Generation of Synthetic Accelerograms Compatible with a Set of Design Specifications

Anas Batou
Assistant professor, Université Paris-Est, Laboratoire Modélisation et Simulation Multi-Echelle, Marne-la-Vallée, France

Christian Soize
Professor, Université Paris-Est, Laboratoire Modélisation et Simulation Multi-Echelle, Marne-la-Vallée, France

ABSTRACT: The research addressed here concerns the generation of seismic accelerograms compatible with a given response spectrum and with other design specifications. The time sampling of the stochastic accelerogram yields a time series represented by a random vector in high dimension. The probability density function (pdf) of this random vector is constructed using the Maximum Entropy (MaxEnt) principle under constraints defined by the available information. In this paper, a new algorithm, adapted to the high stochastic dimension, is proposed to identify the Lagrange multipliers introduced in the MaxEnt principle to take into account the constraints. This novel algorithm is based on (1) the minimization of an appropriate convex functional and (2) the construction of the probability distribution defined as the invariant measure of an Itô Stochastic Differential Equation in order to estimate the integrals in high dimension of the problem.

1. INTRODUCTION

This paper is devoted to the generation of seismic accelerograms that are compatible with some design specifications such as the Velocity Response Spectrum, the Peak Ground Acceleration (PGA), etc. The Maximum Entropy (MaxEnt) principle introduced by Jaynes (1957a,b) in the framework of Information Theory constructed by E.Shannon (1948) is a powerful method which allows us to construct a probability distribution of a random vector under some constraints defined by the available information. This method has recently been applied in Soize (2010) for the generation of spectrum-compatible accelerograms as trajectories of a non-Gaussian non-stationary centered random process represented by a high-dimension random vector for which the probability density function (pdf) is constructed using the MaxEnt principle under constraints relative to (1) the mean value, (2) the variance of the components and (3) the mean value of the Velocity Response Spectrum (VRS). The objective of this paper is to take into account additional constraints related to some design specifications. To achieve this objective, the methodology proposed in Soize (2010) is extended to take into account constraints relative to statistics on (1) the end values for the velocity and the displacement, (2) the PGA, (3) the Peak Ground Velocity (PGV), (4) the envelop of the random VRS and (5) the Cumulative Absolute Velocity (CAV). The MaxEnt pdf is constructed and a generator of independent realizations adapted to the high-stochastic dimension of an accelerogram is proposed. Furthermore an adapted method (see Batou and Soize (2014)) for the identification of the Lagrange multipliers is developed. In Section 2 the MaxEnt principle is used to construct the pdf of the acceleration random vector under constraints defined by the available information. Finally, Section 3 is devoted to an application of the methodology for which the target VRS is constructed following the Eurocode 8 (see CEN (2003)).

2. CONSTRUCTION OF THE PROBABILITY DISTRIBUTION

The MaxEnt principle is used to construct the probability distribution of the random vector associated with a sampled stochastic process under some constraints defined by the available information.

The random acceleration of the soil is modeled by a second-order centered stochastic process $\{A(t), t \in [0,T]\}$. A time sampling of this stochastic process is introduced with a sampling time step Δt such that $T = N\Delta t$, yielding a time series $\{A_1, \ldots, A_N\}$ for which the $\mathbb{R}^N$-valued random vector $\mathbf{A} = (A_1, \ldots, A_N)$ is associated with. The probability distribution of the random vector $\mathbf{A}$ has to be constructed.

2.1. Maximum entropy principle

The objective of this section is to construct the pdf $\mathbf{a} \mapsto p_{\mathbf{A}}(\mathbf{a})$ of the random vector $\mathbf{A}$ using the MaxEnt principle under the constraints defined by the available information relative to random vector $\mathbf{A}$. The support of the pdf is assumed to be $\mathbb{R}^N$. Let $E\{.\}$ be the mathematical expectation. The available information is assumed to be written as

$$E\{\mathbf{g}(\mathbf{A})\} = \mathbf{f}, \tag{1}$$

in which $\mathbf{a} \mapsto \mathbf{g}(\mathbf{a})$ is a given function from $\mathbb{R}^N$ into $\mathbb{R}^\mu$ and where $\mathbf{f}$ is a given (or target) vector in $\mathbb{R}^\mu$. Equation (1) can be rewritten as

$$\int_{\mathbb{R}^N} \mathbf{g}(\mathbf{a}) p_{\mathbf{A}}(\mathbf{a}) d\mathbf{a} = \mathbf{f}. \tag{2}$$

An additional constraint related to the normalization of the pdf $p_{\mathbf{A}}(\mathbf{a})$ is introduced such that

$$\int_{\mathbb{R}^N} p_{\mathbf{A}}(\mathbf{a}) d\mathbf{a} = 1. \tag{3}$$

The entropy of the pdf $\mathbf{a} \mapsto p_{\mathbf{A}}(\mathbf{a})$ is defined by

$$S(p_{\mathbf{A}}) = -\int_{\mathbb{R}^N} p_{\mathbf{A}}(\mathbf{a}) \log(p_{\mathbf{A}}(\mathbf{a})) d\mathbf{a}, \tag{4}$$

where log is the natural logarithm. Let $\mathscr{C}$ be the set of all the pdf defined on $\mathbb{R}^N$ with values in $\mathbb{R}^+$, verifying the constraints defined by Eqs. (2) and (3). Then the MaxEnt principle consists in constructing the probability density function $\mathbf{a} \mapsto p_{\mathbf{A}}(\mathbf{a})$ as the unique pdf in $\mathscr{C}$ which maximizes the entropy $S(p_{\mathbf{A}})$. Then by introducing a Lagrange multiplier $\boldsymbol{\lambda}$ associated with Eq. (2) and belonging to an admissible open subset $\mathscr{L}_\mu$ of $\mathbb{R}^\mu$, it can be shown that the MaxEnt solution, if it exists, is defined by

$$p_{\mathbf{A}}(\mathbf{a}) = c_0(\boldsymbol{\lambda}^{\text{sol}}) \exp(-\langle \boldsymbol{\lambda}^{\text{sol}}, \mathbf{g}(\mathbf{a}) \rangle), \tag{5}$$

in which $\boldsymbol{\lambda}^{\text{sol}}$ is such that Eq. (2) is satisfied and where $c_0(\boldsymbol{\lambda})$ is the normalization constant defined by

$$c_0(\boldsymbol{\lambda}) = \left\{ \int_{\mathbb{R}^N} \exp(-\langle \boldsymbol{\lambda}, \mathbf{g}(\mathbf{a}) \rangle) \, d\mathbf{a} \right\}^{-1}. \tag{6}$$

2.2. Calculation of the Lagrange multipliers

In this section, we propose a general methodology for the calculation of the Lagrange multipliers $\boldsymbol{\lambda}^{\text{sol}}$.

2.2.1. Objective function and methododology

Using Eqs. (2) and (5), vector $\boldsymbol{\lambda}^{\text{sol}}$ is the solution in $\boldsymbol{\lambda}$ of the following set of μ nonlinear algebraic equations

$$\int_{\mathbb{R}^N} \mathbf{g}(\mathbf{a}) \, c_0(\boldsymbol{\lambda}) \exp(-\langle \boldsymbol{\lambda}, \mathbf{g}(\mathbf{a}) \rangle) = \mathbf{f}. \tag{7}$$

A more convenient way to calculate vector $\boldsymbol{\lambda}^{\text{sol}}$ consists in solving the following optimization problem (see Golan et al. (1996)),

$$\boldsymbol{\lambda}^{\text{sol}} = \arg \min_{\boldsymbol{\lambda} \in \mathscr{L}_\mu \subset \mathbb{R}^\mu} \Gamma(\boldsymbol{\lambda}), \tag{8}$$

in which the objective function Γ is written as

$$\Gamma(\boldsymbol{\lambda}) = \langle \boldsymbol{\lambda}, \mathbf{f} \rangle - \log(c_0(\boldsymbol{\lambda})). \tag{9}$$

Let $\{\mathbf{A}_{\boldsymbol{\lambda}}, \boldsymbol{\lambda} \in \mathscr{L}_\mu\}$ be a family of $\mathbb{R}^N$-valued random variables for which the pdf is defined, for all $\boldsymbol{\lambda}$ in $\mathscr{L}_\mu$, by

$$p_{\mathbf{A}_{\boldsymbol{\lambda}}}(\mathbf{a}) = c_0(\boldsymbol{\lambda}) \exp(-\langle \boldsymbol{\lambda}, \mathbf{g}(\mathbf{a}) \rangle). \tag{10}$$

Then the gradient vector $\boldsymbol{\nabla}\Gamma(\boldsymbol{\lambda})$ and the Hessian matrix $[H(\boldsymbol{\lambda})]$ of function $\boldsymbol{\lambda} \mapsto \Gamma(\boldsymbol{\lambda})$ are written as

$$\boldsymbol{\nabla}\Gamma(\boldsymbol{\lambda}) = \mathbf{f} - E\{\mathbf{g}(\mathbf{A}_{\boldsymbol{\lambda}})\}. \tag{11}$$

$$[H(\boldsymbol{\lambda})] = E\{\mathbf{g}(\mathbf{A}_{\boldsymbol{\lambda}})\mathbf{g}(\mathbf{A}_{\boldsymbol{\lambda}})^T\} - E\{\mathbf{g}(\mathbf{A}_{\boldsymbol{\lambda}})\}E\{\mathbf{g}(\mathbf{A}_{\boldsymbol{\lambda}})\}^T\,, \quad (12)$$

in which $\mathbf{u}^T$ is the transpose of $\mathbf{u}$. It is assumed that the constraints defined by Eq. (2) are algebraically independent. Consequently, the Hessian matrix is positive definite and therefore, function $\boldsymbol{\lambda} \mapsto \Gamma(\boldsymbol{\lambda})$ is strictly convex and reaches its minimum for $\boldsymbol{\lambda}^{\text{sol}}$ which is such that $\boldsymbol{\nabla}\Gamma(\boldsymbol{\lambda}) = \mathbf{0}$ for $\boldsymbol{\lambda} = \boldsymbol{\lambda}^{\text{sol}}$. It can then be deduced that the minimum of function $\boldsymbol{\lambda} \mapsto \Gamma(\boldsymbol{\lambda})$ corresponds to the solution of Eq. (7). The optimization problem defined by Eq. (8) is solved using the Newton iterative method

$$\boldsymbol{\lambda}^{\text{i}+1} = \boldsymbol{\lambda}^{\text{i}} - \alpha\,[H(\boldsymbol{\lambda}^{\text{i}})]^{-1}\boldsymbol{\nabla}\Gamma(\boldsymbol{\lambda}^{\text{i}})\,, \quad (13)$$

in which α belongs to $]0\,,1]$ is an under-relaxation parameter that ensures the convergence towards the solution $\boldsymbol{\lambda}^{\text{sol}}$. In general, for the non-Gaussian case, the integrals in the right-hand side of Eqs. (11) and (12) cannot explicitly be calculated and cannot be discretized in $\mathbb{R}^N$. In this paper, these integrals are estimated using the Monte Carlo simulation method for which independent realizations of the random vector $\mathbf{A}_{\boldsymbol{\lambda}}$ are generated using a specific algorithm presented below.

2.2.2. *Generator of independent realizations*

The objective of this section is to provide a generator of independent realizations of the random vector $\mathbf{A}_{\boldsymbol{\lambda}}$ for all $\boldsymbol{\lambda}$ fixed in $\mathscr{L}_\mu$. A generator of independent realizations for MaxEnt distributions has been proposed in Soize (2008, 2010) in the class of the MCMC algorithms. This methodology consists in constructing the pdf of random vector $\mathbf{A}_{\boldsymbol{\lambda}}$ as the density of the invariant measure $p_{\mathbf{A}_{\boldsymbol{\lambda}}}(\mathbf{a})d\mathbf{a}$, associated with the stationary solution of a second-order nonlinear Itô Stochastic differential equation (ISDE). The advantages of this generator compared to the other MCMC generators such as the Metropolis-Hastings (see Hastings (1970)) algorithm are: (1) The mathematical results concerning the existence and the uniqueness of an invariant measure can be used, (2) a damping matrix can be introduced in order to rapidly reach the invariant measure, and (3) there is no need to introduce a proposal distribution which can induce difficulties in high dimension. Below, we directly introduce the generator of independent realizations using a discretization of the ISDE. Details concerning the construction of this generator can be found in Soize (2008, 2010).

The ISDE is discretized using a semi-implicit integration scheme in order to avoid the resolution of an algebraic nonlinear equation at each step while allowing significantly increase of the time step compared to a purely explicit scheme. Concerning the initial conditions of the ISDE, the more the probability distribution of the initial conditions is close to the invariant measure, the shorter is the transient response and then the more efficient is the identification algorithm of the Lagrange multipliers (see Batou and Soize (2014)).

2.2.3. *Estimation of the mathematical expectations*

The generator described hereinbefore allows for constructing n_s independent realizations $\mathbf{A}^1_{\boldsymbol{\lambda}},\ldots,\mathbf{A}^{n_s}_{\boldsymbol{\lambda}}$ of random vector $\mathbf{A}_{\boldsymbol{\lambda}}$ representing the acceleration of the soil. The mean value $E\{\mathbf{g}(\mathbf{A}_{\boldsymbol{\lambda}})\}$ and the correlation matrix $E\{\mathbf{g}(\mathbf{A}_{\boldsymbol{\lambda}})\mathbf{g}(\mathbf{A}_{\boldsymbol{\lambda}})^T\}$ are estimated using the Monte Carlo simulation method by

$$E\{\mathbf{g}(\mathbf{A}_{\boldsymbol{\lambda}})\} \simeq \frac{1}{n_s}\sum_{\ell=1}^{n_s}\mathbf{g}(\mathbf{A}^\ell_{\boldsymbol{\lambda}})\,, \quad (14)$$

$$E\{\mathbf{g}(\mathbf{A}_{\boldsymbol{\lambda}})\mathbf{g}(\mathbf{A}_{\boldsymbol{\lambda}})^T\} \simeq \frac{1}{n_s}\sum_{\ell=1}^{n_s}\mathbf{g}(\mathbf{A}^\ell_{\boldsymbol{\lambda}})\mathbf{g}(\mathbf{A}^\ell_{\boldsymbol{\lambda}})^T\,. \quad (15)$$

3. APPLICATIONS

The acceleration stochastic process is sampled such that the final time $T = 20$ s. The time step is $\Delta t = 0.0125$ s. We then have $N = 1600$ (we assume $A(0) = 0$ ms^{-2} almost surely).

3.1. *Available information*

The available information related to random vector $\mathbf{A}$ is defined by:

(1) The random vector $\mathbf{A}$ is centered.

(2) The standard deviation of each component of random vector $\mathbf{A}$ is imposed. The target values are plotted in Fig. 1.

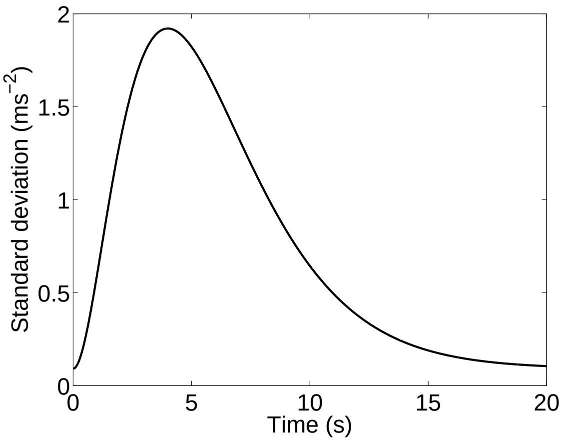

Figure 1: Target for the standard deviations of the components of random vector **A***.*

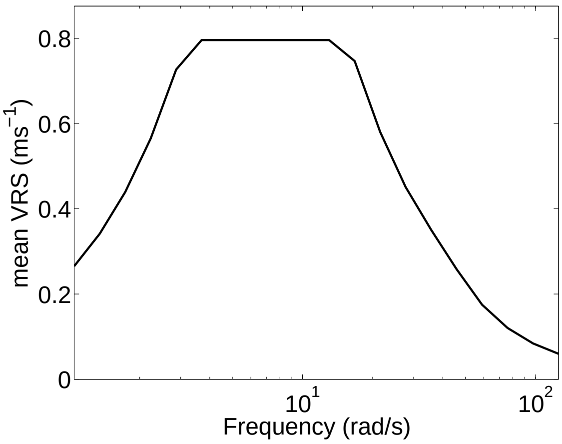

Figure 2: Target of the mean VRS.

(3) The variance of the end-velocity (resulting from a numerical integration of random vector **A**) is zero.

(4) The variance of the end-displacement (resulting from two successive numerical integrations of random vector **A**) is zero.

(5) The target for the mean VRS (see Clough and Penzien (1975)), which is denoted by $\underline{\mathbf{s}}$, is constructed following the Eurocode 8 for a A-type soil and a PGA equal to 5 ms^{-2}. It is defined for a damping ratio $\xi = 0.05$ and for 20 frequencies that are (in rad/s) 1.04, 1.34, 1.73, 2.23, 2.86, 3.69, 4.74, 6.11, 7.86, 10.11, 13.01, 16.74, 21.53, 27.70, 35.64, 45.86, 59.00, 75.91, 97.67 and 125.66. The target of the mean VRS is plotted in Fig. 2.

(6) Let $\mathbf{s}^{\mathrm{low}}$ be the lower envelop defined by $\mathbf{s}^{\mathrm{low}} = 0.5 \times \underline{\mathbf{s}}$ and $\mathbf{s}^{\mathrm{up}}$ be the upper envelop defined by $\mathbf{s}^{\mathrm{up}} = 1.5 \times \underline{\mathbf{s}}$. The probability for random vector **A** of being inside the region delimited by the two envelops is 0.99.

(7) The mean PGA is 5 ms^{-2}.

(8) The mean PGV is 0.45 ms^{-1}.

(9) The mean CAV is 13 ms^{-2}.

3.2. Results

For the ISDE, the number of integration steps is $M = 600$. At each iteration, $n_s = 900$ Monte Carlo simulations are carried out. The methodology developed in Section 2.2.1 is applied using 30 iterations. The under-relaxation parameter is $\alpha = 0.3$. Figure 3 shows two independent realizations of the random vector $\mathbf{A}_{\boldsymbol{\lambda}_{\mathrm{sol}}}$, which is generated using a classical generator for Gaussian random variable and which are representative of two independent realizations of the random accelerogram. The corresponding trajectories of the velocity times series $\mathbb{V}$ and of the displacement times series $\mathbb{D}$ result from two successive numerical integrations of each realization of the random accelerogram and are plotted in Figs. 4 and 5. As expected, it can be seen that the end velocity and the end displacements are both equal to zero. Figure 6 displays a comparison of the estimated standard deviation of the components with the target values.

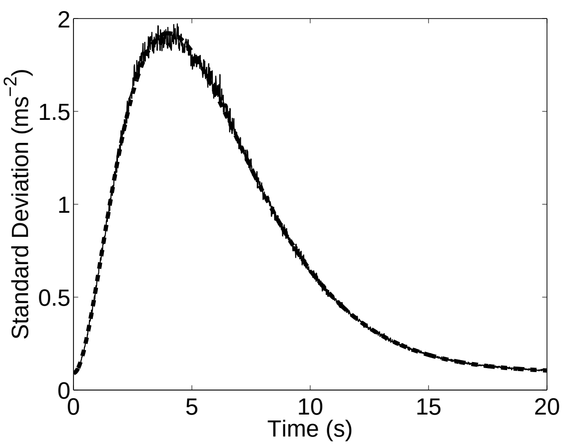

Figure 6: Variance: Target (thick dashed line) and estimation (thin solid line).

Figure 7 shows a comparison of the mean VRS

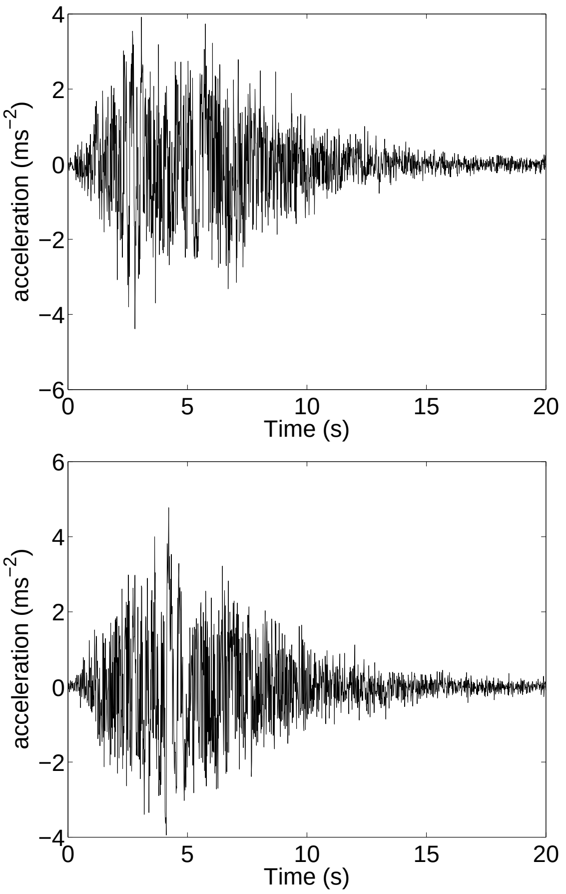

Figure 3: Two independent realizations of the random accelerogram.

with its target.

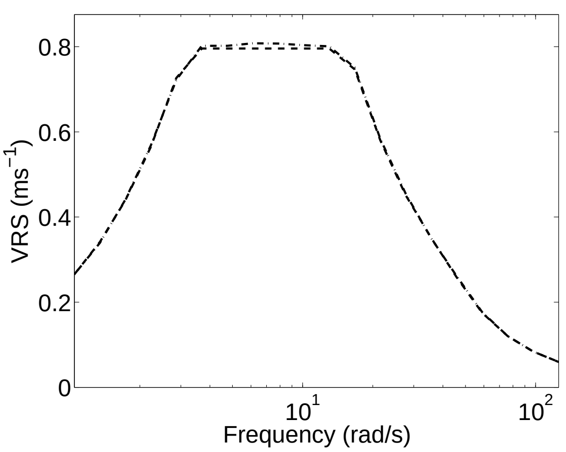

Figure 7: Mean VRS: Target (dashed line), estimation (mixed line).

The Figure 8 shows 100 trajectories of the ran-

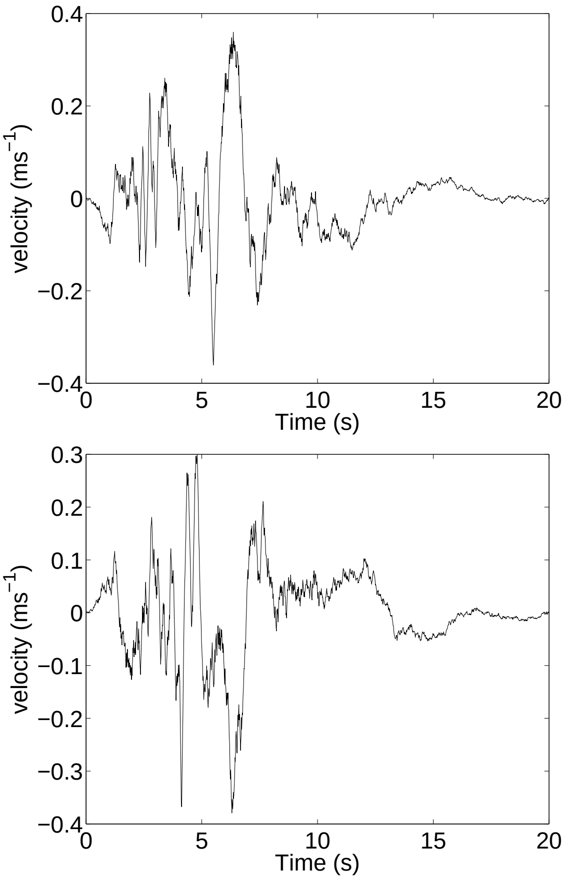

Figure 4: Two independent realizations of the random velocity.

dom VRS and the envelops $\mathbf{s}^{\mathrm{low}}$ and $\mathbf{s}^{\mathrm{up}}$.

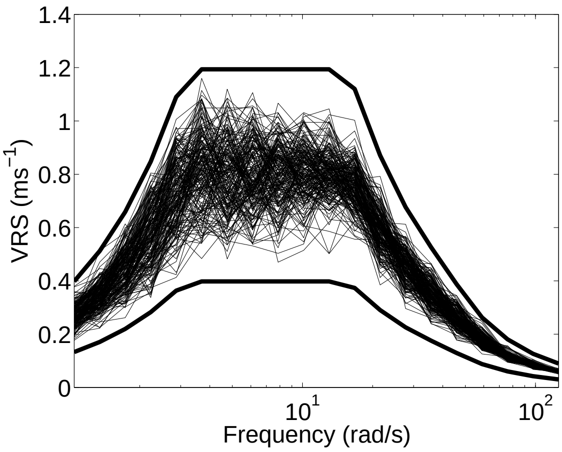

Figure 8: Random VRS: 100 trajectories (thin lines), lower and upper envelop (thick lines).

It can be seen in Figs. 6 to 8 a good matching

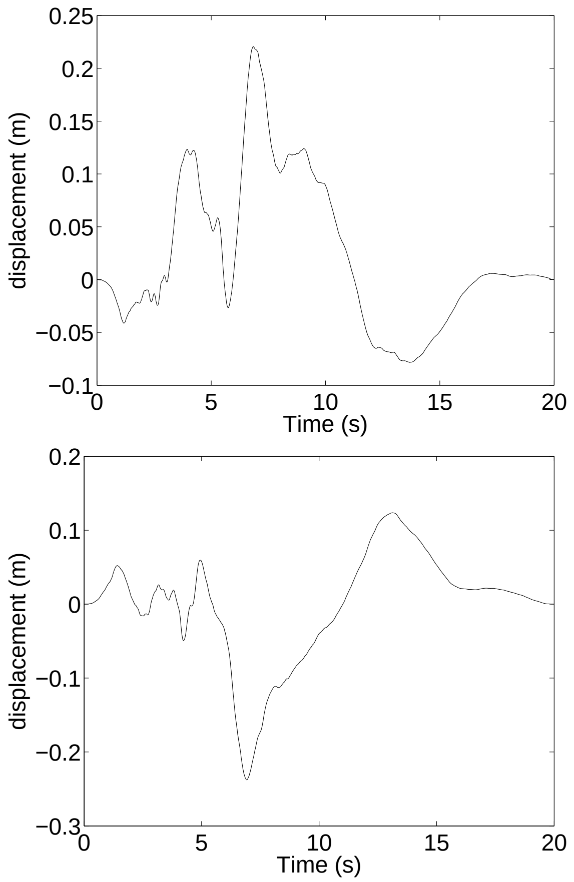

Figure 5: Two independent realizations of the random displacement.

between the estimated values and the target values. Concerning the PGA, the PGV and the CAV, the results are summarized in Table 1. It can be seen a good matching of the estimated mean values for the PGA, the PGV and the CAV with the target values. The results presented in Section **??** show a

Constraint	Estimation	Target
Mean PGA (ms^{-2})	4.98	5
Mean PGV (ms^{-1})	0.46	0.45
Mean CAV (ms^{-1})	13.04	13

Table 1: For the PGA, the PGV, the CAV: comparison of the estimated mean value with the target value.

very good agreement between the target values and the values estimated using the generated accelerograms. Nevertheless, in Figs. 3-5, it can be seen that the generated trajectories are not perfectly natural. For instance, spurious low-frequency content appears at the beginning of the signal. Such low-frequency content should appear later in the signal. The trajectories could be improved by adding contraints related to the nonstationarity of the velocity (or response) spectrum.

4. CONCLUSIONS

A new methodology has been presented for the generation of accelerograms compatible with a given VRS and other properties. If necessary, additional constraints could easily be taken into account in addition to those developed in this paper. The application shows a good matching between the estimated values and the target values. The generated trajectories could be improved by adding contraints related to time dependence of the VRS.

5. ACKNOWLEDGMENT

This research was supported by the "AgenceNationale de la Recherche", Contract TYCHE, ANR-2010-BLAN-0904.

6. REFERENCES

Batou, A. and Soize, C. (2014). “Generation of accelerograms compatible with design specifications using information theory.” *Bulletin of Earthquake Engineering*, 12(2), 769–794.

CEN (2003). *Eurocode 8: Design of Structures for Earthquake Resistance-Part1: General Rules, Seismic Actions and Rules for Buildings. EN 1998-1: 2003 European Committee for Standardization*. Brussels.

Clough, R. W. and Penzien, J. (1975). *Dynamics of Structures*. McGraw-Hill.

E.Shannon, C. (1948). “A mathematical theory of communication.” *Bell System Tech. J.*, 27, 379–423 and 623–659.

Golan, A., Judge, G., and Miller, D. (1996). *Maximum entropy econometrics: robust estimation with limited data*. Wiley, New York.

Hastings, W. K. (1970). “Monte carlo sampling methods using markov chains and their applications.” *Biometrica*, 109, 57–97.

Jaynes, E. T. (1957a). "Information theory and statistical mechanics." *Physical Review*, 106(4), 620–630.

Jaynes, E. T. (1957b). "Information theory and statistical mechanics." *Physical Review*, 108(2), 171–190.

Soize, C. (2008). "Construction of probability distributions in high dimension using the maximum entropy principle. applications to stochastic processes, random fields and random matrices." *International Journal for Numerical Methods in Engineering*, 76(10), 1583–1611.

Soize, C. (2010). "Information theory for generation of accelerograms associated with shock response spectra." *Computer-Aided Civil and Infrastructure Engineering*, 25, 334–347.