A Compressed Sensing Feature Extraction Approach for Diagnostics and Prognostics in Electromagnetic Solenoids

Christian Knoebel[1], Hanna Wenzl[1], Johannes Reuter[1] and Clemens Guehmann[2]

[1] *Institute of System Dynamics, University of Applied Sciences Konstanz, Konstanz, Germany*
cknoebel, hwenzl, jreuter@htwg-konstanz.de

[2] *Chair of Electronic Measurement and Diagnostic Technology, Technische Universität Berlin, Berlin, Germany*
clemens.guehmann@tu-berlin.de

ABSTRACT

One major realm of Condition Based Maintenance is finding features that reflect the current health state of the asset or component under observation. Most of the existing approaches are accompanied with high computational costs during the different feature processing phases making them infeasible in a real-world scenario. In this paper a feature generation method is evaluated compensating for two problems: (1) storing and handling large amounts of data and (2) computational complexity. Both aforementioned problems are existent e.g. when electromagnetic solenoids are artificially aged and health indicators have to be extracted or when multiple identical solenoids have to be monitored. To overcome those problems, Compressed Sensing (*CS*), a new research field that keeps constantly emerging into new applications, is employed. *CS* is a data compression technique allowing original signal reconstruction with far fewer samples than Shannon-Nyquist dictates, when some criteria are met. By applying this method to measured solenoid coil current, raw data vectors can be reduced to a way smaller set of samples that yet contain enough information for proper reconstruction. The obtained *CS* vector is also assumed to contain enough relevant information about solenoid degradation and faults, allowing *CS* samples to be used as input to fault detection or remaining useful life estimation routines. The paper gives some results demonstrating compression and reconstruction of coil current measurements and outlines the application of *CS* samples as condition monitoring data by determining deterioration and fault related features. Nevertheless, some unresolved issues regarding information loss during the compression stage, the design of the compression method itself and its influence on diagnostic/prognostic methods exist.

1. INTRODUCTION

The process of feature extraction (identifying features that represent the health state of the system or component under consideration) is seen as most critical point in data-driven diagnostics and prognostics (Jardine, Daming Lin, & Banjevic, 2006). As the performance of subsequent steps like classification or remaining useful life (*RUL*) estimation is strongly influenced by feature quality (Alkhadafe, Al-Habaibeh, & Lotfi, 2016; Peng, Long, & Ding, 2005), it is necessary to find features that allow differentiation between fault or deterioration states. This is usually achieved by generating features in e.g. time, time-frequency or frequency domain, based on measured data. As these feature sets possibly contain redundant or unnecessary information, a selection has to be performed where only 'meaningful' features are further processed. This reduction step is critical as the amount of necessary training data for classification rises with the number of used features (also known as curse of dimensionality). The whole process of feature extraction and selection is computationally expensive, especially if lots of data has to be analysed. In the application at hand this is for example the case in run-to-failure experiments carried out during development or when multiple identical components have to be monitored, e.g. in a machine. The approach presented in this paper, therefore, tries to improve on two aspects: (a) the amount of data to be handled and stored and (b) the feature extraction and selection step. Step (a) is realised by compressing measured data employing *CS* and step (b) is tackled by performing diagnostics and prognostics directly on compressed samples without additional feature generation.

The remainder of this paper is structured as follows. In section 2 the application and available datasets are briefly explained. Section 3 presents the employed methods and approaches followed by the results in section 4 and a discussion in section 5. The paper concludes with a summary and an outlook on future work.

2. ACTUATOR AND DATA

Figure 1 shows a schematic cross section of an electromagnetic solenoid which is used to transform electric energy via magnetic energy to mechanical energy. By applying a voltage step to the coil, current and therefore magnetic flux build up resulting in an electromagnetic force moving the plunger in the z-direction. Due to back electromagnetic force, caused by magnetic field changes, the plunger movement can be observed in the measured coil current. Especially the end of movement (plunger hits mechanical stop) produces a characteristic spike which is further on denoted as cycle time (time the plunger takes to reach its end position). When the applied voltage is switched off, current and hence the magnetic flux degrade, resulting in a spring driven plunger retraction (magnetic force is exceeded by spring force). During this back and forth movement (denoted as one switching cycle) the solenoid operates e.g. a switch, brake or door lock. As these applications can be safety relevant (elevator brakes, safety switches, etc.) it is desirable to monitor the solenoids health condition and its progression over time. In a run-to-failure experiment

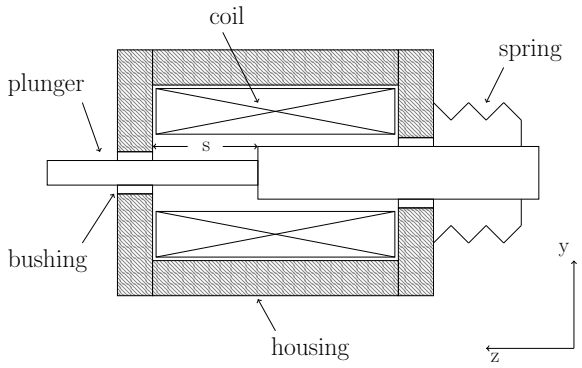

Figure 1. Schematic drawing of an electromagnetic solenoid consisting mainly of: coil, plunger, housing, retaining spring and bushings.

ten solenoids were artificially aged 13 million cycles measuring their coil current every 100 cycles. Using a test-stand, errors were induced influencing plunger movement and supply voltage (see tab. 1). It is assumed that gradual degradation, mainly governed by friction phenomena, and discrete error events can be extracted solely from measured coil current profiles (sampled at $f_S = 10\ kHz$ resulting in $N = 2000$ samples for $t = 0.2\ s$ recording time which corresponds to the *on* time of the solenoid).

Table 1. Emulated fault data with either altered coil supply voltage or reduced stroke. Plunger movement was influenced by employing a mechanical stop.

Mode	Class	Measured Cycles
normal stroke	1	91
reduced supply voltage	2	92
no supply voltage	3	31
reduced stroke 15mm (I)	4	92
reduced stroke 22mm (II)	5	92
no stroke	6	91

3. METHODS

In the following section Compressed Sensing (*CS*) is briefly explained. Furthermore, the usage of *CS* features as input for diagnostic and prognostic methods is outlined.

3.1. Compressed Sensing

Compressed Sensing is a signal processing technique that was first described by (Candes, Romberg, & Tao, 2006) and (Donoho, 2006). The basic idea is to make use of the signal's sparsity in some domains that allows sampling fewer data points than the Shannon-Nyquist theorem dictates. It was shown that the original signal can be reconstructed from these few measurements with high probability using convex optimization methods obeying some constraints and requirements. *CS* received lots of attention due to its applicability in signal and image processing (especially medical imaging) and data compression. During the last two years interesting and promising applications in diagnostics and prognostics kept emerging establishing a new research field (Jayawardhana, Zhu, Liyanapathirana, & Gunawardana, 2017; Liu, Zhang, & Xu, 2017; Wang, Xiang, Mo, & He, 2015).

3.1.1. Compression

Let $x \in \Re^N$ be a k-sparse compressible signal of length N with k non-zero coefficients c and $k \ll N$. Employing a sampling matrix $\Lambda = \Theta\hat{\Lambda}$, that is extracted from $\hat{\Lambda}^{N\times N}$ using a selector matrix $\Theta^{K\times N}$, the *CS* vector (compressed measurements) $y \in \Re^K$ with $K < N$ is given by

$$y = \Lambda x. \tag{1}$$

As most of the signals measured in technical applications are not sparse in time-domain (i.e. *CS* would be not applicable), they have to be expanded in an orthonormal basis, where the necessary sparsity can be achieved (Flandrin & Borgnat, 2010). Here the Wavelet-basis was used to calculate an approximately sparse representation of the signal of interest. Prior to transformation non-stationary signal components are removed by calculating the difference $i_d(t)$ between measured current $i_m(t)$ and a prototype function $f(t, a, b)$ (black curve in Fig. 6).

$$f(t, a, b) = a \cdot (1 - e^{-\frac{t}{b}}) \tag{2}$$

This particular prototype function resembles essentially the inductance step response (solenoid plunger blocked in initial

position) with $a = \frac{V_{cc}}{R}$ and $b = \tau$.

$$i_{step}(t) = \frac{V_{cc}}{R} \cdot (1 - e^{-\frac{t}{\tau}}) \quad \text{with} \quad \tau = \frac{L}{R} \tag{3}$$

Due to the interconnection of inductance L and resistance R, τ can only be identified in its entity. Factor a, however, can be easily calculated. The parameter b was identified using current measurements of a new solenoid, applied supply voltage $V_{cc} = 20V$ and calculated coil resistance $R = \frac{V_{cc}}{i_{end}}$ (with stationary current end value i_{end}). Hence, input to the compression step given in Eq. 1 is not the current signal itself, but the Wavelet transform of $i_s(t)$.

$$i_s(t) = i_{step}(t) - i_m(t) \tag{4}$$

Sparsity of $k = 231$ with an original signal length of $N = 2000$ samples was achieved by a thresholding on the near zero Wavelet-coefficients (which also leads to a denoising of the signal). Using the signal $i_{step}(t)$ as reference is beneficial in three ways: (1) temperature influences and, as already mentioned, non-stationarity are eliminated, (2) all measured signals are compared to the same prototype and (3) it is the worst considerable scenario, since a blocked solenoid plunger can't fulfil its intended task.

3.1.2. Reconstruction

Output of the reconstruction procedure are estimated Wavelet-coefficients $\hat{x}$ that can be used to perform an inverse Wavelet-transform reconstructing $\hat{i}_s(t)$. As $i_{step}(t)$ is known, an estimate of the measured coil current $\hat{i}_m(t)$ can be calculated. The under determined recovery problem $\Lambda\hat{x} = y$ can be solved employing vector norm minimization (ℓ_1-minimization) as shown by (Candes et al., 2006; Donoho, 2006)

$$\underset{\hat{x} \in \Re^N}{min} \parallel \hat{x} \parallel_1 \quad subject\ to \quad \Lambda\hat{x} = y. \tag{5}$$

When x meets the aforementioned criteria and Λ the subsequent ones, reconstruction will be successful with high probability.

3.1.3. Sensing Matrices

The sampling matrix Λ has to meet certain criteria to ensure reconstruction of the original signal x. I.e. the matrix should be designed such that there exists only one *CS* measurement vector y to be matched to at most one x. The *spark* of a sampling matrix Λ is denoted as one key property. It is defined as the smallest number of columns in Λ that are linearly dependent. If Eq. 6 holds, reconstruction can be guaranteed.

$$spark(\Lambda) > 2k \tag{6}$$

But as calculation of $spark(\Lambda)$ is NP-hard, its computation is not feasible for arbitrary sampling matrices. Another key property is matrix coherence $\mu(\Lambda)$ that is defined as the largest inner product of any two columns in Λ.

$$\mu(\Lambda) = \max_{i \neq j} | < \lambda_i, \lambda_j > | \tag{7}$$

The maximum signal sparsity is then determined by:

$$k < \frac{1}{2}\left(1 + \frac{1}{\mu(\Lambda)}\right). \tag{8}$$

As under noisy conditions (measurement noise $y = \Lambda x + \delta$ or noise due to numerical instabilities $\hat{\Lambda} = \Lambda + \Delta$) the aforementioned properties do not hold any more, the Restricted Isometry Property (*RIP*) has to be used to check and ensure stable reconstruction (Candes & Tao, 2005).

$$(1 - \alpha)\|x\|_2^2 \leq \|\Lambda x\|_2^2 \leq (1 + \alpha)\|x\|_2^2 \tag{9}$$

Again, the *RIP* is computationally expensive to calculate for arbitrary sampling matrices. For random Toeplitz-type and random-value-type sampling matrices research results exist showing that both types fulfil the aforementioned properties with high probability allowing reconstruction of the original signals (Haupt, Bajwa, Raz, & Nowak, 2010; Duarte & Eldar, 2011). Hence, they will be used for compression of the Wavelet transformed current measurements.

3.1.4. Optimal Matrix Size *K*

To find an optimal sensing matrix size, i.e. determine the necessary amount of *CS* samples to reconstruct coil current of new as well as deteriorated solenoids, reconstruction was performed with different matrix sizes K and types (random value and random Toeplitz-type) using Monte-Carlo simulations. For every reconstruction step the sum of squared errors (*SSE*) between measured and reconstructed coil current was evaluated.

$$\epsilon(K) = \sum (i_m - \hat{i}_m(K))^2 \tag{10}$$

3.2. Using *CS* Samples as Features

To overcome the problem of massive data amounts occurring during run-to-failure test or when monitoring many identical devices, only *CS* samples of the original signal are stored and not the raw signal itself. As the *CS* data vector can be multiple times smaller not only less space is required but computational costs are also reduced. Reconstruction is performed, whenever desired or necessary. To avoid a separate feature extraction for condition monitoring and remaining useful life estimation (*RUL*) and to make use of the information contained in the *CS* samples they are directly re-utilised as features. I.e. the suitability of *CS* samples for fault detection (do features exist, that allow fault discrimination) and for extraction of wear-related patterns (do features exist that show deterioration and life time dependent behaviour) is evaluated.

For now no further post-processing is performed on the compressed values. Whether such techniques have to be applied to enhance fault separability or wear features, depends on the proceeding results.

4. RESULTS

In the following section compression and reconstructions results for different solenoid ageing stages, sensing matrix types and sizes are given. Furthermore, the usage of *CS* data as features for diagnostic and prognostic methods is outlined.

4.1. Compression Results

Figure 2 shows $K = 250$ *CS* samples of a new and an aged solenoid where compression was performed using a random type matrix. It can be seen how feature values differ from their new-state after approximately $6 \cdot 10^6$ cycles.

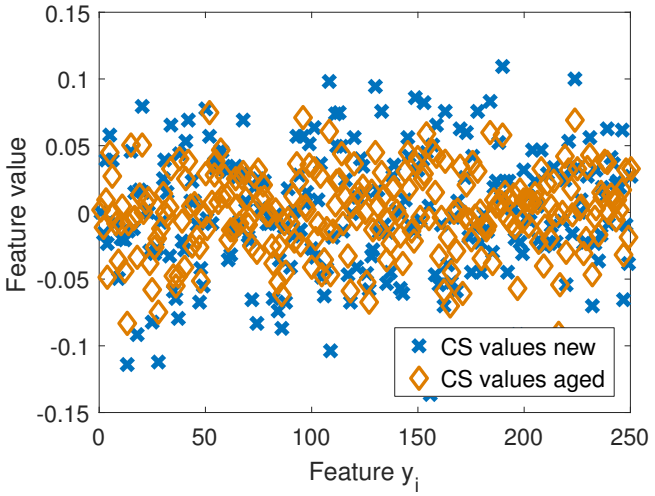

Figure 2. Comparison of $K = 250$ *CS* samples of an aged and a new solenoid.

Normalized trajectories for *CS* samples $y_{50,75,97,108}$ and the changes they undergo during solenoid lifetime are given in Fig. 3. They are compared to the solenoids cycle time t_s that is used as a deterioration indicator (if t_s exceeds a certain threshold for more than ten consecutive measurements, the solenoid is denoted as deteriorated). This particular solenoid failed at approximately $5 \cdot 10^6$ cycles trespassing a cycle time threshold of $t_s = 90ms$. Line breaks are caused by malfunctioning of the used test stand causing incomplete data. It should be noted here, that unlike cycle time, *CS* features start approximately with the same amplitude as before the interruption at about $4.4 \cdot 10^6$ cycles. Figure 4 shows *CS* values $y_{j=24,\ i=20}$ spanning a two dimensional feature space with fault classes 1 to 6. Compression of the current profiles was performed in the same way as described before.

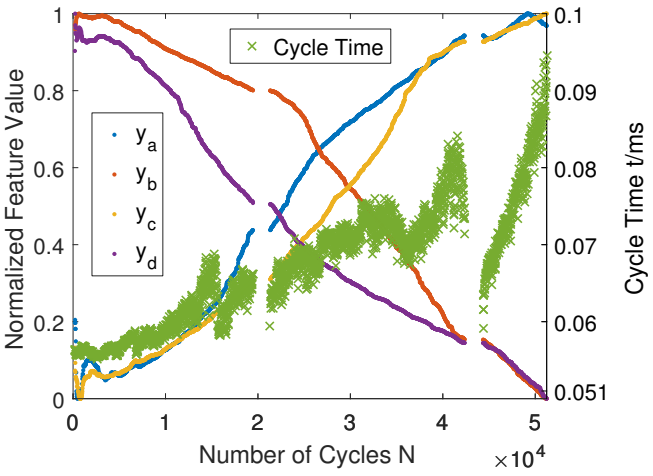

Figure 3. Feature trajectories $y_{a,b,c,d}$ over life time in comparison with cycle time t_s that serves as a deterioration indicator. Gaps in the data were caused by a malfunctioning test stand.

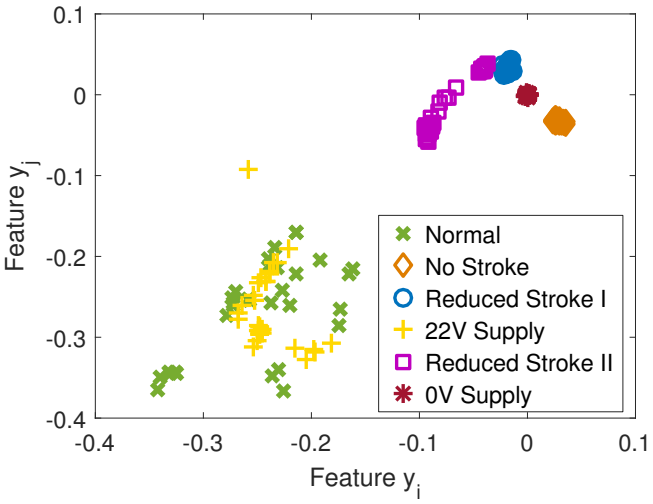

Figure 4. Clustering and class separation of *CS* samples $y_{j=24,\ i=20}$ in two dimensional feature space.

4.2. Optimal Matrix Size *K*

Figure 5 shows the *CS* reconstruction errors and their dependency on sensing matrix size, type and solenoid age. Compression and reconstruction steps were performed for consecutive coil current measurements of one solenoid at three deterioration states. Mean and standard deviation (denoted by error bars) are given for each matrix size K that was evaluated. Reconstruction errors decrease with increasing *CS* samples K for both sensing matrix types converging to error minima for $K_R > 200$ and $K_T > 500$, respectively.

4.3. Reconstruction Results

Figure 6 shows measured coil current of a deteriorated and a new solenoid, their respective reconstructions from *CS* sam-

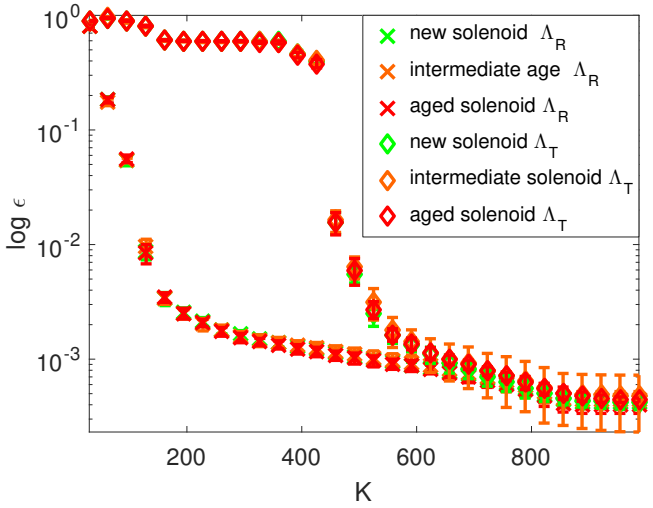

Figure 5. Reconstruction error using a random value sensing matrix Λ_R and a Toeplitz-type sensing matrix Λ_T of increasing size (i.e. taking increasing amounts of *CS* samples). Reconstructions were performed for randomly picked current profiles for each matrix size.

ples, the reference curve as explained in 3.1.1 and the differences between reference and measurement. Compression as well as reconstruction was performed using a random-value sampling matrix with $K = 250$. Reconstruction and original signal are in good accordance with each other for both deterioration stages. I.e. no artefacts are superimposed on the signal that might result from too few *CS* measurements or an inappropriate selection of the sensing matrix.

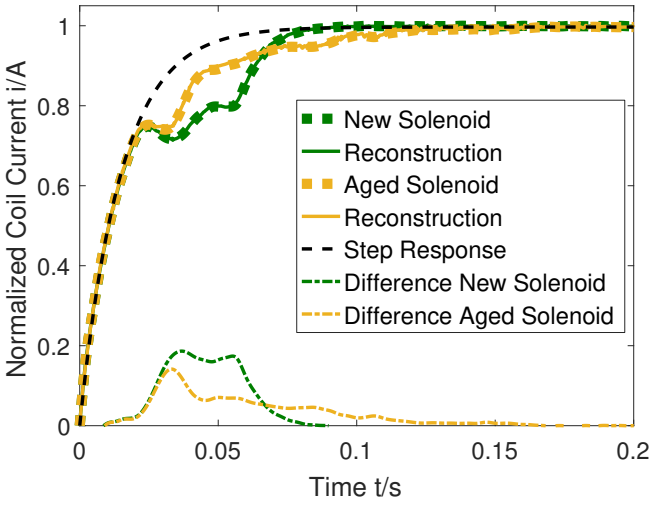

Figure 6. Comparison of reconstructed and measured coil currents of a new and a worn out solenoid. The used reference $i_{step}(t)$ (inductance step response, see section 3.1.1) is displayed as well.

5. DISCUSSION

Figure 2 shows a clear differentiation between a new and an aged solenoid in feature space. By visual examination of the single features it can be seen that some have a higher deviation from their initial state than others and that features tend to drift towards the abscissa with increasing age. This behaviour can be observed for all ten solenoids tested. Nevertheless, feature movement towards end of life is of major interest. Four examples of such feature movements in comparison with the solenoids cycle time are given in Fig. 3. It can be seen, that after some initial oscillations the features start to move nearly monotonically. Missing data was caused by malfunctions of the test stand, where either the solenoids did not toggle or data was not recorded properly. An interruption of solenoid actuation causes the coil to cool down influencing its current profile and thus its cycle time. Once the solenoid heats up again, one would expect the cycle time to approach its previous level again. But, on closer inspection of Fig. 3 it can be seen, that at approximately 4.3 million the cycle time does not continue at a similar level as before the interruption (at approximately 2 million cycles it does).

The *CS* features seem to be invariant and more robust to such disturbances and commence at nearly the same value after an interruption. However, not all features undergo the same changes while approaching the solenoids end of life (*EOL*). Some are correlated with cycle time or show non-monotonic behaviour, hamper additional insight into underlying degradation patterns. To overcome this problem, several approaches might be considered: (a) employing feature selection and rating techniques as mentioned in 3.2 and (b) using data and domain knowledge to enhance the compression step. Approach (a) is state-of-the-art and might work out the best wear dependent features but impose additional computational costs whereas approach (b) improves on the method itself by using structured sensing matrices and more suitable sparsifying transforms to make compression and feature generation more effective and ideally features more expressive.

As it was shown in (Knöbel, Marsil, Rekla, Reuter, & Gühmann, 2015) for the same application, post-processing of fault related features enhances class separation and clustering. Looking at Fig. 4 it can be observed that *CS* based features (here simply two *CS* values y_{20} and y_{24} were selected) already show good clustering and class separability without the need for additional post-processing. The overlapping of class 1 (normal stroke) and class 4 (22V supply voltage) arises from using nearly the same voltage level the coil measurements were acquired with. I.e. at 22V the solenoid works as intended and the slightly lowered voltage can't be detected using only coil current measurements (this particular overlapping of class 1 and 4 exists in all *CS* data features). Fig. 5 shows coil current reconstruction error plotted over *CS* samples employing a random value and a Toeplitz-type sensing matrix for three different solenoid ageing stages. For

both matrix types all three error curves (new, intermediate, aged) are superimposed (showing slightly different standard deviations) indicating that for new solenoids the same compression matrix size can be used as for worn out ones. Furthermore, all curves converge to a lower error bound for $K_{A,T} \gtrapprox 600$, but lower reconstruction errors can be achieved with way fewer *CS* samples for the random value matrix compared to the Toeplitz-type matrix. The graph clarifies in a quite intuitive way, how not only matrix size but also its type influences the performance of *CS*. For the successful current reconstruction shown in Fig. 6 $K_R = 250$ *CS* samples were used which seemed to be a good compromise between matrix size and error. This matrix size is also in good agreement with the achieved sparsity of $k = 231$ (i.e. more *CS* measurements were taken than necessary). Nevertheless, reconstruction errors and losses as well as alternative methods to the employed ℓ_1 minimisation have to be quantified and evaluated.

6. CONCLUSION AND FUTURE WORK

Focus of this paper is the application of Compressed Sensing (*CS*) as a feature extraction approach and the evaluation if and how *CS* features are suitable for diagnostics and prognostics. As results show, compressed coil current measurements contain enough information about wear phenomena and faults to show life time dependent degradation patterns and good clustering regarding different fault cases. The dual and concurrent usage of extracted *CS* data as deterioration/fault indicators as well as for compression constitutes a promising approach in the field of condition based maintenance and provides a solution for both problems outlined in section 1: (a) data handling and (b) feature generation. In future work it has to be evaluated how compression and thus feature extraction can be enhanced by better adapting the compression step to the data. This could either be achieved by applying structured sensing matrices or by appropriate post-processing of the obtained *CS* features. Furthermore, errors occurring during the sparsifying transform and the compression step as well as their influence on diagnostic and prognostic processes have to be quantified and handled accordingly. With regard to concepts like Industry 4.0 and Internet of Things, *CS* is becoming even more interesting. Not only data transfer rates can be kept low by sending compressed rather than raw data, but also refining fault detection and remaining useful life prediction based on fielded assets and components is possible. Ever growing research interest led to vast improvements of the method itself enabling *CS*to emerge to new application fields in the past years. Especially, new developments regarding the aforementioned structured sensing matrices offers significant potential.

ACKNOWLEDGMENT

The support of this work by the German Federal Ministry of Education and Research under grant 03FH034PX3 is gratefully acknowledged.

REFERENCES

Alkhadafe, H., Al-Habaibeh, A., & Lotfi, A. (2016). Condition monitoring of helical gears using automated selection of features and sensors. *Measurement*, *93*, 164–177.

Candes, E. J., Romberg, J., & Tao, T. (2006). Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, *52*(2), 489–509.

Candes, E. J., & Tao, T. (2005). Decoding by linear programming. *IEEE Transactions on Information Theory*, *51*(12), 4203–4215. doi: 10.1109/TIT.2005.858979

Donoho, D. L. (2006). Compressed sensing. *IEEE Transactions on Information Theory*, *52*(4), 1289–1306.

Duarte, M. F., & Eldar, Y. C. (2011). Structured compressed sensing: From theory to applications. *IEEE Transactions on Signal Processing*, *59*(9), 4053–4085.

Flandrin, P., & Borgnat, P. (2010). Time-frequency energy distributions meet compressed sensing. *IEEE Transactions on Signal Processing*, *58*(6), 2974–2982.

Haupt, J., Bajwa, W. U., Raz, G., & Nowak, R. (2010). Toeplitz compressed sensing matrices with applications to sparse channel estimation. *IEEE Transactions on Information Theory*, *56*(11), 5862–5875.

Jardine, A., Daming Lin, & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, *20*(7), 1483–1510.

Jayawardhana, M., Zhu, X., Liyanapathirana, R., & Gunawardana, U. (2017). Compressive sensing for efficient health monitoring and effective damage detection of structures. *Mechanical Systems and Signal Processing*, *84, Part A*, 414–430.

Knöbel, C., Marsil, Z., Rekla, M., Reuter, J., & Gühmann, C. (2015). Fault detection in linear electromagnetic actuators using time and time-frequency-domain features based on current and voltage measurements. In IEEE (Ed.), *Methods and models in automation and robotics (mmar), 2015 20th international conference on methods and models in automation and robotics.*

Liu, Y., Zhang, G., & Xu, B. (2017). Compressive sparse principal component analysis for process supervisory monitoring and fault detection. *Journal of Process Control*, *50*, 1–10.

Peng, H., Long, F., & Ding, C. (2005). Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE transactions on pattern analysis and machine intelligence*, *27*(8), 1226–1238.

Wang, Y., Xiang, J., Mo, Q., & He, S. (2015). Compressed sparse time–frequency feature representation via compressive sensing and its applications in fault diagnosis. *Measurement*, *68*, 70–81.