Fault Detection and Localization in Smart Grid: A Probabilistic Dependence Graph Approach

Miao He and Junshan Zhang
School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ 85287
Email:{Miao.He,Junshan.Zhang}@asu.edu

***Abstract*—Fault localization in the nation's power grid networks is known to be challenging, due to the massive scale and inherent complexity. In this study, we model the phasor angles across the buses as a Gaussian Markov random field (GMRF), where the partial correlation coefficients of GMRF are quantified in terms of the physical parameters of power systems. We then take the GMRF-based approach for fault diagnosis, through change detection and localization in the partial correlation matrix of GMRF. Specifically, we take advantage of the topological hierarchy of power systems, and devise a multi-resolution inference algorithm for fault localization, in a distributed manner. Simulation results are used to demonstrate the effectiveness of the proposed approach.**

I. Introduction

One of the primary concerns on the reliability of power systems has been the issue of large-scale fault events and their impacts on the overall stability of the power grid. However, today's power systems are not equipped with fault diagnosis mechanisms sufficiently against various malicious attacks and natural physical events [1]. Thus, there is an urgent need for quick assessment of fault events so that corrective actions can be taken promptly to avoid cascading events. Notably, fault diagnosis of power systems has recently received a surge of interests [1]–[5].

It is known that fault diagnosis of transmission lines is challenging [1], due to the massive scale, system uncertainty and inevitable measurement errors. It is proposed in [1] that sensor networks could be embedded into the power grid and help to assess the real-time mechanical and electrical conditions of transmission lines. Along a different line, [2]–[6] advocated to take advantage of the rich phasor information, measured by phasor measurement units (PMU), for line outage detection. The basic idea of recent works [2] and [3] is to compute the change of phasor angles across the buses, and compare them with those of candidate fault events to determine the outage on single or double transmission lines. In multiple-event scenarios, the corresponding event space becomes too vast to be tractable, which could make these approaches ineffective. An algorithm based on time series analysis is proposed in [4]. Several patterns are extracted from the normal operation data of transmission lines, and outliers are regarded as abnormal events.

It is clear that deterministic approaches would not work well in some practical scenarios due to many stochastic events in power systems. For example, abnormal trajectories of the phasor angles could be a result of sudden change in the power injections from load or generation, other than the outage of transmission lines. In light of the stochastic nature of power systems, the bus injections and branch flows could be volatile across various time scales, which is be especially true in smart grids that are supposed to integrate a large number of distributed generations.

With this insight, we propose to use probabilistic graphical models [7] for modeling the spatially correlated data from PMUs, and use statistical hypothesis testing for the task of fault diagnosis. The contributions of our study are twofold: 1) based on the DC power flow model [8], we derive a GMRF model for the phasor angles across the buses. Simply put, building on the stochastic characteristics of power flows [9], we treat phasor angles as random variables and examine in depth their dependencies. We also show the relationship between the partial correlations of phasor angles and the physical parameters of power systems. 2) We develop a decentralized algorithm for fault diagnosis by exploiting change detection and localization in the partial correlations of phasor angles.

A. Notation and Basic Definition

In a *Markov random field* (MRF) $\mathbf{X}$ defined on the site set S, the *neighborhood* N_s of site s is a subset of S such that the following Markov property holds [7]:

$$X_s \perp \mathbf{X}_{S\setminus\{s\cup N_s\}} | \mathbf{X}_{N_s} \tag{1}$$

The *dependence graph* of MRF $\mathbf{X}$ is induced by the *minimal neighborhood system* $\mathcal{G}_s$ [10], by placing an edge $\{s,t\} \in E$ between the sites s and t that are neighbors [1]. A *clique* c is a subset of S with cardinality $|c| = 1$ or such that any two sites in c are neighbors. Let C denote the collection of cliques in G, a *potential function on clique* $c \in C$ is a measurable mapping $V_c : \Lambda^S \to R$, which depends only on variables $\{X_s, s \in c\}$. A *irreducible potential function* V_c is either zero everywhere on Λ^S or could not be represented as sum of non-zero potential functions on cliques which are the subsets of c.

According to [11], the pairwise Markov property of GMRF $\mathbf{X} \sim N(\mathbf{u}, \mathbf{J}^{-1})$ implies that $i \not\sim j \Leftrightarrow J_{ij} = r_{ij} = 0$, where $\mathbf{J} = \mathbf{\Sigma}^{-1}$ is the information matrix of $\mathbf{X}$. Another important fact is that GMRFs have only singleton and pairwise cliques

This research was supported in part by the NSF project CNS-0905603 and by the DoD DTRA project HDTRA1-09-1-0032.

[1] We use the standard notation in graph theory, namely $s \sim t$, to denote this adjacency relation in undirected graph, otherwise it is $s \not\sim t$.

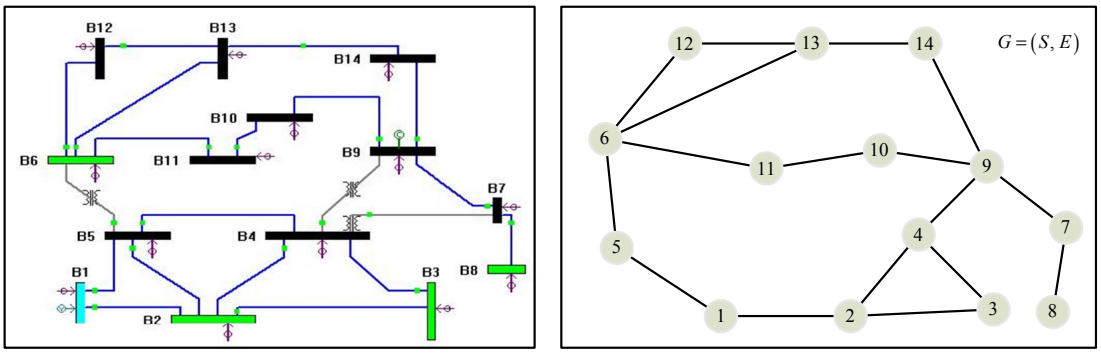

Fig. 1. The IEEE 14-bus system [5] and dependence graph model for the phasor angles

[7], with non-zero entries of $\mathbf{J}$ as the coefficients of the non-zero irreducible potential functions.

II. Problem Formulation

The DC power flow model [8] is often used for analysis of power systems in normal steady-state operations, where the power flow on the transmission line connecting bus i to bus j is given by:

$$Z_{ij} = b_{ij}(X_i - X_j), \tag{2}$$

where X_i and X_j denote the phasor angles at bus i and j, respectively, and b_{ij} denotes the inverse of line inductive reactance. Therefore, the power injection to bus i equals the algebraic sum of the power flowing away from bus i:

$$Z_i = \sum\nolimits_{j\neq i} Z_{ij} = \sum\nolimits_{j\neq i} b_{ij}\left(X_i - X_j\right) \tag{3}$$

In the above formulation, the summation holds since $b_{ij} = 0$ is implied whenever bus i and j are not connected. Thus, it follows that the phasor angle at bus i could be represented as:

$$X_i = \sum\nolimits_{j\neq i} c_{ij} X_j + \frac{1}{\sum_{j\neq i} b_{ij}} Z_i, \tag{4}$$

where $c_{ij} = b_{ij} / \sum_{j\neq i} b_{ij}$.

Existing probabilistic power flow approaches [9], [12] usually model the load flows at buses as random variables, to account for the random disturbances and uncertainty of various load profiles. Some of these approaches [12], [13] model the aggregated load flow injection at buses as Gaussian random variables. Following these approaches, we assume that during the observation window of fault diagnosis applications, the flow injection which originates from the aggregated load requests of a large number of users can be well approximated by Gaussian random variables.

According to [14], branch flows could be expressed as linear combination of the power injection at buses. It follows that branch flow Z_{ij} could also be modeled as a Gaussian random variable. Since the phase difference of buses should not be too large in order to drive branch flows [15], the linear relationship shown in (2) implies that the difference of phasor angles across a bus could be approximated by a Gaussian random variable truncated within $[0, 2\pi)$. In power systems, since the phasor is fixed at the slack bus, we assume that, under steady-state, the phasor angles at non-slack buses could be approximately modeled as Gaussian random variables.

In a nutshell, with $\mathbf{X}_{-i}$ denoting the sites except X_i, the conditional distribution of X_i could be specified in the form of conditional auto-regression (CAR) model [11][2]:

$$X_i \,|\mathbf{X}_{-i} \sim N\left(u_i + \sum\nolimits_{j\neq i} r_{ij}\left(x_j - u_j\right), 1\right) \tag{5}$$

It is shown in [11] that under mild conditions, the joint distribution of the GMRF $\mathbf{X}$ follows $N(\mathbf{u}, (\mathbf{I} - \mathbf{R})^{-1})$, with $\mathbf{R} \triangleq [r_{ij}]$ as the matrix consisting of partial correlation coefficients. Note that for each X_i, its partial correlation coefficients $\{r_{ij}, j \neq i\}$ are proportional to $\{c_{ij}, j \neq i\}$. We have a few key observations in order. 1) The dependency graph of phasor angles agrees with the topology of power systems as illustrated in Fig. 1; 2) Similar to the susceptance matrix $B \triangleq [b_{ij}]$, the partial correlation matrix $\mathbf{R}$ also reflects the electrical distance between buses. Intuitively, *the reduction in the electrical connectivity of buses would result in less partially-correlated phasor angles; further, r_{ij} vanishes if a line outage takes place between buses i and j.*

Let E' be the edge set, $\mathbf{\Sigma}'$ be the covariance matrix of GMRF, and $\mathbf{R}'$ be the partial correlation matrix when the power system is under normal conditions. When fault events take place, some edges might fail and the partial correlation matrix of GMRF would change. Mathematically, the proposed fault detection and localization approach boils down to hypothesis testings on the changes of partial correlations, with null hypothesis given by

$$\mathcal{H}_0\text{: } \{\text{there is no change in } r_{ij},\ \forall \{i,j\} \in E'\ \}$$

One main difficulty in performing the above hypothesis testings originates from the fact that the observations of $\{X_i, X_j\}$ could only help to estimate the correlation coefficient ρ_{ij} between X_i and X_j, rather than the change of r_{ij}. Specifically, observe the walk-sum expression [16]:

$$\rho_{ij} = \left(\left(\mathbf{I} - \mathbf{R}\right)^{-1}\right)_{i,j} = \sum_{l=0}^{\infty} \left(\mathbf{R}^l\right)_{i,j} = \sum_{l=0}^{\infty} \varphi\left(i \xrightarrow{l} j\right), \tag{6}$$

[2]We abuse the notation and use $X_i\,|\mathbf{X}_{-i}$ to denote the conditional distribution of X_i given $\{X_j = x_j,\ j \neq i\}$; and the distribution is normalized to highlight the partial correlations.

where $\varphi\left(i \xrightarrow{l} j\right)$ is defined as the walk on G from site i to j with length l. It is clear that ρ_{ij} depends on the products of partial correlations over all possible walks from i to j, and therefore the change of ρ_{ij} could be a result of some changes in $\mathbf{R}$ other than r_{ij}. Accordingly, it is necessary to obtain an complete estimate of $\mathbf{R}$. Since $\mathbf{I}-\mathbf{R}$ could be computed from $\mathbf{J}$ through normalization, this boils down to estimating $\mathbf{J}$ from the observations.

Another challenge is the requirement on the sparsity of $\hat{\mathbf{J}}$, the estimate of $\mathbf{J}$. Although the sample covariance matrix $\hat{\mathbf{\Sigma}}$ is an unbiased estimate for $\mathbf{\Sigma}$, $\hat{\mathbf{\Sigma}}^{-1}$ may not be a good estimate for $\mathbf{J}$, since $\hat{\mathbf{\Sigma}}^{-1}$ might not have the same sparsity as $\mathbf{J}$, due to noisy observations or small number of samples. Therefore, it is critical for $\hat{\mathbf{J}}$ to have desired sparsity.

In related work [17], the estimation of the information matrix of GMRF is often treated as a constrained optimization problem that maximizes the likelihood:

$$\begin{aligned} \text{maximize} \quad & \log\left|\hat{\mathbf{J}}\right| - tr\left(\hat{\mathbf{J}}\hat{\mathbf{\Sigma}}\right) \\ \text{subject to} \quad & \hat{J}_{ij}=0, \quad (i,j)\notin E' \end{aligned}$$

The solution to the above problem often requires centralized computation and global observations. As noted in [17], the computational complexity could be very high for large-scale problems and current algorithms are not scalable. Worth noting is that the estimation of $\mathbf{J}$ generally requires the number of observations at least to be comparable to the size of $\mathbf{X}$.

III. Decentralized Approach for Fault Detection and Localization

We devise a scheme using the multi-resolution transform on GMRF which could reconstruct the GMRF from the subfields, and by using a multi-scale message-passing procedure, we find a global solution for fault localization. Compared with the centralized algorithm, the proposed approach has the following salient features. 1) with local observations, the number of the samples required is bounded by the size of the largest subfield, rather than the complete GMRF. 2) The dimension of subproblems is much smaller, significantly reducing the computational complexity.

Specifically, for a power system consisting of several subsystems, we decompose the hypothesis testing problem into multiple subproblems, in which the inference can be carried out based on local observations. We note that a direct decomposition of the GMRF, by grouping the sites into K disjoint sub-fields, is not capable of capturing the complete dependence structure of the GMRF. With this insight, we construct an additional sub-field, such that in the dependence graph, all the sites that have neighbors at the other sub-fields are contained in this additional sub-field. In a nutshell, the sites are grouped into *border sites* and *inner sites*, of which the latter are not connected to the other sub-fields. Furthermore, there are three classes of edges: *tie-line edges* which connect different sub-fields, *border-line edges* which connect border sites of the same sub-field, and *inner-line edges* which have at least one end as inner site.

In solving the $K+1$ subproblems, a key challenge is that the graphs of subfields no longer agree with the system topology. Indeed, as discussed in [10], [18], the decomposition on MRF would introduce new edges into the graphs of the subfields. For GMRF, the information matrices of the subfields would have different sparse patterns and non-zero entries from the corresponding diagonal blocks of $\mathbf{J}$. Therefore, the knowledge of local information matrix is not sufficient to identify all of the faults in the concerned subproblem.

To tackle the above challenge, we first show that for the information matrices of the subfields and $\mathbf{J}$, *the entries corresponding to the inner sites, inner-line edges, and tie-line edges remain the same*. Then, we propose to employ message-passing between subfields to "recover" the information about the border sites and border-line edges, lost due to decomposition. We first studied two-scale decomposition of GMRF, and proved that $\mathbf{J}$ *could be reconstructed from the information matrices of the subfields through message-passing*. Finally, we extend the study to the multi-scale decomposition of GMRF.

A. Two-scale Decomposition of GMRF

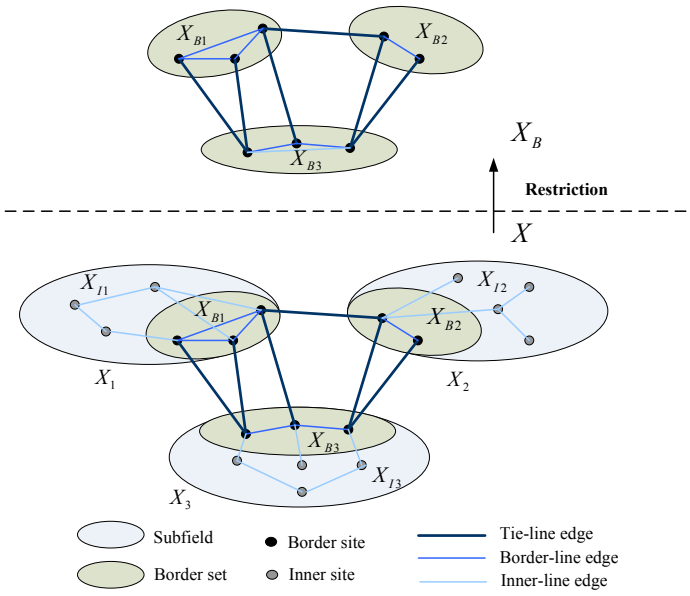

Fig. 2. Two-scale decomposition of GMRF

The basic idea of two-scale decomposition is to group the collection of border sites as the subfield at higher scale, as shown in Fig. 2, which connects the K disjoint subfields at lower scale. Let G, $\tilde{G}_B$, and $\tilde{G}_k$ $k=1,2,\cdots,K$ be the graphs induced by the minimal neighborhood systems of GMRF $\mathbf{X}$, $\mathbf{X}_B$, and $\mathbf{X}_k$, $k=1,2,\cdots,K$, respectively; let C, $\tilde{C}_B$, and $\tilde{C}_k$, $k=1,2,\cdots,K$ be the corresponding collections of cliques. Further, we use $\mathbf{J}$, $\mathbf{J}_B$, and $\mathbf{J}_k$ to denote the respective information matrices.

Since for GMRF, the non-zero entries of information matrix are the coefficients of potential functions, thus we study in depth the graphs and potential functions of the subfields. First,

on the relationship between G, $\tilde{G}_B$, and $\tilde{G}_k$ $k = 1, 2, \cdots, K$, we have the following result.

Lemma 3.1: **a**) For $i \in B_m$, and $j \in B_q$ with $m \neq q$, $i \sim j$ in G_B if and only if $i \sim j$ in G.
b) For $i \in I_k$, and $j \in B_k \cup I_k$ with $i \neq j$, $i \sim j$ in $\tilde{G}_k$ if and only if $i \sim j$ in G .

Next, we give the relationship between the clique functions of $\mathbf{X}$, $\mathbf{X}_B$, and $\mathbf{X}_k$ $k = 1, 2, \cdots, K$.

Lemma 3.2: **a**) For $c \in C$, if it contains two sites connected by tie-line edges, then $c \in \tilde{C}_B$, and the potential function $\tilde{V}c(\mathbf{x}_B) = Vc(\mathbf{x}_B, \mathbf{x}_{\bar{B}})$, in which $\mathbf{x}_{\bar{B}} \in \Lambda^{S \backslash B}$ could be any configuration on $\mathbf{X} \backslash \mathbf{X}_B$.
b) For $c \in C$, if it contains a inner site of $\mathbf{X}_k$, then $c \in \tilde{C}_k$, and the potential function $\tilde{V}c(\mathbf{x}_k) = Vc(\mathbf{x}_k, \mathbf{x}_{\bar{k}})$, in which $\mathbf{x}_{\bar{k}} \in \Lambda^{S \backslash \{B_k \cup I_k\}}$ could be any configuration on $\mathbf{X} \backslash \mathbf{X}_k$.

Remarks: Lemma 3.1 and Lemma 3.2 suggest that: 1) The off-diagonal entries in $\mathbf{J}_B$ and $\mathbf{J}$ corresponding to the same tie-line edge have the same value; 2) For $\mathbf{J}_k$, $k = 1, 2, \cdots, K$ and $\mathbf{J}$, the off-diagonal entries corresponding to the same inner-line edge have the same value, and the diagonal entries corresponding to the same inner site have the same value.

Further, in order to account for the complete reconstruction of $\mathbf{J}$, we consider the border sites and border-line edges and corresponding potential functions. GMRF $\mathbf{X}$ could be organized such that:

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_B \\ \mathbf{X}_I \end{bmatrix}$$

in which

$$\mathbf{X}_B = \begin{bmatrix} \mathbf{X}_{B_1} \\ \cdots \\ \mathbf{X}_{B_K} \end{bmatrix} \quad \mathbf{X}_I = \begin{bmatrix} \mathbf{X}_{I_1} \\ \cdots \\ \mathbf{X}_{I_K} \end{bmatrix}$$

Then, $\mathbf{J}$ has four block matrices $\mathbf{J}_{BB}$, $\mathbf{J}_{BI}$, $\mathbf{J}_{IB}$, and $\mathbf{J}_{II}$, with border-line edges and border sites corresponding to the diagonal blocks of $\mathbf{J}_{BB}$. According to Lemma 3.2, $\mathbf{J}_{BI}$, $\mathbf{J}_{IB}$, and $\mathbf{J}_{II}$ have the same value as corresponding blocks of $\mathbf{J}_k$, $k = 1, 2, \cdots, K$. Another key observation on the blocks of $\mathbf{J}$ and $\mathbf{J}_B$ leads to the following result.

Proposition 3.1: The k-th diagonal blocks of $\mathbf{J_{BB}}$ is given by:

$$[\mathbf{J}_{BB}]_{kk} = [\mathbf{J}_B]_{kk} + [\mathbf{J}_{BI}]_{kk} [\mathbf{J}_{II}]_{kk}^{-1} [\mathbf{J}_{BI}]_{kk}^T, \qquad (7)$$

where the subscript kk denotes the k-th diagonal blocks in the information matrices, respectively.

Remarks: Proposition 3.1 reveals that the entries of $\mathbf{J}$ corresponding to the border-line edges and border sites could be computed by the blocks of the information matrices of subfields.

Combining Lemma 3.2 with Proposition 3.1, we conclude that $\mathbf{J}$ could be completely reconstructed from $\mathbf{J}_B$ and $\mathbf{J}_k$, $k = 1, 2, \cdots, K$, through message-passing between the subfields. For large-scale GMRF, even if the size of each subfield $\mathbf{X}_k$ is moderate, the size of $\mathbf{X}_B$ could still be very large, which makes the estimation of $\mathbf{J}_B$ intractable. This problem could be solved by applying further decomposition on $\mathbf{X}_B$, which gives the multi-scale decomposition of GMRF.

B. Multi-scale Decomposition of GMRF

According to Lemma 3.1, one salient feature of the proposed decomposition of GMRF is that the tie-line structure of GMRF is preserved in $\mathbf{X}^B$, thus the two-scale decomposition could be applied on the subfield at higher scale recursively for multiple times as needed.

In the L-scale decomposition scenario, we use the superscript as the index of scale, and use $\mathbf{X}^2$ to denote the subfield $\mathbf{X}^B$ at the second scale. Suppose that $\mathbf{X}^l$ is obtained by grouping the border sites of the K^{l-1} clusters of $\mathbf{X}^{l-1}$. To obtain $\mathbf{X}^{l+1}$, we re-group the K^{l-1} clusters in $\mathbf{X}^l$ into K^l macro clusters, such that every pair of sites in the same cluster of $\mathbf{X}^{l-1}$ are still contained in the same macro cluster of $\mathbf{X}^l$. For each macro cluster $\mathbf{X}_k^l$, the sites and edges are re-classified according to the same rule. Then, the entire set of edges could be partitioned into disjoint subsets of scale-1 inner-line edges, ..., scale-$(L-1)$ inner-line edges, scale-$(L-1)$ border-line edge, and scale-$(L-1)$ tie-line edges[3]. The set of sites are classified in the same manner as in two-scale decomposition.

Based on previous results, if the information matrices $\mathbf{J}_k^l$ of subfields $\mathbf{X}_k^l$, $k = 1, 2, \cdots, K^l$, at all the L scales are known, then the complete $\mathbf{J}$ could be obtained by performing the reconstruction procedure recursively from top to lowest scale, as stated in Proposition 3.1.

C. Decentralized Fault Detection and Localization

If the fault diagnosis is performed for the entire power system and the phasors at all buses are observable, then the complete estimation of $\mathbf{J}$ is possible. Simply put, we first perform multi-scale decomposition on $\mathbf{X}$. Once the estimation of subfields is done, the complete $\hat{\mathbf{J}}$ could be reconstructed based on message-passing between subfields. For each scale l ($l = 1, 2, \cdots, L$), we assume that there is an inference center at each subfield $\mathbf{X}_k^l$ with $k = 1, 2, \cdots, K^l$, and let $F(k, l)$ be the set of indices for the subfields at the lower scale of $\mathbf{X}_k^{l+1}$. Then, the procedure of the multi-scale estimation is summarized in Algorithm 1.

Algorithm 1 Decentralized estimation of $\mathbf{J}$ with multi-scale message-passing

Local estimation: Estimate $\hat{\mathbf{J}}_k^l$ for each $\mathbf{X}_k^l$ based on the local observations.
Down-top message-passing: For $l = 1, 2, \cdots (L-1)$, the inference centers at $\mathbf{X}_f^l$, $f \in F(k, l)$, submit $\hat{\mathbf{J}}_f^l$ to that of $\mathbf{X}_k^{l+1}$.
Top-down reconstruction: For $l = L-1, L-2, \cdots, 1$, the inference center at $\mathbf{X}^L$ reconstruct $\hat{\mathbf{J}}^l$ from $\hat{\mathbf{J}}^{l+1}$ and $\hat{\mathbf{J}}_k^l$, $k = 1, 2, \cdots, K^l$.
Top-down message-passing: The inference center at $\mathbf{X}^L$ broadcast $\hat{\mathbf{J}}$ to all the subfields.

In some practical scenarios, the power system is not completely observable, e.g., due to the failures of some PMUs.

[3]which are also scale-L edges since we cease clustering for the scale L.

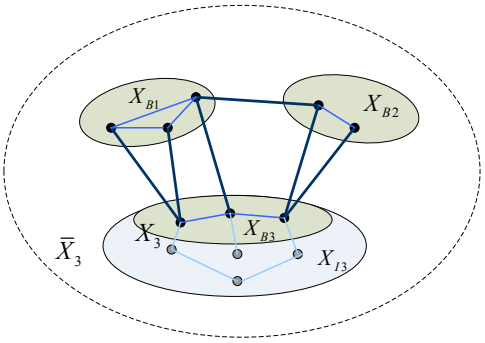

Fig. 3. The extended subfield for fault diagnosis of a sub-system

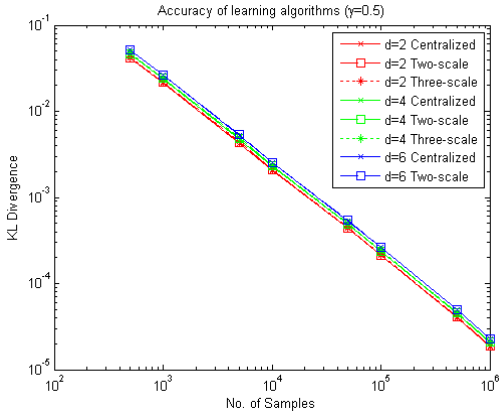

Fig. 4. Estimation algorithms ($\gamma = 0.5$)

As a result, the complete $\hat{\mathbf{J}}$ could not be reconstructed. For this scenario, we show that under mild conditions the fault diagnosis could still be performed for the sub-systems that are observable. Specifically, let $\mathbf{X}_k$ be the subfield of the concerned sub-system, and then it suffices to have the knowledge of $\mathbf{J}_{kk}$, which denotes the k-th diagonal block of $\mathbf{J}$. To deal with this situation, we have the following result.

Corollary 3.1: The observations at $\mathbf{X}_k$ and its neighbor sites form sufficient statistics for the estimation of $\mathbf{J}_{kk}$.

To illustrate further, if fault diagnosis is performed for subfield $\mathbf{X_3}$ in Fig. 2, then we construct an extended subfield $\bar{\mathbf{X}}_\mathbf{3}$ as the concatenation of $\mathbf{X_3}$, and its neighbor sites in $\mathbf{X_{B1}}$ and $\mathbf{X_{B2}}$(suppose they are observable), as shown in Fig. 3. Then, according to Corollary 3.1, $\mathbf{J}_{33}$ is same as the block of the information matrix of $\bar{\mathbf{X}}_\mathbf{3}$ corresponding to $\mathbf{X_3}$.

IV. Numerical Experiment

A. Decentralized Estimation of Information Matrix

In this section, we perform numerical experiments to evaluate the proposed estimation algorithm. We consider a GMRF with $p = 300$ sites uniformly distributed on a $40 * 40$ grid. Any two sites i, j are considered to be neighbors whenever the Euclidean distance $d(i, j)$ is less than a cut-off value d, and the partial correlation coefficients of neighbors are parameterized as: $r_{ij} = \frac{\gamma}{d^2(i,j)}$. We choose $\gamma \in [0, 1]$ such that $(\mathbf{I} - \mathbf{R})$ is positive definite. We generate *i.i.d* samples and perform the estimation procedures. For the multi-scale estimation, we apply the clustering algorithm in [19] and group the GMRF into 16 clusters at the second scale, and 4 clusters at the third scale. The accuracy is quantified by the Kullback-Leibler (KL) distance between the true distribution $N_0(0, \mathbf{J}^{-1})$ and the distribution $N_1(0, \hat{\mathbf{J}}^{-1})$ estimated from samples:

$$D\left(N_0||N_1\right) = \frac{1}{2}\left(\log\det\left(\mathbf{J}\hat{\mathbf{J}}^{-1}\right) + tr\left(\mathbf{J}^{-1}\hat{\mathbf{J}}\right)\right) \qquad (8)$$

The graph density is specified by the value of d. As shown in Fig. 4, the numerical results indicate that for all testing cases, KL distance decreases approximately in power law against the number of samples, and the proposed algorithms have comparable accuracy with the centralized one[4]. It also indicates that the accuracy does not depend on the topology of graph.

[4]For $d = 6$, graph for the subfield at the 2nd scale is so dense that further clustering gives no gains.

In practical applications, we need to take into account the estimation errors when using $\hat{\mathbf{R}}'$ and $\hat{\mathbf{R}}$ for the detection and localization of the faulted lines. In the event of a line outage between the buses i and j, X_i and X_j become partially uncorrelated, and then, we could use the following rule for the identification of line outage:

Line $\{i, j\}$ *failed, if* $\hat{r}'_{ij} \gg \epsilon$ *and* $\hat{r}_{ij} < \epsilon$,

where $\epsilon \ll 1$. For the fault events caused by physical damages, the change of physical parameters usually refers to the increase of line inductive reactance, and thus larger electrical distance between the buses. Therefore, we could use this rule for the identification of such faults:

Line $\{i, j\}$ *with parameter changed, if* $\hat{r}'_{ij}/\hat{r}_{ij} \geqslant \eta$,

where $\eta > 1$ may vary depending on the significant level of the fault events.

B. Fault Localization

In order to evaluate the effectiveness of our algorithms, the proposed approach is tested on the IEEE 300-bus system with simulated data. We use Monte Carlo simulation approaches, and consider the fault diagnosis problem in a 20-second observation window. Since phasor is measured at up to 60 samples per second according to the specification of phasor measurement unit(PMU), thus $n = 1200$ samples are simulated. Power injections to buses are generated, and the branch flows and phasors at buses are solved using the MATPOWER [5] simulation package. We take the solutions of phasors as the measurements collected by PMUs from the power systems.

1) Decomposition of IEEE 300-bus System: We perform two-scale decomposition on the graph of the phasor angles of IEEE 300-bus System. Since IEEE 300 bus system contains three management areas, we take the three areas as the subfields at the lower scale, and the top-scale subfield is obtained by grouping the border sites of the subfields at the lower scale, as shown in Fig. 5.

[5]Http://www.pserc.cornell.edu/matpower/.

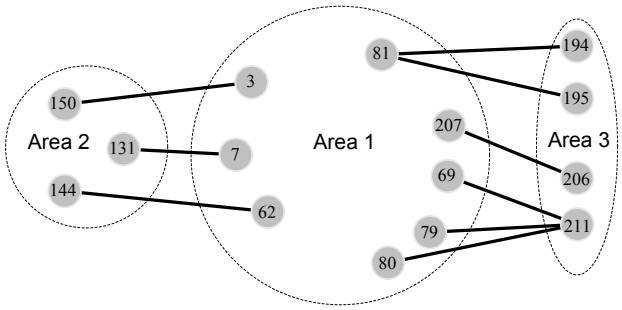

Fig. 5. The top-scale subfield of 300 Bus System

2) Testing cases for line outage: We test the case of line {26,27} outage. To simulate line outage, we exclude the failed line from solving the power flows. After the estimation of the information matrix of phasor angles, the partial coefficients are computed and shown in Table I. It shows that the partial correlation coefficient of bus 26 and 27 turns into small value. Thus the decision rule with $\epsilon = 0.05$ could detect the line outage.

TABLE I
A TESTING CASE FOR LINE {26,27} OUTAGE

Case1	line	Area	Reactance	Par. Cor. Coef.	
				$\hat{r}'$	$\hat{r}$
Bus 26	{25,26}	1	0.071	0.61	0.62
	{26,27}	1	0.12	0.36	0.03
	{26,320}	1	0.13	0.34	0.39
Bus 27	{20,27}	1	0.186	0.23	0.25
	{26,27}	1	0.12	0.36	0.03

3) Testing cases for the change of physical parameter: We test the case that the reactance of line {199,200} changes. To simulate the increase of electrical distance, we double the reactance of the concerned transmission line. The simulated results are shown in Table II. If we choose $\eta = 1.3$, then the fault event could be detected.

TABLE II
A TESTING CASE FOR LINE {199,200} CHANGE

Case1	line	Area	Reactance		Par. Cor. Coef.	
			$1/b'_{ij}$	$1/b_{ij}$	$\hat{r}'$	$\hat{r}$
Bus 199	{199,200}	2	0.135	0.27	0.53	0.38
	{199,210}	2	0.102	0.102	0.68	0.75
Bus 200	{199,200}	2	0.135	0.27	0.53	0.38
	{200,210}	2	0.128	0.128	0.58	0.63
	{200,248}	2	0.22	0.22	0.38	0.41

V. CONCLUSION

We propose a new stochastic fault diagnosis approach for power systems. Specifically, based on existing models and approaches, we develop a GMRF model for the phasor angles across the buses. The fault diagnosis is performed through change detection and localization of the partial correlation matrix of GMRF. In particular, to mitigate the complexity introduced by the estimation of information matrices, we propose decentralized algorithms based on the multi-scale transform.

REFERENCES

[1] R. Leon, V. Vittal, and G. Manimaran, "Application of sensor network for secure electric energy infrastructure," *IEEE Trans. Power Del.*, vol. 22, pp. 1021–1028, April 2007.

[2] J. Tate and T. Overbye, "Line outage detection using phasor angle measurements," *IEEE Trans. Power Syst.*, vol. 23, pp. 1644–1652, Nov. 2008.

[3] J. Tate and T. Overbye, "Double line outage detection using phasor angle measurements," in *IEEE PES 09*, pp. 1–5, July 2009.

[4] N. Zhang and M. Kezunovic, "Improving real-time fault analysis and validating relay operations to prevent or mitigate cascading blackouts," in *Transmission and Distribution Conference and Exhibition, 2005/2006 IEEE PES*, pp. 847–852, May 2006.

[5] J. Zhu and A. Abur, "Improvements in network parameter error identification via synchronized phasors," *IEEE Trans. Power Syst.*, vol. 25, pp. 44–50, Feb. 2010.

[6] N. Katenka, E. Levina, and G. Michailidis, "Robust target localization from binary decisions in wireless sensor networks," Nov. 2008.

[7] M. J. Wainwright and M. I. Jordan, "Graphical models, exponential families, and variational inference," *Found. Trends Mach. Learn.*, vol. 1, no. 1-2, pp. 1–305, 2008.

[8] R. Christie, B. Wollenberg, and I. Wangensteen, "Transmission management in the deregulated environment," *Proceedings of the IEEE*, vol. 88, pp. 170–195, Feb 2000.

[9] R. Allan, B. Borkowska, and C. Grigg, "Probabilistic analysis of power flows," *Electrical Engineers, Proceedings of the Institution of*, vol. 121, pp. 1551 –1556, December 1974.

[10] P. Perez and F. Heitz, "Restriction of a markov random field on a graph and multiresolution statistical image modeling," *IEEE Trans. Inf. Theory*, vol. 42, pp. 180–190, Jan 1996.

[11] H. Rue and L. Held, *Gaussian Markov Random Fields: Theory and Applications*. Chapman & Hall/CRC, 2005.

[12] J. Dopazo, O. Klitin, and A. Sasson, "Stochastic load flows," *IEEE Trans. Power App. Syst.*, vol. 94, pp. 299 – 309, Mar 1975.

[13] A. Schellenberg, W. Rosehart, and J. Aguado, "Cumulant-based probabilistic optimal power flow (p-opf) with gaussian and gamma distributions," *IEEE Trans. Power Syst.*, vol. 20, pp. 773–781, May 2005.

[14] A. Meliopoulos, G. Cokkinides, and X. Chao, "A new probabilistic power flow analysis method," *IEEE Trans. Power Syst.*, vol. 5, pp. 182 –190, Feb 1990.

[15] M. G. Morgan, J. Apt, L. B. Lave, M. D. Ilic, M. Sirbu, and J. M. Peha, "The many meanings of smart grid." http://wpweb2.tepper.cmu.edu/ceic/papers/SmartGrid.htm.

[16] D. M. Malioutov, J. K. Johnson, and A. S.Willsky, "Walk-sums and belief propagation in gaussian graphical models," *J. Mach. Learn. Res.*, vol. 1, no. 7, p. 2031C2064, 2006.

[17] J. Dahl, V. Roychowdhury, and L. Vandenberghe, "Covariance selection for nonchordal graphs via chordal embedding," *Optimization Methods and Software*, vol. 23, no. 4, pp. 501–520, 2005.

[18] S. Krishnamachari and R. Chellappa, "Multiresolution gauss-markov random field models for texture segmentation," *IEEE Transactions on Image Processing*, vol. 6, pp. 251 –267, feb 1997.

[19] M. Girvan and M. Newman, "Community structure in social and biological networks," *Proceedings of the National Academy of Sciences*, vol. 99, no. 1, p. 7821C7826, 2002.