Analyzing and Quantifying the Effect of k-line Failures in Power Grids

Saleh Soltan, Alexander Loh, Gil Zussman
Department of Electrical Engineering
Columbia University, New York, NY
Emails: {saleh,gil}@ee.columbia.edu, al3475@columbia.edu

***Abstract*—Contingency analysis in power grids is one of the most effective ways to improve grids' resilience against failures. The main goal of contingency analysis is to detect probable failures in the grid that result in a critical state and deploy preventive measures to avoid such state. Due to the large number of possibilities, however, high order contingency analysis is computationally expensive and not fully deployed. In order to circumvent this issue, we analytically compute the redistribution of power flows following a k-line failure (i.e., failures in k distinct lines) using the DC power flow model and based on that introduce the *disturbance value* of a k-line failure. We show that this value can be efficiently computed in $O(1)$ for any set of line failures independently of the size of the grid and can be effectively used to filter out noncritical contingencies. The disturbance value can therefore significantly reduce the time complexity of contingency analysis by revealing contingencies that are vital for more in depth analysis and pave the way for the deployment of high order contingency analysis in power grids.**

I. Introduction

Recent large scale power grid blackouts in Turkey (2015), India (2013), and the U.S. (2003) exposed the insufficiency of current control and preventive measures to protect the grid against failures. These events motivated numerous research projects in the past decade aiming to improve grid's resilience and security. One of the most effective ways to improve power grids' resilience is to detect probable failures that result in a critical state and deploy preventive measures to avoid such state [2]. This process is known as the *contingency analysis* in power grids.

Power grids are required to withstand a single component failure, known as the $N-1$ contingency standard. However, one or more lines can often be out of service for various reasons such as maintenance and construction works. This can result in the violation of the $N-1$ standard and vulnerability of the power grid to a single failure, as was the case in Turkey in 2015 [3]. Therefore, higher order contingency analysis is necessary to detect critical events caused by more than a single failure.

Due to the large number of possibilities, high order contingency analysis is computationally expensive. Therefore, most of the previous tools are effective for analyzing contingencies caused by failures in one or two components of the grid [4], [5], [6], [7], [8], [9], [10], [11]. In order to circumvent this issue, in this paper, we introduce the *disturbance value* of a failure. We show that this value can be efficiently computed for any set of line failures independently of the size of the grid and can be effectively used to filter out less crucial contingencies. The disturbance value can therefore significantly reduce the time complexity of contingency analysis by revealing contingencies that are vital for more in depth analysis.

First, we extend and build on the results in [12] which focused on single line failures to analytically compute the effect of *multiple line failures* on redistribution of power flows. We call an event resulting in the failure of k distinct lines, a *k-line failure event*. Similar to the most of the previous work on contingency analysis, we use the linearized DC approximation of the power flows, due to the complexities associated with the AC power flow model [13]. We present an analytical update of the pseudo-inverse of the admittance matrix after a k-line failure event. Our approach is similar to [14], but we use the pseudo-inverse instead of the truncated inverse of the admittance matrix which requires selecting a node as the slack bus. An advantage of using pseudo-inverse is that it allows having a unified formula for all k-line failures regardless of their location and their connectivity to the slack bus.[1] Using this result, we define and analytically compute the *k-line outage distribution matrix* which generalizes the definition of the line outage distribution factors for single line failures [2] as in [14].

While the k-line outage distribution matrix captures all effects of a k-line failure on flow changes, it is not efficient to compute and store this matrix for contingency analysis in large power grids. To overcome this challenge, we use the *matrix of equivalent reactance values* originally defined and used in [12] to efficiently compute the sum of changes in the power flows after a k-line failure and to provide a metric that captures the essence of flow changes after failures. In particular, we define and analytically compute the *disturbance value* of a failure (the weighted sum of squares of the flow changes) and show that this computation can be done for a k-line failure in $O(1)$ as long as k is much smaller than the total number of lines, the case in contingency analysis of the power grids. Hence, the disturbance value of a k-line failure can be computed independently of the size of the grid.

To show that the disturbance values provide a separation

A partial and preliminary version appeared in Proc. IEEE PES-GM'16 [1].

[1] Another advantage of using pseudo-inverse is that our results can be tied to the related notions of *Laplacian* and *resistance distances* in graph theory [15]. This enables further use of existing tools in graph theory for contingency analysis in power grids as in [12].

between failures with higher impact and lower impact, we compute the disturbance values for all possible choices of 3-line failures in the IEEE 118-bus and 300-bus systems. We demonstrate that by ranking cases based on their disturbance values and considering only cases with high disturbance value, we are able to decrease the total number of cases needed to be analyzed for contingency analysis by more than 90%.

To analyze all the k-line failure scenarios, one needs to compute the disturbance values for all possible $\binom{m}{k}$ cases which requires $O(m^k)$ time. To alleviate this issue, we provide an approximation for the disturbance values. We show that by approximating the disturbance value of a k-line failure, one can only focus on the $\widetilde{m} << m$ lines with the highest 1-line disturbance values and analyze only $\binom{\widetilde{m}}{k}$ cases instead of $\binom{m}{k}$ cases. Moreover, we numerically show that the approximation error is below 10% in most of the cases in the 118- and 300-bus systems.

A k-line failure alters the network topology and results in a different flow pattern causing possible consequent line failures. The repetition of this process results in a *cascading failure*. To show that disturbance values effectively rank k-line failures, we numerically compute the relationship between disturbance value of all 3-line failures and the severity of the initiated cascades in the IEEE 118- and 300-bus systems. We show that although the disturbance value does not take into account the capacities of the lines, it can predict the average severity of the cascade initiated by a failure.

Finally, we numerically show the usefulness of the disturbance values in predicting changes under the more detailed AC power flows after a k-line failure. In particular, we compute the disturbance values under the AC power flows and show that the disturbance values under the AC and DC differ only in scaling. We also study the correlation between the voltage changes after a k-line failure and disturbance values, and show that k-line failures with higher disturbance values cause more voltage changes as well.

Since the AC power flow equations are nonlinear, there are k-line failures after which the AC power solver does not converge to a solution. We show that almost all such failures in the IEEE 118- and 300-bus systems have the disturbance values that are between the top 10% of all the k-line failures in terms of the disturbance values.

The main contributions of this paper are: (i) providing an analytical update for the pseudo-inverse of the admittance matrix following a k-line failure and using it to define and compute k-line outage distribution matrices, (ii) introducing the disturbance value of a k-line failure and demonstrating based on simulations that it can distinguish between critical and noncritical contingencies in the grid, and (iii) leveraging the properties of the DC power flow equations to provide a method to compute the disturbance value of a failure efficiently and independently of the size of the grid.

II. Related Work

Many great ideas have been developed for contingency analysis in power grids since the advent of the modern power transmission network. Detecting most important lines and nodes solely based on the topology of the power grid was studied in [16] using network centrality measures. *Current injection* methods were used to analyze the effect of line failures in [2], [5], [8], [11]. In particular, [2] introduced the notion of *line outage distribution factors* that inspired many other studies including the work presented in this paper. In [4] and [12], multiple matrix updates were used to study the effect of single line failures and to speed up the computation of the power flows after line failures. [10], [17] used matrix updates to study the effect of two line failures and used the results to introduce an algorithm for the $N-2$ contingency problem (contingencies caused by failures in at most 2 components of the grid). In a follow up work [18], *contingency and Influence graphs* were introduced to study $N-2$ contingency analysis. More optimization-based techniques for contingency analysis of the grids were explored in [7], [19]. In particular, [7] focused on identifying the most probable failure modes in static load distribution using a linear-program. In a recent innovative paper, probabilistic algorithms were developed to identify collections of multiple contingencies that initiate cascading failure [20].

High order contingency analysis was studied in [6], [21], [22]. In [21], the contingencies were ordered based on their empirical occurrence probabilities and only contingencies with high probability were considered. A mixed-integer model for the $N-k$ contingency problem was presented and used in [6]. However, this method does not scale well as k increases. In [22], the resistance distances are used as in [12] to identify most important lines and to prune the graph. The contingency analysis was then performed on the reduced graph instead of the entire graph in order to reduce the total number of contingencies. However, since resistance distances are computed independently of supply and demand values, such an approach may miss contingencies that are caused by failures in lines carrying large power flows.

The main advantages of using disturbance value for contingency analysis, as proposed in this paper, over previous work is that: (i) using techniques from linear algebra, the disturbance value can be efficiently computed for any number of line failures independently of the size of a grid, and therefore it scales well for high order contingency analysis, (ii) due to its simplicity, it is easy to implement, and (iii) since the disturbance value is based on the power flows as well as the admittance matrix, it captures both the topological and operational properties of the system.

III. Model and Definitions

A. DC Power Flow Model

We adopt the linearized (or DC) power flow model, which is widely used as an approximation for the non-linear AC power flow model [23], [24]. We represent the power grid by a connected directed graph $G = (V, E)$ where $V = \{1, 2, \ldots, n\}$ and $E = \{e_1, \ldots, e_m\}$ are the set of nodes and edges corresponding to the buses and transmission lines, respectively (the definition implies $|V| = n$ and $|E| = m$). Each edge e is a set of two nodes $e = (u, v)$. p_v is the active power *supply* ($p_v > 0$) or *demand* ($p_v < 0$) at node $v \in V$

(for a *neutral node* $p_v = 0$). We assume *pure reactive* lines, implying that each edge $e = (u, v) \in E$ is characterized by its *reactance* $x_e = x_{uv} = x_{vu} > 0$.

Given the power supply/demand vector $\vec{p} \in \mathbb{R}^{n\times 1}$ and the reactance values, a *power flow* is a solution $\vec{f} \in \mathbb{R}^{m\times 1}$ and $\vec{\theta} \in \mathbb{R}^{n\times 1}$ of:

$$\mathbf{A}\vec{\theta} = \vec{p}, \tag{1}$$
$$\mathbf{Y}\mathbf{D}^t\vec{\theta} = \vec{f}, \tag{2}$$

where $\mathbf{A} \in \mathbb{R}^{|V|\times|V|}$ is the *admittance matrix* of G,[2] defined as:

$$a_{uv} = \begin{cases} 0 & \text{if } u \neq v \text{ and } \{u, v\} \notin E, \\ -1/x_{uv} & \text{if } u \neq v \text{ and } \{u, v\} \in E, \\ -\sum_{w\in N(u)} a_{uw} & \text{if } u = v, \end{cases}$$

$\mathbf{D} \in \{-1, 0, 1\}^{n\times m}$ is the *incidence matrix* of G defined as,

$$d_{ij} = \begin{cases} 0 & \text{if } e_j \text{ is not incident to node } i, \\ 1 & \text{if } e_j \text{ is coming out of node } i, \\ -1 & \text{if } e_j \text{ is going into node } i, \end{cases}$$

and $\mathbf{Y} := \text{diag}([1/x_{e_1}, 1/x_{e_2}, \ldots, 1/x_{e_m}])$ is a diagonal matrix with diagonal entries equal to the inverse of the reactance values. It is easy to see that $\mathbf{A} = \mathbf{D}\mathbf{Y}\mathbf{D}^t$.

Since $\mathbf{A}$ is not a full-rank matrix, we follow [12] and use the *pseudo-inverse* of $\mathbf{A}$, denoted by $\mathbf{A}^+$ to solve (1) as: $\vec{\theta} = \mathbf{A}^+\vec{p}$. Once $\vec{\theta}$ is computed, $\vec{f}$, can be obtained from (2).

Notation. Throughout this paper we use bold uppercase characters to denote matrices (e.g., $\mathbf{A}$), italic uppercase characters to denote sets (e.g., V), and italic lowercase characters and overline arrow to denote column vectors (e.g., $\vec{\theta}$). For a matrix $\mathbf{Q}$, q_{ij} denotes its $(i, j)^{\text{th}}$ entry, $\mathbf{Q}^t$ its transpose, and $\text{tr}(\mathbf{Q})$ its trace. We denote the submatrix of $\mathbf{Q}$ limited to the first k columns by $\mathbf{Q}_k$ and the submatrix of $\mathbf{Q}$ limited to the first k rows and columns by $\mathbf{Q}_{k|k}$. For a column vector $\vec{y}$, $\vec{y}^t$ denotes its transpose, and $\vec{y}_k$ denotes the subvector of $\vec{y}$ with its first k entries. We use $\bar{k}$ to show the indices other that 1 to k (e.g., $\vec{f}_{\bar{k}}$ denotes the subvector of $\vec{f}$ with its $k + 1$ to m entries).

B. Failure Model

In this paper, we consider contingencies caused by the set of line failures of size k denoted by $L \subseteq E$. We refer to these failures as *k-line failures*. Without loss of generality, for convenience we assume $L = \{e_1, e_2, \ldots, e_k\}$. We denote the graph after failures by $G' = (V', E')$, in which $E' = E - L$ and $V' = V$. We also assume that removing edges in L from G does not disconnect the graph. Notice that failures that disconnect the grid have the highest priority, since they most likely divide the grid into unbalanced islands in terms of supply/demand values that may result in further failures. Therefore, such failures should always be considered for more in depth contingency analysis. In Section V, we show that some of our results can be used to approximately rank k-line failures that disconnect the grid based on their criticality among themselves.

[2]The admittance matrix $\mathbf{A}$ is also known as the *weighted Laplacian matrix* of the graph [15].

Hence, here we assume G' is connected. Upon failures, the power flows redistributed in G' based on the equation $\mathbf{A}'\vec{\theta}' = \vec{p}$, in which $\mathbf{A}'$ is the admittance matrix of G'. Moreover, we define $\Delta\vec{f}_{\bar{k}} = \vec{f}_{\bar{k}} - \vec{f}'_{\bar{k}}$ to show the flow changes on the lines in $E\backslash L$ after the failure in lines in L.

It is easy to see that $\mathbf{A}' = \mathbf{A} - \mathbf{D}_k\mathbf{Y}_{k|k}\mathbf{D}_k^t$. In Section IV, we use this equation to compute $\mathbf{A}'^+$ and quantify the effect of k-line failures.

C. Matrix of Equivalent Reactance Values

Define matrix $\mathbf{R} \in \mathbb{R}^{m\times m}$ as $\mathbf{R} := \mathbf{D}^t\mathbf{A}^+\mathbf{D}$. It is easy to see that for any $\forall 1 \leq i \leq m : r_{ii}$ is *equivalent reactance* between end buses of the line e_i. Matrix $\mathbf{R}$ is a symmetric matrix and is very useful in quantifying the effect of line failures. In fact, in [12] we used this matrix to quantify the effect of single line failure when all the reactance values are equal to 1. In Section IV, we generalize the idea in [12] for k-line failures.

IV. Failure Analysis

In this section, we study the effect of k-line failures on the flow changes on the other lines. First, in the following lemma, we generalize the results in [12] for single line failures and provide an analytical update of the pseudo-inverse of the admittance matrix following a k-line failure.

Lemma 1: If G' is connected,

$$\mathbf{A}'^+ = \mathbf{A}^+ + \mathbf{A}^+\mathbf{D}_k\mathbf{Y}_{k|k}^{1/2}[\mathbf{I} - \mathbf{Y}_{k|k}^{1/2}\mathbf{D}_k^t\mathbf{A}^+\mathbf{D}_k\mathbf{Y}_{k|k}^{1/2}]^{-1}\mathbf{Y}_{k|k}^{1/2}\mathbf{D}_k^t\mathbf{A}^+.$$

Proof: First, from Lemma A.1 in the Appendix, since G' is connected, $[\mathbf{I} - \mathbf{Y}_{k|k}^{1/2}\mathbf{D}_k^t\mathbf{A}^+\mathbf{D}_k\mathbf{Y}_{k|k}^{1/2}]^{-1}$ is defined. Now to show the equality, it is easy to see that $\mathbf{A}\mathbf{A}^+ = \mathbf{I} - \frac{1}{n}\mathbf{J}$, in which $\mathbf{I}$ is the identity matrix and $\mathbf{J}$ is all 1 matrix (For more details see [12, Theorem 1]). Hence, from [25, Theorem 4.8], since $\mathbf{A}' = \mathbf{A} - \mathbf{D}_k\mathbf{Y}_{k|k}\mathbf{D}_k^t = \mathbf{A} - \mathbf{D}_k\mathbf{Y}_{k|k}^{1/2}(\mathbf{D}_k\mathbf{Y}_{k|k}^{1/2})^t$, the pseudo inverse of $\mathbf{A}'$ can be computed as,

$$\mathbf{A}'^+ = \mathbf{A}^+ + \mathbf{A}^+\mathbf{D}_k\mathbf{Y}_{k|k}^{1/2}[\mathbf{I} - \mathbf{Y}_{k|k}^{1/2}\mathbf{D}_k^t\mathbf{A}^+\mathbf{D}_k\mathbf{Y}_{k|k}^{1/2}]^{-1}\mathbf{Y}_{k|k}^{1/2}\mathbf{D}_k^t\mathbf{A}^+.$$

■

From Lemma 1, the changes in phase angles after k-line failures can be computed as,

$$\begin{aligned}\vec{\theta}' - \vec{\theta} &= (\mathbf{A}'^+ - \mathbf{A}^+)\vec{p} \\ &= \mathbf{A}^+\mathbf{D}_k\mathbf{Y}_{k|k}^{1/2}[\mathbf{I} - \mathbf{Y}_{k|k}^{1/2}\mathbf{D}_k^t\mathbf{A}^+\mathbf{D}_k\mathbf{Y}_{k|k}^{1/2}]^{-1}\mathbf{Y}_{k|k}^{1/2}\mathbf{D}_k^t\mathbf{A}^+\vec{p} \\ &= \mathbf{A}^+\mathbf{D}_k\mathbf{Y}_{k|k}^{1/2}[\mathbf{I} - \mathbf{Y}_{k|k}^{1/2}\mathbf{D}_k^t\mathbf{A}^+\mathbf{D}_k\mathbf{Y}_{k|k}^{1/2}]^{-1}\mathbf{Y}_{k|k}^{-1/2}\vec{f}_k \\ &= \mathbf{A}^+\mathbf{D}_k\mathbf{Y}_{k|k}^{1/2}[\mathbf{I} - \mathbf{Y}_{k|k}^{1/2}\mathbf{R}_{k|k}\mathbf{Y}_{k|k}^{1/2}]^{-1}\mathbf{Y}_{k|k}^{-1/2}\vec{f}_k. \end{aligned} \tag{3}$$

Using (3), we can compute the changes in the flows as,

$$\begin{aligned}\Delta\vec{f}_{\bar{k}} &= \mathbf{Y}_{\bar{k}|\bar{k}}\mathbf{D}_{\bar{k}}^t\mathbf{A}^+\mathbf{D}_k\mathbf{Y}_{k|k}^{1/2}[\mathbf{I} - \mathbf{Y}_{k|k}^{1/2}\mathbf{R}_{k|k}\mathbf{Y}_{k|k}^{1/2}]^{-1}\mathbf{Y}_{k|k}^{-1/2}\vec{f}_k \\ &= \mathbf{Y}_{\bar{k}|\bar{k}}\mathbf{R}_{\bar{k}|k}\mathbf{Y}_{k|k}^{1/2}[\mathbf{I} - \mathbf{Y}_{k|k}^{1/2}\mathbf{R}_{k|k}\mathbf{Y}_{k|k}^{1/2}]^{-1}\mathbf{Y}_{k|k}^{-1/2}\vec{f}_k. \end{aligned} \tag{4}$$

It is important to see that $\mathbf{Y}_{\bar{k}|\bar{k}}\mathbf{R}_{\bar{k}|k}\mathbf{Y}_{k|k}^{1/2}[\mathbf{I} - \mathbf{Y}_{k|k}^{1/2}\mathbf{R}_{k|k}\mathbf{Y}_{k|k}^{1/2}]^{-1}\mathbf{Y}_{k|k}^{-1/2}$ is independent of $\vec{p}$ and solely depends on the structure properties of the network. Hence, following a similar definition in [2] for single line failures, we define this matrix as *k-line outage distribution matrix* and

denote it by $\boldsymbol{\mathcal{L}} := \mathbf{Y}_{\bar{k}|\bar{k}}\mathbf{R}_{\bar{k}|k}\mathbf{Y}_{k|k}^{1/2}[\mathbf{I}-\mathbf{Y}_{k|k}^{1/2}\mathbf{R}_{k|k}\mathbf{Y}_{k|k}^{1/2}]^{-1}\mathbf{Y}_{k|k}^{-1/2}$. Hence, $\Delta\vec{f}_{\bar{k}} = \boldsymbol{\mathcal{L}}\vec{f}_k$.

While k-line outage distribution matrix captures all effects of a k-line failure on the flow changes, it is not efficient to compute and store it for contingency analysis in large power grids. In order to overcome this problem, we use the matrix of equivalent reactance values to efficiently compute the sum of changes in the power flows after k-line failures and to provide a metric to capture the essence of the flow changes after failures. The following lemma is the main step towards this goal. It demonstrates that $\mathbf{Y}^{1/2}\mathbf{R}\mathbf{Y}^{1/2}$ is an idempotent matrix. We use this property, to provide the results in Corollaries 1 and 2.

Lemma 2: $\mathbf{Y}^{1/2}\mathbf{R}\mathbf{Y}^{1/2} = \mathbf{Y}^{-1/2}\mathbf{D}^{+}\mathbf{D}\mathbf{Y}^{1/2}$, and therefore $(\mathbf{Y}^{1/2}\mathbf{R}\mathbf{Y}^{1/2})^2 = \mathbf{Y}^{1/2}\mathbf{R}\mathbf{Y}^{1/2}$.

Proof: We know from before that $\mathbf{R} = \mathbf{D}^t\mathbf{A}^{+}\mathbf{D}$ and $\mathbf{A} = \mathbf{D}\mathbf{Y}\mathbf{D}^t = (\mathbf{D}\mathbf{Y}^{1/2})(\mathbf{D}\mathbf{Y}^{1/2})^t$. Hence,

$$\begin{aligned}\mathbf{Y}^{1/2}\mathbf{R}\mathbf{Y}^{1/2} &= \mathbf{Y}^{1/2}\mathbf{D}^t\mathbf{A}^{+}\mathbf{D}\mathbf{Y}^{1/2} = (\mathbf{D}\mathbf{Y}^{1/2})^t\mathbf{A}^{+}(\mathbf{D}\mathbf{Y}^{1/2})\\ &= (\mathbf{D}\mathbf{Y}^{1/2})^t\big((\mathbf{D}\mathbf{Y}^{1/2})(\mathbf{D}\mathbf{Y}^{1/2})^t\big)^{+}(\mathbf{D}\mathbf{Y}^{1/2})\\ &= (\mathbf{D}\mathbf{Y}^{1/2})^t\big((\mathbf{D}\mathbf{Y}^{1/2})^t\big)^{+}(\mathbf{D}\mathbf{Y}^{1/2})^{+}(\mathbf{D}\mathbf{Y}^{1/2})\\ &= \big((\mathbf{D}\mathbf{Y}^{1/2})^{+}(\mathbf{D}\mathbf{Y}^{1/2})\big)^t(\mathbf{D}\mathbf{Y}^{1/2})^{+}(\mathbf{D}\mathbf{Y}^{1/2}).\end{aligned}$$

From the properties of the pseudo-inverse, $(\mathbf{D}\mathbf{Y}^{1/2})^{+}(\mathbf{D}\mathbf{Y}^{1/2})$ is a symmetric matrix. Moreover, $(\mathbf{D}\mathbf{Y}^{1/2})^{+}(\mathbf{D}\mathbf{Y}^{1/2})(\mathbf{D}\mathbf{Y}^{1/2})^{+} = (\mathbf{D}\mathbf{Y}^{1/2})^{+}$. Therefore,

$$\begin{aligned}\mathbf{Y}^{1/2}\mathbf{R}\mathbf{Y}^{1/2} &= (\mathbf{D}\mathbf{Y}^{1/2})^{+}(\mathbf{D}\mathbf{Y}^{1/2})(\mathbf{D}\mathbf{Y}^{1/2})^{+}(\mathbf{D}\mathbf{Y}^{1/2})\\ &= (\mathbf{D}\mathbf{Y}^{1/2})^{+}(\mathbf{D}\mathbf{Y}^{1/2}) = \mathbf{Y}^{-1/2}\mathbf{D}^{+}\mathbf{D}\mathbf{Y}^{1/2}.\end{aligned}$$

From this,

$$\begin{aligned}(\mathbf{Y}^{1/2}\mathbf{R}\mathbf{Y}^{1/2})^2 &= (\mathbf{Y}^{-1/2}\mathbf{D}^{+}\mathbf{D}\mathbf{Y}^{1/2})^2\\ &= \mathbf{Y}^{-1/2}\mathbf{D}^{+}\mathbf{D}\mathbf{D}^{+}\mathbf{D}\mathbf{Y}^{1/2}\\ &= \mathbf{Y}^{-1/2}\mathbf{D}^{+}\mathbf{D}\mathbf{Y}^{1/2} = \mathbf{Y}^{1/2}\mathbf{R}\mathbf{Y}^{1/2}.\end{aligned}$$

■

Corollary 1: $\mathbf{R}_{k|\bar{k}}\Delta\vec{f}_{\bar{k}} = \mathbf{R}_{k|k}\vec{f}_k$.

Proof: To make equations cleaner in the proof, define $\mathbf{H} := \mathbf{Y}^{1/2}\mathbf{R}\mathbf{Y}^{1/2}$. From Lemma 2, $\mathbf{H}^2 = \mathbf{H}$. Hence, if we use block multiplication, then $\mathbf{H}_{k|k} = \mathbf{H}_{k|k}^2 + \mathbf{H}_{k|\bar{k}}\mathbf{H}_{\bar{k}|k}$. Using this equation,

$$\begin{aligned}&\Delta\vec{f}_{\bar{k}} = \mathbf{Y}_{\bar{k}|\bar{k}}\mathbf{R}_{\bar{k}|k}\mathbf{Y}_{k|k}^{1/2}[\mathbf{I}-\mathbf{Y}_{k|k}^{1/2}\mathbf{R}_{k|k}\mathbf{Y}_{k|k}^{1/2}]^{-1}\mathbf{Y}_{k|k}^{-1/2}\vec{f}_k\\ \Rightarrow&\mathbf{Y}_{k|k}^{1/2}\mathbf{R}_{k|\bar{k}}\Delta\vec{f}_{\bar{k}} = \mathbf{H}_{k|\bar{k}}\mathbf{H}_{\bar{k}|k}[\mathbf{I}-\mathbf{H}_{k|k}]^{-1}\mathbf{Y}_{k|k}^{-1/2}\vec{f}_k\\ \Rightarrow&\mathbf{Y}_{k|k}^{1/2}\mathbf{R}_{k|\bar{k}}\Delta\vec{f}_{\bar{k}} = \mathbf{H}_{k|k}\mathbf{Y}_{k|k}^{-1/2}\vec{f}_k\\ \Rightarrow&\mathbf{Y}_{k|k}^{1/2}\mathbf{R}_{k|\bar{k}}\Delta\vec{f}_{\bar{k}} = \mathbf{Y}_{k|k}^{1/2}\mathbf{R}_{k|k}\vec{f}_k \Rightarrow \mathbf{R}_{k|\bar{k}}\Delta\vec{f}_{\bar{k}} = \mathbf{R}_{k|k}\vec{f}_k.\end{aligned}$$

■

Corollary 1 shows the use of matrix $\mathbf{R}$ in evaluating the effect of k-line failures without computing the flows directly. This equation can be used to estimate the effect of k-line failures. Since the matrix of equivalent reactance values needs to be computed only once, the matrix equation in Corollary 1 can be written for any k-line failures without further computations.

To quantify the effect of k-line failures more efficiently, in the following, we define a metric that captures the effect of k-line failures by a single value and show that it can be computed in $O(1)$. Inspired by the notion of energy in resistive networks, we define $\delta_k(1,2,\ldots,k) := \Delta\vec{f}_{\bar{k}}^{\,t}\mathbf{Y}_{\bar{k}|\bar{k}}^{-1}\Delta\vec{f}_{\bar{k}}$ as the *disturbance value* of a k-line failure. It is easy to see that $y_{ii}^{-1}\Delta f_i$ captures the changes in the phase angle differences between the end buses of a single line. Hence, the disturbance value $\Delta\vec{f}_{\bar{k}}^{\,t}\mathbf{Y}_{\bar{k}|\bar{k}}^{-1}\Delta\vec{f}_{\bar{k}} = \sum_{i=k+1}^{m} y_{ii}^{-1}\Delta f_i^2$ reflects both the big phase difference changes (which is important for the stability of the system) and the big flow changes (which is important for thermal safety of a line). Notice that the disturbance values can be defined also based on the phase angle of the nodes instead of the power flows using (2).

In the following lemma, we provide the key step in computing the disturbance value of a failure analytically and efficiently in Corollary 2.

For convenience in equations, define $\mathbf{B} := [\mathbf{I}-\mathbf{Y}_{k|k}^{1/2}\mathbf{R}_{k|k}\mathbf{Y}_{k|k}^{1/2}]^{-1}$ and $\boldsymbol{\Phi} := \mathbf{Y}_{\bar{k}|\bar{k}}^{1/2}\mathbf{R}_{\bar{k}|k}\mathbf{Y}_{k|k}^{1/2}[\mathbf{I}-\mathbf{Y}_{k|k}^{1/2}\mathbf{R}_{k|k}\mathbf{Y}_{k|k}^{1/2}]^{-1}$. From (4), we know $\mathbf{Y}_{\bar{k}|\bar{k}}^{-1/2}\Delta\vec{f}_{\bar{k}} = \boldsymbol{\Phi}\mathbf{Y}_{k|k}^{-1/2}\vec{f}_k$.

Lemma 3: $\boldsymbol{\Phi}^t\boldsymbol{\Phi} = -\mathbf{I}+\mathbf{B}$.

Proof: To make equations cleaner in the proof, define $\mathbf{H} := \mathbf{Y}^{1/2}\mathbf{R}\mathbf{Y}^{1/2}$. From Lemma 2, $\mathbf{H}^2 = \mathbf{H}$. Hence, if we use block multiplication, then $\mathbf{H}_{k|k} = \mathbf{H}_{k|k}^2 + \mathbf{H}_{k|\bar{k}}\mathbf{H}_{\bar{k}|k}$. Thus,

$$\begin{aligned}\mathbf{H}_{k|k} = \mathbf{H}_{k|k}^2 + \mathbf{H}_{k|\bar{k}}\mathbf{H}_{\bar{k}|k} &\Rightarrow \mathbf{H}_{k|k}[\mathbf{I}-\mathbf{H}_{k|k}] = \mathbf{H}_{k|\bar{k}}\mathbf{H}_{\bar{k}|k}\\ &\Rightarrow \mathbf{H}_{k|k} = \mathbf{H}_{k|\bar{k}}\mathbf{H}_{\bar{k}|k}[\mathbf{I}-\mathbf{H}_{k|k}]^{-1}.\end{aligned}$$

It is easy to see that $\boldsymbol{\Phi} = \mathbf{H}_{\bar{k}|k}[\mathbf{I}-\mathbf{H}_{k|k}]^{-1}$. Hence, using equation above,

$$\begin{aligned}\boldsymbol{\Phi}^t\boldsymbol{\Phi} &= [\mathbf{I}-\mathbf{H}_{k|k}]^{-1}\mathbf{H}_{k|\bar{k}}\mathbf{H}_{\bar{k}|k}[\mathbf{I}-\mathbf{H}_{k|k}]^{-1}\\ &= [\mathbf{I}-\mathbf{H}_{k|k}]^{-1}\mathbf{H}_{k|k}\\ &= [\mathbf{I}-\mathbf{H}_{k|k}]^{-1}\mathbf{H}_{k|k} - [\mathbf{I}-\mathbf{H}_{k|k}]^{-1} + [\mathbf{I}-\mathbf{H}_{k|k}]^{-1}\\ &= -[\mathbf{I}-\mathbf{H}_{k|k}]^{-1}[-\mathbf{H}_{k|k}+\mathbf{I}] + [\mathbf{I}-\mathbf{H}_{k|k}]^{-1}\\ &= -\mathbf{I} + [\mathbf{I}-\mathbf{H}_{k|k}]^{-1} = -\mathbf{I}+\mathbf{B}.\end{aligned}$$

■

Corollary 2: $\Delta\vec{f}_{\bar{k}}^{\,t}\mathbf{Y}_{\bar{k}|\bar{k}}^{-1}\Delta\vec{f}_{\bar{k}} = \vec{f}_k^{\,t}\mathbf{Y}_{k|k}^{-1/2}[-\mathbf{I}+\mathbf{B}]\mathbf{Y}_{k|k}^{-1/2}\vec{f}_k$.

Corollary 2 provides a very important tool for contingency analysis in the power grids. It shows that the disturbance value of a k-line failure can be computed in $O(k^3)$ time which is independent of the size of the network and only depends on the size of the initial failures. Since most of the times in contingency analysis $k << m$, $O(k^3) \approx O(1)$. Corollary 2 can be used for fast ranking of the contingencies based on the disturbance values and pruning most of the cases based on this value. This can significantly reduce the time complexity of the contingency analysis for large k.

Moreover, notice that following Lemma A.1, matrix $\mathbf{B}$ is well defined if, and only if, the k-line failure does not disconnect the grid. Therefore, Corollary 2 can also indicate in $O(1)$ whether a k-line failure disconnects the grid.

We computed the disturbance values for all possible choices of 3-line failures in the IEEE 118-bus and 300-bus systems. Since, the IEEE 118-bus system has 186 lines, it is easy to see that there are $\binom{186}{3} = 1,055,240$ possible choices for the initial set of failures. Using Corollary 2, we could

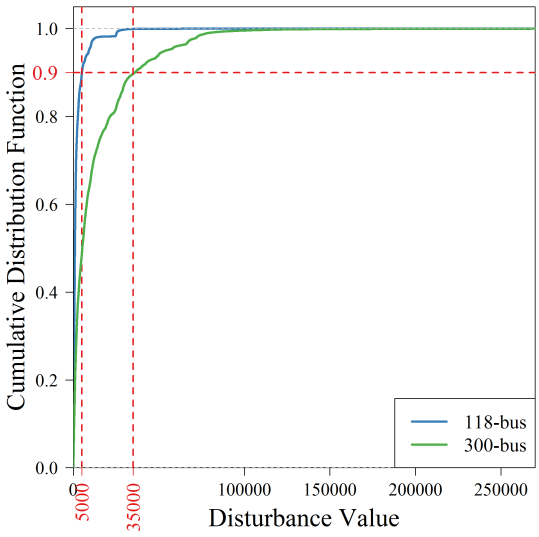


Fig. 1: The cumulative distribution function of the disturbance values for all 3-line failures in the IEEE 118- and 300-bus systems that do not disconnect the grid.

compute the disturbance values for all set of failures (and detect cases that disconnect the grid) in less than a minute. Out of those, $159,591$ of them make the graph disconnected. The cumulative distribution function of the disturbance values for the rest of 895,649 cases are shown in Fig. 1. As can be seen, most of the cases do not result in a high (more than 5,000) disturbance value. Only 10% of the cases have a significant disturbance value. The figure suggests that disturbance values can provide a separation between the failures with higher impact and lower impact.

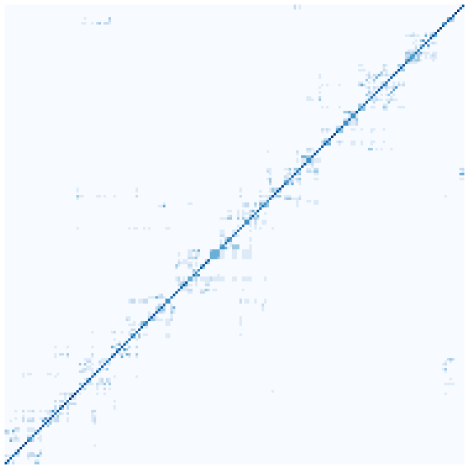
Fig. 1 also shows the cumulative distribution function of the disturbance values for all possible 3-line failures in IEEE 300-bus system that do not disconnect the grid (5,473,725 cases). As can be seen, again most of the cases do not result in a high disturbance value and only less than 10% of the total cases have a significant disturbance value (more than 35,000).[3]

The numerical results support our previous statement that the disturbance values can significantly reduce the total number of cases that are needed to be considered in the contingency analysis of the grids. Moreover, since Corollary 2 provides a very fast way of computing the disturbance values, it can significantly decrease the time complexity of the contingency analysis in power grids.

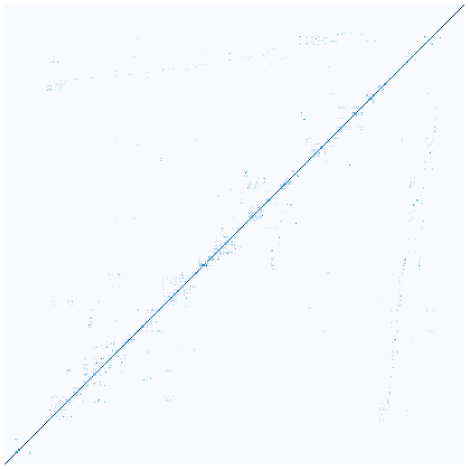
V. Approximating the Disturbance Values

In the previous section, we demonstrated that using Corollary 2, the disturbance value of a k-line failure can be computed efficiently in $O(1)$. However, to analyze all the k-line failure scenarios, one needs to compute the disturbance values for all possible $\binom{m}{k}$ cases which requires $O(m^k)$ time. To alleviate this issue, in this section we provide an approximation for the disturbance values that allows one to focus on the lines with high 1-line disturbance values instead of all the lines.

[3]Notice that the threshold level for "high disturbance value" depends on the level of safety that a system operator likes to maintain. As can be seen in Fig. 1, more than 80% of the cases result in a very small disturbance value. Hence, 1% to 20% of the cases with the highest disturbance values can be selected for deeper analysis depending on the level of safety.



(a) IEEE 118-bus



(b) IEEE 300-bus

Fig. 2: Heatmap for the absolute value of the entries of matrix $\mathbf{R}$ for the IEEE 118- and 300-bus systems. Darker points have higher values.

The idea of our approximation lies in the structure of the matrix $\mathbf{R}$. As shown in Fig. 2, most of the entries of matrix $\mathbf{R}$ in the IEEE 118- and 300-bus systems have very small values compare to its diagonal entries. Hence, we can approximate $\mathbf{R}$ by its diagonal entries as $\mathbf{R} \approx \text{diag}(\mathbf{R}) := \widetilde{\mathbf{R}}$. Using this approximation and Corollary 2, the disturbance values can be computed as:

$$\begin{aligned}
\Delta\vec{f}_{\bar{k}}^{t}\mathbf{Y}_{\bar{k}|\bar{k}}^{-1}\Delta\vec{f}_{\bar{k}} &= \vec{f}_k^{t}\mathbf{Y}_{k|k}^{-1/2}[-\mathbf{I}+[\mathbf{I}-\mathbf{Y}_{k|k}^{1/2}\mathbf{R}_{k|k}\mathbf{Y}_{k|k}^{1/2}]^{-1}]\mathbf{Y}_{k|k}^{-1/2}\vec{f}_k \\
&\approx \vec{f}_k^{t}\mathbf{Y}_{k|k}^{-1/2}[-\mathbf{I}+[\mathbf{I}-\mathbf{Y}_{k|k}^{1/2}\widetilde{\mathbf{R}}_{k|k}\mathbf{Y}_{k|k}^{1/2}]^{-1}]\mathbf{Y}_{k|k}^{-1/2}\vec{f}_k \\
&= \sum_{i=1}^{k} f_i y_{ii}^{-1/2}(-1+\frac{1}{(1-y_{ii}^{1/2}r_{ii}y_{ii}^{1/2})})y_{ii}^{-1/2}f_i \\
&= \sum_{i=1}^{k} f_i^2 \frac{r_{ii}}{(1-y_{ii}r_{ii})}.
\end{aligned}$$

In another words, we approximate the disturbance value of a k-line failure as the sum of the disturbance values of the k single line failures:

$$\delta_k(1,2,\ldots,k) \approx \sum_{i=1}^{k}\delta_1(i). \tag{5}$$

The advantage of this approximation is that to detect k-line failures with high disturbance values, one can focus only on the $\widetilde{m} << m$ lines with the highest 1-line disturbance values. So instead of analyzing $\binom{m}{k}$ cases, one can analyze only $\binom{\widetilde{m}}{k}$ cases.

Another advantage of using the approximation (5) for disturbance values is that it can used to quantify the effect of k-line failures that disconnect the grid as well. Although failures that disconnect the grid may have more devastating effect because of the imbalance they may cause between supply and demand, we believe that this approximation can be used to rank these failures among themselves. Validating the usefulness of approximation (5) for the failures that disconnect the grid is part of the future work.

To quantify the quality of the approximation (5), we define the approximation error percentage (ε):

$$\varepsilon = \frac{\left|\delta_k(i_1, i_2, \ldots, i_k) - \sum_{j=1}^{k} \delta_1(i_j)\right|}{\delta_k(i_1, i_2, \ldots, i_k)} \times 100.$$

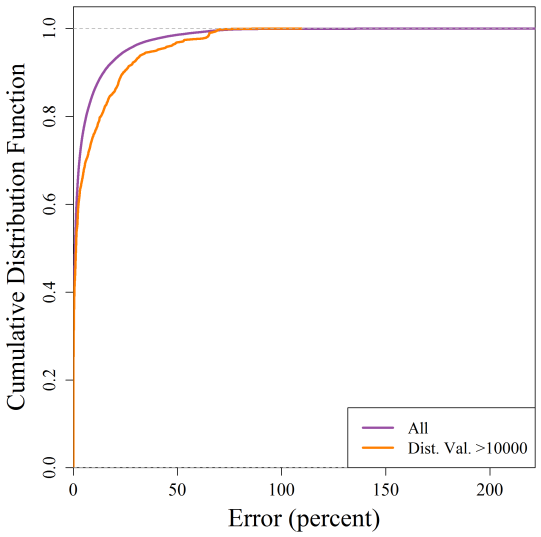
Fig. 3 shows the approximation error in the IEEE 118- and 300-bus systems. As can be seen, in both systems, for most of the cases the error is less than 10%. In particular, the approximation error in the important cases (the ones with high disturbance value) is very low (see orange lines in Fig. 3). Moreover, the correlation between the approximation and actual disturbance values is about 0.98 in both the IEEE 118- and 300-bus systems.

VI. Cascading Failures

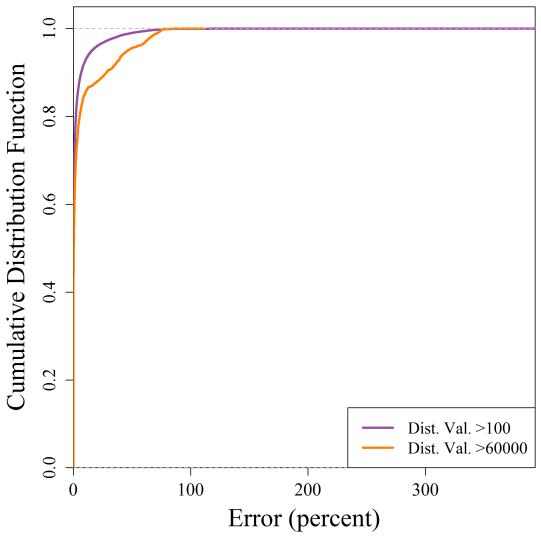
A k-line failure alters the network topology and results in a different flow pattern causing possible consequent line failures. The repetition of this process results in a cascading failure. To show that disturbance values effectively rank k-line failures, we numerically compute the relationship between disturbance value of all 3-line failures and the severity of the initiated cascades in the IEEE 118- and 300-bus systems.

We follow [12], [26] models for cascading failures due to line overloads in power grids with *deterministic outage rule*: namely, a line $\{i, j\}$ fails when the magnitude $|f_{ij}|$ of the flow on that line exceeds its capacity c_{ij}. The line flow capacities are estimated as $c_{ij} = (1 + \alpha) \max\{|f_{ij}|, \overline{f}\}$, where $\overline{f}$ is the *median* of the initial magnitude of line flows and $\alpha = 0.2$ is the lines' *factor of safety*.

When a line fails, it is removed from the network. As a result of this removal, the network topology is changed, and the network can be divided into one or more connected components. We assume that each connected component can operate autonomously. Therefore, within each connected component with non-zero supply and demand, the amounts of the supply and demand are balanced by either scaling down all the supply values (if supply is greater than demand in the connected component) or scaling down all the demand values (if demand is greater than supply in the connected component) within the connected component. If there is either no supply or no demand node within a connected component, all demands or supplies become zero.



(a) IEEE 118-bus



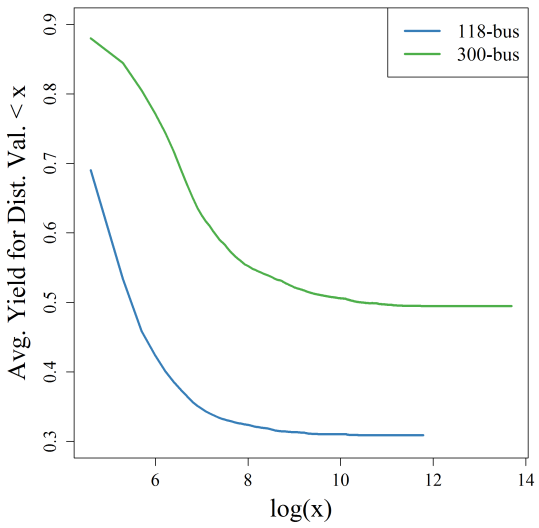
(b) IEEE 300-bus

Fig. 3: Approximation error percentage of the approximate disturbance values for all the 3-line failures than do not disconnect the grid in the IEEE 118- and 300-bus systems.

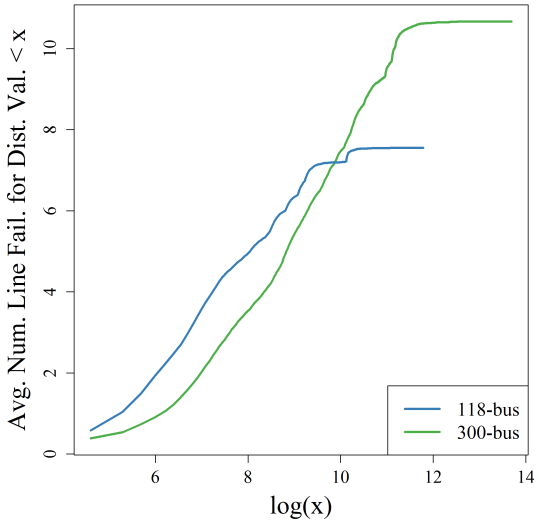
After supply and demand balancing in each component, the power flow equations are solved to compute new flows on the lines. Using the deterministic outage rule, the new set of line failures are then found in all the components, and the cascade continues with the removal of those lines. If there are no overloaded lines in any of the components, the cascade ends.

To measure the severity of a cascade, we compute *yield (the ratio between the demand supplied at the end of a cascade and the original demand) and total number of line failures* at the end of the cascade.

Fig. 4 depicts the relationship between the disturbance values and the average severity of cascades initiated by all the 3-line failures than do not disconnect the grid in the IEEE 118- and 300-bus systems. As can been seen in Fig. 4a, in both systems, the average yield decreases as the disturbance value increases. Moreover, Fig. 4b shows that the average number of line failures at the end of the cascade increases on average,

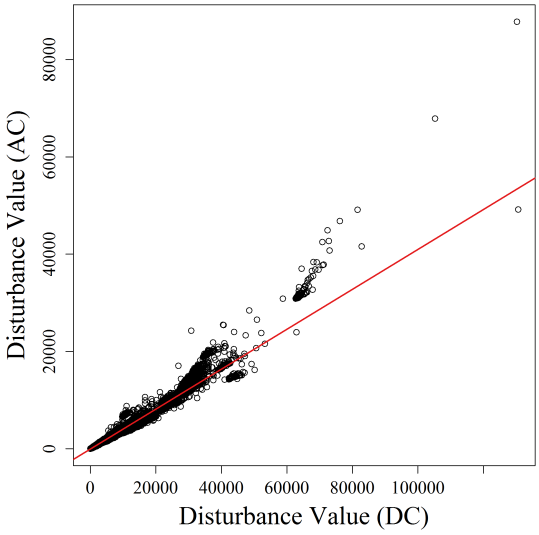


(a) Yield

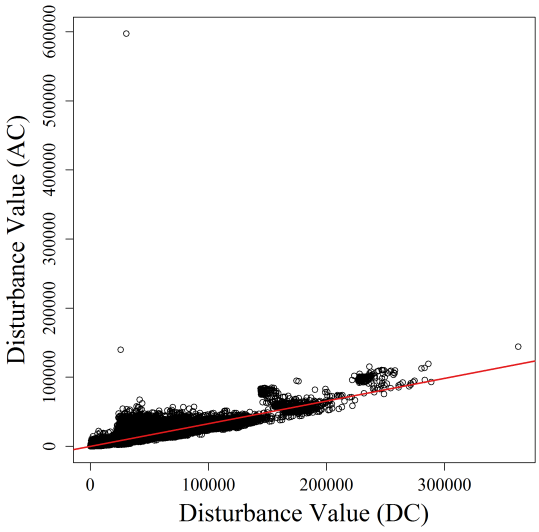


(b) Number of Line Failures

Fig. 4: The relationship between the disturbance values and the average severity of cascades initiated by all the 3-line failures than do not disconnect the grid in the IEEE 118- and 300-bus systems.



(a) IEEE 118-bus



(b) IEEE 300-bus

Fig. 5: Scatter plot of the disturbance values under the AC and DC power flows for all the 3-line failures that do not disconnect the grid in the IEEE 118- and 300-bus systems.

as the disturbance value increases.

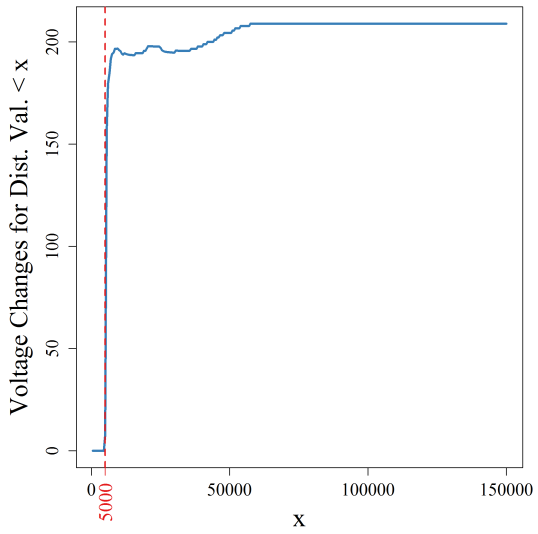
The results in this section suggest that although the disturbance value does not take into account the capacities of the lines, it can predict the average severity of the cascade initiated by a failure.

VII. Disturbance Values under the AC Power Flow Model

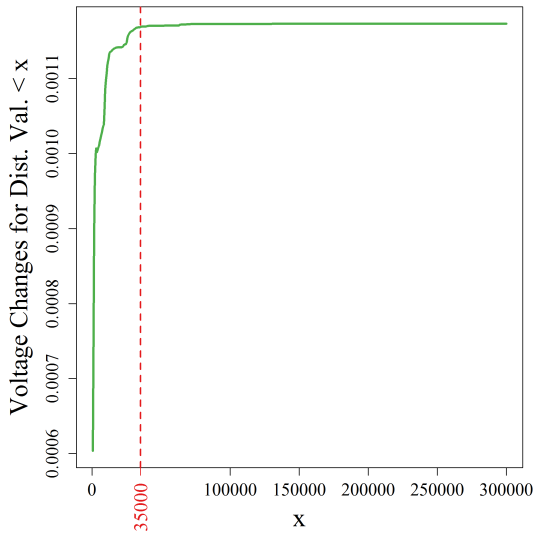
In this section, we numerically show the usefulness of the disturbance values in predicting the changes under the more detailed AC power flows after a k-line failure. The details of the AC power flow equations are provided in Appendix B. To solve the AC power flows, we used the MATPOWER AC power flow solver [27].

We first computed the disturbance values under the AC power flows. To do so, we used the phase angles of the nodes computed before and after a k-line failure under the AC model (recall that the disturbance values can be defined also based on the phase angle). Fig. 5 shows the relationship between the disturbance values computed under DC (using corollary 2) and AC power flows. As can be seen, the disturbance values under the AC and DC are only different in scale. The correlation between these two values is 0.99 and 0.93 in the IEEE 118- and 300-bus systems, respectively. Hence, detecting the important cases based on the disturbance values in the DC is almost similar to the AC model.

Second, we studied the correlation between the voltage changes after a k-line failure and disturbance values. For this purpose, we computed the mean sum squared of the voltage changes after all 3-line failures that do not disconnect the grid in the IEEE 118- and 300-bus systems. As can be seen in Fig. 6, on average, 3-line failures with higher disturbance values cause more voltage changes as well. This demonstrates that the disturbance values which can be computed very efficiently using corollary 2, can also predict the disturbances in the node voltages whose computation is of higher complexity.



(a) IEEE 118-bus



(b) IEEE 300-bus

Fig. 6: Mean sum squared of the voltage changes for the 3-line failures that do not disconnect the grid with the disturbance value less than a certain value in the IEEE 118- and 300-bus systems.

Finally, since the AC power flow equations are nonlinear (see Appendix B), there are 3-line failures after which the AC power solver does not converge to a solution. The CDF of the disturbance values for these 3-line failures is depicted in Fig. 7. As can be seen, for all such failures in the IEEE 118-bus system the disturbance values are above 5000, meaning that they are between the top 10% of all the 3-line failures in terms of the disturbance values (see Fig. 1).

The same is true in the IEEE 300-bus system. Most of the 3-line failures that result in the divergence of the AC power flow solver, are between the top 10-20% cases in terms of the disturbance values.

Overall, the results in this section demonstrate that the disturbance values can be useful for predicting the changes under the AC power flow model. Since following corollary 2 the disturbance values can be computed much faster compared to the computation of steady state changes under the AC power flows (which requires solving a nonlinear set of equations several times), this significantly reduces the computational complexity associated with the contingency analysis in the grid.

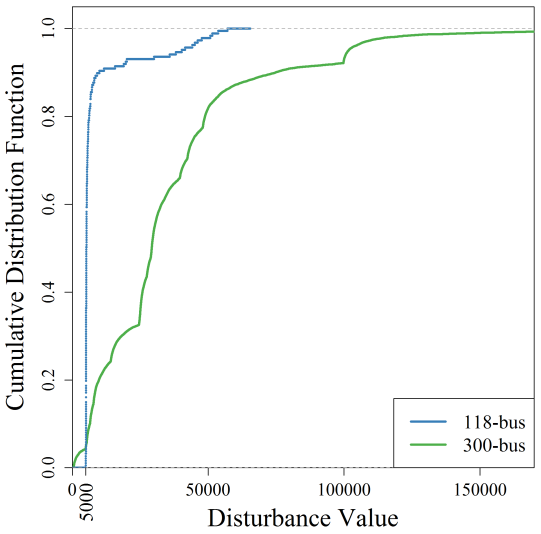


Fig. 7: The CDF of the disturbance values for the 3-line failures that do not disconnect the grid but produce enough instability such that the AC power solver does not converge to a solution.

VIII. Conclusion

The results in this paper provide efficient tools for quantifying the effect of k-line failures. The most unique aspect of our approach is the use of the matrix of equivalent reactance values to efficiently capture the effect of k-line failures. We defined the disturbance value of a failure and show that this metric can be computed for any set of failures in $O(1)$. Moreover, based on the approximation of the disturbance values, the total number of cases that need to be considered can be significantly reduced as well. Our numerical results showed that disturbance values provide a clear separation between the failures with higher impact and lower impact under both the AC and DC power flows.

Despite providing a fast and useful measure for quantifying the severity of a k-line failure, the disturbance value does not capture the thermal capacity of the lines as well as further instabilities caused by separation of the grid into several islands. Hence, the disturbance values alone cannot be used to detect important contingencies. Since thorough contingency analysis for all high order scenarios is intractable, however, disturbance values can be used to reduce the total number of cases needed to be analyzed in contingency analysis and can significantly reduce the computational complexity associated with this analysis. In our future work, we plan to incorporate line capacities in definition of the disturbance value and also study the correlation between disturbance values and other important dynamics of the power grids (e.g., transient stability issues).

Acknowledgement

This work was supported in part by DARPA RADICS under contract #FA-8750-16-C-0054, funding from the U.S. DOE OE as part of the DOE Grid Modernization Initiative, and

DTRA grant HDTRA1-13-1-0021. The work of G. Z. was also supported in part by the Blavatnik ICRC.

REFERENCES

[1] S. Soltan and G. Zussman, "Quantifying the effect of k-line failures in power grids," in *Proc. IEEE PES'16*, July 2016.

[2] A. Wood and B. Wollenberg, "Power generation, operation, and control," 1983.

[3] Project Group Turkey, "Report on blackout in Turkey on 31st March 2015," Sept. 2015, avilable at: https://www.entsoe.eu/Documents/SOC%20documents/Regional_Groups_Continental_Europe/20150921_Black_Out_Report_v10_w.pdf.

[4] M. K. Enns, J. J. Quada, and B. Sackett, "Fast linear contingency analysis," *IEEE Trans. Power App. Syst*, no. 4, pp. 783–791, 1982.

[5] C. Davis and T. Overbye, "Linear analysis of multiple outage interaction," in *Proc. IEEE HICSS'09*, 2009.

[6] D. Bienstock and A. Verma, "The $N-k$ problem in power grids: New models, formulations, and numerical experiments," *SIAM J. Optimiz.*, vol. 20, no. 5, pp. 2352–2380, 2010.

[7] M. Chertkov, F. Pan, and M. G. Stepanov, "Predicting failures in power grids: The case of static overloads," *IEEE Trans. Smart Grid*, vol. 2, no. 1, pp. 162–172, 2011.

[8] C. Davis and T. J. Overbye, "Multiple element contingency screening," *IEEE Trans. Power Syst.*, vol. 26, no. 3, pp. 1294–1301, 2011.

[9] M. J. Eppstein and P. D. Hines, "A "random chemistry" algorithm for identifying collections of multiple contingencies that initiate cascading failure," *IEEE Trans. Power Syst.*, vol. 27, no. 3, pp. 1698–1705, 2012.

[10] K. Turitsyn and P. Kaplunovich, "Fast algorithm for N-2 contingency problem," in *Proc. IEEE HICSS'13*, 2013.

[11] C. Lai and S. H. Low, "The redistribution of power flow in cascading failures," in *Proc. Allerton'13*, 2013.

[12] S. Soltan, D. Mazauric, and G. Zussman, "Analysis of failures in power grids," *to appear in IEEE Trans. Control Netw. Syst. (available on IEEE Xplore Digital Library)*, 2017.

[13] J. D. Glover, M. Sarma, and T. Overbye, *Power System Analysis & Design, 4th Edition*. Cengage Learning, 2011.

[14] T. Güler and G. Gross, "Detection of island formation and identification of causal factors under multiple line outages," *IEEE Trans. Power Syst.*, vol. 22, no. 2, pp. 505–513, 2007.

[15] R. Bapat, *Graphs and matrices*. Springer, 2010.

[16] Z. Wang, A. Scaglione, and R. J. Thomas, "Electrical centrality measures for electric power grid vulnerability analysis," in *in Proc. IEEE CDC'10*, 2010.

[17] P. Kaplunovich and K. Turitsyn, "Fast and reliable screening of N-2 contingencies," *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 4243–4252, 2016.

[18] P. Kaplunovich and K. S. Turitsyn, "Statistical properties and classification of N-2 contingencies in large scale power grids," in *in Proc. HICSS'14*. IEEE, 2014, pp. 2517–2526.

[19] V. Donde, V. López, B. Lesieutre, A. Pinar, C. Yang, and J. Meza, "Severe multiple contingency screening in electric power systems," *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 406–417, 2008.

[20] P. D. Hines, I. Dobson, E. Cotilla-Sanchez, and M. Eppstein, ""dual graph" and "random chemistry" methods for cascading failure analysis," in *Proc. IEEE HICSS'13*, 2013.

[21] Q. Chen and J. D. McCalley, "Identifying high risk N-k contingencies for online security assessment," *IEEE Trans. on Power Syst.*, vol. 20, no. 2, pp. 823–834, 2005.

[22] S. Poudel, Z. Ni, and W. Sun, "Electrical distance approach for searching vulnerable branches during contingencies," *IEEE Trans. Smart Grid*, 2016.

[23] A. R. Bergen and V. Vittal, *Power Systems Analysis*. Prentice-Hall, 1999.

[24] T. J. Overbye, X. Cheng, and Y. Sun, "A comparison of the AC and DC power flow models for LMP calculations," in *Proc. IEEE HICSS'04*, 2004.

[25] A. Albert, *Regression and the Moore-Penrose pseudoinverse*. Academic Press, 1972, vol. 3.

[26] A. Bernstein, D. Bienstock, D. Hay, M. Uzunoglu, and G. Zussman, "Power grid vulnerability to geographically correlated failures - analysis and control implications," in *Proc. IEEE INFOCOM'14*, Apr. 2014.

[27] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, 2011.

[28] J. D. Glover, M. S. Sarma, and T. Overbye, *Power System Analysis & Design, SI Version*. Cengage Learning, 2012.

APPENDIX A

The following lemma is the generalization of [12, Lemma 1] and similar to the idea used in [14, Theorem 2] to detect the connected components after multiple line failures. Our proof is different from the proof of a similar Theorem in [14, Theorem 2].

Lemma A.1: Matrix $\mathbf{I}-\mathbf{Y}_{k|k}^{1/2}\mathbf{D}_k^t\mathbf{A}^+\mathbf{D}_k\mathbf{Y}_{k|k}^{1/2}$ is invertible if, and only if G' is connected.

Proof: First, it is easy to see that $\mathbf{I}-\mathbf{Y}_{k|k}^{1/2}\mathbf{D}_k^t\mathbf{A}^+\mathbf{D}_k\mathbf{Y}_{k|k}^{1/2}$ is invertible if, and only if $\mathbf{I}-\mathbf{Y}_{k|k}\mathbf{D}_k^t\mathbf{A}^+\mathbf{D}_k$ is invertible. Now assume, G' is disconnected. Without loss of generality, assume $C=\{e_1,e_2,\ldots,e_r\}$ is a minimal subset of $\{e_1,e_2,\ldots,e_k\}$ such that $G\backslash C$ is disconnected. Since, C is a minimal subset, $G\backslash C$ has only two connected components G_1 and G_2 and each $e_i\in C$ has one end in G_1 and the other end in G_2. Again with out loss of generality, assume that all the edges in C are directed from G_1 to G_2. We prove that vector $\vec{v}\in\mathbb{R}^k$ defined as $v_i=y_{ii}$ for $i\leq r$ and $v_i=0$ for $i>r$ is an eigenvector of $\mathbf{Y}_{k|k}\mathbf{D}_k^t\mathbf{A}^+\mathbf{D}_k$ associated with the eigenvalue 1. Notice, that if $\vec{p}=\mathbf{D}_k\vec{v}$, then $\theta_i=1$ for $i\in G_1$ and $\theta_i=0$ for $i\in G_2$ gives a solution to DC power flow problem in G. It is easy to see that in this setting $\vec{f_k}=\vec{v}$. On the other hand, $\vec{f_k}=\mathbf{Y}_{k|k}\mathbf{D}_k^t\mathbf{A}^+\vec{p}$, and since $\vec{f_k}=\vec{v}$ and $\vec{p}=\mathbf{D}_k\vec{v}$, therefore $\mathbf{Y}_{k|k}\mathbf{D}_k^t\mathbf{A}^+\mathbf{D}_k\vec{v}=\vec{v}$. Hence, $\mathbf{Y}_{k|k}\mathbf{D}_k^t\mathbf{A}^+\mathbf{D}_k$ has eigenvalue 1 and $\mathbf{I}-\mathbf{Y}_{k|k}\mathbf{D}_k^t\mathbf{A}^+\mathbf{D}_k$ is not invertible.

Now assume, $\mathbf{I}-\mathbf{Y}_{k|k}\mathbf{D}_k^t\mathbf{A}^+\mathbf{D}_k$ is not invertible. Then, $\mathbf{I}-\mathbf{Y}_{k|k}\mathbf{D}_k^t\mathbf{A}^+\mathbf{D}_k$ has an eigenvalue 0 and $\mathbf{Y}_{k|k}\mathbf{D}_k^t\mathbf{A}^+\mathbf{D}_k$ has an eigenvalue 1. Assume $\vec{v}$ is the eigenvector associated with the eigenvalue 1 of $\mathbf{Y}_{k|k}\mathbf{D}_k^t\mathbf{A}^+\mathbf{D}_k$. It is again easy to see that if $\vec{p}=\mathbf{D}_k\vec{v}$, then $\vec{f_k}=\vec{v}$ is the solution to the power flow problem in G. From the flow conservation equations, it is also easy to verify that $\vec{f_{\bar{k}}}=0$. Now, by contradiction assume G' is connected. Then, there should be a path in G' from a node i to node j such that $\theta_i\neq\theta_j$. Therefore, there should be an edge $e=(w,z)$ in this path such that $\theta_w\neq\theta_z$ and thus $f_e\neq 0$. However, since $e\in G'$ and $\vec{f_{\bar{k}}}=0$ we know that $f_e=0$ which is a contradiction. Therefore, G' is not connected.

From the proof it is easy to see that if $\vec{v}$ is an eigenvector associated with the eigenvalue 1 of $\mathbf{Y}_{k|k}\mathbf{D}_k^t\mathbf{A}^+\mathbf{D}_k$, then nodes with the same phase angle values in the solution of the power flow problem in G with $\vec{p}=\mathbf{D}_k\vec{v}$ form a connected component in G'. ■

APPENDIX B
AC POWER FLOWS

In the AC power flows, the status of each node i is represented by its voltage $v_i=|v_i|e^{\mathbf{i}\theta_i}$ in which $|v_i|$ is the voltage magnitude, θ_i is the phase angle at node i, and $\mathbf{i}$ denotes the imaginary unit.

The nonlinear and nonconvex AC power flow problem is the problem of computing the voltage magnitudes and phase angles at each bus in steady-state conditions [28]. In the

steady-state, the injected apparent power s_i at node i equals to

$$s_i = \sum_{\substack{k=1 \\ k \neq i}}^{n} s_{ik} = \sum_{\substack{k=1 \\ k \neq i}}^{n} v_i (g_{ik} + \mathbf{i} b_{ik})^* (v_i^* - v_k^*)$$

where * denotes the complex conjugation, and $g_{ik} + \mathbf{i} b_{ik}$ is the equivalent admittance of the lines from node i to k.

Using the definition of the apparent power $s_{ik} = p_{ik} + \mathbf{i} q_{ik}$ leads to the equations for the active power p_i and the reactive power q_i at each node i:

$$\begin{aligned} p_i &= -\textstyle\sum_{k=1}^{n} |v_i||v_k|(g_{ik} \cos\theta_{ik} + b_{ik} \sin\theta_{ik}) \\ q_i &= -\textstyle\sum_{k=1}^{n} |v_i||v_k|(g_{ik} \sin\theta_{ik} - b_{ik} \cos\theta_{ik}) \end{aligned}$$

where $\theta_{ik} = \theta_i - \theta_k$, and $g_{ii} + \mathbf{i} b_{ii} = -\sum_{j \in N(i)} (g_{ij} + \mathbf{i} b_{ij})$.

In the AC power flow problem, each node i is categorized into one of the following three types:

1) *Slack node*: The node for which the voltage is typically 1.0. For convenience, it is indexed as node 1. The active power p_1 and the reactive power q_1 need to be computed.
2) *Load node*: The active power p_i and the reactive power q_i at these nodes are known and the voltage v_i needs to be computed.
3) *Voltage controlled node*: The active power p_i and the voltage magnitude $|v_i|$ at these nodes are known and the reactive power q_i and the phase angle θ_i need to be computed.

Saleh Soltan is a Ph.D. candidate in the department of Electrical Engineering at Columbia University. He received B.S. degrees in Electrical Engineering and Mathematics (double major) from Sharif University of Technology, Iran in 2011 and the M.S. degree in Electrical Engineering from Columbia University in 2012. He is the Gold Medalist of the 23rd National Mathematics Olympiad in Iran (2005) and the recipient of Columbia University Electrical Engineering Armstrong Memorial Award (2012).

Alexander Loh is currently a PhD student in the department of Electrical Engineering at Columbia University. He received a B.S. in Electrical Engineering from Rutgers University in 2015, and an M.S. in Electrical Engineering from Columbia University in 2016. His research interests include Algorithms, Stochastic Modeling, Network Science, Computer Networks, and Power Systems Analysis. He is the recipient of the Columbia University Presidential Fellowship (2015), NSF IGERT Fellowship (2015), and the NSF GRFP Honorable Mention (2017).

Gil Zussman received the Ph.D. degree in Electrical Engineering from the Technion in 2004 and was a postdoctoral associate at MIT in 2004-2007. He is currently an Associate Professor of Electrical Engineering at Columbia University. He is a co-recipient of 7 paper awards including the ACM SIGMETRICS'06 Best Paper Award, the 2011 IEEE Communications Society Award for Advances in Communication, and the ACM CoNEXT'16 Best Paper Award. He received the Fulbright Fellowship, the DTRA Young Investigator Award, and the NSF CAREER Award, and was a member of a team that won first place in the 2009 Vodafone Foundation Wireless Innovation Project competition.