Received: 7 November 2017 | Revised: 4 June 2018 | Accepted: 14 July 2018

DOI: 10.1002/etep.2693

RESEARCH ARTICLE

WILEY

Sparse representation algorithm applied to power systems signal compression

Leandro Rodrigues Manso Silva | Luciano Manhaes de Andrade Filho | Carlos Augusto Duque

Electrical Engineering, Federal University of Juiz de Fora, Juiz de Fora, Brazil

Correspondence
Leandro Rodrigues Manso Silva, Electrical Engineering, Federal University of Juiz de Fora, Juiz de Fora, MG 36036-900, Brazil.
Email: leandro.manso@engenharia.ufjf.br

Summary

Power systems signal compression is an important task considering the nature of electrical disturbances, which must be analyzed offline in order to extract some useful information. This paper presents a comparative study about greedy algorithms to obtain sparse representation of signals applied to power systems signal compression. The algorithms will be compared according to two parameters: the quality of the compressed signal, in terms of its correlation coefficient related to the original signal, and the number of elements in the representation, which is related to the compression ratio. In addition, a computational complexity, in terms of required flops, will be performed, aiming at identifying which algorithm is most suitable to run in real time. Finally, a comparison with the wavelet transform, using a database of real power system signals, will be shown.

KEYWORDS

orthogonal matching pursuit, power system signal compression, sparse representation, wavelet transform

1 | INTRODUCTION

The analysis of power systems signal has become increasingly important, mainly in the smart grid scenario, where there is a large penetration of distributed generation and nonlinear loads that cause some deformities in the voltage and current signals. Some of these deformities are already known, and there are specific equipment designed to deal with them. Otherwise, some new disturbances, not known yet, may occur and require further studies. In this way, the signal must be stored to a proper analysis. The storage of the voltage and current signals over a long period takes a huge amount of memory and requires the use of a compression technique running in real time.[1]

Although aggregation is useful to data reduction and comparison, deep data analysis should be enabled, considering the possibilities of a full observation.[2] The storage of the raw signal of voltage and current is supported by the fact that offline (more complex) techniques may reveal information that cannot be perceived by online (simpler) techniques. For example, high-frequency disturbances like transient oscillations or switching processes are only visible in the full waveform, while the signature "hidden" in the raw data can be used to predict the breaking of cables.

The signal compression can be divided in two categories: lossless compression and lossy compression. In the former, no information is discarded, and in the latter, some information is discarded in order to achieve higher compression ratio (CR). Lossless techniques based on coding algorithms such Huffman, LZ78, are used in electrical signal processing being able to achieve a CR[3] up to 3. These low CR values occur because no information is discarded.[4] Furthermore, in this context, the power quality (PQ) data interchange format provides a lossless compression.[5] Other techniques are found

List of Symbols and Abbreviations: CR, Compressio Ratio; PQ, Power Quality; WT, Wavelet Transform; MP, Matching Pursuit; OMP, Orhtogonal Mtching Pursuit; ROMP, Regularized Orhtogonal Matching Pursuit; RMP, Regularized Matching Pursuit; StOMP, Stagewise Orhtogonal Matching Pursuit; flop, Floating Point Operations in the literature, such as the use of high-order delta modulation, which is computationally simple and produces a high sample rate as proposed in Zhang et al.[6]

The power systems signal, as well as speech and video signals, has a significant amount of redundancy and useless information in its samples. Therefore, the use of proper signal processing techniques enable to obtain a sparse approximation of the signal reducing the size of data to be stored. The sparse representation must accurately reproduce all the important information in the signal with a smaller dimension than the original one.[7]

Thus, lossy techniques are applied to achieve higher CRs, in which useless information is discarded. The most common techniques used to obtain a sparse representation are orthogonal transforms, such as Fourier transform, wavelet transform (WT), wavelet packet transform, and cosine transform, followed by a thresholding in order to reduce the signal dimension.[8]

The WT is widely used due to its ability to deal with time-varying signals such as the power system signals. The methodologies based on WT usually can be divided in three steps: (1) signal decomposition, (2) thresholding of the coefficients, and (3) coding algorithm.[9]

Other works about compression applied to power systems are found in the literature, such as that uses WT to compress magnitude and frequency signals acquired from the simulation of the IEEE New England 39-bus system[10] and that proposed an real time algorithm to compress phasor measurement unit in order to reduced the amount of data in a wide area monitoring system.[11]

Sparse representation over redundant dictionaries methods has been used in various areas, such as dictionary learning, image processing, image classification, visual tracking,[12] and in power system signal compression.[13,14] The matching pursuit (MP) algorithm, proposed in Mallat and Zhang[15] or some of its variations, is used in most studies of the area.

In this context, where a real-time compression technique is needed, the aim of this paper is to test several greedy algorithms such as MP, orthogonal matching pursuit (OMP), and some of their variations in order to determine which is most suitable to be executed in real time. The algorithms will be tested and compared according to their compression capability, in terms of the CR and representation quality, and the computational complexity.

This paper is divided as follows: Section 2 shows a description of a power system oscillography signal compression system. Section 3 presents a review of the sparse representation theory and the algorithms for sparse representation that will be used for the comparison. In Section 4, dictionary generation for the compression of power systems signal will be presented. Section 5 presents the results related to the comparison of the sparse algorithms and also shows a comparison between the chosen sparse representation technique and the WT. And finally, in Section 6 are the conclusions.

2 | POWER SYSTEM SIGNAL COMPRESSION

As mentioned in Section 1, the WT consists of a technique widely used for power system compression, but it is only a part of a power system signal compression method. A whole compression system is generally based on four stages: (A) segmentation; (B) detection; (C) lossy compression; and (D) lossless compression, as shown in Figure 1.

The segmentation (A) is responsible to divide the signal in windows or frames that will be compressed and stored. The detection (B) is applied in order to determine which frame needs to be stored. If two frames are equal, only one needs to be stored, therefore reaching a compression.

The lossy compression (C) is applied in the detected frames aiming at eliminating the useless information present in the signal. Finally, the lossless compression (D) is performed to reduce even more the size of the data to be stored without discarding any information.

The present work focuses on the lossy compression part, proposing a sparse representation technique, based on a redundant dictionary constituted of the union of Fourier and wavelet basis.

3 | SPARSE REPRESENTATION OF SIGNALS

Before introducing the sparse representation, it is important to understand the atomic signal decomposition that consists in representing the signal using predefined waveforms. Mathematically, a signal $\mathbf{b} \in \mathbb{R}^N$ is represented by a linear combination of M columns of the matrix $\mathbf{A} \in \mathbb{R}^{N \times M}$ as described in (1).

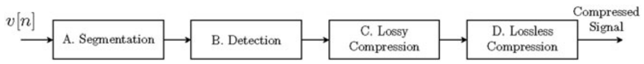

FIGURE 1 Block diagram of a generic power system signal compression method

$$\mathbf{Ax} = \mathbf{b}, \tag{1}$$

where $\mathbf{x} \in \mathbb{R}^M$ is the coefficient vector that maps the signal $\mathbf{b}$ into the subspace spanned by matrix $\mathbf{A}$. In the atomic decomposition context, the matrix $\mathbf{A}$ is called dictionary, and its columns are atoms.

Atomic signal decomposition techniques are used in many areas, such as denoising[16]; harmonic analysis[17]; parameter extraction[18]; and time-frequency decomposition.[19] In atomic decomposition, the matrix $\mathbf{A}$ is considered a redundant dictionary, since it has more elements than the dimension of vector $\mathbf{b}$. In this way, $\mathbf{A}$ has more columns than rows ($M > N$), and the linear system described in (1) has various solutions, and the sparse representation of the signal is the one that uses the lowest number of elements.

Two major issues are involved in the use of sparse representation techniques: (1) given a dictionary matrix, how to find the solution of (1) with less elements and (2) how to build the dictionary.

To solve the first problem, it is necessary to establish a criteria to express the characteristics of the desired solution by writing a generalized optimization problem with a cost function $J(.)$, as follows:

$$(P_J) \ : \ \min_{\mathbf{x}} J(\mathbf{x}) \ \text{subject to} \ \mathbf{b} = \mathbf{Ax}. \tag{2}$$

Therefore, it is desired to minimize the cost function $J(\mathbf{x})$ regarding $\mathbf{x}$, and this $\mathbf{x}$ must be constrained to $\mathbf{b} = \mathbf{Ax}$. The choice of the $J(.)$ function determines the features of the solution to be obtained. It is known that l_p norm, for $p < 1$, leads to sparser solutions, including the l_0 norm, which is the one that best describes the sparsity since its value is the number of nonzero entries in the vector. This way, the problem $P(J)$ turns into the $P(0)$.

$$(P_0) \ : \ \min_{\mathbf{x}} \|\mathbf{x}\|_0 \ \text{subject to} \ \mathbf{b} = \mathbf{Ax}. \tag{3}$$

Because of the discontinuous and discrete nature of l_0, the solution of (3) implies some challenges. It is a combinatorial search problem among all the possible sparse subsystems $\mathbf{b} = \mathbf{A}_S\mathbf{x}_S$, where $\mathbf{A}_S\mathbf{x}_S$ is a matrix containing only $|S|$ columns of $\mathbf{A}$. The complexity of this solution is m exponential; therefore, (P_0) is an NP-hard problem.[20]

Since the direct solution of (P_0) is unfeasible in computational terms, it is important to seek other reliable solution. The task of finding $\mathbf{x}$ can be divided in two stages: (1) finding the support of $\mathbf{x}$, or its nonzero elements, and (2) determine the values of these elements. Therefore, a way of attacking this problem is to focus on the support considering that after being found, its coefficients can be obtained by applying $\hat{\mathbf{x}} = \mathbf{A}^{+}\mathbf{b}$, where $\mathbf{A}^{+}$ is the pseudoinverse of $\mathbf{A}$.[21] This line of thought leads to the family of greedy algorithms, which include the MP[15] and the OMP,[22] highlighted due to their utilization in digital signal processing field.

3.1 | Algorithms for sparse representation

As abovementioned, finding the sparsest solution of an undetermined linear system is a restricted optimization problem. A cost function that favors the sparsity is the l_0 norm originating the (P_0) problem, as described in (3). However, in signal compressing, it is more interesting to apply a relaxed problem, such as the $(P_{0,\epsilon})$, shown in (4), since the signal contains undesired information, such as noise.

$$\left(P_{0,\epsilon}\right) \ : \ \min_{\mathbf{x}} \|\mathbf{x}\|_0 \ \text{subject to} \ \|\mathbf{b} - \mathbf{Ax}\|_2 \le \epsilon, \tag{4}$$

where $\epsilon > 0$ is the error tolerance, allowing a discrepancy between the representation $\mathbf{Ax}$ and the signal $\mathbf{b}$.

A way to solve (P_0) or $(P_{0,\epsilon})$ consists of an exhaustive search for the sparsest solution. For example, if $\mathbf{A}$ has n rows and m columns, and $\mathbf{b}$ is a linear combination of a maximum of k_0 elements, all combinations of k_0 elements of $\mathbf{A}$ should be numbered and tested individually to find the best solution. For this, $O(m^{k_0} n k_0^2)$ flops are required, which makes this solution unfeasible.

The class of algorithm named greedy algorithm aims at solving these problems in a reliably and efficiently way. A greedy algorithm abandons the exhaustive search and works with a series of single-term updates by starting from $\mathbf{x}^0 = \mathbf{0}$ and adding elements iteratively to build an approximation $\mathbf{x}^k$ with k nonzero elements.

3.1.1 | Orthogonal matching pursuit

The OMP is a greedy algorithm that selects the dictionary element with highest orthogonal projection in the residual for each iteration. The selected element is added to the solution, and the coefficients of all elements are updated via linear regression. A description of this algorithm is depicted in Figure 2.

Task: Solve $(P_0): \min_{\mathbf{x}}$ sujeito a $\mathbf{Ax} = \mathbf{b}$

Initialization:

- Initial Solution: $\mathbf{x}^0 = \mathbf{0}$
- Initial Residual: $\mathbf{r}^0 = \mathbf{b} - \mathbf{Ax}^0 = \mathbf{b}$
- Initial Support: $S^0 = Suporte\{\mathbf{x}^0\} = \emptyset$

Main Iteration: $k = k + 1$

- **Sweep:** Compute $\epsilon(j) = \min_{z_j} \left\|\mathbf{a}_j z_j - \mathbf{r}^{k-1}\right\|_2^2$ for all j using the optimal choice $z_j^* = \mathbf{a}_j' \mathbf{r}^{k-1} / \left\|\mathbf{a}_j\right\|_2^2$.
- **Update Support:** Find a minimizer j_0, $\epsilon(j_0) \leq \epsilon(j)$ and update $S^k = S^{k-1} \cup j_0$.
- **Update Provisional Solution:**
  - **OMP**: Compute $\mathbf{x}^k$, the minimizer of $\|\mathbf{Ax} - \mathbf{b}\|_2^2$ sujeito a S^k
  - **MP**: Set $\mathbf{x}^k = \mathbf{x}^{k-1}$, and update the entry $x^k(j_0) = x^k(j_0) + z_j^*$
- **Update Residual:** Compute $\mathbf{r}^k = \mathbf{b} - \mathbf{Ax}^k$
- **Stopping Rule:** If $\left\|\mathbf{r}^k\right\|_2 \leq \epsilon_0$, stop.

Output: The solution $\mathbf{x}^k$ is obtained after k iterations.

FIGURE 2 Orthogonal matching pursuit (OMP) and matching pursuit (MP)[21]

The sweep stage calculates the inner products between the residual $\mathbf{r}^{k-1}$ and each column $\mathbf{a}_j'$ of $\mathbf{A}$; at the update support stage, the index j_0 of the element with the biggest inner product is added to the support; the update of the provisional solution is performed via least squares by using only the $\mathbf{A}$ matrix elements related to the support of $\mathbf{x}$; and, in the end, a new residual $\mathbf{r}^k$ is calculated, and the stopping rule is evaluated.

The computational complexity of the OMP, supposing that the solution has k_0 elements, is $O(k_0 mn)$ flops, which is lower than the $O(m^{k_0} n k_0^2)$ flops required for the exhaustive search.

3.1.2 | Matching pursuit

The MP algorithm is similar to the OMP, the difference is at the update provisional solution stage, where the coefficients that had already been part of the support S^{k-1} are not updated, while the new coefficient related to j_0, chosen as z_j^*. The MP is simpler than the OMP and therefore less accurate. A description of this algorithm is shown in Figure 2.

3.1.3 | Stagewise orthogonal matching pursuit

The stagewise orthogonal matching pursuit (StOMP), proposed in Donoho et al,[23] is based on the OMP. The main difference is that the StOMP builds the solution by adding various elements for each iteration while the OMP adds only one. At the update support stage, the StOMP adds to the support the indexes of all elements that had an inner product higher than a prespecified threshold T. The coefficients are then obtained through least squares. A description of this algorithm is in Figure 3.

An advantage of this method is that it is capable to produce a good representation with few iterations since several elements can be simultaneously added. The strongest difficulty is the setting of the threshold value T, since each iteration has different ranges inner products values. When T is too big, the resulting approximation may not be sufficient, and when T is chosen to prioritize a good approximation, it is possible to add too many elements to the method, which is not good for compression algorithms.

3.1.4 | Regularized orthogonal matching pursuit

An algorithm based on the OMP, which is very similar to the StOMP, is the regularized orthogonal matching pursuit (ROMP),[24] which overcomes the issue of working with a fixed threshold value and use T as the half of the maximum inner product value obtained at the sweep stage. Therefore, the threshold adapts to the inner product values at each iteration. A description of ROMP is shown in Figure 4.

With the adaptive threshold, the algorithm is capable to select the elements that are really significant, providing a high quality representation with less elements than the StOMP.

Task: Solve (P_0) : $\min_{\mathbf{x}}$ sujeito a $\mathbf{Ax} = \mathbf{b}$

Initialization:

- Initial Solution: $\mathbf{x}^0 = \mathbf{0}$
- Initial Residual: $\mathbf{r}^0 = \mathbf{b} - \mathbf{Ax}^0 = \mathbf{b}$
- Initial Support: $S^0 = Suporte\{\mathbf{x}^0\} = \emptyset$
- Threshold choice $= T$.

Main Iteration: $k = k + 1$

- **Sweep:** Compute the inner products $z_j = \mathbf{a}_j' \mathbf{r}^{k-1}$
- **Update Support:** Find the set J constituted by the index of the elements that present inner product greater than T and update the support $S^k = S^{k-1} \cup J$
- **Update Provisional Solution:** Compute $\mathbf{x}^k$ that minimizes $\|\mathbf{Ax} - \mathbf{b}\|_2^2$ sujeito a S^k
- **Update Residual:** Compute $\mathbf{r}^k = \mathbf{b} - \mathbf{Ax}^k$
- **Stopping Rule:** IF $\|\mathbf{r}^k\|_2 \leq \epsilon_0$, stop.

Output: The solution $\mathbf{x}^\mathbf{k}$ is obtained after k iterations.

FIGURE 3 Stagewise orthogonal matching pursuit

Task: Solve (P_0) : $\min_{\mathbf{x}}$ sujeito a $\mathbf{Ax} = \mathbf{b}$

Initialization:

- Initial Solution: $\mathbf{x}^0 = \mathbf{0}$
- Initial Residual: $\mathbf{r}^0 = \mathbf{b} - \mathbf{Ax}^0 = \mathbf{b}$
- Initial Support: $S^0 = Suporte\{\mathbf{x}^0\} = \emptyset$

Main Iteration: $k = k + 1$

- **Sweep:** Compute the inner products $z_j = \mathbf{a}_j' \mathbf{r}^{k-1}$
- **Threshold Calculation:** Find the maximum value among inner products $z_{j_{max}} = \max(|z_j|)$ and set $T = z_{j_{max}}/2$.
- **Update Support:** Find the set J constituted by the index of the elements that present inner product greater than T and update the support $S^k = S^{k-1} \cup J$
- **Update Provisional Solution:** Compute $\mathbf{x}^k$ that minimizes $\|\mathbf{Ax} - \mathbf{b}\|_2^2$ sujeito a S^k
- **Update Residual:** Compute $\mathbf{r}^k = \mathbf{b} - \mathbf{Ax}^k$
- **Stopping Rule:** IF $\|\mathbf{r}^k\|_2 \leq \epsilon_0$, stop.

Output: The solution $\mathbf{x}^\mathbf{k}$ is obtained after k iterations.

FIGURE 4 Regularized orthogonal matching pursuit

4 | SPARSE SIGNAL REPRESENTATION IN POWER SYSTEMS SIGNAL COMPRESSION

An overcomplete dictionary matrix **A** in sparse representation applications can be chosen as a predefined matrix, or it can be constructed adaptively from a set of signal samples, using an automatic learning method such as K-SVD.[25] The choice for a predefined matrix, based on the previous knowledge about the signals, leads to simple and fast algorithms for the evaluation of the sparse representation. Therefore, the choice for a predefined dictionary is more suitable for real-time applications.

In power systems applications, the tools commonly used to analyze and represent signals are the Fourier transform, which corresponds to a linear combination of sinusoidal components and the WT, which uses finite support functions. The signal representation in only one of these bases is not efficient on a sparse representation point of view, since Fourier basis represents well stationary signals and the wavelet basis is capable to represent transient signals more efficiently.[15]

It is known that in power systems, voltage, and current signals are composed of both stationary components (fundamental, harmonic, and interharmonic) and damped transient components, such as the oscillatory transients. The former are well represented on Fourier basis and the latter on the Wavelet basis. This way, power systems signals would be very well represented by a dictionary formed through the union of these two bases.

The choice of the dictionary is very important since it will have direct impact on the sparsity of the approximation. Thus, the dictionary composed of the union of these two bases will constitute a frame, as shown in (5), where **F** represents the Fourier basis and **W** the wavelet basis. Therefore, a signal will not have only one representation over it, while this dictionary is said to be redundant.[26] Because of the redundancy, frames provide many representations to the same signal; it is desired to reach the sparsest possible.

$$\mathbf{A} = [\mathbf{F} \mid \mathbf{W}]_{N \times M} \tag{5}$$

In this work, the signal was segmented in “frames” containing four cycles of the fundamental component. Since the sampling rate used is 7680 Hz, the size of each “frame” is N = 512 samples. The dictionary is composed of $M = 612$ elements, 100 sinusoids of Fourier (**F**) basis, representing the quadrature components for fundamental and harmonic components up to the 50th and 512 wavelet functions obtained from scaling, translating, and modulating a Daubechies 5 mother wavelet[27] in the wavelet (**W**) basis. So the cardinality of the dictionary is then 512×612.

The number of decomposition level for the WT was obtained using the entropy criterion, as described in Coifman and Wickerhauser,[28] where the signal is decomposed until the entropy of the coefficients of the last level is lower than the entropy of the coefficients of previous level. The choice of the mother wavelet was based on the minimum description length criterion proposed in Hamid and Kawasaki.[29] This criterion aims to minimize the relationship between the number of retained coefficients and the error of the represented signal related to the original one.

In the next section, the performance of the algorithms described in Section 3 will be analyzed aiming at finding the best one to be implemented in real time in a signal-compressing system. The chosen algorithm must be capable to represent signals containing PQ disturbances efficiently at a feasible computational cost.

5 | RESULTS

Aiming at testing the representation capability and the compression rate provided by the algorithms, two performance parameters will be evaluated: the correlation coefficient between the original signal and the reconstruction and the number of elements used in the representation. The correlation coefficient is defined by

$$\mathrm{COR} = \frac{\mathbf{x}^T \cdot \hat{\mathbf{x}}}{\mathbf{x}^T \cdot \mathbf{x}}, \tag{6}$$

where $\mathbf{x}$ is the vector notation for the original signal and $\hat{\mathbf{x}}$ is the vector notation of the reconstructed signal.

The CR is calculated as

$$\mathrm{CR} = \frac{N_s}{N_e}, \tag{7}$$

where N_s is the number of samples of the original signal, which would be stored if no compression was applied, and N_e is the number of elements that needs to be stored after the compression. It is used to show the CR as follows: $\mathrm{CR} : 1$, and it means that the size of the data was reduced CR times.

5.1 | Algorithm comparison results

For these tests, the signals were divided in four classes: (1) oscillatory transients; (2) stationary harmonics; (3) time-varying harmonics (harmonics magnitude varies with time); and (4) sags and swells.

The tests were performed as follows: A hundred signals of each disturbance class were generated with random parameters. Each signal was submitted to each algorithm, and the graphics were generated with the mean value for the correlation coefficient and the of number elements for each iteration. The tested algorithms are the greedy MP, OMP, StOMP, ROMP, and a variation without the least squares in the coefficient update stage, named Regularized Matching Pursuit (RMP).

The results will be analyzed using a type of rate distortion curve, that is, a curve that expresses the quality of the reconstructed signal in terms of the percentage of dictionary elements used in the representation. This curve enables to analyze which algorithm is capable to reach better quality with less elements in few iterations. In these curves, each mark represents one iteration of the algorithm.

The performance of the algorithms for oscillatory transients is shown in Figure 5.

In Figure 5, it can be noted that both the OMP and the MP algorithms perform almost in the same manner since each adds one element per iteration, and due to the characteristics of the dictionary matrix that is composed of two orthogonal sets, the orthogonality stage of the OMP does not make difference in the result. It is important to highlight the ROMP

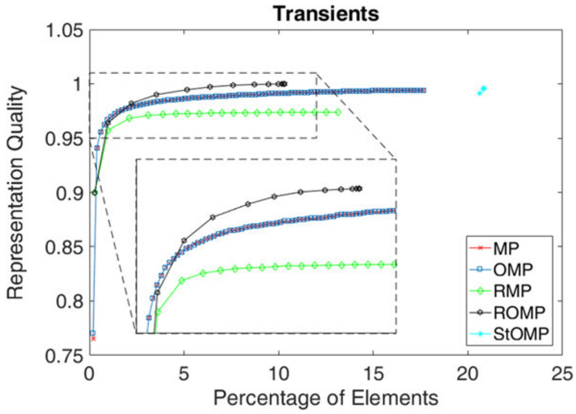

FIGURE 5 Correlation coefficient for oscillatory transient signals. MP, matching pursuit; OMP, orthogonal matching pursuit; ROMP, regularized orthogonal matching pursuit; StOMP, stagewise orthogonal matching pursuit

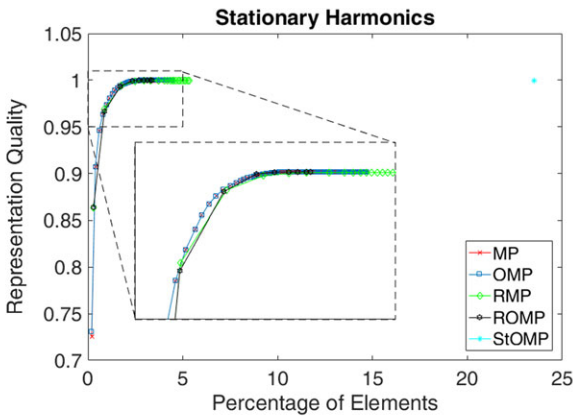

FIGURE 6 Correlation coefficient for stationary harmonic signals. MP, matching pursuit; OMP, orthogonal matching pursuit; ROMP, regularized orthogonal matching pursuit; StOMP, stagewise orthogonal matching pursuit

algorithm for its capacity to well represent the signal with few elements and more importantly, with few iterations. The RMP algorithm did not perform well for this case since it does not update all coefficients at each iteration; therefore, the representation coefficients are not calculated optimally. The StOMP algorithm works by adding all the elements with an inner product higher than a fixed threshold value. If a high value is chosen, the quality of the representation may not be good, while if a small value is chosen, many elements will be added in the representation. As each signal has a different composition, a small threshold value was chosen to prioritize good quality representation. In this way, the StOMP algorithm has added more elements than the other to reach the same quality, making it not suitable to the proposed application.

The results for stationary harmonics are shown in Figure 6. In this case, algorithms present a similar behavior. The RMP algorithm performs closer to the ROMP, which occurs because of to the natural orthogonality in the signal components, composed only of sinusoids at harmonic frequencies.

The same tests were performed for the other classes of disturbances, with the results shown, in Figure 7 for the time-varying harmonics and in Figure 8 for sags and swells. Similar analysis and conclusion can be reached for the algorithms.

Considering all the tested disturbances, it is possible to conclude that the StOMP algorithm is not appropriate for this application since it works with a fixed threshold and the number of elements does not change over the iterations. In this way, it will be disregarded for the next analysis. In addition, it is possible to observe that the ROMP algorithm is the most suitable, for being able to represent all the signals well, by using few elements with less iterations than others.

It is also worth mentioning that the same quality was obtained by the ROMP and OMP, since the two algorithms are almost the same. The only difference is that ROMP select various elements per iteration, while OMP selects only one. The quality obtained by the MP and RMP is lower, and the number of elements is higher since these algorithms do not present the orthogonalization stage that favors the sparsity and the quality of the solution.

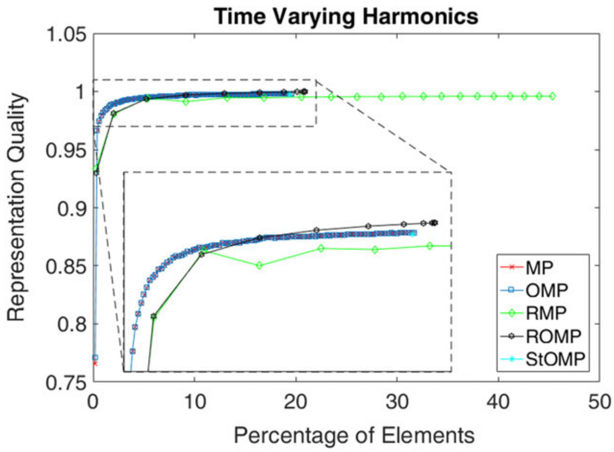

FIGURE 7 Correlation coefficient for time-varying harmonic signals. MP, matching pursuit; OMP, orthogonal matching pursuit; ROMP, regularized orthogonal matching pursuit; StOMP, stagewise orthogonal matching pursuit

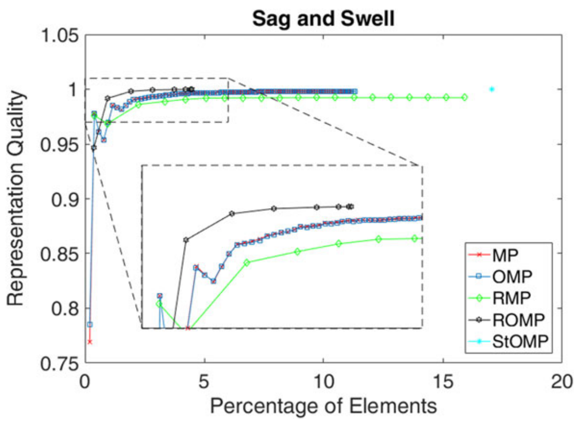

FIGURE 8 Correlation coefficient for sag and swell signals. MP, matching pursuit; OMP, orthogonal matching pursuit; ROMP, regularized orthogonal matching pursuit; StOMP, stagewise orthogonal matching pursuit

TABLE 1 Computational complexity of each algorithm

Method	Stage	
	Sweep	Update
MP	$(2N-1) \cdot M \cdot k$	-
OMP	$(2N-1) \cdot M \cdot k$	$3 \cdot k_0^2 \cdot N + 2N - 1$
RMP	$(2N-1) \cdot M \cdot k$	-
ROMP	$(2N-1) \cdot M \cdot k$	$3 \cdot k_0^2 \cdot N + 2N - 1$

Abbreviations: MP, matching pursuit; OMP, orthogonal matching pursuit; ROMP, regularized orthogonal matching pursuit.

Since the idea is to implement the technique in real time, the computational complexity of each algorithm must be evaluated. At the sweep stage, all of them must perform the same task: Calculate the inner product of the signal vector by each column of A matrix, to such purpose, $(2N-1) \cdot M$ flops (floating point operations) are required for each iteration. The orthogonalization stage in OMP and ROMP is performed through the least-squares application, and for that, $3 \cdot k_0^2 \cdot N$ flops are necessary for the pseudo-inverse calculation and $2N-1$ flops are required for the coefficient update, where k_0 is the number of elements in the representation.

Table 1 shows the number of flops per iteration for each algorithm, where k is the number of iterations and k_0 is the number of elements in the end of the iteration. To illustrate the behavior of each algorithm related to the computational complexity, Figure 9 was made using all disturbance signals.

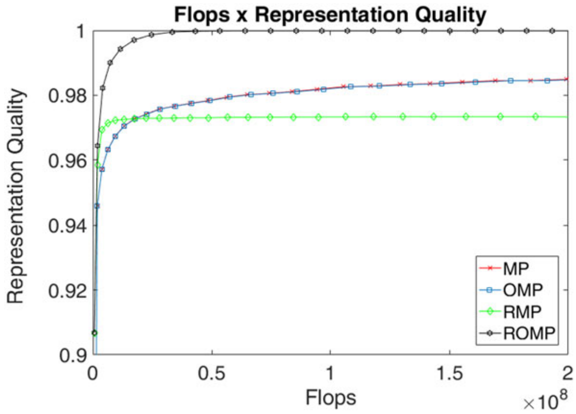

FIGURE 9 Computational complexity curve. MP, matching pursuit; OMP, orthogonal matching pursuit; ROMP, regularized orthogonal matching pursuit

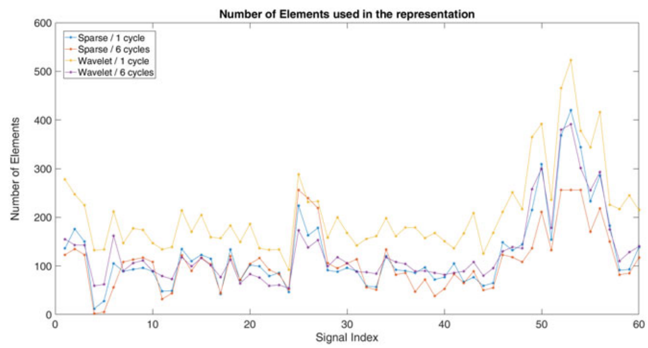

FIGURE 10 Number of elements to represent each signal from IEEE database[30]

In Figure 9, the accumulated number of flops for each iteration for the four algorithms is shown. The curve shows the number of flops versus the reconstruction quality. Despite having higher computational complexity per iteration, the ROMP algorithm needs less iterations to converge, so in the end, it has the lowest computational complexity. The previous analysis enabled to conclude that the ROMP is the most suitable algorithm among all the tested ones, to be implemented in real time applications for providing satisfactory representation quality and lower computational complexity.

5.2 | Power system signal compression results

The previous section showed a comparison of several algorithms of sparse representation applied to power system signals, showing that among the algorithms compared, the algorithm named ROMP had the best performance at representing the signals using the proposed dictionary. This section aims at comparing the sparse representation technique with the classical WT in compressing power system signals.

For the tests in this section, the IEEE Power Quality Database[30] was used. The bank contains 60 signal, sampled at 15.360 kHz, with six cycles each. Therefore, each cycle has 256 samples, and the entire window has 1536 samples. These signals are real signals corrupted by PQ disturbances.

The mother wavelet used for both, WT and dictionary construction, is the Daubechies 5 with five levels of decomposition. Two scenarios were tested; at the first one, a six cycles window was used for the compression. So the dictionary has 1636 elements, being 1536 belonging to the wavelet basis, and 100 to the Fourier basis, corresponding to the first 50 harmonic components. In the second case, a window with a length of one cycle was used, and the dictionary is constituted of 356 elements, 256 of the wavelet basis, and 100 of the Fourier basis.

Figure 10 presents the results obtained by the two methods, each one using one cycle window and six cycles window. As can be noted from Figure 10, the sparse representation method using a six-cycle frame was the one that needed a lower number of components to represent the majority of signals in the data bank; only in five cases the WT with 6 cycles needed less components. The average values of the number of elements/CR obtained by the two methods in both scenarios of test are shown in Table 2. It is important to mention that all the obtained representations has MSE lower than 10^{-4}.

The results shown in Table 2 are presented in as follows: *Ne* / CR, as defined in (7), where *Ne* is the number of elements used in the representation and CR is the compression ratio. From this table, it is possible to note that sparse representation technique using the proposed dictionary performs better when compared with the WT using the same mother wavelet and the same window length. The better results obtained by the sparse representation is due to the dictionary construction that makes possible to use Fourier elements to represent the stationary parts of the signal, while the WT can represent sparsely the transient parts. A good example to show this capability is the signal showed in Figure 11.

TABLE 2 Number of elements used for the compression techniques

	Wavelet	Sparse
1 cycle	208 / 7:1	125 / 12:1
6 cycles	129 / 12:1	109 / 14:1

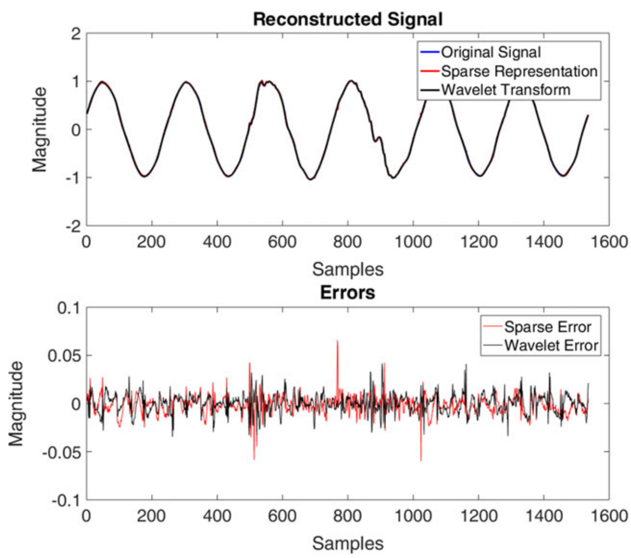

FIGURE 11 Example of a reconstructed signal. The signal was extracted from the IEEE database [30]

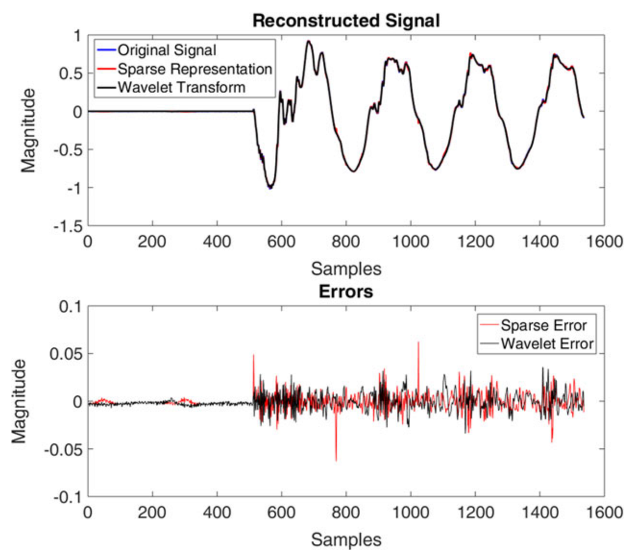

FIGURE 12 Example of a reconstructed signal. The signal was extracted from the IEEE database [30]

This signal is referred to the index 59 in the graph shown in the Figure 10. The sparse technique is better in this case, since few elements of the Fourier basis are necessary to represent the stationary part of the signal while many Wavelet elements would be necessary to represent this part. So the sparsest solution for this case is to use wavelet elements to represent only the transient parts in cycles two and three and Fourier elements to represent the stationary part in all signal.

An example where the sparse representation does not perform better than the WT is shown in Figure 12. This signal is referred to index 26 in the graph shown in Figure 10. In this case, there is a period of interruption followed by a transient, with almost no steady state. In this case, the elements from the Fourier basis do not contribute for the sparsity of the solution, and the wavelet elements are predominant in the the representation.

6 | CONCLUSION

This paper presented a comparison among five sparse representation algorithms in the context of power system signal compression, aiming at determining which one is suitable for a real-time application. The comparison was conducted in terms of the representation quality and the number of elements used. A dictionary formed by the union of the Fourier and wavelet basis was used due to the power system signals characteristics.

Several electrical disturbances were tested; the results enable to conclude that all the algorithms, except StOMP, are capable to represent the signal with good quality but each algorithm needs a different number of iterations and uses a different number of elements to perform the representation. In addition, the results demonstrated that the ROMP algorithm uses the lowest number of elements with lowest number of iterations.

A computational complexity comparison was conducted in terms of required flops, demonstrating that the ROMP needs less flops than the others to reach a better representation quality, despite needing more flops per iteration. Altogether, the results show that ROMP is the most suitable to be applied in real time among all tested algorithms.

The comparison of the CR was conducted between the sparse representation technique using the proposed dictionary and the WT, whose results prove that sparse representation using the proposed dictionary has a better performance than the WT. The authors are working on a field gate programmable array implementation of the algorithm obtaining promising results.

ORCID

Leandro Rodrigues Manso Silva http://orcid.org/0000-0002-6360-0033

REFERENCES

1. Silva L, Kapisch E, Martins C, et al. A gapless power quality disturbance recorder. *IEEE Trans Power Delivery*. 2017;32(2):862-871.
2. Maaß H, Cakmak HK, Bach F, et al. Data processing of high-rate low-voltage distribution grid recordings for smart grid monitoring and analysis. *EURASIP J Adv Signal Process*. 2015;2015(1):1-21.
3. Sayood K. *Introduction to Data Compression*. Amsterdam: Newnes; 2012.
4. Huang S-J, Jou M-J. Application of arithmetic coding for electric power disturbance data compression with wavelet packet enhancement. *IEEE Trans Power Syst*. 2004;19(3):1334-1341.
5. IEEE1159. Ieee recommended practice for the transfer of power quality data. IEEE Std 1159.3-2003; 2004.
6. Zhang D, Bi Y, Zhao J. A new data compression algorithm for power quality online monitoring. In: International Conference on Sustainable Power Generation and Supply; 2009; Nanjing, China.
7. Rebollo-Neira L. Cooperative greedy pursuit strategies for sparse signal representation by partitioning. *Signal Process*. 2016;125:365-375.
8. Panda G, Dash P, Pradhan A, Meher S. Data compression of power quality events using the slantlet transform. *IEEE Trans Power Delivery*. 2002;17:662-667.
9. Huang S-J, Jou M-J. Application of arithmetic coding for electric power disturbance data compression with wavelet packet enhancement. *IEEE Trans Power Syst*. 2004;19(4):1334-1341.
10. Ning J, Wang J, Gao W, Liu C. A wavelet-based data compression technique for smart grid. *IEEE Trans Smart Grid*. 2011;2(1):212-218.
11. Zhang F, Cheng L, Li X, Sun Y, Gao W, Zhao W. Application of a real-time data compression and adapted protocol technique for WAMS. *IEEE Trans Power Syst*. 2015;30(2):653-662.
12. Zhang Z, Xu Y, Yang J, Li X, Zhang D. A survey of sparse representation: algorithms and applications. *IEEE Access*. 2015;3:490-530.
13. Manikandan MS, Samantaray SR, Kamwa I. Simultaneous denoising and compression of power system disturbances using sparse representation on overcomplete hybrid dictionaries. *IET Gener Transm Distrib*. 2015;9(11):1077-1088.
14. Lovisolo L, da Silva EAB, Rodrigues MAM, Diniz PSR. Efficient coherent adaptive representations of monitored electric signals in power systems using damped sinusoids. *IEEE Trans Signal Process*. 2005;53(10):3831-3846.

15. Mallat S, Zhang Z. Matching pursuits with time-frequency dictionaries. *IEEE Trans Signal Process*. 1993.
16. Krim H, Tucker D, Mallat S, Donoho D. On denoising and best signal representation. *IEEE Trans Inf Theory*. 1999;45(7):2225-2238.
17. Gribonval R, Bacry E. Harmonic decomposition of audio signals with matching pursuit. *IEEE Trans Signal Process*. 2003;51(1):101-111.
18. Jaggi S, Karl WC, Mallat S, Willsky AS. High resolution pursuit for feature extraction. *Appl Comput Harmon Anal*. 1998;5(4):428-449.
19. Vera-Candeas P, Ruiz-Reyes N, Rosa-Zurera M, Martinez-Munoz D, Lopez-Ferreras F. Transient modeling by matching pursuits with a wavelet dictionary for parametric audio coding. *IEEE Signal Process Lett*. 2004;11(3):349-352.
20. Van Leeuwen J. *Handbook of Theoretical Computer Science: Algorithms and Complexity*, Vol. 1. Amsterdam: Elsevier; 1990.
21. Elad M. *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. New York: Springer; 2010.
22. Pati Y, Rezaiifar R, Krishnaprasad P. Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition. In: Record of the Twenty-Seventh Asilomar Conference on Signals, Systems and Computers; 1993; Pacific Grove, CA. 40-44.
23. Donoho D, Tsaig Y, Drori I, Starck J-L. Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit. *IEEE Trans Inf Theory*. 2012;58(2):1094-1121.
24. Needell D, Vershynin R. Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit. *Found Comput Math*. 2009;9(3):317-334.
25. Aharon M, Elad M, Bruckstein A. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans Signal Process*. 2006;p. 4311.
26. Eldar YC, Kutyniok G. *Compressed Sensing: Theory and Applications*. Cambridge, UK: Cambridge University Press; 2012.
27. Daubechies I. *Ten Lectures on Wavelets*, Vol. 61. Philadelphia, PA: SIAM; 1992.
28. Coifman RR, Wickerhauser MV. Entropy-based algorithms for best basis selection. *IEEE Trans Inf Theory*. 1992;38(2):713-718.
29. Hamid EY, Kawasaki Z-I. Wavelet-based data compression of power system disturbances using the minimum description length criterion. *IEEE Trans Power Delivery*. 2002;17(2):460-466.
30. IEEE. Ieee pes working group p1433 power quality definitions. 2017 [Online]. http://grouper.ieee.org/groups/1433/. Accessed in 2018.

How to cite this article: Silva LRM, de Andrade Filho LM, Duque CA. Sparse representation algorithm applied to power systems signal compression. *Int Trans Electr Energ Syst*. 2019;29:e2693. https://doi.org/10.1002/etep.2693