Appeared in: *Structural Safety*, 2011, **33**(6): 317–332.

Bayesian network modeling of correlated random variables drawn from a Gaussian random field

Michelle Bensi[1*], Armen Der Kiureghian[1], Daniel Straub[2]
[1] University of California, Berkeley, CA 94720
[2] Technical University of Munich, Germany
[*] Corresponding author, E-mail: mtbensi@gmail.com

1. Introduction

In civil engineering applications, it is often necessary to model vectors of random variables drawn from a random field. For example, in investigating the seismic risk of a lifeline, the earthquake-induced ground motion intensities at the locations of the system components constitute a vector of random variables drawn from the ground motion random field. Similarly, factors determining the progress of deterioration in elements of concrete surfaces are random variables drawn from environmental and material property random fields. Proper modeling of the dependence structure of vectors of random variables is essential for accurate probabilistic analysis. In the special case when the field is Gaussian, or derived from a Gaussian field, the spatial dependence structure of the field is completely defined by the autocorrelation function and the correlation matrix fully defines the dependence structure of the random vector drawn from the field. Typically, this correlation matrix is fully populated. Although this paper only deals with Gaussian random fields, the methods developed are equally applicable to non-Gaussian fields that are derived from Gaussian fields, e.g., [1].

In some applications, including the aforementioned examples, it is of interest to update a probabilistic model in light of available or assumed observations of the random field. For example, in the case of a lifeline subjected to an earthquake, one might be interested in updating the reliability of the system when ground motion intensities at one or more locations are observed, or when evidence is available on the performance of individual components based on the output from structural health monitoring sensors or observations made by inspectors [2]. In the case of a concrete surface subject to deterioration, the reliability of the system can be updated, e.g., when cracking (or no cracking) of the concrete in some of the elements is observed. The Bayesian network (BN) methodology is a powerful tool for such updating purposes, particularly when the available information evolves in time and the updating must be done in (near) real time, see, e.g., [3] [4]. However, there is a limiting characteristic of the BN that poses a challenge when modeling random variables drawn from a random field: due to the full correlation structure of the random variables, the BN becomes densely connected. When combining these random variables with system models that involve additional random variables, the computational and memory demands of the resulting BN rapidly grow with the number of points drawn from the random field. In this paper, we develop approximate methods to overcome this difficulty. Specifically, we present methods for reducing the density of the BN model of the random field by selectively eliminating nodes and links. The aim is to minimize the number of links in the BN while limiting the error in the representation of the correlation structure of the random variables drawn from the Gaussian random field.

When the random field as well as the observed random variables are jointly Gaussian, a well-known analytical solution exists for computing the conditional probabilistic model. However, the random field model often is part of a larger problem involving mixtures of continuous and discrete random variables and fields. For example, in seismic risk assessment of a lifeline, a random field may define the ground

motion intensity across a geographic region, while discrete random variables define the performance or damage states of the lifeline and its constituent components. When evidence is entered on non-Gaussian or discrete random variables in such a model, e.g., the observed damage state of a component, the existing analytical solution for updating the distribution of the Gaussian variables is no more applicable. It is in this context that the BN is useful for modeling and updating of Gaussian random fields.

The paper begins with a brief introduction to BNs as a means for probabilistic inference and describes their advantages and limitations. Next, BN models of random variables drawn from a Gaussian random field are described. Approximation methods are then developed to achieve computationally tractable BN models. Using several generic and systematic spatial configuration models, numerical investigations are performed to compare the relative effectiveness of the proposed approximation methods. Finally, the effects of the random field approximation on estimated reliabilities of example spatially distributed systems are investigated. The paper ends with a set of recommendations for BN modeling of random variables drawn from a random field. More details on development of BN models for random fields and application to infrastructure seismic risk assessment can be found in [2] .

2. **Brief Introduction to Bayesian Network**

Bayesian networks are probabilistic graphical models consisting of nodes representing random variables and directed links describing probabilistic dependencies. Throughout the paper, the terms "node" and "random variable" are used interchangeably. Consider the simple two-node BN in Figure 1a, which models random variables X and Y. The directed link indicates that Y is probabilistically dependent on X. Node Y is a *child* of node X, while node X is a *parent* of node Y. Attached to node Y is the conditional probability distribution of Y given X. Because node X has no parent, a marginal probability distribution is attached to it. If the two variables are discrete, then probability mass functions (PMFs) define their distributions. In particular, a conditional PMF defines the probability that Y is in each of its mutually exclusive states, given each mutually exclusive state of X, i.e., the probabilities $\Pr(Y = y_i|X = x_j)$, where y_i indicates the i^{th} state of Y and x_j indicates the j^{th} state of X. In the BN terminology, this conditional PMF is called a *conditional probability table* (CPT). If node Y has multiple parents, as in Figure 1b, the size of the CPT for node Y becomes large because the PMF of Y must be defined for all combinations of the states of the parent nodes $X_1, \dots, X_n$. If each of the nodes in Figure 1b has m states, then the CPT attached to node Y has m^{n+1} entries. It is seen that the size of the CPT attached to a node grows exponentially as the number of parents increases. These CPTs generally must be stored in memory. Therefore, as n increases, computational bottlenecks are encountered due to physical memory constraints.

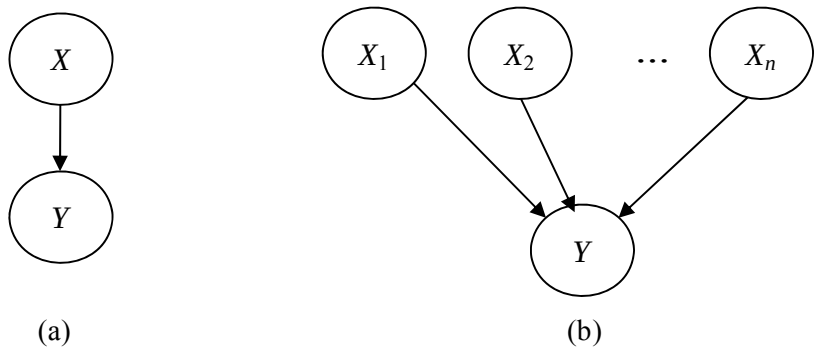


Figure 1: (a) Two-node BN with Y dependent on X, (b) BN with Y dependent on a vector of random variables $\mathrm{X} = \{X_1, \dots, X_n\}$

BNs are most useful in answering probabilistic queries, e.g., determining the posterior distributions of the random variables in the BN, when one or more variables are observed. The process of updating the BN given available evidence is known as probabilistic inference. Updating occurs consistent with the d-separation properties of the BN, which characterize the way in which information flows through different types of connections (see [5,6] for more details). Conceptually, inference may be thought of as efficient application of Bayes Rule and the Total Probability Theorem on a large scale.

Many applications in civil engineering require mixtures of continuous and discrete nodes. For example, as described earlier, in seismic applications the distribution of the ground motion intensity at a site is modeled by a continuous random variable, while the damage state of a component at that location may be represented by a discrete random variable. However, existing exact inference algorithms for BNs and software applications utilizing these algorithms [7,8] impose severe limitations on the use of continuous random variables. Specifically, they only permit linear functions of Gaussian random variables that have no discrete children, for which case analytical solutions exist for computing the conditional probability distribution [9]. The necessity of modeling discrete children of continuous nodes, e.g., the damage state of a component that is causally dependent on the ground motion intensity at its site, prohibits the use of the aforementioned algorithms. Furthermore, inference algorithms for BNs with mixtures of discrete and continuous random variables are computationally far more demanding than those for BNs with only discrete nodes [10,11]. Hence, when using BNs, there is often significant advantage in discretizing all continuous random variables. The examples utilized in this paper are solved using exact inference algorithms that require all nodes to be discretized.

One of the most common exact inference algorithms is Junction Tree [5,12]. When performing exact probabilistic inference using this algorithm, graphical constructs known as *cliques* are formed, which contain subsets of nodes in the BN. Each clique is assigned a *clique table,* created by taking the product of the CPTs of all nodes in the clique. Common general-purpose BN software, such as Hugin [7],[13], require that clique tables formed during probabilistic inference be stored in memory. For BNs with discrete nodes, the memory demand of the Junction Tree algorithm is exponential in the size of the largest clique generated when performing inference. Therefore, it is preferable to work with cliques, and consequently clique tables, that are as small as possible. The sizes of cliques generated when performing inference calculations are related to the sizes of the CPTs associated with nodes in the graph as well as the density of dependency (links) between the nodes. Thus, discrete-node BNs with broadly dependent random variables (i.e. densely connected nodes) and/or nodes with many states (i.e. large CPTs) result in large cliques and, consequently, exponential increases in computational demands. The reader is referred to [5] for more details on clique sizes and the Junction Tree algorithm. Although our interest is in the application of exact inference algorithms, it is noted that small CPTs are also preferable when working with approximate sampling-based algorithms, e.g. [14,15]. Therefore, by reducing CPT sizes, the methods described in this paper are also useful when applying approximate inference algorithms.

A BN representing a vector of random variables drawn from a random field, which has a fully populated correlation matrix, necessarily has densely connected nodes. The exponential increase in computational demand with the number of nodes in a clique can quickly render the BN computationally intractable as the number of random variables increases. Computational efficiency may be achieved by two means: (1) reducing the density of links in the BN; and/or (2) reducing the number of discretized states associated with each node in the BN. Reducing the number of links in the BN introduces inaccuracy in modeling variable interdependencies. Reducing the number of states associated with a node increases discretization error. Therefore, a trade-off exists between computational demand and model accuracy. The focus of this paper is on developing methods for reducing the number of links in the BN rather than on issues related to discretization. Guidance on discretization of continuous random variables is provided in [3,4][4,3].

3. **Bayesian Network Modeling of Random Fields**

Let $Y(\mathbf{x})$, $\mathbf{x} \in \Omega$, be a multi-dimensional Gaussian random field defined within the domain Ω with mean function $\mu_Y(\mathbf{x})$, standard deviation function $\sigma_Y(\mathbf{x})$, and auto-correlation coefficient function $\rho_{YY}(\mathbf{x}_1, \mathbf{x}_2)$, $(\mathbf{x}_1,\mathbf{x}_2) \in \Omega$. Without loss of generality, hereafter we work with the normalized random field

$$Z(\mathbf{x}) = \frac{Y(\mathbf{x}) - \mu_Y(\mathbf{x})}{\sigma_Y(\mathbf{x})} \tag{1}$$

$Z(\mathbf{x})$ is a Gaussian random field with zero mean, unit variance and, because of the linearity of the transformation, auto-correlation coefficient function $\rho_{ZZ}(\mathbf{x}_1, \mathbf{x}_2) = \rho_{YY}(\mathbf{x}_1, \mathbf{x}_2)$.

Consider random variables $Z_i = Z(\mathbf{x}_i)$, $i = 1, \dots, n$, drawn from $Z(\mathbf{x})$ at selected points $\mathbf{x}_i$ within Ω. The Gaussian vector $\mathbf{Z} = [Z_1, Z_2, \dots, Z_n]^{\mathrm{T}}$ has zero means, unit standard deviations and correlation matrix $\mathbf{R} = [\rho_{ij}]$, where $\rho_{ij} = \rho_{YY}(\mathbf{x}_i, \mathbf{x}_j)$, $i, j = 1, \dots, n$. We consider the general case where the correlation matrix is fully populated. Figure 2 shows a BN model of vector $\mathbf{Z}$. The correlation structure implies links between all pairs of Z-nodes. The particular formulation in Figure 2 requires specification of the conditional distribution of each Z_i given its parent nodes $Z_1, \dots, Z_{i-1}$. That is, the conditional probability $\Pr(Z_i = z_i | Z_1 = z_1, \dots, Z_{i-1} = z_{i-1})$ must be specified for each combination of the mutually exclusive states of $Z_1, \dots, Z_{i-1}, Z_i$. It should be clear that the CPT of node Z_n can become extremely large as n increases.

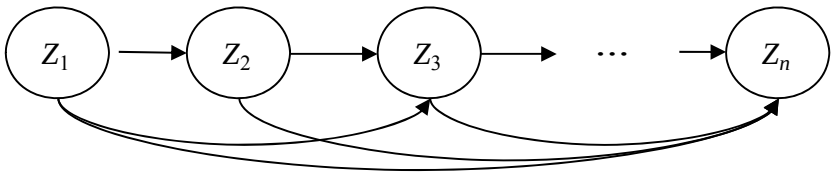


Figure 2: BN model of vector Z drawn from Gaussian random field Z(x)

Vector $\mathbf{Z}$ may be decomposed as a product of an $n \times n$ transformation matrix $\mathbf{T}$ and an $n \times 1$ vector of statistically independent, standard normal random variables $\mathbf{U}$:

$$\mathbf{Z} = \mathbf{TU} = \begin{bmatrix} t_{11} & \cdots & t_{1n} \\ \vdots & \ddots & \vdots \\ t_{n1} & \cdots & t_{nn} \end{bmatrix} \begin{Bmatrix} U_1 \\ \vdots \\ U_n \end{Bmatrix} \tag{2}$$

The transformation matrix $\mathbf{T}$ may be determined using an Eigenvalue (Karhunen–Loève) expansion, Cholesky factorization, or other decomposition methods that diagonalize the covariance matrix [16]. Alternatively, the transformation matrix may be determined approximately via numerical optimization [17,18], as described later in this paper. The BN corresponding to the above transformation is shown in Figure 3, where the latent U-nodes are introduced as parents of the Z-nodes. Here, an element of the transformation matrix, t_{ij}, is interpreted as a factor on the link between U_j and Z_i. A value of $t_{ij} = 0$ corresponds to no link between U_j and Z_i. Due to the unit covariance matrix of $\mathbf{U}$, we have $\mathbf{R} = \mathbf{TT}^{\mathrm{T}}$. The CPTs required by the BN in Figure 3 are easier to specify than those required for the BN in Figure 2, because each Z_i is a deterministic function of its parent U-nodes.

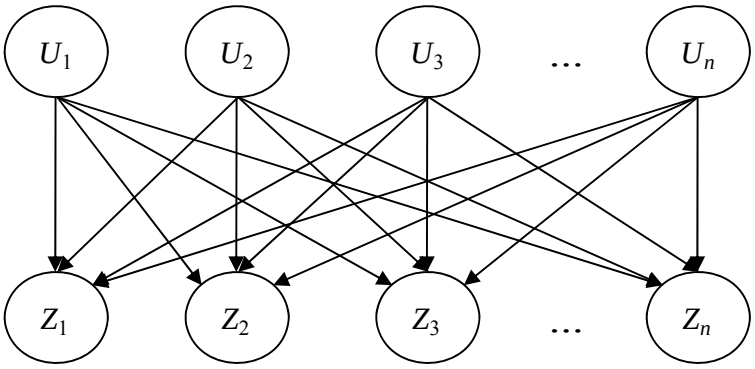


Figure 3: BN model of the decomposition of vector Z drawn from Gaussian random field $Z(\mathrm{x})$

The BNs in Figure 2 and Figure 3 are densely connected and, therefore, exact inference with these BNs becomes computationally intractable as the number n of random variables increases. To achieve computational tractability, the number of links in the BN must be reduced. However, removal of links introduces error. Hence, in developing a procedure for link elimination, the goal is to balance computational efficiency and model accuracy by removing as many links as possible without causing significant loss of accuracy. A procedure must be defined to identify and eliminate links that are least critical for accurately modeling the vector $\mathbf{Z}$.

Define $\widehat{\mathbf{T}} = [\hat{t}_{ij}]$ as an approximate transformation matrix with some of its elements set to zero. Setting $\hat{t}_{ij} = 0$ implies removal of the link connecting U_j and Z_i. If column j of $\widehat{\mathbf{T}}$ has all zero entries, then node U_j has no children and can be eliminated from the BN. The removal of links and nodes in the BN results in an approximation of the covariance matrix of $\mathbf{Z}$, including the on-diagonal variance terms. The errors in the variances are corrected by introducing an additional $n\times1$ vector of statistically independent standard normal random variables $\mathbf{V}$ and a diagonal transformation matrix $\mathbf{S}$,

$$\hat{\mathbf{Z}} = \mathbf{SV} + \widehat{\mathbf{T}}\mathbf{U} = \begin{bmatrix} s_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & s_n \end{bmatrix} \begin{Bmatrix} V_1 \\ \vdots \\ V_n \end{Bmatrix} + \begin{bmatrix} \hat{t}_{11} & \cdots & \hat{t}_{1m} \\ \vdots & \ddots & \vdots \\ \hat{t}_{n1} & \cdots & \hat{t}_{nm} \end{bmatrix} \begin{Bmatrix} U_1 \\ \vdots \\ U_m \end{Bmatrix} \tag{3}$$

where $\hat{\mathbf{Z}}$ denotes the approximated vector. Note that, after elimination of barren U-nodes, $\widehat{\mathbf{T}}$ is an $n\times m$ matrix and $\mathbf{U}$ is and $m \times 1$ vector, where $m \leq n$. To achieve unit variances for $\hat{\mathbf{Z}}$, we set the condition

$$s_i = \sqrt{1 - \sum_{k=1}^{m} \hat{t}_{ik}^2}\,, \quad i = 1, \ldots, n \tag{4}$$

This correction does not affect the off-diagonal terms of the covariance matrix. Furthermore, the approximated correlation coefficients are given by

$$\hat{\rho}_{ij} = \sum_{k=1}^{m} \hat{t}_{ik}\, \hat{t}_{jk}, \quad i = 1, \ldots, n \tag{5}$$

The BN corresponding to the formulation in (3) is shown in Figure 4. We note that the addition of the V-nodes does not significantly increase the computational complexity of the BN.

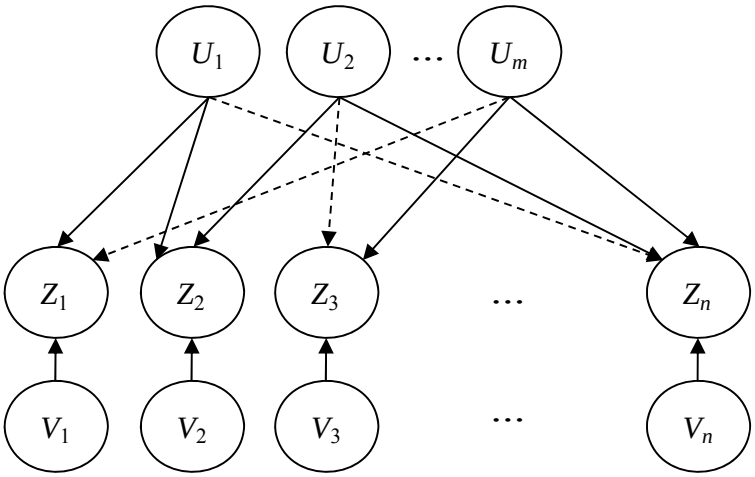


Figure 4: BN model of the approximate decomposition $\hat{\mathrm{Z}} = \mathrm{SV} + \hat{\mathrm{T}}\mathrm{U}$

As suggested by Song and Kang [17] and Song and Ok [18], the transformation in (3) may be regarded as a generalization of the Dunnett-Sobel (DS) class of Gaussian random variables [19]. This class of random variables is defined by

$$Z_i = s_i * V_i + t_i * U, \;\; i = 1 \ldots n \tag{6}$$

where V_i and U are independent standard normal random variables and s_i and t_i are variable-specific coefficients satisfying the conditions $s_i = \left(1 - t_i^2\right)^{1/2}$ and $-1 \leq t_i \leq 1$. Note that U is common to all Z_i and, therefore, is the source of the correlation among them. For this model, one can easily show that Z_i are standard normal random variables having the correlation coefficients $\rho_{ij} = t_i * t_j$ for $i \neq j$ and $\rho_{ij} = 1$ for $i = j$. As a special case, Z_i are equi-correlated when all t_i are identical. The transformation in (6) corresponds to that in (3) with $\hat{\mathbf{T}}$ being an $n \times 1$ vector. The associated BN is similar to that in Figure 4 with a single U-node being a common parent to all the Z nodes as shown in Figure 5.

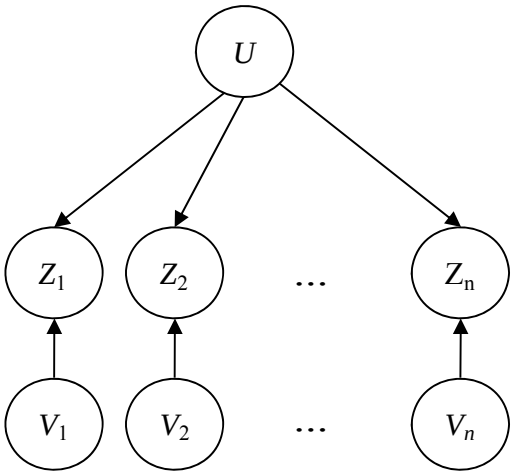


Figure 5: BN model of DS class of random variables

4. Construction of Approximate Transformation Matrix $\widehat{\mathbf{T}}$

As described above, setting elements of the transformation matrix $\widehat{\mathbf{T}}$ to zero corresponds to removing links in the BN. Link removal may be accomplished by three means: (1) selectively removing U-nodes and all associated links by setting entire columns of the transformation matrix to zero; (2) selectively removing links by setting the corresponding elements of the transformation matrix to zero; or (3) combination of the first two, i.e., first reducing the number of U-nodes and then selectively removing links from the remaining U-nodes.We explore all three methods in this paper.

To select nodes and links for elimination, we introduce node and link importance measures. For a transformation matrix $\mathbf{T}$ (or $\widehat{\mathbf{T}}$), the node importance measure (NIM) for node U_i is defined as

$$M_i = \sum_{k=1}^{n} \sum_{l=k}^{n} |\Delta_i\,(k,l)| \tag{7}$$

where $\Delta_i(k,l)$ is the $(k,l)^{\text{th}}$ element of the matrix

$$\boldsymbol{\Delta}_i = \mathbf{t}_i \mathbf{t}_i^{\mathrm{T}} \tag{8}$$

in which $\mathbf{t}_i$ is the i^{th} column of $\mathbf{T}$ or $\widehat{\mathbf{T}}$. $\boldsymbol{\Delta}_i$ quantifies the contribution of U_i to the correlation matrix of $\mathbf{Z}$. Thus, M_i is a measure of the information contained in the correlation matrix $\mathbf{R}$ that is lost by removing U_i. Clearly, eliminating the U-node associated with the smallest NIM will result in the least loss of accuracy. Nodes may be eliminated based on their NIM values until a pre-selected number of nodes remain, or until a pre-defined threshold on NIM is exceeded.

Similarly, a link importance measure (LIM) associated with element t_{ij} of the transformation matrix $\mathbf{T}$ or $\widehat{\mathbf{T}}$ is defined as

$$m_{ij} = \sum_{k=1}^{n} \sum_{l=k}^{n} |\delta_{ij}\,(k,l)| \tag{9}$$

where $\delta_{ij}(k,l)$ is the $(k,l)^{th}$ element of the matrix

$$\boldsymbol{\delta}_{ij} = \mathbf{t}_i \mathbf{t}_i^{\mathrm{T}} - \mathbf{t}_{i,j} \mathbf{t}_{i,j}^{\mathrm{T}} \tag{10}$$

in which $\mathbf{t}_i$ is the i^{th} column of $\mathbf{T}$ or $\widehat{\mathbf{T}}$ and $\mathbf{t}_{i,j}$ is equal to $\mathbf{t}_i$ but with its j^{th} element set equal to zero. Thus, m_{ij} is a measure of the information contained in the correlation matrix that is lost by eliminating the link from node U_j to node Z_i. It follows that eliminating the link associated with the smallest LIM will result in the least loss of accuracy. Links are eliminated based on their LIM until a pre-selected number remain, or until a pre-defined threshold on the LIM is exceeded. The node- and link-based approaches can be combined: U-nodes are first eliminated based on their NIMs, followed by selective elimination of links associated with the remaining nodes based on their LIMs. Below, we describe three methods for constructing the transformation matrix $\mathbf{T}$ and its approximation $\widehat{\mathbf{T}}$ by use of the above measures.

4.1. Decomposition Using Eigenvalue Expansion

Using an eigenvalue expansion, the transformation matrix can be obtained in the form [20]

$$\mathbf{T} = \boldsymbol{\Phi} \boldsymbol{\Lambda}^{\mathbf{1/2}} \tag{11}$$

where $\boldsymbol{\Phi} = [\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_n]$ is the matrix of eigenvectors and $\boldsymbol{\Lambda} = \mathrm{diag}[\lambda_i]$ is the diagonal matrix of eigenvalues obtained by solving the eigenvalue problem

$$\mathbf{R}\boldsymbol{\phi}_i = \boldsymbol{\phi}_i \lambda_i, \qquad i = 1, \ldots, n \tag{12}$$

In this case the transformation matrix is generally full and the resulting BN takes the form in Figure 3. The eigenvalues λ_i may be interpreted as factors associated with the U-nodes, while the elements in the eigenvectors are interpreted as factors on the links. It is well known that the order of contribution of the eigenvectors to the covariance matrix is consistent with the order of magnitudes of the corresponding eigenvalues. (This property is exploited in Principal Component Analysis [20].) Our numerical investiga-

tions reveal that the order of the NIMs generally agrees with the eigenvalue ordering. In the remainder of this paper, the approach in which the transformation is defined via eigenvalue decomposition and U-nodes are eliminated based on their NIMs is referred to as the *node-based eigenvalue approach* (NEA). When the elimination is performed on individual links based on their LIMs, the approach is referred to as a *link-based eigenvalue approach* (LEA). It is noted that the use of a link-based elimination procedure in conjunction with eigenvalue decomposition was first suggested in Straub et al. [21]. However, the importance measure defined there is less general than the one introduced above, as the former is only applicable to transformation matrices obtained via eigenvalue decomposition.

4.2. Decomposition Using Cholesky factorization

The decomposition approach using Cholesky factorization results in a transformation matrix **T** that is triangular [16]. Assuming **T** is a lower triangular matrix, the corresponding BN takes the form shown in Figure 6, where node Z_i has the parents $U_1, \dots, U_i$. By virtue of the triangular form of the transformation matrix, the BN in Figure 6 is less densely connected than the BN is Figure 3. However, the sizes of the largest CPT and the largest clique, $\{U_1, \dots, U_n, Z_n\}$, remain unchanged. Consequently, the order of computational complexity of exact inference associated with the formulations in Figure 3 and Figure 6 are the same. Using a node-elimination approach, the nodes corresponding to the rightmost columns of the transformation matrix are typically associated with the smallest NIMs and are eliminated first. Hereafter, the elimination of nodes based on a transformation matrix determined by Cholesky decomposition is referred to as the *node-based Cholesky approach* (NCA). When individual links are eliminated by zeroing the elements in the Cholesky decomposition matrix based on their LIMs, the approach is referred to as the *link-based Cholesky approach* (LCA).

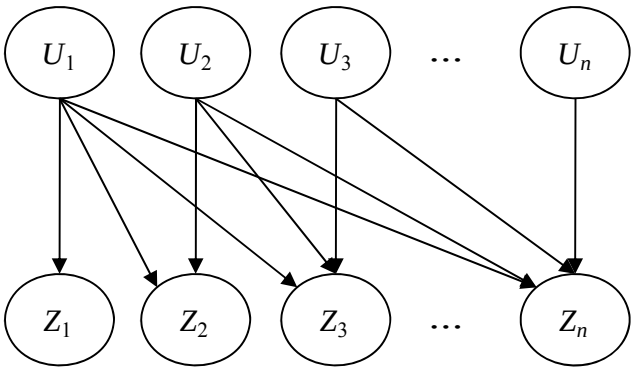


Figure 6: BN corresponding a transformation matrix obtained by Cholesky decomposition

4.3. Node and Link Elimination Using Optimization

Numerical optimization offers an alternative approach to defining the approximate transformation matrix $\hat{\mathbf{T}}$. We propose an optimization-based node-elimination approach consisting of two steps: (1) specification of the number $m \leq n$ of U-nodes to include in the BN; and (2) solution of the nonlinear constrained optimization problem

$$\begin{aligned} &\min \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left[\rho_{ij} - \sum_{k=1}^{m} \hat{t}_{ik} * \hat{t}_{jk} \right]^2 \\ &\text{subject to: } \sum_{k=1}^{m} \hat{t}_{ik}^2 \leq 1 , \qquad i = 1, \dots, n \end{aligned} \tag{13}$$

The objective of the above problem is to minimize the sum of squared deviations between the actual and approximated coefficients in the upper triangle of the symmetric correlation matrix (excluding the diagonal terms, which are later corrected using (3) and (4)). The constraint functions are based on (4). The use of nonlinear constrained optimization to determine the terms of an approximate transformation matrix was earlier suggested by Song and Ok [18] and Song and Kang [17]. Hereafter, construction of the approximate transformation matrix by use of the optimization scheme in (13) is referred to as the node-based optimization approach (NOA).

Analogously, a link-based optimization problem may be formulated, in which the number of links rather than nodes is specified. The result is a mixed-integer, non-linear constrained optimization problem, which is difficult to solve in practice. Problems with convergence and over-sensitivity to initial values were encountered. To avoid solving mixed-integer optimization problems, an alternative might be the use of strategies employed in the field of topology optimization, e.g. [22]. There, elements are allowed to have intermediate values between void (=0) and solid (=1), but these intermediate values are penalized in the optimization. By considering links in the BN as such elements, the binary state of the links (present or not) could be avoided in the optimization. Such an approach is not pursued here but may prove to be effective. Instead, an iterative procedure for link elimination is developed. First, the optimization problem in (13) is solved for a specified m. Then, the term in the resulting transformation matrix with the lowest LIM is set to zero and the remaining terms in the matrix are re-optimized. The link elimination is repeated until a pre-set number of links remain. The iterative algorithm is summarized in Figure 7. Hereafter, this procedure is referred to as the iterative link-based optimization approach (ILOA).

INITIALIZE
Specify:
$m =$ maximum number of U-nodes to include
$L =$ number of links to eliminate
ELIMINATE NODES
1. Determine $n \times m$ matrix $\widehat{\mathbf{T}}_0$ with elements obtained by solving the optimization problem in (13)
ELIMINATE LINKS
for $p = 1, \dots, L$
1. Calculate LIMs for all elements of $\widehat{\mathbf{T}}_{p-1}$
2. Identify p elements in $\widehat{\mathbf{T}}_{p-1}$ with the smallest LIM values. (These include elements previously set equal to zero.)
3. Determine the $n \times m$ matrix $\widehat{\mathbf{T}}_p$ by setting all elements identified in step 2 equal to zero and solving for the remaining elements according to (13)

end

Figure 7: Iterative algorithm for determining $\widehat{\mathbf{T}}$ through node and link elimination

An alternative to the above iterative scheme is as follows: In each iteration step, all elements in the transformation matrix that have LIM values below a threshold are set to zero. The remaining elements in the matrix are then re-optimized according to (13). The procedure is repeated until no link has an LIM below the specified threshold. Hereafter, this procedure is referred to as the alternative iterative link-based optimization approach (AILOA).

For the numerical analysis reported in this paper, the function $fmincon$ in Matlab, which finds the minimum of a constrained non-linear multi-variable function, is used to solve the optimization problem (13).

4.4. Qualitative Comparison of Methods for Constructing $\widehat{\mathbf{T}}$

It is important to recognize a significant distinction between the decomposition (eigen-expansion and Cholesky factorization) and optimization methods described above. In the former methods, after setting selected elements to zero, the remaining elements in the transformation matrix remain unchanged. In con-

trast, in the optimization methods, the matrix $\hat{\mathbf{T}}$ is re-optimized after elimination of each set of nodes or links. In this sense, the optimization approaches are "dynamic," while the decomposition methods are "static." Because of these characteristics, one can expect that the optimization approaches will generally outperform the decomposition methods.

From a practical standpoint, a node-elimination approach is preferable because the effect on the computational complexity of the BN is systematic. The user can easily predict the memory demands of the resulting BN prior to performing inference. The same cannot be said when links are individually eliminated. However, the link-elimination approach is often better able to capture the correlation structure of a given configuration of points drawn from a random field because unimportant links are eliminated selectively rather than in large groups, as is the case with the node-elimination approach. These observations are borne out by numerical investigations that follow.

5. Numerical Investigation of Approximation Methods

In this section, a numerical investigation is performed to determine the relative efficiencies of the proposed methods for approximating the correlation structure of random variables drawn from a Gaussian random field in the context of BN analysis. Four generic configurations for the locations of points drawn from the random field are considered:

(1) Points arranged along a line (Figure 8a)
(2) Points arranged concentrically on a circle (Figure 8b)
(3) Points arranged in a rectangular grid (Figure 8c)
(4) Points arranged in clusters (Figure 8d)

Different sizes of the above configurations are considered by changing the number of points, n, and the distance between the points. The latter is controlled through a distance measure, d, as defined in Figure 8. For the cluster configuration, the points in each cluster are distributed uniformly on the circumference of a circle of radius d, which is centered at a pre-defined coordinate. For the sake of brevity, only results for representative cases of the above configurations are shown and general trends are described.

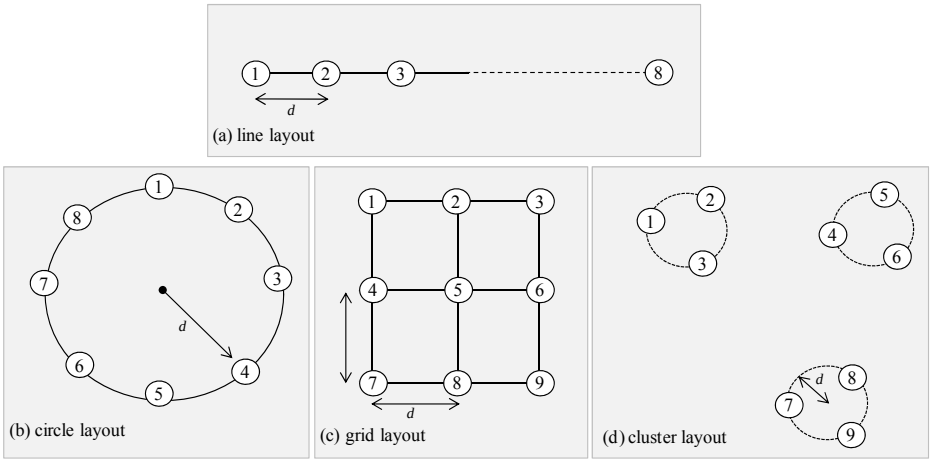


Figure 8: Generic configurations considered in numerical investigation

To measure the relative efficiencies of the aforementioned approximation methods (NEA, LEA, NCA, LCA, NOA, ILOA, and AILOA), the error measure

$$e = \frac{\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}|\hat{\rho}_{ij} - \rho_{ij}|}{\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}|\rho_{ij}|} \tag{14}$$

is used, where ρ_{ij} is the correlation coefficient between Z_i and Z_j and $\hat{\rho}_{ij}$ is the approximated correlation coefficient as defined in (5). Note that the sums are over the upper triangle of the correlation matrix, excluding the diagonal terms. This error measure was selected because the normalization by the sum of absolute values of the correlation coefficients permits comparison of cases with varying sizes of vector $\mathbf{Z}$ and a wide range of correlation values. (Later in the paper, we consider another error measure that is directly relevant to reliability assessment of an infrastructure system.) For each layout, this error is compared against a measure of the computational complexity of the resulting BN. For simplicity and intuitive appeal, we use the number of links in the BN as a proxy for computational complexity. There is not a perfect one-to-one relation between the number of links in a BN and direct measures of computational complexity, such as the tree width, which is equal to one less than the size of the largest clique [23]. However, the number of links is sufficiently indicative to serve our purpose of comparing alternative approximation methods.

For the numerical investigation, the Gaussian random field is assumed to be homogenous and isotropic with the autocorrelation coefficient function

$$\rho_{VV}(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\Delta \mathbf{x}_{ij}}{a}\right) \tag{15}$$

where $\Delta \mathbf{x}_{ij} = |\mathbf{x}_i - \mathbf{x}_j|$ is the distance between sites $\mathbf{x}_i$ and $\mathbf{x}_j$ and a is a measure of the correlation length of the random field with value $a = 6$ in units of distance. This correlation model is taken from [24] and is motivated by our interest in applying the BN methodology to seismic risk assessment of infrastructure systems.

5.1 Results

First consider a line layout system (Figure 8a) with 10 sites. For this system, an exact BN with the eigenvalue expansion method has $10^2 = 100$ links (see Figure 3) and with the Cholesky decomposition method it has $1 + 2 + \cdots + 10 = 55$ links (see Figure 6). Figure 9 shows the error measure versus the number of links included in the BN for $d = 1, 5$ and 10 (corresponding to $d = \frac{1}{6}a, \frac{5}{6}a$ and $\frac{5}{3}a$) using the decomposition and optimization methods described above. All methods exhibit similar trends with increasing d, i.e., the error measure becomes larger and converges to zero at a slower rate. This is partly due to the definition of the error measure: For large d, the correlations are small and, therefore, the error measure is magnified. For all d values and for both node- and link-elimination approaches, the optimization methods consistently achieve smaller errors than the decomposition methods. Within the optimization methods, ILOA and AILOA outperform NOA, particularly with regard to the speed of convergence to zero. For small d, the performances of NEA and LEA are close to those of the corresponding optimization approaches, but the performance of the eigen-expansion approaches rapidly degrade as d increases. This is because, for large d, the correlation coefficients are small and the eigenvalues of the resulting correlation matrix are of similar magnitude. Furthermore, LEA exhibits non-monotonic convergence. Conversely, the approaches based on Cholesky decomposition perform poorly for small distances, but their performance approaches those of the optimization methods as d increases. The above investigation was repeated for 5- and 15-site line layouts and similar trends were observed.

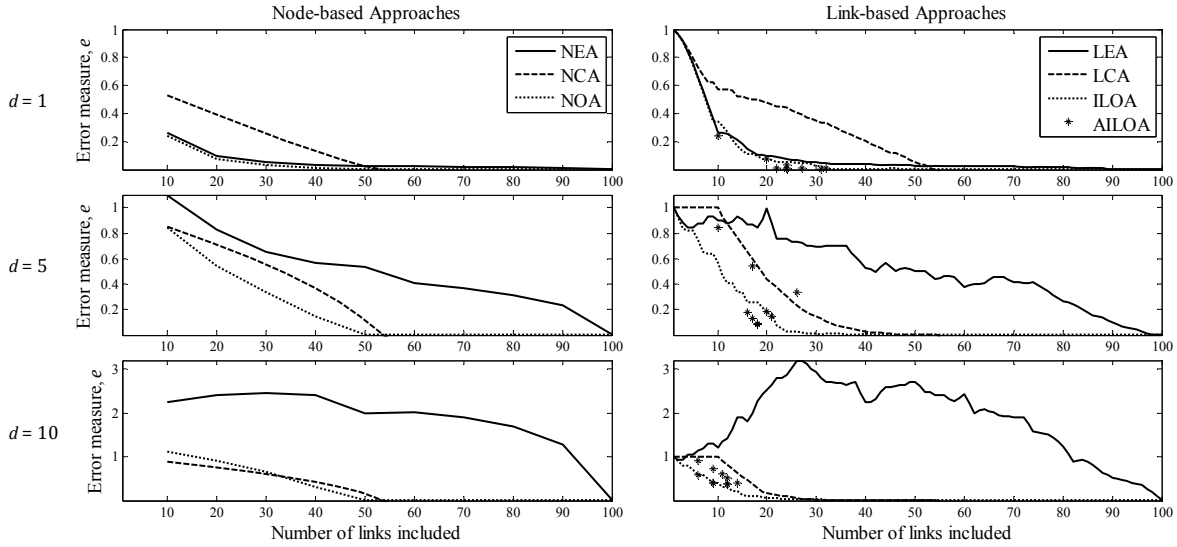


Figure 9: Error in approximating correlation matrix for 10-site line layouts

Figure 10 and Figure 11 respectively show the results for 10-site circle and 9-site grid layouts. As with the line layout, the optimization methods outperform the decomposition methods for all values of d. For small d, all the points in these layouts are closely spaced and, hence, the correlation coefficients are large and relatively uniform in magnitude. As a result, the correlation structure is well captured by a few U-nodes, or even a single U-node. As d increases, the elements of the correlation matrix become less uniform. A larger number of nodes is then required to capture the correlation structure when using node-based approaches. In this situation, the link-based approaches exhibit better performance because they offer versatility in distributing links among different U-nodes.

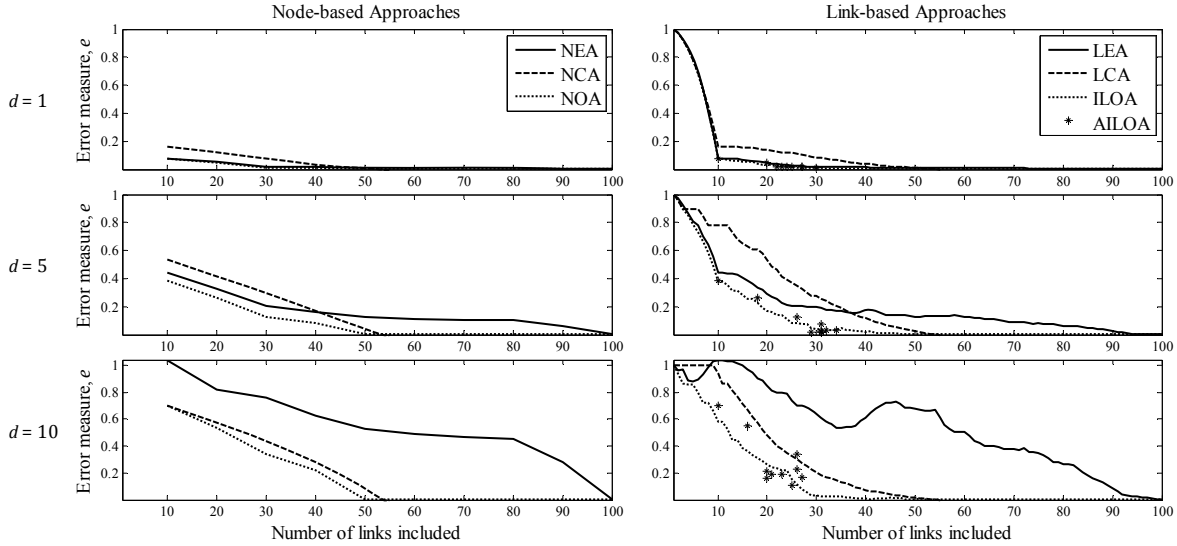


Figure 10: Error in approximating correlation matrix for 10-site circle layouts

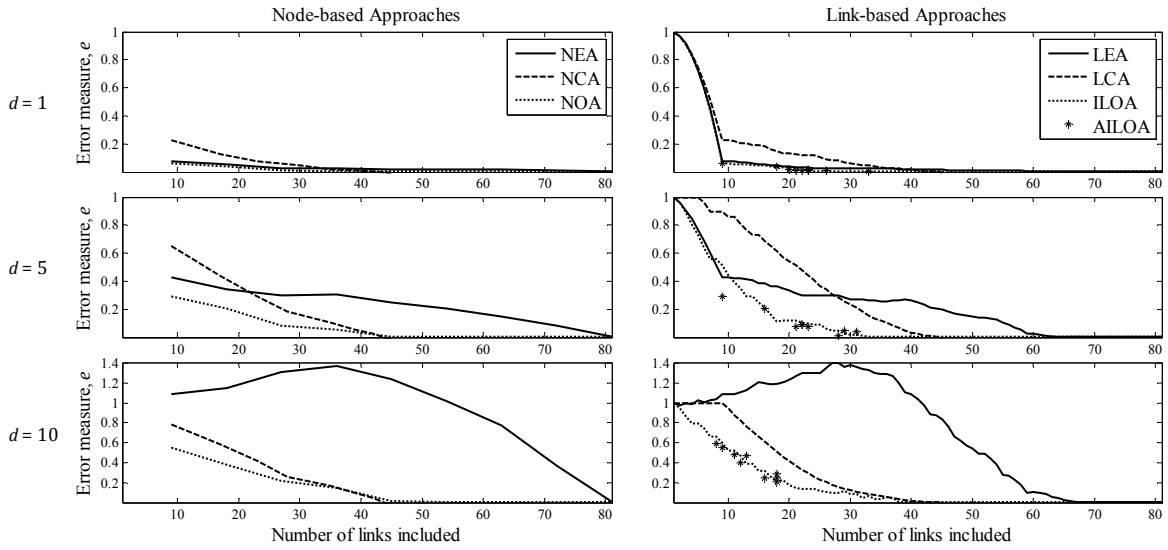


Figure 11: Error in approximating correlation matrix for 9-site grid layouts

Finally, the layout in which points are arranged in clusters is considered. Figure 12 shows examples of 9-site layouts arranged into three three-site clusters centered at coordinates (0,0), (50,50) and (25,75). For each cluster, points are equally distributed around a circle of radius d. Thus, d is a measure of how tightly the nodes in each cluster are arranged.

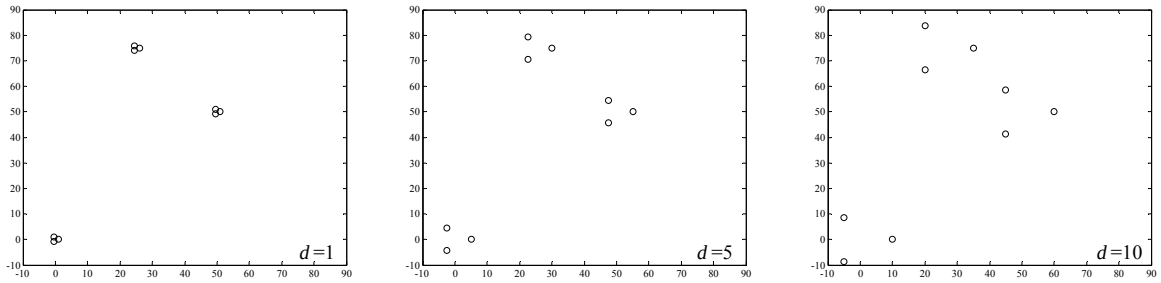


Figure 12: Example cluster layouts with d = 1, 5, 10

Figure 13 shows the errors for the 9-site cluster layouts with $d = 1, 5$ and 10. Among the node-based approaches, NCA and NOA perform similarly while NEA performs poorly, particularly for large d. Overall, the link-based approaches, excluding LEA, offer much better performance for the cluster layouts. This is because link-based approaches are able to distribute links among different U-nodes consistent with the geometry of the clusters. To illustrate this concept, consider the two BNs shown in Figure 14, which represent alternative approximations of the 9-node cluster system with 9 links. The top BN uses the ILOA, which distributes the 9-links such that the points in each cluster are connected to a common U-node and points in different clusters are uncorrelated. The bottom BN uses a node-based approach, in which 9 links connect a single U-node to all the Z-nodes. While both BNs in Figure 14 have the same number of links, ILOA distributes them more efficiently, consistent with the geometry of the layout. If clusters are sufficiently far apart, it is not necessary to include information paths (i.e. common U-nodes) between sites located in different clusters. This is why the link-based methods outperform node-based methods.

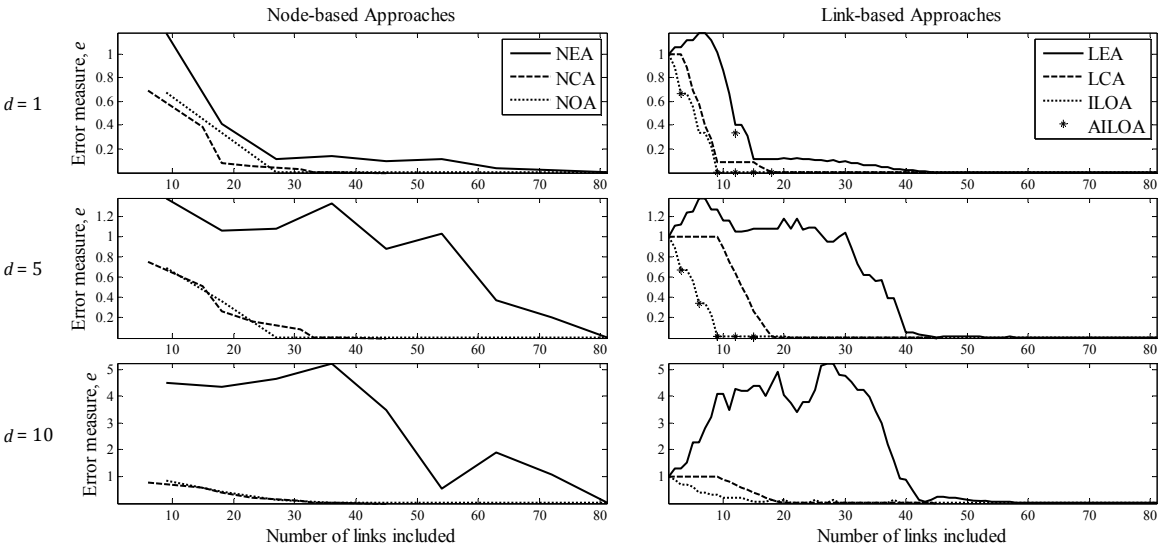


Figure 13:Error in approximating correlation matrix for 9-site cluster layouts

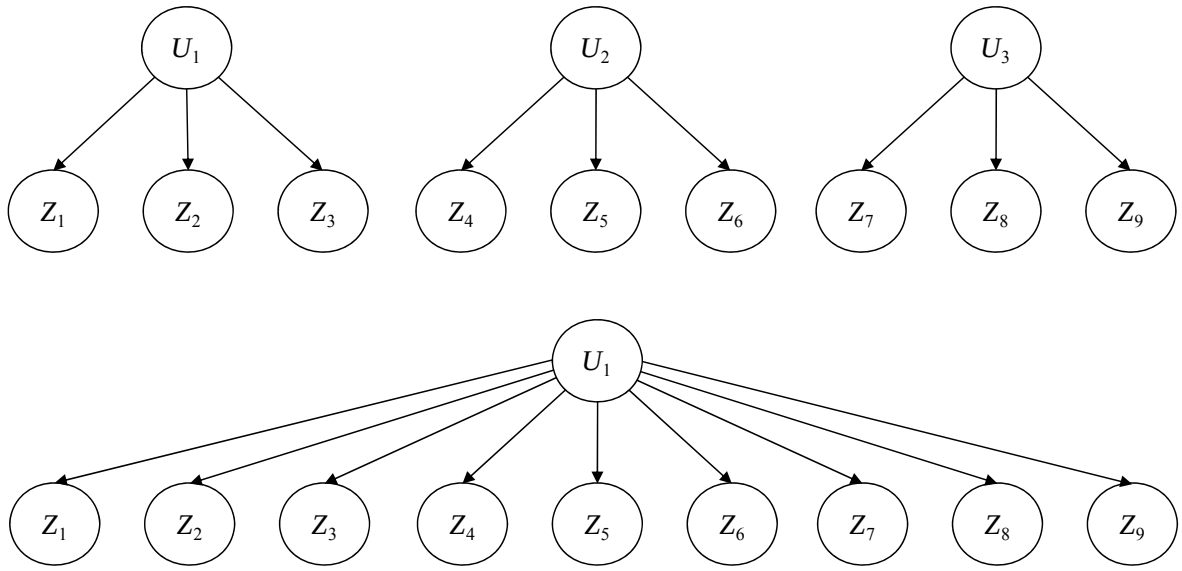


Figure 14: 9-link BN approximations for 9-site cluster layouts using ILOA (top) and NOA (bottom)

For comparison, select BNs obtained using the ILOA for the $n = 10$-site line and circle and $n = 9$-site grid layouts for $d = 5$ are shown next. For the line and circle layouts, each BN contains $2n = 20$ links; the BN corresponding to the grid layout contains $2n = 18$ links. Note that the ILOA is an iterative procedure for identifying important links. The iterative procedure is used in lieu of considering the full link-based optimization problem, which is difficult to solve in practice. Therefore, the BNs obtained from the procedure may not be globally optimal. Geometric interpretations of the resulting BNs for the line, circle, and grid layouts are not as clear as that for the cluster configuration, however trends do exist.

For the line layout the BN using 20 links is shown in Figure 15a. In this BN, links are distributed so that there are over-lapping information paths between nodes in close proximity. Sites located farther apart do not share a common U-node. Figure 15b graphically illustrates the sites that are linked by common U-nodes. Shaded circles in this diagram indicate sites that share a common U-node. The BN for the 10-site circle layout with 20 links is shown in Figure 16a and the corresponding diagram of sites sharing common U-nodes is shown in Figure 16b. Because points are spaced concentrically in this layout, the sites are linked by a single common U-node when as few as 10 links are considered (see the second diagram from the left in Figure 16b). When more than 10 links are included, overlapping information paths around the circle are added, similar to the trend seen for the line-layout. For the grid layout, the BN containing the 18 most important links defined by ILOA is shown in Figure 17a and the diagram of sites sharing common U-nodes is shown in Figure 17b. For this layout with 18 links, the links are distributed so that adjacent sites share at least one U-node.

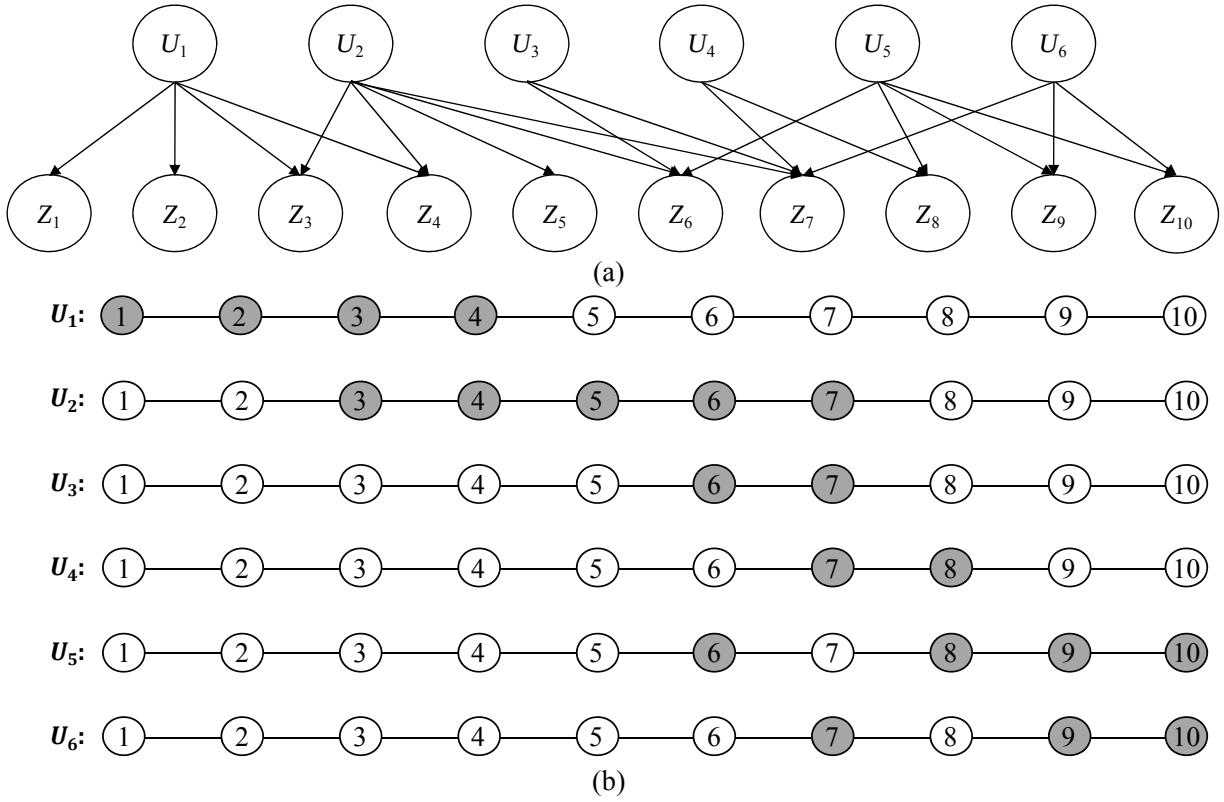


Figure 15: (a) BN approximation using ILOA with 20 links; (b) illustration of sites linked by common *U*-node for 10-site line layout with *d*=5

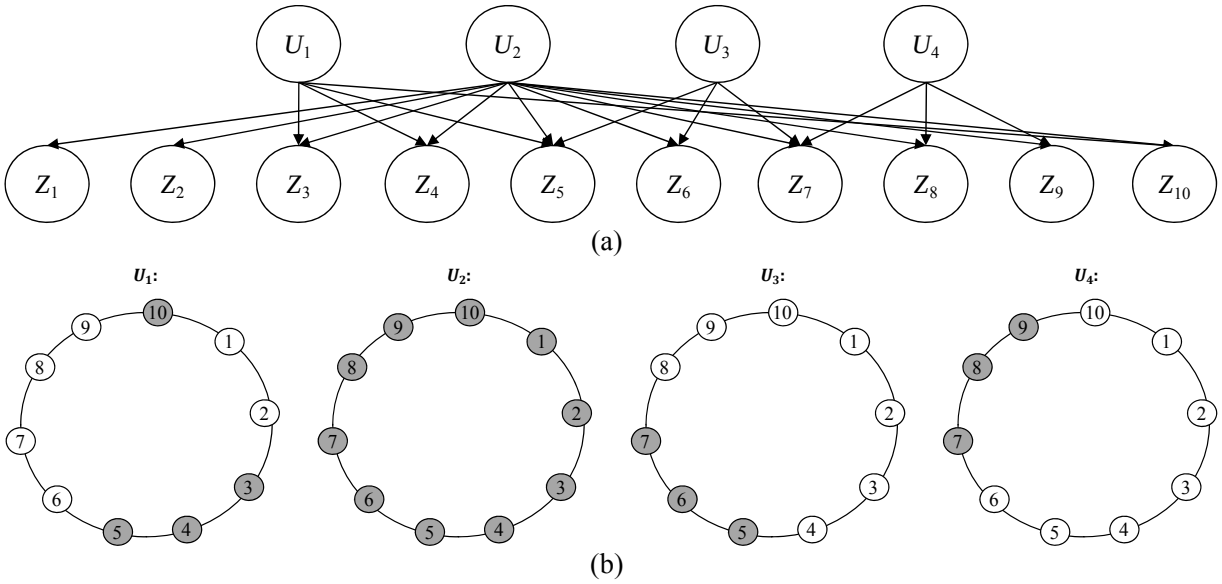


Figure 16: (a) BN approximation using ILOA with 20 links; (b) illustration of sites linked by common *U*-node for 10-site circle layout with *d*=5

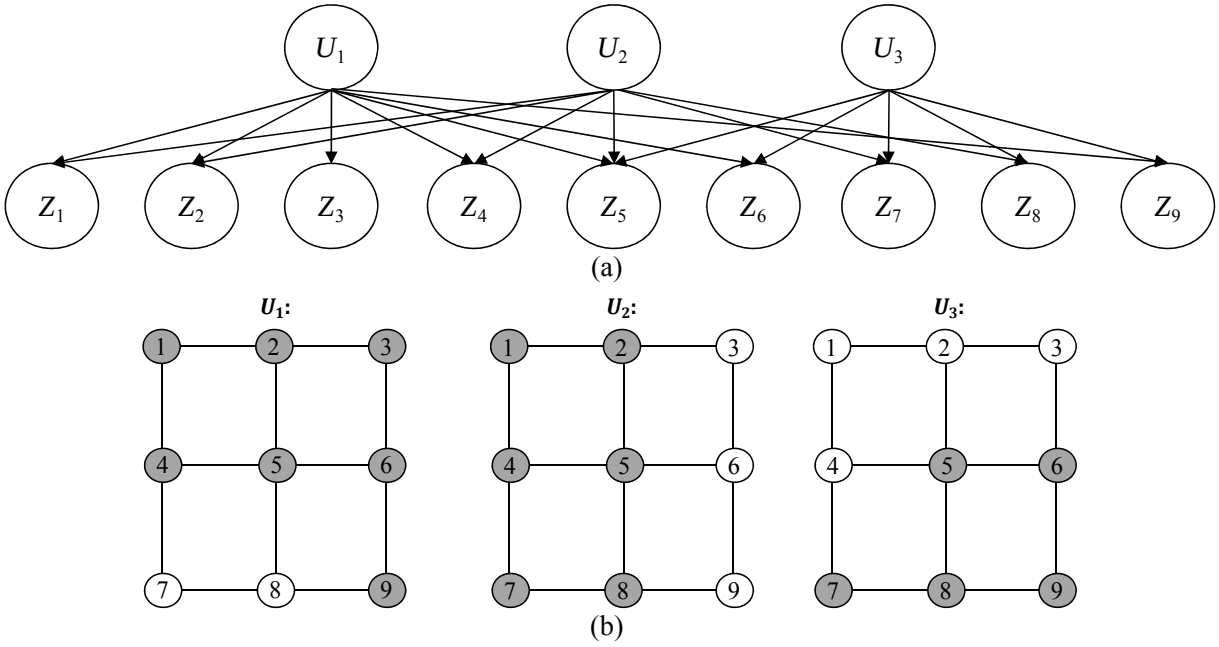


Figure 17: (a) BN approximation using ILOA with 18 links; (b) illustration of sites linked by common *U*-node for 9-site grid layout with *d*=5

6. Effect of Correlation Approximations on System Reliability

The objective in BN analysis is usually some sort of probabilistic assessment involving risk or reliability evaluation, life cycle cost analysis, statistical inference, or expected-utility decision-making. Thus, the effect of approximating the correlation structure of random variables drawn from a Gaussian random field should be evaluated within such a context. Towards that end, we investigate the effect of the correlation approximation on the estimates of reliability of example systems. The points in the layouts in Figure 8 are assumed to represent the locations of components of infrastructure systems subjected to an earthquake demand. Two system performance criteria are considered: (1) all components must survive for the system to survive (series system); (2) only one component needs to survive to ensure survival of the system (parallel system).

A limit-state function is assigned to each component of the system to model its performance. For component i, $i = 1, \dots, n$, the limit-state function has the form

$$g_i = \ln(R_i) - \ln(S_i) \tag{16}$$

where R_i is the capacity of the component and S_i is the demand on the component. The natural logarithms $\ln(R_i)$ are assumed to be statistically independent normal random variables with common means λ_R and common standard deviations ζ_R. The natural logarithm of the demand on component i is expressed as

$$\ln(S_i) = \ln(\bar{S}_i) + \epsilon_C + \epsilon_{S,i} \tag{17}$$

where $\bar{S}_i$ is the median demand, ϵ_C is a normally distributed error term with zero mean and variance σ_C^2 that is common to all components in the system, and $\epsilon_{S,i}$ is a site-specific error term drawn from a homogeneous Gaussian random field with zero mean, variance σ_S^2 and autocorrelation function $\rho_{\epsilon\epsilon}(|\mathbf{x}_i - \mathbf{x}_j|)$ distributed over the spatial domain within which the system is located. It is assumed that $\bar{S}_i$ is the same for all sites in the system. This assumption, as well as the earlier assumption of identical distributions for component capacities, ensures that all components have equal importance with regard to the performance of the system. This allows us to focus solely on the effect of the approximation of the correlation structure. Readers familiar with seismic applications will recognize (17) as the general formulation of a ground-motion prediction equation (also known as an attenuation relation), see, e.g. [25].

For a system of n components, there exist n limit state functions. Component i is in the fail state if $g_i \leq 0$. Thus, for a series system, the failure probability is given by

$$\Pr(F_{series}) = \Pr[(g_1 \leq 0) \cup (g_2 \leq 0) \cup \dots \cup (g_n \leq 0)] = \Phi_n(-\mathbf{M}_g, \mathbf{\Sigma}_{gg}) \tag{18}$$

while for a parallel system the failure probability is given by

$$\Pr(F_{parallel}) = \Pr[(g_1 \leq 0) \cap (g_2 \leq 0) \cap \dots \cap (g_n \leq 0)] = 1 - \Phi_n(\mathbf{M}_g, \mathbf{\Sigma}_{gg}) \tag{19}$$

The right-hand sides of the above equations are exact solutions of the system failure probabilities expressed in terms of the n-variate multinormal cumulative probability function $\Phi_n(\mathbf{M}_g, \mathbf{\Sigma}_{gg})$ with the mean vector $\mathbf{M}_g$ having common elements $\lambda_R - \ln \bar{S}$ and the covariance matrix $\mathbf{\Sigma}_{gg}$ having variances $\zeta_R^2 + \sigma_C^2 + \sigma_S^2$ and covariances $\sigma_C^2 + \rho_{ij}\sigma_S^2$, where $\rho_{ij} = \rho_{\epsilon\epsilon}(|\mathbf{x}_i - \mathbf{x}_j|)$ is the autocorrelation coefficient function of the random field from which $\epsilon_{S,i}$ are drawn. For the latter function, the form in (15) is used. Furthermore, $\lambda_R = -0.9, ln\bar{S} = -1.8, \sigma_m = 0.2$ and $\sigma_r = 0.5$ are used, while for ζ_R two values as described below are selected. The multinormal probabilities are computed by an algorithm available in the general-purpose reliability code CalREL [26]. Figure 18 shows the BN model of the system performance. Each node g_i is binary, indicating failure or survival of component i. It is a child of nodes representing the component capacity (R_i) and demand (S_i). Node S_i is a child of nodes representing the common median demand $\bar{S}$, the common error term (ϵ_C) and the site-specific error term ($\epsilon_{S,i}$). The correlations among the site-specific error terms are accounted for through the latent U- and V-nodes according to the formulation in (3), as exemplified in Figure 4. The performance of the system is represented by a single node Sys, which is a child of all the component limit-state nodes g_i. Note that the inclusion of the discrete state

nodes g_i and Sys results in a problem that is no longer jointly Gaussian. It is noted that the converging structure of links going into node Sys, as shown in Figure 18, is computationally inefficient. We have developed other methods for efficiently constructing the system performance portion of the BN (e.g. see [2]), the description of which is beyond the scope of this paper.

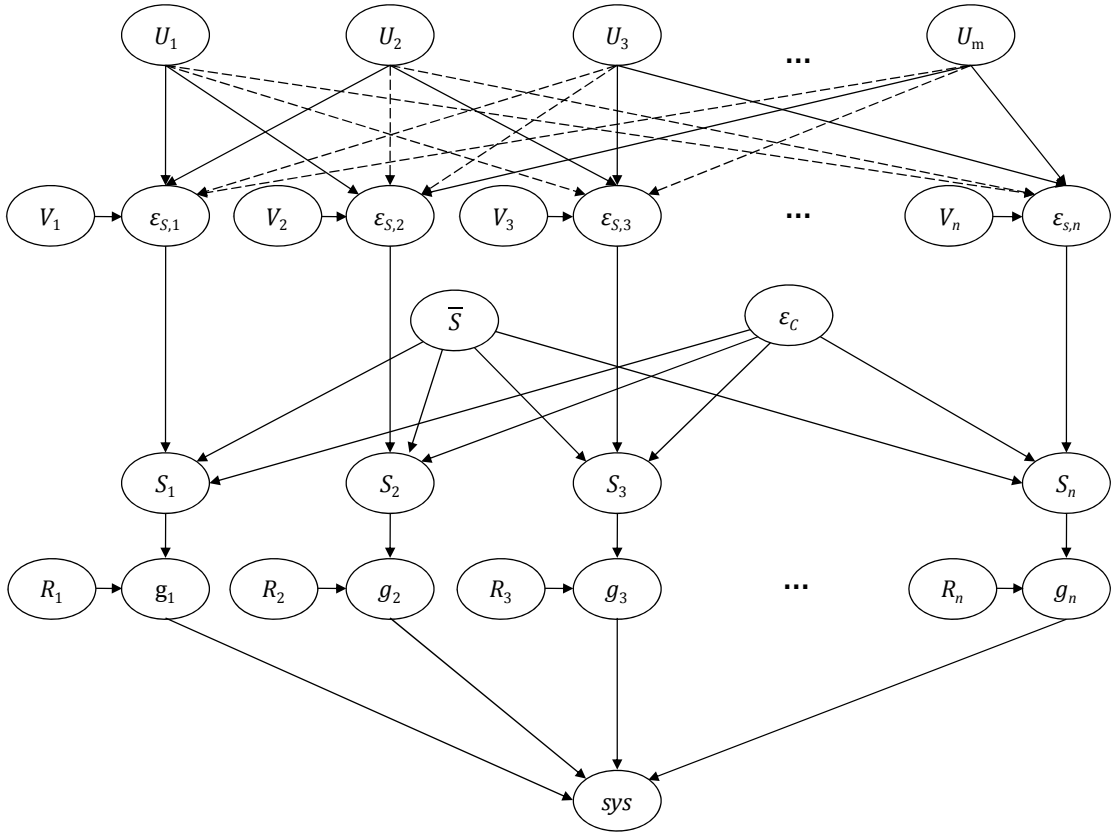


Figure 18: BN model of system performance

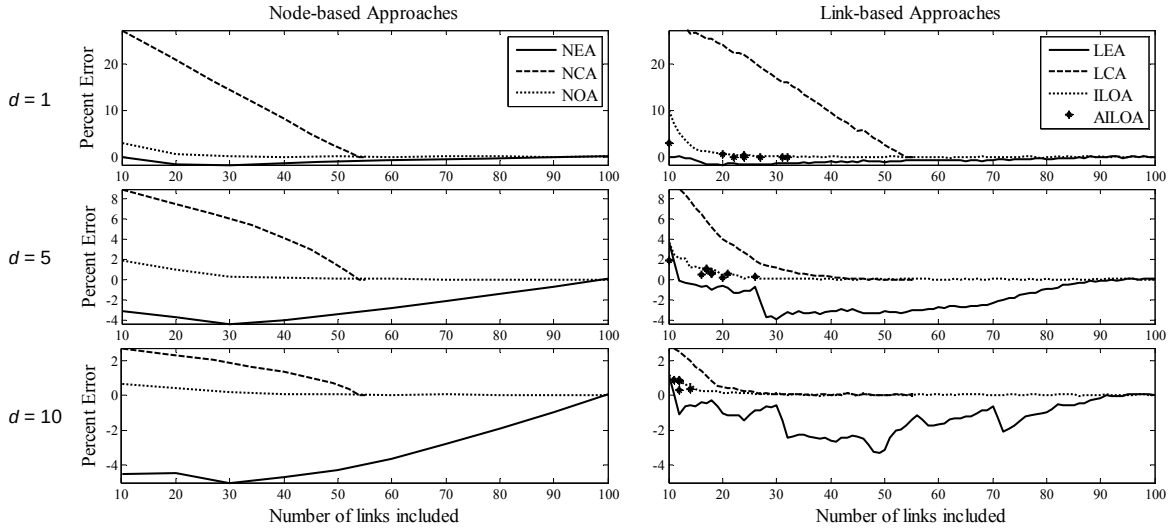
When using the proposed approximation methods, the correlation coefficients ρ_{ij} are replaced by their approximations $\hat{\rho}_{ij}$. In order to avoid mixing the discretization error of the BN with the error due to elimination of nodes and links, failure probabilities are computed by the same formulas as in (18) and (19) but using the approximate covariance matrix. In the following analysis, the error in computing the system failure probability is measured as

$$\text{Percent Error} = \frac{\hat{P}_{f,sys} - P_{f,sys}}{P_{f,sys}} * 100 \tag{20}$$

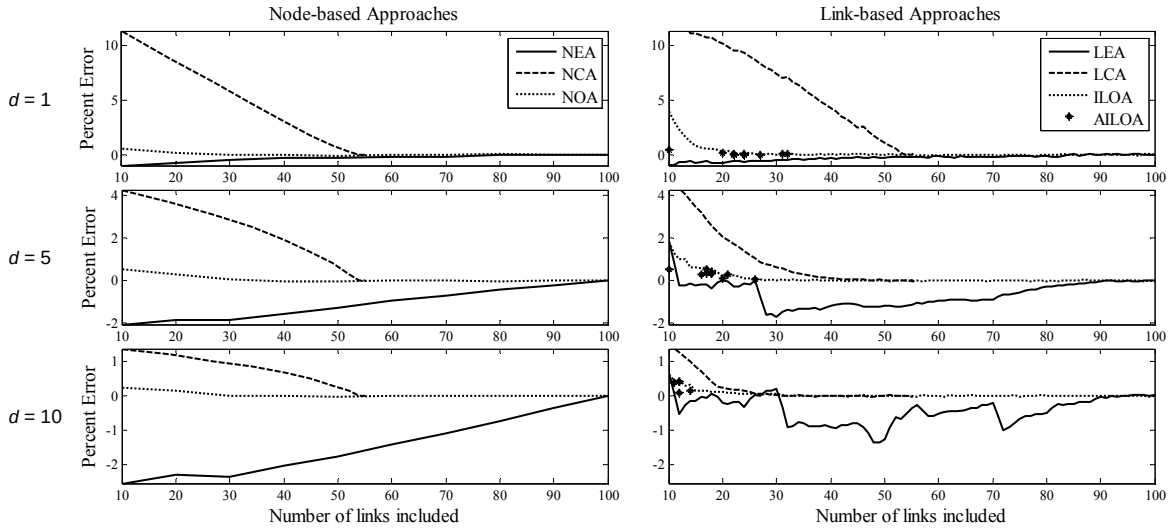
where $P_{f,sys}$ is the true system failure probability and $\hat{P}_{f,sys}$ is the approximate failure probability computed based on the approximate correlation matrix. A negative (positive) error implies underestimation (overestimation) of the system failure probability.

For a 10-site line layout system, Figure 19 and Figure 20 plot the percent errors in estimating the failure probabilities for series and parallel systems, respectively, versus the number of links included for each of the approximation methods. These are shown for two values of ζ_R, which approximately represents the coefficient of variation of the component capacities. Consistent with the results shown in the previous section, the optimization approaches outperform the decomposition approaches for both the parallel and

series configurations. Among the decomposition approaches, the eigen-expansion methods are unconservative for the series system and the Cholesky factorization methods are unconservative for the parallel system. Overall, the errors associated with the series system are significantly smaller than those associated with the parallel system. This is because the series system failure probability is less sensitive to the correlation between component demands than that of a parallel system with the same components -- a fact that has also been observed by other investigators [27]. Comparing the graphs in parts (a) and (b) of Figure 19 and Figure 20, it is evident that the error due to the approximation in the correlation matrix of the component demands becomes less critical when the uncertainty in the component capacities is large ($\zeta_R = 0.6$ versus $\zeta_R = 0.3$). Thus, accurate modeling of the correlation structure of the random field is less critical when important sources of uncertainty other than the random field are present.

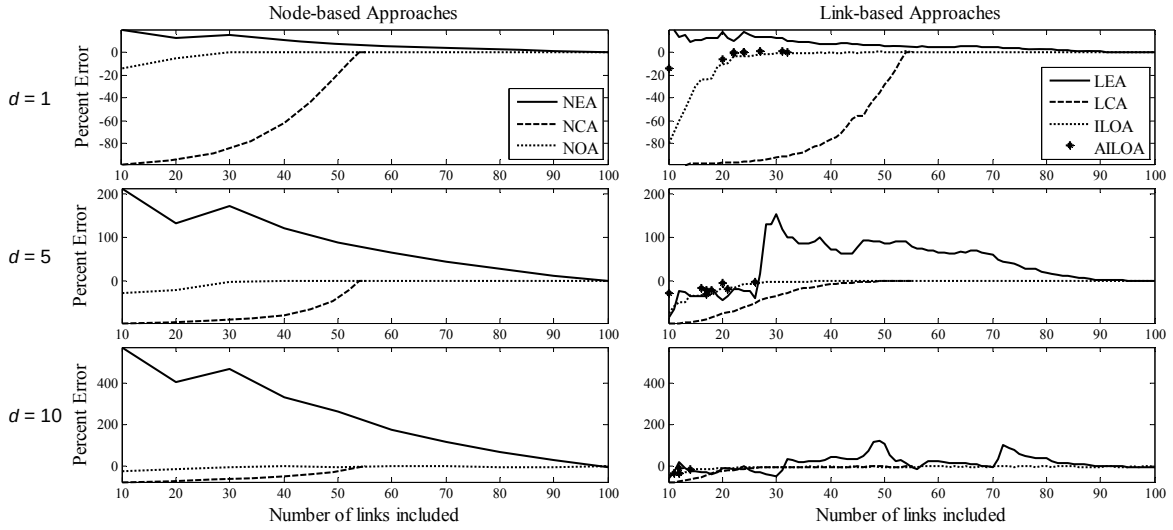


(a)

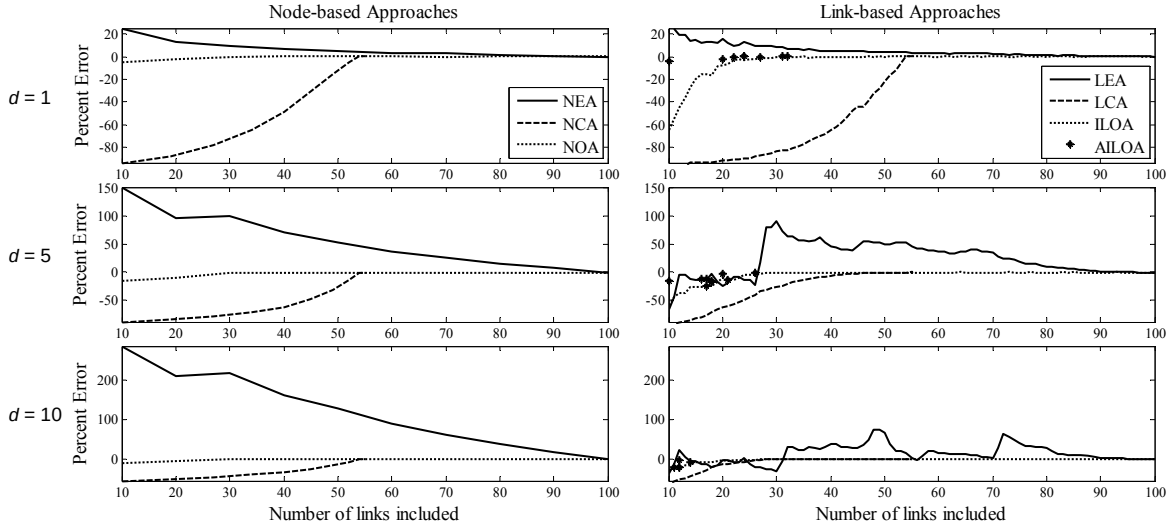


(b)

Figure 19: Percent error in estimating failure probability for a 10-site line series system when (a) $\zeta_R = 0.3$ and (b) $\zeta_R = 0.6$



(a)



(b)

Figure 20: Percent error in estimating failure probability for a 10-site line parallel system when (a) $\zeta_R = 0.3$ and (b) $\zeta_R = 0.6$

To aggregate and compare the results of the numerical investigation for the various layouts, we consider the minimum number of links required in the BN model of each layout to achieve an error less than 10% in the estimate of the system failure probability. Table 1 and Table 2 present the required minimum number of links for series and parallel systems, respectively, for each of the layouts and each of the seven approximation methods considered. Tables on the left list the results for $\zeta_R = 0.3$, while those on the right list the results for $\zeta_R = 0.6$. Results for $d = 10$ are not shown for the series system because, for this relatively long distance, the target accuracy threshold of 10% is achieved without including random field effects. For shorter correlation lengths, similar accuracy is obtainable when neglecting random field effects for some layouts, particularly for $\zeta = 0.6$, as indicated by a symbol (*) in Table 1.

Table 1: Number of links required to achieve less than 10% error in estimate of failure probability for series system when ζ**=0.3** (left) and $\boldsymbol{\zeta}$**=0.6** (right)

(Bold numbers indicate the minimum number of links obtained with any approximation method)

d = 1								
Layout	N	NEA	LEA	NCA	LCA	NOA	ILOA	AILOA
Line	5	**5**	**5**	9	9	**5**	**5**	**5**
Line	10	10	**9**	40	40	10	**9**	10
Circle	5	**5**	**5**	**5**	**5**	**5**	**5**	**5**
Circle	10	**10**	**10**	27	22	**10**	**10**	**10**
Cluster	9	9	**6**	15	7	18	**6**	9
Grid	9	**9**	**9**	17	18	**9**	**9**	**9**

d = 1								
Layout	N	NEA	LEA	NCA	LCA	NOA	ILOA	AILOA
Line	5	5	**4**	5	**4**	5	**4**	5
Line	10	10	**7**	19	21	10	8	10
Circle	5	5	**4**	5	**4**	5	**4**	5
Circle	10	10	**9**	10	**9**	10	**9**	10
Cluster	9	*	*	*	*	*	*	*
Grid	9	9	**8**	9	9	9	**8**	9

d = 5								
Layout	N	NEA	LEA	NCA	LCA	NOA	ILOA	AILOA
Line	5	5	**2**	5	6	5	**2**	5
Line	10	10	**2**	10	11	10	**2**	10
Circle	5	5	**3**	5	6	5	**3**	5
Circle	10	10	**8**	27	24	10	**8**	10
Cluster	9	9	**1**	6	1	9	**1**	3
Grid	9	9	**6**	17	20	9	**6**	9

d = 5								
Layout	N	NEA	LEA	NCA	LCA	NOA	ILOA	AILOA
Line	5	*	*	*	*	*	*	*
Line	10	*	*	*	*	*	*	*
Circle	5	*	*	*	*	*	*	*
Circle	10	10	5	10	8	10	**4**	10
Cluster	9	*	*	*	*	*	*	*
Grid	9	*	*	*	*	*	*	*

*** Percent error below 10% acheiveable when neglecting random field effects.**

Table 2: Minimum number of links required to achieve less than 10% error in estimate of failure probability for parallel system when ζ_R**=0.3** (left) and ζ_R**=0.6** (right)

(Bold numbers indicate the minimum number of links obtained with any approximation method)

d = 1								
Layout	N	NEA	LEA	NCA	LCA	NOA	ILOA	AILOA
Line	5	10	9	14	14	**5**	**5**	**5**
Line	10	50	**14**	52	53	20	21	20
Circle	5	15	9	12	13	**5**	**5**	**5**
Circle	10	**10**	**10**	45	45	**10**	**10**	**10**
Cluster	9	72	35	33	18	27	**9**	**9**
Grid	9	27	13	39	39	**9**	**9**	**9**

d = 1								
Layout	N	NEA	LEA	NCA	LCA	NOA	ILOA	AILOA
Line	5	10	7	12	13	**5**	**5**	**5**
Line	10	30	22	52	53	**10**	19	**10**
Circle	5	10	8	12	13	**5**	**5**	**5**
Circle	10	**10**	**10**	45	44	**10**	**10**	**10**
Cluster	9	63	33	31	17	27	**9**	**9**
Grid	9	18	13	35	37	**9**	**9**	**9**

d = 5								
Layout	N	NEA	LEA	NCA	LCA	NOA	ILOA	AILOA
Line	5	25	20	14	14	**5**	7	**5**
Line	10	100	85	54	39	30	24	**20**
Circle	5	25	19	14	14	**5**	**5**	**5**
Circle	10	90	71	52	50	20	**13**	18
Cluster	9	81	**5**	33	18	27	9	9
Grid	9	81	56	42	41	**9**	19	**9**

d = 5								
Layout	N	NEA	LEA	NCA	LCA	NOA	ILOA	AILOA
Line	5	25	**4**	14	13	5	7	5
Line	10	90	**12**	54	38	30	20	20
Circle	5	25	17	14	14	**5**	**5**	**5**
Circle	10	80	61	52	49	**10**	**10**	**10**
Cluster	9	81	**5**	31	17	27	9	9
Grid	9	72	50	42	40	**9**	18	**9**

d = 10								
Layout	N	NEA	LEA	NCA	LCA	NOA	ILOA	AILOA
Line	5	25	20	14	10	**5**	7	**5**
Line	10	100	13	54	25	30	19	**12**
Circle	5	25	21	14	13	5	9	**5**
Circle	10	100	81	54	47	**10**	23	**10**
Cluster	9	81	**4**	23	17	27	10	■
Grid	9	81	55	42	37	**9**	32	**9**

d = 10								
Layout	N	NEA	LEA	NCA	LCA	NOA	ILOA	AILOA
Line	5	25	**3**	14	9	5	7	5
Line	10	100	13	52	23	20	15	**12**
Circle	5	25	**3**	12	11	5	8	5
Circle	10	100	77	52	44	10	21	**10**
Cluster	9	54	**4**	23	16	9	8	■
Grid	9	81	51	42	34	**9**	31	**9**

■ indicates method was not able to achieve a percent error below specified threshold

It is known that series systems are not strongly sensitive to neglecting correlation in demands. This observation is reflected in Table 1, which demonstrates that the accuracy threshold is met when considering very few links, if any. Note that, for the cases in which the threshold is met without inclusion of random field effects, an asterisk (*) is included in the table. These cases essentially require zero links. Furthermore, for the series system, the optimization-based methods are generally not associated with significant gains in efficiency relative to the decomposition approaches, particularly LEA. Because the optimization approaches are computationally more expensive than the decomposition approaches, it may not be of value to compute transformation matrices using optimization techniques when working with series systems.

Conversely, it is well established that parallel systems are sensitive to inclusion of correlation in demands. This finding is reflected in Table 2 by the relatively large number of links that are needed in each case to achieve the required threshold of accuracy in the estimate of the system failure probability. With only a few exceptions, the optimization approaches are more efficient than the decomposition approaches.

For small values of d (high correlations), the node- and link-based optimization approaches offer similar performance, except in the case of the cluster layout system. For large d values (low correlations), NOA and AILOA are more efficient than ILOA. In most cases, more links are required to achieve a percent error below the specified threshold when $\zeta_R = 0.3$ than when $\zeta_R = 0.6$, particularly for larger systems.

7. Discussion

The BN representation of a Gaussian random field developed in this paper is aimed at engineering applications where observations become available on a spatially distributed system. Consider again the example of assessing the seismic risk of an infrastructure system. The value of using the BN for such a system stems from the ease with which the BN facilitates information updating, particularly in near-real time applications in the immediate aftermath of an earthquake. In that context, let nodes S_i in the BN of Figure 18 correspond to ground motion intensities at the locations of the components of a geographically distributed infrastructure system. The ground motion intensity at a location is a function of a median value $\bar{S}$ (a function of the earthquake magnitude, source-to-site distance, and other geologic characteristics) and inter- and intra-event error terms as shown in (17). Consider that an observation has been made on S_1 in Figure 18, e.g., through a ground motion sensor located near component 1 in the infrastructure system. This information propagates in the forward direction by updating the probability of failure of component 1. The information also propagates in the backward direction by updating the site-specific term $\varepsilon_{s,1}$ as well as the common terms ε_c and $\bar{S}$. The strength of the observation on the posterior distributions of these quantities is governed by the conditional relationships that is a priori specified between them. For the example here, they are governed by attenuation relationships that define the CPTs. Thus, an observation on S_1 provides uncertain information on the value of $\varepsilon_{s,1}$. The updated distribution of $\varepsilon_{s,1}$ in turn updates the distributions of its parent nodes V_1 and U_i. This information propagates through the BN, updating the distributions of the other $\varepsilon_{s,i}$ terms that share a common U-node with $\varepsilon_{s,1}$. Consequently the other S_i and component performance nodes are also updated, yielding updated posterior distributions that take into account the information that is known about the ground motion at the location of the sensor. Ultimately, the system failure probability is updated in light of the updated component failure probabilities. The observation of S_1 need not be exact. By adding additional nodes to the BN, we can easily account for measurement error.

It is emphasized that observations on any selection of random variables can be incorporated in the BN. Examples of other sources of information include observations of component performance (e.g. from a health monitoring instrument or from an inspection report), system performance, as well as information about the earthquake (e.g. observed magnitude and location). The use of BNs is of particular value when observations have been made on non-Gaussian nodes. For example, an observation of component performance (i.e. on a node g_i) updates its ancestor nodes in such a way that they are no longer Gaussian. In these cases, conventional conditional Gaussian matrix equations, which are useful when all variables are jointly Gaussian, are no longer applicable for information updating. Thus, the power of the BN is in the facility it provides for information updating in problems involving mixtures of different types of dependent random variables. The methods developed in this paper for BN representation of random fields are particularly useful for such applications.

8. Summary and Conclusions

Methods for efficient Bayesian network (BN) modeling of correlated random variables drawn from a Gaussian random field, such as those arising in seismic risk assessment of spatially distributed infrastructure systems, are investigated. The modeling of broadly dependent random variables results in a BN that is densely connected. Because exact inference algorithms in densely connected BNs are demanding of computer memory, approximate methods are necessary to make the BN computationally tractable for

large systems. This paper develops methods for reducing the density of connections in a BN by eliminating nodes and/or links, while minimizing the error in the representation of the correlation structure. Methods based on classical decomposition techniques as well as numerical optimization are developed and compared. It is found that optimization methods are able to achieve the best trade-off of accuracy versus computational efficiency. The effects of the approximation methods on estimates of failure probability for idealized infrastructure systems are also considered. It is found that the optimization-based approaches offer significant increases in efficiency when modeling the performance of parallel systems. For series systems, which are known to be less sensitive to the correlation structure of component demands, classical decomposition approaches may suffice. While the work done in this paper has been performed as part of an effort to develop a BN-based framework for seismic infrastructure risk assessment, it is believed that the findings are useful for more general applications involving correlated Gaussian random variables.

Acknowledgement

This research was supported by the State of California through the Transportation Systems Research Program of the Pacific Earthquake Engineering Research Center (PEER). Additional support was provided by the Taisei Chair in Civil Engineering at the University of California, Berkeley. The first author also gratefully acknowledges support from a National Science Foundation Graduate Research Fellowship.

References

[1] Grigoriu M. Non-Gaussian models for stochastic mechanics. Probabilistic Engineering Mechanics 2000;15(1):15-23.
[2] Bensi M, Der Kiureghian A, Straub D. A Bayesian network methodology for infrastructure seismic risk assessment and decision-support, PEER 2011/02. Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA, 2011.
[3] Straub D, Der Kiureghian A. Bayesian network enhanced with structural reliability methods: methodology. Journal of Engineering Mechanics 2010;136(10):1248-1258.
[4] Straub D, Der Kiureghian A. Bayesian network enhanced with structural reliability methods: application. J. Engrg. Mech. 2010;136(10):1259-1270.
[5] Jensen FV, Nielsen TD. Bayesian Networks and Decision Graphs. New York: Springer-Verlag, 2007.
[6] Pearl J. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. San Francisco, CA: Morgan Kaufmann Publishers Inc., 1988.
[7] Hugin Expert A/S. Hugin Researcher API 7.0 (www.hugin.com). Denmark: Hugin Expert A/S, 2008.
[8] DSL. Genie 2.0 by Decision Systems Laboratory (http://genie.sis.pitt.edu/). 2007.
[9] Madsen AL. Belief update in CLG Bayesian networks with lazy propagation. International Journal of Approximate Reasoning 2008;49(2):503-521.
[10] Shenoy PP. Inference in hybrid Bayesian networks using mixtures of Gaussians. In: Proceedings of the 22nd conference on uncertainty in artificial intelligence. 2006. p. 428–436.
[11] Lauritzen SL. Propagation of probabilities, means and variances in mixed graphical association models. Journal of the American Statistical Association 1992;87:1098-1108.
[12] Lauritzen SL, Spiegelhalter DJ. Local computations with probabilities on graphical structures and their application to expert systems. Journal of the Royal Statistical Society. Series B (Methodological) 1988;50(2):157-224.
[13] Andersen SK, Olesen KG, Jensen FV. Hugin—a shell for building Bayesian belief universes for expert systems. In: Proceedings of the 11th international joint conference on artificial

intelligence, vol. 2. Detroit, Michigan: Morgan Kaufmann Publishers Inc., 1989. p. 1080-1085.
[14] Cheng J, Druzdzel M. AIS-BN: an adaptive importance sampling algorithm for evidential reasoning in large Bayesian networks. Journal of Artificial Intelligence Research 2000;13:155-188.
[15] Yuan C, Druzdzel M. Importance sampling algorithms for Bayesian networks: principles and performance. Mathematical and Computer Modelling 2006;43(9-10):1189-1207.
[16] Trefethen LN, Bau D. Numerical linear algebra. Philadelphia: Society for Industrial and Applied Mathematics, 1997.
[17] Song J, Kang W. System reliability and sensitivity under statistical dependence by matrix-based system reliability method. Structural Safety 2009;31(2):148-156.
[18] Song J, Ok S. Multi-scale system reliability analysis of lifeline networks under earthquake hazards. Earthquake Engineering & Structural Dynamics 2009;39(3):259-279.
[19] Dunnett CW, Sobel M. Approximations to the probability integral and certain percentage points of a multivariate analogue of student's t-distribution. Biometrika 1955;42(1/2):258-260.
[20] Jolliffe IT. Principal Component Analysis. New York: Springer-Verlag, 2002.
[21] Straub D, Bensi M, Der Kiureghian A. Spatial modeling of earthquake hazard and infrastructure system performance through Bayesian network. In: Proceedings of inaugural international conference of the engineering mechanics institute (EM'08). Minneapolis, Minnesota: Engineering Mechanics Institute, 2008.
[22] Guest JK, Prévost JH, Belytschko T. Achieving minimum length scale in topology optimization using nodal design variables and projection functions. Int. J. Numer. Meth. Engng 2004;61:238–254.
[23] Wainwright MJ, Jordan M. Graphical models, exponential families, and variational inference. Hanover, MA: now Publishers Inc, 2008.
[24] Park J, Bazzurro P, Baker JW. Modeling spatial correlation of ground motion intensity measures for regional seismic hazard and portfolio loss estimation. In: Applications of statistics and probability in civil engineering. Tokyo, Japan: Taylor & Francis Group, 2007.
[25] Abrahamson N, Atkinson G, Boore D, Bozorgnia Y, Campbell K, Chiou B, Idriss IM, Silva W, Youngs R. Comparisons of the NGA ground-motion relations. Earthquake Spectra 2008;24(1):45-66.
[26] Der Kiureghian A, Haukaas T, Fujimura K. Structural reliability software at the University of California, Berkeley. Structural Safety 2006;28(1-2):44-67.
[27] Grigoriu M, Turkstra C. Safety of structural systems with correlated resistances. Applied Mathematical Modelling 1979;3(2):130-136.