A New Algorithm for Linear and Nonlinear ARMA Model Parameter Estimation Using Affine Geometry

Sheng Lu, Ki Hwan Ju, and Ki H. Chon*

Abstract—A linear and nonlinear autoregressive (AR) moving average (MA) (ARMA) identification algorithm is developed for modeling time series data. The new algorithm is based on the concepts of affine geometry in which the salient feature of the algorithm is to remove the linearly dependent ARMA vectors from the pool of candidate ARMA vectors. For noiseless time series data with *a priori* incorrect model-order selection, computer simulations show that accurate linear and nonlinear ARMA model parameters can be obtained with the new algorithm. Many algorithms, including the fast orthogonal search (FOS) algorithm, are not able to obtain correct parameter estimates in every case, even with noiseless time series data, because their model-order search criteria are suboptimal. For data contaminated with noise, computer simulations show that the new algorithm performs better than the FOS algorithm for MA processes, and similarly to the FOS algorithm for ARMA processes. However, the computational time to obtain the parameter estimates with the new algorithm is faster than with FOS. Application of the new algorithm to experimentally obtained renal blood flow and pressure data show that the new algorithm is reliable in obtaining physiologically understandable transfer function relations between blood pressure and flow signals.

Index Terms—Affine geometry, ARMA model, AR model, arterial blood pressure, blood flow, MA model, model-order selection, parameter estimation, renal autoregulation.

I. INTRODUCTION

IT IS a common practice in the biomedical field to fit certain physiological systems with linear and nonlinear autoregressive (AR) moving average (MA) (ARMA) models. For example, Abdel-Malek *et al.* [1] studied differences in parameters obtained by the ARMA model in a manual-tracking experiment between patients with Parkinson's disease and a normal control group. The ARMA model's popularity can be attributed to the relative ease with which the dynamics of physiological systems can be unveiled, using either transfer function analysis or impulse response functions (IRF) derived from an ARMA model. Of the formal order-estimation methods, perhaps the most well known are the Akaike information criterion (AIC) and the final prediction error (FPE), both introduced by Akaike [2], and the minimum description length (MDL) method of Rissanen [3]. The AIC and the FPE

methods have been shown to be asymptotically equivalent to the F-test by Soderstrom and Stoica, in a book [4] that reviews several order-estimation methods. Accordingly, many novel techniques have been developed to achieve even more accurate ARMA model parameter estimates [5]–[10]. Korenberg has developed a robust algorithm for difference equation modeling based on a fast orthogonal search (FOS) method [5], [6]. The FOS algorithm has been shown to be robust in most cases in obtaining correct ARMA model parameters despite incorrect model-order selection. It relies on the sequential search procedure to extract only the significant ARMA model terms, and discard the insignificant ones. We have recently developed another approach to ARMA model parameter estimation based on the group method of data handling (GMDH) [10]. The main idea of the GMDH is to have the algorithm construct a model of optimal complexity based only on the data; only the candidate terms that best approximate the given data are retained. The GMDH has been shown to be effective in obtaining accurate ARMA model parameter estimates in cases with significant noise contamination as well as *a priori* incorrect model-order selection. In some cases, the GMDH has been shown to provide better parameter estimates than either the FOS or the least-squares under the circumstances outlined in the previous sentence. However, there exists a scenario in which both the FOS and the GMDH methods do not provide accurate ARMA model parameter estimates under noiseless conditions with *a priori* incorrect model-order selection. This is the case when both methods are based on a suboptimal search criterion. An optimal search criterion would search for the minimum error across all possible subsets of ARMA model candidate functions within the candidate function space.

We present a new algorithm for an optimal search criterion, based on the principle of affine geometry, which enables accurate parameter estimation despite incorrect model selection. Affine geometry is a subset of Euclidean geometry. It mainly deals with points, straight lines and incidence (when a point lies on a line), therefore, affine geometry does not consider angles. The FOS algorithm is based on Euclidean geometry because orthogonality of candidate vectors are involved. Unlike the FOS, the proposed algorithm, based on the concepts of affine geometry, utilizes nonorthogonal projection search criteria, and from herein we will refer to the algorithm as the optimal parameter search (OPS). To showcase the efficacy of the OPS, comparisons between the OPS and the FOS are made using various linear and nonlinear ARMA model simulation examples. We chose to compare the performance of the OPS to FOS since FOS is one of the most accurate algorithms available, often superior
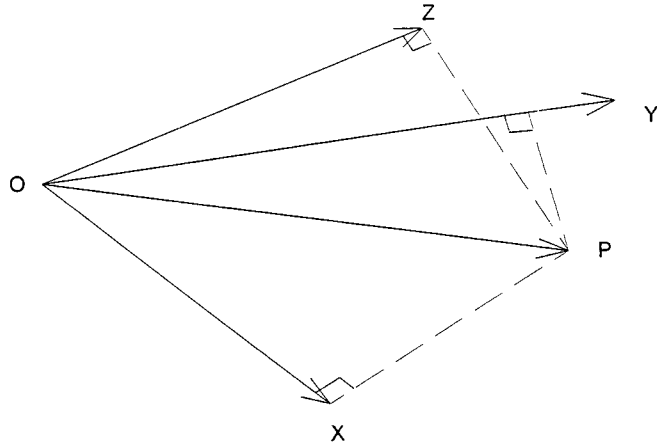
Fig. 1. Orthogonal projection finds the closest point to P in the span of the basis vectors.

to the most widely utilized least-squares algorithm with AIC as the model-order search criteria.

II. METHODS

The key difference between the OPS and the FOS algorithms is that, unlike the FOS, the OPS is based on a nonorthogonal search for model candidate terms. One notable disadvantage of FOS's using an orthogonal search can be seen by the following simple example. A vector P is in the space constructed by vectors X and Y as shown in Fig. 1. If the angle formed by ZOP is smaller than angles YOP and XOP, then the vector Z will be chosen even though the vector Z is not in the space constructed by the vector X and Y. In other words, orthogonal projection finds the closest point to P in the span of the basis vectors, regardless of whether or not the vector belongs to the constructed space. Note that this scenario will produce an erroneous parameter estimate with the FOS when an incorrect model order is chosen *a priori*. With the OPS, because it is based on a nonorthogonal search, this type of scenario will not occur. Note that if the base vectors are all perpendicular to each other, then there is no difference between orthogonal and nonorthogonal search methods.

The first step in the OPS algorithm is to select only the linearly independent vectors from the pool of candidate vectors. To examine how linearly independent vectors are selected, consider an ARMA process of the form

$$y(n) = \sum_{i=1}^{P} a(i)y(n-i) + \sum_{j=0}^{Q} b(j)x(n-j) + e(n) \quad (1)$$

where P and Q represent the maximum AR and MA model orders, respectively. The term $e(n)$ in (1) is considered a noise source or prediction-error term. The parameters $a(i)$ and $b(j)$ represent to-be-estimated coefficients of the AR and MA terms, respectively. The candidate vectors are the following: $y(n-1)$, ..., $y(n-P)$ and $x(n)$, ..., $x(n-Q)$. These candidate vectors can be arranged as the matrix shown in (2) at the bottom of the page, where N is the total number of data points. For a nonlinear ARMA model, the above matrix can be expanded to include products between the input and output itself as well as cross products between the input and output terms. The first step toward achieving linear independence among candidate vectors is to select $y(n-1)$ as the first candidate vector. The next candidate vector $x(n)$, and the first candidate vector $y(n-1)$ are then used to determine linear independence [e.g., use $y(n-1)$ to fit $x(n)$ using the least squares method and calculate the error between $x(n)$ and the estimated vector]. With a perfectly clean signal, linear independence will always be obtained. In the case of correlated noise contamination (error value will not be zero), some preset threshold can be set so that if the error value is smaller than the preset threshold, then the vector $x(n)$, for example, can be selected as an independent candidate vector. For the simulation examples to be considered in Section III, we used a very small preset threshold value (e.g., threshold value $= 0.001$) because it is often difficult to determine *a priori* what that value should be to obtain correct results, especially with experimental data. Once it has been determined that $x(n)$ is a linear independent candidate vector, the vectors $y(n-1)$ and $x(n)$ are used to estimate the candidacy of the linear independence of $y(n-2)$ using the approach just outlined. This procedure is continued until all the linearly independent vectors are selected to form a new candidate vector pool. Let $\phi = [w_0, w_1, w_2, \ldots, w_R]$ where R is the number of selected linearly independent vectors.

With the new candidate pool of linearly independent vectors, least-squares analysis is performed

$$y(n) = \theta_g^T \phi + e(n) \quad (3)$$

where

$$\theta_g = [g_0, g_1, \ldots, g_R]^T. \quad (4)$$

In (4), g_i is the coefficient estimate of the ARMA model. The objective is to minimize the equation error, $e(n)$, in the least-squares sense using the criterion function defined as follows:

$$J_N(\theta_g) = \left[y(n) - \theta_g^T \phi\right]^2. \quad (5)$$

$$\begin{matrix} y(0) & x(1) & y(-1) & x(0) & \cdots & y(1-P) & x(1-Q) \\ y(1) & x(2) & y(0) & x(1) & \cdots & y(2-p) & x(2-Q) \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ y(n-1) & x(n) & y(n-2) & x(n-1) & \cdots & y(n-P) & x(n-Q) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ y(N-1) & x(N) & y(N-2) & x(N-1) & \cdots & y(N-P) & x(N-Q) \end{matrix} \quad (2)$$
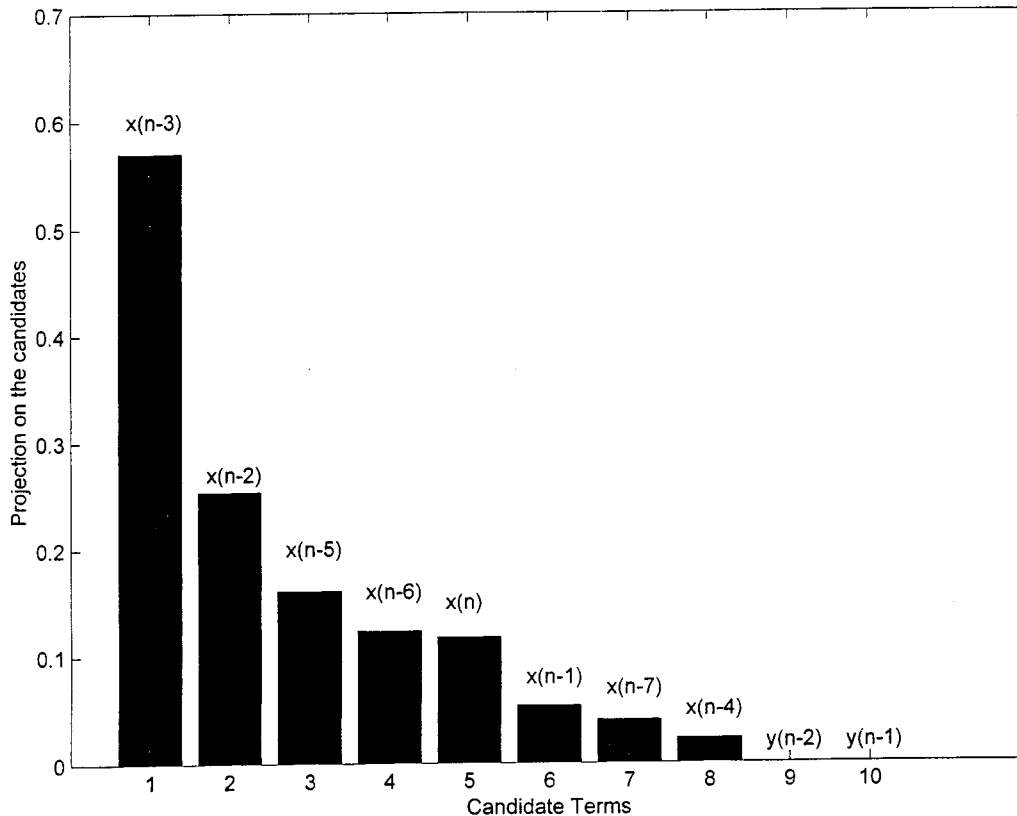
Fig. 2.   Model-order determination via the projection distance of the candidate terms (noiseless case).

The criterion function in (5) is quadratic in θ_g, and can be minimized analytically with respect to θ_g, yielding the following well-known equation:

$$\widehat{\theta}_g = \left[\phi\phi^T\right]^{-1}\phi y(n). \tag{6}$$

With the obtained coefficients, calculate every $\left|\overline{g_m^2 w_m^2}\right|$, and rearrange the w_m in descending order. Note that the overbar represents the time average. At this step of the algorithm we need to choose the number of candidate vectors, w_m, necessary for obtaining proper accuracy. The approach we have taken is to retain only the w_m that reduce the error value significantly. If we observe either negligible decrease or increase in the error value by adding an additional w_m, then those w_m are dropped from the model. Once only those w_m that reduce the error value significantly (only those candidate terms whose projection distance value is greatest) are obtained, the linear and nonlinear ARMA model terms are estimated using the least-squares method. This step is discussed further in Section III (see Figs. 2 and 3).

III. SIMULATION RESULTS

In this section we demonstrate the effectiveness of the developed algorithm for estimating parameters of linear and nonlinear MA and ARMA models. We compare the performance of the OPS to that of the FOS method. For all simulation examples involving linear processes (AR and MA) to follow, we have selected an incorrect model order of ten AR and ten MA terms [ARMA (10,10)] for both the OPS and FOS. For nonlinear models, ten AR, ten MA, and five second-order AR and MA as

well as five cross AR and MA [e.g., x(n − 1)y(n − 1)] model orders were chosen for both the OPS and FOS. The Achilles' heel of ARMA models is to determine *a priori* accurate model orders without knowing the true model order of the system. The model order search for the FOS was obtained by using the automatic model order search criteria as described in [5] and [6]. Succinctly, the automatic model order search criteria of the FOS retains only those candidate terms that reduce the mean-square-error values by a significant amount in conjunction with a statistical 95% interval criterion. The FOS algorithm has been shown to be accurate for various linear and nonlinear ARMA models [5], [6], [8], and often superior to the least-squares approach with the AIC for the model order selection process.

A. MA Model With Additive Noise and Incorrect Model-Order Selection

For the first simulation example, consider the following linear MA model with Gaussian white noise (GWN) as the input, $x(n)$, so that the output, $y(n)$, contains 1000 data points

$$
\begin{aligned}
y(n) = {} & 0.34x(n) - 0.23x(n-1) + 0.5x(n-2) \\
& + 0.75x(n-3) - 0.15x(n-4) - 0.4x(n-5) \\
& + 0.35x(n-6) - 0.2x(n-7).
\end{aligned} \tag{7}
$$

We have purposely selected an incorrect model order of ARMA(10,10) for both methods despite the fact that the above equation does not contain any AR model terms. To subject
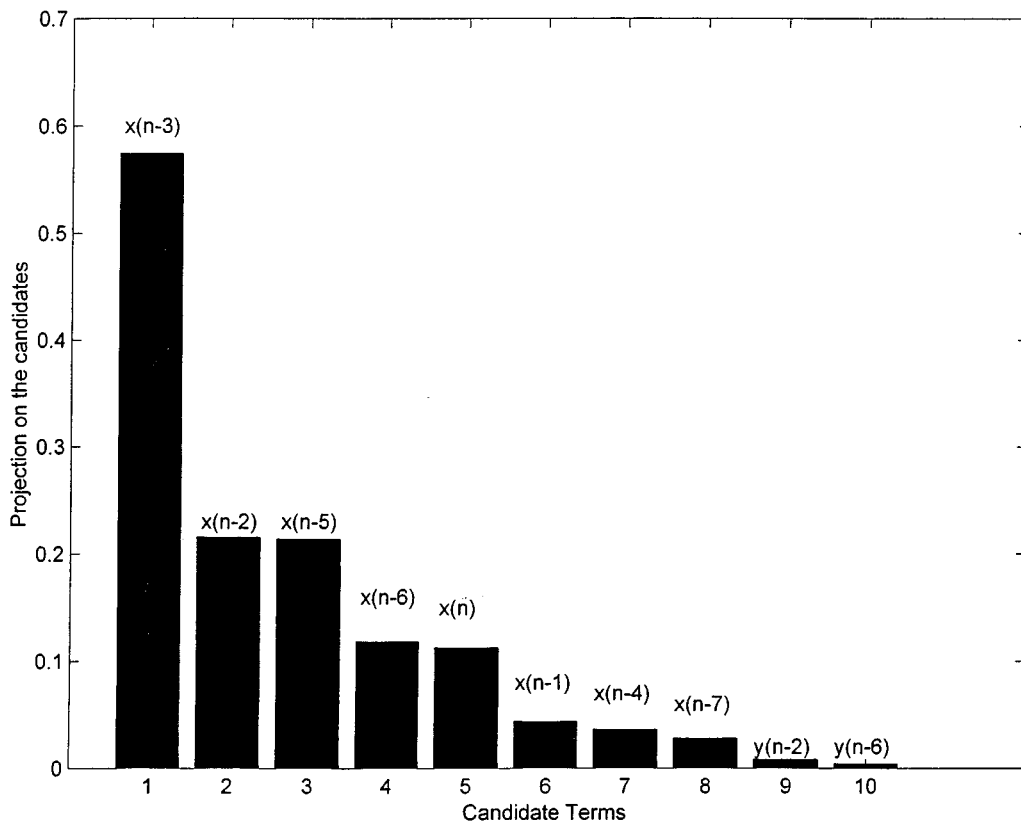
Fig. 3. Model-order determination via the projection distance of the candidate terms (SNR = 10 dB).

TABLE I
COMPARISON OF THE OPS AND FOS WITH ADDITIVE NOISE AND AN INCORRECT MODEL-ORDER SELECTION

Model Terms	$x(n)$	$x(n-1)$	$x(n-2)$	$x(n-3)$	$x(n-4)$	$x(n-5)$	$x(n-6)$	$x(n-7)$	$x(n-8)$	$y(n-1)$
True Value	**0.34**	**-0.23**	**0.5**	**0.75**	**-0.15**	**-0.4**	**0.35**	**-0.20**	**0.00**	**0.00**
OPS (Clean Data)	0.34	-0.23	0.5	0.75	-0.15	-0.4	0.35	-0.20	0.00	0.00
FOS (Clean Data)	0.34	0.0	0.36	1.06	0.32	-0.49	0.10	0.00	-0.12	-0.62
OPS (10 dB Noise)	0.34	-0.22	0.51	0.76	-0.13	-0.42	0.35	-0.20	0.00	0.00
FOS (10 dB Noise)	0.34	-0.18	0.48	0.83	0.00	-0.44	0.30	-0.15	0.00	-0.13
OPS (0 dB Noise)	0.34	-0.21	0.54	0.77	0.00	-0.45	0.36	-0.24	0.00	0.00
FOS (0 dB Noise)	0.33	-0.21	0.54	0.77	0.00	-0.46	0.37	-0.12	0.00	0.00

the algorithms to a more daunting task, we have used additive noise so that signal-to-noise ratios (SNRs) of 10 and 0 dB were obtained for the above MA process. Comparison of the results based on the two methods for the case of noiseless data with an incorrect model order selection, and the cases with noise added (10 and 0 dB) and an incorrect model order selection, are shown in Table I. With a clean signal, the OPS obtained true model terms and coefficients despite the exaggerated incorrect model order selection. The model order-selection process for the OPS is shown in Fig. 2. The ordinate value represents the projection distance of the candidate terms. Note that the projection distance value is zero after the term $x(n-4)$ has been determined. In addition, the projection distance value is the greatest for the term $x(n-3)$ and the lowest for $x(n-4)$ since they have the biggest and smallest coefficient values among the model candidate terms. Thus, without any noise contamination, due to removal of linearly dependent candidate terms, the OPS is able to obtain the correct coefficient and model candidate terms. The model-order search for the FOS was estimated using the automatic model-order search criteria as described above and in [5] and [6]. With this search, however, the FOS completely missed two model terms: $x(n-1)$ and $x(n-7)$. Concomitantly, the FOS incorrectly picked up two additional terms: $x(n-8)$ and $y(n-1)$. With a clean signal, the mean-square-errors (MSEs) are 0.0069 for the FOS and 0.00 for the OPS.

When the signal was corrupted with 10-dB noise, the OPS correctly identified only the true model terms but with a small deviation of the coefficients from the true coefficients, as expected. The MSE is found to be quite small (9.21e–004). The model order was determined with the aid of Fig. 3. Due to additive noise, the terms $y(n-2)$ and $y(n-6)$ erroneously have small projection distance values. In addition, the projection distance is quite negligible compared to the rest of the candidate terms. Thus, only the first eight candidate terms were used to estimate the candidate coefficients. The FOS fared poorly compared to the OPS, as it incorrectly identified additional terms that are not part of the true model. In addition, the identified coefficients deviated from the true coefficients, resulting in a slightly higher MSE (0.0060) value than for the OPS. With increased noise level (SNR of 0 dB), as expected, the performance of both FOS and OPS deteriorated further. Both methods missed the $x(n-4)$ term. It is interesting to note that in many simulations, including the present example, the FOS appears to be more robust with noise corrupted signals than with clean signals. For example, the MSE value is slightly lower with 10-dB noise (MSE = 0.006) than with the clean signal (MSE = 0.0069). In addition, with noise added, the FOS obtained more correct model terms than without noise added.

B. Nonlinear MA Model With Separate Additive or Dynamic Noise and Incorrect Model-Order Selection

The next simulation example consists of an arbitrarily chosen nonlinear MA difference equation of the form

$$
\begin{aligned}
y(n) = {} & 0.23x(n) + 0.4x(n-1) + 0.6x(n-2) \\
& + 0.54x(n-3) + 0.28x(n-4) + 0.8x(n-5) \\
& - 0.76x(n-6) + 0.35x(n-7) - 0.23x(n-8) \\
& + 0.22x(n-9) - 0.65x(n-1)x(n-3).
\end{aligned} \tag{8}
$$

We consider two separate cases of noise corruption, in which the output of (8) is contaminated by additive GWN of the form

$$
z(n) = y(n) + e(n) \tag{9}
$$

and the other in which the output of (8) is disturbed by dynamic noise, which causes (8) to take the form

$$
\begin{aligned}
y(n) = {} & 0.23x(n) + 0.4x(n-1) + 0.6x(n-2) + 0.54x(n-3) \\
& + 0.28x(n-4) + 0.8x(n-5) - 0.76x(n-6) \\
& + 0.35x(n-7) - 0.23x(n-8) + 0.22x(n-9) \\
& - 0.65x(n-1)x(n-3) + e(n-1).
\end{aligned} \tag{10}
$$

Note that additive noise [see (9)] is statically added after the clean output signal has been generated. For dynamic noise, the GWN source, $e(n)$, is fed back to the output so that the current and future output values are dependent on the past states of the input and noise signals. Therefore, the outputs described by (9) and (10) have different values. The SNR for additive noise was obtained for two different levels, 10 and 0 dB. For the dynamic noise, the SNR was 3.5 dB. A model order of ARMA (10,10)

and nonlinear ARMA (5,5) was incorrectly selected to determine the effectiveness of both approaches [the correct model is MA (9) and nonlinear MA (3). With incorrect selection of linear and nonlinear ARMA model orders, the number of parameters to be determined are 21 linear ARMA terms, and 57 nonlinear ARMA model terms, for a total of 78 terms to be searched. This is a daunting task for any algorithms, as (8) contained only 11 model terms, but we are subjecting the FOS and OPS to a superfluous model-order search concomitant with excessive noise corruption in the data signal. The results of additive and dynamic noise for the OPS and FOS are shown in Table II. As in the previous examples, with a clean signal, the OPS is again able to obtain accurate parameter estimates associated with only the true model terms. This result is impressive in itself, since we are not aware of any other algorithm that is able consistently to provide accurate parameter estimates despite incorrect model-order selection even in the case of a clean signal. The FOS, however, is not able to obtain correct parameter estimates and has obtained incorrect model terms [$y(n-1)$, $y(n-2)$, and $y(n-5)$] and missed some of the true model terms [$x(n-7)$, $x(n-8)$ and $x(n-9)$]. With additive noise, either SNR = 10 or 0 dB, the OPS is accurate in providing only the true model terms but the FOS missed a model term [$x(n-8)$, only for SNR = 10 dB] and it inaccurately introduced an additional term [$y(n-2)$]. Moreover, the estimated coefficients with the OPS are closer to the true model coefficients than are those obtained with the FOS. The MSE for the OPS are 0.0044 and 0.0442 for SNR = 10 and 0 dB, respectively. The MSE for the FOS are 0.035 and 0.1038 for SNR = 10 and 0 dB, respectively. It is clear that despite significant noise in the data, and grossly incorrect model-order selection, both methods perform rather nicely. However, it is clear that the OPS outperformed the FOS. The MSE values are approximately three- to ten fold less than those obtained via the FOS.

With dynamic noise, the result is the same as with additive noise; the OPS is more accurate than the FOS. The FOS is unable to identify the term $x(n-9)$, and incorrectly introduced a $y(n-2)$ term. The OPS, although the coefficients associated with the true model coefficients deviate somewhat, is accurate in only producing coefficients related to the true model terms. The MSE also favors the OPS; the MSE values are 0.013 for the OPS and 0.056 for the FOS.

C. The Effect of Incorrect Model-Order Selection and Additive Noise on a ARMA Model

The next simulation example considers the following linear ARMA model with GWN as the input, $x(n)$, so that the output, $y(n)$, contains 1000 data points

$$
\begin{aligned}
y(n) = {} & 0.35y(n-1) + 0.32y(n-2) + 0.1y(n-3) - 0.54x(n) \\
& + 0.34x(n-1) + 0.23x(n-2) + 0.21x(n-3).
\end{aligned} \tag{11}
$$

The objective is, based on only the measured input signal, $x(n)$, and the output signal, $y(n)$, to estimate the parameters of the above equation as accurately as possible. Although the true ARMA model order for the above process is three output lags and three input lags, we purposely selected an incorrect model

TABLE II
COMPARISON OF THE OPS AND FOS WITH ADDITIVE AND DYNAMIC NOISE WITH AN INCORRECT MODEL-ORDER SELECTION

ModelTerms	x(n)	x(n-1)	x(n-2)	x(n-3)	x(n-4)	x(n-5)	x(n-6)	x(n-7)	x(n-8)	x(n-9)	x(n-1)x(n-3)	y(n-1)	y(n-2)	y(n-5)
TrueValue	0.23	0.40	0.60	0.54	0.28	0.80	-0.76	0.35	-0.23	0.22	-0.65	0.0	0.0	0.0
CleanData														
OPS	0.23	0.40	0.60	0.54	0.28	0.80	-0.76	0.35	-0.23	0.22	-0.65	0.0	0.0	0.0
FOS	0.22	0.42	0.61	0.54	0.22	0.72	-0.73	0.00	0.00	0.00	-0.64	-0.14	0.22	0.06
AdditiveNoise														
OPS(10dB)	0.23	0.41	0.62	0.55	0.30	0.78	-0.75	0.35	-0.27	0.25	-0.65	0.00	0.00	0.00
FOS(10dB)	0.23	0.40	0.59	0.49	0.20	0.69	-0.80	0.22	0.00	0.19	-0.65	0.00	0.17	0.00
OPS(0dB)	0.22	0.43	0.66	0.58	0.35	0.72	-0.74	0.35	-0.37	0.30	-0.65	0.00	0.00	0.00
FOS(0dB)	0.00	0.43	0.65	0.57	0.32	0.70	-0.76	0.31	-0.34	0.28	-0.65	0.00	0.05	0.00
DynamicNoise														
OPS(3.5dB)	0.18	0.45	0.66	0.54	0.29	0.79	-0.79	0.35	-0.26	0.17	-0.66	0.00	0.00	0.00
FOS(3.5dB)	0.18	0.45	0.65	0.52	0.29	0.79	-0.81	0.31	-0.22	0.00	-0.66	0.00	0.04	0.00

TABLE III
COMPARISON OF THE OPS AND FOS WITH AN INCORRECT MODEL-ORDER SELECTION AND TWO DIFFERENT LEVELS OF NOISE
CONTAMINATION FOR AN ARMA MODEL

Model Terms	$y(n-1)$	$y(n-2)$	$y(n-3)$	$y(n-4)$	$x(n)$	$x(n-1)$	$x(n-2)$	$x(n-3)$	$x(n-4)$	$x(n-5)$
True Value	**-0.35**	**0.32**	**0.10**	**0.00**	**-0.54**	**0.34**	**0.23**	**0.21**	**0.00**	**0.00**
OPS (Clean Data)	-0.35	0.32	0.10	0.00	-0.54	0.34	0.23	0.21	0.00	0.00
FOS (Clean Data)	-0.24	0.00	0.00	0.11	-0.54	0.40	0.00	0.34	0.03	0.06
OPS (10 dB)	0.00	0.18	0.00	0.00	-0.54	0.54	-0.04	0.26	-0.10	0.08
FOS (10 dB)	-0.13	0.12	0.00	0.09	-0.55	0.47	0.00	0.28	0.00	0.04
OPS (0 dB)	0.00	0.09	0.00	0.00	-0.55	0.56	-0.10	0.28	-0.14	0.12
FOS (0 dB)	0.00	0.09	0.00	0.00	-0.56	0.57	-0.10	0.28	-0.15	0.12

order of ten output lags $(y(n-1), \ldots, y(n-10))$ and ten input lags $(x(n), x(n-1), \ldots, x(n-10))$ for both methods. The efficacy of the ARMA model relies on the algorithm's ability to find the true model order, despite an incorrect *a priori* model-order selection. Comparison of the results based on the two methods is shown in Table III. The OPS algorithm obtained the correct model terms and coefficients but the FOS did not; not only did the FOS miss three true model terms [$y(n-2)$, $y(n-3)$ and $x(n-2)$], but the coefficients obtained for the true model terms deviated from the exact coefficients of the model. In addition, the FOS incorrectly provided three additional model terms, $y(n-4)$, $x(n-4)$ and $x(n-5)$. The MSE of the OPS and the FOS are 0.0 and 4.17e–005, respectively. This is an example which indicates that the goodness of fit according

to the MSE value is not always a good measure, since the MSE of 4.17e–005 obtained with the FOS is a very small error. Note that for chaotic systems modeling, even small differences in the obtained coefficients would result in an exponentially divergent change in the output value. As detailed in Section I, the FOS, because it relies on a suboptimal model-order search, is not always able to obtain correct model terms and coefficients when it is subjected to an incorrect model-order selection even for a clean signal. This has been demonstrated in this and the previous examples.

When (11) was corrupted by additive noise levels of 10 and 0 dB, the OPS and the FOS provided comparable results. The simulations with additive noise were made further challenging with an *a priori* incorrect selection of ten output and ten input

TABLE IV
COMPARISON OF THE OPS AND FOS WITH AN INCORRECT MODEL-ORDER SELECTION AND TWO DIFFERENT LEVELS OF NOISE CONTAMINATION FOR A NONLINEAR ARMA MODEL

Model Terms	$x(n)$	$x(n-1)$	$x(n-2)$	$x(n-3)$	$x(n-4)$	$x(n-5)$	$y(n-1)$	$y(n-2)$	$x(n-1)^2$	$y(n-2)^2$	$x(n-1)y(n-1)$	$y(n-2)y(n-3)$	$y(n-2)y(n-6)$
True Value	0.8	-0.23	0.00	0.50	0.00	0.00	0.32	0.30	-0.20	0.12	-0.18	0.00	0.00
OPS (Clean Data)	0.8	-0.23	0.00	0.50	0.00	0.00	0.32	0.30	-0.20	0.12	-0.18	0.00	0.00
FOS (Clean Data)	0.79	-0.14	0.00	0.51	0.00	0.00	0.25	0.32	-0.31	-0.06	0.14	0.00	0.00
OPS (10 dB)	0.81	0.00	0.00	0.59	0.18	0.00	0.00	0.30	-0.23	0.04	-0.16	0.04	0.00
FOS (10 dB)	0.82	-0.16	0.00	0.51	0.00	0.00	0.27	0.26	-0.23	0.09	-0.14	0.00	0.00
OPS (0 dB)	0.84	0.00	0.12	0.60	0.22	0.11	0.00	0.10	-0.36	0.00	0.00	0.03	0.00
FOS (0 dB)	0.85	0.00	0.00	0.54	0.13	0.00	0.12	0.16	-0.29	0.00	-0.08	0.00	0.05

lags. The advantage of the OPS observed with only the incorrect model-order selection (no noise added) disappeared with the introduction of additive noise as evidenced by Table III. However, in terms of MSE values, the OPS performed slightly better than did the FOS. The MSE values for the OPS for 10 and 0 dB were 0.08 and 0.74, respectively; for the FOS, the corresponding values were 0.09 and 0.88, respectively.

D. Nonlinear ARMA (NARMA) Model With Incorrect Model-Order Selection and Additive Noise

As an arbitrary example of a NARMA model, the following 1000 data point difference equation was generated

$$y(n) = 0.8x(n) - 0.23x(n-1) + 0.5x(n-3)$$
$$+ 0.32y(n-1) + 0.3y(n-2) - 0.2x(n-1)^2$$
$$+ 0.12y(n-2)^2 - 0.18x(n-1)y(n-1). \quad (12)$$

The input, $x(n)$ terms, was generated using GWN. To completely test the features of the general NARMA model, (12) included self-nonlinear input and output terms as well as the cross-nonlinear term. The model order for the two methods compared was selected to be AR = 10, MA = 10, quadratic MA = 5, quadratic AR = 5 and cross nonlinear model order between the input and output = 5 (total number of parameters searched was 78). Table IV shows the results of the estimated coefficients obtained by the two methods for clean data and with two levels of additive noise. For the clean signal, as was the case with the three previous simulation examples, the OPS provided the correct terms and coefficients despite incorrect model-order selection; the FOS, however, shows the ill-effects of incorrect model-order selection as evidenced by all of the coefficients deviating from the true coefficients. When the output of (12) was corrupted by either 10 or 0 dB, the FOS provided comparable coefficient estimates to that of the OPS. However, in terms of MSE, both methods provided similar error values of 0.4% for 10 dB and 2% for 0 dB.

E. Application of the OPS to Experimental Data

In this section, we demonstrate the use of the OPS in analyzing experimentally obtained renal blood pressure and flow data. The aim is not to elucidate the physiological mechanisms involved in renal autoregulatory processes, but to examine if the OPS can provide similar impulse response functions (IRF) to those published and if those IRF's are at all physiologically meaningful [11], [12].

1) Data Acquisition and Experimental Procedure: The data analyzed in this investigation were obtained from a previously published study [11]. Experimental methods are described in detail in [11] and will be briefly summarized. The experimental data were collected from normotensive Sprague–Dawley rats using broadband perturbations of the arterial pressure (input) and measuring the resulting renal blood flow (output). Briefly, operating under halothane anesthesia, the aorta inferior to the renal arteries was cannulated with blood-filled tubing connected to a bellows pump which in turn was driven by a computer-controlled motor. Blood pressure was measured in the superior mesenteric artery with a standard pressure amplifier, and renal blood flow was measured in the left renal artery with an electromagnetic flow probe. The input signal was chosen to be a constant-switching-pace symmetric random signal (CSRS) that exhibited the spectral properties of bandlimited white noise [12]. A unique seed was used for the random number generator in each experiment.

Each of the experimental data records used for analysis was 256 s long, with a sampling rate of two samples per second (Nyquist frequency of 1 Hz), after digital low-pass filtering to avoid aliasing. Each data record, containing 512 data points, was subjected to second-degree polynomial trend removal (which included demeaning) and was normalized to unit variance.

Fig. 4 shows averaged impulse response functions (based on four recordings) computed from the ARMA coefficients obtained from analysis of the OPS [Fig. 4(a)] and from the FOS method [Fig. 4(b)]. For both the OPS and FOS, the model order of ARMA (10,10) was used, which was selected based on our previous work [13]. The dotted lines in the figure represent the standard deviation bounds of the sample mean. We observe that
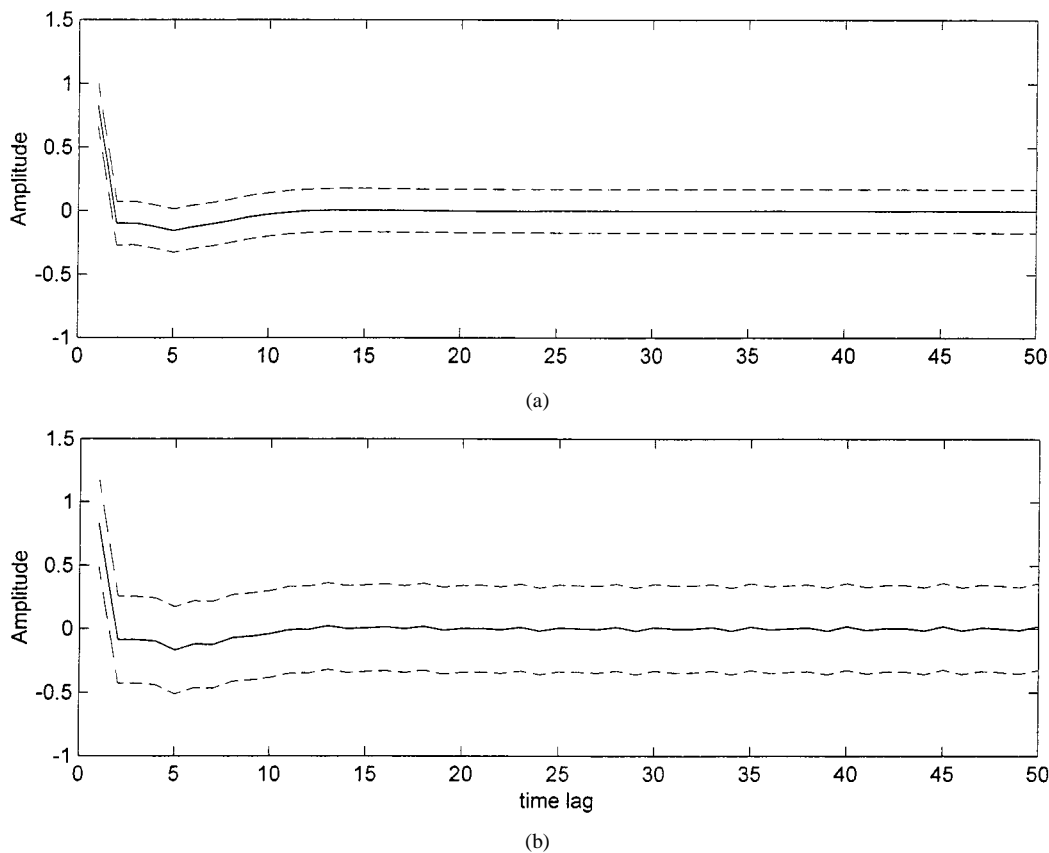
Fig. 4. Averaged impulse response functions of renal blood flow and pressure data obtained via ARMA model for (a) the OPS and (b) the FOS.

the estimated impulse response wave forms are rather consistent for both methods and are similar to those published [11], [13]. However, the impulse response function obtained via the OPS exhibits smoother waveforms and smaller standard deviation bounds than does its counterpart obtained from the FOS. To compare the performance of the two methods quantitatively, model predictions based on the linear ARMA model were computed for both methods. The average MSE obtained for the four datasets for the OPS and the FOS are 3.8% and 4.7%, respectively. As in the simulation examples presented in this paper, better model prediction is obtained with the OPS than with the FOS method.

It should be pointed out that the computational time for both methods is quite fast. The FOS uses a modified Cholesky decomposition to achieve orthogonality rather than inverting the matrix to solve for the least-squares estimation, consequently, the computational speed is enhanced. However, the computational time is faster with the OPS than with the FOS since the orthogonal procedure is not utilized with the OPS. For all of the simulations considered in this paper, the OPS was faster than the FOS, based on code using Matlab software for both methods. All of the simulation examples shown took <1 min each to compute on an Intel Pentium 500-MHz processor.

IV. CONCLUSION

In this paper we introduced a new algorithm, named the OPS. Simulation examples have shown the efficacy of the method and have shown that for clean signals, the OPS is able to extract only the correct parameters despite overdetermined incorrect

model-order selection. The FOS, one of the most accurate algorithms available, is also able to extract correct parameters under similar circumstances but its ability to obtain correct parameters for all noiseless data is not complete. Well-known model-order search criteria such as AIC and MDL also do not always provide accurate parameter estimates for noiseless signals. The OPS, unlike the FOS, does not orthogonalize model terms, resulting in faster computational time. Due to the fact that the OPS achieves linear independence among candidate vectors, it is able to obtain correct parameters despite *a priori* incorrect model-order selection for noiseless signals. In other words, the OPS is an optimal search method, thus, it is able to obtain correct model parameters for noiseless signals. To date, we are not aware of any other algorithm that is able to achieve this kind of remarkable result. For the case of noise contaminated linear and nonlinear MA models, the OPS provides performance superior to that of the FOS, as evidenced by the simulation examples considered in this paper. For noise-contaminated linear and nonlinear ARMA models, both the OPS and FOS provide similar excellent results. In addition, application of the OPS to experimental data indicates the feasibility of the method in obtaining physiologically meaningful transfer function relationships between renal pressure and flow data.

REFERENCES

[1] A. A. Abdel-Malek, C. H. Markham, P. Z. Marmarelis, and V. Z. Marmarelis, "Quantifying deficiencies associated with Parkinson's disease by use of time series analysis," *J. EEG Clin. Neurophys.*, vol. 69, pp. 24–33, 1988.

[2] H. Akaike, "Power spectrum estimation through autoregression model fitting," *Ann. Inst. Stat. Math.*, vol. 21, pp. 407–419, 1969.

[3] J. Rissanen, "A universal prior for the integers and estimation by minimum description length," *Ann. Stat.*, vol. 11, pp. 417–431, 1983.

[4] T. Soderstrom and P. Stoica, *System Identification*. Englewood Cliffs, NJ: Prentice-Hall, 1989.

[5] M. J. Korenberg, S. A. Billings, Y. P. Li, and P. J. McIlroy, "Orthogonal parameter estimation algorithm for nonlinear stochastic system," *Int. J. Contr.*, vol. 48, pp. 193–210, 1988.

[6] M. J. Korenberg, "A robust orthogonal algorithm for system identification and time series analysis," *Biol. Cybern.*, vol. 60, pp. 267–276, 1989.

[7] K. H. Chon and R. J. Cohen, "Linear and nonlinear ARMA model parameter estimation using an artificial neural network," *IEEE Trans. Biomed. Eng.*, vol. 44, pp. 168–174, Mar 1997.

[8] K. H. Chon, M. J. Korenberg, and N. H. Holstein-Rathlou, "Application of fast orthogonal search to linear and nonlinear stochastic systems," *Ann. Biomed. Eng.*, vol. 25, pp. 793–801, 1997.

[9] K. H. Chon, R. J. Cohen, and N. H. Holstein-Rathlou, "Compact and accurate linear and nonlinear autoregressive moving average model parameter estimation using Laguerre functions," *Ann. Biomed. Eng.*, vol. 25, pp. 731–738, 1997.

[10] K. H. Chon and S. Lu, "A new algorithm for ARMA model parameter estimation using group method of data handling," *Ann. Biomed. Eng.*, vol. 29, pp. 92–98, 2001.

[11] N. H. Holstein-Rathlou, A. J. Wagner, and D. J. Marsh, "Tubuloglomerular feedback dynamics and renal blood flow autoregulation in rats," *Amer. J. Physiol.*, vol. 260, pp. F53–F68, 1991.

[12] V. Z. Marmarelis, "A family of quasiwhite random signals and its optimum use in biological system identification. Part I: Theory," *Biol. Cybern.*, vol. 27, pp. 49–56, 1977.

[13] K. H. Chon, Y. M. Chen, N. H. Holstein-Rathlou, D. J. Marsh, and V. Z. Marmarelis, "Jan.," *IEEE Trans. Biomed. Eng.*, vol. 40, pp. 8–20, 1993.

Ki Hwan Ju received the M.Sc.E.E. degree in electrical engineering from Yonsei University, Seoul, Korea, in 1984 and the Ph.D. degree from Keio University, Tokyo, Japan, in 1991.

He is presently a Co-Principal Investigator at City College of The City University of New York. He is also Vice Director at the Institute of Bio-Cybernetics and a Research Associate in Department of Physiology at the University of Tokyo. His main areas of interest are integrative physiology in human and genetic targeted animal model and processing of biological signals and haemodynamic measurement instrumentation.

Dr. Ju is a member of Japan Society of Physiology, Autonomic Nervous System, BME, and SICE.

Sheng Lu received the B.S. degree from the Department of Automatic Control, University of Science and Technology of China, Hefei, China, in 1998. Currently, he is a Ph.D. degree candidate in the Department of Electrical Engineering, the City College of the City University of New York.

His research interests include signal processing, system modeling and artificial intelligence.

Ki H. Chon received the B.S. degree in electrical engineering from the University of Connecticut, Storrs; the M.S. degree in biomedical engineering from the University of Iowa, Iowa City; and the M.S. degree in electrical engineering and the Ph.D. degree in biomedical engineering from the University of Southern California, Los Angeles.

He spent three years as a National Institutes of Healyh (NIH) Post-Doctoral fellow at the Harvard-MIT Division of Health Science and Technology and one year as a Research Assistant Professor in the Department of Molecular Pharmacology, Physiology, and Biotechnology at Brown University, Providence, RI. Currently, he is an Assistant Professor in the Department of Electrical Engineering at the City College of the City University of New York. His current research interests include biomedical signal processing and identification and modeling of physiological systems.