ANS M&C 2021 - The International Conference on Mathematics and Computational Methods Applied to Nuclear Science and Engineering · Raleigh, North Carolina · April 11–15, 2021

SAND2020-14030C

COMPUTATION OF SOBOL'INDICES USING EMBEDDED VARIANCE DECONVOLUTION

James M. Petticrew[1], **Aaron J. Olson**[2]

[1]Atomic Weapons Establishment
Aldermaston, Reading, RG7 4PR, UK

[2]Sandia National Laboratories
Albuquerque, NM 87185, USA

james.petticrew@awe.co.uk, aolson@sandia.gov

ABSTRACT

Sobol sensitivity indices (SI) provide robust and accurate measures of how much uncertainty in output quantities is caused by different uncertain input parameters. These allow analysts to prioritize future work to either reduce or better quantify the effects of the most important uncertain parameters. One of the most common approaches to computing SI requires Monte Carlo (MC) sampling of uncertain parameters and full physics code runs to compute the response for each of these samples. In the case that the physics code is a MC radiation transport code, this traditional approach to computing SI presents a workflow in which the MC transport calculation must be sufficiently resolved for each MC uncertain parameter sample. This process can be prohibitively expensive, especially since thousands or more particle histories are often required on each of thousands or so uncertain parameter samples. We propose a process for computing SI in which only a few MC radiation transport histories are simulated before sampling new uncertain parameter values. We use Embedded Variance Deconvolution (EVADE) to parse the desired parametric variance from the MC transport variance on each uncertain parameter sample. To provide a relevant benchmark, we propose a new radiation transport benchmark problem and derive analytic solutions for its outputs including SI. The new EVADE-based approach is found to converge with MC convergence behavior and be at least an order of magnitude more precise for the same computational cost than the traditional approach for several SI on our test problem.

KEYWORDS: Uncertainty quantification; Sobol sensitivity indices; Monte Carlo; Embedded Variance Deconvolution

1. INTRODUCTION

Global sensitivity analysis (GSA) is the study of how much of the uncertainty in output quantities is caused by different sources of uncertainty in input parameters. GSA is often, but not exclusively, used for systems modelled using physics codes. In applications such as nuclear reactor design and medical physics, GSA can enable decision-makers to prioritize future work to either reduce or better quantify the effects of the most important uncertain input parameters. Sobol sensitivity indices

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

(SI) are variance-based sensitivity measures that provide robust and accurate sensitivity estimates for complex nonlinear systems. Introduced in 1993 alongside a proposed Monte Carlo algorithm for estimating the indices [1], SI have become one of the most popular measures of global sensitivity. Numerous improvements on Sobol's original algorithm have been made, primarily seeking to improve the efficiency of the estimator and thus reduce the number of physics code runs required to estimate SI to a given level of precision [2][3]. However, these approaches still require many code runs, often thousands, to produce reliable results. Consequently, if the physics code is a Monte Carlo (MC) radiation transport simulator, performing enough code runs to compute precise SI estimates can be infeasible, since many thousands, or more, particle histories are often required per MC transport code run. There are also methods for computing SI based on surrogate models, for example using the coefficients from a polynomial chaos expansion [4] or leveraging a Gaussian process emulator [5]. In this paper, we develop a method analogous to the Sobol method and thus do not consider methods based on surrogate models any further.

We sidestep the problem of requiring many MC transport code runs by computing SI leveraging variance deconvolution and frequent, embedded resampling of uncertain parameters every few particle histories. To accomplish this, we used Embedded Variance Deconvolution (EVADE) [6] to parse out parametric variance from the variance due to the MC radiation transport code. We applied the EVADE algorithm, which was originally developed for estimating the variance caused by all the uncertain input parameters, to the estimation of conditional variances, the variance caused by a subset of the uncertain input parameters. We apply our method to MC radiation transport, but it is applicable any time we employ a MC solver to calculate the response for different values of uncertain input parameters.

In order to properly quantify and propogate the uncertainty associated with our SI estimate, and thus understand its accuracy and precision, we needed expressions for the bias and variance of our SI estimator. However, as the SI estimator is a ratio of two quantities estimated separately using MC, no analytic term for its bias and variance exist, and so we expound on a method in the literature [7] in order to estimate for them. In order to benchmark our method, we required a benchmark problem with analytic expressions for SI, and we required those SI to be sufficiently interesting to provide a robust test for our new method. However, we were unable to find such a benchmark problem in the literature, and so we adapted an existing problem [6] to meet these requirements. We then used this problem to benchmark both the traditional numerical method and our EVADE-based method, choosing intuitive parameters for both. Finally, we produced numerical results based on the new benchmark problem to allow us to compare the precision of both the traditional and EVADE-based method as a function of computational cost.

2. ESTIMATION OF CONDITIONAL VARIANCES USING EVADE

SI are useful measures of how uncertainty in the output of a MC transport code can be attributed to each of the code's uncertain parameters and interactions between these parameters. Given a transport code with output of the form $Y = f(x_1, x_2, ..., x_d)$, the first order effect for uncertain parameter $x_i \, \forall \, i \in \{1, \ldots, d\}$ is $V_{x_i}(E_{\mathbf{x}_{\sim \mathbf{i}}}(Y|x_i))$, where $V_{x_i}(\cdot)$ denotes the variance of an argument taken by integrating over the uncertain parameter x_i and $E_{\mathbf{x}_{\sim \mathbf{i}}}(\cdot)$ denotes the mean of an argument taken by integrating over all of the uncertain parameters except x_i. Hence, the main effect SI for uncertain parameter x_i is [2]

$$S_i = \frac{V_{x_i}(E_{\mathbf{x}_{\sim \mathbf{i}}}(Y|x_i))}{V(Y)} = 1 - \frac{E_{x_i}(V_{\mathbf{x}_{\sim \mathbf{i}}}(Y|x_i))}{V(Y)}. \quad (1)$$

S_i can be interpreted as the expected fractional reduction in parametric variance that would be obtained if x_i could be fixed [2].

If $E_{\mathbf{x}_{\sim \mathbf{i}}}(\cdot)$ denotes the mean of an argument taken by integrating over all the uncertain parameters except x_i and $V_{x_i}(\cdot)$ denotes the variance of an argument taken by integrating over the uncertain parameter x_i, then the total effect SI for x_i is [2]

$$S_{T_i} = \frac{E_{\mathbf{x}_{\sim \mathbf{i}}}(V_{x_i}(Y|\mathbf{x}_{\sim \mathbf{i}}))}{V(Y)} = 1 - \frac{V_{\mathbf{x}_{\sim \mathbf{i}}}(E_{x_i}(Y|\mathbf{x}_{\sim \mathbf{i}}))}{V(Y)}, \quad (2)$$

where Eqs. (1) and (2) follow from the the law of total variance: $V(Y) = E\left(V(Y|x)\right) + V\left(E(Y|x)\right)$.

As in [2] and [3], the indices are traditionally estimated from many hundreds or thousands of physics code runs. For MC radiation transport, this means many hundreds or thousands of MC transport code runs, where each run needs to be well resolved, requiring many thousands or more MC particle histories per code run. The computational expense can be prohibitive in practical cases. Consequently, to improve computational efficiency, we needed to employ a better method for estimating SI for MC radiation transport codes with uncertain input parameters.

Recent work [6] at Sandia National Laboratories (SNL) has demonstrated a new approach for estimating parametric variance for MC radiation transport codes with uncertain parameters, yielding significant computational savings. This approach, EVADE, works by simultaneously computing estimates for the average distributional variance due to the MC transport code V_{MC}, and the total distributional variance due to both the transport code and the uncertain parameters V_{tot}, across MC particle histories. However, in order to estimate SI, we not only wish to solve for the parametric variance due to the uncertain parameters $V_P(Y)$ but for the conditional variance for fixed values of x_i $V_P(Y|x_i)$. Hence an adapted approach was required. Rather than using EVADE to solve for $V_P(Y)$, we keep x_i constant, and solve for $V_P(Y|x_i)$ by sampling from the joint PDF of $\mathbf{x}_{\sim \mathbf{i}}$. We then use EVADE to estimate conditional total variance and Monte Carlo variance separately and, hence, solve for conditional parametric variance $V_P(Y|x_i)$. Finally, we compute many estimates of $V_P(Y|x_i)$ to estimate $E(V_P(Y|x_i))$. Given this simple adaptation, for a given number of MC particle histories, we were able to achieve increases in precision of an order of magnitude or more for four of the six tested SI, when compared to traditional SI approaches.

We describe a scheme to estimate S_1 (see Equation (1)), the main effect SI for uncertain input parameter x_1 for a Monte Carlo radiation transport code which has three uncertain parameters $X = \{x_1, x_2, x_3\}$ distributed uniformly on the unit cube. The scheme generalizes to any number of dimensions. We use the notation $Y(\cdot)$ to represent one MC particle history.

With Figure 1 as a visual aid, we describe a sequence of steps to estimate $E(V(Y|x_1))$:

1. Generate a random sample $\xi_{\mathbf{1}} = (x_{1_{\xi_1}}, x_{2_{\xi_1}}, x_{3_{\xi_1}})$ of uncertain input parameters.

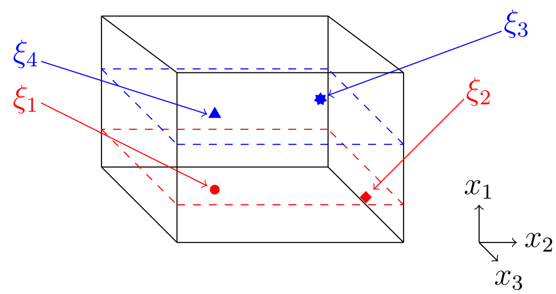

Figure 1: Illustration of sampling scheme on unit cube for estimation of Sobol'indices

2. Evaluate $Y(\xi_\mathbf{1})$ N times, $Y(\xi_\mathbf{1})_1, Y(\xi_\mathbf{1})_2..., Y(\xi_\mathbf{1})_N$. Due to the stochasticity of $Y(\cdot)$, $Y(\xi_\mathbf{1})_i \neq Y(\xi_\mathbf{1})_j$ in general.

3. Tally $Y(\xi_\mathbf{1})_1$ and $(Y(\xi_\mathbf{1})_1)^2$, and use the standard formula for sample variance to compute and tally the distributional Monte Carlo variance

$$V_{MC_1} = \frac{N}{N-1}\left(\frac{1}{N}\sum_{i=1}^{N}(Y(\xi_\mathbf{1})_i)^2 - \left(\frac{1}{N}\sum_{i=1}^{N}Y(\xi_\mathbf{1})_i\right)^2\right). \tag{3}$$

4. Generate a random sample $\xi_\mathbf{2} = (x_{1\xi_1}, x_{2\xi_2}, x_{3\xi_2})$ of uncertain input parameters on the hyperplane $x_1 = x_{1\xi_1}$.

5. Repeat step (2) for $\xi_\mathbf{2}$, computing and tallying a second estimate of the distributional Monte Carlo variance V_{MC_2} and tallying the first and second moments of $Y(\xi_\mathbf{2})_1$ and $(Y(\xi_\mathbf{2})_1)^2$. Generate M samples on the hyperplane $x_1 = x_{1\xi_\mathbf{1}}$ and keep tallies $\sum_{i=1}^{M} Y(\xi_\mathbf{i})_1$, $\sum_{i=1}^{M}(Y(\xi_\mathbf{i})_1)^2$, and $\sum_{i=1}^{M} V_{MC_i}$.

6. Use the standard formula for sample variance to compute the conditional total variance

$$V_{tot}(Y|x_1) = \frac{M}{M-1}\left(\frac{1}{M}\sum_{j=1}^{M}(Y(\xi_\mathbf{j})_1)^2 - \left(\frac{1}{M}\sum_{j=1}^{M}(Y(\xi_\mathbf{j})_1)\right)^2\right). \tag{4}$$

7. Use EVADE to compute a first estimate of the conditional parametric variance

$$V_{P_1}(Y|x_1) = V_{tot}(Y|x_1) - \frac{1}{M}\sum_{j=1}^{M}V_{MC_j}. \tag{5}$$

8. Generate a random sample $\xi_\mathbf{3} = (x_{1_{\xi_3}}, x_{2_{\xi_3}}, x_{3_{\xi_3}})$ of input parameters.

9. Repeat steps 2-6 on the hyperplane $x_1 = x_{1\xi_3}$ to compute a second estimate of the conditional parametric variance $V_{P_2}(Y|x_1)$. Repeat this process p times. For each hyperplane, tally an estimate of the conditional parametric variance, $\sum_{k=1}^{p} V_{P_k}(Y|x_{1_k})$.

10. Compute an estimate of the average conditional parametric variance by averaging over all the conditional parametric variance estimates

$$E(V(Y|x_1)) = \frac{1}{p}\sum_{k=1}^{p} V_{P_k}(Y|x_{1_k}). \tag{6}$$

In the traditional SI approach, it is possible to estimate $V_P(Y)$, the denominators in Eqs. (1) and (2), for "free", as it is easy to accurately estimate the variance over all the well-resolved MC transport code runs. However, using our EVADE-based method, we can also get a "free" estimate of $V_P(Y)$. Since the first random sample of input parameters on each hyperplane (realizations ξ_1 and ξ_3 from steps 1 and 8) are independently and identically distributed (i.i.d.) samples from the joint probability density function of the uncertain input parameters X, we can compute an estimate for $V_P(Y)$ by tallying over these samples. However, since we are estimating the numerator and denominator for the SI expression using EVADE with uncertainty estimates, we must propogate these uncertainties through to the estimate of the SI. We do this in order to show that our proposed estimator converges with the expected MC behavior in the asymptotic limit and, thus, provides an accurate and efficient way of estimating SI.

We compute SI by generating a sequence of i.i.d. estimates for the average conditional parametric variances, the numerators in Eqs. (1) and (2). We then compute a sequence of i.i.d. estimates for the total parametric variance, the denominators in Eqs. (1) and (2). For brevity, we denote these two sequences as $X_1, X_2, ..., X_N$ and $Y_1, Y_2, ..., Y_N$. If the population from which the sequence $X_1, ..., X_N$ is generated has finite mean μ_X and variance σ_X^2, then, by the weak law of large numbers [8], the arithmetic mean of the sequence $\bar{X} = \sum_{i=1}^{N} X_i$ converges in probability to μ_X, that is $\lim_{N\to\infty} P(|\bar{X} - \mu_x| \geq \epsilon) \to 0 \;\forall \epsilon > 0$. We write $\bar{X} \xrightarrow{P} \mu_X$. Similarly, $\bar{Y} \xrightarrow{P} \mu_Y$. Since $\bar{X}$ and $\bar{Y}$ converge in probability, then, by Theorem 2.7 in Owen [7], if we have a sequence of i.i.d. joint random variables (i.e. joint uncertain parameters) $Z_1, ..., Z_N = (X_1, Y_1), ..., (X_N, Y_N)$ then, $\bar{Z} = (\bar{X}, \bar{Y}) \xrightarrow{p} Z = (\mu_X, \mu_Y)$. Note that we require that (X_i, Y_i) and (X_j, Y_j) be independent, but X_i and Y_i need not be independent. Finally, by the Continuous Mapping Theorem [8], $\frac{\bar{X}}{\bar{Y}} \xrightarrow{p} \frac{\mu_X}{\mu_Y}$. This argument serves to convince us that our proposed estimator converges with the expected MC behavior in the asymptotic limit. We can also approximate the bias and variance of our estimator from the sequences $X_1, ..., X_N$ and $Y_1, Y_2, ..., Y_N$. From Owen [7], we use the following approximation for the variance component of the statistical uncertainty for the MC estimator $\frac{\bar{X}}{\bar{Y}}$:

$$\widehat{Var}\left(\frac{\bar{X}}{\bar{Y}}\right) = \frac{1}{N(N-1)\bar{Y}^2}\sum_{i=1}^{N}\left(X_i - \frac{\bar{X}}{\bar{Y}}Y_i\right)^2. \tag{7}$$

For the bias component, we have $Bias\left(\frac{\bar{X}}{\bar{Y}}\right) = \frac{1}{N\mu_Y^2}(\frac{\mu_X}{\mu_Y}\sigma_Y^2 - \rho_{XY}\sigma_X\sigma_Y)$, where ρ_{XY} is the correlation coefficient between X and Y. As the variance component in Eq. (7) is $\mathcal{O}(\frac{1}{\sqrt{N}})$ and the bias component is $\mathcal{O}(\frac{1}{N})$, the bias is negligible compared to the variance for even moderately-sized N.

In order to rigorously study convergence behavior, we seek to demonstrate the performance of the new method on a MC radiation transport problem for which we have analytic expressions for SI.

We adapt the benchmark problem used in [6]. The adaption we propose causes the main effect SI and the total effect SI to be different, which makes the problem interesting as a test for our new method. We illustrate the benchmark problem in Figure 2 and derive analytic expressions for its main and total effect SIs.

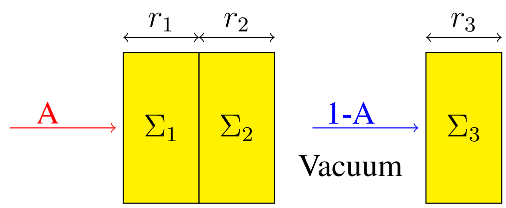

Figure 2: Depiction of analytic test problem

We are interested in the transmission T through a system of three slabs of thicknesses r_1, r_2, r_3 with total absorption cross-sections $\Sigma_1, \Sigma_2, \Sigma_3$, due to two radiation sources: a source of intensity $A \in [0, 1]$ incident on slab 1, and a source of intensity $B = 1 - A$, incident on slab 3. The material cross-sections are the uncertain input parameters, such that $\Sigma_i = \langle\Sigma_i\rangle + \widehat{\Sigma}_i(2\xi_i - 1)$ with $\xi_i \sim U[0, 1]\ \forall i \in \{1, 2, 3\}$. If we define $k = A\exp(-r_1(\langle\Sigma_1\rangle - \widehat{\Sigma}_1) - r_2(\langle\Sigma_2\rangle - \widehat{\Sigma}_2) - r_3(\langle\Sigma_3\rangle - \widehat{\Sigma}_3))$, $j = B\exp(-r_3(\langle\Sigma_3\rangle - \widehat{\Sigma}_3))$ and $m_i = 2r_i\widehat{\Sigma}_i\ \forall i \in \{1, 2, 3\}$, then we have the following expression for T:

$$T(\xi_1, \xi_2, \xi_3) = k\exp(-m_1\xi_1 - m_2\xi_2 - m_3\xi_3) + j\exp(-m_3\xi_3). \tag{8}$$

By integrating T and T^2 over the domains of $\xi_{1,2,3}$ (which we denote $\Omega_{1,2,3}$), we can evaluate $V(T)$. To compute the main effect SI for ξ_1, we need an expression for the variance of the conditional mean, the numerator in Eq. (1), which we evaluate as

$$\int_{\Omega_1}\left(\int_{\Omega_3}\int_{\Omega_2}(T|\xi_1)d\xi_2 d\xi_3\right)^2 d\xi_1 - \left(\int_{\Omega_1}\left(\int_{\Omega_3}\int_{\Omega_2}(T|\xi_1)d\xi_2 d\xi_3\right)d\xi_1\right)^2. \tag{9}$$

To compute the total effect SI for ξ_1, we need an expression for the mean of the conditional variance, the numerator in Eq. (2), which we evaluate as

$$\int_{\Omega_2}\int_{\Omega_3}\left(\int_{\Omega_1}(T|\xi_2\xi_3)^2 d\xi_1 - \left(\int_{\Omega_1}(T|\xi_2\xi_3)d\xi_1\right)^2\right)d\xi_3 d\xi_2. \tag{10}$$

Here we give derived analytic expressions for the first and second moments of T ($\langle T\rangle$ and $\langle T^2\rangle$), the numerators for the main effect indices ($V(E(T|\xi_i))$), and the numerators for the total effect indices ($E(V(T|\xi_i))$), using the constants as defined above:

$$\langle T\rangle = \frac{1}{m_1 m_2 m_3}\exp(-m_1-m_2-m_3)(\exp(m_3)-1)((\exp(m_1)-1)(\exp(m_2)-1)k+ \exp(m_1+m_2)jm_1m_2), \tag{11}$$

$$\langle T^2\rangle = \frac{1}{8m_1m_2m_3}\exp(-2(m_1+m_2+m_3))(\exp(2m_3)-1)(-8\exp(m_1+m_2)(\exp(m_1)-1)jk- (\exp(2m_1)-1)k^2+\exp(2m_2)(-8\exp(m_1)jk-k^2+\exp(2m_1)(8jk+ k^2+4j^2m_1m_2))), \tag{12}$$

$$V(E(T|\xi_{1,2})) = \frac{1}{2m_1^2m_2^2m_3^2}\exp(-2(m_1+m_2+m_3))(\exp(m_{1,2})-1)(\exp(m_{2,1})-1)^2\times (\exp(m_3)-1)^2k^2(2+\exp(m_{1,2})(-2+m_{1,2})+m_{1,2}), \tag{13}$$

$$V(E(T|\xi_3)) = \frac{1}{2m_1^2m_2^2m_3^2}\exp(-2(m_1+m_2+m_3))(\exp(m_3)-1)((\exp(m_1)-1)\times (\exp(m_2)-1)k+\exp(m_1+m_2)jm_1m_2)^2(2+\exp(m_3)\times (-2+m_3)+m_3), \tag{14}$$

$$E(V(T|\xi_{1,2})) = \frac{1}{m_{1,2}^2m_2m_3}2k^2\left(-2+m_{1,2}\coth\left(\frac{m_{1,2}}{2}\right)\right)\sinh\left(\frac{m_{1,2}}{2}\right)^2\sinh\left(m_{2,1}\right)\sinh(m_3)\times (\cosh(m_1+m_2+m_3)-\sinh(m_1+m_2+m_3)), \tag{15}$$

$$E(V(T|\xi_3)) = \frac{1}{8m_1m_2m_3^2}\exp(-2(m_1+m_2+m_3))(\exp(m_3)-1)(-8\exp(m_1+m_2)\times (\exp(m_1)-1)jk-(\exp(2m_1)-1)k^2+\exp(2m_2)(-8\exp(m_1)jk- k^2+\exp(2m_1)(8jk+k^2+4j^2m_1m_2)))(2+\exp(m_3)(-2+m_3)+ m_3). \tag{16}$$

We adapt this function to emulate a MC radiation transport history, by evaluating Eq. (8) for an uncertain input parameter sample Ξ of (ξ_1, ξ_2, ξ_3). We then sample a psuedo-random number Θ from $U[\min(T), \max(T)]$, where $\min(T) = T(0,0,0)$ and $\max(T) = T(1,1,1)$. The particle is either absorbed or transmitted, so, if $\Theta > T(\Xi)$, $\max(T)$ is returned (analogous to a particle transmission), otherwise $\min(T)$ is returned (analogous to a particle absorption). As $A \to 0$, the indices for ξ_3 grow larger and the indices for ξ_1, ξ_2 shrink exponentially. If we choose $A \neq 1$, then this problem has different main and total effect SI for ξ_3 than for $\xi_{1,2}$. We use this adapted benchmark problem in order to demonstrate the performance of our new method.

3. NUMERICAL RESULTS

We compute the precision of our new method for the benchmark problem described in this paper. We set the free parameters in the problem as $A = B = 0.5, r_1 = r_2 = r_3 = 1.0, \langle\Sigma_1\rangle = \langle\Sigma_2\rangle = \langle\Sigma_3\rangle = 0.5$ and $\widehat{\Sigma}_1 = \widehat{\Sigma}_2 = \widehat{\Sigma}_3 = 0.1$. For each set of results, we repeat the calculation of the SI fifty times and compute the Mean Absolute Error (MAE) and statistical uncertainty over these repetitions. We compute estimates of the SI using the traditional method with three different numbers of MC particle histories (1e+04, 1e+05 and 1e+06) per MC transport code run (one code run per uncertain parameter sample), but using the same total number of particle histories per SI estimate. So for example, when we use 1e+04 particle histories per code run, we use 3e+03 code runs per SI estimate, resulting in 3e+07 histories per estimate; when we use 1e+05 MC histories per code run, we use 3e+02 code runs per SI estimate, also resulting in 3e+07 histories per estimate. In previous work, EVADE has been empirically shown to be more efficient when the number of MC particle histories used per variance estimate is small. We therefore choose small numbers of histories when carrying out our numerical analysis for the EVADE-based method: four particle histories per variance estimate (equivalent to two particle pairs per hyperplane) and thirty particle histories per variance estimate (equivalent to fifteen particle pairs per hyperplane). The traditional method benefits from the sharing of MC transport code run results across SI estimates. Thus, in the interest of parity, we plot the MAE against the cost in MC particle histories used to compute all six SI for the test problem for both methods. We present the results in Table 1 and in Figure 3. As evidenced by the numerical results in Table 1 and, as expected from the analytic derivations for the SI, the precision of the estimates for S_1 and S_2 are virtually identical, as are the estimates for S_{T_1} and S_{T_2}. Consequently, for brevity, we omit the results from S_2 and S_{T_2} in Figure 3.

Table 1: Benchmark function numerical results after 3e+07 particle histories

Method	S_1	S_2	S_3	S_{T_1}	S_{T_2}	S_{T_3}
Analytic	0.0635	0.0635	0.8725	0.0638	0.0638	0.8730
1e4 hist. solve	0.0462 ± 0.0250	0.0870 ± 0.0268	0.749 ± 0.0844	0.0644 ± 0.0004	0.0644 ± 0.0003	0.876 ± 0.0045
1e5 hist. solve	0.0920 ± 0.0810	0.0366 ± 0.0812	0.761 ± 0.339	0.0641 ± 0.0009	0.0646 ± 0.0008	0.892 ± 0.0156
1e6 hist. solve	-0.1024 ± 0.2578	0.5619 ± 0.2742	0.9059 ± 1.0310	0.0653 ± 0.0037	0.0623 ± 0.0035	0.9215 ± 0.0371
EVADE 4 hist. plane	0.0627 ± 0.0006	0.0623 ± 0.0009	0.8730 ± 0.0007	0.0638 ± 0.0007	0.0627 ± 0.0009	0.8737 ± 0.0006
EVADE 30 hist. plane	0.0634 ± 0.0007	0.0645 ± 0.0007	0.8729 ± 0.0004	0.0642 ± 0.0005	0.0638 ± 0.0005	0.8722 ± 0.0007

We observe that our method demonstrates MC convergence behavior, which supports the mathematical argument we have given, that the rate of convergence of the ratio estimator is very closely proportional to $1/\sqrt{N}$. We observe an interesting result: the traditional SI estimator consistently produces more precise SI estimates when used with fewer particle histories per transport code run for a given computational cost. This unanticipated and interesting result suggests that using the traditional SI estimator with very few particle histories per transport code run may be more efficient than using the traditional SI estimator with well-resolved transport code runs. For the main effect SI for this problem, the EVADE-based estimators show roughly 2 orders of magnitude greater precision for the same computational cost than the traditional estimators. For estimation of the total effect SI for uncertain parameters 1 and 2, the EVADE-based estimators do not show greater precision than the traditional estimators. For estimation of the total effect SI for uncertain parameter 3, the EVADE-based methods show roughly one order of magnitude greater precision than the traditional method for the same computational cost.

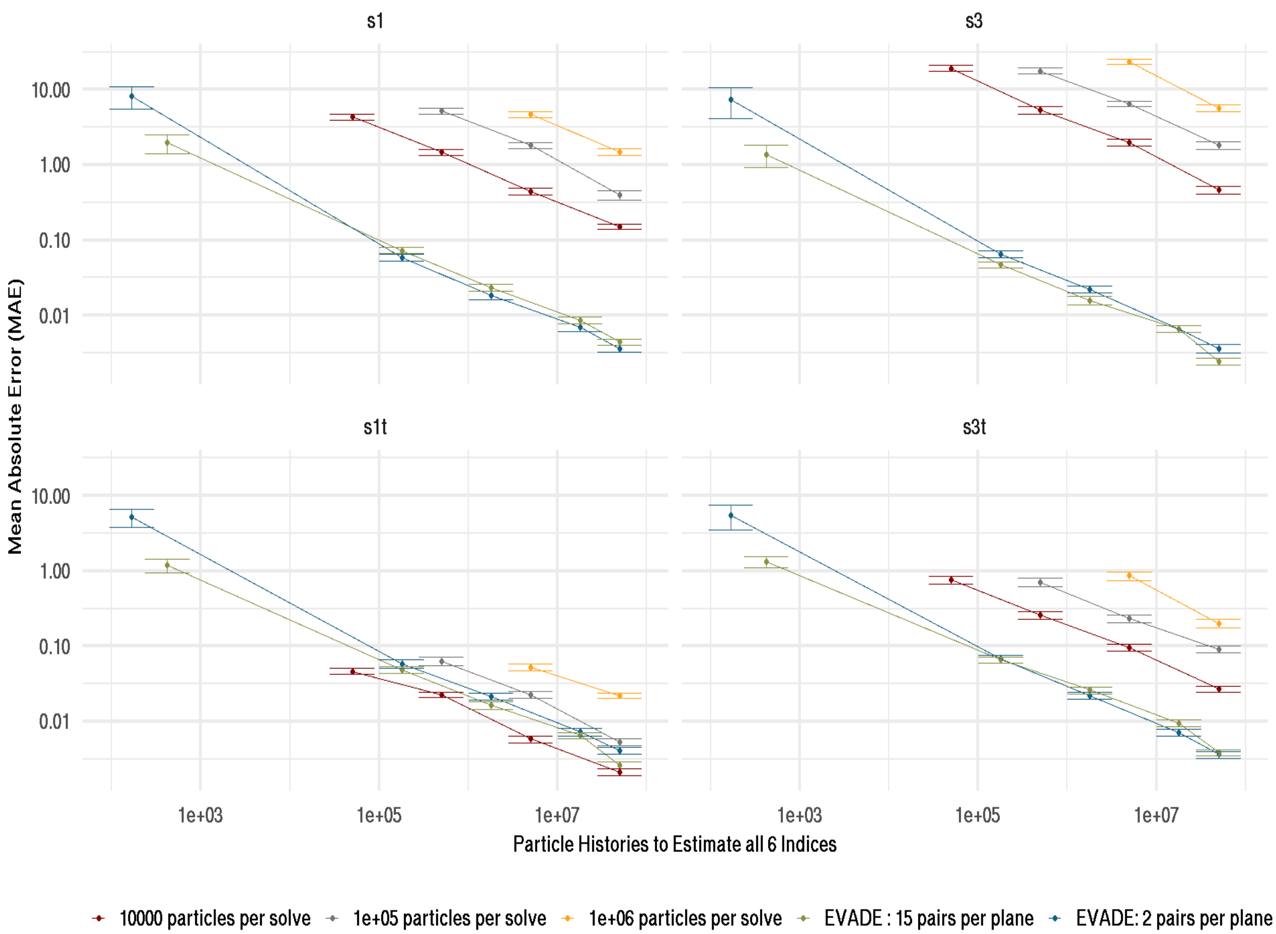

Figure 3: MAE of Sobol'index estimates as a function of total number of particle histories

4. CONCLUSIONS AND FURTHER WORK

We used SNL's EVADE algorithm for estimating parametric variance to also estimate conditional parametric variance. We successfully demonstrated that we are able to leverage EVADE to estimate conditional variances and "free" estimates of the full variance, and that we were able to propogate the uncertainties from both of these through to the SI estimate. To validate our results, we expanded an existing benchmark problem with analytic solutions for variance and SI terms, which can be used as a benchmark function for future uncertainty quantification research. Our results show the expected MC convergence behavior on the new test problem. For four of the six SI estimates, we showed evidence of one to two orders of magnitude greater precision over the traditional estimator for a fixed computational cost for a MC particle transport code with uncertain input parameters. We seek to realize computational gains of such magnitudes when performing GSA by embedding this estimation scheme within a MC radiation transport code.

In our future work, we would like to pursue further computational gains using three mechanisms:

optimizing the distribution of MC particle histories across estimates for the EVADE-based method, exploring the performance of the traditional estimator when used with many poorly resolved MC transport code runs, and, finally, exploring alternative ways to hybridize the traditional method with EVADE. Future work may also include application of the new algorithm when using MC solvers other than MC radiation transport codes.

5. ACKNOWLEDGMENTS

Supported by the Laboratory Directed Research and Development program at Sandia National Laboratories, a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energys National Nuclear Security Administration under contract DENA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government. The authors are grateful to Prof. Jonathan Rougier for technical discussions on asymptotic statistics leading to an understanding of the convergence properties of the ratio estimator. The authors thank Dr. Jacquilyn Weeks for writing consultation in the preparation of this document.

References

[1] I. Sobol. "Sensitivity estimates for nonlinear mathematical models." *Mathematical Modelling and Computational Experiment*, **volume 1**, pp. 407–414 (1993).

[2] A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, and S. Tarantola. "Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index." *Computer Physics Communications*, **volume 181**, pp. 259–270 (2010).

[3] M. Jansen. "Analysis of variance design for model output." *Computer Physics Communications*, **volume 117**, pp. 35–43 (1999).

[4] B. Sudret. "Global sensitivity analysis using polynomial chaos expansion." *Reliability Engineering and System Safety*, **volume 93**, pp. 964–979 (2000).

[5] J. Oakley and A. O'Hagan. "Probabilistic sensitivity analysis of complex models: A Bayesian approach." *Journal of the Royal Statistical Society*, **volume 66**, pp. 751–769 (2004).

[6] A. J. Olson. "Calculation of parametric variance using variance deconvolution." *Transactions of the American Nuclear Society*, **volume 120**, pp. 461–464 (2019).

[7] A. Owen. "Monte Carlo theory, methods and examples." (Cited October 2020). URL http://statweb.stanford.edu/~owen/mc/Ch-intro.pdf.

[8] A. van der Vaart. *Asymptotic statistics*. Cambridge University Press (1998).