Application of the Maximum Entropy Method to X-ray Profile Analysis

Nicholas Armstrong

B. App. Sc. (Hons)

Department of Applied Physics

University of Technology, Sydney

Australia

A thesis submitted for the degree of Doctor of Philosophy

1 March 1999

CERTIFICATE

I certify that this thesis has not already been submitted for any degree and is not being submitted as part of candidature for any other degree.

I also certify that the thesis has been written by me and that any help that I have received in preparing this thesis, and all sources used, have been acknowledged in this thesis.

Signature of Candidate

....

Acknowledgment

The author is deeply grateful to his supervisor Dr W. Kalceff, and co-supervisor Associate Professor R. W. Cheary for their guidance, encouragement, criticisms and generosity over the period of this work. I am also deeply grateful to Dr J. P. Cline from the Ceramics Division, National Institute of Standards and Technology (NIST), Gaithersburg, USA, for his stimulating discussions and encouragement concerning size/strain analysis, and for providing the opportunity for funding and access to resources during the fourteen months spent in his laboratory. I would also like to thank: Dr D. C. Green for his guidance and encouragement concerning the writing of the thesis; Dr A. J. Reuben for his interesting discussions concerning the mathematical issues encountered in this project; Dr M. A. Stevens-Kalceff for the scanning electron microscopy analysis and images of the alumina specimens; Ms Di O' Brien for proof reading the chapters in this thesis. I would like to thank friends and family for their encouragement over the years.

My deepest thanks and gratitude go to my wife Tania Perry for her continual support, patience and encouragement that has seen this work evolve into a thesis.

There is the inherent simplicity and beauty of physics. It's the kind of thing where you can take nature and describe it in terms of mathematical models and the mathematical models allow you to reach far-reaching conclusions and you can test those conclusions by going back in the lab. The whole structure of making progress in this way seems so appealing ... once you've established this connection between experiment and the model of how nature works, you begin to believe it.

Stephen Chu, 1997.

Contents

1 Literature Review **1**

1.1 Introduction 1

1.2 An overview of size- and strain-induced x-ray profile broadening 2

1.2.1 Conditions for diffraction — ideal case 2

1.2.2 Diffraction from real materials 3

1.3 Size and strain broadening of x-ray diffraction profiles 7

1.3.1 Integral-breadth representation of size- & strain-broadened profiles . . 7

1.3.2 Fourier representation of size- & strain- broadened profiles 14

1.4 Analysing broadened diffraction profiles 22

1.4.1 Determining the specimen profile 23

1.4.2 Williamson-Hall method 26

1.4.3 Warren & Averbach method 30

1.4.4 Recent developments 38

1.5 Summary 43

2 Analysis of Deconvolution Methods used in X-ray Diffraction **44**

2.1 Theory 45

2.1.1 Background 45

2.1.2 Unconstrained Inversion/Deconvolution Methods 47

2.1.3 Iterative Deconvolution Method 52

2.1.4 Constrained Deconvolution Method 54

2.2 Computational Results and Discussion 57

2.2.1 Simulated Diffraction Profiles 57

2.2.2 Unconstrained Inversion/Deconvolution Methods 58
2.2.3 Iterative Method 60
2.2.4 Constrained Deconvolution Method 63
2.2.5 Implications for Experimental Data 65
2.3 Summary 67

3 Development of a Maximum Entropy Method for X-Ray Diffraction Profile Analysis 68

3.1 The Maximum Entropy Principle 69
3.1.1 Positive and additive distributions 69
3.1.2 The maximum entropy principle 71
3.2 The Maximum Entropy (MaxEnt) Method 75
3.2.1 *A priori* probability distribution 76
3.2.2 Likelihood probability distribution 77
3.2.3 Bayes' Theorem 79
3.3 Skilling & Bryan MaxEnt Algorithm 80
3.3.1 The rationale of the Skilling-Bryan algorithm 81
3.3.2 Defining the subspace 82
3.3.3 Entropy metric 83
3.3.4 Formulating the problem in the subspace 84
3.3.5 Determining the uncertainty in $\hat{\mathbf{f}}$ 87
3.3.6 Limitations of the Skilling-Bryan Algorithm 88
3.4 Summary 89

4 A Maximum Entropy Method for Determining the Column-length Distribution 91

4.1 Crystallite size broadening and analysis 92
4.1.1 Crystallite size broadening 92
4.1.2 Instrument broadening 94
4.2 Maximum Entropy Method for X-ray Profile Analysis 95
4.2.1 Background estimation of the observed profile 95

4.2.2 Determining the column distribution 96
4.3 Simulated profiles 97
4.3.1 Simulated size broadened profiles 97
4.3.2 Simulated instrument profiles 101
4.3.3 Modelling of "observed" profiles 102
4.4 Computational Results and Discussion 102
4.4.1 Preliminary calculations 103
4.4.2 Using the MaxEnt method to determine $\hat{f}(2\theta)$, $\hat{A}(L)$, $\hat{p}_a(L)$ and size results. 105
4.4.3 The effect of background estimation on $\hat{f}(2\theta)$, $\hat{A}(L)$, $\hat{p}_a(L)$ and size results. 118
4.4.4 Effect of non-ideal instrument profiles on $\hat{f}(2\theta)$, $\hat{A}(L)$, $\hat{p}_a(L)$ and size results. 126
4.5 Summary 130

5 A Maximum Entropy Method for Determining the Strain Distribution 132
5.1 Strain broadening of profile 133
5.1.1 Specimen profile and Fourier coefficients 133
5.1.2 Inverse Cosine Fourier Transform 135
5.2 Maximum Entropy (MaxEnt) method for $p(\epsilon_L)$ 137
5.2.1 First approach: Method of moments 138
5.2.2 Second approach: Using the Fourier coefficients 140
5.3 Simulated strain broadened profiles 142
5.3.1 Strain distribution and specimen profiles 143
5.3.2 Simulated instrument profiles 147
5.3.3 Modelling of "observed" profiles 147
5.4 Computational Results and Discussion 150
5.4.1 Preliminary calculations 150
5.4.2 Case 1: Gaussian strain 153
5.4.3 Case 2: Non-Gaussian strain 163
5.4.4 Strain Analysis 174

5.5 Summary 185

6 Application of the Maximum Entropy Method to Alumina Diffraction Data **188**

6.1 Background to analysis 189

6.1.1 Overview of previous analysis 189

6.1.2 Determining the instrument profile 189

6.1.3 Background-level estimation 190

6.1.4 Determining the specimen profile and column-length distribution . . . 192

6.1.5 Determining the variance and uncertainties 192

6.2 MaxEnt profile analysis using a uniform model 193

6.2.1 Deconvolving the observed profile using a uniform model 193

6.2.2 Integral breadth, $\beta(d^*)$. 198

6.2.3 Fourier coefficients, $A(L)$ 199

6.2.4 Determining the column-length distribution – first attempt 202

6.3 MaxEnt profile analysis using a non-uniform model 204

6.3.1 Developing a non-uniform model for the specimen profile 204

6.3.2 Determining the specimen profile using a non-uniform model 205

6.3.3 Integral breadth, $\beta(d^*)$ 209

6.3.4 Fourier coefficients 210

6.3.5 Determining the column-length distribution – second attempt 211

6.3.6 Electron microscopy 216

6.4 Summary 218

7 Conclusion **221**

7.1 Evaluation of common deconvolution methods 222

7.2 Determining the column-length distribution 222

7.3 Determining the strain distribution 224

7.4 Analysis of SRM 676 x-ray diffraction data 226

7.5 Further research 227

7.5.1 Determining the specimen profile 227

7.5.2 Crystallite/domain size analysis 227
7.5.3 Strain analysis 228
7.5.4 Size/strain analysis 228

A Common Profiles **230**

B The effect of $P(D)$ on $\langle L \rangle_v$ & $\langle L \rangle_a$ **233**

B.1 $\langle L \rangle_v$ & $\langle L \rangle_a$ for a single size 233
B.2 $\langle L \rangle_v$ & $\langle L \rangle_a$ for a distribution 234

C The MaxEnt "all in one" approach **236**

C.1 Overview of the "all in one" approach 236

List of Figures

1.1 Schematic diagram of the Ewald sphere of radius $1/\lambda$. The incident beam is given by $\mathbf{S}_0$ and the scattered beam by $\mathbf{S}$. The angle between these vectors is 2θ and the difference between the vectors defines the diffraction vector by $\mathbf{S} - \mathbf{S}_0 = \mathbf{d}^*$, where $|\mathbf{d}^*| = 1/d$ and d is the atomic-interplanar distance in real space. The Bragg condition is satisfied when the sphere falls on the set of hkl reciprocal-lattice points (small circles). The reciprocal-lattice points have been enlarged to represent the spreading of intensity from a real crystal. For an ideal crystal the reciprocal-lattice points would indeed be represented by a point. Also note that the "cone of diffraction" for a semi-angle 2θ can be constructed from the figure. Adapted from Warren (1969). 4

1.2 A schematic diagram of the intensity region about O and P. As the radius, $|\mathbf{s}|$, increases (for increasing 2θ) an integration over the cross-section of the intensity takes place. This corresponds to the intensity measurement at given 2θ. Adapted from Guinier (1994). . . 6

1.3 A schematic digram of a crystallite (solid line) and its "ghost" (dashed line) shifted a distance L in the $[hkl]$ direction. The $V(L)$ for a crystallite can be related to the integral breadth (1.4) and size-Fourier coefficients (see §1.3.2). 9

2.1 A schematic plot showing the acceptable range for β: (i) When all the λ's are positive (given by crosses), then $0 < \beta < 2/\lambda_{max}^{(+)}$, where $|\lambda_{max}^{(+)}| = \lambda_{max,pos}$; (ii) When all the λ's are negative (given by circles), then $2/\lambda_{max}^{(-)} < \beta < 0$, where $|\lambda_{max}^{(-)}| = \lambda_{max,neg}$. 53

2.2 A schematic diagram showing the surface mapped out by the functions $S(\mathbf{f})$ and $C(\mathbf{f})$. $Q(\mathbf{f})$ is minimized for a given $\hat{\mathbf{f}}$ when the vectors ∇S and ∇C are parallel to each other. Here $\nabla S \equiv grad\, S$ and $\nabla C \equiv grad\, C$. 56

2.3 (a) Normalized specimen profile (solid lines) and instrument profile (dashed lines); (b) Observed profile with noise and background level added. 59

2.4 Plots of the solution from the unconstrained inversion/deconvolutions methods, given by (2.5) & (2.16), rspectivitely: (a) Showing the specimen profile (solid line) and the solution profile (dashed line), $R_f = 43.7\%$, for the matrix inversion method; (b) Showing the specimen profile (solid line) and solution profile (dashed line), $R_f = 43.5\%$, for the Stokes method. 61

2.5 Plots for the terms in (2.32): (a) Showing the misfitting term in (2.32), $E_f(k)$ vs k ; (b) Showing the noise amplification term in (2.32), $E_n(k)$ vs k, for a noise level at 1.6% ; (c) Showing the background error term in (2.32), $E_{\hat{b}b}(k)$ vs k, for a background level 7.5% of the peak maximum; (d) Showing the sum of terms on the right hand side of (2.32). . . . 62

2.6 Plots of the solution profiles from the iterative method (2.24): (a) Showing the specimen profile (solid line) and the solution profile for the optimum number of iterations $k = 100$, $R_f = 11.5\%$; (b) Showing the specimen profile (solid line) and solution profile for $k = 50$ iterations, $R_f = 21.8\%$ (short dash line) and $k = 350$ iterations, $R_f = 17.8\%$ (long dash line). 64

2.7 Plots of the terms in (2.42): (a) Showing the misfitting term in (2.42), $E_f(\alpha)$ vs α (solid line) and the misfitting error of the *a priori* model in (2.42), $E_m(\alpha)$ vs α (dashed line); (b) Showing the misfitting background error term in (2.42), $E_{\hat{b}b}(\alpha)$ vs α; (c) Showing the noise amplification term in (2.42), $E_n(\alpha)$ vs α, for a noise level of 1.6%; (d) Showing the sum of terms on the right hand side of (2.42). The minimum at $\alpha \approx 0.15$ corresponds to the optimum value α. 65

2.8 Plots of the solution profiles from the constrained method (2.37): (a) Showing the sample profile (solid line) and solution for $\alpha = 0.15$ and $R_f = 19.7\%$ (dash line); (b) Showing the specimen profile (solid line) and two other solutions, $\alpha = 0.005$, $R_f = 32.3\%$ (short dash line) and $\alpha = 5.0$, $R_f = 65.8\%$ (long dash line), respectively. 66

4.1 Actual particle size and column-length distributions for spherical crystallites of alumina: (a) Particle size distribution for $\langle D \rangle = 20\,\text{nm}$ (dashed line), 50 nm (solid line) and 100 nm (dots); (b) Column-length distributions for the corresponding particle distributions. All distributions are normalized for unit area. 98

4.2 Simulated size-broadened profile and theoretical Fourier coefficients: (a) Size-broadened profile for the 113 peaks, corresponding to particle sizes $\langle D \rangle = 20\,\text{nm}$ (dashed line), 50 nm (solid line) and 100 nm (dots), over the interval $(2\theta_B \pm 1)\,^\circ 2\theta$. The 226 has similar profiles which are not shown; (b) Theoretical Fourier coefficients for the 113 & 226 profiles, corresponding to particle sizes $\langle D \rangle = 20\,\text{nm}$ (dashed line), 50 nm (solid line) and 100 nm (dots). 106

4.3 Ideal and non-ideal instrument profiles for the 113 & 226 peaks in alumina: (a) The non-ideal instrument profiles for the 113 peak with 100 nm (solid line) and 500 nm (dots) size-broadening, respectively, and the ideal case (dash line); (b) The non-ideal instrument profiles for the 226 peak with 100 nm (solid line) and 500 nm (dots) size-broadening, respectively, and the ideal case (dashed line). 107

4.4 The 113 and 226 simulated ("observed") profiles corresponding to $\langle D \rangle = 20\,\text{nm}$ over the interval of $(2\theta \pm 1)\,^\circ 2\theta$: (a) The 113 profile; (b) The 226 profile. The observed profiles for $\langle D \rangle = 50\,\&100\,\text{nm}$ are not shown. 108

4.5 MaxEnt results for the optimum background level for the 113 peak corresponding to $\langle D \rangle = 20\,\text{nm}$: (a) True 113 specimen profile (solid line), solution specimen profile (dashed line) and lower- & upper-uncertainty regions (dots); (b) Theoretical Fourier coefficients (solid line), Fourier coefficients from the solution profile (dashed line), given in (a), and lower- & upper-uncertainty regions (dots); (c) True column-length distribution (solid line), the solution column-length distribution (dashed line), lower- & upper-uncertainty regions (dots), and the uniform *a priori* model (dash-dot lines); (d) As for (c), but using a non-uniform *a priori* model (dash-dot line). Note: The distributions showing the lower- & upper-uncertainty regions (dots) are not normalized for unit area, while the solution column length distribution and non-uniform *a priori* model are normalized for unit area. 110

4.6 MaxEnt results for the optimum background level for the 226 peak corresponding to $\langle D \rangle = 20\,\text{nm}$: (a), (b), (c), & (d) as in Figure 4.5. 113

4.7 R_f-values for the 113 (diamond) and 226 (squares) solution profiles, determined from the MaxEnt method for different background levels. 119

4.8 Apparent size results from different methods, determined from 113 & 226 solution profiles corresponding to $\langle D \rangle = 20\,\mathrm{nm}$ over a range of background levels: (a) Apparent sizes from the 113 solution profile using the integral breadth method $\langle L \rangle_v^I$, (diamonds + solid line), Fourier method, $\langle L \rangle_a^F$ (cross + dashed line), MaxEnt method using a uniform model, $\langle L \rangle_a^{M,u}$, (squares + small dashed line) and MaxEnt method using a non-uniform *a priori* model, $\langle L \rangle_a^{M,n}$, (triangles + dots); (b) Corresponding differences between the theoretical area- & volume-weighted sizes and apparents sizes given in (a); (c) Apparent sizes from 226 solution specimen profiles. Symbols as described in (a); (d) Showing the corresponding differences between the theoretical area- & volume-weighted sizes and apparents sizes given in (c). 122

4.9 MaxEnt results for the background level of +20% for the 113 peak corresponding to $\langle D \rangle = 20\,\mathrm{nm}$: (a) True 113 specimen profile (solid line), solution specimen profile (dashed line); (b) Theoretical Fourier coefficients (solid line), Fourier coefficients from the solution profile (dashed line), given in (a); (c) True column-length distribution (solid line), the solution column-length (dashed line) and the uniform *a priori* model (dashed line); (d) As for (c), but using a non-uniform *a priori* model (dash-dot line). Dotted line in the above represent the uncertainties in the respective MaxEnt calculations. 125

4.10 Maxent results for the 113 peak corresponding to $\langle D \rangle = 50\,\mathrm{nm}$ when a non-ideal (500 nm size-broadened) instrument profile is used in the deconvolution: (a) The true 113 specimen profile (solid line), solution specimen profile (dashed line); (b) Theoretical Fourier coefficients (solid line), Fourier coefficients from the solution profile (dashed line), given in (a); (c) True column-length distribution (solid line), the solution column-length distribution (dashed line) and the non-uniform *a priori* model (dashed-dotted line). Dotted lines in the above represent the uncertainty in the respective MaxEnt calculations. 129

5.1 Showing the real Fourier coefficients of the specimen profiles using Gaussian strain distributions for different values of d^* from 111 to 420 peaks. 144

5.2 Showing the real Fourier coefficients of the specimen profiles using non-Gaussian (Pearson-VII) strain distributions for different values of d^* corresponding from the 111 to 420 profiles. 146

5.3 Calibration plots for the parameters w and m defining the Pearson-VII profile generated from an (ideal) reference LaB_6 diffraction pattern: (a) w vs 2θ, where $2w = FWHM\,°2\theta$ for the low- (diamonds) & high- (crosses) angle sides of the peak; (b) m vs 2θ, for the low- (diamonds) & high- (crosses) angle sides of the peak. 148

5.4 The 111 and 420 simulated ("observed") profiles for the Gaussian strain: (a) The 111 observed profile over the interval $(2\theta_B \pm 2)°2\theta$; (b) The 420 observed profile over the interval $(2\theta_B \pm 5)°2\theta$. The observed profiles for the 200 to 331 lines are not shown. 149

5.5 The 111 and 420 simulated ("observed") profiles for non-Gaussian strain: (a) The 111 observed profile over the interval $(2\theta_B \pm 2)°2\theta$; (b) The 420 observed profile over the interval $(2\theta_B \pm 5)°2\theta$. The observed profiles for the 200 to 331 lines are not shown. 151

5.6 MaxEnt results for the 111 peak for Gaussian strain: (a) True 111 specimen profile (solid line), solution specimen profile (dashed line) and lower- & upper-uncertainty regions (dots); (b) Theoretical Fourier coefficients (solid line) using (5.21) for $d^* = 1/d_{111}$, Fourier coefficients from the solution profile (dashed line) given in (a), and lower- & upper-uncertainty regions (dots). 155

5.7 MaxEnt results for the 420 peak for Gaussian strain: (a) & (b) as in Figure 5.6 except for $d^* = 1/d_{420}$. 156

5.8 The MaxEnt strain distributions for Gaussian strain using a uniform *a priori* model for $L = 5$, 13 & 20nm over the region of $\epsilon_L \in [0,\, 2FWHM]$: (a) True strain distribution (solid line) for $L = 5$ nm, the solution strain distribution (dashed line), lower- & upper-uncertainty region (dots), and the uniform *a priori* model (dash-dot lines); (b) Solution distribution for $L = 13$ nm; (c) Solution distribution for $L = 20$ nm. Note: The distributions showing the lower- & upper-uncertainty regions (dots) are not normalised for unit area. 160

5.9 MaxEnt strain distributions for Gaussian strain using a non-uniform *a priori* model for $L = 5$, 13 & 20nm over the region of $\epsilon_L \in [0,\, 2FWHM]$: (a) True distribution (solid line) for $L = 5$ nm, the solution strain distribution (dashed line), lower- & upper-uncertainty region (dots), and the non-uniform *a priori* model (dash-dot lines); (b) Solution distribution for $L = 13$ nm; (c) Solution distribution for $L = 20$ nm. Note: The distributions showing the lower- & upper-uncertainty regions (dots) are not normalised for unit area, while the solution distribution and non-uniform *a priori* model are normalised for unit area. . . . 167

5.10 MaxEnt results for the 111 peak for non-Gaussian strain: (a) True 111 specimen profile (solid line), solution specimen profile (dashed line) and lower- & upper-uncertainty regions (dots); (b) Theoretical Fourier coefficients (solid line) using (5.21) for $d^* = 1/d_{111}$, Fourier coefficients from the solution profile (dashed line), given in (a) and lower- & upper-uncertainty regions (dots). 168
5.11 MaxEnt results for the 420 peak for Gaussian strain: (a) & (b) as in Figure 5.10 except for $d^* = 1/d_{420}$. 169
5.12 MaxEnt strain distributions for non-Gaussian strain using a uniform *a priori* model for $L = 10, 30$ & 50nm over the region of $\epsilon_L \in [0, 2FWHM]$: (a) True distribution (solid line) for $L = 10$ nm, the solution strain distribution (dashed line), lower- & upper-uncertainty region (dots), and the uniform *a priori* model (dash-dot lines); (b) Solution distribution for $L = 30$ nm; (c) Solution distribution for $L = 50$ nm. Note: The distributions showing the lower- & upper-uncertainty regions (dots) are not normalised for unit area. 171
5.13 MaxEnt strain distributions for non-Gaussian strain using a non-uniform *a priori* model for $L = 10, 30$ & 50nm over the region of $\epsilon_L \in [0, 2FWHM]$: (a) True distribution (solid line) $L = 10$ nm, the solution strain distribution (dashed line), lower- & upper-uncertainty region (dots), and the non-uniform *a priori* model (dash-dot lines); (b) Solution distribution for $L = 30$ nm; (c) Solution distribution for $L = 50$ nm. Note: The distributions showing the lower- & upper-uncertainty regions (dots) are not normalised for unit area, while the solution distribution and non-uniform *a priori* model are normalised for unit area. 173
5.14 Integral breadth, β, versus d^* for Gaussian and non-Gaussian strain: (a) The integral breadth data from the profiles (diamonds + error bars), theoretical curve (solid line) using (5.23), the Williamson-Hall (dashed lines) and quadratic fitting (short dashed lines) using(5.33); (b) The integral breadth data from the profiles (diamonds + error bars), theoretical curve (solid line) using (5.27) in (5.22), the Williamson-Hall (dashed lines) and quadratic fitting (short dashed lines) using (5.34). 177
5.15 Showing the $\langle\epsilon_L^2\rangle^{\frac{1}{2}}$ results for both the Warren-Averbach and MaxEnt methods: (a) Showing the theoretical $\langle\epsilon_L^2\rangle^{\frac{1}{2}}$ values from (5.4) and the $\langle\epsilon_L^2\rangle^{\frac{1}{2}}$ results from the Warren-Averbach method (diamonds); (b) Showing the theoretical $\langle\epsilon_L^2\rangle^{\frac{1}{2}}$ values from (5.4), the MaxEnt method results for the uniform (diamonds) and non-uniform (crosses) models. 179

5.16 Showing the $\langle\epsilon_L^2\rangle^{\frac{1}{2}}$ results for both the Warren-Averbach and MaxEnt methods: (a) Theoretical $\langle\epsilon_L^2\rangle^{\frac{1}{2}}$ values from (5.26) and the $\langle\epsilon_L^2\rangle^{\frac{1}{2}}$ results from the Warren-Averbach method (diamonds); (b) Theoretical $\langle\epsilon_L^2\rangle^{\frac{1}{2}}$ values from (5.26), the MaxEnt method results for the uniform (diamonds) and non-uniform (crosses) models. 181

5.17 The half-widths or σ_L of the MaxEnt strain distributions: (a) Half-width results from the theoretical curve using (5.25) (solid lines) and the MaxEnt strain distributions using a uniform model (diamonds + error-bars); (b) Half-width results from the theoretical curve using (5.25) (solid lines) and the MaxEnt strain distributions using a non-uniform model (diamonds + error-bars) 184

6.1 Calibration plots for the parameters $FWHM$ and m defining the split Pearson-VII profile generated from an SRM 660 (LaB_6) diffraction pattern: (a) $FWHM$ vs 2θ for the low- (diamonds) & high- (crosses) angle sides of the peak; (b) m vs 2θ, for the low- (diamonds) & high- (crosses) angle sides of the peak. 191

6.2 MaxEnt results for the 012&024 multiple-orders using a uniform model: (a) The 012 specimen profile (solid line) and lower- & upper-uncertainty regions (dots); (b) The 024 specimen profile (solid line) and lower- & upper-uncertainty regions (dots); (c) The 012 (solid line) and 024 (dashed line) Fourier coefficients. NOTE: The uncertainty regions for the Fourier coefficients are not shown, for clarity. 195

6.3 MaxEnt results for the 113&226 multiple-orders using a uniform model: (a) The 113 specimen profile (solid line) and lower- & upper-uncertainty regions (dots); (b) The 226 specimen profile (solid line) and lower- & upper-uncertainty regions (dots); (c) The 113 (solid line) and 226 (dashed line) Fourier coefficients. NOTE: The uncertainty regions for the Fourier coefficients are not shown, for clarity. 197

6.4 Integral breadth $\beta(d^*)$ versus d^* for the MaxEnt specimen profiles using a uniform model (diamonds + error bars) and the linear fit (dots). 198

6.5 The area-weighted size for the Fourier method (diamonds + error-bars), MaxEnt methods (crosses + error-bars), average Fourier size (solid line) and average MaxEnt size (dots). . 202

6.6 MaxEnt results for the 012&024 multiple-orders using a non-uniform model: (a) The 012 specimen profile (solid line), the specimen profile model (dashed line) and lower- & upper-uncertainty regions (dots); (b) The 024 specimen profile (solid line), the specimen profile model (dashed line) and lower- & upper-uncertainty regions (dots); (c) The 012 (solid line) and 024 (dashed line) Fourier coefficients. NOTE: The uncertainty regions for the Fourier coefficients are not shown, for clarity. 207

6.7 MaxEnt results for the 113&226 multiple-order using a non-uniform model: (a) The 113 specimen profile (solid line), the specimen profile model (dashed line) and lower- & upper-uncertainty regions (dots); (b) The 113 specimen profile (solid line), the specimen profile model (dashed line) and lower- & upper-uncertainty regions (dots); (c) The 113 (solid line) and 113 (dashed line) Fourier coefficients. NOTE: The uncertainty regions for the Fourier coefficients are not shown, for clarity. 209

6.8 Integral breadth $\beta(d^*)$ versus d^* for the MaxEnt specimen profiles using a non-uniform model (diamonds + error bars) and the linear fit (dots). 210

6.9 The area-weighted size for the Fourier method (diamonds + error-bars), MaxEnt method (crosses + error-bars), average Fourier size (solid line) and average MaxEnt size (dots). . 211

6.10 The MaxEnt column-length distributions for the 012&024 multiple-orders: (a) MaxEnt 012 column-length distribution (solid line), non-uniform *a priori* model (dashed line), and lower- & upper-uncertainty regions (dots); (b) MaxEnt 024 column-length distribution (solid line), non-uniform *a priori* model (dashed line), and lower- & upper-uncertainty regions (dots). 214

6.11 The MaxEnt column-length distributions for the 113&226 multiple-order: (a) MaxEnt 113 column-length distribution (solid line), non-uniform *a priori* model (dashed line), and lower- & upper-uncertainty regions (dots); (b) MaxEnt 226 column-length distribution (solid line), non-uniform *a priori* model (dashed line), and lower- & upper-uncertainty regions (dots). 215

6.12 SEM images of the alumina SRM 676 sample: (a) An overview of the alumina sample at 2500× magnification; (b) A close-up of the alumina sample at 7500× magnification. . . . 217

B.1 $\frac{\langle L\rangle_a}{\langle L\rangle_v}$ *versus* σ_D for spherical crystallites. In the limit of $\sigma_D \to 0$, $\frac{\langle L\rangle_a}{\langle L\rangle_v} \to \frac{8}{9}$ and the specimen can be considered as consisting of particle of the same shape and size. Numerical values of $\frac{\langle L\rangle_a}{\langle L\rangle_v}$ (diamonds) and the fitted curve (solid line) using a third order polynomial are shown. 235

C.1 MaxEnt "all in one" results for optimum background level for the 113 peak corresponding to $\langle D \rangle = 50\,\text{nm}$: True column-length distribution (solid line), the solution column-length distribution (dashed line), the lower- & upper-uncertainty regions (dots), and the uniform *a priori* model (dash-dot lines). 237

List of Tables

4.1 Average diameters, $\langle D \rangle$, determined using the u, t & r parameters in (4.14); the theoretical area- & volume-weighted sizes, $\langle L \rangle_a^{Th}$ & $\langle L \rangle_v^{Th}$ respectively, determined from (4.16), are also shown. 100

4.2 Apparent size results from the 113 solution profile obtained using different instrument profiles as kernels in the deconvolution of the simulated profile. Shown are the R_f for each solution profile, apparent sizes, $\langle L \rangle_v^I$, $\langle L \rangle_a^F$, $\langle L \rangle_a^{M,u}$ and $\langle L \rangle_v^{M,n}$, and differences between theoretical and calculated results for different particle distributions corresponding to $\langle D \rangle =$ 20, 50 & 100 nm. The superscripts, I, F, (M, u) and (M, n) mean Integral, Fourier, MaxEnt with uniform prior model and MaxEnt with a non-uniform prior model, respectively. . . . 114

4.3 Apparent size results from 226 solution profile obtained using different instrument profiles as kernels in the deconvolution of the simulated profile. Shown are the R_f for each solution profile, apparent sizes, $\langle L \rangle_v^I$, $\langle L \rangle_a^F$, $\langle L \rangle_a^{M,u}$ and $\langle L \rangle_v^{M,n}$, and differences between theoretical and calculated results for different particle distributions corresponding to $\langle D \rangle =$ 20, 50 & 100 nm. Notation is as in Table 4.2. 115

5.1 The parameters and results from applying the MaxEnt method to determine the specimen profile for the Gaussian case: C/C_{aim} is the ratio of the final to the preferred statistic value; R_f is the R-factor from (5.29); $\beta(°2\theta)$ is the integral breadth of the solution profile over 2θ-space; and $diff$ is the percentage difference between the theoretical and calculated integral breadth of the specimen profile. 154

5.2 Parameters and results from applying the MaxEnt method to determine the specimen profile for the non-Gaussian case: C/C_{aim} is the ratio of the final to the preferred statistic value; R_f is the R-factor from (5.29); $\beta(°2\theta)$ is the integral breadth of the solution profile over 2θ-space; and $diff$ is the percentage difference between the theoretical and calculated integral breadths of the specimen profile. 163

5.3 Estimated parameters using the nonlinear Williamson & Hall (WH), Warren & Averbach (WA) and MaxEnt methods: $\text{WH}^{(1)}$ G^2 & c estimated for all data points in Figure 5.14(a); $\text{WH}^{(2)}$ G^2 & c estimated for the first four data points in Figure 5.14(a); $\text{WA}^{(1)}$, G^2 estimated over the region $L \in [1, 34]$ nm from Figure 5.15(a); $\text{WA}^{(2)}$, G^2 is estimated over the region $L \in [5, 22]$ nm from Figure 5.15(a); $\text{MaxEnt}^{(1)}$, G^2 from Figure 5.15(b) using a uniform model in the MaxEnt method; $\text{MaxEnt}^{(2)}$, G^2 from Figure 5.15(b) using a non-uniform model in the MaxEnt method. 175

5.4 The estimated values of σ_0, m, q and percentage difference *diff* for various methods: the Warren & Averbach, (WA); MaxEnt method using a uniform model, $\text{MaxEnt}^{(1)}$; MaxEnt method using a uniform model, $\text{MaxEnt}^{(2)}$. 182

5.5 The estimated values for σ_0 and q using (5.25 from Figures 5.12 & 5.13: $\text{MaxEnt}^{(1)}$ the σ_0 and q results from the MaxEnt method using a uniform model; $\text{MaxEnt}^{(2)}$ the σ_0 and q results from the MaxEnt method using a non-uniform model. 185

6.1 The *a priori* model parameters, $\langle L \rangle_a^m$ and μ_m for hkl, determined by fitting (6.5) to the results of the "low resolution" method. 203

6.2 The model parameters, $\langle L \rangle_a^m$ and μ, for hkl, determined by fitting (6.8) to the Fourier coefficients. 205

Abstract

The analysis of broadened x-ray diffraction profiles provides a useful insight into the structural properties of materials, including crystallite size and inhomogenous strain. In this work a general method for analysing broadened x-ray diffraction profiles is developed. The proposed method consists of a two-fold maximum entropy (MaxEnt) approach.

Conventional deconvolution/inversion methods presently in common use are analysed and shown not to preserve the positivity of the specimen profile; these methods usually result in ill-conditioning of the solution profile. It is shown that the MaxEnt method preserves the positivity of the specimen profile and the underlying size and strain distributions, while determining the maximally noncommital solution. Moreover, the MaxEnt method incorporates any available *a priori* information and quantifies the uncertainties of the specimen profile and the size and strain distributions.

Numerical simulations are used to demonstrate that the MaxEnt method can be applied at two levels: firstly, to determine the specimen profile, and secondly to calculate the size or strain distribution, as well as their average values. The simulations include both size- and strain-broadened specimen profiles. The experimental conditions under which the data is recorded are also simulated by introducing instrumental broadening, a background level and statistical noise to produce the observed profile. The integrity of the MaxEnt results is checked by comparing them with the traditional results and examining problems such as deconvolving in the presence of noisy data, using non-ideal instrument profiles, and the effects of truncation and background estimation in the observed profile. The MaxEnt analysis is also applied to alumina x-ray diffraction data.

It is found that the problems of determining the specimen profile, column-length and strain distributions can be solved using the MaxEnt method, with superior results compared with traditional methods. Finally, the issues of defining the *a priori* information in each problem and correctly characterising the instrument profile are shown to be critically important in profile analysis.

Chapter 1

Literature Review

1.1 Introduction

The development of x-ray powder diffraction in the 1930s and 1940s focussed inpart on the investigation of line-profile broadening (see James 1948), with the realisation that structural imperfections in material specimens caused this broadening. Moreover, by relating the specimen broadening to physical properties of the material specimen, it was proposed that a microscopic "picture" encompassing size, distribution and shape of crystallites, as well as distortions of the crystallite lattice, could be developed.

The implications of developing such an understanding at a microscopic level are immense for both scientific and commercial reasons. For example, x-ray line-broadening provided the first indirect evidence for dislocations in metals (Nabarro et al. 1964). X-ray profile analysis plays an important role in industry. The commercial implication are far reaching, including metallurgy, the strength of metals, alloys and ceramics, understanding the role of dislocations in superconductors, and the fabrication of semiconductors for the computer industry.

The problem of developing a "picture" of structural imperfections is made more difficult by the method of recording intensity distributions, in that the intensity distribution in a three dimensional reciprocal space is mapped into a one-dimensional space. A consequence of this "rolling up" or mapping of dimensions when applied to polycrystalline materials is an averaging over the shape, size and internal distortions. In order to develop a picture of the microscopic properties of the material, we must work backwards from the one-dimensional

space into a three-dimensional space. Generally this transformation does not produce a unique solution and many possible solutions may exit. The problem we face is determining the most plausible or probable of these many possible solutions. This is a problem faced not only in determining the specimen profile, but also the crystallite/domain size distribution and the strain distribution. It is the central problem investigated in this study.

In this chapter an overview of specimen profile broadening is presented and described in reciprocal space (§1.2). The two dominant representations, *viz.* the integral breadth and Fourier representations, for broadened specimen profiles are presented and the physical quantities are defined using this representation (§1.3). The process of determining the specimen profile, crystallite size distributions, and strain distribution from the observed data is discussed by drawing on the existing literature. This will map a path to recent developments and an appreciation for the assumptions, advantages and limitations of the accepted methods (§1.4). More importantly, this lays the necessary foundations to make some observations concerning the underlying problems of determining the specimen profile, crystallite/domain size and strain distribution. On the surface, these problems look different, however it is shown that they can be solved using a unified approach.

1.2 An overview of size- and strain-induced x-ray profile broadening

1.2.1 Conditions for diffraction — ideal case

The condition for Bragg diffraction is best formulated in a reciprocal space by constructing an Ewald sphere. There are many ways to describe the reciprocal space, but here it is useful to think of it as a Fourier transform of the real (atomic) space into an inverse space where a three-dimensional intensity distribution is produced about reciprocal space points hkl. As a first case, the condition for diffraction is established for an ideal crystal, infinite in size and containing no distortions. In the simple picture, the atoms can be represented by delta-functions and in the reciprocal space, the intensity distributions are also represented by delta-functions located at hkl (Guinier 1963).

Suppose we define the incident beam by a vector $\mathbf{S}_0$, where $|\mathbf{S}_0| = 1/\lambda$ and λ is the

wavelength of radiation (see Figure 1.1); $\mathbf{S}_0$ is from the point N such that the end of the vector defines the origin, O. About N a sphere (the Ewald sphere) can be constructed with a radius $|\mathbf{S}_0|$. The vector $\mathbf{S}$ in Figure 1.1 is the scattered beam taken from N to the reciprocal space point, hkl, at P. The scattered and incident beams are joined by the vector $\mathbf{d}^*$. The condition for diffraction is established only if the set of reciprocal space points, hkl, fall on the surface of the Ewald sphere. Given this condition, the following is true

$$\begin{aligned} |\mathbf{S} - \mathbf{S}_0| &= |\mathbf{d}^*| \\ &= |h\,\mathbf{a}_1^* + k\,\mathbf{a}_2^* + l\,\mathbf{a}_3^*| \\ &= \frac{1}{d} \end{aligned} \tag{1.1}$$

where $\mathbf{d}^*$ is called the *diffraction vector*; its magnitude is defined by a set of reciprocal space basis vectors, $\{\mathbf{a}_i^*;\, i = 1, 2, 3\}$ and reciprocal space points hkl; d is the atomic interplanar distance. In (1.1), the magnitude of $\mathbf{d}^*$ must be equal to the inverse of the atomic interplanar distance, d, for diffraction to occur. In the real space of the crystal lattice, the diffraction vector, $\mathbf{d}^*$, is perpendicular to the diffraction planes, hkl, and we can think of these planes as “reflecting” the x-rays.

Under these conditions the angle between the incident and scattered vectors is $2\theta = 2\theta_B$, where $2\theta_B$ is the Bragg angle. From the geometry in Figure 1.1, the length of $\mathbf{d}^*$ is

$$|\mathbf{d}^*| = \frac{2\,\sin\theta_B}{\lambda}. \tag{1.2}$$

By equating the last term of (1.1) with (1.2), the Bragg equation for diffraction follows.

1.2.2 Diffraction from real materials

In the previous section the material was assumed to be an ideal crystal, and it was shown that the diffraction pattern would consist of delta-functions positioned at $2\theta_B$. It is more interesting is to consider the conditions that give rise to broadened x-ray diffraction profiles.

Real materials and structural imperfections

For a real material the intensity distribution in reciprocal space is spread about the point hkl (Langford & Louër 1996). In Figure 1.1, the spreading of the intensity distribution is

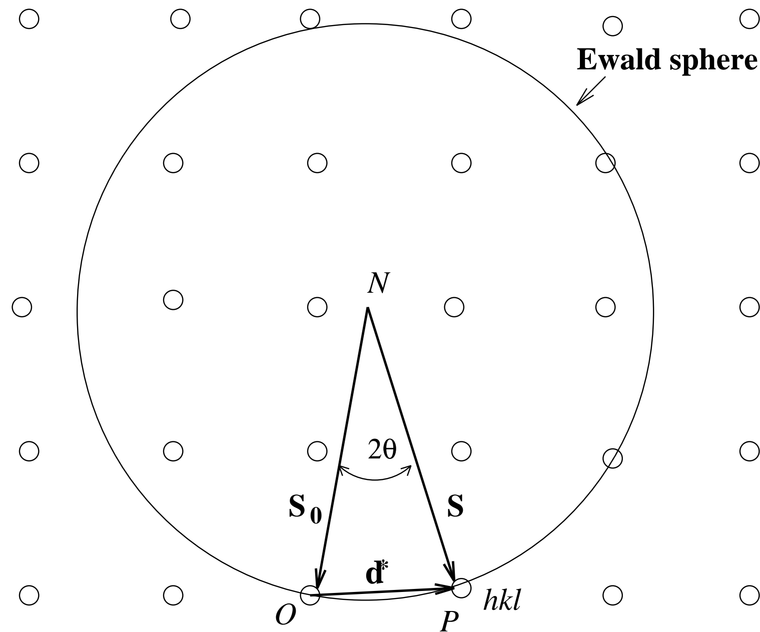


Figure 1.1: Schematic diagram of the Ewald sphere of radius $1/\lambda$. The incident beam is given by $\mathbf{S}_0$ and the scattered beam by $\mathbf{S}$. The angle between these vectors is 2θ and the difference between the vectors defines the diffraction vector by $\mathbf{S} - \mathbf{S}_0 = \mathbf{d}^*$, where $|\mathbf{d}^*| = 1/d$ and d is the atomic-interplanar distance in real space. The Bragg condition is satisfied when the sphere falls on the set of hkl reciprocal-lattice points (small circles). The reciprocal-lattice points have been enlarged to represent the spreading of intensity from a real crystal. For an ideal crystal the reciprocal-lattice points would indeed be represented by a point. Also note that the "cone of diffraction" for a semi-angle 2θ can be constructed from the figure. Adapted from Warren (1969).

schematically represented by circles. The spreading or broadening characterises the structural imperfections of materials[1].

There are various classes of structural imperfection, such as finite size broadening, inhomogeneous strain and dislocations, which give rise to the broadening of the intensity distribution about hkl (Langford & Louër 1996). Finite size broadening refers to Bragg

[1]Here polycrystalline materials or powders are considered. Polycrystalline materials are thought of as simply interlocking single crystals whose crystallographic axes are randomly distributed, while powders are discrete particles in the range of $0.1 - 1000\,\mu$m.

scattering from crystallites or domains in the direction of $\mathbf{d}^*$. This scattering can be related to an average dimension of the crystallite or domain in the direction of $\mathbf{d}^*$, but is independent of the order of diffraction. The domains may be a consequence of discontinuities or boundaries in the crystal lattice; for example, small grain-boundaries produce size broadening and the corresponding dimension represents the distance between boundaries (Langford & Louër 1996). For inhomogenous strain there is distortion of the atomic lattice (twisting, bending etc.) which causes a variation in the d-spacings (Stokes & Wilson 1944a, Langford & Louër 1996) and is dependent on the order of diffraction. In the case of dislocations, both size and strain broadening results: the size broadening may define the mean length between regions of low and high dislocation densities, while the inhomogenous strain is a result of strain fields from the dislocations (van Berkum et al. 1996, Langford & Louër 1996).

Broadened diffraction profiles

Given that structural imperfections produce the spreading of the intensity region in reciprocal-space, it is worth discussing how diffraction profiles, are produced. In Figure 1.2, a schematic representation of an intensity distribution about the hkl point is given. In many ways this can be thought of a "close up" of Figure 1.1.

As was shown in §1.2.1, the distance of OP is $|\mathbf{d}^*|$. Suppose a vector $\mathbf{s}$ is defined at O and is rotated in all directions, defining a sphere of radius $|\mathbf{s}| = \frac{2\sin\theta}{\lambda}$ about O, given by Q in Figure 1.2. As the radius of the sphere increases, it cuts the intensity distribution about P and an integration of the spherical cross-sections takes place. This can be replaced by a plane tt' normal to $\mathbf{d}^*$ and the integration approximated over the planar region that cuts the intensity distribution about hkl (Guinier 1963). For $2\theta < 2\theta_B$, $s - d^* < 0$ and the intensity distribution increases; for $2\theta = 2\theta_B$, $s - d^* = 0$, and the intensity is a maximum, since it corresponds to the maximum cross-section of the intensity distribution; for $2\theta > 2\theta_B$, $s - d^* > 0$ and the intensity distribution decreases. In other words, the diffraction profile is centered about d^* and the intensity measurement at a given 2θ on a diffraction profile is a result of the integration over the sphere surface, Q. Using this representation it can be seen how the diffraction profile from polycrystalline materials can be an averaging of shape, size and lattice distortions. This says that the ensemble of crystallites with varying

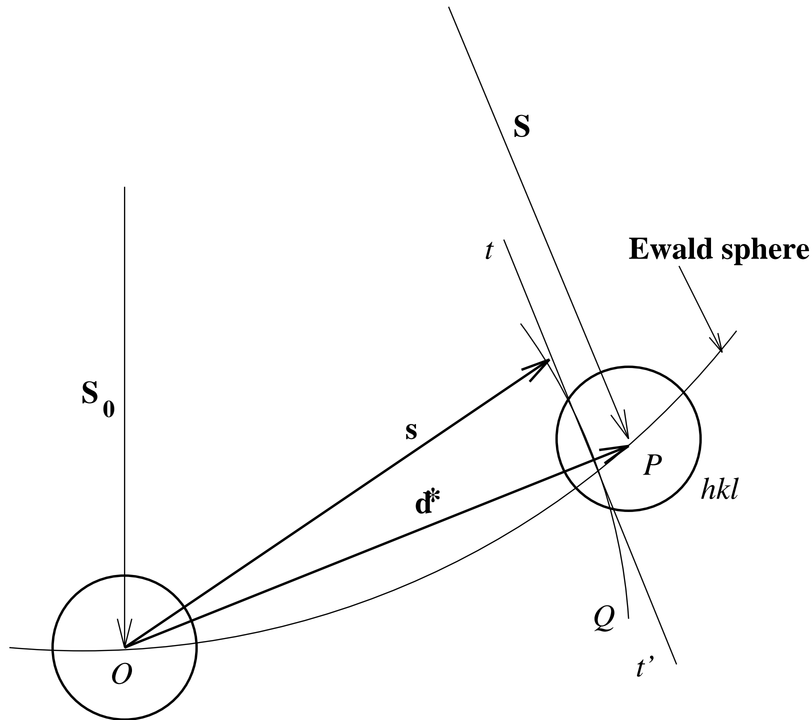


Figure 1.2: A schematic diagram of the intensity region about O and P. As the radius, $|\mathbf{s}|$, increases (for increasing 2θ) an integration over the cross-section of the intensity takes place. This corresponds to the intensity measurement at given 2θ. Adapted from Guinier (1994).

size and shapes and internal distortions contributes to the spreading of the intensity region about hkl. As the sphere Q cuts the intensity region and the integration over its surface is carried out, an averaging of the various contributions from their ensemble is taking place (Guinier 1963, Krivoglaz 1995).

The vital issue that must now be addressed is how to quantify the broadening of the diffraction profiles and relate the measure of broadening to the structural imperfections.

1.3 Size and strain broadening of x-ray diffraction profiles

There are two main approaches to quantifying the specimen broadening of the diffraction profile presented here. The first approach is the integral-breadth representation, which simply provides a measure of the broadening in terms of the area of the profile. It can be considered as a semi-quantitative approach as it does not take into consideration the shape of the line profiles. The second approach is the Fourier representation of the profile. This approach decomposes the line-profile into Fourier coefficients and relates them to broadening from the structural imperfections. By incorporating information concerning the shape and broadening of the profiles, it can be used to make quantitative predictions concerning the shape of diffraction profiles. The integral-breadth and Fourier representations of profile broadening have become the central means of analysing size-, size/strain- and dislocation-broadened profiles and other methods for profile analysis draw on these representations.

1.3.1 Integral-breadth representation of size- & strain-broadened profiles

Integral-breadth description of size-broadening

Stokes & Wilson (1942) expressed the broadening of a line profile from the diffraction of small crystallites in terms of their shape. The broadening was quantified by the integral breadth of the profile (area of profile/profile maximum) and the common-volume (or "ghost") function as

$$\beta_{2\theta} = \frac{\lambda}{\cos\theta_B} \frac{V(0)}{\int V(L)\,\mathrm{d}L} \tag{1.3}$$

where $\beta_{2\theta}$ is the integral breadth determined from the profile in 2θ-space (in units of radians); λ is the wavelength of radiation; $V(0)$ is the volume of the crystallite; and $V(L)$ is the common-volume of the crystallite shifted a distance L, parallel to the diffraction vector, $\mathbf{d}^*$. The limits of integration in (1.3) are over $-L_{max} \leq L \leq L_{max}$, such that $V(L_{max}) = 0$. The integral breadth can be expressed in reciprocal-space units (the preferred usage in this

study) as

$$\beta = V(0)\left[\int_{-L_{max}}^{L_{max}} V(L)\,\mathrm{d}L\right]^{-1}. \tag{1.4}$$

This approach provides a physically meaningful representation of size broadening, in that the shape of the crystallite, given by $V(L)$, has been included. In principle the average dimensions of the crystallite can be determined if enough β-values for a set of hkl-lines can be determined. This has recently been established (for example see Louër et al. 1983, Vargas et al. 1983) and the procedure will be discussed in §1.4. Alternatively the integral breadth can be expressed in terms of the cube-root of the crystallite-volume, sometimes known as the "true" size (Langford & Wilson 1978) *via* the Scherrer constant, which is also dependent on a model for the crystallite shape, expressed as

$$\beta = \frac{\mathcal{K}}{\mathfrak{p}} \tag{1.5}$$

where β is in reciprocal space units; $\mathcal{K}$ is the Scherrer constant and is of order unity; $\mathfrak{p}$ is the cube-root of the crystallite volume. The Scherrer constant is also dependent on a model for the crystalites. Stokes & Wilson (1942) pointed out that determining $\mathcal{K}$ was mathematically demanding, but by determining the common-volume function and relating it to the integral breadth, they in effect "side-stepped" the need to determine the the Scherrer constant.

The integral-breadth, (1.4), can be related to the mean thickness of the crystallite perpendicular to the (hkl)-plane. Moreover, its physical interpretation was only correctly defined by Stokes & Wilson (1942). Using (1.4), we can define the volume-averaged crystallite size as

$$\langle L\rangle_v = \beta^{-1}. \tag{1.6}$$

$\langle L\rangle_v$ is an apparent dimension of the crystallite parallel to the diffraction vector (i.e. perpendicular to the (hkl) planes) and using a suitable crystallite model its overall dimensions can be determined. Figure 1.3 shows a schematic representation of the crystallite and its "ghost" shifted a distance L perpendicular to the (hkl) planes.

By arriving at (1.6), Stokes & Wilson (1942) considered the integral breadth for various shapes. However, they also assumed that the crystal lattice was cubic and all crystallites had the same dimensions. In other words, the size distribution was assumed to be a delta-function. The analysis demonstrated that in the case of size broadening, the integral-breadth

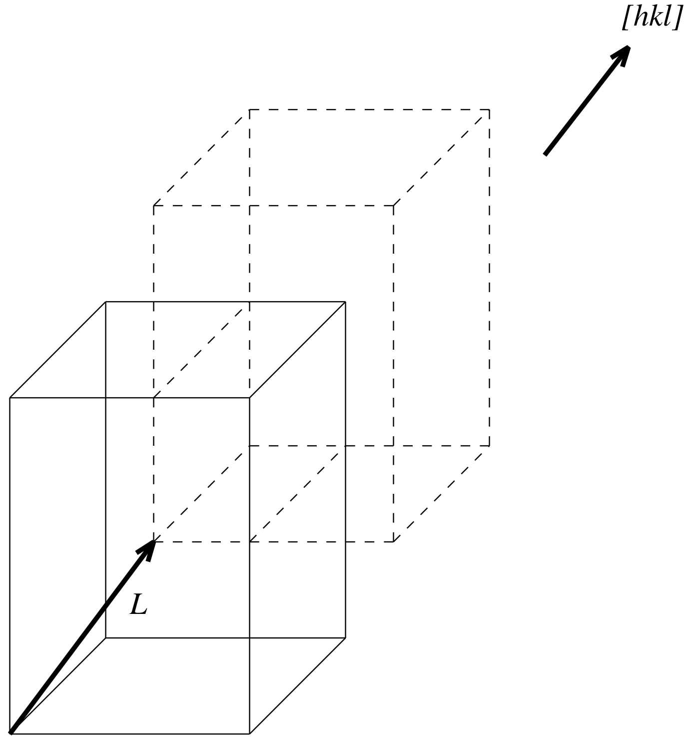


Figure 1.3: A schematic digram of a crystallite (solid line) and its "ghost" (dashed line) shifted a distance L in the $[hkl]$ direction. The $V(L)$ for a crystallite can be related to the integral breadth (1.4) and size-Fourier coefficients (see §1.3.2).

and $\langle L\rangle_v$ were in general dependent or the crystallographic direction, $[hkl]$ and were expressed in terms of the hkl-indices. The shapes considered were regular parallelepipeds, tetrahedra and octahedra. Spherical crystallites were shown to be the exception, in that $\langle L\rangle_v$ is the same (i.e. isotropic) in all $[hkl]$-directions. Stokes & Wilson (1944b) generalised their results for non-cubic lattice-symmetries.

Common-volume function and apparent size

Following the developments of Stokes & Wilson (1942), (1.4) and (1.6) could be used to express $\langle L\rangle_v$ in term of the dimensions of the crystallite. A simple example is a sphere, when $V(L)$ is given by (Stokes & Wilson 1942) as

$$V(L) = \frac{\pi}{12}(D-L)^2(2D+L) \tag{1.7}$$

where D is the diameter of the sphere. On applying (1.4) and (1.6) the volume-weighted size is given as

$$\langle L \rangle_v = \frac{3}{4} D. \tag{1.8}$$

The common-volume of a crystallite can be related to another measure of the crystallite thickness in the direction of the diffraction vector. This dimension is the area-weighted size, $\langle L \rangle_a$ and is given by

$$\frac{1}{\langle L \rangle_a} = -\frac{1}{V(0)} \left(\frac{\mathrm{d}V(L)}{\mathrm{d}L} \right)_{L=0} \tag{1.9}$$

where $\langle L \rangle_a$ is interpreted as the ratio of the crystallite volume projected onto the crystallographic plane (hkl) (Langford & Wilson 1978, Guinier 1963). It can be shown that in general $\langle L \rangle_a$ will be less than $\langle L \rangle_v$. Like $\langle L \rangle_v$, $\langle L \rangle_a$ is an apparent dimension which can be related to the dimensions of the crystallite if a model is assumed. In the case of a sphere, (1.9) can be applied to (1.7) resulting in

$$\langle L \rangle_a = \frac{2}{3} D \tag{1.10}$$

and the ratio of the $\langle L \rangle_a$ to $\langle L \rangle_v$ is $\frac{8}{9}$ (see Appendix B).

Similarly $\langle L \rangle_v$ and $\langle L \rangle_a$ can be related to the thickness parallel to the diffraction vector, if the Scherrer constant is known for each measure of size. Langford & Wilson (1978) reviewed the Scherrer constants for $\langle L \rangle_v$, $\langle L \rangle_a$ using simple shapes. The constants were represented as contour maps. This representation allowed the extrema and saddlepoints of the Scherrer constant to be determined. The resulting diffraction profiles were also determined and compared for different shapes, for a given hkl. It was also shown that an equivalent Scherrer constant can be determined when there exists a distribution of crystallites.

Wilson (1962) and Lele & Anantharaman (1966) developed the Scherrer constants and common-volume functions for various shapes in terms of the indices hkl. They used various crystal lattice systems, such as the hexagonal crystal system for triangular and hexagonal prisms and the tetragonal system for square prisms and cylinders. In order for these results to be applied to other crystal systems, the correct transformations must be applied (Langford & Wilson 1978). Building on the results of Lele & Anantharaman (1966), Wilson (1969) developed the common-volume function for cylinders, prisms and hemispheres. Rather than deriving the common-volume function in terms of hkl, they were expressed in terms of the

angle, φ, between the axis of the crystallite (i.e. the [001] direction) and the diffraction vector. The angle φ depends on the lattice system but this approach is more general and practical.

The application of these common-volume functions to experimental data, where the average dimensions and thickness could be determined, has come quite recently. The notable developments of applying the common-volume functions to experimental data are by Louër et al. (1972), Auffredic et al. (1980), Langford & Louër (1982), Vargas et al. (1983), Louër et al. (1983), Langford (1992) and Langford et al. (1993). Langford & Louër (1982), Vargas et al. (1983) and Louër et al. (1983) applied the common-volume for the cylindrical and hexagonal crystallites to ZnO. The cylindrical case is an approximation to the hexagonal crystallites (Vargas et al. 1983, Louër et al. 1983). Grebille & Bérar (1985) determined the common-volume function for convex-polyhedra and applied it to boehmite (AlOOH) crystallites. This is a particularly challenging problem due to the irregular edges of the polyhedron, but showed that the volume can be broken down into elementary prisms. Recently, Audebrand et al. (1998) have used x-ray techniques to examine the microstructure of nanocrystalline ZnO powders by using $V(L)$ for a cylindrical shape and showing how the dimensions of the crystallite change with temperature.

Integral breadth description of strain broadening

Stokes & Wilson (1944*a*) attempted to relate the integral breadth to the inhomogeneous strain by defining the apparent strain, η. This was analogous to the integral-breadth description of size broadening (1.4). This analysis was based on the broadening of diffraction patterns (Debye-Scherrer lines) from cold-worked metals[2]. Stokes & Wilson (1944*a*) noted that there may be three factors which contribute to the broadening of x-ray diffraction patterns from cold worked metal: (i) scattering domains of size $\lesssim 0.1\mu$m; (ii) grain boundaries having different lattice parameters; (iii) the distortion of large crystallites, $\sim 1\mu$m. In (i), domain-size broadening would result and would be independent of the order of diffraction, d^*. The distinction between points (ii) and (iii), is that in (ii) the different lattice parameters are due to homogeneous strain being applied to the specimen which causes a uniform

[2]The process of cold-worked hardening of metals involves controlling the plastic deformation of a metal well below the recrystallisation temperature (Schlenker 1986).

distortion of the lattice in a particular direction. The diffraction pattern would reveal a shifting of the peak positions as a consequence of the change in lattice parameters. In (iii) the applied strain causes a non-uniform distortion of the crystal lattice (i.e. bending, twisting etc.). This results in the broadening of the intensity region about hkl in reciprocal-space and consequently a broadening of the diffraction profile.

It was also pointed out by Stokes & Wilson (1944a) that the homogeneous and inhomogeneous strain would be dependent on the order of diffraction, d^*. The analysis of Stokes & Wilson (1944a) centered on (iii). However, their treatment can be considered as a first approximation of inhomogeneous deformation caused by dislocations, in that it assumes that a specimen has experienced different homogeneous deformation (Ryaboshapka & Tikhonov 1961a). Their analysis did not take into consideration the strain gradient and its dependency on the distance perpendicular to the diffraction planes.

The integral breadth from a strain-broadened profile was shown by Stokes & Wilson (1944a) to be related to the apparent strain, η, by

$$\eta = \beta_{2\theta} \cot \theta_B \tag{1.11}$$

$$= 2\beta / d^* \tag{1.12}$$

where $\beta_{2\theta}$ is the integral breadth in 2θ-space (expressed in units of radians); β is the strain in reciprocal-space units.

Like the volume-averaged crystallite size, $\langle L \rangle_v$, η in (1.11 & 1.12) is an apparent quantity and is considered as a measure of the maximum strain (Klug & Alexander 1974). This was realized by Stokes & Wilson (1942) and two models were proposed: (i) assumed that all the strain values, between zero and the maximum, were equally probable; (ii) the mean-square strain was drawn from a Gaussian distribution. Model (ii) is appropriate and useful, in that it relates η in terms of the root-mean-square strain drawn from a Gaussian strain distribution with a zero mean. Stokes & Wilson (1944a) expressed the apparent strain in terms of the root-mean-square strain as

$$\eta = 2\sqrt{2\pi \langle \epsilon^2 \rangle}$$

$$\approx 5 \langle \epsilon^2 \rangle^{\frac{1}{2}}. \tag{1.13}$$

The physical interpretation of this model is that the "distortions" are randomly distributed

in the specimen, constant and isotropic at all distances and directions in $[hkl]$. Generally, the assumption that the strain distribution is isotropic in $[hkl]$ is not valid since most materials are elastically anisotropic. The $\langle \epsilon^2 \rangle^{\frac{1}{2}}$ in (1.13) really defines the order of magnitude of the rms-strain.

When Stokes & Wilson (1944*a*) presented their analysis of strain broadening, there was no direct evidence of dislocations and it was realised that these results need to be generalized to include this theory (Wilson 1952). This problem focused essentially on cold-worked metals, where dislocation theory explained many of the properties of cold-worked metals and the broadening observed in the x-ray diffraction profiles (for example see Nabarro 1952, Nabarro et al. 1964). Wilson (1949, 1952), made the initial attempts at incorporating models of bent lammellae (Wilson 1949) and screw dislocations (Wilson 1952) into a kinematic theory of x-ray diffraction. As a result of applying dislocation theory in explaining the broadening of diffraction profiles, competing interpretations emerged (Williamson & Hall 1953, McKeehan & Warren 1953, Williamson & Smallman 1954, Williamson & Smallman 1956, Warren 1959). The debate centred around the nature of the strain distributions and domain/strain broadening from plastically deformed metals.

Krivoglaz & Ryaboshapka (1963) and Krivoglaz (1995) showed that a Gaussian profile resulted from randomly distributed screw dislocations from a cylindrical crystal. Moreover, they demonstrated that β is proportional to $\tan\theta$ and established that in general it would be dependent on the orientation factor, Γ, and the direction of the scattering vector. Ryaboshapka & Tikhonov (1961*a*, 1961*b*) pointed out that the rms-strain would be dependent on the density of dislocations, elastic properties of the material and Γ (e.g. Wilkens 1970*a*, Wilkens 1970*c*, Wilkens 1987). In other words, the rms-strain is a function of the ensemble of dislocations (or "distortions") in the material. van Berkum et al. (1996) showed that β depends strongly on the relative width of the strain field of dislocations. It was shown that for $w_r \to \infty$, (1.12) became applicable; however as $w_r \to 0$, $\beta \propto d^{*2}$. Ungár, Ott, Sanders, Borbély & Weertman (1998), Ungár, Révész & Borbély (1998) and Ungár & Borbély (1996), have shown that β is a function of $d^* \langle C \rangle^{\frac{1}{2}}$, where $\langle C \rangle$ is the contrast (or orientation) factor of the material which describes the broadening for isotropic and anisotropic broadening of line profiles. Wu, Gray & Kisi (1998) and Wu, Kisi & Gray (1998) adapted

the work of Krivoglaz & Ryaboshapka (1963) and Wilkens (1970a) for Rietveld refinement of powder patterns where the dislocation density and slip system can be determined by refining the integral breadth of the profile on a shape parameter.

An alternative way to describe the broadening of the line profile due to dislocation fields and crystallite size is to decompose the specimen profile in terms its Fourier coefficients.

1.3.2 Fourier representation of size- & strain- broadened profiles

Fourier representation of the specimen profile

The Fourier representation was developed by Averbach & Warren (1949) and Warren & Averbach (1950, 1952) and it essentially associated size and strain broadening with the Fourier coefficients independent of d^* in the case of size broadening and dependent d^* in case of strain broadening. By defining the coefficients in this manner, the basis for separating size and strain broadening can be established. It also enables the instrument broadening to be removed by applying methods such as those developed by Stokes (1948) which decompose the observed profile into the Fourier coefficients of the instrument and specimen profiles. The size/strain-broadened specimen profile can be expressed as

$$f(s - s_0,\, d^*) = \sum_{L=-\infty}^{+\infty} A^s(L)\, C^d(L, d^*)\, \exp 2\pi\imath(s - s_0)L \tag{1.14}$$

where $A^s(L)$ and $C^d(L,\, d^*)$ are the exponential Fourier coefficients for the size and distortion (or strain) contribution, respectively; $\imath = \sqrt{-1}$; $s_0 \equiv d^*$ such that $f(s - s_0,\, d^*)$ is centred about d^*. In applying the summation, we note that $C^d(L, d^*) = A^d(L,\, d^*) + \imath B^d(L,\, d^*)$ where $A^d(L,\, d^*) = A^d(-L,\, d^*)$ and $B^d(-L,\, d^*) = -B^d(L,\, d^*)$, and $A^s(L) \in \mathbb{R}$ for all $|L| > 0$; (1.14) reduces to

$$\begin{aligned} f(s - s_0,\, d^*) &= \sum_{L=-\infty}^{+\infty} A^s(L)\, A^d(L,\, d^*)\, \cos 2\pi(s - s_0)L \\ &+ \sum_{L=-\infty}^{+\infty} A^s(L)\, B^d(L,\, d^*)\, \sin 2\pi(s - s_0)L \end{aligned} \tag{1.15}$$

where $A^d(L,\, d^*)$ and $B^d(L\, d^*)$ are the real and imaginary strain coefficients, respectively. By decomposing the specimen profile in terms of Fourier coefficients, the shape of the profile is taken into consideration rather than just its width or breadth.

Equation (1.14) represents the inverse Fourier transform of the product $A^s(L)\,C^d(L,\,d^*)$. Applying the convolution theorem, the specimen profile, $f(s-s_0,\,d^*)$, can be expressed in terms of a "size/domain" profile, $f^s(s-s_0)$ and a "strain" profile $f^d(s-s_0,\,d^*)$ as

$$f(s-s_0,\,d^*) = f^s(s-s_0) * f^d(s-s_0,\,d^*) \tag{1.16}$$

where $*$ is the convolution operator. That is, the strain/size-broadenedspecimen profile can be interpreted as the convolution of a "size profile" with a "strain profile". Furthermore, if $f^s(s-s_0) = \delta(s-s_0)$, the strain broadening will be dominant and the crystallite size will be infinitely large. On the other hand, $f^d(s-s_0,\,d^*) = \delta(s-s_0,\,d^*)$, corresponds to a strain-free specimen where size broadening is dominant. Another observation that can be made of (1.14 & 1.15) is that the size coefficients are independent of the order of diffraction, but dependent on the distance, L, parallel to the diffraction vector, while the strain coefficients are dependent on L and d^* (Warren & Averbach 1952, Eastabrook & Wilson 1952).

If $f(s-s_0,\,d^*)$ is symmetrical, the imaginary coefficients $B^d(L,\,d^*)$ will be zero for all L and d^* and the real coefficients will be non-zero for all L and d^*. In the case of an asymmetrical specimen profile, the imaginary coefficients will be non-zero and as the asymmetry increases the magnitude of these coefficients will increase relative to the real coefficients. The imaginary Fourier coefficients are not usually interpreted physically (van Berkum et al. 1994), however there are some exceptions which Ungár et al. (1989) have discussed.

Size-Fourier coefficients

Warren & Averbach (1950, 1952) and Bertaut (1950, 1952) demonstrated that the size coefficients can be expressed in terms of a column-length distribution, $p_a(L)$ as

$$A^s(L) = \frac{1}{\langle L\rangle_a}\int_{|L|}^{\infty} (L'-|L|)\,p_a(L')\,\mathrm{d}L' \tag{1.17}$$

where $p_a(L)$ has unit area and the apparent crystallite-size, $\langle L\rangle_a$ is given by

$$\langle L\rangle_a = \int_0^{\infty} L\,p_a(L)\,\mathrm{d}L. \tag{1.18}$$

Following on from (1.17), it was also shown by Warren & Averbach (1950, 1952) and

Bertaut (1950, 1952) that the $\langle L\rangle_a$ can be determined from the initial slope of $A^s(L)$ as

$$\frac{1}{\langle L\rangle_a} = -\left(\frac{\mathrm{d}A^s(L)}{\mathrm{d}L}\right)_{L\to 0} \tag{1.19}$$

and taking the second derivative of $A^s(L)$ in (1.17) produces the column-length distribution, $p_a(L)$, as

$$p_a(L) = \frac{1}{\langle L\rangle_a}\frac{\mathrm{d}^2A^s(L)}{\mathrm{d}L^2}. \tag{1.20}$$

We can interpret $A^s(L)$ in terms of the common-volume function, $V(L)/V(0)$ and the integral breadth, (1.4), can be alternatively expressed in terms of the size-Fourier coefficients (Averbach & Warren 1949, Warren & Averbach 1950, Eastabrook & Wilson 1952) as

$$\beta = \left[\int_0^\infty A^s(L)\,\mathrm{d}L\right]^{-1}. \tag{1.21}$$

This also implies that the first and second derivative of the common-volume are directly equivalent to (1.19) and (1.20), respectively.

Some observations about the size distribution

Guinier (1963)[3] related the area-averaged column-length, $p_a(L)$ to the volume-averaged distribution, $p_v(L)$, and established a relation between $\langle L\rangle_a$ and its volume-weighted equivalent, $\langle L\rangle_v$. The relation between the two quantities was established *via* the common-volume function. Guinier (1963) showed that

$$\frac{p_a(L)}{\langle L\rangle_a} = \frac{p_v(L)}{L}. \tag{1.22}$$

In (1.22), either the volume- or area-weighted distribution can be determined from knowing the other.

Wilson (1968, 1971) showed that the area-averaged size, $\langle L\rangle_a$, is the average thickness normal to the reflecting planes. This was established in terms of the common-volume function, in that $V(0)/V'(0)$ was shown to be the area of the crystallite projected onto the reflecting planes. Moreover, Wilson (1968, 1971) showed that $V''(0)/V(0)$ had two interpretations. The first considered it as the rate at which the crystal tapers in the direction perpendicular to the reflecting planes; secondly, it can be considered as the distribution function giving the fraction of the volume of the crystallite with a thickness L and $L + \mathrm{d}L$ given in (1.20). The

[3] Dover (New York) edition, first published in 1994.

resulting crystallite/domain size was interpreted as a crystal made up from all the crystals in the sample and the projected area equal to the sum of projected areas of each crystal. It was also pointed that $\langle L \rangle_a$ will in general not be equal to the average crystallite dimension, $\langle D \rangle$ of the ensemble of crystallites.

Smith (1972) interpreted the area-averaged size in terms of the average-surface area of a catalyst. It was proposed that x-ray profile analysis could be used in determining the total surface-area of the catalyst. Smith (1972) pointed out that $\langle L \rangle_a$ was more appropriate since it defined an area-averaged size, while $\langle L \rangle_v$ represented a volume-weighted size. The area-averaged size was expressed as

$$\begin{aligned} \langle L \rangle_a &= \int L \, \mathrm{d}S \Big/ \int \mathrm{d}S \\ &= V(0)/A_{hkl} \end{aligned} \tag{1.23}$$

where L is the dimension normal to the reflecting plane; dS is the projected area of the crystallite onto the (hkl)-plane; $V(0)$ is the volume of the crystallite and A_{hkl} is the total area of the crystallite projected onto the (hkl)-plane. This can be related to the surface-area of the of the crystallite, $\mathcal{S}$ in terms of the mass, m, and density, ϱ of the catalyst, as

$$\mathcal{S} = 4m/(\varrho \, \langle L \rangle_a) \tag{1.24}$$

where $\mathcal{S}$ is the average surface area of the crystallite. This model assumes that the catalyst has a uniform shape (i.e. spherical). Hence using x-ray diffraction profile analysis the average crystallite size can be determined and used to determine the average surface area.

Smith (1976) also demonstrated the relationship between $p_a(L)$, which behaves as an "apparent-size" distribution, and the true size distribution, $P(D)$, for a given dimension of the crystallite. In the particular case that the fraction of spherical crystallites, $P(D)$ with a diameter between D and $D + \mathrm{d}D$ is mapped into the distribution of columns, $p_a(L)$ is given as

$$p_a(L) = \int_{D_0(L)}^{\infty} G(L,\, D) \, P(D) \, \mathrm{d}D \tag{1.25}$$

where $G(L,\, D)$ is the *shape kernel* defining how the distribution of crystallite diameters (or more general dimension) is mapped into the column-length distribution, $p_a(L)$; $D_0(L)$ is the least value of D and is dependent on L. For a sphere, $D_0 = L$ for all (hkl)-planes.

Smith (1976), showed that in general $\langle L\rangle_a \neq \langle D\rangle$ because $G(L, D)$ smears the particle diameter, $P(D)$. This is the same result Wilson (1968, 1971) showed. The shape kernel can be determined from the second derivative of the volume of the crystallite. For a spherical crystallite, $G(L, D) = \frac{\pi}{2}L$, constant for all D and $L \in [0, L_{max}]$. Smith (1976) demonstrated that a similar relationship can be determined for other shapes such as cubes and tetrahedra. Moreover, Smith (1976) pointed out that the column-length distribution is representative of the crystallite because of the mapping relationship given in (1.25) and that $p_a(L)$ does not correspond to the distribution of crystallites determined from direct methods, such as electron microscopy. LeBail & Louër (1978) showed that a size-broadened profile can be expressed as

$$f(s - s_0) = \sum_{n=0}^{\infty} \frac{\sin^2 \pi(s - s_0)an}{\sin^2 \pi(s - s_0)a}\, p_a(n) \tag{1.26}$$

where $(s - s_0) = \frac{2\sin\theta}{\lambda} - d^*$; a is the step-size in the Fourier space and is given by $a = (\frac{2\sin\theta_2}{\lambda} - \frac{2\sin\theta_1}{\lambda})^{-1}$ for the profile over the region of $[2\theta_1, 2\theta_2]$; n is the harmonic-order of the Fourier order. The distance perpendicular to the diffraction planes is given by $L = a\,n$; that is, the size-broadened profile can be considered as a set of columns of length L diffracting independently to each other. The size-profile corresponds to the weighted sum of intensities from each column. The $\frac{\sin^2 \pi(s-s_0)an}{\sin^2 \pi(s-s_0)a}$ in (1.26) is the scattering kernel from a column-length $a\,n$; as $n \to \infty$, the crystallites are infinitely large and the specimen profile becomes a delta function as described above.

Strain-Fourier coefficients

Warren & Averbach (1950, 1952) and Eastabrook & Wilson (1952) expressed the strain coefficients in (1.14) in terms of the strain distribution, $p(\epsilon_L)$, parallel to the diffraction vector as

$$A^d(L, d^*) = \langle \cos 2\pi L\epsilon_L d^*\rangle \tag{1.27}$$

$$= \int_{-\infty}^{+\infty} p(\epsilon_L)\, \cos 2\pi L\epsilon_L d^*\, \mathrm{d}\epsilon_L \tag{1.28}$$

where $\langle \cos 2\pi L\epsilon_L d^*\rangle$ represents the average quantity. The strain coefficients, (1.27 & 1.28) can be interpreted in terms of the average value of the complex structure factors, FF^*, where F^* is the complex (imaginary) conjugate of F (Stokes & Wilson 1944a, Eastabrook &

Wilson 1952). The distribution, $p(\epsilon_L)$, is dependent on a physical model of strain fields. It attributes a probability distribution to the strain averaged over the distance L perpendicular to the diffraction planes. If there were dislocations in the specimen, ϵ_L and $p(\epsilon_L)$ would be dependent on the ensemble of dislocations and the strain field due to their presence at distance L. For the case of inhomogeneous strain, $p(\epsilon_L)$ will have a zero mean and the rms-strain will be non-zero. (The case where $p(\epsilon_L)$ has a non-zero mean corresponds to homogeneous strain.) The shape of $p(\epsilon_L)$ will also influence the shape of the diffraction profile.

Some observation about $A(L,\, d^*)$, ϵ_L and $p(\epsilon_L)$

It is interesting to notice that (1.27) is a cosine-Fourier transform of the strain distribution, $p(\epsilon_L)$, and in principle $p(\epsilon_L)$ can be determined from the data. However, there are several difficulties with taking the inverse cosine-Fourier transform of (1.27) which will be discussed in the next section (§1.4) and in greater detail in Chapter 5. McKeehan & Warren (1953) and Warren (1959) made this observation and discuss the validity of applying the inverse cosine-Fourier transform to (1.27). Eastabrook & Wilson (1952) looked at a number of cases in which the maximum information concerning the size and strain distribution can be extracted from the specimen profile. They pointed out that the strain Fourier coefficients could be expressed in a general form as

$$A(L,\, d^*) = \exp\left[-|L|\, \mathfrak{f}(d^*)\right] \tag{1.29}$$

where $L \in (-\infty,\, \infty)$ and $\mathfrak{f}(d^*)$ is a function describing the displacements. Eastabrook & Wilson (1952) showed that depending on the form of $\mathfrak{f}(d^*)$, $A(L,\, d^*)$ could represent the Fourier coefficients of many possible functions, including Gaussian and Lorentzian functions. Williamson & Hall (1953), Williamson & Smallman (1954) and Williamson & Smallman (1956) investigated the influence of the strain distribution on the shape of the specimen profile for cold-worked metals. In these contributions it was proposed that dislocations were the source of both size and strain broadening. That is, if the dislocations were arranged in a particular structure which consisted of regions of low-dislocation densities and regions

of high-dislocations densities this would produce incoherent scattering domains[4] and $\langle L \rangle_a$ corresponds to the mean dimension of the domain parallel to the diffraction vector. It was also suggested by these workers that the strain distribution could be characterized by a Lorentzian distribution. The difficulty with this is that the rms-strain, $\langle \epsilon_L^2 \rangle^{\frac{1}{2}}$, is not finite, so a "tunable parameter" was often introduced to produce a finite rms-strain (Williamson & Smallman 1954, Warren 1959, Takahashi 1969). van Berkum (1994) pointed out that if the strain-distribution is describe by a Pearson-VII function (see Appendix A), then $\langle \epsilon_L^2 \rangle^{\frac{1}{2}}$ is finite for the exponent, $m > \frac{3}{2}$. In this region the Pearson-VII function describes an intermediate-Lorentzian. Williamson & Smallman (1956) discussed the influence the shape of the specimen profile has in determining the dislocation density. A range of dislocation densities were determined with the Gaussian and Lorentzian functions defining the two extremes.

Using (1.27) the statistical nature of the strain broadening can be taken into consideration. That is, quantities such as $A(L,\, d^*)$, ϵ_L and $p(\epsilon_L)$ describe the "macroscopic state" of the specimen and each term is dependent on the ensemble dislocations. On the other hand, describing the Fourier coefficients and strain broadening by (1.27), ϵ_L and $p(\epsilon_L)$ provide a phenomenological description of strain broadening (Krivoglaz 1995). That is, it does not account for the "microscopic state" of the specimen or for the underlying mechanisms which result in profile broadening or quantify the type of dislocations, their densities and mean spacing. This was realised by a number of contributors (e.g. Ryaboshapka & Tikhonov 1961*a*, Ryaboshapka & Tikhonov 1961*b*, Krivoglaz & Ryaboshapka 1963, Krivoglaz 1995, Wilkens 1970*c*, Wilkens 1970*b*, Wilkens 1970*a*, Wilkens 1987, Groma et al. 1988, Ungár et al. 1989, Klimanek & Kužel 1988, Kužel & Klimanek 1988, Vermeulen et al. 1995, Wu, Gray & Kisi 1998, Wu, Kisi & Gray 1998, Ungár & Borbély 1996, Ungár, Ott, Sanders, Borbély & Weertman 1998, Ungár, Révész & Borbély 1998). These various contributions have common features: for example, the dislocations are assumed to be randomly distributed throughout the specimen, and the contrast (or orientation) factors for various isotropic and anisotropic materials need to be included into the models as well as the Burgers vectors

[4] Warren (1959) and van Berkum et al. (1996) discuss incoherent scattering domains; essentially the phase of the scattered wave changes randomly between 0 and 2π from one domain to the next. Another term is *uncorrelated scattering* (Cohen 1966).

and slip system of the dislocations. However, the statistical nature of describing diffraction from a polycrystalline material, together with the mathematical difficulties of introducing dislocation theory to describe the profile broadening often results in limiting cases, such as randomly distributed dislocations, being considered (for example see Krivoglaz 1995).

van Berkum et al. (1996) has developed a model in which the Fourier coefficients of the specimen profile are defined in terms of periodically spaced low- & high- regions of dislocations, the mean dimension between the dislocations, width of high-dislocation regions and rms-strain. van Berkum et al. (1996) showed that if the relative width of the strain field (or dislocations), w_r, is infinitely broad, the Fourier coefficients and the resulting strain coefficients in (1.27) are for a Gaussian function. On the other hand, if the $w_r \to 0$ then (1.27) is independent of the order of diffraction, the resulting broadening being interpreted as size broadening. Hence, for intermediate values of w_r both size and strain broadening would occur irrespective of the functional form of $p(\epsilon_L)$.

The Fourier coefficients and rms-strain can describe strain broadening in terms of dislocations and their properties (Wilkens 1970*c*). The strain Fourier coefficients are given as

$$A(L,\, d^*) = \langle \exp 2\pi \imath L d^* \epsilon_L \rangle \tag{1.30}$$

where ϵ_L denotes the strain averaged over L parallel to the diffraction vector and can be expressed as

$$\epsilon_L = \frac{1}{L} \int_{-L/2}^{+L/2} \epsilon_0(\mathbf{r} + s\mathbf{e}_0)\, \mathrm{d}s \tag{1.31}$$

where $\mathbf{r}$ is the position vector and $\mathbf{e}_0$ a unit vector parallel to $\mathbf{d}^*$. In (1.31), ϵ_0 about the core of the dislocation is being smeared as it is averaged over different lengths L. In the limit $L \to \infty$, then $\epsilon_L \to 0$ decreasing monotonically; for $L \to 0$ the $\epsilon_L \to \epsilon_0$. If (1.27) or (1.30) are expanded as a power series we have (Wilkens 1970*c*)

$$A(L,\, d^*) = 1 - \frac{1}{2!}(2\pi d^* L)^2 \, \langle \epsilon_L^2 \rangle + \frac{1}{4!}(2\pi d^* L)^4 \, \langle \epsilon_L^4 \rangle + \cdots + \tag{1.32}$$

where $\langle \epsilon_L^2 \rangle$, $\langle \epsilon_L^4 \rangle$ etc. are the even strain moments. These terms can be related to the dislocation density, ρ, the orientation factor, C and Burgers vector $\mathbf{b}$; in the case of $\langle \epsilon_L^2 \rangle$ this would be given by (Wilkens 1970*c*)

$$\langle \epsilon_L^2 \rangle = (b/2\pi)^2 \, \pi \, \rho \, C \, \mathfrak{f}(L \sin \psi / R_e) \tag{1.33}$$

where $\mathfrak{f}(L \sin\psi/R_e)$ is a function of the effective outer cut-off radius, over the distance L in the xy plane. The orientation factor will depend on the elastic nature of the material and lattice system. The $\langle\epsilon_0^2\rangle$ in (1.31) is the same quantity determined by Ryaboshapka & Tikhonov (1961*a*, 1961*b*), Krivoglaz & Ryaboshapka (1963) and by van Berkum et al. (1996). This approach also enables the above quantities to be compared with transmission electron microscopy (TEM) results. For example Ungár & Borbély (1996), Ungár, Ott, Sanders, Borbély & Weertman (1998), and Ungár, Révész & Borbély (1998) have obtained very good agreements for the dislocation densities between TEM and x-ray analysis (also see Mughrabi 1983, Mughrabi et al. 1986, Ungár et al. 1982, Ungár et al. 1986, Ungár 1994).

In this section the integral and Fourier representations of the line-profile have been reviewed. These representations have been defined in terms of the "size" and "strain" contribution and the underlying physical quantities have been discussed. The central difficulty which is encountered in analysing broadened diffraction profiles is determining the unknown quantities in experimental conditions. This is especially challenging if both size and strain broadening are present. In the next section dominant and most recent methods for analysing broadened profiles are discussed and reviewed.

1.4 Analysing broadened diffraction profiles

The integral breadth and Fourier descriptions of crystallite size and strain broadening gave rise to two dominant methods for analysing the broadened specimen profile, the Williamson & Hall (1953) and Warren & Averbach (1950, 1952) methods, respectively. The application, procedure and interpretation of these methods will now be discussed, together with recent methods that define profile broadening in terms of the microscopic properties of the specimen. However, before any size/strain analysis can be undertaken the instrumental broadening must be removed; various methods for doing this will be outlined.

1.4.1 Determining the specimen profile

The observed profile

The observed profile, g, recorded at the output of a diffractometer can be expressed as the convolution of the specimen profile, f and instrument profile, k, superimposed onto a (slowly varying) background level, b, and statistical noise, n as

$$g(2\theta) = \int k(2\theta - 2\theta')\, f(2\theta')\, \mathrm{d}(2\theta') + b(2\theta) + n(2\theta). \tag{1.34}$$

In (1.34) the specimen profile is being mapped from the $2\theta'$-space into 2θ-space *via* the instrument profile which characterises the finite response of the diffractometer to the radiation source and is a function of the non-ideal optics, and equatorial, axial aberrations (see Wilson 1963, Klug & Alexander 1974). The background level, b, can be made up of various components such as air and thermal diffuse scattering, fluorescenes of the specimen, amorphous content, the substrate and even the specimen holder. The noise is essentially errors in the counting and follows Poisson statistics; for large counts ($\gtrsim 10$) the noise can be approximated by Gaussian statistics (Sivia 1996).

The first problem faced in profile analysis is determining a suitable instrument profile. There are two dominant approaches for determining the instrument profile: the fundamental approach which defines k as the convolution of aberration functions (for example see Klug & Alexander 1974, Cheary & Coelho 1992, Kogan & Kupriyanov 1992, Cheary & Coelho 1998*a*, Cheary & Coelho 1998*b*); and the experimental approach which uses a reference material that has negligible imperfections (for example see Fawcett et al. 1988, Louër & Langford 1988, Rasberry 1989, Scardi et al. 1994). For example, in their fundamental parameters approach, Cheary & Coelho (1992) defined the instrument profile in terms of the Cu $K\alpha$ emission profile and instrumental aberration functions as

$$k = L * k_1 * k_2 * k_3 * \cdots * k_n \tag{1.35}$$

where $L \equiv L(2\theta)$ is the emission profile from Cu $K\alpha$ radiation and $k_i \equiv k_i(2\theta)$ where $i = 1, 2, 3, \ldots, n$. The Cu $K\alpha$ radiation consists of a number of $K\alpha$ components which include the $K\alpha_1$ and $K\alpha_2$ components and a number of satellite lines. Seven instrumental aberration functions were considered by Cheary & Coelho (1992), which included the receiving-slit

width, slit-length, flat-specimen and absorption functions. The instrument profile in (1.35) was fitted with the observed profile and the parameters defining k_i were adjusted until the difference between the calculations and observations were minimised. This approach can also include a specimen profile refined on a particular size and strain model.

In the case of an experimental approach, a reference material is used to characterise the instrumental broadening. The reference material is chosen to have large crystallites and negligible lattice distortions so that broadening in the line-profile consists of the non-ideal optics of the diffractometer. As an example, reference materials used in determining the instrument profile consist of LaB_6 (NIST SRM 676) (Fawcett et al. 1988), annealed-BaF_2 (Louër & Langford 1988) and KCl (Scardi et al. 1994). Systematic errors can be introduced into the specimen profile if the structural imperfections in the reference material are not minimized. For example NIST SRM 676 has recently been shown to contain lattice distortions (see Cheary & Coelho 1998*b*); hence, the instrument profile consists not only of instrumental broadening from the diffractometer but also specimen broadening of the reference material. In deconvolving out the instrumental broadening, it may be over-compensated by the deconvolution method. (This problem is discussed in Chapters 4 & 6.) Furthermore, if there is differential absorption between the reference material and the specimen, the absorption must be corrected in order to eliminate systematic errors in the specimen profile.

Once the diffraction pattern of the reference material is recorded it can be fitted with a suitable analytical profile and the parameters that defined it as a function of angular position, 2θ. For example, Scardi et al. (1994) defined the instrument profile as a pseudo-Voigt function (see Appendix A) and the after fitting it to the KCl line-profiles a set of calibration graphs of the full-width at half maximum (FWHM) and mixing parameters, η, were determined as a function of 2θ. This enabled the instrument profile to be defined at the angular position of the observed profile for the specimen. Usually, the FWHM is fitted using (see Cheary & Cline 1995)

$$FWHM^2 = A\tan^2\theta + B\cot^2\theta + C\tan\theta + D. \quad (1.36)$$

Cheary & Cline (1995) have investigated the dependence of (1.36) (see equations (7) & (8), p81 Cheary & Cline 1995), and the asymmetry on the divergence slit, receiving slit and the number of Soller slits. This analysis can be used as a guide for suitable slits in optimizing

the instrument profile. In addition to determining the specimen profile, it may be necessary to correct for the effect of truncation of the specimen profile or to resolve overlapped profiles. Various studies and methods exist (e.g. Mortier & Costenoble 1973, Rondot & Mignot 1977, de Keijser & Mittemeijer 1980, Zocchi 1980, Delhez et al. 1982, Delhez et al. 1986, Sonneveld et al. 1991, Vermeulen et al. 1991, Vermeulen et al. 1992). Delhez et al. (1982) provide a detailed procedure for the various corrections and suggest that in the case of background estimation that the observed profile be plotted on a logarithm scale. This has the advantage that any distortions in the profile can be highlighted.

Deconvolution methods

The problem of determining the specimen profile from (1.34) involves a deconvolution. A detailed mathematical analysis and simulations of the various deconvolution methods used in profile analysis are discussed in Chapter 2. However, it is worth making some observations about the methods. The problem of removing the instrumental broadening was first realised by Stokes (1948) who applied the Fourier transform to the convolution product in (1.34) (the background level and noise were not considered) and expressed the Fourier transform of the observed profiles as a product of the transforms of the specimen and instrument profiles. Today this method can be applied rapidly by using digital computers and fast Fourier transform algorithms, reducing the number of operation to $\sim N \log_2 N$, as opposed to $\sim N^2$ operations for a discrete Fourier transform (see Arfken & Weber 1996, Vesely 1994). However, this method is very sensitive to noise and usually results in ill-conditioning of the solution (Kalceff et al. 1995). Other common deconvolution methods, such as the iterative deconvolution (Ergun 1968) and constrained deconvolution methods (Louër et al. 1969, Louër & Weigel 1969, Louboutin & Louër 1972) used in profile analysis also suffer similar problems (see Chapter 2). Moreover, these methods do not preserve the positivity of the specimen profile. This aspect is important since the specimen profile represents an intensity distribution.

A direct convolution approach has been developed to overcome these difficulties (e.g. Howard & Snyder 1989, Cheary & Coelho 1992). The direct convolution approach assumes an analytical profile function for the specimen profile; the convolution product between

the instrument and specimen profile is refined, by updating the parameters that define the specimen profile, until the error between the calculate and observed data is minimised. This method is basically a means to an end; that is, it does not ensure that the specimen profile is unique since many different functions can be chosen to fit the data. Moreover, there is often no physical basis for choosing a particular profile function, except that it results in a minimised error. However, since the profile functions are positive definite it ensures that this aspect is preserved.

1.4.2 Williamson-Hall method

Separating size & strain – Integral breadth approach

Williamson & Hall (1953) proposed that the size and strain contribution from a broadened line profile can be resolved if the the total integral breadth for a Lorentzian (or Cauchy) shaped profile is the sum of the size integral breadth, β^s, and strain (or distortion) integral breadth, β^d. Using (1.4) and (1.12) for a Lorentzian profile, we have

$$\beta(d^*) = \beta^s + \beta^d(d^*). \tag{1.37}$$

If the line profile is Gaussian, it can be shown that the square of the total integral breadth is equal to the sum of squares for the size and strain integral breadths

$$\beta^2(d^*) = \beta^{s\,2} + \beta^{d\,2}(d^*) \tag{1.38}$$

where β, β^s and β^d are in reciprocal-space units.

Equations (1.37) and (1.38) simply define the total breadth in terms of the size and strain integral breadths, (1.4) and (1.12), respectively. Both (1.37) and (1.38) have order-independent and order-dependent terms that are needed to describe the size and strain broadening. However, the Williamson & Hall (1953) method assumes a functional form of the profile, either Lorentzian or Gaussian. These cases are seen as defining two extremes which can be quantified by the shape parameter $\phi = FWHM/\beta$ (Langford 1992). For the Lorentzian and Gaussian limits, ϕ, is defined by $0.6366 \leq \phi \leq 0.9394$ (Langford 1992).

Considering the Lorentzian case, (1.37), and using the results (1.4) and (1.12), the intercept can be related to the volume-weighted size, $\langle L \rangle_v$ and the slope is proportional to the

apparent strain, η,

$$\beta(d^*) = \frac{1}{\langle L \rangle_v} + \frac{\eta}{2} d^* \tag{1.39}$$

The Williamson & Hall (1953) method provides a qualitative understanding of the specimen broadening, which includes determining the dominant source of specimen broadening, and whether the specimen broadening is isotropic or anisotropic in hkl (e.g. Langford 1980, Delhez et al. 1982, de Keijser et al. 1982, Louër et al. 1983, Langford et al. 1988, Howard & Snyder 1989, de Keijser et al. 1983, Langford 1992, Langford et al. 1993, Louër 1994). The most notable contributions include Langford (1992) and Langford et al. (1993), where the Williamson & Hall (1953) method was applied qualitatively and quantitatively. Langford (1992) reviewed the application of the Williamson & Hall (1953) method to various samples including tungsten filing and ZnO powder. A positive slope and zero intercept in (1.39) implies that strain broadening is dominant and crystallite size is large. On the other hand, if the slope is zero and has a positive intercept, the specimen broadening is dominated by size broadening. In the case of anisotropic size or strain anisotropy, multiple orders of hkl indices are used to determine the relevant quantities. For example Langford (1992) demonstrated that for ex-hydroxide nitrate ZnO (see "sample B", Figure 3(d)), the β for the $h00$ and $00l$ had a non-zero slope and different values, indicating the presence of anisotropic strain broadening for those indices. Moreover, it was shown that the crystallites had a prismatic shape since the intercept of the $00l$ was less than the intercept of the $h00$ lines.

The Williamson & Hall (1953) method can only be applied quantitatively if a suitable physical model for the strain and size broadening is adopted. In the case of size broadening it generally requires the common-volume function, $V(L)$, to be related to the integral breadth using (1.6) and its overall dimensions can be determined. However, this assumes that the powder consists dominantly of particular shaped crystallites and the crystallite distribution is narrow (see Langford 1992, Langford et al. 1993). Langford et al. (1993) showed that for ex-oxalate ZnO powders a spherical model with average diameter, $\langle D \rangle = 453\,\text{Å}$ was applicable; for another ZnO powder specimen a cylindrical model, with a diameter, $\langle D \rangle = 404\,\text{Å}$ and height $\langle H \rangle = 404\,\text{Å}$ was also applicable. In the case of strain broadening, the apparent strain η is related to the rms-strain by using (1.13), which defines an order of magnitude for the rms-strain. Langford (1992) used (1.13) and (1.38) on tungsten filings

to determine $\langle L \rangle_v = 670\,\text{Å}$ and $\langle \epsilon^2 \rangle^{\frac{1}{2}} = 2.98 \times 10^{-3}$ and a spherical crystallite model was determined resulting in $\langle D \rangle = 860\,\text{Å}$. This was established by noticing that β was isotropic in hkl. However, on applying (1.13), it is assumed the distortions are randomly distributed throughout the crystallite.

Intermediate cases

The Williamson & Hall (1953) method given by (1.37) and (1.38) defines two extremes for Lorentzian and Gaussian profiles, respectively. However, due to a mixture of historical and empirical reasons, size-broadened profiles have been attributed to Lorentzian profiles, while strain-broadened profiles were considered to be Gaussian (see Stokes & Wilson 1944*a*, James 1948). Generally, this is not the case. In the case of size-broadening, (1.25) clearly indicated that the combination of shape and crystallite size distribution will determine the shape of the size profile. In the case of strain, it is the arrangement of dislocations that influences the strain Fourier coefficients (see (1.32)) and in turn determines the strain profile. In (1.16), the size/strain-broadened profile was represented as the convolution of "size" and "strain" profiles. Defining the size profile as a Lorentzian and the strain profile by a Gaussian the convolution product results in a Voigt function (e.g. Ruland 1965, Schoening 1965, Handler & Wanger 1966, Ruland 1968, Langford 1978). Its integral breadth is given by (Langford & Louër 1996)

$$\beta = \frac{\beta_G}{\exp(k^2)\,\left[1 - \text{erf}(k)\right]} \tag{1.40}$$

where $k = \frac{\beta_L}{\sqrt{\pi}\,\beta_G}$ is the Voigt parameter which relates the Gaussian, β_G and the Lorentzian, β_L integral breadth and $\text{erf}(k)$ is the error-integral (see Arfken & Weber 1996).

The Voigt function has been used successfully in integral breadth analysis since it so readily quantifies the specimen broadening in terms of Lorentzian and Gaussian components (e.g. Langford 1980, Langford et al. 1988, de Keijser et al. 1982, de Keijser et al. 1983, Langford 1992, Langford et al. 1993, Louër 1994). It is also suited for rapid analysis (Langford 1978), making it ideal for a commerical environment. Langford (1992) gives a detailed review of the Voigt function and its application in the Williamson & Hall (1953) method. Wu, Gray & Kisi (1998) and Wu, Kisi & Gray (1998) have adapted the Voigt function and Rietveld method to determining the dislocation densities and for studying plastically

deformed isotropic materials.

The almost standard application of digital computers in size-strain and profile analysis has seen the development of analytical functions which account for specimen and instrumental broadening. It has meant that rapid analysis can be applied to a diffraction profile[5]. The application of profile functions ranges from single profile analysis to Rietveld and whole pattern analysis (e.g. Rietveld 1969, Huang & Parrish 1977, Delhez et al. 1982, Benedetti et al. 1988, Enzo et al. 1988, Fawcett et al. 1988, Langford et al. 1986, Howard & Snyder 1989, Vogel 1990, Cheary & Coelho 1992, Scardi et al. 1994, Louër 1994, Cheary & Coelho 1996). Langford & Louër (1996) have reviewed in detail the role of methods such as Rietveld analysis and line profile analysis in characterising the microscopic properties of polycrystalline materials. Balzar & Popović (1996) have shown that systematic errors can arise in determining the crystallite/domain size when applied in Rietveld analysis. Recently Leoni & Langford (1998) have posed the question of whether the Rietveld method and line profile analysis can be combined. They suggests the physical information must be checked with researchers' intuition and the results from other methods.

Theoretical limitations

The Williamson & Hall (1953) method is essentially a semi-quantitative method. It is dependent on assuming a functional form of the specimen profile and a model for the size and strain broadening in order to ascertain the microscopic nature of line-broadening.

For example, its description of the strain is dependent on the model proposed by Stokes & Wilson (1944*a*), which assumes the strain is randomly distributed in a crystallite. At best this determines an order of magnitude of the strain. It does not determine the characteristics or density for the dislocations which give rise to the strain broadening.

In terms of size analysis, no information concerning the size distribution can be determined from the Williamson & Hall (1953) method. The volume-averaged size, $\langle L \rangle_v$ is reliant on the assumptions by Stokes & Wilson (1942), which are that the shape and size of the crystallites are the same in the sample. Generally, it is likely that $\langle L \rangle_v$ will correspond to a distribution.

[5]The common profile functions are described in Appendix A

As a qualitative method for assessing the nature and type of specimen broadening, the Williamson & Hall (1953) method is successful. That is, this qualitative method can be used to build a model for the crystallite size broadening. In the case of strain, it can distinguish between isotropic and anisotropic broadening.

1.4.3 Warren & Averbach method

Separating size & strain – Fourier approach

Building on the results (1.17) and (1.27 & 1.28), Warren & Averbach (1950, 1952) presented a method for separating the size and strain contributions in the specimen profile. That is, using the Stokes (1948) method for determining the instrumental profile, the real and imaginary coefficients of the specimen profile can be determined. From (1.14 & 1.15), it was observed that the real coefficients are the product of the strain and size coefficients given by

$$A(L,\, d^*) = A^s(L)\, A^d(L, d^*). \tag{1.41}$$

Taking the natural logarithm of (1.41), we have

$$\ln A(L,\, d^*) = \ln A^s(L) + \ln A^d(L,\, d^*) \tag{1.42}$$

and substituting (1.27) into the above result, we have

$$\ln A(L,\, d^*) = \ln A^s(L) + \ln\langle \cos 2\pi L \epsilon_L d^* \rangle. \tag{1.43}$$

On expanding the cosine term $\langle \cos 2\pi L\epsilon_L d^*\rangle \approx (1 - 2\pi^2 \langle \epsilon_L^2 \rangle L^2 d^{*\,2})$ and taking the logarithm of the expansion, $\ln(1 - 2\pi^2\langle\epsilon_L^2\rangle L^2 d^{*\,2}) \approx -2\pi^2\langle\epsilon_L^2\rangle L^2 d^{*\,2}$ where the result $\ln(1-x) \approx -x$ for $|x| < 1$ has been applied. Substituting this result into (1.43), we have the Warren & Averbach (1950, 1952) equation for separating size and strain broadening contributions as

$$\ln A(L,\, d^*) = \ln A^s(L) - 2\pi^2 \langle \epsilon_L^2 \rangle L^2 d^{*\,2} \tag{1.44}$$

where $\langle \epsilon_L^2 \rangle$ is the mean-square strain for a given L, perpendicular to the (hkl) plane. The graph of $\ln A(L,\, d^*)$ versus $d^{*\,2}$ for a given L will have an intercept of $\ln A^s(L)$ and a slope equal to $-2\pi^2\langle\epsilon_L^2\rangle L^2$, where $\langle \epsilon_L^2 \rangle$ can be determined from the slope. From $\ln A^s(L)$, the size Fourier coefficients can be determined and the area-weighted size, $\langle L \rangle_a$ determined using

(1.19). In principle $p_a(L)$ can also be determined. Originally the Warren & Averbach (1950, 1952) method was applied to cold-worked metals (e.g. McKeehan & Warren 1953, Warren 1959, Warren 1963), but it can also be applied to a wide range of polycrystalline materials, including ceramic high-T_c superconductors (Balzar et al. 1993).

Fourier analysis

In the same manner that the Williamson & Hall (1953) method can be used in qualitative analysis, the Warren & Averbach (1950, 1952) method can be used to make a qualitative analysis of the broadened profile. In the case of size broadening, this becomes important in determining a suitable crystallite model (see below). For strain broadening, examination of the Fourier coefficients will indicate either isotropic or anisotropic broadening. This information combined with the Williamson & Hall (1953) method analysis can be used to develop the best approach for a quantitative analysis. In the case of isotropic strain broadening, the Fourier coefficients become narrower with increasing d^*. For anisotropic strain broadening no such monotonic relationship exists. In this case the analysis is limited to multiple orders. It also illustrates the limitations of the Warren & Averbach (1950, 1952) method in accounting for anisotropic strain broadening by not describing it in terms of a microscopic picture.

Equation (1.44) will be linear for all L and d^{*2}, only if the strain distribution, $p(\epsilon_L)$, is a Gaussian. Otherwise, the accuracy of (1.44) will be limited for small L. (1.44) can be compared to (1.32), where (1.44) consists only of the second term in the power series. This limitation was noticed by a number of researchers who presented variations or modifications to the original Warren & Averbach (1950, 1952) method (e.g. Pines & Surovtsev 1963, Vogel et al. 1974, Delhez & Mittemeijer 1976, Mignot & Rondot 1977, Delhez et al. 1980, Turunen et al. 1983, Schlosberg & Cohen 1983, Nandi et al. 1984, Hosemann et al. 1985, van Berkum et al. 1993, van Berkum et al. 1994). Delhez & Mittemeijer (1976) proposed a modification for separating the size and strain Fourier coefficients and claimed a 10% improvement in determining the size coefficients and 30% in the $\langle\epsilon_L^2\rangle$-values. They also observed that neglecting higher-order terms in the logarithm expansion introduced large errors for $\langle\epsilon_L^2\rangle$-values, while for large values the error was less pronounced. This modification

differs in not applying the logarithm,

$$A(L,\, d^*) = A^s(L) - 2\pi^2 \langle \epsilon_L^2 \rangle L^2 d^{*\,2}\, A^s(L) \tag{1.45}$$

where by plotting $A(L,\, d^*)$ versus $d^{*\,2}$, the intercept is given by $A^s(L)$ and $\langle \epsilon_L^2 \rangle$ can be found from the slope for a given L. van Berkum et al. (1993, 1994) investigated the region of validity for the Warren & Averbach (1950, 1952) method and a modified method that assumed $p(\epsilon_L)$ was independent of L (see equation(10), p347 van Berkum et al. 1994). The difference between the Warren & Averbach (1950, 1952) and alternative method is that the former method uses the Fourier coefficients of different profiles for a given L, while the latter uses the coefficients for different L-values. It was shown that the Warren & Averbach (1950, 1952) method was reliant on approximately Gaussian strain distributions, while the alternative approach was dependent on small strain gradients and broad crystallite-size distributions.

A number of contributors have proposed that $\langle \epsilon_L^2 \rangle$ is inversely proportional to L for cold-worked metals (see Rothman & Cohen 1971, Schlosberg & Cohen 1983, Adler & Houska 1979), expressed as

$$\langle \epsilon_L^2 \rangle = \frac{G^2(hkl)}{L} \tag{1.46}$$

where $G^2(hkl)$ is a constant of proportionality and is in general dependent on hkl. (Also also note that $G^2(hkl)$ has units of *length*.). For cold worked metals $G^2 \sim 10^{-4}\,\mathrm{nm}$ (Schlosberg & Cohen 1983, Nunzio et al. 1995). Taking the square-root of (1.46), $\langle \epsilon_L^2 \rangle^{\frac{1}{2}} \propto L^{-\frac{1}{2}}$. Adler & Houska (1979) showed that for cold-worked metals, the exponent was in the range of -0.380 to -0.610 and had an average value of -0.463 (also see the appendix, pp3286-3287 Adler & Houska 1979). van Berkum et al. (1994) pointed out that (1.46) corresponds to very large L, where the local strain consists of an ensemble of independent contributions.

In general there will be a direct correlation between the Fourier coefficients, (1.27) and the microscopic properties which give arise to models like (1.46) for specific ranges of L.

Crystallite models in Fourier analysis

The Fourier methods can be used, in a similar way to the Williamson & Hall (1953) method, to develop a qualitative understanding of specimen broadening. In the case of size broadening Louër et al. (1983) determined the Fourier coefficients of ZnO specimen profiles which showed

that size broadening was dominant and anisotropic in hkl. This confirmed the Williamson & Hall (1953) method. Furthermore, comparing the h0l Fourier coefficients showed that $\langle L \rangle_a$ increased with l while the Fourier coefficients for 100 and 110 were isotropic. This suggested that prismatic crystallites with a base which is isotropic in shape. Applying the analytical results from Langford & Louër (1982), the common-volume function for a cylinder was used to relate $\langle L \rangle_a$ to the dimensions of the cylinder. Louër et al. (1983) showed for the ZnO data the diameters and heights to be 154 Å and 213Å, respectively. Vargas et al. (1983) refined the results by applying a hexagonal prism model and calculated the sides to be of length 87 Å and height 213 Å. From Vargas et al. (1983), the average diameter of the hexagons (from face to face) is 151 Å, which is in very close agreement with the cylindrical model.

Vogel (1990) applied an analytical model derived for spherical crystallites to catalysts by expressing the Fourier coefficients in terms of a Hermite polynomial. Using this model, Vogel (1990) determined the crystallites distribution from (1.20). Rao & Houska (1986) used spherical crystallites and determined the column-length distribution assuming a log-normal diameter distribution.

Determining the column-length distribution, $p_a(L)$

Bienenstock (1961, 1963) considered the case of crystallite size broadening and related the Fourier coefficients of the specimen profile in terms of the column-length distribution, $p_a(L)$, *via* a finite difference equation as,

$$p_a(n) = \frac{A^s_{n+1} - 2A^s_n + A^s_{n-1}}{A_0 - A_1} \tag{1.47}$$

where is n is the Fourier harmonic-order and $p_a(n)$ is the column-length distribution expressed in terms of n. The difficulty with (1.47) is that it does not preserve the positivity of the distribution, resulting in ill-conditioning due to systematic and random errors in the Fourier coefficients. Essentially, (1.47) is a numerical method for determining the second derivative given in (1.20).

Young et al. (1967) used (1.47) to examine the effect of background over estimation on the Fourier method and $p_a(L)$. In this classic work, they showed that $\langle L \rangle_a$ and $p_a(L)$ were sensitive to truncation and background estimation. It was shown that 0.5% truncation caused a 5% error in $\langle L \rangle_a$, while a 20% truncation caused a 160% error in $\langle L \rangle_a$. That is, for

increasing truncation and background estimation, a "hook-effect" in the Fourier coefficients was produced. This causes an over-estimation of the $\langle L \rangle_a$. The effect this had on the $p_a(L)$ was to produce negative values, particular for $p_a(0)$. As the truncation or background estimation was increased the negativity in $p_a(0)$ also increased. Moreover, it is noticeable that $p_a(L)$ is shifted towards large L (see Figure 7, p159 Young et al. 1967). The reason for this is that the extended tails in a size-broadened profile characterises crystallite small, while the larger crystallite contribute to the main peak region of the profile. As the truncation or background estimation is increased, information concerning the smaller crystallites is being removed and the remaining information is related to large crystallites. The results for $p_a(L)$ are biased towards large L since the underlying information for the larger crystallites has not been affected by the truncation or background estimation.

Once $p_a(L)$ is known, it can be related to the volume-weighted distribution, $p_v(L)$, by (1.22). Like their respective size quantities , $p_a(L)$ and $p_v(L)$ are apparent size distributions, given in (1.25). That is, the distribution of columns for a single crystallite of a particular shape is "smeared" over the distributions of crystallites.

Developments towards formal and alternative methods in determining the column-length distribution were proposed by de Bergevin & Germi (1972), Păusescu et al. (1974) and Popescu & Benes (1977). These contributions have a number of common threads. The proposed method stems from the following observations

$$\begin{aligned} p_a(L) &\propto \frac{\mathrm{d}^2 A(L)}{\mathrm{d}L^2} \\ &\propto \frac{\mathrm{d}^2 V(L)}{\mathrm{d}L^2} \\ &\propto \int_{-\infty}^{+\infty} (s-s_0)^2\, f(s-s_0)\, \exp[2\pi\imath(s-s_0)L] \end{aligned} \tag{1.48}$$

where (1.48) is the second derivative of Fourier transform[6]. The difficulties encountered with this approach were that the presence of truncation, incorrect background estimation and noise produced spurious oscillations in the solution size-distribution. Various *ad hoc* approaches to removing the oscillations were applied, such as removing negative oscillations and deconvolving out the $\sin x/x$ terms which arise from truncation. These efforts were just treating the symptoms rather attempting to solve the underlying inverse problem.

[6]Given $\frac{\mathrm{d}^n}{\mathrm{d}x^n}\mathcal{F}\{f(x)\} = x^n\,\mathcal{F}\{f(x)\}$ (also see Arfken & Weber 1996)

The contribution by LeBail & Louër (1978) is significant because it treats the problem of determining the column-length distributions, $p_a(L)$, as an inverse problem. Using (1.26), LeBail & Louër (1978) realised that (1.26) can be expressed in terms of a matrix-vector product as

$$\mathbf{f} = \mathbf{K}\mathbf{p}_a \tag{1.49}$$

where $\mathbf{f}$ is the specimen profile expressed as an $[N \times 1]$ vector; $\mathbf{K}$ is the scattering kernel expressed in terms of an $[N \times M]$ matrix, where $M \leq N$; and $\mathbf{p}_a$ is the column-length distribution expressed as an $[M \times 1]$ vector. The temptation is to express the column-length distribution as $\mathbf{p}_a = (\mathbf{K}^T\mathbf{K})^{-1}\mathbf{K}^{\mathbf{T}}\mathbf{f}$. However it was realised by LeBail & Louër (1978) that this produces an ill-conditioned solution.

Transforming the size-broadened profile into (1.49) enabled a regularisation inversion/deconvolution method, applied by Louër et al. (1969), Louër & Weigel (1969) and Louboutin & Louër (1972) in removing the instrumental broadening, to determine the unknown $\mathbf{p}_a$. This involved maximizing the regularisation function, $\mathbf{p}_a^T\mathbf{H}\mathbf{p}_a$ with respect to the misfitting function $(\mathbf{f} - \mathbf{K}\mathbf{p}_a)^T(\mathbf{f} - \mathbf{K}\mathbf{p}_a)$ (Regularisation/deconvolution methods will be discussed in detail in Chapter 2.) It should be noted that the solution to the inverse problem proposed by LeBail & Louër (1978), in general, does not preserved the positivity of $p_a(L)$. This is important because any subsequent calculations assume that $p_a(L)$ is positive definite and any negative values or truncation of $p_a(L)$ to eliminate the negative values may corrupt the final results.

LeBail & Louër (1978) applied regularisation/deconvolution methods successfully in determining the column-length distribution for $Ni(OH)_2$ crystallites and obtained good agreement between the area-averaged column-length size for it and the Fourier method (see Table 3, p54 LeBail & Louër 1978). It is interesting to point out that the regularization inverse/deconvolution method (Louër et al. 1969, Louër & Weigel 1969, Louboutin & Louër 1972) was also applied to removing instrumental broadening, suggesting a "two-fold" approach to determining the size-broadened specimen profile and the corresponding crystallite size. Louër et al. (1983) applied the inversion/deconvolution method to determining the column-length distributions for ZnO crystallites and showed that the column-length distributions in the xy-plane of the ZnO were isotropic, confirming the results from integral breadth and Fourier analyses that the crystallites had an isotropic base.

Guérin et al. (1986) applied a statistical inference method to determine the column–length distribution. This approach is an application of the maximum entropy method developed by Jaynes (1957). The same approach was applied by Frieden (1972) in deconvolving blurred spectra. For discrete data the specimen profile was expressed as

$$f_i = \sum_{j=1}^{M} K_{ij} p_{a\,j} \tag{1.50}$$

for $i = 1, 2, \ldots, M$ and

$$p_{a\,j} = \exp\alpha_0\ \exp\left[-\sum_k \alpha_k K_{jk}\right] \tag{1.51}$$

where $\{\alpha_k; k = 1, 2, \ldots, M\}$ are a set of unknown Lagrangian parameters. Effectively, there are $M+1$ nonlinear equations with $M+1$ unknowns (which includes α_0). Usually numerical methods such as the Newton-Raphson algorithm are applied in similar problems (Burden & Faires 1993). However, Frieden (1972) reported that good estimates of the initial values of $\{\alpha_k\}$ are needed for the Newton-Raphson algorithm to converge onto the correct solution. Once the $\{\alpha_k\}$ are know then $\{p_{a\,j}; j = 1, 2, \ldots, M\}$ can be determined from (1.51). Determining the full set of Lagrangian parameters $\{\alpha_k\}$ is important in ensuring the uniqueness of the solution and a global maximum. The presence of statistical noise in the data further complicates the problem.

However, Guérin et al. (1986) claimed that only three Lagrangian parameters are needed to determine the column-length distribution. Given the above discussion, in general, it seems doubtful that the solutions proposed by Guérin et al. (1986) would represent the underlying physical phenomena. Bonetto et al. (1993) developed a computer program, which determined the set of Lagrangian parameters. Although Guérin et al. (1986) employed the basic results of the maximum entropy method, there was no attempt to include the *a priori* model, apply the method to removing instrumental broadening and to incorporate the statistical noise of the experimental data. Lutterotti & Scardi (1992) attempted to generalise Guérin et al. (1986), by including an analytical expressions for the strain. However, this assumes a particular analytical strain model.

Nunzio et al. (1995) has proposed a Monte-Carlo method for determining the column-length distribution and the $\langle\epsilon_L\rangle^{\frac{1}{2}}$-distribution from the broadened x-ray diffraction profile. Their approach essentially determines a frequency distribution, $W(L_i, \epsilon_i)$, which includes

both the strain and size information. By determining the frequency distribution $W(L_i, \epsilon_j)$, the column-length distribution, volume-averaged size, $\langle L \rangle_v$ and rms-strain $\langle \epsilon_L^2 \rangle^{\frac{1}{2}}$ can be determined (see equations (6-9), p148 Nunzio et al. 1995). The difficulties encountered by Nunzio et al. (1995) were that the Monte Carlo method can produce a number of possible solutions, rather than a unique solution. In order to overcome this, $W(L_i, \epsilon_i)$ was determined which maximizes the size distribution.

Determining the strain distribution, $p(\epsilon_L)$

In comparison to the problem of determining the column-length distribution, determining the strain distribution has received very little attention. Although they are both inverse problems, determining the strain distribution is very challenging. McKeehan & Warren (1953) applied the Warren & Averbach (1950, 1952) method to cold-worked tungsten. The size strain Fourier coefficients were determined. In this case tungsten is an elastically isotropic material, hence strain broadening will also be isotropic (Ryaboshapka & Tikhonov 1961*a*). McKeehan & Warren (1953) noticed that (1.28) was a cosine Fourier transform and the inverse will result in $p(\epsilon_L)$, for a given L

$$p(\epsilon_L) = 2L \int_0^{\infty} \mathrm{d}d^* \, A^d(L,\, d^*) \cos 2\pi L \epsilon_L d^*. \tag{1.52}$$

The strain distributions were determined for L = 125, 150, 175, 200 Å and were argued to be Gaussian in shape; however, some truncation in the tails of the result is noticeable (see Figure 5, p55 McKeehan & Warren 1953). Using the strain distribution the $\langle \epsilon_L^2 \rangle^{\frac{1}{2}}$-values were compared with the values determined from (1.44) and were in agreement with each other. The difficulty of determining $p(\epsilon_L)$ from $A^s(L,\, d^*)$ is the insufficient information. That is, in a diffraction pattern a limited number of line-profiles can be extracted and for a given L the $A^s(L,\, d^*)$ must be interpolated as a function of d^*. This is especially the case for high-symmetry lattice systems. In other words, inverse cosine Fourier transform in (1.52) assumes that the information for $A^s(L,\, d^*)$ is complete for all L and d^*; however, in dealing with experimental data this not the case.

Another difficulty that is faced with trying to determine $p(\epsilon_L)$ is that it is reliant on isotropic strain broadening. In the case of anisotropic strain broadening, at best three orders of diffraction from a conventional diffractometer can be recorded and there is insufficient

data to determine $p(\epsilon_L)$. Eastabrook & Wilson (1952) outline what information could be developed concerning $p(\epsilon_L)$. Warren (1959) also noted the difficultly in determining $p(\epsilon_L)$. Harrison (1966, 1967) and Ungár (1967) determined $p(\epsilon_L)$ using a method of moments of the line-profile. Essentially, the strain moments were determined from the strain coefficients and used to determine the coefficients for Hermite polynomial together with a suitable weight function that represented the strain distribution. Harrison (1966) reported that this approach needed at least three reflections. Ungár (1967) observed that the crystallite size information could also be determined using this approach. It also should be noted that very little work has been done is correlating the strain distribution with the type, density and arrangement of dislocations (see Krivoglaz 1995).

1.4.4 Recent developments

Recent developments in the crystallite and strain analysis has seen the rise of two methods for analysing broadened profiles. Generally, these include the work of van Berkum et al. (1996) and Ungár & Borbély (1996), Ungár, Ott, Sanders, Borbély & Weertman (1998, 1998). These recent development had their origins in the work by Ryaboshapka & Tikhonov (1961*a*, 1961*b*), Wilkens (1970*c*, 1970*b*, 1970*a*, 1987), Groma et al. (1988), Ungár et al. (1989), Krivoglaz (1995), Delhez & Mittemeijer (1976) and van Berkum (1994). However, they represent a clear alternative to the tradational Williamson & Hall (1953) and Warren & Averbach (1950, 1952) method by correlating the integral breadth and Fourier methods to the microscopic picture that gives rise to line-profile broadening.

van Berkum method

van Berkum (1994) based his approach on the kinematic Fourier representation of size and strain broadening given in §1.3.2, but proposed to relate the strain broadening to a "distortion field" projected onto the **x**-axis in the direction of the diffraction vector. The distortions were assumed to be spaced by a distance D_i projected onto the axis, with an average spacing of $\langle D \rangle$. The distortions were averaged over the region by

$$Z_L(x) = \int_{x-L/2}^{x+L/2} \epsilon(x')\,\mathrm{d}x' \tag{1.53}$$

where $\epsilon(x)$ is the strain projeced onto the $\mathbf{x}$-axis. The proposed model relates the Fourier coefficients for a periodically distributed set of distortions. The distortions included screw dislocations and assumed that the specimen was elastically isotropic. The formulation is general in that different dislocations could be included. This resulted in the Fourier coefficients for this model was expressed in terms of the $\langle\epsilon^2\rangle$ and the dimensionless parameters: $L_r \equiv L/\langle D\rangle$, the relative distance parallel to the diffraction vector; $w_r \equiv w/\langle D\rangle$, the relative width of the disortions (see equation (7), p733 van Berkum 1994). The integral breadths were determined using (1.21). As pointed out in §1.3.2, the relative width, w_r, was used to define the region of validity for the method. That is, depending on the value of w_r, Gaussian and non-Gaussian strain broadening, as well as domain broadening, had been taken into consideration. The method proposed by van Berkum (1994) involved fitting the Fourier coefficients and integral breadths to the relevant experimental data and refining on the unknown parameters. The van Berkum (1994) model was applied to ball-milled tungsten which showed isotropic broadening and using all the avaliable line profiles, the mean spacing of the distortion was determined to be $\langle D\rangle = 21\,\text{nm}$; the relative width of the distortions as $w_r = 0.115$; and $\langle\epsilon^2\rangle^{\frac{1}{2}} = 6.8\times10^{-3}$. Using the value for w_r, a prediction was made of the values the Williamson & Hall (1953) and Warren & Averbach (1950, 1952) methods would yield. The Williamson & Hall (1953) method predicted $\langle D\rangle = 33\,\text{nm}$, $\langle\epsilon^2\rangle^{\frac{1}{2}} = 2.2\times10^{-3}$, compared to the experimental values of $62\,\text{nm}$ and 2.5×10^{-3}, respectively. The difference between the calculated and experimental $\langle D\rangle$ was explained by the assumption of spatial periodicity for the distortions in the new method. When applied to predicting the expected Warren & Averbach (1950, 1952) results, the new method gave $\langle D\rangle = 18\,\text{nm}$ and $\langle\epsilon^2\rangle^{\frac{1}{2}} = 1.6\times10^{-3}$ identical to the experimental. In this case the method proposed by van Berkum (1994) was successful in yielding physically meaningful results as to the spacing, width and rms-strain of the distortions; it also predicted how the traditional method would preform.

The difficulty with this appraoch is it assumed the strain broadening is isotropic. It was suggested that in the case of anisotropic strain broadening multiple orders should be used. However, in most cases the strain broadening is likely to be anisotropic, which suggests that a microscopic description of the anisotropic broadening is lacking.

Ungár method

The method developed by Ungár & Borbély (1996), Ungár, Ott, Sanders, Borbély & Weertman (1998) and Ungár, Révész & Borbély (1998) applied to randomly distributed dislocation for isotropic and anisotropic materials. Building on the results of Groma et al. (1988) and Ungár et al. (1989), the Fourier coefficients of the broadened profile are given as

$$\begin{aligned} \ln A(L,\, d^*) \quad &\approx \quad \ln A^s(L) - \rho B L^2 \ln(R_e/L)(d^{*2}\langle C\rangle) \\ &+ \quad Q B^2 L^4 \ln(R_1/L)\ln(R_2/L)(d^{*2}\langle C\rangle)^2 \\ &\pm O(L^6) \end{aligned} \tag{1.54}$$

where ρ is the dislocation density; $\langle C\rangle$ is the average contrast factor; Q is related to the two particle correlations in the dislocation ensemble; $B = \pi b^2/2$ for the Burgers vector $\mathbf{b}$; and R_e, R_1, R_2 are the effective cut-off radii of the dislocations and auxiliary parameters, respectively. The expression for the integral breadth included the same terms (see equation (8), p3694, Ungár, Révész & Borbély 1998). The critical term in (1.54) is the average contrast factor, $\langle C\rangle$, which defines the elastic propertities of the material for a particular dislocation and Burgers vector, $\mathbf{b}$, in the direction of the diffraction vector (see equations (3)-(5), p3694, Ungár, Révész & Borbély 1998). Ungár & Borbély (1996) gave the $\langle C\rangle$-values for copper (see Table 1, p3173 Ungár & Borbély 1996). In the case of a specimen consisting of different types of dislocations, $\langle C\rangle$ will be the weighted sum of the $\langle C\rangle$ from different types of dislocation, where the weight is the relative population of the dislocation in the specimen. This enables the type of dislocations in the specimen to be inferred from the fitting of the experimental data. For isotropic material the dependency on $\langle C\rangle$ is weak, while for anisotropic material the dependency is strong.

By plotting $\ln A(L,\, d^*)$ versus $d^{*2}\langle C\rangle$ for a given L, the unknown parameters can be determined. Similarly, the integral breadths can be plotted as a function of $d^{*2}\langle C\rangle$. Here we notice a distinct difference from the Williamson & Hall (1953) and Warren & Averbach (1950, 1952) methods, in that the constrast factors define a suitable variable to quantify the broadening. The consequence of this is that Fourier or integral breadth results which show no systematic behaviour when graphed as a function of d^{*2}, and traditionally would involve using the multiple-orders, exhibit a systematic behaviour when graphed in terms

of $d^{*2}\langle C\rangle$ (compare Figures 2 & 3 and Figures 6 & 8, pp 3697-3698, Ungár, Révész & Borbély 1998). Ungár & Borbély (1996) applied the method to copper deformed by an angular press, producing a crystallite/domain size in the rangle of 75 to 385 nm and a dislocation density, $\rho = 1.7 \times 10^{15}\,\mathrm{m}^{-2}$. These results compared favourable to the electron transmission results. Ungár, Révész & Borbély (1998) reported that the effective cut-off radius of dislocations, R_e, was greater than the crystallite/domain size. It was explained as a statistical parameter associated with the ensemble of dislocations.

The disadvantage that this method may suffer is the complexity and difficulty in determining the $\langle C\rangle$-values for a range of materials may make it inaccessible for practical calculations. Moreover, the method assumes that the dislocations are randomly distributed on all slip system. However, as van Berkum (1994) has shown, the dislocations can be approximated to have a spatial periodicity. This will also have implications for the physical values determined from the model.

An alternative approach

In the proceding sections, the underlying theory and methods for size/strain analysis of broadened profiles have been discussed. It is interesting to observe a commonality in the expressions for observed profile, (1.26), the size-broadened profile, (1.34), and strain broadening, (1.28). The common feature in these expressions is that the distributions of interest, $f(2\theta)$, $p_a(L)$ and $p(\epsilon_L)$ are all positive definite. Furthermore, each expression is an integral equation of the first kind, where in (1.34) we assume the background level has been removed, and (1.26) is simply a discrete form of the continuous result. Determining these distributions entails numerically solving an inverse problem.

Suppose that these distributions could be determined with a suitable method which makes the least assumptions about the distributions, preserves the positivity, incorporates the statistical noise and (if available) includes *a priori* information of the distributions; this would enable information concerning the specimen profile, column-length distribution, and strain distribution to be determined.Using these distributions a greater understanding of the variation of the strain or crystallite size over a particular range could be developed rather than just relying on the corresponding average values. For example, determining the

specimen profile is the first step in determining microscopic size and/or strain contributions which could be used in the various methods described above. Mathematically speaking, it is a simple case of an integral equation, where the instrument profile is assumed to be shift invariant. Furthermore, from the similarity between the integral equations, it seems plausible that in the case of size-broadend profiles, the inverse method could be applied again to determine the column-length distribution, $p_a(L)$.

This data is useful in the production of catalysts, for example, where the researcher can design the material to have a specific range of crystallite sizes. In the case of strain broadening data, the same method could be used again to determine $p(\epsilon_L)$ and the rms-strain, rather than relying on the expansion of the Fourier coefficients. Although $p_a(L)$ and $p(\epsilon_L)$ are apparent quantities, there is an opportunity to test theoretical models by incorporating them as *a priori* information in the inverse method and testing their plausibility. Using this approach an inference can be made as to the underlying physical quantities. That is, the commonality of (1.26), (1.34) and (1.28) suggests a single method could be developed to determine each distribution.

The constrained/inversion and iterative deconvolution method outlined in §1.4.1, could be used to determine $f(2\theta)$, $p_a(L)$ and $p(\epsilon_L)$. However, these methods will not, in general, preserve the positivity of the $f(2\theta)$, $p_a(L)$ and $p(\epsilon_L)$ and are restricted in the use of *a priori* information. A detailed analysis of the deconvolution methods outlined in §1.4.1 is presented in Chapter 2 for the case of determining $f(2\theta)$. The results from this analysis could also be adapted for the cases of $p_a(L)$ and $p(\epsilon_L)$. An alterative approach to determine the $f(2\theta)$, $p_a(L)$ and $p(\epsilon_L)$ is presented in Chapter 3. This method is the maximum entropy method. The fundamental characteristics of the entropy function ensure that the resulting distributions are positive definite. It is also extended to incorparate the *a priori* information, the statistical nature of the noise, and to quantify the regions of uncertainty in the solution. In Chapters 4 & 5, the maximum entropy method for determining the specimen profile, column-length and strain distributions is developed and tested using simulated broadened profiles. This maximum entropy method is also applied to alumina x-ray diffraction data where information concerning the crystallite size distribution is determined and compared with traditional methods.

1.5 Summary

In this chapter a qualitative overview of broadened line profiles has been developed and it has been outlined how the intensity distribution in a three dimensional space is mapped into one-dimensional (see 1.2.1). The integral breadth and Fourier representations for broadened line-profiles have been developed and it has been shown how the broadened profile is related to the underlying structural imperfections. It is seen from this representation that line broadening is associated with apparent quantities and relies on a model to be adopted in order for any physical interpretaion to be made. This became clear in how the two dominant approaches, which apply the integral breadth and Fourier representations, determine the apparent quantities (see §1.4.2 & 1.4.3). In the case of size broadening, it has been shown how the integral breadth and Fourier representations can be used to develop a qualitative and quantitative picture of the specimen broadening, shape and distribution of crystallites. However, for the case of strain broadening the problem of relating the specimen broadening to the microscopic propertities of the specimen is very difficult. This is due to the ensemble nature of distortions, like dislocations, and corresponding strain fields. It is difficult to interpret the rms-strain unless a suitable model is adopted. Recent developments have attempted to resolve this problem and have been reviewed. The initial results are encouraging; however, due to their mathematical complexity or the limited physical cases that are considered, the practical value may not be fully realised.

Chapter 2

Analysis of Deconvolution Methods used in X-ray Diffraction

Various deconvolution methods were outlined in Chapter 1. It was pointed out that there are two approaches in determining the specimen profile. These are the direct convolution (Howard & Snyder 1989, Cheary & Coelho 1992) and various deconvolution/inversion methods (see 1.4.1). In this chapter, an analysis of the common deconvolution/inversion methods used in x-ray diffraction is presented. This analysis has been adapted from a general discussion by Lagendijk & Biemond (1991). It is also based on the discussion presented by Armstrong & Kalceff (1998), concentrating on an eigensystem analysis of x-ray deconvolution methods which include: unconstrained matrix inversion and deconvolution (Stokes 1948); iterative deconvolution (Ergun 1968) and the constrained deconvolution methods (Louër et al. 1969, Louër & Weigel 1969, Louboutin & Louër 1972). These deconvolution/inversion methods are *linear inversion* methods. That is, they can be expressed as a linear combination of a set of orthonormal vectors. This enables the formulation of an *error-bound* function, which defines the maximum difference between the solution and true specimen profiles.

The principle aim of this chapter is to address an important theoretical issue in profile analysis, concerning the mathematical behaviour of the deconvolution methods and the effect individual terms have on the solution profile. The central issue which emerges from the analysis and numerical results is the *conditioning* of the solution profile. This conditioning raises questions as to the integrity of the solution profile and the physical quantities (i.e.

crystallite size and rms-strain) that can be "extracted" from a specimen profile. Although this chapter focuses on the issue of determining the specimen profile, the analysis is equally applicable in determining the column-length and strain distributions. In this sense the issues discussed here have a greater importance concerning the conditioning of the column-length and strain distributions.

The eigensystem analysis reveals that the error-bound function expresses the sources of misfitting and ill-conditioning as a function of the control parameters for a particular method. In addition, it shows that even for the optimum values of certain control parameters, positivity is not preserved and spurious oscillations appear in the solution profile. The results from the analytical discussion presented in §2.1 are illustrated with a simulation of an instrument-broadened diffraction profile (see §2.2). The implications that the analytical and computational results have on experimental data are discussed in §2.2.5.

2.1 Theory

2.1.1 Background

As discussed in the previous chapter, an observed diffraction profile, g, can be written as the convolution of an instrument function, k, and a specimen profile, f, superimposed onto a background signal, b, and a random noise distribution, n,

$$g(2\theta) = \int k(2\theta - 2\theta')\, f(2\theta')\, \mathrm{d}(2\theta') + b(2\theta) + n(2\theta). \tag{2.1}$$

From this we seek to determine $f(2\theta')$, the specimen profile whose shape and width are characterized by properties of the material, such as microstrain and crystallite or scattering domain size. Physically, it is expected that the positivity of the specimen profile would be preserved and that the integrity of the physical quantities, such as crystallite/domain size and microstrain, would not be compromised by spurious oscillations occurring in the solution profile.

The instrument profile, k, is assumed to be a linear shift-invariant function. That is, the shape of the function does not change over the region of integration and only depends on the difference $(2\theta - 2\theta')$. The function, k, can be thought of as mapping the specimen profile, f,

from a $2\theta'$-space into a 2θ-space. This results in the blurring (i.e. convolution) of f, defined by g, when measured in 2θ-space. In general, k, will be an asymmetrical function about the Bragg angle. However, in this chapter, only symmetrical instrument profiles are considered. The reason for this will become apparent in the analysis that follows.

The convolution equation (2.1) refers to a set of continuous functions associated with a *linear function space* (Wing 1991). However, the measurements of the observed and instrument profiles are made at discrete time intervals. To convey this, we apply a discrete representation of (2.1) (Arfken & Weber 1996, Wing 1991),

$$g_i = \sum_j k_{ij}\, w_j\, f_j + b_i + n_i \tag{2.2}$$

where g_i is the i-th element of observed profile; k_{ij} are the $(i,\ j)$th elements of the instrument profile calculated over a large interval for 2θ and $2\theta'$, such that, $k_{ij} = k(2\theta_i - 2\theta'_j)$ where $i,\ j = 1, 2, 3, \ldots, N$; w_j is the j-th element from a chosen quadrature scheme and accounts for the integration in (2.1); f_j are the j-th elements of the specimen profile, while b_i and n_i are the background level and the noise distribution, respectively.

Alternatively, the discrete form of (2.1), given by (2.2) can be expressed using column vectors and a matrix. That is, the specimen and observed profiles are represented as column vectors and the instrument profile as a convolution matrix by employing the *bra-ket* notation. In the *bra-ket* notation the bra vector, $\langle v|$, represents a row vector, while the ket vector, $|v\rangle$, represents a column vector. The inner product for two ket-vectors, $|v\rangle$ and $|u\rangle$, that span a real, N-dimensional Euclidean space, $\mathbb{E}$, is defined as $\langle v|u\rangle = \sum_{i=1}^{N} v_i\, u_i$ where v_i and u_i are the i-th components of the vectors and $i = 1, 2, 3, \ldots N$. Also, in $\mathbb{E}$, the L_2-norm for $|v\rangle$ is defined as $\|\mathbf{v}\| = |\langle v|v\rangle|^{\frac{1}{2}} = \sqrt{\sum_{i=1}^{N} v_i^2}$.

Thus, (2.1) and (2.2) become

$$|g\rangle = \mathbf{R}|f\rangle + |b\rangle + |n\rangle \tag{2.3}$$

where $|g\rangle$ represents the observed profile, such that $|g\rangle = \{g_i, i = 1, \ldots, N\}$; $\mathbf{R}$ is the convolution matrix which includes the terms generated from $k(2\theta - 2\theta')$ and a quadrature such that $(\mathbf{R})_{ij} = k_{ij} w_j$; $|f\rangle$ is the specimen profile; $|b\rangle$ is the background count; and $|n\rangle$ is the noise distribution. We define $\mathbf{R}$ as a real, nonsingular, symmetric $[N \times N]$ matrix, where $\{|v_i\rangle,\ i = 1, \ldots, N\}$ are the set of orthonormal eigenvectors belonging to $\mathbf{R}$, and λ_i are the

corresponding real and non-zero eigenvalues for i. The eigenvectors of $\mathbf{R}$ define a set of basis vectors in $\mathbb{E}$. Hence, any vector in this space can be expressed as a linear combination of the eigenvectors of $\mathbf{R}$. By decomposing the solution using an eigen-system analysis, the solution profile is simply represented as a linear combination of the eigenvectors, and an understanding of the deconvolution methods in this space can be developed (see equations (2.8), (2.29) & (2.39)). The conditioning of this space can be understood by examining the eigenvalues, $|\lambda_i|$. The eigenvalues can be arranged in a monotonically decreasing sequence, $|\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_N|$. If $|\lambda_i| \to 0$ for $i \to N$, then the matrix R becomes near-singular and its inverse will result in an ill-conditioned matrix. It is this characteristic which contributes to noise amplification (see equations (2.13), (2.32) & (2.41)). Another consequence of $|\lambda_i| \to 0$ as $i \to N$ is that $|v_i\rangle$ becomes linearly dependent, and the uniqueness of the solution is compromised. Defining $\mathbf{R}$ as a real and symmetric matrix is equivalent to saying that the instrument profile is symmetric. However, the generality of the results presented in this chapter can be qualitatively extended to the case of an asymmetric instrument profile.

Before the deconvolution of (2.1) or the inversion of (2.3) is undertaken, the background signal is usually estimated and subtracted from the observed data. The error that results from this estimation propagates into the deconvolution and any subsequent calculations. To reflect this, we define the background-corrected observed data, $|\tilde{g}\rangle$, as

$$|\tilde{g}\rangle = |g\rangle - |\hat{b}\rangle \tag{2.4}$$

where $|\hat{b}\rangle$ is the estimated background level. The specimen profile is usually determined from this corrected data, $|\tilde{g}\rangle$.

2.1.2 Unconstrained Inversion/Deconvolution Methods

Unconstrained Matrix Inversion

The eigen and error-bound function analysis proceeds as follows: suppose that we only know $|\tilde{g}\rangle$ and $\mathbf{R}$, and we make no assumptions about the noise. We write the solution profile, $|\hat{f}\rangle$ in the form

$$\mathbf{R}|\hat{f}\rangle = |\tilde{g}\rangle \tag{2.5}$$

where taking the inverse of $\mathbf{R}$ will produce $|\hat{f}\rangle$. Multiplying (2.5) by the eigenvector $\langle v_i|$ and using the result $\langle v_i|\mathbf{R} = \lambda_i \langle v_i|$ and then multiplying both sides by λ_i^{-1} produces

$$\langle v_i|\hat{f}\rangle = \lambda_i^{-1} \langle v_i|\tilde{g}\rangle. \tag{2.6}$$

We can now express $|\hat{f}\rangle$ as a linear combination of the eigenvectors, $|v_i\rangle$, (see Arfken & Weber 1996, Kreyszig 1993) by multiplying both sides by $|v_i\rangle$ and summing over all i. We arrive at

$$\sum_{i=1}^{N}\langle v_i|\hat{f}\rangle\,|v_i\rangle = \sum_{i=1}^{N}\lambda_i^{-1}\langle v_i|\tilde{g}\rangle\,|v_i\rangle. \tag{2.7}$$

The left hand side of (2.7) is simply $|\hat{f}\rangle$. Substituting (2.3) and (2.4) into the right hand side of (2.7), we have

$$\begin{aligned} |\hat{f}\rangle &= \sum_{i=1}^{N}\lambda_i^{-1}\langle v_i|\left(R|f\rangle + |b\rangle + |n\rangle - |\hat{b}\rangle\right)\,|v_i\rangle && (2.8)\\ &= \sum_{i=1}^{N}\langle v_i|f\rangle\,|v_i\rangle + \sum_{i=1}^{N}(-1)\lambda_i^{-1}\langle v_i|e_{\hat{b}b}\rangle\,|v_i\rangle + \sum_{i=1}^{N}\lambda_i^{-1}\langle v_i|n\rangle\,|v_i\rangle && (2.9) \end{aligned}$$

where $|e_{\hat{b}b}\rangle = |\hat{b}\rangle - |b\rangle$ is the misfitting error between the *estimated* background level and *true* background level. The first term in (2.9) is simply $|f\rangle$. The error between the solution and true specimen profiles is $|\hat{f}\rangle - |f\rangle$. We can express

$$\begin{aligned} |\hat{f}\rangle - |f\rangle &= |e_{\hat{f}f}\rangle \\ &= \sum_{i=1}^{N}(-1)\lambda_i^{-1}\langle v_i|e_{\hat{b}b}\rangle\,|v_i\rangle + \sum_{i=1}^{N}\lambda_i^{-1}\langle v_i|n\rangle\,|v_i\rangle \\ &= \sum_{i=1}^{N}[\lambda_i^{-1}\left((-1)\langle v_i|e_{\hat{b}b}\rangle + \langle v_i|n\rangle\right)]\,|v_i\rangle. && (2.10) \end{aligned}$$

The terms in the square brackets of (2.10) are the components of the error vector. Taking the norm of (2.10) defines the error-bound function (see Kreyszig 1993). We have,

$$\begin{aligned} \|\hat{\mathbf{f}} - \mathbf{f}\| &= \left[\sum_{i=1}^{N}\lambda_i^{-2}\left(\langle v_i|n\rangle + (-1)\langle v_i|e_{\hat{b}b}\rangle\right)^2\right]^{\frac{1}{2}} && (2.11)\\ &\le \left[\sum_{i=1}^{N}\lambda_i^{-2}\langle v_i|n\rangle^2\right]^{\frac{1}{2}} + \left[\sum_{i=1}^{N}\lambda_i^{-2}\langle v_i|e_{\hat{b}b}\rangle^2\right]^{\frac{1}{2}} && (2.12)\\ &\le \sum_{i=1}^{N}\left|\lambda_i^{-1}\right|\,|\langle v_i|n\rangle| + \sum_{i=1}^{N}\left|\lambda_i^{-1}\right|\,|\langle v_i|e_{\hat{b}b}\rangle| && (2.13) \end{aligned}$$

where (2.12) is arrived at by using the Schwarz inequality[1] while the result in (2.13) is arrived at by employing the triangular inequality. For notational convenience we have, $\hat{\mathbf{f}} \equiv |\hat{f}\rangle$ and $\mathbf{f} \equiv |f\rangle$.

The first term in (2.13) represents amplification of the noise, resulting in the ill-conditioning of $|\hat{f}\rangle$. That is, unphysical oscillations and artefacts corrupt the solution profile. This comes about if $|\lambda_i| \to 0$ and the $|\lambda_i^{-1}|$ in (2.13) amplifies the noise term. Similarly, the second term in (2.13) represents the amplification of the misfitting error in the background estimation. A similar argument can be made for the Stokes (1948) deconvolution method.

Stokes (1948) Method

The Stokes (1948) method is also an inversion method, with the unconstrained inversion taking place in the Fourier space of the profiles. It has been shown in numerical studies that this method produces an ill-conditioned specimen profile when deconvolving in the presence of noise (for example see Kalceff et al. 1995). In this section a detailed analysis of the Stokes (1948) method is performed. However, the Fourier basis functions and Fourier coefficients are defined in place of the eigen-vectors and -values.

The Fourier series of a piecewise-continuous function, $f(x)$, is defined over the interval $x \in [-a,\, a]$, and is given by,

$$f(x) \;=\; \sum_m \mathsf{F}_m \,\exp(\imath\pi x m/a) \tag{2.14}$$

where the summation is over $m = 0, \pm1, \pm2, \pm3\ldots$, $\imath = \sqrt{-1}$ and F_m are the complex Fourier coefficients of $f(x)$ given as,

$$\mathsf{F}_m \;=\; \frac{1}{2a}\int_{-a}^{+a} f(x)\,\exp(-\imath\pi x m/a)\,\mathrm{d}x \tag{2.15}$$

for $m = 0, \pm1, \pm2, \pm3\ldots$.

[1]The Schwarz inequality is given by, $\left|\sum_{i=1}^{N} a_i\, b_i\right| \;\le\; \left[\sum_{i=1}^{N} a_i^2\right]^{\frac{1}{2}} \left[\sum_{i=1}^{N} b_i^2\right]^{\frac{1}{2}}$ where $a_i,\, b_i \,\in \mathbb{R}$ for all i (see Kreyszig 1993). The result in (2.12) is arrived at by the following corollary (which requires the Schwarz inequality to prove it), $\sqrt{\sum_{i=1}^{N}(a_i+b_i)^2} \;\le\; \left[\sum_{i=1}^{N} a_i^2\right]^{\frac{1}{2}} + \left[\sum_{i=1}^{N} b_i^2\right]^{\frac{1}{2}}$. This result can be further generalized for cases arising in (2.32) and (2.42).

The Stokes (1948) method can be derived by expressing the instrument and background-corrected profiles in terms of their Fourier coefficients (see Warren 1969). It is given by

$$\hat{\mathsf{F}}_m \mathsf{K}_m = \tilde{\mathsf{G}}_m, \quad \text{for } m = 0, \pm 1, \pm 2, \pm 3 \ldots \tag{2.16}$$

where $\hat{\mathsf{F}}_m$, K_m and $\tilde{\mathsf{G}}_m$ are the complex Fourier coefficients for the solution specimen profile, the instrument profile and the corrected observed profile, respectively. Taking the inverse of K_m in (2.16) will result in the Fourier coefficients for the solution profile. At this point the coefficients are used in the Warren-Averabach method (Warren & Averbach 1950, Warren & Averbach 1952, Warren 1969). The result from the Stokes (1948) method is analogous to (2.5) in its form and the fact that no assumptions concerning the solution or the noise distribution have been made.

In order to determine the error-bound function for the Stokes (1948) method, similar steps as those carried out in the matrix inversion method are followed, with the exception that the space which defines the operations is a linear function space. The next step assumes that the complex Fourier coefficients for $b(x)$, $\hat{b}(x)$ and $n(x)$ exist. In general this will be the case and the complex coefficients for each term in (2.1) and (2.4) are substituted into the right-hand side of (2.16) and by dividing both sides by K_m, where $|\mathsf{K}_m| \neq 0$ for all m, results in

$$\hat{\mathsf{F}}_m = \mathsf{F}_m + (-1)\mathsf{E}_{\hat{\mathsf{B}}\mathsf{B}\,m}/\mathsf{K}_m + \mathsf{N}_m/\mathsf{K}_m \tag{2.17}$$

where $\mathsf{E}_{\hat{\mathsf{B}}\mathsf{B}\,m}$ is the error in the complex Fourier coefficients between the estimated and true background level and N_m are the complex Fourier coefficients for the noise distribution. The error between the solution and true Fourier coefficients of the specimen profile is given by

$$\hat{\mathsf{F}}_m - \mathsf{F}_m = (-1)\mathsf{E}_{\hat{\mathsf{B}}\mathsf{B}\,m}/\mathsf{K}_m + \mathsf{N}_m/\mathsf{K}_m \tag{2.18}$$

where multiplying both sides by $\exp(\imath\pi mx/a)$ and summing over all m transforms the error from the Fourier space and results in

$$\begin{aligned} \hat{f}(x) - f(x) &= e_{\hat{f}f}(x) \\ &= \sum_m [\hat{\mathsf{F}}_m - \mathsf{F}_m] \exp(\imath\pi mx/a) \\ &= \sum_m [(-1)\mathsf{E}_{\hat{\mathsf{B}}\mathsf{B}\,m}/\mathsf{K}_m + \mathsf{N}_m/\mathsf{K}_m] \exp(\imath\pi mx/a) \end{aligned} \tag{2.19}$$

where $e_{\hat{f}f}$ is the error function defined over the range of x. The L_2-norm for (2.19) is defined in a continuous space, given by $\|e(x)\| = \sqrt{\int_{-a}^{a} |e(x)|^2 dx}$. Applying this definition to the results in (2.19) and taking into consideration that the terms in the square-brackets are complex, we arrive at

$$\begin{aligned} \| \hat{f}(x) - f(x) \| &= \Bigg[\sum_{m'} \sum_{m} [(-1)\mathsf{E}_{\hat{\mathsf{B}}\mathsf{B}\,m'}/\mathsf{K}_{m'} + \mathsf{N}_{m'}/\mathsf{K}_{m'}]^* \\ &\times \ [(-1)\mathsf{E}_{\hat{\mathsf{B}}\mathsf{B}\,m}/\mathsf{K}_m + \mathsf{N}_m/\mathsf{K}_m] \ \delta_{mm'} \Bigg]^{\frac{1}{2}} \quad (2.20) \\ &= \Bigg[\sum_{m} |\mathsf{E}_{\hat{\mathsf{B}}\mathsf{B}\,m}/\mathsf{K}_m + \mathsf{N}_m/\mathsf{K}_m|^2 \Bigg]^{\frac{1}{2}} \quad (2.21) \end{aligned}$$

where in (2.20), $\delta_{mm'} = \int_{-a}^{a} \exp[\imath\pi(m - m')x] \, \mathrm{d}x$ such that

$$\delta_{mm'} = \begin{cases} 1 & \text{if } m = m' \\ 0 & \text{if } m \neq m' \end{cases} . \quad (2.22)$$

The result in (2.21) is arrived at by summing over m. Using the same steps that were used to arrive at (2.13), the error-bound function for the Stokes method is given by

$$\| \hat{f}(x) - f(x) \| \leq \sum_{m} |\mathsf{E}_{\hat{\mathsf{B}}\mathsf{B}\,m}/\mathsf{K}_m| + \sum_{m} |\mathsf{N}_m/\mathsf{K}_m| \quad (2.23)$$

Equation (2.23) is essentially the same as the result derived in the matrix inversion method, (2.13), the only difference being that in place of the eigenvalues of $\mathbf{R}$, the Fourier coefficients of the instrument profile have been used. The ill-conditioning of the Stokes (1948) method is a result of $|\mathsf{K}_m| \to 0$ as $|m|$ increases. As in (2.13), there are two terms in (2.23) that will contribute to the ill-conditioning of the solution specimen profile. The first is the Fourier coefficients in the estimation error of the background level and the second is the noise distribution.

The results from this section demonstrate that when carrying out an unconstrained inversion, either in a vector space or in a Fourier space, the solution specimen profile will be ill-conditioned due to the amplification of the noise distribution and the misfitting error in the background level.

2.1.3 Iterative Deconvolution Method

The iterative method applies successive approximations to determine $|\hat{f}\rangle$. It was first introduced into x-ray diffraction by Ergun (1968). Taking the expression for the iterative method (see equ.(4) Ergun 1968), the integral term can be rewritten using a quadrature (see §2.1.1) and expressed in the bra-ket notation. We have,

$$|\hat{f}_{k+1}\rangle = \beta|\tilde{g}\rangle + (\mathbf{I} - \beta\,\mathbf{R})|\hat{f}_k\rangle \tag{2.24}$$

where we have introduced the scalar quantity, β, which controls the rate of convergence (Luenberger 1984) and $\mathbf{I}$ is an $[N \times N]$ identity matrix. In this form of the iterative method, the matrix $\mathbf{R}$ need not be normalized, but limits on the value of β can be defined via the eigenvalues of $\mathbf{R}$. These will be discussed in the latter part of this section. Expressing (2.24) in terms of the eigenvectors, we have

$$|\hat{f}_{k+1}\rangle = \sum_{i=1}^{N} \left[\beta\,\langle v_i|\tilde{g}\rangle + (1 - \beta\,\lambda_i)\langle v_i|\hat{f}_k\rangle\right]\,|v_i\rangle. \tag{2.25}$$

Starting with $|\hat{f}_0\rangle = |0\rangle$, and substituting it into (2.25), we arrive at a result for $|\hat{f}_1\rangle$,

$$\begin{aligned} |\hat{f}_1\rangle &= \sum_i \beta\,\langle v_i|\tilde{g}\rangle\,|v_i\rangle \\ &= \beta|\tilde{g}\rangle. \end{aligned} \tag{2.26}$$

Substituting this result back into (2.25), the expression for $|\hat{f}_2\rangle$ can be written as

$$\begin{aligned} |\hat{f}_2\rangle &= \sum_i \left[\beta\,\langle v_i|\tilde{g}\rangle + \beta(1 - \beta\lambda_i)\,\langle v_i|\tilde{g}\rangle\right]\,|v_i\rangle \\ &= \sum_i \beta\,\left[1 + (1 - \beta\lambda_i)\right]\,\langle v_i|\tilde{g}\rangle\,|v_i\rangle. \end{aligned} \tag{2.27}$$

By continuing to operate iteratively in this manner and taking note of the orthonormal properties of $|v_i\rangle$, we arrive at the following for the $k+1$ term:

$$|\hat{f}_{k+1}\rangle = \sum_{i=1}^{N} \beta \left[\sum_{j=0}^{k} (1 - \beta\lambda_i)^j\right]\,\langle v_i|\tilde{g}\rangle\,|v_i\rangle \tag{2.28}$$

$$= \sum_{i=0}^{N} \lambda_i^{-1}\left[1 - (1 - \beta\lambda_i)^{k+1}\right]\,\langle v_i|\tilde{g}\rangle\,|v_i\rangle \tag{2.29}$$

where the expression in the square bracket of (2.29) is a result of summing the geometric series in (2.28).

At this point it is important to consider the nature of the eigenvalues. There are three cases that arise. The first is where $\lambda_i > 0$ for all i. For convergence to be obtained, $|1 - \beta\lambda_i| < 1$, for all i, or for $0 < \beta < 2/\lambda^{(+)}_{max}$, where $\lambda^{(+)}_{max}$ is the largest positive eigenvalue. The second is where $\lambda_i < 0$ for all i. Again, for convergence $|1 - \beta\lambda_i| < 1$ for all i, which implies that β must be negative and satisfy $2/\lambda^{(-)}_{max} < \beta < 0$, where $\lambda^{(-)}_{max}$ is the largest negative eigenvalue. A schematic diagam of these cases is shown in Figure 2.1. The third case is where some λ_is are positive and others are negative. A β cannot be found such that $|1 - \beta\lambda_i| < 1$ for all i that satisfies both positive and negative eigenvalues, and in this case convergence of the iterative method cannot be achieved. Hence, a very simple method for determining whether convergence can be achieved is provided by determining the eigenvalues of the convolution matrix and inspecting their signs.

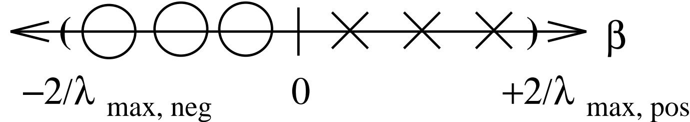


Figure 2.1: A schematic plot showing the acceptable range for β: (i) When all the λ's are positive (given by crosses), then $0 < \beta < 2/\lambda^{(+)}_{max}$, where $|\lambda^{(+)}_{max}| = \lambda_{max,pos}$; (ii) When all the λ's are negative (given by circles), then $2/\lambda^{(-)}_{max} < \beta < 0$, where $|\lambda^{(-)}_{max}| = \lambda_{max,neg}$.

Using (2.3) and (2.4) in (2.28), the difference between the solution and true specimen profiles is given by

$$\begin{aligned} |\hat{f}_k\rangle - |f\rangle &= \sum_i \left[(-1)(1-\beta\lambda_i)^{k+1}\langle v_i|f\rangle\right. \\ &+ \lambda_i^{-1}\left[1-(1-\beta\lambda_i)^{k+1}\right]\langle v_i|n\rangle \\ &+ \left.(-1)\lambda_i^{-1}\left[1-(1-\beta\lambda_i)^{k+1}\right]\langle v_i|e_{\hat{b}b}\rangle\right]|v_i\rangle. \end{aligned} \tag{2.30}$$

Taking the norm of (2.30) and using a similar approach to that carried out in §2.1.2 defines the error-bound function for the iterative method. This is given by

$$\begin{aligned} \| \hat{\mathbf{f}}_{k+1} - \mathbf{f} \| &\le \sum_{i=1}^{N} \left|(1-\beta\lambda_i)^{k+1}\right| |\langle v_i|f\rangle| \\ &+ \sum_{i=1}^{N} \left|\lambda_i^{-1}\right| \left|1-(1-\beta\lambda_i)^{k+1}\right| |\langle v_i|n\rangle| \end{aligned}$$

$$+ \sum_{i=1}^{N} \left|\lambda_i^{-1}\right| \left|1-(1-\beta\lambda_i)^{k+1}\right| \, |\langle v_i | e_{\hat{b}b}\rangle| \tag{2.31}$$

$$\leq E_f(k) + E_n(k) + E_{\hat{b}b}(k). \tag{2.32}$$

Unlike the results in §2.1.2, the error-bound functions for the iterative deconvolution methods are functions of their control parameters. The $E_f(k)$ term in (2.32) represents the restoration error as a function of k, in that it defines the misfitting error between the solution and true specimen profiles after k-iterations have been completed. For instance, if there were no noise or background terms in (2.32), $E_f(k)$ would be the only source of misfitting and the iterative deconvolution method would converge onto the true specimen profile as the number of iterations increased, i.e. as k increases, $E_f(k) \to 0$. The $E_n(k)$ term represents noise amplification as a function of k, since for increasing k and for small $|\lambda_i|$, $E_n(k)$ will dominate. The $E_{\hat{b}b}(k)$ is the misfitting error in the background level, with characteristics very similar to those of $E_n(k)$. The sum of these three terms defines a minimum in the error-bound function. The value of k which minimizes the error-bound function defines the optimum number of iterations, and the solution that corresponds to the optimum number of iterations is considered the optimum solution for a given level of noise.

2.1.4 Constrained Deconvolution Method

The constrained deconvolution method was first introduced into x-ray diffraction profile analysis by Louër et al. (1969). This method has its foundations in the work of Phillips (1962), Twomey (1963) and independently Tikhonov (1963). The starting point for this method is to write down the Lagrangian function $Q(\mathbf{f})$ which includes the regularization function, $S(\mathbf{f})$, and the misfitting function, $C(\mathbf{f})$. The function $S(\mathbf{f})$ can be thought of as a measure of the smoothness of $\mathbf{f}$, while $C(\mathbf{f})$ measures the misfitting of a given $\mathbf{f}$ in (2.4) (Twomey 1977). Following Twomey (1977), we define the regularization function (in the orthodox vector/matrix notation), $S(\mathbf{f})$, as

$$S(\mathbf{f}) = -\frac{1}{2}\left(\mathbf{f}-\mathbf{m}\right)^T \mathbf{H}\left(\mathbf{f}-\mathbf{m}\right) \tag{2.33}$$

where $\mathbf{m}$ is the *a priori* model representing our knowledge/ignorance concerning $\mathbf{f}$, and $\mathbf{H}$ is an $[N \times N]$ smoothing matrix. Here we set $\mathbf{H} = \mathbf{I}$, the identity matrix. The misfitting

function, $C(\mathbf{f})$, is written as

$$C(\mathbf{f}) = (\tilde{\mathbf{g}} - \mathbf{R}\mathbf{f})^T (\tilde{\mathbf{g}} - \mathbf{R}\mathbf{f}) . \tag{2.34}$$

Using the equations for $S(\mathbf{f})$ and $C(\mathbf{f})$, and writing the Lagrangian function in the bra-ket notation, we get

$$\begin{aligned} Q(\mathbf{f}) &= -\alpha\, S(\mathbf{f}) + \frac{1}{2} C(\mathbf{f}) && (2.35) \\ &= \frac{\alpha}{2} \left(\langle f| - \langle m| \right) H \left(|f\rangle - |m\rangle \right) \\ &+ \frac{1}{2} \left(\langle \tilde{g}| - \langle f|R^T \right) \left(|\tilde{g}\rangle - R|f\rangle \right) && (2.36) \end{aligned}$$

where α is the regularization parameter, such that $0 \leq \alpha < \infty$ and can be considered as a measure of the coupling between the regularization and misfitting function. By defining the smoothing function, (2.33) and (2.34), we are seeking to determine the specimen profile which minimizes (2.33) with respect to (2.34). This is embodied in the Lagrange function, $Q(\mathbf{f})$. In order to minimize this function, we want to determine the specimen profile, $\hat{\mathbf{f}}$, which results in $\nabla Q(\mathbf{f}) = 0$. This is equivalent to finding the vectors, $\nabla S(\mathbf{f})$ and $\nabla C(\mathbf{f})$ such that, $\nabla S(\mathbf{f}) \parallel \nabla C(\mathbf{f})$ i.e. $\nabla S(\mathbf{f})$ and $\nabla C(\mathbf{f})$ are parallel to each other as illustrated in Figure 2.2. This property of Lagrangian functions becomes a useful tool in developing a maximum entropy method in the next chapter (see Chapter 3).

In the case where we are ignorant of $|m\rangle$, we assume that it takes on a uniform distribution which is equal to the average value of the observed data. On the other hand, if we have some *a priori* evidence for $|m\rangle$, then this *a priori* information should be used. By minimizing (2.36) (see Twomey 1977), such that $\nabla Q(\mathbf{f})_{f \to \hat{f}} = 0$, we have

$$\left(\mathbf{R}^T \mathbf{R} + \alpha \mathbf{I} \right) |\hat{f}\rangle = \mathbf{R}^T |\tilde{g}\rangle + \alpha |m\rangle \tag{2.37}$$

where taking the inverse of $\left(\mathbf{R}^T \mathbf{R} + \alpha \mathbf{I} \right)$ on the left hand side of (2.37) results in the solution profile, $|\hat{f}\rangle$. It is interesting to compare the results in (2.4) and (2.37). When $\alpha = 0$, (2.37) reduces to (2.5). On the other hand, for a positive and nonzero α, the identity matrix in (2.37) is scaled. The consequence of this is that it offsets the small eigenvalues in the matrix $\mathbf{R}$ and tries to prevent the $\left(\mathbf{R}^T \mathbf{R} + \alpha \mathbf{I} \right)$ becoming ill-conditioned.

To express (2.37) in terms of the eigenvectors of $\mathbf{R}$, we apply similar reasoning to that employed in deriving equations (2.5 - 2.13). We start by multiplying both sides of (2.37) by

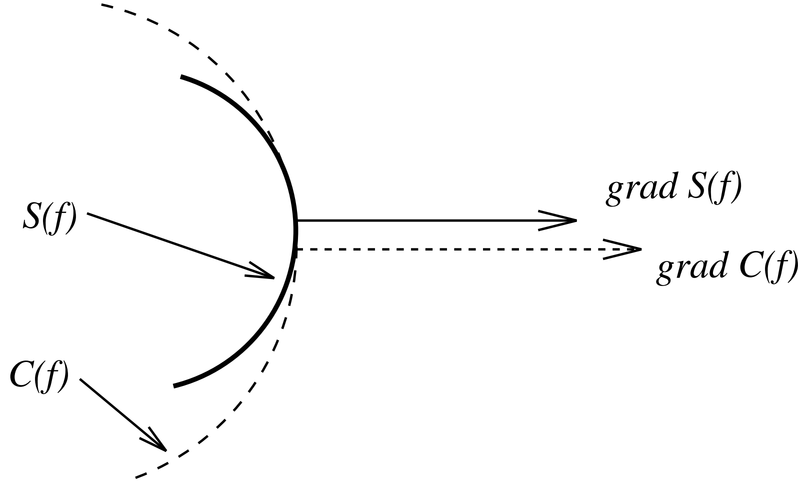


Figure 2.2: A schematic diagram showing the surface mapped out by the functions $S(\mathbf{f})$ and $C(\mathbf{f})$. $Q(\mathbf{f})$ is minimized for a given $\hat{\mathbf{f}}$ when the vectors ∇S and ∇C are parallel to each other. Here $\nabla S \equiv grad\, S$ and $\nabla C \equiv grad\, C$.

$\langle v_i|$ and arrive at,

$$\begin{aligned} \lambda_i^2 \langle v_i|\hat{f}\rangle + \alpha \langle v_i|\hat{f}\rangle &= \lambda_i \langle v_i|\tilde{g}\rangle + \alpha \langle v_i|m\rangle \\ (\alpha + \lambda_i^2) \langle v_i|\hat{f}\rangle &= \lambda_i \langle v_i|\tilde{g}\rangle + \alpha \langle v_i|m\rangle. \end{aligned} \tag{2.38}$$

By dividing through by $(\alpha + \lambda_i^2)$, multiplying both sides by $|v_i\rangle$ and summing over i, we arrive at

$$|\hat{f}\rangle = \sum_{i=1}^{N} \frac{\lambda_i \langle v_i|\tilde{g}\rangle + \alpha \langle v_i|m\rangle}{(\lambda_i^2 + \alpha)} |v_i\rangle. \tag{2.39}$$

After substituting (2.3) and (2.4) into (2.39) and after some manipulation, the error between the solution and true specimen profiles can be expressed

$$\begin{aligned} |\hat{f}\rangle - |f\rangle &= |e_{\hat{f}f}\rangle \\ &= \sum_{i=1}^{N} \frac{(-1)\alpha \langle v_i|f\rangle}{\lambda_i^2 + \alpha} |v_i\rangle + \sum_{i=1}^{N} \frac{\alpha \langle v_i|m\rangle}{\lambda_i^2 + \alpha} |v_i\rangle \\ &+ \sum_{i=1}^{N} \frac{(-1)\lambda_i \langle v_i|e_{\hat{b}b}\rangle}{\lambda_i^2 + \alpha} |v_i\rangle + \sum_{i=1}^{N} \frac{\lambda_i \langle v_i|n\rangle}{\lambda_i^2 + \alpha} |v_i\rangle. \end{aligned} \tag{2.40}$$

Using (2.40), an error-bound function can be written as

$$\| \hat{\mathbf{f}} - \mathbf{f} \| \leq \sum_{i=1}^{N} \left| \frac{\alpha \langle v_i|f\rangle}{\lambda_i^2 + \alpha} \right| + \sum_{i=1}^{N} \left| \frac{\alpha \langle v_i|m\rangle}{\lambda_i^2 + \alpha} \right| + \sum_{i=1}^{N} \left| \frac{\lambda_i \langle v_i|e_{\hat{b}b}\rangle}{\lambda_i^2 + \alpha} \right|$$

$$+ \sum_{i=1}^{N} \left| \frac{\lambda_i \langle v_i | n \rangle}{\lambda_i^2 + \alpha} \right| \tag{2.41}$$

$$\leq E_f(\alpha) + E_m(\alpha) + E_{\hat{b}b}(\alpha) + E_n(\alpha) \tag{2.42}$$

where $E_f(\alpha)$ is the restoration error as a function of α. In the case of no noise or background level, $E(\alpha)$ would be the only source of misfitting for the constrained method. The $E_m(\alpha)$ term represents the uncertainty in our choice of the *a priori* model. As the *a priori* model approaches f, the $E_m(\alpha)$ decreases and the solution profile improves. The $E_{\hat{b}b}(\alpha)$ term is the error resulting from the misfitting in the background level and $E_n(\alpha)$ is the noise amplification term. For small values of α the $E_n(\alpha)$ and $E_{\hat{b}b}(\alpha)$ terms dominate, while for large values of α, the $E_f(\alpha)$ and $E_m(\alpha)$ terms dominate. As was the case in the iterative method, a minimum in the error-bound function can be defined. The optimum α-value which minimizes the error-bound function and the corresponding solution from (2.37) is considered the optimum solution for a given noise level.

2.2 Computational Results and Discussion

2.2.1 Simulated Diffraction Profiles

A simulation of an instrument-broadened x-ray profile was carried out to demonstrate the properties of (2.37) and (2.42) and the corresponding deconvolution methods. The simulations consisted of generating an observed profile from the convolution of an instrument and specimen profile. The instrument profile was assumed to be a symmetrical Lorentzian (Cauchy) function, centred about $2\theta_B = 60\,°2\theta$. The full width at half maximum ($FWHM$) was determined using Cheary & Cline (1995) (see their equation (6) with $A = 3.52246 \times 10^{-4}$; $B = 5.55491 \times 10^{-4}$; $C = 5.25000 \times 10^{-4}$; $D = 6.23203 \times 10^{-5}$). The instrument profile had a $FWHM = 0.03\,°2\theta$, an amplitude of 100 counts, and was determined over the interval $(2\theta_B \pm 0.5)\,°2\theta$. The convolution matrix was determined in the manner discussed in §2.1.1. The elements of the convolution matrix are made up of the instrument profile, $k_{ij} = k(2\theta_i - 2\theta'_j)$ where $i, j = 1, 2, 3, \ldots, 100$, determined over the interval $(2\theta_B \pm 0.5)\,°2\theta$ and a quadrature which was chosen to be the trapezoidal rule (Burden & Faires 1993). The resulting convolution matrix was a $[100 \times 100]$ matrix. The eigenvalues of the convolution

matrix were calculated and found to be all positive, which implies that the convolution matrix is positive definite. The eigenvalue of the convolution matrix decreased monotonically, from the largest positive eigenvalue, $\lambda_{max}^{(+)} = 4.578$ to the smallest value, $\lambda_{min}^{(+)} = 0.0834$. The specimen profile was chosen to be a pseudo-Voigt function with a $FWHM = 0.05\,°2\theta$ and a mixing parameter, $\eta = 0.45$ (see Appendix A). The amplitude of the specimen profile was set to 1361 counts and was calculated over the interval (60.0 ± 0.5) $°2\theta$ with 100 data points. Poisson noise was added to the convolved result. The noise level corresponds to 1.6% of the peak maximum for the convolved result (peak max = 3938 counts). The background level was generated using Poisson noise, with a mean of 295 counts (7.5% of the peak maximum of the convolved profile). The normalized specimen and instrument profiles, as well as the simulated observed profile, are given in Figure 2.3.

In order to determine the individual terms in (2.32) and (2.42), the specimen profile, noise distribution and background, misfitting errors were calculated. The specimen profile used to calculate the observed profile was also used in (2.32) and (2.42). The noise distribution was determined from the difference between the observed profile with noise added to it and the noiseless observed profile. The misfitting error for the background level was determined from the difference between the true background level and the estimated background level. The estimated background level was deliberately under-estimated to demonstrate the effect of the misfitting error of the background level. It was taken to be a constant distribution at 270 counts (or 6.9% of the peak maximum). The solution profiles were also determined using the above methods and were compared with the true specimen profile using an R-factor defined as

$$R_f = \left[\frac{\sum_{i=1}^{N} \left(f_i - \hat{f}_i \right)^2}{\sum_{i=1}^{N} f_i^2} \right]^{\frac{1}{2}} \times 100\%$$

where f_i and $\hat{f}_i$ are the components from the $|f\rangle$ and $|\hat{f}\rangle$, respectively.

2.2.2 Unconstrained Inversion/Deconvolution Methods

The matrix inversion and Stokes (1948) methods were applied to the background-corrected observed profile. In the case of the matrix inversion method, the inverse of the convolution

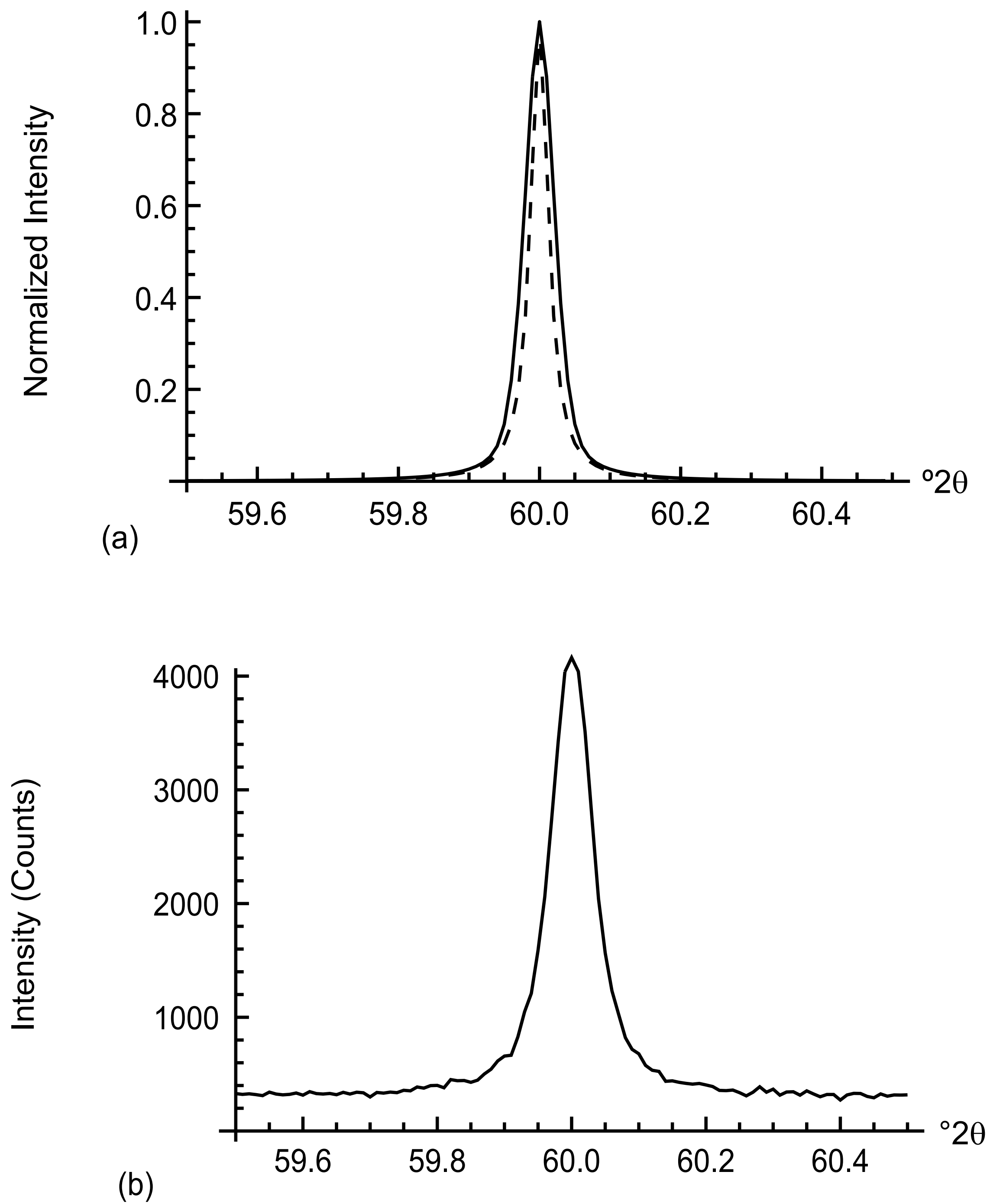


Figure 2.3: (a) Normalized specimen profile (solid lines) and instrument profile (dashed lines); (b) Observed profile with noise and background level added.

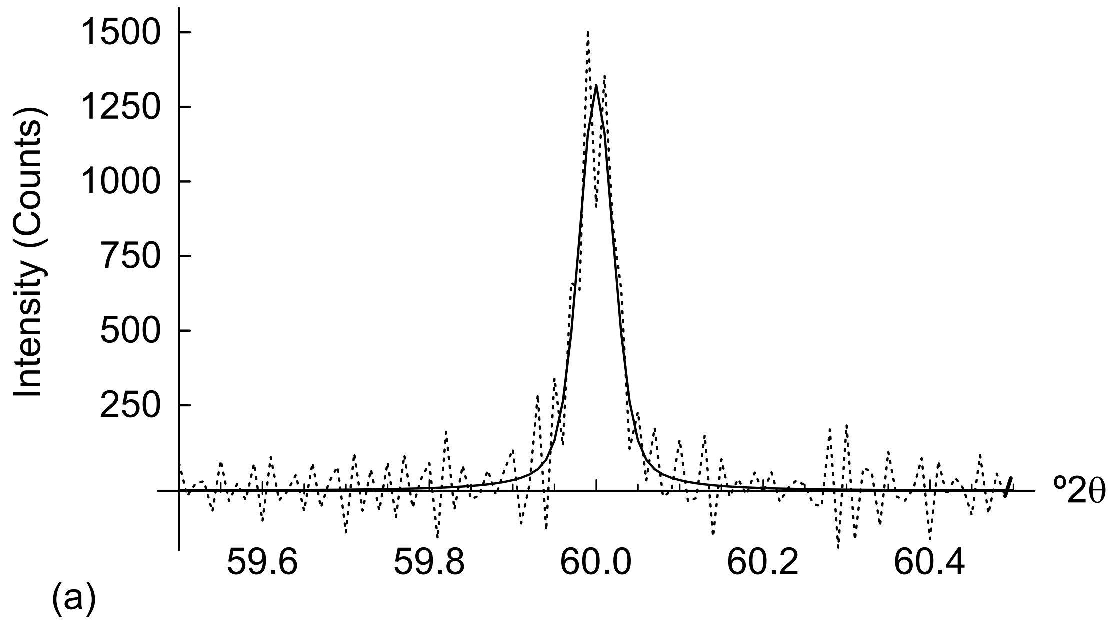
matrix was determined (see (2.5)) and was used to produce the solution profile. The inverse of small eigenvalues in the convolution matrix resulted in the amplification of the noise

and misfitting error in the background level (see (2.13)), hence the ill-conditioning of the solution profile, typically illustrated in Figure 2.4(a). The positivity of the solution profile in Figure 2.4(a) is not preserved. Even if the negative amplitudes were set to zero, the result would still be ill-conditioned, compromising any subsequent calculations. The ill-conditioning is reflected in the large value of $R_f = 43.7\%$.

In the case of the Stokes (1948) method, the Fourier coefficients of the corrected observed and instrument profiles were determined using a fast Fourier transform algorithm. The Fourier coefficients of the instrument profile ranged from $|\mathsf{K}_{max}| = 4.578$ to $|\mathsf{K}_{min}| = 0.084$. In order to determine the Fourier coeffients of the solution profile, the inverse of the instrument profile was taken and multiplied by the coefficients of the corrected observed profile (see (2.16)). The inverse Fourier transform was applied, resulting in the solution profile (see Figure 2.4(b)). In a similar manner as the inverse matrix method, amplification of the noise and misfitting error in the background estimation has occurred due to the small $|\mathsf{K}_m|$ values given in (2.23). The solution profile is corrupted with spurious oscillations. The ill-conditioning of the result in Figure 2.4(b) is also reflected in the large value of $R_f = 43.5\%$. This is almost identical to the inverse matrix result.

2.2.3 Iterative Method

The error-bound function for the iterative method was calculated over the range of 1 to 500 iterations. The eigenvalues of the convolution matrix are positive definite, so convergence is expected. A boundary for the suitable values of β was established as discussed in §2.1.3. The largest positive eigenvalue was 4.511, which implies $0 < \beta < 0.443$. For these calculations $\beta = 0.01$. The individual terms for this function were also determined (see Figure 2.5). It is seen that for a small number of iterations, the $E_f(k)$ term is dominant and decreases as the iterations increase. Conversely, the $E_n(k)$ and $E_{\hat{b}b}(k)$ terms, which have very similar behaviour, become dominant and increase as the number of iterations increases. These terms define the amplification of the noise and background misfitting error as a function of the number of iterations. A minimum occurs when the $E_f(k)$, $E_n(k)$ and $E_{\hat{b}b}(k)$ terms are combined in a sum (see equation (2.32)). This minimum corresponds to the optimum number of iterations; it can be considered as a compromise between the restoration error



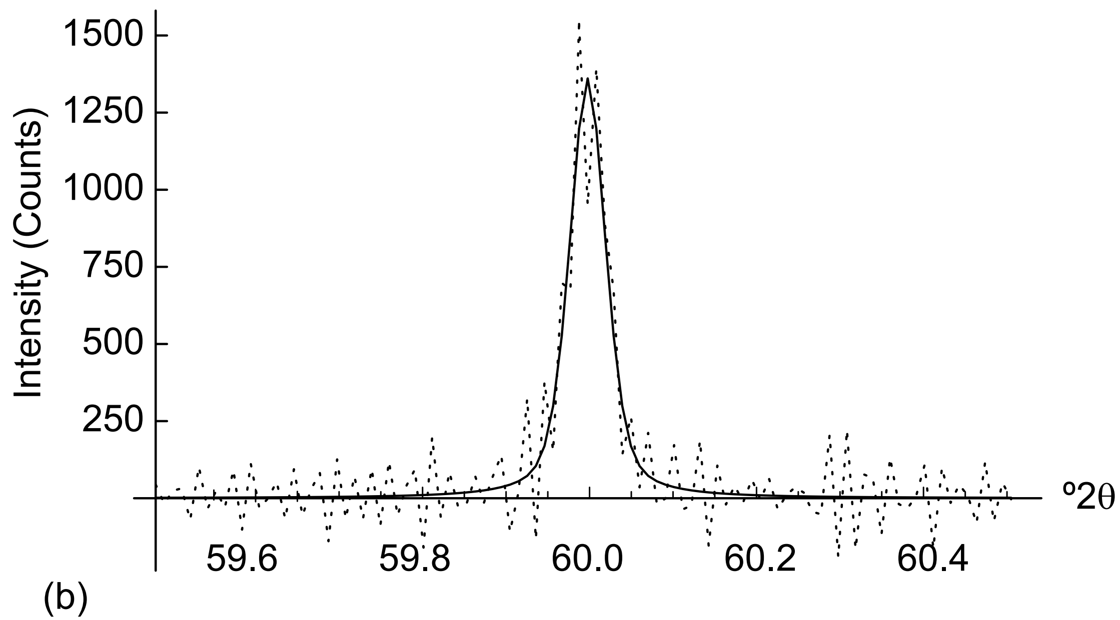


Figure 2.4: Plots of the solution from the unconstrained inversion/deconvolutions methods, given by (2.5) & (2.16), rspectivitely: (a) Showing the specimen profile (solid line) and the solution profile (dashed line), $R_f = 43.7\%$, for the matrix inversion method; (b) Showing the specimen profile (solid line) and solution profile (dashed line), $R_f = 43.5\%$, for the Stokes method.

and the noise/background error terms. The results in Figure 2.5 indicate that if the iterative method were halted before the optimum number of iterations was reached, then the major source of error would be due to the restoration term. On the other hand, if the method were halted after the optimum number of iterations, then the major source of error would be due to the combined influences of the background and noise terms.

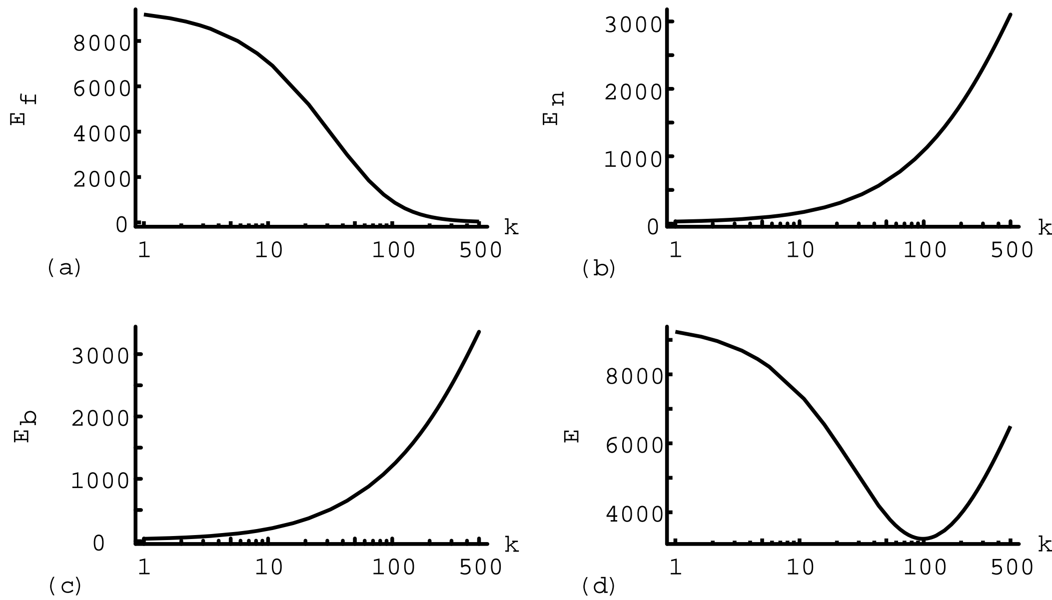


Figure 2.5: Plots for the terms in (2.32): (a) Showing the misfitting term in (2.32), $E_f(k)$ vs k ; (b) Showing the noise amplification term in (2.32), $E_n(k)$ vs k, for a noise level at 1.6% ; (c) Showing the background error term in (2.32), $E_{\hat{b}b}(k)$ vs k, for a background level 7.5% of the peak maximum; (d) Showing the sum of terms on the right hand side of (2.32).

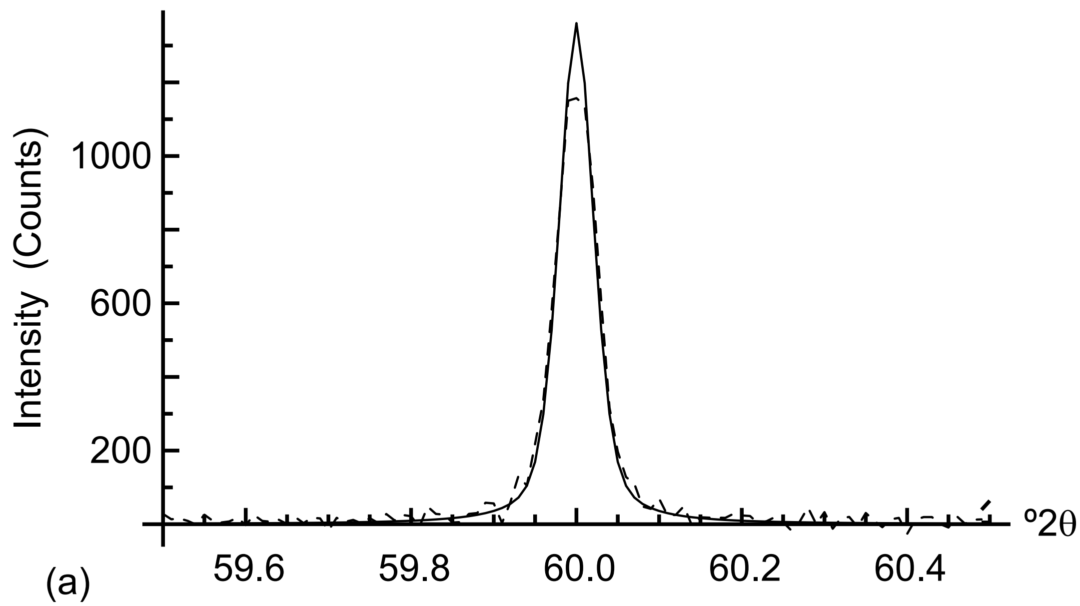
A set of solution profiles for $k = 50$, 100 and 350 iterations was determined (see Figure 2.6). The solution profile for $k = 100$ iterations corresponds to the optimum solution; its $R_f(= 11.5\%)$ was the lowest value for the set of profiles. The solution profile for $k = 50$ iterations demonstrates that at this stage the major source of uncertainty is due to restoration error. This is evident in the under-fitting of the solution profile when compared with the true specimen profile. In addition, it is noticeable that noise amplification has been suppressed. In the case of the solution profile for $k = 350$ iterations, the major source of error and misfitting is due to noise amplification, which results in the negative oscillations about the baseline. These oscillations are unphysical artefacts and will affect any subsequent

calculations such as the crystallite size or microstrain. The case for $k = 100$ iterations can be considered as a compromise between the components in (2.32) (see Figure 2.5). That is, there is less misfitting compared with the solution profile for $k = 50$ iterations and less noise/background amplification compared with the solution profile for $k = 350$ iterations.

2.2.4 Constrained Deconvolution Method

The error-bound function for the constrained deconvolution method (2.42) was calculated over the range of $\alpha \in [10^{-4}, 10]$ (see Figure 2.7). The *a priori* model for the data was set equal to the average of the observed data. In Figure 2.7(a), both the $E_f(\alpha)$ and $E_m(\alpha)$ terms are given as functions of α. These terms increase monotonically with increasing α. The poor choice of the *a priori* model as the average of the observed data is reflected in the large magnitude in $E_m(\alpha)$. On the other hand, we notice that the $E_n(\alpha)$ and $E_{\hat{b}b}(\alpha)$ terms given in Figure 2.7(b & c) decrease with increasing α. For small α, the $E_n(\alpha)$ and $E_{\hat{b}b}(\alpha)$ terms are large, which signals the ill-conditioning of the solution profile. The sum of these terms defines a minimum, which is given in Figure 2.7(d). As was described in the case of the iterative method, this represents an optimum value for the regularization parameter, α. The results shown in Figure 2.7 suggest that for large α-values, the restoration and *a priori* model errors are dominant, while for small α-values, the noise and background misfitting errors are dominant. This is demonstrated in Figure 2.8.

The optimum solution is given in Figure 2.8(a), for $\alpha \approx 0.15$, resulting in the lowest $R_f =$ 19.7%. The solution has negative oscillations about the baseline of the profile, especially on the outer edges. The results for small and large α-values are shown in Figure 2.8(b). In the case $\alpha = 0.005$, there are considerable oscillations about the baseline and misfitting about the peak of the profile. This can be considered as a case where the coupling between $S(\mathbf{f})$ and $C(\mathbf{f})$ in (2.35) is "weak" and $S(f)$ as the smoothing function is ineffectual. In the case of $\alpha = 5.0$, the noise amplification has been suppressed; there is considerable misfitting of the profile about the sides; and its background has been raised. In this case the coupling between $S(\mathbf{f})$ and $C(\mathbf{f})$ is "strong". The $S(\mathbf{f})$ term is dominant, suppressing any noise amplification.



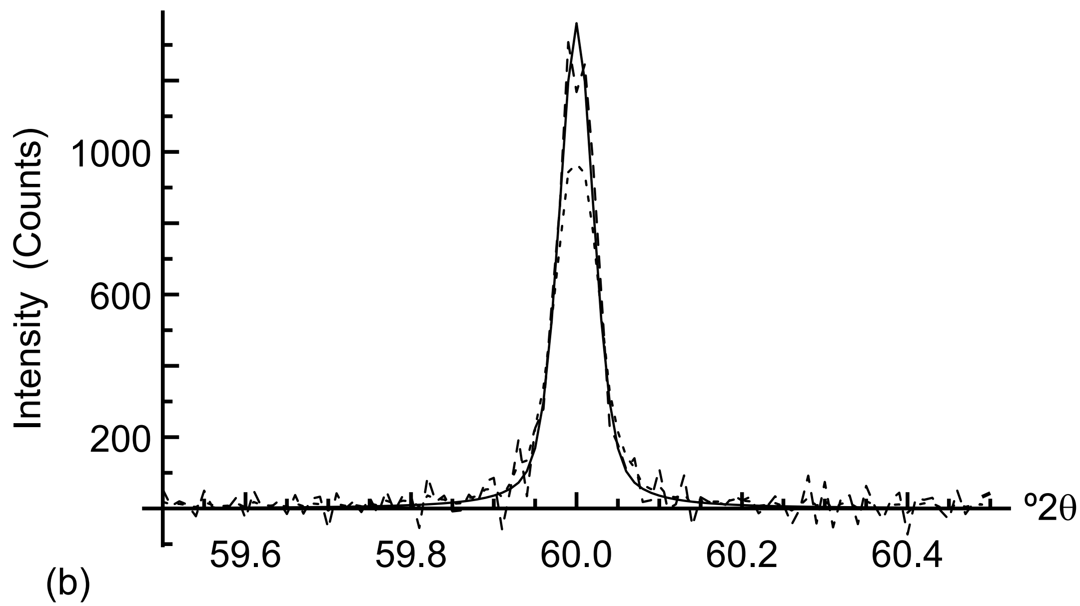


Figure 2.6: Plots of the solution profiles from the iterative method (2.24): (a) Showing the specimen profile (solid line) and the solution profile for the optimum number of iterations $k = 100$, $R_f = 11.5\%$; (b) Showing the specimen profile (solid line) and solution profile for $k = 50$ iterations, $R_f = 21.8\%$ (short dash line) and $k = 350$ iterations, $R_f = 17.8\%$ (long dash line).

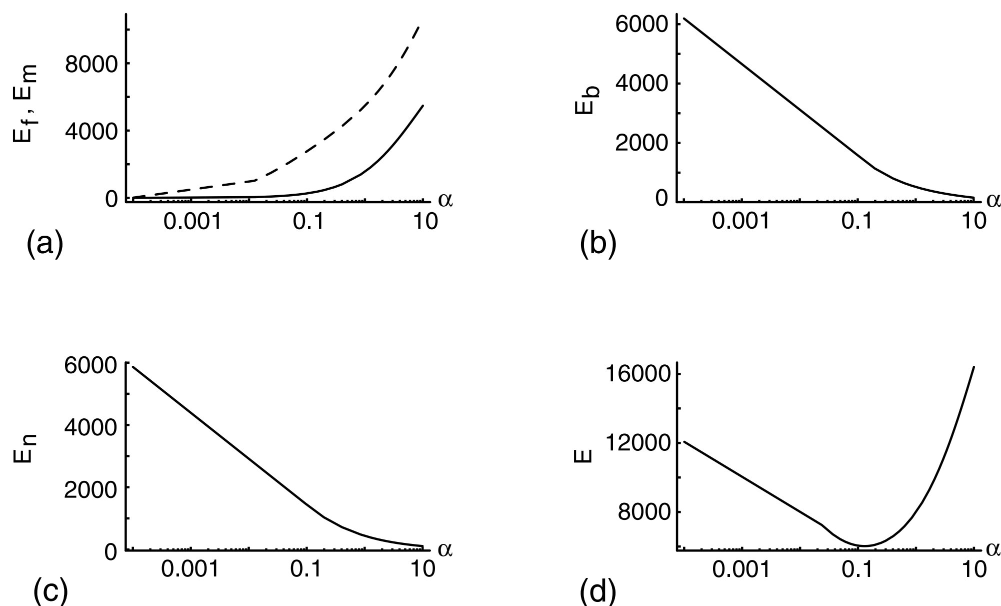
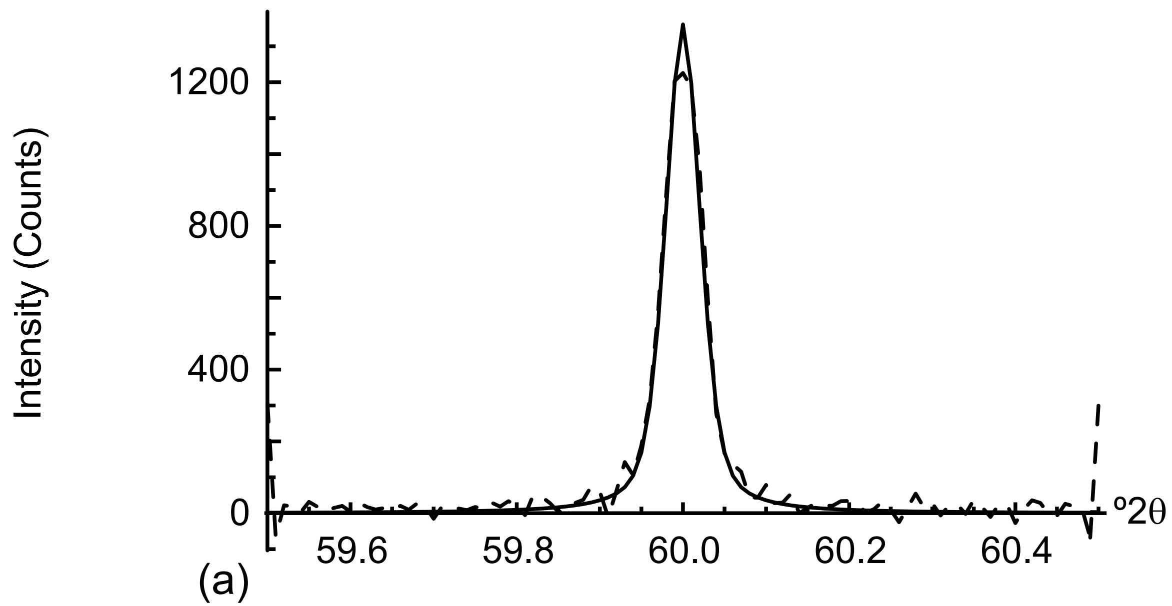


Figure 2.7: Plots of the terms in (2.42): (a) Showing the misfitting term in (2.42), $E_f(\alpha)$ vs α (solid line) and the misfitting error of the *a priori* model in (2.42), $E_m(\alpha)$ vs α (dashed line); (b) Showing the misfitting background error term in (2.42), $E_{\hat{b}b}(\alpha)$ vs α; (c) Showing the noise amplification term in (2.42), $E_n(\alpha)$ vs α, for a noise level of 1.6%; (d) Showing the sum of terms on the right hand side of (2.42). The minimum at $\alpha \approx 0.15$ corresponds to the optimum value α.

2.2.5 Implications for Experimental Data

In developing a deconvolution method that removes the instrument broadening from an observed profile in the presence of random noise, it is expected that the solution profile be physically consistent. That is, two basic requirements are fulfilled: that positivity be preserved; and the solution profile be free of spurious oscillations. The results given for the iterative and constrained deconvolution methods demonstrate that the optimum solution in the simulated case (i.e. the solution that is found when the errors have been minimized) does not fulfil the basic requirement given above. The results also highlight the difficulties in determining an optimum value of the control parameter without any *a priori* information of the noise and background level. The consequence of this is that the physical quantities determined from the solution profile will be compromised.

The implications that these results have on an experiment and subsequent calculations



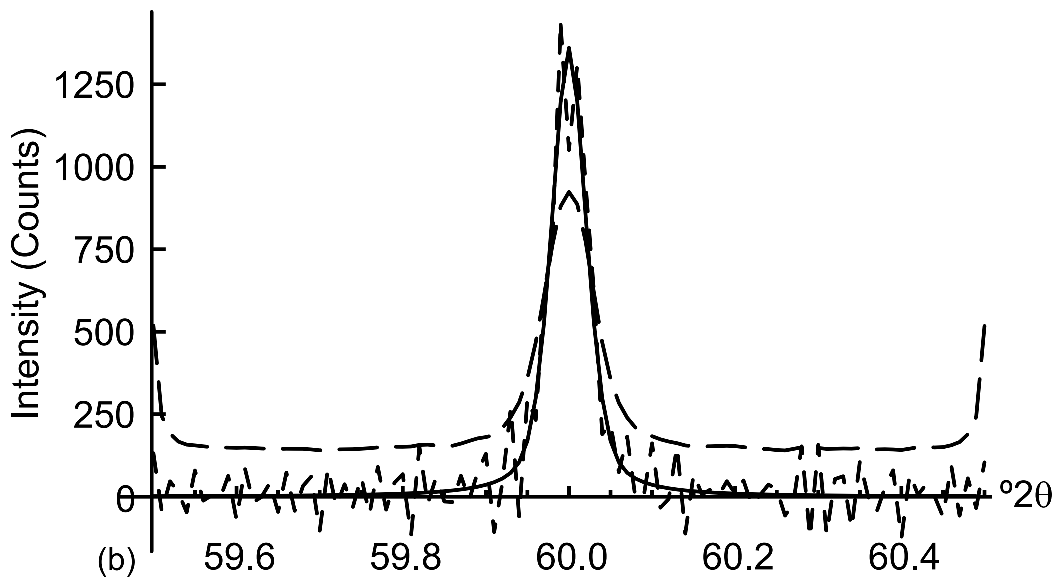


Figure 2.8: Plots of the solution profiles from the constrained method (2.37): (a) Showing the sample profile (solid line) and solution for $\alpha = 0.15$ and $R_f = 19.7\%$ (dash line); (b) Showing the specimen profile (solid line) and two other solutions, $\alpha = 0.005$, $R_f = 32.3\%$ (short dash line) and $\alpha = 5.0$, $R_f = 65.8\%$ (long dash line), respectively.

is that a reasonable understanding of the background level is required to minimize the misfitting errors. In order to reduce the ill-conditioning due to noise, a tolerable noise level

for the deconvolution method must be determined, which in turn places demands on the conditions for collecting the experimental data, such as an optimally adjusted diffractometer and adequate counting times.

2.3 Summary

The common deconvolution methods used in x-ray diffraction have been decomposed using an eigen-system analysis. A comparison of the deconvolution profile and the sample profile was achieved by defining the error-bound function. The unconstrained inversion/deconvolution methods do not have control parameters. However, the error-bound function (which in this case consists of single numbers), from a theoretical point of view, accounts for ill-conditioning in the solution profile. In the case of the iterative and constrained deconvolution methods, the error-bound function was defined in terms of the control parameter. This enabled the various components which define the observed profile and affect its deconvolution, such as the background level and noise distribution, to be studied and quantified.

A simulated instrument-broadened profile, composed of a Poisson noise distribution and a background level, was used to evaluate the error-bound function of each method. The solution profiles for the unconstrained/inversion methods contain spurious oscillations. The positivity for the solution profiles is not preserved. The results for the iterative and constrained methods from the simulated profiles showed that the optimum value of the control parameter can be considered as a compromise between the restoration error and the noise/background estimation error. Under these optimum conditions (which are generally not known *a priori*) the positivity was not preserved and spurious oscillations corrupted the solution profile. It was also shown that the under- or over-estimation of the control parameter resulted in one of two sources of misfitting and ill-conditioning becoming dominant. The results demonstrate that the various sources of misfitting and ill-conditioning can be quantified and related to the corresponding features in the solution profile.

Chapter 3

Development of a Maximum Entropy Method for X-Ray Diffraction Profile Analysis

The maximum entropy (MaxEnt) method has proven to be very successful for data analysis. In many areas of scientific research it has been adopted as the standard approach; in particular, Bayesian and MaxEnt methods have been applied in deconvolving neutron profiles, analysing neutron reflectivity data and determining structure factors in powders (see Sivia 1990, Sivia et al. 1993, Sivia & David 1994, Geoghegan et al. 1996)[1]. The specific aim of this chapter is to present MaxEnt as an alterative method for x-ray diffraction profile analysis. The generality of the MaxEnt method allows the specimen profile, the column-length distribution and the strain distribution to be determined from the observed profile. The particular problem of determining the specimen profile is considered here.

X-ray diffraction profiles are compared with other common distributions that are found in physics. It is shown that they share certain features, in that they are positive and additive distributions. Based on these features a suitable function is determined, the entropy function, which preserves these characteristics and the maximum entropy principle can then be established (see §3.1). Bayesian probability theory is then used to transform the principle

[1]The review by Gilmore (1996) is an interesting and practical review of MaxEnt methods used in crystallography.

into a working method.

To determine the specimen profile (see §3.2), the MaxEnt method incorporates the entropy function, observed profile, instrument profile and counting statistics. The use of the Skilling & Bryan (1984) maximum entropy algorithm enables these elements to be considered in determining the specimen profile. Not only does the MaxEnt method allow the specimen profile to be determined, but it also enables the column-length distribution and strain distribution to be found. In other words, after the specimen profile is known the MaxEnt method can be applied again to calculate the underlying distributions.

The MaxEnt method is shown to be a fully quantitative method in that the uncertainties in the specimen profile can be quantified. This has important significance since uncertainties in subsequent calculations and physical quantities can also be determined. The difficulties and limitations in the Skilling & Bryan (1984) algorithm are outlined (see §3.3).

3.1 The Maximum Entropy Principle

3.1.1 Positive and additive distributions

Many distributions that appear in physics are positive and additive distributions. Some examples are astronomical images (optical and radio), electron density distributions from crystallography, and size distributions from small-angle x-ray scattering. As expected, by positive we mean that the distributions cannot take negative values; in the case of astronomical (optical) images, for example, it is a spatial distribution of light intensity over the surface of a film or detector that is being measured. By additive, we imply that the total sum of the distribution has a physical meaning. Again using the example of optical astronomical images, the total light flux received by the detector may have implications as to the nature of the light source. Another important example of positive and additive distributions in physics relates to *probability distributions*. It is meaningless to have negative probabilities, and the additivity condition ensures that the distribution can be normalized. In the case of statistical mechanics, the normalization term of the probability distribution, or partition function, can be used to determine all the necessary macroscopic physical quantities.

The same observations can be made of x-ray diffraction profiles. The positivity of the

profile is a result of the intensity of the electric field being scattered by a set of atoms in a particular plane of the atomic lattice. This is represented by the square of the magnitude of the structure factor, $|F|^2$. If the intensity is recorded as a function of angular position, 2θ, the positivity of an x-ray diffraction profile is given as

$$I(2\theta) \geq 0, \quad \forall\, 2\theta. \tag{3.1}$$

The additivity condition of diffraction profiles can be seen in the physical meaning of the total integrated intensity. The structure factor is the total contribution of the scattering from individual atoms for a particular crystallographic direction, $[hkl]$. The integrated intensity of a diffraction profile, hkl, for a random powder in terms of the structure factor (see Langford & Louër 1996, Warren 1969) is

$$I_{hkl} = I_0 K \lambda^3 V^{-2} v \mathfrak{m}_{hkl} P L |F_{hkl}|^2 \tag{3.2}$$

where I_0 is the intensity of the primary beam; K depends on x-rays being collected; V is the volume of the unit cell; v is the diffracting volume of the specimen; $\mathfrak{m}_{hkl}$ is the multiplicity of the hkl profile; P is the polarization factor; L is the Lorentz factor; and $|F_{hkl}|^2$ is the structure factor. The important feature in (3.2) is that integrated intensity is related the structure factors $|F_{hkl}|^2$.

These two observations are important in formulating a "method" that can determine both the specimen profile from the observed x-ray diffraction profile and underlying distributions such as the column-length and strain distributions. That is, we expect our method to extract this information from the observed data and produce results which preserve the positivity and additivity of the profile or distribution. These features could be included in some of the deconvolution methods discussed in Chapter 2. For example, in the constrained deconvolution method, (see §2.1.4), an additional Lagrangian parameter could be added to preserve the positivity of the solution profile, $\hat{\mathbf{f}}$. However, to include this additional parameter adds an extra degree of freedom to the problem, in that an extra unknown Lagrangian multiplier must be determined. In addition, from an aesthetic point of view it is somewhat "messy" to arbitrarily add an extra parameter and it implies that the inverse method is "incomplete". Rather, it should be possible to incorporate the properties of positivity and additivity without making additional assumptions about, say, the functional form of the distribution or without having to arbitrarily include parameters in the deconvolution/inverse

method. These conditions ensure that the specimen profile or column-length distribution determined from the observed profile can be interpreted in general terms.

3.1.2 The maximum entropy principle

Using positivity and additivity as conditions for the consistency of a method for determining the specimen profile or column-length distribution, we are presented with a new problem of *assigning values* to the profile or distributions. Again, this is a problem that is often seen in physics. For example, in statistical mechanics, the problem of a system (i.e. canonical ensemble) with a fixed number of particles and a specified average energy is common. How do we assign a probability distribution to all the possible microstates and determine their corresponding energies? Similarly, in x-ray powder diffraction the problem of determining the probability $p(I)\,\mathrm{d}I$ that the intensity I, lies between I and $I + \mathrm{d}I$ is often encountered.

The problem of assigning values to distributions has been embodied in a host of arguments (see Jaynes 1957, Jaynes 1968, Shore & Johnson 1980, Jaynes 1982, Jaynes 1983, Johnson & Shore 1983, Tikochinsky et al. 1984, Skilling & Gull 1985, Skilling 1989*b*, Skilling 1990). The single most important conclusion that can be drawn from these arguments is that the most probable distribution best represents our knowledge of the "system" and that this distribution will have *maximally uncommitted information* or the least assumptions (Wu 1997). To illustrate the discussion, the so called "monkey argument" (see Skilling 1984, Skilling & Gull 1985, Gilmore 1996, Sivia 1996) is presented.

The monkey argument

The following discussion is based on Sivia (1996): suppose we want to assign probabilities to M distinct outcomes, represented by x_i for $i = 1, 2, 3, \ldots, M$, given some background information, $\mathcal{I}$. $\mathcal{I}$ may include properties about x_i or assumptions about the underlying mechanics (Loredo 1990). Suppose we were to employ a team of monkeys to throw a large number, N, of red balls, into M identical boxes, where $N >> M$. The monkeys represent a random and unbiased system and the boxes represent the possibilities, x_i. After they have finished, we examine the number of balls in the boxes. There will be a distribution of the number of balls in the boxes, $\{n_i\}$ or $\mathbf{p}$. There may be n_1 balls in the first box, n_2 red

balls in the second box and similarly we may find n_M balls in the M-th box. We let the monkeys repeat the experiment many times, each time recording the resulting distribution of red balls. The total number of balls will be conserved

$$N = \sum_{i=1}^{M} n_i. \tag{3.3}$$

The probability of n_i red balls being in the ith box is given by

$$p_i = \frac{n_i}{N}, \quad \forall i = 1,\, 2,\, 3,\, \ldots,\, M \tag{3.4}$$

such that $\mathbf{p} = \{p_i;\, i = 1,\, 2,\, 3,\, \ldots,\, M\}$. After a large number of repetitions, we notice that some distributions occur more often than others which are consistent with $\mathcal{I}$. We can define the number of ways a distribution $\mathbf{p}$ or $\{n_i\}$ occurs by determining its frequency,

$$\Omega(\{n_i\}) = \frac{G(\{n_i\})}{M^N} \tag{3.5}$$

where $G(\{n_i\})$ is the number of times the distribution $\{n_i\}$ has appeared and M^N is the number of ways N red balls can be arranged into M boxes. To evaluate the $G(\{n_i\})$, we must first determine the number of ways, say, n_1 red balls can be drawn from N balls, given by the combination, ${}^{N}C_{n_1} = \frac{N!}{n_1!\,(N-n_1)!}$. For n_2, this would be ${}^{N-n_1}C_{n_2} = \frac{(N-n_1)!}{n_2!\,(N-n_1-n_2)!}$, and so on for ${}^{n_M}C_{n_M}$. Using these terms, $G(\{n_i\})$ is given by the product of the combinations,

$$\begin{aligned} G(\{n_i\}) &= {}^{N}C_{n_1} \times {}^{N-n_1}C_{n_2} \times \cdots \times {}^{n_M}C_{n_M} \\ &= \frac{N!}{n_1!\, n_2! \ldots n_M!}. \end{aligned} \tag{3.6}$$

Using this result in (3.5) and taking the logarithm, we have,

$$\begin{aligned} \ln \Omega &= -N \ln(M) + \ln(N!) - \ln(n_1!\, n_2! \ldots n_m!) \\ &= -N \ln(M) + \ln(N!) - \sum_{i=1}^{M} \ln n_i!. \end{aligned} \tag{3.7}$$

For large n, we can apply the Stirling formula,

$$\ln(n!) \approx n \ln(n) - n \tag{3.8}$$

and remembering that $n_i = N\, p_i$, after some manipulation we arrive at the result,

$$\ln \Omega \approx -N \ln(M) - N \sum_{i=1}^{M} p_i \ln p_i \tag{3.9}$$

where $\sum_{i=1}^{M} p_i = 1$, and M and N are constants. Given a distribution, $\mathbf{p}$ or $\{n_i\}$, its frequency of occurrence can be determined from (3.9). The distribution which best represents the system of red balls, that is consistent with $\mathcal{I}$ and constrained by N, will have the greatest frequency. In other words, this distribution will be the most probable distribution. For this condition to be satisfied, $\ln \Omega$ must be a maximum, which implies $-\sum_{i=1}^{M} p_i \ln p_i$ must also be a maximum. We define $-\sum_{i=1}^{M} p_i \ln p_i$ as the entropy function, $S(\mathbf{p})$, of the distribution $\mathbf{p}$,

$$S(\mathbf{p}) = -\sum_{i=1}^{M} p_i \ln p_i. \tag{3.10}$$

where $\sum_{i=1}^{M} p_i = 1$.

The significance of the above discussion is that our best choice of $\mathbf{p}$ was made with the least added information. That is, we did not need to include any additional assumptions about the nature of the distribution, such as its functional form. Also, this distribution was determined with the information supplied. Hence, we can now assign a probability distribution to $\{x_i\}$, by determining the distribution $\mathbf{p}$ which has a maximum entropy. As stated earlier, this distribution will be the maximally noncommittal distribution or the distribution with the least assumptions (Wu 1997); this defines the MaxEnt principle.

We can generalise the monkey problem further by considering the case where the boxes have different sizes (see Sivia 1996). In the above case, all the boxes were assumed to have been the same size, which implies that it is equally likely that a ball will be found in any one box. The *a priori* model is defined as a uniform probability of $1/M$ for all i. In the case where the boxes do not have the same size, the probability distribution, $\mathbf{m} = \{m_i; i = 1, 2, 3, \ldots, M\}$, will be non-uniform. As usual the distribution must sum to unity (the additivity condition),

$$\sum_i m_i = 1. \tag{3.11}$$

The frequency of the $\{n_i\}$, (3.5), is "modulated" by the probability of a ball being found in a particular box, $\mathbf{m}$. This results in a multinomial distribution (Sivia 1996). Following the same steps that produced (3.9), we can show that the best distribution will be when $-\sum_{i=1}^{M} p_i \ln \frac{p_i}{m_i}$ is a maximum. This is similar to (3.10), with the exception that m_i appears

as the denominator in the logarithm. This defines a general entropy function, $S(\mathbf{p})$,

$$S(\mathbf{p}) = -\sum_{i=1}^{M} p_i \ln \frac{p_i}{m_i}. \tag{3.12}$$

In the case where all the boxes have the same size (3.12) reduces (3.10). The difference between the two forms of entropy is that (3.12) allows the entropy function to be transformed. We define $\mathbf{m}$ as a Lebesgue measure, first introduced by Jaynes (1968). It ensures that (3.12) remains invariant under transformation.

The distribution $\mathbf{m}$ also provides us with *a priori* information about the possible distribution of the balls. Suppose for a moment we had not carried out the experiment with the monkeys. By the maximum entropy principle, the distribution which maximized the entropy function would be $\mathbf{m}$ and this would represent our best understanding of the distribution of balls. However, after the experiment the information about the possible distribution of balls, $\mathbf{m}$, is imparted onto the distribution that maximises (3.12). Hence, (3.12) quantifies the transfer of information from $\mathbf{m}$ to $\mathbf{p}$ (Gilmore 1996).

From this we define $\mathbf{m}$ as the *a priori* model. Being able to define and include the *a priori* model in a maximum entropy method becomes very important for determining the column-length distribution and strain distributions from observed diffraction profiles. In the later chapters, it will be demonstrated that the *a priori* model influences the uncertainty region of the solution.

The uniqueness of the entropy function

The monkey argument demonstrates in a particular case that (3.12) can be used to assign values to a distribution. This argument is also applicable to cases where an ensemble of identical systems can be prepared such as in statistical mechanics. However, it has little bearing on the physical reality of recording and analysing an x-ray diffraction profiles. That is, only one pattern is recorded by the diffractometer and we often do not have the liberty of generating a large number of candidate patterns. In recording an x-ray profile and analysing it, we have to work with the available data and information. Accepting that this is the general situation, the following question arises: how can we be sure that there are not other functions, such as $-\sum_i \ln p_i$ or $\sum_i \sqrt{p_i}$, which preserve the positivity and additivity of $\mathbf{p}$ and produce the same or a better result than applying (3.12)?

This question has been addressed in various ways (see Shore & Johnson 1980, Johnson & Shore 1983, Tikochinsky et al. 1984, Gull & Skilling 1984, Skilling 1989*b*, Skilling 1990, Goambo & Gassiat 1997). As was the case in the problem of assigning values to a distribution, the central conclusion that can be made is that (3.12) is consistent with the principles of statistical inference. If another function is used, it may produce inconsistent results, unless it has the same maximum (Shore & Johnson 1980, Johnson & Shore 1983). In the case of reproducible experiments it has been shown that the maximum entropy principle was the only consistent approach to inferring discrete probabilities (Tikochinsky et al. 1984).

In the case of powder x-ray diffraction, it is interesting to note that the problem of determining the probability, $p(I)\,\mathrm{d}I$, for the intensity, I, first presented by Wilson (1949), can be determined independently by using the entropy function. This was demonstrated by Sivia & David (1994) and indicates that the maximum entropy principle can independently produce consistent and reliable solutions.

3.2 The Maximum Entropy (MaxEnt) Method

In the monkey argument we addressed the problem of assigning a probability distribution to a set of outcomes, $\{x_i\}$. We found that the distribution which was consistent with the background information, $\mathcal{I}$, and which represented our best knowledge of the system of red balls, was the distribution that maximised the entropy function (3.12). We can argue that this distribution will have the least added information.

Using (3.12), we want to formulate a "method" that allows us to determine the profile or distribution (i.e. column-length or strain distribution) which has a maximum entropy subject to experimental data. This can be done by appealing to Bayesian probability theory. We apply this theory because it incorporates the process of inductive reasoning. It also provides us with a "grammar" (Loredo 1990) or set of rules to incorporate the entropy function and experimental data. It does this by taking into account the hypothesis (the *a priori* information) with the data (likelihood information) (Loredo 1990, Sivia 1996). The hypothesis and likelihood information are characterised by the *a priori* and likelihood probability distributions, respectively. These distributions are mapped into the *posterior probability* distribution, which determines the plausibility of the hypothesis in the presence

of the experimental data. In applying Bayesian theory, probability is defined as the measure of the plausibility of a proposition when there is incomplete information or evidence to establish with certainty its truth or falsehood (Loredo 1990, Sivia 1996). The measure of plausibility is represented by a mapping onto the real number interval [0, 1], where 0 defines the proposition as being false and 1 defines it as being true (Cox 1946). By introducing a process of inductive reasoning and defining probability in this manner, we can determine the plausibility of a specimen profile or column-length distribution in the presence of the observed profile (i.e. experimental data).

The following section will consider a MaxEnt method for x-ray profile analysis by addressing the problem of deconvolving the observed profile, given in (2.1) and (2.4) such that the specimen profile, $\mathbf{f}$ can be determined. The maximum entropy methods for determining the column-length and strain distributions from the specimen profile are discussed in Chapters 4 & 5, respectively.

The diffraction profiles are assumed to be single profiles from monochromatic radiation.

3.2.1 *A priori* probability distribution

The hypothesis is founded on the positivity and additivity properties of x-ray diffraction profiles and entropy function, (3.12), to infer the values of the specimen profile. This information constitutes our *a priori* information and can be expressed in terms of a probability distribution as

$$p(\mathbf{f}|\,\mathbf{m},\,\alpha,\,\mathcal{I}) = \frac{\exp\left[\alpha\, S(\mathbf{f})\right]}{Z_S} \tag{3.13}$$

where α is a positive constant, such that $\alpha \in [0,\,\infty)$; $S(\mathbf{f})$ is the entropy function, (3.12). Z_S is the normalisation term such that

$$\int p(\mathbf{f}\,|\mathbf{m},\,\alpha,\,\mathcal{I})\,\mathcal{D}\mathbf{f} = 1 \tag{3.14}$$

where the measure $\mathcal{D}\mathbf{f}$ accounts for the volume element in $\mathbf{f}$-space (see Skilling 1989*a*).

The term on the left-hand side of (3.13) is interpreted as the probability of $\mathbf{f}$, conditional (i.e. "|") on the *a priori* model, $\mathbf{m}$, some positive value of α and the background information $\mathcal{I}$. The *a priori* model, $\mathbf{m}$, incorporates our knowledge (or lack of it) about the specimen profile. This information may include theoretical knowledge of the specimen profile or include

information from a Rietveld refinement. In the absence of this information we must assume a uniform model equal to the average number of counts of the observed profile, in which case (3.12) reduces to (3.10). The background information, $\mathcal{I}$, incorporates the knowledge that the profiles are non-overlapping and consist of only monochromatic radiation, say.

3.2.2 Likelihood probability distribution

The hypothesis needs to be tested in the presence of the experimental data (i.e. the observed profile). The likelihood probability distribution determines the plausibility of $\tilde{\mathbf{g}}$ for a particular or *trial* specimen profile, $\mathbf{f}$, which is embodied in the experimental model, such as the convolution equation described by (2.1) and (2.4).

In order to determine this probability distribution, the statistics of the experiment must be considered. The process of recording an x-ray diffraction pattern or profile is simply a counting experiment in which the underlying statistics are governed by a Poisson process. In the case of large counts ($\gtrsim 10$) it can be shown that the Poisson process can be approximated by a Gaussian probability distribution. Following a similar discussion to Sivia (1996), suppose we consider the probability of obtaining the ith datum, $\tilde{g}_i$, given as

$$p(\tilde{g}_i|\,\hat{g}_i,\,\mathcal{I}) = \frac{\hat{g}_i^{\tilde{g}_i}\,\exp(-\hat{g}_i)}{\tilde{g}_i\,!} \tag{3.15}$$

where $\tilde{g}_i$ is conditional on the noiseless data, $\hat{g}_i$ (which has been corrected for the background level) and on $\mathcal{I}$. For the corrected observed profile there are M data points and the corresponding probability distribution becomes

$$p(\tilde{\mathbf{g}}|\,\hat{\mathbf{g}},\,\mathcal{I}) = \prod_{i=1}^{M} \frac{\hat{g}_i^{\tilde{g}_i}\,\exp(-\hat{g}_i)}{\tilde{g}_i\,!}. \tag{3.16}$$

where $\tilde{\mathbf{g}} = \{\tilde{g}_i;\, i = 1,\, 2,\, \ldots,\, M\}$, and similarly for $\hat{\mathbf{g}}$. Where the counts are very large, we can take the logarithm of (3.16),

$$\ln[p(\tilde{\mathbf{g}}|\,\hat{\mathbf{g}},\,\mathcal{I})] = \sum_i \tilde{g}_i \ln \hat{g}_i - \sum_i \hat{g}_i - \sum_i \tilde{g}_i \ln \tilde{g}_i + \sum_i \tilde{g}_i - \frac{1}{2}\sum_i \ln(2\pi\tilde{g}_i) \tag{3.17}$$

where the last three terms are arrived at by applying Stirling's (full) formula on $\ln(\prod_i \tilde{g}_i!)$ for large $\tilde{\mathbf{g}}$. Here we want to determine $\ln[p(\tilde{\mathbf{g}}|\,\hat{\mathbf{g}},\,\mathcal{I})]$ in the region about the noiseless data and

must take into account that $\tilde{g}_i$ will be corrupted by the uncertainty, n_i, such that $\tilde{g}_i = \hat{g}_i + n_i$, where $\langle n_i \rangle = 0$ and $\langle n_i^2 \rangle = \sigma_i^2 = \hat{g}_i \approx \tilde{g}_i$ for all i. Substituting this result for $\tilde{g}_i$ we arrive at,

$$\ln[p(\tilde{\mathbf{g}}|\,\hat{\mathbf{g}},\,\mathcal{I})] = -\sum_i \left(\hat{g}_i + n_i + \frac{1}{2}\right) \ln\left(1 + \frac{n_i}{\hat{g}_i}\right) - \frac{1}{2}\sum_i \ln(2\pi\hat{g}_i). \qquad (3.18)$$

The $\frac{1}{2}$ term in (3.18) can be dropped since $\hat{g}_i + n_i \gg \frac{1}{2}$ and the result can be further approximated by expanding the $\ln\left(1 + \frac{n_i}{\hat{g}_i}\right)$ term as a Taylor series to the second order (Sivia 1996). Substituting the terms of the series into (3.18) and after some manipulation we arrive at

$$\begin{aligned}\ln[p(\tilde{\mathbf{g}}|\,\hat{\mathbf{g}},\,\mathcal{I})] &\approx -\sum_i \frac{n_i^3}{\hat{g}_i^2} - \frac{1}{2}\sum_i \ln(2\pi\hat{g}_i) - \frac{1}{2}\sum_i \left(\tilde{g}_i - \hat{g}_i\right)^2/\hat{g}_i \\ &\approx -\frac{1}{2}\sum_i \ln(2\pi\hat{g}_i) - \frac{1}{2}\sum_i \left(\tilde{g}_i - \hat{g}_i\right)^2/\hat{g}_i \qquad (3.19)\end{aligned}$$

where the $\sum_i \frac{n_i^3}{\hat{g}_i^2}$ term can be dropped since n_i can take on both negative and positive values and will be small after the summation. In the case of large counts, the result $\hat{g}_i = \sigma_i^2$ can be used in (3.19). Ignoring the constant term, (3.19) demonstrates that Poisson statistics can be approximated by a Gaussian process[2]. The noiseless data is defined by the convolution of the trial specimen profile, $\mathbf{f}$, and the instrument profile, such that $\hat{\mathbf{g}} = \mathcal{K}\{f\}$, where $\mathcal{K}$ is the convolution operator. The likelihood probability can be stated in terms of a Gaussian distribution as

$$\begin{aligned}p(\tilde{\mathbf{g}}|\,\mathbf{f},\,\boldsymbol{\sigma}^2,\,\mathcal{I}) &\approx \prod_i \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left[-\frac{1}{2}\sum_i \left(\tilde{g}_i - \hat{g}_i\right)^2/\sigma_i^2\right] \\ &\approx \frac{\exp\left[-\frac{1}{2}C(\mathbf{f})\right]}{Z_C} \qquad (3.20)\end{aligned}$$

where the trial specimen profile, $\mathbf{f}$ and variance, σ^2, have been included on the left hand side of (3.20) to signify that $\tilde{\mathbf{g}}$ will be conditional on $\mathbf{f}$. The left hand side of (3.20) defines the probability of obtaining $\tilde{\mathbf{g}}$ conditional on the trial specimen distribution, $\mathbf{f}$, the noise variance, σ^2, and the background information, $\mathcal{I}$. In (3.20), Z_C is the normalization term, such that $Z_C^{-1} \equiv \prod_i \frac{1}{\sqrt{2\pi\sigma_i^2}}$. The variance has been approximated by $\sigma^2 \approx \tilde{\mathbf{g}}$ when the counts are very large.

[2]It is also useful to point out the above result can be obtained by applying the *central limit theorem.*

$C(\mathbf{f})$, the statistic function is given as

$$C(\mathbf{f}) = \sum_{i=1}^{M} \frac{(\hat{g}_i - \tilde{g}_i)^2}{\sigma_i^2} \tag{3.21}$$

$$\equiv \chi^2 \tag{3.22}$$

This function embodies our understanding of the experiment, the corrected observed profile, $\tilde{\mathbf{g}}$, the trial specimen profile, $\mathbf{f}$ and statistical noise; it is equivalent to the χ^2 used in classical statistics.

3.2.3 Bayes' Theorem

Two important parts of Bayes' theorem have been established, the *a priori* and the likelihood probability distribution given in (3.13) and (3.20), respectively. To unite these two terms, we must formulate the *joint probability distribution*, which is just an application of the product rule in probability theory (Cox 1946) given as

$$p(\mathbf{f},\, \tilde{\mathbf{g}}|\, \mathbf{m},\, \sigma,\, \mathcal{I}) = p(\mathbf{f}|\, \mathbf{m},\, \alpha,\, \mathcal{I})\, p(\tilde{\mathbf{g}}|\, \mathbf{f},\, \sigma,\, \mathcal{I}) \tag{3.23}$$

where the left-hand-side of (3.23) is the joint probability distribution of $\mathbf{f}$ and $\tilde{\mathbf{g}}$ being true, conditional on $\mathbf{m}$, σ and I. There is also a symmetry in the product rule, such that (3.23) can be written as

$$p(\mathbf{f},\, \tilde{\mathbf{g}}|\, \mathbf{m},\, \sigma,\, \mathcal{I}) = p(\tilde{\mathbf{g}}|\, \mathcal{I})\, p(\mathbf{f}|\tilde{\mathbf{g}}\, \mathbf{m},\, \sigma,\, \mathcal{I}). \tag{3.24}$$

Equating the results on the right-hand-side of (3.23) and (3.24) and dividing both sides by $p(\tilde{g}|\mathcal{I})$, where $p(\tilde{g}|\mathcal{I}) > 0$, we arrive at Bayes' theorem,

$$p(\mathbf{f}|\tilde{\mathbf{g}}\, \mathbf{m},\, \sigma,\, \mathcal{I}) = \frac{p(\mathbf{f}|\, \mathbf{m},\, \alpha,\, \mathcal{I})\, p(\tilde{\mathbf{g}}|\, \mathbf{f},\, \sigma,\, \mathcal{I})}{p(\tilde{\mathbf{g}}|\, \mathcal{I})}. \tag{3.25}$$

Thus, Bayes' theorem is simply a corollary of the product rule for probabilities. On the right hand side of (3.25), we have our hypothesis, given by the *a priori* probability distribution, $p(\mathbf{f}|\, \mathbf{m},\, \alpha,\, \mathcal{I})$, combined with the experimental information described by the likelihood probability distribution, $p(\tilde{\mathbf{g}}|\, \mathbf{f},\, \sigma,\, \mathcal{I})$. The result is mapped into a *posterior* probability distribution, to determine the plausibility of the specimen profile conditional on the corrected observed profile, as well as an *a priori* model and the background information, $\mathcal{I}$. This is the process of inductive reasoning.

Introducing (3.13) and (3.20) into (3.25) and ignoring the denominator in (3.25), since it is a proportionality constant, we arrive at

$$p(\mathbf{f}|\tilde{\mathbf{g}}\,\mathbf{m},\,\sigma,\,\mathcal{I}) \propto \exp \alpha\, S(\mathbf{f}) - \frac{1}{2}C(\mathbf{f}). \tag{3.26}$$

Using the same reasoning that was applied in the monkey argument, the best choice of $\mathbf{f}$ is the distribution that maximises (3.26). This is equivalent to maximising $\alpha\, S(\mathbf{f}) - \frac{1}{2}C(\mathbf{f})$, because of the monotonic nature of the exponential function. This is the essence of the MaxEnt method, which consists of determining $\mathbf{f}$ such that it maximises $S(\mathbf{f})$ subject to $C(\mathbf{f})$.

However, maximising $\alpha\, S(\mathbf{f}) - \frac{1}{2}C(\mathbf{f})$ introduces a new problem: the computational problem of developing a robust algorithm that determines the unknowns α and $\mathbf{f}$, while using the nonlinear entropy function, (3.12). This is discussed in the next section and is embodied in the Skilling & Bryan (1984) MaxEnt algorithm.

3.3 Skilling & Bryan MaxEnt Algorithm

A detailed discussion of the Skilling & Bryan (1984) algorithm is presented here. This algorithm has proven to be reliable and robust for solving a variety of inverse problems, deconvolving the observed profile to determine the specimen profile and solving the inverse problems related to crystallite size and strain distributions (see Chapters 4–6). For each problem the MaxEnt method preserves the positivity of the distribution and determines a solution which has a maximum entropy.

The maximum entropy problem can be stated by interpreting $\alpha\, S(\mathbf{f}) - \frac{1}{2}C(\mathbf{f})$ as a Lagrangian function,

$$Q(\mathbf{f}) = \alpha\, S(\mathbf{f}) - \frac{1}{2}C(\mathbf{f}) \tag{3.27}$$

where α is the unknown Lagrangian parameter. Given (3.27), we want to maximise $Q(\mathbf{f})$ with respect to $\mathbf{f}$. That is, we seek the specimen profile, $\hat{\mathbf{f}}$, such that $\nabla\, Q(\hat{\mathbf{f}}) = 0$. The geometrical interpretation of this condition is given in Figure 2.2. The problem of determining α and $\hat{\mathbf{f}}$ is made difficult because of the nonlinearity of (3.27). At this point the significance of the *a priori* model in determining the α and $\hat{\mathbf{f}}$ becomes clear. Given that $\mathbf{m}$ defines our knowledge of the specimen profile, it also contributes to the estimate of α and $\hat{\mathbf{f}}$, which in turn affects the

determination of the uncertainties in the solution. By using a uniform *a priori* model, our lack of knowledge is being incorporated into the MaxEnt method and any solution represents the worst case scenario. In other words, if our estimate of the *a priori* model is accurate or contains information which has "structure", we expect the solution profile to approach the actual profile and the uncertainties in the solution profile to be reduced.

Maximising (3.27) is essentially a nonlinear optimization problem. The use of the logarithmic form of the entropy function is an important feature, because it preserves the positivity of the solution profile and distributions. The nonlinearity also implies that a computational algorithm can only determine the solution iteratively, rather than in an exact form. It is shown in this section that the Skilling & Bryan (1984) algorithm satisfies these requirements.

3.3.1 The rationale of the Skilling-Bryan algorithm

The MaxEnt method is commonly applied to the numerical solution of inverse problems. As we have seen, in x-ray diffraction profile analysis the inverse problems of deconvolving the instrument kernel from the observed data, and determining the crystallite size and strain distributions are expressed as integral equations of the first kind (also see Chapters 4–5).

The discussion in Chapter 2 demonstrated that in the case of deconvolving the observed data, certain methods at best produced ill-conditioned specimen profiles. It was shown that ill-conditioning occurred when the kernel was near singular. In the worst case, the kernel may be singular and not have a unique inverse. To picture this, we define a set of feasible solutions, $\{\mathbf{f}\}$. This is analogous to the monkey argument where a set of feasible distributions, $\{\mathbf{p}\}$ was defined. The aim is to determine the most probable solution, $\hat{\mathbf{f}}$, from the set of feasible solutions. As discussed earlier, the most probable solution is the solution with the maximum entropy.

The first step in achieving this is to define a subset of solutions, $\{\mathbf{f}\}_{subset}$ which are consistent with the experimental data. The solutions which are not consistent with the experimental data are simply rejected. The criteria for defining the subset is given by the statistic function, $C(\mathbf{f})$, by imposing the constraint $C(\mathbf{f}) \leq C_{aim}$, where C_{aim} is some preset critical value. It is normally set equal to the number of data points, M.

In $\{\mathbf{f}\}_{subset}$ there are still many feasible solutions, which are consistent with $C(\mathbf{f}) \leq C_{aim}$. From these, we want to choose a single solution. This is achieved by imposing the entropy function on the subset. The solution that maximises the entropy function is the maximally noncommittal solution.

3.3.2 Defining the subspace

The Skilling & Bryan (1984) algorithm does not attempt to determine the MaxEnt solution in the space mapped by $Q(\mathbf{f})$. This would be a computationally demanding task since $\mathbf{f} = \{f_i; i = 1, 2, 3, \ldots, M\}$, where M can be quite large. It introduces difficulties when attempting to prevent search vectors from becoming linearly dependent. Moreover, in methods such as the Newton-Raphson technique, the Hessian matrix may become ill-conditioned, producing ill-conditioned solutions. To overcome this, a subspace with fewer dimensions, r, than M is defined. This allows the problem to be mapped into the subspace where the solution is determined, then mapped back into the solution space (i.e. $\mathbf{f}$-space). This is a computationally inexpensive task and also has the benefit that the conditioning of the basis vectors, which span the subspace, can be guarded against becoming linearly dependent.

The subspace is defined by a set of basis vectors, e_1,..., e_r, where $3 \leq r \leq 10$. We define the first two basis vectors as,

$$\mathrm{e}_1 = \mathbf{f}\,(\nabla S) \tag{3.28}$$

and

$$\mathrm{e}_2 = \mathbf{f}\,(\nabla C) \tag{3.29}$$

where $\nabla S = \{\partial S/\partial f^i;\ i = 1, 2, 3, \ldots, M\}$ and $\nabla C = \{\partial C/\partial f^i;\ i = 1, 2, 3, \ldots, M\}$. The multiplication in (3.28) and (3.29) is component-to-component, such that $\mathbf{f}\,(\nabla S) = \{f^1\, \partial S/\partial f^1, \ldots,\ f^M\, \partial S/\partial f^M$ and similarly for $\mathbf{f}\,(\nabla C)$. The weighting by $\mathbf{f}$ favours more intense regions of the distribution and prevents negative values. Futhermore it is noticeable that the terms ∇S and ∇C from (3.28) and (3.29) are drawn from ∇Q. This suggests that the next set of vectors for the subspace can be drawn from $\nabla\nabla Q$. Allowing $\nabla\nabla S$ and $\nabla\nabla C$ to operate on e_1 and e_2, respectively, and weighting each term with $\mathbf{f}$, four terms are produced. However, this decreases to two, since $\mathbf{f}\,(\nabla\nabla S) = -\mathbf{I}$, where $\mathbf{I}$ is the identity matrix and nothing new is

obtained. Using the remaining terms, a third vector can be defined as

$$e_3 = |\nabla S|^{-1}\,\mathbf{f}(\nabla\nabla C)\,\mathbf{f}\,(\nabla S) - |\nabla C|^{-1}\,\mathbf{f}(\nabla\nabla C)\,\mathbf{f}\,(\nabla C) \tag{3.30}$$

Additional basis vectors can be determined in a similar manner, but e_1, e_2, e_3 are usually sufficient; however, this does depend upon the complexity of the problem.

At this point it is useful to reinforce the above discussion and place it in a wider context. Usually in the Raphson-Newton method, the Hessian matrix, $\nabla\nabla Q^{-1}$, is determined for each step. However, the choice of the basis vector, (3.28–3.30) "builds up" information about the Hessian matrix (Wu 1997). It should be noted that the Newton-Raphson method assumes the curvature is parabolic. However, this is not the case in the computational problem of determining the maximum entropy solution, where the curvature of $Q(\mathbf{f})$ is dominated by the entropy function. It is this geometrical property which underlies the problem and defines the inner product of vectors.

3.3.3 Entropy metric

The curvature of the solution space is dominated by the entropy metric, defined as,

$$\begin{aligned} g_{ij} &= (-\nabla\nabla S) && (3.31)\\ &= \delta_{ij}/f^i && (3.32)\end{aligned}$$

where δ_{ij} is the Kronecker delta. The entropy metric is a covariant tensor quantity. The transformation from covariant to contravariant metric, $g_{ij} \to g^{ij}$, is given by

$$\begin{aligned} g^{ij} &= \frac{\mathrm{co}(g_{ij})}{\det(g_{ij})} && (3.33)\\ &= \delta_{ij} f^j && (3.34)\end{aligned}$$

where $\mathrm{co}(g_{ij})$ is the cofactor of g_{ij}, determined by eliminating the ith-row and jth-column from g_{ij} and taking the determinant of the remaining matrix; $\det(g_{ij})$ is the determinant of g_{ij}. This simply means that the transformation $g_{ij} \to g^{ij}$ involves taking the inverse of the g_{ij} matrix. The length of the contravariant quantity is given by

$$\begin{aligned} |\mathbf{f}| &= \sqrt{g_{ij}\, f^i\, f^j} && (3.35)\\ &= \sqrt{\sum_j f^j} && (3.36)\end{aligned}$$

and using the contravariant form of the entropy metric, (3.34), the inner product of covariant quantities, such as ∇S and ∇C is given by

$$\begin{aligned} |\nabla S|^2 &= \mathrm{g}^{ij} \, (\nabla S)_i \, (\nabla S)_j \\ &= \sum_j f^j \, (\partial S / \partial f^j)^2 \end{aligned} \tag{3.37}$$

and

$$\begin{aligned} |\nabla C|^2 &= \mathrm{g}^{ij} \, (\nabla C)_i \, (\nabla C)_j \\ &= \sum_j f^j \, (\partial C / \partial f^j)^2 \end{aligned} \tag{3.38}$$

where ∇S and ∇C are covariant vectors. Hence, by using g^{ij} in (3.37) and (3.38), respectively, the vectors are being transformed from covariant to contravariant vectors. This also implies that $\nabla\nabla S$ and $\nabla\nabla C$ must be multiplied by g^{ij} to ensure that the mapping from a covariant to contravariant vector is preserved; it also accounts for the weighting terms in (3.28–3.30).

3.3.4 Formulating the problem in the subspace

Now that the subspace and inner product of quantities have been defined, the problem of determining the $\mathbf{f}$ that maximizes (3.27) is transformed into this subspace. Within the subspace a generalized coordinate system, x^μ where $\mu = 1, 2, \ldots, r$, is adopted and the entropy and statistic functions can be expanded as

$$\tilde{S}(\mathbf{x}) = S_0 + S_\mu \, x^\mu - \frac{1}{2} \, g_{\mu\nu} \, x^\mu \, x^\nu \tag{3.39}$$

and

$$\tilde{C}(\mathbf{x}) = C_0 + C_\mu \, x^\mu + \frac{1}{2} \, M_{\mu\nu} \, x^\mu \, x^\nu \tag{3.40}$$

where, $S_\mu = \mathrm{e}_\mu^T \cdot \nabla S$; $\mathrm{g}_{\mu\nu} = \mathrm{e}_\mu^T \cdot \mathrm{e}_\nu$; $C_\mu = \mathrm{e}_\mu^T \cdot \nabla C$ and $M_{\mu\nu} = \mathrm{e}_\mu^T \cdot \nabla\nabla C \cdot \mathrm{e}_\nu$. The next iteration is given by

$$\mathbf{f}^{new} = \mathbf{f} + x^\mu \mathrm{e}_\mu \tag{3.41}$$

$$= \mathbf{f} + \delta \mathbf{f} \tag{3.42}$$

where the length of the step-size is given by

$$l = \sqrt{\mathrm{g}_{\mu\nu} x^\mu x^\nu}. \tag{3.43}$$

Orthogonalizing the subspace

The vectors spanning the subspace, $\{e_\mu\}$, are not orthogonal and may be linearly dependent. In order to prevent this, the subspace must be orthonormalized (using the Gram-Schimdt process). A consequence of this is that the matrices, $g_{\mu\nu}$ and $M_{\mu\nu}$ are simultaneously diagonalized. This provides a way of guarding against linear dependency. If the eigenvalues of $\mathsf{g}_{\mu\nu}$ are near to zero, they are discarded with their corresponding eigenvectors. Both matrices will be $[r \times r]$ symmetrical matrices. The $M_{\mu\nu}$ matrix will be convex, which implies that its eigenvalues will all be positive. Another consequence of the orthonormalization of the subspace is that the coordinate system is transformed into a Cartesian system and the distinction between covariant and contravariant vectors disappears. In the Cartesian subspace, (3.39) and (3.40) become,

$$\tilde{S}(\mathbf{x}) = S_0 + S_\mu x_\mu - \frac{1}{2} x_\mu x_\mu \tag{3.44}$$

and

$$\tilde{C}(\mathbf{x}) = C_0 + C_\mu x_\mu + \frac{1}{2} \lambda_\mu x_\mu x_\mu \tag{3.45}$$

where λ_μ are the eigenvalues of $M_{\mu\mu}$. The step size in the orthonormalized space is

$$l = \sqrt{x_\mu x_\mu}. \tag{3.46}$$

Basic Control

In the orthonormalized subspace the search for a maximum begins. That is, we maximize $\tilde{S}$ subject to the constraints, $\tilde{C} = C_{aim}$ and $l^2 = l_0^2$. However, the eigenvalues, λ_μ, are positive which implies that the $\tilde{C}(\mathbf{x})$ has a minimum defined by $x_\mu = -C_\mu/\lambda_\mu$, given by

$$\tilde{C}_{min} = C_0 - \frac{1}{2}\lambda_\mu^{-1} C_\mu C_\mu. \tag{3.47}$$

This value may not be attained or may exceed the preferred C_{aim}. In such a case a modest statistic value is chosen, given by

$$\tilde{C}_{aim} = \max\{\frac{2}{3}\tilde{C}_{min} + \frac{1}{3}C_0, C_{aim}\}. \tag{3.48}$$

$\tilde{C}$ decreases monotonically and will eventually settle onto the preferred value. The maximum step-size that is allowed in the subspace is set by

$$l^2 \leq l_0^2 \approx (0.1 \sim 0.5 \sum_j f_j). \tag{3.49}$$

Using (3.44) and (3.45), the Lagrangian function, $\tilde{Q}$, is defined in the subspace as

$$\tilde{Q}(\mathbf{x}) = \alpha\tilde{S}(\mathbf{x}) - \tilde{C}(\mathbf{x}) \tag{3.50}$$

and can be maximized such that, $\partial\tilde{Q}/\partial x_\mu = 0$. This results in

$$x_\mu = \frac{\alpha S_\mu - C_\mu}{\lambda_\mu + \alpha}. \tag{3.51}$$

The result in (3.51) can be substituted into (3.45) and an α can be determined such that $\tilde{C} = \tilde{C}_{aim}$. Using each calculated value of α, (3.51) is then evaluated and checked against (3.49). If the inequality is satisfied, the next increment, $\mathbf{f}^{new}$, can be determined. This is repeated until $\tilde{C} = C_{aim}$.

Additional control

There are cases, especially in the initial iterations, where the preferred step-size, l^2, is over-estimated i.e. when (3.49) does not hold. This threatens the expansion of (3.44) and (3.45) within the subspace. To guard against this, an additional constraint must be included. The Lagrangian function in the subspace is modified, such that,

$$\tilde{Q}(\mathbf{x}) = \alpha\tilde{S}(\mathbf{x}) - \tilde{C}(\mathbf{x}) - \beta l^2 \tag{3.52}$$

where β is an additional Lagrangian multiplier. Maximizing (3.52) $\partial\tilde{Q}/\partial x_\mu = 0$ produces

$$x_\mu = \frac{\alpha S_\mu - C_\mu}{\lambda_\mu + \alpha + \beta}. \tag{3.53}$$

For a given α, (3.53) can be substituted into (3.45) and a β determined such that $\tilde{C} = \tilde{C}_{aim}$. Once β has been calculated, (3.53) can be evaluated and used to determine $\mathbf{f}^{new}$.

Having determined $\mathbf{f}^{new}$, a new set of basis vectors is calculated and the above steps repeated, until $\tilde{C} = C_{aim}$.

Convergence criteria

After each iteration the parallelness of the ∇S and ∇C vectors can be checked against a preset value, δ_{aim}. The geometrical interpretation for this condition is given in Figure 2.2. The parallelness of ∇S and ∇C is given by

$$\delta = \frac{1}{2}\left|\frac{\nabla S}{|\nabla S|} - \frac{\nabla C}{|\nabla C|}\right|^2 \tag{3.54}$$

If $\tilde{C} = C_{aim}$ and $\delta \leq \delta_{aim}$, the algorithm can be halted.

3.3.5 Determining the uncertainty in $\hat{\mathbf{f}}$

The result in (3.26) defines the posterior probability distribution of $\mathbf{f}$ in an M-dimensional space. This distribution has a maximum centered about the solution profile, $\hat{\mathbf{f}}$. In addition, the "spread" in $p(\mathbf{f}|\tilde{\mathbf{g}}\,\mathbf{m},\,\sigma,\,\mathcal{I})$ defines the uncertainty about the $\hat{\mathbf{f}}$. If approximated as a Gaussian distribution in $\mathbf{f}$-space, the variance of $\hat{\mathbf{f}}$ can be determined.

We begin by first substituting (3.27) into (3.26), given as

$$p(\mathbf{f}|\tilde{\mathbf{g}},\,\mathbf{m},\,\sigma,\,\mathcal{I}) \propto \exp\left[Q(\mathbf{f})\right]. \tag{3.55}$$

Expanding (3.55) in a Taylor series about $\hat{\mathbf{f}}$, we have

$$\begin{aligned} p(\mathbf{f}|\tilde{\mathbf{g}},\,\mathbf{m},\,\sigma,\,\mathcal{I}) &\propto \exp\left[Q(\hat{\mathbf{f}}) - \frac{1}{2}(\mathbf{f}-\hat{\mathbf{f}})^T\nabla\nabla Q(\hat{\mathbf{f}})(\mathbf{f}-\hat{\mathbf{f}})\right] \\ &\propto \exp\left[Q(\hat{\mathbf{f}})\right]\,\exp\left[-\frac{1}{2}(\mathbf{f}-\hat{\mathbf{f}})^T\nabla\nabla Q(\hat{\mathbf{f}})(\mathbf{f}-\hat{\mathbf{f}})\right] \end{aligned} \tag{3.56}$$

where $\exp\left[Q(\hat{\mathbf{f}})\right]$ is a constant. In this expansion the first derivative term, $\nabla Q(\hat{\mathbf{f}})$, vanishes, since it indicates the maximum of $p(\mathbf{f}|\tilde{\mathbf{g}},\,\mathbf{m},\,\sigma,\,\mathcal{I})$. The second derivative (i.e. the second term in (3.56)) is a Gaussian centered about $\hat{\mathbf{f}}$. We can therefore conclude that the variance will be given by

$$\sigma^2_{\hat{\mathbf{f}}} = \nabla\nabla Q(\hat{\mathbf{f}})^{-1} \tag{3.57}$$

where (3.57) is evaluated using $\hat{\mathbf{f}}$. If $\nabla\nabla Q^{-1}$ is a diagonal matrix, the elements will define the variance of $\hat{\mathbf{f}}$. However, in general, $\nabla\nabla Q^{-1}$ is not diagonal, rather it defines an $[M \times M]$ correlation matrix. As suggested by Skilling (1990), the uncertainties in $\hat{f}$, $\sigma_{\hat{f}}$, can be determined by evaluating

$$\sigma^2_{\hat{f}} = \mathbf{u}^T \cdot [\nabla\nabla\, Q]^{-1} \cdot \mathbf{u} \tag{3.58}$$

where $\sigma^2_{\hat{f}} = \{\sigma^2_{\hat{f}\,i};\, i = 1,\,2,\,3,\ldots,M\}$; $[\nabla\nabla\, Q]^{-1}$ is the correlation matrix; $\mathbf{u}$ is an M-dimensional vector. In general the correlation matrix will not be diagonal. If it were, determining the uncertainties in $\hat{\mathbf{f}}$ would simply involve taking the diagonal terms; however, ignoring the off-diagonal terms when $[\nabla\nabla\, Q]^{-1}$ is a "full" matrix could result in information about the data not being incorporated into the uncertainties. In order to avoid this, the vector $\mathbf{u}$ in (3.58) is defined and takes into consideration a region of interest which corresponds to the resolution of the data Hansen & Pedersen (1991). By "region of interest" it is meant

the region where the i-th element in $\hat{\mathbf{f}}$ has ones on either side and zeros elsewhere. The uncertainties in (3.58) account for the deconvolution, but do not account for systematic errors that occur in background level estimation. Other approaches of determining the uncertainty regions in the MaxEnt solution have been suggested by Skilling et al. (1991) and Hansen & Wilkins (1994).

The important implication that arises out of the above discussion is that uncertainties in the specimen profile can be used to determine uncertainties in any subsequent calculations and their physical quantities, both of which arise from the Warren-Averbach and Williamson-Hall methods. Another consequence of determining the uncertainties in the specimen profile is that the uncertainty regions in column-length and strain distributions can be determined. This enables column-length and strain distributions to be used in a comparison with theoretical models.

3.3.6 Limitations of the Skilling-Bryan Algorithm

The Skilling & Bryan (1984) MaxEnt algorithm discussed in this section does have some difficulties which are mostly concerned with the theoretical and the epistemological developments of MaxEnt. At the centre of the theoretical difficulties is how the Lagrangian parameter, α, is determined (see Titterington 1985, Skilling 1989*a*, Gull 1989, Bryan 1990, MacKay 1991, Donoho et al. 1992, Strauss et al. 1993, Wolpert & Strauss 1995). More recently, a difficulty has arisen concerning the application of a term known as the *evidence* (i.e. the denominator in (3.25)) (see Strauss et al. 1993, Wolpert & Strauss 1995).

In the discussion of the Skilling & Bryan (1984) algorithm, above, it was pointed out that $C(\mathbf{f})$ is set equal to the number of data points. Under this constraint α is determined. It is worth noting that the $C(\mathbf{f}) = M$ condition is somewhat arbitrary, since $C(\mathbf{f})$ can take on values in the range $M \pm \sqrt{2M}$. Skilling (1990) has pointed out that this criteria is a "frequentist" view and does not constitute a Bayesian method. It assumes on average $C(\mathbf{f}) = M$ would occur after many repeated measurements. However, the specimen profile is only measured once. As pointed out by Titterington (1994), the $C(\mathbf{f}) = M$ criteria can lead to an under-fitting of the data and the solution will be biased towards the observed data in the case of little blurring.

In order to overcome these difficulties, Gull (1989) developed a MaxEnt method based on Bayesian theory. This can be achieved by extending (3.26) to include an *a priori* distribution for α, resulting in a joint probability distribution which includes $\mathbf{f}$ and α. In this distribution, either one of the terms can be integrated out. This derivation results in the entropy function defining the number of "good data", M_g, and the statistic function defining the number of "bad data", M_b, such that $M = M_g + M_b$. This is the critical difference with the Skilling & Bryan (1984) algorithm, in that the good data are used to determine α, rather than using the $C(\mathbf{f}) = M$ criteria. This results in (Gull 1989)

$$-2\hat{\alpha}S = \sum_i \frac{\lambda_i}{\hat{\alpha} + \lambda_i} \tag{3.59}$$

$$= M_g \tag{3.60}$$

where λ_i are the eigen-values from Hessian matrix in the entropy metric (see Gull 1989, Müller et al. 1996, Sivia 1996). The variance, on the other hand, can be determined from M_b. Using (3.60) assumes that the posterior distribution for α is narrow (Wolpert & Strauss 1995, Sivia 1996). This usually corresponds to $M_g \gtrsim 10$ (Sivia 1996). Bryan (1990) has developed an algorithm for the case where $M_g << 10$. It is interesting to note that Geoghegan et al. (1996) have applied the $C(\mathbf{f}) = M$ criteria to neutron reflectometry data.

Nevertheless, in the application to x-ray profile analysis, simulations and calculations have shown that the Skilling & Bryan (1984) algorithm is a reliable and consistent algorithm. This will be demonstrated in Chapters 4, 5 & 6.

3.4 Summary

In this chapter, the MaxEnt method has been presented as a new method for determining the specimen profile from an observed x-ray diffraction profile. This is based on the charactisitics of x-ray diffraction profiles being positive and additive distributions. These characteristics allowed the entropy function to be defined. Based on this, it has been shown that the entropy function is the *only* function which is consistent with the principles of statistical inference. Futhermore, the generalised entropy function incorporates an *a priori* model concerning the unknown specimen profile.

The MaxEnt method was formulated using Bayesian theory which introduces *a priori* information about the characterisitics of the profile and the nature of the experiment. It has also been shown how the uncertainties in the solution profile can be determined. This important extension makes the MaxEnt method a fully quantitative method for analysing x-ray diffraction profiles, enabling the uncertainties in subsequent results to be determined. Moreover, the general formulation of the MaxEnt method allows other inverse problems to be solved, such as determining the column-length and strain distributions from the observed profile.

Chapter 4

A Maximum Entropy Method for Determining the Column-length Distribution

In this chapter, a novel two-fold application of the MaxEnt method is presented; the first application determines the specimen profile from an observed size-broadened x-ray diffraction profile, and the second obtains, from this the area-weighted column length distribution.

The two-fold procedure is important in presenting an altenative approach to removing instrumental broadening and determining the specimen profile and draws upon the discussion of the MaxEnt method presented in Chapter 3. Once the specimen profile is known, conventional methods for crystallite/domain size and microstrain strain analysis can be applied. In the particular case of crystallite-size broadening only (i.e. no strain present), the determination of the specimen profile and column-length distribution can be combined into a single step (see Appendix C). However, in general we cannot know *a priori* that crystallite-size effects are the only source of broadening. The usual approach is to determine the specimen profile first by removing instrumental broadening and then to apply semi-quantitative methods such as the Williamson & Hall (1953) method (see §1.4.2) to them to establish the nature of the broadening (see Langford & Louër 1996). In the two step approach described here, the MaxEnt method determines the set of distributions (*viz.* specimen profile and column-length) that are consistent with the experimental data while maximiz-

ing the entropy function, thus producing the *maximally noncommittal* solution (Skilling & Bryan 1984, Sivia 1996, Wu 1997). The work described in this chapter has been accepted for publication in JAC (see Armstrong & Kalceff 1999*a*).

The MaxEnt method for determining the column-length distribution is presented in §4.2. To evaluate the performance of the MaxEnt method and the integrity of its results, extensive simulations and calculations have been carried out (see §4.3). The simulations report three cases (see §4.4): the first one determines the best background level and examines the effect of truncation on the specimen profile and column distribution, as well as the subsequent results. The second case examines the effects of under- and over-estimation of the background level on the specimen profile, column-length distribution and subsequent results. Finally, we examine the effects of deconvolving with non-ideal instrument profiles and the effect this has on the specimen profile, Fourier coefficients, column-length distribution, and area- and volume-weighted sizes. In all cases the uncertainties in the specimen profile, Fourier coefficients, column-length distribution and size results are determined. The implications of these calculations for experimental conditions used in gathering diffraction data are discussed in §4.5.

4.1 Crystallite size broadening and analysis

4.1.1 Crystallite size broadening

The models for size broadening of x-ray diffraction profiles in polycrystalline materials described by Bertaut (1950, 1952) and Warren & Averbach (1950, 1952) represent the crystallites (or diffracting domains) as consisting of columns perpendicular to the diffracting planes. The columns making up the crystallites diffract independently and the resulting size-broadened profile can be interpreted as a weighted sum of intensities from each column. The weighting terms correspond to the fraction of columns having length L to $L + \mathrm{d}L$ that contributes to the intensity distribution. This can be expressed in a continuous form as

$$f(s - s_0) = \int_0^\infty K(s - s_0,\, L)\, p_a(L)\, \mathrm{d}L \tag{4.1}$$

where $f(s - s_0)$ is the size-broadened profile defined in reciprocal space dimensions, $s - s_0$; $K(s - s_0, L)$ is the scattering kernel that defines the intensity being diffracted from a column

of length L, perpendicular to the diffraction plane over a range of $s - s_0$ in reciprocal space; $p_a(L)\,\mathrm{d}L$ is the area-weighted column-size distribution for columns of length L to $L + \mathrm{d}L$ perpendicular to the diffraction planes. We require that $p_a(L) \geq 0$ for all L, satisfying $\int_0^\infty p_a(L)\,\mathrm{d}L = 1$.

It was also shown by Bertaut (1950, 1952) and Warren & Averbach (1950, 1952) that $\langle L \rangle_a$ can be determined from the initial slope of the Fourier coefficients of $f(s - s_0)$, given by Warren (1969) as

$$\frac{1}{\langle L \rangle_a} = -\frac{\mathrm{d}\,A(L)}{\mathrm{d}L} \tag{4.2}$$

as $L \to 0$. The $\langle L \rangle_a$ term in (4.2) represents the average area-weighted apparent crystallite /domain size; that is, it is the ratio of the total volume of the crystallite to the area projected onto the crystallographic plane. This also corresponds to $V_c(0)/V_c'(0)$, where V_c is the common (or ghost volume) and $V_c'(0) = \mathrm{d}V_c(L)/\mathrm{d}L$ as $L \to 0$ (Langford & Wilson 1978).

The column size-distribution in (4.1) can also be expressed in terms of the second derivative of the Fourier coefficients of $f(s - s_0)$ (see equation (13.31) in Warren 1969). Alternatively, the Fourier coefficients can be expressed in terms of the area-weighted column distribution as

$$A(L) = \frac{1}{\langle L \rangle_a} \int_{|L|}^{\infty} (L' - L)\, p_a(L')\,\mathrm{d}L'. \tag{4.3}$$

The area-weighted size and column length distributions can be related to another measure of the crystallite/domain size, the volume-weighted crystallite size (Guinier 1963),

$$p_v(L) = \frac{L}{\langle L \rangle_a}\, p_a(L) \tag{4.4}$$

where $p_v(L) \geq 0$ for all L, satisfying $\int_0^\infty p_v(L)\,\mathrm{d}L = 1$. The volume-weighted size, $\langle L \rangle_v$ is identical to the result described by Stokes & Wilson (1942), where they relate $\langle L \rangle_v$to the integral breadth (profile area/peak maximum) of a size-broadened profile for a given crystallite shape, expressed as

$$\langle L \rangle_v = \frac{\lambda}{\beta(2\theta)\,\cos\theta} \tag{4.5}$$

where λ is the wavelength of radiation used and $\beta(2\theta)$ is the integral breadth of the size-broadened profile determined in 2θ-space. The integral breadth can also be expressed in terms of the common-volume (or ghost) function, $V_c(L)$, for a crystal translated a distance L perpendicular to the hkl-plane (see equation (1) in Stokes & Wilson 1942).

The area-weighted distribution, p_a, is defined by the direction of the diffraction vector and geometry of the crystallite. It does not correspond to the distribution of crystallites measured by direct methods applied in electron microscopy (Smith 1976). For $p_a(L)$ to be related to the fraction of crystallites, $P(D)\,\mathrm{d}D$, that have a dimension between D and $D + \mathrm{d}D$, it must be related via the shape of the crystallite. This can be expressed as (Smith 1976),

$$p_a(L) = \int_{D_0(L)}^{\infty} G(L,\, D)\, P(D)\, \mathrm{d}D \tag{4.6}$$

where $D_0(L)$ is a function that defines the smallest column and $G(L,\, D)$ is the *shape kernel* that maps the dimensions of the crystallites into the column lengths. $G(L, D)$ is expressed in terms of the second derivative; in the special case of a distribution of spherical crystallites, it is given by Smith (1976) as

$$G(L,\, D) = \frac{\pi}{2} L \tag{4.7}$$

and $D_0(L) = L$. The column distribution is given by

$$p_a(L) = \frac{\pi}{2} L \int_{L}^{\infty} P(D)\, \mathrm{d}D. \tag{4.8}$$

where $P(D)$ is the distribution of diameters. In (4.8), both distributions are independent of crystallographic direction.

The problem of determining f and p_a from experimental data is complicated by the finite response and aberrations of the diffractometer. In the next section the convolution equation that produces the observed profile is discussed.

4.1.2 Instrument broadening

The observed profile, $g(2\theta)$, at the output of a diffractometer is the convolution product of the specimen profile and an instrument profile, superimposed onto a statistical noise and background level according to

$$g(2\theta) = \int k(2\theta - 2\theta')\, f(2\theta')\, \mathrm{d}(2\theta') + b(2\theta) + n(2\theta) \tag{4.9}$$

where $k(2\theta - 2\theta')$ is the instrument profile, and $f(2\theta)$ is the specimen profile containing information about its microscopic properties[1]; $b(2\theta)$ is the background level and $n(2\theta)$ is the

[1] This chapter only considers crystallite size broadening. However, in general the specimen profile will contain both microstrain and crystallite size broadening that will need to be separated (also see Warren

noise distribution.

In removing the instrument profile and determining $f(2\theta)$, it is essential that the solution (i.e. specimen) profile be free from spurious oscillations and for the positivity of $f(2\theta)$ to be preserved, otherwise errors are introduced into subsequent calculations.

4.2 Maximum Entropy Method for X-ray Profile Analysis

The dual problems of removing the instrument broadening from an observed x-ray diffraction profile (by deconvolving (4.9)) and determining the column-distribution in (4.1) are good candidates for the MaxEnt method, in that both are *inverse problems*. Information concerning p_a or f is embedded in (4.1) & (4.9), respectively. In applying the MaxEnt method we extract the maximally noncommittal solutions from (4.1) & (4.9).

4.2.1 Background estimation of the observed profile

Before determining the specimen profile, $f(2\theta)$, the background level of the observed profile, $g(2\theta)$, must be removed. We write

$$\tilde{g} = g - \hat{b} \tag{4.10}$$

where $\hat{b}$ is the estimated background level and $\tilde{g}$ is the background-corrected observed profile.

The effects of poor background correction are well known (see Bertaut 1952, Warren 1959, Pines & Sirenko 1962, Young et al. 1967, Warren 1969). Over-estimating the background produces the so-called "hook effect", in which a negative curvature appears near the origin of a plot of the Fourier coefficients; this is because there is a reduction in the area of the profile, and the zeroth-order coefficient, which is proportional to the area of the profile, therefore decreases relative to the higher order coefficients. When (4.2) is applied, the negative curvature causes an over-estimation of $\langle L \rangle_a$. On the other hand, when there is under-estimation of the background, the total area of the profile is increased and the zeroth-order coefficient is increased relative to the higher order coefficients. Thus, a discontinunity

1969, van Berkum et al. 1996)

or "kink" appears after the zeroth-order in a plot of the Fourier coefficients. When (4.2) is applied, this causes an under-estimation of $\langle L \rangle_a$.

A qualitative method for determining the best background using the MaxEnt method is presented in §4.4.1 & 4.4.3, together with calculations showing the effects of systematic under- and over-estimation of the background level on the MaxEnt solution specimen profile, Fourier coefficients, column-length distribution and apparent size results (see §4.4.2).

The MaxEnt method for determining the specimen profile was presented in Chapter 3. After the background has been removed, the MaxEnt method is applied in order to determine the specimen profile. It essentially involves maximizing (3.27). This is the first level of application. The second level involves determining the column-length distribution.

4.2.2 Determining the column distribution

In deconvolving (4.9), the instrument profile was assumed to be shift-invariant over the interval defining the observed profile. This enabled us to express the trial observed profile in terms of the Fourier convolution operator. The problem of determining the area-weighted column length distribution, $p_a(L)$, is very similar, with the exception that the scattering kernel, $K(s - s_0, L)$, in (4.1) is not linearly shift-invariant.

The entropy function for this problem is similar to (3.12), but with f replaced by p_a:

$$S(p_a) = -\sum_{j=1}^{N} p_{a\,j} \ln(\frac{p_{a\,j}}{m_{a\,j}}) \tag{4.11}$$

where $p_a = \{p_{a\,j};\ j = 1, 2, 3, \ldots, N\}$ is the solution column-length distribution and $m_a = \{m_{a\,j};\ j = 1, 2, 3, \ldots, N\}$ is the *a priori* model for the column-length distribution. The statistic function used in (3.21) now uses the solution specimen profile, $\hat{f}$, as the "observed" profile from which the column-length distribution will be extracted. The scattering kernel given in (4.1) is expressed as an $[M \times N]$ matrix, such that $\mathbf{K} = \{K_{ij}; i = 1, 2, 3, \ldots, M, j = 1, 2, 3, \ldots, N\}$, where $N \leq M$. The statistic function for this problem is given by

$$C(p_a) = \sum_{i=1}^{M} \frac{(\tilde{f}_i - \hat{f}_i)^2}{\sigma_{\hat{f}\,i}^2} \tag{4.12}$$

where $\hat{f}$ is the solution specimen profile determined from the MaxEnt method; $\tilde{f} = \{\tilde{f}_i; i = 1, 2, 3, \ldots, M\}$ is the trial specimen profile given by $\tilde{f}_i = \sum_{j=1}^{N} K_{i\,j}\, p_{a\,j}$; and $\sigma_{\hat{f}\,i}$ are the

uncertainties corresponding to $\hat{f}$ calculated using (3.58).

Using (4.11) and (4.12), the Lagrangian can be expressed similarly to (3.27) and the Skilling & Bryan (1984) algorithm can be modified for the new kernel, $\mathbf{K}$, and used to determine $\hat{p}_a$. As discussed in §3.3.5, the uncertainties in the calculated $\hat{p}_a$ can be determined. The average value, $\langle L \rangle_a$, of the column lengths can be determined using

$$\langle L \rangle_a = \int_0^\infty L \, p_a(L) \mathrm{d}L. \tag{4.13}$$

Using the result from (4.4), $p_v(L)$ can be determined and the volume-weighted result, $\langle L \rangle_v$, can also be evaluated in a similar manner to (4.13).

4.3 Simulated profiles

4.3.1 Simulated size broadened profiles

The size-broadened profiles used were modelled for an alumina sample assuming spherical crystallites and no lattice distortions or microstrain. The modelling did not take into account the absorption of x-rays or other effects such as extinction and diffuse scattering. The 113 and 226 multiple orders were used to test the performance of the MaxEnt method under different conditions.

Particle distribution, $P(D)$

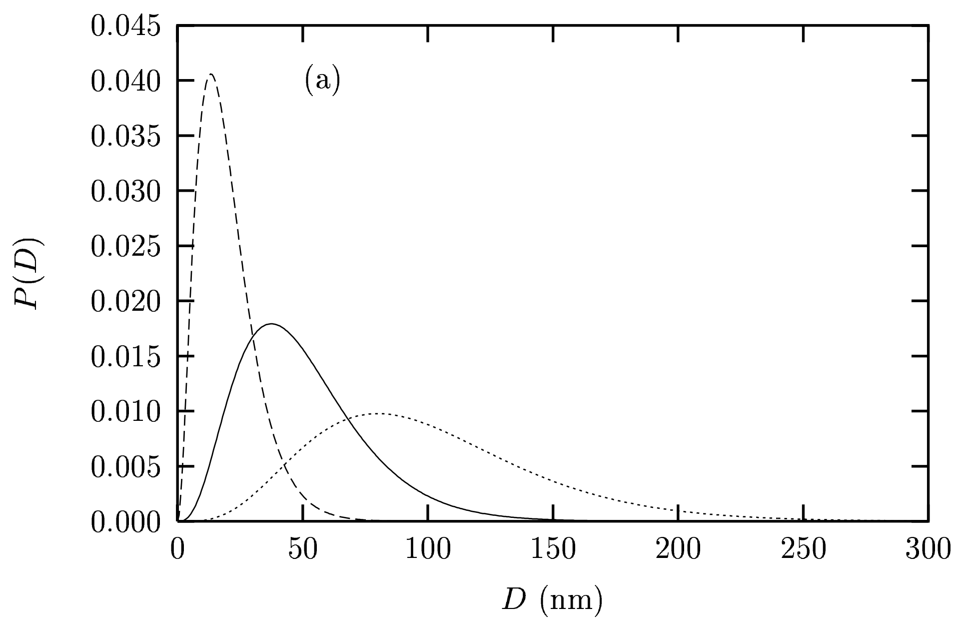
A set of particle distributions with average crystallite diameters corresponding to $\langle D \rangle =$ 20, 50 & 100 nm were calculated using the function (also see Delhez et al. 1982)

$$P(D; u, t, r) = \frac{1}{Z_D(u, t, r)} \, D^r \exp(-uD^t) \tag{4.14}$$

where D is the diameter of the spherical crystallites, u, t and r are parameters, and $Z_D(u, t, r)$ is a normalization term that ensures $\int_0^\infty P(D; u, t, r) \, \mathrm{d}D = 1$, given by

$$Z_D(u, t, r) = t^{-1} \, u^{-\frac{r+1}{t}} \, \Gamma\left(\frac{r+1}{t}\right) \tag{4.15}$$

with $\Gamma(x)$ being the Gamma function. From (4.14), the average diameter can be determined by $\langle D \rangle = \int_0^\infty D \, P(D) \, \mathrm{d}D$. Table 4.1 gives the values of u, t and r, and the



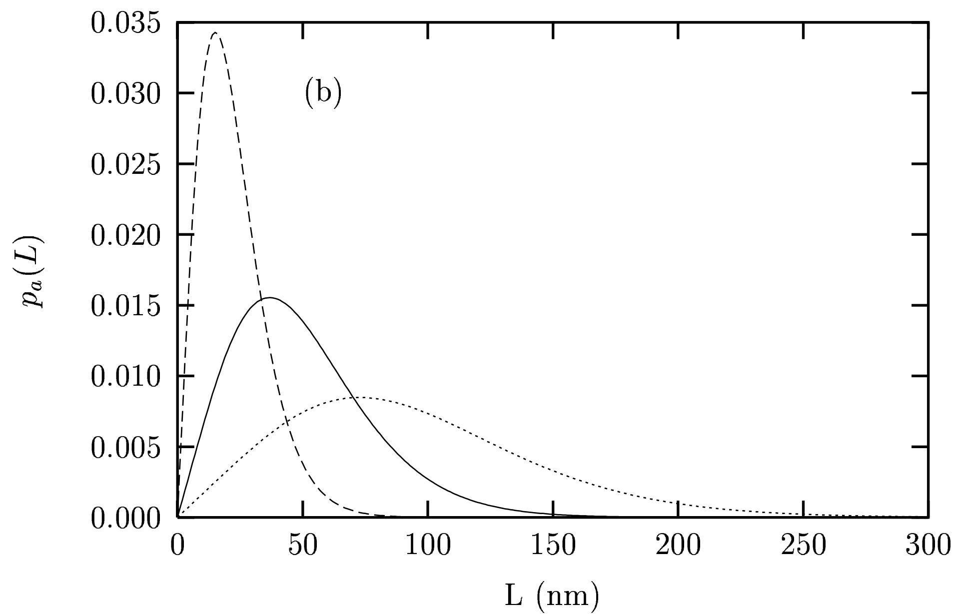


Figure 4.1: Actual particle size and column-length distributions for spherical crystallites of alumina: (a) Particle size distribution for $\langle D \rangle = 20\,\text{nm}$ (dashed line), $50\,\text{nm}$ (solid line) and $100\,\text{nm}$ (dots); (b) Column-length distributions for the corresponding particle distributions. All distributions are normalized for unit area.

corresponding $\langle D \rangle$ values. Also, Figure 4.1(a) shows the particle-size distributions for $\langle D \rangle = 20, 50$ & $100\,\text{nm}$ calculated using (4.14).

Column-length distribution

For spherical crystallites, size-broadening is independent of the diffracting plane. Using (4.7), (4.8) and (4.14), the column-size distribution, $p_a(L)$ was found[2] to be,

$$p_a(L; u, t, r) = \frac{1}{Z_L} \frac{\pi\, L\, \Gamma\left(\frac{1+r}{t},\, uL^t\right)}{\Gamma\left(\frac{1+r}{t}\right)} \tag{4.16}$$

where $\Gamma(x, y)$ is the incomplete Gamma function. The normalization term in (4.16), Z_L, was determined numerically. Figure 4.1(b) shows the column length distributions for spherical crystallites with $\langle D \rangle = 20,\ 50\,\&\,100\,\mathrm{nm}$.

From (4.16), the theoretical values for the area-weighted, $\langle L \rangle_a^{Th}$, and volume weighted, $\langle L \rangle_v^{Th}$, sizes were determined. For $\langle L \rangle_a^{Th}$, the usual result, (4.13) was used. For $\langle L \rangle_v^{Th}$, $p_v(L)$ was first determined from (4.4), and $\langle L \rangle_v^{Th}$ was found in a similar manner to (4.13). Table 4.1 also shows the values for $\langle L \rangle_a^{Th}$ and $\langle L \rangle_v^{Th}$ for different distributions and $\langle D \rangle$values. Appendix B discusses the effect the particle distribution, $P(D)$, has on $\langle L \rangle_a$ and $\langle L \rangle_v$ for a given crystallite shape.

$\langle D \rangle$ (nm)	20	50	100
$u,\ t,\ r$	0.15, 1.00, 2.00	0.08, 1.00, 3.00	0.05, 1.00, 4.00
$\langle L \rangle_a^{Th}$ (nm)	22.2	50	93.3
$\langle L \rangle_v^{Th}$ (nm)	30	65.6	120

Table 4.1: Average diameters, $\langle D \rangle$, determined using the u, t & r parameters in (4.14); the theoretical area- & volume-weighted sizes, $\langle L \rangle_a^{Th}$ & $\langle L \rangle_v^{Th}$ respectively, determined from (4.16), are also shown.

Size-broadened profiles

Modelling of the size-broadened profiles requires the discretisation of terms in (4.1), by dividing the column-length, L, into n unit cells,

$$L = na \tag{4.17}$$

[2]The integration in (4.7) was performed using *Mathematica 3.0*, Wolram Research Inc. 100 Trade Center Drive Champaign, Illinois, 61820-7237.

where n defines the harmonic order and is the number of unit cells parallel to the diffracting plane which makes up the column, and a is the dimension of the unit cell which defines the step-size in Fourier space. The column-length distribution, (4.6), was evaluated in "column-length space" (L-space), and transformed into "unit cell space" (n-space), by using the relationship in (4.17). The consequence of this is that the discretised column-length distribution can be thought of as an N-dimensional column vector. Hence, the scattering kernel is evaluated as an $[M \times N]$ matrix in n-space,

$$K(s - s_0,\, n) = \frac{\sin^2 \pi\, a\, n\, (s - s_0)}{\sin^2 \pi\, a\, (s - s_0)} \tag{4.18}$$

where the reciprocal space variable $s - s_0 = \frac{2}{\lambda}\,(\sin\,\theta - \sin\,\theta_B)$ was divided into M-intervals over the range of $(2\theta_B \pm 5)°2\theta$ at a step-size of $0.01°2\theta$; $a = \frac{\lambda}{2}\,(\sin\theta_2 - \sin\theta_1)^{-1}$; and $n = 0,\, 1,\, 2,\, \ldots,\, N-1$. Figure 4.2, shows the true specimen profiles (normalized) over the interval $(2\theta_B \pm 1)\,°2\theta$ and the Fourier coefficients, corresponding to $\langle D \rangle = 20,\ 50\,\&\,100\,\text{nm}$.

4.3.2 Simulated instrument profiles

The simulation of instrument profiles in this study was based on a LaB_6 standard reference material. The optics of the diffractometer and reference material were modelled using the *XFIT* program (Cheary & Coelho 1992, Cheary & Coelho 1996). The *XFIT* parameters used were: slit length = 15 mm; slit width = 0.1 mm; divergence angle = 1°; linear absorption coefficient = $500\,\text{cm}^{-1}$; target size = 0.04 mm; axial sample length = 15 mm; axial source length = 15 mm; primary Soller angle = 2.3°; secondary Soller angle = 2.3°; radiation source = $CuK_{\alpha 1a} + CuK_{\alpha 1b}$. Three sets of LaB_6 profiles (each numbering 22 peaks) were simulated, including an ideal instrument reference profile with no microstrain or crystallite size broadening, and two others consisting of size broadening from spherical crystallites of 100 nm and 500 nm, respectively. The peaks in each set of instrument profiles were fitted with a split Pearson-VII function; the defining parameters w and m were fitted to a low-order polynomial for each.

The reference instrument profiles were evaluated in 2θ-space over the interval $(2\theta_B \pm 5)\,°2\theta$ at a step-size of $0.01\,°2\theta$. Instrument profiles with residual size broadening (i.e. non-ideal instrument profiles) were used to examine the effect on subsequent calculations.

Figure 4.3 shows the ideal and non-ideal instrument profiles for the 113 and 226 peaks over a range $(2\theta_B \pm 1)\,°2\theta$.

4.3.3 Modelling of "observed" profiles

Simulation of the "observed" alumina 113 and 226 profiles required the evaluation of (4.9) using a fast Fourier algorithm over the range in $(2\theta_B \pm 5)\,°2\theta$ at a step-size of $0.01°2\theta$. The intensities for the 113 and 226 peaks were scaled to reflect the relative intensities expected from their structure factors and Poisson noise was imparted on the convolved product. The background level was modelled on diffuse thermal and air scattering over the range of each peak, $(2\theta_B \pm 5)\,°2\theta$ at a step-size of $0.01°2\theta$. The result from Larson & von Dreele (1994) was used,

$$b = B_1 + \sum_j \left(B_{2j} \frac{Q^{2j}}{j!} + B_{2j+1} \frac{j!}{Q^{2j}} \right)$$

where $Q = 2\pi/d$, d is the interplanar spacing over the 2θ-range and the B_j coefficients were modelled from the experimental data used by Kalceff et al. (1995). In order to reflect the statistical nature of the background, Poisson noise was imposed on the calculated background; the convolved product (with noise) was added to this in order to produce the simulated diffraction profiles, shown in Figure 4.4 for the $\langle D \rangle = 20\,\text{nm}$ specimen.

To simulate the practical situation, where the available range may be restricted by neighbouring profiles, the simulated profiles for the 113 and 226 peaks were truncated to $(2\theta_B \pm 1)\,°2\theta$. Similarly, the instrument profiles used in the deconvolution were only determined over the interval $(2\theta_B \pm 1)\,°2\theta$, also at a step-size of $0.01°2\theta$.

4.4 Computational Results and Discussion

In this section, we evaluate the MaxEnt method for deconvolving the simulated profile and, in the case of size-broadened profiles, for determining the column-length distribution and the average area-weighted size. This is done for the observed profiles corresponding to specimens with $\langle D \rangle = 20,\ 50\,\&\,100\,\text{nm}$.

We begin by describing some preliminary calculations, including the procedures for determining the background level and uncertainties in the solution specimen profile, column-length

distribution and apparent-size results.

4.4.1 Preliminary calculations

Determining the background level

The simulated profiles were corrected for background according to (4.10). The background level was determined iteratively by visual inspection of the Fourier coefficients from the solution profile. First, a linear estimate was made over the end-points of the simulated (observed) profile, g, defined over the truncated interval; this was subtracted from g, and the MaxEnt method was applied to determine the solution profile. The Fourier coefficients were examined for any "distortions" arising from under- or over-estimation. By "distortions", we mean the effects as discussed in §4.2.1. As suggested by Delhez et al. (1982), it was found that plotting the logarithm of the observed profile gave a very good indication of the profile regions affected by the background estimation.

This method of background estimation relies on a deconvolution method that produces well-defined solution profiles, with no corruption of the Fourier coefficients. The MaxEnt method was found to be ideally suited for this.

Determining the specimen profile and column distribution

The Skilling & Bryan (1984) algorithm was programmed using *Mathematica 3.0*. The basic code was taken from Shaw & Tigg (1994) and extended to meet the full requirements discussed by Skilling & Bryan (1984). Determining the specimen profile with the MaxEnt method simply required determining the instrument profile for the corresponding observed profile and the variance (see §4.4). The instrument profile was normalized for unit amplitude and used as the kernel. The trial distribution was defined using a fast Fourier algorithm, with the deconvolution of the observed profile carried out in 2θ-space. The C_{aim} value was taken to be equal to the number of data points in the observed profile, M. In most cases, the MaxEnt method converged onto the preferred C_{aim} and was halted by a preset δ-value (see Chapter 3). The algorithm usually took less than 50 iteration to converge. It was found that the ability of the MaxEnt method to converge onto the preferred C_{aim} depended on the noise level in the observed profiles. After the specimen profile was determined using

the MaxEnt method, its results were transformed from 2θ-space into $(s - s_0)$-space, so that the column-length distribution could be determined. The solution specimen profile, $\hat{f}$, was compared with the true specimen profile, f, by determining R_f given by

$$R_f = \left[\frac{\sum_i (f_i - \hat{f}_i)^2}{\sum_i f_i^2} \right]^{\frac{1}{2}} \times 100\% \tag{4.19}$$

The MaxEnt method was applied again to calculate the column-length distribution, using a matrix kernel as described in (4.18) (see §4.2.2). The resulting column distribution was then used to determine the uncertainties (see below), before being transformed from n-space into L-space, (where L is the physical length of the columns).

The theoretical area- and volume-weighted sizes were compared with the calculated ones by determining $diff$, their percentage difference,

$$diff = \left| 1 - \frac{\langle L \rangle^{cal}}{\langle L \rangle^{Th}} \right| \times 100\% \tag{4.20}$$

Determining the variance and uncertainties

Deconvolution of the observed profiles requires the variance of the noise in the statistic function, (3.21). This was determined using the result that for large counts (>> 10), the Poisson distribution can be approximated by a Gaussian distribution, which implies $\sigma_i^2 = \tilde{g}_i$ for $i = 1, 2, 3, \ldots, M$. After the MaxEnt method had converged onto a solution, the uncertainties in the solution profile were determined in the manner described in §3.3.5. The region of interest in (3.58) consisted of ones over f_{i-1}, f_i and f_{i+1} and zeros elsewhere, and similarly for determining the uncertainties in $\hat{p}_a$.

The un-normalized solution profile and its corresponding uncertainties were used in the MaxEnt method again, to determine the column-length distribution and its uncertainties. The uncertainties in the solution profile, $\hat{f}$, were transformed into variances, $\sigma_{\hat{f}}^2$, and used in the statistic function, $C(p)$, given by (4.12). The uncertainties in the solution column-distribution were determined in the same manner as the uncertainties for the solution profile, while those in the integral-breadth result, $\langle L \rangle_v^I$, were found from the error in the integral breadth in (4.5).

In order to determine the uncertainties in $\langle L \rangle_a^F$, a set of profiles were randomly drawn from the uncertainty region of $\hat{f}$ using a Gaussian distribution. From each profile, the

Fourier coefficients were calculated and the quantity $\langle L\rangle_a^F$ determined. From the set of $\langle L\rangle_a^F$ quantities, the standard error in the mean was calculated and quoted as the error in the quantity. This approach ensured that the quoted error represents the $\pm 1\sigma$ region of a Gaussian distribution. A similar procedure was followed to calculate the uncertainty in $\langle L\rangle_a^M$.

4.4.2 Using the MaxEnt method to determine $\hat{f}(2\theta)$, $\hat{A}(L)$, $\hat{p}_a(L)$ and size results.

In this section, the MaxEnt method is applied to determine the solution profile, Fourier coefficients and column-distribution for the best background level estimate. These calculations are performed on the 113 and 226 simulated profiles corresponding to $\langle D\rangle = 20\,\mathrm{nm}$. In applying the MaxEnt method, the ideal instrument profile was assumed to be known. Both the simulated and instrument profiles were truncated to $(2\theta \pm 1)^{\circ}2\theta$. The effect of this truncation on the column distribution is presented and it is shown how this can be solved when a suitable *a priori* model is chosen.

Deconvolving the 113 & 226 peaks

The results from applying the MaxEnt method for deconvolving the 113 and 226 simulated profile corresponding to $\langle D\rangle = 20\,\mathrm{nm}$ are given in Figures 4.5 & 4.6, respectively. A summary of the R_f-values and apparent size results using an ideal instrument profile are also given in Tables 4.2(a1) & 4.3(a1). In Figures 4.5(a) & 4.6(a), the solution (specimen) profiles are given. From the 113 and 226 solution profiles the following can be determined: the integral breadth and volume-weighted size, given by (4.5), Fourier coefficients (see Figures 4.5(b) & 4.6(b)) and area-weighted size, given by (4.2).

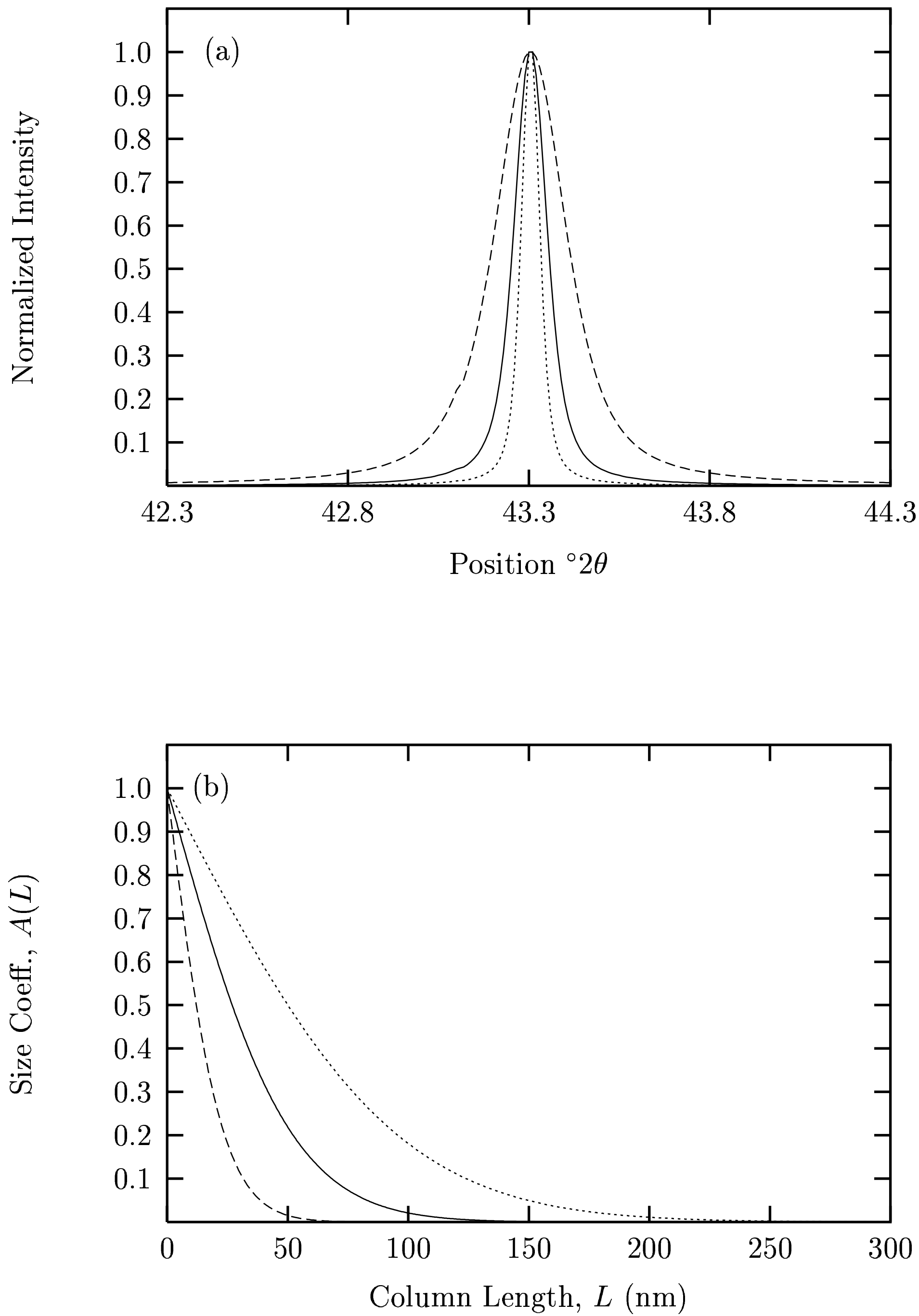
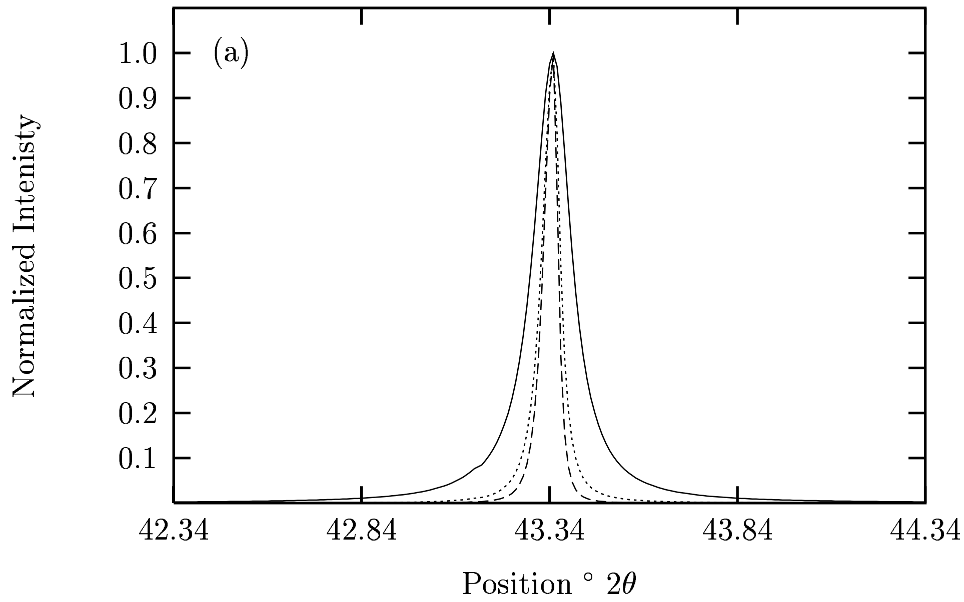


Figure 4.2: Simulated size-broadened profile and theoretical Fourier coefficients: (a) Size-broadened profile for the 113 peaks, corresponding to particle sizes $\langle D \rangle = 20\,\mathrm{nm}$ (dashed line), 50 nm (solid line) and 100 nm (dots), over the interval $(2\theta_B \pm 1)\,°2\theta$. The 226 has similar profiles which are not shown; (b) Theoretical Fourier coefficients for the 113 & 226 profiles, corresponding to particle sizes $\langle D \rangle = 20\,\mathrm{nm}$ (dashed line), 50 nm (solid line) and 100 nm (dots).



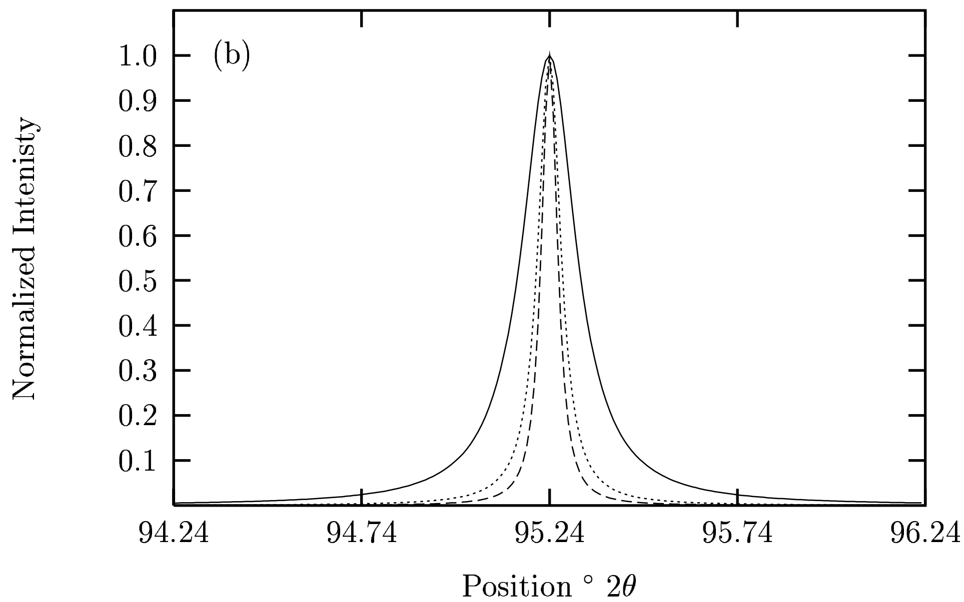
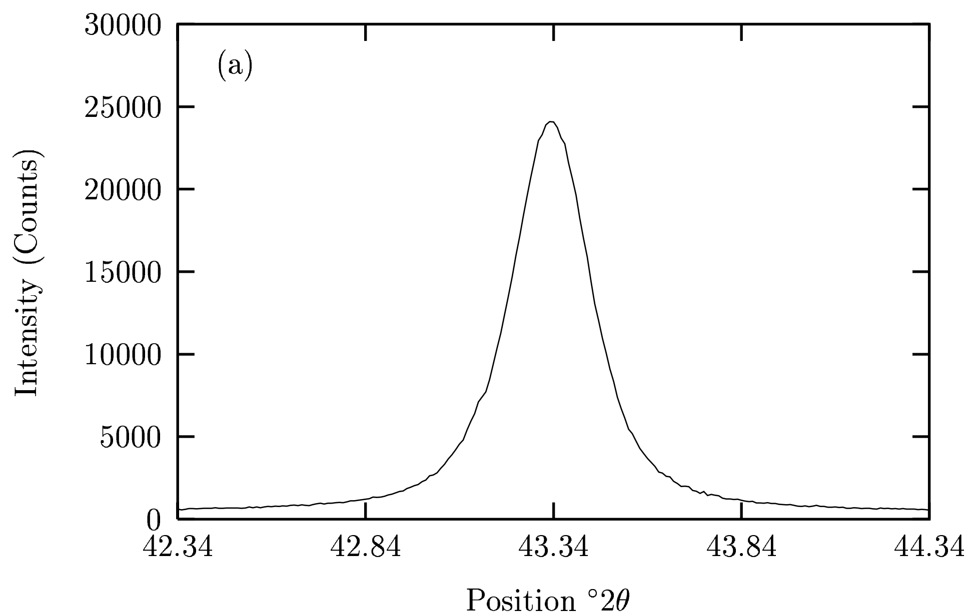


Figure 4.3: Ideal and non-ideal instrument profiles for the 113 & 226 peaks in alumina: (a) The non-ideal instrument profiles for the 113 peak with 100 nm (solid line) and 500 nm (dots) size-broadening, respectively, and the ideal case (dash line); (b) The non-ideal instrument profiles for the 226 peak with 100 nm (solid line) and 500 nm (dots) size-broadening, respectively, and the ideal case (dashed line).



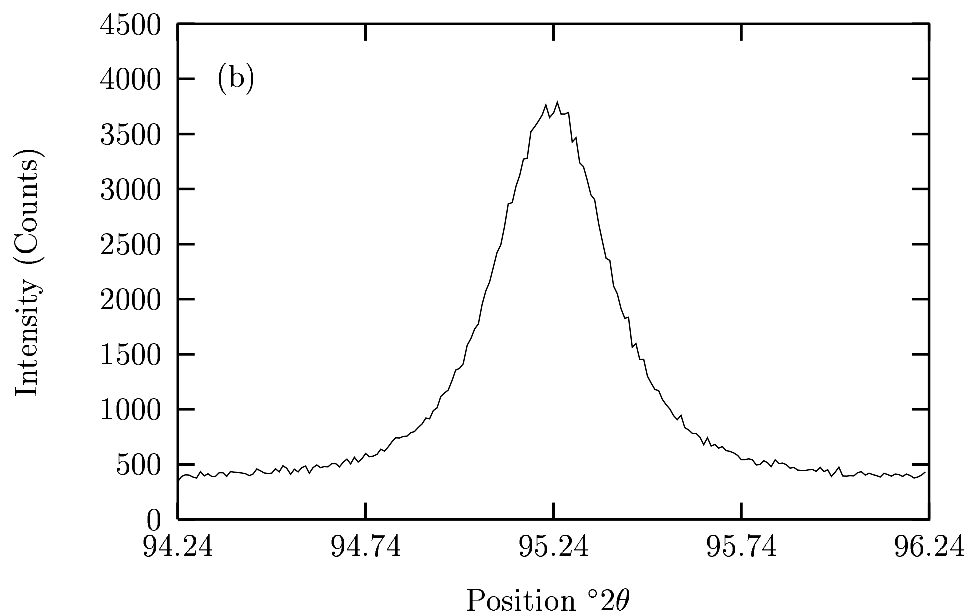
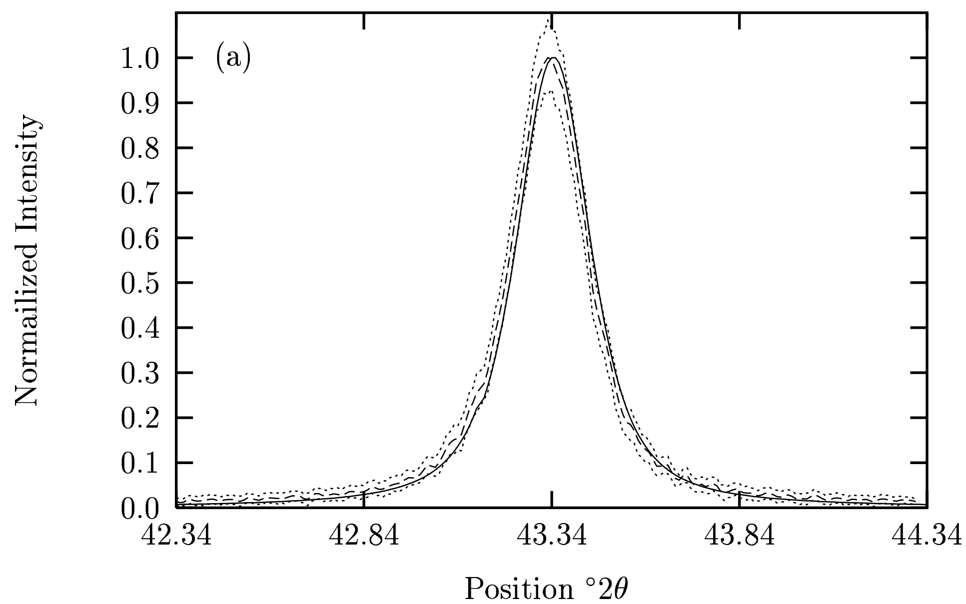


Figure 4.4: The 113 and 226 simulated ("observed") profiles corresponding to $\langle D \rangle = 20\,\text{nm}$ over the interval of $(2\theta \pm 1)\,°2\theta$: (a) The 113 profile; (b) The 226 profile. The observed profiles for $\langle D \rangle = 50\,\&100\,\text{nm}$ are not shown.



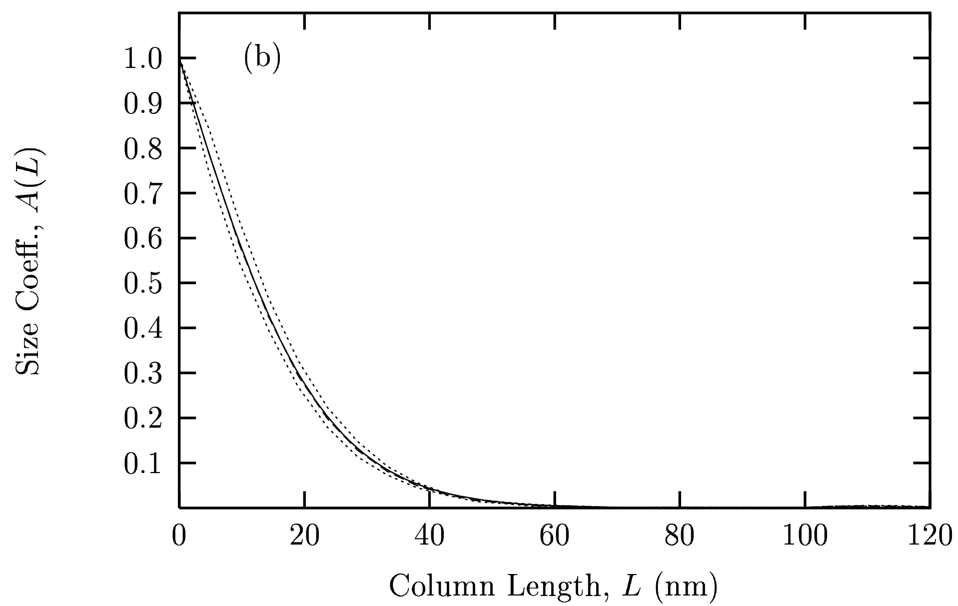
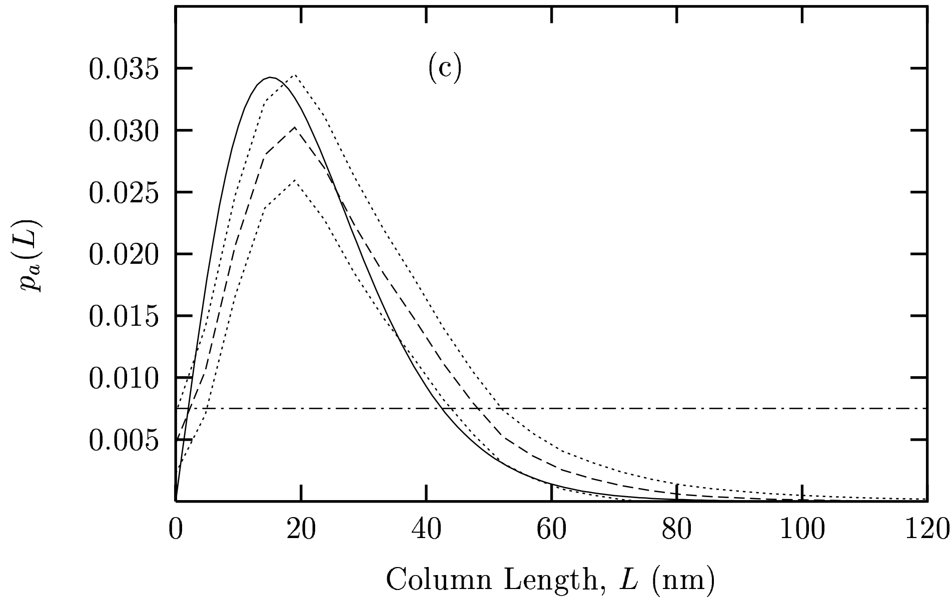


Figure 4.5 continues over...



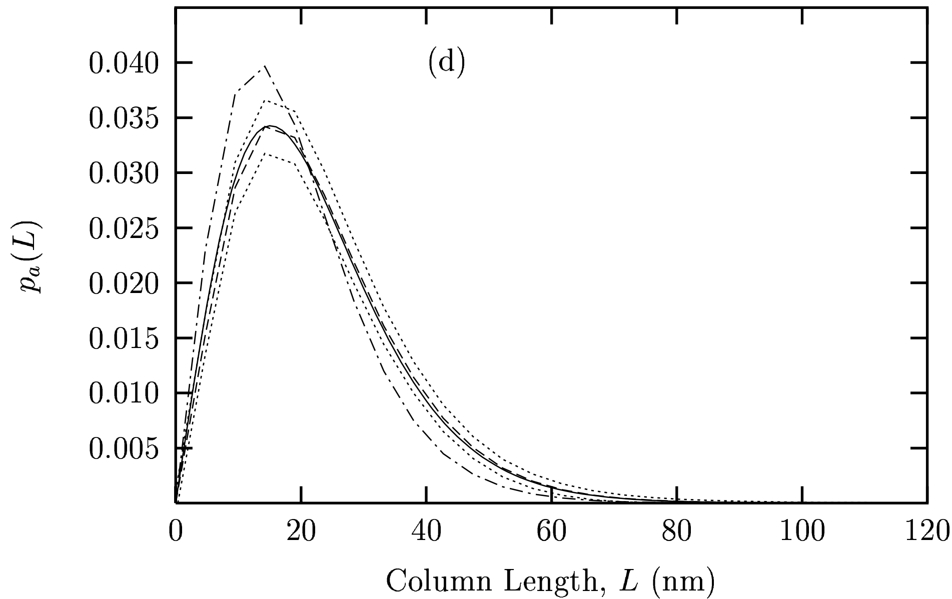
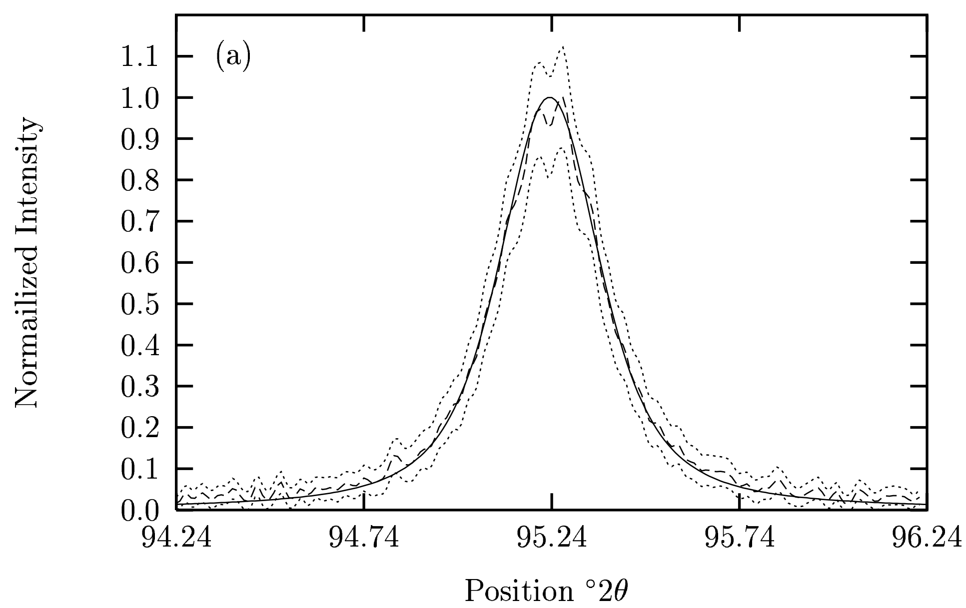


Figure 4.5: MaxEnt results for the optimum background level for the 113 peak corresponding to $\langle D \rangle =$ 20 nm: (a) True 113 specimen profile (solid line), solution specimen profile (dashed line) and lower- & upper-uncertainty regions (dots); (b) Theoretical Fourier coefficients (solid line), Fourier coefficients from the solution profile (dashed line), given in (a), and lower- & upper-uncertainty regions (dots); (c) True column-length distribution (solid line), the solution column-length distribution (dashed line), lower- & upper-uncertainty regions (dots), and the uniform *a priori* model (dash-dot lines); (d) As for (c), but using a non-uniform *a priori* model (dash-dot line). Note: The distributions showing the lower- & upper-uncertainty regions (dots) are not normalized for unit area, while the solution column length distribution and non-uniform *a priori* model are normalized for unit area.

We note that the positivity of each solution profile has been preserved and solutions are resonably well-conditioned. It is interesting to compare the 113 and 226 simulated profiles. The 113 profile is a strong peak, while the 226 profile is much weaker, and consequently has a greater fractional noise level. For example, the 226 profile corresponds to a noise level of $\sim 1.6\%$, while the noise level for the 113 peak is $\sim 0.6\%$. Also, the relatively low number of counts in the 226 profile means the effect of background will be more pronounced. The relatively high 2θ-position also means that the instrumental broadening will be more pronounced. With these factors in mind, some spurious oscillations in the 226 solution profile are noticeable. It should also be noted that an *a priori* model has not been used for the deconvolution of the 113 and 226 simulated profiles. It is expected that if such a model were used (with experimental data), for instance by taking information from other sources, such as Rietveld refinement, then a better result would be obtained.

C/C_{aim} values are shown for the MaxEnt deconvolution and column-length distributions in Tables 4.2 & 4.3. In most cases $C/C_{aim} \sim 1$ for determining $\hat{f}$ and p_a (e.g. see Tables 4.2(a1) & 4.3(a1)). For the cases were $C/C_{aim} >> 1$ (i.e 2.5 to 6.0) there are a number of possible contributing factors, such as incorrect instrument profile, noise-level, background estimation and truncations effects (e.g. see Tables 4.3(b2)).

In addition to the solution profiles given in Figures 4.5(a) & 4.6(a), the uncertainty regions for the solution profiles are given. The uncertainties for the solution profile, $\hat{f}$, were used to determine the uncertainties in the subsequent calculations, such as the Fourier coefficients, column-length distribution and the size results (see parts (b), (c) & (d) in Figures 4.5 & 4.6, & Tables 4.2(a1) & 4.3(a1)).



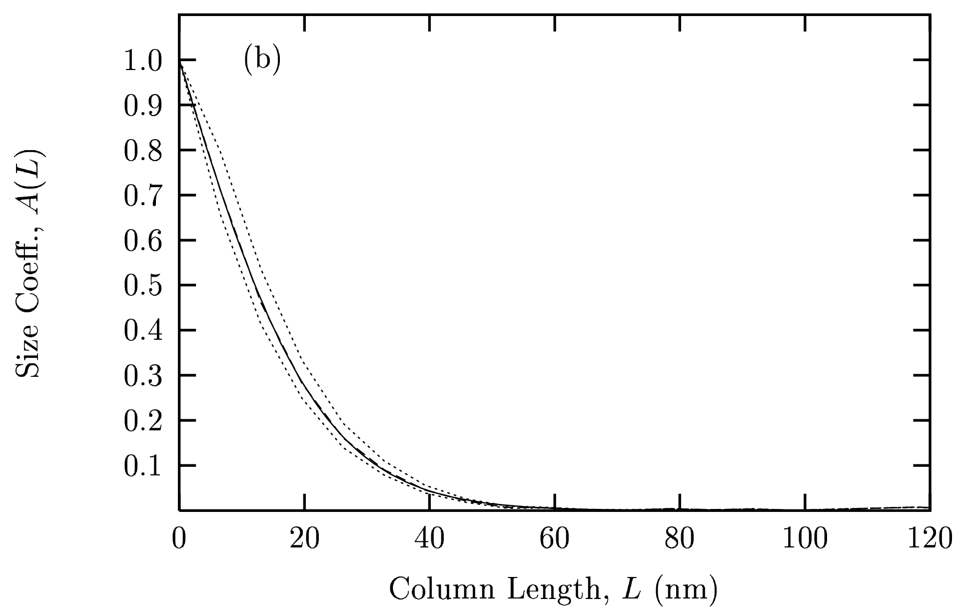
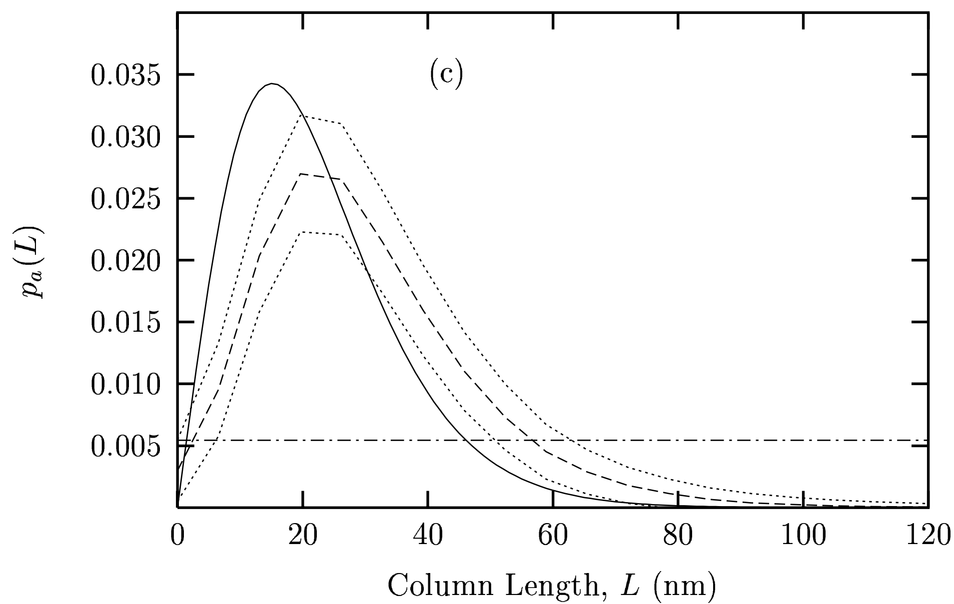


Figure 4.6 continues over...



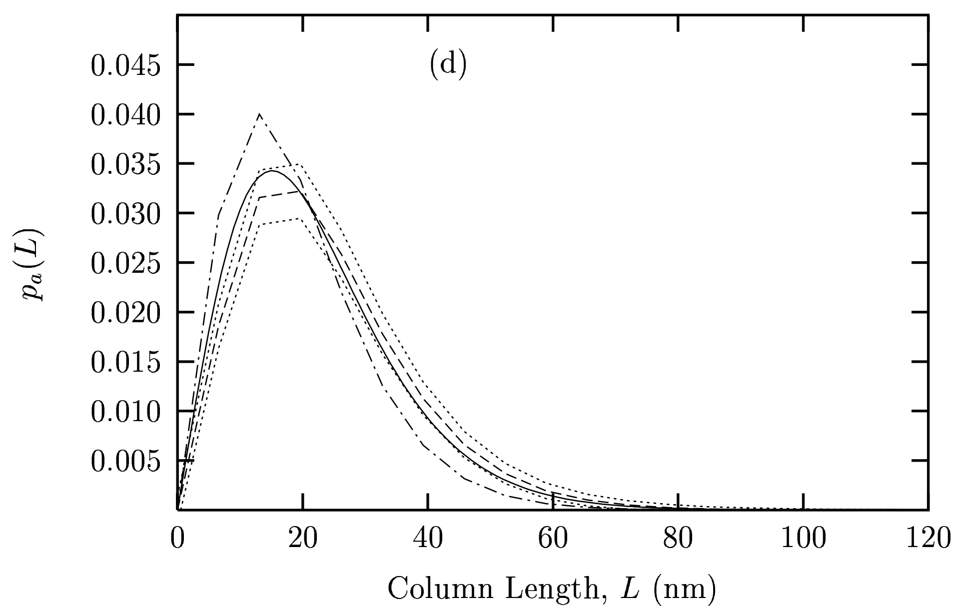


Figure 4.6: MaxEnt results for the optimum background level for the 226 peak corresponding to $\langle D \rangle =$ 20 nm: (a), (b), (c), & (d) as in Figure 4.5.

Determining the column-length distribution

Applying the MaxEnt method again to the 113 and 226 solution profiles, the solution column distribution can be determined (see Figures 4.5(c) & 4.6(c)). For a uniform *a priori* model

Specimen prof.	1. L_aB_6 ref. instr. profile	2. $L_aB_6 + 500\,\mathrm{nm}$ instr. prof.	3. $L_aB_6 + 100\,\mathrm{nm}$ instr. prof.
(a) *spherical cryst.* $\langle D\rangle = 20\,\mathrm{nm}$ $\langle L\rangle_a^{Th} = 22.2\,\mathrm{nm}$ $\langle L\rangle_v^{Th} = 30.0\,\mathrm{nm}$	$R_f = 6.3\%$ $C(f)/C_{aim} = 1.0$ $\langle L\rangle_v^I = (30 \pm 4)\,\mathrm{nm}$ $diff < 0.5\%$ $\langle L\rangle_a^F = (22.9 \pm 0.3)\,\mathrm{nm}$ $diff = 3.2\%$ $\langle L\rangle_a^{M,u} = (26.6 \pm 0.3)\,\mathrm{nm}$ $diff = 19.7\%$ $C(p_a)/C_{aim} = 1.0$ $\langle L\rangle_a^{M,n} = (22.7 \pm 0.1)\,\mathrm{nm}$ $diff = 2.0\%$ $C(p_a)/C_{aim} = 1.0$	$R_f = 7.7\%$ $C(f)/C_{aim} = 1.0$ $\langle L\rangle_v^I = (32 \pm 5)\,\mathrm{nm}$ $diff = 8.1\%$ $\langle L\rangle_a^F = (24.6 \pm 0.3)\,\mathrm{nm}$ $diff = 10.8\%$ $\langle L\rangle_a^{M,u} = (29.4 \pm 0.3)\,\mathrm{nm}$ $diff = 32.1\%$ $C(p_a)/C_{aim} = 1.0$ $\langle L\rangle_a^{M,n} = (23.2 \pm 0.1)\,\mathrm{nm}$ $diff = 4.2\%$ $C(p_a)/C_{aim} = 1.0$	$R_f = 26.3\%$ $C(f)/C_{aim} = 1.0$ $\langle L\rangle_v^I = (45 \pm 13)\,\mathrm{nm}$ $diff = 50.6\%$ $\langle L\rangle_a^F = (37 \pm 2)\,\mathrm{nm}$ $diff = 66.3\%$ $\langle L\rangle_a^{M,u} = (53.0 \pm 0.4)\,\mathrm{nm}$ $diff = 141.3\%$ $C(p_a)/C_{aim} = 1.0$ $\langle L\rangle_a^{M,n} = (24.4 \pm 0.2)\,\mathrm{nm}$ $diff = 9.7\%$ $C(p_a)/C_{aim} = 1.0$
(b) *spherical cryst.* $\langle D\rangle = 50\,\mathrm{nm}$ $\langle L\rangle_a^{Th} = 50.0\,\mathrm{nm}$ $\langle L\rangle_v^{Th} = 65.6\,\mathrm{nm}$	$R_f = 15.0\%$ $C(f)/C_{aim} = 1.0$ $\langle L\rangle_v^I = (68 \pm 23)\,\mathrm{nm}$ $diff = 4.7\%$ $\langle L\rangle_a^F = (51 \pm 3)\,\mathrm{nm}$ $diff = 1.3\%$ $\langle L\rangle_a^{M,u} = (75.6 \pm 1.2)\,\mathrm{nm}$ $diff = 51.3\%$ $C(p_a)/C_{aim} = 1.0$ $\langle L\rangle_a^{M,n} = (51.0 \pm 0.5)\,\mathrm{nm}$ $diff = 1.3\%$ $C(p_a)/C_{aim} = 1.0$	$R_f = 19.6\%$ $C(f)/C_{aim} = 1.35$ $\langle L\rangle_v^I = (80 \pm 32)\,\mathrm{nm}$ $diff = 22.7\%$ $\langle L\rangle_a^F = (58 \pm 4)\,\mathrm{nm}$ $diff = 16.6\%$ $\langle L\rangle_a^{M,u} = N/A$ $C(p_a)/C_{aim} = N/A$ $\langle L\rangle_a^{M,n} = (48.8 \pm 0.3)\,\mathrm{nm}$ $diff = 2.4\%$ $C(p_a)/C_{aim} = 1.0$	N/A
(c) *spherical cryst.* $\langle D\rangle = 100\,\mathrm{nm}$ $\langle L\rangle_a^{Th} = 93.3\,\mathrm{nm}$ $\langle L\rangle_v^{Th} = 120.0\,\mathrm{nm}$	$R_f = 8.1\%$ $C(f)/C_{aim} = 1.5$ $\langle L\rangle_v^I = (120 \pm 30)\,\mathrm{nm}$ $diff < 0.5\%$ $\langle L\rangle_a^F = (100 \pm 7)\,\mathrm{nm}$ $diff = 7.0\%$ $\langle L\rangle_a^{M,u} = (120 \pm 2)\,\mathrm{nm},$ $diff = 28.1\%$ $C(p_a)/C_{aim} = 1.0$ $\langle L\rangle_a^{M,n} = (97.4 \pm 0.8)\,\mathrm{nm}$ $diff = 4.4\%$ $C(p_a)/C_{aim} = 1.0$	$R_f = 19.1\%$ $C(f)/C_{aim} = 2.5$ $\langle L\rangle_v^I = (158 \pm 19)\,\mathrm{nm}$ $diff = 31.4\%$ $\langle L\rangle_a^F = (110 \pm 3)\,\mathrm{nm}$ $diff = 18.2\%$ $\langle L\rangle_a^{M,u} = N/A$ $C(p_a)/C_{aim} = N/A$ $\langle L\rangle_a^{M,n} = (141.5 \pm 0.6)\,\mathrm{nm}$ $diff = 51.6\%$ $C(p_a)/C_{aim} = 6.0$	N/A

Table 4.2: Apparent size results from the 113 solution profile obtained using different instrument profiles as kernels in the deconvolution of the simulated profile. Shown are the R_f for each solution profile, apparent sizes, $\langle L\rangle_v^I$, $\langle L\rangle_a^F$, $\langle L\rangle_a^{M,u}$ and $\langle L\rangle_v^{M,n}$, and differences between theoretical and calculated results for different particle distributions corresponding to $\langle D\rangle = 20, 50\,\&\,100\,\mathrm{nm}$. The superscripts, I, F, (M,u) and (M,n) mean Integral, Fourier, MaxEnt with uniform prior model and MaxEnt with a non-uniform prior model, respectively.

Specimen prof.	1. L_aB_6 ref. instr. profile	2. $L_aB_6 + 500\,\mathrm{nm}$ instr. prof.	3. $L_aB_6 + 100\,\mathrm{nm}$ instr. prof.
(a) *spherical cryst.* $\langle D\rangle = 20\,\mathrm{nm}$ $\langle L\rangle_a^{Th} = 22.2\,\mathrm{nm}$ $\langle L\rangle_v^{Th} = 30.0\,\mathrm{nm}$	$R_f = 5.4\%$ $C(f)/C_{aim} = 1.5$ $\langle L\rangle_v^I = (31 \pm 5)\,\mathrm{nm}$ $diff = 2.3\%$ $\langle L\rangle_a^F = (22.2 \pm 0.4)\,\mathrm{nm}$ $diff < 0.5\%$ $\langle L\rangle_a^{M,u} = (30.2 \pm 0.4)\,\mathrm{nm}$ $diff = 35.7\%$ $C(p_a)/C_{aim} = 1.0$ $\langle L\rangle_a^{M,n} = (23.7 \pm 0.1)\,\mathrm{nm}$ $diff = 6.7\%$ $C(p_a)/C_{aim} = 1.0$	$R_f = 5.1\%$ $C(f)/C_{aim} = 1.5$ $\langle L\rangle_v^I = (32 \pm 10)\,\mathrm{nm}$ $diff = 8.1\%$ $\langle L\rangle_a^F = (24.6 \pm 0.8)\,\mathrm{nm}$ $diff = 10.5\%$ $\langle L\rangle_a^{M,u} = (40.3 \pm 0.7)\,\mathrm{nm}$ $diff = 81.4\%$ $C(p_a)/C_{aim} = 1.0$ $\langle L\rangle_a^{M,n} = (22.9 \pm 0.2)\,\mathrm{nm}$ $diff = 3.4\%$ $C(p_a)/C_{aim} = 1.0$	$R_f = 24.4\%$ $C(f)/C_{aim} = 1.63$ $\langle L\rangle_v^I = (44 \pm 33)\,\mathrm{nm}$ $diff = 47.2\%$ $\langle L\rangle_a^F = (35 \pm 4)\,\mathrm{nm}$ $diff = 59.7\%$ $\langle L\rangle_a^{M,u} = N/A$ $C(p_a)/C_{aim} = N/A$ $\langle L\rangle_a^{M,n} = N/A$ $C(p_a)/C_{aim} = N/A$
(b) *spherical cryst.* $\langle D\rangle = 50\,\mathrm{nm}$ $\langle L\rangle_a^{Th} = 50.0\,\mathrm{nm}$ $\langle L\rangle_v^{Th} = 65.6\,\mathrm{nm}$	$R_f = 5.3\%$ $C(f)/C_{aim} = 2.5$ $\langle L\rangle_v^I = (66 \pm 13)\,\mathrm{nm}$ $diff = 0.7\%$ $\langle L\rangle_a^F = (48.7 \pm 1.2)\,\mathrm{nm}$ $diff = 3.1\%$ $\langle L\rangle_a^{M,u} = (78.2 \pm 0.6)\,\mathrm{nm}$ $diff = 56.4\%$ $C(p_a)/C_{aim} = 2.5$ $\langle L\rangle_a^{M,n} = (52.2 \pm 0.3)\,\mathrm{nm}$ $diff = 4.4\%$ $C(p_a)/C_{aim} = 2.5$	$R_f = 11.2\%$ $C(f)/C_{aim} = 3.5$ $\langle L\rangle_v^I = (77 \pm 15)\,\mathrm{nm}$ $diff = 17.5\%$ $\langle L\rangle_a^F = (59.8 \pm 2.0)\,\mathrm{nm}$ $diff = 19.6\%$ $\langle L\rangle_a^{M,u} = (93.3 \pm 0.7)\,\mathrm{nm}$ $diff = 86.8\%$ $C(p_a)/C_{aim} = 2.5$ $\langle L\rangle_a^{M,n} = (56.9 \pm 0.3)\,\mathrm{nm}$ $diff = 13.9\%$ $C(p_a)/C_{aim} = 2.5$	N/A
(c) *spherical cryst.* $\langle D\rangle = 100\,\mathrm{nm}$ $\langle L\rangle_a^{Th} = 93.3\,\mathrm{nm}$ $\langle L\rangle_v^{Th} = 120.0\,\mathrm{nm}$	$R_f = 4.5\%$ $C(f)/C_{aim} = 2.25$ $\langle L\rangle_v^I = (112 \pm 24)\,\mathrm{nm}$ $diff = 6.1\%$ $\langle L\rangle_a^F = (47 \pm 1)\,\mathrm{nm}$ $diff = 49.1\%$ $\langle L\rangle_a^{M,u} = (102 \pm 2)\,\mathrm{nm}$ $diff = 9.8\%$ $C(p_a)/C_{aim} = 1.0$ $\langle L\rangle_a^{M,n} = (93.1 \pm 0.6)\,\mathrm{nm}$ $diff < 0.5\%$ $C(p_a)/C_{aim} = 1.0$	$R_f = 17.8\%$ $C(f)/C_{aim} = 2.5$ $\langle L\rangle_v^I = (152 \pm 45)\,\mathrm{nm}$ $diff = 26.4\%$ $\langle L\rangle_a^F = (51 \pm 3)\,\mathrm{nm}$ $diff = 45.1\%$ $\langle L\rangle_a^{M,u} = N/A$ $C(p_a)/C_{aim} = N/A$ $\langle L\rangle_a^{M,n} = (95.8 \pm 0.6)\,\mathrm{nm}$ $diff = 2.6\%$ $C(p_a)/C_{aim} = 1.0$	N/A

Table 4.3: Apparent size results from 226 solution profile obtained using different instrument profiles as kernels in the deconvolution of the simulated profile. Shown are the R_f for each solution profile, apparent sizes, $\langle L\rangle_v^I$, $\langle L\rangle_a^F$, $\langle L\rangle_a^{M,u}$ and $\langle L\rangle_v^{M,n}$, and differences between theoretical and calculated results for different particle distributions corresponding to $\langle D\rangle = 20,\ 50\,\&\,100\,\mathrm{nm}$. Notation is as in Table 4.2.

the resulting MaxEnt area-weighted size, $\langle L\rangle_a^{M,u}$ (where the superscript (M,u) represents the MaxEnt method using a uniform *a priori* model) are considerably larger than those for $\langle L\rangle^F$ (see Tables 4.2(a1) & 4.3(a1)). The reason for this is clear from the solution distributions in Figures 4.5(c) & 4.6(c), respectively, which have been shifted or "biased" towards the larger column-lengths.

This shifting of the solution column distribution can be explained by the truncation of the simulated observed profiles. The diffraction from small columns results in the extended tails of a size-broadened profile, while diffraction from the larger columns results in the narrow region about the Bragg angle. When the observed profile is truncated, the tails of the underlying specimen profile are removed and information concerning the small columns is also eliminated, while information related to the large columns is unaffected. Consequently, once the instrument broadening has been removed and the column-length distribution determined using the MaxEnt method, the solution is biased towards the larger columns. The information provided by the uniform *a priori* model is ineffectual, since it assigns a uniform probability over the range of L and basically describes our ignorance about the columns.

The profiles used in these calculation were truncated to $(2\theta_B \pm 1)\,°2\theta$ in order to investigate the effect on the column-length distribution. In the practical case, especially for low symmetry material where there is considerable overlapping of profiles in the diffraction pattern, it may be necessary to truncate the profiles to a smaller interval than $(2\theta_B \pm 1)\,°2\theta$. Based on the above results, the column-length distribution would be biased toward larger columns and produce erroneous results in the apparent size results. In order to minimise this, it would be necessary to separate the overlapped profiles and resolve their tails over a much larger 2θ-interval. In principle, the method developed by Sivia & David (1994) could be used for this purpose.

The effects of truncation can also be reduced by introducing a non-uniform *a priori* model. The definition of the entropy function for the column distribution (4.11) allows *a priori* information to be included in the MaxEnt method. In these sets of calculations, a non-uniform model with the parameters $u = 0.175$, $t = 1.0$ and $r = 2.0$ was aribitrarily chosen and evaluated using (4.14 - 4.16). Physically, this model can be thought of as a set of spherical crystallites with $\langle D\rangle = 17.14\,\text{nm}$ and $\langle L\rangle_a = 19.05\,\text{nm}$. Using this model in

the MaxEnt method, an improvement in the solution column distribution was obtained (see Figures 4.5(d) & 4.6(d)), with the column distribution not biased towards larger columns. Also, considerable improvement in the MaxEnt size results is achieved (see Tables 4.2(a1) & 4.3(a1) compare $\langle L \rangle_a^{M,u}$, $\langle L \rangle_a^{M,n}$ and the *diff* values).

It is interesting to make some observations about the use of *a priori* models in the MaxEnt method for determining the column-length distribution. As mentioned earlier, the uniform *a priori* model is ineffectual since it does not discriminate between the smaller and larger crystallites. It simply defines our ignorance concerning the column-distribution. Additional calculations suggest that as the *a priori* model is modified to include increasing physical information about the crystallites, for example the shape of the crystallites and a suitable particle distribution (see (4.6)), the solution column distribution approaches the true column distribution and the effects of truncation on the column-length distribution and apparent size results are reduced. A carefully determined model would include the shape and common-volume using model-based methods (see Langford & Louër 1996) and applying Bayesian model selection methods (see Sivia et al. 1993) to determine a suitable particle distribution. Another approach would involve using "low resolution" methods (see Bienenstock 1963, LeBail & Louër 1978) to estimate the column-length distribution and fit this distribution to an appropriate function; however, this is dependent upon determining a suitable function and its parameters. A similar problem has been encountered in determining the size distribution in small angle scattering, and the methods used there could be adopted in x-ray profile analysis (see Müller & Hansen 1994, Müller et al. 1996).

In summary, the results from this section demonstrate that with careful consideration of the background level, the solution profile can be determined using the MaxEnt method. Moreover, the solution profile can be used in subsequent calculations to determine $\langle L \rangle_v$ and $\langle L \rangle_a$ by the traditional integral breadth and Fourier methods, respectively. More importantly, the MaxEnt method also provides a better alternative approach not only for determining an average value of the area-weighted size, but also the corresponding column-length distributions and its uncertainity region.

4.4.3 The effect of background estimation on $\hat{f}(2\theta)$, $\hat{A}(L)$, $\hat{p}_a(L)$ and size results.

We now examine the effect of background estimation on the MaxEnt results. The rationale is to assess the MaxEnt method for determining the solution specimen profile, Fourier coefficients, column-length distribution, and for ascertaining the integrity of subsequent physical quantities, such as $\langle L\rangle_v^I$, $\langle L\rangle_a^F$, $\langle L\rangle_a^M$, when the estimated background level is systematically under- or over-estimated.

In the previous section, the best background level estimate was obtained from the observed profile corresponding to spherical crystallites for $\langle D\rangle = 20\,\text{nm}$. For the calculations in this section, this best level was defined as the "zeroth-level" and the estimated background level was systematically lowered and raised relative to it. For the 113 observed profile, the estimated background level ranged from 35% below to 35% above the zeroth level, while for the 226 observed profiles, this was done over the range of -20% to $+20\%$. In these ranges, the negative values correspond to under-estimation, while the positive ones are over-estimates. The deconvolution was carried out using the MaxEnt method and the reference instrument profiles for the 113 and 226 peaks.

Effect of background estimation on the apparent size

In Figure 4.7, the R_f-values for each solution profile were determined over the range of background estimation for the 113 and 226 solution specimen profiles. While the R_f values for the 113 profile vary smoothly as the background level is increased, the factors discussed in the previous section contribute to the result for the 226 specimen profile. That is, the 226 specimen profile is a relatively weak peak and the pronounced instrumental and background effects as well as the increase in noise level all contribute to the observed R_f values in Figure 4.7. For both the 113 and 226 profiles a uniform *a priori* model was assumed. It is expected that a non-uniform *a priori* model could also be used to improve the solution profiles, as was the case with the column-length distributions.

Figure 4.8 shows the apparent size results for the integral breadth method, $\langle L,\rangle_v^I$, the Fourier method, $\langle L\rangle_a^F$, the MaxEnt method for a uniform *a priori* model, $\langle L\rangle_a^{M,u}$, and for the MaxEnt method for a non-uniform *a priori* model, $\langle L\rangle_a^{M,n}$, obtained from the 113 and 226

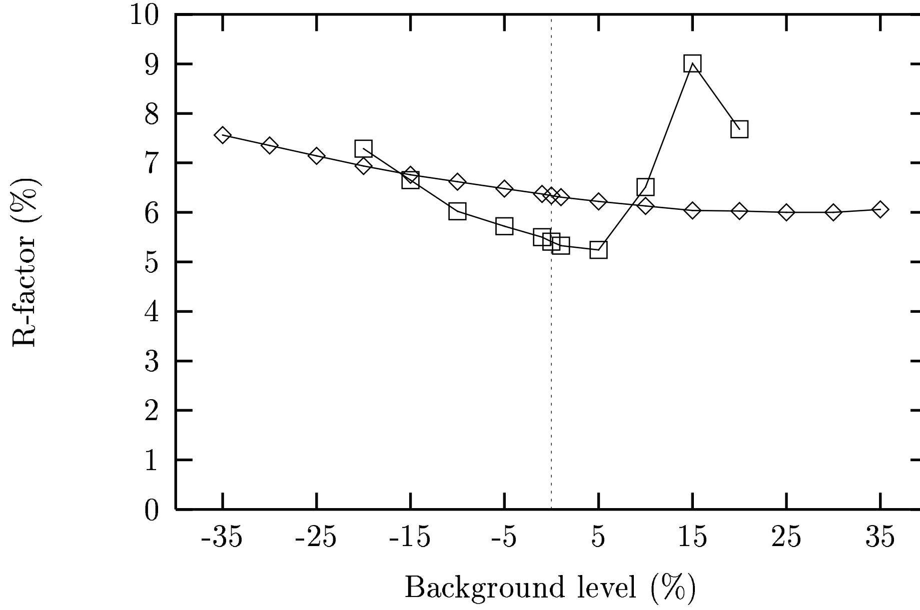


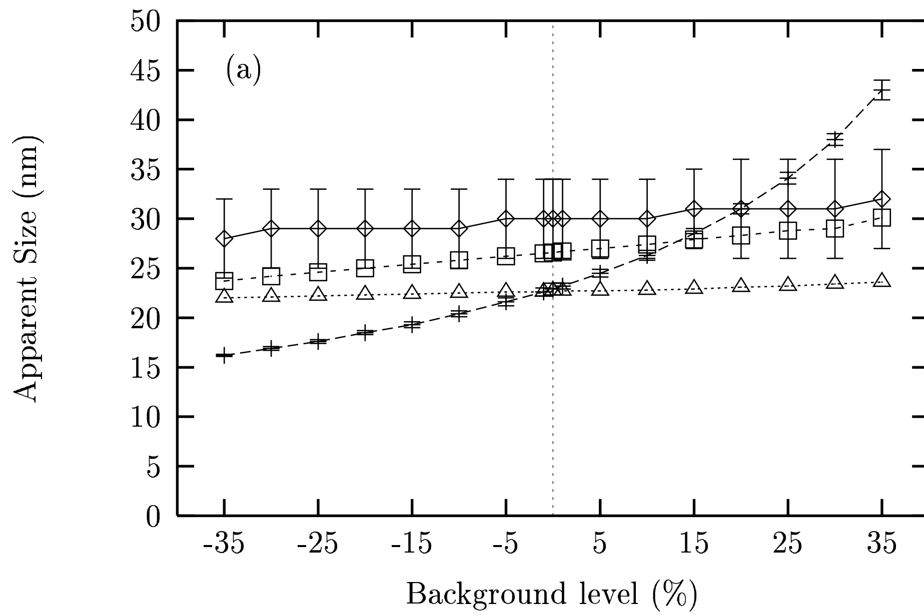
Figure 4.7: R_f-values for the 113 (diamond) and 226 (squares) solution profiles, determined from the MaxEnt method for different background levels.

simulated profiles for different background estimations. This is illustrated in Figures 4.9(a) & 4.9(b) for the 113 solution profile which show the solution profile and Fourier coefficients for an over-estimation of 20%.

For the $\langle L \rangle_v^I$ results, as the estimated background level increases, the integral breadth in the solution profile increases. This results in the $\langle L \rangle_v^I$ values increasing over this region. The relatively large errors in the $\langle L \rangle_v^I$ (about $\pm 13\%$) suggest that the systematic error in the background estimation is being transmitted to the solution profile and integral breadth. However, the difference between the theoretical value, $\langle L \rangle_v^{Th}$, and the calculated one does not go beyond about 10% (see Figures 4.8(b) & (d)).

In general, the results for the 113 and 226 peaks are similar in that they show the same trends over the range of background estimation. This can be seen in the values for $\langle L \rangle_a^F$. In the case of the 113 specimen profile results, the region of under-estimation, from -35% to 0%, the "kink" effect appears in the Fourier coefficients, which produces an under-estimation of the $\langle L \rangle_a^F$ results. That is, $\mathrm{d}A(L)/\mathrm{d}L \to \infty$ at $L = 0$, which implies that the $\langle L \rangle_a^F \to 0$ (see

(4.2)). For this region, the difference between the $\langle L \rangle_a^{Th}$ and $\langle L \rangle_a^F$ values reaches a maximum of $\sim 30\%$ (see Figure 4.8). For the region of over-estimation, from 0% to +35%, the "hook effect" appears in the Fourier coefficients, resulting in an under-estimation of $\langle L \rangle_a^F$. That is, the negative curvature in the Fourier coefficients causes $\mathrm{d}A(L)/\mathrm{d}L \rightarrow 0$ at $L = 0$, therefore $\langle L \rangle_a^F \rightarrow \infty$ (again, see (4.2)). For this region, the difference between the $\langle L \rangle_a^{Th}$ and $\langle L \rangle_a^F$ values reaches a maximum of $\sim 95\%$.



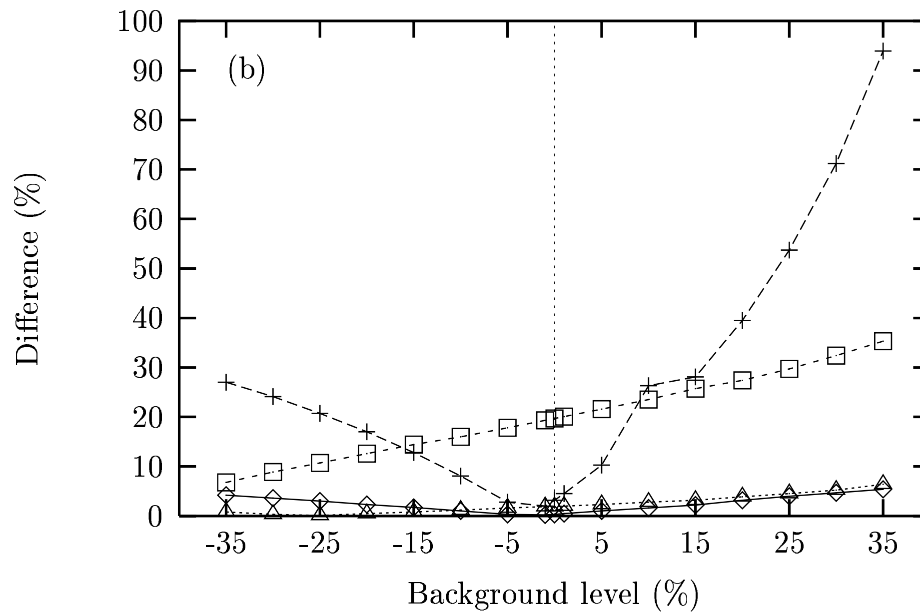
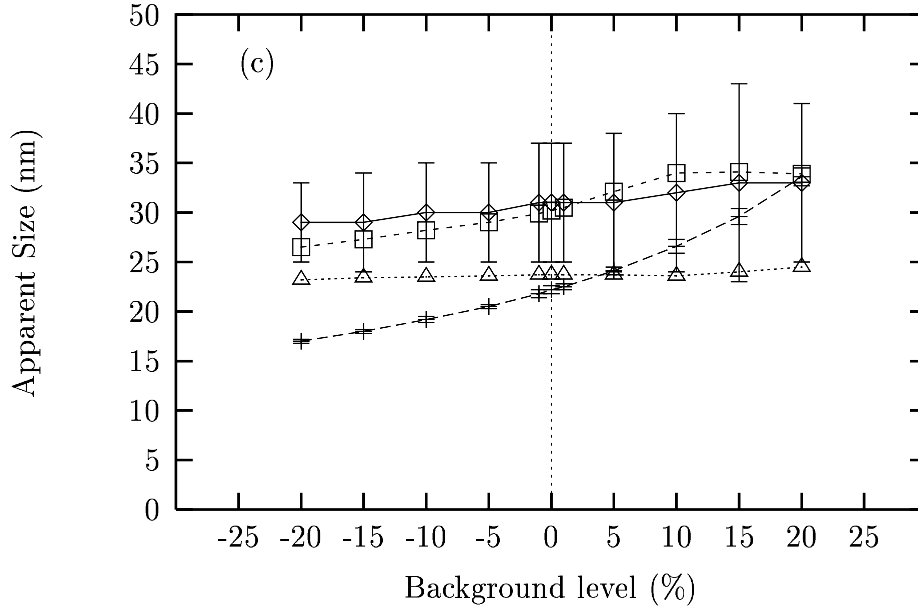


Figure 4.8 continues over. . .



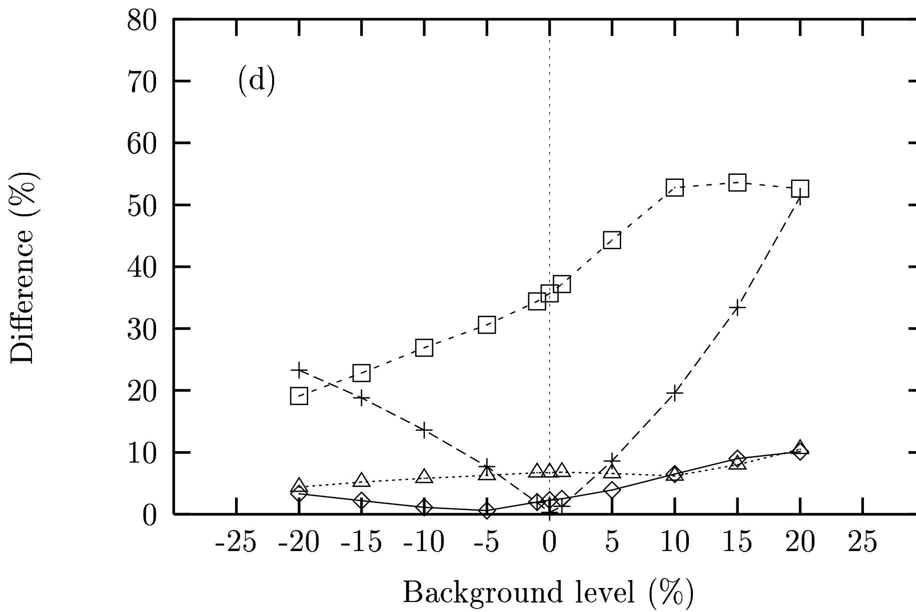
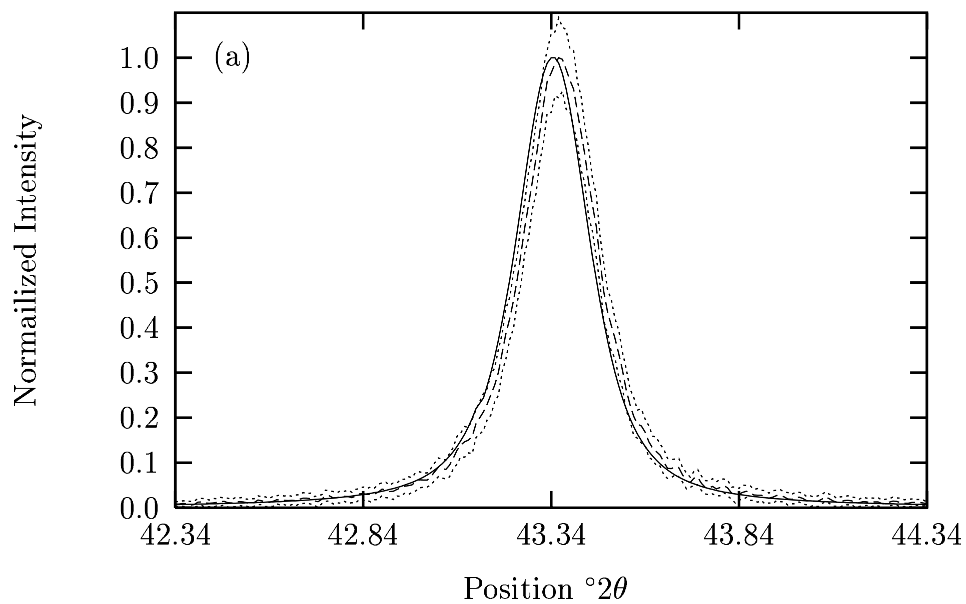


Figure 4.8: Apparent size results from different methods, determined from 113 & 226 solution profiles corresponding to $\langle D \rangle = 20\,\mathrm{nm}$ over a range of background levels: (a) Apparent sizes from the 113 solution profile using the integral breadth method $\langle L \rangle_v^I$, (diamonds + solid line), Fourier method, $\langle L \rangle_a^F$ (cross + dashed line), MaxEnt method using a uniform model, $\langle L \rangle_a^{M,u}$, (squares + small dashed line) and MaxEnt method using a non-uniform *a priori* model, $\langle L \rangle_a^{M,n}$, (triangles + dots); (b) Corresponding differences between the theoretical area- & volume-weighted sizes and apparents sizes given in (a); (c) Apparent sizes from 226 solution specimen profiles. Symbols as described in (a); (d) Showing the corresponding differences between the theoretical area- & volume-weighted sizes and apparents sizes given in (c).

For the MaxEnt method, using a uniform *a priori* model, $\langle L\rangle_a^{M,u}$, the influence of increasing background estimation from -35% to +35% is noticeable, but not as significant as in the $\langle L\rangle_a^F$ results. Generally, for increasing estimated background levels, the simulated profile is being consistently truncated and the information related to the small columns is being removed, causing the solution to be biased towards the larger columns. This can also be seen in the difference between the $\langle L\rangle_a^{Th}$ and $\langle L\rangle_a^{M,u}$ values, where the difference increases from $\sim 10\%$ at -35% background, to $\sim 30\%$ at +35% background. The results from the MaxEnt method, for a non-uniform *a priori* model, $\langle L\rangle_a^{M,n}$, are relatively constant and accurate over the range of -35% to +35%. Even though the background estimation is being increased, resulting in information related to the smaller columns being removed, the non-uniform *a priori* model prevents biasing towards the larger columns by attempting to fill in the missing information. This is clearly demonstrated in Figure 4.9(c). A similar pattern is also found in the apparent size results from the 226 solution profile over the range of -20% to +20% of the estimated background.



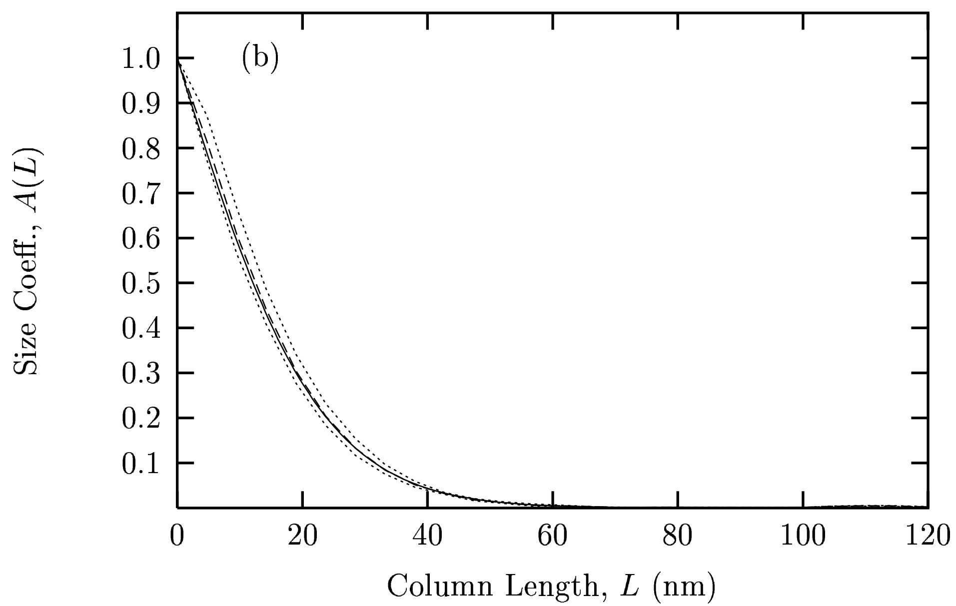
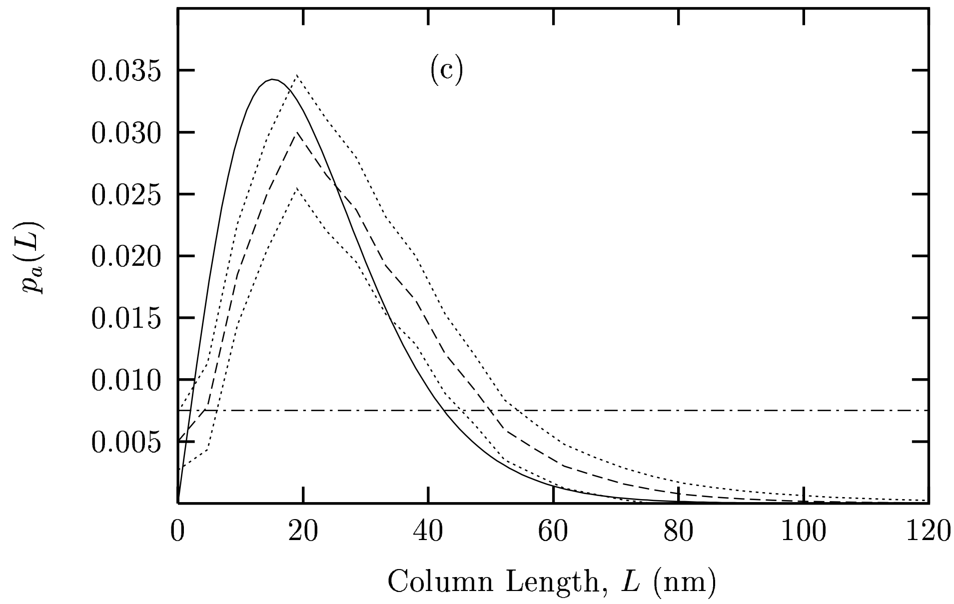


Figure 4.9 continues over...



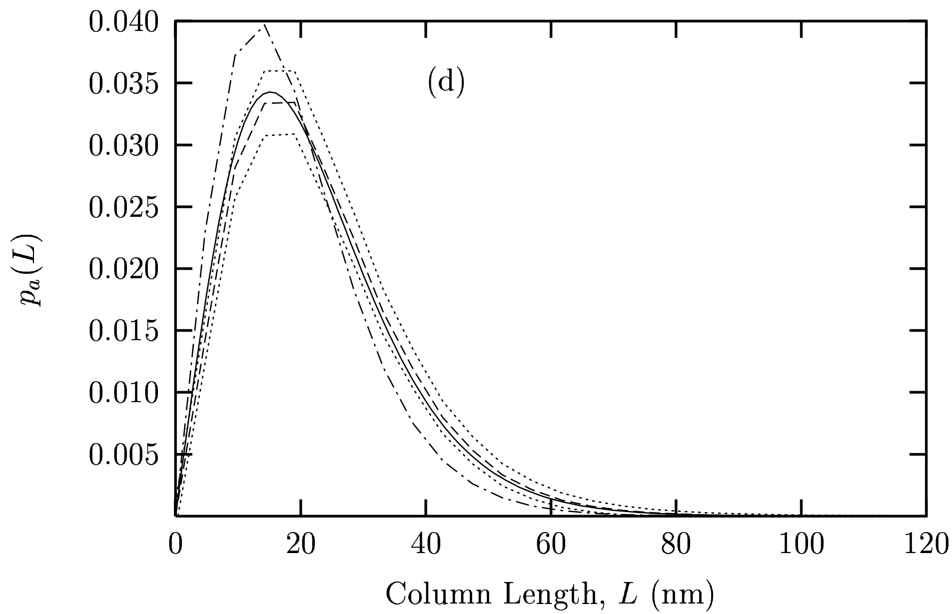


Figure 4.9: MaxEnt results for the background level of +20% for the 113 peak corresponding to $\langle D \rangle$ = 20 nm: (a) True 113 specimen profile (solid line), solution specimen profile (dashed line); (b) Theoretical Fourier coefficients (solid line), Fourier coefficients from the solution profile (dashed line), given in (a); (c) True column-length distribution (solid line), the solution column-length (dashed line) and the uniform *a priori* model (dashed line); (d) As for (c), but using a non-uniform *a priori* model (dash-dot line). Dotted line in the above represent the uncertainties in the respective MaxEnt calculations.

In summary, a set of MaxEnt calculations for varying background levels were carried out to ascertain the effect of background estimation on the solution profile, Fourier coefficients, column length distribution and apparent sizes. For deconvolving the 113 and 226 profiles and determining the specimen profile, the MaxEnt method consistently produced R_f-values below $\sim 10\%$ over the entire range of background estimation. In the case of determining the apparent sizes from the solution profile, it is clear that the error in the background estimation is also being imparted to the size results. The results from the Fourier method, $\langle L\rangle_a^F$, proved to be the most variable, followed by the MaxEnt method using a uniform *a priori* model, $\langle L\rangle_a^{M,u}$. The results from the integral breadth method, $\langle L\rangle_v^I$, are reliable, but the uncertainties in the results imply that the integral breadth is also being compromised by errors in the background estimation. This may be resolved if a suitable *a priori* model were used in the deconvolution of the simulated profile. The results from the MaxEnt method using a non-uniform *a priori* model, $\langle L\rangle_a^{M,n}$, were both reliable and accurate. This can be seen in the uniformity of the $\langle L\rangle_a^{M,n}$ values over the range of estimated background levels.

4.4.4 Effect of non-ideal instrument profiles on $\hat{f}(2\theta)$, $\hat{A}(L)$, $\hat{p}_a(L)$ and size results.

Calculations were performed to examine the effects of deconvolving an observed profile with a *size-broadened instrument profile* on the solution profile, Fourier coefficients, column-length distribution and apparent size results. The rationale for this set of calculation also arises from the experimental situation where the instrument profile is determined from a standard reference material which may have finite-sized crystallites.

Two sets of size-broadened instrument profiles were used as kernels in the MaxEnt method. Both were simulated from a LaB_6 reference material which had no microstrain and consisted of spherical crystallite shapes, whose diameters were 100nm and 500 nm, respectively (see §4.3.2). The instrument profiles were generated for the 113 and 226 positions for alumina over the range of $(2\theta_B \pm 1)\,^\circ 2\theta$. The set of observed profiles for the 113 and 226 peaks, described in §4.3.3, were truncated from $(2\theta_B \pm 5)\,^\circ 2\theta$ to $(2\theta_B \pm 1)\,^\circ 2\theta$. Different non-uniform *a priori* models were also used in determining the column-distribution corresponding to $\langle D\rangle = 50\,\&\,100\,\mathrm{nm}$, respectively. These *a priori* models had the parameters $u = 0.09$,

$t = 1.0$ and $r = 3.0$ (see (4.14) & (4.16)) for the $\langle D \rangle = 50\,\text{nm}$ and $u = 0.06$, $t = 1.0$ and $r = 4.0$ for the $\langle D \rangle = 100\,\text{nm}$. These non-uniform *a priori* models correspond to spherical crystallites with $\langle D \rangle = 44.4$ & $83.3\,\text{nm}$. Calculations using the reference instrument profiles were also performed, to provide a comparison.

Deconvolving with an ideal instrument profiles.

The (ideal) reference instrument profiles (containing no size or microstrain broadening) were used to deconvolve the simulated profiles corresponding to $\langle D \rangle = 20, 50\,\&\,100\,\text{nm}$. The results for the integral breadth, $\langle L \rangle_a^I$, Fourier methods, $\langle L \rangle_a^F$, and the MaxEnt method using a non-uniform *a priori* model, $\langle L \rangle_a^{M,n}$ compare closely with the corresponding theoretical values (see Table 4.2 (b1-c1) & 4.3 (b1-c1)). The results from the MaxEnt method using a uniform *a priori* model, $\langle L \rangle_a^{M,u}$ are disappointing, but can be accounted for by the truncation of the simulated profiles.

Deconvolving with a non-ideal instrument profile.

Unlike the ideal reference instrument profile, the size-broadened instrument profile not only characterizes the diffractometer's optics, but also the size-broadening from the finite LaB_6 crystallites. By deconvolving with a size-broadened profile, we expect the MaxEnt method to *over-compensate* for the instrumental broadening in the profiles. This over-compensation effectively results in removing information which concerns size-broadening in the underlying solution profiles, and produces a narrower solution profile relative to the true specimen profile. The consequence of the over-compensation is that the apparent sizes will, in general, increase relative to their true values. The apparent size results in Table 4.2(a2-c2) & 4.3(a2-c2) for the 500 nm size-broadened instrument profiles demonstrate this point for increasing $\langle D \rangle$. We examine, for example, the values for the $\langle L \rangle_a^F$ and $\langle L \rangle_v^I$ in Table 4.2(a2-c2) for the 113: for $\langle D \rangle = 20\,\text{nm}$, $\langle L \rangle_v^{Th} = 30\,\text{nm}$ and $\langle L \rangle_v^I = (32 \pm 5)\,\text{nm}$, an 8.1% difference, and for $\langle D \rangle = 100\,\text{nm}$, $\langle L \rangle_v^{Th} = 120\,\text{nm}$ and $\langle L \rangle_v^I = (159 \pm 19)\,\text{nm}$, a difference of 31.4%. A similar trend can also be seen for the $\langle L \rangle_a^F$ values. For the 226 line in Table 4.3(a2-c2), much the same results appear, with the exception of $\langle D \rangle = 100\,\text{nm}$, where the MaxEnt method for deconvolving and determining the column-length distribution could not converge onto the

preferred C_{aim} value.

The over-compensation in the solution specimen profile (i.e. narrower solution profiles) also shifts or biases the solution column distribution towards larger columns, in addition to the biasing that occurs due to the truncation of the simulated profile. That is, the MaxEnt method "views" the narrower profile as a result of larger columns contributing to the diffraction. Given this information, it determines the solution which produces larger $\langle L\rangle_a$ values. This can be seen in values for $\langle L\rangle_a^{M,n}$ in Table 4.3(a2-c2) using the 500 nm size-broadened instrument profile. In some cases the MaxEnt method was unable to produce a meaningful result (see parts b3 & c3 in Tables 4.2 & 4.3). This is not a problem with the MaxEnt method, but rather with the information that is being used. However, biasing towards larger columns was reduced when a non-uniform *a priori* model was used in the MaxEnt method. Generally, the use of a non-uniform *a priori* model in the MaxEnt method produced $\langle L\rangle_a^{M,n}$ values that were closer to the theoretical values, $\langle L\rangle_a^{Th}$. Figure 4.10 illustrates this point, where using a non-uniform *a priori* model improves the solution column distribution.

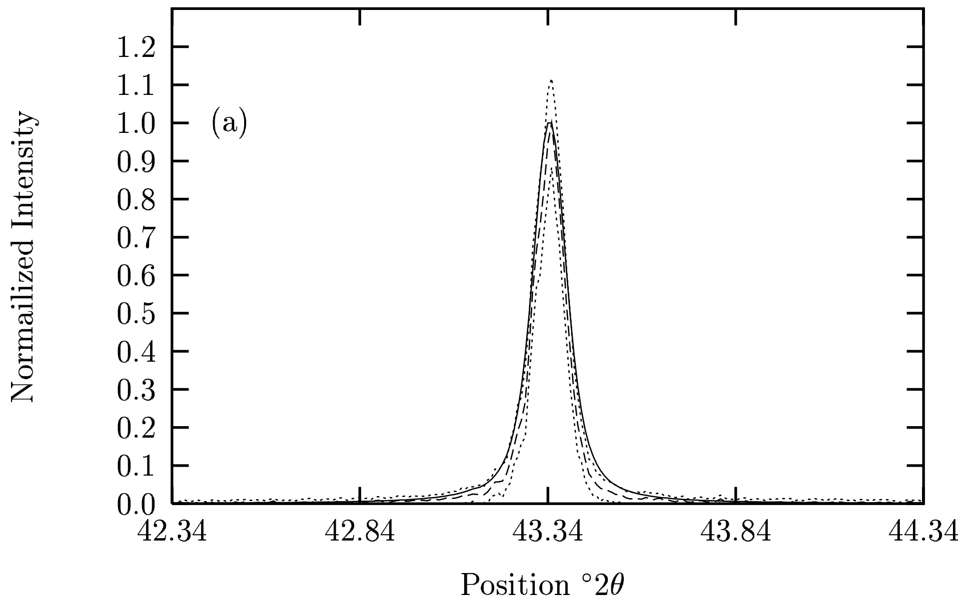


Figure 4.10 continues over...

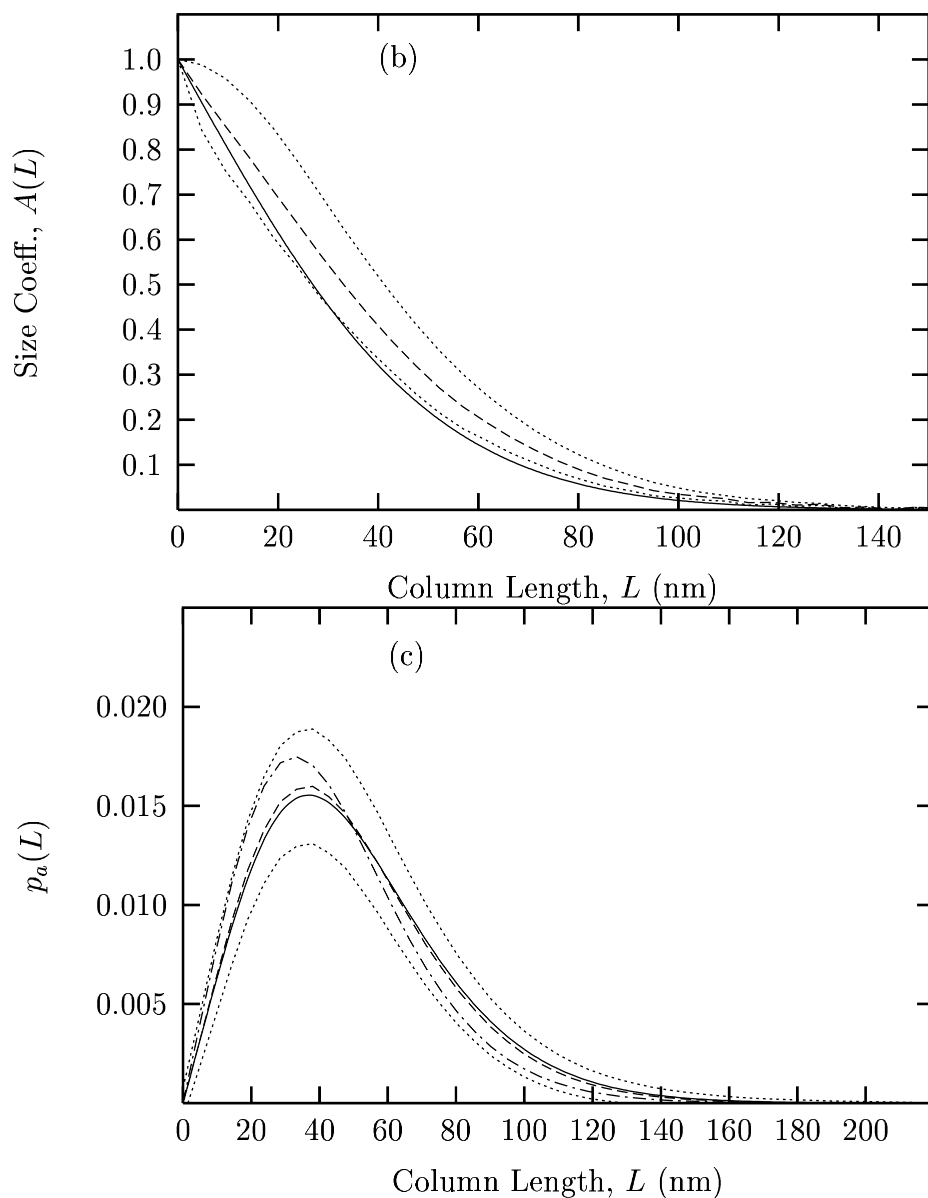


Figure 4.10: Maxent results for the 113 peak corresponding to $\langle D \rangle = 50\,\mathrm{nm}$ when a non-ideal (500 nm size-broadened) instrument profile is used in the deconvolution: (a) The true 113 specimen profile (solid line), solution specimen profile (dashed line); (b) Theoretical Fourier coefficients (solid line), Fourier coefficients from the solution profile (dashed line), given in (a); (c) True column-length distribution (solid line), the solution column-length distribution (dashed line) and the non-uniform *a priori* model (dashed-dotted line). Dotted lines in the above represent the uncertainty in the respective MaxEnt calculations.

In the case of the instrument profile with 100 nm residual-size broadening, the over-compensation of the broadening in the profile defines an upper limit in the MaxEnt method's ability to determine the solution profile and column distribution. The only case, for both the 113 and 226 profiles, in which a result was obtained using the MaxEnt method for the simulated profile corresponding to $\langle D \rangle = 20$ nm (see Table 4.2(a3) & 4.3(a3)). In the case of the simulated profiles corresponding to $\langle D \rangle = 50$ & 100 nm, the over-compensation prevents the MaxEnt method from obtaining a meaningful physical result.

In summary, deconvolving with a non-ideal instrument profile is a problem of deconvolving with incorrect information, in which the MaxEnt method attempts to determine the maximally noncommittal solution. However, the apparent size results from the MaxEnt method, using a non-uniform *a priori* model, suggest the missing information can be filled in. It is plausible that, if a suitable non-uniform *a priori* model were used in the deconvolution of the simulated profile, the effects of over-compensation may be reduced in the solution profiles.

4.5 Summary

We have presented the MaxEnt method as an alternative method for removing instrument broadening in x-ray diffraction profiles. In the case of size-broadened specimen profiles, it was shown that the MaxEnt method can be applied a second time to determine the column-length distribution of the crystallites. This demonstrates the robustness and flexibility of the MaxEnt method, in that it can be applied sequentially to solve a variety of inverse problems. In addition, uncertainties in the solution profile and column distribution can be determined and used in subsequent calculations, such as the integral breadth, Fourier coefficients, column-length distribution and resulting apparent sizes. This makes the MaxEnt method a fully quantitative method for analysing x-ray diffraction profiles.

In addition, the MaxEnt method could be incorporated to include a full Bayesian statistical approach to determine α, a suitable *a priori* model, and background level. While the author did not seek to use this here, the Bayesian technique has been applied in other areas, such as neutron diffraction (see Sivia 1996).

Working with the simulated profiles for $\langle D \rangle = 20$, 50, & 100 nm, the solution specimen

profile, Fourier coefficients, column-length distribution and apparent sizes, $\langle L\rangle_v^I$, $\langle L\rangle_a^F$, $\langle L\rangle_a^{M,u}$ and $\langle L\rangle_v^{M,n}$, were determined for a number of cases. These cases included the best background level (see §4.4.2); the effect of background estimation (see §4.4.3); and deconvolving with non-ideal instrument profiles (see §4.4.4). In each case, the simulated profiles were truncated from $(2\theta_B \pm 5)\,°2\theta$ to $(2\theta_B \pm 1)\,°2\theta$ at a step-size of $0.01\,°2\theta$. In general, where the noise level was relatively high and for varying amounts of instrument broadening, the MaxEnt method was able to determine the solution profile and preserve the positivity of the specimen profile, which enabled subsequent calculations to be performed.

The significance that can be drawn for this study, especially from the last two cases, is that the MaxEnt method can tolerate, to an extent, incorrect or missing information that may arise in the truncation of the observed profile, background estimation and deconvolving with a non-ideal instrument profile. The method produces physically consistent profiles and distributions. Even in the worst case scenarios, the positivity of these profiles and distributions are always preserved. Moreover, when a suitable non-uniform *a priori* model is used, improvements in the final results are obtained. That is, *a priori* information that has "structure" attempts to replace the missing or incorrect information, to produce an accurate and precise result. This was clearly evident when using a non-uniform *a priori* model in the MaxEnt method to determine the column-length distribution and area-weighted size.

As expected, when an instrument profile is used which is itself broadened due to a non-ideal reference material, the ensuing results are compromised. The necessity for a good reference material can not be overstated.

Although a non-uniform *a priori* model was not used in the deconvolution of the profiles, it is the opinion of the author that such a model may reduce the systematic errors introduced in background estimation and using a non-ideal instrument profile on the solution profile and other subsequent results. However, this raises a number of important issues, such as what information should be used in the *a priori* model and how should a suitable model be selected? Although this issue has not been addressed here, it is conceivable that the approach taken by the small-angle scattering community (see Potton et al. 1988*b*, Potton et al. 1988*a*, Hansen & Pedersen 1991, Müller & Hansen 1994, Müller et al. 1996), may provide hints as to how best to resolve these issues.

Chapter 5

A Maximum Entropy Method for Determining the Strain Distribution

In the previous chapter, it was shown how the two-fold maximum entropy method can be applied to determine the specimen profile and the column-length distribution. It was also demonstrated how the MaxEnt method could be used to determine the uncertainties in the solution profile and column-length distribution.

In this chapter, the MaxEnt method is applied to the second problem related to the broadening of x-ray diffraction profiles, *viz.* determining the strain distribution from the broadened profiles. Two extreme cases of strain broadening are considered: the first is Gaussian strain and the second is non-Gaussian strain in the near-Lorentzian limit. As in the previous chapter, the application of the MaxEnt method is two-fold, first determining the specimen profile and then the strain distribution. However, for the strain problem the two-fold application is critically important, because strain is order-dependent (i.e. dependent on hkl) and requires the Fourier coefficients of all available orders of reflection from the x-ray diffraction pattern. In determining the strain distribution using the MaxEnt method, the *a priori* model becomes important in correctly determining the Lagrangian parameter, α, the strain distribution, and in turn its uncertainties. It is demonstrated how a non-uniform *a priori* model can be determined from the data to improve the MaxEnt solution distribution, but issues concerning the information used to determine *a priori* model are raised. This work is in preparation for the literature (see Armstrong & Kalceff 1999*b*).

The background to strain broadening of x-ray diffraction profiles is presented in §5.1 and the difficulties in determining the strain distribution from the available data are discussed here. The two MaxEnt methods for determining strain distributions are presented in §5.2. The first of these is a theoretical approach, for it assumes that the strain moments are known (experimentally this is very difficult to determine). Moreover, it illustrates the role of the entropy function and the *a priori* model or measure in determining the strain distribution, $p(\epsilon_L)$. The second approach is more practical in that it requires the available Fourier coefficients of the profiles in the diffraction pattern and applies the Skilling & Bryan (1984) algorithm. To evaluate the performance of the second MaxEnt method and the integrity of its results, a series of simulations were carried out, as described in §5.3. Two cases of inhomogeneous strain broadening are examined in §5.4 and were modelled on the copper diffraction peaks: the first case considers Gaussian strain, where the MaxEnt method is compared with traditional methods such as the Warren & Averbach and Williamson & Hall methods; the second case examines non-Gaussian strain, where the strain distribution is in the near-Lorentzian limit.

Specimen profiles and simulated observed profiles with instrument broadening, statistical noise and background effects were generated. From these the MaxEnt solution profiles and strain distributions were determined, together with their uncertainties, and used in subsequent calculations of Fourier coefficients, integral breadths and root-mean-square strain. The implications of these calculations for experimental conditions are discussed in §5.5.

5.1 Strain broadening of profile

5.1.1 Specimen profile and Fourier coefficients

A specimen profile can be expressed in terms of a Fourier series as

$$f(s-s_0,\, d^*) = \sum_{L=-\infty}^{+\infty} \left[A(L,\, d^*)\, \cos 2\pi L(s-s_0) \;+\; B(L,\, d^*)\, \sin 2\pi L(s-s_0)\right] \tag{5.1}$$

where $s-s_0$ is the reciprocal-space variable given by $s-s_0 = \frac{2}{\lambda}\sin\theta - d^*$, λ is the wavelength of radiation, d^* is the reciprocal of the atomic interplanar distance and θ is the angular position. L in (5.1) is the column length in real space perpendicular to the diffraction plane

(i.e. parallel to the diffraction vector) and is directly related to the harmonic order, n, of the Fourier coefficients whereby $L = na$, with a defining the "step-size" in real space given as $a = \frac{\lambda}{2}(\sin\theta_2 - \sin\theta_1)^{-1}$ for the angular range $[2\theta_1,\, 2\theta_2]$. Usually (5.1) has a slowly varying scale factor in front of the summation which is set to unity here. $A(L,\, d^*)$ and $B(L,\, d^*)$ are the real and imaginary Fourier coefficients, respectively, which characterise the information concerning the nature of the inhomogeneous lattice strain (order-dependent) distortions and crystallite/domain size (order-independent) broadening contributions. In this chapter, we are only interested in the order-dependent contributions.

Applying the kinematic theory of x-ray diffraction, the Fourier coefficients for the the order-dependent contributions can be expressed as (see Warren 1969)

$$\begin{aligned} A(L,\, d^*) &= \langle \cos 2\pi\epsilon_L L d^* \rangle \\ &= \int_{-\infty}^{\infty} p(\epsilon_L)\, \cos 2\pi\epsilon_L L d^* \, \mathrm{d}\epsilon_L \end{aligned} \tag{5.2}$$

and

$$\begin{aligned} B(L,\, d^*) &= \langle \sin 2\pi\epsilon_L L d^* \rangle \\ &= \int_{-\infty}^{\infty} p(\epsilon_L)\, \sin(2\pi\epsilon_L L d^*) \, \mathrm{d}\epsilon_L \end{aligned} \tag{5.3}$$

where ϵ_L is the strain averaged over the length of column L perpendicular to the diffraction planes, and $p(\epsilon_L)$ is the strain distribution at L. The Fourier coefficients in (5.2) and (5.3) are functions of both L and d^* and represent the averaging over all the distortions in the diffraction volume of the sample (van Berkum et al. 1996). This averaging is represented by $\langle \cdots \rangle$. If $p(\epsilon_L)$ is an even function, then the $B(L,\, d^*)$ is zero for all L and d^*. The resulting specimen profile will be symmetrical about d^*. This implies that the real Fourier coefficients, $A(L,\, d^*)$ are of most interest.

Expressing the Fourier coefficients, (5.2) and (5.3), in terms of $p(\epsilon_L)$ is a phenomenological description of strain broadening of x-ray diffraction profiles. That is, it provides an empirical understanding to strain broadening. A deeper insight can be gained from an understanding of the microscopic arrangement, orientation and type of dislocation (see Wilkens 1970*a*, Wilkens 1970*c*, Groma et al. 1988, Ungár et al. 1989, Ungár & Borbély 1996, Ungár, Ott, Sanders, Borbély & Weertman 1998, Ungár, Révész & Borbély 1998). From an epistemological point of view, there is an analogy with the theory of heat, in that it can be developed from two

directions. The first describes heat in terms of macroscopic variables, while the second applies a microscopic and statistical approach and describes the macroscopic quantities as ensemble averages. The strain distribution, $p(\epsilon_L)$, can be considered as a"signature" of the microscopic description of the strain field in that it assigns a probability distribution to the strain, averaged over a distance L, parallel to the diffraction vector. In other words, $p(\epsilon_L)$ is a function of the ensemble of dislocations in the crystallite. Determining $p(\epsilon_L)$ from the experimental data may provide an approach to testing theoretical models related to dislocation broadening.

From a practical point of view, being able to determine $p(\epsilon_L)$ from strain broadened profiles offers an alternative approach to determining the root-mean-square strain, $\langle\epsilon_L^2\rangle^{\frac{1}{2}}$, over standard methods such as those of Warren & Averbach and Williamson & Hall. Both these methods make specific assumptions about the nature of the strain. In the case of the Warren & Averbach method, the real Fourier coefficients are expanded for a small L. If $p(\epsilon_L)$ is a Gaussian distribution, the Warren & Averbach method is linear for all d^{*2}. The Williamson & Hall method is also based on assumptions about the shape of the specimen profile and the contributions of the crystallite size and strain effects. This is a consequence of attempting to generalize the work of Stokes & Wilson (1942, 1944a). More specifically, in the case of strain Stokes & Wilson (1944a) assume the strain distribution is independent of L, resulting in a linear dependence $\beta(d^*)$ for all d^*. In this chapter, we are not only interested in determining $\langle\epsilon_L^2\rangle^{\frac{1}{2}}$, but also $p(\epsilon_L)$ using a robust and reliable method with a minimum of assumptions.

For a symmetrical $p(\epsilon_L)$, $\langle\epsilon_L\rangle = 0$ for all L, and the mean-square strain, $\langle\epsilon_L^2\rangle$, is nonzero and given by,

$$\langle\epsilon_L^2\rangle = \int_{-\infty}^{\infty} \epsilon_L^2\, p(\epsilon_L)\, \mathrm{d}\epsilon_L \tag{5.4}$$

from which $\langle\epsilon_L^2\rangle^{\frac{1}{2}}$ follows.

5.1.2 Inverse Cosine Fourier Transform

Equations (5.2) and (5.3) are simply cosine and sine Fourier coefficients. An obvious approach to determining $p(\epsilon_L)$ is to apply an inverse cosine Fourier transform to (5.2) for a given L. This approach is analogous to the Stokes (1948) method for deconvolving the ob-

served profile to determine the specimen profile, in that it is a naive approach which does not take into consideration the ill-conditioning of the problem and any limitations in the data. Expressing $p(\epsilon_L)$ in terms of $A(L,\, d^*)$ (see McKeehan & Warren 1953, Warren 1959),

$$p(\epsilon_L) = 2L \int_0^\infty A(L,\, d^*)\, \cos(2\pi L \epsilon_L d^*)\, \mathrm{d}d^* \tag{5.5}$$

Equation (5.5) assumes complete information about $A(L,\, d^*)$ for all L and d^*. However, in the light of experimental data, this is not generally the case and we are presented with several difficulties in determining $p(\epsilon_L)$. For a given L, the number of points for $A(L,\, d^*)$ is limited to the number of lines in the diffraction pattern. For high symmetry materials a few diffraction profiles (typically < 10) are recorded due to the multiplicity (e.g. cubic lattice system), but the profiles are usually well separated. This implies the profiles can be well resolved, but because of the limited number of reflections only a small number of points in $A(L,\, d^*)$ for a given L can be determined. In the case of low symmetry materials, the multiplicity is generally lower, implying a greater number of lines in the diffraction pattern, but the profiles are often overlapped. This makes it difficult to resolve individual profiles and in turn reduces the number of points in $A(L,\, d^*)$ for a given L. There is also the case where the strain is anisotropic, causing the diffraction spot in reciprocal space not to be symmetrical for all crystallographic directions, $[hkl]$. For this case, multiple orders need to be considered. However, for a standard diffractometer, at best, only three multiple orders can be recorded which in turn reduces the number of points for $A(L,\, d^*)$.

The consequence for all these situations is incomplete information for $A(L,\, d^*)$. In addition to these fundamental limitations, the experimental data (i.e. Fourier coefficients) may contain noise and systematic errors. The noise may be imparted from the observed profile, while systematic errors may be due to incorrect background estimation and/or truncation errors affecting the solution profiles. In turn, these factors may contribute to producing erroneous strain distributions or truncation of the distribution due to negative oscillations. This point was not lost on McKeehan & Warren (1953), who demonstrated that in principle $p(\epsilon_L)$ could be determined from (5.5) for strained tungsten diffraction data which is both isotropic and has a cubic atomic lattice system. It was also pointed out by the same authors that in determining $p(\epsilon_L)$ (see pp. 54-55 McKeehan & Warren 1953):

The analysis from which these curves are obtained requires a high degree of accu-

racy in the experimental data, and the reliability of the final curves is somewhat uncertain.

Another observation about (5.5) is that it does not preserve the positivity of $p(\epsilon_L)$ and does not take into account any *a priori* information. These last two points are important. The strain distribution, like any probability distribution, is a positive and additive distribution and we expect any method for determining $p(\epsilon_L)$ to preserve this property. By applying the positivity condition the number of feasible solutions is also reduced. The use of *a priori* information is also an important point, in that the above-mentioned limitations result in incomplete information for $A(L, d^*)$. If *a priori* information is available, however, from either a theoretical model or additional experimental data, it contributes to "filling in" the missing information or overcoming the fundamental or practical limitations. In the next section we consider the MaxEnt method which can incorporate these important features.

5.2 Maximum Entropy (MaxEnt) method for $p(\epsilon_L)$

An alternative approach to determining $p(\epsilon_L)$ that preserves the positivity, quantifies the distribution and incorporates *a priori* information is the maximum entropy (MaxEnt) method. In fact, there are two ways in which $p(\epsilon_L)$ can be determined using the MaxEnt method. The first is interesting from a theoretical point of view. It determines $p(\epsilon_L)$ by maximising the entropy function with respect to the moments of the strain distribution and reveals at a fundamental level the influence to which the measure or *a priori* model has on the solution. This discussion has been adopted from the general discussion presented by Mead & Papanicolaou (1984).

The second MaxEnt method (see §5.2.2) is more practical in that it uses the available data, the Fourier coefficients, and does not rely on the moments of the strain distribution, but rather on resolving the Fourier coefficients of all the available profiles in the diffraction pattern. It also follows on from the approach discussed in Chapter 4: that is, a two-fold application, while treating the problem of determining $p(\epsilon_L)$ as an inverse problem. The first stage applies the MaxEnt method to removing the instrumental broadening and determining the uncertainties in the corresponding solution profile; the second stage consists

of determining the Fourier coefficients of all solution profiles (i.e. for a range of d^* and L values, see (5.2) & (5.3)) and applying the MaxEnt method again to determine $p(\epsilon_L)$ and its corresponding uncertainties. Subsequent results such as $\langle \epsilon_L^2 \rangle$ and their uncertainties can also be determined.

5.2.1 First approach: Method of moments

This approach relies on knowing the moments of the strain distribution. Experimentally this is difficult because is involves expanding the Fourier coefficients, (5.2) and (5.3). From the Warren & Averbach method, only the second moments, $\langle \epsilon_L^2 \rangle$, can be determined for small L and two or more orders must be known. However, it is worth investigating since this approach demonstrates the application of the entropy in such problems.

Suppose that $p(\epsilon_L)$ is a symmetrical distribution centred about $\epsilon_L = 0$. All odd moments will be zero for all ϵ_L. Only the even moments, $\langle \epsilon_L^{2n} \rangle$, for $n = 1, 2, 3, \ldots$ will be nonzero for all L. This can be stated in terms of the strain distribution $p(\epsilon_L)$

$$\langle \epsilon_L^{2n} \rangle = \int_{-\infty}^{+\infty} \epsilon_L^{2n}\, p(\epsilon_L)\, \mathrm{d}\epsilon_L \tag{5.6}$$

where we assume that $p(\epsilon_L)$ is normalised for unit area, and also that a finite number of moments can be determined such that $n = 1, 2, 3, \ldots, N$ and for $n = 0$, $\langle \epsilon_L^0 \rangle = 1$. Strictly speaking, the entropy functional for $p(\epsilon_L)$ can be expressed as

$$S[p(\epsilon_L)] = -\int_{-\infty}^{+\infty} p(\epsilon_L)\, \ln\left[p(\epsilon_L)/m_p(\epsilon_L)\right]\, \mathrm{d}\epsilon_L \tag{5.7}$$

where $m_p(\epsilon_L)$ is the *a priori* model or measure for $p(\epsilon_L)$ (see Chapter 3) and the square brackets in (5.7) distinguish it as a functional[1]. We now want to maximize (5.7) with respect to (5.6). To do this we usually write down the Lagrangian functional

$$Q[p(\epsilon_L)] = -\int_{-\infty}^{+\infty} p(\epsilon_L)\, \ln\left[p(\epsilon_L)/m_p(\epsilon_L)\right]\, \mathrm{d}\epsilon_L - \sum_{n=0}^{N} \alpha_n \left[\int_{-\infty}^{+\infty} \epsilon_L^{2n}\, p(\epsilon_L)\, \mathrm{d}\epsilon_L - \langle \epsilon_L^{2n} \rangle\right] \tag{5.8}$$

[1] A functional takes a function as its argument and returns a number. This is in some ways similar to a function, but a function takes a number as its argument and returns a number. However, unlike a function, a functional is dependent upon the overall properties of the function, in this case $p(\epsilon_L)$. In other words a functional *maps* an infinitely differentiable function *into* a number (Ryder 1985).

where $\{\alpha_n\}$ for $n = 0, 1, 2, 3, \ldots, N$ are the set of positive Lagrangian multipliers. Applying the calculus of variation we have

$$\begin{aligned}\frac{\delta Q[p(\epsilon_L)]}{\delta p(\epsilon_L)} &= -\int_{-\infty}^{+\infty} [\ln[p(\epsilon_L)/m_p(\epsilon)] + 1]\ \delta p(\epsilon_L)\, \mathrm{d}\epsilon_L - \sum_{n=0}^{N} \alpha_n \int_{-\infty}^{+\infty} \epsilon_L^{2n}\, \delta p(\epsilon_L)\, \mathrm{d}\epsilon_L \\ &= -\int_{-\infty}^{+\infty} \left[\ln[p(\epsilon_L)/m_p(\epsilon_L)] + 1 + \sum_{n=0}^{N} \alpha_n\, \epsilon_L^{2n}\right] \delta p(\epsilon_L)\, \mathrm{d}\epsilon_L \\ &= 0,\end{aligned}$$

since $\delta p(\epsilon_L)$ is arbitrary and the term in the square bracket must be zero for $\frac{\delta Q[p(\epsilon_L)]}{\delta p(\epsilon_L)} = 0$. Taking the term in the square brackets, we arrive at the following

$$p(\epsilon_L) = m_p(\epsilon_L)\ \exp\left[-1 - \alpha_0 - \sum_{n=1}^{N} \alpha_n\, \epsilon_L^{2n}\right]. \tag{5.9}$$

However, we require that (5.9) be normalised for all ϵ_L such that

$$\int_{-\infty}^{+\infty} p(\epsilon_L)\, \mathrm{d}\epsilon_L = 1 \tag{5.10}$$

and by substituting (5.9) into (5.10) and after a little manipulation, we can define the normalisation term

$$\begin{aligned}Z(\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_N) &\equiv \exp[1 + \alpha_0] \\ &= \int_{-\infty}^{+\infty} m(\epsilon_L)\ \exp\left[-\sum_{n=0}^{N} \alpha_n\, \epsilon_L^{2n}\right] \mathrm{d}\epsilon_L \end{aligned} \tag{5.11}$$

which is analogous to the partition function encountered in statistical mechanics. $Z(\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_N)$ has many interesting properties, such as being able to express the average quantities in terms of (5.11). (These are not considered here, however.) Using (5.11) in (5.9) we can express $p(\epsilon_L)$ as

$$p(\epsilon_L) = \frac{m_p(\epsilon_L)}{Z(\{\alpha_n\})}\ \exp\left[-\sum_{n=1}^{N} \alpha_n\, \epsilon_L^{2n}\right] \tag{5.12}$$

where the even moments, $\langle \epsilon_L^{2n} \rangle$ for $n = 1, 2, 3, \ldots, N$ can be determined from $p(\epsilon_L)$. This approach requires solving $N+1$ nonlinear simultaneous equations and for practical problems, N is large. Conventional methods for numerically solving nonlinear simultaneous equations, such as the Newton-Raphson method, require a good initial guess of $\{\alpha_n\}$. We also notice in (5.12) the influence the *a priori* model, $m_p(\epsilon)$, has on the solution, in that it "modulates"

the exponential term in (5.12). If $m_p(\epsilon)$ is a uniform distribution, no *a priori* information is conveyed. However, if $m_p(\epsilon)$ is a non-uniform distribution, information concerning the scaling, shape and position of the distribution is conveyed. Later in this chapter, the significance of $m_p(\epsilon)$ is discussed. The uniqueness of this solution is determined by the sequence of moments. In this case, all the even moments, $\langle \epsilon_L^{2n} \rangle$ for $n = 1, 2, 3, \ldots, N$ must be a monotonic sequence (Mead & Papanicolaou 1984). Moreover, the above solution assumes that there is no uncertainty in the moments. Suppose the moments were determined experimentally; then the above solution can be generalised to include the uncertainty in $\langle \epsilon_L^{2n} \rangle$ by introducing a χ^2 statistic (see Wu 1997).

For this version of the MaxEnt method, the positivity of the strain distribution is preserved, *a priori* information concerning $p(\epsilon_L)$ can be included and the method can be generalised to include uncertainties in the moments. However, the moments of $p(\epsilon_L)$ are not usually known and from a numerical point of view, this approach requires numerically solving $N + 1$ nonlinear equations.

5.2.2 Second approach: Using the Fourier coefficients

The second approach to determining the $p(\epsilon_L)$ uses the Fourier coefficients of the specimen profile. By applying the Skilling & Bryan (1984) algorithm, the numerical difficulties of determining the set of Lagrangian multiplier is reduced to determining a single multiplier and the algorithm takes into consideration the uncertainties in the Fourier coefficients. In this approach we can treat the profile and coefficients as discrete data and the entropy function for the strain distribution is

$$S(p) = -\sum_{j=1}^{N} p_j \, \ln(\frac{p_j}{m_{p\,j}}) \tag{5.13}$$

where $p = \{p_j;\ j = 1, 2, 3, \ldots, N\}$ and $m_p = \{m_{p\,j};\ j = 1, 2, 3, \ldots, N\}$ represents the *a priori* model for the strain distribution. A uniform distribution over some region of ϵ_L for the *a priori* model represents our worst choice in that it states our ignorance concerning $p(\epsilon_L)$. However, a theoretical model or *intrinsic model* (Müller & Hansen 1994) determined from the observed data by an independent approach can be used as an *a priori* model for $p(\epsilon_L)$. This is one of the advantages the MaxEnt method has over existing methods for determining

$p(\epsilon_L)$: it incorporates the *a priori* information concerning $p(\epsilon_L)$ and the positivity of $p(\epsilon_L)$ is preserved.

The entropy function is maximized with respect to the statistic function which incorporates the experimental data and uncertainties in $\hat{A}$ as

$$C(p) = \sum_{i=1}^{M} \frac{(\tilde{A}_i - \hat{A}_i)^2}{\sigma^2_{\hat{A}\,i}} \tag{5.14}$$

where $\tilde{A} = \{\tilde{A}_i;\, i = 1, 2, 3, \ldots, M\}$ are the trial real Fourier coefficients; $\hat{A} = \{\hat{A}_i;\, i = 1, 2, 3, \ldots, M\}$ are the real Fourier coefficients determined from the MaxEnt solution profiles for a particular L over a range of $d^* \in [0,\, d^*_{max}]$; $\sigma_{\hat{A}\,i}$ are the corresponding uncertainties of $\hat{A}$ which are determined from the uncertainties in the solution profile, $\hat{f}$ (also see §(3.3.5)).

Maximizing (5.13) with respect to the constraints of $p_i \geq 0$ for all i, (5.14) can be expressed simply in terms of a Lagrangian function, $Q(p)$ as

$$Q(p) = \alpha S(p) - C(p) \tag{5.15}$$

where α is the unknown (positive) Lagrangian multiplier. In maximizing $Q(p)$ we are searching for a solution distribution, $\hat{p}$, which satisfies $\nabla Q(\hat{p}) = 0$. We have found the Skilling & Bryan (1984) algorithm (see §3.3) is well suited to this task. The algorithm determines the Lagrangian multiplier, α, and the solution distribution, $\hat{p}$, when the constraint $C(p) = C_{aim}$, where C_{aim} is set equal to the number of points. The use of the *a priori* model has an important influence in determining α, $\hat{p}$ and in turn the uncertainties in the solution profile (see (3.58)). The algorithm is halted when the vectors ∇S and ∇C are parallel (or near parallel).

Trial data and kernel matrix

The trial coefficients given in (5.14) can be expressed in terms of the kernel and strain distribution, $p(\epsilon_L)$ as

$$\tilde{A}_i = \sum_{j=1}^{N} W_{ij}\, p_j \tag{5.16}$$

where $\tilde{A} = \{\tilde{A}_i;\, i = 1, 2, 3, \ldots, M\}$ are the trial Fourier coefficients; $\mathbf{W} = \{W_{ij};\, i = 1, 2, 3, \ldots, M,\, j = 1, 2, 3, \ldots, N\}$ is the kernel in (5.2) expressed as on $[M \times N]$ matrix

where $N \leq M$. Each term in the kernel matrix is given by

$$W_{ij} = 2\Delta\epsilon_L \cos(2\pi L d_i^* \epsilon_{Lj}) \tag{5.17}$$

where $\Delta\epsilon_L = \epsilon_{Lj+1} - \epsilon_{Lj}$; d^* is divided into M terms over the range $[0, d^*_{max}]$ and ϵ_L in divided into N terms over the range $[0, \epsilon_{Lmax}]$.

The MaxEnt method presented here assumes the imaginary coefficients (5.3) are zero for all L and d^*. This implies that the specimen profile and strain distributions are symmetrical. However, in the case of asymmetrical specimen profiles, the kernel in (5.16) and (5.17) for the MaxEnt method should be re-expressed to take into consideration the nonzero imaginary coefficients.

Determining the uncertainty in $\hat{p}(\epsilon_L)$

The uncertainties in the solution strain distribution, $\hat{p}(\epsilon_L)$ are calculated in the same manner outlined in Chapter 3 and applied in Chapter 4. The uncertainties in $\langle\epsilon_L^2\rangle^{\frac{1}{2}}$ can be determined from the uncertainty region of $\hat{p}(\epsilon_L)$. That is, a set of distributions is randomly drawn from the uncertainty region of $\hat{p}(\epsilon_L)$. From the set of distributions, a set of $\langle\epsilon_L^2\rangle^{\frac{1}{2}}$ values can be determined from which the standard error in the mean can be calculated. The mean value was taken as that value of $\langle\epsilon_L^2\rangle^{\frac{1}{2}}$ determined from the solution distribution, $\hat{p}(\epsilon_L)$.

5.3 Simulated strain broadened profiles

The two-fold MaxEnt method is evaluated by considering two extreme cases of inhomogeneous strain broadening. The first case involves a Gaussian strain distribution from which we expect the standard results from the Warren & Averbach method. This becomes the bench mark against which to compare the MaxEnt method. The second case involves a non-Gaussian distribution modelled on a Pearson-VII function in the near Lorentzian limit. This demonstrates the breakdown of the Warren & Averbach method and the ability of the MaxEnt method to determine the strain distribution in the near-Lorentzian limit. These two cases are considered “extreme” because they define the full “spectrum” of inhomogeneous strain broadening.

Using the distributions for $p(\epsilon_L)$, the Fourier coefficients and the corresponding specimen profiles can be generated. The profiles are modelled on a copper specimen using the six most intense peaks, *viz.* 111, 200, 220, 311, 331 and the 420. The crystallite size is assumed to be infinite. The specimen profiles are convolved with an instrument profile, with a background-level and statistical noise being added.

5.3.1 Strain distribution and specimen profiles

Case 1: Gaussian strain distribution

In the first case, a Gaussian distribution was considered, given by

$$p(\epsilon_L) = \frac{1}{\sqrt{2\pi\langle\epsilon_L^2\rangle}} \exp\left[-\frac{(\epsilon_L - \langle\epsilon_L\rangle)^2}{2\langle\epsilon_L^2\rangle}\right] \tag{5.18}$$

where ϵ_L is the strain averaged over the distance L perpendicular to the diffraction planes; $\langle\epsilon_L\rangle$ is the average strain and is set to zero for all L; $\langle\epsilon_L^2\rangle$ is the mean-square strain averaged over the distance L.

For these simulations a simple model for $\langle\epsilon_L^2\rangle$ for straight dislocations has been applied. $\langle\epsilon_L^2\rangle$ is proportional to L^{-1} (Rothman & Cohen 1971, Adler & Houska 1979, Schlosberg & Cohen 1983, Nandi et al. 1984), given by

$$\langle\epsilon_L^2\rangle = \frac{G_{hkl}^2}{|L|} \tag{5.19}$$

where G_{hkl}^2 is a constant dependent on the crystallographic direction, $[hkl]$, and has units of length; L is the distance perpendicular to the diffraction planes, determined from $L = na$ where a and n have been defined (see §5.1). L can take on both negative and positive values over the region $(-\infty, \infty)$. Here G_{hkl}^2 was made independent of $[hkl]$ and was set to a value of 1.0×10^{-4} nm. This model is often used to describe cold-worked metals and the quoted value is considered typical (Adler & Houska 1979, Nunzio et al. 1995). Using (5.18) and (5.19) the Fourier coefficients can be determined. Equation (5.18) is an even and symmetrical function and the imaginary coefficients (5.3) are zero for all L and d^*. The real Fourier coefficients become

$$A(L,\, d^*) = \exp\left[-2\pi^2\, L^2\, \langle\epsilon_L^2\rangle\, d^{*2}\right]. \tag{5.20}$$

Substituting (5.19) into (5.20) we arrive at

$$A(L,\, d^*) = \exp\left[-2\pi^2\, |L|\, G_{hkl}^2\, d^{*2}\right] \tag{5.21}$$

where $L \in (-\infty,\, +\infty)$. We notice in (5.21) that the Fourier coefficients are for a Lorentzian function. Even though a Gaussian strain distribution has been used, it is the spatial dependence of the distributions that determines the Fourier coefficients. If the strain distribution were constant for all L, then the results from Stokes & Wilson (1944*a*) would apply. Figure 5.1 shows the Fourier coefficients for the six most intense lines in copper using (5.21).

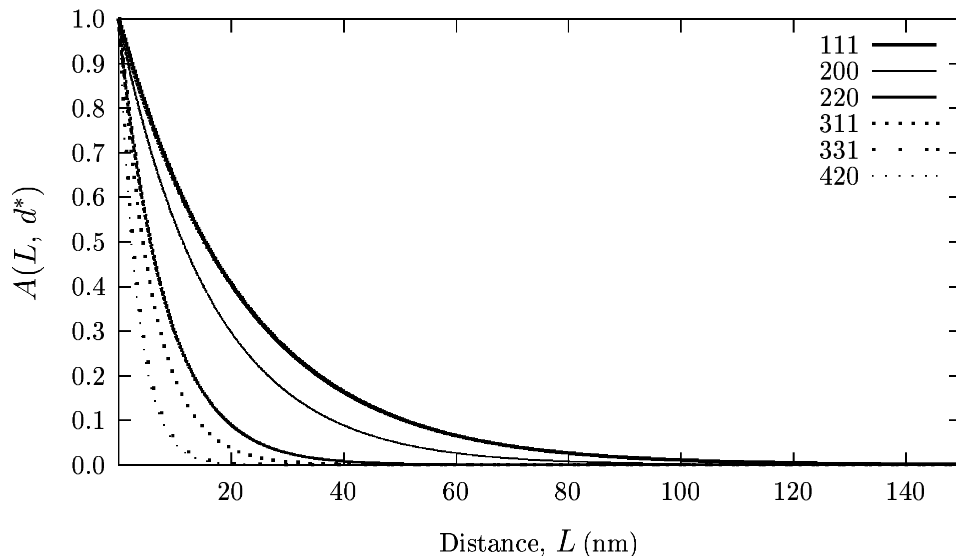


Figure 5.1: Showing the real Fourier coefficients of the specimen profiles using Gaussian strain distributions for different values of d^* from 111 to 420 peaks.

From (5.20) or (5.21) the integral breadth in $(s - s_0)$-space can be determined analytically from

$$\beta(d^*) = \left[\int_{-\infty}^{+\infty} A(L,\, d^*)\ \mathrm{d}L\right]^{-1}. \tag{5.22}$$

Substituting (5.21) into (5.22), we arrive at the following result for the integral breadth

$$\beta(d^*) = \pi^2\, G_{hkl}^2\, d^{*2}. \tag{5.23}$$

We point out that (5.23) is proportional to d^{*2}. The usual linear dependence which is the basis of the Williamson & Hall method, assumes that the strain distribution, $p(\epsilon_L)$, is independent of L.

Applying a Fourier transform, the specimen profile in $(s - s_0)$-space can be determined, and can be transformed (i.e. mapped) into 2θ-space. The modelled profiles assumed isotropic strain in $[hkl]$.

Case 2: Non-Gaussian strain distribution

The non-Gaussian strain distribution used in this case consisted of a Pearson-VII function (also see van Berkum 1994)

$$p(\epsilon_L) = \alpha(\sigma_L, m) \left[1 + \frac{1}{m}(\frac{\epsilon_L}{\sigma_L})^2\right]^{-m} \tag{5.24}$$

where m is the exponent and influences the tails of $p(\epsilon_L)$. For $m = 1$, (5.24) becomes a Lorentzian function and as $m \to \infty$, (5.24) becomes a Gaussian function. In general, m will be dependent on L, but in this case $m = 1.55$ for all L (i.e. independent of L); $2\sigma_L$ is the full width at half maximum ($FWHM$) of the strain distribution and is dependent on L; $\alpha(\sigma_L, m) = \frac{\Gamma(m)}{\sqrt{\pi m}\, \sigma\, \Gamma(m-\frac{1}{2})}$ is the normalization term which ensures the area of (5.24) is equal to unity. The distribution is centred about $\epsilon_L = 0$ and has an average set to zero in (5.24). As was the case for Gaussian strain, a simple model for the dependence of $\langle \epsilon_L^2 \rangle$ and the FWHM, $2\sigma_L$, was used. Following van Berkum (1994) and Adler & Houska (1979), σ_L is given by

$$\sigma_L = \frac{\sigma_0}{|L|^q} \tag{5.25}$$

where σ_0 is a constant with units of $(length)^q$ and had a value of $7.5 \times 10^{-3}\,\mathrm{nm}^{\frac{1}{2}}$. $\frac{1}{q}$ defines the "wavelength" of the distortions (van Berkum 1994) and was set to $q = \frac{1}{2}$. This value of q is often interpreted as an ideal-paracrystal model (see Somashekar & Somashekarappa 1997, Hosemann et al. 1985, Zocchi 1980) and has recently been applied to metallic samples (Somashekar & Somashekarappa 1997). The mean-square strain for (5.24) is only finite for $m > \frac{3}{2}$ and is given by

$$\langle \epsilon_L^2 \rangle = \frac{m\, \sigma_L^2}{2m - 3} \tag{5.26}$$

where (5.25) is used in (5.26).

The real Fourier coefficients using (5.24) are given by (also see van Berkum 1994)

$$A(L, d^*) = \frac{2}{\Gamma(\nu)}\, z^\nu\, K_\nu(2\, z) \tag{5.27}$$

where $L \in (-\infty, +\infty)$, $\nu = m - \frac{1}{2}$; $z = \pi \sqrt{m}\, \sigma_L\, |L|\, d^*$; and $K_\nu(2\, z)$ is a Bessel function of the first kind[2]. Figure 5.2, shows the Fourier coefficients using (5.27), for the six most intense copper peaks. From (5.27) the integral breadth using (5.22) can be determined. Here, we evaluated the integral breadth by numerical integration and were able to described it as

$$\beta(d^*) = \psi(\sigma_0, q, m)\, d^{*2} \tag{5.28}$$

where $\psi(\sigma_0, q, m)$ is a function dependent on the parameters σ_0, q and m. In (5.28) we notice a similar quadratic dependency in the integral breadth. Furthermore, the imaginary Fourier coefficients will be zero for all L and d^* since the distribution $p(\epsilon_L)$ is an even and symmetrical function. As in the approach followed in the Gaussian case, the specimen profiles were determined by taking the Fourier transform of (5.27) and mapping into 2θ-space.

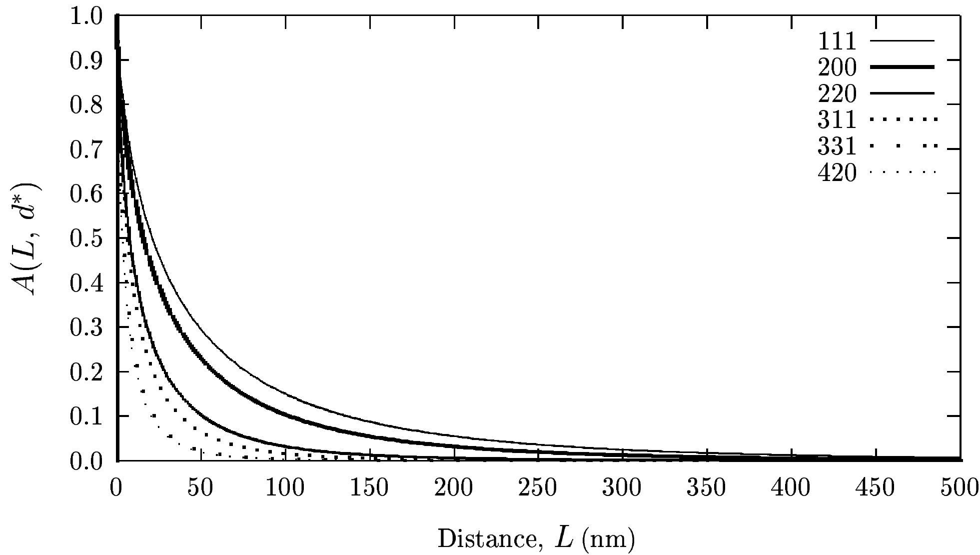


Figure 5.2: Showing the real Fourier coefficients of the specimen profiles using non-Gaussian (Pearson-VII) strain distributions for different values of d^* corresponding from the 111 to 420 profiles.

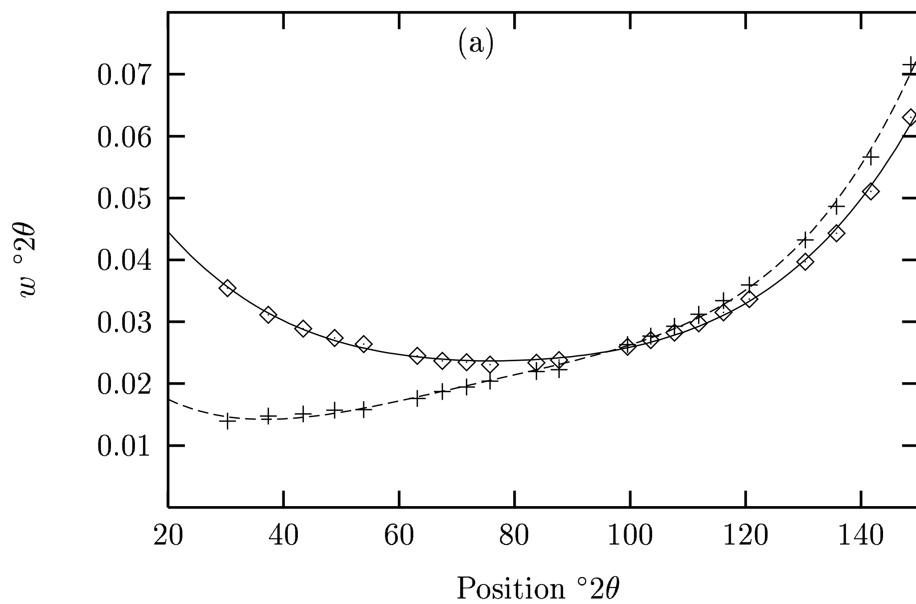
[2]This integration was performed using Mathematica3.0, Wolfram Research Inc., 100 Trade Center Drive Champion, Illinois, 61820-7237.

5.3.2 Simulated instrument profiles

The instrument profile used in this study was based on the "ideal" reference profile modelled on a LaB_6 standard reference specimen. This was applied in §4.3.2. By "ideal", we mean that the reference profile only characterizes the optics the diffractometer and does not include finite size broadening effects or lattice distortions that may arise from the sample. The settings were assumed to be those for a standard diffractometer with a radiation source of $CuK_{\alpha 1a} + CuK_{\alpha 1b}$. (The full set is given in §4.3.2). The peaks of the reference profile were fitted with a split Pearson-VII function. The parameters defining the split-Pearson-VII were fitted to a low order-polynomial. Figure 5.3 shows the w and m parameters for the Pearson-VII function. The instrument profiles were evaluated over the same range as the specimen profiles.

5.3.3 Modelling of "observed" profiles

The modelling of the "observed" copper profiles essentially involved convolving the instrument profile with the specimen profile and including a background level and noise. After the convolution of the instrument profiles and specimen profiles, the intensities of the resulting profiles were scaled to a peak maximum of 6000 counts. This simulated the experimental case where the profiles were recorded to have a particular peak maximum and was done for the six peaks with a relative intensity greater than 5%. Statistical noise was taken from a Poisson distribution and added to the convolved product of the instrument and specimen profile. The background level was assumed to be constant at 10% of the peak maximum over the 2θ-range and Poisson noise was also imparted onto it. An illustration of the low- and high- angle simulated observed profiles representing the two extremes of broadening, for the Gaussian strain case, are shown in Figure 5.4. The corresponding non-Gaussian strain cases are also shown in Figure 5.5. For clarity the 200, 220, 311 and 331 profiles for both cases are not shown.



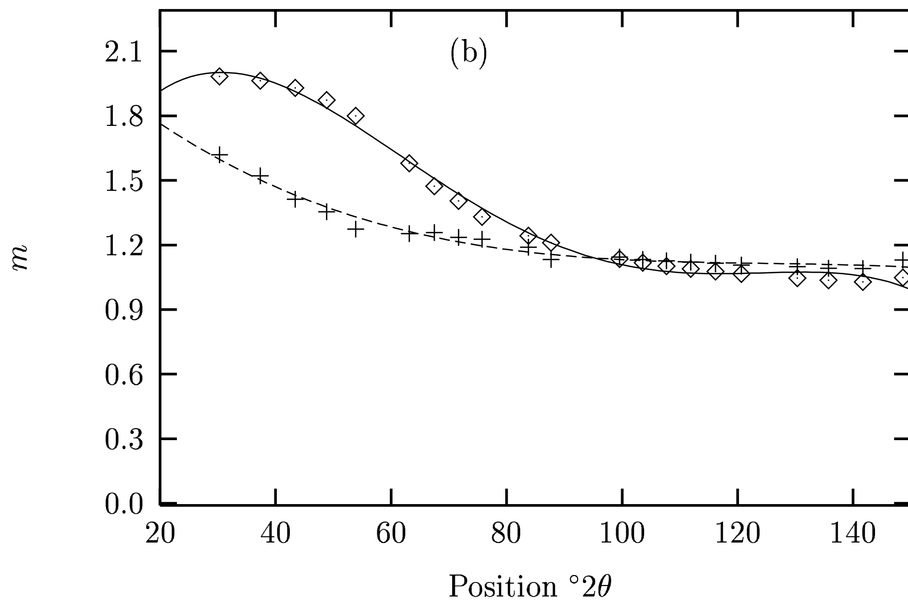
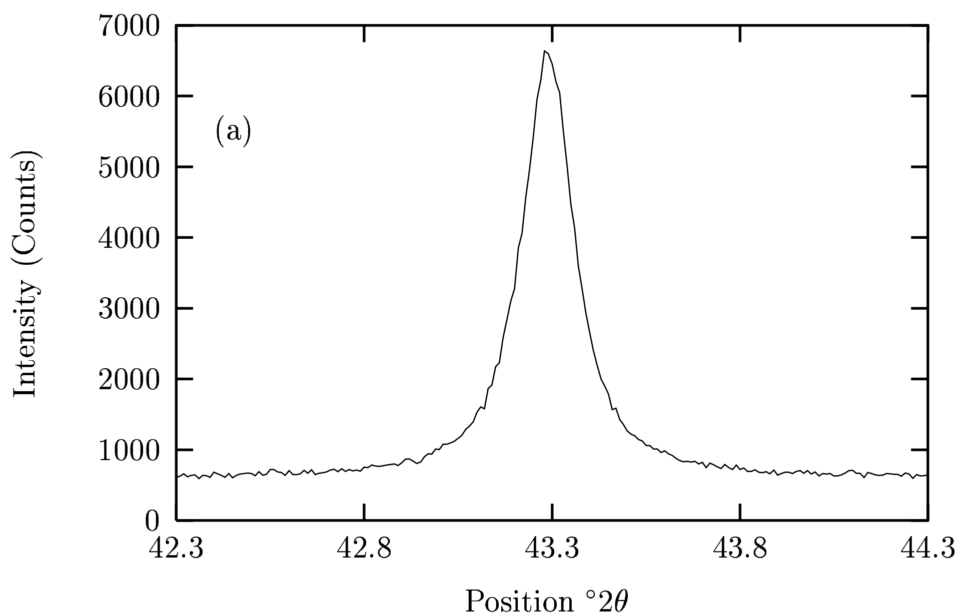


Figure 5.3: Calibration plots for the parameters w and m defining the Pearson-VII profile generated from an (ideal) reference LaB_6 diffraction pattern: (a) w vs 2θ, where $2w = FWHM\ ^{\circ}2\theta$ for the low- (diamonds) & high- (crosses) angle sides of the peak; (b) m vs 2θ, for the low- (diamonds) & high- (crosses) angle sides of the peak.



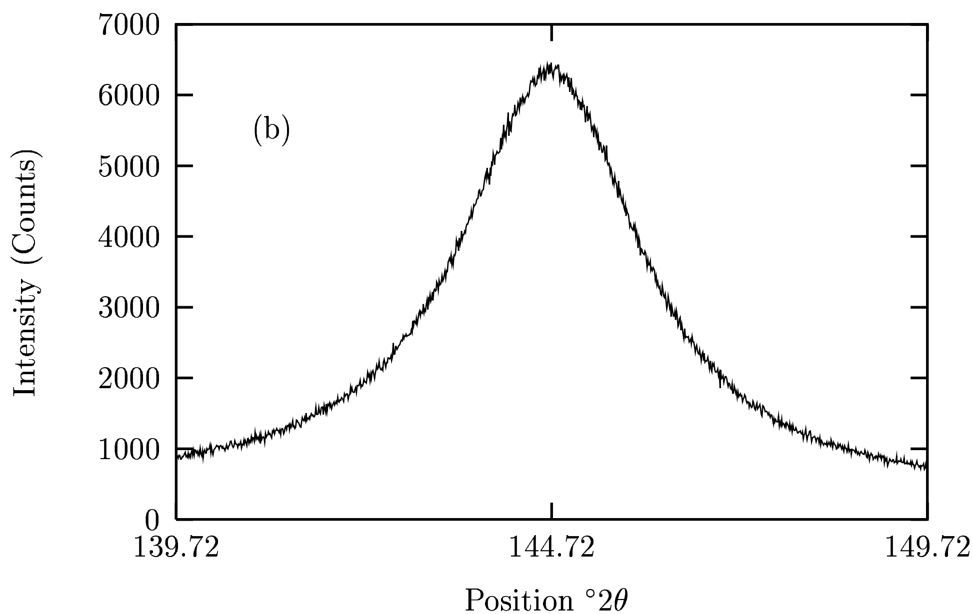


Figure 5.4: The 111 and 420 simulated ("observed") profiles for the Gaussian strain: (a) The 111 observed profile over the interval $(2\theta_B \pm 2)°2\theta$; (b) The 420 observed profile over the interval $(2\theta_B \pm 5)°2\theta$. The observed profiles for the 200 to 331 lines are not shown.

If we were to consider the case of anisotropic inhomogeneous strain, then it would be essential to include the 222 and 400 profiles, because these profiles make up a multiple order.

However, these peaks have a low relative intensity ($< 5\%$) and experimentally it would be impractical to record them as part of a normal full scan using a conventional diffractometer due to the effects of background and statistical noise making the recording times very long.

5.4 Computational Results and Discussion

5.4.1 Preliminary calculations

Background estimation

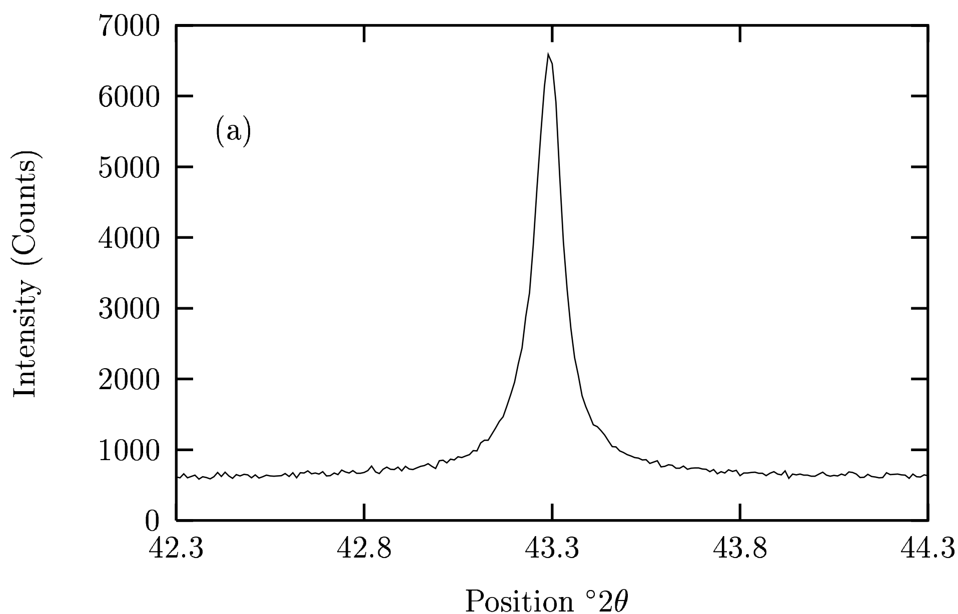
The background estimation for the simulated observed profiles was conducted using the method outlined in §4.4.1. In summary, a linear estimation was made over the profile, the MaxEnt method was applied to determine the specimen profile, and the Fourier coefficients were examined. This was done iteratively, with each iteration involving a lowering or raising of the background level until the distortions (especially about the origin of the Fourier coefficients) were reduced (see §4.2.1). In addition to this procedure, we found that plotting the $\ln A(L, d^*)$ versus L for a given d^* highlighted any distortions.

Determining the specimen profile and strain distribution

For determining the solution profile and strain distribution, the same procedure and MaxEnt algorithm used in Chapter 4 was applied. The deconvolution was carried out in 2θ-space and C_{aim} was set equal to the number of data points in the background-corrected observed profile. The region of deconvolution varied depending on the broadening of the profile. For relatively low angle peaks, such as the 111 and 200 lines, the number of points was 200 or $(2\theta_B \pm 2)\,°2\theta$ at a step-size of $0.01\,2\theta$. This increased to 1000 points or $(2\theta_B \pm 5)\,°2\theta$ at the same step-size of $0.01\,°2\theta$. Deconvolutions for all six most intense profiles were carried out. The solution profiles, $\hat{f}$, and actual profiles, f, were compared by evaluating the R-factor, R_f

$$R_f = \left[\frac{\sum_i (f_i - \hat{f}_i)^2}{\sum_i f_i^2} \right]^{\frac{1}{2}}. \tag{5.29}$$

Once all the solution profiles had been determined, the Fourier coefficients and uncertainties were calculated. This also enabled other methods such as the Warren & Averbach and



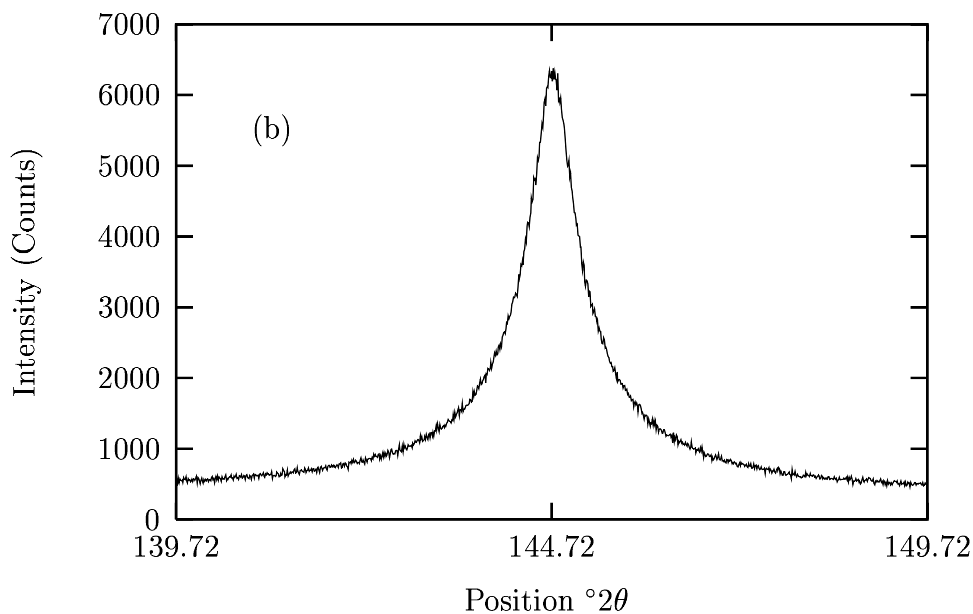


Figure 5.5: The 111 and 420 simulated ("observed") profiles for non-Gaussian strain: (a) The 111 observed profile over the interval $(2\theta_B \pm 2)°2\theta$; (b) The 420 observed profile over the interval $(2\theta_B \pm 5)°2\theta$. The observed profiles for the 200 to 331 lines are not shown.

Williamson & Hall methods to be applied. For a given value of L, the Fourier coefficients and their uncertainties for all available d^* were interpolated. A set of values for $A(L, d^*)$

for a particular value of L were interpolated over the range of $d^* \in [0,\, d^*_{max}]$, where d^*_{max} consisted of the maximum d^* value in the data (in this case the 420 profile). In the case of $d^* = 0\,\mathrm{nm}^{-1}$, the corresponding strain Fourier coefficient was given by $A(L,\, 0) = 1$ for a given L value. Although $A(L,\, 0)$ cannot be measured directly, setting it to unity can be justified from the definition of the strain Fourier coefficients given in (5.2) and (5.3). The solution strain distribution, $\hat{p}(\epsilon_L)$, was normalized for unit area, and the mean-square strain, given by (5.4), was determined.

The difference between the theoretical quantities, $\langle x \rangle_{Th}$, and calculated quantities, $\langle x \rangle_{cal}$, were compared by determining the percentage difference, *diff*,

$$diff = \left| 1 - \frac{\langle x \rangle_{cal}}{\langle x \rangle_{Th}} \right| \times 100\%. \tag{5.30}$$

Determining the variance and uncertainties

For the deconvolution of the simulated observed profiles, the approach discussed in §3.2.2 and applied in §4.4.1 was followed here. The average number of counts in the observed profiles is large ($>> 10$) and the only assumption made about the variance was that $\sigma_i^2 = \tilde{g}_i$. This assumption is valid under the central limit theorem and can be applied to both simulated data and experimental data (see Chapter 6). The uncertainties in the solution profile, $\hat{f}$, were determined as discussed in §3.3.5 and applied in §4.4.1. The covariance matrix that was used to determine the uncertainties for $\hat{f}$ is generally not a diagonal matrix. The uncertainties were determined using a window vector as described in §3.3.5 and given by (3.58). The window vector incorporates the non-diagonal terms in the covariance matrix for the uncertainties in $\hat{f}$. The window vector was centered about $\hat{f}_i$ and ones were set either side of $\hat{f}_i$ with zeros outside the region of interest. The same approach was also used to determine the uncertainties in the solution strain distribution, $\hat{p}(\epsilon_L)$.

The uncertainties in the solution profile were transformed into uncertainties for the Fourier coefficients. These were interpolated over the region of $d^* \in [0,\, d^*_{max}]$. The uncertainty in the strain coefficients at $d^* = 0$ was set arbitrarily to be small, a value of 0.01 being chosen to avoid problems of division by zero in later calculations of error. The uncertainties in the Fourier coefficients were also used in the matrix kernel form of the MaxEnt method.

The uncertainties in the strain distribution, $\hat{p}(\epsilon_L)$, were determined in the same way as the uncertainties for the solution profile. This enabled the uncertainties in $\langle \epsilon_L^2 \rangle^{\frac{1}{2}}$ to be determined using the approached discussed in §4.4.1.

5.4.2 Case 1: Gaussian strain

In this section the MaxEnt method was applied to determine the specimen profiles and Fourier coefficients for the case of Gaussian strain. The deconvolution for the peaks in the copper pattern (the six most intense lines) was carried out over varying regions dependent on the relative broadening of the simulated profiles.

Determining the specimen profile, $\hat{f}(2\theta)$

There are several observations that can be made about the characteristics of the strain and instrumental broadening for this case. First, the strain broadening increases as $\tan^2\theta$ (compare with (5.23)). The implication of this is that for high angle peaks, such as the 331 and 420, the broadening extends over a large range (about $\pm 10\,°2\theta$). This in turn results in a greater noise component in the simulated observed profiles (compare Figures 5.4(a) & 5.4(b)), because the amplitude decreases (to keep the intensity constant) hence noise is a larger fraction of the signal. Secondly, the instrumental broadening relative to the strain broadening is relatively small as demonstrated by overlaying the specimen and observed profiles. Based on these two observations, the central task of the MaxEnt method in this case involves noise removal. Another difficulty that was encountered was background correction. It was very easy to overestimate the background level and introduce truncation effects into the Fourier coefficients of the solution specimen profile. These factors (the insignificant instrumental broadening, noise level, and incorrect background estimation) contributed to the MaxEnt results summarised in Table 5.1. The solution profiles are shown in Figures 5.6 & 5.7.

We notice from the Table 5.1 the inability of the method to converge onto C_{aim}, with C/C_{aim} ideally ~ 1. The C/C_{aim}, while in the range of unity, are all greater than unity, which suggests the solution profile is under-fitted. The failure of the convergence onto C_{aim} can be accounted for by the above mentioned factors. The noise contribution is likely to be

	111	200	220	311	331	420
C/C_{aim}	2.5	2.0	1.5	2.0	1.5	1.5
R_f (%)	5.5	5.9	6.2	5.2	4.1	4.8
$\beta(°2\theta)$	0.21 ± 0.04	0.29 ± 0.04	0.62 ± 0.09	0.99 ± 0.09	2.9 ± 0.3	3.5 ± 0.4
diff (%)	1.6	< 0.5	6.5	4.3	16.1	20.1

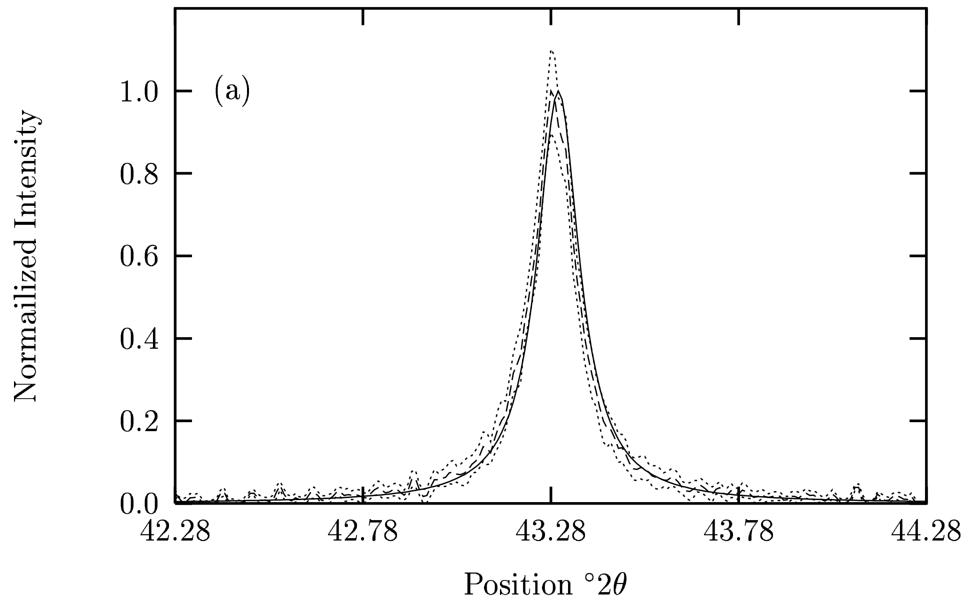
Table 5.1: The parameters and results from applying the MaxEnt method to determine the specimen profile for the Gaussian case: C/C_{aim} is the ratio of the final to the preferred statistic value; R_f is the R-factor from (5.29); $\beta(°2\theta)$ is the integral breadth of the solution profile over 2θ-space; and $diff$ is the percentage difference between the theoretical and calculated integral breadth of the specimen profile.

the main factor. It can be inferred from Figures 5.6 & 5.7 that the noise also affected the solution profile, as spurious oscillations are evident. Despite this the R_f in Table 5.1 are less than 10%.

The large percentage differences in the integral breadths, $\beta(2\theta)$ in Table 5.1 in the 331 and 420 solution profiles can be accounted for by incorrect background estimation. That is, the systematic errors have been propagated from the observed profiles into the 331 and 420 solution profiles.

It should be pointed out that in determining the solution profiles for this case, a uniform model was used in the MaxEnt method, as we have assumed ignorance concerning the specimen profile. The consequence of this choice of an *a priori* model should be pointed out: the choice of the model influences the value of the Lagrangian parameter, α, which in turn influences the solution. This was also demonstrated in Chapter 4, where the choice of *a priori* model played an important role in determining the correct column-length distribution. Moreover, the *a priori* model also has an underlying influence an the magnitude of uncertainties in the solution profiles, since the covariance matrix in (3.58) is determined by the solution profile, $\hat{f}$, and α.

In Figure 5.6(a) the uncertainty region encompasses the solution profile and by visual inspection the region is narrow relative to the $FWHM$ of the solution profile. The greatest uncertainty is about the peak of this distribution. In Figure 5.7(a) the uncertainty region is broad relative to the $FWHM$ of the solution profile. This demonstrated the inadequacy



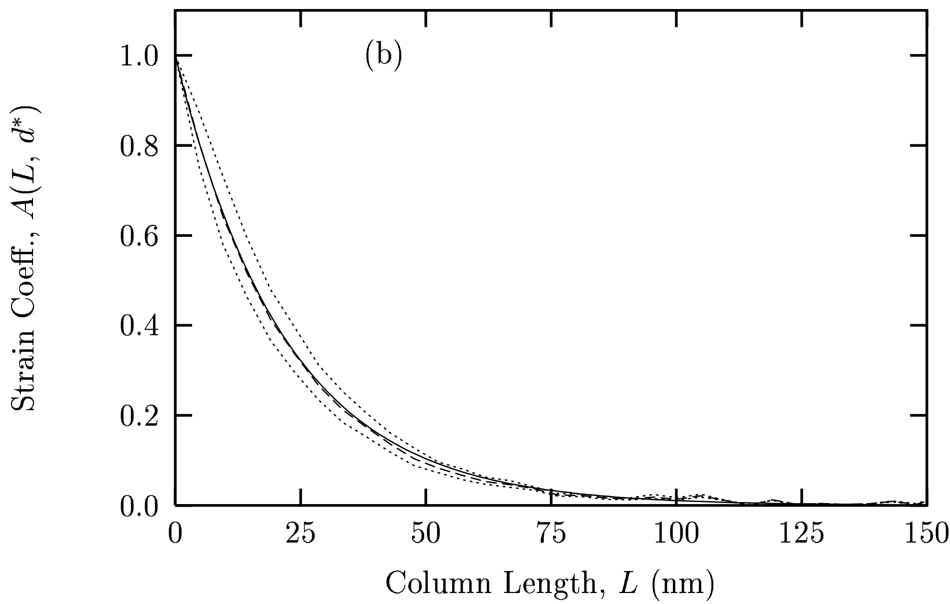
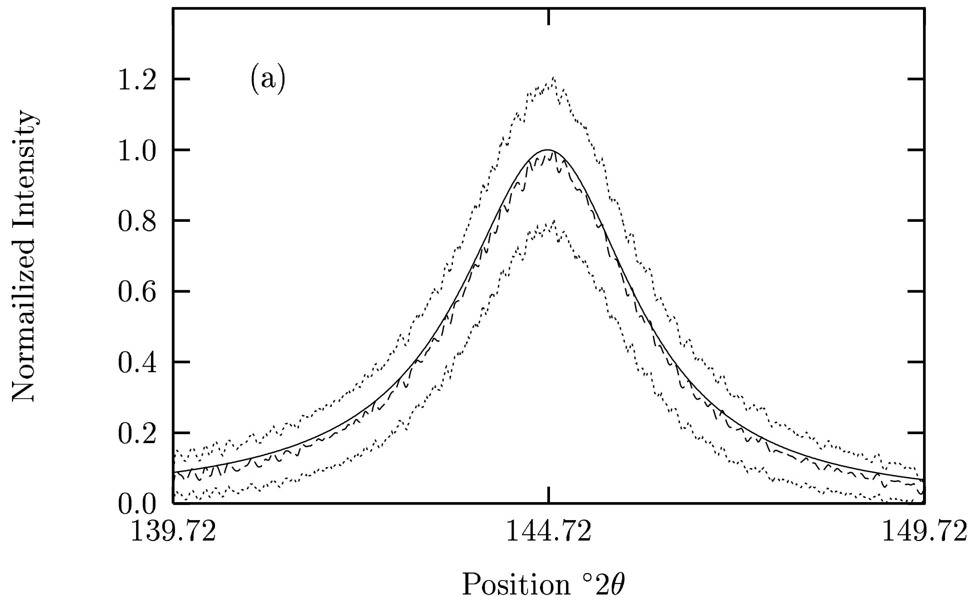


Figure 5.6: MaxEnt results for the 111 peak for Gaussian strain: (a) True 111 specimen profile (solid line), solution specimen profile (dashed line) and lower- & upper-uncertainty regions (dots); (b) Theoretical Fourier coefficients (solid line) using (5.21) for $d^* = 1/d_{111}$, Fourier coefficients from the solution profile (dashed line) given in (a), and lower- & upper-uncertainty regions (dots).

of the *a priori* model. The results in Figures 5.6(a) & 5.7(a) and Table 5.1 represent the worst case scenario in that a uniform *a priori* model was used and no useful information was



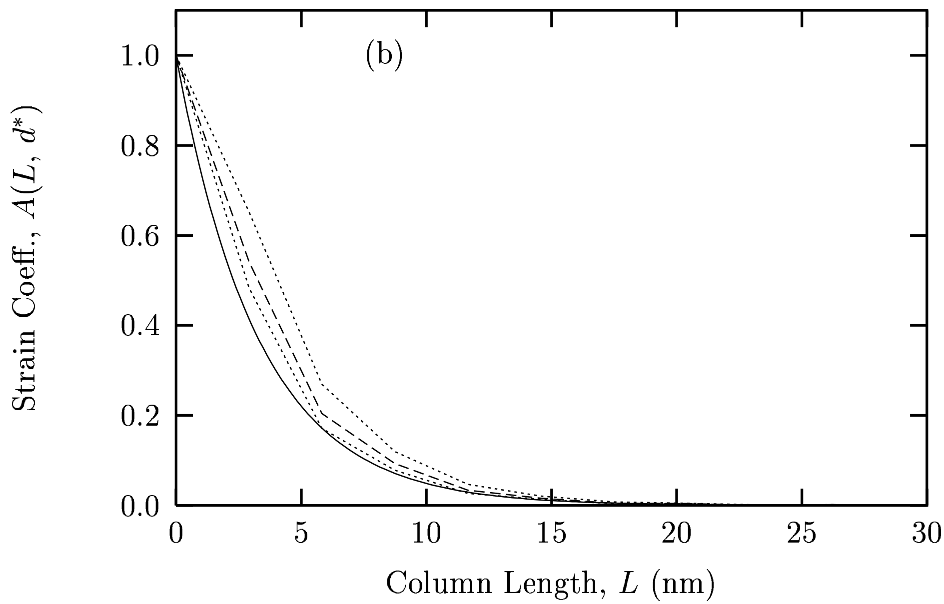


Figure 5.7: MaxEnt results for the 420 peak for Gaussian strain: (a) & (b) as in Figure 5.6 except for $d^* = 1/d_{420}$.

imparted on the solution profile. The Fourier coefficients for the 111 and 420 profiles are given in Figures 5.6(b) & 5.7(b), respectively. In Figure 5.6(b) the Fourier coefficients for the 111 solution profile fit closely with the theoretical coefficients, given by (5.23) for d^*_{111}

over the region of $L \in [0,\, 75]$ nm. Beyond this region there are some oscillations in the tails which are insignificant.

The effect of incorrect background estimation which resulted in the large percentage difference for the integral breadth for the 420 profile is also evident in Figure 5.7(b), shown by the misfitting between the solution Fourier coefficients and the theoretical coefficients, given by (5.23) for d^*_{420}.

Determining $\hat{p}(\epsilon_L)$ – using a uniform model

The Fourier coefficients and their uncertainties for the six profiles were used in determining the strain distribution. This was the second stage of the two fold MaxEnt approach. As a first case the strain distribution was determined using a uniform *a priori* model in the MaxEnt method. It should be pointed out that for all cases involving uniform and non-uniform models, the MaxEnt method converged onto the preferred statistic value, C_{aim}, of 200.

The strain distributions were determined over the range of $L \in [2,\, 20]$ nm. They were also determined over the positive region of ϵ_L, since they are symmetrical about the origin. The kernel matrix, (5.17), was determined as a $[200 \times 200]$ matrix over $\epsilon_L \in [0,\, 4\phi]$ and $d^* \in [0,\, d^*_{max}]$, where ϕ is the period in (5.17) which was found by setting, $\phi = \frac{1}{Ld^*_{max}}$ to be suitable for a given L. The uniform model was defined by

$$m_p = \frac{1}{\epsilon_{L\,max} - \epsilon_{L\,min}} \tag{5.31}$$

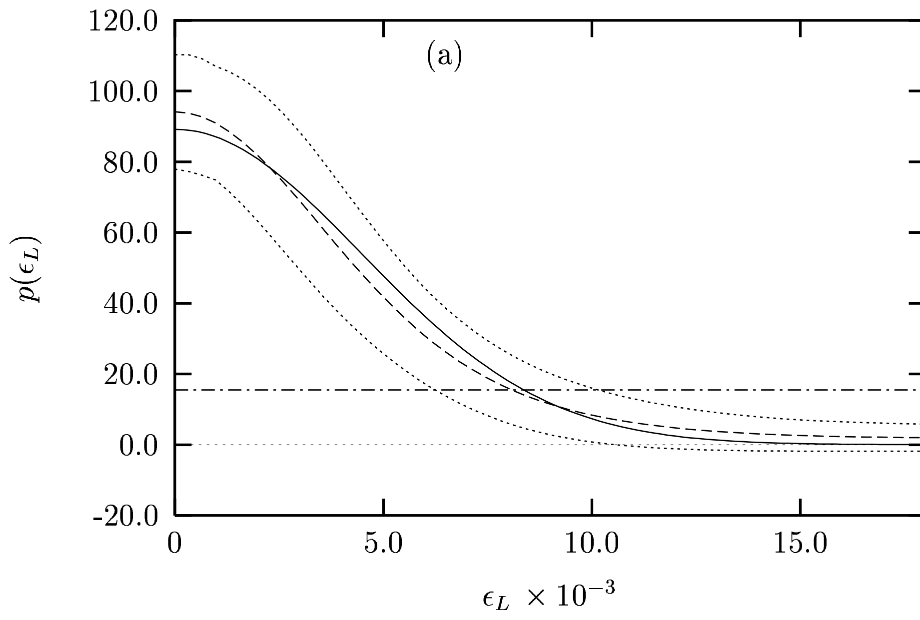
where $\epsilon_{L\,min} = 0$ and $\epsilon_{L\,max} = 4\phi$. The normalising of the solution distribution and determining the rms-strain was carried out over the region $\epsilon_L \in [0,\, 2FWHM]$, where the $FWHM$ is the full width at half maximum of the solution distribution[3].

The MaxEnt strain distributions are shown Figure 5.8, for $L = 5,\, 13,\, \&\, 20$ nm over the region $\epsilon_L \in [0,\, 2FWHM]$. These solutions represent the probability density on the strain at a column-length distance, L, parallel to the diffraction vector. The distribution assigns a probability to the elongation of the column-length as a result of the ensemble of dislocations throughout the crystallite.

[3]This is effectively integrating over $[0,\, 4\sigma]$ and in the Gaussian limit corresponds to a confidence limit of greater than 99.5%.

The solution distributions given in Figure 5.8 demonstrate how ineffectual the uniform model, (5.31), is in the MaxEnt method. There is a misfitting of the solution distribution about the origin, and the tails are extended, rather than decreasing rapidly as expected for a Gaussian distribution. These results show that the uniform model is not well suited for determining the strain distribution, but it does provides an understanding of the solution distribution. The above features are evidence that a better *a priori* model is needed. The use of the uniform model represents our lack of knowledge about the solution. This was demonstrated by the broad uncertainty region about the solution distribution and was reflected in the uncertainty of the $FWHM$ of the solution distributions in Figure 5.8.

Although the solution for a uniform model represents our worst case, the positivity has been preserved, and there are no spurious oscillations in the tails of the solution. This is somewhat surprising given the difficulties in the deconvolution and misfitting in the Fourier coefficients (see Figure 5.6(b)). They appear not to have contributed to the ill-conditioning of the solution distribution.



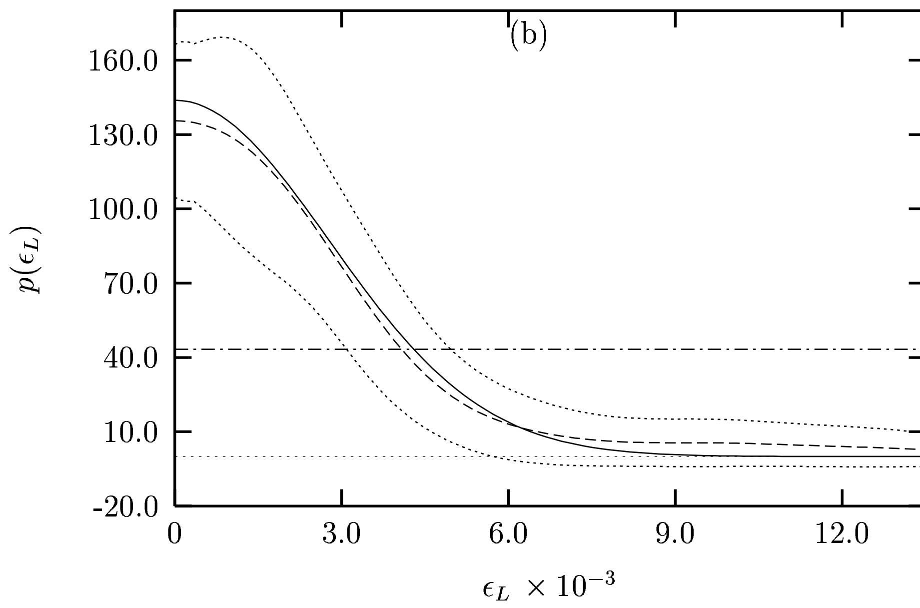


Figure 5.8 continues over...

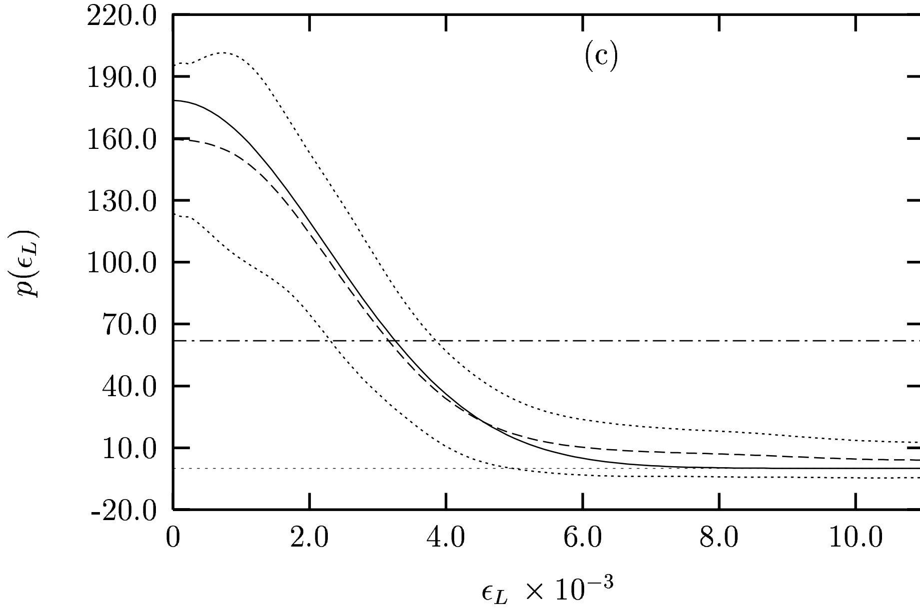


Figure 5.8: The MaxEnt strain distributions for Gaussian strain using a uniform *a priori* model for $L =$ 5, 13 & 20nm over the region of $\epsilon_L \in [0,\ 2FWHM]$: (a) True strain distribution (solid line) for $L = 5\,$nm, the solution strain distribution (dashed line), lower- & upper-uncertainty region (dots), and the uniform *a priori* model (dash-dot lines); (b) Solution distribution for $L = 13\,$nm; (c) Solution distribution for $L = 20\,$nm. Note: The distributions showing the lower- & upper-uncertainty regions (dots) are not normalised for unit area.

Determining $\hat{p}(\epsilon_L)$ – non-uniform model

The features of the solution distribution, shown in Figure 5.8, are a consequence of using a uniform *a priori* model. In order to improve these results a non-uniform *a priori* model was developed; that is, a model determined from the observed data using an independent or low resolution method. In this case we have formulated a non-uniform model using a discrete form of (5.5) given as

$$m_{pj} = 2L\Delta d^* \sum_{i=1}^{N} \hat{A}_i \cos(2\pi L \epsilon_{Lj} d_i^*) \tag{5.32}$$

where $\Delta d^* = d^*_{i+1} - d^*_i$ and was determined over the region of $\epsilon_L \in [0,\ \frac{4}{L\,d^*_{max}}]$ and $d^* \in [0,\ d^*_{max}]$. As discussed above, this approach does not preserve the positivity of the model and it was found that negative spurious oscillations appeared in the non-uniform model. In order to overcome this, the non-uniform model was fitted using a "tunable" function such as

a Pearson-VII , (5.24), to 1% of its amplitude. The parameters from the fitted model were used to generate a model which was positive definite. This fitted model was then used in the MaxEnt method. The MaxEnt method was applied over the reduced region of $2FWHM$s of the fitted model.

The solution strain distributions, using a non-uniform model, were determined over the range $L \in [2, 20]$ nm. In Figure 5.9, the solution distributions for $L = 5$, 13 & 20 nm are given. By comparing these results with Figure 5.8, an improvement in the solution distribution is noticeable. The extended tails and misfitting that are present in Figure 5.8 have been reduced. The uncertainty region for these sets of solution distributions is narrower compared with the distributions in Figure 5.8, while tapering off for increasing ϵ_L. The solution distribution for $L =$ 5nm, Figure 5.9(a), has the greatest misfitting for the solution using a non-uniform model. The reason for this is in the poor quality of the Fourier coefficient data at this column length, L. Generally the positivity of the solutions has been preserved, and the misfitting between the solution and true distributions has been reduced. This clearly demonstrates the influence a non-uniform *a priori* model has upon the solution distribution and uncertainty region. This last point is important, in that it shows the influence of the *a priori* model on the Lagrangian multiplier, α, in (5.15), and on the uncertainties of the solution distribution.

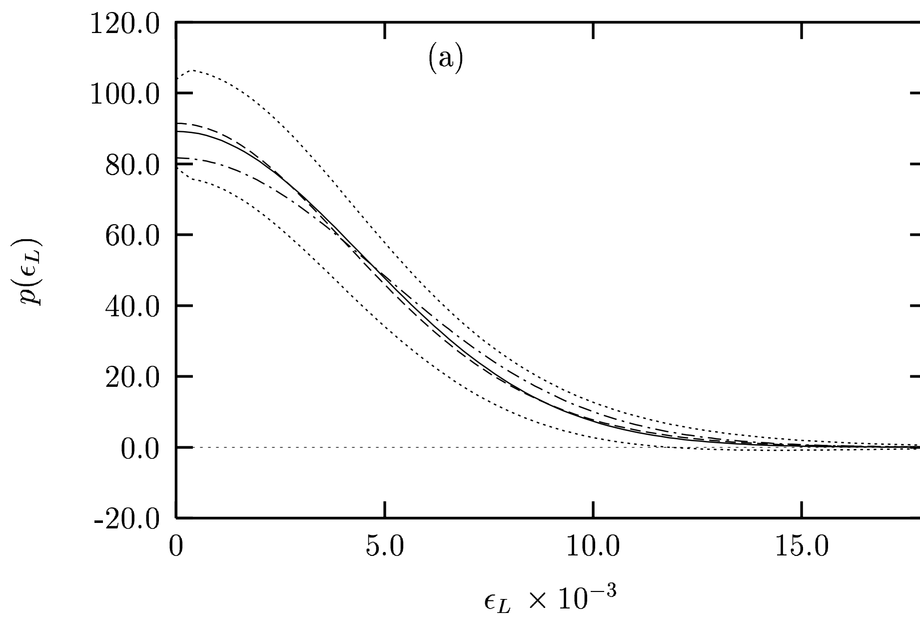


Figure 5.9 continues over...

5.4.3 Case 2: Non-Gaussian strain

The purpose of this case is to examine the MaxEnt two fold approach in the limit of non-Gaussian strain. Furthermore, it is important to develop an understanding of the factors that may affect the MaxEnt method. As in the Gaussian case for uniform and non-uniform models, the MaxEnt method converged onto the preferred statistic value, C_{aim}, of 200.

In determining the solution specimen profile and strain distribution, the same procedure that was applied in the previous section was applied here.

Determining $\hat{f}(2\theta)$

The shapes of the specimen profiles are very different in this case. The tails are extended and the peak is narrow. Table 5.2 summarises the performance of the MaxEnt method in determining the specimen profiles. As was found in §5.4.2, the MaxEnt method was unable to converge onto the preferred C_{aim}, but $C/C_{aim} \sim 1$. To illustrate the MaxEnt solution profiles, the low-angle 111 and high-angle 420 profiles are given in Figures 5.10 & 5.11, respectively.

	111	200	220	311	331	420
C/C_{aim}	2.75	2.0	1.5	1.67	1.5	1.45
R_f (%)	28.8	17.4	15.4	17.4	8.8	3.5
$\beta(°2\theta)$	0.11 ± 0.04	0.13 ± 0.04	0.31 ± 0.07	0.41 ± 0.10	1.2 ± 0.3	1.5 ± 0.3
$diff$ (%)	23.7	8.4	10.8	5.1	18.3	15.9

Table 5.2: Parameters and results from applying the MaxEnt method to determine the specimen profile for the non-Gaussian case: C/C_{aim} is the ratio of the final to the preferred statistic value; R_f is the R-factor from (5.29); $\beta(°2\theta)$ is the integral breadth of the solution profile over 2θ-space; and $diff$ is the percentage difference between the theoretical and calculated integral breadths of the specimen profile.

The solution profile for the 111 peak is given in Figure 5.10. The deconvolution was carried out over a region of $(2\theta \pm 1)\,°2\theta$ at a step-size of $0.01\,°2\theta$. The C_{aim} was set to a value of 200. The main region of the profile has not been correctly resolved and on close inspection the uncertainty regions of the solution profile do not include the true specimen

profile. Also the spurious oscillations in the tails of the solution are an indication that the noise in the observed profile has not been fully removed. The Fourier coefficients of the solution profile fit closely to the true coefficients up to $\sim 60\,\mathrm{nm}$. Beyond this region the misfitting between the coefficients increases and the amplitude of the spurious oscillations increases, which implies that the oscillations in the solution profile are being imparted into the Fourier coefficients.

The results for the 420 specimen profile, given in Figure 5.11(a), are somewhat more encouraging, in that the MaxEnt method has been able to resolve the features of the profile: the amplitude of the spurious oscillation have been reduced; and the uncertainty region of the solution profile encompasses the true solution profile. Although the uncertainty region is broad, which is also indicated by the $FWHM$ of the solution profile, the misfitting of the solution and true specimen profile has been reduced. Overall the results are reasonably good. However, the effects of truncation error are noticeable in the Fourier coefficients, given in Figure 5.11(b). The misfitting between the theoretical and solution Fourier coefficients is considerable and for small L the Fourier coefficient errors fail to encompass the theoretical coefficients. However, unlike the Fourier coefficients in Figure 5.10(b), the are no spurious oscillations.

In the light of these results and the results in §5.4.2 an attempt was made to develop a non-uniform model for determining the specimen profile, but proved to be unsuccessful. Developing the *a priori* model involved smoothing the (background corrected) observed profile using a suitable smoothing (i.e. convolving) distribution and using it in the MaxEnt method. The rationale for smoothing the observed profile was that the resulting model would contain the underlying specimen profile and the process of smoothing the observed profile would reduce the noise in case this produced any ill-conditioning in the solution profile. The instrument profile was used as the smoothing distribution, since there was no theoretical justification for selecting any other smoothing distribution. However, this approach was found to have little influence on the solution profile. The reason for this was the additional broadening in the model. Additional calculations showed that the non-uniform model had to be very "close" to the true specimen profile for any noticeable improvement in the solution profile. However, these models were not used since there is no theoretical or

physical justification for doing so.

From the attempts to develop a non-uniform model for the strain distributions, it seems plausible that a similar approach could be applied to developing a non-uniform model for determining $\hat{f}$. This would consist of applying another deconvolution method, such as the forward deconvolution approach (Howard & Snyder 1989, Cheary & Coelho 1992) since this approach attempts to preserve the positivity of the solution profile, and information about the specimen can be incorporated. This result could then be used as the model in the MaxEnt method.

Determining $\hat{p}(\epsilon_L)$ using a uniform model

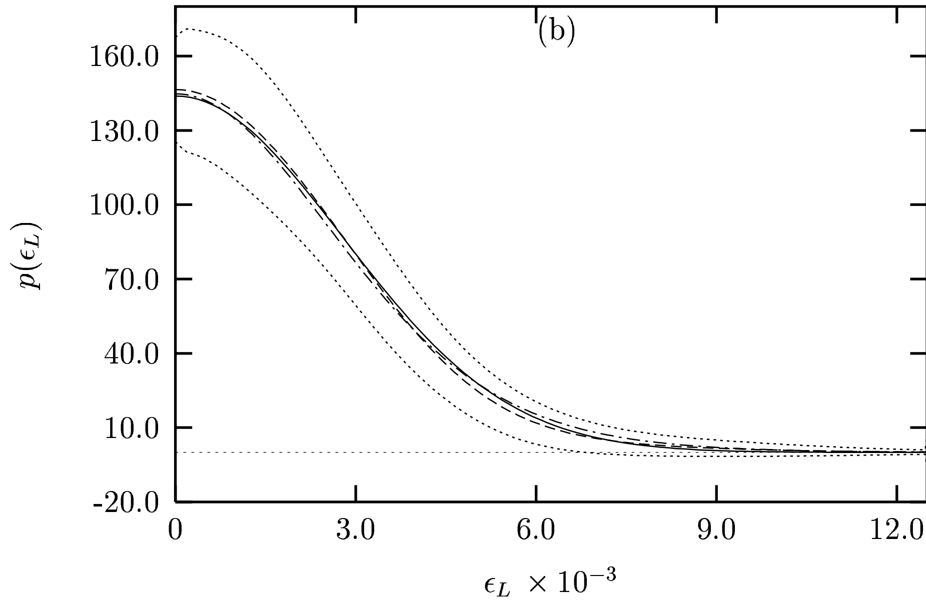
The uniform model was applied in the first case for determining the strain distribution in the near-Lorentzian limit. The procedure applied in §5.4.2 was applied in this section and the uniform model was determined using (5.31). The matrix kernel was evaluated over the region of $\epsilon_L \in [0,\ 5\phi]$ and $d^* \in [0,\ d^*_{max}]$. The normalisation of the solution distribution was carried out over the region of four $FWHM$s of the solution distribution. However, in the solution distribution for $L \geq 46\,\text{nm}$, the integration was applied over $\epsilon_L \in [0,\ 5\phi]$ and was found to be better suited because of the extended tails of the distribution solution.

The solution distributions were determined over the region of $L \in [5,\ 50]\,\text{nm}$. Figure 5.12 shows the strain for the particular column lengths, $L = 10,\ 30,\ \&\ 50\,\text{nm}$. The extended region for L comes about because of the long tails in the Fourier coefficients. For example, compare the Fourier coefficients in Figures 5.7(b) & 5.11(b); the range over which 5.11(b) is plotted is ~ 3 times as great compared to Figures 5.7(b). This increases the working range of the MaxEnt method for determining the strain distribution.

The main difficulty in the result for the strain distribution is the misfitting of the solution distribution about the origin of the distribution. In Figure 5.12(a) the solution distribution fails to fit the true distribution about the origin. The failure is due to the use of uniform *a priori* model. That is, no information about the scale of the distribution and its shape has been specified. There is also some misfitting in the tails of the strain distribution, which is again due to the lack of information about the shape of the strain distribution in the model. In using a uniform model, the solution distribution and its uncertainties represent

the worst case situation. This is evident by visual inspection of the $FWHM$ in the solution distributions. The use of the uniform model influences the value of the Lagrangian multiplier, which in turn influences the magnitude of the uncertainties. This can be seen in examining (3.57). The uncertainty region for the solution strain distribution in Figures 5.12 (b) & (c) encompasses the true distribution. However, in Figure 5.12 (a) the uncertainty region does not encompass the central region of the distribution due to the lack of information in the uniform *a priori* model.

Although the use of a uniform *a priori* model produces misfitting of the solution distribution with the true distribution, it defines the worst case scenario. This situation is also reflected in the broad uncertainty region of the solution distribution.



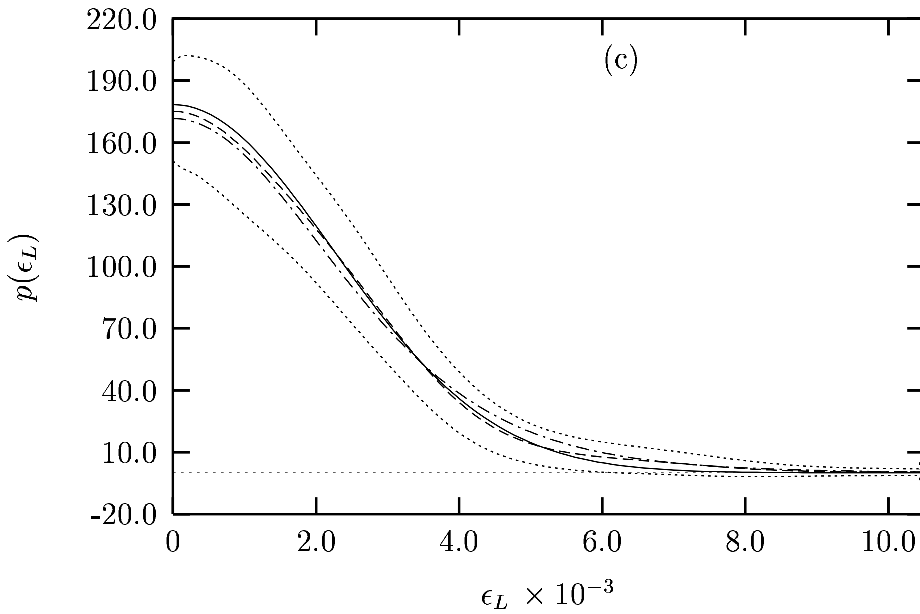
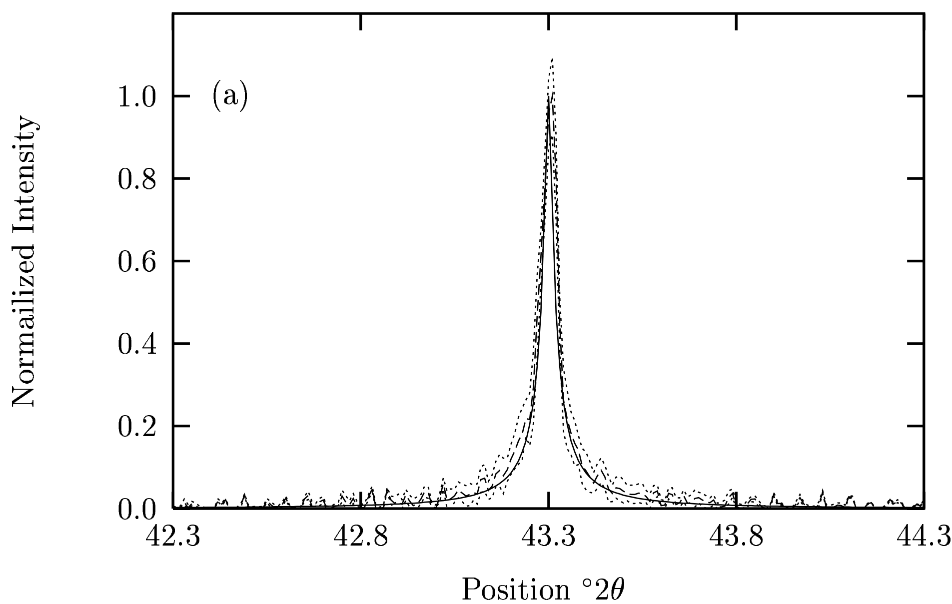


Figure 5.9: MaxEnt strain distributions for Gaussian strain using a non-uniform *a priori* model for $L = 5$, 13 & 20nm over the region of $\epsilon_L \in [0, 2FWHM]$: (a) True distribution (solid line) for $L = 5$ nm, the solution strain distribution (dashed line), lower- & upper-uncertainty region (dots), and the non-uniform *a priori* model (dash-dot lines); (b) Solution distribution for $L = 13$ nm; (c) Solution distribution for $L = 20$ nm. Note: The distributions showing the lower- & upper-uncertainty regions (dots) are not normalised for unit area, while the solution distribution and non-uniform *a priori* model are normalised for unit area.



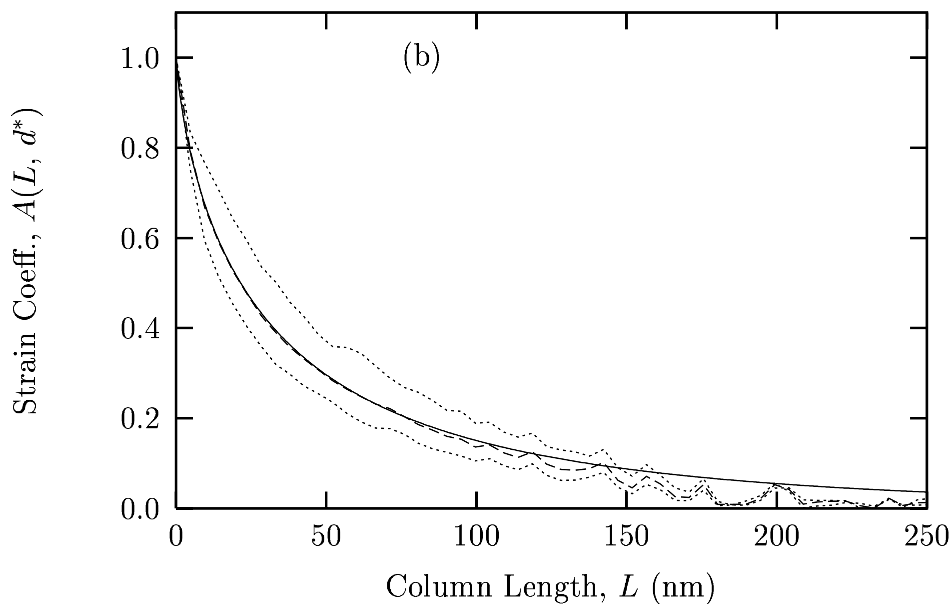
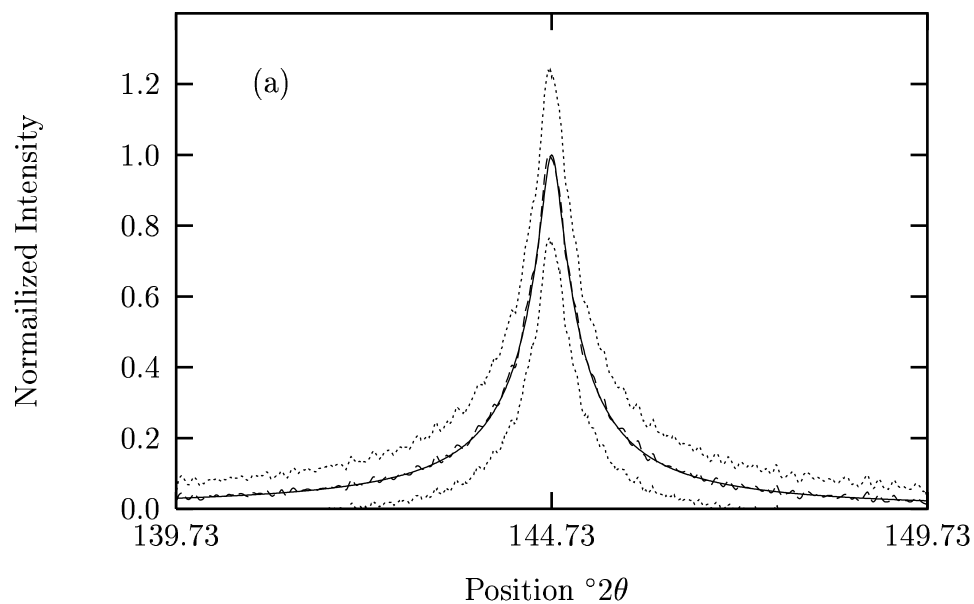


Figure 5.10: MaxEnt results for the 111 peak for non-Gaussian strain: (a) True 111 specimen profile (solid line), solution specimen profile (dashed line) and lower- & upper-uncertainty regions (dots); (b) Theoretical Fourier coefficients (solid line) using (5.21) for $d^* = 1/d_{111}$, Fourier coefficients from the solution profile (dashed line), given in (a) and lower- & upper-uncertainty regions (dots).



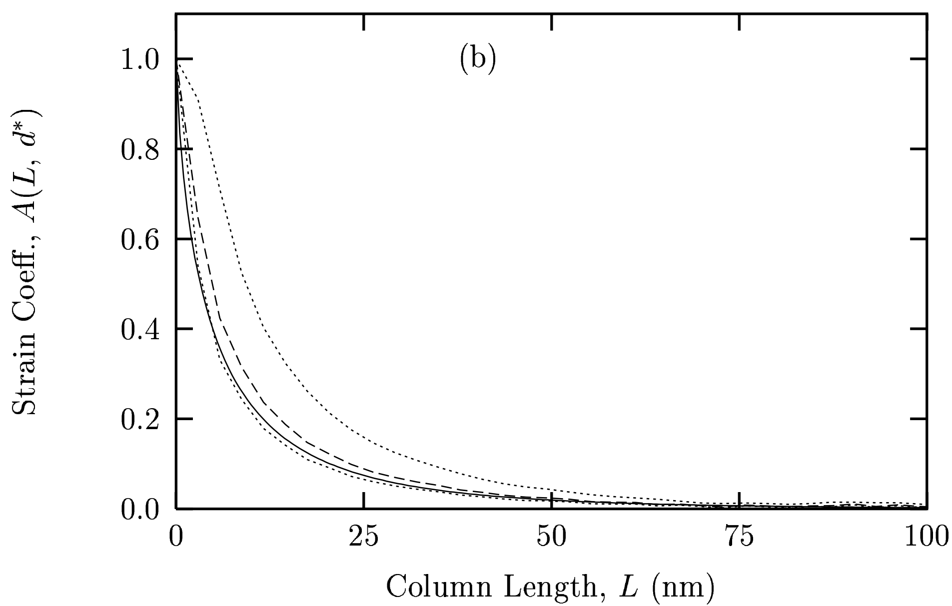
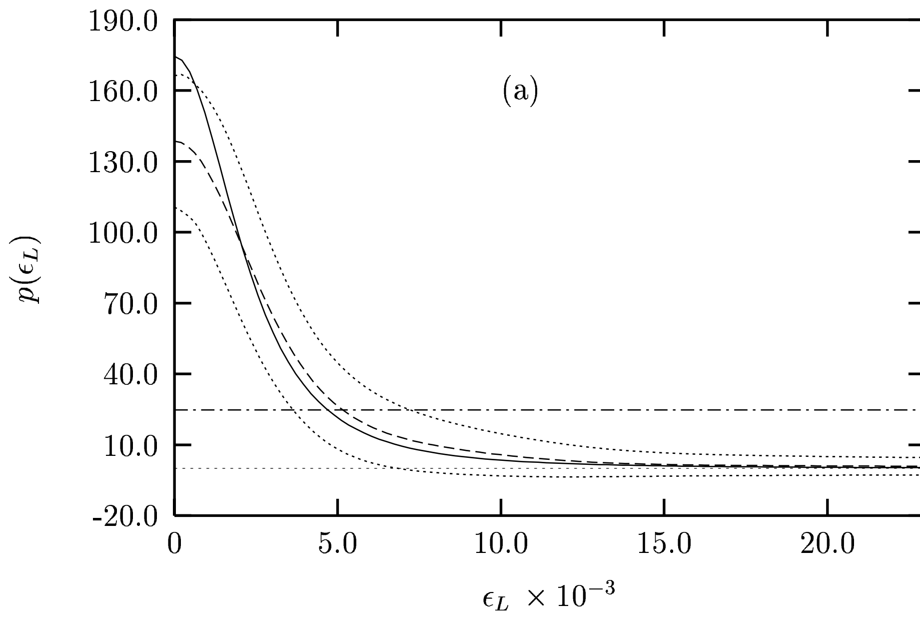


Figure 5.11: MaxEnt results for the 420 peak for Gaussian strain: (a) & (b) as in Figure 5.10 except for $d^* = 1/d_{420}$.



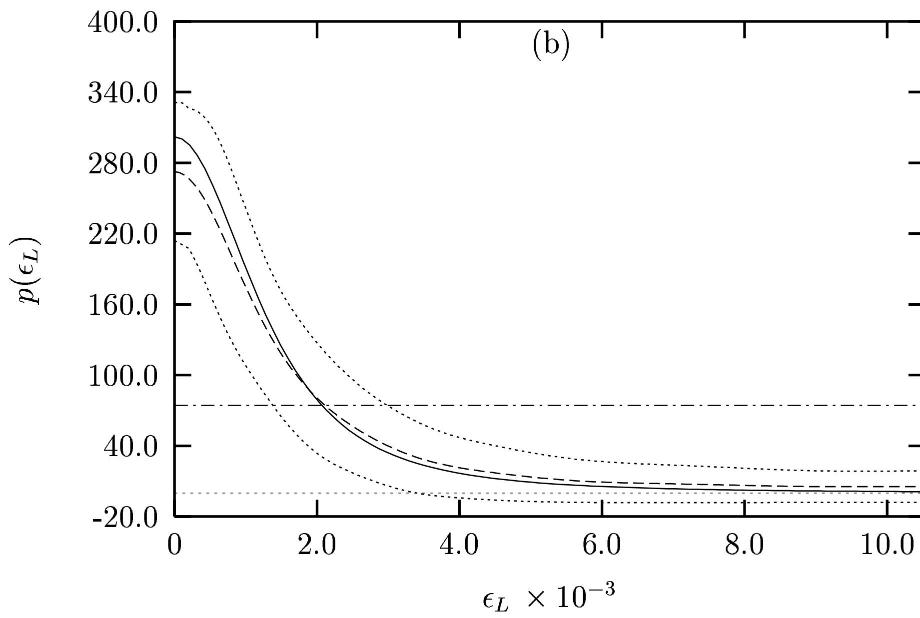


Figure 5.12 continues over...

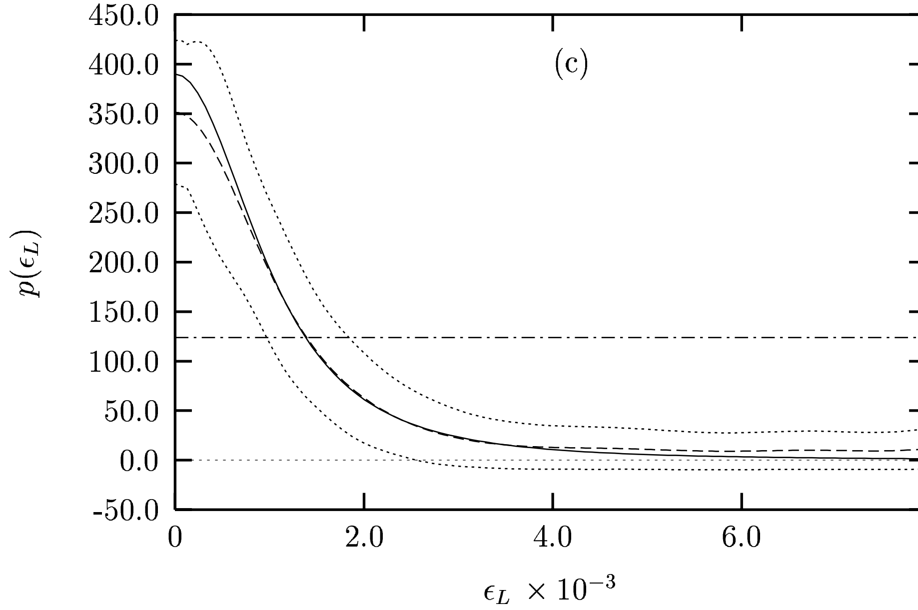


Figure 5.12: MaxEnt strain distributions for non-Gaussian strain using a uniform *a priori* model for $L = 10, 30$ & 50nm over the region of $\epsilon_L \in [0,\ 2FWHM]$: (a) True distribution (solid line) for $L = 10$ nm, the solution strain distribution (dashed line), lower- & upper-uncertainty region (dots), and the uniform *a priori* model (dash-dot lines); (b) Solution distribution for $L = 30$ nm; (c) Solution distribution for $L = 50$ nm. Note: The distributions showing the lower- & upper-uncertainty regions (dots) are not normalised for unit area.

Determining $\hat{p}(\epsilon_L)$ using a non-uniform model

For the case of non-Gaussian strain, a non-uniform model was developed using the same procedure as outlined in §5.4.2 with the exception that the model parameters were fitted to 5% of the amplitude. The reason for this is the large spurious oscillations in the extended tails in the model. The strain distributions were determined over the region of $L \in [10,\ 50]$ nm. The reason for the reduced region compared with the results in §5.4.3, was the of large spurious oscillations in the model using (5.32). Consequently, difficulties were encountered in fitting the Pearson-VII function, (5.24). The matrix kernel for this case was evaluated over four $FWHM$s of the non-uniform *a priori* model. The normalising and subsequent calculations were also carried out over four $FWHM$s of the model.

Figure 5.13 shows the solution distribution for $L = 10,\ 30$ & 50 nm. In general there has been an improvement in the solution distribution compared with the results from the uniform

model. For all cases in Figure 5.13, the largest uncertainties occur about the origin of the strain distribution. The reason for the large uncertainties in the region $\epsilon_L \sim 0$, is due to the information that the *a priori* model is imparting about the scale of the distribution. In all cases in Figure 5.13, the model is unable to correctly determine the amplitude of the strain distribution. This is evident in Figure 5.13(a). In Figures 5.13 (b) & (c), the amplitude of the model approaches the true distribution. This enables the solution distribution to converge onto the true distribution. However, the uncertainty region for $\epsilon_L \sim 0$ is relatively broad.

On the other hand, compared with the results for the uniform model, the improvement in the results is clearly demonstrated. This is particularly evident in the uncertainty region for the $FWHM$ of the solution distribution. The uncertainty region is narrow. Again as discussed in the above solutions, the influence of the *a priori* model on α, on $\hat{p}(\epsilon_L)$, and its uncertainties must not be underestimated.

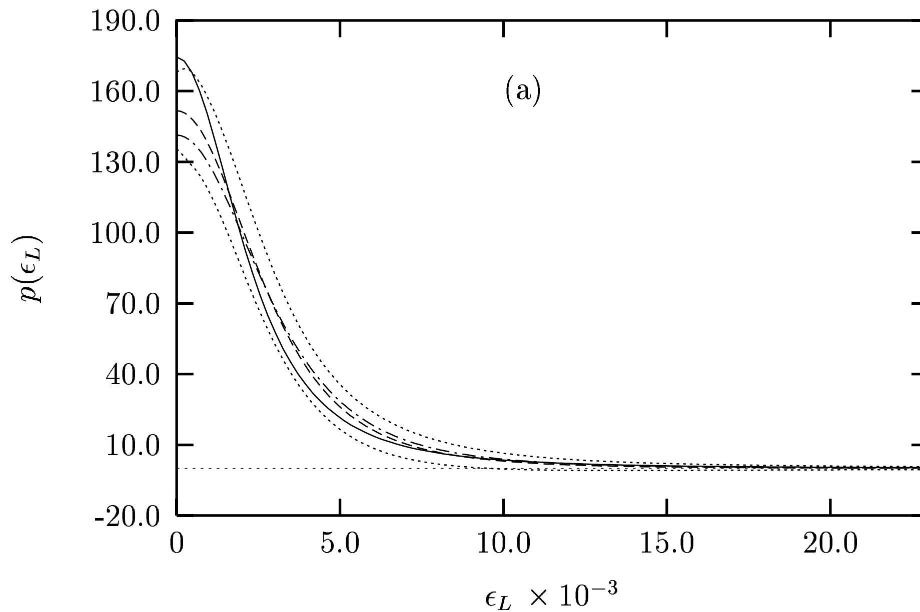


Figure 5.13 continues over...

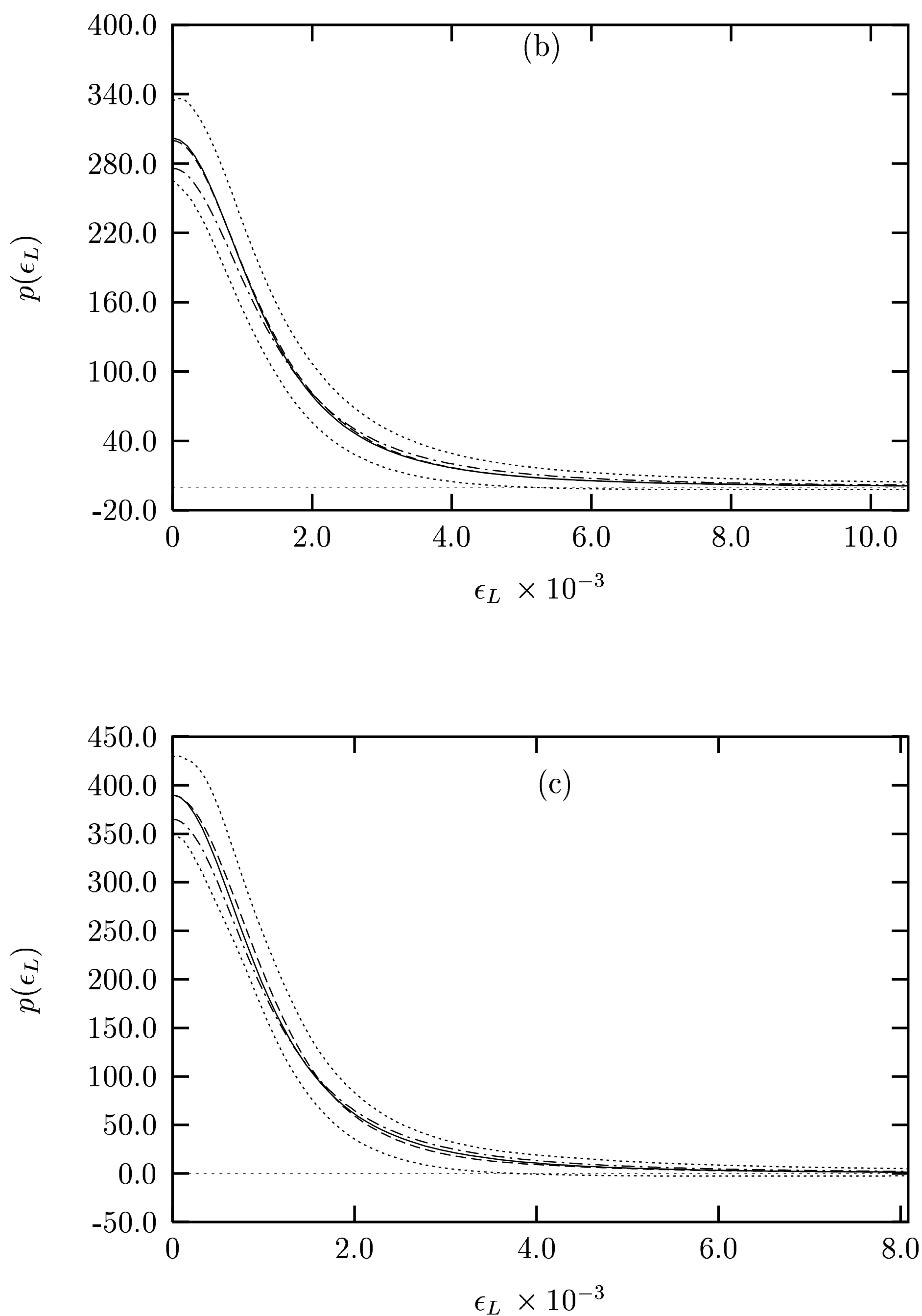


Figure 5.13: MaxEnt strain distributions for non-Gaussian strain using a non-uniform *a priori* model for $L = 10, 30 \,\&\, 50$nm over the region of $\epsilon_L \in [0,\, 2FWHM]$: (a) True distribution (solid line) $L = 10\,$nm, the solution strain distribution (dashed line), lower- & upper-uncertainty region (dots), and the non-uniform *a priori* model (dash-dot lines); (b) Solution distribution for $L = 30\,$nm; (c) Solution distribution for $L = 50\,$nm. Note: The distributions showing the lower- & upper-uncertainty regions (dots) are not normalised for unit area, while the solution distribution and non-uniform *a priori* model are normalised for unit area.

5.4.4 Strain Analysis

In the previous section we demonstrated the two-fold MaxEnt method for determining the specimen profile and the strain distribution. In both cases the uncertainties for the resulting solution profile and distribution were determined. In this section, physical quantities, such as the root-mean-square strain, $\langle \epsilon_L^2 \rangle^{\frac{1}{2}}$, and parameters defining the strain models are determined. These quantities are compared with the standard approaches such of Williamson & Hall (1953) and Warren & Averbach (1950). The uncertainties from the solution profiles and strain distributions were used to determine the uncertainties in the subsequent quantities. This again demonstrates that the MaxEnt method is a fully quantitative method, in that the results and uncertainties from the method can be used in subsequent results.

Integral breadth analysis

The integral breadths of the six most intense profiles for the Gaussian and non-Gaussian cases are shown in Figure 5.14. In both the Gaussian and non-Gaussian cases the classical Williamson & Hall method fails to produce physically consistent results. This is evidenced by the negative intercept of the linear fit. The intercept in the Williamson & Hall method is directly related to the volume-weighted crystallite size (see Chapter 1). The last two points in Figure 5.14(a) & (b), corresponding to the 331 and 420 profiles, are likely to have been affected by systematic errors such as background estimation and truncation. This was pointed out in Tables 5.1 & 5.2. When the 331 and 420 solution profiles were omitted in the linear fitting, a negative intercept was also produced. This further adds weight to the idea that the linear model for strain broadening is not suitable. Unphysical terms such as a negative intercept and consequently unphysical crystallite size still result when eliminating the influence of systematic error in the linear estimation. The reason for this failure is in the underlying assumptions of the Williamson & Hall method. That is, the strain distribution is assumed to be independent of L, resulting in a linear relationship between integral breadth, crystallite size (given by the inverse of the intercept) and strain (proportional to the slope). In general this is not going to be case.

The difference between the theoretical curves, given in §5.3, and the data points in Figure 5.14 for the 331 and 420 lines can only be accounted for by truncation and background

effects. For both cases the first four data points lie close to or on the theoretical curve.

For both cases, the data was found to follow a d^{*2} relationship which produced a positive slope and a nonzero, but positive intercept. Unlike the Williamson & Hall method, plotting β as a function of d^{*2} produces a physically consistent result. For the Gaussian case, given in Figure 5.14(a), data points were found to fit

$$\beta(d^*) = \pi^2 G^2 d^{*2} + c \tag{5.33}$$

where G^2 is the estimated strain constant, independent of hkl, and which defines the mean-square strain in terms of L, see (5.19); c is the intercept and has the units of 1/length. (In Figure 5.14(a), (5.33) has been plotted against d^* rather than d^{*2} for convenience.) It is tempting to relate c to a crystal size dimension, but this result should be interpreted conservatively, the reason being that c is likely to be strongly influenced by systematic errors such as truncation and background effects discussed above. In the context of these simulations, the theoretical value of c is zero (see (5.23)).

	$WH^{(1)}$	$WH^{(2)}$	$WA^{(1)}$	$WA^{(2)}$	$MaxEnt^{(1)}$	$MaxEnt^{(2)}$
$G^2 \times 10^{-5}$ (nm)	7.7 ± 0.7	9 ± 1	6 ± 2	9.1 ± 0.4	14.0 ± 0.4	11.2 ± 0.4
diff (%)	23	10	40	9	40	12
$c \times 10^{-3}$ (nm)$^{-1}$	8 ± 4	1 ± 5	–	–	–	–

Table 5.3: Estimated parameters using the nonlinear Williamson & Hall (WH), Warren & Averbach (WA) and MaxEnt methods: $WH^{(1)}$ G^2 & c estimated for all data points in Figure 5.14(a); $WH^{(2)}$ G^2 & c estimated for the first four data points in Figure 5.14(a); $WA^{(1)}$, G^2 estimated over the region $L \in [1, 34]$ nm from Figure 5.15(a); $WA^{(2)}$, G^2 is estimated over the region $L \in [5, 22]$ nm from Figure 5.15(a); $MaxEnt^{(1)}$, G^2 from Figure 5.15(b) using a uniform model in the MaxEnt method; $MaxEnt^{(2)}$, G^2 from Figure 5.15(b) using a non-uniform model in the MaxEnt method.

Table 5.3 shows the estimated G^2 values for all the methods i.e. Williamson & Hall (more specifically, the G^2 and c results for the non-linear Williamson & Hall, given by (5.33)), Warren & Averbach and the MaxEnt method. In the first case (see $WH^{(1)}$) all six points were used. In the second case (see $WH^{(2)}$) only the first four points were used. By comparing the two cases the effect of systematic errors on the estimation of G^2 and c is

evident. For both cases the uncertainty in c is larger, in the first case it is $\sim 50\%$, while in the second case it is larger than the estimated quantity.

In the non-Gaussian case, shown in Figure 5.14(b), a similar d^{*2} relationship was found. However, the slope of this graph can not be directly related to the parameters q, m and σ_0 from the strain distribution given by (5.24). The slope is a function of these parameters and there are too few equations to determine them. At best we compare the slope of the fitted data with its theoretical value for the values of q, m and σ_0. This fitting was found to be

$$\beta(d^*) = \psi(\sigma_0, m, q)\, d^{*2} + c \tag{5.34}$$

where $\psi(\sigma_0, m, q)$ is the function dependent on q, m and σ_0 and c is the intercept. Using (5.34), the slope ψ was determined as $(3.1 \pm 0.6) \times 10^{-4}\,\text{nm}$. This can be compared to the theoretical value of $4.1 \times 10^{-4}\,\text{nm}$, corresponding to a percentage difference of $\sim 25\%$. The intercept, c, from (5.34) was found to be $(6 \pm 4) \times 10^{-3}\text{nm}^{-1}$, with another large uncertainty, almost 67%. As was the situation in the Gaussian case, the size of the uncertainty suggests that random errors are not the sole cause, and that systematic errors such as truncation and background effects are the most likely culprits. When the integral breadths for the 331 and 420 profiles are not included in the above analysis, ψ was found to be $(4 \pm 1) \times 10^{-4}\,\text{nm}$ and the intercept, c, found to be $(3 \pm 5) \times 10^{-3}\,\text{nm}^{-1}$. As in the Gaussian case, this clearly indicates that systematic errors are the cause with the poor estimate in ψ.

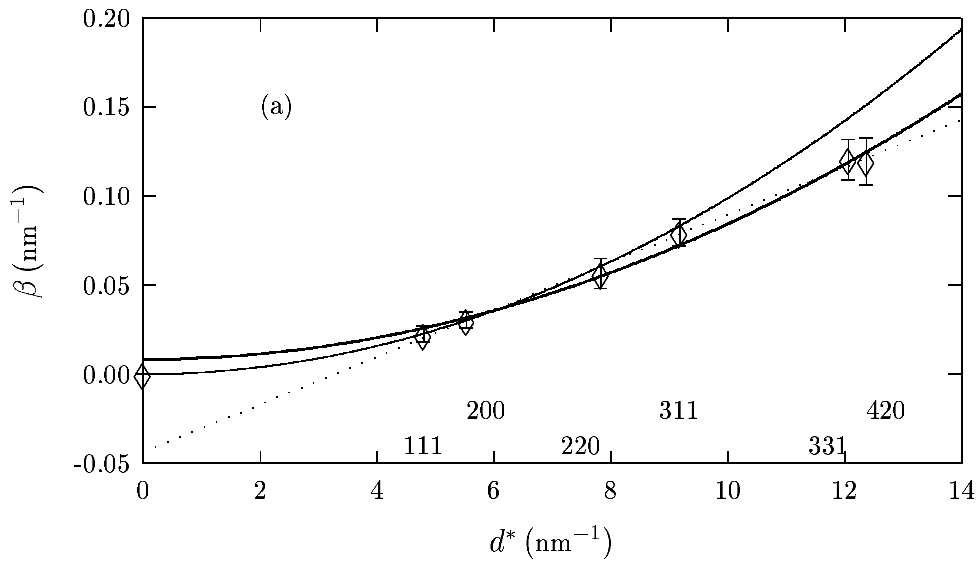
Determining $\langle \epsilon_L^2 \rangle^{\frac{1}{2}}$ distribution — Gaussian case

The Warren & Averbach method was applied to the Fourier coefficients of the solution profile in the Gaussian case. It involved separating the strain and size coefficients by applying

$$\ln A(L,\, d^*) = \ln A^s(L) - 2\,\pi^2\, L^2\, d^{*2}\, \langle \epsilon_L^2 \rangle \tag{5.35}$$

where $\langle \epsilon_L^2 \rangle$ is the mean-square strain for a given L, which can be determined from the slope of the graph $\ln A(L,\, d^*)$ versus d^{*2}. The intercept of (5.35) can be used to determine the size Fourier coefficients and the area weighted size, $\langle L \rangle_a$ using (4.2). The $\langle \epsilon_L^2 \rangle^{\frac{1}{2}}$ distribution can be determined from (5.35) as shown in Figure 5.15(a).

Figure 5.15(a) shows the variation of the strain over a range of L. For small L, the uncertainties in $\langle \epsilon_L^2 \rangle^{\frac{1}{2}}$ are large and there is also considerable misfitting between $L = 1$ to $4\,\text{nm}$.



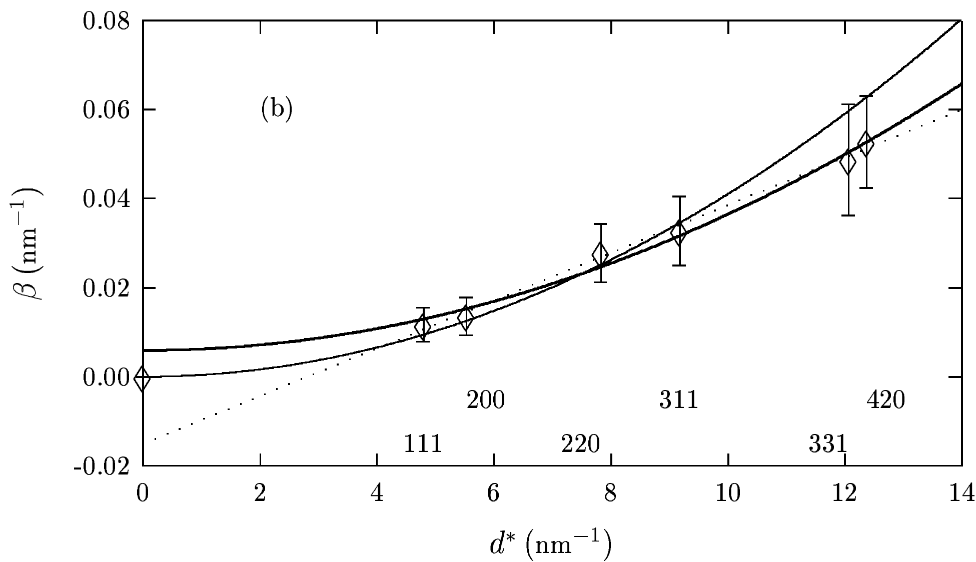
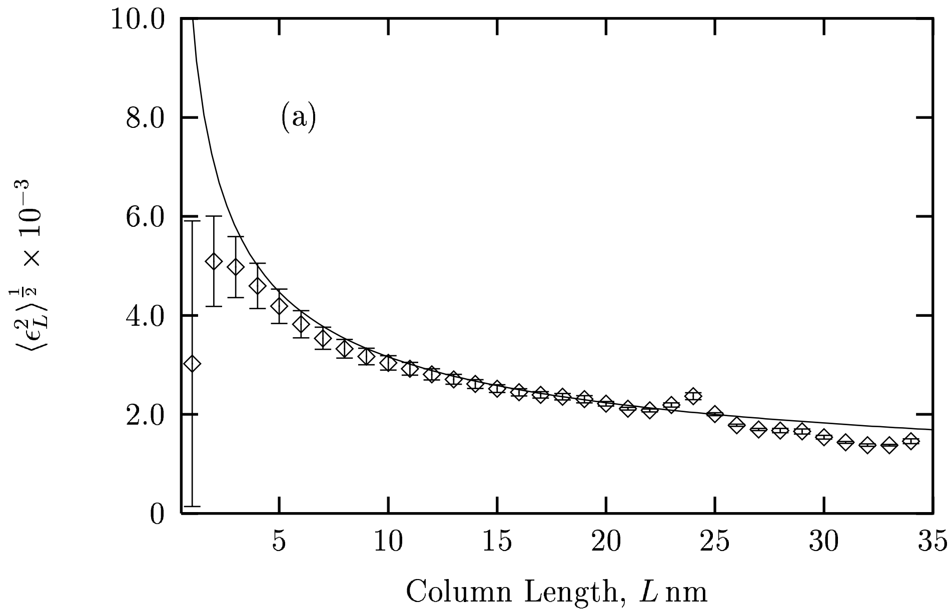


Figure 5.14: Integral breadth, β, versus d^* for Gaussian and non-Gaussian strain: (a) The integral breadth data from the profiles (diamonds + error bars), theoretical curve (solid line) using (5.23), the Williamson-Hall (dashed lines) and quadratic fitting (short dashed lines) using(5.33); (b) The integral breadth data from the profiles (diamonds + error bars), theoretical curve (solid line) using (5.27) in (5.22), the Williamson-Hall (dashed lines) and quadratic fitting (short dashed lines) using (5.34).

The reason for this misfitting is due to the small "hook effect" in the Fourier coefficients, especially in the higher angle peaks. that is, the negative curvature in the Fourier coefficients produced by incorrect background estimation. The consequence of this is that when the Warren & Averbach method is applied, the slope from (5.35) is small, and results in a small $\langle\epsilon_L^2\rangle^{\frac{1}{2}}$ value. Over the range of $L \in [5, 22]$ nm, the $\langle\epsilon_L^2\rangle^{\frac{1}{2}}$ results compare closely with the theoretical distributions given by (5.19). For $L > 22$ nm, the uncertainties in the tails of the Fourier coefficients contribute to the misfitting, like those shown for the 111 and 420 profiles in Figures 5.6 & 5.7. From the results shown in Figure 5.15(a), the estimated G^2 values were determined for two cases, as shown in Table 5.3 (see $\mathrm{WA}^{(1)}$ and $\mathrm{WA}^{(2)}$). For $\mathrm{WA}^{(1)}$ the G^2 values were determined from the fitting (5.4) over the range $L \in [1,\ 34]$ nm. For $\mathrm{WA}^{(2)}$ the G^2 values were determined over the range $L \in [5,\ 22]$ nm. It is evident that in the second case there is an improvement in the estimated G^2 value. As expected, the intercept in (5.35) was zero. This implies that the size coefficients were equal to unity (taking the exponential of $\ln A^s(L)$). The Fourier transform of this is simply a delta-function, and physically signifies an infinitely large crystallite size.

Despite the difficulties in determining the specimen profile, the large noise component and incorrect background estimation, the results for the Warren & Averbach method over $L \in [5,\ 22]$ nm are reasonable.

The rms-strain results for the MaxEnt method are shown in Figure 5.15(b). The solution distributions discussed in §5.4.2 were used to determine $\langle\epsilon_L^2\rangle^{\frac{1}{2}}$ over the range $L \in [2,\ 20]$ nm. The first set of $\langle\epsilon_L^2\rangle^{\frac{1}{2}}$ results was determined using a uniform model in the MaxEnt method — for example see the solution distributions in Figure 5.8. For small L there is a close fit to the theoretical curve given by (5.19), but as L increases misfitting becomes evident. The estimated G^2 values using these uniform $\langle\epsilon_L^2\rangle^{\frac{1}{2}}$ results are given in Table 5.3 for $\mathrm{MaxEnt}^{(1)}$. These results are poor and highlight the role that the *a priori* model has on α and determining $\hat{p}(\epsilon_L)$, and on subsequent calculations. This can be contrasted with the non-uniform *a priori* model which lies reasonably close to the theoretical result — for example, see the solution distributions in Figure 5.9. In other other words, the *a priori* model introduces information into the solution. The estimated G^2 value using these non-uniform $\langle\epsilon_L^2\rangle^{\frac{1}{2}}$ results are given in Table 5.3 for $\mathrm{MaxEnt}^{(2)}$. These result are close to the true the G^2 values of



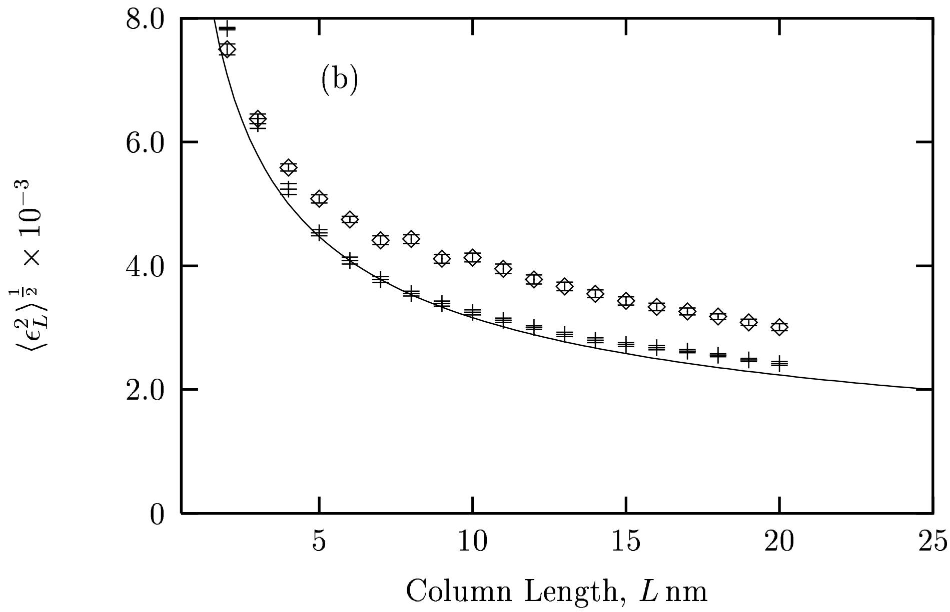


Figure 5.15: Showing the $\langle \epsilon_L^2 \rangle^{\frac{1}{2}}$ results for both the Warren-Averbach and MaxEnt methods: (a) Showing the theoretical $\langle \epsilon_L^2 \rangle^{\frac{1}{2}}$ values from (5.4) and the $\langle \epsilon_L^2 \rangle^{\frac{1}{2}}$ results from the Warren-Averbach method (diamonds); (b) Showing the theoretical $\langle \epsilon_L^2 \rangle^{\frac{1}{2}}$ values from (5.4), the MaxEnt method results for the uniform (diamonds) and non-uniform (crosses) models.

1.0×10^{-4} nm, but could be further improved upon by a better estimate of a non-uniform model.

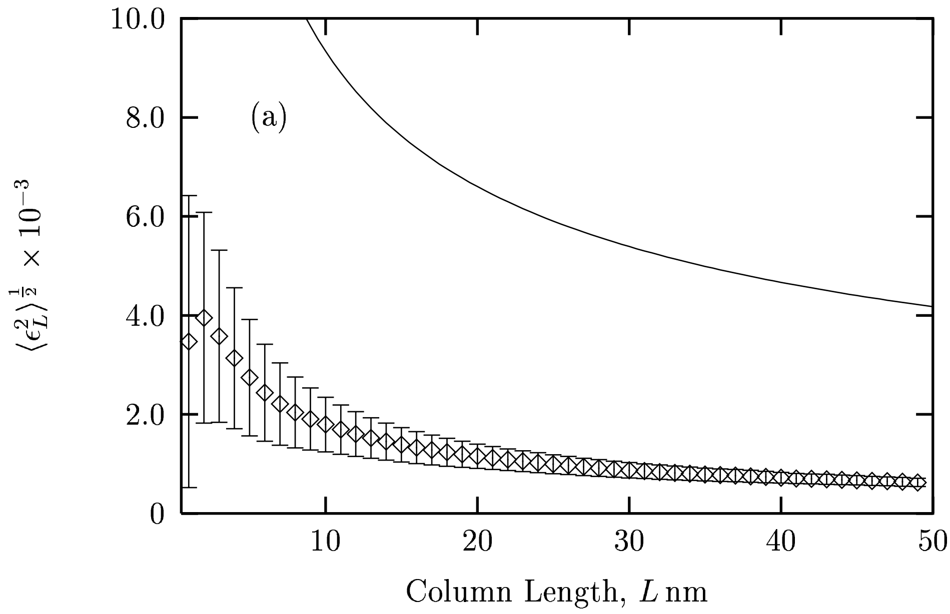
The range over which the MaxEnt method was applied for uniform and nonuniform models was $L \in [2, 20]$ nm. The reason that this cannot go beyond this range is due to the uncertainties in the Fourier coefficients and the useful range of the coefficients — for example, see Figure 5.1.

Determining $\langle \epsilon_L^2 \rangle^{\frac{1}{2}}$ distribution — non-Gaussian case

The second case considered the non-Gaussian strain in the limit of a near-Lorentzian strain distribution, to which the Warren & Averbach, (5.35), and MaxEnt methods were applied. Figure 5.16 shows the results for both methods. The Warren & Averbach method, (5.35), was applied (see Figure 5.16(a)) but has clearly failed to determine the correct $\langle \epsilon_L^2 \rangle^{\frac{1}{2}}$ values.

The failure of the Warren & Averbach method lies in the assumptions of the method. That is, the graph of $\ln A(L, d^*)$ versus d^{*2} is linear for all L and d^{*2}, if and only if $p(\epsilon_L)$ is a Gaussian distribution, resulting in (5.35). As $p(\epsilon_L)$ moves from a Gaussian case, the plot $\ln A(L, d^*)$ versus d^{*2} is linear only for small L. In the limit of a near-Lorentzian $p(\epsilon_L)$, the validity of the assumptions of Warren & Averbach breaks down completely. In this particular case the exponent in (5.24) is $m = 1.55$, which defines the near-Lorentzian limit and lies at the opposite extreme to the Gaussian case where m must be large (i.e. $m \to \infty$). The results in Figure 5.16(a) were used to determined the parameters σ_0, m and q for the strain model (5.26) and are shown in Table 5.4. The large percentage difference between the estimated σ_0 and m parameters are indicators of the failure in the Warren & Averbach method (see WA in Table 5.4).

For this application of the Warren & Averbach method, non-zero size Fourier coefficients were produced. From these coefficients an area-weighted crystallite/domain size, $\langle L \rangle_a^F \sim 13$ nm was determined. It is tempting to interpret this size-dimension as the spatial periodicity of dislocations in the specimen (Warren 1959, van Berkum et al. 1996). However, this dimension should be interpreted conservatively. It is most likely a consequence of the assumptions built into the Warren & Averbach method. That is, in the limit of the Lorentzian distribution the separation of the size and strain coefficients incorrectly attributes



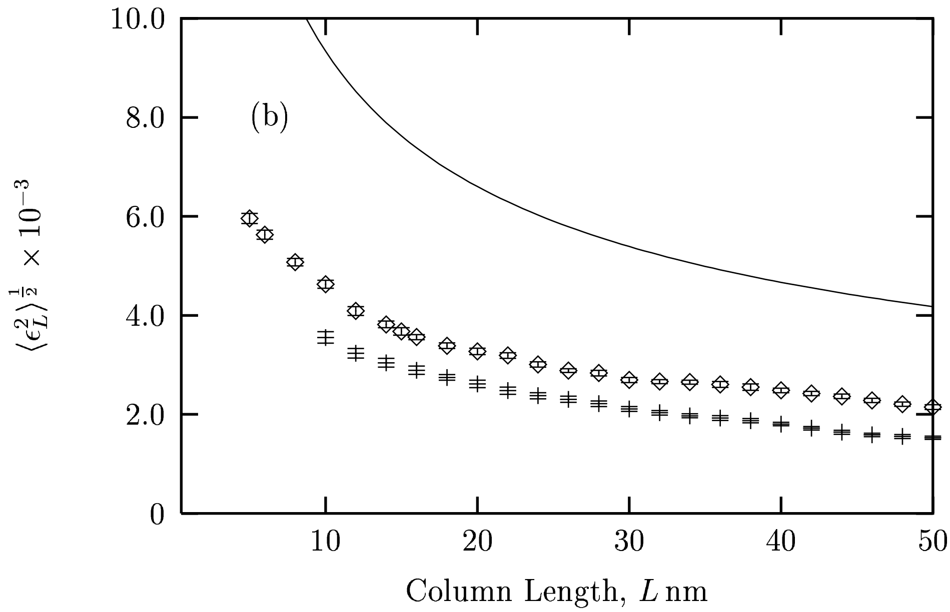


Figure 5.16: Showing the $\langle \epsilon_L^2 \rangle^{\frac{1}{2}}$ results for both the Warren-Averbach and MaxEnt methods: (a) Theoretical $\langle \epsilon_L^2 \rangle^{\frac{1}{2}}$ values from (5.26) and the $\langle \epsilon_L^2 \rangle^{\frac{1}{2}}$ results from the Warren-Averbach method (diamonds); (b) Theoretical $\langle \epsilon_L^2 \rangle^{\frac{1}{2}}$ values from (5.26), the MaxEnt method results for the uniform (diamonds) and non-uniform (crosses) models.

	WA	MaxEnt$^{(1)}$	MaxEnt$^{(2)}$
$\sigma_0 \times 10^{-3}\,(\mathrm{nm}^{\frac{1}{2}})$	5 ± 1	7.5 ± 0.3	7.5 ± 0.4
diff (%)	33	0	0
m	2.9 ± 0.5	1.9045 ± 0.0003	1.99 ± 0.02
diff (%)	87	23	28
q	0.5 ± 0.2	0.48 ± 0.02	0.48 ± 0.02
diff (%)	0	16	4

Table 5.4: The estimated values of σ_0, m, q and percentage difference *diff* for various methods: the Warren & Averbach, (WA); MaxEnt method using a uniform model, MaxEnt$^{(1)}$; MaxEnt method using a uniform model, MaxEnt$^{(2)}$.

size-broadening to the Fourier coefficients and does not correctly determine the strain coefficients. As a result, non-zero size Fourier coefficients are produced. If the $\langle L \rangle_a^F$ were meaningful we would expect a correlation between the integral breadth and Fourier sizes[4]. However the positive intercept determined in the integral breadth results is due to systematic errors. Hence, the $\langle L \rangle_a$ dimension produced in these simulations indicates the breakdown of Warren & Averbach method in the case of a near-Lorentzian strain distribution.

The most surprising outcome is given in Figure 5.16(b). Here the MaxEnt method for both the uniform and non-uniform *a priori* models was used to determine $\langle \epsilon_L^2 \rangle^{\frac{1}{2}}$. If the solution distributions are examined, for example those given by Figures 5.12 & 5.13, we notice the solution strain distributions are reasonably good — especially for the non-uniform model. The Figure 5.16(b) results for the uniform and non-uniform models were fitted to the strain models (5.25) and (5.26). The estimated values for σ_0, m and q parameters are shown in Table 5.4. The percentage differences for the uniform and non-uniform models are less than those for the Warren & Averbach method. The results from Table 5.4 approximate the shape and rate at which the curves decrease in Figure 5.16(b) given by the σ_0 and q parameters. The difference between the calculated and theoretical curves can be explained by the large percentage difference in the m parameter.

[4]Recently van Berkum et al. (1996) show that the $D_{WA} \approx 1.9 D_{WH}$ for experimental data (see experimental results p741 van Berkum et al. 1996).

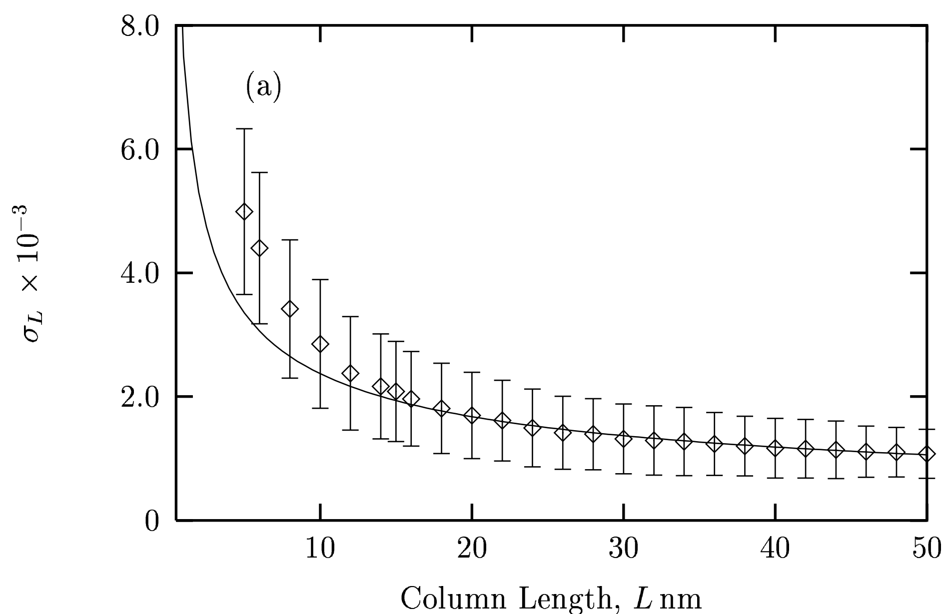
In order to understand this result further, we need to examine (5.26) in terms of the parameters σ_L and m. We notice that (5.26) is dependent on two parameters (unlike the Gaussian case) and if there is an uncertainty in both parameters, these are compounded to produce a large uncertainty in $\langle\epsilon_L^2\rangle^{\frac{1}{2}}$. This can be seen by considering each term: the uncertainty in $\langle\epsilon_L^2\rangle^{\frac{1}{2}}$ with respect to σ_L varies linearly, while the uncertainty in m varies nonlinearly. This makes any error in m more sensitive to small changes (i.e. uncertainties). For example, at $L = 10\,\text{nm}$, $\sigma_L = 2.37 \times 10^{-3}$, taking $m = 1.55$ and using (5.26), $\langle\epsilon_L^2\rangle^{\frac{1}{2}} = 9.34 \times 10^{-3}$. Suppose we introduce an error of 5% into each parameter, then $\langle\epsilon_L^2\rangle^{\frac{1}{2}} = 6.29 \times 10^{-3}$, a percentage difference is of $\sim 33\%$. In the case of 10% error in both parameters, $\langle\epsilon_L^2\rangle^{\frac{1}{2}} = 5.32 \times 10^{-3}$, a percentage difference of $\sim 43\%$. Although the MaxEnt solution strain distributions are visually good, small uncertainties in the parameters produce large uncertainties in $\langle\epsilon_L^2\rangle^{\frac{1}{2}}$.

But what is the underlying cause of these uncertainties in σ_L and m? The best explanation lies in the *a priori* model used. That is, it defines our assumptions and will also include any uncertainties in the parameters that define it. Not only is the "good" information about the strain distribution being imparted into the solution, but so too is "bad" information concerning the an *a priori* model. Perhaps the central issue in defining an *a priori* models is being able to quantify the "good" and "bad" information which it incorporates?

Further examination of the non-Gaussian case

The half-widths at the half-maximum, σ_L, for the solution distributions using uniform and non-uniform models were determined and are shown in Figure 5.17. The influence of the *a priori* model can be seem in comparing Figure 5.17(a) (using a uniform model) with Figure 5.17(b) (using a non-uniform model). The magnitude of the uncertainties in σ_L is reduced and there is misfitting in data points for $L < 12\,\text{nm}$ in Figure 5.17(b). The influence of the model on α, the solution distribution and its uncertainties have been discussed above.

An attempt was made to fit the results in Figure 5.17 to σ_L using (5.25). The estimated values of σ_L and q are shown in Table 5.5. These results are disappointing. The obvious explanation can be seen in the misfitting in the data point for $L < 12\,\text{nm}$ — especially for Figure 5.17(a). The underlying explanation is in the *a priori* models used to determine the



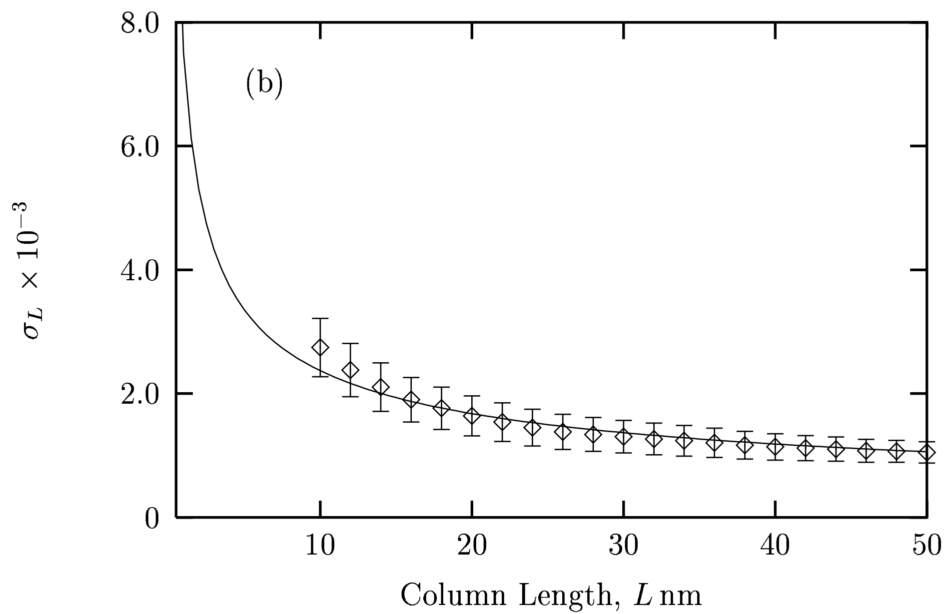


Figure 5.17: The half-widths or σ_L of the MaxEnt strain distributions: (a) Half-width results from the theoretical curve using (5.25) (solid lines) and the MaxEnt strain distributions using a uniform model (diamonds + error-bars); (b) Half-width results from the theoretical curve using (5.25) (solid lines) and the MaxEnt strain distributions using a non-uniform model (diamonds + error-bars)

strain distributions for $L < 12\,\text{nm}$. In Figure 5.17(a) the *a priori* model has no "structure" and no information concerning the scale or shape of the distribution is being transferred into the solution. In Figure 5.17(b) the *a priori* model has "structure", but the information which is being transferred is incorrect, in that the model did not correctly determine the scale of the distribution. This is demonstrated in Figure 5.13.

	MaxEnt$^{(1)}$	MaxEnt$^{(2)}$
$\sigma_0 \times 10^{-3}\,(\text{nm}^{\frac{1}{2}})$	1.6 ± 0.2	1.3 ± 0.2
diff (%)	78.7	82.7
q	0.73 ± 0.05	0.66 ± 0.06
diff (%)	46.0	32.0

Table 5.5: The estimated values for σ_0 and q using (5.25 from Figures 5.12 & 5.13: MaxEnt$^{(1)}$ the σ_0 and q results from the MaxEnt method using a uniform model; MaxEnt$^{(2)}$ the σ_0 and q results from the MaxEnt method using a non-uniform model.

In general the near-Lorentzian or non-Gaussian case has proven to be a challenging problem. The failure of the Warren & Averbach method and its underlying assumptions have been clearly shown. The $\langle \epsilon_L^2 \rangle^{\frac{1}{2}}$ values for the MaxEnt method are disappointing, but the solution distributions are good. Moreover, the half-widths, σ_L, from the solution distribution clearly demonstrate the ability of the MaxEnt method to determine parameters and assign an uncertainty to those quantities.

5.5 Summary

In this chapter the MaxEnt method has been presented as an alternative for determining strain distribution, and has also been used to determine the specimen profile from the simulated observed profiles. This two fold approach revealed the robustness and flexibility of the MaxEnt method. Moreover, developing a MaxEnt method demonstrated how two apparently different problems can be solved under a general and unifying method which preserves the positivity of the solution specimen profiles and strain distribution, takes into consideration the experimental uncertainties and incorporates *a priori* information. It has been

shown that the MaxEnt method for determining the strain distribution makes minimum assumptions concerning the characteristics of the solution, unlike the Warren & Averbach and Williamson & Hall methods, which make specific approximations and were developed from special examples of size and strain broadening.

Using simulations, the cases of Gaussian and near-Lorentzian inhomogeneous strain broadening of diffraction profiles were examined. In each case the simulated observed data consisted of instrumental broadening, background level and statistical noise (see §5.3). For both cases, the MaxEnt method was applied to determining the specimen profile. This enabled a strain analysis to be conducted (see §5.4). The problem of determining the specimen profile proved to be challenging due to the relatively high level of statistical noise, incorrect background estimation and relatively little instrumental broadening (compared to the specimen broadening). However, using a uniform *a priori* model, the positivity of the profiles was preserved and the worst case scenario was defined. Once the specimen profiles were determined the Warren & Averbach and Williamson & Hall methods were applied. The uncertainties determined from the MaxEnt methods were used to quantify the strain broadening in both methods (see §5.4.4). The classical Williamson & Hall method failed to account for the d^{*2} dependence of the integral breadths in both cases and produced physically impossible results, such as a negative intercept. It was also demonstrated that this failure was independent of the systematic errors in the integral breadth. The simulations demonstrated that the underlying assumptions of the Warren & Averbach methods worked as predicted only in the Gaussian case, while in the near-Lorentzian limit it simply failed. The MaxEnt method was used to determine the strain distribution, which in turn was used to determine the root-mean-square strain. The use of a non-uniform *a priori* model proved to be important in determining the strain distribution. This was seen by the visual improvement of obvious characteristics such as the reduced uncertainty region and the shape of the distribution. However, when these results were quantified, by determining the root-mean-square strain and the parameters of the strain model, the results were disappointing. An improved result can only follow with the use of a better *a priori* model.

The central issues that emerge from this study are the ability of the MaxEnt method to remove statistical noise, when there is relatively little instrument broadening, and the use of *a*

priori models in the method in developing a better understanding of the specimen profile and strain distribution. In determining the specimen profile for both cases the MaxEnt method failed to converge to the preferred statistic value, suggesting that the solution profile was over-fitted. This was most likely due the high-level of noise in the simulated data. Furthermore, it indicates that a better approach of noise removal needs to be incorporated into the MaxEnt method. As discussed in Chapter 4, the *a priori* model proved to be important in deducing the effects of background and truncation. The models used in that chapter were assumed to be known and the parameters defining it were arbitrary chosen. In this chapter, however, an attempt was made to determine the model from a "low resolution" method. This approach demonstrated that "good" information concerning the model, such as the scale and shape of the strain distribution, influenced the Lagrangian parameters, the solution distribution and its uncertainties. Also, it demonstrated that the "bad" information, such as the uncertainties in scale and shape of the model, influence the parameters of the strain models. This implies that a method for determining the likelihood of the proposed model and quantifying the "good" and "bad" information which it incorporates is necessary. A basis for such a method could be found in Bayesain statistics by comparing the *posterior* probability of the models conditional on the data, and its parameters quantified (Sivia 1996).

Chapter 6

Application of the Maximum Entropy Method to Alumina Diffraction Data

In this chapter the two-fold MaxEnt method developed in Chapters 3–5 is applied to alumina diffraction data. The diffraction data was taken using the NIST standard reference material (SRM) 676[1] and was first analysed by Kalceff et al. (1994, 1995).

The first level of application uses the MaxEnt method to deconvolve the instrument profile from the observed profile, thus determining the specimen profile. Two cases are considered in determining the specimen profile: the first uses a uniform model (see §6.2), while the second applies a non-uniform model (see §6.3). In both cases the uncertainties in the specimen profile are determined and used in the subsequent analysis. A qualitative analysis is applied to the specimen profiles to determine the nature of the specimen broadening for both cases. It is shown that the specimen broadening is due to size effects. Also, a quantitative analysis is carried out using the traditional integral breadth and Fourier methods to determine the volume-weighted and area-weighted sizes, respectively. At this point the MaxEnt method is applied again (the so-called second level of application) to determine the column-length distribution and the corresponding area-weighted size. The issue of determining a suitable *a priori* starting model is tested for the two cases (see §6.2.4 & 6.3.5). The x-ray analysis is compared with scanning electron microscopy images of the alumina sample (see §6.3.6).

[1]The NIST SRM 676 (α-Al_2O_3) is a material for quantitative analysis or I/I_c determination (Kalceff et al. 1994).

6.1 Background to analysis

6.1.1 Overview of previous analysis

The experimental details of sample preparation and gathering the x-ray diffraction data are discussed in Kalceff et al. (1995). The x-ray diffraction data was recorded using a Siemens D500 diffractometer equipped with a position-sensitive detector (PSD). The NIST SRM 660 (LaB_6) material (Rasberry 1989) was used to determine the instrument profile. It was found that the split Pearson-VII function best fitted the SRM 660 diffraction profile, producing R-factors $1-2\%$ better than other profile functions (Kalceff et al. 1994).

Scanning electron microscopy (SEM) revealed that the SRM 676 has a hexagonal platelet morphology of $1-2\,\mu$m thickness, and $4-5\,\mu$m diameter, with laser scattering data producing a particle size of $1.4\,\mu$m (Kalceff et al. 1995). The Mach-Dollase ratio for the SRM 676 sample was also determined and found to be 0.998 (see Table 1, p 348 Kalceff et al. 1994), this result implying that the particles had no preferred orientation (Larson & von Dreele 1994).

The analysis of Kalceff et al. (1994, 1995) essentially dealt with the issue of determining the instrument profile or instrument profile function (IPF) and removing the instrumental broadening. In particular, they applied the MaxEnt method to remove the instrumental broadening, and the crystallite/domain size analysis used the semi-quantitative Williamson & Hall (1953) method to determine the volume-weighted size[2] as 98 nm. The difference between the volume-weighted size and the SEM size was explained by the presence of small-angle grain boundaries (SAB) within the alumina crystallites. However, this analysis did not correct for the additional broadening due to absorption of x-rays by the alumina specimen.

6.1.2 Determining the instrument profile

The instrument profile was modelled using a split Pearson-VII function. The low- and high-$FWHM$ values and exponents, m, that define the split Pearson-VII were determined from the diffraction data. These parameter data were fitted with functions dependent on the angular position, 2θ, enabling the instrument profile parameters to be determined at the

[2]These are the corrected volume-weighted crystallite/domain size results. The size results determined by Kalceff et al. (1995) contain a factor 2 error (Kalceff 1998).

position of the observed profile.

The low- and high-FWHMs were fitted using (Cheary & Cline 1995)

$$FWHM^2 = A\tan^2\theta + B\cot^2\theta + C\tan\theta + D. \tag{6.1}$$

The low- and high-exponents, m, were fitted to a fourth-order polynomial

$$m = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 \tag{6.2}$$

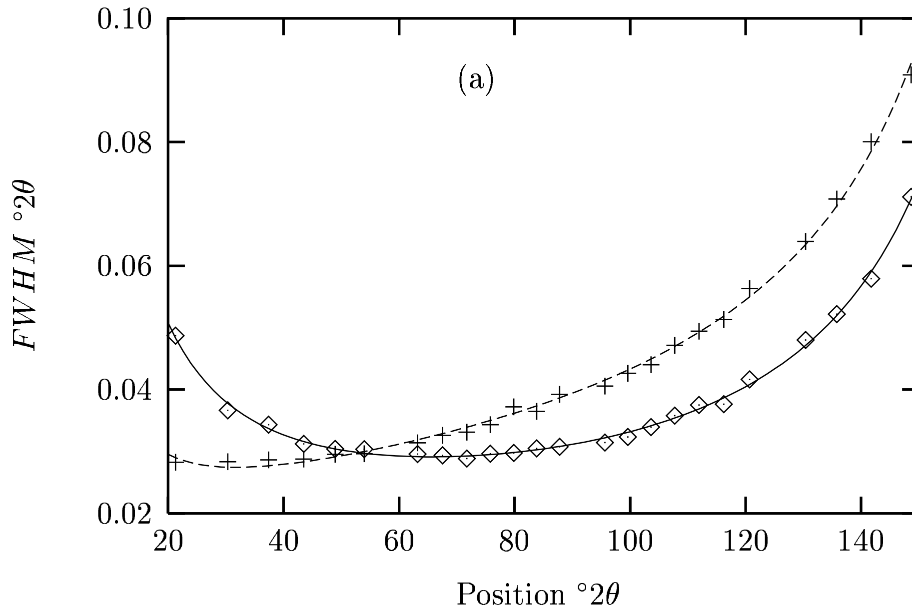
where $x \equiv 2\theta$ and the coefficients $\{a_i;\ i = 0, \ldots, 4\}$ can be determined from the least-squares. Figure 6.1 shows (6.1) and (6.2) and the parameter data from the fitting of the Pearson-VII function to the SRM 660 diffraction data.

The fitting of the low- & high-$FWHM$ data in Figure 6.1(a) follows (6.1) closely. However, there is considerable scattering in the data points for the low- & high-exponents, m. The uncertainties for this parameter were found to be large ($\geq 10\%$).

6.1.3 Background-level estimation

Profiles were extracted from the SRM 676 diffraction pattern, non-overlapped profiles *viz.* 012, 104, 110, 113, 024, 116, 214, 300, 02.10, 226, 21.10, 318 and 410 being used in the analysis. These peaks are generally strong, with a relative intensity greater than 10%. The observed profiles for 012 to 300 were truncated to $2\theta_B \pm 1\,°2\theta$ at a step-size of $0.01\,°2\theta$ (consisting of 200 data points). The remaining observed profiles, 02.10 to 410 were truncated to $2\theta_B \pm 0.8\,°2\theta$ at the same step-size — consisting of 160 data points.

Estimation of the background-level for the profiles followed the procedure discussed in §4.2.1. That is, the MaxEnt method (using a uniform model) was used to determine the specimen profile, $\hat{f}(2\theta)$, and the Fourier coefficients, $\hat{A}(L)$, were examined. In addition, the first and second derivatives of the Fourier coefficients were examined. The distortions that arise in the Fourier coefficients from incorrectly estimating the background-level, such as the hook-effect, propagate into the first and second derivatives (Young et al. 1967). The above steps were repeated, with the estimated background-level being lowered or raised, until distortions in the Fourier coefficients and their derivatives were minimised or had disappeared.



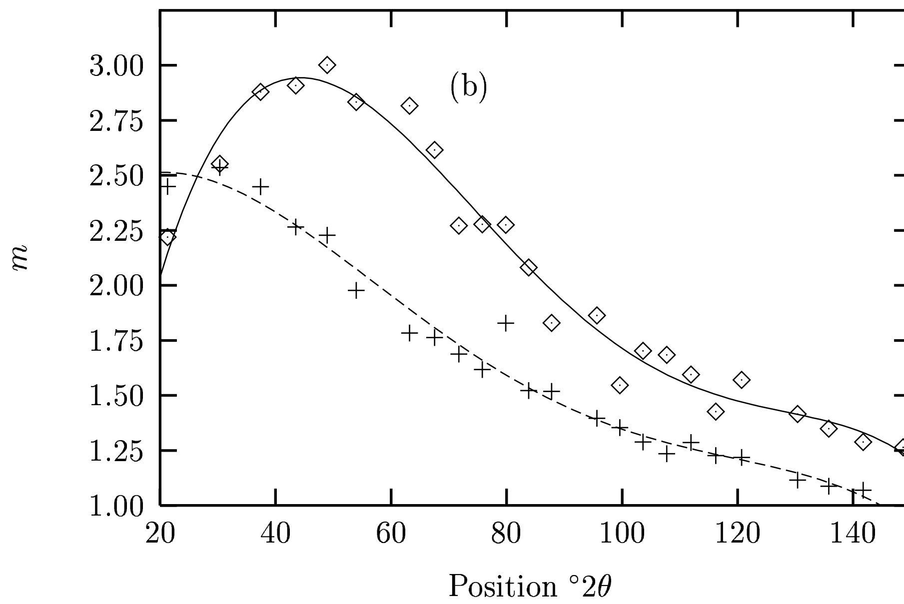


Figure 6.1: Calibration plots for the parameters $FWHM$ and m defining the split Pearson-VII profile generated from an SRM 660 (LaB_6) diffraction pattern: (a) $FWHM$ vs 2θ for the low- (diamonds) & high- (crosses) angle sides of the peak; (b) m vs 2θ, for the low- (diamonds) & high- (crosses) angle sides of the peak.

6.1.4 Determining the specimen profile and column-length distribution

Using the MaxEnt method to determine the specimen profile, two cases were considered: the first applied a uniform *a priori* model; the second used a non-uniform *a priori* model.

The Skilling & Bryan (1984) MaxEnt algorithm (described in Chapter 3 and applied in Chapters 4 & 5) was applied to determine the specimen profile, $\hat{f}$, its uncertainty region, Fourier coefficients, and integral breadth. The C_{aim} value for the MaxEnt method was set equal to the number of data points, M, in the observed profile (see above). The condition $C(f)/C_{aim} = 1$ was satisfied, for the 012 to 300 observed profiles and for the remaining observed profiles $C(f)/C_{aim} \sim 1$ was found to be satisfactory. The Skilling & Bryan (1984) MaxEnt algorithm generally converged within 25 iterations.

The integral breadth and Fourier coefficients from the specimen profiles were determined in order to qualitatively assess the nature of the specimen broadening, thus enabling progression to the next step in the analysis (for example see Louër et al. 1983). This demonstrated that the specimen broadening was a result of crystallite/domain size effects and that the residual-strain was negligible (see §6.2 & 6.3), independently confirming the conclusion of Kalceff et al. (1995).

Once the crystallite/domain size effects had been identified as the major contributors to specimen broadening, the MaxEnt method was applied a second time to determine the column-length distribution. In this second application of the MaxEnt method, the specimen profile and the scattering kernel (4.18) were used in place of the observed and instrument profiles, respectively. As before, C_{aim} was set equal to the number of data points in the specimen profile. For these calculations, a non-uniform *a priori* model was used in the MaxEnt method. Finally, the average area-weighted size, $\langle L \rangle_a^{M,n}$ was determined from the solution column-length distribution.

6.1.5 Determining the variance and uncertainties

The counts in the peaks were observed to be much greater than 10 and the variance was estimated to be $\sigma_i^2 \approx \tilde{g}_i$, where $\tilde{g}_i$ is the background-corrected observed profile and $i = 1, 2, 3, \ldots M$. This approximation had been applied for the simulated data (see §4.4.1 &

5.4.1) and applies equally well to experimental data.

The uncertainty in the MaxEnt specimen profile was determined using (3.58). The region of interest, defined by the vector $\mathbf{u}$ in (3.58), included ones over ± 10 data points about $\hat{f}_i$ and zeros elsewhere. Little change in the uncertainty region was observed for larger regions of interest. The uncertainties for $\hat{f}$ were also used in the MaxEnt method for determining the column-length distribution. Similarly, once $\hat{p}_a(L)$ was determined, its uncertainty region was determined by setting the region of interest to be ones ± 3 points about $\hat{p}_{a\,i}$ and zeros elsewhere (see §4.4.1).

The uncertainties for $\langle L \rangle_a^F$ and $\langle L \rangle_a^M$ were determined using the approach described in §4.4.1, while the uncertainties in the integral-breadth, β, were determined from the uncertainty region in $\hat{f}$.

6.2 MaxEnt profile analysis using a uniform model

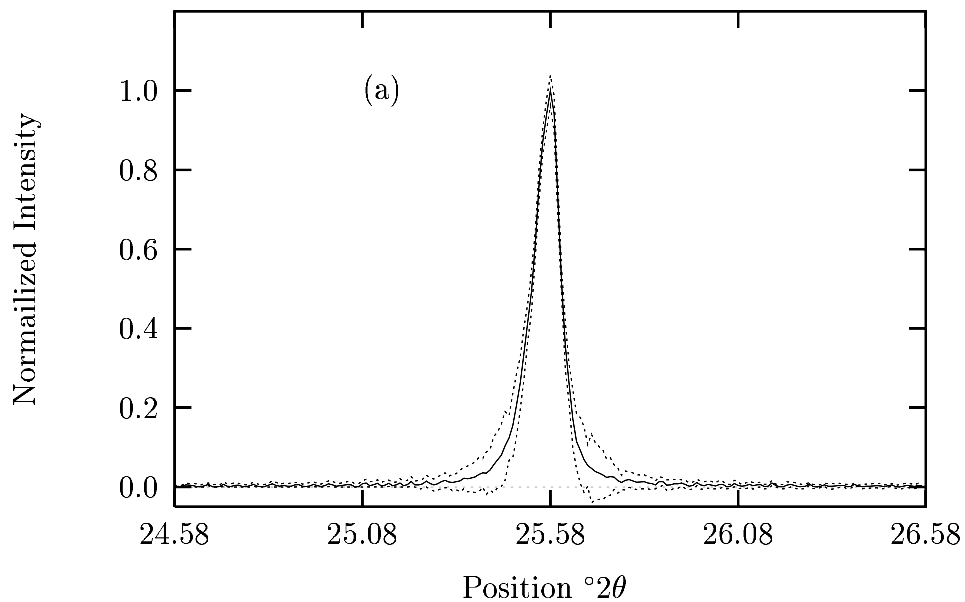
This section presents the SRM 676 diffraction data analysis using a uniform *a priori* model in the MaxEnt deconvolution. A uniform *a priori* model defines our lack of knowledge concerning the shape and broadening of the specimen profile. The MaxEnt method is then applied again to determine the column-length distribution from these specimen profiles and constitutes a “first attempt” (see §6.2.4).

6.2.1 Deconvolving the observed profile using a uniform model

The uniform *a priori* model used in the MaxEnt method was defined by the average number of counts in the background-corrected observed profile. The results of the deconvolution for the multiple orders 012&024 and 113&226 are shown in Figures 6.2(a & b) & 6.3(a & b), respectively. Also shown in these figures are the uncertainty regions of the specimen profiles using (3.58).

In Figures 6.2(a & b) & 6.3(a & b), the MaxEnt solutions suggest that the greatest uncertainty in the specimen profile is about the “shoulders” of the profile. This can also be seen for the 012 and 226 specimen profiles, Figures 6.2(a) & 6.3(a) respectively. The most likely explanation for this comes from the estimation of the exponents, m, for the instrument

profile from Figure 6.1(b). As this parameter determines how rapidly the Pearson-VII function decreases on the low- and high-sides of the Bragg-angle, any uncertainty in estimating m would be imparted to the MaxEnt specimen profile and its uncertainties.



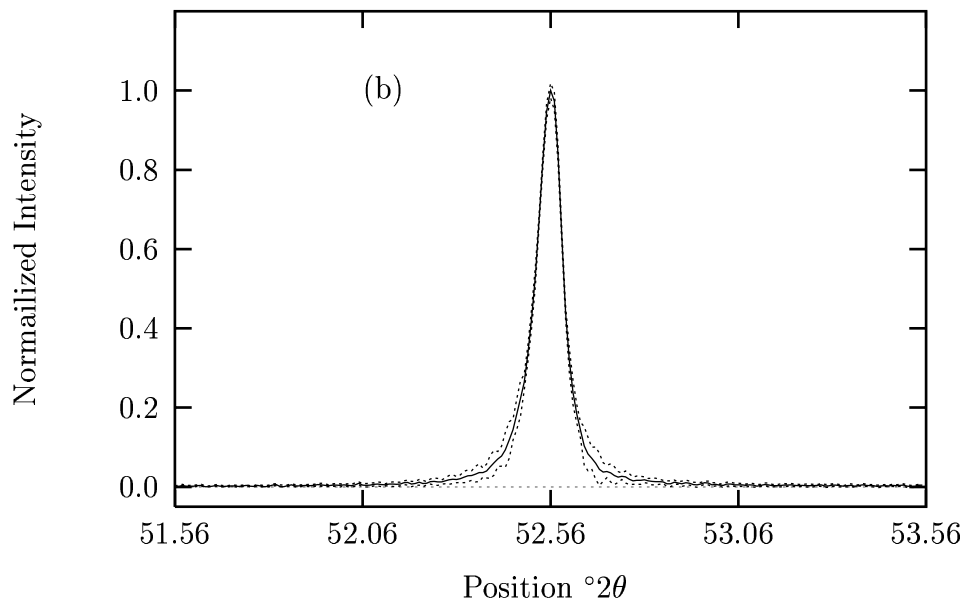


Figure 6.2 continues over...

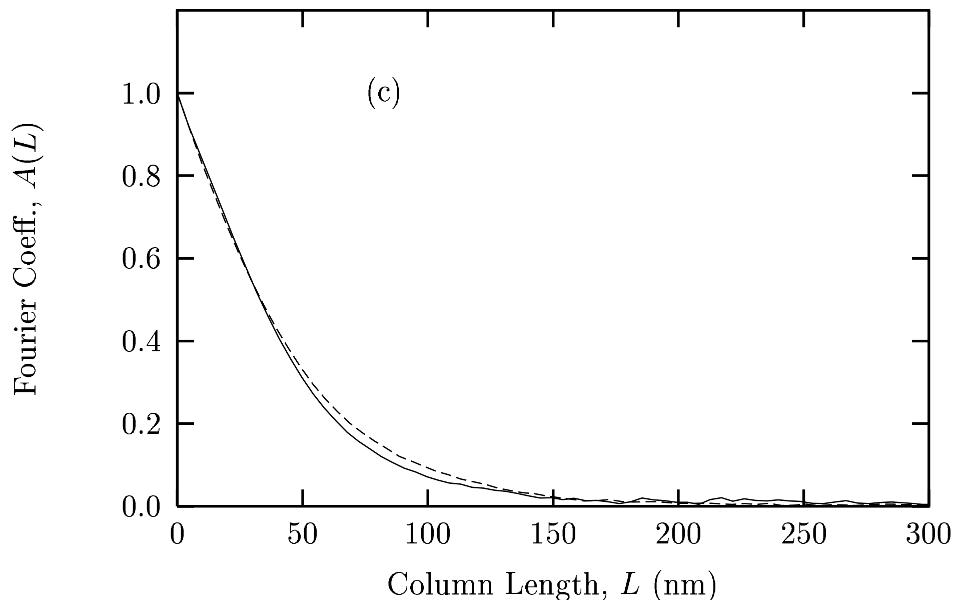
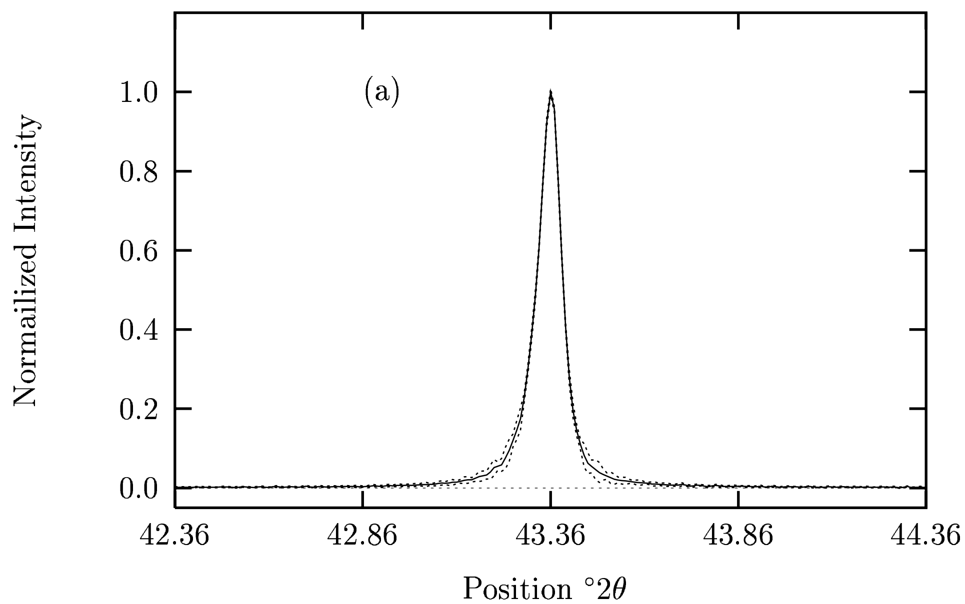


Figure 6.2: MaxEnt results for the 012&024 multiple-orders using a uniform model: (a) The 012 specimen profile (solid line) and lower- & upper-uncertainty regions (dots); (b) The 024 specimen profile (solid line) and lower- & upper-uncertainty regions (dots); (c) The 012 (solid line) and 024 (dashed line) Fourier coefficients. NOTE: The uncertainty regions for the Fourier coefficients are not shown, for clarity.

In Figure 6.3 the MaxEnt specimen profiles for the 113&226 lines are shown. The 113 is the strongest peak, and so the fraction of noise is small. Also, a reliable estimation of the instrument profile results in a narrow uncertainty region in the specimen profile. On the other hand, the 226 profile is a low intensity line (relative intensity 19 %) and the fraction of statistical noise is greater. It is this statistical noise that contributes to the spurious oscillations in Figure 6.3(b), while the uncertainty region indicates the poor reliability of the MaxEnt solution using a uniform *a priori* model.



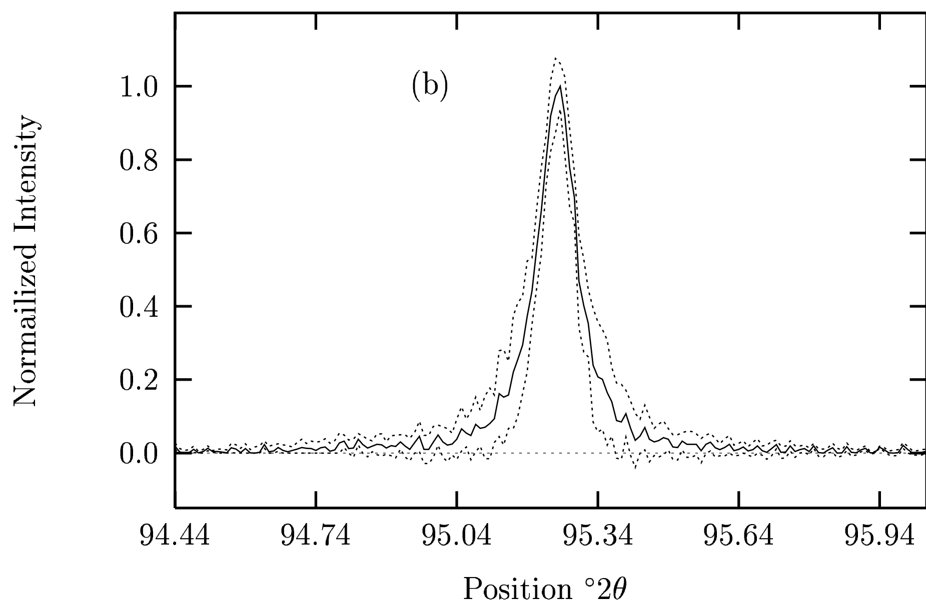


Figure 6.3 continues over...

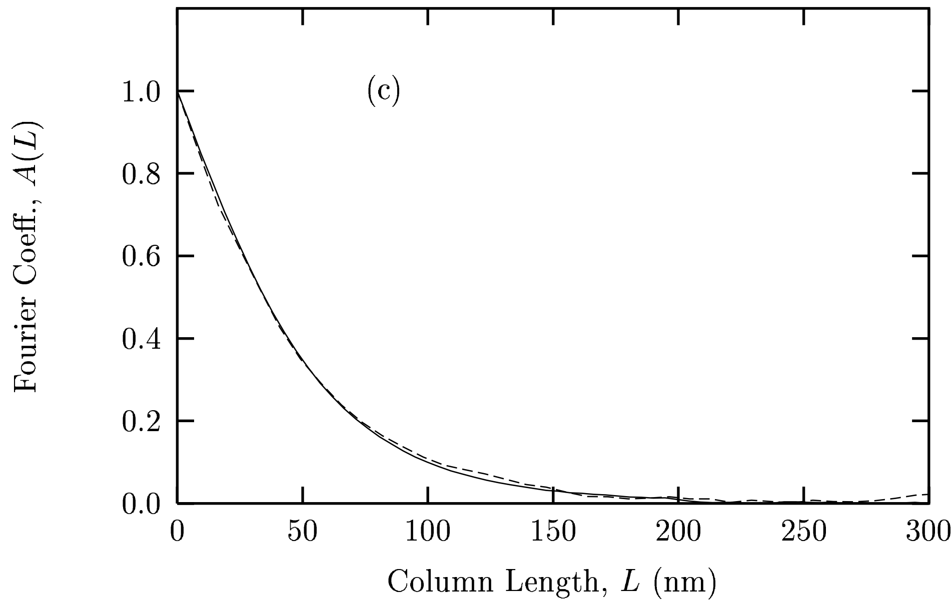


Figure 6.3: MaxEnt results for the 113&226 multiple-orders using a uniform model: (a) The 113 specimen profile (solid line) and lower- & upper-uncertainty regions (dots); (b) The 226 specimen profile (solid line) and lower- & upper-uncertainty regions (dots); (c) The 113 (solid line) and 226 (dashed line) Fourier coefficients. NOTE: The uncertainty regions for the Fourier coefficients are not shown, for clarity.

The MaxEnt results shown in Figures 6.2(a & b) & 6.3(a & b) are characteristic of the specimen profiles from the SRM 676 diffraction data over the low-, mid- and high-2θ range. By applying a uniform *a priori* model in the MaxEnt method, our lack of knowledge concerning the broadening and shape of the specimen profile has been incorporated into the solution method. Given this, the above results could be considered the worst case. In other words, this lack of knowledge about the specimen profile influences the Lagrangian parameter in (3.27) and in turn the uncertainty region, (3.58). However, if any additional information concerning the shape and broadening of the specimen were to be incorporated as a non-uniform *a priori* model, an improvement in the specimen profile would be expected.

6.2.2 Integral breadth, $\beta(d^*)$.

The integral breadth analysis provides a semi-quantitative approach to gathering information of the specimen broadening and crystallite/domain morphology. Figure 6.4 shows the integral-breadth results for each MaxEnt specimen profile (where $\beta(d^*)$ is expressed in reciprocal-space units). The linear fitting provides information on the apparent-strain (given by the slope) and the volume-weighted crystallite/domain size, $\langle L \rangle_v$ (from the inverse of the intercept).

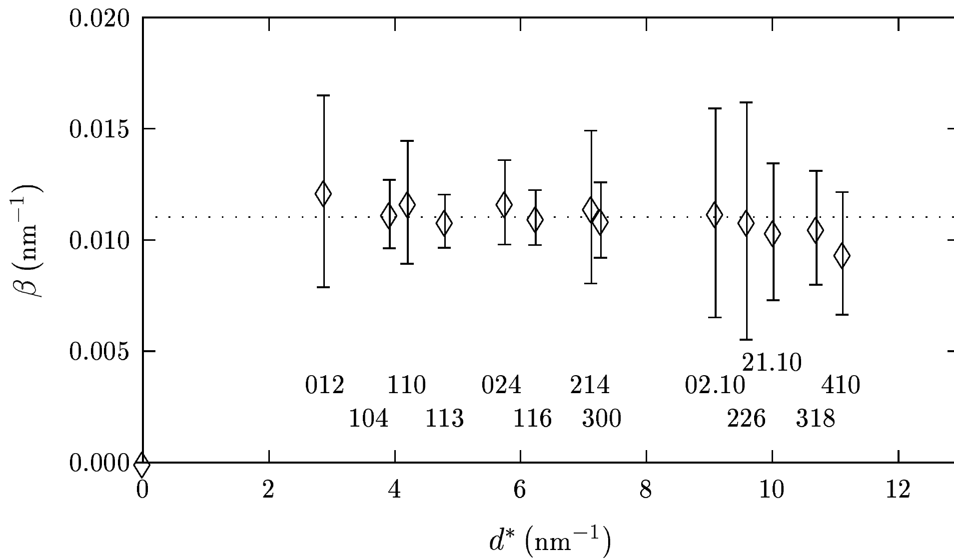


Figure 6.4: Integral breadth $\beta(d^*)$ versus d^* for the MaxEnt specimen profiles using a uniform model (diamonds + error bars) and the linear fit (dots).

Referring to Figure 6.4, several points can be noted: the error-bars reflect the uncertainty in the MaxEnt specimen profiles; these uncertainties are most likely an indication of the errors in estimating the exponent m in Figure 6.1(b) and the statistical noise from the observed profile. The slope of the linear fitting was constrained to zero. The volume-weighted size $\langle L \rangle_v$ was found to be (91 ± 14) nm. The uncertainty in $\langle L \rangle_v$ is large at $\sim 15\%$. This $\langle L \rangle_v$ value is consistent with the (corrected) value determine by Kalceff et al. (1995) (see footnote 2, p 189). The data in Figure 6.4 indicate a negative slope, evident in the decrease of 21.10, 318 and 410 β-values in Figure 6.4. The remaining β-values are relatively constant. The decrease of 02.10, 318 and 410 suggests a systematic error arising from the specimen profiles.

That is, a "weak" dependence on d^* (the β-value for the 318 is slightly greater than the 21.10).

The β-values for 012 to 226 are relatively constant and lie close to the linear-estimation, which implies that the specimen broadening is independent of d^* and is isotropic in $[hkl]$. This can be explained if the specimen broadening is due solely to crystallite/domain effects (independent of d^*), and the crystallites/domains have a near-spherical or shape morphology.

6.2.3 Fourier coefficients, $A(L)$

A qualitative analysis

A qualitative assessment of the Fourier coefficients was undertaken to independently determine the nature of the specimen broadening. The Fourier coefficients for the multiple orders 012&024 and 113&226 are shown in Figure 6.2(c) & 6.3(c), respectively. The multiple-orders are plotted on the same axes, providing a good indication of the specimen broadening. For both sets of multiple-orders, the Fourier coefficients over-lie each other. This suggests that the specimen broadening is independent of the order of diffraction, d^*. Some differences between the 012 and 024 Fourier coefficients are noticeable over the approximate range $50 - 125\,\text{nm}$; similarly there is a slight difference in the 113 & 226 Fourier coefficients over the approximate range $75 - 140\,\text{nm}$.

It is tempting to explain this difference in terms of inhomogeneous strain broadening in the SRM 676 sample. However, if there were strain-broadening, the 024 Fourier coefficients would be narrower relative to the 012 Fourier coefficients and decrease monotonically as L increases. Figures 5.1 & 5.2 provides a good example of the Fourier coefficients from strain-broadened profiles. The results in Figure 6.2(c) & 6.3(c) do not share the same characteristics.

From Figures 6.2(c) & 6.3(c), the maximum crystallite/domain size can be estimated. This dimension refers to the upper-limit in the evaluation of the integral-breadth (Stokes & Wilson 1942). In other words, it defines the L $(= L_{\max})$ when $\hat{A}(L)$ or $V(L)/V(0) \to 0$. From these figures $L_{\max}$ is approximately in the range of $250 - 300\,\text{nm}$.

The Fourier coefficients of all the specimen profiles were plotted on the same axes in order to determine whether the specimen broadening was independent of d^* for all hkl.

The Fourier coefficients from the 012 to 226 specimen profiles were all close to each other and systematic narrowing of the Fourier coefficients for increasing d^* was not observed, eliminating the possibility of isotropic strain broadening. However, this does not exclude the possibly of negligible anisotropic strain broadening, which would cause a non-systematic narrowing of the Fourier coefficients for increasing d^*. Because the coefficients lie close to each other, size broadening can be considered the dominant effect. For the 21.10, 318 and 410 profiles, the Fourier coefficients were observed to be broader than for other peaks and did not indicate the characteristics of inhomogeneous strain as shown in Figures 5.1 & 5.2. As was the case with the β-values, the increase in broadening for the 21.10, 318 and 410 Fourier coefficients followed a "weak" d^* dependence (the 318 were slightly broader than the 21.10 Fourier coefficients).

The Fourier coefficients and integral breadth are related by

$$\beta(d^*) = \left[\int_{-\infty}^{\infty} A(L,\, d^*)\ \mathrm{d}L \right]^{-1}. \tag{6.3}$$

From (6.3), the broadening in the 21.10, 318 and 410 Fourier coefficients would produce a decrease in $\beta(d^*)$. As pointed out above, the broadening of the 21.10, 318 and 410 Fourier coefficients followed a weak d^* dependency and the decrease in $\beta(d^*)$ also follows a weak dependency. This can be seen in Figure 6.4.

A likely explanation for the variation in the Fourier coefficients and β for the 21.10, 318 and 410 profiles is the use of the SRM 660 sample to determine the instrument profile. It is known that SRM 660 contains residual strain broadening at high-2θ (Cheary & Coelho 1998*b*). The low- & high-$FWHM$ and m in Figure 6.1 characterise not only the instrument broadening arising from the diffractometer, but also the residual-broadening inherent in the SRM 660 sample. An analogous problem was investigated in the simulations of §4.4.4, where residual size broadening was introduced into the LaB_6 specimen. From the simulations it was demonstrated that the presence of residual broadening in the instrument profile causes the MaxEnt method to over-compensate for the instrumental broadening and to produce narrower specimen profiles relative to the true specimen profile. This in turn affects the Fourier coefficients, $\langle L \rangle_a$, β and $\langle L \rangle_v$ results. The negative slope in Figure 6.4 reflects the residual strain broadening in the SRM 660 sample, which is manifested at high angles. The reason for this is that strain broadening follows an approximate $\tan\theta$ dependence. The

consequence of the residual strain broadening in the instrument profiles is that any strain broadening that may be present in SRM 676 data, particularly at high 2θ, may be obscured.

From these results it can be stated that the specimen broadening is independent of d^*, especially for the 012 to 226 specimen profiles. Moreover, the results also suggest the specimen broadening is independent of the direction of diffraction, $[hkl]$. That is, the size-broadening (independent in d^*) is isotropic in $[hkl]$, implying that on average the crystallite/domain morphology may be of near-spherical or spherical shape. The maximum dimension (or diameter), $L_{\max}$, determined from examining all the Fourier coefficients was found to be $250 - 300$ nm.

Fourier analysis

The area-weighted size was determined from the Fourier coefficients. This dimension represents the apparent thickness of the crystallite/domain in the direction of the diffraction vector and is defined as the ratio of the volume of the crystallite/domain to the area projected onto the diffraction-planes. The area-weighted size was determined from the initial slope of the coefficients as

$$\frac{1}{\langle L \rangle_a^F} = -\left(\frac{\mathrm{d}A(L)}{\mathrm{d}L}\right)_{L \to 0}. \tag{6.4}$$

The $\langle L \rangle_a^F$ results are shown in Figure 6.5(a). In this figure the $\langle L \rangle_a^F$ have been plotted over the 2θ-range of the profiles. There is considerable scatter in the results and this is much greater than the error in the size results. The size results for the 012 and 024 are in agreement with each other. This is also the case for the 113 and 226 profiles.

There are two possible explanations for the scatter in $\langle L \rangle_a^F$: the first is based on the numerical errors; $\langle L \rangle_a^F$ is sensitive to numerical errors in the Fourier coefficients and the inverse of (6.4) results in the amplification of these numerical errors. The second explanation has a physical basis, in that the scattering may be a consequence of anisotropic size-broadening in $[hkl]$. However, this is doubtful in that the Fourier coefficients and integral breadth from the 012 to 226 profiles suggest the specimen broadening is independent of d^* and of $[hkl]$.

The average size $\langle L \rangle_a^F$ from Figure 6.5 is (56 ± 2) nm, where the uncertainty represents the standard error in the mean. Taking the ratio of $\langle L \rangle_a^F$ and $\langle L \rangle_v$ produces 0.61. If crystallites/domains were all the same size spheres the ratio of the area- to volume- weighted

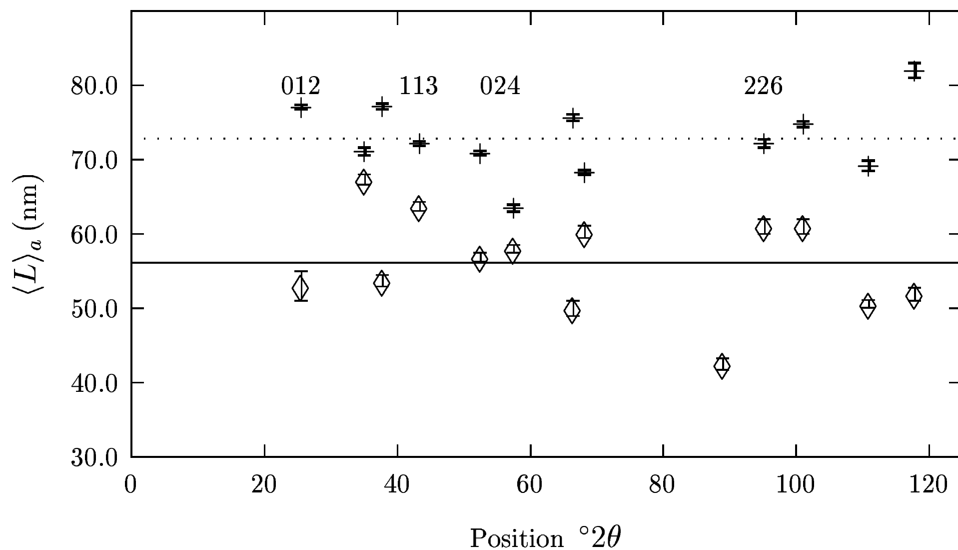


Figure 6.5: The area-weighted size for the Fourier method (diamonds + error-bars), MaxEnt methods (crosses + error-bars), average Fourier size (solid line) and average MaxEnt size (dots).

size would be exactly 8/9 (see Appendix B). However, this result suggests that there is significant range in the sizes of the crystallites/domains.

6.2.4 Determining the column-length distribution – first attempt

The specimen profiles determined using a uniform model were used in determining the column-length distribution. A non-uniform model for the column-length distribution was developed from the specimen profile using a "low resolution" method. This constitutes a first attempt to determine the column-length distribution; an improved model is developed in §6.3 and a second determination of the column-length distribution arising from it is given in §6.3.5.

Using a "low resolution" method to develop a non-uniform model

The low-resolution method that was used to estimate the *a priori* model was the constrained deconvolution/inversion method discussed in §2.1.4, with a smoothing matrix used as described by Twomey (1963). This method does not preserve the positivity of the distribution and large spurious oscillations were observed in the results. In order to produce a positive-

definite model, the estimate was then fitted with a distribution dependent on two parameters μ and $\langle L\rangle_a^m$. The distribution used is given by (see van Berkum et al. 1996),

$$p_m(L|\langle L\rangle_a^m, \mu_m) = Z_m^{-1} \left(\frac{\mu_m L}{\langle L\rangle_a^m}\right)^{\mu_m - 1} \exp\left[-\frac{\mu_m L}{\langle L\rangle_a^m}\right] \tag{6.5}$$

where the "|" in (6.5) reads as "conditional on" and the Z_m is the normalisation term,

$$Z_m = \Gamma(\mu_m)\langle L\rangle_a^m/\mu_m. \tag{6.6}$$

The parameters μ_m and $\langle L\rangle_a^m$ influence the tails of the distribution and estimation of the area-weighted size, respectively. The values of the μ_m and $\langle L\rangle_a^m$ are shown in Table 6.1.

hkl	012	104	110	113	024	116	214	300	02.10	226	21.10	318	410
$\langle L\rangle_a^m$ (nm)	58	63	74	64	58	59	60	57	-	56	59	52	82
μ_m	1.8	3.8	1.4	2.1	1.8	1.7	1.6	2.2	-	1.4	2.1	1.4	1.3

Table 6.1: The *a priori* model parameters, $\langle L\rangle_a^m$ and μ_m for hkl, determined by fitting (6.5) to the results of the "low resolution" method.

The parameter values for the 012 profile were taken to be equal to the 024 parameters, but for the 02.10 case the fitting algorithm did not converge.

The MaxEnt size analysis

Using the non-uniform distribution given in (6.5) and the corresponding parameters in Table 6.1, the MaxEnt column-length distributions were determined. The constraint, $C(p_a)/C_{aim} \sim 1$ was applied (the range of $C(p_a)/C_{aim}$ values was 1.0 to 2.0). The MaxEnt sizes, $\langle L\rangle_a^{M,n}$ are shown in Figure 6.5. The scatter in the $\langle L\rangle_a^{M,n}$ values is less than that of the $\langle L\rangle_a^F$ results. The average $\langle L\rangle_a^{M,n}$ was found to be (73 ± 1) nm. The average offset between the size results is 17 nm. The most likely explanation for the offset in the $\langle L\rangle_a^{M,n}$ results is the large uncertainty in the specimen profile determined in §6.2.1. That is, the MaxEnt method is converging onto the appropriate $C(p_a)/C_{aim}$ values, but producing a larger than expected

$\langle L \rangle_a$. The large uncertainties in the specimen profile are imparted to the column-length distribution and a greater dispersion of the column-length distribution occurs.

The $\langle L \rangle_a^{M,n}$ results are doubtful and little credence can be attributed to them. The results indicated the need to develop a non-uniform *a priori* model for determining the specimen profile, rather than relying on the uniform *a priori* model. As mentioned above, a uniform model represents our ignorance of the shape and broadening of the specimen profile. The large uncertainty regions in the specimen profile describe our lack of knowledge of the specimen profile and this lack of knowledge is also being imparted into the column-length distribution and estimation of $\langle L \rangle_a^{M,n}$.

6.3 MaxEnt profile analysis using a non-uniform model

In the last section the MaxEnt specimen profile, Fourier coefficients and integral breadth were determined using a uniform *a priori* model. Large uncertainties in the specimen profile and integral breadths and a large offset between the $\langle L \rangle_a^{M,n}$ values were observed. In order to improve the specimen profile, integral breadth and $\langle L \rangle_a^{M,n}$ results, a non-uniform model for the specimen profile was developed. Using an analytical expression for the column-length distribution and for the related Fourier coefficients. Using this model a second attempt to determine the column-length distribution was made.

6.3.1 Developing a non-uniform model for the specimen profile

The non-uniform *a priori* model for the specimen profile was developed using (6.5) in

$$A(L) = \frac{1}{\langle L \rangle_a^m} \int_{|L|}^{\infty} (L' - |L|)\, p_m(L' | \langle L \rangle_a^m,\, \mu_m)\, \mathrm{d}L' \tag{6.7}$$

to produce an analytical expression for the Fourier coefficients. The integration[3] was carried for $L \geq 0$ and resulted in

$$A(L) = \frac{\Gamma\left(1 + \mu, \frac{L\,\mu_m}{\langle L \rangle_a^m}\right)}{\Gamma\left(1 + \mu_m\right)} - \frac{L\,\Gamma\left(\mu, \frac{L\,\mu_m}{\langle L \rangle_a^m}\right)}{\langle L \rangle_a^m\, \Gamma\left(1 + \mu_m\right)} \tag{6.8}$$

[3]The integration was done using Mathematica 3.0. The result was reduced to (6.8) using the `PowerExpand` and `Power` commands, respectively.

where $\Gamma(x)$ and $\Gamma(x,y)$ are the Gamma and incomplete-Gamma functions, respectively (see Arfken & Weber 1996). The $\langle L\rangle_a^m$ and μ_m parameters can be determined by fitting the Fourier coefficients from §6.2.3 to (6.8). The values of these parameters are given in Table 6.2.

hkl	012	104	110	113	024	116	214	300	02.10	226	21.10	318	410
$\langle L\rangle_a^m$ (nm)	60	65	61	60	57	59	55	57	49	55	56	-	61
μ	2.5	2.4	2.3	1.9	1.8	1.8	1.7	1.7	1.1	1.5	1.4	-	1.4

Table 6.2: The model parameters, $\langle L\rangle_a^m$ and μ, for hkl, determined by fitting (6.8) to the Fourier coefficients.

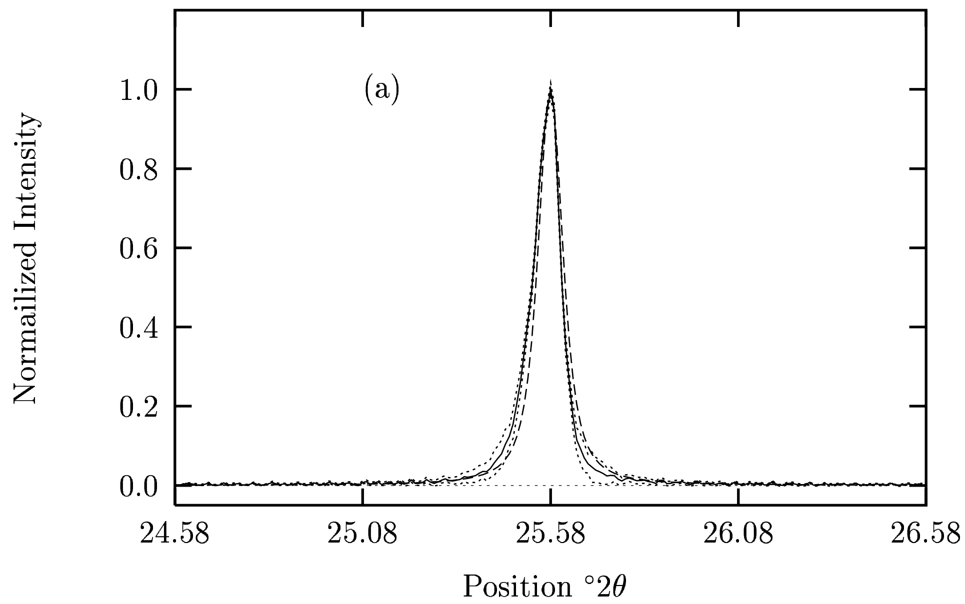
It is interesting to notice the relative uniformity of the $\langle L\rangle_a^m$ with hkl. In the case of the 318 line, the fitting algorithm did not converge. There is some difference between the $\langle L\rangle_a^m$ and μ values for the multiple-orders, suggesting that the parameters are sensitive to small changes in the Fourier coefficients. The uncertainties in the parameters were estimated to be $\geq$ 10%. Using results from Table 6.2 and (6.5), the specimen profile model was evaluated using the procedure outlined in §4.3.1. The models were determined over the range $(2\theta_B \pm 5)\,°2\theta$ at a step-size of $0.01°2\theta$ and then truncated to the same length as the observed profile.

6.3.2 Determining the specimen profile using a non-uniform model

Using the non-uniform model the specimen profile was determined. The condition $C(f)/C_{aim} = 1$ was satisfied for profiles 012 to 300. For the remaining profiles the $C(f)/C_{aim}$ fell in the range of $1.0 - 1.6$.

The MaxEnt specimen profiles for the 012&024 and 113&226 multiple-orders are shown in Figures 6.6(a & b) & 6.7(a & b), respectively. Comparing the MaxEnt specimen profiles from Figures 6.2(a & b) & 6.6(a & b), a reduction in the uncertainty region using the non-uniform model can be seen in the latter figure. The uncertainty region in Figure 6.2(a) was attributed to the uncertainty in determining the exponents m for the instrument profile; despite this,

the non-uniform model for the 012 profile achieved a reduction in the uncertainty region about the shoulders of the profile.



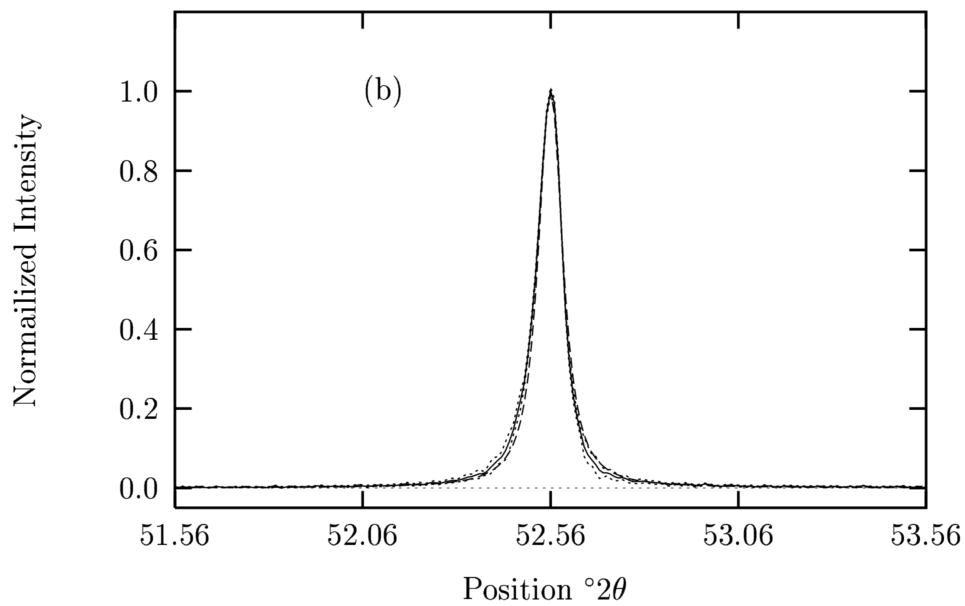


Figure 6.6 continues over...

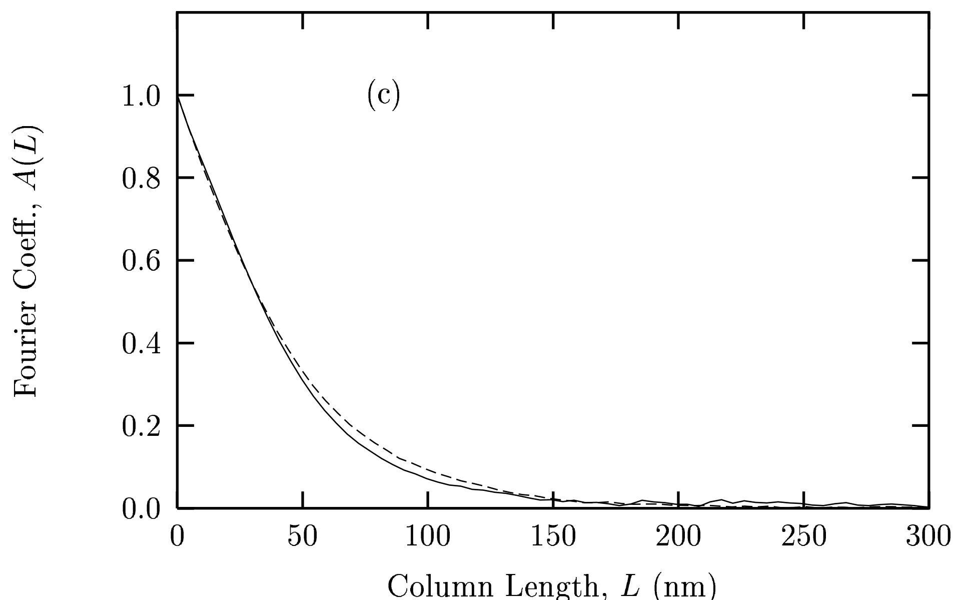
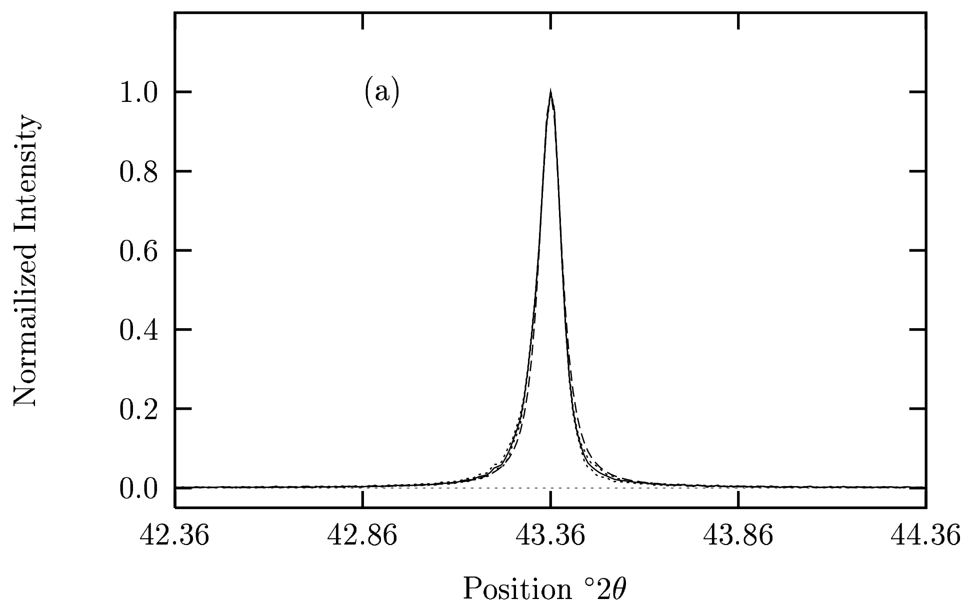


Figure 6.6: MaxEnt results for the 012&024 multiple-orders using a non-uniform model: (a) The 012 specimen profile (solid line), the specimen profile model (dashed line) and lower- & upper-uncertainty regions (dots); (b) The 024 specimen profile (solid line), the specimen profile model (dashed line) and lower- & upper-uncertainty regions (dots); (c) The 012 (solid line) and 024 (dashed line) Fourier coefficients. NOTE: The uncertainty regions for the Fourier coefficients are not shown, for clarity.

Comparing Figures 6.3(a & b) & 6.7(a & b), the influence of the non-uniform *a priori* model in reducing the uncertainty region in the latter results can also be seen. The spurious oscillations shown in Figures 6.3(b) have been eliminated in Figure 6.7(b) by applying the non-uniform model.

In Figures 6.6 & 6.7 asymmetry in the specimen profiles relative to the (symmetrical) specimen model is noticeable. This is due to absorption in the alumina specimen. The asymmetry was not noticed in the observed profiles because the instrumental aberrations over-shadowed the absorption effects.



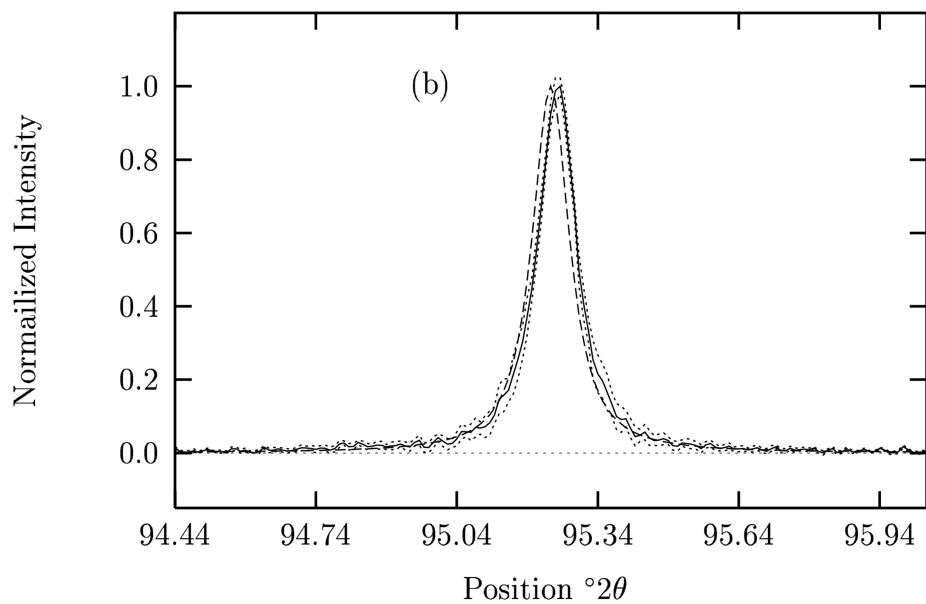


Figure 6.7 continues over...

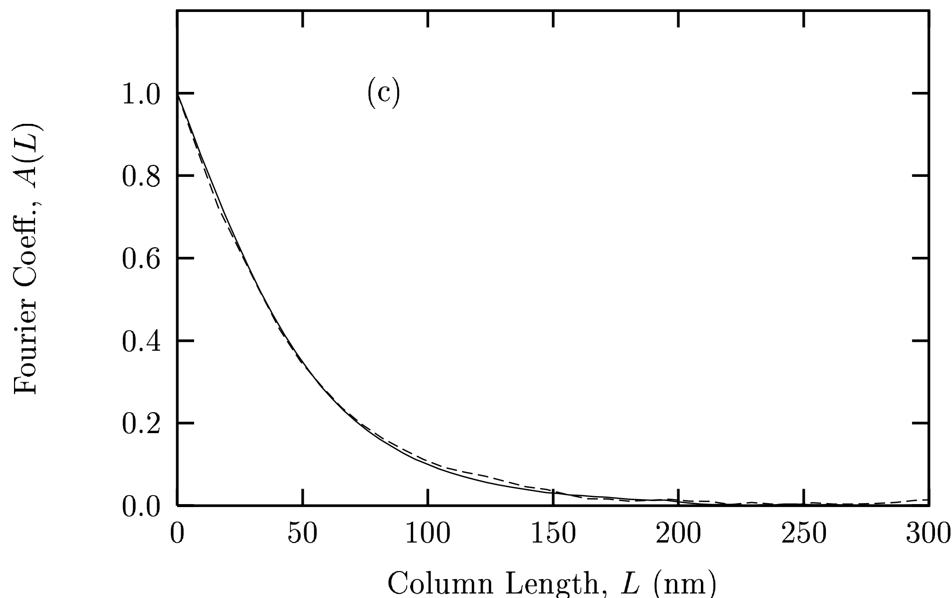


Figure 6.7: MaxEnt results for the 113&226 multiple-order using a non-uniform model: (a) The 113 specimen profile (solid line), the specimen profile model (dashed line) and lower- & upper-uncertainty regions (dots); (b) The 113 specimen profile (solid line), the specimen profile model (dashed line) and lower- & upper-uncertainty regions (dots); (c) The 113 (solid line) and 113 (dashed line) Fourier coefficients. NOTE: The uncertainty regions for the Fourier coefficients are not shown, for clarity.

The MaxEnt specimen profiles given in Figures 6.6(a & b) & 6.7(a & b), demonstrate the influence of the non-uniform model in the MaxEnt method. In this case, information concerning the broadening and shape of the specimen profile has been incorporated by developing a non-uniform *a priori* model using the Fourier coefficients of the specimen profile. This information aids the method in determining the specimen profile with the maximum entropy and reducing the uncertainty region.

6.3.3 Integral breadth, $\beta(d^*)$

Uncertainties in $\beta(d^*)$ are also reduced as a consequence of using a non-uniform model, as can be seen from Figure 6.8.

As was the case in Figure 6.4, the slope is again constrained to zero. The volume-weighted size, $\langle L \rangle_v$, was determined from the inverse of the intercept and found to be (90 ± 8) nm. In this case the uncertainty has been reduce to about 9%. Without constraint, a negative slope

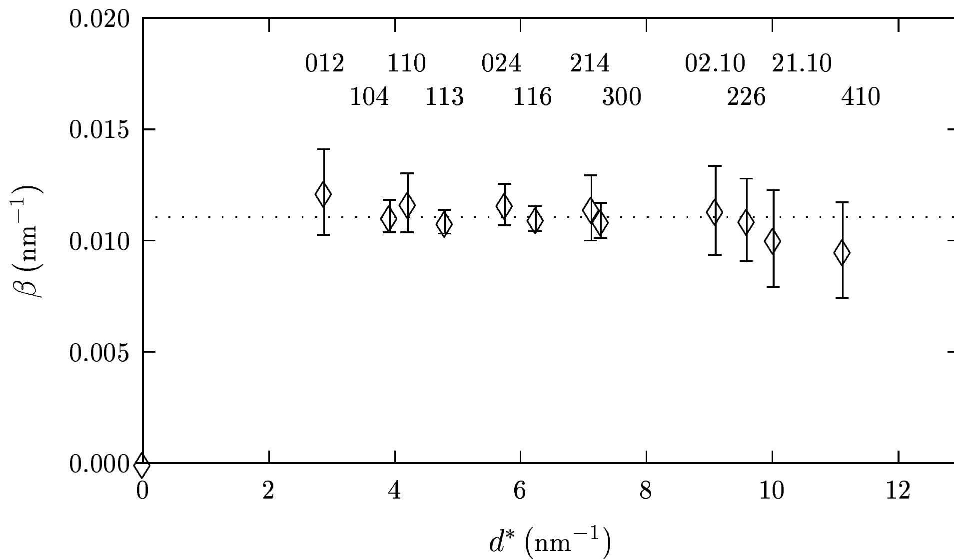


Figure 6.8: Integral breadth $\beta(d^*)$ versus d^* for the MaxEnt specimen profiles using a non-uniform model (diamonds + error bars) and the linear fit (dots).

is indicated in Figure 6.8. Again the reason for this is that the residual broadening in the instrument profile had not been corrected in the specimen profile model. For the profiles from 012 to 226 the integral breadth is constant over d^*, suggesting that the specimen broadening is independent of d^* and $[hkl]$.

6.3.4 Fourier coefficients

A qualitative analysis

The Fourier coefficients for the 012 & 024 and 113 & 226 multiple-order are shown in Figure 6.6(c) & 6.7(c), respectively. Comparing these results with Figures 6.2(c) & 6.3(c), little change in the coefficients is noticeable.

The reason for this is that the $\langle L\rangle_a^m$ and μ parameters were determined by fitting (6.8) with the Fourier coefficients determined in §6.2. The same systematic variation that was observed above was also observed for these Fourier coefficients, further suggesting that the non-uniform model was unsuccessful in accounting for the residual broadening in the instrument profile.

Fourier size analysis

The $\langle L\rangle_a^F$ results using (6.4) were determined from the Fourier coefficients. These results are shown (with others to be discussed in the next section) in Figure 6.9, plotted over the 2θ-range of the diffraction pattern. A close agreement for the 012 and 024 size results can be seen, which is also the case for the 113&226 multiple-orders. The scattering that was observed in Figure 6.5 is also evident in Figure 6.9, probably indicating that the small numerical errors in $A(L)$ about $L \approx 0$ are being amplified in (6.4). The average $\langle L\rangle_a^F$ from the results in Figure 6.9 was determined to be (57 ± 2) nm.

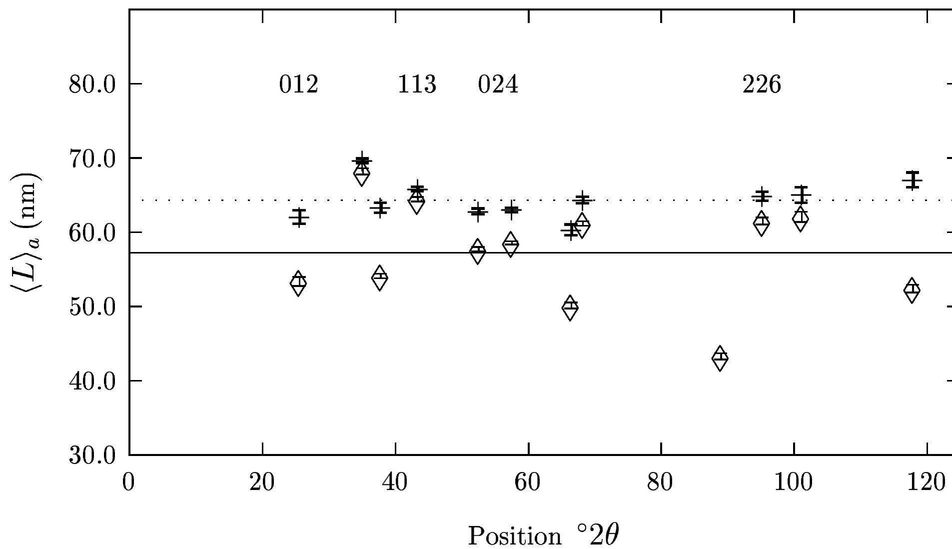


Figure 6.9: The area-weighted size for the Fourier method (diamonds + error-bars), MaxEnt method (crosses + error-bars), average Fourier size (solid line) and average MaxEnt size (dots).

6.3.5 Determining the column-length distribution – second attempt

The specimen profiles determined in §6.3.2 were used in the MaxEnt method for the column-length distribution. The non-uniform model and parameters were taken from (6.5) and Table 6.2, respectively. The column-length distribution and area-weighted sizes were determined.

MaxEnt size analysis

For the MaxEnt size analysis the constraint $C(p)/C_{aim} \sim 1$ was mostly satisfied, but in some cases was found to be much greater than unity (in the range of $2.5 - 5.5$). These results are concerning in that theoretically we expect $C(p)/C_{aim} = 1$. This can be explained in terms of the observed asymmetry in the MaxEnt specimen profiles. Size broadened profiles are symmetrical, as shown by the simulations in Chapter 4. As pointed out above, the asymmetry is due to absorption in the alumina specimen. This asymmetry affects the misfitting of the trial and specimen profiles, in (4.12), and forces the MaxEnt method to converge onto a larger $C(p_a)/C_{aim}$ value. The reason why this was not observed in §6.2.1 is that the uncertainty regions for the specimen profiles were greater. However, in the present results the uncertainty region has been reduced and the asymmetry becomes more apparent in determining the column-length distribution.

The MaxEnt size results, $\langle L\rangle_a^{M,n}$, in Figure 6.9 are uniform over the 2θ-range compared with $\langle L\rangle_a^F$ in the same figure. The average $\langle L\rangle_a^{M,n}$ was found to be (64.3 ± 0.8) nm. The average offset between the $\langle L\rangle_a^{M,n}$ and $\langle L\rangle_a^F$ results is about 7 nm. In these results the offset has been reduced compared with the $\langle L\rangle_a^{M,n}$ and $\langle L\rangle_a^F$ results in Figure 6.5. This indicates that the non-uniform model using (6.5) and Table 6.2 has better described the specimen broadening, but the offset between the size results may suggest the residual broadening in the reference material has not been accounted for in the deconvolution.

MaxEnt column-length distribution

The MaxEnt column-length distributions for the 012, 024, 113 lines and 226 are shown in Figures 6.10 & 6.11.

For the multiple-orders the column-length distributions are expected to be the same, as they represent the distribution of the crystallites in the same crystallographic direction. However, from Figure 6.10 it can be seen that the 012 and 024 MaxEnt column-length distribution are not the same. Figure 6.10(a) shows a large uncertainty region about the solution distribution and spurious oscillations which have no physical meaning. For the 024 column-length distribution, in Figure 6.10(b), the uncertainty region has been reduced and is also well-conditioned. In both of these cases the uncertainty region encompasses the *a*

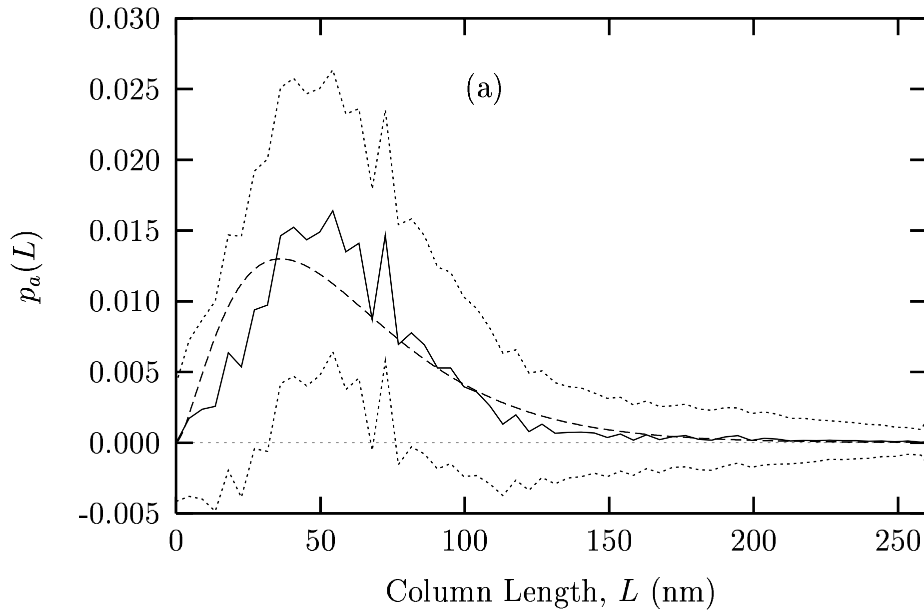
priori model – implying that the actual distribution lies in this region. The distributions in Figure 6.10 produce similar $\langle L\rangle_a^{M,n}$ values of (61 ± 1) nm and (62.8 ± 0.4) nm for the 012 and 024, respectively.

The MaxEnt column-length distributions for the 113 and 226 lines are shown in Figure 6.11. The uncertainty regions encompass the *a priori* model and indicate that the region of greatest uncertainty is about the maximum and the tails of the distribution. Both results are reasonably well-conditioned. The positivity of these results has also been preserved.

The $\langle L\rangle_a^{M,n}$ for these distributions are also very similar, (65.8 ± 0.3) nm and (64.8 ± 0.6) nm for the 113 and 226, respectively. However the distributions are not the same.

It is interesting to observe the similarities and differences between the column-length distributions in Figure 6.10 & 6.11. For example, all distributions have extended tails which decrease to zero at $L \approx 250$ nm indicating the maximum dimensions which contributes to size broadening. Moreover, the 113 and 024 results are very similar. By overlaying the distributions, they are seem to be almost identical. The separation of the diffraction lines is small, only about $9\,°2\theta$. On the other-hand, the 012 and 226 column-length distributions do not share any similarities; the peak positions are separated by $70\,°2\theta$. However, Fourier coefficients and integral breadth analyses indicate that the size broadening is independent of $[hkl]$. This would also imply that the column-length distributions should be the same or very similar to each other. It is also worth noting that relative to the *a priori* model all the distributions have been shifted torward larger L and the tails are extended. This feature in the distribution produces the offset between the $\langle L\rangle_a^{F}$ and $\langle L\rangle_a^{M,n}$ i.e. the average of the distribution is biased towards larger L.

How can the differences with the multiple-order distributions and similarities between the 024 and 113 distributions be explained? We know that the instrument profiles contain residual strain broadening and the strain broadening follows approximately a $\tan\theta$ dependency. In other words, the residual broadening is dependent on the position, 2θ; hence, the difference between the multiple-order distributions, for example 012 and 024 in Figure 6.10, can be explained as follows: in the deconvolution stage, the MaxEnt method over-compensates for the instrumental broadening and produces a narrower specimen profile relative to the actual profile. In the second stage of determining the column-length distribution, the Max-



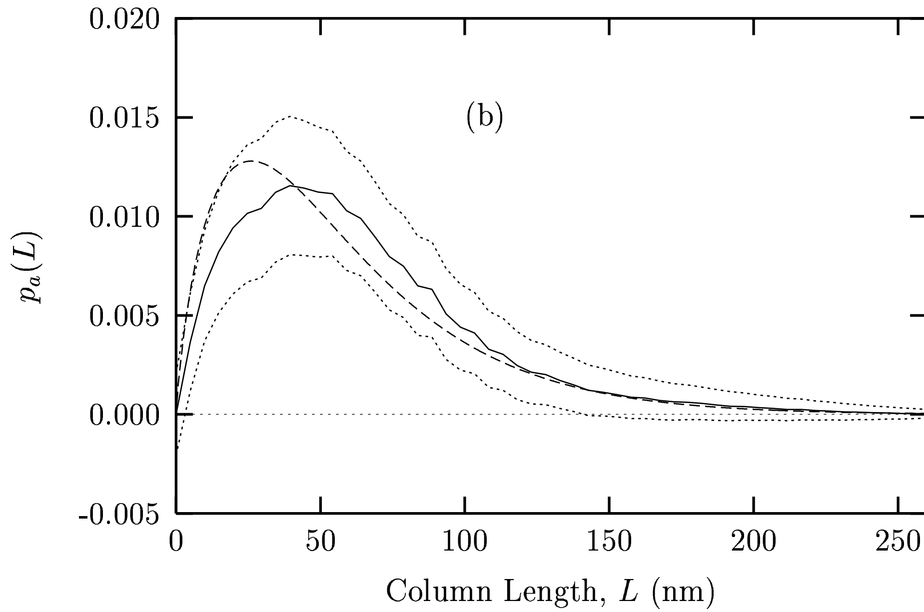
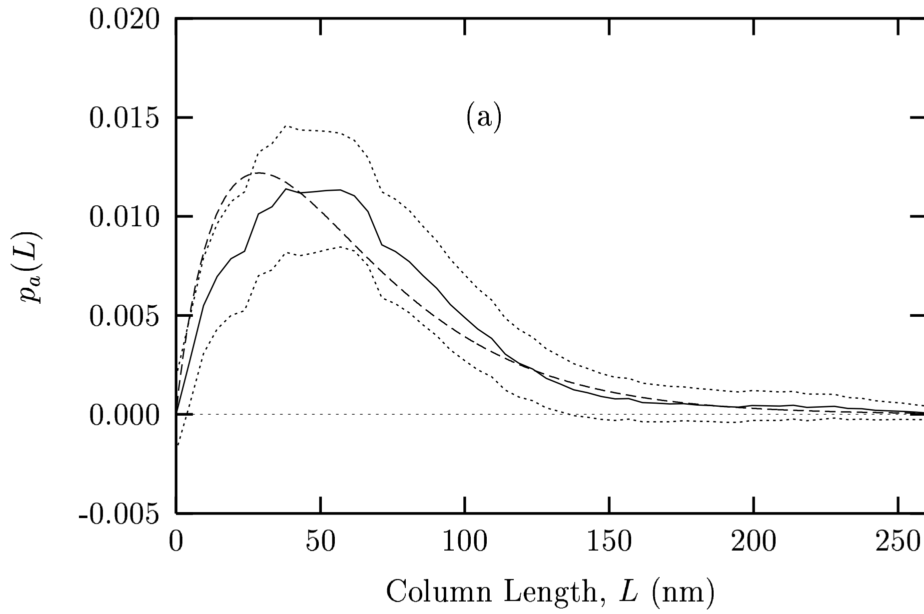


Figure 6.10: The MaxEnt column-length distributions for the 012&024 multiple-orders: (a) MaxEnt 012 column-length distribution (solid line), non-uniform *a priori* model (dashed line), and lower- & upper-uncertainty regions (dots); (b) MaxEnt 024 column-length distribution (solid line), non-uniform *a priori* model (dashed line), and lower- & upper-uncertainty regions (dots).



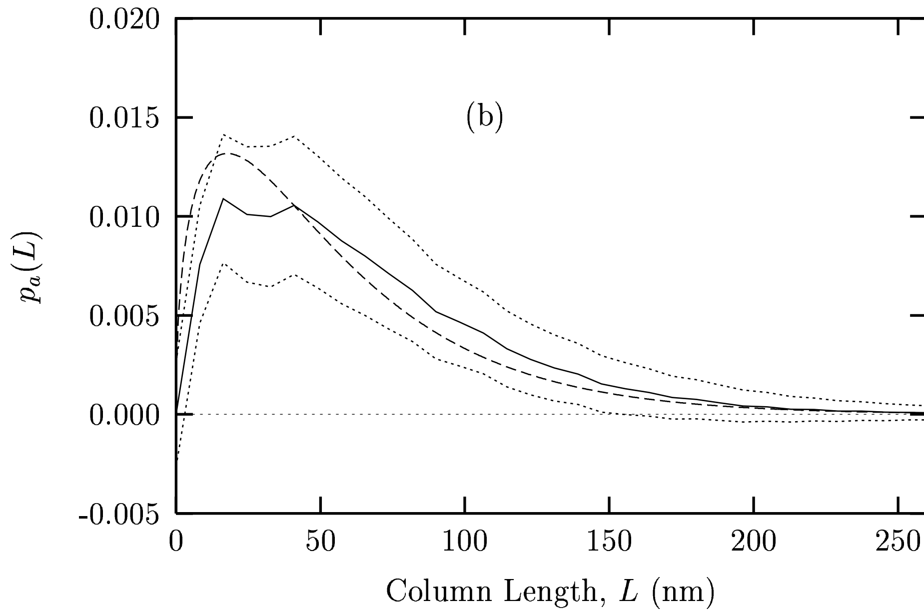


Figure 6.11: The MaxEnt column-length distributions for the 113&226 multiple-order: (a) MaxEnt 113 column-length distribution (solid line), non-uniform *a priori* model (dashed line), and lower- & upper-uncertainty regions (dots); (b) MaxEnt 226 column-length distribution (solid line), non-uniform *a priori* model (dashed line), and lower- & upper-uncertainty regions (dots).

Ent method produces a distribution biased towards larger crystallites in order to satisfy the experimental data, while the separation between the 024 and 113 (only about 9 °2θ) is not significant enough for the residual broadening to make any difference between the distributions. The asymmetry observed in the specimen profile prevents the MaxEnt method from converging on the preferred $C(p_a)/C_{aim}$ value.

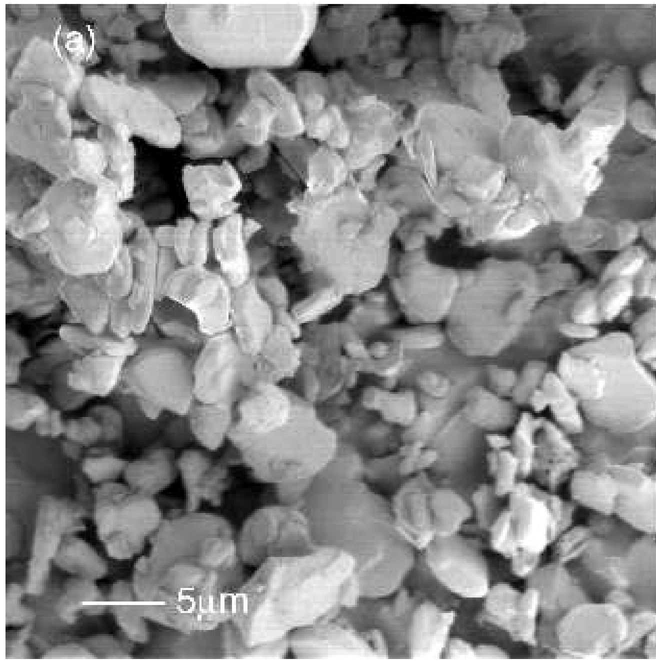
6.3.6 Electron microscopy

A scanning electron microscopy (SEM) examination was carried out on the alumina SRM 676 samples. The sample preparation involved coating the powders with Au/Pd film of about 30 nm thickness. This prevented deflection of the electron-beam and/or the charging of the specimen (Stevens-Kalceff 1999). A Joel 6300 field emission scanning electron microscope was used. The images were taken at 2 keV beam energy and 0.6 nA beam current (Stevens-Kalceff 1999).

"Overview" and "close-up" SEM images of the SRM 676 sample are shown in Figure 6.12. Figure 6.12(a) shows the alumina powder varying in size and shape and giving the appearance of single particles. Most distinctive are the platelets of $4-5\,\mu$m in diameters and thickness of about $1\,\mu$m. These particles have near-hexagonal morphology and may be an indication that the "de-aggregation process" (Kalceff et al. 1995) was incomplete. Also noticeable in Figure 6.12(a) are particles having a range of shapes, from spherical to platelet fragments, with a dimension of about $1-3\,\mu$m.

Figure 6.12(b) shows the alumina powder at higher magnification. It is clear from this image that these smaller particles also have a range of shapes. It is interesting to note there are smaller particles of less than $1\,\mu$m in dimension. These smaller particles are possibly the end-product of the de-aggregation process.

From these images it is unclear what is the exact origin of the specimen broadening in the x-ray diffraction profiles, but it is worth outlining a number of possible sources. If the platelets pointed out in Figure 6.12(a) are truly single particles then very little size-boadening would be expected from them. However, if the platelets consist of domains, these would result in incoherent scattering and contribute to size broadening of the diffraction profiles. In order for this to be the case, the boundaries between the domains would have to



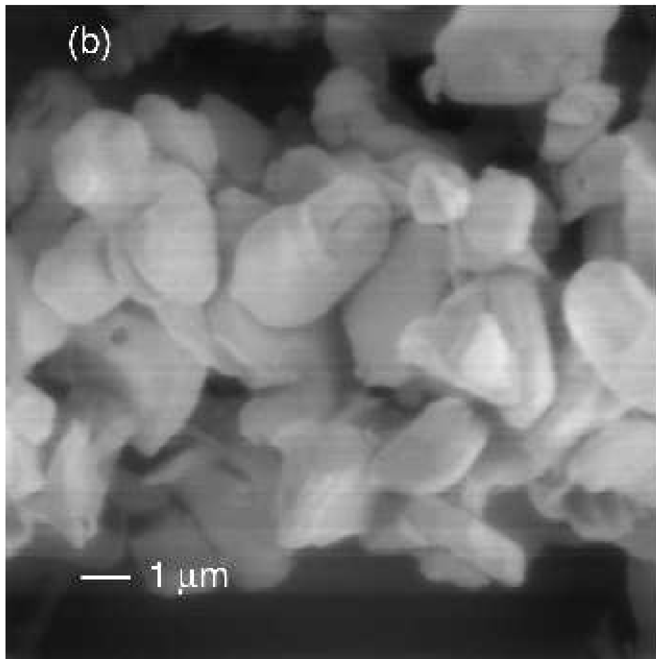


Figure 6.12: SEM images of the alumina SRM 676 sample: (a) An overview of the alumina sample at 2500× magnification; (b) A close-up of the alumina sample at 7500× magnification.

produce discontinuities in the crystal lattice structure. This in turn implies that the strain is concentrated over a very narrow-region to produce the discontinuity. Small-angle grain boundaries proposed by Kalceff et al. (1995) would account for this. This also explains why there is neglible, if any, strain broadening present in the diffraction profiles, since the strain is concentrated in narrow regions. The particles of $\leq 1\,\mu$m pointed out in Figure 6.12(b) could also contribute to size broadening if they are indeed single particles. However, if these particles consist of domains then this too would contribute to incoherent scattering. The extended tails of the column-length distributions (see Figure 6.10 & 6.11) suggest that there is a significant range of domain sizes with a maximum dimension of $250 - 300$ nm and an average between $57 - 64$ nm. The spherical morphology which was suggested by the integral-breadth plots and Fourier coefficients appear to be mostly an averaging over the shape of the crystallites/domains.

Cathodoluminescence examination of the alumina sample was carried out in an attempt to image the domains. However, this proved to be unsuccessful due to the resolution of the SEM, a Joel 3500, which effectively has a probe focus of size $\sim 1\,\mu$m (compared with the FEG-SEM used above, probe focus ~ 1 nm). There could be some defect migration occurring during the cathodoluminescence observations, contributing to further blurring of the image (Stevens-Kalceff 1999).

6.4 Summary

The analysis presented here has applied the maximum entropy (MaxEnt) analysis to alumina x-ray data at two levels. In the first, two cases were considered, where uniform and non-uniform *a priori* models were applied in the MaxEnt method (§6.2 & 6.3, respectively). It was shown that the uniform model, which defines our lack of knowledge concerning the specimen profile, produces large uncertainties in the specimen profile, while the non-uniform model, determined from the Fourier coefficients, reduces the uncertainties.

In the second level of application the specimen profiles were used to determine the column-length distribution. The large uncertainty in the specimen profiles from the first case (uniform model) produced compromised MaxEnt column-length distributions and size results. However, the specimen profiles from the second case (non-uniform model) produced

reasonable MaxEnt column-length distributions, and size results that were comparable to the Fourier size results. The differences between the Fourier and MaxEnt sizes and between the column-length distributions was explained by the residual strain broadening in the instrument profile. This clearly demonstrated the need for an instrument reference material which is free of residual broadening effects. Despite these difficulties, the specimen profiles, column-length distribution and area-weighted sizes determined from the MaxEnt method are most satisfactory and enabled the uncertainties for these quantities to be determined.

The MaxEnt analysis has also attempted to address the problem of determining the *a priori* model. Two approaches were investigated: firstly, the use of a "low resolution" method attempted to estimate the column-length distribution from the specimen profiles, but proved to be unsuccessful because of the ill-conditioning of the estimate of the column-length distribution resulting from large uncertainties in the specimen profile. The second approach modelled the Fourier coefficients of the specimen profile, which contain information about the shape and broadening of the specimen profile, and proved to be successful. However, further development of *a priori* models which incorporate crystallite/domain shape and estimate of the overall dimensions is needed.

The nature of the specimen broadening was established by examining the integral breadth and Fourier coefficients. It was found that broadening was independent of d^* and of $[hkl]$, which implied that the major source of broadening was crystallite/domain size effects. This is also consistent with a spherical or near-spherical crystallite/domain morphology. The Fourier and MaxEnt method estimated the average area-weighted sizes to be in the range of $57 - 64\,\mathrm{nm}$. The size analysis also indicated that there was a large range of sizes, with maximum dimension from $250 - 300\,\mathrm{nm}$. However, the presence of residual strain broadening in the instrument reference, SRM 660, may have over-shadowed any inhomogeneous strain in the alumina sample.

Electron microscopy examinations showed the de-aggregated particles to be in the range of $1 - 5\,\mu\mathrm{m}$. They also showed that the crystallite shape varied from near-spherical to platelets with a six-fold symmetry (the thickness of the platelets was about $1\,\mu\mathrm{m}$). Moreover, the electron microscopy examinations revealed single crystallites of less than $1\,\mu\mathrm{m}$ in dimension with a range of shapes. It is proposed that the specimen broadening is a combination of

crystallites with submicron dimensions and domains existing within the larger crystallites. The spherical or near-spherical morphology inferred from the x-ray data is an averaging over variously shaped crystallites/domains.

Chapter 7

Conclusion

The aim of this study was to develop an general method for analysing x-ray diffraction profiles. This entailed solving three problems: deconvolution of the observed profile; determining the column-length distribution and average size from crystallite/domain size-broadened profiles; and determining the strain distribution and root-mean-square strain from strain broadened profiles. It has been demonstrated the maximum entropy (MaxEnt) method is best suited for solving these problems, as it determines the solution with the least assumptions, preserves the positivity and additivity of the solution, incorporates *a priori* information into the solution procedure and quantifies the reliability of the solution. It also proved to be a robust and flexible method enabling a two-fold procedure to be developed for analysing size and strain broadened profiles, respectively.

The study consisted of numerical simulations and modelling of instrumental, crystallite/domain and strain broadened profiles. It also attempted to simulate the conditions under which the data is collected by introducing statistical noise, a background level and a finite instrument response into the simulated data. Moreover, the analysis examined the issues of deconvolving the observed profile in the presence of noisy data, using a non-ideal instrument profile, and the effects of background estimation and truncation of the observed profile on the physical quantities. These are issues which may affect the integrity of the physical quantities determined and need to be investigated if an alternative method of analysis is to be adopted.

In general the MaxEnt method has proven to be successful in solving these three problems,

and in all cases producing results superior to the conventional methods of deconvolution, size and strain analysis.

7.1 Evaluation of common deconvolution methods

The common deconvolution/inversion methods applied in x-ray diffraction profile analysis were evaluated by applying an eigen-system analysis to formulate the error-bound function (see Chapter 2). In the case of the unconstrained deconvolution/inversion methods it was shown that the error in the background estimation and statistical noise contribute to the ill-conditioning of the solution. This established an analytical result which has been observed in numerical calculations (for example see Kalceff et al. 1995). The significance of this study was in expressing the error-bound function as a function of the control parameters for the iterative and constrained deconvolution/inversion methods. This enabled the individual components, such as the statistical noise, and background estimation to be quantified in terms of the control parameters and an optimum solution defined (i.e. when the error-bound function was a minimum).

Moreover, it was demonstrated that even for the optimum solution the positivity was not preserved, with negative oscillations present. The error analysis provided a means of understanding which components were most likely to contribute to the ill-conditioning of the solution. This analysis not only highlighted the difficulties of deconvolving out the instrument profile, but also, in the case of the iterative and constrained deconvolution/inversion methods, can be extended to determining the column-length and strain distributions. The results from this analysis clearly demonstrated that these methods would not be suited for such a task.

7.2 Determining the column-length distribution

Drawing from a general discussion of the MaxEnt method (see Chapter 3), a novel two-fold procedure was proposed for determining the specimen profile, column-length distribution and area-weighted size. The MaxEnt method was shown to fullfill the basic needs for determining the specimen profile and column-length distribution (see Chapter 4). By determining

the column-length distribution, useful information concerning the distribution of crystallite sizes can be obtained. The two-fold method does not make any assumptions about the specimen broadening, but enables the nature of the specimen broadening to be determined, before proceeding with the second half of the analysis. It was also realised that the two problems of determining the specimen profile and the column-length distribution could be "rolled" into an "all in one" approach (see Appendix C). However, this approach assumes the crystallite/domain size broadening is the only source of broadening. On the other hand, it enables the column-length distribution to be determined directly from the observed data.

The two-fold MaxEnt approach for size analysis was evaluated using simulated diffraction lines modelled on the 113 and 226 alumina peaks, where the specimen consisted of spherical crystallites. Three case were examined: (i) determining the column-length distribution and area-weighted size for the best background level; (ii) the effect of background-estimation on the column-length and average size results; and (iii) the effect of deconvolving with a non-ideal instrument profile on the specimen profile and column-length distribution.

In (i), it was revealed that the consequence of truncation of the observed profile was to cause the column-length distribution to be shifted towards larger crystallites and produce erroneous size results. This was resolved by introducing a non-uniform *a priori* model, which attempts to fill in the missing information. The model also influences the uncertainty levels and the value of the Lagrangian parameters. In turn the uncertainties were usually reduced.

The results from (ii) demonstrate the sensitivity of the Fourier methods to incorrect background estimation. In the case of under-estimation of the background level, this produced an under-estimation of the area-weighted size, due the "kink effect". For over-estimation of the background level, there was an over-estimation of the area-weighted size, due to "hook effects". That is, small distortions to the Fourier coefficients near the origin produced large uncertainties in the area-weighted size. However, these difficulties were overcome using the MaxEnt method by incorporating a non-uniform *a priori* model.

The calculations from (iii) have a direct bearing on the development of instrument reference materials, in that it was demonstrated residual size broadening in the instrument profiles caused the MaxEnt method to over-compensate for the specimen broadening. The solution profiles were generally narrower relative to the actual specimen profiles and erro-

neous size results were produced. Hence, these calculations demonstrated the importance of the instrument profile in accurately characterising the optics of the diffractometer and the need to reduce the contribution of the profile from the reference material.

Although these calculations concentrated on the effects of truncation, background estimation and non-ideal instrument profiles, they did not demonstrate how to determine a suitable non-uniform *a priori* model. It was proposed that "low resolution" methods could be applied to estimate the model. Furthermore, a Bayesian method for determining the background level, Lagrangian parameter and determining a suitable model could be developed. However, the central issue of these calculation was to show how the MaxEnt method could be used to determine the specimen profile and column-length distribution. The deconvolution of the observed profile only considered the use of a uniform model and can be considered the worst case scenario. On defining the worst case scenario, the MaxEnt method is capable of incorporating *a priori* knowledge of the shape and broadening of the specimen profile and the column-length distribution to produce an improved result.

7.3 Determining the strain distribution

The two-fold MaxEnt method was applied to determining the strain distribution from strain-broadened profiles (see Chapter 4). This problem proved to be challenging due to insufficient data, especially in the case of high symmetry lattices. The issue of determining a suitable *a priori* model was also discussed. The MaxEnt strain analysis was compared with the traditional approaches, namely the Williamson & Hall (1953) and Warren & Averbach (1950, 1952) methods. It was shown that the strength of the MaxEnt method over traditional methods was that it does not make specific assumption as to the nature of strain broadening, but enables any available information to be incorporated into the solution procedure. It also provide information concerning the distribution of strain perpendicular to the diffraction planes.

The two-fold approach for strain analysis was evaluated using simulated data. The simulations were modelled for copper and the strain was assumed to be isotropic. The strain models were taken from empirical models proposed in the literature. Two cases were considered: (i) inhomogeneous and isotropic Gaussian strain distributions; (ii) inhomogeneous

and isotropic non-Gaussian (near-Lorentzian) strain distributions. Instrumental broadening, background-level and statistical noise were imparted to produce a simulated observed profiles.

In the case of (i), difficulties were encountered determining the specimen profile using the MaxEnt method due to the high level of statistical noise, incorrect background estimation and relatively little instrumental broadening (compared with the specimen broadening). In the case of the high-angle profile, truncation of the observed profile was a particular problem, which reduced the range of application for the various methods. For these calculations a uniform model was used in determining the specimen profile. Due to relatively little instrument broadening, the problem of determining the specimen profile was essentially a problem of noise removal.

The strain analysis demonstrated the failure of the classical Williamson & Hall (1953) method in not accounting for the d^{*2} dependence of the integral breaths. In this particular case an unphysical negative intercept was produced. The Warren & Averbach (1950, 1952) method proved to be sensitive to the errors in the Fourier coefficients. That is, due to the "hook effect" the rms-strain was under-estimated for small L. A "low resolution" approach was suggested for determining the model and in the case of Gaussian strain it proved to be successful. In applying the MaxEnt method, the *a priori* model proved to be critically important in determining the shape and tails of the strain distribution.

The difficulties encountered in (i), also persisted in case (ii). As in case (i), the classical Williamson & Hall (1953) method simply failed to account for the d^{*2} dependence of the integral breadth. The Warren & Averbach (1950, 1952) method also failed, not for numerical reasons, but because if its assumption of the strain distribution. This resulted in the under-estimation of the rms-strain. As in case (i), using the non-uniform model in determining the strain distribution proved to be important. Visually, the MaxEnt results appeared to be very good. However, on determining the rms-strain, the results were disappointing. It was also demonstrated that uncertainties in determining the model were being imparted into the MaxEnt solution.

The significance of these calculations was that the *a priori* model plays an important role is determining the shape and tails of the strain distribution. The results also suggest

that a Bayesian approach may be necessary in determining a suitable model and underlying parameters. The limitations of this study were that strain distributions were not related to microscopic properties such as the dislocation densities and dislocation type. Also, the problem of anisotropic strain broadening was not considered; it will require an extension of the present work.

7.4 Analysis of SRM 676 x-ray diffraction data

The two-fold MaxEnt method was applied to the NIST SRM 676 x-ray diffraction data (see Chapter 6). This data was first analysed by Kalceff et al. (1995) and the present results were in agreement with that study. However, the present analysis went further in determining the column-length distribution from the x-ray diffraction data, providing information about the range of crystallite sizes.

The analysis used both uniform and non-uniform models in determining the specimen profile and non-uniform models in determining the column-length distribution. In the case of determining the specimen profile, the non-uniform model was determined from Fourier coefficients. The use of this model reduced the uncertainty region in the specimen profile, but also highlighted the need for absorption correction in the diffraction data.

The size analysis consisted of analysing the alumina data using the Williamson & Hall (1953), Fourier and MaxEnt methods. The qualitative analysis revealed that crystallite/domain size broadening was the dominant source of specimen broadening, and also demonstrated that the residual strain broadening in the LaB_6 (NIST SRM 660) instrument profile affected the integral breadth and Fourier coefficients. This aspect of the analysis highlights the need for an instrument reference which is free from structural imperfections. It was proposed that the difference between the Fourier and MaxEnt area-weighted sizes can be attributed to strain in the LaB_6, while the scatter in the Fourier size results can be attributed to numerical errors near the origin of the Fourier coefficients. The presence of strain in the LaB_6 may also overshadow the presence of strain in the alumina data, especially at the high angles. Further analysis using an annealed reference material is required to overcome these difficulties in the analysis. In general, the analysis indicates that the alumina powder consists of scattering domains smaller than the overall dimensions of the powder particle observed by

a scanning electron microscopy examination. The difference between the x-ray and electron microscopy results can be reconciled if the alumina crystallites consist of scattering domains, created by strain concentrated over a narrow region so as to produce a discontinuity in the lattice structure.

The significance of this analysis is that the MaxEnt method when applied to alumina diffraction data preserved the positivity of the specimen profile. It enabled the column-length distribution to be determined and information concerning the distribution of crystallite sizes to be extracted from the diffraction data. Although there were difficulties with the reference material, the MaxEnt method produced physically consistent results.

7.5 Further research

The application of the MaxEnt method for determining the specimen profile, column-length and strain distributions has proven to be successful and can be wided with further research.

7.5.1 Determining the specimen profile

- The analysis of the common deconvolution methods concentrated on defining ill-conditioning in terms of the control parameters of the method. Another issue encountered in profile analysis is the range of reliability for varying degrees of instrumental broadening; this analysis could encompass the direct convolution approaches, as well as the MaxEnt method.

- The issue of determining an *a priori* model for the specimen has proven to be important in this study and requires further work. There is also an opportunity to incorporate a Bayesian/MaxEnt approach for determining the background level and reduce the truncation effects.

7.5.2 Crystallite/domain size analysis

- A systematic study could be undertaken for assess the reliability of *a priori* models in determining the column-length distribution. This could incorporate the common-

volume function of the crystallite to provide information concerning the shape of the crystallites in determining a suitable model for the column-length distribution.

- Bayesian analysis could be applied to determining the average overall dimensions of the crystallites.

- Further development of the "all in one" MaxEnt approach is required. This approach can determine the column-length distribution from the diffraction data. The solution column distribution can then be used to calculate the specimen profile, without needing the deconvolution stage.

7.5.3 Strain analysis

- The present work has highlighted the need to assess the reliability of the *a priori* model in the MaxEnt method for strain analysis. This may require adopting physical models concerning the nature of strain broadening. This would also enable physical quantities for a particular model to be determined from the MaxEnt results.

- The present analysis determines the strain distribution and rms-strain from the x-ray diffraction data. Further research is needed to relate the rms-strain to a dislocation model, where the dislocation density and type can be determined. This would relate the macroscopic description of strain broadening with a microscopic picture of the strain broadening.

- The present study has only considered the problem of isotropic strain broadening. However, to make this analysis applicable to the general case, anisotropic strain broadening must also be considered. This would require a microscopic understanding of the strain broadening and the introduction of contrast factors to "unscramble" the strain Fourier coefficients.

7.5.4 Size/strain analysis

- The problem of separating the size and strain contribution has not been considered in this study. The existing methods such as the Warren & Averbach (1950, 1952) rely on

the expansion of the Fourier coefficients over a limited range. However, the specimen profile can be considered as the convolution of a size profile and strain profile. The difficulty is that the size and strain profiles are unknown. This is essentially a problem of *blind deconvolution*, a problem which Newton (1985) has studied. It is proposed that a blind deconvolution method be developed for separating the size and strain profiles.

- Blind deconvolution could also be combined into an "all in one" approach, by introducing the instrument profile in the blind deconvolution method and determining the size and strain profiles directly from the data.

Appendix A

Common Profiles

The common analytical profile functions are presented in this appendix. These function are used in the fitting of x-ray diffraction profiles and the expressions presented here have been drawn from Langford & Louër (1996).

Notation

$f(0)$	peak maximum minus the background level;
x	position of the profile maximum: angular position, $(2\theta - 2\theta_B)$; reciprocal position, $s - s_0$;
$2w$	full-width at half maximum;
β	integral breadth (peak area/peak maximum);
ϕ	$= 2w/\beta$, the shape factor for the profile;
L	Lorentzian component;
G	Gaussian component.

Lorentzian function

$$f(x) = f(0)\,\frac{1}{1 + \left(\frac{x}{w}\right)^2} \tag{A.1}$$

where the integral breadth is given by

$$\beta = \pi w \tag{A.2}$$

and

$$\phi = 2/\pi. \tag{A.3}$$

Gaussian function

$$f(x) = f(0)\ \exp\left[-\pi\, x^2/\beta^2\right] \tag{A.4}$$

where the integral breadth is given by,

$$\beta = \frac{w}{\sqrt{\ln 2/\pi}} \tag{A.5}$$

and

$$\phi = 2\sqrt{\ln 2/\pi} \tag{A.6}$$

Voigt function

The Voigt function is the convolution of the Gaussian and Lorentzian functions

$$f(x) = \int_{-\infty}^{+\infty} f_L(x-\xi)\, f_G(\xi)\, \mathrm{d}\xi \tag{A.7}$$

and the convolution of two Voigt functions is also a Voigtian (Langford 1978). The analytical expression for the Voigt function is given by (Langford 1978)

$$f(x) = f(0)\beta\, \beta_G^{-1}\Re\{w(z)\} \tag{A.8}$$

where

$$z = \frac{\sqrt{\pi}x}{\beta_G} + \imath k \tag{A.9}$$

and

$$k = \frac{\beta_L}{\sqrt{\pi}\beta_g}. \tag{A.10}$$

Here k is the Voigt parameter and $\Re\{w(z)\}$ is the real part of the complex error function (see equation (21) in Langford 1978). The Gaussian component of the integral breadth, β_G is given by

$$\beta_G = \beta \exp(k^2)\ [1 - \mathrm{erf}(k)]\,. \tag{A.11}$$

The shape paramater, ϕ, for a Voigt function is bounded below and above by the Lorentzian (A.3) and Gaussian (A.6) ϕ-values, respectively (Langford & Louër 1996).

Pseudo-Voigt function

This function is a linear combination of the Gaussian and Lorentzian function

$$f(x) = f(0)\left[\eta f_L(x) + (1-\eta) f_G(x)\right] \tag{A.12}$$

where η is the mixing parameter. For $\eta = 1$, (A.12) becomes a Lorentzain function, while for $\eta = 0$, (A.12) becomes a Gaussian function. Unlike the Voigt function, the convolution of pseudo-Voigt functions is not a Voigtian.

Pearson-VII function

This function is a "tunable" in that the free parameter, w, and the exponent, m, can be determined to fit profiles between the Gaussian and Lorentzian limits

$$f(x) = f(0)\,\frac{1}{(1+Cx^2)^m} \tag{A.13}$$

where

$$C = \frac{2^{1/m}-1}{w^2} \tag{A.14}$$

and

$$\beta = \frac{\pi 2^{2(1-m)}\Gamma(2m-1)w}{(2^{1/m}-1)\Gamma^2(m)}. \tag{A.15}$$

The exponents, m, determine the rate at which the tails decrease. For $m = 1$, (A.13) becomes a Lorentzian function and as $m \to \infty$, (A.13) appproaches a Gaussain function; usually $m = 10$ is satisifactory for approximating a Gaussian. In the case of $m = 1.5$ (A.13) becomes an intermediate-Lorentzian and for $m = 2$ (A.13) is know as a modified-Lorentzian function.

Appendix B

The effect of $P(D)$ on $\langle L \rangle_v$ & $\langle L \rangle_a$

In this appendix the influence of the particle distribution on the two apparent dimensions, the area- and volume-weighted sizes, is discussed using the example of spherical crystallites. Here the particle distribution is considered as the distribution of a given dimension; in the case of a sphere this would correspond to a distribution of diameters, $P(D)$. A similar approach could be applied to other shaped crystallites such as cubes and tetrahedra. However, the apparent dimensions will then be dependent on the direction of the diffraction vector.

B.1 $\langle L \rangle_v$ & $\langle L \rangle_a$ for a single size

The results developed by Stokes & Wilson (1942) introduce the common (or ghost) function for a particular shaped crystallite. However, they assumed that the crystallites all had the same dimensions. Mathematically, this corresponds to a delta function centered about a particular dimension. The apparent dimension can be related to the physical dimensions, such as the diameter and "true" crystallite size by taking into consideration the common-volume function.

For a sphere (1.7) (also see equation (44) in Stokes & Wilson 1942), the area-weighted size becomes

$$\langle L \rangle_a = \frac{2}{3} D. \tag{B.1}$$

Defining the area-weighted Scherrer constant as (Langford & Wilson 1978)

$$\mathcal{K}_a = \frac{\mathfrak{p}}{\langle L \rangle_a} \tag{B.2}$$

where $\mathfrak{p}$ is the cube-root of the crystallite volume, for a sphere, $\mathfrak{p} = \left(\frac{\pi}{6}\right)^{\frac{1}{3}} D$ and $\mathcal{K}_a$ is $\frac{3}{2}\left(\frac{\pi}{6}\right)^{\frac{1}{3}}$.

Similarly the volume-weighted size is

$$\langle L \rangle_v = \frac{3}{4} D \tag{B.3}$$

and the corresponding Scherrer constant is $\frac{4}{3}\left(\frac{\pi}{6}\right)^{\frac{1}{3}}$.

Taking the ratio of the area- to volume-weighted sizes, we have

$$\frac{\langle L \rangle_a}{\langle L \rangle_v} = \frac{\mathcal{K}_v}{\mathcal{K}_a} \tag{B.4}$$

$$= \frac{8}{9}. \tag{B.5}$$

As mentioned above, these results apply only when the crystallites all have the same shape and dimension.

B.2 $\langle L \rangle_v$ & $\langle L \rangle_a$ for a distribution

To investigate how a distribution of spherical particles affects the area- and volume-weighted sizes, we must apply (1.25). In the case of a sphere, the area-weighted distribution, $p_a(L)$, is given by (Smith 1976)

$$p_a(L) = \frac{\pi}{2} L \int_L^{\infty} P(D)\, \mathrm{d}D \tag{B.6}$$

where the shape kernel is taken from the second derivative of (1.7). From (B.6) $\langle L \rangle_a$ can be determined; similarly, (1.22) can be used to detemine the volume-weighted distribution from which $\langle L \rangle_v$ can be calculated. Hence, in the case of a sphere we notice that $P(D) \rightarrow \delta(D - \langle D \rangle)$, $\frac{\langle L \rangle_a}{\langle L \rangle_v} \rightarrow \frac{8}{9}$.

To test this, a set of numerical calculations were conducted. The particle size distibution, $P(D)$, was set as a Gaussian distribution with an average diameter, $\langle D \rangle = 50\,\mathrm{nm}$. Using (B.6) and (1.22), $\langle L \rangle_a / \langle L \rangle_v$ was calculated for a range of standard-deviaions, σ_D. In the case of a Gaussian function as $\sigma_D \rightarrow 0$, $P(D) \rightarrow \delta(D - \langle D \rangle)$. The numerical results are shown in Figure B.1.

From Figure B.1, the intercept with the y-axis produces $0.89 \approx \frac{8}{9}$. Hence, the effect of the particle distribution is a "smearing" of the area- (or volume-)weighted distributions.

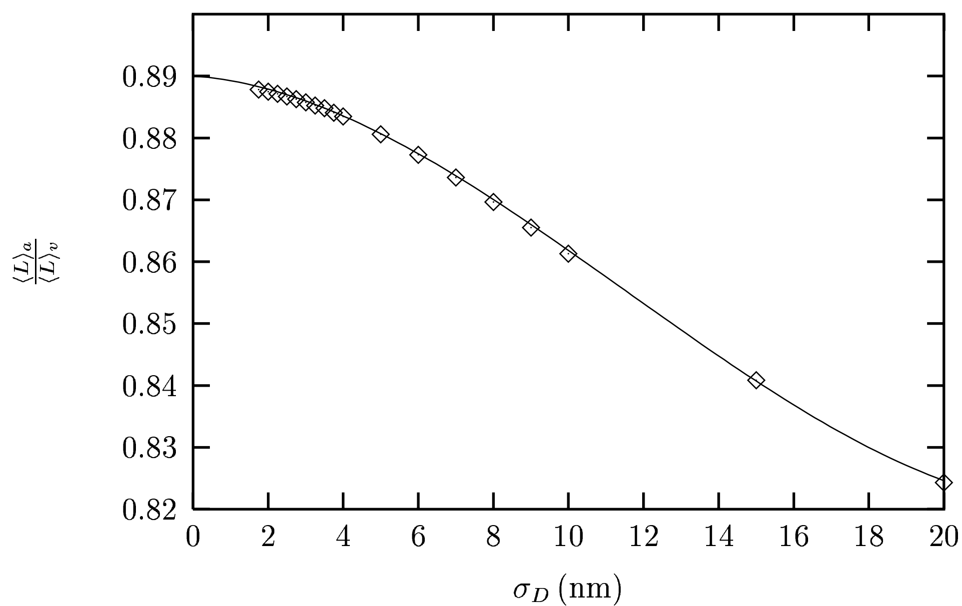


Figure B.1: $\frac{\langle L\rangle_a}{\langle L\rangle_v}$ *versus* σ_D for spherical crystallites. In the limit of $\sigma_D \to 0$, $\frac{\langle L\rangle_a}{\langle L\rangle_v} \to \frac{8}{9}$ and the specimen can be considered as consisting of particle of the same shape and size. Numerical values of $\frac{\langle L\rangle_a}{\langle L\rangle_v}$ (diamonds) and the fitted curve (solid line) using a third order polynomial are shown.

Appendix C

The MaxEnt "all in one" approach

In Chapter 4, a two-fold MaxEnt procedure is proposed to remove the instrumental broadening and determining the column-length distribution. This application does not make any assumptions concerning the nature of the specimen broadening and allows it to be determined before proceeding with the second part of the analysis.

However, if size broadening is the only source of specimen broadening, the two-fold procedure can be "rolled" into a single method. The advantage of this approach is that observed data can be used directly to determine the column-length distribution. Once the column-length distribution in known the specimen profile can be determined, avoiding the difficulty of deconvolution.

C.1 Overview of the "all in one" approach

The "all in one" approach combines the instrument kernel and scattering kernel into a single kernel. The instrument kernel must be mapped from 2θ-space into $(s - s_0)$-space and expressed as a $[M \times M]$ convolution matrix.

Using the scattering kernel in (4.18), the new kernel becomes

$$W_{ij} = \sum_{p=1}^{M} \sum_{q=1}^{M} k_{ip} \, K_{qj} \tag{C.1}$$

where k_{ip} is the instrument kernel; K_{qi} is the scattering kernel; $i = 1, 2, 3 \ldots, M$ and

$j = 1, 2, 3 \ldots, N$ such that $M \geq N$. The statistic function becomes

$$C(p_a) = \sum_i \frac{(\tilde{g}_i - \hat{g}_i)^2}{\sigma_i^2} \tag{C.2}$$

where $\hat{g}_i = \sum_j W_{ij} \, p_{a\,j}$ and the $\sigma_i^2 = \tilde{g}_i$. The entropy function given by (4.11) can be used to fomulate the Lagrangian function and this incorporated into the Skilling & Bryan (1984) algortihm.

As an example of the application of this method, it has been used in the simulated data presented in Chapter 4, with column-length distribution and simulated observed profile for the 113 peak for spherical crystallites with $\langle D \rangle = 50\,\text{nm}$ (also see §4.4.4). A non-uniform *a priori* model was used in this calculation with parameters $u = 0.09$, $t = 1.0$ and $r = 3.0$ (see (4.14) & (4.16)).

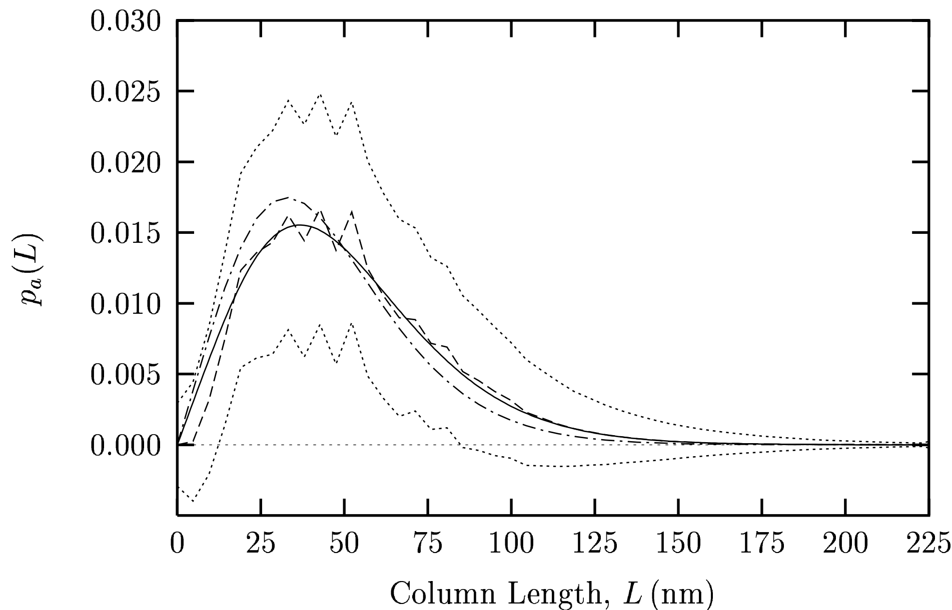


Figure C.1: MaxEnt "all in one" results for optimum background level for the 113 peak corresponding to $\langle D \rangle = 50\,\text{nm}$: True column-length distribution (solid line), the solution column-length distribution (dashed line), the lower- & upper-uncertainty regions (dots), and the uniform *a priori* model (dash-dot lines).

The analysis in Figure C.1 was halted after 40 iteration when it was evident that the algorithm would not converge onto $C(p_a)/C_{aim} \sim 1$. The final $C(p_a)/C_{aim}$ was 10.5 which is unacceptable. Also, the uncertainty regions for the solution distribution are very large. These two results suggest that there are difficulties with the statistic function, (C.1), and possibly

the kernel, (C.2). Further investigation is necessary to refine this approach, but nevertheless the results indicate that the two problems can be combined into a single method.

Bibliography

Adler, T. & Houska, C. R. (1979), 'Simplifications in the x-ray line-shape analysis', *J. Appl. Phys.* **50**(5), 3282–3287.

Arfken, G. B. & Weber, H. J. (1996), *Mathematical Methods for Physicists*, 4th edn, Academic press, San Diego.

Armstrong, N. & Kalceff, W. (1998), 'Eigen-system analysis of x-ray diffraction profile deconvolution methods explains ill-conditioning', *J. Appl. Cryst.* **31**, 453–460.

Armstrong, N. & Kalceff, W. (1999*a*), 'A maximum entropy method for determining column-length distributions from size broadened x-ray diffraction profiles', *accepted by J. Appl. Cryst.* . Manuscript No. HW72.

Armstrong, N. & Kalceff, W. (1999*b*), 'A maximum entropy method for determining the microstrain distribution from broadened x-ray diffraction profiles', *in preparation* .

Audebrand, N., Auffrédic, J. P. & Louër, D. (1998), 'X-ray diffraction study of early stages of growth of nanoscale zinc oxide crystallites obtained from thermal decomposition of four precursors. General concepts on precursor-dependent microstructal properties.', *submitted to J. Appl. Cryst.* .

Auffredic, J. P., Ciosmak, D., Louër, D. & Niepce, J. C. (1980), The relationship between oxide crystallite shape and the crystalline structure of the precursor in thermal decomposition: Case of zinc oxide, *in* 'Reactivity of Solids: Proc. 9th Inter. Sym. React. of Sol.', Vol. 2, Cracow, pp. 826–830.

Averbach, B. L. & Warren, B. E. (1949), 'Interpretation of X-ray patterns of cold-worked metal', *J. Appl. Phys.* **20**, 885–886.

Balzar, D., Ledbetter, H. & Roshko, A. (1993), 'X-ray diffraction peak-broadening analysis of (La-M)$_2$CuO$_4$ high-T_c superconductors.', *Powd. Diffr.* **8**(1), 2–6.

Balzar, D. & Popović, S. (1996), 'Reliability of the simplified integral-breadth methods in diffraction line-broadening analysis', *J. Appl. Cryst.* **29**, 16–23.

Benedetti, A., Fagherazzi, G., Enzo, S. & Battagliarin, M. (1988), 'A profile-fitting procedure for analysis of broadened x-ray diffraction peaks. II. Application and discussion of the methodology', *J. Appl. Cryst.* **21**, 543–549.

Bertaut, E. F. (1950), 'Raies de Debye-Scherrer et répatition des dimensions des domains de Bragg dans les poudres polycristallines', *Acta Cryst.* **3**, 14–18.

Bertaut, E. F. (1952), 'Sur la correction de la transformée de Fourier d'une raie de Debye-Scherrer dans la measure de dimensions cristallines', *Acta Cryst.* **5**, 177–121.

Bienenstock, A. (1961), 'Determination of crystallite size distributions from x-ray line broadening', *J. Appl. Phys.* **32**(2), 187–189.

Bienenstock, A. (1963), 'Calculation of crystallite size distributions from x-ray line broadening', *J. Appl. Phys.* **34**(5), 1391.

Bonetto, R. D., Viturro, H. R. & Alvarez, A. G. (1993), 'XTL-SIZE: a computer program for crystallite-size-distribution from x-ray diffraction line broadening', *J. Appl. Cryst.* **23**, 136–137.

Bryan, R. K. (1990), 'Maximum entropy analysis of oversampled data problems', *Eur. Biophys. J.* **18**, 165–174.

Burden, R. L. & Faires, J. D. (1993), *Numerical analysis*, 5th edn, PWS Pub. Com., Boston.

Cheary, R. W. & Cline, J. P. (1995), 'An analysis of the effect of different instrumental conditions on the shapes of x-ray powder line profiles', *Adv. X-Ray Anal.* **38**, 75–82.

Cheary, R. W. & Coelho, A. (1992), 'A fundamental parameters approach to x-ray line profile fitting', *J. Appl. Cryst.* **25**, 109–121.

Cheary, R. W. & Coelho, A. (1996), Collaborative Computational Project No. 14 in Powder and Single Crystal Diffraction, Programs XFIT and FOURYA. URL://www.dl.ac.uk/CCP/CCP14.

Cheary, R. W. & Coelho, A. A. (1998*a*), 'Axial divergence in a conventional x-ray powder diffractometer. I. Theoretical foundations', *J. Appl. Cryst.* **31**, 851–861.

Cheary, R. W. & Coelho, A. A. (1998*b*), 'Axial divergence in a conventional x-ray powder diffractometer. II. Realization and evaluation in a fundamental-parameter profile fitting procedure', *J. Appl. Cryst.* **31**, 862–868.

Cohen, J. B. (1966), *Diffraction methods in material science*, MacMillan Com., New York.

Cox, R. T. (1946), 'Probability, frequency and reasonable expectation', *Amer. J. Phys.* **14**(1), 1–13.

de Bergevin, F. & Germi, P. (1972), 'Correction des oscillations dans les distributions de tailles de particules obtenues à partir des profils de raies de diffraction', *J. Appl. Cryst.* **5**, 416–420.

de Keijser, T. H., Langford, J. I., Mittemeijer, E. J. & Vogels, A. B. P. (1982), 'Use of the Voigt function in a single-line method for the analysis of x-ray diffraction line broadening', *J. Appl. Cryst.* **15**, 308–314.

de Keijser, T. H. & Mittemeijer, E. J. (1980), 'Notes on the handling of position and broadening errors in deconvoluted x-ray diffraction line profiles', *J. Appl. Cryst.* **13**, 74–77.

de Keijser, T. H., Mittemeijer, E. J. & Rozendaal, C. F. (1983), 'The determination of crystallite-size and lattice-strain parameters in conjunction with the profile-refinement method for the determination of crystal structures', *J. Appl. Cryst.* **16**, 309–316.

Delhez, R., de Keijser, T. H. & Mittemeijer, E. J. (1980), 'Accuracy of crystallite size and strain values from x-ray diffraction line profiles using Fourier series', *Nat. Bur. Stand. Spec. Pub.* **567**, 213–253.

Delhez, R., de Keijser, T. H. & Mittemeijer, E. J. (1982), 'Determination of crystallite size and lattice distortions through x-ray diffraction line profile analysis: Recipes, methods and comments', *Fres. Z. Anal. Chem.* **312**, 1–16.

Delhez, R., de Keijser, T. H., Mittemeijer, E. J. & Langford, J. I. (1986), 'Truncation in diffraction pattern analysis. I. Concept of a diffraction line profile and its range', *J. Appl. Cryst.* **19**, 459–466.

Delhez, R. & Mittemeijer, E. J. (1976), 'The elimination of an approximation in the Warren-Averbach analysis', *J. Appl. Cryst.* **9**, 233–234.

Donoho, D. H., Johnstone, I. M., Hoch, J. C. & Stern, A. S. (1992), 'Maximum entropy and the nearly black object', *J. R. Statist. Soc. B* **54**(1), 41–81.

Eastabrook, J. N. & Wilson, A. J. C. (1952), 'The diffraction of x-rays by distorted-crystral aggregates. III. Remarks on the interpretation of the Fourier coefficients', *Proc. Phys. London* **B65**, 67–75.

Enzo, S., Fagherazzi, G., Benedetti, A. & Polizzi, S. (1988), 'A profile-fitting procedure for analysis of broadened x-ray diffraction peaks. I. Methodology', *J. Appl. Cryst.* **21**, 536–542.

Ergun, S. (1968), 'Direct method of unfolding convolution products–its application to x-ray scattering intensities', *J. Appl. Cryst.* **1**, 19–23.

Fawcett, T. G., Crowder, C. E., Brownell, S. J., Zhang, Y., Hubbard, C. R., Schreiner, W., Hamill, G. P., Huang, T. C., Sabino, E., Langford, J. I., Hamiton, R. & Louër, D. (1988), 'Establishing an instrument peak profile calibration standard for powder diffraction: International round robin conducted by the JCPDS-ICDD and the U.S. National Bureau of Standards', *Powder Diff.* **3**, 209–218.

Frieden, B. R. (1972), 'Restoring with maximum likelihood and maximum entropy', *J. Opt. Soc. Am.* **62**(4), 511–518.

Geoghegan, M., Jones, R. A. L., Sivia, D. S., Penfold, J. & Clough, A. S. (1996), 'Experimental study of surface segregation and wetting films of a partially miscible polymer blend', *Phys. Rev. E* **53**(1), 825–837.

Gilmore, C. J. (1996), 'Maximum entropy and Bayesian statistics in crystallography: A review of practical applications', *Acta Cryst.* **A52**, 561–589.

Goambo, F. & Gassiat, E. (1997), 'Bayesian methods and maximum entropy for ill-posed inverse problems', *Annals Stat.* **25**(1), 328–350.

Grebille, D. & Bérar, J. F. (1985), 'Calculation of diffraction line profiles in the case of a major size effect: Application to boehmite AlOOH', *J. Appl. Cryst.* **18**, 301–307.

Groma, I., Ungár, T. & Wilkens, M. (1988), 'Asymmetric x-ray line broadening of plastically deformed crystals. I. Theory', *J. Appl. Cryst.* **21**, 47–53.

Guérin, D. M. A., Alvarez, A. G., Neira, L. E. R., Plastino, A. & Bonetto, R. D. (1986), 'Statistical inference and crystallite size distributions', *Acta Cryst.* **A42**, 30–35.

Guinier, A. (1963), *X-ray diffraction in crystals, imperfect crystals and amorphous bodies*, W. H. Freeman, San Francisco. Reprinted by Dover (New York), 1994.

Gull, S. F. (1989), Developments in maximum entropy data analysis, *in* J. Skilling, ed., 'Maximum Entopy and Bayesian Methods', Kluwer Acad. Publ., Netherlands, pp. 53–71.

Gull, S. F. & Skilling, J. (1984), 'Maximum entropy image reconstruction', *IEE Proc.* **F131**, 646–659.

Handler, N. C. & Wanger, C. N. J. (1966), 'Separation of particle size and lattice strain in integral breadth measurements', *Acta Cryst.* **20**, 312–313.

Hansen, S. & Pedersen, J. S. (1991), 'A comparison of three different methods for analysing small-angle scattering data', *J. Appl. Cryst.* **24**, 541–548.

Hansen, S. & Wilkins, S. W. (1994), 'On uncertainty in maximum-entropy maps and the generalization of classic MaxEnt', *Acta Cryst.* **A50**, 547–550.

Harrison, J. W. (1966), 'The use of strain moments in determining strain distributions in deformed crystals', *Acta Cryst.* **20**, 390–396.

Harrison, J. W. (1967), 'The determination of strain distribution and particle-size by moments method', *Acta Cryst.* **23**, 175–176.

Hosemann, V. R., Lange, A. & Hentschel, M. P. (1985), 'Mechanische spannungen in mikroparakristallen', *Acta. Cryst.* **A41**, 434–440.

Howard, S. A. & Snyder, R. L. (1989), 'The use of direct convolution products in profile and pattern fitting algorithms. I. Development of the algorithms', *J. Appl. Cryst.* **22**, 238–243.

Huang, T. C. & Parrish, W. (1977), 'Quantative analysis of complicated mixtures by profile fitting x-ray diffraction patterns', *Adv. X-ray Anal.* **21**, 275–288.

James, R. W. (1948), *The optical principles of diffraction of x-rays*, Vol. II, Bell & Hyman.

Jaynes, E. T. (1957), 'Information theory and statistical mechanics', *Phys. Rev.* **106**(4), 620–630.

Jaynes, E. T. (1968), 'Prior probabilities', *IEEE Trans. SSC* **4**(3), 227–240.

Jaynes, E. T. (1982), 'On the rationale of maximum-entropy methods', *Proc. IEEE* **70**(9), 939–952.

Jaynes, E. T. (1983), Where do we stand on maximum entropy?, *in* R. D. Rosenkrantz, ed., 'E. T. Jaynes: Papers on probability, statistics and statistical physics', D. Reidel Pub., Dordrecht, pp. 210–314.

Johnson, R. W. & Shore, J. E. (1983), 'Comments on and correction to "Axiomatic derivation of the principle of maximum entropy and principle of minimum cross-entropy"', *IEEE Trans. IT* **26**(6), 942–943.

Kalceff, W. (1998), Private commuincation.

Kalceff, W., Armstrong, N. & Cline, J. P. (1995), 'An evaluation of deconvolution techniques in x-ray profile broadening analysis and the application of the maximum entropy method to alumina data', *Adv. X-Ray Anal.* **38**, 387–395.

Kalceff, W., Cline, J. P. & von Dreele, R. B. (1994), 'Size/strain broadening analysis of SRM-676 candidate materials', *Adv. X-ray Anal.* **37**, 343–349.

Klimanek, P. & Kužel, R. (1988), 'X-ray diffraction line broadening due to dislocations in non-cubic materials. I. General considerations and the case of elastic isotropy applied to hexagonal crystals', *J. Appl. Cryst.* **21**, 59–66.

Klug, H. P. & Alexander, L. E. (1974), *X-ray diffraction procedures: For polycrystalline and amorphous materials*, Wiley & Sons, New York.

Kogan, V. A. & Kupriyanov, M. F. (1992), 'X-ray powder diffraction line profiles by Fourier synthesis', *J. Appl. Cryst.* **25**, 16–25.

Kreyszig, E. (1993), *Advanced Engineering Mathematics*, 7th edn, Wiley, New York.

Krivoglaz, A. & Ryaboshapka, K. P. (1963), 'Theory of x-ray scattering by crystals containing dislocations. Screw and edge dislocations randomly distributed throughtout the crystal', *Fiz. metal. metalloved.* **15**(1), 18–31.

Krivoglaz, M. A. (1995), *X-ray and neutron diffraction from nonideal crystals*, Springer, Berlin.

Kužel, R. & Klimanek, P. (1988), 'X-ray diffraction line broadening due to dislocations in non-cubic materials. II. The case of elastic anisotropy applied to hexagonal crystals', *J. Appl. Cryst.* **21**, 363–368.

Lagendijk, R. L. & Biemond, J. (1991), *Iterative identification and restoration of images*, Klumer Academic Press, Boston.

Langford, J. I. (1978), 'A rapid method for analysing the breadths of diffraction and spectral lines using Voigt function', *J. Appl. Cryst* **11**, 10–14.

Langford, J. I. (1980), 'Accuracy of crystallite size and strain determined from the integral breadth of powder diffraction lines', *Nat. Bur. Stand. Spec. Pub.* **567**, 255–269.

Langford, J. I. (1992), 'The use of the Voigt function in determining microstructural properties from diffraction data by means of pattern decomposition', *NIST Spec. Pub.* **846**, 110–126.

Langford, J. I., Boultif, A., Auffrédic, J. P. & Louër, D. (1993), 'The use of pattern decomposition to study the combined x-ray diffraction effects of crystallite and stacking faults in ex-oxalate zinc oxide', *Appl. Cryst.* **26**, 22–33.

Langford, J. I., Delhez, R., de Keijser, T. H. & Mittemeijer, E. J. (1988), 'Profile analysis for microcrystalline properties by Fourier and other methods', *Aust. J. Phys.* **41**, 173–87.

Langford, J. I. & Louër, D. (1982), 'Diffraction line profiles and Scherrer constants for materials with cylindrical crystallites', *J. Appl. Cryst.* **15**, 20–26.

Langford, J. I. & Louër, D. (1996), 'Powder diffraction', *Rep. Prog. Phys.* **59**, 131–234.

Langford, J. I., Louër, D., Sonneveld, E. T. & Visser, J. W. (1986), 'Applications of total pattern fitting to a study of crystallite size and strain in zinc oxide powder', *Powd. Diff.* **1**(1), 211–221.

Langford, J. I. & Wilson, A. J. C. (1978), 'Scherrer after sixty years: A survey and some new results in the determination of crystallite size', *J. Appl. Cryst.* **11**, 102–113.

Larson, A. C. & von Dreele, R. R. (1994), *GSAS: General Structure Analysis System*, Los Alamos National Laboratory, Los Alamos.

LeBail, A. & Louër, D. (1978), 'Smoothing and validity of crystallite-size disribution from x-ray line-profile analysis', *J. Appl. Cryst.* **11**, 50–55.

Lele, S. & Anantharaman, T. R. (1966), 'Influence of crystallite shape on particle size broadening of Debye-Scherrer reflections', *Proc. Ind. Acad. Sci.* **A64**, 261–274.

Leoni, M. & Langford, J. I. (1998), 'Line profile analysis and the Rietveld method: crossing paths ?', *CPD Newsletter* **20**, 30.

Loredo, T. J. (1990), From Laplace to Supernova SN-1987A: Bayesian inference in astrophysics, *in* P. F. Fougére, ed., 'Maximum entropy and Bayesian methods', Kluwer Acad. Publ., Netherlands, pp. 81–142.

Louboutin, R. & Louër, D. (1972), 'Méthode directe de correction des profiles de raies de diffraction des rayons X. III. Sur la recherch de la solution optimale lors de la déconvolution', *Acta Cryst.* **A28**, 396–400.

Louër, D. (1994), 'Applications of profile fitting analysis for micro-crystalline properties from total pattern fitting', *Adv. X-ray Anal.* **37**, 27–36.

Louër, D., Auffrédic, J. P. & Langford, J. I. (1983), 'A precise determination of the shape, size and distribution of size of crystallites in zinc oxides by x-ray line-broadening analysis', *J. Appl. Cryst.* **16**, 183–191.

Louër, D. & Langford, J. I. (1988), 'Peak shape and restoration in conventional diffractometry with monochromatic x-rays', *J. Appl. Cryst.* **21**, 430–437.

Louër, D. & Weigel, D. (1969), 'Méthode directe de correction des profiles de raies de diffraction des rayons X. II. Influence de la fente réceptrice sur l'enregistrement d'un profil de raie de diffraction X', *Acta Cryst.* **A25**, 338–350.

Louër, D., Weigel, D. & Langford, J. I. (1972), 'Etude des profils de raises de diffraction de rayons X d'une poudre d'hydroxyde de nickel', *J. Appl. Cryst.* **5**, 353–359.

Louër, D., Weigel, D. & Louboutin, R. (1969), 'Méthode directe de correction des profiles de raies de diffraction des rayons X. I. Méthode numérique déconvolution', *Acta Cryst.* **A25**, 335–338.

Luenberger, D. G. (1984), *Linear and nonlinear programming*, 2nd edn, Addison-Wesley, Massachusetts.

Lutterotti, L. & Scardi, P. (1992), 'Profile fitting by the interference function', *Adv. X-Ray Anal.* **35**, 577–584.

MacKay, D. J. C. (1991), Bayesian interpolation, *in* C. R. Smith, G. J. Erickson & P. O. Neudorfer, eds, 'Maximum Entropy and Bayesian Methods', Kluwer Acad. Pub., pp. 39–66.

McKeehan, M. & Warren, B. E. (1953), 'X-ray study of cold worked in thoriated tungsten', *J. Appl. Phys.* **24**(1), 52–56.

Mead, L. R. & Papanicolaou, N. (1984), 'Maximum entropy in the problem of moments', *J. Math. Phys.* **25**(8), 2404–2417.

Mignot, J. & Rondot, D. (1977), 'Utilisation de méthodes d'approximation pour l'etude des coefficients de développement en série de Fourier des raies de diffraction X', *Acta Cryst.* **A33**, 327–333.

Mortier, W. J. & Costenoble, M. L. (1973), 'The separation of overlapping peaks in x-ray powder patterns with use of an experimental profile', *J. Appl. Cryst.* **6**, 488–490.

Mughrabi, H. (1983), 'Dislocation wall and cell structures and long-range stresses in deformed metal crystals', *Act Metall.* **31**(9), 1367–1379.

Mughrabi, H., Ungár, T., Kienle, W. & Wilkens (1986), 'Long-range stresses and asymmetric X-ray line broadening in tensile-deformed [001]-orientated copper single crystals', *Phil. Mag. A* **53**(6), 793–813.

Müller, J., Hansen, S. & Pürschel, H. V. (1996), 'The use of small-angle scattering and the maximum entropy method for shape-model determination from distance-distribution functions', *J. Appl. Cryst.* **29**, 547–554.

Müller, J. J. & Hansen, S. (1994), 'A study of high-resolution x-ray scattering data evaluation by the maximum-entropy method', *J. Appl. Cryst.* **27**, 257–270.

Nabarro, F. R. N. (1952), 'The mathematical theory of stationary dislocations', *Adv. Phys.* **1**(3), 269–394.

Nabarro, F. R. N., Basinski, Z. S. & Holt, D. B. (1964), 'The plasticity of pure single crystals', *Adv. Phys.* **13**(49-52), 193–323.

Nandi, R. N., Kuo, H. K., Schlosberg, W., Wissler, G., Cohen, J. B. & Crist, B. (1984), 'Single-peak methods for Fourier analysis of peak shapes', *J. Appl. Cryst.* **17**, 22–26.

Newton, T. J. (1985), Blind deconvolution and related topics, PhD thesis, University of Cambridge.

Nunzio, P. E. D., Martelli, S. & Bitti, R. R. (1995), 'A Monte Carlo estimate of crystallite-size and microstrain distribution function from x-ray line broadening', *J. App. Cryst.* **28**, 146–159.

Phillips, D. L. (1962), 'A technique for the numerical solution of certain integral eequations of the first kind', *J. Assoc. Compt. Math.* **9**, 84–97.

Pines, B. Y. & Sirenko, A. F. (1962), 'Derivation of the lattice distortion and spread in block size by means of harmonic analysis applied to x-ray reflections', *Sov. Phys. Cryst.* **7**(1), 15–21.

Pines, B. Y. & Surovtsev, I. Y. (1963), 'The problem of determining the particle size and the microdeformation in a crystal lattice from the data on the harmonic analysis of the shape of the x-ray diffraction lines', *Sov. Phys. Cryst.* **8**(3), 390–393.

Popescu, F. & Benes, L. (1977), 'Elimation of the finite-range effect on the block-size distribution from the Fourier Transform', *Acta Cryst.* **A33**, 323–326.

Potton, J. A., Daniell, G. J. & Rainford, B. D. (1988*a*), 'A new method for the determination of particle size distributions from small-angle neutron scattering measurments', *J. Appl. Cryst.* **21**, 891–897.

Potton, J. A., Daniell, G. J. & Rainford, B. D. (1988*b*), 'Particle size distributions from SANS data using the maximum entropy method', *J. Appl. Cryst.* **21**, 663–668.

Păusescu, P., Mănăilă, R., Popescu, M. & Jijovici, E. (1974), 'Crystallite size and distribution in supported catalysts', *J. Appl. Cryst.* **7**, 281–286.

Rao, S. & Houska, C. R. (1986), 'X-ray particle-size broadening', *Acta Cryst.* **A42**, 6–13.

Rasberry, S. D. (1989), Certificate of analysis, SRM 660, "Instrument line profile and profile shape standard for x-ray powder diffraction", Technical report, NIST, Gaithersburg, Md., USA.

Rietveld, H. M. (1969), 'A profile refinement method for nuclear and magnetic structures', *J. Appl. Cryst.* **2**, 65–71.

Rondot, D. & Mignot, J. (1977), 'Méthode d'obtention d'un profil de raie de diffraction x à partir d'un profil expérimental composé de plusieurs raies superposées', *J. Appl. Cryst.* **10**, 84–90.

Rothman, R. L. & Cohen, J. B. (1971), 'X-ray study of faulting in BCC metals and alloys', *J. Appl. Phys.* **42**(3), 971–979.

Ruland, W. (1965), 'The integral width of the convolution of a Gaussian and a Cauchy distribution', *Acta Cryst.* **18**, 581.

Ruland, W. (1968), 'The separation of line broadening effects by means of line-width relations', *J. Appl. Cryst.* **1**, 90–101.

Ryaboshapka, K. P. & Tikhonov, L. V. (1961*a*), 'Deformation effects due to the presence of typical dislocations in elastically anisotropic metals with cubic symmetry', *Fiz. metal. metalloved.* **12**(1), 1–10.

Ryaboshapka, K. P. & Tikhonov, L. V. (1961*b*), 'Deformation effects in old-deformed metals which are related to characteristics, and the broadening of x-ray diffraction lines. I. The case of elastic isotropism', *Fiz. metal. metalloved.* **11**(4), 489–495.

Ryder, L. H. (1985), *Quantum Field Theory*, Cambridge Uni. Press, Cambridge.

Scardi, P., Lutterotti, L. & Maistelli, P. (1994), 'Experimental determination of the instrument broadening in Bragg-Brentano geometry', *Powd. Diff.* **9**(3), 180–186.

Schlenker, B. R. (1986), *Introduction to material science: SI edition*, 2nd edn, Jacaranda Press, Sydney.

Schlosberg, W. H. & Cohen, J. B. (1983), 'On-line fourier analysis of the shape of x-ray diffaction peaks', *J. Appl. Cryst.* **16**, 304–308.

Schoening, F. R. L. (1965), 'Strain and particle size values from x-ray line breadths', *Acta. Cryst.* **18**, 975–976.

Shaw, W. T. & Tigg, J. (1994), *Applied Mathematica: Getting started, getting done*, Addison-Wesley, Massachusetts.

Shore, J. E. & Johnson, R. W. (1980), 'Axiomatic derivation of the principle of maximum entropy and principle of minimum cross-entropy', *IEEE Trans. IT* **26**(1), 26–37.

Sivia, D. S. (1990), Applications of maximum entropy and Bayesian methods in neutron scattering, *in* P. F. Fougére, ed., 'Maximum Entropy and Bayesian Methods', Kluwer Pub. Acad., Netherlands, pp. 195–209.

Sivia, D. S. (1996), *Data Analysis: A Bayesian Tutorial*, Oxford Science Pub., Oxford.

Sivia, D. S. & David, W. I. F. (1994), 'A Bayesian approach to extracting structure-factor amplitudes from powder diffraction data', *Acta Cryst.* **A50**, 703–714.

Sivia, D. S., David, W. I. F., Knight, K. S. & Gull, S. F. (1993), 'An introduction to Bayesian model selection', *Physica* **D66**, 234–242.

Skilling, J. (1984), 'The maximum entropy method', *Nature* **309**, 748–749.

Skilling, J. (1989*a*), Classic maximum entropy, *in* J. Skilling, ed., 'Maximum Entropy and Bayesian Methods', Kluwer Acad. Publ., Netherlands, pp. 45–52.

Skilling, J. (1989*b*), The eigenvalues of mega-dimensional matrices, *in* J. Skilling, ed., 'Maximum Entropy and Bayesian Methods', Kluwer Acad. Pub., Netherlands, pp. 455–566.

Skilling, J. (1990), Quantified Maximum Entropy, *in* P. F. Fougére, ed., 'Maximum Entropy and Bayesian Methods', Kluwer Acad. Pub., Netherlands, pp. 341–350.

Skilling, J. & Bryan, R. K. (1984), 'Maximum entropy image reconstruction: general algorithm', *Mon. Not. R. astr. Soc.* **211**, 111–124.

Skilling, J. & Gull, S. F. (1985), Algorithms and applications, *in* C. R. Smith & J. W. T. Grandy, eds, 'Maximum-entropy and Bayesian methods in inverse problems', Kluwer Acad. Pub., Dordrecht.

Skilling, J., Robinson, D. R. T. & Gull, S. F. (1991), Probabilistic display, *in* 'Maximum Entropy and Bayesian Methods', Kluwer Academic Publishers, Netherlands, pp. 365–368.

Smith, W. L. (1972), 'Crystallite sizes and surface area of catalysts', *J. Appl. Cryst.* **5**, 127–130.

Smith, W. L. (1976), 'Crystallite shape and the Fourier analysis of diffraction line profiles', *J. Appl. Cryst.* **9**, 187–189.

Somashekar, R. & Somashekarappa, H. (1997), 'X-ray diffraction-line broadening analysis: Paracrystalline methods', *J. Appl. Cryst.* **30**, 147–152.

Sonneveld, E. J., Delhez, R., de Keijser, T. H. & Mittemeijer, E. J. (1991), 'Quality of unravelling of experimental diffraction patterns with artificially varied overlap', *Mat. Sci. For.* **79-81**, 85–90.

Stevens-Kalceff, M. A. (1999), Private communication.

Stokes, A. R. (1948), 'A numerical Fourier analysis method for the correction of width and shapes of the lines on x-ray powder photography', *Proc. Phys. Soc. London* **A61**, 382–391.

Stokes, A. R. & Wilson, A. J. C. (1942), 'A method of calculating the integral breadths of Debye-Scherrer lines', *Proc. Camb. Phil. Soc.* **38**, 313–322.

Stokes, A. R. & Wilson, A. J. C. (1944*a*), 'The diffraction of x-rays by distorted crystal aggregates–I', *Proc. Phys. Soc. Lond.* **56**, 174–181.

Stokes, A. R. & Wilson, A. J. C. (1944*b*), 'A method for calculating the integral breadths for Debye-Scherrer lines: Generalisation to non-cubic crystals', *Proc. Camb. Soc.* **40**, 197–198.

Strauss, C. E. M., Wolpert, D. H. & Wolf, D. R. (1993), Alpha, evidence and the entropic prior, *in* A. Mohammad-Djafari & G. Demoments, eds, 'Maximum entropy and Bayesian methods', Kluwer Acad. Pub., Netherlands, pp. 113–120.

Takahashi, H. (1969), 'Fourier coefficients of line profiles of Debye-scherrer rings broadened by lattice distortions', *J. Phys. Soc. Jap.* **27**(3), 708–712.

Tikhonov, A. N. (1963), 'On the solution of ill-posed problems and the method of regularization', *Soviet Math. Dokl.* **4**, 1035–1038.

Tikochinsky, Y., Tishby, N. Z. & Levine, R. D. (1984), 'Consistent inference of probabilities for reproductive experiments', *Phys. Rev. Lett.* **52**(16), 1357–1360.

Titterington, D. M. (1985), 'General structure of regularization procedures in image reconstruction', *Astron. Astrophys.* **144**, 381–387.

Titterington, D. M. (1994), 'The maximum entropy method for data analysis', *Nature* **312**, 381–382.

Turunen, M. J., de Keijser, T. H., Delhez, R. & van der Pers, N. M. (1983), 'A method for the interpretation of the Warren-Averbach mean-square strains and its application to recovery in aluminium', *J. Appl. Cryst.* **16**, 176–182.

Twomey, S. (1963), 'On the numerical solution of Fredholm integral equations of the first kind by inversion of the linear system produced by quadrature', *J. Assoc. Compt. Math.* **10**, 97–101.

Twomey, S. (1977), *Introduction to the mathematics of inversion in remote sensing and indirect measurements*, Elsevier Sci. Publ., New York.

Ungár, T. (1994), 'Characteristically asymmetric x-ray line broadening, an indication of residual long-range internal stresses', *Mat. Sci. For.* **166-169**, 23–44.

Ungár, T. & Borbély, A. (1996), 'The effect of dislocation contrast on x-ray line broadening: A new approach to line profile analysis', *J. Appl. Lett.* **69**(21), 3173–3175.

Ungár, T., Goma, I. & Wilkens, M. (1989), 'Asymmetric x-ray line broadening of plastically deformed crystals. II. Evaluation procedure and application to [001]-Cu crystals', *J. Appl. Cryst.* **22**, 26–34.

Ungár, T. H. (1967), 'Proposed method of separating the particle size and distortion coefficients in line profile analysis', *Acta Cryst.* **23**, 174–175.

Ungár, T., Mughrabi, H. & Wilkens, M. (1982), 'An x-ray line-broadening study of dislocations near the surface and in the bulk of deformed copper single crystals', *Acta. Metall.* **30**, 1861–1867.

Ungár, T., Ott, S., Sanders, P. G., Borbély, A. & Weertman, J. R. (1998), 'Dislocation, grain size and planar faults in nanostructured copper determined by high resolution x-ray diffraction and a new procedure of peak analysis', *Acta. mater.* **46**(10), 3693–3699.

Ungár, T., Révész, A. & Borbély, A. (1998), 'Dislocations and grain size in electrodeposited nanocrystalline Ni determined by modified Williamson-Hall and Warren-Averbach procedures', *J. Appl. Cryst.* **31**, 554–558.

Ungár, T., Tóth, L. S. & Kovács, I. (1986), 'Dislocation structure and work hardening in polycrystalline ofhc copper rods deformed by torsion and tension', *Acta Metall.* **34**(7), 1257–1267.

van Berkum, J. G., Delhez, R., de Keijser, T. & Mittemeijer, E. J. (1996), 'Diffraction-line broadening due to strain fields in materials: Fundamental aspects and methods of analysis', *Acta. Cryst.* **A52**, 730–747.

van Berkum, J. G. M. (1994), Strain fields in crystalline materials: Methods of analysis based on diffraction-line broadening, PhD thesis, Thesis Technische Universiteit Delft, Delft.

van Berkum, J. G. M., Vermeulen, A. C., Delhez, R., de Keijser, T. H. & Mittemeijer, E. J. (1993), 'Fourier methods for separation of size and strain broadening–validity of the Warren-Averbach and alterative analyses', *Mater. Sc. Forum* **133-136**, 77–82.

van Berkum, J. G. M., Vermeulen, A. C., Delhez, R., de Keijser, T. H. & Mittemeijer, E. J. (1994), 'Applicabilities of the Warren-Averbach analysis and an alternative analysis for separation of size and strain broadening', *J. Appl. Cryst.* **27**, 345–357.

Vargas, R., Louër, D. & Langford, J. I. (1983), 'Diffraction line profiles and Scherrer constants for materials with hexagonal crystallites', *J. Appl. Cryst.* **16**, 512–518.

Vermeulen, A. C., Delhez, R., de Keijser, T. H. & Mittemeijer, E. J. (1995), 'Changes in the densities of dislocations on distinct slip systems during stress relation in thin alumina layers: The interpretation of x-ray diffraction line broadening and the line shift', *J. Appl. Phys.* **77**(10), 5026–5048.

Vermeulen, A. C., Delhez, R., de Keijser, T. H. & Mittemeijer, J. E. (1991), 'A correction for truncation of powder diffraction line profiles', *Mater. Sci. For.* **79-82**, 119–124.

Vermeulen, A. C., Delhez, R., de Keijser, T. H. & Mittemeijer, J. E. (1992), 'Representation of tails of periodic and infinite-range signals: Towards a treatment for truncation', *J. Appl. Phys.* **71**(11), 5303–5309.

Vesely, F. J. (1994), *Computational physics: An introduction*, Plenum Press, New York.

Vogel, W. (1990), 'Size distribution of supported metal catalysts: An analytical x-ray line profile fitting routine', *J. Catal.* **121**, 356–363.

Vogel, W., Haase, J. & Hosemann, R. (1974), 'Linienprofilanalyse von Röentgen-Weitwinkelreflexen mittles Fourier-transformation zur Bestimmung von Mikrospannungen und parakristallinen', *Z. Naturfor.* **A29**, 1152–1158.

Warren, B. E. (1959), 'X-ray studies of deformed metals', *Prog. Metal Phys.* **8**, 147–202.

Warren, B. E. (1963), 'X-ray profiles from cold worked metals', *Acta Metall.* **11**, 995–996.

Warren, B. E. (1969), *X-ray Diffraction*, Addison-Welsy, Massachusetts.

Warren, B. E. & Averbach, B. L. (1950), 'The effect of cold-work distortion on x-ray patterns', *J. Appl. Phys.* **21**, 595–599.

Warren, B. E. & Averbach, B. L. (1952), 'The separation of cold-worked distortion and particle size broadening in x-ray patterns', *J. Appl. Phys.* **23**(4), 497.

Wilkens, M. (1970*a*), 'The determination of density and distribution of dislocations in deformed single crystals from broadened x-ray diffraction profiles', *Phys. Stat. Sol.A.* **2**, 359–370.

Wilkens, M. (1970*b*), The mean square stresses $\langle \sigma^2 \rangle$ for a completely random and restrictedly random distribution of dislocations in a cylindrical body, *in* J. A. Simmons, R. de Witt & R. Bullough, eds, 'Fundamental aspects of dislocation theory', Vol. 2, pp. 1191–1193.

Wilkens, M. (1970*c*), Theoretical aspects of kinematical x-ray diffraction profiles from crystals containing dislocation ditributions, *in* J. A. Simmons, R. de Witt & R. Bullough, eds, 'Fundamental aspects of dislocation theory', Vol. 2, pp. 1195–1221.

Wilkens, M. (1987), 'X-ray line broadening and mean square strains of straight dislocations in elastically anisotropic crystals of cubic symmetry', *Phys. Stat. Sol.* **104**, K1–K6.

Williamson, G. K. & Hall, W. H. (1953), 'X-ray line broadening from filed aluminium and wolfram', *Acta Met.* **1**, 22–31.

Williamson, G. K. & Smallman, R. E. (1954), 'The use of Fourier analysis in the interpretation of x-ray line broadening from cold-worked iron and molybdenum', *Acta Cryst.* **7**, 574–581.

Williamson, G. K. & Smallman, R. E. (1956), 'Dislocation densities in some annealed and cold-worked metals from measurements on the x-ray Debye-Scherrer spectrum', *Phil. Mag.* **1**, 34–46.

Wilson, A. J. (1963), *Mathematical theory of x-ray powder diffractometry*, Philips Publicaations, Eindhoven.

Wilson, A. J. C. (1949), 'The probability distribution of x-ray intensities', *Acta Cryst.* **2**, 318–321.

Wilson, A. J. C. (1952), 'The diffraction of x-rays by distorted-crystal aggregates. IV. Diffraction by a crystal with an axial screw dislocation', *Acta Cryst.* **5**, 318–322.

Wilson, A. J. C. (1962), 'On variance as a measure of line broadening in diffractometry: General theory and small particle size', *Proc. Phys. Soc.* **80**, 286–294.

Wilson, A. J. C. (1968), 'On variance as a measure of line broadening in diffractometry: Effect of distribution of sizes on the apparent crystallite size', *J. Appl. Cryst.* **1**, 194–196.

Wilson, A. J. C. (1969), 'Variance apparent particle sizes of cylinders, prisms and hemispheres', *J. Appl. Cryst.* **2**, 181–183.

Wilson, A. J. C. (1971), 'Some further considerations in particle-size broadening', *J. Appl. Cryst.* **4**, 440–443.

Wing, G. M. (1991), *A Primer on Integral Equations of the First Kind: The Problem of Deconvolution and Unfolding*, SIAM, Philadelphia. With assistance from J. D. Zahrt.

Wolpert, D. H. & Strauss, C. E. M. (1995), What Bayes has to say about the evidence procedure, *in* G. Heidbreder, ed., 'Maximum Entropy and Bayesian Methods', Kluwer Acad. Pub., Netherlands. To be published.

Wu, E., Gray, E. M. A. & Kisi, E. H. (1998), 'Modelling dislocation-induced anistropic line broadening in Rietveld refinements using a Voigt function. I: General principles.', *J. Appl. Cryst.* **31**, 356–362.

Wu, E., Kisi, E. H. & Gray, E. M. A. (1998), 'Modelling dislocation-induced anisotropic line broadening in Rietveld refinements using a Voigt function. II. Application to neutron powder difraction data.', *J. Appl. Cryst.* **31**, 363–368.

Wu, N. (1997), *The Maximum Entropy Method*, Springer-Verlag, Berlin.

Young, R. A., Gerdes, R. J. & Wilson, A. J. C. (1967), 'Propagation of some systematic errors in x-ray line profile analysis', *Acta Cryst.* **22**, 155–161.

Zocchi, M. (1980), 'An improved method for the determination of microstructural parameters by diffraction-profile Fourier analysis', *Acta. Cryst.* **A36**, 164–170.