Fast Probabilistic Seismic Hazard Analysis through Adaptive Importance Sampling

Soung Eil Houng*[1] and Luis Ceferino[1]

ABSTRACT

Probabilistic Seismic Hazard Analysis (PSHA) relies on two widely utilized approaches with high computational demands: (a) Riemann sum and (b) conventional Monte Carlo (MC) integration. The first requires sufficiently fine slices across magnitude, distance, and ground motion, and the second requires extensive synthetic earthquake catalogs to compute seismic hazards accurately. These approaches are notably resource-intensive for low-probability seismic hazards, e.g., up to 10^8 MC samples for a hazard with 10^{-4} probability to achieve coefficient of variation (COV) of 1%. Here, we present a novel framework to compute hazard and deaggregation with unprecedented computational efficiency. We formulate Adaptive Importance Sampling (AIS) PSHA to approximate optimal important sampling (IS) distributions and dramatically reduce the size of synthetic earthquake catalogs (i.e., number of MC samples) to estimate hazards. We evaluate the effectiveness and reliability of our proposed method using comprehensive test problems from the Pacific Earthquake Engineering Research Center (PEER) for PSHA benchmarks, encompassing various seismic source types, including areal, fault, and combined ones. Our findings indicate that this novel approach significantly outpaces Riemann sum and traditional MC methods with computations up to $>10^5$ and 7.8×10^3 times faster, respectively, while maintaining an standard deviation of the estimate below 2%. Moreover, we show theoretically that optimal IS distributions are equivalent to hazard deaggregation distributions. Empirically, we show our approximated optimal IS and the deaggregation distributions are closely alike, e.g., with a Kolmogorov–Smirnovstatistic between 0.017 and 0.113. We developed our methodology to have broad applicability in PSHA practices, especially in cases requiring extensive computational resources to navigate numerous logic tree scenarios addressing epistemic uncertainty.

1. Civil and Environmental Engineering, University of California, Berkeley, California, U.S.A, https://orcid.org/0000-0001-6873-5327 (SEH) https://orcid.org/0000-0003-0322-7510 (LC)

*Corresponding author: shoung@berkeley.edu

Cite this article as Houng, Soung Eil and Luis, Ceferino (2022). Fast Probabilistic Seismic Hazard Analysis through Adaptive Importance Sampling, *Bull. Seismol. Soc. Am.* **0**, 1–35, doi: 00.0000/0000000000.

KEY POINTS

- We introduce an adaptive importance sampling algorithm for Probabilistic Seismic Hazard Analysis (PSHA).
- The proposed algorithm is up to 10^3 and 10^5 times more efficient than conventional Monte Carlo and Riemann sum.
- The algorithm also facilitates the straightforward implementation of hazard deaggregation.

Supplemental Material

INTRODUCTION

Probabilistic Seismic Hazard Analysis (PSHA) has become a foundational method for determining seismic design levels and conducting regional seismic risk analyses since its first inception (Cornell, 1968; U. S. Nuclear Regulatory Commission, 2007; McGuire, 2008; ASCE, 2022). The development of PSHA was driven by the necessity for a probabilistic framework to quantify seismic hazards, acknowledging the unpredictable nature of earthquake location, magnitudes, and ground motion intensities (Esteva, 1967; Cornell, 1968). In PSHA, earthquake location, magnitude, and ground motions are treated as random variables, facilitating the computation of annual exceedance probabilities at various ground motion intensities. Integrating hazard with fragility curves further enables to determine the annual probabilities of structural failures, thereby underscoring the methodology's critical role in risk assessment (Kennedy et al., 1980; Ceferino et al., 2020; Silva et al., 2020; Baker et al., 2021; Papadopoulos and Bazzurro, 2021; Arora and Ceferino, 2023).

Since we cannot solve PSHA analytically due to the complexity in seismic source and ground motion models, numerous researchers have developed computer softwares for PSHA computation (Cornell, 1968; McGuire, 1976; Field et al., 2003; Ordaz et al., 2013; Pagani et al., 2014). According to the PSHA computer code verification project (Thomas et al., 2010; Hale et al., 2018), existing software primarily employs Riemann summation for numerical integration of PSHA, which partitions the earthquake magnitude, location, and ground motion random variable into fine grids to approximate the actual integration. The Riemann summation offers robust PSHA integration with sufficiently dense grids. However, this method generally incurs a significant computational load exponentially increasing with the number of grids and dimensions in multi-dimensional integrations (Philippe and Robert, 2001). Furthermore, the results are highly sensitive to the chosen grid design (e.g., the initial point of the grid, grid spacing), leading to significant deviations from one software to another, especially for low exceedance probabilities (Thomas et al., 2010; Hale et al., 2018).

Alternatively, other softwares adopted Monte-Carlo (MC) integration for PSHA (Assatourians and Atkinson, 2013, 2019). MC integration calculates exceedance probabilities by generating synthetic earthquake catalogs based on the seismic source and ground motion models to evaluate the recurrence of various ground motion intensities (Musson, 2000). MC integration's primary advantage lies in its straightforward concept compared to Riemann summation, without the need to divide the integration range into small slices (Musson, 2000; Dick et al., 2013). Nevertheless, MC framework requires a substantially-

long synthetic catalog to accurately estimate hazard from rare events, making the computation costly, especially for large ground motions with low exceedance probabilities, e.g., $p < 10^{-4}$ /yr (Kroese et al., 2014).

Importance Sampling (IS) can offer a solution to this rare event simulation (Tokdar and Kass, 2010). IS was initially introduced in statistical physics (Hammersley and Morton, 1954) to improve the computational efficiency of rare event simulation that would otherwise require a large sample size with conventional MC. IS relies on identifying an appropriate probability distribution ("IS distribution") to explore low-probability spaces effectively. Researchers use the IS distribution to sample rare events with a higher likelihood than conventional MC and then correct their frequency through weights, significantly reducing the number of samples to compute low probabilities (Robert et al., 1999). However, finding such an appropriate distribution can be challenging because there is no optimal sampling density that is universally applicable; rather, the proper selection of sampling density depends on the problem being solved. Thus, many numerical experiments are often conducted first through trial-and-error to identify IS distributions, which is still computationally expensive.

In regional seismic risk analysis, numerous studies have been conducted to sample hazard-consistent earthquake ground motions (Crowley and Bommer, 2006; Kiremidjian et al., 2007; Jayaram and Baker, 2010; Han and Davidson, 2012; Manzour et al., 2016; Christou et al., 2018; Kavvada et al., 2022). Kiremidjian et al. (2007) first introduced IS distributions that samples large-magnitude earthquakes with a high probability to reduce the computational burden of seismic hazard and risk analyses. Jayaram and Baker (2010) expanded the approach by defining IS distributions to sample high-intensity ground motions. However, Jayaram and Baker (2010) highlighted the computational challenges to identify an effective IS distribution and ended up using K-mean clustering to reduce the number of ground motion samples. Rahimi and Mahsuli (2019) applied system reliability methods to calculate PSHA. They selected the IS sampling density as the normal distribution centered at the "design point" derived from the first- and second-order reliability method.

To find effective IS distributions, computational statisticians have developed a general framework to find them through iterative algorithms denominated "Adaptive Importance Sampling (AIS)" (Bugallo et al., 2017). AIS algorithms leverage the fact that the optimal density is proportional to the integrand of the problem to find the sampling density which minimizes the variance of the estimate. Various AIS algorithms have been introduced, e.g., cross-entropy based AIS, Vegas, Divonne, and Miser (Lepage, 1978; Friedman and Wright, 1981; Rubinstein, 1997; Press and Farrar, 1990; Rubinstein and Kroese, 2004; Bugallo et al., 2017). The application of AIS is widely adopted to solve the integration with high dimensions, such as in the field of statistical physics, finance, reliability engineering, and signal processing (Au and Beck, 2001; Kappen and Ruiz, 2016; Nieto and Ruiz, 2016; Bugallo et al., 2017). Although previous studies suggest the use of importance sampling (IS) for PSHA calculation (Jayaram and Baker, 2010; Rahimi and Mahsuli, 2019), no research has been published regarding the use of AIS for this purpose. AIS can provide a general methodology for identifying an appropriate IS distribution for seismic hazards, eliminating computationally expensive experiments in the regular IS approach, whose efficiency typically depends on the researcher's experience in the field.

Many AIS algorithms have been published (Lepage, 1978; Friedman and Wright, 1981; Rubinstein, 1997; Press and Farrar, 1990; Rubinstein and Kroese, 2004; Bugallo et al., 2017). However, we study the VEGAS algorithm for PSHA to leverage its straightforward mathematical framework and fast convergence. Thus, this study introduces a novel computational method for PSHA curve calculation using Adaptive Importance Sampling (AIS) VEGAS algorithm (Lepage, 1978, 2021). This paper also shows that AIS facilitates hazard deaggregation, the relative contribution of each random variables to the overall hazard (Bazzurro and Cornell, 1999), because the optimal IS density resembles the contributions of each variable—magnitude, distance, and ground motion—to the hazard. Thus, our proposed framework also enhances the computational efficiency of deaggregation estimates, which otherwise would require additional memory resources and complexity in the computer code with traditional calculation methods. We explore three key aspects of the method: 1) the enhancement of computational efficiency that the AIS algorithm offers over traditional methods; 2) the accuracy of the estimates provided by this approach; and 3) the process of obtaining hazard deaggregation through AIS. We present the theoretical background of AIS PSHA and validate the method through numerical examples.

MATHEMATICAL FORMULATION

Probabilistic Seismic Hazard Analysis (PSHA)

At a site of interest, the annual frequency of ground motion exceedance from a single source can be calculated as:

$$\lambda(X > a) = \nu \int_{\varepsilon_{\min}}^{\varepsilon_{\max}} \int_{r_{\min}}^{r_{\max}} \int_{m_{\min}}^{m_{\max}} I(X > a|m, r, \varepsilon) f_{M,R,\mathcal{E}}(m, r, \varepsilon)\, dmdrd\varepsilon \tag{1}$$

, where $\lambda(X > a)$ is annual rate that ground motion, X, exceeds the target ground motion intensity, a, e.g., peak ground acceleration. ν is the annual rate of earthquake occurrence greater than $m_{\min}$ from the source, m is the earthquake magnitude, $m_{\min}$ and $m_{\max}$ are minimum and maximum magnitudes considered for the source, r is the source-to-site distance, $r_{\min}$ and $r_{\max}$ are minimum and maximum source-site distances, ε is a standard normal random variable for generating earthquake ground motion, $\varepsilon_{\min}$ and $\varepsilon_{\max}$ are minimum and maximum ε (generally, $\varepsilon_{\max} \geq 6$ and $\varepsilon_{\min} \leq -6$; Bommer and Abrahamson (2006)), $f_{M,R,\mathcal{E}}(m, r, \varepsilon)$ is joint probability density function (PDF) of m, r, and ε, $I(X > a|m, r, \varepsilon)$ is indicator function that takes 1 when $X > a$, otherwise, 0. The ground motion X given M, R, and $\mathcal{E}$ is generally calculated using ground motion models (Bozorgnia et al., 2014; Goulet et al., 2021). The models usually assume the log-normal distribution for the ground motion given explanatory variables such as M and R. Naturally, these models provide the mean and standard deviation of logarithmic ground motion. Thus, the random ground motion is calculated as:

$$\log X = \mu(M, R) + \mathcal{E}\sigma(M, R) \tag{2}$$

, where μ and σ are mean and standard deviation of logarithmic earthquake ground motion. By taking exponential on both sides of Eq. (2), the ground motion X can be calculated as

$$X = e^{\mu(M,R)+\mathcal{E}\sigma(M,R)}$$

If we assume that the ground motion random variable ε is independent with respect to the m and r (McGuire, 1995), Eq. (1) can be modified as

$$\begin{aligned}\lambda(X > a) &= \nu \int_{\varepsilon_{\min}}^{\varepsilon_{\max}} \int_{r_{\min}}^{r_{\max}} \int_{m_{\min}}^{m_{\max}} I(x > a|m, r, e) f_{\mathcal{E}}(\varepsilon) f_{M,R}(m, r)\, dmdrd\varepsilon \\ &= \nu \int_{\varepsilon_{\min}}^{\varepsilon_{\max}} \int_{R_{\min}}^{R_{\max}} \int_{M_{\min}}^{M_{\max}} I(x > a|m, r, e) f_{\mathcal{E}}(\varepsilon) f_{R|M}(r|m) f_M(m)\, dmdrd\varepsilon \end{aligned} \tag{3}$$

, where $f_M(m)$, $f_{R|M}(r|m)$, $f_{\mathcal{E}}(\varepsilon)$ are PDF of m, r given m, and ε. Under point source assumption, the distance r and magnitude m become independent random variables. Thus, the seismic hazard is given by

$$\lambda(X > a) = \nu \int_{\varepsilon_{\min}}^{\varepsilon_{\max}} \int_{r_{\min}}^{r_{\max}} \int_{m_{\min}}^{m_{\max}} I(x > a|m, r, \varepsilon) f_{\mathcal{E}}(\varepsilon) f_R(r) f_M(m)\, dmdrd\varepsilon \tag{4}$$

The total seismic hazard from multiple seismic sources (e.g., different faults) is the sum of each. Thus,

$$\Lambda(X > a) = \sum_{i=1}^{n} \lambda_i(X > a) \tag{5}$$

, where $\Lambda(X > a)$ is the total annual frequency of exceedance of ground motion, a, i is index for seismic sources, and n is the total number of seismic sources. Under the assumption of Poisson process, the annual probability of exceedance, $\phi(X > a)$, can be converted from annual frequency of exceedance (Λ; Eq. (5)) as

$$\phi(X > a) = 1 - e^{-\Lambda(X>a)} \tag{6}$$

Hazard deaggregation

We can also deaggregate the total hazard (Eq. (6)) to better understand the earthquakes that contribute most to the hazard. Deaggregation is also used to develop the select seismic records (e.g., from the earthquakes that contribute most to the hazard) and conduct non-linear time-history analyses for the design of many critical buildings (Bazzurro and Cornell, 1999; U. S. Nuclear Regulatory Commission, 2007). Mathematically, deaggregation of the hazard is the joint probability distribution of the m, r, and ε conditional on different levels of hazards a to quantify the contributions of each component. The deaggregation of PSHA can be formulated using Bayes' theorem as

$$
\begin{aligned}
P(m,r,\varepsilon|X>a) &= \frac{P(X>a\cap m,r,\varepsilon)}{P(X>a)} \\
&= \frac{P(X>a|m,r,\varepsilon)P(m,r,\varepsilon)}{P(X>a)} \\
&= \frac{P(X>a|m,r,\varepsilon)P(m,r,\varepsilon)}{\sum P(X>a|m,r,\varepsilon)P(m,r,\varepsilon)} \\
&= \frac{I(X>a|m,r,\varepsilon)P(m,r,\varepsilon)}{\sum I(X>a|m,r,\varepsilon)P(m,r,\varepsilon)}
\end{aligned} \tag{7}
$$

, where $P(X>a|m,r,\varepsilon)$ is the probability of ground motion X is greater than a given m, r, and ε, $P(m,r,\varepsilon)$ is joint probability of m, r, and ε, and $P(X>a)$ is the total probability that the ground motion is greater than a, which is is the summation of $P(X>a|m,r,\varepsilon)P(m,r,\varepsilon)$ over all m, r, and ε. Note that $P(X>a|m,r,\varepsilon)$ can be expressed as an indicator function, $I(X>a|m,r,\varepsilon)$, because the probability of ground motion X greater than a can only be 1 or 0 given m, r, and ε (see Eq. (2)).

By replacing probability mass function, $P(m,r,\varepsilon)$, with probability density function, $f_{M,R,\mathcal{E}}(m,r,\varepsilon)$, and change the summation into integration, Eq. (7) can be expressed as:

$$
f(m,r,\varepsilon|X>a) = \frac{I(X>a|m,r,\varepsilon)f_{M,R,\mathcal{E}}(m,r,\varepsilon)}{\iiint I(X>a|m,r,\varepsilon)f_{M,R,\mathcal{E}}(m,r,\varepsilon)dmdrd\varepsilon} \tag{8}
$$

By Eq. (1), the denominator of Eq. (8) equals λ/ν. Thus,

$$
f(m,r,\varepsilon|X>a) = \nu\cdot\frac{I(X>a|m,r,\varepsilon)f_{M,R,\mathcal{E}}(m,r,\varepsilon)}{\lambda} \tag{9}
$$

This equation shows that the contribution of specific m, r, and ε can be represented by the ratio of the partial sum of the given m, r, and ε to the total hazard. Note that λ and ν are constant. Thus, the hazard deaggregation, $f(m,r,\varepsilon|x>a)$, is proportional to $I(x>a|m,r,\varepsilon)f(m,r,\varepsilon)$:

$$
f(m,r,\varepsilon|x>a) \propto I(x>a|m,r,\varepsilon)f(m,r,\varepsilon)
$$

CURRENT NUMERICAL SOLUTIONS

The Riemann summation This method computes PSHA curves by summing the areas of partitioned (m,r,ε) cuboids. The Riemann summation for Eq. (1) can be expressed as:

$$
\lambda(X>a) = \nu\sum_{k=1}^{N_\varepsilon}\sum_{j=1}^{N_r}\sum_{i=1}^{N_m} I(x>a|m_i,r_j,\varepsilon_k)\, f_{M,R,\mathcal{E}}(m_i,r_j,\varepsilon_k)\Delta m\Delta r\Delta\varepsilon \tag{10}
$$

TABLE 1
Comparison of time complexity of various PSHA algorithms

Algorithm	N_m	N_r	N_ε	N_s	N_a
Riemann sum	N	N	N	-	1
Conventional MC	1	1	1	N	1
IS MC	1	1	1	N	1
VEGAS AIS (this study)	1	1	1	N	N

N_m, the number of magnitude grids; N_r, the number of distance grids; N_ε, the number of ε grids; N_s, the number of (m, r, ε) samples; N_a, the number of ground motion intensity of interest (e.g., 0.1 g)

, where Δm, Δr, and $\Delta\varepsilon$ are grid step size for Riemann sum, N_m, N_r, and N_ε are the total number of grids, satisfying $N_x \Delta x = x_{\max}$ - x_{min}, where $x = \{m, r, \varepsilon\}$. Note that $f_{M,R,\mathcal{E}}(m_i, r_j, \varepsilon_k)\Delta m \Delta r \Delta\varepsilon$ is equivalent to the probability at (m, r, ε) such that $|m - m_i| < \Delta m/2$, $|r - r_j| < \Delta r/2$, and $|\varepsilon - \varepsilon_k| < \Delta\varepsilon/2$. The accuracy of the Riemann summation depends on the grid size. Utilizing finer grids enhances the accuracy of the summation. However, the computation time is inversely proportional to the grid step size and thus proportional to the number of grids. For three-dimensional PSHA summation, the computation time scales with $N_m \times N_r \times N_\varepsilon$ (Table 1). Notably, since the distance PDF ($f_R(r)$) cannot be analytically determined in practice, integration often extends over latitude (ϕ), longitude (ψ), and depth (z), increasing the dimensions from three (m, r, ε) to five $(m, \phi, \psi, z, \varepsilon)$. Therefore, Riemann summation for seismic hazard becomes even more computationally intensive, a phenomenon known as "the curse of dimensionality" due to the exponential increase in computation time with the number of dimensions (Novak and Ritter, 1997).

Conventional Monte-Carlo (MC) This method simulates many synthetic earthquake ground motions and calculates PSHA by assessing the frequency with which ground motion intensities exceed a certain threshold. MC PSHA is computed as

$$\begin{aligned}\hat{\lambda}(X > a) &= \frac{\nu}{N}\sum_{i=1}^{N} I(X_i > a|M_i, R_i, \mathcal{E}_i) \\ &= \frac{1}{T}\sum_{i=1}^{N} I(X_i > a|M_i, R_i, \mathcal{E}_i)\end{aligned} \tag{11}$$

, where X_i denotes the simulated ground motions, and N is the total number of samples. $M_i, R_i, \mathcal{E}_i$ are random samples from $f_{M,R,\mathcal{E}}(m, r, \varepsilon)$, T represents the equivalent catalog duration equal to N/ν. MC PSHA is unbiased because the expectation of $\hat{\lambda}(X > a)$ is the same as λ:

$$\begin{aligned}E[\hat{\lambda}] &= E\left[\frac{\nu}{N}\sum_{i=1}^{N} I(X_i > a|M_i, R_i, \mathcal{E}_i)\right] \\ &= \frac{\nu}{N}\sum_{i=1}^{N} E\left[I(X_i > a|M_i, R_i, \mathcal{E}_i)\right] \\ &= \frac{1}{N}\sum_{i=1}^{N} \nu \int I(X_i > a|m_i, r_i, \varepsilon_i) f_{M,R,\mathcal{E}}(m_i, r_i, \varepsilon_i) dm_i dr_i d\varepsilon_i \\ &= \frac{1}{N}\sum_{i=1}^{N} \lambda(X > a) \\ &= \lambda(X > a)\end{aligned} \tag{12}$$

The variance of $\hat{\lambda}$, the variability of each MC estimates, can be obtained as follows (Appendix A):

$$\mathrm{VAR}[\hat{\lambda}] = \frac{\nu\lambda - \lambda^2}{N} \tag{13}$$

Note that $\mathrm{VAR}[\hat{\lambda}]$ is always positive since $\nu \geq \lambda$. Then, the standard deviation of the estimate, $\hat{\lambda}$, can be obtained as:

$$\sigma[\hat{\lambda}] = \sqrt{\frac{\nu\lambda - \lambda^2}{N}} \tag{14}$$

, which provides the absolute uncertainty about MC PSHA estimates. However, we do not want to fully rely on $\sigma[\hat{\lambda}]$ to compare different exceedance probabilities. For example, suppose we are interested in two different exceedance frequencies, $\lambda_1 = 10^{-1}$/yr and $\lambda_2 = 10^{-2}$/yr when $\nu = 1$/yr and $N = 100$. Then, the variance (uncertainty) of the two MC estimates, $\hat{\lambda}_1$ and $\hat{\lambda}_2$, are $\sigma_1 = 3 \times 10^{-2}$ and $\sigma_2 \sim 1 \times 10^{-2}$, respectively (Eq. (14)). Here, someone might argue that uncertainty of $\hat{\lambda}_1$ MC estimate is larger than that of $\hat{\lambda}_2$ because σ_1 is greater than σ_2. However, the uncertainty of λ_1 should be considered smaller than that of λ_2 considering the target true value of each estimate. That is, 0.1 ± 0.03 ($\hat{\lambda}_1$) is better estimate than 0.01 ± 0.01 ($\hat{\lambda}_2$). Instead, it is better to use the coefficient of variation, COV, the standard deviation normalized by its mean

$$COV = \frac{\sigma[\hat{\lambda}]}{E[\hat{\lambda}]} = \sqrt{\frac{\nu - \lambda}{N\lambda}} \tag{15}$$

Smaller COV indicates that the PSHA estimate is more accurate relative to its actual value, e.g., 5% COV means that the hazard estimates are within ±5% of its true value with a 68% probability, assuming normally distributed MC estimates as per the central limit theorem. COV is inversely proportional to the square root of the number of MC samples (N), indicating that increasing N naturally improves the estimate's accuracy.

The target exceedance frequency, λ, also affects the MC accuracy. The lower λ leads to poor accuracy with fixed N and ν. It is intuitively reasonable that sufficiently long earthquake catalog is required to accurately estimate the event from long return period. It is also notable that the variance of MC estimate is proportional to the square root of earthquake occurrence rate, ν, indicating that MC PSHA is more challenging task in region with higher earthquake activity. This can also be explained intuitively. Regions with high seismic activity experience more earthquakes per year than less active areas, therefore, when calculating the ground motion exceedance probability over the same period, more earthquakes need to be considered in these active regions.

An advantage of MC PSHA is that the computational time is dependent on the number of samples N, circumventing the inherent dimensionality problem in Riemann summation (Table 1). In other words, the computation time of MC simulation is independent of the number of grids (N_m, N_r, and N_ε in Eq. (10)). Therefore, we could utilize fine joint probability mass functions for more precise hazard estimation without an increase in computational burden. If closed-form probability density functions are available, PSHA can be implemented without approximated discretization. Also, the computation time of MC is independent of the number of ground motions of interest, N_a (Table 1). We can utilize the generated synthetic ground motion catalog to tally the exceedance events for all the ground motions of interest, though we still should repeat the computation along with the number of MC samples.

Conventional MC faces extreme computational challenges for low probabilities because the number of samples required to achieve low COVs dramatically increases. We can compute the required number of samples by rearranging Eq. (15):

$$N = \frac{1}{(COV)^2} \times \frac{\nu - \lambda}{\lambda} \tag{16}$$

For small λ, i.e., $\nu >> \lambda$, the Eq. (16) can be approximated as :

$$N \sim \frac{1}{(COV)^2} \times \frac{\nu}{\lambda} \tag{17}$$

To demonstrate conventional MC's extreme computational demands, we can calculate N for a target annual exceedance frequency, λ, of 10^{-4} per year and a annual earthquake rate $\nu = 1/\text{yr}$, which is typical values in PSHA practice (Coppersmith et al., 2014). From Eq. (17), we need $N \sim 10^8$ to achieve COV = 1%.

Importance Sampling (IS) Integration

IS is a generalization of MC theory. Consider a random variable x that follows a probability function, $f_X(x)$. The expected value of a function $u(x)$, denoted as S, is defined by

$$S = \int u(x) f_X(x) dx \tag{18}$$

The MC estimate is

$$\hat{S} = \frac{1}{N} \sum_{i=1}^{N} u(X_i) \tag{19}$$

, where X_i is sampled from $f_X(x)$.

By introducing an arbitrary probability function, $q_X(x)$, Eq. (18) can be equivalently expressed as

$$S = \int u(x) \frac{f_X(x)}{q_X(x)} q_X(x) dx \tag{20}$$

We restrict the integration range in Eq. (20) where $f_X(x) \neq 0$ because x such that $f_X(x) = 0$ does not contribute to the integration. Then, $q_X(x)$ can be any distribution with nonzero density in the integration range. Eq. (20) provides important implications in MC estimation. We can get the solution to Eq. (18) by estimating the expected value of $u(x) f_X(x)/q_X(x)$ where x follows the distribution $q_X(x)$. This approach is highly useful in numerical integration, especially when sampling from $f_X(x)$ is challenging or the population of $f_X(x)$ is extremely low in the region of importance, e.g., the exceedance of PGA greater than 1 g is mostly contributed by the ground motion samples from large magnitude (m) earthquakes occurred at close distance (r) with large ε, which all typically correspond to the low probability region.

The IS MC estimate is:

$$\hat{S} = \frac{1}{N} \sum_{i=1}^{N} u(X_i) \frac{f_X(X_i)}{q_X(X_i)} \tag{21}$$

, where X_is are sampled from $q_X(x)$, the proposed (or new) IS sampling density function. If $q_X(x)$ is equal to the original distribution, $f_X(x)$, Eq. (21) simplifies to the conventional Monte-Carlo (Eq. (19)). The ratio $f_X(x)/q_X(x)$, known as the importance weight (w_i), adjusts for the change in sampling distribution.

IS PSHA Researchers have applied IS to PSHA using different sampling functions (Jayaram and Baker, 2010; Rahimi and Mahsuli, 2019). For IS PSHA, Eq. (1) can be reformulated by introducing a new sampling joint density function, $q_{M,R,\varepsilon}(m, r, \varepsilon)$

as follows

$$\lambda(X > a) = \nu \int_{\varepsilon_{\min}}^{\varepsilon_{\max}} \int_{r_{\min}}^{r_{\max}} \int_{m_{\min}}^{m_{\max}} I(x > a|m, r, e) \frac{f_{M,R,\mathcal{E}}(m, r, \varepsilon)}{q_{M,R,\mathcal{E}}(m, r, \varepsilon)} q_{M,R,\mathcal{E}}(m, r, \varepsilon)\, dmdrd\varepsilon \tag{22}$$

The IS MC estimator of equation (22) is

$$\hat{\lambda}(X > a) = \frac{\nu}{N} \sum_{i=1}^{N} I(X_i > a|M_i, R_i, E_i) \frac{f_{M,R,\mathcal{E}}(M_i, R_i, E_i)}{q_{M,R,\mathcal{E}}(M_i, R_i, E_i)}$$

The mean of IS estimator $\hat{\lambda}$ is also unbiased like conventional MC because

$$\begin{aligned} E_q[\hat{\lambda}] &= \frac{\nu}{N} \sum_{i=1}^{N} E_q \left[I(X_i > a|M_i, R_i, \mathcal{E}_i) \frac{f_{M,R,\mathcal{E}}(M_i, R_i, \mathcal{E}_i)}{q_{M,R,\mathcal{E}}(M_i, R_i, \mathcal{E}_i)} \right] \\ &= \frac{1}{N} \sum_{i=1}^{N} \nu \int I(X > a|m, r, \varepsilon) \frac{f_{M,R,\mathcal{E}}(m, r, \varepsilon)}{q_{M,R,\mathcal{E}}(m, r, \varepsilon)} q_{M,R,\mathcal{E}}(m, r, \varepsilon) dmdrd\varepsilon \\ &= \frac{1}{N} \sum_{i=1}^{N} \lambda \\ &= \lambda \end{aligned} \tag{23}$$

Also, the variance of the IS PSHA estimates with respect to the true λ is given by

$$\mathrm{VAR}[\hat{\lambda}] = \frac{1}{N} \left(\nu^2 E \left[\left(I(X_i > a|M_i, R_i, \mathcal{E}_i) \frac{f_{M,R,\mathcal{E}}(M_i, R_i, \mathcal{E}_i)}{q_{M,R,\mathcal{E}}(M_i, R_i, \mathcal{E}_i)} \right)^2 \right] - \lambda^2 \right) \tag{24}$$

Hence, the COV of IS PSHA estimate can be expressed as

$$COV = \sqrt{\frac{\nu^2 E \left[\left(I(X_i > a|M_i, R_i, \mathcal{E}_i) \frac{f_{M,R,\mathcal{E}}(M_i, R_i, \mathcal{E}_i)}{q_{M,R,\mathcal{E}}(M_i, R_i, \mathcal{E}_i)} \right)^2 \right] - \lambda^2}{N\lambda^2}} \tag{25}$$

Equivalency of optimal IS density and hazard deaggregation

From Eq. (24), we can specifically choose a new sampling density $q^*_{M,R,\mathcal{E}}$ that makes VAR[$\hat{\lambda}$]=0,

$$q^*_{M,R,\mathcal{E}} = \nu \frac{I(X_i > a|M_i, R_i, E_i) f_{M,R,\mathcal{E}}(M_i, R_i, E_i)}{\lambda} \tag{26}$$

Using $q^*_{M,R,\mathcal{E}}$, we could compute the true hazard, λ, with only one MC sample because IS MC is unbiased (Eq. (23)) and the variance is zero. We call q^* the optimal IS density for PSHA calculation.

We note that Eq. (26) is exactly the same as Eq. (9), indicating that the optimal density, q^*, is identical to the hazard deaggregation. That is, if we find q^*, we are able to not only dramatically enhance the computational efficiency of seismic hazard estimation but also obtain the hazard deaggregation distributions as a by-product.

In fact, we can also see this profound relationship between the hazard and deaggregation estimates by rearranging Eq. (9):

$$\lambda = \nu \cdot \frac{I(X > a|m, r, \varepsilon) f_{M,R,\mathcal{E}}(m, r, \varepsilon)}{f(m, r, \varepsilon|X > a)}$$

Because this identity holds for any values of (m, r, ε), we can obtain the hazard at ground motion level , a, with any (m, r, ε) triplet if we know $f(m, r, \varepsilon|X > a)$. The term $f_{M,R,\mathcal{E}}(m, r, \varepsilon)/f(m, r, \varepsilon|X > a)$ can be interpreted as the importance weight of IS, and $f(m, r, \varepsilon|X > a)$, the hazard deaggregation, is the optimal density.

In existing numerical methods, the deaggregation can be more computationally expensive than the hazard. In Riemann sum, we must save each sum element in memory and allocate those elements into appropriate deaggregation bins. Also, in conventional MC, we should save the long synthetic ground motion catalog with the corresponding (m, r, ε) triplet to allocate those into the proper bins. These operations necessitate significant computational memory and time. Therefore, the property that the optimal density resembles the hazard deaggregation can be considered a huge benefit for hazard analysts. IS MC with different densities other than q^* could also improve the computational efficiency compared to the conventional MC; however, it cannot give any information on the hazard deaggregation (Rahimi and Mahsuli, 2019; Jayaram and Baker, 2010).

Though the use of optimal density is a major benefit in computation of seismic hazard and hazard deaggregation, obtaining it is not trivial because $I(X_i > a|M_i, R_i, E_i)$ and λ in Eq. (26) are unknowns before calculation of PSHA. Thus, we propose a new PSHA computation method to find it.

ADAPTIVE IMPORTANCE SAMPLING PSHA

In "adaptive" importance sampling (AIS), we iteratively train the IS density to find the optimal one (q^*) by exploring important regions to compute $I(X_i > a|M_i, R_i, E_i)$ and λ with a reduced number of MC samples. AIS must balance different factors in determining how many samples should be used to train optimal IS density. Fewer MC samples can make AIS fail as they will not allow us to explore the important regions effectively. On the other hand, many samples can impose large computational demands, even larger than those from conventional MC. In AIS, we must also consider computational costs are proportional to the number of iterations for convergence. Thus, it is important to select appropriate algorithms that converge fast.

VEGAS Formulation for PSHA VEGAS is a non-parametric adaptive importance sampling (AIS) algorithm that iteratively identifies the optimal proposal density, q^* (Lepage, 1978, 2021). The algorithm has been developed and widely used in com-

putational physics (Kersevan and Richter-Was, 2013; Alwall et al., 2014), and is currently applied to chemistry, astrophysics, finance, and medical statistics (Campolieti and Makarov, 2007; Garberoglio and Harvey, 2011; Ray et al., 2011; Sanders, 2014).

The VEGAS algorithm is conceptually straightforward and is recognized for its rapid convergence, especially when the random variables involved are independent (Lepage, 1978, 2021). In PSHA, the variable ε is always considered an independent variable (Eq. (3)). Additionally, the variables m and r are also treated as independent under point source assumption (Eq. (4)). Notice, however, that when the finite-fault rupture model, in which the rupture dimension changes with magnitude, is adopted, m and r can be correlated, and the distribution of r is conditional on the magnitude m. Also, note that the estimate remains unbiased in this case, because its calculation is still within the IS framework (Eq. (23)).

In three-dimensional integration, which is the case of PSHA, VEGAS employs N^3 cuboids that are independently partitioned. The probability assigned to each cuboid and the total number of partitioned cuboids (N^3) are preserved across the iteration steps. However, the IS sampling density changes because the algorithm updates the cuboid's size depending on its contribution to the integration. If a cuboid's contribution is low, its size grows in the next step, lowering its probability density. Conversely, when a cuboid's contribution is high, it shrinks in the following step, elevating the probability density. Ideally, when every cuboid's contribution to the integration becomes identical, we find the proposed optimal density, q^*, and the algorithm is terminated. The framework to find optimal IS densities for PSHA using VEGAS algorithm is explained with a simple point seismic source example in the following paragraphs.

First, we adopt an independently distributed joint probability function as IS density. Note that we could also include correlations in the IS density, but such approach would increase computational memory and time demands, e.g., the computational complexity increases exponentially with each added dimension, i.e., $O(N^d)$ (Lepage, 1978). In contrast, when we assume independence, the computational complexity grows linearly with the number of dimensions, i.e., $O(Nd)$, making multi-dimensional integration in PSHA exceptionally efficient. Thus,

$$q(m, r, \varepsilon) = q_M(m) q_R(r) q_{\mathcal{E}}(\varepsilon) \tag{27}$$

Then, the integration ranges for each variable—m, r, and ε—are divided into N grids with the same volume. This division is designed to generate cuboids of constant probability:

$$\begin{aligned} &M : m_{i-1} \le m < m_i \ (i = 0, 1, 2, .., N), \Delta m_i = m_i - m_{i-1} \\ &R : r_{j-1} \le r < r_j \ (j = 0, 1, 2, .., N), \Delta r_j = r_j - r_{j-1} \\ &\mathcal{E} : \varepsilon_{k-1} \le \varepsilon < \varepsilon_k \ (k = 0, 1, 2, .., N), \Delta \varepsilon_k = \varepsilon_k - \varepsilon_{k-1} \end{aligned} \tag{28}$$

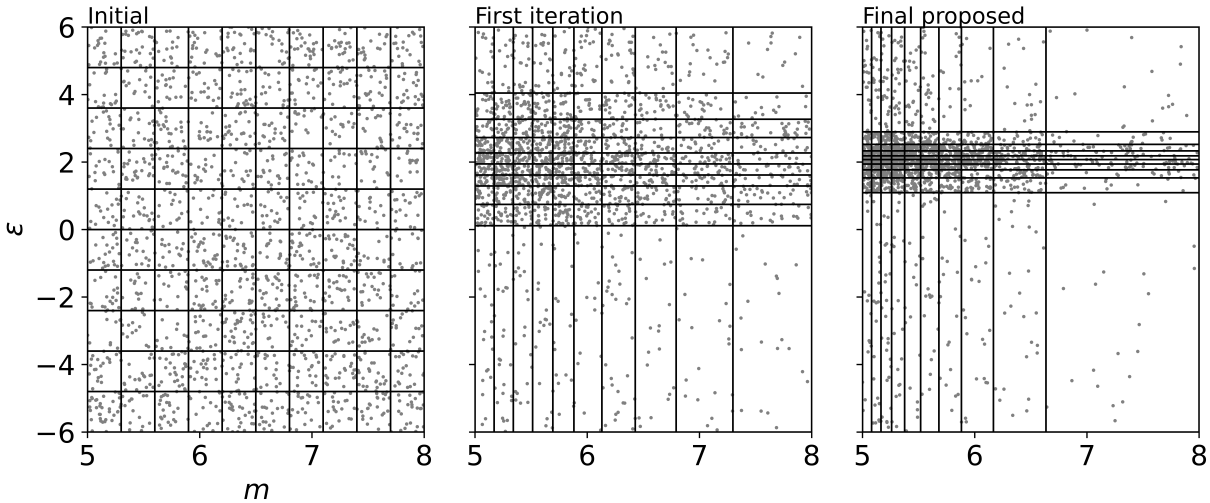

Figure 1. VEGAS iterations of IS density in our AIS PSHA framework. Example for PSHA at PGA = 0.5 g when the site is located 10 km away from a point seismic source ($m_{\min}$ = 5.0, $m_{\max}$ = 8.0, b-value = 1.0, ν = 1.0/yr, GMM = Sadigh et al. (1997)). The total number of MC samples is 2,000, and the number of grids per axis is 10, constituting total 100 rectangles (cuboids in actual three dimensional PSHA integration). The vertical and horizontal black solid lines are boundaries of the m and ε rectangles, and gray dots are the MC samples. Starting from the initial same-size m and ε rectangles, their sizes are adjusted depending on the the contribution of each rectangle to the hazard. The final proposed structure of the rectangles gives a highly concentrated probability density at $m \sim 5$ and $\varepsilon \sim 2$. Note that the hazard estimates using initialized density, density after the first iteration, and the final proposed density are 0.0417, 0.0413, and 0.0378, while the true solution is 0.0385

The number of grids, N, is chosen to be 50 as suggested by Lepage (1978) and because we consider this value makes grids sufficiently fine to capture the actual distribution of the optimal density q^*. The probability for each cuboid from Eq. (28) is set to $1/N^3$ to ensure that the initial IS density function is uniformly distributed across the entire domain. As a result, the initial probability density $q^{(0)}$ for a specific cuboid is

$$\begin{aligned} q^{(0)}(m_{i-1} \le m < m_i,\ r_{j-1} \le r < r_j,\ \varepsilon_{k-1} \le \varepsilon < \varepsilon_k) &= q_M^{(0)}(m)\, q_R^{(0)}(r)\, q_\varepsilon^{(0)}(\varepsilon) \\ &= \frac{1}{N\Delta m_i}\frac{1}{N\Delta r_j}\frac{1}{N\Delta \varepsilon_k} \end{aligned}$$

Figure 1 illustrates an example of initial partitioning when N=10. The gray dots in the figure shows an example of MC samples from this initial q^0. Note that for easier visualization and understanding, Figure 1 illustrates an example in m-ε, the two-dimensional space, and not the three-dimensional m, r, and ε space.

We update q using MC samples to ultimately make it converge to q^* (Eq. (26)). As mentioned earlier, the size of each cuboid is subject to change while the probability of each cuboid remains constant, i.e., probability density changes. This adjustment is facilitated through a "subdivision-and-restoration" process (Lepage, 1978). In this process, ith grid is subdivided into n_i sub-grids, with n_i being proportional to the ith grid's contribution to the overall integration, and restored to the original number, N, by merging N_{subgrid}/N consecutive subgrids, where N_{subgrid} is the total number of sub-grids. Thus, the number

of subdivisions at ith grid is

$$n_i = N_{\text{subgrid}} \times \left(\frac{\overline{H}_i \Delta m_i}{\sum_i \overline{H}_i \Delta m_i} \right), \ (i = 1, 2, ..., N) \tag{29}$$

$$n_j = N_{\text{subgrid}} \times \left(\frac{\overline{H}_j \Delta r_j}{\sum_j \overline{H}_j \Delta r_j} \right), \ (j = 1, 2, ..., N) \tag{30}$$

$$n_k = N_{\text{subgrid}} \times \left(\frac{\overline{H}_k \Delta \varepsilon_k}{\sum_k \overline{H}_k \Delta \varepsilon_k} \right), \ (k = 1, 2, ..., N) \tag{31}$$

N_{subgrid} should be sufficiently larger than N to iterate the IS density effectively, especially when the grids' contributions to the hazard from the previous stage are highly heterogeneous. In this study, we chose N_{subgrid} to be 10,000, which is 200 times greater than N (=50). The second term of Eq. (29)-Eq. (31) represents the portion of each grid's contribution to the hazard.

According to Lepage (1978), $\overline{H}_i$ in Eq. (29) is

$$\overline{H}_i = \sqrt{\sum_{m_{i-1} < M < m_i} \frac{H^2(M, R, \mathcal{E}; a)}{q_r(R) q_\varepsilon(\mathcal{E})}} \tag{32}$$

where

$$H(M, R, \mathcal{E}; a) = I(X > a | M, R, \mathcal{E}) f_{M,R,\mathcal{E}}(m, r, \varepsilon)$$

Note that $H(M, R, \mathcal{E}; a)$ is the integrand of PSHA. Intuitively, $\overline{H}_i$ can be considered as the marginalization of the overall contribution of the magnitude dimension within the ith grid. Note that the summation of H^2s in Eq. (32) is done over all the samples which fall within m_{i-1} and m_i and the division by q_r q_ε can be considered an adjustment for unevenly distributed r and ε samples to purely capture the magnitude contribution.

Similarly, $\overline{H}_j$ and $\overline{H}_k$ in Eq. (30) and Eq. (31) can be calculated as

$$\overline{H}_j = \sqrt{\sum_{r_{i-1} < R < r_i} \frac{H^2(M, R, \mathcal{E}; a)}{q_m(M) q_\varepsilon(\mathcal{E})}}$$

$$\overline{H}_k = \sqrt{\sum_{\varepsilon_{i-1} < \mathcal{E} < \varepsilon_i} \frac{H^2(M, R, \mathcal{E}; a)}{q_r(R) q_m(M)}}$$

In the early iterations, researchers have noticed numerical instabilities due to the relatively poor information on the integrand (Lepage, 1978, 2021). However, researchers have also found effective ways to mitigate it through smoothing (Eq. (33))

and damping (Eq. (34)).

$$d_i := \frac{1}{\sum d_i} \times \begin{cases} (7d_i + d_{i+1})/8 & \text{if } i = 0 \\ (d_{i-1} + 6d_i + d_{i+1})/8 & \text{if } i = 1, 2, ..., N-1 \\ (d_{i-1} + 7d_i)/8 & \text{if } i = N \end{cases} \tag{33}$$

, where d_i is $\overline{H}_i \Delta x_i / \sum_i \overline{H}_i \Delta x_i$, and x are m, r, or ε in Eq. (29)-(31).

$$d_i := \left(\frac{1 - d_i}{\ln(1/d_i)} \right)^{\alpha} \tag{34}$$

, where α is learning rate. We used α as 1.0 as suggested by Lepage (2021). Through numerical experimenting, we observed this value allowed most of the PSHA integration to converge within three iterations (as described later) without causing any numerical instabilities.

Finally, based on calculated n_i, n_j and n_k (Eq. (29)-Eq. (31)), the number of grid is restored to the original size, N, by merging N_{subgrid}/N consecutive subgrids. The restored grid is the updated density, $q^{(1)}$. Figure 1 illustrates this iteration. The sizes of the grids are shrunk at $5 < m < 6$ and $\varepsilon \sim 2$, indicating that the high contributions to the hazard on this range (Eq. (29)-Eq. (31)) in contrast to other less important regions, e.g., at $m > 6.5$ and $\varepsilon < 1$. Note that the probability of each grid is preserved so that the number of MC samples inside each grid is almost the same regardless of the grid size.

The iterative process explained above is repeated until no further improvement is observed in the variance of the hazard estimator (Eq. (24)). An example of the grid structure of the final proposed IS density is shown in Figure 1. The final proposed density exhibits features which is expected based on intuition. First, the low contribution of ε less than 0 is understandable. This is because the simulated ground motion intensity, using Sadigh et al. (1997), never exceeds the target ground motion of 0.5 g at any magnitude, even at $m = 8.0$, when ε is less than 0.1. In addition, the strong contribution of $m \sim 5$ and $\varepsilon \sim 2.2$ to the integration makes sense because the simulated ground motion intensity exceeds the target ground motion of 0.5 g when m and ε reach 5 and 2.2, respectively. The decreasing trend of contribution beyond $m \sim 5$ and $\varepsilon \sim 2.2$ can be interpreted as the exponential decay in the probability of m and ε. This result for a simple PSHA example implies that the proposed algorithm can be utilized for more complex PSHA integration.

NUMERICAL EXAMPLES

We tested our proposed AIS PSHA on the comprehensive benchmark problem sets 1.11 and 2.1 from the PEER PSHA code verification project (Hale et al., 2018). For a thorough comparison, we performed numerical computations using four different algorithms: Riemann sum, conventional Monte Carlo (MC), importance sampling (IS), and adaptive importance sampling (AIS) using the VEGAS algorithm. These computations were conducted across various seismic source settings:

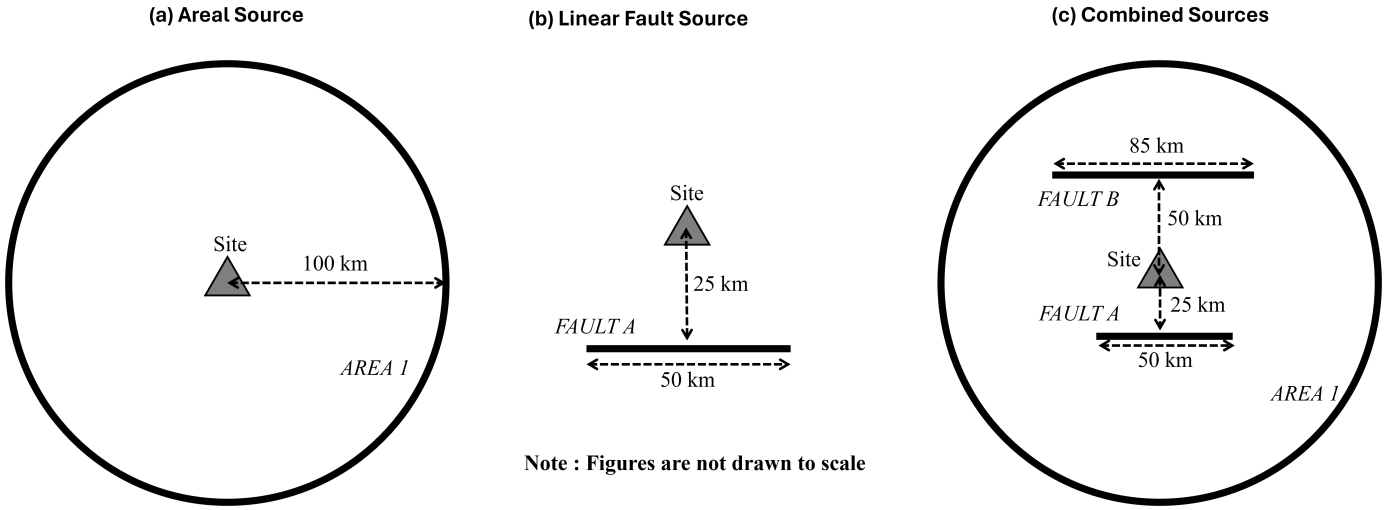

Figure 2. Seimsic source geometry for the three comprehensive numerical examples to test our AIS PSHA framework.

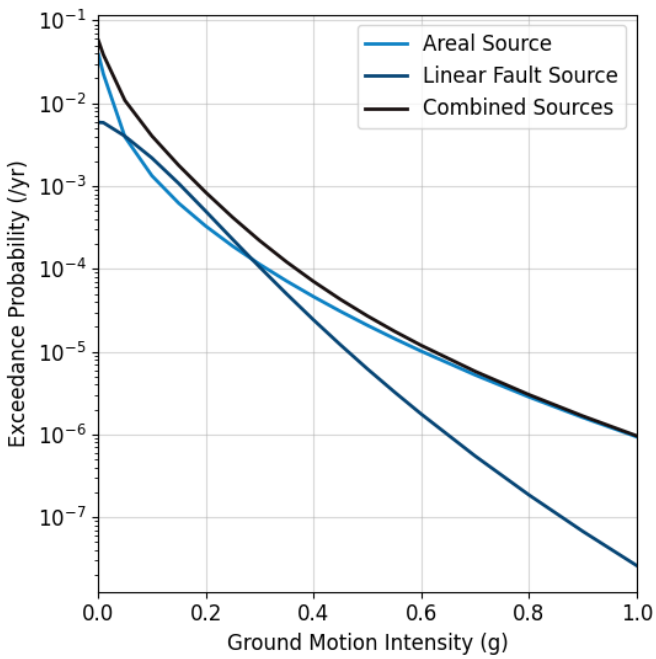

Figure 3. Benchmark PSHA curves for numerical examples 1 (areal source), 2 (linear fault source), and 3 (combined sources)

1) an areal source (Area1), 2) a fault source (FaultA), and 3) a combination of an areal and two fault sources (Area1, FaultA, FaultB). These examples' geometry and seismic activity parameters are detailed in Figure 2 and Table 2. Figure 3 presents the benchmark hazard curves for the examples. The ground motion model from Sadigh et al. (1997) was employed.

For Areal source 1, a circular-shaped areal source with a 100 km radius is considered (Figure 2). The earthquake activity rate of the source, $\nu(M > m_{\min})$, is 0.0395/year with $m_{\min}$ and $m_{\max}$ of 5.0 and 6.5, respectively. The earthquake occurrence model is doubly-truncated exponential with Gutenberg-Richter b-value of 0.9. Seismogenic depth is 5 to 10 km from the surface, and a point source is assumed.

TABLE 2
Seismic source information for comprehensive numerical examples to test our AIS PSHA framework

	Area 1	**Fault A**	**Fault B**
Source Type	Area	Vertical Fault	Vertical Fault
Source Geometry	Circle (R=100 km)	Line (L = 50 km)	Line (L = 85 km)
Earthquake Occurrence Model	Exponential	Characteristic	Characteristic
b-value	0.9	0.9	0.9
$M_{\min}$	5.0	5.0	5.0
$M_{\max}$	6.5	6.75	7.0
M_{char}	-	6.5	6.75
$\nu(M > M_{\min})$ (yr^{-1})	0.0395	-	-
slip rate (mm/yr)	-	1	2
Seismogenic depth (km)	5-10	0-12	0-12
Rupture type	Point	Floating rectanglular rupture	Floating rectanglular rupture

Faults A and B are vertical fault sources (strike = 90°E, dip = 90°) with lengths of 50 and 85 km, respectively. They extend from the surface to 12 km depth. The earthquake activity models are chosen to be characteristic (Youngs and Coppersmith, 1985) with the b-value of 0.9, and $m_{\min}$ is set to be 5.0 for both fault sources. For Fault A, slip rate, $M_{\max}$, M_{char} are 1 mm/year, 6.75, and 6.5; and for fault B, they are 2 mm/year, 7.0, and 6.75. In fault sources, ruptures are assumed to be floating inside the fault with the rupture dimensions following

$$\log_{10}(A) = M - 4 \tag{35}$$

$$\log_{10}(W) = 0.5M - 2.15 \tag{36}$$

$$\log_{10}(L) = 0.5M - 1.85 \tag{37}$$

, where A is rupture area in km^2, W is rupture width in km, L is rupture length in km, and M is earthquake magnitude. The probability of exceedance at eighteen ground motions (PGA), 0.001, 0.01, 0.05, 0.1, 0.15, 0.2, 0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.7, 0.8, 0.9, and 1.0 g, are estimated.

For Riemann summation, we followed the calculation procedure specified in the PEER PSHA code verification project (Hale et al., 2018). Thus, we used magnitude and source-to-site distance step sizes of 0.01 and 0.1 km, respectively. Note that the number of grids of the source-to-site distance distribution ($\sim$ 950) is significantly less than that of source location probability, the uniformly distributed probability of event location across the seismic source region, with the spatial and depth grid spacing of 0.5 km and 1 km ($\frac{\pi \times 100^2}{0.5 \times 0.5} \times 6 \approx 750{,}000$) suggested by Hale et al. (2018). This makes the Riemann sum computation $\sim$ 800 times more efficient. For the ground motion random variable (ε), the grid step size was set to 0.01, with minimum and maximum values of -6 and 6, respectively, beyond which the PSHA curve shows negligible change even at an annual exceedance probability of 10^{-8} (Bommer and Abrahamson, 2006).

In the conventional MC approach, we generate a ground motion catalog based on the probability distributions of m , r , and ε. From this catalog, we calculate the annual exceedance frequency by counting the instances where the ground motion

exceeds the specified threshold, as outlined in Eq. (11). This frequency is then converted to an annual exceedance probability using Eq. (6).

For IS, we adopt uniform IS densities across the entire integration range for magnitude, distance, and ε following Jayaram and Baker (2010). This choice is made due to the absence of prior information regarding which integration range significantly influence the hazard calculation.

For our AIS PSHA with the VEGAS algorithm, we set the initial number of grids for m, r, and ε to 50 (Eq. (28)). The total number of subgrids, N_{subgrid}, is chosen to be 10,000. The learning rate, α, is fixed at 1.0. Algorithm 1 presents the pseudocode. From lines 1 to 5, the generation index t is set to be zero, and the sampling density q is set to be uniform. The main algorithm loop is from lines 6 to 24. The loop is continued until there is no improvement in the coefficient of variation or there is no previous generation (line 6). At line 7, X_m, X_r, X_ε are sampled from the distribution $q^{(t)}$. Then, the probability $\hat{\lambda}$ is estimated in lines 8-10, and its coefficient of variation is also calculated in line 11. From lines 12 to 23 the sampling function $q^{(t)}$ is updated. The steps are repeated over m, r, and ε (lines 12-15). From lines 16 to 20, the contribution of each grid is calculated, at line 21, it is smoothed and dampened, and $q^{(t)}$ is updated by subdivision depending on d_i and restore the number of grids to the original number, N, in line 22. Updated $q^{(t+1)}$ is obtained by multiplying $q_m^{(t+1)}$, $q_r^{(t+1)}$, and $q_\varepsilon^{(t+1)}$ in line 23, and the while loop is ended by increase the generation index t. When main algorithm loop is terminated, it returns the hazard estimate $\hat{\lambda}$ and proposed optimal density q in line 25.

We assessed the accuracy of conventional MC, IS, and AIS probability estimates through the standard deviation of the relative error with respect to the benchmark curve, calculated using the following formula:

$$\sigma = \sqrt{\frac{1}{N}\sum_{n=1}^{N} e_n^2 - \left(\frac{1}{N}\sum_{n=1}^{N} e_n\right)^2}$$

, where e_n is the relative error of the nth MC estimate, and N is the total number of MC exceedance probability estimates. Since MC, IS MC, and VEGAS AIS estimates are unbiased (Eq. (12), Eq. (23)), the sum of e_n when $N \to \infty$ is theoretically zero. The relative error, e_n, is defined as:

$$e_n = \frac{p_n - p_{ref}}{p_{ref}}$$

, where p_n is the exceedance probability estimated at nth MC estimate, and p_{ref} is the benchmark probability.

The analyses were conducted in a Python 3.11 on an Intel Core i7-13700 2100 MHz processor with 64GB RAM.

Example 1 : Areal source

We consider the circular area source with the site at the circle's center in Figure 2. For this example, we show different computational performances to conduct PSHA in Figure 4. Our numerical experiments show that computational times and standard

Algorithm 1 VEGAS adaptive importance sampling PSHA pseudocode

[Parameters and Functions]

a	Ground motion intensity of interest (e.g., $a = 0.1$ g)
N	Number of grids (e.g., N =50)
N_{subgrid}	Number of sub-grids (e.g., N_{subgrid} =10,000)
N_s	Number of MC samples (e.g., $N_s = 1{,}000$)
$\vec{X}_m$	Magnitude sample vector ($\vec{X}_{m,i} = i$th element of $\vec{X}_m$)
$\vec{X}_r$	Distance sample vector
$\vec{X}_\varepsilon$	Ground motion random variable sample vector
$\vec{X}$	$[\vec{X}_m; \vec{X}_r; \vec{X}_\varepsilon]$
$G(\cdot)$	Ground Motion Model
$I(\cdot)$	Indicator function
$f_X(\cdot)$	Original sampling distribution
$q^{(t)}(\cdot)$	Proposed sampling distribution at tth iteration step
$\hat{\lambda}^{(t)}$	Hazard estimate at tth iteration step
$\text{COV}^{(t)}$	COV of the hazard estimate at tth iteration step
ϵ	pre-defined iteration stopping criteria (e.g. COV of 0.02)
i	Index for partitioned grids ($i = 1, 2, \cdots, N$)
j	Index for samples ($j = 1, 2, \cdots, N_s$)

[Algorithm]

$t = 0$
$q_m^{(t)} = \text{U}([m_{\min}, m_{\max}])$
$q_r^{(t)} = \text{U}([r_{\min}, r_{\max}])$
$q_\varepsilon^{(t)} = \text{U}([\varepsilon_{\min}, \varepsilon_{\max}])$
$q^{(t)} = q_m^{(t)}\, q_r^{(t)}\, q_\varepsilon^{(t)}$

while $\text{COV}^{(t-1)} < \epsilon$ **or** t<2 **:**
  $\vec{X} = \{X_m, X_r, X_\varepsilon\}^{N_s}, (X_m \sim q_m^{(t)},\ X_r \sim q_r,\ X_\varepsilon \sim q_\varepsilon^{(t)})$
  $\vec{H}^{(t)} = \{\, I(G(x) > a|x) f_X(x) \mid x \in \vec{X} \,\}$
  $\vec{\lambda}^{(t)} = \vec{H}^{(t)} / q^{(t)}(\vec{X})$
  $\hat{\lambda}^{(t)} = \sum_1^{N_s} \lambda_i^{(t)} / N_s$
  $\text{COV}^{(t)} \leftarrow$ Eq. (25)
  for u **in** $[m, r, \varepsilon]$ **:**
    if $u = m$ **:** $(v, w) \leftarrow (r, \varepsilon)$
    if $u = r$ **:** $(v, w) \leftarrow (\varepsilon, m)$
    if $u = \varepsilon$ **:** $(v, w) \leftarrow (m, r)$
    for i **in** $\{1, 2, ..., N\}$ **:**
      for j **in** $\{1, 2, ..., N_s\}$ **:**
        if $u_{i-1} < X_{u,j} < u_i$ **:** $\overline{H}_i \leftarrow \overline{H}_i + (H_j^2 / \left(q_v(X_{v,j}) \cdot q_w(X_{w,j})\right)$
      $\overline{H}_i \leftarrow \sqrt{\overline{H}_i}$
      $d_i \leftarrow \left(\frac{\overline{H}_i \Delta x_i}{\sum_i \overline{H}_i \Delta x_i} \right)$
      $d_i \leftarrow$ smoothed, dampened d_i (Eq. (33), (34))
    $q_u^{(t+1)} \leftarrow$ subdivision and restoration (Eq. (29)-(31))
  $q^{(t+1)} \leftarrow q_m^{(t+1)}\, q_r^{(t+1)}\, q_\varepsilon^{(t+1)}$
  $t \leftarrow t + 1$
return $\hat{\lambda}^{(t-1)}, q^{(t-1)}$

deviations ("accuracy") are linearly correlated in logarithmic scale, in agreement with the theory, because the required number of MC samples (linearly proportional to the computational time) is inversely proportional to the square of the standard deviation (Eq. (16)).

While the computation time and standard deviation vary depending on the target ground motion of interest, our results show our AIS PSHA generally outperforms the other numerical techniques. AIS PSHA becomes increasingly efficient for larger ground motions. At a low target ground motion of 0.05g, the computation time to achieve a 2 % standard deviation is 0.01 seconds for AIS, while it takes 0.05 and 0.09 seconds for conventional MC and IS estimates, respectively, indicating AIS is 4.8 and 8.5 times faster. However, computational efficiency becomes extreme in high-ground motions. At 1.0 g, the computation time to achieve 2 % standard deviation is estimated to be 0.02 seconds for AIS, while it takes 127 and 1.4 seconds for conventional MC and IS estimates. In this case, AIS is 7,800 and 70 times faster than conventional MC and IS, respectively. Also, note that for 2 % standard deviation case, AIS is $> 10^5$ faster than Riemann sum. We also note that AIS outperforms IS in all ground motion ranges by a factor of 8 to 70 to achieve a 2 % standard deviation. However, this is not always guaranteed because AIS takes t times more computation time than IS with the same N due to the t iterations to find the optimal IS density. Thus, our findings imply AIS PSHA with the VEGAS algorithm can find (close to) optimal IS density quickly.

It is also noteworthy that with fixed number of samples (i.e., constant computational burden), AIS's estimation accuracy is quite similar across different target ground motions, while conventional MC's accuracy decreases sharply for higher ground motions (Figure 5). In fact, if the hazard curve is exponential (Marzocchi and Jordan, 2017), we can show conventional MC decreases its efficiency also exponentially (Eq. (15)). For example, the error in the hazard curve for 1.0g will increase to 100% if it initially was 1 % for 0.001g with exceedance frequency of 0.9/year (when $\nu = 1$/year). From our numerical experiments, we found errors grow from 0 to 5.60 % with $N_{\mathrm{convMC}} = 10^7$ MC samples for these ground motions. In contrast, for a similar computation time, we found AIS had errors ranging from 0.13 to 0.64 %, for this quite different ground motion levels (Figure 5).

This finding is key for PSHA as the computational bottleneck is at the highest ground motion intensity. Using conventional MC PSHA, the hazard analyst has no choice but to largely increase the number of samples to estimate hazard accurately at high ground motions even though such a large number would not be necessary for low ground motions. This makes the conventional MC highly inefficient, and consequently, the efficiency of the conventional MC for lower ground motions cannot be considered a real advantage for PSHA. AIS PSHA overcomes this problem by adopting different optimal densities at different ground motion intensities, making the computational burden to achieve the similar estimation error almost flat for any ground motion intensity, as shown in Figure S1. In terms of the accuracy of AIS PSHA, the PEER PSHA verification project suggests a strict acceptable error range of 5% for reliable PSHA computation codes. AIS PSHA achieves it with only N $\sim$ 10,000 per ground motion (Figure S1).

Another key advantage of our proposed AIS PSHA is the co-production of deaggregation curves at no extra computational cost. From Eq. (26), optimal IS densities are theoretically equivalent to hazard deaggregation distributions. We show AIS can find close-to-optimal IS densities and thus closely match hazard deaggregation curves. We found the benchmark marginal distributions of hazard deaggregation obtained from Riemann sum closely match the iterated IS density from AIS PSHA (Figure 6). We used the Kolmogorov-Smirnov (K-S) D statistic (Kolmogorov, 1933) to quantify their similarities. K-S D statistic measures the maximum difference between two cumulative distribution functions (CDF). If two CDF are identical, D is zero, and its maximum possible value is one. D close to zero indicates that two probability distributions are similar. We calculated D for m, r, and ε at all the ground motion intensities (Figure S2). We found the maximum D values for m, r, ε were 0.032, 0.113, and 0.092, respectively, and the minimum values were 0.019, 0.026, and 0.017, indicating a strong resemblance between the two distributions.

We also estimated the differences of mean values from proposed q^* and Riemann sum hazard deaggregation (Figure S2). Note that the deaggregation distributions and optimal IS densities vary for different ground motion levels. Thus, their mean values also vary. We found the maximum relative differences were 2.5 %, 22.6 %, and 18.5 % for m, r, ε, and the minimum differences were 1.7 %, 1.0 %, and 4.3 %, respectively. Also, the maximum absolute differences were 0.14, 3.9 km, and 0.17 for m, r, ε, and the minimum differences were 0.09, 0.7 km, and 0.0009. The maximum relative difference in distance (r) appears for deaggregation distributions at 0.25 g, where the mean distance obtained from hazard deaggregation is 14.3 km and that from q^* is 17.5 km. Given that we are typically interested in distance ranges on the order of tens of kilometers (e.g., 0-15 km, 15-25 km, 25-50 km, etc.) rather than a single value (U. S. Nuclear Regulatory Commission, 2007), this difference is not crucial in determining the controlling earthquake for critical infrastructures. In addition, IS densities still show small K-S D statistics and their mode almost matches each other, as shown in Figure 6.

Example 2 : Fault source

We consider a 50 km-length vertical fault 25 km away from the site (Figure 2). We adopted a finite-dimension rupture model, which results in a distance distribution dependent on magnitude. We found this dependency can diminish the performance of VEGAS AIS because the VEGAS algorithm assumes the independently distributed random variables. For example, for a ground motion intensity of 0.05 g, AIS is slower than conventional MC (Figure 7), while AIS outperformed conventional MC at the same ground motion intensity when point source assumption was made, as seen earlier (Figure 4). However, as the ground motion intensity increases, the computational gap between the two methods becomes smaller rapidly and closes at 0.2g. For higher ground motions, AIS outperforms conventional MC. At 1.0 g, for a 5% standard deviation, AIS, IS, and conventional MC take 0.08, 3.7, and 166 seconds, respectively, i.e., AIS MC is 48 and 2,162 times faster.

We also present the accuracy of AIS PSHA in comparison to the true hazard calculated by Riemann summation (Figure S3). We observed AIS PSHA estimates approximate the true hazard curve within an acceptable range (5 % error) when N

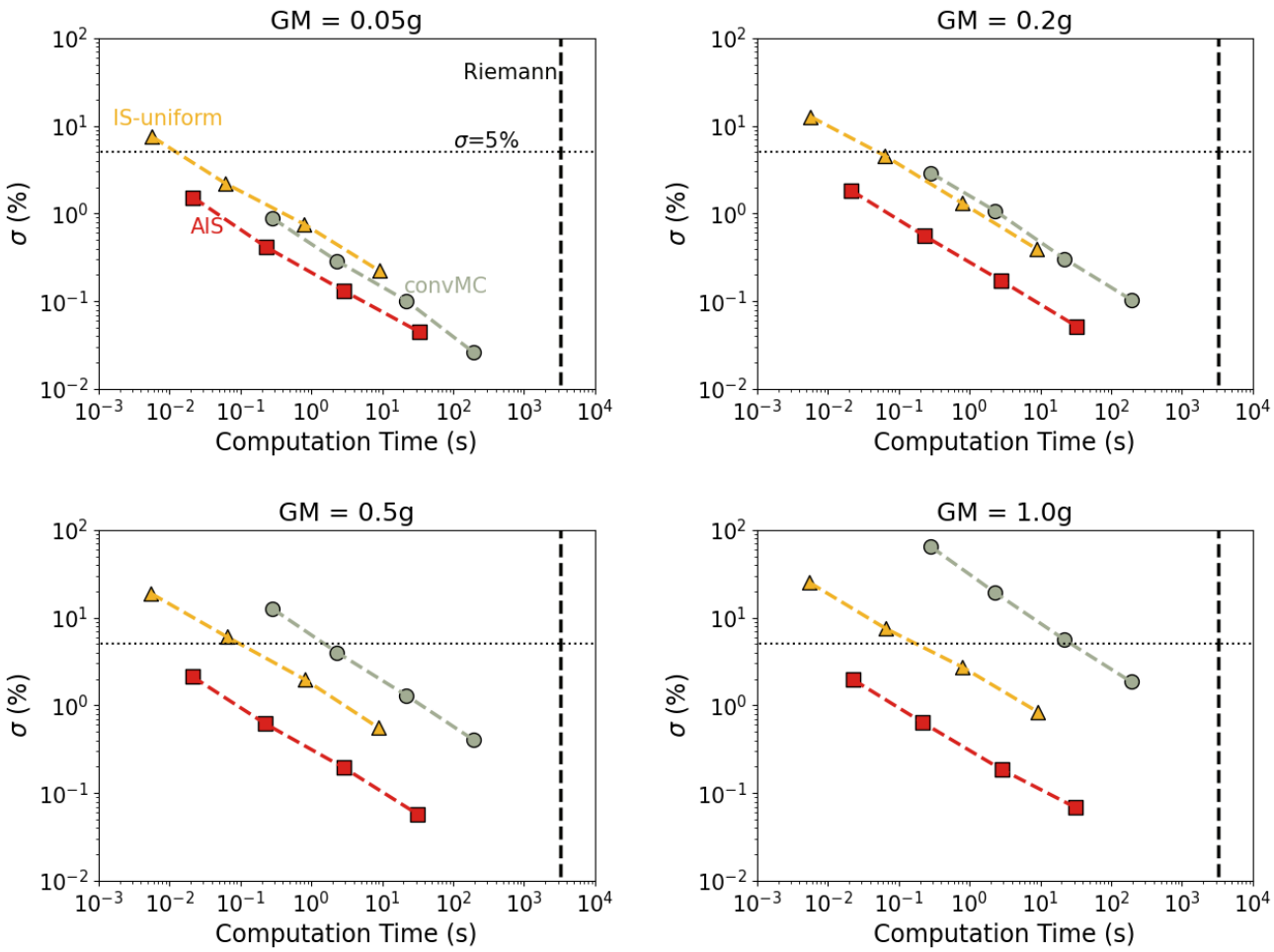

Figure 4. Areal Source Example: Standard deviation of the conventional MC (red), IS MC (blue), AIS MC (black) estimates as functions of computation time at ground motion of 0.05 g, 0.2 g, 0.5 g, and 1.0 g. The computation time required for Riemann summation is presented as green vertical dashed line. At low ground motion, conventional MC outperforms IS and AIS MC, however, as the target ground motion increases, the performance of AIS make a dramatic improvement in terms of both accuracy and computational cost.

is greater than $\sim$ 50,000. Also, note that though we assumed the independently distributed optimal density, AIS PSHA still gives an unbiased hazard estimate that will converge to the true hazard with a sufficiently large number of samples due to the nature of the importance sampling (Figure S3 (d)).

We also compared hazard deaggregation and the iterated IS density (Figure 8) and showed they closely match each other even though the distance distribution depends on the magnitude in this case. It is also noteworthy the iterated IS density can even reproduce complex densities with discontinuities like the large jump within magnitude distribution (for m =6.25) due to the use of a characteristic earthquake occurrence model. The K-S D statistic and mean difference of the two distributions are also presented in Figure S4. The maximum values of D in m, r, and ε are 0.30, 0.68, and 0.13, respectively, and the minimum values are 0.02, 0.31, and 0.04. We found that the largest discrepancies occur in the magnitude distribution, but errors can be considered negligible as the mean magnitude difference is within 6 % error. We also note considerable discrepancies in the distance distribution shape (see K-S D statistics of Figure S4) as curves with concentrated probabilities in narrow ranges are

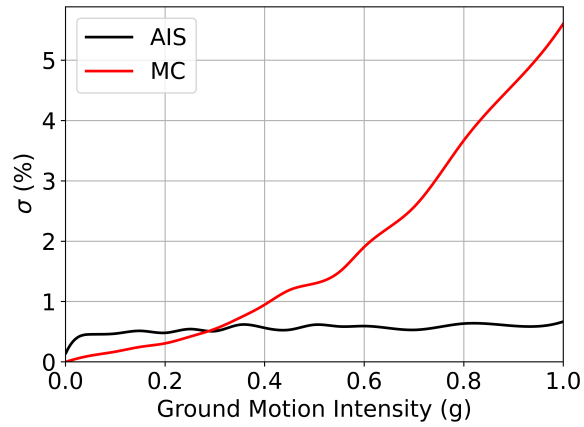

Figure 5. Standard deviation of MC estimates with the ground motion for conventional (red) MC ($N_{\text{convMC}} = 10^7$) and AIS (black) with the similar computation time. The standard deviation of error exponentially increase with the ground motion in conventional MC, while that of AIS remain constant.

harder to estimate for AIS (Figure 8). However, the mean distance is still within a 5 % error range. The ε is generally in good agreement across all ground motion intensities. We observe higher differences at lower ground motion because the mean ε is close to zero. For example, the calculated mean ε at ground motion intensity of 0.01 g is 0.022 and 0.007, which is not a large difference in practice.

The maximum relative differences in mean values were found to be 6.2 %, 5.3 %, and 129 % for m, r, ε, and the minimum differences were 0.006 %, 3.9 %, and 2.2 %, respectively. Note that the 129 % of ε case is corresponding the case where the mean ε is close to zero. The relative difference appears to be slightly higher than previous areal source example, however, the absolute difference is still remain to be significantly small, the maximum absolute differences were 0.36, 1.3 km, and 0.20 for m, r, ε, and the minimum differences were 0.0004, 1.0 km, and 0.07.

Example 3 : Combined sources

In PSHA, we often have multiple seismic sources. We consider one areal and two fault sources around the site to represent this case. This application posits a different mathematical problem than the previous two examples because we must introduce an additional variable to formulate AIS.

First, the probability of earthquake occurrence at ith seismic source can be defined as:

$$P(S = i) = \frac{\nu_i}{\sum_{i=j}^{N_s} \nu_j}$$

, where N_s is the number of seismic sources ($N_s = 3$ in this example), and ν_i is the annual earthquake occurrence rate of ith seismic source. Because the discrete random variables cannot be used in AIS, we define a continuous random variable and its corresponding probability density function as

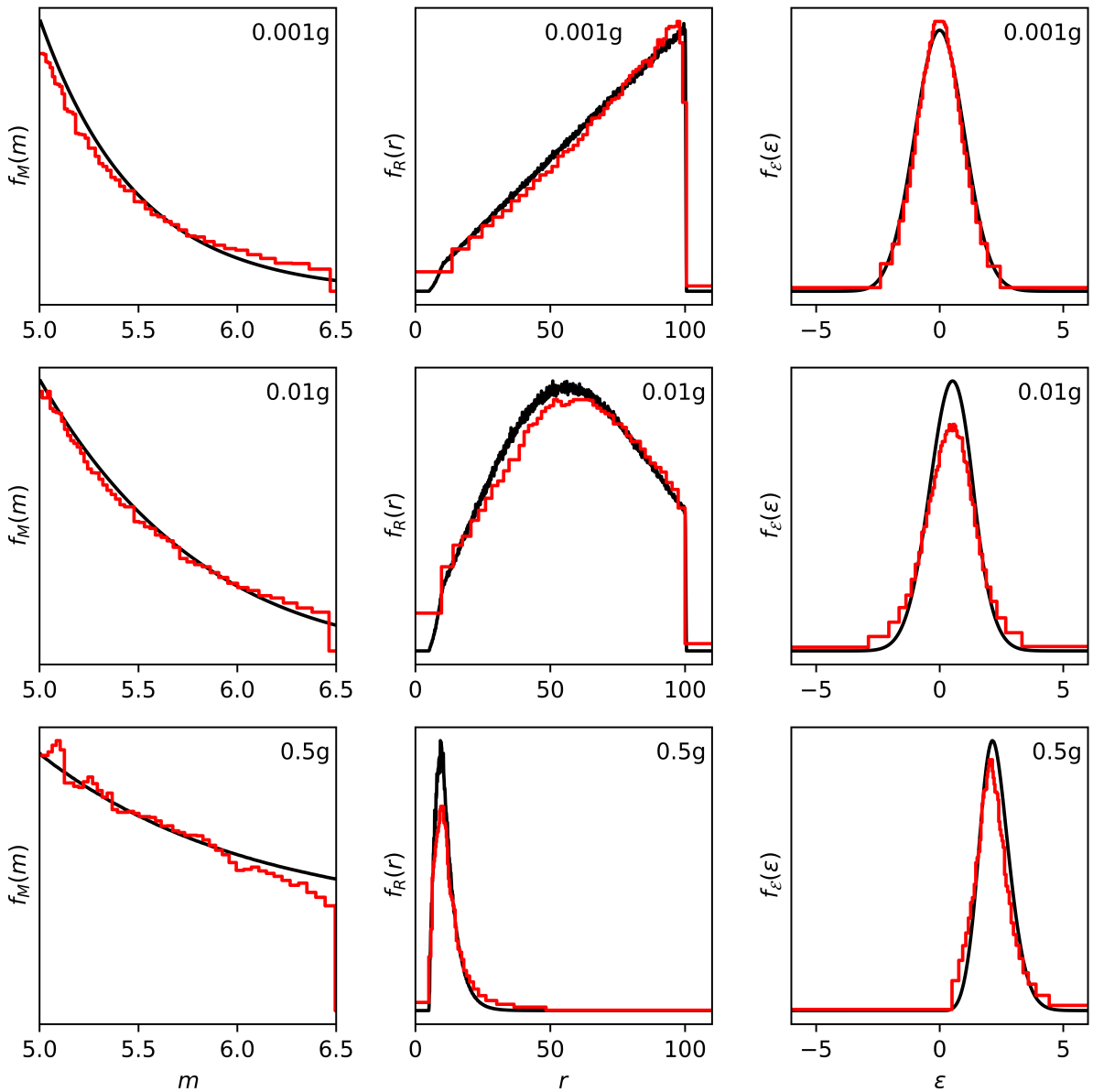

Figure 6. Areal Source Example: The convergence of m, r, and ε iterated IS densities density (red) derived in AIS algorithm (N = 100,000) to marginal distributions of hazard deaggregation (black) at ground motion intensities of 0.001 g, 0.01 g, and 0.5 g.

$$f_S(s) = \begin{cases} 0 & s < 0 \\ P(S=i) & \text{if } i-1 \leq s < i \ (i = 1, 2, ..., N_s), \\ 0 & N_s \leq s \end{cases} \tag{38}$$

$f_S(s)$ is a piece-wise constant function where the heights are proportional to the corresponding sources' earthquake occurrence rates. We introduce $f_S(s)$ into the PSHA integration and obtain

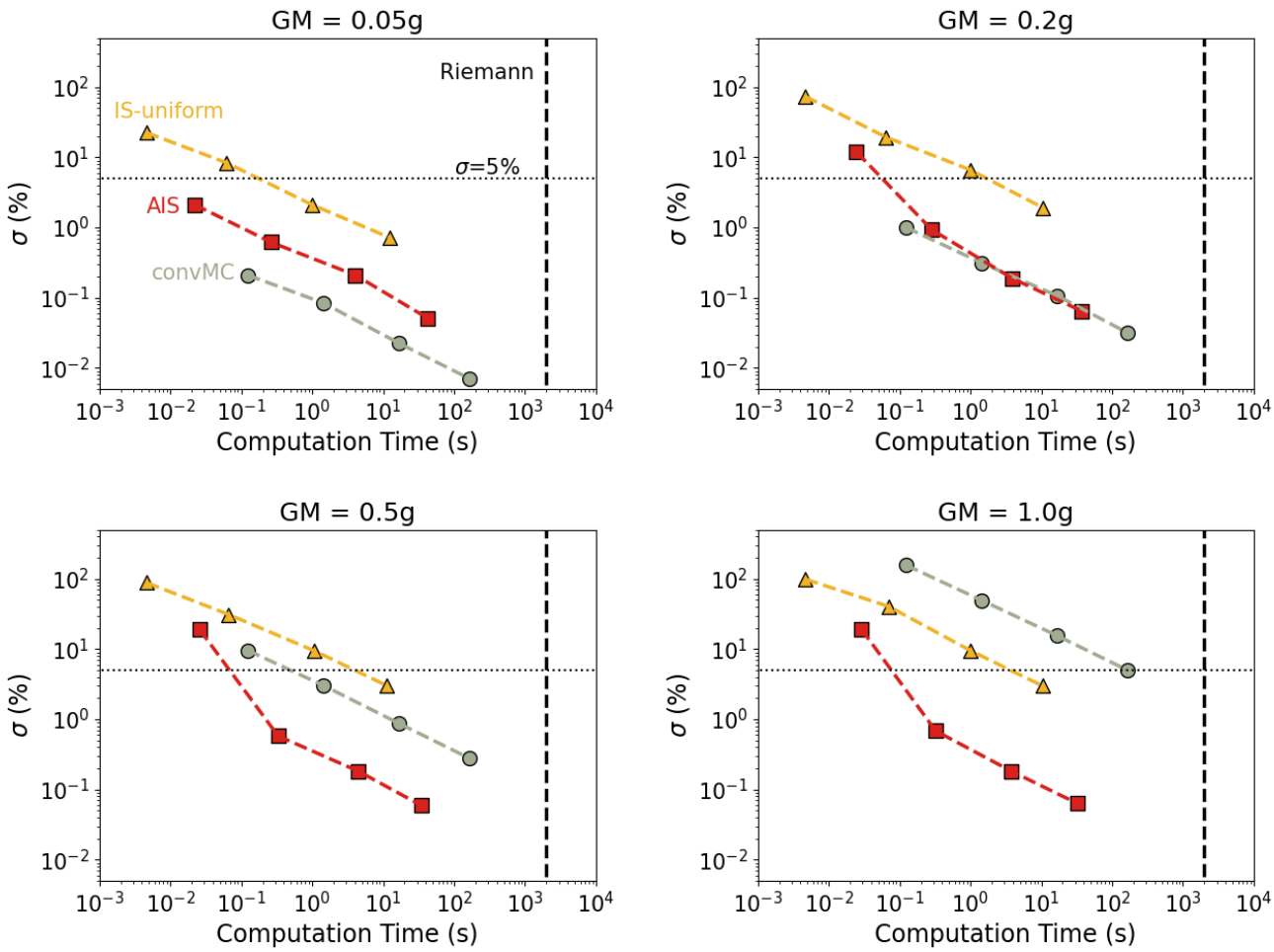

Figure 7. Fault Source Example: Standard deviation of the conventional MC (red), IS MC (blue), AIS MC (black) estimates as functions of computation time at ground motion of 0.05 g, 0.2 g, 0.5 g, and 1.0 g.

$$\Lambda(X > a) = \left(\sum_{i}^{N_{src}} \nu_i \right) \int_s \int_\varepsilon \int_R \int_M I(x > a|s, m, r, \varepsilon) f_{M,R,\varepsilon}(m, r, \varepsilon|s) f_S(s) dm\, dr\, d\varepsilon\, ds \tag{39}$$

, which is the integral version of Eq. (1) and Eq. (5). Solving Eq. (39) posits computational challenges than the single source problems because $f_S(s)$ have the large jumps at $i = 1, 2, \ldots, N_s - 1$. In addition, we introduce additional dependencies in the multi-source case because m and r depend on s, further diminishing the VEGAS algorithm's effectiveness. Thus, we test two AIS PSHA approaches for this case: 1) full AIS approach utilizing the Eq. (39) and 2) partial AIS approach, which is the simple summation of single-source AIS PSHA curves.

We compared the computational performance of MC, IS, and full and partial AIS (Figure 9). Like the previous examples, conventional MC is faster for low ground motions, e.g., 0.05 g, but not high ones. For example, to achieve the σ of 2 % at a ground motion of 0.8 g, conventional MC, IS, and full and partial AIS take 73.8, 10.0, 0.77, and 0.13 seconds, respectively, i.e., partial AIS is the most efficient algorithm, $\sim$ 583, 79, and 6 times faster than conventional MC, uniform IS, and full AIS, respectively.

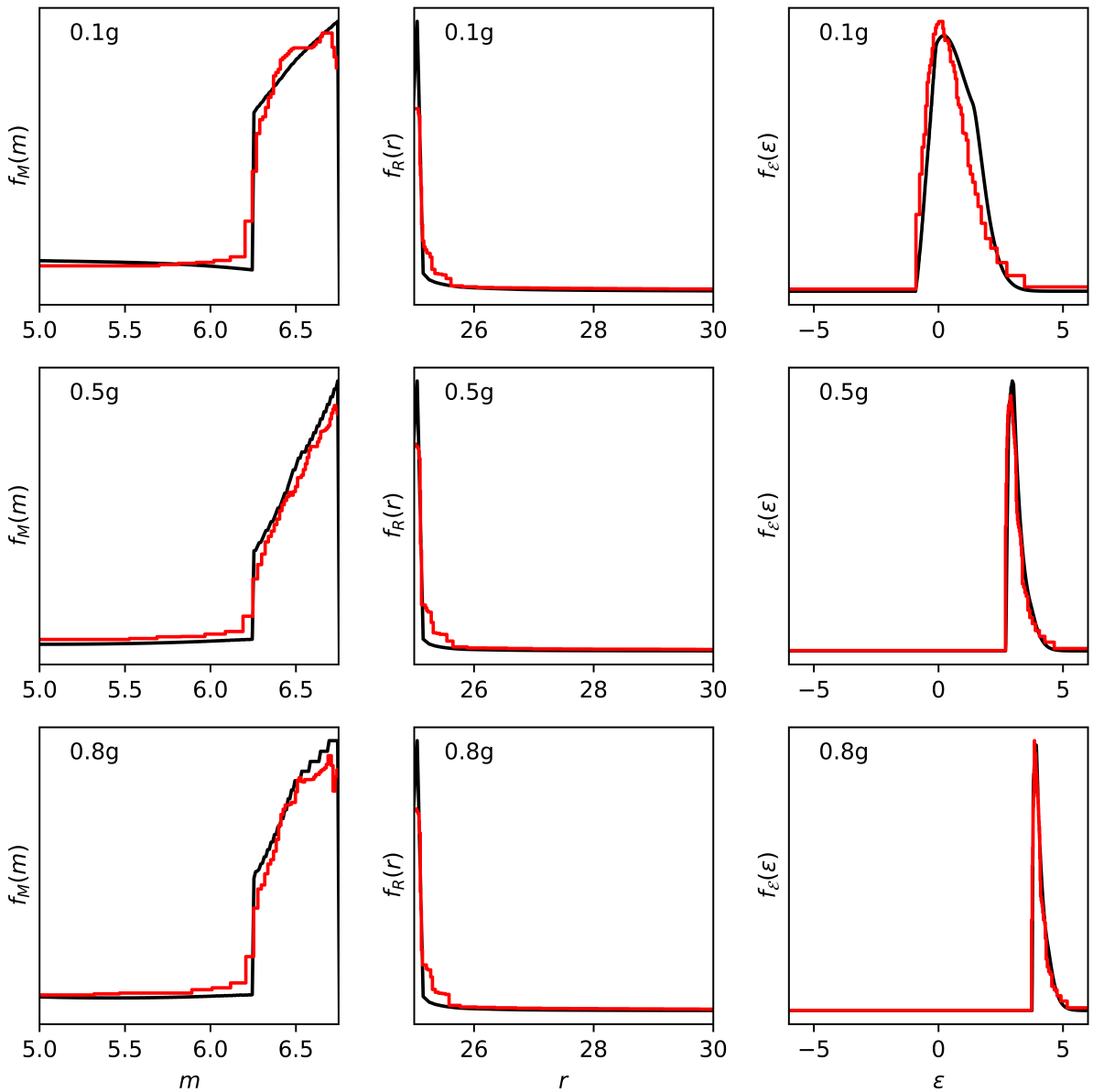

Figure 8. Fault Source Example: The convergence of m, r, and ε iterated IS densities density (red) derived in AIS algorithm (N = 1,000,000) to marginal distributions of hazard deaggregation (black) at ground motion intensities of 0.1 g, 0.5 g, and 0.8 g.

As stated earlier, full AIS is less efficient than partial AIS because of the additional dependencies and jumps introduced by $f_S(s)$. Figures S5 and S6 present the accuracy of both AIS approaches with different sample sizes. The contrast between Figures S5 (c) S6 (c) shows that partial AIS has smaller error than full AIS even with fewer samples (150,000 < 500,000). We don't show the comparison for the partial AIS on ground motions greater than 0.8 g because fault B reaches a numerical instability due to its low exceedance probability ($< 10^{-12}$ /yr) (Figure S5).

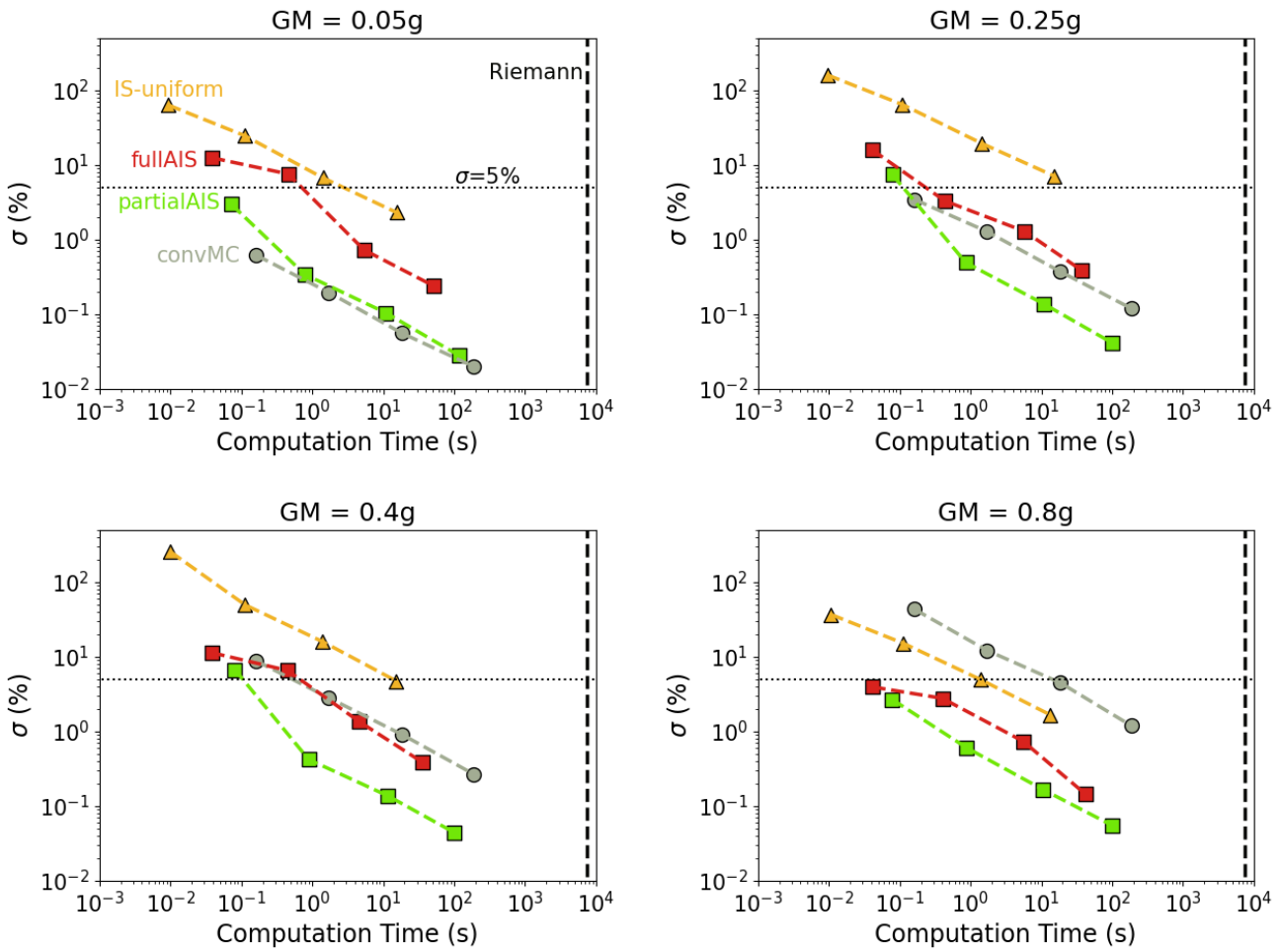

Figure 9. Combined Sources Example: Standard deviation of the conventional MC (red), IS MC (blue), full AIS MC (black), and partial AIS MC (gray) estimates as functions of computation time at ground motion of 0.05 g, 0.25 g, 0.4 g, and 0.8 g

CONCLUSION

We proposed a novel computational framework for PSHA based on an implementation of the VEGAS algorithm. Through comprehensive testbeds, we investigated the computational performance of this new approach, covering widely adopted source types in PSHA practice: 1) areal source, 2) fault source, and 3) combined sources, including both areal and fault sources.

We compared the proposed method to three existing computational frameworks: a) Riemann sum, which has exponentially increasing computational costs for finer grid sizes; b) conventional MC, which requires a substantially long catalog with synthetic earthquakes, particularly for large seismic intensities; and c) importance sampling (IS) with simple (uniform) IS distributions.

Our findings indicate that AIS PSHA outperforms all other computational frameworks. AIS PSHA can dramatically reduce computational times by factors up to $> 10^5$ compared to traditional Riemann summation. AIS PSHA was also 10^3 faster than

the conventional MC while maintaining a 2% standard deviation of error. Additionally, AIS PSHA was up to 70 times faster than IS PSHA, demonstrating that the VEGAS algorithm can approximate optimal IS distributions quickly and well.

We showed that AIS PSHA requires a similar computation time for any ground motion, making its application to larger ground motions with low probability substantially more efficient than conventional MC. In PSHA practice, the computational demands are dominated by large ground motions that need substantially more samples than lower ground motions. In contrast, AIS PSHA only requires a similarly low number of samples for all ground motion levels because it finds optimized IS distributions for each.

We also showed that AIS PSHA finds approximated deaggregation curves at no extra computational cost based on theoretical insights showing that optimal IS densities are equivalent to deaggregation distributions. We show empirically that the hazard deaggregation and iterated IS densities from AIS PSHA are fairly similar by comparing the statistical properties of the two distributions, e.g. K-S D statistics < 0.113 and mean values differences of <4.3 %.

Our study indicated that our AIS PSHA implementation works extremely well for point sources, where the magnitude and distance are independent random variables. We also applied the algorithm to problem employing finite-rupture model, and it still outperformed the pre-existing algorithms up to by factors of $>2{,}000$. It appears to be less effective than point source example because it introduces a dependency between the magnitude and distance. Our implementation of the VEGAS algorithm uses independent random variables for the IS distribution, thus, making it less effective for the finite-rupture case. However, future implementations can also consider other versions of AIS algorithms where the variables are correlated; though computational demands for the AIS iteration would increase.

In the case of combined seismic sources, we proposed two strategies: 1) incorporating the source random variable into the AIS PSHA framework and 2) simple summation of AIS PSHA curves for individual sources. Both strategies outperformed the traditional methods up to by a factor of ~ 580. However, the second strategy was more efficient than the first one by a factor of 6. The first strategy added dependencies to the seismic hazard distribution, making the VEGAS less effective.

In sum, AIS can be applied to any PSHA computation, leading us to expect widespread application of the method. Specifically, we consider our proposed AIS PSHA to be significantly beneficial for large-scale projects that involve numerous logic tree branches and have extreme computational demands.

DATA AND RESOURCES

The source code for computing PSHA using the framework explained in this paper is available at https://github.com/sehoung/ais_psha. Figures were created using Matplotlib (Hunter, 2007) and Microsoft PowerPoint (http://office.microsoft.com; last accessed June 2024).

DECLARATION OF COMPETING INTERESTS

The authors have filed a patent application related to the method discussed in this paper. This application is relevant to the research presented and could represent a potential financial interest.

ACKNOWLEDGMENTS

We acknowledge the financial support by the Tandon School of Engineering at the New York University and the department of Civil and Environmental Engineering at the University of California, Berkeley.

REFERENCES

Alwall, J., R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro (2014). The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *Journal of High Energy Physics* **2014**(7), 1–157.

Arora, P. and L. Ceferino (2023). Probabilistic and machine learning methods for uncertainty quantification in power outage prediction due to extreme events. *Natural Hazards and Earth System Sciences* **23**(5), 1665–1683.

ASCE (2022). Minimum Design Loads for Buildings and Other Structures. ASCE/SEI Standard 7-22. American Society of Civil Engineering, Reston, Virginia.

Assatourians, K. and G. M. Atkinson (2013). EqHaz: An open-source probabilistic seismic-hazard code based on the Monte Carlo simulation approach. *Seismological Research Letters* **84**(3), 516–524.

Assatourians, K. and G. M. Atkinson (2019). Implementation of a smoothed-seismicity algorithm in Monte Carlo PSHA software EQHAZ and implications for localization of hazard in the Western Canada sedimentary basin. *Seismological Research Letters* **90**(3), 1407–1419.

Au, S.-K. and J. L. Beck (2001). Estimation of small failure probabilities in high dimensions by subset simulation. *Probabilistic engineering mechanics* **16**(4), 263–277.

Baker, J., B. Bradley, and P. Stafford (2021). *Seismic Hazard and Risk Analysis.* Cambridge University Press.

Bazzurro, P. and C. A. Cornell (1999). Disaggregation of seismic hazard. *Bulletin of the Seismological Society of America* **89**(2), 501–520.

Bommer, J. J. and N. A. Abrahamson (2006). Why do modern probabilistic seismic-hazard analyses often lead to increased hazard estimates? *Bulletin of the Seismological Society of America* **96**(6), 1967–1977.

Bozorgnia, Y., N. A. Abrahamson, L. A. Atik, T. D. Ancheta, G. M. Atkinson, J. W. Baker, A. Baltay, D. M. Boore, K. W. Campbell, B. S.-J. Chiou, et al. (2014). Nga-west2 research project. *Earthquake Spectra* **30**(3), 973–987.

Bugallo, M. F., V. Elvira, L. Martino, D. Luengo, J. Miguez, and P. M. Djuric (2017). Adaptive Importance Sampling: The past, the present, and the future. *IEEE Signal Processing Magazine* **34**(4), 60–79.

Campolieti, G. and R. Makarov (2007). Pricing path-dependent options on state dependent volatility models with a bessel bridge. *International Journal of Theoretical and Applied Finance* **10**(1), 51–88.

Ceferino, L., J. Mitrani-Reiser, A. Kiremidjian, G. Deierlein, and C. Bambarén (2020). Effective plans for hospital system response to earthquake emergencies. *Nature Communications* **11**(1), 4325.

Christou, V., P. Bocchini, M. J. Miranda, and A. Karamlou (2018). Effective sampling of spatially correlated intensity maps using hazard quantization: Application to seismic events. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering* **4**(1), 04017035.

Coppersmith, K., J. Bommer, K. Hanson, J. Unruh, R. Coppersmith, L. Wolf, R. Youngs, A. Rodriguez-Marek, L. Al Atik, G. Toro, et al. (2014). Hanford Sitewide Probabilistic Seismic Hazard Analysis. PNNL-23361 Pacific Northwest National Laboratory, Richland Washington.

Cornell, C. A. (1968). Engineering seismic risk analysis. *Bulletin of the Seismological Society of America* **58**(5), 1583–1606.

Crowley, H. and J. J. Bommer (2006). Modelling seismic hazard in earthquake loss models with spatially distributed exposure. *Bulletin of Earthquake Engineering* **4**, 249–273.

Dick, J., F. Y. Kuo, and I. H. Sloan (2013). High-dimensional integration: The quasi-Monte Carlo way. *Acta Numerica* **22**, 133–288.

Esteva, L. (1967). Criteria for the construction of spectra for seismic design. In *Third Panamerican Symposium on Structures, Caracas, Venezuela*, Volume 1082.

Field, E. H., T. H. Jordan, and C. A. Cornell (2003). OpenSHA: A developing community-modeling environment for seismic hazard analysis. *Seismological Research Letters* **74**(4), 406–419.

Friedman, J. H. and M. H. Wright (1981). A nested partitioning procedure for numerical multiple integration. *ACM Transactions on Mathematical Software (TOMS)* **7**(1), 76–92.

Garberoglio, G. and A. H. Harvey (2011). Path-integral calculation of the third virial coefficient of quantum gases at low temperatures. *The Journal of chemical physics* **134**(13), 134106.

Goulet, C. A., Y. Bozorgnia, N. Kuehn, L. Al Atik, R. R. Youngs, R. W. Graves, and G. M. Atkinson (2021). Nga-east ground-motion characterization model part i: Summary of products and model development. *Earthquake Spectra* **37**(1), 1231–1282.

Hale, C., N. Abrahamson, and Y. Bozorgnia (2018). Probabilistic Seismic Hazard Analysis Code Verification, PEER report 2018/03, Pacific Earthquake Engineering Research Center, Berkeley, CA.

Hammersley, J. M. and K. W. Morton (1954). Poor Man's Monte Carlo. *Journal of the Royal Statistical Society: Series B (Methodological)* **16**(1), 23–38.

Han, Y. and R. A. Davidson (2012). Probabilistic seismic hazard analysis for spatially distributed infrastructure. *Earthquake Engineering & Structural Dynamics* **41**(15), 2141–2158.

Hunter, J. D. (2007). Matplotlib: A 2d graphics environment. *Computing in science & engineering* **9**(03), 90–95.

Jayaram, N. and J. W. Baker (2010). Efficient sampling and data reduction techniques for probabilistic seismic lifeline risk assessment. *Earthquake Engineering & Structural Dynamics* **39**(10), 1109–1131.

Kappen, H. J. and H. C. Ruiz (2016). Adaptive importance sampling for control and inference. *Journal of Statistical Physics* **162**, 1244–1266.

Kavvada, I., S. Moura, A. Horvath, and N. Abrahamson (2022). Probabilistic seismic hazard analysis for spatially distributed infrastructure considering the correlation of spectral acceleration across spectral periods. *Earthquake Spectra* **38**(2), 1148–1175.

Kennedy, R. P., C. A. Cornell, R. D. Campbell, S. Kaplan, and H. F. Perla (1980). Probabilistic seismic safety study of an existing nuclear power plant. *Nuclear Engineering and Design* **59**(2), 315–338.

Kersevan, B. P. and E. Richter-Was (2013). The monte carlo event generator acermc versions 2.0 to 3.8 with interfaces to pythia 6.4, herwig 6.5 and ariadne 4.1. *Computer Physics Communications* **184**(3), 919–985.

Kiremidjian, A. S., E. Stergiou, and R. Lee (2007). Issues in seismic risk assessment of transportation networks. *Geotechnical, Geological and Earthquake Engineering* **6**, 461–480.

Kolmogorov, A. (1933). Sulla determinazione empirica di una legge didistribuzione. *Giorn Dell'inst Ital Degli Att* **4**, 89–91.

Kroese, D. P., T. Brereton, T. Taimre, and Z. I. Botev (2014). Why the Monte Carlo method is so important today. *Wiley Interdisciplinary Reviews: Computational Statistics* **6**(6), 386–392.

Lepage, G. (1978). A new algorithm for adaptive multidimensional integration. *Journal of Computational Physics* **27**(2), 192–203.

Lepage, G. P. (2021). Adaptive multidimensional integration: vegas enhanced. *Journal of Computational Physics* **439**, 110386.

Manzour, H., R. A. Davidson, N. Horspool, and L. K. Nozick (2016). Seismic hazard and loss analysis for spatially distributed infrastructure in Christchurch, New Zealand. *Earthquake Spectra* **32**(2), 697–712.

Marzocchi, W. and T. H. Jordan (2017). A unified probabilistic framework for seismic hazard analysis. *Bulletin of the Seismological Society of America* **107**(6), 2738–2744.

McGuire, R. K. (1976). FORTRAN computer program for seismic risk analysis. USGS Open-File Report 76-67.

McGuire, R. K. (1995). Probabilistic seismic hazard analysis and design earthquakes: closing the loop. *Bulletin of the Seismological Society of America* **85**(5), 1275–1284.

McGuire, R. K. (2008). Probabilistic seismic hazard analysis: Early history. *Earthquake Engineering and Structural Dynamics* **37**(3), 329–338.

Musson, R. M. W. (2000). The use of Monte Carlo simulations for seismic hazard assessment in the UK. *Annals of Geophysics* **43**(1), 1–9.

Nieto, M. R. and E. Ruiz (2016). Frontiers in var forecasting and backtesting. *International Journal of Forecasting* **32**(2), 475–501.

Novak, E. and K. Ritter (1997). The curse of dimension and a universal method for numerical integration. In *Multivariate approximation and splines*, pp. 177–187. Springer.

Ordaz, M., F. Martinelli, V. D'Amico, and C. Meletti (2013). CRISIS2008: A flexible tool to perform probabilistic seismic hazard assessment. *Seismological Research Letters* **84**(3), 495–504.

Pagani, M., D. Monelli, G. Weatherill, L. Danciu, H. Crowley, V. Silva, P. Henshaw, L. Butler, M. Nastasi, L. Panzeri, M. Simionato, and D. Vigano (2014). Openquake engine: An open hazard (and risk) software for the global earthquake model. *Seismological Research Letters* **85**(3), 692–702.

Papadopoulos, A. N. and P. Bazzurro (2021). Exploring probabilistic seismic risk assessment accounting for seismicity clustering and damage accumulation: Part II. Risk analysis. *Earthquake Spectra* **37**(1), 386–408.

Philippe, A. and C. P. Robert (2001). Riemann sums for MCMC estimation and convergence monitoring. *Statistics and Computing* **11**(2), 103–115.

Press, W. H. and G. R. Farrar (1990). Recursive stratified sampling for multidimensional monte carlo integration. *Computers in Physics* **4**(2), 190–195.

Rahimi, H. and M. Mahsuli (2019). Structural reliability approach to analysis of probabilistic seismic hazard and its sensitivities. *Bulletin of Earthquake Engineering* **17**(3), 1331–1359.

Ray, J., Y. Marzouk, and H. Najm (2011). A bayesian approach for estimating bioterror attacks from patient data. *Statistics in Medicine* **30**(2), 101–126.

Robert, C. P., G. Casella, and G. Casella (1999). *Monte Carlo statistical methods*, Volume 2. Springer.

Rubinstein, R. Y. (1997). Optimization of computer simulation models with rare events. *European Journal of Operational Research* **99**(1), 89–112.

Rubinstein, R. Y. and D. P. Kroese (2004). *The cross-entropy method: a unified approach to combinatorial optimization, Monte-Carlo simulation, and machine learning*. Springer.

Sadigh, K., C. Y. Chang, J. A. Egan, F. Makdisi, and R. R. Youngs (1997). Attenuation relationships for shallow crustal earthquakes based on California strong motion data. *Seismological Research Letters* **68**(1), 180–189.

Sanders, J. L. (2014). Probabilistic model for constraining the galactic potential using tidal streams. *Monthly Notices of the Royal Astronomical Society* **443**(1), 423–431.

Silva, V., D. Amo-Oduro, A. Calderon, C. Costa, J. Dabbeek, V. Despotaki, L. Martins, M. Pagani, A. Rao, M. Simionato, D. Viganò, C. Yepes-Estrada, A. Acevedo, H. Crowley, N. Horspool, K. Jaiswal, M. Journeay, and M. Pittore (2020). Development of a global seismic risk model. *Earthquake Spectra* **36**, 372–394.

Thomas, P., I. Wong, and N. A. Abrahamson (2010). Verification of probabilistic seismic hazard analysis computer programs. PEER Report 2010/106, Pacific Earthquake Engineering Research Center, Berkeley.

Tokdar, S. T. and R. E. Kass (2010). Importance sampling: A review. *Wiley Interdisciplinary Reviews: Computational Statistics* **2**(1), 54–60.

U. S. Nuclear Regulatory Commission (2007). A Performance Based Approach to Define Site Specific Ground Motion. Regulatory Guide 1.208, USNRC.

Youngs, R. R. and K. J. Coppersmith (1985). Implications of fault slip rates and earthquake recurrence models to probabilistic seismic hazard estimates. *Bulletin of the Seismological society of America* **75**(4), 939–964.

APPENDIX A: VARIANCE OF MONTE-CARLO PSHA ESTIMATE

The second moment of $\hat{\lambda}$, $E[\hat{\lambda}^2]$, can be expressed as:

$$
\begin{aligned}
E[\hat{\lambda}^2] &= E\left[\left(\frac{\nu}{N}\sum_i^N I(X_i > a)\right)^2\right] \\
&= \frac{\nu^2}{N^2} E\left[\left(\sum_i^N I(X_i > a)\right)^2\right] \\
&= \frac{\nu^2}{N^2} E\left[\left(\sum_i^N I_i\right)^2\right]
\end{aligned}
$$

Note that $I(X_i > a)$ is denoted as I_i for simplicity in the following derivation. The above equation expands as follows:

$$
\begin{aligned}
E[\hat{\lambda}^2] &= \frac{\nu^2}{N^2} E\left[(I_1 + I_2 + \cdots + I_N)^2\right] \\
&= \frac{\nu^2}{N^2} E[I_1^2 + I_2^2 + \cdots + I_N^2 \; + \; I_1 I_2 + I_1 I_3 + I_1 I_4 + \cdots + I_1 I_N \\
&\qquad + I_2 I_1 + I_2 I_3 + I_2 I_4 + \cdots + I_2 I_N \\
&\qquad + I_3 I_1 + I_3 I_2 + I_3 I_4 + \cdots + I_3 I_N \\
&\qquad \vdots \\
&\qquad + I_N I_1 + I_N I_2 + I_N I_3 + \cdots + I_N I_{N-1}] \\
&= \frac{\nu^2}{N^2} E\left[\sum_i^N I_i^2 + \sum_{i \neq j}^{N^2-N} I_i I_j\right]
\end{aligned}
$$

Here, I_i^2 is the same as I_i since I_i^2 also takes 1 if and only if X_i is greater than a. Also, the expectation operator can go inside the sum terms as it is a linear operator

$$
E[\hat{\lambda}^2] = \frac{\nu}{N^2} \sum_i^N \nu E[I_i] + \frac{1}{N^2} \sum_{i \neq j}^{N^2-N} \nu^2 E[I_i I_j]
$$

Here, $\nu E[I_i]$ is equal to λ. Also, $I_i \; I_j$ takes 1 if and only if both X_i and X_j exceeds a. Given that the ground motion samples (X_i and X_j) are extracted independently, $\nu^2 E[I_i I_j]$ is identical to $\nu E[I_i] \nu E[I_j]$, which is λ^2:

$$
\begin{aligned}
E[\hat{\lambda}^2] &= \frac{\nu}{N^2} \sum_i^N \lambda + \frac{1}{N^2} \sum_{i \neq j}^{N^2-N} \lambda^2 \\
&= \frac{\nu}{N^2} N\lambda + \frac{1}{N^2}(N^2 - N)\lambda^2 \\
&= \frac{\nu\lambda - \lambda^2}{N} + \lambda^2
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\mathrm{VAR}[\hat{\lambda}] &= E[\hat{\lambda}^2] - E[\hat{\lambda}]^2 \\
&= E\left[\hat{\lambda}^2\right] - \lambda^2 \\
&= \frac{\nu\lambda - \lambda^2}{N} + \lambda^2 - \lambda^2 \\
&= \frac{\nu\lambda - \lambda^2}{N}
\end{aligned}
$$

Note that $E[\hat{\lambda}]^2$ is the same as λ^2 since $\hat{\lambda}$ is unbiased estimate.

APPENDIX B:VARIANCE OF IMPORTANCE SAMPLING PSHA ESTIMATE

$$
\begin{aligned}
E[\hat{\lambda}^2] &= E\left[\left(\frac{\nu}{N}\sum_i^N \frac{I(X_i > a) f_{M,R,\mathcal{E}}(M_i, R_i, \mathcal{E}_i)}{q_{M,R,\mathcal{E}}(M_i, R_i, \mathcal{E}_i)}\right)^2\right] \\
&= \frac{\nu^2}{N^2} E\left[\left(\sum_i^N \frac{I_i f_i}{q_i}\right)^2\right]
\end{aligned}
$$

Note that $I(X_i > a)$, $f_{M,R,\mathcal{E}}(M_i, R_i, \mathcal{E}_i)$, and $q_{M,R,\mathcal{E}}(M_i, R_i, \mathcal{E}_i)$ are denoted as I_i, f_i, and q_i for simplicity in the following derivation. As derived in appendix A, the above equation expands as follows:

$$
\begin{aligned}
E[\hat{\lambda}^2] &= \frac{\nu^2}{N^2} E\left[\sum_i^N \left(\frac{I_i f_i}{q_i}\right)^2 + \sum_{i \neq j}^{N^2-N} \left(\frac{I_i f_i}{q_i}\right)\left(\frac{I_j f_j}{q_j}\right)\right] \\
&= \frac{\nu^2}{N^2} \sum_i^N E\left[\left(\frac{I_i f_i}{q_i}\right)^2\right] + \frac{\nu^2}{N^2} \sum_{i \neq j}^{N^2-N} E\left[\left(\frac{I_i f_i}{q_i}\right)\left(\frac{I_j f_j}{q_j}\right)\right] \\
&= \frac{\nu^2}{N} \frac{1}{N} \sum_i^N E\left[\left(\frac{I_i f_i}{q_i}\right)^2\right] + \frac{1}{N^2} \sum_{i \neq j}^{N^2-N} \nu E\left[\left(\frac{I_i f_i}{q_i}\right)\right] \nu E\left[\left(\frac{I_j f_j}{q_j}\right)\right] \\
&= \frac{\nu^2}{N} E\left[E\left[\left(\frac{I_i f_i}{q_i}\right)^2\right]\right] + \frac{1}{N^2}(N^2 - N)\lambda^2 \\
&= \frac{\nu^2}{N} E\left[\left(\frac{I_i f_i}{q_i}\right)^2\right] - \frac{1}{N}\lambda^2 + \lambda^2
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\mathrm{VAR}[\hat{\lambda}] &= E[\hat{\lambda}^2] - E[\hat{\lambda}]^2 \\
&= \frac{\nu^2}{N} E\left[\left(\frac{I_i f_i}{q_i}\right)^2\right] - \frac{1}{N}\lambda^2 + \lambda^2 - \lambda^2 \\
&= \frac{1}{N}\left(\nu^2 E\left[\left(\frac{I_i f_i}{q_i}\right)^2\right] - \lambda^2\right)
\end{aligned}
$$

Manuscript Received 00 Month 0000