Estimating Seismic Performance Uncertainty using IDA with Progressive Accelerogram-wise Latin Hypercube Sampling

D. Vamvatsikos
School of Civil Engineering
National Technical University, Athens, Greece

ABSTRACT: An efficient algorithm is presented that allows the rapid estimation of the influence of model parameter uncertainties on the seismic performance of structures using incremental dynamic analysis (IDA) and Monte Carlo simulation with latin hypercube sampling. The fundamental blocks of this methodology have already been proposed as a means to quantify the uncertainty for structural models with non-deterministic parameters, whereby each model realization out of a predetermined sample size is subjected to a full IDA under multiple ground motion records. However, any practical application is severely restricted due to (a) our inability to determine a priori the required number of samples and (b) the disproportionate increase in the number of analyses when dealing with many random variables. Thus, two fundamental changes are incorporated. First, latin hypercube sampling is applied incrementally by starting with a small sample that is doubled successively until adequate accuracy has been achieved. At the same time, instead of maintaining the same properties for a given model realization over an entire ground-motion record suite, parameter sampling is performed on a record-by-record basis, efficiently expanding the model sample size without increasing the number of nonlinear dynamic analyses. Using a steel moment-resisting frame building as a test case, it is shown that the improved algorithm allows excellent scalability and extends the original methodology to be easily applicable to realistic large-scale applications with hundreds of random variables.

1 INTRODUCTION

The accurate estimation of the seismic demand and capacity of structures stands at the core of performance-based earthquake engineering. While guidelines have emerged (SAC/FEMA 2000) that recognize the need for assessing epistemic uncertainties by explicitly including them in estimating seismic performance, this role is usually left to ad hoc safety factors, or, at best, standardized dispersion values that often serve as placeholders. Still, seismic performance is heavily influenced by both aleatory randomness, e.g., due to natural ground motion record variability, and epistemic uncertainty, owing to modeling assumptions, omissions or errors. While the first can be easily estimated by analyzing a given structure under multiple ground motion records, for example via incremental dynamic analysis (IDA, Vamvatsikos & Cornell 2002), estimating the uncertainty remains a little-explored issue.

Recently, several researchers have proposed applying nonlinear dynamic analysis combined with Monte Carlo simulation (Rubinstein 1981) to quantify the uncertainty for structural models with non-deterministic parameters. For example, Ibarra (2003) actually proposes a method to propagate the uncertainty from model parameters to structural behavior using first-order-second-moment (FOSM) principles verified through Monte Carlo to evaluate the collapse capacity uncertainty. As a performance improvement, Latin Hypercube Sampling (LHS) (McKay et al. 1979) has also been proposed instead of classic random sampling. Kazantzi et al. (2008) used Monte Carlo with LHS to incorporate uncertainty into steel frame fragility curves. Liel et al. (2009) used IDA with Monte Carlo and FOSM coupled with a response surface approximation method to evaluate the collapse uncertainty of a reinforced-concrete building. On a similar track, Dolsek (2009) and Vamvatsikos & Fragiadakis (2010) have proposed using Monte Carlo with efficient Latin Hypercube Sampling (LHS) on IDA to achieve the same goal. However, any practical application of the above methodologies is severely restricted due to two important reasons.

The first is our inherent inability to determine in advance the required number of observations. Due to the nature of LHS, the entire sample has to be decided *a priori*. It is generally not possible to expand or contract a given sample to an arbitrary higher or lower size. For typical random sampling we can stop the simulation before we examine the entire sample if the intermediate results lead us to decide at runtime that the size can be decreased. Doing so for LHS is not possible unless we want to risk a biased estimate. Similarly, if after the end of the simulation we realize that we need more observations, we cannot easily reuse the existing ones by arbitrarily adding to them; the end product will typically not be a proper latin hypercube design. In other words, we are limited to our initial knowledge of the problem to be able to select a proper sample size, which may or may not be correct on the first try, often necessitating the use of more tries to be able to validate our approach.

The second is the problem of the disproportionate increase in the number of analyses when dealing with many random variables. Depending on the correlation structure of the variables, it may become prohibitively expensive to determine the influence of multiple random parameters, as the sample size rises disproportionately. Thus, matters of, e.g., spatial uncertainty and correlation of variables at multiple locations in a structure are quite difficult to resolve as they necessitate a high number of observations that easily runs in the hundreds or thousands. This is what has lead all early attempts (Liel et al. 2009, Dolsek 2009, Vamvatsikos & Fragiadakis 2010) to limit themselves to just a handful of parameters. What compounds all of the above, is that it becomes highly desirable to limit the computational cost of evaluating each sample member. This is bound to lead to a trend of reducing the size of the ground motion record suite used for IDA, making it attractive to use even 10 records rather than a healthier set of 20 or 30.

In our opinion this is all a trap of our own preconceptions about how IDA is applied combined with a desire to separate aleatory from epistemic uncertainty. One is accustomed to using the same model over all accelerograms to determine the record-to-record variability and then attempting to add the epistemic uncertainty influence upon that. Still, when dealing within a performance-based framework such as SAC/FEMA (Cornell et al. 2002) or the Pacific Earthquake Engineering Research Center (Cornell & Krawinkler 2000) framework, it is customary to combine aleatory and epistemic together and use them as a single dispersion parameter (see also Der Kiureghian 2005). Furthermore, it is often forgotten that IDA is itself a record-sampling technique at its core, simply assuming that all records in the set have an equal probability of occurrence.

Therefore, we will attempt to reorganize the application of Monte Carlo with LHS on IDA by performing together the model and record sampling and using incremental sample sizes that have been carefully selected to allow full reuse of the earlier runs performed. The end results is a general algorithm that efficiently upgrades the original to be applicable to large models with hundreds of random variables and without any need of pre-determining sample sizes in any way.

Table 1: The format of the iLHS sample for N parameters and M records.

No.	X_1	X_2	...	X_N	X_{N+1}[1]	X_{N+2}[2]
1	$x_{1,1}$	$x_{1,2}$	...	$x_{1,N}$	ang_1	Rec_1
2	$x_{2,1}$	$x_{2,2}$	...	$x_{2,N}$	ang_2	Rec_2
...	...	...	...	...	...	...
M	$x_{M,1}$	$x_{M,2}$	...	$x_{M,N}$	ang_M	Rec_M
$M+1$	$x_{M+1,1}$	$x_{M+1,2}$	...	$x_{M+1,N}$	ang_{M+1}	Rec_1
$M+2$	$x_{M+2,1}$	$x_{M+2,2}$	...	$x_{M+2,N}$	ang_{M+2}	Rec_2
...	...	...	...	...	...	...

[1] incident angle [2] record index

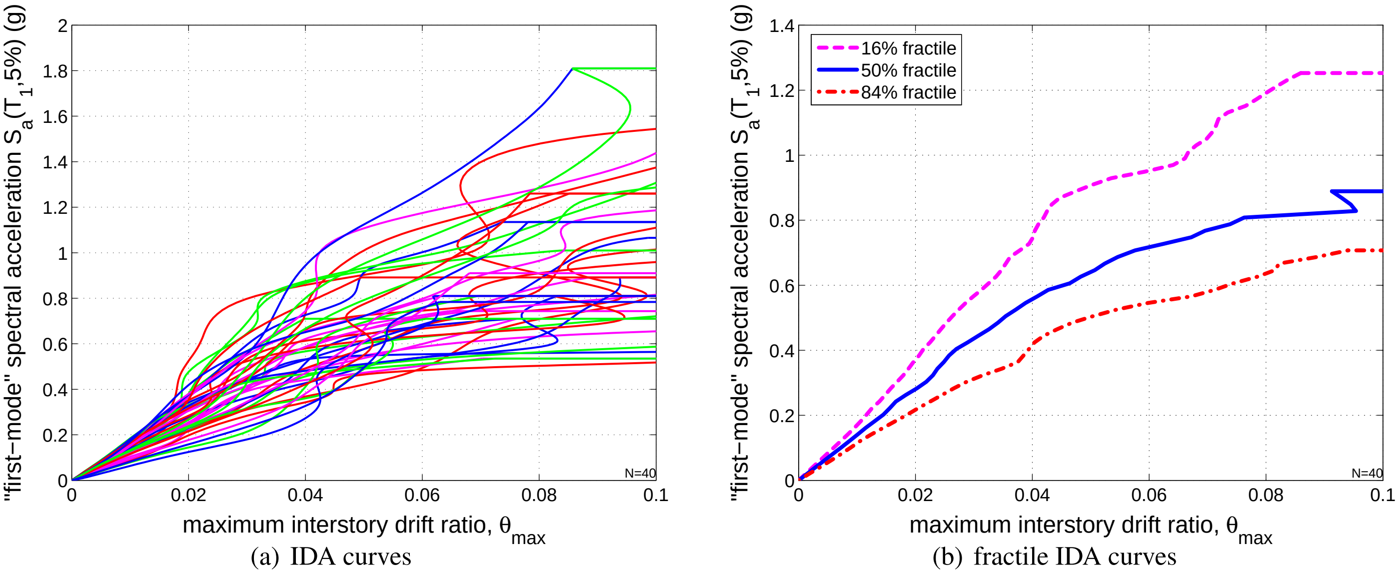
2 INCREMENTAL DYNAMIC ANALYSIS

Incremental Dynamic Analysis (IDA) is a powerful analysis method that offers thorough seismic demand and capacity prediction capability (Vamvatsikos & Cornell 2002). It involves performing a series of nonlinear dynamic analyses under a multiply-scaled suite of ground motion records, selecting proper Engineering Demand Parameters (EDPs) to characterize the structural response and an Intensity Measure (IM), e.g. the 5% damped first-mode spectral acceleration, $S_a(T_1, 5\%)$, to represent the seismic intensity. The results are presented as curves of EDP versus IM for each record (Figure 1(a)). These can be further summarized into the 16,50,84% fractile IDA curves by estimating the respective percentile values given a range of IM or EDP values. Appropriate limit-states can be defined by setting limits on the EDPs and the probabilistic distribution of their capacities can be easily estimated, e.g., for limiting values of the maximum interstory drift by reading off the median and the dispersion of the required $S_a(T_1, 5\%)$ capacity from Figure 1(b). Such results combined with probabilistic seismic hazard analysis (Vamvatsikos & Cornell 2002) allow the estimation of mean annual frequencies (MAFs) of exceeding the limit-states, thus offering a direct characterization of seismic performance.

Nevertheless, IDA comes at a considerable cost, even for simple structures, necessitating the use of multiple nonlinear dynamic analyses that are usually beyond the abilities and the computational resources of the average practicing engineer. Therefore, wherever IDA is involved, searching for an efficient implementation is always desirable.

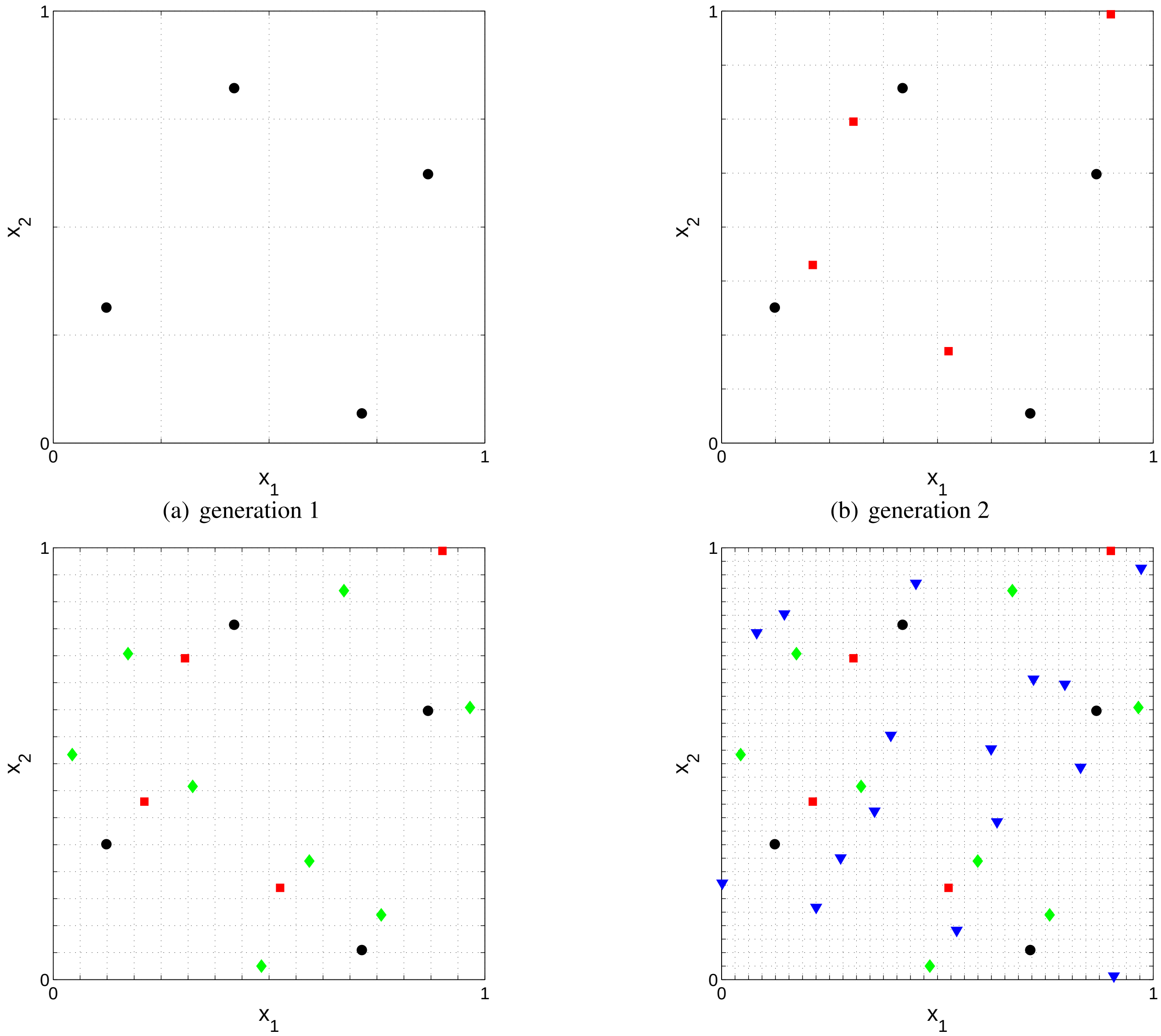
3 INCREMENTAL RECORDWISE LHS

To mitigate the issues related to the typical application of LHS on IDA, we propose using the same two fundamental procedures but essentially redefine the way that they are implemented by incorporating two important changes. First, latin hypercube sampling is applied incrementally by starting with a small sample that is doubled successively until adequate accuracy has been achieved. This is perhaps the only way that one can reuse the results of a previous LHS design, since doubling the size allows a simple way



(a) IDA curves (b) fractile IDA curves

Figure 1: Forty IDA curves and their summarization into 16,50,84% fractile IDA curves.



(a) generation 1 (b) generation 2

(c) generation 3 (d) generation 4

Figure 2: An iLHS design with two variables and four generations. Each successive generation doubles the number of samples.

to insert new observations within the existing sample while maintaining all the properties and advantages of LHS, as shown for example in Figure 2 for two random variables. Thus, by comparing the convergence of the IDA results in successive generations of the LHS design, the development of a rational stopping rule becomes possible. This essentially offers an intuitive way to determine a reasonable sample size, minimizing the waste of runs over repeated tries or the (equally wasteful) tendency to overestimate the size to "get it right" in one step. Actually, the proposed amendment is simple enough that it has probably already appeared in the literature although the author has not been able so far to find a publication that describes it. Still, the use of LHS is so extensive that it is reasonable to surmise that at least something similar must have appeared elsewhere.

Furthermore, by taking advantage of the fact that the IDA is itself a sampling process at equiprobable points (or records), we propose that LHS is performed simultaneously on the structural properties and on the ground motion records. Instead of maintaining the same properties for a given model realization over an entire ground-motion record suite, model parameter sampling is performed on a record-by-record basis, efficiently expanding the number of observations without increasing the number of nonlinear dynamic analyses, a concept that has also been proposed in a different context by Schotanus & Franchin (2004). An example of such a sampling design appears in Table 1, where each row is one structural observation that also corresponds to a single ground motion record on which IDA is performed. As a further bonus, the incident angle of the record may also be varied to allow for including its effect as well. If we need more observations than the *M* records available, the records can be simply recycled, either with the same or a different incident angle. In the customary application of such a procedure, each row of the table would be subject to IDA for the entire record suite, multiplying the number of nonlinear dynamic analyses by a factor of 20–60.

The end result is iLHS, a general algorithm that efficiently upgrades the original to be applicable to large models with hundreds of random variables and without any need of pre-determining sample sizes in any way. On the other hand, iLHS is not without some minor disadvantages that can be easily remedied. Perhaps the most important is that due to the small size of the first generation samples, one cannot use some standard algorithms (Iman & Conover 1982) for imposing the desired correlation structure on the sample. Instead, genetic or evolutionary algorithms need to be employed, such as the one by Charmpis & Panteli (2004). These are more time-consuming but they offer the ability to fine-tune the correlation structure. Another issue is that we cannot use some "accelerated-IDA" techniques, e.g., priority lists (Azarbakht & Dolsek 2007), or SPO2IDA (Fragiadakis & Vamvatsikos 2010) as they assume the same model over an entire record suite. Finally, we cannot distinguish the model uncertainty effects from the record-to-record variability unless additional analyses are performed for the mean-parameter model. Nevertheless, such amendments are easy to provide, and the methodology can be employed with a minimal of programming.

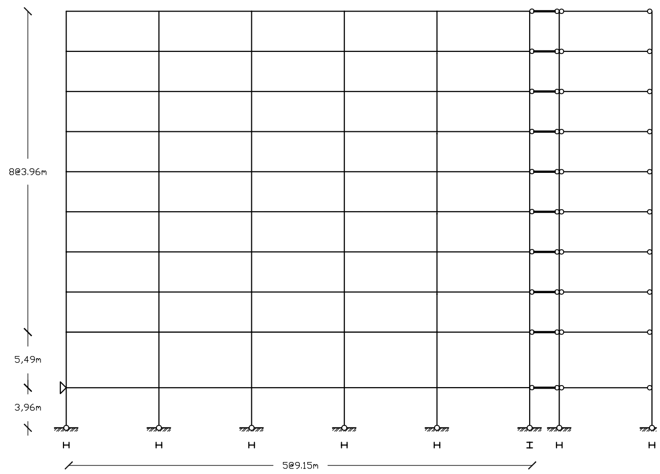


Figure 3: The LA9 steel moment-resisting frame.

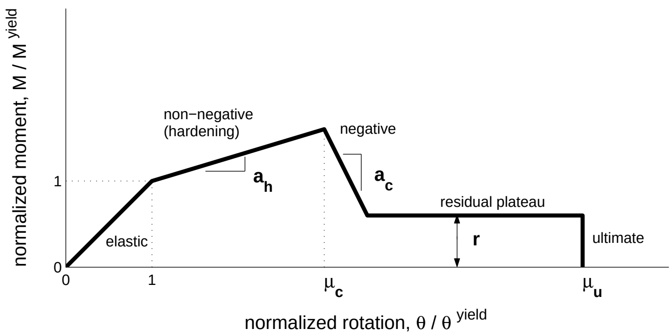


Figure 4: The moment-rotation beam-hinge backbone to be investigated and its six controlling parameters.

Table 2: The distribution properties of the uncertain parameters.

	mean	c.o.v	min	max	type
a_{My}	1.0	20%	0.70	1.30	truncated normal
a_h	0.1	40%	0.04	0.16	truncated normal
μ_c	3.0	40%	1.20	4.80	truncated normal
a_c	-0.5	40%	-0.80	-0.20	truncated normal
r	0.5	40%	0.20	0.80	truncated normal
μ_u	6.0	40%	2.40	9.60	truncated normal

4 EXAMPLE APPLICATION

4.1 *Model description*

The structure selected is a nine-story steel moment-resisting frame with a single-story basement (Figure 3) that has been designed for Los Angeles, following the 1997 NEHRP (National Earthquake Hazard Reduction Program) provisions (Foutch & Yun 2002).

A centerline model with nonlinear beam-column connections was formed using OpenSees (McKenna et al. 2000). It allows for plastic hinge formation at the beam ends while the columns are assumed to remain elastic. The structural model also includes P-Δ effects while the internal gravity frames have been directly incorporated (Figure 3). The fundamental period of the reference frame is $T_1 = 2.35$s and accounts for approximately 84% of the total mass. Essentially this is a first-mode dominated structure that still allows for some sensitivity to higher modes.

The beam-hinges are modeled as rotational springs with a quadrilinear moment-rotation backbone (Figure 4) that is symmetric for positive and negative rotations and employs a moderately pinching hysteresis without any cyclic degradation (Ibarra 2003). The backbone hardens after a yield moment of a_{My} times the nominal, having a non-negative slope of a_h up to a normalized rotation (or rotational ductility) μ_c where the negative stiffness segment starts. The drop, at a slope of a_c, is arrested by the residual plateau appearing at normalized height r that abruptly ends at the ultimate rotational ductility μ_u.

In order to evaluate the effect of uncertainties on the seismic performance of the structure we chose to vary the beam-hinge backbones by assigning truncated normal probabilistic distributions to its six parameters, as shown in Table 2. All distributions were appropriately rescaled to avoid the concentration of high probabilities at the cutoff points (Benjamin & Cornell 1970). The hinges at the end of each individual beam were assumed to be perfectly correlated. Within the same beam, all parameters are independent except the ductilities μ_c and μ_u that share an 80% correlation coefficient. Among different beams in any given story, a 70% correlation was employed for each parameter. Among beams in different stories only 50% correlation was used.

4.2 *Illustrative results*

Having a total of 270 random variables and 60 ordinary ground motion records (i.e., without any soft soil or directivity issues), the incremental iLHS algorithm is applied with a starting size of 10 and it is allowed to run to a total of 8 generations, up to a maximum sample size of 1280. Obviously we have too many correlation coefficients to control: $270(270-1)/2 = 36315$, so it is not possible to get them all right. Even with 1280 observations the maximum absolute error is in the order of 0.25, i.e., all correlation coefficients are matched within a ± 0.25: a 0.5 could be 0.25 or 0.75. In truth, due to the nature of the Charmpis & Panteli (2004) algorithm, this large error is concentrated only at some of the random variables, which are the ones that are left last to be optimized in the process. Therefore, it becomes advantageous to rearrange the sample in such a way that the most important variables, e.g., all the a_{My} and μ_c, appear first. Thus, their correlation is nearly perfectly represented.

The simulation was run in parallel using the IDA running algorithms developed by Vamvatsikos (2011) and using 5 Pentium IV single-core processors for an overall running time of 10hrs. As shown in Figures 5, the 1280 observations are actually too many. The median and the dispersion β (standard deviation of the log of the data) of the $S_a(T_1, 5\%)$ intensity measure (IM) to achieve a certain response value become fairly stable for practically all engineering demand parameters (EDPs) after only 4–5 generations, at 160 or 320 samples, respectively. The results seem to only mildly differ among global or local EDPs considered, e.g., the roof drift θ_{roof} and maximum insterstory drift $\theta_{\max}$, or the individual i-story drifts θ_i.

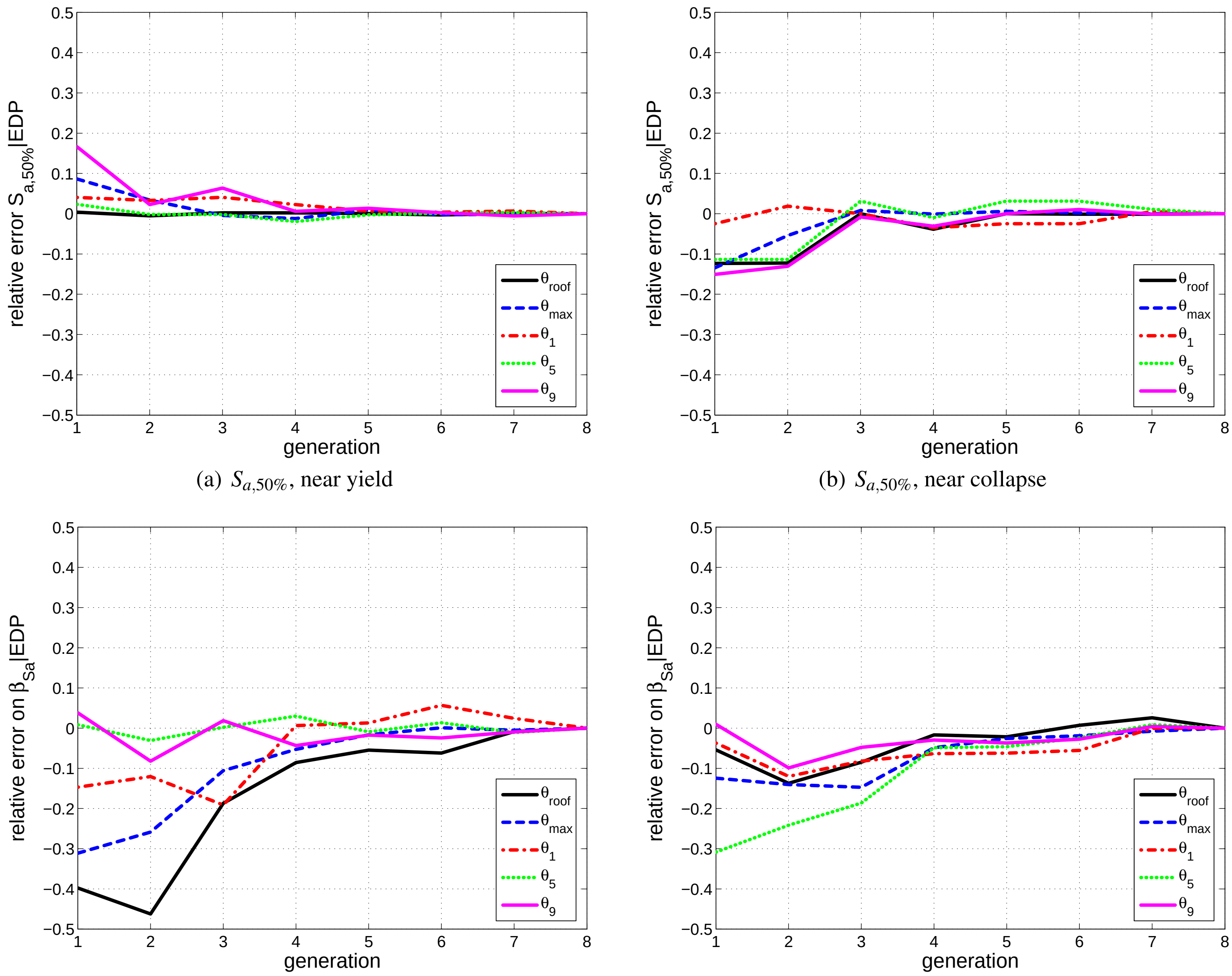
In comparison to a typical analysis considering only the mean-parameter model, the dispersion is found to be similar (Figure 7) but the mean response itself has a prominent bias, which, due to the details of the correlation imposed, appears to be a conservative one (Figure 6). It is also possible to determine the influence of each random variable by measuring its correlation with the estimated response values, following the suggestions of Dolsek (2009). Then, we find that at least for $\theta_{\max}$, the most influential variables in the lower, near-yield, limit-states ($\theta_{\max} = 0.02$) mainly involve the μ_u and r variables at a meager 8-10% correlation. For higher limit-states, closer to collapse, we slowly start to see the presence of a_{My} at the middle stories, mainly the fifth, with the correlation progressively rising up to the order of 22%. Such information can be extracted to any detail and for each EDP type and structural state desired.

5 CONCLUSIONS

The efficient Incremental Accelerogram-wise Latin Hypercube Sampling procedure iLHS has been presented as an algorithm that is capable of efficiently estimating the effect of model parameter uncertainties on the seismic performance of structures. In effect, it builds upon the existing paradigm of incremental dynamic analysis with latin hypercube sampling and further improves it by resolving the problem of sample size determination and by partially mitigating its slow performance, offering an improvement by a factor of 20 at least. The end result is an efficient algorithm that is amenable to parallelization and automated application while it allows excellent scalability and extends the original methodology to be easily applicable to realistic large-scale problems with hundreds of random variables.

REFERENCES

Azarbakht, A. & M. Dolsek (2007). Prediction of the median IDA curve by employing a limited number of ground motion records. *Earthquake Engineering and Structural Dynamics 36*(15), 2401–2421.



(a) $S_{a,50\%}$, near yield

(b) $S_{a,50\%}$, near collapse

(c) β_{Sa} near yield

(d) β_{Sa} near collapse

Figure 5: The relative errors in the median and dispersion in S_a terms as estimated for two limit-states and multiple EDPs.

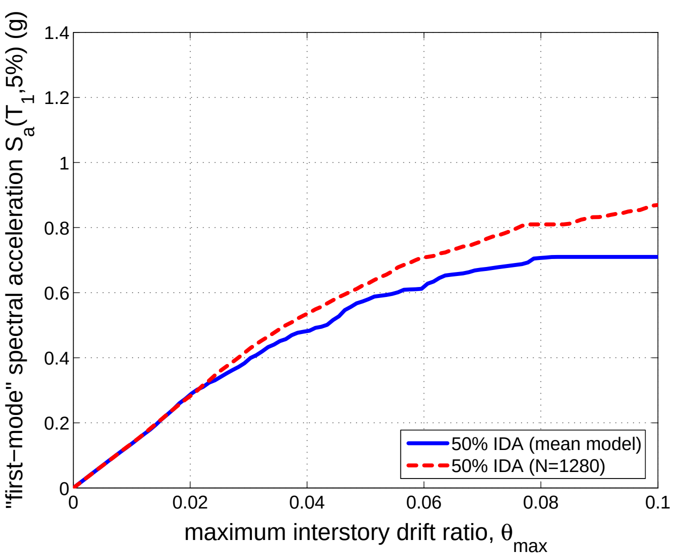


Figure 6: Comparison of median IDAs: iLHS versus the mean model.

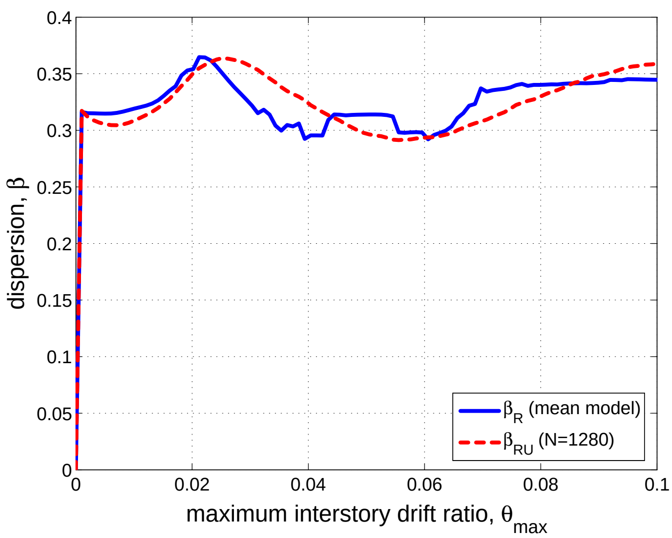


Figure 7: Comparison of dispersion of $S_a(T_1,5\%)$ capacity: iLHS versus the mean model.

Benjamin, J. R. & C. A. Cornell (1970). *Probability, Statistics, and Decision for Civil Engineers*. New York: McGraw-Hill.

Charmpis, D. C. & P. P. Panteli (2004). A heuristic approach for the generation of multivariate random samples with specified marginal distributions and correlation matrix. *Computational Statistics 19*, 283–300.

Cornell, C. A., F. Jalayer, R. O. Hamburger, & D. A. Foutch (2002). The probabilistic basis for the 2000 SAC/FEMA steel moment frame guidelines. *ASCE Journal of Structural Engineering 128*(4), 526–533.

Cornell, C. A. & H. Krawinkler (2000). Progress and challenges in seismic performance assessment. *PEER Center News 3*(2). [Oct 2009].

Der Kiureghian, A. (2005). Non-ergodicity and PEER's framework formula. *Earthquake Engineering and Structural Dynamics 34*(13), 1643–1652.

Dolsek, M. (2009). Incremental dynamic analysis with consideration of modelling uncertainties. *Earthquake Engineering and Structural Dynamics 38*(6), 805–825.

Foutch, D. A. & S.-Y. Yun (2002). Modeling of steel moment frames for seismic loads. *Journal of Constructional Steel Research 58*, 529–564.

Fragiadakis, M. & D. Vamvatsikos (2010). Fast performance uncertainty estimation via pushover and approximate IDA. *Earthquake Engineering and Structural Dynamics 39*(6), 683–703.

Ibarra, L. F. (2003). *Global Collapse of Frame Structures under Seismic Excitations*. PhD Dissertation, Department of Civil and Environmental Engineering, Stanford University, Stanford, CA.

Iman, R. L. & W. J. Conover (1982). A distribution-free approach to inducing rank correlation among input variables. *Communication in Statistics Part B: Simulation and Computation 11*(3), 311–334.

Kazantzi, A. K., T. D. Righiniotis, & M. K. Chryssanthopoulos (2008). Fragility and hazard analysis of a welded steel moment resisting frame. *Journal of Earthquake Engineering 12*(4), 596–615.

Liel, A. B., C. B. Haselton, G. G. Deierlein, & J. W. Baker (2009). Incorporating modeling uncertainties in the assessment of seismic collapse risk of buildings. *Structural Safety 31*(2), 197–211.

McKay, M. D., W. J. Conover, & R. Beckman (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics 21*(2), 239–245.

McKenna, F., G. Fenves, B. Jeremic, & M. Scott (2000). Open system for earthquake engineering simulation. [May 2008].

Rubinstein, R. Y. (1981). *Simulation and the Monte Carlo method*. New York: John Wiley & Sons.

SAC/FEMA (2000). Recommended seismic design criteria for new steel moment-frame buildings. Report No. FEMA-350, SAC Joint Venture, Federal Emergency Management Agency, Washington, DC.

Schotanus, M. & P. Franchin (2004). Seismic reliability analysis using response surface: a simplification. In *Proceedings of the 2nd ASRANet Colloquium*, Barcelona, Spain, pp. 1–8.

Vamvatsikos, D. (2011). Performing incremental dynamic analysis in parallel. *Computers and Structures 89*(1-2), 170–180.

Vamvatsikos, D. & C. A. Cornell (2002). Incremental dynamic analysis. *Earthquake Engineering and Structural Dynamics 31*(3), 491–514.

Vamvatsikos, D. & M. Fragiadakis (2010). Incremental dynamic analysis for estimating seismic performance sensitivity and uncertainty. *Earthquake Engineering and Structural Dynamics 39*(2), 141–163.