IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 13, 2018, accepted September 9, 2018, date of publication September 13, 2018, date of current version October 12, 2018.

*Digital Object Identifier 10.1109/ACCESS.2018.2869923*

# Unsupervised Locality-Preserving Robust Latent Low-Rank Recovery-Based Subspace Clustering for Fault Diagnosis

**JIE GAO[1,2,3], MYEONGSU KANG[4], (Member, IEEE), JING TIAN[5], LIFENG WU[1,2,3], (Member, IEEE), AND MICHAEL PECHT[4], (Fellow, IEEE)**

[1]College of Information Engineering, Capital Normal University, Beijing 100048, China
[2]Beijing Key Laboratory of Electronic System Reliability Technology, Capital Normal University, Beijing 100048, China
[3]Beijing Advanced Innovation Center for Imaging Technology, Capital Normal University, Beijing 100048, China
[4]Center for Advanced Life Cycle Engineering, University of Maryland, College Park, MD 20742, USA
[5]The DEI Group, Millersville, MD 21108, USA

Corresponding author: Lifeng Wu (wulifeng@cnu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 71601022 and Grant 61873175, in part by the Natural Science Foundation of Beijing under Grant 4173074, in part by the Key Project B Class of Beijing Natural Science Fund under Grant KZ201710028028, in part by the Youth Innovative Research Team of Capital Normal University, in part by the Technology Innovation Program [or the Industrial Strategic Technology Development Program (10076392, Development of Vehicle Self Diagnosis System and Service for Automobile Driving Safety Improvement) funded by the Ministry of Trade, Industry and Energy, Korea], and in part by over 100 members of the CALCE Consortium.

**ABSTRACT** With the increasing demand for unsupervised learning for fault diagnosis, the subspace clustering has been considered as a promising technique enabling unsupervised fault diagnosis. Although various subspace clustering methods have been developed to deal with high-dimensional and non-linear data, analyzing the intrinsic structure from the data is still challenging. To address this issue, a new subspace clustering method based on locality-preserving robust latent low-rank recovery (L2PLRR) was developed. Unlike conventional subspace clustering methods, the developed method maps the high-dimensional and non-linear data into a low-dimensional latent space by preserving local similarities of the data with the goal of resolving the difficulty in analyzing the high-dimensional data. Likewise, in the developed L2PLRR method, learned features correspond to low-rank coefficients of the data in the latent space, which will be further used for fault diagnosis (e.g., identification of health states of an object system). The efficacy of the developed L2PLRR method was verified with a bearing fault diagnosis application by comparing with conventional and state-of-the-art subspace clustering methods in terms of diagnostic performance.

**INDEX TERMS** Fault diagnosis, locality-preserving robust latent low-rank recovery, subspace clustering, unsupervised feature learning.

## I. INTRODUCTION

With an ever-increasing demand for reliability in safety-critical devices or systems (e.g., aero engines [1], medical devices [2], and autonomous vehicles [3]), it is important to detect and diagnose impending faults as early as possible to prevent catastrophic failures that can lead to significant economic losses and human casualty. In the recent past, machine learning has evolved to become a key driver of data-driven fault diagnosis methods.

Machine learning methods can be broadly divided into the following two categories depending on the amount and type of supervision they need while training: supervised and unsupervised. In supervised learning, a dataset fed to the machine learning for training includes the desired solutions, called labels, while unsupervised learning methods are trained on an unlabeled dataset. In the field, support vector machines [4], [5], neural networks [6], [7], random forest [8], convolutional neural networks [9], and deep residual networks [10] have been widely used for supervised diagnostic applications.

In real industrial applications, it is difficult to label data because professional knowledge about an object sys-

tem is required. However, there is a relatively massive amount of unlabeled data for fault diagnosis. Thus, it is important to develop unsupervised data-driven fault diagnosis methods. Li *et al*. [11] developed a self-organizing map-based method for the identification of incipient bearing faults. Baraldi *et al*. [12] employed an unsupervised fuzzy c-means algorithm for the identification of operational/faulty transient conditions of a nuclear power plant steam turbine. With the advent of unsupervised deep learning, recent studies have employed restricted Boltzmann machines [13] and autoencoders [14] for unsupervised fault diagnosis with feature learning. Recently, sparse coding and dictionary learning have been used for identifying machine faults in a semi-supervised manner. Jiang *et al*. [15] used a semi-supervised label consistent dictionary learning (SSDL) framework for machine fault classification. Zhang *et al*. [16] introduced a latent label consistent dictionary learning (LLC-DL) model for the sake of classifying salient machine faults.

Subspace clustering (SC), which is used to seek a collection of implicit subspaces to fit a given unlabeled dataset and segment them into different clusters [17], can be an alternative unsupervised learning method for fault diagnosis. To conduct SC, the major two steps are: (1) building an affinity matrix $C$, also called a similarity graph, to represent the affinity of data points, where $C_{ij}$ quantizes the similarity (or closeness) between data points $x_i$ and $x_j$ and (2) clustering data by grouping the eigenvectors of $L = D^{-1/2}AD^{-1/2}$, where $L$ is termed as graph Laplacian, $D$ is with the item $D_{ij} = \sum_j A_{ij}$ and $A = |C| + \left|C^T\right|$. Hence, the clustering performance of this method is largely dependent on the quality of $A$. Likewise, SC has difficulties dealing with the high-dimensional and non-linear data that are distributed in the overlapped subspaces [18]. SC is also sensitive to outliers because they may lead to the wrong segmentation of data points near the intersection of two subspaces [18], [19].

Sparse subspace clustering (SSC) [19] is a variant of SC that clusters data lying in a union of low-dimensional subspaces. The SSC method uses the sparse coefficients of a data point, which ideally corresponds to a combination of data points from its own subspace, as features for clustering [19]. This approach facilitates solving the problems of clustering data near the intersection of subspaces, since sparse optimization automatically picks a few other data points that are not necessarily close to it but belong to the same subspace [20]. Nevertheless, sparse optimization may cause the dispersed distributions of the data in the same subspace, which further leads to unsatisfactory diagnosis performance [20]. Aiming at this problem, Liu *et al*. [20] introduced the SC-based low-rank recovery (LRR) technique, which uses the low-rank coefficients of the data as features. The LRR method captures the global structure of the data, which compacts the distributions of the data into the same subspace. However, the LRR method still has two major drawbacks—it is sensitive to the noisy data and it makes sense only when the data is distributed in independent space (no overlapping or intersect) [21].

Aiming at the first problem (i.e., noisy data), Dyer *et al*. [21] presented a ''greedy'' feature selection method, which seeks the best low-rank representations as features in a ''greedy'' way. Nevertheless, the complexity of iteration limits its use in real applications. From the decomposition perspective, Vidal and Favaro [22] introduced a robust low-rank subspace clustering (LRSC) method. The authors demonstrated that a clean and self-expressive dictionary can be extracted from a corrupted data matrix, and the low-rank coefficients can be learned from this dictionary. In terms of clustering performance, the LRSC method outperformed classical SC methods using LRR.

The second problem (i.e., subspace dependency) is also a challenging issue. High-dimensional data, especially non-linear data in different categories, are often distributed in the non-independent (overlapping) subspaces, and thus the low-rank coefficients calculated from the original data fail to reveal the intrinsic structure of the data, and the discrimination of features is weakened. Searching for a framework that separates the data and reduces the dimensions simultaneously is necessary for improving the clustering performance [23].

To address this issue, Liu and Yan [23] introduced the notion of latent SC. They explored the hidden data located in latent space, which are as important as the observed data for recovering the low-rank coefficients. Through latent space mapping, the overlapping subspaces can be unfolded so that hidden data can be observed. Patel *et al*. [24] investigated a complete framework of latent SC based on SSC (LS3C) to facilitate the dimension reduction and sparse representation, simultaneously. They also offered two mapping functions for the sparse representation: linear mapping and kernel mapping. Wei *et al*. [25] introduced latent SC based on the LRR method to solve the challenge of LRR, which also employs the two mapping functions. However, choosing an applicable mapping function has been a long-standing challenge for latent space sparse or low-rank subspace clustering because both linear mapping and kernel mapping are not appropriate for high-dimensional and non-linear data. Likewise, linear mapping can distort the non-linear structure of the data to some degree, whereas kernel transformation increases the computation complexity, and the performance of clustering is largely determined by the form of the kernel function.

This paper developed a new subspace clustering method based on locality-preserving robust latent low-rank recovery (L2PLRR) to address the problems in the above-mentioned SC methods. The main contributions of this paper are as follows. The developed L2PLRR model extends the LRSC into latent space to recover the low-rank coefficients in that space. Likewise, a new mapping function called ''locality-preserving mapping'' was developed to project a given dataset into a low-dimensional latent space by maintaining the similarity of pairwise data points in a local area (in the same subspace), while also retaining the distinction of data lying in different subspaces (clusters). This new mapping function is effective for revealing the intrinsic structure of non-linear data, which contributes to enhancing feature

discrimination. Two experiments were conducted to verify the effectiveness of the developed method. The first experiment tested the robustness and accuracy of the developed L2PLRR method in terms of clustering performance on a synthetic dataset; the second experiment verified the efficacy of the developed L2PLRR for fault diagnosis (on bearings).

The reminder of the paper is organized as follows. Section II provides the theoretical background of L2PLRR. In Section III, the L2PLRR subspace clustering method and its optimization are presented in detail. Section IV verifies the efficacy of the L2PLRR method for unsupervised clustering by comparing with classical and state-of-the-art SC methods. Section V presents conclusions.

## II. THEORETICAL BACKGROUND OF LOCALITY-PRESERVING ROBUST LATENT LOW-RANK RECOVERY (L2PLRR)

The goal of SC based on LRR is to segment (or cluster) data into clusters with each cluster corresponding to a subspace through extracting the low-rank coefficients of data to build the adjacency matrix, which is used as the input of the spectral clustering [26]. The low-rank coefficients are recovered according to LRR as follows: $\|Z\|_* \, s.t. X = XZ$, where $X \in R^{N\times D}$ is the data matrix whose size is $N$ (observations) $\times D$ (dimensions), and LRR produced low-rank matrix $Z$ by minimizing the nuclear norm: $\|Z\|_* = \sum_i \sigma_i(Z)$, where $\sigma_i(Z)$ is defined as the singular values of the matrix $Z$. The formula means that the data can be represented by linear weight combination of the data itself, and the low-rank coefficients are the weights that measure the degree to which the two data belong to one subspace. Ghergherehchi *et al.* [26] and Zhang *et al.* [18] proved that if the data is seriously corrupted, the low-rank coefficients fail to figure out the subspace information of data because LRR uses the corrupted data as the dictionary to learn the low-rank coefficients, which leads to low clustering performance. Elhamifar and Vidal [19] pointed out that a clean, self-expressive, and low-rank dictionary can be extracted from a corrupted data matrix,

$$\begin{aligned} &\min_{Z,A,G,E} \ \|Z\|_* + \|A\|_* + \frac{\gamma}{2}\|G\|_F^2 + \lambda_1 \|E\|_1 \\ &s.t.\ Z = Z^T, \quad A = AZ, \ and \ X = A + G + E \end{aligned} \tag{1}$$

where $A$ is a new dictionary matrix detached from the data matrix $X \in R^{N\times D}$, which can be self-expressive based on the low-rank matrix $Z$; at the same time, the dictionary $A$ also has a low-rank structure. $G$ and $E$ represent noise and gross error, respectively, both of which should be depressed through keeping Frobenius norm $\|G\|_F^2 = \sum_{ij} G_{ij}^2$ and $L_1$ norm $\|E\|_1 = \sum_{ij} |E_{ij}|$ minimum. Likewise, $\gamma$ and $\lambda_1$ are the parameters to balance the noise and error minimum. Note that the constraint $A = AZ$ is a non-convex problem because both $A$ and $Z$ are unknown, as a result, a convex relaxation framework is put on (1) and the transformation is as follows:

$$\begin{aligned} \min_{Z,A,G,E} \ &\|Z\|_* + \|A\|_* + \frac{\gamma}{2}\|X - A - E\|_F^2 \\ &+ \frac{v}{2}\|A - AZ\|_F^2 + \lambda_1 \|E\|_1 \\ s.t.\ &Z = Z^T \end{aligned} \tag{2}$$

In (2), $v$ is a parameter used to balance the bias of reconstruction. An iterative update strategy was introduced to solve (2), and its steps can be summarized as follows [24]:

$$\begin{aligned} A_{K+1} &= U_K P_{\gamma,v}(\Sigma_K) V_K^T \\ E_{K+1} &= S_{\lambda_1/\gamma}(X - A_{K+1}) \end{aligned} \tag{3}$$

where $U_K \Sigma_K V_K^T$ is the singular value decomposition of $X - E$ at iteration $K$, $\Sigma_K = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)$ are the singular values of $X - E$, and $n$ is the length of singular values. At first, a polynomial thresholding operator $P_{\gamma,v}$ is used to extract the effective singular values of $X - E$ to renovate the dictionary $A$, and a lot of noise can be reduced during the polynomial thresholding operation, which is defined in [25]; in a similar way, $E$ is then updated by thresholding the singular values of $X - A$ through a regular shrinkage-thresholding function $S_{\lambda_1/\gamma}$, defined in [22]. The iteration is started with $E_0 = 0$ and $A_0 = X$, where $K$ is the number of iterations. Finally, the low-rank matrix $Z$ is gained by shrinking the singular values of $A$:

$$Z = V_1 \left( I - \frac{1}{v}\Lambda_1^{-2} \right) V_1^T \tag{4}$$

where $\Lambda = P_{\gamma,v}(\Sigma_K)$.

The matrices $U = [U_1 U_2]$, $\Lambda = diag[\Lambda_1 \Lambda_2]$, and $V = [V_1 V_2]$ are partitioned according to the sets $I_1 = \left\{ i : \lambda_i > 1/\sqrt{v} \right\}$ and $I_2 = \left\{ i : \lambda_i \le 1/\sqrt{v} \right\}$, $Z$ is the optimization low-rank matrix in the end.

## III. DEVELOPED L2PLRR FOR SUBSPACE CLUSTERING

In Section II, the framework of robust LRR for SC was presented. Although the robust LRR is effective for addressing the noisy data for SC, it still fails to overcome the harm of low clustering accuracy if the data is distributed in disjoint and high-dimensional spaces. To address this issue, latent LRR was introduced. As stated in Section I, Patel *et al.* [24] developed a complete framework of latent LRR for SC, which could recover the sparse coefficients and enable dimensionality reduction simultaneously.

However, because there are no SC models that can account for noisy and non-linear data, the L2PLRR method was developed for SC, and its model can be initially expressed as follows:

$$\min_{Z,P,A} \|Z\|_* + \left\| P^T (A - AZ) \right\|_F^2 + \left\| A - PP^T A \right\|_F^2 \tag{5}$$

where $P \in R^{D\times d}$ is a matrix that maps the data matrix $X$ from the original space of dimension $D$ into a latent space of dimension $d$. In (5), the dictionary matrix $Z$ in the latent space still has a low-rank structure, and the last two terms aim at avoiding loss of information of the original data as much as

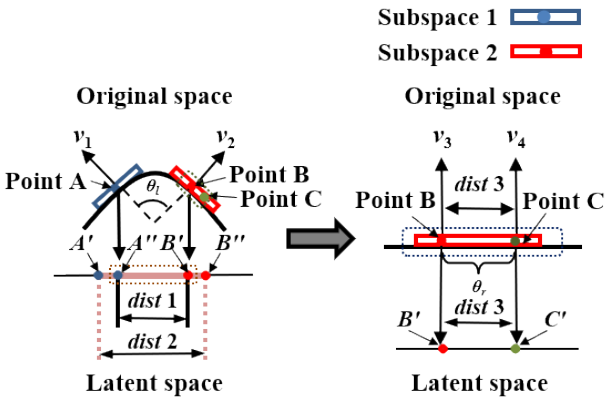



**FIGURE 1.** Effect of mapping data points from different subspaces into the latent space.

possible after latent mapping. The use of the principal component analysis-like mapping function is not appropriate for the non-linear data. Likewise, kernel mapping is not an optimal choice because it suffers from the curse of dimensionality.

As pictorially illustrated in Fig. 1, two data points $A$ and $B$ are in subspaces 1 and 2, respectively, and their linear projections in the latent space are $A^{'}$ and $B^{'}$.

The Euclidean distance between $A'$ and $B'$ is denoted as $dist1$ in Fig. 1. However, because the surface is completely unfolded, the geodesic distance $dist$ 2 (the real discrepancy between actual projections $A''$ and $B''$) is larger than $dist1$. That is, this explains that the difference between data points $A$ and $B$ located in a non-linear subspace is decreased after linear mapping. In contrast, data points $B$ and $C$ in subspace 2 (i.e., the small local area) are assumed to be distributed in a linear space to a great approximation.

The angle between two normal vectors (or data points) $v_i$ and $v_j$ in subspaces $S_i$ and $S_j$ can be used as a metric to determine whether the vectors are located on a curved surface (or different subspaces), defined as:

$$\theta_{ij} = cos^{-1}\left(\frac{\langle v_i, v_j\rangle}{\|v_i\|^2 \|v_j\|^2}\right) \in (0, \pi) \tag{6}$$

where $\langle v_i, v_j\rangle$ is the inner product and $\|\|^2$ is the second norm.

It is worth noting that if $\theta_{ij} \in (0, \pi/2)$, the positive correlation of the subspaces reduces with an increase of $\theta_{ij}$, whereas the negative correlation of the subspaces reduces with an increase of $\theta_{ij}$ if and only if $\theta_{ij} \in (\pi/2, \pi)$. Likewise, if the angle $\theta_{ij}$ is $\pi/2$, the subspaces are assumed to be orthogonal to each other, which implies that the vectors $v_i$ and $v_j$ are placed in different subspaces.

In Fig. 1, $v_1$ and $v_2$ are two normal vectors in subspaces 1 and 2, respectively, where the data points $A$ and $B$ are located, $\theta_l$ indicates the angle between $v_1$ and $v_2$, whereas $\theta_r$ is the angle between normal vectors $v_3$ and $v_4$ in subspace 2, respectively, where the data points $B$ and $C$ are located. Because $v_1$ and $v_2$ are on a curved surface and in different subspaces, $\theta_l$ is in the range of $(0, \pi/2)$. Likewise, because $v_3$ and $v_4$ are parallel and in the same subspace, $\theta_r$ closes to 0.

The subspace structure of the given data points should be preserved in the latent space, which indicates that the data points in the same subspace are tightly agglomerated in the latent space or the data points in different subspaces are alienated from each other in the latent space. Therefore, an additional restraint factor is added to the model described in (5); that is, the developed locality-preserving latent subspace low-rank recovery model is defined as:

$$\begin{aligned} \min_{Z,P,A} \|Z\|_* &+ \left\|P^T(A - AZ)\right\|_F^2 \\ &+ \left\|A - PP^TA\right\|_F^2 + P^TALA^TP \\ s.t.\ P^TASA^TP &= I \end{aligned} \tag{7}$$

After latent space mapping, the data points in the local area remain similar by minimizing the term $P^TALA^TP$ in (7), where $L$ is the Laplacian matrix calculated by the local weight matrix $W$ : $L = S - W$ and $S = \Sigma_j W_{ij}$. This constraint term promotes the symmetry characteristic of the matrix $S$. k-nearest neighbor classification can be used to build the matrix $W$. In fact, the use of a data point's $k$ neighbors is effective for composing a local area. If the $k$ is small enough, it can be said that data points in this area come from the same subspace. Herein, searching for $k$ neighbors of every data point is based on a measure of similarity between angles, defined as:

$$\begin{aligned} W_{ij} &= e^{\frac{\pi/2 - \theta_{ij}|}{\sigma}}, \quad \theta_{ij} \in (0, \pi) \\ \sigma &= \sqrt{\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}(\theta_{ij} - \bar{\theta})}, \\ \bar{\sigma} &= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\theta_{ij} \end{aligned} \tag{8}$$

where $\theta_{ij}$ is the angle of normal vectors $v_i$ and $v_j$ of subspaces $S_i$ and $S_j$. The elements of the matrix $W$ are computed by an exponential function taking the angle of the two normal vectors as an input parameter. Likewise, $\sigma$ is the standard deviation of the angles $\theta_{ij}$. Once the value of $\theta_{ij}$ is $\pi/2$, $W_{ij}$ will be 1, which is the minimum value of the matrix $W$. Then, the intact framework of the L2PLRR is designed by integrating (7) into (2):

$$\begin{aligned} \min_{Z,A,E,P,L} \|Z\|_* &+ \|A\|_* + \frac{\gamma}{2}\|X - A - E\|_F^2 \\ &+ \frac{v}{2}\|A - AZ\|_F^2 + \lambda_1\|E\|_1 + \left\|P^T(A - AZ)\right\|_F^2 \\ &+ \left\|A - PP^TA\right\|_F^2 + P^TALA^TP \\ s.t.\ P^TASA^TP &= I \end{aligned} \tag{9}$$

Algorithm I summarizes the detailed steps of the L2PLRR method.

By carrying out the steps listed in Algorithm I, the optimal low-rank matrix $Z$ is obtained, which will be further used to

**Algorithm 1** Detailed Steps of the Developed L2PLRR Method for SC

**Input:**

A data matrix $X = \{x_1, x_2, \ldots, x_n\}$, where the dimensionality of each data point is $D$.

**Step 1: Initialization**

1. Initialize parameters $\lambda_1$, $\gamma$, and $v$, the number of $k$ neighbors, the dimension of the latent space $d$, the number of iterations $N_{loop1}$ and $N_{loop2}$, and the termination threshold $\varepsilon$.
   $K = 0; E_0 = 0; A_0 = X$
   **FOR** $K = 1: N_{loop1}$
   Calculate the matrices $A_K$ and $E_K$ using (3)
   **END FOR**
2. Obtain the matrix $Z$ using (4)

**Step 2: Update the matrixPby usingZandA**

3. Transform (7) into the generalized Lagrange function, defined as:

$$\Gamma = \left\| P^T (A - AZ) \right\|_F^2 + \left\| A - PP^T A \right\|_F^2 + P^T ALA^T P - \rho \left( P^{\mathrm{T}} ASA^{\mathrm{T}} P - I \right)$$

4. Take the derivative of $\Gamma$ with respect to $P$ and set it to be zero, expressed as:

$$\partial \Gamma / \partial P = \left( (A - AZ)(A - AZ)^{\mathrm{T}} - AA^{\mathrm{T}} + ALA^{\mathrm{T}} \right) = \rho ASA^{\mathrm{T}} P$$

5. Perform $P = V(1:d)$, where $V$ is the generalized eigenvector
6. Perform $A^{'} = PA$

**Step 3: Update the low-rank matrix Z by using A′**

7. Search for the low-rank matrix $Z$ with $A^{'}$

$$\min_{A,Z,E} \|Z\|_* + \lambda_1 \|E\|_1$$
$$s.t.\ A^{'} = A^{'} Z + E$$

8. Solve the equation by employing an augmented Lagrange multiplier to gain the low-rank matrix $Z$.

**Output:**

$Z$ and $P$

build the adjacency matrix (or affinity matrix):

$$M = |Z| + |Z|^T \tag{10}$$

Then, data segmentation (i.e., clustering) is realized by exploiting the adjacency matrix $M$ as an input of SSC.

## IV. EXPERIMENTAL RESULTS

To verify the effectiveness of L2PLRR in clustering, the classical and state-of-the-art SC methods, including SSC [19], LRR [20], LS3C [21], LRSC [22], NLS3C [24], LSRS2 [25], and LATLRR [23], were compared for unsupervised recognition on the synthetic dataset [20] and bearing failure dataset [27]. Additionally, this paper employed the accuracy of clustering as a metric to evaluate the performance of the SC methods because it intuitively shows the performance of clustering, which is defined as follows [24]:

$$\text{Accuracy}(\Omega, C) = \frac{1}{N} \sum_k \max_j \left| \omega_k \cap c_j \right| \tag{11}$$

where $\Omega = \{\omega_1, \omega_2, \ldots, \omega_k\}$ is the true set of the centroids of each cluster, $\omega_k$ is interpreted as the set of documents in the $k^{\text{th}}$ cluster, $C = \{c_1, c_2, \ldots, c_j\}$ is the set of the classes, and $c_j$ is interpreted as the set of documents in the $j^{\text{th}}$ class. Each cluster is assigned to the class that is most frequent in the cluster, and then accuracy is measured by counting the number of correctly assigned documents and dividing by the number of samples $N$, the value of accuracy ranges from 0 to 1.

### A. CLUSTERING RESULTS ON A SYNTHETIC DATASET

To verify the effectiveness of the developed L2PLRR method for SC, a synthetic dataset was generated based on [20]. Four pairwise disjoint subspaces $\{S_i \in \mathbb{R}^{D \times d}\}_{i=1}^4$ of dimension $d = 4$ embedded in $D = 100$ dimensional space were created. More specifically, $C_i$ ($i = 1, 2, 3, 4$) are four bases of these spaces, which can be computed by $C_{i+1} = RC_i$, where the dimensionality of $C_i$ is $100 \times 4$ and $R$ is a rotation matrix that can be independently changed. Then, a $100 \times 120$ dataset $[X_1, X_2, X_3, X_4]$ was constructed by sampling 30 points from each subspace by $X_i = C_i W_i$ with $W_i$ being a $4 \times 30$ weight matrix. In addition, some points could be randomly chosen to be corrupted by adding Gaussian noise.

Fig. 2 illustrates the adjacency (or similarity) matrix recovered by each of the SC methods. Note that the clustering results rely on how well the adjacency matrices are recovered. Likewise, if any of the adjacency matrices show an obvious diagonal block structure on the synthetic data, which means that similarities between data points in the same subspace are larger than those in different subspaces, it can be said that the clustering performance of the method associated with that adjacency matrix is satisfactory. In Fig. 2, it is obvious that the L2PLRR method yielded better clustering performance than the other classical and state-of-the-art SC methods.

Likewise, by comparing with SSC and LRR in Figs. 2(a) and 2(b), it could be possible to identify that SCC resulted in a sparser adjacency matrix, even if the matrix showed a block structure at the frontiers. This is because SSC shows that the sparsest coefficients are also "block-sparse". Namely, the within-cluster similarities are sparse (but nonzero) and the between-cluster similarities are all zeros, which implies that there is no global constraints to its coefficients. Once the data is grossly corrupted, the clustering performance of is largely depressed. It can be observed in Fig 2(b) that low-rank coefficients have a block-impact structure, but the non-diagonal coefficients are nonzero,

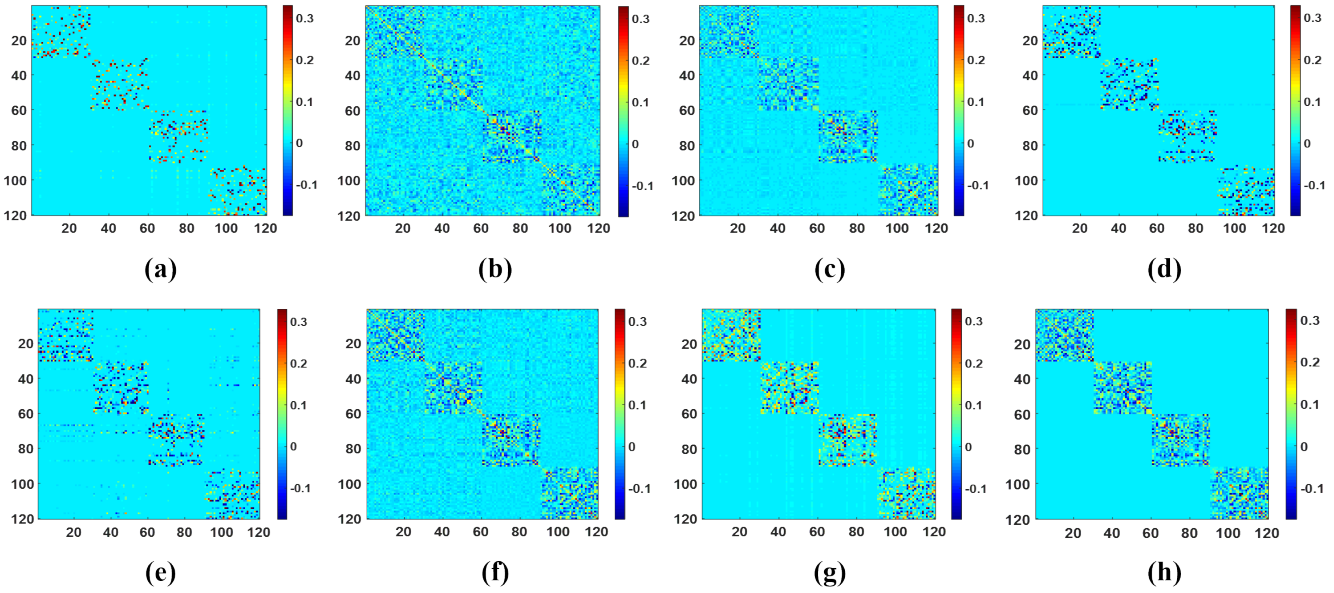



**FIGURE 2.** The adjacency (or similarity) matrixes obtained by eight different algorithms: (a) SSC, (b) LRR, (c) LRSC, (d) LS3C, (e) NLS3C, (f) LSRS2, (g) LATLRR, and (h) L2PLRR.

which means between-cluster similarities are still large and may present an obstacle to clustering. The result of Fig. 2(d) is similar to 2(a)—that is, the elements of the adjacency matrix in a block were not as dense as those obtained by the LRSC (see Fig. 2(c)), LSRS2 (see Fig. 2(f)), and L2PLRR (see Fig. 2(h)), respectively. However, unlike the L2PLRR, both LRSC2 and LRSC yielded scattered impurities in their adjacency matrices.

In summary, the use of low-rank and sparse matrices has pros and cons. More specifically, the inclusion of a low-rank matrix in SC is effective for clustering data into associated clusters but can introduce scattered impurities. However, the inclusion of a sparse matrix in SC suffers from decentralized data in one subspace but produces less-scattered impurities. The L2PLRR method yielded good clustering performance by preserving within-cluster similarities while decreasing the between-cluster similarities as much as possible.

Each of the SC methods was tested on corrupted data to show the robustness of these methods to noise. In the test of clustering pure data, experiments were conducted multiple times and the average accuracy was calculated; in the test of clustering corrupted data, various amounts of Gaussian noise (–8 dB to 13 dB) were added to the data. Maximum, minimum, median, and mean values of clustering accuracy were counted to show that clustering performance varies under different noise strength based on each SC method. The L2PLRR method shows the most satisfactory performance in terms of clustering accuracy, regardless of whether the data is corrupted or not. The accuracy of purely synthetic data is 100%; clustering the corrupted noisy data in different strengths, the median and average accuracy of the L2PLRR is highest, which means that the L2PLRR was effective eliminating the influence of noise when clustering.

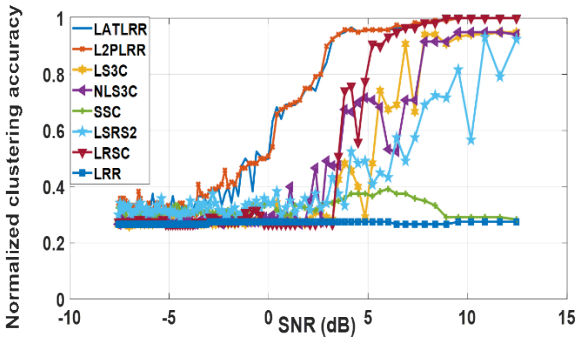



**FIGURE 3.** Clustering accuracy in noisy environments.

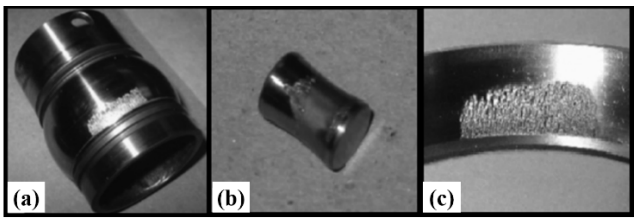



**FIGURE 4.** Snapshots of bearing failures: (a) BFIR, (b) BFOR, and (c) BFRE.

The signal-to-noise ratio (SNR) [20] was used to quantify the amount of noise:

$$SNR = 10log\left(\sigma^2_{signal}\Big/\sigma^2_{noise}\right) \tag{12}$$

The higher the SNR is, the less noise is contained in the data.

Fig. 3 presents the clustering accuracy trends in noisy environments. The clustering accuracy decreases with increase of noise intensities in the data. Among the SC methods, L2PLRR and LATLRR achieved higher accuracies at lower

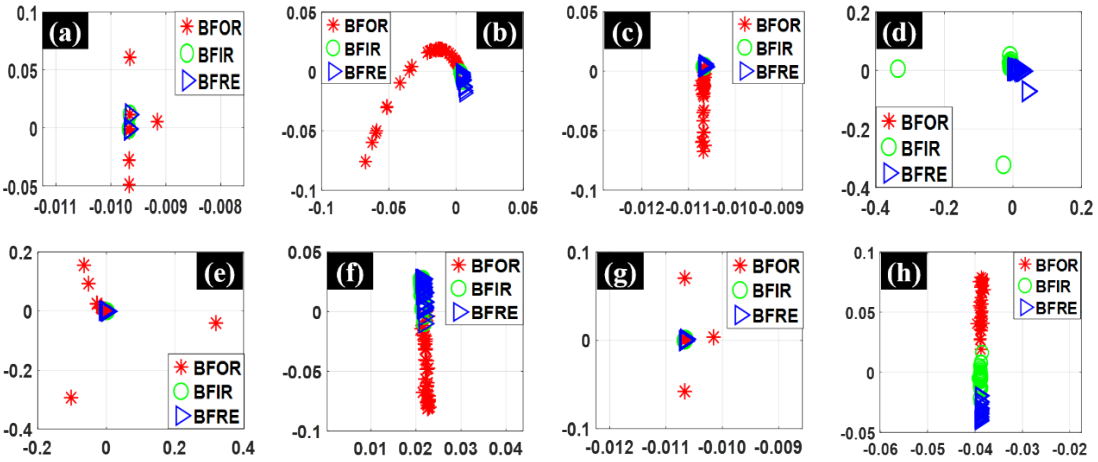



**FIGURE 5.** Visualization of clustering results on dataset 1 in the 2-dimensional latent space: (a) SSC, (b) LRR, (c) LRSC, (d) LS3C, (e) NLS3C, (f) LSRS2, (g) LATLRR, and (h) L2PLRR.

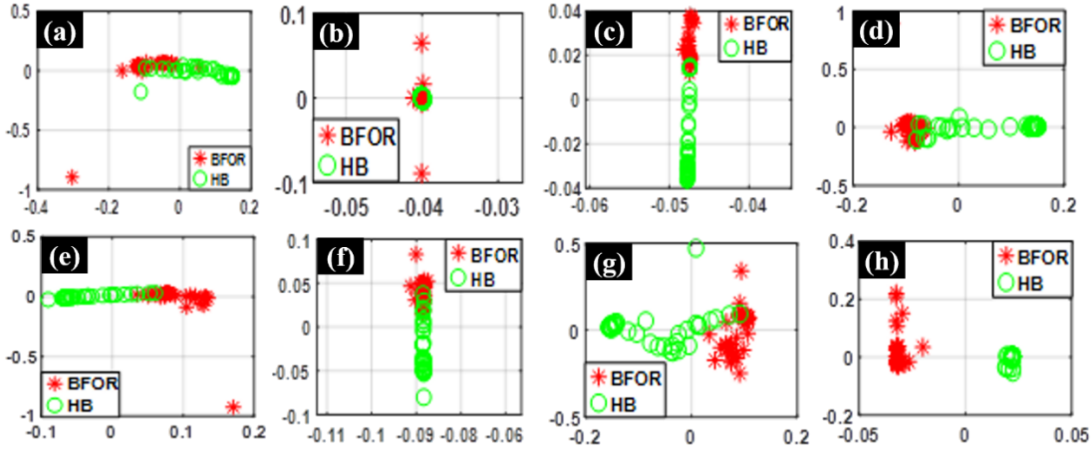



**FIGURE 6.** Visualization of clustering results on dataset 2 in the 2-dimensional latent space: (a) SSC, (b) LRR, (c) LRSC, (d) LS3C, (e) NLS3C, (f) LSRS2, (g) LATLRR, and (h) L2PLRR.

SNRs, indicating that these methods are relatively robust to noise. More specifically, the L2PLRR was capable of linearly improving the clustering accuracy until SNR = 0dB. At above SNR = 0dB, the L2PLRR showed the incisively increased clustering accuracy. The LATLRR's performance pattern was similar to that of the L2PLRR. However, the developed method was more stable than LATLRR at lower SNRs.

### *B. CLUSTERING RESULTS ON BEARING FAILURE DATASETS*

In this study, the SC methods were tested on bearing failure datasets provided by the Center for Intelligent Maintenance Systems (IMS), University of Cincinnati, USA [26]. The IMS conducted tests for 35 days until they found failures on any of bearing's elements (i.e., a cage, an inner race, an outer race, and rolling elements) by inspecting metal debris on the magnetic plugs of the bearings under test. During the run-to-failure tests, the following failures were found on the bearing's inner race (BFIR), outer race (BFOR), and rolling element (BFRE), respectively, as depicted in Fig. 4. In this study, these bearing failures are denoted as BFIR, BFOR, and BFRE, respectively. Further information about test setup and sensor replacement can be found in [26].

The SC methods were tested on two different bearing datasets. The first dataset, denoted as dataset 1, contained bearing run-to-failure data, enabling the observation of bearing degradation. Further, dataset 1 consisted of a number of 1-s vibration signals sampled at 20 kHz for a specific bearing failure—that is, dataset 1 contained 1-s vibration signals for a BFIR, a BFOR, and a BFRE, respectively. On the other hand, the second and third datasets, denoted as dataset 2 and dataset 3, consisted of 1-s vibration signals recorded at different sensors, respectively, for a healthy bearing (HB) and a BFOR.

To verify the efficacy of the SC methods, the first step is to configure a set of features that can be extracted from each of the vibration signals in the datasets. To deal with non-stationarity and non-linearity inherent in the vibration signals, ensemble empirical mode decomposition [28] was applied to

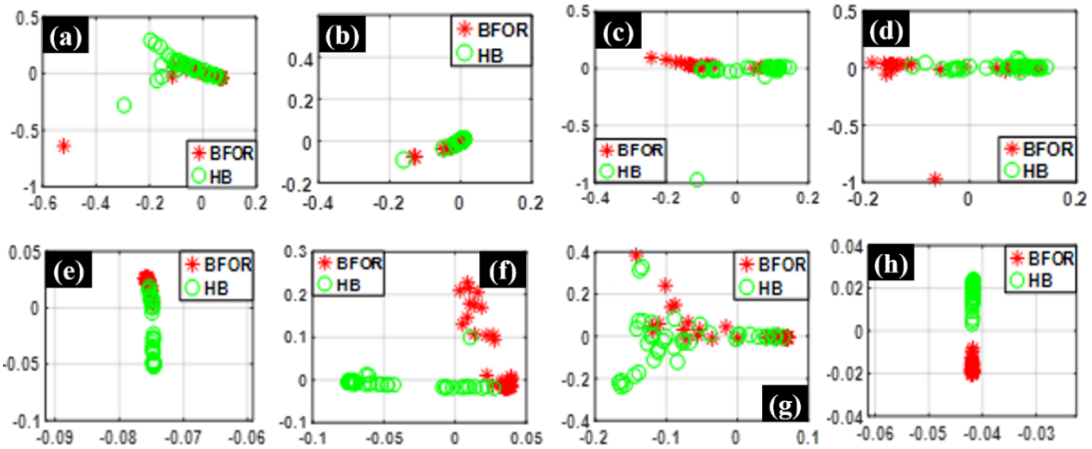



**FIGURE 7.** Visualization of clustering results on dataset 3 in the 2-dimensional latent space: (a) SSC, (b) LRR, (c) LRSC, (d) LS3C, (e) NLS3C, (f) LSRS2, (g) LATLRR, and (h) L2PLRR.

**TABLE 1.** Clustering results for purely synthetic data and corrupted data.

| Method | Purely synthetic data | Corrupted data | | | |
|---|---|---|---|---|---|
| | Average Accuracy | Accuracy | | | |
| | | Max. | Min. | Median | Average |
| SSC | 0.48 | 0.39 | 0.28 | 0.32 | 0.32 |
| LRR | 0.56 | 0.28 | 0.27 | 0.27 | 0.27 |
| LRSC | 0.97 | 0.97 | 0.27 | 0.28 | 0.40 |
| LS3C | 0.91 | 0.95 | 0.26 | 0.28 | 0.37 |
| NLS3C | 0.87 | 0.95 | 0.27 | 0.28 | 0.39 |
| LSRS2 | 0.95 | 0.93 | 0.29 | 0.33 | 0.38 |
| LATLRR | 0.99 | 0.99 | 0.29 | 0.38 | 0.53 |
| L2PLRR | 1.00 | 1.00 | 0.28 | 0.41 | 0.55 |

**TABLE 2.** Fault detection accuracy of various bearing datasets.

| Method | Accuracy | | | Average Accuracy |
|---|---|---|---|---|
| | Dataset 1 | Dataset 2 | Dataset 3 | |
| SSC | 0.54 | 0.53 | 0.44 | 0.50 |
| LRR | 0.56 | 0.51 | 0.46 | 0.51 |
| LRSC | 0.67 | 0.93 | 0.56 | 0.72 |
| LS3C | 0.44 | 0.75 | 0.53 | 0.56 |
| NLS3C | 0.52 | 0.80 | 0.45 | 0.59 |
| LSRS2 | 0.43 | 0.81 | 0.69 | 0.64 |
| LATLRR | 0.53 | 0.85 | 0.50 | 0.62 |
| L2PLRR | 0.94 | 0.99 | 0.99 | 0.97 |

decompose each of the 1-s vibration signals into so-called intrinsic mode functions (IMFs). Because the frontier IMFs tend to involve intrinsic information about bearing failures, statistical features, such as mean, peak-to-peak, and root-mean-square, were computed in the first five IMFs—the total number of features used in this study is 55.

Fig. 5 illustrates the clustering results in a 2-dimensional space on dataset 1. Note that dimensionality reduction is carried out during the process of SC. It can be said that the clustering result is satisfactory if the three bearing failures (i.e., BFIR, BFOR, and BFRE) are clearly separated from each other. In Fig. 5, the SC based on L2PLRR (see Fig. 5(h)) method was superior to the other classical and state-of-the-art clustering methods. Although L2PLRR showed distinct separation among bearing failures, the other methods failed to properly separate BFIR (see Figs. 5(a), 5(b), 5(c), 5(e), 5(g), and 5(f)) or BFOR (see Fig. 5(d)). As illustrated in Figs. 6 and 7, L2PLRR outperformed the other SC methods in facilitating the recognition of a healthy bearing and a BFIR.

Table 2 summarizes the classification accuracy of various bearing datasets. As presented in Table 2, the clustering performance of latent SC was more predominant than the classical LRR and SSC—that is, latent SC was effective for unfolding high-dimensional curled data into a low-dimensional latent space. Likewise, the L2PLRR method achieved 1.35-fold to 1.94-fold performance improvements in

terms of the average of clustering accuracies. This was mainly because of the locality-preserving effect in the L2PLRR.

## V. CONCLUSIONS

As mentioned in Section I, subspace clustering is an effective unsupervised feature learning method that is widely used in intelligent fault diagnosis owing to its ability to classify different kinds of data independent of label information. However, it is hard for conventional subspace clustering methods to extract the discriminative features that reveal the intrinsic structure of high-dimensional and non-linear data (e.g., bearing failure data), thus diminishing diagnostic performance. The developed L2PLRR method addresses this problem by mapping the data into the latent space whose dimension is lower than ones in the original space before clustering, in the meanwhile, the locality-preserving constrained mapping function is provided. Therefore, the learned low-rank coefficients or features contribute to reveal the intrinsic structure of the data.

The result is improved accuracy of bearing fault diagnosis, which is proven by comparing the L2PLRR method with the conventional subspace clustering methods that employ statistical parameters (i.e., LRR, SSC, and LRSC). The experimental results show that the L2PLRR method outperformed the most competitive LRSC and yielded 34.72% performance improvement in terms of the averages of accuracies. Likewise, due to locality-preserving, the L2PLRR method can map the high-dimensional data into low-dimensional latent space without destroying the intrinsic structure of the high-dimensional and nonlinear data that contributes to discrimination. The experimental data of the bearing fault diagnosis shows that the L2PLRR method outperformed the most competitive latent subspace clustering methods, LSRS2 and LATLRR, and yielded 51.56% and 56.45% performance improvements in terms of averages of clustering accuracy, respectively. The variation of accuracy with the increased noise also confirms the high level of robustness of LATLRR. In conclusion, L2PLRR is effective for unsupervised feature learning of complex data, which is the foundation of intelligent fault diagnosis under complicated circumstances such as bearing operation.

Although the effectiveness of the L2PLRR method is verified by bearing fault diagnosis, this method is applicable to general data-driven fault diagnosis with minor changes. The L2PLRR method will be further studied in order to make it appropriate for fault diagnosis of other machinery and electronic components or systems.

## REFERENCES

[1] F. Lu, J, Jiang, J. Huang, and X. Qiu, "Dual reduced kernel extreme learning machine for aero-engine fault diagnosis," *Aerosp. Sci. Technol.*, vol. 71, pp. 742–750, Dec. 2017.

[2] C. C. Oliveira and J. M. da Silva, "Fault diagnosis in highly dependable medical wearable systems," *J. Electron. Test.*, vol. 32, no. 4, pp. 467–479, 2016.

[3] U. Shafi, A. Safi, A. R. Shahid, S. Ziauddin, and M. O. Saleem, "Vehicle remote health monitoring and prognostic maintenance system," *J. Adv. Transp.*, vol. 2018, no. 1, pp. 1–10, Jan. 2018.

[4] M. Kang, J. Kim, J.-M. Kim, A. C. C. Tan, E. Y. Kim, and B.-K. Choi, "Reliable fault diagnosis for low-speed bearings using individually trained support vector machines with kernel discriminative feature analysis," *IEEE Trans. Power Electron.*, vol. 30, no. 5, pp. 2786–2797, May 2015.

[5] J. Uddin, M. Kang, D. V. Nguyen, and J.-M. Kim, "Reliable fault classification of induction motors using texture feature extraction and a multiclass support vector machine," *Math. Problems Eng.*, vol. 2014, no. 3, pp. 1–9, Jun. 2014.

[6] T. de Bruin, K. Verbert, and R. Babuska, "Railway track circuit fault diagnosis using recurrent neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 3, pp. 523–533, Mar. 2017.

[7] S. S. Tayarani-Bathaie, Z. N. S. Vanini, and K. Khorasani, "Dynamic neural network-based fault diagnosis of gas turbine engines," *Neurocomputing*, vol. 125, pp. 153–165, Feb. 2014.

[8] C. Dai, Z. Liu, K. Hu, and K. Huang, "Fault diagnosis approach of traction transformers in high-speed railway combining kernel principal component analysis with random forest," *IET Elect. Syst. Transp.*, vol. 6, no. 3, pp. 202–206, Sep. 2016.

[9] X. Guo, L. Chen, and C. Shen, "Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis," *Measurement*, vol. 93, pp. 490–502, Nov. 2016.

[10] M. Zhao, M. Kang, B. Tang, and M. Pecht, "Deep residual networks with dynamically weighted wavelet coefficients for fault diagnosis of planetary gearboxes," *IEEE Trans. Ind. Electron.*, vol. 65, no. 5, pp. 4290–4300, May 2018.

[11] Z. Li, H. Fang, M. Huang, Y. Wei, and L. Zhang, "Data-driven bearing fault identification using improved hidden Markov model and self-organizing map," *Comput. Ind. Eng.*, vol. 116, pp. 37–46, Feb. 2018.

[12] P. Baraldi, F. Di Maio, M. Rgamonti, E. Zio, and R. Seraoui, "Clustering for unsupervised fault diagnosis in nuclear turbine shutdown transients," *Mech. Syst. Signal Process.*, vols. 58–59, pp. 160–178, Jun. 2015.

[13] X.-H. He, D. Wang, Y.-F. Li, and C.-H. Zhou, "A novel bearing fault diagnosis method based on Gaussian restricted Boltzmann machine," *Math. Problems Eng.*, vol. 2016, Dec. 2016, Art. no. 2957083.

[14] Y. Qi, C. Shen, D. Wang, J. Shi, X. Jiang, and Z. Zhu, "Stacked sparse autoencoder-based deep network for fault diagnosis of rotating machinery," *IEEE Access*, vol. 5, pp. 15066–15079, Jul. 2017.

[15] W. Jiang, Z. Zhang, F. Li, L. Zhang, M. Zhao, and X. Jin, "Joint label consistent dictionary learning and adaptive label prediction for semisupervised machine fault classification," *IEEE Trans. Ind. Informat.*, vol. 12, no. 1, pp. 248–256, Feb. 2016.

[16] Z. Zhang, W. Jiang, F. Li, M. Zhao, B. Li, and L. Zhan, "Structured latent label consistent dictionary learning for salient machine faults representation-based robust classification," *IEEE Trans. Ind. Informat.*, vol. 13, no. 2, pp. 644–656, Apr. 2017.

[17] X. Peng, J. Feng, S. Xiao, W.-Y. Yau, J. T. Zhou, and S. Yang, "Structured autoencoders for subspace clustering," *IEEE Trans. Image Process.*, vol. 27, no. 10, pp. 5076–5086, Oct. 2018.

[18] X. Zhang, L. Jiao, F. Liu, L. Bo, and M. Gong, "Spectral clustering ensemble applied to SAR image segmentation," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 7, pp. 2126–2136, Jul. 2008.

[19] E. Elhamifar and R. Vidal, "Sparse subspace clustering: Algorithm, theory, and applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 11, pp. 2765–2781, Nov. 2013.

[20] G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, and Y. Ma, "Robust recovery of subspace structures by low-rank representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 171–184, Jan. 2013.

[21] E. L. Dyer, A. C. Sankaranarayanan, and R. G. Baraniuk, "Greedy feature selection for subspace clustering," *J. Mach. Learn. Res.*, vol. 14, no. 1, pp. 2487–2517, 2013.

[22] R. Vidal and P. Favaro, "Low rank subspace clustering (LRSC)," *Pattern Recognit. Lett.*, vol. 43, pp. 47–61, Jul. 2013.

[23] G. Liu and S. Yan, "Latent low-rank representation," in *Low-Rank and Sparse Modeling for Visual Analysis*. Cham, Switzerland: Springer, 2014, pp. 23–38.

[24] V. M. Patel, H. V. Nguyen, and R. Vidal, "Latent space sparse and low-rank subspace clustering," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 4, pp. 691–701, Jun. 2015.

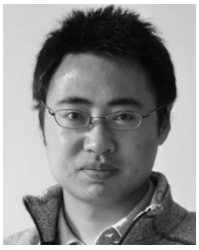
[25] L. Wei, A. Wu, and J. Yin, "Latent space robust subspace segmentation based on low-rank and locality constraints," *Expert Syst. Appl.*, vol. 42, no. 19, pp. 6598–6608, 2015.

[26] M. Ghergherehchi, S. Y. Kim, H. Afarideh, and Y. S. Kim, "RANdom sample consensus (RANSAC) algorithm for enhancing overlapped etched track counting," *IET Image Process.*, vol. 9, no. 2, pp. 97–106, Feb. 2015.
[27] J. J. Lee, H. Qiu, G. Yu, J. Lin, and Rexnord Technical Services. *Bearing Data Set*. Accessed: Jul. 26, 2018. [Online]. Available: http://ti.arc.nasa.gov/project/prognostic-data-repository
[28] Z. P. Feng *et al.*, "Fault diagnosis for wind turbine planetary gearboxes via demodulation analysis based on ensemble empirical mode decomposition and energy separation," *Renew. Energy*, vol. 47, pp. 112–126, Nov. 2012.

**JIE GAO** received the B.S. degree in electrical and information engineering from the Tianjin University of Technology in 2013 and the M.S. degree in communication and information systems from Capital Normal University in 2017. She is currently pursuing the Ph.D. degree with the Institute of Automation, Chinese Academy of Sciences. Her research interests include signal processing, machine learning, system modeling, and statistics for intelligent prognostics and systems health management.

**MYEONGSU KANG** (M'17) received the B.E. and M.S. degrees in computer engineering and information technology and the Ph.D. degree in electrical, electronics, and computer engineering from the University of Ulsan, Ulsan, South Korea, in 2008, 2010, and 2015, respectively.

He passed away in 2018 while he was a Research Scientist with the Center for Advanced Life Cycle Engineering, University of Maryland, College Park, MD, USA. His research interests included data-driven anomaly detection, diagnostics, and prognostics of complex systems, such as automotive, railway transportation, and avionics.



**JING TIAN** received the B.Eng. degree in machinery design and manufacturing from the University of Electronic Science and Technology of China, Chengdu, China, and the Ph.D. degree in mechanical engineering from the University of Maryland, College Park, MD, USA.

In 2010, he was an Engineer and also a Researcher for seven years. He has been a Research Assistant with the Center for Advanced Life Cycle Engineering (CALCE), University of Maryland. He is currently with The DEI Group His research interests include the development of machine learning algorithms and data analysis tools for condition-based maintenance and prognostics and health management.

**LIFENG WU** (M'12) received the B.S. degree in applied physics from the China University of Mining and Technology in 2002, the M.S. degree in detection technology and automation device from Northeast Dianli University in 2005, and the Ph.D. degree in physical electronics from the Beijing University of Posts and Telecommunications in 2010. He is currently an Assistant Professor with Capital Normal University. His research interests include data-driven modeling, estimation and filtering, fault diagnosis, power electronics, and electrical vehicles.

**MICHAEL PECHT** (S'78–M'83–SM'90–F'92) received the B.S. degree in acoustics, the M.S. degrees in electrical engineering and engineering mechanics, and the Ph.D. degree in engineering mechanics from the University of Wisconsin at Madison, Madison, WI, USA, in 1982.

He is the Founder of the Center for Advanced Life Cycle Engineering, University of Maryland, College Park, MD, USA, where he is also a Chair Professor. He has been leading a Research Team in the area of prognostics.

Dr. Pecht is a Professional Engineer and an ASME Fellow. He received the IEEE Undergraduate Teaching Award and the International Microelectronics Assembly and Packaging Society (IMAPS) William D. Ashman Memorial Achievement Award for his contributions in electronics reliability analysis. He served as the Chief Editor of the IEEE TRANSACTIONS ON RELIABILITY for eight years and an Associate Editor for the IEEE TRANSACTIONS ON COMPONENTS AND PACKAGING TECHNOLOGY.