

An Information Theoretic approach to identify Dominant Voltage Influencers for Unbalanced Distribution Systems

Sai Munikoti, *Student Member, IEEE*, Mohammad Abujubbeh, *Student Member, IEEE*, Kumarsinh Jhala, *Member, IEEE*, Balasubramaniam Natarajan, *Senior Member, IEEE*

Abstract—Smart distribution grid with multiple renewable energy sources can experience random voltage fluctuations due to variable generation, which may result in voltage violations. Traditional voltage control algorithms are inadequate to handle fast voltage variations. Therefore, new dynamic control methods are being developed that can significantly benefit from the knowledge of dominant voltage influencer (DVI) nodes. DVI nodes for a particular node of interest refer to nodes that have a relatively high impact on the voltage fluctuations at that node. Conventional power flow-based algorithms to identify DVI nodes are computationally complex, which limits their use in real-time applications. This paper proposes a novel information theoretic voltage influencing score (VIS) that quantifies the voltage influencing capacity of nodes with DERs/active loads in a three phase unbalanced distribution system. VIS is then employed to rank the nodes and identify the DVI set. VIS is derived analytically in a computationally efficient manner and its efficacy to identify DVI nodes is validated using the IEEE 37-node test system. It is shown through experiments that KL divergence and Bhattacharyya distance are effective indicators of DVI nodes with an identifying accuracy of more than 90%. The computation burden is also reduced by an order of 5, thus providing the foundation for efficient voltage control.

Index Terms—voltage influencing score (VIS), voltage influencer nodes, DERs, rank, analytical.

I. INTRODUCTION

THE power grid is evolving with the increasing integration of renewable distributed energy resources (DERs). While offering the benefits of reduced carbon footprint, DERs also impose new technical challenges. Specifically, active consumers with rooftop photovoltaics and distributed generation are expected to alter their generation and usage patterns to follow the trends of time-varying electricity prices [1]. This in turn induces frequent power variations. The uncontrolled operations of DERs under this condition leads to voltage fluctuations in the distribution system. Grid operators have to manage DERs' operations to assure the reliability of the grid. Therefore, various control algorithms have been considered for regulating power injection/consumption across various buses of the network to mitigate voltage fluctuations [2]–[4]. However, traditional methods of voltage control involving

voltage regulators, on-load tap changing transformers are slow and inadequate to deal with bi-directional power flows and fast dynamics. This necessitates the development of computationally efficient and fast dynamic voltage control algorithms that can handle the dynamics of power/voltage variation [5]–[7]. The speed of response to impending voltage issues and the computational efficiency of these dynamic algorithms rely on our ability to select the optimal set of nodes for control that have the highest influence on the system voltage profile.

To this end, we introduce the notion of dominant voltage influencer (DVI) nodes. Throughout this paper, we refer to *actor nodes* as the nodes where power changes (due to DER unit and/or load dynamics) and *observation nodes* represent the set of nodes where voltage state is monitored. Thus, DVI nodes for a particular observation node denote all those actor nodes that have a relatively high impact on the voltage state of that observation node compared to the rest of the actor nodes. Hence, altering the actions of DVI nodes result in the highest reduction in voltage issues at the observation node. The nodes could be DVI because of their size (generation/load capacity) and/or locations in the distribution network. Conventional methods of identifying such DVI nodes involve Monte-Carlo simulations using load flow algorithms. These classical methods possess various drawbacks including (1) high computational complexity, (2) numerical results with no analytical insights, and (3) scenario dependent results with no generalization. These factors limit the applicability of conventional approaches in modern distribution systems. Thus, there is a need to develop an analytical and computationally efficient algorithm for identifying DVI nodes in three phase unbalanced distribution systems, which in turn can be used to develop effective voltage control algorithms.

For identifying DVI nodes, we need to quantify the impact of actor nodes on other nodes of the network in terms of voltage. Therefore, first, we develop a novel metric namely voltage influence score (VIS) for quantifying the voltage influencing capacity of an actor node on any arbitrary observation node. The VIS is based on an analytical method of voltage sensitivity. Thereafter, the VIS is employed to rank the actor nodes for any arbitrary observation node and identify the most influential ones, i.e., DVI nodes. The proposed approach allows us to identify DVI nodes without relying on computationally expensive Monte-Carlo simulations thereby, significantly reducing the computation time.

Knowing the DVI set, one can focus on controlling the

K. Jhala is with the Center for Energy, Environmental, and Economic Systems Analysis in the Energy Systems Division at Argonne National Laboratory (e-mail: kjhala@anl.gov).

S. Munikoti, M. Abujubbeh and B. Natarajan are with Electrical and Computer Engineering, Kansas State University, Manhattan, KS-66506, USA, (e-mail: saimunikoti@ksu.edu, abujubbeh@ksu.edu, bala@ksu.edu)

dominant nodes within a cluster for ensuring the stability of voltage states for all nodes within that cluster. This is beneficial in the sense that minimal control actions would be required compared to any local control scheme such as [8]. Furthermore, the DVI set helps to efficiently assign penalties to controllable nodes through a weight matrix. For example, assume a cluster in an unbalanced distribution system with N nodes. In a nutshell, power injections (ΔS) at the DVI set $n \in N$ can be controlled over the control horizon T by ¹,

$$\text{Min} \sum_{t \in T} \left(S^T R \Delta S \right)$$

where, $R = R^T \succ 0$ is a diagonal weight matrix on control policies. That is, any control actions could have a cost associated with it. Elements that belong to DVI nodes, i.e., $R_{n,n}$, can have lower penalties compared to non-DVI nodes within the cluster. Additionally, there could be a scenario where DVI nodes are not controllable. However, the voltage influencing scores offer greater flexibility in choosing the nodes even if the top dominant nodes are not controllable. Moreover, in the application of network partitioning, non-controllable DVI nodes can also be valuable in defining clusters and thus indirectly participate in voltage control.

Contributions: This paper proposes a novel approach to identify DVI nodes that can be used for network clustering and distributed control mechanisms in unbalanced distribution systems with DERs. The key contributions of this paper include:

- An analytical and computationally efficient method of identifying dominant voltage influencer nodes is proposed which is valid for both balanced and unbalanced distribution system.
- This work introduces a VIS metric that quantifies the voltage influencing capacity of nodes with DERs/active loads, and a computationally efficient method to compute it.
- Compared to [9], this work has three major advancements (1) the probabilistic model of voltage fluctuations is derived for a three phase unbalanced distribution systems; (2) The voltage influencing capacity is quantified for each pair of nodes; (3) In addition to identifying the most dominant node, this work provides ranks of all dominant influencer nodes based on their VIS;
- Reference [10] is our foundational work on deriving an analytical form of voltage change distribution. The VIS metric proposed in this work is a function of the information-theoretic distance between analytical voltage change distributions.
- The effectiveness of proposed method is evaluated on the standard unbalanced IEEE 37-node test system yielding an accuracy of more than 90% with the computation complexity reduced by an order of 5 compared to a classic load flow based approach.

¹Problem constraints and exact formulations are omitted for the sake of simplicity. The aim is to illustrate the usefulness of DVI nodes in efficiently penalizing non-DVI nodes.

The paper is organized as follows. Section II presents a conventional method to identify DVI nodes followed by an analytical method of voltage sensitivity analysis. Section III computes probability distributions of voltage change. Various information theoretic indicators of the DVI nodes are proposed in Section III, which are used to compute VIS. The performance of the proposed approach is tested using the IEEE 37-node test system in Section IV. Conclusions and future-work are presented in Section V.

II. BACKGROUND

This section introduces the conventional approach to identify Dominant voltage influencer (DVI) nodes in a power distribution network. Then, the fundamental analytical expressions of voltage sensitivity are discussed which are later used to model voltage fluctuations for identifying DVI nodes.

A. Related work

Various control algorithms are proposed in the literature for regulating voltage fluctuations in the distribution system. Primarily, the control algorithms can be broadly categorized into two main categories: (1) centralized [2], [3], [11], [12], and (2) decentralized [4], [13], [14] control approaches. Centralized voltage control algorithms are generally based on classical power flow solution, which can be computationally expensive for real time applications. For instance, authors in [11] provide a control algorithm using dominant generator nodes. However [11] only considers deterministic power change scenarios and does not provide any quantification of voltage influence while determining dominant generators. Similarly, authors in [15] propose a new AC optimal power flow formulation that efficiently regulates voltage in distribution systems. It has been shown that the proposed method results in minimal curtailment of real and reactive power while regulating voltages. On the other hand, decentralized voltage control methods trade-off performance for reduced computational complexity. The speed, complexity and performance of both centralized and decentralized algorithms rely on the size and space of control actions. In modern distribution systems, the change in power consumption or injections due to multiple active consumers necessitates the need for analyzing a large number of scenarios, expanding the size and space for determining optimal control actions. This may further increase the computational complexity of traditional voltage control algorithms. Other approaches to bring voltage to safe operational limits in distribution systems involve the installation of additional control devices such as, static-var compensators [16], [17], dynamic voltage restorers [18], and grid-edge volt-var controllers [19]. Despite their effectiveness in voltage control, the installation of additional devices in the system can be expensive. Therefore, a practical and economical solution to regulate voltage fluctuations is to utilize smart grid edge devices like PV inverters. To this end, multiple dynamic algorithms have been proposed to control inverters and loads to regulate voltages [5]–[7]. In [5], authors randomly select and shut down the thermostatically controlled loads during the time of high voltage unbalance. Similarly, [7] proposes a data-driven voltage control algorithm to regulate voltage violations

by minimizing power injections at nodes with over voltage violations. In a nutshell, all these prior works use heuristic based methods to identify nodes whose generations need to be curtailed. However, such irrational selection is not desirable in terms of yielding optimal solution and providing fairness to all generators participating in a transactive energy market.

Furthermore, there are some efforts where system is partitioned into smaller clusters and control is implemented for each cluster in a distributed manner [8], [20]–[22]. The concept of virtual power plants is also shown to help voltage regulation in distribution systems with the presence of DERs and battery storage systems (BSS) [23]. In [20], [21], for instance, authors propose a new clustering method based on voltage sensitivity with respect to power fluctuation due to BSS. However, this method is scenario specific since the location of BSS in the system is fixed. Another sensitivity based zonal voltage control method is proposed in [22] where all nodes in a particular cluster are assumed to participate in power curtailment for voltage control. This is difficult to implement as it is not economical or even feasible for utilities to have access to control all nodes. There are some approaches which also include topological data for system clustering [8], [24], [25]. For the approach in [8] to work, the nodal Q-V sensitivities must be computed to iteratively evaluate the modality index for each cluster [26]. Nevertheless, the computational complexity of this method increases with the size of the system as computing the Q-V based modality index of each cluster is based on forming the adjacency matrix of the system. Additionally, the entire body of prior work on identifying dominant nodes of voltage fluctuations does not consider stochasticity associated with DER power injections or user load variations.

To address these demerits of existing approaches, analytical and computationally efficient approaches to identify DVI nodes that have relatively high influence on the voltage state of critical nodes are needed. The work in this paper addresses this essential research gap using a new information theoretic voltage influencing index that is applicable to three-phase unbalanced systems.

B. Conventional approach to identify DVI nodes

The DVI nodes for an observation node are the nodes that have high impact on the voltage fluctuations at the observation node. An actor node can be a DVI due to association of two factors: (1) location of the actor node, i.e. phase and bus of the distribution network, and (2) generation/load capacity of DER/loads connected at the actor node which influences the variance of power change at that node. Generally, simulation-based scenario analysis is used as a major planning tool to identify DVI nodes. A typical approach involves following steps [9]:

1) *Step 1- Compute variance of voltage change at each phase of observation node due to all actor nodes:* The variance of voltage change at each phase of the observation node is computed by running multiple power flow based Monte-Carlo simulations with varying power, which captures temporal variation of generations.

2) *Step 2- Calculate reduction in variance of voltage change at the observation node due to each actor node by setting power drawn/injected by the actor node as zero:* This step requires repetition of Step 1 for each actor node after setting the variance of actor node as zero.

3) *Step 3- Rank actor nodes based on the reduction in variance caused by the removal of the corresponding actor node::* Actor nodes are ranked in an ascending order with topmost and bottommost rank assigned to those actor node that causes maximum and minimum reduction in variance of voltage change at the observation node, respectively.

This scenario-based method incurs high computational complexity, which grows with the size of the network. Specifically, the complexity arises at two hierarchical levels. At the core of this approach, we have load flow runs, whose computation using NR-based method is of complexity $O(n^3)$ where n represents the size of the distribution network [10]. At a high level, multiple power change scenarios are simulated to obtain voltage change distributions for steps 1 and 2 of the approach. Each such scenario involves a load flow computation, thus one ends up running thousands of load flows. In our experiment, a hundred thousand scenarios are simulated to compute voltage change distribution for each actor node at step 2. The computational effort is also demonstrated via execution times in the Results section. Therefore, to counter computational burden, various information-theoretic metrics are explored in the next section that can identify DVI nodes in a computationally efficient manner and can be used as a tool in various power system operations.

Algorithm 1 Proposed method to identify DVI nodes

Input: Distribution Network with the target observation node.

Output: Ranked actor nodes for the input observation node.

- 1: Compute distribution of voltage change at the input observation node due to each actor node at a time.
 - 2: Compute distribution of voltage change at the input observation node due to aggregate effect of all actor nodes.
 - 3: Find statistical distance between the distributions computed in steps 1 & 2.
 - 4: Compute VIS for each pair of observation and actor nodes.
 - 5: Rank actor nodes in the ascending order based on their respective VIS.
 - 6: **return** ranks and VIS of actor nodes for an input observation node.
-

C. Analytical VSA

In a three phase distribution system, change in power at any one phase of the node causes change in voltage at all phases of all the other nodes. Traditional methods of voltage sensitivity analysis (VSA) are computationally complex and less generic. Therefore, we developed a computationally efficient method of VSA in [10], [27]. For an unbalanced power distribution system, change in complex voltage ΔV_O at an observation

node (O) due to change in complex power at multiple actor nodes can be approximated as:

$$\begin{bmatrix} \Delta V_O^a \\ \Delta V_O^b \\ \Delta V_O^c \end{bmatrix} \approx - \sum_{A \in \tilde{A}} \begin{bmatrix} \frac{\Delta S_A^{a*} Z_{OA}^{aa}}{V_A^{a*}} + \frac{\Delta S_A^{b*} Z_{OA}^{ab}}{V_A^{b*}} + \frac{\Delta S_A^{c*} Z_{OA}^{ac}}{V_A^{c*}} \\ \frac{\Delta S_A^{a*} Z_{OA}^{ba}}{V_A^{a*}} + \frac{\Delta S_A^{b*} Z_{OA}^{bb}}{V_A^{b*}} + \frac{\Delta S_A^{c*} Z_{OA}^{bc}}{V_A^{c*}} \\ \frac{\Delta S_A^{a*} Z_{OA}^{ca}}{V_A^{a*}} + \frac{\Delta S_A^{b*} Z_{OA}^{cb}}{V_A^{b*}} + \frac{\Delta S_A^{c*} Z_{OA}^{cc}}{V_A^{c*}} \end{bmatrix} \quad (1)$$

where a, b and c represent the three phases. V_A^a and ΔS_A^a represent complex voltage and power changes at the phase a of the actor node A , respectively; Z denotes the impedance matrix including self and mutual line impedance of the shared path between observation node and actor node from the source node. $\tilde{A}$ is the set of all actor nodes. Eqn. (1) follows superposition law and voltage change is aggregated due to the power change of each actor nodes. For a single actor node, the voltage change at any phase (say phase a) of an observation node O can be decomposed into real ($\Delta V_{OA}^{a,r}$) and imaginary components ($\Delta V_{OA}^{a,i}$) as:

$$\Delta V_{OA}^{a,r} \approx \left[\left(\frac{(\Delta P_A^a R_{OA}^{aa} + \Delta Q_A^a X_{OA}^{aa})(V_A^{a,r})}{(V_A^{a,r})^2 + (V_A^{a,i})^2} + \frac{(\Delta P_A^a X_{OA}^{aa} - \Delta Q_A^a R_{OA}^{aa})(V_A^{a,i})}{(V_A^{a,r})^2 + (V_A^{a,i})^2} - \dots \right) \right] \quad (2)$$

$$\Delta V_{OA}^{a,i} \approx \left[\left(\frac{(\Delta P_A^a X_{OA}^{aa} - \Delta Q_A^a R_{OA}^{aa})(V_A^{a,r})}{(V_A^{a,r})^2 + (V_A^{a,i})^2} - \frac{(\Delta P_A^a R_{OA}^{aa} + \Delta Q_A^a X_{OA}^{aa})(V_A^{a,i})}{(V_A^{a,r})^2 + (V_A^{a,i})^2} - \dots \right) \right] \quad (3)$$

where ΔP_A^a and ΔQ_A^a are the real and reactive power change at phase- a of the actor node A , R_{OA}^{aa} and X_{OA}^{aa} are resistance and reactance of shared line between observation node O and actor node A . It is proved by Corollary 2 in our fundamental work that the error in our approximation of voltage change (i.e., eqns. (2) and (3)) is always upper bounded [10]. Hence, this approximation scales very accurately with the size of the network and can be leveraged for deriving voltage change distributions, and consequently identifying the DVI nodes in a three-phase unbalanced distribution system.

III. PROPOSED FRAMEWORK

As can be seen from the previous section, the conventional method of identifying DVI nodes is computationally complex, which limits its use in real-time operation of power system. Therefore, we propose the use of information-theoretic distance metrics as potential indicators of DVI nodes. Fundamentally, the proposed approach consists of four steps. In the first step, we obtain probability distributions of voltage change at an observation node due to each actor node as well as due to the aggregate presence of all actor nodes. These distributions are derived in a computationally efficient way by employing the analytical expressions discussed in the

background section. In the second step, we compute distances between the probability distribution due to each actor node and the distribution due to the aggregate effect of all actor nodes. The Third step utilizes the distances to compute voltage influencing score (VIS) for each pair of observation and actor nodes. Finally, for each observation node, all the actor nodes are ranked based on the computed VIS. The actor node whose voltage change distribution is nearest to the aggregate voltage change distribution is deduced as the major influencer of voltage change for that particular observation node. Likewise, all actor nodes are ranked based on the ascending order of their distances. The complete procedure is summarized in the Algorithm 1.

A. Probabilistic model of voltage fluctuations

This section provides a probabilistic model of voltage change in a three-phase unbalanced distribution system, which consist of multiple spatially distributed actor nodes with PVs and active consumers. Random change in power at actor nodes due to intermittent renewable generation causes random voltage fluctuations. Therefore, probability distribution is needed to quantify voltage change under such stochastic scenarios. Here, we implement the first step of the proposed approach, i.e., derive the probability distributions of voltage change at any observation node due to random power change at a single actor node as well as due to the aggregate effect of all actor nodes. Let ΔS_A^a be change in complex power at phase a of the actor node A . Then, using eqn. (1), the voltage change at phase- a of an observation node O can be expressed as:

$$\Delta V_{OA}^a = \Delta V_{OA}^{a,r} + j \Delta V_{OA}^{a,i}, \quad (4)$$

$$\begin{aligned} \Delta V_{OA}^{a,r} &= \sum_{h,u} \frac{-1}{|V_A^h|} [\Delta P_A^h (R_{OA}^u \cos(\omega_A) - X_{OA}^u \sin(\omega_A)) \\ &\quad + \Delta Q_A^h (R_{OA}^u \sin(\omega_A) + X_{OA}^u \cos(\omega_A))], \\ \Delta V_{OA}^{a,i} &= \sum_{h,u} \frac{-1}{|V_A^h|} [\Delta P_A^h (R_{OA}^u \sin(\omega_A) + X_{OA}^u \cos(\omega_A)) + \\ &\quad \Delta Q_A^h (X_{OA}^u \sin(\omega_A) - R_{OA}^u \cos(\omega_A))] \end{aligned} \quad (5)$$

where $h \in \tilde{H}$ and $u \in \tilde{U}$. The sets $\tilde{H}$ and $\tilde{U}$ denote different phases (i.e., a, b, c) and different phase sequences (i.e., aa, ab, ac), respectively. ΔP_A^h and ΔQ_A^h are the active and reactive power changes, respectively. R_{OA}^h, X_{OA}^h are the resistance and reactance of shared path between the observation node O and actor node A from the source node. ω_A is the impedance angle for phase- a . V_A^h denotes the base voltage of actor node A . For brevity, the derivation is shown for phase a . However, the same steps can be followed with the corresponding phase quantities to derive expressions for other phases. The real and imaginary parts of voltage change can

further be simplified as,

$$\begin{aligned}\Delta V_{OA}^{a,r} &= \left[\left(C_{OA}^{aa,r} \Delta P_A^a + D_{OA}^{aa,r} \Delta Q_A^a + C_{OA}^{ab,r} \Delta P_b^b + \right. \right. \\ &\quad \left. \left. D_{OA}^{ab,r} \Delta Q_b^b + C_{OA}^{ac,r} \Delta P_c^c + D_{OA}^{ac,r} \Delta Q_c^c \right) \right] \left(\right. \\ \Delta V_{OA}^{a,i} &= \left[\left(C_{OA}^{aa,i} \Delta P_A^a + D_{OA}^{aa,i} \Delta Q_A^a + C_{OA}^{ab,i} \Delta P_b^b + \right. \right. \\ &\quad \left. \left. D_{OA}^{ab,i} \Delta Q_b^b + C_{OA}^{ac,i} \Delta P_c^c + D_{OA}^{ac,i} \Delta Q_c^c \right) \right] \left(\right.\end{aligned} \quad (6)$$

Eqn. (6) is further written in compact form as,

$$\begin{aligned}\Delta V_{OA}^{a,r} &= \mathbf{C}_{OA}^{a,rT} \Delta \mathbf{S}_A, \\ \Delta V_{OA}^{a,i} &= \mathbf{C}_{OA}^{a,iT} \Delta \mathbf{S}_A,\end{aligned} \quad (7)$$

where,

$$\begin{aligned}\mathbf{C}_{OA}^{a,rT} &= [C_{OA}^{aa,r} \ D_{OA}^{aa,r} \ C_{OA}^{ab,r} \ D_{OA}^{ab,r} \ C_{OA}^{ac,r} \ D_{OA}^{ac,r}] \\ \mathbf{C}_{OA}^{a,iT} &= [C_{OA}^{aa,i} \ D_{OA}^{aa,i} \ C_{OA}^{ab,i} \ D_{OA}^{ab,i} \ C_{OA}^{ac,i} \ D_{OA}^{ac,i}]\end{aligned}$$

The constants C and D are the functions of the line impedances and base voltages as explained in eqn. (4). For brevity, the exact expressions are omitted from here and are provided in the Appendix A. Then, the power change vector can be written as,

$$\Delta \mathbf{S}_A = [\Delta P_A^a, \Delta Q_A^a, \Delta P_A^b, \Delta Q_A^b, \Delta P_A^c, \Delta Q_A^c]^T$$

Thus, because of random power changes at each actor node, the power change vector $\Delta \mathbf{S}_A$ is a random vector with mean vector $\mu_{\Delta \mathbf{S}_A}$ and covariance matrix $\sum_{\Delta \mathbf{S}_A}$. The covariance matrix $\sum_{\Delta \mathbf{S}_A}$ quantifies the correlation of power changes among various phases of a particular actor node A . The diagonal elements denote variances of power change at each phase and off-diagonal elements contain the correlation between the power changes. Thus, it can be seen from (22) that the voltage change at an observation node O due to actor node A is the weighted combination of random vector $\Delta \mathbf{S}_A$, where weights (i.e., $\mathbf{C}_{OA}^{a,rT}$ and $\mathbf{C}_{OA}^{a,iT}$) are constant terms. Invoking the Lindeberg-Feller central limit theorem, it can be shown that the $\Delta V_{OA}^{a,r}$ and $\Delta V_{OA}^{a,i}$ converges in distribution to a Gaussian random variables, i.e.,

$$\begin{aligned}\Delta V_{OA}^{a,r} &\xrightarrow{D} \mathcal{N}(\mu_{OA}^{a,r} = \mathbf{C}_{OA}^{a,rT} \mu_{\Delta \mathbf{S}_A}, \sigma_{OA}^{a,r2} = \mathbf{C}_{OA}^{a,rT} \sum_{\Delta \mathbf{S}_A} \mathbf{C}_{OA}^{a,r}) \\ \Delta V_{OA}^{a,i} &\xrightarrow{D} \mathcal{N}(\mu_{OA}^{a,i} = \mathbf{C}_{OA}^{a,iT} \mu_{\Delta \mathbf{S}_A}, \sigma_{OA}^{a,i2} = \mathbf{C}_{OA}^{a,iT} \sum_{\Delta \mathbf{S}_A} \mathbf{C}_{OA}^{a,i}),\end{aligned} \quad (8)$$

where, $\mu_{OA}^{a,r}$ and $\mu_{OA}^{a,i}$ are the mean vectors of real and imaginary parts of voltage change, respectively. $\sigma_{OA}^{a,r2}$ and $\sigma_{OA}^{a,i2}$ are the variances of real and imaginary parts of voltage change, respectively. For investigating the relationship between the real and imaginary parts of the voltage change, a new bivariate random vector is defined,

$$\begin{bmatrix} \Delta V_{OA}^{a,r} \\ \Delta V_{OA}^{a,i} \end{bmatrix} \sim \mathcal{N} \left[\begin{bmatrix} \mu_{OA}^{a,r} \\ \mu_{OA}^{a,i} \end{bmatrix}, \begin{bmatrix} \sigma_{OA}^{a,r2} & k_{OA}^a \\ k_{OA}^a & \sigma_{OA}^{a,i2} \end{bmatrix} \right], \quad (9)$$

where, $k_{OA}^a = \mathbf{C}_{OA}^{a,rT} \sum_{\Delta \mathbf{S}_A} \mathbf{C}_{OA}^{a,i}$ is the covariance between the real and imaginary parts of the voltage change due to single actor node A . Eqn. (9) provides the probability distribution of voltage change at a particular observation node due to single actor node A . A similar approach can be used to compute individual voltage change distributions due to each actor node in the network. Now, we need to obtain the voltage change distribution due to the aggregate effect of all actor nodes. Using the superposition property in (1), the voltage change due to cumulative effect of power changes at multiple actor nodes can be expressed as [10],

$$\Delta V_O^{a,r} = \sum_{A=1}^L \Delta V_{OA}^{a,r}, \quad (10)$$

where L is the number of actor nodes. By leveraging (22), the net voltage change can be written as,

$$\begin{aligned}\Delta V_O^{a,r} &= \sum_{A=1}^N \left(\mathbf{C}_{OA}^{a,rT} \Delta \mathbf{S}_A \right)_{6 \times 1} \\ &= [\mathbf{C}_{O1}^{a,rT} \ \mathbf{C}_{O2}^{a,rT} \ \dots \ \mathbf{C}_{ON}^{a,rT}] [\Delta \mathbf{S}_1 \ \Delta \mathbf{S}_2 \ \dots \ \Delta \mathbf{S}_N]^T \\ &= \mathbf{C}_O^{a,rT} \Delta \mathbf{S}_{1 \times 6N}^{6N \times 1}\end{aligned} \quad (11)$$

where $\mathbf{C}_O^{a,rT}$ and $\Delta \mathbf{S}$ are the long vectors, composed of a constant term and the power change vector corresponding to each actor node, respectively. Similarly, the imaginary part of voltage change can be written as,

$$\Delta V_O^{a,i} = \sum_{A=1}^N \left(\mathbf{C}_{OA}^{a,iT} \Delta \mathbf{S}_A \right)_{6 \times 1} = \mathbf{C}_O^{a,iT} \Delta \mathbf{S}_{1 \times 6N}^{6N \times 1} \quad (12)$$

The equations (11) and (12) possess a similar form as that of (22), i.e., the net voltage change is the weighted combination of power change vector $\Delta \mathbf{S}$. Here, weight $\mathbf{C}_O^{a,r}$ is a constant vector comprising of line impedances and node base voltages, whereas, $\Delta \mathbf{S}$ comprises of power change at all phases of all actor nodes. Now, invoking the same Lindeberg-Feller central limit theorem, the real $\Delta V_O^{a,r}$ and imaginary $\Delta V_O^{a,i}$ part of aggregate voltage change can be shown to converge in distribution to a Gaussian random variables with the following parameters:

$$\begin{aligned}\Delta V_O^{a,r} &\xrightarrow{D} \mathcal{N}(\mu_O^{a,r} = \mathbf{C}_O^{a,rT} \mu_{\Delta \mathbf{S}}, \sigma_O^{a,r2} = \mathbf{C}_O^{a,rT} \sum_{\Delta \mathbf{S}} \mathbf{C}_O^{a,r}) \\ \Delta V_O^{a,i} &\xrightarrow{D} \mathcal{N}(\mu_O^{a,i} = \mathbf{C}_O^{a,iT} \mu_{\Delta \mathbf{S}}, \sigma_O^{a,i2} = \mathbf{C}_O^{a,iT} \sum_{\Delta \mathbf{S}} \mathbf{C}_O^{a,i})\end{aligned} \quad (13)$$

Similar to the single actor node case, the correlations between the real and imaginary parts of net voltage change is captured by defining a new random vector as:

$$\begin{bmatrix} \Delta V_O^{a,r} \\ \Delta V_O^{a,i} \end{bmatrix} \sim \mathcal{N} \left[\begin{bmatrix} \mu_O^{a,r} \\ \mu_O^{a,i} \end{bmatrix}, \begin{bmatrix} \sigma_O^{a,r2} & k_O^a \\ k_O^a & \sigma_O^{a,i2} \end{bmatrix} \right] \quad (14)$$

where, $k_O^a = \mathbf{C}_O^{a,rT} \sum_{\Delta \mathbf{S}} \mathbf{C}_O^{a,i}$ is the covariance between the real and imaginary parts of net voltage change. Equations (9) and (14) provide the probability distribution due to single actor

node and the aggregation of multiple actor nodes, respectively. It is worth noting that the voltage change expressions in equation (13) are valid for any number of actor nodes. Further, the proposed method to determine voltage change is sufficiently general for PV generation with any arbitrary distribution. This is because the derivation of the voltage change utilizes standard statistical results that are valid for any probability distribution. Precisely, eqn. (12) computes voltage change in the real part of phase a ($\Delta V_O^{a,r}$) as the weighted sum of the power changes in the actor nodes. Here, even if each element of ΔS representing PV generation changes is not Gaussian, we can still invoke Lindeberg-Feller central limit theorem [4], to show that the weighted sum of the elements of ΔS converges in distribution to a Gaussian. Therefore, the distributions of the real and imaginary parts of voltage change will always converge to a normal which eventually leads to a Nakagami distribution for the magnitude of voltage change. The next subsection focuses on computing the statistical distances between these distributions and presents the procedure to rank the actor nodes.

B. Information theoretic metrics as DVI indicators

This sub-section implements the second step of our proposed approach, i.e., calculate statistical distances between the probability distributions (derived in earlier sub-section), and rank the actor nodes based on the computed distances. The information theoretic distance metrics which are potential indicators of DVI nodes are defined first.

1) *Kullback-Liebler distance*: Kullback-Liebler (KL) distance quantifies how much one probability distribution differs from another probability distribution. KL divergence between two multivariate Gaussian distributions ($\mathcal{N}_0$ and $\mathcal{N}_1$) of dimension k with means (μ_0 and μ_1) and covariance matrices (Σ_0 and Σ_1) can be written as:

$$D_{KL}(\mathcal{N}_0||\mathcal{N}_1) = \frac{1}{2} \left[\text{tr}(\Sigma_1^{-1} \Sigma_0) + (\mu_1 - \mu_0)^T \Sigma_1^{-1} (\mu_1 - \mu_0) - k + \ln \frac{|\Sigma_1|}{|\Sigma_0|} \right] \quad (15)$$

where $\text{tr}(\cdot)$ indicates trace of the matrix. Here, the KL distance between the distributions of voltage change at the observation node due to change in power at an actor node A (ΔV_{OA}) and due to change in power at all actor nodes (ΔV_O) is given by

$$D_{KL}(\Delta V_{OA}||\Delta V_O) = \frac{1}{2} \left[\text{tr}(\Sigma_{\Delta V_O}^{-1} \Sigma_{\Delta V_{OA}}) + (\mu_O^a - \mu_{OA}^a)^T \Sigma_1^{-1} (\mu_O^a - \mu_{OA}^a) - 2 + \ln \frac{|\Sigma_{\Delta V_O}|}{|\Sigma_{\Delta V_{OA}}|} \right] \quad (16)$$

where $\Sigma_{\Delta V_{OA}}$ and $\Sigma_{\Delta V_O}$ are the covariances of ΔV_{OA} and ΔV_O , respectively.

2) *Bhattacharyya distance*: Bhattacharyya (BC) distance measures the similarity of two probability distributions. It is related to the Bhattacharyya coefficient which is a measure of the amount of overlap between two statistical samples. BC

distance between the distributions of voltage change at the observation node due to change in power at an actor node A (ΔV_{OA}) and due to change in power at all actor nodes (ΔV_O) can be expressed as:

$$D_{BC}(\Delta V_{OA}||\Delta V_O) = \frac{1}{8} (\mu_O^a - \mu_{OA}^a)^T \Sigma \left(\mu_O^a - \mu_{OA}^a \right) + \frac{1}{2} \ln \left(\frac{|\Sigma|}{\sqrt{|\Sigma_{\Delta V_O}| |\Sigma_{\Delta V_{OA}}|}} \right) \quad (17)$$

where $\Sigma = \frac{\Sigma_{\Delta V_O} + \Sigma_{\Delta V_{OA}}}{2}$

C. Voltage Influencing score (VIS)

The influence of an actor node on an observation node needs to be quantified for identifying its rank among all actor nodes. In this regard, we devise a novel index to quantify voltage influencing capacity, i.e., Voltage influencing score (VIS). For a given network scenario, i.e., the location of actor nodes with the variance of power change from historical data, the distances between voltage change distributions at an observation node due to each actor node and due to aggregate effect of all actor nodes are used to rank the actor nodes. Here, these distances are employed to compute the VIS between any pair of observation and actor node as:

$$VIS(O, A) = \frac{\frac{1}{D(A, O)} - \frac{1}{D(S, O)}}{\frac{1}{D(A', O)} - \frac{1}{D(S, O)}}, \quad (18)$$

where $D(A, O)$ is the statistical distance between the voltage change distribution at an observation node O due to aggregate effect of all actor nodes and when actor node A is solely present in the system. $D(S, O)$ is the statistical distance between source node and observation node. The distance can be computed with any of the metrics described in the earlier subsection. For an observation node O , the lower the distance, the more the actor node A contributes to aggregate voltage change and consequently the more influencing the actor node is and vice-versa. Therefore, the VIS is expressed in terms of inverse of distance. To provide an absolute sense to the score, VIS is normalized with minimum and maximum values. For a particular observation node, the ideal location of actor node A' for minimum distance would be the same observation node location. On the other hand, the maximum distance location would always be the source node as it has minimum influence on voltage fluctuations of any observation node. These minimum and maximum distances are used to normalize VIS as shown in eqn. (18).

In the final step of the proposed approach, VIS is leveraged for ranking the actor nodes. The nodes are ranked in an ascending order based on VIS. The topmost (Rank 1) and bottom-most (Rank L for the case of L actor nodes) ranks will be assigned to the actor nodes in the following way:

$$\begin{aligned} \text{Rank } 1 &: \arg \max_A VIS(O, A) \\ \text{Rank } L &: \arg \min_A VIS(O, A) \end{aligned} \quad (19)$$

It is worth to note that the proposed approach can work with any other information theoretic metric (Frechet distance,

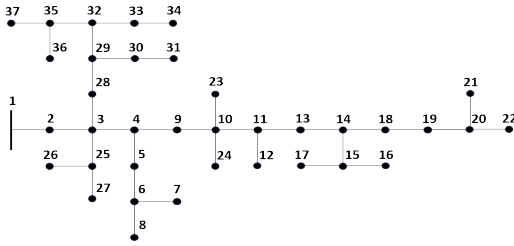


Fig. 1: Modified IEEE 37-node network

Jensen-Shannon distance, etc.), and for illustration purposes, the method is evaluated with two metrics (KL divergence & Bhattacharyya distance) in the results section.

IV. RESULTS AND DISCUSSION

The efficacy of the proposed method in identifying dominant voltage influencer nodes is evaluated in this section. The baseline method is the conventional Monte-Carlo simulation-based approach, and both algorithms are implemented in the modified IEEE 37-node test system as shown in figure 1. This test network is chosen as it denotes a typical unbalanced distribution system and is used by many researchers for illustrating the efficiency of their methods [10], [27]. The nominal voltage of the test system is 4.8 kV. A scenario is generated with 15 actor nodes distributed randomly in the IEEE 37-node network. Change in real and reactive power at 15 actor nodes is modeled as zero-mean Gaussian random vector. Generally, Gaussian distribution is used to validate statistical frameworks, and it has been considered as a common assumption in many prior works related to distribution systems [7], [28]. Moreover, the probability distribution of power/load changes in a power distribution system including PV/wind power changes have also been shown to sometimes follow a Beta distribution [29], [30]. Therefore, the proposed approach is validated for both Gaussian and Beta distribution. It is worth noting that the proposed approach is generic for any choice of actor node and power change distributions, and the simulated case-studies are merely a way to illustrate its performance. Three different PV sizes are considered in this case study. The mean and variance of real and reactive power change of all three kinds of PV capacities along with their location and phase information are as follows:

$$\begin{aligned} \Delta S_A &\sim \mathcal{N} \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1.5 & -0.05 \\ -0.05 & 0.25 \end{bmatrix} \right) \left(\begin{array}{l} A \in \{7^c, 8^a, 12^c, 27^b, 28^c\} \\ \Delta S_A \sim \mathcal{N} \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 & -0.1 \\ -0.1 & 0.5 \end{bmatrix} \right) \left(\begin{array}{l} A \in \{7^b, 18^a, 22^c, 34^b, 36^b\} \\ \Delta S_A \sim \mathcal{N} \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 4.5 & -0.2 \\ -0.2 & 0.75 \end{bmatrix} \right) \left(\begin{array}{l} A \in \{9^c, 14^c, 26^c, 30^a, 31^a\} \end{array} \right) \end{array} \right. \end{aligned} \quad (20)$$

where superscript over actor nodes, i.e., $\{a, b, c\}$ represent respective phases of actor nodes at which power is varying.

The change in power across different actor nodes can be correlated because of environmental factors. The DERs such as PVs and wind turbines are expected to exhibit a similar generation profile due to their geographical proximity. Further, the real and reactive power of inverter-based DERs is negatively correlated. The underlying covariance structure $\sum \Delta S$ can be estimated based on historical or irradiance-related data. The base loads on the test network are the same as mentioned in the IEEE PES Distribution system analysis subcommittee report. For actor nodes, the variance of change in real power and reactive power are present in the diagonal elements of covariance matrices as shown in eqn. (20). Thus, three approximate PV sizes which are simulated in our experiments are 45 kW, 38 kW and 35 kW. The off-diagonal elements of the covariance matrices capture the covariance between real and reactive power. These covariance of each actor nodes are combined to form the covariance matrix of power change vector $\sum \Delta S$ which corresponds to all actor nodes. Here, the correlation coefficient between ΔP 's and ΔQ 's for different actor nodes within the same phase is kept same and covariance between cross-phase terms is assumed to be zero. However, the proposed approach is quite general to accommodate other covariance structures as well. The variance of nodes other than the actor nodes is set to zero.

Table I: Observation node 7 Table II: Observation node 16

Node	MC	KL	BC	Node	MC	KL	BC
7	1	1	1	14	1	1	1
9	0.577	0.163	0.129	22	0.919	0.840	0.808
12	0.543	0.150	0.117	18	0.782	0.170	0.133
22	0.530	0.127	0.096	12	0.613	0.576	0.533
14	0.516	0.137	0.106	9	0.426	0.422	0.373
26	0.441	0.092	0.067	7	0.209	0.290	0.242
28	0.350	0.084	0.060	17	0.146	0.230	0.186
8	0.042	0.065	0.045	8	0.020	0.111	0.083
17	0.026	0.050	0.033	26	0.001	0.176	0.138
27	0.001	0.043	0.029	28	0.001	0.164	0.128

A. VIS for ranking DVI nodes

To assess the performance of the proposed approach in identifying DVI nodes, VIS for two arbitrary observation nodes (i.e., 7, 16) are computed using two different distance metrics namely KL divergence (eqn. (15)) and Bhattacharyya distance (eqn. (16)). The results obtained by using the proposed approach is validated against ground truth values. The baseline approach to compute true DVI nodes is based on Monte-Carlo simulations (MCS) of load flow as explained in Section II-A. Initially, the variance of voltage change magnitude at a given observation node is computed due to the presence of all actor nodes. Multiple power change scenarios are simulated by running 100000 MCS with load flow method. Then, the reduction in the variance of the magnitude of voltage change is determined by setting the variance to zero for each actor node sequentially. Again, for each actor node case, different MCS are executed. Finally, the actor nodes are ranked based on the reduction in variance they brought when the variance of power change at the corresponding actor node is set to zero.

Tables I and II tabulates the VIS for observation nodes 7 and 16, respectively. MC, KL and BC refers to baseline MCS approach, proposed KL divergence and Bhattacharyya distance

based methods, respectively. Node in the tables refers to actor node. It can be inferred that the VIS for the most dominant actor node is 1 and it decreases as we move to lower rank nodes. For observation node-7 with KL distance metric, there is a small difference in VIS between nodes 22 and 14 implying that they exhibit almost equal voltage influencing capacity. However, there is a considerable difference when it comes to nodes 14 and 26. Thus, VIS allows us to quantitatively differentiate between dominant influencer nodes.

B. Ranking of DVI nodes with VIS

DVI nodes are identified by ranking nodes based on VIS. Table III tabulate the top 10 dominant voltage influencer nodes for observation nodes 7 and 16, respectively.

Table III: DVI nodes for observation nodes 7 and 16

Rank → Metric ↓	1	2	3	4	5	6	7	8	9	10
Observation node -7										
MC	7	9	12	22	14	26	28	8	17	27
KL	7	9	12	14	22	26	28	8	17	18
BC	7	9	12	14	22	26	28	8	17	18
Observation node -16										
MC	14	22	18	12	9	7	17	8	26	28
KL	14	22	12	9	7	17	26	18	28	8
BC	14	22	12	9	7	17	26	18	28	8

According to Table III, the MC method indicate that the actor nodes 7 and 14 are Rank-1 nodes for observation node 7 and 16, respectively. The assignment of Rank-1 nodes is carried out correctly by both the metrics of the proposed approach. The reasons for rank-1 allocation are (1) the high variance of power change and (2) proximity of actor nodes to the given observation node in the IEEE 37-node test system.

Apart from rank-1, utilities might be interested in identifying other dominant nodes which are next to rank-1. In this regard, we use Top-N accuracy, which can be defined as,

$$\frac{|\{\text{Predicted Top-N nodes}\} \cap \{\text{True Top-N nodes}\}|}{N}, \quad (21)$$

where N is the desired number based on the applications. The Top-5 and Top-10 accuracy results are presented for observation nodes 7 and 16 in Table IV. The last column of the table denotes the mean accuracy when all the nodes of the IEEE-37 node network act as observation nodes. It can be observed that distance metrics KL and BC have fairly good identification accuracy, and the mean accuracy is more than 90%. Apart from Rank-1 node, our approach also provides the correct sequence of actor nodes when compared to baseline MC approach. However, in some positions, the order of actor nodes is flipped or offset by one or two units. This is because, the corresponding actor nodes have an equal influencing capacity for that particular observation node. For instance, in the case of observation node 7, the position of actor nodes 22 and 14 are flipped. However, this is not a major concern here, as we are more interested in correctly identifying nodes lying in a particular band (i.e., Top-5, Top-10) rather than their exact order. The power at the Top-5 or Top-10 actor nodes can be efficiently regulated to mitigate voltage violations.

Table IV: Identification accuracy of DVI nodes

Node → Accuracy ↓	obs. node-7 (%)	obs. node-16 (%)	Mean all nodes (%)
Top-5	100	80	91
Top-10	90	100	93

Furthermore, to get the overall influencing capacity of actor nodes, one can compute the mean VIS. For each actor node, the mean of VIS is taken across all observation nodes of the network. It can be observed From Table V that actor nodes 17 and 30 have maximum and minimum influencing capacity among all the actor nodes present in the network. This assignment is because of their topological position in the distribution network and the magnitude of power variance due to associated DER units or load variation.

Table V: Average VIS of actor nodes

Nodes	VIS	Nodes	VIS	Nodes	VIS
7	0.220	17	0.444	28	0.156
8	0.220	18	0.304	30	0.154
9	0.212	22	0.349	31	0.189
12	0.258	26	0.163	34	0.378
14	0.300	27	0.166	36	0.427

Table VI: Running time of various approaches

Time Metric	Running time Top-5 (s)
MC	12039
KL	0.92
BC	0.96

C. DVI nodes for Non-Gaussian distribution of power change

The proposed framework is generic and can be applied for power variations following any distribution with finite mean and variance. Therefore, along with the Gaussian, the proposed method is also evaluated for the case of power changes following a Beta distribution. The shape and scale parameters of the Beta distribution are chosen as 2 and 5, respectively. Here, power is varied at 5 actor nodes namely 7, 8, 14, 17, 22. The magnitudes of power variation and correlation structure are kept the same as described earlier for the Gaussian scenario. Table VII represents the ranking of all actor nodes for observation node 8 when power change follows the Beta distribution. It can be seen that the assignment of the Rank-1 node is carried out correctly by both the information-theoretic metrics, which further reinforces the merit of our approach.

Table VII: DVI nodes for observation node-8

Rank → Metric ↓	1	2	3	4	5
MC	8	7	17	22	14
KL	8	7	17	22	14
BC	8	7	17	14	22

D. Computational complexity of proposed method

Apart from correctly identifying the dominant voltage influencing nodes, the proposed approach offers a considerable computational advantage over the conventional approaches. Table VI reports the execution time of various approaches to identify Top-5 actor nodes for the observation node 7. It can be seen that the proposed approach involving any of the two distance metrics is multiple order faster than that of the conventional approach, which takes around 3.2 hrs. All the experiments are conducted on a machine with Intel i7 processor running at 2.2 GHz. This results demonstrates the computational advantage of the proposed approach over conventional methods.

The proposed approach of identifying DVI nodes can enhance voltage control strategies in multiple ways. One of the approaches is based on system partitioning, where a set of effective actor nodes is used for control actions. Our approach helps to identify such set of actor nodes in a computationally efficient manner. Given a particular scenario of power change in the system, one can determine the set of common observation nodes for which a particular actor node induces the maximum impact on voltage profiles. Therefore, the set of common observation nodes can form a cluster and it is cost effective to control power variations on the most dominant actor node to quickly restore voltages to their safe operational limits within that cluster. Along with clustering, our method of ranking nodes can also be used for optimal allocations of DERs or fault current limiters which involves sensitivity analysis. These control and asset allocation strategies will be investigated as part of our future work. Furthermore, the proposed approach is performance (voltage influence score) driven hence can also assist network-based approaches to efficiently identify critical nodes in a distribution network [31], [32].

V. CONCLUSION AND FUTURE WORK

The conventional methods to identify dominant voltage influencer nodes (DVI) are Monte-Carlo simulation-based that are computationally expensive. Therefore, this work focuses on an analytical approach and proposes statistical distance metrics to devise VIS that quantifies voltage influencing capacity of actor nodes. The VIS is then leveraged to identify the DVI nodes. The proposed framework computes VIS not solely on the basis of correlation between change in power at the actor node and change in voltage at the observation node, but it also relies on the magnitude of voltage variation that is being caused due to power change at actor nodes. Effectiveness and computational efficiency of the proposed method are illustrated by comparing the results with conventional method of identifying DVI nodes using Monte-Carlo simulations in IEEE 37-node test system. From the results, it can be inferred that the proposed statistical distance metrics effectively predict the DVI nodes while substantially reducing the execution time. The error in modeling the probability distribution would impact the dominant set. Therefore, the robustness of the proposed approach will be investigated as part of future work. Further, this scalable approach of identifying DVI nodes can find applications in a variety of areas such as in utilizing

DVI nodes for voltage control at strategic locations, identifying most influencer DERs that are responsible for voltage fluctuations, among others. Employing DVI nodes for network partitioning and efficient control applications will be pursued as part of our future work.

ACKNOWLEDGEMENT

This material is based upon work partly supported by the Department of Energy, Office of Energy Efficiency and Renewable Energy (EERE), Solar Energy Technologies Office, under Award # DE-EE0008767 and National science foundation under award # 1855216.

REFERENCES

- [1] K. Jhala, B. Natarajan, A. Pahwa, and L. Erickson, "Real-time differential pricing scheme for active consumers with electric vehicles," *Electric Power Components and Systems*, vol. 45, no. 14, pp. 1487–1497, 2017.
- [2] A. Kulmala, S. Repo, and P. Järventausta, "Coordinated voltage control in distribution networks including several distributed energy resources," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 2010–2020, 2014.
- [3] J. Barr and R. Majumder, "Integration of distributed generation in the volt/var management system for active distribution networks," *IEEE Transactions on Smart Grid*, vol. 6, no. 2, pp. 576–586, 2014.
- [4] W. Tao, C. Fang, S. Lu, Q. Liu, and R. Lin, "Voltage control of distribution network with photovoltaic power generation systems," in *2018 3rd International Conference on Intelligent Green Building and Smart Grid (IGBSG)*. IEEE, 2018, pp. 1–6.
- [5] H. Çimen and N. Çetinkaya, "Mitigation of voltage unbalance in microgrids using thermostatically controlled loads," in *2018 2nd international symposium on multidisciplinary studies and innovative technologies (ISMSIT)*. IEEE, 2018, pp. 1–4.
- [6] S. Acharya, M. S. El-Moursi, A. Al-Hinai, A. S. Al-Sumaiti, and H. H. Zeineldin, "A control strategy for voltage unbalance mitigation in an islanded microgrid considering demand side management capability," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 2558–2568, 2018.
- [7] K. Jhala, V. Krishnan, B. Natarajan, and Y. Zhang, "Data-driven preemptive voltage monitoring and control using probabilistic voltage sensitivities," in *2019 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE, 2019, pp. 1–5.
- [8] B. Zhao, Z. Xu, C. Xu, C. Wang, and F. Lin, "Network partition-based zonal voltage control for distribution networks with distributed pv systems," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 4087–4098, 2017.
- [9] K. Jhala, B. Natarajan, and A. Pahwa, "The dominant influencer of voltage fluctuation (divf) for power distribution system," *IEEE Transactions on Power Systems*, vol. 34, no. 6, pp. 4847–4856, 2019.
- [10] S. Munikoti, B. Natarajan, K. Jhala, and K. Lai, "Probabilistic voltage sensitivity analysis to quantify impact of high pv penetration on unbalanced distribution system," *IEEE Transactions on Power Systems*, pp. 1–1, 2021.
- [11] M. Brenna, E. De Berardinis, L. D. Carpinì, F. Foiadelli, P. Paulon, P. Petroni, G. Sapienza, G. Scrosati, and D. Zaninelli, "Automatic distributed voltage control algorithm in smart grids applications," *IEEE Transactions on Smart Grid*, vol. 4, no. 2, pp. 877–885, 2012.
- [12] M. E. Elkhatabi, R. El-Shatshat, and M. M. Salama, "Novel coordinated voltage control for smart distribution networks with dg," *IEEE transactions on smart grid*, vol. 2, no. 4, pp. 598–605, 2011.
- [13] E. Demirok, P. C. Gonzalez, K. H. Frederiksen, D. Sera, P. Rodriguez, and R. Teodorescu, "Local reactive power control methods for overvoltage prevention of distributed solar inverters in low-voltage grids," *IEEE Journal of Photovoltaics*, vol. 1, no. 2, pp. 174–182, 2011.
- [14] R. Tonkoski, L. A. Lopes, and T. H. El-Fouly, "Coordinated active power curtailment of grid connected pv inverters for overvoltage prevention," *IEEE Transactions on sustainable energy*, vol. 2, no. 2, pp. 139–147, 2010.
- [15] E. Dall'Anese and A. Simonetto, "Optimal power flow pursuit," *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 942–952, 2016.
- [16] M. Parniani and M. Iravani, "Voltage control stability and dynamic interaction phenomena of static var compensators," *IEEE transactions on power systems*, vol. 10, no. 3, pp. 1592–1597, 1995.

- [17] Y. Wan, M. A. A. Murad, M. Liu, and F. Milano, "Voltage frequency control using svc devices coupled with voltage dependent loads," *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 1589–1597, 2018.
- [18] M. Haque, "Compensation of distribution system voltage sag by dvr and d-statcom," in *2001 IEEE Porto Power Tech Proceedings (Cat. No. 01EX502)*, vol. 1. IEEE, 2001, pp. 5–pp.
- [19] R. Moghe, D. Tholomier, D. Divan, J. Schatz, and D. Lewis, "Grid edge control: A new approach for volt-var optimization," in *2016 IEEE/PES Transmission and Distribution Conference and Exposition (T&D)*. IEEE, 2016, pp. 1–5.
- [20] M. Bahramipناه, M. Nick, R. Cherkaoui, and M. Paolone, "Network clustering for voltage control in active distribution network including energy storage systems," in *2015 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*. IEEE, 2015, pp. 1–5.
- [21] M. Bahramipناه, R. Cherkaoui, and M. Paolone, "Decentralized voltage control of clustered active distribution network by means of energy storage systems," *Electric Power Systems Research*, vol. 136, pp. 370–382, 2016.
- [22] K. Mahmoud and M. Lehtonen, "Three-level control strategy for minimizing voltage deviation and flicker in pv-rich distribution systems," *International Journal of Electrical Power & Energy Systems*, vol. 120, p. 105997, 2020.
- [23] E. Dall'Anese, S. S. Guggilam, A. Simonetto, Y. C. Chen, and S. V. Dhople, "Optimal regulation of virtual power plants," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 1868–1881, 2017.
- [24] E. Cotilla-Sanchez, P. D. Hines, C. Barrows, S. Blumsack, and M. Patel, "Multi-attribute partitioning of power networks based on electrical distance," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4979–4987, 2013.
- [25] M. Baranwal and S. Salapaka, "Clustering and supervisory voltage control in power systems," *International Journal of Electrical Power & Energy Systems*, vol. 109, pp. 641–651, 2019.
- [26] M. Girvan and M. E. Newman, "Community structure in social and biological networks," *Proceedings of the national academy of sciences*, vol. 99, no. 12, pp. 7821–7826, 2002.
- [27] S. Munikoti, K. Jhala, K. Lai, and B. Natarajan, "Analytical voltage sensitivity analysis for unbalanced power distribution system," in *2020 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE, 2020, pp. 1–5.
- [28] M. Hassanzadeh, M. Etezadi-Amoli, and M. Fadali, "Practical approach for sub-hourly and hourly prediction of pv power output," in *North American Power Symposium 2010*. IEEE, 2010, pp. 1–5.
- [29] R. Herman and C. T. Gaunt, "Probabilistic modelling of parameter variability for analysing grid-connected lv feeders with dg," in *2013 48th International Universities' Power Engineering Conference (UPEC)*. IEEE, 2013, pp. 1–5.
- [30] M. Zugno, J. M. Morales, P. Pinson, and H. Madsen, "Pool strategy of a price-maker wind power producer," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 3440–3450, 2013.
- [31] A. M. L. da Silva, J. L. Jardim, L. R. de Lima, and Z. S. Machado, "A method for ranking critical nodes in power networks including load uncertainties," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1341–1349, 2015.
- [32] S. Munikoti, L. Das, and B. Natarajan, "Scalable graph neural network-based framework for identifying critical nodes and links in complex networks," *Neurocomputing*, vol. 468, pp. 211–221, 2022.

APPENDIX

The real and imaginary parts of voltage change (equation (6)) can be expressed as,

$$\begin{aligned}\Delta V_{OA}^{a,r} &= C_{OA}^{a,rT} \Delta S_A, \\ \Delta V_{OA}^{a,i} &= C_{OA}^{a,iT} \Delta S_A,\end{aligned}\quad (22)$$

where,

$$\begin{aligned}C_{OA}^{a,rT} &= [C_{OA}^{aa,r} \ D_{OA}^{aa,r} \ C_{OA}^{ab,r} \ D_{OA}^{ab,r} \ C_{OA}^{ac,r} \ D_{OA}^{ac,r}] \\ C_{OA}^{a,iT} &= [C_{OA}^{aa,i} \ D_{OA}^{aa,i} \ C_{OA}^{ab,i} \ D_{OA}^{ab,i} \ C_{OA}^{ac,i} \ D_{OA}^{ac,i}]\end{aligned}$$

The constants C and D are the functions of the line impedances and base voltages which are given by,

$$C_{OA}^{a,r} = \begin{pmatrix} C_{OA}^{aa,r} \\ D_{OA}^{aa,r} \\ C_{OA}^{ab,r} \\ D_{OA}^{ab,r} \\ C_{OA}^{ac,r} \\ D_{OA}^{ac,r} \end{pmatrix} = \begin{pmatrix} \frac{-(R_{OA}^{aa} \cos(\omega_A) - X_{OA}^{aa} \sin(\omega_A))}{|V_A^a|} \\ \frac{-(R_{OA}^{aa} \sin(\omega_A) + X_{OA}^{aa} \cos(\omega_A))}{|V_A^a|} \\ \frac{-(R_{OA}^{ab} \cos(\omega_A) - X_{OA}^{ab} \sin(\omega_A))}{|V_A^b|} \\ \frac{-(R_{OA}^{ab} \sin(\omega_A) + X_{OA}^{ab} \cos(\omega_A))}{|V_A^b|} \\ \frac{-(R_{OA}^{ac} \cos(\omega_A) - X_{OA}^{ac} \sin(\omega_A))}{|V_A^c|} \\ \frac{-(R_{OA}^{ac} \sin(\omega_A) + X_{OA}^{ac} \cos(\omega_A))}{|V_A^c|} \end{pmatrix} \quad (23)$$

Similarly, constants C and D for imaginary part of voltage change are expressed as,

$$C_{OA}^{a,i} = \begin{pmatrix} C_{OA}^{aa,i} \\ D_{OA}^{aa,i} \\ C_{OA}^{ab,i} \\ D_{OA}^{ab,i} \\ C_{OA}^{ac,i} \\ D_{OA}^{ac,i} \end{pmatrix} = \begin{pmatrix} \frac{-(R_{OA}^{aa} \sin(\omega_A) + X_{OA}^{aa} \cos(\omega_A))}{|V_A^a|} \\ \frac{-(X_{OA}^{aa} \sin(\omega_A) - R_{OA}^{aa} \cos(\omega_A))}{|V_A^a|} \\ \frac{-(R_{OA}^{ab} \sin(\omega_A) + X_{OA}^{ab} \cos(\omega_A))}{|V_A^b|} \\ \frac{-(X_{OA}^{ab} \sin(\omega_A) - R_{OA}^{ab} \cos(\omega_A))}{|V_A^b|} \\ \frac{-(R_{OA}^{ac} \sin(\omega_A) + X_{OA}^{ac} \cos(\omega_A))}{|V_A^c|} \\ \frac{-(X_{OA}^{ac} \sin(\omega_A) - R_{OA}^{ac} \cos(\omega_A))}{|V_A^c|} \end{pmatrix} \quad (24)$$

where R_{OA}^h, X_{OA}^h are the resistance and reactance of shared path between the observation node O and actor node A from the source node. h denotes the self or mutual impedance terms of a particular phase, for instance aa, ab, ac of phase a .



Sai Munikoti (S'19) received the B.Tech. degree in electrical engineering from the Birla Institute of Technology Sindri, India, in 2015, and the M.Tech. degree in electrical engineering from Indian Institute of Technology Gandhinagar, India, in 2018. He is currently pursuing Ph.D. with the department of Electrical and Computer engineering at Kansas State University, Manhattan, KS, USA. His research interests spans areas of power distribution system, graph theory, machine learning, and resilience analysis.



Mohammad Abujubbeh (S'20) received his B.Sc. in Electrical & Electronic Engineering from Eastern Mediterranean University, Famagusta, Cyprus in 2016. He was a M.Sc. student of Sustainable Environment and Energy Systems at Middle East Technical University, from 2017 to 2019. He is presently with Kansas State University pursuing his Ph.D. in Electrical and Computer Engineering. His research interests include cyber-physical systems, renewable energy integration, probabilistic loss sensitivity analysis and IoT-enabled smart grids.



Kumarsinh Jhala (M'18) is an energy systems scientist in the Center for Energy, Environmental, and Economic Systems Analysis in the Energy Systems Division at Argonne National Laboratory, Lemont, IL, USA. Dr. Jhala received the Ph.D. and the M.S. degree in Electrical and Computer Engineering from Kansas State University, Manhattan, KS, USA in 2018 and 2015, respectively. He received the B.E. degree in Electronics and Communication Engineering from Gujarat Technological University, Ahmedabad, India in 2013. He has performed extensive

research on distribution system modeling and analysis. His area of expertise spans domains of transactive energy systems, cyber-physical security, machine learning, statistical signal processing, game theory, stochastic modeling, optimization and control theory.



Balasubramaniam Natarajan (SM'08) received the B.E. degree (Hons.) in electrical and electronics engineering from the Birla Institute of Technology and Science at Pilani, Pilani, India, in 1997, the Ph.D. degree in electrical engineering from Colorado State University, Fort Collins, CO, USA, in 2002, and the Ph.D. degree in statistics from Kansas State University, Manhattan, KS, USA, in 2018. He is currently a Clair N. Palmer and Sara M. Palmer Endowed Professor and the Director of the Wireless Communication and Information Processing Research Group, Kansas State University. He has worked on and published extensively on modeling, analysis and networked estimation, and the control of smart distribution grids and cyber physical systems in general. He has published more than 200 refereed journal and conference papers. His research interests include statistical signal processing, stochastic modeling, optimization, and control theories. He has served on the editorial board for multiple IEEE journals including the IEEE Transactions on Wireless communications.

search Group, Kansas State University. He has worked on and published extensively on modeling, analysis and networked estimation, and the control of smart distribution grids and cyber physical systems in general. He has published more than 200 refereed journal and conference papers. His research interests include statistical signal processing, stochastic modeling, optimization, and control theories. He has served on the editorial board for multiple IEEE journals including the IEEE Transactions on Wireless communications.