Blind Identification of Fully Observed Discrete-Time Linear Time-Varying Systems via Sparse Recovery

Roel Dobbe\textsuperscript{a}, Stephan Liu\textsuperscript{a}, Ye Yuan\textsuperscript{b}, Claire Tomlin\textsuperscript{a}

\textsuperscript{a}Department of Electrical Engineering \& Computer Sciences, UC Berkeley, USA
\textsuperscript{b}School of Automation, Huazhong University of Science and Technology, People’s Republic of China

Abstract

Discrete-time linear time-varying (LTV) systems form a powerful class of models to approximate complex dynamical systems or networks with nonlinear dynamics for the purpose of analysis, design and control. Motivated by inference of spatio-temporal dynamics in breast cancer research, we propose a method to efficiently solve an identification problem for a specific class of discrete-time LTV systems, in which the states are fully observed and there is no access to system inputs. In addition, it is assumed that we do not know on which states the inputs act, which can change between time steps, and that the total number of inputs is sparse over all states and over time. The problem is formulated as a compressive sensing problem, which incorporates the effect of measurement and process noise and which has a solution with a partially sparse support. We derive sufficient conditions for the unique recovery of the system model and input values, which lead to practical conditions on the number of experiments and rank conditions on system outputs. Synthetic experiments validate the sensitivity to noise.

Key words: Blind Identification; Discrete-Time; Linear Time-Varying Systems; Networked Systems; Compressive Sensing; Network Inference; Experimental Design

1 Introduction

Many networked dynamical systems, such as power grids or biological systems, exhibit complex nonlinear dynamics. Unfortunately, the formulation of nonlinear system identification is generally hard or intractable unless system structure can be exploited or an efficient black box model structure is used to approximate the system’s dynamics [23]. Nonlinear system dynamics are therefore often approximated by piecewise-affine or discrete-time linear time-varying (LTV) models. In addition, efforts to identify a dynamical model for analysis or control design can be hampered for certain systems or applications, due to a lack of access to the inputs (or disturbances) entering the system. An example is detection and mitigation of malicious attacks on cyber-physical systems [25], or inferring temporal protein-protein interactions in gene regulatory networks (GRNs). The blind system identification problem (BSI) assumes that the values of inputs and disturbances are unknown and aims to retrieve these from outputs of the system dynamics [16].

This work is primarily motivated by inference problems in breast cancer biology that aim to capture protein-protein interactions in gene regulatory networks (GRNs); whether these exist, how strong these are and how these change over time [15]. These effects tend to be nonlinear and time-varying. In earlier work, a discrete-time linear time-varying modeling structure was used to approximate these dynamics [9]. Here, we had access to measurements of concentrations for all proteins in the GRN, so that the output represents a fully observed state vector at every time instance. This assumption is increasingly practical in biological experiments due to rapid developments of measurement technologies, such as Reverse Phase Protein Array (RPPA), which allow high sensitivity and sample throughput of protein level measurements at a reasonable cost per sample [3]. In such settings, the system input, which is the effect that a drug or a mutation has on a GRN, is typically unknown. However, it is generally true that the input is sparse; it affects a relatively small number of states during a small number of time steps. Improving our understanding of how drugs and mutations affect a signaling network is critical for effective and personalized treatment design. As such, the aim of this work is to retrieve both the input effects of drugs and the interaction dynamics between proteins and how these change over time.
Prior Work

BSI is known to be a difficult problem that is generally ill-posed. BSI of time-varying systems is known as a challenging problem, as compared to time-invariant systems. It is well known that in order to reliably retrieve the input and/or system parameters, further information about the system is needed [16]. Hence, all prior works in different areas all impose a certain structure that is rich enough to represent complex system dynamics and simple enough to allow for identification.

Originally, blind identification was well studied for Finite Impulse Response (FIR) systems, in which the filter represents the system’s dynamics. For an extensive overview of BSI for time-invariant systems, the reader is referred to [1,16]. More recently, the advent of ubiquitous sensing and data collection has spurred new efforts to perform BSI for larger-scale multi-input, multi-output systems [24,27]. The BSI literature for systems with time-varying dynamics is sparse. Typically, these systems are modeled as an extension of time-invariant FIR filters using a basis expansion approach [30,31,13]. These results are solely for single-input, multi-output systems. In addition, if a system has arbitrary internal feedback dynamics, a FIR model cannot be used and one requires infinite impulse response (IIR) models. Unfortunately, IIR systems driven by unknown inputs are inherently not identifiable [16]. In control theory, identification methods have been proposed for certain classes of time-varying systems that restrict to certain structures and parameter changes. In [20], a discrete-time LTV state space model is identified, assuming no input and stable dynamics. A widely studied approach is that of Linear Parameter-Varying (LPV) systems, for which identification procedures are proposed by [19], [32] and [4], and for which [29] provides a broad and rigorous overview. Most of these approaches consider continuous-time systems, making the methods more generally applicable in terms of approximating nonlinear models, but much harder to scale [29]. None of these time-varying methods consider scenarios with unknown inputs.

In network inference, sparse recovery theory has been applied and further developed to reconstruct networks from data exploiting the sparsity in network connectivity [22,34,35,14,7]. Most work in this area assumes a linear time-invariant (LTI) network model that governs the dynamic propagation of signals, often without any external inputs. [26] develops a method for inferring autoregressive models with exogenous inputs (ARX), in which the parameter vector changes a limited number of times. The use and identification of time-varying graphical models or stochastic networks are proposed in [2,17,18,33]. [6] gives a general overview of how time-varying graphs and dynamic networks are used in different fields and applications. In [2], a $\ell_1$-regularized logistic regression formalism is used to capture network structure and its changes over time. While scaling well to larger networks, the method does not consider the effect of external inputs. [18] introduces autoregressive time-varying models to describe and infer gene-regulation networks and infers the model using a Reversible Jump Markov Chain Monte Carlo procedure. This model class neatly encodes the time-varying dynamics with a LTV mapping, but does not consider the effect of external inputs.

Contributions

We propose a blind system identification method for discrete-time LTV dynamical systems with four important characteristics: multiple inputs, internal feedback dynamics, a fully observed state vector, and repeated experiments. The first two characteristics address relevant open challenges in the literature of blind identification of time-varying systems. The third characteristic is an assumption that is practical in the context of inferring GRNs, and, in some sense, represents the price to pay to overcome the complexity of the former two characteristics. In addition, we assume that experiments can be repeated multiple times, with the same time-varying dynamics, in order to collect sufficient data for identification. This is a fair assumption for biological studies, in which experiments can be replicated efficiently, but can form a challenge for other applications.

Notation

Denote by $A^\top$ the transpose of a matrix, and by $\text{vec}(\cdot) : \mathbb{R}^{m \times D} \rightarrow \mathbb{R}^{mD}$ the function that vectorizes a matrix column-by-column. The function $\|x\|_0$ (the “$l_0$-norm”) returns the number of nonzero entries in the vector $x$, which is said to be $s$-sparse if at most $s$ of its entries are nonzero: $\|x\|_0 \leq s$. We will use subscript $i = 1,\ldots,n$ to denote the $i$-th entry of a vector in $\mathbb{R}^n$.

2 Problem Formulation

In this section, we first represent an experimental data set as the evolution of a dynamical system using an LTV modeling framework. We then formulate the system identification and input retrieval problem in a sparse recovery framework. Further analysis of the problem then allows us to write it using a more general compressive sensing formulation, allowing for rigorous theoretical analysis later on.

The problem is formulated as an experimental design problem, with the aim to understand necessary and sufficient conditions on perturbations and collected output measurements of system dynamics that guarantee successful inference of parameters related to system dynamics and inputs/disturbances. Consider a series of $q$ experiments, representing different perturbations and samples of the system state $z \in \mathbb{R}^n$. State output measurements are taken at $k_f$ moments, not necessarily equally
sampled through time, but at fixed instances for all experiments. The dynamics of the LTV system during an experiment $j$ are modeled as:

$$z^{(j)}[k+1] = A[k]z^{(j)}[k] + u^{(j)}[k] + w^{(j)}[k], \quad j \in \{1, \ldots, q\}, \quad k \in \{0, \ldots, k_f-1\},$$  

where $z^{(j)}[k], u^{(j)}[k], w^{(j)}[k] \in \mathbb{R}^n$ are the state vector, input vector and noise vector of a single experiment $j$, and $A[k] \in \mathbb{R}^{n \times n}$ is a matrix describing the dynamical interactions between the state variables for the transition from time $k$ to $k+1$. The matrices $A[k]$ are constant across experiments. The state vector $z^{(j)}[k]$ is assumed to be measured at all $k \in \{0, \ldots, k_f\}$. We consider scenarios for which the noise is bounded by $\|w^{(j)}[k]\|_2 \leq \eta_j, \forall j, k$. Note that we allow for the same number of inputs as states - these are both $n$-dimensional. We assume that the inputs are sparse over all $q$ experiments, that is out of all $nk_f q$ input values only $s < nk_f q$ are nonzero. By penalizing sparsity in our problem formulation, we will ensure that the number of nonzero inputs is properly low. The central questions of this work are, given the collection of dynamic output data $z^{(j)}[k]$, how and under what conditions can we correctly infer the parameters in (1): $A[k]$, the sensing matrix and the unknown inputs $u^{(j)}[k]$ (or $s$ values) for $j \in \{1, \ldots, q\}, k \in \{0, \ldots, k_f-1\}$?

For each time step $k$, we stack all our experiments together into matrices for states, inputs and noise vectors:

$$Z_k = [z^{(1)}[k] \ldots z^{(q)}[k]] \in \mathbb{R}^{n \times q},$$  
$$U_k = [u^{(1)}[k] \ldots u^{(q)}[k]] \in \mathbb{R}^{n \times q},$$  
$$W_k = [w^{(1)}[k] \ldots w^{(q)}[k]] \in \mathbb{R}^{n \times q}.$$  

We further organize the data and variables as:

$$Z \triangleq \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_{k_f} \end{bmatrix}, \quad U \triangleq \begin{bmatrix} U_0 \\ U_1 \\ \vdots \\ U_{k_f-1} \end{bmatrix}, \quad W \triangleq \begin{bmatrix} W_0 \\ W_1 \\ \vdots \\ W_{k_f-1} \end{bmatrix}.$$  

We then vectorize $Z$, $U$, $W$ and $A[k]$ as follows: $z = \text{vec}(Z) \in \mathbb{R}^{nk_f q}$, $u = \text{vec}(U) \in \mathbb{R}^{nk_f q}$, $w = \text{vec}(W) \in \mathbb{R}^{nk_f q}$ and $a = \text{vec}(A[k]) \in \mathbb{R}^{nk_f q}$ for all $k = 0, \ldots, k_f-1$ (or $\mathbb{R}^n$ for LTI models). The $i$-th scalar output in (1) can be rewritten as

$$z^{(j)}_i[k+1] = \sum_{l=1}^{k} a_{il}[k]z^{(j)}_l[k] + u^{(j)}_i[k] + w^{(j)}_i[k], \quad i = 1, \ldots, n, \quad j = 1, \ldots, q, \quad k = 0, \ldots, k_f.$$  

for $i = 1, \ldots, n$, $j = 1, \ldots, q$ and $k = 0, \ldots, k_f$. Here, $a_{il}[k]$ denotes the entry in $A[k]$ on the $i$-th row and $l$-th column. Equivalently, the dynamics of one experiment can be formulated as

$$z^{(j)}[k+1] = \begin{bmatrix} 0_{n \times n(k_f(k_f-j)+k-1)} \mid I_n \mid 0_{n \times n(k_f(j-1)+k-1)} \end{bmatrix} \begin{bmatrix} u \end{bmatrix} + \begin{bmatrix} 0_{n \times n} \mid (z^{(j)}[k])^\top \mid 0_{n \times n} \end{bmatrix} a + w^{(j)}[k].$$  

Or in short

$$z^{(j)}[k+1] = \begin{bmatrix} \psi_u^{(j)}[k] \mid \psi_a^{(j)}[k] \end{bmatrix} \begin{bmatrix} u \\ a \end{bmatrix} + w^{(j)}[k].$$  

Here, $\otimes$ denotes the Kronecker product. By stacking this equation vertically for all time steps $k = 1, \ldots, k_f$ and experiments $j = 1, \ldots, q$ we can construct

$$z = [\Psi_u \mid \Psi_a] \begin{bmatrix} u \\ a \end{bmatrix} + w = \Psi \begin{bmatrix} u \\ a \end{bmatrix} + w,$$

where $\Psi_u = I_{nk_f q}$ and

$$\Psi_a = \begin{bmatrix} \text{blkdiag}(I_n \otimes (z^{(1)})^\top[k]) \\ \text{blkdiag}(I_n \otimes (z^{(2)})^\top[k]) \\ \vdots \\ \text{blkdiag}(I_n \otimes (z^{(q)})^\top[k]) \end{bmatrix} \in \mathbb{R}^{nk_f q \times n^2 k_f}.$$  

Here, for each experiment $j = 1, \ldots, q$, $\text{blkdiag}(\cdot)$ constructs a block-diagonal matrix with blocks $I_n \otimes (z^{(j)})^\top[k]$ for $k = 0, \ldots, k_f-1$. Note that in the case of an LTI system, the block diagonal structure collapses, resulting in

$$\Psi \in \mathbb{R}^{nk_f q \times (nk_f q + n^2 k_f)}$$

$\Psi$ denotes the sensing matrix and $w \in \mathbb{R}^{nk_f q}$ a vector with stacked measurement or process noise values. By exploiting prior knowledge about the statistics of $w$, we can determine a bound on $\ell_2$-norm:

$$\|z - \Psi_u u - \Psi_a a\|_2 \leq \eta,$$

where $\eta$ can be formulated as

$$\|z - \Psi_u u - \Psi_a a\|_2 \leq \eta.$$
A compressive sensing formulation for inferring the unknowns \((u, a)\) with noisy measurements now reads

\[
\min_{u, a} \|u\|_1, \quad \text{subject to} \quad \|z - \Psi_u u - \Psi_a a\|_2 \leq \eta. \tag{11}
\]

This formulation attempts to find a partially sparse solution in which sparsity is only enforced on \(u\), and not necessarily on \(a\). Note that some applications could yield sparsity in \(a\), which can be easily addressed in this framework by adding the \(\ell_1\)-norm of \(a\), yielding a compressive sensing problem with a fully sparse support.

### 3 Analysis

Consider the measurement equation

\[
y = \Psi x + w, \tag{12}
\]

where \(y \in \mathbb{R}^m, x \in \mathbb{R}^D\), and \(\|w\|_2 \leq \eta\) is a bounded noise signal. In general, \(m < D\) yields an underdetermined system of equations, with an infinite number of solutions. It turns out that if the sensing matrix \(\Psi\) adheres to certain conditions, \(x\) can be retrieved exactly from far fewer measurements (i.e. \(m \ll D\)) than asserted by the Nyquist sampling theorem [5]. The sparsest solution to the underdetermined system of equations \(y = \Psi x\) can be found by solving:

\[
\min_{x \in \mathbb{R}^D} \|x\|_0, \quad \text{subject to} \quad \|y - \Psi x\|_2 \leq \eta. \tag{13}
\]

**Lemma 3.1** (Unique retrieval of true signal [10, Lemma 2.1]) If the sparsest solution to (13) has \(\|x\|_0 = s\) and \(D \geq 2s\) and all subsets of \(2s\) columns of \(\Psi\) are full rank, then this solution is unique.

In general, (13) is a NP-hard optimization problem that is both combinatorial and non-convex, and hence impractical to solve. In contrast, the Basis Pursuit method [8] solves the convex relaxation of (13) efficiently,

\[
\min_{x \in \mathbb{R}^D} \|x\|_1, \quad \text{subject to} \quad \|y - \Psi x\|_2 \leq \eta. \tag{14}
\]

**Definition 3.1** The spark of a matrix \(\Psi\) is the smallest number of columns of \(\Psi\) that are linearly dependent [11], which is upper bounded by \(\text{rank}(\Psi) + 1\).

Given this definition, the following lemma [5] provides sufficient conditions for equivalence between the compressed sensing problem (13) and its convex relaxation (14).

**Lemma 3.2** (Spark Equivalence Condition [11]) For the system of linear equations \(\Psi x = y\) (\(\Psi \in \mathbb{R}^{m \times D}\) full-rank with \(m < D\)), if a solution \(x\) exists obeying

\[
\|x\|_0 \leq \frac{1}{2} \text{spark}(\Psi),
\]

that solution is both the unique solution to the convex relaxation (14), and the unique solution to the original NP-hard compressive sensing problem (13).

Returning to our central problem (11), the support of the solution is partially sparse due to \(a\) being a potentially dense vector. We determine under what conditions, the true solution \((u, a)\) can be retrieved from some measurements \(z\). Let \(s_u = \|\tilde{u}\|_0\) denote the number of nonzero entries of \(\tilde{u} \in \mathbb{R}^{nk_f q}\). Denote the fraction of nonzero entries in \(\tilde{u}\) as

\[
\rho_u \triangleq \frac{s_u}{nk_f q}.
\]

**Theorem 3.1** Suppose the sparsest solution to (11) has \(\|\tilde{a}\|_0 = n^2 f\), and \(\|\tilde{u}\|_0 = s_u\) with \(\rho_u \leq \frac{1}{2}\). If \(\Psi_a\) is full column rank, then the solution to (11) is unique and equivalent to the solution of the NP-hard \(\ell_0\)-problem.

**Proof 3.1** First note that as \(\tilde{a}\) is a dense vector, we require \(\Psi_a\) to be full column rank in order for all information in \(\tilde{a}\) to be maintained. Now assume \(\tilde{a}\) is known, define the vector \(\tilde{z} = z - \Psi_a \tilde{a}\), and rewrite (11) as

\[
\min_{u} \|u\|_1, \quad \text{s.t.} \quad \|\tilde{z} - \Psi_u u\|_2 \leq \eta. \tag{17}
\]

As \(\Psi_u = I_{nk_f q}\) (identity matrix) is full rank and \(\rho_u \leq \frac{1}{2}\), all subsets of \(2s_u\) columns of \(\Psi_u\) are also full rank. Hence, following Lemma 3.1, any \((s_u + nk_f q)\)-sparse solution to the \(NP\)-hard \(\ell_0\)-problem is unique. In addition, the same full-rank condition yields that \(\text{spark}(\Psi_u) = nk_f q + 1\). With \(\|\tilde{u}\|_0 = s_u \leq \frac{1}{2} nk_f q < \frac{1}{2}(nk_f q + 1) = \frac{1}{2} \text{spark}(\Psi_u)\), the spark equivalence condition in Lemma 3.2 is also satisfied, which guarantees that the solution of our \(\ell_1\)-relaxation in (11) is equivalent to the solution of the corresponding \(\ell_0\)-problem. \(\Box\)

This result suggests that for any true signal \((\tilde{u}, \tilde{a})\) will be uniquely recovered with (11), as long as less than half of the entries in \(\tilde{u}\) are nonzero and the matrix \(\Psi_u\), which is determined by the measurement data, is full column rank. Further inspection of \(\Psi_u\) reveals conditions on the number of experiments and the quality of the measured dynamics in the data.

**Corollary 3.1** (Rank Conditions on Output Matrices) \(\Psi_a\) being full column rank (as required by Theorem 3.1), implies that

1. **LTV case:** For \(k = 0, \ldots, k_f - 1\), the matrix \(Z_k \in \mathbb{R}^{n \times q}\), as defined in (2), is full row rank.
Proof 3.2  (1) LTV case: Due to the structure of $\Psi_a$, as defined in (8), each $k$-th block of $n$ columns has at most $q$ rows with nonzero entries, equivalent to the matrix

$$Z_k^T = [z^{(1)}[k] \cdots z^{(q)}[k]]^T \in \mathbb{R}^{q \times n}, \forall k = 0, \ldots, k_f,$$

(18)
as initially constructed in (2). Every consecutive block of $n$ columns in $\Psi_a$ has its nonzero rows in different rows, due to the blkdiag(·) operation. As such, the full column rank condition proposed for $\Psi_a$ can be reinterpreted as a full column rank condition on each block of $n$ columns, and thus on each matrix $Z_k^T$ for $k = 0, \ldots, k_f - 1$.

(2) LTI case: Here, $a$ is only $n^2$-dimensional and the blkdiag(·) structure in (8) disappears, resulting in (9). Each consecutive block of $n$ columns in $\Psi_a$ has at most $k_f q$ rows with nonzero entries, equivalent to a row permutation of the matrix

$$[Z_0 \cdots Z_{k_f-1}]^T.$$

Every consecutive block of $n$ columns in $\Psi_a$ has its nonzero rows in different rows. As such, the full column rank condition proposed for $\Psi_a$ can be reinterpreted as a full row rank condition on (19). □

The LTV condition implies that each time step we require sufficient (or persistent) excitation in the system state over all experiments $j = 1, \ldots, q$. This result confirms that the number of experiments should at least be equal to or greater than the number of state variables, that is $q \geq n$. The LTI condition implies that we require persistency of excitation in the system over both experiments $j = 1, \ldots, q$ and time steps $k = 0, \ldots, k_f - 1$, which is, unsurprisingly, easier to satisfy than the LTV condition. This confirms that the number of time steps times the number of experiments should at least be greater than or equal to the number of state variables, that is $k_f q \geq n$. Corollary 3.1 provides interesting experimental conditions that are sufficient, which are intuitive from the perspective of system identification, a field that traditionally tries to understand how many and what quality experiments are necessary to guarantee the identifiability of a dynamical system model. The notion of persistency of excitation covers this general challenge, and is typically used to understand if an input signal is able to excite the different dynamic modes of a system [21]. When inputs cannot be designed, Corollary 3.1 can be used to check if the output measurements reflect persistent excitation, and combine the right data to construct a well-posed problem.

Note that in some settings, the condition $q > n$ can be restrictive, for instance for identifying larger systems or networks with thousands of states. In the context of LTI systems, different researchers have addressed this challenge and showed that the use of multiple inputs per experiment can reduce the necessary number of experiment if states are excited simultaneously [28,12]. These principles were further developed for sparse LTI network identification via CS [14]. Extending these result to LTV systems seems relevant and remains an open problem.

4 Experimental Validation

The method is tested via synthetic experiments to study its sensitivity to noise, and to relate numerical results to the theoretical results derived in Section 3. We fix the number of states $n = 10$ and time steps $k_f = 4$, and consider an LTV model. We create datasets and increase the number of experiments $q$ and study the method under increasing levels of noise. The synthetic data sets are constructed using the formulation in (1). The entries of each $A[k]$ are drawn from a standard normal distribution and potentially scaled by a factor $\alpha_A$. The input $w^{(j)}[k] \in \mathbb{R}^n$ is sparse, and acting on one randomly picked state variable at each time point in every experiment, hence $s = 1 \cdot k_f \cdot q = 4q$. The targeted node of each experiment is randomized, such that the nodes are perturbed uniformly throughout all experiments. The nonzero input entries are constructed randomly from a standard normal distribution and scaled by a factor $\alpha_w$. $w^{(j)}[k] \in \mathbb{R}^n$ represents the noise acting on the state variables, sampled from a bounded distribution (either thresholded standard normal or uniform) and scaled by a factor $\alpha_w$. In our experiments we have set $\alpha_A = 1$, $\alpha_w = 1$ and varied $\alpha_w$ for simulating different levels of noise that can be interpreted as noise percentage. Following this setup, we have that the cardinality percentage of the input $\rho_u = \frac{s}{nk_f q} = \frac{1}{10}$. Lemma 3.1 tells us that $\rho_u = \frac{1}{10} \leq \frac{1}{2}$.

To assess the performance of our algorithm, we run a Monte Carlo experiment with $T$ iterations and compare the retrieved signal $\bar{u}^*$ to the true signal $\bar{u}$. We introduce two complementary metrics. First, to assess how many entries in $\bar{u}^*$ were recovered correctly (regardless of magnitude), we compute the Mean Average Percentage Error of the cardinality ($MAPE_{card}$) by computing the number of false positives ($\#FP$) and false negatives ($\#FN$):

$$MAPE_{card} \left(\{\bar{u}^*(t)\}_{t=1}^T\right) = \frac{1}{T} \sum_{t=1}^T \frac{\#FP + \#FN}{nk_f q},$$

(20)

Second, for the $s$ nonzero entries of the true signal $\bar{u}$, coined $\bar{u}_{nz}$, we compute the Average Root Mean Square
Error (ARMSE) to assess the error in the magnitude:

\[
ARMSE_{nz} \left( \{u^*(t)\}_{t=1}^T \right) = \sqrt{\frac{1}{T} \sum_{t=1}^T ||\bar{u}_{nz}(t) - u^*_nz(t)||^2_2}
\]  

(21)

For \( \bar{a}, a^* \), we simply track the standard ARMSE over all entries, unless we try to retrieve a sparse solution.

Figures 1 and 2 present the results of our synthetic experiments. In Figure 1 we took three different levels of measurement noise (0%, 1% and 5%, equivalent to \( \alpha_w \in \{0, 0.01, 0.05\} \)) and increased the number of experiments \( q \). In Figure 2 we fixed the number of experiments (\( q = 30 \)), and increased the noise level. As anticipated by Section 3, for the LTV case, we need \( q > n \) to correctly retrieve the model parameters and input values. Across the different levels of noise, full convergence is reached between 20 to 30 experiments, depending on the necessary accuracy. Since the unknowns are all drawn from a standard normal distribution, the ARMSE and MAPE metrics can be interpreted as a relative percentage-style error. We see that, for sufficient experiments \( q \), the ARMSE and MAPE dive under the noise level added to the dynamics. As expected, higher levels of noise yield a higher asymptotic error, which is clearly visualized in Figure 2. We see that the dynamical system model parameters in \( a \) absorb an entry-wise error that is roughly a factor 5 higher than the entry-wise error in the nonzero input values of \( u \). Lastly, we notice that the ARMSE errors grow gradually and sublinearly, whereas the MAPE error jumps to 5.3% for 1% noise and then increases slightly to 6.2% for 15% noise.

5 Conclusions

This paper developed a method for blind system identification of discrete-time linear time-varying (LTV) models with a fully observed output. A sparse recovery problem was formulated to retrieve the dynamical system parameters and unknown input values by exploiting \( a \) priori knowledge that the effect of unknown inputs is limited to a fixed number of states and time points. An optimization problem was formulated as a compressive sensing problem with a partially sparse support, which allowed analysis via sparse recovery theory. This yielded sufficient conditions stating that the number of experiments should be greater than the number of states, the input should overall affect less than half of the states and time points, and the system output matrices should be of full rank. Assessment of the method with synthetic data confirmed theoretical insights and provided further directions for designing experiments. This work opens the further investigation of the blind system identification problem for time-varying systems. The authors are exploring the incorporation of more structured inputs to reduce the search space and make the overall recovery problem more efficient, and have applied the method to help design experiments in breast cancer research [9].

Acknowledgements

The authors would like to thank Young-Hwan Chang at OHSU and Margaret Chapman for their early involvement and useful feedback, and Joe Gray, James Korkola and Laura Heiser at OHSU for motivating this work through their breast cancer research and experiments.

References

[1] K. Abed-Meraim, W. Qiu, and Y. Hua. Blind system identification. Proceedings of the IEEE, 85(8):1310–1322, Aug. 1997.

[2] A. Ahmed and E. P. Xing. Recovering time-varying networks of dependencies in social and biological studies. Proceedings of the National Academy of Sciences, 106(29):11878–11883, July 2009.

[3] R. Akbani, K. Becker, N. Carragher, T. Goldstein, L. Koning, U. Korf, L. Liotta, G. B. Mills, S. S. Nishizuka, M. Pawlak, E. F. Petricoin, R. B. Pollard, B. Serrels, and J. Zhu. Realizing the Promise of Reverse Phase Protein Arrays for
