Observers for Bilinear State-Space Models by Interaction Matrices

Minh Q. Phan, Francesco Vicario, Richard W. Longman, and Raimondo Betti

***Abstract*— This paper formulates a bilinear observer for a bilinear state-space model. Relationship between the bilinear observer gains and the interaction matrices are established and used in the design of such observer gains from input-output data. In the absence of noise, the question of whether a deadbeat bilinear observer exists that would cause the state estimation error to converge to zero identically in a finite number of time steps is addressed. In the presence of noise, an optimal bilinear observer that minimizes the state estimation error in the same manner that a Kalman filter does for a linear system is presented. Numerical results illustrate both the theoretical and computational aspects of the proposed algorithms.**

I. INTRODUCTION

Bilinear models can be viewed as a bridge between linear and nonlinear models, providing a promising approach to handle various nonlinear identification and control problems. Bilinear models of sufficiently high dimensions can be used to approximate more general nonlinear systems, Ref. [1]. Regardless of the technique used to bilinearize a nonlinear system (e.g. Carleman linearization, Ref. [2], or bilinear system identification, Ref. [3]), the resulting bilinear models have states whose direct measurements are not possible. A state estimator or observer is a crucial element in the implementation of state-feedback controllers designed for such models. In this paper we derive an observer for a bilinear system together with a design technique to optimize its gains based on system identification. Although the general problem of nonlinear state estimation has been previously studied, the connection between the bilinear observer considered here and the interaction matrices in the context of system identification is a new development. The interaction matrices were originally developed for state-space identification of linear models by the Observer/Kalman filter identification algorithm (OKID) Refs. [4], [5], [6]. OKID identifies a state-space model and an associated optimal observer/Kalman filter gain from input-output data without knowledge of the process and measurement noise covariances. Recently, the interaction matrices have been extended to the state-space identification of discrete bilinear systems, Refs. [7], [8], by proving their existence and exploiting them to develop Input-Output-to-State Relationships (IOSRs). In the present work, the connection between the problems of bilinear system identification and state estimation is established and explored as an approach to design both deterministic and stochastic bilinear observers. It is assumed in the current paper that the bilinear state-space model and the process and measurement noise covariances are known, and only the optimal bilinear observer gains are to be designed. The presented results form the first fundamental step of working out a bilinear version of OKID where a bilinear system model and an associated optimal bilinear observer are identified directly from input-output measurements of a nonlinear dynamical system.

M.Q. Phan is with Thayer School of Engineering, Dartmouth College, Hanover, NH 03755, USA. mqphan@dartmouth.edu

F. Vicario is with Department of Mechanical Engineering, Columbia University, New York, NY 10027, USA.

R.W. Longman is with Department of Mechanical Engineering and Department of Civil Engineering and Engineering Mechanics, Columbia University, New York, NY 10027, USA.

R. Betti is with Department of Civil Engineering and Engineering Mechanics, Columbia University, New York, NY 10027, USA.

II. A BILINEAR OBSERVER

Consider the following n-state single-input single-output discrete bilinear system (the extension to the multiple-input multiple-output case can be made without fundamental difficulties),

$$x(k+1) = Ax(k) + Nx(k)u(k) + Bu(k) + w_p(k) \tag{1a}$$
$$y(k) = Cx(k) + Du(k) + w_m(k) \tag{1b}$$

where $w_p(k)$ and $w_m(k)$ are zero-mean random white process and measurement noises with covariances Q and R, respectively. We introduce a bilinear observer of the form,

$$\begin{aligned}\hat{x}(k+1) &= A\hat{x}(k) + Bu(k) - M_1(y(k) - \hat{y}(k)) \\ &\quad + N\hat{x}(k)u(k) - M_2(y(k) - \hat{y}(k))u(k) \\ &= \bar{A}\hat{x}(k) + \bar{N}\hat{x}(k)u(k) + \bar{B}v(k)\end{aligned} \tag{2}$$

where $\hat{y}(k) = C\hat{x}(k) + Du(k)$ is the estimated output based on the estimated state $\hat{x}(k)$. Define

$$\bar{A} = A + M_1C \qquad \bar{N} = N + M_2C \tag{3a}$$
$$\bar{B} = \begin{bmatrix} B + M_1D & -M_1 & M_2D & -M_2 \end{bmatrix} \tag{3b}$$
$$v(k) = \begin{bmatrix} u(k) & y(k) & u^2(k) & u(k)y(k) \end{bmatrix}^T \tag{3c}$$

It can be shown that the dynamics of the state estimation error $e(k) = x(k) - \hat{x}(k)$ is governed by

$$\begin{aligned}e(k+1) &= \bar{A}e(k) + \bar{N}e(k)u(k) + w_p(k) \\ &\quad + M_1w_m(k) + M_2w_m(k)u(k)\end{aligned} \tag{4}$$

The presence of input-dependent terms in the state estimation error equation makes the bilinear observer problem more challenging than the well-known linear case.

This paper concerns with the bilinear observer design problem. In the deterministic case where $Q = 0$, $R = 0$, given A, N, B, C, D, the objective of the problem is to design observer gains M_1 and M_2 so that in the limit as k tends to infinity, the state estimation error $e(k)$ converges to zero. In addition, we also seek to answer the question if a deadbeat bilinear observer exists that causes the state estimation error to converge to zero identically in a minimum number of time steps. In the stochastic case, we assume further that the process and measurement noise covariances Q and R are known, and the design problem is to find the appropriate bilinear observer gains M_1 and M_2 so that the expected value of the squared norm of the steady-state state estimation error, $\mathbb{E}\big[e^T(k)e(k)\big]$, is minimized. This requirement is consistent with that of the Kalman filter. Indeed, when the problem is linear, $N = 0$, the solution presented in this paper produces the well-known steady-state Kalman filter gain K ($M_1 = K$, $M_2 = 0$).

III. A DETERMINISTIC BILINEAR OBSERVER

In the absence of noise, the dynamics of the state estimation error is described by

$$e(k+1) = \bar{A}e(k) + \bar{N}e(k)u(k) \tag{5}$$

From Eq. (5) $e(k)$ can be interpreted as being governed by a linear-time-varying difference equation with a time-varying dynamic matrix $\bar{A} + \bar{N}u(k)$. Convergence of $e(k)$ to zero is in general dependent on $u(k)$, which is exogenous to the observer. Propagating Eq. (5) forward in time produces

$$\begin{aligned}
e(1) &= \bar{A}e(0) + \bar{N}e(0)u(0) \\
e(2) &= \bar{A}^2e(0) + \bar{A}\bar{N}e(0)u(0) \\
&\quad + \bar{N}\bar{A}e(0)u(1) + \bar{N}^2e(0)u(0)u(1) \\
e(3) &= \bar{A}^3e(0) + \bar{A}^2\bar{N}e(0)u(0) \\
&\quad + \bar{A}\bar{N}\bar{A}e(0)u(1) + \bar{A}\bar{N}^2e(0)u(0)u(1) \\
&\quad + \bar{N}\bar{A}^2e(0)u(2) + \bar{N}\bar{A}\bar{N}e(0)u(0)u(2) \\
&\quad + \bar{N}^2\bar{A}e(0)u(1)u(2) + \bar{N}^3e(0)u(0)u(1)u(2) \\
&\dots \\
e(k) &= \mathcal{S}_k(k)e(0)
\end{aligned} \tag{6}$$

Observe that the relationship that expresses $e(k)$ in terms of the previous input values $u(k-1), ..., u(0)$ and the initial error $e(0)$ contain $\bar{A}^k$, $\bar{N}^k$, and all possible products of $\bar{A}$ and $\bar{N}$ whose combined powers add up to k. The sum of such terms (multiplied by appropriate input values) is compactly denoted by $\mathcal{S}_k(k)$ in Eq. (6). The two observer gains M_1 and M_2 can be seen to provide the design freedom for $\mathcal{S}_k(k)$ to converge to zero. In mathematics literature the topic is well-known under the name of *Infinite Product of Matrices*. To our knowledge, no general result is available that guarantees the convergence to zero of the above mentioned products as k tends to infinity regardless of the magnitude of $u(k)$. However, when the magnitude of $u(k)$ is bounded, it is possible to guarantee the existence of the observer gains M_1 and M_2 to ensure that the state estimation error $e(k)$ converges to zero as k tends to infinity. This result can be explained in the context of state-space bilinear identification problem using interaction matrices.

A. Interaction Matrices

The interaction matrices are originally motivated by the problem of system identification of large flexible structures, Refs. [4], [6]. They are later found to be useful in a wide range of applications, due to their ability to provide a unifying framework that connects the state-space model to various input-output representations, Ref. [5].

The interaction matrices are recently extended to the discrete-time bilinear state-space identification problem in Refs. [7], [8]. There they are used to derive Input-Output-to-State Relationships (IOSRs). These IOSRs allow one to express the current state, $x(k)$, in terms of a fixed number of past input and output values, $u(k-1), ..., u(k-p), y(k-1), ..., y(k-p)$. Such relationships are proven to be asymptotically exact for a general bilinear system as p increases. It turns out that the approximation error of the IOSRs in the system identification problem, Ref. [8], has exactly the same structure as the state estimation error of Eq. (6). Therefore, the same theorem in Ref. [8], motivated by Ref. [9], that ensures the validity of the IOSRs in the system identification problem, also guarantees the existence of the observer gains M_1 and M_2 to cause the state estimation error $e(k)$ to converge to 0 as k tends to infinity, provided that (A, C) is a detectable pair, and the magnitude of the input $u(k)$ is bounded below a certain threshold. This result is summarized in the following theorem.

Theorem. *If (A, C) of a bilinear system is a detectable pair then there exists a value γ such that, for $|u(k)| < \gamma$, the state estimation error $e(k)$ of the bilinear observer given in Eq. (2) converges to zero as k tends to infinity.*

The interaction matrix formulation offers a fundamental connection between the system identification problem and the state estimation problem. One can indeed use system identification as a technique to find the desired observer gains. This approach could be more appropriately named *observer identification*. In the bilinear system identification methods of Refs. [7], [8] the interaction matrices are used to derive the desired IOSRs, but they don't need be found explicitly for the identification of the discrete-time bilinear state-space model matrices A, N, B, C, D. This same formulation can now be exploited in reverse order to identify the observer gains M_1 and M_2 given the bilinear state-space model A, N, B, C, D. Here one has the additional advantage of knowing the system state because the model is given. State information is not available in the system identification problem where only input-output data is known.

B. Observer Gain Identification

A special consideration in the deterministic case is to determine if suitable observer gains exist that would cause the state estimation error $e(k)$ to become identically zero after a finite number of time steps. This is analogous to the case of a deadbeat observer for a linear system. The theorem does not imply that all the possible products of $\bar{A}$ and $\bar{N}$ whose powers add up to a certain value vanish identically. In other words, unlike the linear case, no deadbeat observer is guaranteed to exist for a general bilinear system. Deadbeat observers only exist for a very limited class of bilinear systems where A and N satisfy certain restrictive conditions as illustrated in Ref. [7]. These systems are referred to as *ideal* in this work. For *non-ideal* systems, although deadbeat observers do not exist to cause the state estimation error to converge to zero identically in a finite number of time steps, observer gains can still be found to cause the state estimation error to converge to zero asymptotically. The design of these observer gains is described below.

Propagating Eq. (1) forward in time by p time steps,

$$\hat{x}(k) = T_p z_p(k) + \mathcal{S}_p(k)\hat{x}(k-p) \tag{7}$$

See Refs. [7], [8] for the general structure of T_p and $z_p(k)$. For example, for $p = 2$, we have

$$T_2 = \begin{bmatrix}\bar{A}\bar{B} & \bar{N}\bar{B} & \bar{B}\end{bmatrix} \tag{8}$$

$$z_2(k) = \begin{bmatrix}v(k-2) & v(k-2)u(k-1) & v(k-1)\end{bmatrix}^T \tag{9}$$

$$\begin{aligned}\mathcal{S}_2(k) = \bar{A}^2 + \bar{A}\bar{N}u(k-2) + \bar{N}\bar{A}u(k-1)\\ +\bar{N}^2u(k-2)u(k-1)\end{aligned} \tag{10}$$

By the previous theorem, if p is chosen to be sufficiently large then $\mathcal{S}_p(k) \to 0$. For $k \geq p$, Eq. (7) becomes,

$$\hat{x}(k) = T_p z_p(k) \tag{11}$$

Equation (11) expresses the estimated state $\hat{x}(k)$ in terms of p past input and p past output measurements, $u(k-1), ..., u(k-p), y(k-1), ..., y(k-p)$. Following the same terminology used in Ref. [8], Eq. (11) is an observer IOSR. Since $e(k) = x(k) - \hat{x}(k)$, Eq. (11) can be rewritten as

$$x(k) = T_p z_p(k) + e(k) \tag{12}$$

leading to the following matrix relationship

$$X = T_p Z_p + E \tag{13}$$

where

$$X = \begin{bmatrix}x(p) & x(p+1) & x(p+2) & ... & x(l)\end{bmatrix} \tag{14}$$

$$Z_p = \begin{bmatrix}z_p(p) & z_p(p+1) & z_p(p+2) & ... & z_p(l)\end{bmatrix} \tag{15}$$

$$E = \begin{bmatrix}e(p) & e(p+1) & e(p+2) & ... & e(l)\end{bmatrix} \tag{16}$$

where l is the final time step in a data record. This relationship is used to develop a data-based approach to design the bilinear observer gains from a given model of the system as described below.

Using the system state-space model, A, N, B, C, D, state and output data, denoted by $\{x(k)\}$ and $\{y(k)\}$, can be generated from one or more sufficiently long and rich input data records $\{u(k)\}$. The matrices X and Z_p can be formed and the matrix T_p solved for,

$$\tilde{T}_p = X(Z_p)^+ \tag{17}$$

where the superscript $+$ denotes the pseudo-inverse operation. The observer gains M_1 and M_2 are extracted directly from $\tilde{T}_p$ because they appear explicitly in $\bar{B}$ which is in $\tilde{T}_p$. For an ideal bilinear system, T_p can be found to satisfy Eq. (12) exactly with $e(k) = 0$ using Eq. (17). The observer gains extracted from T_p when p is minimum are the deadbeat observer gains that would cause the state estimation error to converge to zero identically in p time steps. For the more general case of a non-ideal system, the pseudo-inverse solution corresponds to an observer that minimizes the Frobenius norm of the state estimation error matrix E. A major drawback of the solution given in Eq. (17) is that the dimension of Z_p grows exponentially with p. This motivates the need for another approach, which is described in the following section for the stochastic case.

IV. A STOCHASTIC BILINEAR OBSERVER

For a linear system, the optimal observer in the presence of noise is the Kalman filter. In this work we are concerned with the steady-state Kalman filter. In the following development, we work out a solution for the linear case, and then extend the result to the bilinear case. As mentioned, the main goal of this solution for the bilinear case is to overcome the high dimensionality associated with Eq. (17). In the linear case, the solution produces the steady-state Kalman filter gain.

A. The Linear Case

In the linear case, $N = 0$, $M_2 = 0$, and the matrices T_p, $z_p(k)$, $\mathcal{S}_p$ become

$$T_p = \begin{bmatrix}\bar{A}^{p-1}\bar{B} & ... & \bar{A}\bar{B} & \bar{B}\end{bmatrix} \tag{18}$$

$$z_p(k) = \begin{bmatrix}v(k-p) & ... & v(k-2) & v(k-1)\end{bmatrix}^T \tag{19}$$

$$\mathcal{S}_p = \bar{A}^p \tag{20}$$

In the linear case, $v(k) = \begin{bmatrix}u(k), y(k)\end{bmatrix}^T$. If p is chosen to be sufficiently large such that $\bar{A}^p$ can be neglected, the least-squares solution $\tilde{T}_p = X(Z_p)^+$ minimizes the Frobenius norm of E which is

$$\gamma = \sqrt{\sum_{k=p}^{l}\sum_{i=1}^{n} e_i(k)^2} \qquad i = 1, 2, ..., n \tag{21}$$

We now argue that $\tilde{T}_p$ contains the steady-state Kalman filter gain. Many approaches exist in literature to derive the steady-state Kalman filter. For example, Ref. [6] finds it as the unique linear filter minimizing the expected value of the squared norm of the state estimation error at any time step of a stationary process of the form given by Eq. (1) with $N = 0$. In other words, the unique observer of the form given by Eq. (2) where $M_2 = 0$, with M_1 as the observer gain, minimizes $\mathbb{E}\left[e^T(k)e(k)\right]$. Such an observer gain can be computed by solving an algebraic Riccati equation and will be denoted by K. Assuming the input-state-output data, denoted by $\{u(k)\}$, $\{x(k)\}$, $\{y(k)\}$, come from the system in question in stationary conditions, the expected value of the squared norm of the state error can be evaluated as

$$\mathbb{E}[e^T(k)e(k)] = \lim_{l\to\infty} \frac{1}{l-p+1}\sum_{k=p}^{l} e^T(k)e(k) \tag{22}$$

$$= \lim_{l\to\infty} \frac{1}{l-p+1}\sum_{k=p}^{l}\sum_{i=1}^{n} e_i^2(k) \tag{23}$$

$$= \lim_{l\to\infty} \frac{\gamma^2}{l-p+1} \tag{24}$$

where the last expression is proportional to the Frobenius norm, minimized by $\tilde{T}_p = X(Z_p)^+$. Therefore, the solution $\tilde{T}_p = X(Z_p)^+$, in the limit as the data record length tends to infinity, minimizes the same objective function minimized by K. Since the Kalman filter is the unique linear filter minimizing $\mathbb{E}\big[e^T(k)e(k)\big]$, it follows that $\tilde{T}_p = X(Z_p)^+$ must contain the state Markov parameters of the Kalman filter, i.e., the products $\bar{A}^j\bar{B}$, $j = 0, 1, ..., p-1$, where $\bar{A}$ and $\bar{B}$ are defined by Eq. (3) with $M_1 = K$ and $M_2 = 0$. Hence, M_1 extracted from $\tilde{T}_p$ converges to K, completing the argument.

B. The Bilinear Case

Although Eq. (17) can be immediately used in the stochastic case for a bilinear model, the high dimensionality of Z_p can be impractical to implement. We now develop an alternate solution. Starting with

$$x(k+1) = \hat{x}(k+1) + e(k+1) \tag{25}$$

and substituting the expression for $\hat{x}(k+1)$ produces

$$x(k+1) = \bar{A}\hat{x}(k) + \bar{N}\hat{x}(k)u(k) + \bar{B}v(k) + e(k+1) \tag{26}$$

Since $\hat{x}(k) = x(k) - e(k)$, Eq. (26) becomes

$$\begin{aligned} x(k+1) = \bar{A}\left[x(k) - e(k)\right] + \bar{B}v(k) \\ +\bar{N}\left[x(k) - e(k)\right]u(k) + e(k+1) \end{aligned} \tag{27}$$

Because we are interested in the batch-form of the solution, it is more convenient to rewrite Eq. (27) as

$$x(k+1) = PV(k) + e(k+1) \tag{28}$$

where the bilinear observer gains M_1 and M_2 are explicitly present in $\bar{B}$ in P,

$$P = \begin{bmatrix} \bar{A} & \bar{N} & \bar{B} \end{bmatrix} \tag{29}$$

$$V(k) = \begin{bmatrix} x(k) \\ x(k)u(k) \\ v(k) \end{bmatrix} - \begin{bmatrix} e(k) \\ e(k)u(k) \\ 0 \end{bmatrix} \tag{30}$$

Equation (28) can be written using all available data as

$$X = P\left(V_X - V_E\right) + E \tag{31}$$

where

$$X = \begin{bmatrix} x(p) & x(p+1) & \cdots & x(l) \end{bmatrix} \tag{32}$$

$$V_X = \begin{bmatrix} x(p-1) & \cdots & x(l-1) \\ x(p-1)u(p-1) & \cdots & x(l-1)u(l-1) \\ v(p-1) & \cdots & v(l-1) \end{bmatrix} \tag{33}$$

$$V_E = \begin{bmatrix} e(p-1) & \cdots & e(l-1) \\ e(p-1)u(p-1) & \cdots & e(l-1)u(l-1) \\ 0 & \cdots & 0 \end{bmatrix} \tag{34}$$

$$E = \begin{bmatrix} e(p) & e(p+1) & \cdots & e(l) \end{bmatrix} \tag{35}$$

Equation (31) is now in a form that existing standard generalized (or extended) least-squares methods can be adapted to in order to find the matrices $\bar{A}$, $\bar{N}$, and $\bar{B}$ that minimize the Frobenius norm of the state estimation error E. A generalized least-squares (GLS) algorithm based on Ref. [10] is described in the next section.

C. Optimal Bilinear Observer Gains by GLS

Let the superscript j denote the iteration number, starting from $j = 1$. Using $\{u(k)\}$, $\{x(k)\}$, $\{y(k)\}$ data that are generated from the given bilinear model and the specified process and measurement noise covariances, we form X and V_X, then solve for $P^{(1)}$ and compute the corresponding error matrix $E^{(1)}$ associated with this solution,

$$P^{(1)} = X\left(V_X\right)^+ \tag{36}$$

$$E^{(1)} = X - P^{(1)}V_X \tag{37}$$

The entries in $E^{(1)}$ are denoted by the superscript (1),

$$E^{(1)} = \begin{bmatrix} e^{(1)}(p) & e^{(1)}(p+1) & \cdots & e^{(1)}(l) \end{bmatrix} \tag{38}$$

$E^{(1)}$ is now used to update the parameter estimate. Since V_E calls for $e(p-1)$, but $e(p-1)$ is not available in E which starts with $e(p)$, the first column of X must now start from $p+1$ instead of p,

$$X^{(1)} = \begin{bmatrix} x(p+1) & x(p+2) & \cdots & x(l) \end{bmatrix} \tag{39}$$

V_X associated with $X^{(1)}$ is then adjusted accordingly so that V_E can start with $e^{(1)}(p)$ which is available in $E^{(1)}$,

$$V_X^{(1)} = \begin{bmatrix} x(p) & \cdots & x(l-1) \\ x(p)u(p) & \cdots & x(l-1)u(l-1) \\ v(p) & \cdots & v(l-1) \end{bmatrix} \tag{40}$$

$$V_E^{(1)} = \begin{bmatrix} e^{(1)}(p) & \cdots & e^{(1)}(l-1) \\ e^{(1)}(p)u(p) & \cdots & e^{(1)}(l-1)u(l-1) \\ 0 & \cdots & 0 \end{bmatrix} \tag{41}$$

The next update $P^{(2)}$ and $E^{(2)}$ can be computed from

$$P^{(2)} = X^{(1)}\left(V_X^{(1)} - V_E^{(1)}\right)^+ \tag{42}$$

$$E^{(2)} = X^{(1)} - P^{(2)}\left(V_X^{(1)} - V_E^{(1)}\right) \tag{43}$$

$E^{(2)}$ is then used to update the parameter estimate. The first entry in $E^{(2)}$ is $e^{(2)}(p+1)$ which is consistent with $X^{(1)}$. For the next iteration, $X^{(2)}$ must start from $x(p+2)$, $V_X^{(2)}$ from $x(p+1), u(p+1), v(p+1)$ so that $V_E^{(2)}$ can start from $e^{(2)}(p+1)$, etc. To avoid losing one data sample at each GLS iteration, an alternate strategy is inserting $e^{(1)}(p-1) = 0$ to Eq. (38), so that $V_X^{(1)}$ can remain the same as V_X. The first column of $V_E^{(1)}$ which now starts with $e^{(1)}(p-1) = 0$ will be zero. Once P is identified, the bilinear observer gains M_1 and M_2 can be directly extracted from it.

V. Examples

Numerical examples are provided to support the theoretical findings and to provide additional insights. State and output data $\{x(k)\}$, $\{y(k)\}$ are generated from random input sequences $\{u(k)\}$ that are uniformly distributed between -0.5 and 0.5. For the stochastic case, zero-mean Gaussian process and measurement noises are added with the following covariances Q and R, respectively,

$$Q = \begin{bmatrix} 0.0025 & 0.005 \\ 0.005 & 0.01 \end{bmatrix} \qquad R = 0.04$$

Three systems are used in the illustration. System I (ideal bilinear) is defined by ($D = 0$)

$$A = \begin{bmatrix} 0 & 0.5 \\ 0.5 & -0.5 \end{bmatrix} \quad N = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad C = \begin{bmatrix} 0 & 1 \end{bmatrix}$$

System II (non-ideal bilinear) is modified from System I with the $(1, 1)$ element of N set to 0.3, which is sufficient to turn it into a non-ideal bilinear model. System III (linear) is modified from System I by setting $N = 0$.

A. Deterministic Bilinear Observer

For System I, $p = 2$ is the smallest value for which T_p can be found to satisfy Eq. (13) exactly ($E = 0$ when T_2 is solved for). This is an indication that a deadbeat observer exists for this system. The identified gains associated with this minimum value of $p = 2$ are extracted from T_2 as

$$M_1 = \begin{bmatrix} 0.5 & -0.5 \end{bmatrix}^T \qquad M_2 = \begin{bmatrix} -1 & -1 \end{bmatrix}^T$$

These bilinear observer gains can be verified to cause $\bar{A}^2 = \bar{A}\bar{N} = \bar{N}\bar{A} = \bar{N}^2 = 0$ identically. The state estimation error converges to zero in 2 time steps for any input (Fig. 1). For System II, the smaller singular values of Z_p decrease gradually as p increases, and a solution for T_p that satisfies Eq. (13) exactly with $E = 0$ does not exist. This is a certain indication that there is no deadbeat observer for this system. Indeed, the identified gains by Eq. (17) do not make $\bar{A}^p$, $\bar{N}^p$, and all possible products of $\bar{A}$ and $\bar{N}$ whose combined powers add up to p, vanish identically. Instead, these gains minimize the Frobenius norm of the state estimation error E in Eq. (13). Although the identified observer is not deadbeat, numerical results suggest that it is the fastest observer when compared to observers whose gains are perturbed from the identified values. The case for $p = 6$ is illustrated in Fig. 2.

B. Stochastic Bilinear Observer

In the stochastic case, we first confirm that the proposed observer identification techniques indeed reproduce the well-known steady-state Kalman filter gain for a linear system (System III). Both the IOSR-based and the iterative GLS techniques identify the Kalman filter gain correctly. The two elements of M_1 are shown in the table below ($M_2 = 0$). Since the design objective is to minimize the expected value of the norm of the state estimation error, an appropriate performance measure is the mean-squared state estimation error $e_{MS} = \frac{1}{l+1} \sum_{k=0}^{l} e^T(k)e(k)$, which in the limit as $l \to \infty$ converges to the expected value of the squared norm of the state estimation error. We use $p = 40$ for the IOSR-based method.

	Kalman	IOSR	GLS
$M_1(1)$	-0.11719372	-0.11715649	-0.11713571
$M_1(2)$	0.08575711	0.08593812	0.08585536
$\|\|M_1\|\|$	0.14521931	0.14529623	0.14523056
e_{MS}	0.017073555	0.017073561	0.017073558

To identify the Kalman filter gain exactly, an infinitely long data set would be necessary. To avoid the obvious computational issues, the above gains (IOSR and GLS) are estimated by averaging the identified gains from 1000 independent data sets of 10^4 samples each. For evaluation, both in this case and later in the bilinear case, another 1000 independent data sets of 10^4 samples are used, and the averaged e_{MS} values are reported. Note that the relatively large value of $p = 40$ used in the IOSR approach guarantees that $\mathcal{S}_p$ is negligible. With the identified observer gain, $\mathcal{S}_{40} = \bar{A}^{40}$ has entries that are of the order of 10^{-7} in magnitude.

In the bilinear case, there is no known result of a bilinear Kalman filter for direct comparison. It could be argued that the identified bilinear observer for System II is in fact that optimal bilinear Kalman filter. We evaluate the effectiveness of the proposed design methods by analyzing the actual performance of the identified bilinear observers. The following table compares the e_{MS} of the observers designed by the IOSR-based technique (again from 1000 data sets of 10^4 samples each) for different values of p.

p	2	3	4	5	6
# Z_p rows	12	28	60	124	252
$100 \times e_{MS}$	3.4761	2.5985	2.4594	2.4147	2.4062

As expected, increasing p leads to better identification, since a larger p ensures better approximation of Eq. (13). For $p = 7$, the identified gains are

$$M_{1,IOSR} = \begin{bmatrix} -0.14499 \\ 0.10556 \end{bmatrix} \qquad M_{2,IOSR} = \begin{bmatrix} -0.31271 \\ -0.21050 \end{bmatrix}$$

In the linear case, the number of rows of Z_p grows linearly with p. In the bilinear case, the growth is exponential, hence the iterative GLS technique is called for. The bilinear observer gains based on the GLS technique are found to be

$$M_{1,GLS} = \begin{bmatrix} -0.14206 \\ 0.10054 \end{bmatrix} \qquad M_{2,GLS} = \begin{bmatrix} -0.30874 \\ -0.20073 \end{bmatrix}$$

The e_{MS} of the observer with gains designed by the GLS technique is 2.4035. For the IOSR design, it is 2.4039. The slightly larger error can be explained by the fact that $p = 7$ is not sufficiently large to make $\mathcal{S}_7(k)$ negligible. For instance, the entries of $(A + M_1C)^7$ and $(N + M_2C)^7$ are of the order of 10^{-2}. A larger p would improve the identification accuracy, but computational and ill-conditioning issues might in practice bound p. As a confirmation we compare the state estimation error of the GLS-derived observer to the observers whose gains are perturbed from the values given above. The resulting e_{MS} values, obtained by averaging 1000 tests of 10^4 samples each, are shown in Fig. 3. The identified observer by the GLS technique performs better than all 100 observers with randomly perturbed gains. In fact, this identified bilinear observer for System II can legitimately be considered as being optimal or very close to it. The state estimation error is minimized, and the output residual, $y(k) - \hat{y}(k)$, is white (Fig. 4). The state estimation error is not expected to be white. These results exactly parallel those of the optimal Kalman filter in the linear case. Finally, fast convergence is observed for the proposed GLS technique. In the above examples, in fewer than 100 iterations, the gains in successive iterations converge with a relative difference of about 10^{-16} in order of magnitude, which is numerically zero by Matlab© double precision calculation.

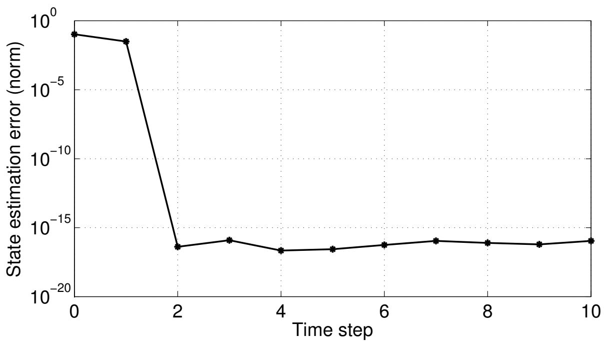

Fig. 1: State estimation error of deadbeat bilinear observer for System I converges to numerically zero in 2 time steps.

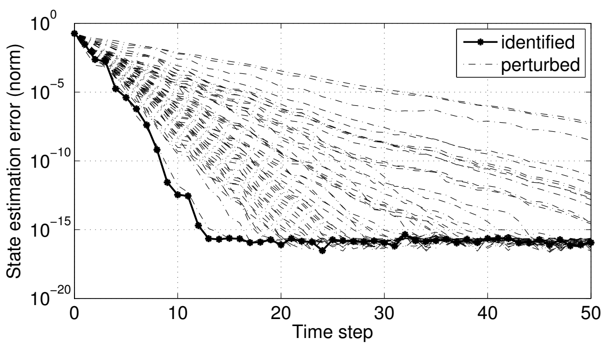

Fig. 2: Observer for System II (deterministic) has the fastest convergence compared to those with perturbed gains.

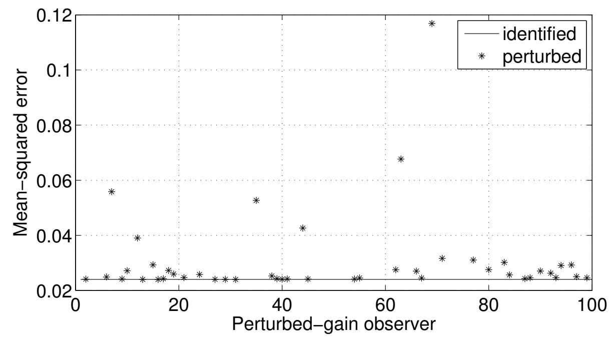

Fig. 3: Observer for System II (stochastic) has minimum state estimation error compared to those with perturbed gains.

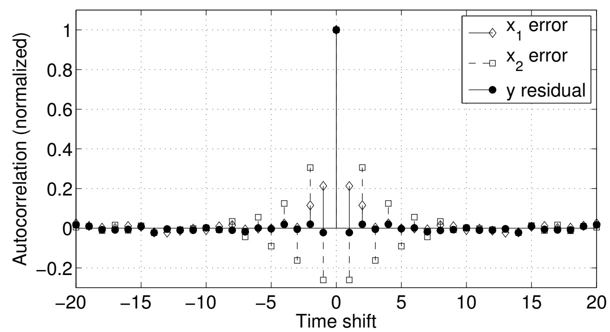

Fig. 4: Autocorrelation of state estimation error and output residual for System II. The output residual is white.

VI. Conclusions

In this paper we have formulated a bilinear observer for a bilinear state-space model. The key feature of the proposed bilinear observer is its connection to the interaction matrices. The interaction matrices are originally developed in the context of the Observer/Kalman filter identification algorithm (OKID) to identify a linear state-space model of the system and an associated optimal observer/Kalman filter gain from input-output data without requiring explicit knowledge of the process and measurement noise covariances. The results presented in this paper serve as a fundamental first step for a bilinear version of OKID. The connection to the interaction matrices is also important because of their linkage to system identification. We take the approach of identifying these observer gains with data generated from the known model and noise covariances (an observer identification problem), instead of finding their closed-form solution. In the absence of noise, a deadbeat bilinear observer, where the estimated state estimation error converges to zero identically in a finite number of time steps, does not exist for a general bilinear model, but it does exist for certain bilinear models. We take the pragmatic approach of using the observer identification algorithms to determine whether these deadbeat observer gains exist, because if they do, the proposed algorithms will find them. In the presence of noise, we formulate an optimal bilinear observer that minimizes the state estimation error in the same manner that the Kalman filter does to a linear system. Numerical examples successfully illustrate both the theoretical and computational aspects of the new results.

References

[1] S. Svoronos, G. Stephanopoulos and A. Rutherford, "Bilinear Approximation of General Non-Linear Dynamic Systems with Linear Inputs," *International Journal of Control*, Vol. 31, No. 1, 1980, pp. 109-126.

[2] K. Kowalski and W.-H. Steeb, *Nonlinear Dynamical Systems and Carleman Linearization*, World Scientific, 1991.

[3] C.-H. Lee and J.-N. Juang, "Nonlinear System Identification - A Continuous-Time Bilinear State Space Approach," *Proceedings of the Kyle T. Alfriend Astrodynamics Symposium*, Monterey, CA, 2010.

[4] J.-N. Juang, M.Q. Phan, L.G. Horta and R.W. Longman, "Identification of Observer/Kalman Filter Markov Parameters: Theory and Experiments," *Journal of Guidance, Control, and Dynamics,* Vol. 16, No. 2, 1993, pp. 320-329.

[5] M.Q. Phan, "Interaction Matrices in System Identification and Control," *Proceedings of the 15th Yale Workshop on Adaptive and Learning Systems*, New Haven, CT, 2011.

[6] J.-N. Juang, *Applied System Identification*, Prentice Hall, Englewood Cliffs, New Jersey, 1994.

[7] M.Q. Phan and H. Čelik, "A Superspace Method for Discrete-time Bilinear Model Identification by Interaction Matrices," *Advances in the Astronautical Sciences*, Vol. 139, 2011, pp. 445-464. *Journal of the Astronautical Sciences* (in press).

[8] F. Vicario, M.Q. Phan, R. Betti and R.W. Longman, "Linear State Representations for Identification of Bilinear Discrete-Time Models by Interaction Matrices," *Proceedings of the 23rd AAS/AIAA Space Flight Mechanics Meeting*, Kauai, HA, 2013, AAS13-337.

[9] K.E. Bouazza, M. Boutayeb and M. Darouach, "State and Output Feedback Stabilization of a Class of Discrete-Time Nonlinear Systems," *Proceedings of the 2004 American Control Conference*, Vol. 4, 2004, pp. 3023-3028.

[10] T.C. Hsia, "On Least Squares Algorithms for System Parameter Identification," *IEEE Transactions on Automatic Control*, Vol. 21, 1976, pp. 104-108.