SAND2019-2691C

Calculation of Parametric Variance using Variance Deconvolution

Aaron J. Olson*

*Sandia National Laboratories, Albuquerque, NM 87185, USA, aolson@sandia.gov

INTRODUCTION

For radiation transport problems involving uncertainty, the mean of a computed quantity is only of some value without some characterization of the spread. Ideally, the spread of each result is fully characterized as a probability density function. In practice, simply computing the parametric variance–the variance of the spread in the response caused by the uncertainty described by the input parameters–often goes a long way. Accurately computing the parametric variance usually requires a collection of well-resolved transport calculations each performed for a different sample of the random variables. In this paper, we present a new approach for computing the parametric variance for problems involving input uncertainty when using a Monte Carlo transport solver in which the total variance is deconvolved into the Monte Carlo and parametric variances and offer a non-intrusive and an intrusive implementation of this idea. We expand on an existing analytic benchmark and use the resulting solutions to assess the accuracy of the new method numerically.

Related work has focused on estimating the statistical uncertainty on mean results [1, 2, 3]. We here instead focus on developing a new way to accurately estimate the parametric variance and the statistical uncertainty on its estimate.

VARIANCE DECONVOLUTION

Randomness described by input parameters causes variance in output parameters–we call this variance parametric variance. Solving with Monte Carlo transport introduces a source of variance independent from the parametric variance who's effects are minimized by solving with many histories–we call this the Monte Carlo variance. We call the sum of these two variances the total variance:

$$V_{tot} = V_P + V_{MC}. \tag{1}$$

The relationship in Eq. (1) enables solution of any one of the terms as a function of the other two. In this section, solution of the parametric variance numerically by first estimating the total and Monte Carlo variances is described.

A sample of random input variables yields a realization of the stochastic problem. The total variance can be estimated by simulating one history on each of a collection of randomly sampled realizations since this process inherently convolves the parametric and Monte Carlo variances. The total sample (as opposed to population) variance for quantity T is then

$$V_{tot} \approx \frac{R_{tot}}{R_{tot}-1}\left(\langle T^2\rangle_{R_{tot}} - \langle T\rangle^2_{R_{tot}}\right) \tag{2}$$

where moment p of T is estimated

$$\langle T^p\rangle_{R_{tot}} \approx \frac{1}{R_{tot}}\sum_{r=1}^{R_{tot}} T_r^p, \tag{3}$$

and r represents one of R_{tot} realizations.

The Monte Carlo variance is different for each realization r, necessitating computation of the average Monte Carlo variance across the possible realizations [2, 4]. The sample Monte Carlo variance is estimated on realization r using N_{MC} histories:

$$V_{MC,r} \approx \frac{N_{MC}}{N_{MC}-1}\left(\langle T^2\rangle_{N_{MC},r} - \langle T\rangle^2_{N_{MC},r}\right). \tag{4}$$

The average Monte Carlo variance is then estimated by averaging on R_{MC} randomly sampled realizations:

$$V_{MC} \approx \frac{1}{R_{MC}}\sum_{r=1}^{R_{MC}} V_{MC,r}. \tag{5}$$

A simple, non-intrusive way to make use of variance deconvolution is to simulate one history ($N_{tot} = 1$) on each of R_{tot} realizations to estimate V_{tot} (using Eq. (2)), separately simulate N_{MC} histories on each of R_{MC} realizations to estimate V_{MC} (using Eqs. (4) and (5)), and take the difference to estimate the parametric variance V_P (using Eq (1)). We call this implementation VADE for VAriance DEconvolution. VADE requires setting up $R_{tot} + R_{MC}$ realizations and simulating $R_{tot} + R_{MC}N_{MC}$ particle histories.

Embedded Variance Deconvolution

We also present an intrusive implementation of variance deconvolution and call it EVADE for Embedded VAriance DEconvolution. In EVADE, tallies are taken to estimate the total variance and Monte Carlo variance on every realization such that $R_{tot} = R_{MC}$. Tallies of the first and second moment used in Eq. (2) to estimate the total variance are taken over the ensemble of realizations only as a function of one history on each realization (we arbitrarily choose the first). Tallies of the first and second moment used in Eq. (4) to estimate the Monte Carlo variance are taken on each realization and after estimating the Monte Carlo variance on that realization are discarded. Thus $\langle T\rangle_{N_{MC},r}$ and $\langle T^2\rangle_{N_{MC},r}$ are tallied within each realization and $\langle T\rangle_{R_{tot}}$, $\langle T^2\rangle_{R_{tot}}$, and $V_{MC,r}$ are tallied over the ensemble of realizations. EVADE requires setting up R_{MC} realizations and simulating $R_{MC}N_{MC}$ particle histories.

In EVADE, we use batches of realizations to provide estimates of the uncertainty (one-sigma standard error) on computed quantities including the parametric variance. Quantities are estimated as the average of the estimates on the batches and the uncertainty is computed in the typical manner as a function of the sample variance of these values:

$$u^2_{V_P} = \frac{1}{B-1}\left(\langle V_P^2\rangle_B - \langle V_P\rangle^2_B\right), \tag{6}$$

where B is the number of batches and u_{V_P} is the uncertainty (standard statistical error) in the parametric variance.

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.

We choose to simulate only two histories on each realization such that as many histories as possible (half of the total histories) are used to contribute to the calculation of the total variance. For both VADE and EVADE, we tally and equally weight all simulated histories to estimate the mean transmittance, i.e., using $\langle T \rangle_{R_{tot}}$ and $\langle T \rangle_{N_{MC},r}$ as applicable for VADE and average values of $\langle T \rangle_{N_{MC},r}$ over ensembles for EVADE.

PROBLEM DESCRIPTION AND BENCHMARKS

Here we solve the stochastic, one-dimensional, neutral-particle, absorption-only, mono-energetic, and steady-state radiation transport equation with a normally incident beam source of magnitude one:

$$\mu \frac{\partial \psi(x,\mu,\omega)}{\partial x} + \Sigma_t(x,\omega)\psi(x,\mu,\omega) = 0, \tag{7a}$$

$$0 \leq x \leq L; \quad -1 \leq \mu \leq 1, \tag{7b}$$

$$\psi(0,\mu) = \delta(1-\mu), \;\; \mu > 0; \quad \psi(L,\mu) = 0, \;\; \mu < 0, \tag{7c}$$

where $\psi(x,\mu,\omega)$ is angular flux, $\Sigma_t(x,\omega)$ is the total cross section, and x, μ, and ω denote spatial, angular, and stochastic dependence. Particle travel is restricted to the forward ($\mu = 1$) direction via the particle flux boundary condition and lack of particle interactions other than absorption. The locations of external material boundaries are fixed, i.e., $x \in [0, L]$, while either total cross sections, internal material boundary locations, or both are varied uniformly. The stochastic total cross section for material m of M material regions is denoted as a function of the average total cross section, $\langle \Sigma_{t,m} \rangle$, a parameter defining the range of possible values, $\widehat{\Sigma}_{t,m}$, and a random variable uniformly distributed from negative one to one, $\xi_m(\omega) \in \mathcal{U}[-1,1]$:

$$\Sigma_{t,m}(\omega) = \langle \Sigma_{t,m} \rangle + \widehat{\Sigma}_{t,m}\xi_m(\omega). \tag{8}$$

The location of stochastic internal material boundary m (of $M-1$) is denoted similarly ($\zeta_m(\omega) \in \mathcal{U}[-1,1]$):

$$x_m(\omega) = \langle x_m \rangle + \widehat{x}_m\zeta_m(\omega) \tag{9}$$

such that material chunk m has width

$$\Delta x_m(\omega) = \begin{cases} x_1(\omega) & \text{if } m = 1 \\ x_m(\omega) - x_{m-1}(\omega) & \text{if } 1 < m < M \\ L - x_{M-1}(\omega) & \text{if } m = M. \end{cases} \tag{10}$$

Analytic and Semi-Analytic Benchmarks

The transmittance $T(\omega) = \psi(L,1,\omega)$ for a sample ω from the stochastic space is solved as a function of the sampled slab's optical thickness $\tau(\omega)$:

$$T(\omega) = \exp\left[-\tau(\omega)\right], \tag{11a}$$

$$\tau(\omega) = \sum_{m=1}^{M} \Sigma_{t,m}(\omega)\Delta x_m(\omega). \tag{11b}$$

The expectation of moment p of the transmittance for a problem involving M independent material chunks is defined by integrating over the probability space:

$$\mathbb{E}[T^p] = \left(\frac{1}{2}\right)^{2M-1} \int_{-1}^{1} \dots \int_{-1}^{1} T^p(\omega) d\xi_1 \dots d\xi_M d\zeta_1 \dots d\zeta_{M-1}. \tag{12}$$

Eq. (12) becomes simpler if only total cross sections vary ($\widehat{x}_m = 0 \forall m$) or if only internal material boundary locations vary ($\widehat{\Sigma}_{t,m} = 0 \forall m$).

In the case that only total cross sections vary, the expectation of moment p of transmittance is solved using the analytic transmittance (Eq. (17)) and random variable definition (Eq. (8)) as in Ref. [5]:

$$\mathbb{E}[T^p] = \prod_{m=1}^{M} \exp\left[-p\langle \Sigma_{t,m} \rangle \Delta x_m\right] \frac{\sinh\left[p\widehat{\Sigma}_{t,m}\Delta x_m\right]}{p\widehat{\Sigma}_{t,m}\Delta x_m}. \tag{13}$$

Similarly, we derive the expectation of moment p for the case involving random internal material boundary locations and constant total cross sections:

$$\begin{aligned} \mathbb{E}[T^p] = &\left(\prod_{m=1}^{M} \exp\left[-p\Sigma_{t,m}\left(\langle x_m \rangle - \langle x_{m-1} \rangle\right)\right]\right) \\ &\cdot \left(\prod_{m=1}^{M-1} \frac{\sinh\left[p\left(\Sigma_{t,m+1} - \Sigma_{t,m}\right)\widehat{x}_m\right]}{p\left(\Sigma_{t,m+1} - \Sigma_{t,m}\right)\widehat{x}_m}\right). \end{aligned} \tag{14}$$

These analytic solutions were attained via the separability of the integrals in Eq. (12). When both coefficients and material boundary locations vary, we no longer have separability and choose instead to numerically integrate Eq. (13) (using scipy):

$$\begin{aligned} \mathbb{E}[T^p] = &\left(\frac{1}{2}\right)^{M-1} \int_{-1}^{1} \dots \int_{-1}^{1} d\zeta_1 \dots d\zeta_{M-1} \\ &\prod_{m=1}^{M} \exp\left[-p\langle \Sigma_{t,m} \rangle \Delta x_m(\omega)\right] \frac{\sinh\left[p\widehat{\Sigma}_{t,m}\Delta x_m(\omega)\right]}{p\widehat{\Sigma}_{t,m}\Delta x_m(\omega)}. \end{aligned} \tag{15}$$

The parametric variance is solved as a function of the first and second moment provided by Eq. (13), (14) , or (15):

$$V_P = \mathbb{E}[T^2] - \mathbb{E}^2[T]. \tag{16}$$

To solve for the total variance, we note that with analog Monte Carlo and one history per realization, all transmittance tallies T_r are either zero or one such that all transmittance moments (Eq. (3)) are equal and that in the limit of many realizations, the Monte Carlo simulation yields the transmittance expectation computed by Eq. (13), (14) , or (15):

$$\langle T \rangle_{R_{tot}} = \langle T^p \rangle_{R_{tot}} \forall p, \tag{17a}$$

$$\lim_{R_{tot}\to\infty} \langle T \rangle_{R_{tot}} = \mathbb{E}[T]. \tag{17b}$$

Equations (17a) and (17b) enable analytic computation of the total variance from Eq. (2) based only on $\mathbb{E}[T]$:

$$V_{tot} = \lim_{R_{tot}\to\infty} \frac{R_{tot}}{R_{tot}-1}\left(\langle T^2 \rangle_{R_{tot}} - \langle T \rangle^2_{R_{tot}}\right) = \mathbb{E}[T] - \mathbb{E}^2[T]. \tag{18}$$

The Monte Carlo variance is analytically solved using variance deconvolution (Eq. (1)) with the analytically solved parametric (Eq. (16)) and Monte Carlo (Eq. (18) variances:

$$V_{MC} = V_{tot} - V_P. \tag{19}$$

NUMERICAL RESULTS

We here apply VADE and EVADE to three different problems and compare against the analytic and semi-analytic solutions defined earlier in the paper. Each problem investigated has the same average cross sections and material boundary locations (listed in Table I), but are differentiated by which values contain uncertainty. The "Random Cross Sections" problem contains the cross-section uncertainty described in Table I while material location uncertainty is zero; the "Random Boundaries" problem contains the random boundary uncertainty described in Table I while cross-section uncertainty is zero; and the third problem contains uncertainty in cross sections and material boundary locations.

TABLE I. Stochastic Attenuation Problem Parameters

	$m=0$	$m=1$	$m=2$	$m=3$
$\langle\Sigma_{t,m}\rangle$	-	0.9	0.15	0.6
$\widehat{\Sigma}_{t,m}$	-	0.7	0.12	0.5
$\langle x_m\rangle$	0.0	2.0	5.0	6.0
$\widehat{x_m}$	-	1.75	0.95	-

We first solve these problems numerically using VADE and EVADE with a total of $N_{TOT} = 2\times10^8$ histories each according to the parameters in Table II and choose the number of batches for EVADE that yields two realizations per batch.

TABLE II. Higher-Fidelity Solver Parameters

VADE		EVADE	
$R_{tot}=10^8$	$R_{MC}=10^4$	$R_{tot}=R_{MC}=10^8$	
$N_{tot}=1$	$N_{MC}=10^4$	$N_{tot}=1$	$N_{MC}=2$
B =N/A		$B=5\times10^7$	
$N_{TOT}=2\times10^8$		$N_{TOT}=2\times10^8$	

Analytic or semi-analytic solutions to four quantities of interest as well as those yielded numerically by VADE and EVADE are listed in Table III. While in most applications the average transmittance $\langle T\rangle$ and variance on that value caused by input uncertainty V_P would be the results of interest, we also provide the total variance V_{tot} and Monte Carlo variance V_{MC} in an effort to help investigate the method. VADE and EVADE were able to provide the correct solutions within statistical uncertainty. While VADE gives no estimate of the uncertainty on the computed parameters, simple convergence studies not shown here suggest Monte Carlo statistical convergence. Through the use of realization batches, EVADE produces estimates of statistical uncertainty; errors are roughly the size of these uncertainty values as expected.

Secondly, we seek to compare the relative precision of these methods as well as test whether the computed uncertainty estimates are accurate. We solve the same three problems 1000 times each, but with 1000 times fewer histories ($N_{TOT} = 2\times10^5$) than in the previous, higher-fidelity solves. Specific solver parameters are given in Table IV. We again choose the number of batches for EVADE that yields batches of size two.

The average error yielded for each quantity by VADE and EVADE is given in Table V as well as the average one-

TABLE III. Higher-Fidelity Simulation Representative Results

Random Cross Sections			
Quantity	Analytic	VADE	EVADE
$\langle T\rangle$	0.083783	0.083981	0.083773 ± 0.000020
V_P	0.005505	0.005637	0.005509 ± 0.000022
V_{tot}	0.076763	0.076749	0.076762 ± 0.000026
V_{MC}	0.071259	0.071112	0.071253 ± 0.000018
Random Boundaries			
Quantity	Analytic	VADE	EVADE
$\langle T\rangle$	0.078277	0.078828	0.078314 ± 0.000020
V_P	0.003731	0.002782	0.003734 ± 0.000021
V_{tot}	0.072150	0.072140	0.072176 ± 0.000025
V_{MC}	0.068419	0.069358	0.068442 ± 0.000017
Random Cross Sections and Boundaries			
Quantity	Semi-An.	VADE	EVADE
$\langle T\rangle$	0.104428	0.103672	0.104398 ± 0.000023
V_P	0.010069	0.011029	0.010097 ± 0.000024
V_{tot}	0.093523	0.093503	0.093532 ± 0.000028
V_{MC}	0.083454	0.082474	0.083435 ± 0.000019

TABLE IV. Lower-Fidelity Solver Parameters

VADE		EVADE	
$R_{tot}=10^5$	$R_{MC}=10^2$	$R_{tot}=R_{MC}=10^5$	
$N_{tot}=1$	$N_{MC}=10^3$	$N_{tot}=1$	$N_{MC}=2$
B =N/A		$B=5\times10^4$	
$N_{TOT}=2\times10^5$		$N_{TOT}=2\times10^5$	

sigma uncertainty provided by EVADE and the number of times the error was less than the uncertainty. For all quantities except the total variance, EVADE was roughly an order of magnitude more precise than VADE. The average error was a little less than the average uncertainty for each problem and quantity of interest, and the error was less than the one-sigma uncertainty about 68% of the time, providing confidence that the uncertainty estimates given by the batches–even batches of only size two–are accurate.

The error in the total variance computed with VADE was a little less than that computed by EVADE even though they each used one history from the same number of realizations. We believe the difference is an effect of estimating the total variance on small batches (only two realizations) and averaging these estimates. Additional investigation not documented here shows that computation of the total variance with EVADE without batches yields the same average error as VADE.

The propagation of uncertainty as a function of R and N derived in Ref. [4] predicts that the precision of the mean is maximized for a set number of histories with $N = 1$, falling off monotonically as a function of larger N after that. We believe that the greater precision achieved by EVADE over VADE for values other than the total variance is caused by this same mechanism suggesting that small values of N are likely

TABLE V. Lower-Fidelity Simulation Ensemble Results

Random Cross Sections				
Quantity	$\langle\epsilon\rangle_{VADE}$	$\langle\epsilon\rangle_{EVADE}$	$\langle u\rangle_{EVADE}$	$\epsilon < u$
$\langle T\rangle$	0.00375	0.00052	0.00064	668/1000
V_P	0.00551	0.00054	0.00068	673/1000
V_{tot}	0.00052	0.00065	0.00081	681/1000
V_{MC}	0.00550	0.00044	0.00055	681/1000
Random Boundaries				
Quantity	$\langle\epsilon\rangle_{VADE}$	$\langle\epsilon\rangle_{EVADE}$	$\langle u\rangle_{EVADE}$	$\epsilon < u$
$\langle T\rangle$	0.00333	0.00049	0.00062	658/1000
V_P	0.00525	0.00053	0.00065	674/1000
V_{tot}	0.00053	0.00065	0.00079	661/1000
V_{MC}	0.00521	0.00043	0.00054	682/1000
Random Cross Sections and Boundaries				
Quantity	$\langle\epsilon\rangle_{VADE}$	$\langle\epsilon\rangle_{EVADE}$	$\langle u\rangle_{EVADE}$	$\epsilon < u$
$\langle T\rangle$	0.00468	0.00058	0.00072	681/1000
V_P	0.00621	0.00059	0.00077	702/1000
V_{tot}	0.00057	0.00069	0.00087	707/1000
V_{MC}	0.00620	0.00047	0.00059	685/1000

optimal. It would be informative to perform an analytic study on this relationship in the context of VADE and EVADE.

CONCLUSIONS AND FUTURE WORK

In this paper, we numerically demonstrated attainment of parametric variance by estimating the total and Monte Carlo variances and taking the difference. A non-intrusive variance deconvolution (VADE) implementation and an intrusive embedded variance deconvolution (EVADE) implementation were demonstrated. While each implementation was accurate, EVADE was roughly an order of magnitude more precise than VADE for all output quantities except total variance for which VADE was marginally more precise. EVADE used batches of realizations to accurately estimate uncertainty on the quantities of interest. EVADE demonstrated use of as few as two particle histories per realization and two realizations per batch to accurately estimate quantities. Precision comparisons between VADE and EVADE suggest that few histories per realization are likely optimal. Both VADE and EVADE were numerically demonstrated on different classes of uncertainty problems: random cross section values, random material boundary locations, and problems involving both sources of uncertainty. In addition, an analytic expression for transmittance moments in a random material boundary attenuation problem was provided as well as an approach to obtain semi-analytic results when both total cross sections and material boundaries contain uncertainty. Finally, expressions were provided for parametric, total, and Monte Carlo transmittance variance as a function of these transmittance moments.

Whereas the parametric variance in the problems in this paper originated from traditional random variables, we would like to apply EVADE to solving for the variance caused by stochastic material mixing. Specifically, we plan to implement EVADE with Conditional Point Sampling (CoPS) [6], a new algorithm for transport in stochastic media. Future work may also include use of EVADE in problems involving both parametric variance originating from traditional random variables and stochastic material mixing, possibly to solve for only one of those contributions at a time. It would be informative to apply Optimal-Cost Monte Carlo (OCMC) [2, 4], which uses Lagrangian optimization to solve for optimal input parameter values in Monte Carlo calculations, to VADE and EVADE, both for the mean and the parametric variance. Future work could also include application of VADE where a different computational model is used for estimating the total variance than the Monte Carlo variance, for example, with stochastic media, estimation of the total variance with Chord Length Sampling [1, 3] and estimation of the Monte Carlo variance with realizations [2, 4].

ACKNOWLEDGMENTS

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.

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 author would like to thank to Ron Kensek for sharing his idea for variance deconvolution and both Ron and Brian Franke for helpful conversations on this topic.

REFERENCES

1. T. J. DONOVAN and Y. DANON, "Application of Monte Carlo Chord-Length Sampling Algorithms to Transport Through a Two-Dimensional Binary Stochastic Mixture," *Nucl. Sci. Eng.*, **143**, *3*, 226–239 (2003).
2. T. J. DONOVAN, T. SUTTON, and Y. DANON, "Implementation of Chord Length Sampling for transport through a binary stochastic mixture," in "Proc. Int. Conf. on Math. and Comp. (M&C2003)," American Nuclear Society, Gatlinburg, Tennessee, April 6-11 (2003).
3. C. LARMIER, A. LAM, P. BRANTLEY, F. MAVAGI, T. PALMER, and A. ZOIA, "Monte Carlo chord length sampling for d-dimensional Markov binary mixtures," *J. Quant. Spectrosc. and Rad. Transfer* (2018).
4. A. J. OLSON and B. C. FRANKE, "An optimal-cost Monte Carlo approach to stochastic media transport calculations," *Trans. Am. Nucl. Soc.*, **118** (2018).
5. A. J. OLSON, A. K. PRINJA, and B. C. FRANKE, "Error convergence characterization for stochastic transport methods," *Trans. Am. Nucl. Soc.*, **116**, 536 (2017).
6. E. H. VU and A. J. OLSON, "Conditional Point Sampling: A novel Monte Carlo method for radiation transport in stochastic media," *Trans. Am. Nucl. Soc.*, **120** (2019).