Proceedings of the 45th IEEE Conference on Decision & Control
Manchester Grand Hyatt Hotel
San Diego, CA, USA, December 13-15, 2006

Model Matching and Filter Design using Orthonormal Basis Functions

Jie Zeng and Raymond de Callafon

***Abstract*— Affine model parametrizations using orthonormal basis functions have been widely used in system identification and adaptive signal processing. The main advantage of using orthonormal basis functions in a (generalized) orthonormal Finite Impulse Response (FIR) filter lies in the possibility of incorporating prior knowledge of the system dynamics into the filter design and approximation process. As a result, more accurate and simplified models can be obtained with a limited number of basis functions. In this paper the linear parameter structure of a generalized FIR filter is used to formulate analytic solutions for model matching problems. Several construction methods of orthonormal basis functions are discussed and a case study using the generalized FIR filter to approximate the dynamics of an optimal feed-forward filter is presented.**

I. INTRODUCTION

Using orthonormal basis functions to parametrize and estimate dynamic systems [1] is a reputable approach in model estimation techniques [2], [3], frequency domain identification methods [4] or realization algorithms [5], [6]. In the development of orthonormal basis functions, Laguerre and Kautz basis functions have been used successfully in system identification and signal processing [7], [8]. A unifying construction proposed in [9] generalized both the Laguerre and Kautz basis in the context of system identification and generalization of these results for arbitrary dynamical systems were also reported in [1]. The generalized orthonormal basis and unifying construction can be used for systems with wide range of dominant modes, i.e, both high frequency and low frequency behavior [3].

In this paper, the use of a generalized FIR filter $F(z,\theta)$ is considered to solve a model matching problem of the form

$$\min_{\theta} \|H_0(z) - F(z,\theta)G_0(z)\|_2 \tag{1}$$

where $H_0(z)$ and $G_0(z)$ are models that are stable but not necessarily have a stable inverse. The model matching in (1) occurs in many applications that require the computation of a stable filter to approximate an unstable system. The problem (1) arises for example in Active Noise Control (ANC), there $H_0(z)$ is primary noise path, $G_0(z)$ is second noise path, and $F(z,\theta)$ is a feedforward controller to be designed [10], [11]. This paper provides analytic solutions to the optimization of (1) in case $F(z,\theta)$ is a generalized FIR filter, parameterized using an orthonormal basis function.

Furthermore, it has been shown in [1] that if the dynamics of the chosen orthonormal basis functions resembles the dynamics of the system to be approximated, the convergence rate of the affine series expansion will increase. As a result, the number of linear combinations of basis functions to accurately approximate the system dynamics can be kept relatively small. Therefore, the choice of the orthonormal basis is an important issue in order to obtain accurate models and in this paper a comparison is made between the approximation results for different orthonormal basis functions that use the same prior knowledge on the system dynamics of $H_0(z)$ and $G_0(z)$.

II. STRUCTURE OF ORTHONORMAL BASIS FUNCTIONS

Considering a linear time invariant stable discrete time system $F_0(z)$ and suppose the functions $V_k(z),\ k=0,1,2,\ldots$ serve as a orthonormal basis $\{V_k(z)\}$ for the set of systems in $\mathcal{H}_2$. Then there exists a unique expansion

$$F_0(z) = \sum_{k=0}^{\infty} L_k V_k(z) \tag{2}$$

where $L_k,\ k=0,1,2,\ldots$ are the generalized orthonormal expansion coefficients for the basis $\{V_k(z)\}$ [12]. Based on this rationale, an *approximate model* of the dynamical system $F_0(z)$ can be represented by a finite length N series expansion

$$F(z,\theta) = \sum_{k=0}^{N-1} L_k V_k(z),\ \theta = [L_0^T, \cdots, L_{N-1}^T]. \tag{3}$$

When the basis function $\{V_k(z)\}$ are chosen as $V_k(z) = z^{-k}$, then (3) simplifies to the conventional Finite Impulse Response (FIR) filter.

Generalizing the notion of a conventional FIR filter, the finite length N series expansion in (3) is dubbed as a generalized FIR filter in the remainder of the paper. The orthornomal basis sequence $\{V_k(z)\}$ can incorporate the possible prior knowledge of the system to be approximated, and the model $F(z,\theta)$ can be more accurate for a smaller finite number of coefficients N compared to a conventional FIR model structure. Obviously, the accuracy of the model $F(z,\theta)$ depends on the choice of the basis $\{V_k(z)\}$ used in (3).

A unifying construction of orthonormal basis functions $V_k(z)$ has been presented in [9] and given by

$$V_k(z) = \left(\frac{\sqrt{1-|\xi_k|^2}}{z-\xi_k}\right) \prod_{i=0}^{k-1} \left(\frac{1-\bar{\xi}_i z}{z-\xi_i}\right) \tag{4}$$

where $\{\xi_i\}_{i=0,1,2,\cdots,N-1}$ is a set of chosen pole locations for the basis $\{V_k(z)\}$. The advantage of using the basis functions $V_k(z)$ in (4) lies in the possibility to include knowledge

Raymond de Callafon is with the Mechanical and Aerospace Engineering, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0411, U.S.A. callafon@ucsd.edu

Jie Zeng is with Zona Technology Inc., 9489 E. Ironwood Square Drive, Suite 100, Scottsdale AZ 85258, U.S.A. jzeng@zonatech.com

of multiple pole locations in the parametrization of the generalized FIR filter $F(z,\theta)$. While the orthonormality of basis $\{V_k(z)\}$ is preserved, the multiple pole locations in $V_k(z)$ allow approximation of $F_0(z)$ using the generalized FIR filter structure $F(z,\theta)$ depicted in Fig. 1.

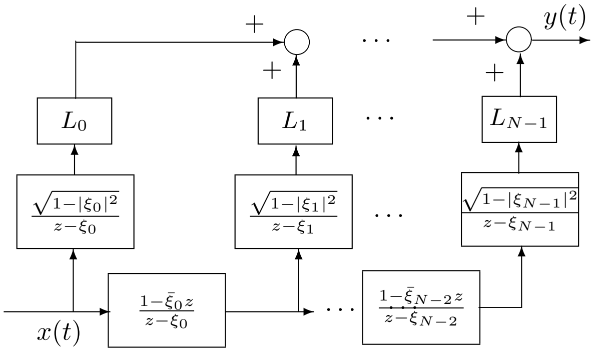

Fig. 1. Illustration of the filter $y(t) = F(g,\theta)x(t)$, where $F(q,\theta)$ is parametrized via unified construction of the basis $\{V_k(z)\}$ in (4).

The set of (generalized) orthonormal basis functions initially published in [1] provides an alternative way to construct an orthonormal basis with all-pass functions. For more details on the construction of the generalized basis functions $V_k(z)$ one is referred to [3]. The following result shows the existence and construction of the inner function which is crucial to create the orthonormal basis $\{V_k(z)\}$.

Proposition 1: Let (A,B) be the state matrix and input matrix of an input balanced realization of a discrete time transfer function $H \in \mathcal{RH}_2^{p\times m}$ ($\mathcal{RH}_2^{p\times m}$ indicates the set of real rational $p\times m$ matrix functions) with a McMillan degree $n > 0$f and with $rank(B) = m$. Then

(a) There exist matrices C, D such that (A,B,C,D) is a minimal balanced realization of a square inner function P.

(b) A realization (A,B,C,D) has the property mentioned in (a) if and only if

$$\begin{array}{rcl} C &=& UB^T(I_n+A^T)^{-1}(I_n+A) \\ D &=& U[B^T(I_n+A^T)^{-1}B - I_m] \end{array} \tag{5}$$

where $U \in \mathbb{R}^{m\times m}$ is any unitary matrix, and where I_n is $n\times n$ identity matrix, and I_m is $m\times m$ identity matrix.

Proof: For the proof, one is referred to [3]. ■

Proposition 1 yields a square $m\times m$ inner transfer function $P(z) = D + C(zI-A)^{-1}B$, where (A,B,C,D) is a minimal balanced realization. With the information obtained in Proposition 1, the orthonormal basis functions can be created with following proposition.

Proposition 2: Let $P(z)$ is a square inner function with McMillan degree $n > 0$ and (A,B,C,D) is a minimal balanced realization of $P(z)$. Define the input to state transfer function $V_0(z) := (zI-A)^{-1}B$ and

$$\begin{array}{rcl} V_k(z) &=& (zI-A)^{-1}BP^k(z) \\ &=& V_0(z)P^k(z) \end{array} \tag{6}$$

then the set of functions $\{V_k(z)\}_{k=0,1,2,\ldots}$ are orthonormal basis functions which have the following property

$$\frac{1}{2\pi j}\oint V_i(z)V_j^T(1/z)\frac{\mathrm{d}z}{z} = \begin{cases} I & i=j \\ 0 & i\neq j \end{cases} \tag{7}$$

Proof: For the proof, one is referred to [3]. ■

Proposition 1 and Proposition 2 show how to use an inner function to construct the orthonormal basis function $V_k(z)$. If an orthonormal basis $\{V_k(z)\}$ with poles at ξ_i, $i = 0,1,2,\cdots,N-1$ is desired, then from Proposition 1 an inner function $P(z)$ with these poles can be created. As a result, a balanced realization (A,B,C,D) of inner function $P(z)$ can be found to form the orthonormal basis $\{V_k(z)\}$ as in (6). In comparison with the filter structure depicted in Fig. 1, the filter $F(q,\theta)$ based on the basis functions $V_k(z)$ of (6) is given in Fig. 2.

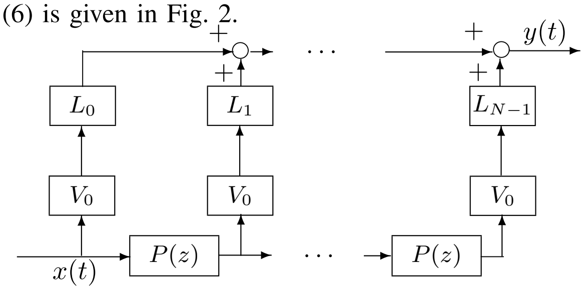

Fig. 2. Filering $y(t) = F(q,\theta)x(t)$, where $F(q,\theta)$ in (3) is parametrized by the basis functions $V_k(z)$ in (6).

The structure in Fig. 1 is equivalent to Fig. 2 if the set of n poles are repeated indefinitely. By the choice of basis functions $V_z(z)$ in (6), the poles of $F(q,\theta)$ will be restricted to a finite (multiple) set $\{\xi_0,\cdots,\xi_{n-1}\}$. The unified construction of basis functions $V_k(z)$ in (4) allows the set of poles of $F(q,\theta)$ to be infinitely large. A more general orthonormal basis $\{V_k(z)\}$ similar as in (6) can incorporate an infinite number of pole locations:

Proposition 3: Consider a sequence of inner function $P_i(z)$, $i = 0,1,2,\cdots$, each $P_i(z)$ having a corresponding balanced realization (A_i,B_i,C_i,D_i) and defining $\Phi_i(z) = (zI-A_i)^{-1}B_i$. Then the set of functions $\{V_i(z)\}_{i=0,1,2,\cdots}$ with

$$\begin{array}{c} V_0(z) = \Phi_0(z), \\ V_i(z) = \Phi_i(z)P_0(z)P_1(z)\cdots P_{i-1}(z) \end{array} \tag{8}$$

is mutually orthonormal.

Proof: The proof is similar to the proof for Proposition 2. ■

Because the set of basis functions $V_i(z)$, $i = 0,1,2,\cdots$ are mutually orthonormal, $\{V_i(z)\}$ constitues an orthonormal basis. To distinguish the construction of this set of basis functions from the unified construction defined in (4) and the general basis functions defined in (6), the basis functions $V_i(q)$ given in Proposition 3 are named (generalized) mutual orthonormal basis functions throughout the remainder of the paper. With $V_i(z)$ in place, the approximation $F(z,\theta)$ of a dynamical system $F_0(z)$ can be represented as

$$F(z,\theta) = \sum_{k=0}^{N-1} L_k\Phi_k(z)P_0(z)P_1(z)\cdots P_{k-1}(z) \tag{9}$$

If $P(z) = P_0(z) = P_1(z) = \cdots = P_{N-2}(z)$, then (8) can be simplified to (6) and therefore the construction of the basis functions in (8) is the generalization of the construction of basis functions in (6).

III. MODEL MATCHING WITH GENERALIZED FIR FILTERS

Given the parametrization of a filter $F(z,\theta)$ based on generalized mutual orthonormal basis functions, the model matching problem defined earlier in (1) will be solved. The model matching in (1) occurs in many applications that require the computation of a stable filter to approximate an unstable system and arises for example in feedforward algorithms for Active Noise Control [10], [11].

For the model matching problem (1), let the dynamical systems $H_0(z)$ and $G_0(z)$ be given by the state space realizations (A_h, B_h, C_h, D_h) and (A_g, B_g, C_g, D_g), where $H_0(z)$ is stable and stable invertible, and $G_0(z)$ is stable but not stably invertible. As a result, the optimal filter $F(z,\theta)$ can not be computed via a trivial solution $H_0(z)G_0^{-1}(z)$, as this would results in an unstable and/or non-causal filter. The objective is to find a stable filter $F(z,\theta)$ such that

$$\|H_0(z) - F(z,\theta)G_0(z)\|_2 \tag{10}$$

is minimized, where $F(z,\theta)$ is parametrized via a linear combination of basis functions $V_k(z)$ given in (8). It is straightforward to show that the minimization of (10) is equivalent to minimizing

$$\|H(z,\theta)\|_2,\ H(z,\theta) := H_0(z) - F(z,\theta)G_0(z)$$

where $H(z,\theta)$ is parametrized by a state-space representation $(A, B, C(\theta), D(\theta))$ in which the output matrix $C(\theta)$ and the feedthrough matrix $D(\theta)$ depends affinely on the parameter θ. To formulate an analytic solution to (1), first an expression for the state space realization $(A, B, C(\theta), D(\theta))$ is derived in case $F(z,\theta)$ is parametrized using generalized mutual orthonormal basis functions.

Proposition 4: Consider a square inner transfer function $P(z)$ with a minimal balanced realization (A_b, B_b, C_b, D_b) and a state dimension $n_b > 0$. Then for any $k > 1$ the realization (A_k, B_k, C_k, D_k) of $P(z)^k$ can be computed with the recursive formulas

$$\begin{aligned} A_k &= \begin{bmatrix} A_{k-1} & 0 \\ B_bC_{k-1} & A_b \end{bmatrix} \\ B_k &= \begin{bmatrix} B_{k-1} \\ B_bD_{k-1} \end{bmatrix} \\ C_k &= \begin{bmatrix} D_{k-1}C_b & C_{k-1} \end{bmatrix} \\ D_k &= D_bD_{k-1} \end{aligned} \tag{11}$$

and is a minimal balanced realization of $P(z)^k$ with state dimension $n \cdot k$.

Proof: For the proof, one is referred to [1]. ■

Proposition 5: Given a filter $F(z,\theta)$ parametrized in the form of

$$F(z,\theta) = D_f + \sum_{k=1}^{n} L_{k-1}V_{k-1}(z),\ \theta = [D_f, L_0^T, \cdots, L_{n-1}^T]$$

where $V_{k-1}(z)$ are obtained from (6). Let the inner function $P(z)^{n-1}$ has a minimal balanced realization $(A_{n-1}, B_{n-1}, C_{n-1}, D_{n-1})$, then $(A_{n-1}, B_{n-1}, C_f, D_f)$ is a realization of $F(z,\theta)$, where $C_f = [L_0^T, L_1^T, \cdots, L_{n-1}^T]$.

Proof: Follows from standard linear algebra. ■

It can be noted here that the state space realization of $F(z,\theta)$ can be found via a similar procedure as described in Proposition 4 and Proposition 5, even if the basis functions $V_k(z)$ are constructed via (8). With the state space realization of $H_0(z)$, $G_0(z)$ and $F(z,\theta)$, the state space realization $(A, B, C(\theta), D(\theta))$ of $H_0(z) - F(z,\theta)G_0(z)$ is given by

$$A = \begin{bmatrix} A_h & 0 & 0 \\ 0 & A_g & 0 \\ 0 & B_{n-1}C_g & A_{n-1} \end{bmatrix}\ B = \begin{bmatrix} B_h \\ B_g \\ B_{n-1}D_g \end{bmatrix} \tag{12}$$
$$C(\theta) = \begin{bmatrix} C_h & -D_fC_g & -C_f \end{bmatrix}\ D(\theta) = [D_fD_g]$$

With these results, the minimization of $\|H(z,\theta)\|_2$ can be computed as an affine function of $\theta = [D_f, L_0^T, \cdots, L_{n-1}^T]$ and the result is summarized in the following theorem.

Theorem 1: Consider the discrete-time system $H(z,\theta)$, Minimization of $\|H(z,\theta)\|_2^2$ is equivalent to

$$\min_{\theta}\{\mathrm{tr}\left[D(\theta)D(\theta)^T + C(\theta)QC(\theta)^T\right]\} \tag{13}$$

where Q is the solution to the Lyapunov (Stein) equation

$$AQA^T - Q + BB^T = 0$$

The optimization in (13) can be solved via a SemiDefinite Programming (SDP) problem

$$\begin{gathered} \min_{\gamma, X_1, X_2, \theta} \gamma \text{ such that} \\ \gamma - \mathrm{tr}\{X_1\} - \mathrm{tr}\{X_2\} > 0,\ \begin{bmatrix} X_1 & C(\theta) \\ C(\theta)^T & Q^{-1} \end{bmatrix} > 0, \\ \begin{bmatrix} X_2 & D(\theta) \\ D(\theta)^T & I \end{bmatrix} > 0 \end{gathered} \tag{14}$$

where γ is a positive real number.

Proof: Since $C(\theta)$ and $D(\theta)$ depend affinely on the parameter θ, the condition

$$\mathrm{tr}\{D(\theta)D(\theta)^T\} + \mathrm{tr}\{C(\theta)QC(\theta)^T\} \leq \gamma$$

can be recasted as a Linear Matrix Inequality (LMI). The LMI will be of the form

$$\begin{gathered} \mathrm{tr}\{X_1\} + \mathrm{tr}\{X_2\} < \gamma,\ \begin{bmatrix} X_1 & C(\theta) \\ C(\theta)^T & Q^{-1} \end{bmatrix} > 0, \\ \begin{bmatrix} X_2 & D(\theta) \\ D(\theta)^T & I \end{bmatrix} > 0 \end{gathered}$$

where $X_1 = X_1^T$ and $X_2 = X_2^T$ are two new (slack) variables of appropriate dimension needed to formulate the convex constraints. As a result, the SDP problem

$$\begin{gathered} \min_{\gamma, X_1, X_2, \theta} \gamma \text{ such that} \\ \gamma - \mathrm{tr}\{X_1\} - \mathrm{tr}\{X_2\} > 0,\ \begin{bmatrix} X_1 & C(\theta) \\ C(\theta)^T & Q^{-1} \end{bmatrix} > 0, \\ \begin{bmatrix} X_2 & D(\theta) \\ D(\theta)^T & I \end{bmatrix} > 0 \end{gathered}$$

can be used to find the value of the parameter θ that minimizes $\|H(z,\theta)\|_2 = \sqrt{\gamma}$. ∎

A more efficient algorithm to compute θ is by exploiting the structure of $C(\theta)$ and $D(\theta)$. Since $H(z,\theta)$ is defined as $H_0(z) - F(z,\theta)G_0(z)$, it follows that $C(\theta) = [C_h \; -D_f C_g \; -C_f]$ and $D(\theta) = D_h - D_f D_g$ where the pairs (C_h, D_h), (C_g, D_g) and (C_f, D_f) are the output matrices of $H_0(z)$, $G_0(z)$ and $F(z,\theta)$ respectively. Defining the parameter θ as $\theta = [D_f \;\; C_f]$, allows the minimization of $\|H(z,\theta)\|_2$ to be written as a weighted Least Squares problem. The result is summarized in the following corollary.

Corollary 1: The analytic solution for the minimization

$$\min_\theta \{\mathrm{tr}\left[D(\theta)D(\theta)^T + C(\theta)QC(\theta)^T\right]\} \tag{15}$$

is given by

$$\theta = [YX^T][XX^T]^{-1} \tag{16}$$

where

$$Y = \begin{bmatrix} D_h & C_h & 0 & 0 \end{bmatrix} H, \; X = \begin{bmatrix} D & C \end{bmatrix} H$$

$$C = \begin{bmatrix} 0 & C_g & 0 \\ 0 & 0 & I \end{bmatrix}, \; D = \begin{bmatrix} D_g \\ 0 \end{bmatrix}$$

and H is found by a Cholesky factorization of the positive definite matrix $\bar{Q} = \begin{bmatrix} I & 0 \\ 0 & Q \end{bmatrix}$.

Proof: Matrices $C(\theta)$ and $D(\theta)$ in (12) can be rewritten as

$$C(\theta) = \begin{bmatrix} C_h & 0 & 0 \end{bmatrix} - \theta C, \; D(\theta) = D_h - \theta D,$$
$$C := \begin{bmatrix} 0 & C_g & 0 \\ 0 & 0 & I \end{bmatrix}, \; D := \begin{bmatrix} D_g \\ 0 \end{bmatrix}$$

Define

$$E = \begin{bmatrix} D_h & C_h & 0 & 0 \end{bmatrix} - \theta \begin{bmatrix} D & C \end{bmatrix}, \; \bar{Q} = \begin{bmatrix} I & 0 \\ 0 & Q \end{bmatrix}$$

then (15) can be rewritten as

$$\|H(z,\theta)\|_2^2 = \mathrm{tr}\{E\bar{Q}E\}$$

where H is found by a Cholesky factorization $\bar{Q} = HH^T$ of the positive definite matrix $\bar{Q}$. With the definition of

$$Y = \begin{bmatrix} D_h & C_h & 0 & 0 \end{bmatrix} H, \; X = \begin{bmatrix} D & C \end{bmatrix} H$$

E is defined as $E = Y - \theta X$ and the minimization of $\mathrm{tr}\{E\bar{Q}E\}$ for $\bar{Q} > 0$ is a standard weighted least squares optimization problem for which the analytic solution can be computed via (16). ∎

IV. Model error bounds

With full knowledge of the dynamics of $H_0(z)$ and $G_0(z)$, the modeling matching problem in (1) can be seen as an H_2-optimal control or filtering problem. In case no restrictions on the parametrization of $F(z,\theta)$ are imposed, the minimization $\|H_0(z) - F(z,\theta)G_0(z)\|_2$ can be solved with standard H_2 optimal control solutions [13]. Including a requirement on the parametrization of $F(z,\theta)$ in the form of a linear combination of orthonormal basis functions only enables the model $F(z,\theta)$ to approach the optimal solution $F_{opt}(z)$. However, the linear parametrization can be exploited in recursive computational tools to adjust for changes in the noise filter $H_0(z)$ as seen for example in Active Noise Control problems.

To formulate error bounds for the model error between $F(z,\theta)$ and $F_{opt}(z)$, the model error is quantified by the 2-norm

$$\|H_0(z) - F(z,\theta)G_0(z)\|_2 \tag{17}$$

and the optimal solution $F_{opt}(z)$ gives a lower bound $\|H_0(z) - F_{opt}(z)G_0(z)\|_2$ for the model error in (17). An upper bound for (17) can be formulated by considering the poles locations ξ_i of the basis functions $V_k(q)$ used for the affine parametrization of $F(z,\theta)$. The result is summarized in the following proposition.

Proposition 6: Let $\{\kappa_i\}$ be the set of stable poles of $H_0(z)$ and stable zeros of $G_0(z)$, let $\{\sigma_i\}$ be the set of unstable zeros of $G_0(z)$ and let the all-pass transfer function $P(z)$ used for the construction of the orthonormal basis functions have poles ρ_j, $j = 1, \cdots, n_p$. Define

$$\lambda := \max_i \prod_{j=1}^{n_p} \left| \frac{\nu_i - \rho_j}{1 - \nu_i \rho_j} \right|, \; \nu_i = \kappa_i \cup \sigma_i^{-1}$$

and denote

$$F(z,\theta) = D_f + \sum_{k=1}^{n} L_{k-1} V_{k-1}(z), \; \theta = [L_0^T, \cdots, L_{n-1}^T].$$

Then there exists a finite constant $c \in \mathbf{R}$ and any $\eta \in \mathbf{R}$, $\lambda < \eta < 1$ such that

$$\|H_0(z) - F(z,\theta)G_0(z)\|_2 \leq c \frac{\eta^{n+1}}{\sqrt{1-\eta^2}}$$

Proof: Let $F_{opt}(z)$ be the solution to (17), where $F_{opt}(z)$ is freely parametrized, e.g. F_{opt} is not restricted to be an affine orthonormal FIR. Then it is straightforward to show that (17) is an LQG control problem that minimizes

$$\sum_{t=1}^{\infty} z^T(t)z(t), z(t) = H_0(q)d(t) + u(t),$$
$$u(t) = F_{opt}(q)G_0(q)d(t)$$

in which only the variance of the control signal $u(t)$ is being penalized. For such a minimum variance controller it was shown in [14] the zeros of $G_0(q)$ are mapped into the unit circle in order to obtain a stable minimum variance controller. As a result, the poles of $F_{opt}(z)$ will include all stable poles of $H_0(z)$, all stable zeros of $G_0(z)$, and all unstable zeros of $G_0(z)$ which are mapped into the unit circle, e.g. $\nu_i = \kappa_i \cup \sigma_i^{-1}$. Therefore, $\min_\theta \|H_0(z) - F(z,\theta)G_0(z)\|_2$ is equivalent to $\min_\theta \|F_{opt}(z) - F(z,\theta)\|_2$. For computation of the upper bound of model error $\|F_{opt} - F(z,\theta)\|_2$, one is referred to the work of [15]. ∎

The above proposition shows that if the poles $\{\rho_j\}$ of $P(z)$ approach the poles $\{\nu_i\}$, then $\lambda = 0$ and the upper bound of model error $\|H_0(z) - F(z,\theta)G_0(z)\|_2$ will decrease drastically. Therefore, from Proposition 6 it can be observed that an appropriate selection of the poles of the all-pass function will have an important contribution to the reduction of the model error $\|H_0(z) - F(z,\theta)G_0(z)\|_2$. This is due to the improvement in the convergence rate of the coefficient L_{k-1} for $k = 1, \ldots, n$ in the affine parametrization of the filter $F(q,\theta)$.

V. CASE STUDY

To illustrate the approximation results using the minimization (13) for model matching, consider a fourth order system $H_0(z)$ and a second order system $G_0(z)$ given by

$$\begin{array}{c} H_0(z) = \dfrac{z^2 - 1.2z + 0.4}{z^2 - 1.8z + 0.8} \cdot \dfrac{z^2 - z + 0.5}{z^2 - z + 0.89}, \\ \text{with poles at } 0.9 \pm 0.3i,\ 0.5 \pm 0.8i \\ G_0(z) = \dfrac{z^2 - 1z + 1.94}{z^2 - z + 0.74}, \\ \text{with poles at } 0.5 \pm 0.7i, \text{zeros at } 0.5 \pm 1.3i \end{array} \quad (18)$$

The objective is to find a filter $F(q,\theta)$ parametrized via a orthonormal basis function expansion

$$\begin{array}{c} F(z,\theta) = D_f + \sum_{k=1}^{n} L_{k-1} V_{k-1}(z), \\ \theta = [L_0^T, \cdots, L_{n-1}^T] \end{array} \quad (19)$$

with a limited number n of parameters, such that

$$\|H_0(z) - F(z,\theta)G_0(z)\|_2 \quad (20)$$

is being minimized. Using H_2 optimal control design technique, an optimal $F_{opt}(z)$ can be obtained as

$$F_{opt}(z) = \frac{b_5 z^5 + b_4 z^4 + b_3 z^3 + b_2 z^2 + b_1 z + b_0}{z^6 + a_5 z^5 + a_4 z^4 + a_3 z^3 + a_2 z^2 + a_1 z + a_0}$$

where $b_0 = -0.08646$, $b_1 = 0.2046$, $b_2 = -0.2522$, $b_3 = 0.1443$, $b_4 = -0.0267$, $b_5 = 0.004005$. $a_0 = 0.4129$, $a_1 = -1.7032$, $a_2 = 3.941$, $a_3 = -5.796$, $a_4 = 5.549$, $a_5 = -3.315$. The poles of $F_{opt}(z)$ are $z_1, \bar{z}_1 = 0.9 \pm 0.3i$, $z_2, \bar{z}_2 = 0.5 \pm 0.8i$ and $z_3, \bar{z}_3 = 0.2577 \pm 0.6701i$. With $F_{opt}(z)$ in place, the lower bound of model error can be computed as $\|H_0(z) - F_{opt}(z)G_0(z)\|_2 = 1.2243$. Since $F_{opt}(q)$ is computed with a feed-through matrix $D_f = 0$, the feed-through term D_f in (19) also be set to 0 for comparison purposes.

For the sake of illustration in this case study, it is assumed only the first two (conjugate) poles $\{z_1, \bar{z}_1, z_2, \bar{z}_2\}$ of the model $F_{opt}(z)$ are known for the construction of basis functions $V_k(z)$. Actually, these poles are also the poles of $H_0(z)$. Obviously, if all pole locations of $F_{opt}(z)$ were known beforehand, n can be set to $n = 1$ as only one unique coefficient L_0 will be able to represent $F_{opt}(z)$. However, in this case incomplete knowledge of the pole locations of $F_{opt}(z)$ is assumed, requiring theoretically an infinite number of parameters for an accurate representation of $F_{opt}(z)$, increasing the McMillan degree of $F(z)$. To make a fair comparison between the usage of different basis functions, n is limited such that $F_{opt}(z)$ will have a McMillan degree that is less than or equal to 20.

The quality of the approximation measured by (20) is compared for different sets of basis functions. The first set of basis functions $V_k(z)$ are constructed via Proposition 2 and are based on a single all-pass function $P(z)$ that uses the knowledge of the two (conjugate) poles $z_1, \bar{z}_1$ and $z_2, \bar{z}_2$ of $F_{opt}(z)$. Since $P(z)$ is a fourth order all-pass function, the parametrization of the first filter $F_1(z)$ is given by

$$F_1(z) = \sum_{k=1}^{5} L_{k-1} V_{k-1}(z),\ V_{k-1}(z) = \Phi_0(z) P(z)^{k-1} \quad (21)$$

where $\Phi_0(z) = (zI - A)^{-1}B$ in which (A, B) are computed from an input balanced state-space realization of $P(z)$. With the fourth order all-pass function $P(z)$, it can be seen from (21) that n has been limited to $n = 5$ to ensure that $F(z)$ has a McMillan degree that is less than or equal to 20.

Instead of a single basis function $V_k(z)$, an alternative approach would be to create mutually orthonormal basis functions on the basis of the two all-pass functions $P_1(z)$ and $P_2(z)$ that separate the knowledge of the (conjugate) poles location at $z_1, \bar{z}_1$ and $z_2, \bar{z}_2$. On the basis of the parametrization given in Proposition 3, the following filters $F_m(z)$ are considered:

$$\begin{array}{l} F_m(z) = \displaystyle\sum_{k=1}^{10} L_{k-1} V_{k-1}(z) \text{ where} \\ V_{k-1}(z) = \begin{cases} \Phi_1(z) P_1(z)^{k-1},\ k = 1, \ldots, m \\ \Phi_2(z) P_1(z)^m P_2(z)^{k-m-1},\ k = 1+m, \ldots, 10 \end{cases} \end{array} \quad (22)$$

where $\Phi_1(z) = (zI - A_1)^{-1}B_1$ in which (A_1, B_1) are computed from an input balanced state-space realization of $P_1(z)$ and $\Phi_2(z) = (zI - A_2)^{-1}B_2$ in which (A_2, B_2) are computed from an input balanced state-space realization of $P_2(z)$.

In case the orthonomal basis functions are simply set to $V_k(z) = z^{-k}$ to obtain a 20th order FIR model $F_{fir}(q,\theta)$, then the model error becomes $\|H_0(z) - F_{fir}G_0(z)\|_2 = 1.2586$. With the 4th order all-pass function $P(z)$ and the construction of the orthonormal basis functions in (21), the computations of 5 coefficients L_i for the 20th order filter $F_1(z)$ reduces the model error to $\|H_0(z) - F_1(z)G_0(z)\|_2 = 1.2257$. This illustrates that a generalized FIR filter can provide much better approximation results than a conventional FIR filter.

Different combinations m of basis functions in the mutual orthonormal basis functions in (22) to construct $F_m(z,\theta)$ will give different model error results. As a final comparison for this case study, the modeling error of $\|H_0(z) - F_m(z,\theta)G_0(z)\|_2$ is calculated and shown in Fig. 3. From Fig. 3, the following observations can be made. Firstly, if only the 2nd order $P_1(z)$, $m = 10$ or $P_2(z)$, $m = 0$ all-pass functions are used to create orthonormal basis functions $V_k(z)$, the approximation result is worse compared to choosing a 4th order basis function $P(z)$ or any linear combination of $P_1(z)$ and $P_2(z)$ as all-pass functions. Hence, higher order basis functions $V_k(z)$ that include more poles of the dynamic system to be approximated is preferable to reach an improvement in model approximation.

Secondly, the smallest model error is obtained when $m = 1$. This implies that the quality of the approximation is not only related to the location of the poles of the basis function,

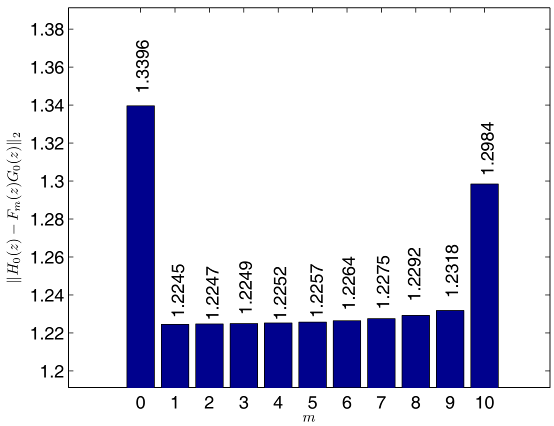

Fig. 3. Comparison of model error for 20th order filter $F_m(q, \theta)$ with different combinations m of mutual orthonormal basis functions in (22).

but is also determined by the number of coefficients used for building the series expansion on the basis of a specific basis function. In this case study, only two poles of F_{opt} are used to create an orthonormal basis. A possible explanation for the approximation results lies in the location of the poles, as indicated by Proposition 6. Since the poles $z_3, \bar{z}_3$ are closer to $z_2, \bar{z}_2$ than to $z_1, \bar{z}_1$ less coefficients $(m = 1)$ are needed to obtain a better approximation.

VI. H_∞ NORM MODEL MATCHING PROBLEM

Next to the H_2-norm based model matching, the linear parametrization of the filter $F(q, \theta)$ in terms of basis functions can also be used to minimize an H_∞-norm based model matching. Since $F(q, \theta) \in \mathbb{R}H_2$ implies $F(q, \theta) \in \mathbb{R}H_\infty$ for a discrete-time filter, a stable filter $F(z, \theta) \in \mathbb{R}H_\infty$ can be parametrized via a linear combination of basis functions $V_k(z)$ given in (8), provided the basis functions $V_k(z)$ have no poles on the unity circle. Given the state space realization $(A, B, C(\theta), D(\theta))$ of $H_0(z) - F(z, \theta)G_0(z)$ in (12), the minimization of $\|H_0(z) - F(z, \theta)G_0(z)\|_\infty$ can be computed by using Linear Matrix Inequalities (LMIs), which is given in the following proposition.

Proposition 7: Given the system matrix $\begin{bmatrix} A & B \\ C(\theta) & D(\theta) \end{bmatrix}$ of a discrete time system $\bar{H}(z, \theta) := H_0(z) - F(z, \theta)G_0(z)$, $\|\bar{H}(z, \theta)\|_\infty < \gamma$ is equivalent to the existence of a positive definite symmetric matrix $P > 0$, such that

$$\begin{bmatrix} A^T PA - P & A^T PB & C(\theta)^T \\ B^T PA & B^T PB - \gamma^2 I & D(\theta)^T \\ C(\theta) & D(\theta) & I \end{bmatrix} < 0 \qquad (23)$$

where $C(\theta)$ and $D(\theta)$ can be written as

$$C(\theta) = \begin{bmatrix} C_h & 0 & 0 \end{bmatrix} - \theta C, \; D(\theta) = D_h - \theta D,$$
$$C := \begin{bmatrix} 0 & C_g & 0 \\ 0 & 0 & I \end{bmatrix}, \; D := \begin{bmatrix} D_g \\ 0 \end{bmatrix}, \; \theta = \begin{bmatrix} D_f & C_f \end{bmatrix}$$

Proof: The bounded-real Lemma states equivalence between $\|\bar{H}(q, \theta)\|_\infty < \gamma$ and the existence of a positive definite matrix P such that

$$\begin{aligned} A^T PA - P + C^T C - (A^T PB + C^T D) \\ (B^T PB + D^T D - \gamma^2 I)^{-1}(B^T PA + D^T C) < 0 \end{aligned} \qquad (24)$$

Via Schur complement on (24) one obtains (23). ■

VII. CONCLUSIONS

In this paper an analytic solution for both H_2- and H_∞-norm based model matching problem is formulated on the basis of an affine model structure parametrized by generalized orthonormal basis functions. The analytic solution is formulated in terms of Semidefinite Programming problem and the solution to model matching is typically found in problems associated to feedforward active noise control.

A model error bound for the model approximation is formulated and using the analytic solution, different orthonormal basis functions for the construction of generalized FIR filter are compared in a case study. The results show that during the construction of the orthornomal basis functions, a high order orthonormal basis function with a small number of coefficients is preferred over a low order orthonormal basis functions with a larger number of parameters.

REFERENCES

[1] P. S. C. Heuberger, P. M. J. Van Den Hof, and O. H. Bosgra, “A generalized orthonormal basis for linear dynamical systems,” *IEEE Transactions on Automatic Control*, vol. 40(3), pp. 451–465, 1995.

[2] B. Ninness, H. Hjalmarsson, and F. Gustafsson, “The fundamental role of generalized orthonormal basis in system identification,” *IEEE Trans. on Automatic Control*, vol. 44, no. 7, pp. 1384–1406, 1999.

[3] P. Heuberger, P. Van Den Hof, and B. Wahlberg, *Modelling and Identification with Rational Orthogonal Basis Functions*. Springer Verlag, 2005.

[4] D. K. de Vries and P. M. J. Van Den Hof, “Frequency domain identification with generalized orthonormal basis functions,” *IEEE Transactions on Automatic Control*, vol. 43(5), pp. 656–669, 1998.

[5] Z. Szabó, P. S. C. Heuberger, J. Bokor, and P. M. J. Van Den Hof, “Extended ho-kalman algorithm for systems represented in generalized orthonormal bases,” *Automatica*, vol. 36, pp. 1809–1818, 2000.

[6] T. J. de Hoog, Z. Szabó, P. S. C. Heuberger, P. M. J. Van Den Hof, and J. Bokor, “Minimal partial realization from generalized orthonormal basis function expansions,” *Automatica*, vol. 38, pp. 655–669, 2002.

[7] B. Wahlberg, “System identification using laguerre models,” *IEEE Transactions on Automatic Control*, vol. AC-36, pp. 551–562, 1991.

[8] ——, “System identification using kautz models,” *IEEE Transactions on Automatic Control*, vol. AC-39, pp. 1276–1282, 1994.

[9] B. Ninness and F. Gustafsson, “A unifying construction of orthonormal bases for system identification,” *IEEE Trans. on Automatic Control*, vol. 42, no. 4, pp. 515–521, 1997, an extended version is available as Technical report EE9433, Department of Electrical and Computer Engineering, University of Newcastle, Australia, 1994.

[10] J. Zeng and R. A. de Callafon, “Feedforward noise cancellation in an airduct using generalized FIR filter estimation,” in *Proc. 42nd IEEE Conference on Decision and Control*, Maui, Hawaii, USA, 2003, pp. 6392–6397.

[11] J. Zeng and R. de Callafon, “Recursive filter estimation for feedforward noise cancellation with acoustic coupling,” *Journal of Sound and Vibration*, vol. 291, pp. 1061–1079, 2006.

[12] P. S. C. Heuberger, T. J. De Hoog, P. M. J. Van Den Hof, and B. Wahlberg, “Orthonormal basis functions in time and frequency domain: Hambo transform theory,” *SIAM Journal on Control & Optimization*, vol. 42, no. 4, pp. 1347–1373, 2003.

[13] K. Zhou and J. C. Doyle, *Essentials of Robust Control*. Prentice-Hall, 1998.

[14] K. J. Åström and B. Wittenmark, “Computer-controlled system: Theory and design,” Englewood Cliffs, New Jersey, 1990.

[15] P. S. C. Heuberger, “On approximate system identification with system based orthonormal functions,” Ph.D. dissertation, Delft University of Technology, The Netherlands, 1991.