IAEA
International Atomic Energy Agency

INDC International Nuclear Data Committee

A Small Guide to Generating Covariances of Experimental Data

Wolf Mannhart

Physikalisch-Technische Bundesanstalt, Braunschweig, Germany

September 2013

A Small Guide to Generating Covariances
of Experimental Data

Wolf Mannhart

Physikalisch-Technische Bundesanstalt, Braunschweig, Germany

September 2013

Contents

Preface

This is a new hard copy and electronic version of the old report PTB-FMRB-84 (June 1981). The original report has been written with a conventional typewriter and the manuscript is meanwhile lost. The proposal to produce a new edition of the report came from Dr. Naohiko Otsuka (IAEA, Vienna) and his support in setting the manuscript with LaTeX was essential and is highly appreciated. The text is mainly the same as in the original. A few mistakes have been corrected and minor modifications were applied to improve the readability and the understanding. In addition, the famous textbook of Don Smith [9], published in 1991, was added to the list of references.

Braunschweig, May 2011

From numerous discussions about data covariances I have learned that there is an general readiness to apply this method. However, as often happens with ideas beyond the familiar, here going further than the familiar methods for the treatment of uncertainties, difficulties arise particularly in understanding how to do it correctly. It is my wish that the present report may contribute in solving some of the basic difficulties. This report is written in a certain sense as a cookery book and it aims to explain the essential rules by means of relatively simple numerical examples which should make it relatively easy to apply the method to more complex and realistic problems. Most of the examples given are based on real experiments. However, these examples are also in some sense "synthetic" since they have been simplified to avoid any unnecessary complexity which may obstruct easy understanding.

This report has mainly been written for experimenters, and so a complete and mathematically rigorous description of the theory was not attempted. In the first chapter the most important rules of statistics are reviewed. The second chapter contains some essential concepts about experimental uncertainties, but it is written rather tentatively. The third chapter demonstrates by way of examples the procedure of generating the covariance matrices of experimental data. The last chapter, concerned with data evaluation procedures, is of more importance to evaluators than to experimenters. Nevertheless it might also be of some interest for experimenters as it explains the importance of quoting an experiment with a complete covariance matrix. Up to this last chapter consistent matrix notation in the examples has been avoided, even when some of them given previously can be most conveniently written in matrix notation, as shown in the Appendix 1. The reason is that all those readers not familiar enough with matrix calculation should not be discouraged too early in the game. In the last chapter a modest transition from non-matrix to matrix notation is made. The well-known least squares equations written in conventional notation compared with the matrix notation should ease the understanding of this chapter.

This work has been strongly influenced by the pioneering work of Francis Perey (Oak

A   Introduction

The continuously increasing amount of nuclear data requires more and more consistent and objective methods of data combination. Two facts have prevented this in the past: firstly, evaluators did not always use those evaluation tools which gave sufficiently objective results. A typical example of this is that the same data set given to different evaluators resulted in different evaluated data. Secondly, and this point influenced the first of course, experimenters failed to express the uncertainties of their results in a sufficiently complete way.

Recently methods have been developed and applied [1, 2] which circumvent these difficulties. The application of these methods permits a consistent treatment of data uncertainties in complex data systems; on the other hand it also defines obligatory requirements for the documentation of experimental results. In contrast to the procedure in the past, apart from the data uncertainties, the correlations of the data too have to be documented. Neglect of the latter is mainly the reason for the above mentioned inconsistencies. Up to now, only few examples exist in the literature [1, 2, 3] which demonstrate complete uncertainty listing of experimental results. The need for sufficiently complete data documentation and the subsequent data combination procedure are the two main topics of this paper.

B Some Definitions

If the "true value" of a physical quantity were known then it would be absurd to carry out an experiment to confirm this value. This automatically implies that the value we know cannot be a true one. I. e., a "true value" is and remains a hypothetical quantity. A "true value" can never be quantified, all experimental physics can do is to determine quantities which approach the "true values" as closely as possible. Each experimental result must be regarded as an estimate of the "true value" and must therefore be accompanied by an uncertainty quotation which expresses the accuracy of the experimental result. The definition of uncertainty is well established in the formalism of mathematical statistics which we shall use in this chapter.

B.1 Single Random Variable

The expectation value of a single random variable, x, associated with a probability density $\phi(x)$, is defined by:

$$\langle x \rangle = \mu = \int_{-\infty}^{+\infty} x\, \phi(x)\, \mathrm{d}x \tag{B-1}$$

The symbol, $\langle \rangle$, is used here to indicate expected values. Eq. (B-1) gives the so-called first moment. Together with the second moment

$$\langle x^2 \rangle = \int_{-\infty}^{+\infty} x^2 \phi(x)\, \mathrm{d}x \tag{B-2}$$

we are able to define the variance of x. By introducing an auxiliary quantity, $\mathrm{d}x$, which describes the deviation between x and the expectation value

$$\mathrm{d}x = x - \mu \tag{B-3}$$

the variance is given as the expected value of the square of Eq. (B-3), as

$$\mathrm{Var}(x) = \langle \mathrm{d}x\mathrm{d}x \rangle = \langle x^2 \rangle - \langle x \rangle^2 \tag{B-4}$$

The square root of Eq. (B-4) is the well-known standard deviation. The first and second moments of a distribution are sufficient to describe the expectation value of the quantity x itself as well as the variance or standard deviation respectively. It should be mentioned that here and in the following no particular assumption about the shape of the probability density is used. Specifically it is not postulated that x is normally distributed (Gaussian shape). The given equations are valid if the integrals exist. i.e. if the values of the integrals are finite.

B.2   Multivariate Probability Density Function

If there are several variables which are statistically not independent then a joint multivariate density must exist. The formalism will be demonstrated for a bivariate case, i.e. with the variables x and y which are not independent. Their joint probability density is given by $\phi(x,y)$. The first and higher moments are defined in an analogous way as above:

$$\langle x \rangle = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x \, \phi(x,y) \, \mathrm{d}x\mathrm{d}y \tag{B-5}$$

and

$$\langle y \rangle = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y \, \phi(x,y) \, \mathrm{d}x\mathrm{d}y \tag{B-6}$$

Besides the variances of x and y

$$\mathrm{Var}(x) = \langle \mathrm{d}x\mathrm{d}x \rangle \, \text{and} \, \mathrm{Var}(y) = \langle \mathrm{d}y\mathrm{d}y \rangle \tag{B-7}$$

there now exists an additional term describing the covariance of x and y:

$$\mathrm{Cov}(x,y) = \langle \mathrm{d}x\mathrm{d}y \rangle = \langle xy \rangle - \langle x \rangle \langle y \rangle \tag{B-8}$$

The variances as well as the covariances are terms of quadratic dimension. In a similar way as one reduces variances to standard deviations, it is convenient to define the correlation coefficient:

$$\mathrm{Corr}(x,y) = \frac{\mathrm{Cov}(x,y)}{\sqrt{\mathrm{Var}(x)} \, \sqrt{\mathrm{Var}(y)}} = \frac{\langle \mathrm{d}x\mathrm{d}y \rangle}{\sqrt{\langle \mathrm{d}x\mathrm{d}x \rangle} \, \sqrt{\langle \mathrm{d}y\mathrm{d}y \rangle}} \tag{B-9}$$

The denominator of Eq. (B-9) is the product of the corresponding standard deviations. The numerical value of the correlation coefficient must be between -1 and +1. The covariance term of Eq. (B-8) must vanish if x and y are independent. In that case the joint probability function becomes

$$\phi(x,y) = \phi_1(x)\phi_2(y) \tag{B-10}$$

Then one obtains

$$\langle xy \rangle = \int_{-\infty}^{+\infty} x \, \phi_1(x) \, \mathrm{d}x \int_{-\infty}^{+\infty} y \, \phi_2(y) \, \mathrm{d}y = \langle x \rangle \langle y \rangle \tag{B-11}$$

Inserting Eq. (B-11) in Eq. (B-8) shows that the covariance becomes zero and vanishes. For completeness it should be mentioned that the converse is not true; a zero covariance does not automatically imply statistical independence, except for a multivariate normal distribution. In all cases where the variables are not independent the variance alone is an incomplete information. The information is complete when variances as well as covariances are regarded.

B.3   Function of Variables

The result of most experiments is not a single variable but a function of several variables:

$$f(x_i) = f(x_1, x_2, \ldots, x_n) \tag{B-12}$$

In general, the expectation value of such a function

$$\langle f \rangle = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(x_i)\, \phi(x_i)\, dx_1 \cdots dx_n \tag{B-13}$$

can only be determined if the multivariate probability density function, $\phi(x_i)$, is known. A special class of functions is given by linear functions for which the mentioned restriction is not valid. Moreover, in the case of non-linear functions this restriction can be circumvented if the function can be approximated by a linear one. The distinction between both types of functions will be shown in the following.

B.3.1   Linear Function

For a linear function

$$f = \sum_i a_i x_i \tag{B-14}$$

the expectation value is given by the linear sum of the expectation values of the different variables:

$$\langle f \rangle = \sum_i a_i \langle x_i \rangle \tag{B-15}$$

The variance of the function is then

$$\text{Var}(f) = \langle df df \rangle = \sum_i \sum_j a_i a_j \langle dx_i dx_j \rangle \tag{B-16}$$

Eq. (B-16) can also be written as

$$
\begin{aligned}
\langle df df \rangle &= \sum_i a_i^2 \langle dx_i dx_i \rangle + \sum_i \sum_j a_i a_j \langle dx_i dx_j \rangle \\
&= \sum_i a_i^2\, \text{Var}(x_i) + \sum_i \sum_{\substack{j \\ (i \neq j)}} a_i a_j\, \text{Cov}(x_i, x_j)
\end{aligned} \tag{B-17}
$$

Due to the symmetry of

$$\langle dx_i dx_j \rangle = \langle dx_j dx_i \rangle \tag{B-18}$$

we can also write

$$\text{Var}(f) = \sum_i a_i^2\, \text{Var}(x_i) + 2 \sum_i \sum_{\substack{j \\ (i < j)}} a_i a_j\, \text{Cov}(x_i, x_j) \tag{B-19}$$

The first term of Eq. (B-17, B-19) represents the well-known "error propagation" rule for independent linear variables. The second term takes into account that the variables are interdependent.

B.3.2 Non-linear Function

In the case of a non-linear function, the interconnection of the expectation value of the function, $\langle f \rangle$, and of the expectation values of the various variables, $\langle x_i \rangle$, normally requires the knowledge of the joint probability density function. Under certain conditions there is a way if the function can be expanded around the expectation values μ_i of the respective variables:

$$\langle x_i \rangle = \mu_i \tag{B-20}$$

In these cases a Taylor expansion (up to the 1st order) is used as an approximation and the function is represented as:

$$f(x_i) \simeq f(\mu_i) + \sum_i \left.\frac{\partial f}{\partial x_i}\right|_\mu (x_i - \mu_i) \tag{B-21}$$

By using the abbreviation

$$a_i = \left.\frac{\partial f}{\partial x_i}\right|_\mu \tag{B-22}$$

one obtains

$$f(x_i) \simeq f(\mu_i) + \sum_i a_i(x_i - \mu_i) \tag{B-23}$$

Under these conditions we can evaluate the expectation value and obtain

$$\langle f(x_i) \rangle \simeq f(\mu_i) = \overline{f} \tag{B-24}$$

With

$$dx_i = x_i - \mu_i \tag{B-25}$$

and

$$df = f - \overline{f} \simeq \sum_i a_i dx_i \tag{B-26}$$

the variance of the function is given by

$$\mathrm{Var}(f) = \langle df df \rangle \simeq \sum_i a_i^2 \langle dx_i dx_i \rangle + \sum_i \sum_{\substack{j \\ (i \neq j)}} a_i a_j \langle dx_i dx_j \rangle \tag{B-27}$$

This means the problem has been linearized with coefficients, a_i, given by Eq. (B-22). Often a non-linear function is represented by a series of products (or quotients):

$$f(x_i) = \prod_i x_i \quad \text{or} \quad \left(f(x_i) = 1/\prod_i x_i \right) \tag{B-28}$$

Then the coefficients of Eq. (B-22) become

$$a_i = (\pm)\frac{\bar{f}}{\mu_i} \tag{B-29}$$

In these cases it is conveniently to introduce relative variances (indicated by the symbol δ instead of d)

$$\text{Rel.Var}(f) = \langle \delta f \delta f \rangle = \left\langle \frac{\mathrm{d}f}{f} \frac{\mathrm{d}f}{f} \right\rangle \tag{B-30}$$

and Eq. (B-27) can be transformed into

$$\langle \delta f \delta f \rangle = \sum_i \langle \delta x_i \delta x_i \rangle + \sum_i \sum_{\substack{j \\ (i \neq j)}} b_i b_j \langle \delta x_i \delta x_j \rangle \tag{B-31}$$

with

$$b_i = \begin{cases} +1 & \text{for a product} \\ -1 & \text{for a quotient} \end{cases} \tag{B-32}$$

It must be kept in mind that Eq. (B-27) and (B-31) are approximations and are only adequate if the Taylor expansion of the 1st order of Eq. (B-21) is valid. This means, independent of the form of the probability density function, the criterion of $\langle \mathrm{d}x_i \mathrm{d}x_i \rangle \ll \mu_i^2$ must always be fulfilled.

In addition the following point should be mentioned. The first terms of Eq. (B-17) as well as of Eq. (B-27) always have a positive sign (quadratic "error propagation"), the sign of the second terms depends on the sign of a_i or b_i, respectively. It is positive only if the corresponding coefficients in the formula show the same sign, otherwise the sign becomes negative.

9

C Estimate of the Parameters of Probability Density Functions

If there is some quantity we want to know extremely well, we often measure it several times, using different methods, with different instruments, for example. Then one can imagine a set of all the possible outcomes of an experiment. This set, called sample space or population, is associated with a probability density. Two parameters of this probability density function are of main interest: the expected value of the quantity itself (see Eq. B-1) and the variance (see Eq. B-4) which defines an interval which embraces the quantity. To obtain estimates for both these parameters one normally draws a sample of limited size, n, from the total population. In the case of a pure random variable, x, the *sample* mean

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \tag{C-1}$$

and the *sample* variance

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 \tag{C-2}$$

can be regarded as best estimates of the population parameters μ and $\text{Var}(x)$, respectively.

Often one tries to distinguish between "statistical" and "systematic" uncertainties. The only real difference between both is that the former can readily be estimated from the laws of probability, Eq. (C-1) and (C-2), while the latter usually arise from the lack of information to exactly correct results for the inability of experimental equipment to measure the desired quantity directly.

Particularly in nuclear experiments, where the statistical components of the uncertainty ("counting" uncertainties) can often be kept small, the second type of "systematic" uncertainty dominates. However, these uncertainties should not be regarded as a different type compared with statistical uncertainties. In an analogous way as for random variables, the probability density represents a real ensemble of possible outcomes; in the case of non-random variables their probability density function can be regarded as hypothetical ensemble of results that could have occurred based on equivalent (but different) experimental efforts. It is from such an ensemble that one can imagine the result to be drawn. I.e., experimental results are considered as drawn from hypothetical populations of potential experiments. Under this aspect, the probability density function represents information from real or hypothetical populations, as well as any other available knowledge of the experimenter.

For the "systematic" uncertainties, the experimental result is usually taken as the estimate of the first momentum of the probability function. The estimate of the corresponding variance (or second central moment) depends strongly on the available supplementary knowledge. Here one can distinguish between two cases. The first case has been discussed

by S. Wagner [4] as follows: If the upper and lower limit of a quantity can be estimated and no additional information exists, then the "principle of equal ignorance" (Bayes and Laplace) requires the assumption of a constant probability density within these limits ("rectangular probability density function"). However, Wagner also states: "If there is any additional information, this should be applied to derive a more specific kind of probability density function". And this brings us to the second case: Very often an experimenter can state it is more likely that a quantity lies within a certain interval than outside it. This is additional information, irrespective of its quantification, and so the "principle of equal ignorance" should not be applied. Here some knowledge of the shape of a probability density function is implied. In practice the principal difference between both methods is of minor importance. No really meaningful reason can be given as to why it should be easier to estimate limits than variances. Moreover, if the problem of the determination of overall uncertainties (including both "statistical'" and "systematic" components") is of some complexity, there is no justification for concentrating the attention on a certain source of uncertainties and ignoring the others. The procedure using variances is of some advantage as these can simply be added independent of the shape of probability density function associated with each component. I.e., the combination of variances does not require specific knowledge of the shape of probability density functions.

The situation becomes rather more complex when one wants to determine confidence intervals for overall uncertainties. That requires a knowledge of the probability density function. In this context it is very useful to remember the Central Limit Theorem which states that, almost regardless of the shape of the component probability density functions, with increasing number of components the combined distribution asymptotically approximates to the Normal distribution. This theorem has, however, some limitations and is not valid if one of the components is dominant. The Central Limit Theorem is restricted to the case of independent components. One may suspect, without proof, that with certain restrictions a modification of this theorem can also be applied to correlated components.

An estimate of confidence intervals can only be made for a specific probability density function. However, in cases where the shape of the density function is not so well-known, a lower limit can be estimated. The so-called "Chebychev Inequality" states that for a random variable x with expectation value μ and variance s^2, independent of the shape of the probability density function, the following is true:

$$P(|x - \mu| \geq \delta) \leq \frac{s^2}{\delta^2} \qquad \text{(C-3)}$$

P is the probability of x being outside the interval $(-\delta, +\delta)$. Eq. (C-3) is valid as long as the variance exists and is finite. From Eq. (C-3) it can easily be estimated that for the worst possible type of probability density function the confidence level corresponding to three standard deviations is 88.9%. For a normal distribution it would be 99.7%. If no detailed information about the probability density function is available, the confidence level must

have a value lying between these two values. With the exception of a few relatively simple cases, practically no information on the detailed shapes of probability density functions is available. In such cases one ought to be led by a "principle of economics" which I would formulate as: one should never try to extract more information from a system than this latter contains.

D   Covariances of Experimental Data

D.1   Linear Function of Data

The first example is taken from length measuring technique. An experimenter, a "Mr.A.",
has to determine two markings on a length scale of distances from a fixed zero point, x_1
= 35 mm and x_2 = 60 mm (see Fig. 1). For this experiment he uses three gauge blocks of

Figure 1:
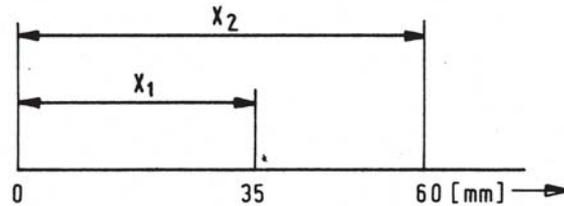
defined length and with well known variances:

Gauge block		Length [mm]	Std.Dev. [μm]	Variance [μm^2]	
l_1	=	50	0.05	Var(l_1)=0.0025	(D-1)
l_2	=	15	0.03	Var(l_2)=0.0009	
l_3	=	10	0.02	Var(l_3)=0.0004	

The first marking is obtained by using the gauge block l_1 and subtracting the length l_2,
while the second marking is reached by adding to l_1 the length of l_3, I.e.:

$$x_1 = l_1 - l_2$$
$$x_2 = l_1 + l_3$$

(D-2)

By using the uncertainty propagation rules:

$$\text{Var}(x_1) = \text{Var}(l_1) + \text{Var}(l_2) = 0.0034$$
$$\text{Var}(x_2) = \text{Var}(l_1) + \text{Var}(l_3) = 0.0029$$

(D-3)

"Mr.A." states his final result as:

$$
\begin{array}{ll}
 & \text{Std.Dev.} \\
x_1 = 35 \text{ mm} & 0.058 \ \mu\text{m} \\
x_2 = 60 \text{ mm} & 0.054 \ \mu\text{m}
\end{array}
$$

(D-4)

15

Figure 2:
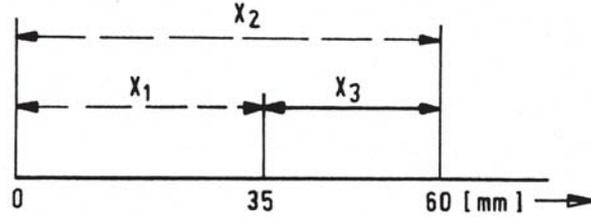
What he has be done appears correct. We will however come back to this point later.

Some time after this experiment one is interested in an additional quantity, x_3, which defines the distance between both markings set by "Mr.A." (see Fig. 2). Two colleagues of "Mr.A.", "Mr.B." and "Mr.C.", are engaged to establishing this new quantity. They proceed using different methods.

"Mr.B." uses the same set of gauge blocks as "Mr.A." and can establish that the distance x_3 can be reproduced by addition of the length of the gauge blocks l_2 and l_3. He obtains

$$x_3 = l_2 + l_3 = 25 \text{ mm}$$

and

$$\text{Var}(x_3) = \text{Var}(l_2) + \text{Var}(l_3) = 0.0013 \ \mu m^2$$

(D-5)

At the same time, "Mr.C." uses the information given by "Mr.A." in Eq. (D-2) and (D-3) and calculates:

$$x_3 = x_2 - x_1 = 25 \text{ mm}$$

and

$$\text{Var}(x_3) = \text{Var}(x_2) + \text{Var}(x_1) = 0.0063 \ \mu m^2$$

(D-6)

At first glance the situation looks frustrating. The two experimenters "Mr.B." and "Mr.C." have both done correct work and obtained the different results of Eq. (D-5) and Eq. (D-6), respectively. What happened? To understand that we have to go back to the first experiment of "Mr.A.". Looking at Eq. (D-2) we recognize that "Mr.A." used in both the determination of x_1 as well as x_2 the length of the gauge block l_1. Further, looking at Eq. (D-3) and (D-4), we find that the information of l_1 being common to both x_1 and x_2 was not communicated in the final results, The whole information is only given if we replace Eq. (D-3) by:

$$\begin{aligned}
\text{Var}(x_1) &= \text{Var}(l_1) + \text{Var}(l_2) \\
\text{Var}(x_2) &= \text{Var}(l_1) + \text{Var}(l_3) \\
\text{Cov}(x_1, x_2) &= \text{Var}(l_1)
\end{aligned}$$

(D-7)

16

The complete covariance matrix (in μm^2) is then:

$$\begin{pmatrix} 0.0034 & 0.0025 \\ 0.0025 & 0.0029 \end{pmatrix} \tag{D-8}$$

With this information the final result of Eq. (D-4) should be written:

$$\begin{array}{ccc} & \text{Std.Dev.} & \text{Correlation Matrix} \\ x_1 = 35 \text{ mm} & 0.058 \ \mu m & \begin{pmatrix} 1.00 & 0.80 \\ x_2 = 60 \text{ mm} & 0.054 \ \mu m & 0.80 & 1.00 \end{pmatrix} \end{array} \tag{D-9}$$

From Eq. (D-9) we recognize that x_1 and x_2 are not independent since their correlation coefficient is 0.80. Due to the symmetry of the correlation matrix it is in most cases sufficient to state only the upper or lower half of this matrix.

Using this complete information the procedure of "Mr.C." shown in Eq. (D-6) changes and "Mr.C." would obtain:

$$x_3 = x_2 - x_1 = 25 \text{ mm}$$

and

$$\begin{aligned} \text{Var}(x_3) &= \text{Var}(x_2) + \text{Var}(x_1) - 2 \ \text{Cov}(x_1, x_2) \\ &= 0.0029 + 0.0034 - 2 \times 0.0025 = 0.0013 \ \mu m^2 \end{aligned} \tag{D-10}$$

What we can learn from the above example is: the complete information about a measuring process is contained in the covariance matrix. To perform correct uncertainty propagation one needs this complete information. With incomplete information (i.e. without a covariance matrix) the risk of incorrect uncertainty propagation arises, as it occurred in Mr.C.'s case in Eq. (D-6). Particularly with increasing complexity of experiments, the statement of a covariance matrix becomes more and more essential to allow further processing of the data. The information required for the covariance matrix is at every experimenter's disposal when estimating the data uncertainties. However, in the past it was usually suppressed and not communicated to others.

D.2 Non-linear Function of Data

D.2.1 Direct (Absolute) Measurement of Neutron Cross Section

Very often the quantities which we can most conveniently measure are not the quantities of basic interest. This is also true for the case of neutron cross section measurement. One of the methods for cross section determination is the activation of defined materials in well-known neutron fields and the subsequent deduction of a cross section σ from the induced radioactivity. The relationship is:

$$\sigma_i = \frac{A_i}{\varepsilon_i} \frac{1}{N_i} \frac{1}{\phi_i} \tag{D-11}$$

with i being an index indicating a specific material and type of reaction. A_i is the measured counting rate, N_i is the number of contributing atoms and ϕ_i is the neutron flux density at the position of the probe during the irradiation process. We assume a mono-energetic neutron field of the energy, E_n, so that $\sigma_i = \sigma_i(E_n)$. The efficiency of the detector, ε_i, used in measuring the radioactivity depends on the energy of the gamma rays produced in a specific neutron reaction. I.e. $\varepsilon_i = \varepsilon(E_{\gamma i})$. The functional form of ε_i is discussed in more detail in section E.4.2.

In our example we assume that, for example, three different reaction cross sections were measured in the same mono-energetic neutron field. i.e., $\phi_1 = \phi_2 = \phi_3$. A complete list of the various components contributing to the uncertainty of Eq. (D-11) of all three experiments is as follows:

Table 1:

Rel.Std.Dev.(in %)			
$i =$	1	2	3
A_i	0.5	1.0	0.3
ε_i	1.6^a	2.2^a	1.3^a
ϕ_i	2.0^b	2.0^b	2.0^b

[a]Correlation coefficients: Corr($\varepsilon_1, \varepsilon_2$)=0.8

Corr($\varepsilon_1, \varepsilon_3$)=0.5

Corr($\varepsilon_2, \varepsilon_3$)=0.9

[b]Fully correlated (same neutron field !)

The uncertainty contribution from N_i is usually very small and can therefore be neglected. In most cases the various components contributing to the uncertainty of a single experiment (fixed i) can be defined in such a way that there are no correlations between components belonging to the same measurement. We see from Table 1 that no vertical correlations exist only horizontal ones, namely between components belonging to different experiments.

Regarding the quoted correlations in more detail, we observe no correlation between the measured counting rates. This indicates that the uncertainty due to counting statistics dominates all other uncertainty components which may contribute to the total uncertainty of the radioactivity measurement. The particular correlations between the various efficiencies are attributed to the relationship describing the efficiency. The neutron flux density

18

is the same for all three experiments, and so must be taken into account by a correlation coefficient of 1.

Using the definitions of Eq. (B-25) and (B-30) we can formally write the relative deviations from the expected values of Eq. (D-11) as

$$\delta\sigma_i = \delta A_i - \delta\varepsilon_i - \delta\phi_i \qquad \text{(D-12)}$$

The estimates of the variances and covariances are given by forming the expected values $\langle \delta\sigma_i\, \delta\sigma_i \rangle$ and $\langle \delta\sigma_i\, \delta\sigma_j \rangle$ respectively, *with due regard* to the signs of Eq. (D-12).

In our example we have no vertical correlations which means that terms of the form of $\langle \delta A_i\, \delta\varepsilon_i \rangle$ and $\langle \delta A_i\, \delta\phi_i \rangle$ and $\langle \delta\varepsilon_i\, \delta\phi_i \rangle$ do not exist.

With $\langle \delta A_i\, \delta A_j \rangle = 0$, i.e. no correlation between the measured count rates we can immediately write:

$$
\begin{aligned}
\text{Rel.Var.}(\sigma_i) = & \quad \langle \delta\sigma_i\, \delta\sigma_i \rangle = \quad \langle \delta A_i\, \delta A_i \rangle + \quad \langle \delta\varepsilon_i\, \delta\varepsilon_i \rangle + \langle \delta\phi_i\, \delta\phi_i \rangle \\
\text{Rel.Cov.}(\sigma_i, \sigma_j) = & \quad \langle \delta\sigma_i\, \delta\sigma_j \rangle = \qquad\qquad\qquad \langle \delta\varepsilon_i\, \delta\varepsilon_j \rangle + \langle \delta\phi_i\, \delta\phi_j \rangle
\end{aligned}
\qquad \text{(D-13)}
$$

The relative covariance matrix (in $\%^2$) is then

$$
\begin{matrix}
i=1 \\
i=2 \\
i=3
\end{matrix}
\begin{pmatrix}
0.5^2 + 1.6^2 + 2.0^2 & & \\
0.8 \times 1.6 \times 2.2 + 2.0^2 & 1.0^2 + 2.2^2 + 2.0^2 & \\
0.5 \times 1.6 \times 1.3 + 2.0^2 & 0.9 \times 2.2 \times 1.3 + 2.0^2 & 0.3^2 + 1.3^2 + 2.0^2
\end{pmatrix}
$$

$$\text{(D-14)}$$

$$
= \begin{pmatrix}
6.81 & & \\
6.82 & 9.84 & \\
5.04 & 6.57 & 5.78
\end{pmatrix}
$$

The final result can also be stated as:

Measured quantity	Rel.Std.Dev. %	Correlation Matrix		
σ_1	2.61	1.00		
σ_2	3.14	0.83	1.00	
σ_3	2.40	0.80	0.87	1.00

$$\text{(D-15)}$$

In forming the ratio σ_2/σ_1, for example, we recognize the importance of quoting a covariance matrix. Without taking the covariances into account we obtain an uncertainty of 4.08% for the given ratio, whereas incorporation of covariances results in an uncertainty of 1.75% for the same ratio. This difference is easily understandable. In forming such a ratio, those terms of Eq. (D-11) which are common to both measurements tend to compensate one another. The covariance or correlation matrix describes the degree to which terms are common to different measurements.

19

From Table 1 we can learn what sort of information is necessary for the generation of a covariance matrix. First a comprehensive list of all the uncertainty components and their values is required. Secondly, the degree of correlation between the various components due to the measuring procedure has to be stated. The first step has usually been satisfied in most advanced measurements in the past. The same is not true for the second step. Since the second step requires a full understanding of the measuring process, it is advantageous when correlations are quoted directly by the experimenter, rather than being estimated by an evaluator who may find it difficult to confirm all the experimental details later.

Often one needs to know how to state a correlation taking any value between zero and unity. For the case of a functional relationship between the uncertainty components, this relationship automatically defines the exact degree of correlation. In other cases it may be somewhat more complex. Let us assume the following covariance matrix for two experimental uncertainty components

$$\begin{pmatrix} 4.0 & \\ 3.0 & 9.0 \end{pmatrix}$$

indicating a correlation coefficient of 50%. In most practical applications one should normally be able to partition the uncertainty components in a sum or product of subcomponents, which are either fully correlated or uncorrelated. The above matrix, for example, may be the sum of two matrices of subcomponents:

$$\begin{pmatrix} 3.0 & \\ 3.0 & 3.0 \end{pmatrix} + \begin{pmatrix} 1.0 & \\ 0.0 & 6.0 \end{pmatrix}$$

where the first subcomponent is 100% correlated and the second subcomponent shows zero correlation. It is therefore strongly recommended to proceed not with ready combined uncertainty components, but to divide such components into a sufficient number of subcomponents whose correlation can be easily defined.

D.2.2 Relative Measurement of Neutron Cross Sections

Many, if not most, cross section measurements are performed as relative ratio measurements. The reason for doing this is the difficulty in determining an exact absolute neutron flux density. This problem can be evaded by measuring a cross section relative to another which is believed more or less well-known. In the literature such ratio measurements are often normalized and it is not always clearly stated that the measurement was in principle a relative one. Because the normalization of a ratio measurement is in a certain sense arbitrary, it is important that the quoted covariance matrix which describes the experimental procedure is the matrix of the actually performed ratio measurement without normalization. Only such a matrix can be propagated correctly. As will be shown in section E, the

normalization is part of the final evaluation procedure and should therefore not be performed at too early a stage. Of course, such ratio measurements can be normalized by the experimenter, yet it is strongly recommended that the stated covariance matrix does not include this normalization as it is information not based on experimental facts.

The example presented here is based on an experiment of Cf-252 spectrum averaged cross sections published by Kobayashi and Kimura [5]. The uncertainty analysis has been re-done based on information from the original experimental protocol [6]. Our example contains a small subset of the complete analysis: the ratio measurement of the ^{27}Al(n,p) cross section relative to the ^{27}Al(n,α) cross section and a ratio measurement of the ^{24}Mg(n,p) cross section relative to the ^{115}In(n,n') cross section. The symbol σ_i stands for the spectrum averaged cross section.

The results were:

$$
\begin{array}{llll}
\sigma_2 \left[^{27}\text{Al(n,p)}\right] & / & \sigma_1 \left[^{27}\text{Al(n,}\alpha\text{)}\right] & = 4.797 \\
\sigma_4 \left[^{24}\text{Mg(n,p)}\right] & / & \sigma_3 \left[^{115}\text{In(n,n')}\right] & = 0.009651
\end{array}
\tag{D-16}
$$

Using the definition of Eq. (D-11) one can write such ratios as with

$$
\sigma_i/\sigma_j = P_i/P_j
\tag{D-17}
$$

with

$$
P_i = \frac{A_i}{\varepsilon_i} \frac{1}{N_i} \prod_l k_i^l
\tag{D-18}
$$

The product of the k_i^l-terms stands for corrective terms necessary to reduce the measured quantities to the simple form of Eq. (D-11). The complete list of all uncertainty contributions and their correlations is given in Table 2. The principle of such measurements is the parallel irradiation of two different foils in the same neutron field and subsequent counting of the induced radioactivities. The experimentally determined quantities P_i of Eq. (D-18) are terms of the form $\sigma_i \phi_i$. The unknown neutron flux density is eliminated by forming the ratio.

In Table 2 the uncertainties of the A_i are regarded as being independent. I.e., uncertainties due to effects such as the determination of photo-peak area or background subtraction, which would establish correlations, were neglected compared with the counting statistics. The efficiency correlations (a) come from the interpolation procedure used to determine the efficiency (more details on that are given in subsection E.4.2). The correction of the geometrical factor is the same for all measurements, and so the corresponding uncertainties show full correlation (b). The uncertainties of the back-scattering correction are different but, as the source of the back-scattering (room walls) is the same, the quantities must be assumed to be fully correlated (e). The measurements no. 1 and no. 2 are based on the mass determination of a common aluminium foil, therefore the correlation is 100% (d).

21

Table 2: List of the Uncertainty Components (from Ref. [6])

Uncertainties (in %) due to	Symbol	Run - 1		Run - 2	
		^{27}Al(n,α)	^{27}Al(n,p)	^{115}In(n,n')	^{24}Mg(n,p)
	No.	1	2	3	4
Counting statistics	A_i	1.6	2.0	1.0	3.2
Efficiency	ε_i	1.1^a	1.4^a	2.2^a	1.1^a
Geometrical factor	k_i^G	2.0^b	2.0^b	2.0^b	2.0^b
Half life	k_i^T	0.2^c	-	-	0.2^c
Mass determination	N_i	0.1^d	0.1^d	0.1	0.2
Back scattering	k_i^B	0.7^e	1.0^e	1.0^e	0.7^e
Irradiation and colling time	k_i^H	0.1	1.8	0.1	0.1
Gamma ray attenuation	k_i^S	0.5^f	0.5	1.0	0.5^f
Gamma ray intensity	k_i^J	0.1^g	1.0	1.0	0.1^g
Others	k_i^R	1.0	1.0	1.0	1.0

a	$\text{Corr}(\varepsilon_1, \varepsilon_4) = 1.00$
	$\text{Corr}(\varepsilon_1, \varepsilon_3) = \text{Corr}(\varepsilon_3, \varepsilon_4) = 0.80$
	$\text{Corr}(\varepsilon_1, \varepsilon_2) = \text{Corr}(\varepsilon_2, \varepsilon_4) = 0.94$
	$\text{Corr}(\varepsilon_2, \varepsilon_3) = 0.95$
b, e	Fully correlated
c, f, g	Fully correlated (same product nucleus)
d	Fully correlated (same foil)

The product nucleus, ^{24}Na, is the same for both reactions ^{27}Al(n,α) and ^{24}Mg(n,p), therefore all uncertainty components depending on the common radioactive decay, half life and mass attenuation must be fully correlated (c, f-g). Finally one should comment that uncertainty sources should normally be specified better than under the title "Others", as given in Table 2.

Formally we can write

$$\delta P_i = \delta A_i - \delta \varepsilon_i - \delta N_i + \sum_l \delta k_i^l \tag{D-19}$$

As Table 2 shows no vertical correlations, the relative covariance of the measured P_i is

$$\langle \delta P_i\, \delta P_j \rangle = \langle \delta A_i\, \delta A_j \rangle + \langle \delta \varepsilon_i\, \delta \varepsilon_j \rangle + \langle \delta N_i\, \delta N_j \rangle + \sum_l \langle \delta k_i^l\, \delta k_j^l \rangle \tag{D-20}$$

i.e. we have no mixed terms between the components of Eq. (D-19).

What we want to determine is the covariance matrix of the ratios

$$\begin{aligned} R_{12} &= P_2/P_1 \\ R_{34} &= P_4/P_3 \end{aligned} \tag{D-21}$$

With

$$\begin{aligned} \delta R_{12} &= \delta P_2 - \delta P_1 \\ \delta R_{34} &= \delta P_4 - \delta P_3 \end{aligned} \tag{D-22}$$

we obtain

$$\begin{aligned} \langle \delta R_{12}\, \delta R_{12} \rangle &= \langle \delta P_2\, \delta P_2 \rangle + \langle \delta P_1\, \delta P_1 \rangle - \ 2\langle \delta P_1\, \delta P_2 \rangle \\ \langle \delta R_{34}\, \delta R_{34} \rangle &= \langle \delta P_4\, \delta P_4 \rangle + \langle \delta P_3\, \delta P_3 \rangle - \ 2\langle \delta P_3\, \delta P_4 \rangle \\ \langle \delta R_{12}\, \delta R_{34} \rangle &= \langle \delta P_1\, \delta P_3 \rangle + \langle \delta P_2\, \delta P_4 \rangle - \ \langle \delta P_1\, \delta P_4 \rangle \ -\langle \delta P_2\, \delta P_3 \rangle \end{aligned} \tag{D-23}$$

We recognize from Eq. (D-23) that we need the complete relative covariance matrix of the P_i to be able to deduce the relative covariance matrix of the ratios. Since Table 2 contains the complete information about the uncertainty components and correlations of the P_i, the generation of this relative covariance matrix is straightforward, as shown in Table 3. With the result of Table 3 we can determine the relative covariance matrix (in %^{2}) of the measured ratios shown in Eq. (D-23) as

$$\begin{pmatrix} \langle \delta R_{12}\, \delta R_{12} \rangle & \langle \delta R_{12}\, \delta R_{34} \rangle \\ \langle \delta R_{34}\, \delta R_{12} \rangle & \langle \delta R_{34}\, \delta R_{34} \rangle \end{pmatrix} = \begin{pmatrix} 13.72 & -1.14 \\ -1.14 & 17.87 \end{pmatrix} \tag{D-24}$$

The final result is then given by:

Ratio		Rel.Std. Dev. %	Correlation Matrix($\times$ 100)
$\sigma_2 \left[^{27}\text{Al(n,p)}\right] \ / \ \sigma_1 \left[^{27}\text{Al(n,}\alpha\text{)}\right]$	$= 4.797$	3.70	$\begin{pmatrix} 100 & \\ -7 & 100 \end{pmatrix}$
$\sigma_4 \left[^{24}\text{Mg(n,p)}\right] \ / \ \sigma_3 \left[^{115}\text{In(n,n')}\right]$	$= 0.009651$	4.23	

Table 3: Combination of the Various Uncertainty Components

$$
\begin{array}{c}
\langle \delta A_i \, \delta A_j \rangle \\
\begin{array}{c}
i = 1 \\ 2 \\ 3 \\ 4
\end{array}
\left(
\begin{array}{cccc}
1.6^2 & & & \\
0 & 2.0^2 & & \\
0 & 0 & 1.0^2 & \\
0 & 0 & 0 & 3.2^2
\end{array}
\right)
\\
j = \quad 1 \qquad 2 \qquad 3 \qquad 4
\end{array}
\; + \;
\begin{array}{c}
\langle \delta \varepsilon_i \, \delta \varepsilon_j \rangle \\
\left(
\begin{array}{cccc}
1.1^2 & & & \\
0.94 \times 1.1 \times 1.4 & 1.4^2 & & \\
0.80 \times 1.1 \times 2.2 & 0.95 \times 1.4 \times 2.2 & 2.2^2 & \\
1.00 \times 1.1 \times 1.1 & 0.94 \times 1.4 \times 1.1 & 0.80 \times 2.2 \times 1.1 & 1.1^2
\end{array}
\right)
\end{array}
$$

$$
+ \;
\begin{array}{c}
\langle \delta k_i^G \, \delta k_j^G \rangle \\
\left(
\begin{array}{cccc}
2.0^2 & & & \\
2.0^2 & 2.0^2 & & \\
2.0^2 & 2.0^2 & 2.0^2 & \\
2.0^2 & 2.0^2 & 2.0^2 & 2.0^2
\end{array}
\right)
\end{array}
\; + \;
\begin{array}{c}
\langle \delta k_i^T \, \delta k_j^T \rangle \\
\left(
\begin{array}{cccc}
0.2^2 & & & \\
0 & 0 & & \\
0 & 0 & 0 & \\
0.2^2 & 0 & 0 & 0.2^2
\end{array}
\right)
\end{array}
\; + \;
\begin{array}{c}
\langle \delta N_i \, \delta N_j \rangle \\
\left(
\begin{array}{cccc}
0.1^2 & & & \\
0.1^2 & 0.1^2 & & \\
0 & 0 & 0.1^2 & \\
0 & 0 & 0 & 0.2^2
\end{array}
\right)
\end{array}
$$

$$
+ \;
\begin{array}{c}
\langle \delta k_i^B \, \delta k_j^B \rangle \\
\left(
\begin{array}{cccc}
0.7^2 & & & \\
0.7 \times 1.0 & 1.0^2 & & \\
0.7 \times 1.0 & 1.0 \times 1.0 & 1.0^2 & \\
0.7 \times 0.7 & 1.0 \times 0.7 & 1.0 \times 0.7 & 0.7^2
\end{array}
\right)
\end{array}
\; + \;
\begin{array}{c}
\langle \delta k_i^H \, \delta k_j^H \rangle \\
\left(
\begin{array}{cccc}
0.1^2 & & & \\
0 & 1.8^2 & & \\
0 & 0 & 0.1^2 & \\
0 & 0 & 0 & 0.1^2
\end{array}
\right)
\end{array}
$$

$$
+ \;
\begin{array}{c}
\langle \delta k_i^S \, \delta k_j^S \rangle \\
\left(
\begin{array}{cccc}
0.5^2 & & & \\
0 & 0.5^2 & & \\
0 & 0 & 1.0^2 & \\
0.5^2 & 0 & 0 & 0.5^2
\end{array}
\right)
\end{array}
\; + \;
\begin{array}{c}
\langle \delta k_i^J \, \delta k_j^J \rangle \\
\left(
\begin{array}{cccc}
0.1^2 & & & \\
0 & 1.0^2 & & \\
0 & 0 & 1.0^2 & \\
0.1^2 & 0 & 0 & 0.1^2
\end{array}
\right)
\end{array}
\; + \;
\begin{array}{c}
\langle \delta k_i^R \, \delta k_j^R \rangle \\
\left(
\begin{array}{cccc}
1.0^2 & & & \\
0 & 1.0^2 & & \\
0 & 0 & 1.0^2 & \\
0 & 0 & 0 & 1.0^2
\end{array}
\right)
\end{array}
$$

$$
= \;
\begin{array}{c}
\langle \delta P_i \, \delta P_j \rangle \\
\left(
\begin{array}{cccc}
9.58 & & & \\
6.16 & 16.46 & & \\
6.64 & 7.93 & 13.86 & \\
6.00 & 6.15 & 6.64 & 17.29
\end{array}
\right)
\end{array}
$$

E   Data Evaluation Methods

Data evaluation methods must involve objective procedures of data combination. This requirement can be fulfilled by the least squares method, which is not confined to simple curve-fitting problems. In particular it is not true that the least squares procedure depends on a normal distribution of the uncertainty components. This erroneous assumption, often believed, is probably due to the fact that, if normally distributed uncertainties are given, it can be shown that the principle of Maximum Likelihood leads to the same solution as does the principle of least squares. The least square procedure can be generalized. Based on off-diagonal weights, it allows the inclusion of available prior information. The consistent application of these principles results in a kind of mechanization of the evaluation process. Notably it is easy to add a new set of data to a prior evaluation without repeating the whole evaluation. Starting from the for two centuries well-known least squares equations the generalization of the procedure is shown and demonstrated by numerical examples.

E.1   Conventional Least Squares

Starting with *n independent* experimental observations (data points)

$$y_i \quad (i = 1, \ldots, n) \tag{E-1}$$

and attributing to each data point a corresponding variance

$$\sigma_i^2 \quad (i = 1, \ldots, n) \tag{E-2}$$

which can also be interpreted as a weight for each data point of

$$w_i = 1/\sigma_i^2 \tag{E-3}$$

one wishes to determine a set of *m* parameters

$$p_j \quad (j = 1, \ldots, m) \quad m < n \tag{E-4}$$

which gives a most consistent description of the data. Very often the fitting function described by the parameter set, p_j, is a nonlinear one and one needs some prior information about the set of parameters given by parameter estimates of

$$p^0 = (p_j^0) \quad (j = 1, \ldots, m) \tag{E-5}$$

Also, if the fit is a linear one it may be of advantage to proceed with so-called reduced parameters

$$x = p_1 - p_1^0, \quad y = p_2 - p_2^0, \ldots, z = p_m - p_m^0 \tag{E-6}$$

25

which are imperative in the nonlinear case. The fitting function is then given by a Taylor expansion of first order around the parameter estimates (see Eq. (B-21) to (B-23)) as:

$$f_i = f_i(p) = f_i(p^0) + a_i x + b_i y + \cdots + c_i z \tag{E-7}$$

$f_i(p^0) = f_i^0$ is the fitting function calculated at the point of the parameter estimates. The coefficients of the above expansion are given by the partial derivatives taken at the point of the parameter estimates:

$$a_i = \frac{\partial f_i}{\partial p_1}(p^0), \quad b_i = \frac{\partial f_i}{\partial p_2}(p^0), \ldots, c_i = \frac{\partial f_i}{\partial p_m}(p^0), \tag{E-8}$$

With the definition of the deviations between the data points and the fitting function

$$v_i = y_i - f_i = r_i - (a_i x + b_i y + \cdots + c_i z) \tag{E-9}$$

and

$$r_i = y_i - f_i^0$$

one can express the minimum condition as

$$\chi^2 = Q = \sum_i w_i (v_i)^2 = \text{Min} \tag{E-10}$$

This condition is equivalent to

$$\frac{1}{2}\frac{\partial Q}{\partial x} = \frac{1}{2}\frac{\partial Q}{\partial y} = \cdots = \frac{1}{2}\frac{\partial Q}{\partial z} = 0 \tag{E-11}$$

and is, in the case of the reduced parameter x, given by:

$$\sum_i w_i v_i \frac{\partial v_i}{\partial x} = \sum_i w_i [r_i - (a_i x + b_i y + \cdots + c_i z)] a_i = 0$$

With the notation

$$\sum_i a_i b_i = [ab] \tag{E-12}$$

one obtains from Eq. (E-11) the complete set of the so-called normal equations:

$$
\begin{aligned}
[waa]x + [wab]y + \cdots + [wac]z &= [war] \\
[wab]x + [wbb]y + \cdots + [wbc]z &= [wbr] \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
[wac]x + [wbc]y + \cdots + [wcc]z &= [wcr]
\end{aligned}
\tag{E-13}
$$

The solution of the linear system of normal equations results in the final values of the reduced parameters:

$$
\begin{aligned}
x &= p_1 - p_1^0 \\
y &= p_2 - p_2^0 \\
&\cdots\cdots\cdots \\
z &= p_m - p_m^0
\end{aligned}
\tag{E-14}
$$

(In the present notation the reduced parameters, x, y and z, are only used as formal quantities and should not be mistaken for the data y_i of Eq. (E-1), for example).

The uncertainties of the final result are given automatically by the least squares procedure as shown below. As in the conventional least squares procedure the *parameter estimates* are regarded as quantities free of any uncertainty, the uncertainties of Eq. (E-14) are equivalent to the uncertainties of the parameters p_j. Also in the case where the data points y_i are regarded as independent, i.e. not correlated, the same is no longer true for the parameters of the fit. The least squares procedure automatically generates correlations between the resulting final parameters. And therefore it is incorrect, as occasionally seen in the literature, to quote only the parameter uncertainties without their correlation. In summary: the least squares procedure when applied to independent data points also gives parameter results which are interdependent.

E.2 Matrix Notation

The above equations can be most conveniently written in a matrix notation which allows easier understanding of the variances and covariances of the result. Before generalizing the least squares method, we will show the same equations as above written in matrix notation. This comparison is also summarized in Table 4. The data points can be expressed by a column vector

$$
D^0 = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}
\tag{E-15}
$$

and their covariances by a matrix V

$$
V = \begin{pmatrix}
\sigma_1^2 & 0 & \cdots & 0 \\
0 & \sigma_2^2 & \cdots & 0 \\
\vdots & \vdots & & \vdots \\
0 & 0 & \cdots & \sigma_n^2
\end{pmatrix}
\tag{E-16}
$$

which in the conventional least squares method only contains elements arranged on the diagonal. In our notation capital letters denote vectors or matrices. In a similar way the

Table 4: Conventional Least Squares

Quantities	Normal Notation	$\longrightarrow$	Matrix Notation
Data	$y_i \ (i=1,\ldots,n)$ "independent"	$D^0 = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$	D^0
Covariances	σ_i^2 "variances only"	$V = \begin{pmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{pmatrix}$	V
Weights	$w_i = 1/\sigma_i^2$		V^{-1}
Parameter	$p_j \ (j=1,\ldots,m)$	$P' = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_m \end{pmatrix}$	P'
Parameter Estimates	p_j^0	$P = \begin{pmatrix} p_1^0 \\ p_2^0 \\ \vdots \\ p_m^0 \end{pmatrix}$	P
Covariances	"not defined"		-
Reduced Parameter	$x = p_1 - p_1^0, \quad y = p_2 - p_2^0$ $\cdots \qquad\qquad z = p_m - p_m^0$		$(P'-P)$
Fit Function	$\begin{aligned} f_i(p) &= f_i(p^0) + a_i x + b_i y \\ &\quad + \cdots + c_i z \end{aligned}$ $\begin{aligned} a_i &= \frac{\partial f_i}{\partial p_1}(p^0), \quad b_i = \frac{\partial f_i}{\partial p_2}(p^0), \\ &\cdots, c_i = \frac{\partial f_i}{\partial p_m}(p^0) \end{aligned}$	$D' = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix}, \ D = \begin{pmatrix} f_1^0 \\ f_2^0 \\ \vdots \\ f_n^0 \end{pmatrix}$ $\underline{G} = \begin{pmatrix} a_1 & b_1 & \cdots & c_1 \\ a_2 & b_2 & \cdots & c_2 \\ \vdots & & & \\ a_n & b_n & \cdots & c_n \end{pmatrix}$	$D' = D + G\,(P'-P)$
Deviations	$v_i = y_i - f_i = r_i - (a_i x + b_i y + \cdots + c_i z)$ $r_i = y_i - f_i^0$		$(D^0 - D')$ $(D^0 - D)$
Minimum Condition	$\chi^2 = Q = \sum_i w_i (v_i)^2 = \min$ $\frac{1}{2}\frac{\partial Q}{\partial x} = \frac{1}{2}\frac{\partial Q}{\partial y} = \cdots \frac{1}{2}\frac{\partial Q}{\partial z} = 0$		$Q = (D^0 - D')^T V^{-1}(D^0 - D')$ $\frac{\partial Q}{\partial (P'-P)} = 0$
Normal Equations	$[waa]x + [wab]y + \cdots + [wac]z = [war]$ $[wab]x + [wbb]y + \cdots + [wbc]z = [wbr]$ $\cdots$ $[wac]x + [wbc]y + \cdots + [wcc]z = [wcr]$	$[ab] = \sum_i a_i b_i$	$G^T V^{-1} G (P'-P)$ $= G^T V^{-1}(D^0 - D)$
Solution	$x = p_1 - p_1^0, \quad y = p_2 - p_2^0$ $\cdots, \qquad\qquad z = p_m - p_m^0$		$(P'-P) = A^{-1} G^T V^{-1}(D^0 - D)$
Covariance Matrix	"correlated solution parameters" (see right hand side)		$\begin{aligned} M' &= A^{-1} \\ A &= G^T V^{-1} G \end{aligned}$

28

parameters and their estimates are given by the vectors:

$$P' = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_m \end{pmatrix} \quad \text{and} \quad P = \begin{pmatrix} p_1^0 \\ p_2^0 \\ \vdots \\ p_m^0 \end{pmatrix} \qquad \text{(E-17)}$$

If the fit function belonging to each data point and the function calculated with the parameter estimates are denoted by

$$D' = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix} \quad \text{and} \quad D = \begin{pmatrix} f_1^0 \\ f_2^0 \\ \vdots \\ f_n^0 \end{pmatrix}$$

and the derivatives of Eq. (E-8) are written as

$$G = \begin{pmatrix} a_1 & b_1 & \cdots & c_1 \\ \vdots & \vdots & & \vdots \\ a_n & b_n & \cdots & c_n \end{pmatrix} \qquad \text{(E-18)}$$

then the fit function can be written as:

$$D' = D + G(P' - P) \qquad \text{(E-19)}$$

The minimum condition is then

$$\chi^2 = Q = (D^0 - D')^T V^{-1}(D^0 - D')$$

with

$$\frac{\partial Q}{\partial (P' - P)} = 0 \qquad \text{(E-20)}$$

where V^{-1} indicates the inverse of the matrix V and the superscript, (T), denotes a transpose. The normal equations derived from Eq. (E-20) are:

$$G^T V^{-1} G(P' - P) = G^T V^{-1}(D^0 - D) \qquad \text{(E-21)}$$

By using the abbreviation

$$A = G^T V^{-1} G$$

the solution of the least squares procedure is given by:

$$(P' - P) = A^{-1} G^T V^{-1}(D^0 - D) \qquad \text{(E-22)}$$

29

and the covariance matrix of the result is

$$M' = A^{-1} \qquad \text{(E-23)}$$

namely the inverse of the matrix given in the normal equations (Eq. E-21) on the left hand side. Here we have implicitly assumed that the elements of the matrix V (Eq. E-16) are realistic variances and not just relative weights. Therefore it is not justifiable to multiply the matrix elements of Eq. (E-23) by the factor $\chi^2/(n-m)$, as is often done when one regards the elements of V only as elements with an arbitrary scale factor.

E.3   Generalized Least Squares (see also Table 5)

It is obvious from Eq. (E-16) that the least squares method can also be applied to data with off-diagonal elements in the matrix V. This means it is not imperative for the data to be independent. One also recognizes that the vector P of the parameter estimates in the conventional least squares procedure is given without a covariance matrix quoting the uncertainties of these estimated parameters. In the data evaluation process one often has prior information (from an earlier evaluation, for example) which can be used for the parameter estimates. This prior information is connected with a covariance matrix, M, which specifies the uncertainties of the parameter vector P. The Bayesian Rule (see Ref. [7], for example) states how to combine a prior information given by the vector P and the corresponding covariance matrix M with the data vector D^0 and its covariance matrix V to obtain a result P' which is most consistent with P and the data D^0.

Applying Bayes' Rule we have to rewrite the minimum condition as:

$$\chi^2 = (D^0 - D')^T V^{-1} (D^0 - D') + (P - P')^T M^{-1} (P - P') \qquad \text{(E-24)}$$

Eq. (E-24) in the given form is only valid as long as there are no correlations between the data P and D^0. The normal equations now become:

$$\left[M^{-1} + G^T V^{-1} G \right] (P' - P) = G^T V^{-1} (D^0 - D) \qquad \text{(E-25)}$$

The matrix G, sometimes called sensitivity matrix, defines how a change in the data is transformed into a corresponding change of the parameters. The elements of this matrix, G_{ij}, are given by the partial derivatives

$$G_{ij} = \frac{\partial D_i}{\partial P_j} \qquad \text{(E-26)}$$

After applying some mathematical transformations on Eq. (E-25), (see ref. [8], for example), we can state the final result as:

$$\begin{aligned} (P' - P) &= M\, G^T \left[N + V \right]^{-1} (D^0 - D) \\ M - M' &= M\, G^T \left[N + V \right]^{-1} G\, M \\ \text{with} \quad N &= G\, M\, G^T \end{aligned} \qquad \text{(E-27)}$$

Table 5: Least Squares Method

Quantities	Conventional	Generalized	Dimension
Data	D^0	D^0	n
Covariances	V (diagonal)	V (off-diagonal)	$n \times n$
Weights	V^{-1}		
Parameter	P'	P'	m
Estimates	P	P	m
Covariances	(not defined)	M (off-diagonal)	$m \times m$
Fit Function	$D' = D + G(P' - P)$	$D' = D + G(P' - P)$ $G_{ij} = \frac{\partial D_i}{\partial P_j}$	n $n \times m$
Minimum Condition	$Q = (D^0 - D')^T V^{-1}(D^0 - D')$ $\frac{\partial Q}{\partial (P' - P)} = 0$	$Q = (D^0 - D')^T V^{-1}(D^0 - D')$ $+ (P - P')^T M^{-1}(P - P')$	1
Normal Equations	$G^T V^{-1} G(P' - P)$ $= G^T V^{-1}(D^0 - D)$	$\left[M^{-1} + G^T V^{-1} G \right](P' - P)$ $= G^T V^{-1}(D^0 - D)$	
Solution	$A = G^T V^{-1} G$ $(P' - P) =$ $A^{-1} G^T V^{-1}(D^0 - D)$	$N = GMG^T$ $(P' - P) =$ $MG^T \left[N + V \right]^{-1}(D^0 - D)$	m
Covariance Matrix	$M' = A^{-1}$	$M - M' = MG^T \left[N + V \right]^{-1} GM$	$m \times m$

E.4   Numerical Examples

E.4.1   Evaluation of Spectrum-Averaged Cross Sections

In this example we assume that two independent experiments exist (performed by different experimenters) of cross sections averaged over a Cf-252 neutron spectrum. Our aim is to combine both these data sets to obtain the most consistent "best" set of data. The first experiment comprises the data of two reactions:

Table 6: Experiment No. 1

Neutron Reaction	Spectrum-Averaged Cross Section	Rel.Std.Dev. %	Correlation Matrix ($\times$ 100)	
^{235}U(n,f)	1215 mb	1.79	100	
^{239}Pu(n,f)	1790 mb	2.26	59	100

In the second experiment the cross section of the reaction ^{235}U(n,f) is measured twice (using a slightly different method in the second case), and, instead of the absolute cross section of ^{239}Pu(n,f) a ratio measurement relative to ^{235}U(n,f) was performed. A first

Table 7: Experiment No. 2

Neutron Reaction (Ratio)	Spectrum-Averaged Cross Section (Ratio)	Rel.Std.Dev. %	Correlation Matrix ($\times$ 100)		
^{235}U(n,f)	1205 mb	2.25	100		
^{235}U(n,f)	1203 mb	3.02	80	100	
^{239}Pu(n,f)/^{235}U(n,f)	1.500	1.33	-19	-5	100

glance at Table 7, makes one astonished why an average value of the first and second measurement was not taken. In principle it is not difficult to calculate such an average (also with regard to existing covariances), as is shown in Appendix 2. However, an average can only be formed if no further correlations to any other data exist of the data contributing to the average. If we look at Table 7 we recognize correlations between the first as well as the second data and the third data. In such a case the average would ignore the existing correlation to the third data. This means that an average, which is (as will be shown later) equivalent with the least squares procedure, cannot be calculated for a subset of the total data contributing to the least squares.

Thus the given example demonstrates two important aspects of the data evaluation process:

a) the combination of data of the same type which have been determined more than once

b) the connection between absolute and relative cross section data.

In our example we regard the data of the first experiment (Table 6) as estimates of the final data (parameters), i,e.:

$$P = \left(\begin{array}{c} \sigma_1 = 1215 \\ \sigma_2 = 1790 \end{array} \right) \quad \text{and} \quad M = \left(\begin{array}{cc} 473.0 & \\ 519.1 & 1636.5 \end{array} \right) \tag{E-28}$$

and the data of the second experiment (Table 7) as the quantities we call (in section E.2) "data":

$$D^0 = \left(\begin{array}{c} 1205 \\ 1203 \\ 1.500 \end{array} \right) \quad \text{with} \quad V = \left(\begin{array}{ccc} 7.351\text{E}+2 & & \\ 7.880\text{E}+2 & 1.320\text{E}+3 & \\ -1.028\text{E}-1 & -3.624\text{E}-2 & 3.980\text{E}-4 \end{array} \right) \tag{E-29}$$

The vector D which represents the fit function at the parameter estimates is then:

$$D = \left(\begin{array}{c} \sigma_1 = 1215 \\ \sigma_1 = 1215 \\ \sigma_2/\sigma_1 = 1.473 \end{array} \right) \tag{E-30}$$

and the G matrix is given by:

$$G = \left(\begin{array}{cc} 1 & 0 \\ 1 & 0 \\ -\sigma_2/\sigma_1^2 & 1/\sigma_1 \end{array} \right) \tag{E-31}$$

The first column of Eq. (E-31) corresponds to the first parameter, σ_1, and the second column to the second parameter, σ_2, while the rows correspond to the data 1 to 3. The data no. 1 and no. 2 are identical (within their uncertainties) with the first parameter, so we have unity as both first elements of the first column and a corresponding zero in the second column. Data no. 3 (the ratio measurement) depends on both parameters which is expressed in the corresponding derivatives in row 3, columns 1 and 2.

Inserting the results of Eq. (E-28) - (E-31) in the least squares formalism shown in section E.3 (or Table 5), we obtain as a final result:

$$P' = \left(\begin{array}{c} 1210 \\ 1805 \end{array} \right) \quad \text{and} \quad M' = \left(\begin{array}{cc} 285.0 & \\ 349.0 & 789.9 \end{array} \right) \tag{E-32}$$

The resulting χ^2 of 0.65 has to be compared with the number of degrees of freedom of $f = n - m = 3 - 2 = 1$ which indicates that the data of Table 6 and Table 7 were sufficiently consistent. We can therefore quote the result of Eq. (E-32) as an "best" set of evaluated data:

Table 8: Evaluated Data

Neutron Reaction	Spectrum-Averaged Cross Section	Rel.Std.Dev. %	Correlation Matrix ($\times$ 100)	
^{235}U(n,f)	1210 mb	1.40	100	
^{239}Pu(n,f)	1805 mb	1.56	74	100

E.4.2 Interpolation of the Efficiency of a Ge(Li)-Detector

This example illustrates a typical curve fitting problem *and* the subsequent uncertainty propagation to estimate the uncertainty of any interpolated value on this curve. The usual procedure of establishing the efficiency of a detector is to measure this efficiency with a limited set of well-known radionuclide sources. This procedure yields a few points in an efficiency versus gamma ray energy diagram. On a log-log scale the slope of the efficiency curve of a germanium detector is approximately given by a straight line. The efficiency ε_x of any radionuclide, whose gamma ray energy E_x does not coincide with that of the radionuclide used in the calibration procedure, can be determined as soon the functional relationship is established (see Fig. 3). The result of a typical efficiency calibration proce-

Figure 3: Efficiency Curve

dure is given in Table 9. Looking at Table 9 we recognize that the efficiencies of the two radionuclide, ^{57}Co and ^{60}Co, which emit more than one gamma line are correlated. This correlation is easy to understand if one looks at the principle of the efficiency calibration represented by:

$$\varepsilon_i = A_i/(N_{12}h_i) \quad i = 1,2 \tag{E-33}$$

Table 9: Experimentally Determined Points of the Efficiency Curve

No.	Radio-nuclide	E [MeV]	$\varepsilon_i(E)$	Rel.Std.Dev. %	Correlation Matrix ($\times$ 100)						
$i=1$	Co-57	0.122	2.616E-1	2.0	100						
2	Co-57	0.1365	2.331E-1	2.0	56	100					
3	Ba-133	0.356	8.683E-2	3.0	0	0	100				
4	Cs-137	0.662	4.383E-2	2.0	0	0	0	100			
5	Mn-54	0.835	3.385E-2	2.0	0	0	0	0	100		
6	Co-60	1.173	2.359E-2	2.5	0	0	0	0	0	100	
7	Co-60	1.332	2.032E-2	2.5	0	0	0	0	0	64	100

The efficiency is given by the ratio of the measured count rate A_i of the photo-peak area, corresponding to the gamma ray energy E_i, relative to the total decay rate N_{12} of the radionuclide, multiplied by the probability, h_i, of emitting this gamma ray energy. In the case of, for example, ^{57}Co, we have the following relative variances and covariances:

$$
\begin{aligned}
\langle \delta\varepsilon_1 \; \delta\varepsilon_1 \rangle &= (2.0)^2 \\
\langle \delta\varepsilon_2 \; \delta\varepsilon_2 \rangle &= (2.0)^2 \\
\langle \delta\varepsilon_1 \; \delta\varepsilon_2 \rangle &= \langle \delta N_{12} \; \delta N_{12} \rangle = (1.5)^2
\end{aligned}
\tag{E-34}
$$

The covariance term of Eq. (E-34) indicates that the variance of the total decay rate of $(1.5\%)^2$ is common to both efficiencies of the 122 keV and 136.5 keV gamma lines. We have further assumed that the uncertainty contribution due to the peak area determination is negligible compared to the counting statistics i.e. no correlations in the A_i. It should also be mentioned that the gamma ray emission probabilities h_i were regarded free of any norm (as $h_1 + h_2 = 1$, for example). Otherwise the covariance term of Eq. (E-34) has to be modified. Thus the relative covariance matrix of both efficiencies of ^{57}Co is given by:

$$
\begin{array}{c}
\%^2 \\
\begin{array}{cc}
\varepsilon_1 \\
\varepsilon_2
\end{array}
\left(
\begin{array}{cc}
4.00 & 2.25 \\
2.25 & 4.00
\end{array}
\right)
\end{array}
\tag{E-35}
$$

which corresponds to the following standard deviations and correlations:

$$
\begin{array}{ccc}
 & \text{Rel.Std.Dev.} & \\
 & \% & \\
\varepsilon_1 & 2.0 & \left(\begin{array}{cc} 100 & \\ 56 & 100 \end{array} \right) \\
\varepsilon_2 & 2.0 &
\end{array}
\tag{E-36}
$$

The information given in Table 9 established our data vector D^0 and its covariance matrix V. As fit function we assume a straight line on a log-log scale. i.e.:

$$
\varepsilon_i = a \, E_i^b
\tag{E-37}
$$

Table 10: Prior Information of the Fit

Parameters	Rel.Std.Dev. %	Correlation Matrix	
$a^0 = 2.761\mathrm{E}-2$	50.0	100	
$b^0 = -1.069\mathrm{E}+0$	50.0	0	100

The information of the vector P is given by

$$P = \begin{pmatrix} a^0 \\ b^0 \end{pmatrix} \tag{E-38}$$

and may, for example, come from a preliminary rough fit. The numerical values of the vector of the parameter estimates P and of the matrix M are summarized in Table 10. As we have no precise information about the covariance matrix of P, it is sufficient to assume large variances and no correlations, as shown in Table 10. The effect of these large variances or standard deviations is to keep the influence of the parameter estimates on the final result as small as possible. In the first stage of a data evaluation this procedure may often be necessary. However, as soon as more precise information is available it has to be used as prior information. The matrix G in our least squares procedure has the dimension: 7 (data points) $\times$ 2 (parameters) and is given by:

$$G = \begin{pmatrix} E_1^{b^0} & \left(a^0 E_1^{b^0} \ln E_1\right) \\ E_2^{b^0} & \left(a^0 E_2^{b^0} \ln E_2\right) \\ \vdots & \vdots \\ E_7^{b^0} & \left(a^0 E_7^{b^0} \ln E_7\right) \end{pmatrix} \tag{E-39}$$

The final result is now:

$$\begin{array}{lccc} & & \text{Rel.Std.Dev.} & \\ & & \% & \\ a = & 2.803\mathrm{E}-2 & 1.26 & \begin{pmatrix} 100 & \\ 67 & 100 \end{pmatrix} \\ b = & -1.0659\mathrm{E}+0 & 1.02 & \end{array} \tag{E-40}$$

The resulting χ^2 is 1.92 compared with the degrees of freedom of 5. This indicates that our fitting model, Eq. (E-37), adequately describes the data. People more familiar with such efficiency curves may be rather astonished about this, since they know that normally there deviations exist from the straight line model over the given gamma energy

range. The reason why one cannot see such deviations here is probably the relatively large variances of the single calibration points which may cover the deviations. So long as no more detailed information is available it is sufficient to define the fitting model as simply as possible. It would be obscure to insert a more complex expression in Eq. (E-37) while the given form is satisfactory. Of course the correlations shown in Eq. (E-40) would change with a change of the model of Eq. (E-37). However, if one uses Eq. (E-37) in the given form for interpolation then the given correlations are obligatory and cannot be replaced by others belonging to another model.

The purpose of fitting a curve to the calibration points is to enable the efficiency to be obtained at points between the calibration points. The efficiency at an arbitrary gamma ray energy E_x is given by

$$\varepsilon_x = a\, E_x^b \tag{E-41}$$

where a and b are the parameters of Eq. (E-40). Introducing deviations, in analogy to Eq. (B-3), we can write

$$d\varepsilon_x = E_x^b\, da + a\, E_x^b\, \ln E_x\, db \tag{E-42}$$

or, by changing to relative deviations:

$$\delta\varepsilon_x = \delta a + b\, \ln E_x\, \delta b \tag{E-43}$$

The expectation value of Eq. (E-43) defines the relative variances or covariances, respectively, of

$$\begin{aligned}\langle \delta\varepsilon_x\, \delta\varepsilon_y\rangle = \langle \delta a\, \delta a\rangle\quad &+\quad b^2\,(\ln E_x)(\ln E_y)\langle \delta b\, \delta b\rangle \\ &+\quad b\,(\ln E_x + \ln E_y)\langle \delta a\, \delta b\rangle\end{aligned} \tag{E-44}$$

and defines the uncertainty propagation rules in our example. Applying the Eq. (E-41) and (E-44) to a few new radionuclide we obtain the result given in Table 11. The table

Table 11: Propagated Uncertainties of Interpolated Values

No. $i =$	Radio- nuclide	E [MeV]	$\varepsilon_i(E)$	Rel.Std.Dev. %	Correlation Matrix		
1'	$^{115}\text{In}^m$	0.336	8.964E−2	2.2	100		
2'	^{27}Mg	0.8438	3.359E−2	1.4	95	100	
3'	^{24}Na	1.3680	2.006E−2	1.1	80	94	100

shows that the log-log interpolation used establishes very strong correlations between the energy-dependent efficiencies. If necessary, Table 11 can be extended to cover other efficiency values.

E.4.3 Evaluation of Energy-Dependent Cross Sections

The shape of a $\sigma(E)$-curve is more or less complex. It can be established when a variety of experimental data measured at various, but nearly mono-energetic, neutron energies is given. The neutron energies chosen in an experiment depend on the experimental equipment as well as on the convenience of their production. Thus, between reaction threshold and an upper limit of 20 MeV, for example, one can have some hundred points all taken at different energies. Covariance matrices of these dimensions can hardly be handled. Therefore it is advantageous to reduce the data system by collapsing the data onto an intelligently chosen energy grid, which on the one hand is fine enough to reproduce the essential structure of the excitation function and on the other hand avoids unnecessary details, i.e. transforming the system into a manageable form. With these rules most excitation functions of neutron threshold reactions can be represented by 20 to 40 points of the energy grid and interpolation rules between these points can be established with sufficient accuracy.

The present example demonstrates the principle of such data reduction. The more important aspects of this example, however, is to show how cross section ratio measurements establish correlations between the absolute cross section data of different reactions. As the number of ratio measurements often surpasses the absolute measurements, it is of special importance to regard these correlations in future evaluations. We start with an experiment which measured the ^{65}Cu(n,2n) cross section relative to the ^{27}Al(n,α) cross section at eight different neutron energies given in Table 12. The relatively low valued

Table 12: Experiment No.1 - Ratio Measurement of ^{65}Cu(n,2n) to ^{27}Al(n,α)

Data No.	Neutron Energy [MeV]	Measured Ratio σ^{Cu}/σ^{Al}	Rel.Std.Dev. %	Correlation Matrix							
1	13.55	6.62	2.9	100							
2	13.62	6.68	2.9	19	100						
3	13.69	6.98	2.5	23	32	100					
4	13.92	7.18	2.4	20	25	21	100				
5	14.05	7.58	2.4	23	23	26	22	100			
6	14.53	8.12	2.8	28	30	26	31	30	100		
7	14.61	8.44	2.5	20	18	24	27	22	26	100	
8	14.68	8.67	2.0	27	23	24	29	25	30	32	100

elements of the correlation matrix describing the experimental procedure show that the formation of a ratio at least partly eliminates some of the uncertainty components. From

an inspection of the shape of the excitation function, as well as of the reaction ^{27}Al(n,α) and also of ^{65}Cu(n,2n), we can establish an energy grid (given in Table 13) which is adequate for both reactions and so also for the ratio. The preliminary shape of the curve of the ratio constructed from the data of Table 12, and additional data not given in Table 12, allows us to establish the slope of this curve and the interpolation rules within this energy grid. This information is summarized in Table 13.

Table 13: Energy Grid and Interpolation Rules

Group No.	Energy Range [MeV]	E_{mesh} [MeV]	Slope of the Ratio [MeV^{-1}]	Interpolation Rules
1	13.5-13.7	13.6	1.67	linear-linear
2	13.9-14.1	14.0	1.60	linear-linear
3	14.5-14.7	14.6	1.78	linear-linear

With the aid of the data of Table 13 we can transform the data of Table 12 to the chosen energy grid. The result is shown in Table 14.

Table 14: Experimental Data in the New Energy Grid

Data No.	E_{mesh} [MeV]	Ratio σ^{Cu}/σ^{Al}
1	13.6	6.70
2	13.6	6.65
3	13.6	6.83
4	14.0	7.31
5	14.0	7.50
6	14.6	8.24
7	14.6	8.42
8	14.6	8.53

It is obvious that the relative covariance matrix of Table 12 is not influenced by this transformation. Therefore the matrix given in Table 12 is also valid for Table 14.

In the evaluation process we now need prior information, namely the absolute cross section of ^{27}Al(n,α) as well as ^{65}Cu(n,2n) in the same energy grid as above. This information may come from other experiments or from prior evaluations. The data used here are summarized in Tables 15 and 16. The data of Table 15 and Table 16 define the vector

Table 15: Experiment (or Evaluation) No.2 - Cross Section of ^{27}Al(n,α)

E_n [MeV]	Symbol	$\sigma(E_n)$ [mb]	Rel.Std.Dev. %	Correlation Matrix		
13.6	σ_1^A	123.3	4.2	100		
14.0	σ_2^A	120.6	5.0	83	100	
14.6	σ_3^A	114.0	4.8	90	73	100

Table 16: Experiment (or Evaluation) No.3 - Cross Section of ^{65}Cu(n,2n)

E_n [MeV]	Symbol	$\sigma(E_n)$ [mb]	Rel.Std.Dev. %	Correlation Matrix		
13.6	σ_1^C	832.0	5.5	100		
14.0	σ_2^C	891.9	6.3	72	100	
14.6	σ_3^C	959.6	5.8	85	88	100

of the parameter estimates

$$P = \begin{pmatrix} P_1 \\ P_2 \end{pmatrix} \tag{E-45}$$

where the sub-vector P_1, for example, is given by

$$P_1 = \begin{pmatrix} \sigma_1^A \\ \sigma_2^A \\ \sigma_3^A \end{pmatrix} \tag{E-46}$$

and in an analogous way the second sub-vector by the data of Table 16. The covariance matrix of the parameter estimates can also be partitioned into sub-matrices

$$M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \tag{E-47}$$

where M_{11} is the absolute covariance matrix calculated from the data of Table 15 and M_{22} the absolute covariance matrix of the ^{65}Cu(n,2n) cross section. We assume that the data of Table 15 and Table 16 are either independent experiments or, in the case of an evaluation,

do not comprise ratio measurements. Then we have:

$$M_{12} = M_{21} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$ (E-48)

The data vector D^0 is given by Table 14 and the corresponding absolute covariance matrix V is obtained from the data of Table 14 and the relative matrix from Table 12. For completeness the transformation matrix is listed again (where the vertical lines indicate identical values):

$$
\begin{array}{c}
\text{Row} = 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ \text{Column} =
\end{array}
\;
G =
\left(
\begin{array}{cccccc}
-\sigma_1^C/(\sigma_1^A)^2 & 0 & 0 & 1/\sigma_1^A & 0 & 0 \\
 & & & & & \\
0 & -\sigma_2^C/(\sigma_2^A)^2 & & 0 & 1/\sigma_2^A & \\
 & & & & & \\
0 & & -\sigma_3^C/(\sigma_3^A)^2 & & 0 & 1/\sigma_3^A \\
 & & & & & \\
 & & & & & \\
1 & 2 & 3 & 4 & 5 & 6
\end{array}
\right)
$$

The result of combining a ratio measurement with two absolute measurements gives a χ^2 of 2.87 and is shown in Table 17 and Table 18.

<div align="center">Table 17: Final Result of the Evaluation</div>

E_n [MeV]	^{27}Al(n,α)		^{65}Cu(n,2n)	
	$\sigma(E_n)$ [mb]	Rel.Std.Dev. %	$\sigma(E_n)$ [mb]	Rel.Std.Dev. %
13.6	123.2	3.4	832.3	3.5
14.0	120.5	3.9	894.4	4.0
14.6	113.9	3.7	961.6	3.8

If we compare the uncertainties of Table 17 with the values given in Table 15 and Table 16 we recognize that the addition of further information, here in form of additional ratio measurements, reduces the uncertainty of the system. The correlation matrix of the final evaluation is given in Table 18.

It is not in the nature of neutron cross sections to be correlated. Correlations are only established due to the measuring process of such data. While the measurement of absolute cross sections introduces correlations between the data belonging to the same

Table 18: Correlation Matrix of the Final Evaluation

Reaction	E_n [MeV]						
	13.6	100					
^{27}Al(n,α)	14.0	82	100				
	14.6	89	79	100			
	13.6	87	72	80	100		
^{65}Cu(n,2n)	14.0	76	90	75	76	100	
	14.6	83	76	90	86	82	100
	E_n [MeV]	13.6	14.0	14.6	13.6	14.0	14.6
	Reaction	^{27}Al(n,α)			^{65}Cu(n,2n)		

cross section, ratio measurements additionally generate correlations between the data of different cross sections. Before the evaluation process we had zero correlation between both reactions ^{27}Al(n,α) and ^{65}Cu(n,2n). The addition of ratio measurements to the prior system establishes large correlations between both reactions, as we can see in Table 18.

The resulting χ^2 of the least squares procedure defines the consistency of the data included in the evaluation process. The probability density of χ^2 has an expectation value which is given by the number of degrees of freedom f and a variance of $2f$. The χ^2 distribution is not symmetrical but as soon as the number of degrees of freedom increases beyond a few, the median of the density function approaches f. As the concept of generalized least squares bases on realistic estimates of the variances, a value of χ^2/f of less than unity never justifies a reduction of the final uncertainties. What happens, however, if a value of χ^2/f greater than unity indicates inconsistencies in the data? The first step must be to check the input data of the least squares procedure again, to correct recognized erroneous data and to eliminate obviously wrong data. This step cannot always be done perfectly because of the lack of sufficient information, In such cases it may become necessary to scale up all covariance elements of the output by a factor of χ^2/f with the implicit assumption that the shape of the evaluated matrix is in principle correct. This procedure should never be done without informing the users of the data. In the example here given we obtained a χ^2 of 2.87 compared to a number of degrees of freedom of 2. Taking this into account would increase the standard deviations given in Table 17 by a factor of $\sqrt{1.435} \simeq 1.2$.

F Summary

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 *not recommended*, that an experimenter evaluates this covariance matrix. The reason for this is that a 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.

The generalized least squares procedure shown in section E.3 and Table 5 permits an easy way of interchanging data D^0 with parameter estimates P. This means new data can easily be combined with an earlier evaluation. However, it must be mentioned that this is only valid as long as the new data have no correlation with any of the older data of the prior evaluation. Otherwise the old data which show correlation with new data have to be extracted from the evaluation and then, together with the new data and taking account of the correlation, have again to be added to the reduced evaluation. In most cases this step cannot be performed and the evaluation has to be completely redone. A partial way out is given if the evaluation is performed step by step and the results of each step are stored. Then the evaluation need only be repeated from the step which contains correlated data for the first time while all earlier steps remain unchanged.

Finally it should be noted that the addition of a small set of new data to a prior evaluation consisting of a large number of data does not require an inversion of a matrix of the dimensions of the large set. As can be seen from Eq. (E-27), the dimension of the matrix which has to be inverted is given by the dimension of the new data set.

References

[1] F.G. Perey: "Covariance Matrices of Experimental Data", Proceedings of an International Conference on Neutron Physics and Nuclear Data for Reactors, Harwell, United Kingdom, (September 1978), 104.

[2] W. Mannhart. F.G. Perey: "Covariance Matrices of Cf-252 Spectrum Averaged Cross Sections", Proc. Third ASTM-EURATOM Symposium on Reactor Dosimetry, Ispra, Italy, Oct. 1-5, 1979, EUR 6813. Vol. II (1980) p. 1016.

[3] J.J. Wagschal, R.E. Maerker, D.M. Gilliam: "Covariances of Fission-Integral Measurements at the NBS ^{252}Cf and ISNF Facilities and at the ORNL-PCA Facility", EUR 6813, Vol. II (1980), p. 683.

[4] S. Wagner: "On the Quantitative Characterization of the Uncertainty of Experimental Results in Metrology", PTB Mitteilungen **89** (1979), 83.

[5] K. Kobayashi, I. Kimura: "Fission Spectrum Averaged Cross Sections with Standard Neutron Fields", EUR 6813, Vol. II (1980), p. 1004.

[6] K. Kobayashi: Priv. communication (November, 1979)

[7] W.C. Hamilton: "Statistics in Physical Science", Ronald Press, New York (1964) pp. 12.

[8] W. Matthes: "Cross-Section Adjustment in the Analysis of Benchmark Experiments", Ann. Nucl. Energy, **6** (1979) 103.

[9] D.L. Smith: "Probability, Statistics and Data Uncertainties in Nuclear Science and Technology", American Nuclear Society, La Grange Park, IL, USA (1991).

Appendix 1: Propagation of Correlated Uncertainties

An *linear* function of data can be represented by a vector

$$F = SX \tag{1}$$

where F is the vector of the values of the function and X the vector of the data. The elements of the matrix S, often called sensitivity matrix, are given by the partial derivatives:

$$S_{ij} = \frac{\partial F_i}{\partial X_j} \tag{2}$$

The connection between the covariance matrix of the function M_F and the covariance matrix of the data M_X, is given by:

$$M_F = S\, M_X\, S^T \tag{3}$$

Eq. (3) establishes the propagation rules for uncertainties.

Applying the equations (1) - (3) to the example given in section D.1 we obtain:

$$F = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad X = \begin{pmatrix} l_1 \\ l_2 \\ l_3 \end{pmatrix} \quad S = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

with

$$M_X = \begin{pmatrix} \mathrm{Var}(l_1) & 0 & 0 \\ 0 & \mathrm{Var}(l_2) & 0 \\ 0 & 0 & \mathrm{Var}(l_3) \end{pmatrix}$$

The final covariance matrix is given by:

$$
\begin{aligned}
M_F &= \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathrm{Var}(l_1) & 0 & 0 \\ 0 & \mathrm{Var}(l_2) & 0 \\ 0 & 0 & \mathrm{Var}(l_3) \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 0 \\ 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} \mathrm{Var}(l_1) + \mathrm{Var}(l_2) & \mathrm{Var}(l_1) \\ \mathrm{Var}(l_1) & \mathrm{Var}(l_1) + \mathrm{Var}(l_3) \end{pmatrix}
\end{aligned}
$$

For the case of a *non-linear* function we have shown in Eq. (B-27) and (B-31) how the uncertainty propagation can be linearized. Instead of a linear function of the data we now have a linear function of the deviations between the data and their expectation values. Following the notation of section B we obtain for the *relative* deviations:

$$(\delta F) = S\,(\delta X) \tag{4}$$

47

(δF) is now the vector of the relative deviations of the function and (δX) the vector of the relative deviations of the data. The matrix S changes from a sensitivity matrix to a relative sensitivity matrix with elements of:

$$S_{ij} = \frac{X_j}{F_i} \frac{\partial F_i}{\partial X_j} \tag{5}$$

In the case of a functional relationship with product or quotient terms, these elements have the simple values +1 or -1. The propagation of the uncertainties is now given by:

$$\langle (\delta F)\,(\delta F)^T \rangle = S \langle (\delta X)(\delta X)^T \rangle S^T \tag{6}$$

The notation of Eq. (6) is an alternative one compared to that of Eq. (3). The product of the vectors $(\delta F)\,(\delta F)^T$ corresponds to the relative covariance matrix of the function with elements of the form $\langle \delta F_i\, \delta F_j \rangle$.

Application of the Eq. (4) to (6) to the example of section D.2.2 results in:

$$(\delta F) = \begin{pmatrix} \delta R_{12} \\ \delta R_{34} \end{pmatrix} \quad (\delta X) = \begin{pmatrix} \delta P_1 \\ \delta P_2 \\ \delta P_3 \\ \delta P_4 \end{pmatrix} \quad S = \begin{pmatrix} -1 & +1 & 0 & 0 \\ 0 & 0 & -1 & +1 \end{pmatrix}$$

The final covariance matrix is given by:

$$\langle (\delta F)\,(\delta F)^T \rangle = \begin{pmatrix} \langle \delta R_{12}\,\delta R_{12} \rangle & \langle \delta R_{12}\,\delta R_{34} \rangle \\ \langle \delta R_{34}\,\delta R_{12} \rangle & \langle \delta R_{34}\,\delta R_{34} \rangle \end{pmatrix}$$

$$= \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 9.58 & 6.16 & 6.64 & 6.00 \\ 6.16 & 16.46 & 7.93 & 6.15 \\ 6.64 & 7.93 & 13.86 & 6.64 \\ 6.00 & 6.15 & 6.64 & 17.29 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 13.72 & -1.14 \\ -1.14 & 17.87 \end{pmatrix}$$

A further application of the Eq. (4) to (6) can be demonstrated for the example given in section E.4.2 with Eq. (E-41) to (E-44). Here we have:

$$(\delta F) = \begin{pmatrix} \delta \varepsilon_{1'} \\ \delta \varepsilon_{2'} \\ \delta \varepsilon_{3'} \end{pmatrix} \quad (\delta X) = \begin{pmatrix} \delta a \\ \delta b \end{pmatrix} \quad S = \begin{pmatrix} 1 & b \ln E_{1'} \\ 1 & b \ln E_{2'} \\ 1 & b \ln E_{3'} \end{pmatrix}$$

It can easily be proved, by inserting the corresponding numerical values, that Eq. (6) produces the relative covariance matrix given in Table 11.

Appendix 2: Off-Diagonal Weighted Averages

The combination of various correlated data of the same quantity can be done in a similar way as for uncorrelated data. While for uncorrelated data the weights in evaluating an average value are given by the diagonal elements of the covariance matrix, for correlated data the off-diagonal matrix elements must also be considered. The rules of the data combination are given by the least squares equations (Eq. (E-21) of section E.2) when we neglect the vector of the parameter estimates P on the left hand side of Eq. (E-21) and correspondingly also the vector D on the right hand side, i.e.:

$$G^T V^{-1} G\, P' = G^T V^{-1} D^0 \tag{7}$$

Now P' represents the average value of the data given by the vector D^0. The solution of Eq. (7) can be written as

$$\bar{y} = \sum_i w_i\, y_i \tag{8}$$

with $\bar{y}$ being the average value of the various y_i. The weights are given by

$$w_i = \frac{\sum_j V_{ji}^{-1}}{\sum_k \sum_l V_{kl}^{-1}} \tag{9}$$

with V_{ij}^{-1} being the elements of the inverse of the absolute covariance matrix of the data y_i. The definition of the w_i includes that $\sum_i w_i = 1$ holds. The variance of Eq. (8) is given by:

$$\mathrm{Var}(\bar{y}) = \sum_{ij} w_i V_{ij} w_j \tag{10}$$

As already mentioned, the Eq. (8) - (10) can only be applied if the data y_i show no correlation to any other data. For the simple case of two correlated data given by the vector

$$D^0 = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \tag{11}$$

and its covariance matrix

$$V = \begin{pmatrix} V_{11} & V_{21} \\ V_{21} & V_{22} \end{pmatrix} \tag{12}$$

Eq. (8) is now:

$$\bar{y} = \frac{(V_{11}^{-1} + V_{21}^{-1})\, y_1 + (V_{22}^{-1} + V_{21}^{-1})\, y_2}{V_{11}^{-1} + V_{22}^{-1} + 2V_{21}^{-1}} \tag{13}$$

As the analytical inversion of a matrix of the dimension 2 can easily be done, we can also rewrite Eq. (13) as

$$\bar{y} = \frac{(V_{22} - V_{21}) \, y_1 + (V_{11} - V_{21}) \, y_2}{(V_{11} + V_{22} - 2 \, V_{21})} \tag{14}$$

and also

$$\text{Var}(\bar{y}) = \frac{V_{11}V_{22} - (V_{21})^2}{V_{11} + V_{22} - 2V_{21}} \tag{15}$$

Eq. (14) and (15) are very useful in illustrating some aspects of the data evaluation process:

Case A: $V_{21} = V_{11}$
Inserting $V_{21} = V_{11}$ in Eq. (14) and (15) we obtain:

$$\bar{y} = y_1 \quad \text{and} \quad \text{Var}(\bar{y}) = V_{11} \tag{16}$$

regardless of the value of V_{22}. At first glance this result is surprising. But it can easily be understood if one recognizes that $V_{21} = V_{11}$ means that the second data y_2 contains no new information compared to the first. It must therefore consistently be neglected in the evaluation process. A special case of the given example is for $V_{11} = V_{21} = V_{22}$: here the matrix V becomes singular, indicating that y_1 and y_2 are two not independent measurements.

Case B: $V_{11} < V_{21} < V_{22}$
In this case we find out that $\bar{y}$ of Eq. (14) is outside the interval (y_1, y_2). If one is not used to work with non-diagonal covariance matrices this result is unfamiliar. It is, however, consistent in view of a complete covariance matrix which describes all correlations.

Case C: $V_{11} = V_{22}$ (Data of equal variance)
The combination of two data of equal variance demonstrates the influence of correlations on the reduction of the uncertainties. We obtain from Eq. (14) and (15):

$$\bar{y} = \frac{y_1 + y_2}{2} \quad \text{and} \quad \text{Var}(\bar{y}) = \frac{V_{11} + V_{21}}{2} \tag{17}$$

This result should be compared with that of uncorrelated data ($V_{21} = 0$):

$$\bar{y} = \frac{y_1 + y_2}{2} \quad \text{and} \quad \text{Var}(\bar{y}) = \frac{V_{11}}{2} \tag{18}$$

In comparing both these results one recognizes that the average value is the same for correlated and uncorrelated data. For uncorrelated data (Eq. (18)) the variance of the

average is given by half of the common variance of both data. For correlated data the variance of the average remains above this value for a positive value of V_{21}. This means a positive correlation of data lessens the uncertainty reduction of an evaluated result. In the past this effect was sometimes responsible for inconsistencies in evaluated data, since the neglect of covariances gave too strong a reduction of the evaluated final uncertainties.