

Application of Bayesian Reasoning and the Maximum Entropy Method to Some Reconstruction Problems

K.W. FORNALSKI^a, G. PARZYCH^a, M. PYŁAK^a, D. SATUŁA^b

AND L. DOBRZYŃSKI^{a,b,*}

^aThe Andrzej Sołtan Institute for Nuclear Studies, 05-400 Otwock-Świerk, Poland

^bFaculty of Physics, University of Białystok, Lipowa 41, 15-424 Białystok, Poland

(Received January 29, 2010; in final form March 9, 2010)

The so-called Bayesian reasoning is applied whenever uncertainty has to be considered as serious factor in interpretation of results. The paper presents analysis of the impact of inaccurate data on the straight line and quadratic relation fittings. This type of analysis is particularly important when one tries to decide on the type of dependence. The paper also shows examples of the Maximum Entropy Methods applied to the reconstruction of the hyperfine parameters distribution from the measured Mössbauer spectra of GaFeO₃ and the electron–positron momentum distribution from the positron annihilation data of Gd.

PACS numbers: 29.85.–c, 78.70.Bj, 76.80.+y

1. Introduction

Experimental data always bear uncertainties. The researcher needs usually either to fit certain function to the measured points or to reconstruct certain distribution which is hidden in the experiment. This happens *e.g.* when one measures a spectrum which usually has smeared details due to the finite resolution of the spectrometer. In the conventional analysis, least-square method is most commonly used. Description of the method can be found in every academic textbooks on data analysis. Moreover, method of curve fitting, is susceptible to ‘outliers’ — the points that accidentally are having values significantly different from other measured points. When such points appear in a single spectrum, it is not particularly difficult to qualify them as outliers and eventually neglect them in the analysis. Much more difficult case is when one collects the data obtained in different laboratories or measured by different techniques. When the scatter of points is substantially larger than their claimed accuracies would allow, the parameters of a curve fitted to the measured points may have very low credibility. In the more difficult case of a reconstruction of a distribution observed through a spectrum smeared by resolution function of measuring instrument, one meets in fact two problems: one is the most accurate deconvolution of the spectrum, the other — carrying out non-parametric model-free reconstruction of the distribution. This is typical inverse problem which is generally difficult to solve. The Maximum Entropy Method is often very useful in solving such problems. In both cases, excel-

lent introduction to the Bayesian and Maximum Entropy methods can be found in [1]. Their bases would only be briefly described in the next sections where it would be necessary.

2. “Outliers” and curve fitting

A good example of the problem was considered by Fornalski and Dobrzyński [2] who analyzed the epidemiological data obtained for the cancer deaths of the nuclear industry workers. The scatter of points presented in original papers [3], that were intended to show that the risk of cancer death increases with the dose absorbed by workers, by far exceeded declared experimental uncertainties. These points exhibited also uneven distribution of points along abscissa axis. In such situation one could immediately say that conventional least-square fitting must lead to the parameters of the fit that may not be reliable. It could be suspected that the conclusions reached in papers [3] were due to a few “outliers” in the data. In fact, in the data considered one cannot directly say which points should be treated as outliers just because of the scatter of the data. Therefore, instead of removing suspected points, the authors of [2] used the method [1], of carrying out the fit under assumption that the values of uncertainties are declared for the points lowest estimate of the true ones.

In accordance to [1], for given uncertainty σ_0 , one assumes that this value is showing rather the lowest estimate, and a probability density that the “true” uncertainty σ is given by so-called Jeffrey’s prior:

$$p(\sigma) = \frac{1}{\ln(\sigma_{\max}/\sigma_{\min})} \frac{1}{\sigma}, \quad (2.1)$$

where σ_{\min} can be set to σ_0 , and σ_{\max} is chosen arbitrar-

* corresponding author; e-mail: ludwik@ipj.gov.pl

ily, or by simpler formula:

$$p(\sigma) = \frac{\sigma_0}{\sigma^2}, \tag{2.2}$$

where σ lies in the limits $[\sigma_0, \infty)$.

The Bayes theorem for conditional probability of an event X given Y and some general information I , which we could have about the studied object or distribution:

$$p(X|Y, I) = \frac{p(Y|X, I)p(X|I)}{p(Y|I)}, \tag{2.3}$$

where the probability $p(X|I)$ should be understood as showing the state of knowledge (“degree-of-belief”) about X in light of the information I . The expression in the nominator is a product of likelihood function (how well the data Y reflect correctness of hypothesis X) and the prior (what do we know about X expressed in terms of a probability value between 0 and 1). The former can be calculated within usual theory of likelihood.

For the case we are interested in, the conditional probability density of having experimental datum E , given its uncertainty σ and expected theoretical value T , is given by:

$$P(E|\sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp[-(T - E)^2/(2\sigma^2)] p(\sigma), \tag{2.4}$$

where the information I was dropped for better clarity of the formula. The probability (for all N experimental points) should be maximized for a number of parameters $\{\alpha_n, n = 1, 2, \dots, M\}$ of expected function (T), which describes the data. In order to find the maximum of:

$$P = \prod_{i=1}^N P_i = \prod_{i=1}^N \int \frac{1}{\sqrt{2\pi}\sigma_i} \times \exp(-(T_i - E_i)^2/(2\sigma_i^2)) p(\sigma_i) d\sigma_i, \tag{2.5}$$

it is convenient to maximize the logarithm $L = \sum_{i=1}^N \ln P_i$ with respect to any of the of parameter fit α_n :

$$\frac{dL}{d\alpha_n} = \sum_{i=1}^N -(T_i - E_i) \frac{dT_i}{d\alpha_n} \frac{1}{P_i} \times \int \frac{1}{\sqrt{2\pi}\sigma_i^3} \exp(-(T_i - E_i)^2/(2\sigma_i^2)) p(\sigma_i) d\sigma_i, \tag{2.6}$$

which can be written in shorter form:

$$\frac{dL}{d\alpha_n} = \sum_{i=1}^N g_i (T_i - E_i) \frac{dT_i}{d\alpha_n}, \tag{2.7}$$

where N is the number of all analyzed data points, and the derivative (2.7) should thus be set to zero. The integral’s range in (2.5) and (2.6) for the prior (2.1) is $[\sigma_{\min}, \sigma_{\max}]$ and for (2.2) — $[\sigma_0, \infty)$. For example, with the prior (2.2), after integrating (2.4) in the above limits, the following expression is obtained for the probability P_i :

$$P_i = \frac{\sigma_{0i}}{(T_i - E_i)^2 \sqrt{2\pi}} \times \left[1 - \exp(-(T_i - E_i)^2/(2\sigma_{0i}^2)) \right]. \tag{2.8}$$

The set of as many equations of the form (2.7) as the number M of parameters α_n has to be solved numerically.

From purely technical point of view, this corresponds to changing the weights $1/\sigma_i^2$ of any i -th point in classical least-square routine to a new value g_i :

$$g_i = \frac{1}{(T_i - E_i)^2} \times \left(2 - \frac{(T_i - E_i)^2}{\sigma_{0i}^2} \frac{1}{\exp((T_i - E_i)^2/2\sigma_{0i}^2) - 1} \right). \tag{2.9}$$

Let’s note that the weights g_i contain the parameters of interest which appear in theoretical functions T_i .

In the case of straight line fitting one deals with $n = 2$ parameters, so $T_i = aD_i + b$, where D_i denotes the experimental datum. This problem was considered in the text-book by Sivia and Skilling [1]. It was shown there that a few ‘outliers’ are effectively not influencing parameters of searched dependence while direct least-square fitting results in producing serious errors of these parameters. It was shown in the paper [2], that such a comfortable situation is encountered only when the number of “outliers” is not too large and not deviating from the proper line (trend) in one direction only as may happen *e.g.* in measurements of low count rates.

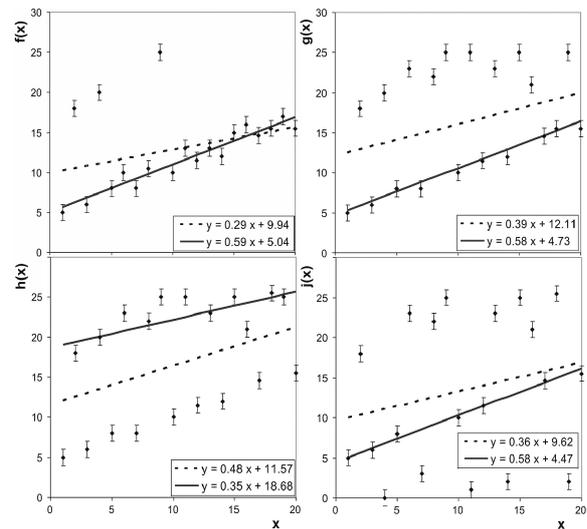


Fig. 1. Four examples of classical χ^2 fitting (dotted lines) and the Bayesian one (solid lines). All points are simulated. The main trend was assumed to be $y = 0.6x + 4.76$.

An example of the straight line fitting is shown in Fig. 1. In four parts of the figure one can see simulated experimental points scattered around assumed original trend $y_i = 0.6x_i + 4.76$, plus a number of points that intentionally deviate markedly from the assumed trend. While the dotted lines present results of the classical least-squares fitting, the solid lines are obtained by means of the Bayesian approach. To show the quality of agreement with the original trend, it is convenient to calculate the value of the following parameter S_A :

$$S_A = \sum_{i=1}^n |y_{d,i} - y_{t,i}|, \tag{2.10}$$

where $y_{d,i}$ denotes the value of the i -th point which follow

from the parameters of the fit, and $y_{t,i}$ its “ideal” value (here $y_{t,i} = 0.6x_i + 4.76$). It is obvious that better reconstruction of the trend must have lower parameter S_A .

TABLE I

Fitted parameters and parameters S_A for both cases presented in Figs. 1 and 2. The least squares (χ^2) and Bayesian methods were used.

Case	Linear (Fig. 1)		Quadratic (Fig. 2)	
	χ^2 method	Bayes	χ^2 method	Bayes
f(x)	$a = 0.29 \pm 0.11$ $b = 9.94 \pm 1.2$ $S_A = 42.6$	$a = 0.59 \pm 0.07$ $b = 5.04 \pm 1.1$ $S_A = 4.1$	$a = 0.17 \pm 0.05$ $b = -1.79 \pm 0.62$ $c = 14.77 \pm 1.6$ $S_A = 28.5$	$a = 0.19 \pm 0.03$ $b = -2.19 \pm 0.52$ $c = 15.21 \pm 1.3$ $S_A = 10.7$
g(x)	$a = 0.39 \pm 0.15$ $b = 12.11 \pm 1.7$ $S_A = 103.1$	$a = 0.58 \pm 0.12$ $b = 4.73 \pm 1.2$ $S_A = 4.0$	$a = 0.14 \pm 0.02$ $b = -1.06 \pm 0.13$ $c = 12.66 \pm 2.1$ $S_A = 57.1$	$a = 0.16 \pm 0.01$ $b = -1.44 \pm 0.09$ $c = 10.87 \pm 1.4$ $S_A = 20.2$
h(x)	$a = 0.48 \pm 0.15$ $b = 11.57 \pm 1.7$ $S_A = 111.6$	$a = 0.35 \pm 0.15$ $b = 18.68 \pm 1.4$ $S_A = 225.9$	$a = 0.07 \pm 0.02$ $b = -0.18 \pm 0.34$ $c = 13.24 \pm 1.3$ $S_A = 89.6$	$a = 0.05 \pm 0.02$ $b = 0.66 \pm 0.27$ $c = 10.60 \pm 1.1$ $S_A = 124.2$
j(x)	$a = 0.36 \pm 0.2$ $b = 9.62 \pm 2.4$ $S_A = 47.2$	$a = 0.58 \pm 0.14$ $b = 4.47 \pm 1.4$ $S_A = 9.2$	$a = 0.12 \pm 0.01$ $b = -1.12 \pm 0.24$ $c = 14.74 \pm 1.7$ $S_A = 56.3$	$a = 0.16 \pm 0.01$ $b = -1.52 \pm 0.14$ $c = 10.75 \pm 0.96$ $S_A = 23.9$

The first function (see Fig. 1) $f(x)$ shows the results of fitting when only three apparent outliers are present. It is a common example of typical experimental event with some outliers. In such a case it would be justified to reject those three points and concentrate on remaining points only. It is shown that the Bayesian approach needs not make such rejection in order to arrive at proper result, what is in good agreement with the results presented in [1]. The $g(x)$ and remaining functions contain much more outliers (such data could come *e.g.* from different laboratories). It is seen that in the first two cases the Bayesian approach results in the fitted line which is very close to the “true” one. In the third example $h(x)$ the number of points which are away from the expected trend is larger than the number of “correct” points. In this particular case one can hardly differentiate between “correct” points and “outliers”. This is also seen in the result of the Bayesian analysis, see parameter S_A in Table I, which favors “outliers”. However, one can also see that in such case the parameters S_A obtained within the maximum likelihood approach and the Bayesian one are not as different as in the other cases. In the last case $j(x)$ although the number of “outliers” is high, they lie below and above the trend. One can see, that from many options of drawing single line, the Bayesian analysis selects

the points that most likely lie on a straight line or are closest to such line. The last three cases are very seldom in practice. In fact, such data with large scatter appear in reality. The authors met such situation when they analyzed epidemiological data on the cancer mortality among nuclear workers [2].

In order to be sure of the usefulness of presented Bayesian reasoning, one should check how such a procedure works in case of more complicated fitted function, *e.g.* quadratic one as the next simplest case. In this case it is necessary to take three fitting parameters, so the function factor is $T = aD^2 + bD + c$. The results of fit are shown in Fig. 2, where the original trend was put as $y = 0.19x^2 - 2.24x + 15.2$. Again, two types of analyses (least-squares — the dotted lines, and Bayesian one — solid lines) are applied to the simulated data. All four cases, $f(x)$, $g(x)$, $h(x)$ and $j(x)$ are intentionally very similar to the previous linear case. The general results indicate again advantages of the Bayesian approach. All the results are summarized in Table I.

The Bayesian analysis can be used in many more situations. For example, one can assess relative reliability of two alternative models of fitted functions, *e.g.* one which assumes a linear dependence (*e.g.* voltage *vs.* electric current), and the other one which assumes quadratic or more

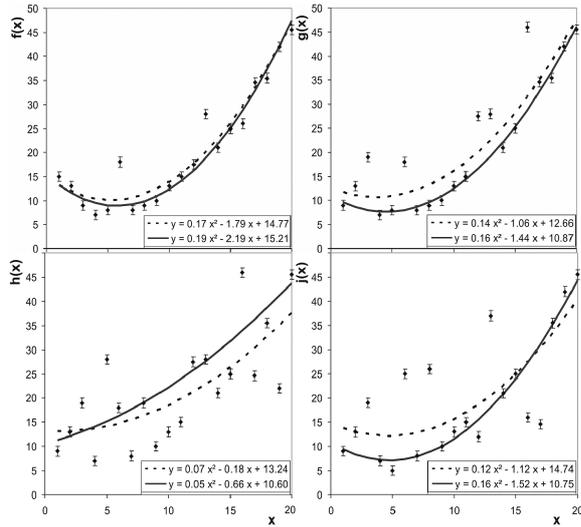


Fig. 2. Four examples of classical χ^2 fitting (dotted lines) and the Bayesian one (solid lines). All points are simulated ones. They should reflect quadratic function $y = 0.19x^2 - 2.24x + 15.2$.

complicated dependences. Generally speaking, it is necessary to find the posterior (confidence) ratio for both models and multiply it by so called ‘‘Ockham’s factor’’, which prevent using of over-complicated model. As it was shown in [1] and [2], if one model A has no parameter to fit, while a model B contains one parameter λ than their relative value can be calculated as follows. If λ turned out to appear with the uncertainty $\delta\lambda$, while prior to the fitting we knew only that it must be contained within the $(\lambda_{\min}, \lambda_{\max})$ limits, the relative superiority of one model A with respect to B is:

$$W_m = \frac{P(A|D, I)}{P(B|D, I)} = \frac{P(A|I)}{P(B|I)} \frac{P(D|A, I)}{P(D|\lambda_0, B, I)} \times \frac{\lambda_{\max} - \lambda_{\min}}{\delta\lambda\sqrt{2\pi}}, \quad (2.11)$$

where the first term describes our evaluation of the relative importance of the models before doing the experiment, the second term is a ratio of likelihood functions, and the last term describes the ‘‘Ockham’s factor’’.

3. Maximum Entropy Method: application to the results of Mössbauer spectroscopy and 2D Angular Correlation of Annihilation Radiation (ACAR)

The Maximum Entropy Method is used to construct the prior in the form of $\exp(\alpha S)$, where α is a parameter, and S is information entropy which is used often in the form of cross-entropy, *i.e.*:

$$S = - \sum_i p_i \log(p_i/m_i), \quad (3.1)$$

where p_i and m_i describe searched and prior distributions, respectively. In the simplest case one can use un-

informative (uniform) prior, *i.e.* $m_i = \text{constant}$. This, however, is known to cause some unwanted features as it is presented *e.g.* in the case of charge-, spin- or electron momentum-density distributions in solids [4].

The Mössbauer spectrum is described by many hyperfine parameters (magnetic hyperfine field B , isomer shift (IS), quadrupole splitting (QS)) that can have their own distributions. In addition, the relative intensities of spectroscopic lines may be determined by possible sample texture and sample thickness effects. When these last two effects are neglected one can use so-called thin-absorber approximation. In the simplest case one gets six Lorentzian lines in the measured spectra appearing at the source velocities:

$$v_1 = B^*(3g_{3/2} - g_{1/2})/2 + QS + IS,$$

$$v_2 = B^*(g_{3/2} - g_{1/2})/2 - QS + IS,$$

$$v_3 = B^*(-g_{3/2} - g_{1/2})/2 - QS + IS,$$

$$v_4 = B^*(g_{3/2} + g_{1/2})/2 - QS + IS,$$

$$v_5 = B^*(-g_{3/2} + g_{1/2})/2 - QS + IS,$$

$$v_6 = B^*(-3g_{3/2} + g_{1/2})/2 + QS + IS, \quad (3.2)$$

where the gyromagnetic factors for ^{57}Fe nucleus are: $g_{3/2} = -0.067897$ mm/s/T, and $g_{1/2} = 0.118821$ mm/s/T. The intensities of the lines are shown in Table II.

TABLE II
Line intensities in Zeeman sextets in the case of unpolarized as well as circularly polarized photons.

Line number	Unpolarized radiation*	Circularly polarized radiation*
1	$I_1 = 3(1 + c_2)/16$	$I_1 = 3(1 + c_2 + 2c_1)/16$
2	$I_2 = (1 - c_2)/4$	$I_2 = (1 - c_2)/4$
3	$I_3 = (1 + c_2)/16 = I_1/3$	$I_3 = (1 + c_2 - 2c_1)/16$
4	$I_4 = I_3$	$I_4 = I_1/3$
5	$I_5 = I_2$	$I_5 = I_2$
6	$I_6 = I_1$	$I_6 = 3I_3$

* — where $c_1 = \langle \cos \Theta \rangle$ and $c_2 = \langle \cos^2 \Theta \rangle$ with Θ being an angle between the direction of a photon and direction of the magnetization of the sample (the averaging runs over all possible grains and domains in the sample [5]).

A given i -th line contributes to the j -th channel in the velocity spectrum the intensity proportional to:

$$J_j = \frac{I_i}{(V(j) - v_i)^2 + (\Gamma/2)^2}, \quad (3.3)$$

where Γ denotes the natural width of the line from Mössbauer source. When the hyperfine parameters B , IS and QS have distributions described by a probability density $P(B, IS, QS)$ the Eq. (3.3) modifies to:

$$J_j = \iiint P(B, IS, QS) \times \frac{I_i}{(V(j) - v_i)^2 + (\Gamma/2)^2} dB dIS dQS, \quad (3.4)$$

where the range of integrals may always be reasonably chosen. In practice, the integral (3.4) is changed to a sum: the whole 3D space is divided into pixels and one is carrying out summation over the pixels seeking for values of P_i , where i denotes a linear parameter of a pixel. From the form of (3.4) it is seen that such a reconstruction is formidable task.

The measured spectra are usually composed of 256 points only, so any kind of least square method can be successfully applied only when the number of pixels is very low, and when one deals with the one dimensional distribution or one can assume certain dependencies between the hyperfine parameters, e.g. that the field B linearly depends on isomer shift and quadrupole splitting. In this situation Maximum Entropy Method is the method of choice. The Maximum Entropy Method can be effectively used for reconstruction of the hyperfine parameters distribution from the Mössbauer spectrum. This was shown in Refs. [6–8] on mimicked spectra. The MEM analysis of experimental case of the experimental spectrum measured for amorphous Fe-B alloy was first presented in the paper [9]. On the example of recently obtained data for GaFeO_3 it is shown in the present paper that three dimensional MEM analysis of the probability distributions of hyperfine parameters can be carried out and the results are shown in Figs. 3–6. The spectrum itself [10] exhibits quite broadened lines. The broadening can be due to the overlapping spectra arising from different lattice sites and possible different environment of the sites. This broadening does not allow to determine uniquely hyperfine field parameters that can characterize each of the sites occupied by iron. The reconstruction by MEM (see Figs. 3–6) shows how much the real distribution is collapsed when one wants to get information about hyperfine magnetic field (B), isomer shifts (IS), and quadrupole splitting (QS) connected with given crystallographic site. Nevertheless, knowing that MEM produces as much diffuse maps as possible. Obtained reconstruction gives certain limits for hyperfine field values and may still be used as a starting point to the description of the measured spectra as a sum of spectra connected with given crystallographic sites. In this case, however, it is not possible to avoid a number of assumptions that may turn out to be invalid. The distributions obtained by MEM are directly making full use of the experimental data and show what we really know without making reference to any model.

It can be noticed that in spite of certain noise present in the figures (that appears mainly at the borders of the analyzed region of hyperfine parameters), they imply that one can hardly speak about linear dependencies between hyperfine parameters. Such correlations are typically assumed in the conventional analysis of the Mössbauer spectra with hyperfine fields distributions. The distributions show that the spectra are most generally interpreted in terms of the hyperfine parameters distributions that can be correlated, but the way they are correlated is not described by a simple linear function. The

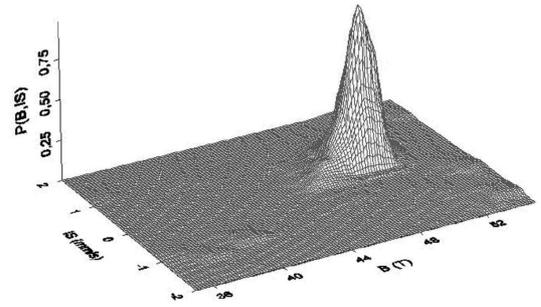


Fig. 3. The distribution of probabilities of hyperfine parameters in GaFeO_3 in B-IS plane from the measurements at 14 K as obtained by MEM. $P(B, IS)$ denotes the probability distribution obtained as a sum of the distributions for all quadrupole splittings.

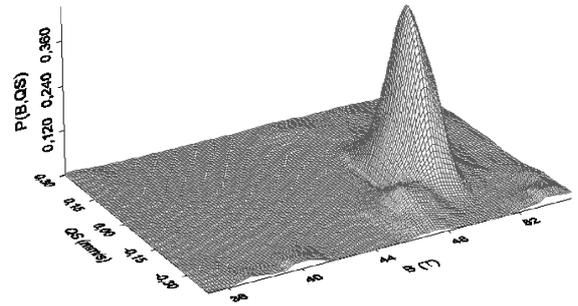


Fig. 4. Similar in B-QS plane at 14 K. The distribution is a sum of the distributions for all IS.

spectrum itself can also be analyzed by fitting four or five individual spectra to the measured spectrum, but, as it follows from the figures, the parameters fitted in this way can be put to doubts. However, if one wishes to do that, the MEM results offer at least the limits of the fit pa-

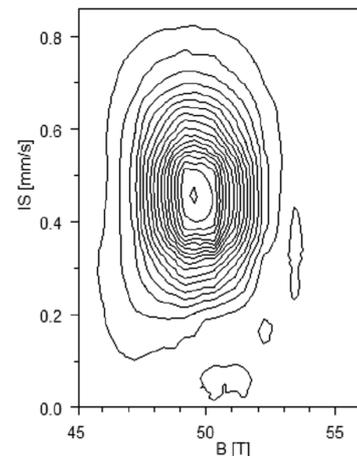


Fig. 5. Reconstruction (by MEM) of the distribution at $T = 14$ K in B-IS plane shown in the 2D contour plot.

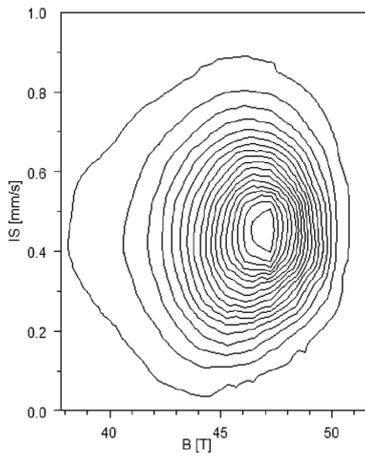


Fig. 6. Reconstruction (by MEM) of the distribution at $T = 100$ K in B-IS plane.

rameters. In the particular case of GaFeO_3 the spectrum quite rapidly broadens with increase of the temperature. The differences of the $P(B, IS)$ distributions at $T = 14$ K and $T = 100$ K are shown in 2D contour plots in the next two figures. Apparent broadening of the spectra at high temperature is clearly seen.

Another novel example of application of MEM to reconstruction of the electron–positron momentum distribution is given below. In the experiment are injected positrons into the material of interest. The positrons are thermalized in the matter and could annihilate with electrons producing two annihilation quanta. If the positron–electron pair had zero momentum, the quanta would be emitted at 180° with respect to each other. Because there is always certain momentum distribution $\rho(\mathbf{p})$ of such pair, one observes distribution of angles through which the annihilation quanta appear. The 2D ACAR data are described by:

$$N(p_x, p_y) = \int \rho(\mathbf{p}) dp_z, \quad (3.5)$$

where \mathbf{p} is momentum of annihilating positron–electron pair. The data are measured for various values of p_z along many directions in (p_x, p_y) -plane. Therefore, one can reconstruct the momentum distribution $\rho(p_x, p_y; p_z)$ in the plane perpendicular to p_z . The MEM analysis was carried out on the data for Gd, measured by R.L. Waspe and R.N. West [11], and deconvoluted (also by MEM [12]) by Fretwell et al. [13].

Maximization of entropy should be carried out under certain constraints $\Phi_k(\rho)$, which, added to Eq. (3.5), result in Lagrangian:

$$L = - \sum_i \rho_i \log(\rho_i/\rho_i^0) - \sum_k \lambda_k \Phi_k(\rho), \quad (3.6)$$

where k runs over the number of constraints and λ_k are Lagrange multipliers. The reconstructed density must also be normalized to certain value A , so $\Phi_0(\rho) = \sum_i \rho_i - A$. The constraints may be chosen so to ensure

that every experimental value E_k differs from the reconstructed one:

$$T_k = \sum_j r_{kj} \rho_j, \quad (3.7)$$

(r_{kj} — transformation matrix elements), by not more than the uncertainty σ_k assigned to the experimental point: $(T_k - E_k)^2 \leq \sigma_k^2$. However, instead of using many Lagrange multipliers one uses mainly only single constraint, namely, that a misfit function:

$$\chi^2 = \sum_k (T_k - E_k)^2 / \sigma^2 \leq \text{const}, \quad (3.8)$$

where const is chosen close to the number of experimental points. Finally, the following equation for density is obtained:

$$\rho_i = A \frac{\rho_i^0 \exp\left(-\frac{1}{2\alpha} \frac{\partial \chi^2}{\partial \rho_i}\right)}{\sum_i \rho_i^0 \exp\left(-\frac{1}{2\alpha} \frac{\partial \chi^2}{\partial \rho_i}\right)}, \quad (3.9)$$

where α denotes Lagrange multiplier appearing in the prior $\exp(\alpha S)$.

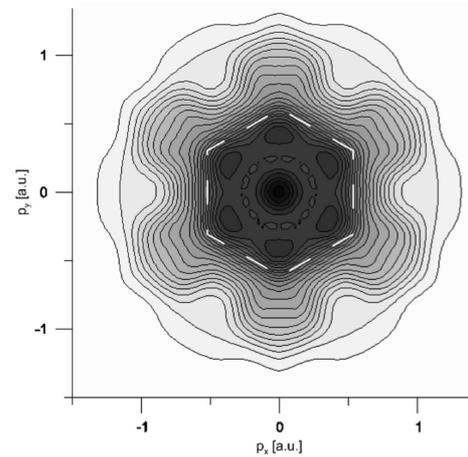


Fig. 7. Electron–positron momentum distribution in Gd for $p_z = 0$. The experimental data used were corrected for resolution [10]. The first Brillouin zone is marked by white dashed line.

Figure 7 shows the reconstruction of the positron–electron momentum distribution from 16 directions measured in $p_z = 0$ plane [12]. The first Brillouin zone (BZ) was also depicted (white line). Again, one notices very clear shape of this distribution and low level of noise. Moreover, one can observe an increase of density of isolines at the borders of BZ.

Good reconstruction of the electron–positron momentum density turns out to be obtained with much smaller number of directions. Figure 8 shows comparison of the results obtained on the grounds of experimental data from 6 projections (6 different directions from ΓM to ΓK) and 16 projections with 6-projection reconstruction treated as a prior. Only little difference between both panels can be observed for medium momentum range. General shape and character of acquired distribution

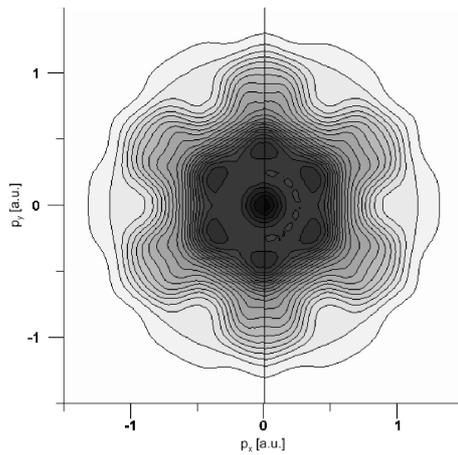


Fig. 8. Electron-positron momentum distribution in Gd for $p_z = 0$. Comparison of reconstructions from 6 (left panel) and 16 projections (right panel).

stayed untouched. Also position of the steepest slope at the BZ is not changed. The difference between these densities, relative to the result obtained from 16-projection is presented in Fig. 9. Only one bigger cusp can be observed, however, its position corresponds to rather low level of density and could not be noticed in Fig. 8, where this area is presented as a background.

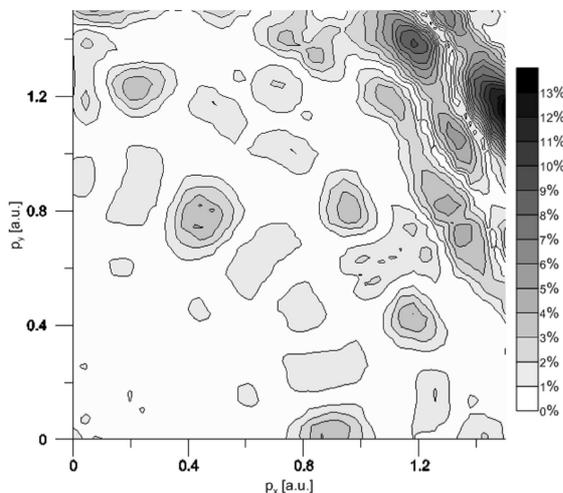


Fig. 9. Relative difference between 6- and 16-projection electron-positron momentum distribution in Gd for $p_z = 0$.

Above situation reveals how useful and powerful is the Maximum Entropy Method technique. Even incomplete set of experimental data (less than half of available projections) is sufficient for correct reconstruction. Naturally, too modest data may lead to improper results and some important details may be lost during reconstruction.

This kind of reconstructions can be used for determination of the shape of Fermi surfaces without the need of fitting a number of parameters when the distribution is represented as a series in symmetry-adapted harmonics, see [14–16].

4. Conclusions

The paper present a few novel applications of the Bayesian reasoning and the Maximum Entropy Method to true experimental problems. In all cases it is shown the powerfulness of these techniques, even in such difficult cases like 3D reconstruction of the hyperfine field parameters from a 1D Mössbauer spectrum.

Acknowledgments

We are very grateful to Prof. R.N. West for making available his experimental 2D ACAR spectra for Gd, to Dr A. Alam and Dr S. Dugdale for their deconvoluted data. We wish to express our gratitude to professors H. Fuess and K. Szymański for their agreement to publish the results of preliminary analysis of the data on GaFeO₃, and to Prof. G. Kontrym-Sznajd for very stimulating discussion of the case of Gd.

References

- [1] D.S. Sivia, J. Skilling, *Data Analysis. A Bayesian Tutorial*, Oxford University Press, 2006.
- [2] K.W. Fornalski, L. Dobrzyński, *Int. J. Low Rad.* **6**, 57 (2009).
- [3] M. Vrijheid, E. Cardis, M. Blettner, E. Gilbert, M. Hakama, C. Hill, G. Howe, J. Kaldor, C.R. Muirhead, M. Schubauer-Berigan, T. Yoshimura, Y.O. Ahn, P. Ashmore, A. Auvinen, J.M. Bae, H. Engels, G. Gulis, R.R. Habib, Y. Hosoda, J. Kurtinaitis, H. Malker, M. Moser, F. Rodriguez-Artalejo, A. Rogel, H. Tardy, M. Telle-Lamberton, I. Turai, M. Usel, K. Veress, *Radiat. Res.* **167**, 361 (2007); I. Thierry-Chef, M. Marshall, J.J. Fix, F. Bermann, E.S. Gilbert, C. Hacker, B. Heinmiller, W. Murray, M.S. Pearce, D. Utterback, K. Bernar, P. Deboodt, M. Eklof, B. Grieciene, K. Holan, H. Hyvonen, A. Kerekes, M.C. Lee, M. Moser, F. Pernicka, E. Cardis, *ibid.*, p. 380; E. Cardis, M. Vrijheid, M. Blettner, E. Gilbert, M. Hakama, C. Hill, G. Howe, J. Kaldor, C.R. Muirhead, M. Schubauer-Berigan, T. Yoshimura, F. Bermann, G. Cowper, J. Fix, C. Hacker, B. Heinmiller, M. Marshall, I. Thierry-Chef, D. Utterback, Y.O. Ahn, E. Amoros, P. Ashmore, A. Auvinen, J.M. Bae, J. Bernar, A. Biau, E. Combalot, P. Deboodt, A. Diez Sacristan, M. Eklöf, H. Engels, G. Engholm, G. Gulis, R.R. Habib, K. Holan, H. Hyvonen, A. Kerekes, J. Kurtinaitis, H. Malker, M. Martuzzi, A. Mastauskas, A. Monnet, M. Moser, M.S. Pearce, D.B. Richardson, F. Rodriguez-Artalejo, A. Rogel, H. Tardy, M. Telle-Lamberton, I. Turai, M. Usel, K. Veress, *ibid.*, p. 396.
- [4] L. Dobrzyński, in *X-ray Compton Scattering*, Eds. M.J. Cooper, P.E. Mijnarends, N. Shiotani, N. Sakai, A. Bansil, Oxford University Press 2004, p. 188.

- [5] K. Szymański, L. Dobrzyński, D. Satuła, B. Kalska-Szostko, in: *Material Research in Atomic Scale by Mössbauer Spectroscopy*, Eds. by M. Mashlan, M. Miglierini, P. Schaaf, Vol. 94, Kluwer Academic Publishers, Dordrecht 2003, p. 317.
- [6] R.A. Brand, G. Le, Caer, *Nucl. Instr. Methods Phys. Res.* **34**, 272 (1988).
- [7] L. Dobrzyński, K. Szymański, D. Satuła, *Nukleonika* **49 Suppl. 3**, S89 (2004).
- [8] L. Dobrzyński, A. Holas, D. Satuła, K. Szymański, in Proc. 26th *International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, Ed. A. Mohamad-Djafari, AIP Conf. Proc. **872**, 511 (2006).
- [9] P.M. Bentley, S.H. Kilcoyne, *J. Phys.; Condens. Matter* **18**, 7751 (2006).
- [10] M. Bakr, A. Senyshyn, H. Wang, G. Parzych, L. Dobrzyński, K. Szymański, H. Fuess, poster presented during XX *International School on Physics and Chemistry of Condensed Matter*, July 4–11, Białowieża, Poland (2009).
- [11] R.L. Waspe, R.N. West, *Positron Annihilation*, Eds. P.G. Coleman, S.C. Sharma, L.M. Diana, North-Holland Publishing Company, 1982 p. 328.
- [12] S.B. Dugdale, Ph.D. Thesis, University of Bristol, 1996 (unpublished).
- [13] H.M. Fretwell, S.B. Dugdale, M.A. Alam, M. Biasini, L. Hoffmann, A.A. Manuel, *Europhys. Lett.* **32**, 771 (1995).
- [14] A. Alam, R.L. Waspe, R.N. West, *Positron Annihilation*, Ed. L. Dorikens-Vanpraet, World-Scientific, Singapore 1988, p. 242.
- [15] G. Kontrym-Sznajd, M. Samsel-Czekała, *Appl. Phys. A* **70**, 89 (2000).
- [16] G. Kontrym-Sznajd, *Fizika Nizkikh Temperatur* **35**, 765 (2009); G. Kontrym-Sznajd *Low Temp. Phys.* **35**, 599 (2009).