Technical Note TN/PCY/5 (2018)

Efficiently Parameterised Models of the Silverbox Nonlinear Circuit

Peter C. Young

Systems and Control,
Lancaster Environment Centre, Lancaster University, UK.

15/5/2018

Abstract

This Technical Note (TN) considers the statistical identification and estimation of the well known Silverbox nonlinear circuit; a real, physical system which has been used for some years as one of the benchmarks for nonlinear modelling in the control and systems community. Unlike a previous TN, whose aim was to identify a model that was as simple as possible for applications such as control system design, the main objective this TN is to identify a model that explains the measured data as well as any other models of the Silverbox that have been proposed heretofore. In doing so, it also addresses the question of an appropriate and efficient parameterisation of the model. The appropriate selection of a model form is an important aspect of the *Data-Based Mechanistic* (DBM) approach to modelling dynamic systems. This is the approach used as a basis for modelling in the TN and it raises questions about whether the most appropriate model of the Silverbox system is 'black box', 'grey box' or of a DBM form. Not surprisingly, given the physical nature of the system (and the author's known views on data-based modelling), it is suggested that an efficiently parameterised differential equation DBM model of the Silverbox has advantages in this regard, reflecting previous experience of DBM modelling in diverse areas of study.

1 Introduction

A previous technical note (Young, 2016b) has discussed the *State-Dependent Parameter Transfer Function* (SDPTF) modelling and control of the Silverbox nonlinear circuit system, concentrating on the simplest, continuous-time differential equation model of this type that provides a good description of the Silverbox experimental data. This has only 5 parameters and relates directly to the originally prescribed differential equation model of the Silverbox. While this model provides a reasonable basis for the initial consideration of automatic control system design, we will see that alternative differential equation models are able to provide much more accurate descriptions of the Silverbox experimental data.

This development of improved models was stimulated by the publication of the paper by Nöel and Schoukens (2018) which investigates a 'grey-box' state space representation of the Silverbox and suggests that the experimental data is best described by a 13 parameter, discrete-time *State Space* (SS) model. The 5 parameter SDPTF model explains the data quite well, with a coefficient of determination, based on the simulated model output, of $R_T^2 = 0.994$, i.e. the simulated model output explains 99.4% of the Silverbox output variance. Such a large R_T^2, very close to unity, is normally adequate for many practical purposes. However, the 13 parameter SS model, as well as other more efficiently parameterised 6 and 8 parameters SDPTF models that are identified in this note, substantially improve the explanation of the data, with an exceptional $R_T^2 = 0.99998$.

So what is the reason for this considerable difference between the explanatory ability of the 5 parameter model and that of these other models? Is it just the additional parameters providing a better description of the nonlinearity in the system, as implied by Nöel and Schoukens (2018)? Or is there another explanation that relates to the ways in which the models are parameterised? The present note attempts to answer these questions and shows that a 6 parameter SDPTF model, only marginally different from the original 5 parameter model, is able to explain the data as well as the 13 parameter SS model.

This raises interesting questions about the parameterisation of models for the Silverbox, in particular, and the modelling of nonlinear systems in more general terms using 'black box', 'grey box' or 'data-based mechanistic' models. By 'parameterisation' I mean here the number and location of the parameters in the model, so that the term 'efficiently parameterised' in the

title of this note means that the model is explained by as few parameters as are necessary to explain the data to a level defined by the modelling objectives.

2 The Silverbox Data Set

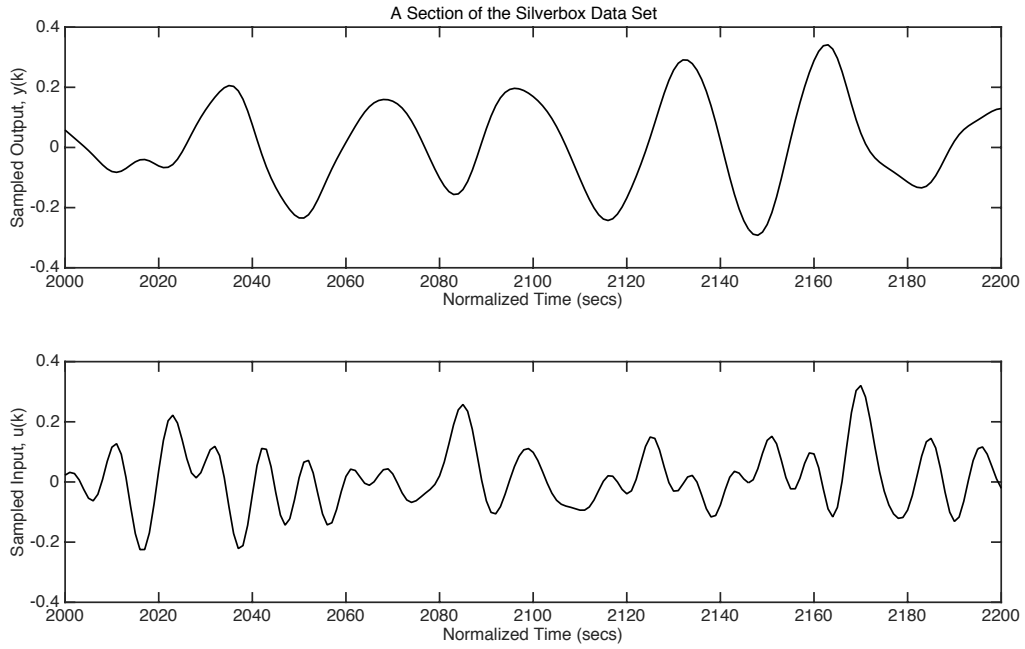

Figure 1: A section of the Silverbox sampled data set plotted in normalized time (actual sampling interval: $\Delta T = 1/2441$ secs.)

The Silverbox data analyzed in this note are taken from those utilized in the paper by Nöel and Schoukens (2018). In particular, they based on the first 20,000 of the 245,760 samples kindly provided by Jean-Philippe Nöel. The first 10,000 samples are used for model identification/estimation and the second 10,000 samples for validation purposes. For optimization purposes, most of the model parameter estimation results were obtained with a criterion function based on the error over the last 7500 samples, in order to ensure that initial condition transients, that have a considerable effect in this example because of the oscillatory dynamics and low damping, do not affect the estimation. In fact, almost exactly the same estimation results

were obtained if the initial conditions were added as additional parameters and optimized simultaneously with the model parameters, thereby allowing for the use of all 10000 samples (see later section 4.3.1).

These data were obtained from experiments on the Silverbox Benchmark system, where the system was excited using random phase multisines (see e.g. Pintelon and Schoukens, 2001) with root-mean-squared (RMS) amplitudes of 5 and 100 mV. The input frequency spectrum is limited to 0-300 Hz, excluding the DC component; and the sampling frequency is 2441 Hz, i.e. a sampling interval of $\Delta t = 1/2441 = 0.00040967$ secs.[1] However, for simplicity in the present analysis, as well as to ensure good numerical conditioning in the computations and associated Simulink$^{\text{TM}}$ simulations, the sampling interval is normalized to unity. A section of the sampled input, $u(k)$, and output, $y(k)$, of these data is plotted with this normalized time axis in Fig.1. It is, of course, straightforward to adjust the continuous-time model estimation results later, taking into account the actual sampling interval.

Note that this temporal normalization is not unusual in DBM modelling, as outlined in the next section 3, because the initial analysis is carried out in a completely 'black-box' manner, in order to avoid any prejudicial assumptions about the nature of the model. The physical interpretation of the model, if required by the objectives of the study, takes place *after this* in order to ensure the credibility of the model to those for whom the model is intended. This is because the recipient of the model who requires a description that can be interpreted in physically meaningful terms (and normally that match the model forms used by the community in which the recipient resides), may well not wish to be confronted by a black box model whose internal functionality is unclear or even opaque. Of course this does not mean that black-box models are not perfectly adequate for some applications (see e.g. some of the practical examples in Young, 2011) and, indeed, can provide a preferable solution in some circumstances.

3 A Brief Review of DBM modelling

DBM modelling is a very general approach to the identification and estimation of stochastic, dynamic models (see Chapter 12 in Young, 2011, and the prior references therein) that follows a *Hypothetico-Inductive* DBM (HI-

[1] These are not the standard Silverbox benchmark data (see e.g. Wigren and Schoukens, 2013; Marconato et al., 2012), where the sampling frequency is less (610.35 Hz).

DBM) philosophy (Young, 2013). This is in contrast to most mechanistic modelling procedures that follow the well known 'hypothetico-deductive' philosophy, as described most clearly by the scientific philospher Karl Popper Popper (1959). Here, the hypothesis, in the form of the mechanistic (or 'conceptual') model structure, is used as the main basis for model identification and estimation. Most often, this involves the optimization of the parameters that characterise the assumed conceptual model structure in order to minimise some specified criterion function.

In contrast to this, the hypothesised model is simply 'borne in mind' in the first stages of HI-DBM modelling so that it is not allowed to dominate the process of statistical inference; namely the data-based *identification* of the model structure and the statistical *estimation* of the parameters that characterise this structure. This process of statistical identification and estimation is purely inductive: the model is identified on the basis of the information in the available data (in all its forms but consisting mainly of measured time series). Only when this is complete is the hypothesised model structure taken into account in order to see if it conforms to that of the identified, data-based model. If so, it is allowed to influence the proceeding stages of model evaluation and development. In this way, HI-DBM modelling avoids the possibility of any prejudice in the selection of the model, as well as the inadvertent choice of models that are more complex than they need to be in order to explain the observable data. Most importantly, it helps to avoid the *over-parameterisation* of the model that afflicts many modelling exercises in the biological, environmental and physical sciences (see e.g. Young, 2018, for a climate modelling example; and some of the models that have been proposed for the Silverbox system, as mentioned later in section 6). It also ensures that the resulting parameter estimates are well defined, *with good estimates of the uncertainty that characterises all models estimated from real data*.

The latest form of HI-DBM modelling (see e.g. Young, 2016a) involves 10 stages of model development. In the case of the Silverbox data, however, only four of these stages are necessary and a description of these, suitably expanded in order to make them particularly relevant within the present context, follows below.

- **Clearly define the modelling objectives**. There is rarely only one possible model of a dynamic system and it is necessary to aim for a model which best suits these objectives. For instance, very large

models are required to describe the mechanisms for the variations in surface temperature over the planet but these models are not the best models for forecasting globally averaged surface temperature changes (as discussed in Young, 2018).

- **Consider whether a linear model will sufficiently explain the data**. If this is not satisfactory, then the model may require time variable parameters; or it may need to be nonlinear. The dtfm routine in the CAPTAIN Toolbox[2] for Matlab™ allows for the investigation of the first possibility; while the sdp routine[3] in CAPTAIN can be used to identify the location and graphical nature of the *State Dependent Parameter* (SDP) nonlinearities in a SDP *Transfer Function* (SDPTF) model.

- **Initial SDPTF model estimation**. The SDPTF model can be either discrete or continuous-time, whichever best suits the nature of the system and the objectives of the study. An initial model order and structure for the SDPTF is often that suggested by linear model analysis. The graphical form of the SDP nonlinearity is defined by any significant state dependent variation of the parameters in the TF model, where these are estimated in a transformed space that converts the rapidly time variable variable SDPs into slowly varying parameters that are estimated by lag-free, recursive *Fixed Interval Smoothing* (FIS). The constant 'hyper-parameters' associated with a *Generalized Random Walk* (GRW) model of the parametric time variation are optimized by *Maximum Likelihood* (ML) using prediction error decomposition (for details of this approach, see Young, 2011, and the prior references therein). Effectively, the SDP estimate is a nonparametric estimate since it exists at every sample in the time series and can be plotted against the state variable or variables on which the SDP is dependent (state variables that are considered in a very general sense but usually based on the input and output variables or other variables that make physical sense).

- **On the basis of the SDP estimation results, identify a suitable parametric model for each SDP**. This should represent the

[2]This can be downloaded free via the web-site `http://captaintoolbox.co.uk/Captain_Toolbox.html/Captain_Toolbox.html`.

[3]The sdp routine is currently limited to single state-dependency

graphical, nonparametric nonlinearity sufficiently well and any method of parameterisation can be utilized: for instance a specific, low dimensional parametric representations of the SDP can be utilised if the identified shape of the SDP is simple and well defined: e.g. a simple power law can often suffice in models of an apparently complex river catchment (see e.g. Young, 2013). But more complex representations are required if the identified shape is more complicated or not too well defined. These could include polynomial expansions or combinations of other basis functions: e.g. radial basis, Takagi-Sugeno Fuzzy inference, etc. (see e.g. Beven et al., 2011). Complete black box models, such as neural networks, are sometimes utilised but these are often inappropriate because they do not reveal sufficiently, if at all, the nature of the nonlinearity. Whatever parametric representation of the SDPs is selected, it is used to form a complete, constant parameter, nonlinear SDP model that is suitable for final constant parameter estimation. This model can be in any appropriate form, such as the SDPTF defined by the identified constant and state dependent parameters; or, if it is more appropriate, an equivalent state space representation of this. In this latter situation, however, care must be taken to avoid the over-parameterisation that is inherent in any SS model that is not specified in a selected canonical (minimally parameterised) form: see e.g. Ljung (1987); Wigren (2006).

- **Initially estimate the parameterised nonlinear model in a suitable manner**: e.g. in the present context, a special 'quasi-linear' form of *Refined Instrumental Variable estimation for Continuous-time TF models* (RIVC) estimation can be used, exploiting the CAPTAIN rivcbjid routine (or its discrete-time equivalent rivbjid), where the nonlinearities are treated as additional inputs.

- **Improve the initial parameter estimates using optimal estimation**: e.g. *Nonlinear Least Squares* (NLS) optimization, initiated by the quasi-linear parameter estimates (or some other source if there are problems with the quasi-linear estimation). Finally, if there is any autocorrelation (colour) in the residuals from this optimiized model that is significant enough to justify a stochastic description (e.g. an ARMA process), *and* if analysis of the residuals then suggests it is worthwhile, re-optimize the model using a stochastically optimal method such as

ML or *Prediction Error Minimization* (PEM).

An interesting previous example of DBM modelling that provides some background to the DBM approach used in present Silverbox example is the data-based modelling of the hydraulic actuator controlling a robot arm. This was considered by Sjöberg et al. (1995) and Hu et al. (2001) and I have discussed it within a DBM context (Young, 2001a), as well as in relation to 'fuzzy' black-box models (Young, 2006). In particular, these latter references show how the DBM approach to modelling nonlinear systems and its use of *State-Dependent Parameter* (SDP) estimation, as discussed in the next section 4, can overcome some of the disadvantages of purely black-box models.

4 DBM State-Dependent Parameter (SDP) Models

SDP estimation is an approach to modelling nonlinear dynamic systems that evolved many years ago (Young, 1968; Mendel, 1969; Young, 1981) and Priestley (1980), who first used the name. The SDP methods used in the the present technical note were developed in the 1990s (see e.g. Young, 2000, 2001b) and, since then, these have been applied successfully to a large number of practical systems, in diverse areas of study.

In their paper describing the Silverbox system, Wigren and Schoukens (2013) actually define the initially hypothesised conceptual model in a state dependent parameter differential equation model form that mimics the behaviour of a mechanical system, namely:

$$m\frac{d^2x(t)}{dt^2} + d\frac{dx(t)}{dt} + k\{x(t)\}.x(t) = u(t)) \tag{1}$$

where $x(t)$ is the model output and,

$$k\{x(t)\} = a + bx^2(t) \tag{2}$$

is the state dependent parameter (although they do not appear to be aware of the term and refer to it, correctly in this mechanical engineering context, as 'a static but position-dependent stiffness'). Using the alternative nomenclature preferred in this note, the associated SDPTF model would be represented as

follows:

$$\begin{aligned} x(t) &= \frac{b_0}{s^2 + a_1 s + a_2\{x(t)\}} u(t); \\ y(k) &= x(k) + \xi(k) \end{aligned} \tag{3}$$

where $a_1 = \frac{d}{m}$; the SDP $a_2\{x(t)\} = \frac{k}{m}\{x(t)\} = a + bx^2(t)$; $b_0 = \frac{1}{m}$; and, in this hybrid continuous-time setting (Young, 2015), $y(k)$ is the sampled value of $x(t)$ at some uniform sampling interval Δt (here 1/2441 secs.), while $\xi(k)$ represents any modelling error and/or additive measurement noise.

In the previous technical note (Young, 2016b), the SDPTF model is identified directly from the Silverbox data in this same form, where now the SDP is a little more complicated, i.e.,

$$a_2\{x(t)\} = k_1 + k_2 x(t) + k_3 x^2(t) \tag{4}$$

This model explains the data well and is in a form that can be used for SDP control system design, which was the primary objective of this earlier technical note.

However, the objective of modelling in the present note is not the use of the model for automatic control system design. Rather it is:

(i) to identify a model that explains the measured data as well as possible with the minimum number of model parameters;

(ii) produce a model that has a mechanistically interpretable form and makes sense in relation to the known nature of the Silverbox system.

In other words, the aim is to produce a mechanistically meaningful model that is rich in its explanatory ability but avoids over-parameterisation. As we shall see, this means that the SDPTF model is identified in two forms, one with 6 parameters and the other with 8 parameters. The former is only a little different from earlier model, with the additional constant parameter appearing in a changed numerator $b_0 s + b_1$ that, surprisingly perhaps, makes a significant difference to its explanatory ability. The latter, on the other hand, takes the more complex form:

$$\begin{aligned} x(t) &= \frac{b_0 s + b_1}{s^2 + a_1\{x(t)\} s + a_2\{x(t)\}} u(t) \\ a_2\{x(t)\} &= k_1 + k_2 x(t) + k_3 x^2(t) \\ a_1\{x(t)\} &= k_4 + k_5 x(t) + k_6 x^2(t); \\ y(k) &= x(k) + \xi(k) \end{aligned} \tag{5}$$

where both SDPTF denominator parameters are quadratic functions of the model output. But which of these is preferable in DBM terms? This is discussed later in sub-section 4.4.

With the above objectives and results in mind, this section is concerned with *the initial stages of model structure identification*, where linear models are rejected because they do not explain the data very well; and time variable parameter models are rejected because the rate of estimated variation is similar to that of the measured variables, suggesting they are probably state dependent. This leads to the first, non-parametric identification stage of SDP modelling that confirms the probability of state-dependency and defines the possible shape and location of the state dependent parameters in the second order nonlinear differential equation model.

4.1 Linear, TVP and initial SDP Model Evaluation

The first step in DBM modelling is to evaluate how well the Silverbox data can be explained by a linear model, as revealed in the results below provided verbatim by the rivcbjid routine in CAPTAIN.

```
------------------------- BEST 12 models (BIC) --------------------------
den num del AR  MA    YIC      Rt2         BIC          S2         condP
 2   3   0   0   0  -2.2779   0.674825   -46980.6664   9.071e-03   6.748e-01
 2   1   0   0   0  -9.2386   0.673969   -46972.7806   9.095e-03   6.740e-01**
 2   2   0   0   0  -3.2534   0.674216   -46971.1578   9.088e-03   6.742e-01**
 3   1   0   0   0  -1.5571   0.674195   -46970.5183   9.089e-03   6.742e-01
 4   1   0   0   0  1.4670    0.674433   -46968.6055   9.082e-03   6.744e-01
 3   4   0   0   0  2.7392    0.674624   -46956.0502   9.077e-03   6.746e-01
 3   2   0   0   0  -4.3315   0.673946   -46953.6521   9.096e-03   6.739e-01
 3   3   0   0   0  -2.9278   0.674180   -46951.6266   9.089e-03   6.742e-01
 4   2   0   0   0  -3.2723   0.673513   -46931.1727   9.108e-03   6.735e-01
```

While these results show that linear models do not provide a very good explanation of the data, with $R_T^2 < 0.7$, they provide some initial information about the order of the system that is useful in subsequent stages of the identification process. In particular, the two second order models with [2 1 0] and [2 2 0] structures (marked with '**') are identified well according to the YIC (see e.g. Young, 2011), and the [2 3 0] model is marginally favoured by the BIC (see e.g. Priestley, 1981). Of the higher order models considered above, the [3 2 0] and [3 3 0] structures provide reasonable results but the additional parameters do not yield any improvement in explanatory ability. Based on these results and considerations of parsimony, it seems reasonable to consider the second order models further.

The relatively poor explanation of the data by the linear model suggests that the model may require time variable parameters or is nonlinear. In the first *Time Variable Parameter* (TVP) case, the dtfm routine in CAPTAIN can be used to investigate if a TVP model is able to improve the explanation of the data and, if so, which parameters in the model appear to be varying most and in what manner. This is not an essential step in the identification process but, as we see below, it can be of assistance. Also, note that the dtfm routine is based on a discrete-time transfer function model and this needs to be taken into account when evaluating the results.

As regards the Silverbox, it will suffice to say that the dtfm routine detects and estimates considerable, rapid variations in the denominator parameters of the [2 2 0] TF which substantially increase R_T^2 to 0.999; but there is little sign of systematic variation in the numerator parameters, so they are constrained to be constant. Moreover, if the estimated rapid variations in the denominator parameters are plotted against the sampled output variable $y(k)$, as shown in Fig.2, then they suggest the denominator parameters are dependent on $y(k)$, so that it is best to consider the model in a nonlinear SDP form.

The possibility that the system is nonlinear could be checked using separate nonlinearity tests (see e.g. Billings, 2013) but this is not carried out here because SDP estimation provides an alternative to such tests. In particular, it can be used to investigate whether there are clear indications of SDP nonlinearity and, if so, where this nonlinearity resides in the model equation. SDP estimation proceeds by stages, at first using the sdp routine in CAPTAIN. In this case, where the model is formulated in continuous-time terms, this has to be preceded by differentiation of the input and output variables using the optimized *Fixed Interval Smoothing* (FIS) approach implemented in the irwsm routine of CAPTAIN, which helps to avoid the usual problems associated with the differentiation of potentially noisy signals (see e.g. Young et al., 1993; Coca and Billings, 1999; Young, 2011; Janot et al., 2017). This works particularly well in the case of the Silverbox data because there appears to be little noise on the measured signals.

The above TVP results have suggested the parameters are dependent on the output variable $y(t)$ and initial SDP estimation analysis confirms this. The sdp analysis in this case is applied to the [2 2 0] SDPTF model and allows for the possibility of state-dependency in all of the four model parameters. The graphical results shown in Fig.3 are generated from the results of unconstrained SDP estimation, with all the four parameters assumed to be

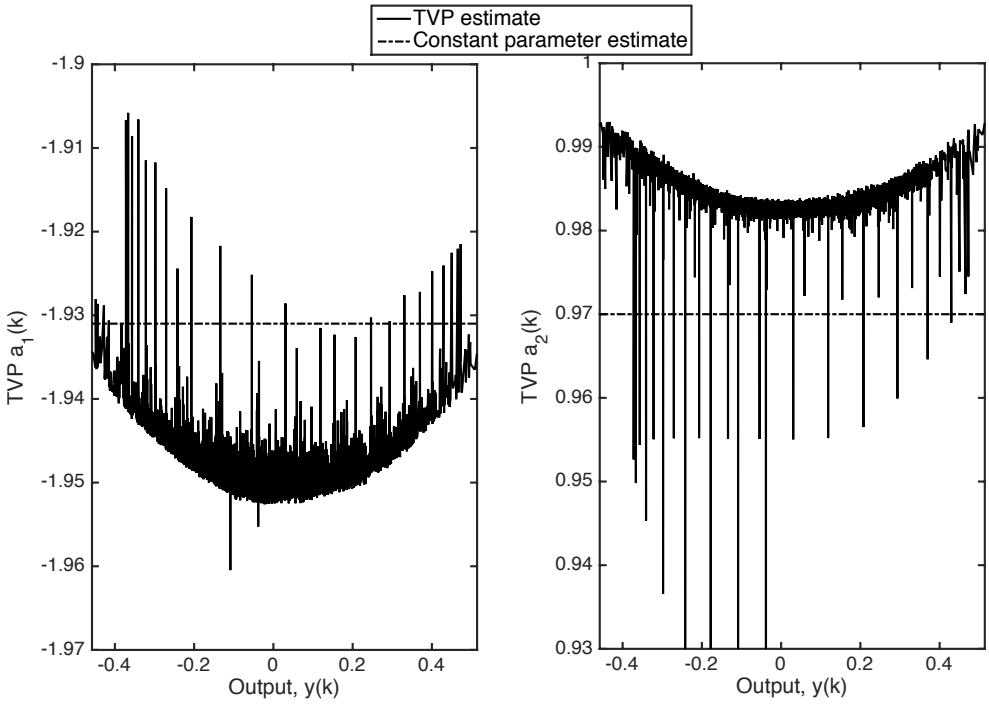

Figure 2: TVP estimates of the denominator parameters in a [2 2 0] discrete-time model, suggesting possible state-dependency.

dependent on $y(t)$. As in the case of the TVP estimation, these plots reveal that there is no significant state-dependency associated with the numerator parameters b_0 and b_1, where the estimated parameters vary a little because the estimation is unconstrained and there is some interaction between the estimates; but there is no large, systematic variation in relation to the constant parameter estimates shown by the red lines. In contrast to this, there appears to be clear state-dependency in the estimates of the a_1 and a_2 parameters, where both show systematic variations away from the red line, suggesting that they are indeed dependent on the output $y(t)$. Notice, however, that the estimated confidence bounds are somewhat larger for $a_1\{y(t)\}$ than for $a_2\{y(t)\}$ and the higher frequency 'noise' on the estimate is greater, so the nonlinearity in this case may not be so well defined, as we see later.

The graphical portrayal of state-dependency can be in three main forms: the SDP estimate itself, as plotted in Fig.3; a plot of the full state dependent nonlinear term in the TF model, e.g. $a_2\{y(t)\} \times y(t)$ in the case of the

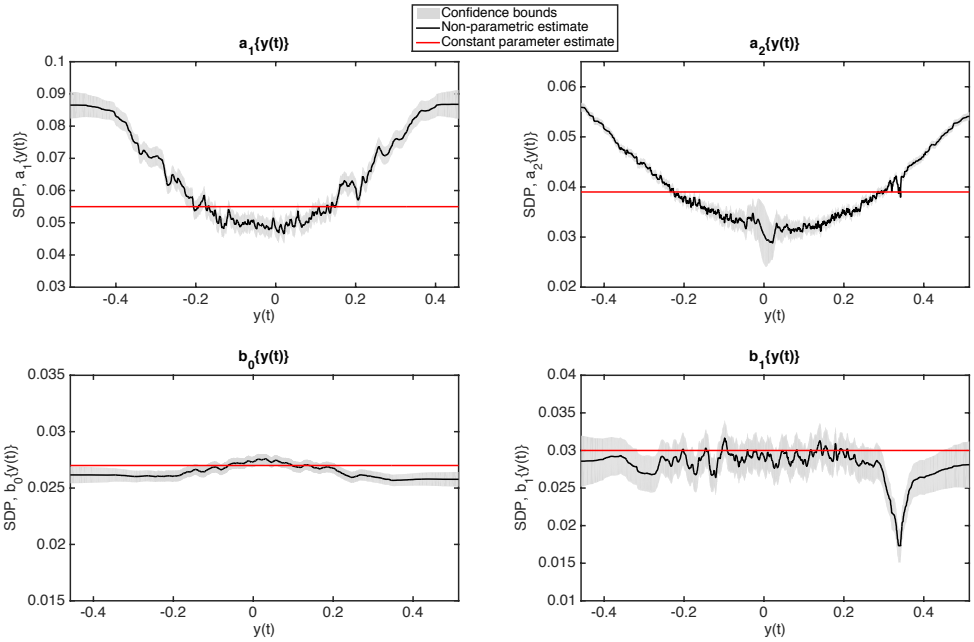

Figure 3: Unconstrained nonparametric estimates of the state dependent parameters in the [2 2 0] SDP model. The red lines show the estimates obtained by the sdp routine assuming constant parameters.

second denominator SDP; or a plot of just the nonlinear terms appearing in the parameterised SDP. These three plots for the second denominator SDP $a_2\{y(t)\}$ are shown in Fig.4. In this figure, the red lines are parameterisations of the curves obtained by linear least squares estimation based on an assumed parametric form, in this case that the SDP $a_2\{y(t)\}$ varies in a quadratic manner with $y(t)$, i.e.,

$$a_2\{y(t)\} = k_1 + k_2 y(t) + k_3 y^2(t) \tag{6}$$

where the least squares estimates of the parameters, i.e.,

$$\begin{aligned} &\hat{k}_1 = 0.0318(0.000011); \hat{k}_2 = -0.0086(0.000052) \\ &\hat{k}_3 = 0.000239(0.00024) \end{aligned} \tag{7}$$

are obtained based on the sorted data used in the SDP estimation (see e.g. chapter 11, section 11.2.1 in Young, 2011). These estimates also define the cubic parameterisation of the dependent nonlinearity $a_2\{y(t)\} \times y(t)$ shown in

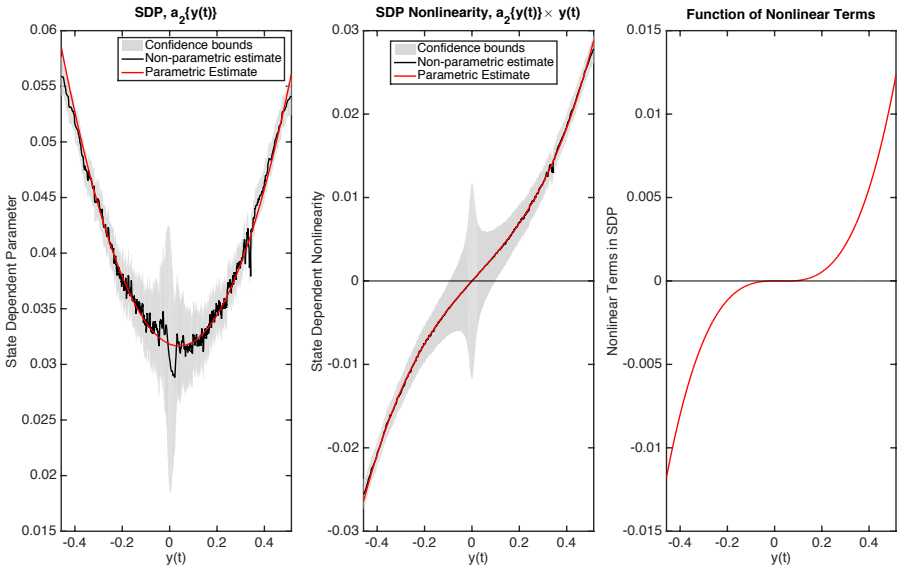

Figure 4: Nonparametric estimates of the SDP $a_2\{y(t)\}$ (left panel) and the state dependent nonlinearity $a_2\{y(t)\} \times y(t)$ associated with $a_2\{y(t)\}$ (centre panel). The red lines are the parameterisations of the non-parametric estimates. The right panel shows the extracted combination of quadratic and cubic nonlinear terms from the parameterisation in the centre panel.

the middle panel of Fig.4; while the parameterisation of extracted nonlinear terms is given by

$$\hat{k}_2 y^2(t) + \hat{k}_3 y^3(t) \tag{8}$$

It is notable how this function appears quite flat over a fairly large section either side of $y(t) = 0$. In fact, as we see later, this region is not flat but has small but interesting curvature.

From the initial SDP modelling discussed above, we can conclude that a possible SDP model takes the following second order form:

$$y(t) = \frac{b_0 s + b_1}{s^2 + a_1\{y(t)\}s + a_2\{y(t)\}} u(t) + \xi(k) \tag{9}$$

or in differential equation form:

$$\frac{d^2y(t)}{dt^2} + a_1\{y(t)\}\frac{dy(t)}{dt} + a_2\{y(t)\}y(t) = b_0\frac{du(t)}{dt} + b_1 u(t) + \mu(t) \tag{10}$$

where $\mu(t)$ is potentially a complex, nonlinear stochastic process.

4.2 Initial Quasi-Linear Estimation of the Nonlinear SDP model

Optimization of SDP models in the next section 4.3 shows that the apparent nonlinearity in the a_1 parameter does not appear to add all that much to the descriptive ability of the SDPTF model and so, at first, it can be assumed to be a constant parameter without any major loss in explanatory power. Taking this into account, the SDPTF model defined in the previous subsection takes the following form:

$$\begin{aligned} y(t) &= \frac{b_0 s + b_1}{s^2 + a_1 s + a_2\{y(t)\}} u(t) + \xi(k) \\ a_2\{y(t)\} &= k_1 + k_2 y(t) + k_3 y^2(t) \end{aligned} \tag{11}$$

or, in differential equation terms:

$$\frac{d^2y(t)}{dt^2} + a_1 \frac{dy(t)}{dt} + a_2\{y(t)\}y(t) = b_0 \frac{du(t)}{dt} + b_1 u(t) + \mu(t) \tag{12}$$

If this model is written in the extended form:

$$\frac{d^2y(t)}{dt^2} = -a_1 \frac{dy(t)}{dt} - k_1 y(t) - k_2 y^2(t) - k_3 y^3(t) + b_0 \frac{du(t)}{dt} + b_1 u(t) + \mu(t) \tag{13}$$

we see that it could be considered as a model with two inputs $y^2(t)$ and $y^3(t)$ in addition to the input $u(t)$. As a result, it can be estimated as a quasi-linear *Multi-Input Single-Output* (MISO) model (see also Laurain et al., 2010, where it is referred to as a LPV-type model) and can be identified as a nominally linear model by RIVC estimation, using the rivcbjid routine in CAPTAIN. However, this must be carried out with introspection and caution because we note that the additional nonlinear inputs are computed from the *measured* output $y(t)$ and are not, as they should be, a function of the internal noise-free output $x(t)$ in the model. So we might expect residuals that have high serial correlation, as well as correlation with the input $u(t)$; and if the noise on $y(t)$ is large, then the estimates of the parameters could be significantly biased. Nonetheless, this quasi-linear approach is quick and easy. And because the noise on the Silverbox measurements is very small, it provides useful insight into the system characteristics although, as we shall see, no real insight into one important aspect of the model; namely, the nature of the numerator dynamics.

The verbatim results produced by rivcbjid for the quasi-linear model (13), using the normalised sampling interval of one second, are as follows:

```
-------------------- BEST 2 models (Rt2) --------------------
den num del AR  MA    YIC      Rt2         BIC          S2        condP
 2   2   0   0   0  -5.4140   0.984206   -77219.0093   4.406e-04   9.842e-01
     1   0
     1   0
 2   1   0   0   0  -11.9715  0.984201   -77224.5816   4.407e-04   9.842e-01
     1   0
     1   0
```

where the results for both the [2 1 0] and [2 2 0] model are presented to show that there appears to be little to choose between the two model structures, with both the YIC and BIC favouring the [2 1 0] structure. For this reason and considerations of parametric efficiency (parsimony), therefore, the previous analysis in Young (2016b) considered only the simpler [2 1 0] model. However, for reasons that will become clear later, we will continue the analysis here with the [2 2 0] model. The rivcbj estimated quasi-linear model can be written in the following 'hybrid' form:

$$\frac{d^2y(t)}{dt^2} = -0.0180\frac{dy(t)}{dt} - 0.03155y(t) + 0.0004818\frac{du(t)}{dt} + 0.03116u(k) - 0.007138y^2(k) + 0.1244y^3(k) + \mu(t) \tag{14}$$

where $y^2(k)$, $y^3(k)$ and $u(k)$ are presented with the index k to emphasise that these are the sampled variables used as inputs in the rivcbj estimation. The TF denominator $A(s) = s^2 + 0.0180s + 0.03155$ and, based on the true sampling interval of 1/2441 seconds, the steady state gain, damping and natural frequency are $G = 0.988$, $\zeta = 0.051$ and $\omega_n = 69$ Hz, respectively. This model explains the data reasonably well with $R_T^2 = 0.984$[4]. For the purposes of control system design, as considered in the first report (Young, 2016b), the model identification and estimation could probably stop at this stage with this model utilised as the basis for the analysis. However, for the present purely modelling purposes, it is essential to proceed with more formal optimization analysis.

4.3 Final Nonlinear Model Optimization

Given the nature of the silverbox data and the model results presented in the previous sub-section, the final model parameter estimation is based on the model shown in the Simulink diagram, Fig.5, where xh=$\hat{x}(t)$ is the simulated

[4]Note that these results were mis-reported in Young (2016b), which showed a larger R_T^2.

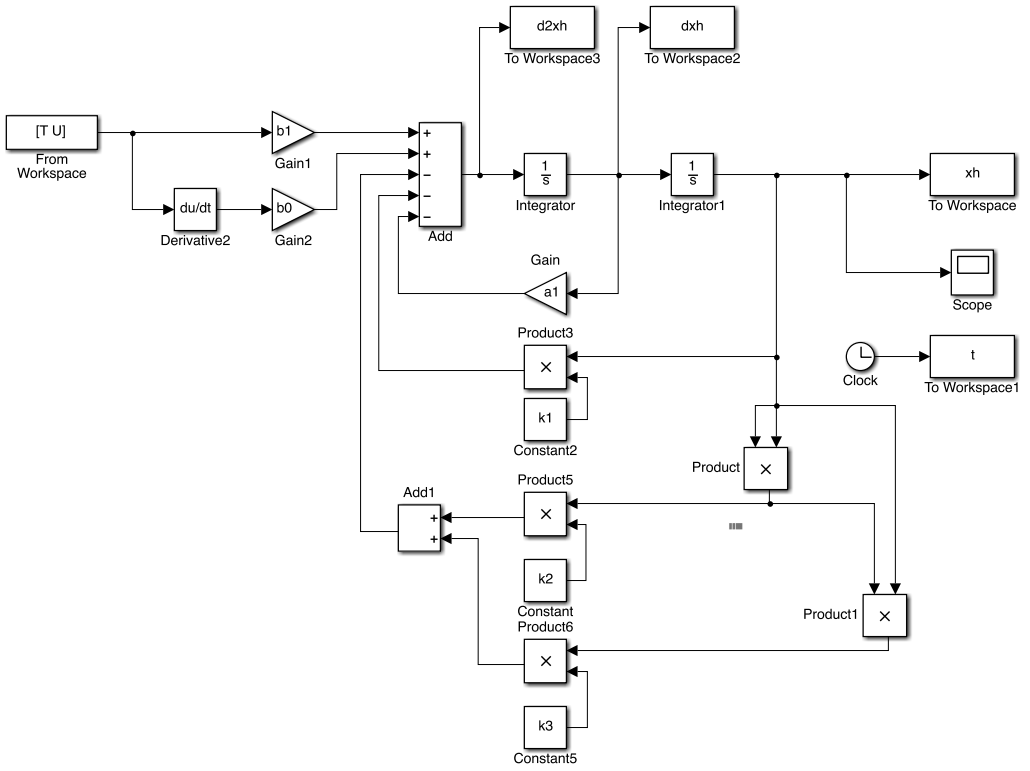

Figure 5: The Simulink model used in the final model parameter estimation using NLS.

deterministic model output; while dxh=$\frac{d\hat{x}(t)}{dt}$ and d2xh=$\frac{d^2\hat{x}(t)}{dt^2}$ are its first and second derivatives, respectively. This model is employed as the basis for estimating the nonlinear model parameters using NLS, exploiting the Matlab lsqnonlin optimization routine. An alternative is stochastic PEM, with the discrete-time residual $\xi(k)$ modelled as an ARMA noise process (see Young, 2016b). This is considered in section 4.3.2, where we see that the PEM optimization makes very little difference to the estimated model parameters and the nature of the model but does allow for improved uncertainty analysis.

4.3.1 Nonlinear Least Squares Optimization

The primary advantage of this approach over the multiple input, quasi-linear estimation outlined in the previous section is that, here, the nonlinear terms are generated within the model, and so *based on the noise-free model output,*

not the noisy measured output. The associated model takes the form:

$$\begin{aligned} \frac{d_2\hat{x}(t)}{dt^2} &= -a_1\frac{d\hat{x}(t)}{dt} - a_{sdp}x(t) + b_0\frac{du(t)}{dt}u(t) + b_1u(t) \\ y(k) &= \hat{x}(k) + \xi(k) \end{aligned} \tag{15}$$

where a_{sdp} is the SDP, i.e.,

$$a_{sdp} = k_1 + k_2\hat{x}(t) + k_3\hat{x}^2(t) \tag{16}$$

The model parameters are optimized by minimizing the sum of squares of the model output error $\xi(k) = y(k) - \hat{x}(k)$ generated by the the Simulink model in Fig.5. This NLS optimization is initiated with the multi-input, quasi-linear parameter estimates obtained previously. The resulting 6 optimized parameter estimates are as follows, with the estimated standard errors, as defined by the square roots of the diagonal elements of the covariance matrix returned by confint, given in parentheses:

$$\begin{aligned} &\hat{a}_1 = 0.01749(2.02\times10^{-6});\ \hat{b}_0 = -0.1678(0.00018) \\ &\hat{b}_1 = 0.03110(2.98\times10^{-6}));\ \hat{k}_1 = 0.03136(8.73\times10^{-7}) \\ &\hat{k}_2 = -0.007782(1.19\times10^{-5});\ \hat{k}_3 = 0.1230(2.35\times10^{-5}) \end{aligned} \tag{17}$$

where it will be noted that the b_0 parameter estimate is very significantly negative, so that *the model has non-minimum phase characteristics*.

This 6 parameter model explains the data extremely well with $R_T^2 = 0.99998$ (residual RMS error of 0.00112 compared with the RMS for $y(k)$ of 0.1677, so a percentage error in these RMS terms of 0.67%). Also, when this model is simulated using the validation set of 10,000 samples, it produces very similar R_T^2 and error statistics to those obtained in this estimation stage. When the error series $\xi(k)$ is analysed, the *Akaike Information Criterion* (AIC) identifies an AR(100) model.

If we consider just the linear terms in this model, then we see that there is an *underlyng linear system* with nominal values of the steady state gain $\hat{b}_1/\hat{k}_1 = 0.992$, natural frequency $\omega_n = \sqrt{\hat{k}_1} = 0.1705$ radians/sampling interval and $2\zeta_n\omega_n = 0.01749$, so that the damping ratio is $\zeta_n = 0.0494$. Or, if the identified model is changed to the true sampling interval $T_s = 1/2441$ seconds, then the the steady state gain, damping ratio naturally remain the same but the natural frequency $\omega_n = 68.70$ Hz. Remember, however, that these are the underlying characteristics and they effectively change as the

output changes in a manner defined by the SDPs. This interpretation of the nonlinear model dynamics is examined further in sub-section 4.4.

One aspect of the model that can be consolidated at this stage is the model order identification. This is because the confint routine in Matlab utilises the results obtained from lsqnonlin, in particular the residual series and the Jacobian matrix J, to produce the confidence intervals and the associated covariance matrix. By providing information on the nature of the cost function-parameter hyper-surface near to the optimized parameter location, they provide a very good indication of whether the model is well parameterised or not. In the present case, for instance, the confidence intervals obtained in this manner are as follows (a verbatim copy from Matlab):

```
-0.16788       -0.16774
 0.031098       0.0311
 0.017492       0.017493
 0.031361       0.031362
-0.0077769     -0.0077864
 0.1230         0.1230
```

and we see that the confidence bounds very tightly encompass the estimated values in (17). If the model happened to be over-parameterised, however, then these bounds can become much wider (see later section 5) and even contain $+\infty$ and $-\infty$ elements.

The SDP nonlinearity $a_2\{y(t)\} \times y(t)$ for the model is plotted in Fig.6, where it is compared with other estimates: namely, the initial, unconstrained, non-parametric estimate; the estimate obtained with all of the parameters of the non-parametric model except $a_2\{y(t)\}$ constrained to be constant; and the initial parametric estimate. We see that, because there is so little noise on the Silverbox data, there is little difference in all these estimates, despite the fact that the non-parametric estimates are dependent on the measured output $y(t)$, while the parametric estimate is dependent on the internally generated, noise free output $x(t)$.

This similarity in the estimates of the nonlinearity demonstrates how well the nonlinearity in the Silverbox circuit is identified, a conclusion that is supported by the blue line on the plot. This blue line is the SDP nonlinearity estimated using the 5 parameter model and reported in Young (2016b). As we see, this is very similar to that of the present optimized 6 parameter model, shown as the dash-dot red line.

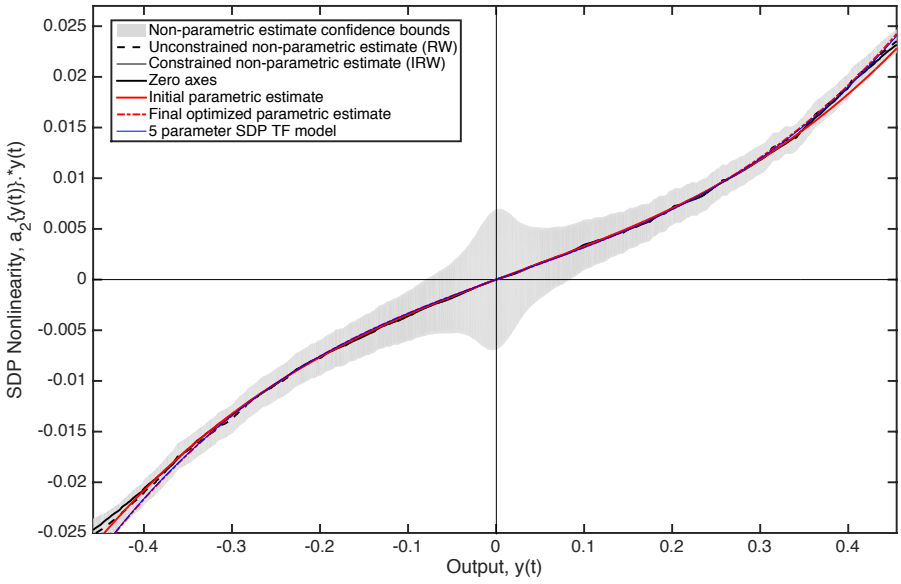

Figure 6: A comparison of the non-parametric and parametric estimates of the state dependent nonlinearity $a_2\{y(t)\} \times y(t)$ associated with the SDP $a_2\{y(t)\}$ in the 6 parameter SDPTF model.

These results suggest that we should compare the estimated parameters of the 6 parameter [2 2 0] model in (17) with those below of the earlier 5 parameter [2 1 0] model:

$$\begin{aligned} &\hat{a}_1 = 0.01704(2.15 \times 10^{-5}); \hat{b}_0 = 0 \\ &b_1 = 0.03086(3.32 \times 10^{-5})); \ \hat{k}_1 = 0.03127(9.64 \times 10^{-6}) \\ &\hat{k}_2 = -0.008093(0.00013); \ \hat{k}_3 = 0.1220(0.00026) \end{aligned} \tag{18}$$

As pointed out earlier the explanatory power of this slightly smaller model is $R_T^2 = 0.994$, which provides a very good explanation of the data but one that is not nearly as good as the exceptional value (for real data) of $R_T^2 = 0.99998$ for the above 6 parameter model (a residual RMS error of 0.0127 compared with 0.00112). However, except for the b_0 estimate (now effectively zero because it is not included in this 5 parameter model), these estimates are very similar to those of the 6 parameter model. Surprisingly, therefore, the superiority of the 6 parameter model does not arise because of any better characterisation of the nonlinearity in the system; it is due almost

entirely to the presence of the negative valued $\hat{b}_0 = -0.1678$ parameter *and the consequent non-minimum phase behaviour of the model.*

Finally, as mentioned previously, it is possible to simultaneously estimate the initial conditions I_1 and I_2 on the integrators in the SDPTF model, so avoiding the need to remove the initial 2500 samples that are affected by the considerable initial condition effects. The resulting parameter estimates, now based on the full 10,000 samples are given below:

$$\begin{aligned}
&\hat{a}_1 = 0.01751(1.62 \times 10^{-6}); \hat{b}_0 = -0.1684(0.00015) \\
&b_1 = 0.03111(2.49 \times 10^{-6})); \ \hat{k}_1 = 0.03136(7.72 \times 10^{-7}) \\
&\hat{k}_2 = -0.007884(9.17 \times 10^{-6}); \ \hat{k}_3 = 0.1231(1.96 \times 10^{-5}) \\
&\hat{I}_1 = -0.06827(2.91 \times 10^{-5}); \hat{I}_2 = 0.03045(0.00037)
\end{aligned} \tag{19}$$

where we see, on comparison with the estimation results in (17) that the standard errors here are smaller because of the larger sample size.

4.3.2 Optimization by Prediction Error Minimization

Discrete-time analysis of the residual noise series $\xi(k)$ from the NLS optimization in the previous sub-section 4.3.1 shows that, although it has an extremely small variance, it is highly correlated. However, the AIC suggests an AR(100) noise model, i.e. $\xi(k)$ in (15) can be modelled as:

$$\xi(k) = \frac{1}{1 + c_1 z^{-1} + c_2 z^{-2} \ldots + c_{100} z^{-100}} e(k); \qquad e(k) = N(0, \sigma^2) \tag{20}$$

This model is able to characterize $\xi(k)$ very well and yield serially uncorrelated, white residuals (one-step-ahead prediction errors), with a reasonably normal distribution and very small RMS value of 0.000052. Using this identified noise model, it is straightforward to optimize the model by PEM and, because of the low noise level, the resulting estimated parameters are very similar to those of the NLS optimized model but with slightly larger standard errors. It is not surprising, therefore, that it also validates well, producing very similar R_T^2 and error statistics.

The only disadvantage of this finally estimated model is that there is some significant cross correlation between the final prediction error series $e(k)$ and the input signal $u(k)$ over lags of up to 50 samples, with a maximum correlation of 0.2870 at zero lag, in relation to a standard error of 0.0115. This suggests that the optimized model has accounted for most, but not

all, of the nonlinearity in the data. Nevertheless, it is reasonable to utilise the estimated statistical properties of the parameter estimates to conduct uncertainty analysis using *Monte Carlo Simulation* (MCS). Such analysis allows us to investigate whether the estimated uncertainty in the parameters accounts for the errors in the model output; and it defines the uncertainty associated with the nonlinearity in the Silverbox circuit.

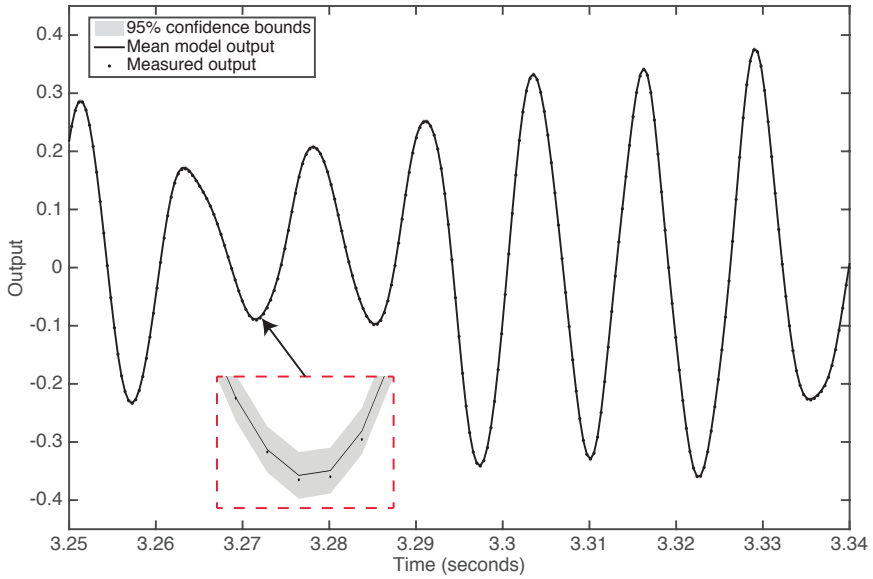

Figure 7: MCS analysis: A section of the measured output data $y(k)$ compared with the ensemble of model outputs $x(t)$. The zoomed inset shows that these very small confidence bounds encompass $y(k)$.

The NLS and PEM optimized models are not significantly different judged by the diagnostic statistics. From a theoretical standpoint, however, we should trust the slightly larger estimated standard errors on the PEM parameter estimates and use the associated error covariance matrix in the MCS analysis. Fig.7, on the next page, shows one of the results that can be obtained in this manner. It is based on MCS analysis using 150 random realizations and compares the model and measured outputs over a short segment of the data. It is difficult to discern the 95% confidence bounds on this plot because they are so small but the zoomed inset, which is typical of the situation throughout the series, shows that that the errors are accommodated well within the confidence bounds, as expected if the covariance matrix returned

by confint is providing reliable statistics.

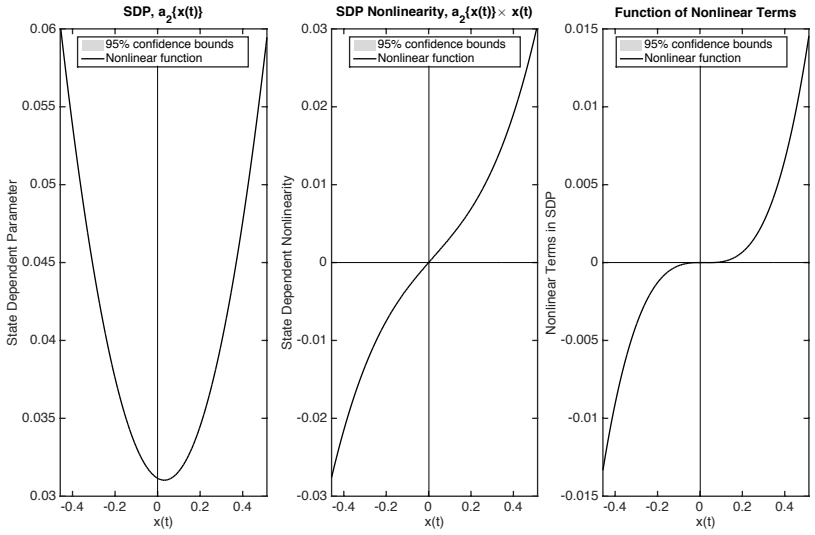

Figure 8: MCS analysis: the finally estimated nonlinear functions with their very small uncertainty bounds.

Fig.8, on the next page, shows the estimates of the three nonlinear functions that provide the best visualisation of the nonlinearity within an SDP context. We see that it compares very well with the earlier Fig.4, so again demonstrating the efficacy of the initial non-parametric estimation stage in this example. In Fig.8, however, the functions are defined parametrically. The left hand panel shows the estimated quadratic SDP function itself, $a_2\{x(t)\} = k_1 + k_2x(t) + k_3x^2(t)$, while the centre and right hand panels portray the nature of the total nonlinearity. The centre panel is the full SDP based nonlinearity $a_2\{x(t)\} \times x(t)$ obtained by multiplying the SDP by $x(t)$; while the right hand panel is the function defined by just the nonlinear terms in the SDP, i.e. $k_2x^2(t) + k_3x^3(t)$.

As pointed out previously, the latter function appears to flatten around the origin but, from the enlargement plotted in Fig.9, it is clear that it dips downwards before recovering and moving upward again. It is interesting to note that this function has almost exactly the same shape as the 'restoring force' referred to by Nöel and Schoukens (2018) in relation to their 13 parameter SS model and shown in their Fig.8. (the scaling of the plots is different, however, because of the different model structures and sampling interval).

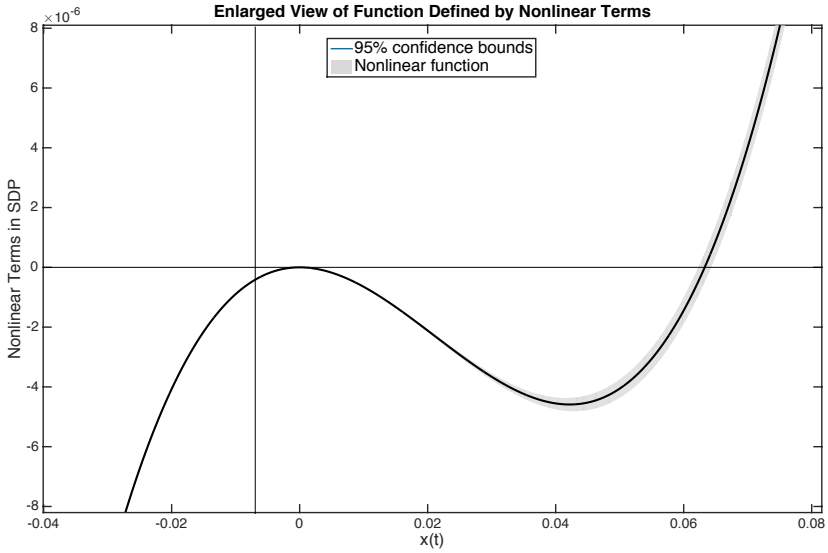

Figure 9: MCS analysis: An enlarged plot of the right hand panel in Fig.8 showing the behaviour near the origin and the presence of the very small confidence bounds.

Fig.9 also reveals the very small confidence bounds, demonstrating the accuracy of the estimated function and that it is a little less accurate once $x(t)$ becomes positive. This will be referred to again in the next sub-section 4.4.

4.4 Any Evidence of More Complex Nonlinearity?

The complexity of the small residual error in the DBM model and its correlation with the input variable probably suggest that there is an additional, very small amount of nonlinearity present that has not been captured by the SDP nonlinear model. This is supported to some degree by the initial non-parametric SDP analysis which identified some evidence of similar nonlinear effects in the first TF denominator parameter a_1. In this situation, it makes sense to investigate whether the introduction of state-dependency into a_1 results in any significant improvement in the explanatory ability of the model (providing, of course, that there is no sign of over-parameterisation that might question any such improvement). The resulting 8 parameter model takes the following form, which can be compared directly with the 6

parameter model (15):

$$\begin{aligned} x(t) &= \frac{b_0 s + b_1}{s^2 + a_1\{x(t)\}s + a_2\{x(t)\}} u(t) + \xi(k) \\ a_2\{x(t)\} &= k_1 + k_2 x(t) + k_3 x^2(t);\ a_1\{x(t)\} = k_4 + k_5 x(t) + k_6 x^2(t); \\ y(k) &= x(k) + \xi(k) \end{aligned} \tag{21}$$

No problems are encountered with optimizing the parameters in this model and the resulting estimates, with their standard errors in parentheses listed below:

$$\begin{aligned} &\hat{b}_0 = -0.1684(0.00017);\ b_1 = 0.03109(2.72 \times 10^{-6})) \\ &\hat{k}_1 = 0.03135(1.04 \times 10^{-6});\ \hat{k}_2 = -0.007956(1.17 \times 10^{-5}) \\ &\hat{k}_3 = 0.1234(2.66 \times 10^{-5});\ \hat{k}_4 = 0.01741(3.41 \times 10^{-6}) \\ &\hat{k}_5 = 0.01083(0.00030);\ \hat{k}_6 = 0.02626(0.00091) \end{aligned} \tag{22}$$

All of the estimates are well defined, with no signs of over-parameterisation and the explanation of the data shows some minor improvement, with a similar $R_T^2 = 0.99998$ to 5 significant figures but a reduced RMS error of 0.00102, compared with 0.00112 for the 6 parameter model. The AIC identified model for the residual error $\xi(k)$ is AR(87), which has white residuals, a reasonably normal distribution and an RMS value of 0.000053 which is slightly larger than for the AR(100) model of the 6 parameter model. However, this residual series is still correlated with the input variable $u(k)$, having a maximum correlation of 0.2052 at a lag of 3 samples, in relation to a standard error of 0.0115. So, disappointingly, we cannot claim that the enlargement of the model has purged much additional nonlinearity from the residual series.

Fig.10 shows the two estimated SDPs: it is clear, on comparison with the left hand panel of Fig.8, that the addition of the second SDP $a_1\{x(t)\}$ has caused no significant change to the estimated $a_2\{x(t)\}$ parameter in the left hand panel. Fig.11, on the next page, compares the estimates of the nonlinear function $k_2x^2(t) + k_3x^3(t)$ for the 8 and 6 parameter SDPTF models. There is only a very small difference apparent in the left hand panel but, on enlargement of the region around the origin, we see that the functions are a little different to the right of the origin and this is clearly related to the increased uncertainty shown by the similar plot in Fig.9, confirming again the efficacy of the uncertainty analysis for the smaller 6 parameter model and enhancing confidence in the presence of the additional SDP nonlinearity in the 8 parameter model.

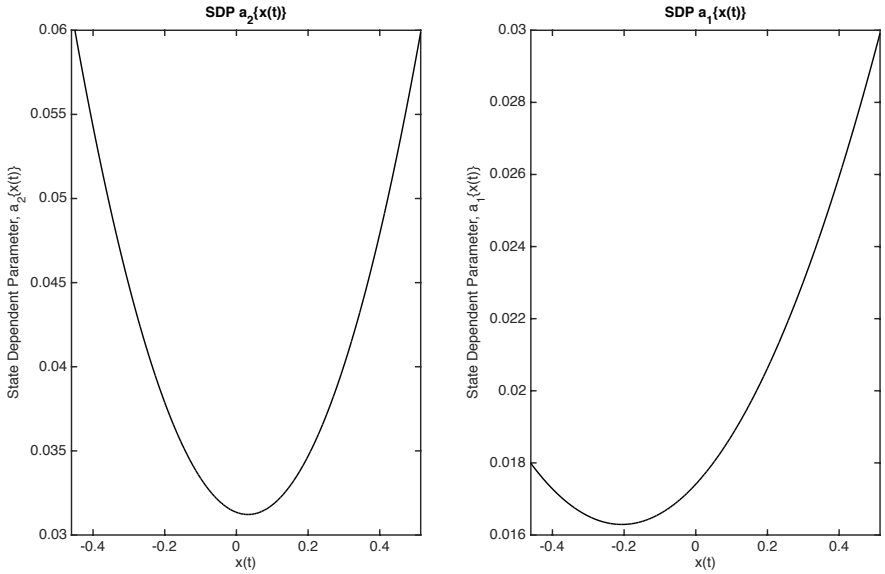

Figure 10: Parametric estimates of the two SDPs in the 8 parameter model (21).

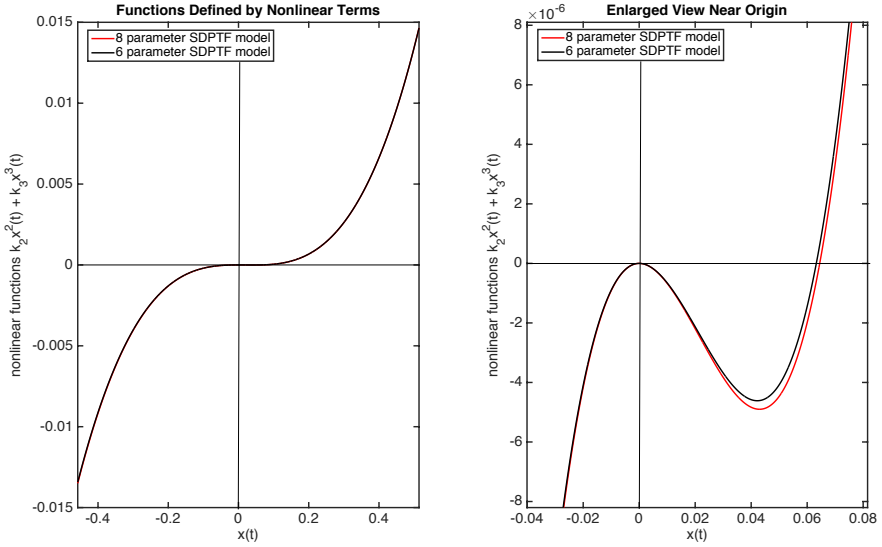

Figure 11: The left panel is a comparison of the nonlinear functions $k_2x^2(t) + k_3x^3(t)$ for the 8 and 6 parameter SDPTF models, with an enlarged view around the origin in the right hand panel.

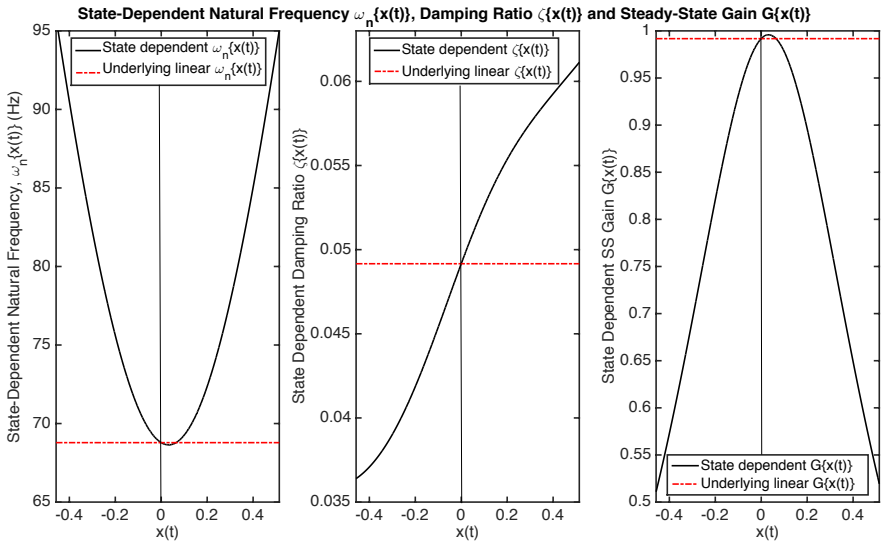

Figure 12: State dependent natural frequency (left panel), damping ratio (centre panel) and steady state gain (right panel) of the 8 parameter SDPTF model.

One interesting interpretative advantage of the continuous-time SDPTF model is that the state-dependency associated with any specified physical properties of the system can be investigated very simply. For example, the natural frequency ω_n, damping ratio ζ and steady state gain G can be computed as for a linear system but defined by the SDPs and so dependent on the output $x(t)$. This is illustrated in Fig.12, on the next page, which shows plots of the resulting state dependent characteristics $\omega_n\{x(t)\}$, $\zeta\{x(t)\}$ and $G\{x(t)\}$, based on the actual sampling interval of 1/2441 secs. Also shown by the red dash-dot lines are the underlying values of these properties defined by the linear part of the model. Extracting this kind of information would not be nearly so easy in the case of a discrete-time state space model, such as that described by equation (24) in the next section 5. And it would be even more difficult in the case of most full black box models.

It is clear that the SDPTF model (21) has the best explanatory power of all the models investigated in this note so far, including the SS models considered in the next section 5, although the difference is very small in all cases. As its parameters are also well defined with very low standard errors, therefore, it is the most acceptable model and superior to the favoured 6 parameter SDPTF model (15). However, this improvement is obtained

with the addition of two more parameters and so whether the quite small reduction in the modelling error is worth the addition of these parameters will depend upon the objectives of the modelling. It would probably be possible to continue to look for other SDP related nonlinearity in the model but I feel that we have reached a stage of diminishing returns and that the more parsimonious 6 parameter model might well be suitable for most modelling objectives.

5 A State Space DBM Model

Nöel and Schoukens (2018) do not specify the details of the 13 parameter discrete time state space model considered in their paper but, on request, Dr. Nöel kindly supplied these. When simulated in Simulink, this model has an $R_T^2 = 0.9988$ and residual RMS error of 0.0056. However, after optimization using lsqnonlin, based on a Simulink state space model of this same form, the resulting model explains the data much better with $R_T^2 = 0.99998$ and RMS error of 0.00107. The latter value is different from that cited in the Nöel and Schoukens (2018) paper but this probably arises from the different methods of model implementation and optimization. More importantly, in the present context, the explanatory ability is very similar to the 6 parameter SDP Model.

The parameters of this re-optimized model

$$\begin{aligned} x(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{B}u(k) + \mathbf{E}\begin{bmatrix} x^2(k) \\ x^3(k) \end{bmatrix} \\ y(k) &= \mathbf{C}\mathbf{x}(k) + \mathbf{D}u(k) + \xi(k) \end{aligned} \tag{23}$$

are as follows[5]:

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} 0.9712 & 0.1828 \\ -0.1665 & 0.9806 \end{bmatrix};\ \mathbf{B} = \begin{bmatrix} -0.04952 \\ -0.1388 \end{bmatrix} \\ \mathbf{E} &= \begin{bmatrix} -0.01228 & 0.1955 \\ -0.03452 & 0.5477 \end{bmatrix};\ \mathbf{C} = \begin{bmatrix} -1.1374 & 0.1833 \end{bmatrix}; \\ \mathbf{D} &= .002649 \end{aligned} \tag{24}$$

These parameter estimates are very close to those provided by Dr. Nöel but the confidence bounds returned by the confint routine, based on the re-

[5]Note that the parameter estimates shown here are rounded to four significant digits.

sults from the lsqnonlin optimization, are extremely wide, as shown verbatim below:

```
-1.7883       3.7306
-1.8221       2.1877
-1.8704       1.5374
-1.7789       3.7401
-7.3829       7.2838
-10.903       10.626
-61.424       59.149
-82.204       82.571
-1.8381       1.8135
-28.754       29.145
-2.7113       2.6422
-41.945       43.041
 0.00261     0.00269
```

suggesting that the 13 parameters are poorly defined statistically and that the model may be over-parameterised. This seems likely because the 6 parameter model has similar explanatory ability with 7 fewer parameters. It is hardly surprising, in any case, because it is well known that linear and nonlinear state space models (see e.g. Ljung, 1987; Wigren, 2006) are inherently over-parameterised unless specified in some canonical form (particularly if, as in Nöel and Schoukens (2018), the optimization is initiated using a completely black box, nonlinear subspace method of model identification that does not yield a model in such a canonical form).

In order to both investigate the possibility of over-parameterisation and compare the results from this type of model with those obtained previously with the SDPTF model, it makes sense to develop and optimize a similar nonlinear state space model in continuous-time form. Using the 5 parameter SDP model parameters as a starting point for identification, it proved fairly straightforward to identify such a model and optimize its parameters, resulting in the following model:

$$\begin{aligned} \frac{d\mathbf{x}(t)}{dt} &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t) + \mathbf{E}\begin{bmatrix} x^2(t) \\ x^3(t) \end{bmatrix} \\ y(k) &= \mathbf{C}\mathbf{x}(k) + \mathbf{D}u(k) + \xi(k) \end{aligned} \tag{25}$$

where the state equations are now in continuous-time differential equation form. The optimized parameters, based on a normalized sampling interval

of unity, as in the case of the previous SDPTF models, are as follows:

$$\begin{aligned}
\mathbf{A} &= \begin{bmatrix} -0.01734 & -0.03143 \\ 0.99692 & -0.0001 \end{bmatrix};\ \mathbf{B} = \begin{bmatrix} 0.9961 \\ -0.2460 \end{bmatrix} \\
\mathbf{E} &= \begin{bmatrix} 0.24528 & -3.9015 \\ 0.05504 & -0.9670 \end{bmatrix};\ \mathbf{C} = \begin{bmatrix} -0.0077 & 0.03152 \end{bmatrix}; \\
\mathbf{D} &= 0.0019
\end{aligned} \tag{26}$$

This model explains the data about the same as the discrete-time model, with $R_T^2 = 0.99998$ and RMS error of 0.00107; and, as expected, the confidence bounds provided by the confint routine are again extremely wide, suggesting that the 13 parameters are not well defined statistically and that the model is over-parameterised.

In order to solve this over-parameterisation problem, the most obvious approach is to consider a canonical form for the SS model where, in the linear part of the state equations, the second state is defined as the derivative of the first state. This not only reduces the number of parameters, it also provides a model where these parameters and the associated state variables have an immediate, more physical interpretation, as required by DBM modelling. The optimized parameters in this case are as follows:

$$\begin{aligned}
\mathbf{A} &= \begin{bmatrix} -0.01744 & -0.03134 \\ 1.0 & 0.0 \end{bmatrix};\ \mathbf{B} = \begin{bmatrix} 1.0 \\ 0.03086 \end{bmatrix} \\
\mathbf{E} &= \begin{bmatrix} 0.2458 & -3.8966 \\ 0.1236 & -2.0637 \end{bmatrix};\ \mathbf{C} = \begin{bmatrix} -0.01631 & 0.03106 \end{bmatrix};
\end{aligned} \tag{27}$$

with D=0, i.e. no direct effect of the input on the output (note that the discrete-time model and the unconstrained continuous-time model (26) have a very small values for this parameter).

The explanation of the data is similar to the unconstrained model case, with $R_T^2 = 0.99998$ and RMS error of 0.00111. Note here that the $b_2 = 0.03086$ is not optimized but set at the value of b_0 in the 5 parameter SDP model, so the model has only 8 optimized parameters. This is possible because b_2 only controls the gain of the system, so that it can be set arbitrarily to any value and the other parameters will be optimized to ensure that the overall gain of the model is maintained to provide a good explanation of the data[6]. However, setting it to value of b_0 provides a nice link between this

[6]this is the case when using SDPTF models of the Hammerstein type, with an input nonlinearity

model and its SDPTF predecessor (15), where the $\hat{a}_1 = 0.01705$, $\hat{b}_0 = 0.03096$ and $\hat{k}_1 = 0.03128$ parameters having similar values to the $\hat{a}_{11} = -0.01744$, $\hat{b}_2 = 0.03086$ and $\hat{a}_{12} = -0.03134$ parameters in the above model.

The 8 optimized parameters in (27), with their associated standard errors in parentheses (as derived again from the covariance matrix provided by the confint routine) are as follows:

$$\begin{aligned} \hat{a}_{11} &= -0.01744(0.000004);\ \hat{a}_{12} = -0.03134(0.000001) \\ \hat{c}_1 &= -0.01631(0.00002);\ \hat{c}_2 = -0.01631(0.000003) \\ \hat{e}_{11} &= 0.2458(0.0004);\ \hat{e}_{12} = -3.8966(0.0011) \\ \hat{e}_{11} &= 0.1236(0.0016);\ \hat{e}_{12} = -2.0637(0.0030) \end{aligned} \tag{28}$$

We see that this canonical approach to the model structure identification has not only considerably reduced the number of parameters from 13 to 8 but also retained the good explanatory ability of the model and provided an estimate of the parametric error-covariance matrix, as required for the uncertainty analysis.

As in the case of the 6 parameter SDP model, the residual error can be modelled as an autoregressive process which, in this case, is identified as an AR(96) model. The final residual $e(k)$ of this model is serially uncorrelated and very small, with a RMS value of 0.000053. However, as expected, there is still significant correlation with the input variable $u(k)$, with a correlation of 0.2828 at a lag of 5 samples, in relation to a standard error of 0.0116.

Now that the parameter estimates are well defined and that the uncertainty in the parameters is available from the covariance matrix, it is possible to conduct Monte Carlo based uncertainty analysis, as in the case of the 6 parameter SDPTF model. This yields quite similar results, although the standard error bounds are a little larger, probably because of the additional 2 parameters in the model.

Finally, note that the input-output transfer transfer function form of the estimated state equations involves the creation of the two states, each with its nonlinear inputs, and the combination of these two states to yield the estimate of the noise-free output variable. This severely complicates the

transfer function representation, which takes the following form:

$$\begin{aligned} x_1(t) &= \frac{B_{11}(s)}{A(s)}u(t) + \frac{B_{12}(s)}{A(s)}x^2(t) + \frac{B_{13}(s)}{A(s)}x^3(t) \\ x_2(t) &= \frac{B_{21}(s)}{A(s)}u(t) + \frac{B_{22}(s)}{A(s)}x^2(t) + \frac{B_{23}(s)}{A(s)}x^3(t) \\ x(t) &= c_1x_1(t) + c_2x_2(t);\ y(k) = x(k) + \xi(k) \end{aligned} \tag{29}$$

Consequently, the two state variables $x_1(t)$ and $x_2(t)$ are both functions of three component transfer functions which have a common denominator $A(s)$ and different, first order, numerators that shape the dynamic effects of the input $u(t)$ and the internally generated nonlinear variables $x^2(t)$ and $x^3(t)$ on the state variables. These are then combined to form the modelled output $x(t)$ and its measured value $y(k)$ of $x(k)$ at sample k. This relative complexity contrasts with the simplicity of the directly estimated SDPTF models in equations (15) and (21).

6 Black-box, Grey-box or Data-based Mechanistic Models?

There is no general answer to the question in the title of this section because this depends on the objectives of the analysis and the nature of the system. Indeed, it could well be that a combination of such methods would provide the best solution in some circumstances. However, I feel that the results of the analysis described in this technical note raise certain application-specific questions about the nature of nonlinear modelling and, in particular, how one should choose a model form in any particular circumstances.

In the case of the Silverbox, Nöel and Schoukens (2018) prescribe the equation of motion on the basis that the circuit is intended to mimic the behaviour of a mechanical system and suggest the following equation

$$M\frac{d2y(t)}{dt^2} + C_v\frac{dy(t)}{dt} + Ky(t) + c_1y^2(t) + c_2y^3(t) = p(t) \tag{30}$$

that reflects this mechanical engineering heritage. It seems most obvious, therefore, that the modelling should proceed on the basis of this equation in

the alternative identifiable form:

$$\begin{aligned} \frac{d2x(t)}{dt^2} + a_1\frac{dx(t)}{dt} + k_1x(t) + k_2x^2(t) + k_3x^3(t) = b_0u(t); \\ y(k) = x(k) + \xi(k) \end{aligned} \tag{31}$$

where $a_1 = C_v/M$, $k_1 = K/M$, $k_2 = c_1/M$, $k_3 = c_2/M$, $b_0 = 1/M$, $u(t) = p(t)$, and $y(k)$ is the sampled measurement of $x(t)$, with $\xi(k)$ representing the model error and any noise on the measurement. Certainly the DBM modelling strategy emphasises the use of a mechanistic model form that, if at all possible, coincides with the kind of general model forms that have been used previously in a similar context: in this case, the differential equation (31). Indeed, this is why the DBM modelling described in the present note is carried out in a continuous-time SDPTF form, because it is the suitably modified SDPTF equivalent of equation (31). We can see, therefore, that the philosophical basis of the 'grey-box' modelling approach used by Nöel and Schoukens (2018) is not the same as that of DBM modelling.

Other modellers of the Silverbox, *using another data set sampled at a different, larger sampling interval* (see e.g. Paduart et al., 2010; Wigren and Schoukens, 2013; Marconato et al., 2012; Nöel and Schoukens, 2018), have used models with parameters numbering from 10 to > 700, some completely black-box (see the section 'Related methods' in Marconato et al., 2012). Many of these are likely to be over-parameterised, some of them excessively so. It would seem, therefore, that the objective in most of these papers was simply to explain the data as well as possible, with not very much regard to the question of efficient parameterisation or mechanistic interpretation. One should ask, I believe, whether this is a reasonable objective when considering the identification of a system that has been constructed as an electronic circuit whose main purpose is to mimic a mechanical system. It is surely more intellectually satisfying to try to identify the system in a form that better matches the mechanistic nature of the system.

7 Conclusions

This technical note describes how the DBM modelling strategy, exploiting nonlinear *State Dependent Parameter Transfer Function* (SDPTF) models, can be applied to the Silverbox data and produce both 6 and 8 parameter models that explain the data very well indeed and much better than the 5

parameter SDPTF model identified in Young (2016b). This improvement in explanatory ability is due to the identification of an additional negative term in the numerator of their TFs that introduces dynamically significant non-minimum phase characteristics into these models. This and the other inferred physical characteristics of both models appear to make sense in relation to the nature of the Silverbox electronic circuit used to generate the data, although it would be advantageous to have more information on the fabrication and nature of this circuit in order to investigate this further.

The 8 parameter model is notable for two reasons: first, that it explains the Silverbox data better than any of the other SDPTF and SS models considered in this note, albeit by a very small amount (The RMS of the model error is 0.00102 compared with a range between 0.00107 and 0.00112 for the rest); second, it includes additional nonlinearity that, in SDP terms, is connected with the first a_1 parameter in the SDPTF. Whether these factors justify the addition of the two parameters, in relation to the smaller 6 parameter SDPTF model, will depend upon the objectives of the modelling.

The equivalent stochastic models estimated by prediction error minimization produce similar results to those obtained by the deterministic NLS optimization but allow better for uncertainty analysis, so confirming that each model's sampled stochastic output (i.e. the deterministic output $\hat{x}(k)$ plus the 95% uncertainty bounds on this response) encompass the measured output $y(k)$, even though these confidence bounds are very small indeed. However, the identification of all the SDPTF models, with their non-minimum phase characteristics, negates the control studies reported in Young (2016b) because SDP approach to control system design is much more difficult if the TF has non-minimum phase characteristics, with its 'unstable' numerator dynamics.

Although the circuit was designed to simulate the behaviour of a mass-spring-damper system, no attempt has been made to interpret it in this manner since the DBM modelling approach is only concerned with modelling the real system and interpreting the final DBM model in these terms. Of course, because the Silverbox circuit was developed in part to simulate a mechanical system, the data can be used to evaluate methodology developed specially for such systems. For instance, a companion technical note (Janot et al., 2016) uses them in this manner, based on the methodology developed in Janot et al. (2017). This methodology has some things in common with the approach used in the present note: in particular, CAPTAIN Toolbox routines irwsm, sdp are used for initial SDP modelling and one of the optimization

methods, IDIM-SDP, exploits these. Also, the rivcbj routine is used in a similar manner to that used in the present note.

The main disadvantage of the two DBM models identified in the present note is that their model error, although very small indeed, has correlation characteristics that are difficult to explain without further access to the Silverbox system (a disadvantage that, it would appear, affects all of the models so far identified for the Silverbox system). In both cases, this error $\xi(k)$ is both serially correlated and correlated with the sampled input signal $u(k)$. In each case, however, $\xi(k)$ can be modelled adequately by a high order AR stochastic process whose extremely small residuals are serially uncorrelated and normally distributed, although still correlated with $u(k)$. Serially correlated model error, requiring such a high AR model is unusual but does appear to occur in electromechanical systems (see e.g. Brunot et al., 2018). It almost certainly means that the error series arises from a very small amount of un-modelled nonlinearity. This conforms with the conclusions of others who have analysed the Silverbox data, in particular Nöel and Schoukens (2018).

Even if there is a small amount of such unexplained nonlinearity, it seems unlikely that it would be important for most practical applications of the models. I feel that it does not justify further investigation unless the model error signal can be shown to have components with an important physical interpretation that needs to be explained by the model. To paraphrase Karl Popper (Popper, 1959), I would contend that these DBM SDPTF models are conditionally validated and can be considered satisfactory for the most modelling objectives until they may be falsified by further experiment and analysis.

8 Acknowledgements

I am very grateful to Dr. J. P Nöel for his assistance in supplying both the Silverbox data set used in this note and the precise details of the model in Nöel and Schoukens (2018), as well as feedback on a number of topics. I have also benefitted by discussion on this subject with Prof. Johan Schoukens. Finally, the research in this note has been influenced very much by my collaboration and discussions with Dr. Alex Janot and Dr. Mathieu Brunot, who have provided valuable advice on the nature and modelling of electro-mechanical systems, as well as the views of the engineering community involved in the modelling and control of such systems.

References

Beven, K. J., D. T. Leedal, P. J. Smith, and P. C. Young (2011). Identification and representation of state dependent non-linearities in flood forecasting using the DBM methodology. In L. Wang, H. Garnier, and A. J. Jakeman (Eds.), *System Identification, Environmetric Modelling and Control*, Berlin. Springer-Verlag.

Billings, S. A. (2013). *Nonlinear system identification : NARMAX methods in the time, frequency, and spatio-temporal domains*. John Wiley & Sons.

Brunot, M., A. Janot, P. C. Young, and F. Carrillo (2018). An improved instrumental variable method for industrial robot identification. *Control Engineering Practice 74*, 107–117.

Coca, D. and S. A. Billings (1999). A direct approach to identification of nonlinear differential models from discrete data. *Mechanical Systems and Signal Processing 13(5)*, 739–755.

Hu, J., K. Kumamaru, and K. Hirasawa (2001). A quasi-ARMAX approach to modelling nonlinear systems. *International Journal of Control 74*, 1754–1766.

Janot, A., P. C. Young, and M. Gautier (2017). Identification and control of electro-mechanical systems using state-dependent parameter estimation. *International Journal of Control 90*(4), 643–660.

Janot, A., P. C. Young, and J. Noël (2016). On the identification of the silverbox system in a electromechanical context. Confidential Technical Report RT 1/25302 DCSD, Systems Control and Flight Dynamics, ONERA, Toulouse, France.

Laurain, V., R. Toth, M. Gilson, and H. Garnier (2010). Refined instrumental variable methods for identification of LPV Box–Jenkins models. *Automatica 46*(6), 959–967.

Ljung, L. (1987). *System Identification. Theory for the User*. Englewood Cliffs, NJ.: Prentice Hall.

Marconato, A., J. Sjöberg, J. Suykens, and J. Schoukens (2012). Identification of the silverbox benchmark using nonlinear state-space models. In

Proceedings of the 16th IFAC Symposium on System Identification, Brussels, Belgium.

Mendel, J. M. (1969). A priori and a posteriori identification of time varying parameters. In *Prodeedings, 2nd Hawaii Conference on System Sciences.*

Nöel, J. P. and J. Schoukens (2018). Grey-box state-space identification of nonlinear mechanical vibrations. *International Journal of Control 91*, 1–22.

Paduart, J., L. Lauwers, J. Swevers, K. Smolders, J. Schoukens, and R. R. Pintelon (2010). Identification of nonlinear systems using polynomial nonlinear state space models. *Automatica 46*, 647–656.

Pintelon, R. and J. Schoukens (2001). *System Identification: A Frequency Domain Approach*. NJ, USA: Piscataway, IEEE Press.

Popper, K. (1959). *The Logic of Scientific Discovery*. London: Hutchinson.

Priestley, M. B. (1980). State-dependent models: A general approach to non-linear time series analysis. *Journal of Time Series Analysis 1*, 47–71.

Priestley, M. B. (1981). *Spectral Analysis and Time Series*. London: Academic Press.

Sjöberg, J., Q. Zhang, L. Ljung, A. Benveniste, B. Deylon, P. Y. glorennec, H. Hjalmarsson, and A. Juditsky (1995). Nonlinear black-box models in system identification : a unified overview. *Automatica 31*(12), 1691–1724.

Wigren, T. (2006). Recursive prediction error identification and scaling of non-linear state space models using a restricted black-box parameterization. *Automatica 42*, 159–168.

Wigren, T. and J. Schoukens (2013). Three free data sets for development and benchmarking in nonlinear system identification. In *Proceedings 2013 European Control Conference*, Zurich, Switzerland, pp. 2933–2938.

Young, P. C. (1968). Adaptive pitch autostabilization of an air-surface missile,. In *Proc. Southwestern IEEE Conf.,San Antonio, Texas*, New York. IEEE.

Young, P. C. (1981). A second generation adaptive autostabilization system for airborne vehicles. *Automatica 17*, 459–470.

Young, P. C. (2000). Stochastic, dynamic modelling and signal processing: time variable and state dependent parameter estimation. In W. J. Fitzgerald, A. Walden, R. Smith, and P. C. Young (Eds.), *Nonlinear and Nonstationary Signal Processing*, pp. 74–114. Cambridge University Press: Cambridge.

Young, P. C. (2001a). Comment on 'a quasi-ARMAX approach to modelling nonlinear systems' by J. Hu *et al. International Journal of Control 74*, 1767–1771.

Young, P. C. (2001b). The identification and estimation of nonlinear stochastic systems. In A. I. Mees (Ed.), *Nonlinear Dynamics and Statistics*, pp. 127–166. Birkhauser: Boston.

Young, P. C. (2006). Non-parametric model structure identification and parametric efficiency in nonlinear state dependent parameter models. In *IEEE International Symposium on Evolving Fuzzy Systems, Ambleside, UK*, pp. 349–354.

Young, P. C. (2011). *Recursive Estimation and Time-Series Analysis: An Introduction for the Student and Practitioner*. Springer-Verlag, Berlin.

Young, P. C. (2013). Hypothetico-inductive data-based mechanistic modeling of hydrological systems. *Water Resources Research 49*(2), 915–935.

Young, P. C. (2015). Refined instrumental variable estimation: Maximum likelihood optimization of a unified Box-Jenkins model. *Automatica 52*, 35–46.

Young, P. C. (2016a). Data-based mechanistic modeling. In V. P. Singh (Ed.), *Handbook of Applied Hydrology*, Chapter 33, pp. 1–12. USA: McGraw Hill.

Young, P. C. (2016b). Data-based mechanistic modelling and control of the silverbox nonlinear circuit. Technical Report TN/PCY/3 (2016), Lancaster Environment Centre, Lancaster University, UK.

Young, P. C. (2018). Data-based mechanistic modelling and forecasting globally averaged surface temperature. *International Journal of Forecasting 34*, 315–334.

Young, P. C., M. Foster, and M. J. Lees (1993). A direct approach to the identification and estimation of continuous-time systems from discrete-time data based on fixed interval smoothing. In G. C. Goodwin and R. J. Evans (Eds.), *12th IFAC World Congress*, Oxford, pp. 27–30. Pergamon Press.