constrained feedback gain set, and the current constrained LQ optimal controller (Section 3)

- Extension of these results to the corresponding stochastic system case and obtaining of the result that the same feedback gain used for optimally constraining the deterministic system constrains the expectation of the state and, furthermore, it minimizes the expectation of the squared constraint error (Section 4).

We confirm these results through numerical examples in Section 5.

Notations: In this paper, matrices will be denoted by upper case boldface (e.g., $\mathbf{A}$), linear spaces are denoted by calligraphic uppercase (e.g., $\mathcal{A}$), column matrices (vectors) will be denoted by lower case boldface (e.g., $\mathbf{x}$), and scalars will be denoted by lower case (e.g., y) or upper case (e.g., Y). For a matrix $\mathbf{A}$, $\mathbf{A}^T$ denotes its transpose and $\mathbf{A}^\dagger$ represents the Moore-Penrose inverse of $\mathbf{A}$ and $\|\mathbf{A}\|_F$ represents the Frobenius norm of $\mathbf{A}$, and $\mathcal{R}(\mathbf{A})$ and $\mathcal{N}(\mathbf{A})$ denote the column and the null space of $\mathbf{A}$. For a symmetric matrix $\mathbf{P}$, $\mathbf{P} > \mathbf{0}$ or $\mathbf{P} \geq \mathbf{0}$ denotes the fact that $\mathbf{P}$ is positive definite or positive semi-definite, respectively. $\mathbf{I}_m$ denotes the $m \times m$ identity matrix. For a random vector $\mathbf{x}$, $\mathcal{E}\{\mathbf{x}\}$ or $\mathcal{E}_x\{\mathbf{x}\}$ represents the expectation of $\mathbf{x}$ and $\mathcal{E}\{\mathbf{x}|\mathbf{y}\}$ or $\mathcal{E}_{x|y}\{\mathbf{x}|\mathbf{y}\}$ denote the conditional mean given $\mathbf{y}$.

2 Constrained System and Control

2.1 Classification of Constrained Systems

Before dealing with the constrained control, we identify the class of constrained systems represented by

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k \tag{1}$$

where $\mathbf{x}_k \in \mathbb{R}^{n\times 1}$ and $\mathbf{u}_k \in \mathbb{R}^m$, and $\mathbf{w}_k \in \mathbb{R}^n$ represent the state, the input, and the process noise, respectively. Here, the process noise $\mathbf{w}_k$ is assumed to have a gaussian distribution of zero-mean and covariance $\mathbf{Q}_e$. Since the allowable space of $\mathbf{x}_k$ with constraint is $\mathcal{N}(\mathbf{D})$, $\mathbf{x}_{k+1}$ given by (1) also must satisfy the constraint, that is, $\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k \in \mathcal{N}(\mathbf{D})$, for which we identify the following possible cases (Ko & Bitmead 2005):

Case 1: $(\mathbf{A}\mathbf{x}_k,\ \mathbf{B}\mathbf{u}_k,\ \mathbf{w}_k) \notin \mathcal{N}(\mathbf{D})$
Since the sum of the three elements must satisfy $\mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k \in \mathcal{N}(\mathbf{D})$ for any $\mathbf{x}_k \in \mathcal{N}(\mathbf{D})$, this case demands that the noise $\mathbf{w}_k$ be correlated with current input $\mathbf{u}_k$ and state $\mathbf{x}_k$. Hence it is not a proper Markovian system model.

Case 2: $(\mathbf{A}\mathbf{x}_k,\ \mathbf{B}\mathbf{u}_k) \notin \mathcal{N}(\mathbf{D})$ but $\mathbf{w}_k \in \mathcal{N}(\mathbf{D})$
Similarly, the sum of the first two elements must satisfy $\mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k \in \mathcal{N}(\mathbf{D})$. This case allows uncorrelated noise sequences $\mathbf{w}_k$ with the input $\mathbf{u}_k$ or the state $\mathbf{x}_k$, but the system cannot maintain the state constraint without corrective action of the input $\mathbf{u}_k$. Hence, this model is suitable for modeling systems having *design constraints*.

Case 3: $(\mathbf{A}\mathbf{x}_k,\ \mathbf{B}\mathbf{u}_k,\ \mathbf{w}_k) \in \mathcal{N}(\mathbf{D})$
This case also allows uncorrelated noise sequences $\mathbf{w}_k$ with the input $\mathbf{u}_k$ or the state $\mathbf{x}_k$ and has a proper form for modeling systems with *physical constraints*, since regardless of the corrective input $\mathbf{u}_k$, the state stays within the constraint surface $\mathcal{N}(\mathbf{D})$. Since for all $\mathbf{x}_k \in \mathcal{N}(\mathbf{D})$, it is required that $\mathbf{A}\mathbf{x}_k \in \mathcal{N}(\mathbf{D})$, $\mathcal{N}(\mathbf{D})$ is $\mathbf{A}$-invariant (Wonham 1979).

For the second and the third cases to hold, the density function of the process noise must have support in a lower dimensional space than the whole state space and, therefore, the covariance matrix $\mathbf{Q}_e$ is a singular matrix. As will be shown in later sections, it is impossible for the second case, a system with design constraints, to constrain the stochastic state inside the constraint subspace, unless the complete state information is available for the state feedback. With incomplete state information via noisy measurements, we cannot constrain the state in a constraint subspace and, therefore, in this paper, we assume that $\mathbf{w}_k$ is independent of $\mathbf{x}_k$ and $\mathbf{u}_k$ and has support on the whole state space.

2.2 Constrained Control

In this section, we deal with constrained control of a system with design constraints considered in Section 2.1 when the process noise is absent. That is, we consider the following problem: for a given discrete system

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k \tag{2}$$

with a design constraint

$$\mathbf{x}_k \in \mathcal{X}_c = \{\mathbf{x} : \mathbf{D}\mathbf{x} = \mathbf{0}\}, \tag{3}$$

find the optimal feedback control law

$$\mathbf{u}_k = -\mathbf{K}_k\mathbf{x}_k \tag{4}$$

which minimizes

$$J = \mathbf{x}_N^T\mathbf{Q}_N\mathbf{x}_N + \sum_{k=0}^{N-1}\left[\mathbf{x}_k^T\mathbf{Q}_c\mathbf{x}_k + \mathbf{u}_k^T\mathbf{R}_c\mathbf{u}_k\right] \tag{5}$$

where $\mathbf{x}_k \in \mathbb{R}^{n\times 1}$ and $\mathbf{u}_k \in \mathbb{R}^{m\times 1}$, and it is assumed that $\mathbf{Q}_N \geq \mathbf{0}$, $\mathbf{Q}_c \geq \mathbf{0}$, $\mathbf{R}_c > \mathbf{0}$, and $\mathbf{D} \in \mathbb{R}^{c\times n}$ has full row rank. If $\mathbf{D}$ is not of full row rank, there exist redundant state constraints. In that case, we can simply remove linearly dependent rows from $\mathbf{D}$.

Definition 1 *Let $\mathcal{X} = \mathbb{R}^n$ and $\mathcal{U} = \mathbb{R}^m$ be the state and the input space, respectively, and $\mathcal{X}_c \subset \mathcal{X}$ be an arbitrary subspace called the constraint subspace, and $\mathcal{K}$ be the set of all $n \times m$ linear state variable feedback gains. A feedback map $\mathbf{K} : \mathcal{X} \mapsto \mathcal{U}$ such that*

$$\mathcal{K}_{\mathcal{X}_c} \triangleq \left\{\mathbf{K} : (\mathbf{A} - \mathbf{B}\mathbf{K})(\mathcal{X}_c) \subset \mathcal{X}_c\right\} \subset \mathcal{K} \tag{6}$$

is called $\mathcal{X}_c$-constrained.

It is shown (Hemami & Wyman 1979) that the set $\mathcal{K}_{\mathcal{X}_c}$ of all $\mathcal{X}_c$-constrained feedbacks for the system pair $[\mathbf{A}, \mathbf{B}]$ is an affine subset of the set of all linear state-variable feedback gains $\mathbf{K} : \mathcal{U} \mapsto \mathcal{X}$. For later use, we need the following Lemma 1.

Lemma 1 (Skelton, Iwasaki & Grigoriadis (1998), Theorem 2.3.1) *Let $\mathbf{A}$, $\mathbf{X}$, $\mathbf{B}$, and $\mathbf{Y}$ be matrices with consistent dimensions. Then the following statements are equivalent:*

(i) *The equation $\mathbf{A}\mathbf{X}\mathbf{B} = \mathbf{Y}$ has a solution $\mathbf{X}$.*
(ii) *$\mathbf{A}$, $\mathbf{B}$ and $\mathbf{Y}$ satisfy*

$$\mathbf{A}\mathbf{A}^\dagger\mathbf{Y}\mathbf{B}^\dagger\mathbf{B} = \mathbf{Y}. \tag{7}$$

(iii) *$\mathbf{A}$, $\mathbf{B}$ and $\mathbf{Y}$ satisfy $(\mathbf{I} - \mathbf{A}\mathbf{A}^\dagger)\mathbf{Y} = \mathbf{0}, \quad \mathbf{Y}(\mathbf{I} - \mathbf{B}^\dagger\mathbf{B}) = \mathbf{0}$.*

In this case, all solutions are

$$\mathbf{X} = \mathbf{A}^\dagger\mathbf{Y}\mathbf{B}^\dagger + \mathbf{G} - \mathbf{A}^\dagger\mathbf{A}\mathbf{G}\mathbf{B}\mathbf{B}^\dagger \tag{8}$$

where $\mathbf{G}$ is an arbitrary matrix with consistent dimension.

Remark 1 *If the equation $\mathbf{A}\mathbf{X}\mathbf{B} = \mathbf{Y}$ is consistent, that is, if the condition* (ii) *of Lemma 1 is satisfied, then $\mathbf{X}$ in* (8) *solves the minimization problem*

$$\min \|\mathbf{A}\mathbf{X}\mathbf{B} - \mathbf{Y}\|_F \tag{9}$$

and the minimum value of $\|\mathbf{A}\mathbf{X}\mathbf{B} - \mathbf{Y}\|_F$ is zero. Therefore, $\mathbf{X}$ given in (8) *can be interpreted as the projection of $\mathbf{G}$ onto the subset given by $\mathcal{F}_\mathbf{X} = \{\mathbf{X} : \mathbf{A}\mathbf{X}\mathbf{B} = \mathbf{Y}\}$.*

The following Lemma 2 provides equivalent existence conditions for the non-empty set $\mathcal{K}_{\mathcal{X}_c}$ in Definition 1.

Lemma 2 *The following statements are equivalent:*

(i) $\mathcal{K}_{\mathcal{X}_c}$ *is non-empty.*
(ii) $\mathbf{A}(\mathcal{X}_c) \subset \mathcal{X}_c + \mathbf{B}(\mathcal{U})$
(iii) *For any basis matrix* $\mathbf{Z}$ *of the subspace* $\mathcal{X}_c$*,*

$$\mathbf{DBKZ} = \mathbf{DAZ}. \tag{10}$$

(iv) $\mathbf{P}_{\mathcal{N}(\mathbf{DB})^T}\mathbf{DAZ} = \mathbf{0}$
(v) *There exist* $c \times c$ *matrix* $\mathbf{H}$ *such that*

$$\mathbf{HD} = \mathbf{D}(\mathbf{A} - \mathbf{BK}). \tag{11}$$

In this case, all $\mathcal{X}_c$*-constrained feedback gains are given by* [2]

$$\mathbf{K} = \mathbf{G} + \Big[\mathbf{G}_0 - \mathbf{P}_{\mathcal{R}(\mathbf{DB})^T}\mathbf{G}\Big]\mathbf{P}_{\mathcal{N}(\mathbf{D})} = \mathbf{G} + \mathbf{P}_{\mathcal{R}(\mathbf{DB})^T}\big(\mathbf{G}_0 - \mathbf{G}\big)\mathbf{P}_{\mathcal{N}(\mathbf{D})} \tag{12}$$

where $\mathbf{G}$ *is an arbitrary matrix with consistent dimension and* $\mathbf{G}_0 \triangleq (\mathbf{DB})^\dagger\mathbf{DA}$. $\mathbf{P}_{\mathcal{R}(\mathbf{DB})^T} \triangleq (\mathbf{DB})^\dagger(\mathbf{DB})$, $\mathbf{P}_{\mathcal{N}(\mathbf{DB})^T} \triangleq \mathbf{I} - (\mathbf{DB})(\mathbf{DB})^\dagger$*, and* $\mathbf{P}_{\mathcal{N}(\mathbf{D})} \triangleq \mathbf{I} - \mathbf{D}^\dagger\mathbf{D}$ *are the orthogonal projection matrices onto the row and the left null space of* $\mathbf{DB}$*, and the null space of* $\mathbf{D}$*, respectively.*

PROOF. Proofs for (i) $\leftrightarrow$ (ii), and (i) $\leftrightarrow$ (v) are given in Hemami & Wyman (1979) and Castelan & Hennet (1992), respectively. For (iii) $\leftrightarrow$ (v), first assume that (v) is true. By multiplying $\mathbf{Z}$ from the right on both sides of (11), we have

$$\mathbf{HDZ} = \mathbf{0} = \mathbf{D}(\mathbf{A} - \mathbf{BK})\mathbf{Z} \tag{13}$$

which is (10). Now, suppose (iii) is true, which implies $\mathcal{N}\left[\mathbf{D}(\mathbf{A} - \mathbf{BK})\right] \supseteq \mathcal{R}\left(\mathbf{Z}\right) = \mathcal{N}\left(\mathbf{D}\right)$ or $\mathcal{R}\left[\mathbf{D}(\mathbf{A} - \mathbf{BK})\right]^T \subseteq \mathcal{R}\left(\mathbf{D}^T\right)$. Therefore, for given $\boldsymbol{\lambda}_i$, there exist $\boldsymbol{\gamma}_i$ such that $\boldsymbol{\lambda}_i^T\mathbf{D}(\mathbf{A} - \mathbf{BK}) = \boldsymbol{\gamma}_i^T\mathbf{D}$, for $i = 1, \cdots, c$. Here, c denotes the number of constraints. Now choosing $\boldsymbol{\lambda}_i^T$ such that $\Big[\boldsymbol{\lambda}_1 \cdots \boldsymbol{\lambda}_c\Big]^T = \mathbf{I}$ yields $\mathbf{D}(\mathbf{A} - \mathbf{BK}) = \Big[\boldsymbol{\gamma}_1 \cdots \boldsymbol{\gamma}_c\Big]^T \mathbf{D} \triangleq \mathbf{HD}$, which is (11). The condition (iv) is obtained from (ii) or (iii) of Lemma 1 by applying it to (10) and also we obtain all $\mathcal{X}_c$-constrained feedback gains (12) by using (8). □

Remark 2 If $(\mathbf{DB})$ is invertible or has full row rank, $\mathbf{P}_{\mathcal{N}(\mathbf{DB})^T} = \mathbf{I} - (\mathbf{DB})(\mathbf{DB})^\dagger = \mathbf{0}$. Then, the condition (iv) of Lemma 2 is always satisfied. Therefore, $\mathcal{K}_{\mathcal{X}_c}$ is non-empty. Specially, if $(\mathbf{DB})$ is invertible, the feedback gain becomes a fixed one

$$\mathbf{u}_k = -\mathbf{K}\mathbf{x}_k = -\mathbf{K}\mathbf{P}_{\mathcal{N}(\mathbf{D})}\mathbf{x}_k = -\mathbf{G}_0\mathbf{P}_{\mathcal{N}(\mathbf{D})}\mathbf{x}_k = -\mathbf{G}_0\mathbf{x}_k = -(\mathbf{DB})^{-1}\mathbf{DA}\mathbf{x}_k.$$

Hence, in this case, we cannot change control law design and need to change design constraints ($\mathbf{D}$) or system input matrix ($\mathbf{B}$), if this unique controller does not show desirable results.

Remark 3 The feedback gain (12) guarantees $\mathbf{x}_k \in \mathcal{X}_c$ for all k, provided that the initial state vector $\mathbf{x}_0$ is chosen from the constraint subspace $\mathcal{X}_c$. Now, suppose that $\mathbf{x}_k \notin \mathcal{X}_c$. Then, from (2), (4), (12), and the condition (iv) of Lemma 2, it can be shown that

$$\mathbf{D}\Big[\mathbf{x}_{k+1}^r - (\mathbf{A} - \mathbf{BG})\mathbf{x}_k^r\Big] = \mathbf{0}, \tag{14}$$

where $\mathbf{x}_k^r$ is the $\mathcal{R}(\mathbf{D}^T)$-component of $\mathbf{x}_k$. Therefore, there exist $\boldsymbol{\lambda}_k \in \mathcal{X}_c = \mathcal{N}(\mathbf{D})$ such that

$$\mathbf{x}_{k+1}^r - \mathbf{P}_{\mathcal{R}(\mathbf{D}^T)}(\mathbf{A} - \mathbf{BG})\mathbf{x}_k^r = \mathbf{P}_{\mathcal{N}(\mathbf{D})}(\mathbf{A} - \mathbf{BG})\mathbf{x}_k^r + \boldsymbol{\lambda}_k. \tag{15}$$

[2] If this gain is multiplied on the right by $\mathbf{x}_k \in \mathcal{N}(\mathbf{D})$, $\mathbf{P}_{\mathcal{N}(\mathbf{D})}$ can be omitted in (12).

Since the left-hand side of (15) is in the row space of $\mathbf{D}$ and the right-hand side is in the null space of $\mathbf{D}$, we have

$$\mathbf{x}_{k+1}^r = \mathbf{P}_{\mathcal{R}(\mathbf{D}^T)}(\mathbf{A}-\mathbf{B}\mathbf{G})\mathbf{x}_k^r. \tag{16}$$

Therefore, for stable $(\mathbf{A}-\mathbf{B}\mathbf{G})$ (which implies that $\mathbf{P}_{\mathcal{R}(\mathbf{D}^T)}(\mathbf{A}-\mathbf{B}\mathbf{G})$ is also stable), $\mathbf{x}_k^r \to \mathbf{0}$, asymptotically. This means, for an initial state not in the constraint subspace, the system satisfies the constraint asymptotically.

3 Deterministic Constrained Optimal Control

Without any state constraint, it is well known that the performance index (5) is minimized by the control law

$$\mathbf{u}_k^u = -\mathbf{K}_k^u \mathbf{x}_k \tag{17}$$

where

$$\mathbf{K}_k^u = (\mathbf{B}^T\mathbf{P}_{k+1}^u\mathbf{B}+\mathbf{R}_c)^{-1}\mathbf{B}^T\mathbf{P}_{k+1}^u\mathbf{A} \tag{18}$$

Here, $\mathbf{P}_k^u$ satisfies the following Riccati Difference Equation, with the initial condition $\mathbf{P}_N^u = \mathbf{Q}_N$,

$$\mathbf{P}_k^u = \mathbf{A}^T\mathbf{P}_{k+1}^u\mathbf{A}+\mathbf{Q}_c-\mathbf{A}^T\mathbf{P}_{k+1}^u\mathbf{B}\left(\mathbf{B}^T\mathbf{P}_{k+1}^u\mathbf{B}+\mathbf{R}_c\right)^{-1}\mathbf{B}^T\mathbf{P}_{k+1}^u\mathbf{A},\ \text{for } k=1,\cdots,N-1. \tag{19}$$

3.1 Projected Linear Quadratic Optimal Control

For the system (2) which satisfies one of the conditions of Lemma 2, an $\mathcal{X}_c$-constrained feedback control can be obtained by projecting the unconstrained Linear Quadratic (LQ) optimal feedback gain (18) onto the set $\mathcal{K}_{\mathcal{X}_c}$ represented by (6) and (12). Then, according to Remark 1, the resulted control is termed the projected LQ optimal control, which satisfies the following theorem.

Theorem 1 (Projected LQ Optimal Control) *The projected LQ optimal control that is obtained by projecting the unconstrained LQ optimal gain* (18) *onto the set* $\mathcal{K}_{\mathcal{X}_c}$ *has the following performance:*

$$J^p \triangleq \mathbf{x}_N^T\mathbf{Q}_N\mathbf{x}_N + \sum_{k=0}^{N-1}\left[\mathbf{x}_k^T\mathbf{Q}_c\mathbf{x}_k + \mathbf{u}_k^{p^T}\mathbf{R}_c\mathbf{u}_k^p\right] = \mathbf{x}_0^T\mathbf{P}_0^p\mathbf{x}_0, \tag{20}$$

where

$$\mathbf{u}_k^p = -\ \mathbf{K}_k^p\mathbf{x}_k \tag{21}$$

with

$$\mathbf{K}_k^p = \mathbf{K}_k^u + \mathbf{P}_{\mathcal{R}(\mathbf{DB})^T}(\mathbf{G}_0 - \mathbf{K}_k^u). \tag{22}$$

Here, $\mathbf{P}_k^p$ *satisfies the following Riccati-like equation with the initial condition* $\mathbf{P}_N^p = \mathbf{Q}_N$

$$\mathbf{P}_k^p = \mathbf{A}^T\mathbf{P}_{k+1}^p\mathbf{A}+\mathbf{Q}_c-\bar{\mathbf{K}}_k^{p^T}\left(\mathbf{B}^T\mathbf{P}_{k+1}^p\mathbf{B}+\mathbf{R}_c\right)\bar{\mathbf{K}}_k^p+\boldsymbol{\nabla}_k^p \tag{23}$$

where

$$\bar{\mathbf{K}}_k^p \triangleq (\mathbf{B}^T\mathbf{P}_{k+1}^p\mathbf{B}+\mathbf{R}_c)^{-1}\mathbf{B}^T\mathbf{P}_{k+1}^p\mathbf{A} \tag{24}$$

$$\boldsymbol{\nabla}_k^p \triangleq (\mathbf{K}_k^p-\bar{\mathbf{K}}_k^p)^T\hat{\mathbf{R}}_k^p(\mathbf{K}_k^p-\bar{\mathbf{K}}_k^p), \tag{25}$$

$$\hat{\mathbf{R}}_k^p \triangleq \mathbf{B}^T\mathbf{P}_{k+1}^p\mathbf{B}+\mathbf{R}_c. \tag{26}$$

PROOF. See Appendix A. □

3.2 Constrained Linear Quadratic Optimal Control

In this section, we derive the LQ optimal feedback control for the system (2) with the constraint $\mathbf{x}_k \in \mathcal{X}_c$, and compare the performance of this with that of the projected controller derived in Section 3.1. With the constaint, the state feedback gain must be of the structure (12) and, for any $\mathbf{x}_k \in \mathcal{X}_c$, the input $\mathbf{u}_k$ is given by

$$\mathbf{u}_k = -\Big[\mathbf{G} + \mathbf{G}_0 - \mathbf{P}_{\mathcal{R}(\mathbf{DB})^T}\mathbf{G}\Big]\mathbf{P}_{\mathcal{N}(\mathbf{D})}\mathbf{x}_k = -\mathbf{G}_0\ \mathbf{x}_k + \mathbf{P}_{\mathcal{N}(\mathbf{DB})}\underbrace{\big(-\mathbf{G}\mathbf{x}_k\big)}_{\triangleq \bar{\mathbf{u}}_k} \tag{27}$$

where $\mathbf{P}_{\mathcal{N}(\mathbf{DB})} = \mathbf{I} - \mathbf{P}_{\mathcal{R}(\mathbf{DB})^T}$ is the projection matrix onto the null space of $\mathbf{DB}$. Therefore with (27), the state equation (2) can be rewritten as

$$\mathbf{x}_{k+1} = \bar{\mathbf{A}}\mathbf{x}_k + \bar{\mathbf{B}}\bar{\mathbf{u}}_k \tag{28}$$

where

$$\bar{\mathbf{A}} \triangleq \mathbf{A} - \mathbf{B}\mathbf{G}_0, \quad \bar{\mathbf{B}} \triangleq \mathbf{B}\mathbf{P}_{\mathcal{N}(\mathbf{DB})}. \tag{29}$$

It can be easily shown that $\mathcal{N}(\mathbf{D})$ is $\bar{\mathbf{A}}$-invariant (if the existence condition of Lemma 2 is satisfied) and also we can see that the column space of $\bar{\mathbf{B}}$ lies in $\mathcal{N}(\mathbf{D})$. Therefore, the new system representation (28) corresponds to a physically constrained system. With the new input $\bar{\mathbf{u}}_k$, we can reconstruct the performance index (5) as

$$J = \mathbf{x}_N^T\mathbf{Q}_N\mathbf{x}_N + \sum_{k=0}^{N-1}\Big[\mathbf{x}_k^T\bar{\mathbf{Q}}_c\mathbf{x}_k + 2\mathbf{x}_k^T\bar{\mathbf{H}}_c\bar{\mathbf{u}}_k + \bar{\mathbf{u}}_k^T\bar{\mathbf{R}}_c\bar{\mathbf{u}}_k\Big] \tag{30}$$

where

$$\bar{\mathbf{Q}}_c \triangleq \mathbf{Q}_c + \mathbf{G}_0^T\mathbf{R}_c\mathbf{G}_0,\ \bar{\mathbf{H}}_c \triangleq -\mathbf{G}_0^T\mathbf{R}_c\mathbf{P}_{\mathcal{N}(\mathbf{DB})},\ \bar{\mathbf{R}}_c \triangleq \mathbf{P}_{\mathcal{N}(\mathbf{DB})}\mathbf{R}_c\mathbf{P}_{\mathcal{N}(\mathbf{DB})}. \tag{31}$$

Note that the new input weighting matrix $\bar{\mathbf{R}}_c$ becomes non-negative definite (singular) for a given positive definite original weighting matrix $\mathbf{R}_c$ due to the projector $\mathbf{P}_{\mathcal{N}(\mathbf{DB})}$ onto the null space of $(\mathbf{DB})$. Therefore, in order to find the optimal feedback gain $\mathbf{G}^*$ such that $\bar{\mathbf{u}}_k^* = -\mathbf{G}^*\mathbf{x}_k$ minimizes J of (30), we have to solve a singular optimal control problem. However, for any discrete time linear-quadratic optimal control problem, regardless of the singularity or otherwise of any matrices in the cost, the associated RDE is well defined and may be solved in a straightforward manner to yield a solution of the optimal control problem (Clements & Anderson 1978), which is summarized by the following definition and lemma:

Definition 2 *The set of admissible weighting matrices, denoted by $\mathcal{S}$, is the set of $n \times n$ symmetric matrices $\mathbf{P}$ such that $\bar{\mathbf{B}}^T\mathbf{P}\bar{\mathbf{B}} + \bar{\mathbf{R}}_c \geq \mathbf{0}$ and $\mathcal{N}(\bar{\mathbf{B}}^T\mathbf{P}\bar{\mathbf{B}} + \bar{\mathbf{R}}_c) \subset \mathcal{N}(\bar{\mathbf{A}}^T\mathbf{P}\bar{\mathbf{B}} + \bar{\mathbf{H}}_c)$.*

Lemma 3 (Clements & Anderson (1978), Theorem II.2.1) *The optimal control problem for the system* (28) *with the performance index* (30) *has a solution on $[0, N]$ for a terminal weighting matrix $\mathbf{Q}_N$ if and only if the $n \times n$ symmetric matrix $\mathbf{P}_{k+1} \in \mathcal{S}$ for each $k = 0, \cdots, N-1$, where $\mathbf{P}_k$ is defined by the recursion, with $\mathbf{P}_N = \mathbf{Q}_N$,*

$$\mathbf{P}_k = \bar{\mathbf{A}}^T\mathbf{P}_{k+1}\bar{\mathbf{A}} + \bar{\mathbf{Q}}_c - \left(\bar{\mathbf{A}}^T\mathbf{P}_{k+1}\bar{\mathbf{B}} + \bar{\mathbf{H}}_c\right)\left(\bar{\mathbf{B}}^T\mathbf{P}_{k+1}\bar{\mathbf{B}} + \bar{\mathbf{R}}_c\right)^{\dagger}\left(\bar{\mathbf{A}}^T\mathbf{P}_{k+1}\bar{\mathbf{B}} + \bar{\mathbf{H}}_c\right)^T. \tag{32}$$

If $\mathbf{P}_k$ is so defined, then the control sequence $\mathbf{U}_l^{*N-1}$ defined by

$$\mathbf{u}_k^* = -\left(\bar{\mathbf{B}}^T\mathbf{P}_{k+1}\bar{\mathbf{B}} + \bar{\mathbf{R}}_c\right)^{\dagger}\left(\bar{\mathbf{A}}^T\mathbf{P}_{k+1}\bar{\mathbf{B}} + \bar{\mathbf{H}}_c\right)^T\mathbf{x}_k \tag{33}$$

achieves the infimum of J defined in (30) *for each $k = l, \cdots, N-1$. That is, with $\mathbf{U}_l^{*N-1} = \left[\mathbf{u}_l^*, \cdots, \mathbf{u}_{N-1}^*\right]$ we have*

$$J_{N-l}^*\left(\mathbf{x}_l, \mathbf{Q}_N\right) = J_{N-l}\left(\mathbf{x}_l, \mathbf{U}_l^{*N-1}, \mathbf{Q}_N\right) = \mathbf{x}_l^T\mathbf{P}_l\mathbf{x}_l. \tag{34}$$

It can be easily verified that for the minimization problem with the new performance index shown in (30) and the system matrices (29), any $n \times n$ symmetric matrix $\mathbf{P}$ is in the admissible set $\mathcal{S}$ and hence we can apply Lemma 3 and obtain the following Theorem 2.

Theorem 2 (Deterministic Constrained LQ Optimal Control) *For the system* (2) *and the state equality constraints* (3) *which satisfy one of the existence conditions in Lemma 2, and the performance index* (5)*, the solution to the LQ optimal control problem is given by the optimal control law*

$$\begin{aligned} \mathbf{u}_k^c &= -\mathbf{G}_0\ \mathbf{x}_k + \mathbf{P}_{\mathcal{N}(\mathbf{DB})}\bar{\mathbf{u}}_k^c \\ &= -\mathbf{K}_k^c \mathbf{x}_k \end{aligned} \tag{35}$$

where

$$\begin{aligned} \hat{\mathbf{R}}_k &\triangleq \mathbf{B}^T\mathbf{P}_{k+1}^c\mathbf{B} + \mathbf{R}_c \\ \hat{\mathbf{R}}_k^{(2)} &\triangleq \mathbf{P}_{\mathcal{N}(\mathbf{DB})}\left[\mathbf{P}_{\mathcal{N}(\mathbf{DB})}\hat{\mathbf{R}}_k\mathbf{P}_{\mathcal{N}(\mathbf{DB})}\right]^\dagger\mathbf{P}_{\mathcal{N}(\mathbf{DB})} = \left[\mathbf{P}_{\mathcal{N}(\mathbf{DB})}\hat{\mathbf{R}}_k\mathbf{P}_{\mathcal{N}(\mathbf{DB})}\right]^\dagger \\ \bar{\mathbf{u}}_k^c &\triangleq -\hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k\big(\bar{\mathbf{K}}_k^c - \mathbf{G}_0\big)\mathbf{x}_k \\ \bar{\mathbf{K}}_k^c &\triangleq \left(\mathbf{B}^T\mathbf{P}_{k+1}^c\mathbf{B} + \mathbf{R}_c\right)^{-1}\mathbf{B}^T\mathbf{P}_{k+1}^c\mathbf{A} = \big(\hat{\mathbf{R}}_k\big)^{-1}\mathbf{B}^T\mathbf{P}_{k+1}^c\mathbf{A} \\ \mathbf{K}_k^c &\triangleq \mathbf{G}_0 + \hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k\big(\bar{\mathbf{K}}_k^c - \mathbf{G}_0\big) \\ \mathbf{G}_0 &\triangleq (\mathbf{DB})^\dagger\mathbf{DA}. \end{aligned} \tag{36}$$

Here, $\mathbf{P}_{k+1}^c$ *satisfies, with the initial condition* $\mathbf{P}_N^c = \mathbf{Q}_N$*,*

$$\begin{aligned} \mathbf{P}_k^c &= \bar{\mathbf{A}}^T\mathbf{P}_{k+1}^c\bar{\mathbf{A}} + \bar{\mathbf{Q}}_c - \left(\bar{\mathbf{A}}^T\mathbf{P}_{k+1}^c\bar{\mathbf{B}} + \bar{\mathbf{H}}_c\right)\left(\bar{\mathbf{B}}^T\mathbf{P}_{k+1}^c\bar{\mathbf{B}} + \bar{\mathbf{R}}_c\right)^\dagger\left(\bar{\mathbf{A}}^T\mathbf{P}_{k+1}^c\bar{\mathbf{B}} + \bar{\mathbf{H}}_c\right)^T && (37) \\ &= \mathbf{A}^T\mathbf{P}_{k+1}^c\mathbf{A} + \mathbf{Q}_c + \boldsymbol{\nabla}_k^c - \mathbf{A}^T\mathbf{P}_{k+1}^c\mathbf{B}\left(\mathbf{B}^T\mathbf{P}_{k+1}^c\mathbf{B} + \mathbf{R}_c\right)^{-1}\mathbf{B}^T\mathbf{P}_{k+1}^c\mathbf{A}^T && (38) \end{aligned}$$

where

$$\boldsymbol{\nabla}_k^c \triangleq \left(\bar{\mathbf{K}}_k^c - \mathbf{G}_0\right)^T\boldsymbol{\Omega}_k^c\left(\bar{\mathbf{K}}_k^c - \mathbf{G}_0\right), \tag{39}$$

with

$$\boldsymbol{\Omega}_k^c \triangleq \hat{\mathbf{R}}_k - \hat{\mathbf{R}}_k\hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k. \tag{40}$$

The optimal value of the performance index is given by

$$J^c = \mathbf{x}_0^T\mathbf{P}_0^c\mathbf{x}_0. \tag{41}$$

The following lemma is useful for comparing relative magnitude of the solutions of the Riccati Difference Equations.

Lemma 4 (De Souza (1989), Bitmead & Gevers (1991)) *Consider two Riccati Difference Equations with the same* $\mathbf{A}$, $\mathbf{B}$ *and* $\mathbf{R}$ *matrices but possibly different* $\mathbf{Q}^1$ *and* $\mathbf{Q}^2$*. Denote their solution matrices* $\mathbf{P}_k^1$ *and* $\mathbf{P}_k^2$ *respectively. Suppose that* $\mathbf{Q}^1 \geq \mathbf{Q}^2$*, and, for some* k *we have* $\mathbf{P}_k^1 \geq \mathbf{P}_k^2$*, then for all* $i > 0$

$$\mathbf{P}_{k+i}^1 \geq \mathbf{P}_{k+i}^2. \tag{42}$$

In (36), $\hat{\mathbf{R}}_k^{(2)}$ represents a (2)-inverse (Campbell & C. D. Meyer 1979) of $\hat{\mathbf{R}}_k$ which satisfies the second condition of the Moore-Penrose inverse and it can be shown (see Appendix B) that

$$\hat{\mathbf{R}}_k \geq \hat{\mathbf{R}}_k\hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k, \tag{43}$$

and, therefore, from $\mathbf{\Omega}_k^c \geq \mathbf{0}$, we have $\mathbf{\nabla}_k^c \geq \mathbf{0}$ in (39). Hence, comparing (38) with the RDE (19) of the unconstrained LQ optimal control yields

$$\mathbf{P}_k^u \leq \mathbf{P}_k^c, \text{ for all } k, \tag{44}$$

through Lemma 4. Therefore, we obtain

$$J^u = \mathbf{x}_0^T \mathbf{P}_0^u \mathbf{x}_0 \leq \mathbf{x}_0^T \mathbf{P}_0^c \mathbf{x}_0 = J^c. \tag{45}$$

This tells us that the optimal performance index of the constrained case is greater than that of the unconstrained case, due to the design constraints on state variables which are given by (3). This is a formal derivation of a self-evident property that constraining the admissible control set worsens performance. Now, by optimality of the constrained LQ optimal control, we have

$$J^c \leq J^p, \tag{46}$$

and therefore by combining (45) with (46), we have the following Theorem 3.

Theorem 3 *For the system* (2) *with the state equality constraint* (3) *satisfying the existence condition in Lemma 2, the ordering of the performance indices among the unconstrained, the projected, and the constrained LQ optimal control is*

$$J^u = \mathbf{x}_0^T \mathbf{P}_0^u \mathbf{x}_0 \leq J^c = \mathbf{x}_0^T \mathbf{P}_0^c \mathbf{x}_0 \leq J^p = \mathbf{x}_0^T \mathbf{P}_0^p \mathbf{x}_0 \tag{47}$$

where $\mathbf{P}_0^u$, $\mathbf{P}_0^c$, *and* $\mathbf{P}_0^p$ *are the solutions of the RDEs* (19), (38), *and* (23), *respectively.*

Remark 4 (Comparison between projected and constrained LQ optimal Control) Without using the notion of optimality, we can directly prove (46) by comparing the related RDEs. From (25) and (39), it can be shown that

$$\mathbf{\nabla}_k^p - \mathbf{\nabla}_k^c = (\mathbf{K}_k^p - \bar{\mathbf{K}}_k^p)^T \hat{\mathbf{R}}_k^p (\mathbf{K}_k^p - \bar{\mathbf{K}}_k^p) - (\bar{\mathbf{K}}_k^c - \mathbf{G}_0)^T (\hat{\mathbf{R}}_k - \hat{\mathbf{R}}_k \hat{\mathbf{R}}_k^{(2)} \hat{\mathbf{R}}_k)(\bar{\mathbf{K}}_k^c - \mathbf{G}_0). \tag{48}$$

Due to the complex structure of (48), here we show $\mathbf{\nabla}_k^p \geq \mathbf{\nabla}_k^c$ explicitly, only for $k = N-1$. Since $\mathbf{P}_N^c = \mathbf{P}_N^p = \mathbf{Q}_N$, we have $\mathbf{K}_{N-1}^u = \bar{\mathbf{K}}_{N-1}^c = \bar{\mathbf{K}}_{N-1}^p \triangleq \mathbf{K}_{N-1}$ and then it can be shown (See Appendix C) that

$$\mathbf{\nabla}_{N-1}^p - \mathbf{\nabla}_{N-1}^c = (\mathbf{K}_{N-1} - \mathbf{G}_0)^T \ \mathbf{P}_{\mathcal{R}(\mathbf{DB})^T} \hat{\mathbf{R}}_{N-1} \Big[\mathbf{P}_{\mathcal{N}(\mathbf{DB})} \hat{\mathbf{R}}_{N-1} \mathbf{P}_{\mathcal{N}(\mathbf{DB})} \Big]^\dagger \hat{\mathbf{R}}_{N-1} \mathbf{P}_{\mathcal{R}(\mathbf{DB})^T} (\mathbf{K}_{N-1} - \mathbf{G}_0) \geq \mathbf{0}, \tag{49}$$

which implies $\mathbf{P}_{N-1}^p \geq \mathbf{P}_{N-1}^c$ by the monotonicity property of the RDE.

Remark 5 (Penalty method) As a conventional approach to reducing the error of the constraint $\mathbf{D}\mathbf{x}_k = \mathbf{0}$, we augment the original state weighting matrix $\mathbf{Q}_c$ with $\alpha \mathbf{D}^T \mathbf{D}$ such that $\mathbf{Q}^a \triangleq \mathbf{Q}_c + \alpha \mathbf{D}^T \mathbf{D}$ where α is any positive real number, and then minimize the performance index (5) with $\mathbf{Q}^a$ in place of $\mathbf{Q}_c$. We will compare this method with the constrained and the projected LQ optimal controller through numerical simulations in Section 5.

4 Stochastic Constrained Optimal Control

Now we consider the optimal control of stochastic systems

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k \tag{50}$$

with state equality constraints represented by (3).

Completing the square as in Åström (1970), the following Lemma 5 is obtained which will be used for generalizing the result of the deterministic constrained LQ optimal control of Theorem 2. Here, note that the state equality constraint has not been used yet in obtaining Lemma 5.

Lemma 5 *Assume that the Riccati Difference Equation* (38) *with the initial condition* $\mathbf{P}_N^c = \mathbf{Q}_N$ *has a solution which is non-negative definite for* $k \in [0, N]$. *Let* $\mathbf{x}_k$ *be the solution of the stochastic difference equation* (50). *Then,*

$$\begin{aligned} J_{sto} &\triangleq \mathbf{x}_N^T \mathbf{Q}_N \mathbf{x}_N + \sum_{k=0}^{N-1} \left(\mathbf{x}_k^T \mathbf{Q}_c \mathbf{x}_k + \mathbf{u}_k^T \mathbf{R}_c \mathbf{u}_k \right) \\ &= \mathbf{x}_0^T \mathbf{P}_0^c \mathbf{x}_0 + \sum_{k=0}^{N-1} \left[\mathbf{w}_k^T \mathbf{P}_{k+1}^c (\mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k) + (\mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k)^T \mathbf{P}_{k+1}^c \mathbf{w}_k + \mathbf{w}_k^T \mathbf{P}_{k+1}^c \mathbf{w}_k + V_k \right] \end{aligned} \tag{51}$$

where

$$\begin{aligned} V_k &= \mathbf{u}_k^T \hat{\mathbf{R}}_k \mathbf{u}_k + (\mathbf{u}_k + \mathbf{K}_k^c \mathbf{x}_k)^T \hat{\mathbf{R}}_k \bar{\mathbf{K}}_k^c \mathbf{x}_k + \mathbf{x}_k^T \bar{\mathbf{K}}_k^{c^T} \hat{\mathbf{R}}_k (\mathbf{u}_k + \mathbf{K}_k^c \mathbf{x}_k) - \mathbf{x}_k^T \mathbf{K}_k^{c^T} \hat{\mathbf{R}}_k \mathbf{K}_k^c \mathbf{x}_k \\ &= \mathbf{x}_k^T \left[\mathbf{K}_k^{c^T} \hat{\mathbf{R}}_k \bar{\mathbf{K}}_k^c + \bar{\mathbf{K}}_k^{c^T} \hat{\mathbf{R}}_k \mathbf{K}_k^c - \mathbf{K}_k^{c^T} \hat{\mathbf{R}}_k \mathbf{K}_k^c \right] \mathbf{x}_k + 2\mathbf{x}_k^T \bar{\mathbf{K}}_k^{c^T} \hat{\mathbf{R}}_k \mathbf{u}_k + \mathbf{u}_k^T \hat{\mathbf{R}}_k \mathbf{u}_k \end{aligned} \tag{52}$$

and $\mathbf{K}_k^c$, $\hat{\mathbf{R}}_k$, $\hat{\mathbf{R}}_k^{(2)}$, $\mathbf{G}_0$, *and* $\bar{\mathbf{K}}_k^c$ *are defined in* (36).

PROOF. See Appendix D. □

Although we have already analyzed the deterministic LQ case in Section 3, we derive the result again here now using Lemma 5, since this procedure will be used for the incomplete state information case in Section 4.2.

4.1 Deterministic Case

For a deterministic system, $\mathbf{w}_k \equiv \mathbf{0}$. Thus, from (51) of Lemma 5, we have

$$J_{sto} = J = \mathbf{x}_N^T \mathbf{Q}_N \mathbf{x}_N + \sum_{k=0}^{N-1} \left[\mathbf{x}_k^T \mathbf{Q}_c \mathbf{x}_k + \mathbf{u}_k^T \mathbf{R}_c \mathbf{u}_k \right] = \mathbf{x}_0^T \mathbf{P}_0^c \mathbf{x}_0 + \sum_{k=0}^{N-1} V_k. \tag{53}$$

Now, in order to incorporate the constraint (3) into the process of minimizing J given by (53), let us express J in terms of $\bar{\mathbf{u}}_k$ by using the $\mathcal{X}_c$-constrained feedback input form (27). Then, we obtain

$$J(\bar{\mathbf{u}}_k) = \mathbf{x}_0^T \mathbf{P}_0^c \mathbf{x}_0 + \sum_{k=0}^{N-1} \bar{V}_k(\mathbf{x}_k, \bar{\mathbf{u}}_k) \tag{54}$$

where

$$\begin{aligned} \bar{V}_k(\mathbf{x}_k, \bar{\mathbf{u}}_k) &\triangleq \mathbf{x}_k^T \left[\mathbf{K}_k^{c^T} \hat{\mathbf{R}}_k \bar{\mathbf{K}}_k^c + \bar{\mathbf{K}}_k^{c^T} \hat{\mathbf{R}}_k \mathbf{K}_k^c - \mathbf{K}_k^{c^T} \hat{\mathbf{R}}_k \mathbf{K}_k^c \right] \mathbf{x}_k + 2\mathbf{x}_k^T \bar{\mathbf{K}}_k^{c^T} \hat{\mathbf{R}}_k \left[-\mathbf{G}_0 \mathbf{x}_k + \mathbf{P}_{\mathcal{N}(\mathbf{DB})} \bar{\mathbf{u}}_k \right] \\ &\quad + \left[-\mathbf{G}_0 \mathbf{x}_k + \mathbf{P}_{\mathcal{N}(\mathbf{DB})} \bar{\mathbf{u}}_k \right]^T \hat{\mathbf{R}}_k \left[-\mathbf{G}_0 \mathbf{x}_k + \mathbf{P}_{\mathcal{N}(\mathbf{DB})} \bar{\mathbf{u}}_k \right] \\ &= \mathbf{x}_k^T \tilde{\mathbf{F}} \mathbf{x}_k + 2\mathbf{x}_k^T \tilde{\mathbf{H}} \bar{\mathbf{u}}_k + \bar{\mathbf{u}}_k^T \tilde{\mathbf{G}} \bar{\mathbf{u}}_k \end{aligned} \tag{55}$$

with

$$\begin{aligned} \tilde{\mathbf{F}} &\triangleq \mathbf{G}_0^T \hat{\mathbf{R}}_k \mathbf{G}_0 + \mathbf{K}_k^{c^T} \hat{\mathbf{R}}_k \bar{\mathbf{K}}_k^c + \bar{\mathbf{K}}_k^{c^T} \hat{\mathbf{R}}_k \mathbf{K}_k^c - \mathbf{K}_k^{c^T} \hat{\mathbf{R}}_k \mathbf{K}_k^c - \mathbf{G}_0^T \hat{\mathbf{R}}_k \bar{\mathbf{K}}_k^c - \bar{\mathbf{K}}_k^{c^T} \hat{\mathbf{R}}_k \mathbf{G}_0 \\ \tilde{\mathbf{G}} &\triangleq \mathbf{P}_{\mathcal{N}(\mathbf{DB})} \hat{\mathbf{R}}_k \mathbf{P}_{\mathcal{N}(\mathbf{DB})} \\ \tilde{\mathbf{H}} &\triangleq \left(\bar{\mathbf{K}}_k^c - \mathbf{G}_0 \right)^T \hat{\mathbf{R}}_k \mathbf{P}_{\mathcal{N}(\mathbf{DB})}. \end{aligned} \tag{56}$$

For minimizing $J(\bar{\mathbf{u}}_k)$ or $\bar{V}_k(\mathbf{x}_k, \bar{\mathbf{u}}_k)$ with respect to $\bar{\mathbf{u}}_k$, we use the following Lemma 6.

Lemma 6 (Clements & Anderson (1978), Lemma IV.2.1) *Consider the quadratic form* $q(\mathbf{z},\mathbf{v}) = \mathbf{z}^T\tilde{\mathbf{F}}\mathbf{z} + 2\mathbf{z}^T\tilde{\mathbf{H}}\mathbf{v} + \mathbf{v}^T\tilde{\mathbf{G}}\mathbf{v}$ *for matrices* $\tilde{\mathbf{F}} = \tilde{\mathbf{F}}^T$, $\tilde{\mathbf{G}} = \tilde{\mathbf{G}}^T$, $\tilde{\mathbf{H}}$, *and vectors* $\mathbf{z}$ *and* $\mathbf{v}$ *of arbitrary but consistent dimensions, and define* $q^*(\mathbf{z}) = \inf_{\mathbf{v}} q(\mathbf{z},\mathbf{v})$. *The following three conditions are equivalent:*

(i) $q^*(\mathbf{z}) > -\infty$ *for each* $\mathbf{z}$
(ii) $\tilde{\mathbf{G}} \geq \mathbf{0}$, $\mathcal{N}(\tilde{\mathbf{G}}) \subset \mathcal{N}(\tilde{\mathbf{H}})$
(iii) *there exists a symmetric matrix* $\mathbf{X}$ *such that*

$$\begin{bmatrix} \tilde{\mathbf{F}} - \mathbf{X} & \tilde{\mathbf{H}} \\ \tilde{\mathbf{H}}^T & \tilde{\mathbf{G}} \end{bmatrix} \geq \mathbf{0}. \tag{57}$$

Moreover, if any one of the above conditions holds, then (iii) *is satisfied by* $\mathbf{X}^* = \tilde{\mathbf{F}} - \tilde{\mathbf{H}}\tilde{\mathbf{G}}^\dagger\tilde{\mathbf{H}}^T$. *In addition,* $\mathbf{X}^* \geq \mathbf{X}$ *for any other* $\mathbf{X}$ *satisfying* (iii). *Finally, if, for each* $\mathbf{y}$, *we set*

$$\mathbf{v}^* = -\tilde{\mathbf{G}}^\dagger\tilde{\mathbf{H}}^T\mathbf{z}, \tag{58}$$

then

$$q^*(\mathbf{z}) = q(\mathbf{z},\mathbf{v}^*) = \mathbf{z}^T\mathbf{X}^*\mathbf{z}. \tag{59}$$

It is easily verified that $\tilde{\mathbf{G}} \geq \mathbf{0}$, $\mathcal{N}(\tilde{\mathbf{G}}) \subset \mathcal{N}(\tilde{\mathbf{H}})$ and hence Lemma 6 can be used for minimizing $\bar{V}_k(\mathbf{x}_k, \bar{\mathbf{u}}_k)$. From (59) together with (56) and (36), we arrive at

$$\bar{V}_k^*(\mathbf{x}_k, \bar{\mathbf{u}}_k^*) \triangleq \min_{\bar{\mathbf{u}}_k} \bar{V}_k(\mathbf{x}_k, \bar{\mathbf{u}}_k) = \mathbf{x}_k^T\mathbf{X}^*\mathbf{x}_k = 0, \tag{60}$$

since $\mathbf{X}^* = \tilde{\mathbf{F}} - \tilde{\mathbf{H}}\tilde{\mathbf{G}}^\dagger\tilde{\mathbf{H}}^T = \mathbf{0}$, for which (36) was used. Using (58), the optimal control $\bar{\mathbf{u}}_k^*$ is given by

$$\begin{aligned} \bar{\mathbf{u}}_k^* &= -\tilde{\mathbf{G}}^\dagger\tilde{\mathbf{H}}^T\mathbf{x}_k \\ &= -\Big[\mathbf{P}_{\mathcal{N}(\mathbf{DB})}\hat{\mathbf{R}}_k\mathbf{P}_{\mathcal{N}(\mathbf{DB})}\Big]^\dagger \mathbf{P}_{\mathcal{N}(\mathbf{DB})}\hat{\mathbf{R}}_k(\bar{\mathbf{K}}_k^c - \mathbf{G}_0)\mathbf{x}_k. \end{aligned} \tag{61}$$

Hence,

$$\begin{aligned} \mathbf{u}_k^* &= -\mathbf{G}_0\mathbf{x}_k + \mathbf{P}_{\mathcal{N}(\mathbf{DB})}\bar{\mathbf{u}}_k^* \\ &= -\Big[\mathbf{G}_0 + \hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k(\bar{\mathbf{K}}_k^c - \mathbf{G}_0)\Big]\mathbf{x}_k \\ &= -\mathbf{K}_k^c\mathbf{x}_k. \end{aligned} \tag{62}$$

By substituting (60) into (54), we obtain the same results as Theorem 2.

4.2 Incomplete State Information

The system we now consider is driven also by the process noise $\mathbf{w}_k$ as shown in (50) and the state information is available only from the measurement

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{v}_k \tag{63}$$

where $\mathbf{v}_k$ is gaussian with zero-mean and covariance $\mathbf{R}_e$. Therefore, in this case, we cannot constrain the state $\mathbf{x}_k$ in $\mathcal{X}_c$ since the exact state information is not available for the $\mathcal{X}_c$-constrained feedback such as $\mathbf{u}_k = -\mathbf{K}_k\mathbf{x}_k$. If we use the Kalman predictor $\hat{\mathbf{x}}_k$ for the feedback

$$\mathbf{u}_k = -\mathbf{K}_k\hat{\mathbf{x}}_k, \tag{64}$$

we have $\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k - \mathbf{B}\mathbf{K}_k\hat{\mathbf{x}}_k + \mathbf{w}_k$ which is generally not in $\mathcal{X}_c$. However, taking expectation yields

$$\mathcal{E}\{\mathbf{x}_{k+1}\} = (\mathbf{A} - \mathbf{B}\mathbf{K}_k)\mathcal{E}\{\mathbf{x}_k\}, \tag{65}$$

resulting in $\mathcal{E}\{\mathbf{x}_{k+1}\} \in \mathcal{X}_c$ for any $\mathcal{E}\{\hat{\mathbf{x}}_k\} = \mathcal{E}\{\mathbf{x}_k\} \in \mathcal{X}_c$. Therefore, for the case of incomplete state information, we can constrain only the expected value of the state in the constraint subspace by using the $\mathcal{X}_c$-constrained feedback obtained for the corresponding deterministic system. In addition to this, as is proved in the following Theorem 4, the $\mathcal{X}_c$-constrained feedback control minimizes the expectation of the squared constraint error.

Theorem 4 *For the stochastic system* (50) *with the measurement equation* (63)*, the state estimate feedback given by* (64) *with the $\mathcal{X}_c$-constrained feedback gain $\mathbf{K}$ form*[3]

$$\mathbf{K} = \mathbf{G} + \Big[\mathbf{G}_0 - \mathbf{P}_{\mathcal{R}(\mathbf{DB})^T}\mathbf{G}\Big] = \mathbf{G}_0 + \mathbf{P}_{\mathcal{N}(\mathbf{DB})}\mathbf{G} \tag{66}$$

has the following properties. Here $\mathbf{G}$ is an arbitrary matrix with consistent dimension.

(i) *$\mathcal{E}\{\mathbf{x}_k\} \in \mathcal{X}_c$ for all k with $\mathcal{E}\{\mathbf{x}_0\} \in \mathcal{X}_c$.*
(ii) *Minimizes the expectation of the squared constraint error defined as*

$$e(\mathbf{L}_k) \triangleq \mathcal{E}\big\{\mathbf{x}_{k+1}^T\mathbf{D}^T\mathbf{D}\mathbf{x}_{k+1}\big|\mathbf{y}_{k-1}\big\} = \mathcal{E}\big\{\mathbf{tr}\big\|\mathbf{D}\mathbf{x}_{k+1}\big\|^2\big|\mathbf{y}_{k-1}\big\} \tag{67}$$

with the assumption that the covariance of the estimate of the Kalman predictor $\hat{\mathbf{X}}_k = \mathcal{E}\{\hat{\mathbf{x}}_k\hat{\mathbf{x}}_k^T|\mathbf{y}_{k-1}\}$ is non-singular. Here, $\mathbf{L}_k$ represents any state estimate feedback gain.

PROOF. See Appendix E.

Since the performance index J_{sto} of (51) is now a random variable, we cannot handle it directly and hence we minimize its expected value. We find

$$\mathcal{E}\Big\{J_{sto}\Big\} = \mathcal{E}\Bigg\{\mathbf{x}_N^T\mathbf{Q}_N\mathbf{x}_N + \sum_{k=0}^{N-1}\Big(\mathbf{x}_k^T\mathbf{Q}_c\mathbf{x}_k + \mathbf{u}_k^T\mathbf{R}_c\mathbf{u}_k\Big)\Bigg\} = \mathcal{E}\Bigg\{\mathbf{x}_0^T\mathbf{P}_0^c\mathbf{x}_0 + \sum_{k=0}^{N-1}\Big[\mathbf{w}_k^T\mathbf{P}_{k+1}^c\mathbf{w}_k + V_k\Big]\Bigg\} \tag{68}$$

where V_k is given by (52). Since in (68) only V_k depends on the input $\mathbf{u}_k$ and the state $\mathbf{x}_k$, we consider only V_k for finding the optimal feedback minimizing (68), which also provides the properties of Theorem 4. For this, we use the following lemmas.

Lemma 7 (Åström (1970), p261) *Assume that the function $f(\mathbf{u}, \mathbf{y}) = \mathcal{E}_{x|y}[J_{sto}(\mathbf{x}, \mathbf{y}, \mathbf{u})|\mathbf{y}]$ has a unique minimum with respect to $\mathbf{u} \in \mathcal{U}$ for all $\mathbf{y} \in \mathcal{Y}$. Let $\mathbf{u}^*(\mathbf{y})$ denote the value of $\mathbf{u}$ for which the minimum is achieved. Then,*

$$\min_{\mathbf{u}(\mathbf{y})} \mathcal{E}_{x,y}\Big\{J_{sto}(\mathbf{x}, \mathbf{y}, \mathbf{u})\Big\} = \mathcal{E}_{x,y}\Big\{J_{sto}(\mathbf{x}, \mathbf{y}, \mathbf{u}^*(\mathbf{y}))\Big\} = \mathcal{E}_y\Big\{\min_{\mathbf{u}} \mathcal{E}_{x|y}\Big[J_{sto}(\mathbf{x}, \mathbf{y}, \mathbf{u})\Big|\mathbf{y}\Big]\Big\}. \tag{69}$$

Lemma 8 (Åström (1970), p262) *Let $\mathbf{x}$ be normal with mean $\mathbf{m}$ and covariance $\mathbf{R}$. Then,*

$$\mathcal{E}\{\mathbf{x}^T\mathbf{S}\mathbf{x}\} = \mathbf{m}^T\mathbf{S}\mathbf{m} + \mathrm{tr}[\mathbf{S}\mathbf{R}]. \tag{70}$$

[3] The only difference between (66) and (12) is the absence of the projector $\mathbf{P}_{\mathcal{N}(\mathbf{D})}$.

Let us denote

$$
\begin{aligned}
f(\mathbf{u}_k, \mathbf{y}_{k-1}) &\triangleq \mathcal{E}_{\mathbf{x}|\mathcal{Y}}\{V_k|\mathbf{y}_{k-1}\} \\
&= \mathcal{E}_{\mathbf{x}|\mathcal{Y}}\Big\{\mathbf{u}_k^T\hat{\mathbf{R}}_k\mathbf{u}_k + (\mathbf{u}_k + \mathbf{K}_k^c\mathbf{x}_k)^T\hat{\mathbf{R}}_k\bar{\mathbf{K}}_k^c\mathbf{x}_k + \mathbf{x}_k^T\bar{\mathbf{K}}_k^{c^T}\hat{\mathbf{R}}_k(\mathbf{u}_k + \mathbf{K}_k^c\mathbf{x}_k) - \mathbf{x}_k^T\mathbf{K}_k^{c^T}\hat{\mathbf{R}}_k\mathbf{K}_k^c\mathbf{x}_k\Big|\mathbf{y}_{k-1}\Big\}.
\end{aligned}
$$

Then, by using Lemma 8, we have

$$
\begin{aligned}
f(\mathbf{u}_k, \mathbf{y}_{k-1}) = \mathbf{u}_k^T\hat{\mathbf{R}}_k\mathbf{u}_k &+ (\mathbf{u}_k + \mathbf{K}_k^c\hat{\mathbf{x}}_k)^T\hat{\mathbf{R}}_k\bar{\mathbf{K}}_k^c\hat{\mathbf{x}}_k + \hat{\mathbf{x}}_k^T\bar{\mathbf{K}}_k^{c^T}\hat{\mathbf{R}}_k(\mathbf{u}_k + \mathbf{K}_k^c\hat{\mathbf{x}}_k) - \hat{\mathbf{x}}_k^T\mathbf{K}_k^{c^T}\hat{\mathbf{R}}_k\mathbf{K}_k^c\hat{\mathbf{x}}_k \\
&+ \operatorname{tr}\Big[\big(\mathbf{K}_k^{c^T}\hat{\mathbf{R}}_k\bar{\mathbf{K}}_k^c + \bar{\mathbf{K}}_k^{c^T}\hat{\mathbf{R}}_k\mathbf{K}_k^c - \mathbf{K}_k^{c^T}\hat{\mathbf{R}}_k\mathbf{K}_k^c\big)\mathbf{\Sigma}_k\Big]
\end{aligned}
$$

where $\mathbf{\Sigma}_k \triangleq \mathcal{E}\{(\hat{\mathbf{x}}_k - \mathbf{x}_k)(\hat{\mathbf{x}}_k - \mathbf{x}_k)^T|\mathbf{y}_{k-1}\}$ satisfies the Kalman predictor RDE

$$
\mathbf{\Sigma}_{k+1} = \mathbf{A}\mathbf{\Sigma}_k\mathbf{A}^T + \mathbf{Q}_e - \mathbf{A}\mathbf{\Sigma}_k\mathbf{C}^T(\mathbf{C}\mathbf{\Sigma}_k\mathbf{C}^T + \mathbf{R}_e)^{-1}\mathbf{C}\mathbf{\Sigma}_k\mathbf{A}^T. \tag{71}
$$

To incorporate the constraint $\mathcal{E}\{\mathbf{x}_k\} \in \mathcal{X}_c$, we use the input of the form

$$
\mathbf{u}_k = -\mathbf{G}_0\hat{\mathbf{x}}_k + \mathbf{P}_{\mathcal{N}(\mathbf{DB})}\bar{\mathbf{u}}_k. \tag{72}
$$

Then, $f(\mathbf{u}_k, \mathbf{y}_{k-1})$ becomes a function of $\bar{\mathbf{u}}_k$

$$
\begin{aligned}
f(\mathbf{u}_k, \mathbf{y}_{k-1}) &= f(\bar{\mathbf{u}}_k, \mathbf{y}_{k-1}) \\
&= \tilde{V}(\bar{\mathbf{u}}_k, \mathbf{y}_{k-1}) + \operatorname{tr}\Big[\big(\mathbf{K}_k^{c^T}\hat{\mathbf{R}}_k\bar{\mathbf{K}}_k^c + \bar{\mathbf{K}}_k^{c^T}\hat{\mathbf{R}}_k\mathbf{K}_k^c - \mathbf{K}_k^{c^T}\hat{\mathbf{R}}_k\mathbf{K}_k^c\big)\mathbf{\Sigma}_k\Big]
\end{aligned}
$$

where

$$
\tilde{V}(\bar{\mathbf{u}}_k, \mathbf{y}_{k-1}) = \hat{\mathbf{x}}_k^T\tilde{\mathbf{F}}\hat{\mathbf{x}}_k + 2\hat{\mathbf{x}}_k^T\tilde{\mathbf{H}}\bar{\mathbf{u}}_k + \bar{\mathbf{u}}_k^T\tilde{\mathbf{G}}\bar{\mathbf{u}}_k \tag{73}
$$

with $\tilde{\mathbf{F}}, \tilde{\mathbf{H}}$ and $\tilde{\mathbf{G}}$ given by (56). Similarly to the deterministic case in Section 4.1, the minimum of $\tilde{V}(\bar{\mathbf{u}}_k, \mathbf{y}_{k-1})$ is zero, which is obtained by

$$
\bar{\mathbf{u}}_k^* = -\Big[\mathbf{P}_{\mathcal{N}(\mathbf{DB})}\hat{\mathbf{R}}_k\mathbf{P}_{\mathcal{N}(\mathbf{DB})}\Big]^{\dagger}\mathbf{P}_{\mathcal{N}(\mathbf{DB})}\hat{\mathbf{R}}_k(\bar{\mathbf{K}}_k^c - \mathbf{G}_0)\hat{\mathbf{x}}_k. \tag{74}
$$

Therefore, by applying Lemma 7 and with (68) and

$$
\begin{aligned}
\mathcal{E}\Big\{\mathbf{x}_0^T\mathbf{P}_0^c\mathbf{x}_0\Big\} &= \bar{\mathbf{x}}_0^T\mathbf{P}_0^c\bar{\mathbf{x}}_0 + \operatorname{tr}\Big[\mathbf{P}_0^c\bar{\mathbf{\Sigma}}_0\Big] \\
\mathcal{E}\Big\{\mathbf{w}_k^T\mathbf{P}_{k+1}^c\mathbf{w}_k\Big\} &= \operatorname{tr}\Big[\mathbf{P}_{k+1}^c\mathbf{Q}_e\Big],
\end{aligned} \tag{75}
$$

we obtain Theorem 5. In (75), $\mathbf{x}_0$ has a gaussian distribution with mean $\bar{\mathbf{x}}_0$ and covariance $\bar{\mathbf{\Sigma}}_0$.

Theorem 5 *Consider the state* (50) *and measurement equation* (63). *Let the admissible control strategies be such that* $\mathbf{u}_k$ *is a function of* $\mathbf{y}_k$. *Assume that* $\mathbf{P}_k^c$*-Riccati equation* (38) *with initial condition* $\mathbf{P}_N^c = \mathbf{Q}_N$ *has a solution* $\mathbf{P}_k^c$ *such that* $\mathbf{P}_k^c$ *is symmetric with* $\mathbf{P}_k^c \in \mathcal{S}$. *Then there exists a unique admissible control strategy*

$$
\mathbf{u}_k = -\mathbf{K}_k^c\hat{\mathbf{x}}_k, \tag{76}
$$

which minimizes the expected performance index (68), *satisfying the equality constraints* $\mathcal{E}\{\mathbf{D}\mathbf{x}_k\} = \mathbf{0}$ *and also minimizing the expectation of the squared constraint error given by* (67). *Here,* $\mathbf{K}_k^c$ *is given by* (36). *The minimal* *value of expected performance index is given by*

$$\mathcal{E}\{J_{sto}^c\} \triangleq \min_{\mathcal{E}\{\mathbf{x}_k\}\in\mathcal{N}(\mathbf{D})} \mathcal{E}\{J\}$$
$$= \bar{\mathbf{x}}_0^T \mathbf{P}_0^c \bar{\mathbf{x}}_0 + \mathrm{tr}\Big[\mathbf{P}_0^c \bar{\mathbf{\Sigma}}_0\Big] + \sum_{k=0}^{N-1} \mathrm{tr}\Big[\mathbf{P}_{k+1}^c \mathbf{Q}_e + \big(\mathbf{K}_k^{c^T} \hat{\mathbf{R}}_k \bar{\mathbf{K}}_k^c + \bar{\mathbf{K}}_k^{c^T} \hat{\mathbf{R}}_k \mathbf{K}_k^c - \mathbf{K}_k^{c^T} \hat{\mathbf{R}}_k \mathbf{K}_k^c\big)\mathbf{\Sigma}_k\Big] \tag{77}$$

where $\mathbf{\Sigma}_k$ *satisfies* (71).

By following the similar approach used for deriving (77), we can obtain the following performance indices of the unconstrained and the projected LQ optimal control for the stochastic system

$$\mathcal{E}\{J_{sto}^u\} = \bar{\mathbf{x}}_0^T \mathbf{P}_0^u \bar{\mathbf{x}}_0 + \mathrm{tr}\Big[\mathbf{P}_0^u \bar{\mathbf{\Sigma}}_0\Big] + \sum_{k=0}^{N-1} \mathrm{tr}\Big[\mathbf{P}_{k+1}^u \mathbf{Q}_e + \big(\mathbf{K}_k^{u^T} \hat{\mathbf{R}}_k^u \mathbf{K}_k^u\big)\mathbf{\Sigma}_k\Big]$$
$$\mathcal{E}\{J_{sto}^p\} = \bar{\mathbf{x}}_0^T \mathbf{P}_0^p \bar{\mathbf{x}}_0 + \mathrm{tr}\Big[\mathbf{P}_0^p \bar{\mathbf{\Sigma}}_0\Big] + \sum_{k=0}^{N-1} \mathrm{tr}\Big[\mathbf{P}_{k+1}^p \mathbf{Q}_e + \big(\mathbf{K}_k^{p^T} \hat{\mathbf{R}}_k^p \bar{\mathbf{K}}_k^p + \bar{\mathbf{K}}_k^{p^T} \hat{\mathbf{R}}_k^p \mathbf{K}_k^p - \mathbf{K}_k^{p^T} \hat{\mathbf{R}}_k^p \mathbf{K}_k^p\big)\mathbf{\Sigma}_k\Big] \tag{78}$$

where $\hat{\mathbf{R}}_k^u \triangleq \mathbf{B}^T \mathbf{P}_{k+1}^u \mathbf{B} + \mathbf{R}_c$. Using optimality yields the following stochastic version of Theorem 3.

Theorem 6 *For the stochastic system represented by* (50) *and* (63) *with the same state and control weighting matrices, the following performance ordering holds:*

$$\mathcal{E}\{J_{sto}^u\} \le \mathcal{E}\{J_{sto}^c\} \le \mathcal{E}\{J_{sto}^p\} \tag{79}$$

5 Numerical Examples

The example demonstrated here is the discretized dynamics (with 0.05 *sec* sampling time) of the biped locomotion used in Hemami & Wyman (1979). The motion is described by 6-dimensional state vector $\mathbf{x}_k$ whose the first three components describe the angles of the first leg, the trunk, and the second leg, respectively, and the last three components are the angular velocities of the first leg, the trunk, and the second leg, respectively. There are 3-dimensional control input $\mathbf{u}_k$ to the biped. The below equation describes the biped with both feet on the ground:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k \tag{80}$$

where

$$\mathbf{A} = \begin{bmatrix} 1.0243 & -0.0086 & -0.0027 & 0.0506 & -0.0002 & -0.0001 \\ -0.0421 & 1.0362 & 0.0047 & -0.0011 & 0.0509 & 0.0001 \\ 0.0337 & -0.0120 & 0.9792 & 0.0008 & -0.0003 & 0.0495 \\ 0.9706 & -0.3445 & -0.1087 & 1.0243 & -0.0086 & -0.0027 \\ -1.6850 & 1.4496 & 0.1887 & -0.0421 & 1.0362 & 0.0047 \\ 1.3482 & -0.4785 & -0.8323 & 0.0337 & -0.0120 & 0.9792 \end{bmatrix}$$
$$\mathbf{B} = \begin{bmatrix} 0.0072 & -0.0198 & 0.0226 \\ -0.0125 & 0.0653 & -0.0702 \\ 0.0100 & -0.0275 & 0.0943 \\ 0.2890 & -0.7906 & 0.9030 \\ -0.5017 & 2.6126 & -2.8078 \\ 0.4014 & -1.0982 & 3.7708 \end{bmatrix} \times 10^{-1}. \tag{81}$$

The state constraints are described by

$$\mathbf{D}\mathbf{x}_k = \begin{bmatrix} 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{bmatrix} \mathbf{x}_k = \mathbf{0}. \tag{82}$$

That is, the constraints are $\mathbf{x}_k(1) = \mathbf{x}_k(3)$ and $\mathbf{x}_k(4) = \mathbf{x}_k(6)$. To see if these constraints can be implemented, consider

$$\mathbf{D}\mathbf{B} = \begin{bmatrix} -0.0028 & 0.0076 & -0.0716 \\ -0.1124 & 0.3076 & -2.8678 \end{bmatrix} \times 10^{-1} \tag{83}$$

and $\mathrm{rank}(\mathbf{D}\mathbf{B}) = 1$. However, since $\mathbf{P}_{\mathcal{N}(\mathbf{D}\mathbf{B})^T}\mathbf{D}\mathbf{A}\mathbf{Z} = \mathbf{0}$, by the condition (iv) of Lemma 2, the constraint can be precisely implemented. The initial condition used here for both deterministic and stochastic cases is

$$\mathbf{x}_0 = \begin{bmatrix} 10 & -10 & 10 & 5 & 0 & 5 \end{bmatrix}^T. \tag{84}$$

5.1 Deterministic Case

For the performance index, the state and control weighting matrices used are

$$\mathbf{Q}_N = \mathbf{Q}_c = \mathbf{I}_6 \text{ and } \mathbf{R}_c = \mathbf{I}_3, \tag{85}$$

respectively. For the unconstrained LQ control with the penalty method as explained in Remark 5, the following augmented state weighting matrix is used in simulation for different values of $\alpha \geq 0$

$$\mathbf{Q}^a = \mathbf{Q}_c + \alpha \mathbf{D}^T\mathbf{D}. \tag{86}$$

However, for the projected and the constrained LQ optimal control, only $\mathbf{Q}_c$ is used for the state weighting matrix. Table 1 compares the optimal performance indices and the average of the squared constraint error of the unconstrained, the projected, and the constrained controller. Here, the average constraint error ϵ is defined as

$$\epsilon \triangleq \frac{1}{N}\sqrt{\sum_{k=0}^{N} \mathbf{x}_k^T \mathbf{D}^T \mathbf{D} \mathbf{x}_k}. \tag{87}$$

As shown in Table 1, Fig. 1 and Fig. 2, as the α increases, the unconstrained and the projected case becomes closer to the constrained case (with being increased J^u and decreased ϵ^u for the unconstrained case and with being decreased J^p and ϵ^p for the projected case.)

5.2 Stochastic Case

The dynamics is disturbed by a process white noise sequence $\mathbf{w}_k$ and the noisy state information is available through the the measurement equation (63) with

$$\mathbf{C} = \mathbf{I}_6. \tag{88}$$

The process and measurement noises are of gaussian distribution with zero mean and covariance of

$$\begin{aligned} \mathbf{Q}_e &= \mathrm{diag}\begin{bmatrix} 1.0, & 0.5, & 1.0, & 0.5, & 0.2, & 0.5 \end{bmatrix}^T \\ \mathbf{R}_e &= 2 \times \mathbf{I}_6, \end{aligned} \tag{89}$$

respectively.

Deterministic	Unconstrained with $\mathbf{Q}^a$ / Constrained and projected Control with $\mathbf{Q}_c$						
LQ Optimal	Penalty factor	$\alpha = 0$	$\alpha = 5$	$\alpha = 10$	$\alpha = 20$	$\alpha = 26.5$	$\alpha = 100$
$J^{u,c,p} \times 10^{-6}$	Unconstrained	1.0913	1.1110	1.1153	1.1184	1.1193	1.1218
	Constrained	1.1228					
	Projected	1.5206	1.1406	1.1290	1.1247	1.1240	1.1229
$\epsilon^{u,c,p}$	Unconstrained	1.3677	0.4597	0.2873	0.1661	0.1306	0.0386
	Constrained/Projected	0					

Table 1

Performance comparison of LQ optimal controllers

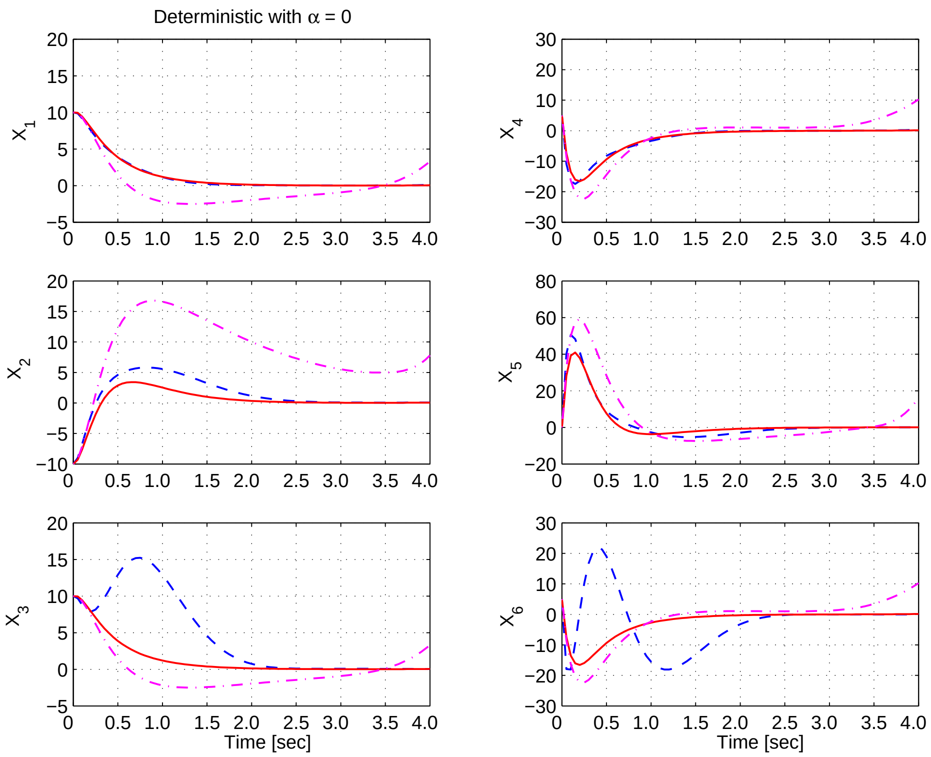

Fig. 1. Time response of a biped locomotion with constraints (deterministic case, dashed: unconstrained with penalty $\alpha = 0$, dash-dot: projected, solid: Constrained)

For the stochastic case, Monte Carlo simulation was performed and the result given in Table 2 was obtained through averaging 1000 simulation runs. Similarly to the deterministic case, as α increases, the expectation of performance index of the unconstrained control increases. Approximately at $\alpha = 26.5$, the expectation of performance indices of the unconstrained and the constrained controller become almost equal, but the average of the squared constraint error of the unconstrained case is much bigger than that of the constrained one (Table 2). Here, $\bar{\epsilon}$ represents the average value of ϵ defined in (87) over 1000 simulation runs. As proved in Theorem 4, we find that the projected and the constrained LQ control have the same constraint error since both have the same feedback gain structure as (66).

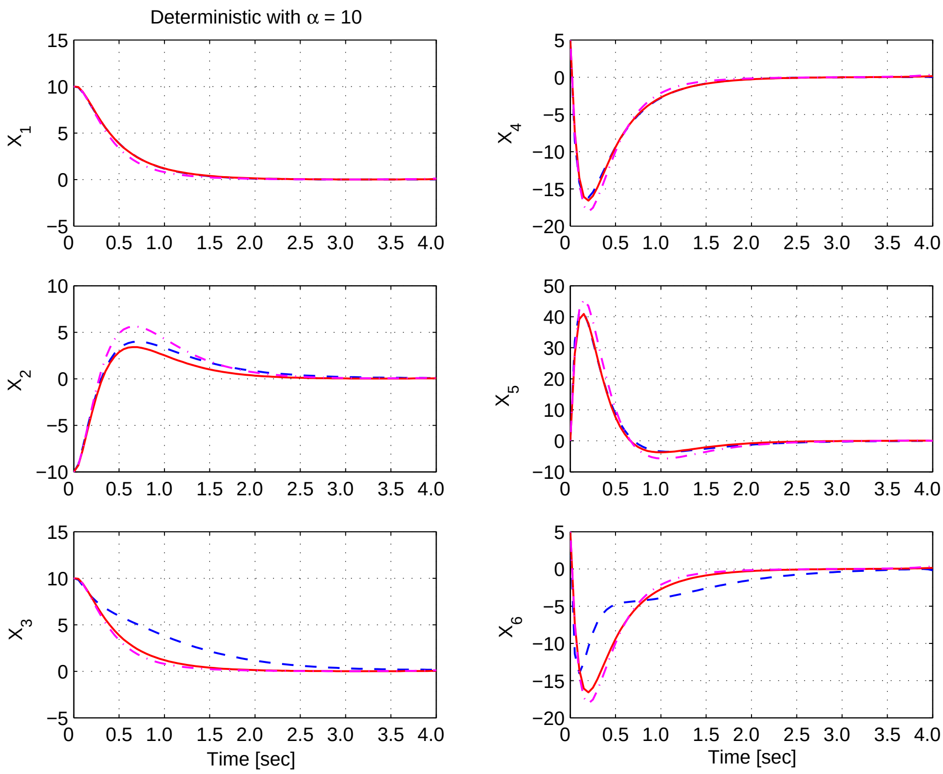

Fig. 2. Time response of a biped locomotion with constraints (deterministic case, dashed: unconstrained with penalty $\alpha = 10$, dash-dot: projected, solid: Constrained)

Stochastic	Unconstrained with $\mathbf{Q}^a$ / Constrained and projected Control with $\mathbf{Q}_c$						
LQ Optimal	Penalty factor	$\alpha = 0$	$\alpha = 5$	$\alpha = 10$	$\alpha = 20$	$\alpha = \mathbf{26.5}$	$\alpha = 100$
$\mathcal{E}\{J_{sto}^{u,c,p}\} \times 10^{-6}$	Unconstrained	2.6184	2.7672	2.8035	2.8554	**2.8823**	3.1330
	Constrained	**2.8845**					
	Projected	5.7759	3.1161	3.0056	2.9560	2.9450	2.9234
$\bar{\epsilon}^{u,c,p}$	Unconstrained	1.4880	0.4849	0.3109	0.1718	**0.1405**	0.0388
	Constrained	0.0472	0.0558	0.0205	0.0184	**0.0184**	0.0384
	Projected	0.0472	0.0558	0.0205	0.0184	0.0184	0.0384

Table 2
Performance comparison of stochastic LQ optimal controllers

6 Concluding Remarks

In this paper, the optimal control problem with state linear equality constraints was considered, first by finding the existence conditions for constraining linear feedback gains and then determining all such gains. By using the results of discrete time singular optimal control, the optimal constraining feedback gain was determined, which is also shown to minimize both the expectation of the performance index and the squared constraint error of the corresponding stochastic system. It is also confirmed that the constrained optimal performance index is increased, due to the constraint. The procedures used for discrete systems here can be similarly extended to the continuous-time case, which can be found in Ko (2005).

Acknowledgements

The research was supported by USA National Science Foundation Grants ECS-0200449 and ECS-0225530. The authors would like to thank Prof. Brian D.O. Anderson of the Australian National University for valuable discussion about this study.

Appendices

A Proof of Theorem 1

We have the identity

$$\mathbf{x}_N^T \mathbf{Q}_N \mathbf{x}_N = \mathbf{x}_N^T \mathbf{P}_N^p \mathbf{x}_N = \mathbf{x}_0^T \mathbf{P}_0^p \mathbf{x}_0 + \sum_{k=0}^{N-1} \underbrace{\left[\mathbf{x}_{k+1}^T \mathbf{P}_{k+1}^p \mathbf{x}_{k+1} - \mathbf{x}_k^T \mathbf{P}_k^p \mathbf{x}_k\right]}_{\triangleq \Delta_k^p} \tag{A.1}$$

Consider the different terms in the sum (A.1) and use (2) and (23). Then

$$\begin{aligned}
\mathbf{x}_{k+1}^T \mathbf{P}_{k+1}^p \mathbf{x}_{k+1} &= (\mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k^p)^T \mathbf{P}_{k+1}^p (\mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k^p) \\
&= \mathbf{x}_k^T \mathbf{A}^T \mathbf{P}_{k+1}^p \mathbf{A}\mathbf{x}_k + \mathbf{u}_k^{p^T} \mathbf{B}^T \mathbf{P}_{k+1}^p \mathbf{B}\mathbf{u}_k^p + \mathbf{x}_k^T \mathbf{A}^T \mathbf{P}_{k+1}^p \mathbf{B}\mathbf{u}_k^p + \mathbf{u}_k^{p^T} \mathbf{B}^T \mathbf{P}_{k+1}^p \mathbf{A}\mathbf{x}_k \\
\mathbf{x}_k^T \mathbf{P}_k^p \mathbf{x}_k &= \mathbf{x}_k^T \Big\{ \mathbf{A}^T \mathbf{P}_{k+1}^p \mathbf{A} + \mathbf{Q}_c - \mathbf{A}^T \mathbf{P}_{k+1}^p \mathbf{B} (\hat{\mathbf{R}}_k^p)^{-1} \mathbf{B}^T \mathbf{P}_{k+1}^p \mathbf{A} + (\mathbf{K}_k^p - \bar{\mathbf{K}}_k^p)^T \hat{\mathbf{R}}_k^p (\mathbf{K}_k^p - \bar{\mathbf{K}}_k^p) \Big\} \mathbf{x}_k.
\end{aligned} \tag{A.2}$$

Introducing (A.2) into Δ_k^p defined in (A.1) yields

$$\begin{aligned}
\Delta_k^p = &-\mathbf{x}_k^T \mathbf{Q}_c \mathbf{x}_k - \mathbf{u}_k^{p^T} \mathbf{R}_c \mathbf{u}_k^p + \mathbf{x}_k^T \Big\{ \mathbf{K}_k^{p^T} \hat{\mathbf{R}}_k^p \mathbf{K}_k^p - \mathbf{A}^T \mathbf{P}_{k+1}^p \mathbf{B}\mathbf{K}_k^p - \mathbf{K}_k^{p^T} \mathbf{B}^T \mathbf{P}_{k+1}^p \mathbf{A} \\
&+ \mathbf{A}^T \mathbf{P}_{k+1}^p \mathbf{B} (\hat{\mathbf{R}}_k^p)^{-1} \mathbf{B}^T \mathbf{P}_{k+1}^p \mathbf{A} - \mathbf{K}_k^{p^T} \hat{\mathbf{R}}_k^p \mathbf{K}_k^p - \bar{\mathbf{K}}_k^{p^T} \hat{\mathbf{R}}_k^p \bar{\mathbf{K}}_k^p + \mathbf{K}_k^{p^T} \hat{\mathbf{R}}_k^p \bar{\mathbf{K}}_k^p + \bar{\mathbf{K}}_k^{p^T} \hat{\mathbf{R}}_k^p \mathbf{K}_k^p \Big\} \mathbf{x}_k
\end{aligned} \tag{A.3}$$

where $\mathbf{u}_k^p = -\mathbf{K}_k^p \mathbf{x}_k$ was used. Now, by using $\bar{\mathbf{K}}_k^p = (\hat{\mathbf{R}}_k^p)^{-1} \mathbf{B}^T \mathbf{P}_{k+1}^p \mathbf{A}$, we obtain $\mathbf{A}^T \mathbf{P}_{k+1}^p \mathbf{B} = \bar{\mathbf{K}}_k^{p^T} \hat{\mathbf{R}}_k^p$ and $\mathbf{B}^T \mathbf{P}_{k+1}^p \mathbf{A} = \hat{\mathbf{R}}_k^p \bar{\mathbf{K}}_k^p$. Then, we have

$$\begin{aligned}
\Delta_k^p &= -\mathbf{x}_k^T \mathbf{Q}_c \mathbf{x}_k - \mathbf{u}_k^{p^T} \mathbf{R}_c \mathbf{u}_k^p \\
&\quad + \mathbf{x}_k^T \Big\{ \mathbf{K}_k^{p^T} \hat{\mathbf{R}}_k^p \mathbf{K}_k^p - \bar{\mathbf{K}}_k^{p^T} \hat{\mathbf{R}}_k^p \mathbf{K}_k^p - \mathbf{K}_k^{p^T} \hat{\mathbf{R}}_k^p \bar{\mathbf{K}}_k^p \\
&\quad + \bar{\mathbf{K}}_k^{p^T} \hat{\mathbf{R}}_k^p (\hat{\mathbf{R}}_k^p)^{-1} \hat{\mathbf{R}}_k^p \bar{\mathbf{K}}_k^p - \mathbf{K}_k^{p^T} \hat{\mathbf{R}}_k^p \mathbf{K}_k^p - \bar{\mathbf{K}}_k^{p^T} \hat{\mathbf{R}}_k^p \bar{\mathbf{K}}_k^p + \mathbf{K}_k^{p^T} \hat{\mathbf{R}}_k^p \bar{\mathbf{K}}_k^p + \bar{\mathbf{K}}_k^{p^T} \hat{\mathbf{R}}_k^p \mathbf{K}_k^p \Big\} \mathbf{x}_k \\
&= -\mathbf{x}_k^T \mathbf{Q}_c \mathbf{x}_k - \mathbf{u}_k^{p^T} \mathbf{R}_c \mathbf{u}_k^p.
\end{aligned} \tag{A.4}$$

Therefore, by substituting (A.4) into (A.1), we obtain the desired result. □

B Proof of (43)

For any positive definite matrix $\mathbf{X}$ and its (2)-inverse $\mathbf{X}^{(2)}$ which satisfies $\mathbf{X}^{(2)}\mathbf{X}\mathbf{X}^{(2)} = \mathbf{X}^{(2)}$,

$$\left(\mathbf{X}^{-1} - \mathbf{X}^{(2)}\right) \mathbf{X} \left(\mathbf{X}^{-1} - \mathbf{X}^{(2)}\right) \geq \mathbf{0}. \tag{B.1}$$

By expanding the left-hand side, we have

$$\mathbf{X}^{-1}\mathbf{X}\mathbf{X}^{-1} + \mathbf{X}^{(2)}\mathbf{X}\mathbf{X}^{(2)} - \mathbf{X}^{-1}\mathbf{X}\mathbf{X}^{(2)} - \mathbf{X}^{(2)}\mathbf{X}\mathbf{X}^{-1} \geq \mathbf{0}. \tag{B.2}$$

Hence, $\mathbf{X}^{-1} \geq \mathbf{X}^{(2)}$, from which, in turn, by multiplying $\mathbf{X}$ from both sides, we have $\mathbf{X} \geq \mathbf{X}\mathbf{X}^{(2)}\mathbf{X}$ and substituting $\mathbf{X} = \mathbf{B}^T\mathbf{P}^c_{k+1}\mathbf{B} + \mathbf{R}_c$ yields (43). $\square$

C Proof of (49) in Remark 4

Denote $\mho \triangleq \hat{\mathbf{R}}_{N\text{–}1}\Big[\mathbf{P}_{\mathcal{N}(\mathbf{DB})}\hat{\mathbf{R}}_{N\text{–}1}\mathbf{P}_{\mathcal{N}(\mathbf{DB})}\Big]^{\dagger}\hat{\mathbf{R}}_{N\text{–}1} + \mathbf{P}_{\mathcal{R}(\mathbf{DB})^T}\hat{\mathbf{R}}_{N\text{–}1}\mathbf{P}_{\mathcal{R}(\mathbf{DB})^T} - \hat{\mathbf{R}}_{N\text{–}1}$. With this notation, we have $\boldsymbol{\nabla}^p_{N\text{–}1} - \boldsymbol{\nabla}^c_{N\text{–}1} = (\mathbf{K}_{N-1} - \mathbf{G}_0)^T\mho(\mathbf{K}_{N-1} - \mathbf{G}_0)$. Observe that

$$\begin{aligned}\mathbf{P}_{\mathcal{N}(\mathbf{DB})}\mho &= \mathbf{P}_{\mathcal{N}(\mathbf{DB})}\hat{\mathbf{R}}_{N\text{–}1}\Big[\mathbf{P}_{\mathcal{N}(\mathbf{DB})}\hat{\mathbf{R}}_{N\text{–}1}\mathbf{P}_{\mathcal{N}(\mathbf{DB})}\Big]^{\dagger}\hat{\mathbf{R}}_{N\text{–}1} + \underbrace{\mathbf{P}_{\mathcal{N}(\mathbf{DB})}\mathbf{P}_{\mathcal{R}(\mathbf{DB})^T}}_{=\mathbf{0}}\hat{\mathbf{R}}_{N\text{–}1}\mathbf{P}_{\mathcal{R}(\mathbf{DB})^T} - \mathbf{P}_{\mathcal{N}(\mathbf{DB})}\hat{\mathbf{R}}_{N\text{–}1} \\ &= \underbrace{\big(\mathbf{P}_{\mathcal{N}(\mathbf{DB})}\hat{\mathbf{R}}_{N\text{–}1}\mathbf{P}_{\mathcal{N}(\mathbf{DB})}\big)\Big[\mathbf{P}_{\mathcal{N}(\mathbf{DB})}\hat{\mathbf{R}}_{N\text{–}1}\mathbf{P}_{\mathcal{N}(\mathbf{DB})}\Big]^{\dagger}}_{=\mathbf{P}_{\mathcal{N}(\mathbf{DB})}}\mathbf{P}_{\mathcal{N}(\mathbf{DB})}\hat{\mathbf{R}}_{N\text{–}1} - \mathbf{P}_{\mathcal{N}(\mathbf{DB})}\hat{\mathbf{R}}_{N\text{–}1} = \mathbf{0},\end{aligned} \tag{C.1}$$

and hence

$$\mho\,\mathbf{P}_{\mathcal{N}(\mathbf{DB})} = \mathbf{0}. \tag{C.2}$$

(C.1) and (C.2) tell us that both column and row spaces of $\mho$ are in the row space of $\mathbf{DB}$. Therefore,

$$\begin{aligned}\mho &= \mathbf{P}_{\mathcal{R}(\mathbf{DB})^T}\mho\mathbf{P}_{\mathcal{R}(\mathbf{DB})^T} \\ &= \mathbf{P}_{\mathcal{R}(\mathbf{DB})^T}\hat{\mathbf{R}}_{N\text{–}1}\Big[\mathbf{P}_{\mathcal{N}(\mathbf{DB})}\hat{\mathbf{R}}_{N\text{–}1}\mathbf{P}_{\mathcal{N}(\mathbf{DB})}\Big]^{\dagger}\hat{\mathbf{R}}_{N\text{–}1}\mathbf{P}_{\mathcal{R}(\mathbf{DB})^T} \\ &\quad + \mathbf{P}_{\mathcal{R}(\mathbf{DB})^T}\mathbf{P}_{\mathcal{R}(\mathbf{DB})^T}\hat{\mathbf{R}}_{N\text{–}1}\mathbf{P}_{\mathcal{R}(\mathbf{DB})^T}\mathbf{P}_{\mathcal{R}(\mathbf{DB})^T} - \mathbf{P}_{\mathcal{R}(\mathbf{DB})^T}\hat{\mathbf{R}}_{N\text{–}1}\mathbf{P}_{\mathcal{R}(\mathbf{DB})^T} \\ &= \mathbf{P}_{\mathcal{R}(\mathbf{DB})^T}\hat{\mathbf{R}}_{N\text{–}1}\Big[\mathbf{P}_{\mathcal{N}(\mathbf{DB})}\hat{\mathbf{R}}_{N\text{–}1}\mathbf{P}_{\mathcal{N}(\mathbf{DB})}\Big]^{\dagger}\hat{\mathbf{R}}_{N\text{–}1}\mathbf{P}_{\mathcal{R}(\mathbf{DB})^T}.\end{aligned} \tag{C.3}$$

Hence $\boldsymbol{\nabla}^p_{N\text{–}1} - \boldsymbol{\nabla}^c_{N\text{–}1} = (\mathbf{K}_{N-1} - \mathbf{G}_0)^T\mathbf{P}_{\mathcal{R}(\mathbf{DB})^T}\hat{\mathbf{R}}_{N\text{–}1}\Big[\mathbf{P}_{\mathcal{N}(\mathbf{DB})}\hat{\mathbf{R}}_{N\text{–}1}\mathbf{P}_{\mathcal{N}(\mathbf{DB})}\Big]^{\dagger}\hat{\mathbf{R}}_{N\text{–}1}\mathbf{P}_{\mathcal{R}(\mathbf{DB})^T}(\mathbf{K}_{N-1} - \mathbf{G}_0) \geq \mathbf{0}$. $\square$

D Proof of Lemma 5

We have the identity

$$\mathbf{x}_N^T\mathbf{Q}_N\mathbf{x}_N = \mathbf{x}_N^T\mathbf{P}^c_N\mathbf{x}_N = \mathbf{x}_0^T\mathbf{P}^c_0\mathbf{x}_0 + \sum_{k=0}^{N-1}\Big(\underbrace{\mathbf{x}_{k+1}^T\mathbf{P}^c_{k+1}\mathbf{x}_{k+1} - \mathbf{x}_k^T\mathbf{P}^c_k\mathbf{x}_k}_{\triangleq\boldsymbol{\Gamma}}\Big). \tag{D.1}$$

Consider the different terms in $\boldsymbol{\Gamma}$ and by using (50) and (38):

$$\mathbf{x}_{k+1}^T\mathbf{P}^c_{k+1}\mathbf{x}_{k+1} = (\mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k)^T\mathbf{P}^c_{k+1}(\mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k) \tag{D.2}$$

$$\mathbf{x}_k^T\mathbf{P}^c_k\mathbf{x}_k = \mathbf{x}_k^T\Big\{\mathbf{A}^T\mathbf{P}^c_{k+1}\mathbf{A} + \mathbf{Q}_c - \bar{\mathbf{K}}_k^{c^T}\hat{\mathbf{R}}_k\bar{\mathbf{K}}^c_k + \big(\bar{\mathbf{K}}^c_k - \mathbf{G}_0\big)^T\big(\hat{\mathbf{R}}_k - \hat{\mathbf{R}}_k\hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k\big)\big(\bar{\mathbf{K}}^c_k - \mathbf{G}_0\big)\Big\}\mathbf{x}_k \tag{D.3}$$

Then

$$\begin{aligned}\boldsymbol{\Gamma} &= \mathbf{w}_k^T\mathbf{P}^c_{k+1}(\mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k) + (\mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k)^T\mathbf{P}^c_{k+1}\mathbf{w}_k + \mathbf{w}_k^T\mathbf{P}^c_{k+1}\mathbf{w}_k \\ &\quad + \mathbf{x}_k^T\mathbf{A}^T\mathbf{P}^c_{k+1}\mathbf{A}\mathbf{x}_k + \mathbf{u}_k^T\mathbf{B}^T\mathbf{P}^c_{k+1}\mathbf{B}\mathbf{u}_k + \mathbf{x}_k^T\mathbf{A}^T\mathbf{P}^c_{k+1}\mathbf{B}\mathbf{u}_k + \mathbf{u}_k^T\mathbf{B}^T\mathbf{P}^c_{k+1}\mathbf{A}\mathbf{x}_k \\ &\quad + \mathbf{u}_k^T\mathbf{R}_c\mathbf{u}_k - \mathbf{u}_k^T\mathbf{R}_c\mathbf{u}_k - \mathbf{x}_k^T\mathbf{A}^T\mathbf{P}^c_{k+1}\mathbf{A}\mathbf{x}_k - \mathbf{x}_k^T\mathbf{Q}_c\mathbf{x}_k + \mathbf{x}_k^T\bar{\mathbf{K}}_k^{c^T}\hat{\mathbf{R}}_k\bar{\mathbf{K}}^c_k\mathbf{x}_k \\ &\quad - \mathbf{x}_k^T\big(\bar{\mathbf{K}}^c_k - \mathbf{G}_0\big)^T\big(\hat{\mathbf{R}}_k - \hat{\mathbf{R}}_k\hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k\big)\big(\bar{\mathbf{K}}^c_k - \mathbf{G}_0\big)\mathbf{x}_k \\ &= -\mathbf{x}_k^T\mathbf{Q}_c\mathbf{x}_k - \mathbf{u}_k^T\mathbf{R}_c\mathbf{u}_k + \mathbf{w}_k^T\mathbf{P}^c_{k+1}(\mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k) + (\mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k)^T\mathbf{P}^c_{k+1}\mathbf{w}_k + \mathbf{w}_k^T\mathbf{P}^c_{k+1}\mathbf{w}_k + \boldsymbol{\Gamma}_1\end{aligned} \tag{D.4}$$

where

$$\begin{aligned}\mathbf{\Gamma}_1 \triangleq\ & \mathbf{u}_k^T\hat{\mathbf{R}}_k\mathbf{u}_k + \mathbf{u}_k^T\hat{\mathbf{R}}_k\bar{\mathbf{K}}_k^c\mathbf{x}_k + \mathbf{x}_k^T\bar{\mathbf{K}}_k^{c^T}\hat{\mathbf{R}}_k\mathbf{u}_k + \mathbf{x}_k^T\bar{\mathbf{K}}_k^{c^T}\hat{\mathbf{R}}_k\bar{\mathbf{K}}_k^c\mathbf{x}_k \\ & - \mathbf{x}_k^T\left(\bar{\mathbf{K}}_k^c - \mathbf{G}_0\right)^T\hat{\mathbf{R}}_k\left(\bar{\mathbf{K}}_k^c - \mathbf{G}_0\right)\mathbf{x}_k + \mathbf{x}_k^T\left(\bar{\mathbf{K}}_k^c - \mathbf{G}_0\right)^T\hat{\mathbf{R}}_k\hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k\left(\bar{\mathbf{K}}_k^c - \mathbf{G}_0\right)\mathbf{x}_k\end{aligned} \tag{D.5}$$

and for the second equality in (D.4), $\mathbf{B}^T\mathbf{P}_{k+1}^c\mathbf{A} = \hat{\mathbf{R}}_k\bar{\mathbf{K}}_k^c$ was used. Now the last term of (D.5) is denoted by $\mathbf{\Gamma}_2$ and it can be rearranged as

$$\begin{aligned}\mathbf{\Gamma}_2 =\ & \mathbf{x}_k^T\left[\mathbf{G}_0 + \hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k\left(\bar{\mathbf{K}}_k^c - \mathbf{G}_0\right)\right]^T\hat{\mathbf{R}}_k\left[\mathbf{G}_0 + \hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k\left(\bar{\mathbf{K}}_k^c - \mathbf{G}_0\right)\right]\mathbf{x}_k \\ & - \mathbf{x}_k^T\mathbf{G}_0^T\hat{\mathbf{R}}_k\hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k\left(\bar{\mathbf{K}}_k^c - \mathbf{G}_0\right)\mathbf{x}_k - \mathbf{x}_k^T\left(\bar{\mathbf{K}}_k^c - \mathbf{G}_0\right)^T\hat{\mathbf{R}}_k\hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k\mathbf{G}_0\mathbf{x}_k \\ & - \mathbf{x}_k^T\mathbf{G}_0^T\hat{\mathbf{R}}_k\mathbf{G}_0\mathbf{x}_k,\end{aligned} \tag{D.6}$$

where, by the definition of (2)-inverse, $\hat{\mathbf{R}}_k\hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k\hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k = \hat{\mathbf{R}}_k\hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k$ was used.

By substituting (D.6) into (D.5), we have

$$\mathbf{\Gamma}_1 = \mathbf{u}_k^T\hat{\mathbf{R}}_k\mathbf{u}_k + \mathbf{x}_k^T\left[\mathbf{G}_0 + \hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k\left(\bar{\mathbf{K}}_k^c - \mathbf{G}_0\right)\right]^T\hat{\mathbf{R}}_k\left[\mathbf{G}_0 + \hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k\left(\bar{\mathbf{K}}_k^c - \mathbf{G}_0\right)\right]\mathbf{x}_k + \mathbf{\Gamma}_3 \tag{D.7}$$

where

$$\begin{aligned}\mathbf{\Gamma}_3 \triangleq\ & -\mathbf{x}_k^T\mathbf{G}_0^T\hat{\mathbf{R}}_k\hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k\left(\bar{\mathbf{K}}_k^c - \mathbf{G}_0\right)\mathbf{x}_k - \mathbf{x}_k^T\left(\bar{\mathbf{K}}_k^c - \mathbf{G}_0\right)^T\hat{\mathbf{R}}_k\hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k\mathbf{G}_0\mathbf{x}_k \\ & - \mathbf{x}_k^T\mathbf{G}_0^T\hat{\mathbf{R}}_k\mathbf{G}_0\mathbf{x}_k - \mathbf{x}_k^T\left(\bar{\mathbf{K}}_k^c - \mathbf{G}_0\right)^T\hat{\mathbf{R}}_k\left(\bar{\mathbf{K}}_k^c - \mathbf{G}_0\right)\mathbf{x}_k \\ & + \mathbf{u}_k^T\hat{\mathbf{R}}_k\bar{\mathbf{K}}_k^c\mathbf{x}_k + \mathbf{x}_k^T\bar{\mathbf{K}}_k^{c^T}\hat{\mathbf{R}}_k\mathbf{u}_k + \mathbf{x}_k^T\bar{\mathbf{K}}_k^{c^T}\hat{\mathbf{R}}_k\bar{\mathbf{K}}_k^c\mathbf{x}_k.\end{aligned} \tag{D.8}$$

Rearranging terms of $\mathbf{\Gamma}_3$ yields

$$\begin{aligned}\mathbf{\Gamma}_3 =\ & \mathbf{u}_k^T\hat{\mathbf{R}}_k\bar{\mathbf{K}}_k^c\mathbf{x}_k + \mathbf{x}_k^T\bar{\mathbf{K}}_k^{c^T}\hat{\mathbf{R}}_k\mathbf{u}_k \\ & + \mathbf{x}_k^T\left[\mathbf{G}_0 + \hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k(\bar{\mathbf{K}}_k^c - \mathbf{G}_0)\right]^T\hat{\mathbf{R}}_k\bar{\mathbf{K}}_k^c\mathbf{x}_k + \mathbf{x}_k^T\bar{\mathbf{K}}_k^{c^T}\hat{\mathbf{R}}_k\left[\mathbf{G}_0 + \hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k(\bar{\mathbf{K}}_k^c - \mathbf{G}_0)\right]\mathbf{x}_k \\ & - 2\mathbf{x}_k^T\bar{\mathbf{K}}_k^{c^T}\hat{\mathbf{R}}_k\hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k\bar{\mathbf{K}}_k^c\mathbf{x}_k - 2\mathbf{x}_k^T\mathbf{G}_0^T\hat{\mathbf{R}}_k\mathbf{G}_0\mathbf{x}_k + 2\mathbf{x}_k^T\mathbf{G}_0^T\hat{\mathbf{R}}_k\hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k\mathbf{G}_0\mathbf{x}_k\end{aligned} \tag{D.9}$$

where it can be shown that the sum of last three terms of $\mathbf{\Gamma}_3$ are equal to

$$\mathbf{\Gamma}_4 \triangleq -2\mathbf{x}_k^T\left[\mathbf{G}_0 + \hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k(\bar{\mathbf{K}}_k^c - \mathbf{G}_0)\right]^T\hat{\mathbf{R}}_k\left[\mathbf{G}_0 + \hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k(\bar{\mathbf{K}}_k^c - \mathbf{G}_0)\right]\mathbf{x}_k. \tag{D.10}$$

By substituting (D.10) into (D.9) and then from (D.7), we thus have

$$\begin{aligned}\mathbf{\Gamma}_1 =\ & \mathbf{u}_k^T\hat{\mathbf{R}}_k\mathbf{u}_k + \mathbf{x}_k^T\left[\mathbf{G}_0 + \hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k\left(\bar{\mathbf{K}}_k^c - \mathbf{G}_0\right)\right]^T\hat{\mathbf{R}}_k\left[\mathbf{G}_0 + \hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k\left(\bar{\mathbf{K}}_k^c - \mathbf{G}_0\right)\right]\mathbf{x}_k \\ & + \mathbf{u}_k^T\hat{\mathbf{R}}_k\bar{\mathbf{K}}_k^c\mathbf{x}_k + \mathbf{x}_k^T\bar{\mathbf{K}}_k^{c^T}\hat{\mathbf{R}}_k\mathbf{u}_k \\ & + \mathbf{x}_k^T\left[\mathbf{G}_0 + \hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k(\bar{\mathbf{K}}_k^c - \mathbf{G}_0)\right]^T\hat{\mathbf{R}}_k\bar{\mathbf{K}}_k^c\mathbf{x}_k + \mathbf{x}_k^T\bar{\mathbf{K}}_k^{c^T}\hat{\mathbf{R}}_k\left[\mathbf{G}_0 + \hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k(\bar{\mathbf{K}}_k^c - \mathbf{G}_0)\right]\mathbf{x}_k \\ & - 2\mathbf{x}_k^T\left[\mathbf{G}_0 + \hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k(\bar{\mathbf{K}}_k^c - \mathbf{G}_0)\right]^T\hat{\mathbf{R}}_k\left[\mathbf{G}_0 + \hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k(\bar{\mathbf{K}}_k^c - \mathbf{G}_0)\right]\mathbf{x}_k \\ =\ & \mathbf{u}_k^T\hat{\mathbf{R}}_k\mathbf{u}_k + (\mathbf{u}_k + \mathbf{K}_k^c\mathbf{u}_k)^T\hat{\mathbf{R}}_k\bar{\mathbf{K}}_k^c\mathbf{x}_k + \mathbf{x}_k^T\bar{\mathbf{K}}_k^{c^T}\hat{\mathbf{R}}_k(\mathbf{u}_k + \mathbf{K}_k^c\mathbf{u}_k) - \mathbf{x}_k^T\mathbf{K}_k^{c^T}\hat{\mathbf{R}}_k\mathbf{K}_k^c\mathbf{x}_k\end{aligned} \tag{D.11}$$

where for the second equality $\mathbf{K}_k^c = \mathbf{G}_0 + \hat{\mathbf{R}}_k^{(2)}\hat{\mathbf{R}}_k\left(\bar{\mathbf{K}}_k^c - \mathbf{G}_0\right)$ was used. From (D.4) and (D.11), we have

$$\begin{aligned}\mathbf{\Gamma} =\ & -\mathbf{x}_k^T\mathbf{Q}_c\mathbf{x}_k - \mathbf{u}_k^T\mathbf{R}_c\mathbf{u}_k + \mathbf{w}_k^T\mathbf{P}_{k+1}^c(\mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k) + (\mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k)^T\mathbf{P}_{k+1}^c\mathbf{w}_k + \mathbf{w}_k^T\mathbf{P}_{k+1}^c\mathbf{w}_k \\ & + \mathbf{u}_k^T\hat{\mathbf{R}}_k\mathbf{u}_k + (\mathbf{u}_k + \mathbf{K}_k^c\mathbf{u}_k)^T\hat{\mathbf{R}}_k\bar{\mathbf{K}}_k^c\mathbf{x}_k + \mathbf{x}_k^T\bar{\mathbf{K}}_k^{c^T}\hat{\mathbf{R}}_k(\mathbf{u}_k + \mathbf{K}_k^c\mathbf{u}_k) - \mathbf{x}_k^T\mathbf{K}_k^{c^T}\hat{\mathbf{R}}_k\mathbf{K}_k^c\mathbf{x}_k.\end{aligned} \tag{D.12}$$

Finally, by substituting (D.12) into (D.1), we get the desired result (51). □

E Proof of Theorem 4

We are to find the optimal $\mathbf{L}_k$ minimizing $e(\mathbf{L}_k)$ given by

$$e(\mathbf{L}_k) = \mathcal{E}\Big\{\mathbf{tr}\big\|\mathbf{D}\mathbf{x}_{k+1}\big\|^2\Big|\mathbf{y}_{k-1}\Big\} = \mathcal{E}\Big\{\mathbf{tr}\big\|\mathbf{D}\left[\mathbf{A}\mathbf{x}_k - \mathbf{B}\mathbf{L}_k\hat{\mathbf{x}}_{k|k-1} + \mathbf{w}_k\right]\big\|^2\Big|\mathbf{y}_{k-1}\Big\}. \tag{E.1}$$

Then

$$\begin{aligned} e(\mathbf{L}_k) &= \mathbf{tr}\ \mathcal{E}\Big\{\mathbf{D}(\mathbf{A}\mathbf{x}_k - \mathbf{B}\mathbf{L}_k\hat{\mathbf{x}}_{k|k-1} + \mathbf{w}_k)(\mathbf{A}\mathbf{x}_k - \mathbf{B}\mathbf{L}_k\hat{\mathbf{x}}_{k|k-1} + \mathbf{w}_k)^T\mathbf{D}^T\Big|\mathbf{y}_{k-1}\Big\} \\ &= \mathbf{tr}\Big[\mathbf{D}\mathbf{A}\mathbf{X}_k\mathbf{A}^T\mathbf{D}^T + \mathbf{D}\mathbf{B}\mathbf{L}_k\hat{\mathbf{X}}_k\mathbf{L}_k^T\mathbf{B}^T\mathbf{D}^T + \mathbf{D}\mathbf{Q}\mathbf{D}^T - \mathbf{D}\mathbf{A}\hat{\mathbf{X}}_k\mathbf{L}_k^T\mathbf{B}^T\mathbf{D}^T - \mathbf{D}\mathbf{B}\mathbf{L}_k\hat{\mathbf{X}}_k\mathbf{A}^T\mathbf{D}^T\Big] \end{aligned} \tag{E.2}$$

where

$$\mathbf{X}_k \triangleq \mathcal{E}\Big\{\mathbf{x}_k\mathbf{x}_k^T\Big|\mathbf{y}_{k-1}\Big\} \text{ and } \hat{\mathbf{X}}_k \triangleq \mathcal{E}\Big\{\hat{\mathbf{x}}_k\hat{\mathbf{x}}_{k|k-1}^T\Big|\mathbf{y}_{k-1}\Big\}, \tag{E.3}$$

and the following was used for the second equality in (E.2):

$$\begin{aligned} \mathcal{E}\Big\{\mathbf{x}_k\hat{\mathbf{x}}_{k|k-1}^T\Big|\mathbf{y}_{k-1}\Big\} &= \mathcal{E}\Big\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})\hat{\mathbf{x}}_{k|k-1}^T\Big|\mathbf{y}_{k-1}\Big\} + \mathcal{E}\Big\{\hat{\mathbf{x}}_{k|k-1}\hat{\mathbf{x}}_{k|k-1}^T\Big|\mathbf{y}_{k-1}\Big\} \\ &= \mathcal{E}\Big\{\hat{\mathbf{x}}_{k|k-1}\hat{\mathbf{x}}_{k|k-1}^T\Big|\mathbf{y}_{k-1}\Big\} = \hat{\mathbf{X}}_k \end{aligned} \tag{E.4}$$

The following lemma can be easily proved.

Lemma 9 *The function $e : \mathcal{K} \mapsto \mathbb{R}^+$ given by* (E.2) *is convex.*

Since the function $e(\mathbf{L}_k)$ is convex, in order to find $\mathbf{L}_k$ which minimizes $e(\mathbf{L}_k)$, we simply seek the gain $\mathbf{L}_k$ at which $\partial e(\mathbf{L}_k)/\partial\mathbf{L}_k = \mathbf{0}$. Then, we have

$$0 = \frac{\partial e(\mathbf{L}_k)}{\partial\mathbf{L}_k} = 2(\mathbf{D}\mathbf{B})^T(\mathbf{D}\mathbf{B})\mathbf{L}_k\hat{\mathbf{X}}_k - 2(\mathbf{D}\mathbf{B})^T\mathbf{D}\mathbf{A}\hat{\mathbf{X}}_k. \tag{E.5}$$

Therefore, finding $\mathbf{L}_k$ which minimizes $e(\mathbf{L}_k)$ needs to solve the following equation

$$(\mathbf{D}\mathbf{B})^T(\mathbf{D}\mathbf{B})\mathbf{L}_k\hat{\mathbf{X}}_k = (\mathbf{D}\mathbf{B})^T\mathbf{D}\mathbf{A}\hat{\mathbf{X}}_k. \tag{E.6}$$

To solve this, we use Lemma 1. To see if the solution $\mathbf{L}_k$ exists, we investigate the condition (ii) of Lemma 1. Since

$$(\mathbf{D}\mathbf{B})^T(\mathbf{D}\mathbf{B})\Big[(\mathbf{D}\mathbf{B})^T(\mathbf{D}\mathbf{B})\Big]^\dagger = (\mathbf{D}\mathbf{B})^T(\mathbf{D}\mathbf{B})(\mathbf{D}\mathbf{B})^\dagger(\mathbf{D}\mathbf{B})^{T^\dagger} = (\mathbf{D}\mathbf{B})^T(\mathbf{D}\mathbf{B})^{T^\dagger}, \tag{E.7}$$

the left-hand side of the condition (ii) of Lemma 1 becomes

$$\text{LHS of (ii)} = (\mathbf{D}\mathbf{B})^T(\mathbf{D}\mathbf{B})^{T^\dagger}(\mathbf{D}\mathbf{B})^T\mathbf{D}\mathbf{A}\hat{\mathbf{X}}_k\hat{\mathbf{X}}_k^\dagger\hat{\mathbf{X}}_k = (\mathbf{D}\mathbf{B})^T\mathbf{D}\mathbf{A}\hat{\mathbf{X}}_k = \text{RHS of (ii)}. \tag{E.8}$$

Therefore, the solution $\mathbf{L}_k$ always exists, and the all solution $\mathbf{L}_k$ is given by, from (8),

$$\mathbf{L}_k = (\mathbf{D}\mathbf{B})^\dagger\mathbf{D}\mathbf{A}\mathbf{P}_{\mathcal{R}(\hat{\mathbf{X}}_k)} + \mathbf{G} - \mathbf{P}_{\mathcal{R}(\mathbf{D}\mathbf{B})^T}\mathbf{G}\mathbf{P}_{\mathcal{R}(\hat{\mathbf{X}}_k)} \tag{E.9}$$

where $\mathbf{G}$ is an arbitrary matrix with consistent dimension. If the covariance $\hat{\mathbf{X}}_k$ is positive definite, the gain becomes

$$\mathbf{L}_k = (\mathbf{D}\mathbf{B})^\dagger\mathbf{D}\mathbf{A} + \mathbf{G} - \mathbf{P}_{\mathcal{R}(\mathbf{D}\mathbf{B})^T}\mathbf{G} = (\mathbf{D}\mathbf{B})^\dagger\mathbf{D}\mathbf{A} + \mathbf{P}_{\mathcal{N}(\mathbf{D}\mathbf{B})}\mathbf{G} \tag{E.10}$$

which is (66). □

References

Bitmead, R. R. & Gevers, M. (1991), Riccati difference and differential equations: Convergence, monotonicity and stability, *in* S. Bittanti, A. J. Laub & J. C. Willems, eds, 'The Riccati Equation', Springer-Verlag, Berlin, Germany, pp. 263–291.

Blanchini, F. (1999), 'Set invariance in control', *Automatica* **35**, 1747–1767.

Campbell, S. L. & C. D. Meyer, J. (1979), *Generalized Inverses of Linear Transformations*, Dover Publications, Inc., New York, NY.

Castelan, E. B. & Hennet, J.-C. (1992), 'Eigenstructure assignment for state constrained linear continuous time systems', *Automatica* **28**(3), 605–611.

Clements, D. J. & Anderson, B. D. (1978), *Singular Optimal Control: The Linear-Quadratic Problem*, Springer-Verlag, Berlin, Germany.

De Souza, C. E. (1989), 'On stabilizing properties of solutions of the riccati difference equation', *IEEE Trans. Automat. Contr.* **34**, 1313–1316.

Hahn, H. (1992), Linear systems controlled by stabilized constraint relations, *in* 'Proceedings of the 31st Conference on Decision and Control', Tucson, Arizona, pp. 840–848.

Hemami, H. & Wyman, B. F. (1979), 'Modeling and control of constrained dynamic systems with application to biped locomotion in the frontal plane', *IEEE Trans. Automat. Contr.* **24**(4), 526–535.

Johnson, C. D. (1973), 'Stabilization of linear dynamical systems with respect to arbitrary linear subspaces', *J. Math. Anal. Appl.* **44**(1), 175–186.

Johnson, C. D. (2000), 'New results in subspace-stabilization control theory', *Mathematical Problems in Engineering* **24**(2-3), 99–124.

Ko, S. (2005), Performance Limitations in Linear Estimation and Control: Quantization and Constraint Effects (being prepared), PhD thesis, Univ. of California, San Diego, La Jolla, CA.

Ko, S. & Bitmead, R. R. (2005), State estimation of linear systems with state equality constraints, *in* '16th IFAC Congress', Prague, Czech. Accepted for publication.

Maciejowski, J. M. (2002), *Predictive Control with Constraints*, Pearson Education Ltd., Essex, England.

McClamroch, N. H. & Wang, D. (1988), 'Feedback stabilization and tracking of constrained robots', *IEEE Trans. Automat. Contr.* **33**(5), 419–426.

Oloomi, H. & Shafai, B. (1997), 'Constrained stabilization problem and transient mismatch phenomenon in singularity perturbed systems', *Int. J. Control* **67**(3), 435–454.

Åström, K. J. (1970), *Introduction to Stochastic Control Theory*, Academic Press, New York, NY.

Skelton, R. E., Iwasaki, T. & Grigoriadis, K. (1998), *A Unified Algebraic Approach to Linear Control Design*, Taylor and Francis, London, UK.

Tournes, C. & Johnson, C. D. (2001), 'Reusable launch vehicle control using linear-adaptive and subspace-stabilization techniques', *J. Guid. Control Dyn.* **24**(2), 278–286.

Tournes, C. & Landrum, B. (2003), 'F-14 aircraft lateral-directional adaptive control using subspace stabilization', *J. Guid. Control Dyn.* **26**(1), 167–169.

Wonham, W. M. (1979), *Linear multivariable control: a Geometric Approach*, Springer-Verlag, New York, NY.

Yu, T.-J., Lin, C.-F. & Müller, P. C. (1996), Design of LQ regulator for linear systems with algebraic-equation constraints, *in* 'Proceedings of the 35th Conference on Decision and Control', Kobe, Japan, pp. 4146–4151.

Yu, T.-J. & Müller, P. C. (1994), 'Design of pole-assignment controller for systems with algebraic-equation constraint', *Systems & Control Letters* **23**, 463–471.