Mimo System Identification Using Orthonormal Basis Functions

Brett Ninness*
Centre for Industrial Control Science
Department of Electrical Engineering
University of Newcastle
Callaghan 2308, Australia.

Juan-Carlos Gómez
Centre for Industrial Control Science
Department of Electrical Engineering
University of Newcastle
Callaghan 2308, Australia.

Steven Weller
Department of Electrical and Electronic Engineering,
University of Melbourne,
Parkville, Australia.

Abstract

There has recently been interest in the use of orthonormal bases for the purposes of SISO system identification. Concurrently, but separately, there has also been vigorous work on estimation of MIMO systems by computationally cheap and reliable schemes. These latter ideas have collectively become known as '4SID' methods. This paper is a contribution overlapping these two schools of thought by showing how general orthonormal bases may be generated to form model structures suitable for identification of MIMO systems using only simple calculations. In contrast to general prediction error methods and in common with 4SID schemes the estimation algorithms involved are computationally simple. However, a distinguishing feature of the orthonormal basis approach described here is that significant prior knowledge about system poles may be included in the estimation problem.

1 Introduction

Recently there has been interest in developing schemes for estimation of SISO systems using so-called orthonormal basis functions. The most common model structure employing these functions is the well known FIR one. However, as explained in [10, 2], the use of such a model structure implicitly involves expanding the true discrete time system transfer function $G(q)$ around poles located at the origin. If the true poles are slow (near the unit circle) and hence very far from the origin then a more sensible approach is to expand $G(q)$ around a pole $q = \xi$ near the dominating slow mode. This amounts to the inclusion of prior knowledge of a single first order response, and may be achieved using the Laguerre orthonormal model structure ([10, 11]). The FIR structure is a special case of the Laguerre structure with $\xi = 0$.

When prior knowledge of a second order mode is to be included (ξ is complex) then a Kautz orthonormal model structure can be used [11]. The Laguerre structure is a special case of the Kautz one when ξ is real.

If prior knowledge is to be included about several modes of different types, then the general orthonormal construction methods of Heuberger, Van den Hof and co-workers [2] or of Ninness and Gustafsson [7, 5] may be employed. The Kautz, Laguerre and FIR structures are all special cases of these methods.

There are several advantages to using these orthonormal basis model structures in system identification. Firstly, they are structures that are linear in the parameters to be estimated. This means that system estimates may be found in a computationally cheap and reliable closed form fashion.

As well, and in contrast to well known prediction error methods employing ARX model structures, no iterative numerical optimisation schemes are necessary to find the estimate so that the attendant problems of convergence to local minima are avoided.

A further convenient feature of the linear structure is that finite data statistical properties of the estimate (such as variances and confidence regions) may be calculated.

A key advantage of orthonormality is that it facilitates theoretical analysis that yields guidelines for the practical implementation of the schemes. For the SISO case, this has been done for FIR structures in [4], for Laguerre in [10], for Kautz in [11] and for more general orthonormal bases in [2, 7, 5].

The final advantage of orthonormal structures is that it is possible to incorporate prior knowledge of the pole locations of $G(q)$ in the estimation problem. Again, this is in contrast to the case of using the popular ARX model structure where even if the user knows the approximate characteristics (time constants, damping ratios, natural frequencies) of $G(q)$, the information cannot be included in the identification process, but must be estimated in a black-box fashion. The utility of including this prior information is that the series expansion of $G(q)$ is then done about poles near the true ones, so that the convergence of the expansion is accelerated and hence as few parameters as possible need to be estimated.

The contribution of this paper is to show how these ideas of using orthonormal bases in SISO system identification can be extended to the multiple input, multiple output (MIMO) problem. Except for the well known

*This author gratefully acknowledges support from the Australian Research Council. The author can be contacted at email:brett@tesla.newcastle.edu.au or at FAX +61 49 21 69 93.

studied to date except for the work in [2] where the authors show how MIMO orthonormal bases can be generated, but do not investigate the use of them in MIMO estimation.

The orthonormal bases employed here are not the ones used in [2], but instead are generated using the methods introduced in [7, 5]. Once these bases have been generated, there is more than one choice as to how the model may be parameterised by them. This paper examines two possible choices, detail how estimates may be found using this choices, and present some preliminary theoretical results pertaining to the quality of the resulting estimates.

A recently popular class of schemes for estimation of MIMO systems has been the new State Space Subspace System Identification (4SID) methods. The attraction with these schemes is the low computational cost involved, and the reliability of the numerical procedures employed (SVD). MIMO estimation using the orthonormal bases of this paper enjoy equivalent numerical and computational cost properties while offering advantages that 4SID methods do not. Namely, the capability to include prior knowledge and to easily calculate statistical properties of estimates. The paper concludes with a short simulation study that illustrates the utility of the estimation schemes proposed.

2 Problem Formulation

This paper addresses problems involving identification of MIMO systems from noise corrupted observations of their response from n inputs to m outputs. To be more specific, the paper assumes the availability of an N point data record of n input sequences $\{u_k^1\}, \{u_k^2\}, \ldots, \{u_k^n\}$, as well as m output sequences $\{y_k^1\}, \{y_k^2\}, \ldots, \{y_k^m\}$ that are related according to

$$y_k^i = \sum_{j=1}^{n} G_{ij}(q)u_k^j + \sum_{s=1}^{r} H_{is}(q)e_k^s, \quad ; i = 1, 2, \cdots, m. \tag{1}$$

The scalar transfer functions $\{G_{ij}(q)\}$ describe the (assumed stable) unknown system dynamics that are to be identified. The output measurements $\{y_k^1\}$, $\{y_k^2\}, \ldots,$ $\{y_k^m\}$ are corrupted by zero mean stationary white noise processes $\{e_k^1\}, \{e_k^2\}, \ldots, \{e_k^r\}$ with finite variances $\sigma^2_{e_k^s} = \mathrm{E}\left\{e_k^{s2}\right\}$, and coloured by stable filters $H_{is}(q)$. A notational simplification is possible by vectorising:

$$y_k \stackrel{\triangle}{=} (y_k^1, y_k^2, \ldots, y_k^m)^T, \quad u_k \stackrel{\triangle}{=} (u_k^1, u_k^2, \ldots, u_k^n)^T,$$

$$e_k \stackrel{\triangle}{=} (e_k^1, e_k^2, \ldots, e_k^r)^T$$

$$G(q) = \begin{bmatrix} G_{11}(q) & G_{12}(q) & \ldots & G_{1n}(q) \\ G_{21}(q) & G_{22}(q) & \ldots & G_{2n}(q) \\ \vdots & \vdots & \ldots & \vdots \\ G_{m1}(q) & G_{m2}(q) & \ldots & G_{mn}(q) \end{bmatrix},$$

$$H(q) = \begin{bmatrix} H_{11}(q) & H_{12}(q) & \ldots & H_{1r}(q) \\ H_{21}(q) & H_{22}(q) & \ldots & H_{2r}(q) \\ \vdots & \vdots & \ldots & \vdots \\ H_{m1}(q) & H_{m2}(q) & \ldots & H_{mr}(q) \end{bmatrix},$$

$$y_k = G(q)u_k + H(q)e_k. \tag{2}$$

The idea of MIMO system identification using orthonormal basis functions is to follow the example of the SISO case and represent the transfer function matrix $G(q)$ as a series expansion in the basis, and then to approximate this expansion with a finite number of terms. In this way it is possible to (approximately) identify the system by estimating a finite number of expansion coefficients.

A feature of using orthonormal bases to represent MIMO systems is that in contrast to the SISO case more than one model structure is possible. Furthermore, the construction of the orthonormal basis depends on the model structure.

3 MIMO Model Structures

The most obvious MIMO model structure to use comes from copying the SISO idea of making a truncated series expansion of $G(q)$ as follows:

$$G(q, \Theta) = \sum_{\ell=0}^{p-1} \theta_\ell^T \mathcal{B}_\ell(q). \tag{3}$$

In the SISO case, the $\{\mathcal{B}_\ell(q)\}$ are orthonormal SISO transfer functions. In the MIMO case this can still be true, or the $\{\mathcal{B}_\ell(q)\}$ can be $m \times n$ matrices of transfer functions.

The construction of the $\{\mathcal{B}_\ell(q)\}$ to obtain orthonormality will be detailed in following sections, as will a discussion of the validity of the p-th order truncation, but for now note that the expansion (3) leads to the convenient linear regressor form:

$$\begin{aligned} G(q, \Theta)u_k &= \underbrace{(\theta_0^T, \theta_1^T, \cdots, \theta_{p-1}^T)}_{\Theta^T} \underbrace{\begin{pmatrix} D_0(q) \\ D_1(q) \\ \vdots \\ D_{p-1}(q) \end{pmatrix} u_k}_{\phi_k} \\ &= \Theta^T \phi_k \end{aligned} \tag{4}$$

where if the bases $\{\mathcal{B}_k(q)\}$ are chosen to be matrices with $\dim\{\mathcal{B}_k(q)\} = m \times n$ then in the above equation $D_k(q) = \mathcal{B}_k(q)$ and $\dim\{\theta_k\} = m \times m$, $\dim\{\Theta\} = mp \times m$ and $\dim\{\phi_k\} = mp \times 1$. On the other hand, if the bases $\{\mathcal{B}_k(q)\}$ are chosen to be scalar, then $D_k(q) = I_n\mathcal{B}_k(q)$ and $\dim\{\theta_k\} = n \times m$, $\dim\{\Theta\} = np \times m$ and $\dim\{\phi_k\} = np \times 1$.

The choice of whether the bases $\{\mathcal{B}_k(q)\}$ are chosen as scalar or matrix depends on the flexibility required in specifying prior information about the individual transfer function entries $\{G_{ij}(q)\}$.

If knowledge exists about the pole locations in each individual $\{G_{ij}(q)\}$ term of $G(q)$, then the elements in matrix $\{\mathcal{B}_k(q)\}$ can be chosen to reflect this knowledge while still preserving orthonormality. This ability to specify poles in individual $\{G_{ij}(q)\}$ is particularly important if there is a wide distribution of time constants in the different $\{G_{ij}(q)\}$ elements.

case, or if less prior knowledge is available, then scalar $\{\mathcal{B}_k(q)\}$ can be used with poles chosen to approximate those in all the $\{G_{ij}(q)\}$ entries in $G(q)$.

In practice, the use of matrix bases gives more accurate estimation due to the greater flexibility of the model structure with (for square systems) the same number of parameters being estimated. Unfortunately, there is a penalty suffered for this in that the theoretical analysis is more complicated than for the scalar case.

4 Parameter Estimation

For the purposes of estimation, the N point data records can be most conveniently handled by adopting the vectorised notation

$$\begin{aligned} Y^T &= (y_0, y_1, \cdots, y_{N-1}) \\ \Phi^T &= (\phi_0, \phi_1, \cdots, \phi_{N-1}) \end{aligned}$$

and then using (4) the model for the N point observed data record can be written as:

$$Y = \Phi\Theta. \tag{5}$$

In this case, the most obvious and simplest scheme for estimating the parameter matrix Θ in the model structure (4) is the ubiquitous least squares method

$$\begin{aligned} \widehat{\Theta} &= \arg\min_{\Theta} \frac{1}{N}\mathrm{Tr}\left\{\sum_{k=0}^{N-1}\varepsilon_k(\Theta)\varepsilon_k^T(\Theta)\right\}, \\ \varepsilon_k(\Theta) &\triangleq y_k - G(q,\Theta)u_k. \end{aligned}$$

Using the linear regressor form in (5) it is well known that $\widehat{\Theta}$ satisfying this criterion may be written in closed form as

$$\widehat{\Theta} = \Phi^\dagger Y \tag{6}$$

where $\Phi^\dagger$ is the left pseudoinverse of Φ. If there is sufficient input excitation for the indicated inverse to exist this will be given by

$$\Phi^\dagger = (\Phi^T\Phi)^{-1}\Phi^T. \tag{7}$$

As is well known, due to considerations of numerical robustness and computation time, in practice $\widehat{\Theta}$ is never calculated according to (6),(7). Rather, the normal equations (5) are solved via an LU or singular value decomposition.

Typically this least squares computation will involve $O(N)$ calculations [1]. In addition the generation of Φ involves deriving N samples of the outputs of $mp \times n$ filters. This will involve approximately $O(N^2)$ calculations. Compared to general prediction error methods this is a very light computational load since no iterative process is required for minimising the prediction error criterion.

Compared to the new 4SID methods of MIMO system identification, the computational load of the methodology of this paper is comparable. There are many suggestions for how 4SID methods can be numerically implemented, but a common thread seems to be that the major computational load comes from 2 SVD's and 2 least squares computations (see [9] for details). This implies [1] $O(N^3)$ calculations, which is of the same order of magnitude of the approach proposed here.

This section describes the construction of the bases employed in (3) which are constructed to be orthonormal with respect to the usual inner product that is useful in discrete time system analysis and which is defined by

$$\langle \mathcal{B}_n, \mathcal{B}_m\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi}\mathcal{B}_n(e^{j\omega})\overline{\mathcal{B}_m(e^{j\omega})}\,d\omega. \tag{8}$$

The idea is that this orthonormality is to be preserved while at the same time incorporating prior knowledge in the estimation problem by choosing the poles $\{\xi_0, \xi_1, \xi_2, \cdots, \xi_{p-1}\}$ of the $\{\mathcal{B}_n\}$ basis function to be near where the poles of the true system $G(q)$ are believed to lie.

The construction of a scalar orthonormal basis that satisfies the above requirements has been developed elsewhere [5, 7] where the following basis functions are presented

$$\mathcal{B}_n(z) = \left(\frac{\sqrt{1-|\xi_n|^2}}{z-\xi_n}\right)\prod_{k=0}^{n-1}\left(\frac{1-\overline{\xi_k}z}{z-\xi_k}\right) \tag{9}$$

Notice that for the choice $\xi_k = 0$ for all k this basis degenerates to that corresponding to an FIR model structure. For the choice of $\xi_k = \xi \in R$ of all poles being the same and real, the Laguerre basis [10] is obtained as a special case. However, it would seem more appropriate to not use this basis formulation in (9) in such a restricted setting, but rather choose the poles $\{\xi_k\}$ in a distributed fashion that most accurately reflects prior knowledge of $G(q)$.

Notice that if any of the poles $\{\xi_k\}$ are chosen as complex, then the formulation (9) has a complex valued impulse response which is inappropriate. In [7, 5] it is shown how this may easily be circumvented by still using the construction (9), but if ξ_n is chosen as complex, then another pole ξ_{n+1} must also be chosen as the complex conjugate $\xi_{n+1} = \overline{\xi}_n$. This leads now to two basis function $\mathcal{B}_n$ and $\mathcal{B}_{n+1}$ with complex valued impulse responses. The idea now is that these may be linearly combined in an infinite variety of ways to yield two new basis function $\mathcal{B}'_n, \mathcal{B}''_n$ which have the same complex valued poles, are orthonormal to one another, and have **real** valued impulse responses. These latter two basis function are the ones used in the identification procedure. The details of this linear combination are unimportant here, but lead to $\mathcal{B}'_n$ having the form

$$\mathcal{B}'_n(z) = \frac{\sqrt{1-|\xi_n|^2}(\beta z+\mu)}{z^2+(\xi_n+\overline{\xi_n})z+|\xi_n|^2}\prod_{k=0}^{n-1}\left(\frac{1-\overline{\xi_k}z}{z-\xi_k}\right)$$

where $x = (\beta, \mu) \in R^2$ is any choice lying on the ellipse

$$x^T M x = |1-\xi_n^2|^2 \tag{10}$$

with

$$M \triangleq \begin{pmatrix} 1+|\xi_n|^2 & 2\mathrm{Re}\{\xi_n\} \\ 2\mathrm{Re}\{\xi_n\} & 1+|\xi_n|^2 \end{pmatrix}.$$

$$\mathcal{B}''_n(z) = \frac{\sqrt{1-|\xi_n|^2}(\beta' z + \mu')}{z^2 + (\xi_n + \overline{\xi_n})z + |\xi_n|^2} \prod_{k=0}^{n-1} \left(\frac{1 - \overline{\xi_k} z}{z - \xi_k} \right)$$

with (β', μ') related to the initial choice of (β, μ)

$$\begin{pmatrix} \beta' \\ \mu' \end{pmatrix} = \frac{1}{\sqrt{1-\alpha^2}} \begin{pmatrix} \alpha & 1 \\ -1 & -\alpha \end{pmatrix} \begin{pmatrix} \beta \\ \mu \end{pmatrix}, \qquad (11)$$

$$\alpha \triangleq \frac{\xi_n + \overline{\xi_n}}{1 + |\xi_n|^2}.$$

A special case of this construction is when only one fixed complex mode $\xi_k = \xi$ is considered and where the following special choice satisfying (10) is made

$$(\beta, \mu) = \left(0, \sqrt{(1-\alpha^2)(1+|\xi_n|^2)} \right)$$

in which case (11) gives

$$(\beta', \mu') = \sqrt{(1+|\xi_n|^2)}(1, -\alpha).$$

which is the Kautz basis used in [11]. Different initial choices for (β, μ) satisfying (10) give an infinite number of second order bases other than this Kautz one.

This provides a general construction for scalar orthonormal $\{\mathcal{B}_k(q)\}$. However, in the context of MIMO estimation, it is possible that more precise prior knowledge is available in the form of p guesses $\{\xi_{ij}^1, \cdots, \xi_{ij}^{p-1}\}$ about likely pole locations in the individual entries $\{G_{ij}(q)\}$ of $G(q)$. In this case, the construction for scalar orthonormal bases can be extended so that matrix $\{\mathcal{B}_k\}$ can be formed as

$$\mathcal{B}_k(q) = \begin{pmatrix} \mathcal{B}_{11}^k(q) & \mathcal{B}_{12}^k(q) & \cdots & \mathcal{B}_{1n}^k(q) \\ \vdots & & & \vdots \\ \mathcal{B}_{m1}^k(q) & \mathcal{B}_{12}^k(q) & \cdots & \mathcal{B}_{mn}^k(q) \end{pmatrix}$$

where

$$\mathcal{B}_{\ell r}^k = \frac{1}{\sqrt{mn}} \frac{\sqrt{1-|\xi_{\ell r}^k|^2}}{q - \xi_{\ell r}^k} \prod_{t=0}^{k-1} \left(\frac{1 - q\overline{\xi_{\ell r}^t}}{q - \xi_{\ell r}^t} \right). \qquad (12)$$

These bases are orthonormal with respect to the usual definition for an inner product on a space of matrix valued functions:

$$\langle \mathcal{B}_m, \mathcal{B}_n \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{Tr} \left\{ \mathcal{B}_m(e^{j\omega}) \mathcal{B}_n^*(e^{j\omega}) \right\} \, d\omega$$

where $\cdot^*$ denotes conjugate transpose. As previously mentioned, use of this matrix basis provides more flexibility in the incorporation of prior knowledge with a concomitant improvement in estimation accuracy. The penalty is an increased difficulty in theoretical analysis.

6 Estimation Accuracy

The section addresses the question of analysing the performance of the system identification scheme proposed. estimation errors can arise from two sources. One of them is a so-called 'undermodelling' induced error that results from the parsimony of the model structure (3) which cannot completely describe the true dynamics $G(q)$. The second source of errors arises from the noise corruption $H(q)e_k$ of the measurements.

In this section, these two errors are quantified separately. Due to the linearity of the model structure (3), and hence the linearity (in the data) of the estimate (6), the total error can be overbounded by the sum of these two error terms.

Attacking the quantification of the undermodelling induced error first, note that one characteristic of this error is that it is such that the gain of the true system is underestimated in an 'on average' sense over all frequencies

Theorem 1. *Define*

$$\Theta_0 = \lim_{N \to \infty} \widehat{\Theta}.$$

Then regardless of whether the $\{\mathcal{B}_\ell(q)\}$ are matrix or scalar

$$\int_{-\pi}^{\pi} Tr \left\{ G(e^{j\omega}, \Theta_0) G^*(e^{j\omega}, \Theta_0) \Phi_u(\omega) \right\} \, d\omega \leq$$

$$\int_{-\pi}^{\pi} Tr \left\{ G(e^{j\omega}) G^*(e^{j\omega}) \Phi_u(\omega) \right\} \, d\omega$$

Proof. Follows using the same methods as employed in Theorem 1 of [6]. □□□

While this gives some insight into how the undermodelling error behaves, it is not a result of very fine structure. What is possibly of more interest is some quantification of error at a particular frequency, rather than averaged over all frequencies.

For the SISO case and for particular orthonormal bases this quantification has been performed in [10, 11, 2]. The methods employed in these works involve establishing an isomorphism between the bases used in those works to the common FIR basis $\{e^{j\omega n}\}$. Standard techniques from Fourier analysis are then employed.

In comparison with the bases studied in [10, 11, 2], the bases (9),(12) studied in this paper enjoy greater flexibility in the possible choice of pole location. Unfortunately, there is a price to be paid for this increased flexibility in that the isomorphic (with respect to FIR) structure is lost.

Nevertheless, it is still possible to develop results equivalent to those in [10, 11, 2], but with considerably more effort. The main idea is to draw inspiration from the Szegö theory of orthonormal polynomials [8] and develop a so-called 'Christoffel-Darboux' formula for the Reproducing Kernel associated with the basis.

Theorem 2. *Christoffel-Darboux formula: Define the Blaschke product*

$$\varphi_p(z) = \prod_{k=0}^{p-1} \frac{1 - \overline{\xi_k} z}{z - \xi_k}.$$

the orthonormal system $\{\mathcal{B}_k\}$ *can be expressed as*

$$\sum_{k=0}^{p-1} \overline{\mathcal{B}_k(\mu^{-1})}\mathcal{B}_k(z^{-1}) = \left\{ \frac{1 - \overline{\varphi_p(\mu^{-1})}\varphi_p(z^{-1})}{1 - \overline{\mu}z} \right\} \overline{\mu}z. \tag{13}$$

Proof. This is too lengthy to be presented here. It does not use the methods of Szegö which rely heavily on the properties of polynomials. Rather the proof is constructive. □□□

In [7] the basis (9) was shown to be H_2 complete if and only if $\sum(1 - |\xi|_k) = \infty$. The Christoffel-Darboux like result in the previous theorem shows that under the same conditions for H_2 completeness, (9) is also complete in the disk algebra; this latter space is commonly employed in the so-called 'Estimation in H_∞' literature.

Perhaps more importantly though, the previous Theorem allows the undermodelling error to be bounded as follows.

Theorem 3. *Suppose* $G_{ij}(z)$ *has partial fraction expansion*

$$G_{ij}(z) = \sum_{\ell=0}^{r-1} \frac{\alpha_\ell}{z - \gamma_{ij}^\ell}$$

where all the poles satisfy $|\gamma_{ij}^\ell| < 1$. *Put* $\widehat{G}_{ij}(z)$ *as the best* H_2 *approximation to* $G_{ij}(z)$ *with respect to the* p *basis functions* $\{\mathcal{B}_0, \mathcal{B}_1, \cdots, \mathcal{B}_{p-1}\}$

$$\widehat{G}_{ij}(z) = \sum_{k=0}^{p-1} \langle G_{ij}, \mathcal{B}_k \rangle \mathcal{B}_k(z).$$

Then

$$|G_{ij}(e^{j\omega}) - \widehat{G}_{ij}(e^{j\omega})| < \sum_{\ell=0}^{r-1} \left| \frac{\alpha_\ell}{e^{j\omega} - \gamma_{ij}^\ell} \right| \left| \prod_{k=0}^{p-1} \frac{\gamma_{ij}^\ell - \xi_k}{1 - \overline{\xi_k}\gamma_{ij}^\ell} \right|$$

Proof. Use Cauchy's Integral Theorem to write $G(\mu^{-1})$ with $|\mu| < 1$ as

$$G(\mu^{-1}) = \frac{1}{2\pi j} \oint_{|z|=1} \frac{G(z^{-1})}{z - \mu}\, dz$$

and then use the previous Theorem together with Cauchy's Residue Theorem and the Triangle inequality. □□□

This shows that the undermodelling error for MIMO estimation with the scalar basis function model structure depends geometrically (in model order p) on the distance between the true pole position γ_{ij}^ℓ and the guess ξ_k of this position. This suggests that the convergence of the expansion (3) can be rapid, and hence the truncation at finite model order p is valid.

A key advantage of the simple linear in the parameter model structure being employed is that it is possible to quantify the noise induced error (under a stochastic model for $\{e_k\}$) as

$$\mathrm{Cov}\left\{\widehat{\Theta}\right\} = \Phi^\dagger E\left\{VV^T\right\}(\Phi^\dagger)^T$$

$$V \triangleq [H(q)e_0, \cdots, H(q)e_{N-1}].$$

If the previous analysis supports the belief that the undermodelling error is negligible, then this quantification can then be used to develop confidence intervals about the parameters.

This sort of calculation does not emphasise how the choice of the basis functions $\{\mathcal{B}_k\}$ affect the noise induced error. To elucidate the latter it is necessary to perform an asymptotic analysis which is akin to the results for scalar FIR bases presented in [12]. Unfortunately, the analysis becomes quite complicated, and we can only provide a result for white and uncorrelated measurement noise with scalar basis in the model structure.

Theorem 4. *Define*

$$V_{\ell,k}(\omega) \triangleq \frac{1}{\beta_p(\omega)} \lim_{N\to\infty} NE\left\{ |G_{\ell,k}(e^{j\omega}, \widehat{\Theta}) - G_{\ell,k}(e^{j\omega})|^2 \right\}.$$

where

$$\beta_p(\omega) = 2\pi \sum_{k=0}^{p-1} |\mathcal{B}_k(e^{j\omega})|^2$$

Then for MIMO estimation using scalar $\{\mathcal{B}_\ell(q)\}$, *for* $H(q) = I$ *and* $E\left\{e_k e_k^T\right\} = \sigma_e^2 I$, *and assuming* p *is so large that the undermodelling error is negligible*

$$\sigma_e^2 \left(\sup_\omega \overline{\sigma}(\Phi_u(\omega)) \right)^{-1} \leq V_{\ell,k}(\omega) \leq \sigma_e^2 \left(\inf_\omega \underline{\sigma}(\Phi_u(\omega)) \right)^{-1}.$$

This shows, that as per the SISO estimation case [7, 5], the error of the estimate at a particular frequency is proportional to the sum of the squared magnitudes of the responses of the basis functions at that frequency.

7 Simulation Example

In this section the utility of the orthonormal basis function approach to MIMO identification is illustrated with a brief simulation study. Suppose an underlying continuous time MIMO system, with $m = 2$ outputs and $n = 2$ inputs, with transfer matrix representations given by:

$$G(s) = \begin{bmatrix} \dfrac{0.1e^{-s}}{(s+1)(s+0.1)} & \dfrac{1}{(s+2)(s+0.5)} \\ \dfrac{0.21}{(s+0.7)(s+0.3)} & \dfrac{0.32e^{-s}}{(s+0.4)(s+0.8)} \end{bmatrix} \tag{14}$$

Notice that there are large discrepancies between the dynamics in the individual entries of $G(s)$ and also there are strong couplings between all inputs and outputs. Furthermore, there are time delays in some terms.

Consider the situation where to estimate the dynamics (14) there is available an observation of $N = 300$ samples spaced 1 second apart of the outputs $\{y_k^1\}, \{y_k^2\}$ of $G(q)$ when the inputs $\{u_k^1\}, \{u_k^2\}$ are unit amplitude square waves of fundamental frequencies 0.02 Hz and 0.05 Hz respectively. The output data is corrupted with stationary and white Gaussian distributed noise of variances $\sigma_{e^1}^2 = \sigma_{e^2}^2 = 0.2$ and with colouring filter $H(q) = I$.

pole locations in (14) is accurate only to within 50 percent so that guesses to the pole locations can be made as follows

$$\begin{aligned}\xi_{11} &= (-2,-0.2,-2,-0.2,-0.5),\\ \xi_{12} &= (-4,-1,-3,-1,-1),\\ \xi_{21} &= (-1.4,-0.6,-1.4,-0.6,-0.1),\\ \xi_{22} &= (-0.8,-1.6,-0.8,-1.6,-0.2)\end{aligned}$$

in the matrix basis function construction (12). These choices corresponds to a fifth order model and twenty scalar parameters being estimated. The results when using this for estimation are shown in figure 1. The estimates are quite accurate even with the effects of time delays and inaccurate prior knowledge. To finish the sim-

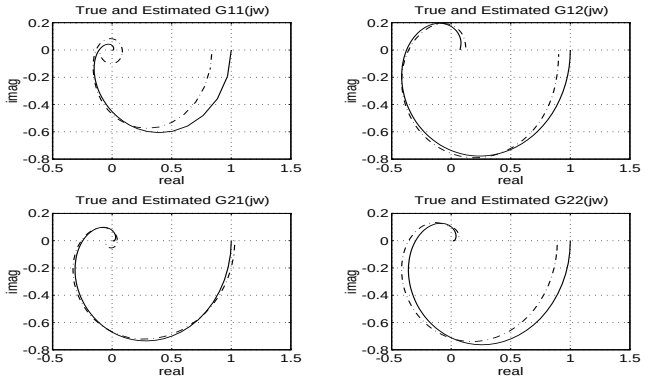

Figure 1: *True and Estimated Frequency Response. The four plots shown are the four possible responses from the 2 inputs to the 2 outputs.*

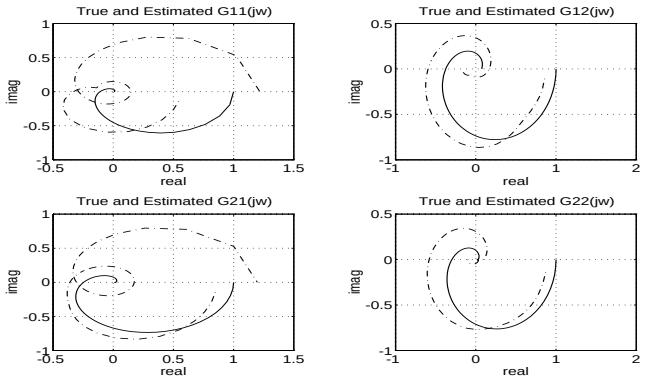

Figure 2: *Results of fitting state space model using prediction error method.*

ulation example, a more conventional approach approach is investigated whereby a state space model structure is estimated using a standard prediction error identification algorithm. For details of how this was implemented see [3].

An eighth order model was used to capture the eight different modes in (14). This implies (after a suitable canonical form for the state space model structure is selected) the estimation of 48 parameters. The same data used to provide the estimates shown in figure 1 was used for identification of the state space model whose the results are somewhat less accurate than for the basis function solution, and involve an order of magnitude more computation time to derive. The lack of accuracy is due to the increased measurement noise induced variability resulting from an increased number of parameters being estimated. On the other hand, the derivation of the estimate in 2 does not require approximate prior knowledge of the location of system modes.

8 Conclusion

This paper has presented a simple and computationally cheap MIMO system identification scheme that is a direct extension of the familiar method of using least-squares with an FIR model structure. A key feature of the extension is that prior knowledge of the poles of the system being identified can be incorporated into the estimation problem in order to increase the estimation accuracy. This is done in such a way that orthonormality is preserved, since this facilitates theoretical analysis to be performed to quantify the undermodelling and measurement noise induced errors in the estimation.

REFERENCES

[1] G. Golub and C. V. Loan, *Matrix Computations*, 1989.

[2] P. Heuberger, P. V. den Hof, and O. Bosgra, *A generalized orthonormal basis for linear dynamical systems*, TAC-40 (1995), pp. 451–465.

[3] L. Ljung, *Identifying state space models with mathwork's system identification toolbox*, in Proceedings of SYSID 1991, pp. 1582–1584.

[4] L.Ljung and Z.D.Yuan, *Asymptotic properties of black-box identification of transfer functions*, TAC-30 (1985), pp. 514–530.

[5] B. Ninness, *Orthonormal bases for geometric interpretations of the frequency response estimation problem*, in Proceedings of SYSID 1994, pp. 591–596.

[6] ——, *Integral constraints on the accuracy of least squares estimation*, Automatica, (To appear, 1996).

[7] B. Ninness and F. Gustafsson, *A unifying construction of orthonormal bases for system identification*, in Proceedings of CDC 1994, pp. 3388–3393.

[8] G. Szegö, *Orthogonal Polynomials*, vol. 23, Colloquium publications - 1959.

[9] M. Viberg, *Subspace methods in system identification*, in Proceedings of SYSID 1994, pp. 1–12.

[10] B. Wahlberg, *System identification using Laguerre models*, TAC-36 (1991), pp. 551–562.

[11] B. Wahlberg, *System identification using Kautz models*, TAC-39 (1994), pp. 1276–1282.

[12] Z. Yuan and L. Ljung, *Black box identification of multivariable transfer functions:Asymptotic properties and optimal input design*, IJC, 40 (1984), pp. 233–256.