The MATCH code – ideas and realization

Karl-Heinz Schmidt, 29. 10. 2014

Concept

Estimating best values

General case

Perturbed-parameter calculations with the GEF code provide event-wise multivariate distributions of fission yields. They represent the fission yields and their correlations given by the physics of the model. From these, covariance matrices are determined. (The diagonal elements represent the variances of the calculated yields.) In the present version of GEF, complete covariance matrices of element (Z) yields and post-neutron fission-fragment mass (A) yields and nuclide (A,Z) yields are available. Eventually, the output may be extended to covariance matrices of cumulative yields.

From any covariance matrix K_{model}, the corresponding analytical multivariate distribution f_{model} can be determined:

$$f_{model}=\frac{1}{(2\pi)^{n/2}\sqrt{\det(K_{model})}}\cdot\exp\left(-\frac{1}{2}(y-m_y)^t K_{model}^{-1}(y-m_y)\right) \tag{1}$$

This multivariate distribution expresses the probability distribution that the values y are compatible with the result of the model. (y is the vector of yields (as a variable), m_y is the vector of the corresponding mean values of the model yields.)

The term $(y-m_y)^t K_{model}^{-1}(y-m_y)$ is the square of the so-called Mahalanobis distance [https://en.wikipedia.org/wiki/Mahalanobis_distance, http://www.real-statistics.com/multivariate-statistics/multivariate-normal-distribution/multivariate-normal-distribution-basic-concepts/]. The Mahalanobis distance is a multi-dimensional generalization of the idea of measuring how many standard deviations away a point $P(y_i,y_j)$ is from the mean of the distribution f_{model}.

From the distribution f_{model}, the following log-likelihood function L_{model} is defined [1]:

$$L_{model}=-\frac{1}{2}(y-m_y)^t K_{model}^{-1}(y-m_y) \tag{2}$$

When measured values y_{exp} of the observables y with their individual uncertainties s_i are available, they define an experimental multivariate distribution:

$$f_{\exp}=\frac{1}{(2\pi)^{n/2}\prod_i s_i}\cdot\exp\left(-\frac{1}{2}\left(\frac{y-y_{\exp}}{s}\right)^2\right) \tag{3}$$

and the corresponding log-likelihood function:

1 Pre-exponential terms are neglected in definitions of this and the other log-likelihood functions.

$$L_{\exp} = -\frac{1}{2}\left(\frac{y - y_{\exp}}{s}\right)^2 \quad (4)$$

In case the experiment also provides covariances, the experimental multivariate distribution reads:

$$f_{\exp} = \frac{1}{(2\pi)^{n/2}\sqrt{\det(K_{\exp})}} \cdot \exp\left(-\frac{1}{2}(y - y_{\exp})^t K_{\exp}^{-1}(y - y_{\exp})\right) \quad (5)$$

with the corresponding log-likelihood function:

$$L_{\exp} = -\frac{1}{2}(y - y_{\exp})^t K_{\exp}^{-1}(y - y_{\exp}) \quad (6)$$

The “best guess” of yield values, considering the constraints by the model and the experiment, are given by maximising the value of the common log-likelihood function

$$L = L_{model} + L_{\exp} \quad (7)$$

with respect to the *y* values.

In order to allow for more flexibility, 3 weighting factors, W_{abs}, W_{corr} and W_{exp} are introduced:

$$L = W_{\mathrm{abs}} \cdot L_{model}^{diagonal} + W_{corr} \cdot L_{model}^{non-diagonal} + W_{\exp} \cdot L_{\exp} \quad (8)$$

($L_{model}^{diagonal}$ includes only the diagonal elements and $L_{model}^{non-diagonal}$ includes only the non-diagonal elements of K_{model}^{-1} .)
This should allow, for example, to reduce the influence of the absolute values of the model calculation by choosing a value $W_{abs} < 1$.

2-dimensional case

For illustration, some equations are explicitly given for the 2-dimensional case.
The analytical multivariate distribution is:

$$f_{model}(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho_{xy}^2}} \cdot \exp\left(-\frac{1}{2(1-\rho_{xy}^2)}\left[\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2} - \frac{2\rho_{xy}(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}\right]\right) \quad (9)$$

The two yields are denoted by *x* and *y*, their mean values by μ_x and μ_y, their standard deviations (square root of the variances) by σ_x and σ_y, and the correlation coefficient between the variables *x* and *y* is given by ρ_{xy} . The correlation coefficient ρ_{xy} is related to the covariance cov_{xy} by

$$cov_{xy} = \rho_{xy} \cdot \sigma_x \cdot \sigma_y \quad (10).$$

The corresponding log-likelihood function is:

$$L_{model} = \frac{1}{2(1-\rho_{xy}^2)}\left[\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2} - \frac{2\rho_{xy}(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}\right] \quad (11).$$

Alternative approach

The above described approach requires the inversion of the covariance matrix. In the case of a large number of dimensions, this is a time-consuming calculation that can be complicated by problems of numerical instability. In order to circumvent these difficulties, an alternative procedure has been developed that is particularly aimed for the case that only the correlations of the model calculations are considered and their absolute values are disregarded.

The approach is based on the geometrical properties of the multivariate distribution of the model results.

The contour lines of a multivariate distribution, projected on the 2-dimensional surface spanned by the coordinates x and y, is shown in figure 1. The (x - μ_x) and (y - μ_y) values are normalized to the corresponding standard deviations σ_x and σ_y. The values of x and y have a positive correlation. Their iso-contour lines are ellipses with an axis ratio ≠ 0. (A value of 2 is chosen in this example). The longer axis of the ellipses has an angle of 45% with respect to the x axis (counter clock wise). When the two variables have a negative correlation, the longer axis of the iso-contour lines has an angle of -45% with respect to the x axis (counter clock wise).

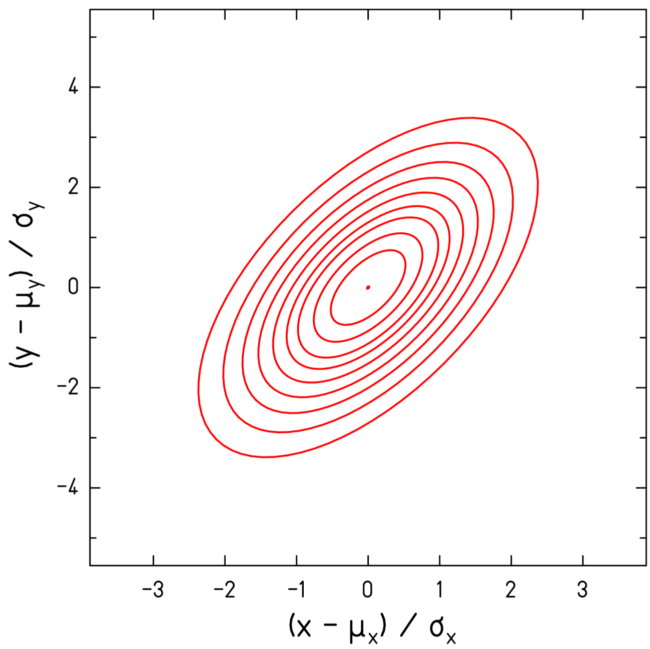

Figure 1: *Iso-contour lines of a normalised two-dimensional multivarite Gaussian distribution with positive correlation. Heights of the contour lines are shown in steps of 0.1 times the maximum value.*

The probabilities along the axes of the ellipses (iso-contour lines) of figure 1 under 45° along the line $(x - \mu_x)/\sigma_x = (y - \mu_y)/\sigma_y$ and under -45° along the line $(x - \mu_x)/\sigma_x = -(y - \mu_y)/\sigma_y$ are shown in figure 2. The ratio of the two curves represents the relative decrease of probability at any point due to its distance $r = (x - \mu_x)/\sigma_x - (y - \mu_y)/\sigma_y$ from the optimum correlation line $(x - \mu_x)/\sigma_x = (y - \mu_y)/\sigma_y$ for any fixed value of $(x - \mu_x)/\sigma_x + (y - \mu_y)/\sigma_y$.

The deviation between the two curves (in and perpendicular to the correlation direction) increases if the correlation between the variables x and y becomes stronger. They coincide if the variables are uncorrelated.

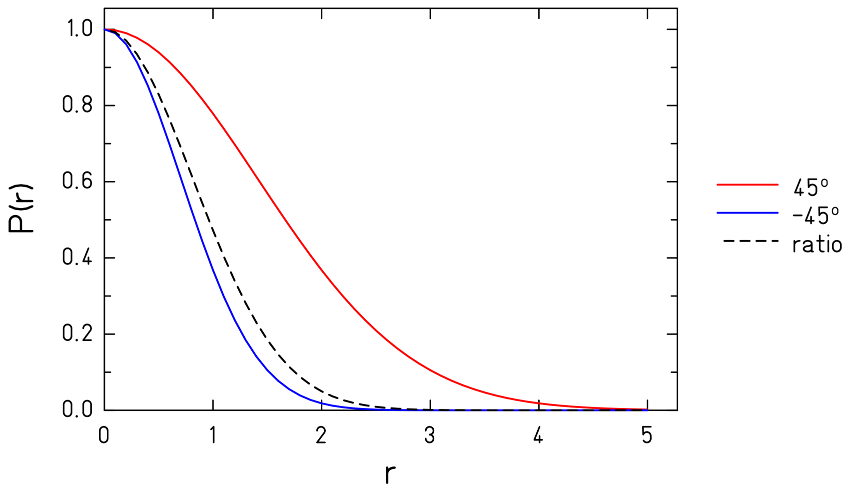

Figure 2: *Probabilities of the multivariate distribution shown in figure 1 along the lines in and perpendicular to the correlation direction. The ratio of the two curves is shown in addition.*

Figure 3 shows the corresponding term for the example given in figures 1 and 2 in the log-likelihood function as a function of $r = (x - \mu_x)/\sigma_x - (y - \mu_y)/\sigma_y$. In the present case of a Gaussian multivariate distribution, this is a parabola.

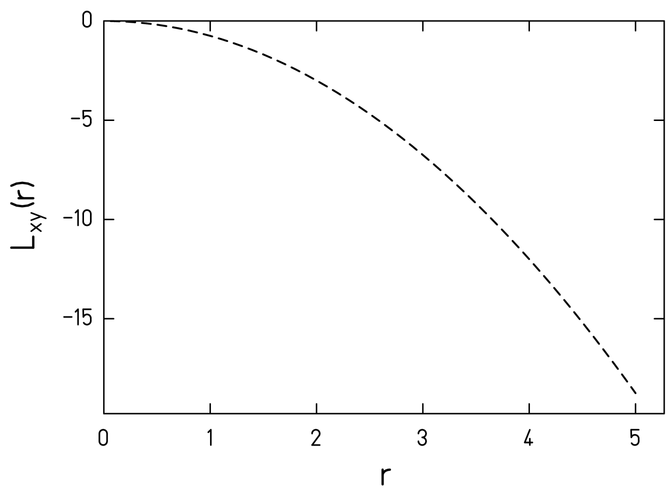

Figure 3: *Log-likelihood function $L_{xy}(r)$ that expresses the likelihood as a function of $r = (x - \mu_x)/\sigma_x - (y - \mu_y)/\sigma_y$ due to a deviation from the optimum correlation of the two variables x and y for the example shown in figure 1.*

According to these ideas, the correlation between any pair of calculated fission-fragment yields x and y is expressed by the term

$$L_{xy} = -A(\rho_{xy}) \cdot ((x - \mu_x)/\sigma_x - (y - \mu_y)/\sigma_y)^2 \quad . \tag{12}$$

The term $L_{model}^{non-diagonal}$ of the log-likelihood function in equation(8) is replaced by the sum over all L_{xy} with $x \neq y$.

The value of $A(\rho_{xy})$ is zero when x and y are uncorrelated, and it grows monotonically with the magnitude of the correlation coefficient ρ_{xy}. Figure 4 shows the function $A(\rho)$ from a numerical calculation and the function $A(\rho)=0.27\cdot\rho/(1-\rho)$ that is implemented in the MATCH code.

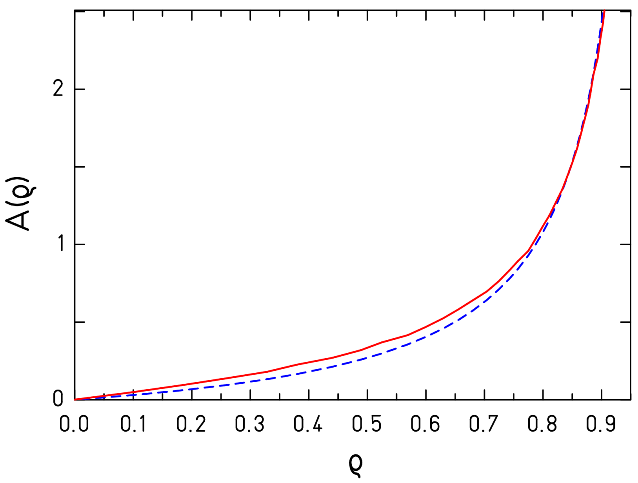

Figure 4: *The function A(ρ) (defined in the text) from a numerical calculation (red line) in comparison with the function* $A(\rho)=0.27\cdot\rho/(1-\rho)$ *(dashed blue line).*

Including different classes of observables

In an evaluation process, all experimental information should be used. Thus, the information from measured Z yields, nuclide yields and eventually cumulative yields must be combined. This is easily possible by constructing a vector y that contains all classes of available observables. The corresponding covariance matrix contains the covariances between all observables, including the covariances between observables of different classes. Then, the same procedure is applied as described above.

Estimating uncertainties and covariances

The combined log-likelihood function L defines the corresponding multivariate distribution

$$f=c\cdot\exp(L). \qquad (??)$$

The constant c is determined by the normalization condition $\int f \, \mathrm{d}y=1$.
From this combined multivariate distribution, the corresponding covariance matrix may be deduced by a suitable analytical or numerical method.

The calculation of uncertainties is still under development.

Practical considerations

Availability of experimental covariances

Experimental values are not only subject to uncertainties, but in most cases they are also more or less correlated, even if the covariances are not given. If only the uncertainties of the experimental values are included in the above procedure, and the correlations are disregarded, the different experimental values are considered to provide independent information. In this case, the correlations provided by the model calculations lead to unrealistically small uncertainties of the evaluated values. Therefore, it is important to include realistic correlations (respectively covariances) of the measured results into the evaluation procedure. This is a problem, if experimental covariances are not available.

If the experimental correlations are not available and they cannot be estimated, a possible conservative solution may be to increase the experimental uncertainties that enter into the above described procedure in a way that the final uncertainties of the results of this procedure are always equal or larger than the nominal experimental uncertainties.

Example

In figures 5 and 6, the result of the GEF code and the result of the MATCH code with different options are compared with the available experimental post-neutron mass distribution for the fission of 241Pu, induced by fast neutrons. For the calculation, a neutron energy of 2.5 MeV was assumed.

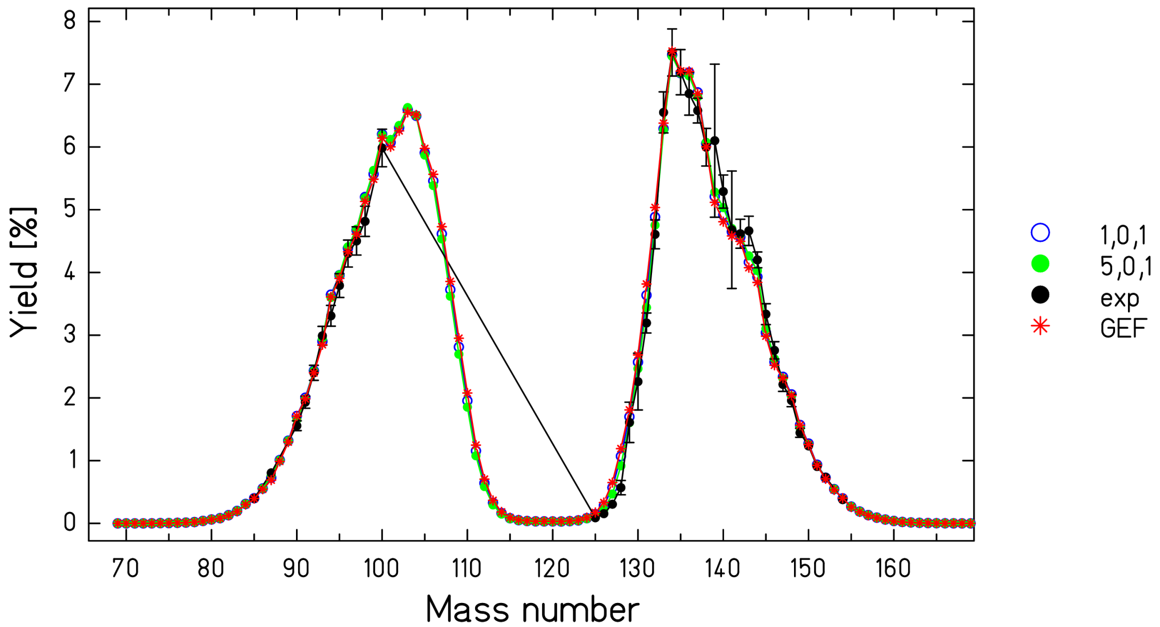

Figure 5: *Adjustment of the fission-fragment mass yields from GEF (red stars) to evaluated data from ENDF/B-VII (black full dots, connected by a line) with the MATCH code for the system ^{241}Pu(n,f), E_n = 2.5 MeV. The experimental data [1] are rather incomplete. The blue open symbols show sets of fission yields that maximise the likelihood function of the measured data and the covariances from GEF. The full green symbols show the result of another calculation with the relative weight of the experimental data increased by a factor of 5.*

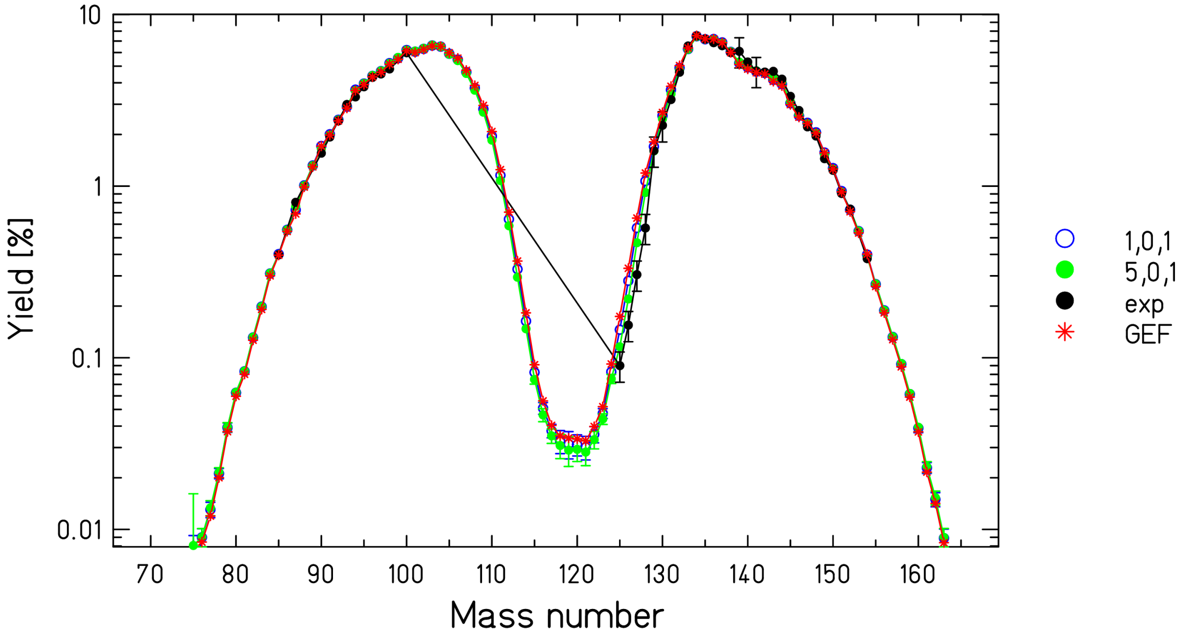

Figure 6: *Like figure 5, but with logarithmic vertical scale.*

The figures show that the MATCH code finds a compromise between the measured values and the result of the GEF model. An increased weight of the experimental data shifts the result of the MATCH code closer to the data, in particular to those with smaller error bars. The fact that the GEF result only enters by the covariances leads to a consistent shift of neighbouring points which have a high degree of correlation. For example, the green points are shifted down in the whole region around symmetry between A = 110 and A = 130 compared to the GEF result, because GEF overestimates the yields in the lower wing of the heavy peak between A = 126 and A = 130.

References:

[1] R. W. Mills, PhD thesis, University of Birmingham, 1995.