Pairwise graphical models for structural health monitoring with dense sensor arrays

Citation: Mohammadi Ghazi, Reza et al. "Pairwise graphical models for structural health monitoring with dense sensor arrays." Mechanical Systems and Signal Processing 93 (September 2017): 578-592 © 2017 Elsevier

As Published: http://dx.doi.org/10.1016/j.ymssp.2017.02.026

Publisher: Elsevier BV

Persistent URL: https://hdl.handle.net/1721.1/124003

Version: Original manuscript: author's manuscript prior to formal peer review

Pairwise graphical models for structural health monitoring with dense sensor arrays

Reza Mohammadi Ghazi, Justin Chen, Oral Büyüköztürk*

Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, 02139 USA

Abstract

Through advances in sensor technology and development of camera-based measurement techniques, it has become affordable to obtain high spatial resolution data from structures. Although measured datasets become more informative by increasing the number of sensors, the spatial dependencies between sensor data are increased at the same time. Therefore, appropriate data analysis techniques are needed to handle the inference problem in presence of these dependencies. In this paper, we propose a novel approach that uses graphical models (GM) for considering the spatial dependencies between sensor measurements in dense sensor networks or arrays to improve damage localization accuracy in structural health monitoring (SHM) application. Because there are always unobserved damaged states in this application, the available information is insufficient for learning the GMs. To overcome this challenge, we propose an approximated model that uses the mutual information between sensor measurements to learn the GMs. The study is backed by experimental validation of the method on two test structures. The first is a three-story two-bay steel model structure that is instrumented by MEMS accelerometers. The second experimental setup consists of a plate structure and a video camera to measure the displacement field of the plate. Our results show that considering the spatial dependencies by the proposed algorithm can significantly improve damage localization accuracy.

*Corresponding author

Email address: rezamg@mit.edu, ju21743@mit.edu, obuyuk@mit.edu (Reza Mohammadi Ghazi, Justin Chen, Oral Büyüköztürk)

Keywords: Structural health monitoring, Damage detection, Graphical models, Ising model, Pairwise graphical model, Sensor network, Video camera, Loopy belief propagation, Gibbs sampling

1. Introduction

Health monitoring functionality is an important aspect of smart structures in dealing with challenges associated with complexity, high operational demands, safety, maintenance, and environmental effects. The main difficulties in developing such systems are: 1) appropriate instrumentation of structures for acquiring informative data sets, and 2) providing efficient and robust data analysis algorithms for reliable decision making and uncertainty quantification (UQ). The instrumentation cost for studying structural dynamics beyond simple modal analysis was previously substantially high. Also, former studies showed that linear characterization of structural dynamics may not be sufficient for reliable damage detection [1, 2, 3, 4]. Therefore, these two issues made it economically difficult to rationalize investing in monitoring systems.

Through advances in sensor technology, it is becoming less costly to instrument structures. In addition to mechanical sensors, recently developed novel video-based techniques, can provide spatially high resolution data sets by using video cameras [5, 6, 7]. Although the first challenge regarding the instrumentation of structures may be managed to an extent by these developments, additional complexities are introduced into decision making and UQ due to the computational demand of analyzing big data sets as well as the possible inducement of more uncertainties associated with the use of more sensors. This issue becomes more difficult to deal with when highly sensitive damage indicative features are used in methods such as the auto regressive (AR) based techniques [8, 9], state-space reconstruction using the delay-coordinate method [10], nonlinear frequency response function technique [11, 12, 13, 14], and energy-based feature extraction using the Hilbert-Huang transform (HHT) [4] are used for SHM. Due to their high sensitivity, these features tend to encode more redun-

dant information. Ignoring the effect of these redundancies may cause problems such as increasing the false-positive (FP) alarms for a given true-positive (TP) rate. This motivates further investigation on new data analysis techniques with the capability of using robust damage sensitive features as well as considering the complex data structure provided by such features.

In this study we address the problem of reducing the FP ratio without significantly affecting the TP ratio in dense sensor networks by considering the spatial dependencies of sensor measurements. Our approach to do that is to use graphical models (GMs) which are strong mathematical tools for representing complex probability distributions and the dependencies between random variables (RV). Appropriately learning and efficiently making inference on GMs are the two main challenges in using these models in practice. Learning GMs is particularly more critical in SHM as it usually involves working with incomplete datasets which do not provide enough information for learning GMs through standard methods [15]. As a solution to this problem, we propose an approximated model for graph parameter learning that works based on the mutual information (MI) between different sensor data. By that, the contributions of this study can be summarized as 1) improving damage localization resolution in dense sensor arrays by using GMs, and 2) proposing an approximated model that uses information analytics to learn the graph parameters. The body of this work starts with formalizing the SHM problem and the objectives of this study. Section 3 reviews undirected GMs as the main mathematical tool that we use in this study and in section 4 we explain the details of our proposed algorithm. As mentioned previously, the main application of this work is to analyze dense sensor networks. Therefore, a discussion on obtaining spatially high resolution datasets with a focus on camera-based data acquisition technique, used in this study, is provided in section 5. Finally, the experimental validation is presented in section 6 by testing various sensor systems and comparing different damage detection algorithms.

2. Formalizing SHM problems

Consider a structure that is instrumented by s number of sensors. Assume the feature matrix associated with the i^{th} sensor data that is obtained by testing the intact structure n_b times is

$$\mathbf{y}_i^b = \left[\underline{y}_{i1}^b, \underline{y}_{i2}^b, \cdots, \underline{y}_{im}^b\right] \tag{1}$$

where m is the total number of damage sensitive features and each $\underline{y}_{il}^b$, $l = 1, \cdots, m$, is a $n_b \times 1$ vector. Note that throughout the paper, RVs are denoted by a sans-serif font, e.g. x, and the realization of those variables are denoted by serifed fonts, e.g. x. Similar to $\mathbf{y}_i^b$, the feature matrix $\mathbf{y}_i^u$ can be obtained from a new set of tests on the structure at an unknown state. If there is information available from some damage scenarios, the corresponding feature matrix is defined similarly and denoted by $\mathbf{y}_i^d$. The superscripts u an d respectively denote *unknown state* and *damaged* . The number of corresponding tests is assumed to be n_u and n_d. In addition to these real-valued quantities, consider a set of discrete RV $\mathbf{x}$ as

$$\mathbf{x} = [\mathsf{x}_1, \mathsf{x}_2, \cdots, \mathsf{x}_s] \tag{2}$$

where x_i, $i = 1, \cdots, s$, takes the values 0 or 1 if the structure is respectively intact or damaged at the i^{th} sensor location. For notation simplicity, define:

$$\mathbf{Y}_S^b = \left\{\mathbf{y}_j^b | j \in S\right\}; S \subseteq \{1, \cdots, s\} \tag{3}$$

which is a set that contains the feature matrices for a subset of sensors, S, for intact state of the structure. Similarly, $\mathbf{Y}_S^d$ and $\mathbf{Y}_S^u$ can be respectively defined by choosing a subset of all feature matrices $\mathbf{y}_i^d$ and $\mathbf{y}_i^u$.

As suggested by other studies on similar problems, consideration of dependencies of RVs may significantly improve the classification results [16, 17]. Therefore, with the above definitions, the objective is considered as computing $p_{\mathsf{x}_i|\mathbf{Y}_{1:s}}, \forall i \in \{1, \cdots, s\}$ which is the probability density of the states at each sensor location given the features at all sensor locations, ($S = \{1, \cdots, s\} = \{1 : s\}$). Fulfilling this objective requires the dependencies of the RVs to be considered.

To do that in this study we use Markov Random Fields (MRF), described in the next section.

3. Graphical models for considering the dependencies of RVs

3.1. Graphical models

GMs are strong mathematical tools for representing complex probability distributions and conditional dependencies between RVs. Depending on the type of dependencies that we want to model, we can use either directed acyclic graphs (DAG), factor graphs, or undirected graphs (UGM). The first two types of GM are mostly used for modeling the induced dependencies between the RVs and their ordering [15]. Such dependencies are not of central importance in this study; thus, we choose UGM or Markov random fields (MRF) as the more appropriate choice to account for the dependencies of RVs in the SHM problem.

An UGM is a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where the vertices (or nodes) $\mathcal{V}$ correspond to RVs and the edges $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ represent the conditional independences as well as the structure of the graph [15]. To further proceed with parametrization of these models we need the notion of a *clique* which is a set of fully connected nodes. A *maximal clique* is a clique which is not a strict subset of another clique. With that, it can be proven that a strictly positive distribution $p_{\mathsf{x}}(\cdot)$ over $\mathsf{x} = \{\mathsf{x}_1, \cdots, \mathsf{x}_s\}$ that satisfies graph separation property [15] can be factorized as [18]

$$p_{\mathsf{x}}(\mathbf{x}) = \frac{1}{Z} \prod_{c \in \mathcal{C}} \pi_{\mathbf{x}_c}(\mathbf{x}_c) \propto \prod_{c \in \mathcal{C}} \pi_{\mathbf{x}_c}(\mathbf{x}_c) \tag{4}$$

where $\mathcal{C}$ is the set of all maximal cliques; Z is the normalizing constant that is called *partition function*, π is a non-negative function, which represents the dependencies between the RVs in its associated maximal clique, referred to as *potentials*, and $\mathbf{x}_c$ is a subset of $\mathbf{x}$ which forms a maximal clique. For more details about the formulation, related terminologies, and comparison of GMs with other multivariate methods the readers are referred to [15, 18, 19]. Sometimes it is

more convenient to reparametrize (4) as

$$p_{\mathbf{x}}(\mathbf{x}) = \frac{1}{Z} \exp\left(\sum_{c \in \mathcal{C}} \psi_{\mathbf{x}_c}(\mathbf{x}_c)\right) \tag{5}$$

where the functions $\psi_{\mathbf{x}_c} = \log \pi_{\mathbf{x}_c}$ are called *log-potential functions* .

The main challenge with the use of such model in practice is the complexity of learning and making inference on GMs which can become computationally intractable as the number of nodes is increased. Therefore, simplifications and approximations are usually involved in making inferences on GMs. One such approximation is considering only the pairwise dependencies between RVs. The resultant model is called pairwise GMs which are particularly used in this study and described in the next section.

3.2. Binary pairwise GMs

A pairwise model is one in which the potential functions are defined over at most two RVs. For further simplification, let's consider binary pairwise models in which the RVs are binary-valued, i.e. $\mathsf{x}_i = \{0, 1\}$. Prior to parameterizing these models, we need to define the *energy of configuration* denoted by $E(\mathbf{x})$ to simplify our formulations. This terminology has been borrowed from statistical physics and is defined as

$$E(\mathbf{x}) = -\sum_{c \in \mathcal{C}} \psi_{\mathbf{x}_c}(\mathbf{x}_c) \tag{6}$$

The energy of binary MRF can be parametrized as

$$E(\mathbf{x}) = -\sum_{i \in \mathcal{V}} \theta_i x_i - \sum_{(i,j) \in \mathcal{E}} \theta_{ij} \left(x_i x_j + (1 - x_i)(1 - x_j)\right) \tag{7}$$

where θ_i and θ_{ij} are respectively called *node potential* and *edge potential*. This specific GM with binary RVs and parametrization of the form (7) is called an *Ising model*. For more information about the properties of binary MRFs, different topological categorizations of Ising models, the exact and approximate inference techniques associated with each category, and learning these models, readers are referred to [15, 20, 21, 19].

4. The proposed algorithm

Our proposed approach to satisfy the above mentioned objective is to use GMs. This approach involves learning an appropriate structure for the GM, estimating its parameters, and efficiently making inference on the learned graph. These tasks will be presented separately in the next three subsections.

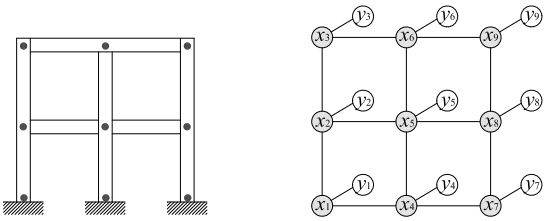
4.1. Graph structure learning

Learning GMs refers to both determining the best structure (configuration) for the graph and estimating the associated parameters. Structure learning of GMs is still an active research area, even with the assumption of availability of complete data. The main challenge in SHM for learning GMs arises from the unavailability of data from all possible damage states, especially for unique structures such as civil infrastructure. Therefore, SHM always involves novelty detection for which, the standard methods of structure learning and parameter estimation [15, 20, 19, 21] do not work.

Throughout this paper, we only consider dense sensor networks for which the configuration of GM can be determined from the geometry of the network. An example of such networks is schematically shown in figure 1(a). In this figure, consider two sensors which are more than one element away from each other. The mechanical waves cannot directly travel between these sensors without passing at least one other sensor location. Therefore, a grid graph, similar to the structure's geometry, over $\mathbf{x}$ could be a reasonable graph configuration which is directly determined based on the physics of the problem, figure 1(b). The intuition for the grid graph over $\mathbf{x}$ is that the effect of damage at each node is first captured at the same node and its neighbors. Clearly, this argument may not be valid for sparse sensor networks where there is more than one direct path between some subsets of the nodes.

In addition to the vertices associated with $\mathbf{x}$, other nodes are drawn in figure 1(b) for the real-valued RV $\mathbf{y}_i$ which are the feature matrices associated with each node. Depending on the state of the system at the i^{th} location, this

RV can be either $\mathbf{y}_i^b$ or $\mathbf{y}_i^d$. Note that for the moment we assume that $\mathbf{y}_i^d$ is known; however, this assumption may not be valid in general and the solution to that will be addressed in section 4.2. The observation $\mathbf{y}_i$ depends on the state x_i and hence, we put an edge between these two. However, there is no edge between $\mathbf{y}_i$ and $\mathbf{y}_j$ for $\forall i, j \in \mathcal{V}$ since the measurements are assumed to be taken independently at different sensor locations.



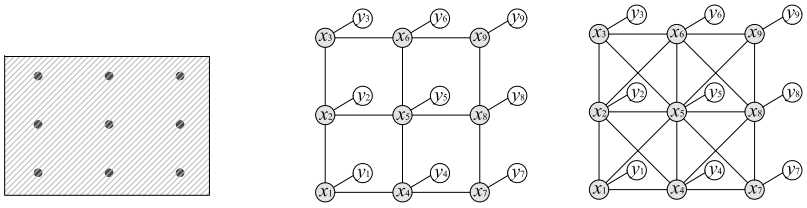
(a) (b)

Figure 1: Simple structure for which the structure of graphical model can be determined by the physics; (a) a frame structure with a sensor node at each connection, (b) corresponding GM for the structure

The sensor network shown in figure 1 may not be widely used in real world frame structures, but similar networks are quite common for monitoring plate structures as shown in figure 2(a). A grid graph as shown in figure 2(b) can be considered for this network analogous to the frame structure. Alternatively, additional cross edges may be added to the grid graph to account for some of the direct paths between sensors in the plate continuum, figure 2(c). This graph is not a standard Ising model since the maximum size of the maximal cliques exceeds two and hence, assuming the same configuration energy for this MRF as in (7) is not exact.

4.2. Graph parameter learning

The parameters associated with a binary pairwise GM are the node and edge potentials. The node potential, θ_i, is in fact a prior about the current state of system at node i. In SHM, there is usually no such information about the possible damage location before analyzing the data and hence, this parameter



(a) (b) (c)

Figure 2: Possible GMs for monitoring plate structures; (a) a typical sensor network for monitoring plate structures, (b) an standard Ising model with a grid graph, (c) an alternative to the standard Ising model for modeling such networks

can be neglected. This implies that all sensors are equally likely to be around the potential damage. The resultant GM is called a *zero-field* Ising model with configuration energy

$$E_{zf}(\mathbf{x}) = -\sum_{(i,j)\in\mathcal{E}} \theta_{ij}\left(x_i x_j + (1-x_i)(1-x_j)\right) \tag{8}$$

where the subscript zf denotes a zero-field model. In presence of information about vulnerability of nodes to damage, non-zero node potentials can be used.

The other parameter for the GMs is edge potential. There are two types of such potentials in the proposed GMs which are shown in figures 1 and 2: 1) the edge potentials for the discrete-valued grid graph over x, 2) the potentials associated with the edges between x_i and $\mathbf{y}_i$. For estimating these parameters, consider the joint distribution $p_{\mathbf{x},\mathbf{Y}_{1:s}}(\cdot,\cdot)$ where $\mathbf{Y}_{1:s} = \{\mathbf{y}_1, \cdots, \mathbf{y}_s\}$. Based on the Bayes' rule, this joint distribution can be factored as

$$p_{\mathbf{x},\mathbf{Y}_{1:s}}\left(\mathbf{x}, \mathbf{Y}_{1:s}\right) = p_{\mathbf{x}}\left(\mathbf{x}\right) p_{\mathbf{Y}_{1:s}|\mathbf{x}}\left(\mathbf{Y}_{1:s}|\mathbf{x}\right) \tag{9}$$

For simplicity, define

$$\psi(x_i, x_j) = \exp\left(\theta_{ij}\left(x_i x_j + (1-x_i)(1-x_j)\right)\right) \tag{10}$$

$\psi(x_i, x_j) = 1$ for $x_i \neq x_j$ and $\psi(x_i, x_j) = e^{\theta_{ij}}$ when the converse is true. Using (5)-(7), (8), and (10) the following proportionality is obtained

$$p_{\mathbf{x}}\left(\mathbf{x}\right) \propto \prod_{(i,j)\in\mathcal{E}} \psi(x_i, x_j) \tag{11}$$

Based on the MRF properties, the conditional density in (9) can be factorized as [20, 16, 17]

$$p_{\mathbf{Y}_{1:s}|\mathbf{x}}(\mathbf{Y}_{1:s}|\mathbf{x}) = \prod_{i\in\mathcal{V}} p_{\mathbf{y}_i|\mathbf{x}_i}(\mathbf{y}_i|x_i) \tag{12}$$

The objective, which is the posterior density of $\mathbf{x}$ given the features, follows the proportionality of

$$p_{\mathbf{x}|\mathbf{Y}_{1:s}} \propto p_{\mathbf{x},\mathbf{Y}_{1:s}} \tag{13}$$

Substituting (12) and (11) into (9) and then considering (13) one obtains

$$p_{\mathbf{x}|\mathbf{Y}_{1:s}}(\mathbf{x}|\mathbf{Y}_{1:s}) \propto \prod_{i\in\mathcal{V}} p_{\mathbf{y}_i|\mathbf{x}_i}(\mathbf{y}_i|x_i) \prod_{(i,j)\in\mathcal{E}} \psi(x_i,x_j) \tag{14}$$

To make an inference on the conditional posterior in (14), both $p_{\mathbf{y}_i|\mathbf{x}_i}, \forall i \in \mathcal{V}$ and $\psi(x_i,x_j), \forall (i,j) \in \mathcal{E}$ should be determined. Note that $p_{\mathbf{y}_i|\mathbf{x}_i}$ should be known for all possible states. For the intact state at each node this conditional density, denoted by $p_{\mathbf{y}_i|x_i=0}$, can be obtained simply by estimating the density of $\mathbf{y}_i^b$ which is defined in (1). If the tests on the intact structure are repeatable, this density can be approximated by a multivariate Gaussian distribution. For a general scenario and depending on the size of data set, either GMM or Kernel density estimate (KDE) can be used for density estimation. The approximated density is respectively denoted by $p^{gmm}_{\mathbf{y}_i|x_i=0}$ and $p^{kde}_{\mathbf{y}_i|x_i=0}$ for GMM and KDE estimations. The main issue in graph parameter estimation for SHM is to estimate $p_{\mathbf{y}_i|x_i=1}$ and $\psi(x_i,x_j)$ in the absence of information from the damaged state of the structure. This issue is discussed next.

4.2.1. Determining $p_{\mathbf{y}_i|x_i=1}$

For mass produced structures, this likelihood function can be obtained by damaging and testing some structures at randomly selected instances during production; however, unique structures, such as much of civil infrastructure, cannot be damaged for testing. In this case, information from similar structures can be used if there is any. If no information is available from the damaged state of the structure, one may consider a uniform distribution for the features over their domain based on the maximum entropy principle. Some features, such as

the one proposed in [4], have a bounded definition domain. For others with an unbounded domain, an interval should be appropriately chosen for the features to ensure that the interval always contains the features. Such an interval can be obtained by modeling the damaged structure and extracting the features from the predicted responses of the model or other numerical techniques such as the ones described in [22]. Note that, what has been mentioned so far was a general recipe and depending on the problem and understanding the dynamics of the structure, more accurate likelihood models can be obtained for the features in damaged state.

4.2.2. Learning $\psi(x_i, x_j)$

Consider a binary GM with only two vertices x_1 and x_2. Based on the proportionality of (11), we have

$$p_{\mathsf{x}_1,\mathsf{x}_2}(x_1, x_2) \propto \psi(x_1, x_2) \tag{15}$$

Noting that $\psi(x_i, x_j)$ only considers the similarity of its input arguments, the following joint probabilities are obtained:

$$\mathbb{P}(1,0) = \mathbb{P}(0,1) = \frac{1}{2\left(1 + e^{\theta_{12}}\right)}$$

$$\mathbb{P}(1,1) = \mathbb{P}(0,0) = \frac{e^{\theta_{12}}}{2\left(1 + e^{\theta_{12}}\right)}$$

Therefore, $\psi(x_1, x_2)$ can be any unnormalized density mass function of the form

$$\psi(x_1, x_2) = \begin{cases} \frac{\alpha e^{\theta_{12}}}{1+e^{\theta_{12}}} & x_1 = x_2 \\ \frac{\alpha}{1+e^{\theta_{12}}} & x_1 \neq x_2 \end{cases} ; \alpha \neq 0 \tag{16}$$

The value of α is not significant for the sake of inference; thus, consider $\alpha = 1$ for simplicity. With this assumption, $\psi(x_1, x_2)$ becomes a probability mass function which implies that with the probability of $e^{\theta_{12}}/(1 + e^{\theta_{12}})$ the state of one node is reflected on its neighboring node. Generalizing this idea for a binary pairwise GM with more than two nodes, the following form for $\psi(x_i, x_j)$ is obtained

$$\psi(x_i, x_j) = \begin{cases} q_{ij} & x_1 = x_2 \\ 1 - q_{ij} & x_1 \neq x_2 \end{cases} ; 0 \leq q_{ij} \leq 1 \tag{17}$$

where q_{ij} is the probability that the two neighboring nodes are in the same state. In the context of SHM, this essentially implies with probability q_{ij}, the damage can be sensed at the neighboring nodes of its actual location.

Estimating q_{ij} in supervised problems are conducted either by a maximum likelihood (ML) approach [15] or through a grid search [17]. Both of these approaches fail in a general SHM problem because of incomplete data from unobserved states. Therefore, the interaction strength between the connected vertices cannot be determined using the available information about $\mathbf{x}$.

Based on the concept of q_{ij} that is a measure of similarity of data in neighboring nodes, we propose using mutual information (MI) between the neighboring nodes in the real-valued feature space to model this similarity. For that, consider $p^{gmm}_{\mathbf{y}_i|x_i=0}$ and $p^{gmm}_{\mathbf{y}_j|x_j=0}$ for two arbitrary nodes i and j where $(i,j) \in \mathcal{E}$ and define the function $g(\cdot,\cdot)$ as

$$g(\mathbf{y}_i, \mathbf{y}_j) = \left(\frac{p^{gmm}_{\mathbf{y}_i,\mathbf{y}_j|x_i=x_j=0}(\mathbf{y}_i, \mathbf{y}_j)}{p^{gmm}_{\mathbf{y}_i|x_i=0}(\mathbf{y}_i)\, p^{gmm}_{\mathbf{y}_j|x_j=0}(\mathbf{y}_j)} \right) \tag{18}$$

where $p^{gmm}_{\mathbf{y}_i,\mathbf{y}_j|x_i=x_j=0}$ is the joint distribution of the features at nodes i and j when both of them are intact. It can be proven that the MI between $\mathbf{y}^b_i$ and $\mathbf{y}^b_j$ can be estimated as [23]

$$I(\mathbf{y}^b_i, \mathbf{y}^b_j) \approx \frac{1}{N_{mi}} \sum_{k=1}^{N_{mi}} \log g\left(\mathbf{y}^{(k)}_i, \mathbf{y}^{(k)}_j\right) \tag{19}$$

where N_{mi} is the number of samples for Monte Carlo (MC) estimation of MI and $\left(\mathbf{y}^{(k)}_i, \mathbf{y}^{(k)}_j\right)$ is the k^{th} sample drawn from $p^{gmm}_{\mathbf{y}_i,\mathbf{y}_j|x_i=x_j=0}$. For convenience, we use normalized MI (NMI) which is defined as

$$R(\mathbf{y}^b_i, \mathbf{y}^b_j) = \frac{I(\mathbf{y}^b_i, \mathbf{y}^b_j)}{\sqrt{H(\mathbf{y}^b_i) H(\mathbf{y}^b_j)}} \tag{20}$$

where $H(\cdot)$ denotes the entropy which can be estimated similar to MI via MC method [23]. The NMI is zero if the two RVs are independent and it becomes unity when the two RVs are deterministic function of one another. When $R(\mathbf{y}^b_i, \mathbf{y}^b_j) = 0$, availability of full information about the state at node i does not reduce the uncertainty about the state at node j and vise versa. Thus, by

knowing the state at either of these nodes the probability of guessing the state at the other one is 0.5. The other extreme case is when $R(\mathbf{y}_i^b, \mathbf{y}_j^b) = 1$ which implies that $\mathbf{y}_i^b$ is a deterministic function of $\mathbf{y}_j^b$. In this case, knowing the state at one node tells the state at the other with probability of 1.0. Therefore, the parameter q_{ij} can be approximately estimated by linear interpolation between these two extreme cases as

$$q_{ij} = 0.5\left(1 + R(\mathbf{y}_i^b, \mathbf{y}_j^b)\right) \tag{21}$$

Note that this approach intrinsically assumes that the similarity between x_i and x_j can be captured by $I(\mathbf{y}_i^b, \mathbf{y}_j^b)$. The first issue with this assumption is associated with approximating the interactions between the discrete-values states $\mathbf{x}$ in the continuous-valued feature space which is not consistent with the graph structures that mentioned in section 4.1. The second issue arises from using only the baseline data for determining the MI which may not result in accurate estimation of this quantity. The first problem is a challenge, but the second one can become less of an issue if we keep enriching the baseline database by taking measurements in various conditions.

4.3. Making inferences on GMs

So far, we have described how to estimate the parameters of GMs with incomplete data. In this section we discuss two inference techniques, which are loopy belief propagation (LBP) and Markov Chain Monte Carlo (MCMC), for computing the posterior probability of the states at each node given the features. LBP [24] is an efficient approximated method which may result in inconsistent marginal distributions [25]. On the other hand, MCMC is asymptotically exact, but computationally intensive for large GMs. The computational cost of MCMC methods may not be an issue for SHM due to the small number of sensors that is used in these problems. Therefore, we use and compare the accuracy of both method in this study.

4.3.1. Loopy belief propagation

Recall that $\mathbf{y}_i^b$ is the baseline feature matrix at node i; $\mathbf{y}_i^d$ is the feature matrix for the damaged state which may not be observed, but its associated conditional density can be obtained by the maximum entropy principal or other techniques; $\mathbf{y}_i^u$ is the feature matrix of the new measurements. Denoting the mean value of $\mathbf{y}_i^u$ for different test by $\overline{\mathbf{y}}_i^u$ the LPB algorithm can be described as follows

1) Initialization:

$$\begin{cases} m_{i\to j}^0(x_j) \propto 1 \\ m_{j\to i}^0(x_j) \propto 1 \end{cases} ; \forall (i,j) \in \mathcal{E} \tag{22}$$

2) Message passing:

$$m_{i\to j}^{t+1}(x_j) = \sum_{x_i} \left(p_{\mathbf{y}_i|\mathsf{x}_i}(\overline{\mathbf{y}}_i^u|x_i)\psi(x_i,x_j) \prod_{(i,k)\in\mathcal{E}, k\neq j} m_{k\to i}^t \right) \tag{23}$$

where the superscript t denotes the iteration. For numerical stability, the messages are normalized at each step, i.e. $\sum_{x_j} m_{i\to j}^t(x_j) = 1$.

3) Marginal probability estimation:

After message passing convergence, the marginal beliefs are computed as

$$p_{\mathsf{x}_i|\mathbf{y}_i}(x_i|\mathbf{y}_i) = c_i p_{\mathbf{y}_i|\mathsf{x}_i}(\overline{\mathbf{y}}_i^u|x_i) \prod_{(i,j)\in\mathcal{E}} m_{j\to i}^t(x_i) \tag{24}$$

where c_i is a normalizing constant and $p(\cdot|x_i)$ is as defined earlier.

4.3.2. MCMC and Gibbs sampling

Among different MCMC techniques, we have chosen Gibbs sampling due to the ease of computing the conditional dependencies for pairwise MRFs. This method starts with a valid realization for $\mathbf{x}$ and randomly choosing x_i. Then sampling the new x_i from the conditional distribution of $p_{\mathsf{x}_i|\mathbf{x}_{-i},\mathbf{Y}}(\cdot|\cdot)$ where $\mathbf{x}_{-i} = [\mathsf{x}_1, \cdots, \mathsf{x}_{i-1}, \mathsf{x}_{i+1}, \cdots, \mathsf{x}_s]$. Using the MRF property that a node is independent of other nodes given its neighbors, this conditional distribution can be written as

$$p_{\mathsf{x}_i|\mathbf{x}_{-i},\mathbf{Y}}(x_i|\mathbf{x}_{-i},\mathbf{Y}) \propto \prod_{(i,j)\in\mathcal{E}} p_{\mathbf{y}_i|\mathsf{x}_i}(\overline{\mathbf{y}}_i^u|x_i)\psi(x_i,x_j) \cdot f(\mathbf{x}_{-i}) \tag{25}$$

where $f(\mathbf{x}_{-i})$ is a function independent of x_i and hence, cancels out when computing the marginal probability of x_i. Thus, the procedure of Gibbs sampling for this problem is as follows:

1) Initialization: Start with a valid realization of $\mathbf{x}$.

2) Choosing the variable to sample from: randomly sample i from $\{1, \cdots, s\}$.

3) Sampling: sample the new x_i from the distribution that is described in (25) and save the new $\mathbf{x}$.

As previously described, these methods are meant to work with dense sensor networks or arrays where the GM can be determined from the geometry of the network, and these inference techniques have varying degrees of exactness and computational speed as a tradeoff. In the next section, we describe one technique for obtaining high spatial density data from a structure.

5. Obtaining spatially high resolution datasets

Dense sensor networks or arrays can be quite expensive to implement, especially with traditional vibration measurement sensors like wired accelerometers. For very small structures, such as plate elements, a large number of accelerometers can be overwhelming and the added mass can alter the dynamics of the system. For large structures, installation and wiring may become difficult. Especially dense networks or arrays can quickly become expensive with the large number of sensors required. More specialized equipment such as laser vibrometers are capable of measuring the velocities at many locations on an object, however multi-beam or scanning mirror setups can cost hundreds of thousands of dollars, reasonable for laboratory equipment, but too expensive for widespread field usage. Video cameras on the other hand, can be relatively inexpensive and collect data from a large scene of interest. Each pixel or group of pixels on a structure can be a pseudo sensor. A major issue is that the video data that is collected is not directly convertible into raw displacements, velocities, or accelerations, and needs some advanced processing to determine the motion of the recorded object. In this paper, a local phase-based optical flow approach

presented in [6] and [7] is used to calculate the displacements.

5.1. Processing videos into displacements

To process videos into displacement data, the general idea is to extract the data from the raw signals in the video, pixel intensities through time. There are two general approaches that are best explained by an analogy to fluid mechanics, Lagrangian or Eulerian. A Lagrangian approach tracks pixels or particles as they move in the scene while an Eulerian approach stays at one pixel or one position and records the flow vectors. An example of a Lagrangian approach is digital image correlation which matches patches of pixel intensity patterns between different frames of the video to compute motion vectors [26]. An example of an Eulerian approach is phase-based optical flow, where the local phase signal is extracted from the video to obtain the optical flow field [27, 28]. In this paper, the phase-based optical flow approach is used to calculate the displacements of the object in the video.

The reason a phase-based approach is used rather than one based on amplitude or pixel intensities is because it is typically more robust to noise. A summary of the procedure is as follows. Filters are used to decompose the video into local amplitude and local phase, quantities analogous to the complex coefficients obtained with a Fourier transform. The complex steerable pyramid filters are used to obtain the local amplitude and phase at multiple orientations and physical length scales [29]. Previous work has shown that the motion of constant phase contours corresponds to motion signals in the video, and we use the decomposed local phase signals to calculate the displacement signal at every pixel in the video [27, 28]. More details about the video decomposition and the calculation of displacements is contained in [6] and [7].

5.2. Assumptions

There are a few important assumptions in calculating the displacement signals from objects in video. The motions must stay small, on the order of one pixel or smaller. The local filters stop working if the motion becomes too large.

Textureless regions do not give reliable motion information for the object; displacements are only well defined at edges or textured regions in the video. In order to obtain displacement information for the object under test, it is spray painted with a speckle pattern to increase the visual texture of the object. To apply the speckle pattern, the object is first painted white, and then a random pattern of black paint is sprayed on in a spotty manner. Without the speckle pattern, the only reasonable displacements would come from edges of the object against a contrasting background.

Another assumption for the measurement is that the lighting is constant. Flickering lights, such as fluorescent lighting, introduce an apparent motion signal into the video with the same variation in time as the lighting. To prevent this from happening, the object under test is flood illuminated with several bright battery powered lamps so that the lighting stays consistent.

As long as these assumptions are taken into consideration, the phase-based approach to measure optical flow reliably gives the displacements of an object. As Ising graphical models were originally designed for processing of image based information, applying them to displacement data from video cameras for the purpose of SHM, is only logical.

It is also noted that with this data acquisition method the measurements can only be analyzed in a centralized manner due to the lack of interconnectivity of pseudo sensors. Thus, the group of pseudo sensors form a sensor array rather than a sensor network.

6. Experimental validation

The methodology presented here is applicable to many different types of structures. Results are presented for a model frame structure and a plate structure.

Three algorithms are compared in this part for monitoring the structures in this section. The first algorithm is a sensor-wise damage detection which performs a likelihood test of new measurements, $\mathbf{y}_i^u$, on the baseline and damaged

distribution model which are $p^{gmm}_{\mathbf{y}_i|x_i=0}$ and $p_{\mathbf{y}_i|x_i=1}$. For this likelihood test we use the notion of high density region (HDR) of multi-modal probability distributions. For more information on that the readers are referred to [30]. The next two algorithms are the GM-based techniques which are proposed in this paper. The second algorithm uses LBP for making inferences on GMs while the third one employs Gibbs sampling.

In order to compare the above-mentioned algorithms, we define an evaluation criteria based on FP and true-positive (TP) rates. The definition of FP and TP is controversial in SHM application as it is unlikely that the damage occurs exactly at the sensor locations. Therefore, we define the FP in SHM as the detection of damage at sensor locations which may neither coincide with the damage location nor the closest neighboring sensor locations. The TP is also defined as detecting the damage at the sensor location which coincides with the damage or its closest neighboring sensor locations. Assume $I = \{1, \cdots, n\}$ with n to be the number of sensors in a network, D_a is the set of sensor locations which are either at the damage location or its closest neighborhood, and D_a^c is the complement of D_a with respect to I. Also assume that D_s and D_g are the set of damaged locations which are detected by the sensor-wise method and the proposed graph -based algorithm, respectively. Then, for a new measurement the FP reduction (FPR) and TP increase (TPI) criteria are defined as

$$FPR = \frac{|D_s \cap D_a^c| - |D_g \cap D_a^c|}{|D_a^c|} \tag{26a}$$

$$TPI = \mathbf{1}_{D_g \cap D_a \neq \emptyset} - \mathbf{1}_{D_s \cap D_a \neq \emptyset} \tag{26b}$$

where $|\ .\ |$ is the cardinality of a set, and $\mathbf{1}_{(\ .\ )}$ is the indicator function. FPR essentially shows how much the FP is reduced by using the proposed method compared to the sensor-wise data analysis. TRI shows which algorithm can/cannot localize the damage. We use $\overline{FPR}$ and $\overline{TPI}$ to show the mean value of the FPR and TPI, respectively, for all tests on each damage scenario.

6.1. Model frame structure

6.1.1. Experimental Setup

A modular steel structure was tested in a 3 story 2 bay configuration. It consists of columns that are 60 cm × 5.08 cm × 64 cm and beam frames of similar dimensions for each story. The parts are bolted together at each connection with four bolts and the whole structure is bolted to a heavy concrete foundation as a reaction mass. The structure is shown in figure 3(a) with the connections numbered, and a typical bolted connection is shown in figure 3b. To measure the vibration response of the structure, 57 MEMS accelerometers, based on the ST Microelectronics LIS344ALH chip, were attached at locations near the 18 connections and mid span between the connections. The acceleration time signals were measured in the flexible direction of the structure, sampled at 2000 Hz. To excite the structure a shaker with a weight of 0.91 kg and a piston weight of 0.17 kg was attached to the top corner of the structure at node 18. The shaker provided a random white Gaussian noise in the frequency range of 5 - 350 Hz in the flexible direction.

In addition to testing the intact structure, there were two separate damage scenarios. The first damage scenario is introduced by substituting an element between nodes 2 and 3 with a reduced cross section. In the second damage scenario, which involves multiple damaged location, the connection bolts at node 7 are loosened in addition to substitution on a reduced cross section element as in the first scenario. These scenarios were tested in both an 18 sensor configuration, using only data from sensors at the connections, and a full 57 sensor configuration. Table 1 summarizes the damage scenarios.

6.1.2. Results

For considering the spatial dependencies between sensor data in this structure, we considered a 3D grid graphical model similar to the geometry of the structure, (figure 4). This configuration for the model structure follows the same rationale as considering a 2D grid graph for the structure shown in figure 1. The damage detection/localization results of the three scenarios listed

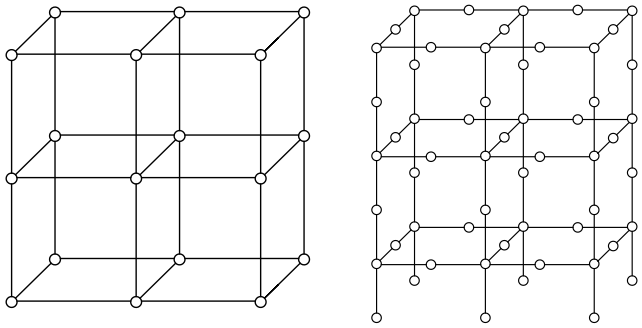
(a) (b)

Figure 3: Pictures showing (a) the laboratory structure, with the connections numbered and shaker labeled with an arrow, and (b) a picture of a typical bolted connection

Table 1: Damage scenarios for the model frame structure

Scenario	Damage Scenario	No. of Sensors
1	Reduced cross section	57
2	Reduced cross section and loose bolts	57
3	Reduced cross section and loose bolts	18

in table 1 are respectively shown in figures 5, 6, and 7. Note that the output of the sensor-wise algorithm, which is shown in part (b) of these figures, are normalized likelihood scores, while the proposed algorithm with GMs provides probability of damage. Table 2 shows these results in terms of $\overline{FPR}$ and $\overline{TPI}$. It follows from these results that all three algorithms are capable for detecting the anomaly; however, the sensor-wise algorithm and GM with LBP suffers from high false-positives rates. Moreover, all three scenarios interestingly show that the approximations in LBP may result in lower localization accuracy compared to sensor-wise damage detection that ignores the spatial dependencies of sensor data. This could be due to either the inherent simplified assumptions in LBP or the 3D configuration of the graph. Contrary to the first two algorithms, using GMs combined with MCMC method for inference provides the most accurate and yet robust approach for damage localization.



(a) (b)

Figure 4: Graph configuration for the model structure

Table 2: $\overline{FPR}$ and $\overline{TPI}$ as the result of using the proposed algorithm for monitoring the frame structure

	LBP		Gibbs Sampling	
Scenario	$\overline{FPR}(\%)$	$\overline{TPI}(\%)$	$\overline{FPR}(\%)$	$\overline{TPI}(\%)$
1	-4.99	0.0	14.21	0.0
2	-11.21	0.0	9.36	0.0
3	1.23	0.0	13.88	0.0

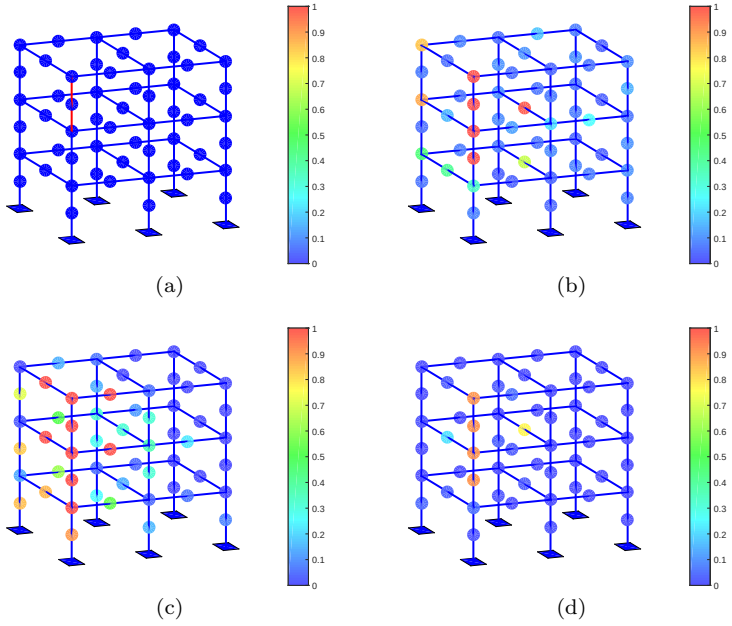


Figure 5: Damage detection result for the first scenario on the model structure. a) the actual location of damage; b) sensor-wise method with likelihood test and GMM; c) damage detection using GM and LBP; d) damage detection using GM and Gibbs sampling

6.2. Plate

6.2.1. Experimental Setup

The experimental setup for measuring the plate involved a shaker, high-speed camera, and extra lighting, as shown in figure 8. The plate, steel with dimensions of 60 cm × 5.08 cm × 0.64 cm, is fixed to a massive concrete base using four bolts, and the shaker is attached to the top of the plate. The shaker excited the plate with a white Gaussian noise waveform in a horizontal direction in the video. After allowing some time for the excitation to reach steady state, 3.5 seconds of video was recorded at 2000 frames per second with a resolution of 1736×244 pixels using the high speed camera. This was repeated for both the intact plate and the damaged plate with a machined crack towards its base. The details of the intact and notched plats, and screenshots from the input videos are shown in figure 9. The displacements were calculated using the procedure as described in section 5, and divided in time into 7 different tests consisting of

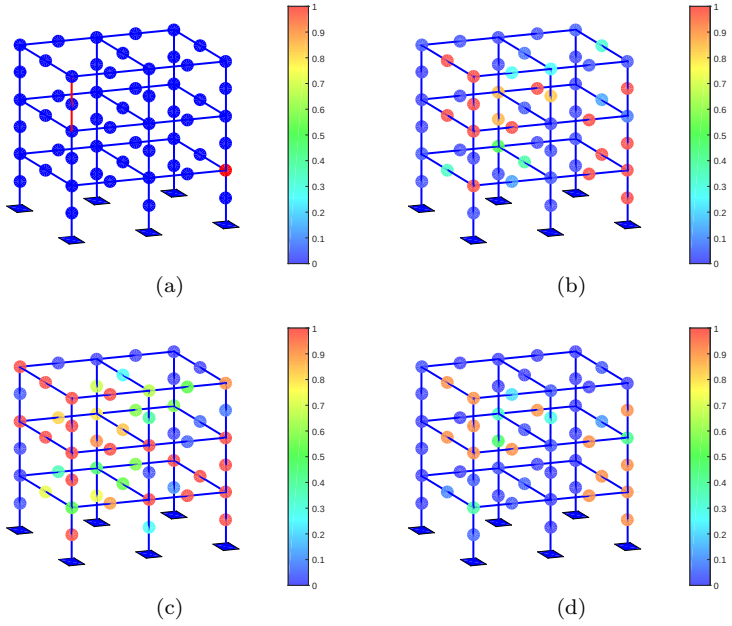


Figure 6: Damage detection result for the second scenario on the model structure. a) the actual location of damage; b) sensor-wise method with likelihood test and GMM; c) damage detection using GM and LBP; d) damage detection using GM and Gibbs sampling

0.5 seconds of video, or 1000 time points each.

6.2.2. Results

Analyzing the extremely high spatial resolution data provided by camera is computationally intractable. Therefore, we picked 100 pixels on a cropped image of the plate to be used as pseudo sensors. The coordinates of these points on both the intact and the notched plates should be roughly the same. Figure 10 shows the location of pseudo sensors on the notched plate as well as the GMs that we have used for considering the spatial dependencies of sensor data. Although the geometry of the two GMs are quite different it turns out that the final results do not have much of a difference; we present only the results for the configuration that is shown in figure 10(b).

The damage localization results of the plate for the three aforementioned algorithms are shown in figure 11. The difference between the algorithms is more obvious for this structure. As shown in figure 11(b), the sensor-wise ap-

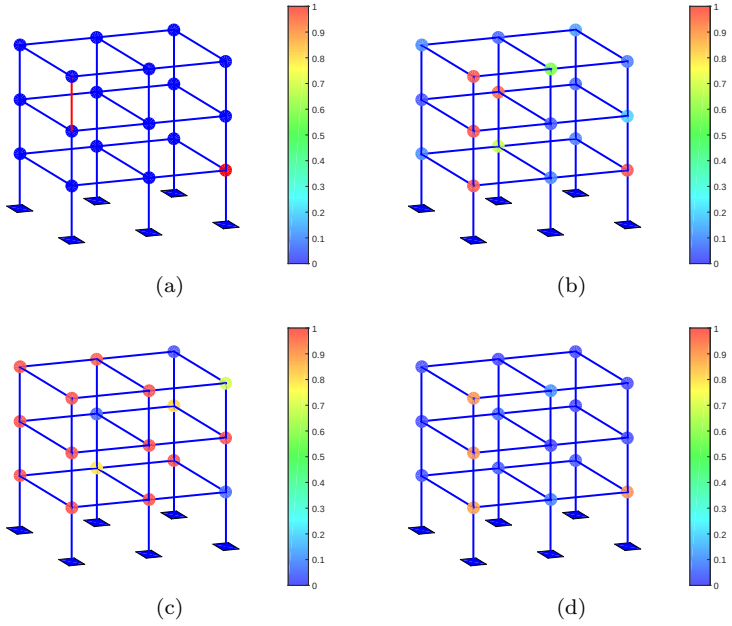


Figure 7: Damage detection result for the third scenario on the model structure. a) the actual location of damage; b) sensor-wise method with likelihood test and GMM; c) damage detection using GM and LBP; d) damage detection using GM and Gibbs sampling

proach does not accurately localize the damage. Using GM combined with the efficient LBP method provides more accurate localization compared to the sensor-wise approach (figure 11(c)). Although the damaged zone predicted by LBP is large, the actual damage location is contained in the predicted damaged zone and distinguishable from the intact zones. Figure 11(d) shows the damage detection results using GM combined with MCMC which provides perfect localization by assigning the highest probability at the tip of the notch. Table 3 also compares the results of the three algorithms in terms of FP reduction and TP improvement. The results show that the damage can be detected by all of the algorithms; however, considering the spatial dependencies of the system response can improve the localization accuracy.

6.3. Discussion

Based on the experiments with different structures and different detection

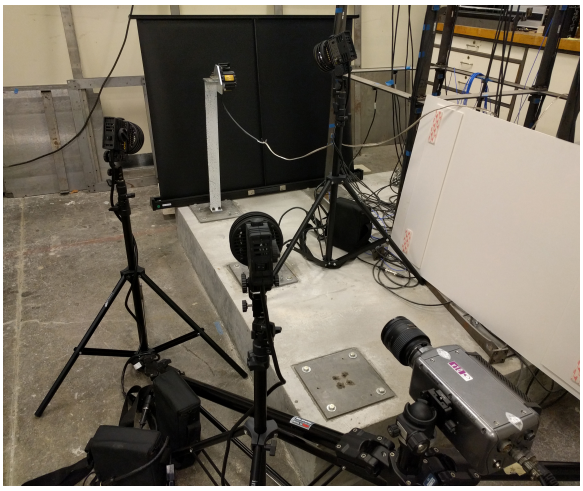


Figure 8: Picture of the experimental setup for the plate, showing the plate fixed to a concrete base, the shaker bolted to the top of the plate, high-speed camera, and extra lighting

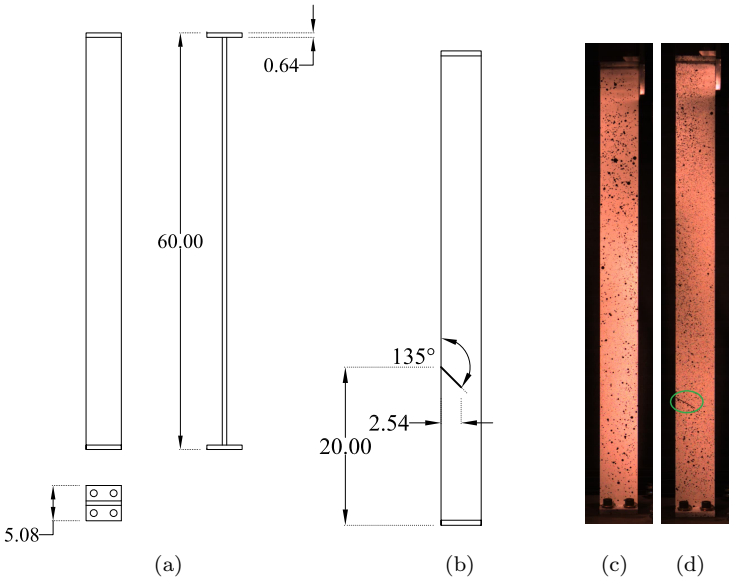
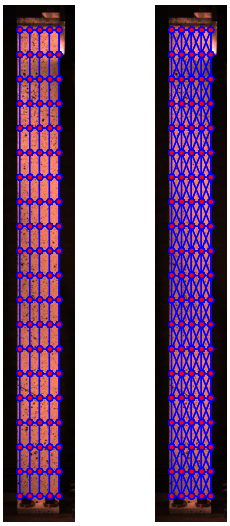
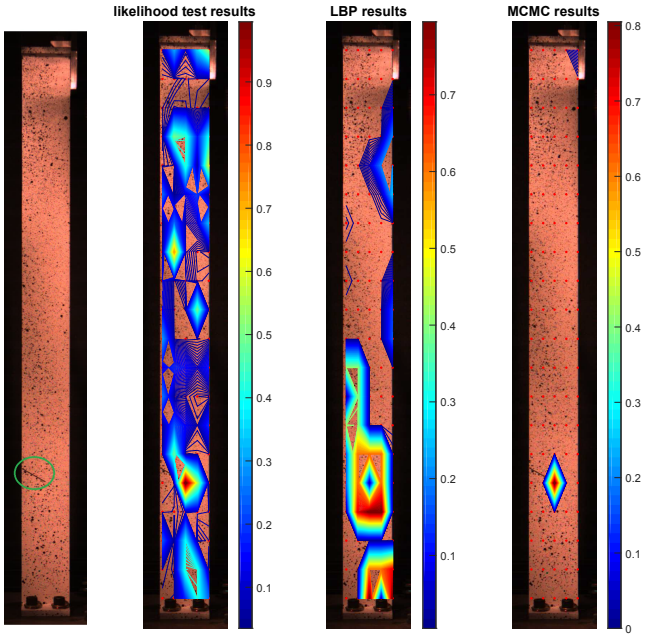


Figure 9: Details of the plate structures (dimensions are in centimeters): a) detail of the plates, b) location of the machined crack on the damaged plate, c) screenshot of input video for the intact plate, (d) screenshot of input video for the damaged plate with crack



(a) (b)

Figure 10: Pseudo sensor locations and different GMs that are considered for the plate (a) simple grid graph, (b) grid graph with cross edges.



(a) (b) (c) (d)

Figure 11: Damage detection result for the plate. a) the actual location of damage; b) sensor-wise method with likelihood test and GMM; c) damage detection using GM and LBP; d) damage detection using GM and Gibbs sampling

Table 3: $\overline{FPR}$ and $\overline{TPI}$ as the result of using the proposed algorithm for damage localization in the notched plate

	LBP		Gibbs Sampling	
Scenario	$\overline{FPR}(\%)$	$\overline{TPI}(\%)$	$\overline{FPR}(\%)$	$\overline{TPI}(\%)$
Notched plate	3.76	0.0	14.11	0.0

cases, the following observations can be made on the performance of the proposed method. 1) LBP on the planar pairwise GM of the plate performs better than the non-planar GM of the steel structure. We believe that the improved performance of the algorithm for the plate is due to the increased number of sensors as well as the configuration of the sensor array. For details of approximate inference algorithms for planar and non-planar GMs the readers are referred to [19, 20]. 2) By increasing the extent of damage, LBP may result in lower performance compared to sensor-wise technique, while Gibbs sampling is not sensitive to the damage extent. 3) The improvement of the proposed method in comparison to the sensor-wise techniques also depends on the damage scenario; e.g. localization accuracy is more improved for the reduced cross-section element than the loosened bolts.

7. Conclusion

In this work, we investigated the effect of spatial dependencies between sensor data on the accuracy of damage localization in dense sensor networks. We proposed an algorithm that employs pairwise GMs to compute the posterior probability of states (intact/damaged) at each sensor location given baseline data sets and new measurements from the entire sensor nodes on the structure. As part of this general approach, we also proposed a method for learning graph parameters with incomplete data and missing states. Finally, experimental validation was conducted by testing two different structures instrumented by MEMS accelerometers and a video camera, respectively. The results of the proposed algorithm were compared with a sensor-wise damage detection method that ignores the dependencies between sensor data.

Based on the experimental results, considering the spatial dependencies of sensor data can significantly improve damage localization accuracy, however, the robustness of the proposed technique depends on the method for making inferences on GMs. The efficient LBP method suffers from lack of robustness in SHM applications as it may negatively affect damage localization accuracy. On the other hand, MCMC and in particular Gibbs sampling, is by far the most robust and yet accurate method for damage localization based on our results. Combining the proposed technique with Gibbs sampling provided almost perfect localization in all tested structures and networks.

The proposed algorithm is appropriate for dense sensor networks or arrays for which the graph structure can be intuitively obtained from the physics of the problem. The limitations of the proposed method includes 1) the lack theoretical guarantee to ensure that such graph structures best model the spatial dependencies of sensor data, and 2) the curse of dimensionality associated with the density estimation required for obtaining the node potentials of the GM. Future work would involve the extension of this method for analyzing general sensor arrays as well as handling problems with small data sets.

8. Acknowledgment

The authors acknowledge the support provided by Royal Dutch Shell through the MIT Energy Initiative, and thank chief scientists Dr. Dirk Smit and Dr. Sergio Kapusta.

References

[1] S. Yin, B. Epureanu, Structural health monitoring based on sensitivity vector fields and attractor morphing, Phylosofical Transactions of Royal Society A 364 (2006) 2515–2538.

[2] K. Worden, C. Farrar, J. Haywood, M. Todd, A review of nonlinear dynamics applications to structural health monitoring, Journal of Structural Control and Health Monitoring 15 (2008) 540–567.

[3] N. Huang, Z. Shen, S. Long, M. Wu, H. Shih, Q. Zheng, N. Yen, C. Tung, H. Liu, The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis, Proceedings of the Royal Society Series A (454) (1998) 903–995.

[4] R. M. Ghazi, O. Buyukozturk, Damage detection with small data set using energy-based nonlinear features, Journal of Structural Control and Health Monitoring`doi:10.1002/stc.1774`.

[5] H.-Y. Wu, M. Rubinstein, E. Shih, J. V. Guttag, F. Durand, W. T. Freeman, Eulerian video magnification for revealing subtle changes in the world., ACM Trans. Graph. 31 (4) (2012) 65.

[6] N. Wadhwa, M. Rubinstein, F. Durand, W. T. Freeman, Phase-based video motion processing, ACM Trans. Graph. (Proceedings SIGGRAPH 2013) 32 (4).

[7] J. G. Chen, N. Wadhwa, Y.-J. Cha, F. Durand, W. T. Freeman, O. Buyukozturk, Modal identification of simple structures with high-speed video using motion magnification, Journal of Sound and Vibration 345 (2015) 58–71.

[8] K. K. Nair, A. Kiremidjian, K. Law, Time series-based damage detection and localization algorithm with application to the asce benchmark structure, Journal of Sound and Vibration 291 (2006) 349–368. `doi:10.1016/j.jsv.2005.06.016`.

[9] K. K. Nair, A. Kiremidjian, Time series based structural damage detection algorithm using gaussian mixtures modeling, Journal of Dynamic Systems, Measurement, and Control 129 (2007) 285–293. `doi:10.1115/1.2718241`.

[10] L. Overbey, M. Todd, Analysis of local state space models for feature extraction in structural health monitoring, Structural Health Monitoring 6 (2007) 145–172.

[11] S. Billings, Z. Lang, A bound for the magnitude characteristics of nonlinear frequency response functions, part 1.: Analysis and computation, International Journal of Control 65 (1996) 309–328.

[12] Z. Lang, S. Billings, Evaluation of nonlinear frequency response function of nonlinear systems under multiple inputs, IEEE Transactions on Circuits and Systems 47 (2000) 28–38.

[13] Z. Lang, S. Billings, Energy transfer properties of nonlinear systems in the frequency domain, International Journal of Control 78 (2005) 345–362.

[14] Z. Peng, Z. Lang, S. Billings, Crack detection using nonlinear frequency response function, Journal of Sound and Vibration 301 (2007) 777–788.

[15] D. Koller, N. Friedman, Probabilistic Graphical Models: Principles and Techniques, The MIT Press, 2009.

[16] Y. Sheikh, M. Shah, Bayesian object detection in dynamic scenes, in: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR05), IEEE, 2005.

[17] S. Mahamud, Comparing belief propagation and graph cuts for novelty detection, in: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR06), IEEE, 2006.

[18] G. R. Grimmett, A theorem about random fields, Bulletin of the London Mathematical Society 5 (1973) 81–84. doi:10.1112/blms/5.1.81.

[19] A. Weller, Methods for inference in graphical models, Ph.D. thesis, Columbia University (2014).

[20] D. Kamenetsky, Ising graphical model, Ph.D. thesis, The Australian National University (2010).

[21] J. K. Johnson, P. Netrapalli, M. Chertkov, Learning planar ising models, arXiv preprint, arXiv:1011.3494v1doi:10.1115/1.2718241.

[22] C. Desir, S. Bernard, C. Petitjean, L. Heutte, One class random forests, Pattern Recognition 46 (2013) 3490–3506.

[23] M. Nilsson, H. Gustafsson, S. V. Andersen, W. B. Kleijn, Gaussian mixture model based mutual information estimation between frequency bands in speech, Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (2002) 525–528.

[24] H. Bethe, Statistical theory of superlattices, Proc. R. Soc. Lond. A 150 (1935) 552575.

[25] J. Yedidia, W. T. Freeman, Y. Weiss, Understanding belief propagation and its generalizations, in: International Joint conference on Artificial Intelligence (IJCAI 2001), 2001.

[26] M. N. Helfrick, C. Niezrecki, P. Avitabile, T. Schmidt, 3d digital image correlation methods for full-field vibration measurement, Mechanical systems and signal processing 25 (3) (2011) 917–927.

[27] D. J. Fleet, A. D. Jepson, Computation of component image velocity from local phase information, Int. J. Comput. Vision 5 (1) (1990) 77–104. doi:10.1007/BF00056772.
URL http://dx.doi.org/10.1007/BF00056772

[28] T. Gautama, M. Van Hulle, A phase-based approach to the estimation of the optical flow field using spatial filtering, Neural Networks, IEEE Transactions on 13 (5) (2002) 1127 – 1136.

[29] E. P. Simoncelli, W. T. Freeman, The steerable pyramid: A flexible architecture for multi-scale derivative computation, in: icip, IEEE, 1995, p. 3444.

[30] R. J. Hyndman, Computing and graphing highest density regions, The American Statistician, Taylor and Francis 50 (1996) 120–126.