State Estimation for the VASIMR Plasma Engine

Shane Lynn[†§]**, John Ringwood**[†] **and Juan Ignacio Del Valle Gamboa**[*]

[†]*Department of Electronic Engineering*
National University of Ireland, Maynooth

E-mail: [§]`shane.a.lynn@eeng.nuim.ie`

[*]*Ad Astra Rocket Company*
Costa Rica

[*]`juan@adastrarocket.com`

Abstract — **This paper presents work on the application of virtual metrology techniques to the VAriable Specific Impulse Magnetoplasma Rocket (VASMIR) engine. The work concentrates on the estimation of internal temperatures of the rocket using state space models and Optical Emission Spectroscopy (OES). These estimations are useful as direct thermal measurements will not be available in the final system design.**

Keywords — **State Estimation, Virtual Metrology, Plasma Systems, VASIMR rocket, Optical Emission Spectroscopy, Principal Component Analysis.**

I Introduction

The VASIMR engine is a relatively new technology that uses a highly ionised plasma as a propelling force for space travel.

The current prototype under development uses 200 kW of power during operation. The thermal management of a system with such high levels of power is a complex and essential task for the projects success. Heat is an unfortunate byproduct of the plasma production process, and this heat must be safely disposed of to avoid damage to the apparatus.

In order to measure temperatures on points of the VASIMR prototypes, thermocouples are used. This allows the temperature of critical components to be monitored in real time. Due to design constraints, and issues with RF interference, thermocouples are not practical for temperature determination on the final engine design.

Optical emission spectrometry (OES) is an optical technique often used to gather more information on luminescent processes in a noninvasive manner. It is hoped that the optical emission from the VASIMR engine may provide information that could be used to aid temperature estimation when thermocouple data is not available.

State estimation, or virtual metrology is a methodology involving the estimation of process variables which are not directly measurable, but where related signals are available. A great deal of research is being undertaken in semiconductor manufacture and virtual metrology [1].

The research reported in this paper attempts to use state estimation techniques, combined with OES measurements, to reconstruct spatial measurements of temperatures in the VASIMR rocket engine.

This paper begins in Section II with a general introduction to the VASIMR system. Sections III, IV and V explain the basics of Optical Emission Spectroscopy (OES), the identification method used to build the models, and the state estimation techniques used, respectively. Finally, in Sections VI and VII, the collection of data is described, and results of the work are presented.

II VASIMR

The VASIMR plasma rocket engine is an alternative space propulsion device that uses a magnetically accelerated plasma as its source of propulsion. The engine uses the widely available gas, argon, as fuel, and has the advantage over traditional chemical rockets in that the specific impulse (I_{sp}) and thrust of the VASIMR engine can be varied during flight, allowing more efficient fuel consumption and minimum transit time [2].

The VASIMR technology can be viewed as three linked magnetic stages. The first stage contains a helicon antenna that ionises neutral propellant gas. The second stage uses an Ion Cyclotron Resonance Heating (ICRH) antenna to further ener-

gise the plasma, and the third stage is a magnetic nozzle, that converts the energy of the plasma into directed flow [3]. A schematic of the plasma engine is shown in figure 1.
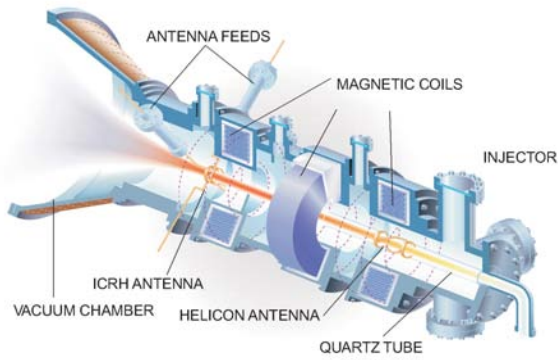

Fig. 1: Schematic of VASIMR engine [2]

The varying of the thrust and specific impulse output of the rocket engine is achieved by varying the total power distribution between the helicon plasma source, and the ICRH antenna [4].

This work concerns the helicon stage of the VASIMR. It consists of a quartz tube, with a water cooled helicon antenna surrounding it (Figure 2). Argon gas is passed through the tube at varying flow rates to be ionised by the antenna, using an RF voltage at 13.56 MHz. The tube is surrounded by electromagnetic coils, that create a magnetic field parallel to the flow of plasma. The plasma flows through the quartz containment tube into a vacuum chamber that simulates the conditions of outer space.
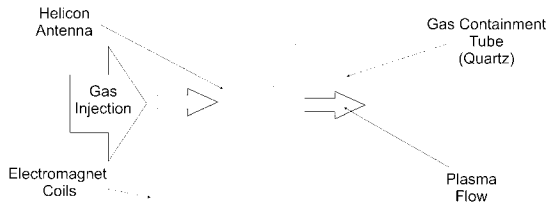

Fig. 2: Schematic of helicon section [2]

The helicon section is visualised as an independent system with three inputs; the antenna power (W), the current in the magnetic coils surrounding the helicon (A) and the flow of argon gas used (sccm). The corresponding output measurements available during experimentation are the OES spectrum taken downstream from the plasma exhaust and the outputs from thermocouples bonded to the surface of the quartz tube. This work aims to find and utilise the relationship between these outputs to assist temperature prediction in real time.

III   OPTICAL EMISSION SPECTROSCOPY

Relaxation and excitation processes in an active plasma discharge cause the plasma to luminesce. The wavelengths of light that are emitted depend on the electron transitions that occur within the atoms and molecules of the plasma. As an electron falls from a higher energy level to a lower one, a photon is released with an energy equal to the energy difference between the levels of the transition.

OES is a technique that is frequently used to monitor and diagnose plasma processes. OES spectrometers can record part of the spectrum of radiation emitted from an active discharge (usually 200-900 nm), through a window with visual access to the plasma. Data is collected in a real time and non-perturbing manner. OES diagnostics are commonly applied in the semiconductor industry [5] [6]. The amount of light recorded per sample by the spectrometer is controlled using the spectrometer integration time.

Over 2000 different wavelength signals can be recorded in every sample of OES data. Due to the computational difficulties associated with large data sets, and as many of the signals are found to be highly correlated, data reduction techniques such as principal component analysis (PCA) are regularly employed [6].

PCA can be seen as a method of transforming an original set of correlated variables into new uncorrelated variables, known as *principal components* (PCs). Each PC is a linear combination of the original variables. They are arranged in order of the variance each one explains in the original dataset [7].

Before PCA is performed on a set of data, X, made up of n samples (rows) and m variables (columns), it is usual to offset each variable to have zero mean, and sometimes to scale each to unit variance. This is useful if the original data has many different scales. Scaling to unit variance also gives all variables equal importance for the analysis. However, in the analysis of OES, many of the variables may contain noise only, and there is no advantage to giving these the same importance as variables with meaningful information. PCA performs an eigenvalue decomposition of the covariance or correlation matrix of X, which decomposes the data matrix X as the sum of the outer product of vectors t_i and p_i plus a residual matrix E. [6]:

$$X = t_1 p_1^T + t_2 p_2^T + ... + t_l p_l^T + E \qquad (1)$$
$$= TP^T + E \qquad (2)$$

where,

$$T = [t_1, t_2, ...t_l] \qquad (3)$$
$$P = [p_1, p_2, ...p_l] \qquad (4)$$

and l is the number of principal components. The vectors t_i are known as the *scores* and $T \in \mathbb{R}^{n \times 1}$ the score matrix; the p_i vectors are the *loadings* and $P \in \mathbb{R}^{m \times 1}$ the loadings matrix. For PCA, the decomposition of X is such that the loading matrix P is orthonormal and the score matrix T is orthogonal. The first PC is the linear combination of the m original variables that explains the greatest amount of variability ($t_1 = Xp_1$). In the m-dimensional variable space, the loading vector p_1 defines the direction of the greatest variance [8]. Overall, loadings represent how the original variables are combined to make the principal components, scores represent the modelled part of the data, and finally, E, the residual, represents the data that is left unrepresented by the model. For a matrix X of rank r, r principal components can be calculated. However, the first k ($k < r$) of these may be sufficient to explain the major variances in the data. If $k = dim(x)$, then $E = 0$, and the representation of the data is exact for the new variables (principal components).

IV   STATE SPACE MODEL IDENTIFICATION

Usually, state space identification is not defined with respect to known state measurements. Rather, outputs (y) and inputs (u) are specified, with the state space representation somewhat arbitrarily chosen to within a similarity transformation, where the states don't necessarily correspond to physical quantities [9]. In this work, we present a method, based on simple least squares, which allows a state space model, with provision of specific state measurements in the training set, to be determined.

It was required to identify the parameters of a state-space model of the form:

$$
\begin{aligned}
x_{k+1} &= Ax_k + Bu_k &(5)\\
y_k &= Cx_k + Du_k &(6)
\end{aligned}
$$

with $x \in \mathbb{R}^n$ representing the temperatures of the system ($n = 18$), $u \in \mathbb{R}^m$ being the inputs to the helicon ($m = 3$), $y \in \mathbb{R}^p$ representing the outputs of the system, the PCs of the OES spectra ($p = 3$)., $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{p \times n}$ and $D \in \mathbb{R}^{p \times m}$.

Data is available, during model building, for x, u and y for a number of different system excitations (see Section VI). Measurements of y and u will be available during operation and we require to determine estimates of the state vector.

For simpler operation using the state space model in feedback estimation form, D is set to zero. We determine the model as by first expand-ing (5), as:

$$
\begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix}_{k+1}
= \begin{pmatrix} a^{11} & a^{12} & \ldots & a^{1n} \\ a^{21} & a^{22} & \ldots & a^{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a^{n1} & a_{n2} & \ldots & a^{nn} \end{pmatrix}
\begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix}_k
$$

$$
\ldots + \begin{pmatrix} b^{11} & b^{12} & \ldots & b^{1m} \\ b^{21} & b^{22} & \ldots & b^{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b^{n1} & b^{n2} & \ldots & b^{nm} \end{pmatrix}
\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{pmatrix}_k \qquad (7)
$$

We can write out the first row of Eq.(7) for each value of k:

$$
\begin{pmatrix} x^1_{k+1} \\ x^1_{k+2} \\ \vdots \\ x^1_{k+N} \end{pmatrix}
= \begin{pmatrix} x^1_k & \ldots & x^n_k & \ldots \\ x^1_{k+1} & \ldots & x^n_{k+1} & \ldots \\ \vdots & \vdots & \vdots & \vdots \\ x^1_{k+N-1} & \ldots & x^n_{k+N-1} & \ldots \end{pmatrix}
$$

$$
\ldots \begin{matrix} u^1_k & \ldots & u^m_k \\ u^1_{k+1} & \ldots & u^m_{k+1} \\ \vdots & \vdots & \vdots \\ u^1_{k+N-1} & \ldots & u^m_{k+N-1} \end{matrix} \Bigg)
\begin{pmatrix} a^{11} \\ \vdots \\ a^{1n} \\ b^{11} \\ \vdots \\ b^{1m} \end{pmatrix} \qquad (8)
$$

where a total of N measurement (time) points are available. This equation is of the form:

$$
Y = \Phi\Theta \qquad (9)
$$

and has the least squares solution [10]:

$$
\hat{\Theta} = (\Phi^T\Phi)^{-1}\Phi^T Y \qquad (10)
$$

A total of n least-squares problems are solved to get all rows of A and B. A similar formulation is required to identify C.

V   STATE ESTIMATION

In order to increase the accuracy and disturbance rejection of state space models such as that described above, it is usual to include some type of feedback that allows the model error to be used to correct estimations. When directly available, state variables are used for this purpose. However, in many cases the state information will not be available during operation.

With the model obtained in Section IV, it is possible to estimate the evolution of the system states in time for a known input sequence, u_k. The accuracy of this type of estimation depends both on the accuracy of the initial conditions (ICs) provided, $\hat{x}_0$, and the accuracy of the model parameters. If the ICs are inaccurate or the model parameters, A, B and C, are incorrect, the estimated state may have a continually growing error [11].

To apply feedback to the state space equations, Equation 5 is adjusted to incorporate the error between the measured system outputs and the predicted outputs.

$$\hat{x}_{k+1} = A\hat{x}_k + Bu_k + L(y_k - C\hat{x}_k) \tag{11}$$

where $\hat{x}_k$ are the estimated states and output, and L is a proportional gain, adjusted to achieve satisfactory error performance. This is shown in Figure 3. L can be manipulated so that $\hat{x}_k$ converges to x_k regardless of the values of $\hat{x}_0$. It should be noted that without a perfect model of the system, errors cannot be reduced completely to zero. However, L can typically be chosen to that errors remain acceptably small.
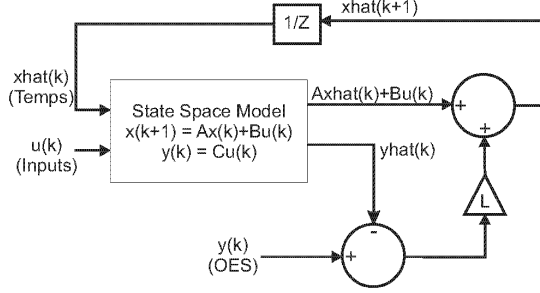

Fig. 3: State Space model in with estimation feedback

The selection of L is chosen with the error dynamics in mind. With the error defined as $e_k \equiv x_k - \hat{x}_k$, these dynamics are found by subtracting the estimate (Equation 11) from the state (Equation 5) to give [11]:

$$e_{k+1} = (A - LC)e \tag{12}$$

which has a characteristic equation:

$$\det[sI - (A - LC)] = 0 \tag{13}$$

With this in mind, we can specify the desired location of the estimator error poles for certain decay speeds and noise rejection properties as:

$$s_i = \beta_1, \beta_2, \beta_3, ...\beta_n \tag{14}$$

which yields are desired characteristic equation:

$$\alpha(s) = (s - \beta_1)(s - \beta_2)(s - \beta_3)....(s - \beta_n) \tag{15}$$

By comparing Equation 15 with Equation 13, the desired value of L can be determined.

An alternative method of determining L for a desired error response is the use of Ackermann's formula in estimator form, given by:

$$L = \alpha(A)O^{-1} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \text{ where } O = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} \tag{16}$$

and O is known as the observability matrix.

VI Data Collection and Experimentation

A large dataset is required so that parameters A, B and C of the state space model can be determined. A Design of Experiment (DOE) type experiment was carried out varying the inputs over specific regions (Table 1). The operating region was kept to a small range, in an effort to find a relatively linear operation region, and also due to equipment restrictions at the time of experimentation.

Value	High	Mid	Low
Antenna Power (W)	800	1100	1400
Magnets (A)	800	1000	1200
Gas Flow Rate (sccm)	100	-	300

Table 1: DOE Input Levels

In total, 18 experiments were run, and the temperatures were allowed to reach steady state values where possible. Temperatures were recorded from 18 thermocouples bonded to the quartz gas tube of the helicon section using ceramic paste. Thermocouples were sampled at 1Hz using a National Instruments ADC interfaced with a LabView control system for the VASIMR. Repeats of four of the experiments were carried out to assess repeatability between operational setpoints. Each setpoint appeared to have a significantly different and repeatable spectrum.

OES measurements were recorded downstream from the plasma using an Ocean Optics S2000 spectrometer, recording 2047 wavelengths from 200-850 nm at a frequency of approximately 1Hz. The integration time used for each spectrum was 200 ms, which was found to be the best choice to avoid saturation while still capturing as many lines as possible. The OES signals were first resampled at the same time instants as the thermocouples, and then represented using three PCs. These explained $\sim 97\%$ of the OES variance. As the transients in the OES PCs were faster than that of the thermocouples, difficulties were found in determining a satisfactory C matrix for the model. In order to assist this effort, the dynamic response of the spectral data was slowed down using an exponentially weighted moving average (EWMA) filter so that the OES (outputs) time constants were of a similar magnitude to those of the thermocouple (states). The equation for the EWMA was:

$$S_{k+1} = \alpha S_k + (1 - \alpha)y_k \tag{17}$$

where α was set to 0.995, S_k represents the newly slowed signal and y_k is the original OES PC signal.

Thermocouple signals with erroneous values were repaired manually where possible, and signals representing the input changes were created. Some thermocouple data was corrupted beyond repair

27

and so discarded. This corruption was caused by RF interference in some high power experiments.

The data set was used to create the state space model which was configured in closed loop estimator form as shown in Figure 3. The model has eighteen states (x) representing temperatures, three outputs (y) that are the OES PCs, and three inputs (u) representing the VASIMR inputs. The estimator transient response was set to be slightly faster than the transients found in the model's A matrix. This was achieved using Ackermann's formula.

VII   RESULTS

Here we examine some of the results from the state estimation process. The diagrams in this section display state or output predictions, and for simplicity only display two randomly chosen states or two system outputs.

a)   Model validation

An important consideration for this application is the accuracy of the model determined, so that it can be relied on for estimation purposes. As a first examination, Figure 4 shows two of the outputs of the model, when driven using the actual temperature values. The actual output is plotted on the same axis for comparison. We can see that the identification technique was relatively successful at determining the linear combination of states (C matrix) that can represent the optical PC outputs.
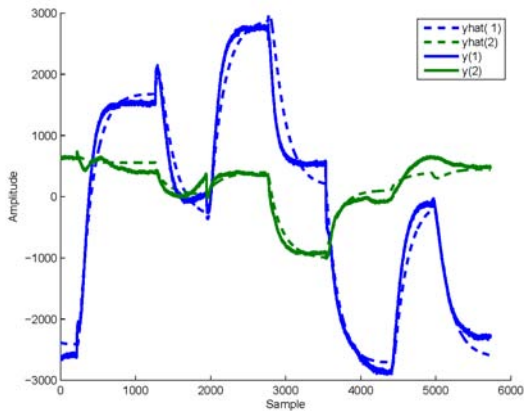
Fig. 4: Real output vs. output of model.

b)   Multi-step Prediction Performance

The open loop performance of the model is satisfactory, once precise initial conditions ($\hat{x}_0$) are provided. The state estimates and the outputs from the model are plotted in Figures 5 and 6 respectively. We can see how with exact initial conditions, the state and output estimates follow the real values with relative accuracy. The error increases over time due to model inaccuracies. With random initial conditions the estimates of temperature and state are much more inaccurate, and do not converge to the real values.
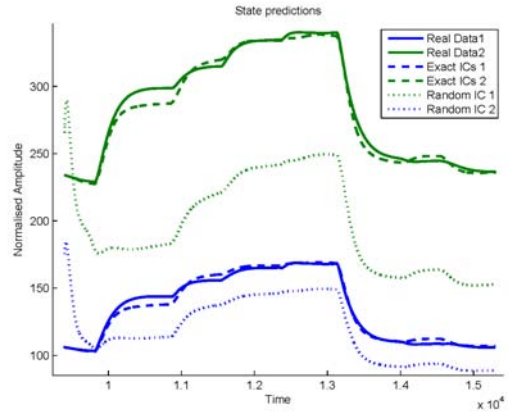
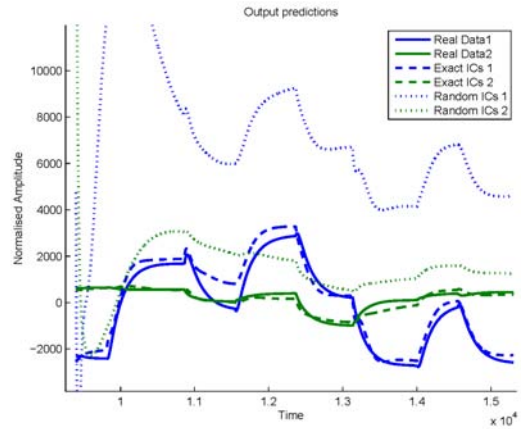Fig. 5: Open loop behaviour of model states.

Fig. 6: Open Loop behaviour of model outputs.

c)   Closed-loop Estimation

In a realistic situation, initial conditions for the system states will be unknown. Figures 7 and 8 show the model behaviour when configured in a closed loop estimation form. It can be clearly seen that both the estimated outputs and the estimated states, although starting with random initial conditions, converge towards the real values over time.

VIII   CONCLUSIONS

This work has demonstrated the successful application of a state estimator to the helicon section of the VASIMR engine. We have shown accurate (within 10 °C) estimation of VASIMR temperatures, using only OES measurements and the system inputs.
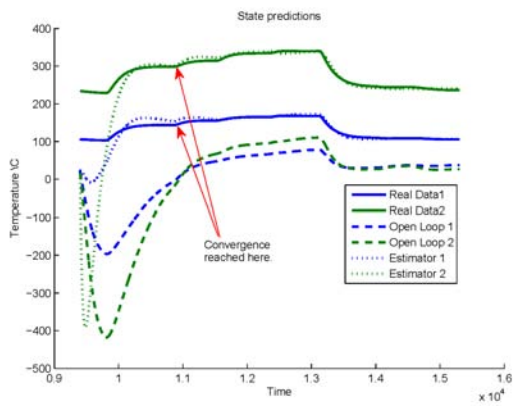
28

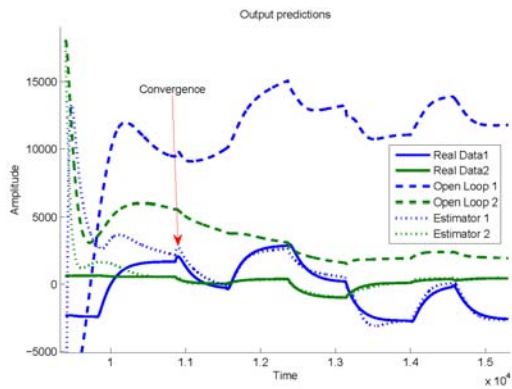Fig. 7: State Estimator performance with random ICs.

Fig. 8: Output predictions for random ICs with closed loop estimation.

This system may be used in future prototypes and final designs of the VASIMR, allowing non invasive determination of temperature, where thermocouple placement is impossible. This will provide a real time estimate of internal temperatures.

One disadvantage to the proposed system is that the full operating region needs to be explored by experiment prior to model building, so that extrapolation is avoided during operation. While the region investigated in this work was assumed to have a linear response, larger regions may not demonstrate such behaviour, requiring more advanced modelling techniques.

Further directions for work include the exploration of larger operating regions of the VASIMR, and the use of alternative heat sensors. The application of Kalman filters to improve the estimation procedure is to be examined, aiming to achieve faster convergence and reduce noise.

IX Acknowledgements

References

[1] P.H Chen, S. Wu, J. Lin, F. Ko, H. Lo, J. Wang, C.H. Yu, and M.S.m Liang. Virtual metrology: a solution for wafer to wafer advanced process control. In *Semiconductor Manufacturing, 2005. ISSM 2005, IEEE International Symposium on*, pages 155–157, 2005.

[2] F. R. Chang-Diaz, J.P. Squire, T. Glover, and A.J. Petro. The VASIMR engine: Project status and recent accomplishments. In *42nd Aerospace Sciences Meeting and Exhibit*, 2004.

[3] F. R. Chang-Diaz. The VASIMR rocket. *Scientific American*, 283(5):90–97, November 2000.

[4] T. Glover, G. F. Franklin, J.P. Squire, Jacobson V.P., Chavers D.G., and Carter M.D. Principal VASIMR results and present objects. In *Space Technology and Applications International Forum*, 2005.

[5] R. Chen, H. Huang, C. J. Spanos, and M. Gatto. Plasma etch modeling using optical emission spectroscopy. *Jrnl of Vacuum Science & Technology A*, 14(3):1901–1906, May 1996.

[6] H. H.Yue, S. J. Qin, J. Wiseman, and A. Toprac. Plasma etching endpoint detection using multiple wavelengths for small open-area wafers. *Journal of Vacuum Science & Technology A*, 19(1):66–75, Jan/Feb 2001.

[7] A. Afifi, V. A. Clak, and S. May. *Computer Aided Multivariate Analysis*. Chapman & Hall/CRC, 2004.

[8] J. Zhang, E.B. Martin, and A.J. Morris. Fault-detection and diagnosis using multivariate statistical techniques. *Chemical Engineering Research & Design*, 74(1):89–96, January 1996.

[9] W. Favoreel, B. De Moor, and P. Van Overschee. Subspace state space system identification for industrial processes. *Jrnl. of Process Control*, 10(2):149–155, 2000.

[10] D. C. Montgomery, G. C. Runger, and N. F. Hubele. *Engineering Statistics*. John Wiley & Sons, Inc., 2nd edition, 2001.

[11] G. F. Franklin, J. D. Powell, and A. Emami-Naeini. *Feedback Control of Dynamic Systems*. Addison-Wesley Publishing Company, 1994.

29