Reduced-Order Models of Static Power Grids based on Spectral Clustering

Mario D. Baquedano-Aguilar*, Sean Meyn*, Arturo Bretas†‡
*Department of Electrical & Computer Engineering, University of Florida, Gainesville, FL, USA
†Distributed Systems Group, Pacific Northwest National Laboratory, Richland, WA, USA
‡G2Elab, Grenoble INP, CNRS, Université Grenoble Alpes, 38000 Grenoble, France
m.baquedanoaguil@ufl.edu, meyn@ece.ufl.edu, arturo.bretas@pnnl.gov

2023 North American Power Symposium (NAPS) | 979-8-3503-1509-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/NAPS58826.2023.10318799

***Abstract*—For large-scale interconnected power systems that cover large geographical areas, certain electrical studies are required so that appropriate decisions ensure system reliability and low cost. For such studies, it is often neither practical nor necessary to model in detail the entire power system, which is increasingly complex due to a more diverse range of grid assets to choose from in both short and long-term planning. The goal of this paper is to present a methodology to reduce the order of large-scale power networks based on spectral graph theory given that current methods for static network reduction are not scalable. A brief analysis of some spectral clustering properties to determine which graph Laplacian matrix should be used and why is included. The analysis shows that the utilization of the normalized graph Laplacian is more advantageous for clustering purposes. Techniques are proposed to approximate cost functions for the aggregated generators. This is done via linear regression. The reduced-order model obtained with the proposed methodology has an accuracy above 94% and solves the scalability issue commonly present in other reduction methods. If the utilization of the reduced-order model is either constrained to load levels above mid-peak demand, or cost functions of aggregated units are approximated via a piece-wise quadratic approach, then the error distribution is in the order of 10^{-3}.**

***Index Terms*—Aggregation of electrical components, function approximation, linear regression, network reduction, spectral graph theory, unit commitment and economic dispatch.**

I. Introduction

Modern electric power systems have increased in size and complexity resulting in faster dynamics due to the rapid growth of renewable generation and widespread of interconnections due to both decarbonization goals and energy independence [1]. Operational scheduling of such systems under uncertain scenarios is required for system reliability and low cost. This is increasingly complex due to a more diverse range of grid assets to choose from either in short or long-term planning [2]. For these reasons, grid operators need to perform several studies (e.g. stability and sensitivity analysis, resource adequacy, unit commitment with economic dispatch (UCED), among others) to consider a wide range of scenarios and make decisions to achieve acceptable reliability levels.

To perform studies for such systems, it is often neither practical nor necessary to model in detail an entire large-scale power network. When the full network model is considered, studies of the sort cannot be quickly performed because the appropriate tuning and validation can take a considerable amount of time [3].

The goal of this paper is to present a methodology to reduce the order of the network models so that operational scheduling and sensitivity studies can be performed relatively quick with accurate enough results as those obtained when using the full network model.

Network reduction is the process of reducing the complexity of a large power system model while (approximately) retaining its steady-state (static) and dynamic characteristics on the reduced system. In this work, we focus on retaining only the static characteristics of the full network model.

In the case of static network reductions two approaches can be highlighted: the Ward equivalent, and the radial, equivalent, and independent (REI) method. In the Ward equivalent, the original network is split into an internal and external system. The main issue is that the behavior of both the internal system (which is accurate) and the external one (which is approximated) cannot be simulated by the same algorithm process [4]. The REI method is a lossless network representation of a set of zero power balance networks, i.e. total demand equals total generation. The issue is that it is a highly time-consuming procedure that works for small test power networks only [5].

Although both of these methods have been widely used for static networks reduction, none of them are scalable, i.e., when large-scale power networks are considered for reduction, none of these methods can be applied [3]. To overcome the scalability issue, we present a methodology for reducing large-scale power networks based on spectral clustering [6]

Besides overcoming the aforementioned issues, other advantages of using the spectral clustering approach is that it needs no internal nor external systems, and neither prior knowledge of the network is required to reduce the order of the networks.

Although spectral clustering has been widely used in power systems to find intentional controlled islanding solutions with minimal power flow disruption [7]–[9], there is no agreement nor clear reasoning among researchers on which graph Laplacian matrix should be used and why [10]. In this work, we attempt to find an answer to this question.

In this work, spectral clustering approach is applied to use the islanding solution (clusters) to reduce the order of the original network via aggregation of electrical components. However, when several units are lumped into a single unit there is a modeling issue associated with the individual cost function of the original units, which are no longer useful to perform

UCED simulations. This implies that a new cost function must be approximated to properly represent the operational cost of the aggregated units in the reduced-order model. For this, different approximation function structures based on linear regression are tested. Once the "best" candidate cost functions are determined, a UCED simulation is performed on both networks, and the error distribution of the daily system cost is computed so that we can compare how good the reduced-order model is.

The specific contributions of this work towards the state-of-the-art are:

- *Normalized over Unnormalized Laplacians.* A brief analysis of spectral clustering properties is provided in Section II which clarifies why *normalized* graph Laplacians are preferred over *unnormalized* ones.
- *Scalability.* The presented methodology can be used for both small and large-scale power networks in contrast to the Ward equivalent and the REI method which only work for small networks.
- *Aggregation of Generators.* Our methodology highlights the importance of performing aggregations considering only units of the same type. Otherwise, static characteristics of the system are lost.

The remaining of the paper is divided as follows. Section II contains a brief description of spectral graph theory in the context of electric power systems. The aggregation of electrical components and the cost function approximation process of aggregated generators are the subject of Section III. Section IV presents a case study of the presented methodology, while conclusions and topics for future research are contained in Section V.

II. Spectral Clustering in Power Systems

A. Power Networks Representation

Power grids can be represented as an undirected graph, $G = (\mathcal{V}, E)$, where $\mathcal{V}$ represents the set of buses, and E represents the set of edges or electrical connections of the network.

The set of buses is defined as $\mathcal{V} = \{1, 2, ..., N\}$ where N is the number of buses, and the set of edges as $E \subset \mathcal{V} \times \mathcal{V}$, where $(i, j) \in E$ represents an edge, i.e., a transmission line or a transformer from bus i to j.

Edge Weights. An *edge weight* is a nonnegative function $\omega : \mathcal{V} \times \mathcal{V}$ such that

1) $\omega(i, j) = \omega(j, i)$ if $i \neq j$ (i.e, edge directions ignored)
2) $\omega(i, j) = 0 \quad$ if $(i, j) \notin E$
3) $\omega(i, i) = \sum_{j=1}^{N} \omega(i, j) = d_i$, is the *weighted vertex degree*.

The *degree matrix* D is defined as the diagonal matrix with weighted vertex degrees $d_1, \cdots, d_N$ on the diagonal, and the weighted adjacency matrix of the graph is the matrix $W = (\omega_{ij})_{i,j} = 1, \cdots, N$.

Edge Weights in Power Systems. Different edge weight functions have been used by researchers to study the functional structure of power systems [10]. For the purpose of this work, power flows through all transmission lines will be used as the edge weight function to create the power flow matrix, W to study the functional structure of the test system.

B. Graph Laplacians.

In the spectral clustering literature, the following three types of matrices that are commonly referred to as graph Laplacians are found [6], [11]:

- The *unnormalized* graph Laplacian, L, is defined as:

$$L = D - W \tag{1}$$

- The *normalized* graph Laplacians, L_n, and L_{rw}. Both matrices are closely related to each other and are defined as:

$$L_n = D^{-1/2} L D^{-1/2} = I - D^{-1/2} W D^{-1/2} \tag{2}$$

$$L_{rw} = D^{-1} L = I - D^{-1} W \tag{3}$$

The first matrix, L_n, is commonly denoted by L_{sym} as it is a symmetric matrix, and the second one is denoted by L_{rw} as it is closely related to a random walk. The eigenvalues of the normalized graph Laplacian (i.e., $\lambda^i_{L_n}$) satisfy the inequality $0 \leq \lambda^i_{L_n} \leq 2$ for all i [12].

In power systems, both *unnormalized* and *normalized* graph Laplacians matrices have been used. Apparently, the latter is preferred among researchers, but the reason of this preference is unknown. With the hope of finding a reasonable answer, we plotted the eigenvalues, the first eigenvectors, and the relative eigengaps to help us analyze which graph Laplacian matrix to use and why. The first row of Fig. 1 shows the eigenvalues and the first five eigenvectors of the *unnormalized* graph Laplacian L, while the second row shows the same parameters for the *normalized* graph Laplacian L_n. In the eigenvalue plot we plot i vs λ_i. In the eigenvector plots of an eigenvector $v = (v_1, \cdots, v_{39})'$ we plot x_i vs. v_i (note that x_i is simply a real number, hence we can depict it on the $x-$axis).

We can see in Fig. 1 that the *unnormalized* graph Laplacian, L, only consists of one connected component. Thus, eigenvalue 0 has multiplicity 1, and the first eigenvector, v^L_1, is the constant vector. The following eigenvectors carry the information about the clusters. For the *normalized* graph Laplacian, L_n, we can see that the first eigenvalue, $\lambda^{L_n}_1 = 0$, and the corresponding eigenvector, $v^{L_n}_1$, is a cluster indicator vector (always positive). The following eigenvectors carry the information about the clusters.

Up to this point, the reason for the preference on using the *normalized* Laplacian remains unclear. Because of this, we will consider analyzing other spectral clustering properties such as the similarity measure and the degree distribution, both of which are the subject of the next section.

C. Which graph Laplacian Should be Used?

To answer this question, at least one of the following options must be considered:

1) The degree distribution either of the similarity graph or the weighted adjacency matrix of the graph. In our case, we will look at the degree distribution of the latter.

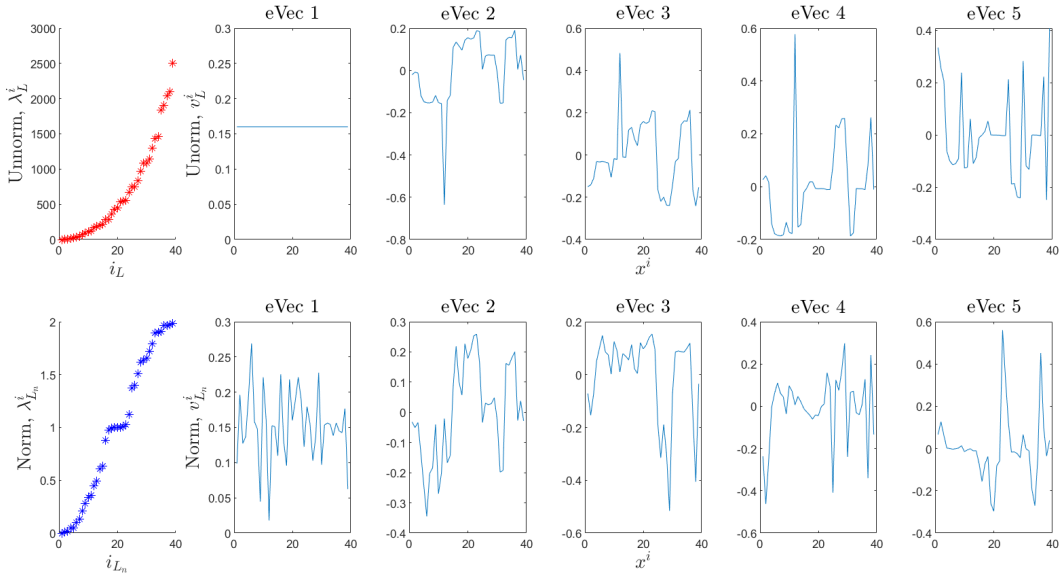

Fig. 1. Eigenvalues (λ^i) and eigenvectors (v^i) of the graph Laplacian matrices L (first row) and L_n (second row).

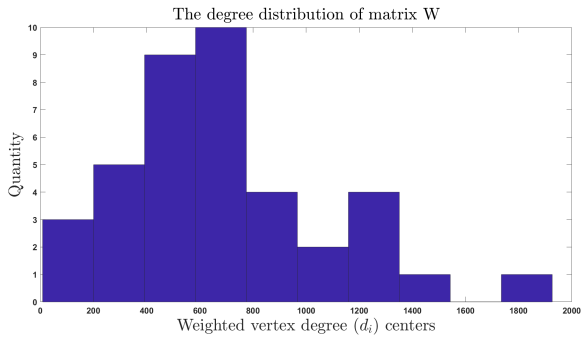

Fig. 2. The degree distribution of the power flow matrix, W.

2) The similarity measure of 1), which is obtained by computing the relative eigengaps with (4).

The Degree Distribution of the Power Flow Matrix. If the graph is very regular[1] and most vertices have approximately the same degree, then all the Laplacians are very similar to each other, and will work equally well for clustering. However, if the degrees in the graph are very broadly distributed, then the Laplacians differ considerably [6].Fig. 2 shows the degree distribution of the power flow matrix, W. As we can see, the degrees in the graph are broadly distributed so that the Laplacians differ considerably. This finding reinforces the importance of choosing the appropriate graph Laplacian matrix.

A Choice of Dimension and Similarity Measure. A common criterion for choosing the number of clusters is by computing the magnitude of the difference between two consecutive eigenvalues relative to their size. This parameter is also used as a similarity measure and is known as *the relative eigengaps*:

$$\gamma_{k,rel} = \frac{|\lambda_{k+1} - \lambda_k|}{\lambda_k} \quad (k \geq 2). \tag{4}$$

A good k-partition exists if the k-*th* eigenvalue of the graph Laplacian, λ_k, is small. A high value of $\gamma_{k,rel}$ indicates that the power network admits a good partition into at least k-zones, and that this will be revealed by the spectral embedding in dimension k.

Fig. 3 shows the relative eigengaps of L (left) and L_n (right) for the power flow based Laplacian of the IEEE 39-bus test system. The highest value of the relative eigengap for the *unnormalized* Laplacian resulted in $\gamma_{k,rel}^{L} \approx 1.2$, while for the *normalized* one resulted in $\gamma_{k,rel}^{L_n} \approx 2.5$.

The violation of the inequality constraint of (4) means that a poor similarity measure is obtained when using the *unnormalized* Laplacian. Because of this, we will use $\gamma_{k,rel}^{L_n} \approx 2.5$, which will be considered as a good similarity measure. This implies that the power network admits a good partition into -at least- $k \approx 3$ clusters, i.e., it can also be $k = 4, \cdots, k_{max}$. In this work, we use $k = 4$ based on the assumption that the system has four coherent groups of generators. However, determining k_{max} and generator coherency identification is out of the scope of this work. Interested readers are referred to [6], [13] for more information.

***Normalized* over *Unnormalized* Graph Laplacians.** Based on the analysis of both the degree distribution of the power flow matrix, W, and the poor similarity measure obtained with the *unnormalized* graph Laplacian, we will choose the *normalized* graph Laplacian L_n to be associated with the undirected weighted simple graph $G = (\mathcal{V}, E, \omega)$. Although the analyses done in this section may be problem dependent, we now have enough reasons to choose the *normalized* over the *unnormalized* graph Laplacian matrix as it appears to be more advantageous for clustering purposes.

[1]In graph theory, a regular graph is a graph where each vertex has the same number of neighbors; i.e. every vertex has the same degree.

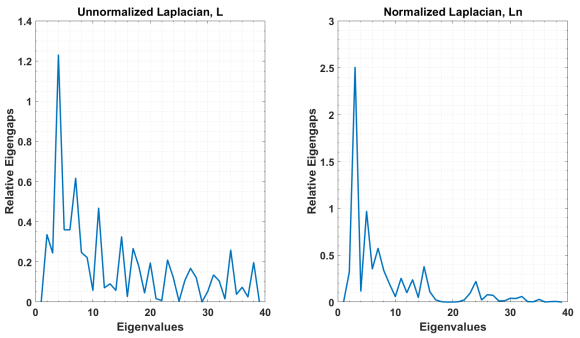

Fig. 3. The relative eigengap, $\gamma_{k,rel}$, of L (left) and L_n (right).

III. Aggregation of Electrical Components and Cost Function Approximations

A. Aggregation of Electrical Components

Building reduced-order models of power grids involves aggregation of loads and generators.

Loads. The aggregated load at time t is computed as the sum of the loads of all buses within each cluster as follows:

$$P^k_{D_{Agg}}(t) = \sum_{i=1}^{N^k_b} P_{D_i}(t); \quad t = \Big\{1, \cdots, 24\Big\} \tag{5}$$

where $P_{D_i}(t)$ is the demand of the ith bus at time t (in hours); $N^k_b \subset \mathcal{V}$ is the number of buses in cluster-k; and $P^k_{D_{Agg}}(t)$ is the total aggregated demand in cluster-k at time t.

Generating Units. For aggregation purposes, only generating units of the same technology will be lumped together to preserve their operational characteristics. For ease of explanation, let's consider two thermal units, G_{1A}, G_{1B}, both of which are of different size and cost functions. The minimum and maximum generation limits of the aggregated units are computed as follows:

$$P^{min}_{Agg,k} = min\Big\{P^{min}_{G_{1A}}, P^{min}_{,G_{1B}}\Big\}; P^{max}_{Agg,k} = \Big\{P^{max}_{G_{1B}} + P^{max}_{G_{1B}}\Big\} \tag{6}$$

However, there is an issue: *their individual cost functions are no longer useful since these units were merged*[2]. Cost approximation of an aggregate of loads is the subject of the next section.

B. Cost Function Approximation

The goal is to approximate total cost over 24 hours for generation units of similar type. Assumed given is a data set $\{G^i\,,\ C^i : 1 \le i \le N\} \subset \mathbb{R}^{24}$ in which G^i_t is generation at time t from the ith sample, and C^i_t is the corresponding cost over one hour (in practice this cost will be a sum of many terms, including fuel cost).

We take a statistical approach: given a function class $\{c^\theta : \theta \in \mathbb{R}^d\}$, with $c^\theta \colon \mathbb{R}^{24} \to \mathbb{R}$ for each θ, and an empirical loss function L, the approximating model is defined by c^{θ^*} with $\theta^* := \arg\min_\theta L(\theta)$. In the numerical results surveyed in Section IV we take a linear function class, and quadratic loss of the form

$$c^\theta = \theta^\intercal \psi\,, \quad L(\theta) = \frac{1}{2}\frac{1}{N}\sum_{i=1}^{N}[c^\theta(G^i) - C^i]^2\,, \quad \theta \in \mathbb{R}^d\,, \tag{7}$$

in which $\psi \colon \mathbb{R}^{24} \to \mathbb{R}^d$ is the vector of basis functions. In this special case the optimizer has the explicit form $\theta^* = \Sigma^{-1}_\psi b$ with

$$\Sigma_\psi := \frac{1}{N}\sum_{i=1}^{N}\psi(G^i)\psi(G^i)^\intercal\,, \qquad b := \frac{1}{N}\sum_{i=1}^{N}\psi(G^i)C^i$$

Three function classes with different dimensions (d) are evaluated:

- Single quadratic approximation with $d = 3$ (Class I).

$$\psi(g) = (1; \sum_t g_t; \sum_t g^2_t)\,, \tag{8}$$

- Multiple quadratics approximation with $d = 26$ *varying the linear term* (Class II).

$$\psi(g) = \big(1, \sum_{j=1}^{24} g^2_j, g_1, g_2, \cdots, g_{24}\big)^T \in \mathbb{R}^{26} \tag{9}$$

- Multiple quadratics approximation with $d = 26$ *varying the quadratic term* (Class III).

$$\psi(g) = \big(1, \sum_{j=1}^{24} g_j, g^2_1, g^2_2, \cdots, g^2_{24}\big)^T \in \mathbb{R}^{26} \tag{10}$$

For example, the approximating model of (7) for Class I gives:

$$c^\theta(g) = \sum_{t=1}^{24}[\theta_1 + \theta_2 g_t + \theta_3 g^2_t] \quad [\$/hr]$$

IV. Case Study

A. Spectral Clustering Setup

A modified version of the IEEE 39-bus test system [14] is used to test our network reduction methodology. Thermal units $G_{39}, G_{31}, G_{32}, G_{33}, G_{35}, G_{36}, G_{38}$ (see Table I) of the original system were modified to represent commercial thermal generators. The quadratic cost function includes the UC costs (including start-up, no-load, and shut-down costs) and dispatch costs of real thermal generators based on the cost coefficients of [15]. Additional generator attributes include ramp rates and minimum up/down times.

Due to lack of space and for ease of explanation, Table I shows a summary of the network partitioning solution for $k = 4$ based on the spectral clustering algorithm proposed in [10] using the power flow weighted adjacency matrix, W. The computation of the spectral clustering algorithm was performed on a desktop computer, AMD Ryzen 9 5900 12-Core Processor, 3.00 GHz, 64 GB RAM. The computational time for determining a partition solution for $k = 4$ using the 39 bus system is ≈ 1.7 s, and ≈ 2.15 s on the 118-bus system.

[2]Note that it cannot be assumed that the cost function of the aggregated units can be obtained by simply adding the individual cost functions of the generators in the clusters.

TABLE I
PARTITION SOLUTION APPLIED ON THE IEEE-39 BUS SYSTEM INTO FOUR CLUSTERS VIA SPECTRAL CLUSTERING.

Cluster No.	Buses	Units Type for Aggregation	No. of Units	Generating Units Capacity
C_1	1,2,3,25,30,37,39	Fuel Oil	2	$G_{39A} = 600$ MW $G_{39B} = 400$ MW
C_2	15,16,17,18,19,20,21, 22,23,24, 33,34,35,36	Fuel Oil	3	$G_{33A} = 600$ MW $G_{33B} = 110$ MW $G_{34} = 600$ MW
		Natural Gas	4	$G_{36A} = 200$ MW $G_{36B} = 600$ MW $G_{35A} = 600$ MW $G_{35B} = 200$ MW
C_3	26,27,28,29,38	None	0	$G_{38} = 600$ MW (Gas) $G_{38} = 350$ MW (Coal)
C_4	4,5,6,7,8,9,10,11 12,13,14,31,32	Natural Gas	2	$G_{31A} = 600$ MW $G_{31B} = 400$ MW
		Coal	2	$G_{32A} = 350$ MW $G_{32B} = 400$ MW

Once the $k = 4$ clusters are determined, the aggregation of electrical components is performed as explained in Section III-A. Fig. 4 shows the structure of the reduced-order model of the IEEE 39-bus system, where the number of buses were reduced to only 4; the total number of generators is shortened from 16 to 9; and the transmission lines are reduced to only five (5). Notice that the spectral clustering algorithm used in this work gives a clustering solution based on minimal power flow disruption [10], i.e., the transmission lines in the reduced-order model carrying the minimum power flow. However, should any specific transmission line needs to be left unchanged, this can be done by adding it as a constraint in the clustering algorithm.

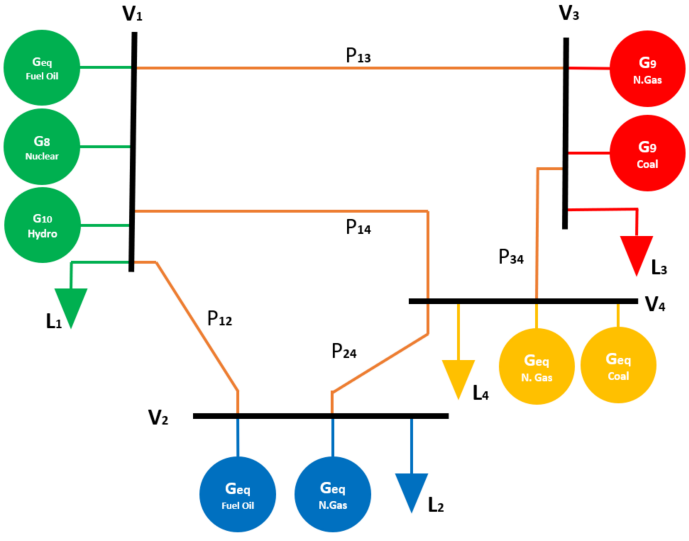

Fig. 4. Reduced-order model of the modified version of the IEEE 39-bus systems, 16 Generators via spectral clustering.

Given the reduced-order model of Fig. 4, we now need to approximate the cost function of the aggregated units following the procedure described in Section III-B.

Function classes I, II, and III were obtained from a training data set and validated with testing data of the same size (50 days). The mean-square error (MSE) distribution by function class for both the training and testing data sets are summarized in Table II, where $|\cdot|$ represents the normalized MSE values, respectively.

TABLE II
MSE BY CLASS: TRAINING DATA (TR) VS TESTING DATA (TE)

	Training Data (Tr)		Testing Data (Te)	
Approach	MSE_{Tr}	$\|\text{MSE}_{\text{Tr}}\|$	MSE_{Te}	$\|\text{MSE}_{\text{Te}}\|$
Fuel Oil Units: Cluster 1				
Class I	212389891.76	0.9976	257993965.21	0.99
Class II	10710425.91	0.05031	32996877.48	0.13
Class III	10005198.54	0.04699	22742746.27	0.09
Fuel Oil Units: Cluster 2				
Class I	94038741.23	0.93116	122484956.54	0.87
Class II	22836522.81	0.22612	52477150.93	0.37
Class III	28887385.98	0.28604	45299754.74	0.32
Natural Gas Units: Cluster 2				
Class I	32065424.20	0.99965	34201397.50	0.0999
Class II	477229.35	0.014878	570450.42	0.02
Class III	707947.77	0.02207	1335084.32	0.04
Natural Gas Units: Cluster 4				
Class I	23923849.31	0.99996	50397370.66	0.99918
Class II	45148.90	0.00188	286462.39	0.00567
Class III	215359.90	0.0090015	2022127.34	0.0401
Coal Units: Cluster 4				
Class I	313.17	1	231.61	1
Class II	8.9e-16	2.8e-18	8.6e-16	3.7e-18
Class III	4.2e-18	1.3e-20	4.5e-18	1.9e-20

MSE Distributions by Class: Training Data vs Testing Data. A summary of the error distribution of the cost function approximation for the aggregated thermal generators (fuel oil) of cluster 1 is shown in Table III.

TABLE III
RANGE OF ERRORS BY DATA SET TYPE.

	Training Data [%]	Testing Data [%]
Class I	-4 to 5	$\approx$ -4 to $\approx$ 5
Class II	-0.7 to 1	-2 to 1.5
Class III	$\pm$ 0.7	$\pm$ 1

Although not shown in this work, similar behavior of the error distribution was obtained from all other cost function approximations for the aggregated units in the other clusters.The quadratic cost function class that minimizes the error with respect to the true data set is chosen to be the new cost function for the aggregated units.

B. Testing the Model via UCED Simulations

UCED simulations are performed on both the full and reduced-order network models to evaluate the accuracy of the network reduction technique proposed. This optimization problem is expressed

$$\min_{P}\left\{C_t^{UC}(u,a,z) + C_t^{Op}(P) + C_t^{lf}(RU,RD) + C_t^{z}(r_z)\right\} \quad (11)$$

where C_t^{UC} represents the costs of the phase known as *Unit Commitment* (UC). It includes the startup, shutdown, and fixed costs for keeping the units synchronized to the grid with a status of available to produce power; $C_t^{Op}(P)$ represents the energy and reserve costs of the system; C_t^{lf} is the cost of load-following ramp reserves; and, C_t^{z} is the cost of zonal reserve requirements. For simplification purposes, our numerical example has no zonal reserve requirements, i.e., $C_t^{z} = 0$.

Standard UCED Constraints. Only power balance equations and transmission flow limits, voltage limits, and standard optimal power flow inequality constraints are considered. The reader is referred to [16] for a detail explanation of the the UCED optimization problem formulation.

Error Distribution Between Network Models. The daily operation costs for both network models were obtained by solving the UCED optimization of (11) on both the IEEE 39-bus system and its reduced-order model (see Fig. 4) using the MATPOWER Optimal Scheduling Tool [17]. A set of different demand profiles over 50 days was considered. The accuracy of the reduced-order model compared to the full network model is defined as:

$$\varepsilon = \sum_{i=1}^{N} \left[\frac{|C^{\theta^\star}(G^i) - C(G^i)|}{C(G^i)} \right] \times 100 \tag{12}$$

where $C^{\theta^\star}(G^i)$ is the approximated daily system cost using the reduced-order network model, i.e., the $4-$bus system; $C(G^i)$ is the daily system cost obtained using the full IEEE 39-bus test system; and N is the number of days with different demand profiles over a 24 hours horizon.

Fig. 5 shows the error distribution of the daily cost between the two network models. It ranges from $6.15\% \leq \varepsilon \leq 7.07\%$. The sources of error are associated with inaccuracies obtained during the approximation of the cost function for the aggregated units, mainly due to high startup costs of thermal units when transitioning from low to mid-peak demand profiles.

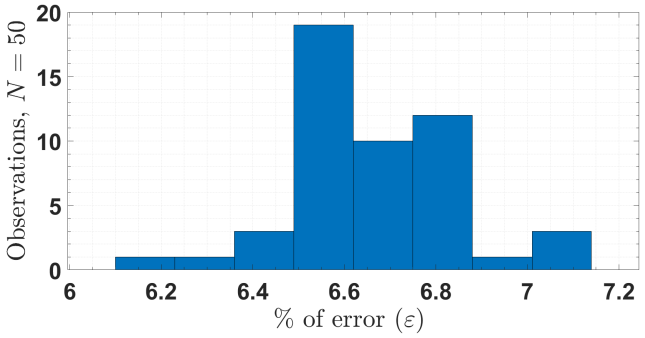

Fig. 5. Daily cost error distribution between the full and reduced-order network models for a data set with 50 days.

V. Conclusions and Future Work

This paper proposes a methodology based on spectral clustering that determines a zone partitioning solution for minimal power flow disruption. The analysis of different spectral clustering properties showed that the utilization of the *normalized* graph Laplacian seems to be more advantageous for clustering purposes. The network reduction was done via aggregation of electrical components. For (same type) aggregated generators, the corresponding cost function is approximated via linear regression. For this, three function architectures classes were tested. The MSE of these function classes suggest that class II and III function structures are more convenient for approximations since the error distributions obtained compared to Class I are smaller. For the "best" candidate cost functions, the error distribution regarding the daily cost obtained (via UCED simulations) with the full and the reduced-order network models is $\approx 5\%$ (i.e., accuracy of $\approx 95\%$). If the utilization of the reduced-order model is constrained to load levels above mid-peak demand, then the error distribution is in the order of 10^{-3}. Although not shown in this work, same high accuracy is also obtained when the cost functions of the aggregated generators are approximated via piece-wise quadratic functions.

Topics for future research. The methodology here proposed will be extended to make the reduced-order model also valid for dynamic studies in low inertia power grids as those presented in [1], [18].

References

[1] 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.

[2] C. J. López-Salgado, O. Añó, and D. M. Ojeda-Esteybar, "Stochastic unit commitment and optimal allocation of reserves: A hybrid decomposition approach," *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 5542–5552, 2018.

[3] F. Milano and K. Srivastava, "Dynamic rei equivalents for short circuit and transient stability analyses," *Electric Power Systems Research*, vol. 79, no. 6, pp. 878–887, 2009.

[4] J. B. Ward, "Equivalent circuits for power-flow studies," *Electrical Engineering*, vol. 68, no. 9, pp. 794–794, 1949.

[5] P. Dimo, "Nodal analysis of power systems," 1975.

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

[7] J. Quirós-Tortós, R. Sánchez-García, J. Brodzki, J. Bialek, and V. Terzija, "Constrained spectral clustering-based methodology for intentional controlled islanding of large-scale power systems," *IET Generation, Transmission & Distribution*, vol. 9, no. 1, pp. 31–42, 2015.

[8] L. Ding, F. M. Gonzalez-Longatt, P. Wall, and V. Terzija, "Two-step spectral clustering controlled islanding algorithm," *IEEE Transactions on Power Systems*, vol. 28, no. 1, pp. 75–84, 2012.

[9] 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.

[10] 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.

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

[12] 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.

[13] J. Machowski, Z. Lubosny, J. W. Bialek, and J. R. Bumby, *Power system dynamics: stability and control*. John Wiley & Sons, 2020.

[14] A. Pai, *Energy function analysis for power system stability*. Springer Science & Business Media, 1989.

[15] D. Krishnamurthy, W. Li, and L. Tesfatsion, "An 8-zone test system based on iso new england data: Development and application," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 234–246, 2015.

[16] A. J. Wood, B. F. Wollenberg, and G. B. Sheblé, *Power generation, operation, and control*. John Wiley & Sons, 2013.

[17] C. E. Murillo-Sánchez, R. D. Zimmerman, C. L. Anderson, and R. J. Thomas, "Secure planning and operations of systems with stochastic sources, energy storage, and active demand," *IEEE Transactions on Smart Grid*, vol. 4, no. 4, pp. 2220–2229, 2013.

[18] 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.