INTENTIONAL ISLANDING OF POWER GRIDS WITH DATA DEPTH

Asim Dey, Yulia R. Gel *

University of Texas at Dallas
Department of Mathematical Sciences
Richardson, TX 75080, USA

H. Vincent Poor

Princeton University
Department of Electrical Engineering
Princeton, NJ 08544, USA

ABSTRACT

A new method for intentional islanding of power grids is proposed, based on a data-driven and inherently geometric concept of data depth. The utility of the new depth-based islanding is illustrated in application to the Italian power grid. It is found that spectral clustering with data depths outperforms spectral clustering with k-means in terms of k-way expansion. Directions on how the k-depths can be extended to multilayer grids in a tensor representation are outlined.

Index Terms— Controlled islanding, power grids, complex networks, data depth, tensor projection depth, clustering, probabilistic geometry

1. INTRODUCTION

Power system vulnerability can be attributed to a wide range of internal and external causes, including but not limited to natural calamities, human errors, component and control failures, and targeted attacks [1, 2]. Partitioning large interconnected power grid networks into smaller subnetworks is a widely adopted and long standing management procedure, primarily employed to isolate parts of the grid that are more prone to failures and, hence, to minimize the risk of cascading blackouts – i.e., the strategy that is often referred to as *intentional*, or *controlled islanding*. In recent years, there has been an increasing interest in adapting modern methods for network community detection and, in particular, spectral clustering with k-means, k-medoids, and kernel k-means for graph-structured data to controlled islanding in power systems (see, e.g., [3, 4, 5]), thus forming an intrinsic linkage of optimal power system fragmentation with a vast research field on unsupervised community detection in complex networks [6].

Analysis of the underlying probabilistic geometry of grid networks provides key information for developing optimal fragmentation strategies and for understanding associated intrinsic properties of subgrid reliability and robustness [4, 7, 8]. To gain insight into geometric and spatial characteristics of the grid, [4] proposes an enhanced version of the spectral clustering approach for controlled islanding of power grids, based on a hierarchical representation of the power grid as a dendrogram with a varying number of islands. In turn, [9] considers agglomerative hierarchical clustering, that is, starting from a high number of partitions and then merging grid segments by minimizing a given loss function.

In this paper we introduce a data-driven and inherently geometric concept of ***data depth*** into intentional islanding of power grids. The key idea of a depth function is to assign a numeric score to each data point to characterize its centrality within a given (multivariate or functional) data cloud or its underlying distribution. By simultaneous accounting for probabilistic geometry of the whole data cloud, depth-based analysis allows us to more systematically assess clusters, outliers and anomalous structures and, as a result, more efficiently recover latent mechanisms behind data formation and system functionality. Data depth is a widely adopted tool in multivariate analysis, high dimensional and functional data studies (see, e.g., [10, 11, 12]), but yet remains a largely unexplored concept in a context of intentional islanding and, in general, for segmentation of complex networks. In this paper we explore the k-depths approach of [13] based on L_1-depth and its extension to other depth functions, as an alternative to conventional k-means clustering for controlled islanding of power grids. We illustrate the utility of depth-based islanding in application to the Italian power grid and discuss the extension of a depth concept to the analysis of multilayer grids in a tensor representation.

2. BACKGROUND ON GRAPHS AND SPECTRAL CLUSTERING

Graph Representation of Power Grids We consider an undirected graph $G = (V, E)$ as a model of a power grid network with node set V and set of edges $E \subset V \times V$, where $(i, j) \in E$

*Yulia R. Gel has been partially supported by NSF DMS 1736368 and NSF IIS 1633331, and H. Vincent Poor has been partially supported by NSF DMS 1736417, NSF CMMI 1435778 and NSF ECCS-1549881. This material is also based in part upon work supported by DARPA. This work was initiated while the second two authors were Visiting Scholars at the Isaac Newton Institute for Mathematical Sciences, Cambridge, UK, under the support of EPSRC grant no EP/K032208/1. The authors would like to thank Martí Rosas-Casals, for providing the data for European power grid networks and Yongwan Chun for help with GIS data extraction.

978-1-5386-1251-4/17/$31.00 ©2017 IEEE

represents an edge (a transmission line or a transformer) from node i to j. We assume that G is undirected, i.e., $(i, j) \in E$ iff $(j, i) \in E$. In this paper, we restrict our analysis to simple graphs, where no loops and no multiple edges are allowed. Here $n = |V|$ is the number of nodes.

Since the topological structure of G does not reflect the functional information about the power grid [1, 4], we can also consider an *edge-weighted* graph, or a pair (G, ω). Here $\omega : V \times V \mapsto \mathbb{R}_{\geq 0}$ is an *edge weight* function such that each edge $e_{uv} \in E$ has a weight ω_{uv}, and $\omega_{uu} = \sum_j \omega_{uv}$. Following [7], we consider *electrical conductance weights* ω_{ij}, defined as a ratio of the number of direct transmission lines between nodes i and j to the geographic distance X_{ij} between the nodes. Hence, elements of the resulting $n \times n$-conductance symmetric adjacency matrix A are defined as $A_{ij} = \omega_{ij}$ if $(i, j) \in E$ and 0 otherwise. The degree f_i of a node i is, hence, defined as $f_i = \sum_j \omega_{ij}$, which in the case of unweighted networks, reduces to a conventional node degree (or number of edges emanating from a node).

Let D be a diagonal matrix of weighted node degrees, i.e., $d_i = D_{ii} = \sum_{j=1}^{n} A_{ij}$. Then, elements of a *normalized graph Laplacian* are defined as follows: L_{ij} is $-\omega_{ij}/\sqrt{d_i d_j}$ if $i \neq j, (i, j) \in E$; 1, if $i = j$, and 0, otherwise.

Spectral Clustering (SC) is one of the most popular tools for controlled islanding of power grids (see recent reviews in [4] and [5]). The key idea of SC is to embed a graph G into a multivariate space $\mathbb{R}^k$. In particular, let k be a given number of islands (clusters), and let $\mathrm{x}_{\cdot j}$, $j = 1, \ldots, k$, be orthogonal eigenvectors of the Laplacian L, corresponding to the k largest eigenvalues. Form an $n \times k$-matrix $X = [\mathrm{x}_{\cdot 1}, \ldots, \mathrm{x}_{\cdot K}]$, where each row of X, $\mathrm{x}_i \equiv \mathrm{x}_{i \cdot}$, provides a representation of a node in $\mathcal{V}$ in $\mathbb{R}^k$. We can now cluster the resulting n sample points in $\mathbb{R}^k$ using any appropriate classifier, such as, for instance, k-means, k-medians, or k-medoids. If the number of islands k is unknown, which is typically the case in most applications, an optimal k can be identified from an eigengap analysis, e.g., a scree plot of leading eigenvalues. Alternatively, hierarchical divisive or agglomerative clustering can be employed.

Let us provide a succinct description of a k-means algorithm that is arguably the most conventional classifier within the SC framework. Given data points $\{\mathrm{x}_i\}_1^n, \mathrm{x}_i \in \mathbb{R}^p$, group the data $\{\mathrm{x}_i\}_1^n$ into k islands $\mathbf{C} = \{C_1, \ldots, C_k\}$ in such a way that the within-cluster sum of squares is minimized, i.e. $\mathrm{argmin}_{\mathbf{C}} \sum_{k=1}^{k} \sum_{\mathrm{x} \in C_k} ||\mathrm{x} - \boldsymbol{\mu}_k||^2$, where $\boldsymbol{\mu}_k$ is the k-th group mean and $||\mathrm{x} - \boldsymbol{\mu}_k||^2$ is the squared Euclidean distance between x and $\boldsymbol{\mu}_k$. (Note that in the case of the SC framework $p = k$.) As a function of means and pair-wise Euclidian distances between x_i and group means, the k-means algorithm is known to be sensitive to outliers and does not account for an intrinsic geometry of the data.

3. INTENTIONAL ISLANDING WITH DATA DEPTH

To get better insight into geometric properties of a power grid network, we propose to employ a concept of *data depth*. A data depth $D(\mathrm{x}, \cdot)$ is a function that measures how closely an observed point $\mathrm{x} \in \mathbb{R}^p$, $p \geq 2$, is located to the "center" of a certain finite set $\mathcal{S} \in \mathbb{R}^p$, or relative to F, a probability distribution in $\mathbb{R}^p$. Depth functions were initially introduced in the setting of nonparametric multivariate analysis, with the goal of defining affine invariant or equivariant versions of quantiles, ranks, and order statistics in multidimensional spaces, where there exists no natural order. Most recently, data depth methods have received a new stimulus due to their versatility for robust and distribution-free analysis, classification, and visualisation of high-dimensional and functional data. Nevertheless, data depth yet remains a largely unexplored tool for analysis of power grids and controlled islanding, in particular. There exists a wide range of depth functions that can be selected based on their desirable properties, geometric interpretation, and computational complexity [10, 11]. In this paper, we primarily focus on the two most commonly used depths, namely, Mahalanobis and L_1 functions:

- *Mahalanobis (MhD) depth* of x with respect to (w.r.t) a set $\mathcal{S}$ is $MhD(\mathrm{x}|\mathcal{S}) = [1 + (\mathrm{x} - \boldsymbol{\mu})'\Sigma^{-1}(\mathrm{x} - \boldsymbol{\mu})]^{-1}$, where $\boldsymbol{\mu}$ and Σ are respectively the (sample) mean vector and covariance matrix of $\mathcal{S}$. The MhD allows one to easily handle the elliptical family of distributions, e.g., a Gaussian case.
- The L_1 *depth* of x w.r.t. a set $\mathcal{S}$ is $LD(\mathrm{x}|\mathcal{S}) = 1 - \max[0, ||\bar{e}(\mathrm{x}|\mathcal{S})|| - f(\mathrm{x}|\mathcal{S})]$. Here $f(\mathrm{x}|\mathcal{S}) = \eta(\mathrm{x})/\sum_i \eta_i$ with $\eta(\mathrm{z}) = \sum_{i=1}^{|\mathcal{S}|} \eta_i I(\mathrm{z} = \mathrm{x}_i)$ and $\bar{e}(\mathrm{x}|\mathcal{S})$ is the average of the unit vectors from a point x to all observations in $\mathcal{S}$ and is defined as $\bar{e}(\mathrm{x}|k) = \sum_{i, \mathrm{x}_i \neq \mathrm{x}} \eta_i e_i(\mathrm{x})/\sum_j \eta_j$, where $e_i(\mathrm{x}) = (\mathrm{x}_i - \mathrm{x})/||\mathrm{x}_i - \mathrm{x}||$. Finally, η_i, $i = 1, \ldots, N$, are viewed as weights or as "multiplicities" of x_i, and $\eta_i = 1$ if the data set has no ties. If $\{\mathrm{x}_i\}_1^n$ has ties, "multiplicities" η_i can be chosen in such a way that it preserves convexity of $C(y)$ [14]. The idea of $1 - LD(\mathrm{x}|k)$ is to quantify a minimal additional weight required to assign x so that x becomes the multivariate L_1-median of the data set $\mathrm{x} \cup \mathcal{S}$ [14]. Hence, L_1 depth provides a robust representation of a topological structure of a data cloud $\mathcal{S}$. Since L_1 is non-zero outside the convex hull of $\mathcal{S}$, it is a feasible depth choice for comparing multiple clusters [15].

In this paper we explore the utility of data depth for intentional islanding of power grids. We consider k-depths for network community detection, based on L_1 depth [13], and extend this idea to other depth functions. The key idea of k-depths is that in order to iteratively find "nearest" clusters, we employ a data depth similarity measure (or a measure of "centrality" with respect to the whole data cloud), instead of pairwise Euclidian distances in the k-means algorithm. The details of the k-depths method are outlined in Alg. 1.

Algorithm 1: Grid islanding with k-depths and arbitrary depth function D.

Input : weighted power grid G, depth function D.
Output: islanding of G.

```
compute a weighted normalized Laplacian L ;
Select a number of islands k, 2 ≤ k ≤ n from the
eigengap analysis of L ;
Form an n × k-matrix X = [x.1, ..., x.K] where
x.j, j = 1, ..., k are k orthogonal eigenvectors of
L, corresponding to the k largest eigenvalues of L;
Normalize rows of X, x_i ≡ x_i. ∈ R^k:
u_i = x_i/||x_i||, 1 ≤ i ≤ n;
Randomly select K points as initial cluster centers;
do
    Given a cluster assignment and depth function
    D, calculate the depth D(x|k) of x, x ∈ R^k
    w.r.t. a k-th cluster;
    Update clusters: C_k = {x_i : D(x_i|k) ≥
    D(x_i|j), ∀j, 1 ≤ j ≤ K};
until the assignment no longer changes;
```

Note similarly to k-means, k-depths does not rely on distributional assumptions on eigenvectors of a graph Laplacian L. This is critical since there yet exists no formal result on asymptotic distribution of eigenvectors of L, even under the simplified conditions of the Erdös-Rényi model, and derivation of such a result remains an open fundamental problem [16, 17]. However, in contrast to k-means, k-depths allows to account for intrinsic geometry of a power grid network.

Extension to Multi-attribute Partitioning and Tensors In many modern applications, particularly involving smart grids, optimal system fragmentation may involve multilayer networks as well as multi edge-attributes, i.e., the same edge can be assigned a set of weights, rather than a single weight ω. Such systems can be naturally represented in a tensor form, that is, an adjacency matrix for a unidimensional power grid network becomes a multilayer adjacency tensor. Let us briefly discuss how the depth-based clustering described above for linear spaces can be then extended to data in multi-linear, or tensor spaces. Most currently available depth functions for data in a tensor form are based on flattening the input tensor data into vectors, which changes the underlying geometry of the data. Recently, [18] proposes a new *tensor projection depth* (TPD) technique that extends the projection depth of [10] directly to tensor spaces, avoiding vectorization. In particular, let $\mathbb{S} \in \mathbb{R}^{\mathbf{m}} \otimes \mathbb{R}^{\mathbf{l}}$ be a set of observed tensors, that is, $\mathbb{S}$ is a set of $m \times l$-matrices. Then the tensor projection depth of $\mathbb{X}$, $\mathbb{X} \in \mathbb{R}^m \otimes \mathbb{R}^l$, w.r.t. $\mathbb{S}$ is

$$\mathrm{TPD}(\mathbb{X}, \mathbb{S}) = \left(1 - \sup_{||\mathbf{v}||=||\mathbf{w}||=1} \frac{\mathbf{v}^{\mathbf{T}}\mathbb{X}\mathbf{w} - \mu(\mathbf{v}^{\mathbf{T}}\mathbb{S}\mathbf{w})}{\sigma(\mathbf{v}^{\mathbf{T}}\mathbb{S}\mathbf{w})}\right)^{-1},$$

where $\mathbf{v} \in \mathbb{R}^m$, $\mathbf{w} \in \mathbb{R}^l$, and $\mu(\cdot)$ and $\sigma(\cdot)$ are location and scale measures in $\mathbb{R}^1$, respectively. The TPD concept can be also extended to higher order tensors [18]. Hence, Alg. 1 can be expanded to islanding of multilayer power grids, via analysis of their (weighted) multilayer adjacency tensor with TPD. Alternatively, we can flatten the (weighted) multilayer adjacency tensor and use Algorithm 1 with MhD or L_1 depth.

4. CASE STUDIES

We illustrate the utility of the k-depths method, in application to intentional islanding of the power grid of Italy. (Data have been obtained from the Union for the Coordination of the Transmission of Electricity (UCTE).) The transmission network of Italy consists of 294 nodes and 252 edges. Since we have no information on node types (i.e., transformers, generators, etc), all nodes are treated equivalently. The edge weight function of the grid is defined as electrical conductance [7].

Following [4], we use *expansion* and k*-way expansion* of islands as a measure of grid zoning performance. That is, intuitively clustering is expected to deliver islands (subnetworks) that are cohesive and well connected internally but sparsely connected externally. For any subset of nodes S, $S \subset V$, the expansion of S is $\phi(S) = d(S)/vol(S)$, where $d(S)$ is a boundary of S (i.e., $d(S) = \sum_{i \in S, j \notin S} \omega_{ij}$), and $vol(S)$ is a volume of S (i.e. $vol(S) = \sum_{i \in S} f_i$). In turn, the k-way expansion constant is $\rho(k) = \min_{S_1 \cdots, S_k} \max_{1 \le i \le k} \phi(S_i)$, and lower values of $\rho(k)$ are preferred. Finally, via a generalization of the Cheeger inequality, $\rho(k)$ can be bounded as $\lambda_k/2 \le \phi_G(k) \le O(k^2)\sqrt{\lambda_k}$, where $0 = \lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$ are eigenvalues of L [19, 20]. The left-hand side of the inequality provides a lower bound on the k-way expansion for any partitioning of G into k zones.

We now apply SC, based on k-means and k-depths with L_1 and Mahalanobis depth functions, to the power grid in Italy. (We also considered k-depths with the projection depth of [10]. The resulting islands were generally similar to that of L_1, and hence are omitted for brevity.) Since all clustering methods exhibit sensitivity to initial centers, we perform 500 replications and report the attained minimum of $\rho(k)$. Fig. 1 depicts examples of partitioning the Italian grid into 5 zones, based on k-means and k-depths with L_1 and MhD functions. Table 1 shows that the SC with k-depths consistently outperforms SC with k-means. In particular, in all cases, geometrically enhanced k-depths islanding provides lower k-way expansion constants $\rho(k)$ than the ones delivered by k-means, with the highest improvement of 27% yielded by k-depths with MhD for k of 3. For k of 4 and 5, the best delivered results of k-depths for $\rho(k)$ are 1.54% and 1.65%, respectively, which are 17% and 15% lower than $\rho(k)$ of k-means, respectively. For the lower number of islands (i.e., k of 3), k-depths with MhD exhibits the most accurate performance, and for k of 4 and 5, k-depths with L_1 depth outperforms all con-

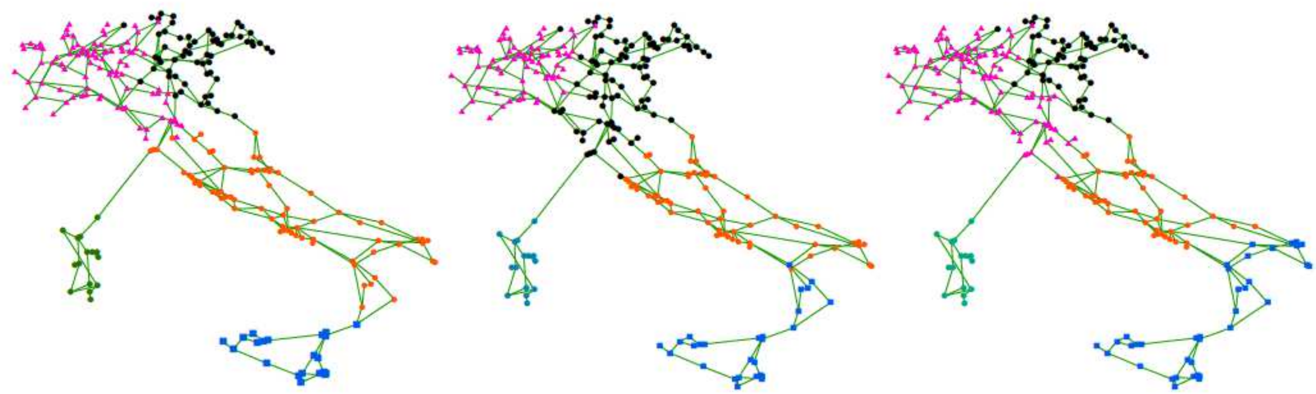

Fig. 1: Examples of islanding the Italian power grid network into 5 zones: with k-means (left panel), k-depths with L_1 (central panel) and k-depths with Mahalanobis (right panel).

Table 1: Summary of the k-way expansion constant $\rho(k)$, $k = 3, 4$ for islanding of the Italian power grid, based on SC with k-means and k-depths with L_1 and MhD depths; $\lambda_k/2$ is a theoretical lower bound from the Cheeger inequality.

# of islands k	Theoretical bound $\frac{\lambda_k}{2}$	k-means	k-depths with L_1	k-depths with MhD
3	0.06%	1.23%	0.96%	0.89%
4	0.18%	1.86%	1.54%	1.64%
5	0.24%	1.95%	1.65%	1.75%

sidered islanding methods. These findings can be potentially explained by the fact that for a lower number of clusters, leading eigenvalues of L tend to more closely follow a Gaussian distribution, and hence k-depths with MhD yields the most competitive results; while for a higher number of islands, the more outlier resistant L_1-depth becomes a preferred choice. In addition, we find that SC with k-depths tends to outperform SC with k-means for sparser power grid networks, and to deliver similar results for denser graphs (results are omitted for brevity). However, even for dense grids, SC with k-means remains sensitive to the choice of initial centers of islands, and performance of k-means can be stabilized if the initial centers are selected via k-depths. Hence, we can conclude that SC with k-depths is a preferred choice for intentional islanding of sparse power grids, and for denser grids, conventional islanding approaches can benefit from initialization via the data depth clustering.

5. CONCLUSION AND DISCUSSION

We have introduced the robust and inherently geometric concept of *data depth* to investigate spatial properties and internal connectivity of power grid networks. We have found that for sparser networks, zone fragmentation using a depth-based (dis)similarity measure, tends to deliver more cohesive islands than conventional methods, based on a Euclidian metric. The current study aims to serve as a starting point for further analysis of probabilistic geometry and its role in power grid functionality. In particular, the depth-based (dis)similarity measure can be naturally combined with the hierarchical approaches of [4, 21]. Furthermore, stability and reliability of a grid fragmentation can be studied w.r.t. inference based on multiple depth functions. Finally, depth tools in multilinear spaces can be expanded to study high-dimensional probabilistic geometry of grids with multi edge-attributes, power grid motif tensors, and multilayer networks of smart grids, following their Laplacians and adjacency tensor embeddings.

6. REFERENCES

[1] L. Cuadra, S. Salcedo-Sanz, J. Del Ser, S. Jiménez-Fernández, and Z. W. Geem, "A critical review of robustness in power grids using complex networks concepts," *Energies*, vol. 8, no. 9, pp. 9211–9265, 2015.

[2] M. Rohden, D. Jung, S. Tamrakar, and S. Kettemann, "Cascading failures in AC electricity grids," *Physical Review E*, vol. 93, pp. 032209, 2007.

[3] A. Peiravi and R. Ildarabadi, "Fast algorithm for intentional islanding of power systems using the multilevel kernel k-means approach," *J. Appl. Sci.*, vol. 9, pp. 2247 – 2255, 2009.

[4] 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 Trans. Power Syst.*, vol. 29, no. 5, pp. 2229–2237, 2014.

[5] A. Esmaeilian and H. Kezunovic, "Prevention of power grid blackouts using intentional islanding scheme," in *IEEE Trans. Ind. Appl.*, 2017, vol. 54, pp. 622–629.

[6] E. Estrada and P.A. Knight, *A First Course in Network Theory*, Oxford University Press, 2015.

[7] Y. Xu, A.J. Gurfinkel, and P.A. Rikvold, "Architecture of the florida power grid as a complex network," *Physica A*, vol. 401, pp. 130 – 140, 2014.

[8] M. Mureddu, G. Caldarelli, A. Damiano, A. Scala, and H. Meyer-Ortmanns, "Islanding the power grid on the transmission level: less connections for more security," *Scientific Reports*, vol. 6, pp. 34797, 2016.

[9] S. Pahwa, M. Youssef, P. Schumm, C. Scoglio, and N. Schulz, "Optimal intentional islanding to enhance the robustness of power grid networks," *Physica A*, vol. 392, no. 17, pp. 3741–3754, 2013.

[10] Y. Zuo and R. Serfling, "General notions of statistical depth function," *Ann. Stat.*, vol. 28, pp. 461–482, 2000.

[11] K. Mosler, "Depth statistics," in *Robustness and Complex Data Structures*, C. Becker, R. Fried, and S. Kuhnt, Eds. 2013, pp. 17–34, Springer.

[12] A. Nieto-Reyes and H. Battey, "A topologically valid definition of depth for functional data," *Stat. Science*, vol. 31, no. 1, pp. 61–79, 2016.

[13] Y. Tian and Y. R. Gel, "Fast community detection in complex networks with a k-depths classifier," in *Big and Complex Data Analysis: Methodologies and Applications*, S. E. Ahmed, Ed. 2017, pp. 139–157, Springer.

[14] Y. Vardi and C.-H. Zhang, "The multivariate l1-median and associated data depth," *PNAS*, vol. 97, no. 4, pp. 1423–1426, 2000.

[15] R. Jörnsten, "Clustering and classification based on the l 1 data depth," *JMVA*, vol. 90, no. 1, pp. 67–89, 2004.

[16] K. Avrachenkov, L. Cottatellucci, and A. Kadavankandy, "Spectral properties of random matrices for stochastic block model," in *PHYSCOMNET 2015*, 2015.

[17] M. Tang and C.E. Priebe, "Limit theorems for eigenvectors of the normalized Laplacian for random graphs," *To appear in Ann. Stat.*, 2017.

[18] Y. Hu, Y. Wang, and Y. Wu, "Tensor-based projection depth," *Bernoulli*, vol. 17, no. 4, pp. 461–482, 2011.

[19] F.R.K. Chung and L. Lu, *Spectral Graph Theory*, vol. 92, American Mathemarical Society, 1997.

[20] J. R. Lee, S.O. Gharan, and L. Trevisan, "Multi-way spectral partitioning and higher-order cheeger inequalities," in *Proc. of the 44th ACM Symp. on Theory of Comp.*, 2012, pp. 1117–1130.

[21] C. Alzate and J.A.K. Suykens, "Hierarchical kernel spectral clustering," *Neural Networks*, vol. 35, pp. 21–30, 2012.