Anomaly Detection in Data-Driven Coherency Identification Using Cumulant Tensor

Bo Sun, *Student Member*, Yijun Xu, *Senior Member*, Qinling Wang, Shuai Lu, *Member*, Ruizhi Yu, Wei Gu, *Senior Member*, Lamine Mili, *Life Fellow*

***Abstract*—As a model reduction tool, coherency identification has been extensively investigated by power researchers using various model-driven and data-driven approaches. Model-driven approaches typically lose their accuracy due to linear assumptions and parameter uncertainties, while data-driven approaches inevitably suffer from bad data issues. To overcome these weaknesses, we propose a data-driven cumulant tensor-based approach that can identify coherent generators and detect anomalies simultaneously. More specifically, it converts the angular velocities of generators into a fourth-order cumulant tensor that can be decomposed to reflect the coherent generators. Also, using co-kurtosis in the cumulant tensor, anomalies can be detected as well. The simulations reveal its excellent performance.**

***Index Terms*—Coherency identification, tensor decomposition, co-kurtosis, anomaly detection, power system.**

I. Introduction

TO SIMPLIFY the dynamic analysis of power systems, coherency identification is widely used to identify the groups of generators that swing together under a contingency. Traditionally, researchers rely on model-driven approaches to achieve this task [1], [2]. Although straightforward, they typically linearize the nonlinear power system model, which inevitably sacrifices some accuracy. Also, uncertainties in the model parameters are inevitable.

Facing these challenges, the data-driven methods that rely on Phasor Measurement Units (PMUs) are gaining increasing attention today. For example, Khalil *et al.* [3] identify coherent generators and electrical areas using frequency deviation signals. Anaparthi *et al.* apply Principal Component Analysis (PCA) to classify generators with common features [4]. Susuki *et al.* [5] initiate a model-free Koopman approach to identify coherent generators using Koopman modes. Typically, these data-driven methods mainly focus on coherency identification but yield biased identification results when bad data appear.

To solve these issues, we propose a novel data-driven coherency identification method using the cumulant tensor for the first time. It only needs to use the generator angular velocity data to achieve cost-effective coherency identification. Furthermore, by decomposing its fourth-order cumulant tensor to analyze its corresponding principal kurtosis vectors, it simultaneously detects anomalies. Simulation results reveal its excellent performance.

B. Sun, Y. Xu, Q. Wang, S. Lu, R. Yu and W. Gu are with the Electrical Engineering Department, Southeast University, Nanjing, Jiangsu 210096, China, (e-mail:{bosun, yijunxu, qlingwang, wgu}@seu.edu.cn).

L. Mili is with the Bradley Department of Electrical and Computer Engineering, Virginia Tech, Northern Virginia Center, Falls Church, VA 22043 USA (e-mail:{lmili}@vt.edu).

This work was supported by the Natural Science Foundation of Jiangsu Province under Grant BK20230851.*(Corresponding author: Yijun Xu.)*

II. Preliminaries

Let us first review a traditional data-driven approach, e.g., the PCA, which serves as the basis of our work, since it has a similar coherency identification criterion to our proposed method [4]. In principle, the PCA method can transform the relevant variables into a set of irrelevant variables arranged in decreasing order of significance. The ones with higher importance are grouped for coherency identification.

More specifically, following Anaparthi *et al.* [4], only the angular velocities of the synchronous generators are utilized for the identification of PCA-based coherency. They are then placed in a matrix $\boldsymbol{X}_{m\times n}$, where m is the number of generators and n is the number of samples. By using Singular Value Decomposition (SVD), we get

$$\boldsymbol{X} = \boldsymbol{U}_x \boldsymbol{\Sigma}_x \boldsymbol{V}_x^T, \quad \boldsymbol{P} = \boldsymbol{U}_x \boldsymbol{\Sigma}_x. \tag{1}$$

Here, each column of $\boldsymbol{P}$ represents the principal component score for each variable. Finally, the matrix is decomposed as

$$\boldsymbol{X} = \begin{pmatrix} p_{1,1} \\ \cdots \\ p_{m,1} \end{pmatrix} \boldsymbol{v}_1^T + \begin{pmatrix} p_{1,2} \\ \cdots \\ p_{m,2} \end{pmatrix} \boldsymbol{v}_2^T + \cdots + \begin{pmatrix} p_{1,m} \\ \cdots \\ p_{m,m} \end{pmatrix} \boldsymbol{v}_m^T. \tag{2}$$

Typically, most of the features of a variable can be captured by the first three principal components. Therefore, we just choose the first three columns of $\boldsymbol{P}$, which can be plotted in a three-dimensional score plot. When the points represented by generators are close together, they are considered coherent.

Remark 1. *Obviously, facing outliers, data-driven approaches,* including the PCA*, produce biased estimation results that could lead to false identification of coherency. Therefore, it is meaningful to develop a data-driven approach that achieves coherent identification and anomaly detection simultaneously.*

III. The Proposed Method

In this section, we elaborate on the cumulant tensor-based approach that not only allows us to identify the coherent generators but also detects anomalies via co-kurtosis.

A. The Construction of Cumulant Tensor

First, let us begin with the traditional Independent Component Analysis (ICA) model. For the aforementioned m-dimension observation matrix, $\boldsymbol{X}$, suppose that we have a q-dimension source sample vector, $\boldsymbol{S} = \left[\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_q\right]^T$, generated by a set of independent non-Gaussian distributed random variables. It follows that

$$\boldsymbol{X} = \boldsymbol{A}\boldsymbol{S} + \boldsymbol{N}, \tag{3}$$

where $\boldsymbol{A} \in \mathbb{R}^{m\times q}$ is a mixing matrix and $\boldsymbol{N} \in \mathbb{R}^{m\times n}$ is the noise matrix. Now, for $\boldsymbol{X}$, let us apply the Taylor expansion to its first characteristic function, $\Phi_{\boldsymbol{X}}(\boldsymbol{t}) = \mathbb{E}(e^{i\boldsymbol{t}^T\boldsymbol{X}})$, and its second characteristic function, $\Psi_{\boldsymbol{X}}(\boldsymbol{t}) = log\Phi_{\boldsymbol{X}}(\boldsymbol{t})$, respectively, [6]. We get

$$\mu'^{\boldsymbol{X}}_{(r)} \stackrel{\text{def}}{=} \mathbb{E}\boldsymbol{X}^r = (-i)^r \frac{\partial^r \Phi_{\boldsymbol{X}}(\boldsymbol{t})}{\partial \boldsymbol{t}}\bigg|_{\boldsymbol{t}=0}, \tag{4}$$

$$\kappa^{\boldsymbol{X}}_{(r)} \stackrel{\text{def}}{=} cum\underbrace{\{\boldsymbol{X},\boldsymbol{X},\ldots,\boldsymbol{X}\}}_{\text{r times}} = (-i)^r \frac{\partial^r \Psi_{\boldsymbol{X}}(\boldsymbol{t})}{\partial \boldsymbol{t}}\bigg|_{\boldsymbol{t}=0}, \tag{5}$$

where $\mathbb{E}$ is the expectation operator; cum is a cumulant operator; r stands for the order of moments and cumulants of $\boldsymbol{X}$; and μ and κ stand for moment and cumulant, respectively. As for $\cdot'$, it indicates that the mean values of the variables are not zero. By expanding the logarithm and merging terms of the same order, we can obtain the relations between moments and cumulants [7]. For example, we have

$$\begin{cases} \mu'_i = \kappa_i, \\ \mu'_{ij} = \kappa_{ij} + \kappa_i\kappa_j, \\ \mu'_{ijk} = \kappa_{ijk} + [3]\kappa_i\kappa_{jk} + \kappa_i\kappa_j\kappa_k, \\ \mu'_{ijkl} = \kappa_{ijkl} + [4]\kappa_i\kappa_{jkl} + [3]\kappa_{ij}\kappa_{kl} \\ \qquad + [6]\kappa_i\kappa_j\kappa_{kl} + \kappa_i\kappa_j\kappa_k\kappa_l. \end{cases} \tag{6}$$

Here, $[\cdot]$ denotes McCullagh's *bracket notation* that acts like a permutation, *i.e.* $[3]\kappa_{ij}\kappa_{kl} = \kappa_{ij}\kappa_{kl} + \kappa_{ik}\kappa_{jl} + \kappa_{il}\kappa_{jk}$ [8]. Now, to get the fourth-order cumulant tensor, namely $\kappa_{(4)}$ or κ_{ijkl}, we can invert the relations between μ and κ in (6). Then, we have

$$\begin{aligned} \kappa_{ijkl} = \mu'_{ijkl} - [4]\mu'_i\mu'_{jkl} - [3]\mu'_{ij}\mu'_{kl} \\ +2[6]\mu'_i\mu'_j\mu'_{kl} - 6\mu'_i\mu'_j\mu'_k\mu'_l. \end{aligned} \tag{7}$$

Now, let us further preprocess the random variables using a normalization procedure [9] to ensure the means are zero, namely $\mu'_i = 0$. Therefore, (7) can be simplified as

$$\kappa_{ijkl} = \mu_{ijkl} - [3]\mu_{ij}\mu_{kl}. \tag{8}$$

For this ICA model, the resulting fourth-order cumulant tensor, $\mathcal{C}_4^{\boldsymbol{X}}$, namely the co-kurtosis, can be denoted as

$$\begin{aligned} [\mathcal{C}_4^{\boldsymbol{X}}]_{ijkl} = \mathbb{E}\left\{\boldsymbol{x}_i\boldsymbol{x}_j\boldsymbol{x}_k\boldsymbol{x}_l\right\} - \mathbb{E}\left\{\boldsymbol{x}_i\boldsymbol{x}_j\right\}\mathbb{E}\left\{\boldsymbol{x}_k\boldsymbol{x}_l\right\} \\ -\mathbb{E}\left\{\boldsymbol{x}_i\boldsymbol{x}_k\right\}\mathbb{E}\left\{\boldsymbol{x}_j\boldsymbol{x}_l\right\} - \mathbb{E}\left\{\boldsymbol{x}_i\boldsymbol{x}_l\right\}\mathbb{E}\left\{\boldsymbol{x}_j\boldsymbol{x}_k\right\}, \end{aligned} \tag{9}$$

where $1 \le i,j,k,l \le q$.

B. Coherency Identification via Cumulant Tensor

Now, we can apply the decomposition to $\mathcal{C}_4^{\boldsymbol{X}}$ to capture its characteristics, which allows us to achieve the identification of coherence. Following Kendall and Stuart [10], $\mathcal{C}_4^{\boldsymbol{X}}$ is expressed as

$$\mathcal{C}^{\boldsymbol{X}}_{ijkl} = \delta(i,j,k,l)\mathcal{C}^{\boldsymbol{X}}_{iiii}, \tag{10}$$

where the values of $\delta(i,j,k,l)$ are zeroes unless all their arguments are equal. Due to this property and (3), we have

$$\begin{aligned} \mathcal{C}^{\boldsymbol{X}}_{ijkl} &= \sum_{1\le a,b,c,d\le q} \delta(a,b,c,d)\mathcal{C}^{\boldsymbol{S}}_{aaaa}\boldsymbol{A}_{ia}\boldsymbol{A}_{jb}\boldsymbol{A}_{kc}\boldsymbol{A}_{ld}, \\ &= \sum_{a=1}^{q}\mathcal{C}^{\boldsymbol{S}}_{aaaa}\boldsymbol{A}_{ia}\boldsymbol{A}_{ja}\boldsymbol{A}_{ka}\boldsymbol{A}_{la}, \end{aligned} \tag{11}$$

where $\boldsymbol{A}_{ia}$ is the element of $\boldsymbol{A}$ in the i-th row a-th column. $\mathcal{C}^{\boldsymbol{S}}_{aaaa}$ is a cumulant of $\boldsymbol{S}$, that is, $\kappa^{\boldsymbol{S}}_{(4)}$. Then, (11) yields

$$\mathcal{C}_4^{\boldsymbol{X}} = \sum_{k=1}^{q}\mathcal{C}^{\boldsymbol{S}}_{kkkk}\boldsymbol{a}_k\otimes\boldsymbol{a}_k\otimes\boldsymbol{a}_k\otimes\boldsymbol{a}_k, \tag{12}$$

where $\otimes$ denotes the tensor product, and $\boldsymbol{a}_k$ is the k-th column of $\boldsymbol{A}$. Now, let us define excess kurtosis, γ, as $\gamma = \frac{\kappa_{(4)}}{\kappa^2_{(2)}}$. When $\boldsymbol{S}$ is normalized, we have $\kappa_{(2)} = \mu_{(2)} = 1$. Then, γ is equal to $\kappa_{(4)}$. Therefore, we have

$$\mathcal{C}_4^{\boldsymbol{X}} = \sum_{k=1}^{q}\gamma_k\boldsymbol{a}_k\otimes\boldsymbol{a}_k\otimes\boldsymbol{a}_k\otimes\boldsymbol{a}_k. \tag{13}$$

Obviously, (13) is a canonical polyadic decomposition that has high computational complexity. Therefore, as suggested by Anandkuma [11], we further unfold $\mathcal{C}_4^{\boldsymbol{X}}$ into a matrix as follows:

$$\mathrm{mat}(\mathcal{C}_4^{\boldsymbol{X}}) = \sum_{k=1}^{q}\gamma_k\boldsymbol{a}_k\otimes\mathrm{vec}(\boldsymbol{a}_k\otimes\boldsymbol{a}_k\otimes\boldsymbol{a}_k), \tag{14}$$

where mat and vec are operations transforming a tensor into a matrix and a vector respectively. Then, we simply perform SVD on the matrix to get

$$\mathrm{mat}(\mathcal{C}_4^{\boldsymbol{X}}) = \boldsymbol{U}_c\boldsymbol{\Sigma}_c\boldsymbol{V}_c^T. \tag{15}$$

Following [4], we truncate the first three columns of matrix $\boldsymbol{P}_c = \boldsymbol{U}_c\boldsymbol{\Sigma}_c$ to form a three-dimensional score plot. For the adjacent data points representing each generator, we consider them coherent.

C. Anomaly Detection via Co-kurtosis

Since outliers can lead to incorrect identification of coherency, it is vital to develop an anomaly detection scheme in the data-driven method. To do so, let us apply the SVD to the unfolded cumulant tensor, $\mathrm{mat}(\mathcal{C}_4^{\boldsymbol{X}})$, we have its principal values $\boldsymbol{U}_c = [\boldsymbol{u}_{c,1}, \boldsymbol{u}_{c,2}, \cdots, \boldsymbol{u}_{c,m}]$, and principal vectors $\boldsymbol{\Sigma}_c = diag(\lambda_{c,1}, \lambda_{c,2}, \cdots, \lambda_{c,m})$. When anomalies occur, they will cause significant changes in the principal values and the directions of the principal vectors. This makes sense since $\mathrm{mat}(\mathcal{C}_4^{\boldsymbol{X}})$ contains the information on kurtosis that serves as a measure of existing outliers or a tendency to produce outliers as articulated by Westfall *et al.* [12].

Using this rule, let us define a metric to quantify this as

$$F_i = \frac{\sum_{k=1}^{m}\lambda_{c,k}(\boldsymbol{e}_i^T\cdot\boldsymbol{u}_{c,k})^2}{\sum_{k=1}^{m}\lambda_{c,k}}, \tag{16}$$

where $\boldsymbol{e}_i^T\cdot\boldsymbol{u}_{c,k}$ denotes the i-th entry in the k-th principle vector, F_i represents the proportion of each principal vector. This means that the sum of all F_i should be unity, that is, $\sum_{i=1}^{m}F_i = 1$. In this way, F_i can reflect the proportion of moments, such as kurtosis, that we are addressing here. When outliers occur, the values of F_i change dramatically, allowing us to directly detect anomalies in coherence identification.

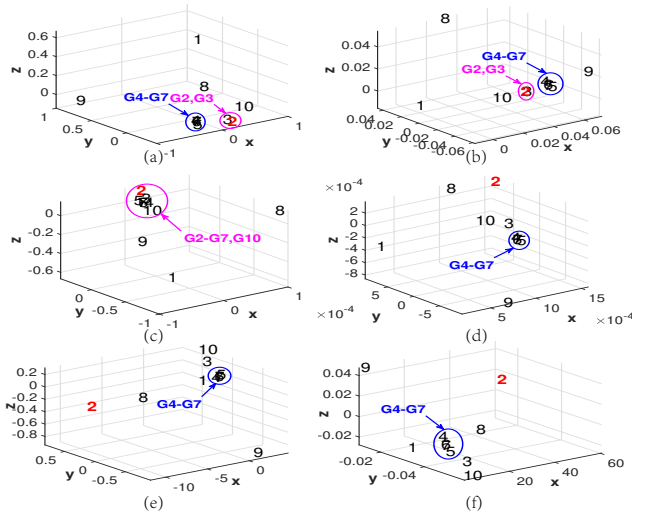

Fig. 1. Coherency identification using PCA and cumulant tensor for generator angular velocities (a)-(b) without anomalies; (c)-(d) with spikes; (e)-(f) with missing data.

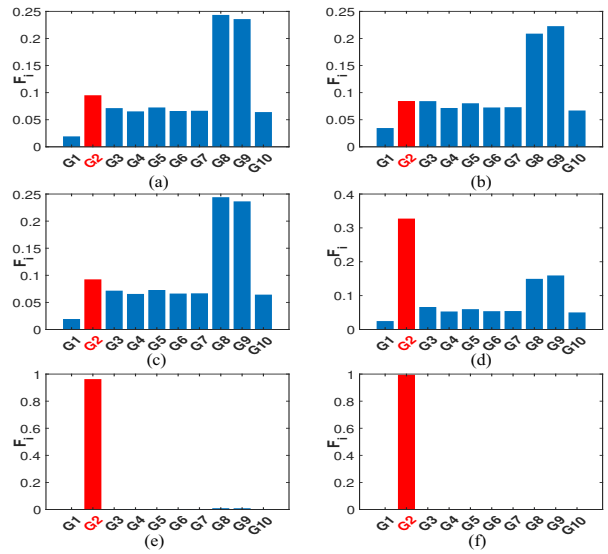

Fig. 2. Plots for anomaly detection where the left is second-moment metrics while the right is fourth-moment one for generator angular velocities (a)-(b) without anomalies; (c)-(d) with spikes; (e)-(f) with missing data.

IV. Case Studies

Let us demonstrate the capability of the proposed method in coherency identification in the New England system. We set anomalies such as spikes and data dropouts for anomaly detection. Lines 11-12 and 25-26 are tripped to stimulate dynamics. The angular velocity of each generator is recorded with a sampling time of $0.02s$. Gaussian noise with a standard deviation of 0.01 is incorporated into all data.

1) Tests without Anomaly: In this test, no anomalies are applied. We tested the proposed method for coherency identification. The results are shown in Fig. 1(a) and (b). The generators in one circle are coherent. Obviously, the coherency identification results obtained using the PCA method and the proposed method are the same, confirming that our approach can accurately achieve coherency identification. Also, from Fig. 2(b), the F_i values of $G8$ and $G9$ are larger than those of other generators, which indicates that their dynamics are different from others and their oscillation amplitudes are relatively large.

2) Tests with Spikes: Now, some observations of $G2$ are set as spikes. The other settings remain unchanged. From Figs. 1(c) and (d), the coherent groups apparently change. PCA generates incorrect results, but our approach can still be identified very well. Then, we use F_i to do anomaly detection. From Figs. 2(c)-(d), the fourth-moment metric of $G2$ becomes large compared with the second-moment metric using PCA. This shows that the outliers in $G2$ are detected successfully. In practice, when spikes occur, we can compare the F_i values of both methods to detect which generator has an anomaly.

3) Tests with Missing Data: This test contains a period of data dropout of about 2 seconds in $G2$. The other settings remain unchanged. From Figs. 1(e) and (f), we see that both results are changed, and those of our method are relatively more obvious. From Figs. 2(e) and (f), we observe that, when missing data occur, both methods can detect a dramatic change in F_i values of $G2$, which indicates that there is indeed an anomaly in the data set $G2$.

4) Further Comparison Tests: We employ more data-driven approaches, including kPCA and ICA. All results are presented in Table I. The results show that when an anomaly occurs, both methods mentioned above are affected.

TABLE I
COHERENCY IDENTIFICATION VIA KPCA AND ICA

	kPCA	ICA
Normal data	{G1},{G2,G3},{G4-G7},{G8},{G9},{G10}	{G1},{G2,G3},{G4-G7},{G8},{G9},{G10}
Data with spikes	{G1},{G2,G3},{G4-G7},{G8},{G9},{G10}	{G1,G6,G7},{G2,G3,G10},{G4},{G5},{G8},{G9}
Data with missing data	{G1},{G2},{G3,G10},{G4-G7},{G8},{G9}	{G1},{G2},{G3},{G4-G7},{G8},{G9},{G10}

Also, to validate our method for real-time applications in large systems, we apply it to the IEEE 54-machine 118-bus system. The computing times are shown in Table II. The results validate its practicality.

TABLE II
COMPUTION TIMES OF DIFFERENT SYSTEMS

	Normal data	Data with spikes	Data with missing data
IEEE 39-bus system	$0.029735s$	$0.026653s$	$0.027036s$
IEEE 118-bus system	$0.278815s$	$0.222124s$	$0.216155s$

In the end, we test different sampling times to investigate whether they have an impact on the results. From Fig. 3, the results exhibit little variation, except for the third set of experiments, because the time periods for missing data selection are different, but do not affect the results of coherency identification and anomaly detection.

The simulations presented above reveal that our proposed method serves as a reliable tool in coherency identification and anomaly detection.

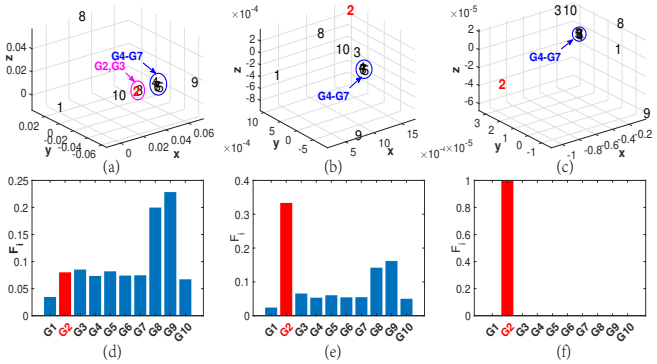

Fig. 3. Coherency identification and anomaly detection when sampling time is $0.1s$ (a) and (d) without anomalies; (b) and (e) with spikes; (c) and (e) with missing data.

V. Conclusion and Future Work

In this paper, we propose a novel data-driven method to simultaneously achieve coherency identification and anomaly detection through a cumulant tensor. The simulation results reveal the excellent performance of the proposed method.

Due to the development of renewable generation sources, the penetration of inverter-based resources is gradually increasing, making the study of coherency identification of inverter networks inevitable [13]. In future work, the proposed method will be applied to the inverter-dominated system.

References

[1] X. Wang, L. Ding, Z. Ma, R. Azizipanah-Abarghooee, and V. Terzija, "Perturbation-based sensitivity analysis of slow coherency with variable power system inertia," *IEEE Transactions on Power Systems*, vol. 36, no. 2, pp. 1121–1129, 2020.

[2] I. Tyuryukanov, M. Popov, M. A. van der Meijden, and V. Terzija, "Slow coherency identification and power system dynamic model reduction by using orthogonal structure of electromechanical eigenvectors," *IEEE Transactions on Power Systems*, vol. 36, no. 2, pp. 1482–1492, 2020.

[3] A. M. Khalil and R. Iravani, "A dynamic coherency identification method based on frequency deviation signals," *IEEE Transactions on Power Systems*, vol. 31, no. 3, pp. 1779–1787, 2015.

[4] K. K. Anaparthi, B. Chaudhuri, N. F. Thornhill, and B. C. Pal, "Coherency identification in power systems through principal component analysis," *IEEE Transactions on Power Systems*, vol. 20, no. 3, pp. 1658–1660, 2005.

[5] Y. Susuki and I. Mezic, "Nonlinear koopman modes and coherency identification of coupled swing dynamics," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 1894–1904, 2011.

[6] T. M. Bisgaard and Z. Sasvári, *Characteristic functions and moment sequences: positive definiteness in probability*. Nova Publishers, 2000.

[7] P. Comon and C. Jutten, *Handbook of Blind Source Separation: Independent component analysis and applications*. Academic press, 2010.

[8] P. McCullagh, "Tensor methods in statistics: Monographs on statistics and applied probability 2018 boca raton chapman and hall," *CRC*, vol. 10, p. 9781351077118.

[9] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *International conference on machine learning*. pmlr, 2015, pp. 448–456.

[10] M. Kendall and A. Stuart, "The advanced theory of statistics. vol. 1: Distribution theory," *London: Griffin*, 1977.

[11] A. Anandkumar, R. Ge, D. Hsu, S. M. Kakade, and M. Telgarsky, "Tensor decompositions for learning latent variable models," *Journal of machine learning research*, vol. 15, pp. 2773–2832, 2014.

[12] P. H. Westfall, "Kurtosis as peakedness, 1905–2014. rip," *The American Statistician*, vol. 68, no. 3, pp. 191–195, 2014.

[13] P. J. Hart, R. H. Lasseter, and T. M. Jahns, "Coherency identification and aggregation in grid-forming droop-controlled inverter networks," *IEEE Transactions on Industry Applications*, vol. 55, no. 3, pp. 2219–2231, 2019.