LA-UR- 12-00709

Approved for public release; distribution is unlimited.

Title:	The application of compressed sensing to long-term acoustic emission-based structural health monitoring
Author(s):	Alessandro Cattaneo, Politecnico di Milano Gyuhae Park, LANL, INST-OFF Charles R. Farrar, LANL, INST-OFF David L. Mascareñas, LANL, INST-OFF
Intended for:	SPIE Smart Structures/NDE, San Diego, California March 11-15, 2012

Los Alamos
NATIONAL LABORATORY
EST. 1943

Form 836 (7/06)

The application of compressed sensing to long-term acoustic emission-based structural health monitoring

Alessandro Cattaneo*[a], David Mascareñas[b], Gyuhae Park[b], Charles Farrar
[a]Politecnico di Milano, Department of Mechanics, Via La Masa 1, Milan, 20156 (MI), Italy;
[b]Los Alamos National Laboratory Engineering Institute, P.O. Box 1663, T001, Los Alamos, NM 87545, USA 505-663-5227

ABSTRACT

The acoustic emission (AE) phenomena generated by a rapid release in the internal stress of a material represent a promising technique for structural health monitoring (SHM) applications. AE events typically result in a discrete number of short-time, transient signals. The challenge associated with capturing these events using classical techniques is that very high sampling rates must be used over extended periods of time. The result is that a very large amount of data is collected to capture a phenomenon that rarely occurs. Furthermore, the high energy consumption associated with the required high sampling rates makes the implementation of high-endurance, low-power, embedded AE sensor nodes difficult to achieve. The relatively rare occurrence of AE events over long time scales implies that these measurements are inherently sparse in the spike domain. The sparse nature of AE measurements makes them an attractive candidate for the application of compressed sampling techniques. Collecting compressed measurements of sparse AE signals will relax the requirements on the sampling rate and memory demands. The focus of this work is to investigate the suitability of compressed sensing techniques for AE-based SHM. The work explores estimating AE signal statistics in the compressed domain for low-power classification applications. In the event compressed classification finds an event of interest, l_1 norm minimization will be used to reconstruct the measurement for further analysis. AE measurements often suffer from the presence of noise sources with band-limited random and/or harmonic contents (e.g. rotating machinery). The suitability of Justice Pursuit to remove unwanted signal components is investigated.

Keywords: acoustic emission, embedded sensing, compressed sampling, damage detection

1. INTRODUCTION

The development of structural health monitoring (SHM) solutions plays a leading role in increasing life-safety of structural and mechanical systems and, at the same time, in reducing economic consequences related to failures. These solutions aim to detect the damage, possibly at its early stages, before it could seriously compromise a particular component functionality or even the entire system integrity. Among the various techniques used for SHM purposes, acoustic emission (AE) monitoring has made it possible to detect evolutionary defects.[1] The physical phenomena behind this technique consists in the release of energy in the form of transitory elastic waves. Surface vibration induced by the elastic waves are then collected by a piezoelectric sensor. Acoustic emission phenomenon (AE) is quite complex as far as acoustic wave propagation in solids involves a series of aspects difficult to model. Propagation of multiple modes with different velocities, reflection, refraction, mode conversion, velocity dispersion and attenuation[2] make the extraction of damage-sensitive features from the measured signal an anything but trivial process. Notwithstanding these undeniable difficulties, several studies[3,4] and practical implementations[5,32] have demonstrated that the appearance of a crack in a material put under certain stress condition, or the growth of a pre-existing crack, cause the emission of acoustic wave and thus AE can be ultimately used to monitor the integrity of structures. From an historical point of view, first studies from 1930s to 1950s had to face with the limited instrumentation capabilities and measurements were made with RMS voltmeters or simple event counters. There was no possibility to record high frequency transient AE waveforms and the monitoring were made by measuring the AE "activity", that is by estimating the number of occurred acoustic events. In the attempt to collect more information, feature based AE systems able to provide parameters for each individual discrete AE signal began to be used. Such parameters are namely arrival time, peak-amplitude, duration, rise time, counts, energy and average frequency content. Measurement system based on AE parameters are still used because they are able to measure and record data at rates of thousands of events/second.[5] It is even worth noticing that approaches for analyzing both AE "activity" and feature base AE data are still used for AE monitoring purposes. As a natural consequence of the

continuous development of measurement hardware capabilities, systems able to record AE waveforms became available in the 1980s. Waveform analysis is appealing because allow more accurate arrival time determination, and consequently a better source location, along with improvements in noise discrimination and source characterization.[5] Currently off-the-shelf sensing technology can offer 18-bit depth resolution, 40 MSamples/sec analog to digital converter (ADC) capable of collecting continuous recording of AE waveforms at up to 10 MSamples/sec rate. This high-performance hardware leads to large data storage demands and high power consumption.
The authors of this work focus on the new possibilities introduced by a novel sensing/sampling technique which goes under the name of compressed sensing (CS).[6,7] Compressed sensing, also known as compressive sampling, tackles the problem of recovering a signal from a sequence of time samples by leveraging properties other than those traditionally exploited by conventional approaches to sampling signals.[8] As a matter of fact standard analog-to-digital converter (ADC) technology exploits the well-known Shannon-Nyquist theorem according to which continuous-time band-limited signals can be stably reconstructed by sampling at a rate at least twice the maximum frequency present in the signal (i.e. Nyquist rate). Recent developments in the statistics community have shown however that the Nyquist-rate criteria is not required for signals that enjoy a high degree of sparsity. The property of sparsity, as it will be clarified in this paper, is related to the concept of signal representation. Probably the most well-known signal representation is the one provided by the Fourier Transform. A signal is said to be sparse if the sorted coefficients used for its representation –for example the FFT coefficients,– decay quickly.[9] Over the years great efforts have been spent in order to provide representations as meaningful as possible for a huge variety of signals. In this attempt solutions like Short Fast Fourier Transform (STFT), wavelet transform, principal component analysis (PCA), and the modern concept of over-complete dictionaries[11] were developed. These solutions made it possible to highlight salient features of the signal (useful for recognition purposes), efficiently separate signal and noise (denoising) and capture large portions of the signal with only a few coefficients (compression).[9] These results demonstrate that a traditional data acquisition process based on the Nyquist criteria is inherently wasteful because it requires high performance hardware to capture the original analog signal. This problem becomes most acute for capturing wideband signals,[8] that only feature a few distinctly non-zero Fourier coefficients. By exploiting sparsity, compressed sensing offers the ability to sensibly reduce the number of samples relative to the traditional approach while still allowing the ability to capture the essential information brought by the analog signal. This paper illustrates the potential benefits offered by compressed sensing in a setting representative of AE structural health monitoring. The following research builds on prior work by the authors that demonstrated the feasibility of compressed sensing for structural health monitoring applications[33]. Section 2 provides the necessary theoretical background in order to understand how CS works. Section 3 focuses on the implementation of the over-complete dictionary used to represent AE signals. Section 4 briefly reviews the concept of the compressed estimator and the compressed classifier which are used to perform some basic signal analysis directly in the compressed domain. Section 5 describes two different sampling models to take into account the effect of sparse noise both when it is already present in the signal prior to sampling and when it is added to each sample as result of the sampling process. Finally section 6 shows the test setup and the result obtained. Section 7 sums up the conclusions and provide directions for future work.

2. COMPRESSED SENSING

The term compressed sensing (CS) denotes a sampling technique capable of capturing the useful information content embedded in a signal while at the same time condensing it into a small amount of data. For this reason CS is a protocol for both sensing and compressing data simultaneously. The design of systems capable of directly collecting compressive measurements of analog signals lies outside the scope of this paper and only a brief overview of the architecture for a compressed sensing analog front-end will be provided in Section 5 in order to better comprehend the effect of the noise added on measurements during the sampling process. This section introduces CS theory. In the attempt to explain how CS works, a generic signal x will be modeled as a finite-length, discrete-time vector.[14] A generic signal has to be considered as it is the result of a traditional sampling process under the assumption that the Shannon theorem is satisfied. Mathematically speaking the signal f is a vector in an n-dimensional Euclidean space, denoted by $\mathbb{R}^N$, i.e. $f \in \mathbb{R}^N$. Compressed sensing behaves as a process that, by collecting a number of linear measurement M typically much smaller than N, is still able to accurately recover the original signal.[12] The acquisition of M linear measurements is represented as the product between an $M \times N$ sensing matrix Φ and the signal vector f

$$y = \Phi f \tag{1}$$

where $y \in \mathbb{R}^M$ is the vector of the compressive measurements and the matrix Φ represents adown projection to a lower dimensional space, i.e. it maps $\mathbb{R}^N$ into $\mathbb{R}^M$ with generally $M \ll N$.[14] The linear measurements extracted through the

sensing matrix Φ are required to ensure that the information embedded in the signal f is preserved when $\mathbb{R}^N$ is mapped into $\mathbb{R}^M$. A detailed study to tackle this problem and explain what conditions the sensing matrix Φ has to satisfy can be found in several works.[14,22,23,24,25,26] Intuitively, the sensing matrix Φ, must be incoherent with the basis in which the signal f has a sparse representation. The results showed in this paper are obtained using a sensing matrix Φ populated with i.i.d. Gaussian entries. In this regard it is worth pointing out that the reconstruction of the original signal $f \in \mathbb{R}^N$ from the compressive measurement $y \in \mathbb{R}^M$ is an ill-posed problem.[12,13] The number M of available measurements is smaller than the dimension N of the signal f, there are potentially infinitely many candidate signals $\tilde{f}$ compatible with the set of compressive measurement, i.e. for which holds $\Phi\tilde{f} = y$. At this point we recognize that that in many cases what is really desired is the solution for which f is most sparse in some representation. This new constraint can be thought of as an application of Occam's Razor. Restated, the goal is to find the simplest solution that describes the data. The reconstruction of the full length signal f starting from what appears to be an incomplete set of measurements is made possible thanks to the property of sparsity. This property, that plays a key role in CS theory, can be described introducing the concept of signal representation. A signal can be analytically represented by mean of a linear combination of elementary waveforms. Such a decomposition[13] is expressed as follow

$$f = \sum_{\gamma \in \Gamma} \alpha_\gamma \psi_\gamma \tag{2}$$

where $(\psi_\gamma)_{\gamma \in \Gamma}$ is a collection of waveforms, with γ a parameter, and α_γ are the coefficients of the decomposition. The collection of waveform is referred to as a dictionary Ψ and each waveform stored in its columns represents an atom of the dictionary,[9] i.e. $\Psi = [\psi_1\ \psi_2 \dots \psi_L] \in \mathbb{R}^{N \times Q}$, with $Q \geq N$.

The property of sparsity is leveraged by CS in order to solve the ill-posed signal recovery problem in the sense that, among the infinite set of candidate signals $\tilde{f}$ satisfying $\Phi\tilde{f} = y$, it seeks the solution that is most sparse[18] The possibility to exploit sparsity has been historically limited by the non-convex, combinatorial nature of the minimum l_0 norm optimization problem that must be solved in order to directly obtain the sparsest solution. The big breakthough of the compressed sensing community was the discovery that the convex of l_1 norm relaxation of the of l_0 norm problem can be used to find the sparsest solution under a wide range of circumstances.[11,13] The problem of recovering a signal by exploiting its sparsity can be formulated by introducing the concept of l_0 norm. The l_0 norm is a function $\|\cdot\|_0$ that simply counts the number of nonzero elements of a vector. The problem of finding the sparsest representation possible in a dictionary Ψ can thus be formally written down as follow

$$P_0: \min_{\alpha} \|\alpha\|_0 \quad s.t. \quad y = \Phi\Psi\alpha \tag{3}$$

where α is the collection of the decomposition coefficients defined as $\alpha = (\alpha_\gamma)$. The problem statement (3) is a combinatorial optimization problem that in general is computationally expensive to solve.[14] Its solution requires the enumeration of all possible k-element collections of Ψ, for $k = 1,2, \dots, N$ looking for the smallest collection permitting representation of the signal.[15] Due to the impossibility of directly searching for the solution of the problem, other approaches are required to be used to solve the problem. A number of greedy algorithms[11,14] have been proposed to find a sparse solution to this problem including the method of frames (MOF), Matching Pursuit (MP), and best orthogonal basis (BOB). The basis-pursuit (BP)[13] approach finds the sparsest decomposition of a signal into a dictionary by replacing $\|\cdot\|_0$ of statement (3) with its convex relaxation $\|\cdot\|_1$:[14]

$$P_1: \min_{\alpha} \|\alpha\|_1 \quad s.t. \quad y = \Phi\Psi\alpha \tag{4}$$

The l_1 norm function $\|\cdot\|_1$, defined for a general vector x occupying $\mathbb{R}^N$ as $\|x\|_1 = \sum_{i=1}^{N} |x_i|$, acts as a sparsity-promoting function, meaning that among all the possible decompositions it leads to obtain the one whose sorted coefficients decay quickly.[9,12,14] A geometric interpretation useful to understand l_1 norm function properties is provided in the following works[14,21] It is worth noticing that contrary to greedy algorithm strategy, that chooses at each iteration a waveform ψ_γ of the dictionary best adapted to reduce the square-error of the signal model,[11,27] BP solves the reconstruction problem in a global way because it minimizes a cost function, i.e. $\|\cdot\|_1$, that is defined over all the coefficients associated to the waveforms ψ_γ of the entire dictionary. It is interesting to note that problem statement (4) can be cast as a linear programming (LP) problem and solved by a variety of methods[14] including interior-point algorithms. These algorithms are of particular interest because, in the recent years, they have made it practical to solve large-scale linear programs as as typically found in implementations of BP problems.[13] Interior-point solvers are actually supported by the package for specifying and solving convex programs[19,20] used by the authors to perform the l_1 norm reconstructions collected in this paper. The measurement model (1) and the optimization problem P_1 of equation (4) will

be further discussed in Section 5 where the effects of noise will be taken into account. The next section deals with the problem of finding a dictionary able to sparsify signals acquired in basic AE experimental tests.

3. OVERCOMPLETE DICTIONARY FOR SPARSE REPRESENTATION

In the era of digital signal processing (DSP) several representations have been developed in the attempt to model in a simple and effective way different class of signals. A signal, as mentioned in the previous section, can be modeled by the superposition of elementary waveforms, called atoms, collected in a dictionary.[9] In order to provide an elementary example, equation (5) describes the dictionary associated to the Fourier transform, i.e. a set of orthogonal waveforms defined as

$$\mathcal{F} = \{\psi_n(x) = e^{inx}\}_{n\in\mathbb{Z}} \tag{5}$$

It clearly appears that the Fourier basis $\mathcal{F}$ consists of a collection of stationary waveforms. As consequence of the lack of localization in time, the representation of transitory signals, though quite feasible, requires the use of several atoms. The borderline case is represented by the ideal impulse whose representation in the frequency domain is obtained only through the totality of the atoms collected in $\mathcal{F}$. The possibility to exploit CS to size AE events requires to find a dictionary capable to model their inner transient nature through a sparse representation. Intuitively the dictionary must have atoms that are highly similar to the structures in the data. In this regard an analytical model capable to provide a time-frequency localization of signals is offered by Gabor atoms.[11] The discrete Gabor atom is defined over a vector n as

$$g_\gamma(n) = g_s(n-p)e^{i(2\pi k/N)n} \tag{6}$$

where $\gamma = (s, p, 2\pi k/n) \in \Gamma$ is the parametric set respectively defining scale, time translation and frequency modulating of each atom. In its turn g_s is the normalized Gaussian $g(t)$ uniformly sampled and periodized over N points in order to obtain, at any scale s, a discrete and periodic signal:

$$g_s(n) = \frac{K_s}{\sqrt{s}}\sum_{p=-\infty}^{p=+\infty} g\left(\frac{n-pN}{s}\right) \tag{7}$$

where the constant K_s is introduced to normalize the discrete norm of g_s. The normalized Gaussian $g(t)$, finally, is a continuous function defined as

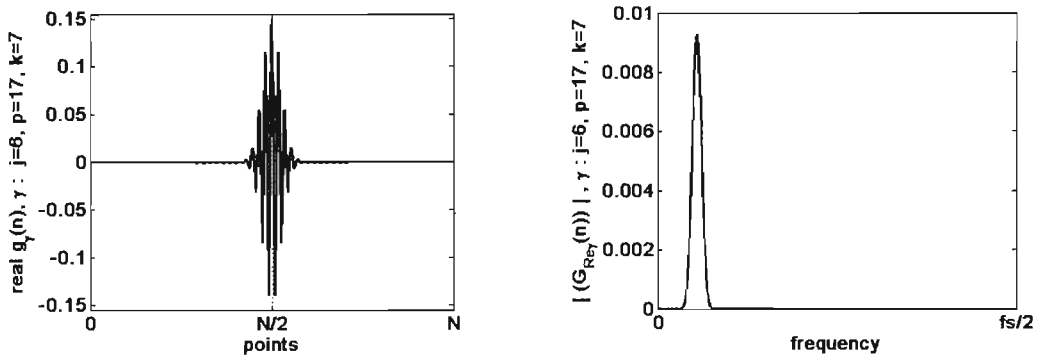
$$g(t) = 2^{1/4}e^{-\pi t^2} \tag{8}$$

It is worth to notice that, according to the Heisenberg uncertainty principle,[28] the normalized Gaussian defined in (8) minimizes the area of the rectangle representing the time-frequency localization of an atom, i.e. the area where is at most concentrated the energy of the atom in the time-frequency plane.[29] The discrete complex Gabor dictionary is defined as the set of all the atom $g_\gamma(n)$ parameterized by $s \in]1, N[$ and p, k integers between 0 and N. This highly redundant dictionary can be limited to a subdictionary indexed by a parametric subset of Γ, i.e. $\gamma \in \Gamma_\alpha \subset \Gamma$.[11] The subset Γ_α is composed of all $\gamma = (a^j, pa^j\Delta u, ka^{-j}\Delta\xi)$ with $a = 2$, $\Delta u = 1/2$, $\Delta\xi = \pi$, $0 < j < log_2 N$, $0 \le p < 2N^{-j+1}$ and $0 \le k < 2^{j+1}$. The subset Γ_α is thus obtained by applying a dyadic sampling to the original set Γ: consecutive values of the discrete scales as well as the corresponding time translation and frequency modulating intervals differ by a factor of two.[30] For the sake of clarity, in Table 1 it is showed the subset Γ_α parameterizing the atoms used to decompose a generic 1024-point length signal. In this case the resulting discrete Gabor dictionary is 36 times overcomplete with an overall dimension equal to $N \times Q = 1024 \times 36864$.

Table 1. Parametric subset Γ_α for a 1024-point length vector

index j (scale)	1	2	3	4	5	6	7	8	9
index p (time translation)	0:1: 1023	0:1: 511	0:1: 255	0:1: 127	0:1: 63	0:1: 31	0:1: 15	0:1: 7	0:1: 3
index k (frequency modulating)	0:1: 3	0:1: 7	0:1: 15	0:1: 31	0:1: 63	0:1: 127	0:1: 255	0:1: 511	0:1: 1023

To illustrate the time-frequency localization property of a generic atom, Figure 1(a) and 1(b) represent, both in the time and frequency domain, the real part of an atom $g_\gamma(n)$ with unitary discrete norm, at scale $j = 6$, time index $p = 17$ and frequency modulating index $k = 7$.



(a) (b)

Figure 1. Real part of a discrete unitary norm atom $g_\gamma(n)$ (a) and absolute value of the associate spectrum $|G_{Re\,\gamma}(n)|$ (b), where fs/2 is the Nyquist frequency, N=1024, parameter set $\gamma : j = 6, p = 17, k = 7$

Clarified how it is defined the discrete Gabor dictionary used for this paper and showed the intrinsic time-frequency localization property of its atoms, this section provides some further information about the dictionary effectively implemented to perform the l_1 norm reconstruction of the signals collected during some basic AE tests. From one hand the use of the discrete Gabor atoms is intended to represent the transient nature of the AE events, from the other hand it is necessary to populate the dictionary with other atoms able to model the possible presence of different structures affecting the collected signals. For this purpose are added the atoms of both the oversampled Discrete Fourier Transform (DFT) and the Dirac basis.[17] The former aims to take into account the effects of stationary components (i.e. steady process like vibrations induced by rotating machinery), the latter the effects of spikes lacking in specific harmonic structures (i.e. signal bursts). The oversampled DFT is defined by sampling the frequencies over equally spaced intervals smaller than those used in the standard DFT, yielding to the following expression

$$\psi_k(n) = \frac{1}{\sqrt{N}} e^{-2\pi i k n / NT} \tag{9}$$

where N is the length of the signal vector n, with $0 \leq n$, T the oversampling factor and $k \leq NT - 1$. The oversampled DFT aims to achieve a sparsest representation for a wide class of stationary signals. As a matter of fact the standard DFT, i.e. the transform (9) defined for $T = 1$ and $k \leq N - 1$, provides a sparse representation only for those signals, a small amount indeed, which are a superposition of the sinusoids collected in the standard DFT dictionary. As regards the Dirac basis, it is simply defined as the identity matrix I.

The overall dictionary Ψ assembled to perform l_1 norm reconstructions carried out in this paper is a 42 times overcomplete dictionary, i.e. with a number of atoms 42 times the length of the vector signal, defined as

$$\Psi = [\Psi_{GABOR} \mid \Psi_{5TIME\ OVERSAMPLED\ DFT} \mid I] \tag{10}$$

Section 2 and Section 3 have respectively offered a fundamental review on CS theory and the definition of the overcomplete dictionary designated to represent the signals acquired during the basic AE test described in Section 7. The next two sections addresses two important issues for an effective use of CS for AE SHM purposes. Section 4 deals with the problem of estimating AE signal statistics in the compressed domain for low-power classification applications. Section 5 explains how the effect of noise can be taken into account.

4. AE SIGNAL STATISTICS IN THE COMPRESSED DOMAIN

This section deals with the problem of inferring some useful information directly from the compressive measurements. The possibility to solve a detection/classification problem in the compressed domain is attractive because it can

potentially reduce the computational effort of signal recovery only on a meaningful subset of collected samples. The detection problem aims to distinguish between two hypothesis:

$$\begin{aligned} \mathcal{H}_0 &: y = \Phi w \\ \mathcal{H}_1 &: y = \Phi(s + w) \end{aligned} \tag{11}$$

where $s \in \mathbb{R}^N$ is a known signal, $w \sim \mathcal{N}(0, \sigma^2 I_N)$ is i.i.d. Gaussian noise, and Φ is the sensing matrix. In literature[14] it is demonstrated how the compressive detector, defined by equation (12), is a sufficient statistic to distinguish between the two hypothesis stated in (11).

$$t := y^T(\Phi\Phi^T)^{-1}\Phi s \tag{12}$$

The statistic t is distributed as a normal distribution. In particular

$$t \sim \begin{cases} \mathcal{N}(0, \sigma^2\|\Phi^T(\Phi\Phi^T)^{-1}\Phi s\|^2) & under\ \mathcal{H}_0 \\ \mathcal{N}(\|\Phi^T(\Phi\Phi^T)^{-1}\Phi s\|^2, \sigma^2\|\Phi^T(\Phi\Phi^T)^{-1}\Phi s\|^2) & under\ \mathcal{H}_1 \end{cases} \tag{13}$$

Defined the false alarm rate $\mathbb{P}_F$ and the detection rate $\mathbb{P}_D$ respectively as

$$\begin{aligned} \mathbb{P}_F &= \mathbb{P}(\mathcal{H}_1\ chosen\ when\ \mathcal{H}_0\ true) \quad and \\ \mathbb{P}_D &= \mathbb{P}(\mathcal{H}_1\ chosen\ when\ \mathcal{H}_1\ true) \end{aligned} \tag{14}$$

for a given level of performance defined as $\mathbb{P}_F = \alpha$, the reference value γ for the statistic test t is computed as

$$\gamma = \sigma\|\Phi^T(\Phi\Phi^T)^{-1}\Phi s\|Q^{-1}(\alpha) \tag{15}$$

where

$$\begin{aligned} \mathbb{P}_F &= P(t > \gamma|\mathcal{H}_0) = Q\left(\frac{\gamma}{\sigma\left\|\Phi^T(\Phi\Phi^T)^{-1}\Phi s\right\|}\right) \\ \mathbb{P}_D &= P(t > \gamma|\mathcal{H}_1) = Q\left(\frac{\gamma - \left\|\Phi^T(\Phi\Phi^T)^{-1}\Phi s\right\|^2}{\sigma\left\|\Phi^T(\Phi\Phi^T)^{-1}\Phi s\right\|}\right) \\ Q(z) &= \frac{1}{\sqrt{2\pi}}\int_z^\infty exp(-u^2/2)du \end{aligned} \tag{16}$$

In Section 6, the detection problem introduced above is used to discriminate from noise a set of events collected in basic AE experimental tests. Results are provided for a different number of compressive measurements, i.e. for different undersampling factors. Experimental data are also exploited to move some initial steps to tackle the problem of classification. A classification problem aims to distinguish between different hypotheses

$$\widetilde{H_i} : y = \Phi(s_i + w) \tag{17}$$

for $i = 1,2, \ldots, R$ where each s_i is a known signal, $w \sim \mathcal{N}(0, \sigma^2 I_N)$ is i.i.d. Gaussian noise and Φ is the $M \times N$ sensing matrix. In the case where each hypothesis is equally likely, the classifier with minimum probability of error selects the $\widetilde{\mathcal{H}_i}$ that minimizes

$$t_i := (y - \Phi s_i)^T(\Phi\Phi^T)^{-1}(y - \Phi s_i) \tag{18}$$

In section 6 the results obtained for two different class of signals are discussed.

5. THE EFFECT OF NOISE

This paper deals with the problem of recovery a sparse signal in presence of well-structured noise. In particular, the setting considered is the one in which also the noise is sparse. The author of work[14] has introduced a measurement noise model in which the noise is supposed to be sparse in a suitable basis, or subspace. His original model is quoted in equation (19):

$$y = \Phi x + \Omega e \tag{19}$$

where Φ is the sensing matrix, x is a vector itself sparse, Ω is an $M \times L$ matrix with $L \leq M$ orthonormal columns, and the vector e is sparse. According to the work done by the author, measurements can be expressed in terms of an $M \times (N + L)$ matrix multiplied by a $(K + \kappa)$-$sparse$ vector:

$$\Phi x + \Omega e = [\Phi\ \Omega] \begin{bmatrix} x \\ e \end{bmatrix} \tag{20}$$

The reconstruction algorithm specifically designed by the author to recovery both the sparse signal x and the sparse noise e from a set of compressive measurements $y = \phi x$ is dubbed Justice Pursuit (JP) and is formulated as

$$\hat{u} = \overset{argmin}{u} \|u\|_1 \quad \text{subject to } [\Phi\ \Omega]u = y \tag{21}$$

where $\hat{u}$ is an intermediate $(N + L) \times 1$ recovery vector. The author analyzes in depth the theoretical conditions under which x and e are perfectly recovered. Furthermore his results[14] shows how JP is blind to the locations and size of the corrupted measurements, i.e. exact recovery can be achieved regardless to which particular elements of the measurement vector y are corrupted and to how large is the amount of noise on the sample y_i. Beside exploiting the possibilities offered by JP, the author demonstrates how sparsity can be also leveraged to remove unwanted component of signal. To better understand this possibility in these paper two different sources of noise are taken into account. Referring to another work[16] , noise is differentiate into a component added to each sample (sampling error) and another component added to the signal prior to sampling (observation error). Recalling what has already been introduced in Section 2, in this paper CS is modeled as a process that collects a set of linear measurements $y = (y_1, y_2, \dots, y_m)^T$ by applying an $M \times N$ sensing matrix Φ to a signal f for which exists a sparse representation in some overcomplete dictionary, i.e. $f = \Psi x$. In this setting the sampling error is represented by a noise vector $w_{obs} = (w_1, w_2, \dots, w_m)^T$ that corrupts the set of linear measurements y, i.e.

$$y = \Phi\Psi x + w_{smp} \tag{22}$$

In work[16] the entries of the vector w_{smp} are represented by random noise w_i. As stated before, this paper specifically addresses the case where the sampling error is sparse in some basis. In this particular setting holds $w_{smp} = \Omega e$ and thus the sampling model (22) leads back to the model (19) used for JP in the more general setting where the signal is not sparse itself, but is sparse when represented in a suitable dictionary Ψ. For the sake of clarity the results of work[14] – results that are equipped with a detailed analysis of the theoretical performance guarantees – are obtained for a signal x itself sparse. According to the author the results can be extended to the more general case where the signal is sparse in some basis Ψ, i.e. an orthogonal matrix with dimension $N \times N$ where N is the length of the signal. The case where the signal is sparse in an overcomplete dictionary is not addressed and consequently there is no possibility to directly inherit the theoretical bounds developed by the author.

The general sampling model (22) can be completed introducing the observation error w_{obs}[16]. So the model (22) turns into the following

$$y = \Phi\Psi(x + w_{obs}) + w_{smp} \tag{23}$$

where w_{smp} is, as previous, the sampling error. It is worth to notice that the observation error w_{obs} is already embedded in the signal before it goes through the sampling process. Under a mathematical point of view holds that $w_{obs} \epsilon \mathbb{R}^N$ whereas $w_{smp} \epsilon \mathbb{R}^M$. Both w_{obs} and w_{smp} are originally populated with i.i.d. elements with zero mean and variance σ_w^2.[16] In this paper the authors consider w_{obs} as an unwanted component of the signal with, once time again, a sparse representation. In particular, whereas the noise is still modeled as $w_{smp} = \Omega e$, the overcomplete dictionary Ψ, and accordingly the vector of the sparse coefficients x, is split respectively in two parts $[\Psi_{signal}\ \Psi_{noise}]$ and $[x_{signal}\ x_{noise}]^T$ so that the model (23) becomes

$$y = \Phi[\Psi_{signal}\ \Psi_{noise}] \begin{bmatrix} x_{signal} \\ x_{noise} \end{bmatrix} + \Omega e \tag{24}$$

The model (24) is intended to provide a sparse representation for the signal and an eventual unwanted component both equipped with a well-defined structure. Defining a matrix $\widetilde{\Phi} = \Phi\Psi = \Phi[\Psi_{signal}\ \Psi_{noise}]$, model (24) is finally recast in

$$y = [\widetilde{\Phi}\ \Omega] \begin{bmatrix} x_{signal} \\ x_{noise} \\ e \end{bmatrix} \tag{25}$$

By leveraging sparsity, model (25) is exploited in order to recover both the signal f and the sampling noise e, as performed with standard JP, with the added capability to denoise the signal from the eventual presence of unwanted component. The reconstruction algorithm is now formulated as

$$\hat{u} = {}^{argmin}_{u}\|u\|_1 \quad \text{subject to } \left[\widetilde{\Phi}\ \Omega\right]u = y \tag{26}$$

where the matrix $\widetilde{\Phi} = \Phi\Psi$ has dimension $N \times K$, with K the number of atoms of the dictionary, and $\hat{u}$ is an intermediate recovery vector with dimension $(K + L) \times 1$, with L the number of columns of Ω.

This section concludes by shedding a light on how the sparse sampling error term w_{smp} and the reconstruction algorithm (26) can be exploited to analyze the robustness of CS hardware in the event that one or more of its sampling channels undergo a failure. In Figure 2 is provided an elementary description of a typical CS architecture in order to introduce the concept of CS channel and explain how the aforementioned goal can be achieved. Further detail on the approach followed and the obtained results will be discussed in Section 6.

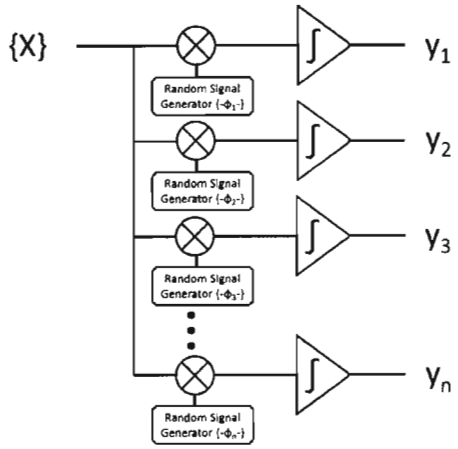


Figure 2. Schematic representation of a CS hardware architecture

A CS acquisition hardware, schematically represented in Figure 2, is a device capable to acquire an analog signal $x(t)$ and return a digital output, i.e. a vector $y(n)$ collecting the compressive measurements. Each element of the digital output $y(n)$ is associated to one of the parallel channels composing the CS system. The continuous-time signal $x(t)$ is correlated with test functions $\Phi_j(t)$ and then integrated. The remarkable advantage of such an architecture is that the sampling process is performed by the analog to digital converter (ADC) on the integrated signal and thus at a frequency much lower than those of a traditional acquisition system. CS architecture can thus potentially yield to realize an acquisition system demanding for low power consumption.[14] Bearing in mind the CS hardware architecture introduced above, the failure of a $i\text{-}th$ CS channel can be modeled with a high burst of noise corrupting the associated sample y_i. Such a kind of noise shows a well-defined structure that has an inherently sparse representation in the canonical basis, i.e. the collection of Dirac delta functions defined as $\psi_k(t) = \delta(t - k)$. For this reason, the sparse sampling noise $w_{smp} = \Omega e$ can be modeled setting $\Omega = \mathrm{I}$, where I is the identity matrix. Results showed in Section 6 prove how gracefully a CS system degrade when a small amount of its channels brake.

6. EXPERIMENTAL TESTS

Experimental tests are performed on the three-story building structure shown in Figure 3. The structure consists of aluminum columns and plates assembled using bolted joints, which slides on rails that allow movements in the x-direction only. At each floor, four aluminum columns (177 x 25 x 6 mm) are connected to the top and bottom aluminum plates (305 x 305 x 25 mm) forming a four degree-of-freedom (DOF) system. An electrodynamic shaker can be turn on to provide a lateral excitation to the base floor along the center line of the structure. The structure has been specifically designed to study in depth several structural health monitoring (SHM) techniques.[13] In this paper the structure is exploited to perform some basic AE test. An AE source is simulated on the top of the 1st floor, as indicated by the dotted arrow in Figure 3, thanks to a pencil lead fracture. This technique is a standard method for simulating AE signals thanks to its fast rise time transient force, close to the behavior of real AE source, and to the remarkably

reproducible resulting waveform.[2] The acoustic waves released by the crack are collected by a piezoelectric accelerometer PCB 336C installed on the side of the 1st floor as indicated by the dashed line circle. Data are sampled at a frequency of 5120 Hz. In this way, from the one hand, the collected signals show an high frequency content related to the AE event and, from the other hand, each AE event can be entirely sized by a short length record. As a matter of fact, as introduced in Section 2, signal recovery performed by solving the optimization problem of equation (4) results in large-size linear program (LP). In the specific case of reconstruction performed in this paper, l_1 norm recovery process is carried out on 1024-point data vector in order to keep computation time within reasonable limits. In this regard it is worth to notice that recovering a 1024-point signal, from a set of 250 compressed measurements seeking for a sparse representation in the overcomplete dictionary defined by equation (10), requires approximately 1400 s by running a non-optimized code on an ordinary desktop personal computer.

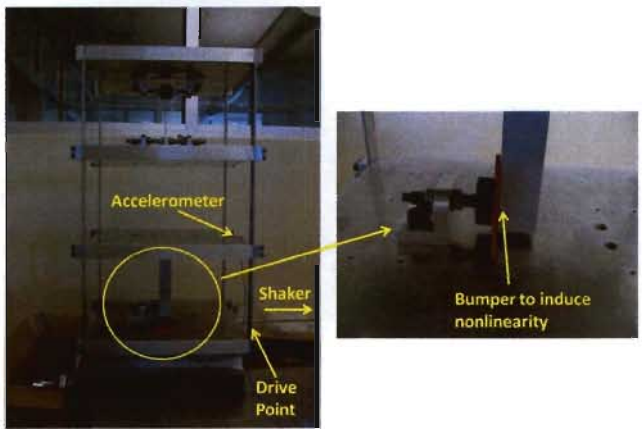
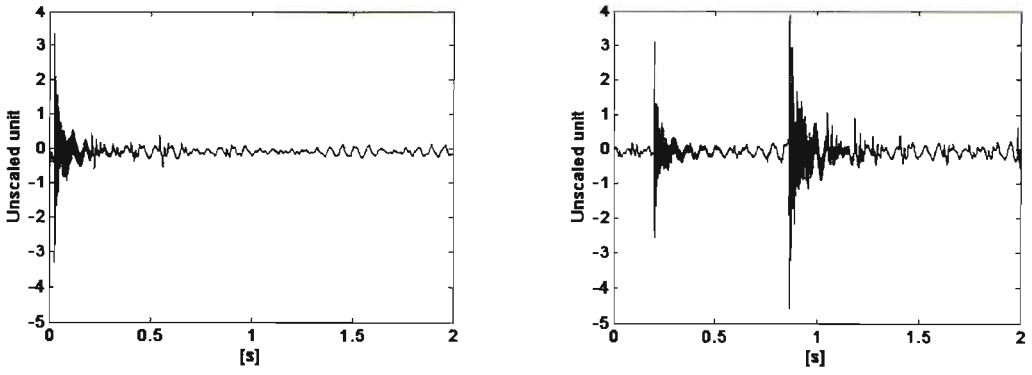


Figure 3. The three-story structure used for experimental tests

Below are discussed the two different types of experimental test carried out. Type A tests are specifically intended to study the suitability of the AE detection technique in the compressed domain described in Section 4 and are performed with the electrodynamic shaker turned off. Type B tests, on the other hand, reproduce a setting in which operational and environmental conditions affect the measured response. Type B tests are thus performed exciting the three-story structure and aim to study, by exploiting the reconstruction algorithm (26), both the denoising capabilities of CS and its robustness to a small amount of sensing channels failures. The excitation is provided by a band-limited random signal in the range of 20-500 Hz in order to avoid the rigid body modes of the structure that are present below 20 Hz.

6.1 Type A tests

In Type A tests, AE events are generated by a pencil lead fracture and the shaker is not running. Two different data sets consisting of 25 records each are collected. One data set presents single AE events like the one depicted in Figure 4(a). The other data set is made of records with the occurrence of two AE events as shown in Figure 4(b).



(a) (b)

Figure 4. Representative signals of the data set with only one AE event (a) and two AE events (b)

The data set with single AE events is used to study the detection problem introduced in Section 4. The hypothesis test stated in (11) is performed with a noise level $\sigma = 0.7$. In this way it is simulated a setting in which the detection of the known signal s is embedded in high background noise w. The resulting signal to noise ratio (SNR), computed as $SNR = 10log_{10}(\|s\|^2/\|w\|^2)$, is reported in Table 2 for each signal s.

Table 2. Probability of detection $\mathbb{P}_D$ for the signals of the data set with only one AE event.

		$\mathbb{P}_D$ with $\alpha = 5\%$			
Signal	**SNR [dB]**	**M/N = 1.25%**	**M/N = 2.5%**	**M/N = 5%**	**M/N = 10%**
s1	-4.0	95.8%	99.9%	100.0%	100.0%
s2	-6.2	80.9%	98.7%	100.0%	100.0%
s3	-9.7	50.3%	82.1%	97.1%	100.0%
s4	-7.1	75.1%	95.8%	100.0%	100.0%
s5	-4.1	97.1%	99.9%	100.0%	100.0%
s6	-4.4	89.9%	100.0%	100.0%	100.0%
s7	-5.9	87.1%	98.5%	100.0%	100.0%
s8	-5.2	87.8%	99.8%	100.0%	100.0%
s9	-4.0	93.8%	100.0%	100.0%	100.0%
s10	-8.2	64.6%	93.6%	99.7%	100.0%
s11	-6.3	83.4%	96.2%	100.0%	100.0%
s12	-6.7	87.3%	94.2%	100.0%	100.0%
s13	-7.5	75.9%	97.5%	99.8%	100.0%
s14	-5.5	92.3%	99.5%	100.0%	100.0%
s15	-5.9	89.0%	99.2%	100.0%	100.0%
s16	-7.9	61.3%	95.6%	99.5%	100.0%
s17	-4.7	95.0%	99.8%	100.0%	100.0%
s18	-7.1	78.3%	97.0%	99.9%	100.0%
s19	-3.2	98.4%	100.0%	100.0%	100.0%
s20	-4.2	96.9%	100.0%	100.0%	100.0%
s21	-6.9	75.9%	97.4%	99.9%	100.0%
s22	-6.8	86.6%	99.2%	100.0%	100.0%
s23	-7.5	84.8%	96.2%	99.9%	100.0%
s24	-5.5	91.6%	99.7%	100.0%	100.0%
s25	-4.5	98.3%	100.0%	100.0%	100.0%

Looking at the results listed in Table 2 it appears that the probability of detection $\mathbb{P}_D$, with $\alpha = 5\%$, improves by increasing the number of compressive measurements and is equal to 1 for all signals when the undersampling factor M/N is equal to 10%.

As far as it concerns the classification problem, signal s_{11} of the data set with single AE event is chosen as reference signal and, according to equation (18) introduced in Section 4, the statistic t_i is computed for all the other signals available, both from the data set with single AE events, and from the other data set with two AE events per record. The values of t_i are represented with a box-plot in Figure 5.

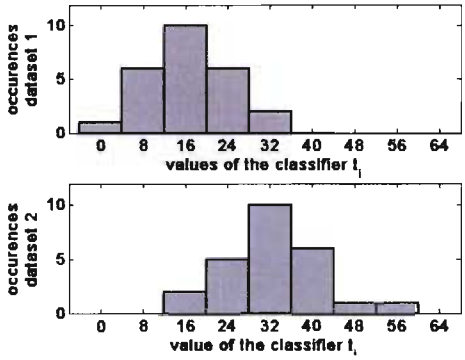


Figure 5. Box-plot representing the values of the classifier t_i for an undersampling factor $M/N = 2.5\%$

The definition (18) of the classifier t_i helps to understand that t_i is minimized for the signal s_i such that Φs_i is closest to the set of compressed measurement y. Computing the classifier for all the signals of the two data set, it results that the signal s_{11} effectively minimizes the classifier t_i to 0. The representation of Figure 5 suggests that the compressive measurements collected on the signals of the two dataset group differently. Taking as reference a generic signal of the data set with only one AE event per record, i.e. s_{11}, the compressive measurements on the signal of the other data set result in larger value of the classifier t_i.

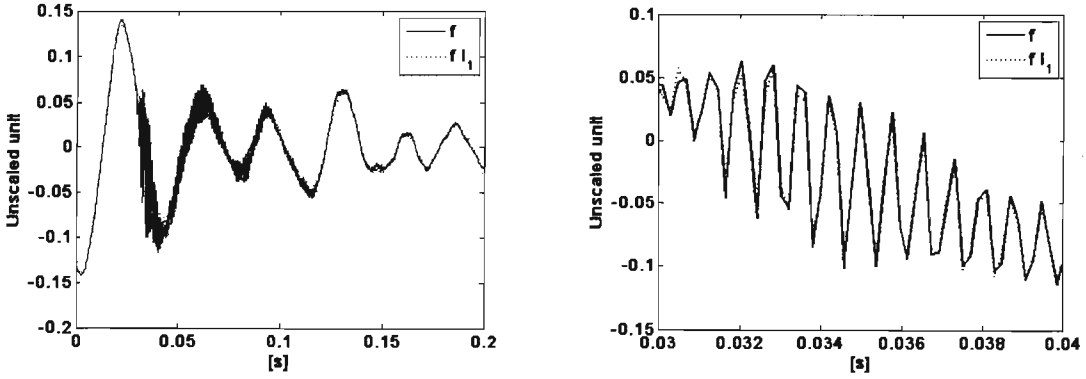
6.2 Type B tests

In type B test, the band-limited random excitation allows to excite all the modes of the structure, with exception of the first that is a rigid body mode. In Table 3 are reported the experimental modal parameters for the 2nd, 3rd and 4th mode identified in a previous work.[31]

Table 3. Experimental modal parameters of the three-story building structure depicted in Figure 4.

Experimental modal parameters	Mode	2	3	4
	Frequency [Hz]	30.7	54.2	70.7
	Damping ratio (%)	6.3	2.0	0.97

Figure 6 depicts the results obtained by reconstructing a 1024-point length data vector by solving the optimization problem P1 of equation (4) with a compressive measurement vector y of length $M = 250$.



(a) (b)

Figure 6. Original f and reconstructed $f\ l_1$ signal (a) and their detailed view (b)

The sensing matrix Φ is formed by sampling i.i.d. entries from the normal distribution with mean 0 and variance $1/M$. The overcomplete dictionary used is the one defined according to equation (10). The overall error of the reconstruction,

computed as $\varepsilon = \frac{\|f - \hat{f}\|}{f}$, is equal to 6.5%. Looking at Figure 6(a) and (b) it clearly appears that the AE event is buried into the signal of the measured response of the structure. The signal associated to the response of the structure induced by the horizontal shaker represents an unwanted component. In order to remove it, both the overcomplete dictionary and the associated sparse coefficient vector can be respectively reorganized in $[\Psi_{signal}\ \Psi_{noise}]$ and $[x_{signal}\ x_{noise}]^T$ as explained in Section 4. In particular, by observing that the mechanical response of the structure is totally concentrated over the frequencies of the three modes, Ψ_{noise} is arranged in order to collect:

- all the atoms of the Gabor dictionary with a center frequency below 100 Hz and thus potentially affected by the low noise frequency content associated to the response of the structure;
- the totality of the oversampled DFT atoms. As a matter of fact a sparse representation is obtained by solving the l_1 norm optimization problem. The eventual presence of few large coefficients associated to the DFT atoms has to be considered representative of the stationary response of the structure under the band-limited excitation provided by the shaker;
- all the atoms of the identity matrix conceived to model the eventual presence of noise bursts without any specific harmonic content.

As consequence of the aforementioned choices, Ψ_{signal} collects all the remaining atoms of the Gabor dictionary. Under the controlled laboratory condition it can be assumed that all the atoms collected in Ψ_{signal} are representative of the AE event. According to this assumption, the unwanted component represented by the response of the structure and eventual sources of spike noise is modeled with $f_{unwanted} = \Psi_{noise} x_{noise}^T$. On the contrary, reconstructed AE event is obtained as $f_{AE} = \Psi_{signal} x_{signal}^T$. Both reconstructed signals are depicted in Figure 7(a) and (b).

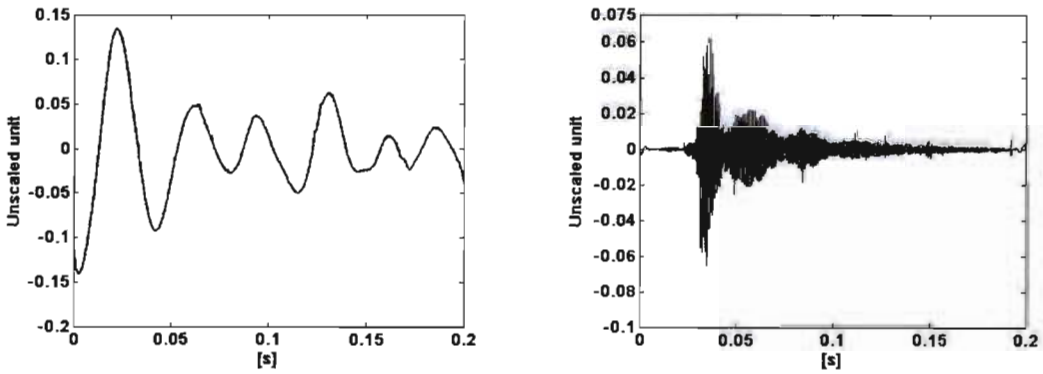


Figure 7. The unwanted component of the signal (a) and the denoised AE event (b)

Results of Figure 7 prove denoising capabilities achieved by rearranging the dictionary atoms and the associated coefficients. By exploiting sparsity, particular structures embedded in the given signal f prior to be sampled can be differentiate. Once explored denoising capability offered by the signal sparsity – property upon which CS relies – the last part of the present paragraph exploits model (25) and the reconstruction algorithm (26) to study the CS hardware robustness to a few of its channel failures. Keeping in mind the CS hardware architecture presented in Section 5, the study about CS robustness is carried out according to the following procedure:

1. a set y of $M = 250$ compressed measurements is collected on a type B test data vector with a sensing matrix Φ formed by sampling i.i.d. entries from the normal distribution with mean 0 and variance $1/M$. The 1024-point length data vector is the same used before to demonstrate CS denoising capability;
2. the failure of a different number of CS channels is simulated. In particular the number of failures ranges, by step of 5, from a minimum of 5 to a maximum of 25;
3. a number of samples equal to the number of channels undergoing a failure is selected uniformly at random among the M measurements available. Broken channels are indexed by bb;

4. a high burst of noise, simulating a CS channel failure, is added on each of the samples picked up at the previous point. Without any further specific knowledge about how CS hardware is practically implemented, all the noise bursts consist of spikes whit amplitude set equal to the RMS value of the original uncorrupted compressive measurements vector y. The sampling error vector w_{smp} of model (20) is thus obtained as

$$w_{smp} = \{w_i\}, \text{ with } w_i = RMS(y) \text{ for } i = bb \text{ and } w_i = 0 \text{ for } i \neq bb \tag{27}$$

5. reconstruction algorithm (26) is solved. To model the high burst of noise, Ω is chosen as $\Omega = c\mathrm{I}$, where I is the identity matrix, and c is a constant to rescale the norm of each columns of Ω to the value of the average norm of the column vectors of $\widetilde{\Phi}$. This choice is dictated by the practical consideration that the l_1 norm optimization problem (26) has to find a sparse vector u seeking among a matrix $[\widetilde{\Phi}\,\Omega]$ which columns should have a comparable energy content. The authors recognize their lack in supporting this choice with a proper mathematical justification and restrict themself to critically discuss the results obtained.

Results of the simulations performed according to the procedure described above are collected in Table 4.

Table 4. Results of the simulations to asses robustness of CS to a small number of its channels failures.

N° of CS channel failures	0	5	10	15	20	25
Reconstruction error [%]	6.5	7.2	9.5	8.9	14.4	15.0

It is worth to notice that the accuracy of the reconstruction worsens with respect to the case of uncorrupted measurements when a small number of channels undergoes a failure. The reconstruction error increases to values that can be considered still acceptable even if the 10% of CS channels fails.

To further understand the results obtained by solving the optimization problem of equation (26), the worst case, the one with 25 channels broken, is analyzed in depth. According to the sampling model (23) and (24), the sparse noise added to the uncorrupt measurement vector y is modeled by

$$w_{smp} = \Omega e = cIe \tag{28}$$

where $\Omega = cI$ is the rescaled identity matrix used to represent the noise bursts and e is the associated sparse coefficient vector. Figure 8 represents the sparse noise w_{smp} as it is reconstructed by solving the optimization problem (26).

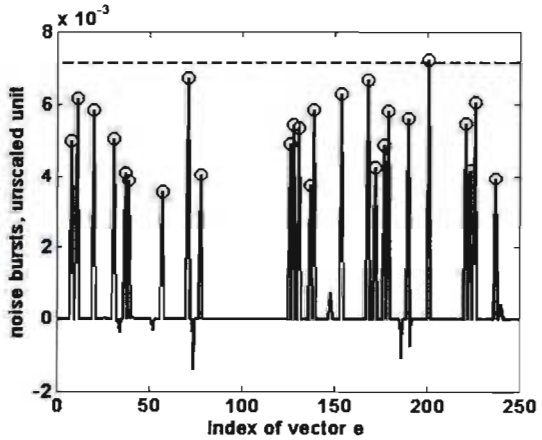


Figure 8. Reconstructed noise bursts (continuous line), burst noise level set for the simulation (horizontal dashed line) and reconstructed noise bursts that match the simulated broken channels (circle).

The reconstructed noise bursts are far to be exactly reconstructed in amplitude. As a matter of fact, according to the simulation procedure described above, all the bursts should have the same amplitude, set equal to the RMS value of the uncorrupted compressive measurements vector y and represented with an horizontal dashed line in Figure 8. Notwithstanding the amplitude error, the spikes associated to the simulated noise bursts can be distinctly detected. To point out this fact, in Figure 8 circles are used to mark the spikes whose index matches the index bb of the simulated broken channels. On the one hand, according to the results collected in Table 4, the sparse coefficients recovered in the vector $[\mathrm{x}_{signal}\ \mathrm{x}_{noise}]^T$ by solving the algorithm of equation (26) still continue to properly model the signal vector f in a setting where the failure of few CS channels is envisaged. On the other hand, the basic CS channel failure simulations carried out suggest that the sparse coefficients e, modeling the sampling error, can be exploited to perform CS hardware diagnosis. As a matter of fact the CS channels indexed by the large noise bursts showed in Figure 8 can be excluded from the CS measurement process. As result of this operation even the algorithm (4), based on the model (1) that is totally blind to any source of noise, can be used to satisfactorily reconstruct the signal. In particular by deleting the rows of the original sensing matrix Φ corresponding to the broken channels, a new sensing matrix Φ_{health} of dimension 225×1024 is obtained. Using Φ_{health} to collect a reduced number of compressive measurements from a subset of healthy CS channels, the reconstruction performed with the algorithm (4) results affected by a tolerable error equal to 7%.

7. CONCLUSIONS

This paper provides an initial step toward the study of compressed sensing (CS) suitability to tackle structural health monitoring related problems. In this specific case a setting with the presence of simulated AE sources has been taken into account. CS protocol has demonstrated its potentiality in providing a satisfactorily representation of AE signals by collecting a number of compressed measurements roughly equal to the 25% of samples required by conventional techniques founded on the Nyquist-Shannon theorem.

In presence of source of noise with a well-defined structure, the property of sparsity has been profitably exploited. Two different models have been discussed in order to take into account both the noise added during the sampling process and the noise embedded in the signal prior to sampling. The former, assuming sampling noise sparsity in the Dirac basis, has allowed to keep the reconstruction error within acceptable levels when a small amount of compressed sensing channels undergoes a failure. The latter has made possible to remove the unwanted component of a signal in a setting representative of a system where the measured response is affected by operational and environmental conditions. In this setting, denoising has been made possible thanks to the different sparse representation of the signal and the unwanted component in an overcomplete dictionary. The possibility to model defective CS channels with sparse noise bursts have furthermore suggested the opportunity to perform basic diagnostic on compressive sensing hardware.

This paper has also studied the significance of computing some statistics directly in the compressed domain. The results obtained, accordingly to the outcomes of other work present in literature, suggests that basic detection/classification problem can be solved relying upon a number of compressed measurements fewer than that required to properly recover the signal. In the ultimate attempt to implement de-centralized, low-power demanding structural health monitoring (SHM) solutions, great attention will need to be paid to the possibility to exploit a reduced amount of compressive measurements to estimate the relevant sufficient statistics to solve a desired interference task.

REFERENCES

[1] Balageas, D., Fritzen C. P., and Guemes A., [Structural Health Monitoring], ISTE Ltd, (2006).
[2] Shull, P. J., [Nondestructive Evaluation Theory, Techniques and Applications], CRC Press, (2002).
[3] Kaiser, J., [Untersuchungen uber das Auftreten Geraushen beim Zugversuch (An Investigation into the Occurrence of Noises in Tensile Tests)], PhD dissertation, Technische Hochschule, Munich Germany, (1950).
[4] Drouillard, T., [Acoustic Emission: A Bibliography with Abstracts], New York, Plenum Data Co., (1979).
[5] Miller, R. K., and McIntire, P., [Nondestructive Testing Handbook, Vol. 5, Acoustic Emission Testing], American Society for Nondestructive Testing, 2-6 (1987).
[6] Candès, E. J., Romberg, J., and Tao, T., "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," IEEE Trans. Inform. Theory, 52(2), 489-509 (2006).
[7] Donoho, D., "Compressed sensing," IEEE Trans. Inform. Theory, 52(4), 1289-1306 (2006).
[8] Tropp, J. A., Laska, J. N., Duarte, M. F., Romberg, J. K., and Baraniuk, R. G., "Beyond Nyquist: Efficient Sampling of Sparse Bandlimited Signals," IEEE Trans. Inform. Theory, 56(1), 520-544 (2010).
[9] Rubinstein, R., Bruckstein, A.M., and Elad, M., "Dictionaries for Sparse Representation Modeling," Proc. IEEE 98(6), 1045-1057 (2010).

[10] Cooley, J. W., and Tukey, J. W., "An algorithm for the machine calculation of complex Fourier series," Math. Comput., 19(90), pp. 297–301, 1965.
[11] Mallat, S., and Zhang, Z., "Matching pursuits with time-frequency dictionaries," IEEE Trans. Signal Process., 41(12), 3397–3415 (1993).
[12] Candès, E. J., and Wakin, M. B., "An Introduction to Compressive Sampling," IEEE Signal Processing Magazine, 25(2), 21-30 (2008).
[13] Chen, S. S., Donoho, D. L., and Saunders, M. A., "Atomic Decomposition by Basis Pursuit," SIAM Review, 43(1), pp. 129-159 (2001).
[14] Davenport, M. A., [Random Observations on Random Observations: Sparse Signal Acquisition and Processing], Ph.D. Dissertation, Rice University, Houston TX, USA (2010)
[15] Donoho, D.L., Elad, M., and Temlyakov, V.N., "Stable Recovery of Sparse Overcomplete Representations in the Presence of Noise," IEEE Trans. Inform. Theory, 52(1), 6-18 (2006).
[16] Reeves, G., [Sparse Signal Sampling using Noisy Linear Projections], Technical Report No. UCB/EECS-2008-3, Electrical Engineering and Computer Sciences University of California at Berkeley, Berkeley CA, USA (2008).
[17] Candès, E. J., Eldar, Y. C., Needell, D., and Randall, P., "Compressed Sensing with Coherent and Redundant Dictionaries," Applied and Computational Harmonic Analysis, 31(1), 59-73 (2011)
[18] Elad, M., "Optimized Projections for Compressed Sensing," IEEE Transactions on Signal Processing, 55(12), 5695- 5702 (2007).
[19] Grant, M., and Boyd, S., CVX: Matlab software for disciplined convex programming, version 1.21, April 2011
[20] Grant, M., and Boyd, S., [Graph implementations for nonsmooth convex programs, Recent Advances in Learning and Control (a tribute to M. Vidyasagar)], V. Blondel, S. Boyd, and H. Kimura, editors, Lecture Notes in Control and Information Sciences, Springer, 95-110 (2008) http://stanford.edu/~boyd/graph_dcp.html.
[21] Candès, E.J., Wakin, M.B., and Boyd, S.P., "Enhancing Sparsity by Reweighting l(1) Minimization," Journal of Fourier Analysis and Applications, 14(5-6), 877-905 (2008).
[22] Candès, E. J., and Tao., T., "Decoding by linear programming," IEEE Trans. Inform. Theory, 51(12), 4203-4215 2005.
[23] Baraniuk, R., Davenport, M., DeVore, R., and Wakin, M., "A simple proof of the restricted isometry property for random matrices," Constr. Approx., 28(3), 253-263, 2008.
[24] Candès, E. J., and Tao, T., "Near optimal signal recovery from random projections: Universal encoding strategies?," IEEE Trans. Inform. Theory, 52(12), 5406-5425 (2006).
[25] Mendelson, S., Pajor, A., and Tomczak-Jaegermann, N., "Uniform uncertainty principle for Bernoulli and subgaussian ensembles," Constr. Approx., 28(3), 277-289 (2008).
[26] Rudelson,M., and Vershynin, R., "On sparse reconstruction from Fourier and Gaussian measurements," Comm. Pure Appl. Math., 61(8), 1025-1045 (2008).
[27] Sturm, B. L., Roads, C., McLeran, A., and Shynk, J., "Analysis, Visualization, and Transformation of Audio Signals Using Dictionary-based Methods", Journal of New Music Research, 38(4), 325-341 (2009)
[28] Gröchenig, K., [Foundations of Time-Frequency Analysis], Birkhäuser, Boston, MA, (2001)
[29] Hlawatsch, F., Auger, F., [Time-Frequency Analysis], ISTE Ltd and John Wiley & Sons, Inc., (2008)
[30] Rao, R.M., Bopardikar, A.S., [Wavelet Transforms: Introduction to Theory and Applications], Addison-Wesley Pub. Co, (1998)
[31] Figueiredo, E., Park, G., Figueiras, J., Farrar, C., and Worden,K., "Structural Health Monitoring Algorithm Comparisons Using Standard Data Sets", Los Alamos National Laboratory, (2009)
[32] Di Benedetti, M., Loreto, G., Matta, F., and Nanni, A., "Continuous acoustic emission monitoring of reinforced concrete under accelerated corrosion", JMCE (Submitted 2012).
[33] Mascarenas, D., Hush, D., Theiler, J., Farrar, C., "The application of compressed sensing to detecting damage in structures," 8th International Workshop on Structural Health Monitoring, Palo Alto, CA, 2011.