OPTIMAL HANKEL-NORM APPROXIMATION OF IIR BY FIR SYSTEMS

Li Chai[§†], *Jingxin Zhang*[§], *Cisheng Zhang*[‡] *and Edoardo Mosca*[*]

[§]Department of Electrical and Computer Systems Engineering Monash University, Clayton, VIC3800, Australia

[†]Department of Automation, Hangzhou Dianzi University, Hang Zhou, P.R. China

[‡]School of Electrical and Electronic Engineering
Nanyang Technological University, Singapore

[*]Dipartimento di Sistemi e Informatica, Universityá di Firenze, Italy

ABSTRACT

This paper presents a constructive method to (sub)optimal finite impulse response (FIR) approximation of a given infinite impulse response (IIR) model. The method minimizes the Hankel norm of approximation error by using the explicit solution of norm-preserve dilation problem. It has the advantage over the existing methods that it provides an explicitly constructive solution and allows the tradeoff between the Chebyshev and least square criteria. The lower and upper bounds on the l^2 and Chebyshev norms of approximation error are given. The effectiveness and properties of the proposed algorithm are demonstrated through a computation example.

Keywords: Hankel-norm; FIR approximation; norm-preserve dilation; mixed norm design; IIR filters

1. INTRODUCTION

Finite impulse response (FIR) models have the advantage of intrinsically stable properties and easy implementation, thus are more preferred to infinite impulse response (IIR) models [1, 2]. However, we often end up with IIR models in system and signal modelling, filter and controller design, etc. Therefore, effective methods are required to approximate an IIR model by FIR model. Generally, the approximation problem can be stated as follows:

Given $G(z)$, a stable rational transfer function, find

$$F(z) = f_0 + f_1 z^{-1} + \cdots + f_{m-1} z^{-m+1}$$

such that the norm of the error $\|G(z) - F(z)\|$ is minimized, where $||\cdot||$ could be different norms corresponding to different design criteria.

The early methods to the approximation use direct truncation of impulse response that minimizes the least-square error criterion, or equivalently the l^2 error norm $\|G(z) - F(z)\|_2$. In [3, 4, 5], the minimum Chebyshev error criterion, or equivalently, the Chebyshev (H) error norm $\|G(z) - F(z)\|$ is used. In [4, 5], a method called Nehari Shuffle is proposed and upper and lower bound on the approximation error are derived. However, the Nehari Shuffle doesn't provide the optimal solutions with respect to Chebyshev norm. A direct Chebyshev norm optimization approach is given by the powerful tool of linear matrix inequalities (LMIs) [3].

As pointed out in [6], the least square criterion is appropriate if the input signal is narrow-band, and Chebyshev criterion is appropriate if the input signal is wide-band and distributed approximately uniformly in the frequency. Thus, there are situations where neither the Chebyshev criterion nor the least square criterion is appropriate, and where we call for alternative design methods with trade-off between least square and Chebyshev criteria [6, 7]. The trade-off design issues are studied extensively by Adams group, see [6, 8] and references therein. The least p-power error design is discussed in [7].

In this paper, the Hankel norm of the error is chosen to be minimized. Hankel-norm approximation is extensively used in model reduction since the remarkable work of Glover [9, 10]. However, the problem here is different from that of [9], which is to find a lower order IIR model for a given high order IIR model. The resulting method of this paper has the following advantages.

- It allows the tradeoff between least-square criterion and Chebyshev criterion.
- The design algorithm is constructive, and only involves algebraic manipulations, therefore no iteration and convex optimization program (as LMIs) are needed.
- No need to carry out balanced realization and truncation as [5].
- Lower and Upper bounds on l^2 norm and Chebyshev norm are also provided.

The remainder of this paper is as follows. Section 2 provides some necessary background material on Hankel operators and norm-preserve dilations, Section 3 develops the approximation algorithm, Section 4 gives a computation example, and Section 5 presents conclusions. Due to space limit, all the proofs of the lemmas and theorems are omitted, except that of Theorem 3.

2. PRELIMINARY

This section introduces the notations and some preliminary results used in the sequel. For a matrix X, let X^* denote its complex conjugate transpose, (X) its eigenvalue, and (X) its singular value. Denote the spectrum norm of X as $||X|| = (\bar{\ } (X^*X))^{\frac{1}{2}}$, where $\bar{\ }$ denotes the largest eigenvalue of X. For a positive definite matrix X, we use $X^{\frac{1}{2}}$ to denote its Hermitian square root, that is, $X^{\frac{1}{2}}X^{\frac{1}{2}} = X$ and $(X^{\frac{1}{2}})^* = X^{\frac{1}{2}}$.

2.1 Spaces, norms and Hankel Operators

Definition 1 *Given a causal transfer function $G(z)$, (A,B,C,D) is called a state space realization if $G(z) = D + C(zI - A)^{-1}B$, where $A \in R^{n\times n}$, $B \in R^{n\times 1}$, $C \in R^{1\times n}$ and $D \in R$.*

Definition 2 *For a stable transfer function with state space realization $G(z) = D + C(zI - A)^{-1}B$, the controllability and observability Gramian, denoted by P and Q, is defined by $P = \ _{k=0} A^k BB^* A^{*k}$ and $Q = \ _{k=0} A^{*k} C^* C A^k$.*

It is well known that P and Q can be computed from the following Lyapunov equations respectively

$$APA^* - P + B^*B = 0 \quad (1)$$
$$A^*QA - Q + C^*C = 0. \quad (2)$$

The realization is minimal if P and Q are nonsingular.

For a stable and causal $G(z) = \ _{k=0} g_k z^{-k}$, the l^2 norm of $G(z)$, denoted by $||G(z)||_2$, is given by

$$||G(z)||_2^2 = \ _{k=0} g_k^2 = \frac{1}{2} \int_{-} G^*(e^{j}) G(e^{j}) d \ . \quad (3)$$

The Chebyshev norm (or H norm) of $G(z)$, denoted by $||G(z)||$, is given by

$$||G(z)|| \quad = \quad \max_{\in(- \ , \ ]} |G(e^{j})|. \tag{4}$$

Note that we assume the right hand sides of (3) and (4) are all well-defined for simplicity. For the rigorous definitions, please refer to [11]. The Hankel operator of G, denoted by $_G$, is defined as

$$_G = \begin{bmatrix} g_1 & g_2 & g_3 & \cdots \\ g_2 & g_3 & g_4 & \cdots \\ g_3 & g_4 & g_5 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

The Hankel singular values of $G(z)$, denoted by $_i(\ _G)$, $i = 1, \ldots, n$ are the ith singular values of $_G$. The Hankel norm of G, denoted by $||\ _G||$ is defined to be the largest singular value of $_G$, i.e. $||\ _G|| = \ _1(\ _G)$. The following can be used to compute the Hankel-norm of a transfer function, see [9, 11] for details.

Lemma 1 *For the above $G(z)$, we have*

$$\begin{aligned} {}^2_i(\ _G) &= \ _i(QP) = \ _i(Q^{\frac{1}{2}}PQ^{\frac{1}{2}}) = \ _i(P^{\frac{1}{2}}QP^{\frac{1}{2}}) \\ ||\ _G||^2 &= \bar{\ }(QP) = \bar{\ }(Q^{\frac{1}{2}}PQ^{\frac{1}{2}}) = \bar{\ }(P^{\frac{1}{2}}QP^{\frac{1}{2}}) \end{aligned}$$

where P and Q are controllability and observability Gramians respectively.

2.2 Norm-preserve Dilations

Consider the block matrix $\begin{bmatrix} X & R \\ S & T \end{bmatrix}$, where X, R, S and T are matrices of compatible dimensions, and denote

$$(X) = \left\| \begin{bmatrix} X & R \\ S & T \end{bmatrix} \right\|.$$

The norm-preserve dilation problem is to find X such that (X) is minimized for given matrices R, S, and T. Denote

$$_0 = \min_X \ (X). \tag{5}$$

The following results play a very important role in our development [12].

Lemma 2 *The minimum $_0$ in (5) is given by*

$$_0 = \max \left\{ ||[\ S \quad T\ ]||, \left\| \begin{matrix} R \\ T \end{matrix} \right\| \right\}.$$

Moreover, assume $\ \geq \ _0$, then the solution set X such that $(X) \leq$ can be characterized by

$$X = -YT^*Z + \ (I - YY^*)^{1/2}W(I - Z^*Z)^{1/2} \tag{6}$$

where W is an arbitrary contraction ($||W|| \leq 1$) and Y and Z are contractions satisfying

$$\begin{aligned} R &= Y(\ ^2I - T^*T)^{1/2} && (7) \\ S &= (\ ^2I - TT^*)^{1/2}Z. && (8) \end{aligned}$$

The following lemma gives a more explicit formula when $||T|| <$.

Lemma 3 *Assume that $\ \geq \ _0$ and $||T|| <$. Then the solution set X such that $(X) \leq$ can be characterized by*

$$\begin{aligned} X = &-R(\ ^2I - T^*T)^{-1}T^*S + \ [I - R(\ ^2I - T^*T)^{-1}R^*]^{\frac{1}{2}} \\ &W[I - S^*(\ ^2I - TT^*)^{-1}S]^{\frac{1}{2}}. \end{aligned} \tag{9}$$

The norm-preserve dilation problem is solved independently by Parrot and Davis et. al. For more detail, please refer to [12].

3. HANKEL-NORM FIR APPROXIMATION

In this section, an algorithm is developed to solve the (sub)optimal Hankel-norm FIR approximation of a given IIR model. First, we present a basic theorem from which the approximation can be converted to a matrix norm-preserving dilation problem. Then a constructive algorithm is developed step by step. Finally, some properties of the resulting FIR approximation are discussed and the bounds on error norms are given.

The problem to be considered in this section is as follows. Given an IIR model $G(z) = D + C(zI - A)^{-1}B$, find an FIR model $F(z) = f_0 + f_1z^{-1} + \cdots + f_{m-1}z^{-m+1}$ that minimizes $||\ _E||$, the Hankel norm of the approximation error $E(z) = z^{-1}(G(z) - F(z))$. The reason we put a delay term z^{-1} in $E(z)$ is due to the fact that the Hankel norm of a system is unrelated to the constant term. The relation of Hankel norm, l^2-norm and Chebyshev-norm are given in the following lemma

Lemma 4 *Let $E(z) = \ _{i=1} e_iz^{-i}$ satisfy $||\ _E|| <$. Then we have*

$$||E(z)||_2 \leq ||\ _E|| \leq ||E(z)|| \ \leq 2 \sum_{i=1}^{N} \ _i(\ _E)$$

where $_i(\ _E)$ is the ith singular value of $E(z)$ and N=rank$(\ _E) =$ McMillan degree of $E(z)$.

The first two inequalities are shown in [13] and the last inequality is shown in [11, 5].

Lemma 4 tells us that the Hankel-norm can be seen as the trade-off between l^2-norm and Chebyshev-norm. The following theorem is important to the development of our algorithm.

Theorem 1 *For $G(z) = D + C(zI - A)^{-1}B$, define $H(z) = z^{-1}G(z)$. Then we have*

$$||\ _H|| = \left\| \begin{matrix} D & CP^{\frac{1}{2}} \\ Q^{\frac{1}{2}}B & Q^{\frac{1}{2}}AP^{\frac{1}{2}} \end{matrix} \right\|$$

where P and Q are solutions of Lyapunov equations (1) and (2) respectively.

Theorem 2 *Given a transfer function $G(z) = C(zI - A)^{-1}B$ with $||\ _G|| = \ _0$, define $H(z) = z^{-1}(D + G(z))$ for a scalar D. Then we have*
(i) $||\ _H|| \geq \ _0$ for any D.
(ii) There exist D's such that $||\ _H|| = \ _0$, and all such D's can be characterized by

$$D = -YP^{\frac{1}{2}}A^*Q^{\frac{1}{2}}Z + \ _0(I - YY^*)^{\frac{1}{2}}(I - Z^*Z)^{\frac{1}{2}}w \tag{10}$$

where $|w| \leq 1$ and Y and Z are contractions satisfying

$$\begin{aligned} CP^{\frac{1}{2}} &= Y\left(\ ^2_0I - P^{\frac{1}{2}}A^*QAP^{\frac{1}{2}}\right)^{\frac{1}{2}} \\ Q^{\frac{1}{2}}B &= \left(\ ^2_0I - P^{\frac{1}{2}}A^*QAP^{\frac{1}{2}}\right)^{\frac{1}{2}}Z. \end{aligned}$$

(iii) For any $\ > \ _0$, all D's such that $||\ _H|| \leq$ are given by

$$D = \ + \sqrt{\ }\ w \tag{11}$$

where $|w| \leq 1$, and

$$\begin{aligned} &= -C(\ ^2P^{-1} - A^*QA)^{-1}A^*QB && (12) \\ &= \ ^2\left[1 - C(\ ^2P^{-1} - A^*QA)^{-1}C^*\right] \\ &\quad \left[1 - B^*(\ ^2Q^{-1} - APA^*)^{-1}B\right]. && (13) \end{aligned}$$

We are now ready to present the main result of this section. Before presentation, we recall the following well known fact [5]: a causal transfer function $G(z) = D + C(zI - A)^{-1}B$ can be written in the form $G(z) = G_1(z) + z^{-m+1}G_m(z)$, where

$$G_1(z) = \sum_{i=0}^{m-1} g_i z^{-i} \quad (14)$$

with g_i being the first m impulse responses of $G(z)$, and $G_m(z)$ is a strictly proper (rational function of z) transfer function.

Theorem 3 *Given a stable and causal transfer function $G(z) = G_1(z) + z^{-m+1}G_m(z)$ with $\|\ _{G_m}\| = \ _0$ and a positive number $\ \geq\ _0$, a sequence of numbers $e_0, \ldots, e_{m-1}$ can be found to construct*

$$E(z) = \sum_{i=0}^{m-1} e_i z^{-i} + z^{-m+1}G_m(z)$$

such that $\|\ _E\| \leq\ $.

Proof. We will prove the theorem in a constructive manner by showing that e_{m-i-1} can be computed if e_{m-i-1} is obtained. Denote $E_m(z) = G_m(z)$ and

$$E_{m-i-1}(z) = z^{-1}(e_{m-i-1} + E_{m-i}(z)) \quad (15)$$

for $i = 0, \ldots, m-1$. Now assume that a state-space realization for $E_{m-i}(z)$ is given by

$$E_{m-i}(z) = C_{m-i}(zI - A_{m-i})^{-1}B_{m-i}. \quad (16)$$

Then the controllability and observability Gramians P_{m-i} and Q_{m-i} can be computed from equations (1) and (2) respectively. Since $\ \geq\ _0$, it then follows from Theorem 2 that there exist e_{m-i-1} such that $\|\ _{E_{m-i-1}}\| \leq\ $ where $E_{m-i-1}(z)$ is defined by (15). Moreover, if $\ >\ _0$, then those e_{m-i-1} are given by

$$e_{m-i-1} = \ _{m-i} + \sqrt{\ _{m-i}}\, w_{m-i} \quad (17)$$

where

$$\begin{aligned} _{m-i} &= -C_{m-i}P_{m-i} \\ &\cdot(\ ^2 I - A^*_{m-i}Q_{m-i}A_{m-i}P_{m-i})^{-1}A^*_{m-i}Q_{m-i}B_{m-i} \end{aligned} \quad (18)$$

$$\begin{aligned} _{m-i} &= \\ ^2&[1 - C_{m-i}P_{m-i}(\ ^2 I - A^*_{m-i}Q_{m-i}A_{m-i}P_{m-i})^{-1}C^*_{m-i}] \\ &[1 - B^*_{m-i}Q_{m-i}(\ ^2 I - A_{m-i}P_{m-i}A^*_{m-i}Q_{m-i})^{-1}B_{m-i}] \end{aligned} \quad (19)$$

and $|w_{m-i}| \leq 1$. It is easy to check that a state space realization for $E_{m-i-1}(z)$ is given by

$$E_{m-i-1}(z) = C_{m-i-1}(zI - A_{m-i-1})^{-1}B_{m-i-1} \quad (20)$$

where $A_{m-i-1} = \begin{bmatrix} A_{m-i} & 0 \\ C_{m-i} & 0 \end{bmatrix}$, $B_{m-i-1} = \begin{bmatrix} B_{m-i} \\ e_{m-i} \end{bmatrix}$ and $C_{m-i-1} = \begin{bmatrix} 0 & I \end{bmatrix}$. The controllability and observability Gramians P_{m-i-1} and Q_{m-i-1} for the state space realization (20) are as follows

$$\begin{aligned} P_{m-i} &= \begin{bmatrix} P_{m-i+1} & A_{m-i+1}P_{m-i+1}C^*_{m-i+1} + B_{m-i+1}e_{m-i} \\ C_{m-i+1}P_{m-i+1}A^*_{m-i+1} + e_{m-i}B^*_{m-i+1} & C_{m-i+1}P_{m-i+1}C^*_{m-i+1} + e^2_{m-i} \end{bmatrix} \\ Q_{m-i} &= \begin{bmatrix} Q_{m-1} & 0 \\ 0 & 1 \end{bmatrix}. \end{aligned}$$

The proof is then completed by noting that we can now compute e_{m-i-2} by Theorem 2 again. □

The proof of Theorem 3 provides us an algorithm to compute the m-length FIR approximation of a given IIR model $G(z) = D + C(zI - A)^{-1}B$ which achieves a suboptimal Hankel-norm. This algorithm is summarized below.

Algorithm 1

1) Set $G_m(z) = C_m(zI - A_m)^{-1}B_m$, where $C_m = CA^{m-1}$, $A_m = A$ and $B_m = B$.

2) Obtain P_m and Q_m by solving the Lyapunov equations (1) and (2) and compute $P_m^{\frac{1}{2}}$ and $Q_m^{\frac{1}{2}}$.

3) Compute the Hankel norm of $G_m(z)$ by any of the following equations

$$\|\ _{G_m}\| = \bar{\ }(Q_m P_m) = \bar{\ }(Q_m^{\frac{1}{2}} P_m Q_m^{\frac{1}{2}}) = \bar{\ }(P_m^{\frac{1}{2}} Q_m P_m^{\frac{1}{2}})$$

4) Obtain e_{m-1} by the following equation

$$e_{m-1} = -C_m P_m(\ ^2 I - A^*_m Q_m A_m P_m)^{-1}A^*_m Q_m B_m. \quad (21)$$

5) Obtain a state space realization of $E_{m-1}(z)$ from (20) and obtain P_{m-1} and Q_{m-1}.

6) Repeat step 3) and 4) to find $e_{m-2}, \ldots, e_0$.

7) The optimal Hankel-norm approximant $F(z)$ is then given by

$$F(z) = G_1(z) - \sum_{i=0}^{m-1} e_i z^{-i} = \sum_{i=0}^{m-1}(g_i - e_i)z^{-i} := \sum_{i=0}^{m-1} f_i z^{-i}.$$

The following Corollary gives the lower and upper bounds on the l^2 and Chebyshev norms of the approximation error of the above algorithm.

Corollary 1 *For $G(z) = G_1(z) + z^{-m+1}G_m(z)$, let $F(z)$ be obtained by algorithm 1. Then the following holds for the approximation error $E(z) = G(z) - F(z)$.*

$$\begin{aligned} \|\ _{E_0}\| &\leq \|E(z)\| \leq 2\sum_{i=1}^{N}\ _i(\ _E) \\ \|G_m(z)\|_2 &\leq \|E(z)\|_2 \leq \|\ _{E_0}\|. \end{aligned}$$

Corollary 1 tells us that the upper bound on the Chebyshev-norm of approximation error is $2\sum_{i=1}^{N}\ _i(\ _E)$. Actually we can achieve a tighter upper bound simply by another choice of f_0. The result is as follows, see [9, 11] for details.

Corollary 2 *For $G(z) = G_1(z) + z^{-m+1}G_m(z)$, let $f_1 \cdots f_{m-1}$ be chosen as in algorithm 1. If f_0 is chosen such that $\|G(z) - F(z)\|$ is minimized, then we have $\|G(z) - F(z)\| \leq \sum_{i=1}^{N}\ _i(\ _E)$.*

4. COMPUTATION EXAMPLE

Given below is a 6th order IIR model $G(z)$. This is the model of spindle vibration we obtained at a hot steel rolling mill for prediction and reduction of mechanical failure [14]. The model is non-minimum phase and has a pole very close to unit circle. Hence, it is prone to numerical error and not suitable for DSP implementation. To overcome this implementation difficulty, an FIR approximation is required.

$$\begin{aligned} G(z) &= \frac{-0.1242z^5 + 0.1581z^4 + 0.5273z^3}{z^6 - 1.095z^5 + 1.299z^4 - 1.113z^3} \\ &\quad \frac{+0.2154z^2 - 0.0647z^1 + 0.6889}{+1.028z^2 - 0.6043z + 0.426} \end{aligned}$$

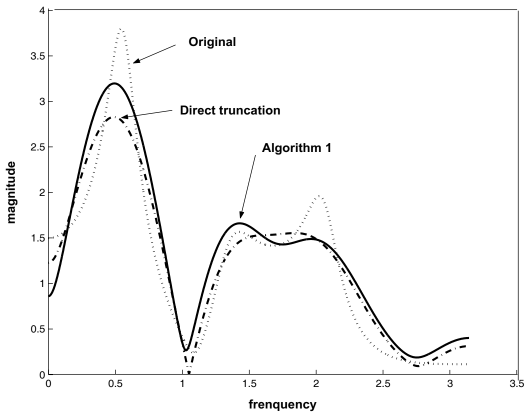

Figure 1: Comparison of frequency response for $m = 12$

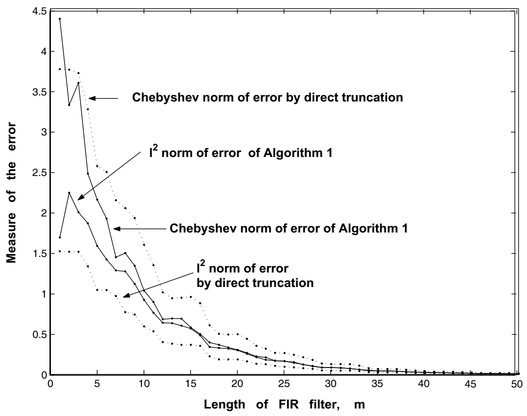

Figure 2: Bounds of l^2 and Chebyshev norms for error systems

As shown in Figure 1, the model's frequency response spikes at about $\ = 0.6, 1.4, 2.1$. These spikes, particularly those two at $\ = 0.6, 1.4$ cause mechanical damage to the spindle [14]. Thus, for this particular application, we need an FIR approximation that better captures these two spikes. Now we use Algorithm 1 to find an FIR approximation of the model with length $m = 12$.

Figure 1 compares the frequency responses of the original IIR model and those of the 12-length FIR approximations obtained by Algorithm 1 and by direct truncation of impulse response. We can see from the figure that the FIR approximation of Algorithm 1 better captures the frequency spikes at $\ = 0.6, 1.4$, whereas that of direct truncation tends to smooth out these spikes. Compared with the IIR model, the 12-length FIR approximation of Algorithm 1 has the same arithmetic complexity and very similar responses in the frequency range $\leq \ \leq 1.75$ that is critical to the application. But it is numerically more robust since its intrinsic stability.

Figure 2 compares the Chebyshev and l^2 norms of approximation errors achievable by Algorithm 1 with those of direct truncation. As can be seen from the figure, the Chebyshev error norms are above the l^2 error norms, and the Chebyshev (l^2) error norm achievable by Algorithm 1 is below (above) that of direct truncation. These agree with the analysis of Corollary 2, and demonstrate that Algorithm 1 truly provides a trade off between the Chebyshev and l^2 approximation criteria.

5. CONCLUSION

A constructive method is presented to obtain the optimal FIR Hankel norm approximation for a given IIR model. This method can provide a trade-off design between the worse case Chebyshev criterion and the least square criterion. Lower and upper bounds on the l^2 and Chebyshev error norms are provided for the Hankel norm approximate. The effectiveness and properties of the proposed algorithm are demonstrated through a computation example. The algorithm can be extended to MIMO systems directly which may provide potential application to filter banks design.

Acknowledgments: This research was supported in part by the Australian Research Council under grant DP03430457 and in part by the National Natural Science Foundation of China under grant 60304011 and 60434020.

REFERENCES

[1] S. K. Mitra, *Digital Signal Processing: A Computer-Based Approach*, McGraw-Hill, Irwin, 2nd edition, 2001.

[2] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, Englewood Cliffs, NJ: Prentice-Hall, 2nd edition, 1989.

[3] Y. Yamamoto, B.D.O. Anderson, M. Nagahara, and Y. Koyanagi, "Optimizing FIR approximation for discrete-time IIR filters," *IEEE Signal Processing Letters*, vol. 10, no. 9, pp. 273–276, September 2003.

[4] P. J. Kootsookos, R. R. Bitmead, and M. Green, "The Nehari shuffle: FIR(q) filter design with guaranteed error bounds," *IEEE Trans. Signal processing*, vol. 40, pp. 1876–1883, August 1992.

[5] P. J. Kootsookos and R. R. Bitmead, "The Nehari shuffle and Minimax FIR filter design," in *Control and Dynamic Systems*, vol. 64, pp. 239–298. Orlando, FL: Academic Press, 1994.

[6] J. W. Adams, "FIR digital filters with least-squares stopbands subject to peak-gain constraints," *IEEE Trans. Circuits Syst.*, vol. 39, pp. 376–388, April 1991.

[7] C. S. Burrus and J. Barreto, "Least p-power error design of FIR filters," in *Proc. IEEE Int. Symp. Circuits Syst.*, San Diego, CA, 1992, pp. 545–548.

[8] J. W. Adams and J. L. Sullivan, "Peak-constrained least-squares optimization," *IEEE Trans. Signal Processing*, vol. 46, pp. 306–321, February 1998.

[9] K. Glover, "All optimal Hankel-norm approximations of linear multivariable systems and their l -error bounds," *Int. J. Control*, vol. 39, pp. 1115–1193, June 1984.

[10] G. Obinata and B. D. O. Aderson, *Model Reduction for Control System Design*, Berli, Germany: Spring-Verlag, 2001.

[11] K. Zhou, J. C. Doyle, and K. Glover, *Robust Optimal Control*, Prentice-Hall, 1996.

[12] C. Davis, W. M. Kahan, and H. F. Weinberger, "Norm-preserving dilitions and their applications to optimal error bounds," *SIAM Journal on Numerial Analysis*, vol. 19, no. 3, pp. 445–469, June 1982.

[13] C. K. Chui and G. Chen, *Discrete H Optimization: With Application in Signal Processing and Control Systems*, Springer-Verlag, 2nd edition, 1997.

[14] K. Zhang, J. Zhang and S. Nahavandi, Mechanical Failure Reduction of Hot Steel Rolling Mills via Modeling and Control, *The 4th Asian Control Conference*, pp. 1630-1635, Singapore, Sept. 2002.