

Importance sampling based active learning for parametric seismic fragility curve estimation

Clément Gauchy^{a,b}, Cyril Feau^a, Josselin Garnier^b

^a *Université Paris-Saclay, CEA, Service d'Études Mécaniques et Thermiques, 91191, Gif-sur-Yvette, France*

^b *CMAP, École Polytechnique, Institut Polytechnique de Paris, 91128 Palaiseau Cedex, France*

Abstract

The key elements of seismic probabilistic risk assessment studies are the fragility curves which express the probabilities of failure of structures conditional to a seismic intensity measure. A multitude of procedures is currently available to estimate these curves. For modeling-based approaches which may involve complex and expensive numerical models, the main challenge is to optimize the calls to the numerical codes to reduce the estimation costs. Adaptive techniques can be used for this purpose, but in doing so, taking into account the uncertainties of the estimates (via confidence intervals or ellipsoids related to the size of the samples used) is an arduous task because the samples are no longer independent and possibly not identically distributed. The main contribution of this work is to deal with this question in a mathematical and rigorous way. To this end, we propose and implement an active learning methodology based on adaptive importance sampling for parametric estimations of fragility curves. We prove some theoretical properties (consistency and asymptotic normality) for the estimator of interest. Moreover, we give a convergence criterion in order to use asymptotic confidence ellipsoids. Finally, the performances of the methodology are evaluated on analytical and industrial test cases of increasing complexity.

Keywords: Computer experiments, probabilistic risk assessment, importance sampling, statistical learning

1. Introduction

The notion of fragility curve appeared in the 1980s as a key element of Seismic Probabilistic Risk Assessment (SPRA) studies (see e.g. [1, 2, 3]). A fragility curve expresses the probability of failure of a mechanical structure subjected to earthquake excitations conditional to a seismic Intensity Measure (IM), that may be the Peak Ground Acceleration (PGA) or the Pseudo-Spectral Acceleration for a given frequency (PSA). Fragility curves are also part of Performance-Based Earthquake Engineering (PBEE) framework [4, 5] and are not limited to seismic loading. For example, they are also used for wind loading [6] or to address the problem of multi-hazard storm conditions [7].

In earthquake engineering, various data sources can be exploited to estimate these curves, namely: expert judgments supported by test data [1, 2, 3, 8], experimental data [3, 9], post-earthquake damage results called empirical data [10, 11] and analytical results given by more or less refined numerical models (most of the works cited in this article fall into this category). Moreover, in practice, three families of procedures coexist to evaluate seismic fragility curves: Incremental Dynamic Analysis (IDA), Multiple Stripes Analysis (MSA) and Cloud Analysis (CA). In most cases, these approaches are also coupled with a parameterization of the fragility curve since, compared to non-parametric estimations, parametric ones require small sample sizes. The lognormal model historically introduced in the SPRA framework [1, 2] is currently the most used (see e.g. [12, 13, 14, 15, 16]), even if its validity is questionable (see e.g. [17, 18]).

In a few words, IDA is based on scaled accelerograms until the failure threshold of interest. Fragility curves are then considered as empirical cumulative distribution functions. The main disadvantage of IDA is that excessive scaling can lead to signals with unrepresentative frequency content and duration, which can imply biased results in nonlinear structural responses [19, 20]. Strong evidence against scaling accelerograms is provided in [21] considering a more theoretical approach. Although not recommended, this approach is still implemented (see e.g. [15, 22]). MSA is based on multiple accelerograms selected or scaled to match specific IMs. Thus, for each IM value, the structural analyses provide a fraction of the ground motions that cause failure. Finally, CA is a kind of

generalization of MSA in the sense that it is based on a single accelerogram for each IM value.

35 Depending on the context, different techniques can be employed to estimate the fragility curves. For example, for parametric estimation, we distinguish: the method of moments (which is mainly used for the IDA-based methodology), the Maximum Likelihood Estimation (MLE) by assuming the independence of the observations (which can be questionable when empirical data are
40 concerned [10]), and the minimization of the Sum of Squared Errors (SSE). For non-parametric estimation, kernel smoothing can be used (see e.g. [11, 17]) as well as other methodologies (see e.g. [18, 20]). Note that most of these strategies are compared in [11, 17, 23] and [11] gives a clear presentation of the advantages and disadvantages of each of them.

45 Beyond these methods, techniques based on machine learning can also be used, including: linear regression or generalized linear regression [11], classification-based techniques [7, 24, 25], kriging [26], polynomial chaos expansion [27], artificial neural networks [25, 28, 29]. Some of them were coupled with adaptive techniques to reduce the number of calculations to be performed [25, 26]. The
50 Bayesian framework is also relevant in this context since it allows either (i) to fit numerical models (metamodels, mathematical expressions based on engineering judgments, etc.) to experimental data to directly estimate the fragility curves [9] or (ii) to use empirical data or analytical data to fit the parametric models of the fragility curves [10].

55 The major drawback of most of the work of the literature is that it does not address the issue of Confidence Intervals (CIs) related to the size of the samples used, in order to reflect the estimation uncertainty. This question arises, among others, when implementing computationally expensive numerical models of complex structures because very little data are then available. When data
60 are assumed to be from an independent and identically distributed population, the bootstrap method can be used (see e.g. [12, 16, 30]). However, when little data are involved, bootstrap samples can lead to unrealistic fragility curves such as unit-step functions. As a result, this can lead to excessively large CIs.

Moreover, when active learning techniques are implemented, the samples are
65 no longer independent and the bootstrap method cannot be used. It should
be mentioned that several works have tackled the question of the convergence
of the estimations according to the number of available data, but only from
an empirical point of view since, in most cases, it seems that there is no other
possibility (see e.g. [18, 22, 25, 30, 31]).

70 For the modeling-based approaches, the aim of this work is to address the
issue of optimized fragility curve estimations (i.e. based on few data) within
a rigorous mathematical framework. To do this, we propose and implement
an active learning methodology based on adaptive importance sampling [32]
in a statistical learning context [33], called Importance Sampling based Active
75 Learning (IS-AL).

Adaptive importance sampling was introduced for classical Monte Carlo in-
tegral approximation in [34], and later studied in [35]. Moreover, adaptive
importance sampling is also used in industrial applications and have already
been discussed, implemented and tested for probability estimation of rare event
80 (e.g. failure state) in reliability analysis [36, 37]. By applying it to the para-
metric estimations of the fragility curves, we show by asymptotic analysis and
numerical simulations that IS-AL allows for (i) a rapid convergence of the es-
timated fragility curve towards the true (unknown) fragility curve and (ii) a
rigorous quantification of the estimation uncertainty. It gives asymptotic CIs
85 and confidence ellipsoids for the quantities of interest as well as statistical tests
to determine whether the asymptotic regime has been reached and whether
asymptotic CIs and confidence ellipsoids can be used.

The proposed methodology relies on parametric approximations of fragility
curves for any IM of interest. Although the validity of parametric models is
90 both questionable and difficult to assess (see e.g. [17, 18, 38]), some numerical
experiments based on the seismic responses of simple mechanical systems - i.e
few degrees of freedom systems - suggest that the choice of an appropriate
IM makes it possible to reduce the potential biases between reference fragility
curves - that can be obtained by massive Monte Carlo simulations - and their

95 parametric approximations. This point is illustrated in the application section 6
of this paper. Remember, however, that in practice, the selection of an optimal
IM is not a trivial matter (see e.g. [39, 40]) and Machine Learning techniques
can be used for this purpose (e.g. [25]), knowing that the references [41] and
[42] give optimality criteria for selection of such IM.

100 In this work, the methodology is applied to different test cases and compared
with more traditional approaches such as MLE often used by practitioners (see
e.g. [10, 11, 12, 16, 17, 23, 30]). In order to avoid the scaling of the accelero-
grams, the stochastic model of modulated and filtered white-noise process de-
fined in [43] is used to enrich a set of real ground motion records selected in a
105 database using magnitude and distance criteria. This stochastic model is chosen
because it well encompasses the temporal and spectral non-stationarities of real
seismic signals. Additionally, it has already been used in several works (see e.g.
[17, 25, 44]).

The paper is organized as follows. In section 2 the statistical framework
110 is defined for any parametric fragility curve model and any IM. Section 3 is
dedicated to the presentation of the IS-AL algorithm applied to seismic fragility
curves estimation for the lognormal model. Section 4 summarizes the main
theoretical results of this work, which are proved in the appendices. These
results concern a criterion for evaluating the convergence of the IS-AL strategy
115 and the definition of asymptotic confidence ellipsoids for the fragility curve
parameters. Section 5 presents the performance metrics used in this work to
compare IS-AL, random sampling and MLE strategies. Finally, in section 6, IS-
AL performance is assessed on analytical and industrial test cases of increasing
complexity.

120 **2. Parametric seismic fragility curve estimation: a statistical learning framework**

2.1. General framework

We consider the following situation. Let \mathcal{X} be a compact set of \mathbb{R} , X a \mathcal{X} -valued random variable and $S \in \{0, 1\}$ a random label. In SPRA studies $X = \log \text{IM}$ – more generally we can define $X = \psi(\text{IM})$ where ψ is an increasing function of IM such as a Box-Cox transform [45] – and S is the indicator variable of the failure of the structure. The pair (X, S) has the probability distribution P over $\mathcal{X} \times \{0, 1\}$:

$$P(dx, ds) = [\mu(x)\delta_1(ds) + (1 - \mu(x))\delta_0(ds)]p(x)dx , \quad (1)$$

where δ_j is the Dirac distribution at j , p is the marginal probability density function (pdf) of X , and the fragility curve $\mu(x)$ is the conditional expectation of S (conditional probability of failure or fragility curve):

$$\mu(x) = \mathbb{E}[S|X = x] . \quad (2)$$

The aim of the paper is to estimate the curve $\mu(x)$ from datapoints $(X_i, S_i)_{i=1}^n$ that may be independent and identically distributed with the distribution P or that may be selected by a more appropriate scheme. As mentioned in the introduction, it is a classical assumption to use a parametric form for the fragility curve μ to tackle the need for time consuming mechanical simulations, we thus consider the space of functions $\mathcal{F} = \{f_\theta, \theta \in \Theta\}$, where $x \mapsto f_\theta(x)$ is a function from \mathbb{R} to $[0, 1]$ for any θ and $\Theta \subset \mathbb{R}^m$. The goal is to minimize the quadratic risk:

$$g(\theta) = \mathbb{E}[(\mu(X) - f_\theta(X))^2] , \quad (3)$$

in order to find (provided it exists and is unique):

$$\theta = \underset{\theta \in \Theta}{\operatorname{argmin}} g(\theta) . \quad (4)$$

Unfortunately, the observable data are $(X_i, S_i)_{i=1}^n$, we do not observe directly $\mu(X_i)$. But considering that:

$$\begin{aligned} \mathbb{E}[(S - f_\theta(X))^2] &= \mathbb{E}[(S - \mu(X))^2] + \mathbb{E}[(\mu(X) - f_\theta(X))^2] \\ &\quad + 2\mathbb{E}[(\mu(X) - f_\theta(X))(S - \mu(X))] \\ &= \mathbb{E}[\mu(X)(1 - \mu(X))] + \mathbb{E}[(\mu(X) - f_\theta(X))^2] \end{aligned} \quad (5)$$

because $\mathbb{E}[S^2|X] = \mathbb{E}[S|X] = \mu(X)$, we can observe that the minimization with respect to θ of $\mathbb{E}[(S - f_\theta(X))^2]$ is equivalent to the minimization of $\mathbb{E}[(\mu(X) - f_\theta(X))^2]$. Hence, we will consider the quadratic risk

$$r(\theta) = \mathbb{E}[(S - f_\theta(X))^2]. \quad (6)$$

In the context of classical learning, when we observe n datapoints $(X_i, S_i)_{i=1}^n$ drawn independently from the probability distribution $P(dx, ds)$ over $\mathcal{X} \times \{0, 1\}$, the expectation can be approximated by the empirical mean:

$$\widehat{R}_n(\theta) = \frac{1}{n} \sum_{i=1}^n (S_i - f_\theta(X_i))^2. \quad (7)$$

The corresponding passive estimator (the term passive is used to highlight the absence of any particular sampling strategy) is then:

$$\widehat{\theta}_n = \underset{\theta \in \Theta}{\operatorname{argmin}} \widehat{R}_n(\theta). \quad (8)$$

Conversely to classical learning, active learning aims at selecting the most useful numerical experiments to be carried out in order to form the learning set.

125 In the passive strategy, the datapoints X_i are sampled from the original probability distribution with pdf p drawn from a stochastic ground-motion model. In the same way as in [32], we propose an active learning strategy, called Importance Sampling based Active Learning (IS-AL). It consists to draw the datapoints X_i from an instrumental probability distribution with pdf q that is chosen in an
130 adaptive way. In our context, it is straightforward to use a rejection method applied to the stochastic ground-motion model in order to generate seismic loads with a desired intensity measure distribution. Let us recall in fact that the label S_i (which gives the failure state of the structure) is, in our case, expensive

135 to obtain because it comes from complex numerical simulations of mechanical structures while the artificial seismic signals are inexpensive to generate.

The main objective of this procedure is to reduce the variance implied by the empirical approximation of the quadratic risk $r(\theta)$. Importance Sampling is a classical variance reduction technique for Monte Carlo estimation used in structural reliability [37, 46]. If the $(X_i)_{i=1}^n$ are sampled with the pdf q and $(S_i)_{i=1}^n$ are the labels obtained from n calls to the mechanical model, then the importance sampling estimator of the empirical quadratic risk is:

$$\widehat{R}_n^{\text{IS}}(\theta) = \frac{1}{n} \sum_{i=1}^n \frac{p(X_i)}{q(X_i)} (S_i - f_\theta(X_i))^2. \quad (9)$$

In the rest of the paper, we will denote by $r(\theta) = \mathbb{E}_{(X,S) \sim P}[\ell_\theta(X, S)]$ with $\theta \mapsto \ell_\theta(x, s)$ a positive loss function for the sake of generalization. For the numerical applications, only the case of the quadratic loss

$$\ell_\theta(x, s) = (s - f_\theta(x))^2, \quad (10)$$

will be considered.

2.2. Problem regularization for the parametric lognormal model

For applications to seismic fragility curves estimation, a classical space of functions to approximate μ is $\mathcal{F} = \{\Phi(\frac{\log(\text{IM}/\alpha)}{\beta}), (\alpha, \beta) \in \Theta\}$ where Φ is the cumulative distribution function of the standard Gaussian distribution [1], $\theta = (\alpha, \beta)^T$, and Θ a compact set of $(0, +\infty)^2$ (therefore $m = 2$). Compactness of Θ is a common assumption in our applications. From an engineer perspective, it is possible to bound α and β . However, in practice, the lower bound for β may be reached by the different estimators. Consequently, inspired by Bayesian inference theory [47], we introduce a regularization term $\Omega(\theta; \beta_{\text{reg}})$ to tackle this issue (we will take $\Omega(\theta; \beta_{\text{reg}}) = \beta_{\text{reg}}/\beta$ below). The squared loss (7) is then replaced by:

$$\widehat{R}_{n,\text{reg}}(\theta; \beta_{\text{reg}}) = \frac{1}{n} \sum_{i=1}^n \ell_\theta(X_i, S_i) + \frac{\Omega(\theta; \beta_{\text{reg}})}{n}. \quad (11)$$

The derivation of the importance sampling estimator of the regularized square loss is straightforward:

$$\widehat{R}_{n,\text{reg}}^{\text{IS}}(\theta; \beta_{\text{reg}}) = \frac{1}{n} \sum_{i=1}^n \frac{p(X_i)}{q(X_i)} \ell_{\theta}(X_i, S_i) + \frac{\Omega(\theta; \beta_{\text{reg}})}{n}. \quad (12)$$

This regularization is motivated by the intrinsic difficulty of estimating the standard deviation β of the lognormal model when β is small [48]. Fragility curves with small β are hard to distinguish due to the convergence towards a degenerate 0 – 1 fragility curve.

3. Principles of the IS-AL strategy

This section focuses on the choice of an optimal density q (section 3.1) as well as on the description of the IS-AL strategy (section 3.2).

3.1. Choice of an optimal density q

The heuristic used to find a good instrumental probability distribution family is presented in [32]. The first idea would be to minimize the variance of the importance sampling risk estimator (12):

$$\text{Var}(\widehat{R}_{n,\text{reg}}^{\text{IS}}(\theta)) = \frac{1}{n} \left\{ \iint_{\mathcal{X} \times \mathcal{S}} \frac{p(x)}{q(x)} \ell_{\theta}(x, s)^2 P(dx, ds) - r(\theta)^2 \right\}, \quad (13)$$

with respect to q within the set of all pdfs. If we denote by $\tilde{\ell}_{\theta}^2(x) = \mathbb{E}[\ell_{\theta}(X, S)^2 | X = x]$ the squared loss averaged on S :

$$\tilde{\ell}_{\theta}^2(x) = \mu(x) \ell_{\theta}(x, 1)^2 + (1 - \mu(x)) \ell_{\theta}(x, 0)^2, \quad (14)$$

the variance of the importance sampling risk estimator (12) can be expressed as

$$\text{Var}(\widehat{R}_{n,\text{reg}}^{\text{IS}}(\theta)) = \frac{1}{n} \left\{ \int_{\mathcal{X}} \frac{p(x)^2}{q(x)} \tilde{\ell}_{\theta}^2(x) dx - r(\theta)^2 \right\},$$

and we look for

$$q_{\theta} = \underset{q}{\text{argmin}} \int_{\mathcal{X}} \frac{p(x)^2}{q(x)} \tilde{\ell}_{\theta}^2(x) dx. \quad (15)$$

Using Jensen’s inequality [49, Theorem 3.12], we can solve the optimization problem and we can find that the optimal sampling pdf is of the form

$$q_\theta(x) \propto \tilde{\ell}_\theta(x)p(x) ,$$

which depends on μ because $\tilde{\ell}_\theta$ depends on μ [Here and below \propto means equality up to a multiplicative constant]. Hence an approximation step is made by replacing μ by f_θ in (14):

$$\tilde{\ell}_\theta^2(x) \approx f_\theta(x)\ell_\theta(x,1)^2 + (1 - f_\theta(x))\ell_\theta(x,0)^2 . \quad (16)$$

Hence the instrumental density becomes:

$$q_\theta(x) \propto p(x)\sqrt{f_\theta(x)\ell_\theta(x,1)^2 + (1 - f_\theta(x))\ell_\theta(x,0)^2} . \quad (17)$$

Note that the instrumental distribution depends on θ , the parameter we aim to estimate. Moreover, using IS-AL with the instrumental density q_θ directly could increase the variance if the density has light tails. We propose finally a defensive strategy as illustrated in [50, 51]. The instrumental density becomes

$$q_{\theta,\varepsilon}(x) = \varepsilon p(x) + (1 - \varepsilon)q_\theta(x), \quad (18)$$

with $\varepsilon \in [0, 1]$. ε is a mixing parameter, between the original marginal pdf $p(x)$ and the instrumental one $q_\theta(x)$, meaning that one time out of $1/\varepsilon$ the element is drawn from the pdf $p(x)$. This distribution allows to bound the likelihood ratio:

$$\frac{p(x)}{q_{\theta,\varepsilon}(x)} = \frac{1}{\varepsilon + (1 - \varepsilon)\frac{q_\theta(x)}{p(x)}} < \frac{1}{\varepsilon} . \quad (19)$$

Thus the defensive strategy bounds the variance even if the likelihood ratio $p(x)/q_\theta(x)$ is large.

3.2. Description of the IS-AL strategy

3.2.1. Algorithm

150 The procedure for computing the IS-AL estimator $\widehat{\theta}_n^{\text{IA}}$ is described in Algorithm 1. Its main objective is to use an updated instrumental density $q_{\theta,\varepsilon}$ at

each step. Note that (i) the algorithm needs to start from a certain parameter value $\widehat{\theta}_0^{\text{IA}}$ and (ii) we choose $\Omega(\theta; \beta_{\text{reg}}) = \beta_{\text{reg}}/\beta$ for the regularization term in equation (12).

Algorithm 1 Importance Sampling based Active Learning (IS-AL)

1. Choice of $\widehat{\theta}_0^{\text{IA}}$ (section 3.2.2) and estimations of β_{reg} and ε (section 3.2.3).
2. For $i = 1, \dots, n$:
 - (a) Draw X_i from the distribution with pdf $q_{\widehat{\theta}_{i-1}^{\text{IA}}, \varepsilon}$.
 - (b) Call the mechanical simulation at point X_i to get label S_i
 - (c) Compute

$$\widehat{\theta}_i^{\text{IA}} = \underset{\theta \in \Theta}{\text{argmin}} \widehat{R}_{i,\text{reg}}^{\text{IA}}(\theta; \beta_{\text{reg}}), \quad (20)$$

$$\widehat{R}_{i,\text{reg}}^{\text{IA}}(\theta; \beta_{\text{reg}}) = \frac{1}{i} \sum_{j=1}^i \frac{p(X_j)}{q_{\widehat{\theta}_{j-1}^{\text{IA}}, \varepsilon}(X_j)} \ell_{\theta}(X_j, S_j) + \frac{\beta_{\text{reg}}}{n\beta}. \quad (21)$$

155 Additionally, a convergence criterion is presented in section 4.2 and an asymptotic confidence ellipsoid for θ centered on $\widehat{\theta}_n^{\text{IA}}$ is defined by equation (36).

3.2.2. Initialization and choice of $\widehat{\theta}_0^{\text{IA}}$

Regarding the initialization, as expected, the closer $\widehat{\theta}_0^{\text{IA}}$ is from the true parameter θ the faster IS-AL is in asymptotic normal regime. A naive approach
160 is to get a small sample of size n_0 (e.g. $n_0 = 20$) $(X_i, S_i)_{i=1}^{n_0}$ from the original marginal density p of X and to compute the passive learning estimator $\widehat{\theta}_{n_0}$ (equation (8)). This crude estimation can be used as the initial parameter $\widehat{\theta}_0^{\text{IA}}$ to start IS-AL.

A better approach is to consider a metamodel - in the broad sense - of the
165 mechanical simulation. As often used by practitioners, a numerical resolution based on a modal base projection can be implemented to get an estimate of the fragility curve corresponding to the linear behavior of the structure of interest. It is then possible to get a huge amount of datapoints of the reduced model (e.g. an independent and identically distributed sample of $n_{\text{red}} = 10^3\text{--}10^5$ pairs

170 $(X_i, S_{\text{red},i})_{i=1}^{n_{\text{red}}}$ where X_i is sampled with the original pdf p and $S_{\text{red},i}$ is the associated label obtained with the reduced model). The initial parameter $\hat{\theta}_0^{\text{IA}}$ is then chosen to be equal to $\hat{\theta}_{n_{\text{red}}}$. Statistical metamodels could also be used such as Gaussian Processes [52] or Support Vector Machines [25].

In our applications reduced models are only used to give us prior knowledge
 175 on the fragility curve shape, encapsulated in the initial parameter of the IS-AL procedure. We then initialize IS-AL with a small sample of 20 datapoints with the instrumental density $q_{\hat{\theta}_0^{\text{IA}}, \varepsilon}$ (equation (18)). In other words, in Step 2 of Algorithm 1, we do not update $\hat{\theta}_i^{\text{IA}}$ during the first 20 steps.

3.2.3. Estimations of β_{reg} and ε

180 The regularization parameter, called $\hat{\beta}_{\text{reg}}^{\text{IA}}$, is determined by minimizing the Leave One Out error on the initialization sample (see previous section).

Regarding the choice of the defensive parameter value ε , it is cumbersome and there is no direct methodology for its estimation. Moreover, its value depends strongly of the problem studied as shown in [53]. Nevertheless, in section
 185 6.2, we propose a benchmark in order to evaluate the "optimal" value of ε for the class of structures considered in this study.

4. Theoretical results

This section summarizes the main theoretical results of this work. Section 4.1 addresses the issue of the consistency and asymptotic normality for the IS-AL
 190 estimator. Then, in section 4.2, a convergence criterion is proposed in order to be able to use the asymptotic confidence ellipsoids defined in section 4.3. A discussion is finally proposed about the practical use of the convergence criterion in section 4.4.

4.1. Consistency and asymptotic convergence of the IS-AL estimator

195 We derive some theoretical properties for the estimator $\hat{\theta}_n^{\text{IA}}$, consisting in its consistency towards the parameter θ defined by (4) and its asymptotic normality by adapting several proofs of [54] about asymptotic optimality of

adaptive importance sampling. Detailed proofs of the following results are given in the Appendix. The proofs are given in a more general context of empirical risk minimization, instead of IS-AL specifically. Indeed, we consider that these theoretical results can be used in a broader manner for other kind of applications.

We first prove in Appendix A the consistency of the IS-AL estimator $\widehat{\theta}_n^{\text{IA}}$ using Algorithm 1:

$$\widehat{\theta}_n^{\text{IA}} \xrightarrow[n!+1]{} \theta \text{ in probability .} \quad (22)$$

Then, we prove in Appendix B the convergence of $\sqrt{n}(\widehat{\theta}_n^{\text{IA}} - \theta)$ to a Gaussian random variable with mean zero and covariance matrix:

$$G_{\theta_*, \varepsilon} = \ddot{r}(\theta)^{-1} V(q_{\theta_*, \varepsilon}) (\ddot{r}(\theta)^{-1})^T, \quad (23)$$

where

$$V(q_{\theta_*, \varepsilon}) = \mathbb{E} \left[\frac{p(X)}{q_{\theta_*, \varepsilon}(X)} \ell_{\theta_*}(X, S) \nabla f_{\theta_*}(X) \nabla f_{\theta_*}(X)^T \right], \quad (24)$$

and $\ddot{r}(\theta)$ is the Hessian of $r(\theta)$ at θ .

A straightforward corollary of equation (23) is that, if $G_{\theta_*, \varepsilon}$ is nonsingular (which we assume from now on), then for any $\xi \in (0, 1)$:

$$\mathbb{P} \left(n(\widehat{\theta}_n^{\text{IA}} - \theta)^T G_{\theta_*, \varepsilon}^{-1} (\widehat{\theta}_n^{\text{IA}} - \theta) < q_{\chi^2(m)}^\xi \right) \xrightarrow[n!+1]{} \xi, \quad (25)$$

with $q_{\chi^2(m)}^\xi$ the ξ -quantile of the $\chi^2(m)$ distribution (remember that $\theta = (\alpha, \beta)^T$ and $m = 2$ for the lognormal model). Remark that the matrix $G_{\theta_*, \varepsilon}$ depends on the unknown parameter θ . It is thus possible to use a plug-in estimator:

$$\widehat{G}_n = \widehat{\ddot{r}}_n(\widehat{\theta}_n^{\text{IA}})^{-1} \widehat{V}_n(\widehat{\theta}_n^{\text{IA}}) (\widehat{\ddot{r}}_n(\widehat{\theta}_n^{\text{IA}})^{-1})^T, \quad (26)$$

with

$$\widehat{\ddot{r}}_n(\theta) = \frac{1}{n} \sum_{i=1}^n \frac{p(X_i)}{q_{\widehat{\theta}_{i-1}^{\text{IA}}, \varepsilon}(X_i)} \ddot{\ell}_\theta(X_i, S_i), \quad (27)$$

$$\widehat{V}_n(\theta) = \frac{1}{n} \sum_{i=1}^n \frac{p(X_i)^2}{q_{\theta, \varepsilon}(X_i) q_{\widehat{\theta}_{i-1}^{\text{IA}}, \varepsilon}(X_i)} \dot{\ell}_\theta(X_i, S_i) \dot{\ell}_\theta(X_i, S_i)^T, \quad (28)$$

and $\ddot{\ell}_\theta(x, s)$ the Hessian of $\ell_\theta(x, s)$ with respect to θ . We have:

$$\widehat{G}_n^{-1} \rightarrow G_{\theta_*, \varepsilon}^{-1} \text{ in probability.} \quad (29)$$

The proof is in Appendix C. Using asymptotic normality of $\hat{\theta}_n^{\text{IA}}$, we can show that: $n(\hat{\theta}_n^{\text{IA}} - \theta)^T G_{\theta_*, \varepsilon}^{-1}(\hat{\theta}_n^{\text{IA}} - \theta) \rightarrow \chi^2(m)$. Using Slutsky's lemma, we have the following convergence in distribution:

$$n(\hat{\theta}_n^{\text{IA}} - \theta)^T \hat{G}_n^{-1}(\hat{\theta}_n^{\text{IA}} - \theta) \xrightarrow{n! + 1} \chi^2(m). \quad (30)$$

4.2. Convergence criterion using a statistical hypothesis test

The estimation of the generalization error without a validation set is often based on Cross Validation. When IS-AL is used, the data points (X_i, S_i) are no longer independent and identically distributed. We propose to use a convergence criterion that ensures that asymptotic normality is reached. Consider two independent datasets $\mathcal{D}_1 = (X_{i,1}, S_{i,1})_{i=1}^n$ and $\mathcal{D}_2 = (X_{i,2}, S_{i,2})_{i=1}^n$ generated with IS-AL. Let $\hat{R}_{n,\text{reg},j}^{\text{IA}}$ be the weighted loss for \mathcal{D}_j for $j = 1, 2$ defined as in (20). Denote:

$$\hat{\theta}_{n,j}^{\text{IA}} = \underset{\theta \in \Theta}{\text{argmin}} \hat{R}_{n,\text{reg},j}^{\text{IA}}(\theta; \beta_{\text{reg}}), \quad j = 1, 2.$$

Then we have:

$$\sqrt{n}(\hat{R}_{n,\text{reg},1}^{\text{IA}}(\hat{\theta}_{n,2}^{\text{IA}}; \beta_{\text{reg}}) - \hat{R}_{n,\text{reg},2}^{\text{IA}}(\hat{\theta}_{n,1}^{\text{IA}}; \beta_{\text{reg}})) \xrightarrow{L} \mathcal{N}(0, 8V(q_{\theta_*, \varepsilon}, \dot{\ell}_{\theta_*})) \quad (31)$$

as $n \rightarrow +\infty$. Denote

$$\widehat{W}_n = \frac{n}{8} (\hat{R}_{n,\text{reg},1}^{\text{IA}}(\hat{\theta}_{n,2}^{\text{IA}}; \beta_{\text{reg}}) - \hat{R}_{n,\text{reg},2}^{\text{IA}}(\hat{\theta}_{n,1}^{\text{IA}}; \beta_{\text{reg}}))^T \widehat{V}_{n,12}^{-1} (\hat{R}_{n,\text{reg},1}^{\text{IA}}(\hat{\theta}_{n,2}^{\text{IA}}) - \hat{R}_{n,\text{reg},2}^{\text{IA}}(\hat{\theta}_{n,1}^{\text{IA}})), \quad (32)$$

$$\widehat{V}_{n,12} = \frac{1}{2} (\widehat{V}_{n,1}(\hat{\theta}_{n,1}^{\text{IA}}) + \widehat{V}_{n,2}(\hat{\theta}_{n,2}^{\text{IA}})), \quad (33)$$

with $\widehat{V}_{n,j}$ the empirical estimator in equation (28) for the j -th IS-AL dataset \mathcal{D}_j for $j = 1, 2$.

By equation (31) and by Slutsky's lemma, \widehat{W}_n converges weakly to $\chi^2(m)$. It is, therefore, possible to define a convergence criterion inspired by statistical test theory to check the asymptotic normality of $\hat{\theta}_n^{\text{IA}}$. Our convergence criterion is equivalent to the hypothesis test:

$$(\mathcal{H}_0) : \widehat{W}_n \text{ follows } \chi^2(m) \quad \text{against} \quad (\mathcal{H}_1) : \widehat{W}_n \text{ does not follow } \chi^2(m). \quad (34)$$

For $\xi \in (0, 1)$, we then consider the statistical test which rejects (\mathcal{H}_0) if:

$$\widehat{W}_n > q_{1-\xi}^{\chi^2(m)}, \quad (35)$$

where $q_{1-\xi}^{\chi^2(m)}$ denotes the $(1-\xi)$ -quantile of the $\chi^2(m)$ distribution. Hence, this statistical test is of asymptotic level ξ .

4.3. Asymptotic confidence ellipsoid

Thanks to the equation (30), it is possible to construct an asymptotic confidence ellipsoid of level $\xi \in (0, 1)$ defined by:

$$\mathcal{E}_{n,\xi}^{\text{IA}} = \{\theta : n(\theta - \widehat{\theta}_n^{\text{IA}})^T \widehat{G}_n^{-1}(\theta - \widehat{\theta}_n^{\text{IA}}) < q_{1-\xi}^{\chi^2(m)}\}, \quad (36)$$

with:

$$\mathbb{P}(\theta \in \mathcal{E}_{n,\xi}^{\text{IA}}) \xrightarrow[n \rightarrow \infty]{} \xi.$$

Because the convergence criterion \widehat{W}_n indicates when the estimator follows the asymptotic Gaussian distribution, it also indicates at which sample size n the value $\mathbb{P}(\theta \in \mathcal{E}_{n,\xi}^{\text{IA}})$ is close to its theoretical value ξ .

4.4. Discussion about the practical use of the convergence criterion

An apparent drawback of this convergence criterion is that it doubles the computational cost, due to the necessity of having two independent IS-AL estimators $\widehat{\theta}_{n,1}^{\text{IA}}$ and $\widehat{\theta}_{n,2}^{\text{IA}}$ to compute \widehat{W}_n . It is, however, possible to use the estimator

$$\widehat{\theta}_{n,12}^{\text{IA}} = \frac{\widehat{\theta}_{n,1}^{\text{IA}} + \widehat{\theta}_{n,2}^{\text{IA}}}{2}, \quad (37)$$

which has an asymptotic variance that is half the one of $\widehat{\theta}_{n,1}^{\text{IA}}$ and $\widehat{\theta}_{n,2}^{\text{IA}}$. Indeed, it is straightforward that $\sqrt{n}(\widehat{\theta}_{n,12}^{\text{IA}} - \theta)$ converges in distribution to a zero mean Gaussian random variable with covariance matrix $G_{\theta^*, \varepsilon}/2$. It is, therefore, possible to define an asymptotic confidence ellipsoid which exploits all the data points used to build the estimator $\widehat{\theta}_{n,12}^{\text{IA}}$ of θ :

$$\mathcal{E}_{n,12,\xi}^{\text{IA}} = \{\theta : 2n(\theta - \widehat{\theta}_{n,12}^{\text{IA}})^T \widehat{G}_{n,12}^{-1}(\theta - \widehat{\theta}_{n,12}^{\text{IA}}) < q_{1-\xi}^{\chi^2(m)}\},$$

with $\widehat{G}_{n,12} = \widehat{r}_{n,12}^{-1} \widehat{V}_{n,12} (\widehat{r}_{n,12}^{-1})^T$, $\widehat{V}_{n,12}$ defined by (33), $\widehat{r}_{n,12} = \frac{1}{2} \widehat{r}_{n,1}(\widehat{\theta}_{n,1}^{\text{IA}}) + \frac{1}{2} \widehat{r}_{n,2}(\widehat{\theta}_{n,2}^{\text{IA}})$ and $\widehat{r}_{n,j}$ defined as (27) with the dataset \mathcal{D}_j , $j = 1, 2$.

215 **5. Performance evaluation of the IS-AL strategy compared to the
random sampling and MLE strategies**

This section explains how to assess the performance of IS-AL with respect to Random Sampling (RS) and MLE strategies. In section 5.1, RS and MLE principles are briefly summarized. Performance metrics inspired from [55, 56] to check the quality of IS-AL strategy are detailed in section 5.2. Finally, the statistical procedure used to assess the quality of the IS-AL asymptotic confidence ellipsoid compared to that of a classical approach such as MLE is given in section 5.3.

220 *5.1. RS and MLE principles*

RS strategy consists in applying the IS-AL algorithm with the proposal probability density q being the marginal probability density p of the intensity measure. This boils down to classical empirical risk minimization for supervised learning. The RS estimator $\hat{\theta}_n^{\text{RS}}$ is then defined by:

$$\hat{\theta}_n^{\text{RS}} = \underset{\theta \in \Theta}{\operatorname{argmin}} \hat{R}_{n,\text{reg}}^{\text{RS}}(\theta; \beta_{\text{reg}}), \quad (38)$$

$$\hat{R}_{n,\text{reg}}^{\text{RS}}(\theta; \beta_{\text{reg}}) = \frac{1}{n} \sum_{i=1}^n \ell_{\theta}(X_i, S_i) + \frac{\beta_{\text{reg}}}{n\beta}. \quad (39)$$

225

As mentioned in the introduction, MLE is a classical estimation method in the field of seismic probabilistic risk assessment and fragility curve estimation (see e.g. [10, 11, 12, 16, 17, 23, 30]). It is defined by the estimator $\hat{\theta}_n^{\text{MLE}}$ that maximizes the likelihood given a dataset $(X_i, S_i)_{i=1}^n$ that is sampled at random from the original marginal density p of X :

$$\hat{\theta}_n^{\text{MLE}} = \underset{\theta \in \Theta}{\operatorname{argmax}} \sum_{i=1}^n S_i \log(f_{\theta}(X_i)) + (1 - S_i) \log(1 - f_{\theta}(X_i)). \quad (40)$$

The initializations of the RS and MLE algorithms are based on 20 data points drawn at random from the original distribution p . For the RS algorithm, the regularization parameter, called $\hat{\beta}_{\text{reg}}^{\text{RS}}$, is computed using Leave One Out cross validation as for the IS-AL algorithm.

230 *5.2. Performance metrics for the numerical benchmarks*

This section aims to provide performance metrics, inspired from [55, 56], to assess IS-AL performances, in comparison with the RS and MLE strategies, on test cases.

5.2.1. *Performance metrics based on the training errors*

235 For the IS-AL strategy, the training error is called $\widehat{R}_n^{\text{IA}} = \widehat{R}_{n,\text{reg}}^{\text{IA}}(\widehat{\theta}_n^{\text{IA}}; \beta_{\text{reg}}^{\text{IA}})$ and is defined by equation (21).

For the RS and MLE strategies, the training errors are respectively called $\widehat{R}_n^{\text{RS}} = \widehat{R}_{n,\text{reg}}^{\text{RS}}(\widehat{\theta}_n^{\text{RS}}; \beta_{\text{reg}}^{\text{RS}})$ and $\widehat{R}_n^{\text{MLE}} = \widehat{R}_{n,\text{reg}}^{\text{MLE}}(\widehat{\theta}_n^{\text{MLE}}; \beta_{\text{reg}}^{\text{RS}})$, and are defined by :

$$\widehat{R}_{n,\text{reg}}(\theta; \beta_{\text{reg}}^{\text{RS}}) = \frac{1}{n} \sum_{i=1}^n \ell_{\theta}(X_i, S_i) + \frac{\beta_{\text{reg}}^{\text{RS}}}{n\beta}$$

where \bullet is for RS or MLE. Note that for MLE the penalization $\widehat{\beta}_{\text{reg}}^{\text{RS}}$ is only used to define similar training errors as for IS-AL and RS algorithms, in order to compare the same quantity.

240 Thus, the performance metrics are :

- the *Relative Standard Deviation*

$$\text{RSD}_n = \frac{\sqrt{\text{V}[\widehat{R}_n]}}{\text{E}[\widehat{R}_n]}, \quad (41)$$

where \bullet is for IA, RS and MLE.

- the *Relative Bias*

$$\text{RB}_n = \frac{|\text{b} - \text{E}[\widehat{R}_n]|}{\text{b}}, \quad (42)$$

where \bullet is for IA, RS and MLE, and $\text{b} = \text{E}[\mu(X)(1 - \mu(X))]$.

- The *efficiency*

$$\nu_n = \frac{\text{V}[\widehat{R}_n]}{\text{V}[\widehat{R}_n^{\text{IA}}]}, \quad (43)$$

where \bullet is for RS and MLE. A value of $\nu_n > 1$ shows that IS-AL has a smaller loss variance than RS or MLE.

245 The above metrics are empirically calculated using R replications of the three procedures (IS-AL, RS and MLE) and b is estimated using a massive Monte Carlo estimator of μ . In practice, this is only possible when the numerical models are not numerically expensive. This is the case for the first two numerical test cases addressed in Section 6, but not for the third industrial case.

250 *5.2.2. Performance metrics based on the testing errors*

Additionally, we define the testing error by:

$$\widehat{Q}_{n,\text{reg}}(\theta; \beta_{\text{reg}}) = \frac{1}{n_t} \sum_{i=1}^{n_t} \ell_{\theta}(X_i^{(t)}, S_i^{(t)}) + \frac{\beta_{\text{reg}}}{n_t \beta},$$

where $(X_i^{(t)}, S_i^{(t)})_{i=1}^{n_t}$ is a testing set (independently and identically distributed with the original distribution P). The associated performance metrics are calculated by replacing $\widehat{R}_{n,\text{reg}}(\theta; \beta_{\text{reg}})$ with $\widehat{Q}_{n,\text{reg}}(\theta; \beta_{\text{reg}})$ (also called \widehat{Q}_n in the following) as defined in section 5.2.1.

255 *5.3. Benchmark on the confidence ellipsoids IS-AL and MLE*

This section aims to propose a procedure to evaluate the quality of the IS-AL asymptotic confidence ellipsoid, $\mathcal{E}_{n,\xi}^{\text{IA}}$, compared to that of the MLE, $\mathcal{E}_{n,\xi}^{\text{MLE}}$. This procedure is based on the use of R replications of the IS-AL and MLE algorithms, as for the evaluation of the performance metrics. So, we first define 260 the empirical estimator of $\mathbb{P}\left(\theta \in \mathcal{E}_{n,\xi}^{\text{IA}}\right)$ (resp. $\mathbb{P}\left(\theta \in \mathcal{E}_{n,\xi}^{\text{MLE}}\right)$), namely the Coverage Probability (CP), in order to numerically (i) verify the definitions of the ellipsoids and (ii) evaluate their convergences with respect to the size n of the samples. Then, to quantify the effectiveness of the IS-AL strategy on reducing the variance of the estimate of the fragility curve, compared to that of 265 the MLE, we define and compare their Confidence Ellipsoid Volumes (CEVs).

Section 5.3.1 concerns the definitions of the confidence ellipsoid and the coverage probability for MLE. As the IS-AL confidence ellipsoid is defined in section 4.3, section 5.3.2 deals only with the associated CP. Finally section 5.3.3 defines the CEVs for both procedures.

270 5.3.1. Confidence ellipsoid and coverage probability for MLE

In order to define the asymptotic confidence ellipsoid for the MLE and to compute the associated CP, we use (i) the asymptotic normality of the MLE estimator [57] and (ii) the independence property of the samples that allows the use of the bootstrap method.

We first consider R replications of MLE estimator $\hat{\theta}_n^{\text{MLE}}$ for different sample size n and build B bootstrap samples of size n for each replication in order to compute a bootstrap covariance:

$$\hat{V}_{n,r}^{\text{MLE}} = \frac{1}{B} \sum_{b=1}^B n(\theta_{b,r}^{\text{MLE}} - \hat{\theta}_{n,r}^{\text{MLE}})(\theta_{b,r}^{\text{MLE}} - \hat{\theta}_{n,r}^{\text{MLE}})^T, \quad (44)$$

275 where $\hat{\theta}_{n,r}^{\text{MLE}}$ is the MLE estimator for the r -th replication of size n and $\theta_{b,r}^{\text{MLE}}$ the bootstrap MLE estimator for the b -th bootstrap sample of the r -th replication.

Thus, the bootstrapped confidence ellipsoid for MLE is defined by:

$$\mathcal{E}_{n,\xi}^{\text{MLE},r} = \{\theta : n(\theta - \hat{\theta}_{n,r}^{\text{MLE}})^T (\hat{V}_{n,r}^{\text{MLE}})^{-1} (\theta - \hat{\theta}_{n,r}^{\text{MLE}}) \leq q_1^{\chi^2(m)}\}, \quad (45)$$

while the bootstrap CP writes:

$$\text{CP}_n^{\text{MLE},r} = \frac{1}{R} \sum_{r=1}^R \mathbf{1}_{\theta_* \in \mathcal{E}_{n,\xi}^{\text{MLE},r}}. \quad (46)$$

5.3.2. Coverage probability for IS-AL

The IS-AL asymptotic confidence ellipsoid $\mathcal{E}_{n,\xi}^{\text{IA}}$ is defined in section 4.3 by equation (36). So, as for MLE, the associated CP is computed by considering R replications of IS-AL, namely:

$$\text{CP}_n^{\text{IA},r} = \frac{1}{R} \sum_{r=1}^R \mathbf{1}_{\theta_* \in \mathcal{E}_{n,\xi}^{\text{IA},r}},$$

280 where $\mathcal{E}_{n,\xi}^{\text{IA},r}$ is the asymptotic confidence ellipsoid of the r -th replication of the IS-AL procedure of size n .

5.3.3. Confidence ellipsoid volumes for IS-AL and MLE

A qualitative criterion to measure the sharpness of a confidence ellipsoid is its volume [58]. So, to evaluate the effectiveness of the IS-AL strategy on the reduction of the variance of the fragility curve estimations, we define the CEVs, for respectively the MLE and IS-AL strategies, as follows:

$$\text{CEV}_n^{\text{MLE},r} = \det \left(\frac{\widehat{V}_{n,r}^{\text{MLE}}}{n} \right), \quad (47)$$

and

$$\text{CEV}_n^{\text{IA},r} = \det \left(\frac{\widehat{G}_{n,r}}{n} \right), \quad (48)$$

where $\widehat{G}_{n,r}$ is the estimated covariance matrix (26) of the r -th replication of IS-AL procedure of size n .

6. Numerical results

285 To evaluate IS-AL efficiency, a numerical benchmark has been performed with three test cases with increasing complexity:

- 1) a synthetic test case with known fragility curve and probability distribution of the seismic log-intensity measure X ,
- 2) a nonlinear elasto-plastic oscillator with kinematic hardening subjected to 290 synthetic signals generated from the modulated and filtered white-noise ground-motion model [43], as in [25],
- 3) an industrial test case of a nuclear facility's pipeline-system, submitted to the same artificial signals.

For test cases 2 and 3, 97 acceleration records selected from the European 295 Strong Motion Database [59] in the domain $5.5 < M < 6.5$ and $R < 20\text{km}$ - where M is the magnitude and R the distance from the epicenter - are considered in order to identify the parameters of the ground-motion model. 10^5 realizations of synthetic signals are then generated to form the unlabeled pool.

The oscillator test case aims to evaluate the effectiveness of the IS-AL strat- 300 egy before its application to an industrial test case which is numerically much

more costly. Moreover, since it well represents the essential features of the non-linear responses of a large variety of real structures subjected to earthquakes, this test case allows to determine the value of the hyperparameter ε - thanks to a numerical benchmark - because there is no ad hoc procedure to do this.

305 *6.1. Synthetic test case*

Here we benchmark our methodology while having full knowledge of the true fragility curve. We generate 30,000 datapoints (X_i, S_i) with the fragility curve $\mu(x) = \Phi\left(\frac{x - \log(\alpha_*)}{\beta_*}\right)$ with $(\alpha, \beta) = (0.3, 0.4)$. The original marginal distribution of X is here a Gaussian distribution with mean $\log\left(\frac{\alpha_*}{5}\right)$ and variance 1.69.
 310 The parameters have been chosen so that the data generated are qualitatively close to the nonlinear oscillator test case presented in section 6.2. The unlabeled pool consists of 20,000 datapoints X_i . 10,000 datapoints (X_i, S_i) will be our validation set for testing error estimation, using crude Monte Carlo.

Figure 1 shows (i) the target fragility curve μ in dashed red line, (ii) a kernel density estimation of the density p based on the whole dataset in green and
 315 (iii) a kernel density density estimation q of the 120 datapoints X_i obtained by IS-AL in red.

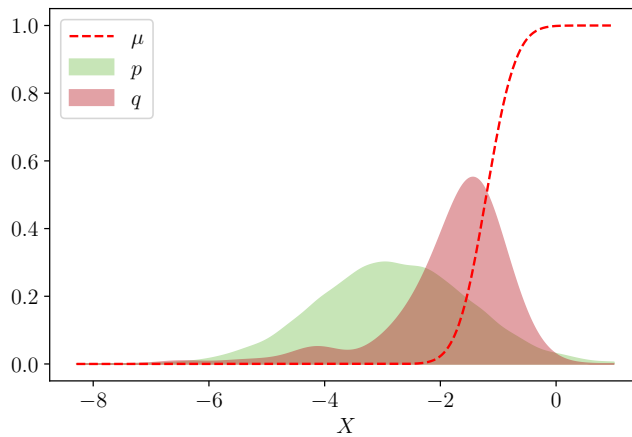


Figure 1: Synthetic test case with lognormal fragility curve with parameters $(\alpha, \beta) = (0.3, 0.4)$ and $X \sim N\left(\frac{\alpha_*}{5}, 1.69\right)$. Comparison of the original marginal density p of X with the empirical density q of the $n = 120$ datapoints X_i obtained by IS-AL.

Figure 2 shows the training and testing errors for $R = 500$ replications of the IS-AL, RS and MLE algorithms. The algorithms are initialized with 20 datapoints and $n = 100$ datapoints are extracted from the unlabeled pool with the three procedures. The regularization parameters $\beta_{\text{reg}} \in (10^{-4}, 10^{-1})$ were determined by cross validation with the 20 datapoints used for initialization for each replication of the IS-AL and MLE strategies. We also use a defensive parameter value ε of 10^{-3} (see section 6.2 for justification).

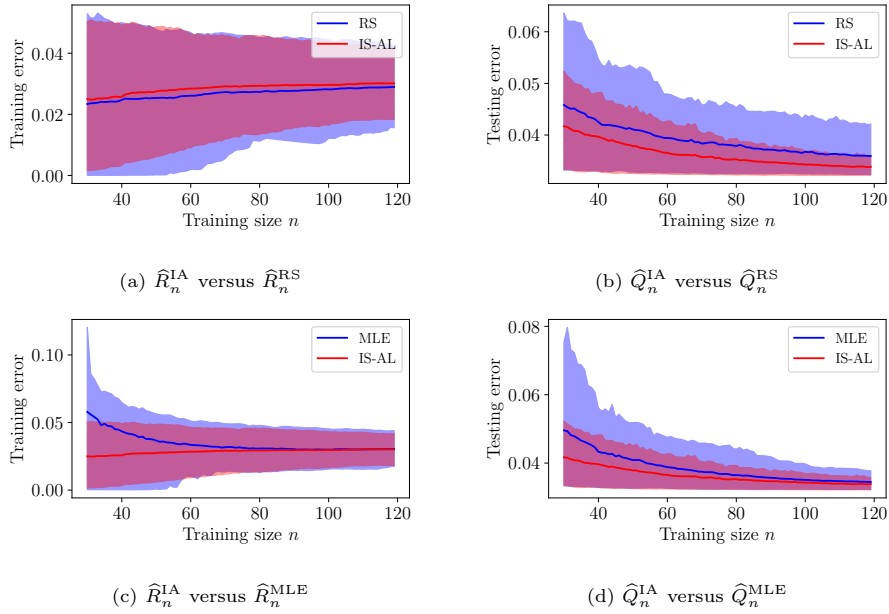


Figure 2: Results of the numerical benchmark for the synthetic test case: the thick lines represent the mean loss for $R = 500$ replications, the shaded areas represent the ranges between the quantiles at 90% and 10% of the 500 replications of the IS-AL, RS and MLE procedures. For this case, the bias is known and is equal to $E[\mu(X)(1 - \mu(X))] = 0.032$.

As depicted by Figure 2 and Table 1, IS-AL does not seem to reduce the training error. This result is normal because IS-AL selects seisms whose intensity measures maximize $\tilde{\ell}_\theta$, which can be seen as a marginalized training loss variance of the observations. In other words, as illustrated in Figure 1 with the density $q(x)$, IS-AL selects "difficult" points - typically values of x for which $\mu(x)$ takes values between 0 and 1 - and therefore the training error can be large because

Table 1: Performance metrics for the synthetic test case for $n = 120$

•	Train			Test		
	RS	MLE	IS-AL	RS	MLE	IS-AL
RSD ₁₂₀ (%)	38	36	34	12	9	8.5
ν_{120}	1.2	1.2	×	2.3	1.1	×
RB ₁₂₀ (%)	8.6	4.2	5	13	9	6.7

it is not representative of the generalization error as the testing one. RS, MLE and IS-AL strategies really distinguish themselves on the testing error, which is smaller for IS-AL. Moreover, IS-AL quickly converges to the known bias equal to $\mathbb{E}[\mu(X)(1 - \mu(X))] \simeq 0.032$. In comparison with RS and MLE strategies, the variance of IS-AL is smaller after 120 iterations: ν_{120}^{MLE} is smaller than ν_{120}^{RS} , meaning that MLE is competitive with IS-AL in this synthetic case.

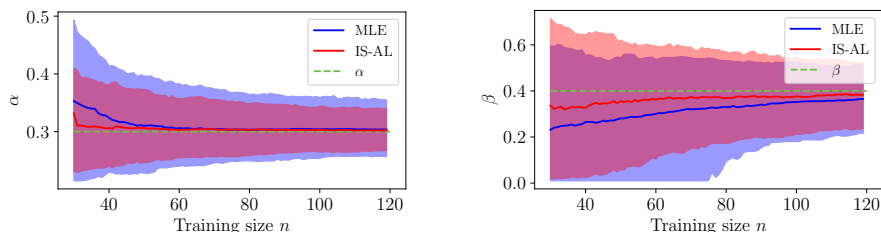


Figure 3: Results of the numerical benchmark for the synthetic test case: empirical distributions of the parameters α and β are represented by ranges between the empirical 10% and 90% quantiles of 500 replications. The shaded blue and red areas correspond respectively to MLE and IS-AL. The dashed green lines correspond to the true parameters α and β .

Figure 3 compares the distributions of the parameters α and β for several sample sizes using the 500 replications of MLE and IS-AL estimators and they are similar when $n > 100$. Indeed, the statistical model is in this case well specified (i.e. failure events follow a Bernoulli distribution with a lognormal probability of failure) and thus MLE is supposed to perform well as shown in [57]. Note that up to $n = 80$, the MLE strategy can produce degenerate

fragility curves because $\beta \simeq 0$. The IS-AL algorithm avoids this pitfall due to the regularization parameter.

345 *6.2. A nonlinear oscillator*

This test case aims to validate the overall strategy developed in this work on a simple but representative case, because this is not possible for complex structures like the one in section 6.3. This section is therefore particularly comprehensive, from the initialization of the IS-AL algorithm to the estimations
350 of the fragility curves, via the choice ε and the numerical verification of the theorems.

6.2.1. Presentation of the oscillator

This second test case - illustrated in Figure 4 - relates to a single degree of freedom elasto-plastic oscillator which exhibits kinematic hardening. It has been used in previous studies such as [18, 25]. For a unit mass m , its equation of motion is:

$$\ddot{z}(t) + 2\zeta\omega_L\dot{z}(t) + f_{NL}(t) = -s(t) ,$$

with $s(t)$ an artificial seismic signal. $\dot{z}(t)$ and $\ddot{z}(t)$ are respectively the velocity and the acceleration of the mass while ζ is the damping ratio and ω_L the pulsa-
355 tion of the oscillator. The nonlinear force f_{NL} is governed by two parameters: the post-yield stiffness, a , and the yield displacement, Y .

With this model, the quantity of interest is the maximum displacement of the mass, $D = \max_{t \in [0, T]} |z(t)|$, where T is the duration of the seismic excitation. The failure state is then defined by the $\{0, 1\}$ -valued variable $S = \mathbf{1}_{(D > C)}$, where
360 $C = 2Y$ is chosen to be approximately the 90% quantile of the maximal linear displacement of the unlabeled pool of size 10^5 .

In order to check the performances of the IS-AL algorithm, the unlabeled training set consists in $9 \cdot 10^4$ seismic signals and the testing set is composed of 10^4 signals. The benchmark study consists in $R = 500$ replications with $n = 120$
365 sampled seismic signals using IS-AL (that includes the initial 20 points) and 120 for the RS and MLE strategies.

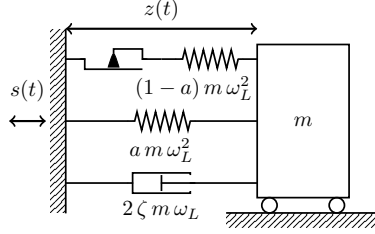


Figure 4: Elasto-plastic mechanical oscillator with kinematic hardening, with parameters $f_L = 5$ Hz and $\zeta = 2\%$. The yield limit is $Y = 5.10^{-3}$ m and the post-yield stiffness is 20% of the elastic stiffness, hence $a = 0.2$.

6.2.2. Initialization of the IS-AL procedure

In this test case, for IS-AL initialization, we use the underlying elastic oscillator as a cheap model. The initialization parameter $\hat{\theta}_0^{\text{LA}}$ is approximated by $\hat{\theta}_{10^5}^{\text{RS}}$ (equation (8)) using the 10^5 -sized dataset. In addition, the PGA is first considered as IM. Even if the PGA is not known to be the best indicator, doing so helps to verify the relevance of the methodology in a "less favorable" case. Note that the influence of the IM on the results is discussed in section 6.2.8. As shown in Figure 5, the parameter $\hat{\theta}_0^{\text{LA}}$ could be considered "close to" the true parameter θ . Thus, 20 datapoints are queried on the nonlinear oscillator with the instrumental density $q_{\hat{\theta}_0^{\text{LA}}, \varepsilon}$ (equation (18)) before launching the adaptive strategy.

6.2.3. Choice of ε

As mentioned in section 3.2.3, there is no direct methodology for the choice of the ε value. One thus benefits from this simple test case to implement a numerical benchmark in order to obtain a reasonable value of ε for the class of structures for which the oscillator represents the global nonlinear behavior under seismic excitation. This benchmark consists in evaluating the IS-AL efficiency with respect to the RS strategy, ν_n^{RS} (equation (43)), as a function of ε when IM = PGA and $n = 120$. Results are given in Table 2.

They show that ν_{120}^{RS} does not change between $\varepsilon = 10^{-2}$ and $\varepsilon = 10^{-3}$. ν_{120}^{RS} is smaller when $\varepsilon = 10^{-1}$ meaning that this value is too conservative because

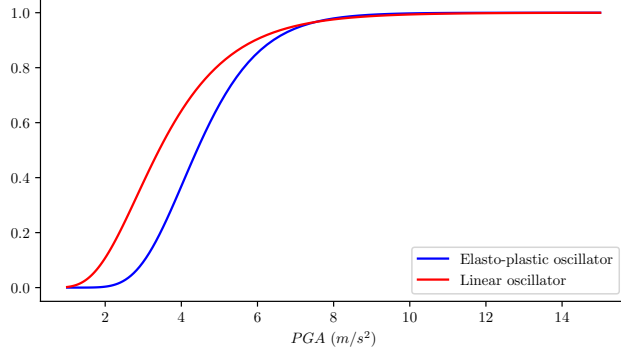


Figure 5: Lognormal fragility curves of the linear elastic and the nonlinear oscillators obtained by using least squares minimization on the total 10^5 synthetic seismic signals of the dataset.

Table 2: Defensive parameter ε influence on ν_{120}^{RS} when IM = PGA.

ε	Train	Test
10^{-1}	1.3	1.2
10^{-2}	2.1	3.9
10^{-3}	2.2	3.3

there are too many elements drawn from the pdf p . Accordingly, all the results will presented hereafter with a defensive parameter $\varepsilon = 10^{-3}$. This implies
 390 that the defensive strategy plays essentially no role here, but gives theoretical convergence guarantees.

6.2.4. Performance metrics

Figure 6 compares the IS-AL, MLE and RS training and testing errors as functions of n . The mean training loss for the 500 replications is higher for IS-AL
 395 than for RS. Indeed, the instrumental density is chosen to sample seismic signals that maximize the loss variance, resulting in a high training error. Moreover, the mean testing error of IS-AL is also significantly smaller than for RS and quickly converges to the "minimal" error related to the term $\mathbb{E}[\mu(X)(1 - \mu(X))]$ in (5).

This is shown in Table 3 by a significantly smaller value of relative bias RB_{120}^{IA} (1%) for the testing error than with RS (12%). With respect to the MLE, one cannot make equivalent remarks insofar as the two errors are "artificial" (see definitions in section 5.2.1) and only plotted for illustration purpose. However, Table 3 shows that the IS-AL strategy has overall better performance than the other two strategies.

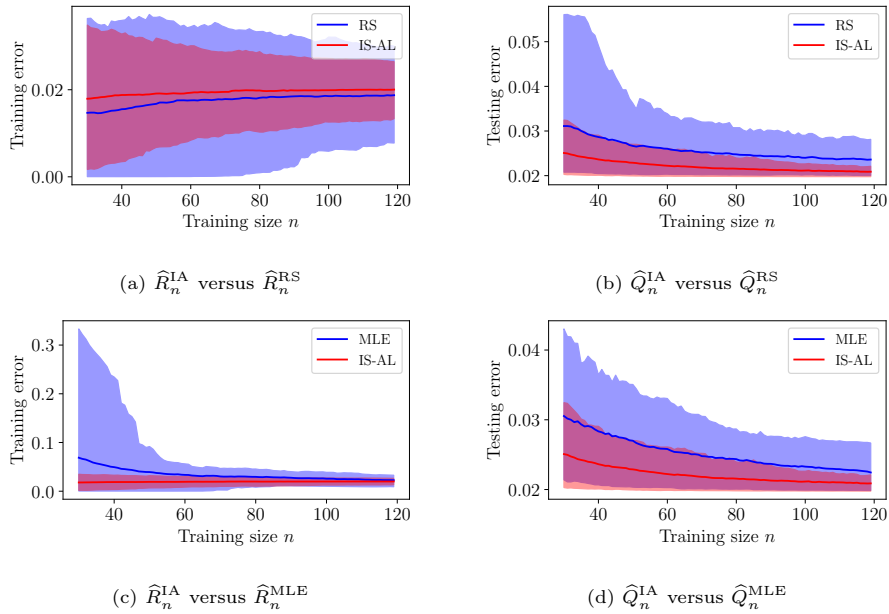


Figure 6: Results of the numerical benchmark for the elasto-plastic mechanical oscillator with the PGA as IM (same notations as for the synthetic test case). The empirical distributions of the training and testing errors are represented by the range between the empirical 90% and 10% quantiles of the 500 replications.

6.2.5. Empirical distributions of the parameters α and β

Figure 7 shows the empirical distributions of the parameters α and β for several sample sizes using 500 replications of MLE and IS-AL estimators. Remark in this case that IS-AL performs better than the MLE by reducing the variances of the parameters' estimators. The effects are particularly visible for the parameter β , when the active learning strategy and the regularization play their

Table 3: Performance metrics for the elasto-plastic oscillator for $n = 120$ when IM = PGA.

•	Train			Test		
	RS	MLE	IS-AL	RS	MLE	IS-AL
RSD_{120} (%)	47	65	30	19	14	11
ν_{120}	2.2	5.9	×	3.3	1.8	×
RB_{120} (%)	11.1	7.6	5	12	6.4	1

role in reducing the standard deviations of the estimators without increasing bias. Indeed, MLE performances are downgraded when the model is not well specified [57]. We remark that parameter estimation is quite unstable for IS-AL for low sample sizes. Indeed, the number of failure events for low sample sizes is often 0, which makes impossible a correct estimation of the fragility curve's parameters.

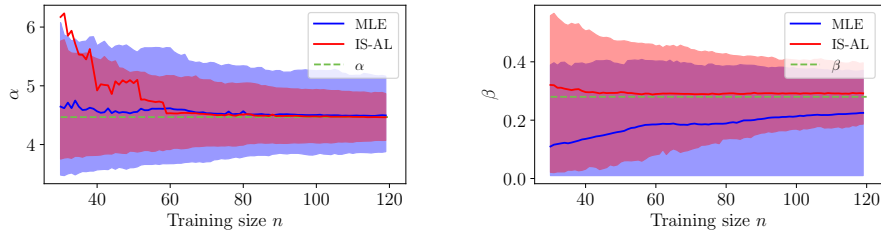


Figure 7: Results for the elasto-plastic mechanical oscillator with the PGA as IM: the empirical distributions of the parameters α and β are represented by the empirical 90% and 10% quantiles of the 500 replications and correspond to the shaded blue and red areas for MLE and IS-AL, respectively. The dashed green lines correspond to the values α and β , which have been here approximated by $\hat{\alpha}_N$, $\hat{\beta}_N$ for $N = 10^5$.

Figure 8 helps to visualize how IS-AL reduces the uncertainty of the fragility curve estimation: IS-AL is designed to sample seismic ground motions in the transition zone between 0 and 1 of the fragility curve, this phenomenon is responsible for the uncertainty reduction.

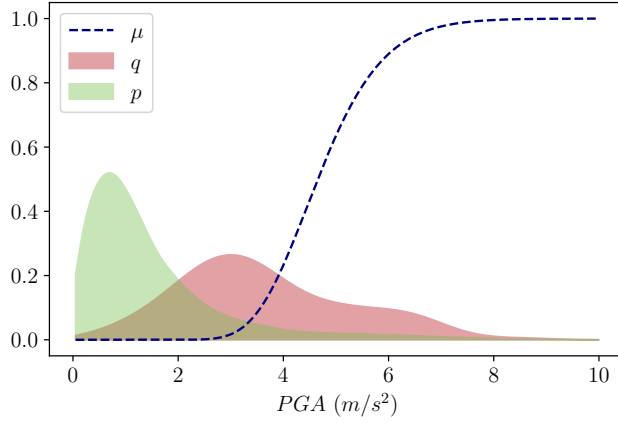


Figure 8: Comparison of the original marginal density p of PGA with the estimated density q of sampled PGA using IS-AL with $n = 100$ datapoints for the nonlinear oscillator. Fragility curve is approximated by $\mu(PGA) = \Phi\left(\frac{\log(PGA/\hat{\alpha}_N)}{\hat{\beta}_N}\right)$ for $N = 10^5$.

6.2.6. Convergence criterion

Figure 9 shows the value of the test statistics \widehat{W}_n (see section 4.2) for two independent IS-AL realizations. This result expresses that the IS-AL algorithm achieves asymptotic normality from $n = 100$ because the value of \widehat{W}_n is less than the quantile 90% of the distribution $\chi^2(2)$.

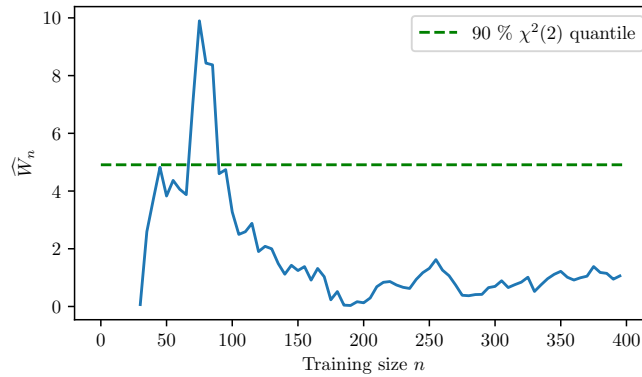


Figure 9: Values of the test statistic \widehat{W}_n for two independent IS-AL realizations, when IM = PGA.

6.2.7. CPs and CEVs

Figure 10 shows the CP values for the nonlinear oscillator for a training size n between 100 and 500 for the fragility curve estimation by MLE or IS-AL. The true parameter θ for this case has been approximated by $\hat{\theta}_N$ for $N = 10^5$.
 430 The numerical results show that the CP values are close to the theoretical and nominal value of 90%, which validates the theoretical results of the section 4.

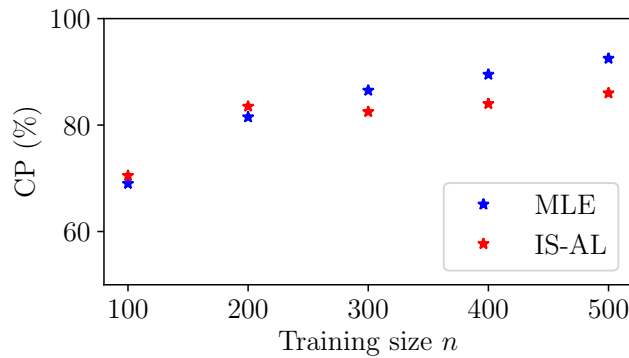


Figure 10: CP values (for the confidence ellipsoid with level $1 - \xi = 0.9$ for θ) as a function of the training size n of IS-AL and MLE when IM = PGA. $R = 100$ IS-AL replications are used to estimate the CP for each training size n . $B = 200$ bootstrap samples are generated for the MLE to build the confidence ellipsoid at level $1 - \xi = 0.9$ for n between 200 and 500, $B = 300$ bootstrap samples are generated for $n = 100$ due to numerical instabilities.

Figure 11 shows the CEVs for the MLE and IS-AL estimators. For $R = 200$ replications, these results show that for all the values of n considered $CEV_n^{IA} < CEV_n^{MLE}$. This indicates that MLE and IS-AL succeed in generating confidence
 435 ellipsoids that have the required coverage probability but MLE does so by generating ellipsoids that are much larger than the ones generated by IS-AL. We can then conclude that IS-AL is much more efficient.

We emphasize that the convergence criterion \widehat{W}_n , illustrated in Figure 9, gives us at which sample size the IS-AL reaches asymptotic normality and thus
 440 at which sample size asymptotic confidence ellipsoid can be used. Even though CP_{100}^{IA} is less than the theoretical 90%, 70% is considered as acceptable in prac-

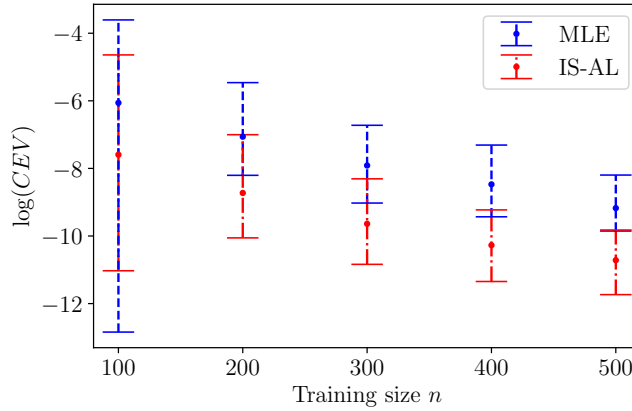


Figure 11: CEVs (for the confidence ellipsoid with level $1 - \xi = 0.9$ for θ) as a function of the training size n for IS-AL and MLE strategies when IM = PGA. The points are the medians over the $R = 200$ replications while the vertical lines are the ranges between the 10% and 90% quantiles.

tice.

6.2.8. Empirical distributions of the fragility curves and influence of the IM value

445 The choice of the seismic IM is crucial for the accuracy of fragility curves estimates, especially when parametric models are concerned. So, empirical distributions of the fragility curves for IS-AL and RS methods are shown in Figure 12 when IM = PGA, and in Figure 13 when IM is the spectral acceleration (SA) at 5 Hz and 2% damping ratio. The parametric fragility curves estimated with a
 450 dataset of 10^4 seismic ground motions, are also shown in order to validate both the model choice and the uncertainty reduction provided by IS-AL.

With the PGA, a bias between the lognormal fragility curve and the k-means nonparametric fragility curve, called μ_{MC} , can be seen in Figure 12. This phenomenon could be explained by the small correlation between maximal
 455 displacement of the oscillator during the seismic excitation and the PGA, which conveys small information about the seismic ground motion [40].

The results presented in Figure 13 show a reduction of the bias between the

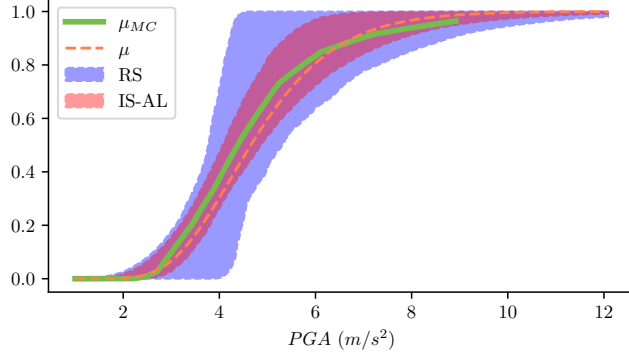


Figure 12: Empirical distribution of the fragility curves estimated by RS and IS-AL for the nonlinear oscillator. The dashed orange line and solid green line are respectively the parametric estimation μ using 10^4 seismic ground motions and the k-means nonparametric estimation of the fragility curve using 10^5 seismic ground motions μ_{MC} . The red and blue shaded areas correspond respectively to the 90% to 10% quantile ranges for the fragility curve dataset computed with IS-AL or RS. Remark that the nonparametric fragility curve is only plotted for $PGA < 10m/s^2$ due to the lack of seismic signals with PGA above that threshold.

nonparametric and the parametric fragility curve. This illustrates that, for the class of structures and for the seismic signal generator considered in this study, the parametric lognormal model has a better fit with the reference SA-based
460 fragility curve than with reference PGA-based fragility curve.

6.2.9. Confidence interval for parametric fragility curves : towards the engineering practice

After assessing the validity of the asymptotic confidence ellipsoid for IS-
465 AL thanks to the computation of the CP values in section 6.2.7, we can use the asymptotic Gaussian distribution to construct the CI of the parametric fragility curve, as in the engineering practice. Thus, using a single run of the IS-AL procedure, we estimate the asymptotic covariance matrix \hat{G}_n and sample fragility curve parameters from the asymptotic distribution $\mathcal{N}\left(\hat{\theta}_n^{IA}, \frac{\hat{G}_n}{n}\right)$. For
470 the sake of comparison, we also construct the CI on a single replication of the MLE procedure using the bootstrap technique.

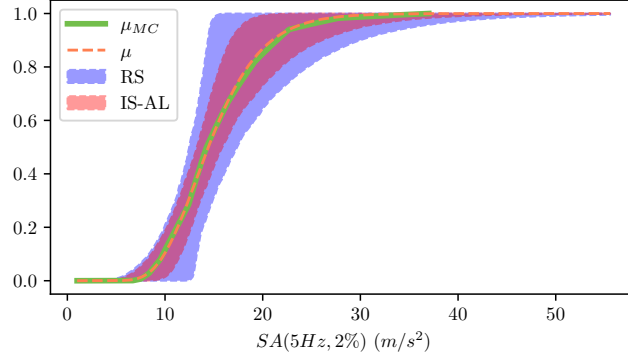


Figure 13: Empirical distributions of the fragility curves of RS and IS-AL for the nonlinear oscillator when IM is the spectral acceleration at 5 Hz and 2% damping ratio. The notations are the same as those in the Figure 12. Remark that the bias between the nonparametric fragility curve μ_{MC} and the parametric fragility curve μ is smaller than the one obtained when using the PGA as the intensity measure (compare with Figure 12).

Figure 14 represents the CIs for the fragility curve at level 90% for IS-AL and MLE strategies, for a single replication of size $n = 200$ of each procedure. Remark that the fragility curves estimated by MLE can be degenerated (i.e. as a unit step function), which implies that the CI for MLE is too conservative. This is consistent with the results of the figures 7 and 11.

6.2.10. Synthesis

In this section, we have shown that the IS-AL-based methodology is (i) efficient to reduce the variance of the fragility curve estimation and (ii) can be applied regardless of the IM of interest. However, in practice, it is more suitable to use an IM as correlated as possible to the response of the structure to minimize potential biases due to the use of a parametric model. In addition, we have shown that, if the computation times allow it, it is possible to know when to stop the IS-AL algorithm, in order to build asymptotic confidence ellipsoids.

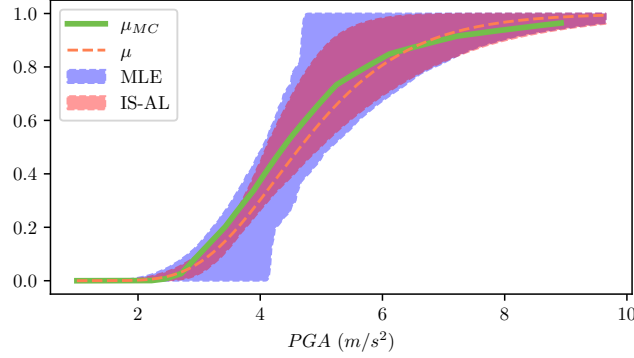


Figure 14: CIs of the parametric fragility curves of the elasto-plastic oscillator obtained with 500 samples of the parameter asymptotic Gaussian distribution for IS-AL and 500 bootstrapped estimators with MLE. The red and blue shaded areas correspond to the ranges between the 95% and 5% quantiles for respectively IS-AL and MLE. The solid green line corresponds to the Monte Carlo estimation (k-means nonparametric estimation) of the fragility curve based on a dataset of 10^5 of synthetic seismic signals.

6.3. Industrial test case: safety water supply pipe of a pressurized water nuclear reactor

6.3.1. Description of the piping system

The following test case corresponds to a piping system which is a simplified
 490 part of a secondary line of a French Pressurized Water Reactor. The numerical
 model was validated based on seismic tests performed on the shaking table
 Azalee of the EMSI laboratory of CEA/Saclay. The experimental program,
 called ASG program, and the main results are outlined in Ref. [60]. In Figure 15a
 a view of the mock-up mounted on the shaking table is shown. The Finite
 495 Element (FE) model, based on beam elements, is depicted in Figure 15b.

The mock-up is a 114.3 mm outside diameter and 8.56 mm thickness pipe
 with a 0.47 elbow characteristic parameter, in carbon steel TU42C, filled with
 water without pressure. It contains three elbows and a mass modeling a valve
 (120 kg) which corresponds to more than 30% of the specimen total mass. As
 500 shown in Figure 15b, one end of the mock-up is clamped whereas the other is

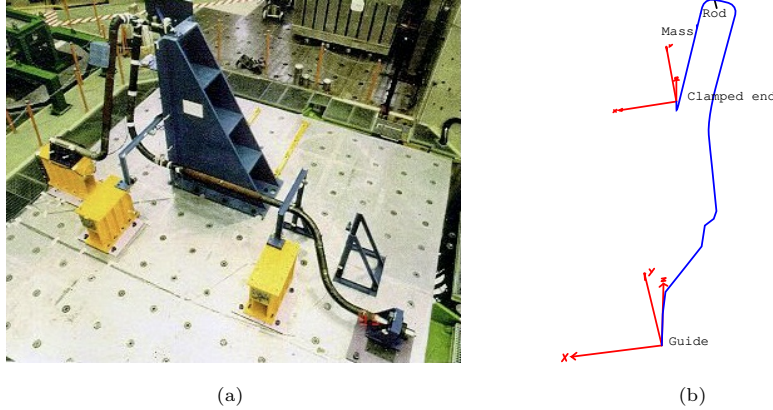


Figure 15: (a) Overview of the ASG mock-up on the CEA's shaking table and (b) ASG FE model.

supported by a guide in order to prevent the displacements in the X and Y directions. Additionally, a rod is placed on the top of the specimen in order to limit the mass displacements in the Z direction. In the tests, excitation act in the X direction.

505 Numerical comparisons are carried out with the homemade FE code CAST3M [61]. Concerning the FE model, the boundary conditions are adjusted in order to obtain shapes and frequencies similar of those of the first two eigenmodes of the mock-up in the X and Y directions, respectively at 5.1 Hz and 6.6 Hz. As measured in the experiments, a critical damping ratio of 1% is considered
 510 for these two eigenmodes with a damping Rayleigh assumption. Finally, regarding the nonlinear constitutive law of the material, a bilinear law exhibiting kinematic hardening is used to reproduce the overall nonlinear behavior of the mock-up with satisfactory agreement compared to the results of seismic tests [60].

515 In the context of this test case, the yield stress of the bilinear law is equal to $3 \cdot 10^8$ Pa, the Young modulus is equal to $1.92 \cdot 10^{11}$ Pa whereas the hardening modulus is equal to $4.3 \cdot 10^8$ Pa. Moreover, since for the synthetic signals considered in this work (the same as those used in the reference [25] and in the

second test case of this paper) the piping system remains in the linear domain,
 520 they are filtered by a fictitious linear single-mode building at 5 Hz and damped
 at 2%. Finally, we consider excessive out-of-plane rotation of the elbow located
 near the clamped end of the mock-up as failure criterion, as recommended in
 [62]. Since the weight of the mass is not completely taken up by the mechanical
 assembly, the overall behavior of the mock-up exhibits ratcheting.

525 In the following, the random variable R_e corresponds to the maximum of
 the out-of-plane rotation of the elbow. The binary variable which indicates the
 failure state is defined by $S = \mathbf{1}_{R_e > C}$ where C is the admissible rotation in
 degree. In our case, $C = 4.38$. This value is the 90%-level quantile from a
 sample of 2000 mechanical simulations.

530 6.3.2. Performance metrics

For this test case, the numerical benchmark is based on 50 replications of
 120 signals sampled using IS-AL (that includes the initial 20 points) with a
 defensive parameter $\varepsilon = 10^{-3}$ and 120 signals for the RS and MLE strategies.

The IS-AL procedure is initialized by considering the linear FE model of the
 535 ASG mock-up and a numerical resolution based on a modal base projection.
 Thus, the initialization parameter $\hat{\theta}_0^{IA}$ is approximated by $\hat{\theta}_{2000}^{RS}$ using a 2000-
 sized dataset randomly selected from the 10^5 synthetic seismic ground motions.
 Then, 20 datapoints are queried with the instrumental density $q_{\hat{\theta}_0^{IA}, \varepsilon}$ before
 launching the adaptive strategy. For the training, 100 signals are then chosen in
 540 a pool of 1500 CAST3M computations while 500 additional computations are
 carried out in order to compute the testing error.

Figure 16 compares the IS-AL, MLE and RS training and testing errors as
 functions of n . Remark that the training loss of MLE is greater than the training
 loss of IS-AL. This numerical artifact is essentially due to the regularization
 545 term β_{reg} : if the β parameter estimated by MLE is small, the penalization term
 β_{reg}/β can be very high.

Table 4 shows that the IS-AL strategy has overall better performance than
 the other two strategies.

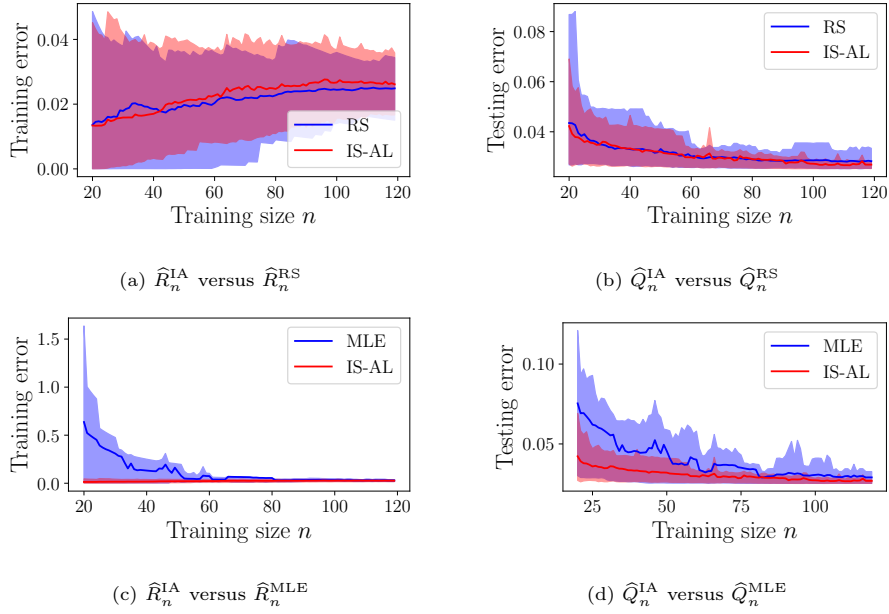


Figure 16: Numerical benchmark of the ASG piping system. The empirical distributions of the training and testing errors are computed, the red and blue shaded areas correspond to the area between the empirical quantiles of levels 10% and 90% of the 50 replications for respectively IS-AL, RS and MLE.

Table 4: Performance metrics for the ASG piping system for $n = 120$ when IM = SA

•	Train			Test		
	RS	MLE	IS-AL	RS	MLE	IS-AL
RSD ₁₂₀ (%)	40.5	46	34	24.1	28	12
ν_{120}	0.93	2.4	×	1.3	5.8	×
RB ₁₂₀ (%)	7.2	8.6	5.5	18	8.4	0.3

6.3.3. Empirical distributions of the parameters α and β

550 Figure 17 compares the distributions of parameters α and β for several sample sizes between MLE and IS-AL using 50 replications. As with the nonlinear oscillator, the β parameter estimated with IS-AL is less likely to be close to 0

than when it is estimated with MLE. This motivates further the use of active
learning to have a better accuracy for fragility curves parameters estimates with
555 the same computational cost as state of the art estimation methods.

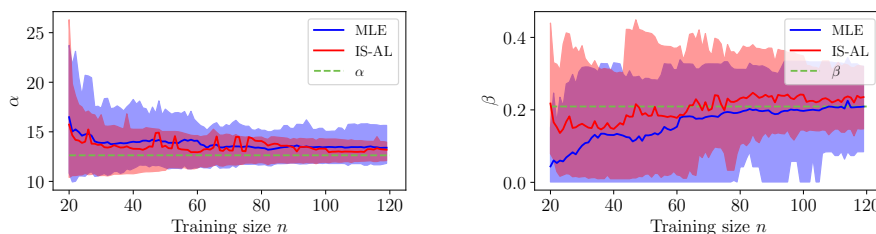


Figure 17: Results for the ASG piping system: the empirical distributions of parameters α and β are represented by the empirical 90% and 10% quantiles of 50 replications and correspond to the shaded blue and red areas respectively for MLE and IS-AL. The dashed green lines correspond to the values α and β , which have been here approximated by $\hat{\alpha}_N$, $\hat{\beta}_N$ for $N = 2000$.

6.3.4. Fragility curve estimations

Figure 18 illustrates the uncertainty reduction provided by IS-AL on the
fragility curve estimate. Motivated by the results obtained for the nonlinear
oscillator, the fragility curve of the piping system is here expressed as a function
560 of the pseudo-spectral acceleration of the initial set of synthetic signals (i.e not
filtered signals), calculated at 5 Hz and 1% damping ratio.

Figure 19 represents the confidence interval on the fragility curve for IS-AL
and MLE using a single replication of 200 CAST3M computations (20 computa-
tions used for initialization and 180 computations for IS-AL), obtained with the
565 same methodology as that presented in section 6.2.9. Remark that the Monte
Carlo estimation of the fragility curve (i) belongs to the confidence interval of IS-
AL for seisms with relatively small spectral acceleration (ii) is not accurate for
high spectral accelerations due to the lack of seismic signals of such intensities.

As for the nonlinear oscillator, the figures 18 and 19 suggest that for the RS
570 and MLE strategies, even with $n = 120$ points, it is possible to obtain samples
for which a β estimate is close to 0, which IS-AL avoids.

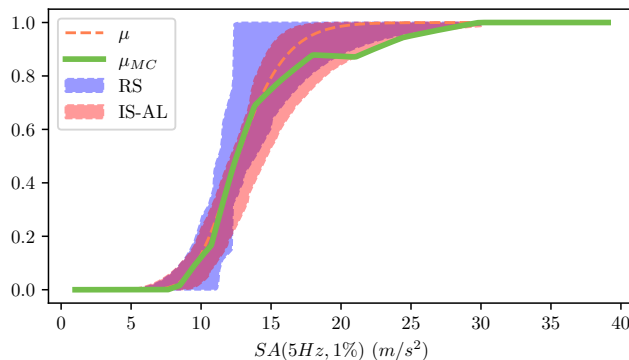


Figure 18: Empirical distributions of the fragility curves of RS and IS-AL for the ASG piping system. The red and blue areas correspond respectively to the ranges between the 10% and 90% quantiles of the fragility curve dataset generated with IS-AL and RS with $n = 120$ training datapoints (that includes the initialization points). The dashed orange line corresponds to a parametric fragility curve estimation using least squares minimization with a dataset of 2000 seismic ground motions and FE simulations of the piping system. The solid green line corresponds to a Monte Carlo estimation (k-means nonparametric estimation) of the fragility curve using the same 2000-sized dataset.

7. Conclusion

In this paper, we have introduced an original methodology to improve the accuracy of parametric fragility curve estimation without increasing the sample size, thanks to an active learning strategy based on importance sampling. Defensive strategy has been implemented to control the likelihood ratio and the possible increase of the training loss variance in the early steps. We use a penalized least square loss to avoid an identifiability issue of the standard deviation of the lognormal model. We define a convergence criterion that indicates asymptotic normality of the estimator and provide asymptotic confidence intervals and ellipsoids. We illustrate our active learning procedure in numerical examples, from a synthetic case to a FE mechanical simulation of a piping system of a French Pressurized Water Reactor. In comparison with the engineering practice based on the joint use of the MLE and the bootstrap techniques, the proposed methodology is more efficient. For the same number of calculations, the IS-

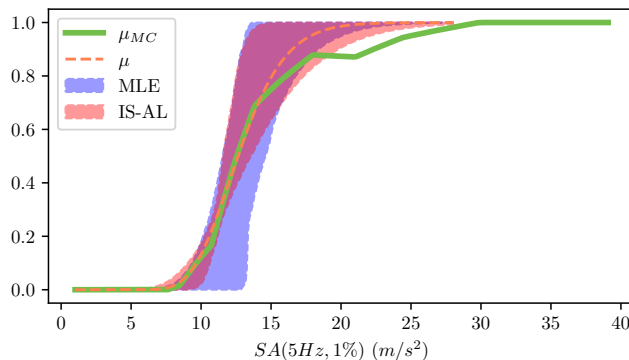


Figure 19: Parametric fragility curve of the ASG piping system confidence interval obtained with 500 samples of the parameter asymptotic Gaussian distribution for IS-AL and 500 bootstrapped estimators with MLE. The red and blue shaded areas correspond to the ranges between the 95% and 5% quantiles for respectively IS-AL and MLE. The solid green line corresponds to a Monte Carlo estimation (k-means nonparametric estimation) of the fragility curve using our 2000 sized dataset of CAST3M computations.

AL procedure reduces the variance of the parametric estimation of the fragility curve and gives theoretical guarantees on the convergence of the estimations.

Acknowledgements

This research was supported by CEA (French Alternative Energies and Atomic Energy Commission) and SEISM Institute (www.institut-seism.fr/en/).
590

Appendix A. Proof of Equation (22)

Throughout the appendix, the IS-AL estimator $\hat{\theta}_n^{\text{IA}}$ is denoted by $\tilde{\theta}_n$ and the loss \hat{R}^{IA} by \tilde{R} .

The proof for the consistency is based on Theorem 2 of [54]. We precise the needed assumptions in a very general way, with a parametric family $\mathcal{F} = \{f_\theta, \theta \in \Theta\}$, loss function ℓ_θ and instrumental density q_θ . We will then check that the needed assumptions are satisfied by IS-AL. Set $L(x, s) =$

$\sup_{\theta \in \Theta} \ell_\theta(x, s)$. Assume that Θ is a compact set, $\theta^* = \operatorname{argmin}_{\theta \in \Theta} r(\theta)$ exists and is unique and that:

$$\begin{aligned} \iint L(x, s)P(dx, ds) &< +\infty, \\ \sup_{\theta \in \Theta} \iint \frac{L(x, s)^2 p(x)}{q_\theta(x)} P(dx, ds) &< +\infty, \\ \forall \theta \neq \theta^*, \iint \ell_\theta(x, s)P(dx, ds) &> \iint \ell_{\theta^*}(x, s)P(dx, ds). \end{aligned}$$

and for any $(x, s) \in \mathcal{X} \times \{0, 1\}$, $\theta \in \Theta \mapsto \ell_\theta(x, s)$ is continuous. Thus, we can apply Theorem 2 of [54] in order to prove the consistency of $\tilde{\theta}_n = \operatorname{argmin}_{\theta \in \Theta} \frac{1}{n} \sum_{i=1}^n \frac{p(X_i)}{q_{\tilde{\theta}_{i-1}}(X_i)} \ell_\theta(X_i, S_i)$. More precisely, these assumptions are verified for IS-AL. Indeed, the regularized squared loss is bounded for the variables θ, x, s when $\theta = (\alpha, \beta)^T$ is in a compact set of $(0, +\infty)^2$. Moreover, the likelihood ratio $\frac{p(x)}{q_{\theta, \varepsilon}(x)}$ with the defensive instrumental density is bounded for $x \in \mathcal{X}$. Concerning the regularization, we have:

$$\left| \tilde{R}_n(\theta) - r(\theta) \right| < \left| \frac{1}{n} \sum_{i=1}^n \frac{p(X_i)}{q_{\tilde{\theta}_{i-1}^{\text{IA}}, \varepsilon}(X_i)} \ell_\theta(X_i, S_i) - r(\theta) \right| + \frac{\Omega(\theta; \beta_{\text{reg}})}{n}. \quad (\text{A.1})$$

595 Thus the condition (17) of Theorem 2 in [54] is still valid.

Appendix B. Proof of the asymptotic normality of $\hat{\theta}_n^{\text{IA}}$

In the same way as in the proof of the consistency of $\tilde{\theta}_n$, we provide a general proof of asymptotic normality. Assume that $\theta \mapsto \ell_\theta$ is three times differentiable at θ^* for all x, s and that the matrix $\ddot{r}(\theta^*)$ exists and is nonsingular. Assume
600 that the third-order derivatives of $\theta \mapsto \ell_\theta(x, s)$ are dominated in a neighborhood of θ^* by a function that is integrable with respect to P . Assume also that the following conditions are satisfied:

1. The hypotheses needed for the consistency of $\tilde{\theta}_n$ are satisfied,
2. $\exists \eta > 0$ such that $\sup_{\theta \in \Theta} \iint \left\| \frac{p(x) \dot{\ell}_{\theta^*}(x, s)}{q_\theta(x)} \right\|^{2+\eta} P(dx, ds) < +\infty$,
- 605 3. $\sup_{\theta \in \Theta} \iint \frac{p(x) \ddot{\ell}_{\theta^*}(x, s) \ddot{\ell}_{\theta^*}(x, s)^T}{q_\theta(x)} P(dx, ds) < +\infty$,
4. there exists a neighborhood \mathcal{B} of θ^* such that $\forall (x, s) \in \mathcal{X} \times \{0, 1\}$,
 $\sup_{\theta \in \mathcal{B}} \frac{p(x) \dot{\ell}_{\theta^*}(x, s) \dot{\ell}_{\theta^*}(x, s)^T}{q_\theta(x)} < +\infty$.

The asymptotic normality of an estimator built such as $\tilde{\theta}_n$ is based on the following arguments highlighted in Theorem 5.41 of [63]:

- 610 • **(P1)** The random function $\sqrt{n}\Psi_n(\theta)$, with $\Psi_n(\theta) = \dot{R}_n(\theta) - \dot{r}(\theta)$, converges in law to a centered Gaussian distribution with covariance V_{θ_*} .
- **(P2)** The random function $\dot{\Psi}_n(\theta)$ converges in probability to $E[\dot{\ell}_{\theta_*}(X, S)]$
- **(P3)** The random function $\ddot{\Psi}_n(\theta_n)$ is bounded in probability for θ_n a deterministic sequence in a neighborhood of θ .

615 Of course, we need all the quantities above to be properly defined, hence we have to restrict ourselves to a loss function $\theta \mapsto \ell_\theta$ that is smooth enough, such as the quadratic loss. We use Theorem 1 of [54] to prove proposition **(P1)**. Theorem 2.18 in [64] ensures that **(P2)** and **(P3)** are verified by the assumptions 2), 3) and 4) so that $\dot{\Psi}_n(\theta)$ converges toward the matrix $\ddot{r}(\theta)$. The sequence
620 $\sqrt{n}(\tilde{\theta}_n - \theta)$ is asymptotically normal with mean zero and covariance matrix $\ddot{r}(\theta)^{-1}V_{\theta_*}(\ddot{r}(\theta)^{-1})^T$. For IS-AL, the functions $\dot{\ell}_\theta, \ddot{\ell}_\theta, \ddot{\ddot{\ell}}_\theta$ are continuous for variables θ, x on a compact set and thus are bounded for variable θ , in the same way as for the consistency, the likelihood ratio for the defensive instrumental density is bounded for $x \in \mathcal{X}$. Concerning the regularization, the third deriva-
625 tive $\ddot{\ddot{\Omega}}(\theta; \beta_{\text{reg}})$ is continuous on Θ which is compact, hence bounded. Naturally, we have **(P3)** verified. Because $\frac{\dot{\Omega}(\theta; \beta_{\text{reg}})}{n}$ converges in probability to 0, **(P2)** is also verified. Using Slutsky's lemma, **(P1)** is verified.

Appendix C. Proof of Lemma Equation (29)

First of all, we precise the needed assumptions for a general proof. Set
630 $L_{1,k,l}(x, s) = \sup_{\theta \in \Theta} \ddot{\ell}_\theta(x, s)_{k,l}$ and $L_{2,k,l}(x, s) = \sup_{\theta \in \Theta} \frac{p(x)}{q_\theta(x)} (\dot{\ell}_\theta(x, s) \dot{\ell}_\theta(x, s)^T)_{k,l}$
 $\forall k, l = 1, \dots, m$ and assume that:

1. $\inf_{(\theta, x, s) \in \Theta \times \mathcal{X} \times \mathcal{S}} \frac{p(x)}{q_\theta(x)} \ddot{\ell}_\theta(x, s)_{k,l} > -\infty \forall k, l = 1, \dots, m.$
2. $\inf_{(\theta, x, s) \in \Theta \times \mathcal{X} \times \mathcal{S}} \left(\frac{p(x)}{q_\theta(x)} \right)^2 \dot{\ell}_\theta(x, s) \dot{\ell}_\theta(x, s)^T_{k,l} > -\infty \forall k, l = 1, \dots, m.$
3. $\iint L_{i,k,l}(x, s) P(dx, ds) < +\infty, \forall i \in \{1, 2\}, \forall k, l = 1, \dots, m.$

$$635 \quad 4. \sup_{\theta \in \Theta} \iint \frac{L_i(x,s)^2 p(x)}{q_\theta(x)} P(dx, ds) < +\infty, \quad i \in \{1, 2\}.$$

The result comes from the uniform convergence of $\widehat{G}_n(\theta) = \widehat{r}_n(\theta)^{-1} \widehat{V}_n(\theta) (\widehat{r}_n(\theta)^{-1})^T$ to G_θ for θ in a neighborhood of θ . It boils down to prove uniform convergence of $\widehat{r}_n(\theta)$ and $\widehat{V}_n(\theta)$. The proof is in the same spirit as in Appendix A. We proceed coordinate by coordinate defining $H_i(\theta)_{k,l} = \frac{p(X_i)}{q_\theta(X_i)} \ddot{\ell}_\theta(X_i, S_i)_{k,l} -$
 $640 \quad \inf_{(\theta, x, s) \in \Theta \times \mathcal{F} \times \mathcal{G}} \frac{p(x)}{q_\theta(x)} \ddot{\ell}_\theta(x, s)_{k,l}$ to prove uniform convergence of $\widehat{r}_n(\theta)$ and $H_i(\theta)_{k,l} = \left(\frac{p(X_i)}{q_\theta(X_i)} \right)^2 \dot{\ell}_\theta(X_i, S_i) \dot{\ell}_\theta(X_i, S_i)_{k,l}^T - \inf_{(\theta, x, s) \in \Theta \times \mathcal{F} \times \mathcal{G}} \left(\frac{p(x)}{q_\theta(x)} \right)^2 \dot{\ell}_\theta(x, s) \dot{\ell}_\theta(x, s)_{k,l}^T$ for $\widehat{V}_n(\theta)$. Assumptions 3) and 4) ensure the uniform convergence using the proof technique of Theorem 1 of [54].

Appendix D. Proof of Equation (31)

The proof relies on the Taylor expansions of $\dot{R}_{n,1}(\tilde{\theta}_{n,2})$ and $\dot{R}_{n,2}(\tilde{\theta}_{n,1})$ around the parameter value θ :

$$\begin{aligned} \dot{R}_{n,1}(\tilde{\theta}_{n,2}) &= \dot{R}_{n,1}(\theta) + \ddot{R}_{n,1}(\theta)(\tilde{\theta}_{n,2} - \theta) + o\left(\|\tilde{\theta}_{n,2} - \theta\|\right). \\ \dot{R}_{n,2}(\tilde{\theta}_{n,1}) &= \dot{R}_{n,2}(\theta) + \ddot{R}_{n,2}(\theta)(\tilde{\theta}_{n,1} - \theta) + o\left(\|\tilde{\theta}_{n,1} - \theta\|\right). \end{aligned}$$

Using the asymptotic normality of $\tilde{\theta}_n$, we can apply Appendix B.3 in [54] to prove the convergence of $\ddot{R}_{n,1}(\theta)$ and $\ddot{R}_{n,2}(\theta)$ to $\ddot{r}(\theta)$ in the same spirit as for the proof of Appendix A. We proceed coordinate by coordinate, defining $H_i(\theta)_{k,l} = \frac{p(X_i)}{q_{\tilde{\theta}_{i-1}(X_i)}} \ddot{\ell}_{\theta_*}(X_i, S_i)_{k,l} - \inf_{(\theta, x, s) \in \Theta \times \mathcal{F} \times \mathcal{G}} \frac{p(x)}{q_\theta(x)} \ddot{\ell}_{\theta_*}(x, s)_{k,l}$. Remark that $H_i(\theta)_{k,l} \geq 0$, hence we can apply Appendix B.3 in [54] to obtain the desired convergence. Moreover, the Taylor expansions of $\dot{R}_{n,1}(\tilde{\theta}_{n,1})$ and $\dot{R}_{n,2}(\tilde{\theta}_{n,2})$ write:

$$\begin{aligned} 0 &= \dot{R}_{n,1}(\tilde{\theta}_{n,1}) = \dot{R}_{n,1}(\theta) + \ddot{r}(\theta)(\tilde{\theta}_{n,1} - \theta) + o\left(\|\tilde{\theta}_{n,1} - \theta\|\right), \\ 0 &= \dot{R}_{n,2}(\tilde{\theta}_{n,2}) = \dot{R}_{n,2}(\theta) + \ddot{r}(\theta)(\tilde{\theta}_{n,2} - \theta) + o\left(\|\tilde{\theta}_{n,2} - \theta\|\right). \end{aligned}$$

Finally, the Taylor expansion of $\sqrt{n}(\dot{R}_{n,1}(\tilde{\theta}_{n,2}) - \dot{R}_{n,2}(\tilde{\theta}_{n,1}))$ writes:

$$\begin{aligned} & \sqrt{n}(\dot{R}_{n,1}(\tilde{\theta}_{n,2}) - \dot{R}_{n,2}(\tilde{\theta}_{n,1})) \\ &= \sqrt{n}(\dot{R}_{n,1}(\theta) - \dot{R}_{n,2}(\theta) + \ddot{r}(\theta)(\tilde{\theta}_{n,2} - \theta) - \ddot{r}(\theta)(\tilde{\theta}_{n,1} - \theta)) + o_P(1) \\ &= 2\sqrt{n}(\ddot{R}_{n,1}(\theta) - \ddot{R}_{n,2}(\theta)) + o_P(1), \end{aligned} \tag{D.1}$$

645 because $\sqrt{n}\|\tilde{\theta}_{n,j} - \theta\| = O_P(1)$ for $j = 1, 2$. The right-hand side of equation (D.1) weakly converges towards the centered Gaussian distribution with covariance matrix $8V(q_{\theta_*}, \dot{\ell}_{\theta_*})$.

References

- [1] R. Kennedy, C. Cornell, R. Campbell, S. Kaplan, H. Perla, Probabilistic
650 seismic safety study of an existing nuclear power plant, *Nuclear Engineering and Design* 59 (2) (1980) 315 – 338.
- [2] R. Kennedy, M. Ravindra, Seismic fragilities for nuclear power plant risk studies, *Nuclear Engineering and Design* 79 (1) (1984) 47–68.
- [3] Y. Park, C. Hofmayer, N. Chokshi, Survey of seismic fragilities used in pra
655 studies of nuclear power plants, *Reliability Engineering & System Safety* 62 (3) (1998) 185–195. doi : [https://doi.org/10.1016/S0951-8320\(98\)00019-2](https://doi.org/10.1016/S0951-8320(98)00019-2).
- [4] A. Ghobarah, Performance-based design in earthquake engineering: state of development, *Engineering Structures* 23 (8) (2001) 878–884.
- 660 [5] K. Porter, R. Kennedy, R. Bachman, Creating fragility functions for performance-based earthquake engineering, *Earthquake Spectra* 23 (2) (2007) 471–489. doi : 10.1193/1.2720892.
- [6] A. Quilligan, A. O’Connor, V. Pakrashi, Fragility analysis of steel and concrete wind turbine towers, *Engineering Structures* 36 (2012) 270—282.

- 665 [7] C. Bernier, J. E. Padgett, Fragility and risk assessment of aboveground storage tanks subjected to concurrent surge, wave, and wind loads, *Reliability Engineering & System Safety* 191 (2019) 106571. doi:<https://doi.org/10.1016/j.res.2019.106571>.
- [8] I. Zentner, M. Gündel, N. Bonfils, Fragility analysis methods: Review of existing approaches and application, *Nuclear Engineering and Design* 323
670 (2017) 245–258.
- [9] P. Gardoni, A. D. Kiureghian, K. M. Mosalam, Probabilistic capacity models and fragility estimates for reinforced concrete columns based on experimental observations, *Journal of Engineering Mechanics* 128 (10) (2002)
675 1024–1038. doi:10.1061/(ASCE)0733-9399(2002)128:10(1024).
- [10] D. Straub, A. Der Kiureghian, Improved seismic fragility modeling from empirical data, *Structural Safety* 30 (4) (2008) 320–336. doi:<https://doi.org/10.1016/j.strusafe.2007.05.004>.
- [11] D. Lallemand, A. Kiremidjian, H. Burton, Statistical procedures for developing earthquake damage fragility curves, *Earthquake Engineering & Structural Dynamics* 44 (9) (2015) 1373–1389.
680
- [12] M. Shinozuka, M. Q. Feng, J. Lee, T. Naganuma, Statistical analysis of fragility curves, *Journal of Engineering Mechanics* 126 (12) (2000) 1224–1231.
- 685 [13] B. R. Ellingwood, Earthquake risk assessment of building structures, *Reliability Engineering & System Safety* 74 (3) (2001) 251–262.
- [14] S.-H. Kim, M. Shinozuka, Development of fragility curves of bridges retrofitted by column jacketing, *Probabilistic Engineering Mechanics* 19 (1) (2004) 105–112, fourth International Conference on Computational
690 Stochastic Mechanics. doi:<https://doi.org/10.1016/j.probengmech.2003.11.009>.

- [15] T. K. Mandal, S. Ghosh, N. N. Pujari, Seismic fragility analysis of a typical indian PHWR containment: Comparison of fragility models, *Structural Safety* 58 (2016) 11–19. doi : <https://doi.org/10.1016/j.strusafe.2015.08.003>.
695
- [16] F. Wang, C. Feau, Influence of Input Motion's Control Point Location in Nonlinear SSI Analysis of Equipment Seismic Fragilities: Case Study on the Kashiwazaki-Kariwa NPP, *Pure and Applied Geophysics* doi : <https://doi.org/10.1016/j.engstruct.2018.02.024>.
- 700 [17] C. Mai, K. Konakli, B. Sudret, Seismic fragility curves for structures using non-parametric representations, *Frontiers of Structural and Civil Engineering* 11 (2) (2017) 169–186.
- [18] K. Trevelopoulos, C. Feau, I. Zentner, Parametric models averaging for optimized non-parametric fragility curve estimation based on intensity measure data clustering, *Structural Safety* 81 (2019) 101865.
705
- [19] N. Luco, P. Bazzurro, Does amplitude scaling of ground motion records result in biased nonlinear structural drift responses?, *Earthquake Engineering & Structural Dynamics* 36 (13) (2007) 1813–1835. doi : <https://doi.org/10.1002/eqe.695>.
- 710 [20] D. Altieri, E. Patelli, An efficient approach for computing analytical non-parametric fragility curves, *Structural Safety* 85 (2020) 101956. doi : <https://doi.org/10.1016/j.strusafe.2020.101956>.
- [21] M. Grigoriu, To scale or not to scale seismic ground-acceleration records, *Journal of Engineering Mechanics* 137 (4) (2011) 284–293. doi : [10.1061/\(ASCE\)EM.1943-7889.0000226](https://doi.org/10.1061/(ASCE)EM.1943-7889.0000226).
715
- [22] Y.-G. Zhao, M.-J. Qin, Z.-H. Lu, L.-W. Zhang, Seismic fragility analysis of nuclear power plants considering structural parameter uncertainty, *Reliability Engineering & System Safety* 216 (2021) 107970. doi : <https://doi.org/10.1016/j.res.s.2021.107970>.

- 720 [23] J. W. Baker, Efficient analytical fragility function fitting using dynamic structural analysis, *Earthquake Spectra* 31 (1) (2015) 579–599. doi : 10.1193/021113EQS025M.
- [24] J. Kiani, C. Camp, S. Pezeshk, On the application of machine learning techniques to derive seismic fragility curves, *Computers & Structures* 218 (2019) 108–122. doi : <https://doi.org/10.1016/j.compstruc.2019.03.004>.
725
- [25] R. Saint, C. Feau, J.-M. Martinez, J. Garnier, Efficient methodology for seismic fragility curves estimation by active learning on support vector machines, *Structural Safety* 86 (2020) 101972.
- 730 [26] I. Gidaris, A. A. Taflanidis, G. P. Mavroeidis, Kriging metamodeling in seismic risk assessment based on stochastic ground motion models, *Earthquake Engineering & Structural Dynamics* 44 (14) (2015) 2377–2399.
- [27] C. Mai, M. Spiridonakos, E. Chatzi, B. Sudret, Surrogate modeling for stochastic dynamical systems by combining nonlinear autoregressive with exogenous input models and polynomial chaos expansions, *Int. J. Uncertainty Quant.* 6 (2016) 313–339.
735
- [28] C. C. Mitropoulou, M. Papadrakakis, Developing fragility curves based on neural network ida predictions, *Engineering Structures* 33 (12) (2011) 3409–3421. doi : <https://doi.org/10.1016/j.engstruct.2011.07.005>.
- 740 [29] Z. Wang, N. Pedroni, I. Zentner, E. Zio, Seismic fragility analysis with artificial neural networks: Application to nuclear power plant equipment, *Engineering Structures* 162 (2018) 213–225. doi : <https://doi.org/10.1016/j.engstruct.2018.02.024>.
- 745 [30] I. Zentner, Numerical computation of fragility curves for NPP equipment, *Nuclear Engineering and Design* 240 (6) (2010) 1614–1621. doi : <https://doi.org/10.1016/j.nucengdes.2010.02.030>.

- [31] P. Gehl, J. Douglas, D. M. Seyedi, Influence of the number of dynamic analyses on the accuracy of structural response estimates, *Earthquake Spectra* 31 (1) (2015) 97–113. doi : <https://doi.org/10.1193/102912EQS320M>.
- 750 [32] W. Chu, M. Zinkevich, L. Li, A. Thomas, B. Tseng, Unbiased online active learning in data streams, in: *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '11*, Association for Computing Machinery, New York, NY, USA, 2011, p. 195–203.
- 755 [33] T. Hastie, R. Tibshirani, J. Friedman, *The elements of statistical learning: data mining, inference and prediction*, Springer, 2009.
- [34] T. Kloek, T. Kloek, H. Van Dijk, Bayesian estimates of equation system parameters: An application of integration by monte carlo, *Econometrica* 46 (1978) 1–19.
- 760 [35] M.-S. Oh, J. O. Berger, Adaptive importance sampling in monte carlo integration, *Journal of Statistical Computation and Simulation* 41 (3-4) (1992) 143–168.
- [36] C. Gong, W. Zhou, Importance sampling-based system reliability analysis of corroding pipelines considering multiple failure modes, *Reliability Engineering & System Safety* 169 (2018) 199–208.
- 765 [37] I. Papaioannou, S. Geyer, D. Straub, Improved cross entropy-based importance sampling with a flexible mixture model, *Reliability Engineering & System Safety* 191 (2019) 106564.
- [38] I. Zentner, A general framework for the estimation of analytical fragility functions based on multivariate probability distributions, *Structural Safety* 770 64 (2017) 54–61. doi : <https://doi.org/10.1016/j.strusafe.2016.09.003>.

- [39] M. Hariri-Ardebili, V. Saouma, Probabilistic seismic demand model and optimal intensity measure for concrete dams, *Structural Safety* 59 (2016) 67–85. doi : <https://doi.org/10.1016/j.strusafe.2015.12.001>.
775
- [40] M. Ciano, M. Giofrè, M. Grigoriu, The role of intensity measures on the accuracy of seismic fragilities, *Probabilistic Engineering Mechanics* 60 (2020) 103041.
- [41] N. Luco, C. A. Cornell, Structure-specific scalar intensity measures for near-source and ordinary earthquake ground motions, *Earthquake Spectra* 23 (2) (2007) 357–392. doi : [10.1193/1.2723158](https://doi.org/10.1193/1.2723158).
780
- [42] J. E. Padgett, B. G. Nielson, R. DesRoches, Selection of optimal intensity measures in probabilistic seismic demand models of highway bridge portfolios, *Earthquake Engineering & Structural Dynamics* 37 (5) (2008) 711–725. arXiv : <https://onlinelibrary.wiley.com/doi/pdf/10.1002/eqe.782>, doi : [10.1002/eqe.782](https://doi.org/10.1002/eqe.782).
785
- [43] S. Rezaeian, A. Der Kiureghian, Simulation of synthetic ground motions for specified earthquake and site characteristics, *Earthquake Engineering & Structural Dynamics* 39 (10) (2010) 1155–1180.
- [44] N. Kwong, A. Chopra, Evaluation of the exact conditional spectrum and generalized conditional intensity measure methods for ground motion selection, *Earthquake Engineering & Structural Dynamics* 45 (2015) 757–777.
790
- [45] G. E. P. Box, D. R. Cox, An analysis of transformations, *J. R. Stat. Soc. B* 26 (1964) 211—252.
- [46] M. M. Zuniga, A. Murangira, T. Perdrizet, Structural reliability assessment through surrogate based importance sampling with dimension reduction, *Reliability Engineering & System Safety* 207 (2021) 107289.
795
- [47] G. Box, G. Tiao, *Bayesian Inference in Statistical Analysis*, Addison-Wesley, 1973.

- 800 [48] M. Keller, A. Popelin, N. Bousquet, E. Remy, Nonparametric estimation of the probability of detection of flaws in an industrial component, from destructive and nondestructive testing data, using approximate bayesian computation, *Risk Analysis* 35 (2015) 1595–1610.
- [49] C. P. Robert, G. Casella, Monte Carlo Statistical Methods (Springer Texts in Statistics), Springer-Verlag, Berlin, Heidelberg, 2005.
- 805 [50] A. Owen, Y. Zhou, Safe and effective importance sampling, *Journal of the American Statistical Association* 95 (449) (2000) 135–143.
- [51] T. Hesterberg, Weighted average importance sampling and defensive mixture distributions, *Technometrics* 37 (2) (1995) 185–194.
- 810 [52] B. Echard, N. Gayton, M. Lemaire, N. Relun, A combined importance sampling and kriging reliability method for small failure probabilities with time-demanding numerical models, *Reliability Engineering & System Safety* 111 (2013) 232–240.
- [53] J. Bect, R. Sueur, A. Gerossier, L. Mongellaz, S. Petit, E. Vazquez, Echantillonnage préférentiel et méta-modèles : méthodes bayésiennes optimale et défensive, in: 47èmes Journées de Statistique de la SFdS (JdS 2015), 2015.
- 815 [54] B. Delyon, F. Portier, Asymptotic optimality of adaptive importance sampling, in: Proceedings of the 32nd International Conference on Neural Information Processing Systems, NIPS’18, Curran Associates Inc., Red Hook, NY, USA, 2018, pp. 3138—3148.
- 820 [55] V. Chabridon, M. Balesdent, J.-M. Bourinet, J. Morio, N. Gayton, Evaluation of failure probability under parameter epistemic uncertainty: application to aerospace system reliability assessment, *Aerospace Science and Technology* 69 (2017) 526–537.
- 825 [56] J. Morio, M. Balesdent, Estimation of rare event probabilities in complex aerospace and other systems: a practical approach, Woodhead Publishing, 2015.

- [57] F. Bachoc, Cross validation and maximum likelihood estimations of hyper-parameters of gaussian processes with model misspecification, *Comput. Stat. Data Anal.* 66 (2013) 55—69. 830
- [58] F. Golestaneh, P. Pinson, R. Azizipanah-Abarghooee, H. B. Gooi, Ellipsoidal prediction regions for multivariate uncertainty characterization, *IEEE Transactions on Power Systems* 33 (4) (2018) 4519–4530. doi : 10.1109/TPWRS.2018.2791975.
- [59] N. Ambraseys, P. Smit, R. Berardi, D. Rinaldis, F. Cotton, C. Berge, Dissemination of european strongmotion data, cD-ROM collection. European Commission, Directorate-General XII, Environmental and Climate Programme, ENV4-CT97-0397, Brussels, Belgium (2000). 835
- [60] F. Touboul, P. Sollogoub, N. Blay, Seismic behaviour of piping systems with and without defects: experimental and numerical evaluations, *Nuclear Engineering and Design* 192 (2) (1999) 243–260. 840
- [61] T. Charras, J. Kichenin, Développer dans CAST3M (2011).
URL <http://www-cast3m.cea.fr/>
- [62] F. Touboul, N. Blay, P. Sollogoub, S. Chapuliot, Enhanced seismic criteria for piping, *Nuclear Engineering and Design* 236 (1) (2006) 1–9. 845
- [63] A. W. van der Vaart, *Asymptotic Statistics*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, 1998.
- [64] P. Hall, C. Heyde, Z. Birnbaum, E. Lukacs, *Martingale Limit Theory and Its Application, Communication and Behavior*, Elsevier Science, 2014.