2011 American Control Conference
on O'Farrell Street, San Francisco, CA, USA
June 29 - July 01, 2011

Exact Topology Identification of Large-Scale Interconnected Dynamical Systems from Compressive Observations

Borhan M. Sanandaji, Tyrone L. Vincent, and Michael B. Wakin

***Abstract*— In this paper, we consider the problem of identifying the exact topology of an interconnected dynamical network from a limited number of measurements of the individual nodes. Within the network graph, we assume that interconnected nodes are coupled by a discrete-time convolution process, and we explain how, given observations of the node outputs, the problem of topology identification can be cast as solving a linear inverse problem. We use the term *compressive observations* in the case when there is a limited number of measurements available and thus the resulting inverse problem is highly underdetermined. Inspired by the emerging field of Compressive Sensing (CS), we then show that in cases where network interconnections are suitably *sparse* (i.e., the network contains sufficiently few links), it is possible to perfectly identify the topology from small numbers of node observations, even though this leaves a highly underdetermined set of linear equations. This can dramatically reduce the burden of data acquisition for problems involving network identification. The main technical novelty of our approach is in casting the identification problem as the recovery of a *block-sparse* signal $\mathbf{x} \in \mathbb{R}^N$ from the measurements $\mathbf{b} = A\mathbf{x} \in \mathbb{R}^M$ with $M < N$, where the measurement matrix A is a block-concatenation of Toeplitz matrices. We discuss identification guarantees, introduce the notion of *network coherence* for the analysis of interconnected networks, and support our discussions with illustrative simulations.**

I. Introduction

Many problems of current interest to the controls community involve large-scale interconnected dynamical systems. In this paper, we focus on systems with a large number of observable variables, where the relationships between these variables can be described by a signal flow graph with nodes of low maximum degree. Examples of such systems come from thermal modeling of buildings [1]–[3], biology [4], and economics [5]. While there has been quite a bit of work to date on the analysis and control of networked systems (see e.g., [6]–[8]), such analysis typically requires knowledge of the network topology, which is not always available a priori. Thus, there is a need for effective "topological identification" procedures [9]–[11] which, given measurements of the nodes of an interconnected dynamical system over a finite time interval, can determine the correct interconnection topology.

The topology identification problem has been solved by Materassi and Innocenti [11] in the case that the interconnection graph has a tree structure and enough data is available to form reliable estimates of cross-power spectral densities. In this paper, we consider a more general setting, allowing arbitrary interconnections (including trees, loops, and self-loops) between nodes in the network, but we assume that the interconnection graph is sparse in the sense that each node has a relatively low degree. Figure 1 shows one such example network. Given this constrained structure to the network graph, we are interested in identifying the network topology from as few observations of the nodes as possible.

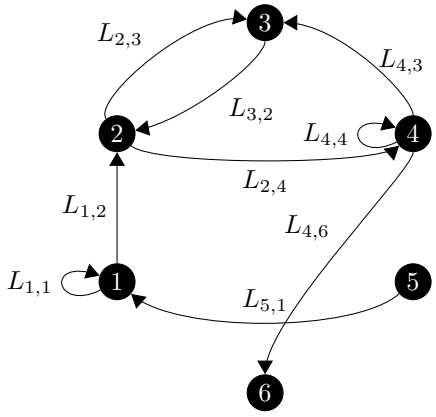

Fig. 1. Network model of 6 interconnected nodes.

We cast the identification problem in the context of Compressive Sensing (CS) [12]–[15], which is concerned with the recovery of sparse signals from limited numbers of measurements. In particular, we explain how the identification problem can be formulated as the recovery of a *block-sparse* signal $\mathbf{x} \in \mathbb{R}^N$ from the measurements $\mathbf{b} = A\mathbf{x} \in \mathbb{R}^M$ with $M < N$, where the measurement matrix A is a block-concatenation of Toeplitz matrices.

The connection between CS and network topology identification has been noted before in [10], where a greedy algorithm was proposed, although no recovery guarantees were presented. We propose to perform the identification using the Block Orthogonal Matching Pursuit (BOMP) algorithm [16]–[18] from CS. Based upon the block-coherence metrics that guarantee the *exact* recovery of a block-sparse signal from its compressive measurements [16]–[18], we introduce the notion of *network coherence* for the analysis of interconnected networks where the interconnections couple the nodes based on a discrete-time convolution process. We derive inequality bounds for the network coherence in certain simple networks, and we show that the coherence metrics are bounded below by a non-zero value that depends on the link impulse responses. We also illustrate the performance of the BOMP identification algorithm on graphs that include loops.

All authors are with Division of Engineering, Colorado School of Mines, Golden, CO 80401, USA. Email:{bmolazem,tvincent,mwakin}@mines.edu.
This work was partially supported by AFOSR Grant FA9550-09-1-0465, NSF Grant CCF-0830320 and NSF Grant CNS-0931748.

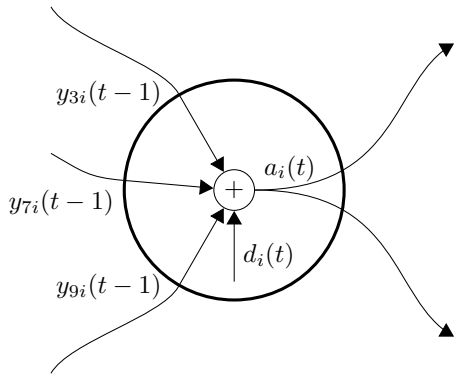

Fig. 2. Single-node model. Node i sums the signals that terminate upon it (in this illustration, nodes 3, 7, and 9), plus a node specific input term d_i.

II. Notation

Denote the set of real numbers by $\mathbb{R}$. All signals are discrete time, defined over a finite non-negative time interval, and represented equivalently as either the function of integer time $x(t)$ or grouped into vector form using the boldface $\mathbf{x}$. Given a finite sequence $\mathbf{x}$, define the following mapping to a Toeplitz matrix

$$\mathcal{T}(\mathbf{x})_i^j := \begin{bmatrix} x(0) & 0 & \cdots & 0 \\ x(1) & x(0) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ x(i-1) & x(i-2) & \vdots & x(i-j) \end{bmatrix}$$

where zeros are applied if the index goes outside the defined range of $\mathbf{x}$.

III. Problem Setup

A. Network Model

Given an interconnected network of P nodes, let node i be associated with the time series $a_i(t)$, $t = 1, 2, \ldots, M$. An example network model is illustrated in Fig. 1. Any types of interconnections between the nodes such as trees, loops, and self-loops are allowed in the network topology.

An edge in the graph, labeled $L_{j,i}$, represents a dynamic system that filters the output of node j (that is, $a_j(t)$) and passes the result (which we call $y_{ji}(t)$) as an input to node i. Let $\mathcal{N}_i$ denote the set of nodes whose outputs are processed and fed to node i. As shown in Fig. 2, we assume that each node i simply sums the signals that terminate upon it $\{\mathbf{y}_{ji}\}_{j\in\mathcal{N}_i}$ (with a one sample delay) and adds a node-specific input term $\mathbf{d}_i$ that may or may not be known. In other words, the output of node i is given by

$$a_i(t) = \sum_{j\in\mathcal{N}_i} y_{ji}(t-1) + d_i(t), \tag{1}$$

for $t = 1, 2, \ldots, M$.

In this paper, each system $L_{j,i}$ is modeled as a causal Finite Impulse Response (FIR) filter with impulse response $\mathbf{x}_{ji} \in \mathbb{R}^n$, so that $\mathbf{y}_{ji} = \mathbf{a}_j * \mathbf{x}_{ji}$. Assuming $a_j(t) = 0$ for $t \le 0$, the convolution can be written as

$$y_{ji}(t) = \sum_{s=1}^{t} x_{ji}(s) a_j(t-s+1), \tag{2}$$

for $t = 1, 2, \ldots, M-1$. Note that according to Eq. (1) we have assumed no feedthrough term, $y_{ji}(0) = 0$, and $x_{ji}(s)$ is defined for $1 \le s \le n$. Setting $\mathbf{x}_{ji} = \mathbf{0}$ for $j \notin \mathcal{N}_i$, then according to Eq. (1), the output of each node $\mathbf{a}_i \in \mathbb{R}^M$ can be written as

$$\mathbf{a}_i = \sum_{j=1}^{P} A_j \mathbf{x}_{ji} + \mathbf{d}_i, \qquad i = 1, 2, \cdots, P, \tag{3}$$

where $A_j = \mathcal{T}(\mathbf{a}_j)_M^n$ is an $M \times n$ Toeplitz matrix, $\mathbf{x}_{ji} \in \mathbb{R}^n$ and $\mathbf{d}_i \in \mathcal{R}^M$. Equation (3) can be further rewritten as

$$\mathbf{a}_i = \underbrace{\begin{bmatrix} A_1 & \cdots & A_j & \cdots & A_P \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \mathbf{x}_{1i} \\ \vdots \\ \mathbf{x}_{ji} \\ \vdots \\ \mathbf{x}_{Pi} \end{bmatrix}}_{\mathbf{x}^i} + \mathbf{d}_i, \tag{4}$$

or equivalently as

$$\mathbf{a}_i = A\mathbf{x}^i + \mathbf{d}_i, \qquad i = 1, 2, \cdots, P, \tag{5}$$

where $\mathbf{a}_i \in \mathbb{R}^M$, $\mathbf{x}^i \in \mathbb{R}^{Pn}$, and $A \in \mathbb{R}^{M\times Pn}$ is a matrix formed by the concatenation of P Toeplitz matrices.

B. Topology Identification

The problem of identifying the network topology can be viewed as recovering the set of interconnection impulse responses $\{\mathbf{x}^i\}_{i=1}^P$ given experimental data. We assume the inputs are decomposed into $\mathbf{d}_i = \widehat{\mathbf{d}}_i + \widetilde{\mathbf{d}}_i$, where $\widehat{\mathbf{d}}_i$ is known and $\widetilde{\mathbf{d}}_i$ is unknown. Therefore, the measurements available to us consist of all outputs $\{\mathbf{a}_i\}_{i=1}^P$ and the known components $\{\widehat{\mathbf{d}}_i\}_{i=1}^P$ of the inputs. Our goal is to estimate the impulse responses $\widehat{\mathbf{x}}_{ji}$ that best match these observations in an appropriate sense; we then determine that a link exists whenever $\|\widehat{\mathbf{x}}_{ji}\|$ exceeds some threshold (we set the threshold to zero when striving for perfect recovery of all $\widehat{\mathbf{x}}_{ji}$).

To solve this problem, we will utilize the following minimization problem:

$$\min_{\{\mathbf{x}^i\}_{i=1}^P} \quad \sum_{i=1}^{P} \|A\mathbf{x}^i - (\mathbf{a}_i - \widehat{\mathbf{d}}_i)\|_2^2. \tag{6}$$

Equivalently, the objective function in (6) can be minimized by solving

$$\min_{\mathbf{x}^i} \quad \|A\mathbf{x}^i - (\mathbf{a}_i - \widehat{\mathbf{d}}_i)\|_2^2 \tag{7}$$

separately for each node in the network. We note that the same matrix A will be used for recovery of all $\mathbf{x}^i$. For simplicity and without loss of generality, we will suppose henceforth that $\mathbf{a}_i - \widehat{\mathbf{d}}_i = \mathbf{b}$ and $\mathbf{x}^i = \mathbf{x}$ for each specific node and focus on the optimization problem

$$\min_{\mathbf{x}} \quad \|A\mathbf{x} - \mathbf{b}\|_2^2, \tag{8}$$

where by letting $N = Pn$, we have $\mathbf{b} \in \mathbb{R}^M$, $\mathbf{x} \in \mathbb{R}^N$, and $A \in \mathbb{R}^{M \times N}$ is a matrix consisting of a concatenation of P Toeplitz matrices.

IV. Exploiting Sparse Topological Structures for Network Identification

From standard arguments in linear algebra, we know that if we collect $M \geq N$ measurements and if A is full rank, then exact recovery of $\mathbf{x}$ when $\widetilde{\mathbf{d}}_i = \mathbf{0}$ is possible from

$$\mathbf{x}^\star = A^\dagger \mathbf{b}, \tag{9}$$

where $A^\dagger = (A^T A)^{-1} A^T$ is the Moore-Penrose pseudoinverse of A. Note that N depends on the number of nodes in the network. Thus, large numbers of measurements are required for exact recovery of large-scale networks in this way.

However, we note that under the assumption of *sparsity* of the node interconnections (that is, assuming only a few nodes contribute to the output of each node), there will be a distinct structure to the solutions $\mathbf{x}$ that we are searching for. In particular, a typical vector $\mathbf{x}$ under our model assumptions will have very few non-zero coefficients, and these non-zero coefficients will be clustered into blocks of FIR coefficients. In the field of CS, such structure is known as *block-sparsity* [16]–[18], and for signals obeying block-sparse models it is known that we can recover $\mathbf{x} \in \mathbb{R}^N$ exactly from measurements $\mathbf{b} = A\mathbf{x} \in \mathbb{R}^M$ even when $M \ll N$. Thus, it may be possible to dramatically reduce the amount of data that must be collected in order to solve the network identification problem. In this and the following sections, we will develop these ideas more formally. For simplicity, we assume $\widetilde{\mathbf{d}}_i = \mathbf{0}$, but it is possible to extend our arguments from exact recovery in noise-free settings to robust recovery in noisy settings.

A. CS Notation

First introduced by Candès, Romberg and Tao [12]–[14], and Donoho [15], CS is a paradigm which enables the recovery of an unknown vector from its underdetermined set of measurements under the assumption of sparsity of the signal and under certain conditions on the measurement matrix A. The CS recovery problem can be viewed as recovery of a K-sparse signal $\mathbf{x} \in \mathbb{R}^N$ from its observations $\mathbf{b} = A\mathbf{x} \in \mathbb{R}^M$ where $A \in \mathbb{R}^{M \times N}$ is the measurement matrix with $M < N$ (in many cases $M \ll N$). A K-sparse signal $\mathbf{x} \in \mathbb{R}^N$ is a signal of length N with K non-zero (significant) entries where $K < N$. The notation $K := \|\mathbf{x}\|_0$ denotes the sparsity level of $\mathbf{x}$. Since the null space of A is non-trivial, there are infinitely many candidate solutions to the equation $\mathbf{b} = A\mathbf{x}$; however, CS recovery algorithms exploit the fact that, under certain conditions on A, only one candidate solution is suitably sparse.

CS recovery algorithms can be classified into two main types: greedy algorithms such as Orthogonal Matching Pursuit (OMP) [19] and Compressive Sampling Matching Pursuit (CoSaMP) [20], and convex optimization algorithms such as Basis Pursuit (BP) [21]. In this paper, we focus on using greedy algorithms for recovery of block-sparse signals.

B. The BOMP Recovery Algorithm

Several extensions of the standard CS recovery algorithms have been proposed to account for additional structure in the sparse signal to be recovered [18], [22]. Among these, the BOMP (Block OMP) algorithm [16]–[18] is designed to exploit block sparsity. We will consider BOMP for the topology identification problem due to its ease of implementation and its flexibility in recovering block-sparse signals of different sparsity levels.

To find a block-sparse solution to the equation $\mathbf{b} = A\mathbf{x}$, the formal steps of the BOMP algorithm are listed in Algorithm 1. The basic intuition behind BOMP is as follows. Due to the block sparsity of $\mathbf{x}$, the vector of observations $\mathbf{b}$ can be written as a succinct linear combination of the columns of A, with the selections of columns occurring in clusters due to the block structure of the sparsity pattern in $\mathbf{x}$. BOMP attempts to identify the participating indices by correlating the measurements $\mathbf{b}$ against the columns of A and comparing the correlation statistics among different blocks. Once a significant block has been identified, its influence is removed from the measurements $\mathbf{b}$ via an orthogonal projection, and the correlation statistics are recomputed for the remaining blocks. This process repeats until convergence.

Eldar et al. [18] proposed a sufficient condition for BOMP to recover any sufficiently concise block-sparse signal $\mathbf{x}$ from compressive measurements. This condition depends on the properties of A, as described in the next section.

C. Block-Sparsity and Block-Coherence

Consider $\mathbf{x} \in \mathbb{R}^N$ as a concatenation of P vector-blocks $\mathbf{x}_i \in \mathbb{R}^n$ where $N = Pn$, i.e.,

$$\mathbf{x} = [\mathbf{x}_1^T \cdots \mathbf{x}_i^T \cdots \mathbf{x}_P^T]^T. \tag{10}$$

Also consider a matrix $A \in \mathbb{R}^{M \times N}$ as a concatenation of P matrix-blocks $A_i \in \mathbb{R}^{M \times n}$ as

$$A = [A_1 \cdots A_i \cdots A_P]. \tag{11}$$

A signal $\mathbf{x}$ is called block K-sparse if it has $K < P$ nonzero blocks. We also assume that there is a unique block K-sparse signal $\mathbf{x}$ that satisfies $\mathbf{b} = A\mathbf{x}$.

1) Block-Coherence: Assume for the moment that matrix A has columns of unit norm. The block-coherence [16]–[18] of A is defined as

$$\mu_{\text{block}}(A) := \max_{i,j \neq i} \frac{1}{n} \|(A_i^T A_j)\|_2 \tag{12}$$

where $\|A\|_2$ is the spectral norm of matrix A. In the case where $n = 1$, this matches the conventional definition of coherence [19], [23],

$$\mu(A) := \max_{i,j \neq i} |\mathbf{a}_i^T \mathbf{a}_j|, \tag{13}$$

where $\{\mathbf{a}_i\}_{i=1}^P$ are the columns of matrix A. While μ_{block} characterizes the intra-block relationships within matrix A, the inter-block properties can be quantified by the sub-coherence [16]–[18] of A as

$$\mu_{\text{sub-block}}(A) := \max_k \max_{i,j \neq i} |\mathbf{a}_{k_i}^T \mathbf{a}_{k_j}| \tag{14}$$

where $\mathbf{a}_{k_i}, \mathbf{a}_{k_j}$ are columns of the matrix-block A_k.

Algorithm 1 The BOMP algorithm for recovery of block-sparse signals

Require: matrix A, measurements $\mathbf{b}$, block size n, stopping criteria
Ensure: $\mathbf{r}^0 = \mathbf{b}$, $\mathbf{x}^0 = \mathbf{0}$, $\Lambda^0 = \emptyset$, $l = 0$
repeat
1. match: $\mathbf{e}_i = A_i^T \mathbf{r}^l, \quad i = 1, 2, \ldots, P$
2. identify support: $\lambda = \arg\max_i \|\mathbf{e}_i\|_2$
3. update the support: $\Lambda^{l+1} = \Lambda^l \cup \lambda$
4. update signal estimate: $\mathbf{x}^{l+1} = \arg\min_{\mathbf{z}:\text{supp}(\mathbf{z}) \subseteq \Lambda^{l+1}} \|\mathbf{b} - A\mathbf{z}\|_2$,
where $\text{supp}(\mathbf{z})$ indicates the blocks on which $\mathbf{z}$ may be non-zero
5. update residual estimate: $\mathbf{r}^{l+1} = \mathbf{b} - A\mathbf{x}^{l+1}$
6. increase index: $l = l + 1$
until stopping criteria true
output: $\widehat{\mathbf{x}} = \mathbf{x}^l$

2) Recovery Condition: In [18, Theorem 3], a sufficient condition is provided that guarantees recovery of any block K-sparse signal $\mathbf{x}$ from the measurements $\mathbf{b} = A\mathbf{x}$ via BOMP. This condition is stated in terms of the block-coherence metrics, μ_{block} and $\mu_{\text{sub-block}}$ of the matrix A.

Theorem 1: [18] If $\mathbf{x}$ is block-sparse with K non-zero blocks of length n, then BOMP will recover $\mathbf{x}$ from the measurements $\mathbf{b} = A\mathbf{x}$ if

$$Kn < \mu_T, \tag{15}$$

where

$$\mu_T = \frac{1}{2}\left(\mu_{\text{block}}^{-1} + n - (n-1)\,\frac{\mu_{\text{sub-block}}}{\mu_{\text{block}}}\right). \tag{16}$$

When $n = 1$, Eq. (15) is equivalent to the exact recovery condition using OMP [19], namely, $K < \frac{1}{2}(\mu^{-1} + 1)$. What Theorem 1 tells us is that, for a given matrix A with certain block-coherence metrics (μ_{block} and $\mu_{\text{sub-block}}$), BOMP is guaranteed exact recovery of block-sparse signals of a limited sparsity level. The smaller the block-coherence metrics, the higher the permitted value of K, and the broader the class of signals that can be recovered via BOMP.

V. Network Coherence

In previous sections, we explained how we can cast the topology identification of a large-scale interconnected network as a CS recovery problem where the signal to be identified has a block-sparse structure and the measurement matrix is a block-concatenation of Toeplitz matrices. Since the block-coherence metrics (12) and (14) give a sufficient condition for recovery via BOMP, it is of interest to examine the particular effect of the network interconnection structure on the block-coherence metrics of A. To highlight the important role that these metrics play in the context of our topology identification problem, where the coupling between node outputs is based on a discrete-time convolution process, we will collectively refer to μ_{block} and $\mu_{\text{sub-block}}$ as the *network coherence* metrics. In order to give some insight into how the network coherence relates to the network topology, we focus in this section on networks with very simple interconnection structures.

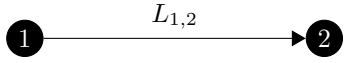

Fig. 3. A simple network for our study of network coherence.

To begin, let us consider the simple network shown in Fig. 3 and assume that the input $d_i(t)$ is Gaussian white noise with unit variance. We would like to estimate the block-coherence and sub-coherence of the matrix A associated with this network. For this network configuration, we can write the output of each node as

$$\mathbf{a}_1 = \mathbf{d}_1 \quad \text{and} \quad \mathbf{a}_2 = A_1\mathbf{x}_{12} + \mathbf{d}_2. \tag{17}$$

It is easy to see that $\mathbf{a}_2$ can be rewritten as

$$\mathbf{a}_2 = G_{12}\mathbf{d}_1 + \mathbf{d}_2 \tag{18}$$

where $G_{12} = \mathcal{T}(\mathbf{x}_{12})_M^M$. Using the down-shift operator $S_M \in \mathbb{R}^{M\times M}$ defined as

$$S_M = \begin{bmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ 1 & 0 & \ddots & \ddots & \vdots \\ 0 & 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix},$$

$G_{12} \in \mathbb{R}^{M\times M}$ can be rewritten as

$$G_{12} = x_{12}(1)S_M + x_{12}(2)S_M^2 + \cdots + x_{12}(n)S_M^n. \tag{19}$$

Similarly, the matrix $A \in \mathbb{R}^{M\times 2n}$ can be written as

$$A = [A_1 \ A_2] = \left[S_M\mathbf{a}_1 \ S_M^2\mathbf{a}_1 \ \cdots S_M^n\mathbf{a}_1 \ S_M\mathbf{a}_2 \ S_M^2\mathbf{a}_2 \ \cdots S_M^n\mathbf{a}_2\right]. \tag{20}$$

Note that $\mathbf{E}\left[\|\mathbf{a}_1\|_2^2\right] = M$ and $\mathbf{E}\left[\|\mathbf{a}_2\|_2^2\right] = M(1+\|\mathbf{x}_{12}\|_2^2)$. Using concentration of measure inequalities, it can be shown that as $M \to \infty$, $\|\mathbf{a}_1\|_2^2$ and $\|\mathbf{a}_2\|_2^2$ are highly concentrated around their expected values [24]. Normalizing by these expected column norms, let $\widehat{A}_1$ and $\widehat{A}_2$ be A_1 and A_2 with approximately normalized columns, and let $\widehat{A} = [\widehat{A}_1 \ \widehat{A}_2]$. Therefore, a reasonable estimate of the block-coherence

$\mu_{\text{block}}(A)$ is simply given by the spectral norm of $\widehat{A}_1^T \widehat{A}_2$. We define such an estimate:

$$\widetilde{\mu}_{\text{block}}(A) := \mu_{\text{block}}(\widehat{A}) = \frac{1}{n}\|\widehat{A}_1^T \widehat{A}_2\|_2. \quad (21)$$

In order to derive a lower bound on $\mathbf{E}\left[\widetilde{\mu}_{\text{block}}(A)\right]$, we use the result of Lemma 1 which states lower and upper bounds on $\|\mathbf{E}\left[\widehat{A}_1^T \widehat{A}_2\right]\|_2$.

Lemma 1: Assume $M > n$. Considering the configuration of the network shown in Fig. 3, we have

$$\frac{\|\mathbf{x}_{12}\|_2}{\sqrt{1+\|\mathbf{x}_{12}\|_2^2}} \leq \|\mathbf{E}\left[\widehat{A}_1^T \widehat{A}_2\right]\|_2 \leq \frac{\|\mathbf{x}_{12}\|_1}{\sqrt{1+\|\mathbf{x}_{12}\|_2^2}}. \quad (22)$$

Proof: See Appendix A. ∎

We are especially interested in deriving lower bounds on the expected value of the network coherence metrics. However, deriving upper bounds would also be of interest. Using Lemma 1, we can state the following theorem.

Theorem 2: For the network of Fig. 3, $\mathbf{E}\left[\widetilde{\mu}_{\text{block}}(A)\right]$ is bounded from below as

$$\mathbf{E}\left[\widetilde{\mu}_{\text{block}}(A)\right] \geq \frac{\|\mathbf{x}_{12}\|_2}{n\sqrt{1+\|\mathbf{x}_{12}\|_2^2}}. \quad (23)$$

Proof: From Jensen's inequality applied for convex functions and Eq. (21), we have the following lower bound for $\mathbf{E}\left[\widetilde{\mu}_{\text{block}}(A)\right]$ as

$$\mathbf{E}\left[\widetilde{\mu}_{\text{block}}(A)\right] = \frac{1}{n}\mathbf{E}\left[\|\widehat{A}_1^T \widehat{A}_2\|_2\right] \geq \frac{1}{n}\|\mathbf{E}\left[\widehat{A}_1^T \widehat{A}_2\right]\|_2 \quad (24)$$

where we use the fact that the spectral norm of a matrix $\|\cdot\|_2$ is a convex function. Combining Eq. (24) and (22), we have

$$\mathbf{E}\left[\widetilde{\mu}_{\text{block}}(A)\right] \geq \frac{\|\mathbf{x}_{12}\|_2}{n\sqrt{1+\|\mathbf{x}_{12}\|_2^2}}. \quad (25)$$

∎

A similar approach can be carried out for the analysis of other types of network elements. For example, we can show that the network coherence of Fig. 4 is bounded by

$$\mathbf{E}\left[\widetilde{\mu}_{\text{block}}(A)\right] \geq \frac{1}{n}\max\left\{\frac{\|\mathbf{x}_{21}\|_2}{\sqrt{1+\|\mathbf{x}_{21}\|_2^2}}, \frac{\|\mathbf{x}_{31}\|_2}{\sqrt{1+\|\mathbf{x}_{31}\|_2^2}}\right\}. \quad (26)$$

We can follow the same steps and derive a bound for the sub-coherence of the simple network of Fig. 3. We simply state the result here: letting $\widetilde{\mu}_{\text{sub-block}}(A) := \mu_{\text{sub-block}}(\widehat{A})$, we have

$$\mathbf{E}\left[\widetilde{\mu}_{\text{sub-block}}(A)\right] \geq \|[\mathfrak{R}_{\mathbf{x}_{12}}(1)\cdots\mathfrak{R}_{\mathbf{x}_{12}}(n-1)]^T\|_\infty \quad (27)$$

where

$$\mathfrak{R}_{\mathbf{x}_{12}}(\tau) := \sum_{i=1}^{n-\tau} \mathbf{x}_{12}(i)\mathbf{x}_{12}(i+\tau)$$

denotes the un-normalized sample autocorrelation function of $\mathbf{x}_{12} \in \mathbb{R}^n$. While finding network coherence bounds for more complicated interconnected networks (e.g., networks with loops and nodes of high out-degree) is a harder task, we observe the following important characteristics:

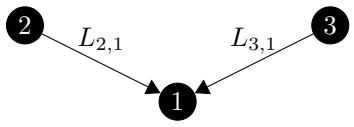

Fig. 4. Three nodes used in our recovery simulations. In the connected case, we suppose node 1 has a subsequent connection to all other 29 nodes in the network. In the disconnected case, node 1has no subsequent connections.

1) In the limit, the network coherence metrics are independent of the number of measurements.
2) The network coherence metrics are bounded below by a non-zero value that depends on the link impulse responses.

The latter phenomenon may suggest an ultimate limitation of the coherence metrics in the analysis of interconnected dynamical networks. Nevertheless, our simulations in the network topology problem do indicate that as the number of measurements M increases, recovery remains possible for a range of interesting problem sizes. The asymptotic behavior of the network coherence metrics is contrary to the conventional behavior in CS, in which increasing the number of rows of a dense matrix (number of measurements M) populated with independent and identically-distributed (i.i.d.) Gaussian random variables will make the coherence approach a zero value, guaranteeing the recovery of signals with more and more non-zero coefficients.

VI. Numerical Simulations

In this section, we test the proposed method for recovering the topology of a dynamical network based on compressive observations with random but known inputs, and we observe how the probability of successful recovery changes for different nodes in the network based on the local sparsity. In all of these simulations, we consider a network of 32 nodes with second order ($n = 2$) interconnecting links.

To begin, in order to highlight the influence of network coherence in the recovery success rate of BOMP, consider the three nodes illustrated in Fig. 4. We suppose these nodes are part of a larger 32-node network, and we consider two possible scenarios: in one case, node 1 has a subsequent connection to all other 29 nodes in the network (we call this the "connected" case), while in the other, "disconnected" case, node 1 has has no subsequent connections to other nodes in the network. In both scenarios, the in-degree of node 1 is 2, while its out-degree is 29 in the connected case and 0 in the disconnected case. In either scenario, we are interested in recovering the incoming links that contribute to node 1. As a function of the number of measurements M, Fig. 5 plots the coherence measures for the two cases where curves are averaged over 1000 realizations of the network. As would be expected from our analysis on the network coherence in Section V, the coherence metrics are bounded from below by a non-zero value that depends on the link impulse responses, namely expressed in Eq. (26) for the disconnected network. The connected network, however, has higher (worse) block- and sub-coherence measures. Although coherence is only a

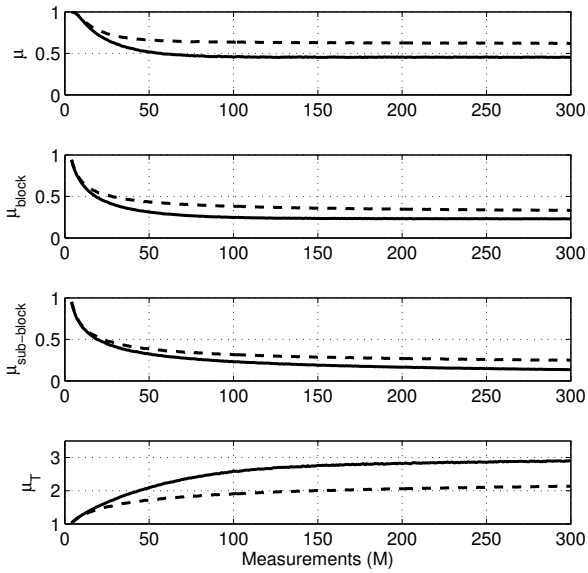

Fig. 5. Coherence metrics for the connected (dashed lines) and disconnected (solid lines) networks. The curves are averaged over 1000 realizations of the networks. Note that the coherence metrics approach a non-zero asymptote as the number of measurements M increases.

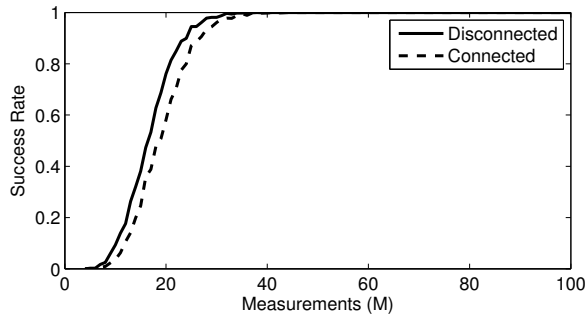

Fig. 6. Recovery rate comparison of node 1 ($L_{2,1}$ and $L_{3,1}$) for connected and disconnected networks. For each measurement, 1000 realizations of the network are carried out and the recovery rate is calculated.

sufficient condition for recovery, simulation results do show weaker recovery performance for the connected network, as shown in Fig. 6. For each value of M, 1000 realizations of the network are carried out and the recovery rate is calculated.

For the sake of comparison, we compute the same coherence metrics for matrices populated with random Gaussian entries in either an unstructured format or in a Toeplitz block format. The results in Fig. 7 show that for $A \in \mathbb{R}^{M\times N}$, the coherence measures approach zero as M increases. In contrast, as we have seen from Fig. 5, in an interconnected network of dynamical systems, the coherence measures have an asymptotic behavior. This prevents the predicted recovery performance from growing as M increases (see the plot of μ_T in Fig. 5).

Finally, to examine the performance of BOMP in recovering the topology of a network containing loops, we consider the more complicated network shown in Fig. 8. As a function of the number of measurements M, the coherence measures of this network are plotted in Fig. 9 where curves are averaged over 1000 realizations of the network.

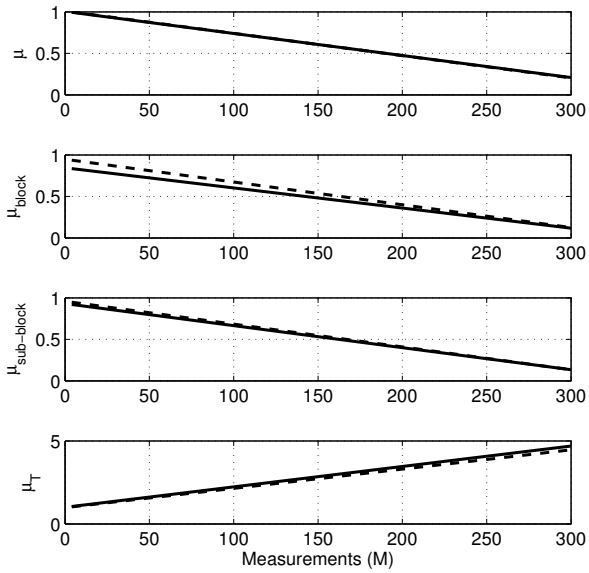

Fig. 7. Coherence metrics for matrices with i.i.d. Gaussian entries (solid lines) and matrices which are block-concatenations of Toeplitz matrices with Gaussian entries. The curves are averaged over 1000 realizations of these types of matrices. Note that the coherence metrics approach zero as the number of measurements M increases.

These measures again approach a non-zero asymptote as the number of measurements increases. Within the network graph, BOMP has different recovery performance for different nodes mainly based on their block-sparsity level (in-degree of the coming links to a node). Fig. 10 shows the recovery performance of the BOMP algorithm for nodes 3, 26 and 28. These links are distinguished from each other by their in-degrees. The single in-connection to node 3 is more likely to be recovered than the 5 in-connections to node 26; we cannot recover these nodes with higher in-degrees until M is large enough that the network coherence metrics reach a suitably small level, and if such a level were never to be reached, we might be limited in the degree of nodes that we could recover by this technique.

VII. Conclusions

In this paper, we have considered the problem of identifying the exact topology of an interconnected dynamical network based on compressive measurements of the individual nodes. We have shown that exact topology identification is indeed possible form compressive node measurements under the assumption that the network contains nodes of low maximum degree. To do this, we have cast the topology identification in the context of CS and employed the BOMP algorithm to recover the block-sparse solutions that encode the network

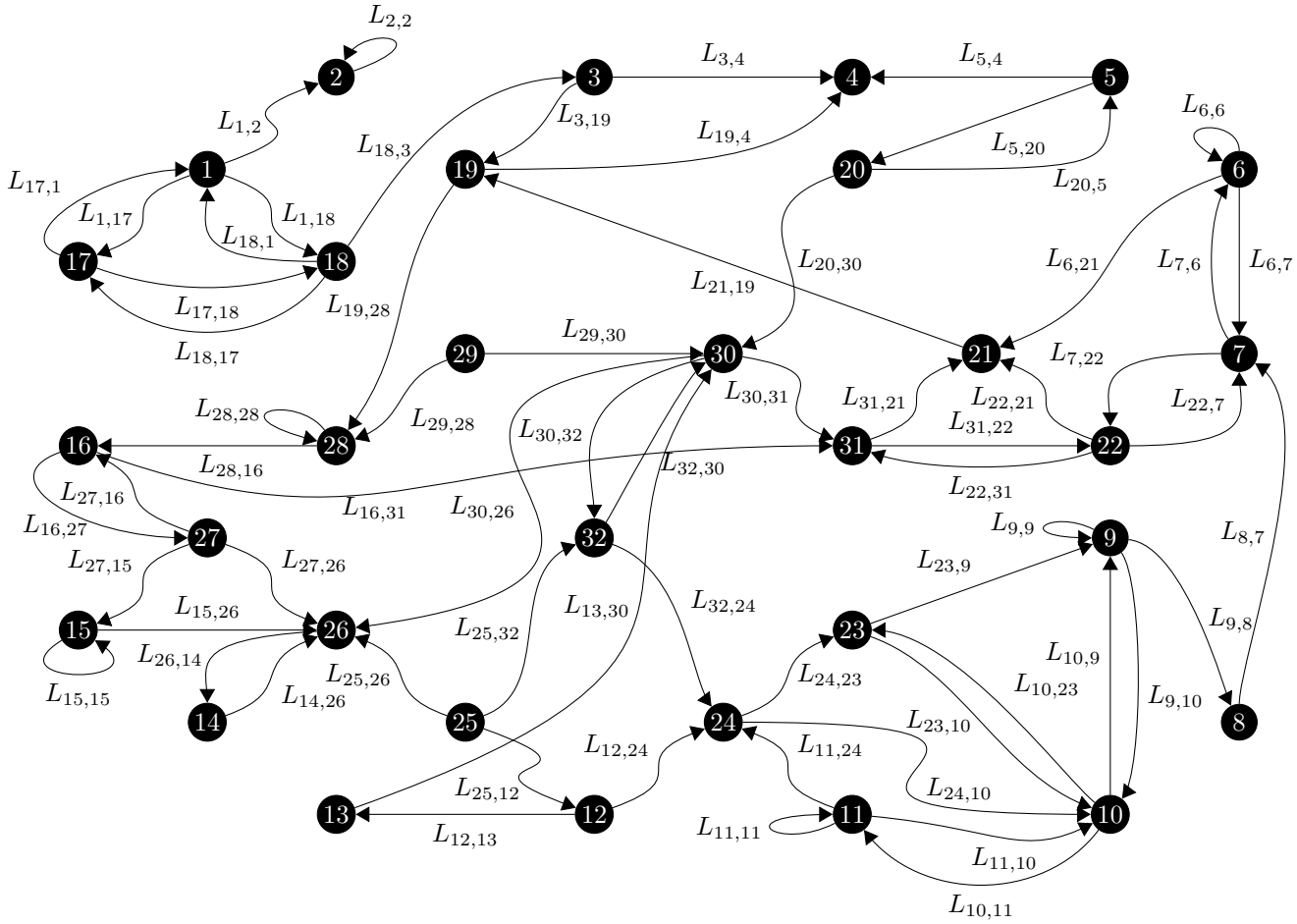

Fig. 8. A complicated network of interconnected dynamical systems including trees, loops and self-loops.

topology. We have characterized the recovery conditions of this algorithm in terms of the network coherence metrics and observed that, as the number of measurements increases, these metrics approach a non-zero value that depends on link impulse responses. Similar approaches can be used to extend our analysis to more complicated networks such as those containing loops. As our simulations have indicated, however, the basic phenomena (including the asymptotic behavior to a non-zero coherence) are likely to generalize.

APPENDIX

A. Proof of Lemma 1

We can show that for $M > n$

$$\mathbf{E}\left[\widehat{A}_1^T \widehat{A}_2\right] = \frac{1}{\sqrt{1+\|\mathbf{x}_{12}\|_2^2}} \cdot \mathcal{T}(\mathbf{x}_{12})_n^n. \tag{28}$$

Using a conservative lower bound for $\|\mathcal{T}(\mathbf{x}_{12})_n^n\|_2 \geq \|\mathbf{x}_{12}\|_2$, from Eq. (28) we get

$$\|\mathbf{E}\left[\widehat{A}_1^T \widehat{A}_2\right]\|_2 \geq \frac{\|\mathbf{x}_{12}\|_2}{\sqrt{1+\|\mathbf{x}_{12}\|_2^2}}. \tag{29}$$

We can derive an upper bound for $\|\mathbf{E}\left[\widehat{A}_1^T \widehat{A}_2\right]\|_2$ using the Cauchy Interlacing Theorem [25]. Let $X_{12} = \mathbf{E}\left[\widehat{A}_1^T \widehat{A}_2\right]$. We have

$$\|X_{12}\|_2^2 = \max_i \lambda_i(X_{12}^T X_{12}) \leq \max_i \lambda_i(\widetilde{X}_{12}^T \widetilde{X}_{12}), \tag{30}$$

where $\widetilde{X}_{12}$ is a $(2n-1)\times(2n-1)$ circulant matrix with

$$\widetilde{\mathbf{x}}_{12} = \frac{1}{\sqrt{1+\|\mathbf{x}_{12}\|_2^2}}[0,\ldots,0,x_{12}(n-1),\ldots,x_{12}(2),x_{12}(1)]$$

as its first row [24]. Since $\lambda_i(\widetilde{X}_{12}^T\widetilde{X}_{12}) = |\lambda_i(\widetilde{X}_{12})|^2$, an upper bound for $\|X_{12}\|_2^2$ is provided by the maximum eigenvalue of $\widetilde{X}_{12}$. Because $\widetilde{X}_{12}$ is circulant, $\lambda_i(\widetilde{X}_{12})$ simply equals the un-normalized length-$(2n-1)$ Discrete Fourier Transform (DFT) of the first row of $\widetilde{X}_{12}$. As a result,

$$\lambda_i(\widetilde{X}_{12}) = \frac{1}{\sqrt{1+\|\mathbf{x}_{12}\|_2^2}} \sum_{k=1}^{n-1} x_{12}(k) e^{-j2\pi(i-1)k/(2n-1)}. \tag{31}$$

From Eq. (31) and by applying the triangle inequality, we get

$$|\lambda_i(\widetilde{X}_{12})| \leq \frac{1}{\sqrt{1+\|\mathbf{x}_{12}\|_2^2}} \sum_{k=1}^{n-1} |x_{12}(k)|. \tag{32}$$

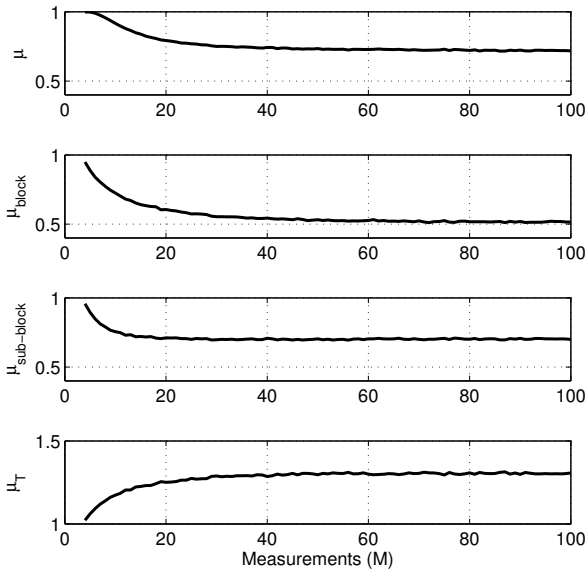

Fig. 9. Coherence metrics for the network of Fig. 8. The curves are averaged over 1000 realizations of the network. Note that the coherence metrics approach a non-zero value and are independent of the number of measurements, in the limit.

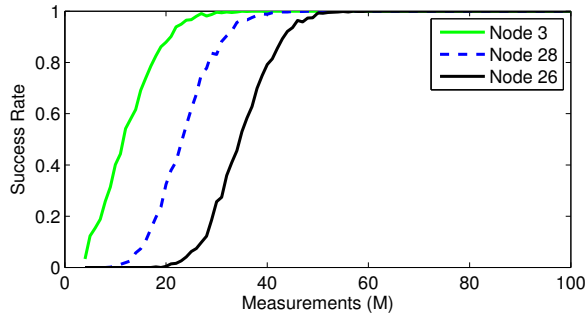

Fig. 10. Recovery rate comparison. For each value of M, 1000 realizations of the network are carried out and the recovery rate is calculated. Note that node 3 has in-degree one, node 28 has in-degree 3, and node 26 has in-degree 5, according to the network shown in Fig. 8.

Therefore, combining Eq. (30) and Eq. (32), we have

$$\|X_{12}\|_2 \leq \frac{1}{\sqrt{1+\|\mathbf{x}_{12}\|_2^2}} \sum_{k=1}^{n-1} |x_{12}(k)| \leq \frac{\|\mathbf{x}_{12}\|_1}{\sqrt{1+\|\mathbf{x}_{12}\|_2^2}}.$$

REFERENCES

[1] J. Pakanen and S. Karjalainen, "Estimating static heat flows in buildings for energy allocation systems," *Energy and Buildings*, vol. 38, no. 9, pp. 1044–1052, 2006.

[2] N. Chaturvedi and J. Braun, "An inverse gray-box model for transient building load prediction," *International Journal of Heating, Ventilating, Air-Conditioning and Refrigerating Research*, vol. 8, no. 1, pp. 73–100, 2002.

[3] B. M. Sanandaji, C. Liu, and T. L. Vincent, "Identification of the thermal dynamics of a building as an interconnected system," *submitted to IEEE Multi-Conference on Systems and Control*, 2011.

[4] E. Ravasz, A. Somera, D. Mongru, Z. Oltvai, and A. Barabási, "Hierarchical organization of modularity in metabolic networks," *Science*, vol. 297, no. 5586, pp. 1551–1555, 2002.

[5] R. Mantegna, "Hierarchical structure in financial markets," *The European Physical Journal B*, vol. 11, no. 1, pp. 193–197, 1999.

[6] S. Hara, H. Shimizu, and T. Kim, "Consensus in hierarchical multi-agent dynamical systems with low-rank interconnections: Analysis of stability and convergence rates," *Proceding of American Control Conference*, pp. 5192 – 5197, 2009.

[7] R. Olfati-Saber and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[8] A. Rahmani, M. Ji, M. Mesbahi, and M. Egerstedt, "Controllability of multi-agent systems from a graph-theoretic perspective," *SIAM Journal on Control and Optimization*, vol. 48, no. 1, pp. 162–186, 2009.

[9] G. Innocenti and D. Materassi, "A modeling approach to multivariate analysis and clusterization theory," *Journal of Physics A: Mathematical and Theoretical*, vol. 41, 2008.

[10] D. Materassi, G. Innocenti, and L. Giarré, "Reduced complexity models in the identification of dynamical networks: links with sparsification problems," *Proceedings of 48th IEEE Conference on Decision and Control*, pp. 4796–4801, 2009.

[11] D. Materassi and G. Innocenti, "Topological identification in networks of dynamical systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 8, pp. 1860–1871, 2010.

[12] E. Candès, "Compressive sampling," *Proceedings of the International Congress of Mathematicians*, vol. 1, no. 3, 2006.

[13] E. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.

[14] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[15] D. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[16] Y. Eldar and M. Mishali, "Robust recovery of signals from a structured union of subspaces," *IEEE Transactions on Information Theory*, vol. 55, no. 11, pp. 5302–5316, 2009.

[17] ——, "Block-sparsity and sampling over a union of subspaces," *Proceedings of the 16th International Conference on Digital Signal Processing*, pp. 1–8, 2009.

[18] Y. C. Eldar, P. Kuppinger, and H. Bölcskei, "Block-sparse signals: uncertainty relations and efficient recovery," *IEEE Transactions on Signal Processing*, vol. 58, no. 6, pp. 3042–3054, 2010.

[19] J. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2231–2242, 2004.

[20] D. Needell and J. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, 2009.

[21] S. Chen, D. Donoho, and M. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1999.

[22] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, "Model-based compressive sensing," *IEEE Transactions on Information Theory*, vol. 56, no. 4, pp. 1982–2001, 2010.

[23] D. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Transactions on Information Theory*, vol. 47, no. 7, pp. 2845–2862, 2001.

[24] B. M. Sanandaji, T. L. Vincent, and M. B. Wakin, "Concentration of measure inequalities for compressive Toeplitz matrices with applications to detection and system identification," *Proceedings of 49th IEEE Conference on Decision and Control*, pp. 2922–2929, 2010.

[25] S. Hwang, "Cauchy's Interlace Theorem for Eigenvalues of Hermitian Matrices," *The American Mathematical Monthly*, vol. 111, no. 2, pp. 157–159, 2004.