Available online at www.sciencedirect.com

SciVerse ScienceDirect

Nuclear Data Sheets 113 (2012) 3006–3053

Nuclear Data Sheets

www.elsevier.com/locate/nds

Experimental Nuclear Reaction Data Uncertainties: Basic Concepts and Documentation

D.L. Smith[1,*] and N. Otuka[2]

[1]*Argonne National Laboratory, 1710 Avenida Del Mundo #1506, Coronado, CA 92118, USA*
[2]*Nuclear Data Section, International Atomic Energy Agency, Wagramerstraße 5, A-1400 Wien, Austria*

(Received 23 January 2012; revised received 10 June 2012; accepted 18 July 2012)

This paper has been written to provide experimental nuclear data researchers and data compilers with practical guidance on dealing with experimental nuclear reaction data uncertainties. It outlines some of the properties of random variables as well as principles of data uncertainty estimation, and illustrates them by means of simple examples which are relevant to the field of nuclear data. Emphasis is placed on the importance of generating mathematical models (or algorithms) that can adequately represent individual experiments for the purpose of estimating uncertainties in their results. Several types of uncertainties typically encountered in nuclear data experiments are discussed. The requirements and procedures for reporting information on measurement uncertainties for neutron reaction data, so that they will be useful in practical applications, are addressed. Consideration is given to the challenges and opportunities offered by reports, conference proceedings, journal articles, and computer libraries as vehicles for reporting and documenting numerical experimental data. Finally, contemporary formats used to compile reported experimental covariance data in the widely used library EXFOR are discussed, and several samples of EXFOR files are presented to demonstrate their use.

Editorial Note: *All papers in the present issue, including this paper, are printed in black and white. Color version is available online at www.sciencedirect.com, see also www.elsevier.com/locate/nds.*

CONTENTS

* Corresponding author, electronic address:
Donald.L.Smith@anl.gov

doi:10.1016/j.nds.2012.11.004

I. INTRODUCTION

A. Overview

Physics is an observational science. All that is known about our physical environment has been acquired by observing, compiling, and striving to understand the results of observations. Consequently, experimental investigation is a fundamental aspect of the physical sciences. However, observations and compilation of the results are insufficient to make adequate progress toward understanding Nature. There is an important role for theoretical studies that aim to make sense of what is observed. Experiment and theory complement each other in seeking to understand Nature.

Many contemporary theories are so sophisticated that they can predict, with considerable confidence, certain physical phenomena which have yet to be measured, or that may even be impossible to measure. An example is the energy and angular correlations of emitted radiations, such as gamma rays and neutrons. Such information is essential for certain applications, *e.g.*, homeland security technologies. Nevertheless, with few exceptions, nuclear theories alone cannot provide the detail and numerical accuracy required for important physical parameters. Yet, experiments, by their nature, are imperfect. They are designed to determine specific numerical values for important physical properties that are as close to being correct and enduring as possible. However, these inherent values can never be determined exactly because the results from experimentation are inevitably "fuzzy" and, to varying degrees, biased from the truth. They are uncertain due to limitations of measurement procedures, as well as to statistical fluctuations that are uncontrollable. In this paper the term "uncertainty", rather than "error", is used to describe this phenomenon. The word "error" refers to a mistake or blunder, whereas "uncertainty" refers to imperfections associated with well-intentioned attempts to determine true values for the parameters of Nature.

Numerical values for physical parameters predicted by contemporary theories are also uncertain. However, uncertainties associated with predictions from theory, or the modeling of physical systems, are different from experimental uncertainties, even though they are often treated similarly in evaluating data used for practical applications. Theory and modeling yield well-defined numerical results, within the precisions provided by contemporary computational technology. Scientists who calculate the same physical quantity using the same methods, algorithms, input parameters, and computational tools, will likely produce numerical results that agree to a high degree of precision. That does not mean that they are accurate (*i.e.*, close to the true values)! The uncertainties are attributable to limitations of the theories, algorithms, and input parameter values employed for the calculations. Therefore, these uncertainties are not statistical in the usual sense of the term. Nevertheless, progress has been made recently toward estimating these uncertainties, and utilizing this information in practical applications [1].

More is known about the nature of experimental uncertainties than about theoretical uncertainties. A large body of literature exists on the subject, ranging from instructive guides to sophisticated mathematical treatments [2–5]. Still, experimenters frequently do not provide sufficient information about the uncertainties in their experiments to satisfy contemporary requirements. This may result from unawareness of the need for this information, or from a lack of practical understanding on how to go about estimating and specifying uncertainties. The topic of experimental data uncertainties is often given rather limited attention in academic curricula. Also, when nuclear experiments are performed with pure physics objectives in mind, detailed knowledge of uncertainties in the results may be of lesser importance than offering interpretations of the data in the context of theories. Nevertheless, these data are often compiled in databases where they may influence practical applications [6].

It is unfair to criticize experimenters for shortcomings in estimating and reporting uncertainties in their data without also assigning some of the blame to compilers who prepare the databases that serve as sources upon which data users have come to rely. As an example, a compiled cross-section data set for ^{55}Mn$(n,\gamma)^{56}$Mn (taken from the EXFOR library [6]) is shown in the following text. In this example, the line number designations that appear in files downloaded from EXFOR have been eliminated so that a larger font could be used to enhance readability of the essential information.

```
ENTRY            11130     860513
SUBENT        11130001     860513
BIB                  7          14
INSTITUTE  (1USATNC)
REFERENCE  (R,WADC-TN-59-107,59)
AUTHOR     (N.A.BOSTROM,I.L.MORGAN,J.T.PRUD'HOMME,P.L.OKHUYSEN,
           O.M.HUDSON JR)
TITLE      NEUTRON INTERACTIONS IN LITHIUM, CARBON, NITROGEN,
           ALUMINUM, ARGON, MANGANESE, YTTRIUM, ZIRCONIUM,
           RADIOLEAD AND BISMUTH.
INC-SPECT  INCIDENT ENERGY CORRECTED DUE TO RECALCULATION OF
            THRESHOLD.
STATUS     (SCSRS)
HISTORY    (760628T) TRANSLATED FROM SCISRS
           (800814A) CONVERTED TO REACTION FORMALISM
           (850117A) DELETED SANS 012,020,025.
           (860513A) BIB CORRECTION.
ENDBIB              14
NOCOMMON             0           0
ENDSUBENT           17
SUBENT        11130008     800814
BIB                  2           2
REACTION   (25-MN-55(N,G)25-MN-56,,SIG)
METHOD     (ACTIV)    ACTIVATION
ENDBIB               2
NOCOMMON             0           0
DATA                 4           2
EN         EN-RSL     DATA        DATA-ERR
MEV        MEV        B           B
 4.0    -02 8.     -03 3.9     -02 1.0     -02
 3.5    +00 4.     -01 1.7     -03 6.      -04
ENDDATA              4
ENDSUBENT           11
ENDENTRY             2
```

Although cross sections and uncertainties are given, there are few details upon which to judge the quality of these data. Perhaps the experimenters were thorough in documenting their work in the original report, but the EXFOR data compiler simply failed to digest this material into usable form in this EXFOR entry. Or, maybe the information was never available. The chances of a contemporary data user being able to locate the original report (issued in 1959) are small. Unfortunately, the situation illustrated here is relatively common.

This paper is organized as follows. First, the roles played by nuclear data uncertainties in nuclear science are discussed, with emphasis on a few contemporary nuclear applications. Next, the subjects of random variables and their statistical properties are examined, since measured nuclear parameters can be treated mathematically as random variables. A description of several probability functions that play roles in analyses of experimental nuclear data uncertainties is included. The process of identifying the primary parameters of experiments, and assessing their uncertainties, is explored, and simple examples are provided. The need to develop models of experiments that describe how derived physical quantities, such as cross sections, are extracted from what is actually measured is stressed. Characteristics of the uncertainties for various types of primary parameters that influence the outcome of experiments are described. Attention is devoted to documenting data uncertainties by means of reports, conference proceedings, journal articles, and computer libraries of numerical values. The main challenge is the large volume of numerical data that is often required to adequately document this information. Finally, contemporary formats used in compiling covariances for reported experimental data are discussed, and their use is illustrated by samples from the EXFOR library [6].

B. Roles for Nuclear Data Uncertainties

Some experimenters may not realize how detailed specification of uncertainty data can be beneficial in practice. A few examples are presented here to illustrate the importance of nuclear data uncertainties in applications.

1. *Nuclear data evaluations*

Applied nuclear data users seldom rely on numerical data extracted directly from original publications. Instead, they utilize evaluated (recommended) data. These are values distilled from consideration by experienced evaluators of the body of existing experimental and theoretical numerical results. Evaluators strive to produce the most reliable possible recommendations, based on critical assessments of the quality of data available from the reported scientific studies. This applies to both nuclear structure and nuclear reaction data (*e.g.*, reaction cross sections). The emphasis in this paper is on nuclear reaction data, in particular on reactions initiated by neutrons. It is an area of nuclear science where attention to data uncertainties currently appears to be the most concentrated. Hopefully, similar attention will be devoted in the future to uncertainties for other types of nuclear data (*e.g.*, nuclear structure and charged-particle reaction data). The principles and techniques discussed here are also applicable to these areas.

In earlier times, the evaluation of nuclear reaction data often involved drawing smooth eye guides through plots of measured experimental data extracted from the literature. The "error" bars (uncertainties) shown in these plots inevitably influenced evaluators, but only subjectively. This approach is no longer viewed as acceptable. More sophisticated and objective data evaluation procedures have evolved during the last several decades. Weights are assigned to nuclear reaction data sets, according to their perceived quality. Judgments on quality are based on both the uncertainties quoted by authors and evaluator impressions as to the reliability of these estimates. Rigorous statistical techniques are ultimately applied to produce the evaluated results [7–9]. If a data set is poorly documented, provides limited uncertainty information, or offers uncertainty values that appear to be unrealistic, it is likely to be rejected by evaluators. From the experimenter's point of view, the data would then have limited impact on the advancement of applied nuclear science. So, it is important for experimenters to provide reasonable and well-documented uncertainty data for their reported results, regardless of which physical parameters are measured. The responsibility of experimenters for assigning credible uncertainties is especially great for processes where, for practical reasons, it is likely that only a single set of measured data will be available to directly influence evaluations. An example would be measurements at neutron white-source facilities of total, scattering, capture, and fission cross sections in the resolved and unresolved resonance regions. Evaluated values of the resonance parameters for these energy regions usually emerge directly from the procedures used to analyze the measured data. They depend strongly on the uncertainties assigned to the raw experimental information [10].

The significant pieces of information which evaluators of reaction cross section data sets need to have can be summarized as follows: i) incident particle energies (usually for neutrons); ii) energy resolutions; iii) cross-section values; iv) total uncertainties in these cross sections; v) standard reactions used in the measurements; vi) actual values of the standard cross sections at the specified energies; and vii) estimates of the various uncorrelated and correlated uncertainty components included in the given total uncertainties (so that overall uncertainty correlations between the individual data points in the set can be assessed). As an alternative to items (iii) and (vi), values of the measured ratios of the unknown and specified standard reaction cross sections, at the various energies, can be provided. Experimenters should utilize well-established and documented standard reactions whenever the outcomes of their experiments depend upon reference cross sections rather than their being based on absolute measurements. Then, if there are changes in the recommended values for these standards, evaluators will be able to adjust the reported experimental data accordingly. Additional, important pieces of information for evaluation purposes are the sample properties and detector characteristics, as well as the procedures used for their determination. Failure to provide this bare minimum of information is likely to lead to the experimental data being rejected, improperly used, or not accorded the appropriate weightings that are merited in the evaluation process. Experiments are costly, time consuming, and difficult to conduct, so experimental data tend to be relatively limited in quantity compared with the copious results that are routinely generated from theoretical calculations. Since experimental data are precious, experimenters should be encouraged to properly document both measured values and the corresponding estimated

uncertainties for their work.

2. *Nuclear technology applications*

The uses that nuclear data find in applications, and the emphases placed on these data, evolve over time. It is very difficult to envision what uses the results of a measurement may find many decades hence. For example, consider the compiled data entry in EXFOR for ^{55}Mn$(n,\gamma)^{56}$Mn mentioned above. This cross section value was measured in 1959. More than fifty years later, it played a role in evaluating the cross section for a reactor dosimetry standard [11]. For illustrative purposes, three nuclear technologies are selected where evaluated neutron reaction data (including uncertainties) are currently playing important roles: neutron radiation dosimetry, nuclear criticality safety assessment, and the development and conceptual design of advanced nuclear reactors (mainly fast reactors). Many other important nuclear applications that benefit from reliable nuclear data, *e.g.*, accelerator shielding, medical science, space technologies, fusion energy, defense, homeland security, and nuclear forensics, could have been mentioned.

Neutron radiation dosimetry is an area where nuclear data uncertainties have played a significant role for many years. Neutron dose determinations are important in assessing the safety and life expectancies of nuclear power plants. In particular, extended exposure of reactor pressure vessels to neutron radiation can induce brittleness that threatens their structural integrity. Neutron reactions with various properties (energy thresholds, *etc*.) are used for dosimetry measurements. In addition to determining total neutron dosages, the integral data are used to adjust (formerly referred to as "unfold") a trial estimate of the spectrum shape that has been derived from theoretical modeling. This adjustment process, often carried out using the least-squares method, relies, in part, on specifying the uncertainties of the dosimeter integral responses and the uncertainties of the evaluated microscopic differential cross sections for the utilized dosimetry reactions [12].

Assurance of safe handling, transport, and storage of fissionable nuclear materials is a crucial issue in the chain of processes associated with assuring that nuclear fission is a safe and reliable energy option. The avoidance of criticality accidents is one of the most important objectives in this context [13]. Nuclear fission is viable as an energy source if a chain reaction is maintained within the reactor core, *i.e.*, if the system is critical. This critical point is achieved when the production and removal rates for neutrons within a system involving nuclear fission are in equilibrium. Assessment of the critical state of any nuclear energy system involving fissionable material centers upon consideration of a parameter referred to as k-eff. Equilibrium between the production and escape or absorption of neutrons in a system results in k-eff being exactly equal to unity. In sub-critical systems (k-eff < 1.000), the neutron inventory rapidly dies away when the driving source of neutrons is removed (or switched off). In a super-critical system (k-eff > 1.000), there is an uncontrollable chain reaction (explosion), even after the initiating neutron source is eliminated. The four-significant-figure precision shown here for k-eff emphasizes the extreme sensitivity of nuclear system behavior to departures of k-eff from unity. Insurance of sub-criticality is essential to insure criticality safety whenever materials that contain fissionable elements are handled, transported, processed, or stored outside of a reactor core [13]. Computer software is employed in the analysis of system criticality for a wide variety of system geometries [14]. These calculations utilize extensive evaluated nuclear data. Knowledge of the nuclear data uncertainties enables uncertainties in calculated values of k-eff to be estimated for these systems.

Uncertainty information for evaluated differential neutron reaction cross sections, and for measured integral quantities, is used in design and performance studies of conceptual advanced reactor systems [15]. The objective is to estimate uncertainties in calculated integral parameters such as k-eff (once again), burn-up and build-up of individual isotopic species during operation of the reactors (minor actinides, fission products, *etc*.), sodium void effects on reactivity, *etc*. Knowledge of nuclear data uncertainties enables determinations to be made of the design margins that are needed for safe and economical operation of these systems. This task can be accomplished in a couple of ways. First, uncertainties in evaluated nuclear data can be propagated directly to determine calculated system parameter uncertainties. A less direct approach is to use this uncertainty information to create adjusted nuclear data libraries for specific applications. The procedure involves merging general-purpose evaluated differential data and high-quality integral data that are pertinent to the system under investigation [15]. The latter approach enables system simulation calculations to be performed with greater accuracy than would be the case using evaluated nuclear data libraries based solely on differential data. Data libraries that have been adjusted to optimize their use for a particular nuclear system should not be used in applications other than the one intended due to the potential for generating biases. General-purpose evaluated libraries are based almost entirely on differential data. They can be utilized in a wide variety of nuclear applications with minimal concern about such biases [7–9].

II. BASIC CONCEPTS

A. Probability Distributions

In the physical sciences, it is assumed that the numerical values obtained from performing measurements of parameters are governed by the laws of probability and statistics. In particular, measurable physical parameters can be represented by random variables. Sometimes the term "stochastic variable" is used. This assumption enables nuclear parameters to be treated using mathematical laws. Random variables come in two forms: discrete and continuous. In the nuclear context, an example of a discrete random variable would be the spin "ψ" of a nuclear level. For example, when it is uncertain whether the value of that spin ψ should be either 1/2 or 3/2, but it is absolutely certain that there are no other possibilities, the decision might be to assume that $P(\psi = 1/2) = 0.5$ and $P(\psi = 3/2) = 0.5$. In other words, these two outcomes

are treated as equally likely. The sum of these two probabilities is unity (*i.e.*, certainty). Parameters such as cross sections are observed to be described by continuous random variables rather than discrete ones. Measurements of these quantities could produce an uncountable number of different values if it were possible to measure them an uncountable number of times. If this could be done, and if it were possible to keep track of the outcomes, they would be found to be distributed in predictable ways that can be described by continuous scalar functions called probability density functions.

This section focuses mainly on the general properties of continuously distributed parameters. Some mathematical concepts are discussed, first for single random variables, then for multiple random-variable sets. The emphasis is on practical aspects, *e.g.*, describing how experimenters should report their measured data (including uncertainties), and on how they should be compiled, rather than on rigorous statistical theory. Details are provided in this paper for only five single-variable (univariate) probability functions, and two multiple-variable (multivariate) functions, that are commonly encountered in nuclear applications. Certain other probability functions, *e.g.*, those that might be more appropriate in addressing specialized situations, such as small samples of data or rare events, are not treated here. They are encountered relatively infrequently in nuclear data applications. The interested reader can refer to the following three additional references for further insight on this subject: a BIPM document [2], and monographs by Smith [3] and Drosg [5].

1. *Single random variables*

With no loss of generality, let x be a single, continuous random variable that represents an arbitrary physical parameter. Let $p(x)$ be the probability density function that governs the statistical behavior of x. It is important to always distinguish conceptually between "x" as the random variable and "x" as a specific numerical value for this variable. In this paper it is assumed that this distinction can always be made based on context. Sampling (*i.e.*, measuring) x a large number of times (say K times, where K is a large integer), while keeping the experimental conditions under control to assure stability of the underlying sampling probability distribution, would generate a collection of K numerical values $\{x_k\}_K$, *i.e.*, $(x_1, x_2, \cdots, x_{k-1}, x_k, x_{k+1}, \cdots, x_{K-1}, x_K)$. The properties of this set are governed by the probability density function $p(x)$. We then define a derived real number $\langle x\rangle_K$ as

$$\langle x\rangle_K = \left(\sum_{k=1}^{K} x_k\right)/K. \tag{1}$$

Clearly, $\langle x\rangle_K$ is the well-known average of a finite collection of numbers $\{x_k\}_K$. In statistics it is called the sample mean value for the collection $\{x_k\}_K$. The actual numerical value $\langle x\rangle_K$ depends on K, as well as on the collection of sample values x_k. Two different sequences of K measurements of x will produce distinct averages (sample mean values) due to the stochastic nature of the random variable x, the finiteness of the sample set, and the nature of the measurement procedure used to generate the sample-value sets. Clearly, $\langle x\rangle_K$ is not a unique quantity even though $p(x)$ is well defined.

Next, define

$$v_{xK} = \left[\sum_{k=1}^{K} (x_k - \langle x\rangle_K)^2\right]/K. \tag{2}$$

The quantity v_{xK} is called the mean-square deviation from the mean for the collection $\{x_k\}_K$ or, more commonly, the variance of the collection $\{x_k\}_K$. The standard deviation of $\{x_k\}_K$ is closely related to v_{xK}, and it is defined by

$$s_{xK} = (v_{xK})^{1/2} = \left\{\left[\sum_{k=1}^{K} (x_k - \langle x\rangle_K)^2\right]/K\right\}^{1/2}. \tag{3}$$

The advantage of considering s_{xK} rather than v_{xK}, when discussing uncertainties, is that the dimensions (*i.e.*, units) of s_{xK} are identical to those of x_k and $\langle x\rangle_K$ (*e.g.*, they are measured in the units "barn" in the case of cross sections). Eqs. (1) - (2) define specific moments of the collection $\{x_k\}_K$. They are the most important moments for practical applications. The value K appearing in the denominators of Eqs. (2) and (3) is sometimes replaced by $K - 1$ in statistical analyses, for technical reasons [2]. As long as K is very large, the distinction is of little practical importance. The mean value $\langle x\rangle_K$ represents a central location around which the sampled values x_k tend to be encountered, while v_{xK} and s_{xK} measure the degree to which the values x_k tend to scatter or concentrate (peak) in the vicinity of $\langle x\rangle_K$. Thus, if a particular physical parameter is measured repetitively, under controlled conditions, the mean value and standard deviation can be treated as the "best value" and its corresponding "uncertainty". Alternative definitions are sometimes used by statisticians to represent best values and their uncertainties [2]. One example is the median value (*i.e.*, the value midway between two extreme possibilities) instead of the mean value. This paper considers only mean values and standard deviations. They are the most widely used in nuclear science applications.

When size K of the sample sets approaches infinity the sample moments of all these sets approach the same numerical values, namely those obtained from integrations that involve the function $p(x)$. Thus,

$$x_0 = \langle x\rangle = \int_{\Im} x\, p(x)\, \mathrm{d}x, \tag{4}$$

$$v_x = \left(\int_{\Im} x^2 p(x)\mathrm{d}x\right) - x_0^2, \tag{5}$$

$$s_x = v_x^{1/2}, \tag{6}$$

$$\int_{\Im} p(x)\mathrm{d}x = 1. \tag{7}$$

The integrations are shown symbolically as spanning the domain $\Im$ of all possible values of x that might be obtained from random sampling. Eq. (7) imposes the condition of normalization to unity on the probability density function $p(x)$. It should be noted that probability density is not a true probability, but

leads to one only as a consequence of its integration over a selected continuous range of possible values for x. Higher-order moments of p can also be defined. They correspond to integration of x^3 (third moment), x^4 (fourth moment), *etc*., weighted by $p(x)$, if these integrals over the domain $\mathfrak{S}$ are finite.

The five single-variable probability distribution functions considered in this paper are: Poisson distribution, continuous uniform distribution, normal (Gaussian) distribution, lognormal distribution, and chi-square distribution [2, 3, 5].

a. Poisson The Poisson distribution is a function of the discrete, non-negative integer random variable x. It takes the form

$$P(x) = \lambda^x \exp(-\lambda)/x! \quad (x = 0, 1, 2, \cdots, \infty;\ \lambda > 0). \quad (8)$$

Capital "P" is used to highlight the fact that it is a discrete function. The mean value is $x_0 = \lambda$ and the standard deviation is $s_x = \lambda^{1/2}$. Since the sum of terms $\lambda^x/x!$ (x from 0 to ∞) equals $\exp(\lambda)$, this function is normalized for all positive λ [3, 16]. Thus,

$$\sum_{x=0}^{\infty} P(x) = 1. \quad (9)$$

Fig. 1 is a plot of this distribution for two choices of parameter λ (Lambda): $\lambda = 4$ and $\lambda = 36$. When x is large, it is difficult to calculate λ^x or $x!$. Computational programs are available from the Internet to determine $P(x)$ for individual cases (*e.g.*, [17]), and computer software can also be employed. The normal (Gaussian) distribution can be used to approximate the Poisson function when $\lambda > 30$. Fig. 1 demonstrates this point. It also shows that the Poisson function is noticeably asymmetric for small values of λ, but it approaches symmetry as λ becomes larger. The underlying Poisson distribution mean value λ and uncertainty $\lambda^{1/2}$ are often assumed to be N and $N^{1/2}$, respectively, when N counts are observed in an experiment. If N is large, this is an adequate approximation. However, for small values of N, this approximation is unacceptable.

b. Continuous uniform Unlike the other probability distributions discussed in this paper, the continuous uniform distribution is unlikely to reflect the outcome from actual measurements of physical parameters. Nevertheless, it is still very useful for certain practical applications. It is defined as [3, 16]

$$p(x) = \begin{cases} 1/(b-a) & \text{if } a < x < b, \\ 0 & \text{otherwise.} \end{cases} \quad (10)$$

Parameters a and b are usually taken to be positive numbers. Any value of x within the interval (a, b) is equally likely to occur in sampling, while the probability of finding x outside these limits is zero. The mean value is $x_0 = (a + b)/2$. It is positive if a and b are positive. The standard deviation is $s_x = [(b-a)^2/12]^{1/2}$. This probability distribution is useful if it is suspected that the value of a particular random variable should be somewhere within a specific range.

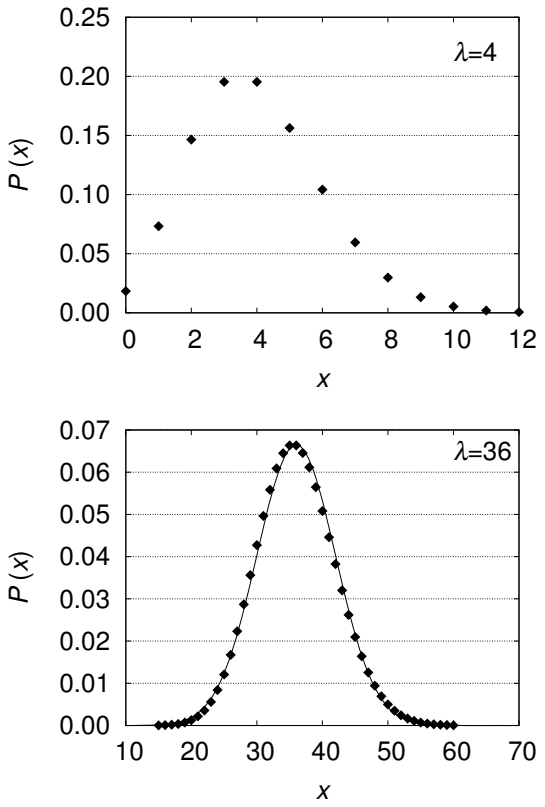

FIG. 1. Examples of the Poisson distribution for two distinct values of the parameter Lambda (λ). Top: $\lambda = 4$. Bottom: $\lambda = 36$. The corresponding normal distribution approximation for the case $\lambda = 36$ is also shown here as a smooth curve.

Consider this scenario: An experimenter might be tempted to suggest a value for a measured physical quantity, and then estimate that its uncertainty (*i.e.*, standard deviation) should be in the range 5% to 10%. Unfortunately, this unwillingness to quote a specific uncertainty value will confuse users of these data. They will have difficulty in assessing its quality. Should they assume 5%, or 10%, or perhaps even 7.5% (splitting the difference) as the uncertainty? The experimenter has an obligation to be more definite about this matter. A better approach is to suggest a likely range of possible values for the measured result, and stipulate a continuous uniform distribution. This satisfies a desire to be somewhat vague, yet it also offers a statistically acceptable option for data users. A well-defined mean value and standard deviation can then be deduced. A numerical example illustrates this point. Suppose the experimenter believes that a measured parameter x should have a value somewhere in the range 16 to 20 ($a = 16$, $b = 20$). This assumption produces a mean value $x_0 = 18$ and standard deviation $s_x = 1.15470$ (to six significant figures), corresponding to a fractional uncertainty of $\approx 6.4\%$.

c. Normal (Gaussian) The normal (Gaussian) function is the most widely employed statistical distribution in nuclear

applications [3, 16]. It is a continuous function of x that involves two parameters, ν and s_x. It takes the form

$$p(x) = \exp\left[-(x-\nu)^2/(2s_x^2)\right]/(2\pi s_x^2)^{1/2}. \quad (11)$$

The random variable x can assume any real number. The mean value is $x_0 = \nu$. It can also be any real number. The standard deviation, s_x, must be a positive real number. Function $p(x)$ in Eq. (11) is normalized, since integration from $x = -\infty$ to $+\infty$ yields unity. This function is symmetric about its mean value, and it is often referred to as the "bell-shaped curve" because of its characteristic shape (see Fig. 1). In nuclear applications, the convention that measured parameters x are positive is generally adopted. This tends to restrict practical use of this distribution to cases where $x_0 = \nu > 0$ and $s_x \ll x_0$, *i.e.*, when the fractional uncertainty $f_x = (s_x/x_0)$ is fairly small. In practice, the normal distribution can be useful for most applications in nuclear science when $(s_x/x_0) < 0.3$ [18, 19]. The likelihood of obtaining negative values in random sampling will then be minimal, without having to artificially truncate the sampling space.

The Central Limit Theorem states that the normal distribution governs random events that are the consequences of relatively large numbers of small, unrelated linear disturbances (*e.g.*, [2, 3]). This property can be demonstrated by examining the result of dropping a large number of balls from a single location through a maze of uniformly distributed barriers. The balls bounce randomly as they travel through the maze. The distribution of accumulated balls below the maze approaches a normal distribution centered on the drop point, when the number of balls dropped becomes very large (*e.g.*, see the following website for an animated demonstration of the phenomenon: http://www.squadron13.com/games/balldrop/balldrop.htm). Other distributions mentioned in this paper, *i.e.*, the Poisson, lognormal, and chi-square distributions, approach the normal distribution under limiting conditions. The Maximum Entropy Principle, introduced by Shannon [20], and further elaborated by Jaynes [21], states that if all one knows about the statistical properties of a random variable are estimates of the mean value and standard deviation, the best assumption about the unknown probability distribution is that it is likely to be normal. An important exception is when the fractional standard deviation is large (see above). Then, the lognormal distribution is more appropriate.

d. Lognormal The lognormal function is defined for positive real number random variables x as [3, 16]

$$p(x) = \exp\left[-(\ln x - \nu)^2/(2\gamma^2)\right]/(2\pi\gamma^2 x^2)^{1/2}. \quad (12)$$

Parameter ν can be any real number, while γ must be positive. The mean value x_0 and standard deviation s_x for the lognormal function are

$$x_0 = \exp\left[\nu + (\gamma^2/2)\right], \quad (13)$$

$$s_x = \left[\exp(2\nu + 2\gamma^2) - \exp(2\nu + \gamma^2)\right]^{1/2}. \quad (14)$$

This function is asymmetric for large values of the standard deviation relative to the mean value. In the limit of parameter choices for ν and γ such that $s_x \ll x_0$, the lognormal distribution approaches the symmetric normal distribution. This is illustrated in Fig. 2. Furthermore, if a transformation is made from a lognormally distributed random variable x to random variable y, where $y = \ln x$ and "ln" signifies the natural logarithm function, y will be normally distributed. Then, ν is the mean value of y and γ is its standard deviation. Since x ranges from 0 to $+\infty$, y ranges from $-\infty$ to $+\infty$. The lognormal distribution can be useful in representing random variables with large uncertainties. Negative values are never encountered in random sampling of x, even for very large standard deviations. This can occur for some parameters in nuclear astrophysics [19]. It can be demonstrated that any random variable which is the consequence of a sequence of a relatively large number of small, independent, positive, multiplicative random disturbances will be lognormally distributed [22]. This can be understood as follows. Suppose that a variable x is represented by a product (Π) of J variables ω_j, *i.e.*, $x = \Pi_{j=1,J}\,\omega_j$. Then, $y = \ln x = \Sigma_{j=1,J} \ln \omega_j$. Thus, multiplicative disturbances in ordinary space become additive disturbances in logarithm space.

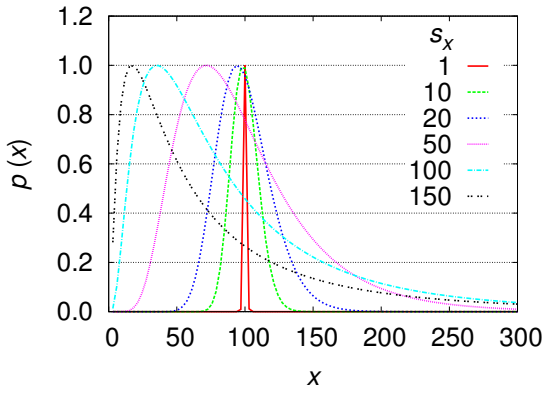

FIG. 2. Comparison of shapes of the lognormal distribution for $x_0 = 100$ and $s_x = 1$, 10, 20, 50, 100, and 150 [19]. These distributions are normalized so that $p(x_m) = 1$; where $x_m = \exp(\nu - \gamma^2)$ is the mode of the distribution.

e. Chi-square The chi-square (χ^2) probability distribution plays a special role in statistics that differs in several respects from the other distributions mentioned here [3, 16]. It is applied in tests of goodness of fit, or assessments of the significance of observed differences (scatter) among individual experimental and/or theoretically calculated data, relative to a stated hypothesis that these data should be normally distributed, with a specified set of distribution parameters. The chi-square distribution is a member of the gamma distribution family [3]. For all positive real numbers x, and integers $n = 1, 2, 3, \cdots$, the normalized chi-square probability density function is given by

$$p(x) = x^{[(n/2)-1]} \exp(-x/2)/\left[2^{(n/2)}\Gamma(n/2)\right], \quad (15)$$

$$\Gamma(n/2) = \int_0^\infty t^{[(n/2)-1]} \exp(-t)\, dt. \quad (16)$$

$\Gamma(n/2)$ represents the value of the gamma function $\Gamma(\alpha)$ for the argument $\alpha = n/2$. The mean value is $x_0 = n$ and the standard deviation is $s_x = (2n)^{1/2}$. The integer parameter n denotes degrees of freedom [3]. Values for the gamma function cannot be expressed in closed form for arbitrary values of the argument $\alpha > 0$, so it must be calculated numerically. Computer software or on-line computational programs can be used to perform this task (*e.g.*, [23]). In the limit of very large n, the chi-square distribution approaches the normal distribution, also as a consequence of the Central Limit Theorem [3].

The cumulative probability $P(\beta; n)$ for the chi-square probability density function $p(x)$, and a specific real number $\beta > 0$, must be considered in applying the chi-square test to data sets. $P(\beta; n)$ is

$$P(\beta; n) = \int_0^{\beta} p(x)\mathrm{d}x. \tag{17}$$

It is the probability, for a particular data set, of encountering a value x (called the test statistic) $< \beta$. Since probability must be normalized, $0 < P(\beta; n) < 1$. The certain event, $0 < x < +\infty$, has a probability of unity. The number "β" is often denoted by "χ^2" (chi-square). The distinction between parameter and distribution name must be based on context. In this paper, "β" is employed exclusively to avoid any confusion regarding this matter. In practical applications, it is convenient to consider the quantity $Q(\beta; n) = 1 - P(\beta; n)$. $Q(\beta; n)$ is the probability of encountering a value of $x > \beta$. Plots of $Q(\beta; n)$, as a function of β, and for various choices of n, are given in Fig. 3 (Top). Eqs. (15) and (16), as well as Fig. 3, indicate that the values of $P(\beta; n)$ and $Q(\beta; n)$ for a particular choice of β depend strongly on the degree-of-freedom parameter n. For $\beta \ll n$, $P(\beta; n) \ll 1$. For $\beta \gg n$, $Q(\beta; n) \ll 1$. The probabilities of encountering values of β either much smaller, or much larger, than the mean value n are small. This is true regardless of the value of n. Tables of values for $P(\beta; n)$ or $Q(\beta; n)$ are sometimes provided only for $n = 1$, *i.e.*, $P(\beta; 1)$ or $Q(\beta; 1)$, to save space. In some cases, the information provided in these tables can be applied for other values of n, based on the assumption $P(\beta; n) \approx P(\beta/n; 1)$ or $Q(\beta; n) \approx Q(\beta/n; 1)$. The validity of these assumptions can be judged from the plots shown in Fig. 3, in particular by examining the bottom plot which shows the ratio $Q(\beta/n; 1)/Q(\beta; n)$ as a function of β/n. The assumption that this ratio should be approximately unity holds only for assumed values of β in the vicinity of parameter "n" or, alternatively, for $\beta/n \approx 1$. Since computational algorithms are readily available these days to produce values of $P(\beta; n)$ or $Q(\beta; n)$, for arbitrary β and n, it is advisable to avoid using this approximation when applying the chi-square test. The degree-of-freedom parameter "n" equals the number of data points N being tested for consistency, but only if these data are independent, *i.e.*, if there are no constraints involved. Otherwise, the degrees of freedom must be reduced by the number of constraints ν, *i.e.*, $n = N - \nu$. The following example illustrates the chi-square test.

Example 1 The stated hypothesis is that a particular collection of independent sampling data should be described by a univariate normal distribution with mean value x_0 and standard deviation s_x. A collection of n statistically independent sampled data values, $\{x_i\} = (x_1, x_2, \cdots, x_i, \cdots, x_n)$, is produced (*e.g.*, from an experiment). These data are then examined for consistency with the stated hypothesis. The appropriate parameter β is

$$\beta = \sum_{i=1}^{n} \left[(x_i - x_0)^2 / s_x^2\right]. \tag{18}$$

Eq. (18) is an example of a quadratic form. Quadratic forms play important roles in applied mathematics, including chi-square data consistency tests such as the present one. In this example, there are n degrees of freedom, since there are n independent sample values with no constraints. If $\beta \approx n$ (the mean value for the chi-square distribution) and $Q(\beta, n) \approx 0.5$, then the data set $\{x_i\}$ is said to be reasonably consistent with the hypothesis. If $\beta \ll n$, and $Q(\beta, n)$ is fairly close to 1, the scatter of the sampled values x_i relative to x_0 is far less than might be anticipated, based on the hypothesis. If $\beta \gg n$, and $Q(\beta, n)$ is fairly small, the scatter of the various x_i relative to x_0 is far greater than would be expected, based on the assumed hypothesis. In the two latter cases, there is cause to suspect that the sampling data do not support the stated hypothesis. It would then be appropriate to first propose different values of x_0 and s_x for the normal distribution that defines the hypothesis. Alternatively, least-squares procedures could be used to try and optimize these parameters [24]. However, taking this last step would require considering a degree-of-freedom value n smaller than the data set, since data fitting introduces constraints.

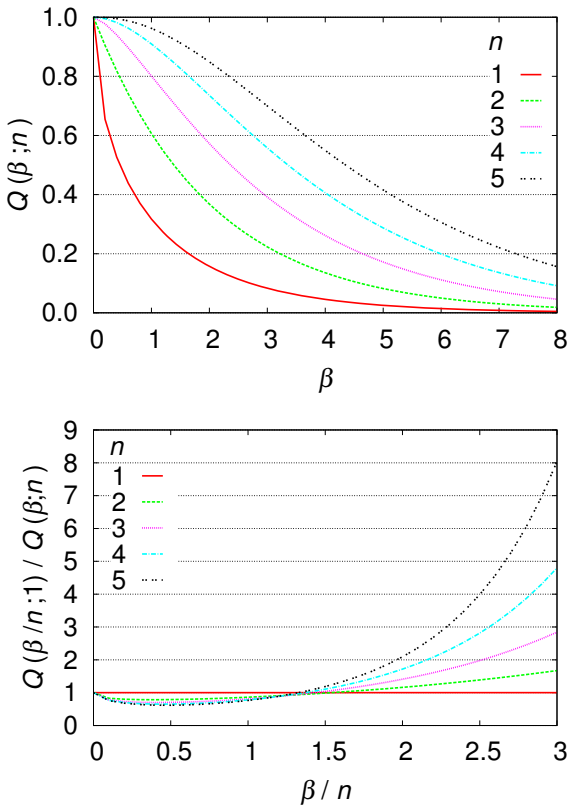

FIG. 3. Top: Curves correspond to $Q(\beta; n) = 1 - P(\beta; n)$ versus β, for various n. Bottom: Curves correspond to $Q(\beta/n; 1)/Q(\beta; n)$ versus β/n, for various n.

If the test results are still inconsistent, based on the computed value of β, the sampled data are probably not normally distributed. It is instructive to consider a simple numerical example involving five independently sampled values: $x_1 = 105$, $x_2 = 97$, $x_3 = 101$, $x_4 = 95$, and $x_5 = 111$. The hypothesis (stated prior to the sampling) is that $x_0 = 100$ and $s_x = 4$. When all five sample values are considered, Eq. (18) yields $\beta = 11.3125$. Referring to the plot of $Q(\beta; n)$ in Fig. 3, it is seen that it is very unlikely ($Q(\beta; 5) < 0.1$) that these five values support the stated hypothesis regarding the mean value and standard deviation for sampling data that are assumed to be normally distributed. However, if x_5 is eliminated, and the chi-square test is repeated, Eq. (18) yields $\beta = 3.75$. This is a reasonable outcome from this test that supports the given hypothesis. It would appear quite likely that x_5 is a discrepant outlier, while the other sampled values collectively support the indicated hypothesis.

More complicated expressions for β (generally of the quadratic form variety) are used in chi-square testing for multiple-variable situations, and for those cases that involve correlated (*i.e.*, constrained) data. Extensions to this theory to deal with such situations are not explored in this paper.

The chi-square test of statistical significance is strictly valid only for normally distributed data, although it is frequently applied rather indiscriminately by investigators. What often happens is that a particular data point included in the chi-square test appears to be an outlier (as illustrated in Example 1). This outlier can contribute to failure of the data set to be consistent with a normal distribution. It is usually suspected that this might be due to an error on the part of the experimenter. However, statistical tests cannot prove that a particular data point is wrong. They can only serve to suggest that it is more likely that it is wrong than correct, based on the information available to the analyst at the time the chi-square test is performed. Of course, there is always a chance, albeit seemingly small, that the apparently discrepant point is actually closer to the correct value than the others if, unbeknownst to the analyst, or possibly even to the experimenters, an important factor that was not considered in the majority of the measurements, was taken into account in the experiment that produced the outlier. Discrepancies of this nature do tend to become resolved in the long term, but they can plague data analysts for many years before the underlying cause of the problem is eventually discovered.

2. *Multiple random variables*

Texts dealing with probability, statistics, and data uncertainties often focus on treating single variables, with limited mention made of multiple variable sets (*e.g.*, [2]). Nuclear experimenters generally measure and report more than the value and uncertainty for just a single physical parameter, *e.g.*, they might report a set of cross sections measured at several different neutron energies. It is actually relatively straightforward to extend the discussion of the statistical properties of single random variables to collections of random variables. Assume that n distinct variables $\{x_i\} = (x_1, x_2, \cdots, x_i, \cdots, x_n)$ are to be considered. This should not be confused with n samples of a single random variable, as discussed earlier. This collection can be represented by the n-dimensional vector $\boldsymbol{x}$. Bold-face symbols are used in this paper for vectors and matrices.

We let $p(\boldsymbol{x})$ denote a normalized scalar probability density function of n continuous random variables. The mean value for each of the n random variables is

$$x_{0i} = \langle x_i \rangle = \int_{\mathfrak{S}} x_i \, p(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \quad (i = 1, n). \tag{19}$$

The vector of mean values x_{0i} is $\boldsymbol{x}_0$; it is referred to as the mean-value vector. Integration encompasses an n-dimensional domain $\mathfrak{S}$ inclusive of the ranges of all the random variables in the array $\boldsymbol{x}$. The quantity "$\mathrm{d}\boldsymbol{x}$" is an n-dimensional differential volume element (voxel) in this space. Integration of $p(\boldsymbol{x})$ over domain $\mathfrak{S}$ yields unity. The corresponding variances are

$$v_{xii} = \int_{\mathfrak{S}} x_i^2 \, p(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} - x_{0i}^2 \quad (i = 1, n). \tag{20}$$

The significance of the double subscript "ii" used in labeling these variances is clarified below. The standard deviation for each of the random variables is

$$s_{xi} = (v_{xii})^{1/2} \quad (i = 1, n). \tag{21}$$

Additional second-order moments can be defined for the probability function $p(\boldsymbol{x})$. They are

$$v_{xij} = \int_{\mathfrak{S}} x_i \, x_j \, p(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} - x_{0i} \, x_{0j} \quad (i, j = 1, n; \; i \neq j). \tag{22}$$

The v_{xij} are known as the covariances for the probability function $p(\boldsymbol{x})$. The elements defined by Eqs. (20) and (22) form a symmetric $n \times n$ matrix $\boldsymbol{V}_x$, called the covariance matrix [3]. The symmetric $n \times n$ matrix $\boldsymbol{C}_x$, with elements c_{xij}, is the correlation matrix. Its elements are

$$c_{xij} = v_{xij} / (v_{xii} \, v_{xjj})^{1/2} \quad (i, j = 1, n). \tag{23}$$

The diagonal elements of $\boldsymbol{C}_x$ are unity and the off-diagonal elements must satisfy the requirement

$$-1 < c_{xij} < +1 \quad (i, j = 1, n; \; i \neq j). \tag{24}$$

Eq. (24) is related to the Cauchy-Schwarz inequality [16]. This paper adopts the convention of specifying only the elements of the lower triangle for $\boldsymbol{V}_x$ or $\boldsymbol{C}_x$, with the understanding that the upper triangle is a mirror image. The dimensions of the covariance matrix elements are equal to squares of the variables, while the correlation matrix elements are dimensionless. This paper addresses only those properties of covariance matrices that are essential to understanding the relationship between the mathematical concepts and actual uncertainties of physical parameters. Information on additional mathematical properties of covariances matrices can be found in the literature (*e.g.*, [3]). Scientific investigators frequently consider uncertainties, but fail to consider the effects of nonzero correlations. When data sets are partially correlated, there is a degree of "stiffness" that can strongly influence their impact in applications.

Two specific multivariate probability distributions are discussed in this paper: the multivariate normal distribution and the multivariate lognormal distribution.

a. Multivariate normal This distribution is widely employed in nuclear science (*e.g.*, [3]), *e.g.*, it is the basis for many contemporary nuclear data evaluation schemes (*e.g.*, [25]). The normalized multivariate normal distribution is

$$p(\boldsymbol{x}) = (2\pi)^{-n/2}\left[\det(\boldsymbol{V}_x)\right]^{-1/2} \times \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}_0)^{+}\boldsymbol{V}_x^{-1}(\boldsymbol{x}-\boldsymbol{x}_0)\right]. \quad (25)$$

The mean-value vector is $\boldsymbol{x}_0$ and the covariance matrix is $\boldsymbol{V}_x$. The superscript symbol "+" signifies matrix transposition. The inverse matrix $\boldsymbol{V}_x^{-1}$ will exist if $\boldsymbol{V}_x$ is positive definite [3].

b. Multivariate lognormal The relationship between the multivariate lognormal and multivariate normal distributions is conceptually similar to that for their single-variable counterparts (*e.g.*, [26]). This distribution is applied in economics, medical science, and atmospheric studies (*e.g.*, see [27, 28]), but it is not very commonly seen in nuclear science applications. The transformation between corresponding normally and lognormally distributed multiple variable sets follows from the normalization of probability, as in the univariate case, but the mathematics is more involved (*e.g.*, see [26, 27]). The multivariate lognormal distribution for an n-dimensional variable set $\boldsymbol{x}$ is (*e.g.*, see [27])

$$p(\boldsymbol{x}) = (2\pi)^{-n/2}\left[\det(\boldsymbol{U})\right]^{-1/2}\left(\prod_{i=1}^{n} x_i^{-1}\right) \times \exp\left[-\frac{1}{2}(\mathbf{ln}\,\boldsymbol{x}-\boldsymbol{\nu})^{+}\boldsymbol{U}^{-1}(\mathbf{ln}\,\boldsymbol{x}-\boldsymbol{\nu})\right]. \quad (26)$$

Matrix $\boldsymbol{U}$ must be nonsingular in order for its inverse $\boldsymbol{U}^{-1}$ to exist. The notation "$\mathbf{ln}\,\boldsymbol{x}$" represents the n-dimensional array $(\ln x_1, \cdots, \ln x_i, \cdots, \ln x_n)$. The n-dimensional vector $\boldsymbol{\nu}$ (with elements ν_i) and $n \times n$ matrix $\boldsymbol{U}$ (with elements u_{ij}) appearing in Eq. (26) do not correspond to the mean-value vector $\boldsymbol{x}_0$ and covariance matrix $\boldsymbol{V}_x$ for the multivariate lognormal distribution. The elements of $\boldsymbol{x}_0$ and $\boldsymbol{V}_x$ are given by (*e.g.*, see [29])

$$x_{0i} = \exp\left[\nu_i + (u_{ii}/2)\right], \quad (27)$$

$$v_{xij} = \exp\left\{\nu_i + \nu_j + \left[(u_{ii}+u_{jj})/2\right]\right\}\left[\exp(u_{ij}) - 1\right]. \quad (28)$$

3. *Confidence intervals*

It is often desired to know the absolute probability $P(x_{\rm low}, x_{\rm high})$ that random sampling of a single-variable (*i.e.*, univariate) probability function will produce a value x within the interval $(x_{\rm low}, x_{\rm high})$. Clearly,

$$P(x_{\rm low}, x_{\rm high}) = \int_{x_{\rm low}}^{x_{\rm high}} p(x)\mathrm{d}x \quad (29)$$

for a continuous probability density function p. That probability depends on the range $(x_{\rm low}, x_{\rm high})$ and the characteristics of function p. This probability is unity for the entire sampling space, since p is assumed to be normalized. Confidence intervals correspond to particular ranges of x in the vicinity of the mean value x_0. For symmetric distributions they involve upper and lower limits located equidistant above and below the mean value.

For the continuous uniform distribution, as defined by Eq. (10), the probability of sampling x within one standard deviation from the mean value, *i.e.*, in the interval $(x_0 - s_x, x_0 + s_x)$, is always 0.577350 (to six significant figures), independent of the parameters a and b. However, the fractional uncertainty $f_x = (s_x/x_0)$ can be very small when $(b-a) \ll x_0$. Therefore, it can be quite useful to apply this distribution in estimating the impact of uncertainties assumed for certain physical parameters. One advantage is that sampling of a variable governed by this distribution can be performed easily using an ordinary random number generator. It is worth noting that larger confidence intervals, *e.g.*, two-sigma or greater, are meaningless for this distribution, since they extend to regions of zero probability.

Confidence intervals associated with the normal distribution are of great interest for statistical applications in nuclear science. The normal distribution is symmetric, so these confidence intervals are centered on the mean value. The regions above and below these intervals involve equal probabilities that samples will not be drawn from within the confidence intervals. Table I shows probabilities for the univariate normal distribution that correspond to one-, two-, three-, and four-sigma confidence intervals.

TABLE I. Confidence interval probabilities for a single-variable (univariate) normal distribution (to six significant figures). The values were obtained from [30].

Confidence Interval	Probability
$(x_0 - s_x, x_0 + s_x)$	0.682689
$(x_0 - 2s_x, x_0 + 2s_x)$	0.954500
$(x_0 - 3s_x, x_0 + 3s_x)$	0.997300
$(x_0 - 4s_x, x_0 + 4s_x)$	0.999937

Since the lognormal distribution is asymmetric, and approaches symmetry only in the limit of very small standard deviations (see Fig. 2), any interval centered about the mean value will not result in equal sampling probabilities above and below such an interval. In order to observe equal probabilities above and below the interval, corresponding to nonoccurrence of a designated event, it is necessary that the interval not be symmetrically situated with respect to the mean value. Uncertainty bars for lognormal distributions are asymmetric, with unequal upper and lower lobes. If an experimenter believes the reported results ought to correspond to sampling from a lognormal distribution, then uncertainty bars with unequal upper and lower lobes relative to the mean values should be provided. If experimental data are quoted with uncertainty bars having unequal upper and lower lobes, but no explicit specification is made as to the underlying probability distribution, there is a good possibility that these data are lognormally distributed. However, since there are other asymmetric probability distributions, it would be wrong to automatically assume that the lognormal distribution applies. Further details on this point can be found in the literature, along

with formulas that define the uncertainty bars for various confidence levels, when the distribution is lognormal (*e.g.*, [31]).

The concept of confidence intervals can be applied in a similar manner to multivariate probability functions. These intervals are not as simple to visualize as in the univariate case, nor are they as commonly considered. For example, one could define a region of vector space near the mean-value vector $\boldsymbol{x}_0$, and integrate the probability density function over this region. This volume could be quite arbitrary as long as it is fully connected topologically. In particular, one could choose to define the region in such a manner that each random variable is limited to excursions of no more than one standard deviation from its mean value.

Investigators should strive to estimate uncertainties for all their measured parameters that correspond to the same level of confidence. It is usual to assume that experimental data correspond to sampling from univariate or multivariate normal distributions, and that the estimated uncertainties reflect a one-sigma level of confidence [2, 3]. Then, roughly 68% of the measurements of these parameters are likely to yield results falling within one-sigma confidence intervals (according to Table I). Unfortunately, this guidance is often ignored by experimenters. This can lead inadvertently to overestimation of some uncertainty components and underestimation of others.

B. Covariance Matrices and Correlations

This section discusses several properties of covariance matrices, and describes a few useful techniques for dealing with them. The reader can refer to the literature for more information on this topic (*e.g.*, [3, 16]).

1. *Features of covariance matrices*

a. Uncertainties and correlations An element of covariance matrix $\boldsymbol{V}_x$ can be expressed as $v_{xij} = s_{xi}\, c_{xij}\, s_{xj}$, where s_{xi} and s_{xj} are standard deviations and c_{xij} is an element of the correlation matrix $\boldsymbol{C}_x$. Matrix algebra formalism can also be used, *i.e.*, $\boldsymbol{V}_x = \boldsymbol{S}_x^{+}\boldsymbol{C}_x\boldsymbol{S}_x$. Symbol "$^+$" signifies matrix transposition. $\boldsymbol{S}_x$ is a $n \times n$ diagonal matrix with elements $(\boldsymbol{S}_x)_{ij} = \delta_{ij}\, s_{xi}$, and δ_{ij} is the Kronecker Delta (equal to 1 if $i = j$ and 0 otherwise) [3]. Since $\boldsymbol{S}_x$ is diagonal, $\boldsymbol{S}_x^{+} = \boldsymbol{S}_x$. So, the matrix transposition notation is not needed. There are advantages to this representation, *e.g.*, as shown in the following example.

Example 2 Suppose that $\boldsymbol{V}_x$ is a covariance matrix which has an inverse $\boldsymbol{V}_x^{-1}$. From the rules of matrix algebra [3],

$$\boldsymbol{V}_x^{-1} = (\boldsymbol{S}_x\boldsymbol{C}_x\boldsymbol{S}_x)^{-1} = \boldsymbol{S}_x^{-1}(\boldsymbol{S}_x\boldsymbol{C}_x)^{-1} = \boldsymbol{S}_x^{-1}\boldsymbol{C}_x^{-1}\boldsymbol{S}_x^{-1}. \tag{30}$$

Since $\boldsymbol{S}_x$ is diagonal, its inverse exists, with elements δ_{ij}/s_{xi}. The s_{xi} are non-zero since all reasonable standard deviations should be positive. The elements of $\boldsymbol{V}_x^{-1}$ are therefore given by $(\boldsymbol{V}_x^{-1})_{ij} = (\boldsymbol{C}_x^{-1})_{ij}/(s_{xi}\, s_{xj})$. So, it may be advantageous for the elements of $\boldsymbol{V}_x^{-1}$ to be determined by inverting $\boldsymbol{C}_x$ and dividing the elements of $\boldsymbol{C}_x^{-1}$ by products of two standard deviations.

b. Relative covariance matrices It is often useful to work with the relative covariance matrix $\boldsymbol{R}_x$. It is constructed from the covariance matrix elements v_{xij} or, alternatively, the correlation matrix elements c_{xij} and the standard deviations s_{xi} and s_{xj}, along with the mean values x_{0i} and x_{0j}. The elements of $\boldsymbol{R}_x$ are

$$r_{xij} = v_{xij}/(x_{0i}\, x_{0j}) = (s_{xi}/x_{0i})\, c_{xij}\, (s_{xj}/x_{0j}) = f_{xi}\, c_{xij}\, f_{xj}. \tag{31}$$

The factor $f_{xi} = (s_{xi}/x_{0i})$ is the fractional uncertainty in x_i. Relative covariance matrices are useful in tabulating uncertainty data in documents and compiled experimental or evaluated nuclear data libraries [6, 32, 33].

c. Covariance matrix properties Standard deviations should always be positive. If s_{xk} were zero that would suggest that the variable x_k is perfectly known. This cannot happen in reality. However, when a physical parameter is considered to have negligible uncertainty in a particular context (*e.g.*, the speed of light), compared with the other variables, it is reasonable to treat it as a fixed constant of the experiment, like the numbers "2" or "π" (pi). It should be excluded from the variable set $\boldsymbol{x}$. A covariance matrix $\boldsymbol{V}_x$ that represents uncertainty information for a measured data set $\{x_i\}$ with dimension n should generally be nonsingular. It should have rank n, not $n-1$, nor $n-2$, *etc*. If $\boldsymbol{V}_x$ has rank n, its determinant will satisfy the inequality $\det(\boldsymbol{V}_x) > 0$, and it will be positive definite [3]. If the rank of $\boldsymbol{V}_x$ is less than n, the determinant will not be positive, $\boldsymbol{V}_x$ will be singular, and its inverse $\boldsymbol{V}_x^{-1}$ will not exist. One consequence is that the multivariate normal probability function, which contains the factor $[\det(\boldsymbol{V}_x)]^{1/2}$ in the denominator, and requires existence of the inverse $\boldsymbol{V}_x^{-1}$, is ill defined. So, experimenters should insure that covariance matrices that represent uncertainties for their data are positive definite, unless there is a reason why this should not be the case, *e.g.*, when constraints are mandated by the nature of the variable set. Relationships will then exist between the x_i, and this signifies redundancy [3].

Unfortunately, covariance information provided by data measurers and evaluators, even when properly generated, can be corrupted by limitations of the formats used to represent these data in libraries, or by the manner in which users manipulate these data [3]. Problems can also arise due to numerical precision issues. Discussion of various ways that these anomalies can develop is beyond the scope of this paper. Nevertheless, measurers and evaluators should be aware of this issue. Fortunately, quality assurance procedures used to test candidate evaluations submitted to the ENDF/B library do test for such effects (*e.g.*, [34]).

Uncertainty correlations tend not to be well understood by many nuclear researchers. The following example provides some insight as to their meaning.

Example 3 Consider two random variables, x_1 and x_2, with standard deviations s_{x1} and s_{x2}, and one distinct off-diagonal correlation coefficient c_{x21}. Collectively, they form a covariance matrix $\boldsymbol{V}_x$. Since $-1 < c_{x21} < +1$, it is instructive to represent c_{x21} by the cosine function, since it spans the same numerical range. Thus, $c_{x21} = \cos\theta$. If

$\theta = 90°$, $c_{x21} = 0$ and x_1 and x_2 are said to be orthogonal (*i.e.*, completely independent). If $\theta = 0°$, $c_{x21} = 1$ and x_1 and x_2 are fully correlated. Other values of the angle θ between 0° and 90° correspond to partial positive correlation. Anti-correlations (negative correlations) can be represented by values of θ in the range 90° to 180°. Next, envision two new random variables, y_1 and y_2, that are linear combinations of x_1 and x_2. They have standard deviations, s_{y1} and s_{y2}, corresponding to a new covariance matrix $\boldsymbol{V}_y$. We choose to require that y_1 and y_2 be uncorrelated (*i.e.*, completely independent), so the off-diagonal correlation coefficient is $c_{y21} = 0$. Such a transformation, from variables $\boldsymbol{x}$ to $\boldsymbol{y}$, is called a linear orthogonal transformation [35]. This is possible if $\boldsymbol{V}_x$ is positive definite, and this is assured if c_{x21} satisfies the conditions stated above, *i.e.*, $0° < \theta < 90°$ or $90° < \theta < 180°$. For two variables, this transformation can be seen as a rotation of the coordinate system. For larger random-variable sets, such a simplistic geometric picture cannot be provided. However, linear orthogonal transformations are possible for all dimensions $n \geq 2$ as long as $\boldsymbol{V}_x$ is positive definite.

Diagonalization of a positive definite $n \times n$ matrix $\boldsymbol{V}_x$, to produce an $n \times n$ diagonal matrix $\boldsymbol{V}_y$, is accomplished by a linear orthogonal transformation of variables from $\boldsymbol{x}$ to $\boldsymbol{y}$. In general terms, $\boldsymbol{V}_y = \boldsymbol{P}^{-1}\boldsymbol{V}_x\boldsymbol{P}$ where $\boldsymbol{P}$ is the $n \times n$ matrix that performs the transformation. As long as an invertible matrix $\boldsymbol{P}$ can be found, $\boldsymbol{V}_x$ is said to be diagonalizable. Otherwise, $\boldsymbol{V}_x$ is defective [35]. The diagonal elements of $\boldsymbol{V}_y$ are the eigenvalues of $\boldsymbol{V}_x$. If a covariance matrix is found to be defective, the problem is usually due to flaws in the method used to generate it, or possibly to numerical precision issues. Since the diagonal elements of $\boldsymbol{V}_x$ correspond to variances, *i.e.*, squares of total uncertainties estimated by the experimenter, they are unlikely to be problematic. That leaves the correlations as the most likely candidates for producing defective covariance matrices. It is advisable to derive the correlation matrix $\boldsymbol{C}_x$ and inspect the off-diagonal elements as a first step toward resolving this matter.

2. *Dealing with covariance matrices*

a. Large covariance matrices Covariance matrices can be very large when large data sets are involved. A covariance matrix that provides uncertainty information for n different random variables has the dimension $n \times n$, so there are n^2 elements. Since covariance matrices are symmetric, there are actually only $n(n+1)/2$ distinct elements. This can still be a large number if n is large, as illustrated in the following example.

Example 4 A recent evaluation of neutron reaction data for ^{235}U considered 3,193 resonances for (n,γ) and (n,f) reactions in the resolved-resonance (RR) region [36]. The analysis involved estimating mean values for $3,193 \times 5 = 15,965$ resonance parameters (E_R, Γ_n, Γ_γ, Γ_{f1}, and Γ_{f2}). The covariance matrix to represent the uncertainty information for these parameters therefore contained 1.2745×10^8 distinct elements!

Such large matrices can be cumbersome to manipulate. It is important to find ways to deal with them by making some adjustments or approximations that will render them manageable without sacrificing essential information. This paper discusses some techniques to address this problem.

b. Ordering and labeling variables Although an $n \times n$ covariance matrix $\boldsymbol{V}_x$ could have $n(n+1)/2$ distinct elements, there may be correlation coefficients that are either zero or very close to zero. Also, certain groups of variables may have non-zero correlations within individual groups, but no correlations between variables belonging to other groups. If certain correlation elements are zero, or whole regions off the diagonal consist of zero elements, it may be possible to alter variable labels without changing the information content. The full matrix is then a collection of sub-matrices. This is referred to as partitioning. An example demonstrates the concept.

Example 5 Suppose an $n \times n$ covariance matrix $\boldsymbol{V}_x$ can be partitioned into 3 sub-matrices $\boldsymbol{V}_{xk}$ ($k = 1, 3$). Then

$$\boldsymbol{V}_x = \begin{pmatrix} \boldsymbol{V}_{x1} & & \\ \boldsymbol{0} & \boldsymbol{V}_{x2} & \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{V}_{x3} \end{pmatrix}. \tag{32}$$

The zero blocks are indicated by the symbol "$\boldsymbol{0}$". These off-diagonal blocks need not be square or even have the same dimensions. A benefit of partitioning can be seen when calculating the inverse matrix $\boldsymbol{V}_x^{-1}$. It is easier to invert several smaller matrices than a single large one. Thus,

$$\boldsymbol{V}_x^{-1} = \begin{pmatrix} \boldsymbol{V}_{x1}^{-1} & & \\ \boldsymbol{0} & \boldsymbol{V}_{x2}^{-1} & \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{V}_{x3}^{-1} \end{pmatrix}. \tag{33}$$

If the dimensions of the sub-matrices are $n_1 \times n_1$, $n_2 \times n_2$, and $n_3 \times n_3$, with $n = n_1 + n_2 + n_3$, the number of elements of $\boldsymbol{V}_x$ to be considered is $\sum_{k=1}^{3} n_k(n_k+1)/2$, which is smaller than $n(n+1)/2$. How much smaller depends on the specific values n_1, n_2, and n_3. Suppose $n_1 = n_2 = n_3 = n/3$ (assuming n is divisible by 3). The number of elements to consider is $n(n+3)/6$. For large n, storage volume is thus smaller by about a factor of 3.

Very small correlations can often be set to zero with no significant consequences. The altered matrix will approximate the original matrix $\boldsymbol{V}_x$, but that may be adequate for practical purposes. It is a matter of judgment whether such a technique can be employed effectively.

Sometimes, it is possible to reformulate a problem by resorting to algebraic manipulations that reduce the dimensions of certain matrices that have to be manipulated. This technique has been exploited in algorithms employed by some nuclear data analysis codes, *e.g.*, SAMMY [10].

c. Numerical precision Numerical precision can be a concern when covariance matrices are manipulated, *e.g.*, multiplied or inverted. Modern computational assets can perform these operations quite accurately, but limitations of precision might still be encountered. They could generate biases that are impossible to trace. Covariance matrices that are positive definite in their original form might become singular due to round-off errors. Precision issues can also arise when data

span a numerical dynamic range of several orders of magnitude, *e.g.*, energy-dependent cross sections for an (n,γ) reaction can range from thousands of barns at thermal energies to just a few millibarn at higher energies. It may be helpful to rescale data by changing units or transforming variables. An example of this is transformation to natural logarithm form. While this technique might solve a dynamic range problem, it could also lead to unacceptable distortions in the results computed using these transformed data [37]. There may also be advantages to carrying out matrix operations using correlation matrices rather than covariance matrices. Matrix inversion is one such instance, as illustrated in Example 2.

Matrix size and computational precision issues may be intertwined in such a way that tradeoffs are required to avoid conflicts. One approach that can be used to reduce data storage needs is to represent correlation coefficients by integers. Integers require less storage capacity than do comparable real (decimal) numbers, and correlations need not be expressed with high precision. What is often done is to scale correlation value to the range (-100, +100), which provides two-significant-figure precision, or even to the range (-1,000, +1,000), which provides three-significant-figure precision. When performing arithmetical operations using these altered values, actual correlations are obtained by dividing the stored integer values by 100 or 1,000. Since variances (or standard deviations) are estimates of uncertainty, it is also unnecessary to specify them to high precision. However, it is advisable to store them as real numbers because they can span wide ranges of values. The potential for reducing storage requirements by these techniques can be significant. The extent to which significant computational accuracy might be lost by this approach should be investigated in specific cases, but it unlikely to be problematic. However, when resorting to such approaches, it is essential to preserve the integrity of the covariance matrices, *e.g.*, to avoid introducing singularities.

d. Uncertainty attributes It may be possible, in certain situations, to adequately represent the information provided in a covariance matrix $\boldsymbol{V}_x$ by a sum of several distinct matrix components, each one corresponding to a specific type of uncertainty or category of uncertainties. This is referred to as decomposing $\boldsymbol{V}_x$ according to distinct uncertainty attributes. If there are Q distinct uncertainty attributes, then

$$\boldsymbol{V}_x \approx \sum_{q=1}^{Q} \boldsymbol{V}_{xq}. \tag{34}$$

The sign " $\approx$ " is used, rather than strict equality, since approximations are usually involved. Each matrix $\boldsymbol{V}_{xq}$ has dimension $n \times n$, which is identical to that for $\boldsymbol{V}_x$. A particular matrix $\boldsymbol{V}_{xq}$ may involve uncertainty components that are common to all the variables represented in the full matrix $\boldsymbol{V}_x$. The process of decomposing a covariance matrix is quite different from partitioning it. The sub-matrices associated with partitioning have smaller dimensions than the full matrix, and the different sets of variables corresponding to the distinct partitions have no common uncertainty components. It is convenient to introduce the terminology: "micro-correlation" and "macro-correlation". Micro-correlation coefficients are correlation coefficients of the individual correlation matrices $\boldsymbol{C}_{xq}$ that correspond to the matrices $\boldsymbol{V}_{xq}$. Macro-correlation coefficients are correlation coefficients of the correlation matrix $\boldsymbol{C}_x$ associated with the matrix $\boldsymbol{V}_x$. An advantage to substituting Q matrices, comprised of $Q \times n^2$ parameters, for one matrix, with n^2 elements, is illustrated by the following example.

Example 6 Suppose that $n = 3$ and $Q = 2$. Then, Eq. (34) indicates that $\boldsymbol{V}_x \approx \boldsymbol{V}_{x1} + \boldsymbol{V}_{x2}$. Let x_1, x_2, and x_3 identify the three parameters. Their mean values are x_{01}, x_{02}, and x_{03}. It is not possible to specify the uncertainties using only three numbers (variances or standard deviations). Six values are needed: e_{x11}, e_{x12}, and e_{x13} for $\boldsymbol{V}_{x1}$, and e_{x21}, e_{x22}, and e_{x23} for $\boldsymbol{V}_{x2}$. Two sets of micro-correlation coefficients are also required: c_{x121}, c_{x131}, and c_{x132} for $\boldsymbol{V}_{x1}$, and c_{x221}, c_{x231}, and c_{x232} for $\boldsymbol{V}_{x2}$. The micro-correlation coefficients not listed are unity diagonal values and those equal to listed off-diagonal ones by symmetry, *e.g.*, $c_{x112} = c_{x121}$. The labeling convention used here is that the first integer subscript of each listed parameter indicates whether it belongs to $\boldsymbol{V}_{x1}$ or $\boldsymbol{V}_{x2}$. For this example, it is assumed that $c_{x121} = c_{x131} = c_{x132} = 0$. This would apply if the first attribute represented statistical counting uncertainties. Next, assume that the second attribute involves a common, fully-correlated fractional uncertainty f for x_1, x_2, and x_3 (*e.g.*, it might be 0.05 or 5%). Thus, $c_{x221} = c_{x231} = c_{x232} = 1$. This would occur if all counting were done using the same detector, with a common efficiency. The absolute uncertainty components are $e_{x21} = f x_{01}$, $e_{x22} = f x_{02}$ and $e_{x23} = f x_{03}$. It is assumed that mean values are provided elsewhere for this data set. In this example, only four parameters need to be specified instead of 12. They are e_{x11}, e_{x12}, e_{x13}, and f. The data user should be informed that the first uncertainty attribute is uncorrelated, while the second uncertainty attribute is fully correlated. For large data sets, and additional uncertainty attributes, this approach can lead to significant savings in uncertainty data specification and storage requirements. The total uncertainties for these three data values are $s_{xi} = (e_{x1i}^2 + e_{x2i}^2)^{1/2}$ for $i = 1, 3$. $\boldsymbol{V}_x$ can be constructed according to Eq. (34).

The sub-matrices $\boldsymbol{V}_{xq}$ are unlikely to share the same properties as the full covariance matrix $\boldsymbol{V}_x$. There may be duplicate rows and columns, leading these matrices to be non-positive definite. If $\boldsymbol{V}_x$ is formulated this way, it must be insured that it fulfills all the requirements of a proper covariance matrix. If experimental covariance data can be represented by uncertainty attribute decomposition, then the complete data for an experiment can be specified by giving mean values (measurement results), partial uncertainties, and micro-correlation values. This approach has been used in nuclear data compilations such as EXFOR [6].

It is common to consider the total uncertainty as consisting of a "statistical component" and a "systematic component" (*e.g.*, [2]). In fact, the assumption that variables have only "statistical" and "systematic" uncertainties is usually excessively simplistic. It is misleading to refer to "statistical" and "systematic" uncertainties as if they are different species governed by distinct mathematical laws [38]. All uncertainties should be interpreted as standard deviations of random variables, governed by probability distributions, and differentiated only by the roles they play in particular situations. This paper refers only to "uncorrelated" and "correlated" (or "par-

tially correlated") uncertainty components (*i.e.*, uncertainty attributes), and avoids using "systematic" in discussing data uncertainties.

Actual "errors" are often committed in experiments, *e.g.*, the incorrect calibration of a detector. Then, an experimenter does not measure what was intended, due to the erroneous detector calibration, but he is unaware of the mistake. Such "errors" are often the sources of discrepancies observed in compiled databases of experimental values. Since these "errors" were unrecognized, the assigned uncertainties appear to have been underestimated [3, 6]. When undertaking nuclear experiments, it is essential that consideration of the possible sources of "errors" and uncertainties be given adequate attention.

e. Approximate micro-correlations Provision of some uncertainty correlation information is better than giving no correlation information at all, whenever the circumstances of an experiment strongly suggest that non-zero correlations should be present. The absence of provided correlation values can lead data users (*e.g.*, evaluators) to assume that all the correlations are zero, with potentially undesirable, and largely untraceable, consequences (*e.g.*, [39, 40]). The absence of correlation information for differential data can lead to unrealistically small uncertainties being determined for integral results computed using these data (*e.g.*, [34]).

In some cases, it can be extremely difficult to do a credible job of estimating uncertainties and their correlations. If a particular uncertainty attribute is expected to make only a minor contribution to the total uncertainties in an experiment, it may be justified to provide estimates that are relatively crude, yet nevertheless plausible, especially regarding correlations. An example would be the effects associated with neutron scattering perturbations to neutron cross-section data. Experimenters should try to keep such corrections as small as possible through the design of their experiments, even if the need for corrections cannot be completely eliminated.

This section discusses three *ad hoc* approaches that have been used by experimenters in estimating uncertainty correlations for their data at the micro-correlation level. While it may appear that the correlation values these methods yield are "plucked out of thin air", there is a degree of rationality associated with these approaches. If the potential benefits of attempting more objective techniques to generated the covariance data do not appear to warrant the amount of effort that would be required, these may be the most practical options. Still, such crude approaches to estimating correlations should be considered only as a last resort, when more objective methods, based on thorough analyses of the experimental details, are impractical.

The first approach assumes that the micro-correlation corresponding to uncertainty attribute "q", for two variables x_i and x_j, is $c_{xqij} = \pm 0.25$ (weak), or $c_{xqij} = \pm 0.5$ (moderate), or $c_{xqij} = \pm 0.75$ (strong). The possibility that these might be either positive or negative is considered, although experience suggests that it is generally more difficult to envision negative correlations than positive ones. The choice $c_{xqij} = \pm 0.5$ is the most reasonable option if it is suspected that there should be some correlation, but there is no justification for assuming that it should be either weak or strong. This default choice is "moderate", with a magnitude half way between 0 (none) and 1 (full). Weak correlation is assigned $c_{xqij} = \pm 0.25$, since this magnitude is half way between 0 (none) and 0.5 (moderate). Strong correlation is assigned $c_{xqij} = \pm 0.75$, since this magnitude is half way between 0.5 (moderate) and 1 (full). These assumptions, applied in the absence of any evidence other than the experimenter's intuition, stem from a general principle known as Occam's Razor [41]. It states that, in the absence of concrete information, one makes choices (or decisions) that appear to involve a minimum number of arbitrary prior assumptions.

A second approach to estimating micro-correlations can be described as follows: Consider a data set $\{x_i\}$ which entails an intrinsic parameter that can be used to characterize each member of the set (*e.g.*, neutron energy E_i for neutron cross sections). Estimated micro-correlations based on the concept of "range" can be specified. For example, suppose that $E_n \gg E_1$ and $E_{i+1} > E_i$ for all the data points, and that neutron-scattering corrections are considered. It is then assumed that $c_{xqij} \approx +\{1 - \mathrm{abs}[(E_i - E_j)/(E_n - E_1)]\}$. Notice that only positive correlations are considered in this approach. The correlation is strong between points with neighboring energies, and tapers to zero linearly for the points most widely separated in energy. Alternatively, either the exponential function or the normal function can be used. However, details of such assumed "range" functions should not be taken too seriously since they are very subjective.

A third approach toward representing uncertainty correlations incorporates the concept of short-, medium- and long-range correlations (*e.g.*, see [37]). Implementation of this concept is frequently found in evaluated nuclear data files, and the ENDF-6 formats [33] support its use. In this approach it is postulated that the covariance matrix for a set of experimental data is comprised of three distinct partial uncertainty components, with magnitudes to be specified. One component (short-range) involves correlations only between neighboring data points (*e.g.*, using energy separation to measure "distance"). The usual assumption is that it is strong, *e.g.*, $c_{xqij} = +0.75$. The second component is assumed to be medium-range, involving more than just the "nearest neighbor" data points, but no distant points, and it is usually assumed to be moderate, *e.g.*, $c_{xqij} = +0.5$. The third component is usually taken to be weak, *e.g.*, $c_{xqij} = +0.25$. It introduces correlations through the whole data set. Notice once again that only positive correlations are considered in this approach.

Smith [42] demonstrated that the consequences of making crude estimates of the micro-correlations for relatively small uncertainties are rarely damaging. If there is a fairly large number of uncertainty attributes involved, there is a tendency to average out over-estimates or under-estimates of micro-correlation strengths for particular components. This is another manifestation of the central limit concept. Nevertheless, such subjective approaches to generating correlation data should not be employed if more objective information about the actual correlations of the various considered uncertainty attributes can be determined.

Two final comments are worth mentioning: Uncertainties corresponding to several attributes can sometimes be lumped together in quadrature to reduce storage requirements if their correlation properties are all identical, *e.g.*, either random or fully correlated (*e.g.*, [43]). It should be re-emphasized that it is a mistake to assume that $c_{xqij} = 0$, when it is strongly suspected that some degree of correlation exists, simply because it is too difficult to generate an objective estimate of the actual correlation.

C. Uncertainty Propagation Methods

Most of the physical quantities of interest in nuclear science are not measured directly, but rather are derived indirectly from parameters that are actually measured or otherwise determined. For example, neutron cross sections are not measured directly, but instead detector counts, detector calibrations, standards, *etc*., are determined. The cross sections are then derived from these primary (*i.e.*, lower-level) parameters by means of formulas or data analysis algorithms. It is necessary to evaluate the uncertainties of derived quantities based on the uncertainties of measured or otherwise determined primary parameters.

An important property of random variables is that any continuous, differentiable real function of a continuous, real-number random variable is also a random variable. It will be governed by the laws of probability, and can be treated using mathematical tools that are described in preceding sections of this paper. For example, if x is a primary random variable, and y is related to x through the function $y = f(x)$, then one can define its mean value y_0, variance v_y, and standard deviation s_y as

$$y_0 = \langle y \rangle = \int_{\Xi} y\, p(x)\, \mathrm{d}x = \int_{\Xi} f(x) p(x) \mathrm{d}x, \quad (35)$$

$$v_y = \int_{\Xi} y^2 p(x)\, \mathrm{d}x - y_0^2 = \int_{\Xi} [f(x)]^2\, p(x)\, \mathrm{d}x - y_0^2, \quad (36)$$

$$s_y = v_y^{1/2}. \quad (37)$$

The probability distribution that governs the primary random variable x is $p(x)$. Extension of this formalism to a collection of m derived random variables $\boldsymbol{y} = (y_1, \cdots, y_k, \cdots, y_m)$ is straightforward. Each y_k involves a distinct function f_k of the n random variables $\boldsymbol{x} = (x_1, \cdots, x_i, \cdots, x_n)$, *i.e.*, $y_k = f_k(\boldsymbol{x})$ or, using vector notation, $\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x})$. An interesting example is a single random variable y, defined by $y = f(\boldsymbol{x}) = \Sigma_{i=1,n}\, x_i$. If n is large, and the x_i are independent, then the probability distribution for y converges to a normal distribution, consistent with the well-known Central Limit Theorem (*e.g.*, [3]). Furthermore, the Lyapunov Theorem states that this holds true regardless of the nature of the distributions of the individual x_i (they all may be different), as long as the mean values, variances, and third moments exist for these distributions (*e.g.*, [44]). For this particular case, $v_y = \Sigma_{i=1,n}\, v_{xii}$. Variances are seen to exhibit an additive property when several distinct sources of uncertainty are involved. However, standard deviations generally cannot be added linearly to determine total uncertainties. They can be added in quadrature.

Eqs. (35) and (36) are symbolic, so there is a need to establish how the moments of y should be calculated when the moments of x are known. There are two distinct ways of accomplishing this task: deterministic and stochastic.

1. *Deterministic approach*

If f is a linear function of x, *i.e.*, $y = f(x) = a + bx$, with "a" and "b" as constants, the relationship between the moments of x and y can be determined easily. Thus,

$$y_0 = a + bx_0, \quad (38)$$

$$v_y = b^2 v_x = b^2 s_x^2. \quad (39)$$

The translation constant "a" and scale constant "b" both affect the mean value y_0. Only the scale constant "b" affects the variance v_y. The relationship between y and x is rarely linear, so it is frequently necessary to linearize the problem, *i.e.*,

$$y \approx f(x_0) + f'(x_0)(x - x_0). \quad (40)$$

This expression involves the first two terms of a Taylor series expansion. The first derivative of function f with respect to x, evaluated at $x = x_0$, is $f'(x_0)$. The mean value of y is $y_0 \approx f(x_0)$. Rearranging Eq. (40), squaring both sides of the equation, averaging over the function $p(x)$, and applying the definition of variance given in Eq. (36), yields

$$v_y \approx \langle [y - f(x_0)]^2 \rangle = [f'(x_0)]^2 \langle (x - x_0)^2 \rangle = [f'(x_0)]^2\, v_x. \quad (41)$$

The standard deviation is $s_y = v_y^{1/2}$. This is commonly referred to as the deterministic approach. It provides a reasonable approximation for y_0 and v_y only if $v_x \ll x_0^2$.

The formalism can be extended to collections of nonlinear functions, $\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x})$, of a primary variable set $\boldsymbol{x}$. The mean values are given by $\boldsymbol{y}_0 \approx \boldsymbol{f}(\boldsymbol{x}_0)$, and the covariance matrix $\boldsymbol{V}_y$ can be obtained from the expression

$$\boldsymbol{V}_y \approx \boldsymbol{T} \boldsymbol{V}_x \boldsymbol{T}^{+}. \quad (42)$$

$\boldsymbol{V}_y$ is an $m \times m$ matrix, $\boldsymbol{V}_x$ is an $n \times n$ matrix, $\boldsymbol{T}$ is an $m \times n$ matrix known as the sensitivity matrix, and $\boldsymbol{T}^{+}$ is its transpose, with dimension $n \times m$. The elements of $\boldsymbol{T}$ are the partial derivatives $t_{ki} = (\partial f_k / \partial x_i)|_{\boldsymbol{x}=\boldsymbol{x}_0}$ $(k = 1, m;\ i = 1, n)$, *i.e.*, derivatives evaluated at $\boldsymbol{x} = \boldsymbol{x}_0$. Eq. (42) is known as the Law of Error Propagation [3]. It provides an approximate expression for propagating uncertainties from a primary variable set $\boldsymbol{x}$ to a derived variable set $\boldsymbol{y}$, based on linearization of the functional relationships between $\boldsymbol{y}$ and $\boldsymbol{x}$.

If the partial derivatives cannot be expressed in analytic form, it is necessary to calculate approximate values for these derivatives numerically. For a single variable

$$f'(x_0) \approx [f(x_0 + \Delta x) - f(x - \Delta x)] / 2\Delta x. \quad (43)$$

A carefully chosen small increment, $\Delta x > 0$, of the variable x will usually provide an adequate approximation to the value of

the derivative at $x = x_0$, if the function is reasonably "smooth" in the vicinity of x_0, and there are no fluctuations. The increment Δx must not be too small (to avoid computational instability) nor too large (resulting in a poor approximation to the derivative). Convergence can be sought by trying various choices of Δx. For multivariate situations, analogous expressions can be used to approximate the partial derivatives seen in matrix $\boldsymbol{T}$.

a. Attributes of derived uncertainties It was indicated that an $n \times n$ matrix $\boldsymbol{V}_x$ can sometimes be approximated by a sum of $n \times n$ matrices, each corresponding to a distinct uncertainty attribute. In accordance with Eq. (34), the formula for the covariance matrix elements produced by this matrix decomposition is

$$(\boldsymbol{V}_x)_{ij} = v_{xij} \approx \sum_{q=1}^{Q} c_{xqij}\, e_{xqi}\, e_{xqj}. \tag{44}$$

This concept can also be applied to derived variables, and exploited in tabulating the uncertainties and correlations for the derived quantities produced by an experiment.

A particular decomposition of $\boldsymbol{V}_y$ can be generated for cases where the covariance matrix $\boldsymbol{V}_x$ for the primary variables can be partitioned, as shown in Eq. (32). The primary variables are relabeled so that they form groups that may be internally correlated within each group, but are uncorrelated between distinct groups. It is assumed that there are Q such distinct groups of variables, and that $\boldsymbol{V}_x$ can be partitioned so that the essential covariance information is contained in a collection of sub-matrices $\boldsymbol{V}_{xq}$ ($q = 1, Q$). Thus,

$$\boldsymbol{V}_y \approx \boldsymbol{T}\boldsymbol{V}_x\boldsymbol{T}^{+} = \sum_{q=1}^{Q} \boldsymbol{T}_q\boldsymbol{V}_{xq}\boldsymbol{T}_q^{+} = \sum_{q=1}^{Q} \boldsymbol{V}_{yq}. \tag{45}$$

This expression can be understood by considering Eqs. (32) and (42). Since $\boldsymbol{y}$ consists of m elements, each term in the sum appearing on the right-hand side of Eq. (45) must be an $m \times m$ matrix. If a particular sub-matrix $\boldsymbol{V}_{xq}$ has the dimension $\alpha \times \alpha$, then the dimension of $\boldsymbol{T}_q$ must be $m \times \alpha$ and its transpose $\boldsymbol{T}_q^{+}$ has dimension $\alpha \times m$. The various $\boldsymbol{T}_q$ appearing in Eq. (45) are sub-matrices of the original sensitivity matrix $\boldsymbol{T}$. Due to the partitioning of $\boldsymbol{V}_x$, and the fact that portions of that matrix consist entirely of zeroes, many of the elements of $\boldsymbol{T}$ will amount to "deadwood" for the computations indicated by Eq. (45). Only the elements of the sub-matrices $\boldsymbol{T}_q$ enter into these calculations. Other sensitivity coefficients from $\boldsymbol{T}$ are multiplied by zeroes and play no roles in the analysis. The elements of $\boldsymbol{V}_y$ are

$$\begin{aligned} v_{yk\theta} &= (\boldsymbol{V}_y)_{k\theta} \approx \sum_{q=1}^{Q} (\boldsymbol{T}_q\boldsymbol{V}_{xq}\boldsymbol{T}_q^{+})_{k\theta} \\ &= \sum_{q=1}^{Q} (\boldsymbol{V}_{yq})_{k\theta} = \sum_{q=1}^{Q} e_{yqk}\, c_{yqk\theta}\, e_{yq\theta}. \end{aligned} \tag{46}$$

The $c_{yqk\theta}$ that appear in Eq. (46) are micro-correlation coefficients, and the e_{yqk} are the partial uncertainties for the derived parameters y_k associated with the uncertainty attribute q. The standard deviation s_{yk} of y_k is

$$s_{yk} = \left(\sum_{q=1}^{Q} e_{yqk}^2 \right)^{1/2}. \tag{47}$$

b. Effects of constraints Covariance matrices generally ought to be positive definite so they can be inverted and their determinants calculated. However, there are cases where they must be singular. The following example illustrates this point.

Example 7 Let $\sigma_{\rm tot}$, $\sigma_{\rm el}$, and $\sigma_{\rm non}$ represent the neutron total, elastic, and non-elastic cross sections at a specific energy. A physical requirement is that $\sigma_{\rm tot} = \sigma_{\rm el} + \sigma_{\rm non}$. So, these three parameters cannot vary independently. For convenience, we write $\sigma_{\rm non} = \sigma_{\rm tot} - \sigma_{\rm el}$. Then, assume that $\sigma_{\rm tot}$ and $\sigma_{\rm el}$ were measured independently and their uncertainties estimated. The uncertainty in $\sigma_{\rm non}$ is to be determined. Also, a single covariance matrix that includes uncertainty information for all three cross sections is to be determined. For convenience, let $x_1 = \sigma_{\rm tot}$ and $x_2 = \sigma_{\rm el}$. The corresponding standard deviations are s_{x1} and s_{x2}. These two variables are independent, so the correlation coefficient is $c_{x21} = c_{x12} = 0$. This defines a 2×2 covariance matrix $\boldsymbol{V}_x$ which is clearly positive definite. Next, define three random variables: $y_1 = x_1$, $y_2 = x_2$, and $y_3 = x_1 - x_2$. Here, y_3 represents the non-elastic cross section, $\sigma_{\rm non}$. The objective is to derive a 3×3 covariance matrix $\boldsymbol{V}_y$. The Law of Error Propagation, *i.e.*, Eq. (42), is applied. Since y_3 is defined as a linear combination of x_1 and x_2, no approximations are involved. Matrix algebra yields the following results for the variances of $\boldsymbol{V}_y$: $v_{y11} = s_{x1}^2$, $v_{y22} = s_{x2}^2$, and $v_{y33} = s_{x1}^2 + s_{x2}^2$. The corresponding covariances are $v_{y21} = v_{y12} = 0$, $v_{y31} = v_{y13} = s_{x1}^2$, and $v_{y32} = v_{y23} = -s_{x2}^2$. The uncertainty of the elastic cross section therefore is anti-correlated (negatively correlated) with that for the non-elastic cross section, since the parameter variations are constrained by $\sigma_{\rm tot} = \sigma_{\rm el} + \sigma_{\rm non}$. Sums of the elements in both the second row and second column of $\boldsymbol{V}_y$ are zero. Finally, $\det(\boldsymbol{V}_y) = 0$ since the matrix $\boldsymbol{V}_y$ is degenerate (singular).

Experimenters should take care to insure that their covariance matrices are always positive definite when there are no constraints among the included variables. However, augmented covariance matrices for constrained variables should be constructed by propagating uncertainties and correlations for the redundant variable set in the manner shown in the preceding example. Often, constraints among sets of random variables will introduce anti-correlations between certain members of the set. This is very obvious in the case of two variables. An increase in one of them must lead to a decrease in the second one to satisfy the constraint.

Consider another example that demonstrates the effect of constraints on sets of random variables.

Example 8 A measured neutron spectrum is represented by energy-group flux parameters $\{\varphi_i\}$, for $i = 1, n$. Neutron spectra are generally represented in normalized form so that a sum over all group fluxes yields unity. Normally, measured spectrum values would not be normalized. So, these data must be normalized by defining a new variable set $\{\psi_i\}$

$$\psi_i = \varphi_i / \Phi. \tag{48}$$

Here, $\Phi = \Sigma_{i=1,n}\,\varphi_i$. Clearly, $\Sigma_{i=1,n}\,\psi_i = 1$. We assume that a proper covariance matrix $\boldsymbol{V}_\varphi$ has been generated by considering various sources of uncertainty in the experiment. It is then possible to compute a covariance matrix $\boldsymbol{V}_\psi$ for the normalized set $\{\psi_i\}$ by applying the Law of Error Propagation, *i.e.*, Eq. (42). Expressions for the elements of the $n \times n$ sensitivity matrix $\boldsymbol{T}$ are needed. They are given by

$$t_{ki} = (\partial\psi_k/\partial\varphi_i) = (\Phi\delta_{ki} - \varphi_k)/\Phi^2 \quad (k = 1, n;\ i = 1, n). \qquad (49)$$

Here, δ_{ki} is the Kronecker Delta function. It can be shown that all rows and columns of $\boldsymbol{V}_\psi$ sum to zero, and $\det(\boldsymbol{V}_\psi) = 0$ [3, 35].

There are advantages and disadvantages to deterministic uncertainty propagation.

Advantages

- It is transparent, and the calculations yield well-defined and, for many cases, sufficiently accurate results.
- It tends to be relatively fast by comparison with stochastic (Monte Carlo) approaches.

Disadvantages

- If uncertainties are large, and the functions that relate primary to derived variables are nonlinear, then deterministic uncertainty propagation may produce unacceptably biased values. These deficiencies may be partially compensated by including higher-order terms in series expansions that represent the functional relationships, but in practice this approach can be awkward to implement.
- There are limited opportunities for further analysis. For example, estimation of higher-order moments of the probability distributions, such as skewness or kurtosis, tends to be impractical to carry out this way.
- Uncertainties for discrete variables that are governed by discrete probability functions, *e.g.*, nuclear spin parameters, cannot be analyzed deterministically.

2. *Monte Carlo approach*

The second approach to uncertainty propagation is a stochastic one. Its use for certain applications in the nuclear data field has been rendered feasible (and often practical) as a consequence of recent advances in computational power [45]. In this approach, $\boldsymbol{x}$ is a vector of primary variables, with dimension n, and $\boldsymbol{y}$ is the vector of derived variables, with dimension m. The variables of $\boldsymbol{y}$ are derived from those of $\boldsymbol{x}$ through a collection of m functional relationships, $\boldsymbol{f} = \{f_k\}$, such that $y_k = f_k(\boldsymbol{x})$ for $k = 1, m$. The uncertainties associated with $\boldsymbol{x}$ are represented by the covariance matrix $\boldsymbol{V}_x$. The method involves producing a large collection of Λ vectors $\boldsymbol{x}_\lambda\,(\lambda = 1, \Lambda)$ by randomly varying each component x_i of $\boldsymbol{x}$ in accordance with the probability function $p(\boldsymbol{x})$ governing $\boldsymbol{x}$. A set of vectors, $\boldsymbol{y}_\lambda = \boldsymbol{f}(\boldsymbol{x}_\lambda)\,(\lambda = 1, \Lambda)$, is generated using the random collection of primary variables.

Correlations between component variables of $\boldsymbol{x}$ should be taken into account in these sampling exercises. Algorithms for performing this task are available. One of these is the Metropolis-Hasting scheme [46, 47]. Another approach involves carrying out an orthogonal transformation of the original variable set $\boldsymbol{x}$. Each transformed component variable is then sampled independently, and the entire set is transformed back to the original variable space [3, 16]. This process is repeated Λ times. It may be revealed that one or more of the eigenvalues of the primary parameter matrix $\boldsymbol{V}_x$ are either zero or even negative, so the matrix is not positive definite. This could occur due to numerical round-off. Usually, the magnitudes of these pathological eigenvalues, if they are negative, will be small. A possible "fix" in such circumstances (with minimal consequences) is to set all zero or negative eigenvalues to small positive values before transforming back to the original variable space. Fortunately, in many situations it suffices to assume that the components of $\boldsymbol{x}$ are independent, and that correlations between the components of $\boldsymbol{y}$ are introduced by the functional relationships between $\boldsymbol{y}$ and $\boldsymbol{x}$. The stochastic sampling of the variables $\boldsymbol{x}$ is then more straightforward.

The collection of Λ generated sample vectors $\boldsymbol{y}_\lambda\,(\lambda = 1, \Lambda)$ can be used to estimate sample mean values, *i.e.*,

$$\langle y_k\rangle_\Lambda = \left(\sum_{\lambda=1}^{\Lambda} y_{k\lambda}\right)/\Lambda \quad (k = 1, m). \qquad (50)$$

Eq. (50) is similar in form to Eq. (1). The numerical values obtained depend on Λ, and two different sequences of Λ generated values of $\boldsymbol{y}_\lambda$ will lead to distinct mean values [46–48]. Sample variances and covariances are obtained in a similar manner, *i.e.*,

$$(v_{yk\theta})_\Lambda = \left(\sum_{\lambda=1}^{\Lambda} y_{k\lambda}\,y_{\theta\lambda}\right)/\Lambda - \langle y_k\rangle_\Lambda\langle y_\theta\rangle_\Lambda \quad (k, \theta = 1, m). \qquad (51)$$

The sample standard deviations are $(s_{yk})_\Lambda = [(v_{ykk})_\Lambda]^{1/2}$. The collection of elements $(v_{yk\theta})_\Lambda$ forms a sample covariance matrix $(\boldsymbol{V}_y)_\Lambda$ that depends on sample size and stochastic history. As Λ becomes large, all these results will converge to the collection of values $\boldsymbol{y}_0$ and $\boldsymbol{V}_y$, determined by the primary variable probability distribution $p(\boldsymbol{x})$.

The stochastic (Monte Carlo) approach also offers both advantages and disadvantages.

Advantages

- Eqs. (50) and (51) are consistent with an intuitive understanding of mean values and covariances.
- This approach is impervious to nonlinear effects, since there is no need to linearize the functions f_k. This is an important advantage since there may be instances where such nonlinear effects cannot be neglected in uncertainty propagation or related applications.
- It is flexible, since collections of randomly derived vectors, $\{\boldsymbol{y}_\lambda\}_\Lambda$, can be utilized for a wide variety of purposes. These include calculating higher-order moments

of probability distributions and contemporary reactor physics applications (*e.g.*, [49]).

- This method can be used in modeling experiments, where data analyses may involve complex numerical algorithms as well as analytical functions.

Disadvantages

- The Monte Carlo approach is computationally intensive, and its application can be impractical in many circumstances. However, such computational barriers are rapidly disappearing as computer hardware and software capabilities become more powerful.
- Unique results are not obtained, since outcomes always depend on Λ and the stochastic history. However, since Monte Carlo techniques are widely used with great success in many areas of nuclear science, this lack of exact predictability for the obtained results is acceptable for most practical applications.
- It is difficult, but not impossible, to obtain estimates for parameter sensitivity coefficients by this approach. In some applications such sensitivity information is useful.

The choice of whether to use the deterministic or stochastic approach is generally left to the discretion of the practitioner. It may also be dictated by the computational tools that are available to deal with the problems at hand. Many of the computer programs that are currently used for nuclear applications, *e.g.*, for data evaluation or reactor physics calculations, are based on the deterministic approach, largely because it has been in use for a much longer time. However, new computational procedures and computer codes that incorporate Monte Carlo methods for dealing with uncertainties are under development [49].

3. *Some simple examples*

It is instructive to consider several simple examples in order to clarify the concepts discussed in this paper.

Example 9 Consider the simple linear expression $y = a + bx$. Let $a = 1$ and $b = 2$. Then, let x be governed by the uniform distribution over the interval (4.5, 5.5). Therefore, $x_0 = 5$, $v_x = 1/12 = 0.083333$, and $s_x = 0.28868$. The fractional uncertainty in x is $f_x = 0.057735$, *i.e.*, $\approx$ 5.8% (a modest uncertainty). Eqs. (38) and (39) lead to the deterministic results $y_0 = a + bx_0$ and $v_y = b^2 v_x$. Therefore, $y_0 = 11$, $v_y = 0.33333$, and $s_y = 0.57735$. The fractional uncertainty $f_y = 0.052486$ ($\approx$ 5.2%). An analysis of this problem by Monte Carlo, tracing 1,000 histories ($\Lambda = 1,000$), generated the results $\langle y\rangle_\Lambda = 11.013$ and $v_{y\Lambda} = 0.33521$. A second set of results from 1,000 different Monte Carlo histories yielded the values $\langle y\rangle_\Lambda = 11.017$ and $v_{y\Lambda} = 0.33722$. Several additional tests were performed, and each yielded values that were comparable (within anticipated statistical uncertainties) to the first two trials, and to the values that were obtained deterministically. Adequate Monte Carlo convergence is observed for this simple, linear situation for $\Lambda = 1,000$ histories. Good agreement is also seen for the results from both the deterministic and stochastic methods.

In the following example, the variables x and y are not linearly related, and the uncertainty in x is substantial.

Example 10 Consider the following functional relationship: $y = A\exp(-tx)$. This nonlinear function describes radioactive decay, where y is the activity at time $t > 0$, x is the decay constant, and A is the activity at time $t = 0$ [3]. Then, suppose that x is governed by the uniform distribution over the range (0.05, 0.15), with $x_0 = 0.1$, $v_x = 0.01/12 = 0.00083333$, and $s_x = 0.028868$. The fractional uncertainty is $f_x = 0.28868\,(\approx 28.9\%)$. Let $A = 1$ and $t = 10$. The deterministic solution for the comparable derived parameters yields are $y_0 \approx 0.36788$, $v_y \approx 0.011278$, $s_y \approx 0.10620$, and $f_y \approx 0.28868\,(\approx 28.9\%)$. The Monte Carlo approach is then applied. Table II gives results for three distinct sampling exercises of 1,000 histories each. Good convergence is provided by $\Lambda = 1,000$ histories. The Monte Carlo values are a bit larger than the deterministic results for both the mean value ($\approx$ 5% larger) and standard deviation ($\approx$ 3% larger). But, these differences are relatively minor considering the uncertainties involved, and the fact that y decreases in magnitude by a factor of 2.7183 as x increases from 0.05 to 1.5.

TABLE II. Mean values, variances, standard deviations, and fractional uncertainties derived from three separate Monte Carlo exercises.

$\langle y\rangle$	v_y	s_y	f_y(%)
0.38508	0.011955	0.10934	28.4
0.38768	0.012049	0.10977	28.3
0.38869	0.011856	0.10889	28.0

The following example examines the propagation of uncertainties for ratios of random variables, and it demonstrates the effect of correlations.

Example 11 Consider two variables, x_1 and x_2, and let $y = f(x_1, x_2) = x_1/x_2$. The standard deviations for the two primary variables are s_{x1} and s_{x2}, and c_{x21} is their correlation. If $x_{10} = 1$, $x_{20} = 2$, $s_{x1} = 0.3\,(30\%)$, $s_{x2} = 0.6\,(30\%)$, and $c_{x12} = c_{x21} = -0.5$, a deterministic analysis yields the results $y_0 = 0.5$, $v_y = 0.0675$, $s_y = 0.25981$, and $f_y = 0.51962\,(\approx 52\%)$. The negative correlation (anti-correlation) between x_1 and x_2 leads to enhancement in the uncertainty of y. The opposing variations for these two primary variables accentuate the variation for the derived ratio. Other possible outcomes are as follows: $c_{x21} = -1$ yields 60% uncertainty in y; $c_{x21} = 0$ yields $\approx$ 42% uncertainty in y; and $c_{x21} = +0.5$ yields 30% uncertainty in y.

Correlations for derived variables can arise from the functional relationships between the derived and primary variables, even if the primary variables are uncorrelated.

Example 12 Consider two primary variables x_1 and x_2. The standard deviations of the primary variables are s_{x1} and s_{x2}, and we assume $c_{x12} = c_{x21} = 0$. Then consider two derived variables $y_1 = x_1 + x_2$ and $y_2 = x_1 - x_2$. The derived mean values are $y_{01} = x_{01} + x_{02}$

and $y_{02} = x_{01} - x_{02}$. The covariance matrix can be determined from an application of Eq. (42). Its elements are $v_{y11} = v_{y22} = s^2_{x1} + s^2_{x2}$ and $v_{y21} = v_{y12} = s^2_{x1} - s^2_{x2}$. If $s^2_{x1} \neq s^2_{x2}$, then $v_{y21} \neq 0$, and y_1 and y_2 are at least partially correlated. However, if $s^2_{x1} = s^2_{x2}$, the covariance matrix is diagonal. Only then is $c_{y12} = c_{y21} = 0$. This unique occurrence does not suggest that y_1 and y_2 should be treated as uncorrelated. If two variables are truly uncorrelated, then their covariance matrix is always diagonal. However, the opposite statement is not true. If a covariance matrix happens to be diagonal, this does not imply that the variables are uncorrelated. This exemplifies the "necessary and sufficient" condition often encountered in mathematics.

III. MODELS OF EXPERIMENTS

A. Mathematical Models

The following words were written by a prominent academic economist, in collaboration with a journalist.

> *"Knowing what to measure and how to measure it makes a complicated world much less so. If you learn how to look at data in the right way, you can explain riddles that otherwise might have seemed impossible. Because, there is nothing like the sheer power of numbers to scrub away layers of confusion and contradiction."*[*]
>
> [*]Steven D. Levitt and Stephen J. Dubner, *Freakonomics: A Rogue Economist Explores the Hidden Side of Everything*, Harper Collins Publishers, New York (2005).

An experimenter might ask: "Why should I spend extra time estimating the uncertainties in my experiment rather than devoting that time to doing the best possible job I can at obtaining the actual experimental results?" The answer is: By spending time considering the sources of uncertainty associated with an experiment, an investigator will gain a much better understanding of that experiment and, as a consequence, be more likely to achieve better results than might otherwise have been the case. Careful attention to the likely sources of uncertainty in an experiment, before performing the measurements, will likely lead to a better designed and more efficient enterprise. An understanding and acceptance of the benefits of including uncertainty analysis as an integral part of experimental work should serve to adequately motivate experimenters to devote proper attention to this matter.

Experiments often involve false steps, repetition, performing certain measurement tasks out of natural order, *etc*. However, in spite of the superficially chaotic appearance of typical experiments, they can be organized into sequences of reasonably well-defined steps: i) establish the objectives of the experiment; ii) design the experiment to optimize chances for success, considering the constraints of available time, physical and financial resources, and contemporary methodologies; iii) decide specifically what should be measured; iv) conduct the actual measurements; v) analyze the measured data; vi) interpret the experimental results; vii) document these results and their interpretation; and viii) suggest improvements for future experiments. Uncertainty analysis is not listed as a separate task, since the consideration of uncertainties should be an integral part of all of the above mentioned steps. It should not be viewed as a separate chore to be addressed as an afterthought to conform to professional conventions, or to satisfy a mandated requirement. Approaches taken in performing the measurements tend to be guided by how the acquired data will be analyzed, while the data analysis procedures are clearly influenced by what could (or could not) be measured in the given circumstances. What is eventually documented, regarding details and results of an experiment, depends on what has been measured and derived from the measured data through analysis. Finally, the course followed in an experiment will likely be influenced by how the experimenter intends to document the work. That could be affected by an experimenter's perception of how the results might eventually be used. Advanced planning and modeling of an experiment are key factors in producing good results. They also establish a framework for properly estimating the experimental uncertainties. Attributes of an experiment that are likely to contribute significantly to uncertainties in the data should, to a large extent, be foreseen as a consequence of good planning. Analysis of data measured in nuclear experiments is usually accomplished using software developed for this purpose. Many components of this software can be developed before the measurements actually begin. The modeling of experiments and development of data analysis software are closely related tasks, and both of them involve the consideration of uncertainties.

1. *Formulas and algorithms*

In planning a nuclear experiment, it is very useful to develop a model that reflects the significant features of that experiment. This is similar to the way reactor physicists use mathematical models in designing nuclear reactors prior to their construction. A distinction needs to be made between use of the term "model" in the present context and the more common usage of this term in the nuclear data field. The latter usage refers to theoretical models of physical processes that are used to calculate numerical values of physical parameters for comparison with experimental results. These models can also be employed to generate nuclear data beyond the realm where experiments are currently capable of yielding results. The distinction between these two usages of the term "model" has become somewhat blurred due to the growing sophistication of modern nuclear science. For example, information based on theoretically derived models is often used in contemporary experiments to calibrate detectors, to correct for perturbing effects in experiments (such as neutron scattering), *etc*. Models of experiments can help in defining the parameters to be measured, and in clarifying the relationships between measured parameters and those derived through analyses, based on the objectives of the experiments. In complex experiments, the process of developing such models might well entail undertaking auxiliary measurements and analytical studies, *e.g.*, calibration of apparatus, trials to established likely detector count rates, *etc*. Their

purpose is to insure that the developed models conform to reality. Models provide the opportunity to simulate some, or all, aspects of an experiment, and to vary certain parameters (and their likely uncertainties), to determine whether the experiment has the potential to achieve the desired objectives. Modeling can enable the exploration of various possible measurement scenarios at much less cost of time, effort, and treasure than would be expended in actually performing the measurements. Such exercises can enhance the efficiency of an experiment, because knowledge gained from simulation provides the experimenter with insight as to which observables are the most important to measure carefully, measurement times needed to achieve the desired accuracy, *etc*. Mathematical models of experiments facilitate analyses of measured data and determinations of uncertainties.

Next, we examine the process of modeling experiments from a mathematical perspective. Let the variables $\boldsymbol{x}$ represent a collection of primary parameters of the model, with covariance matrix $\boldsymbol{V}_x$. It includes variables to be measured and those obtained from other sources, *e.g.*, the literature. The vector $\boldsymbol{y}$, with covariance matrix $\boldsymbol{V}_y$, represents the results to be derived from the experiment. Among the primary parameters of an experiment, there will be certain members that contribute noticeably to the uncertainties of the final experimental results, and others where the contributions are considered to be small enough to neglect. Both categories of primary parameters are needed in modeling experiments, but they should be distinguished when assessing the uncertainties. We will assume that only those primary parameters which contribute significantly to the uncertainties in the derived results should appear in the collection $\boldsymbol{x}$. Parameters which the experimenter knows can be treated as constants, but that are needed to derive the desired results, are included in a vector $\boldsymbol{\eta}$. Examples of parameters in the latter category might be the speed of light, the mass of the neutron, *etc*. Experimental planning should focus on choosing the parameters in sets $\boldsymbol{x}$ and $\boldsymbol{\eta}$ so as to optimize the chances of obtaining accurate values for the derived variables $\boldsymbol{y}$, with the smallest possible uncertainties, as represented by $\boldsymbol{V}_y$. The final step of an experiment is then the interpretation and documentation of these results.

a. Primary random variables The choice of a parameter set $\boldsymbol{x}$ is rarely unique. The experimenter usually has considerable latitude in what to measure, and how to carry out the experiment. Experimenters should seek to identify primary physical parameters $\boldsymbol{x}$ at the most elementary level possible. That is, they should correspond to those physical entities that will actually be measured, or otherwise be determined directly, and that are as uncorrelated to each other as possible so that in many cases they can be represented by independent variables. This may mean that $\boldsymbol{x}$ could be a rather large set of parameters, perhaps numbering in the hundreds, or even many thousands, for a modern nuclear experiment. If experimenters make an effort to choose independent primary parameters to characterize their experiments, they will be rewarded by not having to be as concerned with estimating their correlations. Propagation of uncertainties from primary parameters to derived parameters will then be relatively straightforward. However, when it is impossible to avoid certain correlations, orthogonal transformations can be used to facilitate the analysis procedure.

It is understood from statistical theory that mean values $\boldsymbol{x}_0$ correspond to weighted averages of primary random variables $\boldsymbol{x}$ with respect to known probability distributions, or at least to averages calculated from reasonably large numbers of sampled values. Unfortunately, it is rarely possible to perform sufficient repetitions in nuclear experiments to fulfill this requirement. An investigator's knowledge of a particular primary parameter frequently stems from a single experimental determination. The mean value x_{0i} for variable x_i is then taken to be the single measured value x_{mi} of that variable, *i.e.*, $x_{0i} \approx x_{mi}$, or in vector notation, $\boldsymbol{x}_0 \approx \boldsymbol{x}_m$. This is certainly the case for primary parameters like detector counts, sample masses, sample thicknesses, detector efficiencies, *etc*. In discussing experimental data, the vector $\boldsymbol{x}_0$ is therefore equated to $\boldsymbol{x}_m$, with the understanding that rarely is it a collection of mean values in the statistical sense, but instead it corresponds to a collection of specific measured results that are assumed to be reasonably comparable to the mean values.

Standard deviations s_{xi} for these primary variables are often estimated quite subjectively, rather than being derived from statistical analyses. The experimenter stipulates that the postulated uncertainties for certain primary variables of the experiment are reasonable based on experience [3]. Not all uncertainty estimates are necessarily this subjective. For measured detector counts N, the usual approach, based on Poisson statistics, is to adopt the value $N^{1/2}$ as an approximation to the standard deviation. This is acceptable provided that N is sufficiently large [3, 16]. For detector calibrations, objective uncertainty estimates can often be deduced from detector calibration procedures. It is important, regardless of the method used, that experimenters strive to provide standard deviation estimates that correspond to one-sigma uncertainties, *i.e.*, to a confidence level of $\approx$ 68% that the quoted values x_{0i} should lie within the ranges $x_{0i} \pm s_{xi}$. Larger or smaller estimates of uncertainty, especially with different levels of confidence for different primary parameters of the experiment, should be avoided because this corrupts the entire process of generating covariances $\boldsymbol{V}_y$ for the derived results $\boldsymbol{y}$ that can be interpreted by data users as satisfying the one-sigma uncertainty condition. Data users are likely to be misled in assessing the quality of the reported experimental values if there is not a tacit understanding that a "best effort" was made by the experimenter to adhere to the one-sigma rule.

b. Modeling an experiment Modeling an experiment, and selecting the primary parameters to characterize it, are closely related aspects of the experimental process. For example, if it is intended to measure a neutron total cross section by transmission, it is necessary to specify density and thickness of the transmission sample. They are primary parameters of the model for this particular experiment. However, the process of experiment modeling involves more than just specifying the primary parameters. Mathematical expressions that specify the relationships between derived variables

$\boldsymbol{y}$ and primary variables $\boldsymbol{x}$ are required when analyzing the measured data. Just as the choice of primary parameters is not unique, the development of an experiment model that relates them to the derived parameters, and dictates how the data are to be analyzed, is also not unique. Some experiment models may be very detailed while others are rather sketchy. In general, the more detailed the model of an experiment, the better the outcome, since the experimenter is less likely to overlook factors whose neglect could lead to biases (errors) in the derived results. For example, some of the cross sections archived in EXFOR [6], for certain neutron reactions at energies $\approx$ 14 MeV, often measured many years ago, are considered to be excessively large. This effect can be observed by using an on-line EXFOR data plotting program, *e.g.*, that found at http://www.nndc.bnl.gov/exfor/. A particular case to examine involves the ^{55}Mn(n,γ)^{56}Mn reaction. The procedure is: Check the boxes and enter the following parameters into the template provided by this program: Target (MN-55), Reaction (N,G), Quantity (CS), Energy from (10) to (20) MeV. Then, click on "Submit". Finally, check the "All" and "Quick Plot" boxes on the next screen, and click on "Retrieve" to generate a survey plot. It is suspected that the data analysis procedures employed by many early investigators neglected to treat the additional reaction yield produced by neutrons scattered in the laboratory environment in analyzing their measured data. Measurements of cross sections for reactions with no threshold (*e.g.*, certain fission and all capture cross sections), or reactions with thresholds well below 14 MeV (*e.g.*, certain (n,p) and (n,α) processes) are vulnerable to this effect.

For simple experiments, or individual aspects of more complicated experiments, developing workable models can be fairly straightforward, and they can be quite comprehensive. In other situations, it is necessary, for practical reasons, to develop experiment models that involve simplifying approximations regarding the experimental setup. Generally, the models used to derive the results $\boldsymbol{y}_0$ from the measured data $\boldsymbol{x}_0$ (and fixed parameters $\boldsymbol{\eta}$) tend to be more detailed than those used for corresponding uncertainty analyses. The reasons are largely pragmatic ones. In order to obtain the best possible values for the derived parameters, it is necessary to take into account as many details of the experiment as possible, even if some of their effects are relatively small. However, models used for uncertainty estimation generally focus on aspects of an experiment that are believed to generate the most significant sources of uncertainty. Small uncertainty components, attributable to minor aspects of the experiment, are often estimated rather crudely, or they may even be neglected. Identification of primary parameters that are likely to involve significant sources of uncertainty in a particular experiment must be done by original investigators, since they are the persons best acquainted with the experimental details. Evaluators, regardless of their experience, are not equipped to do as good a job as the experimenters.

c. Data analysis and uncertainties Given a set of parameters $\boldsymbol{x}_0$ (actually, $\boldsymbol{x}_m \approx \boldsymbol{x}_0$), the covariance matrix $\boldsymbol{V}_x$, and the fixed parameters $\boldsymbol{\eta}$, the next step is to determine $\boldsymbol{y}_0$ and $\boldsymbol{V}_y$ using the model (or models) developed for data analysis. For simplicity, it is assumed that the same model can be used to obtain both $\boldsymbol{y}_0$ and $\boldsymbol{V}_y$, although, as mentioned above, this may not be the case for complicated experimental situations. Some models, especially those representing simple experiments, will involve a collection of functions $\{f_k\}$. Then, the derived values $\boldsymbol{y}_0$ can be calculated using analytical formulas: $y_{0k} = f_k(\boldsymbol{x}_0, \boldsymbol{\eta})$. Where appropriate, the deterministic Law of Error Propagation, as given in Eq. (42), can be used to determine the covariance matrix $\boldsymbol{V}_y$.

However, the data analysis procedures dictated by the experiment model are often so complex that some, or all, of the relationships between primary and derived variables must be represented by numerical algorithms rather than analytical functions. Then, it is more appropriate to use the notation $\mathfrak{F}\{\boldsymbol{x}_0, \boldsymbol{V}_x, \boldsymbol{\eta}\} \rightarrow \{\boldsymbol{y}_0, \boldsymbol{V}_y\}$ to signify the fact that $\boldsymbol{y}_0$ and $\boldsymbol{V}_y$ are obtained by applying the algorithm $\mathfrak{F}$ to $\boldsymbol{x}_0$, $\boldsymbol{V}_x$, and $\boldsymbol{\eta}$. When the experiment model involves algorithms such as $\mathfrak{F}$, the Monte Carlo method may be simpler to use for evaluating the covariance matrix $\boldsymbol{V}_y$ than the deterministic approach. A routine can be prepared to perform repetitive applications of the algorithm $\mathfrak{F}$, in which some or all of the parameters of $\boldsymbol{x}$ are varied randomly within ranges governed by $\boldsymbol{x}_0$ and $\boldsymbol{V}_x$. If only a few primary variables belonging to $\boldsymbol{x}$ are varied, and the remaining ones are treated as constants, this would correspond to using a simplified model of the experiment for the uncertainty analysis. Then, the same algorithm $\mathfrak{F}$ could be used to conveniently determine both $\boldsymbol{y}_0$ and $\boldsymbol{V}_y$. This may be adequate for all practical purposes. The Monte Carlo approach to generating the covariance matrix $\boldsymbol{V}_y$ may be computationally intensive, but that time is likely to be relatively modest compared to the time actually spent in carrying out the measurements and analyzing the data.

2. *Simple experiment examples*

Three simple examples of nuclear data experiments are presented here for demonstration purposes.

Example 13 This experiment measures the neutron total cross sections σ_{t1}, σ_{t2}, and σ_{t3} for aluminum at three distinct energies, E_1, E_2, and E_3. Aluminum is mono-isotopic (^{27}Al) and chemically stable. A pure metallic sample with uniform thickness δ is used. Let n_{27} represent its atomic density. Homogeneous, parallel beams of monoenergetic neutrons are provided, and a single, stable detector, with efficiencies ε_1, ε_2, and ε_3 for the indicated neutron energies, is used in the measurements. Neutron total cross sections are usually measured by comparing sample-in and sample-out count rates. Suppose the neutron flux on the sample is constant in time and the same measurement time, τ, is selected for the measurements. Also, assume that time τ can be established so accurately that there is negligible uncertainty from this source. Let N_{Ik} represent the sample-out counts at energy E_k, and N_{Tk} the corresponding sample-in counts due to the neutrons transmitted through the sample ($k = 1, 2, 3$). Subscript "T" indicates "transmitted", and subscript "I" indicates "incident". The mathematical model for this experiment is

$$\varphi_{Tk} = \varphi_{Ik} \exp(-n_{27}\, \delta\, \sigma_{tk}) \quad (k = 1, 2, 3). \tag{52}$$

Here, φ_{Ik} is the incident neutron flux, and φ_{Tk} is the transmitted neutron flux through the sample at energy E_k. Detector counts, not fluxes,

are actually measured. Thus, $N_{Ik} = \varphi_{Ik}\varepsilon_k\tau$ and $N_{Tk} = \varphi_{Tk}\varepsilon_k\tau\,(k = 1,2,3)$. The efficiencies ε_k of the detector are assumed to be count-rate independent. Thus, Eq. (52) reduces to

$$N_{Tk} = N_{Ik}\exp(-n_{27}\,\delta\,\sigma_{tk}) \quad (k = 1,2,3). \tag{53}$$

The time and efficiency parameters cancel, leading to a simple model for the experiment. Transmission measurements of total cross sections are self-normalizing, since no standard cross section information is required. Eq. (53) can be rewritten to provide an explicit formula for the derived total cross section σ_{tk} in terms of the measured parameters, *i.e.*,

$$\sigma_{tk} = [1/(n_{27}\delta)]\ln(N_{Ik}/N_{Tk}) \quad (k = 1,2,3). \tag{54}$$

This equation provides a mathematical representation of the model for this experiment that can be used for analyzing the experimental data.

In this experiment there are eight measured parameters and three derived ones. For convenience, let $x_1 = N_{I1}$, $x_2 = N_{T1}$, $x_3 = N_{I2}$, $x_4 = N_{T2}$, $x_5 = N_{I3}$, $x_6 = N_{T3}$, $x_7 = n_{27}$, and $x_8 = \delta$. Likewise, let $y_1 = \sigma_{t1}$, $y_2 = \sigma_{t2}$, and $y_3 = \sigma_{t3}$. The eight primary parameters in this experiment are clearly independent. This will not be the case for the derived parameters since all three of them involve two common primary parameters, x_7 and x_8. Using the relabeled variables, Eq. (54) is replaced by the three equations

$$y_1 = f_1(\boldsymbol{x}) = [1/(x_7x_8)]\ln(x_1/x_2), \tag{55}$$

$$y_2 = f_2(\boldsymbol{x}) = [1/(x_7x_8)]\ln(x_3/x_4), \tag{56}$$

$$y_3 = f_3(\boldsymbol{x}) = [1/(x_7x_8)]\ln(x_5/x_6). \tag{57}$$

The components of $\boldsymbol{y}_0$ can be obtained directly from these formulas by substituting values from the vector $\boldsymbol{x}_0$ for those from $\boldsymbol{x}$ in Eqs. (55) - (57).

The uncertainties associated with total cross section measurements are usually modest, so nonlinear effects can be ignored and the deterministic approach represented by Eq. (42) can be used. The elements of the 3×8 sensitivity matrix $\boldsymbol{T}$, defined by $t_{ki} = (\partial f_k/\partial x_i)|_{\boldsymbol{x}=\boldsymbol{x}_0}$ $(k = 1,3;\ i = 1,8)$, need to be calculated. Twelve of these matrix elements are seen to be zero as follows: $t_{13} = t_{14} = t_{15} = t_{16} = 0$; $t_{21} = t_{22} = t_{25} = t_{26} = 0$; $t_{31} = t_{32} = t_{33} = t_{34} = 0$. The partial derivatives with respect to x_7 and x_8 are easily obtained. They are $t_{k7} = -y_{0k}/x_{07}\,(k = 1,2,3)$ and $t_{k8} = -y_{0k}/x_{08}\,(k = 1,2,3)$. This leaves six additional sensitivity parameters to evaluate. They are $t_{11} = 1/(x_{01}x_{07}x_{08})$, $t_{12} = -1/(x_{02}x_{07}x_{08})$, $t_{23} = 1/(x_{03}x_{07}x_{08})$, $t_{24} = -1/(x_{04}x_{07}x_{08})$, $t_{35} = 1/(x_{05}x_{07}x_{08})$, and $t_{36} = -1/(x_{06}x_{07}x_{08})$.

According to Eq. (42), the elements of covariance matrix $\boldsymbol{V}_y$ are

$$v_{yk\theta} = \sum_{i,j=1}^{8}(\boldsymbol{T})_{ki}\,v_{xij}\,(\boldsymbol{T}^+)_{j\theta} = \sum_{i,j=1}^{8} t_{ki}\,v_{xij}\,t_{\theta j}. \tag{58}$$

By the definition of matrix transposition, $(\boldsymbol{T}^+)_{j\theta} = t_{\theta j}$. $\boldsymbol{V}_x$ is diagonal, so $v_{xij} = 0$ if $i \neq j$. The double sum in Eq. (58) reduces to a sum over the index "i", *i.e.*,

$$v_{yk\theta} = \sum_{i=1}^{8} t_{ki}\,v_{xii}\,t_{\theta i} = \sum_{i=1}^{8} t_{ki}\,s_{xi}^2\,t_{\theta i}. \tag{59}$$

Six of the standard deviations that appear in Eq. (59) can be deduced from Poisson statistics. Thus, $s_{xk} \approx x_{0k}^{1/2}$ $(k = 1,6)$. That leaves only two standard deviations to be estimated by other means, *i.e.*, s_{x7} and s_{x8} for the two primary variables x_7 and x_8. The experimenter must estimate the uncertainties in n_{27} (x_7) and δ (x_8). These will generally be rather small for an experiment designed to yield results of high accuracy. The diagonal elements (variances) of $\boldsymbol{V}_y$ are deduced from Eq. (59)

$$v_{y11} = \sum_{i=1}^{8} t_{1i}\,s_{xi}^2\,t_{1i} = \sum_{i=1}^{8} t_{1i}^2\,s_{xi}^2, \tag{60}$$

$$v_{y22} = \sum_{i=1}^{8} t_{2i}\,s_{xi}^2\,t_{2i} = \sum_{i=1}^{8} t_{2i}^2\,s_{xi}^2, \tag{61}$$

$$v_{y33} = \sum_{i=1}^{8} t_{3i}\,s_{xi}^2\,t_{3i} = \sum_{i=1}^{8} t_{3i}^2\,s_{xi}^2. \tag{62}$$

Expressions for t_{ki} given above are substituted into these equations. Also, fractional uncertainties in x_k are defined as $f_{xk} = s_{xk}/x_{0k}$ $(k = 1,8)$. This yields the following formulas for the variances:

$$v_{y11} = \left[1/(x_{07}\,x_{08})^2\right]\left(f_{x1}^2 + f_{x2}^2\right) + y_{01}^2\left(f_{x7}^2 + f_{x8}^2\right), \tag{63}$$

$$v_{y22} = \left[1/(x_{07}\,x_{08})^2\right]\left(f_{x3}^2 + f_{x4}^2\right) + y_{02}^2\left(f_{x7}^2 + f_{x8}^2\right), \tag{64}$$

$$v_{y33} = \left[1/(x_{07}\,x_{08})^2\right]\left(f_{x5}^2 + f_{x6}^2\right) + y_{03}^2\left(f_{x7}^2 + f_{x8}^2\right). \tag{65}$$

Off-diagonal elements of $\boldsymbol{V}_y$ (covariances) are obtained from

$$v_{y21} = y_{01}\,y_{02}\left(f_{x7}^2 + f_{x8}^2\right), \tag{66}$$

$$v_{y31} = y_{01}\,y_{03}\left(f_{x7}^2 + f_{x8}^2\right), \tag{67}$$

$$v_{y32} = y_{02}\,y_{03}\left(f_{x7}^2 + f_{x8}^2\right). \tag{68}$$

By symmetry, $v_{y12} = v_{y21}$, $v_{y13} = v_{y31}$, and $v_{y23} = v_{y32}$. As expected, the off-diagonal elements of $\boldsymbol{V}_y$ can be attributed entirely to the uncertainties in x_7 and x_8. Correlation matrix $\boldsymbol{C}_y$ can be calculated from $c_{yk\theta} = v_{yk\theta}/(v_{ykk}\,v_{y\theta\theta})^{1/2}$ $(k,\theta = 1,3)$, using the formulas from Eqs. (63) - (68). Eqs. (63) - (65) show that each variance is a sum of four terms. The origins of these terms are exhibited in Table III. For convenience the notation $\omega = 1/(x_{07}\,x_{08})$ is used.

TABLE III. Origins of the components of the variances for each derived quantity y_k in terms of the primary variables x_i of the experiment.

	y_1	y_2	y_3
x_1	$\omega^2 f_{x1}^2$	0	0
x_2	$\omega^2 f_{x2}^2$	0	0
x_3	0	$\omega^2 f_{x3}^2$	0
x_4	0	$\omega^2 f_{x4}^2$	0
x_5	0	0	$\omega^2 f_{x5}^2$
x_6	0	0	$\omega^2 f_{x6}^2$
x_7	$y_{01}^2 f_{x7}^2$	$y_{02}^2 f_{x7}^2$	$y_{03}^2 f_{x7}^2$
x_8	$y_{01}^2 f_{x8}^2$	$y_{02}^2 f_{x8}^2$	$y_{03}^2 f_{x8}^2$

It is intuitively more challenging to understand the form of the off-diagonal covariance matrix elements given in Eqs. (66) - (68). Each expression in these equations is the sum of two terms, one corresponding to the uncertainty in x_7 and the second to the uncertainty in x_8. Recall that $v_{yk\theta} = \langle(y_k - y_{0k})(y_\theta - y_{0\theta})\rangle$ from the definition of covariance, where $\langle\cdots\rangle$ signifies averaging with respect to an assumed probability distribution. The deterministic linear approximation is embodied in the expressions $y_k \approx y_{0k} + (\partial f_k/\partial x_7)_0\,(x_7 - x_{07}) + (\partial f_k/\partial x_8)_0\,(x_8 - x_{08})$ and $y_\theta \approx y_{0\theta} + (\partial f_\theta/\partial x_7)_0\,(x_7 - x_{07}) + (\partial f_\theta/\partial x_8)_0\,(x_8 - x_{08})$ where, *e.g.*, $(\partial f_k/\partial x_7)_0$ represents the partial derivative of f_k with respect to x_7 evaluated at $\boldsymbol{x} = \boldsymbol{x}_0$. Furthermore, $(\partial f_k/\partial x_7)_0 = -y_{0k}/x_{07}$ and $(\partial f_k/\partial x_8)_0 = -y_{0k}/x_{08}$. If these

factors are considered, and we recall that both variables x_7 and x_8 are fully correlated for all the y_k, but are uncorrelated to each other, then $v_{yk\theta} \approx y_{0k}\, y_{0\theta} \left[(s_{x7}/x_{07})^2 + (s_{x8}/x_{08})^2\right] = y_{0k}\, y_{0\theta}(f_{x7}^2 + f_{x8}^2)$ for $k \neq \theta$, as in Eqs. (66) - (68). Note that only the variables x_7 and x_8 are considered in this treatment since we are only interested in the off-diagonal elements of the covariance matrix where the correlations are introduced.

It is instructive to examine some numerical results that can be obtained using the present formalism. The following analysis is based on Eqs. (63) - (68). For simplicity, assume that the total cross section values at each of the three energies are about the same. Actually, we are not interested here in actual values of these cross sections, but rather just the total and correlated fractional uncertainties. Consider just one data point (*e.g.*, the first one), and specify the fractional uncertainties in percent. Suppose $x_{01} = 100,000$ and $x_{02} = 80,000$, *i.e.*, 80% transmission. These are the recorded incident and transmitted detector counts. Then, $f_{x1} \approx 0.32\%$ and $f_{x2} \approx 0.45\%$, if statistical counting uncertainties are considered to be the only source of uncertainty in the measured neutron transmission. This is a reasonable assumption for the present example, but it is oversimplified in the case of actual experiments, where detector background effects need to be considered. Finally, assume that the sample density uncertainty is $f_{x7} = 0.1\%$ and the sample thickness uncertainty is $f_{x8} = 0.2\%$. The correlated uncertainty, which depends only on these sample properties, is therefore only 0.22% for this data point. It is independent of the number of recorded counts. The resulting total uncertainty in the measured total cross section (for the indicated numbers of detector counts) is $f_{y1} \approx 4.9\%$. This is rather poor accuracy for this type of experiment. How can the overall uncertainty be reduced? Suppose the total counts recorded are increased by a factor of 10, while maintaining the same transmission property (*e.g.*, by using the same sample), *i.e.*, $x_{01} = 1,000,000$ and $x_{02} = 800,000$. The resulting total uncertainty is $f_{y1} \approx 1.6\%$, which is considerably improved. Finally, if the sample thickness is increased so that the transmitted flux is 50%, *i.e.*, if $x_{02} = 500,000$, then the total uncertainty is only $\approx 0.6\%$. Then, the correlation coefficient is ≈ 0.13 between values at the three distinct energies. The achievable accuracy in a total cross section measurement is thus strongly dependent on the numbers of recorded counts and the sample transmission factor.

Considerable analysis is required to determine mean values and a covariance matrix for derived parameters, given the measured results and their covariance matrix. This is evident from Example 13, which describes a quite simple experiment (with its simple model). Other approaches to designing a total cross section experiment might have led to much more complicated models, *e.g.*, if different times had been used for each measurement and time uncertainty had proved to be a significant consideration. Corrections might have been needed to account for neutron beams that were not completely parallel, for neutron background effects that were not entirely negligible, for dispersive neutron-energy spectra rather than purely mono-energetic neutrons, *etc*. Nevertheless, the basic approach described above in Example 13 for analyzing the data, including the determination of covariances, would remain the same. Clearly, it behooves experimenters to design their experiments to be as simple as possible, consistent with achieving the intended objectives, and with the resources (time, materials, equipment, personnel, *etc*.) at their disposal. Experienced experimenters know that seemingly innocuous complications introduced during the course of an experiment can greatly complicate the data analysis procedures that follow completion of the measurements. These complications may also lead to larger uncertainties and/or difficulties in estimating them. This further justifies carefully considering the experimental procedures to be employed before the measurements actually begin, and assessing their potential impact on analyzing the accumulated data. The surest way for an experimenter to propagate uncertainties from primary to derived parameters, and to generate a proper covariance matrix for the experimental results, is to use the deterministic formalism (illustrated in the preceding example) or, alternatively, the Monte Carlo approach. Mistakes can be made easily when shortcuts are taken, especially when determining off-diagonal elements of covariance matrices.

The next example demonstrates use of the Monte Carlo approach in generating the covariance matrix for a set of results derived from an experiment.

Example 14 A sample is irradiated in a steady neutron beam in order to measure a neutron activation cross section. The sample is not radioactive prior to irradiation. The irradiation proceeds until the sample approaches saturation activity, *i.e.*, the formation rate of activated atoms by the considered reaction is exactly balanced by the decay rate of already radioactive atoms. At this point the neutron source is switched off abruptly. Immediately thereafter the sample activity is measured for a fixed time period. These assumptions are idealistic compared to what actually happens in most experiments of this genre, but they serve to provide a relatively transparent example that demonstrates use of the stochastic approach (Monte Carlo) to estimate the uncertainties. Further simplification is achieved by avoiding corrections for background, neutron multiple scattering, radiation absorption, depletion of sample atoms by transmutation, sample irregularities, *etc*.

The differential equation that governs the formation and decay of radioactivity in the irradiated sample is

$$\mathrm{d}N_{\mathrm{act}}/\mathrm{d}t(t) = N_{\mathrm{nuc}}\, \varphi_n\, \sigma_R - \lambda\, N_{\mathrm{act}}(t). \tag{69}$$

$N_{\mathrm{act}}(t)$ is the number of active atoms in the sample at time $t > 0$, N_{nuc} is the target nucleus density in the sample, φ_n is the steady incident neutron flux rate (neutrons per unit area per unit time), σ_R is the reaction cross section (units of area), and λ is the radioactivity decay constant (reciprocal of the mean lifetime). The solution to this differential equation, for $t \geq 0$ and the constraint $N_{\mathrm{act}}(t = 0) = 0$, is

$$N_{\mathrm{act}}(t) = (N_{\mathrm{nuc}}\, \varphi_n\, \sigma_R/\lambda)\left(1 - e^{-\lambda t}\right). \tag{70}$$

Clearly, $N_{\mathrm{act}}(t = \infty) = (N_{\mathrm{nuc}}\varphi_n\sigma_R/\lambda)$ is the saturation number of atoms in the sample. In practice, if the neutron source is switched off at a time $t_0 \gg \lambda^{-1}$, then $N_{\mathrm{act}}(t_0) \approx N_{\mathrm{act}}(t = \infty)$. In this experiment, the sample activity is counted for a time ΔT that begins immediately after the neutron source is switched off. Thus, the total number of atoms that will have decayed during this counting time interval is

$$\Delta N_{\mathrm{act}} \approx (N_{\mathrm{nuc}}\, \varphi_n\, \sigma_R/\lambda)\left(1 - e^{-\lambda \Delta T}\right). \tag{71}$$

This is very similar to Eq. (70). The number of decay events measured by the emitted-radiation detector is $N_R = \Delta N_{\mathrm{act}} \varepsilon_R$, where ε_R is the detector efficiency. A 100% decay branching factor is assumed. The total number of counts recorded by the detector that monitors the neutron flux during the irradiation period t_0 is $N_\varphi = \varphi_n \varepsilon_\varphi t_0$, where ε_φ

is the efficiency of that detector. Substitution of the expressions involving $\Delta N_{\rm act}$ and φ_n into Eq. (71), leads to the following expression for the cross section σ_R:

$$\sigma_R \approx \left[(N_R\,\lambda\,\varepsilon_\varphi\,t_0)/(N_\varphi\,N_{\rm nuc}\,\varepsilon_R)\right]\left(1-e^{-\lambda\Delta T}\right)^{-1}. \quad (72)$$

Three determinations are made of the reaction cross section, *i.e.*, σ_{R1}, σ_{R2}, and σ_{R3}. So, three yield measurements with the activation detector are required, *i.e.*, N_{R1}, N_{R2}, and N_{R3}, as well as three comparable measurements with the neutron flux monitor, *i.e.*, $N_{\varphi1}, N_{\varphi2}$, and $N_{\varphi3}$. Thus, there are twelve primary variables x_i ($i = 1, 12$), as defined in Table IV.

TABLE IV. Primary parameters and their assumed uncertainties (standard deviations) for a simple activation experiment.

Parameter	Physical quantity	Uncertainty[a]
x_1	N_{R1}	4% (uncorr)
x_2	N_{R2}	4% (uncorr)
x_3	N_{R3}	4% (uncorr)
x_4	$N_{\varphi1}$	3% (uncorr)
x_5	$N_{\varphi2}$	3% (uncorr)
x_6	$N_{\varphi3}$	3% (uncorr)
x_7	λ	1% (corr)
x_8	$N_{\rm nuc}$	0.5% (corr)
x_9	ε_R	2% (corr)
x_{10}	ε_φ	3% (corr)
x_{11}	ΔT	Negligible
x_{12}	t_0	Negligible

[a] Uncertainties for the primary parameters in this table are typical of what might be expected in such an experiment. Each uncertainty corresponds to an independent uncertainty attribute. However, they are labeled as either "uncorr" (uncorrelated) or "corr" (fully correlated) depending on how their individual effects impact on the three derived reaction cross sections.

The following three derived variables serve to represent the three derived cross sections: $y_1 = \sigma_{R1}$, $y_2 = \sigma_{R2}$, and $y_3 = \sigma_{R3}$. These definitions, those in Table IV, and Eq. (72), lead to the following formulas:

$$y_1 = (x_1\,x_7\,x_{10}\,x_{12}/x_4/x_8/x_9)\left[1-\exp(-x_7x_{11})\right]^{-1}, \quad (73)$$

$$y_2 = (x_2\,x_7\,x_{10}\,x_{12}/x_5/x_8/x_9)\left[1-\exp(-x_7x_{11})\right]^{-1}, \quad (74)$$

$$y_3 = (x_3\,x_7\,x_{10}\,x_{12}/x_6/x_8/x_9)\left[1-\exp(-x_7x_{11})\right]^{-1}. \quad (75)$$

It is evident from Eqs. (73) to (75) that correlated uncertainties in y_1, y_2, and y_3 will be introduced from uncertainties in the primary variables x_7, x_8, x_9, and x_{10}. The time variables x_{11} (ΔT) and x_{12} (t_0) are assumed to have negligible uncertainty. ΔT plays a part in the uncertainty analysis through the factor $(1-e^{-\lambda\Delta T})^{-1}$. For simplicity, it is assumed that the mean value for each of the ten primary variable parameters is unity. ΔT is also set to equal to unity, and t_0 is assumed to equal 10 (so that the condition $t_0 \gg \lambda^{-1}$ is satisfied). The goal of this exercise is to determine cross-section fractional uncertainties, and their correlations, not mean values or absolute variances, so it is evident from Eqs. (73) to (75) that there is no loss of generality in making these simplifying assumptions.

The Monte Carlo procedure involves randomly sampling the ten primary variable parameters independently many times, according to their expressed uncertainties (listed in Table IV). The continuous uniform distribution is used in this exercise. A collection of randomly derived values y_1, y_2, and y_3 is produced using Eqs. (73) to (75). A statistical analysis is then performed to produce the desired covariance matrix. Ten separate stochastic exercises were performed, each one comprised of 1,000 histories. Averages of these results were determined to demonstrate the possible outcome from one exercise of 10,000 histories. Six distinct quantities specify the covariance matrix for the derived cross sections. The numerical ranges of the standard deviations (in percent) obtained from the ten Monte Carlo exercises (given to two significant figures) are s_{y1} (6.0 to 6.3%), s_{y2} (6.0 to 6.3%), and s_{y3} (5.9 to 6.2%). The results for the off-diagonal correlation coefficients are c_{y21} (0.31 to 0.36), c_{y31} (0.31 to 0.38), and c_{y32} (0.32 to 0.37). Average values for these quantities (corresponding to 10,000 histories), with the correlation coefficients multiplied by 100 for more convenient visualization, are

	Std Dev	Correlations		
y_1	6.2%	100		
y_2	6.1%	33	100	
y_3	6.1%	35	34	100

This exercise demonstrates that 1,000 histories are not sufficient to achieve adequate convergence for the experiment in this example. However, 10,000 histories are sufficient. The following simple approach to estimating these uncertainties and their correlations was applied to test this conjecture. It can be shown that the uncertainty in the cross section attributable to the multiplicative factor $g(\lambda) = \lambda(1-e^{-\lambda\Delta T})^{-1}$ that appears in Eq. (72), is $\approx 0.4\%$, if the uncertainty in λ is 1% and $\Delta T = 1$ (with no uncertainty). The fully-correlated cross section uncertainties (in percent), obtained from combining individual component uncertainty values listed in Table IV in quadrature, is $s_{\rm ycorr} \approx 3.7\%$. If this correlated component is then combined in quadrature with the two random uncertainty components for each cross section, the total standard deviation values (in percent) are $s_{y1} = s_{y2} = s_{y3} \approx 6.2\%$. These values are almost identical to the Monte Carlo results. Furthermore, if the correlation coefficients are computed using the expression $c_{yij} \approx s^2_{\rm ycorr}/(s_{yi}\,s_{yj})$ the result $c_{y21} = c_{y31} = c_{y32} \approx 0.36$ is obtained. This value also agrees quite well with the stochastic results.

Why should an experimenter bother to pursue the more computationally laborious Monte Carlo approach if a simple deterministic "hand calculation", as illustrated in Example 14, yields essentially the same results? The short answer is that one should resort to simplified approaches to estimating covariance data only when the conditions warrant their use. Most realistic experiments are considerably more complex than the one in Example 14. Deterministic uncertainty propagation can certainly be used in such complicated situations, if the uncertainties are modest, but Monte Carlo analyses yield more reliable estimates of the uncertainties for the derived results whenever the primary variables have large uncertainties and nonlinear effects come into play. Under these circumstances, the Monte Carlo approach is an appealing option. Also, it may lead to simplifications in the analysis of experimental uncertainties, since the bookkeeping may very well be more straightforward.

The techniques discussed in this section can be useful to data compilers as well as to experimenters. The following

example discusses how a covariance matrix can be generated from information provided for an actual experiment in documents prepared by the original investigators [50, 51]. The work in question was completed over 27 years ago. The laboratory notes and raw data have long since been discarded, the laboratory itself has closed, one of the experimenters is deceased, the whereabouts of another experimenter (who was then a student) are unknown, and the third experimenter, who led this experiment, and is one of the authors of the present paper, has forgotten many of the experimental details. It would ordinarily be necessary in this particular situation for a contemporary compiler (in 2012) to rely completely on the available documentation for this experiment. This is a very common set of circumstances encountered in compiling (or evaluating) older data. Fortunately, the data for this particular experiment were compiled relatively soon after the experiment itself was completed [6]. It is advantageous for experimental data to be compiled as soon as possible following completion of the experiment. Unfortunately, this is not always the case. An examination of the documentation reveals that the experimenters made an effort to document the experimental procedures and sources of uncertainty (including correlations) [50, 51]. This simplified the compiler's task when the EXFOR compilation was prepared (see Accession Number 12898 in the EXFOR data library [6]).

Example 15 The activation method was used to measure ^{51}V(n,p)^{51}Ti reaction cross sections relative to the ^{238}U neutron fission cross section, from near threshold at 2.856 MeV up to 9.267 MeV. Forty-five approximately mono-energetic values for the ratio were obtained, with FWHM resolutions of 0.08 to 0.1 MeV below 4.9 MeV and 0.14 to 0.28 MeV above 4.6 MeV. Data corresponding to both these categories were acquired in the overlap region, 4.6 to 4.9 MeV. The ^{7}Li(p,n)^{7}Be reaction was used as a neutron source for the lower energies, and the ^{2}H(d,n)^{3}He reaction provided neutrons at the higher energies. Both of these neutron sources were used in the overlap region. Various details were considered by the investigators in converting measured data to cross-section ratios. ^{51}V(n,p)^{51}Ti cross sections were then derived using evaluated ^{238}U fission cross sections. The experimental details, including the uncertainty analysis, are documented in a laboratory report [50] and a journal article [51]. Attention is given in this example to the measured ratios and corresponding uncertainties. Nineteen distinct uncertainty attributes were examined. They are summarized in Table V.

Seven are uncorrelated (None), eight are fully correlated (100%), q = 15 and 17 are partially correlated (Par), and for q = 14 and 16 correlation is irrelevant (Irrel) since the uncertainties are negligible. The investigators in this experiment did not propagate uncertainties from primary to derived variables as recommended in the present paper. This task could not be performed today since insufficient information is available about the primary variables and algorithms that were originally used to produce these data. The experimenters estimated uncertainties and correlations at the level of the derived variables, in a manner similar to that used to provide the simple estimates discussed in Example 14. A covariance matrix can be constructed using Eq. (46) to combine partial uncertainties e_{yqk} and micro-correlations $c_{yqk\theta}$ obtained from information in the documentation [50, 51] (partly summarized in Table V). Fig. 4 shows combined uncorrelated uncertainties (Uncorrelated), combined correlated uncertainties (Correlated), and total uncertainties (Total) for m = 45 measured cross-section ratio values. Fig. 5 exhibits the macro-correlation matrix for this data set. The experimenters made a distinction between data measured using the two different neutron sources. This is important in determining the micro-correlation parameters for some uncertainty attributes considered in this work. In treating the uncertainty attributes q = 15 and 17, the linear "range" rule, $c_{yqk\theta} \approx 1 - \mathrm{abs}\,[(E_k - E_\theta)/(E_m - E_1)]$, was assumed by the experimenters for the micro-correlation parameters, but only for data points involving the same neutron source. For different neutron sources, $c_{yqk\theta} \approx 0$ for these two attributes. The information originally in EXFOR for this experiment was quite detailed, but still incomplete by contemporary standards. Combined uncorrelated and combined correlated uncertainties were documented for each experimental point. However, information on the partial uncertainties and micro-correlations for the individual attributes was limited. No formats were available in EXFOR in 1982 for specifying detailed micro-correlation data [6]. Expanded EXFOR format options have been developed and approved recently (2012) to enable detailed covariance data to be included in future compilations. Detailed covariance information cannot be provided in compilations such as EXFOR unless it is generated and documented by the original experimenters. When critical details about a particular experiment are omitted - or become lost - it is virtually impossible for compilers to speculate, with any degree of reliability, as to what they might have been.

TABLE V. Sources of uncertainty in the measurement of ^{51}V(n,p)^{51}Ti cross sections relative to ^{238}U neutron fission.

q	Description	% Unc	Correl
1	Time factors	0.2	None
2	γ-ray yield	0.3-47.8	None
3	Fission yield	0.7-1.5	None
4	Extrapolation[a]	1-2	None
5	(n,f) background	Neg[d]-3	None
6	(n,p) background	0.2-1.2	None
7	Geometry	1.5	None
8	^{51}Ti half life	0.1	100%
9	U sample assay	2	100%
10	V sample assay	0.2	100%
11	U deposit[b]	0.8	100%
12	Count efficiency	2.4	100%
13	γ-ray branching	1	100%
14	Sample orientation	Negligible	Irrel
15	Neutron source	2	Par
16	Environmental[c]	Negligible	Irrel
17	Scattering effects	1.4-2.1	Par
18	Geometry	1.5	100%
19	Neutron energy	0.5-19.5	100%

[a] Fission events lost in the fission deposit.
[b] U deposit thickness correction.
[c] Room-return fission events.
[d] Neg = Negligible.

The material in Example 15 originated from an experiment that is much better documented than most others found in the literature, in spite of its imperfections. The situation is less favorable for most of the experiments represented in EXFOR. This poses a major problem for nuclear data evaluators that must be addressed in the years ahead.

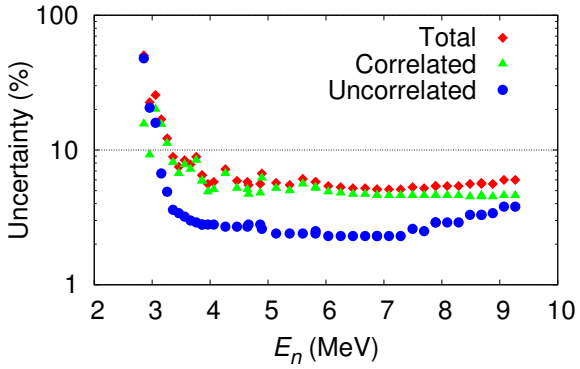

FIG. 4. Plots of uncertainty components for a set of measured ^{51}V$(n,p)^{51}$Ti-to-^{238}U fission cross-section ratios, as generated from information extracted from the original documentation of this experiment [50, 51].

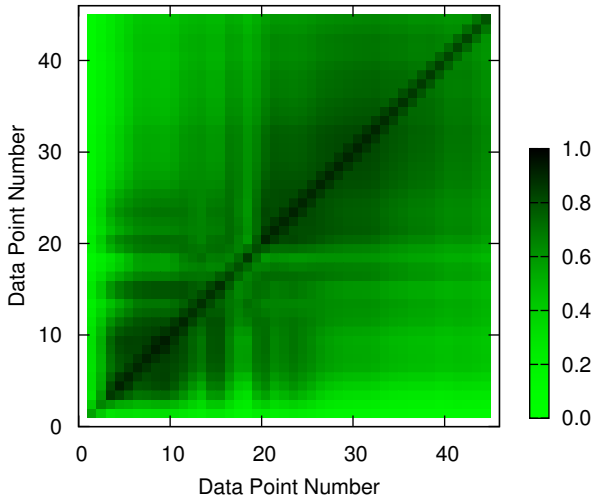

FIG. 5. Uncertainty correlation pattern for a set of 45 measured ^{51}V$(n,p)^{51}$Ti-to-^{238}U fission cross-section ratios generated from data found in the original reference [50, 51]. The x and y axes are keyed to the incident neutron energy, with the lowest energy given index "1" and increasing in neutron energy. The color scale indicates degree of correlation in percent. The low correlations seen between the first few low-energy data points and all the others are attributed to large statistical event counting uncertainties near threshold.

B. Special Situations

This section addresses some special situations involving the determination of covariance matrices for physical quantities (such as cross sections) that are derived from a collection of primary random variables. These cases find many useful applications in practice, and they can serve as useful templates for experimenters to consider when analyzing uncertainties in their experiments. For convenience in the following discussions, the subscripts "$_0$" will generally not be shown as they have been previously. However, it should be assumed that all numerical calculations involving the variables, and/or derivatives of functions with respect to these variables, will correspond to mean values.

1. *Product functions*

Consider a collection of n uncorrelated random variables x_i, that form a vector $\boldsymbol{x}$. They constitute the primary variables of an experiment, and their $n \times n$ covariance matrix $\boldsymbol{V}_x$ is diagonal. Then, define a vector $\boldsymbol{y}$ consisting of m elements y_k, *i.e.*, $\{y_k\}$. They are derived from the primary variables $\boldsymbol{x}$ through a collection of functional relationships $\boldsymbol{f}$, *i.e.*, $\{f_k\}$, according to $y_k = f_k(\boldsymbol{x})$. The covariance matrix $\boldsymbol{V}_y$ for $\boldsymbol{y}$ can be obtained by applying the deterministic Law of Error Propagation, *i.e.*, Eq. (42). In general, there are no conditions on the individual functions f_k other than requiring that they be smooth, differentiable functions of all, or subsets, of the primary variable set $\boldsymbol{x}$. Now, we consider as models for an experiment the class of those functions that have the general multiplicative form

$$y_k = f_k(\boldsymbol{x}) = \prod_{i=1}^{n} g_{ki}(x_i) \quad (k = 1, m). \tag{76}$$

Only three options are considered for the individual functions g_{ki}: $g_{ki}(x_i) = a_{ki}\, x_i$, $g_{ki}(x_i) = a_{ki}/x_i$, or $g_{ki}(x_i) = a_{ki}$, where a_{ki} is a constant. Also, a particular primary variable x_i is assumed to appear at most once in any one of the functions f_k, and it may not appear at all in some of them. The choice of g_{ki} is dictated by the role that the variable x_i is expected to play in the function f_k. Each derived variable y_k is formed by multiplying or dividing variables from $\boldsymbol{x}$. If x_i does not play a role in determining y_k, $g_{ki}(x_i) = 1$ is assumed. The primary variables that actually appear in a function f_k will usually be a subset of the entire set of variables from the complete vector $\boldsymbol{x}$. If a task in an experiment is to measure the yield of radiation twice, there will be two independent counts of that radiation using a calibrated detector. Each count is assigned its own primary random variable. The detector efficiency is common to both, so it can be represented by a single primary random variable. These three variables are mutually independent by their definitions.

A simple example will illustrate the nature of Eq. (76). Suppose $\boldsymbol{x} = (x_1, x_2, x_3, x_4)$, then two values y_1 and y_2 might be derived using the formulas $y_1 = f_1(\boldsymbol{x}) = x_1 \times (1) \times x_3 \times (1/x_4)$ and $y_2 = f_2(\boldsymbol{x}) = (1) \times x_2 \times x_3 \times (1/x_4)$, where the symbol "$\times$" represents scalar multiplication and all the $a_{ki} = 1$. Variable x_1 appears in f_1 but not in f_2, while x_2 appears in f_2 but not in f_1. The multiplicative factors "(1)" serve as "place keepers", for clarity. Normally it is unnecessary to do this, as long as their presence is clearly understood.

Since $\boldsymbol{V}_x$ is diagonal, Eq. (42) yields

$$v_{yk\theta} = \sum_{i=1}^{n} (\partial f_k/\partial x_i)\, s_{xi}^2\, (\partial f_\theta/\partial x_i) \quad (k, \theta = 1, m). \tag{77}$$

From the definitions given above, the indicated partial derivatives in Eq. (77) are as follows: $(\partial f_k/\partial x_i) = f_k(\boldsymbol{x})/x_i$, if

$g_{ki}(x_i) = a_{ki}x_i$; $(\partial f_k/\partial x_i) = -f_k(\boldsymbol{x})/x_i$, if $g_{ki}(x_i) = a_{ki}/x_i$; and $(\partial f_k/\partial x_i) = 0$, if $g_{ki}(x_i) = a_{ki}$. Thus, Eq. (77) can be rewritten as

$$v_{yk\theta} = \sum_{i\{k,\theta\}} \left\{ \pm \left[f_k(\boldsymbol{x})/x_i \right] s_{xi}^2 \left[f_\theta(\boldsymbol{x})/x_i \right] \right\} \quad (k, \theta = 1, m). \tag{78}$$

The notation "$i\{k, \theta\}$" in Eq. (78), signifies that the sum over "i" includes only those terms in which both of the partial derivatives involved (*i.e.*, for both f_k and f_θ) are non-zero. The notation "±" signifies that either sign might appear in particular terms of the sum. The appropriate sign, "+" or "-", in front of individual terms depends on the relative locations (numerator or denominator) of x_i in g_{ki} and $g_{\theta i}$. If x_i appears either in the numerator of both functions, or in the denominator of both functions, then the "+" sign is used. If x_i appears in the numerator of one function, but in the denominator of the other function, then the "-" sign is appropriate. The occurrence of negative terms in Eq. (78) can happen only for off-diagonal elements of the covariance matrix $\boldsymbol{V}_y$. The diagonal elements of $\boldsymbol{V}_y\,(\theta = k)$ will generally have more terms in their sums than the off-diagonal elements ($\theta \neq k$). The reason is that each diagonal term v_{ykk} includes a component in the sum for each variable x_i represented in the function f_k, *i.e.*, those where $(\partial f_k/\partial x_i) \neq 0$, whereas for each off-diagonal element of the covariance matrix, two conditions have to be satisfied simultaneously. Both of the partial derivatives appearing in Eq. (78) must be non-zero. This occurs only if a specific primary variable x_i is common to both f_k and f_θ. It happens that these are just the primary variables that introduce correlations between derived variables such as y_k and y_θ.

If both sides of Eq. (78) are divided by the product $y_k \times y_\theta = f_k(\boldsymbol{x}) \times f_\theta(\boldsymbol{x})$, and it is recalled that the fractional uncertainty in x_i is defined as $f_{xi} = (s_{xi}/x_i)$, then

$$r_{yk\theta} \equiv v_{yk\theta}/(y_k y_\theta) = \sum_{i\{k,\theta\}} \pm f_{xi}^2 \quad (k, \theta = 1, m). \tag{79}$$

The signs for the terms are governed by the same conditions indicated for Eq. (78). Recall that $r_{yk\theta}$, according to Eq. (79), is just an element of the relative covariance matrix $\boldsymbol{R}_y$, as defined in Eq. (31). The diagonal elements of $\boldsymbol{R}_y$ correspond to the squares of the fractional uncertainties for the derived quantities $\boldsymbol{y}$, *i.e.*, $r_{ykk} = f_{yk}^2 \quad (k = 1, m)$. The off-diagonal elements of $\boldsymbol{R}_y$ can be written as: $r_{yk\theta} = f_{yk}\, c_{yk\theta}\, f_{y\theta}$. Here, $c_{yk\theta}$ is the corresponding off-diagonal element of the correlation matrix $\boldsymbol{C}_y$. These definitions, and Eq. (79), lead to the following expression for the correlation coefficients:

$$c_{yk\theta} = \left(\sum_{i\{k,\theta\}} \pm f_{xi}^2 \right) / (f_{yk}\, f_{y\theta}) \quad (k, \theta = 1, m). \tag{80}$$

For the class of functions f_k defined by Eq. (76), it is relatively straightforward to calculate the derived variables and their fractional uncertainties and correlations, given the fractional uncertainties for the uncorrelated primary variables. All one needs to know is which of the primary variables are included in defining each of the derived variables, and which position (numerator or denominator) they occupy in the model, as expressed by Eq. (76).

Although this formalism might appear to be rather restrictive, it actually finds many useful applications. In the analysis of many experiments to determine cross sections, certain measurable primary variables often appear as multiplying or dividing factors (*i.e.*, scale factors) in the formulas used to determine cross sections. Examples of these are: detector counts, detector efficiencies, sample atoms, branching factors, *etc*. Furthermore, corrections to measured data to account for radiation absorption, second-order geometric effects, neutron scattering, *etc*., are often applied as multiplicative factors in deriving the cross sections. Under such circumstances, this mathematical formalism can provide a template, even if sometimes an approximate one, for analyzing experimental data uncertainties. However, for some complicated experiments this approach may be too crude. Better quality covariance data could be obtained if these experiments were modeled in greater detail, and the deterministic or stochastic methods described earlier were applied to generate the covariances. The simple "hand calculation" mentioned in Example 14 yielded covariance data that agreed quite closely with the results of a detailed Monte Carlo analysis involving 10,000 histories because the data analysis model used for this example is essentially consistent with the template just described. In this example there are three values y_1, y_2, and y_3 that are derived from twelve primary variables. Except for x_7 and x_{11}, which appear only in the factor $[1 - \exp(-x_7 x_{11})]^{-1}$, all the remaining primary variables correspond to scale parameters, thereby conforming to the experiment model embodied in Eq. (76). Since x_{11} is specified to be a constant (no uncertainty), the factor $[1 - \exp(-x_7 x_{11})]^{-1}$ can be equated to a single primary, scale variable. This was effectively done in Example 14, by noting that the uncertainty in this factor is $\approx 0.4\%$ if the uncertainty in x_7 is 1%. So, the "hand calculation" provided in Example 14 actually corresponds to applying Eqs. (79) and (80).

2. *Cross-section standards*

This paper recommends that experimenters report those results that they actually measure, *i.e.*, either primary parameters or quantities that can be derived directly from measured data through formulas or data reduction algorithms that are sufficiently well documented to enable users of the results to completely understand the essential features of the experiment. While some experiments yield "absolute" cross sections, *e.g.*, transmission measurements to determine neutron total cross sections, most neutron reaction experiments involve the measurement of ratios of "unknown" (or lesser known) cross sections to cross sections that are sufficiently well-known to be treated as standards [37]. Doing this avoids the difficult task of having to determine absolute neutron fluences. In these situations, experimenters should report measured ratios rather than derived cross sections. They should also provide sufficient details regarding how the standard reaction was used in the

experiment to enable data users to compute the corresponding "unknown" cross-section values from the given ratios using contemporary values for the standard cross sections. If the experimenter does choose to report derived cross sections, then the standard reaction used should be specified, and the actual standard cross-section values employed in deriving the reported "unknown" cross sections should be tabulated, along with the assumed uncertainties in these values. A data user will then be able to make adjustments to the derived cross-section results to account for any changes in the standard values during the intervening time between when the measurements were performed and when the data are used.

It could be argued that this should not be necessary if the standard cross sections correspond to values from well-established evaluated data libraries, *e.g.*, from ENDF/B-VII.1 [7], since it is assumed that these values will be available from data centers for many years after their release. However, experience has shown that these libraries are not always available indefinitely. For example, the three earliest versions of the ENDF/B library are currently unavailable from the U.S. Nuclear Data Center (NNDC) at Brookhaven National Laboratory [52]. ENDF/B-II and -III have been located in the data archives of Bettis Atomic Power Laboratory, and administrative procedures have been initiated to authorize the release of these two libraries to the NNDC for unrestricted distribution [53]. Unfortunately, it is likely that ENDF/B-I is lost forever unless, by chance, it happens to be discovered in some yet-to-be-identified laboratory's archives. This demonstrated fragility of long-term access to evaluated data libraries strongly supports the recommendation in this paper that experimenters document the actual values they use for standard cross sections. Of course, if a reported data set is very large, *e.g.*, numbering several thousand values, it may not be practical to provide corresponding values for the standard cross sections when it is reasonable to anticipate that, in most instances, they will be available electronically well into the future from libraries maintained by data centers. So, it is ultimately left to the judgment of an experimenter whether or not to tabulate cross-section standard values for reported large data sets.

In the following discussion, it is assumed that ratio data, including uncertainties, have been reported by the experimenter, and one seeks to calculate cross sections and their uncertainties for the unknown reaction, taking into account the uncertainties in the standard values as well as in the experimentally determined cross-section ratios. Let $\boldsymbol{x}$ represent a collection of $2n$ variables, and $\boldsymbol{V}_x$ signify its covariance matrix. The collection $\boldsymbol{x}$ is assumed to include the measured cross-section ratios and the corresponding values for the standard cross sections used in the experiment. Thus, $\boldsymbol{x} = (\boldsymbol{\rho}, \boldsymbol{\sigma}_s)$, where $\boldsymbol{\rho}$ is the collection of n measured ratios and $\boldsymbol{\sigma}_s$ is the corresponding collection of n standard cross sections. It is then assumed that the collection of measured ratios $\boldsymbol{\rho}$ is completely independent from the evaluated standard cross sections $\boldsymbol{\sigma}_s$. In most realistic situations this is a very reasonable assumption. This means that the covariance matrix $\boldsymbol{V}_x$ can be partitioned, as shown in Eq. (32), leading to the expression

$$\boldsymbol{V}_x = \begin{pmatrix} \boldsymbol{V}_\rho & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{V}_{\sigma s} \end{pmatrix}. \tag{81}$$

We then define $\boldsymbol{\sigma}_u$ to be the collection of n unknown cross sections generated by considering both the measured ratios $\boldsymbol{\rho}$ and the corresponding standard cross sections $\boldsymbol{\sigma}_s$. The covariance matrix $\boldsymbol{V}_u$ for the unknown cross sections $\boldsymbol{\sigma}_u$ can be obtained deterministically by applying the Law of Error Propagation, *i.e.*, Eqs. (42) and (45). This yields the result

$$\boldsymbol{V}_{\sigma u} \approx \boldsymbol{T}\boldsymbol{V}_x\boldsymbol{T}^+ \approx \boldsymbol{T}_\rho\boldsymbol{V}_\rho\boldsymbol{T}_\rho^+ + \boldsymbol{T}_{\sigma s}\boldsymbol{V}_{\sigma s}\boldsymbol{T}_{\sigma s}^+. \tag{82}$$

The two terms are a consequence of the partitioning of $\boldsymbol{V}_x$. Finally, it is assumed that the experimenter has provided a covariance matrix $\boldsymbol{V}_\rho$ for the experimentally determined ratios $\boldsymbol{\rho}$, and that the standard cross sections $\boldsymbol{\sigma}_s$ and their covariance matrix $\boldsymbol{V}_{\sigma s}$ are available from an evaluated data library (or from the literature). The "unknown" cross sections are related to the ratio and standard values by

$$\sigma_{ui} = \rho_i\sigma_{si} \quad (i = 1, n). \tag{83}$$

The dimensions of each of the collections of variables $\boldsymbol{\rho}, \boldsymbol{\sigma}_s$, and $\boldsymbol{\sigma}_u$ are identical and equal to n. Their covariance matrices have the dimensions $n \times n$. The measured ratios and standard cross sections used to derive the "unknown" cross sections, in accordance with Eq. (83), must be comparable. The two values must correspond to the same experimental conditions, *e.g.*, the neutron energy. Achieving this is not a straightforward matter, as the following discussion shows.

Quoted experimental energies are usually average neutron energies E_i. They are averages since measurements are rarely performed with purely mono-energetic neutron sources. The ratio data reported by an experimenter will consist of the collection of pairs of values (E_1, ρ_1), (E_2, ρ_2), $\cdots$, (E_i, ρ_i), $\cdots$, (E_n, ρ_n). The covariance matrix $\boldsymbol{V}_\rho$ for this experiment corresponds to these same energies. It is assumed to be complete, not simply a collection of standard deviations $s_{\rho i}$ with no correlation information provided. To apply Eq. (83), an experimenter also needs a set of appropriate standard reaction cross-section values, σ_{s1}, σ_{s2}, $\cdots$, σ_{si}, $\cdots$, σ_{sn}, corresponding precisely to the average energies E_1, E_2, $\cdots$, E_i, $\cdots$, E_n of the experiment. If cross sections for the standard reaction are available from an evaluated data library in ENDF-6 formats, then the interpolation rule(s) specified in the evaluated file can be used for this purpose [33]. It is not as straightforward to obtain the appropriate covariance matrix $\boldsymbol{V}_{\sigma s}$ for the standard cross section. Cross-section covariance data in ENDF-6 formats can be expressed using a variety of approaches [33]. Explicit covariance matrices, relative covariance matrices, or various component covariance sub-matrices corresponding to assorted energy-grid structures, may be found in evaluated data libraries. Point covariance data are not provided in ENDF-6, but rather uncertainty and correlation data based on defined energy-group structures are given [33]. That is, uncertainty and correlation data corresponding to energy intervals that are defined in specific covariance files are given. The

experimenter must calculate standard deviations and correlations corresponding to the appropriate experimental energy-grid structure, based on data provided in the evaluated covariance data file for the standard, in order to generate $\boldsymbol{V}_{\sigma s}$. The experimenter will generally need to employ a computer program to accomplish this task. There are options: actually write one independently, perhaps find relevant free software on-line, or even ask a colleague who happens to possess a licensed version of an appropriate program such as NJOY [54] or PUFF [55] to help out. Regardless of the approach chosen, the resulting matrix $\boldsymbol{V}_{\sigma s}$ could turn out to not be positive definite unless steps are taken to avoid this problem. For example, this might occur in situations where two or more of the experimental energies correspond to the same energy-group interval, as specified in the ENDF-6 formatted covariance data file for the evaluated standard reaction cross section. This will manifest itself in occurrence of duplicate rows and columns in the derived correlation matrix, resulting in the covariance matrix being singular. The *ENDF-6 Formats Manual* suggests possible ways to remedy this problem [33].

The sensitivity parameter matrices that appear in Eq. (82) can be evaluated by referring to Eq. (83). The elements of $\boldsymbol{T}_{\rho}$ and $\boldsymbol{T}_{\sigma s}$ are given by

$$t_{\rho ij} = (\partial\sigma_{ui}/\partial\rho_j) = \delta_{ij}\,\sigma_{si} \text{ and } t_{\sigma sij} = (\partial\sigma_{ui}/\partial\sigma_{sj}) = \delta_{ij}\,\rho_i. \tag{84}$$

In this formula, δ_{ij} is the Kronecker Delta function that equals 1 if $i = j$ and is 0 otherwise. Elements of covariance matrix $\boldsymbol{V}_{\sigma u}$ can be derived from Eqs. (82) and (84) as follows:

$$v_{\sigma uij} = \sigma_{si}\,v_{\rho ij}\,\sigma_{sj} + \rho_i\,v_{\sigma sij}\,\rho_j \quad (i, j = 1, n). \tag{85}$$

Each element of $\boldsymbol{V}_{\sigma u}$ is just the sum of two terms, one originating from the uncertainties in the measured ratios and the other from the uncertainties in the standard cross sections.

3. *External parameters*

Many physical parameters that are not measured directly by experimenters are used in reducing experimental data to the reported results. These may include values of decay half-lives, radioactive decay branching factors, sample material densities, isotopic abundances, and other physical quantities that are tabulated in the literature or handbooks. Even detector calibrations may be acquired by experimenters from secondary sources. For convenience these are referred to here as "external parameters". Whether or not such external parameters are treated separately from the analysis of uncertainties in an experiment is a matter best left to the judgment of the experimenter. In many situations it will not be practical to separate them. However, the accepted values for these parameters may change over time, and this can create problems for users of these data, *e.g.*, evaluators. If a particular external parameter is a scale factor, such as a decay branching ratio, then corrections to cross sections for such changes are easy to apply. In other situations, it may be more difficult to make the corrections. An example would be the half-life of radioactive decay. The sensitivity of derived cross sections to the decay half-life will depend on the counting times. The data can still be corrected for such changes in external parameter values and uncertainties as long as the actual values and uncertainties used by the experimenter, and the manner in which they were used, are specified in documenting the experiment.

C. Well-designed Experiments

There are choices that can be made by investigators in designing their experiments that will minimize some of the difficulties associated with the analysis of their measured data, including the assessment of uncertainties. Attention to such matters will not only simplify the job of analyzing the results from these experiments, but it will also insure that the results can be understood more easily by compilers for the preparation of EXFOR entries, and for more effective use in the evaluation of data for libraries such as ENDF/B.

It has been said that the ability to identify and understand the varied issues and problems that one encounters in life is a manifestation of good intelligence, while wisdom is the capacity to recognize which of these issues and problems one can realistically hope to influence favorably by one's own actions, and which of them are either beyond one's power to deal with, or represent situations where meddling could possibly lead to unfavorable outcomes. The intent of this section is to point out some considerations that an experimenter should keep in mind in planning well-designed experiments. Certain factors associated with the planning and execution of an experiment can, if addressed thoughtfully, have a beneficial influence on reducing the associated uncertainties. This subject is approached here from the perspective of data uncertainties, but the consequences go far beyond this single issue. There can be a favorable impact on reducing the cost of conducting an experiment, as well as on the quantity and quality of results that it will produce. It is possible to only touch the surface of this important topic. Each experiment is unique, or at least involves some unique features, so this discussion must be rather general in nature as well as incomplete. The approach taken here is to stimulate investigators to think about how to go about designing and conducting their experiments in an efficient manner, consistent with the intended objectives, rather than to introduce a litany of detailed steps that should be followed.

The exercise of intelligence and wisdom should begin as early in the experimental process as possible, preferably in the planning stage. Important factors that influence the conduct and outcome of an experiment often have their origins in decisions made before the measurements even begin. Once the measurements have begun, it is often too late to make very many significant changes in the way the experiment is conducted. Of course, some alterations during the course of an experiment, to deal with unforeseen circumstances, are inevitable. Wisdom entails flexibility, and a lack of flexibility can lead to failure or, at the very least, to missed opportunities. So, a well-designed experiment is one that is thoughtfully planned and executed from its conception to eventual docu-

mentation of the results, but that permits sufficient flexibility to contend with unforeseen problems, and to benefit from opportunities.

The following principles should guide the choice, design, and conduct of a successful nuclear data experiment:

(a) Select a topic considered to be worthwhile.

(b) Establish realistic goals for the experiment.

(c) Make efficient use of the available resources.

(d) Make sensible choices about what to measure.

(e) Strive to balance quantity versus quality of data.

(f) Aim for simplicity wherever possible.

(g) Seek to minimize experimental uncertainties.

Items (a) and (b) are not addressed in this paper since they are beyond its scope. Attention is focused on discussing items (c) through (g) in the context of the nuclear data field.

Those aspects of an experiment that, for practical reasons, lie beyond an investigator's control are the fixed (unalterable) features of the experiment, around which adjustments have to be made to the variable factors in planning the experiment. These fixed attributes cannot be altered when modeling an experiment, (although their uncertainties must be considered). Possible variations of those factors that can be adjusted should be explored during the planning phase, with the intent of optimizing the experimental process so as to achieve good results, minimize uncertainties, *etc*. Computer simulations, using models of the experiment, are very useful for this purpose.

Several decades ago a seminar for physics graduate students was presented by a Nobel laureate on the experiment that earned the Prize. This lecturer showed a slide of the apparatus used, and then offered the following apology: "If I had known at the time that this experiment would win the Nobel Prize, I would have attempted to make the apparatus look more attractive so it would have photographed better." No one can claim that making an experimental setup visually attractive will guarantee that the experiment will be successful and productive, or even win prizes, but a case can be made that there is a positive correlation between the appearance of an experimental setup and success of the experiment. If the experimental apparatus is visually presentable, it is more likely than not that the experimenter is giving attention to items (c) through (g) than might be the case if there is an appearance of sloppiness in the apparatus and approach to the measurements.

Some important ingredients of well-designed experiments are discussed in the following sections.

1. Efficient use of resources

In broadest terms, the principal "resources" of an experiment are time, people, and materials (samples, facilities, apparatus, *etc*.). By "efficient use" it is meant that these resources are not wasted or under-utilized. Only the individual experimenters can make these judgments for their particular experiments.

Let's consider the resource of time. The date when the experiment can begin and the total time available for a measurement program are often beyond the experimenter's control. An experimenter will generally be permitted access for a certain time period to a measurement facility, and perhaps will also have the use of specific sample materials, equipment, and other requisite resources for a limited time. The time available for subsequent analysis of measured data is generally more flexible, but even then time constraints often enter. Younger scientists need to complete their thesis or post-doctoral work in a timely fashion in order to proceed with their careers. More established scientists usually have additional professional responsibilities, *e.g.*, teaching or administrative work, that limit the time they can devote to analyzing experimental data and documenting their work. A typical nuclear researcher probably cannot lay out in advance a precise time line that governs all aspects of a particular experiment, unlike experiments performed during manned spacecraft missions, where time is precisely budgeted, and many months of "dry run" training and simulations are involved. Nevertheless, some careful planning in advance of the experiment will usually provide the nuclear investigator with a reasonable estimate of the time needed to successfully carry out the various aspects of an experiment, including the assessment of uncertainties. Planning can reduce the likelihood that "panic situations" will emerge, where the experimenter "runs out of time" before completing all the intended tasks of the experiment. This often is a consequence of devoting excessive time to certain tasks early in the measurement process, perhaps even out of proportion to their relative importance, simply because the eventual deadline is perceived as some vague point in the future. The phenomenon of "running out of time", and the inevitable sloppy shortcuts taken in such circumstances, is probably one of the major factors leading to either failed experiments or those with suboptimal outcomes. Also, the assessment of experimental data uncertainties is likely to be left undone or at best treated inadequately.

A simple example is provided to illustrate the concept of optimizing choices of measurements time in experiments.

Example 16 The experiment involves determining a particular reaction yield (Y) by measuring the counts (N) recorded in a calibrated detector with efficiency ε. The mathematical model is: $Y = N/\varepsilon$. It is assumed that the detector efficiency is known to 5%. Event counting uncertainty is governed by Poisson statistics, so it is assumed that the uncertainty for N recorded counts is $\approx N^{1/2}$. The counting and calibration uncertainties are clearly uncorrelated. How many counts should the experimenter acquire to achieve a desired accuracy of ≈6% for Y, and how much time ought to be spent achieving the results sought? Table VI gives the counting uncertainties, calibration uncertainty, and total uncertainties in percent to two significant figures for various possible experimental scenarios involving choices of accumulated detector counts N. The counting times are given for an anticipated count rate of 10 counts per second (cps).

Which counting option in Table VI constitutes the optimal choice for a well-designed experiment? Barring other considerations, the best choice would appear to be the one where $N \approx 1,000$. To reduce the total uncertainty to 5.0% would require between 10,000 and 50,000 counts. Then the counting statistical uncertainty would be negligible compared to the calibration uncertainty. This would require between 20 m to 1 and 1/2 h of counting time, rather than ≈

TABLE VI. Total counts and counting times required to achieve various levels of accuracy for a simple experiment.

Counts (N)	Count Unc	Calib Unc	Total Unc	Time (s)	Time (m)	Time (h)
100	10.0%	5.0%	11.2%	10	0.2	<0.1
500	4.5%	5.0%	6.7%	50	0.8	<0.1
1,000	3.2%	5.0%	5.9%	100	1.7	<0.1
5,000	1.4%	5.0%	5.2%	500	8.3	0.1
10,000	1.0%	5.0%	5.1%	1,000	16.7	0.3
50,000	0.4%	5.0%	5.0%	5,000	83.3	1.4
100,000	0.3%	5.0%	5.0%	10,000	166.7	2.8
200,000	0.2%	5.0%	5.0%	20,000	333.3	5.6

2 m. If limited measurement time is not a consideration, the experimenter could choose to employ longer measurement times, just to be "on the safe side". However, experienced investigators know that time limitations are always factors in experimental work. If the experimenter is striving for higher accuracy, it would be advisable to explore other options, *e.g.*, reducing the detector calibration uncertainty to < 5% (if possible).

People arguably constitute the most important of all resources, and the one that it is almost always limited. Attention should be given to insuring that the time spent by, and roles played by, people participating in an experimental team are wisely planned, with contingency options well thought out in advance. It is important that everyone involved in planning and conducting an experiment (including technical support personnel) have at least a rudimentary understanding of the importance of estimating experimental uncertainties and, wherever possible, taking steps to minimize them. They should be committed to carefully documenting the details that are needed for this purpose, throughout the course of the experiment. They should be encouraged, through formal training and "hands-on" experience, to develop an intuitive understanding of various aspects of the experiment that contribute to uncertainty, and to follow procedures aimed at minimizing these uncertainties and avoiding blunders. It should be appreciated that the consideration of uncertainties is an integral part of the experimental process, not an afterthought to follow completion of the measurements and data analysis.

The resource category "materials" refers to choosing sample properties, experimental apparatus, *etc*., so as to enhance, to the extent allowed by constraints, the quantity and quality of the data produced by the experiment, and to minimize uncertainties. The following example, taken from the literature, illustrates this point.

Example 17 Considerable attention is paid by researchers at Rensselaer Polytechnic Institute to selecting transmission sample thicknesses and measurement times (subject to practical constraints), in order to minimize the statistical uncertainties in their white-source measurements of neutron total cross sections. Details of their work, including the derivation of formulas and related numerical procedures they apply, have been published [56], so their approach is discussed here briefly in general terms. These investigators have developed approaches to determining optimal measurement conditions for both reasonably constant cross sections (as would be encountered with light nuclei) and strongly resonating cross sections (for heavier nuclei). For the latter case, the emphasis is on finding the best conditions for deducing resonance widths from the measured data. Their approach involves two distinct steps. First, given a fixed total measurement time, a determination is made of the optimal relative times to be spent on sample-in, sample-out, and background measurements. Optimal partitioning of the total available measurement time depends on the background-to-signal ratio as well as on the "optical thickness" of the sample (*i.e.*, the total cross section multiplied times the areal density of sample atoms). Once having determined these time allocations, the optimal optical thickness of the sample to be employed can be deduced from a numerical analysis. It has been determined from this work that even in situations where the background dominates the total numbers of counts, the sample transmission should not exceed 0.37. On the other hand, when the background is low, the sample transmission should not be lower than 0.08. According to these investigators, the optimal sample-out measurement time fraction is relatively insensitive to the relative intensity of the background, and it should occupy about 20% of the total measurement time. When the background is large, the time needed for background measurement needs to increase at the expense of the time spent on the sample-in measurement. As the background-to-signal ratio increases, the optimal sample transmission should be larger, implying that thinner samples should be used. In experiments to measure resonance parameters, it has been determined that minimization of the statistical uncertainty in the measured area under the cross-section curve of a resonance leads to reduced uncertainty in the deduced resonance strength. Extensive experimental and simulation studies by these investigators have substantiated predictions of the analytical formulas and numerical procedures to be used in determining optimal conditions for their measurements of neutron total cross sections [56].

2. *Deciding what to measure*

A challenge every investigator faces is deciding specifically what to measure to best achieve the experimental objectives, within existing constraints. In the nuclear data field, the experimenter is rarely exploring completely new territory, or seeking to discover previously unknown physical phenomena. Instead, these experiments are usually motivated by quantitative considerations. The experimenter intends to produce a data set that will provide improved information on the magnitude of certain physical parameters, usually cross sections, in a particular domain (*e.g.*, a range of energies). The experimenter should have in mind an accuracy goal, so that the results of the work will have a significant impact in this field, and not simply add more data points to a crowded body of pre-existing results.

An important consideration in carrying out an experiment is deciding how many data points to measure to adequately define the parameters in question for the chosen physical process. Experimenters know that this is governed by two fundamental factors. One is the inherent structure (*i.e.*, energy dependence) of the physical quantity (which generally is not completely understood), and the other is limitations in the

resolution (*i.e.*, energy resolution) achievable in the experiment. If the experiment is capable of high resolution, and it is suspected that the cross section for the physical process has considerable structure that might be of some practical consequence for applications, then the experimenter should measure as many data points as time and circumstances allow. Otherwise, the decision may be to choose to measure fewer data points with broader resolution, in order to obtain a better understanding of the average normalization of the cross section in question.

Similar deliberations are involved in the allocation of resources to various domains of the physical quantities to be measured, the most common being defined by neutron energy. For example, consider a threshold neutron reaction. Knowledge of the shape of the cross section near threshold is very important for the use of these data in neutron dosimetry applications [12]. The experimenter may decide that it is worthwhile focusing most of the available time (and other resources) to the often difficult task of measuring the very small cross sections encountered in the threshold region, especially if there are few existing experimental data points available there. Then, it may be decided to measure only a few data points at higher energies, with the remaining time and resources, to validate the normalization of the experimental procedure, and to provide a comparison with results previously reported in this region by other experimenters.

3. *Quality versus quantity*

The decision to focus on quality as opposed to quantity, or the converse, will be driven largely by the current status of the database for a particular reaction process. If the experimenter has developed a measurement capability, not previously available in the field, that will permit the exploration of a physical process or domain where only theoretical results were available up to that time, it may be decided to focus on quantity rather than quality, in order to "scope out" the region, and by that means provide a test of the theoretical predictions. The uncertainties associated with theoretically determined values are generally rather large, so an extensive collection of experimental data of modest quality could be valuable. On the other hand, if the experimenter is aiming to resolve a long-standing question concerning the magnitude of a physical quantity in a particular domain, and abundant (and unfortunately often discrepant) data are already available from the literature, the choice may be to perform an experiment to generate a few high-quality data points, *i.e.*, reliable values with small uncertainties, in an attempt to resolve the discrepancies.

4. *Virtue of simplicity*

Experiments should be designed to be as simple as possible. Some reasons why this is beneficial are:

- It is easier to carry out the measurements and analyze the data, without making mistakes (errors), when an experiment is as simple as possible.
- Members of the research team will be more likely to thoroughly understand and follow the experimental and analytical procedures involved if they are simple.
- Simple experiments can be represented by simple mathematical models and data analysis procedures. Estimating experimental uncertainties can be arduous if the experiment is unnecessarily complicated.
- Simpler experiments tend to lead to smaller data uncertainties. When many variables are involved, the uncertainties tend to accumulate, and their combined impact on the final results can become significant.

Two examples are mentioned here to clarify these points. Consider a neutron activation experiment where data are acquired over an energy range where the cross section is anticipated to be flat and fairly smooth. The data analysis procedure can be simplified, and uncertainties reduced, by attempting to measure each data point in essentially the same way, *e.g.*, using similar, steady beam intensities for the neutron producing source reaction, similar exposure times, similar waiting (cooling) times, similar radiation measurement times, *etc*. If this can be achieved, then non-linear relationships in the model relating the primary and derived variables will tend to exhibit similar behavior for the various data points. This tactic will simplify estimation of correlated uncertainties. While it might not always be possible to do this, due to unforeseen circumstances, it is nevertheless worthwhile to strive toward this goal. Likewise, considerable simplification in the analysis of measured data can be achieved if all data points are measured using like samples in similar geometries. Then, the application of corrections for effects such as radiation scattering and absorption can be applied more consistently, and correlated uncertainties can be estimated more readily.

Of course, a plausible opposing argument can be made for avoiding reliance on excessive uniformity in the design and conduct of an experiment. For example, it might be justified to measure a particular parameter more than once, under several rather different conditions, in order to expose any unforeseen systematic errors associated with a single experimental technique. An experimental approach that embraces several distinct methods might be inefficient, but that may be warranted if the experiment involves new and not well validated aspects, *e.g.*, poorly characterized sample materials, untested equipment, or recently developed data correction procedures. However, under conventional circumstances, where the experimental techniques are well tested, simplicity is probably the best option.

5. *Minimizing uncertainties*

Regarding uncertainties, one author expressed the following sentiment in a journal paper several years ago:

> *"... This work is dedicated to the memory of my dear mother Alice Müller-Schmid (1902-1978) who, however, might have found it of little use*

> *as she always tried to adhere to the principle of avoiding errors rather than estimating them ..."**
>
> *Jörg Müller [38].

While avoidance of errors in experiments is important, effort should also be devoted to estimating uncertainties. Robert Peelle expressed this idea as follows [3, 57]:

> *"What is the value of a collection of experimental results if no attempt has been made to estimate and document their uncertainties?"*

In addition to estimating uncertainties, effort should be expended in attempting to minimize them. A general way to minimize uncertainties is to focus the main effort in an experiment on carefully measuring those quantities that will have the greatest impact on the results to be derived from the acquired data. Once the mathematical model to be used for data analysis and uncertainty assessment has been developed, an experimenter can use the methods of sensitivity analysis to guide this process. Knowledge of the underlying theory of the physical processes being investigated can also be beneficial in this regard. The following example illustrates this point.

Example 18 Angular distributions for scattering and emission of radiation in nuclear processes can be represented by Legendre polynomial expansions. The number of terms required depends on the angular momentum involved, based on quantum-mechanical considerations. In the $(n,n'\gamma)$ reaction, gamma rays are emitted symmetrically with respect to 90° in the center-of-mass system (where 0° defines the incident neutron beam direction). Also, this is normally very nearly true in the laboratory frame of reference. Symmetry follows automatically from the fact that the appropriate Legendre expansions include only even-order terms. If the angular momentum involved in the gamma transition is L, then the highest-order polynomial to be included is of order $2L$. $E1$ or $M1$ transitions can be represented by $Y(\theta) = A_0 + A_2P_2(\theta)$, since $P_0(\theta) = 1$ for all θ. Angular distributions are usually measured to determine these Legendre coefficients (*e.g.*, A_0 and A_2). These parameters are extracted from the data by least-squares fitting of the measured gamma-ray yields, Y, at various angles θ. A well-designed experiment is one that involves Y measurements at angles θ specifically selected to provide the greatest measurement sensitivity for determining individual Legendre coefficients, with minimal interference from the remaining coefficients. Reduction of uncertainties in the measured yields will then translate to reduction of uncertainties in the derived Legendre coefficients. The uncertainty in Y, in relation to the uncertainties in A_0 and A_2, follows from considering the differential sensitivity expression $\delta Y = \delta Y_0 + \delta Y_2 = (\partial Y/\partial A_0)\delta A_0 + (\partial Y/\partial A_2)\delta A_2 = \delta A_0 + P_2(\theta)\delta A_2$. Since $P_2(\theta) = 0$ for $\theta \approx 55°$ and 125°, a measurement should be made at one or both of these angles to maximize the impact on determining A_0. Then, A_2 will not influence the result. The impact of A_2 on the measured yield will be greatest at $\theta = 0°$ or 180°, but these two angles are usually not conveniently accessible to measurement apparatus. A practical choice would be $\theta \approx 30°$ or 150°. Clearly, the effect of A_0 can never be eliminated in determining A_2. So, a set of four gamma-ray yield measurements performed at $\theta \approx 30°$, 55°, 125°, and 150° should provide an adequate data set for use in a least squares fit to determine A_0 and A_2, with the added bonus of providing a test for possible asymmetry of the experimental apparatus. An example of carefully positioning gamma-ray detectors to optimize determination of Legendre coefficients can be found in the work of Deleanu et al. [58]. They suggest placing detectors at 110° and 150°, since these are nodes of the fourth-degree Legendre polynomial function. They further quote Mihailescu et al. [59] to the effect that these various choices of detector positions can facilitate angle integration for gamma transitions with multipolarities up to $L = 3$.

Nuclear data experiments are often conducted to benchmark nuclear models. Nuclear models are characterized by theoretical functional relationships and variable parameters. Consequently, experimental data are used to determine reasonable values and uncertainties for these nuclear-model parameters. The concept of sensitivity can be applied to select specific measurements that provide the greatest impact in determining values for these model parameters, with minimal uncertainties. Suppose that $\{\sigma_i\}$ represents a collection of cross sections to be calculated using a nuclear model, and $\{p_k\}$ is a set of variable parameters of this model. Dimensionless sensitivity parameters, defined by $S_{ik} = (\delta\sigma_i/\sigma_i)/(\delta p_k/p_k)$, can be useful in deciding the best measurements to perform [60]. This approach was investigated by Fessler and Smith [61] and put to use by Plompen et al. [62].

6. *Additional comments*

Young scientists should learn to approach experimental work in a disciplined way, and to appreciate the importance of uncertainties. They should understand that the consideration of uncertainties should be an essential and integral aspect of experimental investigations if the results are to play useful roles in contemporary applications.

It is important for investigators to understand that the basic assumptions about what an experiment is intended to accomplish inevitably drive the procedures used to acquire and analyze the data. If these assumptions eventually prove to be flawed, the derived results may fail to fulfill expectations, or they may be misleading. For example, if the investigator is erroneously convinced that a particular gamma-ray transition should be $L = 1$ (when it actually it is $L = 2$) he may fail to make measurements at sufficient angles to unequivocally establish the correct transition multipolarity. If actual measured data (values of primary variables) are discarded along the way, and only the derived results are retained, it will be impossible to resurrect the primary information and subject it to an alternative data analysis scheme (*i.e.*, a revised model of the experiment). Therefore, it behooves all experimenters to retain as much content as possible from the collection of original measured values, along with estimates of their uncertainties. In earlier times, when data storage options were limited (printout, punched cards, *etc.*), this was often impractical. Since several terabytes of information can be stored these days at very low cost, there is no longer any reason, in most nuclear data experiments, to discard original measured values. Measured data are precious, because of the resources required to produce them. Experimenters should keep this point in mind

and not discard their raw data indiscriminately, while retaining only the analyzed results.

IV. SOURCES OF UNCERTAINTY

In this section we examine a few of the more common sources of uncertainty encountered in experiments, and discuss how uncertainty correlations are introduced by the analysis of measured data to produce the derived results.

A. Event Counting Uncertainties

A single detector count that influences only one derived physical quantity among several will introduce no correlations within the collection of derived results. However, correlations are introduced if this value influences several derived quantities, as the following example shows.

Example 19 Suppose that the counting of radiation decay events is affected by background. Background counts must be subtracted to obtain the desired results. Two measurements are made using the same detector and measurement conditions (time, geometry, *etc*.). Let N_1 be the raw counts from the first measurement and N_2 the raw counts from the second measurement. A single, separate measurement is made of the background using the same detector and conditions, leading to the value N_B. This value is used to correct both raw detector counts. Let Y_1 and Y_2 represent these corrected counts. The model for this experiment is $Y_1 = N_1 - N_B$ and $Y_2 = N_2 - N_B$. Detector counts are governed by Poisson statistics, so the standard deviations for Y_1 and Y_2 are: $s_1 = (N_1 + N_B)^{1/2}$ and $s_2 = (N_2 + N_B)^{1/2}$. The uncertainties in Y_1 and Y_2 are correlated due to the presence of the common background, and the fact that a single measurement of that background was used in deriving both Y_1 and Y_2. The correlated uncertainty component is $N_B^{1/2}$, so the covariance matrix $\boldsymbol{V}_Y$ for the derived values $\boldsymbol{Y} = (Y_1, Y_2)$ is given by Eq. (86). It is evident that the correlation coefficient is: $c_{Y21} = c_{Y12} = N_B/[(N_1+N_B)^{1/2}(N_2+N_B)^{1/2}]$. Therefore,

$$\boldsymbol{V}_Y = \begin{pmatrix} N_1 + N_B & N_B \\ N_B & N_2 + N_B \end{pmatrix}. \tag{86}$$

Background-to-foreground ratios cannot be reduced by counting samples for longer time periods. To reduce the background and associated uncertainties in this experiment requires that the experimental conditions be altered. This might be done by using better shielding, larger samples (enhancing the useful radioactivity yield), or various other means. The experimenter must be careful to avoid introducing additional problems while attempting to solve an existing one. For example, when larger samples are used, the corrections for radiation scattering and absorption are larger. This introduces larger uncertainties for these effects.

B. Detector-related Uncertainties

Calibration is usually the principal source of uncertainty in experiments that employ radiation detectors. If a detector is calibrated for a particular condition, and the calibration factor is applied to an entire set of derived experimental results, the corresponding uncertainty is fully correlated between these data. It is then relatively straightforward to include this uncertainty information in determining a covariance matrix for the experiment. It is more difficult to determine the correlations if each derived quantity involves a different detector calibration value (even if the same detector is used), and the correlations are partial rather than 100%. The following formalism can be applied in a relatively common category of calibration scenarios.

Let $\boldsymbol{y} = \{y_i\} \quad (i = 1, n)$ be a collection of physical quantities derived from an experiment. For convenience, it is assumed that $y_i = F_i \varepsilon_i$, where F_i represents the numerical factor corresponding to all other aspects of the experiment that enter into determining y_i, and ε_i is the corresponding detector efficiency. The derived results are directly proportional to the detector efficiencies, so efficiency is a scale factor. This is a common scenario in most experiments. It is then assumed that $\boldsymbol{V}_F$ is the covariance matrix for the factors F_i, and $\boldsymbol{V}_\varepsilon$ is the covariance matrix for the calibration values ε_i. Our objective is to determine the covariance matrix $\boldsymbol{V}_y$ for the set of derived quantities $\boldsymbol{y}$. This situation resembles that of normalizing measured ratio data to an evaluated cross-section standard to obtain absolute cross sections. The formalism represented in Eq. (85) is applicable. Thus, the elements of $\boldsymbol{V}_y$ are given by

$$v_{yij} = \varepsilon_i \, v_{Fij} \, \varepsilon_j + F_i \, v_{\varepsilon ij} \, F_j \quad (i, j = 1, n). \tag{87}$$

It is assumed that the factors F_i and covariance matrix $\boldsymbol{V}_F$ are obtained from procedures already discussed. At issue here is how experimenters might determine the collection of efficiency values $\{\varepsilon_i\}$ and their covariance matrix $\boldsymbol{V}_\varepsilon$. There is no unique way to approach this task. The details depend on characteristics of the detector and the calibration procedures involved.

Detector calibrations, and estimates of corresponding uncertainties, can sometimes be obtained by using standard radioactive sources that provide measured values at specific points within a range of interest. It is then assumed that the actual calibration curve throughout this range varies smoothly as a function of the measurement parameters. An empirical curve (*e.g.*, a polynomial function) is fitted to the actual measured points using the least-squares method [4, 24, 63]. This approach yields calculated efficiency values and a corresponding covariance matrix, with uncertainties that reflect the scatter of the measured calibration points relative to corresponding calculated points on the curve. However, whenever detector efficiencies at various energies are represented by polynomials with few parameters, strong correlations are introduced.

A more sophisticated calibration approach involves the use of deterministic or Monte Carlo detector simulation algorithms, *e.g.*, GEANT [64], that consider detector geometry, as well as properties of radiation interaction with matter, to calculate efficiencies, instead of relying on fitted empirical curves. These calculations can also be validated by measurements using calibrated sources wherever possible. While simulation codes can be effective in calculating detector efficiencies ε_i, estimation of the associated uncertainties and their correla-

tions, needed to generate the covariance matrix $\boldsymbol{V}_\varepsilon$, is more challenging. Reasonable estimates of the uncertainties can be obtained by comparing calculated efficiencies with the measured values obtained using standard sources. This can be done at only a few points, so uncertainty estimates elsewhere on the calibration curve must be generated by interpolation. Plausible estimates of the correlations can be provided using approximation techniques discussed earlier.

C. Geometric and Materials Uncertainties

One would tend to assume that primary parameters, such as geometrical dimensions of the experimental apparatus and sample material properties, are so well characterized that their uncertainties are negligible compared with most other sources of uncertainty. Often this is not the case, especially when the geometries and compositions for small samples of rare materials are not well established. Nevertheless, an experimenter will usually be able to provide reasonable estimates of the uncertainties in such primary parameters. If they enter into determining the derived results as scale factors, it is easy to propagate their uncertainties using the approach described earlier. If the relationships between certain primary parameters and the derived results are more involved, and complicated algorithms are required to model the experiment, then either deterministic sensitivity or Monte Carlo techniques must be used to propagate these particular uncertainties. Therefore, investigators should identify which primary parameters of the experiment model are scale factors and which enter into the model in more complicated ways. A hybrid approach can be pursued to generate the experimental data covariance matrix. Deterministic or Monte Carlo calculations can be employed to construct that portion of the covariance matrix involving those parameters implicated in complicated functional relationships, while the simpler approach mentioned earlier can be applied in propagating the uncertainties for scale parameters.

D. Uncertainties from Data Corrections

Factors C_q with values close to unity are often applied in multiplicative fashion to correct preliminary results that are derived using a simplified experimental data analysis model. These corrections account for various relatively small, unrelated experimental effects that are not considered in the experiment model. Typical corrections of this type are those for neutron multiple scattering, for secondary neutrons produced by sources with spectra distinct from the primary neutron spectrum, from radiation absorption, from second-order geometric effects, *etc*. In a well-designed experiment, these corrections should be either unnecessary, or kept as small as possible. However, most experiments require modest corrections of this nature to be applied in order to produce accurate results. The relationships between preliminary and final values can be expressed as $(\text{Final Result})_i = (\text{Preliminary Result})_i \times \Pi_q C_{qi}$. Additional contributions to the total uncertainties that accrue from applications of these correction factors can be included in the covariance matrix for the final results in the manner discussed earlier. It might be thought that the uncertainties in these correction factors could be estimated by applying the Monte Carlo method, where all primary parameters included in calculations of these corrections (*i.e.*, geometric parameters, cross sections, angular distributions, *etc*.) are randomly varied in the same manner used to generate uncertainty information for the preliminary results. However, this approach is often impractical, since the numbers of primary parameters involved in determining most of these corrections are too large, the uncertainties (and possibly their correlations) involved in dealing with these primary parameters are frequently rather poorly known, and the computational algorithms employed are usually quite complicated. The corrections, even though they may depart by only a few percent from unity, are usually calculated using sophisticated simulation codes that are not amenable to investigating the effects of perturbations of their input values. These values usually are evaluated, and processed cross-sections are taken from nuclear data libraries. Experimenters therefore resort to making plausible estimates of the uncertainties in the corrections, and even cruder estimates of the correlations. The following example demonstrates how this might be done in a specific situation.

Example 20 Suppose that corrections are required for the effects of scattered neutrons when measuring a neutron activation cross section relative to ^{235}U neutron fission. Assume this activation reaction has an energy threshold of 0.5 MeV, and that cross-section ratios $R_{af} = (\sigma_a/\sigma_f)$ are measured at several energies in the range $\approx$ 1 to 10 MeV. The cross-section ratio at each energy is proportional to the ratio of the true counts n_a and n_f in the detectors used to measure decay activity and fission events, *i.e.*, $R_{af} = F(n_a/n_f)$. These counts are attributable only to primary source neutrons that are unperturbed by scattering effects. F is an energy-dependent proportionality factor. The counts actually measured in the two detectors are given by $N_a = \zeta_a n_a$ and $N_f = \zeta_f n_f$, where ζ_a and ζ_f are correction factors to account for counting events produced by scattered neutrons. Neutron scattering adds to the activation and fission measured counts, so $\zeta_a \geq 1$ and $\zeta_f \geq 1$. Clearly, $R_{af} = F \times [(N_a\zeta_f)/(N_f\zeta_a)]$. Scale factors to apply these corrections are normally defined by expressions of the form $(R_{af})_{\text{corrected}} = C_a C_f (R_{af})_{\text{uncorrected}}$. Thus, $C_a = (1/\zeta_a) \leq 1$ and $C_f = \zeta_f \geq 1$. It is assumed that these correction factors are calculated using a Monte Carlo code that models the experimental geometry and primary neutron spectrum, as well as the scattered neutron environment, and that these procedures utilize cross sections from evaluated nuclear data libraries. For example, suppose that these calculations yield values of the correction factor C_f ranging from about 1.1 at the lower energies to about 1.3 at the higher energies. The corresponding correction factor C_a ranges from 1.0 near threshold (*i.e.*, no correction) to about 0.9 at the highest energies. The challenge is to provide plausible estimates of the uncertainties in the correction factors C_f and C_a. The scattering of neutrons in an experimental environment is largely governed by neutron elastic and inelastic scattering, along with neutron absorption. For instance, assume that an overall uncertainty in characterizing the secondary neutron spectra associated with these processes is $\approx$10%. One might then expect uncertainties in the differences from unity of the computed correction factors to be roughly of this same order. For the hypothetical situation stated

above, the anticipated uncertainties in the correction factor C_f should be in the range 1 - 3% and those for the correction factor C_a to be in the range 0 - 1%. There is some cancellation of the scattered-neutron corrections to the derived cross-sections, since the fission and activation counts appear as ratios in the cross-section derivation formulas. But, that cancellation is incomplete owing to the different magnitudes of these corrections. Regarding correlations, the correction factor uncertainties for data points closely spaced in primary energy are likely to be strongly correlated, while those widely separated in energy should be more weakly correlated. Approximate expressions for the energy dependence of these correlations, such as those provided by the prescriptions discussed earlier, could be applied for this purpose. The uncertainties in ζ_f and ζ_a should be positively correlated to each other since these corrections arise from the same source of scattered neutrons, and there is an increase in both fission and activation events due to the presence of these scattered neutrons. What should the magnitude of the correlations be? One can surmise roughly that they will be substantial, but not necessarily 100%. A plausible choice might be ≈ 0.75 (a strong correlation). The uncertainties in C_a and C_f will be negatively correlated by this amount, by virtue of their definition. The net effect of the corrections C_a and C_f on the derived cross-section ratios tends to be relatively small, since their effects partially cancel owing to this negative correlation. This treatment of correlations does not consider that scattering-correction uncertainties are energy dependent, so these correlations are likely to be variable.

Obviously, it is desirable to keep such correction factors as small as possible, so their uncertainty contributions will also be small. There are various ways to accomplish this through careful design of an experiment. For example, the smallest possible activation samples could be used, consistent with other experimental considerations. Extraneous material in the environment could be minimized wherever possible, *e.g.*, by using a low-mass target assembly for the accelerator-produced neutron source. Room-return background could be reduced by irradiating the samples as far from laboratory walls, vacuum pump assemblies, *etc*., as possible. When such measures are not feasible, an experimenter must do as good a job as possible in calculating the required corrections and estimating their uncertainties. It may be necessary to accept larger uncertainties in the experimental results, as a consequence of requiring these corrections, in order to compensate for unavoidably less than optimal experimental conditions.

E. "Psychological" Errors

Experienced experimenters no doubt will have encountered the following scenario at least once, and perhaps several times, during their research careers: An experiment is completed. The measured data have been analyzed, and the results from the investigation have been determined. The experimenter will undoubtedly proceed to compare these results with previously reported values. If the results happen to disagree by a "noticeable" amount from other reported, comparable data, the investigator will face the uncomfortable situation of having to decide what to do next. The anxiety experienced will be even more severe if the new data appear to be inconsistent with more than just one other data set. The temptation will be great for the experimenter to pursue one or both of the following steps: i) seek additional corrections to apply to the new data, or enhance existing corrections, often with shaky justification, so as to achieve better agreement with the other results; or, ii) arbitrarily enhance the uncertainties for the new data to reduce the discrepancy. Both steps are usually ill-advised. This is not an issue of "uncertainty", but rather one of "error". Poenitz referred to this as committing a "psychological error" [3, 65].

This section suggests a more reasonable (and certainly more ethical) approach to dealing with this situation. It is necessary to establish what is meant here by "discrepancy". Table I provides some guidance, since it is generally reasonable to assume that comparable measured data should be normally distributed. The likelihood that any single value should differ from the ensemble of comparable results by more than three combined standard deviations (*i.e.*, considering the total uncertainties of the alleged discrepant point as well as those to which it is being compared) is less than 1%. Since three-sigma occurrences are rather unlikely for a normal distribution, one must suspect that something could be wrong if a deviation of this magnitude is observed. As mentioned earlier, this statistical outcome does not automatically prove that the new data are flawed. Nevertheless, the burden is on the present experimenter to do something; the situation clearly cannot be ignored. Reviewers will likely not let the matter pass unresolved when the reported results are submitted for publication. The experimenter should review the measurement and data analysis procedures used in the experiment, as objectively as possible, to determine if something significant has been overlooked, or if some mistake has been made. This is difficult to do under the circumstances, but it is an important responsibility in experimental work (although certainly not a pleasant one). It is worthwhile to have an impartial colleague, whose experience and professional judgment the experimenter values, take a fresh look at the situation. If this process does not resolve the problem, or if it yields only a modest reduction in the discrepancy, the investigator should examine details of the results reported by the other experimenters, wherever possible. This could be done by studying the available documentation or discussing the issue with the authors directly. It is especially important to acquire a thorough understanding of the similarities and differences of the various experimental methods, *e.g.*, the apparatus employed, the calibration standards used, the sample properties, *etc*. For example, this might uncover the fact that different decay half-life and/or branching-ratio values were used in analyzing the various data. If none of these steps uncover the source of the problem, then the wisest course of action for the investigator is probably to go ahead and report his results as originally obtained, without arbitrarily applying dubious corrections or enhanced uncertainties. However, it is essential that the experimenter make an effort to carefully document the details of this work, in order to guide future attempts to resolve the discrepancy. Such discrepancies eventually do become resolved, and valuable insight is gained along the way to guide future work in the field.

A situation of some interest in the present context was reported in the May 2011 issue of *Nuclear News*. The following sentences are taken from an article in this magazine:

"FermiLab's Tevatron, in its final few months before closure, has produced something that does not conform to the consensus understanding of the fundamental nature of the universe."

...

"The study of collisions that produced W bosons (associated with weak nuclear force) and jets of smaller particles showed the later production of electrons and muons well in excess of what would have come from the decay of the W bosons, and also from Z bosons (the collisions produce what are referred to as 'diboson' events). Inferred from the electrons and muons is a particle with mass and other characteristics that are not predicted by the standard model."

"A key to this report is the extent to which error and uncertainty can be determined. If an anomaly can be shown to be at least five-sigma confidence level, it is counted as a discovery. If it is one-sigma, it is probably an artifact of experimental error. The CFD researchers have found that what they have observed is three-sigma. If the collisions in the Tevatron turn out to be ideal for creating diboson events that give rise to this presumed new particle, the accelerator's closure in four months[†] could leave this phenomenon in a frustrating three-sigma limbo, neither confirmed nor refuted".[*]

[*] *Nuclear News*, American Nuclear Society, May 2011, p. 69.

[†] Note: Fermilab's Tevatron was shut down on September 30, 2011.

The distinction between "error" and "uncertainty" is not made clear in this article. The Fermilab research team apparently obtained an experimental result which differed by three standard deviations from what would have been expected from the standard model, based on their estimated experimental "uncertainties" (not "errors"). In "Big Science" (research on important fundamental physical questions at large, expensive facilities), the criterion for being able to claim an unexpected discovery is very strict. It requires that a measured result differ by at least five standard deviations, based on estimated uncertainties, from what would normally be expected, in order to be considered a true "discovery". Furthermore, the possibility that this result could be the consequence of a random outcome (a fluke) should be ruled out to a comparably high degree of certainty. Research on fundamental physics issues of this nature is a "high stakes" activity that is usually associated with large budgets, professional career reputations, big egos, and many technical support jobs being "on the line". Therefore, it is not surprising that a five-sigma requirement of confidence is needed to convince the physics community that a discovery that could win the Nobel Prize is actually real. For practical reasons, it is reasonable to abide by the three-sigma test for identifying discrepancies in the nuclear data field, rather than the more severe five-sigma criterion just mentioned.

V. DOCUMENTING UNCERTAINTIES

Experimental results in the nuclear field are typically documented in reports, conference proceedings, journal articles, and electronic communications and compilations. Each of these (and experimenters will often resort to using more than one for reporting their data) offers both opportunities and limitations. For convenience the various possibilities for documenting experimental results are grouped into two categories: "publications" and "data compilations." In the first category, the material usually appears as printed hardcopy (although these days electronic versions will normally also be available). Included here are journal articles, conference proceedings, and reports. The second category is almost always electronic (although a few printed handbooks provide extensive tabulated nuclear data information, *e.g.*, [66]). It also includes formal data compilations as well as informal communications, *e.g.*, those privately distributed or posted on the Internet.

These two categories are distinguished in part by limitations in the volume of information that can be presented. Covariance data can be quite voluminous, since as many as $n\,(n+1)/2$ numerical values may be required to specify the uncertainties for n reported experimental parameters. The volume of material that can be provided in the category "publications" is quite limited for practical reasons, whereas in principle there are essentially no such limitations in the category "data compilations".

Another distinction between these categories pertains to data accessibility. Convenient access to mean-value and covariance data in electronic form is important for nuclear data evaluators, since they usually employ computer algorithms designed to manipulate large volumes of numerical data. Data that are available in printed form, but not electronically, are much less useful for such applications, since it is difficult to transcribe numerical information reliably from printed pages to computer files. However, plots of these data (including 2-D correlation plots) do tend to enlighten evaluators regarding the quality of the data, and to reveal peculiarities.

Another important issue, and one over which experimenters generally have little control, has to do with how long reported data will be available to users. Experienced experimenters are aware that many older printed reports are very difficult, if not impossible, to obtain. As for journal articles, the major ones are generally available from libraries, or on-line, for indefinite periods, and they can be obtained without too much difficulty. That is often not the case for lesser known journals. As for electronic sources, on-line data must reside on a server at some location. Whether or not they will continue to be available far into the future is unclear. It will depend on how well the databases are maintained on these servers. Fortunately, progress is being made in formally compiling and archiving experimental data, and making it available on-line from data centers, *e.g.*, the EXFOR system [6].

It is important to indicate the essential information experimenters should be responsible for reporting, so that the results of their work can be of potential value to data evaluators and data users in the future. These include:

- A sufficiently detailed description of the measurement and data analysis procedures to enable data evaluators and users to interpret and properly utilize the data.
- A clear specification of the standard reaction(s) used, where applicable, and actual numerical values for these standards. It is useful for experimenters to also provide the derived cross sections based on the standard employed, as a supplement to the mandated reporting of cross-section ratios (if that is what is determined in the experiment). That way, adjustments for future revisions in the standards may be applied easily, and the effects of these changes can be ascertained. Also of great value to data compilers and evaluators is specification of external parameters, such as decay branching ratios, half-lives, *etc*., that may be applicable to the interpretation and utilization of reported experimental results. Adjustments to the data for changes in these parameters can be made if the experiment documentation is sufficiently detailed and clear.
- A description of the distinct sources of uncertainty.
- Tables of the partial and total uncertainties (most conveniently expressed in percent).
- Either a complete correlation matrix, or information on the nature of any micro-correlations involved for the various partial uncertainties, so that anyone who is interested can construct a covariance matrix using the formulas and procedures presented in this paper.

These requirements are consistent with those stated in a report by Mannhart [4]. The relevant passage is:

> *"A complete description of the uncertainties of an experiment can only be realized by a detailed list of all the uncertainty components, their value and a specification of existing correlations between the data. Based on such information the covariance matrix can be generated, which is necessary for any further proceeding with the experimental data. It is not necessary, and (is) not recommended, that an experimenter evaluates this covariance matrix. The reason for this is that an incorrectly evaluated final covariance matrix can never be corrected if the details are not given. (Such obviously wrong covariance matrices have recently occasionally been found in the literature). Hence quotation of a covariance matrix is an additional step which should not occur without quoting a detailed list of the various uncertainty components and their correlations as well. It must be hoped that editors of journals will understand these necessary requirements."*

As mentioned above, the extent to which these requirements can be fulfilled through the various available conduits for documenting experimental results depends strongly on the volume of information involved. This issue is discussed in more detail in the following section.

A. Publications

Most scientists these days work in a "publish or perish" cultural environment. There is a hierarchy of prestige associated with the various documentation options. Journal publication is at the top, conference papers in the middle, and laboratory reports below that. Even in the category "journal publication", there exists a hierarchy of prestige that is well understood by workers in the field. Contemporary perceptions of the relative worth of these documentation options in furthering a researcher's career stem from long standing traditions, as well as from the perceived benefits to science of the peer review process. All journal articles and many conference contributions are peer-reviewed, whereas informal reports rarely are.

Unfortunately, the opportunities for reporting extensive uncertainty information are rather limited for journal articles, and they are especially limited for most conference papers. If an experimenter wishes to gain the credit for reported work that is afforded by publication in a refereed journal or conference proceedings, and insufficient space is available for documenting the experimental details and numerical data, then a supplementary report can be prepared to document this information for the benefit of interested users, including data compilers. The space limitations are not as severe for reports, since there is greater flexibility permitted in the manner of presentation. If space limitation happens to be a problem for a printed report, then the results can be made available electronically on the Internet using a readily accessible format. An example of this would be the archived reports prepared by the IAEA Nuclear Data Section (http://www-nds.iaea.org/publications/). References to detailed documentation can be provided in a journal article, conference paper, or widely distributed summary report.

If the number of data points produced by an experiment is ≤ 20, the reporting requirements specified above can generally be satisfied in a journal article (or perhaps even a conference paper). If the number of data points is ≤ 100, it may still be possible to document the necessary information in a journal article, provided that techniques discussed in this paper are used to minimize the number of uncertainty correlation parameters required to define a covariance matrix. An example is the work of Smith, Meadows, and Kanno (Example 15 of this paper) [50, 51]. In conference papers, experimenters will normally have to resort to plots in order to present the data generated by their work. Since this is of little practical value to data users, it is imperative that the authors also make their results available in more detail in a journal article, laboratory report, or on the Internet. Accessibility is a key factor in determining the potential usefulness of supplemental documents, *i.e.*, those that are not formally published. Laboratory reports that are available only in printed form, with limited numbers of copies, or reports or Internet contributions where access is impeded by administrative restrictions, tend to be of limited value. The reporting of numerical values for data sets exceeding $\approx 1,000$ data points is likely to be prohibitive for all printed documents and most electronic versions. Even for a few hundred to 1,000 data points, steps must be taken to minimize the quantity of numbers required to specify the

uncertainty correlations.

B. Numerical Data Compilations

There is no alternative to electronic documentation for experiments involving more than $\approx 1,000$ data points. These data can be distributed to users through local networks or the Internet. Access to retrieval of this material can range from open to tightly-restricted. Informal data sets are distinguished from formal data compilations in that they have not been archived by experienced data compilers using standard formats. As long as these numerical data are available as tables of energies, cross sections, and various uncertainty components (including correlations), using ASCII coding, users may find the information to be useful for plotting and other applications. Informal data collections also provide a convenient way for experimenters to back up the results of their experiments in private files, to share their results with selected colleagues prior to formal release, or to transmit the data to compilers at data centers. However, informal data collections will generally not be very highly valued by data evaluators for two reasons. First, they may not be easy to find, either locally or from the Internet, except through private communication with the authors. Second, evaluators generally prefer to work with data that have been vetted by experienced data compilers, and that have been archived in libraries using standard formats so they can be retrieved and handled with well-established computer routines. Experimenters should understand that their results will be of value in contemporary applications only if they are "packaged" in a way that enables data users to readily access and manipulate them. Data users need to have confidence that the data they use are of decent quality. For this reason, they turn to formal data compilations, *e.g.*, EXFOR [6], in order to be assured that the data they use have been subjected to some degree of quality control applied by experienced data compilers.

C. An Overview of EXFOR

A widely-available and unrestricted compilation of experimental nuclear reaction data is the EXFOR (EXchange FORmat) library. This library was developed to provide stewardship for the database of experimental nuclear-reaction cross sections [6]. It is maintained and updated as a cooperative endeavor involving several nuclear data centers, *i.e.*, the International Network of Nuclear Reaction Data Centres (NRDC) [67]. The world is divided into geographical regions, with different data centers assigned responsibility for providing nuclear data services to users in their respective regions, and also for compiling experimental data generated in these regions for inclusion in EXFOR. These centers constitute the aforementioned NRDC (*e.g.*, see http://www-nds.iaea.org/nrdc/) [67]. Producers of experimental data should contact the appropriate data center for their region when submitting new data to be compiled. Anyone seeking nuclear data can easily access a wide range of information, and benefit from other provided nuclear data services, at those data centers that offer open access to their data and other services via the Internet. By mutual agreement, members of the NRDC maintain up-to-date mirror images of the EXFOR data library content at their Internet sites. The EXFOR library continues to grow in content, and improve in quality, as new data are measured, reported, and compiled, as errors in archived data are discovered and corrected, and as new options for compiling and disseminating these data are developed and approved by the NRDC network. Some features of the EXFOR data library, especially those pertaining to experimental nuclear reaction data uncertainties (covariances), are discussed in the following three sections.

Some experimenters may not be familiar with how the flow of information from data producers to eventual data users occurs, within the contemporary environment where the EXFOR library plays a role. Fig. 6 is a flow chart of the process that is discussed in more detail below.

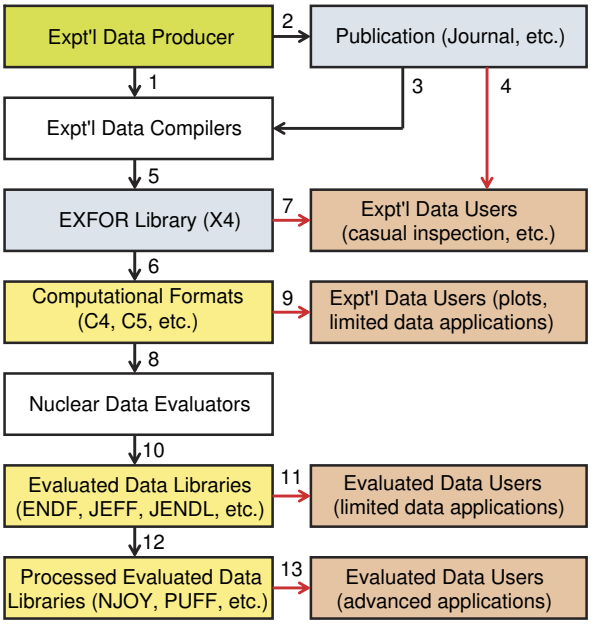

FIG. 6. The flow of data from experiment to users. Green: Experimental data are produced from measurements. Blue/Grey: Experimental data are archived as provided. White: Data are compiled and/or evaluated. Yellow: Original data are manipulated, evaluated, and processed. Orange/Brown: Data are available to potential users.

When an experiment is completed, an experimenter has various options for disseminating the generated information. It could be included in a formal data compilation. Fig. 6 implies that this would be the EXFOR library [6]. The data could be transmitted directly to the appropriate regional nuclear data center [67] where it would be compiled (Path 1) and archived in the EXFOR library (Path 5). Here, the term EXFOR is used to refer to both the library and the format system used to compile data in this library. Sometimes "X4" is used as a shorthand symbol to signify EXFOR. The EXFOR formats are described in the *EXFOR Formats Manual* [6]. Interested users can access EXFOR data on-line directly from

data centers (Path 7). It is also likely that these data will be published by the experimenter, either in a report, conference paper, journal article, or possibly in more than one such outlet (Path 2). If these data are not transmitted directly from the experimenter to a data center, then compilers at these centers can - and normally will - extract the data from publications and compile them in EXFOR (Path 3). Surveying the literature for new data is an on-going task for data center compilers. Users can also acquire data directly from publications if they choose to do so (Path 4). Within the EXFOR system there exist utility codes that can process data from the EXFOR format to produce the Computational formats, either C4 formatted files [68, 69] or, more recently, files in the newer C5 format [70]. C4 includes values of parameters and their total uncertainties (if available). C5 includes the same information as C4, but there are two additional columns, where total statistical and total correlated (systematic) uncertainties can be provided (if available). An approximate covariance matrix can be constructed from C5 data entries by making the simplified assumption that the total systematic uncertainties are fully (100%) correlated. Experimental data in the C4 and C5 formats are more convenient for automated operations such as plotting, evaluation, *etc*., than data in the EXFOR format (Path 6) [69]. Much of the text information recorded in EXFOR is stripped out in converting data from the basic EXFOR compilation format to C4 or C5, leaving only the essential numbers and descriptions (*e.g.*, experimenters, reaction types, energies, cross sections, and uncertainty information) needed for plotting, evaluations, and other applications. Users can also obtain files of processed data from the data centers (in either the C4 or C5 formats) in the same manner as for EXFOR formatted data (Path 9). Nuclear reaction data evaluators consider both experimental data and nuclear-model-calculated data in performing evaluations, *i.e.*, in generating recommended values. Acquisition and manipulation of experimental data from the data centers (in either C4 or C5 formats) is often handled automatically by the software used in the evaluation process (Path 8). The results from these evaluation efforts will eventually appear in libraries of evaluated nuclear data, *e.g.*, ENDF/B [7], JEFF [8], or JENDL [9] (Path 10). These evaluated data are benchmarked for quality assurance purposes by comparisons of calculated (C) and experimental (E) integral quantities (C/E ratios) for well-known and carefully characterized integral nuclear systems (data validation). Information from these evaluated libraries can also be acquired by users directly from the data centers. For example, evaluated data files can be downloaded from the Internet, or these data can be examined through on-line utilities (Path 11). Most advanced nuclear data users (*e.g.*, reactor physicists) require evaluated data that have been further processed (*i.e.*, manipulated into specialized group or point data formats) so that they can be employed by the computer codes which are employed to simulate complex nuclear systems. Examples of such nuclear systems would be nuclear power reactors or zero-power integral nuclear assemblies. Data processing codes such as NJOY [54] or PUFF [55] perform the task of producing these processed evaluated-data libraries. These processed libraries are derived from evaluated nuclear data obtained from the nuclear data centers (Path 12). Processed data libraries are usually included with software (*e.g.*, MCNP [71]), acquired from code centers such as RSICC [72] (Path 13).

Clearly, the paths from measured data to advanced applications of these data can be quite complicated, involving considerable data manipulation and assessment by compilers and evaluators along the way. The point made earlier in this paper that experimental results are rarely used directly in contemporary applications is substantiated by the process outlined in Fig. 6.

D. EXFOR Covariance Formats

The current EXFOR formats applicable to experimental nuclear data correlations were initially designed by the NRDC as an outcome of discussions held during the 1980's [73–75], following the *Workshop on Evaluation Methods and Procedures* held at Brookhaven National Laboratory during September 1980. The approved formats and procedures that evolved from this work are defined in the *EXFOR Formats Manual* and *EXFOR Compiler's Manual* (LEXFOR) [32]. These documents are readily available on-line from the IAEA Nuclear Data Center, and they are updated periodically as required. The reader can examine these documents, in conjunction with reading this section of the present paper, to gain a broader view of the data compilation capabilities available within the framework of this system. This paper focuses mainly on format issues associated with the compilation of nuclear reaction data uncertainties. They are illustrated using entries taken from the EXFOR library.

The earlier format options (*i.e.*, those in force prior to 2012) enabled experimental data-point values (cross sections, ratios, *etc*.) to be compiled along with corresponding constant or variable "statistical" uncertainties (under the data heading ERR-S) as well as several "systematic" uncertainty components (under the data headings ERR-1, ERR-2, ...). Furthermore, it has been possible to record constant micro-correlation coefficients $c_{y_q k_\theta}$, defined in Eq. (46), under the keyword ERR-ANALYS. If a variable macro-correlation coefficient was provided, it was recorded under the keyword COVARIANCE in the same EXFOR entry, or in a separate covariance file.

However, as seen in Example 15, information on uncertainties for individual data sets compiled in the EXFOR library have not always been complete. Some problems encountered by compilers are:

- Values of "total uncertainty" have often been supplied by experimenters to compilers without providing partial uncertainties.
- Lists of the ranges of partial uncertainties (*e.g.*, as in Table V) have often been submitted by the original experimenters without providing partial uncertainty values for each of the data points. Consequently, some compilers have coded the upper and lower boundaries of these given uncertainty ranges as just two constant partial uncertainties. This is wrong and very misleading. It demonstrates how the original intent of experimenters

can be perverted in compiling the data due to faulty or nonexistent communication between data producers and data compilers.

- Only one data heading (ERR-S) was available earlier in the EXFOR format options for "statistical uncertainty". Of course, there can be two or more distinct sources of statistical uncertainty for most experiments, *e.g.*, counting statistics for the reaction sample as well as for the monitor sample. This is not a fatal limitation of EXFOR, since statistical uncertainties from various experimental sources can be combined into one value for a single data point without creating any problems in constructing a full covariance matrix. However, data center compilers are not permitted - according to agreed-upon policy - to combine several uncorrelated uncertainties for individual data points into overall uncorrelated uncertainties without obtaining explicit permission from the experimenters.

- Only constant values for micro-correlation coefficients could be stored in a computer-readable form in the earlier EXFOR formats. This option has been exploited for the compilation of only two experimental works [76, 77], where 0.5 (moderate correlation) was assigned for some sources of uncertainty correlation, according to the "Occam's Razor" principle [41]. This is further evidence of the fact that very few experimenters appear to be generating detailed uncertainty correlation information for their data sets.

- If a particular correlation coefficient is specified by an experimenter to vary across a given data set, knowledge of this fact can be recorded under the keyword COVARIANCE as free text. However, this information is not amenable to automatic access from the EXFOR library by computer programs.

- Variations in the usage and interpretation of uncertainty terminology often create difficulties for compilers in deciding how to handle uncertainty data provided by experimenters. The earlier EXFOR formats defined two types of uncertainty sources: "statistical" and "systematic". However, a variety of terms relating to uncertainties appear in actual publications, *e.g.*, "random", "relative", "absolute", "shape", *etc*. Compilers are then faced with judging whether a given uncertainty value is "statistical" or "systematic" for the purpose of data entry into the EXFOR library. Also, usage of the term depends on the individual experimenters. For example, some experimenters use "relative uncertainty" in the context of the fractional uncertainty, but this is not so for other investigators. These ambiguities led to adoption of the generic data headings DATA-ERR, DATA-ERR1, DATA-ERR2, *etc*., with the consequence that the given information in EXFOR is generally not sufficiently informative for detailed uncertainty analyses, *i.e.*, construction of the full data covariance matrices.

The first two problems could be addressed - consistent with the constraints of the earlier EXFOR formats - if experimenters simply provided actual values for all the partial uncertainties associated with each data point. Resolution of the third and fifth problems mentioned above required that some modifications be made to the EXFOR formats. Also, the C5 format was introduced to augment the C4 option. These changes were approved by the NRDC network and adopted in 2012.

However, the final problem cited above can only be addressed (even considering the recent enhancements to the EXFOR formats) by direct contact between individual experimenters and data center compilers. It is clear, from the material presented in this paper, that the detailed characteristics of specific micro-correlations generated by experimenters should be specified and communicated to compilers in order to enable construction of proper full covariance matrices. Otherwise, EXFOR compilers are forced to make assumptions about these correlations. Furthermore, the EXFOR formats must allow recording of this provided information as coded data (*i.e.*, not simply as free descriptive text). Those modifications to the EXFOR formats that have been adopted in 2012, although they do not address every possible situation that might be encountered, do significantly enhance the capabilities of the EXFOR system to compile experimental data uncertainty information. Also, they have been designed so that they do not jeopardize the ability to handle data already compiled under the earlier EXFOR formats. In other words, the changes made provide enhancements but they do not introduce incompatibilities with the existing options (*i.e.*, they are "backward compatible"). The main changes that have been made are:

- The designations "statistical" and "systematic" should be avoided in future compilations. Instead, various uncertainty sources for an experiment are to be characterized by their correlation properties, as mentioned above. All partial uncertainties e_{yqk} in Eq. (46) will be coded in EXFOR under the data headings ERR-1, ERR-2, *etc*., unless more specific data headings are defined in the dictionary (*e.g.*, MONIT-ERR for the uncertainty in the standard cross section, ERR-D for uncertainty due to digitization by the compiler, *etc*.).

- However, to retain compatibility with the existing EXFOR coded entries (prior to 2012), the ERR-S and ERR-SYS options are retained. In the revised EXFOR formats they mean "total statistical" uncertainty or "total systematic" uncertainty, without additional details being provided about their origins. Similar steps have been taken to insure compatibility between the C5 formatted files and the C4 formatted files.

- Correlation properties in the updated EXFOR formats are indicated under the fourth field of the keyword ERR-ANALYS by the following correlation property flag options: F ("Fully correlated), P ("Partially correlated"), C ("Correlated", but its strength is unknown), or U ("Uncorrelated"). Note that the three flag options, F, P, and U, correspond to $c_{yqk\theta} = 0$, $0 < c_{yqk\theta} < 1$, and $c_{yqk\theta} = 1$ in Eq. (46). The first and third flags were coded as constant micro-correlation factors in the earlier EXFOR formats.

- The indicated correlation properties described in these formats are valid within one particular data set (typically, a data table in one subentry). That is, two data tables compiled by two different REACTION identification codes, and/or two different subentries, will be treated as independent.

When a partial uncertainty is defined with the correlation property flag P, the energy-dependent micro-correlation coefficients (or partial covariance matrix) $c_{yqk\theta}$ are required to construct macro-correlation coefficients properly, and they can be coded under COVARIANCE. Note that flags U and F signify that all micro-correlation coefficients are 0 and 1, respectively, so it is not necessary to record these numerical coefficient values explicitly. In addition to the micro-correlation coefficients, macro-correlation coefficients can also be coded under COVARIANCE. This is especially useful when the authors give only macro-correlation coefficients, without information on micro-correlations. Finally, when the data set contains n data points, the size of each matrix coded under COVARIANCE is always n^2. However, only lower triangle elements are coded in EXFOR, since covariance matrices are always symmetric.

If uncertainty values e_{yqk} are defined with the flag F, P, or U, and $c_{yqk\theta}$ is explicitly given for the all sources defined with P, EXFOR users may construct the full covariance matrix according to Eq. (46).

It should be stressed that it is not a requirement for individual experimenters to submit their results to the data centers in EXFOR format. The regional centers normally receive the numerical data files from experimenters, and the compilers then prepare EXFOR entries. There are no specific format requirements for submission of these data other than for them to be easily readable by compilers, preferably electronically. However, an on-line template is available at the IAEA Nuclear Data Section website to facilitate uploading the data directly by experimenters who are served by this particular center [70]. Although compilation of data for the EXFOR library is the explicit responsibility of data center compilers, some experimenters may wish to code their data using EXFOR formats (guided by the *EXFOR Formats Manual* [32]) prior to submission to their respective data center, so as to minimize potential misunderstandings. This, of course, is not discouraged, but experimenters should be aware of the fact that their submissions, regardless of the approach chosen, will be examined by EXFOR compilers at the center to which they submit their data. Furthermore, all EXFOR coded files are screened by checking codes for format and physics errors (*e.g.*, conservation laws, consistent dimensions, *etc.*) for all components of the submitted data, not just uncertainty information, before inclusion in the EXFOR library.

Further enhancements to the EXFOR and C5 formats are currently being discussed by the data compiler community and the software developers, but these have not been adopted by the publication time of this paper [78].

E. Samples from EXFOR

Five coding samples are presented in this section to illustrate the features and general appearance of typical EXFOR entries generated using the new EXFOR formats [32]. In all of the following samples, the line number designations that appear in files downloaded from EXFOR have been eliminated so that a larger font could be used to enhance readability of the essential information. The reader should also be aware that the sample entries appearing below are often incomplete in the interest of saving space.

Sample 1: Ratio of $^{51}V(n,p)^{51}Ti$-to-$^{238}U(n,f)$ cross sections for fast neutrons [50, 51]

This EXFOR coding sample involves $^{51}V(n,p)^{51}Ti$ activation cross sections measured relative to the $^{238}U(n,f)$ standard at 45 incident neutron energies (18 points using the $^{7}Li(p,n)^{7}Be$ neutron source and 27 points using the $^{2}H(d,n)^{3}He$ neutron source). This experiment is discussed in more detail in Example 15 of this paper. These data are compiled with the total uncertainty given, as well as 17 partial uncertainties, of which seven sources are treated as uncorrelated, eight sources are treated as fully correlated, and two sources are treated as partially correlated. For the last two sources, micro-correlation coefficients are expressed as functions of the magnitude of the energy difference between two data points (ΔE), as specified by the experimenters, under the keyword COVARIANCE. Evaluators can therefore construct the full covariance matrix based on this clearly stated assumption.

```
SUBENT        12898002   20120710
BIB                4         32
REACTION   ((23-V-51(N,P)22-TI-51,,SIG)/(92-U-238(N,F),,SIG))
...
FLAG       (1.) 7Li(p,n)7Be source
           (2.) 2H(d,n)3He source
ERR-ANALYS Uncertainty due to orientation of sample for counting
            was treated as negligible.
           (ERR-T)      Total uncertainty
           (ERR-S)      Total random uncertainty in ratio
           (ERR-SYS)    Total systematic uncertainty
           (ERR-1,,,U)  Exposure, waiting and counting times(0.2%)
           (ERR-2,,,U)  0.320-MeV gamma ray yield      (0.3-47.8%)
           (ERR-3,,,U)  Fission yield                   (0.7-1.5%)
           (ERR-4,,,U)  Extrapolation correction            (1-2%)
           (ERR-5,,,U)  Background fission correction        (<3%)
           (ERR-6,,,U)  Background activation           (0.2-1.2%)
           (ERR-7,,,U)  Geometric corrections               (1.5%)
           (ERR-8,,,F)  51Ti decay half life                (0.1%)
           (ERR-9,,,F)  238U content of monitor deposit       (2%)
           (ERR-10,,,F) 51V content of samples              (0.2%)
           (ERR-11,,,F) Uranium deposit thickness correction(0.8%)
           (ERR-12,,,F) Gamma-ray counting efficiency       (2.4%)
           (ERR-13,,,F) 51Ti gamma-ray decay branch factor    (1%)
           (ERR-14,,,P) Neutron source properties             (2%)
           (ERR-15,,,P) Neutron scattering corrections  (1.4-2.1%)
           (ERR-16,,,F) Geometric corrections               (1.5%)
           (ERR-17,,,F) Average neutron energy         (0.5-19.5%)
COVARIANCE ERR-14:
            No correlation between p+7Li and and d+D points.
            Otherwise correlation coefficient=100-10 dE
            (dE: energy difference in MeV)
           ERR-15:
            Correlation coefficient=100-10 dE
            (dE: energy difference in MeV)
...
ENDBIB              32          0
COMMON              10          6
ERR-1      ERR-7      ERR-8      ERR-9      ERR-10     ERR-11
ERR-12     ERR-13     ERR-14     ERR-16
```

```
PER-CENT   PER-CENT   PER-CENT   PER-CENT   PER-CENT   PER-CENT
PER-CENT   PER-CENT   PER-CENT   PER-CENT
 0.2        1.5        0.1        2.         0.2        0.8
 2.4        1.         2.         1.5
ENDCOMMON            6          0
DATA                14         45
EN         EN-RSL-FW  DATA       ERR-S      ERR-SYS    ERR-T
ERR-2      ERR-3      ERR-4      ERR-5      ERR-6      ERR-15
ERR-17     FLAG
MEV        MEV        NO-DIM     PER-CENT   PER-CENT   PER-CENT
PER-CENT   PER-CENT   PER-CENT   PER-CENT   PER-CENT   PER-CENT
PER-CENT   NO-DIM
  2.856     0.095      9.075 E-06 47.9       15.6       50.4
 47.8       0.8        2.0                              2.0
 14.9       1.
  2.957     0.094      1.966 E-05 20.6        9.2       22.6
 20.4       0.8        2.0                              2.1
  7.9       1.
  3.057     0.094      2.575 E-05 15.9       20.1       25.6
 15.7       0.8        2.0                              2.1
 19.5       1.
  3.158     0.092      7.369 E-05  6.7       15.5       16.9
  6.2       0.7        2.0                              2.0
 14.8       1.
  3.258     0.090      1.441 E-04  4.9       11.2       12.2
  4.2       0.7        2.0                              1.9
 10.2       1.
  3.359     0.087      2.354 E-04  3.6        8.1        8.9
  2.5       0.7        2.0                              1.9
  6.6       1.
  3.459     0.087      3.210 E-04  3.4        6.7        7.5
  2.2       0.7        2.0                              1.9
  4.9       1.
  3.560     0.087      4.130 E-04  3.2        7.8        8.4
  1.8       0.7        2.0                              1.9
  6.3       1.
  3.661     0.087      6.140 E-04  3.0        7.2        7.8
  1.5       0.7        2.0                              1.9
  5.5       1.
  3.761     0.082      7.917 E-04  2.9        8.4        8.9
  1.2       0.7        2.0                              1.9
  7.0       1.
  3.861     0.084      1.281 E-03  2.8        5.9        6.5
  1.0       0.7        2.0                              1.9
  3.6       1.
  3.861     0.084      1.282 E-03  2.8        5.9        6.5
  1.0       0.7        2.0                              1.9
  3.6       1.
  3.962     0.081      1.348 E-03  2.8        4.9        5.6
  1.0       0.7        2.0                              1.9
  1.7       1.
  4.063     0.081      1.552 E-03  2.8        5.1        5.8
  0.9       0.7        2.0                              1.9
  2.1       1.
  4.264     0.075      1.967 E-03  2.7        6.7        7.2
  0.8       0.7        2.0                              1.9
  4.9       1.
  4.464     0.076      2.692 E-03  2.7        5.2        5.9
  0.7       0.7        2.0          0.1                 2.0
  2.3       1.
  4.643     0.276      4.105 E-03  2.7        5.1        5.8
  1.1       1.5        1.0          0.5       0.3        2.0
  2.0       2.
  4.664     0.076      4.136 E-03  2.8        4.7        5.5
  0.7       0.9        2.0          0.2                 2.0
  0.5       1.
  4.865     0.076      4.996 E-03  2.8        4.8        5.6
  0.7       1.0        2.0          0.2                 2.0
  1.3       1.
  4.893     0.248      5.315 E-03  2.6        6.2        6.7
  1.0       1.5        1.0          0.5       0.3        1.9
  4.2       2.
  5.139     0.211      6.049 E-03  2.4        5.2        5.7
  0.8       1.3        1.0          0.5       0.3        1.9
  2.4       2.
  5.374     0.194      6.696 E-03  2.4        5.0        5.5
  0.7       1.2        1.0          0.6       0.2        1.9
  1.9       2.
  5.600     0.180      7.571 E-03  2.4        5.6        6.1
  0.7       1.2        1.0          0.7       0.2        1.9
  3.2       2.
  5.819     0.174      8.559 E-03  2.4        5.3        5.8
  0.6       1.2        1.0          0.8       0.3        1.8
  2.6       2.
  5.822     0.167      8.870 E-03  2.5        5.2        5.8
  0.6       1.2        1.0          1.0       0.3        1.8
  2.5       2.
  6.040     0.162      9.600 E-03  2.3        4.9        5.4
  0.5       1.1        1.0          0.8       0.2        1.8
  1.8       2.
  6.254     0.153      1.006 E-02  2.3        4.8        5.3
  0.5       1.0        1.0          0.8       0.3        1.7
  1.5       2.
  6.465     0.146      9.435 E-03  2.3        4.7        5.2
  0.5       1.0        1.0          0.8       0.3        1.7
  1.3       2.
  6.675     0.149      9.317 E-03  2.3        4.7        5.2
  0.4       0.9        1.0          0.9       0.3        1.7
  1.2       2.
  6.881     0.144      1.003 E-02  2.3        4.6        5.1
  0.4       0.8        1.0          1.0       0.3        1.6
  1.1       2.
  6.882     0.141      9.969 E-03  2.3        4.6        5.1
  0.4       0.9        1.0          1.0       0.3        1.6
  1.1       2.
  7.087     0.141      1.003 E-02  2.3        4.6        5.1
  0.4       0.8        1.0          1.0       0.3        1.6
  1.1       2.
  7.290     0.144      1.089 E-02  2.3        4.6        5.1
  0.4       0.8        1.0          1.0       0.3        1.6
  1.0       2.
  7.494     0.144      1.182 E-02  2.6        4.6        5.3
  0.3       0.8        1.0          1.5       0.5        1.6
  0.9       2.
  7.694     0.152      1.189 E-02  2.5        4.6        5.2
  0.4       0.7        1.0          1.5       0.5        1.6
  0.9       2.
  7.893     0.155      1.285 E-02  2.9        4.6        5.4
  0.3       0.7        1.0          2.0       0.6        1.6
  0.9       2.
  7.893     0.156      1.337 E-02  2.9        4.6        5.4
  0.3       0.8        1.0          2.0       0.6        1.5
  0.9       2.
  8.092     0.161      1.331 E-02  2.9        4.6        5.4
  0.3       0.7        1.0          2.0       0.7        1.5
  0.9       2.
  8.290     0.169      1.403 E-02  2.9        4.6        5.4
  0.3       0.7        1.0          2.0       0.8        1.5
  0.9       2.
  8.486     0.173      1.452 E-02  3.3        4.5        5.6
  0.3       0.8        1.0          2.5       0.9        1.5
  0.8       2.
  8.683     0.181      1.571 E-02  3.3        4.6        5.7
  0.3       0.8        1.0          2.5       0.9        1.4
  1.1       2.
  8.684     0.179      1.578 E-02  3.3        4.5        5.6
  0.3       0.7        1.0          2.5       0.9        1.4
  1.0       2.
  8.879     0.184      1.598 E-02  3.4        4.5        5.6
  0.3       0.8        1.0          2.5       1.0        1.4
  1.0       2.
  9.071     0.189      1.778 E-02  3.8        4.6        6.0
  0.3       0.9        1.0          3.0       1.0        1.4
  1.1       2.
  9.267     0.195      1.850 E-02  3.8        4.6        6.0
  0.3       0.9        1.0          3.0       1.2        1.4
  1.1       2.
ENDDATA            141          0
ENDSUBENT          185          0
```

Sample 2: $^{24}Mg(n,p)^{24}Na$ cross sections for fast neutrons [43]

This EXFOR coding sample shows $^{24}Mg(n,p)^{24}Na$ activation cross sections, with the total uncertainty as well as one fully correlated uncertainty component given. The authors calculated total uncertainty (ERR-T) by quadratic summation of the individual components, and their fully correlated portion, 2.11% (labeled "common" uncertainty in the report),

is coded as a fully correlated partial uncertainty under the heading ERR-1. The macro-correlation coefficient between the highest and lowest neutron energy data points can be calculated as $(2.11)^2/(4.26)/(3.15) = 0.332$ as discussed earlier in this report. Note that only uncertainties utilized to generate the covariance matrix are treated as coded information (*i.e.*, computer-readable), and other values (*e.g.*, uncertainty in the monitor cross section) are treated as free text information. The authors treat all off-diagonal elements of the ^{238}U(n,f) monitor cross section relative covariance matrix as negligible, and the uncertainty (1.5%) is treated as an uncorrelated uncertainty. Namely, the covariance matrix of the primary cross-section ratios is obtained by subtracting the 1.5% uncertainty component from the diagonal elements of the relative covariance matrix of the absolute cross section.

```
SUBENT        22976002   20120710
BIB                 3         14
REACTION   (12-MG-24(N,P)11-NA-24,,SIG)
...
ERR-ANALYS (EN-ERR) Error of the energy scale (20 keV)
           (ERR-T,,,P) Total uncertainty
            - 238U(n,f) monitor cross section   (1.5%)
            - 238U deposit mass                 (0.6%)
            - Photon detector calibration       (1.5%)
            - 24Mg isotopic content            (0.04%)
            - 25Mg isotopic content            (0.01%)
            - Half-life                   (0.0012 hr)
           (ERR-1,,,F) Fully correlated uncertainty (2.11%)
COVARIANCE The fully correlated portion of the uncertainty is
            given as ERR-1 and enables, in combination with the
            ERR-T values, the generation of the complete
            covariance matrix
...
ENDBIB              14          0
COMMON               2          3
EN-ERR     ERR-1
KEV        PER-CENT
 20.        2.11
ENDCOMMON            3          0
DATA                 4         26
EN         EN-RSL-FW  DATA       ERR-T
MEV        MEV        MB         PER-CENT
 14.740     0.091      166.3      4.26
 14.147     0.093      197.6      4.17
 13.609     0.096      210.5      4.03
 13.425     0.097      213.5      3.93
 13.187     0.099      182.7      3.79
 12.976     0.099      194.2      3.78
 12.703     0.100      188.7      3.67
 12.446     0.101      196.7      3.63
 12.273     0.101      171.4      3.56
 12.116     0.101      164.5      3.58
 11.839     0.104      168.4      3.43
 11.654     0.106      190.6      3.47
 11.399     0.109      168.7      3.41
 11.223     0.110      164.6      3.42
 10.998     0.112      153.8      3.29
 10.751     0.113      159.0      3.26
 10.551     0.115      152.9      3.22
 10.264     0.120      155.2      3.23
 10.069     0.123      147.4      3.28
  9.837     0.126      132.9      3.29
  9.547     0.131      134.8      3.20
  9.265     0.136      117.0      3.22
  9.111     0.139      127.6      3.24
  8.907     0.144      121.0      3.19
  8.556     0.150      113.6      3.21
  8.334     0.155      118.1      3.15
ENDDATA             28          0
ENDSUBENT           51          0
```

Sample 3: ^{241}Am(n,2n)^{240}Am cross sections for fast neutrons [79, 80]

This EXFOR coding sample pertains to ^{241}Am$(n,2n)^{240}$Am activation cross sections measured at nine incident neutron energies, with the total uncertainty as well as nine partial uncertainties provided. Four sources are treated as uncorrelated, three sources are treated as fully correlated, and two sources are treated as partially correlated. For the last two sources (MONIT-ERR and ERR-5), matrix elements of micro-correlation coefficients are explicitly given under the keyword COVARIANCE. Therefore, EXFOR users can derive the full covariance matrix. In addition to micro-correlation coefficients, the macro-correlation coefficients are also coded explicitly. Derivation of macro-correlation coefficients from micro-correlation data are described in Appendix B of Ref. [80].

```
SUBENT        23114002   20120710
BIB                 3         43
REACTION   (95-AM-241(N,2N)95-AM-240,,SIG)
...
ERR-ANALYS (ERR-T,,,P) Total uncertainty
           (MONIT-ERR,,,P) 27Al(n,a) standard x-section (1.6-5.4%)
           (ERR-1,,,U) Counting of 240Am activity       (1.4-6.3%)
           (ERR-2,,,U) Counting of 24Na activity        (0.7-2.0%)
           (ERR-3,,,F) Intensity of 240Am gamma line    (1.2%)
           (ERR-4,,,U) Number of 27Al in sample         (0.1%)
           (ERR-5,,,P) Number of 241Am in sample        (0.3%)
           (ERR-6,,,F) 24Na/240Am efficiency ratio      (3.0%)
           (ERR-7,,,F) Correction for decay of 240Am    (0.4-0.9%)
           (ERR-8,,,U) Correction for second. neutron   (<1.4%)
COVARIANCE  (XY,9,EN,MEV) incident neutron energy
             8.34 9.15 13.33  16.1 17.16  17.9 19.36 19.95 20.61
            (Z,45,COR:ERR-T,PER-CENT) macro-correlation
             100
              35   100
              37    42   100
              38    43    53   100
              40    45    57    58   100
              41    45    57    59    84   100
              21    24    30    31    39    39   100
              30    34    44    45    58    59    51   100
              20    22    29    30    40    42    39    65   100
            (Z,45,COR:MONIT-ERR,PER-CENT) micro-corr.(standard)
             100
              43  100
               0    0  100
               0    0    6  100
               0    0    9   12  100
               0    0   11   12  100  100
               0    0   11   11   40   40  100
               0    0   11   11   40   40  100  100
               0    0   11   11   40   40  100  100  100
            (Z,45,COR:ERR-5,PER-CENT)   micro-corr. (sample mass)
             100
               0  100
               0  100  100
               0  100  100  100
               0    0    0    0  100
             100    0    0    0    0  100
               0    0    0    0  100    0  100
               0    0    0    0    0    0    0  100
             100    0    0    0    0  100    0    0  100
...
ENDBIB              43          0
COMMON               4          3
ERR-3      ERR-4      ERR-5      ERR-6
PER-CENT   PER-CENT   PER-CENT   PER-CENT
 1.2        0.1        0.3        3.0
ENDCOMMON            3          0
DATA                 6          9
EN         DATA       ERR-T      MONIT-ERR  ERR-1      ERR-2
MEV        MB         PER-CENT   PER-CENT   PER-CENT   PER-CENT
  8.34      96.8       6.5        1.9        5.0        1.0
  9.15     162.9       5.7        1.9        4.0        1.0
 13.33     241.8       4.6        1.6        2.5        1.0
 16.1      152.4       4.6        2.         2.1        1.0
 17.16     116.1       4.4        2.         1.5        1.0
 17.9      105.7       4.4        2.2        1.3        0.7
 19.36      89.5       8.2        3.1        6.3        2.0
 19.95     102.1       5.8        4.1        1.4        1.0
```

```
 20.61      77.9       8.8        5.4        5.7        1.6
ENDDATA             11          0
ENDSUBENT           63          0
```

Sample 4: Ratio of $^{232}Th(n,f)$-to-$^{235}U(n,f)$ cross sections for fast neutrons [81]

This EXFOR coding sample expresses data for ^{232}Th(n,f) cross sections relative to the ^{238}U(n,f) standard at four incident neutron energies, with the total uncertainty and four partial uncertainties provided for each point. Correlation properties are not given by the experimenters, and only the upper boundaries of uncertainties are stated for two sources. Macro-correlation coefficients are coded under the keyword COVARIANCE.

```
SUBENT        22282002   20120723
BIB                3         14
REACTION   ((90-TH-232(N,F),,SIG)/(92-U-235(N,F),,SIG))
...
ERR-ANALYS (ERR-T,,,P)    Total uncertainty
           (MONIT-ERR)    235U(nf) monitor cross section (4%)
           (ERR-1)        Number of U-235 atoms       (1.47%)
           (ERR-2)        Number of Th-232 atoms      (1.64%)
           (ERR-3,,0.887) Fission rate ration       (<0.887%)
           (ERR-4,,0.276) Correction factor         (<0.276%)
COVARIANCE  (XY,4,EN,MEV)
            13.47  14.00 14.46 14.89
            (Z,10,COR:ERR-T,PER-CENT) Macro-corr. coefficients
             100
              87    100
              86     87    100
              87     87     87    100
...
ENDBIB              14          0
COMMON               3          3
MONIT-ERR  ERR-1      ERR-2
PER-CENT   PER-CENT   PER-CENT
 4.         1.47       1.64
ENDCOMMON           3          0
DATA                4          4
EN         EN-ERR     DATA       ERR-T
MEV        MEV        NO-DIM     PER-CENT
 13.47      0.18       0.150      2.41
 14.00      0.06       0.158      2.38
 14.46      0.16       0.166      2.37
 14.89      0.29       0.181      2.38
ENDDATA             6          0
ENDSUBENT          29          0
```

Sample 5: $^{nat}Cd+n$ transmission data in the resonance region [82]

This EXFOR coding sample gives ^{nat}Cd+n high-resolution transmission data in the resonance region, with the total uncertainty as well as five partial "uncertainties" provided, of which one source is treated as uncorrelated, while five sources are treated as fully correlated. Five energy-dependent "uncertainties" are AGS code vectors [83]. It is possible to construct the full covariance matrix considering one source (ERR-2) as uncorrelated and the other five partial uncertainty sources as fully correlated. This can be accomplished utilizing the prescriptions described in Ref. [83]. More detailed explanations about the time-of-flight spectra and their covariances in the EXFOR library are seen in Ref. [84–86].

```
SUBENT        23077002   20110128
BIB                4         11
REACTION   (48-CD-0(N,TOT),,TRN)
...
MISC-COL   (MISC1) Width of time-of-flight bin
           (MISC2) Uncorrelated uncertainty squared
ERR-ANALYS (ERR-T,,,P) Total uncertainty
           (ERR-1,,,F) Normalization uncertainty (0.5%)
           (ERR-2,,,U) Uncorrelated uncertainty
           (ERR-3,,,F) Correlation dead time correction  (sample)
           (ERR-4,,,F) Correlation background correction (sample)
           (ERR-5,,,F) Correlation dead time correct. (open beam)
           (ERR-6,,,F) Correlation background correct.(open beam)
COVARIANCE Compiled in ERR-2 to ERR-6 in the AGS format
...
ENDBIB              11          0
COMMON               1          3
ERR-1
PER-CENT
 0.5
ENDCOMMON            3          0
DATA                12          3
EN          TOF-MIN    TOF-MAX    MISC1      DATA       ERR-T
ERR-2       MISC2      ERR-3      ERR-4      ERR-5      ERR-6
EV          NSEC       NSEC       NSEC       NO-DIM     NO-DIM
NO-DIM      NO-DIM     NO-DIM     NO-DIM     NO-DIM     NO-DIM
 4.79930E+0   873301.2   873429.2      128.0 1.14780E+0 6.65562E-2
 6.65376E-2 4.42725E-3 1.33447E-5-9.12646E-4-1.37145E-5 1.28552E-3
 4.79790E+0   873429.2   873557.2      128.0 9.70250E-1 5.63176E-2
 5.63021E-2 3.16993E-3 1.09942E-5-8.47529E-4-1.16680E-5 1.00913E-3
 4.79649E+0   873557.2   873685.2      128.0 1.04716E+0 6.01241E-2
 6.01078E-2 3.61295E-3 1.19256E-5-8.59307E-4-1.26779E-5 1.10425E-3
ENDDATA             10          0
ENDSUBENT           30          0
```

VI. SUMMARY

This paper discusses the basic principles of uncertainty analysis, as applied to measured nuclear reaction data. Properties of underlying probability distributions that influence the measurement of primary physical parameters are considered. Two approaches to propagating uncertainties from measured variables to derived results are described: deterministic and stochastic (Monte Carlo). Various schemes and approximations that can be used to deal with uncertainty data are mentioned. The application of these techniques to estimating and specifying uncertainties in experimental results is discussed and illustrated by simple examples. The importance of documenting experiments is stressed. Issues associated with documentation are addressed, including a discussion of features of various information dissemination options. Processes whereby experimental data are compiled, evaluated, and ultimately utilized in contemporary nuclear applications, are outlined. Special attention is given to the EXFOR library, since it is a widely used vehicle for documenting experimental nuclear reaction data. Samples are given of covariance data recorded in EXFOR.

ACKNOWLEDGMENTS

The authors are indebted to the following agencies for the support they provided that enabled this project to be undertaken: U.S. Department of Energy (D.L.S) and the Interna-

tional Atomic Energy Agency (N.O.). The authors are grateful to V. Zerkin for suggestions and assistance in dealing with data handling utilities associated with the EXFOR system. R. Haight pointed out the quotation from J. Müller that is included in this paper [38].

[1] P. Obložinský (ed.), Special Edition, "Workshop on Neutron Cross Section Covariances", *Nuclear Data Sheets* **109**, preface page (2008).
[2] "Evaluation of Measurement Data - Guide to the Expression of Uncertainty in Measurement", Report JCGM 100:2008 (Report GUM 1995 with minor corrections), First Edition, JCGM, Paris, September 2008. See: http://www.iso.org/sites/JCGM/GUM-JCGM100.htm.
[3] D.L. Smith, *Probability, Statistics, and Data Uncertainties in Nuclear Science and Technology*, American Nuclear Society, LaGrange Park, IL, USA (1991).
[4] W. Mannhart, "A Small Guide to Generating Covariances of Experimental Data", Report INDC(NDS)-0588, IAEA Nuclear Data Section (2011).
[5] M. Drosg, *Dealing with Uncertainties: A Guide to Error Analysis, 2nd Edition*, Springer, Berlin (2009).
[6] Experimental Nuclear Reaction Data Library (EXFOR), IAEA Nuclear Data Section. See: http://www-nds.iaea.org/exfor/ or, for the NNDC at Brookhaven National Laboratory, the mirror site is http://www.nndc.bnl.gov/exfor/.
[7] M.B. Chadwick et al., "ENDF/B-VII.1 Nuclear Data for Science and Technology: Cross Sections, Covariances, Fission Product Yields and Decay Data", *Nuclear Data Sheets* **112**, 2887 (2011).
[8] A. Koning et al., "The JEFF-3.1 Nuclear Data Library", JEFF Report 21, OECD Nuclear Energy Agency (2006).
[9] K. Shibata et al., "JENDL-4.0: A New Library for Nuclear Science and Engineering", *J. Nucl. Sci. Technol.* **48**, 1 (2011).
[10] N.M. Larson, "Updated Users' Guide for SAMMY: Multilevel R-Matrix Fits to Neutron Data Using Bayes' Equations", Report ORNL/TM-9179/R8, Oak Ridge National Laboratory (2008).
[11] O. Bersillon et al., "International Reactor Dosimetry File 2002 (IRDF-2002)", Technical Report Series No. 452, International Atomic Energy Agency (2006).
[12] *Reactor Dosimetry: Radiation Metrology and Assessment*, J.G. Williams, D.W. Vehar, F.H. Ruddy, and D.M. Gilliam (eds.), American Nuclear Society, LaGrange Park, IL, USA (2001).
[13] B.L. Broadhead et al., "Sensitivity-and-Uncertainty-Based Criticality Safety Validation Techniques", *Nucl. Sci. Eng.* **146**, 340 (2004).
[14] M.L. Williams and B.T. Rearden, "SCALE-6 Sensitivity/Uncertainty Methods and Covariance Data", *Nuclear Data Sheets* **109**, 2796 (2008).
[15] M. Salvatores et al., "Uncertainty and Target Accuracy Assessment for Innovative Systems Using Recent Covariance Data Evaluations", Report NEA/WPEC-26, Working Party for International Evaluation Cooperation, OECD Nuclear Energy Agency (2008).
[16] E. Parzen, *Modern Probability Theory and Its Applications*, John Wiley & Sons, Inc., New York (1960).
[17] "Poisson Calculator: Online Statistical Table", StatTrek.com (2011). See: http://stattrek.com/Tables/poisson.aspx.
[18] D.L. Smith, "An Approach for Dealing with Large Errors", Report ANL/NDM-154, Argonne National Laboratory (2001).
[19] D.L. Smith et al., "Large Errors and Severe Conditions", *Nuclear Instruments and Methods in Physics Research* **A488**, 342 (2002).
[20] C.E. Shannon, "A Mathematical Theory of Communication", *Bell System Technical Journal* **27**, 379 & 623 (1948).
[21] E.T. Jaynes, "Prior Probabilities", *IEEE Trans. on Systems Science and Cybernetics* **4**, 227 (1968).
[22] D.L. Smith, "A Demonstration of the Lognormal Distribution", Report ANL/NDM-156, Argonne National Laboratory (2003).
[23] "Solve My Math", SolveMyMath.com (2011). See: http://www.solvemymath.com/online_math_calculator/statistics/continuous_distributions/gamma/pdf_gamma.php.
[24] D.L. Smith, "A Least-Squares Computational 'Tool Kit'", Report ANL/NDM-128, Argonne National Laboratory (1993).
[25] D.L. Smith, "A Unified Monte Carlo Approach to Fast Neutron Cross Section Data Evaluation", Report ANL/NDM-166, Argonne National Laboratory (2008).
[26] G. Žerovnik et al., "Correlated Random Sampling for Multivariate Normal and Log-normal Distributions", *Nuclear Instruments and Methods in Physics Research* **A690**, 75 (2012).
[27] S.J. Fletcher and M. Zupanski, "A Hybrid Multivariate Normal and Lognormal Distribution for Data Assimilation", *Atmospheric Science Letters* **7**, 43 (2006).
[28] I. Bebu and T. Mathew, "Comparing the Means and Variances of a Bivariate Lognormal Distribution", *Statistics in Medicine* **27**, 2684 (2008).
[29] A.M. Law and W.D. Kelton, *Simulation Modeling and Analysis, 3rd Edition*, McGraw Hill, Boston (2000).
[30] "Normal Distribution", Wikipedia Free Encyclopedia (2011). See: http://en.wikipedia.org/wiki/Normal_distribution.
[31] D.L. Smith and D.G. Naberejnev, "Confidence Intervals for the Lognormal Probability Distribution", *Nuclear Instruments and Methods in Physics Research* **A518**, 754 (2004).
[32] O. Schwerer (ed.), "EXFOR Formats Description for Users (EXFOR Basics)", Report IAEA-NDS-206, IAEA Nuclear Data Section (2008); N. Otuka (ed.), "EXFOR Formats Manual", Report IAEA-NDS-207 Rev.2011/11, IAEA Nuclear Data Section (2011); N. Otuka (ed.), "LEXFOR (EXFOR Compiler's Manual)", Report IAEA-NDS-208 Rev. 2011/01, IAEA Nuclear Data Section (2011).
[33] M. Herman and A. Trkov (eds.), "ENDF-6 Formats Manual: Data Formats and Procedures for the Evaluated Nuclear Data File ENDF/B-VI and ENDF/B-VII", Report BNL-90365-2009, National Nuclear Data Center, Brookhaven National Laboratory (2009).
[34] D.L. Smith, "Evaluated Nuclear Data Covariances: The Journey from ENDF/B-VII.0 to ENDF/B-VII.1", *Nuclear Data Sheets* **112**, 3037 (2011).
[35] F.B. Hildebrand, *Methods of Applied Mathematics*, Prentice-Hall, Inc. (1954).
[36] L. Leal et al., "Resonance Region Covariance Analysis Method and New Covariance Data for ^{232}Th, ^{233}U, ^{235}U, ^{238}U, and ^{239}Pu", *Nuclear Data Sheets* **109**, 2868 (2008).
[37] A.D. Carlson et al., "International Evaluation of Neutron Cross Section Standards", *Nuclear Data Sheets* **110**, 3215 (2009).

[38] J.W. Müller, "Some Second Thoughts on Error Statements", *Nuclear Instruments and Methods* **163**, 241 (1979).

[39] J.J. Wagschal, "How Important are Response-Parameter Correlations?", *Nuclear Data Sheets* **109**, 2880 (2008).

[40] D.W. Muir, "The Contribution of Individual Correlated Parameters to the Uncertainty of Integral Quantities", *Nuclear Instruments and Methods in Physics Research* **A644**, 55 (2011).

[41] "Occam's Razor", Wikipedia Free Encyclopedia (2011). See: http://en.wikipedia.org/wiki/Occam's_razor.

[42] D.L. Smith, "On the Relationship Between Micro and Macro Correlations in Nuclear Measurement Uncertainties", *Nuclear Instruments and Methods in Physics Research* **A257**, 365 (1987).

[43] W. Mannhart and D. Schmidt, "Measurement of Neutron Activation Cross Sections in the Energy Range from 8 MeV to 15 MeV", Report PTB-N-53, PTB Braunschweig, Germany (2007). EXFOR 22976.

[44] M. Fisz, *Probability Theory and Mathematical Statistics, 3rd Edition*, John Wiley & Sons, New York (1963).

[45] D.L. Smith, "Covariance Matrices for Nuclear Cross Sections Derived from Nuclear Model Calculations", Report ANL/NDM-159, Argonne National Laboratory (2004).

[46] N. Metropolis et al., "Equations of State Calculations by Fast Computing Machines", *J. Chem. Phys.* **21**, 1087 (1953).

[47] W.K. Hastings, "Monte Carlo Sampling Methods Using Markov Chains and Their Applications", *Biometrika* **57**, 97 (1970).

[48] R. Capote and D.L. Smith, "An Investigation of the Performance of the Unified Monte Carlo Method of Neutron Cross Section Data Evaluation", *Nuclear Data Sheets* **109**, 2768 (2008).

[49] A.J. Koning and D. Rochman, "Towards Sustainable Nuclear Energy: Putting Nuclear Physics to Work", *Annals of Nuclear Energy* **35**, 2024 (2008).

[50] D.L. Smith et al., "Measurement of the ^{51}V(n,p)^{51}Ti Reaction Cross Section From Threshold to 9.3 MeV by the Activation Method", Report ANL/NDM-85, Argonne National Laboratory (1984). EXFOR 12898.

[51] D.L. Smith et al., "Measured Activation Cross Sections Below 10 MeV for the ^{51}V(n,p)^{51}Ti and ^{51}V(n,α)^{48}Sc Reactions", *Annals of Nuclear Energy* **11**, 623 (1984). EXFOR 12898.

[52] M.W. Herman, National Nuclear Data Center, Brookhaven National Laboratory, private communication, March 2012.

[53] M.L. Zerkle, Bettis Atomic Power Laboratory, private communication, March 2012.

[54] R.E. MacFarlane and A.C. Kahler, "Methods for Processing ENDF/B-VII with NJOY", *Nuclear Data Sheets* **111**, 2739 (2010).

[55] D. Wiarda et al., "Recent Advances with the AMPX Covariance Processing Capabilities in PUFF-IV", *Nuclear Data Sheets* **109**, 2791 (2008).

[56] Y. Danon and R.C. Block, "Minimizing the Statistical Error of Resonance Parameters and Cross-Sections Derived from Transmission Measurements", *Nuclear Instruments and Methods in Physics Research* **A485**, 585 (2002).

[57] R.W. Peelle, "Peelle's Pertinent Puzzle", privately circulated informal memorandum, Oak Ridge National Laboratory, dated 13 October 1987.

[58] D. Deleanu et al., "The Gamma Efficiency of the GAINS Spectrometer", *Nuclear Instruments and Methods in Physics Research* **A624**, 130 (2010).

[59] L.C. Mihailescu et al., "A New HPGe Setup at Gelina for Measurement of Gamma-ray Production Cross-sections from Inelastic Neutron Scattering", *Nuclear Instruments and Methods in Physics Research* **A531**, 375 (2004).

[60] D.L. Smith, "Concept and Significance of a Dimensionless Sensitivity Matrix in Applications of Generalized Least-Squares Analysis", *Nuclear Instruments and Methods in Physics Research* **A339**, 626 (1994).

[61] A. Fessler and D.L. Smith. "Parameter Sensitivities in Nuclear Reaction Cross-Section Calculations", *Annals of Nuclear Energy* **29**, 363 (2002).

[62] A.J.M. Plompen et al., "Neutron Activation Cross-Section Measurements From Threshold to 20 MeV for the Validation of Nuclear Models and Their Parameters", Report NEA/WPEC-19, Working Party for International Evaluation Cooperation, OECD Nuclear Energy Agency (2005).

[63] J.W. Meadows and D.L. Smith, "Gamma-Ray Detector Calibration Methods Utilized in the Argonne FNG Group Activation Cross Section Measurement Program", Report ANL/NDM-60, Argonne National Laboratory (1984).

[64] Geant4 Collaboration, "Introduction to Geant4", documentation for version Geant4 9.5.0 (2011). See: http://geant4.web.cern.ch/geant4/UserDocumentation/ UsersGuides/IntroductionToGeant4/fo/BookIntroToGeant4.pdf.

[65] W.P. Poenitz, "Data Interpretation, Objective Evaluation Procedures and Mathematical Techniques for the Evaluation of Energy-Dependent Ratio, Shape and Cross Section Data", *Proceedings of the Conference on Nuclear Data Evaluation Methods and Procedures*, Brookhaven National Laboratory, Upton, USA, September 22-25, 1980, B.A. Magurno and S. Pearlstein (eds.), Report BNL-NCS-51363, Brookhaven National Laboratory, Vol. 1, 249 (1981).

[66] S.F. Mughaghab, *Atlas of Neutron Resonances: Resonance Parameters and Thermal Cross Sections Z= 1-100 5th Edition*, Elsevier, Amsterdam (2006).

[67] N. Otuka and S. Dunaeva (eds.), "International Network of Nuclear Reaction Data Centres", Report INDC(NDS)-0401 Rev.5, IAEA Nuclear Data Section (2010); N. Otuka and S. Dunaeva (eds.), "NRDC Protocol", Report IAEA-NDS-215 Rev.2010/11, IAEA Nuclear Data Section (2010). See: http://www-nds.iaea.org/nrdc/.

[68] D.E. Cullen and A. Trkov, "Program X4TOC4 (Version 2001-3): Translation of Experimental Data from the EXFOR Format to a Computation Format", Report IAEA-NDS-80 (Rev. 1), IAEA Nuclear Data Section (2001).

[69] V. Zerkin and A. Trkov, "Development of IAEA Nuclear Reaction Databases and Services", *Proceedings of the International Conference on Nuclear Data for Science and Technology (Vols. 1 and 2)*, Nice, France, April 22-27, 2007, O. Bersillon, F. Gunsing, E. Bauge, R. Jacqmin, and S. Leray (eds.), CEA, EDP Sciences, ISBN: 978-2-7598-0090-2, 769 (2008).

[70] V. Zerkin, IAEA Nuclear Data Section, Vienna, Austria, private communication (2011).

[71] X-5 Monte Carlo Team, "MCNP - A General Monte Carlo N-Particle Transport Code, Version 5, Volume I: Overview and Theory", Report LA-UR-03-1987, Los Alamos National Laboratory, (2003). Revised 2/1/2008. See also: http://mcnp-green.lanl.gov/.

[72] Radiation Safety Information Computational Center (RSICC), Oak Ridge National Laboratory, Oak Ridge, TN, USA. See: http://www-rsicc.ornl.gov/.

[73] G.J. Wyant and A.M. Daly (eds.), "Report on the Fifth IAEA Consultants' Meeting of Nuclear Reaction Data Centers", Report INDC(NDS)-125, IAEA Nuclear Data Section (1981).

[74] M. Lammer and O. Schwerer (eds.), "Report on the Sixth IAEA Consultants' Meeting of Nuclear Reaction Data Centers", Report INDC(NDS)-141, IAEA Nuclear Data Section (1982).

[75] H.D. Lemmel (ed.), "Report on the Seventh IAEA Consultants' Meeting of Nuclear Reaction Data Centres", Report INDC(NDS)-154, IAEA Nuclear Data Section (1984).

[76] G. Winkler et al., "Measurement of Cross Sections for the ^{63}Cu(n,α)^{60}Co Reaction from Threshold to 10 MeV", *Nucl. Sci. Eng.* **76**, 30 (1980). EXFOR 10921.

[77] P. Guenther et al., "Total, Scattering, and Gamma Production Cross Sections for Few-MeV Neutrons on Elemental Copper", *Nucl. Phys.* **A488**, 280 (1986). EXFOR 12869.

[78] This comment is based on discussions the authors of the present paper held with V. Zerkin (IAEA Nuclear Data Section) in December 2011. He developed the software used to implement earlier EXFOR formats and the C4 format, as well as the recent revisions to the EXFOR formats and the newly developed C5 format (2012). Future upgrades to the EXFOR formats (beyond 2012) may incorporate additional options that allow the specification of constant full-correlation micro-correlation parameters (other than 1), and expanded user-provided micro-correlation capabilities.

[79] C. Sage et al., "High Resolution Measurements of the ^{241}Am(n,2n) Reaction Cross Section", *Phys. Rev.* **C81**, 4604 (2010). EXFOR 23114.

[80] C. Sage, Doctoral Thesis (in French), Université de Strasbourg, Strasbourg, France (2009). EXFOR 23114.

[81] F. Manabe et al., "Measurements of Neutron Induced Fission Cross Section Ratios of ^{232}Th, ^{233}U, ^{234}U, ^{236}U, ^{238}U, ^{237}Np, ^{242}Pu and ^{243}Am Relative to ^{235}U Around 14 MeV", *The Technology Reports of the Tohoku University* **52**, 97 (1988). EXFOR 22282.

[82] S. Kopecky et al., "The Total Cross Section and Resonance Parameters for the 0.178 eV Resonance of ^{113}Cd", *Nuclear Instruments and Methods in Physics Research* **B267**, 2345 (2009). EXFOR 23077.

[83] C. Bastian, "General Procedures and Computational Methods for Generating Covariance Matrices", *Proceedings of the International Symposium on Nuclear Data Evaluation Methodology*, Brookhaven National Laboratory, October 12-16, 1992, C.L. Dunford (ed.), World Scientific, Singapore, 642 (1992); C. Bastian et al., "AGS, A Computer Code for Uncertainty Propagation in Time-of-Flight Cross Section Data", *Proceedings of the Conference on Advances in Nuclear Analysis and Simulation (PHYSOR-2006)*, Vancouver, Canada, September 10-14, 2006, C013.

[84] N. Otuka et al., "Database for Time-of-Flight Spectra Including Covariances", *J. Korean Physical Society* **59**, 1314 (2011).

[85] P. Schillebeeckx et al., "Determination of Resonance Parameters and Their Covariances from Neutron Induced Reaction Cross Section Data", *Nuclear Data Sheets*, this issue.

[86] B. Becker et al., "Data Reduction and Uncertainty Propagation of Time-of-Flight Spectra with AGS", in preparation.