Identification of multivariable linear parameter varying models, a frequency domain approach ⋆

B. Vergauwen * **J. Lataire** ** **R. Pintelon** ** **B. De Moor** *

* *KU Leuven, Department of Electrical Engineering (ESAT) Stadius Center for Dynamical Systems, Signal Processing and Data Analytics (e-mail: bob.vergauwen@esat.kuleuven.be; bart.demoor@esat.kuleuven.be)*
** *Vrije Universiteit Brussel, Department of Fundamental Electricity and Instrumentation (ELEC) (e-mail: jlataire@vub.ac.be; rik.pintelon@vub.ac.be)*

Abstract: In this paper an extension is made from the one dimensional frequency-domain least-squares estimator for identifying linear, continuous-time, parameter-varying dynamical systems to a multivariable setting. All the estimators operate entirely in the frequency domain, this allows for a full control over the frequency band of interest. Working in the frequency domain is also a natural representation to include non-parametric noise models in the estimation algorithms. Much of the results of the one dimensional parameter varying system theory can be generalised to the multivariable case by an extension of the basis function representation of the model. In this paper the family of the total least squares and maximum likelihood estimators is extended to estimate the presented model structure for multi-input, multi-output systems. This extension is made for a general parameter invariant additive noise model. To demonstrate the results and concepts of this new set of estimators an example on simulation data has been worked out.

Keywords: LPV, Frequency domain, Maximum likelihood, MIMO, ODE input output model, Flutter.

1. INTRODUCTION

Some systems are inherently time-varying due to their nature. Modelling these systems with a correct time or parameter varying class of models will in most cases increase the model accuracy significantly compared to a time invariant model class. For single-input single-output (SISO) linear parameter varying (LPV) systems, a set of estimators has already been developed by Goos et al. (2017). In this paper we present the extension of these estimators to the class of multi-input multi-output (MIMO) systems. From an identification perspective most of the concepts for the single-input single-output systems can be extended to MIMO systems by extending the basis representation of the SISO model. The biggest challenge is to correctly include the spectral characteristics of the noise to make the estimator consistent. In this paper we restrict ourselves to a time invariant coloured additive noise source. The extension to parameter varying noise models is not made. In the second part of this paper the model description is derived to represent parameter varying models. The emphasis on the linearity of the unknowns and the vector space representation of the model is made. This valuable insight will make the transition from a SISO description to a MIMO context much more clear.

In the third section the idea of the Total Least Squares (TLS) estimator will shortly be stated. This simple estimator is not consistent when the signals are disturbed by noise. At the basis of deriving consistent estimators is the maximum likelihood principle and the error covariance matrix. The derivation of the error covariance matrix is done in the fourth section.

In the fifth section all of the previous concepts are combined to end up with the generalisation of the known consistent estimators for LPV systems. In the last section a descriptive simulation example is presented to illustrate the various concepts of the estimation process.

2. MIMO LPV SYSTEMS

In this section the model representation of linear parameter varying multi input multi output systems is presented. To derive this model structure we start with a SISO frequency domain model derived by Lataire and Pintelon (2011). To extend this SISO model a set of matrices is presented to represent the MIMO structure of the system. The key observation is that the models represented in this paper are linear in the unknown parameters. This allows the parameters of the model to be estimated using a least squares method.

⋆ This research was sponsored by the Belgian Federal Science Policy Office: IUAP P7/19 (DYSCO, Dynamical systems, control and optimization, 2012-2017)

2.1 System representation

Over the years a multitude of different system representation have been derived to describe linear parameter varying systems. In this paper we restrict ourselves to an input output model representation with parameter varying coefficients. The model equation corresponding to this class is given by,

$$\sum_{n}^{N_a} \mathbf{A}_n(\mathbf{p}(t)) \frac{d^n \mathbf{y}}{dt^n} = \sum_{n}^{N_b} \mathbf{B}_n(\mathbf{p}(t)) \frac{d^n \mathbf{u}}{dt^n}. \tag{1}$$

In this equation the bold font represents a matrix or vector quantity. In a MIMO setting with n_u inputs and n_y outputs, the matrices $\mathbf{A}_n$ en $\mathbf{B}_n$ are parameter varying matrices whose value depend on the scheduling function $\mathbf{p}(\cdot)$. The function $\mathbf{p}(\cdot)$ is allowed to be multi dimensional, in that case the system of (1) will depend on several scheduling parameters. It's also possible to allow the matrices $\mathbf{A}_n$ and $\mathbf{B}_n$ to dynamically depend on the scheduling function $\mathbf{p}(\cdot)$, e.g. be a function of $d\mathbf{p}(t)/dt$. The case of a dynamic dependency of the coefficients in (1) can be modelled as a static one. This is done by extending the parameter function $\mathbf{p}(\cdot)$ to explicitly include the dynamic behaviour of the parameter. In our example we would extend the parameter $\mathbf{p}(t)$ to include the derivative $d\mathbf{p}(t)/dt$ explicitly. Using this method $\mathbf{p}(\cdot)$ can be extended such that the coefficients $\mathbf{A}_n$ and $\mathbf{B}_n$ only depend in a static way on $\mathbf{p}(\cdot)$. One consequence of working with a MIMO system description is that (1) effectively consists of a system of n_y different equations. These equations are in general coupled but can be decoupled by choosing an appropriate basis to represent the model.

Common denominator One specific LTI input output system representation for MIMO systems, popularized by Kailath (1980), is the common denominator description. In this system representation the constant matrix $\mathbf{A}_n|_{\mathbf{p}}$ is decomposed as the product of a scalar function and the unity matrix for each value of $\mathbf{p}$. This representation couples the different equations in (1) and imposes that all the elements of the transfer matrix have the same poles. Although the focus of this paper lies on a common denominator representation, the theory and estimation methods can be applied to a general model representation.

2.2 Basis function representation

The idea first proposed in Lataire (2011a) is to write the model equations for a linear parameter varying systems as linear combinations of parameter varying and Laplace basis functions. In this paper, we extend this concept to a multi variate setting by introducing a new basis for the matrix representation of $\mathbf{A}_n(\cdot)$ and $\mathbf{B}_n(\cdot)$. Under the mild condition that each individual element of the matrices $\mathbf{A}_n(\mathbf{p}(t))$ and $\mathbf{B}_n(\mathbf{p}(t))$ is a continuous function of the parameter $\mathbf{p}(\cdot)$ over the closed domain of $\mathbf{p}(t)$. Weierstrass convergence theorem ensures that there exists a basis of the vector space of analytic functions ($C^\infty(\mathbf{p})$) such that each element can be approximated arbitrary well in two norm. Furthermore each of the matrices $\mathbf{A}_n|_p$ can be represented as an element in a matrix vector space. The complete space to represent every matrix $\mathbf{A}_n(\cdot)$ and $\mathbf{B}_n(\cdot)$ is the product space of these two spaces.

$$\mathbf{A}_n(\cdot) \in C^\infty \times \mathbb{R}^{n_y \times n_y},\ \mathbf{B}_n(\cdot) \in C^\infty \times \mathbb{R}^{n_y \times n_u}$$

The elements of the matrix basis are denoted by $\mathbf{M}_l$. For the common denominator basis, only one basis element is needed to represent the matrix basis for $\mathbf{A}_n$. This element is the unity matrix of dimension $n_y \times n_y$. To represent the matrices $\mathbf{B}_n$ a basis is constructed containing the matrices with only one element different from zero. The basis functions to represent the parameter variation are represented by α_p, here p is an index and not the parameter. A good choice for this basis is the set of orthogonal Legendre polynomials. The only restriction on these basis functions is that they need to construct a basis of the individual elements of the matrices $\mathbf{A}_n$ and $\mathbf{B}_n$.

Laplace basis function Together with each matrix polynomial $\mathbf{A}_n(\cdot)$ and $\mathbf{B}_n(\cdot)$ respectively the derivative operator d^n/dt^n is associated. The property of linearity for the derivative operator implies that one can construct a vector space to represent these derivative operators. We define one such basis of this space to be $\langle \Psi_n \rangle$.

In the Laplace domain these basis vectors are represented by polynomials in the Laplace variable s. The individual basis elements $\langle \Psi_n \rangle$ can be constructed as a linear combination of derivative operators up to order n. The reasoning for choosing a general basis to represent the derivative operator is to improve the conditioning of the final algorithms. After an estimation in this general basis in s a simple linear transformation can be used to transform the model back to the form of (1). The general discussion on this topic is given in chapter 6 of Lataire (2011a). The important conclusion of this section is that there are three different vector spaces needed to represent each side of the model in (1). One vector space to represent the parameter variation, one vector space to represent the differential operators and one to represent the MIMO (matrix) structure. Using this representation and making the abstraction of the specific basis, (1) can be written in a compact form,

$$\sum_{\mathbf{i}} a_{\mathbf{i}} \mathbf{k}_{\mathbf{y}} \langle \mathbf{i} \rangle \mathbf{y} = \sum_{\mathbf{j}} b_{\mathbf{i}} \mathbf{k}_{\mathbf{u}} \langle \mathbf{j} \rangle \mathbf{u}. \tag{2}$$

Here $\mathbf{i}$ is called the generalised index of the basis. This index is a combination of the three indices corresponding to the three different vector spaces. The operators $\mathbf{k}_{\mathbf{y}} \langle \mathbf{i} \rangle$ and $\mathbf{k}_{\mathbf{u}} \langle \mathbf{j} \rangle$ are a product of the three different basis elements.

2.3 Frequency domain representation

The identification algorithm presented in this paper operates directly in the frequency domain. For this the model of (1) has to be transformed to the frequency domain by applying the Fourier transformation. In general this transformation leads to a transient effect. The effect of the transient can partially be eliminated by using a periodic scheduling or it could be included in the estimation process by adding a smooth transient function $\mathbf{T}(\omega)$. By transforming (2) to the frequency domain and adding a transient term to keep the equality we end up with the following set of equations,

$$\mathcal{F}\left\{\sum_{\mathbf{i}} a_{\mathbf{i}} \mathbf{k}_{\mathbf{y}} \langle \mathbf{i} \rangle \mathbf{y}\right\}_\omega + \mathbf{T}(\omega) = \mathcal{F}\left\{\sum_{\mathbf{i}} b_{\mathbf{i}} \mathbf{k}_{\mathbf{u}} \langle \mathbf{i} \rangle \mathbf{u}\right\}_\omega. \tag{3}$$

For each angular frequency ω we have a set of n_y equations. To make this model computable the last step is to discretise these equations.

2.4 Complete basis representation

For the completeness and to illustrate a computationally effective way to calculate the model equations, the complete basis representation of the model is given. Equation (3) can be written as,

$$\sum_{n}^{N_a} \mathcal{F}\left\{\mathbf{A}_n(\mathbf{p}(t))\mathcal{F}^{-1}\left\{\Psi_n(j\omega)\mathbf{Y}(\omega)\right\}\right\} + \mathbf{T}(\omega) = \sum_{n}^{N_b} \mathcal{F}\left\{\mathbf{B}_n(\mathbf{p}(t))\mathcal{F}^{-1}\left\{\Psi_n(j\omega)\mathbf{U}(\omega)\right\}\right\}. \quad (4)$$

In this equation the input and output signals are transformed to the frequency domain. This transformation allows the calculation of the derivatives to be replaced by its equivalent multiplication in the Laplace domain. This product is then transformed back to the time domain to calculate the product of the parameter varying basis functions.

The next step is to write the matrix polynomials $\mathbf{A}_n(\cdot)$ and $\mathbf{B}_n(\cdot)$ as a linear combination of basis functions. By using the place holder $\mathbf{x}$ to represent the input $\mathbf{u}$ or the output $\mathbf{y}$, $\mathbf{k_x}\langle\mathbf{i}\rangle$ from (2) can be rewritten as,

$$\mathbf{k_x}\langle\mathbf{i}\rangle\,\mathbf{x} = \mathbf{M_l}\alpha_p(\mathbf{p})\mathcal{F}^{-1}\left\{\Psi_n(j\omega)\mathbf{X}(j\omega)\right\}, \mathbf{i} = (l,p,n) \quad (5)$$

System equation for sampled systems For any practical implementation (4) has to be discretised. This implies replacing the continuous signals $\mathbf{y}$ and $\mathbf{u}$ by sampled versions and replacing the continuous time Fourier operator with the discrete $N-$point `DFT`. For the transformation to be reversible the Nyquist condition has to be fulfilled for each term of the sum in equation (4). This substitution results in,

$$T_s\texttt{DFT}\left\{\sum_{\mathbf{i}} a_{\mathbf{i}}\mathbf{M}_l\alpha_p(\mathbf{p})\texttt{IDFT}\left\{\Psi_n(j\omega_k)\mathbf{Y}(\omega_k)\right\}\right\}_{\omega_k} + \sum_i w_i\mathbf{T}_i(\omega_k) = T_s\texttt{DFT}\left\{\sum_{\mathbf{i}} b_{\mathbf{i}}\mathbf{M}_l\beta_p(\mathbf{p})\texttt{IDFT}\left\{\Psi_n(j\omega_k)\mathbf{U}(\omega_k)\right\}\right\}_{\omega_k} \quad (6)$$

The most important aspect of this model representation is that it's linear in its unknowns $a_{\mathbf{i}}$ and $b_{\mathbf{i}}$. The main difference between this model representation and the single-input, single-output model is the dimension of these equations. For every discrete frequency ω_k there are n_y equations to solve. The linearity in the unknowns $a_{\mathbf{i}}$, $b_{\mathbf{i}}$ and w_i implies that (6) can be rewritten as a matrix multiplication. Suppose that we select a band containing N_f frequencies where we want to solve the equations, this results in a system of $n_y \times N_f$ equations, such that

$$e(\mathbf{\Theta}) = \mathbf{K}\mathbf{\Theta} \text{ with } \mathbf{\Theta} = [a_{\mathbf{i}}\ b_{\mathbf{i}}\ w_i]^T \text{ and } \mathbf{K} = [\mathbf{K_y}\ -\mathbf{K_u}\ \mathbf{T}].$$

Where $e(\mathbf{\Theta})$ is called the model error, $\mathbf{\Theta}$ is the vector of the model parameter and $\mathbf{K}$ is the regression matrix of the problem. The submatrices $\mathbf{K_y}, \mathbf{K_u}$ and $\mathbf{T}$ contain respectively the vectors who contain the output, input and transient contributions. Each column of this regression matrix contains the n_y equations corresponding to the vector $\mathbf{k_x}\langle\cdot\rangle$ or the transient. To deal with the MIMO structure of the problem we place the n_y equations underneath each other in the regression matrix. This implies that the size of the matrix is $(n_y.n_f) \times n_p$. Here n_p is the total number of parameters and n_f the total number of selected frequencies.

The columns of the submatrix $\mathbf{K_y}$ and $\mathbf{K_u}$ can be calculated by using (5) with the corresponding substitution The columns of $\mathbf{T}$ contain the smooth contributions of the transient.

2.5 Rank condition

To guarantee that there exists a unique solution to equation (6) the regression matrix $\mathbf{K}$ must be of rank $n_p - 1$. This condition is known as the rank condition and does place restrictions on the model class as well as on the input and the output signals.

Formulating the rank condition in a formal way would not contribute to the general understanding of the algorithms in this paper. Informally it implies three things. First and foremost, all of the inputs should de sufficiently different from each other. Secondly, all of the signals must excite the selected frequency domain sufficiently. Lastly, the model complexity should not be taken too high. There is one condition where the user has little or no control over, that is the output signal. When for example two outputs are identical to each other (e.g. as a result of a bad experimental set up) the rank condition could be jeopardized. This can be solved by using the common denominator basis.

3. TLS ESTIMATOR

After showing that the model for the multivariate linear parameter varying system is linear in its unknowns it is clear that a simple total least squares estimator can be applied to find an estimate for $\mathbf{\Theta}$. The cost function for this linear estimator is given by,

$$\hat{\mathbf{\Theta}}_{TLS} = \underset{\mathbf{\Theta}}{\text{argmin}} = e(\mathbf{\Theta})^H e(\mathbf{\Theta}) = \mathbf{\Theta^H}(\mathbf{K^H}\mathbf{K})\mathbf{\Theta} \quad \text{s.t.} ||\mathbf{\Theta}||_2 = 1.$$

This simple estimator is not consistent in the presence of white or colored noise corrupting the data. The TLS estimator is only consistent when the disturbances on the columns of the regression matrix are mutually uncorrelated with the same variance. Due to the application of the basis transformations on the data, this assumption does not hold in general. More robust estimator are derived based on the maximum likelihood principle.

4. NOISE MODEL

To derive a general maximum likelihood estimators, the influence of the different noise sources on the uncertainty of the prediction error $e(\mathbf{\Theta})$ has to be calculated. The error covariance matrix on the estimation error is given by,

$$C_e = \mathbb{E}\left\{\Delta e(\mathbf{\Theta})\Delta e(\mathbf{\Theta})^H\right\} \in \mathbb{R}^{(n_f.n_y)\times(n_f.n_y)}$$

In this formula Δe represents the variance on the equation error. For LPV systems the most general noise model

would be parameter varying by nature. However we restrict the analysis to a *stationary noise model with a full correlation matrix*. Each of the noise realizations can be *coloured* and the noise is assumed to be *additive* and independent of the deterministic part of the signal (no feedback). We do not present techniques to estimate a noise model for the LPV system we are trying to estimate. The details on how such a model could be obtained can be found in Pintelon et al. (2015).

The first step in the derivation of the noise covariance on the equation error is to calculate the frequency depending noise on the columns of the regression matrix. Recall (5) to calculate the columns of the regression matrix, by applying basis element $\mathbf{i}$ to the input or output data. The noise on the spectrum $\mathbf{X}(\omega)$ is denoted by $\Delta_{\mathbf{x}}(\omega)$. The frequency depending variance of each column of $\mathbf{K}$ is than given by,

$$\Delta\mathbf{K}_{\mathbf{X},l,n,p} = \mathbf{M}_l \mathtt{DFT}\left\{b_p(\mathbf{p}) * \left\{\Psi_n(j\omega_k)\Delta_{\mathbf{X}}(j\omega_k)\right\}\right\}.$$

This expression can be derived by removing the deterministic part from the columns. In this first step we place a restriction on the structure fo the matrix $\mathbf{M}_l$. Recall that dimension of $\Delta_{\mathbf{X}}$ is equal to dimension of $\mathbf{X}$ (respectively the number of inputs or outputs). By multiplying the result of the DFT by a general (full) matrix $\mathbf{M}_l$ the sum of noise contributions must be taken into account. To avoid this extra difficulty in the notation we place a restriction on the structure of $\mathbf{M}_l$ to only contain one non-zero element in each row. This ensures that the variance on each column of $\mathbf{K}$ only depends on exactly one noise source. A second result of the matrix multiplication is that the dimension of the frequency depending noise variance on each column is equal to the number of outputs of the system. This result adds to the complexity of the noise description of the MIMO system. As for the regression matrix, this extra dimension is handled by placing the different outputs of the system under each other in the variance matrix.

At this point we can calculate the covariance matrix between two columns $\mathbf{i}$ and $\mathbf{j}$ of the matrix $\mathbf{K}$. The covariance is given by,

$$.C(\mathbf{i},\mathbf{j}) = \mathbb{E}\left\{\Delta\mathbf{K}_{\mathbf{i}}\Delta\mathbf{K}_{\mathbf{j}}^H\right\} \in \mathbb{R}^{(N_f.N_y)\times(N_f.N_y)} \quad (7)$$

This matrix can be interpreted as consisting of $N_y \times N_y$ block matrices. Each of the blocks containing the covariance between the N_y different equations. We index each block with a subscript (l,k). For example the matrix $C_{\mathbf{K}}(\mathbf{i},\mathbf{j})_{(l,k)}$ contains the frequency depending covariance between column $\mathbf{i}$ and $\mathbf{j}$ and equation l and k. By previously assuming that each matrix $\mathbf{M}_l$ has only one element different from zero on each row we limit each block to depend exactly on two noise signals. This reduces the computation of each block $C(\mathbf{i},\mathbf{j})_{(l,k)}$ to be exactly the same as for the SISO case.

To calculate the covariance between two columns on frequency bin k_1 and k_2 we have,

$$C(\mathbf{i},\mathbf{j})(k_1,k_2) = \mathbb{E}\{\mathbf{M}_l\mathtt{DFT}\left\{b_p(\mathbf{p}) * \left\{\Psi_n\Delta_{\mathbf{X}}\right\}\right\}|_{k_1} \cdot \overline{\mathbf{M}_{l'}\mathtt{DFT}\left\{b_{p'}(\mathbf{p}) * \left\{\Psi_{n'}\Delta_{\mathbf{X}'}\right\}\right\}|_{k_2}}\} \quad (8)$$

Every block of this matrix can be calculated in the same way as for the SISO case, due to the restriction on the matrix basis. This implies that $C(\mathbf{i},\mathbf{j})_{(l,k)}$ is calculated by limiting the matrix basis to respectively the $l-$th and $k-$row. When the noise is uncorrelated over the frequencies e.g. ,$\mathbb{E}\left\{\Delta_{\mathbf{X}}(\nu)\Delta_{\mathbf{X}'}(\nu')\right\} = 0$ for $\nu \neq \nu'$ each block of (8) is approximated by,

$$C(\mathbf{i},\mathbf{j})_{(l,k)}(k_1,k_2) \approx \left(\frac{Ts}{N}\right)\sum_{\nu=1}^{N_f-1}\sigma^2_{\mathbf{X}_l,\mathbf{X}'_k}(\nu)\Psi_n(\nu)\overline{\Psi_{n'}(\nu)} \mathtt{DFT}\left\{b_p(t)\right\}|_{k_1-\nu}\overline{\mathtt{DFT}\left\{b_{p'}(t)\right\}|_{k_2-\nu}}. \quad (9)$$

Where $\sigma^2_{\mathbf{X}_l,\mathbf{X}'_k}(\nu)$ is the correlation between the signal $\mathbf{X}'$ and $\mathbf{X}$ after applying the matrix product (effectively selecting one particular in- or output) and between equation l and k. This expression involves a lot of different indices but in essence it's really an extension of the SISO case. With each generalised index $\mathbf{i}$ corresponds a matrix basis which assigns a signal to some equation of the model structure. The subscripts l and k pick respectively the $l-$th and the $k-$th equation.

In order to calculate the full error covariance matrix we introduce the matrix $C(k_1,k_2)_{(l,k)}$ as

$$C(k_1,k_2)_{(l,k)} = \begin{bmatrix} C(\mathbf{1},\mathbf{1})_{(l,k)}(k_1,k_2) & C(\mathbf{1},\mathbf{2})_{(l,k)}(k_1,k_2) & .. \\ C(\mathbf{2},\mathbf{1})_{(l,k)}(k_1,k_2) & & \ddots \\ \vdots & & \end{bmatrix}$$

Put in words, $C(k_1,k_2)_{(l,k)}$ is a matrix of size $n_p \times n_p$ which gives the frequency depending correlation between equation l and k for each of the parameters. For each pair of equations l and k the error covariance matrix is given by,

$$C_{e_{(l,k)}}(k_1,k_2) = \mathbf{\Theta}^H C(k_1,k_2)_{(l,k)}\mathbf{\Theta}$$

When organizing these $n_y \times n_y$ matrices to match the structure of $\mathbf{K}$ we get the final expressions for the total error covariance matrix,

$$C_e = \begin{bmatrix} C_{e_{(1,1)}} & C_{e_{(1,2)}} & \cdots \\ C_{e_{(2,1)}} & \ddots & \\ \vdots & & \end{bmatrix}$$

We can also calculate the column covariance matrix of $\mathbf{K}$, this matrix is defined by,

$$C_{\mathbf{K}} = \mathbb{E}\left\{\Delta\mathbf{K}^H\Delta\mathbf{K}\right\} \in \mathbb{R}^{N_p\times N_p}$$

One element of this matrix is the total covariance between two columns $\mathbf{i}$ and $\mathbf{j}$ of $\mathbf{K}$. When ignoring the correlation of the noise between the different frequencies, one element of the matrix $C_{\mathbf{K}}$ is given by [1],

$$C_{\mathbf{K}}(\mathbf{i},\mathbf{j}) = \sum_{m,n}^{n_y,n_y}\sum_{k\in\mathbb{K}} C(\mathbf{i},\mathbf{j})_{(m,n)}(k,k).$$

This formula is the sum of the diagonal elements of each block matrix of equation (7).

5. MAXIMUM LIKELIHOOD ESTIMATOR

The maximum likelihood estimator derived by Pintelon et al. (1998), is defined as,

[1] This is only an approximation, even when the measurement and the input noise is uncorrelated. Correlation of the different frequencies in the noise on the columns of $\mathbf{K}$ is the result of the different basis operators who are applied to the signals.

$$\boldsymbol{\Theta}_{ML} = \underset{\boldsymbol{\Theta}}{\text{argmin}} \;\; e(\boldsymbol{\Theta})^H C_e^{-1}(\boldsymbol{\Theta}) e(\boldsymbol{\Theta})$$
$$\text{s.t.} ||\boldsymbol{\Theta}||_2 = 1$$

This equation is non linear in its arguments. This asks for easier to compute estimators.

5.1 GTLS

The first approximation of the ML estimator is the generalised total least squares estimator. The cost function is given by,

$$V_{GTLS} = \left\| \left(\mathbf{K} C_{\mathbf{K}}^{\frac{-1}{2}} \right) \left(C_{\mathbf{K}}^{\frac{1}{2}} \Theta \right) \right\|_2, \;\; \text{s.t} \;\; \left\| C_{\mathbf{K}}^{\frac{1}{2}} \Theta \right\|_2 = 1$$

This estimator is generalised total least squares estimator is derived by Pintelon and Schoukens (2012). In this reference it has been shown that this estimator is consistent in the presence of noise.

5.2 BTLS

Closely related to the GTLS cost function is the BTLS Pintelon et al. (1998) cost function. In this estimator the GTLS cost function is iteratively solved. In each iteration the weights of each row of $\mathbf{K}$ are updated according to the calculated uncertainty based on the previous estimate. The cost function is given by,

$$V_{BTLS}(\Theta^n) = \left\| \left(\text{diag}(C_e(\Theta^{n-1}))^{-2} \mathbf{K} C_{\mathbf{K}}^{\frac{-1}{2}} \right) \left(C_{\mathbf{K}}^{\frac{1}{2}} \Theta \right) \right\|_2,$$
$$\text{s.t} \;\; \left\| C_{\mathbf{K}}^{\frac{1}{2}} \Theta \right\|_2 = 1$$

This cost function is consistent for all noise models. However in the presence of noise correlation between the different in- and outputs some efficiency is lost.

5.3 WNLS

The WNLS for single-input single-output systems LPV systems has been defined in Goos et al. (2017). The cost function for MIMO case is,

$$V_{WNLS} = e^H \text{diag}(C_e(\Theta))^{-1} e(\Theta) = \sum_{k \in \mathcal{K}} \sum_{i=1}^{p} \frac{\|e_i(\Theta, k)\|_2^2}{\sigma_{e_i}^2(\Theta, k)}$$

As for the BTLS, the efficiency of this estimator decreases in the presence of noise correlation between the different in- and outputs. The reason for this is that only the diagonal elements of the error covariance matrix are taken into account.

6. FLUTTER APPLICATIONS

In this section we illustrate the performance of the estimators on a simulation example. The simulation model is based on the physical behaviour of a flutter system where one pole pair becomes unstable for some value of the external schedule function. Flutter is a phenomenon that often occurs in mechanical system where there is a feedback between the structural dynamics and the wind or fluid flowing around the structure. The velocity of the wind is the external scheduling parameter. A full description of the phenomenon of flutter, together with an experimental set up and estimation is given by Ertveldt et al. (2014). Based on these results a MIMO fourth order model is constructed where the trajectory of the poles are predefined to mimic flutter behaviour. Extending these experiments to MIMO will greatly improve the data acquisition rate. For a complex system it is possible that some poles will not get picked up by only one sensor. The reason for this is that vibrational modes can be orthogonal to each other in space.

6.1 Identification signal and noise assumptions

For the identification signal a multi sine was used. The above methods are not bound to any specific kind of identification signal. However, each input signal must satisfy the two following conditions. The first restriction on the input signal is the rank condition, as described earlier. Different input signals will affect the condition number of the regression matrix. This condition number is important for the accuracy of the different algorithms [2]. Using orthogonal input signals, e.g. random phase multi sines will increase the accuracy of the algorithms. A second natural restriction is the Nyquist sampling theorem. For a parameter varying system this condition is a bit more strict than in the case of an LTI system. Due to the parameter variation higher frequency components appear in the different columns of the regression matrix. Every term of the summation of equation (4) must satisfy the Nyquist condition. When using a multi sine excitation signal, both conditions can easily be satisfied. It's also easy to select a particular frequency band and maximize the SNR in this region.

In the simulation set up the full knowledge of the noise colour and correlation is assumed. To simplify the discussion and calculations we used white noise on all of the input and output signals.

The output spectrum of the flutter system is shown in Figure 1. For the input signal only 22 frequencies were excited, this makes visualisation of the data easier. For practical experiments more frequencies are excited. The effect of the parameter variation is clearly visible in the output spectrum in the form of skirts that occur around the excited frequencies, a general explanation of this phenomena is explained in chapter 3 of Lataire (2011b).

6.2 Calculating the error covariance matrix

Selecting the model order is the first step in the identification process. The user has full control over the model order, each of the different orders can be chosen separately from each other, this can be done by removing specific columns of the regression matrix. After the selection of the model orders the regression matrix is calculated by using (5).

The bottleneck in the estimation process is by far the calculation of the error covariance matrix. This matrix can become very large, the dimension of which scale with the number of measurements and the number of parameters. To calculate the column covariance matrix equation (9) is used. When the colour of the noise correlation is the

[2] The condition number of the matrix has to be calculated by excluding the smallest singular value. The reason for this is that the rank of the matrix is only equal to $n_p - 1$

same for some of the block matrices, it is sufficient to only calculate one matrix and use a correctly rescaled version of this matrix for the other blocks. For a simple noise model where al the noise is white, calculating the covariance matrix is as fast as for the SISO case.

6.3 Estimation results

To identify the flutter system a third order parameter variation was used. The model order itself was chosen to be four (to estimate two pole pairs). To estimate the transient a 10-th order basis was used. The total number of parameters for this model is 117. We reduced the number of parameters by making the coefficient of the highest order derivative time invariant.

Two estimations were performed. The first estimation was performed with only partial knowledge of the error covariance matrix and only the variance was used. In the second estimation the full information of the correlation matrix was used. Using the full correlation matrix in the estimation yielded a small increase in the accuracy of the estimation. The estimated trajectories of the poles together with the analytic solutions are shown in Figure 2.

7. CONCLUSION

In this paper we extended the frequency domain estimation algorithm for parameter varying single-input single-output system to the family of multi-input multi-output systems. This extension was done through the introduction of a new vector space to represent the various matrices. There are two advantages for opting for a MIMO approach. The first one is the increased data acquisition rate. Measurements can be done in parallel. This decreases the measurement times significantly. The second benefit of using MIMO experiments is that the system can be better excited and it is possible that more poles can be picked up by the sensors. Compared to the SISO case, in the MIMO case the size of the matrices grow with the number of in- and outputs.

In total 4 estimators were presented, the TLS, GTLS, BTLS and WNLS. The difference in the estimation performance between the TLS and the maximum likelihood estimation is not to be overstated. Even when the unity matrix is used for the correlation matrix the improvement is significant. The reason for this is that derivative operators act like a high pass filter and the algorithms will give to much weight to the high frequency components of the noise.

The personal preference of the author is to use the BTLS estimator when the noise model is well estimated. When the SNR of the signal of is high or the uncertainty on the noise model big, the GTLS estimator gives the best results.

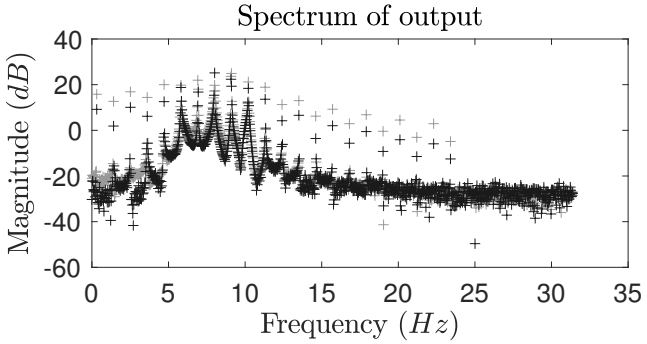


Fig. 1. Output spectrum of the simulation system. The different colours of grey indicate the spectrum of the two different outputs.

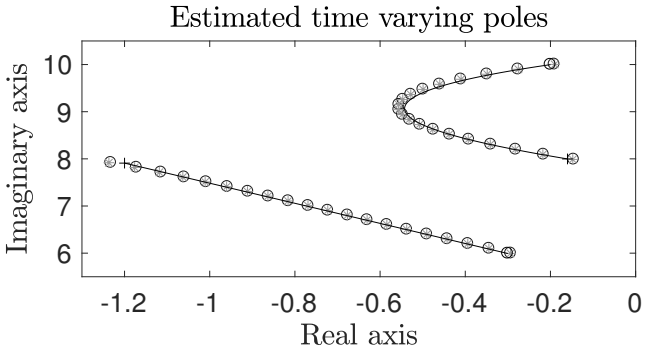


Fig. 2. Comparison of the estimated poles and the analytic solution. For both estimates the BTLS estimator was used. The circles are the estimation result when using the full correlation matrix, the stars use only the noise variance.

REFERENCES

Ertveldt, J., Lataire, J., Pintelon, R., and Vanlanduit, S. (2014). Frequency-domain identification of time-varying systems for analysis and prediction of aeroelastic flutter. *Mechanical Systems and Signal Processing*, 47(1), 225–242.

Goos, J., Lataire, J., Louarroudi, E., and Pintelon, R. (2017). Frequency domain weighted nonlinear least squares estimation of parameter-varying differential equations. *Automatica*, 75, 191–199.

Kailath, T. (1980). *Linear systems*, volume 1. Prentice-Hall Englewood Cliffs, NJ.

Lataire, J. (2011a). *Frequency domain measurement and identification of linear, time-varying systems.* Ph.D. thesis, PhD thesis, Vrije Universiteit Brussel, University Press, Zelzate.

Lataire, J. and Pintelon, R. (2011). Frequency-domain weighted non-linear least-squares estimation of continuous-time, time-varying systems. *IET Control Theory and Applications*, 5, 923–933(10).

Lataire, J. (2011b). Frequency domain measurement and identification of linear, time-varying systems.

Pintelon, R., Guillaume, P., Vandersteen, G., and Rolain, Y. (1998). Analyses, development, and applications of TLS algorithms in frequency domain system identification. *SIAM journal on matrix analysis and applications*, 19(4), 983–1004.

Pintelon, R., Louarroudi, E., and Lataire, J. (2015). Nonparametric time-variant frequency response function estimates using arbitrary excitations. *Automatica*, 51, 308–317.

Pintelon, R. and Schoukens, J. (2012). *System identification: a frequency domain approach.* John Wiley & Sons.