Proactive Frequency Stability Scheme: A Distributed Framework Based on Particle Filters and Synchrophasors

Gian Paramo ¹ and Arturo Bretas ¹, ²*,

¹ Department of Electrical & Computer Engineering, University of Florida, Gainesville, FL 32608, USA; gparamo@ufl.edu
² Pacific Northwest National Laboratory, Richland, WA 99352, USA
* Correspondence: arturo.bretas@pnnl.gov

Abstract: The reactive nature of traditional under-frequency load shedding schemes can lead to delayed response and unnecessary loss of load. This work presents a proactive framework for power system frequency stability. Bayesian filters and synchrophasors are leveraged to produce predictions after disturbances are detected. By being able to estimate the future state of frequency corrective actions can be taken before the system reaches a critical condition. This proactive approach makes it possible to optimize the response to a disturbance, which results in a decrease in the amount of compensation utilized. The framework is tested via Matlab simulations based on Kundur’s Two-Area System, and the IEEE 14-Bus System. Performance metrics are provided and evaluated against other contemporary solutions found in literature. During testing this framework outperformed other solutions by drastically reducing the amount of load dropped during compensation.

Keywords: frequency stability; particle filter; phasor measurement units; power systems

1. Introduction and Motivation

Traditional frequency stability methods are inherently reactive, and do not have the ability to adapt to the dynamic behavior seen in the modern grid [1]. Islands and frequency thresholds are currently identified via simulations [2]. This means that contingency plans are created for only a finite number of scenarios [3]. This practice leads to undesirable results including delayed response, overshedding, and uncoordinated response [4]. In this work, the term traditional under frequency load shedding (UFLS), refers to techniques requiring human interaction, and those based on fixed settings. Given their reactive nature, traditional UFLS methods operate only after the system has entered a critical state [5]. Mitigating this limitation of traditional schemes is the motivation for this work.

In recent years a number of novel solutions have been proposed to mitigate the shortcomings of traditional UFLS methods. However, despite the novelty of the solutions found in literature, many of them still take a reactive approach—corrective actions are taken after the system has entered a critical state. On the other hand, solutions that take a proactive approach can be divided into two groups: Those based on non-RoCoF approaches and those based on artificial intelligence (AI). The limitation of the first is that they are based on complex calculations, while AI-based solutions rely on historical data that might not always be available.

This work presents a proactive frequency control scheme based on phasor measurement units (PMUs) and the particle filter (PF). The contributions of this work to the state of the art lie in the use of the PF and PMUs to produce future estimates of the state of frequency after disturbances are detected. This proactive approach makes it possible to optimize corrective actions. Additionally, this technique eliminates the need to perform simulations in order to establish contingencies for emergency conditions. This is because the scheme is able to adapt to the dynamics of the system. Fundamentally, an empirical
model is derived from measurements and a frequency target of 60 Hz is used for reference. Future states are controlled by either actuating distributed energy resources (DERs) or shedding load. These control actions are optimized and executed iteratively. The concept and its steps are summarized as follows:

1. Disturbances are detected using the rate of change of frequency (RoCoF).
2. The state of frequency is predicted one to three seconds into the future.
3. A load excess factor is determined using the predicted state of frequency.
4. Using loading data collected at the feeder level, the method finds a suitable combination of load (feeders) that will be dropped in stages to meet the excess load factor and regain load-generation balance.

This paper is organized as follows: Section 2 provides a brief overview of the state-of-the-art. Basic theory is presented in Section 3. An overview of the proposed solution including an illustrative example are given in Section 4. Comparative case studies are found in Section 5. Finally, concluding remarks are given in Section 6.

2. State-of-the-Art

A solution that predicts the state of frequency after a disturbance is presented in [6]. This method derives an approximate model from PMU readings and when a disturbance is detected it predicts a new steady state operating point. Using the predicted frequency value, the amount of load to be shed is calculated. One of the highlights of this technique is that very few assumptions are made in the context of inertia and generator governors. However, the performance of [6] is highly dependent on system observability. Also, convergence times are highly variable, and in some cases it takes up to two seconds for the algorithm to provide a load excess factor. The uncertainty produced by variable processing times is an important limitation. For this reason, this solution is suited for systems with high inertia where processing time requirements are not as stringent.

In [7], curves are projected a few seconds into the future to estimate the future state of frequency. Polynomial curve fitting is used to produce these predictions. One of the main appeals of this technique is its simplicity; however, the technique has several limitations. First, high PMU sampling rates are used. This is a critical parameter that determines the feasibility of the technique, as higher sampling rates place a significant burden on communication systems [8]. Second, in [7] the algorithm alternates between first and second degree polynomials based on assumptions. Finally, the authors assume that deviations will follow a mostly linear pattern. These assumptions limit the type of systems where this technique can be implemented.

In [9], frequency is monitored continuously through a feedback controller. Load excess calculations are updated in real time reducing overshedding. A customized version of a PMU is developed in [10]. Measurements captured at fast sampling rates are used to detect and correct issues early. A multilevel load shedding solution is presented in [11]. This technique leverages neural networks to predict the trajectory of frequency after a disturbance. A similar technique based on deep Q learning, developed for micro-grids, is presented in [12]. An index based approach is introduced in [13]. Historical data is used to calculate a capacity index which is then used for load shedding. Performance issues in micro-grid UFLS related to lack of communication and dynamic inertia are addressed in [14]. The proposed technique analyzes the droop characteristics of the system and makes decisions when values deviate from the reference. A frequency stability control method based on topology control is presented in [15]. The goal of this method is to prevent load shedding by switching transmission lines.

Power measurements and power injections have also been proposed as a means of monitoring and controlling power system stability. In [16], DERs are actuated to compensate generation-load imbalances. The method uses the extended Kalman filter to drive DERs in real-time. A frequency stability solution based on monitoring the active power injections from synchronous condensers (SCs) is presented in [17]. A novel aspect of this method is that it offers an alternative to RoCoF detection. The premise of this method is that valuable
information such as the loss of a generating unit can be retrieved from the change in the active power injection from an SC. It is suggested that disturbances can be detected earlier by monitoring power injections from SCs compared to RoCoF based methods. In [18], a predictive transient stability mitigation method is presented. After a disturbance occurs, the center of oscillations (CoO) produced by the disturbance is located. This is done by comparing the frequencies at the two ends of the line. Once the location of the CoO has been identified, a simplified equivalent model is created. By monitoring the potential and the total energy (the sum of potential and kinetic energies) of the system, the affected generator can be seen drifting into instability before any thresholds are surpassed. A similar set of ideas is applied in [19] to develop an UFLS scheme based on RoCoF and the center of inertia (CoI). Inflections in the projection of the estimated CoI with respect to the RoCoF are used to detect disturbances.

3. Background

3.1. Particle Filter

The particle filter (PF), is a type of Bayesian filter that estimates probability distributions via particles. The PF combines data from statistical models with data produced by an underlying system model [20]. At each time step, predictions and corrections are made similar to the Kalman filtering process [21]. The PF is not limited to Gaussian distributions and has an inherent ability to quantify uncertainty [22–25].

Mathematically, the PF combines elements from the Monte Carlo method and Bayesian estimation. The estimated state along with uncertainties are calculated from probability density functions (PDFs) [26]. Being a Bayesian filter, the PF updates its prior and posterior at each time step. Given a process model, the prior can be updated using the following equation:

$$p(x_k|z_{1:k-1}) = \int p(x_k|x_{k-1})p(x_{k-1}|z_{1:k-1})dx_{k-1}$$

(1)

Here, $p$ is the posterior obtained from a PDF with inputs $u_{1:k}$ and measurements $z_{1:k}$. Inputs $u_k$, and the previous state $x_{k-1}$ are used to provide an updated state $x_k$ at time step $k$. This prior is a guess based on information received up to time step $k-1$. With this prior, Bayes’ theorem can be utilized to obtain the new posterior.

The steps described so far are very similar to the steps involved in Kalman filtering. A key feature that differentiates the PF from the KF, is the PF’s use of the Monte Carlo method. When searching for an optimal Bayesian solution, the PF evaluates a sum of weighted samples, known as particles:

$$p(x_{0:k}|z_{1:k}) \approx \sum_{i=1}^{N_s} w_i \delta(x_{0:k}-x_{0:k}^i)$$

(2)

In Equation (2), $N_s$ represents the number of samples in a set. $\delta(\cdot)$ refers to the Dirac delta function. $w_i$ represents the weight of each particle $x_{0:k}^i$.

Subsequent mathematical manipulations of Equation (2) yield Equation (3):

$$p(x_k|z_{1:k}) \approx \sum_{i=1}^{N_s} w_i \delta(x_k-x_k^i)$$

(3)

An important principle of the PF can be observed through (3), which is that the solution approaches the real values as the number of particles is increased [26].

Another pivotal element of the PF process is the concept of resampling. Resampling in the PF is analogous to the correction step in the KF. The PF resamples with the goal of correcting imbalances in particle weights [20]. The estimation process conducted by the PF is illustrated by Figure 1.

The system model used in this work is a simple sinusoidal. This underlying model was chosen because it is consistent with the oscillatory behavior of power system frequency. For the interested reader, an in-depth look at the theory behind the PF and its mathematical formulation are provided in [26]. Furthermore, multiple variations of Bayesian filters, including the PF are compared in [27].
3.2. Synchrophasors

Synchrophasors are electrical measurements sharing a common reference. These measurements are commonly synchronized via GPS clocks. Figure 2 offers a basic depiction of the framework.

Device that support synchrophasors are referred to as phasor measurement units or PMUs. PMUs have relatively high sampling rates, in some cases up to 120 frames per second (fps) [28]. These sampling rates combined with the synchronization of measurements, make it possible to derive input-output pairs that can be used to build mathematical models [29].

One common assumption made by PMU-based solutions is that high PMU sampling rates will be viable. Sampling rate limitations are in part driven by the capacity of communication networks. High PMU sampling rates could overwhelm communication networks in terms of data transfer and data storage [8,29]. With these limitations in mind, 30 fps seems to offer a balance between practicality and estimation accuracy. Estimation across multiple sampling rates is performed in [30], with 30 fps delivering an acceptable performance relative to higher sampling rates.

3.3. Disturbance Detection

Techniques that utilize derivatives or rate of change, to detect frequency disturbances are referred to as semi-adaptive techniques [31]. Semi-adaptive techniques offer a good balance between security and dependability [32]. In this work, the rate of change of frequency (RoCoF), represented by $R$, is used as the primary means of disturbance detection. $R$ can be computed as:

$$ R = \frac{f_2 - f_1}{dt} \tag{4} $$

In Equation (4), $f_1$ refers to the frequency at the start of the measurement period, and $f_2$ is frequency at the end of the period. $dt$ represents the duration of the period in seconds. One of the benefits of semi-adaptive detection is that it allows for intentional delays to be part of the scheme. These delays give the system a brief opportunity to self-correct issues.
before corrective actions are taken. The duration of these delays is inversely proportional to $R$. In this work, the thresholds and their respective delays were taken from [33].

3.4. Prediction

This work introduces a simple, yet effective way to predict the future state of frequency using the particle filter. By default the PF will make predictions; however, these predictions extend only $k + 1$ time steps into the future [26]. The time horizon of these predictions can be extended without significant modifications to the PF algorithm if measurements are augmented with artificial data points. The idea here is to use measurements to derive artificial data points, which are then processed by the PF. The output of the PF are estimates based on historical data, the underlying system model, and artificial data. From the perspective of the user, since these estimates are in the future, they can be considered a prediction. The artificial data points (ADPs) are computed as follows:

$$ADP_i = ADP_{i-1} + t_s f' + t_s^2 f''$$  \hspace{1cm}(5)

where $ADP_{i-1}$ is the previous ADP, or the last measurement when $i = 1$. $f'$ and $f''$ are the first and second frequency derivatives, respectively. $t_s$ represents the time window of the derivatives. These derivatives play a key role in determining the degree of influence system dynamics have on the ADPs. Derivatives calculated using only a few measurements produce a set of ADPs that is very dynamic (sensitive to changes), which can be useful in some situations, this is explored in the case studies. The number of measurements used in the derivation of the ADPs can be adjusted to increase (or decrease) the sensitivity of the algorithm to system dynamics as desired. This is similar to the tuning of $Q$ and $R$ parameters in the KF. These parameters are tuned to find a balance between new and historical data [34]. A compromise was found by averaging the change of the last ten measurements and then dividing this value by $t_s$. One can think of this as an average derivative. ADP computation is a sequential process that starts with the last measurement received. The derivatives are then used to adjust the previous ADP to produce the next one. The result is a vector containing the ADPs that will be the basis of predictions.

The time horizon of the prediction can be adjusted based on the number of ADPs used. For a desired time horizon, $t_p$, the desired number of ADPs can be found per:

$$N_{ADP} = t_p f_s$$  \hspace{1cm}(6)

The number of ADPs, $N_{ADP}$, is a function of the sampling rate $f_s$, and the time horizon of the prediction, $t_p$, which is measured in seconds. The process of ADP generation is summarized as follows, see Algorithm 1:

Algorithm 1 ADP Generation

Initialization:

$ADP_{-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:   $ADP_i = ADP_{i-1} + t_s f'$
3:   $f' = f' + t_s f''$
4: end for

In this work, after an initial prediction is made, subsequent predictions are made with an emphasis on new data. Before a subsequent load shedding stage is executed, synchronization between the machines is considered. Surprisingly, synchronization is often overlooked in frequency stability solutions in literature [29]. In this work, machine rotor angles are assumed to be tracked via relays at the same sampling rate used for frequency.
The difference between the rotor angles, $\Delta \theta$, together with the predicted state of frequency are used to determine the need for additional load shedding stages. This aspect of the solution will be explored in the case studies presented on Section 5.

3.5. Excess Loading Equations

After a prediction indicating the need for corrective actions is made, predicted frequency values are used to calculate the corresponding load excess $L$. The load excess can be calculated via the following expression, which is itself derived from the swing equation [35]:

$$
L = \frac{RH(1 - \frac{f_2}{f_1})}{p(f_2 - f_1)}
$$

(7)

Here, parameter $f$ refers to frequency measurements. $f_1$ is the frequency at the beginning, while $f_2$ is the frequency at the end of the measurement period. $R$ is the RoCoF found via Equation (4). $p$ is the power factor, while the inertia coefficient is represented by $H$. When $L$ is calculated using predicted values, a modified version of Equation (7) is used to calculate the corresponding predicted load excess $L_p$:

$$
L_p = \frac{R_p H_{est}(1 - \frac{f_2}{f_1})}{p(f_p - f_1)}
$$

(8)

In Equation (8), $f_p$ is the predicted frequency. $R_p$ is still the RoCoF, but now this parameter is calculated as the change from $f_1$, to the predicted frequency $f_p$, over the prediction period. Power factor $p$ is calculated the same before, meanwhile $H_{est}$ is now an estimated coefficient of inertia. The process of obtaining $H_{est}$ is described in the next section.

3.6. Online $H$ Estimation

The sampling rates and synchronization of PMUs make it possible to utilize system identification algorithms to estimate power system parameters [29]. One such approach, subspace estimation, has delivered encouraging results in power systems applications [36–38]. This work adopts an $H$ estimation approach presented in [39], where $H$ is estimated online from frequency and power measurements. As presented in [39], the online $H$ estimation algorithm is not compatible with the framework developed in this paper. The original formulation utilizes high PMU sampling rates (100 fps), and the necessary data is collected over long time windows. Therefore, two simple modifications are made:

1. The sampling rate is decreased to 30 fps, a much more manageable data transfer rate [8,40].
2. Estimation of $H$ is carried out periodically as opposed to in real-time. Historical data and forecasts can be used to reduce the number of times $H$ is estimated. This work uses dynamic loading (modeled from real life data available at [41]), and the procedure is run the equivalent of six times per day. For the case studies presented on Section 5, $H$ is assumed to have been estimated roughly two hours before the events take place.

3.7. Optimizing the Response

Mixed integer linear programming (MILP) is used to optimize the response. An objective function is minimized subject to a set of constraints. The algorithm iterates until it finds an optimal solution. The iteration process can be carried out in different ways depending on the requirements defined by the user. MILP offers advantages in terms of ease of implementation due to its relatively low complexity, and its ability to incorporate relaxations [42]. In terms of cost, in this work, DERs are given the lowest penalty among the three sources of compensation. This is followed by non-critical loads (i.e., residential loads). Sensitive
loads carry the highest cost as these represent industrial operations where an outage can lead to significant losses in revenue or equipment damage. Critical or life-safety loads were not considered in this work as these loads are usually not dropped intentionally [35]. A general MILP problem can be formulated as [43]:

\[
\begin{align*}
    \text{Minimize } J &= c^T x + d^T y \\ 
    \text{s.t. } A x + B y &\geq b \\ 
    (x, y) &\in \mathbb{R}^n \times \mathbb{Z}^p \\
    x_j &\in \mathbb{R} \text{ for } j \in C
\end{align*}
\]

where \( A \) is an \( m \times n \) matrix, \( B \) is an \( m \times p \) matrix, \( b \) is an \( m \)-dimensional vector, \( c \) is an \( n \)-dimensional vector, and \( d \) is an \( p \)-dimensional vector. \( c \) and \( d \) define the cost associated with selecting an agent from \( A \) and \( B \), respectively. In this context, the term agent refers to sources of load excess compensation (i.e., load shedding or distributed energy resource (DER) actuation). These ideas are illustrated by Figure 3.

When deciding which agents to use, in addition to cost, the algorithm also considers factors such as distance from the disturbance, and the number of customers affected. For instance, it might be more desirable to drop one large industrial load instead of a large number of residential customers. Additional parameters can be added as required, and all parameters can be modified to meet a desired operational standard.

The information fed to the algorithm (capacity and loading information), is already collected at the substation level and transmitted to control centers [44]. This step would not place a significant burden on communication networks since all it requires is a periodical loading and capacity update. The process of optimizing the response can be summarized as follows, see Algorithm 2:

**Algorithm 2 Load Balance Optimization**

**Initialization:**
- \( A \) = DER capacity available
- \( B \) = Sensitive loads available to be shed (feeder level)
- \( C \) = Non-critical loads available to be shed (feeder level)
- \( c_1^T \) = Cost of DER actuation
- \( c_2^T \) = Cost of shedding sensitive loads
- \( c_3^T \) = Cost of shedding non-critical loads
- \( b \) = Calculated load excess

1. **Minimize** \[ J = c_1^T x + c_2^T y + c_3^T z \]
   \[ \text{s.t. } A x + B y + C z \geq b \]
2. **Return** Selected agents in \( A \), \( B \), and \( C \)
As previously stated, the optimization routine can be easily modified to include additional agent types, costs, or constraints. After the optimal combination of DERs and loads for compensation has been identified, a trip signal is sent to the selected feeder breakers.

4. Overview in Distributed Form

This method is intended to operate as a distributed solution overseeing the frequency stability of an area of the grid. PMUs at the edge of an area share information with selected PMUs at neighboring areas, as depicted in Figure 4. In this work, each area is previously defined. Figure 4 illustrates a distributed framework with 4 areas, considering the IEEE 14-Bus system [45].

Figure 4. Overview of the architecture in distributed form.

Figure 5 shows the modules that make up this proposed solution.

Figure 5. Overview of the algorithm and its modules.

Related techniques in literature assume a processing time of roughly 500 ms. This includes 400 ms due to communication delays, and 100 ms that account for relay processing times and breakers operating times [6,7]. In this work a conservative delay of one full second is used.

Many of the operating parameters of this scheme are customizable, facilitating compliance with industry standards such as NERC PRC-006 [2]. Some of the ways this framework follows PRC-006 guidance include among others:

- Adjustable frequency thresholds.
- Adjustable number of load-shedding stages.
• Customizable islands (based on PMU availability).
• Additional parameters such as voltage levels, frequency overshoot, rotor angles, breaker status, and loading of power lines can be integrated as constraints as long as the networks can support it.
• This method eliminates the need for simulations to establish underfrequency load-shedding (UFLS) contingencies.

A flow diagram of the process is depicted by Figure 6.

![Data flow diagram of the method.](image)

5. Case Studies

This section contains three case studies that showcase the advantages of the presented method under different operating conditions and constraints. Each case study includes key performance metrics as well as comparisons with other methods found in literature. Before proceeding to the test cases, an illustrative example is presented. The purpose of the example is to summarize the topics and steps discussed in the previous sections.

5.1. Illustrative Example

Figure 7 illustrates the complete response of a four-machine system after a disturbance is introduced and mitigated using the scheme presented in this work.

![Multi-stage compensation based on predictions.](image)

The disturbance starts at 0.5 s, and it is detected via the RoCoF from Equation (4). At 1 s, a prediction is made as shown in Figure 8. Dotted vertical lines in red, considering Figure 7, represent compensation stages derived with the solution. To account for processing and communication delays, corrective actions are taken one full second after a prediction is made.
Figure 8. First particle filter prediction.

The method estimates the state of the system frequency one second into the future. These estimates are then used to calculate the loading excess via Equation (8), which is then followed by a compensation stage at 2 s, as seen in Figure 7. A second prediction is made immediately after the first stage of compensation, this is shown in Figure 9.

Figure 9. Second particle filter prediction.

Based on this new prediction, a second stage of compensation is executed at 3 s, which is then followed by a third prediction as illustrated in Figure 10.

Figure 10. Third particle filter prediction.

The third prediction shows that the frequency is trending back into stability. Since frequency is expected to return to its normal range, a third stage of compensation is not executed. In this case the system regains stability with minimal overshoot as seen in Figure 11.

Figure 11. Complete response estimated by the particle filter.

Subsequent stages of compensation are not executed when three conditions are met:
1. The PF predicts a rise in frequency.
2. The RoCoF is positive.

The goal of the method is to first stop the decay in frequency, and second, to send it back to a range where generator governors can stabilize it.
5.2. Case Study I

This case study contains four scenarios and it is based on Kundur’s Two-Area System. This model can be found at [46]. The locations where excesses in load are introduced are numbered (1 through 4) in Figure 12. Loading, capacity, and other operational parameters are presented in Tables 1 and 2, and also in Table A1 in the Appendix A. Considering Figure 12, buses are referred by numbers, loads by the letter ‘L’ followed by a number, capacitor banks by the letter ‘C’ followed by a number, and generators by the letter ‘G’ followed by a number. Again, to account for processing and communication delays, corrective actions are taken one full second after a prediction is made. Gaussian noise with a variance of 0.025 is added to the measurements.

![Figure 12. One line diagram of Case Study I.](image)

Table 1. Load flow data.

<table>
<thead>
<tr>
<th>Bus</th>
<th>V (pu)</th>
<th>P_G (MW)</th>
<th>P_L (MW)</th>
<th>Q_L (MVar)</th>
<th>Q_C (MW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.03</td>
<td>615</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>1.01</td>
<td>700</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>1.03</td>
<td>719</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>1.01</td>
<td>700</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>-</td>
<td>967</td>
<td>100</td>
<td>200</td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td>-</td>
<td>1767</td>
<td>100</td>
<td>350</td>
</tr>
</tbody>
</table>

Table 2. Generator dynamic data

<table>
<thead>
<tr>
<th>Generator</th>
<th>Rating (MVA)</th>
<th>X_d (pu)</th>
<th>X_d’ (pu)</th>
<th>X_d'' (pu)</th>
<th>T_d0 (s)</th>
<th>T_d0' (s)</th>
<th>T_d0'' (s)</th>
<th>X_q (pu)</th>
<th>X_q' (pu)</th>
<th>X_q'' (pu)</th>
<th>T_q0 (s)</th>
<th>T_q0' (s)</th>
<th>T_q0'' (s)</th>
<th>H (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>900</td>
<td>1.8</td>
<td>0.3</td>
<td>0.25</td>
<td>8</td>
<td>0.03</td>
<td>1.7</td>
<td>0.25</td>
<td>0.4</td>
<td>0.05</td>
<td>6.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2</td>
<td>900</td>
<td>1.8</td>
<td>0.3</td>
<td>0.25</td>
<td>8</td>
<td>0.03</td>
<td>1.7</td>
<td>0.25</td>
<td>0.4</td>
<td>0.05</td>
<td>6.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3</td>
<td>900</td>
<td>1.8</td>
<td>0.3</td>
<td>0.25</td>
<td>8</td>
<td>0.03</td>
<td>1.7</td>
<td>0.25</td>
<td>0.4</td>
<td>0.05</td>
<td>6.175</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G4</td>
<td>900</td>
<td>1.8</td>
<td>0.3</td>
<td>0.25</td>
<td>8</td>
<td>0.03</td>
<td>1.7</td>
<td>0.25</td>
<td>0.4</td>
<td>0.05</td>
<td>6.175</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The scenario where the load excess is placed at bus 1 in Figure 12, is run twice. First, a fixed three stage compensation procedure is used. The first compensation stage is set at 50% of the calculated load excess. Stages 2 and 3 compensate 25% of the load excess each, for a total of 100%, similar to the approach used by [7]. The load imbalance starts at 0.5 s, and each compensation stage is illustrated through vertical dotted red lines in Figure 13.
In this case the system was successful in first, arresting the decline in frequency, and second, in providing a smooth transition back to steady state. Three predictions were made by the particle filter, they are presented in Figures 14–16:

This same scenario was performed using a version of the algorithm referred to as adaptive compensation. Adaptive compensation works as follows: The initial stage compensates for 50% of the estimated excess load factor. After the first stage is executed, the number of samples used in obtaining the derivatives in Equation (5) is decreased, making the particle filter more receptive to system dynamics. Compensation with variable stages is illustrated in Figure 17 through a red dotted vertical line.
Figure 17. Scenario 1, PF compensation with adaptive stages.

In this case, the PF predicts that frequency will be returning to levels close to the reference after only one stage of compensation. This is shown in Figure 18.

Figure 18. PF prediction after the first stage.

With frequency expected to return to its normal range, and $\Delta \theta$ not trending towards loss of synchronism, further compensation stages are avoided. Being able to restore frequency stability while compensating only a portion of the calculated load excess happens for two reasons: First, being able to take corrective actions early means that compensation can begin before the system reaches a critical state. Second, generator governors and other controllers are not considered during compensation calculations, and while this might lead to some inaccuracies, it also gives the scheme a safety margin, which translates into a lower amount of load shed. Realistically, a highly accurate load compensation value is difficult to obtain given the time constraints and the dynamics of both the loads and the sources. Moreover, it has been suggested that in the context of frequency stability, early action trumps accuracy [7].

This scenario is performed once again with the adaptive compensation procedure, but this time, polynomial curve fitting (PCF) as presented in [7], is used to obtain predictions instead of the PF. After the first stage of compensation has been executed, a follow up prediction is made as illustrated by Figure 19.

Figure 19. Prediction via PCF.
The prediction obtained via polynomial curve fitting is unable to leverage the dynamics of the system the way the PF can, leading to further stages of compensation.

Three more disturbances are introduced and both the PF based method and the PCF technique are used to mitigate them. A performance comparison is presented in Table 3.

**Table 3.** Key metrics for Case Study I.

<table>
<thead>
<tr>
<th>Disturbance Location</th>
<th>Method</th>
<th>Number of Stages</th>
<th>% of Load Excess Compensated</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PF</td>
<td>1</td>
<td>50%</td>
</tr>
<tr>
<td></td>
<td>PCF</td>
<td>3</td>
<td>100%</td>
</tr>
<tr>
<td>2</td>
<td>PF</td>
<td>2</td>
<td>75%</td>
</tr>
<tr>
<td></td>
<td>PCF</td>
<td>3</td>
<td>100%</td>
</tr>
<tr>
<td>3</td>
<td>PF</td>
<td>2</td>
<td>75%</td>
</tr>
<tr>
<td></td>
<td>PCF</td>
<td>3</td>
<td>100%</td>
</tr>
<tr>
<td>4</td>
<td>PF</td>
<td>1</td>
<td>50%</td>
</tr>
<tr>
<td></td>
<td>PCF</td>
<td>3</td>
<td>100%</td>
</tr>
</tbody>
</table>

For the four scenarios presented in Table 3, the method presented in this paper was able to bring the system back into stability while shedding 25% to 50% less load. This is, as previously discussed, due to the ability of the prediction model to leverage the dynamics of the system as frequency begins to return to a stable condition. Performance metrics are examined in Table 4, where frequency overshoot during recovery is measured.

**Table 4.** Overshoot values for Case Study I.

<table>
<thead>
<tr>
<th>Disturbance Location</th>
<th>Number of Stages</th>
<th>Overshoot</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>-0.83%</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1.13%</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>0.91 %</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>2.1%</td>
</tr>
</tbody>
</table>

As expected, more stages lead to a smoother transition back into steady state. It is worth noting that even when only one compensation stage was used, the overshoot was still fairly low.

5.3. Case Study II

A second case study was conducted using the IEEE 14-Bus System as reference. Following the same process as in the first case study, four scenarios are presented. The locations where the load excesses are introduced are presented in Figure 20, through numbers. Loading, capacity, and other operational parameters are presented in Table A2 in the Appendix A.

Once again, each scenario is run twice, one time using PCF to build a base line and then using the PF. Key metrics are presented in Table 5.

The results of this case study are consistent with the results of the first one, and they show that the PF based method facilitates the recovery of frequency while greatly reducing the amount of compensation used. Overshoot is also used as a performance metric for this case study, and the results are presented in Table 6.
Table 5. Key metrics for Case Study II.

<table>
<thead>
<tr>
<th>Disturbance Location</th>
<th>Method</th>
<th>Number of Stages</th>
<th>% of Load Excess Compensated</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PF</td>
<td>1</td>
<td>50%</td>
</tr>
<tr>
<td></td>
<td>PCF</td>
<td>3</td>
<td>100%</td>
</tr>
<tr>
<td>2</td>
<td>PF</td>
<td>2</td>
<td>75%</td>
</tr>
<tr>
<td></td>
<td>PCF</td>
<td>3</td>
<td>100%</td>
</tr>
<tr>
<td>3</td>
<td>PF</td>
<td>1</td>
<td>50%</td>
</tr>
<tr>
<td></td>
<td>PCF</td>
<td>3</td>
<td>100%</td>
</tr>
<tr>
<td>4</td>
<td>PF</td>
<td>1</td>
<td>50%</td>
</tr>
<tr>
<td></td>
<td>PCF</td>
<td>3</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 6. Overshoot values for Case Study II.

<table>
<thead>
<tr>
<th>Disturbance Location</th>
<th>Number of Stages</th>
<th>Overshoot</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>2.3%</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0.84%</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>−1.72%</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>1.93%</td>
</tr>
</tbody>
</table>

The observed overshoot in this scenario is once again fairly low, which is consistent with the results obtained in the first case study.

5.4. Case Study III

The goal of this final test case is to highlight the flexibility of the method and use it to drive distributed energy resources (DERs) in real-time. This test was run under a similar set of assumptions as those made in [16], where a KF-based compensation method is presented. In order to test the algorithm under demanding conditions, faster frequency deviations than those seen in [16] were generated. Most importantly, the total delay time involved in the processing of data and actuation of DERs was increased to 500 ms; up from the 40 ms time delay used in [16]. That is a response time over ten times slower.
Frequency deviations start at 1 s, with a significant loss of generation at 1.5 s. As illustrated in Figure 21, DER actuation takes place 0.5 s after the frequency deviations.

![Figure 21. Real-time mitigation via DERs.](image)

As before, measurements are made via PMUs at 30 fps. The power mismatch is calculated continuously in this test case per Equations (4) and (8) in the form of a controller. As shown in Figure 21 above, the frequency and power mismatch follow virtually the same trend but in opposite directions. When frequency deviates from the 60 Hz reference, a corresponding current output is seen from the DERs. The compensation is based on the power mismatch calculated by the controller. Despite having a 0.5 s delay, the system successfully mitigates the frequency deviations.

6. Conclusions

This work presented a distributed scheme for proactive frequency stability. By being able to estimate the future state of frequency, corrective actions can be taken before the system reaches a critical condition. This proactive approach makes it possible to optimize the response to a disturbance thereby reducing the amount of compensation utilized. The method was tested via simulations based on Kundur’s Two-Area System, and the IEEE 14-Bus System. The results were promising and outperformed other contemporary solutions by drastically reducing the amount of load dropped during compensation. Future work will focus on streamlining the solution by reducing requirements in terms of data and the number of PMUs utilized. Alternative optimization methods as well as improving the robustness of the particle filter will also be investigated.

Author Contributions: Conceptualization, G.P. and A.B.; methodology, G.P.; software, G.P.; validation, G.P.; formal analysis, G.P.; investigation, G.P.; resources, G.P.; data curation, G.P.; writing—original draft preparation, G.P.; writing—review and editing, G.P. and A.B.; visualization, G.P.; supervision, A.B.; project administration, G.P.; funding acquisition, A.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.
Abbreviations

DER Distributed Energy Resource
KF Kalman Filter
MILP Mixed Integer Linear Programming
PMU Phasor Measurement Unit
PF Particle Filter
RoCoF Rate of Change of Frequency
UFLS Underfrequency Load-Shedding

Appendix A

The tables below present the loading conditions for each case study. Total DER Capacity refers to the total additional power DERs can supply (rated power minus power supplied). This number is given as a percentage of the total rated power of the generators. Excess loading is also given as a percentage of the total rated power of the generators. A detailed list of parameters for the first case study can be found at [46]. For the second case study, the IEEE 14-Bus System was used with default parameters. In both cases the loading was divided into groups so the algorithm described in Section 3.7 could be used.

Table A1. System loading and capacity for Case Study I.

<table>
<thead>
<tr>
<th>Disturbance Location</th>
<th>Total DER Capacity</th>
<th>Excess Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5%</td>
<td>18%</td>
</tr>
<tr>
<td>2</td>
<td>3%</td>
<td>21%</td>
</tr>
<tr>
<td>3</td>
<td>6%</td>
<td>20%</td>
</tr>
<tr>
<td>4</td>
<td>8%</td>
<td>22%</td>
</tr>
</tbody>
</table>

Table A2. System loading and capacity for Case Study II.

<table>
<thead>
<tr>
<th>Disturbance Location</th>
<th>Machine</th>
<th>Loading %</th>
<th>Total DER Capacity</th>
<th>Excess Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>80%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>85%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>88%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>90%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>85%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>80%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>90%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>85%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>85%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>80%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>90%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>88%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>80%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>80%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>85%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>85%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>85%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>85%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>85%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>85%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>