Sparse Signal Reconstruction on Fixed and Adaptive Supervised Dictionary Learning for Transient Stability Assessment

Raoult Teukam Dabou [1] , **Innocent Kamwa [1],*** , **Jacques Tagoudjeu [2]** and **Francis Chuma Mugombozi [3]**

[1] Electrical and Computer Science Engineering Department, Laval University, Quebec, QC G1V 0A6, Canada; raoult.teukam-dabou.1@ulaval.ca

[2] Department of Mathematics and Physical Science, National Advanced School of Engineering of Yaoundé, University of Yaoundé I, Yaoundé P.O. Box 8390, Cameroon; jacques.tagoudjeu@univ-yaounde1.cm

[3] Department of Power Systems Simulation and Evolution, Research Institute of Hydro Québec/IREQ, Varennes, QC J3X 1S1, Canada; mugombozi.chumafrancis@hydroquebec.com

* Correspondence: innocent.kamwa.1@ulaval.ca

Abstract: Fixed and adaptive supervised dictionary learning (SDL) is proposed in this paper for wide-area stability assessment. Single and hybrid fixed structures are developed based on impulse dictionary (ID), discrete Haar transform (DHT), discrete cosine transform (DCT), discrete sine transform (DST), and discrete wavelet transform (DWT) for sparse features extraction and online transient stability prediction. The fixed structures performance is compared with that obtained from transient K-singular value decomposition (TK-SVD) implemented while adding a stability status term to the optimization problem. Stable and unstable dictionary learning are designed based on datasets recorded by simulating thousands of contingencies with varying faults, load, and generator switching on the IEEE 68-bus test system. This separate supervised learning of stable and unstable scenarios allows determining root mean square error (RMSE), useful for online stability status assessment of new scenarios. With respect to the RMSE performance metric in signal reconstruction-based stability prediction, the present analysis demonstrates that [DWT], [DHT | DWT] and [DST | DHT | DCT] are better stability descriptors compared to K-SVD, [DHT], [DCT], [DCT | DWT], [DHT | DCT], [ID | DCT | DST], and [DWT | DHT | DCT] on test datasets. However, the K-SVD approach is faster to execute in both off-line training and real-time playback while yielding satisfactory accuracy in transient stability prediction (i.e., 7.5-cycles decision window after fault-clearing).

Keywords: sparse signal decomposition; supervised dictionary learning; dictionary impulsion; singular value decomposition; discrete cosine transform; discrete Haar transform; discrete wavelet transform; transient stability assessment

1. Introduction

Dynamic state estimation (DSE) is a fast-developing tool for stability monitoring and control [1]. According to the IEEE TF on DSE [2], enhanced system observability using DSE based internal angle and speed estimates will lead to several breakthroughs in Wide-Area System Integrity Protection Systems (WA-SIPS) functions, such as faster out-of-step detection, more realistic location of runaway generator and minimal amount of generation/load to be shed in order to preserve system integrity without knowing the topology accurately. However, instrumental in these achievements is the ability to predict fault-induced instability from DSE based information in real-time, reliably, securely, and fast enough to enable timely and effective countermeasures [3]. In the context of the integration of renewable energy, development of smart grids and artificial intelligence, the authors in [4] proposed a literature review of power system restoration to analyze dynamic decision-making. More generally, application of AI techniques including supervised

machine learning in the context of power systems event detection and diagnosis is a field of growing interest [5].

1.1. Background on Sparse Dictionaries Learning

Sparse signals are characterized by a few non-zero coefficients in one of their transformation domains [6]. Therefore, sparse signal decomposition consists in modelling data vectors as sparse linear combinations of basis elements. The basis set used in the decomposition is also called a dictionary. Sparse dictionary learning has attracted widespread scientific attention in signal processing and power system areas. Sparse signal decomposition on hybrid dictionaries is used in [7] for detecting and classifying single and multi-power quality disturbances. In [8], discrete wavelet transform (DWT) and deep neural networks (DNN) are developed for detecting series arc faults (with accuracy recognition rate of 97.75%) in low-voltage residential power distribution. Alternatively, several approaches are proposed in [9] for sparse signal decomposition based on DCT+K-SVD, DCT+ method of optimal directions (MOD), DWT+K-SVD, and DWT+MOD. In [10], new automatic transient detection and localization approaches are proposed for power quality analysis based on an over-complete dictionary (OCD) and $\ell - 1$ norm minimization. Dictionary learning sparse decomposition is implemented in [11] for accurate and fast classification of power quality (PQs) disturbances generated from the IEEE Power Energy Society database. Damped-sinusoidal functions integrated in a redundant dictionary are used in [12] to detect and characterize power systems oscillatory transients in the portion of the time-domain signal based on matching pursuit. A Stockwell transform (ST) approach is proposed in [13] to extract significant disturbance features, used as the input dataset of K-nearest neighbor algorithms (K-NN), decision trees (DT) and support vector machines (SVM). A compressive-information sparse representation based on sparse recovery with a new high-dimensional convex hull is developed in [14] for fast and reliable PQs event classification due to the new challenge of smart grids. The sparse representation and fully connected neural networks based on sparse coding, intelligent features learning, and neural networks are developed in [15] for extracting and classifying idiographic signal residential fault. A fault diagnosis approach based on sparse representation and SVM is used in [16] for computing sub-dictionaries and representing the testing sample while in [17], K-SVD, compressive sampling matching pursuit, and Gauss random matrix are used for reconstructing power quality events. For a signal-to noise ratio of 30 dB, a K-SVD-CoSaMP algorithm reconstructs PQs better than the DCT algorithm.

According to [18], linear sensitivity distribution factor is performed based on injection shift factor of the line with respect to active power injection at all buses. The phasor data angle and DC linear power flow reformulation are used in [19] for identifying multiple line outages based on solving a sparse representation problem via coordinate descent iteration. The cluster-based sparse coding (CSC) algorithm is proposed in [20] for multiple events detection, recognition, and temporal localization in large-scale power systems while in [21], sparse representation classification and random dimensionality reduction projection are used to extract, reduce feature dimensionality and classify power system faults. The singular value decomposition (SVD) and total least square estimation via rotational invariance techniques are used in [22], to analyze and extract amplitude, frequency, attenuation coefficient, initial phase of combined PQs. Among all dictionaries-based learning schemes, which built on K-SVD and OMP algorithms, used e.g., in [23,24] for short-term prediction of solar power fluctuations, is often considered as the most powerful computationally and accuracy wise. In [25], the features extracted from active power are incorporated into a dictionary learning based defense scheme for understanding the cyberattack process in smart grids. As alternatives to dictionary learning, recurrent neural networks (RNN), long short-term memory (LSTM), gated recurrent unit (GRU) and convolutional neural network-LSTM (CNN-LSTM) have also been proposed for time series forecasting of solar irradiance and photovoltaic power [26]. In the above work, the LSTM achieved the best performance in terms of the root-mean square error evaluation error (RMSE). As a further application to

time series, [27] used a nonlinear dictionary learning to decompose electricity signals and learn the most representative temporal features introduced by operating devices.

1.2. Contributions of the Present Paper

The previous literature review established clearly that dictionary-based learning applications in power systems have been focused hitherto on PQs analysis and time-series forecasting and characterization. A timely contribution of this paper is therefore to extend dictionary-based learning to transient stability assessment. Even though machine learning has been studied extensively on this problem [28], there is no reported application of K-SVD + OMP dictionary-based learning of post-disturbance time-responses for fast stability prediction in the context of WA-SIPS or contingency screening in control rooms. Furthermore, the paper investigates performance comparisons of single, hybrid, and adaptive dictionaries for wide-area stability assessment which has not yet been addressed in power system literature. In order to determine which dictionary learning approach is the best, according to metrics such as accuracy, reliability, and security, in addition to on-line and off-line computational requirements, this paper proposes to design twenty stable and unstable supervised dictionary learning schemes for sparse reconstruction and classification of stability signal responses based on generator time-series recorded by PMU in post-disturbance stage. Our findings highlight the reconstruction and classification performance of single, hybrid (with two and three sub-dictionaries) and learned dictionaries on the test dataset (rotor speed and corresponding stability status).

The rest of this paper is organized as follows. Section 2 presents design approaches of fixed and adaptive overcomplete dictionary learning for sparse signal decomposition. Section 3 presents performance measures of supervised dictionary learning. Section 4 presents a top-down approach of sparse signal decomposition on fixed and adaptive dictionary learning. Section 5 presents the experimental results of ten supervised dictionary learning algorithms for TSA. Finally, some conclusions and perspectives are given in Section 6.

2. Design of Fixed and Adaptive Supervised Dictionary Learning

Given $Y_{rs} = [y_1, \ldots, y_N] \in \mathbb{R}^{n \times N}$ the post-contingency rotor speed, which contains N input signals of n dimension. The sparse decomposition of Y_{rs} on dictionaries $D_{rs} \in \mathbb{R}^{n \times K}$ ($K \succ n$, over-complete dictionary) is computed by:

$$Y_{rs} = \sum_{i=1}^{K} D_{irs} X_{irs} + \varepsilon_{irs} \tag{1}$$

where: X_{rs} and ε_{irs} denote sparse coding and reconstruction errors of Y_{rs} sparse signal decomposition. The orthogonal matching pursuit (OMP) is the greedy method for sparse coding mostly used in supervised dictionary learning, for minimizing the response reconstruction over a finite number of iterations [29]. These dictionaries are used to evaluate the transient stability which is the ability of the power system to remain in synchronism immediately following a disturbance [30]. The architecture of singles, hybrids, and adaptive supervised dictionary learning for reconstructing and classifying the signals and rapidly predicting the transient stability status immediately after clearing a fault is presented in Figure 1.

2.1. Fixed Single Supervised Dictionary Learning

From the PQD and time-series forecasting experiences, fixed dictionaries can be used to reconstruct rotor speeds and predict the transient stability of generator responses. For this purpose, the cosine transform, sine transform, wavelet transform and Haar transform are simply applied to each signal to extract sparse and redundant features for stability of signals reconstruction.
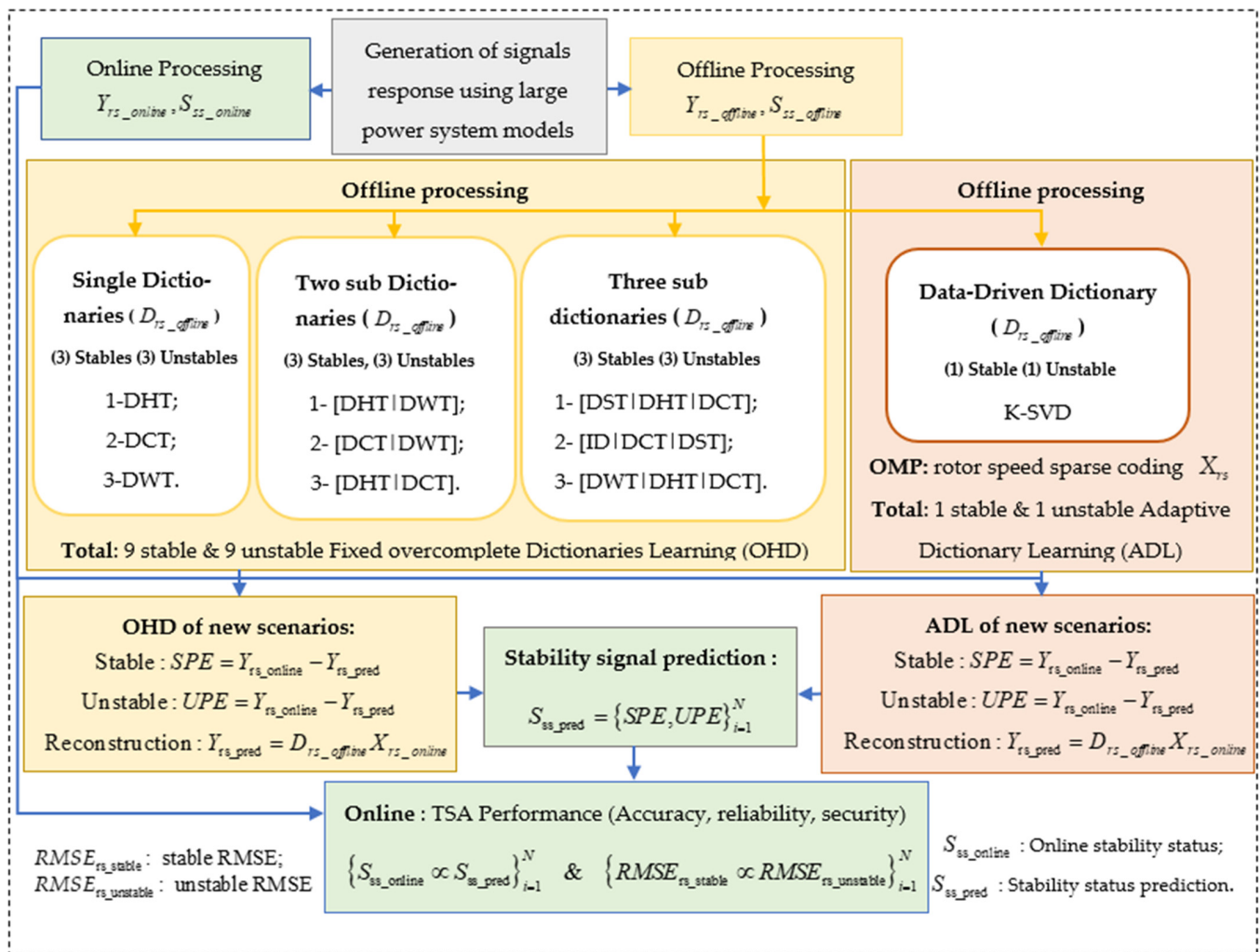
Figure 1. Proposed flowchart of sparse signal decomposition on supervised dictionaries for transient stability assessment.

2.1.1. Discrete Cosine Transform

The discrete cosine transform (DCT) expresses a time-series generator response in terms of a sum of cosine functions oscillating at different frequencies introduced by the fault occurrence. The DCT kernel projection is defined by:

$$\text{DCT} = D_{rs} = \sqrt{\tfrac{2}{n}} \left[\xi_i \cos\left(\tfrac{\pi(2j+1)i}{2n} \right) \right]$$

$$i, j = 0, 1, 2, 3, \ldots, K - 1$$

(2)

where $\xi_i = 1/\sqrt{2}$ for $i = 0$, otherwise $\xi_i = 1$. The DCT proposes optimal de-correlation coefficients of the stability signals, while grouping energy contained in the signals in low frequency coefficients [7].

2.1.2. Discrete Sine Transform

The discrete sine transform (DST) derived from the DCT, allows expressing the signal as a sum of sinusoids with different frequencies and amplitudes. The DST kernel projection is computed by:

$$\text{DST} = D_{rs} = \sqrt{\tfrac{2}{n}} \left[\xi_i \sin\left(\tfrac{\pi(2j+1)(i+1)}{2n} \right) \right]$$
$$i, j = 0, 1, 2, 3, \ldots, K - 1$$

(3)

where $\xi_i = 1/\sqrt{2}$ for $i = n - 1$, otherwise $\xi_i = 1$. The DCT and the DST are almost similar operations to the discrete Fourier transform (DFT). The only difference lies in the projection kernels, which give real coefficients for the DCT, imaginary coefficients for the DST and complex coefficients for the DFT [7].

2.1.3. Discrete Haar Transform

Discrete Haar transform (DHT) consists of orthogonal switched rectangular waveforms, which can take zero value and sample points in subintervals of $t \in [0, 1]$ and where the amplitude can differ from one function to another as follows [31]:

$$\text{DHT} \equiv D_{rs} = \frac{1}{\sqrt{N}} \begin{cases} 2^{m/2}, \frac{k-1}{2^m} \leq t < \frac{k-(1/2)}{2^m} \\ -2^{m/2}, \frac{k-(1/2)}{2^m} \leq t < \frac{k}{2^m} \\ 0, \text{ otherwise in } [0, 1) \end{cases} \tag{4}$$
$$r = 2^m + k - 1, \ t = n/N, \ 0 \leq r, n \leq N - 1$$

where: m and k represent the integer decomposition of the index r. The time-frequency Haar function is unitary and invariant under circulant shift.

2.1.4. Discrete Wavelet Transform

The discrete wavelet transform (DWT) is calculated by convolving the post-contingency signal with a mother wavelet transform $\Psi_{m,n}$ [32].

$$\text{DWT} = D_{rs} = \frac{1}{\sqrt{a_0^m}} \sum_k \Psi_{m,n} \left(\frac{k - nb_0 a_0^m}{a_0^m} \right) \tag{5}$$

where: n and m are positive integers. The DWT allows extracting information content at different positions and scales for subsequently reconstructing post-contingency stability signals.

2.1.5. Impulse Dictionary

The impulse dictionary (ID) allows expressing signal as a linear combination of the Dirac basis vectors representing the frequency response pic [7]. It is constructed as:

$$\text{ID} = D_{rs} = [I]_{K \times K} \tag{6}$$

2.2. Fixed Hybrids Supervised Dictionaries Learning

The fixed hybrid supervised dictionary learning is developed from the concatenation of DCT, DHT, DST, DWT, and ID, with the aim of improving online reconstruction and classification of the single dictionaries.

2.2.1. Two Sub-Dictionaries Concatenation

Three concatenations of two sub-dictionaries are constructed for sparse features extraction and classification. Equations (7)–(9) represent these strings of dictionaries joining single structure end-to end.

$$Y_{rs_{n \times N}} = D_{rs} X_{K \times N} = [DHT|DCT]_{n \times K} X_{K \times N} \tag{7}$$

$$Y_{rs_{n \times N}} = D_{rs} X_{K \times N} = [DWT|DHT]_{n \times K} X_{K \times N} \tag{8}$$

$$Y_{rs_{n \times N}} = D_{rs} X_{K \times N} = [DCT|DWT]_{n \times K} X_{K \times N} \tag{9}$$

2.2.2. Three Sub-Dictionaries Concatenation

To improve the singles and hybrids structures with two sub-dictionaries, three string concatenations are developed, see for instances Equations (10)–(12). The fixed predefined overcomplete dictionary learning enables online TSA.

$$Y_{rs_{n\times N}} = D_{rs}X_{K\times N} = [DST|DHT|DCT]_{n\times K}X_{K\times N} \tag{10}$$

$$Y_{rs_{n\times N}} = D_{rs}X_{K\times N} = [ID|DCT|DST]_{n\times K}X_{K\times N} \tag{11}$$

$$Y_{rs_{n\times N}} = D_{rs}X_{K\times N} = [DWT|DHT|DCT]_{n\times K}X_{K\times N} \tag{12}$$

2.3. Adaptive Supervised Dictionary Learning

The K-SVD is a generalization of a k-means clustering algorithm introduced by [24] for dictionary learning after estimation of the representation matrix. During the first learning step of K-SVD, dictionary $D_{rs} = [d_1 \ldots d_K] \in \mathbb{R}^{n\times K}$ ($K \succ n$, over-complete dictionary) is fixed, and the best sparse coding $X_{rs} = [x_1 \ldots x_N] \in \mathbb{R}^{K\times N}$ is found, under the sparsity constraint T. The K-SVD approach is evaluated based on solving an optimization problem defined in Equation (13).

$$\langle D_{rs}, X_{rs} \rangle = \operatorname*{argmin}_{D_{rs},X_{rs}} \|Y_{rs} - D_{rs}X_{rs}\|_2^2 \text{ subject to } \forall_i \|x_i\|_0 \leq T \tag{13}$$

The OMP is used for sparse coding fixed, hybrids, and adaptive overcomplete dictionary learning (ODL). Given D_{rs}, the sparse coding $X_{rs} \in \mathbb{R}^{K\times N}$, is evaluated using the ℓ_0 norm by:

$$x_{irs} = \operatorname*{argmin}_{x_{irs}} \|Y_{rs} - D_{rs}X_{rs}\|_2^2 \text{ subject to } \forall i, \|x_i\|_0 \leq T \tag{14}$$

where: $D_{rs} \in \mathbb{R}^{n\times K}$ and T denote overcomplete dictionaries and sparsity constraint respectively. The update of each d_i atom necessarily leads to that of the non-zero entries in a row vector $x_i^T \neq 0$. The dictionary is learned by solving to following minimization problem:

$$\min_{d_i, x_i} \|E_i - d_i x_i^T\|_2^2 \text{ subject to } \|d\|_2^2 = 1 \tag{15}$$

where: $E_i = Y_{rs} - \sum_{j\neq i} d_j x_j^T$ denotes the approximation error matrix, which can be easily decomposed into $U\Delta V^T$. The solution d_j is a column of U and the coefficient vector x_j^T is a column of $V \times \Delta(1, 1)$ [33].

3. Performance Measures of Supervised Dictionary Learning

The confusion matrix is used to evaluate the classification performance of the dictionary learning algorithms on the testing database. Three main metrics are calculated from Table 1.

Table 1. Confusion matrix for TSA classification.

Observation		Predicted Class	
		Positive (Secure)	Negative (Insecure)
Actual	Positive	*TP*	*FN*
Class	Negative	*FP*	*TN*

The accuracy metric defines the ability of learned dictionaries for correctly classifying stable and unstable cases in the testing database:

$$Accuracy \; (\%) = \frac{TP + TN}{TP + TN + FN + FP} \tag{16}$$

The reliability metric in Equation (17) allows for evaluating the performance of learned dictionaries on the unstable cases:

$$Reliability \; (\%) = \frac{TN - FN}{TN} \tag{17}$$

The security metric in (18) makes it possible to define the capacity of learned dictionaries to predict the stable case:

$$Security \; (\%) = \frac{TP - FP}{TP} \tag{18}$$

The RMSE defined in (19) is used to evaluate the difference between the predicted $Y_{\text{rs\_pred}}$ and observed $Y_{\text{rs\_online}}$ signal responses:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Y_{i\text{rs\_online}} - Y_{i\text{rs\_pred}}\right)^2} \tag{19}$$

4. Top-Down Approach of Sparse Signal Decomposition on Fixed and Adaptive Dictionaries Learning

The simulation of a hundred faults on the line and generator buses of the IEEE test system allowed generating generator rotor-speed responses and their stability status. Figure 2 describes top-down architecture of the proposed system along with the details of the experiment for an even greater clarity of SLOD.

Seventy percent of signals are used for offline-supervised dictionary learning while 30% are kept aside for online TSA. For an offline learning approach, the training signals are labelled and separated into two databases: stable rotor speeds (Database 1) and unstable rotor speeds (Database 2). Database 1 is used as input data for implementing 9 fixed dictionaries D_{rs} for stable cases reconstruction using: DHT, DCT, DWT, [DHT|DWT], [DWT|DCT], [DHT|DCT], [DST|DHT|DCT], [ID|DCT|DST] [7], and [DWT|DHT|DCT] dictionaries. Similarly, Database 1 is used as input data for developing an adaptive stable dictionary D_{rs} based on K-SVD approach. The second Database 2 is used as input for establishing fixed dictionaries D_{rs} for unstable case reconstruction using: DHT, DCT, DWT, [DHT|DWT], [DWT|DCT], [DHT|DCT], [DST|DHT|DCT], [ID|DCT|DST] [7], and [DWT|DHT|DCT] dictionaries. Similarly, database 2 is used as input data for realizing an adaptive unstable dictionary D_{rs} based on K-SVD approach.

The orthogonal matching pursuit algorithm is developed to carry out sparse coding X_{rs} of rotor speeds from the test database (Database 3). The test database contains both stable and unstable rotor speed signals Y_{rs}. The stable and unstable sparse coding X_{rs} are individually projected onto the 20 learned dictionaries D_{rs} to determine playback degree of stability. According to the projection result, if RMSE index of the signal obtained from a stable dictionary is lower than RMSE index predicted by the unstable dictionary, the test signal is targeted as being stable (stability status = "0"). In the opposite case, the signal is labelled as being unstable (stability status = "1"). This analysis criterion is used for evaluating accuracy, reliability (false positive success rate), and security (true negative success rate) of each supervised dictionary-learning scheme on the testing database.
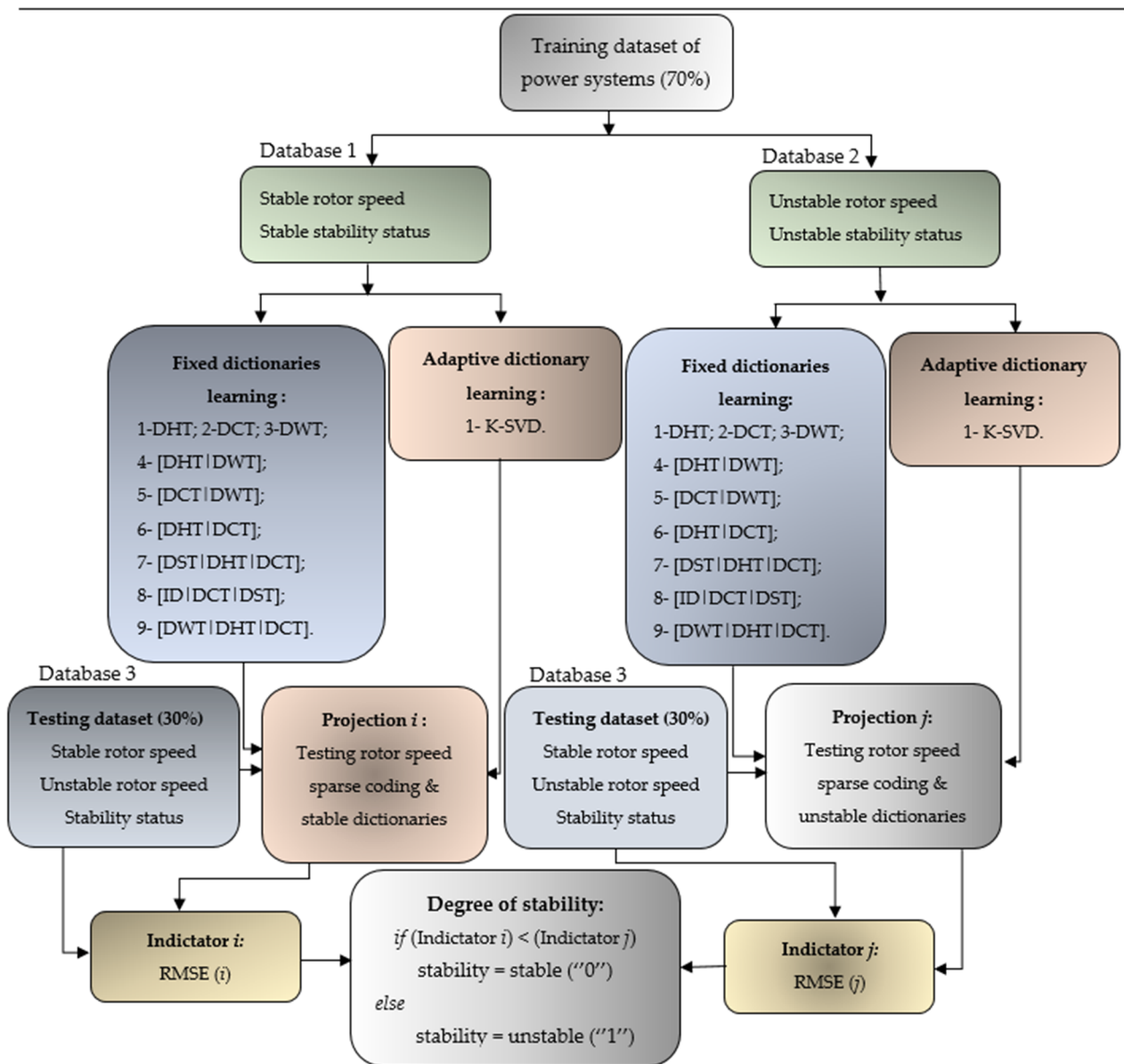
Figure 2. Top-down approaches of supervised dictionary learning for TSA.

5. Experimental Results

The analysis of supervised single ([DCT], [DHT], [DWT]) dictionaries, hybrids with two ([DHT|DCT], [DCT|DWT], [DHT|DWT]) and hybrids with three ([DST|DHT|DCT], [ID|DCT|DST], and [DWT|DHT|DCT]) sub-dictionaries for TSA was performed on the IEEE 68-bus test system. The $75 \times 20,216$ (i.e., 75×8664 allocated for online testing) post-contingency speed signals used to train dictionary learning are generated by simulating various types of fault with wide-range of duration applied on each line and at each generator terminal bus of the test system under various power dispatch and topologies conditions. The generated generator responses used for learning are summarized in stable training set (i.e., $75 \times 19,477$) and unstable training set (i.e., 75×739). The unstable cases thus represent a mere 3.8% of the cases, which means that the dataset is highly skewed toward stable cases. These two dictionaries learned separately will allow classifying a joint database containing stable and unstable scenarios. Each of the signals will be projected on each of the two stable and unstable dictionaries. The learned stable or unstable dictionaries

that give the lowest RMSE will define the stability status of the new scenario as either stable or unstable.

5.1. Description of Dataset for TSA

To generate a diversity of cases, fault duration is gradually increased (per step of 0.5 cycle) until critical loss-of synchronism limit is reached along each line and close to each generator bus of the test system. For each study case, the numerical simulation duration is set to 10 s and transient fault occurs at $t = 1$ s. However, for the unstable scenarios, numerical simulation duration varies according to the time instant of loss-of-synchronous. For each simulation, only 75 post-contingency samples (sampling rate: 600 samples/s) per rotor speed are used for dictionary learning which means a time-window of 125 ms for making decision (7.5 cycles of fundamental). The re-sampling start time of the generator signal responses varies according to the simulation fault clearing time instant.

In the presence of a disturbance, the differences in the absolute values of rotor angles of all the combinations (i, m) of the generators are calculated and compared to $180°$. If the rotor angle does not exceed $180°$, or if the disturbance causing the rotor swing is promptly removed, the generator may remain in synchronism with the system. Hence, during fault simulations, the stability of the power system can take two possible status after fault clearing, "$S_{ss} = 0$" corresponding to the stable status, or "$S_{ss} = 1$" corresponding to the unstable status [34,35]. Figure 3 presents an example of stable (a) and unstable (b) speed recorded during the search of critical fault clearing time on line 1 of the IEEE 68-bus test system.
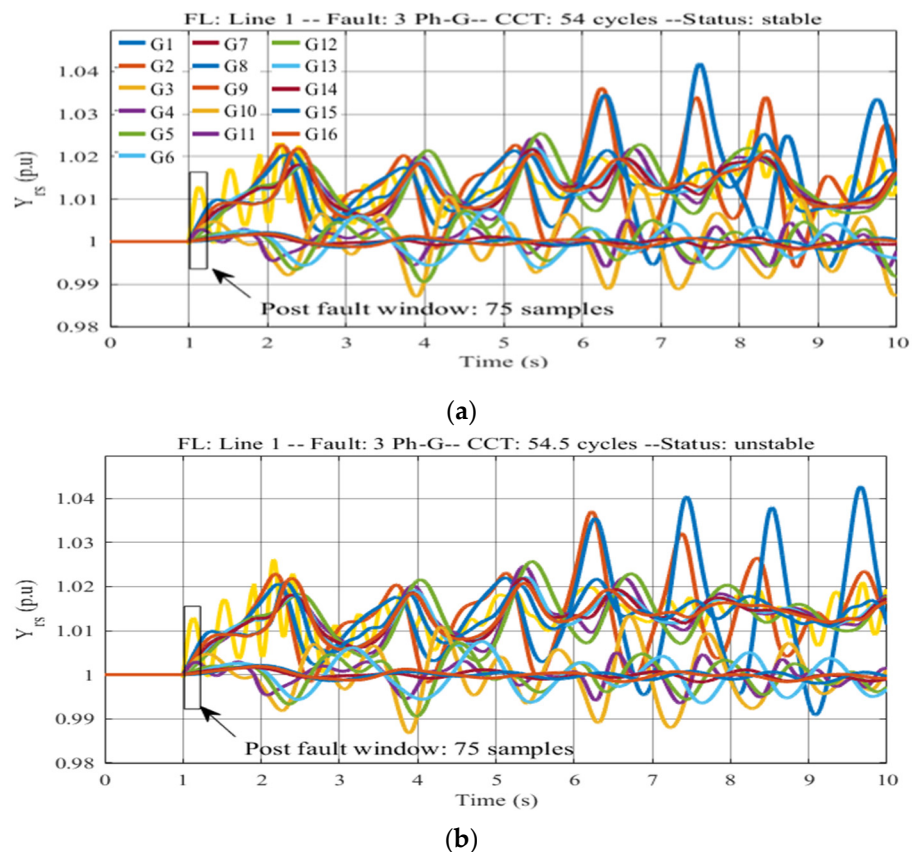

(a)

(b)

Figure 3. Stable (**a**) and unstable (**b**) rotor speed obtained by simulating three-phase ground fault.

Table 2 presents the dataset configuration used for reconstructing and classifying the unknown testing database (containing 8664 signals: 8341 stable scenarios and 323 unstable scenarios i.e., 3.72% of the cases) based on supervised learning of overcomplete dictionaries.

Table 2. Configuration of stable and unstable sparse signal decomposition for TSA.

Sparse Signal	Offline	Online
Decomposition	Features Extraction	Reconstruction and TSA
Adaptive Dictionary Learning (ADL)	Rotor speed	Prediction error
	(For D_{rs} and X_{rs} computed)	(For RMSE computed)
Fixed Dictionaries Learning (1,2,3H OHD)	Rotor speed	Prediction error
	(For D_{rs} and X_{rs} computed)	(For RMSE computed)

5.2. Fixed and Adaptive Sparse Features Extraction

The number of fixed and hybrid features (i.e., 25 indicators) extracted per signal depends on the number of iterations defined by the orthogonal matching pursuit [36]. However, the adaptive dictionary learning allows extraction of 75 features per generator response for online transient stability assessment. Table 3 summarizes the sparse signal decomposition developed off-line and used for online TSA.

Table 3. Parameters used to train single, hybrids and adaptive supervised dictionary learning.

Configuration	Learning Approaches	Dictionaries/ Feature Size	Sparse Expansion Coefficient Size
Adaptive dictionary (ADL)	K-SVD	$75 \times 20,216$	$8700 \times 20,216$
Single dictionary (SOD)	DHT DCT DWT		$75 \times 1,617,280$ $75 \times 1,516,200$ $75 \times 1,556,632$
Two sub-dictionaries (2H-OHD)	[DHT｜DWT] [DCT｜DWT] [DHT｜DCT]	$25 \times 20,216$	$75 \times 3,133,480$ $75 \times 3,153,696$ $75 \times 2,304,624$
Three sub-dictionaries (3H-OHD)	[DST｜DHT｜DCT] [ID｜DCT｜DST] [7] [DWT｜DHT｜DCT]		$75 \times 3,800,608$ $75 \times 4,589,032$ $75 \times 3,820,824$

The stable K-SVD & OHD dictionaries are initialized and trained with stable signals as input, while unstable K-SVD & OHD dictionaries are trained with unstable responses only. For the fixed sparsity (i.e., matrix of numbers that includes many zeros or values that will not be significant) $T = 10$, the coefficients are computed using OMP and the maximum of number iterations in set as 90. K-means is the method used for defining the maximum iterations able to converge the learning algorithm. It allows defining the threshold beyond which no point changes during the sparse signal decomposition. This method supposes that the last iterations have the least contribution to the percent of correct representation [37]. From the online sparse coding extracted using OMP, stable, and unstable K-SVD & OHD dictionary learning will allow reconstructing and predicting transient stability signals. Sample K-SVD Figure 4a and OHD Figure 4b atoms extracted from the learning of rotor speed responses. Figure 4 presents some examples of adaptive K-SVD (a) and fixed OHD (b) atoms randomly extracted in the offline rotor speed Y_{rs}.

Adaptive K-SVD atoms well describe the non-stationary and transient response of generators signals immediately after clearing fault. However, usually sinusoidal fixed OHD atoms (except DHT) are not satisfactory for analysis of the transient behavior of generators. To illustrate the atoms performance extracted from the training database, we selected two signals from the testing database, which are respectively stable and unstable. Figure 5 presents sparse decomposition of testing signals ID 1 (in blue color) & ID 23 (in red color) based on a separate stable and unstable K-SVD dictionary learning. The K-SVD

demonstrates here its ability for an almost perfect fit of the signals in each of the dictionary. This pattern of success is repeated for all signals in the training dataset.
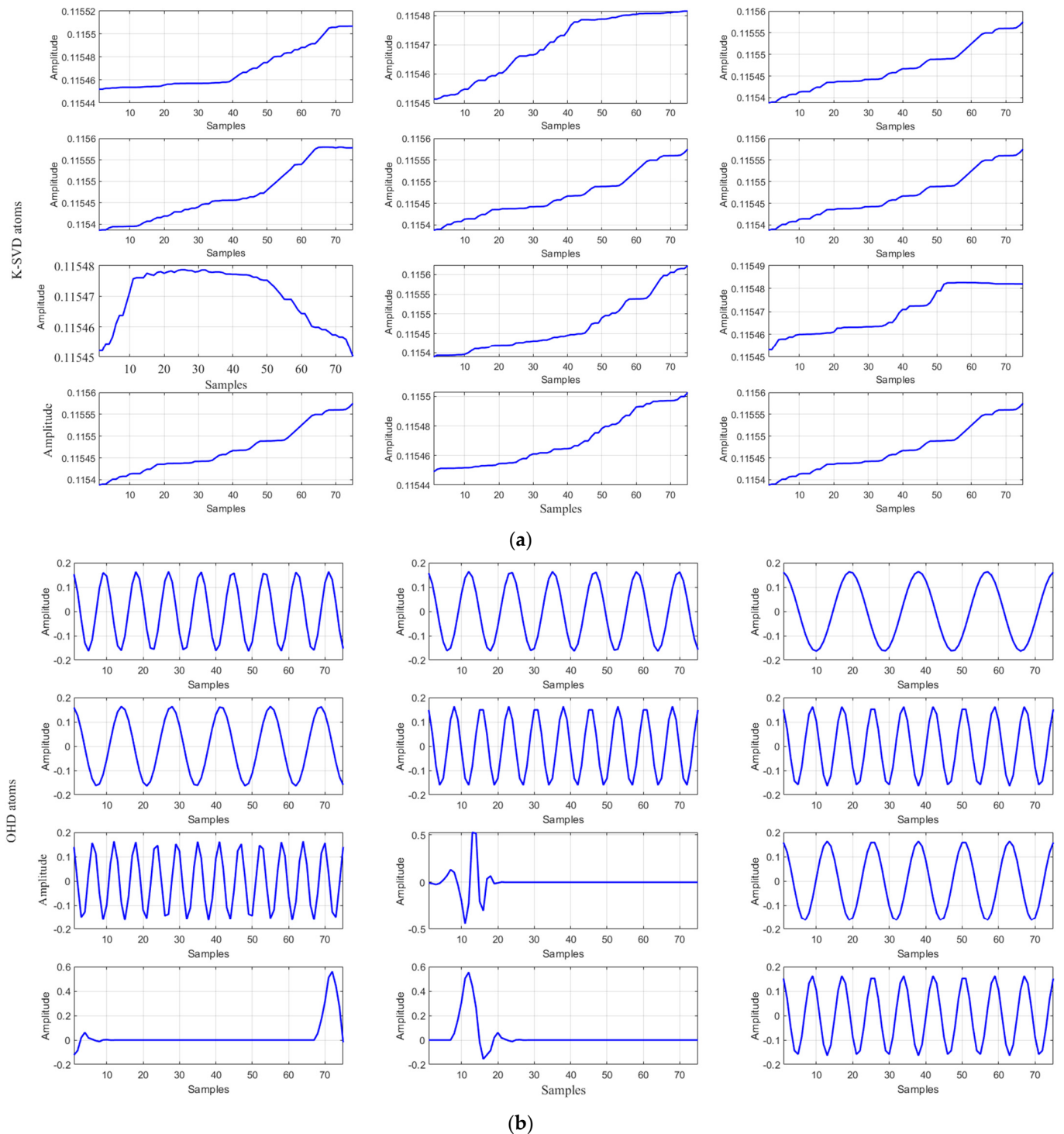

(**a**)

(**b**)

Figure 4. K SVD (**a**) and OHD (**b**) atoms used for TSA reconstruction and classification.

The OHD atoms, such as single (DCT, DWT, DHT) dictionaries, two ([DHT | DCT], [DCT | DWT], [DHT | DWT]) and three ([DST | DHT | DCT], [ID | DCT | DST], [DWT | DHT | DCT]) sub-dictionaries are next used for reconstructing stable (blue color) & unstable (red color) rotor speed and tracking online (black color) response, see Figure 6.
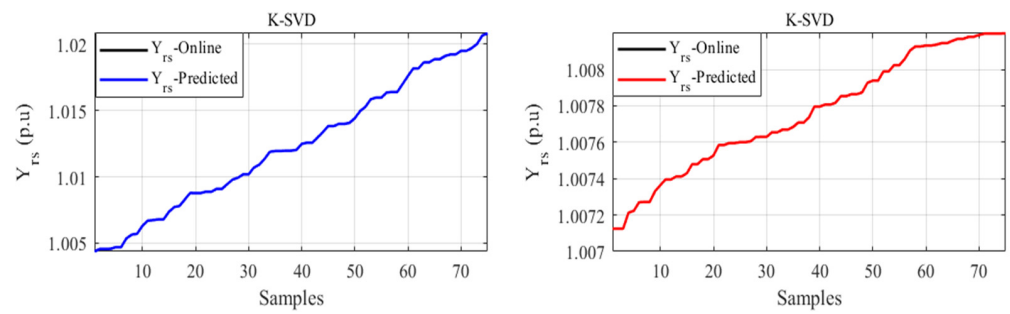
Figure 5. Online & K-SVD reconstruction of stable (blue) and unstable (red) of sample signal's response in the test dataset.
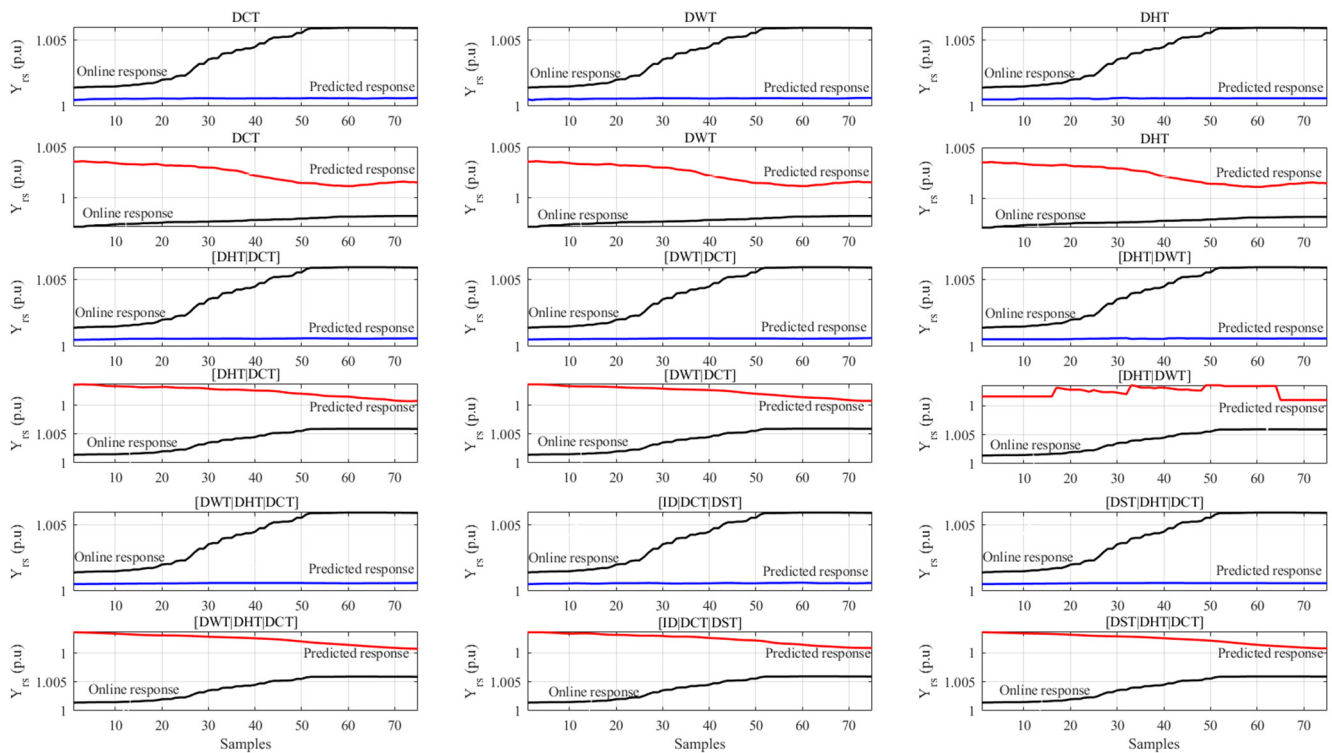

Figure 6. Online & OHD reconstruction of stable (blue) and unstable (red) of sample signal's response in the test dataset.

5.3. Online Sparse Features Classification Based Overcomplete Dictionaries Learning

The projection of signals ID 1 & ID 23, on stable and unstable dictionary learning (SOD, 2H-OHD, 3H-OHD & K-SVD), will allow evaluation of the online stability status according to the atoms structure which gives the lowest RMSE. Figure 7 presents RMSE prediction value obtained by projecting signals ID 1 (blue) and ID 23 (red) on stable and unstable SOD and K-SVD respectively.

According to SOD, the DWT is the best descriptor of stable rotor speed, while for the unstable scenarios prediction DHT is most indicated. The RMSE resulting from hybrids of [DHT | DCT], [DCT | DWT] & [DHT | DWT] are similarly used to perform fixed dictionary learning for online TSA. The projection of signals ID 1 (blue) & ID 23 (red) on two hybrids dictionaries learned separately, allows for a correct prediction of the online stability status based on the lowest RMSE, see Figure 8.
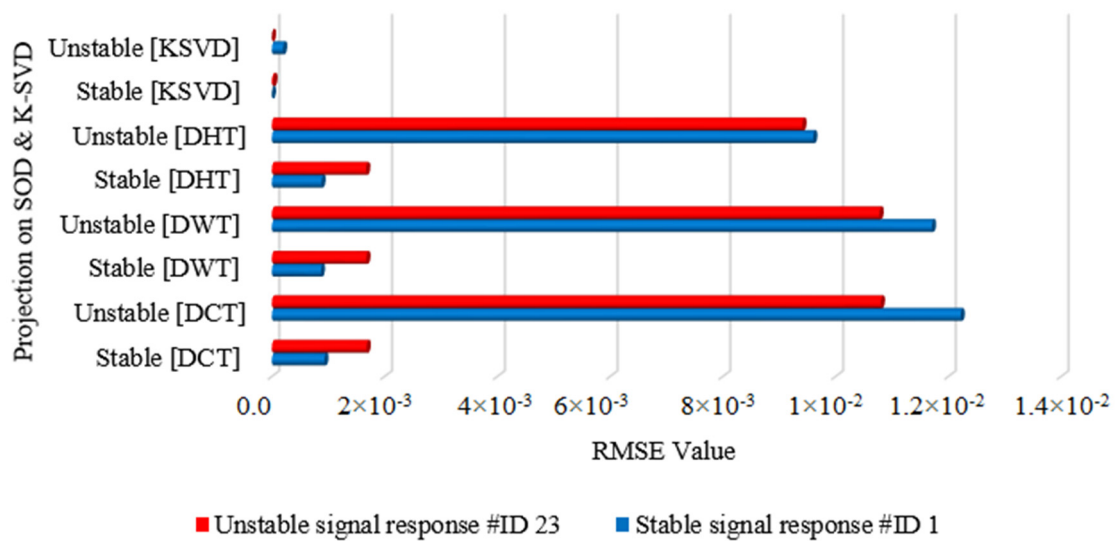
Figure 7. Projections of signals ID 1 and ID 23 on stable and unstable SOD and ADL.
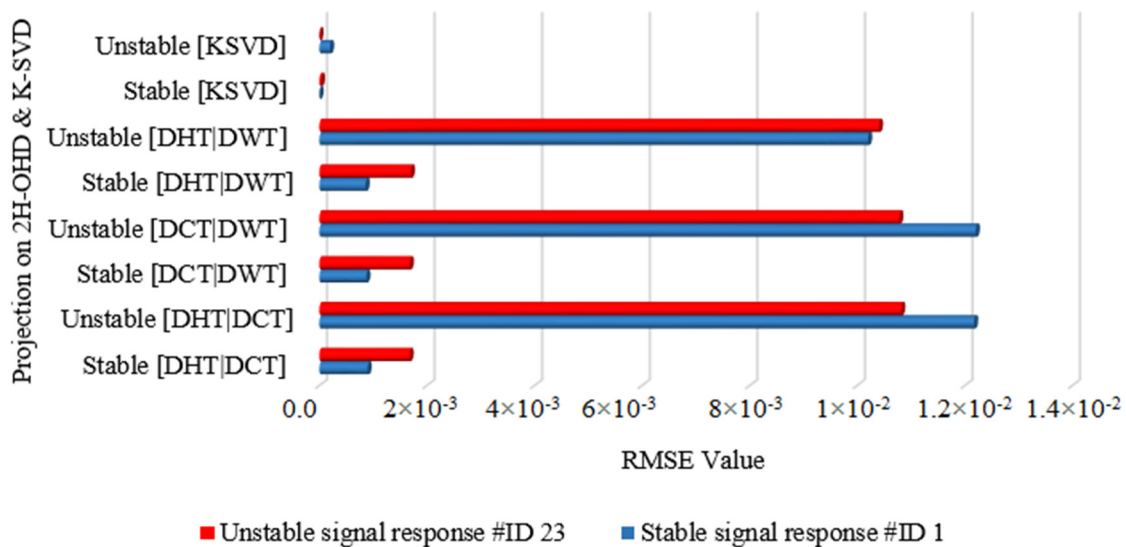

Figure 8. Projections of signals ID 1 and ID 23 on stable and unstable 2H-OHD and ADL.

Regarding 2H-OHD, the [DHT|DWT] has lower RMSE predictors for the stable and unstable stability status responses compared to [DHT|DCT], [DCT|DWT]. The projection of signals ID 1 (blue) & ID 23 (red) on three hybrids dictionary learning allows effective prediction of online TSA, with a slightly improved performance compared to single dictionaries & hybrids with two sub-dictionaries, see Figure 9. Regarding 3H-OHD, the [DST|DHT|DCT] has lower RMSE predictors for the stable and unstable stability status responses compared to [ID|DCT|DST] and [DWT|DHT|DCT], which means a crisper separation between stable and unstable cases.

Figure 10 presents the stability degree of signals ID 1 and ID 23, developed based on stable and unstable dictionary learning. By definition, it is computed as the ratio between RMSE of unstable dictionary divided by the RMSE of stable dictionary-based reconstruction. Therefore, a small RMSE results in a high degree of stability meaning a higher probability of the case being stable while a small value means that the case is likely classified as unstable. For the stable case ID 1, the K-SVD gives a very large stability degree compared to [DCT], [DHT], [DWT], [DHT|DCT], [DCT|DWT], [DWT|DHT], [DST|DHT|DCT],

[ID | DCT | DST], and [DWT | DHT | DCT]. Similarly, K-SVD gives a very low stability degree of ID 23, which confirms a good stability separability of K-SVD based dictionary.
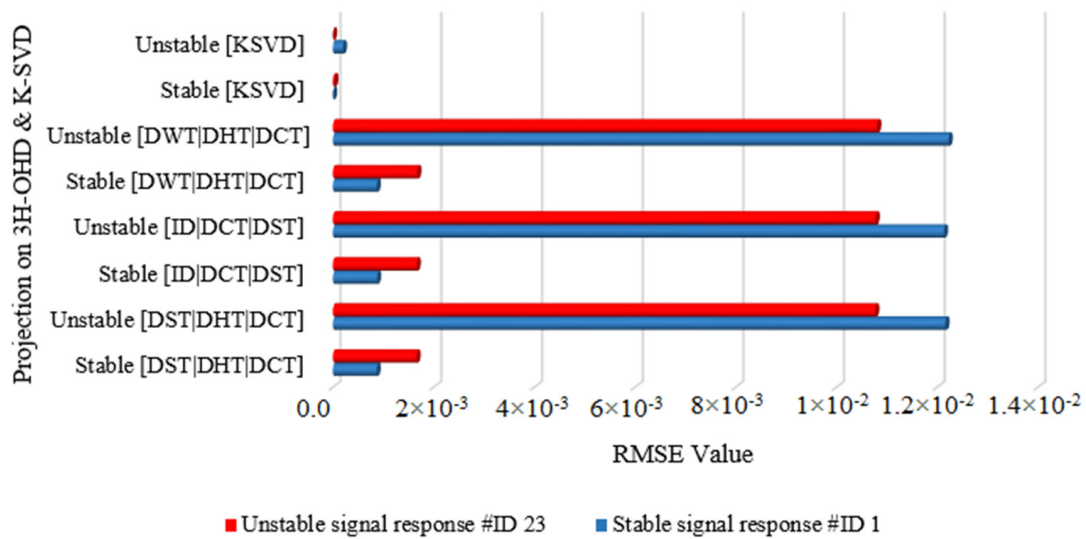

Figure 9. Projections of signals ID 1 and ID 23 on stable and unstable 3H-OHD and ADL.
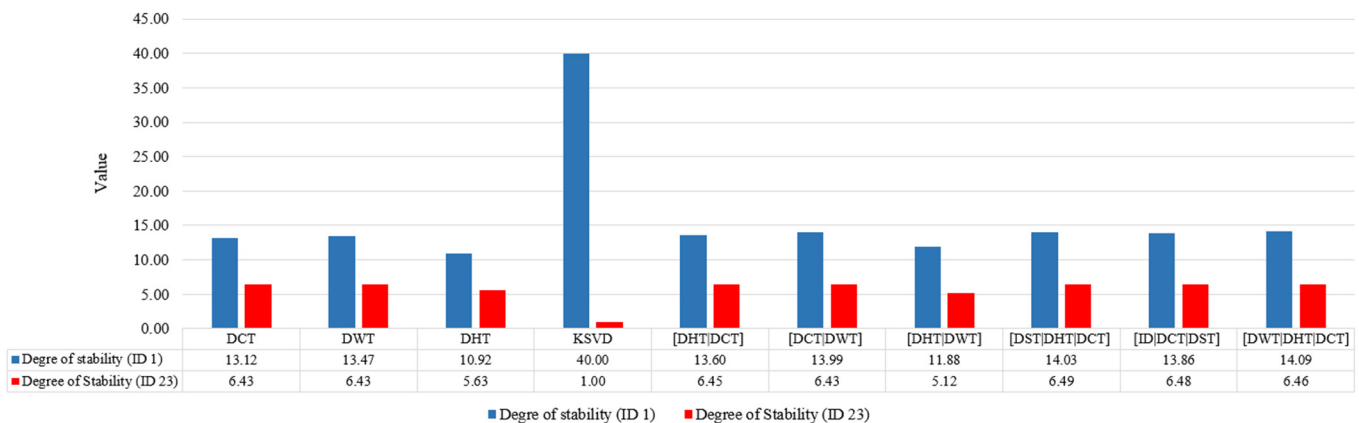

Figure 10. Stability degree of adaptive, single, hybrids with two and three overcomplete dictionary learning.

5.4. Performance Metrics of Sparse Decomposition on Fixed and Adaptive Overcomplete Dictionaries Learning

The metrics (namely: accuracy, reliability, and security) are used to evaluate the performance of fixed and adaptive overcomplete dictionary learning. The accuracy metric defines the ability of learned dictionaries for correctly classifying stable (i.e., 8341 scenarios) and unstable (i.e., 323 scenarios) in the testing database (i.e., 8664 scenarios). The reliability metric makes it possible to evaluate the performance of learned dictionaries to predict unstable scenarios. The security metric allows defining the capacity of learned dictionaries to predict the stable scenarios. They broadly succeed in projecting the new transient scenarios on each stable and unstable dictionary and evaluating the probability for each of them to have a low RMSE reconstruction. For each signal, the stability status classification is confirmed or not, according to the absolute values of generators rotor angles. Table 4 summarizes the online TSA effectiveness of each OHD and ADL developed based on separated training datasets. Moreover, the performance of OHD and K-SVD based stable/unstable dictionaries are compared to the supervised learning of overcomplete dictionaries (SLOD) developed in [34], based on both rotor speed and stability status taken as the training input and using a single dictionary containing both stable and unstable cases. It appears from

Table 4 that all fixed dictionaries provided a decent performance with right to all metrics using separate stable/unstable datasets in contrast to K-SVD dictionary whose reliability is pretty-weak.

Table 4. TSA metric comparisons of fixed and adaptive supervised dictionary learning.

Configuration	Learning Approaches	Accuracy	Reliability	Security
Adaptive dictionary (ADL) (with a separate training dataset)	K-SVD	93.42	81.68	93.35
SLOD [34] (with a joint training dataset)	TK-SVD	99.99	99.99	99.99
Single dictionary (SOD) (with a separate training dataset)	DCT	94.76	93.05	94.52
	DHT	95.79	94.10	95.66
	DWT	94.96	93.40	94.74
Two sub-dictionaries (2H-OHD) (with a separate training dataset)	[DHT | DWT]	95.41	93.75	95.24
	[DCT | DWT]	94.79	93.05	94.56
	[DHT | DCT]	94.82	93.05	94.59
Three sub-dictionaries (3H-OHD) (with a separate training dataset)	[DST | DHT | DCT]	94.88	93.05	94.66
	[ID | DCT | DST] [7]	94.88	93.05	94.66
	[DWT | DHT | DCT]	94.81	93.05	94.58

The single dictionary based on DHT seems in average to be the best post-disturbance stability predictor with a 94.1% reliability success rate and 95.79% overall accuracy. By contrast, the reliability of the ADL (K-SVD) algorithm is only 81.68%. This compared poorly with the SLOD (TK-SVD) algorithm proposed in [34], which provided 99.99% success rate across all performance metrics.

5.5. Computational Efficiency of Fixed and Adaptive Overcomplete Dictionaries Learning

The computational performance of fixed and adaptive dictionary learning was assessed on a DELL computer configured with the Intel processor i7-7700HQ 4-core running at 2.80 GHz with 16 GB of RAM. Although the CPU time is relatively significant, the actual code is in Matlab scripting language and can therefore be made faster using C-programming. Table 5 summarizes all CPU computational time for the offline learning and online playback for TSA. The proposed adaptive overcomplete dictionaries take more offline computation time (i.e., 298,475.89 s or about 83 h) for sparse features extraction. However, K-SVD sparse features enable a very quick online transient stability prediction (i.e., 4.45 s computation time) based on the generator signal responses (i.e., 8664 scenarios). Compared to the best dictionary learning algorithm, the SLOD (TK-SVD), which has a 409,744.06 s or about 114 h offline training time and 6.59 s online playback time, the ADL (K-SVD) is 37% and 48% faster respectively. Moreover, several paths for improving the performance metrics of the ADL (K-SVD) to fully benefit from this improved computational performance remains to be investigated, such as a longer data window for decision-making and enhanced data normalization before dictionary learning, i.e., by using for instance rotor speed deviation instead of the absolute speed.

Table 6 provides a summary of the advantages and disadvantages of supervised dictionary learning for transient stability assessment.

The different methods are further compared in Figure 11 using data from Table 4. The bottom line is that the reconstruction-based TSA is better done with fixed dictionaries. The DHT turned out to be the best fixed dictionary in terms of both reliability and security although the difference is not huge across all types of fixed dictionaries.

Table 5. CPU time (in sec.) of fixed and adaptive supervised dictionary learning.

Configuration	Learning Approaches	Offline Processing Time/s	Online Processing Time/s
Adaptive dictionary (ADL) (with a separate training dataset)	K-SVD	298,475.89	4.45
SLOD [34] (with a joint training dataset)	TK-SVD	409,744.06	6.59
Single dictionary (SOD) (with a separate training dataset)	DHT	16,518.44	6061.06
	DCT	42,290.29	17,547.32
	DWT	45,882.93	17,366.59
Two sub-dictionaries (2H-OHD) (with a separate training dataset)	[DHT | DWT]	31,535.18	13,854.68
	[DCT | DWT]	52,812.70	18,519.29
	[DHT | DCT]	55,830.41	20,278.90
Three sub-dictionaries (3H-OHD) (with a separate training dataset)	[DST | DHT | DCT]	33,949.11	14,489.51
	[ID | DCT | DST] [7]	28,617.70	12,320.68
	[DWT | DHT | DCT]	40,062.59	16,521.75

Table 6. Advantages and disadvantages of fixed and adaptive dictionary learning for TSA.

| Approaches	Advantage	Disadvantage
ADL	—Satisfactory accuracy and security; —Good online CPU time; —Implemented based on data-driven algorithm (K-SVD); —Used separate approach for ADL.	—Not satisfactory reliability. —Takes longer time for offline learning
SLOD [34]	—Good accuracy, reliability, security; —Satisfactory online CPU time; —Implemented based on data-driven algorithm (K-SVD); —Used joint approach for SLOD.	Takes longer time for offline learning.
SOD	—Satisfactory accuracy security and reliability; —Rectangular waveforms, which can take zero value and sample points in subintervals of $t \in [0, 1]$; —Grouping energy contained in the signals in low frequency coefficients; —Extracting information content at different positions and scales for subsequently reconstructing post-contingency; —Used separate approach for SOD.	—Not satisfactory online CPU time; —Not suitable for transient signal reconstruction.
2H-OHD	—Satisfactory accuracy security and reliability; —Allows extracting simultaneously the non-sinusoidal information, the frequencies of the oscillations and a representation in a unitary sub-interval, invariant by shifts. —Allows extracting simultaneously the real low frequency coefficients, the non-sinusoidal information and the oscillation frequencies; —Allows extracting simultaneously, a representation in a unit sub-interval, invariant by shifts and the low frequency real coefficients; —Used separate approach for SOD.	—Not satisfactory online CPU time; —Not suitable for transient signal reconstruction.
3H-OHD	—Satisfactory accuracy security and reliability; —Allows extracting simultaneously in the stability signals, the imaginary coefficients, the representations in a unitary sub-interval, invariant by shifts and the low frequency real coefficients; —Allows extracting frequency peaks, real low frequency coefficients and imaginary coefficients; —Allows extracting in a way, the real coefficients low frequency, the non-sinusoidal information, scales, representations in a sub-unit interval, invariant by shifts and the imaginary coefficients; —Used separate approach for 2H-OHD.	—Not satisfactory online CPU time; —Not suitable for transient signal reconstruction.
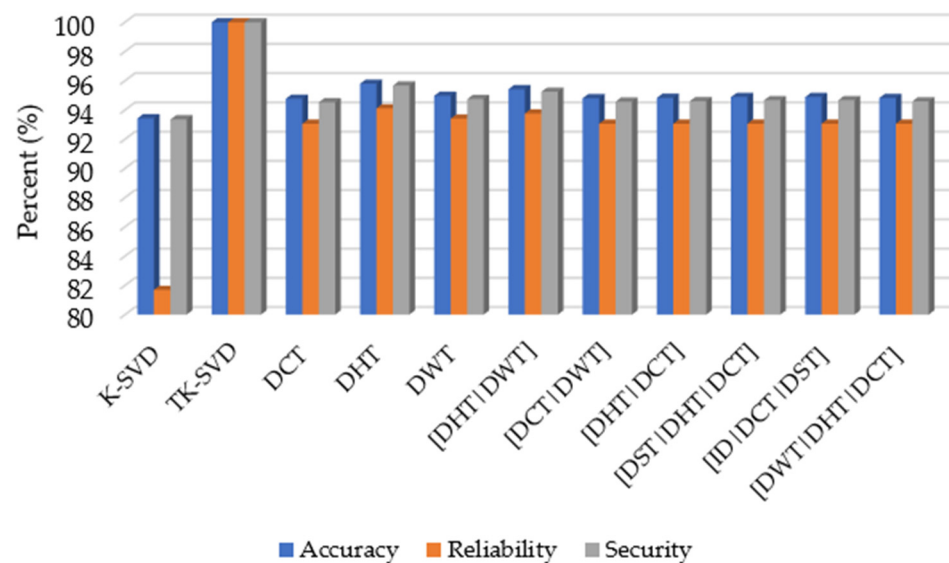
Figure 11. Comparison of adaptive, singles, and hybrids with two and three overcomplete dictionary learning.

Giving that DHT takes only about 50% of the computation time of the 3H-OHD from [7] in both supervised learning and online playback, it is definitely more promising and warrants further investigation in the TSA context. Additionally, given the short prediction window (75 ms) and the highly skewed two-data classes characterizing machine learning based TSA (with the unstable class typically representing only 3 to 10% of the total data set), the reconstruction-based classification using K-SVD appears unsuitable and rather, the TK-SVD based supervised learning algorithm of [34] should be further investigated and compared with the reconstruction-based classification using fixed DHT dictionary. These two options emerged as the best dictionary-based TSA prediction methods from the present study.

6. Conclusions

This work aims at reconstructing and predicting transient stability assessment based on fixed and adaptive overcomplete dictionary learning for online stability signal responses reconstruction. Practically speaking, this study develops and implements new arrangement of atoms able to improve wide-area stability monitoring and control using a dual-dictionary signal reconstruction approach. To this end, fixed and adaptive structures have been investigated based on DCT, DHT, and DWT dictionary learning for rotor speeds sparse reconstruction. Afterwards, stability status is determined based on the RMSE of rotor speed prediction using each stable/unstable dictionary to design the level of belonging to each class. The ratio of RMSE from these two separated predictions is proposed as measured of stability (or instability) of the given observed post-disturbance signal over a short 4.5-cyles data window.

Several concatenations of two (i.e., [DHT | DCT], [DCT | DWT], and [DHT | DWT]) and three (i.e., [DST | DHT | DCT], [ID | DCT | DST], and [DWT | DHT | DCT]) fixed sub-dictionaries are used to perform similar sparse decomposition on dual-dictionaries. Alternatively, the K-SVD adaptive approach is implemented using dual-dictionaries learned from the stable/unstable generator response datasets. While the adaptive K-SVD demonstrated a better reconstruction performance, in terms of prediction errors, its reliability performance was not satisfactory due to the highly skewed dataset derived from the IEEE 68-bus test system (3.8% of stable cases only) and short data frames (75 ms). Overall, this study proved that [DWT], [DHT | DWT], and [DST | DHT | DCT] fixed dictionaries are better stability indicators compared to adaptive K-SVD, [DHT], [DWT], [ID | DCT | DST], and [DWT | DHT | DCT] but none of these methods could match the good performance

of the method reported in [35]. We will further evaluate the performance of the K-SVD and DHT based dictionary learning methods for TSA on the combined database of IEEE 39 and 68 -bus test systems in the presence of multi-faults and for various simultaneous disturbances. Moreover, to reduce the offline training time, we also plan to develop parallel computer-based dictionary learning of PMU based generator signal responses to enable online TSA on large-scale power systems.

Author Contributions: Conceptualization, R.T.D. and I.K.; methodology, I.K., J.T. and F.C.M.; software, I.K.; validation, I.K., J.T. and F.C.M.; formal analysis, R.T.D. and I.K.; investigation, R.T.D. and I.K.; resources, I.K.; data curation, R.T.D., I.K., J.T. and F.C.M.; writing—original draft preparation, R.T.D., I.K.; writing—review and editing, R.T.D. and I.K.; visualization, I.K.; supervision, I.K.; project administration, I.K.; funding acquisition, I.K. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data are available on request to corresponding authors.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

K-SVD	K-singular value decomposition
TK-SVD	Transient K-singular value decomposition
DST	Discrete sine transform
DCT	Discrete cosine transform
ADL	Adaptive dictionary learning
DWT	Discrete wavelet transform
DHT	Discrete Haar transform
ID	Impulse dictionary
UPE	Unstable prediction error
SPE	Stable prediction error
OHD	Overcomplete hybrid dictionaries
SOD	Single overcomplete dictionaries
2H-OHD	Two overcomplete hybrid dictionaries
3H-OHD	Three overcomplete hybrid dictionaries
[DHT|DCT]	Concatenation of discrete Haar and cosine transforms
[DCT|DWT]	Concatenation of discrete wavelet and cosine transforms
[DHT|DWT]	Concatenation of discrete Haar and wavelet transforms
[DST|DHT|DCT]	Concatenation of discrete sine Haar and cosine transforms
[ID|DCT|DST]	Concatenation of discrete impulse, cosine and sine transforms
[DWT|DHT|DCT]	Concatenation of discrete wavelet, Haar and cosine transforms
Stable DHT	DHT dictionary learned from stable dataset
Unstable DHT	DHT dictionary learned from unstable dataset
Stable DWT	DWT dictionary learned from stable dataset
Unstable DWT	DWT dictionary learned from unstable dataset
Stable DCT	DCT dictionary learned from stable dataset
Unstable DCT	DCT dictionary learned from unstable dataset
Stable [DHT|DCT]	[DHT|DCT] dictionary learned from stable dataset
Unstable [DHT|DCT]	[DHT|DCT] dictionary learned from unstable dataset
Stable [DCT|DWT]	[DCT|DWT] dictionary learned from stable dataset
Unstable [DCT|DWT]	[DCT|DWT] dictionary learned from unstable dataset
Stable [DHT|DWT]	[DHT|DWT] dictionary learned from stable dataset
Unstable [DHT|DWT]	[DHT|DWT] dictionary learned from unstable dataset

Stable [DST	DHT	DCT]	[DST	DHT	DCT] dictionary learned from stable dataset
Unstable [DST	DHT	DCT]	[DST	DHT	DCT] dictionary learned from unstable dataset
Stable [ID	DCT	DST]	[ID	DCT	DST] dictionary learned from stable dataset
Unstable [ID	DCT	DST]	[ID	DCT	DST] dictionary learned from unstable dataset
Stable [DWT	DHT	DCT]	[DWT	DHT	DCT] dictionary learned from stable dataset
Unstable [DWT	DHT	DCT]	[DWT	DHT	DCT] dictionary learned from unstable dataset

References

1. Kamwa, I.; Jain, A.; Samantaray, S.R.; Geoffroy, L. Synchrophasors data analytics framework for power grid control and dynamic stability monitoring. *IET Eng. Technol. Ref.* **2016**, 1–22. [CrossRef]
2. Liu, Y.; Singh, A.K.; Zhao, J.; Meliopoulos, A.P.S.; Pal, B.; Ariff, M.A.B.M.; Van Cutsem, T.; Glavic, M.; Huang, Z.; Kamwa, I.; et al. Dynamic state estimation for power system control and protection. *IEEE Trans. Power Syst.* **2021**, *36*, 5909–5921. [CrossRef]
3. Paul, A.; Kamwa, I.; Joos, G. PMU Signals Responses-Based RAS for Instability Mitigation Through On-The Fly Identification and Shedding of the Run-Away Generators. *IEEE Trans. Power Syst.* **2020**, *35*, 1707–1717. [CrossRef]
4. Liu, Y.; Fan, R.; Terzija, V. Power system restoration: A literature review from 2006 to 2016. *J. Mod. Power Syst. Clean Energy* **2016**, *4*, 332–341. [CrossRef]
5. Yadav, R.; Pradhan, A.K.; Kamwa, I. Real-Time Multiple Event Detection and Classification in Power System Using Signal Energy Transformations. *IEEE Trans. Ind. Inform.* **2019**, *15*, 1521–1531. [CrossRef]
6. Stanković, L.; Sejdić, E.; Stanković, S.; Daković, M.; Orović, I. A Tutorial on Sparse Signal Reconstruction and Its Applications in Signal Processing. *Circuits. Syst. Signal Process.* **2019**, *38*, 1206–1263. [CrossRef]
7. Manikandan, M.S.; Samantaray, S.R.; Kamwa, I. Detection and Classification of Power Quality Disturbances Using Sparse Signal Decomposition on Hybrid Dictionaries. *IEEE Trans. Instrum. Meas.* **2015**, *64*, 27–38. [CrossRef]
8. Yu, Q.; Hu, Y.; Yang, Y. Identification Method for Series Arc Faults Based on Wavelet Transform and Deep Neural Network. *Energies* **2020**, *13*, 142. [CrossRef]
9. Qayyum, A.; Malik, A.S.; Naufal, M.; Saad, M.; Mazher, M.; Abdullah, F.; Abdullah, T.A.R.B.T. Designing of overcomplete dictionaries based on DCT and DWT. In Proceedings of the 2015 IEEE Student Symposium in Biomedical Engineering & Sciences (ISSBES), Shah Alam, Malaysia, 4 November 2015. [CrossRef]
10. Kathirvel, P.; Manikandan, M.S.; Maya, P.; Soman, K.P. Detection of power quality disturbances with overcomplete dictionary matrix and ℓ1-norm minimization. In Proceedings of the International Conference on Power and Energy Systems, Chennai, India, 22–24 December 2011; pp. 1–6.
11. Cai, D.; Li, K.; He, S.; Li, Y.; Luo, Y. A highly accurate and fast power quality disturbances classification based on dictionary learning sparse decomposition. *Trans. Inst. Meas. Control.* **2019**, *41*, 145–155. [CrossRef]
12. Zhu, T.X. Detection and Characterization of Oscillatory Transients Using Matching Pursuits With a Damped Sinusoidal Dictionary. *IEEE Trans. Power Deliv.* **2007**, *22*, 1093–1099. [CrossRef]
13. Bravo-Rodríguez, J.C.; Torres, F.J.; Borrás, M.D. Hybrid Machine Learning Models for Classifying Power Quality Disturbances: A Comparative Study. *Energies* **2020**, *13*, 2761. [CrossRef]
14. Babakmehr, M.; Sartipizadeh, H.; Simoes, M.G. Compressive Informative Sparse Representation-Based Power Quality Events Classification. *IEEE Trans. Ind. Inform.* **2020**, *16*, 909–921. [CrossRef]
15. Wang, Y.; Zhang, F.; Zhang, S. A New Methodology for Identifying Arc Fault by Sparse Representation and Neural Network. *IEEE Trans. Instrum. Meas.* **2018**, *67*, 2526–2537. [CrossRef]
16. Ren, L.; Lv, W.; Jiang, S.; Xiao, Y. Fault Diagnosis Using a Joint Model Based on Sparse Representation and SVM. *IEEE Trans. Instrum. Meas.* **2016**, *65*, 2313–2320. [CrossRef]
17. Liu, C.; Liu, J. Research on power quality signals reconstruction method based on K-SVD dictionary learning. In Proceedings of the Chinese Control Conference, Shenyang, China, 27–29 July 2020; pp. 2930–2934.
18. Chen, Y.C.; Dominguez-Garcia, A.D.; Sauer, P.W. A Sparse Representation Approach to Online Estimation of Power System Distribution Factors. *IEEE Trans. Power Syst.* **2015**, *30*, 1727–1738. [CrossRef]
19. Zhu, H.; Giannakis, G. Sparse Overcomplete Representations for Efficient Identification of Power Line Outages. *IEEE Trans. Power Syst.* **2012**, *27*, 2215–2224. [CrossRef]
20. Song, Y.; Wang, W.; Zhang, Z.; Qi, H.; Liu, Y. Multiple Event Detection and Recognition for Large-Scale Power Systems Through Cluster-Based Sparse Coding. *IEEE Trans. Power Syst.* **2017**, *32*, 4199–4210. [CrossRef]
21. Cheng, L.; Wang, L.; Gao, F. Power system fault classification method based on sparse representation and random dimensionality reduction projection. In Proceedings of the IEEE Power & Energy Society General Meeting, Denver, CO, USA, 26–30 July 2015; pp. 1–5.
22. Xiao, H.; Wei, J.; Li, Q. Identification of Combined Power Quality Disturbances Using Singular Value Decomposition (SVD) and Total Least Squares-Estimation of Signal Parameters via Rotational Invariance Techniques (TLS-ESPRIT). *Energies* **2017**, *10*, 1809. [CrossRef]
23. Chatterjee, A.; Yuen, P.W.T. Rapid estimation of orthogonal matching pursuit representation. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26–28 October 2020.
24. Aharon, M.; Elad, M.; Bruckstein, A. K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation. *IEEE Trans. Signal Process.* **2006**, *54*, 4311–4322. [CrossRef]

25. Xiang, Z.; Huang, K.; Deng, W.; Yang, C. Blind topology identification for smart grid based on dictionary learning. In Proceedings of the IEEE Symposium Series on Computational Intelligence, Xiamen, China, 6–9 December 2019; pp. 1319–1326.

26. Rajagukguk, R.A.; Ramadhan, R.A.; Lee, H.-J. A Review on Deep Learning Models for Forecasting Time Series Data of Solar Irradiance and Photovoltaic Power. *Energies* **2020**, *13*, 6623. [CrossRef]

27. Khodayar, M.; Wang, J.; Wang, Z. Energy Disaggregation via Deep Temporal Dictionary Learning. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *31*, 1696–1709. [CrossRef] [PubMed]

28. Kamwa, I.; Samantaray, S.R.; Joos, G. On the Accuracy Versus Transparency Trade-Off of Data-Mining Models for Fast-Response PMU-Based Catastrophe Predictors. *IEEE Trans. Smart Grid* **2012**, *3*, 152–161. [CrossRef]

29. Rubinstein, R.; Bruckstein, A.M.; Elad, M. Dictionaries for Sparse Representation Modeling. *Proc. IEEE* **2010**, *98*, 1045–1057. [CrossRef]

30. Kundur, P. *Power System Stability and Control*; McGraw-Hill: New York, NY, USA, 1994.

31. Akansu, A.; Haddad, R. *Multiresolution Signal Decomposition*, 2nd ed.; Academic Press: Cambridge, MA, USA, 2000; pp. 1–499.

32. Ma, B.; Zou, X. Discrete wavelet transform for signal processing in weight-in-motion system. In Proceedings of the International Conference on Electrical and Control Engineering, Washington, DC, USA, 25–27 June 2010; pp. 4668–4671.

33. Elad, M. *Sparse and Redundant Representations: From Theory to Application in Signal and Image Processing*; Springer: New York, NY, USA, 2010; pp. 1–376.

34. Pavella, M.; Ernst, D.; Ruiz-Vega, D. *Transient Stability of Power Systems: A Unified Approach to Assessment and Control*; Springer: New York, NY, USA, 2000; pp. 1–256.

35. Dabou, R.T.; Kamwa, I. Rapid design method for generating power system stability databases in SPS for machine learning. In Proceedings of the IEEE Canadian Conference on Electrical and Computer Engineering, London, ON, Canada, 30 August–2 September 2020; pp. 1–6.

36. Dabou, R.T.; Kamwa, I.; Tagoudjeu, J.; Mugombozi, C.F. Supervised learning of overcomplete dictionaries for on-line response based transient stability prediction. *IEEE Trans. Power Syst.* **2021**. submitted.

37. Lachiheb, O.; Gouider, M.S.; Said, L.B. An Improved MapReduce design of Kmeans with iteration reducing for clustering stock exchange very large datasets. In Proceedings of the 2015 11th International Conference on Semantics, Knowledge and Grids (SKG), Beijing, China, 19–21 August 2015; pp. 252–255.