Automatica 48 (2012) 2189–2193

Contents lists available at SciVerse ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

Hammerstein system identification using nuclear norm minimization☆

Younghee Han [1], Raymond A. de Callafon

University of California at San Diego, Mechanical and Aerospace Engineering, 9500 Gilman Drive, La Jolla, CA, 92093-0401, USA

ARTICLE INFO

Article history:
Received 11 August 2011
Received in revised form
19 December 2011
Accepted 13 March 2012
Available online 30 June 2012

Keywords:
Hammerstein system identification
Block-oriented system identification
Rank minimization
Semidefinite programming problem

ABSTRACT

This paper presents a new method for the identification of Hammerstein systems. The parameter estimation problem is formulated as a rank minimization problem by constraining a finite dimensional time dependency between signals. Due to the unknown intermediate signal, the rank minimization problem cannot be solved directly. Thus, the rank minimization problem is reformulated as an intermediate signal construction problem. The main assumption used in this paper is that static nonlinearity is monotonically non-decreasing in order to guarantee a unique combination of a static nonlinear block and a Finite Impulse Response (FIR) linear block. The rank minimization is then relaxed to a convex optimization problem using a nuclear norm. The main contribution of this paper is that the proposed method extends the rank minimization approach to Hammerstein system identification, and does not need a bilinear parametrization and singular value decomposition (SVD), which are commonly used in two-step approaches for Hammerstein system identification.

1. Introduction

A Hammerstein system has a block oriented structure where a static input nonlinearity and a linear dynamic system are separated, as shown in Fig. 1. A comprehensive overview of Hammerstein nonlinear system identification can be found in Giri and Bai (2010). Regarding Fig. 1, the objective of this paper is to formulate a procedure that allows the characterization and identification of the nonlinear static function $f(\cdot)$ and the linear dynamic system $G(q)$ individually based on the input $u(t)$ and the output $y(t)$ observation. This is done in a novel way by the reconstruction of the intermediate signal $x(t)$ with conditions on the finite dimensional dynamic representation of the linear systems $G(q)$ and the memoryless static nonlinearity $f(\cdot)$. A similar approach was also considered in Zhang, Iouditski, and Ljung (2007). The main contribution of this paper is that the proposed method extends the rank minimization approach to Hammerstein system identification, and does not need a bilinear parametrization and singular value decomposition (SVD), which are commonly used in two-step approaches for Hammerstein system identification. The order of a finite dimensional model can be expressed as the rank of a matrix that is filled with input and output measurement. If the set of feasible models is described by convex constraints, then choosing the simplest model can often be expressed as a rank minimization problem (Fazel, Hindi, & Boyd, 2004). Based on this idea, the rank minimization approach is used to formulate a convex parameter estimation problem via nuclear norm relaxation, where the nuclear norm of H is defined as the summation of its singular values as $\|H\|_* = \sum_{i=1}^{r} \sigma_i(H)$. The use of nuclear norm approximation with application to system identification can be found in Liu and Vandenberghe (2009) and with application to Hammerstein systems in Falck, Suykens, Schoukens, and De Moor (2010).

2. System description

2.1. Hammerstein system

A rank constrained Semidefinite Programming (SDP) problem will be formulated in such a way that $x(t)$ and $u(t)$ are related via a memoryless static nonlinearity and that $x(t)$ and $y(t)$ are related via a linear dynamical system with the smallest McMillan degree.

Condition 1. I. *The static nonlinear function (the relation between $u(t)$ and $x(t)$) has no memory.*
II. *The linear dynamic system has a finite, but unknown McMillan degree n, relating a finite number of the past input samples to the past output samples.*

The properties in Condition 1 are used to formulate a procedure to reconstruct the unmeasurable intermediate signal $x(t)$ based on rank minimization.

☆ The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Er-Wei Bai under the direction of Editor Torsten Söderström.

E-mail addresses: y3han@ucsd.edu (Y. Han), callafon@ucsd.edu (R.A. de Callafon).

[1] Tel.: +1 858 2329039; fax: +1 858 8223107.

0005-1098/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2012.06.013

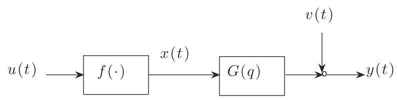

Fig. 1. Hammerstein system consists of a static nonlinear block followed by a linear dynamic block.

2.2. Modeling of static nonlinearity

It is well known that the static nonlinear function can be approximated as a linear combination of a finite set of basis functions as $f(u(t)) = \sum_{m=1}^{M} \theta_m \phi_m(u(t))$, where θ_m are weighting parameters to be estimated and $\phi_m(\cdot)$ are user chosen basis functions. In this paper, a piecewise linear approximation of the static nonlinearity $f(\cdot)$ using piecewise triangle functions is used. Using triangle basis functions $f_m(\cdot)$, the static nonlinearity $f(\cdot)$ is assumed to satisfy the following condition

$$\sup_{u(t)\in[u_{min},u_{max}]} \lim_{M\to\infty} \sum_{m=1}^{M} |\theta_m f_m(u(t)) - f(u(t))| = 0, \tag{1}$$

where the center location vector $m = [m_1 \cdots m_M]^T$ spans the amplitude of the input vector $u(t)$ and the amplitude vector $\theta = [\theta_1 \cdots \theta_M]^T$ is to be estimated. The condition in (1) indicates that the static nonlinearity $f(\cdot)$ can be approximated arbitrarily well with a dense grid of triangular basis functions. In order to define a piecewise linear approximation of the static nonlinearity $f(\cdot)$, a finite value M in (1) can be chosen, whereas the points $m_1, \ldots, m_M$ of a grid over $[u_{min}, u_{max}]$ can be chosen linearly spaced or at strategic locations. Each triangle function $f_m(u(t))$ in (1) has nonzero values through two segments and zeros elsewhere except for the first and the last intervals of the grid.

$$f_m(u(t)) = \begin{cases} \dfrac{u(t) - m_{l-1}}{m_l - m_{l-1}} & \text{for } m_{l-1} \le u(t) < m_l \\ \dfrac{m_{l+1} - u(t)}{m_{l+1} - m_l} & \text{for } m_l \le u(t) < m_{l+1} \\ 0 \quad \text{Otherwise} \end{cases}$$

$$f_1(u(t)) = \begin{cases} \dfrac{m_2 - u(t)}{m_2 - m_1} & \text{for } m_1 \le u(t) < m_2 \\ 0 \quad \text{Otherwise} \end{cases}$$

$$f_M(u(t)) = \begin{cases} \dfrac{u(t) - m_{M-1}}{m_M - m_{M-1}} & \text{for } m_{M-1} < u(t) \le m_M \\ 0 \quad \text{Otherwise.} \end{cases}$$

In each segment of the m-axis, the resulting linear function is defined by two overlapping triangle functions in the segment. Let $\hat{x}(t) = f(u(t), \theta)$ be the approximation of $x(t)$ and θ is the amplitude parameter

$$\theta = \begin{bmatrix} \theta_1 & \cdots & \theta_M \end{bmatrix}^T. \tag{2}$$

Then, $\hat{x}(t)$ can be written as

$$\hat{x}(t) = \rho(u(t))\theta, \tag{3}$$

where $\rho(u(t))$ is defined as

$$\rho(u(t)) = \left[\cdots 0 \frac{m_{k+1} - u(t)}{m_{k+1} - m_k} \frac{u(t) - m_k}{m_{k+1} - m_k} 0 \cdots \right]$$
$$\text{for } m_k \le u(t) < m_{k+1},$$

where m_k and m_{k+1} are the center locations of the triangle basis functions. There could be many possible combinations of a static nonlinear block and a Finite Impulse Response (FIR) linear block that satisfy Condition 1 and (3). In order to limit the number of possible selections for a linear block and a static nonlinear block, in this paper, a monotonically non-decreasing static nonlinearity with the user chosen maximum slope of 1 is considered as follows:

Condition 2. I. *The static nonlinear function is monotonically non-decreasing with the maximum slope of* 1*:*

$$(\hat{x}(i) - \hat{x}(j))(\hat{x}(i) - \hat{x}(j) - u(i) + u(j)) \le 0 \quad \forall i > j.$$

Without loss of generalization, Condition 2 guarantees a solution for an FIR linear system and serves as a normalization condition on the static nonlinearity. A Hammerstein system with a monotonically non-decreasing static nonlinear function can be used to model many control, mechanical, electrical, chemical, and biological systems with various static nonlinear functions, such as saturation, deadzone, quantization, etc. The examples can be found in Giri and Bai (2010).

2.3. The input–output map of the dynamical system

Let $g(i),\ i = 0, 1, \ldots$ be the causal sequence of unit impulse responses for $G(q)$. The relationship between the intermediate signal $x(t)$ and the output $y(t)$ can be described by the convolution as

$$y(t) = \sum_{i=0}^{\infty} g(i)x(t-i) + v(t),$$

where $v(t)$ is noise. Due to Condition 1 (finite McMillan degree), for a finite data sequence of $N = n_1 + n_2$ data points and a zero initial condition, the relationship between the intermediate signal $x(t)$ and the output $y(t)$ can be described by

$$Y = HX_p + TX_f + V, \tag{4}$$

where Y is the data matrix, including the future output $y(t)$, defined by

$$Y = \begin{bmatrix} y(1) & \cdots & y(n_2) \\ \vdots & \ddots & \vdots \\ y(n_1) & \cdots & \cdots y(n_1 + n_2 - 1) \end{bmatrix}, \tag{5}$$

X_p is the data matrix, including the past intermediate signal, defined by

$$X_p = \begin{bmatrix} x(0) & x(1) & \cdots & x(n_2 - 1) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & x(0) \end{bmatrix},$$

X_f is the data matrix, including the future intermediate signal, defined by

$$X_f = \begin{bmatrix} x(1) & x(2) & \cdots & x(n_2) \\ \vdots & \vdots & \ddots & \vdots \\ x(n_1 + 1) & x(n_1 + 2) & \cdots & x(n_1 + n_2) \end{bmatrix}, \tag{6}$$

H is the Hankel matrix defined by

$$H = \begin{bmatrix} g(1) & \cdots & g(n_2) \\ \vdots & \ddots & \vdots \\ g(n_1) & \cdots & \cdots g(n_1 + n_2 - 1) \end{bmatrix},$$

T is the Toeplitz matrix defined by

$$T = \begin{bmatrix} g(0) & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ g(n_1 - 1) & g(n_1 - 2) & \cdots & g(0) & 0 \end{bmatrix},$$

and V is the matrix, including noise data, defined by

$$V = \begin{bmatrix} v(1) & \cdots & v(n_2) \\ \vdots & \ddots & \vdots \\ v(n_1) & \cdots & \cdots v(n_1 + n_2 - 1) \end{bmatrix}.$$

The order of the linear dynamical system is determined by the $rank(H)$ as H is simply the product of the extended observability and controllability matrices (Goethals, Pelckmans, Suykens, & De Moor, 2005). A lower order model, consistent with the input and output signals can be estimated by minimizing the rank of H.

2.4. Problem formulation

The effect of X_f to Y in (4) can be removed by the orthogonal projection of Y onto the null space of X_f. With the projection matrix $X_f^{\perp} = I - X_f^T(X_f X_f^T)^{\dagger} X_f$, the effect of X_f is removed (Ljung, 1999). Then,

$$YX_f^{\perp} = HX_p X_f^{\perp} + VX_f^{\perp}.$$

In order to remove the effect of noise, the projection Y can be subsequently weighted by matrices W_1 and W_2 such that

$$W_1 YX_f^{\perp} W_2 = W_1 HX_P X_f^{\perp} W_2 + W_1 VX_f^{\perp} W_2 \tag{7}$$

in which W_1 and W_2 are chosen to be rank-preserving and such that $W_1 VX_f^{\perp} W_2 \to 0$ as the number of samples $N \to \infty$. Details of the role and choice of weighting matrices W_1 and W_2 can be found in Van Overschee and De Moor (1996). The result in (7) indicates that the rank minimization problem for H can be rewritten as the rank minimization problem for $YX_f^{\perp}$. Unfortunately, the rank minimization problem for $YX_f^{\perp}$ cannot be solved directly since X is unknown. In this section, the rank minimization problem of $YX_f^{\perp}$ is reformulated using LQ decomposition of data matrix $\begin{bmatrix} X_f \\ Y \end{bmatrix}$ so that the rank minimization problem can be solved without knowing $X_f^{\perp}$. Let the LQ decomposition of the data matrix $\begin{bmatrix} X_f \\ Y \end{bmatrix}$ be given by

$$\begin{bmatrix} X_f \\ Y \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} Q_1^T \\ Q_2^T \end{bmatrix}, \tag{8}$$

where L_{11}, L_{22} are lower triangular and Q_1, Q_2 are orthogonal. From (8), Y can be written as follows:

$$Y = L_{21} L_{11}^{-1} X_f + L_{22} Q_2^T. \tag{9}$$

The first term in (9) is spanned by the row vectors in X_f and the second term is orthogonal to it. Thus, the orthogonal projection of Y onto the null space of X_f can be written as (Goethals et al., 2005; Ljung, 1999)

$$YX_f^{\perp} = L_{22} Q_2^T. \tag{10}$$

Since Q_2 is orthogonal, (10) indicates

$$rank(YX_f^{\perp}) = rank(L_{22}).$$

As a result, the rank minimization problem for $YX_f^{\perp}$ can be rewritten as the rank minimization problem for L_{22}. Using the signal $\hat{x}(t)$ in (3), X_f in (6) can be reconstructed using $\hat{x}$ as

$$\hat{X}_f = U_2 \Theta, \tag{11}$$

where Θ is the block diagonal matrix, including θ, defined by

$$\Theta = \begin{bmatrix} \theta & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \theta & \cdots & \mathbf{0} \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ddots & \theta \end{bmatrix}, \tag{12}$$

where $\mathbf{0}$ is a zero vector and U_2 is the data matrix including the input $u(t)$, defined by

$$U_2 = \begin{bmatrix} \rho(u(1)) & \cdots & \rho(u(n_2)) \\ \vdots & \ddots & \vdots \\ \rho(u(n_1+1)) & \cdots & \rho(u(n_1+n_2)) \end{bmatrix}. \tag{13}$$

From (8), we have

$$rank\left(\begin{bmatrix} X_f \\ Y \end{bmatrix}\right) = rank\left(\begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix}\right).$$

Let $L = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix}$. The following lemma explains the rank inequality condition for the block matrix L.

Lemma 1. *The rank of the block matrix L satisfies the following inequality:*

$$rank\left(\begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix}\right) \geq rank(L_{11}) + rank(L_{22}).$$

With Lemma 1, we have

$$rank(L_{11}) + rank(L_{22}) \leq rank\left(\begin{bmatrix} X_f \\ Y \end{bmatrix}\right). \tag{14}$$

Since $rank(L_{11}) = rank(X_f)$ from (8), (14) can be written as

$$rank(X_f) + rank(L_{22}) \leq rank\left(\begin{bmatrix} X_f \\ Y \end{bmatrix}\right).$$

From (11), we have

$$rank(X_f) = rank(U_2 \Theta) = constant.$$

In this section, the rank minimization problem for L_{22} is relaxed to the upper bound minimization problem for $rank(L_{22})$, which is equivalent to the minimization problem for $rank\left(\begin{bmatrix} X_f \\ Y \end{bmatrix}\right)$. As a result, system parameters for a lower order model will be estimated by minimizing $rank\left(\begin{bmatrix} X_f \\ Y \end{bmatrix}\right)$ under the constraints developed based on Condition 2.

3. Rank minimization for intermediate signal reconstruction

With the parametrization and constraints explained in Section 2, an optimization problem can be written as

Optimization Problem 1.

Consider

variable θ in (2)

to create $\hat{x}$ in (3) and Θ in (12),

Minimize

$$rank\left(\begin{bmatrix} \hat{X}_f \\ Y \end{bmatrix}\right),$$

where $\hat{X}_f = U_2 \Theta$, and Y is given in (5),
with U_2 defined in (13),
subject to

$$(\hat{x}(i) - \hat{x}(j))(\hat{x}(i) - \hat{x}(j) - u(i) + u(j)) \leq 0 \quad \forall i > j.$$

Optimization Problem 1 results in the optimal solution for the system parameter θ that is used to construct the intermediate signal x using the relationship in (3), automatically satisfying Condition 1-I. Unfortunately, the rank constraint in Optimization Problem 1 is not convex. Minimizing the nuclear norm instead of the rank of the matrix is a convex relaxation of the rank minimization problem (Fazel, Hindi, & Boyd, 2001; Fazel et al., 2004). The motivation for this nuclear norm relaxation will be explained in Section 3.1 by showing that the nuclear norm is the convex envelope of the rank function on the set of matrices with norms less than 1.

3.1. Convex envelope of rank

Lemma 2 (*Fazel et al. 2001*)**.** *The convex envelope of the function $\phi(X) = Rank(X)$, on $C = \{X \in \Re^{m\times n} \mid \|X\| \leq 1\}$, is $\phi_{env}(X) = \|X\|_*$.*

Proof. Can be found in Fazel et al. (2001). □

Lemma 2 has the following implications. Suppose the feasible set is bounded by Q, i.e., for all $X \in C$, we have $\|X\| \leq Q$. The convex envelop of $RankX$ on $\{X|\ \|X\| \leq Q\}$ is given by $\frac{1}{Q}\|X\|_*$. In particular, for all $X \in C$, we have $RankX \geq \frac{1}{Q}\|X\|_*$. Thus, by solving the nuclear norm minimization problem, we obtain a lower bound on the optimal value of the original rank minimization problem (Fazel et al., 2001). Using the nuclear norm relaxation for rank minimization, Optimization Problem 1 will be reformulated as a convex problem. First, let us express the constraints in Optimization Problem 1 as Linear Matrix Inequalities (LMIs). Let $\delta x = [\delta x(1) \ \cdots \ \delta x(k_{max})]^T$, where $\delta x(k) = \hat{x}(i) - \hat{x}(j)$ and $\delta u = [\delta u(1) \ \cdots \ \delta u(k_{max})]^T$, where $\delta u(k) = u(i) - u(j)$ for all $i > j$ and $k_{max} = \sum_{k=1}^{N-1} k$. Let $\Delta X = diag(\delta x)$ and $\Delta U = diag(\delta u)$, where $diag(\delta x)$ is the diagonal matrix whose diagonal elements are the elements of δx. Then, the constraints in Optimization Problem 1 can be written as $\Delta X(\Delta X - \Delta U) \leq 0$. Nuclear norm minimization coupled with linear constraints lead to a Semidefinite Programming (SDP) problem. This is easier to solve and the solution is close to the solution of the original non-convex problem (Fazel et al., 2001, 2004). Using SDP relaxation, Optimization Problem 1 can be rewritten as the following convex optimization problem:

Optimization Problem 2.

Consider

variable θ in (2)

to create $\hat{x}$ in (3) and Θ in (12),

Minimize

$$\left\| \begin{bmatrix} \hat{X}_f \\ Y \end{bmatrix} \right\|_*,$$

where $\hat{X}_f = U_2\Theta$, *and* Y *is given in* (5),
with U_2 *defined in* (13),
subject to

$$\Delta X(\Delta X - \Delta U) \leq 0$$

where

$$\|H\|_* = \sum_{i=1}^{r} \sigma_i(H).$$

As a summary, the intermediate signal $x(t)$ in Fig. 1 is parametrized by (3) and estimation of the unknown coefficients θ_i, $i = 2, \ldots, M$ in (3) is solved by computing the solution to the SDP problem given in Optimization Problem 2. The optimization guarantees that $\hat{x}(t)$ and $u(t)$ are related via a memoryless static nonlinearity and guarantees that $\hat{x}(t)$ and $y(t)$ are related via a linear dynamical system with the smallest McMillan degree. Once $\hat{x}(t)$ has been reconstructed, the identification of $G(q)$ from $\hat{x}(t)$ and $y(t)$ can be solved with a standard Prediction Error (PE) identification method (Ljung, 1999).

4. Numerical example

In this section, a numerical example of the Hammerstein system identification using the proposed identification method is presented. An excitation signal $u(t)$ is zero mean white noise with a standard deviation of 3. The output disturbance $v(t)$ is filtered white noise, where the filtering properties are unknown. For the system identification, a set of data (ten different measurements) is generated from the Hammerstein system. In the data set, SNR is greater than 20 dB. $m = [\min(u(t)) \ -3 \ -1 \ 0 \ 1 \ 3 \ \max(u(t))]$, and

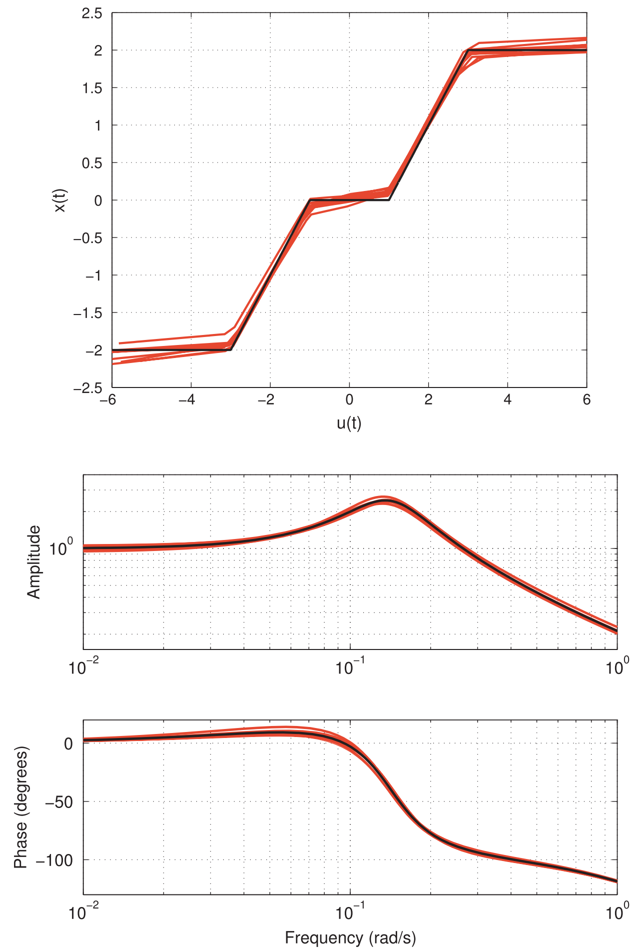

Fig. 2. The plot of the identified static nonlinearity function (top). The Bode plot of the identified linear dynamic system (bottom). The black line indicates the real Hammerstein system. The red (shaded) lines indicate estimated systems by using ten different data. The SNR of each data in the set is greater than 20 dB. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

$n_a = 2$ and $n_b = 2$ are used to model the static nonlinear function and the linear dynamic system respectively. $W_1 = W_2 = I$ are chosen in this example. In order to solve the SDP problem (Optimization Problem 2), SeDuMi (Self-Dual-Minimization) Sturm (1999) and YALMIP (Yet Another LMI Parser) Löfberg (2004) are used. The specifications of the Hammerstein system are as follows:

$$\begin{cases} \text{Linear dynamical system:} \\ G(q) = \dfrac{0.1994q^{-1} - 0.1804q^{-2}}{1 - 1.886q^{-1} + 0.9048q^{-2}} \\ \text{Static nonlinearity:} \\ \quad f(u(t)) = \begin{cases} 2 & \text{if } x(t) > 3 \\ u(t) + 1 & \text{if } 1 < x(t) \leq 3 \\ 0 & \text{if } |x(t)| \leq 1 \\ u(t) - 1 & \text{if } -3 \leq x(t) < -1 \\ -2 & \text{if } x(t) < -3 \end{cases} \\ \text{Noise dynamics:} \\ H(q) = \dfrac{1 + 0.5q^{-1}}{1 - 0.85q^{-1}} \end{cases}$$

The estimation results are shown in Figs. 2 and 3. As shown in Figs. 2 and 3, the proposed algorithm provides excellent identification results for data with SNR greater than 20 dB. The pole location and the characteristics of the static nonlinearity are very well captured.

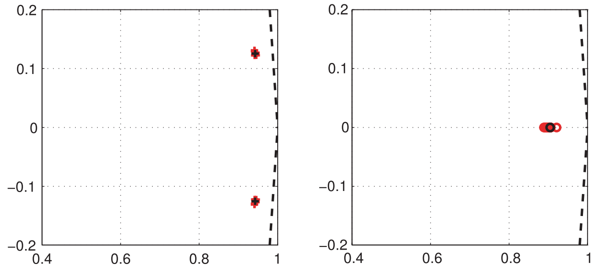

Fig. 3. Pole (left) and zero (right) locations of the identified linear dynamical system. The black cross and circle indicate the real linear dynamical system. The red (shaded) crosses and circles indicate estimated linear dynamical systems. SNR is greater than 20 dB. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

5. Conclusion

In this paper, the Hammerstein system parameter identification problem is formulated as a nuclear norm minimization problem. First, the system parameter identification problem is formulated as a rank minimization problem to reconstruct the intermediate signal between the static nonlinearity and the linear dynamics in a Hammerstein system. This non-convex optimization problem is then reformulated as a convex optimization problem using a nuclear norm relaxation. Once the system parameters for the static nonlinearity are estimated, an intermediate signal can be created to facilitate the identification of the linear dynamic system. The main assumption used in this paper is that static nonlinearity is monotonically non-decreasing in order to guarantee a unique combination of a static nonlinear block and a Finite Impulse Response (FIR) linear block. The proposed identification method is applied to simulation data from a Hammerstein system. The numerical simulation result shows the effectiveness of the proposed identification method.

References

Falck, T., Suykens, J.A.K., Schoukens, J., & De Moor, B. (2010). Nuclear norm regularization for overparametrized Hammerstein Systems. In *49th IEEE conference on decision and control* (pp. 7202–7207).

Fazel, M., Hindi, H., & Boyd, S. (2001). Rank minimization heuristic with application to minimum order system approximation. In *Proc. American control conference* (pp. 4734–4739).

Fazel, M., Hindi, H., & Boyd, S. (2004). Rank minimization and applications in system theory. In *Proc. American control conference* (pp. 3273–3278).

Giri, F., & Bai, E. W. (2010). *Block oriented nonlinear system identification*. Berlin: Springer-Verlag.

Goethals, I., Pelckmans, K., Suykens, J. A. K., & De Moor, B. (2005). Subspace identification of hammerstein systems using least squares support vector machines. *IEEE Transactions on Automatic Control*, *50*, 1509–1519. Special issue on system identification.

Liu, Z., & Vandenberghe, L. (2009). Interior-point method for nuclear norm approximation with application to system identification. *SIAM Journal on Matrix Analysis and Applications*, *31*, 1235–1256.

Ljung, L. (1999). *System identification: theory for the user*. Upper Saddle River, New Jersey: Prentice-Hall, Inc.

Löfberg, J. (2004). YALMIP: A toolbox for modeling and optimization in MATLAB. In *IEEE international symposium on CACSD conference* (pp. 284–289).

Sturm, J. F. (1999). Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, *11-12*, 625–653.

Van Overschee, P., & De Moor, B. (1996). *Subspace identification for linear systems: theory, implementation, applications*. Kluwer Academic Publishers.

Zhang, Q., Iouditski, A., & Ljung, L. (2007). Identification of Wiener system with monotonous nonlinearity. In *14th IFAC symposium on system identification* (pp. 1092–1097).

Younghee Han is a Ph.D. candidate in the Department of Mechanical and Aerospace Engineering (MAE) at the University of California, San Diego (UCSD). She received her M.Sc. (2006) in Mechanical Engineering from the University of Kentucky. In 2007, she started her Ph.D. study in the Dept. of MAE at UCSD and is currently affiliated with the System Identification and Control Laboratory (SICL) in the Cymer Center for Control Systems and Dynamics (CCSD) at UCSD. Younghee Han's current research interests lie in modeling, estimation and control of dynamic systems. In particular, she is interested in designing and analyzing (block-oriented) nonlinear system identification techniques.

Raymond A. de Callafon is a Professor with the Department of Mechanical and Aerospace Engineering (MAE) at the University of California, San Diego (UCSD). Raymond de Callafon received his M.Sc. (1992) and his Ph.D. (1998) degrees in Mechanical Engineering from the Delft University of Technology in the Netherlands. Since 1998 he has been with the Dept. of MAE and he is currently directing the System Identification and Control Laboratory (SICL) and is an affiliated faculty of the Center for Magnetic Recording Research (CMRR) directing the CMRR servo laboratory. His research interests include topics in the field of experiment-based approximation modeling, control relevant system identification and recursive/adaptive control.