Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Closed-Loop Identification of Hammerstein Systems with Application to Gas Turbines

Chad M. Holcomb, Raymond A. de Callafon and Robert R. Bitmead

Department of Mechanical and Aerospace Engineering
University of California San Diego
9500 Gilman Drive, La Jolla, CA, 92093-0411, USA

Abstract: Many practical applications, such as the fuel control of a gas turbine engine, can be modeled by a feedback connection of a linear controller in series with a Hammerstein system, where the nonlinearity provides a representation of the control element or actuator. An iterative gradient-based method is proposed to simultaneously identify the nonlinear fuel valve characteristic and a low-order linear plant model in gas turbine applications that leverages *a priori* knowledge of both the nonlinearity and engine dynamics. The identification is a nonlinear prediction error minimization method in a closed-loop Hammerstein model framework. It is applied to data from a high-fidelity simulation of a 5 megawatt $Taurus^{TM}$ 60 industrial gas turbine.

Keywords: closed-loop identification, Hammerstein systems, gas turbine

1. INTRODUCTION

The performance of the fuel control system in a gas turbine engine is critical to maintain stability and achieve performance targets. A digital feedback controller meters fuel into the combustion chamber using measurements of shaft speed, stage temperatures, pressures, and power. The fuel control valve(s) typically possess a nonlinear position to flow area relationship. The control system requires knowledge of this nonlinear characteristic to accurately regulate fuel flow. Uncertainty or degradation of the physical fuel valve's flow characteristic can lead to instability or operational limitations of turbine engine. Sources of uncertainty vary from manufacturing variability to contamination due to sulfur deposits on the valve's control surface [Cézac et al. (2008)]. Maximization of machine availability is essential to operators and the cost of unplanned service interruption is typically greater than the cost of preventative maintenance and returning the unit to service. The motivation of this paper is the identification of uncertainty in a nonlinear actuator characteristic in closed-loop operation.

Breikin et al. (2004) demonstrated that low-order linear plant models effectively capture the relationship from fuel flow to output power and Dai and Wang (2006) presented similar results for the relationship from fuel flow to shaft speed. Measurement and simulation data are only available from closed-loop operation. The nonlinear flow control valve, assumptions on linear behavior of the turbine engine over limited operating range, and closed-loop data fit nicely into a closed-loop Hammerstein model framework. The Hammerstein model structure comprises an input nonlinearity in series with a linear dynamic model. This model structure can be used to identify the uncertainty in the actuator nonlinearity and approximate the dynamics of the turbine engine via a linear plant model.

Identification of closed-loop Hammerstein systems has focused on instrumental variable (IV) based methods as they mitigate the bias due to the correlation of output noise and the input and output signals. Laurain et al. (2009) presented an iterative refined IV identification algorithm for LTI systems and later for LPV systems [Laurain et al. (2010)]. Han and De Callafon (2011) applied iterative IV identification to the problem using piecewise triangle basis functions to parametrize the nonlinear function. Laurain states that the IV methods provide a good initialization for use in statistically optimal prediction error methods that are sensitive to the initialization step. The IV methods offer consistent parameter estimates on average, although without optimality properties or convergence guarantees.

Prediction error minimization (PEM) methods offer an alternative. De Bruyne et al. (1999) developed generalized gradient expressions for prediction error minimization in linear closed-loop systems and also noted that exact gradient expressions can be developed for closed-loop nonlinear systems where the controller is smooth and the system is bounded-input, bounded-output (BIBO) stable around a stable trajectory. Van Pelt and Bernstein (2000) used piecewise linear static maps to parametrize the nonlinearities for system identification in open-loop and closed-loop Hammerstein frameworks. Narendra and Gallman (1966) applied an iterative gradient descent algorithm for the open-loop case that motivates the closed-loop formulation here.

Given that that the input nonlinearity (fuel valve) is smooth and partially known and linear dynamic models have been shown to accurately capture the dynamic response of gas turbines over a small operating range during closed-loop control, there is an opportunity to expand the application of systematic closed-loop identification Hammerstein systems. We seek to apply closed-loop Hammerstein system identification that exploits *a priori* knowledge of the input actuator (control valve) for a targeted identification of uncertainty in the flow characteristic of the fuel valve and to evaluate the method on a high-fidelity first principles simulation of a gas turbine generator control system. Results are presented from closed-loop data from a high-fidelity simulation of a $Taurus^{TM}$ 60 conventional combustion gas turbine generator.

2. PROBLEM FORMULATION

The objective of this paper is to identify a low order linear dynamic system with possible static nonlinearity in the form of a Hammerstein model to capture the dynamics of a high-fidelity nonlinear turbine model with a non-linear, but static fuel valve characteristic. The nonlinear thermodynamic Matlab/Simulink model of a gas *Taurus*^{TM} 60 turbine generator and feedback control system is used to generate data for use in identification. The high-fidelity model contains a full model of the Brayton cycle that includes each stage of the cycle in addition to the full fuel control, actuator and sensor models. The engine portion of the simulation used in this discussion, depicted in Figure 1, is comprised of the major engine sub-assembly models, i.e. compressor, burner, rotor, power turbine and exhaust, and their interconnections.

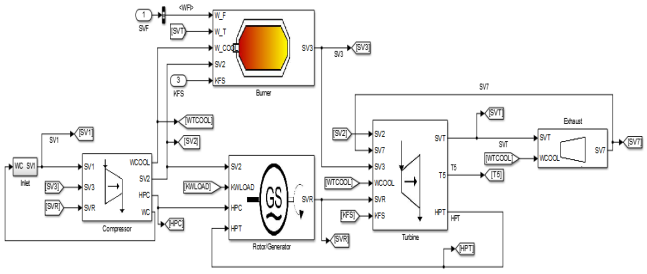

Fig. 1. High-fidelity engine simulation implemented in Matlab / Simulink.

3. MODEL DESCRIPTION

For the sake of the approximation problem formulated in this paper, the fuel control of a gas turbine is represented by the closed-loop BIBO stable nonlinear system model $\mathcal{M}$, shown in Figure 2. The known controller $K(q)$, and static nominal actuator mapping $f_0(\cdot)$, with output $w(t)$ is explicitly included in the model. The uncertain static memoryless nonlinearity $\delta(\cdot)$, in the series connection of $f_0(\cdot)$ with linear dynamics $G(q)$ *jointly* capture any deviation of the fuel valve characteristic from $f_0(\cdot)$ and nonlinear behavior of the gas turbine. The problem is to identify the unknown nonlinear map $\delta(\cdot)$, and linear dynamics $G(q)$. For identification, an additive and persistently exciting reference signal is applied to the shaft speed set point $r(t)$. The data set contains the uniformly sampled input-output signals of $r(t)$, $u(t)$, and $y(t)$ with sampling time T_s over N samples. System identification applies a two stage iterative gradient-descent procedure within a prediction error minimization framework to estimate $\delta(\cdot)$ and a low-order linear dynamic plant $G(q)$.

Since $\delta(\cdot)$ jointly captures uncertainty in $f_0(\cdot)$ and the nonlinearity of the turbine plant, we introduce $\delta_w(t)$ in the series connection of the static nonlinearity and linear plant dynamics to be identified. The noise $v(t)$, is assumed as inherent to the physical system and in this context, the noise model is not important to the identification objective. Since an output error (OE) model structure is used for identification, the noise is assigned a zero mean sequence $v(t) \sim N(0,\lambda)$ as a matter of convention and a noise model is not estimated, i.e. $H_0(q) = 1$. The signal $w(t)$, represents the flow area of the control valve is a known function of the controller output $u(t)$ that is given by,

$$w(t) = f_0(u(t)). \tag{1}$$

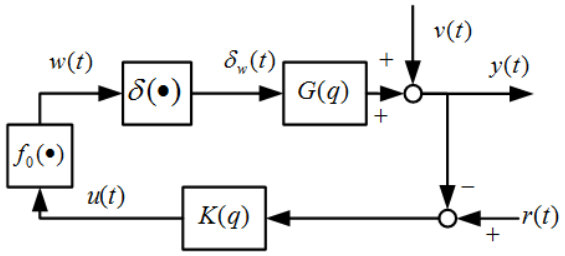

Fig. 2. Closed-loop Hammerstein system model $\mathcal{M}$.

To facilitate parameter estimation, $\delta(\cdot)$ is approximated by a set of orthogonal basis functions that allow $\delta_w(t)$ to be written

$$\delta_w(t) = \sum_{j=1}^{M} \rho_j(w(t))\mu_j. \tag{2}$$

The linear dynamic process $G(q)$ is the linear time invariant plant,

$$G(q) = q^{-t_d}\frac{B(q)}{A(q)}, \tag{3}$$

of polynomials $A(q)$ and $B(q)$, with the time shift operator q^{-1}, and input time delay t_d. Similarly, the controller may be written

$$K(q) = \frac{D(q)}{C(q)}. \tag{4}$$

The following assumptions apply throughout the discussion:

- A1: The reference input $r(t)$ is known, control output $u(t)$, and noisy output $y(t)$ are measured for identification.
- A2: The system is closed-loop BIBO stable.
- A3: The nominal $f_0(\cdot)$ in (1) is a known, monotonic, continuously differentiable function.
- A4: The input time delay t_d, to the linear plant is known.
- A5: The reference input $r(t)$, is persistently exciting for the identification of $G(q)$.

Note: It is not assumed that $u(t)$ excites the full input range of the static nonlinearity $\delta(\cdot)$ and plant. We will specifically use the series connection of $\delta(\cdot)$ and $G(q)$ to identify $\delta(\cdot)$ over a limited range.

4. PARAMETRIZATION

4.1 Static nonlinearity

The nonlinear mapping in (2) is written as a linear combination of orthogonal basis functions $\rho_j(w(t))$, with weights μ_j. The weighting vector μ, of function $\delta(\cdot)$, in this basis is an M-vector parameter to be identified. The basis uses the grid

$$m = [\, m_1 \ \cdots \ m_M \,]^T, \tag{5}$$

to define the center locations of the basis functions and satisfies $[m_1 \le w(t) \le m_M]$, $\forall\, t \in [1..N]$. In practice, the entire range of $u(t)$, and therefore $w(t)$, may not be able to be excited due to operational constraints on the physical system. The choice of the M-vector of basis functions $\rho(w(t))$, is closely related to the identification objective and structure of the nonlinearity. A basis that facilitates a good approximation with parsimony in the parameters is desirable. Han and De Callafon (2011), for example, apply a set of piecewise triangular basis functions to the problem. In this study, we seek specifically to locally identify $\delta(\cdot)$ in the series connection of a smooth valve characteristic $f_0(\cdot)$, and linear dynamics $G(q)$. Lippmann (1991) discusses

various basis function choices and states that Gaussian bases excel at characterizing local properties and based on this, a Gaussian basis is used here. The radial basis functions at grid node m_j, are defined as

$$\rho_j(w(t)) = \exp\left(\frac{-(w(t)-m_j)^2}{\sigma_j^2}\right)$$

where σ_j controls the spread of the basis functions. The choice of σ_j is a design parameter and is taken to be constant. The center values m_j, are taken on a uniformly spaced grid. The vector of basis functions $\rho(w(t))$, and parameter μ, are written

$$\rho(w(t)) = [\,\rho_1(w(t))\ \cdots\ \rho_M(w(t))\,]^T, \tag{6}$$

$$\mu = [\,\mu_1\ \cdots\ \mu_M\,]^T \tag{7}$$

The vector μ, specifies the weights of the basis functions at the grid points in (5) and $\delta(w(t))$ can be written in vector notation as

$$\delta_w(t) = \rho^T(w(t))\mu. \tag{8}$$

4.2 Closed-Loop Model Parametrization

The model $G(q)$ is parameterized by $G(q,\eta) = B(q,\eta)/A(q,\eta)$ where

$$A(q,\eta) = 1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a},$$
$$B(q,\eta) = b_0 + b_1 q^{-1} + \cdots + b_{n_b} q^{-n_b}.$$

The stacked plant parameter vector η, is constructed from

$$\eta_a = [\,a_1\ \cdots\ a_{n_a}\,]^T\ ,\ \eta_b = [\,b_0\ \cdots\ b_{n_b}\,]^T$$

as

$$\eta = [\,\eta_a\ \eta_b\,]^T \tag{9}$$

The predictor of $y(t)$, using an OE model structure of the linear plant, is given by

$$\hat{y}(t) = \frac{B(q,\eta)}{A(q,\eta)}\rho^T(\hat{w}(t-t_d))\mu. \tag{10}$$

Substituting the expression for $w(t)$ from (1) allows the set of closed-loop signal relations, from Figure 2 to be written:

$$\mathcal{M}(\eta,\mu)\begin{cases} \hat{y}(t) &= \dfrac{B(q,\eta)}{A(q,\eta)}\rho^T(\hat{w}(t-t_d))\mu \\ \hat{w}(t) &= f_0(\hat{u}(t)) \\ \hat{u}(t) &= K(q)\,(r(t)-\hat{y}(t)) \end{cases} \tag{11}$$

The prediction in (10) can be written as a linear combination of the parameter dependent *regressor* $\phi^T(t,\theta)$, and a non-minimal parameter θ, as

$$\hat{y}(t,\theta) = \phi^T(t,\theta)\theta. \tag{12}$$

The parameter θ, and noise free data *regressor* $\varphi(t,\theta)$, vectors are written

$$\theta = \begin{bmatrix} \theta_a \\ \theta_{b\mu} \end{bmatrix} \text{ and } \varphi(t,\theta) = \begin{bmatrix} \varphi_a(t,\theta) \\ \varphi_{b\mu}(t,\theta) \end{bmatrix}. \tag{13}$$

The elements of (13) are given by

$$\begin{aligned} \theta_a &= [\,a_1\ \cdots\ a_{n_a}\,]^T \\ \theta_{b\mu} &= \left[\,b_0\mu^T\ \cdots\ b_{n_b}\mu^T\,\right]^T, \\ \varphi_a(t,\theta) &= [\,-\hat{y}(t-1)\ \cdots\ -\hat{y}(t-n_a)\,]^T, \\ \varphi_{b\mu}(t,\theta) &= \left[\,\rho^T(\hat{w}(t-t_d)))\ \cdots\ \rho^T(\hat{w}(t-t_d-n_b)\,\right]^T. \end{aligned} \tag{14}$$

The *regressor* matrix built from N data samples is given by

$$\Phi^T(\theta) = \begin{bmatrix} \Phi_a^T(\theta) & \Phi_{b\mu}^T(\theta \end{bmatrix} \tag{15}$$

$$= \begin{bmatrix} \varphi_a^T(1,\theta) & \varphi_{b\mu}^T(1,\theta) \\ \vdots & \vdots \\ \varphi_a^T(N,\theta) & \varphi_{b\mu}^T(N,\theta) \end{bmatrix} \in \Re(N \times n_{\mathcal{M}}) \tag{16}$$

where $n_{\mathcal{M}} = n_a + (n_b+1)M$.

5. OPTIMIZATION

5.1 Prediction Error and Objective Function

The prediction error $\varepsilon(t,\theta)$, is given by

$$\varepsilon(t,\theta) = y(t) - \varphi^T(t,\theta)\theta. \tag{17}$$

In batch estimation with N data points, the measurement vector Y, prediction vector $\hat{Y}(\theta)$, and RMS prediction error vector $E(\theta) = \frac{1}{\sqrt{N}}[Y - \hat{Y}(\theta)]$, are given by:

$$\begin{aligned} Y &= [\,y(1)\ \cdots\ y(N)\,]^T \\ \hat{Y}(\theta) &= [\,\hat{y}(1,\theta)\ \cdots\ \hat{y}(N,\theta)\,]^T \\ E(\theta) &= \frac{1}{\sqrt{N}}\,[\,\varepsilon(1,\theta)\ \cdots\ \varepsilon(N,\theta)\,]^T \end{aligned} \tag{18}$$

We seek $\hat{\theta}$ to minimize the quadratic cost function

$$V_{OE}^N(\theta) = \frac{1}{2N}\sum_{t=1}^{N}\varepsilon^2(t,\theta) = \frac{1}{2}E(\theta)^T E(\theta). \tag{19}$$

The minimization of (19) is an OE minimization where $V_{OE}^N(\theta)$ is inherently non-convex in θ due to the dependency of $\varphi(t,\theta)$ in (17) on θ. Furthermore, $V_{OE}^N(\theta)$ is over-parameterized in θ due to the multiplicative elements in $\theta_{b\mu}$ from (14). Despite these difficulties, we develop an iterative gradient-based algorithm that is robust to over-parameterization and ensures convergence to at least a local minima. Details on the iterative gradient-based algorithm are presented in the next subsection.

5.2 Dealing With Overparametrization

The non-minimal structure of the parameter $\theta_{b\mu}$ is common to identification of Hammerstein systems with an arbitrary gain that must be distributed between the nonlinearity and linear plant. Bai (1998) presents an optimal rank one method for decomposing $\theta_{b\mu}$ with freedom to obtain uniqueness in the parameters by normalizing $\|b_i\|_2$ or $\|\mu_j\|$ to one and applying a SVD procedure. While this yields a unique estimate, it might not conform to physical reality. In our analysis, we seek to identify a local range of $\delta(\cdot)$ and dynamic plant model, and will achieve uniqueness of the parameter estimates by constraining a single function value of $\delta(\cdot)$. The excitation is additive and symmetric about a constant speed reference r_0. This corresponds to a steady-state operating point $[r_0, u_0, w_0, y_0]$. Define $\bar{u}(t) = u_0$ and with (1), define the grid (5) such that it contains a central grid point $m_i = \bar{w}(t) = f_0(\bar{u}(t))$ and set $\delta(\bar{w}(t)) = 0$. This constraint allows us to use a priori knowledge of $f_0(\cdot)$ within the model structure as it corrects for the nonlinearity as the system moves from the operating point. This allows the linear dynamic plant to capture the steady state gain as well as the dynamics of the physical system. In addition, this enables the comparison of $\delta(\cdot)$ from successive batches of data.

From the definition of the parameter vector in (13), we define an auxiliary parameter matrix

$$\Gamma \triangleq \begin{bmatrix} b_0\mu_1 & b_0\mu_2 & \cdots & b_0\mu_M \\ b_1\mu_1 & b_1\mu_2 & \cdots & b_1\mu_M \\ \vdots & \vdots & \ddots & \vdots \\ b_{n_b}\mu_1 & b_{n_b}\mu_2 & \cdots & b_{n_b}\mu_M \end{bmatrix}.$$

The optimal rank one estimate of the parameter vectors $\hat{\eta}_b$ and $\hat{\mu}$ from the parameter estimate $\hat{\theta}^T_{b\mu}$, are obtained by minimizing

$$[\hat{\mu}, \hat{\eta}_b] = \underset{\mu \in \Re^M,\ \eta_b \in \Re^{n_b+1}}{\arg\min} ||\Gamma - \hat{\Gamma}(\theta^T_{b\mu})||^2, \tag{20}$$

where $\hat{\Gamma}(\hat{\theta}^T_{b\mu})$ is assembled using the *blockvec* operator as

$$\begin{aligned}\hat{\Gamma}(\hat{\theta}^T_{b\mu}) =& \text{blockvec}(\hat{\theta}_{b\mu}) \\ =& \begin{bmatrix} \hat{\theta}^T_{b\mu}(1:M) \\ \hat{\theta}^T_{b\mu}(M+1:2M) \\ \vdots \\ \hat{\theta}^T_{b\mu}(n_bM+1:(nb+1)M) \end{bmatrix}.\end{aligned}$$

To assert the constraint $\delta(\bar{w}(t)) = 0$, first set an initial value of $\hat{b}_0 = 1$ and $\hat{\mu}_0 = \hat{\theta}^T_{b\mu}(1:M)$, which gives $\hat{\delta}_0(w(t)) = \rho^T(w(t))\hat{\mu}_0$ and $\hat{\delta}_0(\bar{w}) = \rho^T(\bar{w})\hat{\mu}_0$. Define Ω as the mapping of M points in $[\, w_1 \cdots w_M \,]$ within the range $[m_1, m_M]$ as

$$\Omega = \begin{bmatrix} \rho_1(w_1) & \cdots & \rho_M(w_1) \\ \vdots & \ddots & \vdots \\ \rho_1(w_M) & \cdots & \rho_M(w_M) \end{bmatrix} \in \Re^{(MxM)}. \tag{21}$$

A unique inverse of Ω is guaranteed to exist by its construction as an orthogonal set of basis functions. If the correction

$$\hat{\mu} = \hat{\mu}_0 - \Omega^{-1}\hat{\delta}_0(\bar{w}) \tag{22}$$

is applied to the weighting parameter vector, then $\delta(w(t)) = \rho^T(w(t))\hat{\mu}$ satisfies $\delta(\bar{w}(t)) = 0$ and $\hat{b}_i$ is updated using $\hat{\mu}$ as

$$\hat{b}_i = \frac{\hat{\mu}^T \hat{\Gamma}(i,:)}{\hat{\mu}^T\hat{\mu}}, \tag{23}$$

where the i^{th} row of the $\hat{\Gamma}$ matrix is $\hat{\Gamma}(i,:)$ in this notation. For an estimate $\hat{\theta}^N$, $\hat{\eta}_a$, $\hat{\eta}_b$, and $\hat{\mu}$ are created from (9), and (7) respectively.

5.3 Initialization

The cost function in Eqn. (19) will likely contain multiple local minima, which stresses the importance of a good initial parameter estimate. Define n^*_a and n^*_b as such that the linear plant model $G(q,\hat{\eta})$ capture the plant dynamics. Breikin et al. (2004) offers insight that first order linear models perform adequately around local operating points. For initialization, we increase the orders of n_a and n_b to minimize the bias from the nonlinear distortions and unmodeled dynamics on the dynamic plant in the initialization. Define n_e as the model order increase such that the orders of the initial estimate follow $n_{a,0} = n^*_a + n_e$ and $n_{b,0} = n^*_b + n_e$ while the parameter μ retains dimension M. A least squares initial estimate $\hat{\theta}^N_0$, is calculated generate $\hat{\eta}_0$and $\hat{\mu}_0$.

A subsequent model reduction via singular perturbation procedure applied to $G(q,\hat{\eta}_0)$ to generate a new linear plant estimate $\hat{\eta}_1$ of orders n^*_a and n^*_b [Liu and Anderson (1989)]. Subsequently, $\hat{\theta}^N_1$ is constructed with $\hat{\mu}_0$ and $\hat{\eta}_1$.

The model order reduction reduces the computational expense of the iterative gradient descent algorithm due to the significant reduction in parameters. The matrix inverse in the iterative update, detailed in the next section, dominates the computational cost and is of order $O(n^3_M)$. The model reduction enables the use a high order initial estimate, while reducing the cost of each subsequent iteration by $O((n_eM + n_e)^3)$.

5.4 Iterative Gradient Descent Estimation

The iterative gradient descent algorithm requires a parameter estimate θ_s, and gradient of the cost function (19) with respect to the parameter estimate, at each iteration step s. The gradient is commonly defined as $\Psi(\theta) = -\nabla_\theta V^N(\theta)$. The non-minimal parametrization contains multiplicative elements of η_b and μ. To relax the parametrization's constraint on the gradient direction we apply a two-step parameter update procedure. The procedure first takes an iteration step in the μ direction, with η held constant, that generates $\hat{\theta}_{\mu,s}$, and a step in the η direction, with μ held constant, that generates $\hat{\theta}_{s+1}$. The gradient descent estimation requires that the cost function gradients $F_\mu(\hat{\theta}_s)$ and $F_\eta(\hat{\theta}_s)$, to be calculated at each iteration step.

Gradient Expressions The gradient expressions for $F_\mu(\hat{\theta}_s)$ and $F_\eta(\hat{\theta}_s)$ are developed in the following two lemmas. The gradient is simply a function of $theta$, and the evaluation of the gradient at $\hat{\theta}_s$ is found by substitution of $\theta = \hat{\theta}_s$.

Lemma 1. The gradient of the cost function (19) with respect to μ, with η held constant, is given by

$$F_\mu(\theta) = -\nabla_\mu V^N(\theta)\big|_\eta = \Psi^T_\mu(\theta)E(\theta), \tag{24}$$

where $\nabla_\mu = \begin{bmatrix} \frac{\partial}{\partial\mu_1} & \cdots & \frac{\partial}{\partial\mu_M} \end{bmatrix}$ and $\Psi_\mu(\theta)$ is the gradient of the prediction error vector $\nabla_\mu E(\theta)$. For a time sample t,

$$\psi_\mu(t,\theta) = \begin{bmatrix} \frac{\partial\hat{y}(t)}{\partial\mu_1} & \cdots \frac{\partial\hat{y}(t)}{\partial\mu_k} & \cdots & \frac{\partial\hat{y}(t)}{\partial\mu_M} \end{bmatrix} \tag{25}$$

and $\Psi_\mu(\theta) \in \Re(N \times M)$ is written

$$\Psi_\mu(\theta) = \frac{1}{\sqrt{N}}\begin{bmatrix} \psi^T_\mu(1,\theta) & \cdots & \psi^T_\mu(N,\theta) \end{bmatrix}^T. \tag{26}$$

The prediction error gradient $\Psi_\mu(\theta)$, is calculated by filtering $\hat{w}(t)$ using the parameters μ, η, and closed-loop signal relations from (11).

Proof. With $[r(t), \hat{u}(t), \hat{w}(t), \hat{y}(t)]$ defined in (11),

$$\begin{aligned}\frac{\partial\hat{y}(t)}{\partial\mu_k}\bigg|_\eta =& \frac{1}{A(q,\eta)}\frac{\partial\hat{y}(t)}{\partial\mu_k} + \frac{B(q,\eta)}{A(q,\eta)}\rho^T(\hat{w}(t-t_d)) \\ &+ \frac{B(q,\eta)}{A(q,\eta)}\left(\frac{\partial\rho^T(\hat{w}(t-t_d)}{\partial\hat{w}(t-t_d))}\frac{\partial\hat{w}(t-t_d)}{\partial\mu_k}\right)\mu.\end{aligned} \tag{27}$$

With $\hat{w}(t)$ defined in (1),

$$\frac{\partial\hat{w}(t)}{\partial\mu_k} = -\left(\frac{\partial f_0(\hat{u}(t))}{\partial\hat{u}(t)}\right)\left(\frac{D(q)}{C(q)}\right)\left(\frac{\partial\hat{y}(t)}{\partial\mu_k}\right), \tag{28}$$

and

$$\frac{\partial\rho^T(\hat{w}(t))}{\partial\hat{w}(t)} = -\frac{2}{\sigma^2}\begin{bmatrix} (\hat{w}(t)-m_1)\rho_1(\hat{w}(t)) \\ \vdots \\ (\hat{w}(t)-m_M)\rho_M(\hat{w}(t)) \end{bmatrix}^T, \tag{29}$$

for $[m_1 < \hat{w}(t) < m_M]$. With $\hat{w}(t)$, (29) can be calculated from $t = 0, 1, \cdots, N$ and $\psi_\mu(t)$ is then be calculated with (27) and (28) to assemble $\Psi_\mu(\theta)$.

Lemma 2. The gradient of the cost function (19) with respect to η, with μ held constant, is given by

$$F_\eta(\theta) = -\nabla_\eta V^N(\theta)\big|_\mu = \Psi^T_\eta(\theta)E(\theta) \tag{30}$$

where

$$\nabla_\eta = \begin{bmatrix} \frac{\partial\hat{y}(t)}{\partial a_1} & \cdots \frac{\partial\hat{y}(t)}{\partial a_{n_a}} & \frac{\partial\hat{y}(t)}{\partial b_0} & \cdots & \frac{\partial\hat{y}(t)}{\partial b_M} \end{bmatrix}$$

and $\Psi_\eta(\theta)$ is the gradient of the prediction error vector $\nabla_\eta E(\theta)$. For a time sample t,

$$\psi_\eta(t,\theta) = \nabla_\eta \hat{y}(t) = \left[\frac{\partial \hat{y}(t)}{\partial a_1} \ \cdots \frac{\partial \hat{y}(t)}{\partial a_{n_a}} \ \frac{\partial \hat{y}(t)}{\partial b_0} \ \cdots \ \frac{\partial \hat{y}(t)}{\partial b_M} \right], \quad (31)$$

and $\Psi_\eta(\theta) \in \Re(N \times (n_a + n_b + 1))$ is written

$$\Psi_\eta(\theta) = \frac{1}{\sqrt{N}} \left[\psi_\eta(1) \ \cdots \ \psi_\eta(N) \right]^T. \quad (32)$$

The prediction error gradient $\Psi_\eta(\theta)$, is calculated by filtering $\hat{w}(t)$ using the parameters μ, η, and closed-loop signal relations from (11).

The proof of Lemma 2 follows the same format as Lemma 1 and is omitted due to space constraints.

Iterative Parameter Update: The parameter update is written in terms of an iteration step in the μ direction that generates $\hat{\theta}_{\mu,s}$, and a step in the η direction that generates $\hat{\theta}_{s+1}$, using the gradient expressions (26) and (32) and prediction error in (18). Using the subscript s to denote the iteration step,

$$\hat{\theta}_{\mu,s} = \hat{\theta}_s - \gamma(\hat{\theta}_s) N_\mu(\hat{\theta}_s)^{-1} F_\mu(\hat{\theta}_s), \quad (33)$$

$$\hat{\theta}_{s+1} = \hat{\theta}_{\mu,s} - \gamma(\hat{\theta}_{\mu,s}) N_\eta(\hat{\theta}_{\mu,s})^{-1} F_\eta(\hat{\theta}_{\mu,s}), \quad (34)$$

where $N_\mu(\hat{\theta}_s)$ and $N_\eta(\hat{\theta}_s)$ are approximations of the Hessians given by,

$$N_\mu(\hat{\theta}_s) = \left[I + \Psi_\mu(\hat{\theta}_s)\gamma(\hat{\theta}_s)\Psi_\mu^T(\hat{\theta}_s) \right],$$
$$N_\eta(\hat{\theta}_{\mu,s}) = \left[I + \Psi_\eta(\hat{\theta}_{\mu,s})\gamma(\hat{\theta}_{\mu,s})\Psi_\eta^T(\hat{\theta}_{\mu,s}) \right].$$

In the parameter updates $\gamma(\hat{\theta}_{\mu,s})$ and $\gamma(\hat{\theta}_{\eta,s})$ are the Levenberg-Marquardt step sizes. $F_\mu(\hat{\theta}_s)$ and $N_\mu(\hat{\theta}_s)$ are calculated via (24) to update $\hat{\theta}_{\mu,s}$. Then new estimates of $\hat{u}_{\mu,s}(t)$ and $\hat{y}_{\mu,s}(t)$ are generated via (11). $F_\eta(\hat{\theta}_s)$ and $N_\eta(\hat{\theta}_s)$ are calculated using $\hat{u}_{\mu,s}(t)$ and $\hat{y}_{\mu,s}(t)$ and (30). The parameter update $\hat{\theta}_{s+1}$ is calculated via (33). The two-step parameter updates in (33) are repeated until the parameter converges using the stopping criterion, for a user defined τ by,

$$\|\hat{\theta}_s - \hat{\theta}_{s-1}\| / \|\hat{\theta}_{s-1}\| > \tau. \quad (35)$$

6. EXPERIMENTAL RESULTS

We apply the proposed procedure to closed-loop time domain data from a first principles dynamic simulation of a 5 megawatt *Taurus*^{TM} 60 simple-cycle conventional combustion gas turbine. In this experiment, the gas turbine is driving an electric generator for power generation in an "island mode" operation. In island mode, the frequency, and therefore shaft speed varies, as there is no utility grid to regulate the frequency. The fuel control is operating in closed-loop to maintain the output frequency at 60Hz ($r_0 = 100[\%]$). The system is uniformly sampled at 10Hz and for a total of 16384 data samples of the shaft speed reference input $r(t)$ in [%], controller output $u(t)$ in [%], and shaft speed $y(t)$ in [%] are recorded. Each experiment used $N = 8184$ data points for identification and $N = 8184$ for validation. The excitation is an additive band-limited uniformly distributed noise signal $r_e(t) \in [\,-4,\ 4\,]$ chosen to excite a significant span of the fuel valve's command input range such that $r(t) = r_0 + r_e(t)$.

The nominal fuel valve characteristic $f_0(\cdot)$ is a 4th order polynomial with known coefficients α_p, and the contamination is represented by Δ_f that has been constructed to represent an approximate 10% reduction in the local flow area around a command of 55 and 60%.

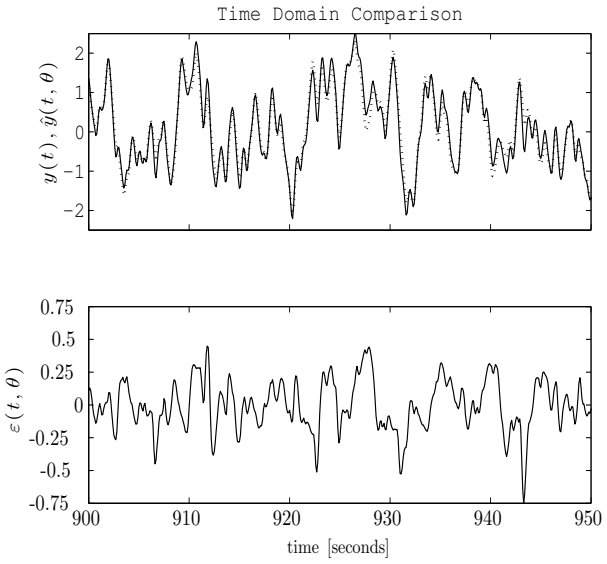

Fig. 3. (Top) Comparison of shaft speed $y(t)$ from first principles simulation and $\hat{y}(t,\hat{\theta}_s^N)$ for $n_a^* = 3$ and $n_b^* = 2$. (Bottom) Estimation error $\varepsilon(t,\hat{\theta}_s)$.

$$f_0(\cdot) = \sum_{p=1}^{4} \alpha_p u(t)^p \qquad \Delta_f = \kappa_i \exp\left(\frac{-(u(t)-c_i)^2}{\chi_i} \right)$$
$$\kappa_i = [0.02, 0.03], \quad c_i = [55, 60], \qquad \chi_i = [20, 10]$$

In the estimation there are several design parameters available to the user, those related to identification of the linear dynamics $\left[n_a^*, n_b^*, n_e\right]$, and those related to identification of the nonlinearity,$[m, \sigma]$. Table 1 compares the performance for variation in the linear plant parametrization.

Figure 3 demonstrates the prediction accuracy of the low-order model,$\hat{\theta}_s^N$, of shaft speed and the frequency response is of the identified linear plant is compared to the frequency response function (FRF) of the first principles simulation data in Figure 4 for . Identification of the nonlinearity $\delta(\cdot)$, is demonstrated in Figure 5 for an uncontaminated fuel valve and a contaminated fuel valve.

	$\sigma = 0.35$						
	N	s	n_a^*	n_b^*	n_e	M	$V^N(\theta)$
$\hat{\theta}_0^N$	8192	0	3	2	5	9	0.302
$\hat{\theta}_s^N$	8192	5	3	2	0	9	0.214
$\hat{\theta}_0^N$	8192	0	3	2	3	9	0.423
$\hat{\theta}_s^N$	8192	10	3	2	0	9	0.314
$\hat{\theta}_0^N$	8192	0	3	2	1	9	0.433
$\hat{\theta}_s^N$	8192	10	3	2	0	9	0.343

Table 1. Experimental results illustrating the benefit of the high order initializion on $V^N(\theta)$ and the number of iterations required for convergence.

7. CONCLUSION

The discussion demonstrates that a low-order linear dynamic model within a Hammerstein model structure successfully captures the nonlinear behavior of a full nonlinear model of a gas

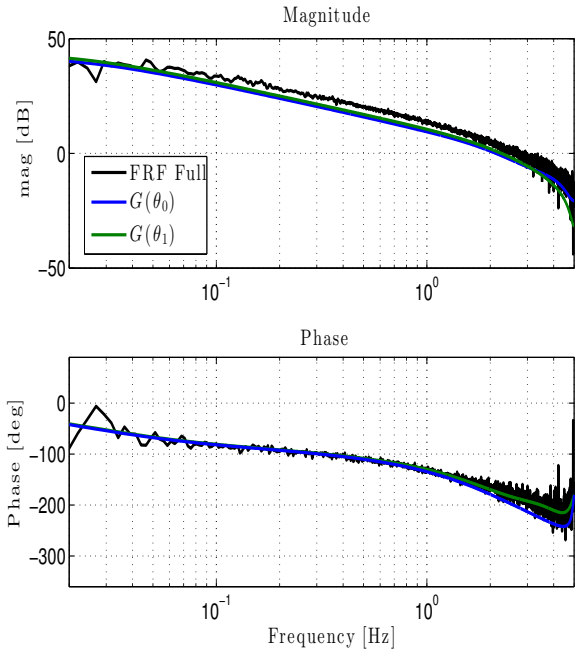

Fig. 4. Frequency Response Estimates for (Input) $\hat{\delta}_w(t)$ (Output) - Shaft Speed $y(t)$ for the FRF and magnitude plots of (Input) - $\hat{\delta}_w(t)$ and (Output) - $\hat{y}(t)$ of initial high-order $G(\hat{\theta}_0^N)$ and low-order $G(\hat{\theta}_s^N)$ resembling the FRF.

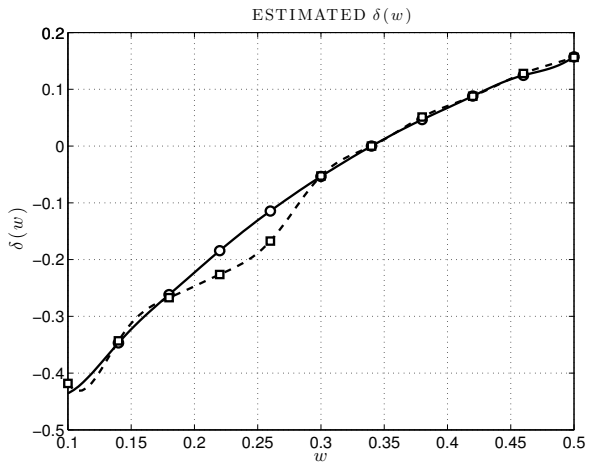

Fig. 5. Identified $\delta(\cdot)$ for perfect knowledge of the fuel valve characteristic $f(\cdot) = f_0(\cdot)$ (solid, ◯ marker) and a contaminated fuel valve $f(\cdot) = f_0(\cdot) + \Delta_f$ (dashed, □ marker).

turbine. The high-order parameter initialization and subsequent model reduction procedure allows for a superior initial estimate and parsimonious linear dynamic plant model compared to methods where the model orders are held constant. The gradient expressions are presented for the closed-loop system may be used in a variety of optimization methods and nonlinearities that may be approximated with a continuously differentiable function. The approach to this specific problem should readily extend to many applications where the system design is well understood and similar condition monitoring objectives are needed. This excitation level used in the simulation example is typically not achievable in practice, although the use of simulation to provide benchmark models for use on physical plants in operation shows promise. The low order linear plant models produced in this algorithm are realistic for use in operation with lower levels of excitation. The iterative procedure is effective at computing a low-order linear model in the Hammerstein framework and while it is computationally intensive, with new estimates of the noise free signal estimates as well as gradient updates at each step, it is less computationally demanding than computing the gradients using the non-minimal parametrization.

ACKNOWLEDGEMENTS

This work was supported through funding from Solar Turbines Inc. under grant P1100040.

REFERENCES

Bai, E.W. (1998). An optimal two-stage identification algorithm for Hammerstein-wiener nonlinear systems. *Automatica*, 34, 333–338.

Breikin, T., Kulikov, G., and Arkov, V. (2004). Regularisation approach for real-time modelling of aero gas turbines. *Control Engineering Practice*, 12, 401–407.

Cézac, P., Serin, J.P., Reneaume, J.M., Mercadier, J., and Mouton, G. (2008). Elemental sulphur deposition in natural gas transmission and distribution networks. *The Journal of Supercritical Fluids*, 44, 115–122.

Dai, X. andBreikin, T. and Wang, H. (2006). An algorithm for identification of reduced-order dynamic models of gas turbines. In *First International Conference on Innovative Computing, Information and Control*, 134–137.

De Bruyne, F., Anderson, B.D.O., Linard, N., and Gevers, M. (1999). Gradient expressions for a closed-loop identification scheme with a tailor-made parametrization. *Automatica*, 35, 1867–1871.

Han, Y. and De Callafon, R. (2011). Closed-loop identification of Hammerstein systems using iterative instrumental variables. In B. Sergio (ed.), *18th IFAC World Congress*, 13930–13935. Milano, Italy.

Laurain, V., Gilson, M., and Garnier, H. (2009). Refined instrumental variable methods for identifying hammerstein models operating in closed loop. In *Proceedings of the 48th IEEE Conference on Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009*, 3614 –3619.

Laurain, V., Gilson, M., Tth, R., and Garnier, H. (2010). Refined instrumental variable methods for identification of LPV BoxJenkins models. *Automatica*, 46(6).

Lippmann, R. (1991). A critical overview of neural network pattern classifiers. In *IEEE Neural Networks for Signal Processing Workshop*, 266–275.

Liu, Y. and Anderson, B.D.O. (1989). Singular perturbation approximation of balanced systems. In , *Proceedings of the 28th IEEE Conference on Decision and Control, 1989*, 1355–1360 vol.2.

Narendra, K.S. and Gallman, P.G. (1966). An iterative method for the identification of nonlinear systems using a Hammerstein model. *IEEE Transactions on Automatic Control*, 11, 546–550.

Van Pelt, T.H. and Bernstein, D.S. (2000). Nonlinear system identification using Hammerstein and nonlinear feedback models with piecewise linear static maps - part 1: Theory. In *Proc. American Control Conference*, 225–229. Chicago, Illinois.