Spectral Network Approach for Multi-channel Profile Data Analysis with Applications in Advanced Manufacturing

Chen Zhang†, Linmiao Zhang‡, Nan Chen†

†Department of Industrial and Systems Engineering, National University of Singapore, Singapore

‡F10 Data Science, Micron Technology, Singapore

Abstract - In the advanced manufacturing, a lot of sensors are used to collect real-time process signals for statistical monitoring. Motivated by the complex correlation structures of these mutli-channel profile signals, this paper proposes a monitoring scheme for their cross-correlations with the help of spectral network approaches. In particular, we first construct a network model for multi-channel profiles by extracting their features based on the multi-channel functional PCA. The topological structure of the network can represent the cross-correlations of multi-channel profiles. Then we propose to monitor the topological structure using its spectrum information. Numerical studies in a certain fabrication process demonstrate the applicability and efficiency of the proposed methodology.

Keywords - **Multi-channel profiles; Spectral network approach; Statistical process control; Advanced manufacturing**

I. Introduction

Advanced manufacturing has maintained a high growth rate over the past few years. To guarantee product reliability and quality, efficient monitoring and anomaly detection of the manufacturing process is very crucial. An reactor may operate at faulty conditions and produce off-specification products without being detected until off-line metrology tests are performed. Furthermore, off-line metrology tests for every single product are time-consuming. Therefore, on-line efficient monitoring of the advanced manufacturing process is highly desirable to remove these yield limitations. Fortunately, advanced automatic data collection and inspection techniques have been widely adopted in the advanced manufacturing nowadays, which generate large amounts of high-dimensional data streams from multiple sensors characterizing the reaction properties. Those multiple sensing data provide opportunities for on-line process monitoring and anomaly detection. However, to develop such an efficient on-line monitoring system, several challenges have to be addressed.

The most crucial challenge is how to process the high-dimensional, cross-correlated streaming data. In the advanced manufacturing process, each chamber has more than two hundred sensors that measure the reactor temperatures, reactor pressures, reaction gas flows, and reaction electricity environment parameters in millisecond intervals. The data produced by each sensor are non-stationary and strongly time-dependent. They are actually profile (functional) data with auto-correlations over time. Furthermore, different sensor profiles may also be cross-correlated. This means that though, on one hand, monitoring each sensor separately may reduce the dimensionality, this method neglects the cross-correlation structure between sensors and will eventually sacrifice the detection power. For example, while separate monitoring might be still useful for detecting change patterns occurred in each single sensor profile, this method is ineffective in detecting changes of the correlation structure of different sensors. However, on the other hand, the multiple profiles that come from different sources may exhibit quite heterogeneous features. This feature heterogeneity or diversity leads their cross-correlations to be sparse to some degree.

As a powerful tool to describe complex cross-correlations of different random variables, network models are commonly used for multivariate data analysis and anomaly detection. Most of the research follows the logical procedure that first constructs a network for multivariate data based on either some prior information or their inter-relationships, and then extracts aggregated measures of the network topology. Then by treating these measures as univariate or multivariate data, conventional SPC approaches can be applicable [1], [2]. Among different topology information metrics, spectral theoretical approaches have proved to be an efficient way. They were first used predominantly to describe topological relationships in image processing [3], [4]. Later applications of graph theoretic approaches for general signal processing become a nascent domain. Readers may refer to [5], [6] for general reviews.

In particular, [7], [8], [9] proposed to embed high-dimensional data as an undirected graph, and subsequently projected the data into the eigenvector space of the graph Laplacian for process monitoring. [10] applied directed graphs (digraphs) for the analysis of multidimensional vector cardiogram (VCG) signals by converting the VCG signal into a directed graph based on Markov transition probabilities. Then various features, including spectral graph features, of the Markov transition probability matrix were extracted and used for anomaly identification. [11] proposed a weighted directed network approach for characterizing dynamical systems by ordering the data points within a sliding window according to their ordinal and amplitude information.

However, all the methods mentioned above focus on multivariate variables, while their extensions to multivariate functions or profile data is not easy. So

978-1-5386-0948-4/17/$31.00 ©2017 IEEE

far to our best knowledge, few papers focus on this research field. This may be caused by the lack of efficient network models to describe the inter-relationships of multivariate functions. Recently, [12] extended the graphical LASSO [13] to multivariate functional data by constructing a penalized log-Gaussian likelihood on the functional PCA scores. However, the functional PCA is conducted for every profile separately. Consequently the PCA scores lose the information of the cross-correlations of different functions.

Motivated by the wide applications of multi-channel profile data and the infancy of reasonable network monitoring frameworks for them, this paper further explores this field with a twofold contribution. First, we improve the work of [12] by using multi-channel FPCA methods to construct the network. In this way, their cross-correlations can be fully extracted and represented in the topological structure of the network. Second, we propose to monitor the topological structure based on spectral methods. In particular, the eigenvectors of the adjacency matrix is used to construct the monitoring statistic for quantifying the connectivity change of the network. Finally, we apply the proposal into a certain fabrication process.

The remainder of the paper is organized as follows. Section II introduces the network model in detail. Section III presents the monitoring scheme based on the network spectral approach. Section IV uses some numerical studies in a certain fabrication process to demonstrate applicability and efficiency of the proposal. Finally Section V concludes this paper with remarks.

II. Functional Network Model

A. Network Model via Multi-channel Functional PCA

Assume that we have m_0 independent and identical distributed (*i.i.d*) multi-channel profile samples $\{\mathbf{Y}_i(t), i = -m_0+1,\dots,0\}(t \in \mathcal{T})$. Each sample $\mathbf{Y}_i(t)(t \in \mathcal{T})$ contains p-channel profiles which are all square integrable random processes, denoted as $\mathbf{Y}_i(t) = [Y_{i1}(t),\dots,Y_{ip}(t)]\,(t \in \mathcal{T})$. Without loss of generality, we assume that $\mathcal{T} = [a,b], -\infty < a < b < \infty$, and the profiles have zero mean curves, i.e., $\mathrm{E}(\mathbf{Y}_{ij}(t)) = 0$. To analyze their cross-correlations, for each pair of functions $\{Y_{ij}, Y_{il}\}(j \neq l)$ and any $(t,s) \in \mathcal{T}^2$, we first introduce the concept of conditional cross-covariance function as

$$C_{jl}(t,s) = \mathrm{Cov}(Y_{ij}(t), Y_{il}(s)|Y_{ik}(u), k \neq j,l, \forall u \in \mathcal{T}),$$

which represents the covariance between $Y_{ij}(t)$ and $Y_{il}(s)$ conditional on the remaining $p-2$ functions. In another word, $Y_{ij}(t)$ and $Y_{il}(s)$ are conditional uncorrelated if and only if $C_{jl}(t,s) = 0$ for all $(t,s) \in \mathcal{T}^2$.

To analyze the conditional cross-correlations of multi-channel profiles using network modeling, we assume these functions belong to an undirected graph $G = (V,E)$ with vertex set $V = \{1,\dots,p\}$ and edge set E, which is defined as

$$E = \left\{(j,l) : C_{jl}(t,s) \neq 0, \exists s,t \in \mathcal{T}^2, (j,l) \in V^2, j \neq l\right\}.$$

To estimate the edge set E, here we first use multi-channel functional PCA (MFPCA) [14] to represent the multi-channel profiles. In particular, define $\Gamma(t,s) = \mathrm{Cov}(\mathbf{Y}_i(t), \mathbf{Y}_i(s))$ and the covariance operator $(\Gamma f)(t) = \int_a^b f(s)\Gamma(t,s)ds$. Under the assumption that $\Gamma(t,s)$ is continuous over $\mathcal{T}$, this operator Γ has orthonormal eigen-functions, $v_k(t)(t \in \mathcal{T}, k = 1,2,\dots)$, with non-increasing eigenvalues λ_k, satisfying $\Gamma v_k = \lambda_k v_k$. Then according to the Karhunen-Loeve expansion, $\mathbf{Y}_i(t)$ is assumed to be represented as

$$\mathbf{Y}_i(t) = \sum_{k=1}^{\infty} v_k(t)\boldsymbol{\xi}_{ik}', \tag{1}$$

where $\boldsymbol{\xi}_{ik} \in \mathcal{R}_{p\times 1}$ follows a p-dimensional distribution with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Phi}_k$, and has an explicit representation that $\boldsymbol{\xi}_{ik} = \int_a^b \mathbf{Y}_i(t)' v_k(t)dt$.

Generally, the majority of variation in the data is contained in the subspace spanned by the first few eigen-functions (i.e., PCA loadings, features or variation patterns) of (1). Furthermore, in practice, every underlying sample $\mathbf{Y}_i(t)(t \in \mathcal{T})$ is usually recorded at a grid of points. In particular, here we assume that the grid points are the same for all the samples, and dense and equally spaced at $\{t_l, 1 \leq l \leq n\}$. Then we may reformulate (1) and get the following rank-d MFPCA model as

$$\mathbf{Y}_i(t_l) = \sum_{k=1}^{d} v_k(t_l)\boldsymbol{\xi}_{ik}' + \mathbf{e}_i(t_l), l = 1,\dots,n. \tag{2}$$

In (2), $\mathbf{Y}_i \in \mathcal{R}_{n\times p}$, and $\mathbf{Y}_i(t_l)$ are the l^{th} row (grid point) of $\mathbf{Y}_i$. $v_k \in \mathcal{R}_{n\times 1}$, and $v_k(t_l)$ is the l^{th} component of v_k; $\mathbf{e_i}(t_l) = [e_{i1}(t_l),\dots,e_{ip}(t_l)] \in \mathcal{R}_{1\times p}$, and $\mathbf{e}_i(t_l)$ is independent noise with mean $\mathbf{0}$ and constant covariance matrix $\sigma^2\mathbf{I}_{p\times p}$, for $l = 1,\dots,n$. Define the score matrix $\boldsymbol{\Xi}_i = [\boldsymbol{\xi}_{i1}^T,\dots,\boldsymbol{\xi}_{ip}^T]^T \in \mathcal{R}^{d\times p}$, and reformulate $\boldsymbol{\Xi}_i$ by stacking its every column as $\boldsymbol{\xi}_i = \mathrm{vec}(\boldsymbol{\Xi}_i) \in \mathcal{R}^{dp\times 1}$. Then if $Y_{ij}(t)$ is a realization of a Gaussian process for $j = 1,\dots,p$, $\boldsymbol{\xi}_i$ will have a multivariate Gaussian distribution with covariance matrix $\boldsymbol{\Sigma}^d = (\boldsymbol{\Theta}^d)^{-1}$. If d is large enough, then the approximation of $\mathbf{Y}_i$ by the first d principle functions will be close enoug to the true one. Therefore, we may recover the edge set E based on $\boldsymbol{\xi}_i$ and $\boldsymbol{\Theta}^d$.

Lemma 1: For $(j,l) \in V^2$, let $\boldsymbol{\Theta}_{jl}^d$ be the $d\times d$ matrix corresponding to the $(j,l)-th$ submatrix of $\boldsymbol{\Theta}^d$. Then

$$E = \{(j,l) : ||\boldsymbol{\Theta}_{jl}^d||_F \neq 0, (j,l) \in V^2, j \neq l\}$$

where $||\cdot||_F$ is the Frobenius norm.

The proof of Lemma 1 is similar to [12]. Lemma 1 suggests that the problem of recovering E can be reduced to one of accurately estimating the block sparsity structure $\boldsymbol{\Theta}^d$, which will be introduced in the following section.

B. Functional Graphical LASSO

In this section we first introduce the estimation of $\boldsymbol{\xi}_i$ and then introduce the method to estimate the edge set E.

The estimation of $\boldsymbol{\xi}_{ik}(k = 1, \ldots, d)$ is based on its explicit form $\boldsymbol{\xi}_{ik} = \int_a^b \mathbf{Y}_i(t) v_k(t) dt$ where $v_k(k = 1, \ldots, d)$ are the eigenfunctions of the covariance function of $\mathbf{Y}_i(t)$, i.e., $c(t,s) = \mathrm{E}(\mathbf{Y}_i(t)\mathbf{Y}_i(s))$, with the corresponding eigenvalues $\lambda_k(k = 1, \ldots, d)$. In particular, $c(t,s)$ can be estimated using the m_0 reference samples as

$$c(t,s) = \frac{1}{m_0} \sum_{i=-m_0+1}^{0} \sum_{j=1}^{p} Y_{ij}(t) Y_{ij}(s).$$

Then the corresponding estimators of $v_k(t)$ and λ_k are

$$\int_a^b c(t,s) v_k(s) ds = \lambda_k v_k(t), t \in \mathcal{T}.$$

Then based on $\boldsymbol{\xi}_{ik}, i = -m_0 + 1, \ldots, 0, k = 1, \ldots, d$, we can estimate the sample covariance matrix of $\boldsymbol{\xi}$, defined as $\mathbf{S}$. Then motivated by the functional graphic group LASSO of [12], we define

$$\hat{\boldsymbol{\Theta}}^d = \arg\max_{\boldsymbol{\Theta}^d} \left\{ \log\det \boldsymbol{\Theta}^d - \mathrm{tr}(\mathbf{S}\boldsymbol{\Theta}^d) - \sum_{j \neq l} ||\boldsymbol{\Theta}^d||_F \right\}, \tag{3}$$

where $\boldsymbol{\Theta}^d \in \mathcal{R}^{dp \times dp}$ is symmetric positive definite and γ is a non-negative tunning parameter. This group lasso penalty in (3) forces the elements of $\boldsymbol{\Theta}^d_{jl}$ to either all be zero (a sparse solution) of all non-zero (a connected edge between $Y_{ij}(t)$ and $Y_{il}(t)$). As such, the final estimated edge set is defined as

$$\hat{E} = \{(j,l) : ||\hat{\boldsymbol{\Theta}}^d_{jl}||_F \neq 0, (j,l) \in V^2, j \neq l\}. \tag{4}$$

The detailed estimation of (4) can follow the estimation procedure of [12]. Note though $\hat{\boldsymbol{\Theta}}, \mathbf{S}, \hat{E}$ and $\boldsymbol{\xi}_i$ all depend on d, for notation simplicity, we omit the superscripts d when there is no confusion hereafter.

After estimating E, we may further define the weight of the edge between (j,l) as $||\boldsymbol{\Theta}_{jl}||_F$ with the network adjacent matrix A as

$$A(j,l) = \begin{cases} ||\boldsymbol{\Theta}_{jl}||_F, & \text{if } (j,l) \in E \\ 0, & \text{Otherwise} \end{cases}$$

III. A Spectral Approach for Network Anomaly Detection

Now based on the constructed network A, we can extract its relevant topological information using graph theoretic approaches, and detect the cross-correlation change of different profiles by monitoring the network topological structure.

First, the degree d_j of a node j is computed as

$$d_j = \sum_{l=1}^{p} A(j,l).$$

The diagonal degree matrix D structured from d_j is obtained as $D = \mathrm{diag}(d_1, \ldots, d_p)$. Then we can get the normalized Laplacian matrix $\mathcal{L} = I - D^{-0.5} \times A \times D^{-0.5}$, and compute its Eigen spectrum as

$$\mathcal{L}\mathbf{v}_j = \lambda_j \mathbf{v}_j, \quad j = 1, \ldots, p$$

If $\mathcal{L}$ is symmetric positive semidefinite, its eigenvalues $\lambda_j, j = 1, \ldots, p$ are nonnegative and bounded between 0 and 2, i.e., $0 \leq \lambda_j \leq 2$. The smallest nonzero eigenvalue λ_2 is termed as the Fiedler number and the corresponding eigenvector $\mathbf{v}_2$ as the Fiedler vector. The Fiedler vector and values are topological invariant with respect to the transformation $A \to kA$. As such, they can be used as a discriminant for quantifying the connectivity of the network. Therefore, we can use the Fiedler vector to detect the topological change of the network, as the inter-relationships of multi-channel profiles change.

In this study, we propose to use the Fiedler vector to construct the monitoring scheme. In particular, based on the reference samples $\mathbf{Y}_i, i = -m_0 + 1, \ldots, 0$, we can estimate the IC MFPCA loadings and get the scores $\boldsymbol{\xi}_i(i = -m_0, \ldots, 0)$, and the corresponding $\mathbf{S}_0$ and $\hat{\boldsymbol{\Theta}}_0$. Then we get calculate the IC network structure E_0, A_0, and the Fiedler information $\mathbf{v}_0$.

Then for the online testing sample $\mathbf{Y}_i, i = 1, \ldots$, we first project it into the extracted MFPCA loadings by the IC data to get its MFPCA scores $\boldsymbol{\xi}_i, i = 1, \ldots$, then we estimate the online graph by incorporating the exponential window moving average (EWMA) technique. In particular, define

$$\mathbf{S}_i = \rho \boldsymbol{\xi}_i \boldsymbol{\xi}_i^T + (1-\rho)\mathbf{S}_{i-1},$$

$$\boldsymbol{\Theta}_i = \arg\max_{\boldsymbol{\Theta}} \left\{ \log\det \boldsymbol{\Theta} - \mathrm{tr}(\mathbf{S}_i\boldsymbol{\Theta}) - \gamma \sum_{j \neq l} ||\boldsymbol{\Theta}||_F \right\},$$

where ρ is the tuning parameter and $\mathbf{S}_0$ is the estimated IC sample covariance matrix. Based on $\boldsymbol{\Theta}_i$, we further estimate the online graph E_i and A_i, and compute its Fiedler vector $\mathbf{v}_i$. Then we construct the monitoring statistic Z_i as

$$Z_i = (\mathbf{v}_i - \mathbf{v}_0)^T (\mathbf{v}_i - \mathbf{v}_0). \tag{5}$$

If Z_i is bigger than a thresholding l, the monitoring scheme triggers an OC alarm for sample i. The thresholding is calculated based on a pre-specific false alarm rate.

IV. A Real Application in Advanced Manufacturing

In this section, we apply the proposed approach into a real case application in the advanced manufacturing. Due to confidential reasons, we reserve the background of the process, but only focus on the process data. In particular, there are nine process sensors, denoted as S1 to S9, located in the reactor. Each sensor collects observations every 0.01 second during the fabrication, and totally 240 time points are recorded for the whole process. In the dataset, we totally have 125 IC product samples from one batch.

We first use MFPCA to exact features from these nine sensors. To capture more than 90% sample variation, $d = 10$ principle components are considered in our scenario. The recovering of the nine profiles based the extracted 10 principle components for one sample is shown in Figure 1. We can see that the trend of the profiles can be described very well, with some undescribed tight fluctuations due to the system inherent vibration.

Then based on the MFPCA scores, we construct the network model. The network topological structure is represented in Figure 2. In particular, the edge weight is visualized according to its thickness. The thicker the edge, the higher their weight is, indicating a stronger partial cross-correlation. As observed from Figure 2, S2, S3 and S6 have strong partial cross-correlations with each other. This is consistent with their profile patterns in Figure 2. Similarly, S1 and S5 have strong partial cross-correlations with each other, and S7 and S8 have strong partial cross-correlations. In constrast, S4 and S9 have weak cross-correlations with all the other profiles, since their patterns are quite diverse with others as Figure 1 shows. All the above analysis demonstrates the constructed network can efficiently represent the inter-relationships of different profiles. As such, the monitoring statistic based on the topological structure of the network is hoped to detect their inter-relationship changes efficiently, which is demonstrated as follows.

Our experiment mimics the sequential monitoring scheme. In particular, we randomly draw $m_0 = 100$ IC samples with replacement as reference samples from the 125 IC samples, then we draw the subsequent observations sequentially as on-line testing samples. We set ARL_0=200 and $\lambda = 0.1$. We first calculate the control limit l by drawing the on-line samples still from the IC samples. For every test sample we calculate its Z_i, and define if $Z_i > l$, the monitoring scheme triggers an OC alarm at sample i. Then we compute the IC ARL of the monitoring scheme with l based on a large number of simulation replications (say, 5,000 in our manuscript). If the computed IC ARL is smaller than the nominal one, we increase the value of l, otherwise we decrease it. In this way, we can achieve the nominal ARL_0 with the final $l = 0.3$.

Then we evaluate the detection power of the monitoring scheme. We generate the OC pattern purposely from the IC samples as OC samples for testing. This generated OC pattern is inspired by the true anomaly patterns in the fabrication process. In particular, we consider S1 and S2 have a magnitude shift of $\text{std}(Y_{ij}), j = 1, 2$ in the first half time duration, i.e., $Y_{ij}^{OC}(t) = Y_{ij}^{IC}(t) + \text{std}(Y_{ij})$ for $j = 1, 2, t \in [1, 120]$. We assume the change point occurs at the $\tau = 21$ samples. To be more specific, in the simulation we still draw the first $\tau = 20$ on-line testing samples from the IC samples, and then we draw the later samples after τ from the generated OC ones. The monitoring statistic $Z_i, i = 1, \ldots$ are shown in Figure 3. We can see that before τ, Z_i is stable and consistently smaller than the control limit. Later after τ, Z_i has a significant increase and finally goes out of the control limit at $i = 24$, triggering an OC alarm with a four sample delay.

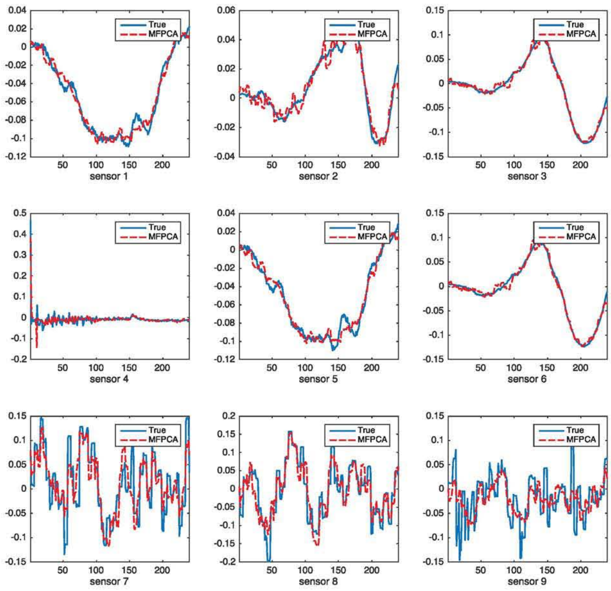

Fig. 1: Profile fitting based on the first 10 principle components of MFPCA

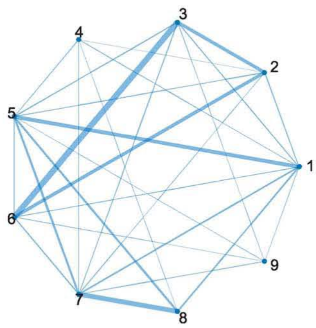

Fig. 2: Graph representation of multi-channels

V. Conclusion

This paper proposes a monitoring scheme for cross-correlations of multi-channel profile data with particular applications in the advanced manufacturing. In particular, we first construct a network model for multi-channel profiles by using the features extracted from the multi-channel FPCA. In this way, their cross-correlations can be fully extracted and represented in the topological structure of the network. Second, we propose to monitor the topological structure based on spectral methods using the Fiedler information. Some numerical studies in the a certain fabrication process

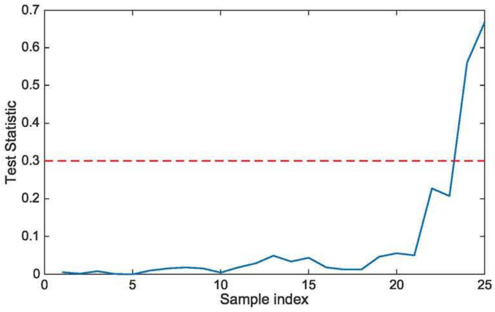

Fig. 3: Test statistic for OC samples starting from $\tau = 20$

demonstrate the applicability and efficiency of the proposal.

Along this research field, there are several potential valuable extensions. Firstly, even if the process is IC, the cross-correlations of multi-channel profiles are still changing along with time. Then how to capture the evolution of the cross-correlations is interesting. Dynamic network models may be useful. Second, there are many sequential stages in the advanced manufacturing, then how to extend our model to multi-stages also deserves more research.

REFERENCES

[1] Y. Park, C. E. Priebe, and A. Youssef, "Anomaly detection in time series of graphs using fusion of graph invariants," *IEEE journal of selected topics in signal processing*, vol. 7, no. 1, pp. 67–75, 2013.
[2] J. Neil, C. Hash, A. Brugh, M. Fisk, and C. B. Storlie, "Scan statistics for the online detection of locally anomalous subgraphs," *Technometrics*, vol. 55, no. 4, pp. 403–414, 2013.
[3] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, no. 8, pp. 888–905, 2000.
[4] U. Von Luxburg, "A tutorial on spectral clustering," *Statistics and computing*, vol. 17, no. 4, pp. 395–416, 2007.
[5] D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 129–150, 2011.
[6] A. Sandryhaila and J. M. Moura, "Discrete signal processing on graphs," *IEEE transactions on signal processing*, vol. 61, no. 7, pp. 1644–1656, 2013.
[7] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural computation*, vol. 15, no. 6, pp. 1373–1396, 2003.
[8] ——, "Semi-supervised learning on riemannian manifolds," *Machine learning*, vol. 56, no. 1-3, pp. 209–239, 2004.
[9] Q. He and S. Zhou, "Discriminant locality preserving projection chart for statistical monitoring of manufacturing processes," *International Journal of Production Research*, vol. 52, no. 18, pp. 5286–5300, 2014.
[10] T. Q. Le, S. T. Bukkapatnam, B. A. Benjamin, B. A. Wilkins, and R. Komanduri, "Topology and random-walk network representation of cardiac dynamics for localization of myocardial infarction," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 8, pp. 2325–2331, 2013.
[11] X. Sun, M. Small, Y. Zhao, and X. Xue, "Characterizing system dynamics with a weighted and directed network constructed from time series data," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 24, no. 2, p. 024402, 2014.
[12] X. Qiao, G. James, and J. Lv, "Functional graphical models," Technical report, University of Southern California, Tech. Rep., 2015.
[13] M. Yuan and Y. Lin, "Model selection and estimation in the gaussian graphical model," *Biometrika*, pp. 19–35, 2007.
[14] K. Paynabar, C. Zou, and P. Qiu, "A change-point approach for phase-i analysis in multivariate profile monitoring and diagnosis," *Technometrics*, vol. 58, no. 2, pp. 191–204, 2016.