Localization & Mitigation of Cascading Failures in Power Systems, Part I: Spectral Representation & Tree Partition

Linqi Guo, Chen Liang, Alessandro Zocca, Steven H. Low, and Adam Wierman

***Abstract*— Cascading failures in power systems propagate non-locally, making the control of outages extremely difficult. In this work, we propose a new framework that offers strong analytical guarantees on both the localization and mitigation of cascading failures in power systems. The key component of this framework leverages the concept of *tree partition*, which characterizes regions of a power network inside which line failures are automatically localized. In Part I of this paper we establish a mathematical theory that underlies all the performance guarantees of tree partition as well as its failure localization properties. This theory consists of a set of tools developed using the Laplacian matrix of the transmission network and reveals a novel perspective that precisely captures the Kirchhoff's Law in terms of topological structures. Our results show that the distribution of different families of subtrees of the transmission network plays a critical role on the patterns of power redistribution, and motivates tree partitioning of the network as a strategy to eliminate long-distance propagation of disturbances. These results are used in Parts II and III of this paper to design strategies to localize and mitigate line failures.**

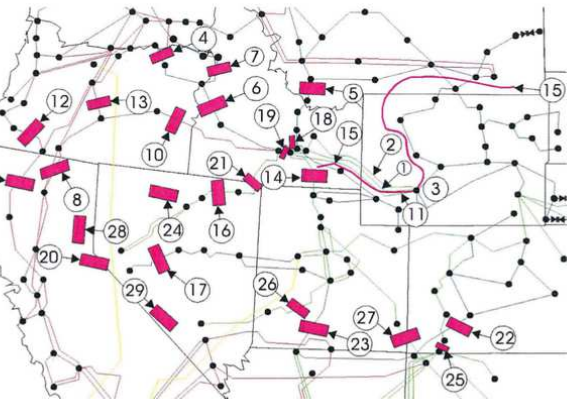

Fig. 1: The sequence of events, indexed by the circled numbers, that lead to the Western US blackout in 1996 from [1], [2].

I. Introduction

Cascading failures in power systems propagate non-locally, making their analysis and mitigation difficult. This fact is illustrated by the sequence of events leading to the 1996 Western US blackout (as summarized in Fig. 1 from [1], [2]), in which successive failures happened hundreds of kilometers away from each other (e.g. from stage ③ to stage ④ and from stage ⑦ to stage ⑧). Non-local propagation makes it challenging to design distributed controllers that reliably prevent and mitigate cascades in power systems. In fact, such control is often considered impossible, even when centralized coordination is available [3], [4].

Current industry practice for mitigating cascading failures mostly relies on simulation-based contingency analysis, which focuses on a small set of most likely initial failures [5]. Moreover, the size of the contingency set which is tested (and thus the level of security guarantee) is often constrained by computational power, undermining its effectiveness in view of the enormous number of components in power networks. After a blackout event, a detailed study typically leads to a redesign of such contingency sets, potentially together with physical network upgrades and revision of system management policies and regulations [4].

The limitations of the current practice have motivated a large body of research on cascading failure; see e.g. [6] for a recent review with extensive references. In particular, the literature on analytical properties of cascading failure can be roughly categorized as follows: (a) applying Monte-Carlo methods to analytical models that account for the steady state power redistribution using DC [7]–[10] or AC [11]–[13] power flow models; (b) studying pure topological models built upon simplifying assumptions on the propagation dynamics (e.g., failures propagate to adjacent lines with high probability) and inferring component failure propagation patterns from graph-theoretic properties [14]–[16]; (c) investigating simplified or statistical cascading failure dynamics [2], [17]–[19].

In all these approaches, it is often difficult to make general inferences about failure patterns. For example, power flow over a specific transmission line can increase, decrease and even reverse direction as cascading failure unfolds [20]. The failure of a line can cause another line that is arbitrarily far away to trip [21]. Load shedding instead of mitigating a cascading failure, can actually increase the congestion on certain lines [22]. This lack of structural properties is a key challenge in the modeling, control, and mitigation of cascading failures in power systems.

In this work, we take a different approach that leverages the spectral representation of transmission network topology to establish several structural properties. The spectral view is powerful as it reveals surprisingly simple characterizations of complex system behaviors, e.g., on controllability and observability of power system dynamics [23], on dynamic properties of frequency control [24], [25], and on monotonicity properties and power flow redistribution [26]. In the context of cascading failures, our spectral approach [26] motivates the use of the network *tree partition* to eliminate long-distance propagation of disturbances, and allows us to develop a new control framework that offers strong analytical guarantees in both the mitigation and localization of cascading failures.

This work has been supported by Resnick Fellowship, Linde Institute Research Award, NWO Rubicon grant 680.50.1529, NSF through grants CCF 1637598, ECCS 1619352, ECCS 1931662, CNS 1545096, CNS 1518941, CPS ECCS 1739355, CPS 154471.

LG, CL, SHL, AW are with the Department of Computing and Mathematical Sciences, California Institute of Technology, Pasadena, CA, 91125, USA. Email: {lguo, cliang2, slow, adamw}@caltech.edu. AZ is with the Department of Mathematics of the Vrije Universiteit Amsterdam, 1081HV, The Netherlands. Email: a.zocca@vu.nl.

Contributions of Part I of this paper: *We establish a mathematical theory that underlies the performance guarantees of the tree partition of a transmission network and the failure localization properties that tree partitions bring to the proposed control framework.* This theory starts from the Laplacian matrix of a transmission network and unveils a close connection between power redistribution patterns and the distribution of different families of (sub)trees of the power network topology. This connection uncovers new topological structures of several important and well-studied quantities in power system contingency analysis, such as the generation shift sensitivity factor and the line outage redistribution factor. Further, in contrast to pure graphical models such as those in [14], [15], [27], such topological interpretations capture Kirchhoff's Law in a precise way and do not rely on any assumptions or simplifications on the failure propagation dynamics.

The key result that relates the power redistribution to graphical structures is given in Proposition 4, which states that the distribution of different families of subtrees of the transmission network fully determines the system state under a given set of injections. In Section III, we establish a new set of graphical representations of generation shift sensitivity factors and line outage distribution factors in contingency analysis. This novel graph-theoretical viewpoint enables us to derive precise algebraic properties of power redistribution using purely graphical arguments, and shows that disturbances propagate through "subtrees" in a power network. Using this framework, in Section III-E we derive the *Simple Loop Criterion* that fully determines whether the failure of one line can impact another line in a given network topology.

In order to prove these results, we exploit the celebrated Kirchhoff Matrix Tree Theorem as well as its generalization known as the All Minors Matrix Tree Theorem. Further, we make use of some novel properties of the Laplacian matrix derived in the context of DC power flow and characterize certain quantities of interest by different families of spanning forests in the transmission network.

Our graphical interpretation of power redistribution naturally suggests that we can eliminate long-distance propagation of system disturbances by refining the (possibly trivial) tree partition of a transmission network. In Section IV, we formally define the notion tree partition and show that the "finest" tree partition of a general graph is unique and can be computed in linear time. In Section V, we demonstrate how tree partitioning localizes the impacts of a line failure to the region of the network where the failure happens. The rigorous proof of such localization properties and how they can be leveraged to provide analytical guarantees for failure mitigation are presented in Part II and Part III of this paper, respectively.

II. Preliminaries

In this section, we present the DC power flow model, explain its network Laplacian matrix, and introduce some of its basic properties that will be used throughout this paper.

A. Power network and DC power flow

We describe a transmission network using a graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, whose node set $\mathcal{N} = \{1, \dots, n\}$ models the $n = |\mathcal{N}|$ buses and whose edge set $\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$ models the $m = |\mathcal{E}|$ transmission lines. We without loss of generality assume the graph is simple and use the terms bus/node and line/edge are used interchangeably. An edge in $\mathcal{E}$ between node i and j is denoted either as e or (i, j). We assign an arbitrary orientation over $\mathcal{E}$ so that if $(i, j) \in \mathcal{E}$ then $(j, i) \notin \mathcal{E}$. The line susceptance (weighted by nodal voltage magnitudes [28]) of line $e = (i, j)$ is denoted as $B_e = B_{ij}$ and the susceptance matrix is the $m \times m$ diagonal matrix $B := \operatorname{diag}(B_e : e \in \mathcal{E})$. Let f be the m-dimensional vector consisting of all branch flows, with f_e denoting the branch flow on line e. The node-edge incidence matrix of $\mathcal{G}$ is the $n \times m$ matrix C defined as

$$C_{ie} = \begin{cases} 1 & \text{if node } i \text{ is the source of } e, \\ -1 & \text{if node } i \text{ is the target of } e, \\ 0 & \text{otherwise.} \end{cases}$$

We introduce the n-dimensional vectors p and θ, where p_i and θ_i are the power injection and voltage phase angle at bus i, respectively. With the above notation, the DC power flow model can be written as

$$p = Cf \tag{1a}$$
$$f = BC^T\theta, \tag{1b}$$

where (1a) is the flow conservation (Kirchhoff's) law and (1b) is the Ohm's laws. Given an injection vector p that is balanced over the network $\sum_{j \in \mathcal{N}} p_j = 0$, the DC model (1) has a solution (θ, f) where θ is unique up to an arbitrary reference angle. Without loss of generality, we choose node n as a reference node and set $\theta_n = 0$. With this convention, the solution (θ, f) is unique.

When a line e is tripped, the power redistributes according to the DC model (1) on the newly formed graph $\mathcal{G}' := (\mathcal{N}, \mathcal{E} \setminus \{e\})$. If $\mathcal{G}'$ is still connected, then the flow change on a line $\hat{e}$ with the same power injection p can be given in terms of the *line outage distribution factor* $K_{\hat{e}e}$ from e to $\hat{e}$ as

$$\Delta f_{\hat{e}} = K_{\hat{e}e} f_e.$$

It is known that these distribution factors depend solely on the network topology, in the sense that they are independent of the power injections p and can be computed directly from the matrices B and C [29]. In Section III-D, we derive a new formula for $K_{\hat{e}e}$ that precisely captures how the Kirchhoff's law redistributes branch flows along subtrees of $\mathcal{G}$ and that is the foundation for our results on the localization properties of network tree partitions.

B. Laplacian matrix

The DC power flow equations (1) imply that

$$p = CBC^T\theta.$$

In other words, the phase angles θ are determined from the power injection p via the matrix CBC^T. This matrix, which we denote as $L = CBC^T$, is known as the Laplacian matrix of $\mathcal{G}$ and plays a central role in our analysis. Since for any $v \in \mathbb{R}^n$ we have that

$$v^T L v = \sum_{(i,j) \in \mathcal{E}} B_{ij} (v_i - v_j)^2 \geq 0, \tag{2}$$

the matrix L is positive semidefinite. Moreover, equality is attained in (2) if and only if $v_i = v_j$ for every $(i, j) \in \mathcal{E}$.

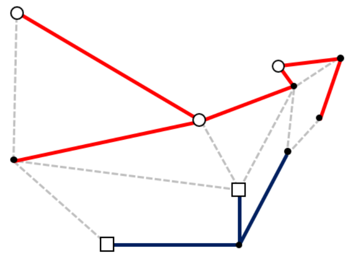

Fig. 2: An example element in $\mathcal{T}(\mathcal{N}_1, \mathcal{N}_2)$, where circles correspond to elements in $\mathcal{N}_1$ and squares correspond to elements in $\mathcal{N}_2$. The two trees containing $\mathcal{N}_1$ and $\mathcal{N}_2$ are highlighted as solid lines.

Thus the Laplacian matrix L is singular and the eigenspace of L corresponding to 0, i.e., the kernel of L, consists of vectors that take the same value on each of the connected components of $\mathcal{G}$. The following well-known result summarizes a special case.

Lemma 1. *If the graph $\mathcal{G}$ is connected, then the kernel of L is* $\operatorname{span}(\mathbf{1})$*, the set of vectors with uniform entries.*

This result implies that the Laplacian matrix L of a connected graph has rank $n-1$ and, hence, contains an invertible submatrix of size $(n-1) \times (n-1)$. Consider the submatrix $\overline{L}$ obtained from L deleting the last row and the last column (i.e., those corresponding to the reference node n). The Kirchhoff's Matrix Tree Theorem (stated below, see [30] for more details) relates the determinant of $\overline{L}$ to a weighted sum of spanning trees of $\mathcal{G}$.

Proposition 2. *If the graph $\mathcal{G}$ is connected, the determinant of $\overline{L}$ is given by*

$$\det(\overline{L}) = \sum_{E \in \mathcal{T}_{\mathcal{E}}} \prod_{e \in E} B_e,$$

where $\mathcal{T}_{\mathcal{E}}$ is the set of all spanning trees of $\mathcal{G}$.

This result implies, in particular, that $\overline{L}$ is invertible since the set of spanning trees $\mathcal{T}_{\mathcal{E}}$ is non-empty for a connected graph and $B_e > 0$ for all e. Define the $n \times n$ matrix A as follows:

$$A = \begin{bmatrix} (\overline{L})^{-1} & \mathbf{0} \\ \mathbf{0} & 0 \end{bmatrix}. \tag{3}$$

The entries of A encode information on the topology of $\mathcal{G}$ and have a clear graph-theoretical interpretation. More specifically, as we show in Section III, all entries of A are closely related to the subtree distributions of $\mathcal{G}$ and suggest how the DC power flow equations (1) inherit features from the structure of the graph $\mathcal{G}$.

III. Power Redistribution and Trees

In this section, we show how the entries of the matrix A are related to specific families of subtrees of the power network $\mathcal{G}$. This relation reveals new perspectives on many important and well-studied quantities in contingency analysis and it is at the core of a new criterion, we call the Simple Loop Criterion, that characterizes whether the failure of a line impacts other lines. This new representation in terms of network subtrees is extensively used to develop further results in Parts II and III of this paper.

A. Spectral representation

Proposition 2 relates $\det(\overline{L})$ to the spanning trees of $\mathcal{G}$ and hints at the fact that the subtree families of $\mathcal{G}$ determine certain algebraic properties of the matrix A. We now present a finer-grained result that explicitly characterizes each entry of A using the tree structure of $\mathcal{G}$.

We first introduce some more notation. Given a subset $E \subseteq \mathcal{E}$ of edges, we denote by $\mathcal{T}_E$ the set of spanning trees of $\mathcal{G}$ with edges from E (which can possibly be empty if E is too small or if $\mathcal{G}$ is disconnected). In particular, $\mathcal{T}_{\mathcal{E}}$ is the set of spanning trees on $\mathcal{G}$. For any pair of subsets $\mathcal{N}_1, \mathcal{N}_2 \subset \mathcal{N}$, we define $\mathcal{T}(\mathcal{N}_1, \mathcal{N}_2)$ to be the set of spanning forests of $\mathcal{G}$ consisting of exactly two trees (necessarily disjoint) that contain $\mathcal{N}_1$ and $\mathcal{N}_2$, respectively (see Fig. 2 for an illustration of $\mathcal{T}(\mathcal{N}_1, \mathcal{N}_2)$). Given a subset $E \subseteq \mathcal{E}$ of edges, we write

$$\chi(E) := \prod_{e \in E} B_e.$$

Then, the All Minors Matrix Tree Theorem [30] applied to the matrix $\overline{L}$ yields the following result.

Proposition 3. *The determinant of the matrix $\overline{L}^{ij}$, obtained from $\overline{L}$ by deleting the i-th row and j-th column, is equal to*

$$\det\left(\overline{L}^{ij}\right) = (-1)^{i+j} \sum_{E \in \mathcal{T}(\{i,j\},\{n\})} \chi(E).$$

This result leads to our graphical interpretation of the elements of A, formally stated as follows (its proof is presented in Appendix I):

Proposition 4 (Spectral Representation)**.** *For any $i, j \in \mathcal{N}$, we have*

$$A_{ij} = \frac{\sum_{E \in \mathcal{T}(\{i,j\},\{n\})} \chi(E)}{\sum_{E \in \mathcal{T}_{\mathcal{E}}} \chi(E)}. \tag{4}$$

In the expression (4) for A_{ij}, the denominator is a normalization constant common for all entries of A. The sum at the numerator is over the trees in $\mathcal{T}(\{i,j\},\{n\})$, which means that A_{ij} is proportional to the (weighted) number of trees that connect i to j and can be interpreted as the "connection strength" between the buses i and j in $\mathcal{G}$.

The matrix A fully determines the phase angles θ (and thus the branch flows f) from the power injections p as prescribed by (1), more specifically

$$\theta = Ap.$$

In view of this fact, Proposition 4 implies that the power redistribution under DC power flow model (1) can be described using the distribution of subtree families over the transmission network. In particular, we can deduce analytical properties of the branch flows using purely graphical structures, as we do in the following corollary.

Corollary 5. *For all $i, j \in \mathcal{N}$ we have*

$$A_{ij} \geq 0,$$

and equality holds if and only if every path from i to j contains the reference node n.

Proof. The statement trivially holds when i or j coincides with n, so in the rest of the proof we focus on the case $i, j \neq n$. In this case, since $\chi(E) \geq 0$ for all E, we clearly have $A_{ij} \geq 0$ and equality holds if and only if the set $\mathcal{T}(\{i,j\},\{n\})$ is empty.

If every path from i to j contains n, then since any tree containing $\{i,j\}$ induces a path from i to j, we know that this tree also contains n. As a result, $\mathcal{T}(\{i,j\},\{n\}) = \emptyset$. Conversely, if $\mathcal{T}(\{i,j\},\{n\}) = \emptyset$, then any path from i to j must contain n, since for any path from i to j not passing through n, we can iteratively add edges that do not have n as an endpoint to obtain a spanning tree over the nodes set $\mathcal{N}\setminus\{n\}$. This tree together with the node n itself is an element of $\mathcal{T}(\{i,j\},\{n\})$. We thus have shown that $\mathcal{T}(\{i,j\},\{n\})$ is empty if and only if every path from i to j contains n. □

In the following subsections, by making use of Proposition 4, we derive novel expressions for well-studied quantities in power system analysis, unveiling a new graphical meaning for each of them. These results allow us to derive the Simple Loop Criterion, which is the core motivation behind using the *tree partitioning* of power networks for failure localization.

B. Generation Shift Sensitivity Factor

Consider a pair of buses i and j, which may or may not be adjacent, i.e., (i,j) may or may not be an edge in $\mathcal{E}$. Suppose the injection at bus i is increased by Δ_{ij}, the injection at bus j is reduced by Δ_{ij}, and all other injections remain unchanged so that the new injections remain balanced. Under such changes, the branch flow change $\Delta f_{\hat{e}}$ on an edge $\hat{e}$ is determined by the DC power flow equations (1). The *generation shift sensitivity factor* between the pair of buses i,j and the edge $\hat{e}$ is defined as the ratio [29]

$$D_{\hat{e},ij} := \frac{\Delta f_{\hat{e}}}{\Delta p}.$$

When $e := (i,j) \in \mathcal{E}$ is an edge in the network, we also write $D_{\hat{e},ij}$ as $D_{\hat{e},e}$. The factor $D_{\hat{e},ij}$ is fully determined by the matrix A, and, if $\hat{e} = (w,z)$, it can be explicitly computed as (see [29]):

$$D_{\hat{e},ij} = B_{\hat{e}}(A_{iw} + A_{jz} - A_{iz} - A_{jw}).$$

Combining this formula with Proposition 4 yields the following result (proved in Appendix II):

Corollary 6. *For $i,j \in \mathcal{N}, \hat{e} = (w,z) \in \mathcal{E}$, we have*

$$D_{\hat{e},ij} = \frac{B_{\hat{e}}}{\sum_{E\in\mathcal{T}_{\mathcal{E}}}\chi(E)} \times \left(\sum_{E\in\mathcal{T}(\{i,w\},\{j,z\})} \chi(E) - \sum_{E\in\mathcal{T}(\{i,z\},\{j,w\})} \chi(E) \right).$$

Despite its complexity, this formula carries a clear graphical meaning, as we now explain. The two sums appearing at the numerator are over the spanning forests $\mathcal{T}(\{i,w\},\{j,z\})$ and $\mathcal{T}(\{i,z\},\{j,w\})$ respectively. Each element in $\mathcal{T}(\{i,w\},\{j,z\})$, as illustrated in Fig. 3, specifies a way to connect i to w and j to z through disjoint trees and captures a possible path for i,j to "spread" impact to (w,z). Similarly, elements in $\mathcal{T}(\{i,z\},\{j,w\})$ captures possible paths for i,j to "spread" impact to (z,w), which counting orientation, contributes negatively. Therefore, Corollary 6 implies that the impact of shifting generations from j to i propagates to the edge $\hat{e} = (w,z)$ through all possible spanning forests that connect the endpoints i,j,w,z (accounting for orientation). The relative strength of the trees in these two families determines the sign of $D_{\hat{e},ij}$.

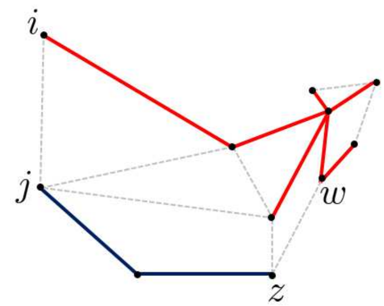

Fig. 3: An example element in $\mathcal{T}(\{i,w\},\{j,z\})$. The spanning trees containing $\{i,w\}$ and $\{j,z\}$ are highlighted as solid lines.

C. Effective Reactance

The Laplacian matrix L appears in circuit analysis as the admittance matrix (with a different weight), which explicitly relates the voltage and current vectors in an electrical resistive network [31]. In particular, given a network of resistors, it is shown in [31] that the *effective resistance* between two nodes i and j can be computed as

$$R_{ij} := L^{\dagger}_{ii} + L^{\dagger}_{jj} - L^{\dagger}_{ij} - L^{\dagger}_{ji}, \tag{5}$$

where $L^{\dagger}$ is the Penrose-Moore pseudoinverse of L. With an analogous calculation, we can show that in our setting (5) gives the *effective reactance* between the buses i and j. This means that, assuming we connect the buses i and j (with phase angles θ_i and θ_j respectively) to an external probing circuit, when there are no other injections in the network, the branch flow f_{ij} (from the external circuit) into bus i and out of bus j (into the external circuit) is given as

$$f_{ij} = \frac{\theta_i - \theta_j}{R_{ij}}.$$

Therefore, the network can be equivalently reduced to a single line with reactance R_{ij}. If $(i,j) \in \mathcal{E}$, the reactance of (i,j) is given as $X_{ij} := 1/B_{ij}$. Then, the physical intuition suggests that $R_{ij} \le X_{ij}$, as additional paths from i to j in the network should only decrease the overall reactance. We now give a graphical interpretation for the difference $X_{ij} - R_{ij}$ and use it to prove this result rigorously.

Corollary 7. *For all $(i,j) \in \mathcal{E}$ we have*

$$X_{ij} - R_{ij} = X_{ij} \cdot \frac{\sum_{E\in\mathcal{T}_{\mathcal{E}\setminus\{(i,j)\}}}\chi(E)}{\sum_{E\in\mathcal{T}_{\mathcal{E}}}\chi(E)}.$$

In particular, $X_{ij} \ge R_{ij}$ and the inequality is strict if and only if the graph $\mathcal{G}$ is still connected after removing (i,j).

The proof of this result is presented in Appendix III. From Corollary 7, we see that for an edge (i,j), the reduction ratio of its reactance coming from the network is precisely the portion of weighted spanning trees not passing through (i,j) among all spanning trees. Thus, more connections from the network leads to more reduction in the effective reactance on (i,j), which agrees with our physical intuition.

We remark that this reactance reduction ratio is closely related to the *spanning tree centrality measure* [32]. Indeed,

$$\frac{\sum_{E\in\mathcal{T}_{\mathcal{E}\setminus\{(i,j)\}}}\chi(E)}{\sum_{E\in\mathcal{T}_{\mathcal{E}}}\chi(E)} + c_{(i,j)} = 1,$$

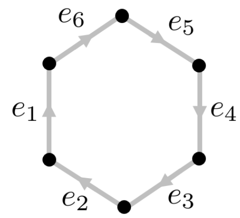

Fig. 4: A ring network with clockwise orientation. Line e_1 can only spread "negative" impacts to other lines.

where $c_{(i,j)}$ is the spanning tree centrality of (i,j). As a result

$$R_{ij} = X_{ij} c_{(i,j)},$$

which means that in a power network the spanning tree centrality $c_{(i,j)}$ of a transmission line (i,j) precisely captures the ratio between its effective reactance R_{ij} and its physical line reactance X_{ij}.

D. Line Outage Distribution Factor

As we discussed in Section II, when a line e is tripped from a power network $\mathcal{G}$, the line outage distribution factor $K_{\hat{e}e}$ captures the ratio between the flow change over line $\hat{e}$ with respect to the original branch flow on e before it is tripped. Writing $e = (i,j), \hat{e} = (w,z)$ with $i,j,w,z \in \mathcal{N}$, the constant $K_{\hat{e}e}$ can be computed as [29]:

$$K_{\hat{e}e} = \frac{X_e}{X_{\hat{e}}} \cdot \frac{A_{iw} + A_{jz} - A_{jw} - A_{iz}}{X_e - (A_{ii} + A_{jj} - A_{ij} - A_{ji})}.$$

This formula only holds if the graph $\mathcal{G}' := (\mathcal{N}, \mathcal{E} \setminus \{e\})$ is connected, as otherwise its denominator is 0 by Corollary 7. An edge of the graph $\mathcal{G}$ whose removal disconnects the graph is known as a *bridge* in the literature. We discuss this concept and its relationship to tree partitions in more detail in Section IV.

Using our previous results, we can readily derive the following new formula for $K_{\hat{e}e}$.

Theorem 8. *Let $e = (i,j)$, $\hat{e} = (w,z)$ be edges such that $\mathcal{G}' := (\mathcal{N}, \mathcal{E} \setminus \{e\})$ is connected. Then $K_{\hat{e}e}$ is given by*

$$B_{\hat{e}} \times \frac{\sum_{E \in \mathcal{T}(\{i,w\},\{j,z\})} \chi(E) - \sum_{E \in \mathcal{T}(\{i,z\},\{j,w\})} \chi(E)}{\sum_{E \in \mathcal{T}_{\mathcal{E} \setminus \{(i,j)\}}} \chi(E)}. \tag{6}$$

Proof. This result readily follows by dividing the identity in Corollary 6 by that in Corollary 7. □

Similar to our discussions in Section III-B, each term in (6) also carries clear graphical meanings: (a) The numerator of (6) states that the impact of tripping e propagates to $\hat{e}$ through all possible trees that connect e to $\hat{e}$, counting orientation. (b) The denominator of (6) sums over all spanning trees of $\mathcal{G}$ that do not pass through $e = (i,j)$, and each tree of this type specifies an alternative path that power can flow through if (i,j) is tripped. When there are more trees of this type, the network has a better ability to "absorb" the impact of (i,j) being tripped, and the denominator of (6) precisely captures this effect by saying that the impact of e being tripped on other lines is inversely proportional to the sum of all alternative tree paths in the network. (c) The $B_{\hat{e}}$ in (6) captures the intuition that a line with a larger reactance tends to be more robust against failures of other lines.

This graphical interpretation of identity (6) allows us to make general inferences on $K_{\hat{e}e}$ using only knowledge of the network topology. For example, in the ring network shown in Fig. 4, by inspecting the graph, we conclude that

$$K_{e_s e_1} < 0, \quad s = 2, 3, 4, 5, 6,$$

as e_1 can only spread "negative" impacts to other lines (the positive term in (6) vanishes).

From Theorem 8, we recover the following result from [33], whose original proof is much longer.

Corollary 9. *For adjacent lines $e = (i,j)$ and $\hat{e} = (i,k)$, we have*

$$K_{\hat{e}e} \geq 0.$$

Proof. For such e and $\hat{e}$, the second term in the numerator of (6) is over the empty set and thus equals to 0. □

Given two edges e and $\hat{e}$, $K_{\hat{e}e}$ captures how the branch flow on $\hat{e}$ changes if e is tripped. In particular, if $K_{\hat{e}e} = 0$, then the failure of e cannot lead to successive failure of $\hat{e}$ (in one step) as the branch flow on $\hat{e}$ is not impacted. The following result shows that the converse also holds if we are allowed to choose the injection and line capacities adversarially (see Appendix IV for its proof).

Proposition 10. *Given $e, \hat{e} \in \mathcal{E}$ with $K_{\hat{e}e} \neq 0$, there exist injections p and line capacities $\overline{f}$, such that before e trips*

$$|f_u| \leq \overline{f}_u \quad \forall u \in \mathcal{E},$$

while, after e trips,

$$|f'_{\hat{e}}| > \overline{f}_{\hat{e}},$$

where $f'_{\hat{e}}$ is the new flow on $\hat{e}$ over $\mathcal{G}' = (\mathcal{N}, \mathcal{E} \setminus \{e\})$.

In other words, for e and $\hat{e}$ such that $K_{\hat{e}e} \neq 0$, there always exists a scenario where the failure of e leads to the failure of $\hat{e}$. As such, $K_{\hat{e}e}$ determines whether it is possible for the failure of e to cause the failure of $\hat{e}$. The factors $K_{\hat{e}e}$'s can thus be interpreted as indicators on the system's ability to localize failures, and it is reasonable to optimize the topology $\mathcal{G}$ so that $K_{\hat{e}e} = 0$ for as many pairs of edges as possible.

E. Simple Loop Criterion

The formula (6) shows that the spanning forests $\mathcal{T}(\{i,w\},\{j,z\})$ and $\mathcal{T}(\{i,z\},\{j,w\})$ fully determine if $K_{\hat{e}e}$ is 0 or not. In other words, whether tripping e has any impact on $\hat{e}$ depends on how these edges are connected by subtrees in $\mathcal{G}$. We now establish an equivalent criterion that can be directly verified from the graph.

If $K_{\hat{e}e} \neq 0$, then (6) implies that at least one of the spanning forest sets $\mathcal{T}(\{i,w\},\{j,z\})$ and $\mathcal{T}(\{i,z\},\{j,w\})$ is nonempty. Without loss of generality, let us assume $\mathcal{T}(\{i,w\},\{j,z\}) \neq \emptyset$. For any element in $\mathcal{T}(\{i,w\},\{j,z\})$ (see Fig. 3), the tree containing $\{i,w\}$ induces a path from i to w, and the tree containing $\{j,z\}$ induces a path from j to z. By adjoining the edges $e = (i,j)$ and $\hat{e} = (w,z)$ to these two paths, we obtain a simple loop[1] containing both e and $\hat{e}$. As a result, $K_{\hat{e}e} \neq 0$ implies that we can find a simple loop in $\mathcal{G}$ which contains both e and $\hat{e}$.

[1]A loop is simple if it visits each node in the loop at most once.

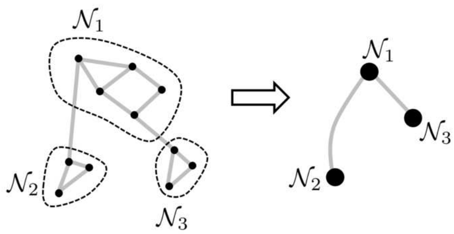

Fig. 5: The construction of $\mathcal{G}_{\mathcal{P}}$ from $\mathcal{P}$.

The converse, unfortunately, in general does not hold because of certain systems with high symmetry. That is, there exist pathological systems where a simple loop containing e and $\hat{e}$ exists, yet $K_{\hat{e}e} = 0$. Nevertheless, for such systems, by perturbing the line susceptances B_e with an arbitrarily small noise, we can "break" the symmetry and show that $K_{\hat{e}e} \neq 0$ almost surely. The detailed technical treatments for this perturbation analysis is presented in Part II of this paper.

The following proposition formally summarizes the discussions above:

Theorem 11 (Simple Loop Criterion)**.** *For $e = (i, j), \hat{e} = (w, z) \in \mathcal{E}$ such that $\mathcal{G}' := (\mathcal{N}, \mathcal{E} \setminus \{e\})$ is connected, we have $K_{\hat{e}e} \neq 0$ "if" and only if there exists a simple loop in $\mathcal{G}$ that contains both e and $\hat{e}$.*

We name this criterion "Simple Loop" both because it involves simple loops in $\mathcal{G}$ and because it is an intuitive, hence simple, criterion depending on loops. The "if" part of Theorem 11 should be interpreted as a probability one event under proper perturbations (see Part II of this paper for more details). This proposition shows that a simple loop containing e and $\hat{e}$ must exist if tripping e can possibly cause a successive failure of $\hat{e}$, since otherwise the branch flow on $\hat{e}$ is not impacted by the tripping of e. As a result, by clustering the network in smaller regions connected in a loop-free manner, we can prevent long-distance propagation of line failures. With this motivation in mind, in the next section we define precisely a *tree partition* of a power network and study its properties.

IV. Tree Partitions of Power Networks

For a power network $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, a collection $\mathcal{P} = \{\mathcal{N}_1, \mathcal{N}_2, \cdots, \mathcal{N}_k\}$ of subsets of $\mathcal{N}$ is said to form a *partition* of $\mathcal{G}$ if $\bigcup_{i=1}^k \mathcal{N}_i = \mathcal{N}$ and $\mathcal{N}_i \cap \mathcal{N}_j = \emptyset$ for $i \neq j$.

Given a partition $\mathcal{P} = \{\mathcal{N}_1, \mathcal{N}_2, \cdots, \mathcal{N}_k\}$ of $\mathcal{G}$ we define a *reduced multi-graph* $\mathcal{G}_{\mathcal{P}}$ as follows (see Fig. 5). First, $\mathcal{G}_{\mathcal{P}}$ has k nodes, one for each subset $\mathcal{N}_i$; we often use $\mathcal{N}_i$ to identify the nodes of $\mathcal{G}_{\mathcal{P}}$. Second, there is an undirected edge connecting the nodes $\mathcal{N}_i$ and $\mathcal{N}_j$ of $\mathcal{G}_{\mathcal{P}}$ for every pair of nodes $v \in \mathcal{N}_i$, $w \in \mathcal{N}_j$ in the original graph $\mathcal{G}$ that are adjacent in $\mathcal{G}$, i.e., $(v, w) \in \mathcal{E}$. There are multiple edges between nodes $\mathcal{N}_i$ and $\mathcal{N}_j$ of $\mathcal{G}_{\mathcal{P}}$ when multiple pairs of such nodes (v, w) exist in the original graph $\mathcal{G}$. Unlike the graph $\mathcal{G}$ (to which we have assigned an arbitrary orientation), the reduced multi-graph $\mathcal{G}_{\mathcal{P}}$ is undirected.

Definition 12. *A partition $\mathcal{P} = \{\mathcal{N}_1, \mathcal{N}_2, \cdots, \mathcal{N}_k\}$ of $\mathcal{G}$ is a **tree partition** if the reduced graph $\mathcal{G}_{\mathcal{P}}$ is a tree. The sets $\mathcal{N}_i$ are referred to as **regions** of $\mathcal{P}$. An edge $e \in \mathcal{E}$ with both endpoints inside the same region $\mathcal{N}_i$ is said to be **within** $\mathcal{N}_i$ and it is called a **bridge**[2] otherwise. We denote by $\mathcal{E}_i$ the set of edges within $\mathcal{N}_i$ and by $\mathcal{E}_b$ the set of bridges, so that $\mathcal{E} = \mathcal{E}_b \cup \bigcup_{i=1}^k \mathcal{E}_i$.*

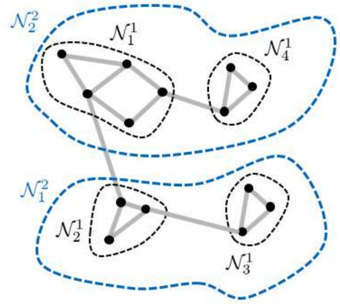

Fig. 6: An illustration of the partial order $\succeq$ over tree partitions. The partition $\mathcal{P}^1 = \{\mathcal{N}_1^1, \mathcal{N}_2^1, \mathcal{N}_3^1, \mathcal{N}_4^1\}$ is finer than $\mathcal{P}^2 = \{\mathcal{N}_1^2, \mathcal{N}_2^2\}$.

A tree partition is exhibited in Fig. 5 for the specific graph. Other tree partitions of the same graph exist: for instance, one can always consider the trivial partition $\mathcal{P}_0 = \{\mathcal{N}\}$, in which all the nodes are collapsed into a single region. The tree partition of a network $\mathcal{G}$ is thus not unique in general. Nevertheless, if we require the tree partition to be as "fine" as possible, such a partition is unique.

Specifically, given a graph $\mathcal{G}$, we define a partial order $\succeq$ over the set of all tree partitions of $\mathcal{G}$ (which is nonempty as it always contains the trivial partition $\mathcal{P}_0$) as follows. For two tree partitions $\mathcal{P}^1 = \{\mathcal{N}_1^1, \mathcal{N}_2^1, \cdots, \mathcal{N}_{k_1}^1\}$ and $\mathcal{P}^2 = \{\mathcal{N}_1^2, \mathcal{N}_2^2, \cdots, \mathcal{N}_{k_2}^2\}$, we say $\mathcal{P}^1$ is *finer* than $\mathcal{P}^2$, denoted as $\mathcal{P}^1 \succeq \mathcal{P}^2$, if for any $i = 1, 2, \ldots, k_1$, there exists some $j(i) \in \{1, 2, \ldots, k_2\}$ such that $\mathcal{N}_i^1 \subseteq \mathcal{N}_{j(i)}^2$. That is, $\mathcal{P}^1$ is finer than $\mathcal{P}^2$ if each region in $\mathcal{P}^1$ is contained in some region in $\mathcal{P}^2$ (see Fig. 6). It is easy to check that $\succeq$ is in fact a partial order over all possible tree partitions of $\mathcal{G}$.

Definition 13. *A tree partition $\mathcal{P}$ is said to be **irreducible** if $\mathcal{P}$ is maximal with respect to the partial order $\succeq$.*

In other words, an irreducible tree partition is a tree partition of $\mathcal{G}$ that cannot be made finer.

Proposition 14. *For any graph $\mathcal{G}$, the irreducible tree partition of $\mathcal{G}$ is unique.*

See Appendix V for a proof. We remark that our proof of Proposition 14 not only shows that the irreducible tree partition of $\mathcal{G}$ is unique, but also implies that the problem of computing this unique irreducible tree partition reduces to finding all the bridges of $\mathcal{G}$. We can thus adapt Tarjan's bridge-finding algorithm [34] and devise an algorithm that computes the irreducible tree partition of $\mathcal{G}$ in $O(n + m)$ time, as we summarize in Algorithm 1 below (see also the proof of Proposition 14 in Appendix V for more details).

V. Guaranteed Localization

In the previous section, motivated by the Simple Loop Criterion, we formally introduce the notion of tree partition of a power network. In this section we demonstrate three

[2] We remark that our definition of bridges agrees with the classical definition of bridges in graph theory (i.e., the removal of any such edge disconnects the original graph) in the sense that if the tree partition $\mathcal{P}$ is irreducible (see Definition 13 later) any bridge defined in our sense is a bridge in the classical sense, and vice versa.

Algorithm 1 Irreducible Tree Partition Finding Algorithm

1: Execute Tarjan's bridge-finding algorithm [34] on $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ to compute the set of bridges $\mathcal{E}_b$.
2: Remove edges in $\mathcal{E}_b$ from $\mathcal{E}$ to form the partitioned graph $(\mathcal{N}, \mathcal{E}\backslash\mathcal{E}_b)$.
3: Breadth-first search on the partitioned graph $(\mathcal{N}, \mathcal{E}\backslash\mathcal{E}_b)$ to compute its set of connected components $\mathcal{P} := \{C_1, C_2, \ldots, C_k\}$. Return $\mathcal{P}$.

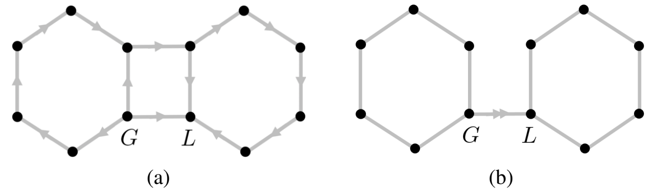

Fig. 7: (a) A double-ring network with G as generator bus and L as load bus, with arrows representing the original branch flow. (b) The new network after removing an edge, with arrows representing the new branch flow.

localization benefits of tree partitioning: (a) It confines the impact of line outages to within their own tree-partition regions; (b) It confines the impact of certain power injection disruptions to within their own tree partition-regions; and (c) Tree partitioning can sometimes reduce, rather than increase, line congestions. We demonstrate benefits (a) and (c) using a stylized example here, and both aspects will be more extensively studied in Part II of the paper. Benefit (b) is demonstrated by a more general result, which will also be used in Part III to design our real-time mitigation strategy for cascading failure.

A. Localization of Line Failures

Consider the double-ring network in Fig. 7(a), which contains exactly one generator and one load bus, denoted by G and L respectively. This network admits only the trivial tree partition and the branch flows are shown in Fig. 7(a). Consider a new network obtained by switching off the upper tie-line, creating a finer tree partition with two regions: the left ring and the right ring. This new network and the redistributed branch flows are shown in Fig. 7(b).

It is easy to check that in the original network Fig. 7(a) there is a simple loop containing any pair of edges. The Simple Loop Criterion implies that the impact of any single-line failure in this topology is global, i.e., all the other branch flows will change almost surely. Hence, after the single-line failure, every remaining line may carry a new branch flow that exceeds its capacity and subsequently trips (see Proposition 10). In contrast, for the network in Fig. 7(b), there are only two loops (corresponding to the two rings). The Simple Loop Criterion guarantees that the line failures inside one ring do not impact branch flows in the other and, thus, line failures are localized within their own regions.

Such localization, however, only applies in the first stage of a cascading failure, as further failures may involve bridges in the graph, to which the Simple Loop Criterion no longer applies. In fact, as we show in Part II of this paper, tripping a bridge almost always has a global impact and, therefore, failures may still propagate from one region to another after multiple steps. Further, the newly created bridge in Fig. 7(b) is a single-point vulnerability, whose failure disconnects the system into two islands. In Part III, we discuss how these drawbacks can be overcome by adopting a new control approach for fast-timescale frequency regulation.

B. Localization of Injection Disruptions

Besides the ability to localize the impact of line failures, the tree partition of a power network also ensures that disruptions in power injections in one region do not impact another region. This is due to a "decoupling" structure of the solution space of Laplacian equations, as we now describe. Given a tree partition $\mathcal{P} = \{\mathcal{N}_1, \mathcal{N}_2, \cdots, \mathcal{N}_l\}$ of $\mathcal{G}$, for $k = 1, 2, \ldots, l$, let

$$\partial\mathcal{N}_k := \{j : j \notin \mathcal{N}_k, \exists i \in \mathcal{N}_k \text{ s.t. } (i,j) \in \mathcal{E} \text{ or } (j,i) \in \mathcal{E}\}$$

and $\overline{\mathcal{N}}_k = \mathcal{N}_k \cup \partial\mathcal{N}_k$. The set $\partial\mathcal{N}_k$ defined above consists of the *boundary* buses of $\mathcal{N}_k$ in $\mathcal{G}$ and $\overline{\mathcal{N}}_k$ can be interpreted as the *closure* of $\mathcal{N}_k$. As an example, in Fig. 7(b), the closure of the left ring is given by the ring itself together with the bus L, and the closure of the right ring is given by the right ring itself together with the bus G.

Lemma 15. *Let $\mathcal{P} = \{\mathcal{N}_1, \mathcal{N}_2, \cdots, \mathcal{N}_l\}$ be a tree partition of $\mathcal{G}$ and consider a vector $b \in \mathbb{R}^n$ such that $b_j = 0$ for all $j \in \mathcal{N}_1$ and $\sum_{j\in\mathcal{N}_k} b_j = 0$ for $k \neq 1$. Then the Laplacian equation*

$$Lx = b \tag{7}$$

is solvable, and any solution x of (7) satisfies $x_i = x_j$ for all $i, j \in \overline{\mathcal{N}}_1$.

The proof of this result is presented in Appendix VI. Lemma 15 is used in Part III of this paper to design mitigation strategy that adjusts power injections in a way that localizes the impact of initial failures. It implies the following region independence property in the DC power flow context: If b represents perturbations in power injections, then x is the resulting changes in phase angles. Lemma 15 implies that the changes in phase angles are the same at every node in $\overline{\mathcal{N}}_1$. Therefore the angle difference across every line in or incident to $\mathcal{N}_1$ is zero and the branch flows on these lines are unchanged, i.e., the branch flows in $\mathcal{N}_1$ are independent of the perturbations in power injections in other regions as long as they are balanced.

This result holds only if the regions form a tree partition. In fact, the graph in Fig. 7(a) provides a counter-example when the network is not tree-partitioned: for this graph, if we set $\mathcal{N}_2 := \{G, L\}$ and collect the remaining buses in $\mathcal{N}_1$, then the injections satisfies the condition that $b_j = 0$ for all $j \in \mathcal{N}_1$ and $\sum_{j\in\mathcal{N}_2} b_j = 0$, yet the phase angle differences across edges in $\mathcal{N}_1$ are not zero.

C. Congestion Reduction

It is reasonable to expect that switching off lines to refine a tree partition may increase the stress on the remaining lines and, in this way, worsens network congestion. Nonetheless, by comparing the branch flows in Fig. 7(a) and (b), we notice that creating a nontrivial tree partition as in Fig. 7(b) actually can potentially remove the circulating flows and hence globally reduce network congestion. In fact, stronger result holds in this stylized example: the tree partition in Fig. 7(b) further minimizes the sum of absolute branch flows over all possible topologies on this network where G and L are connected. To see this, let

$$\mathfrak{C} := \{\mathcal{G}' = (\mathcal{N}, \mathcal{E}') : G \text{ and } L \text{ are connected in } \mathcal{G}'\}$$

be the collection of all graphs $\mathcal{G}' = (\mathcal{N}, \mathcal{E}')$ with edge set $\mathcal{E}' \subseteq \mathcal{N} \times \mathcal{N}$ (which do not need to be a sub-graph of Fig. 7(a)) such that G is connected to L (but G and L are not necessarily adjacent to each other). Let p' be the vector of injections described earlier, namely $p'_G = d$, $p'_L = -d$ for some $d > 0$, and $p'_j = 0$ for all the other buses. For each $\mathcal{G}' = (\mathcal{N}, \mathcal{E}') \in \mathfrak{C}$, define the metric

$$\Psi(\mathcal{G}') := \sum_{e \in \mathcal{E}'} |f'_e|,$$

where f'_e is the branch flow on e under the injection p'. Then we have the following result (see Appendix VII for its proof):

Proposition 16. *The graph in Fig. 7(b) minimizes the sum of absolute branch flow $\Psi(\cdot)$ over $\mathfrak{C}$.*

Hence not only may switching off lines *reduce* congestion, the tree partition in Fig. 7(b) in fact achieves the best possible congestion reduction in terms of $\Psi(\cdot)$. We revisit this phenomenon with case studies on more complex IEEE test systems in Part II of this paper.

VI. Conclusion

In Part I of this work, we develop a spectral theory using the transmission network Laplacian matrix that precisely captures the Kirkhhoff's Law in terms of graphical structures. Our results show that the distributions of different families of subtrees play an important role in understanding power redistribution and enables us to derive algebraic properties using purely graphical arguments. By considering a double ring network, we demonstrate that creating a non-trivial tree partition not only helps localize impacts of line failures, but also potentially reduces system congestion. The results in this part underlie the performance guarantees of a tree partition as well as its localization properties, which we present in Parts II and III of this paper, respectively.

References

[1] NERC, "1996 system disturbances: Review of selected 1996 electric system disturbances in North America," The Disturbance Analysis Working Group, North American Electric Reliability Council, Princeton Forrestal Village, Princeton, NJ, Report, 2002.

[2] P. D. Hines, I. Dobson, and P. Rezaei, "Cascading power outages propagate locally in an influence graph that is not the actual grid topology," *IEEE TPS*, vol. 32, no. 2, pp. 958–967, 2017.

[3] D. Bienstock and S. Mattia, "Using mixed-integer programming to solve power grid blackout problems," *Discrete Optimization*, vol. 4, no. 1, pp. 115 – 141, 2007.

[4] P. Hines, S. Talukdar *et al.*, "Controlling cascading failures with cooperative autonomous agents," *International journal of critical infrastructures*, vol. 3, no. 1, p. 192, 2007.

[5] R. Baldick, B. Chowdhury, I. Dobson, Z. Dong, B. Gou, D. Hawkins, H. Huang, M. Joung, D. Kirschen, F. Li *et al.*, "Initial review of methods for cascading failure analysis in electric power transmission systems IEEE PES CAMS task force on understanding, prediction, mitigation and restoration of cascading failures," in *2008 IEEE Power and Energy Society General Meeting-Conversion and Delivery of Electrical Energy in the 21st Century*. IEEE, 2008, pp. 1–8.

[6] H. Guo, C. Zheng, H. H.-C. Iu, and T. Fernando, "A critical review of cascading failure analysis and modeling of power system," *Elsevier Renewable and Sustainable Energy Reviews*, vol. 80, pp. 9–22, 2017.

[7] B. A. Carreras, V. E. Lynch, I. Dobson, and D. E. Newman, "Critical points and transitions in an electric power transmission model for cascading failure blackouts," *Chaos: An interdisciplinary journal of nonlinear science*, vol. 12, no. 4, pp. 985–994, 2002.

[8] M. Anghel, K. A. Werley, and A. E. Motter, "Stochastic model for power grid dynamics," in *HICSS*. IEEE, 2007, pp. 113–113.

[9] J. Yan, Y. Tang, H. He, and Y. Sun, "Cascading failure analysis with DC power flow model and transient stability analysis," *IEEE TPS*, vol. 30, no. 1, pp. 285–297, 2015.

[10] A. Bernstein, D. Bienstock, D. Hay, M. Uzunoglu, and G. Zussman, "Power grid vulnerability to geographically correlated failures: Analysis and control implications," in *IEEE INFOCOM*, 2014, pp. 2634–2642.

[11] D. P. Nedic, I. Dobson, D. S. Kirschen, B. A. Carreras, and V. E. Lynch, "Criticality in a cascading failure blackout model," *International Journal of Electrical Power & Energy Systems*, vol. 28, no. 9, pp. 627–633, 2006.

[12] M. A. Rios, D. S. Kirschen, D. Jayaweera, D. P. Nedic, and R. N. Allan, "Value of security: modeling time-dependent phenomena and weather conditions," *IEEE TPS*, vol. 17, no. 3, pp. 543–548, 2002.

[13] J. Song, E. Cotilla-Sanchez, G. Ghanavati, and P. D. Hines, "Dynamic modeling of cascading failure in power systems," *IEEE TPS*, vol. 31, no. 3, pp. 2085–2095, 2016.

[14] C. D. Brummitt, R. M. DSouza, and E. A. Leicht, "Suppressing cascades of load in interdependent networks," *Proceedings of the National Academy of Sciences*, vol. 109, no. 12, pp. E680–E689, 2012.

[15] Z. Kong and E. M. Yeh, "Resilience to degree-dependent and cascading node failures in random geometric networks," *IEEE TIT*, vol. 56, no. 11, pp. 5533–5546, Nov 2010.

[16] P. Crucitti, V. Latora, and M. Marchiori, "A topological analysis of the Italian electric power grid," *Physica A: Statistical mechanics and its applications*, vol. 338, no. 1-2, pp. 92–97, 2004.

[17] I. Dobson, B. A. Carreras, and D. E. Newman, "A loading-dependent model of probabilistic cascading failure," *Probab. Eng. Inf. Sci.*, vol. 19, no. 1, pp. 15–32, Jan. 2005.

[18] Z. Wang, A. Scaglione, and R. J. Thomas, "A Markov-transition model for cascading failures in power grids," in *HICSS*. IEEE, 2012, pp. 2115–2124.

[19] M. Rahnamay-Naeini, Z. Wang, N. Ghani, A. Mammoli, and M. M. Hayat, "Stochastic analysis of cascading-failure dynamics in power grids," *IEEE TPS*, vol. 29, no. 4, pp. 1767–1779, 2014.

[20] D. Mazauric, S. Soltan, and G. Zussman, "Computational analysis of cascading failures in power networks," in *Proceedings of the ACM SIGMETRICS/International Conference on Measurement and Modeling of Computer Systems*, ser. SIGMETRICS '13. New York, NY, USA: ACM, 2013, pp. 337–338.

[21] A. Bernstein, D. Bienstock, D. Hay, M. Uzunoglu, and G. Zussman, "Power grid vulnerability to geographically correlated failures 2014: analysis and control implications," in *IEEE INFOCOM*, April 2014, pp. 2634–2642.

[22] D. Bienstock and A. Verma, "The $n-k$ problem in power grids: New models, formulations, and numerical experiments," *SIAM Journal on Optimization*, vol. 20, no. 5, pp. 2352–2380, 2010.

[23] L. Guo and S. H. Low, "Spectral characterization of controllability and observability for frequency regulation dynamics," in *CDC*. IEEE, 2017, pp. 6313–6320.

[24] L. Guo, C. Zhao, and S. H. Low, "Graph laplacian spectrum and primary frequency regulation," in *CDC*. IEEE, 2018, pp. 158–165.

[25] ——, "Cyber network design for secondary frequency regulation: A spectral approach," in *PSCC*. IEEE, 2018.

[26] L. Guo, C. Liang, and S. H. Low, "Monotonicity properties and spectral characterization of power redistribution in cascading failures," *55th Annual Allerton Conference*, 2017.

[27] A. E. Motter and Y.-C. Lai, "Cascade-based attacks on complex networks," *Phys Rev E*, Dec. 2002.

[28] C. Zhao, U. Topcu, N. Li, and S. Low, "Power system dynamics as primal-dual algorithm for optimal load control," *arXiv preprint arXiv:1305.0585*, 2013.

[29] A. Wood and B. Wollenberg, *Power Generation, Operation, and Control*. Wiley-Interscience, 1996.

[30] S. Chaiken, "A combinatorial proof of the all minors matrix tree theorem," *SIAM Journal on Algebraic Discrete Methods*, vol. 3, no. 3, pp. 319–329, 1982.

[31] F. Dorfler and F. Bullo, "Kron reduction of graphs with applications to electrical networks," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 60, no. 1, pp. 150–163, Jan 2013.

[32] A. S. Teixeira, P. T. Monteiro, J. A. Carriço, M. Ramirez, and A. P. Francisco, "Spanning edge betweenness," in *Workshop on mining and learning with graphs*, vol. 24, 2013, pp. 27–31.

[33] C. Lai and S. H. Low, "The redistribution of power flow in cascading failures," in *2013 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Oct 2013, pp. 1037–1044.

[34] R. E. Tarjan, "A note on finding the bridges of a graph," *Information Processing Letters*, vol. 2, no. 6, pp. 160–161, 1974.

[35] K. Menger, "Zur allgemeinen kurventheorie," *Fundamenta Mathematicae*, vol. 10, no. 1, pp. 96–115, 1927.

[36] D. B. West *et al.*, *Introduction to graph theory*. Prentice hall Upper Saddle River, NJ, 1996, vol. 2.

Appendix I
Proof of Proposition 4

Recall that without loss of generality, we choose the node n as reference node and defined the matrix A accordingly via (3).

If $i = n$ or $j = n$, it is easy to see that $\mathcal{T}(\{i,j\},\{n\}) = \emptyset$ so that $A_{ij} = 0$. Now suppose $i, j \neq n$ and let A_j denote the j-th column of A after removing the reference node. Note from the definition of A that

$$\overline{L}A_j = e_j,$$

where $e_j \in \mathbb{R}^{n-1}$ is the vector with 1 as its j-th component and 0 elsewhere. Cramer's rule gives

$$A_{ij} = \frac{\det\left(\overline{L}_j^i\right)}{\det\left(\overline{L}\right)},$$

where $\overline{L}_j^i$ is the matrix obtained by replacing the i-th column of $\overline{L}$ by e_j. Now, by Proposition 3, we have

$$\det\left(\overline{L}_j^i\right) = (-1)^{i+j}\det\left(\overline{L}^{ij}\right) = \sum_{E\in\mathcal{T}(\{i,j\},\{n\})}\chi(E),$$

and by the Kirchhoff's Matrix Tree Theorem

$$\det\left(\overline{L}\right) = \sum_{E\in\mathcal{T}_\mathcal{E}}\chi(E).$$

This finishes the proof. □

Appendix II
Proof of Corollary 6

Proposition 4 implies that

$$\begin{aligned}&\left(\sum_{E\in\mathcal{T}_\mathcal{E}}\chi(E)\right)(A_{iw}+A_{jz}-A_{iz}-A_{jw})\\ &=\sum_{E\in\mathcal{T}(\{i,w\},\{n\})}\chi(E)+\sum_{E\in\mathcal{T}(\{j,z\},\{n\})}\chi(E)\\ &-\sum_{E\in\mathcal{T}(\{i,z\},\{n\})}\chi(E)-\sum_{E\in\mathcal{T}(\{j,w\},\{n\})}\chi(E).\end{aligned} \tag{8}$$

We can decompose the set $\mathcal{T}(\{i,w\},\{n\})$ based on the tree to which node j belongs. This leads to the identity

$$\mathcal{T}(\{i,w\},\{n\}) = \mathcal{T}(\{i,j,w\},\{n\}) \sqcup \mathcal{T}(\{i,w\},\{j,n\}),$$

where $\sqcup$ means disjoint union. Similarly, we also have

$$\begin{aligned}\mathcal{T}(\{j,z\},\{n\}) &= \mathcal{T}(\{i,j,z\},\{n\}) \sqcup \mathcal{T}(\{j,z\},\{i,n\}),\\ \mathcal{T}(\{i,z\},\{n\}) &= \mathcal{T}(\{i,j,z\},\{n\}) \sqcup \mathcal{T}(\{i,z\},\{j,n\}),\\ \mathcal{T}(\{j,w\},\{n\}) &= \mathcal{T}(\{i,j,w\},\{n\}) \sqcup \mathcal{T}(\{j,w\},\{i,n\}).\end{aligned}$$

Substituting the above decompositions into (8) and simplifying, we obtain

$$\begin{aligned}&\left(\sum_{E\in\mathcal{T}_\mathcal{E}}\chi(E)\right)(A_{iw}+A_{jz}-A_{iz}-A_{jw})\\ &=\sum_{E\in\mathcal{T}(\{i,w\},\{j,n\})}\chi(E)+\sum_{E\in\mathcal{T}(\{j,z\},\{i,n\})}\chi(E)\\ &-\sum_{E\in\mathcal{T}(\{i,z\},\{j,n\})}\chi(E)-\sum_{E\in\mathcal{T}(\{j,w\},\{i,n\})}\chi(E).\end{aligned} \tag{9}$$

Furthermore, the following set of identities hold:

$$\begin{aligned}&\mathcal{T}(\{i,w\},\{j,n\})\\ &=\mathcal{T}(\{i,w\},\{j,z,n\}) \sqcup \mathcal{T}(\{i,w,z\},\{j,n\}),\\ &\mathcal{T}(\{j,z\},\{i,n\})\\ &=\mathcal{T}(\{j,z\},\{i,w,n\}) \sqcup \mathcal{T}(\{j,w,z\},\{i,n\}),\\ &\mathcal{T}(\{j,w\},\{i,n\})\\ &=\mathcal{T}(\{j,w\},\{i,z,n\}) \sqcup \mathcal{T}(\{j,w,z\},\{i,n\}),\\ &\mathcal{T}(\{i,z\},\{j,n\})\\ &=\mathcal{T}(\{i,z\},\{j,w,n\}) \sqcup \mathcal{T}(\{i,w,z\},\{j,n\}).\end{aligned}$$

Substituting these into (9) and rearranging yields

$$\begin{aligned}&\left(\sum_{E\in\mathcal{T}_\mathcal{E}}\chi(E)\right)(A_{iw}+A_{jz}-A_{iz}-A_{jw})\\ &=\sum_{E\in\mathcal{T}(\{i,w\},\{j,z,n\})}\chi(E)+\sum_{E\in\mathcal{T}(\{j,z\},\{i,w,n\})}\chi(E)\\ &-\sum_{E\in\mathcal{T}(\{j,w\},\{i,z,n\})}\chi(E)-\sum_{E\in\mathcal{T}(\{i,z\},\{j,w,n\})}\chi(E)\\ &=\sum_{E\in\mathcal{T}(\{i,w\},\{j,z\})}\chi(E)-\sum_{E\in\mathcal{T}(\{j,w\},\{i,z\})}\chi(E),\end{aligned}$$

where the last equality follows from

$$\begin{aligned}&\mathcal{T}(\{i,w\},\{j,z\})\\ &=\mathcal{T}(\{i,w\},\{j,z,n\}) \sqcup \mathcal{T}(\{j,z\},\{i,w,n\})\end{aligned}$$

and

$$\begin{aligned}&\mathcal{T}(\{j,w\},\{i,z\})\\ &=\mathcal{T}(\{j,w\},\{i,z,n\}) \sqcup \mathcal{T}(\{i,z\},\{j,w,n\}\,.\end{aligned}$$

This completes the proof. □

Appendix III
Proof of Corrollary 7

The proof uses the following relation between $L^\dagger$ and A:

Lemma 17. *For every edge $(i,j) \in \mathcal{E}$ we have*

$$L^\dagger_{ii} + L^\dagger_{jj} - L^\dagger_{ij} - L^\dagger_{ji} = A_{ii} + A_{jj} - A_{ij} - A_{ji}.$$

Proof. We know that if the equation $p = L\theta$ is solvable, the solution θ is unique after quotienting away the kernel of L, which is given by $\mathrm{span}(\mathbf{1})$. Noting that the latter coincides with the kernel of C^T, we conclude that the vector $f = BC^T\theta$ is uniquely determined.

To prove the desired identity, we will equate two alternative formulas for the branch flows f. The first one relies on the fact that $L^\dagger p$ always gives a feasible θ and, therefore,

$$f = BC^T L^\dagger p. \tag{10}$$

The second way to calculate f is to set the phase angle at the reference node to zero, which indirectly implies that

$$\tilde{\theta} = \begin{bmatrix}\overline{\theta}\\ 0\end{bmatrix} = \begin{bmatrix}\overline{L}^{-1}\overline{p}\\ 0\end{bmatrix} = Ap,$$

where $\overline{\theta}$ and $\overline{p}$ are the vector of non-reference node phase angles and injections. We then have

$$f = BC^T\tilde{\theta} = BC^T Ap. \tag{11}$$

Now take $i, j \in \mathcal{N}$. Setting the injections $p_i = -p_j = 1$ and zero elsewhere, by equating the branch flow f_{ij} computed from (10) and (11), we obtain that

$$L^\dagger_{ii} + L^\dagger_{jj} - L^\dagger_{ij} - L^\dagger_{ji} = A_{ii} + A_{jj} - A_{ij} - A_{ji}. \quad \square$$

In other words, we can replace the $L^\dagger$'s in equation (5) by the matrix A, which allows us to apply Proposition 4.

Now by taking $w = i$, $z = j$ in Corollary 6 and applying Lemma 17, we see that

$$R_{ij} = A_{ii} + A_{jj} - 2A_{ij} = \frac{\sum_{E\in\mathcal{T}(\{i\},\{j\})}\chi(E)}{\sum_{E\in\mathcal{T}_\mathcal{E}}\chi(E)}.$$

For each forest in $\mathcal{T}(\{i\},\{j\})$, we can add the edge (i, j) to form a spanning tree passing through (i, j). Conversely, each spanning tree passing through (i, j) after removing the edge (i, j) produces a forest in $\mathcal{T}(\{i\},\{j\})$. By definition of $\chi(E)$, this fact implies

$$\sum_{E\in\mathcal{T}(\{i\},\{j\})}\chi(E) = X_{ij}\sum_{E\in\mathcal{T}'}\chi(E),$$

where $\mathcal{T}'$ denotes the set of all spanning trees passing through (i, j). Therefore,

$$\begin{aligned} X_{ij} - R_{ij} &= X_{ij}\cdot\frac{\sum_{E\in\mathcal{T}_\mathcal{E}}\chi(E) - \sum_{E\in\mathcal{T}'}\chi(E)}{\sum_{E\in\mathcal{T}_\mathcal{E}}\chi(E)} \\ &= X_{ij}\cdot\frac{\sum_{E\in\mathcal{T}_{\mathcal{E}\setminus\{(i,j)\}}}\chi(E)}{\sum_{E\in\mathcal{T}_\mathcal{E}}\chi(E)}, \end{aligned}$$

and the proof is completed. $\square$

Appendix IV
Proof of Proposition 10

In order to prove the desired result, we first establish a lemma that pertains the branch flow directions on all edges under a specific type of injections. More concretely, given an edge $e = (i, j) \in \mathcal{E}$, we can define its "characteristic" injection p^e by

$$p^e_k = \begin{cases} 1, & k = i, \\ -1, & k = j, \\ 0, & \text{otherwise.} \end{cases}$$

The following result shows that the sign of the branch flow on $\hat{e}$ under p^e is determined by $D_{\hat{e}e}$, the generation shift sensitivity factor between e and $\hat{e}$ (see Section III-B).

Lemma 18. *Under the injections p^e, for any $\hat{e} = (w, z) \in \mathcal{E}$ we have*

$$f_{\hat{e}} = D_{\hat{e}e}.$$

Proof. From the DC power flow equations we have $f_{\hat{e}} = B_{\hat{e}}(\theta_w - \theta_z)$, where $\theta = Ap$ is the phase angles vector (with $\theta_n = 0$ for the reference bus n). By expanding the matrix products and noting that A is symmetric, we see that

$$f_{\hat{e}} = B_{\hat{e}}(A_{iw} + A_{jz} - A_{iz} - A_{jw}) = D_{\hat{e}e}. \quad \square$$

In particular, for the edge e, we know $f_e = D_{ee} = B_e R_e > 0$, where R_e is the effective reactance of e, which agrees with our intuition.

Now let us consider the case $K_{\hat{e}e} > 0$. When this holds, under the injection p^e, we have $f_e > 0$ and $f_{\hat{e}} > 0$ from Lemma 18. Choose a line capacity vector $\overline{f}$ so that the capacity at $\hat{e}$ is exactly at $f_{\hat{e}}$ and the capacities for other edges are sufficiently large in the sense that these lines operate safely before and after tripping e. One example of $\overline{f}$ is given by

$$\overline{f}_u := \begin{cases} f_{\hat{e}}, & u = \hat{e} \\ (1 + \|K\|_\infty)\,\|f\|_\infty\,, & \text{otherwise,} \end{cases}$$

where $\|\cdot\|_\infty$ denotes the sup norm. With this choice of capacities, before tripping e, it is easy to see that f is a safe operating point under $\overline{f}$. Further, after we trip e from $\mathcal{G}$, for any edge u other than $\hat{e}$, the new flow satisfies

$$\begin{aligned} |f'_u| &= |f_u + K_{ue}f_e| \\ &\le |f_u| + |K_{ue}|\,|f_e| \\ &\le \|f\|_\infty + \|K\|_\infty\,\|f\|_\infty \\ &= \overline{f}_u, \end{aligned}$$

and for $\hat{e}$ we have

$$|f'_{\hat{e}}| = |f_{\hat{e}} + K_{\hat{e}e}f_e| = f_{\hat{e}} + K_{\hat{e}e}f_e > \overline{f}_{\hat{e}},$$

where the second equality is because $f_e, f_{\hat{e}}, K_{\hat{e}e} > 0$. In other words, after power redistributes over the new graph $\mathcal{G}' := (\mathcal{N}, \mathcal{E}\setminus\{e\})$, the branch flow on $\hat{e}$ is over its capacity yet all other lines stay within the safe range. As a result, $\hat{e}$ would be tripped in the next stage.

The case $K_{\hat{e}e} < 0$ is analogous. In particular, for this case we have $f_e > 0$, $f_{\hat{e}} < 0$ and hence

$$f_{\hat{e}} + K_{\hat{e}e}f_e < f_{\hat{e}}.$$

The remaining argument follows a similar construction to the case $K_{\hat{e}e} > 0$. The desired result then follows. $\square$

Appendix V
Proof of Proposition 14

For $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, let $\mathcal{E}_b$ be the set of edges that all spanning trees of $\mathcal{G}$ pass through. In other words, $\mathcal{E}_b$ is the set of all bridges in classical graph theory. Let $\mathcal{G}_b = (\mathcal{N}, \mathcal{E}\setminus\mathcal{E}_b)$ denote the graph obtained from $\mathcal{G}$ by removing all edges in $\mathcal{E}_b$, and let $\mathcal{P}^* = \{C_1, C_2, \cdots, C_k\}$ be its connected components, where $k = |\mathcal{E}_b| + 1$. We claim $\mathcal{P}^*$ is an irreducible tree partition of $\mathcal{G}$.

First we show that the reduced graph $\mathcal{G}_{\mathcal{P}^*}$ is a tree. Assume not, then there is a loop in $\mathcal{G}_{\mathcal{P}^*}$. Without loss of generality, let us assume the loop is $(C_1, C_2, \cdots, C_l)$ for some $2 \le l \le k$. Then by the construction of $\mathcal{G}_{\mathcal{P}^*}$, there exist nodes $n^t_1, n^s_1, n^t_2, n^s_2, n^t_3, \cdots, n^t_l, n^s_l$ such that:

1) For each i, the nodes $n^s_i, n^t_i \in C_i$;
2) For each i, the edge $e_{i,i+_l 1} := (n^s_i, n^t_{i+_l 1}) \in \mathcal{E}$, where $+_l$ denotes the addition modulo l.

For any i, since C_i is connected, we can find a path P_i from n^t_i to n^s_i. It is then clear that the concatenated path

$$(P_1, e_{1,2}, P_2, e_{2,3}\ldots e_{l-1,l}, P_l, e_{l,1})$$

forms a loop in the original graph $\mathcal{G}$. Consequently, not all spanning trees pass through $e_{1,2}$ and thus $e_{1,2} \notin \mathcal{E}_b$, which leads to a contradiction.

Next we prove $\mathcal{P}^*$ is irreducible by showing that $\mathcal{P}^*$ is finer than any tree partition $\mathcal{P} := \{\mathcal{N}_1, \mathcal{N}_2, \cdots, \mathcal{N}_{k'}\}$ of $\mathcal{G}$. Consider a region in $\mathcal{P}^*$, say C_1. Since both $\mathcal{P}^*$ and $\mathcal{P}$ are partitions of $\mathcal{G}$, there must be some region in $\mathcal{P}$, say $\mathcal{N}_1$, such that $C_1 \cap \mathcal{N}_1 \ne \emptyset$. We claim that $C_1 \subset \mathcal{N}_1$. Otherwise, there exists another region in $\mathcal{P}$, say

$\mathcal{N}_2$, such that $C_1 \cap \mathcal{N}_2 \neq \emptyset$. Pick $n_1 \in C_1 \cap \mathcal{N}_1$ and $n_2 \in C_1 \cap \mathcal{N}_2$. Then $n_1 \neq n_2$ because $\mathcal{N}_1 \cap \mathcal{N}_2 = \emptyset$. Now since $n_1, n_2 \in C_1$, and C_1 does not contain any bridge (in classical graph theory sense), by Menger's Theorem [35], there exists a cycle (which is not necessarily simple) in C_1 containing both n_1 and n_2. By collapsing adjacent nodes in this cycle that belong to common regions, we can find regions $\mathcal{N}^1_{l_1}, \mathcal{N}^1_{l_2}, \cdots, \mathcal{N}^1_{l_{p_1}}, \mathcal{N}^2_{l_1}, \mathcal{N}^2_{l_2}, \cdots, \mathcal{N}^2_{l_{p_2}}$ so that the path from n_1 to n_2 in this cycle is given by

$$\left(P^1_1, e^1_{1,l_1}, P^1_{l_1}, e^1_{l_1,l_2}, \cdots, e^1_{l_{p_1},2}, P^1_2\right)$$

and the path from n_2 to n_1 in this cycle is given by

$$\left(P^2_2, e^2_{2,l_1}, P^2_{l_1}, e^2_{l_1,l_2}, \cdots, e^2_{l_{p_2},1}, P^2_1\right),$$

where $e^1_{i,j}, e^2_{i,j}$ are edges with source nodes in $\mathcal{N}_i$ and target nodes in $\mathcal{N}_j$ and P^1_i, P^2_i are paths contained in $\mathcal{N}_i$. Therefore,

$$\left(\mathcal{N}_1, \mathcal{N}^1_{l_1}, \cdots, \mathcal{N}^1_{l_{p_1}}, \mathcal{N}_2, \mathcal{N}^2_{l_1}, \cdots, \mathcal{N}^2_{l_{p_2}}\right)$$

forms a loop in $\mathcal{G}_\mathcal{P}$. This implies that $\mathcal{P}$ is not a tree partition, contradicting our assumption.

We have shown that $\mathcal{P}^*$ is an irreducible tree partition of $\mathcal{G}$. Moreover, for any other irreducible tree partition $\overline{\mathcal{P}}$, the above proof shows that

$$\mathcal{P}^* \succeq \overline{\mathcal{P}}.$$

Since $\overline{\mathcal{P}}$ is irreducible and thus maximal with respect to $\succeq$, we see $\overline{\mathcal{P}} = \mathcal{P}^*$. In other words, any irreducible tree partition of $\mathcal{G}$ must coincide with $\mathcal{P}^*$. This completes our proof. $\square$

Appendix VI
Proof of Lemma 15

As we discussed in Section II, the Laplacian matrix $L := CBC^T$ of a connected graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ has rank $n-1$, and $Lx = b$ is solvable if and only if $\mathbf{1}^T b = 0$, where $\mathbf{1}$ is the vector with a proper dimension that consists of ones. Moreover, the kernel of L is given by $\operatorname{span}(\mathbf{1})$.

If $\mathcal{N}_1$ is the only region in $\mathcal{P}$, then $b = 0$ since $b_j = 0$ for all $j \in \mathcal{N}_1$. We thus know the solution space to $Lx = b$ is exactly the kernel of L, and the desired result holds.

If $\mathcal{N}_1$ is not the only region in $\mathcal{P}$, then we can find a bus that does not belong to $\mathcal{N}_1$, say bus z. Without loss of generality, assume the bus $z \in \mathcal{N}_k$ and corresponds to the last row and column in L. Consider a solution x to $Lx = b$. Since the kernel of L is $\operatorname{span}(\mathbf{1})$, we can without loss of generality assume that the last component of x is 0. Let $\overline{L}$ be the submatrix of L obtained by removing its last row and last column, and similarly let $\overline{x}$ and $\overline{b}$ be the vectors obtained by removing the last component of x and b, respectively. Then $\overline{L}$ is invertible, and we have

$$\overline{L}\,\overline{x} = \overline{b}.$$

Denote the matrix obtained by deleting the l-th row and i-th column of $\overline{L}$ by $\overline{L}^{li}$, then by Proposition 4 we have

$$\det\left(\overline{L}^{li}\right) = (-1)^{l+i} \sum_{E \in \mathcal{T}(\{l,i\},\{z\})} \chi(E), \tag{12}$$

where $\chi(E) = \prod_{e \in E} B_e$ and $\mathcal{T}(\{l,i\},\{z\})$ is the set of spanning forests of $\mathcal{G}$ that consists of exactly two trees containing $\{l,i\}$ and $\{z\}$, respectively.

To state some useful results derived from (12), we introduce the following definition of directly connected regions:

Definition 19. *For a tree partition $\mathcal{P} = \{\mathcal{N}_1, \mathcal{N}_2, \cdots, \mathcal{N}_k\}$ of $\mathcal{G}$, we say $\mathcal{N}_v$ and $\mathcal{N}_w$ are* **directly connected without** *$\mathcal{N}_l$ if the path from $\mathcal{N}_v$ to $\mathcal{N}_w$ in $\mathcal{G}_\mathcal{P}$ does not contain $\mathcal{N}_l$.*

The path from $\mathcal{N}_v$ to $\mathcal{N}_w$ in the above definition is unique since $\mathcal{G}_\mathcal{P}$ forms a tree. As an example, in Fig. 5, $\mathcal{N}_1$ and $\mathcal{N}_2$ are directly connected without $\mathcal{N}_3$, while $\mathcal{N}_2$ and $\mathcal{N}_3$ are not directly connected without $\mathcal{N}_1$.

In the following proofs we need to refer to paths in both the original graph $\mathcal{G}$ and the reduced graph $\mathcal{G}_\mathcal{P}$. To clear potential confusions, we introduce the following terminology. Given two sets of nodes $\mathcal{N}_v$ and $\mathcal{N}_w$ (that can be different from the tree-partition regions in $\mathcal{P}$) of $\mathcal{G}$, a path in $\mathcal{G}$ from $\mathcal{N}_v$ to $\mathcal{N}_w$ refers to a path consisting of nodes (and lines) from the original graph $\mathcal{G}$ whose starting node belongs to $\mathcal{N}_v$ and ending node belongs to $\mathcal{N}_w$. Given two tree-partition regions $\mathcal{N}_v$ and $\mathcal{N}_w$, a path in $\mathcal{G}_\mathcal{P}$ from $\mathcal{N}_v$ to $\mathcal{N}_w$ refers to a path consisting of nodes (and lines) from the reduced graph $\mathcal{G}_\mathcal{P}$ whose starting node is $\mathcal{N}_v$ and ending node is $\mathcal{N}_w$. Since there is a natural correspondence between bridges in $\mathcal{G}$ and lines in $\mathcal{G}_\mathcal{P}$, if a line e in $\mathcal{G}_\mathcal{P}$ is contained in a path P in $\mathcal{G}_\mathcal{P}$, we also say the corresponding bridge $\tilde{e}$ from $\mathcal{G}$ is contained in P.

Lemma 20. *Assume $\mathcal{N}_2$ and $\mathcal{N}_k$ are not directly connected without $\mathcal{N}_1$. If $l_1, l_2 \in \mathcal{N}_2$, $z \in \mathcal{N}_k$ and $i \in \overline{\mathcal{N}}_1$, then*

$$\mathcal{T}(\{l_1, i\}, \{z\}) = \mathcal{T}(\{l_2, i\}, \{z\}).$$

Proof. The path from $\mathcal{N}_1$ to $\mathcal{N}_2$ in $\mathcal{G}_\mathcal{P}$ contains a bridge in $\mathcal{G}$ that incidents to $\mathcal{N}_1$. Denote this bridge as $\tilde{e}$ and let w be the endpoint of $\tilde{e}$ that is not in $\mathcal{N}_1$. Then it is easy to check that w is a cut node that any path from $\overline{\mathcal{N}}_1$ to $\mathcal{N}_2$ in $\mathcal{G}$ must contain.

Since $\mathcal{N}_2$ and $\mathcal{N}_k$ are not directly connected without $\mathcal{N}_1$, the path from $\mathcal{N}_2$ to $\mathcal{N}_k$ in $\mathcal{G}_\mathcal{P}$ passes through $\mathcal{N}_1$. In other words, any path in $\mathcal{G}$ from $\mathcal{N}_2$ to $\mathcal{N}_k$ must pass through a certain node in $\mathcal{N}_1$, and thus contains a sub-path in $\mathcal{G}$ from $\mathcal{N}_1$ to $\mathcal{N}_2$. This implies that w is contained in any path in $\mathcal{G}$ from $\mathcal{N}_2$ to $\mathcal{N}_k$.

Note that any tree containing $i \in \overline{\mathcal{N}}_1$ and $l_1 \in \mathcal{N}_2$ induces a path in $\mathcal{G}$ from $\overline{\mathcal{N}}_1$ to $\mathcal{N}_2$ and thus contains w. Further, any tree containing $l_2 \in \mathcal{N}_2$ and $z \in \mathcal{N}_k$ induces a path from $\mathcal{N}_2$ to $\mathcal{N}_k$ in $\mathcal{G}$, and thus also contains w. As a result, these two types of trees always share a common node w and cannot be disjoint:

$$\mathcal{T}(\{l_1, i\}, \{l_2, z\}) = \emptyset.$$

Similarly ,

$$\mathcal{T}(\{l_2, i\}, \{l_1, z\}) = \emptyset.$$

Therefore,

$$\begin{aligned} &\mathcal{T}(\{l_1, i\}, \{z\}) \\ &= \mathcal{T}(\{l_1, l_2, i\}, \{z\}) \sqcup \mathcal{T}(\{l_1, i\}, \{l_2, z\}) \\ &= \mathcal{T}(\{l_1, l_2, i\}, \{z\}) \sqcup \mathcal{T}(\{l_2, i\}, \{l_1, z\}) \\ &= \mathcal{T}(\{l_2, i\}, \{z\}), \end{aligned}$$

where $\sqcup$ means disjoint union. The desired result then follows. $\square$

Lemma 21. *Assume $\mathcal{N}_2$ and $\mathcal{N}_k$ are directly connected without $\mathcal{N}_1$. If $l \in \mathcal{N}_2$ and $i_1, i_2 \in \overline{\mathcal{N}_1}$, then*

$$\mathcal{T}(\{l, i_1\}, \{z\}) = \mathcal{T}(\{l, i_2\}, \{z\}).$$

Proof. The path from $\mathcal{N}_1$ to $\mathcal{N}_k$ in $\mathcal{G}_\mathcal{P}$ (denoted as P_1) contains a bridge in $\mathcal{G}$ that incidents to $\mathcal{N}_1$. Denote this bridge as $\tilde{e}$ and let w be the endpoint of $\tilde{e}$ that does not belong to $\mathcal{N}_1$. Then it is easy to check that w is a cut node that any path in $\mathcal{G}$ from $\overline{\mathcal{N}}_1$ to $\mathcal{N}_k$ must pass through.

We claim that if $\mathcal{N}_2$ and $\mathcal{N}_k$ are directly connected without $\mathcal{N}_1$, then any path from $\overline{\mathcal{N}}_1$ to $\mathcal{N}_2$ in $\mathcal{G}$ must also contain w. Indeed, suppose not, then the path from $\mathcal{N}_1$ to $\mathcal{N}_2$ in $\mathcal{G}_\mathcal{P}$ (denoted as P_2) contains a bridge in $\mathcal{G}$ that incidents to $\mathcal{N}_1$, and this bridge is different from $\tilde{e}$. If P_1 and P_2 do not have any common super nodes, then concatenating the two paths induces a path in $\mathcal{G}_\mathcal{P}$ from $\mathcal{N}_2$ to $\mathcal{N}_k$ that passes through $\mathcal{N}_1$. In other words, the path from $\mathcal{N}_2$ to $\mathcal{N}_k$ in $\mathcal{G}_\mathcal{P}$ passes through $\mathcal{N}_1$, contradicting the assumption that $\mathcal{N}_2$ and $\mathcal{N}_k$ are directly connected without $\mathcal{N}_1$. Therefore, P_1 and P_2 share a common node, say $\mathcal{N}_3$. However, P_1 and P_2 induce two different sub-paths in $\mathcal{G}_\mathcal{P}$ from $\mathcal{N}_1$ to $\mathcal{N}_3$, contradicting the assumption that $\mathcal{G}_\mathcal{P}$ forms a tree. We thus have proved the claim.

Finally, note that any tree containing $i_1 \in \overline{\mathcal{N}}_1$ and $l \in \mathcal{N}_2$ induces a path in $\mathcal{G}$ from $\overline{\mathcal{N}}_1$ to $\mathcal{N}_2$ and thus contains w. Further, any tree containing $i_2 \in \overline{\mathcal{N}}_1$ and $z \in \mathcal{N}_k$ induces a path in $\mathcal{G}$ from $\overline{\mathcal{N}}_1$ to $\mathcal{N}_k$ and thus contains w. Therefore these two types of trees always share a common node w and cannot be disjoint:

$$\mathcal{T}(\{l, i_1\}, \{i_2, z\}) = \emptyset.$$

Similarly

$$\mathcal{T}(\{l, i_2\}, \{i_1, z\}) = \emptyset.$$

As a result,

$$\begin{aligned}
&\mathcal{T}(\{l, i_1\}, \{z\}) \\
&= \mathcal{T}(\{l, i_1, i_2\}, \{z\}) \sqcup \mathcal{T}(\{l, i_1\}, \{i_2, z\}) \\
&= \mathcal{T}(\{l, i_1, i_2\}, \{z\}) \\
&= \mathcal{T}(\{l, i_1, i_2\}, \{z\}) \sqcup \mathcal{T}(\{l, i_2\}, \{i_1, z\}) \\
&= \mathcal{T}(\{l, i_2\}, \{z\}).
\end{aligned}$$

□

Now since $b_k = \overline{b}_k = 0$ for all $k \in \mathcal{N}_1$, by Cramer's rule, we have

$$x_i = \overline{x}_i = \frac{\sum_{l \notin \mathcal{N}_1} (-1)^{l+i} b_l \det\left(\overline{L}^{li}\right)}{\det\left(\overline{L}\right)} \tag{13}$$

for all i.

Let $\mathcal{P}_1$ be set of the regions in $\mathcal{P}$ that are not directly connected to $\mathcal{N}_k$ without $\mathcal{N}_1$ and let $\mathcal{P}_2$ be the remaining regions. For a region $\mathcal{N}_l \in \mathcal{P}_1$, let

$$\chi(\mathcal{N}_l) := \sum_{E \in \mathcal{T}(\{\tilde{l}, i\}, \{z\})} \chi(E),$$

where $\tilde{l}$ is an arbitrary bus in $\mathcal{N}_l$. $\chi(\mathcal{N}_l)$ is well-defined by Lemma 20. This fact, together with the assumption $\sum_{j \in \mathcal{N}_l} b_j = 0$, then implies

$$\begin{aligned}
&\sum_{\tilde{l} \in \mathcal{N}_l} (-1)^{\tilde{l}+I} b_{\tilde{l}} \det\left(\overline{L}^{\tilde{l}i}\right) \\
&= \sum_{\tilde{l} \in \mathcal{N}_l} b_{\tilde{l}} \left(\sum_{E \in \mathcal{T}(\{\tilde{l}, i\}, \{z\})} \chi(E) \right) = \sum_{\tilde{l} \in \mathcal{N}_l} b_{\tilde{l}} \chi(\mathcal{N}_l) \\
&= \chi(\mathcal{N}_l) \sum_{\tilde{l} \in \mathcal{N}_l} b_{\tilde{l}} = 0.
\end{aligned}$$

As a result

$$\begin{aligned}
&\sum_{\tilde{l} \notin \mathcal{N}_1} (-1)^{\tilde{l}+i} b_{\tilde{l}} \det\left(\overline{L}^{\tilde{l}i}\right) \\
&= \sum_{\mathcal{N}_l \in \mathcal{P}_1} \sum_{\tilde{l} \in \mathcal{N}_l} (-1)^{\tilde{l}+i} b_{\tilde{l}} \det\left(\overline{L}^{\tilde{l}i}\right) \\
&\quad + \sum_{\mathcal{N}_l \in \mathcal{P}_2} \sum_{\tilde{l} \in \mathcal{N}_l} (-1)^{\tilde{l}+i} b_{\tilde{l}} \det\left(\overline{L}^{\tilde{l}i}\right) \\
&= \sum_{\mathcal{N}_l \in \mathcal{P}_2} \sum_{\tilde{l} \in \mathcal{N}_l} (-1)^{\tilde{l}+i} b_{\tilde{l}} \det\left(\overline{L}^{\tilde{l}i}\right) \\
&= \sum_{\mathcal{N}_l \in \mathcal{P}_2} \sum_{\tilde{l} \in \mathcal{N}_l} b_{\tilde{l}} \left(\sum_{E \in \mathcal{T}(\{\tilde{l}, i\}, \{z\})} \chi(E) \right),
\end{aligned}$$

which by Lemma 21 takes the same value for all $i \in \overline{\mathcal{N}}_1$. In other words, equation (13) takes the same value for all $i \in \overline{\mathcal{N}}_1$, completing the proof. □

Appendix VII
Proof of Lemma 16

Thanks to the conservation constraints (1a) in DC power flow equations, for any $\mathcal{G}' = (\mathcal{N}, \mathcal{E}') \in \mathfrak{C}$ the injection p' and the branch flow f' satisfies

$$\sum_{e \in N(j)} f'_e = \begin{cases} d, & j = G, \\ -d, & j = L, \\ 0, & \text{otherwise}, \end{cases}$$

where $N(j)$ is the set of edges incident to j. In particular, p' and f' can be considered as a single "flow" (see [36] for the rigorous definition of such "flow") from G to L with volume d. Thus by the Max-Flow-Min-Cut Theorem we know for any edge cut $E \subset \mathcal{E}'$,

$$\sum_{e \in E} |f_e| \geq d,$$

and therefore, $\Psi(\mathcal{G}') \geq d$. For the graph shown in Fig. 7(b), the sum of absolute branch flow is exactly d. The desired result then follows. □