DECONVOLUTION BASED ON EXPERIMENTALLY DETERMINED APPARATUS FUNCTIONS

V. DOSE, R. FISCHER AND W. VON DER LINDEN
Max-Planck-Institut für Plasmaphysik, EURATOM Association
POB 1533, D-85740 Garching, Germany†

Abstract.
Deconvolution of experimental measurements in e.g. electron or ion scattering can result in considerable resolution enhancement. It is usually assumed that the apparatus function which we wish to remove from the experimental signal is either known exactly or with a much higher precision than the signal. This assumption is not valid in general. In fact many situations are conceivable where measurement of the apparatus function requires the same effort as measurement of the signal. We have performed a rigorous Bayesian analysis for this general case and present applications to Rutherford backscattering from thin films.

Key words: Deconvolution, Inverse Problem, Apparatus Function, Likelihood, Image Processing

Deconvolution remains at the center of interest in the analysis of experimental data. Increasing the instrumental resolution to the desired value may not be possible, too cumbersome or expensive. The question then arises whether a sufficiently accurate measurement together with a sufficiently precise knowledge of the apparatus function suffices to arrive at the desired resolution by appropriate deconvolution procedures. The problem is old and has so far withstood a sustainable solution.

In this paper we shall choose for purposes of illustration the particular experiment of Rutherford backscattering (RBS). RBS is a surface analytical technique whose importance derives from its quantitative nature. Consider a beam of ions of mass m_0, usually either protons or He nuclei with well defined energy E_0 in the range of some MeV incident on a target of atomic mass number M_B with a thin film overlayer M_A. We shall assume - in line with the actual data - that M_A is very much larger than M_B. Ions undergoing an elastic Coulomb collision with a deflection by an angle Θ are recorded in a solid state detector. Since the projectile target interaction is coulombic, the cross section associated with the scattering is the Rutherford cross section, which is quantitatively known. All backscattered ions have energy $E_1 < E_0$. E_1 depends on E_0, m_0, $M_{A,B}$, Θ and is, as a function of

†Email: rainer.fischer@ipp.mpg.de

target mass, larger for scattering from M_A than for scattering from M_B since we have assumed $M_A >> M_B$. A typical situation is shown schematically in fig. 1.

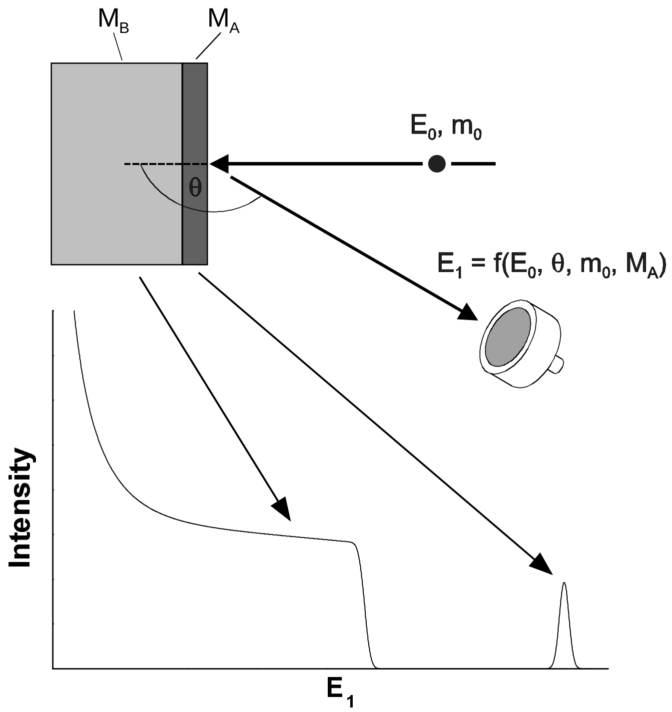

Figure 1. A schematic of a Rutherford backscattering (RBS) experiment.

Scattering from species A which forms only a thin film on top of the bulk substrate material B results therefore in a spectral spike slightly broadened by the apparatus transfer function while scattering from species B produces a step function. The high energy side of this step results from scattering from the topmost B-layer. Particles which have penetrated to some depth into material B before undergoing large angle scattering loose energy on their way into the scattering center and out to the detector and appear in the spectrum at progressively lower energies E_1. Note that the apparatus broadening is reflected in the rounded leading edge of the B-species spectrum.

Let us now turn to the various mechanisms contributing to the overall apparatus function. Numbers given for the various contributing effects are best estimates for the particular experimental setup which we used to obtain the data we are going to analyze here. The energy chosen for the analysis of Cu and Co films on a silicon substrate was 2.6 MeV with full width at half maximum of 13 keV. The solid state detector resolution at $E_1 \approx 1$ MeV is 15 keV wide. The electronic noise amounts to 5 keV while the contributions from detector solid angle, beam spot size and finite film thickness are estimated to be 3 keV. The resulting overall resolution is then approximately 20.7 keV. Data from a thin cobalt film on a silicon substrate

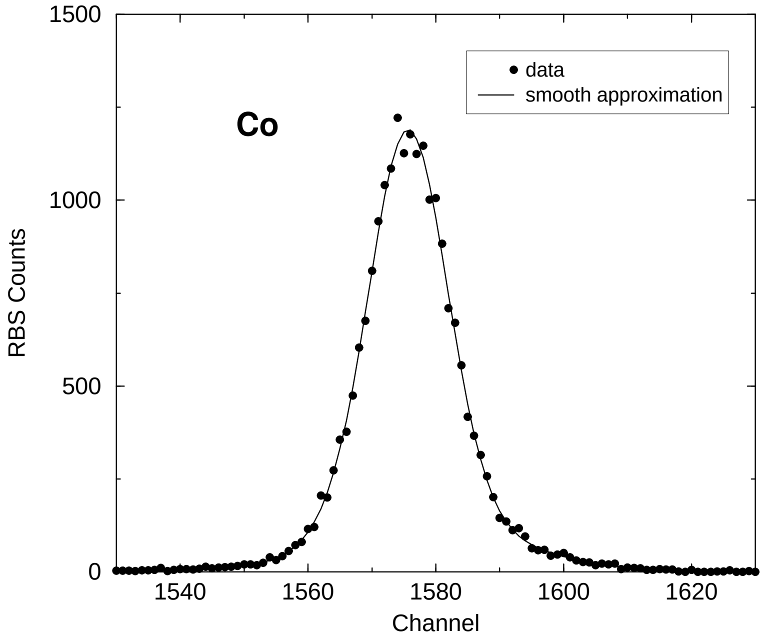

Figure 2. RBS spectrum from a thin cobalt layer on a silicon substrate using 2.6 MeV He^{+} ions. Since cobalt is a single isotope element the measurement represents the apparatus function.

displayed in fig. 2 are comparable with this estimate and yield the more precise overall resolution of 19 keV (FWHM). The thin continuous line running through the data points is a parameter free smooth density estimate of the experimental data shown as full dots [1]. The target material cobalt was chosen because it is a single isotope element. The Rutherford backscattering spectrum from a monolayer of cobalt obtained with an ideal apparatus is accordingly a δ-spike, the real data in fig. 2 represent consequently the overall apparatus function.

We shall now proceed to write down the likelihood function . The data vector $\boldsymbol{d}$ (the i-th component of $\boldsymbol{d}$ corresponds to the i-th energy channel) is then modeled as

$$\boldsymbol{d} = \boldsymbol{A}\,\boldsymbol{f} + \boldsymbol{N_d}, \quad < N_{d_i}^2 >= \sigma_i^2. \tag{1}$$

where $\boldsymbol{A}$ is the apparatus broadening matrix , $\boldsymbol{f}$ the desired spectroscopic feature and $\boldsymbol{N_d}$ the vector of noise deteriorating $\boldsymbol{d}$. According to the principle of maximum entropy we then have

$$P(\boldsymbol{d}|\boldsymbol{f}, \boldsymbol{A}, \boldsymbol{\sigma}, I) = \frac{1}{\prod_{i=1}^{N_d} \sqrt{2\pi\sigma_i^2}} \exp\left[-\frac{1}{2}\sum_i \left(\frac{d_i - (\boldsymbol{A}\,\boldsymbol{f})_i}{\sigma_i}\right)^2\right]. \qquad (2)$$

Since we are dealing with a counting experiment the diagonal matrix $\boldsymbol{\sigma}$ is easily estimated from the measurement. If we assume, that the apparatus matrix $\hat{\boldsymbol{A}}$ is also known for example as the continuous curve in fig. 2, then the desired spectral distribution $\boldsymbol{f}$ results from Bayes theorem

$$P(\boldsymbol{f}|\boldsymbol{d}, \hat{\boldsymbol{A}}, \boldsymbol{\sigma}, I) = \frac{P(\boldsymbol{d}|\boldsymbol{f}, \hat{\boldsymbol{A}}, \boldsymbol{\sigma}, I)\, P(\boldsymbol{f}|I)}{P(\boldsymbol{d}|\hat{\boldsymbol{A}}, \boldsymbol{\sigma}, I)}. \qquad (3)$$

We have offered two powerful procedures to obtain the answer to Eqn. (3) at the 1996 conference on maximum entropy and Bayesian methods [2,3]. One is the adaptive kernel method and the other the spline based adaptive resolution image reconstruction. Since both are well documented in the last years proceedings we shall abandon a discussion of the actual inverse problem (3) but rather consider the effects, if we replace the assumed known apparatus matrix by the real world pointwise measurement shown in fig. 2 as full dots. Let $\boldsymbol{A}$ represent this matrix and let a_{ij} be its elements. We then have to calculate $P(\boldsymbol{d}|\boldsymbol{f}, \boldsymbol{A}, \boldsymbol{\sigma}, \boldsymbol{\Delta}, I)$ by first introducing the true apparatus function $\hat{\boldsymbol{A}}$ and subsequently integrating it out,

$$P(\boldsymbol{d}|\boldsymbol{f}, \boldsymbol{A}, \boldsymbol{\sigma}, \boldsymbol{\Delta}, I) = \int d\,\hat{a}_{ij}\; P(\boldsymbol{d}|\boldsymbol{f}, \hat{\boldsymbol{A}}, \boldsymbol{\sigma}, \boldsymbol{\Delta}, I)\, P(\hat{\boldsymbol{A}}|\boldsymbol{A}, \boldsymbol{\Delta}, I), \qquad (4)$$

where we have suppressed all logically irrelevant conditional information. $\boldsymbol{\Delta}$ is the matrix of standard deviations of the measurement of a_{ij} and again easily estimated since we deal with a counting experiment. Accordingly the prior distribution $P(\hat{\boldsymbol{A}}|\boldsymbol{A}, \boldsymbol{\Delta}, I)$ is

$$P(\hat{\boldsymbol{A}}|\boldsymbol{A}, \boldsymbol{\Delta}, I) \propto \exp\left[-\frac{1}{2}\sum_{ij}\left(\frac{\hat{a}_{ij} - a_{ij}}{\Delta_{ij}}\right)^2\right]. \qquad (5)$$

Since both factors in the integral (4) are quadratic forms in the variables $\hat{a}_{ij}$ we proceed in the usual way: find the maximum and complete the square. This is somewhat massy but elementary. Assuming further that the likelihood function (2) is sufficiently localized such that all integrations can be extended to $-\infty < \hat{a}_{ij} < \infty, \forall i, j$ we arrive at

$$P(\boldsymbol{d}|\boldsymbol{f}, \boldsymbol{A}, \boldsymbol{\sigma}, \boldsymbol{\Delta}, I) = \frac{1}{\prod_{i=1}^{N_d} \sqrt{2\pi\sigma_{\mathrm{eff},i}^2}}\, exp\left(-\frac{1}{2}\sum_{i=1}^{N_d}\frac{(d_i - (\boldsymbol{A}\,\boldsymbol{f})_i)^2}{\sigma_{\mathrm{eff},i}^2}\right). \qquad (6)$$

The most important difference between (6) and (2) is the replacement of $\boldsymbol{\sigma}$ by $\boldsymbol{\sigma}_{\mathbf{eff}}$, again a diagonal matrix given by

$$\sigma_{\mathrm{eff},i}^2 = \sigma_i^2 + \sum_{j=1}^{N} \Delta_{ij}^2 f_j^2. \qquad (7)$$

with σ_i^2 being the contribution to the likelihood variances from the data measurement and $\sum_{j=1}^{N} \Delta_{ij}^2 f_j^2$ being the contribution from the finite precision of the measurement of the apparatus matrix $\boldsymbol{A}$.

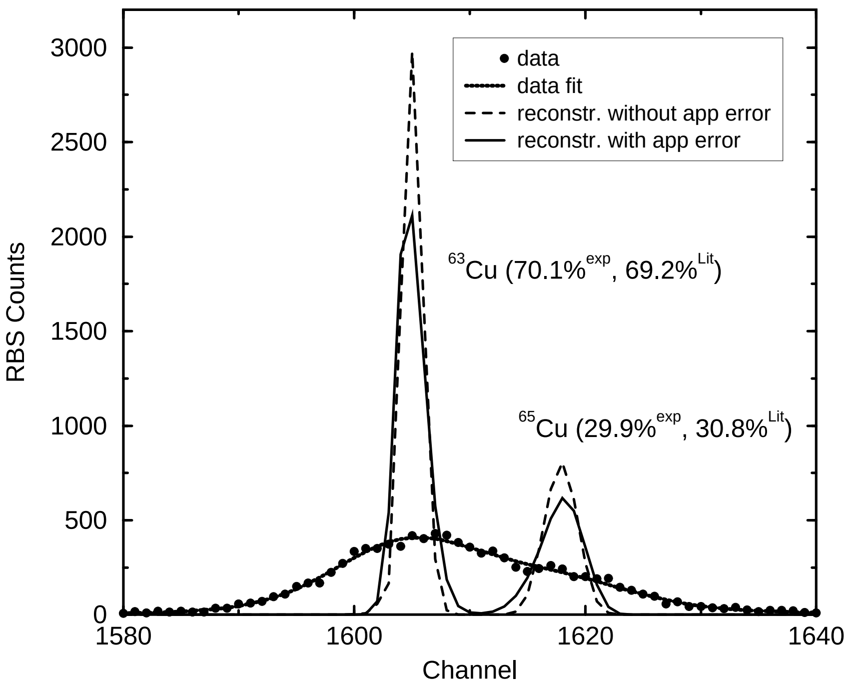

Figure 3. RBS spectrum from a thin copper layer on silicon using 2.6 MeV He^+ ions. The primary data do not separate the isotopes at $A_1 = 63$ amu and $A_2 = 65$ amu. Reconstruction using the smooth and the pointwise function of fig. 2 are given. Note the excellent agreement of the isotope abundances in comparison to tabulated data.

The data which we have processed were obtained from a thin copper film deposited again on silicon and are shown in fig. 3. Copper unlike cobalt has two different isotopes $A_1 = 63$ amu with a natural abundance of 69.2% and $A_2 = 65$ amu with a natural abundance of 30.8%. The unprocessed RBS spectrum of the copper thin film shown as full dots in fig. 3 does not resolve these two isotopes. The dashed line is obtained upon assuming that the apparatus function is identical to the smooth curve approximation of the measurement shown in fig. 2 using (3). We see a dramatic gain in resolution. The continuous curve through the data points represents the data fit. If we now replace the likelihood function (2) by the likelihood function (6) which is based on the pointwise measured cobalt spectrum in fig. 2 we obtain the continuous peaks and obviously, we have lost resolution. The difference between the results based on (2) and (6) respectively can easily be reconciled recalling that the variances of apparatus function $\boldsymbol{A}$ and spectral feature $\boldsymbol{f}$ add up under the convolution operation to the variance of the signal $\boldsymbol{d}$. Thus

$$\mathrm{var}(\boldsymbol{d}) \quad = \quad \mathrm{var}(\boldsymbol{A}) + \mathrm{var}(\boldsymbol{f}). \tag{8}$$

The question arises which measurement accuracy is sufficient for the desired resolution. For the sake of arguments we assume that the variances of the peaks of the data $\boldsymbol{d}$, the apparatus function $\boldsymbol{A}$ and the image $\boldsymbol{f}$ are well described by their widths W_d, W_A and W_f. With $W_f = \alpha \times W_d$, where α is the inverse of the resolution enhancement, $\alpha < 1$, eqn. 8 transforms into

$$\frac{W_A}{W_d} \quad \approx \quad (1-\alpha^2)^{\frac{1}{2}} \approx 1 - \epsilon \tag{9}$$

From conservation of the area of the images under convolution we estimate $\alpha = 1/7.5$ leading to $\epsilon = 0.009$ for the dashed line spectrum in fig. 3 and $\alpha = 1/5.4$ with $\epsilon = 0.017$ for the continuous line spectrum in fig. 3. The difference in the width of the images based on either (2) or (6) translates accordingly into a difference in width of the data of 0.8%. This is exactly the uncertainty of the width of the pointwise given apparatus function in fig. 2 if we make a quick and dirty calculation of its second moment and the error thereof. But it is not only the difference in width that matters, it is also necessarily the pointwise apparatus function which produces the smaller resolution enhancement since the smooth curve representation is always one with as low a curvature as possible when obeying the data constraints. Its width represents therefore systematically an upper limit to the width of the true apparatus function, thereby overdoing the deconvolution.

In addition, ϵ is a measure for the width accuracies necessary for the desired resolution. For an ideal measurement of a δ-peak $\frac{W_d}{W_A} = 1$ whereas for a real measurement $\frac{\Delta W_d}{W_A} = \epsilon$. For a resolution enhancement of 10 (100) the joint uncertainty of the widths of the data and the apparatus function has to be smaller than 0.5% (0.005%). This crude assessment reveals the limits of appropriate deconvolution procedures in cases where we deal with noisy data.

References

1. R. Fischer, M. Mayer, W. von der Linden, and V. Dose, "Enhancement of the energy resolution in ion-beam experiments with the maximum-entropy method," *Phys. Rev. E*, **55**, p. 6667, 1997.
2. R. Fischer, W. von der Linden, and V. Dose, "Adaptive kernels and occam's razor in inversion problems," in *MAXENT96 - Proceedings of the Maximum Entropy Conference 1996*, M. Sears, V. Nedeljkovic, N. E. Pendock, and S. Sibisi, eds., p. 21, NMB Printers, Port Elizabeth, South Africa, 1996.
3. W. von der Linden, V. Dose, and R. Fischer, "How to separate the signal from the background," in *MAXENT96 - Proceedings of the Maximum Entropy Conference 1996*, M. Sears, V. Nedeljkovic, N. E. Pendock, and S. Sibisi, eds., p. 146, NMB Printers, Port Elizabeth, South Africa, 1996.