Computers and Structures 182 (2017) 26–40

Contents lists available at ScienceDirect

Computers and Structures

journal homepage: www.elsevier.com/locate/compstruc

Compressive sensing with an adaptive wavelet basis for structural system response and reliability analysis under missing data

L. Comerford [a,*], H.A. Jensen [b], F. Mayorga [b], M. Beer [a,c,d], I.A. Kougioumtzoglou [e]

[a] *Institute for Risk and Reliability, Leibniz University Hannover, Germany*
[b] *Dept. of Civil Engineering, Santa Maria University, Valparaiso, Chile*
[c] *Institute for Risk and Uncertainty, University of Liverpool, Liverpool L69 3GH, UK*
[d] *School of Civil Engineering & Shanghai Institute of Disaster Prevention and Relief, Tongji University, China*
[e] *Department of Civil Engineering and Engineering Mechanics, Columbia University, USA*

ARTICLE INFO

Article history:
Received 16 June 2016
Accepted 16 November 2016

Keywords:
Compressive sensing
Harmonic wavelets
Power spectrum
Missing data
Advanced simulation techniques
Reliability analysis

ABSTRACT

The challenge of determining response and reliability statistics of large-scale structural systems under earthquake induced stochastic excitations is considered where the source load data records are incomplete. To this aim, a compressive sensing based framework in conjunction with an adaptive wavelet basis is presented for reconstructing the samples with missing data and estimating the underlying process EPS. In this regard, novel insights are provided whereas certain conceptual, numerical, and practical implementation aspects of the technique are presented in detail. A numerical example pertaining to the stochastic response and reliability analysis of an eight floor reinforced concrete building structural system demonstrates the effectiveness of the proposed methodology.

1. Introduction

Numerical simulations for the analysis of structural systems subject to random dynamic excitations require realistic stochastic models of the system excitation processes. Such systems may be highly sensitive to the nature of these excitations and so simulation accuracy is dependant upon a reliable excitation process model.

A reliable spectral model providing frequency dependant information can be of significant importance in investigating the response of an engineering system to stochastic input such as earthquakes. Further, spectral models may be utilized for generating stochastic process records, fitting with the frequency dependant statistics of the given model, for use in numerical Monte-Carlo analyses e.g. [1–3]. However, a basic spectral model such as that based on a non-windowed discrete Fourier transform (DFT) may only describe a stationary process, i.e. one in which the spectral content does not change over time. This assumption of stationarity may give a poor approximation of the true process, especially in the case of earthquake excitations in which the frequency content can change dramatically over their duration. Hence, in many cases, realization of time-dependant properties of stochastic processes is also considered central to defining reliable spectral models. In this regard, the concept of the evolutionary power spectrum (EPS) [4,5] provides an appealing model for capturing the statistics and the time-varying frequency content of the underlying non-stationary stochastic processes. Further, they can be used as a basis for joint time-frequency system response analysis [6,7], or efficient stochastic simulation utilizing advanced Monte Carlo techniques.

In an ideal scenario, such a model could be avoided entirely in the case where extremely large data banks of real recorded excitation processes of interest were available. Unfortunately, particularly in the field of earthquake engineering, this is seldom the case. Instead, process models are often estimated, based partially or entirely on a small set of relevant recorded processes. Numerous approaches exist for EPS estimation based on time records including short-time Fourier transforms, wavelet [8–10] & chirplet [11] transforms. Harmonic wavelets [12,13] are concentrated on in this paper due in part to their box-shaped frequency spectrum, ideal for identifying specific bands of energy and for the fact that they constitute an orthogonal basis, which is ideal for the compressive sensing approach applied herein.

It is logical to assume that the more data upon which such a model is based, the more statistically accurate/relevant numerical simulation results are likely to be. As the available data may be quite limited, it is important that it is utilized to the fullest extent, which in some cases includes working with problematic data sets.

* Corresponding author.
E-mail address: comerford@irz.uni-hannover.de (L. Comerford).

http://dx.doi.org/10.1016/j.compstruc.2016.11.012
0045-7949/© 2016 Published by Elsevier Ltd.

In this regard, when analyzing real earthquake excitation data, coupled with the problem of limited numbers of samples or shorter than ideal sample lengths, is the potential major issue of missing data. Practical reasons for having limited data include for example, equipment failure (if a sensor becomes damaged, perhaps even as a result of the process itself, data may be lost) and sensor thresholding limitations (high fidelity sensors with a wide operational range can be expensive, and so in some cases the equipment used to record a process may not be able to capture extreme features). Numerous other issues including sensor maintenance, bandwidth limitations, usage & data acquisition restrictions as well as data corruption may also lead to missing data.

Under these conditions, when working with limited and/or missing data, standard Fourier techniques for spectral estimation, will frequently demonstrate poor performance. Although there exist many algorithms and procedures in the literature that provide spectral estimates in the presence of missing data, these alternatives come with certain drawbacks and often impose significant assumptions on the statistics of the underlying stochastic process. For instance, autoregressive methods may be applied under the assumption that source time-histories are relatively long and that the missing data are grouped [14,15]. Further, least-squares sinusoid fitting and zero-padded gaps [16–18] offer efficient solutions for re-constructing the Fourier spectrum in the presence of missing data but suffer, in general, from falsely detected peaks, spectral leakage and significant loss of power as the number of missing data increases. Similar issues are faced when applying these methods to wavelet transforms in the case of EPS estimation, and specific approaches for non-stationary signal reconstruction are uncommon. However, recent developments have been made in the area of EPS estimation subject to missing data including applications of artificial neural networks [19] and compressive sensing (CS) [20]. The latter is applied herein, utilizing the relative band-limited nature of evolutionary earthquake spectra. To further improve the spectrum estimation, the CS approach is applied in conjunction with an adaptive basis re-weighting procedure, building on ideas introduced in [21], which is useful in the case where process record ensembles are available.

The organization of this contribution is as follows. A brief introduction to CS theory is provided in Section 2 with references to further reading. Section 3 highlights the differences between the methods applied herein and established CS theory. The novel adaptive basis re-weighting procedure for signal reconstruction from multiple records is detailed in Section 4. In Section 5, discrete orthogonal Harmonic Wavelets are introduced briefly, along with commonly encountered practical wavelet transform issues, before being set in the context of the CS-based reconstruction problem. Section 6 deals with the estimation of ground excitation power spectra from simulated earthquake records with various missing data configurations for which response statistics and system reliability of a large structural model subject to such excitations are compared. The work closes with some conclusions and final remarks.

2. Compressive sensing

CS [22,23] is a signal reconstruction method that is commonly used in image processing and becoming a widely used tool in civil and mechanical engineering. CS, when applied to missing data problems requires several important assumptions to be made concerning the nature of the process of interest. However, in many problem cases, especially those related to environmental processes (and in particular spectral representation of earthquake excitations), these assumptions can be made with confidence. In the group of missing data problems for which CS is applicable, significant gains in spectrum estimation accuracy and computational efficiency can be achieved over alternative reconstruction methods.

2.1. CS background

The Shannon-Nyquist theorem states that a time-dependent signal with maximum frequency f can be completely determined when sampled at time intervals of $\frac{f}{2}$ or smaller. This maximum sampling frequency is commonly known as the Shannon-Nyquist rate. Compressive sensing is a signal processing technique that allows for signal reconstruction even if the maximum frequency f present in the signal is greater than half the signal's sampling rate [24].

2.2. CS requirements

For robust compressive sensing, several properties concerning the source signal and transformation basis are required. Most importantly the signal must be sparse in a known basis, and obey properties of incoherence and restricted isometry (RIP). The last two requirements are discussed in detail in any introductory text on CS theory (eg. [25]). For clarity and completeness in notation a brief description of sparsity is provided in the following subsections.

2.2.1. Signal sparsity

For a sampled signal to exhibit sparsity in some known basis, it must be possible to represent that signal with far fewer coefficients than the number determined by the Shannon-Nyquist rate. A discrete time signal, x may be viewed as an N *by* 1 column vector. Given an orthogonal N *by* N basis matrix A in which the columns A_i are the basis functions, x may be represented in terms of this basis via a set of N *by* 1 basis coefficients y, i.e.,

$$x = \sum_{i=1}^{N} A_i y_i, \tag{1}$$

The vector x is said to be K-sparse in the basis A if y has K non-zero entries and $K < N$, i.e.,

$$x = \sum_{i=1}^{K} A_{n_i} y_{n_i}, \tag{2}$$

where n_i are the integer locations of the K non-zero entries in y. Hence y is an N *by* 1 column vector with only K non-zero elements. Therefore,

$$|y|_{L0} = K, \tag{3}$$

where $|.|_{Lp}$ denotes the Lp norm defined as

$$|y|_{Lp} = \left(\sum_i |y_i|^p\right)^{\frac{1}{p}}. \tag{4}$$

The $L0$ norm used in Eq. (3) is defined as the limit of the Lp norm as $p \to 0$. In general the $L0$ norm is the total number of non zero elements in a vector,

$$|y|_{L0} = \sum_i \begin{cases} 1 & y_i \neq 0 \\ 0 & otherwise \end{cases} \tag{5}$$

It is important to note that for real signals, it is highly unlikely that they are exactly sparse in any orthogonal basis. Even a minimal amount of random noise on top of an otherwise K-sparse signal will produce non zero coefficients for all N. However, a large number of