

Multi-objective Control Design for Discrete Time Periodic Systems via Convex Optimization

Richard Conway

Computer Mechanics Laboratory
Mechanical Engineering
University of California, Berkeley, CA, USA

January 21, 2009

Abstract

This report considers the design of controllers for linear discrete time systems with a periodic state space realization. Two types of cost functions are considered here—the ℓ_2 semi-norm, which generalizes the \mathcal{H}_2 norm, and the ℓ_2 induced norm, which generalizes the \mathcal{H}_∞ norm for stable systems. The theory for the ℓ_2 semi-norm and ℓ_2 induced norm is rigorously developed for systems with periodic state space realizations. These results are then used to construct a methodology for multi-objective control design which utilizes convex optimization. Finally, this methodology is applied to track-following control design for hard disk drives. For this particular application, it is shown that the methodology is much less conservative than in the general case. Multi-objective design is compared to standard single-objective design and the effects of introducing multi-rate sampling and actuation characteristics (i.e. periodicity) into the plant model is examined. It is shown that both multi-objective and multi-rate design greatly improve the closed-loop performance of the system.

Chapter 1

Introduction

In the past several decades, optimal output feedback control has been the subject of much research. For a long time, this research focused on methods whose solutions could be easily solved using Riccati equations, such as \mathcal{H}_2 optimal control and \mathcal{H}_∞ optimal control. These control design methodologies have solutions which are easy to compute, guarantee high performance of the closed-loop system, and when formulated correctly guarantee robust stability of the closed-loop system. However, all of these methodologies have the drawback of not being very flexible.

In more recent years, with the rapid increase of available computational power and efficient numerical optimization algorithms, this view of control design has begun to shift towards the view that control design problems are actually constrained optimization problems of the form

$$\begin{aligned} \min_x \quad & c^T x \\ \text{s.t.} \quad & F(x) \succeq 0 \end{aligned}$$

where the notation $F(x) \succeq 0$ means that $F(x)$ is positive semi-definite. Since this is, in general, an NP-hard problem, the best we can hope for is finding a local minimum of our problem. However, in the special case when F is affine in x , the optimization problem becomes a semi-definite program (SDP), for which there are numerous algorithms and available software packages to find the solution. In addition, since SDP's are convex optimizations, these algorithms are guaranteed to find the globally optimal solution in polynomial time. Although these optimization-based control design methods are not as computationally efficient as those based on Riccati equations, they have the benefit of being very flexible.

One area in which the optimization-based approach to control design is particularly beneficial is multi-objective control, in which multiple cost functions are simultaneously considered and constraints are imposed

on the system [11]. An example of multi-objective control design would be to maximize the performance of a system subject to the system having “enough” robustness. Clearly, these types of formulations allow for much more general control design problems than the methodologies which use Riccati equations. Thus, multi-objective formulations allow the control designer to deal with design tradeoffs and constraints in a much more explicit manner, which in turn reduces the need for design iteration.

This report considers a certain class of discrete time multi-objective control problems for linear periodic time-varying (LPTV) systems and shows how they can be represented as SDP’s. Chapter 2 presents the theory associated with this class of problems. Then, in Chapter 3, this methodology is applied to the problem of finding a track-following dual-stage hard disk drive controller which has multi-rate sampling and actuation characteristics.

Chapter 2

Periodic Discrete Time Systems

In this chapter, we start off by introducing the notation that will be used in this report and the class of systems that will be examined. To motivate this class of systems, we will show that when multi-rate sampling and actuation characteristics are added to a linear time-invariant (LTI) system, LPTV systems arise [2].

We will then define the ℓ_2 semi-norm of a linear operator and show a computationally feasible method of computing it when the operator can be realized as an LPTV state space model. At this point, the relationship between the ℓ_2 semi-norm of a linear operator and the \mathcal{H}_2 norm of a LTI system will be discussed. Finally, a sharp upper bound on the ℓ_2 semi-norm of an operator will be presented in terms of matrix inequalities for the case when the operator can be realized as an LPTV state space model.

After presenting the ℓ_2 semi-norm of a linear operator, we will define the ℓ_2 induced norm of a linear operator and state some of its relevant properties when the operator can be realized as an LPTV state space model. At this point, the relationship between the ℓ_2 induced norm of a linear operator and the \mathcal{H}_∞ norm of an LTI system will be discussed. Finally, a sharp upper bound on the ℓ_2 induced norm of a linear operator will be presented in terms of matrix inequalities for the case when the operator can be realized as an LPTV state space model.

Finally, we will present an extension of the methodology in [11] for formulating a general class of multi-objective control design problems as SDP's with little conservatism. A few notes on the implementation and efficiency of this optimization will then be made.

Many of the analytical developments related to the ℓ_2 semi-norm and the ℓ_2 induced norm in this section are drawn from [10]. However, in that work, the matrices in their state space realizations are taken to be infinite-dimensional block-diagonal linear operators. In an effort to simplify the analysis, we will instead interpret the state space matrices as being a function of time.

2.1 Preliminaries

To begin, we will introduce some definitions for linear operators. First we define for finite-dimensional matrices the norm

$$\|M\| := \sup_{x \neq 0} \frac{(Mx)^T Mx}{x^T x}.$$

Note that this norm is, by definition, equivalent to the largest singular value of a matrix and thus has the properties

$$\|MN\| \leq \|M\| \cdot \|N\|$$

$$\|M^T\| = \|M\|.$$

The former of these properties is called the sub-multiplicative property. We define the spectral radius of M as

$$\rho(M) := \lim_{n \rightarrow \infty} \|M^n\|^{1/n}$$

which can be shown to satisfy

$$\rho(M) = \max_k |\lambda_k(M)|$$

where $\lambda_k(M)$ is the k^{th} eigenvalue of M . For all matrices considered in this report, unspecified entries are zero and bullets represent entries whose values can be inferred by symmetry of the matrix. We will denote that a symmetric matrix M is positive definite by $M \succ 0$.

Now, note that linear systems can of the form

$$y_k = \sum_{j=-\infty}^{\infty} G_{k,j} u_j$$

be represented as infinite-dimensional linear operators of the form

$$\begin{bmatrix} \vdots \\ y_0 \\ y_1 \\ \vdots \end{bmatrix} = \begin{bmatrix} \ddots & \vdots & \vdots & \vdots \\ \cdots & G_{0,0} & G_{0,1} & \cdots \\ \cdots & G_{1,0} & G_{1,1} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} \vdots \\ u_0 \\ u_1 \\ \vdots \end{bmatrix}$$

where u_k and y_k are respectively the input and output at time k . This motivates the notation

$$y := \begin{bmatrix} \vdots \\ y_0 \\ y_1 \\ \vdots \end{bmatrix}, \quad G := \begin{bmatrix} \ddots & \vdots & \vdots & \vdots \\ \cdots & G_{0,0} & G_{0,1} & \cdots \\ \cdots & G_{1,0} & G_{1,1} & \cdots \\ \cdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad u := \begin{bmatrix} \vdots \\ u_0 \\ u_1 \\ \vdots \end{bmatrix}.$$

In this framework, causal operators are ones that have the property that

$$G_{k,j} = 0, \quad j > k.$$

Now we define the vector space

$$\ell_2 := \left\{ u : \sum_{k=-\infty}^{\infty} u_k^T u_k < \infty \right\}$$

and a norm on that vector space given by

$$\|u\|_{\ell_2} := \sqrt{\sum_{k=-\infty}^{\infty} u_k^T u_k}.$$

With this in mind, we will say that G is bounded if $Gu \in \ell_2, \forall u \in \ell_2$.

Now that we've finished introducing the operator notation, we will introduce the state space notation and definitions. We will say that a causal linear operator has the realization

$$G \sim \left(\begin{array}{c|c} A_k & B_k \\ \hline C_k & D_k \end{array} \right) \tag{2.1}$$

if

$$\begin{bmatrix} x_{k+1} \\ y_k \end{bmatrix} = \begin{bmatrix} A_k & B_k \\ C_k & D_k \end{bmatrix} \begin{bmatrix} x_k \\ u_k \end{bmatrix}.$$

We will further say that this is an LPTV realization if

$$\begin{bmatrix} A_k & B_k \\ C_k & D_k \end{bmatrix} = \begin{bmatrix} A_{k+N} & B_{k+N} \\ C_{k+N} & D_{k+N} \end{bmatrix}, \quad \forall k.$$

A causal operator which has an LPTV realization will be called an LPTV operator or an LPTV system. The realization (2.1) will be called uniformly exponentially stable (UES) if there exist constants $c > 0$ and

$\beta \in [0, 1)$ such that

$$\|A_{k+l-1} \cdots A_k\| \leq c\beta^l, \quad \forall k \in \mathbb{Z}, l \in \mathbb{N}$$

where \mathbb{Z} denotes the set of integers and \mathbb{N} denotes the set of positive integers. (Note that this is unrelated to the condition that $\rho(A_k) < 1, \forall k$.) Finally, we have the following lemma [1].

Lemma 1. *If a realization is UES and $\exists M > 0$ such that $\|B_k\|, \|C_k\| < M, \forall k$, then the operator that it realizes is bounded.*

2.2 Multi-rate Sampling and Actuation

Before we develop the theory behind LPTV operators, we will first motivate why this is an important class of operators by presenting a particular instance in which LPTV operators arise. Suppose we have an LTI model with measurements and control inputs respectively given by

$$y_k = \begin{bmatrix} y_k^1 \\ \vdots \\ y_k^{n_y} \end{bmatrix} \quad u_k = \begin{bmatrix} u_k^1 \\ \vdots \\ u_k^{n_u} \end{bmatrix}$$

where y_k^i and u_k^i are scalars. We may want to measure the components of y_k at different rates, i.e. we measure y_k^i only once every N_i time steps, where N_i is a positive integer. This situation could arise, for example, when measurement rates of certain signals are limited by the physical system. For completeness, we now further suppose that we want to actuate the plant at different rates, i.e. we change the value of u_k^i only once every M_i time steps, where M_i is a positive integer. We now define N to be the least common multiple of $N_1, \dots, N_{n_y}, M_1, \dots, M_{n_u}$. There are thus two goals. First, we would like to create a model for a multi-rate sampler which will take y_k as its input and only let the measurements pass which are being measured at that particular time step. Second, we would like to create a model for a multi-rate hold which will take its inputs from a controller and hold these inputs for the necessary number of time steps in order to create actuation signals which change at different rates than the LTI model actuation rate.

We will begin by considering the multi-rate sampler. Since we only measure y_k^i once every N_i time steps, we are interested in constructing the signal

$$\bar{y}_k^i := \begin{cases} y_k^i, & \frac{k - n_i}{N_i} \in \mathbb{Z} \\ 0, & \text{otherwise} \end{cases}$$

where $n_i \in \{0, \dots, N_i - 1\}$ represents one particular time step at which we measure y_k^i . This can be realized as just a time-varying gain

$$\bar{y}_k^i = S_k^i y_k^i$$

$$S_k^i := \begin{cases} 1, & \frac{k - n_i}{N_i} \in \mathbb{Z} \\ 0, & \text{otherwise} \end{cases}.$$

It is clear that S_k^i is periodic and has period N , i.e. $S_k^i = S_{k+N}^i, \forall k$. Now we can stack these up to form the time-varying gain

$$S_k := \begin{bmatrix} S_k^1 & & \\ & \ddots & \\ & & S_k^{n_y} \end{bmatrix}.$$

This S_k is our multi-rate sampler model. Since S_k^i has period N , it is obvious that S_k also has period N . Thus, our multi-rate sampler is just a time-varying diagonal gain matrix with period N whose entries are zeros and ones.

Now we turn our attention to finding a model for the multi-rate hold. Since we only change the value of u_k^i once every M_i time steps, we are interested in constructing the signal

$$\bar{u}_k^i := \begin{cases} u_k^i, & \frac{k - m_i}{M_i} \in \mathbb{Z} \\ \bar{u}_{k-1}^i, & \text{otherwise} \end{cases}$$

where $m_i \in \{0, \dots, M_i - 1\}$ represents one particular time step at which we change the value of u_k^i . When $M_i = 1$, clearly $\bar{u}_k^i = u_k^i$, which can be realized as the LTI static gain, $H^i = 1$. When $M_i \neq 1$ it is easily verified that the dynamics from u_k^i to \bar{u}_k^i have the LPTV realization

$$H^i \sim \left(\begin{array}{c|c} 1 - \delta_k & \delta_k \\ \hline 1 - \delta_k & \delta_k \end{array} \right)$$

$$\delta_k = \begin{cases} 1, & \frac{k - m_i}{M_i} \in \mathbb{Z} \\ 0, & \text{otherwise} \end{cases}.$$

Note that in both cases, the state space realization has period N . We now stack these LPTV realizations

up to form

$$H := \begin{bmatrix} H^1 & & \\ & \ddots & \\ & & H^{n_u} \end{bmatrix}.$$

Since each of the realizations of H^i have period N , the realization of H also has period N .

When we hook up our system to the multi-rate sampler and hold, the resulting system is still linear, but its state space realization is periodically time varying. Hence, multi-rate measurement sampling and control actuation results in LPTV systems.

2.3 ℓ_2 Semi-Norms of Operators

We first define the ℓ_2 semi-norm of the bounded operator G as

$$\|G\|_2^2 := \limsup_{l \rightarrow \infty} \frac{1}{2l+1} \sum_{k=-l}^l \sum_{j=-\infty}^{\infty} \text{tr} \left\{ G_{k,j} (G_{k,j})^T \right\}. \quad (2.2)$$

Here, the factor of $2l+1$ is introduced to ensure that the ℓ_2 semi-norm is bounded for all admissible G and the limit supremum guarantees existence of the limit in the right-hand side of (2.2). In a sense, the ℓ_2 semi-norm of an operator can be thought of as being analogous to the Frobenius norm for matrices, except with a normalizing factor to ensure that it is bounded for bounded operators. It should also be noted that the ℓ_2 semi-norm is not, in general, a norm. For instance, if an operator had only one non-zero entry in its infinite-dimensional representation, it would not be the zero operator, but the ℓ_2 semi-norm of that operator would be zero.

2.3.1 ℓ_2 Semi-Norms of LPTV Operators

Since we are interested in operators with time-varying state space realizations, we now present some results which allow us to calculate the ℓ_2 semi-norm of these operators. We begin by looking at an equation which is analogous to a discrete time Lyapunov equation for an LTI system.

Lemma 2. *Suppose that a realization given by (2.1) is UES and a given sequence, W_k , satisfies*

$$W_k \succeq 0, \quad \|W_k\| \leq M \quad \forall k$$

for some scalar $M > 0$. Then the equation

$$L_{k+1} = A_k L_k A_k^T + W_k \quad (2.3)$$

has a unique bounded solution given by

$$L_k = W_{k-1} + \sum_{j=1}^{\infty} (A_{k-1} \cdots A_{k-j}) W_{k-j-1} (A_{k-1} \cdots A_{k-j})^T. \quad (2.4)$$

Moreover, $L_k \succeq 0$.

The proof of this lemma will be deferred to Appendix A.1. There are two ways to interpret (2.3): as an infinite-dimensional linear equation or as an update equation in increasing k . For this second interpretation, the unique bounded solution can be thought of as being the result of the following limiting process. Define the sequence of sequences

$$L_k^{[i]} := \begin{cases} 0, & k < -i \\ A_{k-1} L_{k-1}^{[i]} A_{k-1}^T + W_{k-1}, & k \geq -i \end{cases}.$$

Notice that for each i , $L_k^{[i]}$ can be viewed as an update equation in increasing k where the initial condition is taken to be $L_{-i-1}^{[i]} = 0$. The unique bounded solution is then arrived at by simply taking the limit

$$L_k = \lim_{i \rightarrow \infty} L_k^{[i]}.$$

Just as the Lyapunov equation for LTI systems plays an important role in computing the \mathcal{H}_2 norm of a system, this equation will play an important role in finding the ℓ_2 semi-norm of a operator with a time-varying realization. The next lemma will highlight this fact.

Lemma 3. *If a realization given by (2.1) is UES and $\exists M > 0$ such that*

$$\|B_k\|, \|C_k\| < M, \quad \forall k,$$

then

$$\|G\|_2^2 = \limsup_{l \rightarrow \infty} \frac{1}{2l+1} \sum_{k=-l}^l \text{tr} \{D_k D_k^T + C_k L_k C_k^T\} \quad (2.5)$$

where L_k is unique bounded solution of

$$L_{k+1} = A_k L_k A_k^T + B_k B_k^T. \quad (2.6)$$

Proof. First note that by Lemma 1, the operator corresponding to this realization is bounded. Now note that

$$\begin{aligned}
y_k &= D_k u_k + C_k x_k \\
&= D_k u_k + C_k \left(B_{k-1} u_{k-1} + \sum_{j=2}^{\infty} A_{k-1} \cdots A_{k-j+1} B_{k-j} u_{k-j} \right) \\
\Rightarrow G_{k,k-j} &= \begin{cases} 0, & j < 0 \\ D_k, & j = 0 \\ C_k B_{k-1}, & j = 1 \\ C_k A_{k-1} \cdots A_{k-j+1} B_{k-j}, & j > 1 \end{cases}.
\end{aligned}$$

Defining $W_k = B_k B_k^T \succeq 0$ gives

$$\begin{aligned}
\sum_{j=-\infty}^{\infty} \text{tr} \{ G_{k,j} G_{k,j}^T \} &= \text{tr} \left\{ D_k D_k^T + C_k \left[W_{k-1} \right. \right. \\
&\quad \left. \left. + \sum_{j=2}^{\infty} (A_{k-1} \cdots A_{k-j+1}) W_{k-j} (A_{k-1} \cdots A_{k-j+1})^T \right] C_k^T \right\}.
\end{aligned}$$

By Lemma 2, the term in square brackets is the unique L_k corresponding to a bounded solution of (2.6). Plugging this into (2.2) then gives the desired result. \square

Although this result is computationally intractable for a general time-varying realization, the next theorem will show that for LPTV realizations, the limit supremum in (2.5) is equivalent to the limit and the infinite sum becomes a finite sum.

Theorem 4. *If an LPTV realization given by (2.1) is UES, then*

$$\|G\|_2^2 = \frac{1}{N} \sum_{k=1}^N \text{tr} \{ D_k D_k^T + C_k L_k C_k^T \} \tag{2.7}$$

where

$$L_{k+1} = A_k L_k A_k^T + B_k B_k^T. \tag{2.8}$$

Moreover,

$$L_{k+N} = L_k, \quad \forall k.$$

Proof. Since we only have a finite number of values of B_k and C_k , it is trivial to verify that $\exists M > 0$ such

that $\|B_k\|, \|C_k\| < M, \forall k$. Therefore, (2.8) must have a unique bounded solution. Now, for notational convenience, we define $R_k := \text{tr}\{D_k D_k^T + C_k L_k C_k^T\}$. Note that if we define $\bar{L}_k := L_{k+N}$, we get that

$$\begin{aligned}\bar{L}_{k+1} &= A_{k+N} \bar{L}_k A_{k+N}^T + B_{k+N} B_{k+N}^T \\ &= A_k \bar{L}_k A_k^T + B_k B_k^T.\end{aligned}$$

Thus, by uniqueness of L_k (Lemma 2),

$$\begin{aligned}L_{k+N} &= \bar{L}_k = L_k \\ \Rightarrow R_k &= R_{k+N}.\end{aligned}$$

Now let

$$l = Nn_l + j$$

where n_l is the largest integer such that $Nn_l \leq l$. Now we can express

$$\begin{aligned}\|G\|_2^2 &= \limsup_{l \rightarrow \infty} \frac{1}{2Nn_l + 2j + 1} \sum_{k=-Nn_l-j}^{Nn_l+j} R_k \\ &= \limsup_{l \rightarrow \infty} \left\{ \frac{1}{2Nn_l + 2j + 1} \left(\sum_{i=-n_l}^{n_l-1} \sum_{k=iN+1}^{(i+1)N} R_k + \sum_{k=-Nn_l-j}^{-Nn_l} R_k + \sum_{k=Nn_l+1}^{Nn_l+j} R_k \right) \right\} \\ &= \limsup_{l \rightarrow \infty} \left\{ \frac{1}{2Nn_l + 2j + 1} \left(2n_l \sum_{k=1}^N R_k + \sum_{k=N-j}^N R_k + \sum_{k=1}^j R_k \right) \right\}.\end{aligned}$$

Note that by letting $n_l \gg j$, we can see that the limit (not limit supremum) of the term in curly braces is equal to the desired result. Since the limit supremum is equal to the limit when it exists, this completes the proof. \square

This theorem lends itself to a computationally tractable means of computing the ℓ_2 semi-norm of LPTV operators. To apply this theorem, since the solution of (2.8) is periodic, we only to solve a finite set of linear matrix equations for the finite number of matrices (L_1, \dots, L_N) which uniquely solve them and plug those values into a finite summation.

2.3.2 Interpretation of ℓ_2 Semi-Norms of LPTV Operators

To build intuition regarding these ℓ_2 semi-norm of LPTV operators, we will now discuss several interpretations of Theorem 4. First of all, it should be noted that when $N = 1$ (i.e. the realization is LTI), the ℓ_2

semi-norm agrees with the \mathcal{H}_2 norm for its LTI realization. Since the \mathcal{H}_2 norm for an LTI system can be related to the output covariance of a system, we would like to see if there is a similar interpretation for the ℓ_2 semi-norm. Suppose that G is driven by u where

$$\begin{aligned} E[u] &= 0 \\ E[uu^T] &= I \end{aligned}$$

i.e. u_k is zero-mean white noise with covariance I . This implies that

$$\begin{aligned} E[y] &= E[G u] = G E[u] = 0 \\ E[yy^T] &= E[G u u^T G^T] = G E[u u^T] G^T = G G^T. \end{aligned}$$

Thus, we can say that the covariance of y_k is

$$E[y_k y_k^T] = \sum_{j=-\infty}^{\infty} G_{k,j} (G_{k,j})^T.$$

Now, recall from the proof of Theorem 4 that the limit supremum in (2.2) could be replaced by a limit. Thus, when y_k is scalar, we see that

$$\|G\|_2^2 = \lim_{l \rightarrow \infty} \frac{1}{2l+1} \sum_{k=-l}^l E[y_k^2].$$

Thus, we can interpret the ℓ_2 semi-norm in this case as the RMS (over time) standard deviation of y_k . Note that as hoped, this interpretation is similar to the interpretation that the \mathcal{H}_2 norm of an LTI system with a scalar output corresponds to the standard deviation of that output when driven by white noise whose covariance is I .

We now extend this comparison of the ℓ_2 semi-norm to the \mathcal{H}_2 norm by using a technique which “lifts”

the LPTV system to an LTI system with similar characteristics [3]. We first define

$$\begin{aligned}
\mathcal{Z} &:= \begin{bmatrix} 0 & I & & 0 \\ & \ddots & \ddots & \\ & & \ddots & I \\ I & & & 0 \end{bmatrix} & \mathcal{A} &:= \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_N \end{bmatrix} & \mathcal{B} &:= \begin{bmatrix} B_1 & & \\ & \ddots & \\ & & B_N \end{bmatrix} \\
\mathcal{L} &:= \begin{bmatrix} L_1 & & \\ & \ddots & \\ & & L_N \end{bmatrix} & \mathcal{C} &:= \begin{bmatrix} C_1 & & \\ & \ddots & \\ & & C_N \end{bmatrix} & \mathcal{D} &:= \begin{bmatrix} D_1 & & \\ & \ddots & \\ & & D_N \end{bmatrix}
\end{aligned} \tag{2.9}$$

where, as previously mentioned, all unspecified elements are zero. With this new notation in mind, the following lemma (whose proof will be deferred to Appendix A.2) gives an computationally tractable method for determining whether or not an LPTV realization is UES.

Lemma 5. *An LPTV realization is UES $\Leftrightarrow \rho(\mathcal{Z}^T \mathcal{A}) < 1$, where \mathcal{Z} and \mathcal{A} are as in (2.9).*

Now we will show a means by which to calculate the ℓ_2 semi-norm of an LPTV realization. We begin by defining

$$\bar{\mathcal{L}} := \mathcal{Z} \mathcal{L} \mathcal{Z}^T = \begin{bmatrix} L_2 & & & \\ & \ddots & & \\ & & L_N & \\ & & & L_1 \end{bmatrix}$$

so that (2.8) can be written in block diagonal form as

$$\bar{\mathcal{L}} = \mathcal{A} \mathcal{L} \mathcal{A}^T + \mathcal{B} \mathcal{B}^T.$$

Exploiting the fact that \mathcal{Z} is unitary gives

$$\begin{aligned}
\mathcal{L} &= \mathcal{Z}^T (\mathcal{A} \mathcal{L} \mathcal{A}^T + \mathcal{B} \mathcal{B}^T) \mathcal{Z} \\
&= (\mathcal{Z}^T \mathcal{A}) \mathcal{L} (\mathcal{Z}^T \mathcal{A})^T + (\mathcal{Z}^T \mathcal{B}) (\mathcal{Z}^T \mathcal{B})^T.
\end{aligned}$$

Furthermore, (2.7) can be written

$$\|G\|_2^2 = \frac{1}{N} \text{tr} \{ \mathcal{D} \mathcal{D}^T + \mathcal{C} \mathcal{L} \mathcal{C}^T \}.$$

With all of this in place, it is easily verified that

$$\|G\|_2^2 = \frac{1}{N} \left\| \left(\begin{array}{c|c} \mathbf{Z}^T \mathcal{A} & \mathbf{Z}^T \mathcal{B} \\ \hline \mathcal{C} & \mathcal{D} \end{array} \right) \right\|_{\mathcal{H}_2}^2 \quad (2.10)$$

where the norm on the right-hand side of the equation is the usual \mathcal{H}_2 norm for LTI systems. Interestingly enough, the LTI realization on the right-hand side of the equation is *not* a realization of G . However, if we wanted to leverage existing software packages to compute the ℓ_2 semi-norm of an LPTV operator, using this LTI system for computational purposes would work reasonably well.

It should be noted, though, that the `MATLAB` `ss` object does not support sparse state space matrices. Furthermore, most of the algorithms for `ss` objects do not exploit sparsity. For instance, if you tried to find \mathcal{L} by computing the controllability gramian of the LTI system in (2.10), the solution will not, in general, be exactly block diagonal—there will be some small numerical errors. Furthermore, these algorithms will tend to be inefficient for large N . A better approach would be to instead consider the “lifted” system

$$\left(\begin{array}{c|cccc} \bar{A}_{N,1} & \bar{A}_{N,2}B_1 & \bar{A}_{N,3}B_2 & \cdots & \bar{A}_{N,N}B_{N-1} & B_N \\ \hline C_1 & D_1 & & & & \\ C_2A_1 & C_2B_1 & D_2 & & & \\ \vdots & \vdots & C_3B_2 & \ddots & & \\ C_{N-1}\bar{A}_{N-2,1} & C_{N-1}\bar{A}_{N-2,2}B_1 & \vdots & \ddots & D_{N-1} & \\ C_N\bar{A}_{N-1,1} & C_N\bar{A}_{N-1,2}B_1 & C_N\bar{A}_{N-1,3}B_2 & \cdots & C_NB_{N-1} & D_N \end{array} \right)$$

where $\bar{A}_{i,j} = A_i \cdots A_j$. If we consider this system to have input \bar{u}_k and output \bar{y}_k which are given by

$$\bar{u}_k := \begin{bmatrix} u_{N(k-1)+1} \\ \vdots \\ u_{Nk} \end{bmatrix} \quad \bar{y}_k := \begin{bmatrix} u_{N(k-1)+1} \\ \vdots \\ u_{Nk} \end{bmatrix},$$

it is straightforward to check that this system realizes G . (The only difference is that the block partitioning is a bit different.) Since this system has a factor of N fewer states than the system in (2.10), it is a lot more efficient and reliable to compute the \mathcal{H}_2 norm of this system. Since this is an LTI realization of G , its \mathcal{H}_2 norm is equivalent to the ℓ_2 semi-norm of G . However, this LTI formulation is only useful for computing the norm of an LPTV operator; it is not convenient for control design.

2.3.3 Upper Bound on ℓ_2 Semi-Norms of LPTV Operators

In this section, we present two formulations of an upper bound on the ℓ_2 semi-norm of an LPTV operator. Both of these upper bounds are shown to be sharp. Although this seems like a counter-intuitive development because we can directly compute the actual value of the ℓ_2 semi-norm using Theorem 4, it will be essential to the control design developments in section 2.5.

Theorem 6. *For LPTV operators, the following conditions are equivalent:*

1. $\|G\|_2^2 < \gamma$
2. $\exists W_1, \dots, W_N, P_1, \dots, P_N$ such that

$$\frac{1}{N} \sum_{k=1}^N \text{tr } W_k < \gamma, \quad \begin{bmatrix} W_k & C_k P_k & D_k \\ \bullet & P_k & 0 \\ \bullet & \bullet & I \end{bmatrix} \succ 0, \quad \begin{bmatrix} P_{k+1} & A_k P_k & B_k \\ \bullet & P_k & 0 \\ \bullet & \bullet & I \end{bmatrix} \succ 0$$

holds $\forall k \in \{1, \dots, N\}$ where P_{N+1} is taken to be P_1 .

3. $\exists W_1, \dots, W_N, Q_1, \dots, Q_N$ such that

$$\frac{1}{N} \sum_{k=1}^N \text{tr } W_k < \gamma, \quad \begin{bmatrix} W_k & C_k & D_k \\ \bullet & Q_k & 0 \\ \bullet & \bullet & I \end{bmatrix} \succ 0, \quad \begin{bmatrix} Q_{k+1} & Q_{k+1} A_k & Q_{k+1} B_k \\ \bullet & Q_k & 0 \\ \bullet & \bullet & I \end{bmatrix} \succ 0$$

holds $\forall k \in \{1, \dots, N\}$ where Q_{N+1} is taken to be Q_1 .

The proof of this theorem will be deferred to Appendix A.3. Note that the two matrices and the trace in condition (2) are all affine in $W_1, \dots, W_N, P_1, \dots, P_N$. With a little manipulation, these conditions can be represented in the form $F(x) \succ 0$, where $F(x)$ varies linearly with x and x is a vector which contains all of the entries of the upper triangular part of the matrices $W_1, \dots, W_N, P_1, \dots, P_N$. (Since all of these matrices are symmetric, only their upper triangular entries are required to determine their values.) Thus, finding the ℓ_2 semi-norm of an LPTV operator can be represented as an SDP in which γ is minimized. Similarly, condition (3) can be used to construct a SDP to find the ℓ_2 semi-norm of an LPTV operator.

2.4 ℓ_2 Induced Norms of Operators

In this section, we look at the ℓ_2 induced norm of LPTV operators. First we define the ℓ_2 induced norm of a bounded operator to be

$$\|G\|_\infty := \sup_{u \in \ell_2 \setminus \{0\}} \frac{\|Gu\|_{\ell_2}}{\|u\|_{\ell_2}}.$$

In a sense, the ℓ_2 induced norm of an operator can be thought of as being analogous to the maximum singular value of a matrix. The basic result for calculating the ℓ_2 induced norm for an operator with a state space realization is given by the following lemma, which is proven in [10].

Lemma 7. *If G has the realization (2.1), then $\|G\|_\infty < 1 \Leftrightarrow \exists M > 0, Y_k \succ 0$ such that $\|Y_k\| < M \forall k$,*

$$Y_{k+1} = A_k Y_k A_k^T + B_k B_k^T + (A_k Y_k C_k^T + B_k D_k^T) U_k^{-1} (A_k Y_k C_k^T + B_k D_k^T)^T, \quad (2.11)$$

and A_k^L is UES where

$$\begin{aligned} U_k &= I - D_k D_k^T - C_k Y_k C_k^T \succ 0 \\ A_k^L &= A_k + (A_k Y_k C_k^T + B_k D_k^T) U_k^{-1} C_k. \end{aligned}$$

When we were analyzing the ℓ_2 semi-norm, the uniqueness of the solution to (2.3) was important because it allowed us to say that the solution was periodic for LPTV realizations, which greatly simplified analysis. Thus, we would like to see whether or not the Y_k mentioned in Theorem 7 is unique. The following lemma tells us that it is, indeed, unique.

Lemma 8. *If $\|G\|_\infty < 1$, then the Y_k mentioned in Theorem 7 is unique.*

Proof. Let Y_k, U_k, A_k^L satisfy the conditions in Theorem 7. Also let $\bar{Y}_k, \bar{U}_k, \bar{A}_k^L$ satisfy the conditions in Theorem 7. Note that the following identity holds:

$$\begin{aligned} U_k &= \bar{U}_k + C_k (Y_k - \bar{Y}_k) C_k^T \\ \Rightarrow \bar{U}_k^{-1} &= U_k^{-1} + \bar{U}_k^{-1} C_k (Y_k - \bar{Y}_k) C_k^T U_k^{-1} \\ \Rightarrow \bar{U}_k^{-1} + \bar{U}_k^{-1} C_k Y_k C_k^T U_k^{-1} &= U_k^{-1} + \bar{U}_k^{-1} C_k \bar{Y}_k C_k^T U_k^{-1} = 0. \end{aligned}$$

Using this identity, it is straightforward to show that

$$\begin{aligned}
Y_{k+1} - \bar{Y}_{k+1} &= \bar{A}_k^L (Y_k - \bar{Y}_k) (A_k^L)^T \\
\Rightarrow Y_k - \bar{Y}_k &= \left(\bar{A}_{k-1}^L \cdots \bar{A}_{k-j}^L \right) (Y_{k-j} - \bar{Y}_{k-j}) (A_{k-1}^L \cdots A_{k-j}^L)^T \\
&\Rightarrow \|Y_k - \bar{Y}_k\| \leq \bar{c} \bar{\beta}^j \|Y_{k-j} - \bar{Y}_{k-j}\| c \beta^j.
\end{aligned}$$

Since Y_{k-j} and \bar{Y}_{k-j} are bounded, the only way this can be true is when $Y_k = \bar{Y}_k$. \square

Now, we look at the LPTV case. Our approach will be to first find a “lifted” LTI system [3] whose \mathcal{H}_∞ norm is equal to the ℓ_2 induced norm of the LPTV operator and then use the standard results for finding the \mathcal{H}_∞ norm of an LTI system to find an upper bound on the ℓ_2 induced norm of the LPTV operator. We begin by stating and proving the following lemma, which relates the ℓ_2 induced norm of an LPTV operator to the \mathcal{H}_∞ norm of a “lifted” LTI system.

Lemma 9. *If G is a bounded LPTV operator, then*

$$\|G\|_\infty < \gamma \quad \Leftrightarrow \quad \left\| \left(\begin{array}{c|c} \mathcal{Z}^T \mathcal{A} & \mathcal{Z}^T \mathcal{B} \\ \hline \mathcal{C} & \mathcal{D} \end{array} \right) \right\|_{\mathcal{H}_\infty} < \gamma.$$

Proof. First note that it suffices to prove the result when $\gamma = 1$ because, for any norm, $\|\alpha G\|_x = |\alpha| \cdot \|G\|_x$, i.e. we can always scale an operator by a factor of γ^{-1} to recover the desired result. Since the solution to the Riccati equation (2.11) is unique, it is easily verified using the same methodology as in the proof of Theorem 4 that $Y_{k+N} = Y_k$. Thus if we define

$$\mathcal{Y} = \begin{bmatrix} Y_1 & & \\ & \ddots & \\ & & Y_N \end{bmatrix} \quad \mathcal{U} = \begin{bmatrix} U_1 & & \\ & \ddots & \\ & & U_N \end{bmatrix}$$

the conditions in Lemma 7 for $\|G\|_\infty < 1$ can be written

$$\begin{aligned}
\mathcal{Z} \mathcal{Y} \mathcal{Z}^T &= \mathcal{A} \mathcal{Y} \mathcal{A}^T + \mathcal{B} \mathcal{B}^T + (\mathcal{A} \mathcal{Y} \mathcal{C}^T + \mathcal{B} \mathcal{D}^T) \mathcal{U}^{-1} (\mathcal{A} \mathcal{Y} \mathcal{C}^T + \mathcal{B} \mathcal{D}^T)^T \succ 0 \\
\rho \left(\mathcal{Z}^T \mathcal{A} + (\mathcal{Z}^T \mathcal{A} \mathcal{Y} \mathcal{C}^T + \mathcal{Z}^T \mathcal{B} \mathcal{D}^T) \mathcal{U}^{-1} \mathcal{C} \right) &< 1
\end{aligned}$$

where

$$\mathcal{U} = I - \mathcal{D} \mathcal{D}^T - \mathcal{C} \mathcal{Y} \mathcal{C}^T \succ 0.$$

Multiplying the Riccati equation on the left by \mathcal{Z}^T and on the right by \mathcal{Z} reveals that these conditions are equivalent to saying that the \mathcal{H}_∞ norm of the LTI system above is < 1 . \square

It should be noted that when $N = 1$ (i.e. the realization is LTI), the ℓ_2 induced norm agrees with the \mathcal{H}_∞ norm for its LTI realization. Also, it is noteworthy that this “lifted” LTI system is the same as the one that arose in the analysis of the ℓ_2 semi-norm of a LPTV system. Now, with this lemma in place, we can now state and prove the final result of this section.

Theorem 10. *For UES LPTV operators, the following conditions are equivalent:*

1. $\|G\|_\infty^2 < \gamma$
2. $\exists P_1, \dots, P_N$ such that

$$\begin{bmatrix} \gamma I & 0 & C_k P_k & D_k \\ \bullet & P_{k+1} & A_k P_k & B_k \\ \bullet & \bullet & P_k & 0 \\ \bullet & \bullet & \bullet & I \end{bmatrix} \succ 0$$

holds $\forall k \in \{1, \dots, N\}$ where P_{N+1} is taken to be P_1 .

3. $\exists Q_1, \dots, Q_N$ such that

$$\begin{bmatrix} \gamma I & 0 & C_k & D_k \\ \bullet & Q_{k+1} & Q_{k+1} A_k & Q_{k+1} B_k \\ \bullet & \bullet & Q_k & 0 \\ \bullet & \bullet & \bullet & I \end{bmatrix} \succ 0$$

holds $\forall k \in \{1, \dots, N\}$ where Q_{N+1} is taken to be Q_1 .

The proof of this theorem will be deferred to Appendix A.4. Note that the matrix in condition (2) is affine in P_1, \dots, P_N . With a little manipulation, these conditions can be represented in the form $F(x) \succ 0$, where $F(x)$ varies linearly with x and x is a vector which contains all of the entries of the upper triangular part of the matrices P_1, \dots, P_N . Thus, finding the ℓ_2 induced norm of an LPTV operator can be represented as an SDP in which γ is minimized. Similarly, condition (3) can be used to construct a SDP to find the ℓ_2 induced norm of an LPTV operator.

2.5 Output Feedback Control Design via SDP's

In this section, we will apply the results of Theorems 6 and 10 to control design problems of the form

$$\min_{K, \Gamma} c^T u \text{ subject to:} \quad (2.12)$$

$$\begin{bmatrix} \|L_1 G_{cl}(K) R_1\|_{q_1}^2 \\ \vdots \\ \|L_M G_{cl}(K) R_M\|_{q_M}^2 \end{bmatrix} \leq \Gamma, \quad M\Gamma \leq \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

where inequalities are elementwise, c and M have nonnegative entries, and $G_{cl}(K)$ is the closed-loop LPTV operator as a function of the dynamic controller K . (In this formulation, the entries in Γ serve as upper bounds on each relevant ℓ_2 semi-norm and ℓ_2 induced norm.) A few special cases of this optimization formulation are \mathcal{H}_2 control, mixed $\mathcal{H}_2/\mathcal{H}_\infty$ control, \mathcal{H}_2 control with \mathcal{H}_2 constraints, and their counterparts for LPTV systems. This formulation is also able to handle much more general control design problems. For instance, it is possible to impose a constraint on the squared ℓ_2 semi-norm of a relevant input-output pair added to the squared ℓ_2 induced norm of a different input-output pair.

To solve this optimization problem, we will take the following approach: we will find an expression for the realization of the closed-loop LPTV operator, apply a change of variables, and finally change the realization of the operator so that the resulting matrix inequalities are all affine in a set of optimization parameters from which the optimal controller can be extracted.

2.5.1 Problem Formulation

First, we assume that our plant model is in the form

$$\begin{bmatrix} x_{k+1} \\ z_k \\ y_k \end{bmatrix} = \begin{bmatrix} A_k & B_k^1 & B_k^2 \\ C_k^1 & D_k^{11} & D_k^{12} \\ C_k^2 & D_k^{21} & 0 \end{bmatrix} \begin{bmatrix} x_k \\ w_k \\ u_k \end{bmatrix}$$

where z_k is the performance output of our system and w_k is the disturbance input of the system. Our controller will have the form

$$\begin{bmatrix} x_{k+1}^K \\ u_k^K \end{bmatrix} = \begin{bmatrix} A_k^K & B_k^K \\ C_k^K & D_k^K \end{bmatrix} \begin{bmatrix} x_k^K \\ y_k^K \end{bmatrix}.$$

A realization for the closed-loop system from w_k to z_k with the state vector $[x_k^T \ (x_k^K)^T]^T$ is

$$G_{cl} \sim \left(\begin{array}{c|c} A_k^{cl} & B_k^{cl} \\ \hline C_k^{cl} & D_k^{cl} \end{array} \right)$$

where

$$\left[\begin{array}{c|c} A_k^{cl} & B_k^{cl} \\ \hline C_k^{cl} & D_k^{cl} \end{array} \right] = \left[\begin{array}{c|c} A_k & 0 & B_k^1 \\ \hline 0 & 0 & 0 \\ \hline C_k^1 & 0 & D_k^{11} \end{array} \right] + \left[\begin{array}{c|c} 0 & B_k^2 \\ \hline I & 0 \\ \hline 0 & D_k^{12} \end{array} \right] \left[\begin{array}{c|c} A_k^K & B_k^K \\ \hline C_k^K & D_k^K \end{array} \right] \left[\begin{array}{c|c} 0 & I & 0 \\ \hline C_k^2 & 0 & D_k^{21} \end{array} \right]. \quad (2.13)$$

We now make the following partitions of Q_k from condition (3) of Theorems 6 and 10 and its inverse:

$$Q_k = \begin{bmatrix} Y_k & N_k \\ (N_k)^T & T_k \end{bmatrix} \quad Q_k^{-1} = \begin{bmatrix} X_k & M_k \\ (M_k)^T & R_k \end{bmatrix}. \quad (2.14)$$

This allows us to define the matrix

$$\Pi_k := \begin{bmatrix} X_k & I \\ (M_k)^T & 0 \end{bmatrix}$$

and the change of variables

$$\begin{bmatrix} \hat{A}_k & \hat{B}_k \\ \hat{C}_k & \hat{D}_k \end{bmatrix} := \begin{bmatrix} N_{k+1} & Y_{k+1} B_k^2 \\ 0 & I \end{bmatrix} \begin{bmatrix} A_k^K & B_k^K \\ C_k^K & D_k^K \end{bmatrix} \begin{bmatrix} (M_k)^T & 0 \\ C_k^2 X_k & I \end{bmatrix} + \begin{bmatrix} Y_{k+1} A_k X_k & 0 \\ 0 & 0 \end{bmatrix}. \quad (2.15)$$

Note that when $Q_k \succ 0$,

$$Q_k^{-1} = \begin{bmatrix} Y_k^{-1} + Y_k^{-1} N_k \Delta_k N_k^T Y_k^{-1} & -Y_k^{-1} N_k \Delta_k \\ -\Delta_k N_k^T Y_k^{-1} & \Delta_k \end{bmatrix}$$

where $\Delta_k = (S_k - N_k^T Y_k^{-1} N_k)^{-1}$. Thus, in this case, $\det M_k = 0 \Leftrightarrow \det N_k = 0$. Therefore, when $Q_k \succ 0$ and $\det M_k \neq 0$ we can invert the change of variables (2.15) by using

$$\begin{bmatrix} A_k^K & B_k^K \\ C_k^K & D_k^K \end{bmatrix} = \begin{bmatrix} N_{k+1}^{-1} & -N_{k+1}^{-1} Y_{k+1} B_k^2 \\ 0 & I \end{bmatrix} \begin{bmatrix} \hat{A}_k - Y_{k+1} A_k X_k & \hat{B}_k \\ \hat{C}_k & \hat{D}_k \end{bmatrix} \begin{bmatrix} (M_k)^{-T} & 0 \\ -C_k^2 X_k (M_k)^{-T} & I \end{bmatrix}. \quad (2.16)$$

We now present several identities. First, note that

$$Q_k \Pi_k = \begin{bmatrix} I & Y_k \\ 0 & (N_k)^T \end{bmatrix}.$$

Using this fact, it is straightforward to prove the following identities:

$$\begin{aligned}
(\Pi_{k+1})^T Q_{k+1} A_k^{cl} \Pi_k &= \begin{bmatrix} A_k X_k + B_k^2 \hat{C}_k & A_k + B_k^2 \hat{D}_k C_k^2 \\ \hat{A}_k & Y_{k+1} A_k + \hat{B}_k C_k^2 \end{bmatrix} \\
(\Pi_{k+1})^T Q_{k+1} B_k^{cl} &= \begin{bmatrix} B_k^1 + B_k^2 \hat{D}_k D_k^{21} \\ Y_{k+1} B_k^1 + \hat{B}_k D_k^{21} \end{bmatrix} \\
C_k^{cl} \Pi_k &= \begin{bmatrix} C_k^1 X_k + D_k^{12} \hat{C}_k & C_k^1 + D_k^{12} \hat{D}_k C_k^2 \end{bmatrix} \\
D_k &= \hat{D}_k \\
(\Pi_k)^T Q_k \Pi_k &= \begin{bmatrix} X_k & I \\ I & Y_k \end{bmatrix}.
\end{aligned}$$

The right-hand side of all of these identities are all affine in $X_k, Y_k, Y_{k+1}, \hat{A}_k, \hat{B}_k, \hat{C}_k, \hat{D}_k$ and all other matrices are known from the plant model. This lends itself to two theorems. The first of these theorems, Theorem 11, gives an upper bound on achievable closed loop squared ℓ_2 semi-norms. The second of these theorems, Theorem 12, gives an upper bound on achievable closed loop squared ℓ_2 induced norms.

Theorem 11. *If L^j and R^j are matrices and G is an LPTV operator, the following conditions are equivalent:*

1. $\exists A_k^K, B_k^K, C_k^K, D_k^K, W_k, Q_k$ for $k \in \{1, \dots, N\}$ such that M_k^{-1} exists and

$$\begin{aligned}
\frac{1}{N} \sum_{k=1}^N \text{tr} W_k < \gamma_j, \quad & \begin{bmatrix} W_k & L^j C_k^{cl} & L^j D_k^{cl} R^j \\ \bullet & Q_k & 0 \\ \bullet & \bullet & I \end{bmatrix} \succ 0, \\
& \begin{bmatrix} Q_{k+1} & Q_{k+1} A_k^{cl} & Q_{k+1} B_k^{cl} R^j \\ \bullet & Q_k & 0 \\ \bullet & \bullet & I \end{bmatrix} \succ 0
\end{aligned}$$

holds $\forall k \in \{1, \dots, N\}$ where Q_{N+1} is taken to be Q_1 , M_k is as defined in (2.14), and $A_k^{cl}, B_k^{cl}, C_k^{cl}, D_k^{cl}$ are as defined in (2.13).

2. $\exists W_k, X_k, Y_k, \hat{A}_k, \hat{B}_k, \hat{C}_k, \hat{D}_k$ for $k \in \{1, \dots, N\}$ such that

$$\frac{1}{N} \sum_{k=1}^N \text{tr} W_k < \gamma_j \quad (2.17a)$$

$$\begin{bmatrix} W_k & L^j (C_k^1 X_k + D_k^{12} \hat{C}_k) & L^j (C_k^1 + D_k^{12} \hat{D}_k C_k^2) & L^j (D_k^{11} + D_k^{12} \hat{D}_k D_k^{21}) R^j \\ \bullet & X_k & I & 0 \\ \bullet & \bullet & Y_k & 0 \\ \bullet & \bullet & \bullet & I \end{bmatrix} \succ 0 \quad (2.17b)$$

$$\begin{bmatrix} X_{k+1} & I & A_k X_k + B_k^2 \hat{C}_k & A_k + B_k^2 \hat{D}_k C_k^2 & (B_k^1 + B_k^2 \hat{D}_k D_k^{21}) R^j \\ \bullet & Y_{k+1} & \hat{A}_k & Y_{k+1} A_k + \hat{B}_k C_k^2 & (Y_{k+1} B_k^1 + \hat{B}_k D_k^{21}) R^j \\ \bullet & \bullet & X_k & I & 0 \\ \bullet & \bullet & \bullet & Y_k & 0 \\ \bullet & \bullet & \bullet & \bullet & I \end{bmatrix} \succ 0 \quad (2.17c)$$

holds $\forall k \in \{1, \dots, N\}$ where X_{N+1} and Y_{N+1} are respectively taken to be X_1 and Y_1 .

Proof. It is easy to see that condition (2) is equivalent to the existence of the relevant variables such that

$$\frac{1}{N} \sum_{k=1}^N \text{tr} W_k < \gamma_j$$

$$\begin{bmatrix} I & & & & \\ & \Pi_k^T & & & \\ & & I & & \end{bmatrix} \begin{bmatrix} W_k & L^j C_k^{cl} & L^j D_k^{cl} R^j \\ \bullet & Q_k & 0 \\ \bullet & \bullet & I \end{bmatrix} \begin{bmatrix} I & & & & \\ & \Pi_k & & & \\ & & I & & \end{bmatrix} \succ 0$$

$$\begin{bmatrix} \Pi_{k+1}^T & & & & \\ & \Pi_k^T & & & \\ & & I & & \end{bmatrix} \begin{bmatrix} Q_{k+1} & Q_{k+1} A_k^{cl} & Q_{k+1} B_k^{cl} R^j \\ \bullet & Q_k & 0 \\ \bullet & \bullet & I \end{bmatrix} \begin{bmatrix} \Pi_{k+1} & & & & \\ & \Pi_k & & & \\ & & I & & \end{bmatrix} \succ 0.$$

Since $\det \Pi_k = 0 \Leftrightarrow \det M_k = 0$, condition (2) is equivalent to condition (1). \square

Theorem 12. If L^j and R^j are matrices and G is an LPTV operator, the following conditions are equivalent:

1. $\exists A_k^K, B_k^K, C_k^K, D_k^K, W_k, Q_k$ for $k \in \{1, \dots, N\}$ such that M_k^{-1} exists and

$$\begin{bmatrix} \gamma I & 0 & L^j C_k^{cl} & L^j D_k^{cl} R^j \\ \bullet & Q_{k+1} & Q_{k+1} A_k^{cl} & Q_{k+1} B_k^{cl} R^j \\ \bullet & \bullet & Q_k & 0 \\ \bullet & \bullet & \bullet & I \end{bmatrix} \succ 0$$

holds $\forall k \in \{1, \dots, N\}$ where Q_{N+1} is taken to be Q_1 , M_k is as defined in (2.14), and $A_k^{cl}, B_k^{cl}, C_k^{cl}, D_k^{cl}$ are as defined in (2.13).

2. $\exists W_k, X_k, Y_k, \hat{A}_k, \hat{B}_k, \hat{C}_k, \hat{D}_k$ for $k \in \{1, \dots, N\}$ such that

$$\begin{bmatrix} \gamma I & 0 & 0 & L^j (C_k^1 X_k + D_k^{12} \hat{C}_k) & L^j (C_k^1 + D_k^{12} \hat{D}_k C_k^2) & L^j (D_k^{11} + D_k^{12} \hat{D}_k D_k^{21}) R^j \\ \bullet & X_{k+1} & I & A_k X_k + B_k^2 \hat{C}_k & A_k + B_k^2 \hat{D}_k C_k^2 & (B_k^1 + B_k^2 \hat{D}_k D_k^{21}) R^j \\ \bullet & \bullet & Y_{k+1} & \hat{A}_k & Y_{k+1} A_k + \hat{B}_k C_k^2 & (Y_{k+1} B_k^1 + \hat{B}_k D_k^{21}) R^j \\ \bullet & \bullet & \bullet & X_k & I & 0 \\ \bullet & \bullet & \bullet & \bullet & Y_k & 0 \\ \bullet & \bullet & \bullet & \bullet & \bullet & I \end{bmatrix} \succ 0$$

holds $\forall k \in \{1, \dots, N\}$ where X_{N+1} and Y_{N+1} are respectively taken be X_1 and Y_1 .

Proof. It is trivial to see that condition (2) is equivalent to the existence of the relevant variables such that

$$\begin{bmatrix} I \\ \Pi_{k+1}^T \\ \Pi_k^T \\ I \end{bmatrix} \begin{bmatrix} \gamma I & 0 & L^j C_k^{cl} & L^j D_k^{cl} R^j \\ \bullet & Q_{k+1} & Q_{k+1} A_k^{cl} & Q_{k+1} B_k^{cl} R^j \\ \bullet & \bullet & Q_k & 0 \\ \bullet & \bullet & \bullet & I \end{bmatrix} \begin{bmatrix} I \\ \Pi_{k+1} \\ \Pi_k \\ I \end{bmatrix} \succ 0.$$

Since $\det \Pi_k = 0 \Leftrightarrow \det M_k = 0$, condition (2) is equivalent to condition (1). \square

The conditions in these theorems are very close to the condition that there exists a controller such that

$$\|L^j G_{cl} R^j\|_{q_j}^2 < \gamma_j \quad (2.18)$$

is satisfied where q_j is either ‘2’ or ‘ ∞ ’. Clearly, if the conditions of the appropriate theorem are satisfied, then (2.18) must be satisfied. The only conservatism in these theorems is the constraint that the controller must allow M_k to be invertible for some choice of Q_k which satisfies condition (1) of either theorem. To understand what conservatism this introduces, we first look at the LTI case, i.e. when $N = 1$. When M_1 is not invertible, it is always possible to change the realization of the controller such that the corresponding \bar{Q}_1 has the form

$$\bar{Q}_1 = \begin{bmatrix} \hat{Q}_1 \\ \tilde{Q}_1 \end{bmatrix}$$

where the dimension of \tilde{Q}_1 is less than the dimension of x_k^K . It is possible to show that in this controller realization, it is possible to throw away the states corresponding to \tilde{Q}_1 without any loss in closed-loop performance with respect to either the ℓ_2 semi-norm or the ℓ_2 induced norm. Thus, from this, we can extrapolate the interpretation for the LPTV case. The constraint that the controller must admit a choice of Q_k which gives invertible M_k means that we do not allow controllers with states at particular time steps which can be thrown away with no loss in performance. This is a reasonable assumption in the LPTV case because throwing away states at particular time steps would result in non-square A_k matrices.

With these theorems in mind, we can therefore construct an SDP corresponding to the control design problem (2.12) by using condition (2) of these theorems. This SDP approach will, in general, be a little bit conservative for two reasons. First, we require the values of X_k and Y_k to be the same for each set of relevant matrix inequalities. This corresponds to using the same values of Q_k in each of the inequalities in condition (3) of Theorems 6 and 10. Second, we can only consider closed-loop systems which admit a Q_k such that M_k is invertible, as discussed in the previous paragraph.

Once we solve this SDP, we need to reconstruct our controller from the optimal values of $\hat{A}_k, \hat{B}_k, \hat{C}_k, \hat{D}_k, X_k, Y_k$. However, in order to apply the inverse change of variables (2.16), we need to have values for N_k and M_k . Since condition (2) of Theorems 11 and 12 imply that

$$\begin{aligned} & \begin{bmatrix} X_k & I \\ I & Y_k \end{bmatrix} \succ 0 \\ \Rightarrow & X_k - Y_k^{-1} \succ 0, \quad Y_k \succ 0 \\ \Rightarrow & \det(X_k Y_k - I) \det(Y_k^{-1}) \neq 0 \end{aligned}$$

we know that $I - X_k Y_k$ is nonsingular. Also, from examining the upper left block of $Q_k^{-1} Q_k$, we know that

$$\begin{aligned} & X_k Y_k + M_k (N_k)^T = I \\ \Rightarrow & M_k (N_k)^T = I - X_k Y_k. \end{aligned}$$

It is easily shown that whenever M_k and N_k are chosen to meet this condition, it is possible to choose R_k and T_k such that

$$\begin{bmatrix} X_k & M_k \\ (M_k)^T & R_k \end{bmatrix} \begin{bmatrix} Y_k & N_k \\ (N_k)^T & T_k \end{bmatrix} = I$$

i.e. there are no additional constraints that need to be met when choosing values of M_k and N_k . Thus, we can find M_k and N_k using any matrix factorization technique, such as the QR decomposition. Then, once we've chosen values of M_k and N_k , we can reconstruct our controller (i.e. $A_k^K, B_k^K, C_k^K, D_k^K$) using (2.16).

2.5.2 Heuristics for Improved Numerics

In some cases, the quantity $X_k Y_k - I$ is ill-conditioned. In this case, there can be large numerical errors when we try to reconstruct the optimal controller by using (2.16). Thus, we would like some heuristics to improve the conditioning of the controller reconstruction. We will present two heuristics which both operate

using the same methodology:

1. Solve the SDP's associated with the control design problem.
2. Pick a slightly sub-optimal value of the objective function and add the constraint to the SDP that the objective function must be less than this value.
3. Redefine the cost function of the SDP so that it is trying to improve the numerics of the system. Add any related variables and constraints to the SDP.
4. Solve the new SDP and reconstruct the controller from these optimal values.

The two heuristics presented in this report will only differ in the details of step (3) of this methodology.

We begin by looking at a heuristic presented in [11]. Consider the matrix inequality

$$\begin{bmatrix} X_k & tI \\ tI & Y_k \end{bmatrix} \succ 0. \quad (2.19)$$

Note that using Schur complements this is equivalent to

$$\begin{aligned} X_k - t^2 Y_k^{-1} &\succ 0, \quad Y_k \succ 0 \\ \Leftrightarrow Y_k^{1/2} X_k Y_k^{1/2} &\succ t^2 I, \quad Y_k \succ 0. \end{aligned}$$

Now let λ_j be an eigenvalue of $Y_k^{1/2} X_k Y_k^{1/2}$. Then, since $Y_k \succ 0$

$$\begin{aligned} \det \left(Y_k^{1/2} X_k Y_k^{1/2} - \lambda_j I \right) &= 0 \\ \Leftrightarrow \det \left[Y_k^{1/2} (X_k Y_k - \lambda_j I) Y_k^{-1/2} \right] &= 0 \\ \Leftrightarrow \det (X_k Y_k - \lambda_j I) &= 0. \end{aligned}$$

From this we can say that the eigenvalues of $X_k Y_k$ are the same as the eigenvalues of $Y_k^{1/2} X_k Y_k^{1/2}$, which are real. Also, we can say that

$$\min_j \lambda_j = \sup_t t^2 \text{ s.t. } \begin{bmatrix} X_k & tI \\ tI & Y_k \end{bmatrix} \succ 0.$$

Thus, for this heuristic, step (3) in the methodology outlined above corresponds to adding the variable t and the constraint (2.19) to the SDP and redefining the cost function to be $-t$ (i.e. we maximize $+t$). This heuristic has the effect of pushing the minimum eigenvalue of $X_k Y_k$ as far away from 1 as possible.

However, what we want is not for the smallest eigenvalue of $X_k Y_k$ to be large, but for $X_k Y_k - I$ to be well-conditioned. Thus, a good heuristic would also make use of an upper bound. With this in mind, we now construct the second heuristic for improving the numerics of the controller reconstruction. Consider the matrix inequality

$$\begin{bmatrix} 2z & X_k + Y_k \\ X_k + Y_k & z \end{bmatrix} \succ 0. \quad (2.20)$$

Clearly this implies that

$$\begin{aligned} & 2z^2 I - (X_k + Y_k)^2 \succ 0 \\ \Rightarrow & 2z^2 I - X_k Y_k - Y_k X_k \succ X_k^2 + Y_k^2 \succ 0 \\ \Rightarrow & (X_k Y_k - z^2 I) + (X_k Y_k - z^2 I)^T \prec 0. \end{aligned}$$

Note that this last equation looks like a Lyapunov equation for a stable continuous time LTI system. Thus, since the eigenvalues of $X_k Y_k$ are real, we can say that the eigenvalues of $X_k Y_k - z^2 I$ are all strictly negative. Thus, the maximum eigenvalue of $X_k Y_k$ is less than z^2 . Thus, for this heuristic, step (3) in the methodology outlined above corresponds to adding two variables (t and z) and three constraints ((2.19), (2.20), and $t > 1$) to the SDP and redefining the cost function to be $\alpha z - t$, where $\alpha > 0$ is a design parameter which trades off the importance of minimizing the upper bound on $X_k Y_k$ and maximizing the lower bound on $X_k Y_k$. This heuristic has the effect of squeezing the maximum and minimum eigenvalues of $X_k Y_k$ together to make it more well-conditioned. This, in turn, squeezes the maximum and minimum eigenvalues of $X_k Y_k - I$ together to make it more well-conditioned. Through experience, it was found that a reasonable choice of α is 10^{-10} .

Chapter 3

Track-Following Control of Dual-Stage Hard Disk Drives

For several decades now, the areal storage density of hard drives has been doubling every 18 months, as predicted by Moore's law. The current areal storage density of hard drives, as reported by Hitachi GST, is 345 gigabits/in² [4]. As the storage density is pushed higher, the concentric tracks on the disk which contain data must be pushed closer together, which requires much more accurate control of the read/write head. This report will consider track-following control of the read/write head.

The current goal of the magnetic recording industry is to achieve an areal storage density of 1 terabit/in². The predicted track width required to achieve this data density is 46 nm. A good rule of thumb for track-following control design is that three times the standard deviation of the position error signal (PES) between the read/write head and the track center should be less than 10% of the track width [7]. Thus, the servo design goal is to keep the 3σ value of the PES less than 4.6 nm.

There are several types of disturbances which act on the system. The first type, torque disturbances, includes D/A quantization noise, power-amp noise, bearing imperfection and nonlinearity, and the effects of high-frequency airflow turbulence. The second type, track runout, includes disk flutter, eccentricity due to disk slippage, and imperfection of track circles. The third type, noises, includes demodulation noise and A/D quantization noise [5].

A schematic of a typical hard drive is shown in Figure 3.1. A traditional hard drive setup uses only the voice coil motor (VCM) to control the position of the read/write head. However, this places a strong limitation on the closed-loop performance that can be achieved. In an effort to improve the achievable closed-loop performance, we consider dual-stage hard drives, i.e. ones which have a secondary actuator. The

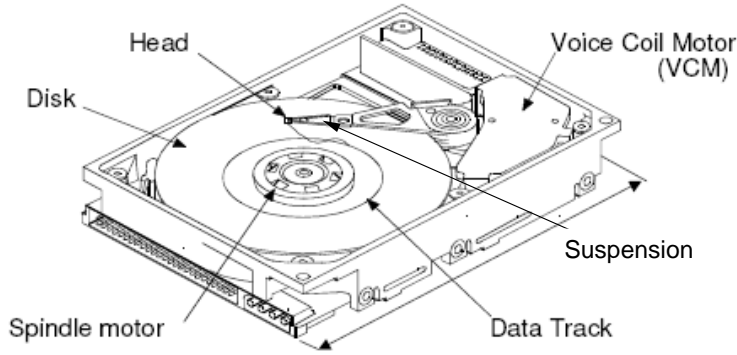


Figure 3.1: Hard drive schematic

secondary actuator considered here is a microactuator (MA) which directly actuates the slider upon which the read/write head sits [8]. This MA has a built-in capacitive sensor which measures the MA displacement. This signal is also called the relative position error signal (RPES). In addition, this system uses two PZT sensors which are symmetrically placed on the suspension to measure the strain in the suspension which corresponds to off-track head motion [6].

3.1 Modeling

Figure 3.2 shows the block diagram of the continuous time hard drive model where each signal is as described in Table 3.1. Here, w_v captures the effect of airflow turbulence and w_r captures the rest of the disturbances

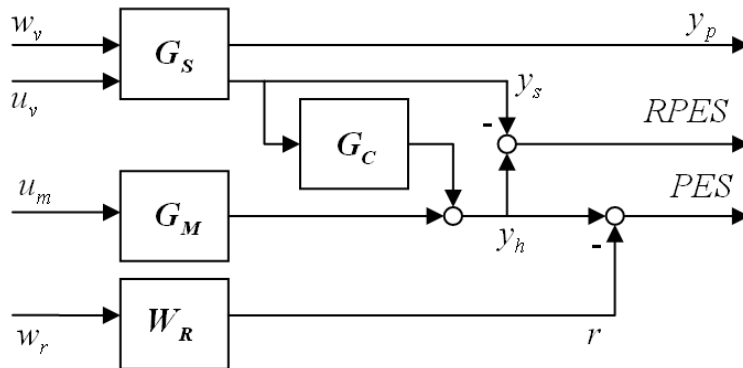


Figure 3.2: Block diagram of the hard drive model

mentioned in the previous section. Also, w_v and w_r have been normalized so that their variance is 1. The suspension strain measurements have been scaled so that they correspond to off-track displacement of the

Table 3.1: Summary of signals and units in Figure 3.2

Signal	Description	Units
u_v	Control input into VCM	V
u_m	Control input into MA	mV
w_v	Airflow disturbance on VCM	normalized
w_r	White noise which drives W_R to produce runout signal	normalized
r	Runout signal	nm
y_s	Suspension tip displacement	nm
y_h	Head position relative to track center	nm
y_p	Uncontaminated suspension strain measurement	nm
$RPES$	Uncontaminated RPES measurement	nm
PES	Uncontaminated PES measurement	nm

head. The models for each block are given by

$$\begin{aligned}
G_S(s) &:= \sum_{j=1}^3 \left(M_j^{out} \frac{\omega_j^2}{s^2 + 2\zeta_j \omega_j s + \omega_j^2} M_j^{in} \right) \\
G_M(s) &:= \frac{\omega_0^2}{s^2 + 2\zeta_0 \omega_0 s + \omega_0^2} \\
G_C(s) &:= \frac{2\zeta_0 \omega_0 s + \omega_0^2}{s^2 + 2\zeta_0 \omega_0 s + \omega_0^2} \\
W_R(s) &:= \frac{(12 \times 10^4)s^2 + (2.89 \times 10^9)s + (5.298 \times 10^{12})}{s^3 + (2.684 \times 10^3)s^2 + (1.756 \times 10^6)s + (4.703 \times 10^8)}
\end{aligned}$$

where the model parameters are specified in Table 3.2 and G_S is regarded as taking $[w_v \quad u_v]^T$ as its inputs and producing $[y_p \quad y_s]^T$ as its outputs. It is assumed here that w_v and w_r are both zero-mean Gaussian

Table 3.2: Hard drive model parameters

j	Parameter			
	ζ_j	ω_j	M_j^{out}	M_j^{in}
0	0.2	1.4137×10^4	N/A	N/A
1	0.05	376.99	$[0 \quad 10^4]^T$	$[0 \quad 1.4]$
2	0.015	4.6496×10^4	$[0 \quad 1]^T$	$[0.48 \quad -0.84]$
3	0.015	6.7218×10^4	$[0.96 \quad -0.5]^T$	$[-0.816 \quad 0.6]$

white noises. The relevant frequency responses of the model are shown in Figures 3.3 and 3.4.

At this point, we put the model into a form more suitable for control. We assumed that our measurements of the PES, the RPES and y_p are respectively contaminated by the sensor noises n_{PES} , n_{RPES} , and n_p , which are all zero-mean white Gaussian noises with standard deviation of 1 nm. With a bit of manipulation, the model was put into the form

$$\begin{bmatrix} z \\ y \end{bmatrix} = G_{ct}(s) \begin{bmatrix} w \\ u \end{bmatrix}$$

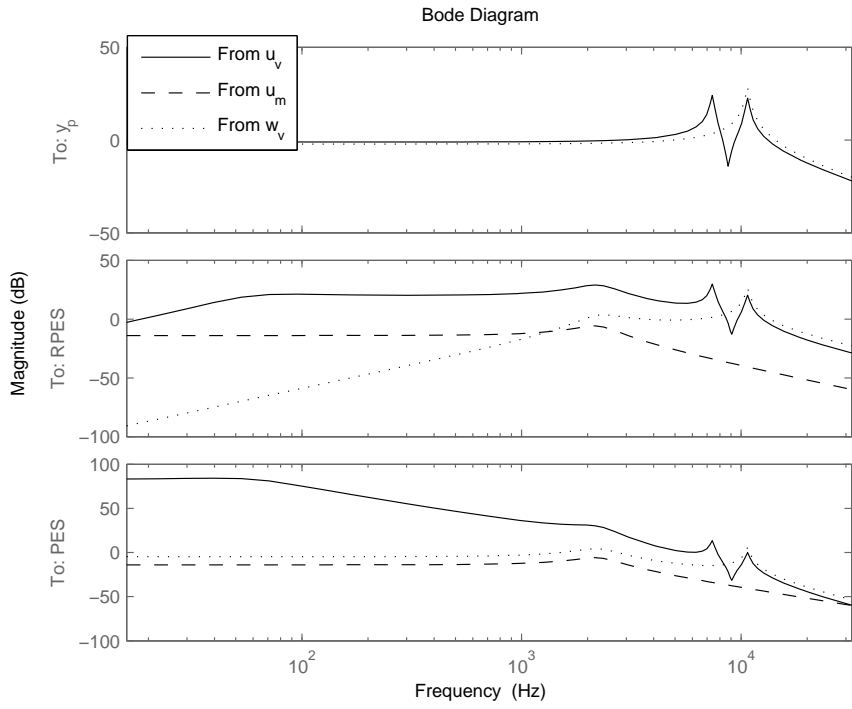


Figure 3.3: Frequency response of the hard drive model

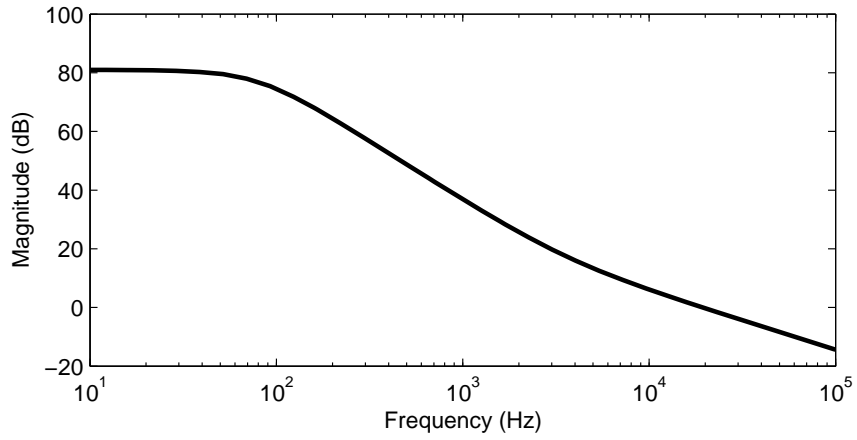


Figure 3.4: Frequency response of the runout model

where the performance outputs, z , measurements, y , disturbances, w , and control inputs, u are given by

$$z := \begin{bmatrix} PES \\ RPES \\ u_v \\ u_m \end{bmatrix} \quad w := \begin{bmatrix} w_r \\ w_v \\ n_{PES} \\ n_{RPES} \\ n_p \end{bmatrix}$$

$$y := \begin{bmatrix} PES + n_{PES} \\ RPES + n_{RPES} \\ y_p + n_p \end{bmatrix} \quad u := \begin{bmatrix} u_v \\ u_m \end{bmatrix}.$$

Note that w is zero-mean white Gaussian noise whose covariance matrix is I , so the \mathcal{H}_2 norm from w to any element of z can be interpreted as the standard deviation of that element of z .

Since we only have access to our PES measurements at the rate of 25 kHz, we must discretize our system. We will consider two types of sampling and actuation schemes in this report. In the first scheme, we discretize the system using a zero-order hold at 25 kHz and assume that all sampling and actuation occurs at that frequency. In this case, since there is no need for a multi-rate sampler or a multi-rate hold, the resulting discrete time system is LTI. In the second scheme, the system is discretized at 50 kHz using a zero-order hold. In this case, we assume that we only have measurements of the PES for $k \in \{\dots, -1, 1, 3, \dots\}$ whereas all other signals can be sampled and actuated at all time steps. Thus, we need a multi-rate sampler, but not a multi-rate hold. The multi-rate sampler in this case is given by

$$S_k = \begin{cases} I, & k - \text{odd} \\ \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, & k - \text{even} \end{cases}.$$

Incorporating S_k into the discrete time plant model in this case results in an LPTV system with period 2.

3.2 Control Design

This section will consider the control design for the hard drive model described in the previous section. The goal of the control design is to minimize the PES while keeping u_v , u_m , and the RPES reasonably small. In

particular, we would like the following inequalities to hold:

$$\begin{bmatrix} \sigma_{RPES} \\ \sigma_{u_v} \\ \sigma_{u_m} \end{bmatrix} \leq \frac{1}{3} \begin{bmatrix} 100 \text{ nm} \\ 5 \text{ V} \\ 20 \text{ V} \end{bmatrix}.$$

In the LPTV case, σ refers to the RMS standard deviation over time. We want the 3σ value of the RPES to be smaller than 100 nm because that's the range in which the capacitive sensor works best. This also guarantees that the MA doesn't saturate. We want the 3σ values of u_v and u_m to be respectively smaller than 5 V and 20 V because those are the maximum allowable control inputs. This report will compare two approaches to meeting these closed-loop specifications. The first method applies a static weight to the elements of z and minimizes the ℓ_2 semi-norm from w to z . In particular, we solve

$$\min \left\| \begin{bmatrix} 1 & & & \\ & 0 & & \\ & & 1 & \\ & & & 0.01 \end{bmatrix} G_{cl} \right\|_2^2.$$

This will be called the unconstrained control design. The second method uses the multi-objective formulation to minimize the ℓ_2 semi-norm from w to the PES while maintaining the inequalities

$$\begin{bmatrix} \|[0 \ 1 \ 0 \ 0]G_{cl}\|_2 \\ \|[0 \ 0 \ 1 \ 0]G_{cl}\|_2 \\ \|[0 \ 0 \ 0 \ 1]G_{cl}\|_2 \end{bmatrix} \leq \frac{1}{3} \begin{bmatrix} 100 \text{ nm} \\ 5 \text{ V} \\ 20 \text{ V} \end{bmatrix}.$$

This will be called the constrained or multi-objective design.

For the constrained design, there is a simplification that arises in the SDP formulation of the problem. When we formulate the problem, we have four identical copies of the matrix inequality (2.17c). Thus, we can reduce the number of matrix inequalities that we need to satisfy. In addition, it is easy to show using the methodology of Appendix A.3 that there is no conservatism introduced by requiring Q_k to be the same for all of the required inequalities. Thus, the control design methodology described in Section 2.5 will be much less conservative when applied to this problem than in the general case.

Thus, since we have two types of sampling/actuation schemes and two types of control design formulations, we have four controller designs to consider. These four optimizations were carried out using the function `mincx`, which is a part of the Robust Control Toolbox for MATLAB. The results of all four controller designs are summarized in Table 3.3. Note that all four designs meet the desired constraints on the RMS 3σ values of u_v , u_m , and the RPES. (In the unconstrained case, this just means that we picked a reasonable

Table 3.3: RMS 3σ values for signals in z for each controller design

Design Type		PES (nm)	$RPES$ (nm)	u_v (V)	u_m (V)
Single Rate	Unconstrained	23.19	65.82	3.66	0.96
	Constrained	21.18	100.00	5.00	2.61
Multi-Rate	Unconstrained	15.54	45.45	2.55	0.75
	Constrained	12.93	100.00	5.00	5.43

weighting for the elements in z .)

At this point, we compare the constrained control design results to the unconstrained control design results. The first thing to note is that for both sampling/actuation schemes, the constrained control design did significantly better than the corresponding unconstrained control design in terms of the 3σ PES value. This improved PES performance, however, came at the cost of saturating u_v and the RPES. Interestingly enough, u_m did not saturate. The reason for this is that increasing the input into the MA tends to increase the MA displacement, i.e increasing u_m tends to increase the RPES. Thus, at a certain point, u_m can't be increased any further without violating the RPES constraint. This is why the constrained control designs are unable to further increase performance by increasing u_m .

Now we examine the effect of introducing multi-rate sampling and actuation into the control design. In the multi-rate case, since the controller measures y_p and the $RPES$ twice as often, it has more information about the plant. Also, since it can change the value of u_v and u_m twice as often, it has more control over the plant. Therefore, we expect dramatic closed-loop performance improvements in the multi-rate designs. This is verified in Table 3.3. However, since the unconstrained design and constrained design have very different interpretations of "performance," introducing multi-rate sampling and actuation has a very different effect on the two designs. In the unconstrained design, the cost function is a trade-off between the RMS standard deviations of the relevant signals. Thus, the optimization tries to reduce the values of those signals in a uniform way. In this particular example, the RMS standard deviation of *all* of those signals is reduced. However, in the constrained design, the cost function includes *only* the PES RMS standard deviation. Thus, the optimization only tries to reduce this value. Thus the PES performance increase is more dramatic in the constrained design. An interesting effect to note is that RMS 3σ value of u_m actually *increases* when the multi-rate sampling and actuation is introduced into the design. One way to interpret this is that the controller uses some of its additional control over the plant to keep the RPES within its bound while applying increased u_m .

Thus, in summary, the multi-rate and constrained designs respectively achieved better closed-loop performance than the single-rate and unconstrained designs. And, when the constrained multi-rate design was used, a 44% improvement of the RMS PES standard deviation was achieved.

Chapter 4

Conclusion

This report presented a framework and methodology for multi-objective controller design for discrete-time LPTV systems. The objectives could include squared ℓ_2 semi-norms (which were shown to generalize the \mathcal{H}_2 norm) and squared ℓ_2 induced norms (which were shown to generalize the \mathcal{H}_∞ norm). The theory for the ℓ_2 semi-norm and the ℓ_2 induced norm was rigorously developed and their interpretation in the LPTV case was discussed.

This control design methodology was then applied to track-following control of dual-stage hard drives. Two types of sampling and actuation schemes were considered. The first, a single rate scheme in which all sampling and actuation was performed at the same frequency, could use LTI control design techniques. The second scheme, which used multi-rate sampling and actuation required use of LPTV control design techniques. It was shown that the multi-rate scheme achieve much better closed-loop performance than the single rate scheme. Also, two types of control design objectives were considered. The first type was an unconstrained design which minimized the ℓ_2 semi-norm of the closed-loop system with a static weight on the its output. The second type was a constrained design which minimized the RMS PES variance subject to the RPES and the control inputs being “small enough.” This constrained design, which is a multi-objective control problem, achieved better closed-loop performance than the unconstrained (i.e. single objective) design. It was also shown that using the constrained design with the multi-rate sampling and actuation resulted in a 44% decrease in RMS PES standard deviation.

Bibliography

- [1] Brian D.O. Anderson. Internal and external stability of linear time-varying systems. *SIAM Journal on Control and Optimization*, 20(3):408–413, May 1982.
- [2] R.E. Crochiere and L.R. Rabiner. Interpolation and decimation of digital signals—a tutorial review. *Proceedings of the IEEE*, 69(3):300–331, March 1981.
- [3] G.E. Dullerud and S. Lall. A new approach for analysis and synthesis of time-varying systems. *IEEE Transactions on Automatic Control*, 44(8):1486–1497, August 1999.
- [4] Hitachi GST. Perpendicular magnetic recording, 2006. http://www.hitachigst.com/hdd/research/recording_head/pr/index.html.
- [5] X. Huang, R. Nagamune, and R. Horowitz. A comparison of multirate robust track-following control synthesis techniques for dual-stage and multi-sensing servo systems in hard disk drives. In *Proceedings of the 2006 American Control Conference*, June 2006.
- [6] S. Kon, K. Oldham, R. Ruzicka, and R. Horowitz. Design and fabrication of a piezoelectric instrumented suspension for hard disk drives. In *Proceedings of SPIE, Smart Structures and Materials 2006: Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems*, 2006.
- [7] Yunfeng Li. *Dual-Stage Servo Control and Active Vibration Compensation in Magnetic Hard Disk Drives*. PhD thesis, University of California, Berkeley, 2003.
- [8] K. Oldham, X. Huang, and R. Horowitz. Design, fabrication, and control of a high-aspect ratio microactuator for vibration suppression in a hard disk drive. In *Proceedings of the 16th IFAC World Congress*, July 2005.
- [9] M. C. De Oliveira, J. C. Geromel, and J. Bernussou. Extended H_2 and H_∞ norm characterizations and controller parametrizations for discrete-time systems. *International Journal of Control*, 75(9):666–679, June 2002.

- [10] M. A. Peters and P. A. Iglesias. *Minimum Entropy Control for Time-Varying Systems*. Birkhäuser, Boston, 1997.
- [11] C. Scherer, P. Gahinet, and M. Chilali. Multiobjective output-feedback control via LMI optimization. *IEEE Transactions on Automatic Control*, 42(7):896–911, July 1997.

Appendix A

Additional Proofs

A.1 Proof of Lemma 2

Lemma. *Suppose that a realization given by (2.1) is UES and a given sequence, W_k , satisfies*

$$W_k \succeq 0, \quad \|W_k\| \leq M \quad \forall k$$

for some scalar $M > 0$. Then the equation

$$L_{k+1} = A_k L_k A_k^T + W_k \tag{A.1}$$

has a unique bounded solution given by

$$L_k = W_{k-1} + \sum_{j=1}^{\infty} (A_{k-1} \cdots A_{k-j}) W_{k-j-1} (A_{k-1} \cdots A_{k-j})^T. \tag{A.2}$$

Moreover, $L_k \succeq 0$.

Proof. Define the sequence

$$\begin{aligned} L_k^{[0]} &:= W_{k-1} \\ L_k^{[i]} &:= A_{k-1} L_{k-1}^{[i-1]} A_{k-1}^T + W_{k-1}. \end{aligned}$$

It is easy to see that

$$L_k^{[i]} = W_{k-1} + \sum_{j=1}^i (A_{k-1} \cdots A_{k-j}) W_{k-j-1} (A_{k-1} \cdots A_{k-j})^T$$

and, moreover, that $L_k^{[i]} \succeq 0$, $\forall i \geq 0, k$. Fixing $l > i$ gives

$$\begin{aligned} \left\| L_k^{[l]} - L_k^{[i]} \right\| &= \left\| \sum_{j=i+1}^l (A_{k-1} \cdots A_{k-j}) W_{k-j-1} (A_{k-1} \cdots A_{k-j})^T \right\| \\ &\leq \sum_{j=i+1}^l \|A_{k-1} \cdots A_{k-j}\|^2 \|W_{k-j-1}\|. \end{aligned}$$

Using the assumption that the realization is UES then gives

$$\begin{aligned} \left\| L_k^{[l]} - L_k^{[i]} \right\| &\leq c^2 M \sum_{j=i+1}^l (\beta^2)^j \\ &\leq \frac{c^2 M \beta^{2(i+1)}}{1 - \beta^2}. \end{aligned}$$

Thus $\forall \epsilon > 0 \exists N \in \mathbb{N}$ such that $i, j \geq N \Rightarrow \|L_k^{[j]} - L_k^{[i]}\| < \epsilon$. Therefore, this sequence converges, which in turn implies

$$\lim_{i \rightarrow \infty} \left(L_{k+1}^{[i]} - A_k L_k^{[i]} A_k^T - W_k \right) = 0$$

i.e. this sequences converges to a solution of (A.1). We now denote

$$L_k = \lim_{i \rightarrow \infty} L_k^{[i]}.$$

Clearly

$$\begin{aligned} \|L_k\| &\leq \|W_{k-1}\| + \sum_{j=1}^{\infty} \left\| (A_{k-1} \cdots A_{k-j}) W_{k-j-1} (A_{k-1} \cdots A_{k-j})^T \right\| \\ &\leq M + \frac{c^2 M \beta^2}{1 - \beta^2} \end{aligned}$$

i.e. this solution is bounded. It only remains to show that this solution of (A.1) is the unique bounded

solution. Let \bar{L}_k be any other solution. Then

$$\begin{aligned} L_k - \bar{L}_k &= A_{k-1} (L_{k-1} - \bar{L}_{k-1}) A_{k-1}^T \\ &= (A_{k-1} \cdots A_{k-j}) (L_{k-j} - \bar{L}_{k-j}) (A_{k-1} \cdots A_{k-j})^T \\ \Rightarrow \|L_k - \bar{L}_k\| &\leq c^2 \beta^{2j} \|L_{k-j} - \bar{L}_{k-j}\|. \end{aligned}$$

Thus

$$\begin{aligned} \|L_{k-j}\| + \|\bar{L}_{k-j}\| &\geq \|L_{k-j} - \bar{L}_{k-j}\| \\ \Rightarrow \|\bar{L}_{k-j}\| &\geq \frac{\|L_k - \bar{L}_k\|}{c^2 \beta^{2j}} - \|L_{k-j}\|. \end{aligned}$$

Since L_{k-j} is bounded, this implies that either $L_k = \bar{L}_k$ or \bar{L}_{k-j} becomes unbounded as $j \rightarrow \infty$. \square

A.2 Proof of Lemma 5

Lemma. *An LPTV realization is UES $\Leftrightarrow \rho(\mathcal{Z}^T \mathcal{A}) < 1$, where \mathcal{Z} and \mathcal{A} are as in (2.9).*

Proof. (\Rightarrow)

A realization is UES if $\exists c > 0, \beta \in [0, 1)$ such that

$$\sup_{k \in \mathbb{Z}} \|A_k \cdots A_{k-l+1}\| \leq c \beta^l, \quad \forall l \in \mathbb{N}.$$

In the LPTV case, it suffices to replace \mathbb{Z} with $\{1, \dots, N\}$. To write this in a more convenient form, note that two important properties of \mathcal{Z} are

$$\begin{aligned} \mathcal{Z} \mathcal{Z}^T &= \mathcal{Z}^T \mathcal{Z} = I \\ \|\mathcal{M} \mathcal{Z}\| &= \|\mathcal{M}\|. \end{aligned}$$

(The second of these is proved by using the first property and the submultiplicative property of the norm.)

The UES condition can now be written

$$\begin{aligned}
& \left\| (\mathcal{Z}^T \mathcal{A} \mathcal{Z}) ((\mathcal{Z}^T)^2 \mathcal{A} \mathcal{Z}^2) \cdots ((\mathcal{Z}^T)^l \mathcal{A} \mathcal{Z}^l) \right\| \leq c \beta^l, \quad \forall l \in \mathbb{N} \\
& \Rightarrow \left\| (\mathcal{Z}^T \mathcal{A})^l \mathcal{Z}^l \right\| \leq c \beta^l, \quad \forall l \in \mathbb{N} \\
& \Rightarrow c^{-1/l} \left\| (\mathcal{Z}^T \mathcal{A})^l \right\|^{1/l} \leq \beta, \quad \forall l \in \mathbb{N} \\
& \Rightarrow \rho(\mathcal{Z}^T \mathcal{A}) = \lim_{l \rightarrow \infty} \left\| (\mathcal{Z}^T \mathcal{A})^l \right\|^{1/l} \leq \beta < 1.
\end{aligned}$$

(\Leftarrow)

Define

$$\begin{aligned}
\alpha & := \rho(\mathcal{Z}^T \mathcal{A}) = \lim_{l \rightarrow \infty} \left\| (\mathcal{Z}^T \mathcal{A})^l \right\|^{1/l} < 1 \\
\beta & := \frac{\alpha + 1}{2}.
\end{aligned}$$

Since $\beta \in (\alpha, 1)$, we know that $\exists N$ such that $\left\| (\mathcal{Z}^T \mathcal{A})^l \right\|^{1/l} < \beta, \forall l > N$. Now define

$$c := \max \left\{ 1, \max_{l \in \{1, \dots, N\}} \frac{\left\| (\mathcal{Z}^T \mathcal{A})^l \right\|^{1/l}}{\beta^l} \right\}.$$

It is easily checked that these choices of β and c imply that the realization is UES. □

A.3 Proof of Theorem 6

Theorem. *For LPTV operators, the following conditions are equivalent:*

1. $\|G\|_2^2 < \gamma$
2. $\exists W_1, \dots, W_N, P_1, \dots, P_N$ such that

$$\frac{1}{N} \sum_{k=1}^N \text{tr} W_k < \gamma, \quad \begin{bmatrix} W_k & C_k P_k & D_k \\ \bullet & P_k & 0 \\ \bullet & \bullet & I \end{bmatrix} \succ 0, \quad \begin{bmatrix} P_{k+1} & A_k P_k & B_k \\ \bullet & P_k & 0 \\ \bullet & \bullet & I \end{bmatrix} \succ 0$$

holds $\forall k \in \{1, \dots, N\}$ where P_{N+1} is taken to be P_1 .

3. $\exists W_1, \dots, W_N, Q_1, \dots, Q_N$ such that

$$\frac{1}{N} \sum_{k=1}^N \text{tr} W_k < \gamma, \quad \begin{bmatrix} W_k & C_k & D_k \\ \bullet & Q_k & 0 \\ \bullet & \bullet & I \end{bmatrix} \succ 0, \quad \begin{bmatrix} Q_{k+1} & Q_{k+1}A_k & Q_{k+1}B_k \\ \bullet & Q_k & 0 \\ \bullet & \bullet & I \end{bmatrix} \succ 0$$

holds $\forall k \in \{1, \dots, N\}$ where Q_{N+1} is taken to be Q_1 .

In these conditions, a bullet represents an entry in the matrix which follows from symmetry.

Proof. (1) \Rightarrow (2):

In this part of the proof, we will construct the necessary variables which satisfy the given conditions in (2).

First we define

$$\epsilon := \gamma - \|G\|_2^2 > 0.$$

Then, we define L_k and P_k^I to satisfy

$$L_{k+1} := A_k L_k A_k^T + B_k B_k^T$$

$$P_{k+1}^I := A_k P_k^I A_k^T + I.$$

By (2.4), $L_k \succeq 0$ and $P_k^I \succ 0$. Now we define

$$\alpha := \frac{\epsilon}{4 \max\{1, \text{tr}\{C_k P_k^I C_k^T\}\}} > 0$$

$$P_k := L_k + \alpha P_k^I \succ 0$$

$$W_k := C_k P_k C_k^T + D_k D_k^T + \frac{\epsilon}{4n} I$$

where n is the number of rows (and columns) in the A_k matrix. Note that P_k satisfies

$$P_{k+1} = A_k P_k A_k^T + B_k B_k^T + \alpha I$$

$$\Rightarrow P_{k+1} - A_k P_k A_k^T - B_k B_k^T \succ 0$$

$$\Rightarrow \begin{bmatrix} P_{k+1} & A_k P_k & B_k \\ \bullet & P_k & 0 \\ \bullet & \bullet & I \end{bmatrix} \succ 0.$$

The last implication followed using Schur complements. Also note that P_k and W_k satisfy

$$\begin{aligned} W_k - C_k P_k C_k^T - D_k D_k^T &\succ 0 \\ \Rightarrow \begin{bmatrix} W_k & C_k P_k & D_k \\ \bullet & P_k & 0 \\ \bullet & \bullet & I \end{bmatrix} &\succ 0 \end{aligned}$$

by Schur complements. Finally, note that

$$\begin{aligned} \sum_{k=1}^N \text{tr} W_k &= \sum_{k=1}^N \text{tr} \left\{ \alpha C_k P_k^I C_k^T + C_k L_k C_k^T + D_k D_k^T + \frac{\epsilon}{4n} I \right\} \\ &= \sum_{k=1}^N \text{tr} \left\{ \alpha C_k P_k^I C_k^T + C_k L_k C_k^T + D_k D_k^T \right\} + \frac{N\epsilon}{4} \\ &\leq \sum_{k=1}^N \text{tr} \left\{ C_k L_k C_k^T + D_k D_k^T \right\} + \frac{N\epsilon}{2} \\ &= N \|G\|_2^2 + \frac{N\epsilon}{2} \\ \Rightarrow \frac{1}{N} \sum_{k=1}^N \text{tr} W_k &< \|G\|_2^2 + \epsilon = \gamma. \end{aligned}$$

This completes the proof of (1) \Rightarrow (2).

Proof of (1) \Leftarrow (2):

Using Schur complements, we know that

$$\begin{aligned} P_{k+1} - A_k P_k A_k^T - B_k B_k^T &= \Delta_k \succ 0 \\ \Rightarrow P_{k+1} &= A_k P_k A_k^T + (B_k B_k^T + \Delta_k) \\ \Rightarrow P_{k+1} - L_{k+1} &= A_k (P_k - L_k) A_k^T + \Delta_k \end{aligned}$$

where L_k is as defined in the previous part of the proof. Thus, by (2.4), $P_k - L_k \succ 0$. Also, we know that

using Schur complements gives

$$\begin{aligned}
W_k &\succ C_k P_k C_k^T + D_k D_k^T \\
&\Rightarrow \text{tr } W_k > \text{tr} \{C_k (P_k - L_k) C_k^T + C_k L_k C_k^T + D_k D_k^T\} \\
&> \text{tr} \{C_k L_k C_k^T + D_k D_k^T\} \\
&\Rightarrow \frac{1}{N} \sum_{k=1}^N \text{tr } W_k > \frac{1}{N} \sum_{k=1}^N \text{tr} \{C_k L_k C_k^T + D_k D_k^T\} = \|G\|_2^2.
\end{aligned}$$

Thus

$$\|G\|_2^2 < \frac{1}{N} \sum_{k=1}^N \text{tr } W_k < \gamma.$$

This completes the proof of (1) \Leftrightarrow (2).

Proof of (2) \Leftrightarrow (3):

$$\begin{aligned}
0 &\prec \begin{bmatrix} P_{k+1} & A_k P_k & B_k \\ \bullet & P_k & 0 \\ \bullet & \bullet & I \end{bmatrix} \\
&\Leftrightarrow 0 \prec \begin{bmatrix} P_{k+1}^{-1} & & \\ & P_k^{-1} & \\ & & I \end{bmatrix} \begin{bmatrix} P_{k+1} & A_k P_k & B_k \\ \bullet & P_k & 0 \\ \bullet & \bullet & I \end{bmatrix} \begin{bmatrix} P_{k+1}^{-1} & & \\ & P_k^{-1} & \\ & & I \end{bmatrix} \\
&= \begin{bmatrix} P_{k+1}^{-1} & P_{k+1}^{-1} A_k & P_{k+1}^{-1} B_k \\ \bullet & P_k^{-1} & 0 \\ \bullet & \bullet & I \end{bmatrix}.
\end{aligned}$$

Also,

$$\begin{aligned}
0 \prec & \begin{bmatrix} W_k & C_k P_k & D_k \\ \bullet & P_k & 0 \\ \bullet & \bullet & I \end{bmatrix} \\
\Leftrightarrow 0 \prec & \begin{bmatrix} I & & \\ & P_k^{-1} & \\ & & I \end{bmatrix} \begin{bmatrix} W_k & C_k P_k & D_k \\ \bullet & P_k & 0 \\ \bullet & \bullet & I \end{bmatrix} \begin{bmatrix} I & & \\ & P_k^{-1} & \\ & & I \end{bmatrix} \\
= & \begin{bmatrix} W_k & C_k & D_k \\ \bullet & P_k^{-1} & 0 \\ \bullet & \bullet & I \end{bmatrix}.
\end{aligned}$$

Defining $Q_k = P_k^{-1}$ completes the proof. □

A.4 Proof of Theorem 10

Theorem. *For UES LPTV operators, the following conditions are equivalent:*

1. $\|G\|_\infty^2 < \gamma$
2. $\exists P_1, \dots, P_N$ such that

$$\begin{bmatrix} \gamma I & 0 & C_k P_k & D_k \\ \bullet & P_{k+1} & A_k P_k & B_k \\ \bullet & \bullet & P_k & 0 \\ \bullet & \bullet & \bullet & I \end{bmatrix} \succ 0$$

holds $\forall k \in \{1, \dots, N\}$ where P_{N+1} is taken to be P_1 .

3. $\exists Q_1, \dots, Q_N$ such that

$$\begin{bmatrix} \gamma I & 0 & C_k & D_k \\ \bullet & Q_{k+1} & Q_{k+1} A_k & Q_{k+1} B_k \\ \bullet & \bullet & Q_k & 0 \\ \bullet & \bullet & \bullet & I \end{bmatrix} \succ 0$$

holds $\forall k \in \{1, \dots, N\}$ where Q_{N+1} is taken to be Q_1 .

Proof. (1) \Rightarrow (2):

By Lemma 9, condition (1) is equivalent to

$$\left\| \left(\begin{array}{c|c} \mathcal{Z}^T \mathcal{A} & \mathcal{Z}^T \mathcal{B} \\ \hline \mathcal{C} & \mathcal{D} \end{array} \right) \right\|_{\mathcal{H}_\infty} < \gamma.$$

Since this is an LTI system, we can apply a well-known fact [9] which allows us to equivalently say that $\exists \bar{\mathcal{P}}$ such that

$$\begin{aligned} 0 &\prec \begin{bmatrix} \bar{\mathcal{P}} & \mathcal{Z}^T \mathcal{A} \bar{\mathcal{P}} & \mathcal{Z}^T \mathcal{B} & 0 \\ \bullet & \bar{\mathcal{P}} & 0 & \bar{\mathcal{P}} \mathcal{C}^T \\ \bullet & \bullet & I & \mathcal{D}^T \\ \bullet & \bullet & \bullet & \gamma I \end{bmatrix} \\ \Leftrightarrow 0 &\prec \begin{bmatrix} & & I & \\ \mathcal{Z} & & & \\ & I & & \\ & & I & \end{bmatrix} \begin{bmatrix} \bar{\mathcal{P}} & \mathcal{Z}^T \mathcal{A} \bar{\mathcal{P}} & \mathcal{Z}^T \mathcal{B} & 0 \\ \bullet & \bar{\mathcal{P}} & 0 & \bar{\mathcal{P}} \mathcal{C}^T \\ \bullet & \bullet & I & \mathcal{D}^T \\ \bullet & \bullet & \bullet & \gamma I \end{bmatrix} \begin{bmatrix} \mathcal{Z}^T & & & \\ & I & & \\ & & I & \\ I & & & \end{bmatrix} \\ &= \begin{bmatrix} \gamma I & 0 & \mathcal{C} \bar{\mathcal{P}} & \mathcal{D} \\ \bullet & \mathcal{Z} \bar{\mathcal{P}} \mathcal{Z}^T & \mathcal{A} \bar{\mathcal{P}} & \mathcal{B} \\ \bullet & \bullet & \bar{\mathcal{P}} & 0 \\ \bullet & \bullet & \bullet & I \end{bmatrix} \end{aligned}$$

where we have used that \mathcal{Z} is unitary. We now partition $\bar{\mathcal{P}}$ as

$$\bar{\mathcal{P}} = \begin{bmatrix} P_1 & & * \\ & \ddots & \\ * & & P_N \end{bmatrix}$$

where the asterisk denotes entries which are unimportant. Note that all of our matrices are now partitioned into $n \times n$ blocks. Now we define

$$T_1 := \begin{bmatrix} I_n & \\ 0_{n(N-1) \times n} & \end{bmatrix}.$$

Note that $T_1^T M T_1$ extracts the 1st $n \times n$ block diagonal entry of M . Since T has trivial null space, we can

say that the matrix inequality above implies that

$$\begin{aligned}
0 \succ & \begin{bmatrix} T_1^T & & & \\ & T_1^T & & \\ & & T_1^T & \\ & & & T_1^T \end{bmatrix} \begin{bmatrix} \gamma I & 0 & \mathcal{C}\bar{\mathcal{P}} & \mathcal{D} \\ \bullet & \mathcal{Z}\bar{\mathcal{P}}\mathcal{Z}^T & \mathcal{A}\bar{\mathcal{P}} & \mathcal{B} \\ \bullet & \bullet & \bar{\mathcal{P}} & 0 \\ \bullet & \bullet & \bullet & I \end{bmatrix} \begin{bmatrix} T_1 & & & \\ & T_1 & & \\ & & T_1 & \\ & & & T_1 \end{bmatrix} \\
= & \begin{bmatrix} \gamma I & 0 & C_1 P_1 & D_1 \\ \bullet & P_2 & A_1 P_1 & B_1 \\ \bullet & \bullet & P_1 & 0 \\ \bullet & \bullet & \bullet & I \end{bmatrix}.
\end{aligned}$$

Since we can define T_k so that $T_k^T M T_k$ extracts the k^{th} $n \times n$ block diagonal entry of M , we can use the same methodology to verify that

$$\begin{bmatrix} \gamma I & 0 & C_k P_k & D_k \\ \bullet & P_{k+1} & A_k P_k & B_k \\ \bullet & \bullet & P_k & 0 \\ \bullet & \bullet & \bullet & I \end{bmatrix} \succ 0$$

where P_{N+1} is taken to be P_1 . Therefore, condition (2) holds.

Proof of (1) \Leftarrow (2):

Each inequality in condition (2) can be expressed as

$$\begin{bmatrix} \gamma I - D_k D_k^T - C_k P_k C_k^T & -(A_k P_k C_k^T + B_k D_k^T)^T \\ \bullet & P_{k+1} - A_k P_k A_k^T - B_k B_k^T \end{bmatrix} \succ 0$$

by using Schur complements. Using Schur complements again, each inequality in condition (2) can be expressed as

$$\begin{aligned}
P_{k+1} - A_k P_k A_k^T - B_k B_k^T - (A_k P_k C_k^T + B_k D_k^T) U_k^{-1} (A_k P_k C_k^T + B_k D_k^T)^T & \succ 0 \\
U_k = \gamma I - D_k D_k^T - C_k P_k C_k^T & \succ 0 \\
P_k & \succ 0.
\end{aligned}$$

Expressing these inequalities in block diagonal form gives

$$\begin{aligned} \mathcal{Z}\mathcal{P}\mathcal{Z}^T - \mathcal{A}\mathcal{P}\mathcal{A}^T - \mathcal{B}\mathcal{B}^T - (\mathcal{A}\mathcal{P}\mathcal{C}^T + \mathcal{B}\mathcal{D}^T)\mathcal{U}^{-1}(\mathcal{A}\mathcal{P}\mathcal{C}^T + \mathcal{B}\mathcal{D}^T)^T &> 0 \\ \mathcal{U} = \gamma I - \mathcal{D}\mathcal{D}^T - \mathcal{C}\mathcal{P}\mathcal{C}^T &> 0 \\ \mathcal{P} &> 0 \end{aligned}$$

which can be rewritten using Schur complements as

$$\begin{aligned} 0 &< \begin{bmatrix} \gamma I & 0 & \mathcal{C}\mathcal{P} & \mathcal{D} \\ \bullet & \mathcal{Z}\mathcal{P}\mathcal{Z}^T & \mathcal{A}\mathcal{P} & \mathcal{B} \\ \bullet & \bullet & \mathcal{P} & 0 \\ \bullet & \bullet & \bullet & I \end{bmatrix} \\ \Leftrightarrow 0 &< \begin{bmatrix} I & & & \\ & \mathcal{Z}^T & & \\ & & I & \\ & & & I \end{bmatrix} \begin{bmatrix} \gamma I & 0 & \mathcal{C}\mathcal{P} & \mathcal{D} \\ \bullet & \mathcal{Z}\mathcal{P}\mathcal{Z}^T & \mathcal{A}\mathcal{P} & \mathcal{B} \\ \bullet & \bullet & \mathcal{P} & 0 \\ \bullet & \bullet & \bullet & I \end{bmatrix} \begin{bmatrix} I & & & \\ & \mathcal{Z} & & \\ & & I & \\ & & & I \end{bmatrix} \\ &= \begin{bmatrix} \gamma I & 0 & \mathcal{C}\mathcal{P} & \mathcal{D} \\ \bullet & \mathcal{P} & \mathcal{Z}^T\mathcal{A}\mathcal{P} & \mathcal{Z}^T\mathcal{B} \\ \bullet & \bullet & \mathcal{P} & 0 \\ \bullet & \bullet & \bullet & I \end{bmatrix}. \end{aligned}$$

From the first part of the proof, this is equivalent to condition(1).

Proof of (2) \Leftrightarrow (3):

$$\begin{aligned}
0 &\prec \begin{bmatrix} \gamma I & 0 & C_k P_k & D_k \\ \bullet & P_{k+1} & A_k P_k & B_k \\ \bullet & \bullet & P_k & 0 \\ \bullet & \bullet & \bullet & I \end{bmatrix} \\
\Leftrightarrow 0 &\prec \begin{bmatrix} I & & & \\ & P_{k+1}^{-1} & & \\ & & P_k^{-1} & \\ & & & I \end{bmatrix} \begin{bmatrix} \gamma I & 0 & C_k P_k & D_k \\ \bullet & P_{k+1} & A_k P_k & B_k \\ \bullet & \bullet & P_k & 0 \\ \bullet & \bullet & \bullet & I \end{bmatrix} \begin{bmatrix} I & & & \\ & P_{k+1}^{-1} & & \\ & & P_k^{-1} & \\ & & & I \end{bmatrix} \\
= &\begin{bmatrix} \gamma I & 0 & C_k & D_k \\ \bullet & P_{k+1}^{-1} & P_{k+1}^{-1} A_k & P_{k+1}^{-1} B_k \\ \bullet & \bullet & P_k^{-1} & 0 \\ \bullet & \bullet & \bullet & I \end{bmatrix}.
\end{aligned}$$

Defining $Q_k := P_k^{-1}$ completes the proof. □