Overcoming Communication Delays in Distributed Frequency Regulation

Thiagarajan Ramachandran, *Student Member, IEEE,* Masoud H. Nazari, *Member, IEEE,* Santiago Grijalva, *Senior Member, IEEE,* Magnus Egerstedt, *Fellow, IEEE,*

***Abstract*—This paper proposes a general framework for determining the effect of communication delays on the convergence of certain distributed frequency regulation (DFR) protocols for prosumer-based energy systems, where prosumers are serving as balancing areas. DFR relies on iterative and distributed optimization algorithms to obtain an optimal feedback law for frequency regulation. But, it is, in general, hard to know beforehand how many iterations suffices to ensure stability. This paper develops a framework to determine a lower bound on the number of iterations required for two distributed optimization protocols. This allows prosumers to determine whether they can compute stabilizing control strategies within an acceptable time frame by taking communication delays into account. The efficacy of the method is demonstrated on two realistic power systems.**

***Index Terms*—Communication delays, Distributed frequency control, Inter-area oscillations, NERC reliability criteria, Prosumer, Power system frequency control.**

I. Introduction

In today's electricity industry, frequency regulation is performed by balancing areas in a unilateral way by taking into account local frequency measurements and power interchange deviations as part of Automatic Generation Control (AGC). Tie-lines are either neglected or modeled as static generators with output equal to active and reactive tie-line flows and voltage set-points equal to bus voltage magnitudes. Each balancing area is only able to control its own system footprint, but there is no real-time coordination between neighboring utilities/control-areas. This lack of coordination can create inter-area oscillation problems, which can cause system-wide instability and in the worst case scenario can lead to blackout [1], [2].

Regulating frequency in the future grid will be even more challenging. It is envisioned that in the future the grid will be populated by multiple electric sub-systems with local energy production as well as consumption. These so-called prosumers can produce, consume, and/or store electricity and make strategic decisions empowered by a cyber-layer superposed on top of the physical grid. Prosumers can be arranged in a hierarchical (nested) or flat organization. Prosumers can be as small as an electric vehicle (EV) or a smart building with home energy management systems, or as large as a utility or an independent system operator (ISO). All prosumers can share the operating task of balancing generation and demand at different time scales. But only large-scale prosumers are able to operate as a frequency control area. Indeed, small-scale prosumers, such as EVs, inside a utility-level prosumer cannot be considered as a balancing area. But, these prosumers can participate in frequency restoration by adjusting their set-points based on the overall policy of the larger prosumer. The main focus of this paper is overcoming communication delays for distributed frequency regulation or distributed power balancing between larger prosumers, which serve as balancing areas. Although the performance of the algorithms under consideration is important, the framework itself is the main contribution.

The state-of-the-art in frequency regulation or secondary frequency control adopts one of two different architecture designs: 1) the current approach, which is unilateral or fully decentralized; 2) centralized with wide area monitoring and (closed-loop) control systems (WAMCS). Under the decentralized architecture, each control area measures its local power deviations and deviations on tie-line flows and adjust internal frequency regulators in response to these deviations. Theoretically, this is similar to decomposing a large-scale optimal control problem into sub-problems and solving each sub-problem separately without considering coupling constraints [3].

The main drawback of the current approach is a lack of coordination between frequency regulators. In order to reduce this problem, different methods have been proposed, such as implementing power system stabilizers (PSS) and/or FACTS devices, to enhance the damping of oscillatory modes [4], [5], and [6]. Unfortunately, none of the proposed methods could guarantee frequency stabilization.

The second architecture for frequency regulation relies on WAMCS systems to collect information from different parts of the grid [7]. Although this architecture has the potential to obtain near-optimal control strategies without creating inter-area oscillations, it needs a large-scale centralized control/communication infrastructure and has a single point of failure, which makes the system vulnerable to cyber attacks. Note that currently WAMCS systems are only monitoring the state of the grid and not providing closed-loop automatic control.

To overcome these challenges, in [8], a distributed framework for frequency regulation of prosumer-based energy systems is proposed. Distributed frequency regulation (DFR) in smart power grids is a fairly new topic. In [9], a dis-

Thiagarajan Ramachandran, Santiago Grijalva, and Magnus Egerstedt are with the School of Electrical and Computer Engineering at the Georgia Institute of Technology, Atlanta, GA USA.

Masoud H. Nazari is with the Department of Electrical Engineering at California State University Long Beach, CA USA.

E-mail: tramachandran3@gatech.edu, masoud.nazari@csulb.edu, sgrijalva@ece.gatech.edu, and magnus@gatech.edu.

tributed model predictive control (MPC) framework is proposed for Automatic Generation Control (AGC) of power systems. In [10], a distributed control algorithm for frequency control of electrical power systems is presented. But, the DFR framework, proposed in [8] and extended in this paper, only requires one-hop communication between prosumers and allows prosumers to obtain optimal control strategies through a consensus-based ADMM (Alternating Direction Method of Multipliers) method.

However, the computational burden of the DFR algorithm increases as the system grows in scale and/or communication delays increases. In this paper, a new framework is developed to determine the lower bound on the number of iterations for obtaining guaranteed stabilizing control strategies for distributed frequency regulation.

The rest of the paper is organized as follows: in section II-A an overview of the one-step DFR framework is presented. This is followed by developing a new framework for computing an iteration budget for the DFR algorithm in Section III. The proposed framework is simulated on two realistic power systems in Section IV, and in Section V the paper concludes with discussions of the overall findings.

II. Overview of Distributed Frequency Regulation

A. DFR Framework

Frequency regulation includes bringing frequency deviations to the desired value, 60 or 50 Hz depending on the country [1], using minimal control effort. This is indeed an optimal control problem whose objective is to drive the power deviations to zero using minimal control effort. This paper bases its analysis on the frequency regulation problem for prosumer-based energy systems as it is formulated in [8]. For convenience, the problem is summarized as follows:

$$\min_{u} J(x(t_c), u) = \min_{u} \sum_{i \in N} p_i x_i(t_c+1)^2 + r_i u_i^2, \quad (1)$$

subject to coupling constraints

$$\begin{aligned} x_i(t_c+1) = \; & a_{ii}x_i(t_c) + b_{ii}u_i(t_c) \\ & + \sum_{j \in N_i} a_{ij}x_j(t_c) + b_{ij}u_j \end{aligned} \quad (2)$$

where $u = \left[u_1, \dots u_n\right]^T$ and $x(t_c) = \left[x_1(t_c), \; \dots x_n(t_c)\right]^T$, N is the set of all prosumers ($n = |N|$ is the number of prosumers), x_i and u_i are the power deviation and control variable of prosumer i, $P = diag(p_i)$ and $R = diag(r_i)$ are cost coefficients, and N_i is the set of prosumer i's neighbors. In addition, the system matrices are $A = [a_{ij}]$ and $B = [b_{ij}]$, which have the same sparsity structure as the Laplacian of the grid.

In today's industry, this problem is solved by neglecting coupling between prosumers. Therefore, the problem becomes much simpler and each prosumer solves its sub-problem in a fully decentralized way as:

$$\begin{aligned} \min_{u_i} \quad & p_i x_i(t_c+1)^2 + r_i u_i^2 \\ s.t. \quad & x_i(t_c+1) = a_{ii}x_i(t_c) + b_{ii}u_i \end{aligned} \quad (3)$$

As discussed in the preceding section, neglecting coupling can cause critical technical problems for the grid. Advanced frequency regulators take into account the effect of tie-line flows [3], but still neglect the effect of neighbors' control strategy.

In [8], a distributed framework for frequency regulation is proposed, under which prosumers have a perception of the decision variables of their neighbors and through a consensus-based ADMM method they achieve agreement on their control strategy. This framework is denoted as "One-Step DFR", because only one-hop communication between prosumers is sufficient to achieve stabilizing optimal control strategies. Under the one-step DFR framework, the frequency regulation problem is recast as follows [8]:

$$\begin{aligned} \min_{U_1, \cdots U_n} \quad & \sum_{i=1}^{n} \left(p_i \left[A_i^T X_i + B_i^T U_i\right]^2 + r_i U_{ii}^2\right) \\ \text{s.t.} \quad & U_{ij} = U_{jj}, \; \forall i \in N, j \in \mathcal{N}_i, \end{aligned} \quad (4)$$

where A_i, B_i, and U_i are the ith rows of A, B and U matrices, where $U = [U_{ij}]$, $\forall i \in N$ and $\forall j \in \mathcal{N}_i \cup \{i\}$, and U_{ij} is the perception of prosumer i from the control action of its neighbor, prosumer j. In addition, X_i is a column vector, which includes x_j, $j \in \mathcal{N}_i \cup \{i\}$.

In order to solve the DFR problem, the constraints are augmented in the objective function and the ADMM method is used to produce the augmented Lagrangian function as [8], [11], and [12]:

$$\begin{aligned} \mathcal{L}_{\rho,i}(U_i, \bar{U}_i^h, \lambda_i^h) = \; & p_i \left[A_i^T X_i + B_i^T U_i\right]^2 + r_i U_{ii}^2 \\ & + {\lambda_i^h}^T (U_i - \bar{U}_i^h) + \frac{\rho}{2} \|U_i - \bar{U}_i^h\|_2^2, \end{aligned} \quad (5)$$

where, $\rho > 0$ is a given penalty factor, and $\bar{U}_i^h$ is a column vector, which includes the average control strategy of prosumer i and that of its neighbors, defined as:

$$\bar{U}_{ij}^h := \frac{\sum_{l \in \mathcal{N}_j \cup \{j\}} U_{lj}^k}{|\mathcal{N}_j| + 1}, \; \forall j \in \mathcal{N}_i \cup \{i\}. \quad (6)$$

In each iteration, prosumer i computes its optimal control strategy by solving a self-contained problem of the following form:

$$U_i^{h+1} = \underset{U_i}{\operatorname{argmin}} \; \mathcal{L}_{\rho,i}(U_i, \bar{U}_i^h, \lambda_i^h), \quad (7)$$

Next, prosumers share their perceptions with their neighbors and continue this process until errors in power deviations and errors in perceptions become smaller than a desired value.

As shown in [8], DFR is a distributed method for frequency regulation, which can guarantee system-wide stability using minimal control effort. It can also address inter-area oscillations problems because coupling between prosumers is considered in computing minimizing control strategies.

[1] Note that most countries operates at 50Hz.

TABLE I: Convergence of the DFR algorithm for three practical energy systems

Number of prosumers	Number of buses	Number of iterations	Duality gap (pu)
3 (Flores Island)	46	10	10^{-4}
10 (IEEE System)	24	47	10^{-4}
15 (Sao Miguel Island)	1900	151	10^{-4}

B. Modelling communication delays

DFR relies on communication between prosumers, which can potentially pose limitations for the convergence of the algorithm, if the communication delays become large. In general, the communication architecture in smart grids supports the functionalities of the DFR algorithm, as the cyber layer has a similar sparsity pattern as the prosumer-based power grid and it connects neighboring prosumers, which are located at seperate geographical regions. The communication networks for prosumer-based energy systems have three classes: ISO-level network for communication between independent system operators (ISO); utility-level network to connect various devices within utilities and interconnect neighboring utilities; and, micro-level network to form a backbone for communication between microgrids, facilities, homes, etc [13].

Timing is critical for communication between prosumers, particularly for supporting the functionalities of the DFR algorithm. In fact, NERC A1 criterion requires that a prosumer brings power and frequency deviations (area control error, ACE) to zero once every 10 minutes and NERC B2 criterion requires that a prosumer begins to return ACE to zero within 1 minute after the beginning of a disturbance [14]. These reliability criteria enforce a clear time limit for the DFR algorithm. If a prosumer cannot compute stabilizing control strategy within 1 minute, it is a violation, which can lead to system-wide stability problems.

Increasing the communication delay between prosumers increases the risk of violating the NERC reliability criteria as it slows down the convergence of the DFR algorithm (iterations take more time). It can also be noted as the size of the grid increases, the number of iterations required to reach the minimizing control action increases for the DFR algorithm. This has been illustrated in Table I for three practical prosumer-based energy systems with different size and connectivity.

IEEE and the International Electrotechnical Commission (IEC) have defined rigorous standards for communication delay requirements in smart grids in order to ensure reliable operation of the grid and avoid potential stability problems [15]. Table II illustrates a summery of the expected packet delays in different communication categories. It is shown that the communication networks are responsible for delivering diverse categories of messages. In addition, some of the messages, such as the monitoring and control information, have critical delay requirements [13].

In reality, the communication networks are not always able to meet the strict communication delay requirements of IEEE and IEC. For instance, experimental results on communication delays between substations, reported in [15] and [16] and summarized in Table III, show that in many scenarios the packet delays exceed the maximum required limit for the most critical messages.

TABLE II: IEEE and IEC standards for communication timing requirements in smart grids

Information category	Delay requirment
Protection	4 ms
Monitoring and control	16 ms
Medium speed control functions	100 ms
Slow speed auto-control functions	500 ms
Operations and maintenance	1 s

TABLE III: Experimental delay measurement for messages requiring immediate actions [15], [16], and [13]

Test scenario for critical messages	Delay rang (ms)
1	0.2 - 0.7
2	3.2 - 17
3	12 - 86
4	32 - 173
5	18 - 97

In order to overcome the limitations of the communication delays, in the next section, a new framework is proposed to estimate a lower bound on the number of DFR iterations. This algorithm may not be the most efficient but it allows us to explicitly ask questions about the computation budget. The proposed method allows prosumers to obtain a system-dependent budget for DFR iterations.

III. Obtaining a Lower Bound for DFR Iterations

A. Steepest Descent-based DFR

In this section, a gradient descent-based approach is proposed to obtain the computation budget of DFR iterations. The general structure of the gradient descent-based DFR algorithm is formulated as follows.

$$u^{l+1} = -\frac{\gamma}{2}\frac{\partial J(x(t_c),u)}{\partial u} + u^l = Cx(t_c) + Du^l \qquad (8)$$

where γ_{t_c} is the step size at time t_c and C and D matrices are defined as:

$$C = -\gamma B^T P A \qquad (9)$$

$$D = I - \gamma\left(R + B^T P B\right) \qquad (10)$$

It follows from (8), (9), and (10) that the predicted control strategy is related to the square of the Laplacian of the grid, which implies that each prosumer needs to communicate with its neighbors and neighbors' neighbors to estimate its control strategy for the next step.

Equation (11) shows the predicted control strategy for two

iterations and L iterations scenarios.

$$\begin{aligned} u(t_c)^2 &= (C+DC)x(t_c) + D^2u(t_c-1) \\ &\vdots \\ u(t_c)^L &= \sum_{j=0}^{L-1} D^jCx(t_c) + D^Lu(t_c-1), \end{aligned} \tag{11}$$

In order to determine how many iterations need to be taken to obtaining stabilizing control strategies, one needs to first calculate power deviations at time t_c+1 based on the predicted control strategy at time t_c

$$x(t_c+1) = \left(A + B\sum_{j=0}^{L-1} D^jC\right)x(t_c) + BD^Lu(t_c-1), \tag{12}$$

combining (11) and (12) leads to:

$$\begin{aligned} \begin{bmatrix} x(t_c+1) \\ u(t_c) \end{bmatrix} &= A_{steepest}\begin{bmatrix} x(t_c) \\ u(t_c-1) \end{bmatrix} \\ &= \begin{bmatrix} A + B\sum_{j=0}^{L-1} D^jC & BD^L \\ \sum_{j=0}^{L-1} D^jC & D^L \end{bmatrix}\begin{bmatrix} x(t_c) \\ u(t_c-1) \end{bmatrix}, \end{aligned} \tag{13}$$

where $A_{steepest}$ is defined as the composite system matrix whose spectral properties determines the stability of the system.

In fact, it is possible to define a formal characterization of the sufficient number of iterations based on the stability of the composite system matrix. The number of iterations (L) needs to be large enough to satisfy the following condition:

$$\left|eig(A_{steepest}(L))\right| < 1 \tag{14}$$

B. Stability Condition for Steepest Descent-based DFR

Since (14) guarantees the stability of the composite system, it is important to first understand whether there exists any L for which the stability condition holds. Intuitively, it should be possible to find L if the optimal solution to DFR stabilizes the system. Note that when performing gradient descent starting from an arbitrary initial point, the distance to the optimal solution after L iterations depends on the step size and the distance between the initial estimate and optimal solution.

The following theorem shows that as long as the step-size for the gradient descent process is chosen appropriately, there exists L such that the composite system is always stable regardless of how the initial estimates for the gradient descent process is chosen.

Theorem 1. *Recalling from* (9) *and* (10)*, if γ is such that the spectral radius $\rho(D) < 1$ and the spectral radius $\rho(I - B(R + B^TPB)^{-1}B^TP)A) < 1$, there exists L such that $\rho(A_{steepest}(L)) < 1$.*

Note that the spectral radius of D determines the stability of the gradient descent process, while the spectral radius of $I - B(R + B^TPB)^{-1}B^TP)A$ determines the stability of the closed-loop system.

Proof. The key idea behind the proof lies in the following observation. Since $\rho(D) < 1$, the expression $\sum_{i=0}^{L-1} D^i$ corresponds to a convergent geometric sum and therefore converges to $(I-D)^{-1}$ as L approaches ∞. The same assumption also implies that D^L must converge to 0.

Therefore,

$$\begin{bmatrix} A + B\sum_{j=0}^{L-1} D^jC & BD^L \\ \sum_{j=0}^{L-1} D^jC & D^L \end{bmatrix} \rightarrow \begin{bmatrix} A + B(I-D)^{-1}C & 0_{n\times n} \\ (I-D)^{-1}C & 0_{n\times n} \end{bmatrix} \tag{15}$$

as L approaches ∞, where $(I-D)^{-1} = \frac{1}{\gamma}(R + B^TPB)^{-1}$. Substituting for $(I-D)^{-1}$ and C in (15), we obtain

$$A_\infty = \begin{bmatrix} A - B(R + B^TPB)^{-1}B^TPA & 0_{n\times n} \\ -(R + B^TPB)^{-1}B^TPA & 0_{n\times n} \end{bmatrix} \tag{16}$$

The eigenvalues of the block lower triangular matrix A_∞ are the eigenvalues of $A - B(R+B^TPB)^{-1}B^TPA$ and $0_{n\times n}$ (due to the zero matrix on the bottom right corner of A_∞). Thus, all the eigenvalues of A_∞ are contained in the unit circle as the spectral radius of $A - B(R+B^TPB)^{-1}B^TPA$ is less than 1 by assumption [8].

This implies that $\rho(A_\infty) < 1$. Since the spectral radius of a matrix is a continuous function of its entries, we have $\rho(A_{steepest}(L)) \rightarrow \rho(A_\infty) < 1$. This shows that for large enough L, $\rho(A_{steepest}(L)) < 1$.

□

The above theorem guarantees that as long as a "large enough" L budget is chosen, the system represented by (13) would be asymptotically stable. However, the size of the L budget is quite dependent on the optimization procedure and the spectral characteristics of the system matrix (A). This implies that the L budget, found by the gradient descent-based method, can be conservative for many real-world power grids.

In the next section, an alternative and much faster approach is proposed, called Nesterov's accelerated method, which converges to the optimal solution with a quadratic rate as opposed to the steepest descent, which has linear convergence.

C. Nesterov's Accelerated-based DFR

In this section, the Nesterov's accelerated gradient descent method is applied to obtain an effective L budget for the DFR algorithm. The Nesterov method is a variation of the gradient descent, which uses a variable step-size to accelerate convergence.

The following equations outline the theory of the accelerated gradient method:

$$y^{l+1} = u^l - \gamma\nabla J(u^l) \tag{17}$$

$$u^{l+1} = \eta_l y^l + (1-\eta_l)y^{l+1} \tag{18}$$

where $u^l, y^l \in \mathbb{R}^n$ (n is the dimension of the system), γ is

the step-size and the sequence η_l is defined as:

$$\eta_0 = 0 \tag{19}$$

$$a_l = \frac{1+\sqrt{1+4\eta_{l-1}^2}}{2} \qquad \eta_l = \frac{1-a_{l-1}}{a_l} \tag{20}$$

The process is initialized such that $y^0 = u^0$. Next, it will be shown that y^l converges to the minimum of the DFR cost function (J) for all initial estimates u^0. Recalling from (8), the gradient of the cost function is recast as:

$$\nabla J(x(t_c), u) = B^T P A x(t_c) + (R + B^T P B)u \tag{21}$$

Therefore, the equations (17) and (18) can be re-formulated as follows:

$$y^{l+1} = Du^l + Cx(t_c) \tag{22}$$

$$u^{l+1} = \eta_l y^l + (1-\eta_l)y^{l+1} \tag{23}$$

Equations (22) and (23) constitute a time-varying linear system, which can be expanded as:

$$\begin{bmatrix} w^{l+1} \\ y^{l+1} \\ u^{l+1} \end{bmatrix} = M_l \begin{bmatrix} w^l \\ y^l \\ u^l \end{bmatrix} + N_l x(t_c) \tag{24}$$

where the state variable w_l is used to keep track of previous values of y^l. In addition, matrices M_I and N_1 are defined as:

$$M_l = \begin{bmatrix} 0_{n\times n} & I_n & 0_{n\times n} \\ 0_{n\times n} & 0_{n\times n} & C \\ \eta_l & 0_{n\times n} & (1-\eta_l)C \end{bmatrix} \tag{25}$$

and

$$N_l = \begin{bmatrix} 0_{n\times n} \\ D \\ (1-\eta_l)D \end{bmatrix}. \tag{26}$$

Using (22) to (26), the Nesterov method can be casted as a linear time-varying system driven by a constant input $x(t_c)$. The response of such a system at time L is given by

$$\begin{bmatrix} w^L \\ y^L \\ u^L \end{bmatrix} = \Phi(0,L) \begin{bmatrix} w^0 \\ y^0 \\ u^0 \end{bmatrix} + F_L x(t_c) \tag{27}$$

where $\Phi(0,L)$ is the state transition matrix and F_L is the discrete time convolution operator, defined as follows:

$$\Phi(0,L) = M_{L-1}M_{L-2}\dots M_0 I_n \text{ when } L > 1 \tag{28}$$

$$F_L = \sum_{k=0}^{L-1} \prod_{i=0}^{L-1-k} M_{L-1-i}N_k \tag{29}$$

Recalling from Section III.A, the composite system matrix for the Nesterov-based DFR algorithm takes the following structure:

$$\begin{aligned} \begin{bmatrix} x(t_c+1) \\ u(t_c) \end{bmatrix} &= A_{Nesterov} \begin{bmatrix} x(t_c) \\ u(t_c-1) \end{bmatrix} \\ &= \begin{bmatrix} A+BPF_L & BT\Phi(0,L)G \\ PF_L & T\Phi(0,L)G \end{bmatrix} \begin{bmatrix} x(t_c) \\ u(t_c-1) \end{bmatrix} \end{aligned} \tag{30}$$

where

$$T = \begin{bmatrix} 0_{n\times n} & I & 0_{n\times n} \end{bmatrix} \tag{31}$$

and

$$G = \begin{bmatrix} I \\ 0_{n\times n} \\ I \end{bmatrix}. \tag{32}$$

The matrix G is used to generate initial conditions for the Nesterov's update equations and T is used to recover the vector of interest (i.e y^L). Since y^L converges to the minimizer of the DFR cost function as L approaches ∞, it can be shown that for large enough L, the composite system will stabilize to the origin. The system given in (30) is asymptotically stable if the spectral radius of the Nesterov's composite matrix is less than 1. Equation (33) illustrates a formal characterization for the stability of the Nesterov method.

$$\rho\left(\begin{bmatrix} A+BPF_L & BT\Phi(0,L)G \\ PF_L & T\Phi(0,L)G \end{bmatrix}\right) < 1 \tag{33}$$

IV. Simulation Results

In this section, the Steepest descent-based and Nesterov's accelerated-based DFR algorithms are demonstrated on two practical power systems. The first system is the electric power system on Sao Miguel Island, the capital of Azores Archipelago, and the second system is the IEEE 24-bus system. The results show that the L budget depends on the optimization procedure, the spectral characteristics of the system matrix, and the size of the grid.

A. Computing L budget for Sao-Miguel Island

Sao Miguel is the largest and capital of Azores Archipelago, islands of Portugal. The electric power system on Sao Miguel has more that 2000 lines, around 1900 buses, and 15 generators. The average demand of the island is 70 MW. The detailed description of the Sao Miguel system is presented in [17], [18].

In this paper, the power system of Sao Miguel is clustered into a prosumer-based structure, where each prosumer represents a control area for frequency regulation. Figure 1 illustrates the schematics of the equivalent power grid on Sao Miguel, in which each node represents a prosumer, which has a generator and a load. The loads are representing the equivalent demand on the prosumers [19].

The DFR cost is chosen such that the minimizer to the cost, which takes on the form of $u^* = -Kx$, becomes an stabilizing control strategy for the system. The state transition matrix and the convolution matrix corresponding to the gradient descent process for different values of L are computed using the following recursive equations:

$$\Phi(0,k) = D\Phi(0,k-1) \tag{34}$$

$$F_k = DF_{k-1} + C \tag{35}$$

As shown in Figure 2, it takes at least 4300 iterations for the gradient descent process to obtain a stabilizing control strategy, which can bring the spectral radius of $A_{steepest}(L)$ to less than 1. On the other hand, for the Nesterov's accelerated method the number of iterations (shown in Figure 3) is drastically less

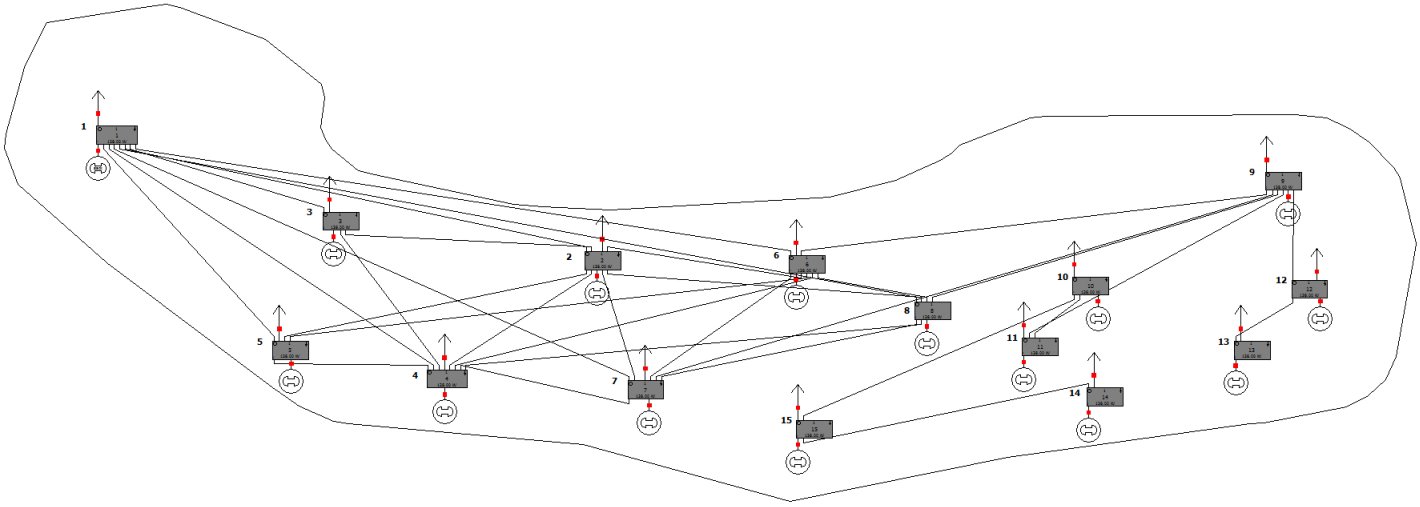

Fig. 1: Schematics of the equivalent power grid on Sao Miguel Island

(L is approximately 700) due to the quadratic convergence of the Nesterov process[2].

Assuming that the expected communication delay for each iteration is δ, by increasing the number of iterations the overall time[3] taken to attain a stablizing controller is $\Delta = L\delta$.

Recalling from Section II-A, NERC B2 criterion requires that prosumers start regulating frequency within 1 minute after the disturbance. Therefore, if $\Delta \geq 1$ minute, prosumers will violate the NERC reliability criteria and the DFC algorithm will fail to converge. Computing L allows prosumers to estimate whether they are able to stabilize power and frequency deviations after any arbitrary perturbations within the acceptable time window.

B. Computing L budget for the IEEE 24-bus system

The next case study is the IEEE 24-bus system, which has 38 lines and 32 generators. The average demand of the system is 2,577 MW. The detailed description of the IEEE 24-bus system is presented in [20].

The power system is clustered into 10 prosumers, where each prosumer represents a utility or area balancing authority. Figure 4 illustrates the schematics of the power grid of the IEEE 24-bus system and Figure 5 demonstrates the cyber-physical network of the prosumer-based IEEE 24-bus system. It is shown in Figure 5 that the cyber-layer has the same sparsity structure as the physical grid.

Figures 6 and 7 illustrate the results of applying the Steepest descent-based and Nesterov's accelerated-based DFR algorithms to the IEEE 24-bus system. It can be observed that the Nesterov's accelerated gradient method outperforms the gradient descent-based approach by a large margin.

[2]Unlike gradient descent, Nesterov's accelerated gradient descent is not a descent method and exhibits oscillations, called Nesterov's ripples, around the optimal solution. This is reflected in the osciallatory behaviour of the spectral radius of the $A_{Nesterov}$ matrix.

[3]Note that the computation delays are considered to be negligible compared to the communication delays. This is a reasonable assumption as the update law just requires taking linear combinations of state measurements, which can be done quite quickly.

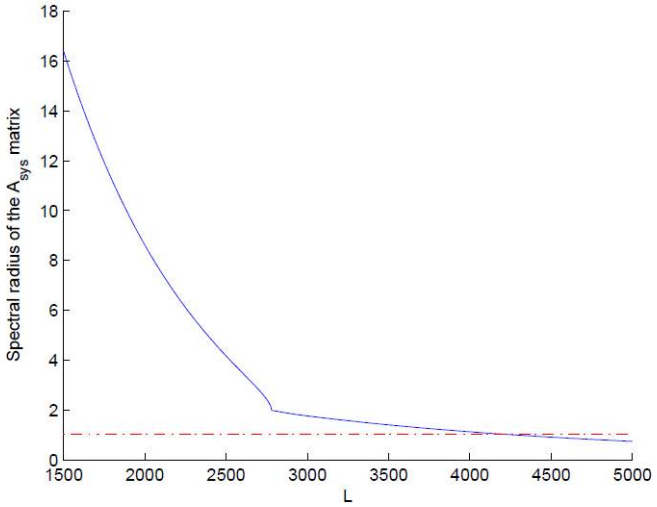

Fig. 2: The plot of the spectral radius of the $A_{steepest}$ matrix for the Sao-Miguel island system.

Note that, the number of iterations required by the gradient descent DFR for the IEEE 24-bus system is approximately 3300, while it takes more than 4300 iterations to find an stabilizing control strategy for the Sao Miguel system. This is mainly due to the fact that the IEEE system has 10 prosumers and the Sao Miguel Island has 15 prosumers. A similar trend can be observed when comparing Nesterov's accelerated method for the two test systems.

The findings also show that the Nesterov method provides an acceptable lower bound for the computation budget of the test systems. Assuming that the cyber networks satisfy the communication delay requirements of IEEE ($\delta < 16$ ms), the overall delay for the convergence of DFR for both systems would be with in the acceptable time window ($\Delta < 11.2$ s for Sao Miguel and $\Delta < 5$s for the 24-bus system).

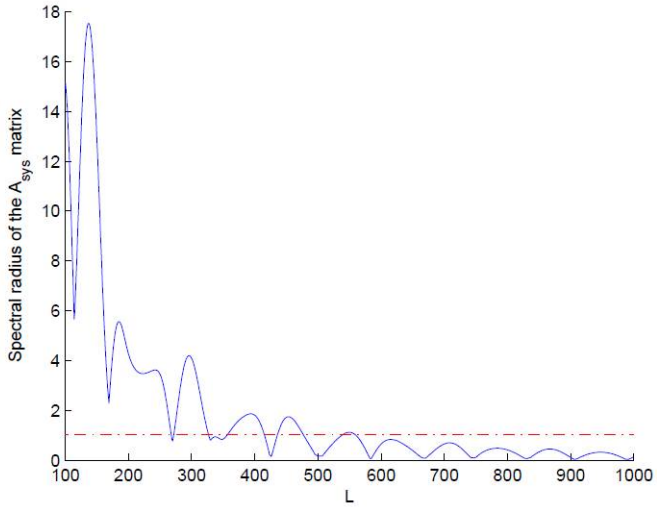

Fig. 3: The plot of the spectral radius of the $A_{Nesterov}$ matrix for the Sao-Miguel island.

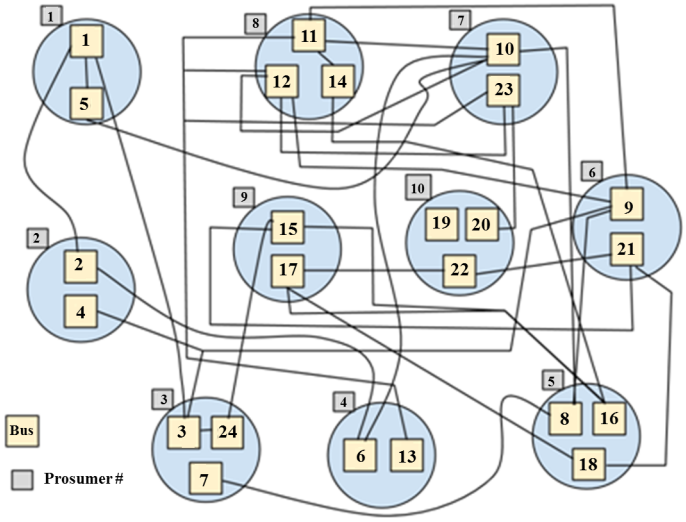

Fig. 5: Schematics of the cyber-physical grid of the prosumer-based IEEE 24-bus system

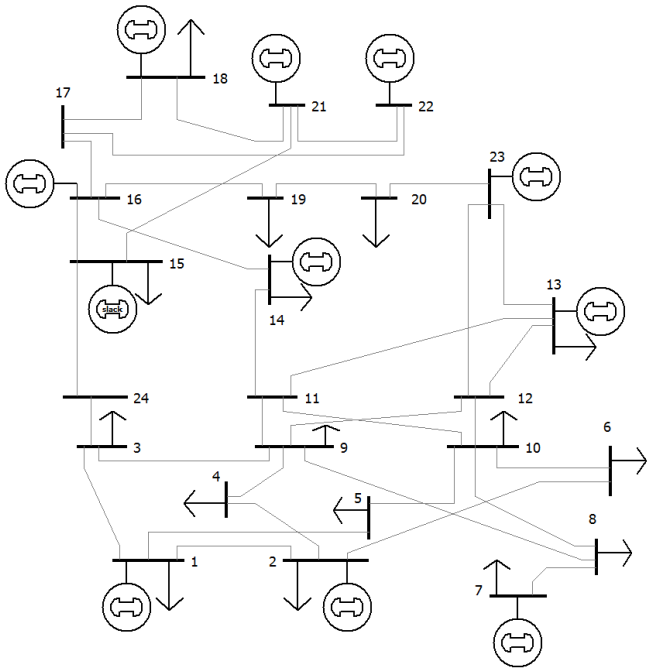

Fig. 4: Schematics of the power grid of the IEEE 24-bus system [20]

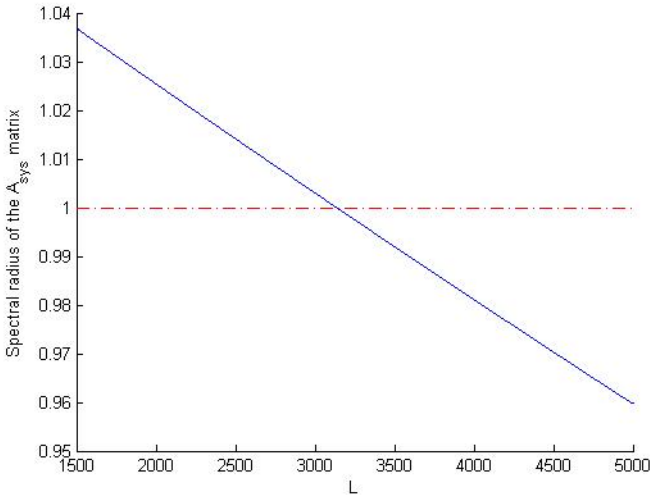

Fig. 6: The plot of the spectral radius of the $A_{steepest}$ matrix for the IEEE 24-bus system.

C. Frequency Regulation Performance

The spectral radius of the $A_{steepest}$ or $A_{Nesterov}$ matrices determines the rate at which the power deviations decay down to zero (closer the spectral radius is to 1, slower the convergence) and has a direct impact on the performance. At an execution level, each prosumer improves its initial estimate of the control action by executing L steps of a pre-determined optimization protocol and then applies the improved control action to the system. This process is repeated until the power deviations, and consequently the frequency deviations, are reduced to zero. The simulation results presented thus far demonstrates that the spectral radius of $A_{steepest}$ and $A_{Nesterov}$ depends directly on the number of iterations (L) spent improving the initial estimate.

In this section, further simulation results are presented to illustrate and compare the performance of the optimization protocols at different values of L on the IEEE 24-bus system. The dynamics used to simulate the evolution of the power deviations and the input vector is given by (13) (steepest descent) and (30) (Nesterov's accelerated gradient descent). The plots presented in this section track the evolution of power deviations assuming a communication delay δ of 16 ms.

According to Figure 7, the spectral radius of the $A_{Nesterov}$ matrix for the IEEE 24-bus system dips below 1 for $L = 560$. Figure 8 shows the norm of the power deviation when the number of optimization steps used to compute the control action is $L = 560$ when the optimization protocol used is Nesterov's accelerated gradient descent. It takes $L\delta = 8.96$

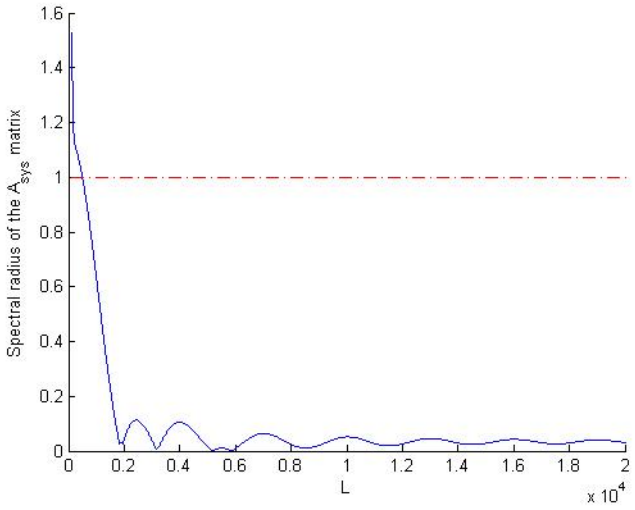

Fig. 7: The plot of the spectral radius of the $A_{Nesterov}$ matrix for the IEEE 24-bus system.

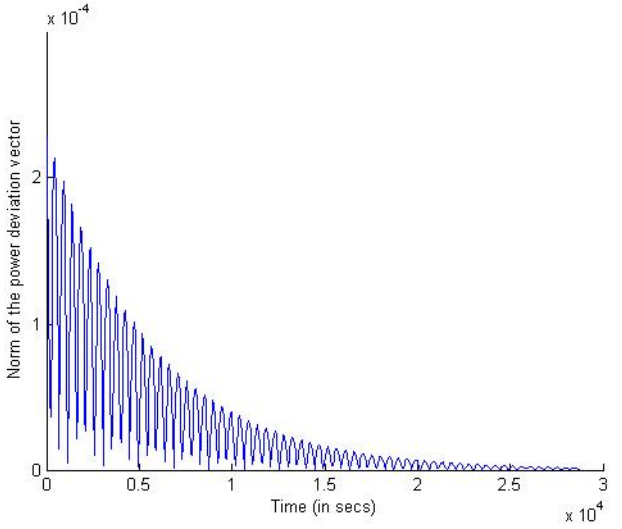

Fig. 9: Plot of power deviations ($L = 3600$, optimization procedure = steepest descent)

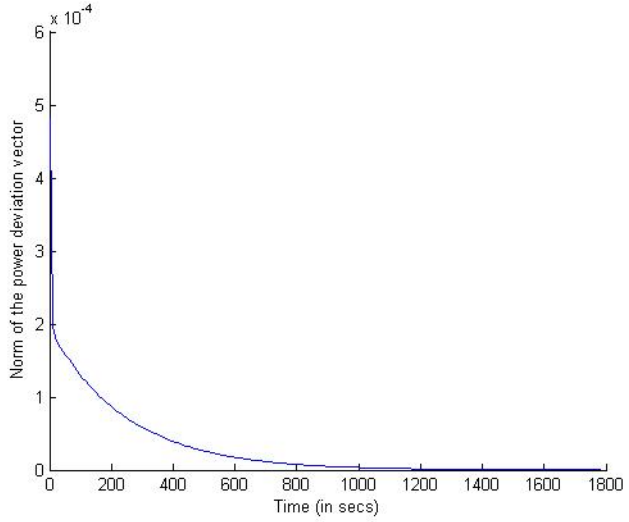

Fig. 8: Plot of power deviations ($L = 560$, optimization procedure = Nesterov's accelerated descent)

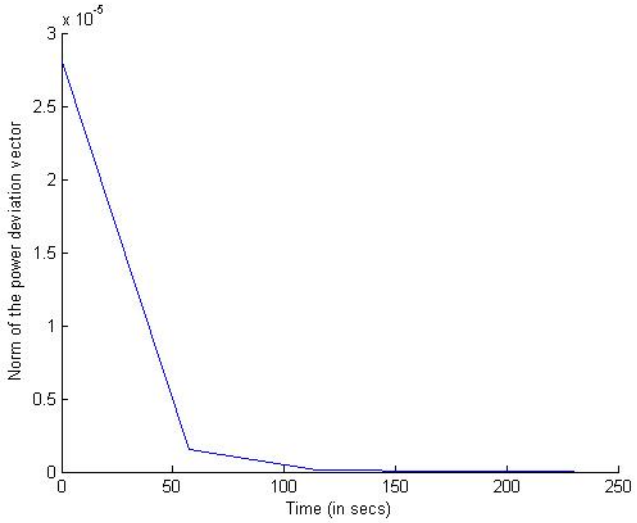

Fig. 10: Plot of power deviations ($L = 3600$, optimization procedure = Nesterov's accelerated descent)

seconds to compute a control action. The largest eigenvalue of the matrix $A_{Nesterov}$ determines the rate at which the power deviations converge to zero. This can be seen in Figure 8. The spectral radius of $A_{Nesterov}$ is 0.9646 and as such it takes roughly 1200 seconds (20 minutes) for the system to converge to zero.

Figure 9 is a plot of the power deviations when the number of optimization steps is $L = 3600$ when using the steepest descent method. Figure 10 shows a plot of power deviations when the number of optimization steps is $L = 3600$, but for Nesterov's gradient descent method. With $L = 3600$, the delay between the application of control actions is $L\delta = 57.6$s which is still within the one minute limit, imposed by the NERC criteria.

Note that it takes about 500 minutes (8.5 hrs) for the power deviations to stabilize to zero when using steepest descent (Figure 9) as opposed to 4 minutes (approximately) required by Nesterov's accelerated gradient descent (Figure 10). This is due to the fact that the spectral radius of the $A_{steepest}$ matrix is 0.9898, which slows down the convergence. The oscillatory behavior is due to the fact that the largest eigenvalue of the system matrix $A_{steepest}$ happens to be complex when using steepest descent. The fast convergence exhibited in (Fig 10) is due to the extremely small spectral radius of the $A_{Nesterov}$ matrix (0.0783) when $L = 3600$.

It should be stressed that the 8.5 hrs required for the steepest descent algorithm to stabilize is absolutely not realistic. Nesterov's accelerated gradient algorithm, on the other hand, does stabilize sufficiently fast. However, this paper does not claim that other algorithms, such as ADMM, would not do better. But, the paper has indeed shown that it is possible to connect the computing budget (L) to the system performance

in an explicit way. The search for even better algorithms will be left to future endeavors.

V. Conclusions

This paper introduces a novel method to estimate a lower bound for the computation budget of the DFR algorithm. Under the proposed method, prosumers are able to predict how many iterations they need to take to obtain a stabilizing control strategy in a distributed manner. Note that, knowing this information by itself does not ensure stability. Only after executing and implementing the optimal control action, frequency and power stability can be achieved.

The paper shows that the computation budget depends on the optimization method; the size of the grid, and; the spectral characteristics of the system matrix. In fact, the Nesterov accelerated-based method has faster convergence rate comparing with the gradient descent-based approach and therefore can provide an appropriate lower bound for the L budget.

In addition, the IEEE and IEC standards impose strict communication delay requirements, which can satisfy the expected requirements of the DFR algorithm. The proposed method allows prosumers to overcome these challenges by estimating whether they are able to achieve convergence within the acceptable time window. The method is simulated on two realistic power systems. The results show that the lower bound computed by the Nesterov method can satisfy the NERC reliability criterion, assuming that the communication networks satisfy the requirements of IEEE.

The findings of this study gives rise to several new research questions. For instance, what is the impact of large disturbances such as generators/prosumers failure and/or lines disconnection on the behavior of DFR? How can prosumers be persuaded to participate in the DFR framework? Do prosumers need to see it as a system responsibility or should a policy mechanism be implemented to incentivize them? Do prosumers with renewable energy sources require operating below potential to release some regulating reserve? If so, what mechanisms should be implemented to cover the economic losses? These interesting questions are out of the scope of this paper and should be part of future research endeavors.

Acknowledgments

The financial support of this project is provided by the Department of Energy through Advanced Research Projects Agency-Energy (ARPA-E) program. The authors greatly appreciate this support.

References

[1] N. Popli and M. Ilic. *Modeling and Control Framework to Ensure Intra-dispatch Regulation Reserves, in Chapter 14 of Engineering IT-Enabled Sustainable Electricity Services*. Springer, 2013.

[2] M. Ilic, H. Allen, W. Chapman, C. King, J. Lang, and E. Litvinov. Preventing Future Blackouts by Means of Enhanced Electric Power Systems Control: From Complexity to Order. *Proceedings of The IEEE*, 93:1920–1941, 2005.

[3] M. Ilic, X. Liu, B. Eidson, C. Vialas, and M. Athans. A structure-based modeling and control of electric power systems. *Automatica*, 33(4):515 – 531, 1997.

[4] M. Klein, G. Rogers, and P. Kundur. A fundamental study of inter-area oscillations in power systems. *Power Systems, IEEE Transactions on*, 6(3):914–921, 1991.

[5] D. Ke, C. Chung, and X. Yusheng. An eigenstructure-based performance index and its application to control design for damping inter-area oscillations in power systems. *Power Systems, IEEE Transactions on*, 26(4):2371–2380, 2011.

[6] D. Rai, S. Faried, G. Ramakrishna, and A. Edris. Damping inter-area oscillations using phase imbalanced series compensation schemes. *Power Systems, IEEE Transactions on*, 26(3):1753–1761, 2011.

[7] R. Avila-Rosales and J. Giri. Wide-area monitoring and control for power system grid security. *In Proc. PSCC*, pages 1–7, 2005.

[8] M. H. Nazari, Z. Costello, M.J. Feizollahi, S. Grijalva, and M. Egerstedt. Distributed frequency control of prosumer-based electric energy systems. *IEEE Transactions on Power Systems*, 29(6):2934–2942, Nov 2014.

[9] A.N. Venkat, I.A. Hiskens, J.B. Rawlings, and S.J. Wright. Distributed mpc strategies with application to power system automatic generation control. *Control Systems Technology, IEEE Transactions on*, 16(6):1192–1206, Nov 2008.

[10] M. Andreasson, D.V. Dimarogonas, K.H. Johansson, and H. Sandberg. Distributed vs. centralized power systems frequency control. In *Control Conference (ECC), 2013 European*, pages 3524–3529, July 2013.

[11] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 3(1):1–124, 2011.

[12] M. J. Feizollahi, M. Costley, S. Ahmed, and S. Grijalva. Large-scale decentralized unit commitment. *International Journal of Electrical Power & Energy Systems*, 73:97–106, 2015.

[13] Wenye Wang, Yi Xu, and Mohit Khanna. A survey on the communication architectures in smart grid. *Computer Networks*, 55(15):3604 – 3629, 2011.

[14] Howard F Illian. Frequency control performance measurement and requirements. *Lawrence Berkeley National Laboratory*, 2011.

[15] Veselin Skendzic and A Guzma. Enhancing power system automation through the use of real-time ethernet. In *Power Systems Conference: Advanced Metering, Protection, Control, Communication, and Distributed Resources, 2006. PS'06*, pages 480–495. IEEE, 2006.

[16] IEC 61850-5. Communication networks and systems in substations–part 5: Communication requirements for functions and device models. 2003.

[17] M. H. Nazari. *Electrical Networks of the Azores Archipelago, in Chapter 3 of Engineering IT-Enabled Sustainable Electricity Services*. Springer, 2013.

[18] M. H. Nazari. *Small-Signal Stability Analysis of Electric Power Systems on the Azores Archipelago, in Chapter 17 of Engineering IT-Enabled Sustainable Electricity Services*. Springer, 2013.

[19] M. H. Nazari. *Making the Most out of Distributed Generation without Endangering Normal Operation: A Model-Based Technical-Policy Approach*. Phd thesis, Carnegie Mellon University, 2012.

[20] Cliff Grigg, Peter Wong, Paul Albrecht, Ron Allan, Murty Bhavaraju, Roy Billinton, Quan Chen, Clement Fong, Suheil Haddad, Sastry Kuruganty, et al. The ieee reliability test system-1996. a report prepared by the reliability test system task force of the application of probability methods subcommittee. *Power Systems, IEEE Transactions on*, 14(3):1010–1020, 1999.