Identification of Distributed Systems with Identical Subsystems

Mukhtar Ali [*] **Andrey P. Popov** [*] **Herbert Werner** [*]
Hossam S. Abbas [**]

[*] *Institute of Control Systems, Hamburg University of Technology,*
21073 Hamburg, Germany (e-mail:
mukhtar,andrey.popov,h.werner@tuhh.de).
[**] *Electrical Engineering Department, Faculty of Engineering, Assiut*
University, Assiut, Egypt (e-mail: h_abbas@aun.edu.eg).

Abstract: This article presents an identification technique for distributed systems with identical units using linear recurrent neural networks and exploiting the replicated structure of the units inside the system. The proposed method is applicable both to open-loop and closed-loop identification, takes into consideration boundary conditions and available information about the structure of the system, and is capable of identifying systems with heterogeneous units. The approach provides parameters estimate with minimum bias for unstable plant models when there is additive colored noise in the data. The method is described for two-dimensional systems (one for time and one for space), but is equally applicable to systems having more dimensions in space. The effectiveness of the method is demonstrated by two examples.

1. INTRODUCTION

In the recent years there has been a renewed interest in distributed control of complex engineering systems that are multi-dimensional and composed of identical subsystems which interact with each other. Examples of such systems include flexible structures [Wu and Yildizoglu, 2005], fluid and heat flow [Chughtai and Werner, 2008b], multi-agent systems [Olfati-Saber et al., 2007], satellite formations, systems characterized by the same class of partial differential equations, etc.

There exists a vast literature on controller synthesis for such distributed systems. A small sample of it includes distributed controller synthesis techniques for spatially invariant systems in Fourier domain [Bamieh et al., 2002], a separability-based method for systems with temporal and spatial dimensions [Gorinevsky et al., 2008] as well as various LMI techniques for spatially invariant systems [D'Andrea, 1999; D'Andrea and Dullerud, 2003; Chughtai and Werner, 2008a; Massioni and Verhaegen, 2009a; Popov and Werner, 2009].

However, in order to reliably apply any of the above synthesis techniques an accurate system model is needed. If accurate distributed models are not available from first principles it is important to have a method to identify such a model from measured data. Relatively few results are available in the literature for identification of two-dimensional (2-D) and multi-dimensional distributed systems. Identification of transfer function models of 2-D causal systems is presented in [Chen and Kao, 1979], while identification of 2-D and multi-dimensional non-causal rational transfer functions is given in [Arun et al., 1987], [Krogmeier and Arun, 1989]. However the latter ones are difficult to apply due to their dependence on the impulse response of the system. State-space-based identification

methods for distributed interconnected systems have been proposed recently in [Massioni and Verhaegen, 2008] and [Massioni and Verhaegen, 2009b], but these apply only to a special class of distributed systems, namely distributed decomposable systems or circulant systems. Moreover, extending all these methods to identify systems with noisy data or identifying them in closed-loop is questionable.

A more general least-squares-based identification method for 2-D open-loop systems has been proposed in [Ali et al., 2009]. For identification of unstable systems a closed-loop identification method is needed [Ljung, 1999]. The indirect methods proposed in [Li et al., 2006] and [d. Hof and d. Callafon, 1996] do not suffer from bias due to noise correlated with the input signal, as the input signal for identification is taken to be an external reference signal, but they need to impose restrictive assumptions on the plant and controller in order to separate them. Finally a linear recurrent neural networks (LRNN) based method has been proposed in [Ali et al., 2010a]. The method identifies each subsystem separately and then uses iterations until the parameters of all subsystems converge to the same values. Due to the use of LRNN the plant model can be directly extracted, but the method suffers from bias, since the interconnection signals from neighbors are considered as external signals and obtained from noisy measured data.

A way to avoid the above mentioned problems is to identify the distributed system as a whole, where the available knowledge of the system is used to define the connections between the units/subsystems. In this case the interconnecting signals will be estimated internally and will be not corrupted by measurement noise. However, by doing so the identified system will not necessarily be a connection of identical units, as during the neural networks (NN) training process the weights will be adjusted in such

a way that the output error is minimized, but without requiring them to be identical.

In this paper we present a LRNN-based method for identification of distributed systems that have identical subsystems. The method uses the available knowledge of the system in that one defines the interconnection topology and the identical structure of the subsystems. Then a model for the complete system is identified under the constraint of identical units. Finally, the model of a single unit can be extracted from the obtained model of the system. The advantages of this approach are fourfold.

- The approach is applicable to both open-loop and closed-loop identification. The main advantage is the identification of unstable plant models by using the approach as indirect closed-loop identification tool.
- An accurate model is obtained, as the interconnection signals between the subsystems are not considered as external, noise corrupted, measured signals.
- Consistent identification can be provided under colored noise condition even without identifying a noise model.
- The method exploits the complete knowledge of the system, including boundary conditions and structural "anomalies" (e.g., heat propagation in a metal plate with a hole – the hole representing the anomaly in the otherwise uniform structure).
- The method allows identifying distributed and interconnected systems with heterogeneous subsystems (e.g., a supported beam, part of which is constructed of one material and other part of a different material).

In order to simplify the notation the paper discusses 2-D systems that have one temporal and one spatial dimension (e.g., supported beams, heat propagation in rods, etc.), but the proposed methods are valid also for systems with a higher number of spatial dimensions, as well as space invariant systems.

The rest of the paper is organized as follows. Section 2 offers a brief introduction to 2-D distributed systems. Section 3 presents the proposed LRNN identification approach and the required changes to the standard backpropagation training algorithm for NN. Section 4 illustrates the approach on both open-loop and closed-loop numerical examples. Finally conclusions are drawn in Section 5.

The following standard notation is used throughout the paper. I_n denotes the $n \times n$ identity matrix; $\mathbb{R}^{n \times m}$ is the set of $n \times m$ real matrices; A_{ij} denotes element (ij) of the matrix A.

2. PRELIMINARIES

Let $r(k,j)$ be the two-dimensional discrete-time input signal to a linear invariant 2-D data generating single input single output (SISO) system, as the one shown in Fig. 1 and Fig. 2, where k and j are independent variables representing temporal and spatial indices respectively. Then the output $y(k,j)$ of the system can be represented in difference equation form, as a linear combination of weighted input-output values [Glentis et al., 1994], as

$$y(k,j) = - \sum_{(i_a,i_b) \in M_0^y} a_{i_a,i_b} \, y(k - i_a, j - i_b)$$
$$+ \sum_{(i_a,i_b) \in M^r} b_{i_a,i_b} \, r(k - i_a, j - i_b) + v(k,j). \quad (1)$$

Here M^y and M^r denote support regions (masks) for output and input terms, respectively with $M_0^y = M^y \backslash (0,0)$ and $v(k,j)$ is colored noise (Filtered 2-D zero-mean white-noise with normal distribution). The support region is defined as a subset of the two-dimensional space in which the indices of the coefficients of input and output terms in the difference equation lie. A general support region for 2-D systems lies in the 2-D plane. For causal systems the support region is a subset of first quadrant and for semi-causal (causal in time, but non-causal in space) the support region is a subset of the right half plane. For the details of masks description see [Ali et al., 2010b].
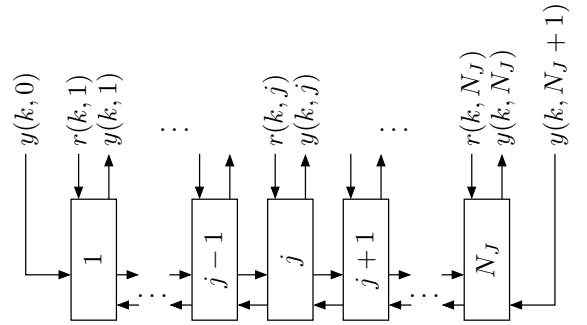

Figure 1. Interconnection of a 2-D system with J units

In 2-D, equation (1) and Fig. 2 the noise free process can be represented in a transfer function form as

$$G(q,p) = \frac{B(q,p)}{A(q,p)}, \quad (2)$$

with

$$B(q,p) = \sum_{(i_a,i_b) \in M^r} b_{i_a,i_b} q^{-i_a} p^{-i_b} \quad (3)$$

$$A(q,p) = 1 + \sum_{(i_a,i_b) \in M_0^y} a_{i_a,i_b} q^{-i_a} p^{-i_b}, \quad (4)$$

where q and p are forward shift operators in the first and second dimension, respectively (i.e., time and space).

Let θ represent the vector of model parameters, which is constructed by stacking columnwise $a_{i_a,i_b} \, \forall (i_a, i_b) \in M_0^y$, $b_{i_a,i_b} \, \forall (i_a, i_b) \in M^u$, and $\hat{y}(k,j,\theta)$ is the predicted output. Then the prediction error is

$$\epsilon(k,j,\theta) = y(k,j) - \hat{y}(k,j,\theta).$$

In Fig. 2 $H(q,p)$ is 2-D linear stable filter having monic numerator and denominator polynomials with specific masks. The problem of identifying a model for the system can then be stated as follows:

Given a set of input and output data

$$Z^{N_K \times N_J} = \left\{ [r(k,j), y(k,j)], \quad \begin{matrix} k = 1, \ldots, N_K \\ j = 1, \ldots, N_J \end{matrix} \right\}$$

estimate a parameter vector θ by minimizing the quadratic criterion

$$V(Z, \theta) = \frac{1}{2N_K N_J} \sum_{j=1}^{N_J} \sum_{k=1}^{N_K} \epsilon^2(k, j, \theta). \qquad (5)$$

This scheme is a prediction error method (PEM), [Ljung, 1999], as the objective is to minimize a sum over a norm of the prediction error.
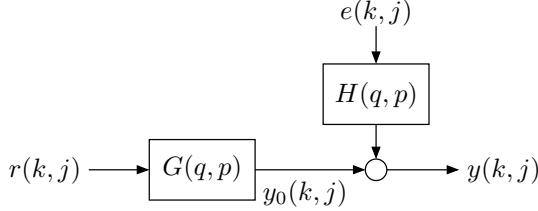

Figure 2. Two-dimensional data generating system

3. IDENTIFICATION OF 2-D INTERCONNECTED SYSTEMS

In this section we present a method to identify 2-D distributed systems based on linear recurrent neural networks (LRNN). The method has the advantage that an unbiased parameter estimate is obtained even if there is additive output colored noise in the data. By using LRNN we are solving the problem as pseudo-linear [Ljung, 1999] as the regressor contains terms which depend on the parameter to be estimated. We also discuss the consistency of the approach in this section. A further motivation to use LRNN is the fact that it offers an indirect closed-loop identification of 2-D distributed systems – the plant and controller models are obtained separately. Furthermore there is no restriction on the controller structure.

We propose a method for training a LRNN with identical weights for several layers, in a closed-loop identification framework where open-loop identification can be viewed as a special case. Hence, assume that each unit $j = 1, \ldots, N_J$ is controlled/stabilized locally and the controller coefficients are fixed and known. A reference input $r(k, j)$ is applied to the system and the outputs $y(k, j)$ are measured. The system interconnection is shown in Fig.1. Fixing the signals $y(k, 0)$ and $y(k, N_J + 1)$ represents boundary conditions to the system.

Under the assumption of linearity, a single unit/subsystem of the system could be modeled via LRNN, as shown in Fig. 3, where for simplicity a proportional controller (only static gains) is shown. Clearly, a controller with higher order temporal or/and spatial dynamics, as well as 3-D and higher dimensional systems, can also be modeled in a similar manner. In case that there is no local controller present the controller block will be empty and reference signal r will be entering the system directly (open-loop case). Given the input signal $r(k, j)$ and output signal $y(k, j)$, $k = 1, \ldots, N_K$, $j = 1, \ldots, N_J$, one could construct a LRNN network corresponding to the interconnection in Fig. 1, where each unit is as in Fig. 3. A successive training of the network will adapt the weights $b_{i_b,0}^j$, $a_{i_a,+1}^j$, $a_{i_a,-1}^j$ and $a_{i_f,0}^j$, $\forall i_b \in \{0, \ldots, n_b\}$, $i_a \in \{0, \ldots, n_a\}$, $i_f \in \{1, \ldots, n_a\}$, (that is, all weights with a white arrow) in order to minimize the performance index (5). Once the training of these $M = N_J(3n_a + n_b + 3)$ weights is

completed, a model of the complete system is obtained. As, in the general, the training is performed with imperfect measurements the optimal coefficients of the N_J units are different. In other words, one has obtained a model, which might provide a good description of the overall system, but from which no unique model of a single unit can be extracted.

This problem can be remedied by exploiting the identical nature of the units during the training process. In order to do so, one has to impose that the corresponding weights in the layers of the N_J units are indeed identical, i.e., $\forall j \in \{1, \ldots, N_J\}$, the following should hold

$$b_{i_b,0}^j = b_{i_b,0}, \qquad \forall i_b \in \{0, \ldots, n_b\}, \qquad (6)$$
$$a_{i_a,+1}^j = a_{i_a,+1}, \qquad \forall i_a \in \{0, \ldots, n_a\}, \qquad (7)$$
$$a_{i_f,0}^j = a_{i_f,0}, \qquad \forall i_f \in \{1, \ldots, n_a\}, \qquad (8)$$
$$a_{i_a,-1}^j = a_{i_a,-1}, \qquad \forall i_a \in \{0, \ldots, n_a\}, \qquad (9)$$

Here for simplicity of notation it has been assumed that the weights corresponding to the boundary input channels are the same as the weights for the inputs from the corresponding neighbor. Hence, now the goal is to train only these $m = 3n_a + n_b + 3$ weights. Note, that in the multivariable case these weights of the NN are matrices,
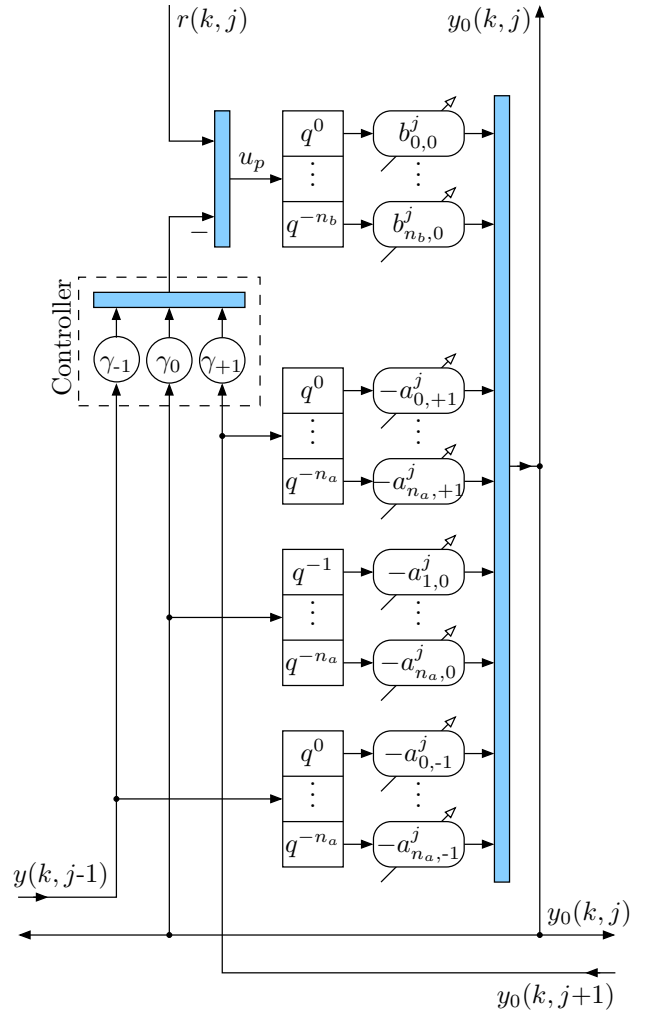

Figure 3. Closed-loop LRNN of unit j of a 2-D system

but as this would unnecessarily complicate our notations here we derive our result in the case of scalar signals.

Let $\Theta \in \mathbb{R}^{M \times 1}$ denote a vector of the stacked weights subject to training of the complete closed-loop interconnection, i.e., excluding weights that are not subject to training, such as controller gains. Let, further, $\theta \in \mathbb{R}^{m \times 1}$, denote a vector of the stacked unique weights in the closed-loop interconnection. Then equations (6)–(9) can be imposed, by writing

$$\Theta = P\theta, \qquad (10)$$

where $P \in \mathbb{R}^{M \times m}$ is a matrix, each row of which has a single identity element and all other elements zero. The position of the identity element is such that the correct element of θ is substituted in Θ, according to (6)–(9). Formally this can be stated as: each row of P is equivalent to e_h for some h, where e_h is row h of I_m. For example in Fig. 1 and Fig. 3 this equality reads

$$\underbrace{\begin{bmatrix} b_{0,0}^1 \\ b_{1,0}^1 \\ \vdots \\ b_{n_b,0}^1 \\ a_{0,+1}^1 \\ \vdots \\ a_{n_a,-1}^1 \\ b_{0,0}^2 \\ \vdots \\ b_{0,0}^{N_J} \\ \vdots \\ a_{n_a,-1}^{N_J} \end{bmatrix}}_{\Theta} = \underbrace{\begin{bmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 1 & \dots & 0 \\ \vdots & & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 1 \\ 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 1 \end{bmatrix}}_{P} \underbrace{\begin{bmatrix} b_{0,0} \\ b_{1,0} \\ \vdots \\ b_{n_b,0} \\ a_{0,+1} \\ \vdots \\ a_{n_a,-1} \end{bmatrix}}_{\theta}.$$

It should be noted that in general the structure of P is not as simple as the one for the example presented here. However in all cases P has full column rank, as each element of θ is involved in at least one element of Θ.

Now, the true weights to be trained are the elements of θ. For any particular θ one can obtain the corresponding weights of the complete LRNN by substituting in (10). Then the training process of the unconstrained LRNN (see, e.g., [Hagan and Menhaj, 1994], [Narendra and Parthasarathy, 1990] and [Narendra and Parthasarathy, 1991]) can readily compute the gradient of (5) w.r.t. the elements of Θ and the necessary update step $d\Theta$ of Θ. The following lemma provides the necessary update $d\theta$ of the trained coefficients.

Lemma 1. If the update step $d\Theta$ of $\Theta = P\theta$ is known, then the update step $d\theta$ is

$$d\theta = (P^\top P)^{-1} P^\top d\Theta. \qquad (11)$$

Proof. Recall (see, e.g., [Petersen and Pedersen, 2008]) that the derivative of Θ is

$$d\Theta = (\nabla_\theta \Theta)\, d\theta,$$

where $\nabla_\theta \Theta = P$ is the gradient of Θ w.r.t. θ. Then since P has a full column rank it follows that $(P^\top P) \in \mathbb{R}^{m \times m}$ is invertible. Applying the Moor-Penrose inverse leads to the desired result.
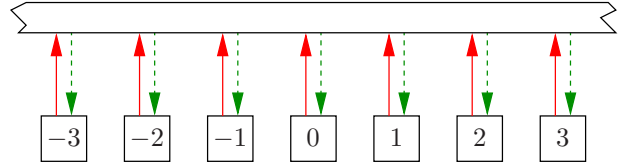

Figure 4. Portion of a rod. Dashed arrows show sensor signals, solid lines show actuator signals to the segments.

4. ILLUSTRATIVE EXAMPLES

The presented approach is illustrated by two simulation examples.

4.1 Example 1: Open-Loop Identification

As a practical example of the proposed approach, a heat conduction in a rod of length one meter is considered. This represents a distributed parameter system and has also been used in [Augusta et al., 2007]. Fig. 4 shows a schematic diagram of a part of the rod with an array of temperature sensors and actuators. The indices of the sensor-actuator units indicate an offset w.r.t the center of the rod and are meant to emphasize that the system is non-causal in space. The system is described by the heat equation

$$\frac{\partial u(t,x)}{\partial t} = \kappa \frac{\partial^2 u(t,x)}{\partial x^2} + Q(t,x), \qquad (12)$$

where $u(t,x)$ denotes temperature [°C]; t and x denote time [s] and space [m] coordinates, respectively; κ is a constant [m²s⁻¹] and Q is the input heat energy [°Cs⁻¹].

Equation (12) is discretized using the central difference method to approximate the partial derivatives as

$$\left(\frac{\partial u(t,x)}{\partial t}\right)_{k,j} = \frac{u(k+1,j) - u(k,j)}{T},$$

$$\left(\frac{\partial^2 u(t,x)}{\partial x^2}\right)_{k,j} = \frac{u(k,j-1) - 2u(k,j) + u(k,j+1)}{h^2},$$

where T is the sampling time and h is the spatial sampling distance between two nodes along the rod. Then at instance (k,j) equation (12) can be approximated as

$$u(k+1,j) = \frac{T}{h^2}u(k,j-1) + (1 - 2\frac{T}{h^2})u(k,j) + \frac{T}{h^2}u(k,j+1) + Q(k,j) \quad (13)$$

Here we are assuming that $\kappa = 1$ and the input is $TQ(k,j)$. The difference equation (13) can be represented in transfer function form as

$$G(q,p) = \frac{b_{1,0}q^{-1}}{1 + a_{1,-1}q^{-1}p + a_{1,0}q^{-1} + a_{1,1}q^{-1}p^{-1}}, \quad (14)$$

with

$$b_{1,0} = 1,$$

$$a_{1,-1} = -\frac{T}{h^2}, \quad a_{1,0} = -1 + 2\frac{T}{h^2}, \quad a_{1,1} = -\frac{T}{h^2}.$$

We assume that the rod is divided spatially into $N_J = 9$ nodes, i.e., $h = 1/9$, and that $T = 0.005$ s. Noisy data is generated from this system by considering a structure as in Fig. 2 with $H(q,p) = 1$, where $y(k,j)$ represents noisy output. The input $r(k,j)$ and $e(k,j)$ are taken as 2-D zero-mean normally distributed white-noises. For different signal-to-noise ratios (SNR), Monte-Carlo simulations

were carried out with 100 samples. In each simulation run $N_K = 1000$ measurements were made, i.e., $Z^{N_K \times N_J} \in \mathbb{R}^{1000 \times 9}$. For each measurement set the identification was carried out both with the proposed approach and the one presented in [Ali et al., 2010a], based on iterative algorithm based on linear recurrent neural networks (IA-LRNN).

The obtained results are reported in Tables 1 and 2, where the bias and variance norms are defined as $\|\theta_0 - E[\hat{\theta}]\|_2$ and $\|E[(\hat{\theta} - E[\hat{\theta}])^2]\|_2$, respectively, where θ_0 is the true parameter vector and $\hat{\theta}$ is the estimated one. One can easily see, both from the numerical values of the coefficients in Table 1 and the bias norm in Table 2, that the proposed approach obtains values closer to the true parameters than the iterative algorithm based approach, by achieving similar variance of the results.

Table 1. Mean of the estimated parameters at SNR = 10 dB for example 1.

Parameter	True values	Estimated values	
		IA-LRNN	Proposed
$b_{1,0}$	1.0000	0.9940	0.9997
$a_{1,-1}$	−0.4050	−0.3724	−0.4048
$a_{1,0}$	−0.1900	−0.2222	−0.1909
$a_{1,1}$	−0.4050	−0.3745	−0.4045

Table 2. Estimation bias and variance norm at different SNR for example 1.

SNR dB	Bias Norm		Variance Norm	
	IA-LRNN	Proposed	IA-LRNN	Proposed
15	0.0247	0.0007	1.5×10^{-5}	1.8×10^{-5}
10	0.0554	0.0011	6.4×10^{-6}	1.1×10^{-5}

The measured noisy output, noise-free output and simulated output of the identified model for unit $j = 5$ are shown in Fig. 5.
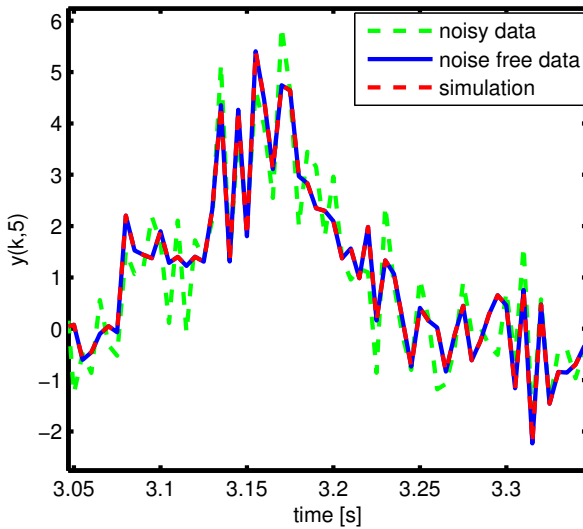

Figure 5. Example 1:Response at $j = 5$ for $SNR = 10$ dB

4.2 Example 2: Closed-Loop Identification

Our second example considers the closed-loop identification of an unstable spatially distributed system. For the purpose of illustration we take a simple model with appropriately chosen coefficients, that result in an open-loop unstable system. The system dynamics are described by

$$G(q,p) = \frac{b_{1,0}q^{-1}}{1 + a_{1,-1}q^{-1}p + a_{1,0}q^{-1} + a_{1,1}q^{-1}p^{-1}}, \quad (15)$$

where $N_J = 9$, $T = 0.005$ s, and the selected coefficient values are

$$b_{1,0} = 1$$
$$a_{1,-1} = 0.595, \quad a_{1,0} = 0.81, \quad a_{1,1} = 0.595.$$

In order to stabilize the system a simple, spatially distributed controller is selected,

$$C(q,p) = \gamma_{-1}p + \gamma_0 + \gamma_1 p^{-1}, \quad (16)$$

where $\gamma_{-1} = -0.75$, $\gamma_0 = -0.75$ and $\gamma_1 = -0.75$. The closed loop system with the above plant and controller is shown in Fig. 6. The measurement data is generated by exciting the system with 2-D random white noise with normal distribution. Thus the output signal is corrupted with a $v(k, j)$ which is filtered 2-D white-noise $e(k, j)$ with normal distribution. The measurements are made with $N_J = 9$ and $N_K = 1000$.
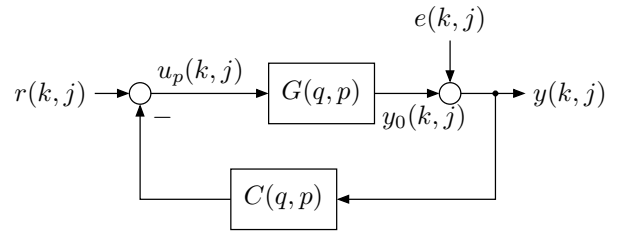

Figure 6. Two-dimensional closed-loop system

A LRNN as shown in Fig. 3 is then constructed. The network is trained using the modified back-propagation algorithm discussed in Section 3, where only the plant weights are trained. As before, Monte-Carlo simulations with 100 runs are performed and the results are presented in Tables 3 and 4.

Table 3. Mean of the estimated parameters at SNR = 10dB for example 2.

Parameter	True values	Estimated values	
		IA-LRNN	Proposed
$b_{1,0}$	1.0000	0.9997	1.0000
$a_{1,-1}$	0.5950	0.6165	0.5949
$a_{1,0}$	0.8100	0.8107	0.8102
$a_{1,1}$	0.5950	0.6142	0.5959

5. CONCLUSIONS

In this paper we have presented a method for the identification of transfer function models for distributed systems composed of identical subsystems. The approach is based on structured LRNN and exploits the replicated nature of

Table 4. Estimation bias and variance norm at different SNR for example 2.

SNR	Bias Norm		Variance Norm	
dB	IA-LRNN	Proposed	IA-LRNN	Proposed
15	0.0101	0.0005	3.5×10^{-6}	1.1×10^{-5}
10	0.0288	0.0009	1.1×10^{-5}	3.4×10^{-5}

the subsystems. The method allows open-loop and closed-loop identification, provides an unbiased parameter estimate as compared to the existing methods and allows identification of systems with heterogeneous units and non-uniform structure, while at the same time naturally providing a way for handling boundary conditions. The main contribution of this note is that it can identify unstable plant model with minimum bias in the parameters estimate. Two numerical examples, one in open-loop and one in closed-loop, illustrate the approach and demonstrated its superiority over an iterative algorithm based on LRNN.

Future application of the proposed method will include the identification of input-output form models of fluid flows with Reynolds number greater than 5800. Such flows are unstable spatially distributed system [Bewley and Liu, 1998] the control of which is an active area of research (see, e.g., [Chughtai and Werner, 2008a]).

REFERENCES

Ali, M., Abbas, H., Chughtai, S., and Werner, H. (2010a). Identification of spatially interconnected systems using neural networks. In *Proc. of 49th IEEE Conference on Decision and Control*, 6938–6943.

Ali, M., Chughtai, S., and Werner, H. (2009). Identification of spatially interconnected systems. In *Proc. of 48th IEEE Conference on Decision and Control*, 7163–7168. Shanghai, China.

Ali, M., Chughtai, S., and Werner, H. (2010b). Consistent identification of two-dimensional systems. In *Proc. of American Control Conference*, 3464–3469. Baltimore, USA.

Arun, K., Krogmeier, J., and Potter, L. (1987). Identification of 2-d noncausal systems. In *Proc. of IEEE Conference on Decision and Control*. Los Angeles, CA.

Augusta, P., Hurak, Z., and Rogers, E. (2007). An algebraic approach to the control of spatially distributed systems- the 2-d systems case with a physical application. In *Proc. of the 3rd IFAC Symposium on Systems, Structure and Control*. Brazil.

Bamieh, B., Paganini, F., and Dahleh, M. (2002). Distributed control of spatially invariant systems. *IEEE Transactions on Automatic Control*, 47(7), 1091–1107.

Bewley, T. and Liu, S. (1998). Optimal and robust control and estimation of linear paths to transition. *Journal of Fluid Mechanics*, 365(12), 305–349.

Chen, C. and Kao, Y. (1979). Identification of two-dimensional transfer function from finite input-output data. *IEEE Transactions on Automatic Control*, 24(5), 748–752.

Chughtai, S. and Werner, H. (2008a). Distributed control for a class of spatially interconnected discrete-time systems. In *Proc. of 17th IFAC World Congress*, 7761–7766. Seoul, South Korea.

Chughtai, S. and Werner, H. (2008b). Transition control of plane Poiseuille flow - a spatially interconnected model. In *Proc. of 47th IEEE Conference on Decision and Control*. Cancun, Mexico.

d. Hof, P.M.V. and d. Callafon, R.A. (1996). Multivariable closed-loop identification: From indirect identification to dual-youla parametrization. In *Proc. of the 35th IEEE Conference on Decision and Control*.

D'Andrea, R. (1999). Linear matrix enequalities, multidimensional system optimization, and control of spatially distributed systems an example. In *Proc. of American Control Conference*, 2713–2717.

D'Andrea, R. and Dullerud, G.E. (2003). Distributed control design for spatially interconnected systems. *IEEE Transactions on Automatic Control*, 48(9), 1478–1495.

Glentis, G., Slump, C., and Herrmann, O. (1994). Efficient two-dimensional ARX modeling. In *Proc. of International Conference on Image Processing (ICIP)*. 605-609.

Gorinevsky, D., Boyd, S., and Stein, G. (2008). Design of low-bandwidth spatially distributed feedback. *IEEE Transactions on Automatic Control*, 53(1), 257–272.

Hagan, M. and Menhaj, M. (1994). Training feedforward networks with the mardurdt algorithm. *IEEE Transactions on Neural networks*, 5(6), 989–993.

Krogmeier, J. and Arun, K. (1989). A comparative study of causal and non-causal models for multidimensional spectrum estimation. In *Maple Press*.

Li, P., Kruger, U., and Irwin, G.W. (2006). Identification of dynamic systems under closed-loop control. *Intern. Journ. of Systems and Science*, 37(3), 181–195.

Ljung, L. (1999). *System Identification, Theory for the User*. Prentice-Hall Inc. USA, 2nd edition edition.

Massioni, P. and Verhaegen, M. (2008). Subspace identification of circulent systems. *Automatica*, 44(11), 2825–2833.

Massioni, P. and Verhaegen, M. (2009a). Distributed control for identical dynamically coupled systems: A decomposition approach. *IEEE Transactions on Automatic Control*, 54(1), 124–135.

Massioni, P. and Verhaegen, M. (2009b). Subspace identification of distributed, decomposible systems. *in Proc. of 48th IEEE Conference on Decision and Control*, 3364–3369.

Narendra, K. and Parthasarathy, K. (1990). Identification and control of dynamical systems using neural networks. *IEEE Transactions on Neural networks*, 1(1), 4–27.

Narendra, K. and Parthasarathy, K. (1991). Gradient methods for the optimization of dynamical systems containing neural networks. *IEEE Transactions on Neural Networks*, 2(2), 252–262.

Olfati-Saber, R., Fax, J., and Murray, R. (2007). Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, 1(95), 215–233.

Petersen, K.B. and Pedersen, M.S. (2008). *The Matrix Cookbook*. URL http://matrixcookbook.com/.

Popov, A. and Werner, H. (2009). A robust control approach to formation control. In *Proc. of the European Control Conference*, 4428–4433. Budapest, Hungary.

Wu, F. and Yildizoglu, S.E. (2005). Distributed paramter dependent modelling and control of flexible structures. *Transactions of the AMSE*, 127.