

THE CONTRIBUTION OF  
INDIVIDUAL CORRELATED PARAMETERS  
TO THE UNCERTAINTY  
OF INTEGRAL QUANTITIES

*D. W. Muir*  
*Special Term Appointee*  
*Argonne National Laboratory*

# INTRODUCTION

A basic tool for modern neutronics analysis is a library of multigroup cross sections, accompanied by a covariance matrix describing the uncertainty of these data and their correlations. A common application of such a covariance library is to provide input for the calculation of the variance  $D(z)$  due to the data, of a calculated integral quantity  $z$ .

Beyond this single number, it is often of interest to inquire as to the magnitude of the contributions to the variance from just a portion of the data field. For example, one may want to know the contribution of the uncertainty in the energy dependent cross section  $\sigma(E)$  for the  $^{16}\text{O}(n,p)$  reaction to the uncertainty of  $k_{\text{eff}}$  for a given reactor. Knowing this would clearly be helpful in judging the benefits of performing a new measurement, calculation, or evaluation of  $\sigma(E)$ .

Below, we propose a methodology for determining the contribution from an individual parameter, or an identified group of parameters, to the variance of  $z$ . We show that the magnitude of the contribution depends, in an important way, on the extent of data correlations.

# PARTIAL CONTRIBUTIONS TO THE TOTAL VARIANCE

Suppose there are  $k$  data parameters. A single parameter might correspond to the cross section for a specified nuclear reaction in a specified energy group. We begin with the simple case where there are no correlations present in the covariance library. If we neglect higher order derivatives, the total variance  $D(z)$  of an integral quantity  $z$  is

$$D(z) = S_1^2 D(a_1) + S_2^2 D(a_2) + \dots + S_k^2 D(a_k), \text{ where} \quad (1)$$

$$S_i = \partial z / \partial a_i \quad (2)$$

In the uncorrelated case, the size of the contribution of the uncertainty of one of the parameters, say  $a_k$ , to  $D(z)$  is obvious. It is just the  $k$ -th term in the sum. The situation becomes considerably more complex when there are non-zero correlations in the data. In this case,  $D(z)$  is no longer a simple weighted sum of variances of the parameters. If there are  $k$  parameters,

$$D(z) = \sum_{\substack{i=1, k \\ j=1, k}} S_i S_j \text{cov}(a_i, a_j) \quad (3)$$

In this more general case, it is not at all clear which term or terms should be identified as providing "the contribution of the  $k$ -th parameter to  $D(z)$ ."

One way to proceed is to introduce the concept of the "uncertainty profile,"

$$D(z) = \sum_{i=1, k} L(z,i), \text{ where} \quad (4)$$

$$L(z,i) = \sum_{j=1, k} S_i S_j \text{ cov}(a_i, a_j) \quad (5)$$

If one forms a  $k \times k$  square matrix square matrix  $\mathbf{U}$ , having the  $k^2$  elements  $u_{ij} = S_i S_j \text{ cov}(a_i, a_j)$ , then  $L(z,i)$  is the sum of the elements of the  $i$ -th row of  $\mathbf{U}$ .

Some authors use the uncertainty profile  $L(z,i)$  to quantify the contribution of parameter  $i$  to  $D(z)$  (see, for example, Ref. 2). This is an attractive approach for three reasons: First, the profile is straightforward to calculate. Secondly, it provides a useful indicator of where the "action" is in the variance summation. Finally, as shown in Eq. (4), the quantities  $L(z,i)$  sum up to the variance  $D(z)$ .

These advantages, however, are offset by the following disadvantages. Although  $L(z,i)$  has the units of variance (barns-squared), it is **not** a variance. For this reason, it cannot be guaranteed to be a positive quantity. The square root of  $L(z,i)$  is not a standard deviation and can even be imaginary.

# VARIANCE PENALTY

In this talk, we propose an alternative to the use of uncertainty profiles. The proposed approach fits with the common sense definition of "contribution" and, at the same time, avoids the mathematical problems just mentioned.

**We propose to quantify the contribution of the uncertainty in  $a_k$  to the variance of  $z$  in terms of the "variance penalty" associated with our less-than-perfect knowledge of  $a_k$ .**

The variance penalty is directly related to the usual reason for wanting to know the contribution of parameter  $a_k$  to the variance of  $z$ , namely, to learn which parameters have the most leverage in improving the accuracy of our knowledge of  $z$ .

In this simplest case where the parameters are uncorrelated, all three of the mentioned procedures for calculating the contribution of  $a_k$  to  $D(z)$  give the same result, namely,  $S_k^2 D(a_k)$ .

We now describe a procedure for computing the variance penalty in the interesting, and important, case where there are finite correlations.

The procedure involves

- (1) calculating the current value of the variance  $D(z)$ ,
- (2) imagining that the parameters of interest are re-measured with near-perfect precision,
- (3) using the covariance matrix of this hypothetical re-measurement as input to an evaluation of the updated variance  $D(z')$  using the method of minimum-variance estimation, and
- (4) calculating the variance penalty as  $D(z) - D(z')$ .

Fortunately, to use this approach, we do not actually have to go to the laboratory and do the near-perfect re-measurement, because in minimum variance-estimation, the actual numerical results of measurements do not have any effect on the updated variance. For the present purposes, a crucial feature of minimum-variance estimation is that the variance penalty  $D(z) - D(z')$ , thus defined, is itself a variance and, hence, is guaranteed to be a positive number.

A convenient reference to the methodology of updating an evaluation, using minimum-variance estimation, is Ref. 1. Following the notation there, suppose that there exists a data set corresponding to the prior measurement (or evaluation) of  $k$  different physical parameters, here represented as a column vector  $\mathbf{a}$ , with elements  $(a_j, j=1, \dots, k)$ .

Each measurement is assumed to differ from the true value  $\mathbf{x}$  by a random error  $\mathbf{e}$

$$\mathbf{a} = \mathbf{x} + \mathbf{e} \quad (6)$$

The errors  $\mathbf{e}$  are assumed to have zero mean, so that

$$E(\mathbf{a}) = \mathbf{x}, \quad (7)$$

Here  $E(\ )$  denotes the expectation operator. It is also assumed that the errors  $\mathbf{e}$  have finite second moments. These second moments form the elements of the covariance matrix  $D(\mathbf{a})$ ,

$$D(\mathbf{a}) = \text{cov}(\mathbf{a}, \mathbf{a}) = E(\mathbf{e} \mathbf{e}^T), \quad (8)$$

Here  $(^T)$  denotes the matrix transpose.

We now consider the impact of the results of a set of new measurements, represented by a column vector  $\mathbf{b}$  with elements  $(b_i, i=1, \dots, m)$ . The expected or true value of  $\mathbf{b}$ ,  $E(\mathbf{b})$ , is a known function of the true values of the parameters,  $E(\mathbf{a})$ . All such functions are assumed here to be linear, with any constant terms absorbed into the definition of the vectors  $\mathbf{a}$  and  $\mathbf{b}$ ,

$$E(\mathbf{b}) = \mathbf{R} E(\mathbf{a}) = \mathbf{R} \mathbf{x} \quad (9)$$

In Eq. (11),  $R_{ij}$  is the  $ij$ -th element of the  $m \times k$  sensitivity matrix  $\mathbf{R}$ . As above, a vector of random errors  $\mathbf{f}$ , with zero mean, causes  $\mathbf{b}$  to differ from the expected value

$$\mathbf{b} = \mathbf{R} \mathbf{x} + \mathbf{f} \quad (10)$$

For the present purposes, it is helpful to partition the parameter matrix  $\mathbf{a}$  into  $n$  "passive" parameters  $\mathbf{v}_{n,1}$  and  $m$  "active" parameters  $\mathbf{w}_{m,1}$  (the parameters of particular interest).

$$\mathbf{a} = \begin{bmatrix} \mathbf{v}_{n,1} \\ \mathbf{w}_{m,1} \end{bmatrix} \quad (11)$$

Clearly,  $n + m = k$ . Note that, for the assumed precise re-measurements, the sensitivity matrix  $\mathbf{R}$  has the special partitioned form,

$$\mathbf{R} = \begin{bmatrix} \mathbf{0}_{m,n} & \mathbf{I}_{m,m} \end{bmatrix} \quad (12)$$

Here  $\mathbf{0}_{m,n}$  is an  $m \times n$  rectangular matrix containing only zeroes, and  $\mathbf{I}_{m,m}$  is the  $m \times m$  identity matrix. Hence,

$$\mathbf{R} \mathbf{a} = \mathbf{w} \quad (13)$$

The covariance matrix of the precise re-measurements  $\mathbf{b}$  is given by  $\mathbf{D}(\mathbf{b}) = \mathbf{E}(\mathbf{f} \mathbf{f}^T)$ . In the present work, we assume that these measurements are uncorrelated with each other,

$$\mathbf{D}(\mathbf{b}) = \Phi_{m,m} \quad (14)$$

where  $\Phi_{m,m}$  denotes a square matrix of dimension  $m$ , with very small elements  $\phi_{ii}$  on the diagonal and zeroes elsewhere.



"Cross-type" covariances, indicating correlations between the prior and new measurements, can be defined analogously,  $\text{cov}(\mathbf{a}, \mathbf{b}) = E(\mathbf{e} \mathbf{f}^T)$  and  $\text{cov}(\mathbf{b}, \mathbf{a}) = E(\mathbf{f} \mathbf{e}^T)$ . However, in the present context, we assume that the precise re-measurements are uncorrelated with the prior measurements, so that

$$\text{cov}(\mathbf{a}, \mathbf{b}) = \text{cov}(\mathbf{b}, \mathbf{a})^T = \mathbf{0}_{k,m} \quad (15)$$

We note that the vector  $\mathbf{R} \mathbf{a}$  contains the values of the newly measured quantities  $\mathbf{b}$ , as calculated with the prior evaluation of the parameters. We introduce the notation  $\mathbf{p}$  for the "discrepancy" vector

$$\mathbf{p} = \mathbf{b} - \mathbf{R} \mathbf{a} \quad (16)$$

The covariance matrix for  $\mathbf{p}$  can be written immediately,

$$D(\mathbf{p}) = \text{cov}(\mathbf{p}, \mathbf{p}) = \text{cov}[(\mathbf{b} - \mathbf{R} \mathbf{a}), (\mathbf{b} - \mathbf{R} \mathbf{a})] \quad (17)$$

$$= D(\mathbf{b}) - \text{cov}(\mathbf{b}, \mathbf{a}) \mathbf{R}^T - \mathbf{R} \text{cov}(\mathbf{a}, \mathbf{b}) + \mathbf{R} D(\mathbf{a}) \mathbf{R}^T \quad (18)$$

For convenience, we introduce the inverse of  $D(\mathbf{p})$

$$\mathbf{G} = D^{-1}(\mathbf{p}) \quad (19)$$

Substituting from Eqs. (14) and (15) into Eq. (18), we obtain

$$D(\mathbf{p}) = \Phi_{m,m} + \mathbf{R} D(\mathbf{a}) \mathbf{R}^T \quad (20)$$

In the limit as the  $\phi_{ii}$  approach zero,  $D(\mathbf{p})$  approaches

$$\mathbf{R} D(\mathbf{a}) \mathbf{R}^T = D(\mathbf{R} \mathbf{a}) = D(\mathbf{w}) \quad (21)$$

Recall that our object is to determine the variance penalty of the integral quantity  $z$  associated with our imperfect knowledge of  $\mathbf{w}$ . In matrix notation,

$$\mathbf{z} = \mathbf{S} \mathbf{a} \quad (22)$$

where  $\mathbf{S}$  is a  $k$  element row vector containing the sensitivities discussed above in Eq. (2), and  $z$  is a scalar.

The impact of a new measurement on the variance (or covariances) of an integral quantity  $z$  (or a set of such quantities) is derived in Ref. 1. In particular, Eq. (12) of that reference supplies the updated variance  $D(z')$  as a function of the prior variance  $D(z)$  and other known quantities.

$$D(z') = D(z) - D[ \text{cov}(z,\mathbf{p}) \mathbf{G} \mathbf{p} ] \quad (23)$$

We are now in a position to define the variance penalty  $P(z,\mathbf{w})$ ,

$$P(z,\mathbf{w}) = D(z) - D(z') \quad (24)$$

$P(z,\mathbf{w})$  is the amount by which the current variance  $D(z)$  exceeds the smaller variance  $D(z')$  attainable by means of an ideal re-measurement of  $\mathbf{w}$ . This is what we mean by the "variance penalty" associated with our current knowledge. From the two previous equations, we have

$$P(z, \mathbf{w}) = D[ \text{cov}(z, \mathbf{p}) \mathbf{G} \mathbf{p} ] \quad (25)$$

From Eq. (25) we see that **the variance penalty is itself a variance**. Since it is a variance, it is guaranteed to be a positive quantity. This was a primary objective in formulating our approach to the problem. Since it is positive, we can also define a standard deviation penalty, in units of barns, as

$$\Delta(z, \mathbf{w}) = P(z, \mathbf{w})^{1/2} \quad (26)$$

We now proceed to obtain an explicit expression for  $P(z, \mathbf{w})$  that is convenient for computation. From the above expressions for  $\mathbf{p}$  and  $z$ , Eqs. (16) and (22), we have

$$\text{cov}(z, \mathbf{p}) = \text{cov}[ \mathbf{S} \mathbf{a}, (\mathbf{b} - \mathbf{R} \mathbf{a}) ] = \mathbf{S} \text{cov}(\mathbf{a}, \mathbf{b}) - \mathbf{S} D(\mathbf{a}) \mathbf{R}^T \quad (27)$$

$$\text{cov}(z, \mathbf{p}) = - \mathbf{S} D(\mathbf{a}) \mathbf{R}^T \quad (28)$$

From Eq. (25), we have then

$$P(z, \mathbf{w}) = \text{cov}(z, \mathbf{p}) \mathbf{G} D(\mathbf{p}) \mathbf{G} \text{cov}(z, \mathbf{p})^T \quad (29)$$

From the definition of  $\mathbf{G}$ , we have

$$\mathbf{G} D(\mathbf{p}) \mathbf{G} = \mathbf{G} \quad (30)$$

Now combining Eqs. (28) - (30), we have

$$P(z, \mathbf{w}) = \mathbf{S} D(\mathbf{a}) \mathbf{R}^T \mathbf{G} \mathbf{R} D(\mathbf{a}) \mathbf{S}^T = \mathbf{S} [ D(\mathbf{a}) \mathbf{R}^T \mathbf{G} \mathbf{R} D(\mathbf{a}) ] \mathbf{S}^T \quad (31)$$

$$P(\mathbf{z}, \mathbf{w}) = \mathbf{S} \mathbf{H} \mathbf{S}^T, \text{ where} \quad (32)$$

$$\mathbf{H} = [ \mathbf{R} \mathbf{D}(\mathbf{a}) ]^T \mathbf{D}^{-1}(\mathbf{w}) [ \mathbf{R} \mathbf{D}(\mathbf{a}) ] \quad (33)$$

In this last result, we have replaced the matrix  $\mathbf{G} = \mathbf{D}^{-1}(\mathbf{p})$  by the high-precision limit  $\mathbf{D}^{-1}(\mathbf{w})$ . It is useful now to partition the covariance matrix of the parameters. It is easily shown that

$$\mathbf{D}(\mathbf{a}) = \begin{bmatrix} \mathbf{D}(\mathbf{v})_{n,n} & \text{cov}(\mathbf{v}, \mathbf{w})_{n,m} \\ \text{cov}(\mathbf{w}, \mathbf{v})_{m,n} & \mathbf{D}(\mathbf{w})_{m,m} \end{bmatrix} \quad (34)$$

We next calculate the intermediate matrix product  $\mathbf{R} \mathbf{a}$ . Recalling Eq. (12),

$$\mathbf{R} \mathbf{D}(\mathbf{a}) = \begin{bmatrix} \mathbf{0}_{m,n} & \mathbf{I}_{m,m} \end{bmatrix} \begin{bmatrix} \mathbf{D}(\mathbf{v})_{n,n} & \text{cov}(\mathbf{v}, \mathbf{w})_{n,m} \\ \text{cov}(\mathbf{w}, \mathbf{v})_{m,n} & \mathbf{D}(\mathbf{w})_{m,m} \end{bmatrix} \quad (35)$$

$$= \begin{bmatrix} \text{cov}(\mathbf{w}, \mathbf{v})_{m,n} & \mathbf{D}(\mathbf{w})_{m,m} \end{bmatrix} \quad (36)$$

Substituting this result into Eq. (33), and now omitting the dimensions, we have

$$\mathbf{H} = \begin{bmatrix} \text{cov}(\mathbf{v}, \mathbf{w}) \\ \mathbf{D}(\mathbf{w}) \end{bmatrix} \mathbf{D}^{-1}(\mathbf{w}) \begin{bmatrix} \text{cov}(\mathbf{w}, \mathbf{v}) & \mathbf{D}(\mathbf{w}) \end{bmatrix} \quad (37)$$

$$\mathbf{H} = \begin{bmatrix} \text{cov}(\mathbf{v}, \mathbf{w}) \mathbf{D}^{-1}(\mathbf{w}) \text{cov}(\mathbf{w}, \mathbf{v}) & \text{cov}(\mathbf{v}, \mathbf{w}) \\ \text{cov}(\mathbf{w}, \mathbf{v}) & \mathbf{D}(\mathbf{w}) \end{bmatrix} \quad (38)$$

For convenience, we recall from Eq. (24) that the integral  $z$  is related to the parameters  $\mathbf{a}$  by

$$z = \mathbf{S} \mathbf{a} \quad (39)$$

and from Eq. (33), that

$$\mathbf{P}(z, \mathbf{w}) = \mathbf{S} \mathbf{H} \mathbf{S}^T \quad (40)$$

Eqs. (38-40) summarize our **recommended calculational procedure**.

If the data set  $\mathbf{a}$  contains only data subsets  $\mathbf{a}^T = [\mathbf{g}_1 \ \mathbf{g}_2 \ \mathbf{g}_3 \ \dots \ \mathbf{g}_j]$  that are uncorrelated with each other (i.e., if the parameter covariance matrix  $\mathbf{D}(\mathbf{a})$  is block diagonal), it is easy to show that

$$\mathbf{D}_{\text{bd}}(z) = \sum_{i=1, j} \mathbf{P}_{\text{bd}}(z, \mathbf{g}_i) \quad (41)$$

In this special case (and only in this case), the sum over  $i$  of the variance penalties  $\mathbf{P}_{\text{bd}}(z, \mathbf{g}_i)$  is equal to the variance of the integral quantity  $z$ .

# DISCUSSION AND CONCLUSION

We note from our solution, Eqs. (38-40), that the variance penalty associated with the uncertainties in the data subset  $\mathbf{w}$  involves, as expected, the variances of the elements of  $\mathbf{w}$  and the correlations among them, i.e., the entire submatrix  $D(\mathbf{w})$ , as well as the sensitivity of the integral quantity  $z$  to these parameters.

More interesting is the fact that the variance penalty also depends on the inverse of this matrix  $D^{-1}(\mathbf{w})$  and on the covariance data occupying the upper right and lower left "corners" of  $D(\mathbf{a})$  in Eq. (34), namely,  $\text{cov}(\mathbf{v}, \mathbf{w})$  and  $\text{cov}(\mathbf{w}, \mathbf{v})$ .

In contrast to the uncertainty profile approach, where  $L(z, i)$  is computed as a weighted sum of  $k$  covariances, the calculation of the variance penalty  $P(z, \mathbf{w})$  requires a matrix inversion and several matrix multiplications.

Nevertheless, the mathematical clarity of the variance penalty approach, with its guarantee of a positive result, makes it attractive for general use in characterizing **the contribution of uncertainty in data subset  $\mathbf{w}$  to the variance of  $z$ .**

## REFERENCES

1. D.W. Muir, "Evaluation of Correlated Data Using Partitioned Least Squares: A Minimum-Variance Derivation," Nucl. Sci. Eng. 101, 88-93 (1989). A related report is available online at <http://home.comcast.net/~douglasmuir/LA-UR-2365REV.doc>.
2. H. Hiruta, G. Palmiotti, M. Salvatores, R. Arcilla Jr., P. Oblozinsky, R. D. McKnight, " Few group collapsing of covariance matrix data based on a conservation principle," Workshop on Neutron Cross Section Covariances, Port Jefferson, NY, June 24-27, 2008.