# Proactive Frequency Stability Scheme Based on Bayesian Filters and Spectral Clustering

Gian Paramo, Mario D Baquedano-Aguilar, Arturo Bretas, Sean Meyn

# Proactive Frequency Stability Scheme Based on Bayesian Filters and Spectral Clustering

Gian Paramo*, Mario D. Baquedano-Aguilar*, Arturo Bretas†‡, Sean Meyn*
*Department of Electrical & Computer Engineering, University of Florida, Gainesville, FL, USA
†Electric Grid Security, Sandia National Laboratories, Albuquerque, NM, USA
‡G2Elab, Grenoble INP, CNRS, Université Grenoble Alpes, 38000 Grenoble, France
gparamo@ufl.edu, mariodaniel_ba@ieee.org, asbreta@sandia.gov, meyn@ece.ufl.edu

***Abstract*—This work presents a proactive distributed model for power system frequency stability. High-level penetration of renewable energy sources into the grid have introduced unforeseen and unmodeled system dynamics. Underfrequency load shedding state-of-the-art solutions are reactive in design, with efficiency constrained by the modeling error. Being able to detect unstable conditions early makes it possible to generate optimized corrective actions. In this work, phasor measurement units are used to predict frequency values. When a disturbance is detected, the state of frequency is predicted a few seconds into the future via a particle filter algorithm. Corrective actions are modeled through a mixed integer linear programming algorithm within system areas established through spectral clustering. The solution is implemented on *Matlab*, considering *IEEE* test systems. The proactive design of the method combined with its multiple layers of optimization deliver results that outperform state-of-the-art solutions. Easy-to-implement model, without hard-to-derive parameters, highlights potential aspects towards real-life implementation.**

***Index Terms*—phasor measurement units, spectral clustering, dynamic state estimation, mixed-integer linear programming, particle filter, frequency stability.**

## I. Introduction

As the electric grid transitions to a smaller carbon footprint, it has also become increasingly complex to operate, protect and control. These complexities are testing the limits of traditional solutions based on quasi-static models. The electric grid of today requires that protection and control solutions deliver satisfactory levels of performance under a wide-range of operating conditions unforeseen and unmodeled. Another limitation of state-of-the-art protection and control solutions is that they are reactive by design. In most cases corrective actions are taken only after the system has already entered a critical condition. With this in mind, the goal of this work is put forth a method that overcomes some of the limitations of quasi-static models in the area of power systems protection.

The specific contributions of this work are:

- The application of Bayesian filters in the context of frequency stability.
- The development of an adaptive distributed model for optimized corrective actions.

The remaining of the paper is organized as follows. A review of relevant works is presented in Section II. Section III presents a theoretical background. Proactive distributed model is presented in section IV. Case studies and their results are discussed in Section V. Finally, concluding remarks and future work are presented in Section VII.

## II. Literature Review

In order to address the limitations of traditional underfrequency load-shedding (UFLS) schemes such as delayed responses and over-shedding [1], [2], research efforts have investigated the use of adaptive and proactive schemes. For instance in [3] measurements captured by phasor measurement units (PMUs) are used to generate autoregressive models and produce predictions. In [4] a technique based on model predictive control is developed to optimize the operation of microgrids. In [5] PMUs are used monitor power flows and predictions are made via linearized models. The work in [6] makes predictions based on the power output of synchronous condensers. Finally, [7] uses polynomial curve-fitting to predict future frequency values.

These techniques deliver interesting results and have certainly inspired this work, however, most of them lack generality. Some of them rely on high PMU sampling rates, while others are based on specialized equipment such as synchronous condensers. Several of them base their decision-making on complex calculations, which require long computation periods to find potential solutions, alongside the challenge of detecting real-time threats due to the substantial volume of measured data. These factors place constraints on the accessibility and the applicability of these solutions, effectively limiting them to niche scenarios. For example, [8] may be inadequate since areas for load shedding are kept fixed for all operating conditions. These areas are predetermined for deterministic scenarios through offline analysis or by the power grid operator's knowledge of the system.

To address these issues, graph-theory-based network clustering algorithms such as spectral clustering have made it possible to determine areas (clusters) for load shedding in PMU time [9].

Spectral clustering algorithms are renowned for their deterministic and polynomial-time solutions, distinguishing them from traditional clustering algorithms like $k$-means [10], [11] and $k$-means++ [12], and are equivalent to graph min-cut problems [13], [14].

Traditionally, the process of decomposing large interconnected networks into loosely connected areas, for easier man-

agement, has relied on the concept of electrical cohesiveness. This cohesion is often quantified by the electrical distance between network nodes [15].

However, in spectral clustering the electrical cohesiveness is treated as a static edge weight function, see Section III-F, as the clustering solution remains unchanged regardless of operating conditions due to the fixed parameter. To be able to capture significant variations and fast dynamics introduced by renewable generation, dynamic edge weight functions can be utilized instead. These edge weight functions incorporate variable parameters such as power flows, generator synchronizing coefficients, distribution factors, and line ratings.

For the purpose of this study, it is essential to employ an electrical parameter capable of capturing actual operating conditions, particularly to account for the fast dynamics introduced by renewable generation in modern power systems. Therefore, power flow measurements, as discussed in [16]–[18], are used as dynamic edge weight functions since they are contingent upon the power system's current operating conditions.

In some cases, finding the complete solution space is unnecessary, so the spectral clustering can be constrained. In [19], two main types of constrained spectral clustering techniques are identified based on constraint enforcement. The first method involves manipulating the graph Laplacian or affinity matrix based on constraints and then applying unconstrained spectral clustering to the modified graph. The second type of method uses constraints to restrict the feasible solution space by defining centroids with different variations discussed in [19].

In this work, only unconstrained hierarchical spectral clustering is employed, which will be explained in Section III-F3.

## III. Theoretical Background

### A. Particle Filter

The particle filter (PF) is an estimation algorithm that combines elements from Bayesian estimation with the Monte Carlo method. The goal is to utilize a set of particles to approximate the probability distribution driving the behavior of the system being observed. This is an iterative process with updates and corrections at each time step, similar to those seen in popular algorithms such as the Kalman Filter (KF). An intuitive example is now presented in order to provide a general understanding of the algorithm. For a problem where the goal is to track an object moving a cross a 2-D plane, the PF starts by placing particles around the expected location of the particle. At this point these particles are simply educated guesses of where the object could be located. Measurements are then used to assign weights to the particles. Particles closer to the measured values are given higher weights. The average of the weighted particles is then used to estimate the location of the object being tracked. It should be noted that once weights are assigned to the particles, the collection or cluster of particles, effectively becomes a probability distribution, this is an important feature of the PF which will be discussed in more detail later in this work. For the following time

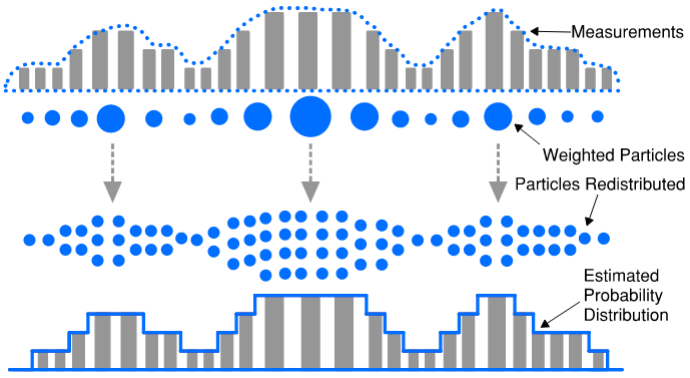


Fig. 1. Estimated Probability Distribution

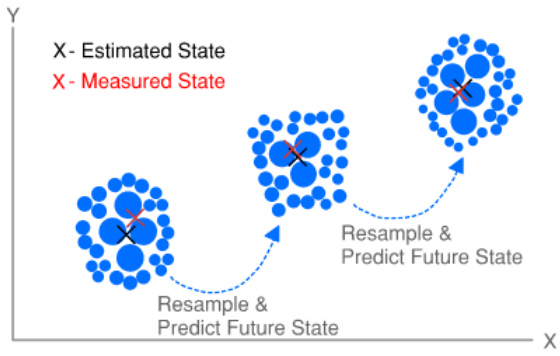


Fig. 2. Particle Based Sampling

step, the entire cluster must be moved according to a system model. For this example, the system model can be a simple equation relating the speed of the object to time. Once the cluster has been moved, particles are redistributed. Particles that were assigned lower weights during the previous time step are placed near particles of higher weight. At this point, measurements are once again used to update the weights of the particles, and a new estimate is produced. These ideas are illustrated in figures 1 and 2. The task of tracking the frequency of a power system is somewhat similar to the example just described. In this case the object moving across a 2-D is the frequency, with time on the x-axis and the frequency value (in Hertz) on the y-axis.

As previously mentioned, the ability of the PF to approximate a probability distribution via particles is an important feature of the algorithm, and it's also one of its main advantages. For instance, in its original formulation, the application of the KF is limited to data that follows a Gaussian distribution. This limitation could be problematic in the context of power systems, as research has shown that electrical measurements take on a diverse set of distributions [20]. An in-depth mathematical discussion of the PF can found in [21]. Comparisons between the PF and other filters can be found in [22].

### B. Predictions

One of the challenges of using a Bayesian filter for predictions is that in their original formulation these filters predict $K+1$ time steps into the future. Depending on sampling rates and processing times, the duration of each time step can be as fast as a fraction of a second. For the purposes of this work, a prediction a fraction of a second into the future isn't very useful. In order to overcome this limitation an equation was developed to extend the time horizon of the predictions. The equation is part of an iterative routine that creates a vector of artificial data points (ADPs) which are then fed to the particle filter. The equation is similar to the Taylor series and is defined as:

$$ADP_i = ADP_{i-1} + t_s f' + t_s^2 f'' \tag{1}$$

Here $ADP_{i-1}$ represents the previous $ADP$. When $i = 1$, $ADP_{i-1}$ refers to the last measurement. $t_s$ represents the duration of each measurement cycle. $f'$ and $f''$ represent the first and second derivatives of frequency respectively. These derivatives are included to account for the dynamics of the system.

The number of $ADPs$ required or $N_{ADP}$, is a function of the measurement periods, and how far into the future the prediction is set to be.

$$N_{ADP} = t_p f_s \tag{2}$$

Here $f_s$ is the reporting rate of the measurements, which corresponds to the inverse of the measurement period. $t_p$ is the time horizon of the prediction in seconds. As an example, if a prediction 3 seconds into the future is required with a reporting rate of 30 fps, then 90 $ADPs$ will be needed. The process that generates the $ADP$ vector that is then processed by the PF is summarized below.

**Algorithm 1** ADP Generation

*Initialisation*:
$ADP_{i-1}$ = Last measurement
$f'$ = Average first derivative in last 10 measurements
$f''$ = Average second derivative in last 10 measurements
$N_{ADP}$ = Number of ADPs required
1: **for** $i = 1$ to $N_{ADP}$ **do**
2: $ADP_i = ADP_{i-1} + t_s f'$
$f' = f' + t_s f''$
3: **end for**

### C. Detection

These predictions are triggered by changes in the frequency that surpass rate-of-change-of-frequency (RoCoF) thresholds. The RoCoF, $R$ in the equation that follows, is the change in frequency over a period of time.

$$R = \frac{f_2 - f_1}{dt} \tag{3}$$

Here $f_1$ and $f_2$ are the frequency values at the beginning and at the end of the measurement period respectively. The time window of this measurement is represented by $dt$.

### D. Load Excess Calculations

After the predictions are made, the predicted values are used to estimate the amount of compensation needed to bring the system back into stability. Load excess values are produced by equations derived from the swing equation. These equations were modified to be compatible with predicted values. The predicted load excess, $L_p$, is defined as:

$$L_p = \frac{R_p H_{est}(1 - \frac{f_p^2}{f_1^2})}{p(f_p - f_1)} \tag{4}$$

Where the frequency value at the start of the estimation period is represented by $f_1$. In most cases this is the current frequency value. Meanwhile $f_p$ represents the predicted frequency value produced by the PF. $p$ is the power factor of the system. $R_p$ is the RoCoF from $f_1$ to $f_p$. Finally $H_{est}$ is the estimated inertia constant. $H_est$ is calculated through an algorithm presented in [8], which is based on the work presented in [23].

### E. Optimizing the Response

Being able to foresee an unstable condition a couple of seconds before it occurs makes it possible to deliver an optimized response. This solution includes an optimization module that finds a combination of compensation agents that counters the excess loading found via eq. 4. The goal is to meet the needs of the system while adhering to the constraints set by the user. In this work three types of compensation agents are used: critical loads, sensitive loads, and non-critical loads. Each load type carries a different penalty if shed, with critical loads having the highest penalty, while non-critical loads have the lowest. A mixed integer linear programming (MILP) algorithm is used to scan the feeders in the area to find a suitable combinations of feeders. These ideas can be summarized as follows:

**Algorithm 2** Load Balance Optimization

*Initialisation*:
$A$ = Critical loads available for shedding (feeder level)
$B$ = Sensitive loads available for shedding (feeder level)
$C$ = Non-critical loads available for shedding (feeder level)
$c_1^T$ = Cost of shedding critical loads
$c_2^T$ = Cost of shedding sensitive loads
$c_3^T$ = Cost of shedding non-critical loads
$b$ = Calculated load excess
1: **Minimize** J = $c_1^t x + c_2^t y + c_3^t z$
s.t. $Ax + By + Cz \geq b$
2: **Return** Selected agents in $A$, $B$, and $C$

Additional constraints can be added to the optimization routine to meet the operating standard defined by the user.

### F. Spectral Clustering in Power Systems

Spectral clustering techniques make use of the spectrum (eigenvalues and corresponding eigenvectors) of the similarity matrix of the data to perform dimensionality reduction before clustering in fewer dimensions. The similarity matrix is provided as an input and consist of a quantitative assessment of the relative similarity of each pair of points in the data set.

In this work, we employ a spectral clustering methodology to assess the optimal regions within the power grid for preemptive load shedding, depending on varying operating conditions such as high, moderate, and low demand.

The identification and ongoing updates of these regions occur swiftly, within a time frame of less than $t < 90$ ms, aligning with the specified range of PMU processing times [9]. To mitigate communication challenges during fault events, the model relies on the latest complete set of PMU measurements preceding a disturbance to construct the graph representation of the network, a crucial step for determining the relevant regions.

*1) Determining Areas for Load-Shedding:* To identify regions for potential load shedding based on current operating conditions, the initial step involves constructing the graph Laplacian matrix from the measurement data set $W$.

The matrix $W$ is derived by representing the power grid as an undirected graph, denoted as $G = (\mathcal{V}, E)$. The set of buses is represented by $\mathcal{V} = 1, 2, ..., n$, where $n$ corresponds to the total number of buses. Additionally, the set of edges is defined as $E \subset \mathcal{V} \times \mathcal{V}$, where $(i, j) \in E$ indicates an edge connecting buses $i$ and $j$. These edges may represent transmission lines or transformers.

**The Power Flow Weighted Adjacency Matrix.** Since the topological structure of the graph does not capture the changing operational conditions of the power grid, we use edge weights. In this work, power flow measurements are used to define edge weights. For a network with $n$ buses, the weighted adjacency matrix of the graph is

$$W = \begin{pmatrix} 0 & \omega_{12} & \cdots & \omega_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ \omega_{n1} & \omega_{n2} & \cdots & 0 \end{pmatrix} \tag{5}$$

where $\omega_{ij}$ represents an edge weight as a non-negative function $\omega : \mathcal{V} \times \mathcal{V} \to \Re^+$ such that

- $\omega_{i,j} = 0$ if $(i, j) \notin E$, i.e., vertices $i$ and $j$ are not connected by and edge.
- $\omega_{i,j} = \omega_{j,i}$ if $i \neq j$, i.e, edge directions are ignored since $G$ is undirected.

**Graph Laplacians.** In the field of spectral clustering, three graph Laplacian matrices are well-documented: one is referred to as the unnormalized Laplacian, denoted as $L$; and the other two are recognized as the normalized versions, namely a symmetric matrix $L_n$ and a rectangular matrix associated with a random walk $L_{rw}$.

In our work, we opt for the *normalized* symmetric graph Laplacian matrix $L_n$, commonly utilized for clustering purposes (refer to [13], [14], [24]–[26] and related literature), to mitigate the impact of specific weights on clustering accuracy caused by fluctuating operating conditions, as depicted in the eigenvalues magnitudes of the unnormalized Laplacian matrix $L$ shown in Fig. 3 (left) for three operating conditions. This necessitates the transformation of the resulting weighted adjacency matrix $W$ into the normalized Laplacian and symmetric matrix $L_n$, which provide more stable and accurate results (see Fig. 3-right). The interested reader is referred to [24] for a detailed explanation of the reasons and advantages of using any of the normalized Laplacian matrices over the unnormalized version in the power systems area.

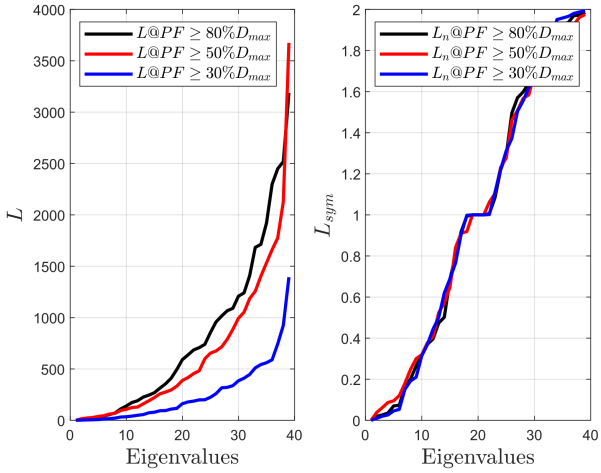


Fig. 3. Eigenvalues of both the *unnormalized* and the *normalized* Laplacian matrices for three ranges of system operating conditions with respect to maximum demand ($D_{max}$): high demand ($\geq 80\%D_{max}$); mid-demand ($\geq 50\%D_{max}$); and low demand ($\geq 30\%D_{max}$).

The entry-wise definition of the *unnormalized* graph Laplacian matrix follows:

$$L_{ij} = \begin{cases} d_i & \text{if } i = j \\ -1 & \text{if } i \text{ is adjacent to } j \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

where $d_i = \omega(i, i) = \sum_{j=1}^{n} \omega(i, j)$ is the degree of the $i$th vertex. This is closely related to the adjacency matrix and is sometimes written as $L = D - W$, where $D$ is defined as the diagonal matrix with weighted vertex degrees $d_1, \cdots, d_n$ on the diagonal. Then, the *normalized* and symmetric graph Laplacian matrix is defined as:

$$L_n = I - D^{-1/2} \cdot W \cdot D^{-1/2}. \tag{7}$$

The eigenvalues of $L_n$ satisfy the inequality $0 \leq \lambda_i^{L_n} \leq 2$ for all $i$ [13], [26], [27].

*2) Choice of Dimension for the Spectral Embedding:* A typical heuristic used to determine how many non-zero eigenvectors of the graph Laplacian matrix ($L$ or $L_n$) can be used for clustering purposes is to compute the absolute value of the ratio between the difference of two consecutive

eigenvalues and their size. In the spectral clustering literature, this parameter is known as the relative eigengap:

$$\gamma_{r,k} = \frac{|\lambda_{k+1} - \lambda_k|}{\lambda_k} \quad (k \ \geq 2). \tag{8}$$

The presence of favorable $k$-eigenvectors to be used for clustering purposes is indicated by a small value of the $k$th eigenvalue of the graph Laplacian, denoted as $\lambda_k$. Then, a high value of $\gamma_{r,k}$ means that a minimum of $k$-eigenvectors can be used to determine the number of system areas ($\mathcal{C}_k$) we seek to find.

*3) Determining the Number of Areas ($\mathcal{C}_k$):* Authors in [16], [28], suggest that the number of regions can be determined by knowing the number of groups of coherent generators in the system. However, in situations where such information is limited, two other approaches are recommended in the technical literature. One involves calculating the relative eigengap (see Section III-F2) as discussed in [17], [24], where the highest one determines the number of areas for load shedding. The other approach involves using agglomerative hierarchical clustering algorithms [29], where a dendrogram helps determine the number of areas that can be built based on the hierarchy of the similarity matrix. In this work, we consider the latter approach following the work of [25].

**The dendrogram.** A dendrogram is a tree diagram (see Fig. 4) that encodes the hierarchical structure in the spectral embedding. The bottom 'leaves' represent the individual vertices (buses), each considered as an initial cluster of unit size. At each step up the tree, the closest clusters are merged together, with the distance between clusters measured as the shortest pairwise distance between points in different clusters.

The dendrogram represents the hierarchical structure of the network at all levels simultaneously, and a specific $k$-partition, for some $\mathcal{C}_k$ areas, can be recovered by 'cutting' the dendrogram at level $k$ from the root.

Figure 4 illustrates the number of areas identified in the IEEE 39-bus system, determined from the hierarchy of the similarity graph obtained using the $k$-eigenvectors of the normalized Laplacian matrices corresponding to three sets of power flow measurements under different operating conditions. For instance, under high demand (left), $\mathcal{C}_k = 2$ regions are naturally visible, while under mid and low demand (center and right, respectively), $\mathcal{C}_k = 3$ regions can be observed. It is important to note that the elements within these clusters vary based on the operating condition.

*4) The Algorithm:* In this work, we utilize the normalized spectral clustering algorithm proposed in [24], [30] to identify the most effective areas for load shedding based on the actual operating conditions of the IEEE 39-bus system.

**Algorithm 3** The Spectral Clustering Algorithm.

INPUT: Matrix $\mathbf{W} \in \Re^{\mathbf{n}\times\mathbf{n}}$; parameter $\mathcal{C}_k$.
**Compute the normalized Laplacian** $L_{\text{sym}}$.
**Compute the eigenvectors** $v_1, \cdots, v_k$ of the generalized eigenproblem $L_{\text{sym}}v = \lambda Dv$ and take the first $k-$ eigenvectors.
**Let** $V \in \Re^{n\times k}$ be the matrix containing the vectors $v_1, \cdots, v_k$ as columns.
**For** $i = 1, \cdots, n$ **let** $y_i \in \Re^k$ be the vector corresponding to the $i$th row of $V$.
**Cluster the points** $(y_i)_{i=1,\cdots,n}$ in $\Re^k$ using a dendrogram into regions $C_1, \cdots, C_k$.
OUTPUT: **Regions** $A_1, \cdots, A_k$ with $A_i = \{j | y_j \in C_i\}$

## IV. Proactive Distributed Model

The goal of this section is to illustrate how the modules and techniques described in the previous sections come together. Figure 5 presents a data-flow chart of the model. Frequency and other relevant measurements are acquired through PMUs at a sampling rate of 30fps. Using this information inertia and areas of stability are created and updated periodically. The RoCoF is monitored in real time to identify disturbances requiring corrective actions. If RoCoF thresholds are surpassed, a prediction is made. The time horizon of the prediction is adjustable, but this work has found that 1 second offers a balance between the reliability of the predictions and having enough time to react. If the prediction suggests that frequency will continue to decline then corrective actions are taken based on these predicted values. First the load excess is computed per eq. 4, then a mixed integer linear programming (MILP) algorithm scans the feeders within its area to find a suitable combination of feeders to drop. Compensation is carried out in stages. The number of stages is an adjustable parameter defined by the user. In this work four stages are used, however in order to avoid unnecessary loss of load, provisions are made to stop pending stages of compensation when they are no longer necessary. This is done by producing a new prediction immediately after a stage of compensation is completed. This new prediction is compared with real-time RoCoF and rotor angle measurements. If the prediction suggests that frequency will be returning to normal, the RoCoF is positive, and rotor angles are stable level then further stages of compensation are not carried out. The ability to leverage to dynamics of the system allows the PF based solution to outperform other predictive techniques found in literature. This is explored in [8].

## V. Case Study

Simulations are conducted in Matlab and Simulink with a HP i7-1355U 32 GB RAM. In this case study, the spectral clustering technique described in Section III-F is applied to the IEEE 39-bus system to determine $\mathcal{C}_k$- areas using power flow (PF) measurements for three operating conditions with respect to maximum demand ($D_{max}$):

- High demand: PF $\geq 80\% D_{max}$.

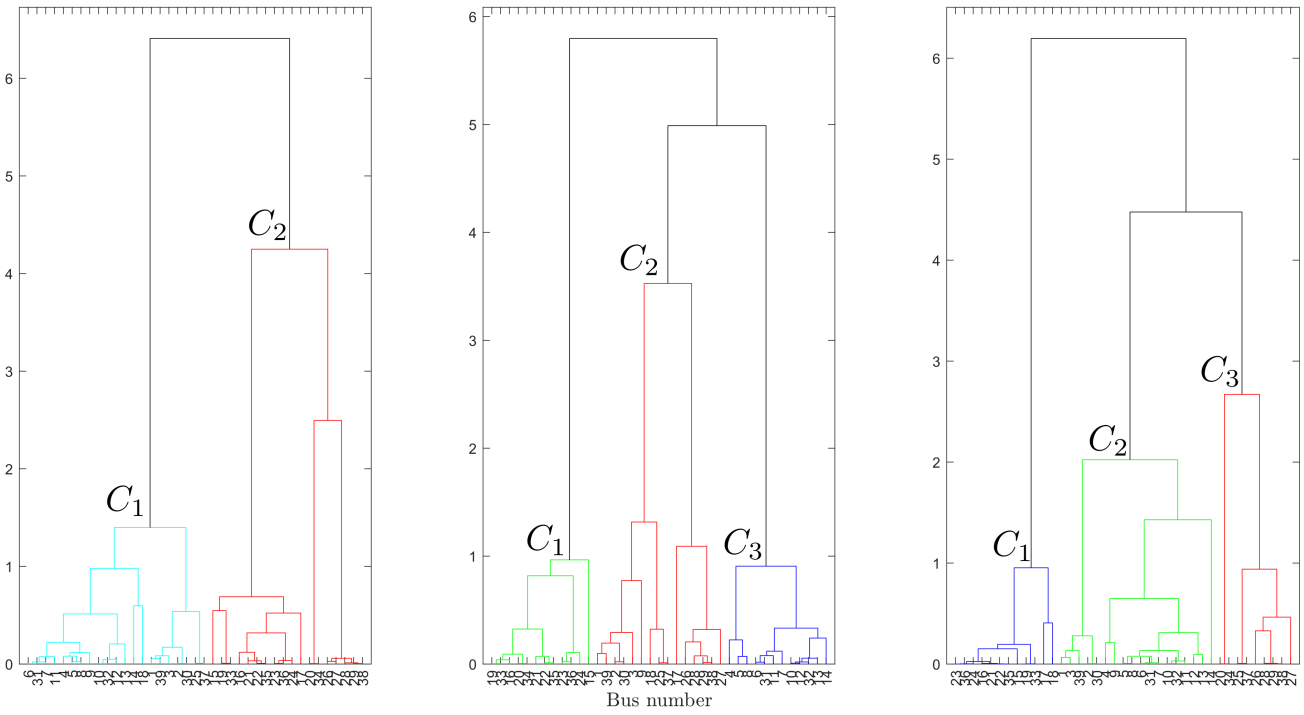


Fig. 4. A dendrogram for the similarity graph of $L_n$ for three operating conditions: high (left), mid (center), and low demand (right). Two natural clusters are evident for high demand, while three clusters emerge for mid and low demand conditions based on the hierarchy established from the similarity graph of $L_n$.

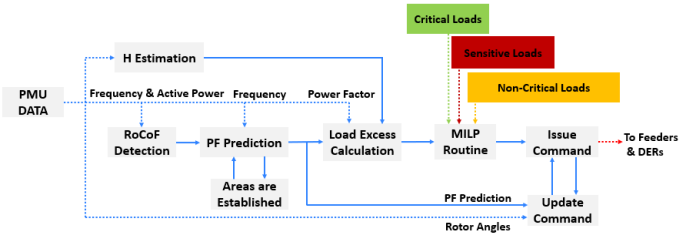


Fig. 5. Solution Overview

- Mid demand: PF $\geq 50\%D_{max}$.
- Low demand: PF $\geq 30\%D_{max}$.

**Obtaining k-Eigenvectors for the Spectral Embedding.** The power flows in the system are used to update the weighted adjacency matrix $W$ described in Section III-F1. The dendrograms of Fig. 4 show that at least $\mathcal{C}_k = 2, 3$ areas can be found based on the natural hierarchy of the similarity graphs for high, mid, and low demand operating conditions of the 39-bus system.

The relative eigengaps of $L$, and $L_n$ are computed to determine how many eigenvectors are needed for clustering purposes as described in Section III-F2. Figure 6 shows the relative eigengaps of $L$ (left) and $L_n$ (right) for the power flow-based Laplacians for all operating conditions considered.

It can be observed that the maximum values of $\gamma_{r,k}$ for the *unnormalized* and *normalized* Laplacians among operating

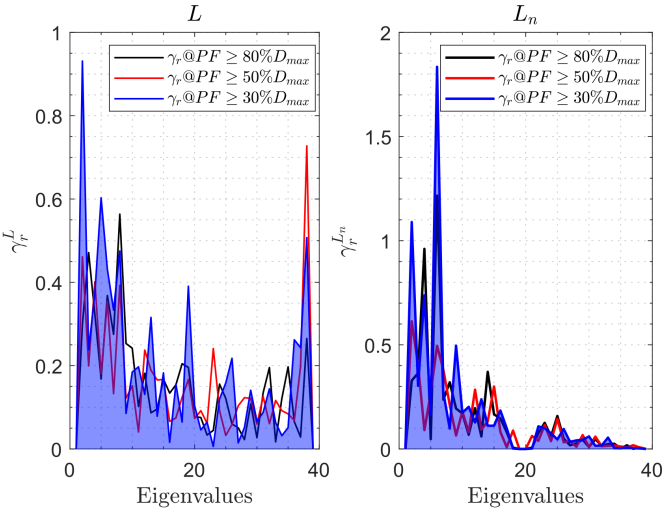


Fig. 6. The relative eigengap, $\gamma_{k,rel}$, of $L$ (left) and $L_n$ (right).

conditions are approximately $\gamma_{r,k}^{L} \approx 1$, and $\gamma_{r,k}^{L_n} \approx 2$, respectively. After several simulations performed, the clustering solutions using only $k = 1$ eigenvector of $L$ yield spurious results. In contrast, when using $k \geq 2$ eigenvectors of $L_n$, consistent results were obtained, aligning with findings in [13], [17], [24], [25]. Then, $\gamma_{r,k}^{L_n} \geq 2$ suggests that a minimum of $k = 2$ eigenvectors are needed to effectively cluster the power network into $\mathcal{C}_k$-areas.

**Determining the Number of Areas.** As explained in Sec-

tion III-F3, at least $\mathcal{C}_k = 2, 3$ areas can be found when employing the hierarchical spectral clustering approach (see the dendrogram of Fig. 4).

**Determining the Buses by Area.** By knowing the number of eigenvectors needed to determine the $\mathcal{C}_k$-areas we seek to find, then the spectral clustering algorithm described in Section III-F4 can be implemented to determine the system buses that belong to each of the $\mathcal{C}_k$-areas.

*1) Clustering Solution for $\mathcal{C}_k = 2$:* The resulting areas to shed load for high, mid, and low demand are depicted in Fig. 7 and specified in Table I.

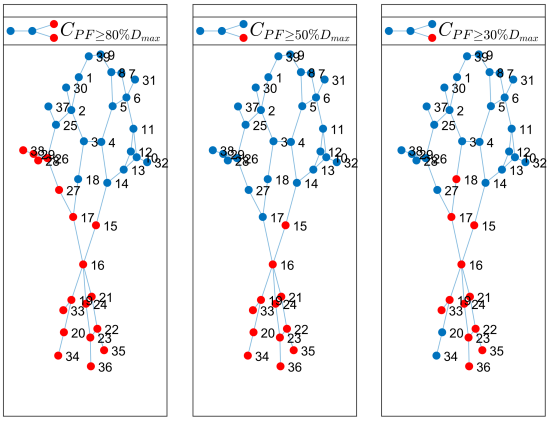


Fig. 7. For high, mid, and low demand operating conditions (left, center, and right, respectively), $\mathcal{C}_k = 2$ regions are identified in the IEEE 39-bus system with the hierarchical spectral clustering technique.

It can be observed from Fig. 7 that $\mathcal{C}_k = 2$ regions ($A_1 =$ blue, and $A_2 =$ red) for high and mid demand conditions can be effectively determined. Nevertheless, the solution obtained for the low demand condition is not consistent. Notice that its blue region $A_1$ (right) has buses $B_{20}, B_{34}$ disconnected from the rest of the buses in the region.

TABLE I
SPECTRAL CLUSTERING SOLUTION FOR $\mathcal{C}_k = 2$ REGIONS FOR HIGH, MID, AND LOW DEMAND CONDITIONS.

| **Operating Condition** | $A_1$ | $A_2$ |
|---|---|---|
| **PF** $\geq 80\%D_{max}$ | $1-14, 18, 25,$ $30-32, 37, 39$ | $15-17, 19-24,$ $26-29, 33-36, 38$ |
| **PF** $\geq 50\%D_{max}$ | $1-14, 17, 18,$ $25-32, 37-39$ | $15, 16, 19-24,$ $33-36$ |
| **PF** $\geq 30\%D_{max}$ | $1-14, 20, 25-32,$ $34, 37-39$ | $15-19, 21-24,$ $33, 35, 36$ |

*2) Clustering Solution for $\mathcal{C}_k = 3$:* The resulting load shedding areas for all operating conditions are depicted in Fig. 8 and specified in Table II.

Figure 8 shows that $\mathcal{C}_k = 3$ areas ($A_1 =$ blue, $A_2 =$ red, and $A_3 =$ green) can be effectively determined for mid-demand conditions only. The solutions obtained for both the high and the low-demand conditions are not consistent. Notice that their red regions $A_3$ (left and right) have buses $B_{20}, B_{34}$ disconnected from the rest of buses within the corresponding regions.

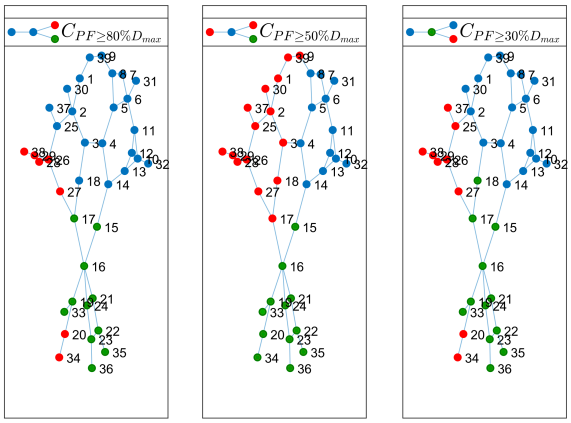


Fig. 8. For high, mid, and low demand operating conditions (left, center, and right, respectively), $\mathcal{C}_k = 3$ regions are identified in the IEEE 39-bus system with the spectral clustering technique

TABLE II
SPECTRAL CLUSTERING SOLUTION FOR $\mathcal{C}_k = 3$ REGIONS FOR HIGH, MID, AND LOW DEMAND CONDITIONS.

| **Operating Condition** | $A_1$ | $A_2$ | $A_3$ |
|---|---|---|---|
| **PF** $\geq 80\%D_{max}$ | $1-14, 18, 25,$ $30-32, 37, 39$ | $15-17, 19, 21-24,$ $33, 35, 36$ | $20, 26-29$ $34, 38$ |
| **PF** $\geq 50\%D_{max}$ | $1-3, 9, 17, 18,$ $25-30, 37-39$ | $15, 16, 19-24, 33-36$ | $4-8, 10-14,$ $31, 32$ |
| **PF** $\geq 30\%D_{max}$ | $1-14, 30-32, 39$ | $15-19, 21-24,$ $33, 35, 36$ | $20, 25-29,$ $34, 37, 38$ |

*3) Clustering Solution for $\mathcal{C}_k = 4$:* The resulting areas to shed load for all operating conditions are depicted in Fig. 9 and specified in Table III. Figure 9 shows that $\mathcal{C}_k = 4$ areas ($A_1 =$ blue, $A_2 =$ red, $A_3 =$ green, and $A_4 =$ black) can be effectively determined for all operating conditions.

Although the solution is consistent for both high and low-demand conditions, buses $B_{20}, B_{34}$ seem to be unnecessarily clustered into different regions as they appear to be part of their neighboring black regions. Notice that the bus $B_{20}$ has

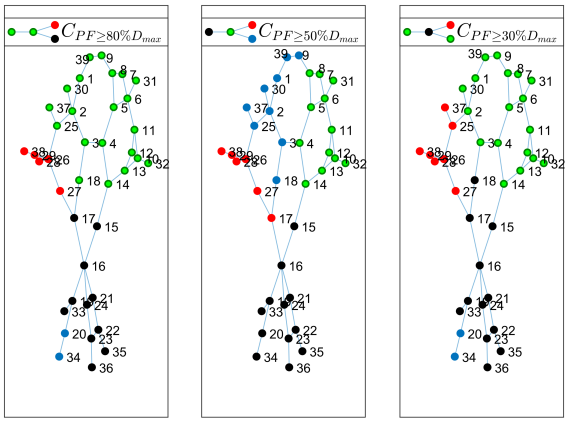


Fig. 9. For high, mid, and low demand operating conditions (left, center, and right, respectively), $\mathcal{C}_k = 4$ regions are identified in the IEEE 39-bus system with the spectral clustering technique.

a big load, which is served by the generator connected to bus $B_{34}$. However, the clustering solution is considered consistent.

TABLE III
SPECTRAL CLUSTERING SOLUTION FOR $\mathcal{C}_k = 4$ REGIONS FOR HIGH, MID, AND LOW DEMAND CONDITIONS.

| Operating Condition | $A_1$ | $A_2$ | $A_3$ | $A_4$ |
|---|---|---|---|---|
| **PF** $\geq 80\%D_{max}$ | $1-14, 18, 25, 30-32, 37, 39$ | $15-17, 19, 21-24, 33, 35, 36$ | $26-29, 38$ | $20, 34$ |
| **PF** $\geq 50\%D_{max}$ | $1-3, 9, 18, 25, 30, 37, 39$ | $15, 16, 19-24, 33-36$ | $17, 26-29, 38$ | $4-8, 10-14, 31, 32$ |
| **PF** $\geq 30\%D_{max}$ | $1-14, 30-32, 39$ | $15-19, 21-24, 33, 35, 36$ | $25-29, 37, 38$ | $20, 34$ |

In the next section, we will test the proactive load-shedding methodology for frequency stability enhancement using particle filters with the clustering solution for $\mathcal{C}_k = 4$ areas.

A delay of half a second is included from the moment a prediction is made to the moment a stage of compensation is executed. This is consistent with similar techniques found in literature [7]. It is assumed that the corresponding $H$ parameters have been estimated shortly before the start of the test. In order to showcase the capabilities of the PF based compensation technique and also the robustness of the clustering solution, the selected area in each test case is overloaded by 15% over the maximum capacity of the area.

### A. *Compensation for* **PF**$\geq 80\%D_{max}$

Area $A_2$ is selected for this test. The complete response of the area is illustrated in fig. 10. The disturbance is applied 1 second into the simulation. As frequency starts to decline, RoCoF thresholds are surpassed and this triggers the prediction shown in fig. 11.

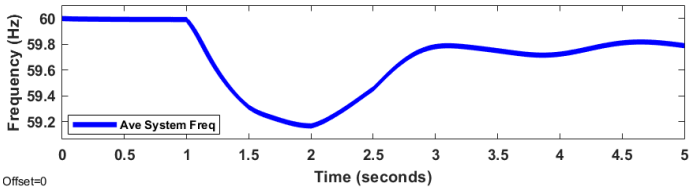


Fig. 10. Complete System Response

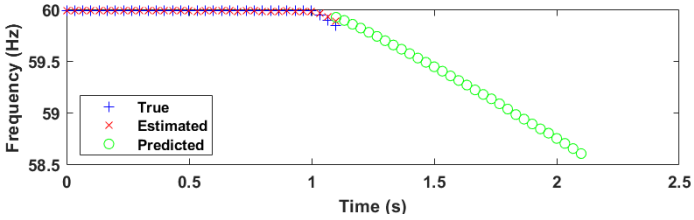


Fig. 11. Initial Prediction

Since the prediction suggests frequency will continue declining and the RoCoF is negative, compensation calculations are carried out. The predicted values shown in fig. 11, are used with eq. 4 to estimate the amount of compensation required. These values are then processed by the optimization algorithm described in Section III-E. The first stage of compensation takes place at 1.5 seconds as shown in fig. 12.

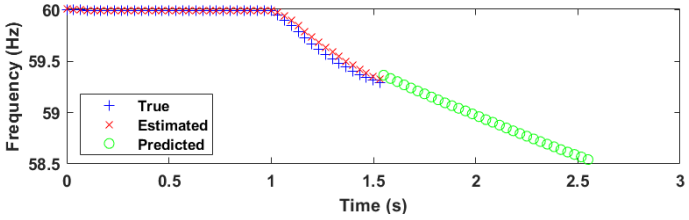


Fig. 12. Second Prediction

Shortly after the stage of compensation is completed, a new prediction is made. In this case the predicted values indicate that frequency is still on the decline, and therefore a new stage of compensation is performed. This is illustrated in fig. 13.

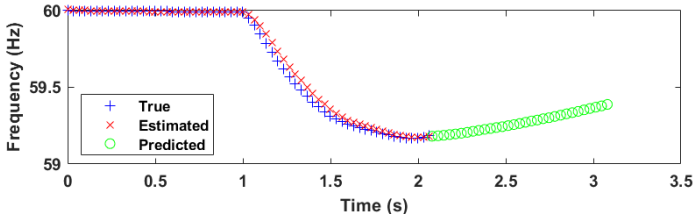


Fig. 13. Third Prediction

This second stage of compensation is completed at around 2 seconds. Once again, a new prediction is made. This time, however, frequency is expected to increase. This prediction combined with a positive RoCoF and stage rotor angles indicate that the system is returning to stable conditions, and as a result further stages of compensation are canceled. The rotor angles in this area are displayed in fig. 14 where the machine at bus 33 is used as reference.

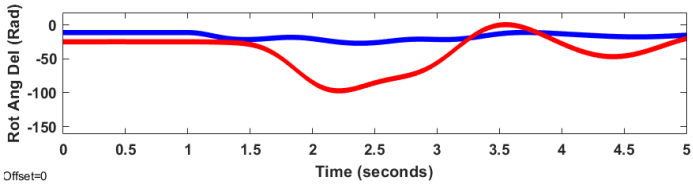


Fig. 14. Rotor Angle Deviation

### B. *Compensation for* **PF**$\geq 50\%D_{max}$

Area $A_1$ is selected for this test. The complete response of the area is illustrated in fig. 15. The disturbance is applied 1 second into the simulation. As frequency starts to decline, RoCoF thresholds are surpassed and this triggers the prediction shown in fig. 16.

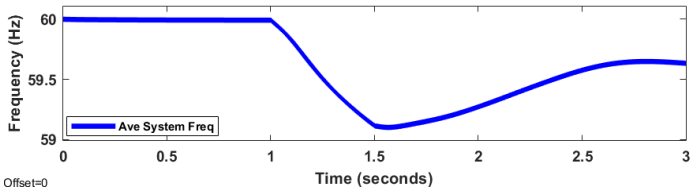


Fig. 15. Complete System Response

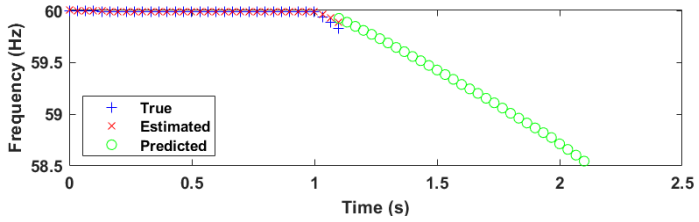


Fig. 16. Initial Prediction

Following the steps described in the previous test, the prediction leads to the stage of compensation illustrated in fig. 17.

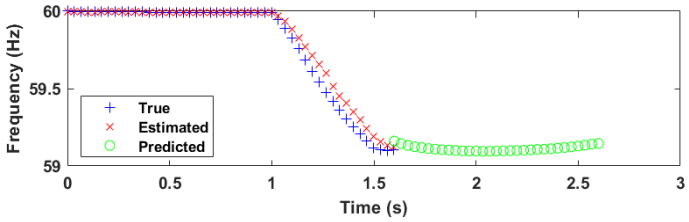


Fig. 17. Second Prediction

A new prediction is made immediately after the stage of compensation is completed. Since the prediction suggests that frequency will be returning to normal ranges, the algorithm stops after only one stage of compensation. Rotor angles with the machine at bus 30 used as reference are shown in fig. 18.

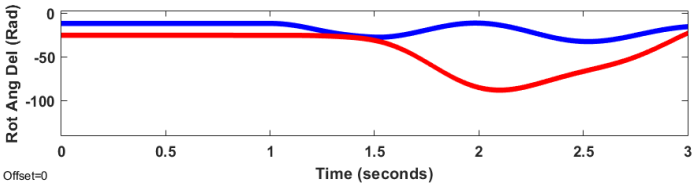


Fig. 18. Rotor Angle Deviation

### C. Compensation for $\boldsymbol{PF} \geq 30\% D_{max}$

Area $A_3$ is selected for this test. The complete response of the area is illustrated in fig. 19. The disturbance is applied 1 second into the simulation. As frequency starts to decline, RoCoF thresholds are surpassed and this triggers the prediction shown in fig. 20. This is followed by a series of stages of compensations and new predictions until a rise in frequency is predicted in fig. 22.

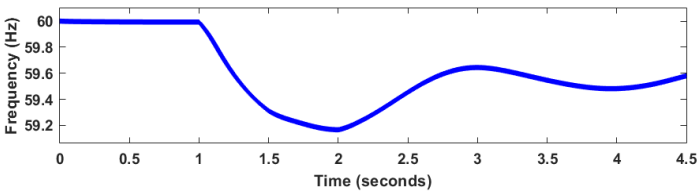


Fig. 19. Complete System Response

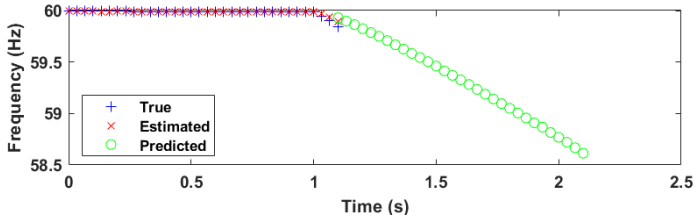


Fig. 20. Initial Prediction

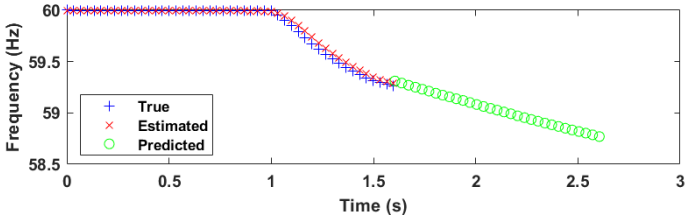


Fig. 21. Second Prediction

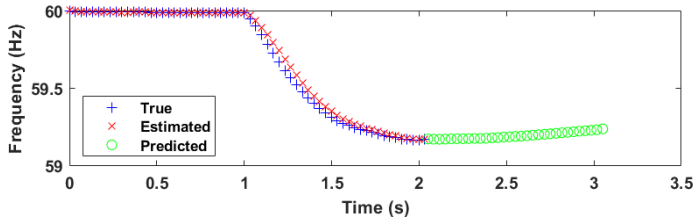


Fig. 22. Third Prediction

Rotor angles are depicted in fig. 23. The machine located at bus 37 is used as reference.

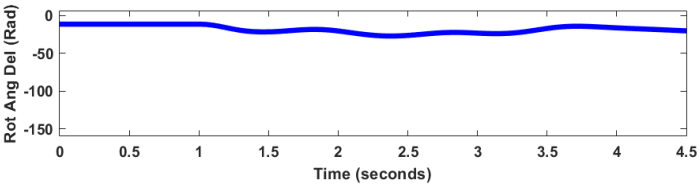


Fig. 23. Rotor Angle Deviation

## VI. Conclusions and Future Work

This work presented a proactive load-shedding model based on PMUs, Bayesian filters, and spectral clustering. Being able to predict an unstable condition before the system enters a critical state makes it possible to deliver an optimized response that mitigates the impact of disturbances. Bayesian filters were used to track the system's frequency via PMU measurements. As disturbances are detected predictions are made, and these predicted quantities are then used to formulate an optimized response. This work found that early corrective actions carried out in stages often overcome challenges presented by inaccuracies and uncertainty. The solution was tested via simulations that involve severe operating conditions. The scheme was able to successfully stabilize the system in all of these cases highlighting the robustness and flexibility of the solution. This work also addresses the limitations of communication networks by developing a solution based on

a distributed architecture. This alleviates concerns related to the transmission of large amounts of data. Additional case studies and comparisons with contemporary solutions found in literature can be found in [8]. On the one hand, the hierarchical spectral clustering method swiftly identifies clustering solutions within a short time frame of 9ms, which aligns with previous research findings [16], [24], [25], [30]. The results illustrated in Figures 7 to 9 suggest that fewer areas are consistently determined under high-demand conditions, while more areas need to be identified under low-demand conditions, which underscores the importance of avoiding fixed areas, as it could lead to instability issues as highlighted in [31]. Special attention should be given to low-inertia power systems [32]. This observation aligns with reliability assessments of power systems, where systems tend to exhibit higher reliability during periods of elevated demand due to the increased availability of generating units, compared to periods of low demand, which often rely on fewer and more dispersed generators to meet power requirements. While the areas identified for proactive load shedding are clearly delineated through the natural hierarchy of the similarity matrix, it may be computationally intensive for real-world power systems. Consequently, one of the future challenges of this research is to develop methods to accelerate the computation of the similarity matrix when employing hierarchical spectral clustering. Future research will also concentrate on validating the approach on larger systems and enhancing the efficiency and cohesiveness of the scheme

## VII. Acknowledgments

The authors gratefully acknowledge the support from the U.S. Department of Energy (DoE), Office of Cybersecurity, Energy Security, and Emergency Response (DOE-CESER), and Sandia National Laboratories, for the funding of this research.

## References

[1] S. Horowitz, A. Phadke, and J. Niemira, *Power System Relaying*. Wiley, 2013. [Online]. Available: https://books.google.com/books?id=odHDAQAAQBAJ

[2] L. Sigrist, L. Rouco, and F. M. Echavarren, "A review of the state of the art of ufls schemes for isolated power systems," *International Journal of Electrical Power & Energy Systems*, vol. 99, pp. 525–539, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0142061517323967

[3] N. Bretas and A. Phadke, "Real time instability prediction through adaptive time series coefficients," in *IEEE Power Engineering Society. 1999 Winter Meeting (Cat. No.99CH36233)*, vol. 1, 1999, pp. 731–736 vol.1.

[4] A. Parisio, E. Rikos, and L. Glielmo, "A model predictive control approach to microgrid operation optimization," *IEEE Transactions on Control Systems Technology*, vol. 22, no. 5, pp. 1813–1827, 2014.

[5] M. Larsson and C. Rehtanz, "Predictive frequency stability control based on wide-area phasor measurements," in *IEEE Power Engineering Society Summer Meeting,*, vol. 1, 2002, pp. 233–238 vol.1.

[6] A. Sauhats, A. Utans, J. Silinevics, G. Junghans, and D. Guzs, "Enhancing power system frequency with a novel load shedding method including monitoring of synchronous condensers' power injections," *Energies*, vol. 14, no. 5, 2021. [Online]. Available: https://www.mdpi.com/1996-1073/14/5/1490

[7] U. Rudez and R. Mihalic, "Wams-based underfrequency load shedding with short-term frequency prediction," *IEEE Transactions on Power Delivery*, vol. 31, no. 4, pp. 1912–1920, 2016.

[8] G. Paramo and A. Bretas, "Proactive frequency stability scheme: A distributed framework based on particle filters and synchrophasors," *Energies*, vol. 16, no. 11, 2023. [Online]. Available: https://www.mdpi.com/1996-1073/16/11/4530

[9] "IEEE Standard for Synchrophasor Measurements for Power Systems," *IEEE Std C37.118.1-2011 (Revision of IEEE Std C37.118-2005)*, pp. 1–61, 2011.

[10] A. Likas, N. Vlassis, and J. J. Verbeek, "The global k-means clustering algorithm," *Pattern recognition*, vol. 36, no. 2, pp. 451–461, 2003.

[11] F. Wang, C. Ding, and T. Li, "Integrated kl (k-means–laplacian) clustering: a new clustering approach by combining attribute data and pairwise relations," in *Proceedings of the 2009 SIAM International Conference on Data Mining*. SIAM, 2009, pp. 38–48.

[12] D. Arthur, S. Vassilvitskii *et al.*, "k-means++: The advantages of careful seeding," in *Soda*, vol. 7, 2007, pp. 1027–1035.

[13] U. Von Luxburg, "A tutorial on spectral clustering," *Statistics and computing*, vol. 17, pp. 395–416, 2007.

[14] A. Ng, M. Jordan, and Y. Weiss, "On spectral clustering: Analysis and an algorithm," *Advances in neural information processing systems*, vol. 14, 2001.

[15] S. Blumsack, P. Hines, M. Patel, C. Barrows, and E. C. Sanchez, "Defining power network zones from measures of electrical distance," in *2009 IEEE Power & Energy Society General Meeting*. IEEE, 2009, pp. 1–8.

[16] M. D. Baquedano-Aguilar, A. Bretas, and S. Meyn, "Coherency-Constrained Spectral Clustering for Proactive Load Shedding: A Tri-Level Model," in *2024 PES GM*, Seattle (USA), Washington, United States, Jul. 2024. [Online]. Available: https://hal.science/hal-04667966

[17] M. D. Baquedano-Aguilar, A. Bretas, N. Aljohani, and S. Meyn, "Tri-Level Linear Programming Model for Automatic Load Shedding Using Spectral Clustering," in *2024 Conference on Innovative Smart Grid Technologies, North America (ISGT NA 2024)*. Washington, United States: IEEE, Feb. 2024. [Online]. Available: https://hal.science/hal-04471128

[18] H. Li, G. W. Rosenwald, J. Jung, and C.-C. Liu, "Strategic power infrastructure defense," *Proceedings of the IEEE*, vol. 93, no. 5, pp. 918–933, 2005.

[19] X. Wang and I. Davidson, "Flexible constrained spectral clustering," in *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2010, pp. 563–572.

[20] S. Wang, J. Zhao, Z. Huang, and R. Diao, "Assessing gaussian assumption of pmu measurement error using field data," *IEEE Transactions on Power Delivery*, vol. 33, no. 6, pp. 3233–3236, 2018.

[21] J. Elfring, E. Torta, and R. van de Molengraft, "Particle filters: A hands-on tutorial," *Sensors*, vol. 21, no. 2, 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/2/438

[22] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, 2002.

[23] F. Zeng, J. Zhang, G. Chen, Z. Wu, S. Huang, and Y. Liang, "Online estimation of power system inertia constant under normal operating conditions," *IEEE Access*, vol. 8, pp. 101 426–101 436, 2020.

[24] M. D. Baquedano-Aguilar, S. Meyn, and A. Bretas, "Reduced-order models of static power grids based on spectral clustering," in *2023 North American Power Symposium (NAPS)*, 2023, pp. 1–6.

[25] R. J. Sánchez-García, M. Fennelly, S. Norris, N. Wright, G. Niblo, J. Brodzki, and J. W. Bialek, "Hierarchical spectral clustering of power grids," *IEEE Transactions on Power Systems*, vol. 29, no. 5, pp. 2229–2237, 2014.

[26] F. R. Chung, *Spectral graph theory*. American Mathematical Soc., 1997, vol. 92.

[27] J. R. Lee, S. O. Gharan, and L. Trevisan, "Multiway spectral partitioning and higher-order cheeger inequalities," *Journal of the ACM (JACM)*, vol. 61, no. 6, pp. 1–30, 2014.

[28] J. H. Chow, *Power system coherency and model reduction*. Springer, 2013, vol. 84.

[29] D. Peña *et al.*, "Análisis de datos multivariantes," 2002. [Online]. Available: https://www.researchgate.net/profile/Daniel-Pena/publication/40944325_Analisis_de_Datos_Multivariantes/links/549154880cf214269f27ffae/Analisis-de-Datos-Multivariantes.pdf

[30] M. D. Baquedano-Aguilar, A. Bretas, and S. P. Meyn, "Coherency-constrained spectral clustering for power network reduction," in *IEEE*

*Open Access Journal of Power and Energy (Under Review)*, 2024, pp. 1–10.

[31] M. D. Baquedano-Aguilar, D. G. Colomé, E. Agüero, and M. Molina, "Impact of increased penetration of large-scale PV generation on short-term stability of power systems," in *2016 IEEE 36th Central American and Panama Convention (CONCAPAN XXXVI)*, 2016, pp. 1–6.

[32] M. D. Baquedano-Aguilar, N. Aljohani, S. P. Meyn, and A. Bretas, "Implementational aspects for the characterization of low-inertia in power systems," in *2022 North American Power Symposium (NAPS)*. IEEE, 2022, pp. 1–6.