A Unified Gradient Approach to Performance Optimization Under Pole Assignment Constraint

Tingshu Hu & Zongli Lin
Department of Electrical Engineering
University of Virginia
Charlottesville, Virginia 22903
U.S.A.
Email: th7f, zl5y@virginia.edu

James Lam
Department of Mechanical Engineering
University of Hong Kong
Pokfulam, Hong Kong
China
Email: jlam@hku.hk

Abstract

General closed-loop performance optimization problems with pole assignment constraint are considered in this paper under a unified framework. By introducing a free parameter matrix and a matrix function based on the solution of a Sylvester equation, the constrained optimization problem is transformed into an unconstrained one, thus reducing the problem of closed-loop performance optimization with pole placement constraint to the computation of the gradient of the performance index with respect to the free parameter matrix. Several classical performance indices are then optimized under the pole placement constraint. The effectiveness of the proposed gradient method is illustrated with an example.

1 Introduction

It is well known that the dynamics of a linear system is closely related to its pole locations. Although the relation is very complicated, some rules have been developed from analysis and experience for the selection of closed-loop pole locations that would result in certain desired dynamics. For a controllable system that has more than one input, there will be infinitely many feedback controllers that assign the closed-loop poles to the same set of locations. A classical design problem is to choose among these feedback laws to optimize other performances. One of the earliest and most widely studied performance associated with pole assignment is the sensitivity of the eigenvalues of the closed-loop system matrix $A+BF$, where A is the open-loop system matrix, B is the input matrix, and F is the feedback gain matrix. This is also referred to as robust pole assignment (see, e.g., [3, 4, 7, 12, 13, 14, 16, 20, 22, 24] and references therein). It is known [19, 23] that this sensitivity is directly related to the condition number of the eigenvector matrix V of $A+BF$. Some numerical methods were proposed in [13, 14, 16, 21] to minimize this condition number. Other indices used in the robust pole assignment include $\mathrm{tr}(I-V^*V)^2$ [4], $\|A+BF\|_F$ [6], and some weighted indices [11, 13].

In this paper, a unified gradient method is proposed to optimize a general performance index under pole assignment constraints. This gradient method was first used in [9, 10]. In particular, it was used in [9] to improve the stability margin. The effectiveness was further illustrated in [10] in the minimization of the condition number of V and the $\mathcal{H}_2$ norm. By introducing a free parameter matrix and a matrix function based on the solution of a Sylvester equation, the constrained optimization problem is transformed into an unconstrained one, thus reducing the problem of closed-loop performance optimization with pole placement constraint to the computation of the gradient of the performance index with respect to the free parameter matrix. We will derive the computational formulae for the gradient of various performance indices with respect to the free parameter matrix. These performance indices include the conventional condition number, the $\mathcal{H}_2$ norm, the $\mathcal{H}_\infty$ norm, and two indices for measuring the overshoot.

It should be emphasized that these pole assignment optimization problems are generally non-convex and hence may possess many local minima. Therefore, as in any gradient-based optimization algorithms, one should expect that different initial starting points have to be used. Numerical tests on a benchmark system in the pole assignment literature have indicated that the gradient approach has the potential for achieving a performance significantly better than that could be achieved by the existing methods.

This paper is organized as follows. In Section 2, a general pole assignment optimization problem is formulated and a unified gradient approach to its solution is presented. This approach is then utilized to optimize the $\mathcal{H}_2$ norm, the $\mathcal{H}_\infty$ norm, and a new index for overshoot in Sections 3, 4, and 5 respectively. Section 6 provides a design example and makes some interesting comparisons among different design approaches. A brief concluding remark is made in Section 7. The $\mathcal{H}_2$ norm as a performance index of the closed-loop system was studied in [10]. We have briefly recalled the main results of [10] in Section 2 for completeness and for the use in the comparison study of the example in Section 6.

2 The Unified Gradient Approach

2.1 The General Problem

Given $A \in \mathbf{R}^{n\times n}$, $B \in \mathbf{R}^{n\times m}$. Assume throughout the paper that (A, B) is controllable, $m > 1$ and B has full column rank m. Let $\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$ be a set of self-conjugate complex numbers corresponding to the set of desired poles. Assume that there are n' complex conjugate pairs, $\lambda_{2i-1}, \lambda_{2i} = \alpha_i \pm j\beta_i,\ i = 1, 2, \ldots, n'$, then one can define the following real block diagonal matrix:

$$\Lambda := \text{blkdiag}\left\{\begin{bmatrix} \alpha_1 & \beta_1 \\ -\beta_1 & \alpha_1 \end{bmatrix}, \ldots, \begin{bmatrix} \alpha_{n'} & \beta_{n'} \\ -\beta_{n'} & \alpha_{n'} \end{bmatrix}, \lambda_{2n'+1}, \ldots, \lambda_n\right\}. \tag{1}$$

We also assume that the eigenvalues of Λ are distinct. The problem of pole assignment by state feedback is to choose a matrix F such that

$$V^{-1}(A + BF)V = \Lambda \tag{2}$$

for some nonsingular V. A state feedback matrix F is said to be *admissible* if the pole assignment constraint (2) is satisfied. Since $m > 1$, there are infinitely many admissible F and the corresponding V. The freedom in the choice of F and V is then exploited to achieve other design objectives, such as robust stability, $\mathcal{H}_\infty$ and $\mathcal{H}_2$ sensitivity reduction. Let $J(F, V)$ be a general performance index, as a function of F and V, the problem to be studied in this paper is then

$$\inf_{F,V} \quad J(F, V) \tag{3}$$

$$\text{s.t.} \quad V^{-1}(A + BF)V = \Lambda. \tag{4}$$

2.2 Parameterization of the Pole Placement Constraint

In this section, we follow the idea of [1, 4] to parameterize all the feedback matrices F and the eigenvector matrices V that satisfy (4) as a function of a free parameter $U \in \mathbf{R}^{m\times n}$. This is achieved by solving for V a Sylvester equation parameterized in U and then recovering the feedback matrix via $F = UV^{-1}$. In this way, the general performance index becomes a function of the free parameter U. Explicit formulae for computing the gradient of the performance index with respect to the free parameter U can be derived.

Given a controllable pair (A, B) and a real block diagonal matrix Λ in the form of (1) such that A and Λ have no common eigenvalues, then a function $f : U \to (F, V)$ is defined as follows. For $U \in \mathbf{R}^{m\times n}$, solve

$$AV - V\Lambda = -BU \tag{5}$$

for V and if V is nonsingular, let $F = UV^{-1}$. The function is denoted as $(F, V) = f(U)$. The domain of f is

$$\mathcal{D}_f := \{U \in \mathbf{R}^{m\times n} \mid V \text{ in (5) is nonsingular}\}$$

and the range of f is $\mathcal{R}_f = f(\mathcal{D}_f)$.

Remark 2.1 *The assumption on the eigenvalues of A and Λ, i.e., the open-loop and the closed-loop poles are entirely different, ensures that (5) always has a unique solution for each U. This assumption is not a severe restriction in general. If certain open-loop poles are required to be retained in the closed-loop system, a partial pole assignment approach may be taken [15, 18]. It is also easy to see that V depends linearly on U and F is in fact a rational function of the elements of U.*

The following result justifies the use of the parameter U as a means to optimize the performance index under pole assignment constraints.

Theorem 2.1 [1, 8]

(a) $\mathcal{D}_f$ is a dense open set in $\mathbf{R}^{m\times n}$;

(b) $\{(F,V) : V^{-1}(A+BF)V = \Lambda\} = \mathcal{R}_f = f(\mathcal{D}_f)$.

Since the performance index $J(F,V)$ is uniquely determined by F and V, it is a function of the free parameter U. Consequently, it can be expressed as $\hat{J}(U)$. As $(F,V) = f(U)$ is a rational function and $\mathcal{D}_f$ is an open set, so F and V are differentiable with respect to U for all $U \in \mathcal{D}_f$. Thus $\frac{\partial \hat{J}}{\partial U}$ exists if J has partial derivative with respect to F and V.

Remark 2.2 *The method of parameterizing the pole assignment constraint as described above can also be found in [13, 14]. The ideas employed are basically the same, but with different descriptions. The description in this paper facilitates the derivation of the gradients of the performance index with respect to U.*

2.3 Computation of the Gradient

Using the parameterization method in Section 2.2, we have

$$(F,V) = f(U) : AV - V\Lambda = -BU, \quad F = UV^{-1}, \quad U \in \mathcal{D}_f. \tag{6}$$

Denote $\hat{J}(U) = J[f(U)]$, then the optimization problem (3) can be simply formulated as

$$\inf_{U\in\mathcal{D}_f} \hat{J}(U). \tag{7}$$

In what follows, we derive a formula for expressing the gradient $\frac{\partial \hat{J}}{\partial U}$ in terms of $\frac{\partial J}{\partial F}$ and $\frac{\partial J}{\partial V}$. To this end, we need to first establish the following lemma, which will also be used in several other occasions throughout the paper.

Lemma 2.1 *If M, N, Q, R, X, $Y \in \mathbf{R}^{n\times n}$ satisfy*

$$MX + XN = Q, \tag{8}$$

$$YM + NY = R, \tag{9}$$

then $\mathrm{tr}(RX) = \mathrm{tr}(QY)$.

Proof. Multiplying from the right of (8) and (9) by Y and X respectively, we have,

$$MXY + XNY = QY, \tag{10}$$

$$YMX + NYX = RX. \tag{11}$$

The result then follows from the facts that $\operatorname{tr}(MXY) = \operatorname{tr}(YMX)$ and $\operatorname{tr}(XNY) = \operatorname{tr}(NYX)$. □

The following theorem gives a unified gradient formula. Some separate gradient formulae can be found in [10].

Theorem 2.2 *For $U \in \mathcal{D}_f$, let $(F, V) = f(U)$. Then the gradient of $\hat{J}(U)$ is given in terms of $\frac{\partial J}{\partial F}$ and $\frac{\partial J}{\partial V}$ as*

$$\frac{\partial \hat{J}}{\partial U} = \left(\frac{\partial J}{\partial F}\right)(V^{-1})^{\mathrm{T}} + B^{\mathrm{T}} Y^{\mathrm{T}}, \tag{12}$$

where Y is the unique solution to the Sylvester equation,

$$YA - \Lambda Y = V^{-1}\left(\frac{\partial J}{\partial F}\right)^{\mathrm{T}} F - \left(\frac{\partial J}{\partial V}\right)^{\mathrm{T}}. \tag{13}$$

Proof. Denote the ith basis vector of $\mathbf{R}^m$ as e_i, the jth basis vector of $\mathbf{R}^n$ as ε_j and the (i,j) element of U as u_{ij}, then

$$\frac{\partial U}{\partial u_{ij}} = e_i \varepsilon_j^{\mathrm{T}}$$

and

$$\frac{\partial F}{\partial u_{ij}} = \frac{\partial U V^{-1}}{\partial u_{ij}} = \frac{\partial U}{\partial u_{ij}} V^{-1} - U V^{-1} \frac{\partial V}{\partial u_{ij}} V^{-1} = e_i \varepsilon_j^{\mathrm{T}} V^{-1} - F \frac{\partial V}{\partial u_{ij}} V^{-1},$$

where $\frac{\partial V}{\partial u_{ij}}$, by (6), satisfies

$$A \frac{\partial V}{\partial u_{ij}} - \frac{\partial V}{\partial u_{ij}} \Lambda = -B e_i \varepsilon_j^{\mathrm{T}}. \tag{14}$$

Denote the (p,q) element of F as f_{pq} and that of V as v_{pq}, then,

$$\begin{aligned}
\frac{\partial \hat{J}}{\partial u_{ij}} &= \sum_{p=1}^{m}\sum_{q=1}^{n} \frac{\partial J}{\partial f_{pq}} \frac{\partial f_{pq}}{\partial u_{ij}} + \sum_{p=1}^{n}\sum_{q=1}^{n} \frac{\partial J}{\partial v_{pq}} \frac{\partial v_{pq}}{\partial u_{ij}} \\
&= \operatorname{tr}\left[\left(\frac{\partial J}{\partial F}\right)^{\mathrm{T}} \frac{\partial F}{\partial u_{ij}} + \left(\frac{\partial J}{\partial V}\right)^{\mathrm{T}} \frac{\partial V}{\partial u_{ij}}\right] \\
&= \operatorname{tr}\left[\left(\frac{\partial J}{\partial F}\right)^{\mathrm{T}} \left(e_i \varepsilon_j^{\mathrm{T}} V^{-1} - F \frac{\partial V}{\partial u_{ij}} V^{-1}\right) + \left(\frac{\partial J}{\partial V}\right)^{\mathrm{T}} \frac{\partial V}{\partial u_{ij}}\right] \\
&= \varepsilon_j^{\mathrm{T}} V^{-1} \left(\frac{\partial J}{\partial F}\right)^{\mathrm{T}} e_i - \operatorname{tr}\left\{\left[V^{-1}\left(\frac{\partial J}{\partial F}\right)^{\mathrm{T}} F - \left(\frac{\partial J}{\partial V}\right)^{\mathrm{T}}\right] \frac{\partial V}{\partial u_{ij}}\right\}.
\end{aligned}$$

By Lemma 2.1, we have

$$-\operatorname{tr}\left\{\left[V^{-1}\left(\frac{\partial J}{\partial F}\right)^{\mathrm{T}} F - \left(\frac{\partial J}{\partial V}\right)^{\mathrm{T}}\right] \frac{\partial V}{\partial u_{ij}}\right\} = \operatorname{tr}\left[\left(B e_i \varepsilon_j^{\mathrm{T}}\right) Y\right] = \varepsilon_j^{\mathrm{T}} Y B e_i,$$

where Y is the unique solution of (13). Consequently,

$$\frac{\partial \hat{J}}{\partial u_{ij}} = \varepsilon_j^{\mathrm{T}} \left[V^{-1} \left(\frac{\partial J}{\partial F} \right)^{\mathrm{T}} + YB \right] e_i = e_i^{\mathrm{T}} \left[V^{-1} \left(\frac{\partial J}{\partial F} \right)^{\mathrm{T}} + YB \right]^{\mathrm{T}} \varepsilon_j$$

and the result follows. □

From Theorem 2.2, we see that, given $\frac{\partial J}{\partial F}$ and $\frac{\partial J}{\partial V}$, the gradient $\frac{\partial \hat{J}}{\partial U}$ is very easy to compute. With formula (12), one can optimize the performance index under the pole assignment constraint by simply applying a gradient algorithm (for instance, 'fminu' in the Matlab Optimization Toolbox).

For a given U, the computation of $\hat{J}(U)$ and $\frac{\partial \hat{J}}{\partial U}$ is summarized as follows:

$\hat{J}(U)$: Solve $AV - V\Lambda = -BU$ for V, let $F = UV^{-1}$, then $\hat{J}(U) = J(F,V)$.

$\frac{\partial \hat{J}}{\partial U}$: Compute F and V as above. Compute $\frac{\partial J}{\partial F}$ and $\frac{\partial J}{\partial V}$. Solve (13) for Y, then compute $\frac{\partial \hat{J}}{\partial U}$ from (12).

The above unified procedure applies to any performance index $J(F,V)$. For a specific performance index, the task of optimization reduces to the computation of $\frac{\partial J}{\partial F}$ and $\frac{\partial J}{\partial V}$. This will be dealt with in the following sections for several important $J(F,V)$'s.

3 The $\mathcal{H}_2$ Problem

Consider the following system with disturbance

$$\begin{aligned} \dot{x} &= Ax + B_1 w + B_2 u, \\ z &= Cx + D_{12} u, \\ y &= x, \end{aligned} \tag{15}$$

where $w \in \mathbf{R}^r$ is the exogenous disturbance, $u \in \mathbf{R}^m$ the control input, $y \in \mathbf{R}^p$ the measured output and $z \in \mathbf{R}^q$ the controlled output. Under state feedback $u = Fx$, the transfer function from w to z is

$$T(s,F) = (C + D_{12}F)(sI - A - B_2 F)^{-1} B_1.$$

Suppose that $A + B_2 F$ is stable, then the $\mathcal{H}_2$ norm of $T(s,F)$ is given by

$$\|T(s,F)\|_2^2 = \mathrm{tr}(B_1^{\mathrm{T}} P B_1) =: J_2(F),$$

where $P = P^{\mathrm{T}} \geq 0$ satisfies

$$(A + B_2 F)^{\mathrm{T}} P + P(A + B_2 F) = -(C + D_{12}F)^{\mathrm{T}} (C + D_{12}F). \tag{16}$$

In the case where the poles of $A + B_2 F$ are not fixed *a priori*, F may become unbounded in order to make the $\mathcal{H}_2$ norm small. Therefore, to regularize the problem, we consider the following $\mathcal{H}_2$ optimization under the constraint of pole assignment,

$$\inf_F J_2(F) \tag{17}$$

$$\mathrm{s.t.} V^{-1}(A + B_2 F)V = \Lambda. \tag{18}$$

In [5], the same $\mathcal{H}_2$ norm optimization under regional pole assignment constraint was considered. The original problem is not convex and is transformed into the solution of certain LMIs with some conservatism. The LMIs are easy to solve but only a suboptimal solution can be obtained. We will see in an example that the resulting suboptimal H_2 norm can be significantly greater than the optimal one. Clearly, (17) falls into our unified framework of (3)-(4). With Theorem 2.2, the remaining task is to compute $\frac{\partial J_2}{\partial F}$.

Proposition 3.1

$$\frac{\partial J_2}{\partial F} = 2\left[B_2^{\mathrm{T}} P + D_{12}^{\mathrm{T}}(C + D_{12}F)\right] X, \tag{19}$$

where X is the unique solution to the Lyapunov equation,

$$(A + B_2F)X + X(A + B_2F)^{\mathrm{T}} = -B_1B_1^{\mathrm{T}}. \tag{20}$$

Proof. See [10]. □

For the $\mathcal{H}_2$ performance index J_2, the gradient $\frac{\partial J_2}{\partial U}$ exists for all $U \in \mathcal{D}_f$. Numerical results show that the gradient algorithm always terminates at a stationary point with $\frac{\partial \hat{J}_2}{\partial U} = 0$. It also reveals that the optimization problem may possess different local minima.

4 The $\mathcal{H}_\infty$ Problem

In this section, the $\mathcal{H}_\infty$ norm of a transfer matrix is used as an index of robustness or an index of sensitivity. The robustness problem is first considered below.

Robust pole assignment

The uncertain system is

$$\dot{x} = (A + G\Delta H)x + Bu \tag{21}$$

where $A \in \mathbf{R}^{n\times n}$, $B \in \mathbf{R}^{n\times m}$, $G \in \mathbf{R}^{n\times l}$ and $H \in \mathbf{R}^{l\times n}$ are given constant matrices. Δ is the uncertainty, possibly time-varying, with some specified structure,

$$\Delta = \mathrm{blkdiag}\{\Delta_1, \cdots, \Delta_N\}.$$

Denote

$$\mathbf{D} := \{D = \mathrm{blkdiag}[d_1I, \cdots, d_NI] : d_i > 0\},$$

where for each $i = 1$ to N, d_iI has the dimension of Δ_i. Under the state feedback $u = Fx$, the closed-loop system is given by,

$$\dot{x} = (A + BF + G\Delta H)x. \tag{22}$$

It is well known from [25] that system (22) is stable for all Δ with the specified structure and $\overline{\sigma}(\Delta) \le \gamma$ if

$$\inf_{D\in\mathbf{D}} \|D^{-1}H(sI - A - BF)^{-1}GD\|_\infty < \frac{1}{\gamma},$$

where and elsewhere $\overline{\sigma}(\cdot)$ and $\underline{\sigma}(\cdot)$ denote respectively the maximal and minimal singular values of a matrix.

Let

$$J_{\infty 1}(F, D) := \|D^{-1}H(sI - A - BF)^{-1}GD\|_\infty,$$

then a robust pole assignment problem can be formulated as follows,

$$\begin{aligned} \inf \quad & J_{\infty 1}(F, D) \\ \text{s.t.} \quad & V^{-1}(A + BF)V = \Lambda, \; D \in \mathbf{D}. \end{aligned} \tag{23}$$

Sensitivity reduction with pole assignment

The system with disturbance of finite energy is described as

$$\begin{aligned} \dot{x} &= Ax + B_1 w + B_2 u \\ z &= Cx + D_{11} w + D_{12} u \\ y &= x \end{aligned} \tag{24}$$

where w is the disturbance. Under state feedback $u = Fx$, the transfer function from w to z is

$$T(s, F) = (C + D_{12}F)(sI - A - B_2F)^{-1}B_1 + D_{11}.$$

Denote $J_{\infty 2}(F) := \|T(s, F)\|_\infty$. The $\mathcal{H}_\infty$ sensitivity reduction with pole assignment problem is:

$$\begin{aligned} \inf \quad & J_{\infty 2}(F) \\ \text{s.t.} \quad & V^{-1}(A + B_2F)V = \Lambda. \end{aligned} \tag{25}$$

Both optimization problems (23) and (25) fall into the framework of (3)-(4), except that at some points F, the functions $J_{\infty 1}(F, D)$ or $J_{\infty 2}(F)$ may be nondifferentiable. Although the gradient method is hard to apply at nondifferentiable points, computation experience have shown that the gradient method can improve the performance significantly. With Theorem 2.2, again the remaining tasks are to compute $\frac{\partial J_{\infty 1}}{\partial F}$, $\frac{\partial J_{\infty 2}}{\partial F}$ and $\frac{\partial J_{\infty 1}}{\partial D}$.

In order to proceed on developing the derivative formulae, we consider a general transfer function depending on an arbitrary parameter P:

$$T(s, P) = C(P)[sI - A(P)]^{-1}B(P) + D(P),$$

where P is a real valued matrix containing all the design parameters. Assume that $A(P)$, $B(P)$, $C(P)$ and $D(P)$ are continuously differentiable with respect to P. The following intuitive result is first established.

Proposition 4.1 *For a given P_0, if*

(a) $\|T(s, P_0)\|_\infty = \sup_{\omega \in \mathbf{R}^+} \bar{\sigma}[T(j\omega, P_0)]$ *is attained at* $\omega_0 < \infty$ *and*

$$\bar{\sigma}[T(j\omega_0, P_0)] > \bar{\sigma}[T(j\omega, P_0)], \quad \textit{for all } \omega \neq \omega_0;$$

(b) $\sigma_0 = \bar{\sigma}[T(j\omega_0, P_0)]$ *is a distinct singular value of* $T(j\omega_0, P_0)$.

Then,

$$\left.\frac{\partial \|T(s,P)\|_\infty}{\partial P}\right|_{P=P_0} = \left.\frac{\partial \bar{\sigma}[T(j\omega,P)]}{\partial P}\right|_{\substack{P=P_0\\ \omega=\omega_0}}.$$

Proof. If condition (a) is satisfied, then there exist $\delta_1, \omega_1 > 0$ such that

$$\|T(s,P)\|_\infty = \max_{0\le\omega\le\omega_1} \bar{\sigma}[T(j\omega,P)], \quad \forall P \in \mathcal{N}_P(P_0,\delta_1).$$

Here and in the sequal, we use $\mathcal{N}_P(P_0,\delta)$ to denote the δ neighborhood of P_0 and $\mathcal{N}_\omega(\omega_0,\eta)$ to denote the η neighborhood of ω_0. If (b) is satisfied, then $\left.\frac{\partial \bar{\sigma}[T(j\omega,P)]}{\partial \omega}\right|_{\substack{P=P_0\\ \omega=\omega_0}} = 0$ and $\bar{\sigma}[T(j\omega,P)]$ is analytic in some neighborhoods $\mathcal{N}_\omega(\omega_0,\eta)$, $\mathcal{N}_P(P_0,\delta)$. δ can be made sufficiently small such that for any $P \in \mathcal{N}_P(P_0,\delta)$, there exists a unique $\omega_m \in \mathcal{N}_\omega(w_0,\eta)$ such that

$$\left.\frac{\partial \bar{\sigma}[T(j\omega,P)]}{\partial \omega}\right|_{\omega=\omega_m} = 0$$

and this ω_m also attains the global maximum of $\bar{\sigma}[T(j\omega,P)]$. Hence, a function $\omega_m(P)$, $P \in \mathcal{N}_P(P_0,\delta)$ can be determined such that

$$\left.\frac{\partial \bar{\sigma}[T(j\omega,P)]}{\partial \omega}\right|_{\omega=\omega_m(P)} = 0, \quad \omega_m(P_0) = \omega_0, \;\; \omega_m(P) \in \mathcal{N}_\omega(\omega_0,\eta)$$

and $\omega_m(P)$ is the unique frequency at which $\bar{\sigma}[T(j\omega,P)]$ attains the maximum. In other words,

$$\|T(s,P)\|_\infty = \bar{\sigma}[T(j\omega_m(P),P)], \quad \forall P \in \mathcal{N}_P(P_0,\delta).$$

Therefore,

$$\begin{aligned}
\left.\frac{\partial \|T(s,P)\|_\infty}{\partial P}\right|_{P=P_0} &= \left.\frac{\partial \bar{\sigma}[T(j\omega,P)]}{\partial \omega}\right|_{\substack{P=P_0\\ \omega=\omega_m(P_0)}} \left.\frac{\partial \omega_m(P)}{\partial P}\right|_{P=P_0} + \left.\frac{\partial \bar{\sigma}[T(j\omega,P)]}{\partial P}\right|_{\substack{P=P_0\\ \omega=\omega_m(P_0)}} \\
&= \left.\frac{\partial \bar{\sigma}[T(j\omega,P)]}{\partial \omega}\right|_{\substack{P=P_0\\ \omega=\omega_0}} \left.\frac{\partial \omega_m(P)}{\partial P}\right|_{P=P_0} + \left.\frac{\partial \bar{\sigma}[T(j\omega,P)]}{\partial P}\right|_{\substack{P=P_0\\ \omega=\omega_0}} \\
&= \left.\frac{\partial \bar{\sigma}[T(j\omega,P)]}{\partial P}\right|_{\substack{P=P_0\\ \omega=\omega_0}},
\end{aligned}$$

upon noticing that $\left.\frac{\partial \bar{\sigma}[T(j\omega,P)]}{\partial \omega}\right|_{\substack{P=P_0\\ \omega=\omega_0}} = 0.$ □

In the following, we give the derivatives $\frac{\partial J_{\infty 1}}{\partial F}$ and $\frac{\partial J_{\infty 2}}{\partial F}$ associated with optimization problems (23) and (25). First, notice that $\left.\frac{\partial \bar{\sigma}[T(j\omega,P)]}{\partial P}\right|_{\substack{P=P_0\\ \omega=\omega_0}}$ can be computed in terms of the singular vectors of $T(j\omega_0, P_0)$. Let u and v be the singular vectors of $T(j\omega_0, P_0)$ corresponding to σ_0, such that $T(j\omega_0, P_0)v = \sigma_0 u$, then it can be shown that

$$\left.\frac{\partial\bar{\sigma}[T(j\omega, P)]}{\partial p_{ij}}\right|_{\substack{P=P_0\\ \omega=\omega_0}} = \mathrm{Re}\left[u^* \left.\frac{\partial T(j\omega, P)}{\partial p_{ij}}\right|_{\substack{P=P_0\\ \omega=\omega_0}} v\right]. \tag{26}$$

Now, we apply Proposition 4.1 and (26) to $J_{\infty 1}(F, D)$. Given F, D, with ω_0, u, v defined as above, we have

$$\begin{aligned}\frac{\partial J_{\infty 1}(F, D)}{\partial f_{ij}} &= \mathrm{Re}\,[u^* D^{-1} H (j\omega_0 I - A - BF)^{-1} B e_i \varepsilon_j^{\mathrm{T}} (j\omega_0 I - A - BF)^{-1} G D v] \\ &= \mathrm{Re}\,[\varepsilon_j^{\mathrm{T}} (j\omega_0 I - A - BF)^{-1} G D v u^* D^{-1} H (j\omega_0 I - A - BF)^{-1} B e_i] \\ &= e_i^{\mathrm{T}}\, \mathrm{Re}\,[(j\omega_0 I - A - BF)^{-1} G D v u^* D^{-1} H (j\omega_0 I - A - BF)^{-1} B]^{\mathrm{T}} \varepsilon_j.\end{aligned}$$

Hence,

$$\frac{\partial J_{\infty 1}(F, D)}{\partial F} = \mathrm{Re}\,[(j\omega_0 I - A - BF)^{-1} G D v u^* D^{-1} H (j\omega_0 I - A - BF)^{-1} B]^{\mathrm{T}}. \tag{27}$$

Let $E_i = \mathrm{blkdiag}\{0, 0, \ldots, I, 0, \ldots, 0\}$ have the same block diagonal structure with $D \in \mathbf{D}$, with all the blocks zero except for the ith diagonal block. Then

$$\begin{aligned}\frac{\partial J_{\infty 1}(F, D)}{\partial d_i} &= \mathrm{Re}\,\{u^*[-D^{-1} E_i D^{-1} H (j\omega_0 I - A - BF)^{-1} G D + D^{-1} H (j\omega_0 I - A - BF)^{-1} G E_i] v\} \\ &= \mathrm{tr}\,(Q E_i), \end{aligned} \tag{28}$$

where

$$Q = \mathrm{Re}\,[-D^{-1} H (j\omega_0 I - A - BF)^{-1} G D v u^* D^{-1} + v u^* D^{-1} H (j\omega_0 I - A - BF)^{-1} G].$$

Following a similar procedure, we have

$$\frac{\partial J_{\infty 2}(F)}{\partial F} = \mathrm{Re}\,\{(j\omega_0 I - A - B_2 F)^{-1} B_1 v u^* [D_{12} + (C + D_{12} F)(j\omega_0 I - A - B_2 F)^{-1} B_2]\}^{\mathrm{T}}. \tag{29}$$

With $\frac{\partial J_{\infty 1}(F,D)}{\partial F}$ and $\frac{\partial J_{\infty 2}(F)}{\partial F}$, we can easily obtain the derivatives $\frac{\partial \hat{J}_{\infty 1}}{\partial U}$ and $\frac{\partial \hat{J}_{\infty 2}}{\partial U}$ by applying Theorem 2.2.

5 Transient Behavior

Consider the closed-loop system

$$\dot{x} = (A + BF)x. \tag{30}$$

In general, even with the desired closed-loop eigenvalues, an acceptable transient behavior still cannot be guaranteed. We can improve the transient behavior by exploiting the extra freedom of F. Let's first introduce some indices for transient behavior. Define the maximum overshoot as

$$J_o(F) := \sup\left\{\left\|e^{(A+BF)t} x(0)\right\| : \ \|x(0)\| = 1, t \geq 0\right\}.$$

where $x(0)$ is the initial state. In the state regulation problem, one would like to minimize $J_o(F)$. Clearly, we have

$$J_o(F) = \sup_{t \geq 0} \bar{\sigma}\left(e^{(A+BF)t}\right).$$

Unfortunately, the index involves taking the supremum over the positive real line which is computational demanding. Thus, a simpler index is desirable. With the assumption that $A + BF$ has distinct eigenvalues, let V be an eigenvector matrix such that $V^{-1}(A + BF)V = \Lambda$, where Λ takes the form in (1). It is easy to see that

$$J_o(F) \leq \frac{\bar{\sigma}(V)}{\underline{\sigma}(V)}. \tag{31}$$

Therefore, an upper bound on the overshoot is given by the condition number of V. Since the eigenvector matrix V of $A + BF$ is not unique, we consider

$$J_v(F) := \inf_V \frac{\bar{\sigma}(V)}{\underline{\sigma}(V)}, \quad \text{s.t.} \quad V^{-1}(A + BF)V = \Lambda.$$

Clearly,

$$J_o(F) \leq J_v(F). \tag{32}$$

The upper bound given by (32) can still be very conservative as will be revealed in the numerical example. In what follows, we consider another bound that has been used in the literature (see, e.g., p.66 of [2]), and we will show that it is tighter than (32).

For $P > 0$, if

$$(A + BF)^{\mathrm{T}}P + P(A + BF) \leq 0, \tag{33}$$

then $x^{\mathrm{T}}(t)Px(t) \leq x^{\mathrm{T}}(0)Px(0)$. It follows that,

$$\|x(t)\| \leq \left(\frac{\overline{\sigma}[P]}{\underline{\sigma}[P]}\right)^{\frac{1}{2}} \|x(0)\|. \tag{34}$$

By definition, we have $J_o(F) \leq \left(\frac{\overline{\sigma}[P]}{\underline{\sigma}[P]}\right)^{\frac{1}{2}}$. Hence, the condition number of P also provides an indication on the overshoot. As P satisfying (33) is not unique, the upper bound in (34) can be further tighten by letting

$$J_p(F) := \inf \left(\frac{\overline{\sigma}[P]}{\underline{\sigma}[P]}\right)^{\frac{1}{2}}, \quad \text{s.t.} \quad P \text{ satisfies (33)}.$$

Then, we have

$$J_o(F) \leq J_p(F). \tag{35}$$

Remark 5.1 *Given F, $J_p(F)$ can be easily obtained by solving the LMIs,*

$$\inf \gamma, \quad (A + BF)^{\mathrm{T}}P + P(A + BF) \leq 0, \quad I \leq P \leq \gamma I.$$

Then the minimal γ equals to $J_p^2(F)$ (see p.65 of [2]).

Remark 5.2 *Notice that*

$$\Omega(P) := \{x \in \mathbf{R}^n : x^{\mathrm{T}} P x \leq 1\}$$

is an invariant set. If $x(0) \in \Omega(P)$, then $x(t) \in \Omega(P)$ for all t. When $P = I$, $\Omega(P)$ is a unit ball and there will be no overshoot. If the condition number of P is small, the invariant set is close to the unit ball and the overshoot will be small.

Proposition 5.1 $J_p(F) \leq J_v(F)$.

Proof. Let V be an eigenvector matrix of $A + BF$, that is,

$$V^{-1}(A + BF)V = \Lambda,$$

then

$$V^{\mathrm{T}}(A + BF)^{\mathrm{T}}(V^{-1})^{\mathrm{T}} + V^{-1}(A + BF)V = 2\,\mathrm{Re}\,\Lambda < 0.$$

Multiply the above from the left with $(V^{-1})^T$ and from the right with V^{-1}, we obtain

$$(A + BF)^{\mathrm{T}}(V^{-1})^{\mathrm{T}}V^{-1} + (V^{-1})^{\mathrm{T}}V^{-1}(A + BF) < 0.$$

Let $P = (V^{-1})^{\mathrm{T}}V^{-1}$, then

$$\frac{\overline{\sigma}[P]}{\underline{\sigma}[P]} = \left(\frac{\overline{\sigma}[V^{-1}]}{\underline{\sigma}[V^{-1}]}\right)^2 = \left(\frac{\overline{\sigma}[V]}{\underline{\sigma}[V]}\right)^2.$$

This shows $J_p(F) \leq J_v(F)$. □

Numerical examples have indicated that $J_p(F)$ can be much smaller than $J_v(F)$ and hence a better index to assess the overshoot of the closed-loop system. In the pole assignment context, we consider minimizing $J_p(F)$ under the constraint of pole assignment. The problem is

$$\inf J_p^2(F), \quad \text{s.t. } V^{-1}(A + BF)V = \Lambda. \tag{36}$$

Since $J_p(F)$ itself involves taking the infimum, for simplicity, define

$$J_{p1}(F, L) := \frac{\overline{\sigma}[P]}{\underline{\sigma}[P]},$$

where P solves

$$(A + BF)^{\mathrm{T}}P + P(A + BF) = -L^{\mathrm{T}}L,$$

then (36) is equivalent to

$$\inf \quad J_{p1}(F, L), \tag{37}$$

$$\text{s.t.} \quad V^{-1}(A + BF)V = \Lambda. \tag{38}$$

In the following, we will deal with (37) instead of (36) and, as before, by Theorem 2.2, the task of computing $\frac{\hat{J}_{p1}}{\partial U}$ reduces to the computation of $\frac{\partial J_{p1}}{\partial F}$. It should be noted that L is a variable independent of F and U. Since $P > 0$, its singular values are the same as its eigenvalues.

Proposition 5.2 *Suppose that the singular values $\bar{\sigma}(P)$ and $\underline{\sigma}(P)$ of P are distinct. Let v_1, v_n be the eigenvectors of P, normalized to unit length, corresponding to $\bar{\sigma}(P)$ and $\underline{\sigma}(P)$. Then*

$$\frac{\partial J_{p1}}{\partial F} = 2B^{\mathrm{T}}PY, \quad \frac{\partial J_{p1}}{\partial L} = 2LY,$$

where Y is the unique solution to the Lyapunov equation,

$$(A+BF)Y + Y(A+BF)^{\mathrm{T}} = -\frac{1}{\underline{\sigma}^2(P)}[\underline{\sigma}(P)v_1v_1^{\mathrm{T}} - \bar{\sigma}(P)v_nv_n^{\mathrm{T}}]. \tag{39}$$

Proof. It can be easily established that

$$\frac{\partial J_{p1}}{\partial f_{ij}} = \frac{1}{\underline{\sigma}^2(P)}\left[v_1^{\mathrm{T}}\frac{\partial P}{\partial f_{ij}}v_1\underline{\sigma}(P) - v_n^{\mathrm{T}}\frac{\partial P}{\partial f_{ij}}v_n\bar{\sigma}(P)\right] = \mathrm{tr}\left\{\frac{1}{\underline{\sigma}^2(P)}\left[v_1v_1^{\mathrm{T}}\underline{\sigma}(P) - v_nv_n^{\mathrm{T}}\bar{\sigma}(P)\right]\frac{\partial P}{\partial f_{ij}}\right\},$$

where $\frac{\partial P}{\partial f_{ij}}$ is the solution to

$$(A+BF)^{\mathrm{T}}\frac{\partial P}{\partial f_{ij}} + \frac{\partial P}{\partial f_{ij}}(A+BF) = -\left[\left(B\frac{\partial F}{\partial f_{ij}}\right)^{\mathrm{T}}P + PB\frac{\partial F}{\partial f_{ij}}\right] = -\left[\left(Be_i\varepsilon_j^{\mathrm{T}}\right)^{\mathrm{T}}P + PBe_i\varepsilon_j^{\mathrm{T}}\right].$$

Since Y satisfies (39), by Lemma 2.1, we have

$$\frac{\partial J_{p1}}{\partial f_{ij}} = \mathrm{tr}\left[\left(Be_i\varepsilon_j^{\mathrm{T}}\right)^{\mathrm{T}}P + PBe_i\varepsilon_j^{\mathrm{T}}\right]Y = 2e_i^{\mathrm{T}}B^{\mathrm{T}}PY\varepsilon_j.$$

It then follows that

$$\frac{\partial J_{p1}}{\partial F} = 2B^{\mathrm{T}}PY.$$

Similarly, we obtain $\frac{\partial J_{p1}}{\partial L} = 2LY$. □

6 A Numerical Example

Example. A dynamical model is considered for sensitivity reduction as described in (24) with

$$A = \begin{bmatrix} -0.1094 & 0.0628 & 0 & 0 & 0 \\ 1.3060 & -2.1320 & 0.9807 & 0 & 0 \\ 0 & 1.5950 & -3.1490 & 1.5470 & 0 \\ 0 & 0.0355 & 2.6320 & -4.2570 & 1.8550 \\ 0 & 0.00227 & 0 & 0.1636 & -0.1625 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 0 & 0 \\ 0.0638 & 0 \\ 0.0838 & -0.1396 \\ 0.1004 & -0.2060 \\ 0.0063 & -0.0128 \end{bmatrix}, \tag{40}$$

$$B_1 = I, \quad C = I, \quad D_{11} = 0, \quad D_{12} = 0.$$

Matrices A and B_2 are taken from a nominal distillation model. In the pole assignment literature, this system has been used as a benchmarch example for demonstrating and comparing various design approaches (see, e.g., [3, 14, 16, 17, 24]).

Assume that the disturbances enter each channel independently, then we have $B_1 = I$. Equal weights are considered for all the states, hence $C = I$. With state feedback $u = Fx$, the transfer matrix from w to z is

$$T(s,F) = (sI - A - B_2F)^{-1}.$$

The open-loop poles are at -0.077324, -0.014232, -0.89531, -2.8408 and -5.9822. The desired closed-loop poles are $-1 \pm j, -0.2$, -0.5 and -1. One may then choose

$$\Lambda = \text{blkdiag}\left\{ \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix}, -0.2, -0.5, -1 \right\}.$$

A. $\mathcal{H}_2$ sensitivity reduction

In [5], an LMI method for dealing with $\mathcal{H}_\infty$ and $\mathcal{H}_2$ design with regional pole assignment constraints were presented. The original optimization problems are non-convex and are transformed into some LMIs which provide suboptimal solutions. Although the LMI approach is very efficient, it turns out that the $\mathcal{H}_\infty$ or $\mathcal{H}_2$ norm achieved by the suboptimal controllers can be significantly greater than the actual optimal value. On the other hand, it is also difficult to determine the actual optimal value. Numerical example has indicated that the gradient method proposed in this paper leads to achievable performances with exact pole assignment that are much better than the suboptimal value provided by LMI method with regional pole assignment.

The optimal $\mathcal{H}_2$ norm of $T(s,F)$ without a pole assignment constraint is 2.52, which can be approached only by high gain feedback. Let's consider the $\mathcal{H}_2$ optimization with regional pole assignment, where the convex pole assignment region is

$$\Omega = \{z \in \mathbf{C} : -1 \leq \operatorname{Re} z \leq -0.2, \operatorname{Im} z \leq |\operatorname{Re} z|\}.$$

This region is depicted in Figure 1. All the desired eigenvalues, indicated by $*$ in Figure 1, lie within this region. By solving the $\mathcal{H}_2$ sensitivity reduction problem with regional pole assignment constraint Ω (based on Matlab LMI Toolbox), the suboptimal state feedback

$$F_{s2} = \begin{bmatrix} -48.8290 & 109.5537 & -228.5706 & 191.0731 & -51.0602 \\ -22.8624 & 36.7713 & -59.4091 & 41.7404 & -4.9388 \end{bmatrix},$$

is obtained with $\|(sI - A - B_2F_{s2})^{-1}\|_2 = 10.7068$. The closed-loop eigenvalues are $-0.9970 \pm j0.9749$, -0.7629, -0.3916, -0.2059.

On the other hand, when the gradient approach proposed in this paper is applied to minimize $J_2(U)$, several local minima of $J_2(U)$ are obtained. The smallest one is $J_2(U^*)^{\frac{1}{2}} = \|(sI - A - B_2F^*)^{-1}\|_2 = 6.0516$, with the initial value

$$U_0 = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$

and the optimal value

$$U^* = \begin{bmatrix} 6.4771 & -1.7967 & 0.5273 & 0.7232 & 1.1381 \\ -1.4939 & -1.6194 & 0.2663 & -0.0275 & 0.4836 \end{bmatrix}.$$

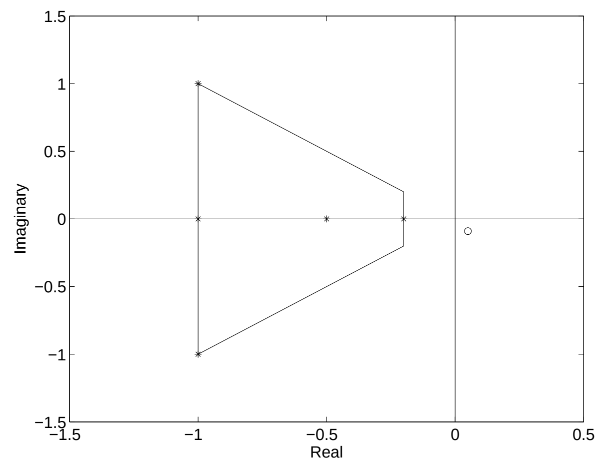

Figure 1: Region for pole assignment

The corresponding optimal feedback gain F^* is

$$F^* = \begin{bmatrix} -41.8857 & 89.2184 & -180.9924 & 151.6352 & -42.3689 \\ -16.7450 & 37.2976 & -49.4027 & 30.4931 & -0.9877 \end{bmatrix}.$$

Note that the state feedback F obtained via $F = UV^{-1}$ always assign the closed-loop eigenvalues to the desired locations. Since the desired eigenvalues are within the region Ω, the optimal $\mathcal{H}_2$ norm with regional pole assignment constraint must be no greater than $J_2(U^*)^{\frac{1}{2}} = 6.0516$. However, the LMI method only gives a suboptimal value 10.7068. We would like to note that, in general, the proposed gradient method requires more computation than the LMI method. For this particular example, the computations were performed on an Ultra-10 Sun workstation. The CPU time used by the LMI method is 2.61 seconds while that by our gradient method is 3.4 seconds.

B. $\mathcal{H}_\infty$ sensitivity reduction

With the same distillation model, the optimal $\mathcal{H}_\infty$ norm without a pole assignment constraint is 8.35, which can be approached only by high gain feedback. By imposing the regional pole assignment constraint Ω to the $\mathcal{H}_\infty$ sensitivity reduction problem and solving this with LMI toolbox, we obtain the suboptimal state feedback

$$F_{s\infty} = \begin{bmatrix} -63.1316 & 89.5281 & -184.7355 & 152.8467 & -38.7542 \\ -33.5325 & 21.5918 & -26.4262 & 12.7533 & 4.2949 \end{bmatrix},$$

with $\|(sI - A - BF_{s\infty})^{-1}\|_\infty = 16.2610$. The eigenvalues of $A + BF_{s\infty}$ are $-0.9864 \pm j0.8919$, -0.7418, -0.5291, -0.2266. The CPU time for the LMI method is 2.49 seconds.

Now we impose the exact pole placement constraint. By using the gradient method proposed in this paper, we found that the gradient algorithm usually terminates at some nondifferentiable points. However, it should be stressed that $J_{\infty 2}(U)$ is more difficult to minimize than $J_2(U)$ due to its nonsmoothness. Also choose

$$U_0 = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{bmatrix}.$$

We get $J_{\infty 2}(U_0) = 55.6604$ and after 6.15 second CPU time, the gradient algorithm outputs an improved performance $J_{\infty 2}(U) = 8.9531$. By starting with different initial values U_0 and choosing the best results, we obtain

$$U^* = \begin{bmatrix} -0.3083 & 1.4352 & 7.7678 & -2.3307 & -1.3210 \\ 4.5134 & -5.9718 & -13.8690 & 3.5832 & 0.0238 \end{bmatrix},$$

with

$$F^* = \begin{bmatrix} -26.8199 & 78.8613 & -137.3429 & 100.1628 & -44.7653 \\ -13.6914 & 18.2826 & -7.5218 & -8.9219 & 5.8129 \end{bmatrix},$$

and $J_{\infty 2}(U^*) = \|(sI - A - B_2F^*)^{-1}\|_\infty = 8.8946$. This value is very close to the unconstrained optimal $\mathcal{H}_\infty$ norm but here we have the desired eigenvalues and the feedback gain is comparatively smaller. As before, the optimal $\mathcal{H}_\infty$ sensitivity with regional pole assignment ought to be no greater than $J_{\infty 2}(U^*) = 8.8946$. However, the LMI approach gives 16.2610.

C. Overshoot reduction

Consider the same distillation model. The closed-loop system matrix is $A + B_2F$. In this case, the indices $J_p(F)$ and $J_v(F)$ are minimized and compared. For the numerical trials, the gradient algorithm terminates at nondifferentiable points which corresponds to the equality of the first two singular values of P or V. Therefore, the attainment of global infimum is not guaranteed. By starting the algorithm from different initial U_0 and selecting the best result, we see that the transient response of the closed-loop system is significantly improved.

Let an initial value U_0 be

$$U_0 = \begin{bmatrix} 1 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix},$$

then the corresponding initial state feedback F_0 is

$$F_0 = \begin{bmatrix} -36.9350 & 53.0168 & -102.0848 & 81.1492 & -23.8017 \\ -47.5685 & -16.1156 & 47.2416 & -49.4003 & 23.1897 \end{bmatrix}.$$

Using the gradient algorithms, we determined two state feedbacks,

$$F_1^* = \begin{bmatrix} -80.4804 & 90.3607 & -193.0810 & 161.1071 & -40.7114 \\ -47.7592 & 21.8898 & -31.7364 & 18.3788 & 2.9646 \end{bmatrix},$$

and

$$F_2^* = \begin{bmatrix} -37.5812 & 83.6861 & -174.2750 & 147.4140 & -54.5433 \\ -15.2826 & 38.3116 & -51.7317 & 30.7113 & -1.7976 \end{bmatrix},$$

which correspond respectively to the smallest $J_v(F)$ and $J_p(F)$, with $J_v(F_1^*) = 31.4998$, $J_p(F_2^*) = 5.5435$. The transient responses $\|x(t)\|$ of the closed-loop system using state feedbacks F_0, F_1^* and F_2^* are compared in Figure 2 for 50 randomly generated initial states x_0 normalized to unit length. The three plots from left to the right correspond to F_0, F_1^* and F_2^*, respectively.

The time response $\|x(t)\|$ with $\|x(0)\| = 1$ is governed by $\bar{\sigma}\left(e^{(A+BF)t}\right)$, i.e,

$$\|x(t)\| \le \bar{\sigma}\left(e^{(A+BF)t}\right).$$

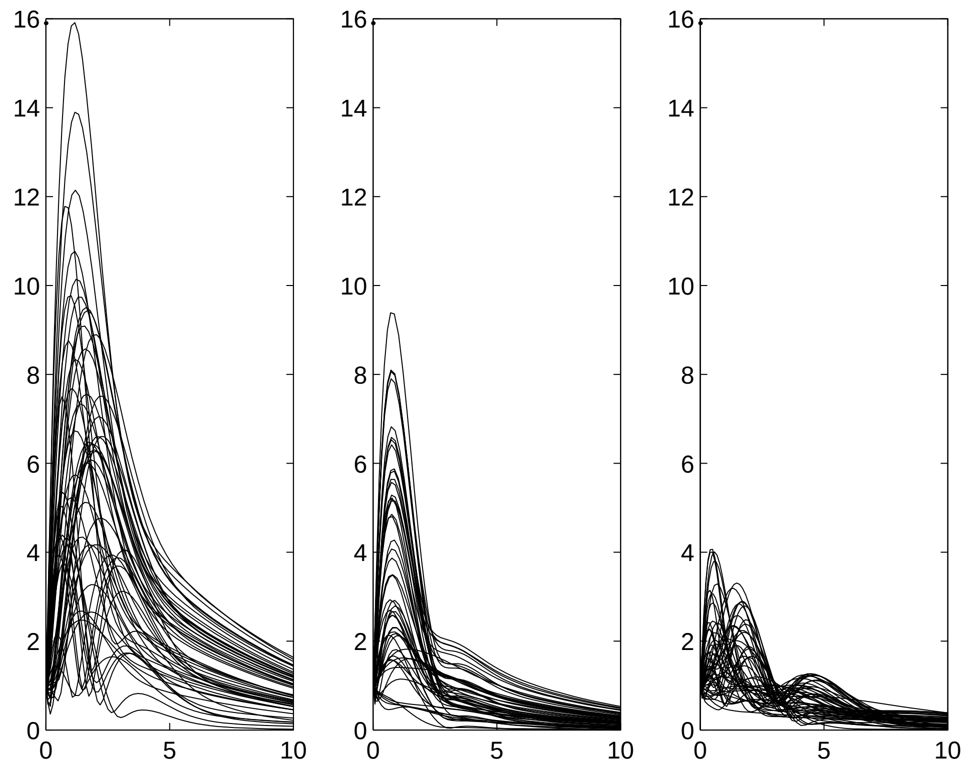

Figure 2: Comparison of the transient responses

See Figure 3 for a comparison of $\bar{\sigma}\left(e^{(A+BF)t}\right)$ with $F = F_0, F_1^*, F_2^*$, where the solid curve, the dashed curve and the dotted curve correspond to F_0, F_1^* and F_2^*, respectively. The maximum overshoots are

$$J_o(F_0) = 15.9341, \quad J_o(F_1^*) = 9.4326, \quad J_o(F_2^*) = 4.2928.$$

We see that the transient response is significantly improved by decreasing $J_p(F)$, even though we cannot guarantee to locate the global minimum of the index function. Obviously, the condition number of P as an index of overshoot is much better than the condition number of V, as is illustrated in the following table.

Feedback Gain	$J_v(F)$	$J_p(F)$	$J_o(F)$
F_0	86.0101	20.3194	15.9341
F_1^*	31.4998	12.0735	9.4326
F_2^*	183.7992	5.5435	4.2928

Table 1. Comparison of index function values

Clearly $J_p(F)$ gives us a much tighter upper bound for the overshoot than $J_v(F)$. The overshoot corresponding to F_2^* is the smallest but the prediction from $J_v(F)$ is fairly large.

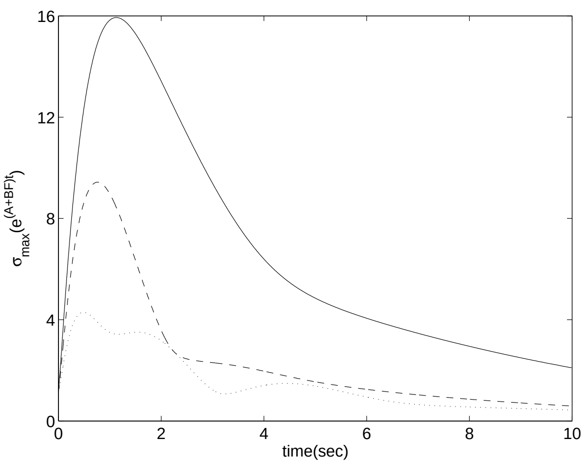

Figure 3: Comparison of $\bar{\sigma}\left(e^{(A+BF)t}\right)$

7 Conclusions

We have presented a unified approach for optimizing performance indices under the constraint of pole assignment. Taking this unified framework, several performance index optimization problems with pole placement constraint were solved. Even though the gradient method does not guarantee globally optimal solutions, examples show that the approach can improve the performance significantly. For instance, in $\mathcal{H}_2$ and $\mathcal{H}_\infty$ norm optimization, comparison reveals that the gradient approach has the potential to achieve performances better than what could be achieved by some convex yet suboptimal optimization.

References

[1] Bhattacharyya, S.P. and de Souza, E., 1982, "Pole assignment via Sylvester equation," *Systems & Control Letters*, 1, pp. 261-263.

[2] S. Boyd, L. El Ghaoui, E. Feron and V. Balakrishnan, *Linear Matrix Inequalities in Systems and Control Theory*, SIAM Studies in Appl. Mathematics, Philadelphia, 1994.

[3] Byers, R. and Nash, S.G., 1989, Approaches to robust pole assignment, *Int. J. Control*, 49, pp. 97-117.

[4] Cavin III, R. K. and Bhattacharyya, S. P., 1983, "Robust and well conditioned eigenstructure assignment via Sylvester's equation," *Optimal Control: Applic. and Methods*, 4, pp. 205-212.

[5] Chilali, M and Gahinet, P, 1996, "$\mathcal{H}_\infty$ design with pole placement constraints: an LMI approach", *IEEE Transactions on Automatic Control*, 41, pp. 358-367.

[6] Dickman, A., 1987, On the robustness of multivariable linear feedback systems in state-space representation, *IEEE Transactions on Automatic Control*, 32, pp. 407-410.

[7] Gourishankar, V. and Ramer, K., "Pole assignment with minimum eigenvalue sensitivity to plant parameter variations," *Int. J. Contr.*, 23, pp. 493-540, 1976.

[8] Hu, T., and Shi, S., 1989, "The set of feedback matrices that assign the poles of a system," *Proc. of MTNS-89: Progress in Systems and Control Theory 4*, pp. 129-135, Birkhauser, 1990.

[9] Hu, T., and Lam, J., "Improvement of parametric stability margin under pole assignment," *IEEE Trans. Automat. Contr.*, to appear. (see also *Proc. of 1998 ACC* , pp. 2797-2801.)

[10] Hu, T., and Lam, J., "On optimizing performance indices with pole assignment constraints," *Proceedings of IME Part I – Journal of Systems and Control Engineering*, Vol. 212, pp. 327-337, 1998.

[11] Keel, L.H., Fleming, J.A., and Bhattacharyya, S. P., 1985, "Minimum norm pole assignment via Sylvester's equation," *Contemporary Mathematics Series (American Mathematical Society)*, 47, pp. 265-272.

[12] Keel, L.H., and Bhattacharyya, S. P., "State space design of low-order stabilizers," *IEEE Trans. Automat. Contr.*, 35, pp. 182-186, 1990.

[13] Kautsky, J., and Nichols, N.K., 1990, "Robust pole assignment in systems subjected to structured perturbations," *Systems & Control Letters*, 15, pp. 373-380.

[14] Kautsky, J., Nichols, N.K., and van Dooren, P., 1985, "Robust pole assignment in linear state feedback systems," *Int. J. Control*, 41, pp. 1129-1155.

[15] Lam, J. and H. K. Tam, 1997, "Robust partial pole-placement via gradient flow," *Optimal Control: Applic. and Methods*, 18, pp.371-379.

[16] Lam, J. and Yan, W., 1995, "A gradient flow approach to robust pole-placement problem," *Int. J. Robust and Nonlinear Control*, 5, pp. 175-185.

[17] Oh, M, et al., "Robust pole assignment in a specified region using output feedback," *Optim. Contr. Appl. Meth.*, Vol. 14, 57-66, 1993.

[18] Saad, Y., 1988, "Projection and deflation methods for partial pole assignment in linear state feedback," *IEEE Transactions on Automatic Control*, 33, pp. 290-297.

[19] Stewart, G. W., and Sun, J. G., *Matrix Perturbation Theory*, San Diego, CA: Academic, 1990.

[20] Sun, J. G., "On numerical methods for robust pole assignment in control design," *J. Computational Math.*, 5, pp. 119-134, 1987.

[21] Tits, A.L., and Yang, Y., "Globally convergent algorithm for robust pole assignment by state feedback,", *IEEE Trans. Automat. Contr.,* 41, pp. 1432-1452, 1996.

[22] White, B.A, 1995, "Eigenstructure assignment: a survey," *Proc. I.Mech.E, Part I*, 209, pp. 1-11.

[23] Wilkinson, J. H., *The Algebraic Eigenvalue Problem*, Oxford: Charendon, 1965.

[24] Yang Y. and A.L. Tits, "On robust pole assignment by state feedback," *Proc. American Control Conf.*, pp. 2765-2766, 1993.

[25] Zhou, K., with Doyle, J. C. and Glover, K., 1996, *Robust and Optimal Control*, Prentice Hall, Upper Saddle River, NJ.