Optimal Singular Value Decomposition Based Big Data Compression Approach in Smart Grids

Naser Hashemipour, Jamshid Aghaei, *Senior Member, IEEE*, Abdollah Kavousi-fard, Taher Niknam, Ladan Salimi, Pedro Crespo del Granado, Miadreza Shafie-Khah, *Senior Member, IEEE*, Fei Wang, *Senior Member, IEEE*, and João P. S. Catalão, *Senior Member, IEEE*

***Abstract*—The smart grid is a fully automatic delivery grid for electricity power with a two-way reliable flow of electricity and information among different equipment on the grid. Smart meters and sensors monitoring the system provide a huge amount of data in various part of smart grid. To logically manage this trouble, a new lossy data compression approach for big data compression is proposed. The optimal singular value decomposition (SVD) is applied to a matrix that achieves the optimal number of singular values to the sending process, and the other ones will be neglected. This goal is done due to the quality of retrieved data and the compression ratio. In the presented scheme, to implement the optimization framework, various intelligent optimization methods are used to determine the number of optimal values in the elimination stage. The efficiency and capabilities of the proposed method are examined using a wide range of data types, from electricity market data to image processing benchmarks. The comparisons show that the compression level obtained by the proposed method can dominate the points given by the existing SVD rank reduction methods. Also, as the other finding of this article, the performance of the rank reduction methods depends on the application and data types. It means that a rank reduction method can reveal a good performance in one application and performs unacceptably for another purpose. So, the optimized rank reduction can pave the way toward a robust and reliable performance.**

***Index Terms*—Big data, data compression, optimization, singular value decomposition (SVD), smart grid.**

Manuscript received July 29, 2020; revised December 26, 2020 and March 3, 2021; accepted March 22, 2021. Date of publication April 15, 2021; date of current version July 16, 2021. Paper 2020-IDC-0212.R2, presented at the 2020 IEEE Industry Applications Society Annual Meeting, Detroit, MI USA, Oct. 10–16, and approved for publication in the IEEE TRANSACTIONS ON INDUSTRY APPLICATIONS by the Energy Systems Committee of the IEEE Industry Applications Society. The work of João P. S. Catalão was supported in part by FEDER funds through COMPETE 2020 and in part by Portuguese funds through FCT, under POCI-01-0145-FEDER-029803 (02/SAICT/2017). *(Corresponding author: Fei Wang.)*

Naser Hashemipour, Abdollah Kavousi-fard, and Taher Niknam are with the Shiraz University of Technology, 71557-13876 Shiraz, Iran (e-mail: n.hashemipour@sutech.ac.ir; kavousi@sutech.ac.ir; niknam@sutech.ac.ir).

Jamshid Aghaei is with the Department of Electrical and Electronics Engineering, Shiraz University of Technology, Shiraz, Iran, and also with the Department of Electrical Engineering, School of Energy Systems, Lappeenranta University of Technology, 53850 Lappeenranta, Finland (e-mail: aghaei@sutech.ac.ir).

Ladan Salimi is with the Razi University, Kermanshah 67144-14971, Iran (e-mail: salami.ladan@stu.razi.ac.ir).

Pedro Crespo del Granado is with the Norwegian University of Science and Technology Trondheim, 7491 Trondheim, Norway (e-mail: pedro@ntnu.no).

Miadreza Shafie-Khah is with the School of Technology and Innovations, University of Vaasa, 65200 Vaasa, Finland (e-mail: mshafiek@univaasa.fi).

Fei Wang is with the Department of Electrical Engineering, North China Electric Power University, Baoding 071003, China, with the State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, North China Electric Power University, Beijing 102206, China, and also with the Hebei Key Laboratory of Distributed Energy Storage and Microgrid, North China Electric Power University, Baoding 071003, China (e-mail: feiwang@ncepu.edu.cn).

João P. S. Catalão is with the Faculty of Engineering, University of Porto and INESC TEC, Porto 4200-465, Portugal (e-mail: catalao@fe.up.pt).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TIA.2021.3073640.

Digital Object Identifier 10.1109/TIA.2021.3073640

NOMENCLATURE

A	Original data.
$\bar{A}$	Reconstructed data.
C_r	Compression ratio.
D_r	Elements of remained data.
D_d	Elements of deleted data.
D_o	Elements of original data.
GAm	Genetic algorithm with mutation.
m	Rows of the original matrix.
n	Columns of the original matrix.
N_t	Threshold of Euclidean norm for comparison of original and retrieved data.
$N_2(Z)$	Euclidean norm of matrix Z.
p	Number of deleted singular values.
S_d	Number of deleted singular values.
α	Weight coefficient for N_t

I. INTRODUCTION

A. Data in Smart Grids

THE smart grid is an intelligent electricity grid that optimizes the generation, distribution, and consumption of electricity through the introduction of information and communication technologies on the electricity grid that includes smart meters and various sensors in different parts. The measurement and monitoring instruments to gathering the information in the transmission system and medium-voltage level distribution system are managed by supervisory control and data acquisition and wide-area monitoring system (WAMS). Similarly, in the level of consumers, advanced metering infrastructure and automatic meter reading systems are employed for data gathering in the smart grid. Phasor measurement units (PMUs) are among the other units used in the smart grid to measure the required information and send it through a communications platform.

Fig. 1 shows the general structure of the WAMS system in the smart grid. Information for each PMU is transmitted through public switched telephone networks, fiber optic cables, low

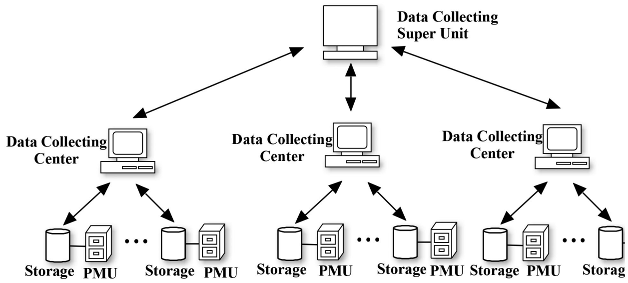

Fig. 1. WAMS structure in a smart grid.

altitude satellites, power line carriers, or microwave links. As a result, a huge amount of multisource varied data is stored in the smart grids. These data, if exploited properly, can reveal much information about the customers and generating units and improve the power quality and smart grid efficiency.

A challenge in this way is the huge volume of the transmitted information and the limited bandwidth for the data transfer. In this situation, data compression techniques can bring great benefit to the smart grid. In compression methods, the initial goal is to reduce the data size. But, it can happen as long as the compressed data contain the main features of the original information [1]. Data compression is broadly classified into two categories of lossless and lossy [2], and various researchers are actively engaged to propose efficient methods of data compression. The following sections review the important techniques that are being used commonly in data compression.

B. Literature Review

Different lossy compression schemes have been developed for smart grid applications based on wavelet decomposition (WD) [3]–[6], discrete cosine transform (DCT) [7], fuzzy-based methods [8], compressed sensing theory [9], and singular value decomposition (SVD)-based approaches [10]–[15]. There also is ongoing research on spatial domain methods like neural network (NN)-based methods [16], [17] and deep stacked autoencoders [18]. Also, different methods of lossless compression for data compression applications have been emerged [19], [20].

Khan *et al.* [3] have introduced a novel method for simultaneous signal compression and denoising in smart grids. According to [3], the wavelet packet decomposition (WPD) is more accurate than WD. In this method, the WD tree has been converted to a fully binary tree using a cost function, and the best tree has been selected from several WPD bases. Besides, reconstruction of a noisy signal can easily be done by setting a threshold. This work provides an acceptable compression ratio and a good de-noising tool for signals. Ning *et al.* [4] have suggested another compression technique using wavelet transform (WT) and multiresolution analysis. In the decomposition process, the Daubechies filter is considered as the mother signal. Experimental results illustrate that the WT-MRI can only deal with the white noise of the signals and cannot compress the data adequately. So, it must be combined with other algorithms. Also, Khan *et al.* [5] presented an approach based on embedded zero tree WT (EZWT) that depresses the noisy elements of the grid signals. Since EZWT does not require tables and codebooks for signal recovery, it is a simple and efficient method. EZWT allows us to carry out both compression and denoising of the PMU and power system data. A similar approach in [6] with the WD has been deployed for signal compression. In this method, after performing wavelet packet, dynamic bit allocation is carried out by calculating the neural shape estimator (NSE) to estimate the spectral shape that is necessary to eliminate data redundancy and implementing the entropy coding. The results showed that NSE is more successful than the other estimators that provide an acceptable ratio for compression of waveforms. Gadde *et al.* [7] have presented a cascade technique in PMU data compression using intrinsic correlation that discovers spatial and temporal redundancies. In this work, the principal component analysis is defined. Then, a DCT associated with each component is modeled. The required compression parameters have been adjusted using efficient statistical techniques. The results show that this approach can be employed for phasor data concentrators fed from any number of PMUs. A research in [8] suggested fuzzy-based approaches to save the required memory and bandwidth, which reduces the computational burden for smart grid data analysis. Some works using compressed sensing theory for smart grid applications have been developed, such as reference [9] that provided the compression technique for electricity datasets. At the decoder side, an iterative threshold algorithm has been employed to reconstruct the compressed bitstream. A good performance for both compression/decompression of the considered data was concluded from the results. A linear algebra based technique, such as SVD, is widely used as another tool for data compression. The main purpose of SVD-based methods is to approximate the original data with a rank reduced matrix. Many efforts [10]–[14] have been made on the development of data compression using SVD decomposition. de Souza *et al.* [10] proposed an algorithm that performs the SVD data compression on a power system dataset. The level of the compression is determined by the bottleneck of the communication links of the grid. Then, an accurate enough compression level that satisfies the bottleneck constraint is found by a trial and error procedure. A similar method based on the combination of SVD and wavelet difference reduction (WDR) is proposed in [11]. In this study, at first, the approximation of the original image is achieved through iterative rounds of test and error, and then WDR has been applied to reduced data. Indeed, WDR has been added to the SVD decomposition to enhance the compression ratio. Thresholding techniques based on the energy information are among the SVD rank reduction methods. Padhy *et al.* [12] selected a thresholding technique based on multiscale root fractional energy contribution of the singular values to reduce the dimensionality of orthogonal and singular value matrices for Electrocardiogram(ECG) compression. In this manner, at first, discrete wavelet transformation (DWT) is applied to the original matrix. Then, energy variations in different subbands are calculated. The singular values of each subband that are greater than the threshold are retained. Dixit *et al.* [13] have worked on SVD-DCT based compression of images. In this article, the authors test the quality of the image with different ranks between 70 and 150 and make a tradeoff between the quality and the compression. Yu *et al.* [14] presented a lossy compression image based on SVD. Their approach to select the rank of the image is similar to [13], but in the range between 10 and 200. Mukhopadhyay *et al.* [15] have worked on biosignals compression algorithm using a combination of downsampling, SVD, and American Standard Code for Information Interchange(ASCII) compression method.

In the spatial domain, artificial intelligence based methods, such as NNs [16] and intelligent measurement techniques [17], have also been used in recent years. Barrosa *et al.* [16] proposed a compression method based on genetic algorithm (GA) and NN for electrical power signals. The GA is used to select the best samples of the signal, and then NN is deployed to compress the remaining samples and reconstruct the signal. The rate of compression is 2.5%–10% for an installed recorder in a 230-kV electrical power system. In [17], an approach has been suggested for the compression of electricity load data, based on the intelligent measurement. As another application of compression in the electrical load dataset, Huang *et al.* [18] provided a technique based on deep stacked autoencoders.

In addition to the compression application, SVD-based approximation is used in other areas, such as noise reduction [21]–[23], image reconstruction [24], shot boundary detection in the video [25], and simultaneous compression and denoising [26]. In several SVD-based rank reduction research works, such as those in [10], [11], and [13]–[15], the number of singular values that contain the main information of the matrix is determined in a trial and error procedure. It means that the operator starts with an initial guess for the number of singular values to retain. If it is good enough, this guess is chosen. Otherwise, it should be changed to get a better solution. It is worth noting that the knee point in the singular value diagram can be considered as a good starting point. In addition to these approaches, there are some methodologies, such as the Guttman–Kaiser criterion [26], [27], Cattell's scree test [26], and entropy-based methods [28], to determine rank of the truncated SVD. In the first method of the Guttman–Kaiser criterion, all singular values smaller than one are ignored. The second method of the Guttman–Kaiser criterion keeps enough number of singular values such that the squared summation of them covers 90% of the squared sum of all of the singular values. Both of these methods are based on arbitrary thresholds. The scree test is a subjective decision on the rank based on the shape of the scree plot. Indeed, the operator can determine an appropriate matrix rank based on the knee point in the singular value diagram. But, in some data types, the Singular Value(SV) curve may be nonconvex (more than a knee region) and it can be more challenging to decide between various knee points. In entropy-based methods, relative contribution of singular values determines the rank of the matrix.

In noise reduction, some of the singular values with greater energy have been kept and the other ones are considered as noisy singular values [21]. For example, Liu and Liu [22] presented an SVD-based denoising system that removes noisy elements from the data by rank reduction of the original matrix in the frequency domain based on the screen test method. Image reconstruction by inpainting is a popular area of research. The damaged region of the image is constructed by the rank reduced approximation of the main image. Here again, decision making about the proper rank of the image becomes a matter of controversy. Alilou and Yaghmaee [24] determined the matrix rank according to the structural similarity index (SSIM). Indeed, they try to find a good enough matrix rank that keeps the SSIM higher than a prespecified value. The rank reduction based on the second method of Guttman–Kaiser is employed by Youssef *et al.* [25] for the feature extraction in video shot boundary detection. Schanze *et al.* [26] divided the original data into two parts of useful and noisy elements to do compression and denoising, simultaneously. Determining the proper cutoff number of k that is the threshold to data division was the main challenge there. The scree was employed to deal with this problem.

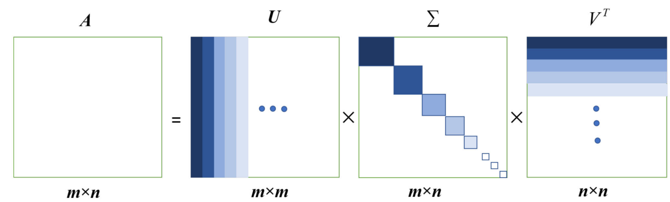

Fig. 2. SVD form of a matrix.

II. Compression in Smart Grids

The generation, transmission, and distribution of power in smart power systems are deeply impressed by data analysis. Therefore, a considerable increase in data exchange and the required memory is likely to occur, and the required data storage and bandwidth of the communication links in the smart grids have a growing trend. Besides, to obtain accurate and real-time information of the smart grid, the frequency of sampling should be increased. Accordingly, the importance of data compression in the smart grid will be more highlighted. The proposed compression method is presented in the following. This method can be employed effectively in different points of the grid where the volume of the sent and received data is high.

A. SVD Decomposition

The SVD is a computational tool for approximating a matrix by three other matrices. Indeed, it decomposes the matrix A into U, V, and Σ. Let us assume m and n be arbitrary, and A is a matrix. An SVD of A is a factorization, as can be seen in (1).

$$A = U\Sigma V^T \tag{1}$$

where U is a $m \times m$ real or complex unitary matrix, Σ is a $m \times n$ rectangular diagonal matrix with nonnegative real numbers on the diagonal, and V is an $n \times n$real or complex unitary matrix. The diagonal entries σ_i of Σ are known as the singular values of A. Briefly [29] the following statements hold:

1) U: is $m \times m$ unitary (the left singular vectors of A);
2) V: is $n \times n$ unitary (the right singular vectors of A);
3) Σ: is $m \times n$ diagonal (the singular values of A)

where $\Sigma_{(m\times n)} = \mathrm{diag}(\sigma_1, ..., \sigma_n)$ with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$.

In Fig. 2, the SVD decomposition on a matrix has been illustrated. As already mentioned, Σ is a diagonal matrix whose elements on its original diameter are singular values that are placed in descending order. Each singular value is involved in the retrieving process of the original matrix. In other words, (1) can be rewritten in the form of the following equation [29]:

$$A = \sum_{i=1}^{m} u_i \sigma_i v_i^T \tag{2}$$

where u_i and v_i are the left and right singular vectors of the matrix A, respectively, and σ_i is the ith singular value. As can be seen, the smaller singular values play a smaller role in the building of the original data. Thus, the low-rank matrix approximation can be obtained by the elimination of the smaller values and the original information can be retrieved as can be seen in (3a). According to (3b), Σ is decomposed into a submatrix including the important singular values ($\overline{\sum}$) and three nonimportant submatrices that are replaced by zero matrices with the same dimension

$$\overline{A} = \overline{U}\overline{\sum}\overline{V}^T \tag{3a}$$

$$\begin{aligned} \Sigma &= \begin{bmatrix} \bar{\Sigma}_{(n-p)*(n-p)} & 0_{(n-p)*p} \\ 0_{(m-n+p)*(n-p)} & 0_{(m-n+p)*p} \end{bmatrix} \\ U &= \begin{bmatrix} \bar{U}_{m*(n-p)} & \bar{U}_{m*m-(n+p)} \end{bmatrix} \\ V &= \begin{bmatrix} \bar{V}_{n*(n-p)} & \bar{V}_{n*p} \end{bmatrix}. \end{aligned} \tag{3b}$$

As can be seen in the following equation, $\overline{A}$ is a low-ranked approximation $m \times n$ matrix, the matrix $\overline{U}$ is $m \times\ (n-p)$, $\overline{V}$ is $n \times\ (n-p)$, and $\overline{\sum}$ is $(n-p) \times (n-p)$:

$$A_{m*n} = \begin{bmatrix} \bar{U}_{m*(n-p)} & \bar{U}_{m*m-(n+p)} \end{bmatrix} \times \begin{bmatrix} \bar{\Sigma}_{(n-p)*(n-p)} & 0_{(n-p)*p} \\ 0_{(m-n+p)*(n-p)} & 0_{(m-n+p)*p} \end{bmatrix} \times \begin{bmatrix} \bar{V}_{n*(n-p)} & \bar{V}_{n*p} \end{bmatrix}^T. \tag{4}$$

Since $\overline{U}$, $\overline{V}$, and $\overline{\sum}$ can provide an acceptable approximation of the original data (A) and can be sent instead of A, their dimension specifies the compression ratio.

Therefore, the compression ratio is defined as follows:

$$C_r = \frac{m \times n}{(n-p)(m+n+1)}. \tag{5}$$

The compression is done when the ratio is more than 1. The following equations provide a lower bound on the number of neglected singular values:

$$\frac{m \times n}{(n-p)(m+n+1)} > 1 \tag{6}$$

$$(n-p) < \frac{m \times n}{(m+n+1)} \tag{7}$$

$$p > n - \frac{m \times n}{(m+n+1)}. \tag{8}$$

Along with the compression ratio, the redundancy of data can also be calculated by the following equation:

$$\%R = 1 - \frac{1}{C_r}. \tag{9}$$

In this process, according to (4), we have

$$\begin{aligned} C_r &= \frac{D_o}{D_r} = \frac{D_r + D_d}{D_r} \Rightarrow C_r = 1 + \frac{D_d}{D_r} \\ &\Rightarrow D_d = D_r (C_r - 1). \end{aligned} \tag{10}$$

To find the redundant data, both sides of (10) is divided by the original data, so

$$R = \frac{D_d}{D_o} = \frac{D_r}{D_o} \times (C_r - 1) \overset{\frac{D_r}{D_o} = \frac{1}{C_r}}{\Rightarrow} \%R = \left(1 - \frac{1}{C_r}\right) \times 100. \tag{11}$$

Therefore, the percentage of redundant data is calculated by (11).

Another important issue refers to the calculation of the retrieved data precision in the decoding process. The Euclidean norm criteria in the following equation has been used to check the accuracy and proximity between original and retrieved data:

$$N_2(A) = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{m} (a_{ij})^2}. \tag{12}$$

Of course, it can be replaced by other measures like mean square error.

B. Optimization Framework

Determining the number of singular values to retain can significantly affect the performance of the SVD-based data compressions. The more singular values would lead to more accuracy, and of course, less compression ratio. As mentioned in the literature review, there are various SVD rank reduction methods. Here, the proposed method presents the optimal SVD rank reduction that maximizes the compression ratio, subject to the accuracy constraint. Before explaining the optimal SVD in Section III, the general form of an optimization problem that can be seen in the following equation is reviewed:

$$\begin{aligned} &\min F(X) \\ &\text{s.t. } H(X) = 0\, G(X) \le 0. \end{aligned} \tag{13}$$

where $F$is the objective function, X is the set of decision variables, H is the equality constraints, and G represents the inequality constraints [30], [31]. Taking the feasible region of the problem into account, the minimization problem aims to find the optimal point which satisfies the following equation:

$$\forall x \in X \Rightarrow F(x^*) \le F(x). \tag{14}$$

Based on (14), the point given by the proposed method dominates the other points in the feasible region. Evolutionary algorithms have been used to solve the formulated optimization problem [32], [33].

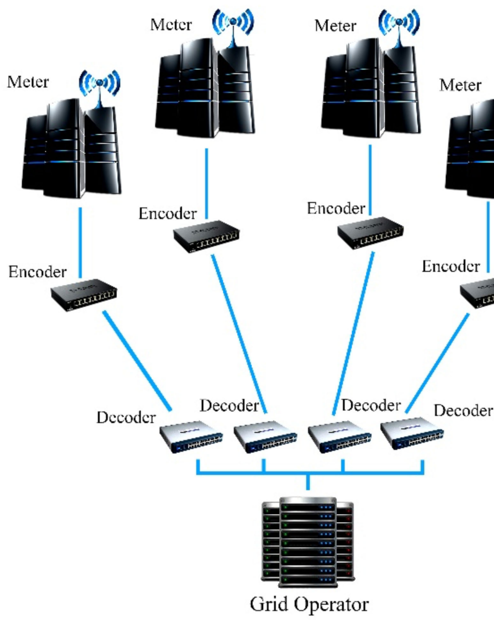

Fig. 3. Data communication channel and transmission process.

III. Proposed Method

As mentioned, data exchange between various regions of the smart grid is happening increasingly. The PMU data, bids, and offers of the participants in a P2P market, weather data including solar, wind, humidity, and temperature data, etc., that are very important for optimal, resilient, and reliable operation are among the most omnipresent examples of data production and exchange. Generally, the data exchange process is divided into encoding and decoding phases. In fact, in the encoding stage, data get prepared for transmission by some processes. After this operation, the data are transmitted through the communication channel and will be recovered to the initial form in the decoding phase. This process is shown in Fig. 3.

In the proposed method, the original data matrix is decomposed into U, V, and Σ by the SVD decomposition. Then, the optimal rank reduction is specified and three rank-reduced matrices will be ready to be sent. The decision variable (x) in this problem is the number of singular values that will be eliminated. In fact, the number of eliminated singular values influences the compression ratio and data quality. Since the main goal of the problem is data compression, the inverse of the compression ratio in the following equation can be considered as the objective function in a minimization problem:

$$F\ (x) = \frac{1}{C_r\,(x)} = \frac{(n-x)\,(m+n+1)}{m \times n}. \tag{15}$$

By minimizing (15), the compression ratio will be maximized.

Alternatively, the objective function can be replaced by minus x, as can be seen in the following equation:

$$F(x) = -x. \tag{16}$$

Indeed, more compression ratio can be gained by maximizing the number of neglected singular values.

As an important consideration, the compression is valuable if the information can be retrieved with an acceptable accuracy. Therefore, the following constraint reinforces the proximity of two matrices with Euclidean norm criteria:

$$N_2\left(A-\overline{A}\right) \leq N_t. \tag{17}$$

In other words, introducing (17) indicates the Euclidean norm of the difference between the original and recovered matrix is less than the N_t value.

In this step, an important issue refers to determining the upper bound of the constraint (N_t). N_t can be determined based on the matrix. For further information, please refer to https://data.london.gov.uk/dataset/smartmeter-energy-use-data-in-london-households as can be seen in the following equation:

$$N_t = \alpha \times N_2\,(A). \tag{18}$$

In this way, we will have a better insight into this threshold. In this regard, the upper bound of (17) is restricted to a fraction of the Euclidean norm of the original matrix.

To convert a constrained minimization problem to an unconstrained one, the objective function can be penalized for any violation of the constraint. Accordingly, the fitness function is calculated as follows:

$$\text{fit} = F\,(x) + K \times \max\left\{0, N_2\left(A-\overline{A}\right) - N_t\right\} \tag{19}$$

where K is the penalty coefficient.

As mentioned, the number of singular values ignored is the decision variable. So, x can change between the lower bound based on the following equation and the upper bound that is the total number of the singular values:

$$\begin{aligned} x > n - \frac{m \times n}{(m+n+1)} &\Rightarrow x_{\min} \\ &= round\left(n - \frac{m \times n}{(m+n+1)}\right) + 1. \end{aligned} \tag{20}$$

In some cases, the communication network bottlenecks play the main role in determining the compression ratio [10]. So, the lower bound expression can be modified based on the communication network requirements, as shown in the following equations:

$$\frac{m \times n}{(n-x)\,(m+n+1)} > \text{cr}_0 \tag{21}$$

$$x > n - \frac{1}{\text{cr}_0}\,\frac{m \times n}{(m+n+1)} \tag{22}$$

where cr_0 is the required compression ratio that is determined based on the communication network status.

To sum up, minimizing the objective function (16), subject to the constraints (17) and (22), returns the optimal number of singular values to retain. It is worth noting that the proposed optimization framework contains one decision variable (x) regardless of the dimension of the original data. Besides, the proposed optimization is applied to the output of the SVD decomposition. It means that before the algorithm starts, the SVD decomposition should be applied to the original data. Then, the proposed method optimizes the compression ratio using simple operations like the multiplication of U, V, and Σ. As a result, it can be solved by the existing heuristic optimization algorithms efficiently.

IV. Numerical Results and Discussions

Various datasets are employed to demonstrate the effectiveness of the proposed method. At first, a data matrix is employed to investigate the performance of the existing heuristic algorithms in solving the defined optimization problem (*Case 1*). In this case, a dataset from Day-Ahead Energy Market of New England's wholesale electricity marketplace on January 1, 2018, for 315 participants has been tested. It should be noted that each agent must send the five-segment (price–power curve) for 24 h. Hence, the matrix has 315 rows (equal to the number of market participants) and 240 columns (5 segments of power and 5 segments of the price for 24 h). This dataset is available in [34]. After this case, the proposed methodology is compared with the first and second methods of Guttman–Kaiser rank reduction [26], [27] as well as the proposed method by de Souza *et al.* [10] to analyze the accuracy and optimality of the algorithm. Also, the knee point on the singular value diagram is considered as another potential point in the comparisons. To this end, a variety of data types such as energy consumption[1] and renewable production[2] of 80 houses in the U.K. [35], [36] for nine months with half an hour time-step (*Case 2*) and the electrical data over a single 24-h period from 443 unique houses on February 4, 2011 [37] (*Case 3*) are selected. Moreover, some of the known image processing benchmarks like Lena and Cameraman images [38] (*Case 4*) are added to the comparison cases to cover more data types and have more concrete results.

The presented scheme has been implemented using MATLAB 2018a on the market data. All implementations have been done on a PC Intel core i7 processor 1.8 GHz, with 8GB RAM.

Case 1. Differential evolution (DE) [39], simulated annealing (SA) [40], teaching learning based optimization (TLBO) [41], particle swarm optimization (PSO) [42], as well as the GA with and without the mutation (GA-M and GA) [43] are employed to investigate the ability of the existing evolutionary algorithms in solving the proposed optimization framework. Table I shows the average of the obtained objective function over 20 runs. As can be seen, except for the GA, all other algorithms could converge to the optimal solution (or very close to the optimal solution) with zero standard deviation that implies the robustness of the algorithms. The number of iterations in each run and the corresponding run-time are shown in this table, as well. According to this table, PSO is the fastest algorithm that can find the optimal solution among all algorithms used in this case study. Along with PSO, TLBO converges to the optimal solution with robust performance. DE is an elementwise algorithm and usually needs higher iterations to converge the optimal solution.

[1]For further information, please refer to http://www.soda-pro.com/webservices/radiation/helioclim-3-archives-for-free and https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/.

[2]For further information, please refer to https://data.london.gov.uk/dataset/smartmeter-energy-use-data-in-london-households.

TABLE I
Obtained Objective Values by Evolutionary Algorithms

	DE [39]	SA [40]	TLBO [41]	PSO [42]	GA-M [43]	GA [43]
Mean	-209	-209	-210	-210	-210	≈-206
Standard deviation	0	0	0	0	0	4.51
Iteration	14	150	5	6	6	100
Run time (second)	1.58	2.76	1.17	0.77	1.82	4.37

TABLE II
Calculation of Compression Ratio and Percentage of Redundant Data in *Case 1*

Algorithm	C_r	R%
DE [39]	4.3861	77.20
SA [40]	4.3861	77.20
TLBO [41]	4.5323	77.93
PSO [42]	4.5323	77.93
GA_m [43]	4.5323	77.93
GA [43]	3.999	74.99

So, it most probably reaches the exact optimal solution with high enough iterations. But it converges to a quite close point to the global optima with 14 iterations. It also is seen that the optimization framework even can be solved with a local search algorithm like SA. GA is the only one that does not provide as satisfying results as the others. But, its performance surges with introducing the mutation operator to this algorithm and converges to the global solution like PSO and TLBO.

Table II shows the compression ratio and the percentage of redundant data for the obtained solution by each algorithm. Of course, the algorithms with the same objective value propose the same compression ratio. According to the percentage of redundant data in this table, solving the proposed optimization framework leads to a considerable compression of data. But, the goal of this case study is to check the solvability of the proposed optimization problem. The quality of the compression scheme is evaluated in the next cases through comparisons with the other rank reduction methods.

Case 2. Since (17) constrains the recovering error of the data compression, three related cases to α that determines the N_t are addressed, as are shown in Fig. 4(a)–(c). In the first case, it is assumed that the maximum tolerable error must be lower than 0.5% of the Euclidean norm of the original data ($\alpha = 0.5\%$). The performance of various methods is illustrated in Fig. 4 (a). The obtained point by the second method of the Guttman–Kaiser corresponds to the error of 11.66% of the Euclidean norm of the original data that is considerably higher than the tolerable error. This percentage is equal to 0.48 in the proposed method. It means that the proposed method compresses the data as much as possible concerning the maximum allowed error. It is worth noting that the first method of the Guttman–Kaiser and the method by de Souza *et al.* [10] satisfy the accuracy constraint with a lower compression ratio. If we move on to the second one

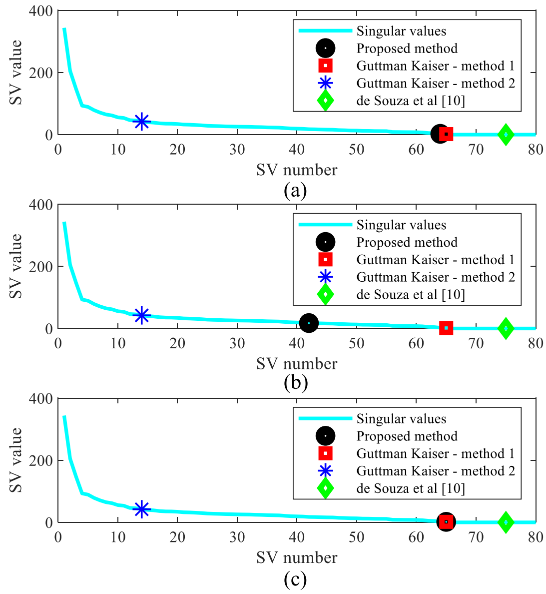

Fig. 4. Comparison of the rank reduction methods with upper bound of the error equal to (a) 0.005, (b) 0.05 and (c) 0.0001 of the Euclidean norm of the original data.

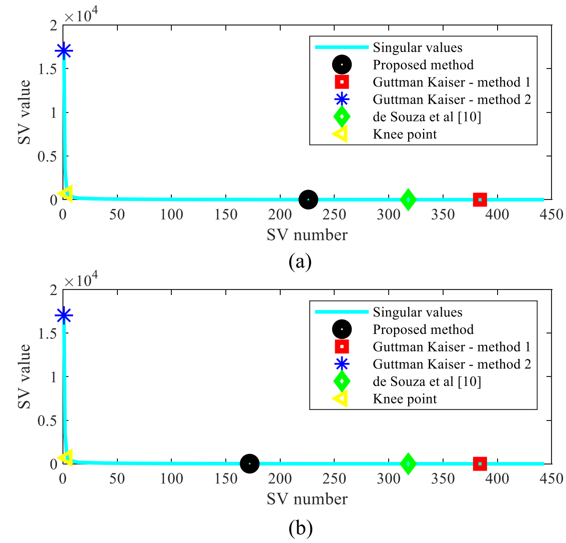

Fig. 5. Impact of the singular values in recovering error with (a) $\alpha = 0.0001$ and (b) $\alpha = 0.001$.

with $\alpha = 5\%$, the proposed method increases the compression ratio, as can be seen in Fig. 4(b). This point is equivalent to $\alpha = 4.99\%$. It means that if one more singular value is eliminated, the error goes above the prespecified value. However, the solutions given by the other methods remain unchanged. In the third case, if we are very strict with the accuracy of the recovered data and, for instance, set $\alpha = 0.01\%$, the proposed method suggests the same point that is obtained by the first method of the Guttman–Kaiser. Of course, it might be different in other case studies. This feature enhances the flexibility of the proposed method.

The performance of the SVD-based data compression can be affected by the data type. For example, two matrices of the same size may require a different number of SVs to meet the same error threshold, depending on the dependence of the rows or columns of the matrices. Therefore, to cover various data types and have a concrete conclusion, more comparisons on three other case studies are presented in the following. The first one covers the different features of each home in a microgrid for a 24-h period. The size of the database is 1440×443 and is available in [37]. The next two are selected from the image processing benchmarks, as they might have different characteristics or different applications from the power system data.

Case 3. When the rows of the matrix are dependent, or there is a correlation between them, decision making gets more complicated. In these cases, most of the SVs have small values, and suddenly after a knee point, their value increases sharply. In this situation, it is a challenging problem to determine the number of singular values to retain, based on the methods that do not react to the accuracy constraint; because there is a huge chance for compression due to the significant number of small singular values. On the other hand, the recovering error caused by ignoring the SVs might be out of the acceptable range after a point far from the knee point in the singular values diagram. Fig. 5(a) and (b) illustrates the importance of the proposed method in this situation. In some rank reduction applications, it is suggested to retain the singular values that are bigger than the knee point that is shown by yellow triangular in Fig. 5(a) and (b). It should be noted that the reconstruction error of the data at this point is 4%. This error can lead to a significant change in the recovered data. On the other side of the SV graphs of Fig. 5(a) and (b), there are the points proposed by the first method of the Guttman–Kaiser and de Souza *et al.* They provide good accuracy. But, they do not compress the data when still there is a good error margin. The other issue about the first method of Guttman–Kaiser is that the point by this method violates the lower bound of the ignored SVs as introduced in (22). According to this equation, at least 105 singular values must be ignored to have a compression ratio higher than one ($\mathrm{cr}_0 = 1$, $m = 1440$, $n = 443$). Therefore, this method might be a good choice for other applications of SVD rank reduction like data denoising. Similar to the previous case study, the proposed method maximizes the compression ratio within the feasible region of the problem. Fig. 5(a) and (b) represents the optimal points of the proposed method with $\alpha = 0.0001$ and $\alpha = 0.001$, respectively. Since the matrix is too big to show, Table III provides a comparison of four elements of the data (small to large) in compression levels obtained by various methods. The proposed answers by the first method of Guttman–Kaiser, de Souza *et al.* [10], and the proposed method are accurate and satisfy the accuracy constraint. However, the compression ratio of the proposed method is higher than the other ones. As long as the accuracy constraint is satisfied, the higher compression ratio is preferred. Accordingly, the proposed method dominates the other methods in this case study, as well. It should be noted that the accuracy constraint can be set according to the applications. So, in the case with higher accuracy requirements, the upper bound in (17) can be set to a new threshold. Also, Table III shows that the knee point does not provide an accurate compression level in this case. However, it may reveal a good performance

TABLE III
COMPARISON OF RANK REDUCTION METHODS ON ARBITRARY ELEMENTS OF ORIGINAL DATA

Elements (X,Y)	(1,1)	(1,8)	(1,15)	(1,29)
Original data	3.06	0	58.97	3.87
Knee point	9.37	5.78e-17	49.69	6.37
Guttman-Kaiser (1^{th})	3.06	1.3e-13	58.97	3.87
Guttman-Kaiser (2^{th})	3.82	1.4e-29	24.47	2.95
de Souza et al. [10]	3.06	1.6e-13	58.97	3.87
Proposed method	3.07	9.5e-14	58.97	3.89

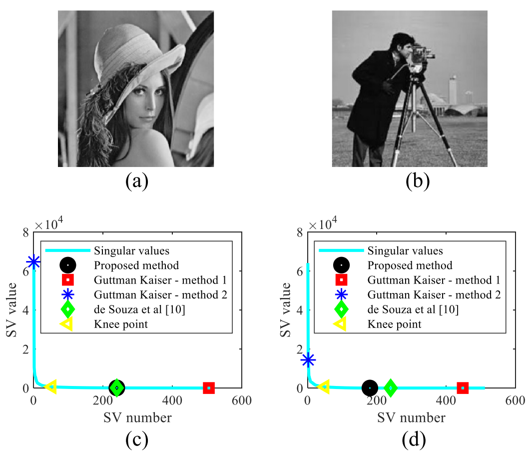

Fig. 6. Image processing benchmarks. (a) Lena image. (b) Cameraman image. The SV diagram and performance of the algorithms on (c) Lena image and (d) Cameraman image.

Fig. 7. Compressed images. (a) Lena and (b) Cameraman with the proposed method. (c) Lena and (d) Cameraman with the knee point.

in image compression applications. It is investigated in the next case on two images shown in Fig. 6(a) and (b).

Case 4. The performance of various algorithms, as well as the proposed method on each image, can be seen in the subplot below it. The quality of the obtained points with various methods can be analyzed similarly to the previous cases. Here, we are going to compare the performance of the proposed method in comparison to the knee point that is a popular approach in image compression. Indeed, we consider the knee as a point where the size of singular values increases rapidly. Fig. 7(a) and (b) shows the compressed images based on the proposed method. Also, the compressed images based on knee point detection are shown in Fig. 7(c) and (d). As can be seen, both methods keep the main information of the images, as they are clearly visible. However, the quality of the compressed images by the proposed method is higher. After comparing the performance of the optimal rank reduction with other methods in Section IV, more discussion and conclusion of this article is presented in the next section.

V. CONCLUSION

This article proposed an optimization formulation for the rank reduction in SVD-based data compression. The SVD-based data compression is basically a tradeoff between the data accuracy and the compression level and like any tradeoff problem, decision making plays a crucial role. In this situation, the quality of the decision can affect the efficiency and the problem. Various methods decide for the number of the singular values in rank reduction problems and some of them reveal good performance in data compression. Here, some points must be considered. Along with the data compression, there are a lot of applications for matrix rank reduction like image reconstruction [24], shot boundary detection in videos [25], and signal denoising [26]. Indeed, these methods are general rank reduction methods and might show acceptable performance in some applications like noise reduction. But, the objective function of the proposed method is specially designed for data compression. That is why the proposed method provides better performance in the case studies. However, in the proposed method, the objective function can be defined according to the specific application. For instance, the compression ratio [or the equivalent function in (16)] is considered for the objective function of the data compression problem as it is discussed in this article. An important feature of the proposed method is that the objective function can easily be changed to find the optimal rank of the matrix for other applications. So, the same implemented code for the data compression can be used for other applications. In this case, the main structure of the code remains unchanged. Replacing the objective function is the main change that is required. Along with this advantage of the optimized rank reduction, it facilitates the decision making in problems with more than one objective function by changing it to a multiobjective optimization problem. This can be complicated in other methods. Besides, the results of the case studies show that the optimal points dominate the points given by the other methods. It is because the optimization problem searches for the optimal point in the feasible region. While the other methods, for example de Souza *et al.* [10], find a good enough point, not the optimal one. As another example, the first method of Guttman–Kaiser does not compress the data and might show a good performance in the other applications of the rank reduction.

The computational complexity of the algorithm is another important issue that needs to be investigated. Generally, the computational complexity of the heuristic algorithms depends on the population, the number of the decision variables, and the number of iterations. Also, the population size is determined based on the number of decision variables. If we consider the proposed method, it contains one decision variable, regardless of the matrix size. Therefore, the computational complexity of the heuristic methods for solving the optimal rank reduction problem is proportional to the number of iterations. On the other hand, the computation complexity of the SVD decomposition implemented in MATLAB is in the order of $O(\max(m,n)2)$. So, the complexity of the SVD determines the complexity of the whole problem. It is because the algorithm is applied to the output of the SVD decomposition and the decomposition needs to be run before the optimization algorithm starts. Also, the other SVD-based methods at least have the complexity of the SVD decomposition. So, in terms of the computational complexity, the proposed method is in the same situation as the other SVD rank reduction algorithms are in.

The difference in the application of the data compression makes a huge difference in approaches to doing it [44]. For instance, the reconstruction error must be kept very small in power system applications; because the error can change the schedule of the grid and impose a higher cost to it. The situation can be a bit different in image compression. The accuracy constraint can be less strict in some image processing applications, as long as the compressed image contains the main features of the original image. That is why the knee point can play a role in numerous applications of image processing. As it was seen in the results, the knee point contains very important features of the images and it is good enough in some applications. But, the same point in the market data did impose a high error to the compressed data, as is shown in Table III. In a conclusion, an optimization method that finds the optimum point regarding the accuracy constraint is required in data compression, especially for the applications that are sensitive to the error like energy market or smart grid data.

To sum up, the presented method can individually solve the issues appearing by the big volume of data such as required bandwidth and data storage by reducing the number of data elements. The proposed framework achieves the higher compression ratio as well as the satisfaction of accuracy constraint simultaneously and the redundant section is cut down. It is simple and efficient and could be utilized by market operators [45], load aggregators [46]–[50], electricity retailers [51]–[54], and Combined Heat and Power(CHP) and microgrid operation [55], [56].

Answering the question "What is the value of the optimal data compression on the grid applications like state estimation?" can be an interesting area for future research. Along with this question, providing the optimization frameworks for the other data compression methods like DCT or DWT is another area that the authors are going to investigate in the future.

References

[1] K. M. Aishwarya, R. Ramesh, P. M. Sobarad, and V. Singh, "Lossy image compression using SVD coding algorithm," in *Proc. IEEE Int. Conf. Wireless Commun., Signal Process. Netw.*, 2016, pp. 1384–1389.

[2] K. Sayood, *Introduction to Data Compression*, 4th ed. San Mateo, CA, USA: Morgan Kaufmann, 2012.

[3] J. Khan, S. H. Bhuiyan, G. Murphy, and J. Williams, "Data denoising and compression for smart grid communication," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 2, no. 2, pp. 200–214, Jun. 2016.

[4] J. Ning, J. Wang, W. Gao, and C. Liu, "A wavelet-based data compression technique for smart grid," *IEEE Trans. Smart Grid.*, vol. 2, no. 1, pp. 212–219, Mar. 2011.

[5] J. Khan, S. M. A. Bhuiyan, G. Murphy, and M. Arline, "Embedded-Zerotree-Wavelet-Based data denoising and compression for smart grid," *IEEE Trans. Ind. Appl.*, vol. 51, no. 5, pp. 4190–4200, Oct. 2015.

[6] J. Cormane and F. Nascimento, "Spectral shape estimation in data compression for smart grid monitoring," *IEEE Trans. Smart Grid*, vol. 7, no. 3, pp. 1214–1221, May 2016.

[7] P. H. Gadde, M. Biswal, S. Brahma, and H. Cao, "Efficient compression of PMU data in WAMS," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2406–2413, Sep. 2016.

[8] V. Loia, S. Tomasiello, and A. Vaccaro, "Fuzzy transform based compression of electric signal waveforms for smart grid," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 47, no. 1, pp. 121–132, Jan. 2017.

[9] Y. Sun, C. Cui, J. Lu, and Q. Wang, "Data compression and reconstruction of smart grid customers based on compressed sensing theory," *Int. J. Elect. Power Energy Syst.*, vol. 83, pp. 21–25, 2016.

[10] J. C. S. de Souza, T. M. L. Assis., and B. Pal, "Data compression in smart distribution systems via singular value decomposition," *IEEE Trans. Smart Grid*, vol. 8, no. 1, pp. 275–284, Jan. 2017.

[11] A. M. Rufai, G. H. Anbarjafari, and H. Demirel, "Lossy image compression using singular value decomposition and wavelet difference reduction," *Digit. Signal Process.*, vol. 24, pp. 117–123, 2014.

[12] S. Padhy, L. N. Sharma, and S. Dandapat, "Multilead ECG data compression using SVD in multiresolution domain," *Biomed. Signal Process. Control*, vol. 23, pp. 10–18, 2016.

[13] M. M. Dixit and C. Vijaya, "Effects of hybrid SVD–DCT based image compression scheme using variable rank matrix and modified vector quantization," in *Innovations in Computer Science and Engineering (Lecture Notes in Networks and Systems)*, vol. 32, H. Saini, R. Sayal, A. Govardhan, and R. Buyya, Eds. Singapore: Springer, 2019.

[14] C. Yu, H. Li, and X. Wang, "SVD-based image compression, encryption, and identity authentication algorithm on cloud," *IET Image Process.*, vol. 13, no. 12, pp. 2224–2232, 2019.

[15] S. K. Mukhopadhyay, M. O. Ahmad, and M. N. S. Swamy, "SVD and ASCII character encoding-based compression of multiple biosignals for remote healthcare systems," *IEEE Trans. Biomed. Circuits Syst.*, vol. 12, no. 1, pp. 137–150, Feb. 2018.

[16] F. Barrosa, W. Fonsecab, U. Bezerraa, and M. Nunesaa, "Compression of electrical power signals from waveform records using genetic algorithm and artificial neural network," *Elect. Power Syst. Res.*, vol. 142, pp. 207–214, 2017.

[17] A. Unterweger and D. Engel, "Resumable load data compression in smart grids," *IEEE Trans. Smart Grid*, vol. 6, no. 2, pp. 919–929, Mar. 2015.

[18] X. Huang, T. Hu, C. Ye, G. Xu, X. Wang, and L. Chen, "Electric load data compression and classification based on deep stacked auto-encoders," *Energies*, vol. 12, no. 4, pp. 653. Feb. 2019.

[19] J. Kraus, P. Stephan, and L. Kukacka, "Optimal data compression techniques for smart grid and power quality trend data," in *Proc. 15th IEEE Int. Conf. Harmon. Qual. Power*, 2012, pp. 707–712.

[20] M. P. Tcheou *et al.*, "The compression of electric signal waveforms for smart grids: State of the art and future trends," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 291–302, Jan. 2014.

[21] J. Li, Y. Li, X. Wang, and P. Zhang, "De-noising method research on RF signal by combining wavelet transform and SVD," in *Proc. 28th Conf. Spacecr. TT&C Technol. China*, R. Shen and G. Dong, Eds., 2018, vol. 445, pp. 479–486.

[22] B. Liu and Q. Liu, "Random noise reduction using SVD in the frequency domain," *J. Petrol. Explor. Prod. Technol.*, vol. 10, pp. 3081–3089, 2020.

[23] N. Khan and K. V. Arya, "Two stage image de-noising by SVD on large scale heterogeneous anisotropic diffused image data," *Multimedia Tools Appl.*, vol. 77, pp. 22543–22566, 2018.

[24] K. Alilou and Y. Yaghmaee, "Exemplar-based image inpainting using SVD-based approximation matrix and multi-scale analysis," *Multimedia Tools Appl.*, vol. 76, pp. 7213–7234, 2017.

[25] B. Youssef, E. Fedwa, A. Driss, and S. Ahmed, "Shot boundary detection via adaptive low rank and SVD-updating," *Comput. Vis. Image Understanding*, vol. 161, pp. 20–28, 2017.

[26] T. Schanze, "Compression and noise reduction of biomedical signals by singular value decomposition," *IFAC-PapersOnLine*, vol. 51, no. 2, pp. 361–366, 2018.
[27] D. A. Jackson, "Stopping rules in principal components analysis: A comparison of heuristical and statistical approaches," *Ecology*, vol. 74, pp. 2204–2214, 1993.
[28] M. Bivalkar, D. Singh, and H. Kobayashi, "Entropy-Based low-rank approximation for contrast dielectric target detection with through wall imaging system," *Electronics*, vol. 8, 2019, Art. no. 634.
[29] D. C. Lay, J. Goldstein, and D. Schneider, *Linear Algebra and Its Applications*, 4th ed. Richardson, TX, USA: Univ. Texas Dallas, 1993.
[30] S. Zaferanlouei, M. Korpås, J. Aghaei, H. Farahmand, and N. Hashemipour, "Computational efficiency assessment of multi-period AC optimal power flow including energy storage systems," in *Proc. IEEE Int. Conf. Smart Energy Syst. Technol.*, 2018, pp. 1–6.
[31] N. Hashemipour *et al.* "A linear multi-objective operation model for smart distribution systems coordinating tap-changers, photovoltaics and battery energy storage," in *Proc. IEEE Power Syst. Comput. Conf.*, 2018, pp. 1–7.
[32] S. M. Islam, S. Das, S. Ghosh, S. Roy, and P. N. Suganthan, "An adaptive differential evolution algorithm with novel mutation and crossover strategies for global numerical optimization," *IEEE Trans. Syst., Man Cybern.*, vol. 42, no. 2, pp. 482–500, Apr. 2012.
[33] Y. L. Li, Z. H. Zhan, Y. J. Gong, W. N. Chen, J. Zhang, and Y. Li, "Differential evolution with an evolution path: A DEEP evolutionary algorithm," *IEEE Trans. Cybern.*, vol. 45, no. 9, pp. 1798–1810, Sep. 2015.
[34] "Day-ahead and real-time energy markets," vol. 3, May 2018. [Online]. Available: https://www.iso-ne.com/markets operations/markets/dart-energy-markets
[35] A. Lüth, J. M. Zepter, P. Crespo del Granado, and R. Egging, "Local electricity market designs for peer-to-peer trading: The role of battery flexibility," *Appl. Energy*, vol. 229, pp. 1233–1243, 2018.
[36] J. M. Zepter, A. Lüth, P. Crespo del Granado, and R. Egging, "Prosumer integration in wholesale electricity markets: Synergies of peer-to-peer trade and residential storage," *Energy Buildings*, vol. 184, pp. 163–1176, 2019.
[37] "Smart* Data Set for Sustainability," May–Aug. 2018. [Online]. Available: http://traces.cs.umass.edu/index.php/Smart/Smart
[38] L. Salimi, A. Haghighi, and A. Fathi, "A novel watermarking method based on differential evolutionary algorithm and wavelet transform," *Multimedia Tools Appl.*, vol. 79, pp. 11357–11374, 2020.
[39] A. K. Qin, V. L. Huang, and P. N. Suganthan, "Differential evolution algorithm with strategy adaptation for global numerical optimization," *IEEE Trans. Evol. Comput.*, vol. 13, no. 2, pp. 398–417, Apr. 2009.
[40] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 1983.
[41] R. V. Rao, V. J. Savsani, and D. P. Vakharia, "Teaching–learning-based optimization: An optimization method for continuous non-linear large scale problems," *Inf. Sci.*, vol. 183, no. 1, pp. 1–15, 2012.
[42] M. Clerc and J. Kennedy," The particle swarm-explosion, stability and convergence in a multidimensional complex space," *IEEE Trans. Evol. Comput.*, vol. 6, no. 1,pp. 58–73, Feb. 2002.
[43] M. Srinivas and L. M. Patnaik, "Genetic algorithms: A survey," *IEEE Comput.*, vol. 27, no. 6, pp. 17–26, Jun. 1994.
[44] S. N. Hashemipour *et al.* "Big data compression in smart grids via optimal singular value decomposition," in *Proc. 2020 IEEE Ind. Appl. Soc. Annu. Meeting*, 2020, pp. 1–8.
[45] H. Xu, Y. Lin, X. Zhang, and F. Wang, "Power system parameter attack for financial profits in electricity markets," *IEEE Trans. Smart Grid*, vol. 11, no. 4, pp. 3438–3446, Jul. 2020.
[46] F. Wang *et al.* "Smart households' aggregated capacity forecasting for load aggregators under incentive-based demand response programs," *IEEE Trans. Ind. Appl.*, vol. 56, no. 2, pp. 1086–1097, Mar./Apr. 2020.
[47] X. Lu, K. Li, H. Xu, F. Wang, Z. Zhou, and Y. Zhang, "Fundamentals and business model for resource aggregator of demand response in electricity markets," *Energy*, vol. 204, May 2020, Art. no. 117885.
[48] K. Li *et al.* "Capacity and output power estimation approach of individual behind-the-meter distributed photovoltaic system for demand response baseline estimation," *Appl. Energy*, vol. 253, Nov. 2019, Art. no. 113595.
[49] W. Huang, N. Zhang, C. Kang, M. Li, and M. Huo, "From demand response to integrated demand response: Review and prospect of research and application," *Protection Control Mod. Power Syst.*, vol. 4, 2019, Art. no. 12. [Online]. Available: https://doi.org/10.1186/s41601-019-0126-4
[50] F. Wang *et al.* "Association rule mining based quantitative analysis approach of household characteristics impacts on residential electricity consumption patterns," *Energy Convers. Manage.*, vol. 171, pp. 839–854, Sep. 2018.
[51] S. Yan *et al.* "Time-frequency feature combination based household characteristic identification approach using smart meter data," *IEEE Trans. Ind. Appl.*, vol. 56, no. 3, pp. 2251–2262, May/Jun. 2020.
[52] K. Li *et al.* "Impact factors analysis on the probability characterized effects of time of use demand response tariffs using association rule mining method," *Energy Convers. Manage.*, vol. 197, no. August, Oct. 2019, Art. no. 111891.
[53] F. Wang *et al.* "Impact analysis of customized feedback interventions on residential electricity load consumption behavior for demand response," *Energies*, vol. 11, no. 4, Apr. 2018, Art. no. 770.
[54] F. Wang, X. Ge, K. Li, Z. Mi, P. Siano, and N. Duic´, "Day-ahead optimal bidding and scheduling strategies for DER aggregator considering responsive uncertainty under Real-time pricing," *Energy*, vol. 213, 2020, Art. no. 118765.
[55] V. Murty and A. Kumar, "Multi-objective energy management in microgrids with hybrid energy sources and battery energy storage systems," *Protection Control Mod. Power Syst.*, 2020.
[56] F. Wang, L. Zhou, H. Ren, and X. Liu, "Search improvement process-chaotic optimization-particle swarm optimization-elite retention strategy and improved combined cooling-heating-power strategy based two-time scale multi-objective optimization model for stand-alone microgrid operation," *Energies*, vol. 10, no. 12, Dec. 2017, Art. no. 1936.