Mechanical Systems and Signal Processing ■ (■■■■) ■■■–■■■

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/ymssp

# Modeling and diagnosis of structural systems through sparse dynamic graphical models

Luke Bornn [a,*], Charles R. Farrar [b], David Higdon [b,1], Kevin P. Murphy [c]

[a] Department of Statistics, Harvard University, United States
[b] Los Alamos National Labs, United States
[c] Google Research, United States

ARTICLE INFO

*Article history:*
Received 18 June 2014
Received in revised form
10 September 2015
Accepted 3 November 2015

*Keywords:*
Damage detection
Changepoint detection
Covariance estimation
Bayesian vector autoregression
Graphical models

ABSTRACT

Since their introduction into the structural health monitoring field, time-domain statistical models have been applied with considerable success. Current approaches still have several flaws, however, as they typically ignore the structure of the system, using individual sensor data for modeling and diagnosis. This paper introduces a Bayesian framework containing much of the previous work with autoregressive models as a special case. In addition, the framework allows for natural inclusion of structural knowledge through the form of prior distributions on the model parameters. Acknowledging the need for computational efficiency, we extend the framework through the use of decomposable graphical models, exploiting sparsity in the system to give models that are simple to fit and understand. This sparsity can be specified from knowledge of the system, from the data itself, or through a combination of the two. Using both simulated and real data, we demonstrate the capability of the model to capture the dynamics of the system and to provide clear indications of structural change and damage. We also demonstrate how learning the sparsity in the system gives insight into the structure's physical properties.


## 1. Introduction

As the number and complexity of mechanical and structural systems increases, there grows a need for automated tools to diagnose damage and other anomalies. For instance, airlines may be interested in maximizing the lifespan and reliability of their jet engines, or governmental authorities might like to monitor the condition of bridges and other civil infrastructures in an effort to develop cost-effective lifecycle maintenance strategies. These examples indicate that the ability to efficiently and accurately monitor all types of structural systems is crucial for both economic and life-safety issues. This issue of detecting and explicating damage in engineering structures, known as structural health monitoring (SHM), is facing increasing challenges as existing approaches reach their limit due to growing streams of data from multifarious systems across diverse environmental conditions. Methods are needed, therefore, which tackle the torrent of data and diverse application environments in a manner which separates natural variability from variability caused by damage. This work attempts to address this problem by developing statistical methodologies to detect damage in the face of huge data sources

* Corresponding author. Now at: Department of Statistics and Actuarial Science, Simon Fraser University, Canada and Department of Statistics, Harvard University, United States. Fax: +1 617 496 8057.
*E-mail address:* bornn@stat.harvard.edu (L. Bornn).
[1] Now at: Virginia Tech University, United States.

http://dx.doi.org/10.1016/j.ymssp.2015.11.005
0888-3270/

across broad environmental variability. The proposed methods hold potential to improve economic and life-safety issues associated with all types of aerospace, civil and mechanical infrastructure through improved lifecycle management.

One important automated diagnosis technique is vibration-based damage detection, which operates under the premise that damage will manifest itself through a change in the structure's dynamic response [9,14]. This area of structural health monitoring (SHM) has received much attention, and several detailed reviews have been written [10,34]. These techniques generally fall into two categories – those that use physical process-based models to understand the structural system, and those that use statistical techniques to quantify the various sources of uncertainty present [13]. Both these two paradigms have been employed under a multitude of circumstances and frameworks. We propose to use statistical, in particular Bayesian, techniques to directly model the vibration data, while incorporating knowledge of the physical system through the use of prior distributions and specified sparsity in the statistical model.

This work builds off that of Fugate et al. [17] by modeling the vibration sensor output using autoregressive (AR) techniques. The linear response assumption inherent in AR models is usually justified, as most high-capital expenditure engineering systems are designed to respond in a linear manner to their postulated operational and environmental loading conditions when they are not damaged. Once a model is fit to the vibration data from the structure in its undamaged state, this model is then subsequently used to obtain an assessment of the model's fit to future data. The logic behind this approach is that damage or other structural anomalies will present themselves through changes in the vibration output that will result in a poor fit to the original model. There are many tools to assess the model's fit, the most common being the prediction residuals – the difference between the fitted model and the observed data. If these residuals are large and/or correlated, this indicates poor model fit, and hence structural anomalies. Under the proposed methodology, we can also use the marginal likelihood of the data to detect damage. Specifically, we can measure how likely the observed data is under the model, and conclude that there is damage if this likelihood becomes considerably small. To determine thresholds for determining damage, one can use a sequential probability ratio test [1], control charts [17] or alternative techniques [5].

Most approaches to date have built a single model for the output of each vibration sensor, using heuristic schemes to indicate damage [19]. More recently statistical tools have been used to combine prediction residuals from individual sensors in a statistically rigorous manner to allow for statistical testing and inference [5]. The proposed research makes significant gains by jointly modeling all sensor outputs in a multivariate autoregressive framework. The benefits of such an approach include more accurate modeling of the system, as the statistical model can borrow information from adjacent sensors to provide a more faithful representation of the structural dynamics. While most previous methods could only detect damage that manifested itself in any given sensor, our multivariate framework will also detect anomalies that appear through a change in the correlation structure of the system. One can imagine a situation, such as a crack or loosened bolt, where a system is damaged in such a way that the individual sensor outputs are only slightly affected, but the correlation in the system changes drastically. While previous methods are insensitive to such damage, our approach naturally handles damage that results in changes to the individual sensor outputs or the correlation between sensors. After damage has occurred, one can then use the estimated change in correlation to determine the location of the damage within the sensor network's spatial distribution. It is worth emphasizing that multivariate modeling has seen application in the structural health monitoring literature for some time [11,7,2,27,29,32]. However, the work presented herein extends significantly beyond these studies by adapting Bayesian autogressive models coupled with graphical models to the SHM problem. Graphical models have been used in sensor design and fusion, where the interest is in combining output from multiple sensors [28,30], but have not been applied to SHM. Also, non-Bayesian vector autoregressive methods have recently been proposed for damage detection [15,24,16], with the goal of building hypothesis tests for detecting signal novelty. These methods have ignored system structure, and as such do not naturally allow for the inclusion of physics-based knowledge of the system under study. Likewise for examining the correlation structure of the data from the sensor network; there is very little in the SHM literature on this approach. As such, the proposed multivariate techniques hold significant promise for improving the state-of-the-art in the SHM community.

The traditional AR approaches taken to date have also ignored structure in the system. A Bayesian approach is natural for including knowledge about the system, as there is always some system knowledge to inform the creation of a prior distribution, even if only weak knowledge. This knowledge can come from a variety of sources ranging from numerical models developed during the design process to previously acquired data from the structure when it is in a known condition. Knowledge of the system structure can also be used to introduce sparsity into the model parameters; for instance, if you know that two sensors will be uncorrelated, it is logical to set their correlation to zero instead of using valuable degrees of freedom to estimate it. This is especially important in the multivariate framework, which as a penalty for more accurate modeling capabilities has an increased number of parameters to estimate. Specifically, past approaches modeled each of $K$ sensor individually with an AR model of lag $p$, requiring a total of $pK$ parameters, whereas the joint multivariate framework requires $O(pK^2)$ parameters. This becomes a significant concern due to the large number of sensors ($K$) being generated by massive modern SHM implementations which are becoming economically practical with the continuing evolution of low-cost sensing hardware. As an example, various bridges in Hong Kong currently have monitoring systems with sensors counts ranging from several hundred to more than 1500 [39]. Further, processing is typically performed directly on this hardware (an example of which is shown in Fig. 1), making reduced computation a high priority even for moderate $K$. We cope with this problem by using graphical models to enforce sparsity on the parameters in a way that agrees with knowledge of the system (see for example [22,25,23]).

The remainder of the paper is structured as follows: we begin by showing how Bayesian vector autoregressive (BVAR) models – a tool developed and used primarily in econometrics [26] – are a natural method for modeling SHM systems; we demonstrate this through the use of a simulated example. We then propose the use of graphical models to exploit sparsity in

Please cite this article as: L. Bornn, et al., Modeling and diagnosis of structural systems through sparse dynamic graphical models, Mech. Syst. Signal Process. (2015), http://dx.doi.org/10.1016/j.ymssp.2015.11.005

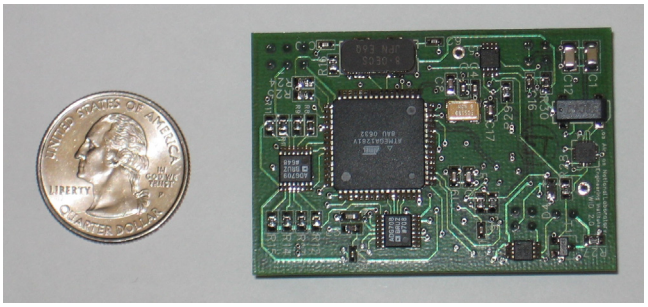

**Fig. 1.** Typical wireless sensor node developed specifically for SHM applications [12].

the system to increase computational efficiency and model estimability in the face of massive data sources. In addition, we discuss methods for specifying this sparsity, including automated methods, expert opinion, and combinations of the two in order to take advantage of knowledge of the structure under study. We show incredibly promising early results with these graphical model-based BVAR methods, making them a natural foundation for detecting damage in SHM systems. Next, we look at the methodology's application to real-world structures, in particular a laboratory test structure designed to replicate the nonlinearities that result from damage in structural systems, comparing our result to traditional AR approaches.

## 2. SHM with Bayesian vector autoregressive models

Our primary goal in performing statistical structural health monitoring is to obtain a feature or set of features from the data which are sensitive to damage, and hence may be used to indicate the presence of damage. One approach is to fit an AR model to sensor output from the undamaged system and monitor the residuals from the model's predictions of subsequent data. An AR model with $p$ autoregressive terms, AR($p$), applied to the $k$th sensor, may be written as

$$x_t^k = \sum_{j=1}^{p} \beta_j^k \cdot x_{t-j}^k + \epsilon_t^k \tag{1}$$

where $x_t^k$ is the measured sensor output from sensor $k$ at discrete times $t$, $\beta_j^k$ are the model parameters, and $\epsilon_t^k$ is an unobservable noise term. Much work has been done on learning the appropriate choice of the model order $p$. For example, one can look at partial autocorrelation plots, measure out of sample prediction performance, or employ a penalized likelihood criterion; see [17,5] for examples in the SHM context.

It is straightforward to see that by expanding our notation we can perform a multivariate analysis of all sensors simultaneously [20]. Let $\boldsymbol{x_t}$ be the $K \times 1$ column vector of observations at time $t$, i.e. $\boldsymbol{x_t} = [x_t^1, x_t^2, \ldots, x_t^K]$ for $K$ sensors. The vector autoregressive (VAR) model may then be written as

$$\boldsymbol{x_t} = \sum_{j=1}^{p} \boldsymbol{B}_j \cdot \boldsymbol{x}_{t-j} + \boldsymbol{\epsilon}_t \tag{2}$$

where each $\boldsymbol{B}_j$ is a $K \times K$ matrix of model parameters and $\epsilon_t$ is a $K \times 1$ noise term coming from a zero-mean Gaussian distribution with covariance parameter $\boldsymbol{\Sigma}$. While these parameters may be estimated using traditional maximum likelihood approaches or Yule–Walker equations, we now turn to Bayesian analysis, which allows us to include structure knowledge through the use of prior distributions [38].

The premise of Bayesian statistics is that parameters are treated as random rather than fixed quantities which can be estimated through the use of Bayes theorem [18]. In this framework, the formula in (2) is the data likelihood; specifically,

$$\boldsymbol{x_t} | \boldsymbol{B}_1, \ldots, \boldsymbol{B}_p, \boldsymbol{\Sigma} \sim N\left( \sum_{j=1}^{p} \boldsymbol{B}_j \cdot \boldsymbol{x}_{t-j}, \boldsymbol{\Sigma} \right). \tag{3}$$

As (3) specifies the likelihood, all that remains to be specified is the prior distributions on $\boldsymbol{B}_1, \ldots, \boldsymbol{B}_K, \Sigma$. Let $\boldsymbol{B} = [\boldsymbol{B}_1 \boldsymbol{B}_2 \ldots \boldsymbol{B}_p]$ be the concatenated matrix of regression parameters. A convenient choice for a prior distribution for $\boldsymbol{B}$ is the conjugate Gaussian prior [26,37]. Using the standard notation with $Vec(\boldsymbol{B})$ the vector of stacked columns of $\boldsymbol{B}$, the prior for $\boldsymbol{B}$ is

$$\boldsymbol{B} \sim N(Vec(\boldsymbol{B}_0), \Omega_0). \tag{4}$$

The prior mean $Vec(\boldsymbol{B}_0)$ and covariance $\Omega_0$ can be specified to match known system knowledge. For instance, if a given sensor is expected to operate as a highly correlated AR(1) process, the corresponding element of $Vec(\boldsymbol{B}_0)$ might be set close to 1. If two sensors are expected to have similar properties, the corresponding elements of $\Omega_0$ might be quite large to reflect the fact that we expect the regression parameters to be similar. Further, we can also encode information about the model order by putting priors centered at zero (with small variance) for the elements of $\boldsymbol{B}$ corresponding to long lags. Another

Please cite this article as: L. Bornn, et al., Modeling and diagnosis of structural systems through sparse dynamic graphical models, Mech. Syst. Signal Process. (2015), http://dx.doi.org/10.1016/j.ymssp.2015.11.005

point worth making is that if preliminary studies have been run, or pilot data has already been collected, we may center these prior parameters at the sample values obtained from this pilot data [18]. Alternatively, if computer model simulations are available (finite element models, for example), one could fit a preliminary BVAR model to this simulated data, using the subsequently estimated values as the prior for the subsequent experiment.

We also must specify a prior distribution for $\Sigma$. The conjugate prior for this parameter is the inverse Wishart distribution [18]. Assigning $\Sigma$ an inverse Wishart prior with mean $\Psi$ and degrees of freedom $\nu$, the distribution takes the form

$$\Sigma \propto |\Sigma|^{(\nu+K+1)/2}\exp(-\mathrm{tr}(\Psi\Sigma^{-1})/2). \tag{5}$$

If two sensors are expected to evolve in the same manner, for instance if they are adjacent and the system remains in an undamaged linear state, we might set the corresponding element of $\Psi$ to be large. It is straightforward to use Bayes' theorem to conclude that the posterior distribution for $\boldsymbol{\theta}=(\boldsymbol{B},\boldsymbol{\Sigma})$ are Gaussian and inverse Wishart, respectively. Further, due to the conjugacy of the model, the marginal likelihood $\pi(\boldsymbol{x}_{1:T})=\int \pi(\boldsymbol{x}_{1:T}|\boldsymbol{B},\Sigma)\pi(\boldsymbol{B},\Sigma)\,d\boldsymbol{B}\,d\Sigma$ (the likelihood after integrating out the model parameters) is also available in closed form. This is quite important, as the marginal likelihood may be used to compare graphical models as well as measure model fit and performance.

Since the posterior distribution $\pi(\boldsymbol{\theta}|\boldsymbol{x}_{1:T})$ is tractable, we can proceed without the use of computationally costly numerical methods. It is worth reiterating this point, as real-world SHM implementations often require computation to be performed locally within the sensor network. As such, computationally expensive methods such as Monte Carlo, or even most optimization methods, are not feasible. In addition, the Bayesian approach provides us with a natural way to perform inference online. Because of the conjugacy of the prior distributions and the sequential nature of SHM, one may simply use the posterior distribution at time $t-1$ as the prior at time $t$, i.e. $\pi(\boldsymbol{\theta}|\boldsymbol{x}_{1:t})\propto\pi(\boldsymbol{\theta}|\boldsymbol{x}_{1:(t-1)})\times\pi(\boldsymbol{x}_t|\boldsymbol{\theta})$. Intuitively, we may proceed by adding in data one time step at a time (using the posterior from the current time as the prior for the next time). In fact, proceeding recursively in this manner is equivalent to performing batch inference; see [18,38] for further details on sequential updating of Bayesian dynamic models.

## 3. Graphical models for structural health monitoring

As inexpensive sensors make vast SHM networks feasible across a wide range of engineering structures, the resulting data generated by such systems is massive (see the example sensor in Fig. 1) requiring methods which naturally filter and make use of the information contained within this data. To model the full system in networks with hundreds or thousands of vibration sensors is not feasible due to the huge number of parameters requiring estimation. By exploiting the interconnectivity of the engineering structure, we can induce parsimonious models which scale naturally with growing networks.

Statistical theory dictates that with enough data, the BVAR parameters will converge to the true value, but that rate of convergence depends on the number of parameters to be estimated. Because the BVAR methodology requires the estimation of at least $(p+1)K^2$ parameters, the convergence is relatively slow. If we suspect that several parameters are zero, we can directly fix these parameters to zero and focus our inference on the remaining parameters. This is especially useful in engineering systems, where knowledge of the system is available through engineering analyses (e.g. finite element simulations), previously acquired data, or more qualitative sources such as expert opinion that might be used to inform the choice of sparsity.

One possible approach to specifying sparsity in a BVAR model is to create a binary vector of length $(p+1)K^2$ indicating whether each variable is zero or non-zero. Specifically, we create an indicator for every element of each of the $K\times K$ matrices $\boldsymbol{B}_1,\ldots,\boldsymbol{B}_p,\boldsymbol{\Sigma}$. A simple way to specify this sparsity is to represent the observed data as vertices in a graph, and use the presence or absence of edges between vertices as the indicator of sparsity. Let $V$ be the set of all vertices. In a BVAR there will be $(p+1)K$ such vertices. Let $E$ be the set of $|V|\times|V|$ possible edges connecting the vertices in the graph. In the BVAR framework where we model recursively through time, we assume at a given time point past lags are uncorrelated, and hence we restrict ourselves to $K^2+pK$ possible edges. The graph $G=(V,E)$ describes the sparsity of the system. In addition, each edge is either directed or undirected; for the purpose of our discussion, it suffices to consider the directed edges as the mean parameters and the undirected edges as the covariance parameters. It is possible to model the system using only undirected or directed edges or various combinations of the two, but the above choice allows for the most natural interpretation for BVAR models as well as for conjugate (and hence computationally efficient) posterior distributions. Intuitively, one may think of the mean parameters (directed edges) as driving the system dynamics, while the covariance parameters (undirected edges) capture the correlation between sensors. In cases where certain sensor output is non-Gaussian, or nonlinear relationships exist, more general graphical model frameworks could be employed [31].

### 3.1. Sparse directed graphs

Sparsity in the AR coefficients is represented by the presence and absence of directed edges in the graph. For example, in a one-dimensional AR model with lag length 4, the graphical model in Fig. 2 would indicate that the AR coefficients for lag lengths 2 and 3 are zero. The model indicated in Fig. 2 would require the estimation of only 3 parameters (the two regression coefficients, as well as the error variance), compared to 5 parameters for the full model. It is worth noting that

Please cite this article as: L. Bornn, et al., Modeling and diagnosis of structural systems through sparse dynamic graphical models, Mech. Syst. Signal Process. (2015), http://dx.doi.org/10.1016/j.ymssp.2015.11.005

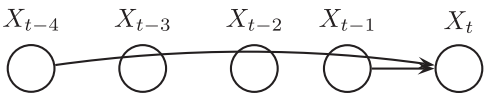


**Fig. 2.** AR(4) example of graphical model.

due to the time invariant nature of AR models, we only display the model for a single time-slice, looking solely at the variables influencing the current time point. It is straightforward to see that this sparsity results in $Vec(\boldsymbol{B})$ having reduced length as parameters that are fixed to zero can be removed from the inference, allowing us to maintain tractable and scalable inference. Thus inference proceeds straightforwardly with only minor modifications once the graph $G$ is specified. Another interesting features of graphical models is that we can view the AR model order, $p$, as the lag beyond which there are no more directed edges. As such, learning of the model order $p$ can be considered as part of the graphical model learning process.

### 3.2. Sparse undirected graphs

Specifying the undirected portion of the graph corresponding to sparsity in the covariance matrix $\Sigma$ requires several graph theoretic notions to maintain tractability; we now formally describe the notion of decomposable graphs – a subset of graphical models which are both tractable and computationally efficient. Focusing solely on the undirected portion of the graph, a graph is decomposable if any cycle of length 4 or greater is connected by an edge connecting two non-consecutive vertices in the cycle. We say that a set of vertices of $G$ forms a complete subgraph of $G$ if every pair of nodes in the set are connected by an edge. A complete subgraph that is not a subset of a larger complete subgraph is called a clique. A set of vertices through which all paths between two cliques must pass is called a separator of those cliques. A triple $(C_1, C_2, S)$ of disjoint nonempty subsets of $V$ forms a decomposition of $G$ if $S$ separates $C_1$ and $C_2$, $V = C_1 \cup C_2 \cup S$, and $G_S = (S, E \cap (S \times S))$, the subgraph of $G$ induced by $S$ is complete. Thus we can obtain an alternative definition of decomposability. A graph $G$ is decomposable if it is complete, or if there exists a decomposition $(C_1, C_2, S)$ into decomposable subgraphs $G_{C_1 \cap S}$ and $G_{C_2 \cap S}$.

Assuming the undirected portion of the BVAR graph is decomposable, we can then redefine our prior distribution for $\Sigma$ to a form maintaining both sparsity and tractability. The resulting distribution is termed a hyper-inverse Wishart distribution [6], which maintains the marginal prior (and posterior) distribution on each clique as hyper-inverse Wishart. Let $C(G)$ and $S(G)$ denote the set of cliques and separators of $G$, respectively. Then densities on the decomposable graph factorize [25] as

$$P(X) = \frac{\prod_{c \in C(G)} P(X_c)}{\prod_{s \in S(G)} P(X_s)}. \tag{6}$$

In other words all densities, including the prior and posterior, can be computed by taking the ratio of the density calculated over cliques to the density calculated over separators.

### 3.3. Learning directed graphical model structure

We treat the learning of the directed and undirected portions of the graph separately. Firstly, learning the directed portion of the model reduces to determining which elements of the matrix $\boldsymbol{B}$ should be set to zero – a regression-style variable selection problem – and hence standard tricks for such models can be employed. One option is the Lasso [36], which uses penalized least squares to induce sparsity. Simply running the regression problem on one of the many tools available to solve the optimization required for Lasso will lead to a sparse solution [31]. In addition, the level of sparsity can be controlled through cross-validation, selecting the level of sparsity to optimize some model fit criterion. Here we employ the graphical lasso [33], using cross-validation to select the appropriate level of sparsity.

Because the underlying model is inherently multivariate normal (and hence analytically tractable), the model parameters $\boldsymbol{B}$ may be analytically integrated out during the learning of the graphical models. Specifically, all that is required is calculation of the marginal likelihood $\pi(\boldsymbol{x}_{1:T})$; see [31] for details. Subsequently, conditional on the underlying graph $\boldsymbol{B}$ is itself multivariate normal [18,31].

### 3.4. Learning decomposable graphical model structure

Estimating the covariance matrix $\Sigma$ through a decomposable undirected graph takes more effort. On small problems, say with $K < 15$, all possible $2^K$ graphs can be enumerated and checked for decomposability. The decomposable graphs can then be searched over to pick the optimal decomposable graph. While there are numerous metrics available by which to select the best graph, we use the marginal likelihood $\pi(\boldsymbol{x}_{1:T})$ of the data under each graph to select the best-fitting graph, which given the Normal-inverse-Wishart model structure, is a multivariate $t$-distribution evaluated on the cliques over the separators as in Eq. (6). In situations with $K$ large, it is highly inefficient to enumerate all $2^K$ graphs, and instead we prefer to only search over decomposable graphs from the start. There has been much recent work in this area, for instance [8], though we focus on the simple perturbation scheme of [35]. By starting with a decomposable graph (for instance the complete

Please cite this article as: L. Bornn, et al., Modeling and diagnosis of structural systems through sparse dynamic graphical models, Mech. Syst. Signal Process. (2015), http://dx.doi.org/10.1016/j.ymssp.2015.11.005

graph), this method allows one to find the decomposable neighbors. In this way we can search over the space of graphs, proceeding in a greedy manner to select the best graph by starting from some initial structure estimate, as in [21].

To jointly learn the regression and covariance structure, one can iterate between the two problems; in practice we have found that such iteration is not necessary – specifically, learning of the directed structure is not sensitive to the undirected structure and vice versa.

### 3.5. Simulated example

We now demonstrate the BVAR methodology on a simulated example, generating a multivariate AR system where each time series is independent of the history of the other series. Specifically, $\boldsymbol{B}_j$ is diagonal for all $j$, and the values of these diagonals are set to $(-2)^{-j}$ with the length $p$ of the AR process set to 5 (i.e. $j=1,\ldots,5$). We include correlation between series through $\Sigma$; the generating mechanism uses a true value $\Sigma_{True}$ of $\Sigma$,

$$\Sigma_{True} = \begin{matrix} 1 & \rho & \rho & \rho & 0 & 0 & 0 & 0 \\ \rho & 1 & \rho & \rho & 0 & 0 & 0 & 0 \\ \rho & \rho & 1 & \rho & \rho & \rho & 0 & 0 \\ \rho & \rho & \rho & 1 & \rho & \rho & 0 & 0 \\ 0 & 0 & \rho & \rho & 1 & \rho & \rho & \rho \\ 0 & 0 & \rho & \rho & \rho & 1 & \rho & \rho \\ 0 & 0 & 0 & 0 & \rho & \rho & 1 & \rho \\ 0 & 0 & 0 & 0 & \rho & \rho & \rho & 1 \end{matrix} \tag{7}$$

In other words, the graph is decomposable, and the series are correlated in a block structure [3]. We will explore this structure in more detail in later sections.

We simulate a realization from the above model with $\rho=0.2$ using $T=10{,}000$ time steps, then attempt to learn the underlying graph structure. Due to the small model space, we are able to enumerate each possible undirected graph. We recover all of the original directed edges of the graph, as well as several additional spurious edges, and one missing undirected edge between $X_5$ and $X_7$. Fig. 3 shows the resulting estimated structure. Solid edges are the true edges and the dashed edges are the spurious edges found by the method. Repeating the above simulation from a different random seed in the simulation will potentially result in different spurious and missed edges. It is worth noting that if we increase $T$ significantly, we will no longer detect spurious (or missing) edges. Similarly, with much less data, the noise of the system will result in more false positives and false negatives. There is a significant literature about estimating graphical models; see [23,31] for more discussion.

In addition to the computational benefits of restricting the model to a sparse representation, the reduced number of parameters to estimate will result in improved inference, particularly for small training data. In SHM systems, where pre-damage data can be expensive to obtain, efficient use of the data in this way can lead to significant improvements in damage detection fidelity. We explore this in the simulated example monitoring convergence of the parameters using the Bayesian

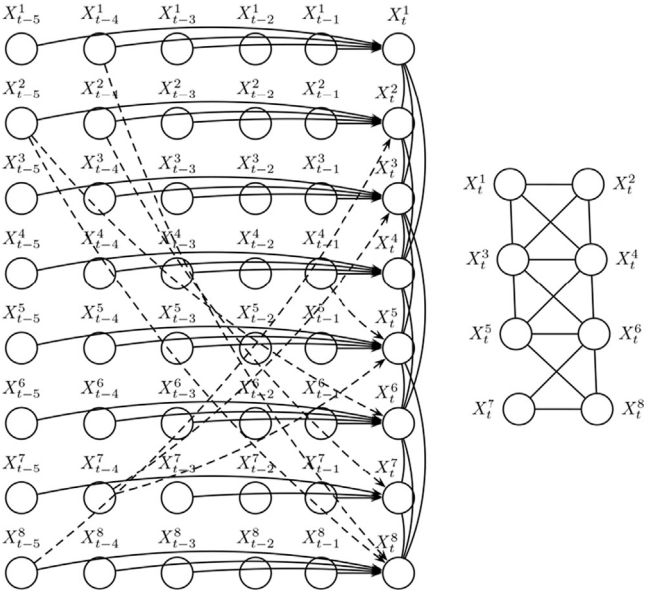


**Fig. 3.** *Left*: learned structure of simulated system. Extraneous edges shown with dotted lines. *Right*: close-up of learned undirected structure.

Please cite this article as: L. Bornn, et al., Modeling and diagnosis of structural systems through sparse dynamic graphical models, Mech. Syst. Signal Process. (2015), http://dx.doi.org/10.1016/j.ymssp.2015.11.005

approach for the full and sparse model. We use a noninformative prior, specifically a zero-mean prior with diagonal covariance with values 100 for the regression parameters. The prior mean for $\Sigma$ is the identity matrix and the degrees of freedom $\nu$ was set to $K$, a conservative choice which reflects limited prior input. Using a recursive Bayesian formula, we are able to monitor the distance between the posterior mean and the true generating value of both the regression and covariance parameters at each time point. For each of the 1000 repetitions of the simulation, we generate new data and store the resulting root mean squared error (RMSE) of the posterior means. In Fig. 4 we show the average convergence of the RMSE over the 1000 simulations as well as the 5% and 95% quantiles to indicate the varying rates of convergence for the full and sparse models. From this, it is clear that the sparse model, by not using degrees of freedom to estimate the zero elements, results in improved inference, particularly for the regression parameters $\boldsymbol{B}$.

## 4. Evaluation and laboratory testing

We now turn to real-world testing, exploring the graphical model-based BVAR in a laboratory structure designed to replicate the non-linear response resulting from damage. The three-story aluminum structure under study (Fig. 5) consists of plates separated by columns and an electro-dynamic shaker that excites the structure through its base. To induce non-linear behavior, a column is suspended from the top floor and a bumper placed on the second floor. Contact of the bumper and column results in non-linearity similar to what might be observed in the opening and closing of a crack or rattling of a loose connection in a damaged structure. The undamaged state of the system occurs when the bumper and the column are set sufficiently apart to prevent contact, and the damaged state corresponds to a static distance of 0.05 mm between the two objects. The shaker excites the structure in both its damaged and undamaged state, and accelerometers mounted on each floor capture the system's response, for a total of $K=4$. The data and more information on the structure can be obtained from http://institutes.lanl.gov/ei. As our test data we concatenate the damaged to the undamaged data, each of which was originally of length 8192.

We now demonstrate the method's ability to detect damage in real-time on our laboratory test structure. After fitting the sparse BVAR on the first 6000 time points, we monitor the marginal likelihood from points 7000 to 9384 in the hopes of detecting the occurrence of damage at time 8192 (Fig. 6). This figure shows that the marginal likelihood is clearly sensitive to the presence of damage, and makes an excellent candidate as a feature to be used in the damage-detection process. Fig. 7 shows the residuals from the fitted BVAR model, as well as a traditional AR model, which fits an AR model to each individual sensor. Here the traditional AR model was fit with maximum likelihood, with the model order $p$ selected with the autocorrelation function. It is evident that the BVAR provides a much better fit to the undamaged data, and hence is much more sensitive to damage. Because we are able to model the correlation structure between the sensors, the presence of damage is detected in the residuals for all 4 sensors, compared to only 2 in the traditional AR model. This can be seen in Fig. 7, where the AR model does not detect any change in the first 2 sensors. As all of the residual plots are on the same scale, it is apparent that the BVAR model provides a much more natural foundation from which to build features for damage detection. In recent work using autoregressive support vector machines, Bornn et al. [5,4] were able to demonstrate improved performance over a traditional AR, but were still unable to identify damage from the first two sensors. A comparison of our

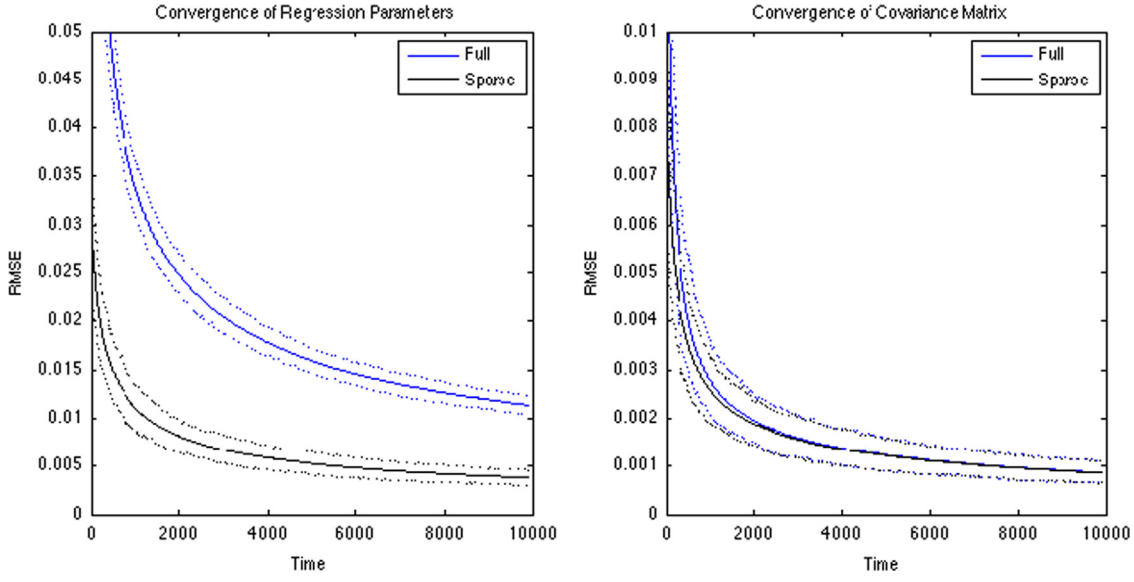


**Fig. 4.** Convergence of posterior mean using BVAR on simulated example. 5% and 95% quantiles shown with dotted lines.

Please cite this article as: L. Bornn, et al., Modeling and diagnosis of structural systems through sparse dynamic graphical models, Mech. Syst. Signal Process. (2015), http://dx.doi.org/10.1016/j.ymssp.2015.11.005

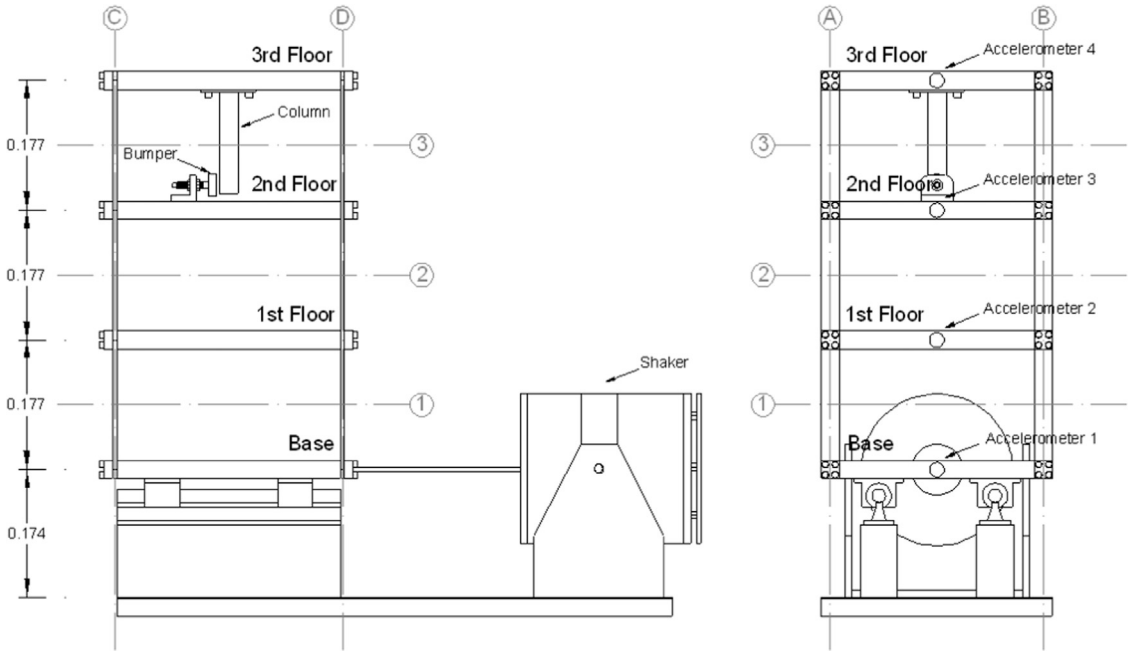


**Fig. 5.** Diagram of laboratory structure.

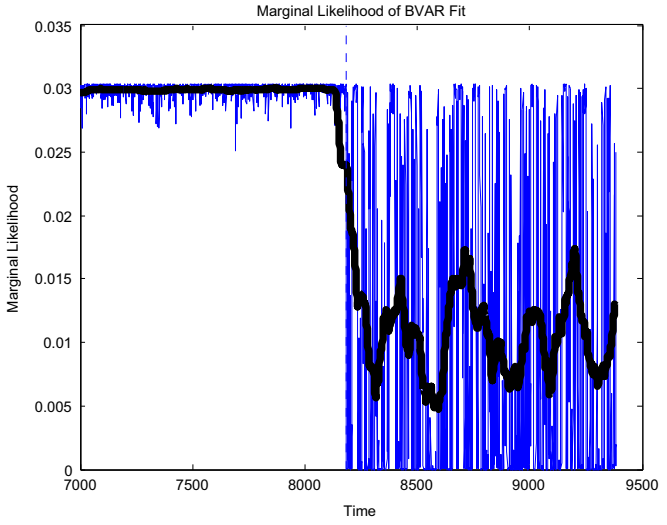


**Fig. 6.** Marginal likelihood of BVAR model with 100 point moving average. Simulated damage occurs at time 8192 (dashed line). Solid black line is the result obtained with a width 100 exponential smoothing window.

results with their results shows that the BVAR approach provides better indication of damage, likely a result of this system behaving linearly in its initial state.

In addition to the BVAR method providing improved prediction at similar cost to the standard AR model, we are also able to use the model's features to understand the damage source. By comparing the estimated structure before and after damage is indicated, we can obtain insight into the location of the damage. Fig. 8 shows the resulting changes in estimated graph structure before and after damage. In the structure, damage is induced by creating contact in the bumper, which is located between sensors 3 and 4. As such, the relationships have generally been broken between the first 3 sensors and sensor 4, as the damage created by the bumper reduces the impact of the other sensors on sensor 4. From this, we can see that by studying the model's structure before and after damage occurs, we can isolate the damage location. In large structures, where finding the damage location consumes time and resources, having a model which provides preliminary estimates of damage location has the potential to cut down on maintenance costs.

Please cite this article as: L. Bornn, et al., Modeling and diagnosis of structural systems through sparse dynamic graphical models, Mech. Syst. Signal Process. (2015), http://dx.doi.org/10.1016/j.ymssp.2015.11.005

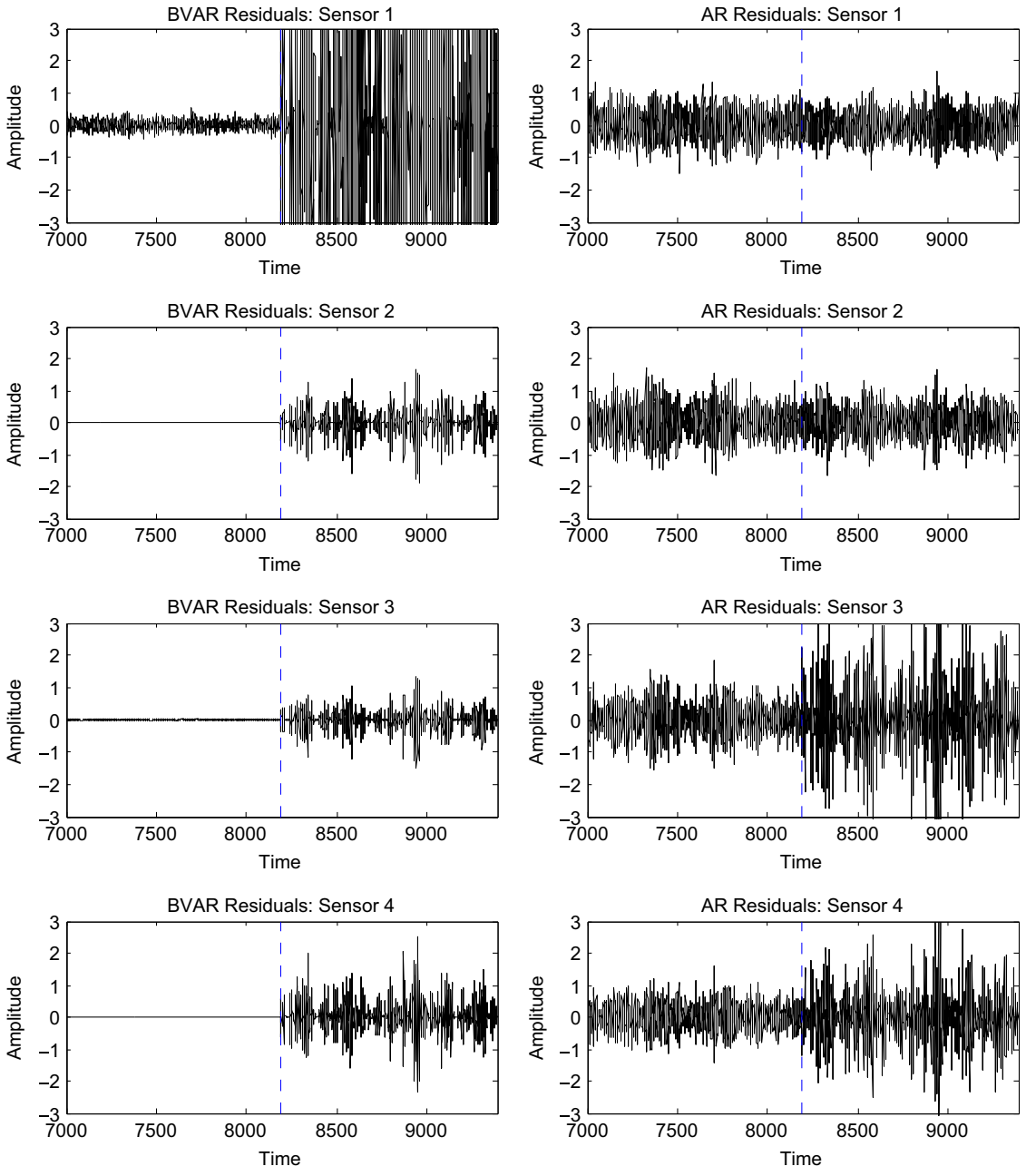


**Fig. 7.** Comparison of residuals from BVAR and AR models. Simulated damage occurs at time 8192 (dashed line). It is clear that the BVAR approach, which models all sensors jointly, provides a solid foundation on which to build statistical SHM methodology.

## 5. Conclusion

In this paper we have demonstrated that the BVAR methodology holds considerable power in modeling structural health monitoring systems. After discussing the merits and challenges associated with the use of BVAR models, namely the considerable number of parameters, we develop a framework for using graphical models to exploit sparsity within a BVAR model. Through a simulated example it was shown that sparsity allows for more accurate modeling of multivariate systems, including more efficient use of the data.

We then demonstrated the applicability and use of our methodology on a physical structure designed to simulate the traits of a damaged structural system. The graphical models approach enables us to examine the system from a different angle, comparing the correlation structure of sensors within the system. Through implementation on the laboratory

Please cite this article as: L. Bornn, et al., Modeling and diagnosis of structural systems through sparse dynamic graphical models, Mech. Syst. Signal Process. (2015), http://dx.doi.org/10.1016/j.ymssp.2015.11.005

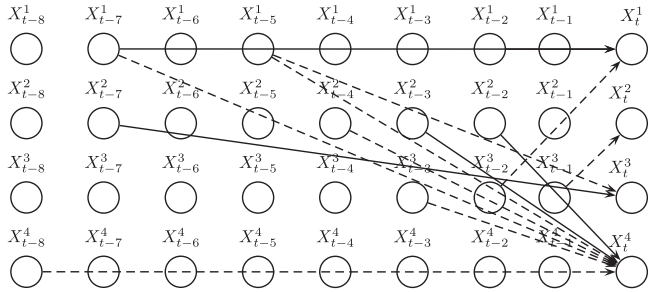

**Fig. 8.** Change in regression parameter structure after damage occurs. Solid arrows are edges added after damage, while dashed lines are edges removed. We see many edges removed between the first 3 sensors and sensor 4, which suggests that damage has occurred near that sensor, clearly matching with our knowledge of the structure.

structure dataset the sparse BVAR methodology was shown to outperform traditional approaches in modeling the system and detecting damage. In addition, through the marginal likelihood we have a powerful and easily calculated single dimensional feature to use in the damage detection process.

## References

[1] D. Allen, H. Sohn, K. Worden, C. Farrar, Utilizing the sequential probability ratio test for building joint monitoring, In: Proceedings of the SPIE Smart Structures Conference, San Diego, 2002.
[2] J.-B. Bodeux, J.-C. Golinval, Modal identification and damage detection using the data-driven stochastic subspace and armav methods, Mech. Syst. Signal Process. 17 (1) (2003) 83–89.
[3] L. Bornn, F. Caron, Bayesian clustering in decomposable graphs, Bayesian Anal. 6 (4) (2011) 829–846.
[4] L. Bornn, C. Farrar, G. Park, Damage detection in initially nonlinear systems, Int. J. Eng. Sci. 48 (10) (2010) 909–920.
[5] L. Bornn, C. Farrar, G. Park, K. Farinholt, Structural health monitoring with autoregressive support vector machines, J. Vib. Acoust. 131 (2009). 021004–1.
[6] A. Dawid, S. Lauritzen, Hyper-Markov laws in the statistical analysis of decomposable graphical models, Ann. Stat. 3 (1993) 1272–1317.
[7] A. De Stefano, D. Sabia, L. Sabia, Structural identification using armav models from noisy dynamic response under unknown random excitation, In: Proceedings of DAMAS, 1997, pp. 419–428.
[8] A. Deshpande, M. Garofalakis, M. Jordan, Efficient stepwise selection in decomposable models, In: Proceedings of the Seventeenth conference on Uncertainty in artificial intelligence, Morgan Kaufmann Publishers Inc, Burlington, MA, 2001, pp. 128–135.
[9] S. Doebling, C. Farrar, M. Prime, D. Shevitz, Damage Identification and Health Monitoring of Structural and Mechanical Systems from Changes in Their Vibration Characteristics: A Literature Review, Technical Report, Los Alamos National Lab., NM, United States, 1996.
[10] S. Doebling, C. Farrar, M. Prime, D. Shevitz, A review of damage identification methods that examine changes in dynamic properties, Shock Vib. Dig. 30 (2) (1998) 91–105.
[11] X. Fang, H. Luo, J. Tang, Structural damage detection using neural network with learning rate improvement, Comput. Struct. 83 (25) (2005) 2150–2161.
[12] K. Farinholt, S. Taylor, T. Overly, G. Park, C. Farrar, Recent advances in impedance-based wireless sensor nodes, in: ASME Conference Proceedings, 2008, pp. 57–65.
[13] C. Farrar, T. Duffey, S. Doebling, D. Nix, A statistical pattern recognition paradigm for vibration-based structural health monitoring, Struct. Health Monitor. 1999 (2000) 764–773.
[14] C. Farrar, K. Worden, An introduction to structural health monitoring, Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci. 365 (1851) (2007) 303–315.
[15] S.D. Fassois, J.S. Sakellariou, Time-series methods for fault detection and identification in vibrating structures, Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci. 365 (1851) (2007) 411–448.
[16] R. Fuentes, A. Halfpenny, E. Cross, R. Barthorpe, K. Worden, An approach to fault detection using a unified linear gaussian framework, In: Proceedings of Structural Health Monitoring, 2013.
[17] M. Fugate, H. Sohn, C. Farrar, Vibration-based damage detection using statistical process control, Mech. Syst. Signal Process. 15 (4) (2001) 707–721.
[18] A. Gelman, G. Carlin, H. Stern, D. Rubin, Bayesian Data Analysis, 2 ed., Chapman & Hall/CRC, London, 2003.
[19] J. Herzog, J. Hanlin, S. Wegerich, A. Wilks, High performance condition monitoring of aircraft engines, In: Proceedings of the GT 2005 ASME Turbo Expo, Reno, NV, 2005.
[20] S. Johansen, Likelihood-Based Inference in Cointegrated Vector Autoregressive Models, Cambridge Univ Press, New York, 1995.
[21] B. Jones, C. Carvalho, A. Dobra, C. Hans, C. Carter, M. West, Experiments in stochastic computation for high-dimensional graphical models, Stat. Sci. 20 (4) (2005) 388–400.
[22] M. Jordan, Graphical models, Stat. Sci. (2004) 140–155.
[23] D. Koller, N. Friedman, Probabilistic Graphical Models: Principles and Techniques, The MIT Press, Cambridge, MA, 2009.
[24] F.P. Kopsaftopoulos, S.D. Fassois, Scalar and vector time series methods for vibration based damage diagnosis in a scale aircraft skeleton structure, J. Theoret. Appl. Mech. 49 (3) (2011) 727–756.
[25] S. Lauritzen, Graphical Models, vol. 17, Oxford University Press, USA, 1996.
[26] R. Litterman, Forecasting with Bayesian vector autoregressions–five years of experience, J. Bus. Econ. Stat. 4 (1) (1986) 25–38.
[27] S.G. Mattson, S.M. Pandit, Statistical moments of autoregressive model residuals for damage localisation, Mech. Syst. Signal Process. 20 (3) (2006) 627–645.
[28] S. Mensah, N.O. Attoh-Okine, A. Faghri, Sensor fusion in transportation infrastructure systems using belief functions, New Dev. Sens. Technol. Struct. Health Monitor. (2011) 187–203.
[29] A. Mosavi, D. Dickey, R. Seracino, S. Rizkalla, Identifying damage locations under ambient vibrations utilizing vector autoregressive models and mahalanobis distances, Mech. Syst. Signal Process. 26 (2012) 254–267.
[30] S.C. Mukhopadhyay, New Developments in Sensing Technology for Structural Health Monitoring, vol. 96, Springer, New York, 2011.
[31] K.P. Murphy, Machine Learning: A Probabilistic Perspective, MIT Press, Cambridge, MA, 2012.
[32] N.M. Okasha, D.M. Frangopol, D. Saydam, L.W. Salvino, Reliability analysis and damage detection in high-speed naval craft based on structural health monitoring data, Struct. Health Monitor. 10 (4) (2011) 361–379.
[33] M. Schmidt, Graphical model structure learning with l1-regularization (Ph.D. thesis), University of British Columbia, Vancouver, 2010.

Please cite this article as: L. Bornn, et al., Modeling and diagnosis of structural systems through sparse dynamic graphical models, Mech. Syst. Signal Process. (2015), http://dx.doi.org/10.1016/j.ymssp.2015.11.005

[34] H. Sohn, C. Farrar, F. Hemez, D. Shunk, D. Stinemates, B. Nadler, J. Czarnecki, A Review of Structural Health Monitoring Literature, Laboratory Los Alamos, New Mexico, 2004, 1996–2001.
[35] A. Thomas, P. Green, Enumerating the decomposable neighbors of a decomposable graph under a simple perturbation scheme, Comput. Stat. Data Anal. 53 (4) (2009) 1232–1238.
[36] R. Tibshirani, Regression shrinkage and selection via the lasso, J. R. Stat. Soc.: Ser. B 58 (1) (1996) 267–288.
[37] G. Welch, G. Bishop, An Introduction to the Kalman Filter, Technical Report TR 95-041, University of North Carolina at Chapel Hill, 1995.
[38] M. West, J. Harrison, Bayesian Forecasting and Dynamic Models, Springer, New York, 1996.
[39] K. Wong, Y. Ni, Integrating bridge structural health monitoring and condition-based maintenance management, In: Proceedings of the Civil Structural Health Monitoring Workshop, 2012.

Please cite this article as: L. Bornn, et al., Modeling and diagnosis of structural systems through sparse dynamic graphical models, Mech. Syst. Signal Process. (2015), http://dx.doi.org/10.1016/j.ymssp.2015.11.005