Deconvolution of positron annihilation coincidence Doppler broadening spectra using iterative projected Newton method with non-negativity constraints

K.F.Ho[(1)], C.D. Beling[(1)], S. Fung[(1)*], K.W. Cheng[(1)], Michael K. Ng [(2)] and A.M. Yip[(3)]

(1) Department of Physics, The University of Hong Kong, Pokfulam Road, Hong Kong, PR China.
(2) Department of Mathematics, The University of Hong Kong, Pokfulam Road, Hong Kong, PR China
(3) Department of Mathematics, University of California, Los Angeles, 405 Hilgard Avenue, CA 90095, U.S.A.

[*] Email: sfung@hkucc.hku.hk

Abstract

A generalized least-square method with Tikonov-Miller regularization and non-negativity constraints has been developed for deconvoluting two dimensional Coincidence Doppler Broadening Spectroscopy spectra. A projected Newton algorithm is employed to solve the generalized least-square problem. The algorithm has been tested on Monte Carlo generated spectra to find the best regularization parameters for different simulated experimental conditions. Good retrieval of the underlying positron-electron momentum distributions in the low momentum region is demonstrated. The algorithm has been successfully used to deconvolute experimental CDBS data from Aluminum.

PACS number(s): 78.70.Bj, 07.05.Kf, 07.05.Rm

1. Introduction

Angular Correlation of Annihilation Radiation (ACAR) and Doppler Broadening Spectroscopy (DBS) provide information on the Electron Momentum Distribution (EMD) in the material under investigation through measurements of the Momentum Density of electron-positron Annihilating Pairs (MDAP) [1]. In regular solids the MDAP can be used to check theoretical band structure calculations, and provides details on Fermi surfaces in more complex alloys [2]. The higher resolving power of ACAR is normally required for these studies. On the other hand, in defected solids where the MDAP is normally more representative of the positron trapping sites than the bulk solid, the DBS technique can be quickly and usefully employed to distinguish different defects by plotting the S(valence) parameter against the W(core) parameter (the S-W plot) [3]. The DBS technique, while capable of a fast data accumulation rate, suffers from the poor intrinsic resolution of HP(Ge) gamma ray spectroscopy systems (typically ~5mrad ACAR equivalent, in comparison to ~1mrad for typical ACAR systems). DBS on its own is thus normally considered insufficient for accurate MDAP measurement in regular solids and with regard to defect studies DBS is often perceived to have reached its limits of usefulness in the S-W plot.

An extension of the conventional DBS, Coincidence Doppler Broadening Spectroscopy (CDBS) has recently been opening up new horizons [4]. In CDBS, two simultaneous measurements are made of the Doppler shift on a pair of annihilation photons; a procedure having the advantage of providing very low background levels which permits core annihilations to be accurately observed and thus providing information on the chemical environment of the positron [5]. Moreover, with relevance in MDAP measurement, CDBS also has the intrinsic property of improving the instrumental resolution by factors of up to $\sqrt{2}$ ($\sqrt{2}$ applying only when the detectors have equal resolution) [6], together with the possibility of high quality deconvolution being made possible by the fact that an almost perfect resolution function of the system exists; namely in the form of the energy spectrum of the 514 keV gamma ray line of ^{85}Sr where photons from the source are observed in pseudo-coincidence [7]. Britton *et al.* [7] demonstrated both the importance of the $\sqrt{2}$ improvement and the effect of deconvolution by building a CDBS system with an effective resolution of 386 eV in full-width-at-half-maximum (FWHM) unit (~1.5 mrad ACAR equivalent). The main motivation of the present work has been similar, namely that of trying to investigate if ACAR quality spectra can be obtained using CDBS. CDBS remains a desirable method over ACAR because of its simplicity of operation and

the added usefulness of the chemical environment sensitivity provided by its ability to observe high momentum core electron momenta.

The use of suitable deconvolution algorithms is an important issue in improving the quality of CDBS spectra in the low momentum range. A major factor in the success of any deconvolution venture is the quality of the input spectrum itself. Assuming the spectrum has been perfectly stabilized against electronic drift effects, there are still the uncertainties due to noise arising from the stochastic nature of the counting process. Thus the more counts in the spectrum, the more true to the convoluted functional shape, i.e. the less noisy, it becomes. The number of channels into which the CDBS spectrum is discretized is also a factor effecting spectral quality [8]. A well set up CDBS system could be expected to record good annihilation photon coincidences at a rate of 10^3cps, which over a period of a few days could give total number of counts (N_{cts}) ~ 10^9 events [8]. Moreover a good modern nuclear analogue-to-digital-converter (ADC) can digitize into number of channels (N_{chn}) ~ 16,000 giving CDBS image data of 16000×16000 pixels. Improvements in deconvolution are not linear in N_{cts} and N_{chn} but tend to be logarithmic [8]. The question which naturally arises as to what constitutes a "good enough" deconvolution under the presently available hardware and computational resources is one of the issues considered in the present work.

A variety of algorithms have been adopted for deconvolution studies on the 1D annihilation line spectra of DBS in the past. Many different algorithms have been used such as the Stokes method [7, 9-12], the maximum entropy method [13, 14], iterative methods [15-18], optimized linear filtering [19] and the method of generalized least squares [20-24]. It is the latter method which is presented in detailed form in this paper. Although the methods listed above have all given promising results, none of them excepting reference [20] incorporates non-negativity constraints on the deconvoluted spectra. A major objective of the present work is to discuss the feasibility of implementing such constraints through the use of the projected Newton method [25]. Another objective is to explore the more general use of deconvolution in the two dimensional image data produced in CDBS.

The outline of the paper is as follows: First, with reference to the intrinsically ill-conditioned image restoration problem, the use of Tikhonov regularization will be illustrated. *A priori* knowledge of non-negativity will be introduced into the generalized least square formulation. A projected Newton-based algorithm will then be described to solve the least-square problem. The effectiveness of this algorithm is investigated as a function of the number of counts, matrix size (number of channels), order of derivative to be regularized and the

regularization parameter. The effective system resolution of a CDBS system which incorporates a deconvolution code is then considered. Finally the projected Newton algorithm is applied to real experimental CDBS data for polycrystalline Aluminum.

2. Theory

2.1 The CDBS "image" data

In the CDBS technique the energies E_1 and E_2 of both annihilation photons are measured by two HP(Ge) detectors in back-to-back orientation and E_2 is plotted against E_1 to form a 2D histogram. These energies may be written as:

$$E_1 = m_o c^2 - \tfrac{1}{2}E_n + \Delta E + \delta_1$$
$$E_2 = m_o c^2 - \tfrac{1}{2}E_n - \Delta E + \delta_2 \qquad (1)$$

where m_o is the electron rest mass, E_n is the binding energy of electrons in the n^{th} shell, ΔE is the Doppler shift and δ_1 and δ_2 are the independent measurement errors for each detector. Annihilation events as given by (1) are thus seen as lying on the lines $E_1 + E_2 = 2m_o c^2 - E_n$ with the lines broadened by s_1 and s_2 in both coordinates, the latter being the standard deviations of instrumental errors δ_1 and δ_2.

In comparison with conventional DBS method, the CDBS momentum signal is doubled since:

$$E_1 - E_2 = p_{//}c = 2\Delta E + \delta_1 + \delta_2 \qquad (2)$$

By taking $\sigma_1 = \sigma_2$, one has, assuming δ_1 and δ_2 to be Gaussian in distribution, a standard deviation on $E_1 - E_2$ of:

$$s = \sqrt{s_1^2 + s_2^2} = \sqrt{2}s_1 \qquad (3)$$

From (2) and (3), it is seen that the signal strength is doubled but the error on the signal has only increased by a factor of $\sqrt{2}$ ~40% [6]. In practice the condition $\sigma_1 = \sigma_2$ is seldom met (the detectors being of different intrinsic resolution) so that the $\sqrt{2}$ improvement in momentum measurement is only approached, but never fully attained.

Figure 1a shows an experimental CDBS taken for polycrystalline Aluminum. We can see that there are vertical and horizontal cross-pieces at $E_1 = m_0c^2$, $E_2 = m_0c^2$. These are produced on the high energy side by pulse pile-up and on the low energy side by incomplete charge collection. In the case of the positron source being ^{22}Na, events are present on the low and high energy sides due to coincident Compton events from the 1.27MeV gamma ray associated with

this source. However the cross-pieces need not be of undue concern since in the first place their intensity is much less than that of the real annihilation events on the diagonal, and second, under the proviso that we have "subtracted-off" any Compton background, the remaining counts can in some sense be considered as part of the instrumental resolution function, the features being connected with the main signal in the same way for both the CDBS spectrum and the instrumental "blurring" function. The presence of similar "cross-pieces" indeed may be seen in the resolution spectrum shown in Figure 1b obtained from a ^{85}Sr source taken under pseudo-coincidence conditions.

2.2 The deconvolution problem

The convoluted 2D image g_{ij} of a true 2D image function f_{ij} ($1 = i = m,\ 1 = j = m$) with an instrumental function h_{ij} ($1 = i = m,\ 1 = j = m$) 4) with additive noise may be written [26] as follows

$$g_{ij} = \sum_{i'=1}^{m} \sum_{j'=1}^{m} h_{i-i'j-j'} f_{i'j'} + n_{ij} \tag{4}$$

where n_{ij} is an unknown noise and m is the size of the image (N_{chn} in our case). The matrix form becomes:

$$\boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{n}. \tag{5}$$

It is well known that the matrix $\boldsymbol{H}$ is ill-conditioned as a result of its averaging effect. Tikhonov [27] postulated an equation

$$\min_{f} \left\| \boldsymbol{g} - \boldsymbol{H}\boldsymbol{f} \right\|_2^2 + a \left\| \boldsymbol{f}^{(k)} \right\|_2^2 . \tag{6}$$

Here $\| \cdot \|_2$ denotes the Euclidean norm. In the minimization problem associated with (6), the quality of the solution is controlled by the choice of the regularization parameter $\alpha>0$ in the second "Tikhonov regularization" term. Here $\| \boldsymbol{f}^{(k)} \|$ provides a measure of the total energy of $\boldsymbol{f}$ (when $k=0$) or the k^{th} derivatives of $\boldsymbol{f}$ (when $k>0$) depending upon the particular choice of k. Large values of α yield less noisy solutions for $\boldsymbol{f}$ but with some loss of information on "sharp" features, while if α is set too low the solution for $\boldsymbol{f}$ may be too noisy [21]. The formal solution of this equation is, for a given α,

$$\boldsymbol{f}_a = \left(\boldsymbol{H}^T \boldsymbol{H} + a \boldsymbol{L}^{(k)} \right)^{-1} \boldsymbol{H}^T \boldsymbol{g} . \tag{7}$$

Here $\boldsymbol{H}^T$ denotes the transpose of $\boldsymbol{H}$ and $\boldsymbol{L}^{(k)}$ is the regularization matrix corresponding to the k^{th} derivative of $\boldsymbol{f}$. If the instrumental function is spatially shift invariant, then the deconvoluted image $\boldsymbol{f}_a$ can be obtained efficiently by using fast cosine or Fourier transforms. However, when

the instrumental function is spatially shift variant, the equation (7) can be solved iteratively to obtain $\boldsymbol{f}_a$ [28].

Both the calculated spectrum $\boldsymbol{f}$ and the observed spectrum $\boldsymbol{g}$ should not contain negative elements. This is imposing known *a priori* knowledge of $\boldsymbol{f}$ on the solution. The constraints on non-negative elements in the minimization problem are essential and turn out to be highly effective as a means of regularization. With non-negativity included the minimization problem (6) may be written as:

$$\min_{f \geq 0} \left\| \boldsymbol{H}\boldsymbol{f} - \boldsymbol{g} \right\|_2^2 + a \left\| \boldsymbol{f}^{(k)} \right\|_2^2 . \tag{8}$$

The projected Newton method [25] can be used to solve the above nonlinear minimization problem. The j^{th} iteration of the projected Newton algorithm can be described as follows:

$$[\boldsymbol{D}_A + \boldsymbol{D}_I(\boldsymbol{H}^T\boldsymbol{H} + a\ \boldsymbol{L}^{(k)})\boldsymbol{D}_I]\boldsymbol{s}_{j+1} = \boldsymbol{z}(\boldsymbol{f}_j)$$

$$\boldsymbol{f}_{j+1} = [\boldsymbol{f}_j + \boldsymbol{s}_{j+1}]_+ \tag{9}$$

where $\boldsymbol{z}(\boldsymbol{f}_j) = (\boldsymbol{H}^T\boldsymbol{H} + a\boldsymbol{L})\boldsymbol{f}_j - \boldsymbol{H}^T\boldsymbol{g}$, $([\boldsymbol{x}]_+)_i = x_i$ if $x_i \geq 0$, and 0 otherwise, and, $\boldsymbol{D}_A$ is a diagonal matrix whose i^{th} diagonal entry is 1 if the i^{th} entry of both $\boldsymbol{f}_j$ and $\boldsymbol{z}(\boldsymbol{f}_j{}^k)$ are non-negative, and is 0 otherwise. Similarly, $\boldsymbol{D}_I$ is a diagonal matrix given by $\boldsymbol{I} - \boldsymbol{D}_A$. In each of the projected Newton iterations, the linear system in (9) can be solved by the conjugate gradient method effectively [29].

2.3 Monte-Carlo simulation of CDBS data

There has been a significant amount of research which aims at determining the optimal regularization parameter in generalized least square problems [30]. In this work, this parameter is optimized by interactive selection as shown schematically in Figure 2. First a Monte-Carlo (M.C.) CDBS spectra is generated. As outlined below this technique closely mimics the response of a CDBS spectrometer to annihilation photons having a typical MDAP characteristic of a metal. The deconvoluted function $\boldsymbol{f}_a$ is finally compared with the underlying MDAP of the material $\boldsymbol{f}$ through observation of the weighed residuals $(\boldsymbol{f}_a - \boldsymbol{f}) \cdot \boldsymbol{w}$, $\boldsymbol{w}$ being the column vector with elements $1/\sqrt{f_i}$. The 'optimal' regularization parameter can thus be found so as to bring the closest visual match of $\boldsymbol{f}_a$ to $\boldsymbol{f}$ or alternatively by defining a weighted "chi-squared" error

$$c^2 = \sum_{i=1}^{N} \frac{(f_{a,i,w} - f_i)^2}{f_i} . \tag{10}$$

The method of producing the simulated data has already been described in a previous article [24]. Only a brief summary will be given here. Seven variates in the range of 0 to 1 are required to register a single event on the simulated 2D CDBS histogram. Considering the sample to be a simple metal (Lithium is used here) the first variate decides whether the positron is to annihilate with a core or conduction band electron. If the annihilation is from a core electron the next two variates are used to position its energy $E_{\gamma 1}$ according to the Gaussian distribution:

$$P(E_{g1}) = \frac{1}{a_C\sqrt{2p}} e^{-\frac{(E_{g1}-E_0)^2}{2a_C^2}}. \qquad (11)$$

where $E_0=mc^2$ and a_C is the standard deviation of the distribution. This distribution is obtained using the standard MC formulation for Gaussian generation [31, 32]:

$$E_{g1} = E_0 + a_C\sqrt{2\ln\frac{1}{x_2}}\cos 2px_3 \qquad (12)$$

$?_2$ and $?_3$ being the two random variates. Alternatively, in the case of an event coming from the conduction band then E_{g1} is thrown according to the inverted parabola.

$$P(E_{g1}) = \frac{3}{4a_F^3}.[a_F^2 - (E_{g1}-E_0)^2] \qquad (13)$$

where a_F represents Doppler shift corresponding to the Fermi momentum. The MC method for generating (14) is not well documented, but this distribution can simply be obtained by throwing events in momentum space out to a radius of a_F (which also closely mimics the real annihilation process in a metal). That is one writes [33]:

$$E_{g1} = E_0 + a_F\sqrt[3]{x_2}(2x_3 - 1) \qquad (14)$$

$?_2$ and $?_3$ being the two random variates. With the $E_{\gamma 1}$ energy determined, energy conservation through equation (1) demands that:

$$E_{g2} = 2E_0 - E_{g1}. \qquad (15)$$

The next step is to incorporate the resolution broadening associated with both detectors and their associated electronics. This is implemented by shifting the energies of the annihilation event to $E_{?1}$ and $E_{?2}$ to new values according to independent randomized Gaussian distributions. Four variates are used at this stage – two for each Gaussian. The instrumental "blurring" function can be obtained easily using the same MC code by putting $a_C = a_F = 0$ (i.e. by replacing $\boldsymbol{h}$ with a delta function --- see Figure 2). Finally the events are cumulated and binned.

Various sizes of 2D matrices have been investigated (256×256, 512×512 and 1024×1024). MC spectra were thrown from 10^4 to 10^8 counts. To approximately mimic the

ACAR momentum distribution, parameters pertinent to Lithium data [34] were chosen (i.e. a_c = 7.5mrad, a_F = 4.4mrad). The σ value of the detector was taken as 2 mrad (corresponding to a FWHM unit of 1.2keV and 4.7mrad of ACAR equivalent). Simulated CDBS data, assuming the MDAP for Li are shown in Figure 3a. The spectrum is quite ideal when compared to a real CDBS spectrum. In particular it has no low level random background and none of the characteristic "cross" events at $E_{?1} = E_{?2} = 511$ keV. An example of the "blurring" function is shown in Figure 3b.

3. Results and Discussion

3.1 Improvement with non-negativity constraints

The deconvolution of the simulated spectrum without non-negativity constraints is shown in Figure 4. While the resulting deconvoluted image is indeed sharper it is seen to suffer from ripples, possessing negative values, and a fragmented background. Variation of the regularization parameter was unable to reduce this rippling effect. It was found, however, that introduction of the non-negativity constraints into the deconvolution algorithm could largely reduce these undesirable features. The deconvoluted spectrum of a simulated CDBS data for Lithium metal using non-negativity constraints with $\alpha = 10^{-3}$ is shown in Figure 5. The $\boldsymbol{f}_a$ data lie close to the line $E_1 + E_2 = 2E_0$ (= 512 channels) in a narrow band without any of the negative ripples present in the "single shot" method. The quality of the retrieval was assessed by taking a restricted channel cut along the diagonal line. The retrieved CDBS spectrum $\boldsymbol{f}_a$ is shown in Figure 6 together with the true MDAP function $\boldsymbol{f}$ and experimental 1D-ACAR data for Li metal [34]. The good agreement between $\boldsymbol{f}$ and $\boldsymbol{f}_a$ suggests that the non-negativity constraints are highly effective at producing accurate regularized solutions for the deconvolution. Although there is still some rippling present, $\boldsymbol{f}_a$ does not significantly deviate from the true $\boldsymbol{f}$. Such reasonable agreement would be sufficient for a wide range of positron annihilation investigations.

The first problem to address in optimal deconvolution is to decide on the best cut for the channel width 'δ'. Figure 7 shows a plot of c^2 against δ. For low-count spectra, i.e. $N_{cts} = 10^4$ and 10^5, the minimum c^2 is attained when $\delta = 3.5$ (in FWHM units). For high-count spectra with $N_{cts} = 10^6$ and 10^8, the δ for c^2 minimum is slightly larger at ~ 4.0 FWHM unit. These results are in good agreement with the work of Gebauer et al. [35], indicating that a choice of channel width

of around 2 FWHM unit is suitable for a wide range of N_{cts}. The value of δ = 2 FWHM was taken in the following optimization studies.

3.2 Optimizing the regularization parameter and norm

In the present study N_{cts} was set at either 10^4, 10^5, 10^6, and 10^8 while image matrix sizes ($N_{chn} \times N_{chn}$) were set to 256×256, 512×512 and 1024×1024 --- our computational limit. In all the studies $N_{cts} = 10^8$ was chosen for $\boldsymbol{h}$ spectra and the tolerances of errors on the conjugate gradient method was taken as 10^{-3}.

Figure 8 shows c^2 as a function of the regularization parameter a. It is seen that a minimum occurs when $N_{cts}=10^4$ while for $N_{cts}=10^5$, 10^6 and 10^8 the c^2 decreases monotonically as a is reduced. If a is too small, then insufficient regularization will cause c^2 to be high. Conversely if a is too large, regularization errors will occur again forcing c^2 to be high [36]. In the $N_{cts} = 10^4$ case, the minimum occurs at a = 0.1. The absence of a minimum for the high-count spectra ($N_{cts} = 10^5$, 10^6 and 10^8) indicates the strong regularizing action of the non-negativity constraint. At $N_{cts} = 10^5$, signal-to-noise ratio is so small that the effect of regularization becomes insignificant and thus by putting a = 0, we can still obtain optimal deconvolution.

The quality of the deconvolution as expressed by c^2 improves as N_{cts} increases since the statistical noise level reduces for large N_{cts} to allow better function retrieval. Larger matrix sizes are also expected to yield better deconvolution results as more information is present in the spectra on both $\boldsymbol{g}$ and $\boldsymbol{h}$ functions. Figure 9a and Figure 9b where c^2 is plotted for different N_{chn} against a for both $N_{cts} = 10^4$ and 10^8 respectively. Figure 9a shows as already discussed, the regularization parameter α is still effective, the value of a = 0.1 giving optimal regularization irrespective of N_{chn}. The important point, however, is that the c^2 value decreases with N_{chn}, indicating that some improvement in deconvolution is to be obtained by choosing higher binning levels. Irrespective of matrix sizes, the high statistics data ($N_{cts} = 10^8$), as shown in Figure 9b, produces the same "no minimum" behaviour as remarked on above. Once again the $N_{cts} = 10^4$ data gives a poorer quality deconvolution. The general trend of the larger N_{chn} giving smaller c^2 is clear.

In Figures 10a and 10b equivalent c^2 versus a plots are shown for the cases of k = 1 and k = 2 respectively. Comparison of Figures 8 and 10a shows very little difference between k = 0

and $k = 1$ norms. However for the $k = 2$ norm shown in Figure 10b, while being very similar to $k = 0$ and 1 for low a ($<10^{-2}$), gives a markedly improved regularization for large regularization parameters a ($>10^{2}$). For $N_{cts} = 10^{5}$, there exists a c^2 minimum now at a = 0.1, and a better deconvolute is obtained than the case of zero a. From this evidence it is to be generally concluded that the second derivative is a slightly better choice of norm [23]. However, for high spectral content CDBS data $N > 10^{5}$ in which the regularization parameter $\alpha \leq 10^{-4}$, there is no noticeable difference in the quality of deconvolution, all derivatives essentially yielding the same result. This finding is the same as that of Chambless and Broadway [21, 22] who also used non-negativity constraints in deconvolution.

3.3 The Residual Instrumental Function (RIF)

With many modern spectroscopy systems the distinction between hardware and software components becomes less distinct, since both contribute to the final data quality. Certainly within the context of the present discussion of deconvolution a CDBS spectroscopy system may most appropriately be considered to consist of the hardware electronics and software as a single system. This being the case, the single system may be expected to have an effective resolution function, termed the Residual Instrumental Function (RIF) [19], which is sharper than that of the hardware alone, namely the Hardware Instrumental Function (HIF). In order to address this question quantitatively, a deconvolution improvement factor F is defined as:

$$F = \frac{FWHM(HIF)}{FWHM(RIF)}. \qquad (16)$$

Comparison of the narrower width of deconvoluted CDBS data from the peak along the $E_1 - E_2 = 0$ diagonal in Figure 4 and the raw data in Figure 1a reveals an improvement of effective system resolution. This follows from equation (1) since the detection errors d_1 and d_2 add in quadrature along this diagonal making the FWHM in this direction an accurate indicator of the system resolution. Some caution must be taken, however, in generally assessing the system resolution from the FWHM value along the $E_1 - E_2 = 0$ (positive) diagonal since the shape of the combined resolution curve along this diagonal is not strictly speaking the same as that along the $E_1 - E_2 = 2mc^2$ (negative) diagonal. As discussed by Britton *et al.* [7] the resolution function along the negative diagonal (which is the resolution function of importance in CDBS) is the correlation integral of the two separate detector resolution functions $R_1(E)$ and $R_2(E)$, while that along the positive diagonal is the convolution of $R_1(E)$ and $R_2(E)$. In the case of $R_1(E)$ and $R_2(E)$ being both symmetric functions, as is the case in the present testing on simulated spectra, the combined resolution function is the same along both diagonals. Thus in the present study we are permitted

to take FWHM values of HIF and RIF from the profiles of the raw and deconvoluted data along the positive diagonal as a means of assessing the quality of the deconvolution.

The $E_1 - E_2 = 0$ diagonal cross-sections of both $\boldsymbol{g}$ and $\boldsymbol{f}_a$ data are shown in the inset of Figure 11. The value of F obtained from these cross sections is plotted against the regularization parameter α for the simulated 512×512 channel Li CDBS spectra of different N_{cts} in Figure 11. For $\alpha<10^{-3}$, F remains essentially constant. (The lowering of F for large α results from the regularization error tends to broaden $\boldsymbol{f}_a$ through over-smoothing.) Apart from the $N_{cts} = 10^3$ case, the saturation value of F is ~3. An increase N_{cts} improves F as expected. However, the reason for the small improvement (i.e. about 3.0 to 3.1 for $N_{chn} = 512$) and the small increase in RIF quality against a large increases of N_{cts} is not well understood. In a similar study, Beling *et al*. [8] using Monte-Carlo data with a much simpler Stoke's scheme found empirically that:

$$F = 1.5 + 0.067 \ln(N_{cts} N_{chn}) \qquad (17)$$

This relationship gives a stronger dependency of F on $\ln(N_{cts})$ than that is found in the present work, although the magnitude, i.e. F~3 is similar. It is easy to obtain F values of around 3 from deconvoluted spectrum and this leads to some optimism for approaching typical ACAR resolution performance using CDBS. In our case, the effect of dependence on N_{cts} and N_{chn} on the quality of deconvolution is less significant, and the realistic gains in instrumental resolution to be obtained using the non-negativity constraint by increasing N_{cts} and N_{chn} are more limited.

Deconvoluted spectra of effective system resolution (FWHM) 320 eV (or 1.25 mrad ACAR equivalent) are obtained with F ~ 3 from the present simulated data which started with HP(Ge) detector resolution of (FWHM) 1200 eV. The comparison with the resolving power of a typical ACAR spectrometer ~0.5-1.0 mrad (i.e. 128-256 eV) is encouraging. It must also be stressed that while the non-negativity method is good, it is probably not the ultimate optimal way of performing CDBS spectral deconvolutions. Further improvements may well be obtained with more sophisticated algorithms currently under development in our research group.

3.4 Deconvoluting experimental CDBS spectra

The CDBS spectrometer used in this work consisted of two HPGe detectors, with 80% and 60% relative efficiencies and 1.6keV and 2.0keV FWHM resolutions at 514keV. The net effective resolution of the system was ~ 4.0 mrad. The number of counts in the CDBS spectrum $\boldsymbol{g}$ and the resolution function spectrum $\boldsymbol{h}$ (obtained from ^{85}Sr in quasi-coincidence) were 2×10^7 and 5×10^7 respectively. Digital spectrum stabilization was employed on both E_1 and E_2 channels.

The experimental CDBS data for annealed polycrystalline Aluminum are shown in Figure 12a, where background counts due to Compton events from the 1.27MeV gamma ray associated with ^{22}Na have been subtracted off by fitting a Gaussian to the "cross" data on the high energy sides of both E_1 and E_2. (This background was not considered to be part of the experimental signal in contrast to the "tailing" counts on the low energy side of the "cross" which may be considered as part of the instrumental resolution. The background was similarly subtracted from the $\boldsymbol{h}$ spectrum). The deconvoluted $\boldsymbol{f}_a$ is shown in Figure 12b. The improved resolution is clearly seen from the narrowed width of the data in the E_1-E_2=0 direction. Moreover there is no sign of rippling in the near vicinity of this CDBS diagonal. Unfortunately there is some residual signal belonging to the low energy "cross" portions of E_1 and E_2. This effect indicates that the low energy "cross" portions in the function $\boldsymbol{h}$ give a perfect description of the instrumental function. The reasons for this are unclear at present. This residual, however, is unlikely to be affecting the data on the main CDBS diagonal, which is seen to be nicely symmetric.

The quality of the deconvolute can be assessed by comparing it with the known 2? momentum density as obtained using 1D-ACAR data for Al. This comparison is shown in Figure 13. The ACAR data were taken from the early work of Stewart [34] and are used here for comparison being the only data available for the polycrystalline metal. The agreement with the ACAR data, while not perfect is seen to be very good. Indeed we believe that for many research applications this quality of 2? momentum density may prove quite adequate.

Acknowledgements:

The work described in this paper is partially supported by the grants from the Research Grant Council of the Hong Kong Special Administrative Region, China (under project nos. HKU7103/02P, HKU7104/98P, HKU7091/00P and HKU1/00C) and the Hung Hing Ying Physical Science research fund. The research of one of us (M. K. Ng) is supported in part by RGC Grant Nos. 7130/02P and 7046/03P and Science Faculty Collaborative Seed Grant, 2002-2003. The authors wish to thank Mr. C. M. Woo for writing the MPI code on high performance computer cluster and Ms. P. Y. Kwan for preparing the graphs.

References:

[1] M. Sob, H. Sormann and J. Kuriplach, *Adv. Quant. Chemistry* **42**, 77 (2003)

[2] R. N. West, *Positron Spectroscopy of Solids*, edited by A. Dupasquier and A. P. Mills jr. (IOS Press, Ohmsha, 75, 1995)

[3] L. Liszkay, C. Corbel, L. Baroux, P. Hautojarvi, M. Bayhan, A. W. Brinkman, S. Tararenko, *Appl., Phys. Lett.* **64**, 380 (1994)

[4] M. Alatalo, H. Kauppinen, K. Saarinen, M. J. Puska , J. Makinen, P. Hautojarvi and R. M. Nieminen, *Phys Rev.* B **51,** 4176 (1995)

[5] P. Asoka-Kumar, M. Alatalo, V. J. Ghosh, A. C. Kruseman, B. Nielsen and K. G. Lynn, *Phys Rev. Lett.*, **77**, 2097, (1996)

[6] J. R. MacDonald, K. G. Lynn, R. A. Boie and M. F. Robbins, *Nucl. Instr. Meth* ., **153**, 189(1977)

[7] D. T. Britton, W. Junker, P. Sperr, *Materials Science Forum* **105**, 1845 (1992)

[8] C. D. Beling, M. Li, Y. Y. Shan, S. H. Cheung, S. Fung, B. K. Panda and A. P.Seitsonen, *J Phys: Condens. Matter*, **10**, 10475 (1998)

[9] A. R. Stokes, *Proc. Phys. Soc*., **61**, 382 (1948)

[10] R. Chang R., B. Williams, M. Copper, *Phil. Mag* ., **23**, 115 (1971)

[11] J. P. Schaffer, E. J. Shaughnessy and P. L. Jones, *Nucl. Instr. Meth* . **B 5** , 75 (1984)

[12] J. P. Schaffer and P. L. Jones , *J. Phys F:Met. Phys*. **16**, 1885 (1986)

[13] Y. Kong and K. G. Lynn, *Nucl. Instr. Meth*., **A302**, 145 (1991)

[14] A. Schukla , M. Peter and L. Hoffmann, *Nucl. Instr. Meth* . **A 335** , 310 (1993)

[15] Shizuma Kiyoshi, *Nucl. Instr. Meth.* **150**, 447 (1978)

[16] Y. Ichioka, Y. Takubo, K. Matsuoka and T. Suzuki, *J. Optics (Paris)* , **12**, 35 (1981)

[17] S. Dannefaer and D. P. Kerr, *Nucl. Instr. Meth* . **131** , 119 (1975)

[18] L. Calderin and J. J. Diaz, *Nucl. Instr. and Meth* ., **B117**, 457 (1996)

[19] A. A. Manuel, *Proceedings of the international school of physics <<ENRICO FERMI>>, Course CXXV, Positron Spectroscopy of Solids, SOCIETA ITALIANA DI FISICA BOLOGNA-ITALY IOS Press*, p. 155 (1995)

[20] P. Paatero, S. Manninen and T. Paakkari, Report series in Physics, University of Helsinki, No. 75 (1974)

[21] D. A. Chambless and J. A. Broadway, *IEEE Transactions on Nuclear Science*, NS-28, **2**, 1938 (1981)

[22] D. A. Chambless and J. A. Broadway, *Nucl. Instr. Meth.* **179**, 563 (1981)

[23] B. K. Panda, S. Fleischer, C. C. Ling, C. D. Beling, S. Fung, S. Panda, *Appl. Surf. Sci.*, **85**, 182 (1997)

[24] K.F. Ho, K. P. Ng, C. D. Beling, S. Fung, K. L. Chan and H. W. Tang, *Materials Science Forum*, **363**, 673 (2001)

[25] D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, New York, pg.76, 221-246 (1982)

[26] J. H. Jansson, *Deconvolution of Images and Spectra* , Academic Press (1984)

[27] A. Tikhonov and V. Arsenine , *Methodes de resolution de problemes mal poses* (Mir, Moscow, 1976)

[28] M. K. Ng, R. H. Chan and W. C. Tang, *SIAM J Sci. Comput*, **21**, 851 (1999)

[29] Vogel C. R., *Scientific Computing, Springer*, Singapore, p148 (1997)

[30] G. Wahba, *SIAM, J. Numer. Anal*, **14**, 651(1977)

[31] L. L. Carter and E. D. Cashwell, *Particle Transport Simulation with the Monte Carlo Method*, (Oak Ridge, TN: USERDA Technical Information Center) (1975);

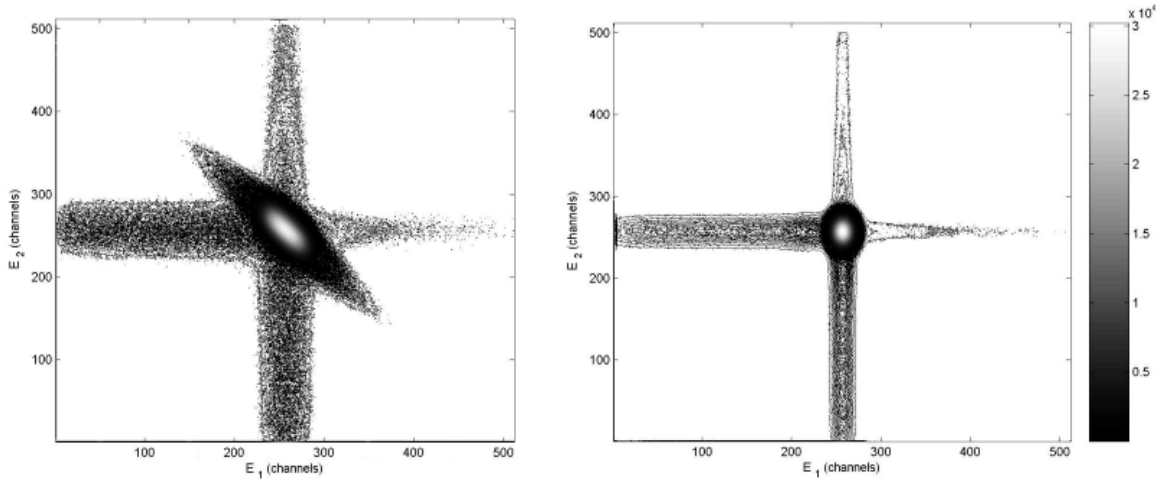
[32] G. S. Fisherman, *Principles of Discrete Event Simulation* (New York; Wiley-Interscience) (1978)

[33] P. K. MacKeown and C. D. Beling, Private communication (2000)

[34] A. T. Stewart, *Can. J. Phys.*, **35** 168 (1957); Phys. Rev. A **133** (1964) 1651

[35] J. Gebauer, R. Krause-Rehberg, S. Eichler, F. Borner, *Applied Surface Science* **149**,110-115 (1997)

[36] R. L. Lagendijk and J. Biemond, *Iterative identification and restoration of images*, (Boston: Kluwer Academic Publishers, 1991)



(a) (b)

Figure 1. (a) Experimental CDBS 2D histogram spectrum taken for polycrystalline Al metal (N_{cts} = 2×10^7). (b) Effective CDBS 2D histogram resolution function as obtained using a ^{85}Sr source ($E_?$=514keV) taken with random events (Coincidence time window = 2μs, N_{cts}= 5×10^7). The energy calibration for both spectra is 75eV/channel. As can be seen, the resolution of the E_1 detector (1.2keV) is superior to that of the E_2 detector (2.0keV).

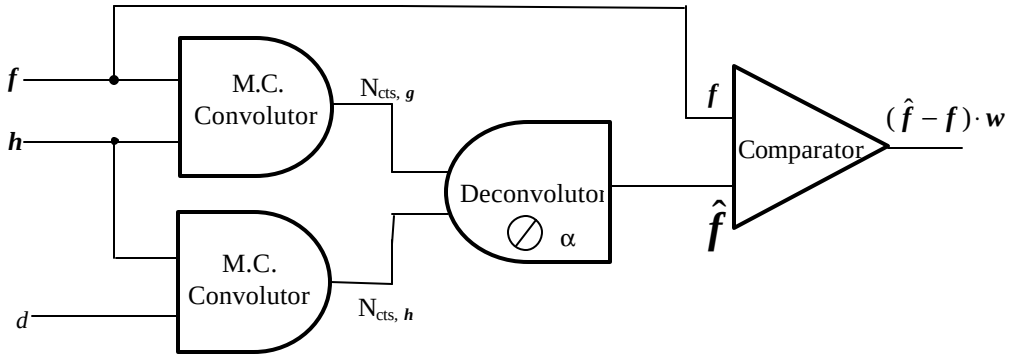
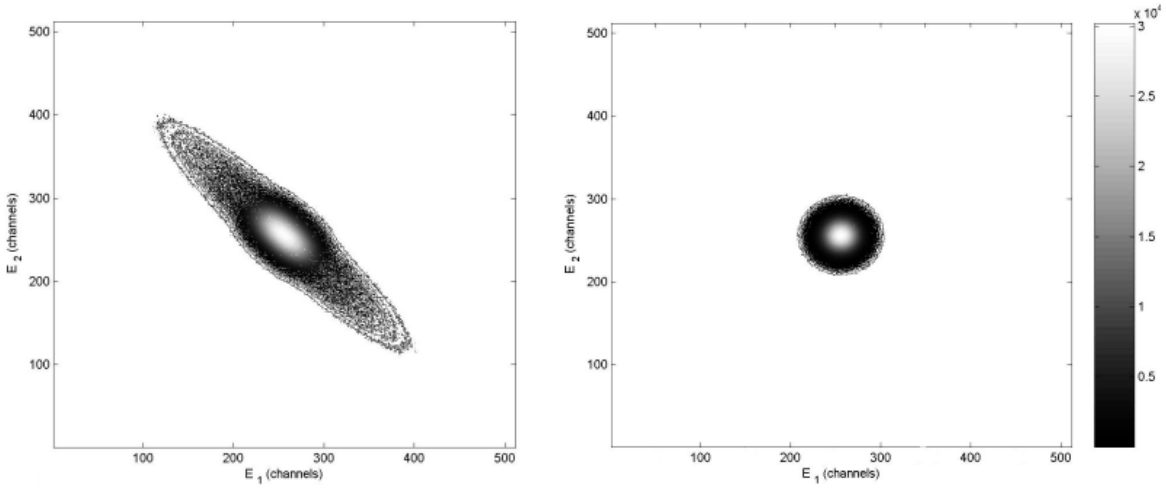


Figure 2. Schematic diagram shows the methodology used for optimizing the deconvolution algorithm. Two runs of a MC (Monte Carlo) program simulates both the experimental processes producing a synthetic CDBS spectrum of known MDAP $\boldsymbol{f}$ and resolution function $\boldsymbol{h}$. This data is then fed into the deconvolution program (2DNNGLSD) to produce the deconvoluted $\boldsymbol{f}_a$. The final weighted residual spectrum between $\boldsymbol{f}$ and $\boldsymbol{f}_a$ is constructed for purposes of assessing the fidelity of the deconvolution.



(a) (b)

Figure 3. Typical synthetic spectra produced by the MC simulation program. (a) CDBS spectrum, $\boldsymbol{g}$, approximately corresponding to Li metal. (b) Resolution function $\boldsymbol{h}$.

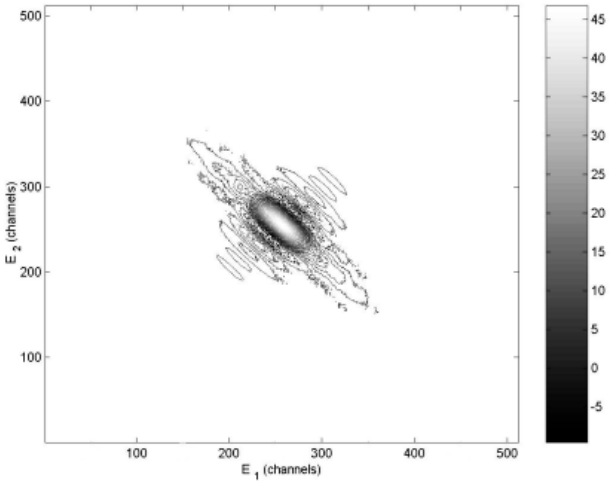


Figure 4. Result of deconvoluted CDBS $\boldsymbol{f}_a$ with no non-negativity constraint has been employed. The regularization parameter has been optimized but still negative portions exist with considerable rippling.

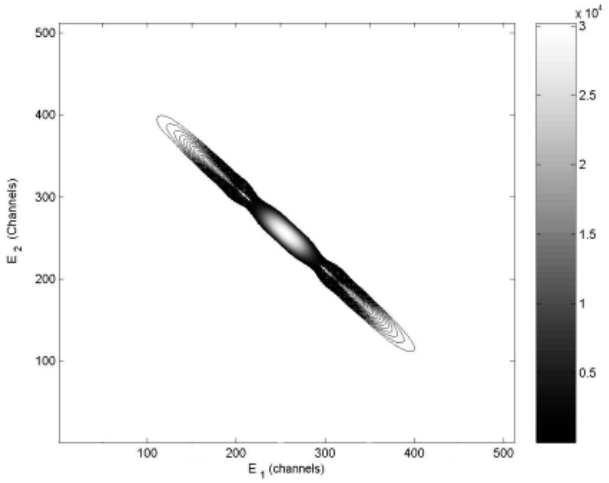


Figure 5. Result of deconvoluted CDBS $\boldsymbol{f}_a$ with non-negativity constraint has been employed. The amount of rippling in the deconvoluted result is seen to be very much reduced. For this deconvolution a = 10^{-3} and the zeroth norm regularizer were used.

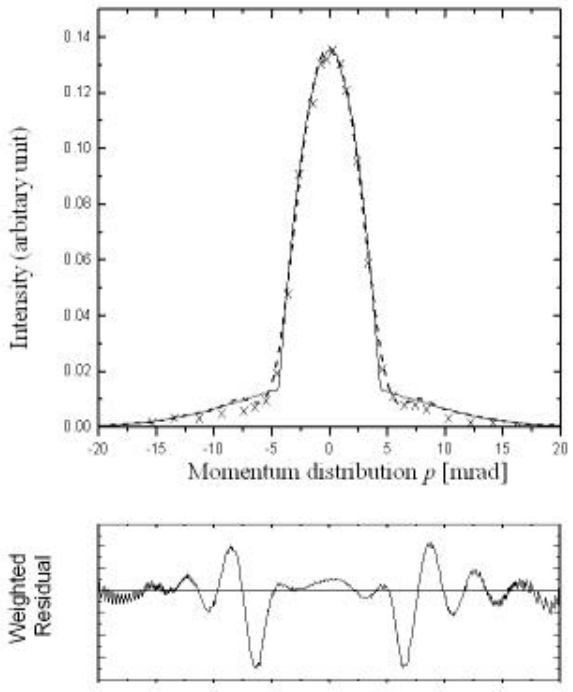


Figure 6. 1D CDBS spectrum $\boldsymbol{f}_a$ (dashed line) taken from a diagonal cut of 31 channels (2 system FWHMs) compared with the true MDAP $\boldsymbol{f}$. The crosses show ACAR – MDAP data taken from Stewart [34].

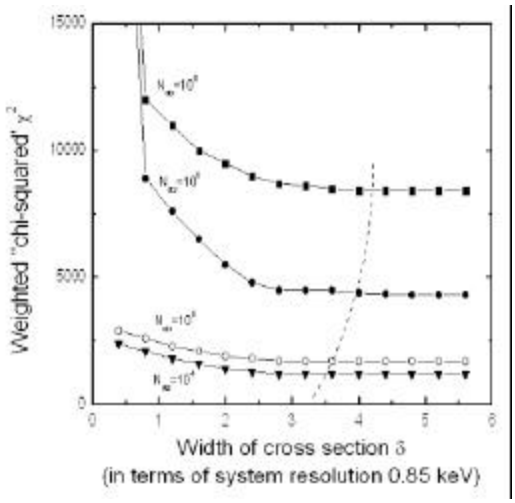


Figure 7. Variation of summed weighted squared residuals, "chi-squared", c^2 plotted as a function of the cutting width d along the CDBS diagonal for spectra having N_{cts} = 10^4, 10^5, 10^6 and 10^8 counts. The dotted line indicates the value of d for which c^2 is minimum.

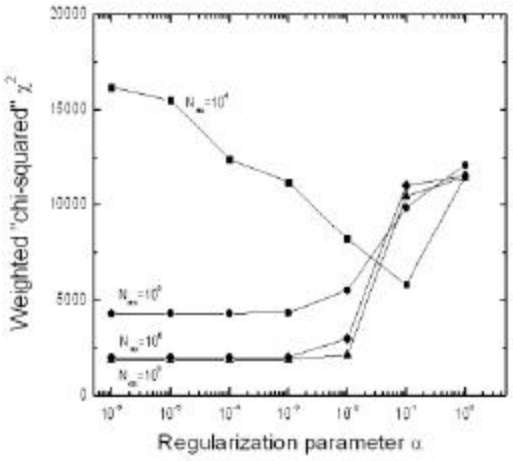
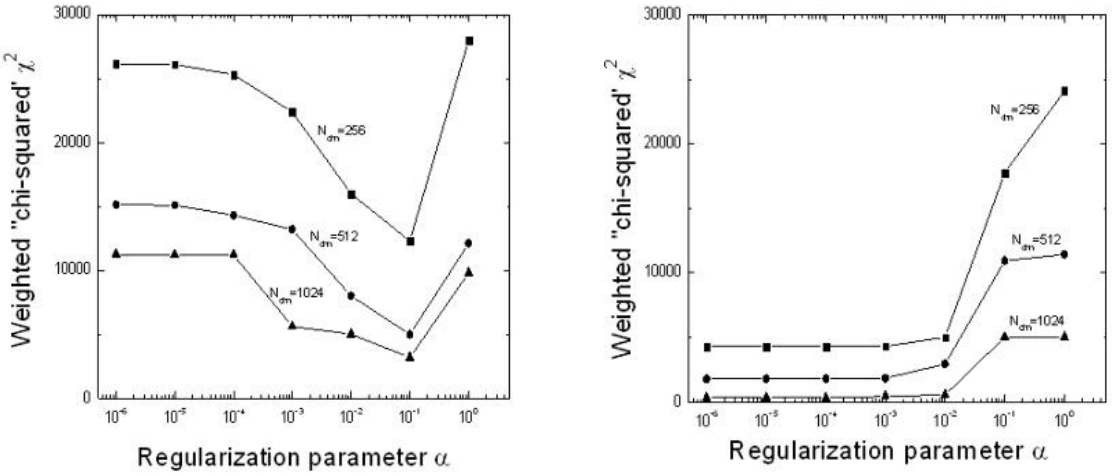
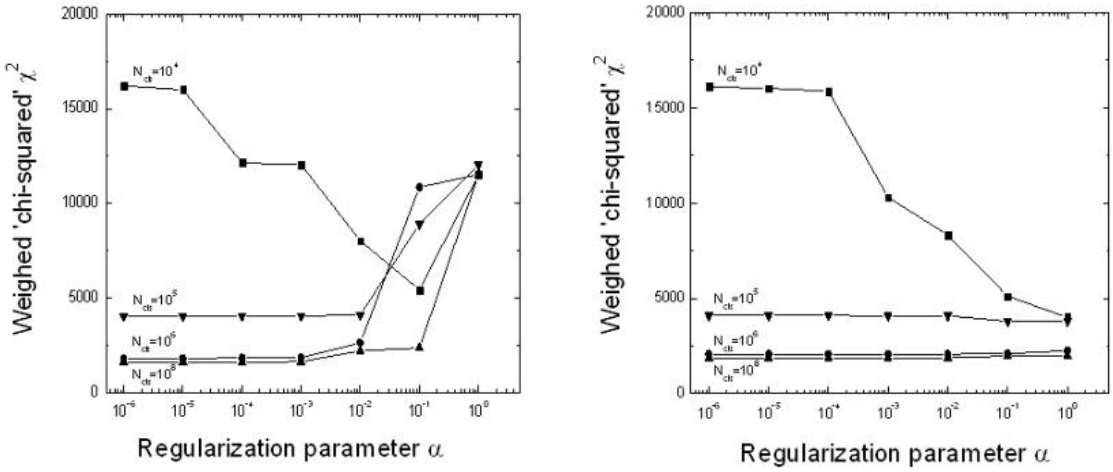


Figure 8. Variation of the "chi-squared" c^2 of residuals plotted against the regularization parameter a for synthetic CDBS spectra having N_{cts} = 10^4, 10^5, 10^6 and 10^8 counts, deconvoluted using the 2DNNGLSD algorithm. The zeroth norm regularization was used, with N_{chn} = 512.



(a) (b)

Figure 9. Variation of the "chi-squared" c^2 of residuals plotted against the regularization parameter a for synthetic CDBS spectra obtained using the 2DNNGLSD algorithm; (a) with N_{cts} = 10^4 and (b) with N_{cts} = 10^8. The data are shown for the different CDBS matrix sizes or N_{chn} = 256, 512 and 1024.



(a) (b)

Figure 10. Variation of the "chi-squared" c^2 of residuals plotted against the regularization parameter *a* for synthetic CDBS spectra with N_{chn} = 512 channels obtained using the 2DNNGLSD algorithm, N_{cts} = 10^4, 10^5, 10^6 and 10^8 counts, (a) with 1st derivative norm regularizer and (b) with 2nd derivative norm regularizer.

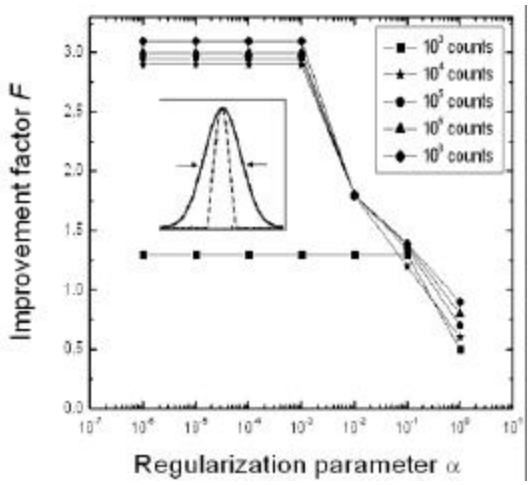
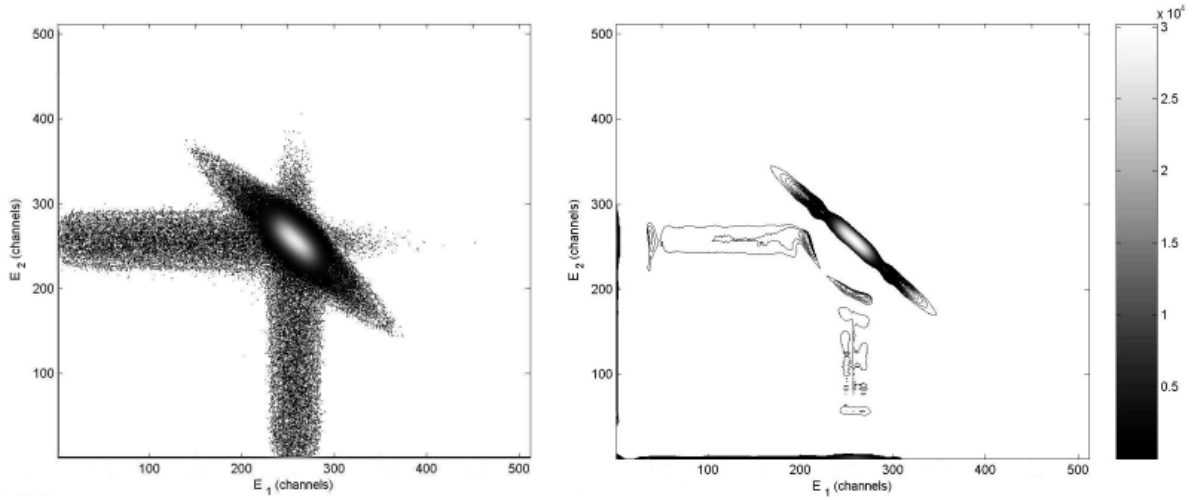


Figure 11. Deconvolution improvement factor *F* plotted as a function of regularization parameter a for zeroth norm regularization obtained using the 2DNNGLSD algorithm. The data are shown for N_{cts} = 10^3, 10^4, 10^5, 10^6 and 10^8 counts. The insert shows the cross-sections through the CDBS diagonal for both the $\boldsymbol{g}$ and $\boldsymbol{f}_a$ functions (the ratio's of which FWHM's define *F*).



(a) (b)

Figure 12. (a) Experimental CDBS spectrum for polycrystalline Al metal, with high energy backgrounds subtracted across the whole spectrum. (b) The same data after having been deconvoluted using the 2DNNGLSD algorithm. The $\boldsymbol{h}$ function (not shown) was that of Fig 1b, but again the "cross-pieces" were partially subtracted based upon the data on the high energy sides. The zeroth norm with $a=10^{-3}$ were used in this deconvolution. The number of spectral counts $N_{cts} = 2\times10^7$cts.

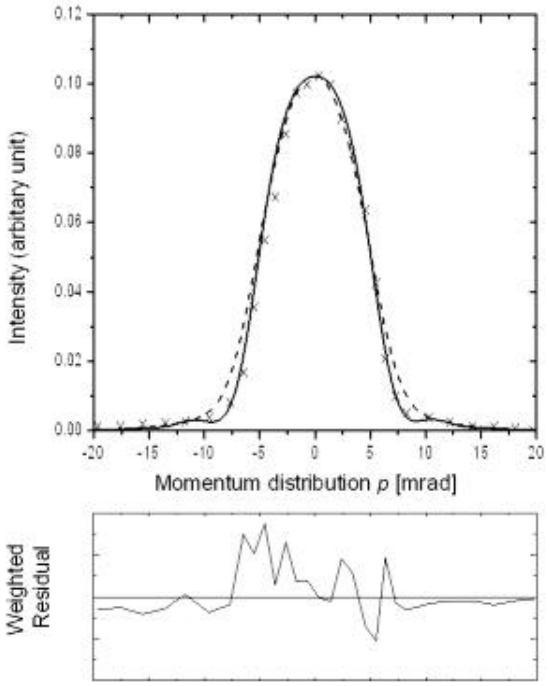


Figure 13. 1D CDBS cut (with channel width = 2 system FWHMs) for the deconvoluted Al data shown in Fig. 12b. The dotted line shows the CDBS cut prior to deconvolution. The solid line shows the deconvoluted f_a. The crosses show the 1D-ACAR data of Stewart [34] for polycrystalline Al. The residual plot shows an accurate MDAP is produced over the whole momentum range after deconvolution.