Preprints of the 16th IFAC Symposium on System Identification
The International Federation of Automatic Control
Brussels, Belgium. July 11-13, 2012

Identification of a Wiener System via Semidefinite Programming

Younghee Han and Raymond A. de Callafon

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

Abstract: This paper presents a new method for the identification of Wiener systems in the presence of output noise. The Wiener system identification problem is formulated as a convex Semidefinite Programming (SDP) problem by constraining a finite dimensional time dependency between signals. The main contribution of this paper is that the proposed method is robust to output noise and neither the Gaussian assumption of the input signal nor the invertibility of the static nonlinearity is necessary. The main assumption used in this paper is that static nonlinearity is monotonically non-decreasing. In the proposed identification method, the linear dynamical system is parametrized as a Finite Impulse Response (FIR) model and a nonparametric identification method is used to create the noise free output signal. Because both the intermediate signal and the noise free output signal are unknown, an over-parametrization technique is used. Once parameters are estimated, a Singular Value Decomposition (SVD) is used to separate the linear system parameters and the noise free output signal. The proposed identification method is applied to simulation data from a Wiener system. The effectiveness and accuracy of the proposed method are verified via numerical simulations.

Keywords: Nonlinear system identification, block-oriented nonlinear model, semidefinite programming.

1. INTRODUCTION

A Wiener system has a block oriented structure where a linear dynamical system and a static output nonlinearity are separated, as shown in Figure 1. The identification of Wiener systems involves estimating the parameters describing the linear dynamical and the static nonlinear blocks from the measured input and output data. A comprehensive overview of block-oriented nonlinear system identification, including Wiener systems, can be found in Giri and Bai [2010]. The most common assumptions used in Wiener system identification are the Gaussian assumption of the input signal and the invertibility of the static nonlinearity. These assumptions are popular because, if the input signal is Gaussian noise, the identification of the linear dynamical block can be separated from the identification of the static nonlinear function based on separability assumption [Greblicki [1992] Enqvist and Ljung [2005]] and parameterization of the output static nonlinearity is possible for the inverse of the given static nonlinearity. However, the Gaussian input assumption is too restrictive for practical application and the invertibility of the static nonlinearity assumption excludes hard nonlinearities, such as saturation, common in control systems.

Recently, a system identification method was introduced based on the sector bound property of static nonlinearity using Quadratic Programming (QP) in Zhang et al. [2006]. This monotonicity assumption on the unknown static nonlinearity guarantees a solution for an Finite Impulse Response (FIR) linear system and leads to possible nonparametric identification of static nonlinear function [Reyland [2011]].

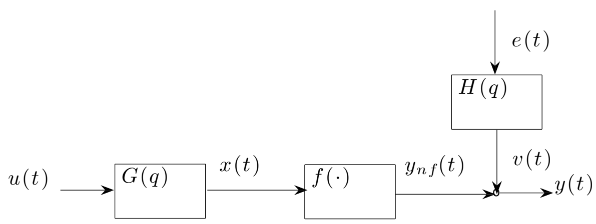

Fig. 1. Wiener system with output noise.

In this paper, the monotonicity assumption on the unknown static nonlinearity is utilized for nonparametric identification of static nonlinear function. The main contribution of this paper is that the proposed method is robust to output noise and neither the Gaussian assumption of the input signal nor the invertibility of the static nonlinearity is necessary. Regarding Figure 1, the objective of this paper is to formulate a procedure that allows for the characterization and identification of the nonlinear static function $f(\cdot)$ and the linear dynamical 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 $x(t)$ and the noise free output signal $y_{nf}(t)$ with conditions on the finite dimensional dynamical representation of the linear systems $G(q)$ and the memoryless static nonlinearity $f(\cdot)$. The output disturbance $v(t)$ is filtered zero-mean white noise independent of the input signal, where the filtering properties are unknown.

2. SYSTEM DESCRIPTION

The system to be modeled is a Wiener system as shown in Figure 1. The purpose of this study is to propose a method to identify the unknown linear dynamical systems $G(q)$ and a static nonlinear function $f(\cdot)$ from a finite number of observations of the data $u(t)$ and $y(t)$. The order of a finite dimensional linear dynamical 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 et al. [2004]]. Based on this idea, in this paper, the rank minimization problem is used to formulate a convex optimization problem via Semidefinite Programming (SDP) relaxation. The system parameters will be estimated by finding a feasible model consistent with the input and output data, and satisfying the following basic properties of the Wiener system:

Condition 1.

I. The static nonlinear function has no memory: The current output $y_{nf}(t)$ only depends on the current input $x(t)$.

II. The linear dynamical system has a finite, but unknown, McMillan degree n:
$x(t) = \phi^T(t)\theta$, where
$\phi^T(t) = [u(t) \ \cdots \ u(t-n_b) \ x(t-1) \ \cdots \ x(t-n_a)]$,
θ is the linear system parameter,
and $n \leq max(n_b - 1, n_a)$.

The intermediate signal $x(t)$ and the noise free output signal $y_{nf}(t)$ in Figure 1 are not measurable. The unknown signals will be parametrized and the estimation of the unknown coefficients will be formulated as a SDP problem. Let $\hat{x}(t)$ be the reconstructed signal of $x(t)$ and $\hat{y}_{nf}(t)$ be the reconstructed signal of $y_{nf}(t)$. The SDP problem will be formulated in such a way that $\hat{x}(t)$ and $\hat{y}_{nf}(t)$ are related via a memoryless static nonlinearity, $u(t)$ and $\hat{x}(t)$ are related via a linear dynamical system with the smallest McMillan degree, and $||y - \hat{y}_{nf}||_2$ is minimized under Condition 1. Once $\hat{x}(t)$ and $\hat{y}_{nf}(t)$ have been reconstructed, the identification of $G(q)$ from $u(t)$ to $\hat{x}(t)$ can be solved with a standard Prediction Error (PE) identification method in Ljung [1999] and the identification of $f(\cdot)$ from $\hat{x}(t)$ to $\hat{y}_{nf}(t)$ can be solved via the Least Squares (LS) method. The proposed identification method deals with Wiener systems in this paper, but the idea of constraining rank for signal reconstruction can be extended to Hammerstein, Wiener-Hammerstein, or Hammerstein-Wiener systems.

3. SYSTEM PARAMETRIZATION

3.1 The input-output map of the linear dynamical system

Let $h(k), \ k = 0, 1, \cdots$ be a causal sequence of unit impulse responses of $G(q)$. The relationship between the input $u(t)$ and the intermediate signal $x(t)$ can be described by the convolution as

$$x(t) = \sum_{k=0}^{\infty} h(k)u(t-k). \tag{1}$$

Due to Condition 1 (finite McMillan degree), the Hankel matrix defined as

$$H = \begin{bmatrix} h(1) & \cdots & h(N/2) \\ h(2) & \cdots & h(N/2+1) \\ \vdots & \ddots & \vdots \\ h(N/2) & \cdots\cdots & h(N-1) \end{bmatrix}, \tag{2}$$

has a $rank(H) \leq n$. 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 et al. [2005]]. A lower order model, consistent with the input and output signals can be estimated by minimizing the rank of H. Let

$$\hat{x} = [\hat{x}(1) \ \hat{x}(2) \ \cdots \ \hat{x}(N)]^T$$

and

$$U = \begin{bmatrix} u(1) & u(0) & \cdots & u(2-N) \\ u(2) & u(1) & \cdots & u(1-N) \\ \vdots & \vdots & \ddots & \vdots \\ u(N) & u(N-1) & \cdots & u(1) \end{bmatrix}. \tag{3}$$

With

$$h = [h(0) \ h(1) \ \cdots \ h(N-1)]^T, \tag{4}$$

the finite sequence of the input

$$u = [u(1) \ u(2) \ \cdots \ u(N)]^T$$

and the estimate of the intermediate signal x can be written as

$$\hat{x} = Uh. \tag{5}$$

3.2 Characteristics of static nonlinearity

Let $\hat{y}_{nf}$ be the noise free output defined as

$$\hat{y}_{nf} = [\hat{y}_{nf}(1) \ \cdots \ \hat{y}_{nf}(N)]^T. \tag{6}$$

Due to the memoryless relationship between $\hat{x}(t)$ and $\hat{y}_{nf}(t)$, the cross-covariance function between $\hat{x}(t)$ and $\hat{y}_{nf}(t)$,

$$R_{yx}(\tau) = \frac{1}{N}\sum_{t=1}^{N} \hat{y}_{nf}(t)\hat{x}(t-\tau),$$

must only depend on the auto covariance of $\hat{x}(t)$,

$$R_x(\tau) = \frac{1}{N}\sum_{t=1}^{N} \hat{x}(t)\hat{x}(t-\tau),$$

and the static nonlinearity [Nuttall [1958]]. There could be many combinations of $\hat{x}(t)$ and $\hat{y}_{nf}(t)$ that satisfy this memoryless relationship. In this paper, a monotonically non-decreasing static nonlinearity with the 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:
$(\hat{y}_{nf}(i) - \hat{y}_{nf}(j))(\hat{y}_{nf}(i) - \hat{y}_{nf}(j) - \hat{x}(i) + \hat{x}(j)) \leq 0$
$\forall i > j$.

In Condition 2,

$\hat{y}_{nf}(i) - \hat{y}_{nf}(j) \geq 0 \Rightarrow \hat{y}_{nf}(i) - \hat{y}_{nf}(j) \leq \hat{x}(i) - \hat{x}(j)$
or
$\hat{y}_{nf}(i) - \hat{y}_{nf}(j) \leq 0 \Rightarrow \hat{y}_{nf}(i) - \hat{y}_{nf}(j) \geq \hat{x}(i) - \hat{x}(j)$.
In both cases,
$\hat{x}(i) - \hat{x}(j) = 0 \Rightarrow \hat{y}_{nf}(i) - \hat{y}_{nf}(j) = 0$
or
$\hat{x}(i) - \hat{x}(j) \neq 0 \Rightarrow \dfrac{\hat{y}_{nf}(i) - \hat{y}_{nf}(j)}{\hat{x}(i) - \hat{x}(j)} \leq 1.$

Condition 2 implies that once $\hat{x}(t)$ is chosen, $\hat{y}_{nf}(t)$ is determined as

$$\hat{y}_{nf} = \alpha(t)\hat{x}(t),\ 0 \leq \alpha(t) \leq 1.$$

This implies that the cross-covariance function between $\hat{x}(t)$ and $\hat{y}_{nf}(t)$ only depends on the static nonlinearity, characterized by $\alpha(t)$, and the auto-covariance of $\hat{x}(t)$, not τ as

$$\begin{aligned} R_{yx}(\tau) &= \frac{1}{N}\sum_{t=1}^{N} \hat{y}_{nf}(t)\hat{x}(t-\tau) \\ &= \frac{1}{N}\sum_{t=1}^{N} \alpha(t)\hat{x}(t)\hat{x}(t-\tau). \end{aligned}$$

Thus, Condition 2 guarantees that the intermediate signal $\hat{x}(t)$ and the output $\hat{y}_{nf}(t)$ are related by a static nonlinear function.

4. PROBLEM FORMULATION

In this section, a rank minimization problem with the memoryless constraint on the static nonlinearity for the reconstruction of the intermediate signal $x(t)$ and the noise free output signal $y_{nf}(t)$ in Figure 1 is summarized and the optimization problem is constructed. With the parametrization and constraints explained in the previous section, an optimization problem can be written as follows:

Optimization problem 1.

Consider
  variables h in (4) and $\hat{y}_{nf}$ in (6)
Define
  $\hat{x} = Uh$, with U in (3)
Minimize
  $w_1 \cdot ||y - \hat{y}_{nf}||_2 + w_2 \cdot rank\ H$, with H in (2)
subject to
  $(\hat{y}_{nf}(i) - \hat{y}_{nf}(j))(\hat{y}_{nf}(i) - \hat{y}_{nf}(j) - \hat{x}(i) + \hat{x}(j)) \leq 0$
  $\forall i > j$
where
  w_1 and w_2 are weighting factors

Optimization Problem 1 results in the optimal solution for the system parameter h that is used to construct the intermediate signal x and the noise free output y_{nf}. In Optimization Problem 1, the reconstructed signals $\hat{x}$ and $\hat{y}_{nf}$ are generated in such a way that a static nonlinear function satisfies the monotonically non-decreasing condition, the linear dynamical system has the minimum order, and the prediction error is minimized under the chosen weighting and constraints.

Unfortunately, the rank condition and the constraint in Optimization Problem 1 are not convex. In this paper, a new variable Θ is defined in order to convert the non-convex optimization problem to an approximated convex optimization problem, resulting in a Semidefinite Programming (SDP) problem. This SDP problem is easier to solve and the solution is close to the solution of the original non-convex problem [Fazel et al. [2004]].

Let $\theta = [h^T \hat{y}_{nf}^T]^T$. With θ, let us define a positive semidefinite symmetric matrix $\Theta = \theta\theta^T$ as

$$\Theta = \begin{bmatrix} h(0)h(0) & \cdots\cdots & h(0)\hat{y}_{nf}(N) \\ \vdots & \vdots\ \vdots & \vdots \\ h(N)h(0) & \vdots\ \vdots & h(N)\hat{y}_{nf}(N) \\ \hat{y}_{nf}(1)h(0) & \cdots\cdots & \hat{y}_{nf}(1)\hat{y}_{nf}(N) \\ \vdots & \vdots\ \cdots & \vdots \\ \hat{y}_{nf}(N)h(0) & \cdots\cdots & \hat{y}_{nf}(N)\hat{y}_{nf}(N) \end{bmatrix}. \quad (7)$$

Based on its structure, it is clear that Θ is a *rank* 1 matrix if there is no noise in the data. For cases where there is noise, system parameters will be found by minimizing $rank(\Theta)$. Because Θ is a square positive semidefinite matrix, minimizing its trace is the closest approximation of the rank minimization that can be efficiently solved. Without loss of generalization, the maximum slope 1 of the static nonlinearity combined with the minimization of $trace(\Theta)$ serves as a normalization condition on the static nonlinearity, so that the static gain of the Wiener system is modeled by the static gain of the linear system $G(q)$. Due to the over-parametrization of Θ, it is impossible to access $\hat{y}_{nf}$ directly through Θ. However, Θ contains information of $\hat{y}_{nf}\hat{y}_{nf}^T$. Thus, minimizing $||y - \hat{y}_{nf}||_2$ is relaxed to minimizing $||yy^T - \hat{y}_{nf}\hat{y}_{nf}^T||_F$, where $||\cdot||_F$ is a Frobenius norm. With Θ, let us express the quadratic constraints in Optimization Problem 1 as Linear Matrix Inequalities (LMIs). Let

$$\delta Y = \begin{bmatrix} \hat{y}_{nf}(2) - \hat{y}_{nf}(1) & \cdots & 0 \\ 0 & \cdots & 0 \\ 0 & \vdots & 0 \\ 0 & \cdots & \hat{y}_{nf}(N) - \hat{y}_{nf}(1) \end{bmatrix}$$

be a diagonal matrix whose diagonal entries include $\hat{y}_{nf}(i) - \hat{y}_{nf}(j),\ \forall i > j$ and

$$\delta X = \begin{bmatrix} \hat{x}(2) - \hat{x}(1) & 0 & \cdots & 0 \\ 0 & \hat{x}(3) - \hat{x}(2) & \cdots & 0 \\ 0 & 0 & \vdots & 0 \\ 0 & 0 & \cdots & \hat{x}(N) - \hat{x}(1) \end{bmatrix}$$

be a diagonal matrix whose diagonal entries include $\hat{x}(i) - \hat{x}(j),\ \forall i > j$. Then $diag(\delta Y) = \Delta Y y$, where

$$\Delta Y = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0\cdots & 0 \\ 0 & 0 & 1 & 0\cdots & 0 \\ \vdots & \cdots & \cdots & \cdots & \vdots \\ \vdots & 0 & 0\cdots & \cdots 0 & 1 \\ \vdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \cdots 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & \vdots \\ 1 & 0 & 0 & \cdots & \vdots \\ 0 & \cdots & \cdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 & \vdots \\ 0 & \vdots & \cdots & \cdots & \vdots \\ 1 & 0 & \cdots & \cdots & 0 \end{bmatrix}$$

and $diag(\delta X) = \Delta X h$, where

$$\Delta X = \begin{bmatrix} u_2 & u_1 & 0 & \cdots & 0 \\ u_3 & u_2 & u_1 & 0\cdots & 0 \\ u_3 & u_2 & u_1 & 0\cdots & 0 \\ \vdots & \cdots\cdots & & \cdots & \vdots \\ u_N & \cdots\cdots & & \cdots & u_1 \end{bmatrix} - \begin{bmatrix} u_1 & 0 & \cdots & \cdots & 0 \\ u_2 & u_1 & 0 & \cdots & 0 \\ u_1 & 0 & \cdots & \cdots & 0 \\ u_3 & u_2 & u_1 & 0\cdots & 0 \\ u_2 & u_1 & 0\cdots & \cdots & 0 \\ u_1 & 0 & \cdots & \cdots & 0 \\ \vdots & \cdots & \cdots & \cdots & \vdots \\ u_1 & 0 & \cdots & \cdots & 0 \end{bmatrix}$$

where u_i is used instead of $u(i)$ for notational brevity. Then, the constraints in Optimization Problem 1 can be written as

$$\delta Y^T \delta Y - \delta Y^T \delta X \leq 0. \tag{8}$$

where

$$\delta Y^T \delta Y = diag(diag(\Delta Y \tilde{\tilde{\Theta}} \Delta Y^T))$$
$$\text{where } \tilde{\tilde{\Theta}} = \Theta(N+1:2N, N+1:2N)$$

and

$$\delta Y^T \delta X = diag(diag(\Delta X \tilde{\Theta} \Delta Y^T))$$
$$\text{where } \tilde{\Theta} = \Theta(N+1:2N, 1:N)$$

where the notation $(k,:)$ and $(:,k)$ are used to denote the k^{th} row and the k^{th} column in a matrix respectively. Here $diag(x)$ indicates a square matrix with the elements of a vector x on the diagonal, and $diag(X)$ indicates the main diagonal of a matrix X.

Using SDP relaxation, Optimization Problem 1 can be rewritten as the following convex optimization problem:

Optimization problem 2.

Consider
variable symmetric Θ
Minimize
$w_1 \cdot ||\tilde{\tilde{\Theta}} - yy^T||_F + w_2 \cdot trace(\Theta)$
subject to
$\delta Y^T \delta Y - \delta Y^T \delta X \leq 0$
$\Theta \geq 0$
where
$\delta Y^T \delta Y = diag(diag(\Delta Y \tilde{\tilde{\Theta}} \Delta Y^T))$
$\delta Y^T \delta X = diag(diag(\Delta X \tilde{\Theta} \Delta Y^T))$
where
$\tilde{\tilde{\Theta}} = \Theta(N+1:2N, N+1:2N)$
$\tilde{\Theta} = \Theta(N+1:2N, 1:N)$
$||\cdot||_F$ is a Frobenius norm
where
w_1 and w_2 are weighting factors.

As a summary, the intermediate signal $x(t)$ in Figure 1 is parametrized by (5), and estimation of the unknown coefficients h in (4) and the noise free signal $\hat{y}_{nf}$ in (6) are solved by computing the solution to the SDP problem given in Optimization Problem 2. The optimization guarantees that $\hat{x}(t)$ and $\hat{y}_{nf}(t)$ are related via a memoryless static nonlinearity, and guarantees that $u(t)$ and $\hat{x}(t)$ are related via a linear dynamical system with the smallest McMillan degree. Once $\hat{x}(t)$ and $\hat{y}_{nf}(t)$ have been reconstructed, the identification of $G(q)$ from $u(t)$ to $\hat{x}(t)$ can be solved with a standard Prediction Error (PE) identification method in Ljung [1999].

4.1 Parameter separation

Due to the over-parametrization used to define Θ in (7), we need to separate the parameters of the linear dynamical system h and the noise free output $\hat{y}_{nf}$. Singular Value Decomposition (SVD) is used in this paper to separate the system parameters. The SVD of Θ is given as

$$\Theta = U \Sigma V^T \tag{9}$$

where $U_{2N\times 2N}$ and $V_{2N\times 2N}$ are orthogonal matrices, $U_{2N\times 2N} = V_{2N\times 2N}$ due to the structure of Θ, and $\Sigma_{2N\times 2N}$ is a rectangular diagonal matrix. The positive diagonal entries of Σ are called singular values. From (9), the parameter vector $\theta = [h^T \hat{y}_{nf}^T]^T$, where $\Theta = \theta\theta^T$ can be calculated by

$$\begin{aligned} \theta &= \sqrt{\sigma_1} U(:,1) \\ h &= \theta(1:N) \\ \hat{y}_{nf} &= \theta(N+1:2N) \end{aligned} \tag{10}$$

providing an optimal *rank* 1 approximation of Θ.

5. NUMERICAL EXAMPLE

In this section, a numerical example of Wiener system identification using the proposed identification method is presented. A Pseudo Random Binary Sequence (PRBS) excitation signal, defined as

$$u(t) = 4 \cdot sign(randn(N,1))$$

where

$$u(t) = \begin{cases} +1 & w.p. \ \frac{1}{2} \\ -1 & w.p. \ \frac{1}{2} \end{cases},$$

is used as the input. The output disturbance $v(t) = H(q)e(t)$ is filtered zero-mean white noise independent of the input signal, where the filtering properties, $H(q)$, are not estimated or not need to be known. For the system identification, twenty sets of estimation data with 100 samples are generated from the Wiener system with the following specifications:

$$\begin{cases} \text{Linear dynamical system:} \\ G(q) = \dfrac{0.0997q^{-1} - 0.0902q^{-2}}{1 - 1.886q^{-1} + 0.9048q^{-2}} \\ \text{Static nonlinearity:} \\ \quad f(x(t)) = \begin{cases} .5 & \text{if} \quad x(t) > .5 \\ x(t) & \text{if} \quad |x(t)| \leq .5 \\ -.5 & \text{if} \quad x(t) < -.5 \end{cases} \\ \text{Noise dynamics:} \\ H(q) = \dfrac{1 + 0.5q^{-1}}{1 - 0.85q^{-1}} \end{cases}.$$

The input and output signals are shown in Figure 2. In order to solve the Semidefinite Programming (SDP) problem (Optimization Problem 2), SEDUMI [Sturm [1999]] and YALMIP [Löfberg [2004]] are used. The estimation results are shown in Figure 3, Figure 4 and Figure 5. As shown in Figure 3 and Figure 4, both pole and zero locations are

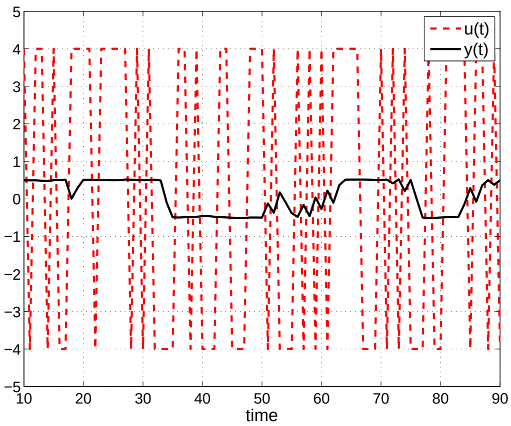

Fig. 2. The input and output signals.

well estimated. As shown in Figure 5, ±.5 saturation is well identified.

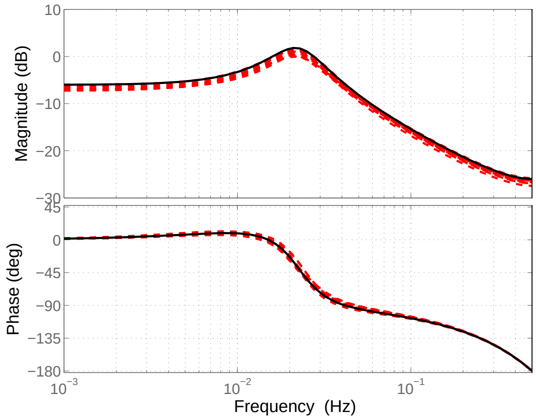

Fig. 3. The Bode plot of the identified linear dynamical system. The black solid line indicates the real linear dynamical system. The (colored) dashed lines indicate estimated linear dynamical systems by using twenty different sets of data. The SNR of each data set is greater than 50dB.

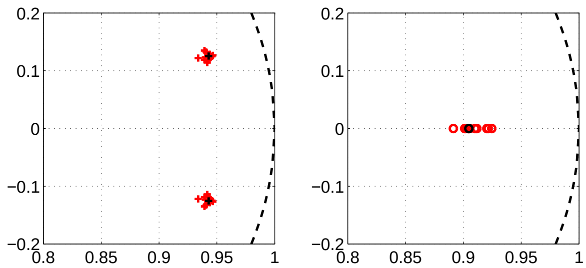

Fig. 4. Pole (left figure) and zero (right figure) locations of the identified linear dynamical system. The black cross and circle indicate the real linear dynamical system. The colored crosses and circles indicate estimated linear dynamical systems. The SNR of each data set is greater than 50dB.

6. CONCLUSION

In this paper, the Wiener system identification problem is formulated as a Semidefinite Programming (SDP) problem to reconstruct the intermediate signal and noise free output. The system parameter identification problem is formulated as a rank minimization problem by imposing the monotonically non-decreasing condition on the static nonlinear function. This non-convex optimization problem is then reformulated as a convex optimization problem via SDP relaxation by using over-parametrization. The proposed method is robust to output noise and neither the Gaussian assumption of the input signal nor the invertibility of the static nonlinearity is necessary. Singular Value Decomposition (SVD) is used to separate the linear system parameters and the noise free output signal. Once the intermediate signal and noise free output signal are reconstructed, the identification of the linear dynamical system and the static nonlinear function become trivial. The proposed identification method is applied to simulation data from a Wiener system. The numerical simulation

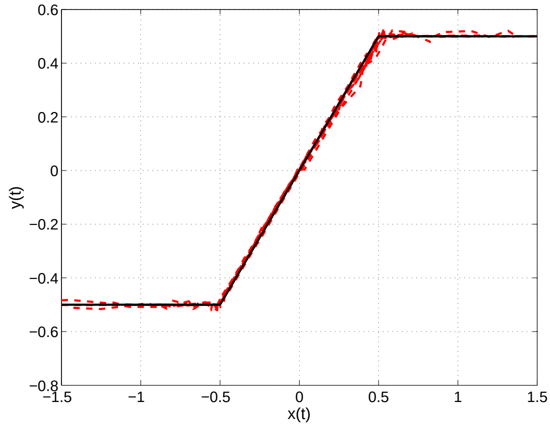

Fig. 5. The plot of the identified static nonlinear function. The black solid line indicates the real static nonlinear function. The (colored) dashed lines indicate estimated static nonlinear functions by using twenty different sets of data. The SNR of each data set is greater than 50dB.

result shows the effectiveness of the proposed identification method.

REFERENCES

M. Enqvist and L. Ljung. Linear approximations of nonlinear FIR systems for separable input process. *Automatica*, 41:459–473, 2005.

M. Fazel, H. Hindi, and S. Boyd. Rank minimization and applications in system theory. In *Proc. American Control Conference*, pages 3273–3278, Boston, Massachusetts, 2004.

F. Giri and E.W. Bai. *Block Oriented Nonlinear System Identification.* Springer-Verlag, Berlin, 2010.

I. Goethals, K. Pelckmans, J.A.K. Suykens, and B. De Moor. Subspace identification of hammerstein systems using least squares support vector machines. *IEEE Transactions on Automatic Control, Special Issue on System Identification*, 50:1509–1519, 2005.

W. Greblicki. Nonparametric identification of Wiener systems. *IEEE Transactions on information theory*, 38: 1487–1493, 1992.

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

J. Löfberg. YALMIP: A toolbox for modeling and optimization in MATLAB. In *IEEE International Symposium on CACSD Conference*, pages 284–289, Taipei, Taiwan, 2004.

A. H. Nuttall. *Theory and application of the separable class of random processes.* Ph.D. Thesis, MIT., 1958.

J. M. Reyland. *Towards Wiener system identification with minimum a priori information.* Ph.D. Thesis, University of Iowa., 2011.

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

Q. Zhang, A. Iouditski, and L Ljung. Identification of Wiener system with monotonous nonlinearity. In *14th IFAC Symposium on System Identification*, pages 166–171, Newcastle, Australia, 2006.