



## A probabilistic approach for the construction of regional earthquake response spectra

E. Kallinikidou<sup>a,\*</sup>, S.F. Masri<sup>a</sup>, R.L. Nigbor<sup>a</sup>, A.W. Smyth<sup>b</sup>, K.B. Olsen<sup>c</sup>

<sup>a</sup> Viterbi School of Engineering, University of Southern California, Los Angeles, CA 90089-2531, USA

<sup>b</sup> School of Engineering & Applied Science, Columbia University, New York, NY 10027, USA

<sup>c</sup> Department Geological Sciences, San Diego State University, San Diego, CA 92182-1010, USA

### ARTICLE INFO

#### Article history:

Received 11 October 2007

Received in revised form

4 March 2009

Accepted 6 March 2009

Available online 18 March 2009

#### Keywords:

Regional earthquake spectra

Probabilistic earthquake spectra

Karhunen–Loève expansion

Regional seismic risk

Proper orthogonal decomposition

### ABSTRACT

Five scenario earthquakes plausible for the Los Angeles metropolitan region, and one numerical approximation of the 1994 Northridge  $M_w$  6.7 event, provide the database of the proposed methodology that is applied for the construction of regional earthquake response spectra. The methodology involves two main stages of data compaction. In the first stage, the Karhunen–Loève (K–L) decomposition of the excitation temporal covariance matrix is performed. In the second stage, the dominant eigenvectors are analytically approximated with Chebyshev polynomials, thus being converted from eigenvectors to eigenfunctions. This compact analytical representation of the nonstationary excitation data provides an exact closed-form solution for the nonstationary response of linear multi-degree-of-freedom systems. Furthermore, statistical inference analysis for the response variables is conducted, which leads to the construction of regional probabilistic response spectra based on the Log-Normal probability model for the response variables.

© 2009 Elsevier Ltd. All rights reserved.

### 1. Introduction

A key feature of the ground motion generated by earthquakes is its transient nature. A typical accelerogram is characterized by a period of intensity growth, followed by an interval of almost steady strong shaking and finally a period of decay. For an accurate representation of a system's dynamic response to earthquake excitation, the earthquake loads should be formulated as a nonstationary stochastic process. A considerable amount of research into the simulation of nonstationary random processes was performed in the past. Representative publications in the field of applied mechanics, which deal with the stochastic response of dynamic systems, include the work of Caughey and Stumpf [1], Corotis and VanMarcke [2], Masri [3], Spanos and Lutes [4], Debchaudhury and Gasparini [5], Shinozuka and Deodatis [6], Iwan and Hou [7], Soong and Grigoriu [8], Lin and Cai [9], and Lutes and Sarkani [10].

A popular procedure for analytically representing stochastic earthquake ground motion time histories is to model them as zero-mean, white noise processes modulated by a deterministic

envelope function representing the variation of the mean-square intensity with time. While this approach is a step in the right direction, it is not an authentic representation of actual earthquake spectral contents. On the other hand, when the earthquake ground motion is modeled as filtered white noise – where the filter may have constant or time-varying parameters – the nonstationary solutions do not have a closed form (Conte and Peng [11]; Papadimitriou and Beck [12]).

Masri and Miller [13] proposed a procedure for the compact covariance kernels' probabilistic representation (zero-mean, Gaussian probability density function modulated by a deterministic exponential function) in a form that results in the analytical evaluation for the transient, mean-square response of a linear single-degree-of-freedom (SDOF) system. This procedure was later improved by using the orthogonal Karhunen–Loève (KL) expansion and the least-squares approach to develop an approximate, analytical fit for the KL-eigenvectors of the random excitation process (Traina et al. [14]). A few years later, a relatively small ensemble of ground motion records (66) from the Los Angeles region for the 1971, San Fernando  $M_w$  6.6 earthquake served as the application database of the method (Masri et al. [15]). Finally, high-quality ground motion records from the Los Angeles region due to the 1994, Northridge  $M_w$  6.7 earthquake furnished an excellent excitation stochastic process for a multi-degree-of-freedom (MDOF) response solution analysis (Masri et al. [16]).

In this paper, the potential of the analytical method cited in the previous studies to provide a procedure for the construction

\* Corresponding author. Tel.: +1 310 3092891.

E-mail addresses: [Elena.Kallinikidou@3ds.com](mailto:Elena.Kallinikidou@3ds.com), [ekallinikidou@solidworks.com](mailto:ekallinikidou@solidworks.com) (E. Kallinikidou), [masri@usc.edu](mailto:masri@usc.edu) (S.F. Masri), [nigbor@ucla.edu](mailto:nigbor@ucla.edu) (R.L. Nigbor), [smyth@civil.columbia.edu](mailto:smyth@civil.columbia.edu) (A.W. Smyth).

of regional response spectra is addressed. Synthesized acceleration records from several scenario earthquakes plausible for the Los Angeles basin serve as the excitation stochastic database.

### 1.1. Motivation

In recent years, major earthquake research projects have been focusing on developing large-scale computational simulations of earthquake scenarios for geographic regions that span hundreds of kilometers. Long-term goals of such simulations include a comprehensive physics-based understanding of complex earthquake phenomena, reliable prediction of ground velocities that are expected to shake the infrastructure above ground, and identification of high-risk areas expected to sustain the worst impact. The output of such simulations yields large data files with size that reaches terabytes, making the tasks of data-management and data-archiving particularly challenging. Representative publications in the field of large-scale computational earthquake simulations include the work of Maechling et al. [17], Faerman et al. [18], Olsen et al. [19] and [20], and Benites and Olsen, [21].

This paper addresses the challenge of incorporating the emerging knowledge of this state-of-the-art earthquake research into building codes that engineers can apply to the performance-based design of earthquake-resistant structures.

### 1.2. Scope

As the number of available strong motion records from future earthquakes will keep increasing, due to the expansion of seismological instrumentation worldwide, the proposed compact probabilistic method can be extensively applied and its efficiency significantly enhanced. The data-processing methodology is not only a useful data-archiving and earthquake feature-extraction tool, but can also accurately quantify the average seismic risk in a probabilistic format over a spatially extended area. Eventually, the methodology can lead to the construction of regional, probabilistic design spectra that will take into account the nonstationarity of both the earthquake excitation and the response.

Synthetic accelerograms of five scenario earthquakes for the Los Angeles basin, and one numerical approximation of the 1994 Northridge  $M_w$  6.7 event provide an excellent database to test the methodology. The extreme root-mean-square (rms) spectra for a linear SDOF system are directly calculated from the analytical solution of the nonstationary response, and compared with ensemble mean response quantities derived from the conventional response spectrum analysis. To construct probabilistic response spectra with a prescribed confidence interval, the appropriate probability model that could best describe the distribution characteristics of the response spectra variables (accelerations, velocities, and displacements) needs to be determined. A statistical inference analysis of the response parameters is performed towards that goal.

The material is organized as follows:

- Section 2: Analytical formulation for nonstationary excitation propagation. Derivation of closed-form solution for the response covariance of an MDOF system due to nonstationary support excitation.
- Section 3: Description of the Los Angeles basin model, and the earthquake scenarios' simulation data.
- Section 4: Derivation and comparison of synchronous and asynchronous earthquake acceleration covariance matrices. Orthogonal decomposition of synchronous acceleration covariances, and eigenvector approximation with Chebyshev polynomials.
- Section 5: Response spectra construction for each earthquake scenario based on (a) extreme standard deviation values, and (b) mean response values across the records' ensemble.

- Section 6: Investigation of appropriate probability model for response spectra variables. Construction of probabilistic response spectra. Superposition of extreme rms and percentile curves.
- Section 7: Conclusions and recommendation for future research.

## 2. Formulation

The data compaction of the system's input excitation is performed in two steps:

(1) The spectral decomposition of the input covariance matrix is calculated, and only the dominant eigenvectors are retained. (2) The dominant eigenvectors are then least-squares fitted with a series of Chebyshev orthogonal polynomials. This data compaction method permits the closed-form solution for a linear, dynamic system's nonstationary response to random excitation (Traina et al. [14]; Masri et al. [16]). Following is a brief overview of this procedure.

### 2.1. Eigenvector expansion of the covariance matrix

The covariance kernel  $[C]$  for a system with support acceleration  $\ddot{s}(t)$  is a symmetric, square matrix with values defined as

$$C_{\ddot{s}\ddot{s}}(t_1, t_2) = E[(\ddot{s}(t_1) - \mu_{\ddot{s}}(t_1))(\ddot{s}(t_2) - \mu_{\ddot{s}}(t_2))] \quad (1)$$

where  $E[\cdot]$  is the expectation operator, and  $\mu_{\ddot{s}}(t)$  is the mean value of the support acceleration at time  $t$ . Using the Karhunen–Loève expansion, the spectral representation of the  $[C]$  matrix of order  $n$  by  $n$  may be expressed as

$$[C] = \sum_{i=1}^n \lambda_i \mathbf{p}_i \mathbf{p}_i^T \quad (2)$$

where  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$  are the real and positive eigenvalues of  $[C]$ , and  $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n$  are the corresponding normalized eigenvectors such that  $\mathbf{p}_i^T \mathbf{p}_j = \delta_{ij}$ , for  $i = 1, 2, \dots, n$ . In most cases, only a small fraction  $k$  out of a total  $n$  eigenvalues are likely to reconstruct the covariance matrix  $[C_k]$ . Matrix  $[C_k]$  represents the reconstruction of the covariance matrix by using the first largest eigenvalues  $k$  out of a total  $n$  eigenvectors, and matrix  $[E_k]$  contains the error due to the eigenvector truncation:

$$[C] = \sum_{i=1}^k \lambda_i \mathbf{p}_i \mathbf{p}_i^T + \sum_{i=k+1}^n \lambda_i \mathbf{p}_i \mathbf{p}_i^T = [C_k] + [E_k]. \quad (3)$$

### 2.2. Least-squares fit of eigenvectors

The second step in the data condensation procedure involves the approximation of the truncated set of  $k$  discrete eigenvectors with continuous analytical functions. Chebyshev polynomials were chosen for the least-squares approximation because, in addition to being orthogonal, they have the desirable feature of equal-error approximation within an interval of interest, as well as being convenient for defining analytical solutions for the dynamic system response.

The approximated covariance matrix  $[\hat{C}_k]$  can then be reconstructed, and expressed in terms of a Chebyshev polynomial series (Masri et al. [15]):

$$[\hat{C}_k(t_1, t_2)] = \sum_{i=1}^k \lambda_i \sum_{j=0}^{m_i-1} \sum_{\ell=0}^{m_i-1} H_{ij} H_{i\ell} T_j(t'_1) T_\ell(t'_2), \quad (4)$$

where the  $T$ 's are the Chebyshev polynomials,  $H_{ij}$  is the coefficient of the Chebyshev polynomial of order  $j$  associated with the eigenvector  $\mathbf{p}_i$ ,  $m_i$  is the number of Chebyshev polynomials used, and  $0 \leq t_i \leq t_{\max}$  for  $i = 1, 2$ .

### 2.3. Nonstationary excitation propagation

The analytical representation of the input acceleration covariance matrix (Eq. (4)) permits the closed-form response solution for a multi-degree-of-freedom (MDOF), linear, viscously damped (and not necessarily of the proportional type), dynamic system. The system's response is arranged in a  $2n$ -dimensional vector of the form

$$\{y(t)\} = \begin{Bmatrix} x(t) \\ \dot{x}(t) \end{Bmatrix} \quad (5)$$

where  $\{y(t)\}$  is the state vector, and defines the state of the system uniquely for any time  $t$ . In this application, the  $n$ -dimensional vectors  $x(t)$  and  $\dot{x}(t)$  are the nodal, relative displacements and velocities with respect to the support position. The system of differential equations of motion for an MDOF linear system excited by a base excitation  $\ddot{s}(t)$  is given in matrix form by

$$[M]\{\ddot{x}(t)\} + [C]\{\dot{x}(t)\} + [K]\{x(t)\} = -[M]\{e\}\ddot{s}(t) \quad (6)$$

where  $[M]$ ,  $[C]$ , and  $[K]$  are the  $n \times n$  mass (diagonal), damping, and stiffness matrices, respectively, and  $\{e\}$  is the  $n$ -dimensional unit vector.

To derive the closed-form solution for the response covariance matrix of a MDOF system due to a support nonstationary excitation, analytical representation of Chebyshev polynomials and integrals in a finite series form are constructed. Since the focus of this paper is the application of existing analytical tools to develop seismic response spectra, the intermediate steps of the theoretical derivation are omitted. A detailed treatment of the closed-form solution of the nonstationary response can be found in the thesis by Smyth [22] and in the work of Masri et al. [16]. The development of the seismic response spectra in this paper is based on an SDOF representation of the response covariance.

## 3. Application

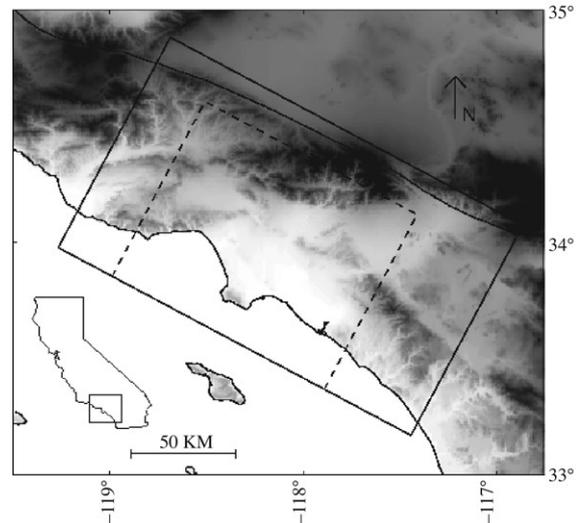
### 3.1. Los Angeles basin model and earthquake scenarios

The metropolitan area of the greater Los Angeles region lies on top of a deep sedimentary basin. During past earthquakes, such as the 1985,  $M_w$  8.1 Michoacan, and the 1989,  $M_w$  6.9 Loma Prieta events, it was observed that ground motions in basins get significantly amplified, as was the case in Mexico City, and the Marina district of San Francisco. To obtain a confident estimate of a particular site's basin response, all the possible regional earthquake scenarios need to be included in a probabilistic seismic hazard analysis. In this study, the following five scenario earthquakes, all plausible for the Los Angeles region, and a numerical approximation of the 1994,  $M_w$  6.7 Northridge event are examined:

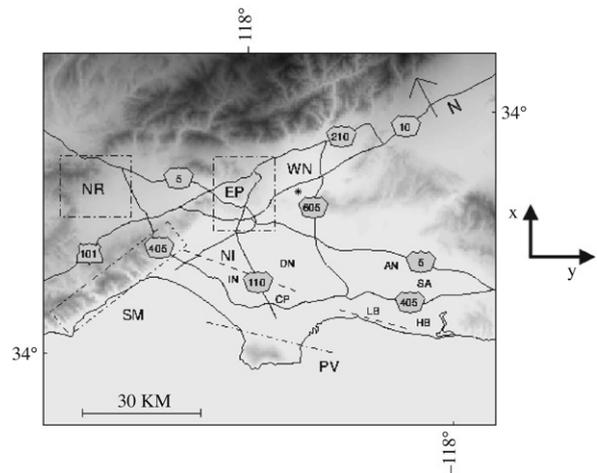
(1) A blind thrust on the Elysian Park fault (EP); (2) Thrust on the Santa Monica fault (SM); (3) Northwest propagating rupture on the Newport–Inglewood fault (NI); (4) Southeast propagating rupture (SAFSE); and (5) Northwest propagating rupture on a 170 km long stretch of the San Andreas fault (SAFNW); (6) Approximation of the 1994,  $M_w$  6.7 Northridge (NOR) earthquake.

These earthquake simulations were developed by Olsen [23] to explore the ground motion amplification of seismic waves in the deep Los Angeles basin. They are all within the “geological reasonable scenarios” defined by Dolan et al. [24]. The rupture parameters for the six earthquake simulations are given in Table 1.<sup>1</sup>

<sup>1</sup> A comprehensive treatment of the underlying physics of the wave propagation in the Los Angeles basin for the earthquake simulations can be found in Olsen [25].



**Fig. 1.** Topographic map of Southern California. The larger rectangular area depicts the extent of the model for the San Andreas Fault (SAF) scenarios. The smaller rectangle shows the region used for modeling the remaining scenarios.



**Fig. 2.** Topographic map of the Los Angeles basin area with the surface projections of the fault planes associated with the earthquake scenarios: Santa Monica (SM), Elysian Park (EP), Newport–Inglewood (NI), Palos Verdes (PV), 1994  $M$  6.7 Northridge (NR).

A topographic map of southern California is illustrated in Fig. 1. The larger rectangular area depicts the extent of the model for the San Andreas Fault (SAF) scenarios. Within the smaller rectangular area, which is shown in Fig. 2 along with the surface projections of the fault planes, lies the region for which the remaining scenarios are simulated. The reference frame of the numerical model was rotated  $28^\circ$  counterclockwise from the North, to minimize the computational requirements for the simulations. Throughout the present study, the horizontal ground motions and the subsequent response parameters are analyzed along the azimuths of  $118^\circ$  ( $x$ -axis), and  $28^\circ$  ( $y$ -axis). The model of the extended Los Angeles basin spreads over an area of  $155 \text{ km} \times 134 \text{ km} \times 34 \text{ km}$ , and is discretized with a grid spacing of approximately  $0.4 \text{ km}$  for a total of  $388 \times 337 \times 85$  ( $= 11,114,260$ ) grid points. It includes the Los Angeles basin, the smaller San Fernando basin to the northwest separated from the Los Angeles basin by the Santa Monica mountains, and the Ventura basin continuing to the North. A relatively small basin is located below the San Gabriel Valley, and the larger, but shallower

**Table 1**  
Earthquake rupture parameters (after Olsen).

	EP	SM	NI	SAFNW	SAFSE	NOR
Hypocenter						
Longitude (deg)	−118.20	−118.64	−118.24	−118.66	−117.09	−118.53
Latitude (deg)	34.10	34.06	33.88	34.79	34.09	34.20
Depth (km)	14.8	10.4	7.6	4.0	15.2	17.6
Fault parameters						
Width (km)	21	17	15	16	16	24
Length (km)	16	37	41	170	170	18
Depth to top (km)	10.5	0.4	0.4	0.4	0.4	5.0
Dip (deg)	25	65	90	90	90	40
Strike (deg)	298	260	138	118	118	118
Rake (deg)	90	90	180	180	180	101
Moment magnitude	6.75	6.75	6.75	7.5	7.5	6.7

**Table 2**  
Los Angeles basin 3D modeling parameters (after Olsen).

Spatial discretization (km)	0.4
Temporal discretization (s)	0.25
Lowest <i>P</i> -wave velocity (km/s)	2.41
Lowest <i>S</i> -wave velocity (km/s)	1.0
Lowest density (kg/m <sup>3</sup> )	2070
Number of time steps	360,500 (Northridge)
Total simulation time (s)	90,125 (Northridge)

Chino basin, which extends the LA basin to the East. The topography above sea level was not included. The three-dimensional (3D) modeling parameters are listed in Table 2.

### 3.2. Limitations of the Los Angeles basin amplification study

According to Olsen [25], the main limitations of the average amplification map for the Los Angeles basin area were primarily due to the relatively small number of scenario earthquakes (nine in the original study) that were used in the research study and the maximum frequency (0.5 Hz) used for the ground motion simulation, and secondarily, to uncertainties concerning the accuracy of the basin model, and the omission of surface layers with shear velocities less than 1 km/s. It is important to incorporate more simulations of geologically reasonable scenarios, in order to reduce the uncertainty of the average amplification ratios,<sup>2</sup> and also to increase the highest frequencies, to capture the entire spectrum of amplification effects. Considering the recent explosive increase in computational power, it may be feasible to construct more accurate amplification maps in future studies with broader frequency content derived from smaller model grid spacing. These maps should include the effects of realistic rupture parameters, and anelastic attenuation as well.

### 3.3. Velocity records

The synthetic velocity seismograms were low-pass filtered to frequencies less than 0.5 Hz (Butterworth filter with four poles and two passes). There were three sets of records – 118° (*X*-direction), 28° (*Y*-direction), and vertical (*Z*-direction) – for each simulation, which brings the total number of processed sets to 18. Each velocity record corresponds to a grid point of the basin's modeled region that covers a total area of 115 km × 95 km. The total number of grid points on the ground surface is 301 × 216, resulting in 195,048 particle velocity surface records for each earthquake

<sup>2</sup> Amplification is quantified as the peak velocity obtained from the 3D simulation divided by that predicted using a regional one-dimensional (1D) crustal model.

scenario simulation. Each record has a 90 s duration (125 s for the Northridge records), and a sampling rate  $dt = 0.25$  s.

The maximum absolute horizontal particle velocity magnitudes range for the horizontal directions from 0.767 m/s (San Andreas NW) to 4.168 m/s (Newport–Inglewood) and for the vertical from 0.286 m/s (San Andreas NW) to 4.278 m/s (Santa Monica). Since the narrow-band frequency content of the simulated time histories is limited to frequencies less than 0.5 Hz, and the near-surface *S*-wave velocity's upper bound is 1.0 km/s, these values are estimated to be lower bounds of the expected ground motion velocities in the near source area. Future simulations that would result from signals with a wider frequency content, and more realistic values for the surface *S*-wave velocity, would give rise to higher ground velocities and accelerations.

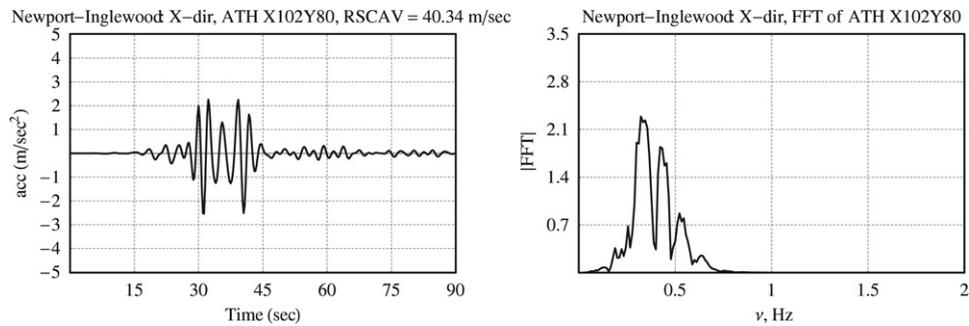
### 3.4. Acceleration records

To derive the ground acceleration time histories, a numerical differentiation by fitting a moving three-point quadratic polynomial ( $y = a_0 + a_1x + a_2x^2$ ) to the ground velocity record data was performed. A three-point analytical fit was chosen instead of a five-point fit ( $y = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4$ ), which, intuitively, would have resulted in higher accuracy, because the accuracy of the three-point fit proved to be adequate when applied to the low frequency velocity signals (4 samples/s).

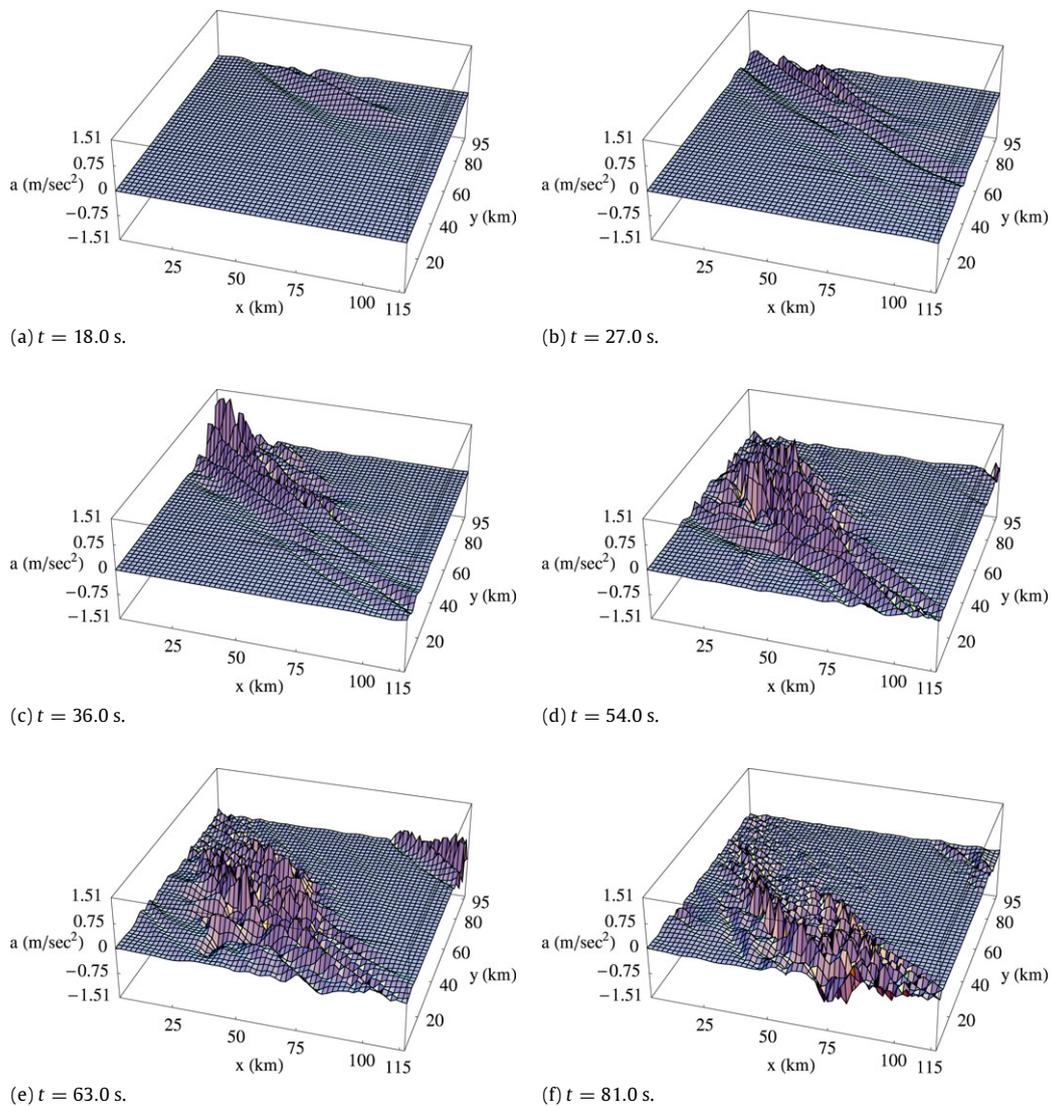
The cumulative absolute velocity (CAV), a severity indicator of an acceleration time history, was calculated for each record. The CAV, defined as the area under the absolute acceleration versus time duration curve, is given by  $CAV = \int_0^{t_{\max}} |\alpha(t)| dt$ . The three acceleration records with the maximum root square of the sum of the squares (RSSS) CAV ( $\sqrt{CAV_X^2 + CAV_Y^2 + CAV_Z^2}$ ) magnitude of 40.34 m/s resulted from the Newport–Inglewood dataset. The *X*-direction acceleration record and the corresponding fast Fourier transform (FFT) diagram are given in Fig. 3. The narrow-band nature of the signal (< 0.5 Hz) is evident in the FFT plot. Particle acceleration propagation snapshots of the San Andreas, SE propagating rupture (*X*-direction) are shown in Fig. 4.

## 4. Synchronous and asynchronous covariances

The temporal covariance matrices of order 360 × 360 (500 × 500, for the Northridge sets) extending over a time span of 90 s (125 s, for the Northridge sets) were constructed from the velocity and acceleration record ensembles. In general, the order of the covariance matrices is not constant, but depends on the number of samples in the signal. The covariance matrices were calculated twice: first, from the unprocessed, unsynchronized records, and second, from records that were synchronized at the trigger threshold value of 1% rms for each ensemble.



**Fig. 3.** Acceleration time history (left side) and its FFT diagram (right side) of the record with the largest root square of the sum of the squares (RSSS) cumulative absolute velocity (CAV) value. The record is from the Newport–Inglewood X-direction dataset.



**Fig. 4.** Six 3D particle acceleration snapshots from the San Andreas, SE propagating rupture (X-direction). The horizontal axes are the basin's kilometric distances and the vertical axis is the acceleration amplitude.

The velocity ground motion records were asynchronous because they captured the wave propagation effect through the extended Los Angeles region. The synchronization of records was performed by shifting the velocities and accelerations in the time domain. The rms value for each record ensemble was calculated, and the first sample ( $t = 0$ ) of each record was set to 1% of the ensemble rms. It was expected the record synchronization would result in a conservative estimate for the covariance amplitudes.

The synchronization of the ground motion records, indeed, gave rise to slightly larger covariance amplitudes as expected. This is clearly evident in Fig. 5(a–f), where the two variance time histories which depict the time variations of the mean-square accelerations (combined X, Y, and Z direction) are plotted for all six simulations. The covariance amplitudes from the synchronized records, besides being amplified, were also time-shifted in the case of the San Andreas NW (Fig. 5a), and San Andreas SE (Fig. 5b) scenarios.

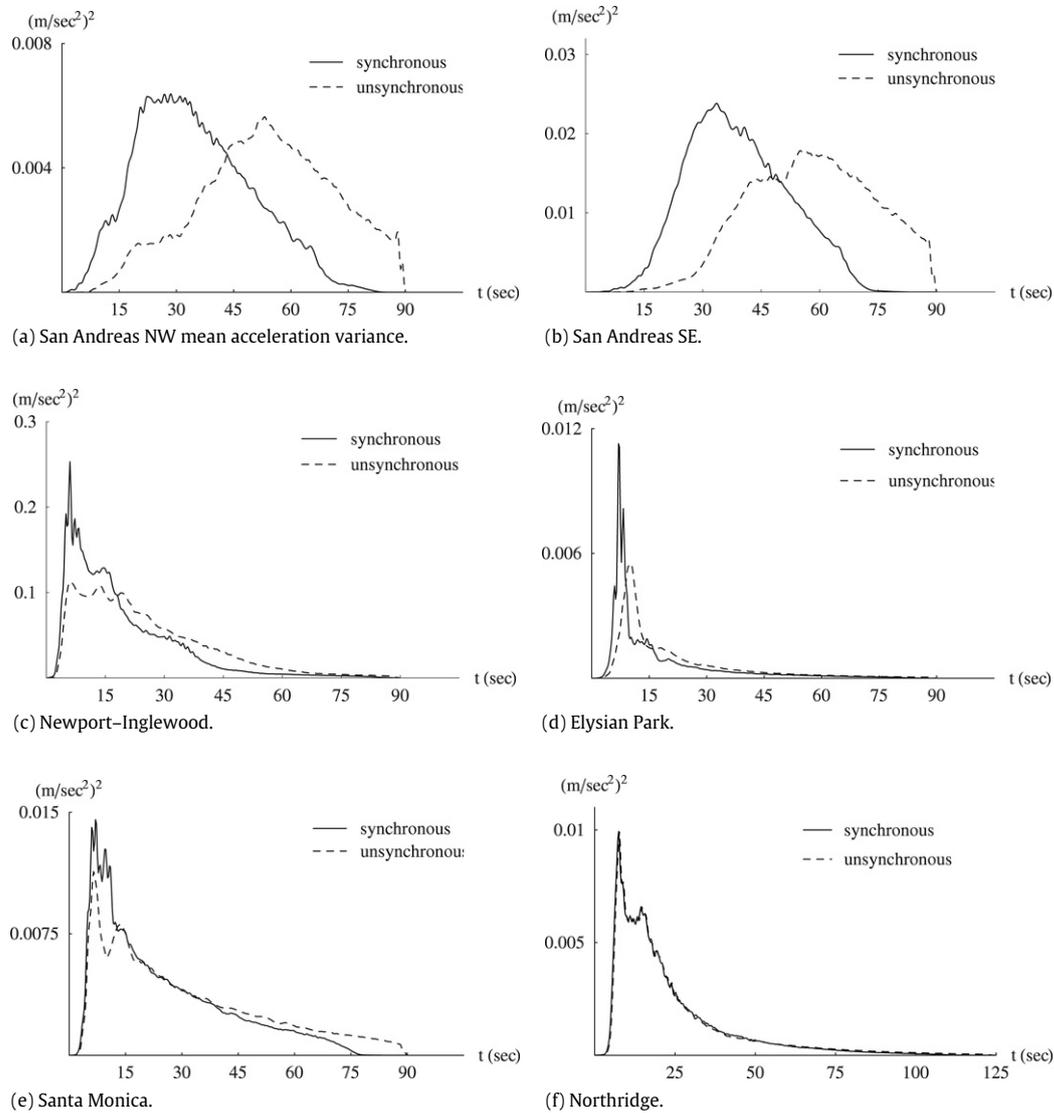


Fig. 5. Comparison of the synchronous (solid lines) and asynchronous (dashed lines) acceleration variances for all six scenario earthquakes.

These synchronization effects, which are directly related to the wave propagation patterns in the basin model, were due to the extended rupture duration, in excess of 60 s, that resulted in a much prolonged basin ground motion excitation (Olsen [25]).

On the other hand, for the Northridge simulation (Fig. 5f), where the duration of shaking in the basin was short, with most of the energy being dissipated in the first 30 s of excitation, the two variance curves coincided. The largest mean variance amplitude of the particle accelerations ( $0.228 \text{ (m/s}^2\text{)}^2$ ) was calculated from the Newport–Inglewood scenario datasets (Fig. 5c). The particle acceleration covariance matrices – synchronous and asynchronous – for the San Andreas SE X, Y, and Z datasets are plotted in Fig. 6(a–f). As expected, the covariance surfaces are symmetrical with respect to their diagonal, where the largest amplitudes are located; the off-diagonal terms converge rapidly to zero.

#### 4.1. Eigenvector expansion of the covariance matrices

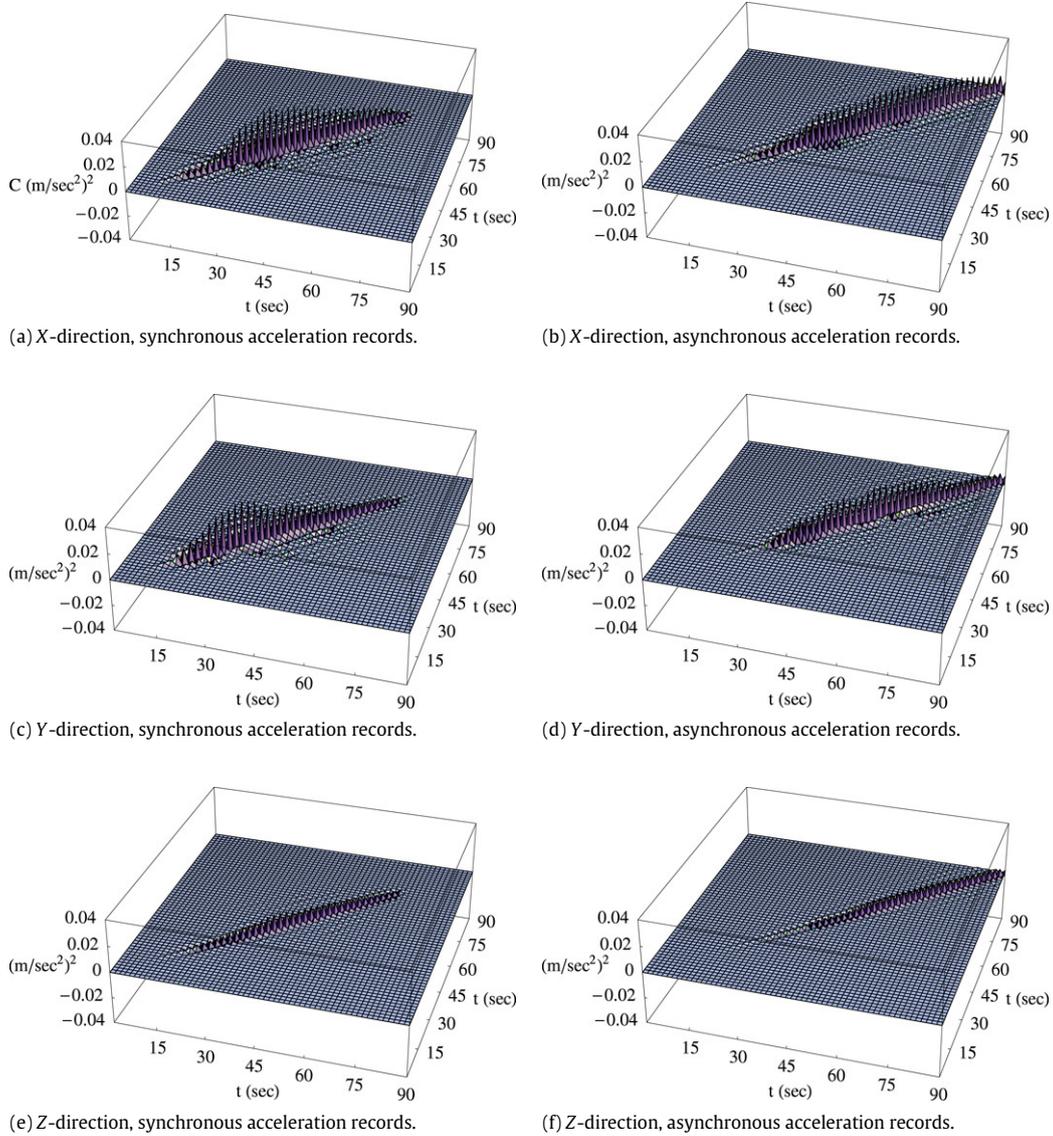
The spectral decomposition of the temporal particle velocity and acceleration covariances resulted in 360 eigenvalues  $\lambda_k$  (500 eigenvalues for the Northridge case), and their corresponding eigenvectors  $e_k$ . Although the main focus of the following discussion is placed on the acceleration covariances – subsequently, the

acceleration covariances are included in the non-stationary response solution for an MDOF linear system formulation – similar observations could be applied to the analytical behavior of the velocity covariances.

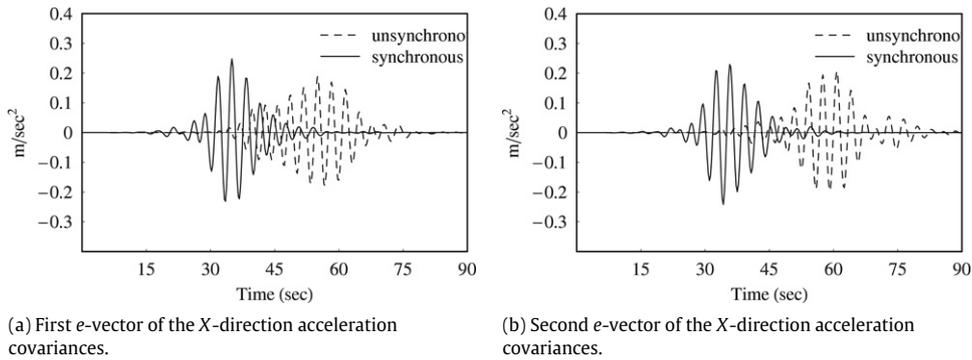
The first two eigenvectors of the synchronous and asynchronous acceleration covariances from the San Andreas SE record ensemble (X-direction) are presented in Fig. 7(a, b). The magnitude amplification and the time-shift observed by the comparison of the two covariance surfaces (Fig. 6) are also apparent in the eigenvector plots.

Furthermore, the eigenvalue series  $\lambda_k, k = 1360$  converges to zero after the first  $\approx 50$  values (Fig. 8). All the eigenvalues  $\lambda_k$  with indices  $k > 30$  have magnitudes less than 10% of the first eigenvalue  $\lambda_1$ , as illustrated by the steep convergence curve of the normalized eigenvalues—normalized with respect to the first (largest) eigenvalue. Therefore, for the first step of data compaction, only the significant eigenvalues and eigenvectors, out of the original 360, are needed to reconstruct the acceleration covariance matrices.

Truncated eigenvalue series  $\lambda_k, k = 160$  from the six earthquake datasets are presented in Fig. 9. Note that different amplitude (ordinate) scales are used for the various earthquakes. All the eigenvalues beyond the cutoff index of 40 (15 for the Elysian Park case) are practically zero. The total area under



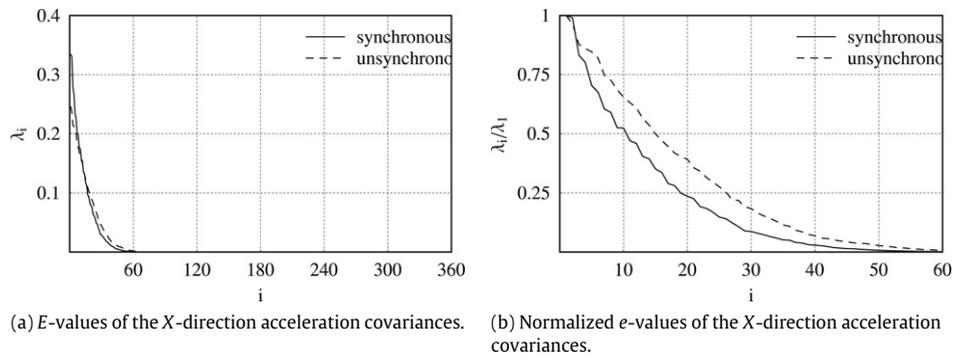
**Fig. 6.** Comparison of the covariance matrices for the synchronous (left side: a, c, e) and asynchronous (right side b, d, f) acceleration records for the San Andreas, SE propagating rupture scenario earthquake.



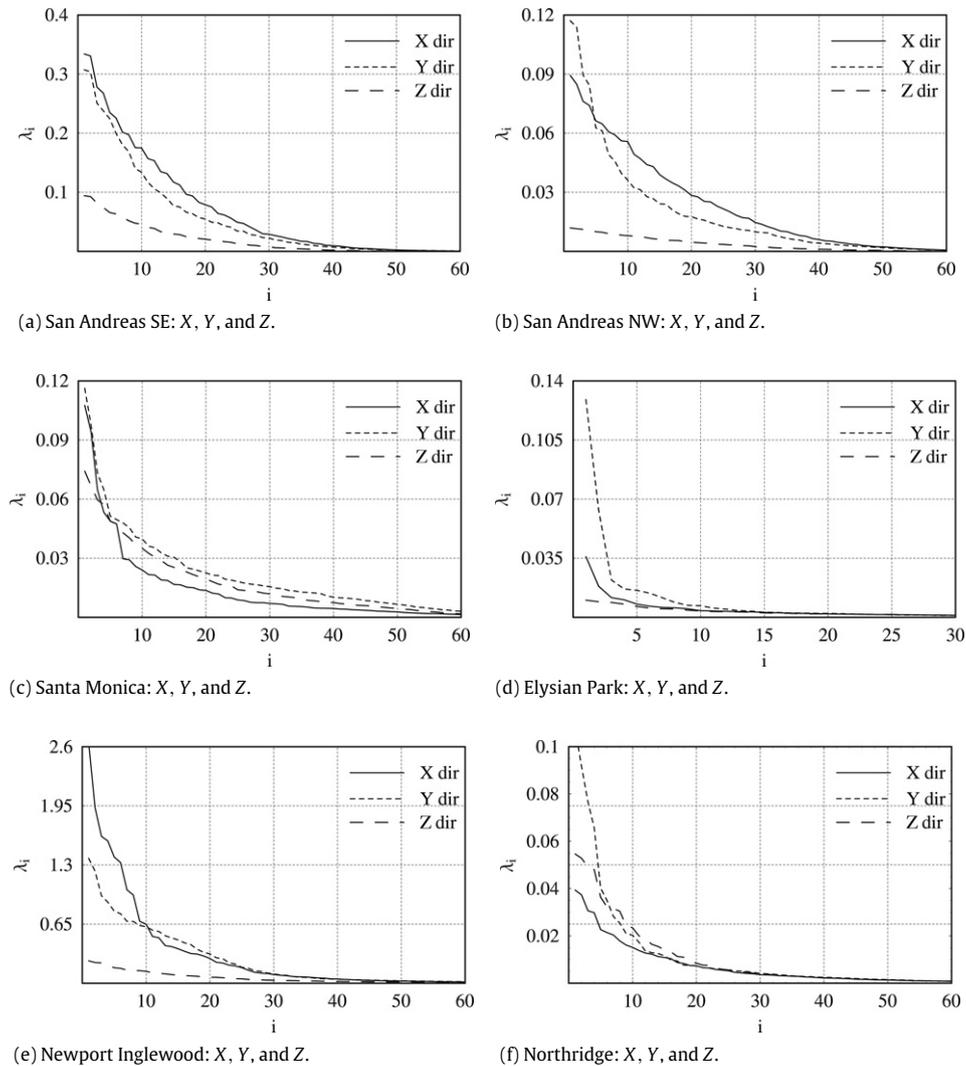
**Fig. 7.** (a) and (b): The time histories of the first two eigenvectors of the synchronous (solid lines) and asynchronous (dashed lines) acceleration covariances from the San Andreas SE (X-direction) record ensemble.

the individual eigenvalue curves for each X, Y, and Z direction is a qualitative estimator of the three-dimensional earthquake energy distribution in the basin. For example, the superimposed covariance eigenvalue plots from the Santa Monica scenario

(Fig. 9c) depict an equipartition of energy for the basin’s horizontal directions. In addition, significant energy distribution in the vertical Z direction is observed for the Santa Monica (Fig. 9c) and Northridge (Fig. 9f) simulations.



**Fig. 8.** Part (a) shows the magnitudes of the eigenvalues of the synchronous and asynchronous acceleration covariances from the San Andreas SE ( $X$ -direction) record ensemble. Fig. (b) shows the first 60 normalized eigenvalues.



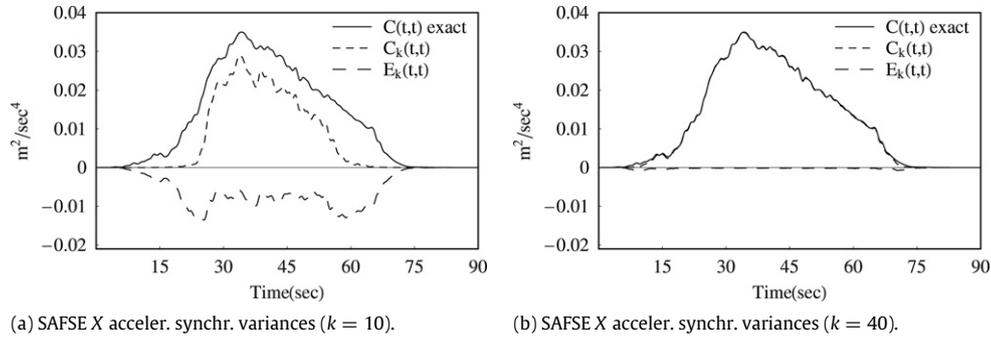
**Fig. 9.** Comparison of eigenvalues' amplitudes of the  $X$ ,  $Y$ , and  $Z$  direction synchronous acceleration covariances for the six earthquake datasets plotted as a function of the eigenvalue index  $i$ . Note that different ordinate scales are used for the different scenarios.

**4.2. Covariance reconstruction and eigenvector approximation with Chebyshev polynomials**

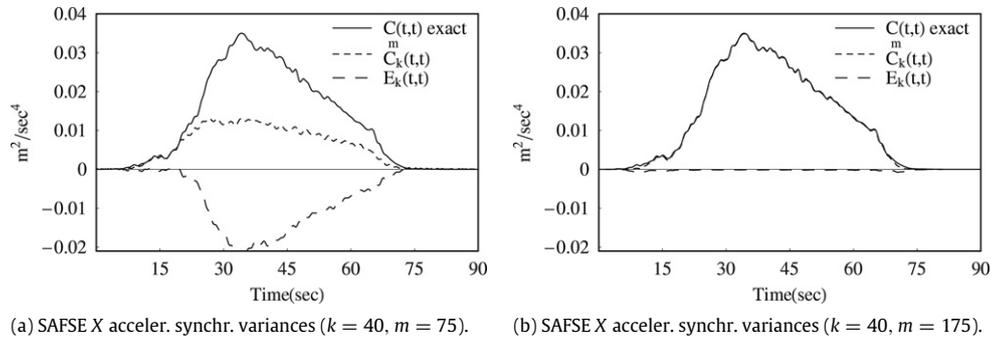
To demonstrate the quality of fit for the covariance reconstruction, the diagonals of the original acceleration covariance matrix  $[C]$ , and the reconstructed  $[C_k]$ , where  $k$  is the number of eigenvalues/eigenvectors that define the truncation level, are plotted in Fig. 10(a, b). As  $k$  increases ( $k = 10$  for Fig. 10a, and  $k = 20$

for Fig. 10b), the quality of the variance fit improves, and the error function  $E_k = [C] - [C_k]$  gradually decreases; for  $k = 40$  the error function  $E_k$  approximates zero, and the two variances coalesce. The inclusion of eigenvectors/eigenvalues beyond the first 40 would not have further improved the quality of the reconstructed covariance.

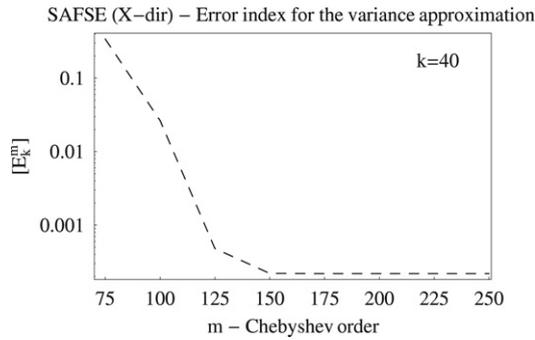
The second data compaction step includes the least-squares fit of the truncated eigenvectors series with Chebyshev polynomials.



**Fig. 10.** Time histories of exact variances  $C(t, t)$ , reconstructed variances  $C_k(t, t)$ , and error plots  $E_k(t, t)$  plotted for eigenvector truncation parameter  $k = 10$  (a) and  $k = 40$  (b), for the San Andreas SE X-direction synchronous acceleration records.



**Fig. 11.** Time histories of exact variances  $C(t, t)$ , reconstructed variances  $C_k^m(t, t)$ , and error plots  $E_k^m(t, t)$  plotted for the same eigenvector truncation parameter  $k = 40$ , and different values of the Chebyshev order  $m$  ( $m = 75$  (a),  $m = 175$  (b)), for the San Andreas SE X-direction synchronous acceleration records.



**Fig. 12.** Logarithmic plot showing the correlation of the normalized error index between the exact and reconstructed variance for increasing values of the Chebyshev polynomial order  $m$  (San Andreas SE, X-direction dataset).

The optimality of the Chebyshev polynomial fit is illustrated in Fig. 11(a–b), where the exact  $C(t, t)$  and the reconstructed variances  $C_k^m(t, t)$ ,  $k = 40$  are plotted for increasing values of the polynomial order ( $m = 75$ , and 175). The optimal value for the order of the polynomial fit was  $m = 175$ , as indicated in Fig. 12, where the log of the summation error function  $[E_{k=40}^m] = \Sigma ([C] - [C_k^m])^2$  versus the polynomial order  $m$  is plotted. If the eigenvectors had a wider frequency content, a higher Chebyshev polynomial order would have been required to achieve the optimum quality of fit. For the present study, an increase of the current Chebyshev order  $m = 175$  would not have improved the analysis accuracy, but would only have increased the computational time.

### 5. Construction of response spectra

The transient mean-square response time histories of a single-degree-of-freedom (SDOF) system with a critical damping ratio of  $\zeta = 0.01, 0.05$ , and  $0.10$  were obtained from the closed-form analytical expression for the response covariance by setting

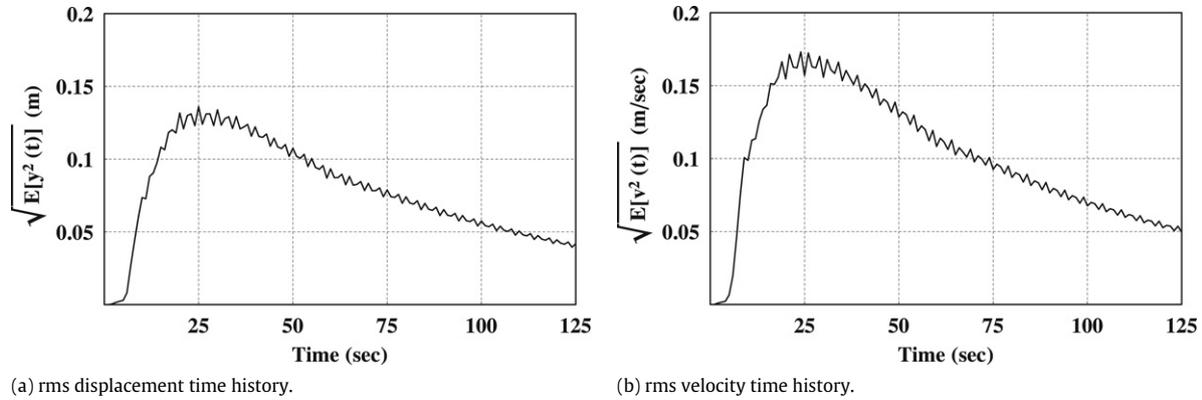
$t_1 = t_2 = t$ . Note that the analytical solution provides a measure of the variance (i.e., uncertainty with respect to the mean). If the ensemble mean is zero, then the root-mean-square (rms) is the same as the response standard deviation. The six scenario earthquake approximating covariances, used as input, were reconstructed with the first 40 eigenvectors of the original covariances that had been least-squares fitted by Chebyshev polynomials of order  $m = 175$ . A sample root-mean-square time history of a representative SDOF with a natural frequency of 0.2 Hz and critical damping ratio of  $\zeta = 0.01$  is shown in Fig. 13 (a,b). Fig. 13 (a) depicts the rms values of the relative displacements, and Fig. 13 (b) depicts the rms values of the relative accelerations. The maximum values of the rms response of a damped SDOF system are given by

$$R_d(T, \zeta) \equiv \sqrt{\max_t \{E[y^2(t)]\}} \quad (7)$$

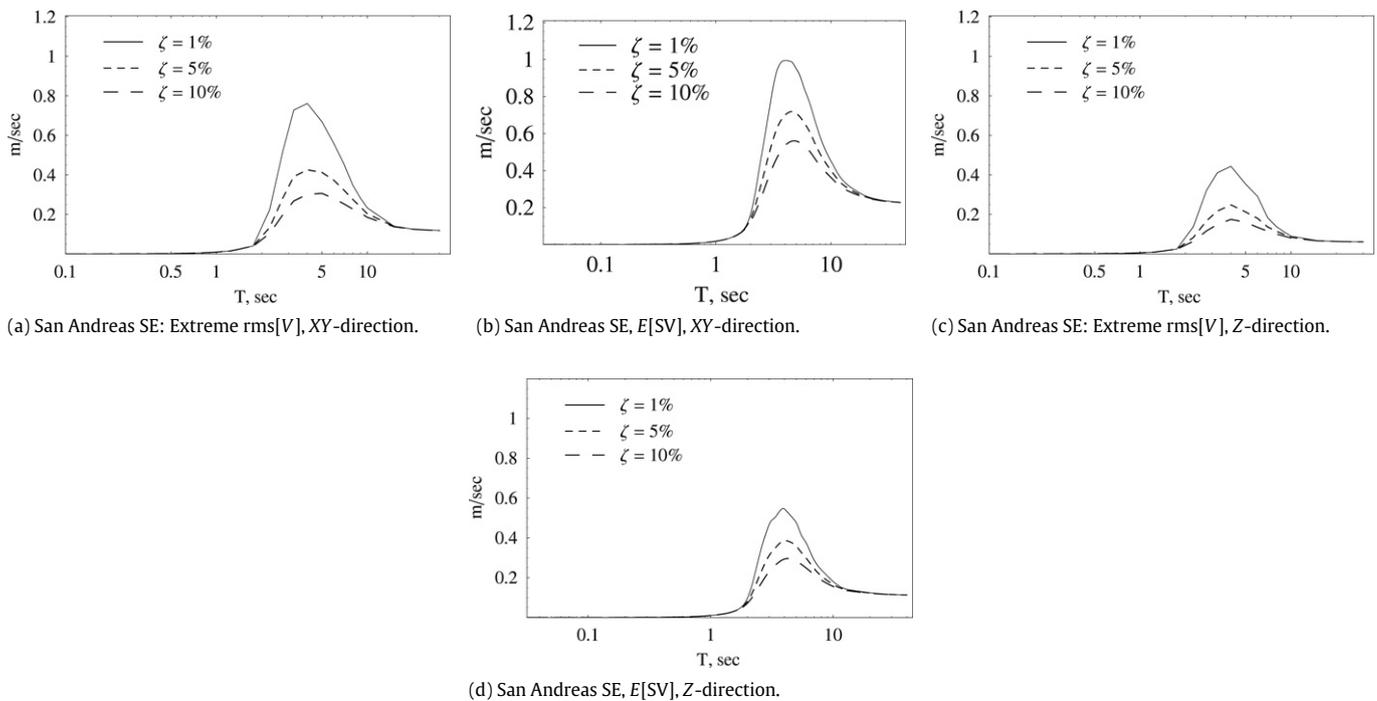
$$R_v(T, \zeta) \equiv \sqrt{\max_t \{E[\dot{y}^2(t)]\}} \quad (8)$$

where  $E[y^2(t)]$  and  $E[\dot{y}^2(t)]$  are the mean-square displacement, and velocity response time histories, respectively.

For each scenario earthquake, the  $E[y^2(t)]$  and  $E[\dot{y}^2(t)]$  response time histories were calculated with input acceleration covariances derived from the horizontal (combined X and Y direction), and vertical (Z direction) datasets. The computed maximum values  $R_d$  and  $R_v$  of the rms responses were plotted versus period values  $T$ . A total of 24 maximum rms spectral plots (6 scenarios  $\times$  2 directions  $\times$  2 response variables) were obtained. Representative plots for the San Andreas SE scenario are shown on the left side column in Figs. 14 and 15. The three curves in each plot correspond to damping  $\zeta = 0.01, 0.05$ , and  $0.10$ . The amplitudes of the response values for all spectral plots are significant for periods larger than 2 s. This result is expected due to the limited frequency bandwidth ( $n < 0.5$  Hz) of the input acceleration records.



**Fig. 13.** Transient root-mean-square response of an SDOF system with frequency of 0.2 Hz, and critical damping ratio of 1%, for the Northridge (X-direction) scenario earthquake.



**Fig. 14.** Response spectra for maximum rms velocities rms[V] (left side), and mean peak velocities SV (right side) for the combined X and Y horizontal (a and b), and vertical Z (c and d) datasets from the San Andreas SE scenario earthquake.

Furthermore, for each synthetic acceleration record, the response of an SDOF linear system subjected to input acceleration was calculated by solving analytically the classic differential equation of motion in a computationally efficient fashion<sup>3</sup>:

$$\ddot{x} + 2\zeta\omega\dot{x} + \omega^2x = f(t). \quad (9)$$

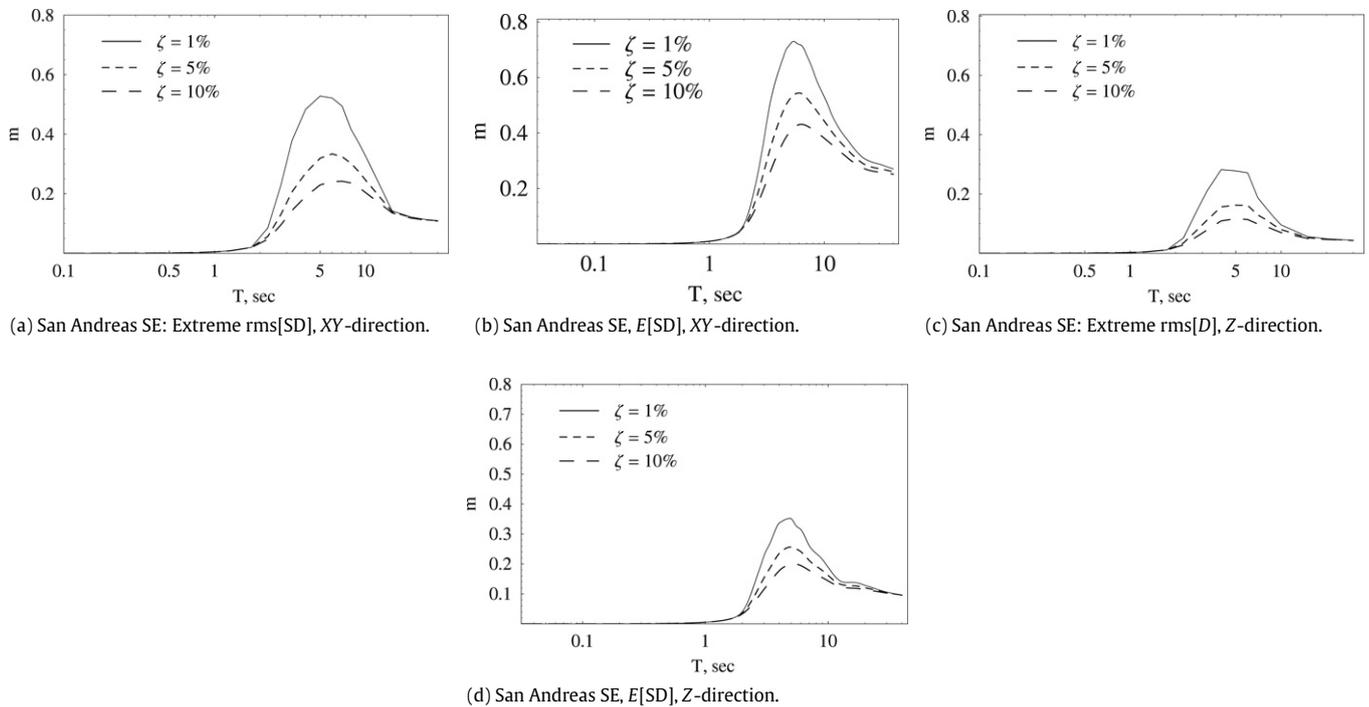
The peak absolute values of the response time histories (SA: absolute acceleration, SV: relative velocity, and SD: relative displacement) were plotted versus period values  $T$ , and damping ratios  $\zeta = 0.01, 0.05$ , and  $0.10$  to form the response spectra. For each of the six scenario record ensembles, the mean SA, SV, and SD values were thus calculated. The horizontal datasets (X and Y direction) were combined in one group to comply with the data processing scheme followed for the maximum rms spectra calculation. The size of each record ensemble (X, Y, and Z direction)

was  $n = 65,016$  for the Northridge, and  $n = 68,256$  for the remaining scenarios. Representative mean response spectra plots for the San Andreas SE scenario are shown on the right side columns in Figs. 14 and 15. For ease of comparison, identical amplitude and period scales are used for the two types of response plots.<sup>4</sup>

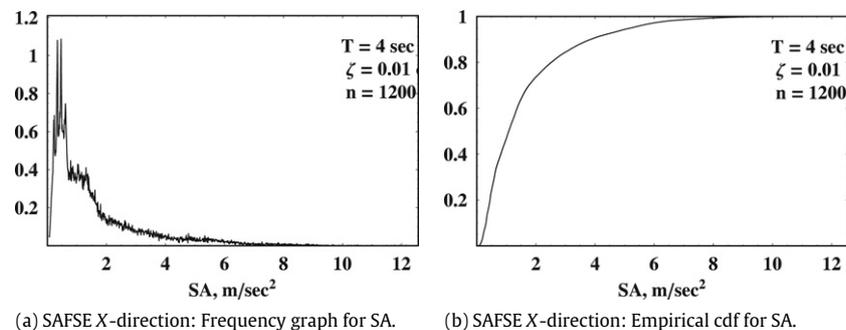
It is worth keeping in mind that the two response plots (left and right columns of Figs. 14 and 15) are qualitatively different: one tracks the maximum rms values of the nonstationary responses obtained at one specific time instant (or else the maximum standard deviation of the response), while the other provides an average of the response peaks, each one occurring at a different time. The question that arises is which of the two plots an engineer could use to estimate the expected peak responses (velocities, displacements or accelerations) for his design. In general, one

<sup>3</sup> The analytical solution of the second-order differential equation is given in Nigam and Jennings [26]

<sup>4</sup> Plots for the mean peak and maximum root-mean-square responses are available for the remaining scenarios, but are not included in this paper due to size constraints.



**Fig. 15.** Response spectra for maximum rms displacements rms[D] (left side), and mean peak displacements SD (right side) for the combined X and Y horizontal (a and b), and vertical Z (c and d) datasets from the San Andreas SE scenario earthquake.



**Fig. 16.** Frequency and cumulative distribution frequency (cdf) diagrams for the peak accelerations of a linear oscillator with a period of 4 s and damping  $\zeta = 0.01$  for the San Andreas SE X-direction dataset).

would like to multiply either the rms or the average peak response with amplification factors to get an estimate of the expected maximum response. But without attaching a probability distribution model to the response peaks, it is not possible to establish confidence intervals for the expected value of the maximum response.

## 6. Statistical inference of the response spectra variables

In order to construct probabilistic response spectra with a prescribed confidence interval, first the appropriate probability model that could best describe the distribution characteristics of the peak response variables (accelerations, velocities, and displacements) needs to be determined. Frequency and cumulative distribution diagrams of the peak response variables for  $k = 1200$  intervals were constructed for all data samples. The diagrams of a linear oscillator having a natural period of 4 s ( $\zeta = 0.01$ ) are presented in Fig. 16. An initial visual inspection of the sample histograms indicates that the assumption of a Log-Normal probability distribution model for the peak response variables is plausible. Similar histogram shapes (that are not presented herein) were observed for the peak responses of oscillators with other

period values. Furthermore, the observed response sample data were plotted versus the linear probability scales of the following assumed distributions: Log-Normal, Weibull, Frechet, and Gamma. The data points should follow a linear trend, if the assumed distribution is an appropriate model for the data.

Based on the quality of the linearity that these probability plots exhibited, the following distributions were chosen as the most promising candidates for a potential suitable probability model for the spectral response variables of interest: (a) Two-parameter Log-Normal; (b) Three-parameter Log-Normal; (c) Two-parameter Weibull; (d) Two-parameter Gamma.

### 6.1. Parameter estimation for the candidate distributions

The optimal parameter estimators of the chosen distributions were determined with the *maximum likelihood* (ML) method (Ang and Tang [27]).<sup>5</sup> The final parameters of the distributions were

<sup>5</sup> The initial estimation of the distribution parameters for each sample of the observed frequencies was based on the formulas given in [27].

**Table 3**  
Estimated parameters and normalized least-square errors for the Northridge (X-direction) response spectra variables.

T (s)	Earthquake approximation: Northridge (X-direction)									
	Log-Normal (two-parameter)					Log-Normal (three-parameter)				
	$\mu$	$\sigma$	$\chi$ (a)	$\epsilon$ (b)	$\delta$	$\mu$	$\sigma$	$\chi$	$\epsilon$	
SA (1% damping)										
1	-2.614	1.136	102.262	0.121	0.005	-2.700	1.210	95.057	0.111	
2	-1.244	1.100	11.454	0.053	0.014	-1.329	1.171	10.451	0.05	
3	-1.128	0.990	12.183	0.037	0.029	-1.216	1.063	11.569	0.038	
4	-1.526	0.999	14.855	0.038	0.023	-1.592	1.062	13.276	0.029	
5	-1.986	1.034	56.865	0.144	0.010	-2.079	1.137	54.858	0.127	
6	-2.275	0.928	29.128	0.035	0.018	-2.39	1.034	27.632	0.028	
7	-2.678	0.906	65.923	0.053	0.009	-2.847	1.039	43.946	0.02	
8	-3.132	0.886	95.185	0.040	0.005	-3.300	1.042	65.317	0.019	
9	-3.503	0.849	218.087	0.073	0.004	-3.676	0.996	289.162	0.109	
10	-3.879	0.844	315.602	0.052	0.002	-4.027	0.954	291.101	0.056	
SV (1% damping)										
1	-5.462	1.142	1403.72	0.114	0.0003	-5.562	1.234	1312.550	0.109	
2	-2.486	1.108	34.881	0.047	0.004	-2.563	1.186	32.367	0.046	
3	-1.864	0.993	25.386	0.034	0.011	-1.943	1.064	24.357	0.036	
4	-1.939	1.002	22.683	0.037	0.007	-2.015	1.074	20.447	0.028	
5	-2.136	1.040	77.20	0.160	0.009	-2.227	1.154	77.230	0.147	
6	-2.196	0.941	34.383	0.054	0.008	-2.305	1.041	32.036	0.05	
7	-2.394	0.921	43.697	0.058	0.012	-2.563	1.067	32.612	0.035	
8	-2.629	0.942	78.356	0.095	0.008	-2.784	1.060	61.673	0.075	
9	-2.761	0.923	108.037	0.093	0.005	-2.894	1.032	107.575	0.109	
10	-2.915	0.947	106.378	0.053	0.005	-3.057	1.055	96.735	0.055	
SD (1% damping)										
1	-6.278	1.124	4091.5	0.123	0.0001	-6.381	1.213	3781.684	0.109	
2	-3.540	1.100	113.252	0.053	0.001	-3.611	1.171	103.326	0.046	
3	-2.603	0.990	51.847	0.037	0.007	-2.690	1.064	50.571	0.036	
4	-2.440	0.999	36.817	0.037	0.004	-2.505	1.085	34.137	0.028	
5	-2.450	1.034	90.477	0.143	0.007	-2.542	1.137	87.064	0.147	
6	-2.369	0.928	32.569	0.035	0.016	-2.483	1.035	30.693	0.05	
7	-2.455	0.894	52.876	0.052	0.012	-2.636	1.039	35.430	0.035	
8	-2.652	0.899	59.358	0.042	0.012	-2.803	1.003	39.388	0.075	
9	-2.789	0.862	106.748	0.076	0.007	-2.963	0.996	137.151	0.109	
10	-2.948	0.844	121.766	0.051	0.009	-3.084	0.926	108.996	0.055	

(a)  $\chi = \sum_{i=1}^n \frac{(p_i - e_i)^2}{p_i}$ ; (b)  $\epsilon = \frac{\sum_{i=1}^n (p_i - e_i)^2}{\sum_{i=1}^n p_i^2}$

chosen from a parametric optimization process which minimized the *chi-square* quantity

$$\chi = \sum_{i=1}^k \frac{(n_i - e_i)^2}{e_i} \tag{10}$$

where  $n_1, n_2, \dots, n_k$  ( $k = 1200$ ) were the observed frequencies, and  $e_1, e_2, \dots, e_k$  ( $k = 1200$ ) the assumed theoretical frequencies. For each tested distribution, the error quantity  $\chi$  was calculated for the parameter values taken from the following intervals:

Log-Normal:  $0.85 \hat{\sigma} \leq \sigma \leq 1.15 \hat{\sigma}$ ;  $0.85 \hat{\mu} \leq \mu \leq 1.15 \hat{\mu}$ ;  
 $\hat{\delta} \leq \delta \leq 2.0 \hat{\delta}$

Weibull:  $0.7 \hat{\sigma} \leq \sigma \leq 1.0 \hat{\sigma}$ ;  $0.7 \hat{\lambda} \leq \lambda \leq 1.0 \hat{\lambda}$

Gamma:  $0.5 \hat{\sigma} \leq \sigma \leq 2.0 \hat{\sigma}$ ;  $0.5 \hat{\lambda} \leq \lambda \leq 2.0 \hat{\lambda}$

( $\hat{\sigma}, \hat{\mu}, \hat{\lambda}$ : initial parameter estimators;  $\hat{\delta}$ : min. of response spectra variable samples).

The optimization procedure used in this task was based on a trial-and-error process. The interval limits were chosen through an iterative process, which predetermined the best plausible range of values for minimizing the quantity  $\chi$ . The optimum parameter estimators for a Log-Normal (two and three parameters) model, along with the error quantities  $\chi$  and  $\epsilon$ , for the Northridge X-direction response spectra variables are presented in Table 3. For all response samples, the smallest values of the error quantities were for the period values between 2 and 4 s. In particular, by comparing the minimum error quantities, among the assumed candidate distributions – Log-Normal, Weibull, and Gamma – the

Log-Normal was found to be the most suitable for the observed frequencies of all three response samples: spectral accelerations, velocities, and displacements. The application of a three-parameter Log-Normal distribution reduced the error quantities, mostly for the  $2 \text{ s} \leq T \leq 5 \text{ s}$  period range, but overall did not significantly improve the quality of fit.

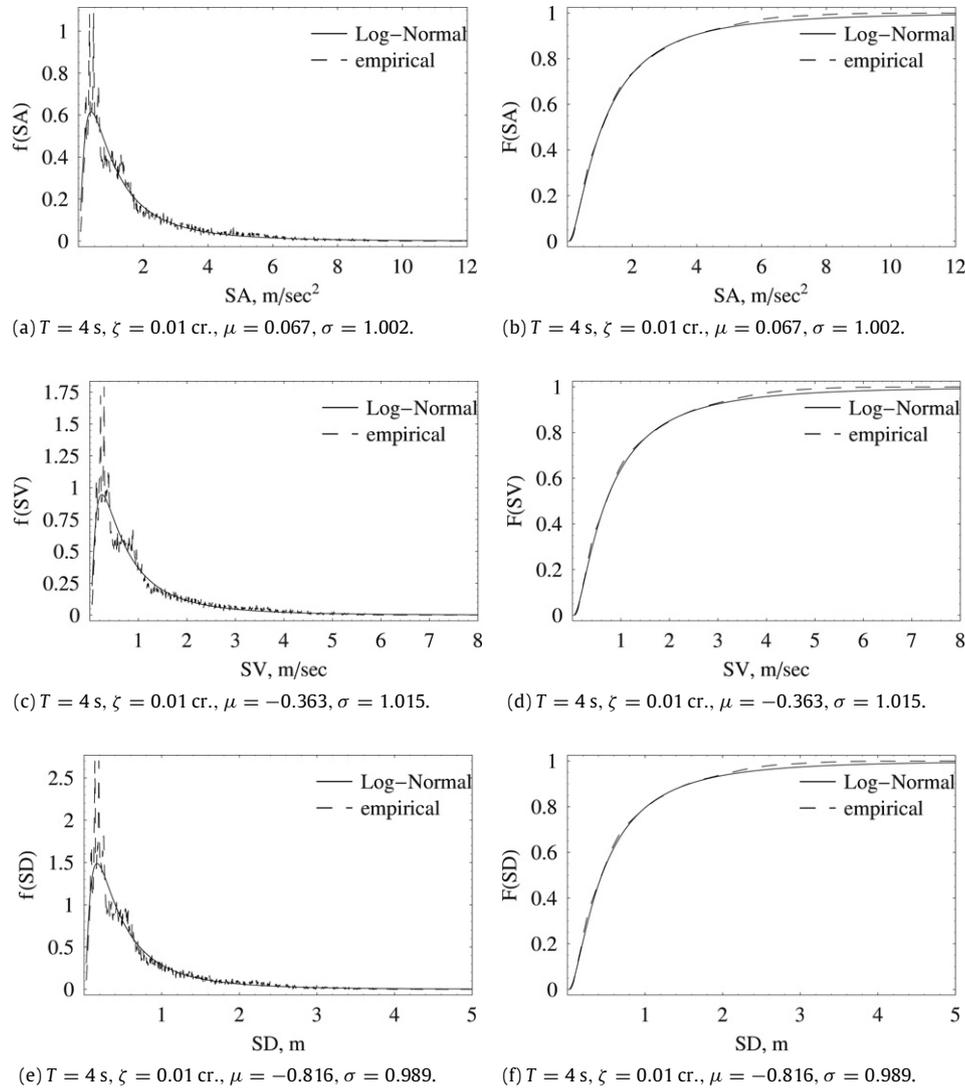
Furthermore, the validity of the distributions was statistically verified by the *chi-square* test. The distribution of the quantity  $\chi$  Eq. (12) as  $k \rightarrow \infty$  approaches the chi-square ( $\chi_f^2$ ) distribution with ( $f = k - 1$ ) degrees of freedom, where  $k$  is the number of pairs for the observed and expected frequencies (Hoel, Chapter 9) [28];  $k = 1200$  in the present study. If an assumed distribution yields

$$\chi = \sum_{i=1}^k \frac{(n_i - e_i)^2}{e_i} < c_{1-\alpha, f} \tag{11}$$

where  $c_{1-\alpha, f}$  is the value of the appropriate  $\chi_f^2$  distribution at the cumulative probability  $(1 - \alpha)$ , the assumed theoretical distribution is an acceptable model, at the *significance level*  $\alpha$ . The  $c_{1-\alpha, f}$  values of the  $\chi_f^2$  distribution with  $f = 1197$ , and  $f = 1196$  degrees of freedom for a two-parameter and three-parameter assumed distribution respectively are presented in Table 4.<sup>6</sup>

All the assumed distributions yielded  $\chi$  values that were less than the  $\chi_f^2$  ( $f = 1196$  and  $1197$ ) variables of Table 4. Therefore,

<sup>6</sup> The degree of freedom  $f$  of the  $\chi_f^2$  distribution is reduced by one for every unknown parameter that must be estimated [27].



**Fig. 17.** Comparison between Log-Normal, and empirical probability density function (pdf) and cumulative distribution function (cdf) plots for the spectral response measures of a linear oscillator having a natural period of 4 s (damping  $\zeta = 0.01$ ; the dataset is San Andreas SE X-direction). The left sides of the plots correspond to the pdfs of the indicated spectral response measures, while the right sides of the plots correspond to the respective cdfs. The first row shows spectral accelerations SA; second row, spectral velocities SV; third row, spectral displacements SD.

**Table 4**  
 $\chi_f^2$  cumulative distribution values.

	$\chi_{f=1,196}^2$			$\chi_{f=1,197}^2$		
$\alpha$	20%	10%	5%	20%	10%	5%
$c_{1-\alpha, f}$	1236.96	1259.09	1277.57	1237.97	1260.12	1278.6

the four tested distributions are shown to be valid models for the spectral accelerations (SA), velocities (SV), and displacements (SD) of a linear SDOF oscillator under the excitation of the Northridge, and San Andreas SE record ensembles.

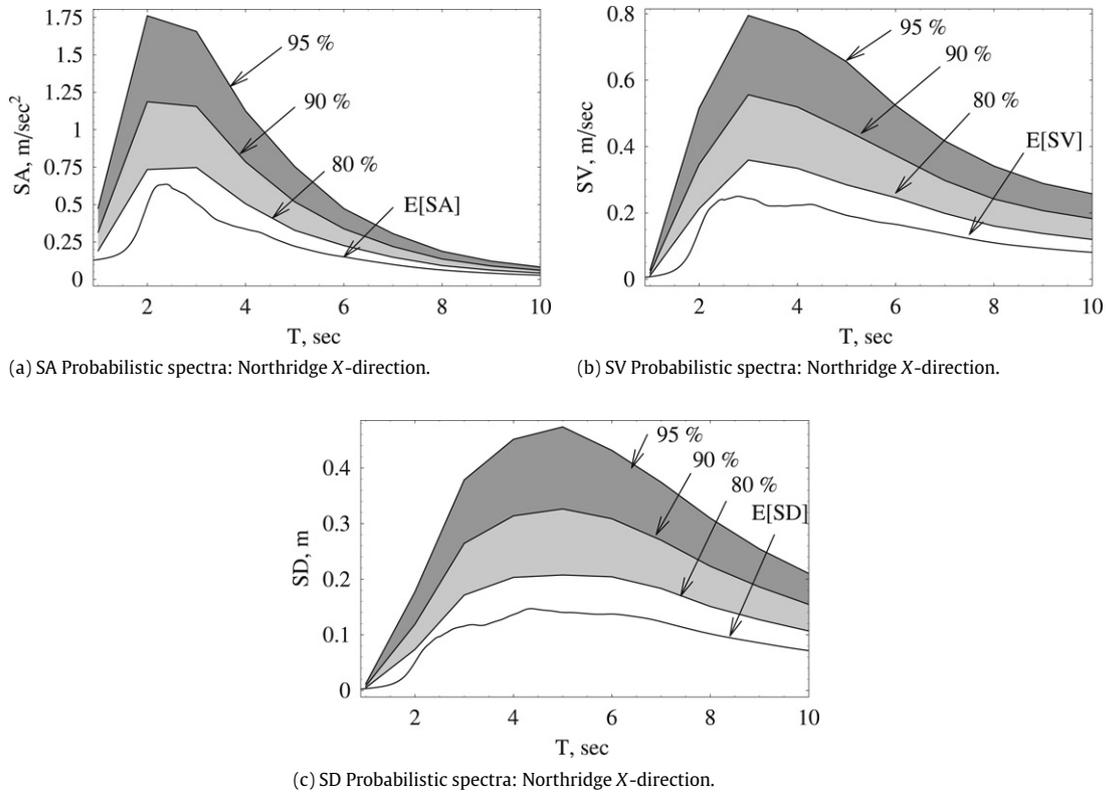
6.2. Probabilistic response spectra based on the Log-Normal family of distributions

The two-parameter Log-Normal model yielded the minimum  $\chi$  values over the period spectrum for both earthquake ensembles, and was selected to formulate the probabilistic curves of the peak response variables. In a related study, a Log-Normal model for the response quantities of a linear SDOF oscillator was proposed by Chopra ([29], Section 6.9). A comparison between the empirical

and the Log-Normal cumulative distribution functions for a 4 s SDOF oscillator validates the choice of the Log-Normal model (Fig. 17).

Furthermore, the 80%, 90%, and 95% percentile curves of the response variables (SA, SV, SD) based on the families of Log-Normal distributions were plotted for each record ensemble (Figs. 18 and 19). The three probabilistic response curves follow similar shape trends with the ensemble mean curves— $E[SA]$ ,  $E[SV]$ ,  $E[SD]$ . It should be noted that confidence intervals of the parameter estimators have not been calculated. For instance, the  $(1 - \alpha)$ -level confidence interval of the Log-Normal sample mean  $\mu$  estimator is given by

$$\left( \mu + t_{n-1, \alpha/2} \frac{\sigma}{\sqrt{n}}, \mu - t_{n-1, \alpha/2} \frac{\sigma}{\sqrt{n}} \right), \tag{12}$$



**Fig. 18.** Response curves of peak accelerations SA, peak velocities SV, and peak displacements SD over period values for the Northridge X-direction dataset. The curves correspond to mean peak values ( $E[SA]$ ,  $E[SV]$ ,  $E[SD]$ ) and percentiles: 80%, 90%, 95% of non-exceedance (damping  $\zeta = 0.01$ ).

where  $t_{n-1,q}$  denotes the  $q$ -quantile of the  $t$ -distribution with  $(n-1)$  degrees of freedom, and  $\sigma$  is the sample standard deviation. For example, for a 90% confidence interval for the sample mean  $\mu$ , the 5% Student- $t$  distribution quantile for  $n = 10,000$  dof is  $t(9999, 0.05) = -1.645$ ; thus, the 90% confidence interval of  $\mu$  is

$$\left( \mu - 1.645 \frac{\sigma}{\sqrt{9999}}, \mu + 1.645 \frac{\sigma}{\sqrt{9999}} \right). \quad (13)$$

Similarly, the  $(1-\alpha)$ -level confidence intervals for the Log-Normal sample standard deviation  $\sigma$  estimator is obtained from

$$\left( \sigma \sqrt{\frac{n-1}{\chi_{n-1,1-\alpha/2}^2}}, \sigma \sqrt{\frac{n-1}{\chi_{n-1,\alpha/2}^2}} \right) \quad (14)$$

where  $\chi_{n-1,q}^2$  is the  $q$ -quantile of the  $\chi^2$  distribution with  $(n-1)$  degrees of freedom. For example, for a 90% confidence interval for the sample standard deviation  $\sigma$ , the 5% and 95%  $\chi^2$  distribution quantiles for  $n = 10,000$  dof are  $\chi^2(9999, 0.05) = 9767.54$ , and  $\chi^2(9999, 0.95) = 10,232.7$ , respectively; thus, the 90% confidence interval of  $\sigma$  is

$$\left( \sigma \sqrt{\frac{9999}{10,232.7}}, \sigma \sqrt{\frac{9999}{9767.54}} \right). \quad (15)$$

Consequently, for a complete probabilistic approach to the construction of linear response spectra, the lower and upper bounds of a predetermined confidence interval (i.e., 90%, 95%, etc.) should be included in the probabilistic plots (Figs. 18 and 19).

A potential application of the regional probabilistic spectra in the area of seismic design is the derivation of amplification factors for the peak response variables. Amplification factors derived from regional probabilistic spectra can contribute to the evaluation of seismic design spectra associated with a certain level of confidence,

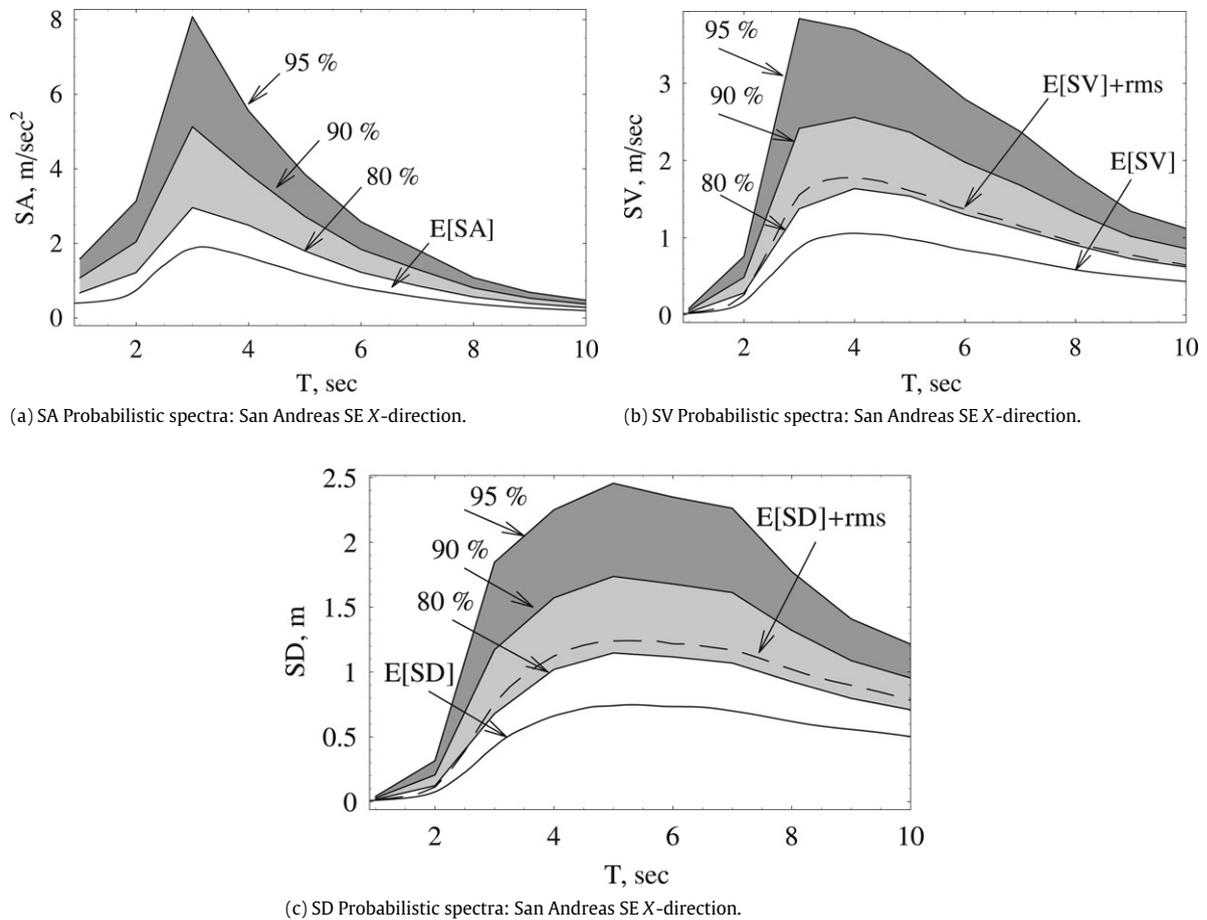
when multiplied by the peak response values derived, for example, from a single available earthquake record. A crude estimate for the peak response amplification factors is presented in Table 5. These amplification factors were based on the San Andreas SE X-direction ( $\zeta = 0.01$ ) probabilistic spectra plots. They were computed by adding the peak response values (SA, SV, and SD) corresponding to a given percentile curve (80%, 90%, and 95%) over the set of period values  $T = 2, 3, 4, 5, 6, 7$ , and  $8$ , s, and dividing by the sum of the mean peak values ( $E[SA]$ ,  $E[SV]$ , and  $E[SD]$ ) for the same set of period values. For example, the amplification factor of the peak acceleration SA for the 80% percentile curve ( $F_{SA}^{80\%}$ ) is

$$F_{SA}^{80\%} = \frac{SA_{T=2}^{80\%} + SA_{T=3}^{80\%} + \dots + SA_{T=8}^{80\%}}{E[SA]_{T=2} + E[SA]_{T=3} + \dots + E[SA]_{T=8}}. \quad (16)$$

Additional processing of the remaining earthquake scenarios would not contribute any further to the clarification of the underlying probability models for the response variables due to the simulated acceleration records' limited frequency spectrum, which yielded similar response spectra shapes for all the scenario datasets in this study. The response amplification factors are based on records from sites overlying diverse geological strata with sedimentary layers, rocks, etc. The spatial dependency of the amplification factors should be addressed in future studies by classifying the ground motion records in groups from sites that have similar geological properties.

### 6.3. Probabilistic association of rms response measures

To inquire about a probabilistic association between the maximum rms response quantities derived from the analytical solution of the nonstationary response and the peak response quantities derived from the statistical inference analysis, curves of multiple rms values ( $k * rms$ ,  $k = 1, 2, 3, 4, 5$ ) were superimposed



**Fig. 19.** Response curves of peak accelerations  $SA$ , peak velocities  $SV$ , and peak displacements  $SD$  over period values for the San Andreas SE X-direction dataset. The curves correspond to mean peak values ( $E[SA]$ ,  $E[SV]$ ,  $E[SD]$ ) and percentiles: 80%, 90%, 95% of non-exceedance. The dashed lines in Fig.(b) and Fig.(c) are the mean plus rms values for the peak velocities  $SV$ , and peak displacements  $SD$  respectively (damping  $\zeta = 0.01$ ).

**Table 5**

Amplification factors for response spectral parameters based on the San Andreas SE X-direction probabilistic spectra (damping factor  $\zeta = 0.01$ ).

Spectral variables	Percentile curve		
	80%	90%	95%
SA (acceleration)	1.56	2.49	3.67
SV (relative velocity)	1.57	2.47	3.60
SD (relative displacement)	1.55	2.37	3.37

on the peak response percentile plots that are based on the Log-Normal distribution (Fig. 20). The rms values in Fig. 20 are the maximum values of the transient mean-square responses, and are determined from Eqs. (7) and (8) (see Fig. 13 for a sample transient mean-square response).

These graphs were plotted only for the peak velocities  $SV$  and peak displacements  $SD$  of the San Andreas SE X-direction dataset, and for one level of damping  $\zeta = 0.01$ . The graphs indicate the expected level of probability confidence, when the rms value (or a multiple value of rms) is used for a performance-based design of earthquake-resistant structures. For example, based on Fig. 19(b), by selecting the value of  $E[SV] + rms$  as the design threshold for peak relative velocities, there is an 80% probability that this value will not be exceeded. Similarly, based on Fig. 20(a), by selecting the value of  $5 * rms$  (i.e., five times the maximum rms response) as the design threshold for peak relative velocities  $SV$ , there is a 95% probability that this value will not be exceeded.

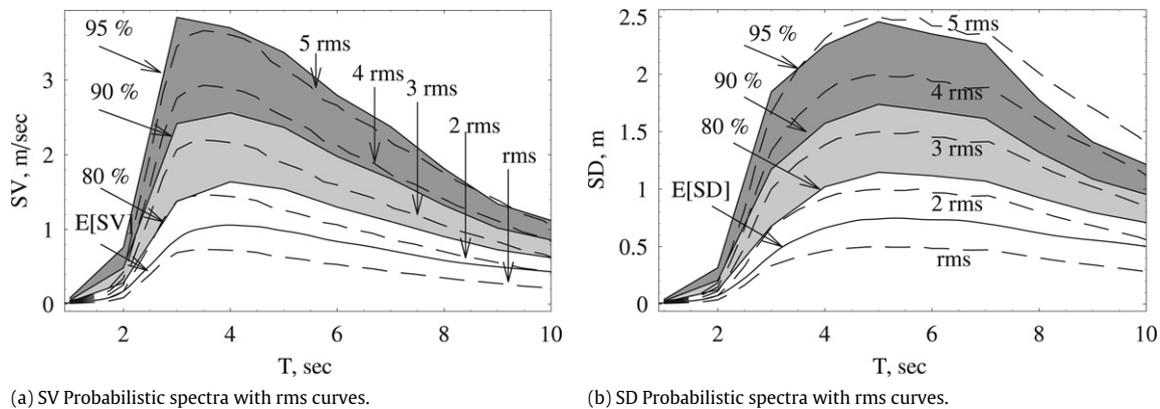
## 7. Conclusions

To construct regional earthquake response spectra from thousands of synthesized acceleration records derived from several scenario earthquakes plausible for the Los Angeles basin, a procedure with two stages of analysis is proposed.

The first stage of analysis includes the construction of the excitation temporal covariance matrices, their Karhunen–Loève (K–L) decomposition, and subsequent conversion of the dominant K–L eigenvectors to eigenfunctions with Chebyshev polynomials. Based on the exact closed-form solution for the nonstationary response of linear MDOF systems, response spectra for different damping values are constructed for the maximum values of the rms response quantities.

For the second stage of analysis, the response of an SDOF system subjected to acceleration excitation is calculated by solving analytically the differential equation of motion in a computationally efficient fashion for processing large size datasets. Response spectra for different damping values are constructed for the mean values of the ensemble response peaks. In addition, statistical inference analysis for the response variables leads to the construction of response spectra associated with percentile values of non-exceedance based on the Log-Normal probability model for the response variables.

Analysis results from both stages are combined by superimposing the percentile curves of the response quantities and the curves of multiple rms values. These graphs associate the rms values with confidence intervals attached to a certain probability value of not being exceeded.



**Fig. 20.** Superposition of percentile curves (solid lines) and maximum rms (dashed lines) for the spectral velocities SV (a) and spectral displacements SD (b) of the San Andreas SE X-direction scenario (damping is  $\zeta = 0.01$ ). The solid lines correspond to ensemble spectral means ( $E[SV]$  and  $E[SD]$ ), and 80%, 90%, and 95% values of non-exceedance. The dashed lines correspond to the maximum  $k * rms$  values for  $k = 1, 2, 3, 4, 5$ .

A significant and useful contribution of this study is the establishment of a quantitative relationship between the analytically determined maximum values of the nonstationary system response, and the probabilistic response spectra directly constructed from the ensemble statistics of the available earthquake records, so as to estimate the confidence levels of non-exceedance of specific spectral response levels. Such information can be useful in conducting performance-based designs matching certain reliability constraints.

The main limitations for the application of the proposed methodology for the construction of regional earthquake spectra are primarily due to the relatively small number of scenario earthquakes considered in this study and the maximum frequency (0.5 Hz) used for the ground motion simulation, and secondarily, to uncertainties concerning the accuracy of the basin model, and the omission of surface layers with shear velocities less than 1.0 km/s. Ground motion simulations with minimum shear velocities 300 m/s and maximum frequency 1.5 Hz are implemented in Benites and Olsen [21]. Future simulations that would incorporate signals with a wider frequency content, and more realistic values for the surface S-wave velocity, would give rise to higher ground velocities and accelerations, and as a result, to higher response spectra amplitudes.

## Acknowledgment

This study was supported in part by grants from the US National Science Foundation.

## References

- [1] Caughey TK, Stumpf HJ. Transient response of a dynamic system under random excitation. *ASME J Appl Mech* 1961;28:563–6.
- [2] Corotis RB, VanMarcke EH. Time-dependent spectral content of system response. *ASCE J Eng Mech Div* 1975;101(5):623–37.
- [3] Masri SF. Response of a multidegree-of-freedom system to nonstationary random excitation. *J Appl Mech* 1978;45(Sept.):649–56.
- [4] Spanos PTD, Lutes LD. Probability of response to evolutionary process. *ASCE J Eng Mech Div* 1980;106(2):213–24.
- [5] Debchaudhury A, Gasparini DA. Response of mdf systems to vector random excitation. *ASCE J Eng Mech Div* 1982;108(2):367–85.
- [6] Shinozuka M, Deodatis G. Stochastic process models for earthquake ground motion. *Probab Eng Mech* 1988;3(3):114–23.
- [7] Iwan WD, Hou ZK. Explicit solutions for the response of simple systems subjected to nonstationary random excitation. *Structural Safety* 1989;6: 77–86.
- [8] Soong TT, Grigoriu M. *Random Vibration of Mechanical and Structural Systems*. Englewood Cliffs, NJ: Prentice-Hall; 1993.
- [9] Lin YK, Cai GQ. *Probabilistic Structural Dynamics: Advanced Theory and Applications*. New York, NY: McGraw-Hill; 1995.
- [10] Lutes LD, Sarkani S. *Stochastic Analysis of Structural and Mechanical Vibrations*. Englewood Cliffs, NJ: Prentice-Hall; 1997.
- [11] Conte JP, Peng BF. An explicit closed-form solution for linear systems subjected to nonstationary random excitation. *Probab Eng Mech* 1996;11:37–50.
- [12] Papadimitriou C, Beck J.L. Nonstationary stochastic characterization of strong-motion accelerograms, in: *Proc. of 4th US National Conference on Earthquake Engineering*, vol. 1, 1990, pp. 477–486.
- [13] Masri SF, Miller RK. Compact probabilistic representation of random processes. *ASME J Appl Mech* 1982;49:871–6.
- [14] Traina MI, Miller RK, Masri SF. Orthogonal decomposition and transmission of nonstationary random processes. *Probab Eng Mech* 1986;1(3):136–49.
- [15] Masri SF, Miller RK, Traina MI. Probabilistic representation and transmission of earthquake ground motion records. *Earthq Eng Struct Dyn* 1990;19:1025–40.
- [16] Masri SF, Smyth AW, Traina MI. Probabilistic representation and transmission of nonstationary processes in multi-degree-of-freedom systems. *ASME J Appl Mech* 1998;65:398–409.
- [17] Maechling P, et al. Using the scec computational platforms for seismic hazard analysis research, in: *Geoinformatics*, USGS Reston VA, May 2006.
- [18] Faermi M., et al., Managing large scale data for earthquake simulations, in: *Proceedings of International Workshop on High-Performance Data Management in Grid Environments*, Rio de Janeiro, July 2006.
- [19] Olsen KB. 3d ground-motion estimation in Rome, Italy. *Bull. Seis. Soc. Am.* 2006;95:133–46.
- [20] Olsen KB. Strong shaking in Los Angeles expected from southern San Andreas earthquake. *Geophys Res Lett* 2006;33(L07305):doi:10.1029/2005GL025472.
- [21] Benites R, Olsen KB. Modeling strong ground motion in the Wellington metropolitan area, New Zealand. *Bull Seism Soc Am* 2005;95:2180–96.
- [22] Smyth A.W. Experimental and analytical studies in nonlinear system identification and modeling for structural control, Ph.D. Thesis, University of Southern California, Los Angeles, CA, May 1998.
- [23] Olsen K.B. Simulation of three-dimensional wave propagation in the Salt Lake Basin, Ph.D. Thesis, University of Utah, Salt Lake City, Utah, 1994.
- [24] Dolan JF. Prospects for larger or more frequent earthquakes in the Los Angeles metropolitan region. *Science* 1995;267:199–205.
- [25] Olsen KB. Site amplification in the Los Angeles basin from three-dimensional modeling of ground motion. *Bull Seism Soc Am* 2000;90(6B):S77–94.
- [26] Nigam NC, Jennings PC. Calculation of response spectra from strong-motion earthquake records. *Bull Seism Soc Am* 1969;59(2):909–22.
- [27] Ang A, Tang W. *Probability Concepts in Engineering Planning and Design*. John Wiley & Sons, Inc.; 1975.
- [28] Hoel PG. *Introduction to Mathematical Statistics*. 2nd edition. John Wiley & Sons Inc.; 1961.
- [29] Chopra AK. *Dynamics of Structures, Theory and Applications to Earthquake Engineering*. Prentice Hall; 1995.