

# State-Space Model and Kalman Filter Gain Identification by a Superspace Method

Ping Lin, Minh Q. Phan, and Stephen A. Ketcham

**Abstract** This paper describes a superspace method to identify a state-space model and an associated Kalman filter gain from input-output data. Superstate vectors are simply vectors containing input-output measurements, and used directly for the identification. The superstate space is unusual in that the state portion of the Kalman filter becomes completely independent of both the system dynamics and the input and output noise statistics. The system dynamics is entirely carried by the measurement portion of the superstate Kalman filter model. When model reduction is applied, the system dynamics returns to the state portion of the state-space model.

## 1 Introduction

Finding a state-space model of a system and an associated optimal observer/Kalman filter gain in the presence of unknown process and measurement noises is a well-known and important system identification problem. The identification problem is nonlinear because it involves the product of the system matrix and the state, both of which are unknown. A method known as Observer/Kalman filter identification (OKID) bypassed the need to determine the system states by working through an intermediate set of parameters called Observer/Kalman filter Markov parameters [1]. These parameters are related to input-output data linearly, and can thus be found by a simple linear least-squares solution. The nonlinear step is handled by the Eigensystem Realization Algorithm (ERA), which offers an analytical (exact) solution to the problem of recovering a state-space model representation from the identified Markov parameters [2]. Under appropriate conditions, OKID identifies a state-space model of the system and an associated Kalman filter gain that is optimal with respect to the unknown process and measurement noises embedded in the input-output data. There are numerous extensions of the observer-based method

---

P. Lin • M.Q. Phan (✉)

Thayer School of Engineering, Dartmouth College, Hanover, NH 03755, USA

e-mail: [ping.lin@dartmouth.edu](mailto:ping.lin@dartmouth.edu); [minh.q.phan@dartmouth.edu](mailto:minh.q.phan@dartmouth.edu)

S.A. Ketcham

Cold Regions Research and Engineering Laboratory (CRREL), Hanover, NH 03755, USA

e-mail: [stephen.a.ketcham@erdc.usace.army.mil](mailto:stephen.a.ketcham@erdc.usace.army.mil)

[3–5], such as residual whitening [6], and methods that are derived from interaction matrices [7].

Another important class of methods that solved this problem is known as “subspace methods” [8, 9]. These methods put the emphasis on the recovery of the system states from input-output measurements by various oblique projection techniques. A fundamental attraction of this approach is that once the states are known, the identification of the state-space model becomes linear. These methods are also capable of recovering the optimal Kalman filter gain when the input-output measurements are corrupted by unknown process and measurement noises. There are numerous variations of the subspace technique [10–12]. The method used in this paper for comparison purpose is N4SID [13].

This paper describes a recently developed class of methods which can be referred to as “superspace methods”. A superstate vector is made up of input and output measurements. These superstate vectors are treated as the states of the system, and used directly in the identification of the state-space model and an associated Kalman filter gain. The superspace method bypasses the need to recover the states of the system as required in a subspace method. It also sidesteps the need to work through the Markov parameters as in OKID-based methods. In this paper, we further show that in the space of the superstates, the system matrices that define the state portion of the Kalman filter are made up entirely of 1’s and 0’s. They do not need to be identified, and even more interestingly, they are independent of the system dynamics and the process and measurement noise statistics. All system dynamics are carried in the output measurement portion of the Kalman filter. When model reduction is applied, the dynamics of the system returns to the state portion of the model as one would expect in a state-space model. The superspace idea has also been recently applied successfully to the bilinear system identification problem [14].

## 2 Mathematical Formulation

The state-space identification problem has many related or equivalent forms. We will use one that is both common in literature and convenient for stating our algorithm.

### 2.1 Problem Statement

Suppose a set of input-output measurements,  $\{u_0, u_1, \dots, u_s\}$ ,  $\{y_0, y_1, \dots, y_s\}$ , of an  $m$ -input,  $q$ -output system corrupted by unknown process and measurement noises, is given. The objective of the problem is to find a state-space representation  $\{A, B, C, D\}$  and a steady-state Kalman filter gain  $K$  such that the input-output measurements are related to each other by the following innovation form of the state-space model,

$$\hat{x}_{k+1} = A\hat{x}_k + Bu_k + Ke_k \quad (1)$$

$$y_k = C\hat{x}_k + Du_k + e_k \quad (2)$$

where  $\hat{x}_k$  denotes the (unknown) Kalman filter state,  $K$  is the (unknown) corresponding steady-state Kalman filter gain, and  $e_k$  is the (unknown) Kalman filter residual that is unique for the given system, input-output measurements, and noises in the system,  $e_k = y_k - \hat{y}_k = y_k - (C\hat{x}_k + Du_k)$ . The innovation form is derived from the conventional process form,

$$\hat{x}_{k+1} = A\hat{x}_k + Bu_k + w_k \quad (3)$$

$$y_k = C\hat{x}_k + Du_k + n_k \quad (4)$$

The process noise  $w_k$  and measurement noises  $n_k$  are assumed to be independent, white, zero-mean, and Gaussian. The Kalman filter gain  $K$  is a function of the system state-space model and the covariances of the process and measurement noises. In the identification problem, only input-output measurements are known. The noise covariances are unknown.

## 2.2 A Superstate Vector Definition

A superstate vector  $z_k$  is defined from input-output data as follows,

$$z_k = \begin{bmatrix} v_{k-p} \\ \vdots \\ v_{k-2} \\ v_{k-1} \end{bmatrix} \quad v_k = \begin{bmatrix} u_k \\ y_k \end{bmatrix} \quad (5)$$

From the given input-output measurements, these superstates can be easily created, and used in the subsequent superspace identification method.

## 3 A Superspace Identification Algorithm

A superspace identification algorithm is summarized below.

- Choose a value for  $p$  such that  $(m+q)p$  defines the dimension of the superstate vector, which is typically larger than the true minimum state dimension of the system being identified.
- Form the following matrix  $Z$  from the available input-output data,

$$Z = \begin{bmatrix} v_0 & v_1 & \cdots & v_{s-p} & v_{s-p+1} \\ v_1 & v_2 & \cdots & v_{s-p+1} & v_{s-p+2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ v_{p-2} & v_{p-1} & \cdots & v_{s-2} & v_{s-1} \\ v_{p-1} & v_p & \cdots & v_{s-1} & v_s \end{bmatrix} \quad (6)$$

Define  $Z_0$  as  $Z$  with its last column removed, and  $Z_1$  as  $Z$  with its first column removed. Furthermore, define

$$U_p = [u_p \ u_{p+1} \ \cdots \ u_s] \quad Y_p = [y_p \ y_{p+1} \ \cdots \ y_s] \quad V_p = \begin{bmatrix} U_p \\ Y_p \end{bmatrix} \quad (7)$$

- Solve for  $\bar{A}^*$ ,  $\bar{B}^*$ ,  $C^*$ , and  $D^*$  by least-squares from

$$Z_1 = \bar{A}^* Z_0 + \bar{B}^* V_p \quad (8)$$

$$Y_p = C^* Z_0 + D^* U_p + E_p^* \quad (9)$$

It turns out that there is no need to solve for  $\bar{A}^*$ ,  $\bar{B}^*$  from (8), as they are simply matrices made up of 0's and 1's,

$$\bar{A}^* = \left[ \begin{array}{c|c} 0_{(p-1)b \times b} & I_{(p-1)b \times (p-1)b} \\ \hline 0_{b \times b} & 0_{b \times (p-1)b} \end{array} \right] \quad \bar{B}^* = \left[ \begin{array}{c} 0_{(p-1)b \times b} \\ \hline I_{b \times b} \end{array} \right] \quad (10)$$

where  $b = m + q$ . A key point to observe here is that  $\bar{A}^*$ ,  $\bar{B}^*$  are completely independent of the system dynamics and the process and measurement noise statistics. Information about the system is completely contained in  $C^*$  and  $D^*$ , which can be solved by least-squares,

$$[C^* \mid D^*] = Y_p \left[ \begin{array}{c} Z_0 \\ U_p \end{array} \right]^\dagger \quad (11)$$

The  $\dagger$  denotes the Moore-Penrose pseudo-inverse.  $Z_1$  is not needed in (11), but will be used later in establishing the optimality of the algorithm.

- Lastly, a representation of  $\{A, B, C, D, K\}$  denoted by  $\{A^*, B^*, C^*, D^*, K^*\}$  can be recovered from  $\bar{A}^*$ ,  $\bar{B}^*$ ,  $C^*$ ,  $D^*$  based on the following relationship,

$$\bar{A}^* = A^* - K^* C^* \quad \bar{B}^* = [B^* - K^* D^* \mid K^*] \quad (12)$$

Recall that  $C^*$  and  $D^*$  are obtained from (11),  $\bar{A}^*$  and  $\bar{B}^*$  are given in (10). The matrix  $\bar{B}^*$ , which is defined in (10), has two partitions according to (12): the left partition is  $B^* - K^* D^*$  of dimensions  $pb \times m$ , the right partition is  $K^*$  of dimensions  $pb \times q$ . Because  $\bar{B}^*$  is made up of entirely 0's and 1's

and known beforehand, these partitions are known.  $B^*$  can be recovered from the first partition of  $\bar{B}^*$  because  $K^*$  and  $D^*$  are known. Similarly,  $A^*$  can be recovered from  $\bar{A}^*$  because  $K^*$  and  $C^*$  are known. Having obtained a full set of  $\{A^*, B^*, C^*, D^*, K^*\}$ , standard model reduction techniques can be applied to reduce the dimension of  $\{A^*, B^*, C^*, D^*, K^*\}$  to the correct minimum state dimension of the system being identified.

## 4 Optimality of Superspace Identification

We now establish the optimality of the combination  $\{A^*, B^*, C^*, D^*, K^*\}$  by proving that Markov parameters of the combination  $\{A^*, B^*, C^*, D^*, K^*\}$  match the Markov parameters of the optimal Kalman filter given in (1) and (2). First, we need to eliminate  $e_k$  from the state portion of the Kalman filter by solving for  $e_k$  in (2) and substituting it into (1) to produce

$$\hat{x}_{k+1} = \bar{A}\hat{x}_k + \bar{B}v_k \quad (13)$$

$$y_k = C\hat{x}_k + Du_k + e_k \quad (14)$$

where  $\bar{A} = A - KC$ ,  $\bar{B} = [B - KD \mid K]$ , and  $v_k$  is defined in (5). Define

$$\hat{X}_p = [\hat{x}_p \cdots \hat{x}_s] \quad \hat{X}_{p+1} = [\hat{x}_{p+1} \cdots \hat{x}_{s+1}] \quad E_p = [\hat{e}_p \cdots \hat{e}_s] \quad (15)$$

Equations (13) and (14) can be written for all available time steps,

$$\hat{X}_{p+1} = \bar{A}\hat{X}_p + \bar{B}V_p \quad (16)$$

$$Y_p = C\hat{X}_p + DU_p + E_p \quad (17)$$

where  $V_p = [v_p \ v_{p+1} \ \cdots \ v_s]$ . Next, we express  $\hat{X}_p$  and  $\hat{X}_{p+1}$  in terms of input and output measurements. As long as  $p$  is sufficiently large such that  $\bar{A}^p \approx 0$ , then by repeated substitution, the Kalman filter state can be expressed in terms of input and output measurements, and when packaged together, can be put in the form,

$$[\hat{x}_p \ \hat{x}_{p+1} \ \cdots \ \hat{x}_{s+1}] = [\bar{A}^{p-1}\bar{B} \ \cdots \ \bar{A}\bar{B} \ \bar{B}] \begin{bmatrix} v_0 & v_1 & \cdots & v_{s-p+1} \\ v_1 & v_2 & \cdots & v_{s-p+2} \\ \vdots & \vdots & \vdots & \vdots \\ v_{p-1} & v_p & \cdots & v_s \end{bmatrix} \quad (18)$$

Define  $\mathcal{C}_p = [\bar{A}^{p-1}\bar{B}, \dots, \bar{A}\bar{B}, \bar{B}]$ . It follows from (18) that

$$\hat{X}_p = \mathcal{C}_p Z_0 \quad \hat{X}_{p+1} = \mathcal{C}_p Z_1 \quad (19)$$

where  $Z_1$  and  $Z_2$  are the two partitions of  $Z$  as previously defined in (6). Substituting (19) into (16) produces

$$\mathcal{C}_p Z_1 = \bar{A} \mathcal{C}_p Z_0 + \bar{B} V_p \quad (20)$$

$$Y_p = C \mathcal{C}_p Z_0 + D U_p + E_p \quad (21)$$

This is the relation that the ideal optimal system matrices need to satisfy with the given input-output measurements. Notice that  $E_p$  here is made of the innovation sequence  $\{e_k\}$ . In the superspace algorithm, we impose (8) and (9). To relate (20) and (21) to (8) and (9), premultiplying (8) by  $\mathcal{C}_p$  produces

$$\mathcal{C}_p Z_1 = \mathcal{C}_p \bar{A}^* Z_0 + \mathcal{C}_p \bar{B}^* V_p \quad (22)$$

$$Y_p = C^* Z_0 + D^* U_p + E_p^* \quad (23)$$

In the superspace identification algorithm,  $C^*$  and  $D^*$  are solved from (11) by least-squares. This step ensures that the residual  $E_p^*$  is minimized and orthogonal to the input and output data. These are the conditions that the optimal Kalman filter residual must satisfy, hence  $E_p^* = E_p$ . Furthermore, if the input-output data set is sufficiently rich such that the matrix formed by  $Z_0$  and  $V_p$  is full rank, we also have

$$\bar{A} \mathcal{C}_p = \mathcal{C}_p \bar{A}^* \quad (24)$$

$$\bar{B} = \mathcal{C}_p \bar{B}^* \quad (25)$$

$$C \mathcal{C}_p = C^* \quad (26)$$

$$D = D^* \quad (27)$$

As long as  $p$  is sufficiently large such that  $\bar{A}^p \approx 0$ , our choices of  $\bar{A}^*$  and  $\bar{B}^*$  in (10) indeed satisfy (24) and (25). Furthermore, it can be shown that the Markov parameters of the identified Kalman filter match the Markov parameters of the optimal Kalman filter. For example, the first Markov parameter can be shown to match,

$$C^* \bar{B}^* = (C \mathcal{C}_p) \bar{B}^* = C(\mathcal{C}_p \bar{B}^*) = C \bar{B} \quad (28)$$

Similarly, the second Markov parameter can also be shown to match,

$$C^* \bar{A}^* \bar{B}^* = C(\mathcal{C}_p \bar{A}^*) \bar{B}^* = C(\bar{A} \mathcal{C}_p) \bar{B}^* = C \bar{A}(\mathcal{C}_p \bar{B}^*) = C \bar{A} \bar{B} \quad (29)$$

and so on. The two sets of system matrices have the same Markov parameters. This result establishes the optimality of the identified Kalman filter.

We discuss briefly how the superspace method is different from the subspace method. Although there are several variants of the subspace identification method, the well-known N4SID is used for the present discussion. The N4SID method first estimates the optimal states by an oblique projection, then estimates all the system state-space matrices by least-squares. The superspace method bypasses the state estimation step by forming the superstates, and uses them directly in the identification. The least-squares calculation is used to obtain the coefficients of the measurement equations. The state portion of the Kalman filter are made up of 1's and 0's, and they do not need to be computed. More fundamentally, in deriving the subspace method, the process form of the state-space model is used to establish the input-output relationship, whereas in the superspace method the innovation form which involves the Kalman filter is used. The subspace method is thus process form oriented, whereas the superspace method is innovation form oriented. A consequence of using the innovation form is that the superspace identification method simultaneously recovers the steady-state Kalman filter gain together with the system state-space model, whereas in the subspace method, the Kalman filter gain is typically computed as an additional step.

## 5 Illustrative Examples

This section demonstrates the effectiveness of the superspace identification method using both simulated and experimental data. Comparison of the results to those obtained with the subspace N4SID algorithm will also be provided.

### 5.1 A Simulated System

The following system is driven by random input excitation, and the resultant output is recorded for system identification,

$$A = \begin{bmatrix} -0.35 & -0.5 \\ 1 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad C = [0 \ 0.5] \quad D = 0$$

In order to show the convergent behavior, the input-output data record is deliberately chosen to be large ( $2^{16}$  samples). The input and measurement noise covariances are  $Q = 0.025$ ,  $R = 0.025$  corresponding to the input and measurement noises shown in Figs. 1 and 2. Using  $p = 8$  in the identification, the final model order is reduced to 2. Figure 3 shows the Kalman filter Markov parameters,  $C(A - KC)^k B$  and  $C(A - KC)^k K$ , constructed from the identified state-space model and the identified Kalman filter gain by both methods. These Markov parameters are compared to those of the optimal Kalman filter whose gain is computed from perfect

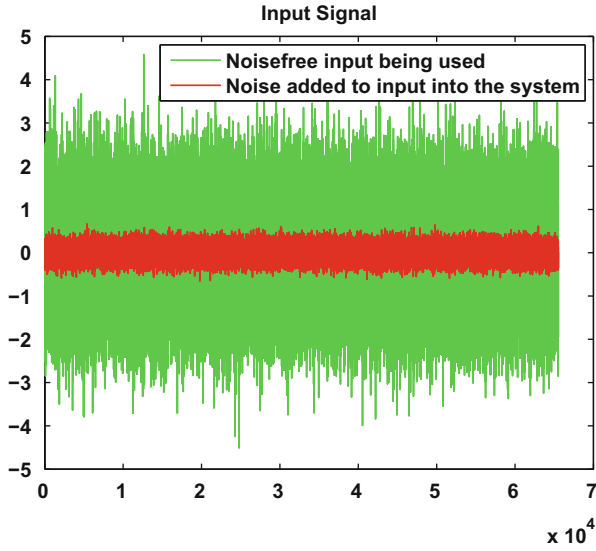


Fig. 1 Input data with added noise

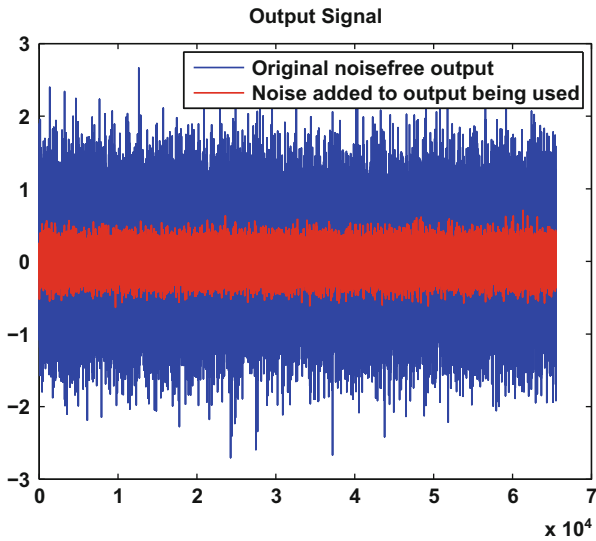
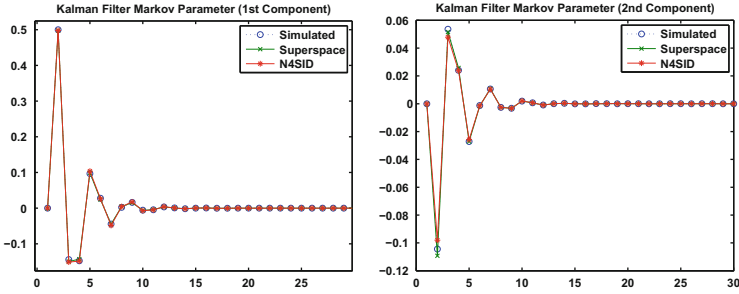


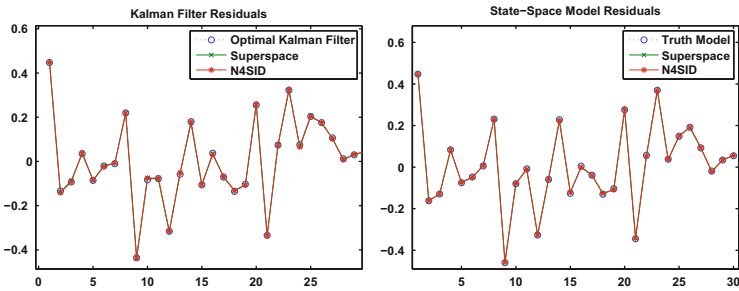
Fig. 2 Output data with added noise

knowledge of the system model and the input and measurement noise covariances. Figure 4 shows the residuals of the identified Kalman filters and the state-state models by both methods (superspace and N4SID) matching the residuals of the





**Fig. 3** Kalman filter Markov parameters by two identification methods (superspace and N4SID) compared to the Kalman filter Markov parameters computed from perfect knowledge of the system model and noise statistics



**Fig. 4** Comparison of identified filter residuals to optimal Kalman residual (*left*), and identified state-space model residuals to truth model residual (*right*)

optimal filter and of the truth model point-wise. The results confirm that both methods produce optimal identification results as expected.

### 5.2 CD Player Arm

A set of experimental data of a CD player arm is used in this example [15]. The system has two inputs and two outputs. The input-output data record used for identification is shown in Figs. 5 and 6 (2,048 samples). Using  $p = 6$  in the identification, the final system order is reduced to 12 before the identification results are compared. Figure 7 shows a comparison of the identified Kalman filter outputs to the measured outputs. Figure 8 shows a comparison of the identified state-space model outputs to the measured outputs. Overall, both methods appear to capture the dynamics of the mechanism relatively well with this set of input-output data.

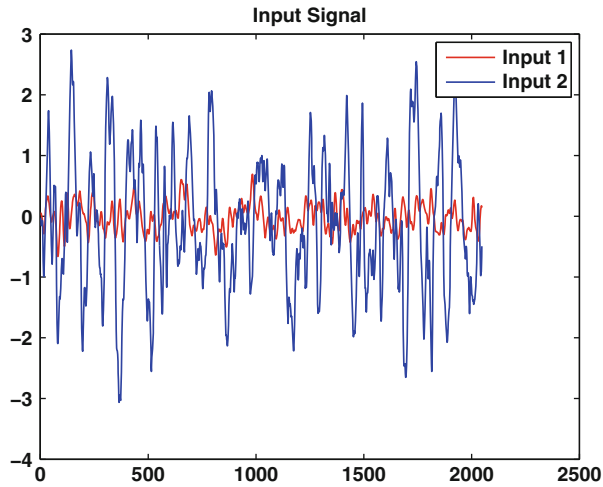


Fig. 5 Input signals of CD player arm

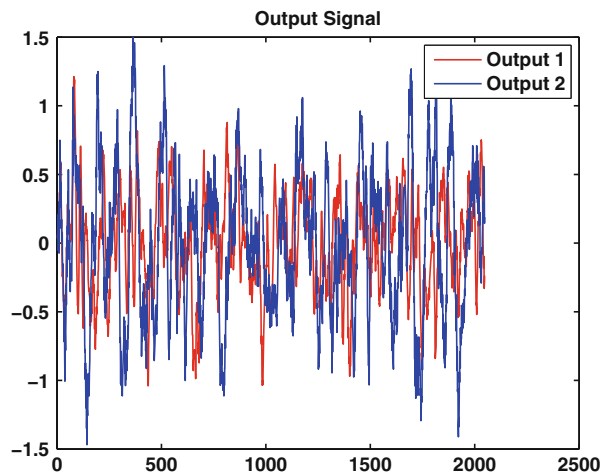
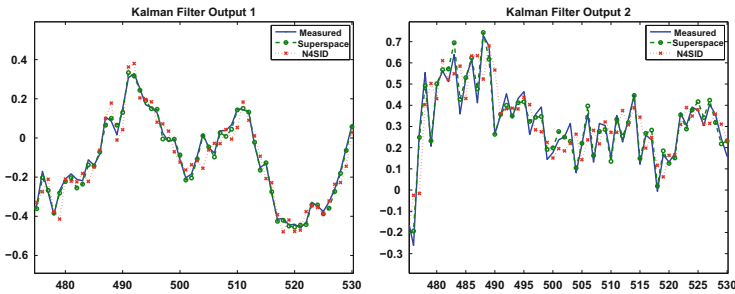
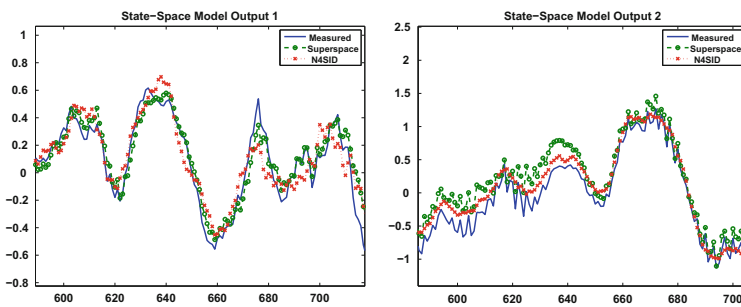


Fig. 6 Output signals of CD player arm



**Fig. 7** Comparison of Kalman filter outputs to measured outputs by two identification methods (superspace and N4SID)



**Fig. 8** Comparison of state-space model outputs to measured outputs by two identification methods (superspace and N4SID)

## 6 Conclusions

A superspace method for identification of a system state-space model and its associated Kalman filter gain has been formulated. It is found that in the space of the superstates, which are vectors of input and output measurements, the matrices that define the state portion of the Kalman filter are made up entirely of 0's and 1's. These matrices are known in advance, and do not need to be identified. Because they are known in advance, they are completely independent of the actual system dynamics and the noise statistics. This is a highly intriguing and very counter-intuitive result. Moreover, in the superstate space, the system dynamics are contained the measurement equation, not the state portion, of the Kalman filter. When model reduction is applied, the actual system dynamics returns to the state portion of the reduced-order model as one would expect in a state-space model. Optimality of the proposed superspace identification method is also established in theory and confirmed in numerical simulation. The Kalman filter identified from input-output measurements by the superspace technique is found to match the optimal Kalman filter derived from perfect knowledge of the system and perfect knowledge of the noise statistics, both in their Markov parameters and their output

residuals. When applied to experimental data of a CD arm mechanism, the method produces excellent results when compared to an established subspace identification method.

**Acknowledgements** This research is supported by a STTR Phase II contract from the US Army Corps of Engineers Cold Regions Research and Engineering Laboratory (CRREL) to Dartmouth College and Sound Innovations, Inc.

## References

1. Juang, J.-N., Phan, M.Q., Horta, L.G., Longman, R.W.: Identification of observer/Kalman filter Markov parameters – theory and experiments. *J. Guid. Control Dyn.* **16**(2), 320–329 (1993)
2. Juang, J.-N., Pappa, R.S.: An eigensystem realization algorithm for modal parameter identification and model reduction. *J. Guid. Control Dyn.* **8**, 620–627 (1985)
3. Phan, M.Q., Juang, J.-N., Longman, R.W.: Identification of linear multivariable systems by identification of observers with assigned real eigenvalues. *J. Astronaut. Sci.* **15**(1), 88–95 (1992)
4. Phan, M.Q., Horta, L.G., Juang, J.-N., Longman, R.W.: Linear system identification via an asymptotically stable observer. *J. Optim. Theory Appl.* **79**(1), 59–86 (1993)
5. Phan, M.Q., Horta, L.G., Juang, J.-N., Longman, R.W.: Improvement of observer/Kalman filter identification (OKID) by residual whitening. *J. Vib. Acoust.* **117**, 232–238 (1995)
6. Juang, J.-N.: *Applied System Identification*. Prentice-Hall, Upper Saddle River (2001)
7. Phan, M.Q.: Interaction matrices in system identification and control. In: *Proceedings of the 15th Yale Workshop on Adaptive and Learning Systems*, New Haven (2011)
8. Van Overchee, P., De Moor, B.: *Subspace Identification for Linear Systems*. Kluwer Academic, Boston (1996)
9. Qin, S.J.: An Overview of Subspace Identification. *Comput. Chem. Eng.* **30**(10–12), 1502–1513 (2006)
10. Verhaegen, M., Dewilde P.: Subspace model identification part 1: the output error state-space model identification class of algorithms. *Int. J. Control* **56**(5), 1187–1210 (1992)
11. Van Overchee, P., De Moor, B.: A unifying theorem for three subspace system identification algorithms. *Automatica* **31**(12), 1853–1864 (1995)
12. Jansson, M., Wahlberg, B.: A linear regression approach to state-space subspace system identification. *Signal Process.* **52**, 103–129 (1996)
13. Van Overchee, P., De Moor, B.: N4SID: subspace algorithms for the identification of combined deterministic-stochastic systems. *Automatica* **30**(1), 75–93 (1994)
14. Phan, M.Q., Čelik, H.: A superspace method for discrete-time bilinear model identification by interaction matrices. *J. Astronaut. Sci.* **59**(1–2), 433–452 (2012)
15. De Moor, B., De Gerssem, P., De Schutter, B., Favoreel, W.: DAISY: a database for identification of systems. *Journal A* **38**(3), 4–5 (1997)