M. Filipowicz*, V. M. Bystritsky**, J. Woźniak***

MONTE CARLO FITTING OF DATA FROM MUON CATALYZED FUSION EXPERIMENTS IN SOLID HYDROGEN

Applying the classical chi-square fitting procedure for multiparameter systems is in some cases extremely difficult due to the lack of an analytical expression for the theoretical functions describing the system. This paper presents an analysis procedure for experimental data using theoretical functions generated by Monte Carlo method, each corresponding to definite values of the minimization parameters. It was applied for the E742 experiment (TRIUMF, Vancouver, Canada) data analysis with the aim to analyze data from Muon Catalyzed Fusion experiments (extraction muonic atom scattering parameters and parameters of pd fusion in $pd\mu$ molecule).

Keywords: *data fitting, Monte Carlo simulation, chi-square estimator, gradient calculation, diffusion simulation, muonic atoms and molecules*

DOPASOWYWANIE METODĄ MONTE CARLO DANYCH Z EKSPERYMENTÓW KATALIZY MIONOWEJ W ZESTALONYM WODORZE

Zastosowanie klasycznej procedury fitowania przy użyciu estymatora chi-kwadrat jest w pewnych przypadkach bardzo trudne z powodu braku analitycznych wyrażeń na teoretyczne funkcje opisujące system. W pracy zaprezentowano procedurę analizy danych eksperymentalnych przy użyciu teoretycznych funkcji generowanych metodą Monte Carlo, odpowiadających zdefiniowanym wartościom szukanych parametrów. Analiza została zastosowana dla eksperymentu E742 (TRIUMF, Vancouver, Kanada) w celu opisu danych pochodzących z syntezy jądrowej katalizowanej mionami (uzyskanie parametrów opisujących rozpraszanie atomów mionowych i parametrów syntezy pd w molekule $pd\mu$).

Słowa kluczowe: *dopasowywanie danych, symulacja Monte Carlo, estymator chi-kwadrat, obliczenia gradientu, symulacja dyfuzji, atomy i molekuły mionowe*

* Faculty of Fuels and Energy, AGH University of Science and Technology, Krakow, Poland, filipow@agh.edu.pl

** Laboratory of Nuclear Problems, Joint Institute for Nuclear Research, Dubna 141980, Russia, bystvm@jinr.ru

*** Faculty of Physics and Applied Computer Science, AGH University of Science and Technology, Krakow, Poland, wozniak@ftj.agh.edu.pl

1. Introduction

For a wide range of physical problems the only applicable way to compare experiment with theory is via the Monte Carlo (MC) method. Problems of this type are often multiparameter, with nontrivial interdependencies between the parameters, such as averaging arising from spatially discrete effects. Thus, only an exact simulation of the experimental system allows us a possible analysis, and hence we rely on the MC method.

However, the MC method has several limitations, mostly related to the calculation time and nature of random sampling. Provided the simulation has been correctly established to analyze all competitive processes, the modeling of events which are fairly rare and hidden inside many other processes often requires long calculation time to establish sufficient statistics for comparison. The intrinsic nature of random numbers and the generators presently in use cause that any two simulations of the same system can give different results. Normally, for large numbers of generated statistics these differences are small and completely insignificant. Such limitations are not very important when we carry on a single simulation for one experiment, i.e., when there are no variable parameters and one output result is sufficient. However, the limitations are amplified when we apply the MC method to fitting procedures.

As is fairly well known, finding the best fit parameters describing a multitude of data requires repeated calculation of some statistical estimator. Most often, the estimator is χ^2, which is defined as the difference between the theoretical description and the experimental data, see Eq. (3). The theoretical description of the data depends on several parameters and thus χ^2 is calculated as a function of those parameters. Finally, the result is the set of parameters for which the χ^2 is minimal.

Classical fitting algorithms, e.g., the minimization package MINUIT [1], calculate χ^2 from the model parameters (see Fig. 1). When the χ^2 minimum depends on two or more variables, the error determination on the parameters as well as the study of the possible occurrence from several minima require calculations of several thousands theoretical functions, and hence, the calculations becomes extremely time-consuming. However, the MC evaluation of the theoretical function, just for one set of parameters, is very time exhaustive (measured in hours or even days): thousands of iterations are not possible. Even if the calculation time were acceptable, the intrinsic nature of MC simulations makes such an approach impossible since instabilities will arise resulting from the statistical nature of the results. When fitting, the minimization procedure examines the behaviour of differences in χ^2 for differing values of the parameter set. The minimization procedure then calculates the internal gradient of χ^2 and uses it to control the minimum searching procedure. The gradient is obtained from a set of partial derivatives for each variable parameter where the derivatives are calculated numerically from difference quotients. Normally, the minimum should be reached when all the gradients converge to zero. Clearly, the statistical fluctuations of the MC method can cause entirely false gradients, and thus such a minimization procedure is not suitable for our problem.

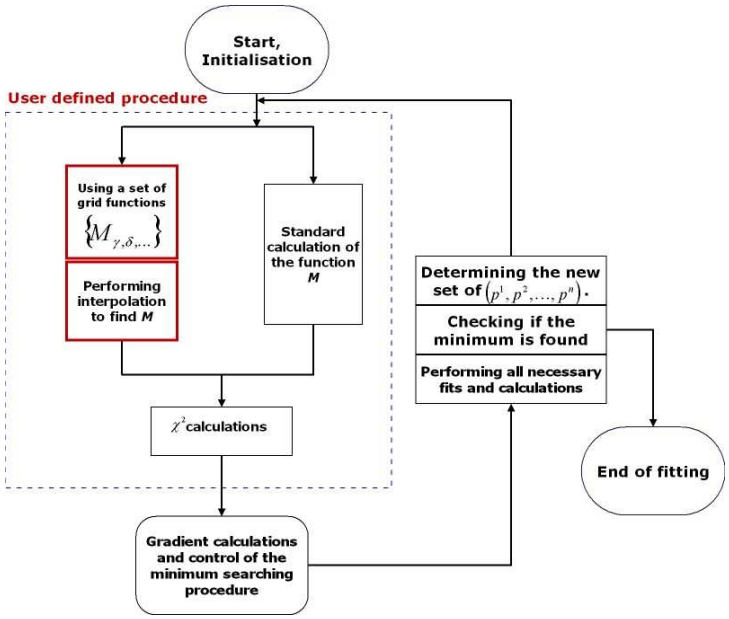

Fig. 1. Diagram of fitting procedure using grids. Modifications are marked with by red line. The dashed line denotes the procedure defined by users

The work of Zech [2] presents methods for comparing MC generated histograms to experimental data when the analytic distribution is known. In our case we compare the experimental data with MC simulations [3] which including all parameters of the apparatus, such as the spatial separation of processes, detector resolution, dead time, etc, and therefore, we can directly compare experimental and Monte Carlo spectra.

If we ask "given a data distribution, and a set of MC distributions, what is the best estimate of the fraction of each MC distribution present in the data distribution?" a standard set of subroutines [4], are available to solve that problem. However, the experimental histograms in our case are not equivalent to assuming of discrete MC spectra, and the method above cannot be applied. The aim our paper is to describe such algorithms and to show via example that it is fully applicable. As examples the scattering of muonic atoms on a structure of crystalline hydrogen and fusion in a $pd\mu$ molecule are presented.

2. Description of Method

2.1. Modified fitting procedure

We proposed a modification to the calculation of the theoretical functions $M(\vec{p})$ which describe the data for a given parameter vector $\vec{p} = (p_1, p_2, \ldots, p_n)$. Before fitting, one generates a set of theoretical functions $\{M\} = \{M_{\gamma,\delta,\ldots,\phi},\ M_{\gamma',\delta',\ldots,\phi'},\ \ldots\}$ for

all permutations of a *chosen* discrete set of parameter values $(p_1^{\gamma}, p_2^{\delta}, \ldots, p_n^{\phi})$ where $\gamma, \delta, \ldots, \phi$ are the function indexes in the set. The parameter vector $\vec{p}$ is allowed to assume only discrete values which gives the grid of theoretical functions a size $\gamma \times \delta \times \ldots \times \phi$, and means that the time-consuming calculations are only executed for a selected and limited parameter set $\{p\} = \{(p_1^{\gamma}, p_2^{\delta}, \ldots, p_n^{\phi}), (p_1^{\gamma'}, p_2^{\delta'}, \ldots, p_n^{\phi'}), \ldots\}$. The resulting set of function values, $\{M\}$, is used to calculate any theoretical function M for any arbitrary parameter vector $\vec{p}$ (provided all p_i values in $\vec{p}$ are between some calculated values of p_i^{γ} and $p_i^{\gamma'}$ contained in the grid) using the interpolation procedure described in Sec. 2.3. Once the $\{M\}$ set is known, the interpolation is relatively fast, and the results, $M(\vec{p})$, can be used to calculate the χ^2. Note that none of the above precludes M from depending on other variables, such as time or space, and hence the generated $M_{\gamma,\delta,\ldots,\phi}$ could just as well be written $M_{\gamma,\delta,\ldots,\phi}(t, \vec{x})$, so the generated functions may very well, themselves, be multidimensional. Figure 1 presents schematically the fitting procedure with these modifications.

The number of functions in the set $\{M\}$ depends on each analysis case and should depend on the behaviour of the function $M(\vec{p})$ for a given parameter p_i. One should note that using too few grid points will give only a weak expression of the function's behaviour, whereas using a too fine division will, for small parameter changes, falsify the gradient calculations due to the statistical MC fluctuations. Properly chosen distances between grid points eliminate the statistical fluctuations of the theoretical function because the values between grid points are interpolated.

2.2. Description of the χ^2 calculation

We choose MC statistics on average about ten times greater than the statistical uncertainty in experimental data (less than that and we are insensitive to our parameters while fitting; more we use more MC calculation time for essentially no gain in sensitivity). Therefore, we neglect the statistical errors connected with the MC and use the classical χ^2 definition where the fits of the analytical functions are applied. Very sophisticated definitions of χ^2, including MC statistical fluctuations, are presented in Ref. [5], however, they are most useful in the case where experiment and simulation have similar statistics.

It is possible to use many sets of data from different experimental conditions, provided they can all be modeled by the same $\vec{p}$ parameters, and define the total χ^2 as the sum of the individual χ_k^2 calculated separately for a single data set k. Thus, the total χ^2, calculated when we perform simultaneously fits to m sets of data, is:

$$\chi^2 = m \sum_k w_k \cdot \chi_k^2, \tag{1}$$

where k is the histogram index running over a single set of data, and w_k is the corresponding weight.

The weights, w_k are calculated as a count ratio in each histogram relative to the total counts in all histograms, such that histograms with more counts give a greater share in the total χ^2:

$$w_k = \frac{\sum\limits_i N^i(k)}{\sum\limits_{l=1}^{m} \sum\limits_i N^i(l)}, \tag{2}$$

where N^i is the number of events in channel i of the experimental spectrum.

The partial χ^2_k is calculated as:

$$\chi^2_k = \sum_i \frac{\left[c_k \cdot M^i(k) - N^i(k)\right]^2}{N^i(k)}, \tag{3}$$

where c_k is a factor matching the k^{th} experimental N^i with its corresponding MC histograms M^i and is given by:

$$c_k = \frac{\sum\limits_i N^i(k)}{\sum\limits_i M^i(k)}. \tag{4}$$

2.3. The interpolation method

The grids are generated only for a finite and discrete set of parameters for all permutations of the parameters. However, as follows from the minimization procedure, theoretical functions are necessary from a continuous parameter space $(p_1, p_2, \ldots, p_n)$ and the interpolation procedure using the grids is applied to generate such functions. A well known method called 'the cubic spline interpolation' is used. In the presented analysis, the second derivatives for the first and last values of M'' are set to zero, the so-called natural cubic spline interpolation. This method has ability to reproduce the M function at any $[p_1, p_2]$ inside the four neighboring points (for two-dimensional interpolation). In our method we had to perform a multidimensional interpolation, which was reduced to series of one dimensional interpolations. Each one-dimensional interpolation is well defined.

Detailed description of the used method can be found, for example, in [6, 7].

3. Application of the method

3.1. Description of the experiment

In this section we apply our analysis method to the data obtained in the E742 experiment performed at TRIUMF, Vancouver (Canada). The experiment was dedicated to the study of μ-atomic processes occurring in solid hydrogen isotopes. The aim of the measurement was:

- reconstruction (theoretically calculated [8]) energy dependence of the elastic scattering cross-sections for muonic atoms in the process: $d\mu + p \rightarrow d\mu + p$ on crystalline hydrogen at a temperature of 3 K,
- determination $pd\mu$ molecule formation rate and rate of fusion in this molecule.

To analyze the experiment, we used MC simulations because other processes are competing with the scattering and $pd\mu$ molecule formation, as well as the complications coming from the geometry of the experiment. Figure 2 presents the diagram of the simulation of the processes taken into account [3]. The initial time is given when a muon enters the apparatus. The outputs of the simulation are:

- a X–ray time spectrum (see 'processes in external layers'),
- spectra of γ and converted μ (see 'muonic processes'),

which can be directly compared with the experimentally measured ones.

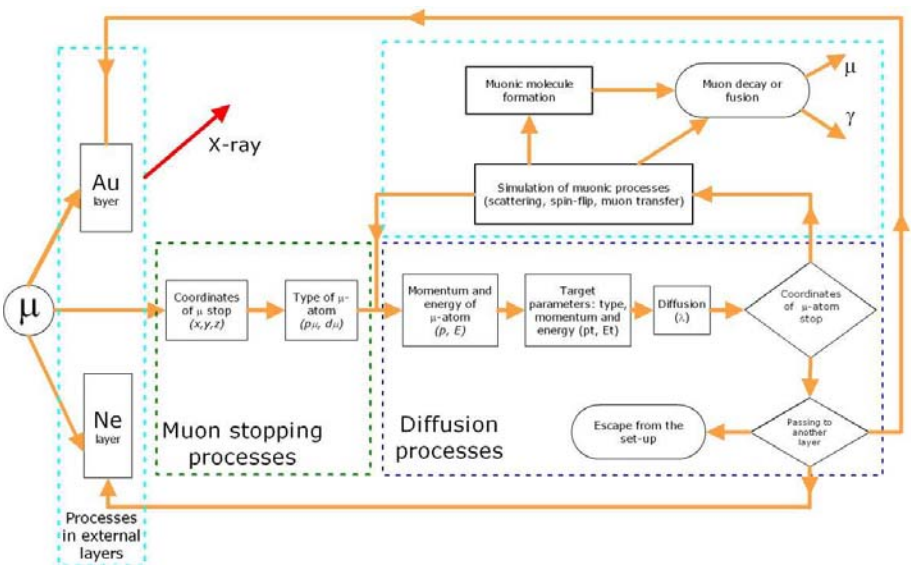

Fig. 2. Diagram of FOW program used to simulate the E742 experiment. Four process blocks are shown, namely the processes in external layers, the muon stopping, the diffusion, and the muonic processes. The output of the simulation: X-ray, μ and γ is visible

Detailed information about the experiment can be found in Ref. [9] and references therein.

3.2. Grid construction

One-dimensional grid. This grid concerns the formation rate of the $pd\mu$ molecule, the formation rate was rescaled by scaling factor in range 0.0–2.0 with step equal to 0.2. The grid size is 11 and $\vec{p} = r$ [10].

Two-dimensional grid. We assumed only that the general shape of the cross-section curve was valid. We vary only the depth and the position of the minimum of the R–T region. The theoretical dependence of the cross-section from the collision energy was

parametrized using parameter vector defined as: $\vec{p} = (\Delta E, s)$ (energy shift ΔE and a scaling factor s).

The grid size is $\gamma \times \delta = 11 \times 11$. The shift values, $p_1 = \Delta E$, were defined as 11 points between -0.5 eV to 0.5 eV, in steps of 0.1 eV. The curves for the depth parametrization have a rescaled minimum cross-section $\sigma(E_R) = s \cdot \sigma(E_R)$. The rescaling took place in the energy range 0.5–6 eV, with 11 values of the rescaling parameter $p_2 = s$: $s \in \{0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.25, 1.4, 1.7, 2.0\}$.

The parametrisation example can be seen in Fig. 3a.

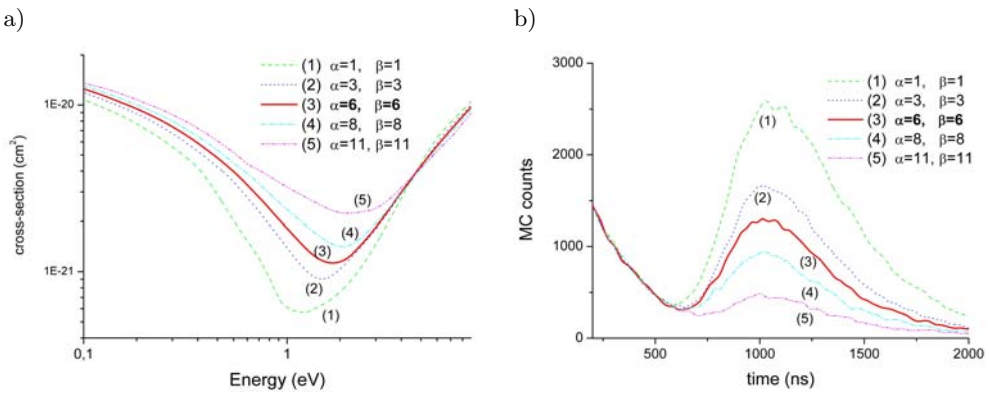

Fig. 3. An example of two-dimensional parametrization of the R-T minimum depth (a). In this case the parameter vector $\vec{p} = (\Delta E, s)$. Some values of ΔE and s, which concern the minimum cross section $\sigma_{\Delta E,s}(E_R)$ are indicated. An example of the grid – $M_{\alpha,\beta}(\Delta E, s)$ function, for the parametrisation given in the left, for chosen and indicated values of α and β (b)

To preserve the smooth form of the cross sections, values within the 0.5–6 eV range were not globally scaled by s, but were scaled by a factor which ranged from 1 at the borders, to s at the energy of the R–T minimum. The s values (except the minimum value) were selected numerically to reproduce the characteristic shape of the cross-section function inside the R–T region.

An example of grid obtained for such a parametrisation is presented in Fig. 3b.

In the presented simulations it was assumed that good events should be about 50–100 thousand (then, it is possible to accept that statistical error of MC spectrum is considerably smaller than in the case of experimental spectra). This necessitated to simulate for one single analysis (one point on the "grid") very large number of events (we simulated 5 million events). It should be mentioned, that interesting events have small probability of incident, for example: the probability of $d\mu$ escaping from US layer is order of 0.05 and the probability of reaching the neon layer (on DS) is ca. 0.01 and the $pd\mu$ formation probality is ca. 0.07. The computational time necessary for realization of such a single simulation was equal to ca. 2h of the CPU on SGI2800 computer, for programme in Fortran 77, in scalar version, optimized under calculations speed.

For the grid with size 11×11 the time of generation would take theoretically ca. two weeks of the CPU time, but simultaneous run of the 3 jobs was sometime possible, so practically real time of one grid generation was ca. one week. For the final analysis, it was necessary to have generated six such grids (three for hydrogen-deuterium mixture and three for hydrogen-tritium mixture). Additionally, it was necessary to generate some test grids in order to choose the simulation statistics and the number and values of grid points.

3.3. The results

An example of the experimental data fits using the MC time spectra (see the example in Fig. 3) is presented in Figure 4 (example for one-parameter fitting) and in Figure 5 (example for two-parameter fitting).

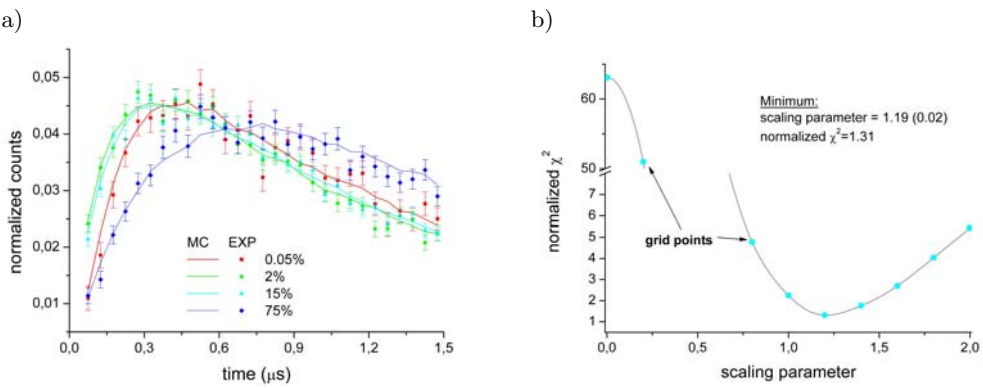

Fig. 4. Example of a one-dimensional fit (for one set of data) (a). Example of a χ^2 scan, the grid points are marked by blue circles (b)

For the presented examples we obtained the following results:

- One parameter fitting: fusion rates in $pd\mu$ molecule (for details see [10]). The formation rate were scaled and the obtained value is $6.7 \cdot 10^6$ s^{-1} (scaling factor equal 1.19) (see Fig. 4). The results are consistent (within the error range) with the simplified analytical calculation.
- Two parameter firring: $\Delta E = 0.19 \pm 0.03$ eV and $s = 1.09 \pm 0.03$.

The errors of the fitting are connected with low experimental statistics, some background problems, and also with the grid steps (0.10 eV for ΔE and 0.1–0.2 for s) and finally with the interpolation procedure.

4. Conclusion

The method allowed us to perform a correct comparison of experimental data with theoretical predictions based on MC calculations. Although the experimental data

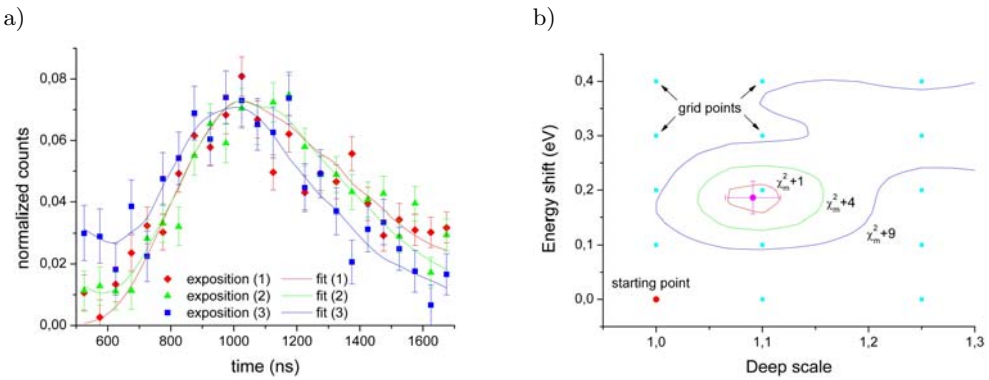

Fig. 5. Example of a two-parameter fit. In this case three sets of experimental data (called expositions 1-3) were fitted simultaneously (the scattering cross-section does not depend on experimental conditions and fitting spectra obtained at three different conditions with the same parameters should reduce systematics) (a). Example of a contour (b)

was obtained only in a few weeks of muon beam usage, the grid construction was a time-consuming step requiring more than six months of calculation (SGI2800 in ACK Krakow). The fitting procedure was quick (on PC computer one single fitting took ca. half an hour) and allowed us to prepare many fits for any combination of the data and perform a more complex analysis of the data themselves, e.g., χ^2-contour and error calculations.

To establish a more precise set of mathematical rules, which test the correctness of this method, one needs to perform further studies but such was not the aim of this work. This paper do not present a formal mathematical proof of the method, however the procedure based of the method was able to give us useful results. The grid method, as demonstrated here, is fully acceptable for analyzing systems with a complex multiparameter dependences and we believe that can be very useful for such class of data analysis. Moreover, our procedure was internally checked by comparing results of fitting single and summed data, and in some cases, comparison was performed with analytical calculations.

As a further step we plan to perform analysis of the errors of this procedure and determine errors connected e.g. with: limited number of the grid points, distance between the grid points or number of simulated events. Test of some modification of classical χ^2 estimator will be also performed.

Acknowledgements

The authors thank ACK CYFRONET AGH in Krakow for allowing them to use their supercomputers.

This work was supported by grant WPiE/AGH No. 10.10.210.52/7.

References

[1] James F., Roos M.: *MINUIT – system for function minimization and analysis of the parameter errors and correlations*. Comp. Phys. Comm. 10 (1975), 343

[2] Zech G., *Comparing statistical data to monte carlo simulation – parameter fitting and unfolding*. DESY, 95–113

[3] Woźniak J., Bystritsky V. M., Jacot-Guillarmod R., Mulhauser F.: *Study of muonic hydrogen transport in TRIUMF Experiment 742 by the Monte Carlo method*. Hyp. Interact. 101/102 (1996), 573–582

[4] Application Software Group, *HBOOK, Reference Manual*. CERN, Program Library Long Writeup Y250 (1994)

[5] Kortner O., Zupancic C.: *Fitting simulated random events to experimental histograms by means of parametric models*. Nucl. Instrum. Methods A 503 (2003) 625–648 [Erratum in Nucl. Instr. and Meth. A 515 (2003), 894]

[6] Filipowicz M., Bystritsky V. M., Knowles P. E., Mulhauser F., Woźniak J.: *Method of Monte-Carlo grid for data analysis*. Nucl. Instr. Meth. A547(2005), 652–662

[7] Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P.: *Numrical Recipes in Fortran: The Art of Scientific Computing*. Cambridge University Press, New York, 1992

[8] Chiccoli C. *et. al.*: *The atlas of the cross sections of mesic atomic processes III. The process* $p\mu + (d,t)$, $d\mu + (p,t)$ *and* $t\mu + (p,d)$. Muon Catal. Fusion 7 (1992), 87–153

[9] Mulhauser F. *et. al.*: *Ramsauer-Townsend effect in solid hydrogen*. Hyp. Interact. 138 (2001), 41–46

[10] Filipowicz M., Bystritsky V. M., Gerasimov V. V., Woźniak J.: *Kinetics of muon catalyzed fusion processes in solid H/D mixture*. Eur. Phys. J. D 47 (2008), 157–170