THE VITAL IMPORTANCE OF PRIOR INFORMATION FOR THE DECOMPOSITION OF ION SCATTERING SPECTROSCOPY DATA

R. Fischer, W. von der Linden, and V. Dose
Max-Planck-Institut für Plasmaphysik,
POB 1533, D-85740 Garching, Germany
EURATOM Association

ABSTRACT. The ubiquitous spectroscopic problem of decomposing overlapping lines is solved for ion scattering spectroscopy employing Maximum Entropy. The chosen example of Pd adsorption on a Ru surface is particularly challenging because 13 partially overlapping isotopes contribute to the total scattering signal. Proper decomposition using appropriate prior information enables accurate coverage determination.

1. Introduction

Ion scattering spectroscopy (ISS) is among the most important and successful techniques to determine the nuclear mass composition of a surface and the geometric surface structure in real space on an atomic scale [1]. ISS data, like experimental data in physics quite generally, constitute the convolution of a broadening function with the desired physical quantity. Yet the inverse problem of deriving the underlying physics is in most cases not unique and often corrupted, even ruined, by noise [2]. A frequently adopted way out of this problems is given if we know a reliable model function for the solution in terms of a few parameters. These may then be determined by well established least squares fitting techniques. The prior knowledge, which we use in this procedure, is quite substantial. The solution must be of the model form; the reliability of this assumption may be subject to a χ^2 test. Reliable model functions are only rarely available, however, and may even obscure small effects present in the data. The way out of this dilemma is offered by Bayesian reasoning with Maximum Entropy (ME) prior information [3, 4, 5]. The ME concept imposed on the infinite manifold of solutions which agree after convolution with the broadening function with the experimental data within error bars chooses the single unique solution containing the least amount of information exceeding our prior knowledge. The incorporation of prior knowledge differs, however, radically from the model function approach. Prior knowledge in ME constitutes a *soft* constraint, which can always be overruled by the measured data. This paper exemplifies this procedure with ion scattering spectroscopy data from Pd adsorbed on Ru(001). Special attention is given to the problem of incorporating prior knowledge and the reliability of the ME solutions.

In ISS an approximately monoenergetic beam of ions, typically noble gas ions, is directed at the surface in some well-defined directions. The energy of the primary scattered ions is measured at a well-defined scattering angle. In the binary collision model the energy of the backscattered ions is determined by the scattering geometry, the masses of the incident ion and the backscattering surface atom, and the incident energy. The backscattered ion intensity yields the information of the quantitative composition of the surface. Relative coverages can be determined if the signals originating from different masses are well separated. Fitting procedures are necessary if masses are not resolved, and, especially, if the overlapping signals contain isotopes. The system Pd/Ru consists of 6 Pd (A=102-110) and 7 Ru (A=96-104) isotopes which overlap in their mass numbers A. Since no straight forward theory of the line shape for scattering of a noble gas ion off a single Pd isotope and a single Ru isotope is available, we proceed to a form-free reconstruction of these lines with the help of ME.

This brings us back to the deconvolution problem. While in general one attempts a signal reconstruction removing the smearing effect of a known broadening, decomposition of a spectrum into its constituents requires the determination of the unknown broadening. The latter comprises in the present case apparatus effects, all contributions from the binary collision event and from various possible effects like multiple scattering and/or inelastic mechanisms.

2. Formalism

First, we will specify the mathematical structure of the problem imposed by ion scattering experiments. The scattering of ions with energies ranging from 0.1 to 10 keV off surfaces can be described very well by binary collisions between the incident ion and surface atoms. The energy E_1 of an ion of mass m_1 and incident energy E_0 scattered through an angle ϑ by an atom of mass m_2 is given by [6]

$$\frac{E_1}{E_0} = \left(\frac{\cos\vartheta + \sqrt{m_2^2/m_1^2 - \sin^2\vartheta}}{1 + m_2/m_1} \right)^2 \tag{1}$$

An energy spectrum of the scattered ions bears the information of the surface composition since for a fixed scattering angle ϑ the energy depends only on the mass ratio m_2/m_1. The energy spectrum $F(E)$ can be treated as a sum of energy spectra $f(E)$ of the various scattering atoms each modeling the convolution of the apparatus function with the energy distribution of the backscattered ions.

$$F(E) = \sum_{\mu} a_\mu \sum_{\nu} b_\mu^\nu f_\mu(E + E_\mu^\nu) \tag{2}$$

The abundance of the atom μ is a_μ and the relative abundance of isotope ν of atom species μ at the surface is b_μ^ν. b_μ^ν is taken from isotope tables of the elements. The energy distributions of the ions backscattered off the isotopes of atom species μ are assumed to be identical in their form but shifted in energy by E_μ^ν according to their masses. The scattering signals f_μ are allowed to vary between different atom species μ. Such differences can arise from energy losses in course of internal excitation of the scattering partners which will of course depend on details of their electronic structure. f_μ in our procedure will also comprise apparatus broadening and the energy distribution of the incident noble gas ions.

Let G_l represent the measurement of F at energy E_l. If the data G_l are assumed independent and normally distributed with error σ_l then the likelihood function of the data

G_l, given the solution F_l is

$$P(G_l|F_l,\lambda) \propto \exp(-\frac{1}{2}\chi^2) \qquad \text{with} \qquad \chi^2 = \sum_l \frac{1}{\sigma_l^2}(G_l - F_l)^2. \tag{3}$$

What we need is, however, $P(F_l|G_l,\lambda)$, the posterior probability of a particular solution F_l in view of the measured data G_l, which we want to maximize. λ stands for a set of further parameters for example a_μ, the abundance of the atom μ, in (2) or the incident energy E_0. We choose the prior probability of F_l, $P(F_l|\lambda)$, to be entropic relative to a certain prior knowledge m

$$P(F_l|\lambda) \propto \exp(\alpha S) \tag{4}$$

with regularization parameter α and entropy S.

$$S = \sum_\mu \sum_l f_{\mu l} - m_{\mu l} - f_{\mu l} \log \frac{f_{\mu l}}{m_{\mu l}} \tag{5}$$

An entropic prior is appropriate for positive additive distributions such as the present ISS spectra. According to Bayes' theorem we have to maximize

$$P(F_l|G_l,\lambda) \propto \exp(-\frac{1}{2}\chi^2 + \alpha S) \tag{6}$$

α determines the competition between a reconstruction close to the data ($\alpha = 0$) and a reconstruction close to the default model m (α sufficiently large). A selfconsistent choice of α is obtained by marginalizing $P(\alpha|G_l) = \int DF \cdot P(F_l,\alpha|G_l)$ [7] or the evidence of the data $P(G_l|\lambda) = \int DF \cdot P(F_l,G_l|\lambda)$ [8] over the reconstruction F_l . These quantities are proportional to each other providing the prior probability of α is uninformative, i.e. constant. The result F_l will then agree with the experimental data G_l and simultaneously approximate the default model m as closely as possible, in other words the default model is updated by the smallest amount required by the data. It is clear that the default model should be chosen such that it incorporates as much prior information as possible. For example the requirement that only positive functions f_μ are physically acceptable requires that $m_\mu > 0$. If this constitutes all our prior knowledge then m_μ is taken a small positive number such that f_μ will tend to m_μ where the data do not provide different information. If the data are not of sufficient precision this will generally lead to “soft” solutions with corresponding large error bars. The reliability of the solution can be very poor as will be shown in the following sections.

3. The clean Ru surface

Experimental data G_l shown as error bars in fig. 1 were obtained by scattering of Ar^+ ions of nominally 5 keV energy by an angle of 165°. Experimental details have been published previously [6]. The continuous solid line marks the final ME data fit F_l. The dotted curves indicate the decomposition of the overall spectrum into seven single isotope scattering functions f_l of the Ru surface.

Initially the data reduction was performed with a flat uninformative default model. The result for the single isotope scattering signal $\vec{f}_{\rm Ru}$ resulting from this choice is displayed in

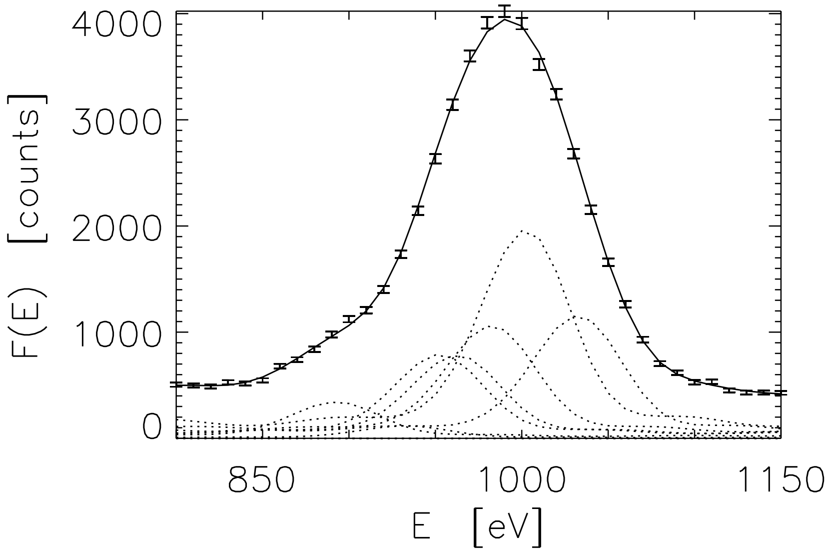

Figure 1: ISS spectrum of the clean Ru(001) surface (error bars), the ME reconstruction (solid line), and the decomposition into the 7 Ru isotopes (dotted lines). The incident energy of the Ar^+ ions $E_0 = 5.1$ eV is determined via evidence analysis. The Ru isotopes cause a broad nearly unstructured backscattering signal.

the upper panel of fig. 2. This is clearly unacceptable from our ISS experience in cases without isotope complication. It is, however, worth mentioning that the reconstruction of the overall signal from this unacceptable function fits the data in fig. 1 even closer than shown by the continuous solid line. The unacceptable shape of $\vec{f}_{\mathrm{Ru}}$ is the price we pay by pretending to be more ignorant then we actually are. Practically all data points provide information beyond our default model with the consequence of overfitting the data and little entropic smoothing. More precisely, the number of data which provides information differing from the default model approaches the number of measurements: $N \approx 2\alpha S$. Since the result of our calculation is not only the most probable function or more specifically vector $\vec{f}$ but rather its full probability distribution we can proceed and calculate errors of amplitude and position of the various structures in $\vec{f}$. They are given in fig. 2 and demonstrate that the disappointing ringing region of $\vec{f}$ is mostly insignificant.

From experience in image reconstruction it is well known that the quality of the ME solution depends progressively on the choice of default models as the noise level of the data increases [9]. We therefore proceed to the construction of a more refined default model. First of all we have obviously to account for a constant background on top of which the Ru scattering signal appears. The next point is more subtle. For the clean Ru surface equation (2) reduces to

$$F(E) = \sum_{\nu} b^{\nu} f(E + E^{\nu}). \qquad (7)$$

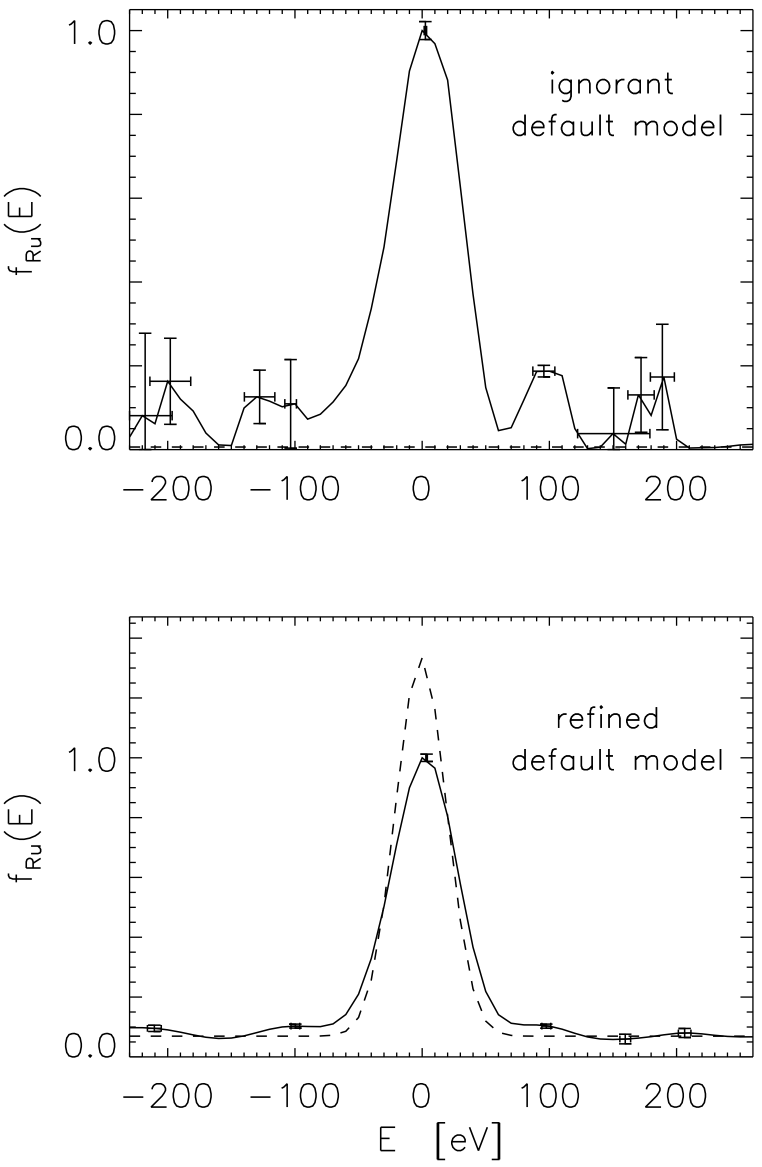

Figure 2: ME solutions (solid lines) for the broadening function resulting in the ISS spectrum shown in fig. 1 with different default models (dashed lines). In the upper panel the value of the default model m is small resulting in "ringing" with great errors in position and intensity of the structures. In the lower panel the default model is refined with a decomposition into a constant background and a Gaussian for the backscattering signal (see text).

We now take the first three moments

$$\begin{aligned} M^n(F) &= \int E^n F(E) dE \\ &= \sum_\nu b^\nu \int E^n f(E+E^\nu) dE, \quad n=0,1,2 \end{aligned} \tag{8}$$

of F and approximate the left-hand side by the corresponding moments computed from the experimental data

$$M^n(F) \approx M^n(G). \tag{9}$$

By change of variables and binomial expansion (8) transforms into a system of linear equations for the moments $M^n(f), n = 0,1,2$. Our ME solution should of course obey these "sum rules". But the moments tell even more. It can be shown by ME itself that the functional

$$\Phi = \int \left[f(x) - m(x) - f(x) \log \frac{f(x)}{m(x)} \right] dx \tag{10}$$

subject to the conditions that the first three moments are given numbers is maximum if $f(x)$ is a Gaussian [10]. This is of course quite satisfactory since experience from simple ion scattering spectra comply with this solution. Accordingly it is a reasonable choice for the default model and we shall now examine its influence on the single isotope scattering signal f. The lower panel of fig. 2 shows the Gaussian default function as determined from the moments of the experimental data in fig. 1 as dashed curve. The solid line is the ME solution on the basis of this default model. It has the same maximum position and the same half width as the upper panel solution and is, of course, much more satisfactory. The unphysical ringing in the wings of the line is greatly reduced. Note the error bars showing the confidence of the improved solution in these regions. Furthermore, $2\alpha S$ has decreased to 80% of its previous value indicating that more measured data accord with the default model. A minor point worth mentioning is that the reconstruction in fig. 2 is for an optimal incident energy $E_0 = 5.1$ keV determined via evidence analysis rather than at the nominal 5 keV, or in case of inelastic losses even below 5 keV. This discrepancy must be blamed on the calibration of the effective accelerating voltage of the primary ion beam and/or on minute deviations from the nominal scattering angle of 165°.

The single isotope scattering signal summarizes all contributions to the unknown broadening. The effects of apparatus broadening, multiple scattering events, inelastic energy losses or even surface defects, which are always present on real surfaces, are not considered in the binary collision formula (1). Since the apparatus broadening is not known further discrimination into the different physical effects is conjectural. Nevertheless, our main task of decomposing ISS spectra with overlapping signals into the constituents can be done without any bias.

4. The composite system Pd/Ru(001)

The treatment of the data from the clean Ru surface has already given a flavor of the importance to choose a default model incorporating all our reliable knowledge in a complicated case like the present one. Experimental data for scattering of Ar^+ off a Ru surface partially covered with Pd are shown in fig. 4 (note the different energy axis as compared

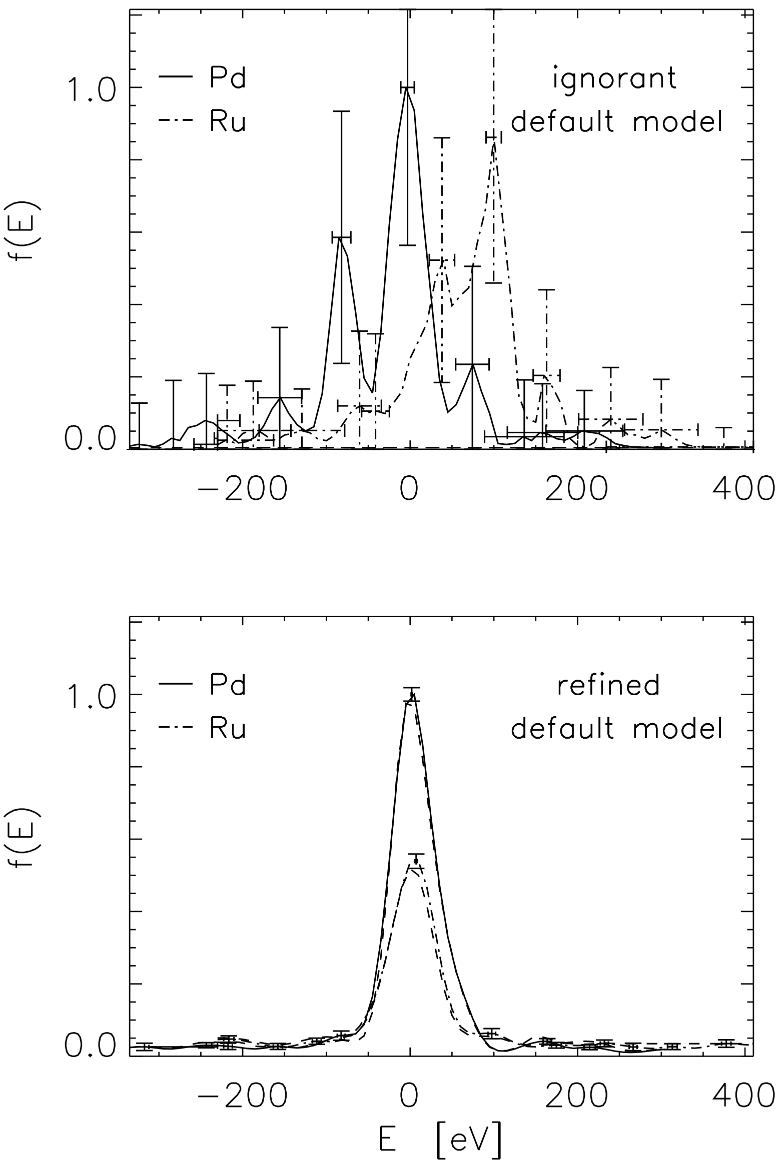

Figure 3: ME solution for the broadening function resulting in the ISS spectrum shown in fig. 4 for ions backscattered off Pd isotopes (solid line) and Ru isotopes (dash-dotted line) with different default models (dashed lines). In the upper panel the value of the default model is chosen small resulting in a useless solution with great errors in position and intensity of the structures. In the lower panel the default models are the solutions for the cases of a clean Ru surface and a thick Pd film. Note the considerably smaller error bars. The optimal incident energies for ions backscattered off Ru and Pd is 5100 eV and 5075 eV, respectively.

to fig. 1!). Proceeding as before with a flat default model we obtain the single isotope scattering functions $\vec{f}_{\mathrm{Ru}}$ (dash-dotted) and $\vec{f}_{\mathrm{Pd}}$ (solid) in the upper panel of fig. 3. They

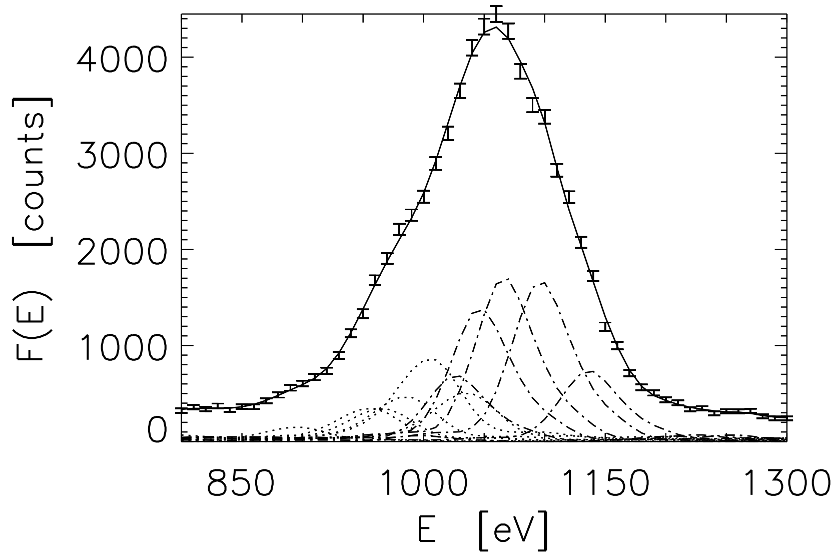

Figure 4: ISS spectrum of some amount of Pd adsorbed on the Ru(001) surface (error bars), the ME reconstruction (solid line), and the decomposition into the 7 Ru (dotted lines) and 6 Pd isotopes (dash-dotted lines). The ME solutions are as depicted in the lower panel of fig. 3.

are even more unacceptable than the solution for the clean Ru surface. Both functions show excessive oscillations which are even less significant according to the error bars than previously. Again a much better choice of the default function is mandatory. For Ru we choose the final reconstruction of the single isotope scattering function as derived for the clean Ru surface. For Pd we go through a similar procedure with auxiliary scattering data obtained from a thick Pd film such that no backscattering from Ru occurs. The two default models obtained are shown as dashed lines in the lower panel of fig. 3. Analysis of the scattering data from the partially covered Ru surface on the basis of these default models results in the dash-dotted curve for Ru for an optimal incident energy $E_0 = 5100$ eV as before and the continuous line for Pd for $E_0 = 5075$ eV. The energy difference must be due to different average inelastic losses in Ar^+Pd and Ar^+Ru collisions. Note that the single isotope scattering functions hardly deviate from their respective default models. The data for the partially covered surface support no additional information beyond the information we obtained for the clean Ru surface or the thick Pd film. Such a coincidence is expected on the basis of the binary collision model for ISS. On the other hand our results do not prove the binary collision model conclusively. Data of insufficient precision would lead to the same behavior.

The similarity of the here determined single isotope scattering functions to Gaussians may be taken as a justification for Gaussian fits to ISS spectra. Nevertheless the remaining deviations (see fig. 2, lower panel) are sizable and large enough that a Gaussian fit to a spectrum as complicated as in fig. 1 results in systematic deviations from the data. In other words, the data support more information as given by their first three moments $M^n(G), n = 0, 1, 2$.

The analysis of the Pd/Ru spectra requires the determination of still another parameter, namely $a_{\mathrm{Pd}}/a_{\mathrm{Ru}}$, the fractional coverage. This is determined by evidence analysis in much the same way as the selfconsistent determination of α. The fractional Pd coverage follows then from the backscattering yields observed on the all Ru and all Pd surfaces. A more extensive discussion of Pd adsorption on Ru(001) will be published elsewhere [11].

Let us finally look at the data synthesis on the basis of the single isotope scattering functions of fig. 3. This is shown as the solid line in fig. 4. In view of the meaning of Poisson errors the fit is quite good. Once again the solutions shown in the upper panel of fig. 3 would have lead to unreasonable overfitting. Dotted and dash-dotted lines indicate the contributions of Ru and Pd, respectively, to the overall spectrum.

In conclusion, the present application of ME techniques to experimental data has underlined their character as a quantitative measure of strength of belief. The reliability of the ME solutions must be accessed carefully before proceeding to physical interpretations. ME processing requires elaborate consideration and incorporation of prior knowledge which can only be supplied by an experienced user of the subject. It appears unlikely and dangerous for the unexperienced user to develop the ME procedure to a black box all purpose data processing package.

ACKNOWLEDGMENTS

We are grateful to N. Memmel and A. Steltenpohl for supplying the ISS data and helpful comments.

References

[1] E. Taglauer, "Investigation of the Local Atomic Arrangement on Surfaces Using Low-Energy Ion Scattering", *Appl. Phys.*, Vol. A 38, pp: 161–170, 1985.

[2] W. von der Linden, M. Donath and V. Dose, "Unbiased Access to Exchange Splitting of Magnetic Bands Using the Maximum Entropy Method", *Phys. Rev. Lett.*, Vol. 71, pp: 899–902, 1993.

[3] J. N. Kapur, "Maximum-Entropy Models in Science and Engineering", Wiley, New York, 1989.

[4] G. J. Daniell, "Of Maps and Monkeys: an Introduction to the Maximum Entropy Method", in: "Maximum Entropy in Action", Eds. B. Buck and V. A. Macaulay, Oxford University Press Inc., New York, p: 1, 1991.

[5] P. W. Anderson, "The Reverend Thomas Bayes, Needles in Haystacks, and the Fifth Force" *Phys. Today*, Vol. 45 (Jan.), pp: 9–11, 1992.

[6] Th. Fauster, "Surface geometry determination by large-angle ion scattering", *Vacuum*, Vol. 38, pp: 129–142, 1988.

[7] S. F. Gull, "Developments in Maximum Entropy Data Analysis", in: "Maximum Entropy and Bayesian Methods", Eds. J. Skilling, Kluwer Academic Publishers, Norwell, MA, pp: 53–71, 1989.

[8] J. Skilling, "Quantified Maximum Entropy", in: "Maximum Entropy and Bayesian Methods", Eds. P. F. Fougère, Kluwer Academic Publishers, Norwell, MA, p: 341, 1990.

[9] S. F. Gull and J. Skilling, "Maximum entropy method in image processing", *IEE Proc.*, Vol. 131 F, pp: 646–659, 1984.

[10] L. R. Mead and N. Papanicolaou, "Maximum entropy in the problem of moments", *J. Math. Phys.*, Vol. 25 (8), pp: 2404–2417, 1984.

[11] N. Memmel et al., unpublished.

KEYWORDS

- "Ion scattering spectroscopy" on p. 1
- "Prior information" on p. 1
- "Default model" on p. 3