Linear Parameter Varying System Identification for Joint Impedance

by

M.M. Kerklaan

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Friday January 26, 2018 at 10:00 AM.

Student number:	4095723	
Project duration:	October 16, 2016 – January 26, 2018	
Thesis committee:	Dr. ir. A.C. Schouten,	TU Delft, supervisor
	Dr. ir. J. W. van Wingerden,	TU Delft
	Dr. ir. W. Mugge,	TU Delft
ME-BMD MSc Project:	ME51032	

TU Delft

Acknowledgements

Before submitting my thesis to the infinite of the repository I would like to take the time to thank everybody that has helped on my journey.

Firstly I would like to thank my supervisors Alfred, Jan-Willem and Winfred. You guys taught me a lot about being a researcher. I was always amazed how much you guys knew about everything. It was a very comforting thought to know there is still so much more to know.

Next I would like to thank the people in Chicago. It was an amazing experience where I met loads of new and wonderful people. I would like to thank: Carlos, Costa, Daniel, Emma, Eric, Hongchul, Rosalind, Tim, Wil and in particular Michel. I will never forget the slow hungover summer days we spend at your house after a night of drinking.

I would like to thank my good friends Bas and Karel who were always there to listen and push me further. I could not have wished for a more awesome and ridiculous time as a student. You guys made that happen.

Thanks to my family and other friends for the support. You were not always able to help me with the content of my research but knowing you were there if everything went south was always good.

Lastly and in particular I would like to thank my girlfriend Martine. Without you I probably would not be able to write two whole sentences or finish a thought in one paragraph. You were always there when I needed you and you always listened to the endless nonsense that falls out of my head. I love you.

Martijn Kerklaan
4095723
Delft
11 Januari 2018

Contents

Acronyms

CT Continuous Time

DT Discrete Time

LPV Linear Parameter Varying System

LPV-NC-PBSID Linear Parameter Varying Non-Causal Predictor Based Subspace IDentification

LPV-PBSID Linear Parameter Varying Predictor Based Subspace IDentification

LTI Linear Time Invariant

MIMO Multiple Input Multiple Output

MVC Maximum Voluntary Contraction

MVT Maximum Voluntary Torque

PRBS Pseudo Random Binary Signal

SISO Single Input Single Output

SS State Space

SVD Singular Value Decomposition

TI Time Invariant

TV Time Varying

Symbols

$\Phi_{f,k}$	Transition matrix
Γ^f	Extended Observability matrix
$\theta(s)$	Angle
μ_k	Scheduling variable
B	Damping
$\breve{B}_k$	Packed input observer matrix
e_k	Innovation sequence
$\overline{E}, \overline{A}, \overline{B}, \overline{K}, \overline{C}, \overline{D}$	System matrices of the generalized state space description
f	Future window size
I	Inertia
k	Time step
K	Stiffness
$K(t)$	Time-varying stiffness
$\mathcal{K}$	Extended controllability matrix
$\overline{\mathcal{K}}_k^f$	Time-varying extended controllability matrix
l	Number of outputs
$\mathcal{L}_j$	Combination matrix
m	Number of local models
n	Order of the system
n_c	Order of the causal system
n_{ac}	Order of the anti-causal system
N_k^f	Combined order matrix
p	Past window size
$P_{j\|k}$	Matrix containing the structure of the known time varying part
$\bar{q}$	Number of rows in the extended controllability matrix
u_k	Input
U	Stacked input matrix
r	Number of inputs
t	Time
$T(s)$	Torque
y_k	Output
Y	Stacked output matrix
x_k^c	State causal part
X	State sequence
$\mathcal{X}$	States generalized state space description
x_k^{ac}	State anti-causal part
$\bar{z}_k^f$	Stacked input output vector
Z	Stacked time-varying input and output matrix

Linear Parameter Varying System Identification for Joint Impedance Estimation

Martijn Kerklaan*

Delft University of Technology

January 11, 2018

The dynamic relation between the displacement and reaction torque of the human joint is known as joint impedance. Properly quantifying joint impedance has medical potential in the diagnosis, understanding and modelling of movement disorders associated with neuromuscular conditions like cerebral palsy, stroke, dystonia and old age. The identification of joint impedance is often done with Linear Time Invariant (LTI) methods which lack the complexity to fully capture joint impedance over large operating ranges and over time. In this report a novel algorithm was developed which is able to identify joint impedance as a linear parameter varying system. This system description overcomes some of the limitations of the LTI methods. The algorithm was successfully tested in a simulation study in which it identifies a time-varying impedance model with a 5dB signal to noise ratio. Also, the developed method was applied on a force task with position perturbations done with the ankle and wrist. However, these data sets did not show sufficient time-varying behaviour and therefore the algorithm did not lead to better results compared to LTI methods. The reason the time-varying behaviour was not sufficiently excited was because of a faulty experimental protocol where the input was the main culprit.

I. Introduction

Joint impedance is the dynamic relation between displacements of the joint and the torque response that is generated. Besides being a dynamical relationship, it is also used as a descriptor for the physiology and pathology of muscles and joints (Kearney and Hunter, 1990) as well as biomimetics. For example, in Scholtes et al. (2006) joint impedance was used to assess spasticity in children with cerebral palsy. Currently joint impedance is assessed via tests like the Ashworth scale. These tests manually assess the resistance to fast and slow movements on an ordinary scale. However, these tests are not objective and it is hard to distinguish between neural and tissue components (Scholtes et al., 2006; Pandyan et al., 1999). To distinguish between neural and tissue components is important because it determines the treatment that the patients will receive. Therefore, there is a need for objective measures of joint impedance which can be used for the assessment of neuromuscular diseases.

Previously many Linear Time Invariant (LTI) methods have been used to quantify impedance or its inverse: admittance (Mugge et al., 2010; Schouten, 2004; Van Der Helm et al., 2002; Kearney et al., 1997). However, LTI methods are restricted to time invariant, linear systems around specific operating points. Therefore LTI methods are not able to fully quantify joint impedance. A logical next step in joint impedance identification is to find methods that can overcome these limitations. A solution is offered by ensemble-based methods which require a lot of repetitions and a long recording time, for example Lee and Hogan (2015), but this is not desirable when working with humans due to variability and fatigue. Besides a long recoding time they assume that each repetition has the same time-varying behaviour which is violated for long recording times. A promising class of methods are the Linear Parameter Varying (LPV) methods. LPV systems are systems whose parameters change depending on a Scheduling Variable (SV). Imagine a rocket leaving orbit; the dynamics of a rocket change with its mass, so the amount of fuel can be used as an SV. LPV methods will be able to capture time-varying behaviour as well as describing behaviour over a larger operating range than LTI methods. Another advantage of LPV systems is that they identify a general model which can predict output for novel trajectories of the scheduling variable, unlike the ensemble-based methods.

A promising LPV identification method to use for the estimation of joint impedance is the LPV Predictor

*correspondence to: martijn.kerklaan@gmail.com

Based Subspace IDentification (LPV-PBSID) algorithm by (van Wingerden and Verhaegen, 2009). This method is promising because it is able to deal with closed-loop systems; multiple scheduling variables, for example joint angle and velocity; needs a single measurement to identify a model and can easily be extended to multiple input/output (MIMO) systems. This algorithm has also been used in practise for the identification and control of wind turbines (van Wingerden, 2008), using wind velocity as an SV.

However, state space descriptions are not able to describe improper systems. Joint impedance identification experiments are often done with position perturbations, these experiments lend themselves for estimating systems using the following system description.

$$T(s) = (Is^2 + Bs + K)\theta(s) \tag{1}$$

With $T(s)$ torque, $\theta(s)$ angle, I Inertia, B the viscosity, and K the stiffness. This description is a problem because the LPV-PBSID is a state space method. To overcome this problem, the LPV-PBSID algorithm was combined with the generalized state space system

$$\begin{aligned} \overline{E}x_{k+1} &= \overline{A}x_k + \overline{B}u_k, \\ y_k &= \overline{C}x_k + \overline{D}u_k. \end{aligned}$$

The addition of $\overline{E}$ allows for the description of improper systems. Improper systems do not only occur in biomechanics but also in economical systems and electrical networks (Lewis, 1986). Ideas on how to implement the generalized state space description were taken from Verhaegen (1996).

The rest of the paper will have the following structure; the LPV non-causal identification algorithm is derived in Section II. With the algorithm derived, Section III consists of a simulation study to verify the derived algorithm with a discussion of the result. The algorithm was subsequently applied to real life data of both the ankle and wrist. These results are shown and discussed in Section IV and V. In Section VI criticisms against LPV methods for the identification of joint impedance are discussed. This led to some recommendations in Section VII and lastly the conclusions were drawn in Section VIII.

II. Linear Parameter Varying System Identification

In this section a closed-loop version of the LPV non-causal subspace algorithm is derived. Important to note is that this version was derived after the open-loop version and is therefore not used in the other sections of this paper. Theoretically this version is superior to the open-loop version derived in Appendix B. Unfortunately no time was found to implement and test this version with an appropriate simulation.

i. Generalized State Space system

Given the following LTI generalized state space system in predictor form:

$$\begin{aligned} \overline{E}x_{k+1} &= \overline{A}x_k + \overline{B}u_k + \overline{K}e_k, \\ y_k &= \overline{C}x_k + \overline{D}u_k + e_k \end{aligned}$$

with input $u_k \in \mathbb{R}^r$, output $y_k \in \mathbb{R}^l$, zero mean white innovation sequence $e_k \in \mathbb{R}^l$ and state $x_k \in \mathbb{R}^n$. The matrices $\overline{A}, \overline{B}, \overline{K}, \overline{C}$ and $\overline{D}$ have appropriate dimensions. The predictor form can be rewritten as:

$$\begin{aligned} \overline{E}x_{k+1} &= \overline{A}x_k + \overline{B}u_k + \overline{K}(y_k - \overline{C}x_k - \overline{D}u_k) \\ \Rightarrow \overline{E}x_{k+1} &= (\overline{A} - \overline{K}\,\overline{C})x_k + (\overline{B} - \overline{K}\,\overline{D})u_k + \overline{K}y_k \\ \Rightarrow \overline{E}x_{k+1} &= \tilde{\overline{A}}x_k + \tilde{\overline{B}}u_k + \overline{K}y_k \\ y_k &= \overline{C}x_k + \overline{D}u_k + e_k. \end{aligned}$$

Now, assuming that the matrix pencil $z\overline{E} - \tilde{\overline{A}}$ is regular i.e. there is at least one value of z such that $det(z\overline{E} - \tilde{\overline{A}}) \neq 0$, the generalized state space system in predictor form can be transformed into the so-called Kronecker canonical form (Gerdin, 2004; Verhaegen, 1996) which has the following structure:

$$z\begin{bmatrix} I & 0 \\ 0 & E \end{bmatrix} - \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix}.$$

Subsequently, splitting the non-causal system in a causal and anti-causal part:

$$x^c_{k+1} = Ax^c_k + B^c u_k + K^c y_k \qquad \text{causal part,} \tag{2}$$

$$Ex^{ac}_{k+1} = x^{ac}_k - B^{ac}u_k - K^{ac}y_k \qquad \text{anti-causal part,} \tag{3}$$

$$y_k = \begin{bmatrix} C^c & C^{ac} \end{bmatrix} \begin{bmatrix} x^c_k \\ x^{ac}_k \end{bmatrix} + Du_k + e_k. \tag{4}$$

where E is possibly singular and in some cases nilpotent. E is singular or close to singular because else E would be invertible. If E is invertible, the system would be causal. For a nilpotent matrix, all $eigenvalues = 0$ and $E^h = 0$, where h is an integer which is $h \leq n$, with n the order of the matrix.

Next, the time-varying or LPV case is considered where E, A, B^{ac} and B^{c} were made time-varying:

$$\begin{aligned} x^c_{k+1} &= A_k x^c_k + B^c_k u_k + K^c_k y_k && \text{causal part,} \\ E_k x^{ac}_{k+1} &= x^{ac}_k - B^{ac}_k u_k - K^{ac}_k y_k && \text{anti-causal part,} \end{aligned} \tag{5}$$

$$y_k = \begin{bmatrix} C^c & C^{ac} \end{bmatrix} \begin{bmatrix} x^c_k \\ x^{ac}_k \end{bmatrix} + D u_k + e_k. \tag{6}$$

With affine dependency such that:

$$\begin{aligned} E_k &= \sum_{i=1}^{m} \mu_k^{(i)} E^{(i)} \\ \mu_k &= \begin{bmatrix} 1 & \mu_k^{(2)} & \dots & \mu_k^{(m)} \end{bmatrix}^T \end{aligned} \tag{7}$$

with $\mu_k^{(i)} \in \mathbb{R}$ and m local models. By splitting the generalized state space system into a causal and anti-causal part a PBSID identification scheme can be derived for the anti-causal part and merged with the causal part which was described in van Wingerden and Verhaegen (2009) to create the non-causal identification scheme.

ii. Anti-Causal LPV identification

The derivation of the causal LPV part using PBSID is already described in van Wingerden and Verhaegen (2009). The steps taken in this derivation were used as an inspiration to derive the estimation of the anti-causal part. Assume the following system description where the superscript ac is dropped for clarity.

$$\begin{aligned} E_k x_{k+1} &= x_k - B_k u_k - K_k y_k \\ \Rightarrow x_k &= E_k x_{k+1} + B_k u_k + K_k y_k \\ \Rightarrow x_k &= E_k x_{k+1} + \breve{B}_k z_k \\ y_k &= C x_k + D u_k + e_k \end{aligned} \tag{8}$$

where $x_k \in \mathbb{R}^n$, $u_k \in \mathbb{R}^r$ and $y_k \in \mathbb{R}^l$ are the state, input and output, with $k = \{1, 2, \dots, N\}$. Here z_k is the stacked vector $z = \begin{bmatrix} u_k^T & y_k^T \end{bmatrix}^T$. The matrices $E_k \in \mathbb{R}^{n\times n}$, $B_k \in \mathbb{R}^{n\times r}$, $K_k \in \mathbb{R}^{n\times l}$, $C \in \mathbb{R}^{l\times n}$ and $D \in \mathbb{R}^{l\times r}$ are the system, input, observer, output and feed through matrices. Where $\breve{B}_k = \begin{bmatrix} B_k & K_k \end{bmatrix}$ is the packed input observer matrix. It is important to note that now $E_k x_{k+1}$ is defined instead of $E_{k+1} x_{k+1}$ which was done in the open-loop case as can be seen in Appendix B. Defining $E_k x_{k+1}$ as such allows for a better definition of $P_{p|k}$, which was less elegantly defined in the previous derivation.

The time-varying matrices are dependent in an affine way with m local models:

$$E_k = \sum_{i=1}^{m} \mu_k^{(i)} E^{(i)} \tag{9}$$

$$\mu_k = \begin{bmatrix} 1 & \mu_k^{(2)} & \dots & \mu_k^{(m)} \end{bmatrix}^T \tag{10}$$

with $\mu_k^{(i)} \in \mathbb{R}$. It is important to know that the matrix E_k is possibly singular and in some cases nilpotent.

iii. Assumptions and notations

First the transition matrix is defined which is slightly different from the one used for the causal LPV-PBSID:

$$\phi_{f,k} = E_k E_{k+1} \dots E_{k+f-1} \tag{11}$$

with a future window f and the following stacked input output vector:

$$\bar{z}_k^f = \begin{bmatrix} z_k \\ \vdots \\ z_{k+f-1} \end{bmatrix}.$$

The state sequence is assumed to be:

$$X = \begin{bmatrix} x_k, & \dots, & x_{N-f} \end{bmatrix}$$

and has full row rank. Notice how the anti-causal system sequence is again slightly different from the one used in the causal estimation. The matrix:

$$\Gamma^f = \begin{bmatrix} C(E^{(1)})^{f-1} \\ \vdots \\ CE^{(1)} \\ C \end{bmatrix} \tag{12}$$

is the extended observability matrix of the first local model and has full rank. Note that no assumptions have been made on the correlation of the input and the noise and therefore this description is suitable for closed-loop identification.

iv. Factorization

In this subsection the time-varying extended controllability matrix, $\overline{\mathcal{K}}_k^f$, is factorized in the unknown state matrices and the known scheduling sequence. The time-varying extended controllability matrix is given by:

$$\overline{\mathcal{K}}_k^f = \begin{bmatrix} \phi_{f-1,k}\breve{B}_{k+f-1}, & \dots, & \phi_{1,k}\breve{B}_{k+1}, & \breve{B}_k \end{bmatrix},$$

with $\overline{\mathcal{K}}_k^f \in \mathbb{R}^{n\times(r+l)f}$.

Define $\mathcal{K}^f$: First $\mathcal{K}^f$ is defined, the extended controllability. This matrix can be seen as the time-invariant controllability matrix of the LPV system or as the matrix containing all used combinations of affine matrices. In the LTI case the controllability matrix is defined as

$$\mathcal{K} = \begin{bmatrix} E^{n-1}B, & \dots, & B \end{bmatrix}.$$

Since the LPV E and B matrices contain m local models, the product $EB = (E^{(1)} + E^{(2)} + \dots + E^{(m)})(B^{(1)} + B^{(2)} + \dots + B^{(m)})$ results in m^2 combinations and the product E^2B in m^3 combinations, etc. All these combinations are defined in $\mathcal{K}^f$. Separating the combinations from the weights given to them by the scheduling variables is an important step in the algorithm. This procedure is explained in more detail in Appendix A. To make all possible combinations, the following definitions are needed:

$$\mathcal{L}_j = \begin{bmatrix} E^{(1)}\mathcal{L}_{j-1}, & \dots, & E^{(m)}\mathcal{L}_{j-1} \end{bmatrix},$$

with

$$\mathcal{L}_1 = \begin{bmatrix} \breve{B}^{(1)}, & \dots, & \breve{B}^{(m)} \end{bmatrix}$$

and $\mathcal{L}_j \in \mathbb{R}^{n \times (r+l)m^j}$.

With these variables the extended controllability matrix can be defined:

$$\mathcal{K}^f = \begin{bmatrix} \mathcal{L}_1, & \dots, & \mathcal{L}_f \end{bmatrix}, \qquad \in \mathbb{R}^{n \times \tilde{q}}$$

with $\tilde{q} = (r+l)\sum_{j=1}^{f} m^j$. Notice how $\mathcal{K}^f$ is build up from 1 to f instead of from f to 1 as was done in the causal case.

Define $P_{j|k}$: Next the known time-varying part is defined, this gives weight to the combinations of state matrices in $\mathcal{L}_j$.

$$P_{j|k} = \mu_k \otimes \dots \otimes \mu_{k+j-1} \otimes I_{r+l}$$

with $P_{j|k} \in \mathbb{R}^{m^f(r+l) \times (r+l)}$ and $\otimes$ denoting the Kronecker product (Brewer, 1978).

Define N_k^f: The weight matrices $P_{j|k}$ can be combined in a bigger matrix which gives the correct weights to the different combinations made in $\mathcal{K}^f$

$$N_k^f = \begin{bmatrix} P_{1|k} & & & 0 \\ & P_{2|k} & & \\ & & \ddots & \\ 0 & & & P_{f|k} \end{bmatrix}$$

with $N_k^f \in \mathbb{R}^{\tilde{q} \times f(r+l)}$ such that.

$$\overline{\mathcal{K}}_k^f = \mathcal{K}^f N_k^f$$

Proof The proof is shown Appendix A. It follows from extending x_k and capturing the structure that appears.

v. Closed-loop identification

The first objective in order to identify the closed-loop system is to reconstruct the state x_k:

$$x_k = \phi_{f,k} x_{k+f} + \mathcal{K}^f N_k^f \bar{z}_k^f.$$

If f is chosen big enough such that $\phi_{f,k} \approx 0$ or in the case of a nilpotent matrix we can guarantee $\phi_{f,k} = 0$ given $f \geq h$. With a big enough f a definition which approximates the state can be given by:

$$x_k \approx \mathcal{K}^f N_k^f \bar{z}_k^f. \tag{13}$$

In the case of a nilpotent E_k it can be guaranteed that $x_k = \mathcal{K}^f N_k^f \bar{z}_k^f$. With a definition of the state the input-output behaviour can be approximated.

$$y_k \approx C\mathcal{K}^f N_k^f \bar{z}_k^f + Du_k + e_k \tag{14}$$

Next, the stacked matrices can be defined:

$$U = \begin{bmatrix} u_1, & \dots, & u_{N-f+1} \end{bmatrix}, \tag{15}$$

$$Y = \begin{bmatrix} y_1, & \dots, & y_{N-f+1} \end{bmatrix}, \tag{16}$$

$$Z = \begin{bmatrix} N_1^f \bar{z}_1^f, & \dots, & N_{N-f+1}^f \bar{z}_{N-f+1}^f \end{bmatrix}, \tag{17}$$

Notice how Y starts at the first sample but cannot estimate the last samples because future values of the input are required. If $[Z^T, \quad U^T]^T$ has full row rank, the matrices $C\mathcal{K}^f$ and D can be estimated by solving the following minimization problem:

$$\min_{C,\mathcal{K}^f,D} ||Y - C\mathcal{K}^f Z - DU||_F^2, \tag{18}$$

where $||\dots||_F$ is the Frobenius norm (Golub and Loan, 1996).

vi. Observability matrix times controllability matrix

Next the approximation of the product of the extended observability matrix and controllability matrix can be constructed. Because of the slightly different notations this will become a lower-block triangular matrix. This was done to combine it with the upper-block triangular matrix of the causal case. In the end this was only possible for the LTI case because $Z^{ac} \neq Z^{ac}$ due to the scheduling variable appearing in a different order.

$$\begin{aligned}\Gamma^f \mathcal{K}^f &= \begin{bmatrix} C(E^{(1)})^{f-1} \\ \vdots \\ CE^{(1)} \\ C \end{bmatrix} \begin{bmatrix} \mathcal{L}_1, & \dots, & \mathcal{L}_f \end{bmatrix} \\ &\approx \begin{bmatrix} C(E^{(1)})^{f-1}\mathcal{L}_1 & 0 & & 0 \\ \vdots & \ddots & & \\ CE^{(1)}\mathcal{L}_1 & \dots & CE^{(1)}\mathcal{L}_{f-1} & 0 \\ C\mathcal{L}_1 & \dots & C\mathcal{L}_{f-1} & C\mathcal{L}_f \end{bmatrix}\end{aligned} \tag{19}$$

The zeros appear if f is chosen big enough such that $\phi_{f,k} \approx 0$. In the case that E_k is nilpotent the block triangular matrix can be guaranteed because $\phi_{f,k} = 0$.

Using the following relations:

$$\begin{aligned} C\mathcal{K}^f &= \begin{bmatrix} C\mathcal{L}_1, & \dots, & C\mathcal{L}_f \end{bmatrix}, \\ C\mathcal{L}_f &= \begin{bmatrix} CE^{(1)}\mathcal{L}_{f-1}, & \dots, & CE^{(m)}\mathcal{L}_{f-1} \end{bmatrix}, \end{aligned} \tag{20}$$

and the estimate of $C\mathcal{K}^f$ from the minimization problem enables the construction of $\Gamma^f \mathcal{K}^f$. Subsequently, $\Gamma^f \mathcal{K}^f Z$ can be constructed which is by definition $\Gamma^f X$. Now, assuming both X and Γ^f have full rank and f is big enough to satisfy the assumption $\phi_{f,k} \approx 0$, the state sequence can be estimated up to a similarity transform using the Singular Value Decomposition (SVD):

$$\widehat{\Gamma^f \mathcal{K}^f Z} = \begin{bmatrix} \mathcal{U} & \mathcal{U}_\perp \end{bmatrix} \begin{bmatrix} \Sigma_n & 0 \\ 0 & \Sigma \end{bmatrix} \begin{bmatrix} V \\ V_\perp \end{bmatrix}. \tag{21}$$

Where Σ_n is the diagonal corresponding to the n largest singular values and V corresponds to the row space of those singular values. With this the estimate of the state is given by:

$$\hat{X} = \Sigma_n V \tag{22}$$

With the state found (8) can be used to estimate the unknown system matrices.

Note that this method suffers from the curse of dimensionality. This refers to the fact that the numbers of rows in Z grows exponentially with f. Recommendations have been made to overcome this problem, see Section VII.

LPV-AC-PBSID Algorithm The algorithm can be summarized as follows:

1. Build Z, Y and U using ((15)-(17)) to find $\mathcal{CK}^f$ and D.
2. Solve the minimization problem given in (18).
3. Use relations (20) and Z to construct $\Gamma^f \mathcal{K}^f Z$.
4. Use SVD to find an estimate of X
5. With the state estimate the system matrices can be found up to a similarity transform using (8)

vii. Combining anti-causal and causal

To combine both the causal and anti-causal description and create the Linear Parameter Varying Non-Causal Predictor Based Subspace IDentification (LPV-NC-PBSID) algorithm, the following stacked matrices from the input, output and SV are made:

$$U = \begin{bmatrix} u_p, & \dots, & u_{N-f+1} \end{bmatrix}, \tag{23}$$

$$Y = \begin{bmatrix} y_p, & \dots, & y_{N-f+1} \end{bmatrix}, \tag{24}$$

$$Z^{ac} = \begin{bmatrix} N^f_{ac,p}\bar{z}^f_p, & \dots, & N^f_{ac,N-f+1}\bar{z}^f_{N-f+1} \end{bmatrix}, \tag{25}$$

$$Z^c = \begin{bmatrix} N^p_{c,1}\bar{z}^p_1, & \dots, & N^p_{c,N-p-f+1}\bar{z}^p_{N-p-f+1} \end{bmatrix}, \tag{26}$$

where $N^p_{c,k}$ and $N^f_{ac,k}$ are the weighting matrices for the causal and anti-causal part respectively. Note that the causal algorithm uses the same definition for $\bar{z}^{p/f}_k$. With the stacked matrices defined the following minimization problem can be solved

$$\min_{C,\mathcal{K}^p,D} ||Y - \begin{bmatrix} C^c & C^{ac} \end{bmatrix} \begin{bmatrix} \mathcal{K}^p_c & 0 \\ 0 & \mathcal{K}^f_{ac} \end{bmatrix} \begin{bmatrix} Z^c \\ Z^{ac} \end{bmatrix} - DU||^2_F. \tag{27}$$

The solution of this problem is an estimate of $\begin{bmatrix} C^c\mathcal{K}^p_c & C^{ac}\mathcal{K}^f_{ac} & D \end{bmatrix}$ with $C^c\mathcal{K}^p_c \in \mathbb{R}^{l\times\tilde{q}}$, $C^{ac}\mathcal{K}^f_{ac} \in \mathbb{R}^{l\times\tilde{q}}$ and $D \in \mathbb{R}^{l\times r}$. This solution can be used to find estimates of the causal and anti-causal states. Notice how the causal part has a past window size p and the anti-causal part has a future window size f because their respective state estimations are either dependent on past or future inputs/states.

LPV-NC-PBSID Algorithm The algorithm can be summarized as follows:

1. Build the stacked matrices using ((23)-(26)).

2. Solve the minimization problem given in (27) to find $\begin{bmatrix} C^c\mathcal{K}_c^p & C^{ac}\mathcal{K}_{ac}^f & D \end{bmatrix}$.
3. Use relations (20), the causal equivalent from (van Wingerden and Verhaegen, 2009), Z^{ac} and Z^c to construct $\Gamma^f\mathcal{K}^f Z^{ac}$ and $\Gamma^p\mathcal{K}^p Z^c$.
4. Use SVD to find an estimate of X^{ac} and X^c
5. With the state estimate the system matrices can be found up to a similarity transform using (5)

III. Simulation study

In this section the open-loop LPV-NC-PBSID algorithm derived in Appendix B, which is an open-loop version of the algorithm in Section II, was used in the estimation of a mass spring damper model with time-varying stiffness to show its viability.

i. Model

For the simulation study a mass spring damper model with time-varying stiffness was used. The simulation has the same model, input and noise as used in Ludvig and Perreault (2012). It can be represented by the following equation:

$$T(s) = H(s,t)^{-1}\theta(s)$$
$$T(s) = (Is^2 + Bs + K(t))\theta(s)$$

with $H(s,t)$ the admittance model from Figure 1, I the inertia equal to $I = 0.13\,\mathrm{kgm}^2$, B the viscosity equal to $B = 2.2\,\mathrm{N\,m\,s\,rad^{-1}}$ and $K(t)$ the time-varying stiffness equal to $K(t) = 100 + 50sin(2\pi t)\mathrm{N\,m\,rad^{-1}}$. Because the simulation of improper systems uses derivatives, which is undesirable, it was done by simulating the proper system seen in Figure 1. By using x_{ankle} and $T_{measure}$ as input and output an improper system is still estimated.

This system was simulated in Simulink, Matlab 2016a, using a fixed time step of 4×10^{-4} s. The input x_{per} is a PseudoRandom Binary Sequence (PRBS) which is rate limited to ± 0.5 rad/s and has an amplitude of 0.03 rad. Subsequently the input is filtered by an 8th order Bessel filter with a cutoff frequency at 8Hz before being fed into the system. This type of input is not only used in Ludvig and Perreault (2012), but also in other time invariant and time variant joint impedance identification studies (Kearney et al., 1997; Mirbagheri et al., 2000; Ludvig et al., 2011). The disturbance from Figure 1 is

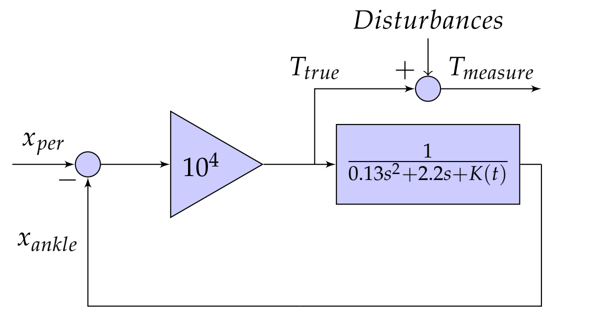


Figure 1: *The human connected in closed-loop with the machine, represented by a stiff gain of* 10^4

white noise filtered with a 3[th] order Bessel low-pass filter with a cutoff frequency of 3Hz. This filtered noise represents the voluntary torque the participant would apply. The noise was scaled such that the signal to noise ratio(SNR) would be 5dB. Where the SNR was calculated using the root mean square of the signal and noise, like this:

$$SNR = 20\log_{10}\left(\frac{\sqrt{mean(T_{true}^2)}}{\sqrt{mean(Disturbance^2)}}\right)$$

ii. Estimation and Analysis

The system was simulated a 100 times for 30s to get an estimate of the error bound, each simulation had a newly generated noise and perturbation signal. Subsequently, the data was decimated from 2.5kHz to 100Hz. For the estimation, the x_{ankle} and $T_{measure}$ were used as input and output respectively. The time-varying signal $\sin(2\pi t)$ was used as SV. All 30s of data was used for the estimation of the system. The algorithm assumed that A_k,E_k,B_k^c and B_k^{ac} were time-varying. A past and future window of $p_c = f_{ac} = 2$ and an order selection of $n_c = 1$ and $n_{ac} = 2$ gave the best results. The quality of the estimation was quantified by looking at the Variance Accounted For (VAF). The VAF was calculated as follows

$$VAF(T_{measure}, \hat{T}) =$$
$$max\left(0, \left(1 - \frac{var(T_{measure} - \hat{T})}{var(T_{measure})}\right)\right) \times 100\%$$

In this case the measured torque and the predicted torque, $\hat{T}$, were used as an example. Besides calculating the VAF, the parameters of the system were also estimated.

The parameters were estimated for 50 points along 1 period of the scheduling variable. At each point along the SV the LTI state space model was calculated and simulated with 1×10^6 points of white noise. From this input and output a frequency response function was generated and a 2^{nd} order transfer function was fitted to the data using all frequencies up to 20Hz. From this 2^{nd} order transfer function the parameters were extracted. This was done with the help of the inbuilt Matlab functions *tfestimate.m* and *tfest.m*. For all the parameters the range of the standard deviation is given and in the case of the damping and inertia the mean is given as well. How these Matlab commands were applied can be seen in Appendix H.

iii. Results

The Figures 2 and 3 show the input and output signal used for identification and the predicted output.

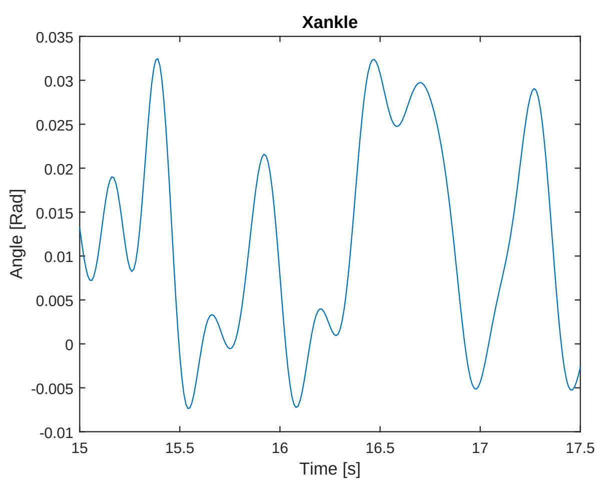


Figure 2: *Example of x_{ankle} used for identification*

These input and outputs of every simulation were used to predict the output. This prediction resulted in a $VAF(T_{true}, \hat{T})$, $VAF(T_{measure}, \hat{T})$ and $VAF(T_{measure}, T_{true})$. The result is shown in Table 1.

$VAF(X,Y)$	$(T_{measure}, T_{true})$	$(T_{measure}, \hat{T})$	$(T_{true}, \hat{T})$
LPV	68.4(1.45)%	68,2(1.52)%	99.1(0.67)%

Table 1: *The average VAF with standard deviation comparing the measured with true, measured with predicted and true with predicted. It can be seen that the LPV-NC-PBSID has great noise rejection capabilities.*

Besides estimating the output torque, the parameters of the model were also estimated. On average the 2nd order system fitted on the found model had a 98.2%(0.53) fit percentage. A plot showing the fit percentage and standard deviation for each trial can be seen in Appendix C. Below in Figure 4 an example of two FRF estimates is shown. These FRF's are from the same simulation but they made for different values of the SV.

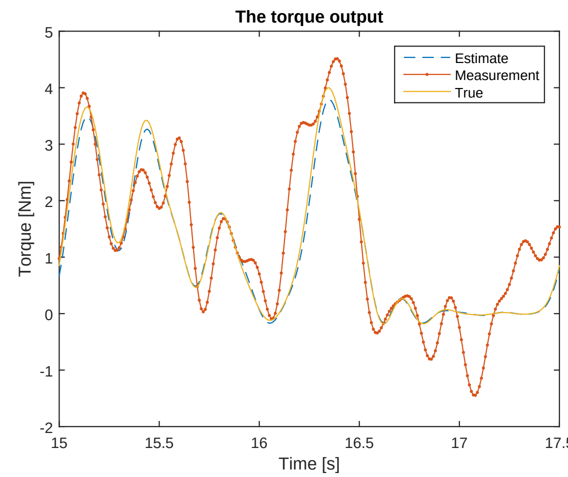


Figure 3: *Example of the measured torque output used for identification, the predicted and true torque output*

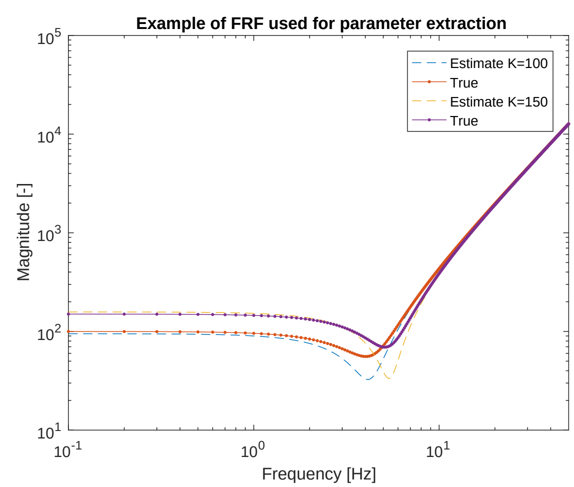


Figure 4: *Notice how the stiffness and inertia are fitted well but the damping is underestimated*

The result of the stiffness, damping and inertia estimate can be seen in the Figures 5, 6, 7.

The damping and inertia estimates were also time-varying even if their true parameters are not. The mean and range of the standard deviation is given for each estimate in Table 2.

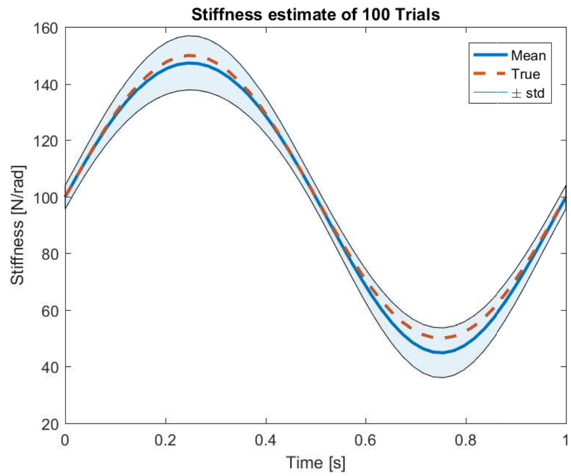


Figure 5: *The time-varying stiffness estimate of a 100 trials. It can be seen that the mean almost follows the true stiffness and that the true stiffness is within the standard deviation.*

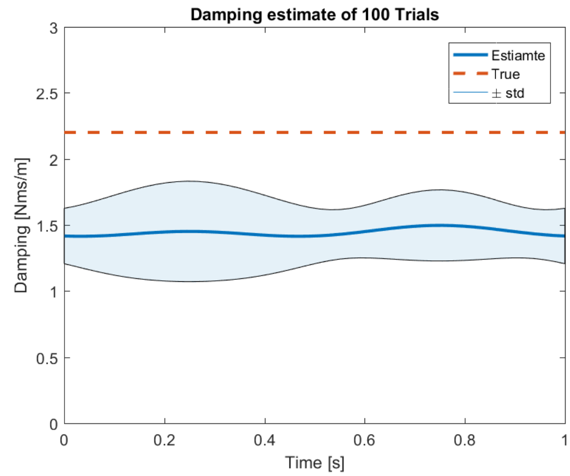


Figure 6: *The time-varying damping estimate of a 100 trials. Notice how the damping estimate is time-varying even though the true damping is not.*

	Mean	**Range std**
Stiffness	97.97 $\mathrm{N\,m\,rad^{-1}}$	4.15-9.61$\mathrm{N\,m\,rad^{-1}}$
Damping	1.449 $\mathrm{N\,m\,s\,rad^{-1}}$	0.187-0.380$\mathrm{N\,m\,s\,rad^{-1}}$
Inertia	0.135 $\mathrm{kgm^2}$	0.0023-0.0048$\mathrm{kgm^2}$

Table 2: *The mean and standard deviation of all the retrieved parameters*

Lastly, an example is given of the prediction error in the frequency domain, $T_{measure} - \hat{T}$. This gives an idea if all frequencies were equally weighted in the estimation of the model and where the noise is most present.

iv. Discussion

The first results obtained from the simulation were the VAFs shown in Table 1. From this table it can be seen that the predicted torque accurately represents the true torque and has roughly the same VAF with the measured torque as the true torque. Together this indicates the algorithm accurately estimated the underlying model even with a 5dB SNR.

The second result obtained by the simulation was the parameter estimations. As can be seen in Figures 5, 6 and 7, the stiffness and inertia were accurately estimated while the damping is biased. The most probable reason why the damping is biased, is because the disturbance is a coloured noise. The current system description, (28) in Appendix B, is not able to properly deal with coloured noise. This means that even though the bias is only clearly present in the damping it could also be in the stiffness and inertia estimates. This problem can be solved by using the predictor form of the state space description which was used in Section II and (van Wingerden and Verhaegen, 2009) which is able to deal with coloured noise.

Another reason why the damping could be biased is because the frequencies needed for this estimation are attenuated and therefore harder to distinguish from the noise. Where a proper system would have a resonance peak at its cutoff frequency, an improper system has an anti-resonance valley at this frequency. The valley can be seen in Figure 9 which appears to be an improper system at low frequencies. This anti-resonance valley causes the frequencies which give information about the damping to be attenuated and harder to distinguish from the noise. However, this explanation seems unlikely; with the addition of noise one would expect the valley to become less pronounced. This would lead to a bias of the damping in such a way that there would be more damping, not less.

What might be surprising is that the damping and inertia were estimated as time-varying parameters as well. The reason why this happens is because the simulated system is a transfer function while the estimated model is a state space system. In LTI systems switching between system description is not a problem, while it is in the LPV case. For a more elaborated discussion on this subject the reader is referred to (Tóth et al., 2012; Tóth, 2007). In the case of this simulation the transform between the LPV transfer function and LPV state space

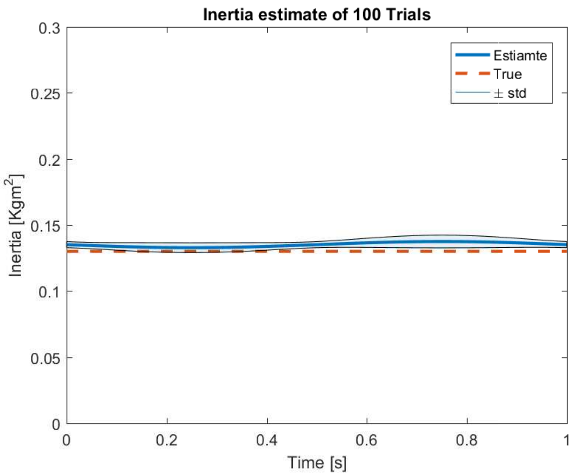


Figure 7: *The time-varying inertia estimate of a 100 trials. Notice how the damping estimate is time-varying even though the true damping is not.*

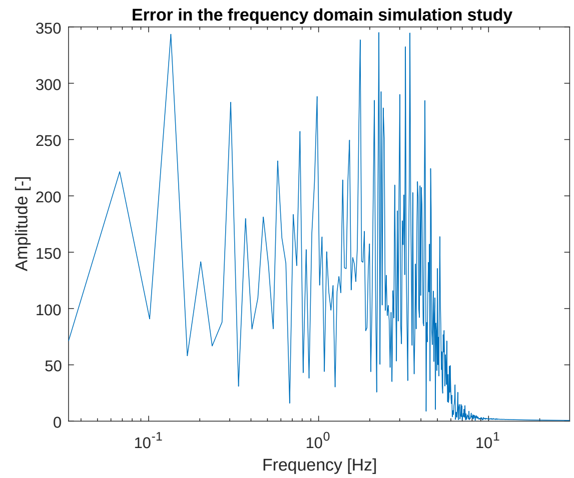


Figure 8: *The prediction error in the frequency domain of a single run.*

system is determined analytically in Appendix G. This derivation shows that the LPV transfer function cannot be 100% accurately represented by the LPV state space system. The discrepancy between these descriptions also explains why all parameters are estimated time-varying and could also explain part of the bias.

Another reason that the parameters vary over time is because the combined causal anti-causal system is a 3$^{\text{th}}$ order system. Looking at (6) it can be seen that it is essentially $y_k = y_c + y_{ac}$, this is clearly shown in Appendix E, where the analytic solution of an LTI-NC state space system to a transfer function is given. Adding systems results in $n = n_c + n_{ac}$. In order to illustrate more clearly that summing lower order systems creates a higher order system, an example with two first order systems is shown below.

$$\frac{s-a}{s-b} + \frac{s-c}{s-d} = \frac{(s-a)(s-c)}{(s-b)(s-d)}$$

This also explains the choice for $p_c = f_{ac} = 2$ as estimation settings. Using these window sizes might seem low for the estimation of a second order system because p needs to be $\geq n$. By using $p_c = f_{ac} = 2$ actually a $p = 4$ has been used which can accurately estimate a 3$^{\text{th}}$ order system.

In the introduction it is stated that the LPV-NC-PBSID algoritm was developed to deal with improper systems. One might be tempted to estimate a system from x_{per} to $T_{measure}$ to solve this problem, but this will not help. The system only becomes proper at a high frequency, as can be seen in Figure 9. So using low frequencies for x_{per} will still result in an improper system. If one would add high frequency content to x_{per} to estimate a proper system, the high frequency content would dominate the output and thus drown out the time-varying low frequency part. Simply put, an LTI method could be used to correctly predict the output because the time invariant high frequency content is dominant. Even if a

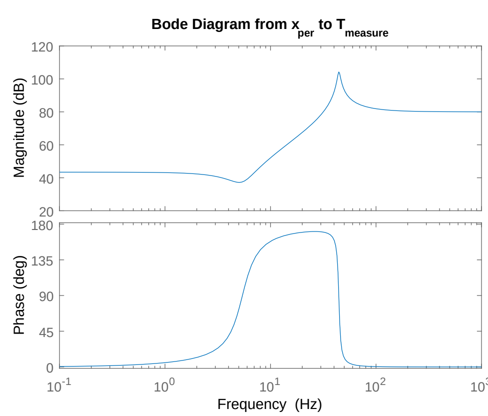


Figure 9: *The bode diagram of the system shown in Figure 1 from x_{per} to $T_{measure}$ with $K = 150\,\mathrm{Nm/rad}$*

proper input could be designed to estimate the system from x_{per} to $T_{measure}$ the stiffness estimate would be biased because the machine is not infinitely stiff. This becomes clear when one looks at the transfer function of the bode plot shown in Figure 1.

$$P(s) = \frac{10^4 H(s)}{1+KH(s)} = \frac{10^4(Is^2+Bs+K)}{Is^2+Bs+K+10^4},$$
$$P(0) = \frac{10^4 K}{K+10^4} = \frac{K}{K/10^4+1}.$$

Theoretically the machine would be infinitely stiff, this would give the correct stiffness. However, in practise it is not infinitely stiff and thus the estimated stiffness will be biased. Therefore, x_{ankle} and $T_{measure}$ were used for the identification.

Using x_{ankle} and $T_{measure}$, one might be tempted to estimate the system from $T_{measure}$ to x_{ankle} and estimate a proper system in such a way. This approach is problematic because it introduces input noise. Input noise biases the estimate of the system and is hard to deal with. Since there already is a lot of noise present with the estimation of joint impedance this is not recommended.

In Figure 8 the prediction error in the frequency domain can be seen. The error is present up to around 3Hz, as was expected because 3Hz is the cut-off frequency of the disturbance. The error is present there because the model did not fit the noise.

Lastly, the model used in this simulation has a weakness. The weakness is that the stiffness modulation and the torque generated are represented as separate entities, while they are related either as reflexive or as visco-elastic properties of the muscle (Nichols and Houk, 1976; Zhang and Rymer, 1997). In practise the scheduling function would have to be retrieved from the measurements. This means that the scheduling variable is never available without noise. The search for the right scheduling variable was not taken into account in this simulation.

IV. Analysis of B1 ROBIN Data

In 2013 a little over a thousand experiments were conducted, these were evenly split between 15 healthy elderly subjects. 129 of these experiments were called B1 Torque Control experiments. These experiments were designed to elicit time-varying behaviour in the ankle during a force task with position perturbations. However, the behaviour was not time-varying, as will be shown in this section.

i. B1 Protocol

The B1 experiment was a force task experiment with a position perturbation. The experiment was performed by fifteen healthy elderly, aged 64.0(10.2) years, 8 male. The participant needed to apply a torque with their ankle following a chirp signal from 0.02-0.25Hz. These experiments have a duration ranging from 45s for the first 3 subjects to 90s for the other 12 subjects. The torque range differs within subjects, but on average(standard deviation) it was 8.87(3.45)Nm Planarflexion and 8.71(3.72)Nm Dorsiflexion, which was 25% of their Maximum Voluntary Contraction (MVC).

Of the 129 B1 experiments, roughly half was done with 70° and the other half with 20° knee flexion, as can be seen in Figure 10. The experiments were done at two measurement moments, where the second measurement was always three weeks after the first. During the first measurement moment the participants did the experiment twice but with different experimental supervisors. The initial position of the motor was such that the foot and lower leg would be at a 90° angle with each other.

Figure 10: *Measurement setup showing a person with either the 70° or 20° flexion of the knee. This figure is copied from Sloot et al. (2015).*

The torque that the participant applied was filtered with a 2nd order Butterworth filter with a cutoff frequency of 1.25Hz and visualized. The applied torque was displayed together with the torque trajectory the participant needed to follow, in this case a chirp. This process has been schematically represented in Figure 11. During the task the participants were perturbed with a low-pass filtered multisine. The multisine consisted of a crested sum of sines with frequencies from 1-15Hz in steps of 0.1Hz. The signal was filtered with a second order low-pass filter with a damping ratio of $\frac{1}{2}\sqrt{2}$ and a natural frequency of 10Hz. The perturbation had an amplitude of ± 0.015 rad and was applied by the motor. The data was recorded with a sampling rate of $f_s = 1024$Hz.

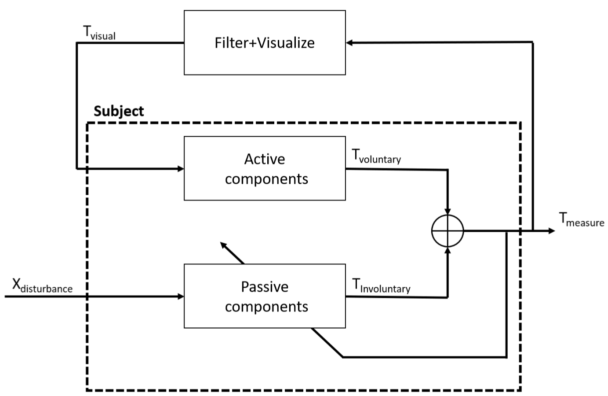


Figure 11: *Experimental Setup of the B1 experiment, notice how the response is due to active compnents triggered by following the trajectory and passive components triggered by the pertubation signal*

ii. Analysis with LPV, LTI and FRF

To analyse the data three methods were used, the LPV-NC-PBSID, LTI-NC-PBSID and an LTI-FRF method. The VAF was used to asses the quality of each estimation. To asses whether the results of the different methods varied between each other, statistical tests were used. First the Kolmogorov-Smirnov test was used to asses normality. Subsequently, the parametric or non-parametric correlation and one way ANOVA were used, as well as ad hoc tests that were needed.

For all methods the measured data from $t = 5s$ until $t = 35s$ was used in case of the $45s$ experiments and the data from $t = 10s$ until $t = 70s$ was used in the case of the $90s$ experiments. From all measured signals the mean was removed.

For the LPV-NC-PBSID and LTI-NC-PBSID the measured position and the measured torque were used as input and output respectively. Both signals were decimated from 1024Hz to 128Hz. Subsequently, both signals were bandpass filtered with a 6th order Butterworth filter with cutoff frequencies at 2Hz and 40Hz. In Figure 11 it can also be seen how the measured torque is partly active and partly passive the bandpass filter was used to separate these responses. For the LPV method the measured torque was low-pass filtered (2nd order Butterworth, 0.25Hz) and used as the SV. This filtering approach was also applied in Van Eesbeek et al. (2013) and aims to separate voluntary and involuntary torques. For both the LPV-NC and LTI-NC a window size of $p_c = f_{ac} = 2$ was used and $n_c = n_{ac} = 2$, furthermore both algorithms assumed $D = 0$. In the LPV case, it was assumed that A_k, E_k, B_k^c and B_k^{ac} were all time-varying.

The LTI-FRF was made by estimating a Frequency Response Function (FRF) using the unfiltered raw measured torque and the measured position as input and output respectively. The FRF from 1Hz-15Hz with $\Delta 0.1 Hz$ was used to fit a second order transfer function using the Matlab function *tfest.m*. The torque was used as input because Matlab is not able to simulate improper transfer functions. The position was estimated from the torque and the VAF was calculated accordingly. The FRF was calculated using the *tfestimate.m* function of Matlab. How this was implemented and what settings were used can be seen in Appendix H.

iii. Results

In Figures 12, 13 the input, output and predicted output are shown.

The VAF was calculated for each experiment and method; the results are displayed in Figure 14. The Kologorov-Smirnov test was significant for the LPV results (D(129)= 1, p<0.05), significant for LTI-NC (D(129) = 1, p<0.05) and significant for the LTI-FRF (D(129) = 0.99, p<0.05). Because all the sets were non-normal, a Kruskal-Wallis test was used to asses if there is any differences between the groups. The Kruskal-Wallis test showed that there was no significant differences between groups (H(2.08), p=0.35) therefore no post hoc tests were conducted.

A Pearson's correlation test was used because it does not assume normality and VAF is on a rational scale. There was a strong correlation found between LPV-NC and LTI-NC of (r(127) = 0.995, p<0.05), as well as be-

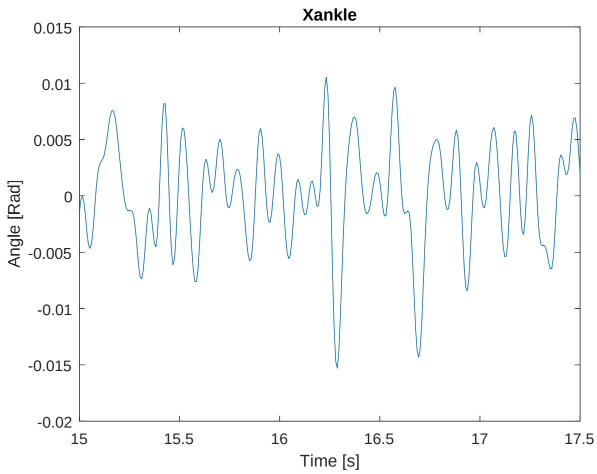


Figure 12: *Example of an input, named x_{ankle}, used for identification of the B1 experiment*

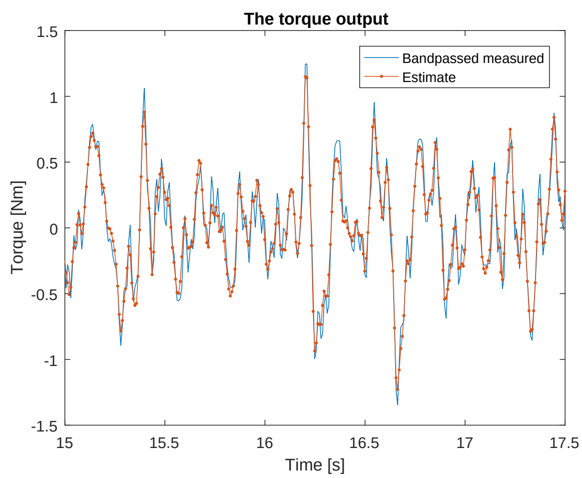


Figure 13: *Example of the measured and bandpass filtered output torque used for identification and the prediction*

tween LPV-NC and LTI-FRF (r(127) = 0.869, p<0.05), and between LTI-NC and LTI-FRF (r(127) = 0.8714, p<0.05).

In Figure 15 an example of the second order fit to the FRF is shown, using the raw input and output data.

Figure 16 shows an example of the prediction error in the frequency domain. This is the prediction error of the first trial of the first subject. This trial had a VAF of 83.2%.

Graphs displaying the VAF of all methods for each trial are shown in Appendix F. Besides the healthy elderly analysed in this section, Appendix F also contains similar graphs for elderly with a stroke. The elderly with a stroke are split into two groups: CVA and stroke. Stroke

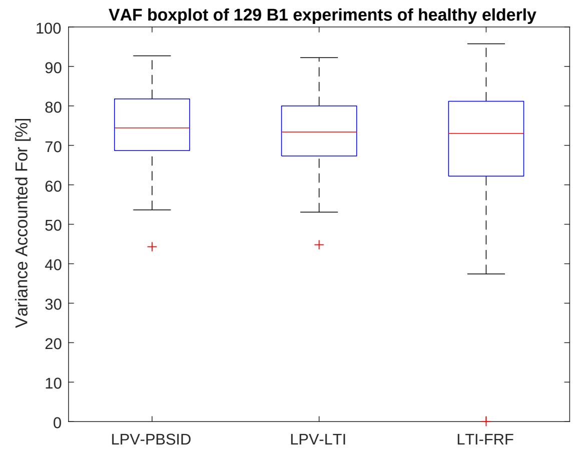


Figure 14: *Box plot of the VAF for each method. The medians of each box were LPV-NC 74%, LTI-NC 73% and LTI-FRF 73%*

and CVA are groups of different elderly who all had the same condition, stroke. Besides graphs of the B1 experiment, the appendix also contains the same graphs for all three groups of the C1 position task experiment. The C1 experiment is similar to the B1 experiment but with position trajectory and torque perturbations. The results of the groups Stroke and CVA as well as the C1 experiments are not analysed in this report.

iv. Discussion

One of the first observations is that the whiskers of the boxplot from Figure 14 cover a large range. The reason why the VAF has such a large variance is mostly likely because one setting was used for all experiments. The window size, order of the system, filtering and amount of the data used for identification: all these variables can have a big impact on the variance of the VAF. Another possibility is that, since all experiments were treated the same, possible faulty data was also used for identification.

To show the influence of the time-varying behaviour, both the LTI and LPV state space method had the same identification settings (window size, order, etc.). The LTI-FRF method was added to determine if any differences between LTI- and LPV-NC-PBSID would be due to an LTI method or because of the PBSID algorithm. The results show that the VAFs of all methods differ very little from each other and are strongly correlated. It can therefore be concluded that the experiment

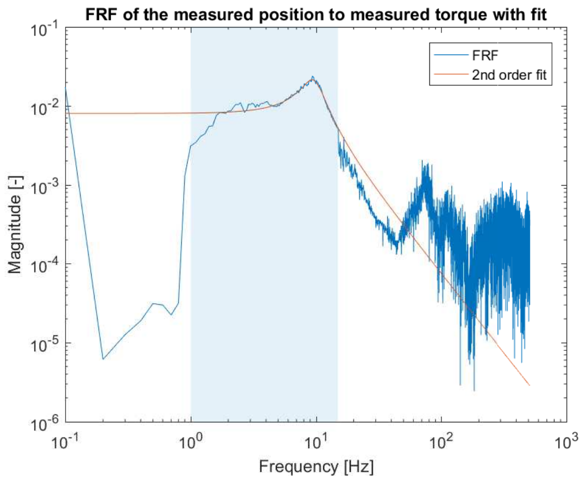


Figure 15: *An example of an FRF with the LTI second order fit from the second subject, during the second measurement moment with 70° flexion. This 2nd order fit had a VAF of 94% with x_{ankle}. The area between 1-15Hz is marked*

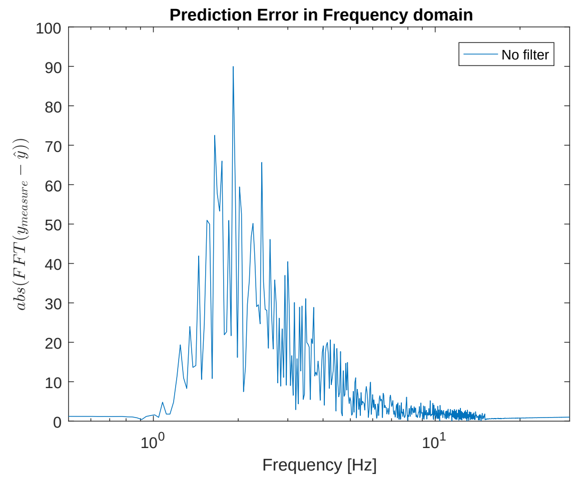


Figure 16: *The prediction error in the frequency domain of the first run of the first subject.*

did not elicit significant time-varying behaviour. This conclusion leads to following advice: in any joint time-varying identification experiment the data needs to be estimated with both the time-varying and time invariant methods to assure the researcher(s) and reader(s) that there is actually time-varying behaviour present. This is needed because a time-varying method will also accurately estimate an LTI system.

In Figure 15 the FRF of the raw data can be seen, and this example particular had a very good fit. It can be seen that the fit is only accurate in the range of 2-15Hz. The fit and FRF separate after 15Hz, which is the frequency where the disturbance signal had no power any more. In the low frequencies, <1Hz, there is a big drop in the FRF. This drop is there because the measured torque had power at those frequencies but the disturbance did not. The output, $T_{measure}$ had power at frequencies <1Hz due to satisficing and due to the SV. Satisficing occurs because people search for a solution that is good enough, not the optimal solution (SIMON, 1976). They follow the trajectory good enough and correct only when they get too far from the desired trajectory and this introduces low frequency noise. Satisficing was also the reason $T_{measure}$ and x_{ankle} were bandpass filtered starting from 2Hz for the state space methods. The SV elicits torques between 0.02-0.25Hz, which are frequencies not present in the disturbance signal. This explains part of the drop in the FRF. Lastly, from Figure 15 it can be seen that the stiffness is most dominant at frequencies <5Hz, where there were few excited frequencies.

The reason the B1 experiment does not elicit any time-varying behaviour is most likely due to a faulty B1 protocol. This has been deduced as follows. It has long been known that the ankle demonstrates time-varying behaviour using only 15% MVC (Kirsch and Kearney, 1997), therefore the problem cannot be with the ankle. The algorithm works, as was shown in the simulation section, therefore the problem is not with the algorithm. What remains, is that the problem must lie with the protocol. The B1 protocol consisted of an experimental setup, a task and instruction, and a perturbation signal. The experimental setup has also been used in other studies, for example Sloot et al. (2015). The task seems clear and also appears to be followed fairly well by the subjects, but that does not mean it is optimal in eliciting time-varying behaviour. Lastly the perturbation signal remains and is probably the biggest cause as will be explained in the next paragraph.

The most probable reason why time-varying behaviour was not observed is because this behaviour mainly occurs at low frequencies, <5Hz, which is dominated by stiffness. The used disturbance signal had little power available at low frequencies. Besides having little power at low frequencies the output behaviour is dominated by the inertia, which adds power to the higher frequencies. The reason the inertia is dominant is because an improper system is identified, see equation (1). The B1 experiment was designed such that the the voluntary and involuntary torque were separated by putting them in 0.02-0.25Hz and 1-15Hz frequency ranges respectively. The problem with this approach is that by separating those torques low frequency power needs to be sacrificed. For example, in the ensemble based approaches of (Kirsch and Kearney, 1997; Ludvig et al., 2011; Bennett et al., 1992) a PRBS was used for the time-varying identification. The PRBS has frequency content below 1Hz. Another difference in those approaches is that they treated the voluntary torque as noise. In the study of Van Eesbeek et al. (2013) a perturbation signal from 2-20Hz was used in combination with the LPV-PBSID algorithm to identify a wrist. In the study the same filtering approach to find the SV and involuntary output was used. The study concluded that it accurately found time-varying behaviour. The findings from this section cast doubts on their result. However, a wrist was identified instead of an ankle which could explain the differences. If the study would have performed a verification with an LTI method this doubt could be put to rest.

In Appendix D a simulation study is described of the B1 experiment using the input and scheduling variable used in the B1 experiment. This simulation can be seen as a B1 experiment with a known time-varying system where the participants perfectly followed the trajectory. Analysing the B1 simulation resulted in the same problems as the analysis of the B1 experiment. This confirms that the problem of the B1 experiment lies with the perturbation signal.

From the previously mentioned studies two experimental paradigms for time-varying identification can be distinguished. One paradigm in which the frequency content of the voluntary torque is separated from perturbations and one in which they are mixed. The advantage of the separated approach is that it has to deal with less noise due to voluntary actions, but it lacks power at low frequencies. The mixed approach has to deal with more noise, but has power in low frequencies. The problem when using the mixed approach for the LPV-NC algorithm is that very noisy data is used. Estimating with noisy data makes it hard to determine whether the correct underlying model is found. As can be seen in table 1, on average a VAF of around 70% was obtained. Is this result trustworthy enough? Probably not, in which case extra measures are needed to assure validity of the model.

In Figure 16 the prediction error in the frequency domain can be seen. The error seems mainly present in the low frequency ranges. This could be interpreted in two ways. Firstly, the system is estimated well but it will not fit the low frequency noise. Therefore, the error is mainly in the low frequencies as was the case in the simulation study. Secondly, it could be that the error is not equally distributed over all excited frequencies. Therefore, not all frequencies were weighted equally in the estimation. The error is biggest in the frequencies that were given the lowest weight. Since the time-varying behaviour is mainly present in the low frequency range this could mean that: pre-filtering to give equal weights to all frequencies could still result in the identification of an LPV system which performs better than an LTI system. These different interpretations of the error have been investigated in Appendix D and it was showed that the first interpretation of the prediction error is correct.

Lastly, humans will have voluntary torque in frequencies that are not present in the disturbance signal or the SV due to satisficing. Both the mixed and separate approach do not deal with the satisficing problem, which is a big source of noise. To circumvent this problem one could design an experiment in which satisficing is reduced. The filtering approach for finding the SV can help with experimental design where the specific trajectory of the SV is not known but it is known that it is a slowly varying signal. Another way to deal with the satisficing problem might be to design a disturbance signal with a bandstop where the voluntary torque should be. This bandstop could then be closer to the satisficing frequency such that those overlap but there is still low frequency content for estimating the stiffness. However, even if the satisficing problem is reduced, you still have the problem that the voluntary torque contribution is in a low frequency range and it cannot be shifted to any desired frequency range.

V. **Wrist data analysis**

In this section an experiment with a torque task and position perturbations of the wrist is analysed. It was performed and analysed in a similar way as the B1 ankle experiment.

i. Wrist protocol

The wrist experiment was a torque task with a position perturbation. The experiment was performed by two healthy subjects: 1 male, 30 year and 1 female, 31 years. The participants needed to apply a torque with their wrist following a sinusoidal signal of 0.05Hz. These experiments have a duration of 50s. During the first and last 5s of the experiment, no perturbations were applied to the subjects. The subjects were requested to modulate their flexion torque between $5-20\%$ of their Maximum Voluntary Torque (MVT). Each participant performed the task 5 times on the experimental setup depicted in Figure 17.

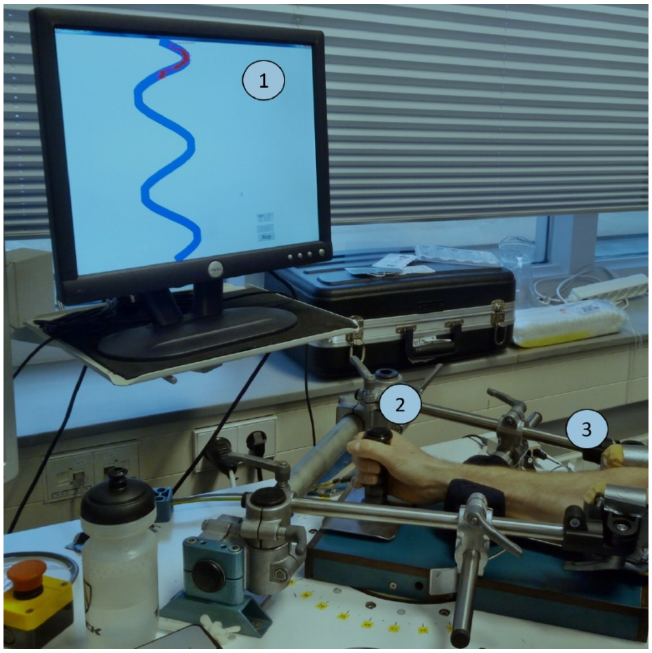


Figure 17: *Picture and caption from Cavallo (2017), showing the experimental setup. 1: screen, showing the desired torque trajectory in blue, allowing for an error of around ±2%, and the applied torque in red. 2: handle of the manipulator. 3: armrest around the forearm of the subject.*

The torque that the participants applied was low-pass filtered with a cutoff of 0.6Hz and visualized on the screen as seen in Figure 17. The applied torque was displayed together with the torque trajectory the participant needed to follow. During the task the participants were perturbed with a low-pass filtered random phase multisine. The random phase multisine consists of a sum of sines with frequencies from 0.1-19.3Hz in steps of 0.8Hz. The signal was filtered with a second order low-pass filter which was constant up to 6Hz. The perturbation had an amplitude of ± 0.02 rad which was applied by the motor and all data was recorded with a sampling rate of $f_s = 2500$Hz.

ii. Analysis with LPV, LTI and FRF

In order to analyse the data three methods were used: the LPV-NC-PBSID, LTI-NC-PBSID and an LTI-FRF method. The first trial of each subject was used to estimate a model which was used to estimate the output of all the trials. The VAF was used to assess the quality of each estimation. For all methods the measured data from $t = 5s$ until $t = 45s$ was used and from all measured signals the mean was removed.

For the LPV-NC-PBSID and LTI-NC-PBSID the measured position and the measured torque were used as input and output respectively. Both signals were decimated from 2500Hz to 100Hz. Subsequently, both signals were bandpass filtered with a 6$^{\text{th}}$ order Butterworth filter with cutoff frequencies at 2Hz and 20Hz, respectively. For the LPV method the measured torque was low-pass filtered (2$^{\text{nd}}$ order Butterworth, 0.1Hz) and used as the SV. For both the LPV-NC and LTI-NC a window size of $p_c = f_{ac} = 2$ was used and $n_c = n_{ac} = 2$, also both algorithms assumed $D = 0$. In the LPV case, it was assumed that A_k, E_k, B_k^c and B_k^{ac} were all time-varying.

The LTI-FRF was made by first estimating an FRF using the unfiltered raw measured torque and the measured position as input and output respectively. The torque was used as input because Matlab is not able to simulate improper transfer functions. From the torque, the position was estimated and the VAF was calculated accordingly. This FRF was calculated with the use of the *tfestimate.m* function of Matlab. The FRF from 0.1Hz-19.3Hz with Δ0.8Hz was used to fit a second order transfer function using the Matlab function *tfest.m*. How this was implemented with specific settings can be seen in Appendix H.

iii. Results

In Tables 3 and 4 the VAF of each estimation for each participant is shown. Along side the VAFs, the mean and standard deviation of each method is given as well. In the tables LPV-NC-PBSID and LTI-NC-PBSID are short handed to LPV and LTI.

The VAF of each method are shown using different trials as initial trial to estimate a model in Appendix I.

	T1	T2	T3	T4	T5	Mean(Std)
LPV	86%	87%	87%	85%	84%	86(1)%
LTI	94%	94%	96%	96%	96%	95(1)%
FRF	76%	77%	83%	78%	80%	79(3)%

Table 3: *Participant 1*

	T1	T2	T3	T4	T5	Mean(Std)
LPV	94%	96%	96%	95%	95%	95(1)%
LTI	98%	96%	96%	98%	98%	97(1)%
FRF	65%	30%	35%	66%	76%	54(21)%

Table 4: *The VAF values of participant 2 using her first trial to build a model and estimate the other trials. T1 is the first trial, T2 the second, etc.*

iv. Discussion

From Tables 3 and 4 it can be seen that the LTI-NC-PBSID method performed best and could accurately predict the output for both participants. In Appendix I the same conclusion can be drawn if other trials were used. The LPV-NC-PBSID also performed well but overall a little less then the LTI equivalent. Suprisingly enough the FRF method performed a lot weaker than the other methods. This could be due to the fact that a proper transfer function was estimated which used input noise. From these results it can be concluded that no significant time-varying behaviour is elicited by the experimental protocol. The reasons for not eliciting time-varying behaviour are the same as described in the discussion of Section IV concerning the ankle. Mainly a perturbation signal which does not have enough power at the low frequencies and an experimental setting in which the time-invariant inertia is dominant.

With the addition of this experiment it can now be conclusively said that the experiment in Van Eesbeek et al. (2013) did not elicit time-varying behaviour. Therefore the results presented in the paper are misleading but not necessarily false. In Van Eesbeek et al. (2013) the identification using the causal LPV-PBSID resulted in a VAF of 91(2.6)% which is in the range of results obtained by the LPV-NC-PBSID as can be seen in Tables 3 and 4, and Appendix I.

Because a time-varying model is estimated, time-varying parameters can also be extracted. In Van Eesbeek et al. (2013) the estimated stiffness displays time-varying behaviour while the inertia does much less so. A reason why the LTI method can accurately estimate the output is because the inertia is the dominant component, as can be seen from equation (1). Even if the stiffness is time-varying, an LTI system can describe much of the behaviour because the inertia is dominant. A solution would be to do a position task with torque perturbations such that stiffness would be the dominant component.

These results also show the importance of using linear methods to estimate supposed time-varying systems because time-varying systems will fit on time invariant ones.

VI. Discussion LPV as identification method

Some studies like Ludvig and Perreault (2014) show that the stiffness displays a more complex behaviour than the sinusoidal varying position and torque during voluntary movement. In the discussion of the simulation study in Section III it was mentioned that the SV is not always perfectly available. These arguments could lead one to think that: it is impossible to find the right SV for joint impedance and therefore LPV methods cannot be used. This idea sells LPV methods short and multiple arguments will be presented why.

Firstly, for LPV state space systems it is more accurate to think of linear varying models instead of parameters. This is already hinted at by the fact that the system description in equation (5) has m local models. In fact, the LPV state space system is a linear weighted combination of different local models where the weights can change in accordance to the SV. Because models are varied instead of parameters, the parameters can change along a different trajectory than the SV. This can be seen from the estimated damping and inertia shown in the results of Section III.

Another advantage of using a state space model is that similarity transforms exist which might be able to handle an altered version of the scheduling variable and still accurately reproduce the output behaviour.

Another possibility is to add more SVs to give the algorithm more freedom to fit the model. However, this approach is not recommended because the SVs will lose their physical interpretation and it can cause overfitting.

Lastly, a scaled version of the SV can always be used because the SV is used to give relative weight to local models. Because of these reasons the filtering approach to find the SV is very powerful and circumvents the problem of finding the 'true' scheduling variable. These statements about the SV can also easily be checked in a simulation using either the LPV-NC-PBSID or the LPV-PBSID.

VII. **Recommendations**

Firstly, implementing the algorithm derived in Section II would allow for the identification of closed-loop systems which is important for joint impedance identification. Closed-loop identification was one of the reasons the PBSID algorithm was chosen in the first place.

An improvement of the algorithm can be done by adding the kernel method (van Wingerden and Verhaegen, 2009). The LPV-NC-PBSID algorithm, like many subspace algorithms, suffers from the *curse of dimensionality*. The curse of dimensionality refers to the fact that the number of rows in Z of equation (17) increases exponentially with the size of the window, f. This can be seen from the equation of the number of rows in Z:

$$rows = (r + l) \sum_{j=1}^{f} m^j.$$

The kernel is a dimension reduction which contains some information of the system. For example, the dot product of a vector with itself is its kernel. This reduces the dimension of the pieces needed to calculate the solution. The advantage of this improvement is that it reduces the computational load of the algorithm and allows for handling bigger systems and more data. The disadvantage is that the solution becomes ill conditioned, but this can be solved with regularization as was done in van Wingerden and Verhaegen (2009).

The last algorithm related recommendation is to expand the algorithm to include Hammerstein systems (a system with a static non-linear component followed by a linear component). The response of the joint to a perturbation is partly due to visco-elastic properties of the tissues around the joint and partly due to the reflexes that are triggered by the disturbance. The reflexive component can be modelled by a Hammerstein system as done in (Mirbagheri et al., 2000; Jalaleddini et al., 2016). These studies use a parallel cascade description to model the joint, the non-causal system description lends itself very well for this description. The Hammerstein system could even be made LPV to incorporate different reflexive behaviour.

Next two experimental recommendations are made. Firstly, it is recommended to experiment with the different experimental paradigms (mixed vs. separated) to see what will elicit the most time-varying behaviour, handles satisficing the best and is easiest to implement. This could for example be done by carrying out a simulation study to find the limitations of using the filtering approach to find the correct SV. Secondly, it is recommended to design an experiment with a position task and force perturbations. In such an experiment an admittance model, in which stiffness is dominant, can be estimated and the LPV-PBSID algorithm from van Wingerden and Verhaegen (2009) can be used. The advantage of this algorithm is that it is already usable for closed-loop systems and uses the kernel method.

Lastly, it is recommended to identify presumed time-varying data with one and preferably multiple LTI methods to make sure it is time-varying.

VIII. **Conclusion**

The theoretical framework for a closed-loop linear parameter varying non-causal predictor based subspace identification (LPV-NC-PBSID) was developed. The algorithm's open-loop version has successfully been implemented and used to identify a time-varying system with a low SNR.

It was shown that both the ankle B1 torque control experiments performed in 2013 and similarly recorded data on the wrist did not elicit significant time-varying behaviour. This was most likely due to a poorly designed perturbation signal which had little power <5Hz where the time-varying behaviour is present. Thus more work is needed to find the correct experimental paradigm for LPV techniques.

Lastly, the results from this report can cast reasonable doubt on the conclusions drawn in Van Eesbeek et al. (2013). To prevent these doubts in the future it is strongly recommended to always identify supposed time-varying data with a time-invariant method.

References

Bennett, D. J., Hollerbach, J. M., Xu, Y. and Hunter, I. W. (1992), 'Time-varying stiffness of human elbow joint during cyclic voluntary movement', *Exp. Brain Res.* **88**(2), 433–442.

Brewer, J. W. (1978), 'Kronecker Products and Matrix

Calculus in System Theory', *IEEE Trans. Circuits Syst.* **25**(9), 772–781.

Cavallo, G. (2017), Skirt Decomposition Method for the Identification of Linear Time-Varying Human Joint Admittance, PhD thesis, Delft University of Technology.
URL: *https://repository.tudelft.nl/islandora/object/uuid:3187dfa9-8728-452e-88ed-a7af1e0a9f3a?collection=education*

Gerdin, M. (2004), *Parameter Estimation in Linear Descriptor Systems*, number 1085.

Golub, G. H. and Loan, C. F. V. (1996), *Matrix Computations*, Vol. 10, John Hopkins University Press.

Jalaleddini, K., Sobhani Tehrani, E. and Kearney, R. (2016), 'A Subspace Approach to the Structural Decomposition and Identification of Ankle Joint Dynamic Stiffness.', *IEEE Trans. Biomed. Eng.* **9294**(c).

Kearney, R. E. and Hunter, I. W. (1990), 'System Identification of Human Joint Dynamics', *Crit. Rev. Biomed. Eng.* **18**(1), 55–87.

Kearney, R. E., Stein, R. B. and Parameswaran, L. (1997), 'Identification of intrinsic and reflex contributions to human ankle stiffness dynamics', *IEEE Trans. Biomed. Eng.* **44**(6), 493–504.

Kirsch, R. F. and Kearney, R. E. (1997), 'Identification of time-varying stiffness dynamics of the human ankle joint during an imposed movement', *Exp. Brain Res.* **114**(1), 71–85.

Lee, H. and Hogan, N. (2015), 'Time-varying ankle mechanical impedance during human locomotion', *IEEE Trans. Neural Syst. Rehabil. Eng.* **23**(5), 755–764.

Lewis, F. L. (1986), 'A Survery of Linear Singular Systems', *Circuits Syst. Signal Process.* **5**(I), 3–36.

Ljung, L. (1998), *System Identification: Theory for the User*, Prentice-Hall.

Ludvig, D. and Perreault, E. J. (2012), 'System identification of physiological systems using short data segments', *IEEE Trans. Biomed. Eng.* **59**(12), 3541–3549.

Ludvig, D. and Perreault, E. J. (2014), 'Task-relevant adaptation of musculoskeletal impedance during posture and movement', *Proc. Am. Control Conf.* pp. 4784–4789.

Ludvig, D., Visser, T. S., Giesbrecht, H. and Kearney, R. E. (2011), 'Identification of time-varying intrinsic and reflex joint stiffness', *IEEE Trans. Biomed. Eng.* **58**(6), 1715–1723.

Mirbagheri, M. M., Barbeau, H. and Kearney, R. E. (2000), 'Intrinsic and reflex contributions to human ankle stiffness: Variation with activation level and position', *Exp. Brain Res.* **135**(4), 423–436.

Mugge, W., Abbink, D. A., Schouten, A. C., Dewald, J. P. A. and Van Der Helm, F. C. T. (2010), 'A rigorous model of reflex function indicates that position and force feedback are flexibly tuned to position and force tasks', *Exp. Brain Res.* **200**(3-4), 325–340.

Nichols, T. and Houk, J. (1976), 'Improvement in linearity and regulation of stiffness that results from actions of stretch reflex.', *J Neurophysiol* **39**(1), 119–142.

Pandyan, A. D., Johnson, G. R., Price, C. I. M., Curless, R. H., Barnes, M. P. and Rodgers, H. (1999), 'A review of the properties and limitations of the Ashworth and modified Ashworth Scales as measures of spasticity', *Clin. Rehabil.* **13**(5), 373–383.

Scholtes, V. a. B., Becher, J. G., Beelen, A. and Lankhorst, G. J. (2006), 'Clinical assessment of spasticity in children with cerebral palsy: a critical review of available instruments.', *Dev. Med. Child Neurol.* **48**(1), 64–73.

Schouten, A. C. (2004), *Proprioceptive reflexes and neurological disorders*.
URL: *http://resolver.tudelft.nl/uuid:b34e0469-39eb-4a76-98b0-fb0d29fa06b5*

SIMON, H. A. (1976), *Administrative behavior*, Macmillan.

Sloot, L. H., van der Krogt, M. M., de Gooijer-van de Groep, K. L., van Eesbeek, S., de Groot, J., Buizer, A. I., Meskers, C., Becher, J. G., de Vlugt, E. and Harlaar, J. (2015), 'The validity and reliability of modelled neural and tissue properties of the ankle muscles in children with cerebral palsy', *Gait Posture* **42**(1), 7–15.

Tóth, R. (2007), *Modeling and Identification of linear Parameter-Varying Systems*, Vol. 360, Springer.

Tóth, R., Abbas, H. S. and Werner, H. (2012), 'On the state-space realization of lpv input-output models: Practical approaches', *IEEE Trans. Control Syst. Technol.* **20**(1), 139–153.

Van Der Helm, F. C. T., Schouten, A. C., De Vlugt, E. and Brouwn, G. G. (2002), 'Identification of intrinsic and reflexive components of human arm dynamics during postural control', *J. Neurosci. Methods* **119**(1), 1–14.

Van Der Veen, G., Van Wingerden, J. W. and Verhaegen, M. (2013), 'Global identification of wind turbines using a hammerstein identification method', *IEEE Trans. Control Syst. Technol.* **21**(4), 1471–1478.

Van Eesbeek, S., Van Der Helm, F., Verhaegen, M. and De Vlugt, E. (2013), 'LPV subspace identification of time-variant joint impedance', *Int. IEEE/EMBS Conf. Neural Eng. NER* pp. 343–346.

van Wingerden, J.-w. (2008), Control of wind turbines with'smart'rotors: Proof of concept & LPV subspace identification, PhD thesis, Technical University of Delft.

van Wingerden, J.-W. and Verhaegen, M. (2009), 'Subspace identification of Bilinear and LPV systems for open- and closed-loop data', *Automatica* **45**(2), 372–381.

Verhaegen, M. (1996), 'A subspace model identification solution to the identification of mixed causal, anti-causal LTI systems', *Siam J. Matrix Anal. Appl.* **17**(2), 332–347.

Verhaegen, Michel; Verdult, V. (2007), *Filtering and system identification: a least squares approach*, Cambridge University Press.

Zhang, L. Q. and Rymer, W. Z. (1997), 'Simultaneous and nonlinear identification of mechanical and reflex properties of human elbow joint muscles', *IEEE Trans. Biomed. Eng.* **44**(12), 1192–1209.

Appendices

A. Closed-loop LPV-NC-PBSID derivation

Full written derivations

In this appendix the computations will be shown such that the definitions of $\phi_{p,k}$, $\mathcal{K}^p$, N_k^p and $P_{p|k}$ for the anti-causal system can easily be found.

The first step is to expand $x_k = E_k x_{k+1} + \breve{B} z_k$ such that a structure becomes visible.

$$\begin{aligned}
x_1 &= E_1 x_2 + \breve{B}_1 z_1 = E_1 E_2 x_3 + E_1 \breve{B}_2 z_2 + \breve{B}_1 z_1 = E_1 E_2 E_3 x_4 + E_1 E_2 \breve{B}_3 z_3 + E_1 \breve{B}_2 z_2 + \breve{B}_1 z_1 \\
x_2 &= E_2 x_3 + \breve{B}_2 z_2 = E_2 E_3 x_4 + E_2 \breve{B}_3 z_3 + \breve{B}_2 z_2 \\
x_3 &= E_3 x_4 + \breve{B}_3 z_3
\end{aligned}$$

Notice how the state is defined with future inputs. When only looking at x_1 and generalize it to x_k a pattern starts to emerge this pattern can be described as

$$\begin{aligned}
x_k &= E_k x_{k+1} + \breve{B}_k z_k \\
&= E_k E_{k+1} x_{k+2} + E_k \breve{B}_{k+1} z_{k+1} + \breve{B}_k z_k \\
&\vdots \\
&= E_k \dots E_{k+p-1} x_{k+p} + E_k \dots E_{k+p-2} \breve{B}_{k+p-1} z_{k+p-1} + E_k \dots E_{k+p-3} \breve{B}_{k+p-2} z_{k+p-2} + \dots + \breve{B}_k z_k
\end{aligned}$$

From this example we can now define $\phi_{p,k}$ as can be seen above

$\phi_{p,k} = E_k E_{k+1} \dots E_{k+p-1}$

To find $\mathcal{K}^p$, N_k^p and $P_{p|k}$ a specific example will be written out more. This will show how we can factorize the unknown state matrices from the known time varying behaviour. Assume $m = 2$, $p = 3$, $k = 1$

$$\begin{aligned}
x_1 &= E_1 x_2 + \breve{B}_1 z_1 \\
&= (E^{(1)} + E^{(2)} \mu_1^{(2)}) x_2 + (\breve{B}^{(1)} + \breve{B}^{(2)} \mu_1^{(2)}) z_1 \\
&= (E^{(1)} + E^{(2)} \mu_1^{(2)})(E^{(1)} + E^{(2)} \mu_2^{(2)}) x_3 + (E^{(1)} + E^{(2)} \mu_1^{(2)})(B^{(1)} + B^{(2)} \mu_2^{(2)}) z_2 + (B^{(1)} + B^{(2)} \mu_1^{(2)}) z_1 \\
&= (E^{(1)} + E^{(2)} \mu_1^{(2)})(E^{(1)} + E^{(2)} \mu_2^{(2)})(E^{(1)} + E^{(2)} \mu_3^{(2)}) x_4 + \dots \\
&\dots (E^{(1)} + E^{(2)} \mu_1^{(2)})(E^{(1)} + E^{(2)} \mu_2^{(2)})(B^{(1)} + B^{(2)} \mu_3^{(2)}) z_3 + \dots \\
&\dots (E^{(1)} + E^{(2)} \mu_1^{(2)})(B^{(1)} + B^{(2)} \mu_2^{(2)}) z_2 + (B^{(1)} + B^{(2)} \mu_1^{(2)}) z_1
\end{aligned}$$

Now we substitute $\phi_{3,1}$ to increase the readability.

$$\begin{aligned}
x_1 &= \phi_{3,1} x_4 + ((E^{(1)})^2 + E^{(1)} E^{(2)} \mu_2^{(2)} + E^{(2)} E^{(1)} \mu_1^{(2)} + (E^{(2)})^2 \mu_1^{(2)} \mu_2^{(2)})(B^{(1)} + B^{(2)} \mu_2^{(2)}) z_3 + \dots \\
&\dots (E^{(1)} + E^{(2)} \mu_1^{(2)})(B^{(1)} + B^{(2)} \mu_2^{(2)}) z_2 + (B^{(1)} + B^{(2)} \mu_1^{(2)}) z_1
\end{aligned}$$

Next we define

$$\begin{aligned}
\mathcal{L}_1 &= \begin{bmatrix} \breve{B}^{(1)} & \breve{B}^{(2)} \end{bmatrix} \\
\mathcal{L}_2 &= \begin{bmatrix} E^{(1)} \breve{B}^{(1)} & E^{(1)} \breve{B}^{(2)} & E^{(2)} \breve{B}^{(1)} & E^{(2)} \breve{B}^{(2)} \end{bmatrix} \\
&= \begin{bmatrix} E^{(1)} \mathcal{L}_1 & E^{(2)} \mathcal{L}_1 \end{bmatrix} \\
\mathcal{L}_3 &= \begin{bmatrix} (E^{(1)})^2 \mathcal{L}_1 & E^{(1)} E^{(2)} \mathcal{L}_1 & E^{(2)} E^{(1)} \mathcal{L}_1 & (E^{(2)})^2 \mathcal{L}_1 \end{bmatrix} \\
&= \begin{bmatrix} E^{(1)} \mathcal{L}_2 & E^{(2)} \mathcal{L}_2 \end{bmatrix}
\end{aligned}$$

Or more generally we can define

$$\mathcal{L}_j = \begin{bmatrix} E^{(1)}\mathcal{L}_{j-1}, & \dots, & E^{(m)}\mathcal{L}_{j-1} \end{bmatrix},$$
with,
$$\mathcal{L}_1 = \begin{bmatrix} \breve{B}^{(1)}, & \dots, & \breve{B}^{(m)} \end{bmatrix}$$

With $\mathcal{L}_j$ defined we can separate the time varying behaviour from the unknown state matrices.

$$x_1 = \phi_{3,1}x_4 + \mathcal{L}_3 \begin{bmatrix} 1 \\ \mu_3^{(2)} \\ \mu_2^{(2)} \\ \mu_2^{(2)}\mu_3^{(2)} \\ \mu_1^{(2)} \\ \mu_1^{(2)}\mu_3^{(2)} \\ \mu_1^{(2)}\mu_2^{(2)} \\ \mu_1^{(2)}\mu_2^{(2)}\mu_3^{(2)} \end{bmatrix} \otimes I_{r+l}z_3 + \mathcal{L}_2 \begin{bmatrix} 1 \\ \mu_2^{(2)} \\ \mu_1^{(2)} \\ \mu_1^{(2)}\mu_2^{(2)} \end{bmatrix} \otimes I_{r+l}z_2 + \mathcal{L}_1 \begin{bmatrix} 1 \\ \mu_1^{(2)} \end{bmatrix} \otimes I_{r+l}z_1$$
$$x_1 = \phi_{3,1}x_4 + \mathcal{L}_3(\mu_1 \otimes \mu_2 \otimes \mu_3 \otimes I_{r+l})z_3 + \mathcal{L}_2(\mu_1 \otimes \mu_2 \otimes I_{r+l})z_2 + \mathcal{L}_1(\mu_1 \otimes I_{r+l})z_1$$

Now we can see how the time varying behaviour is build up, so we can define $P_{j|k}$

$$P_{j|k} = \mu_k \otimes \dots \otimes \mu_{k+j-1} \otimes I_{r+l}$$

With these definitions x_1 simplifies to

$$\begin{aligned} x_1 &= \phi_{3,1}x_4 + \mathcal{L}_3 P_{3|1} z_3 + \mathcal{L}_2 P_{2|1} z_2 + \mathcal{L}_1 P_{1|1} z_1 \\ &= \phi_{3,1}x_4 + \mathcal{L}_1 P_{1|1} z_1 + \mathcal{L}_2 P_{2|1} z_2 + \mathcal{L}_3 P_{3|1} z_3 \\ &= \phi_{3,1}x_4 + \begin{bmatrix} \mathcal{L}_1 P_{1|1} & \mathcal{L}_2 P_{2|1} & \mathcal{L}_3 P_{3|1} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \\ &= \phi_{3,1}x_4 + \begin{bmatrix} \mathcal{L}_1 & \mathcal{L}_2 & \mathcal{L}_3 \end{bmatrix} \begin{bmatrix} P_{1|1} & 0 & 0 \\ 0 & P_{2|1} & 0 \\ 0 & 0 & P_{3|1} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \end{aligned}$$

And lastly we define $\mathcal{K}^p$ and N_k^p as

$$\mathcal{K}^p = \begin{bmatrix} \mathcal{L}_1, & \dots, & \mathcal{L}_p \end{bmatrix}, \quad N_k^p = \begin{bmatrix} P_{1|k} & & & 0 \\ & P_{2|k} & & \\ & & \ddots & \\ 0 & & & P_{p|k} \end{bmatrix}$$

such that we find a general definition for x_1 and x_k

$$\begin{aligned} x_1 &= \phi_{3,1}x_4 + \mathcal{K}^3 N_1^3 \bar{z}_1^3 \\ x_k &= \phi_{p,k}x_{k+p} + \mathcal{K}^p N_k^p \bar{z}_k^p \end{aligned}$$

B. Open-loop LPV-NC-PBSID derivation

Derivation of the LPV-AC-PBSID

To describe the improper systems the following system description is used taken from (Verhaegen, 1996).

$$
\begin{aligned}
x^c_{k+1} &= A_k x^c_k + B^c_k u_k && \text{(Causal)} \\
E_k x^{ac}_{k+1} &= x^{ac}_k - B^{ac}_k u_k && \text{(Anti-Causal)} \\
y_k &= \begin{bmatrix} C^c & C^{ac} \end{bmatrix} \begin{bmatrix} x^c_k \\ x^{ac}_k \end{bmatrix} + D u_k + e_k
\end{aligned}
\tag{28}
$$

where $x^c_k \in \mathbb{R}^{nc}$, $x^{ac}_k \in \mathbb{R}^{nac}$, $u_k \in \mathbb{R}^r$, $e_k \in \mathbb{R}^l$ and $y_k \in \mathbb{R}^l$ are the states, input, white noise and output. The matrices $A_k \in \mathbb{R}^{nc \times nc}$, $E_k \in \mathbb{R}^{nac \times nac}$, $B^c_k \in \mathbb{R}^{nc \times r}$, $B^{ac}_k \in \mathbb{R}^{nac \times r}$, $C^c \in \mathbb{R}^{l \times nc}$, $C^{ac} \in \mathbb{R}^{l \times nac}$ and $D \in \mathbb{R}^{l \times r}$ are the system, input, output and feed through matrices. The time-varying matrices are dependent in an affine way with m local models.

$$
\begin{aligned}
E_k &= \sum_{i=1}^{m} \mu^{(i)}_k E^{(i)} \\
\mu_k &= \begin{bmatrix} 1 & \mu^{(2)}_k & \dots & \mu^{(m)}_k \end{bmatrix}^T
\end{aligned}
$$

In this section only the derivation of the anti-causal part will be shown, closely following the steps made in (van Wingerden and Verhaegen, 2009). The derivation of the causal part is not shown because is follows very similar steps to the anti-causal part and is already done in (van Wingerden and Verhaegen, 2009). From now on only the anti-causal part will be looked at and the "ac" script will be dropped for clarity, so $n = nac$ and $x_k = x^{ac}_k$.

The anti-causal part can be rewritten in the following way.

$$
\begin{aligned}
E_{k+1} x_{k+1} &= x_k - B_k u_k \\
\Rightarrow x_k &= E_{k+1} x_{k+1} + B_k u_k \\
y_k &= C x_k + D u_k + e_k
\end{aligned}
\tag{29}
$$

It is important to know that the matrix E_k is possibly singular and in some cases nilpotent. A nilpotent matrix has all $eigenvalues = 0$ and $E^h_k = 0$ were h is an integer which is $h \leq n$, the order of the matrix.

Assumptions and notations

First the transition matrix is defined, which is slightly different from the one used in the original LPV-PBSID.

$$
\phi_{f,k} = E_{k+1} E_{k+2} \dots E_{k+f} \tag{30}
$$

We define a future window f and the following stacked input vector

$$
\bar{u}^f_k = \begin{bmatrix} u_k \\ \vdots \\ u_{k+f-1} \end{bmatrix}
$$

We assume that the state sequence

$$
X = \begin{bmatrix} x_k, & \dots, & x_{N-f} \end{bmatrix}
$$

is full rank. Notice how the anti-causal system sequence is again slightly different from the one used in the causal estimate of the LPV-PBSID. The matrix:

$$\Gamma^f = \begin{bmatrix} C(E^{(1)})^{f-1} \\ \vdots \\ CE^{(1)} \\ C \end{bmatrix} \tag{31}$$

is the extended observability matrix of the first local model and has full rank.

Factorization

Here we will define the factorization of the time-varying extended controllability matrix, $\bar{\mathcal{K}}_k^f$, in the unknown state matrices and the know scheduling sequence.

Define $\mathcal{K}^f$: We start by defining $\mathcal{K}^f$, the extended controllability. This matrix can be seen as the time-invariant controllability matrix of the LPV system or as the matrix containing all used combinations of affine matrices. We define the matrix with all possible combinations

$$\begin{aligned} \mathcal{L}_j &= \left[E^{(1)}\mathcal{L}_{j-1}, \quad \dots, \quad E^{(m)}\mathcal{L}_{j-1} \right], \\ \text{with} & \\ \mathcal{L}_1 &= \left[B^{(1)}, \quad \dots, \quad B^{(m)} \right] \end{aligned}$$

and $\mathcal{L}_j \in \mathbb{R}^{n \times rm^j}$.

With these variables the extended controllability matrix can be defined

$$\mathcal{K}^f = \left[\mathcal{L}_1, \quad \dots, \quad \mathcal{L}_f \right], \qquad \in \mathbb{R}^{n \times \tilde{q}}$$

with $\tilde{q} = r \sum_{j=1}^{f} m^j$. Notice how $\mathcal{K}^f$ is build op from 1 to f instead of f to 1 as was in the causal case.

Define $P_{f|k}$: Next we define the known time-varying part, this gives weight to the combinations of state matrices.

$$\begin{aligned} P_{f|k} &= \mu_{k+1} \otimes \mu_{k+2} \otimes \dots \otimes \mu_{k+f-1}(\otimes \mu_{k+f-1} \otimes I_{inputs}) \\ P_{1|k} &= \mu_k \otimes I_{inputs} \end{aligned} \tag{32}$$

with $P_{f|k} \in \mathbb{R}^{m^f r \times r}$

Define N_k^f: Now we can define the matrix which gives the correct weigths to the different combinations made in $\mathcal{K}^f$

$$N_k^f = \begin{bmatrix} P_{1|k} & & & 0 \\ & P_{2|k} & & \\ & & \ddots & \\ 0 & & & P_{f|k} \end{bmatrix}$$

with $N_k^f \in \mathbb{R}^{\tilde{q} \times fr}$ such that.

$$\bar{\mathcal{K}}_k^f = \mathcal{K}^f N_k^f$$

Proof The proof can be found in the Appendix. It will show the computations that will lead to these results.

Open-loop identification

Now we can do an open-loop identification of the system, the first objective is to reconstruct the state x_k:

$$x_k = \phi_{f,k} x_{k+f} + \mathcal{K}^f N_k^f \bar{u}_k^f$$

if f is chosen big enough such that $\phi_{f,k} \approx 0$ or in the case of the nilpotent matrix we can guarantee $\phi_{f,k} = 0$ given $f \geq h$. Now we have a definition which approximates the state

$$x_k \approx \mathcal{K}^f N_k^f \bar{u}_k^f \tag{33}$$

In the case of a nilpotent E_k we can guarantee $x_k = \mathcal{K}^f N_k^f \bar{u}_k^f$. With a definition of the state the input-output behaviour can be approximated.

$$y_k \approx C\mathcal{K}^f N_k^f \bar{u}_k^f + Du_k + e_k \tag{34}$$

Now we define the stacked matrices

$$U = \begin{bmatrix} u_1, & \dots, & u_{N-f+1} \end{bmatrix}, \tag{35}$$
$$Y = \begin{bmatrix} y_1, & \dots, & y_{N-f+1} \end{bmatrix}, \tag{36}$$
$$Z = \begin{bmatrix} N_1^f \bar{u}_1^f, & \dots, & N_{N-f+1}^f \bar{u}_{N-f+1}^f \end{bmatrix}, \tag{37}$$

Notice how Y starts at the first sample and ends at y_{N-f+1}, this is because f future values of the input are required to estimate the output. If $[Z^T, \quad U^T]^T$ has full row rank, the matrices $C\mathcal{K}^f$ and D can be estimated by solving the following minimization problem

$$\min_{C,\mathcal{K}^f,D} ||Y - C\mathcal{K}^f Z - DU||_F^2 \tag{38}$$

Observability matrix times controllability matrix

Now we can construct the approximation of the product of the extended observability matrix and controllability matrix. Because of the slightly different notations this will become a lower-block triangular matrix instead of an upper-block triangular in the causal case. This is important because it will provide the possibility to combine the product of the controllability and extended observability matrix in one matrix for the causal and anti-causal case. Unfortunately this is only possible in the LTI case because then $Z^c = Z^{ac}$. The extended observability matrix times the controllability matrix looks like.

$$\begin{aligned} \Gamma^f \mathcal{K}^f &= \begin{bmatrix} C(E^{(1)})^{f-1} \\ \vdots \\ CE^{(1)} \\ C \end{bmatrix} \begin{bmatrix} \mathcal{L}_1, & \dots, & \mathcal{L}_f \end{bmatrix} \\ &\approx \begin{bmatrix} C(E^{(1)})^{f-1}\mathcal{L}_1 & 0 & & 0 \\ \vdots & \ddots & & \\ CE^{(1)}\mathcal{L}_1 & \dots & CE^{(1)}\mathcal{L}_{f-1} & 0 \\ C\mathcal{L}_1 & \dots & C\mathcal{L}_{f-1} & C\mathcal{L}_f \end{bmatrix} \end{aligned} \tag{39}$$

The zeros appear if we chose f big enough such that $\phi_{f,k} \approx 0$ and in the case that E_k is nilpotent we can guarantee this block triangular matrix.

Using the following relations

$$C\mathcal{K}^f = \begin{bmatrix} C\mathcal{L}_1, & \dots, & C\mathcal{L}_f \end{bmatrix}, \tag{40}$$

$$C\mathcal{L}_f = \begin{bmatrix} CE^{(1)}\mathcal{L}_{f-1}, & \dots, & CE^{(m)}\mathcal{L}_{f-1} \end{bmatrix} \tag{41}$$

and the estimate of $C\mathcal{K}^f$ from the minimization problem we can construct $\Gamma^f\mathcal{K}^f$. With this we can construct $\Gamma^f\mathcal{K}^f Z$ which is by definition $\Gamma^f X$. Now assuming both X and Γ^f have full rank and f is big enough to satisfy our assumption $\phi_{f,k} \approx 0$. We can estimate the state sequence using the Singular Value Decomposition (SVD).

$$\widehat{\Gamma^f\mathcal{K}^f Z} = \begin{bmatrix} \mathcal{U} & \mathcal{U}_\perp \end{bmatrix} \begin{bmatrix} \Sigma_n & 0 \\ 0 & \Sigma \end{bmatrix} \begin{bmatrix} V \\ V_\perp \end{bmatrix} \tag{42}$$

where Σ_n is the diagonal corresponding to the n largest singular values and V corresponds to the row space of those singular values. With this the estimate of the state is given by:

$$\hat{X} = \Sigma_n V \tag{43}$$

With the state found (28) can be used to estimate the unknown system matrices.

LPV-AC-PBSID Algorithm The algorithm can be summarized to the following steps:

1. Build Z, Y and U using ((35)-(37)).
2. Solve the minimization problem given in (38).
3. Use relations (40)-(41) and Z to construct $\Gamma^f\mathcal{K}^f Z$.
4. Use SVD to find an estimate of X
5. With the state estimate we can find the system matrices using (29)

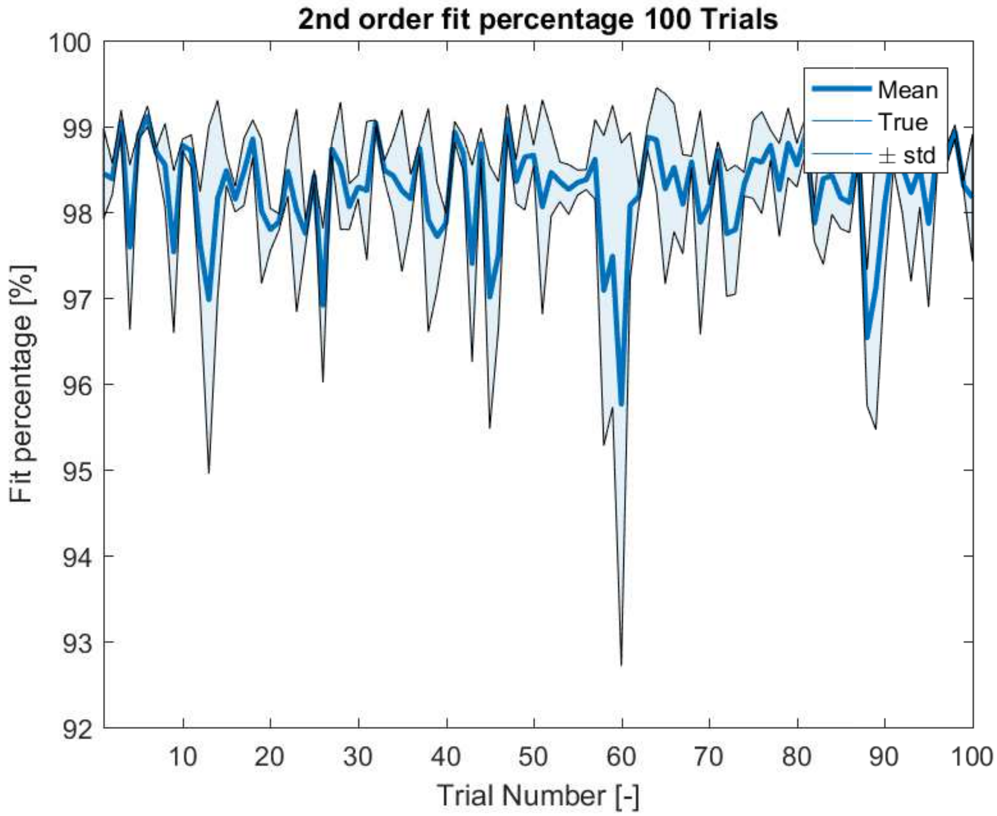
Combining both approaches the open-loop LPV-NC-PBSID

Combining both algorithms leads to solving the following minimization problem.

$$\min_{C,\mathcal{K}^f,D} ||Y - \begin{bmatrix} C^c & C^{ac} \end{bmatrix} \begin{bmatrix} \mathcal{K}_c^f & 0 \\ 0 & \mathcal{K}_{ac}^f \end{bmatrix} \begin{bmatrix} Z^c \\ Z^{ac} \end{bmatrix} - DU||_F^2 \tag{44}$$

We finally find $\begin{bmatrix} C^c\mathcal{K}_c^f & C^{ac}\mathcal{K}_{ac}^f & D \end{bmatrix}$ Which can be used in the algorithm described above to find the states of the causal and anti-causal system and estimate their system matrices.

C. Simulation Study



D. Simulation and Error analysis B1 data

In this appendix the simulation of Section III is adapted to mimic the B1 experiment. This is done to gain more insight in where the problem lies with the B1 experiment. Besides simulating the system some data of the experiment is pre filtered to compensate for errors in the frequency domain. It will be shown that the simulation results support the conclusions drawn in the discussion of the B1 data section and that pre filtering to compensate for errors in the frequency domain will decrease the quality of the estimate for both the LPV and LTI method.

Simulating the B1 experiment

Below the simulation used to test the LPV-NC-PBSID algorithm is adapted to mimic the B1 experiment. This is done by using the disturbance signal and scheduling variable that were used for the B1 experiment.

Model and Analysis The model of Figure 1 was used in combination with the inputs and scheduling variables of the participants. The simulation was run 129 times and for each simulation the multisine input from the experiment was used for x_per (see Figure 1. For the identification, the x_{ankle} and $T_{measure}$ were used as input and output respectively. As scheduling variable the chirp signal from the B1 experiment was used, an example can be seen in Figure 18. The SV was scaled between 0 and 1 such that $K(t) = 100 + 50 * chirp(t)$. The simulation data was decimated from 2.5kHz to 100Hz before being used for identification. The LPV algorithm assumed that A_k,E_k,B_k^c and B_k^{ac} were time-varying. A past and future window of $p_c = f_{ac} = 2$ and an order selection of $n_c = 2$ and $n_{ac} = 2$ gave the best results. The quality of the estimation was quantified by looking at the Variance Accounted For (VAF). The disturbance was the same as was used in the simulation of Section III.

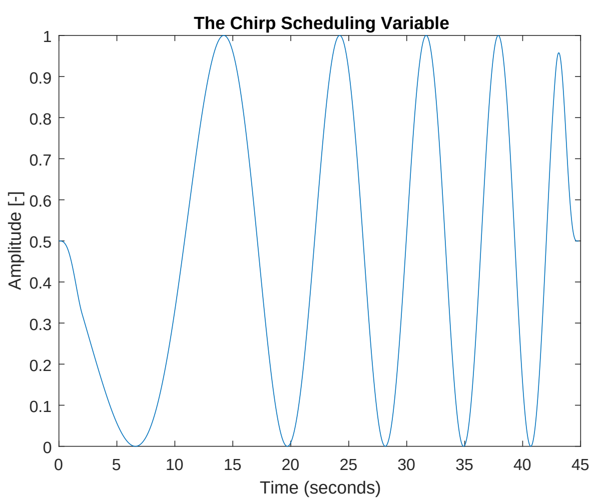


Figure 18: *The trajectory the participant had to follow (the scheduling variable)*

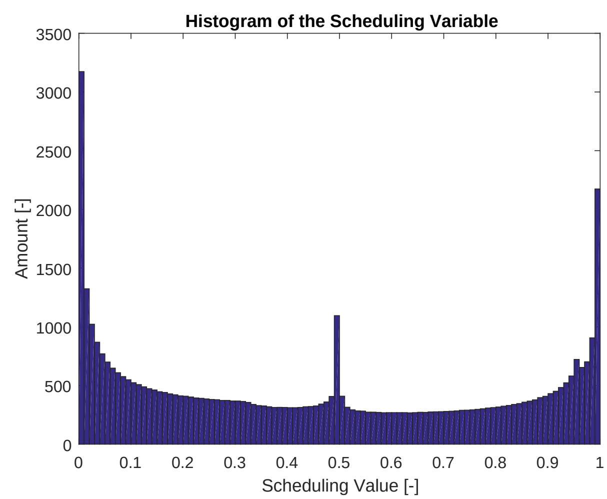


Figure 19: *Historgram of the scheduling variable*

Results The VAF was calculated for each experiment with the noisy and true signal, these are displayed in Figure 20. The LPV method had two outliers both, at 0%. The average and standard deviation have been calculated without the two outliers and are shown in Table 5.

Discussion This simulation can be seen as a B1 experiment where the participants perfectly followed the instructions, the machine was perfect and the underlying system was time-varying. The simulation results show

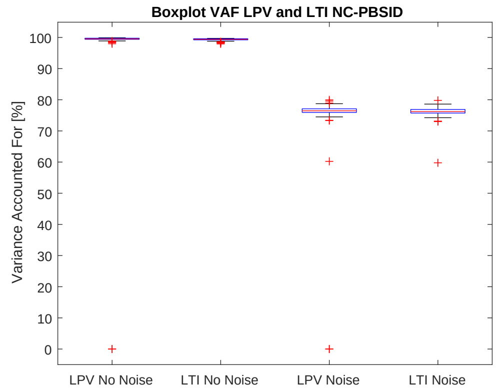


Figure 20: *Boxplot of all the calculated VAF's with two outliers for the LPV method on top of each other at 0%*

	With True	**With Noise**
LPV	99.5(0.3)%	76.4(1.8)%
LTI	99.4(0.3)%	76.2(1.8)%

Table 5: *The average VAF and standard deviation without the two outliers.*

the same problems as the B1 experiment results. The LTI and LPV method provide similair results. Therefore, it can confidantly be stated that the problem of the B1 protocol lies with its disturbance design.

From Figure 19 it can be seen that most time is spend at the edges of the scheduling variable which indicates that a LTI system should not be able to estimate the system. The fact that most time is spend in the two extreme regions further shows that the LTI content is dominant in this experiment.

Errors in frequency domain

An advantage of estimating improper or not strictly proper systems is that they weigh frequencies differently than strictly proper systems, which weigh low frequencies more heavily. To show what this means a slight detour in LTI system theory is made.

Theory When the LTI equivalent of equation (14) is written and it is assumed that f has been chosen sufficiently big that $\phi_{f,k} = 0$. The equivalence to the vector-ARX(VARX) model structure can be seen as follows:

$$\begin{aligned} y_k &= C\mathcal{K}^f \bar{z}_k^f + Du_k + e_k \\ \Rightarrow A(q)y_k &= B(q)u_k + e_k \\ \Rightarrow y_k &= \frac{B(q)}{A(q)}u_k + \frac{1}{A(q)}e_k \\ \Rightarrow y_k &= W_0(q)u_k + L_0(q)e_k. \end{aligned}$$

A more elaborate explanation can be found in (Van Der Veen et al., 2013). Now the input output behaviour is captured in a higher order VARX model which is equivalent to the state space predictor model; the LTI version of equation (2) (see example 10.11 (Verhaegen, Michel; Verdult, 2007)). Given the estimates of the system $W(q,\theta), L(q,\theta)$ where θ are the parameters, equation (27) can be given in the frequency domain:

$$\min_{\theta} \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{|W_0(e^{j\omega}) - W(e^{j\omega},\theta)|^2 \Phi_u(\omega) + \Phi_e(\omega)]}{|L(e^{j\omega},\theta)|^2} d\omega.$$

This is the well known formula 8.66 from Ljung (1998), with $\Phi_u(\omega)$ and $\Phi_e(\omega)$ as the spectral density of the input and noise respectively. From formula 8.66 it can be seen that some frequencies of the estimated model, $W(e^{j\omega}, \theta)$, will add more weight than others. For example, if a strictly proper system is estimated the low frequencies will weigh heavier than the high frequencies. In other words $|W_0(e^{j0}) - W(e^{j0}, \theta)|^2 > |W_0(e^{j\infty}) - W(e^{j\infty}, \theta)|^2$.

Because the NC-PBSID either estimates proper (not strictly proper) or improper systems different behaviour will occur. Proper systems will not infinitely amplify high frequencies like improper systems. In other words, a proper system can either have:

$$|W_0(e^{j0}) - W(e^{j0}, \theta)|^2 > |W_0(e^{j\infty}) - W(e^{j\infty}, \theta)|^2 \quad \text{or}$$
$$|W_0(e^{j0}) - W(e^{j0}, \theta)|^2 < |W_0(e^{j\infty}) - W(e^{j\infty}, \theta)|^2.$$

An example of $|W_0(e^{j0}) - W(e^{j0}, \theta)|^2 < |W_0(e^{j\infty}) - W(e^{j\infty}, \theta)|^2$ can be seen in Figure 9. The order of the poles and zeros determines the weight of the frequencies. Where the proper system in Figure 9 first has two zeros and then two poles another proper system could have two poles and then the two zeros. This would change what frequencies weigh heavier. In contrast to improper systems which will always have a heavier weight on the high frequencies because for an improper system $|W(e^{j\infty}, \theta)| = \infty$.

The noise filter $L(e^{j\omega}, \theta)$ influences the frequency weighing as well. Because the human interjects low frequency noise, $L(e^{j\omega}, \theta)$, can be seen as a low-pass filter. Therefore $1/|L(e^{j\omega}, \theta)|$ will increase the weight of the high frequencies.

One could conclude from this that if a pre-filter was added to the data, such that all frequencies would get equal weighting in the minimization problem, the LPV method would perform better. The LPV method would perform better because the time-varying part is present in the low frequency range which would then be as important as the time-invariant high frequency part. However, the following paragraph will show that this is not the case.

Analysis Now a closer look will be taken at the frequency error of the B1 experiment. The input and output data will get an extra filter besides the bandpass filter. The input and output will be pre-filtered with a 2nd order butterworth filter with a cutoff frequency of 6,8 and 10Hz before being used for the estimation of the system. After the system is estimated the output will be predicted using the non-pre-filtered input and compared to the non-pre-filtered output. In Figure 21 and 22 the frequency content of the error $y - \hat{y}$ from the first trial of the first participant is given and in Table 6 the VAF of each prediction is given.

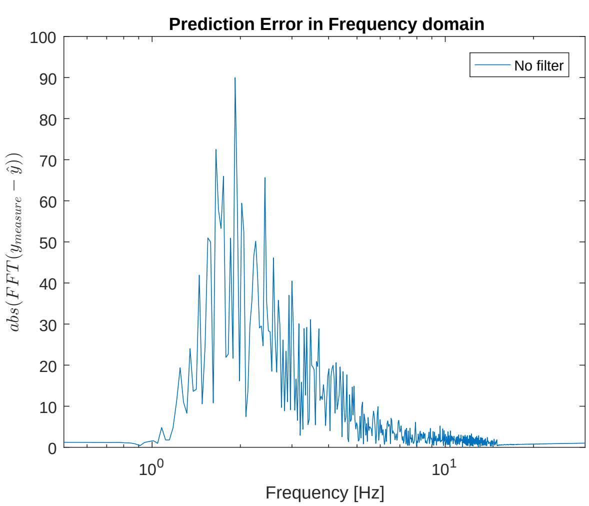


Figure 21: *The prediction error in the frequency domain*

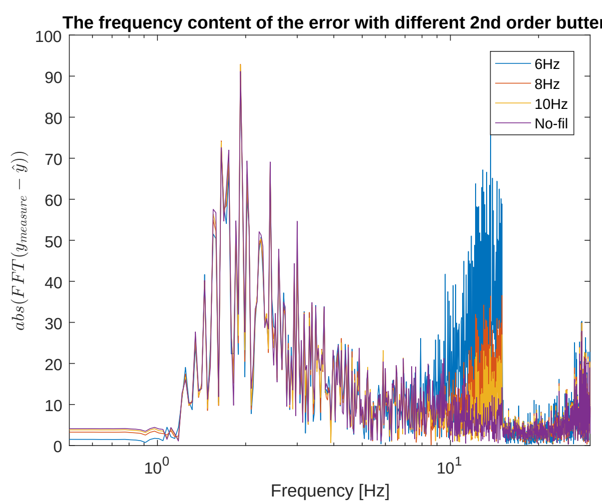


Figure 22: *Error in frequency domain with multiple fitlers. Notice how filtering increases the error in at the high frequencies*

	No filter	12Hz	10Hz	8Hz	6Hz
LPV	78%	71%	70%	70%	69%
LTI	77%	70%	68%	68%	68%

Table 6: *VAF of participant 1 applying different prefilters besides the bandpass filter.*

In Figure 21 it can be seen that the bulk of the frequency content of the error is in the low frequencies as was expected. Subsequently, low-pass filters were added to decrease the weight of the high frequencies. In Figure 22 it can be seen that this does indeed increase the error on the high frequencies. However, Table 6 shows that the VAF of the LPV and LTI method are not diverging from each other. By equalizing the weights, the error of the high frequency content went up but the error on the low frequencies did not decrease. One would expect the error in the low frequencies to decrease because they are given more weight by the pre-filtering. The most likely explanation why this does not happen is that the method is not fitting the noise at those frequencies and therefore the error does not decrease. Because the VAF's do not diverge and because of the presence of low frequency noise, weighting of the frequencies did not have the desired effect of extracting the time-varying behaviour.

Next all three butterworth filters were applied to all the B1 experiment data. This resulted in 129 VAF's for each filter. The results are displayed in Figure 23, the mean and standard deviation can be found in Table 7

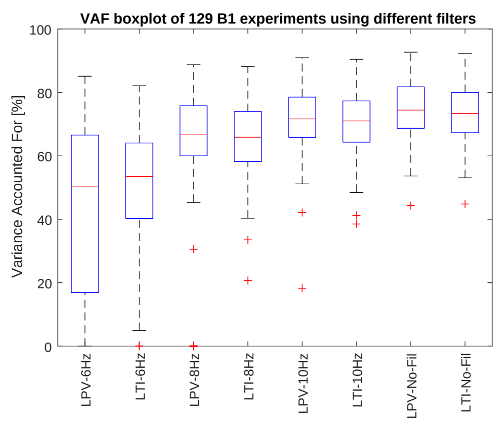


Figure 23: *The Boxplots containing the VAF of each filter for each experiment.*

	6Hz	8Hz	10Hz	No Filter
LPV	43.1(27.7)%	63.0(21.0)%	71.2(10.6)%	74.1(9.2)%
LTI	50.1(18.5)%	65.7(11.5)%	70.1(9.9)%	73.1(9.14)%

Table 7: *The average VAF and standard deviation for the VAF's of each filter for both the LPV and LTI method.*

Lastly, a Mann-Whitney test was applied to see if the filtering created any significant difference between the LPV or LTI results. The Mann-Whitney was used to see if each of the LPV-LTI groups differed from each other using a 5% significance. The 6Hz group did not significantly differ in the VAF the produced, $p = 0.24$, $r = -0.10$. The 8Hz group did not significantly differ in the VAF the produced, $p = 0.49$, $r = 0.06$. The 10Hz group did not significantly differ in the VAF the produced, $p = 0.28$, $r = 0.09$. The "No-filter" group did not significantly differ in the VAF the produced, $p = 0.34$, $r = 0.08$.

With these results it can conclusively be stated that: even if all frequencies were weighed equally, it would not have separated the LPV and LTI method. Another conclusion that can be drawn is that giving more equal weights to all frequencies does not improve the VAF.

Summary

In this Appendix two analysis have been done to further support the conclusions drawn from the results in the main text. The simulation study of the B1 experiment confirms that the problem of the protocol is the perturbation signal. The analysis of the error in the frequency domain confirmed that pre-filtering the data to equalize the error in the frequency domain will not separate the LPV and LTI estimate.

E. LTI mixed causal anti-causal state space to transfer function

In this appendix it will be shown how to analytically transform between the LTI mixed causal anti-causal state space form and transfer function. This is mainly added for completion.

Continuous transform

First the transform is done in continuous time. If we have the following system

$$\begin{aligned} \dot{x}^c &= Ax^c + B^c u, \\ E\dot{x}^{ac} &= x^{ac} - B^{ac}u, \\ y &= C^c x^c + C^{ac}x^{ac} + Du. \end{aligned}$$

Next the Laplace transformation is applied to the system and get a definition of $X(s)$ for both the causal and anti-causal part and a continuous time transfer function can be made.

$$\begin{aligned} sX(s)^c &= AX(s)^c + B^c U(S) \\ EsX(s)^{ac} &= X(s)^{ac} - B^{ac}U(S) \\ \Rightarrow \begin{matrix} (Is-A)X(s)^c \\ (Es-I)X(s)^{ac} \end{matrix} & \begin{matrix} = B^c U(S) \\ = -B^{ac}U(S) \end{matrix} \\ \Rightarrow \begin{matrix} X(s)^c \\ X(s)^{ac} \end{matrix} & \begin{matrix} = (Is-A)^{-1}B^c U(S) \\ = -(Es-I)^{-1}B^{ac}U(S) \end{matrix} \\ \Rightarrow Y(s) &= (C^c(Is-A)^{-1}B^c - C^{ac}(Es-I)^{-1}B^{ac} + D)U(s). \end{aligned}$$

Discrete time

This straightforward derivation can be done as well for the continuous time case. If we have the following system

$$\begin{aligned} x^c_{k+1} &= Ax^c_k + B^c u_k, \\ Ex^{ac}_{k+1} &= x^{ac}_k - B^{ac}u_k, \\ y_k &= C^c x^c_k + C^{ac}x^{ac}_k + Du_k. \end{aligned}$$

Next the Z-transform is applied such that the set of equations can be rewritten into a discrete time transfer function.

$$\begin{aligned} zX(z)^c &= AX(z)^c + B^c U(z, \\ EzX(z)^{ac} &= X(z)^{ac} - B^{ac}U(z) \\ \Rightarrow \begin{matrix} (zI-A)X(z)^c \\ (Ez-I)X(z)^{ac} \end{matrix} & \begin{matrix} = B^c U(z) \\ = -B^{ac}U(z) \end{matrix} \\ \Rightarrow \begin{matrix} X(z)^c \\ X(z)^{ac} \end{matrix} & \begin{matrix} = (zI-A)^{-1}B^c U(z) \\ = -(Ez-I)^{-1}B^{ac}U(z) \end{matrix} \\ \Rightarrow Y(z) &= (C^c(zI-A)^{-1}B^c - C^{ac}(Ez-I)^{-1}B^{ac} + D)U(z). \end{aligned}$$

This concludes the derivations.

F. Figures C1 and B1 Study

In this Appendix figures that resulted from the analyses of the B1 and C1 study are shown.

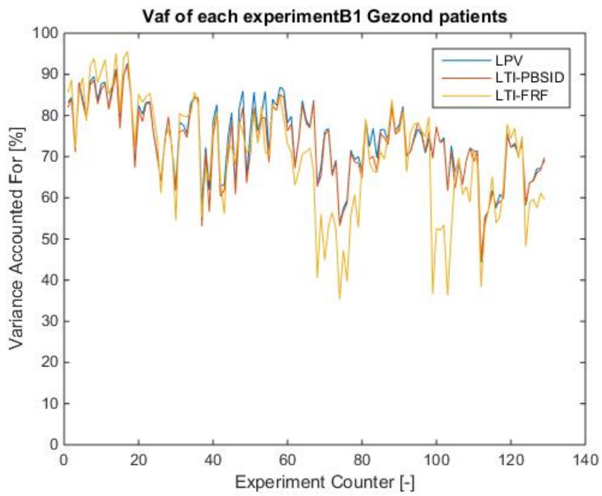


Figure 24: *VAF for each experiment of B1 healthy elderly*

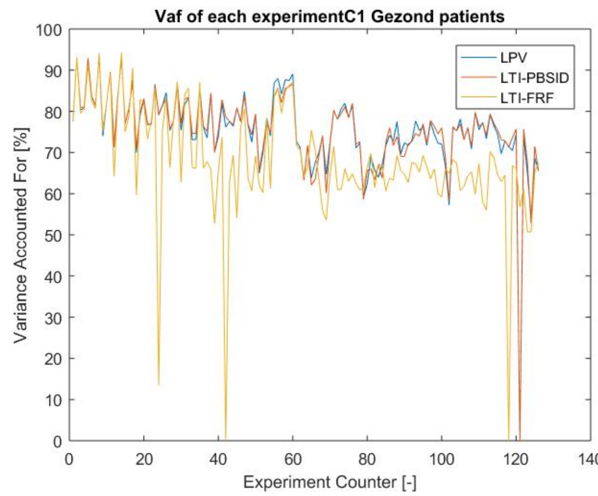


Figure 25: *VAF for each experiment of C1 healthy elderly*

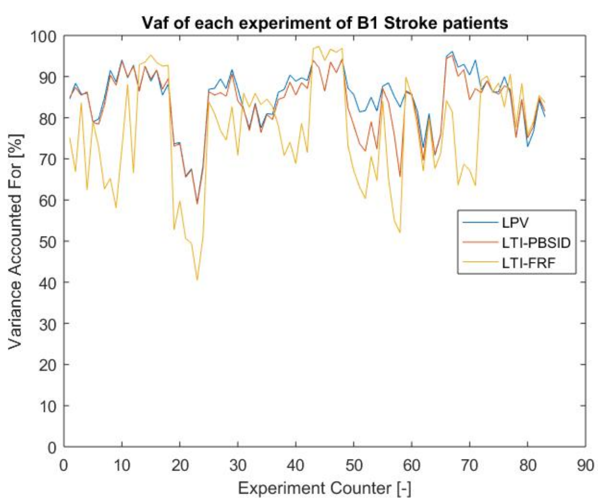


Figure 26: *VAF for each experiment of B1 stroke patients group 1*

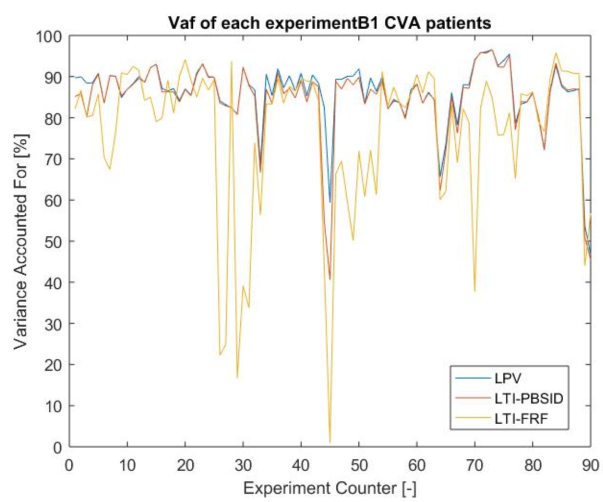


Figure 27: *VAF for each experiment of B1 stroke patients group 2*

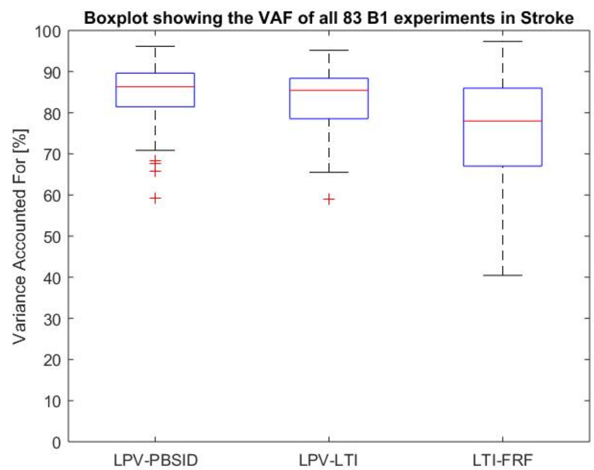


Figure 28: *Box plot of B1 experiments for stroke patients group 1*

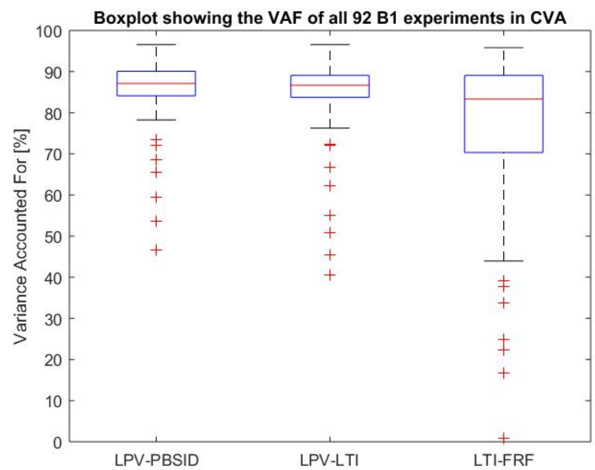


Figure 29: *Box plot of B1 experiments for stroke patients group 2*

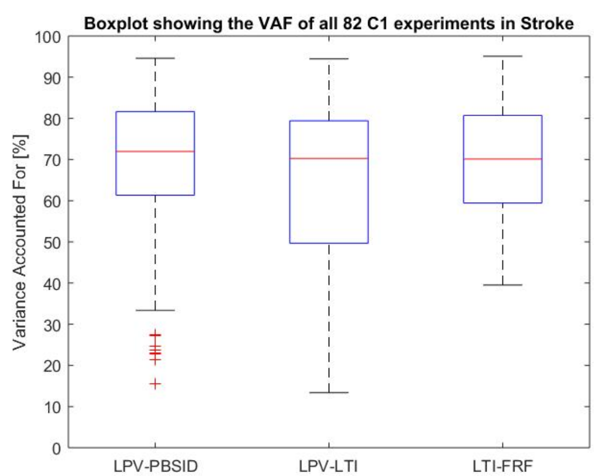


Figure 30: *Box plot of C1 experiments for stroke patients group 1*

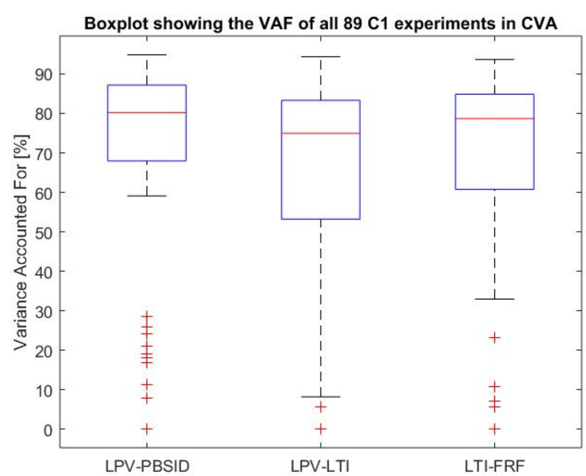


Figure 31: *Box plot of C1 experiments for stroke patients group 2*

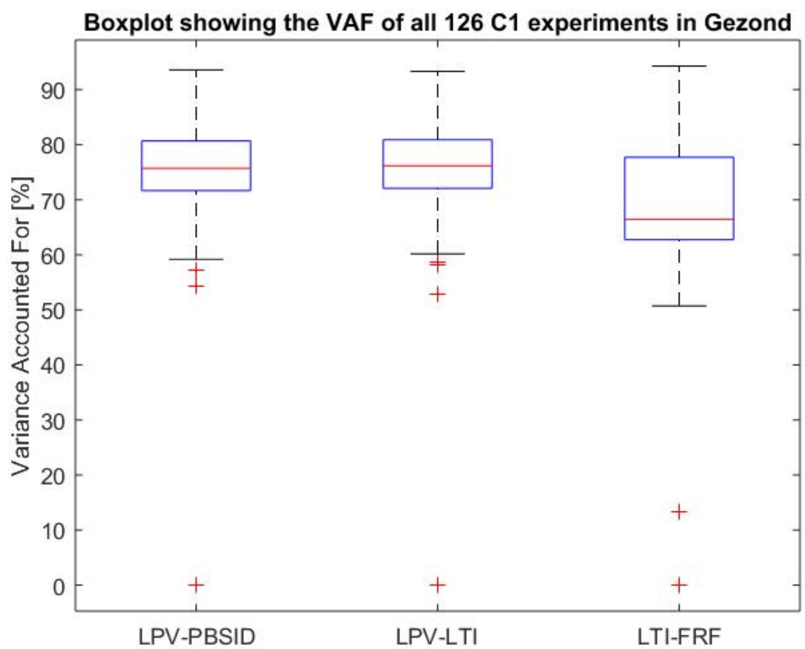


Figure 32: *Box plot of C1 experiments for healthy elderly*

G. Analytical Approach to Mass Spring Damper System Description

In this appendix it will analytically be shown that there is no one to one translation of the time varying mass spring damper system in the form of a transfer function (TF) to an LPV state space (SS) description. Here an admittance model is use to prove this claim.

The Derivation

First the time varying mass spring damper system which was used in Section 3 of the report is slightly rewritten.

$$H(s) = \frac{1}{Is^2 + Bs + k_1 + k_2\mu} \rightarrow \frac{1/I}{s^2 + \frac{B}{I}s + \frac{k_1+k_2\mu(t)}{I}} \tag{45}$$

For LTI systems there exist certain standard descriptions to change TF's into SS systems and back. This is done because the state space description is not unique, these descriptions are called canonical forms. It will be assumed that the LPV transfer function can be written into the observability canonical form with an affine dependency.

$$\begin{aligned}
\dot{x}(t) &= \begin{bmatrix} 0 & -\frac{k_1+k_2\mu}{I} \\ 1 & -\frac{B}{I} \end{bmatrix} x(t) + \begin{bmatrix} \frac{1}{I} \\ 0 \end{bmatrix} u(t) \\
&= \left(\begin{bmatrix} 0 & -\frac{k_1}{I} \\ 1 & -\frac{B}{I} \end{bmatrix} + \mu \begin{bmatrix} 0 & -\frac{k_2}{I} \\ 0 & 0 \end{bmatrix} \right) x(t) + \begin{bmatrix} \frac{1}{I} \\ 0 \end{bmatrix} u(t) \\
\dot{x}(t) &= A(\mu)x(t) + Bu(t) \\
y(t) &= \begin{bmatrix} 0 & 1 \end{bmatrix} x(t) + 0u(t) \\
y(t) &= Cx(t)
\end{aligned} \tag{46}$$

To prove this is state space model correctly represents $H(s)$ we calculate $H(s)$ from the found state space system.

$$\begin{aligned}
H(s) &= C(sI - A)^{-1}B + D \\
&= \begin{bmatrix} 0 & 1 \end{bmatrix} \left(\begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix} - \begin{bmatrix} 0 & -\frac{k_1+k_2\mu}{I} \\ 1 & -\frac{B}{I} \end{bmatrix} \right)^{-1} \begin{bmatrix} \frac{1}{I} \\ 0 \end{bmatrix} + 0 \\
&= \begin{bmatrix} 0 & 1 \end{bmatrix} \left(\begin{bmatrix} s & \frac{k_1+k_2\mu}{I} \\ -1 & s + \frac{B}{I} \end{bmatrix} \right)^{-1} \begin{bmatrix} \frac{1}{I} \\ 0 \end{bmatrix} \\
&= \frac{1}{s(s+\frac{B}{I}) + \frac{k_1+k_2\mu}{I}} \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} s + \frac{B}{I} & -\frac{k_1+k_2\mu}{I} \\ 1 & s \end{bmatrix} \begin{bmatrix} \frac{1}{I} \\ 0 \end{bmatrix} \\
&= \frac{1}{s^2 + \frac{B}{I}s + \frac{k_1+k_2\mu}{I}} \begin{bmatrix} 1 & s \end{bmatrix} \begin{bmatrix} \frac{1}{I} \\ 0 \end{bmatrix} \\
&= \frac{1/I}{s^2 + \frac{B}{I}s + \frac{k_1+k_2\mu}{I}} = \frac{1}{Is^2 + Bs + k_1 + k_2\mu} = H(s)
\end{aligned}$$

This shows that (46) is the correct state space representation of (45).

Discretize

To discretize the system the bilinear or Tustin approach is used. This approach has been taken from (Van Wingerden 2008, van Wingerden et al. 2010).

$$\begin{aligned}
A_d(\mu, \frac{1}{\mu}, ...) &= \left(I + \frac{T_s}{2}A(\mu)\right)\left(I - \frac{T_s}{2}A(\mu)\right)^{-1} \\
B_d(\mu, \frac{1}{\mu}, ...) &= \sqrt{T_s}\left(I - \frac{T_s}{2}A(\mu)\right)^{-1} B \\
C_d(\mu, \frac{1}{\mu}, ...) &= \sqrt{T_s}C\left(I - \frac{T_s}{2}A(\mu)\right)^{-1} \\
D_d(\mu, \frac{1}{\mu}, ...) &= \frac{T_s}{2}C\left(I - \frac{T_s}{2}A(\mu)\right)^{-1} B + D
\end{aligned} \tag{47}$$

With the use of the symbolic toolbox of Matlab this leads to the following discrete state space system

$$\begin{aligned}
A_d(\mu_1, \mu_2) &= \begin{bmatrix} 4I - T_s^2 k_1 + 2BT_s & -4T_s k_1 \\ 4ITs & 4I - T_s^2 k_1 - 2BT_s \end{bmatrix}\mu_1 + \begin{bmatrix} -T_s^2 k_2 & -4T_s k_2 \\ 0 & -T_s^2 k_2 \end{bmatrix}\mu_2 \\
B_d(\mu_1, \mu_2) &= \begin{bmatrix} \frac{2\sqrt{T_s}(2I+BT_s)}{I} \\ 2\sqrt{T_s}T_s \end{bmatrix}\mu_1 + \begin{bmatrix} 0 \\ 0 \end{bmatrix}\mu_2 \\
C_d(\mu_1, \mu_2) &= \begin{bmatrix} 2I\sqrt{T_s}T_s & 4I\sqrt{T_s} \end{bmatrix}\mu_1 + \begin{bmatrix} 0 & 0 \end{bmatrix}\mu_2 \\
D_d(\mu_1, \mu_2) &= T_s^2\mu_1 + 0\mu_2 \\
\mu_1(\mu) &= \frac{1}{(4I + T_s^2 k_1 + 2BT_s) + T_s^2 k_2 \mu} \\
\mu_2(\mu) &= \mu_1\mu = \frac{\mu}{(4I + T_s^2 k_1 + 2BT_s) + T_s^2 k_2 \mu}
\end{aligned}$$

From this analytical expression it can be seen that all the system matrices are now dependent on some variant of the original scheduling variable μ. The affine structure dependent on μ is lost but two new scheduling functions μ_1 and μ_2 are found which preserve the affine structure.

Discussion

To accurately represent $H(s)$ the discrete time LPV state space model will need all system matrices dependent on some variation of the scheduling variable μ. What can be seen in the derivation is that the new scheduling variables μ_1 and μ_2 are not only dependent on the scheduling variable but also on system parameters. So to perfectly estimate the system the system parameters need to be know such that the correct scheduling functions can be used. Since this is impossible, there will always be some difference between the simulated continuous time LTV transfer function and estimated discrete time LPV state space model.

H. Matlab Code Implementation

Below the Mcode used to analyse the data is shown. This code would be inside one or multiple loops which loaded the data that was going to be analysed. Here can be seen how Matlab functions as *tfestimate.m* and *tfest.m* are used.

```
%% Extract data
Time = OutputData.AchData(:,1);
Mpos = OutputData.AchData(:,8);
Mvel = OutputData.AchData(:,9);
Mtorque = OutputData.AchData(:,10);
MtorqComp = OutputData.AchData(:,11);

%% Filter to get the SV
[B,A] = butter(2,0.25/512);
TVtraj = filtfilt(B,A,Mtorque);

%% Decimate and detrend the data
fs = 1024;
Deci = 8; % Resulteerd in fs = 128
fnew = fs/Deci;

TimeD = downsample(Time,Deci);
MtorqueD = -1*decimate(detrend(Mtorque,'constant'),Deci);
MposD = decimate(detrend(Mpos,'constant'),Deci);
TrajTVD = decimate(TVtraj,Deci); % Is already detrended
MvelD = decimate(detrend(Mvel,'constant'),Deci);

%% Bandpass filter output data to only have involuntary torque
[Bh,Ah] = butter(6,[1.5/(fnew/2) 40/(fnew/2)]);
MtorqueDfil = filtfilt(Bh,Ah,MtorqueD);
MposDfil = filtfilt(Bh,Ah,MposD);
MvelDfil = filtfilt(Bh,Ah,MvelD);

%% Select ID data
%Select samples
if Time(end)<40
    continue
end
if Time(end)>40
    T0 = 5;
    Tend = 35; % Select as much of the usable data as possible
end
if Time(end)>90
    T0 = 10;
    Tend = 70;
end

SamplesID = 1+T0*fnew:Tend*fnew;

nSV = 1; % Number of derivatives of the SV
Syn = 0; % Scale the SV y=1 n=0

%% Select data to ID
% In = [MposDfil(SamplesID).'; MvelDfil(SamplesID).'];
In = MposDfil(SamplesID).';
Out = MtorqueDfil(SamplesID).';
SV = TrajTVD(SamplesID).';
%SV_Scale_Derivs,m can calculate derivatives of the SV and scale them.
Use = SV_Scale_Derivs(SV,nSV,fnew,Syn,-1,1);
mu = [ones(1,length(SamplesID)); Use.'];

%% Do frequency estimation
```

```
dt = 1/1024;
[txy, f2] = tfestimate(-1*detrend(Mtorque,'constant')),Mpos,hanning(10/dt),5/dt,10/dt,1/dt);

FRdata = frd(txy(11:151),2*pi*f2(11:151));
Sysid = tfest(FRdata,2,0);

[mag,¬] = bode(Sysid,2*pi*f2);
Ytf = lsim(Sysid,Out,(0:length(Out)-1)*1/fnew);

figure
loglog(f2,abs(txy)); hold all
semilogx(f2,squeeze(mag))
xlabel('Frequency [Hz]')
ylabel('Magnitude [-]')
title('FRF of the measured position to measured torque with fit')
legend('FRF','2nd order fit')

%% ID search
pend = 2;
r = 1; % #Input
l = 1; % #Output
m = nSV+1; % #Local Models
nac = 2; % ORder Causal
nc = 2; % Order A-Causal

[Ai,Ei,Bci,Baci,Cci,Caci,Di] = LPV_ACC_PBSID(In,Out,mu,pend,r,l,m,[0 0],[0 0],nc,nac);
[AiL,EiL,BcL,BacL,CcL,CacL,DiL] = LTI_ACC_PBSID(In,Out,pend,r,l,nac,nc);

%% Simulate and compare
f = 40;

[ysim, ¬,¬,¬,¬] = LPV_ACC_SIM(Ai,Ei,Bci,Baci,Cci,Caci,Di,In,mu,f,r,l,nc,nac);
[YL,¬,¬,¬,¬] = LTI_ACC_SIM(AiL,EiL,BcL,BacL,CcL,CacL,DiL,In,(0:length(In)-1)*1/fnew,f);

display('VAF for the data used for the identification')
All_That_VAF(count,:) = [vaf(Out(1:end-f),ysim) vaf(Out(1:end-f),YL) vaf(In,Ytf)]
```

I. Wrist Study Tables

	T1	T2	T3	T4	T5	Mean(Std)
LPV	86	87	87	85	84	86(1)
LTI	94	94	96	96	96	95(1)
FRF	76	77	83	78	80	79(3)

Table 8: *The VAF's for each method and trial of Participant 1 using Trial 1 as estimator*

	T1	T2	T3	T4	T5	Mean(Std)
LPV	89	90	91	88	88	89(1)
LTI	94	94	96	96	96	95(1)
FRF	76	77	83	78	80	79(3)

Table 9: *The VAF's for each method and trial of Participant 1 using Trial 2 as estimator*

	T1	T2	T3	T4	T5	Mean(Std)
LPV	90	91	92	89	89	90(1)
LTI	94	94	96	96	96	95(1)
FRF	75	75	83	78	81	78(4)

Table 10: *The VAF's for each method and trial of Participant 1 using Trial 3 as estimator*

	T1	T2	T3	T4	T5	Mean(Std)
LPV	90	91	92	90	90	91(1)
LTI	94	94	96	97	97	96(2)
FRF	68	68	78	80	81	75(6)

Table 11: *The VAF's for each method and trial of Participant 1 using Trial 4 as estimator*

	T1	T2	T3	T4	T5	Mean(Std)
LPV	93	93	94	93	93	93(0)
LTI	94	94	96	96	97	95(1)
FRF	68	69	79	80	81	75(6)

Table 12: *The VAF's for each method and trial of Participant 1 using Trial 5 as estimator*

	T1	T2	T3	T4	T5	Mean(Std)
LPV	94	96	96	95	95	95(1)
LTI	98	96	96	98	98	97(1)
FRF	65	30	35	66	76	54(21)

Table 13: *The VAF's for each method and trial of Participant 2 using Trial 1 as estimator*

	T1	T2	T3	T4	T5	Mean(Std)
LPV	91	95	94	91	91	92(2)
LTI	97	98	97	97	97	97(0)
FRF	47	80	73	49	51	60(15)

Table 14: *The VAF's for each method and trial of Participant 2 using Trial 2 as estimator*

	T1	T2	T3	T4	T5	Mean(Std)
LPV	94	97	96	95	95	95(1)
LTI	97	97	97	97	97	97(0)
FRF	45	78	72	48	50	59(15)

Table 15: *The VAF's for each method and trial of Participant 2 using Trial 3 as estimator*

	T1	T2	T3	T4	T5	Mean(Std)
LPV	97	97	97	97	98	97(0)
LTI	98	96	96	98	98	97(1)
FRF	64	25	31	65	75	52(22)

Table 16: *The VAF's for each method and trial of Participant 2 using Trial 4 as estimator*

	T1	T2	T3	T4	T5	Mean(Std)
LPV	97	97	97	97	97	97(0)
LTI	98	96	96	98	98	97(1)
FRF	56	0	3	59	72	38(34)

Table 17: *The VAF's for each method and trial of Participant 2 using Trial 5 as estimator*