



University  
of Glasgow

Mallik, Procheta Chandra Vasu (2010) *Diagnostics of solar flare energetic particles: neglected hard X-ray processes and neutron astronomy in the inner heliosphere*. PhD thesis, University of Glasgow.

<http://theses.gla.ac.uk/1510/>

Copyright and moral rights for this thesis are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

# Diagnostics of Solar Flare Energetic Particles: Neglected Hard X-ray Processes and Neutron Astronomy in the Inner Heliosphere

Procheta Chandra Vasu Mallik, B.Sc. Hons.

Astronomy and Astrophysics Group  
Department of Physics and Astronomy  
Kelvin Building  
University of Glasgow  
Glasgow, G12 8QQ  
Scotland, U.K.



University  
of Glasgow

Presented for the degree of  
Doctor of Philosophy  
The University of Glasgow  
October 2009

*“To my parents, Sushma and Dipankar, and my brother, Proteep”*

# Abstract

For work on my thesis dissertation (cf. [Mallik et al. 2009b](#)), we have been studying some energetic processes in solar flares. On our work on hard X-ray (HXR) emission from flares, we have shown that non-thermal recombination emission can compare with the bremsstrahlung HXR flux for certain flare conditions. In this thesis, we show spectral features characteristic of non-thermal recombination HXR emission and suggest how it plays a significant role in the flare HXR continuum, something that has been ignored in the past. It is important to note that these results could demand a reconsideration of the numbers of accelerated electrons since recombination can be much more efficient in producing HXR photons than bremsstrahlung. We go on to show that although non-thermal recombination is not likely to dominate the total HXR flux unless we consider extreme parameter regimes, it can still form a significant proportion of the HXR flux for typical flare conditions, thereby remaining important for both spectral inversion and low energy electron cut-off diagnostic capabilities.

In related work on diagnosing particle acceleration in flares (cf. [Mallik et al. 2009a](#)), we also have an interest in studying solar neutrons. To this end, this thesis presents our work done with new-age neutron detectors developed by our colleagues at the University of New Hampshire. Using laboratory and simulated data from the detector to produce its response matrix, we then employ regularisation and deconvolution techniques to produce encouraging results for data inversion. As a corollary, we have been reconsidering the role of inverse Compton scattering (ICS) of photospheric photons.  $\gamma$ -ray observations clearly show the presence of 100 MeV electrons and positrons in the solar corona, by-products of GeV energy ions. We present results of ICS of solar flare photons taking proper account of radiation field geometry near the solar surface. If observed, such radiation would let us determine the number of secondary positrons produced in large flares, contributing to a full picture of ion acceleration and to predicting neutron fluxes to be encountered by future inner heliosphere space missions.

*“Astronomy compels the soul to look upwards and leads us from this world to another.”*

– Plato

# Contents

List of Tables	iv
List of Figures	xiii
Preface	xiv
Acknowledgements	xvi
List of Abbreviations	xx
<b>1 Review of Energetic Processes in Solar Flares</b>	<b>1</b>
1.1 The Observations: an Overview . . . . .	3
1.1.1 Radio and microwave observations . . . . .	8
1.1.2 EUV Observations . . . . .	9
1.1.3 X-ray observations . . . . .	10
1.1.4 $\gamma$ -ray observations . . . . .	13
1.1.5 Observations of energetic particles . . . . .	13
1.2 Review of Relevant Theory . . . . .	16
1.2.1 X-ray emission . . . . .	16
1.2.2 Particle acceleration . . . . .	22
1.3 Review of Neutron Astronomy in the Inner Heliosphere . . . . .	24
1.3.1 Motivation . . . . .	24
1.3.2 Introduction to neutron detectors . . . . .	27
<b>2 Recombination as a Neglected Source of Hard X-rays</b>	<b>33</b>
2.1 Introduction . . . . .	33
2.2 Bremsstrahlung and recombination HXR spectra . . . . .	35
2.2.1 General considerations . . . . .	35
2.2.2 Bremsstrahlung . . . . .	35

2.2.3	Recombination Radiation . . . . .	36
2.2.4	Element parameters and flare plasma ionisation . . . . .	40
2.3	Local thin-target approximation for non-thermal electron distributions . . . . .	42
2.3.1	Basic expressions for $j_B, j_R$ . . . . .	42
2.3.2	Ratio of $j_R$ to $j_B$ . . . . .	43
2.3.3	Typical results in limiting regimes . . . . .	44
2.4	Whole flare thin-target, thermal and thick-target expressions . . . . .	48
2.4.1	Thin Target Coronal Loop . . . . .	48
2.4.2	Hot Coronal Loop Thermal Emission (in the Kramers approximation) . . . . .	49
2.4.3	Thick target (dense loop or footpoint) f-f and f-b emission spectra . . . . .	52
2.5	Practical case study results derived from a real flare . . . . .	54
2.6	Inference of electron spectra by inversion . . . . .	59
2.7	Recombination efficiency and different forms of $F(E)$ . . . . .	62
2.7.1	Comparison of the efficiency of f-b versus f-f HXR yield . . . . .	62
2.7.2	Ratio of $j_R$ to $j_B$ for an example of a smooth $F(E)$ with no cut-off . . . . .	63
2.8	Discussion and conclusions . . . . .	64
<b>3</b>	<b>Major Contribution of Non-thermal Recombination to Flare Hard X-rays</b> . . . . .	<b>67</b>
3.1	Introduction . . . . .	67
3.2	Contribution to deka-keV emission . . . . .	68
3.3	Implications of non-thermal recombination dominance . . . . .	73
3.4	Applications: coronal and footpoint sources . . . . .	74
3.4.1	Coronal HXR Sources . . . . .	74
3.4.2	Chromospheric Footpoint HXR Sources . . . . .	76
3.5	Conclusions . . . . .	77
<b>4</b>	<b>Inverse Compton Scattering</b> . . . . .	<b>80</b>
4.1	Introduction to ICS . . . . .	80
4.2	ICS versus NTB . . . . .	84
4.3	In the solar context and geometry . . . . .	85
4.4	ICS from relativistic flare electrons and positrons . . . . .	89
4.4.1	ICS from fast electrons . . . . .	89
4.4.2	ICS from relativistic positrons . . . . .	93
4.5	A comment on observations . . . . .	96

---

4.6	Conclusions and discussion . . . . .	99
<b>5</b>	<b>FNIT - the Fast Neutron Imaging Telescope</b>	<b>101</b>
5.1	Introduction . . . . .	101
5.2	Detection technique . . . . .	107
5.3	Laboratory and simulation data and results . . . . .	109
5.4	Simulations . . . . .	116
5.5	Characterising detector response . . . . .	120
<b>6</b>	<b>Data Inversion Techniques for Solar Neutron Detectors</b>	<b>125</b>
6.1	Relevance, importance and significance of inversion . . . . .	125
6.2	Inversion techniques and response matrix . . . . .	127
6.3	Choice of regularisation method . . . . .	129
6.4	Tikhonov regularisation and results . . . . .	130
6.5	Testing Tikhonov regularisation . . . . .	132
6.6	A preliminary study of blind deconvolution and results . . . . .	141
6.7	Conclusions . . . . .	143
<b>7</b>	<b>Conclusions and Future Work</b>	<b>146</b>
7.1	Recombination . . . . .	146
7.2	Neutrons and ICS . . . . .	147
	<b>Bibliography</b>	<b>150</b>
<b>A</b>	<b>Appendix to Chapters 2 and 3</b>	<b>164</b>
A.1	Dielectronic Recombination . . . . .	164
A.2	“Inverse” Recombination . . . . .	165
<b>B</b>	<b>Appendix to Chapter 4</b>	<b>168</b>

# List of Tables

2.1	Elements with their coronal abundances and ionisation potentials at $T \gg 10^8$ K . . . . .	38
2.2	Ionic species of iron at 20 MK . . . . .	38
2.3	Significance of Fe 25+ and 26+ at 4 different temperatures . . . . .	56

# List of Figures

1.1	A schematic representation of the different phases of a solar flare ( <a href="#">Kane 1974</a> ) . . . . .	4
1.2	The Sun as seen in various wavelengths by various modern instruments. Courtesy NASA/JPL-Caltech . . . . .	6
1.3	Schematic diagram of the solar radio spectrum from a “typical” large flare ( <a href="#">Dulk 1985</a> ). This simplified diagram can be quite different from observed flare radio spectra. . . . .	9
1.4	Observed RHESSI flux spectrum showing the distinct thermal and non-thermal components as well as the $\gamma$ -ray line and continuum emission extending to higher energies. Courtesy E. Kontar . . . . .	12
1.5	Typical solar X-ray flux spectrum showing the distinct Maxwellian thermal component and the non-thermal power-law tail . . . . .	17
1.6	Schematic diagram showing bremsstrahlung emission. Courtesy S. Bank, University of Wisconsin-Madison . . . . .	18
1.7	Schematic diagram showing synchrotron emission. Courtesy S. Bank, University of Wisconsin-Madison . . . . .	19
1.8	Schematic diagram showing inverse Compton emission. Courtesy S. Bank, University of Wisconsin-Madison . . . . .	20
1.9	Proximity to the Sun is essential to get a good signal of incident neutrons in the 1-20 MeV range. Courtesy UNH . . . . .	25
1.10	Upper-limit Quiet Sun neutron distributions assuming initially isotropic fast ion distributions at various distances from the Sun ( <a href="#">MacKinnon &amp; Ryan 2009</a> ). The plots are for a power-law primary ion distribution with $\delta = 2$ (left) and $\delta = 6$ (right). The four curves for both graphs are for four different distances from the Sun: 0.25, 0.33, 0.5 and 1.0 AU from top to bottom. . . . .	26

1.11	ToF spectrum showing the visible distinction between $\gamma$ -rays (peaking around 0 ns) and neutrons ( $\sim 10$ ns); back-scattered neutrons have a ‘negative’ ToF . . . . .	29
1.12	Schematic of a double scatter with the event cone and event circle/disc clearly represented. Courtesy UNH . . . . .	30
1.13	Source location identification using event circles for a GRB detected by COMPTEL. Courtesy UNH . . . . .	31
1.14	The SONTRAC instrument prototype. Courtesy UNH . . . . .	31
2.1	Actual shapes of non-thermal f-b and f-f spectra for different temperature regimes and non-thermal electron parameters. The solid curves represent non-thermal cool (red), hot (green) and ultra-hot (blue) f-b spectra as well as the f-f (black) spectrum. The dashed curves represent the total spectrum multiplied by 10. . . . .	44
2.2	Photon flux ratio of non-thermal f-b to f-f emission for different temperature regimes and parameters. Line styles have the same meaning as in Figure 2.1. . . . .	45
2.3	Thin-target non-thermal f-b:f-f ratio for $E_c = 5$ keV and $\delta = 7$ for 20, 30, 40 and 50 MK. Each incremental increase in the temperature produces a larger f-b:f-f ratio from the Fe edge at about 14 keV ( $= E_c + V_{Fe25+,26+}$ ) . . . . .	46
2.4	Spatially localised spectra from a loop with fully ionised plasma for two values of $E_c$ . The left plot shows a very distinct iron edge at $\approx 24$ keV ( $= E_c + V_{Fe26+}$ ) and a break at 15 keV ( $= E_c$ ), whereas the second plot shows an iron edge ( $\approx 28$ keV) and a break at $E_C = 21$ keV. This shows the value of recombination as an $E_c$ diagnostic. The ‘edges’ appear to be of finite slope because of the finite (1 keV) resolution used. . . . .	47
2.5	Non-thermal f-f and f-b spectra for the thick target case (Equations 2.29 and 2.30) shown for 2 different temperatures: 20 MK that is pertinent to events such as the 2002 April 14 event and 10 MK, which is more in the range of ‘microflare’ temperatures. The red curve represents the non-thermal f-f spectrum and the blue is for non-thermal f-b. The dashed black curve is the total non-thermal emission multiplied by 10. It is interesting to note the three distinct energy regimes for the f-b spectrum, namely: $\epsilon < V_{Fe}$ ; $V_{Fe} \leq \epsilon \leq V_{Fe} + E_c/n^2$ ; $\epsilon > V_{Fe} + E_c/n^2$ . Clearly f-b is very important in the 10-50 keV range, precisely where albedo issues are also important. . . . .	49

2.6	The thermal model spectra (blue: thermal f-f; red: thermal f-f) compared with Chianti's (green for f-f and black for f-b). Courtesy E. Landi for the Chianti spectra. . . . .	51
2.7	The spectral components for 2 different hypothetical situations. We show these spectra by varying the parameters around the results in the <a href="#">Veronig &amp; Brown (2004)</a> paper that analyses the coronal thick target 2002 April 14 event. In all cases we keep the same values of $\delta = 6.7$ , $E_c = 15$ keV and $T = 19.6$ MK. The curves represent the non-thermal f-f (green), non-thermal f-b (black), thermal f-f (blue) and thermal f-b (red) spectra. The magenta curve shows the total spectrum multiplied by 10. Plot A is for the thick-target coronal case with the actual event parameters $n_p, \mathcal{F}_{oc}$ according to Veronig and Brown. Plot B was obtained for the same event parameters but with $n_p$ reduced 25 times to make the loop collisionally thin above 10 keV and with footpoint emission occulted. The injection rate is the same as Plot A so the density fraction of fast electrons is 25 times higher. The non-thermal emission is down by 25 times while the thermal is down by a factor of 625. . . . .	55
2.8	Spectral components for a resolution of 0.1 keV. The f-b edges of all elements involved are clearly noticeable. The parameters are $T = 19.6$ MK, $E_c = 15$ keV and $\delta = 6.7$ . This plot can be compared to Plot A of Figure <a href="#">2.7</a> , which has the same parameters but for 1 keV binning resolution. . . . .	56
2.9	The thin-target spectra for 4 different temperatures with $f_c = 0.01$ . The colour-coding is the same as in Figure <a href="#">2.7</a> . . . . .	57
2.10	The revised thin-target spectra for 4 different temperatures with $f_c = 0.1$ . The colour-coding is the same as in Figure <a href="#">2.7</a> . . . . .	58
2.11	Fractional error ( $\Delta G/G$ ) in G (Equation <a href="#">2.36</a> ) as discussed in <a href="#">§2.6</a> for $E_* = 5$ (blue), 10 (green) and 20 (red) keV for a shifted power-law due to inference of G from H ignoring the presence of recombination. . . . .	60
2.12	The f-b electron efficiency compared to f-f for the 4 elements discussed in <a href="#">§2.7</a> (Equation <a href="#">2.41</a> ) assuming full ionisation. It is evident from the graph that, if present, highly ionised Fe is the most efficient source of f-b HXRs in terms of the $F(E)$ needed followed by Si, O and Mg. Note that the curves for each element begin at their ionisation potential since f-b emission below the ionisation potential is zero. . . . .	64

2.13	The $\Psi_{smooth}$ as discussed in §2.7 and Equation 2.41. It is the ratio of $j_R$ to $j_B$ for the smooth $F(E) \propto E(E + E_*)^{-\delta-1}$ for $E_* = 5$ (blue), 10 (green) and 20 (red) keV. . . . .	65
2.14	Inferred electron spectra using the photon flux as shown in Figure 2.8 with no noise included. The $x$ -axis is electron energy in keV and the $y$ -axis is electron flux. The black curve denotes the inferred spectrum if f-b emission is included whereas the green curve shows the inferred electron spectrum and the effects were f-b emission to be ignored. . . . .	66
3.1	The ratios $R_{NTR}(\epsilon, T)$ (blue) and $R_{Th}(\epsilon, T)$ (red) versus $\epsilon$ and $T$ for $\delta = 5$ and $f_c = 0.01$ . . . . .	70
3.2	Regime plot showing the relevant areas of importance in the $(\epsilon, T)$ domain for $E_c = 10$ keV, $\delta = 5$ and $f_c = 0.01$ . . . . .	71
3.3	Regime plot showing the relevant areas of importance in the $(\epsilon, T)$ domain for $E_c = 10$ keV, $\delta = 5$ and $f_c = 0.1$ . Note that the area of importance is larger as we go to bigger values of $f_c$ , as can be seen by comparing this plot to Figure 3.2 . . . . .	72
3.4	Regime plot showing the relevant areas of importance in the $(\epsilon, T)$ domain for $E_c = 10$ keV, $\delta = 7$ and $f_c = 1$ ; an extreme case. This shows that NTR may still be a dominant source of HXR in some cases. . . . .	73
3.5	The factor $\Phi(T)$ by which $EM_c$ is reduced when NTR and NTB are included compared to NTB alone for $\delta = 5$ . . . . .	75
4.1	Scattering angle in the ERS as given by Blumenthal & Gould (1970). Note that we use $\epsilon'_1$ and $\epsilon'_2$ instead of $\epsilon'$ and $\epsilon'_1$ for the incident and scattered photon energy respectively. . . . .	82
4.2	Schematic diagram showing the geometry used to describe the radiation field at the solar surface with the relevant angles and vectors. $\phi$ lies in the solar surface plane. . . . .	87
4.3	Top view of the Sun-Earth ICS geometry system. Note that the Earth is effectively at $\infty$ to the right and hence electrons/photons from any point on the Sun are parallel to the horizontal axis as viewed from Earth. As $\lambda$ , or solar longitude, increases from $0^\circ$ to $90^\circ$ , $\eta_{max}$ increases from $90^\circ$ to $180^\circ$ , which is the most energetic geometric configuration. . . . .	88

- 4.4 Photon spectra at the Sun (photons per keV per second per source electron) from fast electrons with a power-law energy distribution  $E^{-3}$ , where the thin-blue curve is the flux from disc centre ( $\sin \lambda = 0$ ), medium-red for  $\sin \lambda = 0.5$  and thick-black for the solar limb ( $\sin \lambda = 1$ ). Fluxes are normalised to one electron above 0.5 MeV and for an incident photon of energy 2 eV. . . . . 90
- 4.5 Photon spectra at the Sun (photons per keV per second per source electron) from fast electrons with a power-law energy distribution  $E^{-3}$ , where the thin-blue curve is the flux from disc centre ( $\sin \lambda = 0$ ), medium-red for  $\sin \lambda = 0.5$  and thick-black for the solar limb ( $\sin \lambda = 1$ ). Fluxes are normalised to one electron above 0.5 MeV and for an incident solar black-body photon spectrum. . . . . 91
- 4.6 Photon spectra at the Sun (photons per keV per second per source electron) from limb fast electrons with different power-law distributions: thin-blue is for  $\delta = 5$ , medium-red for  $\delta = 2$  and thick-black for  $\delta = 3$  with an incident photon energy of 2 eV. . . . . 92
- 4.7 Photon spectra at the Sun (photons per keV per second per source electron) from fast electrons with a power-law energy distribution  $E^{-3}$  for an incident photon energy of 200 eV. . . . . 93
- 4.8 Photon spectra at the Sun (photons per keV per second per proton) from relativistic positrons produced by protons with a power-law energy distribution  $E^{-3}$ , where the thin-blue curve is the flux from disc centre ( $\sin \lambda = 0$ ), medium-red for  $\sin \lambda = 0.5$  and thick-black for the solar limb flux ( $\sin \lambda = 1$ ). Fluxes are normalised to one proton above 1 MeV and for an incident photon energy of 2 eV. . . . . 94
- 4.9 Photon spectra at the Sun (photons per keV per second per proton) from limb relativistic positrons produced by protons with different power-law distributions: thin-blue is for  $\delta = 5$ , medium-black for  $\delta = 3$  and thick-red for  $\delta = 2$  with an incident photon energy of 2 eV. . . . . 95
- 4.10 Photon spectra at the Sun (photons per keV per second per proton) from relativistic positrons produced by protons with a power-law energy distribution  $E^{-3}$  for an incident photon energy of 200 eV. . . . . 96

4.11	Figure taken from <a href="#">Krucker et al. (2008b)</a> showing the imaging spectroscopy of the 2005 January 20 flare. The spectrum of the coronal source is given in red and that of the footpoint source is given in blue; their sum is shown in black with the spatially integrated high resolution spectrum shown in grey. As you can see, the spectral index of the coronal source is hard, $\approx 1.5$ . . . . .	97
4.12	TRACE image of the 2005 January 20 flare taken from <a href="#">Krucker et al. (2008b)</a> overlaid with RHESSI contours (50%, 70% and 90%) showing the regions of non-thermal emission. A strong HXR coronal source is clearly observed . . . . .	98
5.1	Picture of the first FNIT prototype with a plastic slab engraved with a grid of optic fibres. Courtesy UNH . . . . .	102
5.2	Schematic showing the principle of the first FNIT instrument with a stack of slabs. Courtesy UNH . . . . .	103
5.3	The new FNIT prototype with 3 rods, each filled with liquid plastic scintillator . . . . .	104
5.4	A schematic of the new FNIT prototype. Courtesy UNH . . . . .	105
5.5	The double n-p scatter principle for the FNIT prototype with the neutron source location known. Courtesy UNH . . . . .	106
5.6	The efficiency of BC-519 as established by GEANT4 simulations. Courtesy UNH . . . . .	106
5.7	Energy resolution of the FNIT scintillator bar was obtained from $^{137}\text{Cs}$ $\gamma$ -rays scattering in the detector at $90^\circ$ . The central energy is 374 keV and the resolution was found to be about 10%. Courtesy UNH . . . . .	109
5.8	Collimated $^{137}\text{Cs}$ photons located at 3.75 cm away from the centre of the bar (0 cm) on either side. Resolution was found to be 0.85 cm at these positions. Courtesy UNH . . . . .	110
5.9	Scatter plot of pulse height vs. pulse shape clearly shows the separation between $\gamma$ -ray-induced and neutron-induced events. Courtesy UNH . . . . .	111
5.10	Angular resolution measure of 2 MeV neutrons produced with beam data from CNL. The ARM was found to be $5.3^\circ$ at this energy. Courtesy UNH . . . . .	112
5.11	Event circles and a composite image of the 2 MeV neutron beam as compiled from measurements at a range of detector rotations . . . . .	113

5.12	Neutron source imaging using FNIT. This image was obtained by reconstructing the event circles (Figure 5.11) of each neutron and projecting them on to an image plane. The height of each bin in the histogram represents that number of circles that intersect at that location on the image plane. Courtesy UNH . . . . .	114
5.13	$^{252}\text{Cf}$ spectra: input theoretical distribution (black), uncorrected measured (red) and simulated (blue). Courtesy UNH . . . . .	118
5.14	Scatter plot representing FNIT response for all double scatter events. Courtesy UNH . . . . .	121
5.15	Scatter plot representing FNIT response with ToF cuts imposed. Courtesy UNH . . . . .	121
5.16	Scatter plot representing FNIT response for double scatters with ToF cuts and one elastic n-p scatter in each detector imposed. Courtesy UNH	122
5.17	Scatter plot representing FNIT response for double scatters with ToF cuts, one elastic n-p scatter in each detector and zero total n-C scatters imposed. Courtesy UNH . . . . .	123
5.18	Scatter plot representing FNIT response for ‘smoothed’ double scatters with ToF cuts imposed. Courtesy UNH . . . . .	124
6.1	Neutron emissivity of the solar flare of 1991 June 15 as a function of neutron energy, derived from COMPTEL data using a Maximum Entropy deconvolution of the instrument response, and compared with the earlier results of <a href="#">Kocharov et al. (1998)</a> , obtained using just the diagonal elements of the response matrix. From <a href="#">Toner et al. (2001)</a> . . . . .	128
6.2	SVD regularisation of the inverted solution using the UNH neutron telescope response matrix. The red curve is the original data generated on MATLAB and the blue curve is the reconstructed spectrum. A random Poisson noise of 10% was added to the generated data before the reconstruction. The $x$ -axis denotes the energy bin number of $E_{in}$ and the $y$ -axis is the data/energy value in arbitrary units. . . . .	133

- 6.3 SVD regularisation of the inverted solution using the FNIT response matrix as obtained from the ‘smoothed’ data (Figure 5.18) using linear binning. The red curve is the original data generated on MATLAB and the blue curve is the reconstructed spectrum. A random Poisson noise of 10% was added to the generated data before the reconstruction. The  $x$ -axis denotes the energy bin number of  $E_{in}$  and the  $y$ -axis is the data/energy value in arbitrary units. . . . . 134
- 6.4 SVD regularisation of the inverted solution using the FNIT response matrix as obtained from the ‘smoothed’ data (Figure 5.18) using logarithmic binning. The red curve is the original data generated on MATLAB and the blue curve is the reconstructed spectrum. A random Poisson noise of 10% was added to the generated data before the reconstruction. The  $x$ -axis denotes the energy bin number of  $E_{in}$  and the  $y$ -axis is the data/energy value in arbitrary units. . . . . 135
- 6.5 SVD regularisation of the inverted solution using the FNIT response matrix as obtained from the ToF cuts only data (Figure 5.15) using linear binning. The red curve is the original data generated on MATLAB and the blue curve is the reconstructed spectrum. A random Poisson noise of 10% was added to the generated data before the reconstruction. The  $x$ -axis denotes the energy bin number of  $E_{in}$  and the  $y$ -axis is the data/energy value in arbitrary units. . . . . 136
- 6.6 SVD regularisation of the inverted solution using the FNIT response matrix as obtained from the ToF cuts only data (Figure 5.15) using log binning. The red curve is the original data generated on MATLAB and the blue curve is the reconstructed spectrum. A random Poisson noise of 10% was added to the generated data before the reconstruction. The  $x$ -axis denotes the energy bin number of  $E_{in}$  and the  $y$ -axis is the data/energy value in arbitrary units. . . . . 137
- 6.7 GSVD Tikhonov regularisation of the inverted solution using the FNIT response matrix as obtained from the ToF cuts only data (Figure 5.15) using log binning. The red curve is the original data generated on MATLAB and the blue, magenta and black curves are the zeroth, first and second order reconstructed spectra respectively. A random Poisson noise of 10% was added to the generated data before the reconstruction. The  $x$ -axis denotes the energy bin number of  $E_{in}$  and the  $y$ -axis is the data/energy value in arbitrary units. . . . . 138

- 6.8 GSVD Tikhonov regularisation of the inverted solution using the FNIT response matrix as obtained from the ToF cuts only data (Figure 5.15) using log binning. The red curve is the smooth quadratic test data and the blue, magenta and black curves are the zeroth, first and second order reconstructed spectra respectively. A random Poisson noise of 10% was added to the generated data before the reconstruction. The  $x$ -axis denotes the energy bin number of  $E_{in}$  and the  $y$ -axis is the data/energy value in arbitrary units. . . . . 139
- 6.9 Blind Deconvolution of the inverted solution using the FNIT response matrix as obtained from the ToF cuts only data (Figure 5.15) using linear binning. The red curve is the original data generated on MATLAB and the blue curve is the reconstructed spectrum. A random Poisson noise of 10% was added to the generated data before the reconstruction. The  $x$ -axis denotes the energy bin number of  $E_{in}$  and the  $y$ -axis is the data/energy value in arbitrary units. . . . . 142
- 6.10 Blind Deconvolution of the inverted solution using the FNIT response matrix as obtained from the ToF cuts only data (Figure 5.15) using log binning. The red curve is the original data generated on MATLAB and the blue curve is the reconstructed spectrum. A random Poisson noise of 10% was added to the generated data before the reconstruction. The  $x$ -axis denotes the energy bin number of  $E_{in}$  and the  $y$ -axis is the data/energy value in arbitrary units. . . . . 143

# Preface

This thesis deals with two distinct yet connected aspects of solar physics. In the first half, we look at neglected mechanisms for the production of hard X-rays (HXRs) in solar flares. And in the second half, we delve into the world of neutron astronomy in the inner heliosphere. Studying energetic radiation processes as well as diagnosing the accelerated particles they produce underlie the basis of this thesis. Both halves of this thesis are integrated by this very principle: how can we better understand energetic processes in solar flares by analysing the radiation and accelerated particles they produce?

Chapter 1 reviews the necessary background relevant to this thesis. We introduce the most pertinent questions that interest solar astronomers these days. Various observations in all wavelengths and observations of energetic particles have been reviewed. We then discuss the corresponding theoretical models and interpretations to explain these observations. We also introduce neutron astronomy and the typical instruments required to measure solar neutrons.

In Chapter 2 we introduce the concept of non-thermal recombination (NTR) and conclude that it has wrongly been neglected as an important source of HXRs in solar flares. We provide the theory and use our model to make NTR HXR predictions. We also show that for typical flare parameters, NTR can be crucial and play an important role in electron cut-off diagnostics and have implications for spectral inversion inference.

We go further to show that NTR can be more efficient in producing HXRs than bremsstrahlung in Chapter 3. We look at various parameter regimes where NTR is a significant source of HXRs and conclude that in most solar flares, there is very possibly an energy and temperature domain where NTR is important, thereby requiring a possible rethink on the accelerated electron number problem as well as the flare energy budget problem, especially in extreme parameter regimes.

Like NTR, inverse Compton scattering (ICS) has been largely ignored as a significant source of HXRs in solar flares. However, we decided to revisit this problem in Chapter 4. Our purpose here was not to show that ICS is a dominant source of HXRs

for a large number of flares, but to be able to determine whether ICS in the solar context can be observed in some select flares. If so, it would give us an estimate of energetic particle distributions in solar flares and that would allow us to determine neutron fluxes likely to be encountered by future inner heliosphere space missions. We show that although ICS fluxes are likely to be low, there are those coronal HXR sources that can be interpreted using ICS quite viably.

Chapter 5 details our collaborative work with the Universities of New Hampshire and Bern to analyse the performance of a new-age neutron telescope that has been built to detect neutrons in the 1-20 MeV range. We describe the Fast Neutron Imaging Telescope (FNIT) instrument design and neutron detection technique as well as outlining the limitations and challenges that such an instrument entails. Using both laboratory and simulated data, we show how FNIT is an ideal candidate for a space mission into the inner heliosphere to detect solar neutrons.

In Chapter 6 we provide solutions to deal with data inversion from an instrument such as FNIT. We introduce and test several regularisation methods using the FNIT response matrix that we generated from simulated data. The performance of these deconvolution methods have been analysed and we conclude that zeroth order Tikhonov regularisation seems to be the most optimally suited for FNIT's needs. Undoubtedly, more work needs to be done to fine-tune the instrument as well as the data inversion strategies to be used, but our preliminary results are highly encouraging.

In conclusion, Chapter 7 summarises our main results in this thesis and provides the direction for future work to be carried out in the diverse range of topics elaborated on in this thesis.

# Acknowledgements

And so ends one of the most remarkable and enjoyable chapters of my life. The people who have made me feel this way would require several pages written in their honour, but I shall try and be as brief as possible to thank most of those people who have made my PhD possible, plausible and palpable.

Firstly, I would like to thank my supervisors John and Alec. Most people have a ‘first’ supervisor and a ‘second’ supervisor, but that was never the case for me. John and Alec have both contributed equally to my development not only as a research student and a scientist but as a person in general. Their infinite patience and meticulous attention to detail will never be forgotten and I shall cherish that for the rest of my life. In Indian culture, one’s teacher, or *guru*, is in likeness to the Divine. I am no religious person, but to me my gurus are my Gods. John and Alec have always been bubbling with ideas and every meeting that I have had with them has inspired me further in the field of science. I shall truly be indebted to them forever.

I would never have come to Glasgow or known about the people here had it not been for my undergraduate supervisor and my de-facto international PhD adviser, Jim. At UNH, Jim was virtually the first person I met and then worked in his lab for three years. During the summer of 2004, I got the opportunity to apply for an International Research Opportunities Programme scholarship, and Jim recommended I come to Glasgow to pursue our work on inversion techniques for solar neutron detectors. I jumped at the opportunity, successfully got an IROP scholarship and was able to undertake 2 months of research here. Since then, there was only ever going to be one place that I wanted to come to do my PhD, but it was never going to be that simple.

Hence here I would like to thank the University and the Science and Technology Facilities Council (then PPARC) for awarding me a Dorothy Hodgkin’s Postgraduate Fellowship. Without this financial support I would not have been able to do my PhD at Glasgow University and enjoy all the support to attend conferences and carry out field-work that was essential to my research. I consider the opportunity to travel the world and present my work at various venues as one of the enormous advantages of

having carried out my PhD in this Group and University. I would also like to thank UNH for their support on the field-trip I made with them to the Crocker Nuclear Lab in California.

I would like to thank many many more people within the Group as well. For often helping with a lot of the science, Eduard and Lyndsay were great. Similarly, excellent discussions and input have also come from colleagues at the University of California, Berkeley and UNH, especially those in Jim's group like Rich, John and Mark. From Strathclyde University, Nigel was of great help debugging some of the atomic physics in our work. Coming back to the A&A Group, for all matters related to LaTeXing or computing in general, Norman was the go-to man. Without him, none of all this would ever be in print-form! Hugh and Tobia were often of great help to deal with the wonder that is MATLAB! Daphne first and then Rachael have been brilliant in providing all the support necessary for making my life sane at most times! I would also like to thank Ross for inducting me into the Astronomy Outreach Team as that experience has made my education here all the more complete. Graham, thanks for putting me as a Demonstrator for A2 for three years; that was lovely. I enjoyed my time as Coffee Captain/King, whatever you want to call it! And all the exciting discussions I had at coffee with folk like Fiona, Iain, Declan, Craigs, Nicolas, Ed, Martin and everyone else in the Group I am forgetting to name! A special mention has to also be made for my wonderful office-mates, Robert and Russell. I really enjoyed their company and I hope I was not troubling or irritating to them. I would also like to thank Hazel, for she and I were true contemporaries! We could always share our PhD experiences as we have been in the same boat virtually throughout. Hamish has also been outstanding company and someone I feel most free with to engage in conversations of any kind. And for discussing all things from cricket to politics to culture, Luis was the man that made my daily existence here anything but mundane and something to look forward to!

Uttam, my friend for 20+ years and co-PhD student at Imperial College has been of invaluable support and company as trips to London or the Highlands have never been too far away with our other mate Hinal. I would also like to thank all my other friends, from Bangalore and UNH, who have taken the trouble of visiting me over the last three years and enjoying the Scottish experience. And my aunt in London, whom I can visit any time (literally!) I want.

Last but never ever the least and always the most, I would of course like to thank my family. Without my parents' and brother's unstinting support, counsel and love, I could never ever dream of accomplishing or experiencing what I have. I cannot even

imagine how difficult it must have been for my parents to first see my brother leave home and country as a teenager to pursue his studies and then see me follow suit shortly after. For the last 13 years, our lives have been lived on 3 continents with annual get-togethers and eternal farewells. Nothing could be more emotionally draining for any of us, but we have all become stronger from it and cherish the wonder that is Home and Family. I have also been lucky enough to have my uncles and aunts visit me here as well, and their good wishes have forever blessed me. My grandparents keep telling me that I make them proud, but I have only ever seen it the other way around. Since starting my PhD, I have also been extremely lucky to have welcomed my wonderful sister-in-law as the newest addition to our family. I guess that made it 4 PhDs out of 5, so I was never going to be far away in trying to convert that to a 100% record!

So thanks to one and all.

Let the thesis begin, and let a man who inspires me beyond any measure have the final say.

*“Live as if you were to die tomorrow. Learn as if you were to live forever.”*

– Mohandas Karamchand Gandhi

---

This thesis is my own composition except where indicated in the text. No part of this thesis has been submitted elsewhere for any other degree or qualification.

**Copyright © 2009 by Procheta C.V. Mallik**

1 October 2009

---

# List of Abbreviations

ARM	Angular Resolution Measure
BATSE	Burst and Transient Source Experiment
BBSO	Big Bear Solar Observatory
BD	Blind Deconvolution
CCD	Charged Coupled Device
CERN	European Organisation (Council) for Nuclear Research
CGRO	Compton Gamma Ray Observatory
CME	Coronal Mass Ejection
CNL	Crocker Nuclear Laboratory
COMPTEL	Compton Telescope
CTTM	Cold Thick Target Model
EGRET	Energetic Gamma Ray Experiment Telescope
EIT	EUV Imaging Telescope
ENDF	Evaluated Nuclear Data File
ERS	Eelectron Rest System
EUV	Extreme Ultra-violet
FGST	Fermi Gamma-ray Space Telescope
FIP	First Ionisation Potential
FNIT	Fast Neutron Imaging Telescope
GCR	Galactic Cosmic Ray
GLAST	Gamma-ray Large Area Space Telescope
GRB	Gamma Ray Burst
GRS	Gamma Ray Spectrometer
GSVD	Generalised Singular Value Decomposition
HXIS	Hard X-ray Imaging Spectrometer
HXM	Soft X-ray Monitor
HXR	Hard X-ray
HXRBS	Hard X-ray Burst Spectrometer

---

HXT	Hard X-ray Telescope
ICS	Inverse Compton Scattering
JPL	Jet Propulsion Laboratory
LS	Laboratory System
MDI	Michelson Doppler Imager
ME	Maximum Entropy
MHD	Magnetohydrodynamics
NASA	National Aeronautics and Space Administration
NM	Neutron Monitor
NoRH	Nobeyama Radioheliograph
NSO	National Solar Observatory
NTB	Non-thermal Bremsstrahlung
NTR	Non-thermal Recombination
OSO	Orbiting Solar Observatories
OSSE	Oriented Scintillation Spectrometer Experiment
PMT	Photomultiplier Tube
PNNL	Pacific Northwest National Laboratory
PSD	Pulse Shape Discrimination
RHESSI	Ramaty High Energy Solar Spectroscopic Imager
SCR	Solar Cosmic Ray
SMM	Solar Maximum Mission
SNM	Special Nuclear Material
SoHO	Solar and Heliospheric Observatory
SONG	Solar Neutron and Gamma-ray
SONTRAC	Solar Neutron Tracking Telescope
SVD	Singular Value Decomposition
SXR	Soft X-ray
SXT	Solar X-ray Telescope
TB	Thermal Bremsstrahlung
TGSVD	Truncated Generalised Singular Value Decomposition
ToF	Time of Flight
TR	Thermal Recombination
TRACE	Transition Region and Coronal Explorer
TSVD	Truncated Singular Value Decomposition
UNH	University of New Hampshire
UV	Ultraviolet

# Chapter 1

## Review of Energetic Processes in Solar Flares

The Sun is an enigmatic object. It is the reason that the Earth and the life it supports exist, but yet there is a lot about the Sun we do not understand. For solar astronomers, it is a laboratory par excellence. Its proximity to us, compared to other stars, allows us to observe it in detail and further the science of stars in general. Studying the Sun - an average main sequence star - gives us a good idea of phenomena that occur on other potentially life-giving stars in the Universe. Our search for extra solar planets and extra-terrestrial life involves looking for Sun-like stars and Earth-like planets. One of the biggest endeavours through human history has been to look at the heavens and try to understand far-away objects that we will never be able to reach but that continuously provide us with information. This is what makes astronomy unique among the sciences: it is primarily an observational science where we will forever be limited in scope by the distance of the objects we study. As a curious species, humans have endlessly strived to bridge this gap by analysing these observed data and using our theoretical knowledge to explain what we observe. The marvellous thing about this is that we know the objects we study are perfect distant laboratories that we will never be able to manipulate. Whatever information they give us is *always* natural. We just strive to be good enough to make sense of this nature!

We begin by reviewing some general features of the Sun, described in more detail in e.g. [Phillips \(1995\)](#). The Sun has a mass of around  $2 \times 10^{33}$  g and a radius of about  $7 \times 10^{10}$  cm. It is about 150 million km away from us and is a G2 star with 70% of its mass composed of hydrogen, 28% helium and the rest made up of heavier elements. However average a star the Sun might be, it still is a star. And that means an ever-changing ball of hot gas powered by nuclear fusion that has an environment at the extremes of our

---

imagination. High temperatures and strong dynamic magnetic fields mean that the Sun often produces the entire electromagnetic spectrum and accelerates some of the particles that make up its atmosphere by ejecting them into deep space as solar wind and cosmic rays. The core of the Sun, where nuclear fusion occurs, is at an astonishing temperature of 15 MK. The temperature and density reduce as we move through the following ‘layers’ of the Sun - the radiative zone (through which the energy produced in the core travels as electromagnetic radiation) and the convection zone (where the energy transport is by the means of convection) - till the temperature is about 5800 K at the photosphere. This is the surface we see with our naked eye or with optical telescopes and the temperature means we see it as a yellow-ish star. Quite remarkably, as we traverse the solar atmosphere above the photosphere, the temperature starts to rise again even as the density continues to drop. The chromosphere, a reddish layer about 2000 km thick lies above the photosphere and can be as hot as 20,000 K whereas the corona, even further up, is 1-5 MK. During solar flares, which will be reviewed in detail later, both the chromosphere and corona become much hotter: up to 10s of millions of degrees. During these flare events, not only does the solar atmosphere around the region of the flare become much hotter, but a lot of solar material, light, and other particles are ejected and emitted at high energies. Typical flares last for a few hours and an increased flux of emitted particles and light of all wavelengths is observed. The focus of this thesis will be to study some of these energetic particles and radiation to try and interpret what they tell us about the Sun, flares and the energetic processes that produce them.

One of the most enigmatic aspects of the Sun is how these energetic particles and photons are produced. With a broader understanding and improved technology, we aspire to understand these phenomena better and better, but that also means we observe more detail and require yet more complex analysis to explain the observations. Our understanding is no doubt improving, but some major questions remain unanswered and more new questions keep cropping up. The work presented here, in trying to probe some of these questions, has been like a drop in the ocean, but the ocean is made up of many drops. In this thesis, it is our aim to elucidate the work we have carried out over the last three years and put it down in words as effectively as possible. The general topic that ties the entire work together is probing particle acceleration at the Sun and our hope is that some of the aspects of this research are useful to present and future solar and stellar astronomers as well as scientists at large.

Our interest in energetic solar radiation has involved two aspects: hard X-ray (HXR) emission in solar flares and solar neutron astronomy in the inner heliosphere.

Both probe accelerated particle distributions at the Sun. And although both are mostly about flares, the inner heliosphere would also be the best place to look for quiet Sun neutrons. In the rest of this review chapter, we intend briefly to introduce these topics and then describe in detail our work on them in the subsequent chapters. Most of our work pertains to the physics of the solar corona and an easy and relatively comprehensive read on the relevant topics, some of which shall be briefly discussed in this introductory review, can be seen in more detail in the book by [Aschwanden \(2004\)](#).

## 1.1 The Observations: an Overview

Solar flares are produced when the continuously changing magnetic field structure at and above the solar surface experiences immense stress - due to the rotational and convective flow of the ionised solar gases - which then releases huge amounts of energy, up to  $10^{33}$  ergs, when these stressed field-lines re-orient in a different spatial arrangement. Regions on the Sun where this dynamic magnetic activity occurs are commonly called *active regions*. There are several models that describe what the energy release mechanism may be and how these field lines might re-orient themselves (e.g. [Sweet 1958](#); [Parker 1963](#); [Petschek 1964](#); [Brown & Priest 2001](#)). The largest of flares occur over an area of about  $4 \times 10^{19}$  cm<sup>2</sup> and their main body (soft X-ray loops) reaches heights of not much more than 10 Mm above the chromosphere, but the consequences are felt throughout the interplanetary medium in the form of phenomena such as the solar wind, magnetic storms and coronal mass ejections. A typical flare can be broadly divided into three stages: a precursor phase (which is sometimes not present), an impulsive phase and a gradual decay phase. Most of the energy is released in the impulsive phase (lasting tens to hundreds of seconds) and decay phase (lasting minutes to hours). This energy is shared about equally between producing heating, expelling masses of material and accelerating particles.

Since the atmosphere of the Sun is hot, most of the matter exists in an ionised plasma state. This means that individual charged particles are ‘freely’ zipping around the atmosphere interacting with other charged particles and the solar magnetic field to produce radiation. The sudden increase in radiation during flares is observed in all energy bands. A typical time-profile for a flare is given in [Figure 1.1 \(Kane 1974\)](#). Of course, the range, quantity and quality of these observations have been governed purely by technology. Flares were first observed in optical white-light emission by Carrington and Hodgson in 1859 ([Carrington 1859](#); [Hodgson 1859](#)) but are observed today in all wavelengths (radio, Extreme Ultra-violet (EUV), soft X-rays (SXR), HXR and  $\gamma$ -

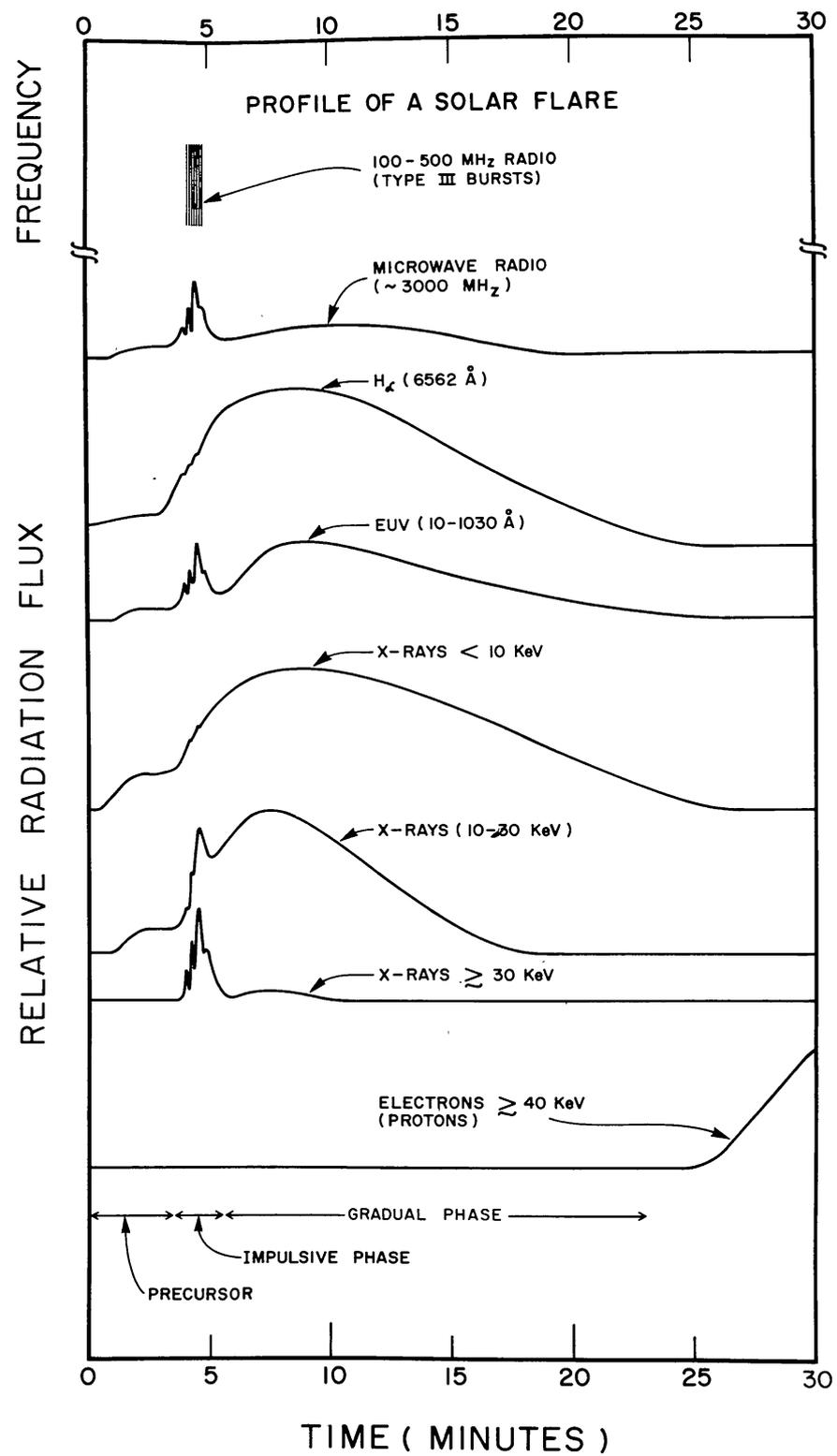
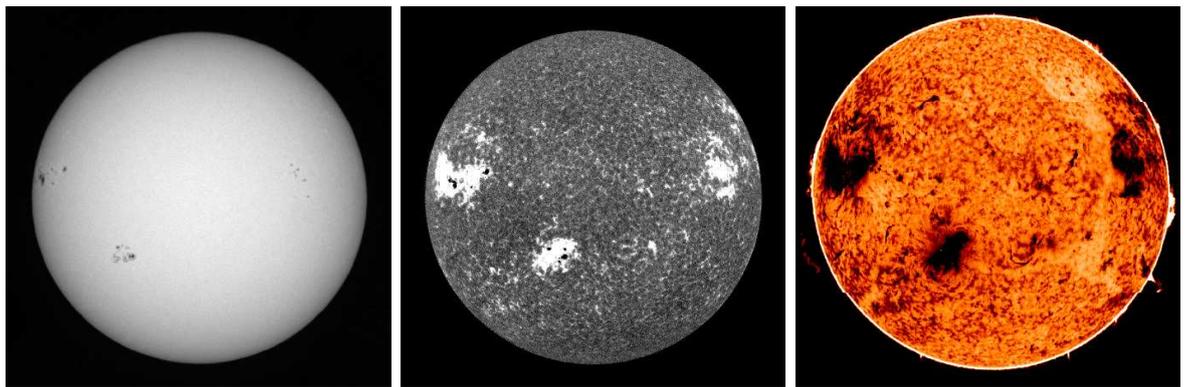


Figure 1.1: A schematic representation of the different phases of a solar flare (Kane 1974)

rays). We also detect particles, like neutrons and electrons, that have been accelerated in flares. In the last 150 years, we truly have come a long way and it can be safely said that it is studying flares in these non-optical energy bands as well as the energetic particles they produce that is most interesting and gives us the most diagnostic capabilities.

One of the biggest questions in solar physics is to know *how* the corona is heated to such high temperatures (millions of Kelvin) and even more perplexingly, *how* particles are accelerated to energies far above their thermal energy. Of course, there will always be some particles of arbitrarily high energies, but these very high energy particles are orders of magnitude more abundant in flares than expected just from Maxwell-Boltzmann statistics. But what we do know is that the corona *is* very hot and a lot of the radiation produced in flares *is* non-thermal. Electrons and ions both produce radiation when they interact with each other as well as when they interact with electric and magnetic fields. All of these interactions cause electrons and ions to accelerate - but since ions are so much heavier, the electromagnetic radiation they produce is negligible. However, when thermal and non-thermal electrons in solar flares interact with ions around them, they produce X-ray emission. The same electrons also produce microwave radiation when interacting with the surrounding magnetic field through gyrosynchrotron emission. However, the focus of our research has been to look at the mechanisms that produce the X-rays, in particular hard X-rays (HXR), which can be broadly defined as emission between 10-1000 keV. Our interest in this is motivated by the fact that the most direct way to trace accelerated non-thermal electrons is to study the HXR emission they produce. Nevertheless, since X-rays do not penetrate the Earth's atmosphere, solar X-rays have been observed best only since the space age with the advent of space telescopes and instruments. Although balloon-borne and rocket-borne instruments had detected solar X-rays as early as 1951 ([Friedman et al. 1951](#); [Chubb et al. 1957](#); [Kazachevskaya & Ivanov-Kholodnyi 1959](#)), it was only several decades later, with the launch of the Solar Maximum Mission (SMM), that we started seeing the Sun at a much higher resolution in X-rays and  $\gamma$ -rays.

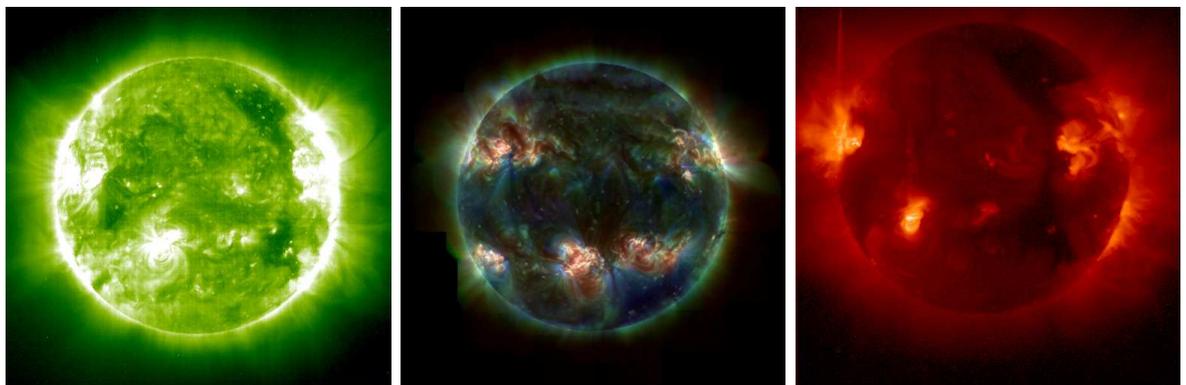
Modern solar spectroscopy and imaging have revolutionised the field of solar astronomy with stunning pictures of the Sun in radio, infra-red, visible, EUV, X-rays,  $\gamma$ -rays and even neutrons (Figure 1.2). Some of the most successful missions and instruments in the recent past include the Hard X-ray Imaging Spectrometer (HXIS) ([van Beek et al. 1980](#)) and the Hard X-ray Burst Spectrometer (HXRBS) ([Orwig et al. 1980](#)) aboard the SMM; the Hard/Soft X-ray Monitor (HXM) spectrometer and the Solar X-ray Telescope (SXT) onboard the Japanese Hinotori (Astro-A) satellite ([Tanaka 1983](#));



(a) Visible: White light  
BBSO

(b) Visible: Ca-K BBSO

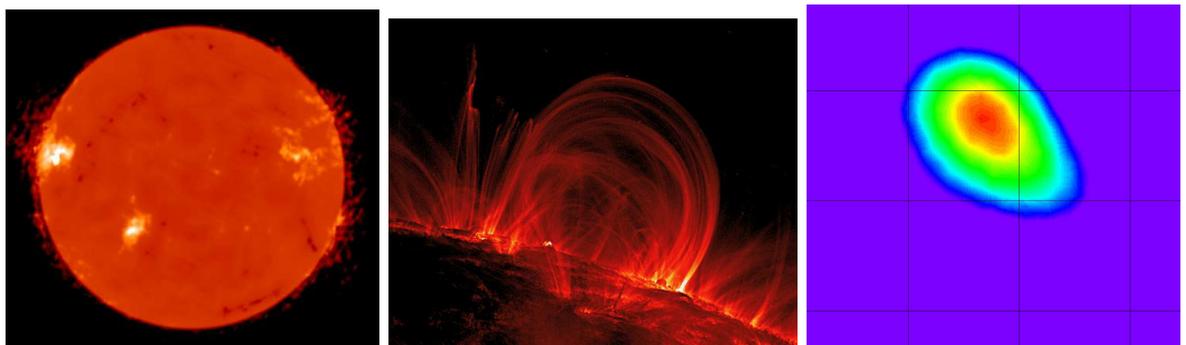
(c) Infrared: NSO



(d) UV: SoHO EIT

(e) EUV: TRACE full disc

(f) X-ray: Yohkoh



(g) Radio: NoRH

(h) EUV: TRACE

(i) Gamma Ray: Neutrons  
COMPTEL

Figure 1.2: The Sun as seen in various wavelengths by various modern instruments. Courtesy NASA/JPL-Caltech

the Compton Telescope (COMPTEL) (Schoenfelder et al. 1993) aboard the Compton Gamma Ray Observatory (CGRO); the Hard X-ray Telescope (HXT) (Kosugi et al. 1991) onboard the Yohkoh (Solar-A) satellite; the EUV Imaging Telescope (EIT) (Delaboudinière et al. 1995) aboard the Solar and Heliospheric Observatory (SoHO); the Transition Region and Coronal Explorer (TRACE) (Strong et al. 1994; Handy et al. 1999); the Ramaty High Energy Solar Spectroscopic Imager (RHESSI) (Lin et al. 2002); and Hinode (Solar-B) (Kosugi et al. 2007). Each of these missions has broken new ground and shown us the fine structure of magnetic fields and flares, ribbons and prominences, filaments and jets with ever improving spatial and temporal resolution. Some of the typical spectra and images observed by these instruments, as well as modern ground-based telescopes like the National Solar Observatory (NSO), Big Bear Solar Observatory (BBSO) and Nobeyama Radioheliograph (NoRH) (Nakajima et al. 1994; Nishio et al. 1994), are shown in Figures 1.2 and 1.4.

The impulsive phase of the flare is when all the fun really starts! This is observed as sudden bursts of emission in all energies. Most of the energy in this phase is widely thought to be released as fast electrons, although various theories to the contrary also exist, (e.g. Simnett 1995), who argues that quite often most of the energy goes in accelerating protons. These accelerated particles travel along the magnetic field lines down from the acceleration site, which is deduced from X-ray and radio observations, to the lower more dense regions of the atmosphere where they produce emission in other wavelengths like EUV and visible light. The most energetic flares also produce  $\gamma$ -ray emission from which one can deduce the energy in the accelerated ions. An excellent complement to  $\gamma$ -ray observations is to study neutrons, which are emitted in nuclear reactions involving these accelerated particles. Majority of the neutrons emitted in solar nuclear reactions are in the energy range of 1-20 MeV. Studying neutrons gives us information on ions in the 100s of MeV range. Being neutral, like light, neutrons are unaffected by ambient electric and magnetic fields although they may scatter off other particles before they are finally detected. Unlike light, of course, neutron  $v \neq c$ ! However, since heavier ions do not produce narrow  $\gamma$ -ray lines, studying neutrons produced by collisions involving these heavier ions is considered a more accurate method to probe heavier ion energy distributions.

Some of the energy of the flare also goes into heating electrons in the corona. This energy is quite often considerably less than the energy used to accelerate the electrons (Saint-Hilaire & Benz 2002). In the same paper, the authors also show that the thermal and kinetic energy in the EUV-observed jet is less than the energy imparted to accelerating the electrons. Some flares are also associated with CMEs, or

Coronal Mass Ejections, where coronal plasma of  $10^{15}$  g or more is ejected from the Sun into space at velocities greater than  $1000 \text{ km s}^{-1}$ . The energy carried by CMEs for a typical flare is about  $10^{32}$  ergs. CMEs are of particular interest in solar physics, not least because they are expansive events whose effects can be felt throughout the interplanetary medium.

### 1.1.1 Radio and microwave observations

As is sadly quite often the case, scientific discovery and technology is often driven by war. A lot of rapid progress in science has been made during war-time and some of the motivation for funding science even today has an underlying military aspect to it. In any case, by a quirk of fate, solar radio emission was first detected in 1942 by J.S. Hey working for the British Army Operational Research Group. While studying the jamming of army radar, he observed intense radio emission and concluded it must be from the Sun (Hey et al. 1948). Solar radio emission is often observed as these intense metric-wavelength ‘bursts’, divided into five different types which shall be briefly introduced here. At higher energies, microwaves are produced through cyclotron and gyrosynchrotron emission by thermal and non-thermal electrons as they accelerate due to the prevalent magnetic field. Excellent reviews on solar radio emission include Dulk (1985) and Bastian et al. (1998). A schematic diagram taken from Dulk (1985) is shown in Figure 1.3.

Type I bursts are not associated with flares and are long-lasting (several hours to days). They are thought to be related to a continuous process like plasma oscillations created by accelerated electrons. Type II bursts are related to large flares, most likely produced by a shock-wave that propagates from a flare (e.g. Wild 1950; Nelson & Melrose 1985). Type III bursts occur over a short time period but over a large frequency range. This is because they are thought to arise from relativistic electrons moving along magnetic field lines away from the corona into space and exciting plasma emission at the plasma frequency in the form of Langmuir waves. As the density of the plasma reduces, so does the frequency of the emitted radiowave since the electrons move into more tenuous reaches of space. They typically start in the corona at frequencies of the order of 100 MHz and quite often appear as harmonic bands differing in frequency by a factor of  $\sim 2$  (e.g. Suzuki & Dulk 1985). It is now also accepted that Type III bursts are associated with electron beams (Lin et al. 1981, 1986). Type IV bursts are caused either by gyrosynchrotron emission or plasma emission from non-thermal particles and are shock-associated. This emission is of a broadband nature. Moving Type IV bursts

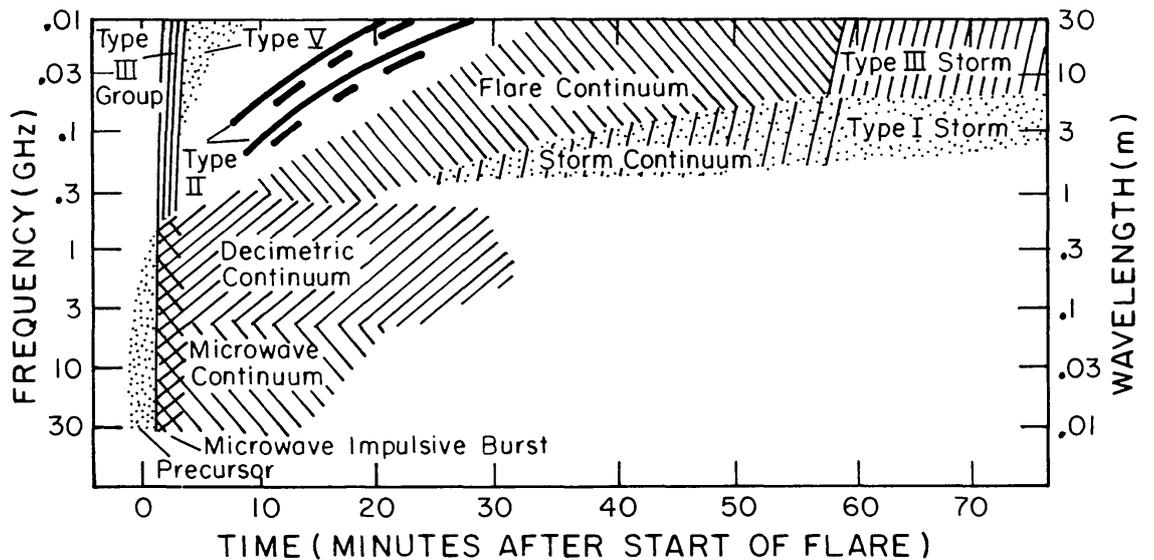


Figure 1.3: Schematic diagram of the solar radio spectrum from a “typical” large flare (Dulk 1985). This simplified diagram can be quite different from observed flare radio spectra.

can be observed from an expanding loop or a CME. Langmuir waves produced by Type III bursts produce Type V bursts by forward or counter-streaming, which is why Type V bursts are observed shortly after Type III bursts, but are broadband and consistent with electrons extending over a range of heights.

### 1.1.2 EUV Observations

Flare and coronal EUV outputs are from thermal emission of spectral lines. The EIT and TRACE instruments have produced some of the best pictures of the Sun in EUV. EIT produces full-disc images of the Sun whereas TRACE has a much higher spatial resolution and takes images in blocks of  $8.5 \times 8.5$  arc minutes. The TRACE images in Figure 1.2 show one of these block images as well as an image where several of these blocks have been collated to produce a full-disc representation of the Sun. EIT shows spectral line images of He II ( $304\text{\AA}$ ), Fe XV ( $284\text{\AA}$ ), Fe XII ( $195\text{\AA}$ ) and Fe IX ( $171\text{\AA}$ ), thereby showing plasma temperatures from 0.08 MK to 2 MK, whereas TRACE also shows lines at  $1700\text{\AA}$ ,  $1550\text{\AA}$  and  $1216\text{\AA}$  as well as white-light images. Fast electrons propagate down magnetic loops from the acceleration site in the corona to the denser chromosphere, where they heat chromospheric material. This material then ‘evaporates’ back into the coronal loops and glows as EUV emission. Hence this EUV emission can be thought of as a response of the lower atmosphere to the fast

electrons, which produced HXR bursts at the beginning of the flare. The loops appear to rise through the duration of a flare through the corona before radiatively cooling down after the flare and eventually draining under the force of gravity. These post-flare loops reveal the magnetic structure in the corona after a flare. Reviews of observations in EUV include those by e.g. [Bray et al. \(1991\)](#); [Benz \(2008\)](#).

### 1.1.3 X-ray observations

Reviews of interpreting HXR observations have been given by various authors over the last forty years, starting with the seminal papers of [Sturrock & Coppi \(1966\)](#), [Sweet \(1969, 1971\)](#), [Kane \(1974\)](#) [Brown \(1975\)](#) and [Kane et al. \(1980\)](#) to the more recent reviews such as those by [Krucker et al. \(2008a\)](#). Although the observations have improved dramatically over that time period, the basic interpretation of the data remains intact. In essence, microwave, EUV, X-ray and  $\gamma$ -ray observations have complemented each other and reveal various aspects of solar flares. Microwave radio bursts seem to follow similar time-profiles to HXR bursts; EUV radiation, also reflecting the same time profiles, is often observed in narrow emission line bands that give a wonderful temperature diagnostic for the solar atmosphere;  $\gamma$ -ray observations sometimes mirror similar time profiles, in which cases the ions needed to produce the  $\gamma$ -rays must be accelerated at about the same time as the electrons needed to produce HXR bursts (e.g. [Chupp et al. 1973](#)) (arguments for and against this idea are discussed in §1.1.5). Accelerated electrons produce HXR bursts by interacting with ambient ions (the mechanisms will be described in the Theory section following this) and microwaves through synchrotron emission by gyrating along the prevailing magnetic field lines.

The solar flare X-ray spectrum (Figure 1.5), from instruments such as RHESSI (Figure 1.4), shows a distinct thermal component ( $< \sim 10$  keV, depending on the plasma temperature and density) and a non-thermal component above  $\sim 10$  keV. The thermal component is characterised by the form  $I(\epsilon) \propto e^{-\epsilon/kT}$  produced by ‘thermal’ or a Maxwellian distribution of electrons. The non-thermal component seems to be roughly that of a power-law of the form  $I(\epsilon) \propto \epsilon^{-\gamma}$  produced by a ‘non-thermal’ or roughly power-law distribution of electrons. The value of  $\gamma$  usually ranges between 2 (hard) and 6 (soft) or so. That is, when the relative proportion of more energetic photons is larger, we call that a ‘harder’ or flatter spectrum. Whereas a steep spectrum corresponds to a relatively larger proportion of less energetic photons, i.e. a ‘soft’ spectrum. The non-thermal electron distribution is typically considered to have a low- and high-energy cut-off or roll-over energy. Various forms of these cut-off energies

have been postulated, some physical and some not-so physical but which are easier to interpret mathematically. The basic premise is that the distribution of electrons changes or goes to zero at these cut-off energies. The question and concept of non-thermal electron cut-off energies is one of the most eagerly debated topics in solar X-ray physics. As a convention, the thermal-dominated component is thought of as the SXR spectrum and the non-thermal-dominated component as the HXR spectrum. When studying a flare time profile in HXRs, the *typical* flare seems to follow a soft-hard-soft (Kane & Anderson 1970) or a simply soft-hard (Hoyng et al. 1975) development, reflecting the transition between the phases of the flare. This makes sense since one would expect the spectrum to harden during the most energetic phase - the impulsive phase - of the flare when most of the energy is released in a short duration. The fact that you end with a hard spectrum or some softening is also logical since a large amount of energy is released in the gradual decay phase of the flare as well. Also, low energy particles collisionally lose their energy first and so too contribute to the hardening of the spectrum.

The mechanisms for producing X-rays (1-1000 keV) will be explained in §1.2.1, but the essential point to make here is that to produce X-rays by free-free or free-bound emission requires the electrons to be of a similar energy to the radiation they produce, namely in the 1-1000 keV range. This is of course not true for other radiation mechanisms, like inverse Compton (which requires relativistic electrons to produce X-rays) or gyrosynchrotron (where even relativistic electrons produce only radio emission) - see Chapter 4 for details. Large emission measures and temperatures in the region of 10 MK in the corona during flares mean that thermal electrons dominate the production of X-rays up to about 10-20 keV whereas it is the non-thermal electrons that produce the harder X-rays and even continuum  $\gamma$ -rays from relativistic electrons. The X-ray continuum is also embellished with line and edge features at lower energies. Line features are due to spectral emission lines of highly ionised species of elements such as Fe or Ni and occur at SXR energies whereas edges may be produced due to radiative recombination of electrons on to the same highly ionised species and can even occur at higher HXR energies as will be explained later in the thesis. Throughout the history of studying X-ray spectra, it has been assumed that free-free bremsstrahlung emission has been the predominant source of X-ray production, but since this thesis seriously challenges that assumption, we shall refrain from making the same statements.

RHESSI is one of the best instruments to study X-ray emission because of its excellent spatial resolution as well as its dynamic range, allowing it to study radiation from the SXR range to  $\gamma$ -rays (3 keV - 18 MeV). However, several instruments, as mentioned

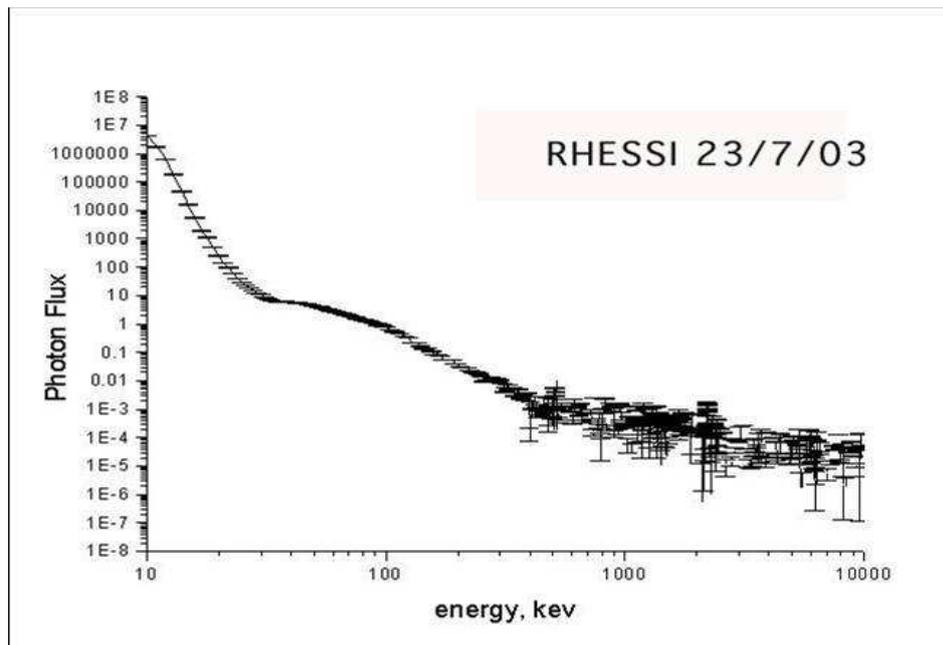


Figure 1.4: Observed RHESSI flux spectrum showing the distinct thermal and non-thermal components as well as the  $\gamma$ -ray line and continuum emission extending to higher energies. Courtesy E. Kontar

before, have provided astronomers with excellent X-ray data. The HXRBS observed between 20-300 keV, HXIS 3.5-30 keV, the Burst and Transient Source Experiment (BATSE) (Fishman et al. 1985) aboard the CGRO had a range of 25-300 keV and the Hard X-ray Telescope (HXT) aboard Yohkoh observed in several HXR energy bands down from 15 keV to about 100 keV. Also, the Yohkoh HXT instrument and HXIS were imagers, whereas the BATSE detectors performed high-resolution spectroscopy.

Like EUV radiation, SXR's have been observed by SXT in the loops of flares, which continue to emit for several hours after the flare. Since it is hard for highly ionised plasma material to move across strong magnetic field lines, SXR emitting regions are confined to flare and post-flare loops. This provides us with excellent diagnostics for flare density, size and temperature. HXR's are usually observed in the 'footpoints' of flares as well as in some cases in loop-top coronal sources. Electrons are thought to be accelerated in the high corona (which may be why we sometimes see HXR's in the corona) and move down the field lines into the dense chromosphere, losing energy throughout this journey, and producing HXR's in the chromospheric footpoints. A comprehensive review of flare observations from radio-waves to X-rays, explaining all the above phenomena, is given by Benz (2008).

### 1.1.4 $\gamma$ -ray observations

Some of the more energetic flares produce  $\gamma$ -rays (radiation above 1 MeV). This is because particles are accelerated to relativistic energies and so produce more energetic radiation. The continuum  $\gamma$ -ray flux can be attributed to relativistic electrons, basically a continuation of the HXR continuum. However, several line features are also observed at  $\gamma$ -ray energies and these provide excellent diagnostics for accelerated ions in flares. Nuclei bombarded by protons and  $\alpha$ -particles produce de-excitation lines as energy is released as a  $\gamma$ -ray photon. The positron annihilation and neutron capture line, which often seem to be the most prominent  $\gamma$ -ray lines, are also present at 511 keV and 2.223 MeV respectively.

Using the OSO-7 satellite (Clark 1973),  $\gamma$ -ray emission from solar flares was first detected in 1972 (Chupp et al. 1973), where lines at 0.511, 2.223, 4.4 and 6.1 MeV were detected as well as continuum emission in the 0.35-8 MeV range. Subsequent observations of solar  $\gamma$ -rays include those of Chambon et al. (1978); Hudson et al. (1980) and from instruments such as the Gamma Ray Spectrometer (GRS) (Forrest et al. 1980) on SMM in the 10 keV - 160 MeV range and the Oriented Scintillation Spectrometer Experiment (OSSE) (Johnson et al. 1993) (50 keV - 10 MeV) and COMPTEL (0.75-30 MeV) onboard CGRO.

The main importance of  $\gamma$ -ray observations is that they give us a number for energetic nucleons, mainly protons, in flares. The fact that  $\gamma$ -rays seem to be produced nearly at the same time as HXRs (within 3 minutes in Chupp et al. (1973)) implies impulsive phase acceleration of  $\sim 10^{33}$  particles to energies of  $> 30$  MeV, deduced from the 2.223 MeV neutron capture line. The 511 keV positron annihilation line may give us more information on the density and temperature in the emitting region (Murphy et al. 1997). Although the widely held belief is that observations indicate the energy in the energetic protons is much less than that in accelerated electrons, this is not necessarily the case (e.g. Simnett 1995; Vilmer & MacKinnon 2003). Hence emission signatures of ions as well as those by electrons, viz. HXRs, should be used as diagnostic tools to determine the primary mechanism of energy release in flares.

### 1.1.5 Observations of energetic particles

So far we have talked about observing photons (i.e. electromagnetic radiation) of various energies. But we know from these observations that they are produced by particles that are accelerated in a flare. Direct observations of accelerated particles have also been made, but observing particles has even more constraints than observing light.

For one, photons have no rest-mass and are neutrally charged and so detecting them is generally simpler as they are unperturbed by the strong magnetic and electric fields that they encounter on their way towards the Earth. On the other hand, electrons, protons and other heavy nuclei are charged particles and have non-zero rest-mass thereby being a completely different kettle of fish when it comes to detecting them. Neutrons, nevertheless, are massive but are neutral and so end up being a good complement to  $\gamma$ -ray observations. Since a sizeable portion of this thesis concerns doing neutron astronomy, more details of this will be given in this and later chapters. However, in this section, we shall briefly talk about observations of other energetic particles as well as give some background on neutrons.

As has been alluded to, only a small fraction of charged particles accelerated in flares is likely to escape the magnetic field near the Sun and actually reach the Earth to be detected. However, some charged particles do escape on open field lines. Interpretation of how they were accelerated in the flare is clouded by the fact that they would have interacted with the plasma and magnetic field before being detected by us. Hence their original information is lost, but we can still study and interpret them. Properties of observed relativistic and non-relativistic electrons have been reviewed extensively by [Lin \(1974, 1998\)](#) and [Simnett \(1974, 2001\)](#) respectively. The distribution of non-relativistic electrons (5-100 keV) tends to be a power-law, with index between 2 and 5, before steepening considerably above 100 keV for flares where electrons are the only charged particles that are detected. However, there are flares where protons are also detected and these flares have the electron distribution extending well into the relativistic regime. This is to be expected because the more energetic flares would accelerate heavier particles as well as accelerate these particles to higher and even relativistic energies. However, the fact that these relativistic particles tend to arrive at about the same time if not *after* non-relativistic electrons also points to the fact that there may not only be a difference in the acceleration mechanism that produces them (this has been refuted as well - see below) but also that the mean-free path of non-relativistic electrons through the interplanetary medium is  $\gtrsim 1$  AU as they travel ‘scatter-free’ along the magnetic field lines ([Lin 1974](#)) whereas relativistic electrons and protons propagate diffusively, thereby producing radio Type II bursts, often associated with these more energetic ‘proton’ flares ([Švestka & Fritzová-Švestková 1974](#); [Cane et al. 2002](#)). This variability in the behaviour of non-relativistic electrons can be attributed to the spectrum of the magnetic irregularities at those wavelengths that correspond to the gyroradii of these electrons ([Lin 1974](#)). Other explanations such as the difference in location of acceleration sites may also be considered. About one in

every forty flares has any sort of non-relativistic electrons detected; proton flares are a further six times rarer. For more discussion on observations of electrons, protons and heavier nuclei, please refer to the reviews mentioned above as well as e.g. [Ramaty et al. \(1980\)](#); [Cliver & Ling \(2007\)](#).

Neutrons are created by nuclear reactions of high energy protons ([Biermann, Haxel, & Schlüter 1951](#); [Lingenfelter 1969](#); [Ryan, Lockwood, & Debrunner 2000](#)). Further evidence of acceleration of charged particles with specific reference to the production of neutrons and  $\gamma$ -rays has been reviewed by [Chupp \(1971, 1984\)](#); [Hudson & Ryan \(1995\)](#); [Reames \(1999\)](#). And neutrons emitted by a solar flare were first detected by the GRS on SMM in June 1980 ([Chupp et al. 1982](#)). This breakthrough and subsequent research on energetic processes in solar flares has probed deeper into the acceleration processes that govern flares. Early  $\gamma$ -ray observations, such as those by [Forrest & Chupp \(1983\)](#), claim to show that electrons and ions are accelerated almost simultaneously and can be explained by a single process and challenged the widely accepted 2-phase acceleration theory alluded to earlier. Recently, solar  $\gamma$ -rays have been detected by a variety of instruments (e.g. [Hurford et al. 2003](#); [Vilmer et al. 2003](#)). Since  $\gamma$ -ray detectors tend to detect neutrons and vice-versa, the instruments - and not only the science - pertaining to  $\gamma$ -rays and neutrons is often complementary and supplementary. Instruments such as GRS and COMPTEL were mainly  $\gamma$ -ray detectors, but their success has paved the way for neutron astronomy and modern detectors (see Chapter 5), where neutrons are the primary particle of interest. Neutrons are produced in the same reactions that produce nuclear  $\gamma$ -rays, like when protons and  $\alpha$ -particles interact with heavier nuclei, e.g. C, N, O, and produce  $\gamma$ -ray lines. Hence any energetic flare that produces  $\gamma$ -rays also produces free neutrons. The main differences in their production reactions involve threshold energies. For  $\gamma$ -production of de-excitation lines, the nucleus of the target species has to be excited and so the threshold is not particularly high since no new particle is being expelled in the reaction. However, to produce free neutrons via pion production in a proton-proton interaction, for example, the threshold energy is much higher, about 300 MeV, since you have the pion carry away the positive charge of the proton. However, heavier nuclei produce broader  $\gamma$ -ray lines and so studying neutrons becomes even more important because they end up being better diagnostic tools for understanding the acceleration processes of these heavier nuclei. The main problem with neutrons is that free neutrons have a lifetime of just under 15 minutes. This means that only the most energetic and highly relativistic neutrons, in the range of 100s of MeV or more, can be detected on Earth or in a near-Earth orbit. Hence only a handful of flares in the last several decades have produced neutrons energetic enough to

be detected and studied by us. Apart from near-Earth orbit space-based instruments such as GRS and COMPTEL, neutrons have also been detected on Earth by neutron monitors like those at Jungfraujoch in Switzerland, Lomnitsky Stit in Slovakia, Mt. Washington in USA and many more located at similar high elevations in Tibet, Russia and so on. An observation made by one such Earth-based neutron monitors is given by e.g. [Muraki et al. \(1992\)](#). Solar neutron observations in general have also been reviewed extensively by e.g. [Ryan et al. \(2000\)](#). Solar neutron astronomy has been discussed in greater detail later in this review chapter (§1.3).

Increasing the area of an Earth-based neutron detector is not too beneficial either since neutrons below 150 MeV rarely reach the Earth and those below 500 MeV get scattered so much in the atmosphere that they lose any useful information before they can be detected. Hence today, there is great interest and a push for studying neutrons in the 1-200 MeV range with new-age neutron telescopes that are being developed for space flights. Since most neutrons will be in this energy range, the most we can learn from neutron astronomy will be when we study these neutrons at 0.2 or 0.3 AU orbits around the Sun. It is hoped that in the near future, such instruments (one of which will be introduced in Chapter 5) will be put into orbit on inner heliospheric space missions thereby expanding our knowledge and understanding of solar flares and the Sun in general. For excellent and relatively recent reviews of solar energetic particles, see [Hudson & Ryan \(1995\)](#); [Ryan et al. \(2000\)](#).

## 1.2 Review of Relevant Theory

In this section, we shall briefly talk about the theory pertinent to this thesis. Since a substantial part of the thesis is about Hard X-ray emission mechanisms (Chapters 2, 3 and 4), in the following subsection I will introduce the relevant theories, which will be elaborated on in the aforementioned chapters. In the subsection on particle acceleration, the emphasis will be slightly different since this thesis deals with diagnosing accelerated particles and not in particular on the possible processes that accelerate and/or heat these particles. Hence, this subsection will give a brief overview of the processes that have been suggested to contribute to this acceleration process.

### 1.2.1 X-ray emission

Theoretically, there are four ways of producing HXR emission: free-free bremsstrahlung emission, free-bound recombination emission, gyrosynchrotron (or magnetobremstrahlung);

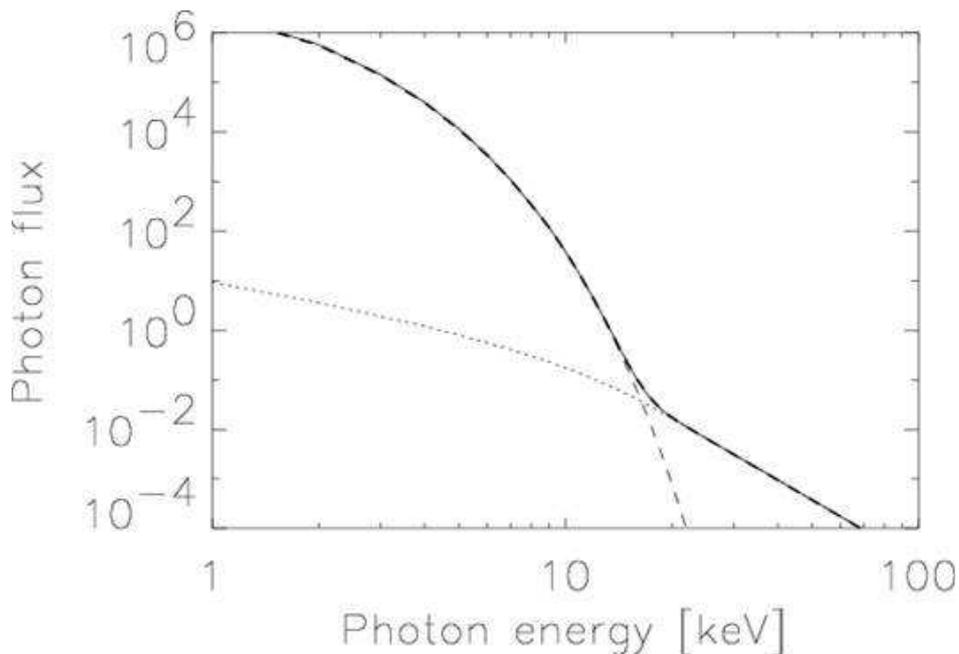


Figure 1.5: Typical solar X-ray flux spectrum showing the distinct Maxwellian thermal component and the non-thermal power-law tail

in this section we shall just refer to this process by the former name) emission and inverse Compton scattered emission.

Among these four methods, bremsstrahlung has been considered the primary source of the HXR's we observe from solar flares (cf. [Korchak 1967, 1971](#)). However, since half this thesis talks about how recombination emission and inverse Compton scattering can also be significant sources of HXR's for certain conditions prevalent in solar flares, we shall not make the above assumption and here will instead just elucidate what each process means.

Bremsstrahlung is German for “braking radiation” (Figure 1.6). In the context of the Sun, the particles that are “braking” or decelerating and producing this emission are electrons. The primary particles responsible for this deceleration are other electrons, which slow the electrons collisionally; and protons - which are also abundant in the solar atmosphere - off which the electrons scatter and produce the free-free bremsstrahlung radiation. Hence whenever electrons interact with protons, or other charged particles for that matter, their energy changes and photons in a range of energies are released. In the context of bremsstrahlung emission, it is the e-p interaction that is the primary source of the radiation. The photon energy depends primarily on the initial energy of the electron and obviously the degree to which it is decelerated by the proton. We know this to be dependent on parameters such as the collisional cross-section, distance

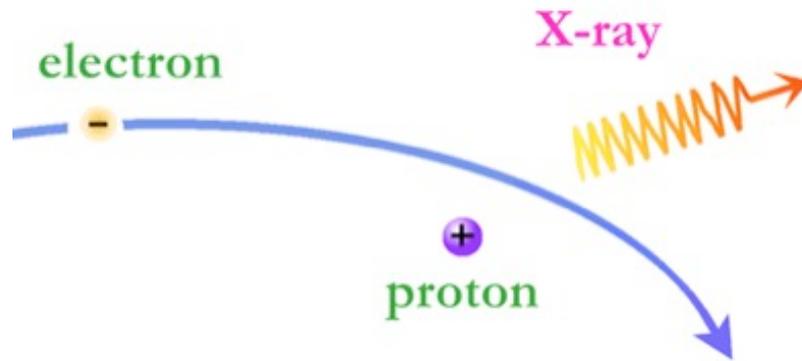


Figure 1.6: Schematic diagram showing bremsstrahlung emission. Courtesy S. Bank, University of Wisconsin-Madison

of interaction, plasma density and so on. There is a spread of photon energies, but the energy of the emitted photon can never be greater than that of the incident electron and so an electron of a particular energy can produce photons of all energies less than the electron energy. Subsequently, when we look at a distribution of electrons, they will always produce a distribution of photons that is an integral of the electron energies. As mentioned earlier in this chapter, it is by observing this photon spectrum (e.g. Figure 1.5) that we deduce an electron spectrum. The photon spectra also tell us that there seem to be two distinct populations of electrons: thermal and non-thermal. Thermal electrons reflect the temperature of the plasma and hence have the form of a Maxwellian distribution ( $\propto e^{-\lambda E/T}$ ) whereas non-thermal electrons seem to have a power-law distribution ( $\propto E^{-\delta}$ ), completely independent of temperature, hence “non”-thermal. To deduce the photon distribution produced by an electron distribution, one has to essentially integrate the electron distribution over all energies (and of course include the relevant parameters such as the cross-section, density etc.) and normalise the result. Detailed equations are given in Chapter 2, so we shall not repeat them here.

Exactly the same electrons that produce bremsstrahlung emission may also produce free-bound recombination emission (Chapters 2 and 3). In a hot plasma (several million K, which is typical of flares), other heavy nuclei besides H also exist in an ionised state, and so electrons can interact with these nuclei as well. In particular, electrons can recombine with any of these nuclei and thereby produce emission. In practice, the electron can recombine with any positively charged ion. But as will be made clear in the body of this thesis, it is recombining with the heavier nuclei, especially iron, that produces the most significant free-bound HXR emission. The main differences between bremsstrahlung and recombination are that in the latter the energy of the



Figure 1.7: Schematic diagram showing synchrotron emission. Courtesy S. Bank, University of Wisconsin-Madison

emitted photon is unique for a given electron energy and is *greater* than that of the incident electron because of the ionisation potential of the ionic level into which the electron recombines. That is, the energy of the emitted photon is equal to the energy of the incident electron *plus* the ionisation potential of the relevant ion. This produces edges in the spectrum not seen in the smooth bremsstrahlung spectra and of course, the edges are largest and most noticeable for the heaviest nuclei present, e.g. iron, nickel and sometimes even oxygen. These results are explained and discussed in detail in Chapters 2 (Brown & Mallik 2008) and 3 (Brown & Mallik 2009).

Charged particles can also interact with the prevalent magnetic field in a solar flare to produce gyrosynchrotron emission (Figure 1.7). As the name suggests, this is when ions or electrons of relativistic energies gyrate along magnetic field lines due to the Lorentz force and thereby produce emission as they accelerate. However, to produce gyrosynchrotron radiation at X-ray energies would require particles of unrealistically high energies, for which there is no evidence. Non-relativistic electrons radiate at the gyrofrequency  $\nu_B = 2.8 \times 10^6 B$ , where the frequency is in Hz and the magnetic field strength  $B$  in Gauss, plus a few more harmonic frequencies (Krucker et al. 2008a). Even for sunspot magnetic fields in the kG range, this cyclotron radiation is in the radio frequency. For relativistic electrons of velocity  $\beta c$ , the gyrofrequency is shifted by a factor  $\gamma = (1 - \beta^2)^{-1/2}$  and the harmonics become so closely spaced due to the relativistic effects that they merge to form a continuum radiation given by the synchrotron frequency (Ginzburg & Syrovatskii 1969; Blumenthal & Gould 1970). This has an equivalent photon energy of  $\epsilon_s = 1.74 \times 10^{-11} \gamma^2 B$  keV. Likely coronal kG fields would require electrons in the 10s of GeV energy range to produce X-rays by synchrotron emission. We know that relativistic electrons exist in the 10s of MeV range (Moses et al. 1989) in the interplanetary medium and highly relativistic positrons (of order  $\sim 100$  MeV) produce gyrosynchrotron emission in the radio and sub-mm wavelengths (Silva et al. 2007), but there is nothing that shows us the presence of

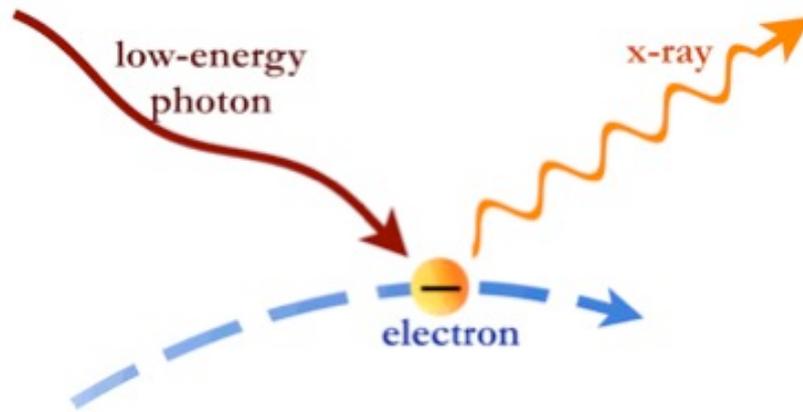


Figure 1.8: Schematic diagram showing inverse Compton emission. Courtesy S. Bank, University of Wisconsin-Madison

GeV energy electrons or positrons. Hence gyrosynchrotron has always been considered a major source of radio emission in flares but has never been considered a viable mechanism to generate X-rays.

Inverse Compton scattering (ICS) (Figure 1.8), on the other hand, has been largely ignored since [Korchak \(1967, 1971\)](#) suggested that bremsstrahlung is likely to be the main source of HXRs, but he never ruled out ICS completely. He concluded that in low-density regions of the flare where there also happened to be highly relativistic particles (electrons and positrons in particular), then ICS could be a significant source of HXRs. This is the primary focus of our Chapter 4. In Compton scattering, a photon interacts with an electron, for example, which is at rest or has negligible energy compared to the photon. In this case, the photon loses some of its energy to the electron thereby being scattered to a lower energy. However, in astrophysical cases, and in this context in flares, electrons and positrons are often accelerated to relativistic energies, much more than the energy of a visible or even EUV photon. In such a case, if a photon scatters off such a particle, it is likely to be “up-scattered” to a higher energy and hence this process is called “inverse” Compton scattering. From observations, we know that highly relativistic electrons and positrons do exist in solar flares and so there may be those situations where ICS is important. Like in the case of bremsstrahlung and recombination, ICS spectra are likely to be power-law shaped since electrons at that energy are non-thermal. However, positrons are secondary particles produced by power-law protons, but they themselves do not have power-law distributions. Hence ICS spectra produced by positrons would be quite different from those produced by electrons. The details of this and much more are in Chapter 4 ([MacKinnon & Mallik 2009](#)). ICS is already known to be a major contributor to energetic emission from

cosmic rays in intergalactic and interstellar space, but here we apply ICS to the solar context.

Since past studies have centred mainly on bremsstrahlung emission, they have been able to consider various theoretical models that would explain the spectra observed by considering only that particular process. We challenge several of these assumptions and requirements in this thesis, but for the purposes of this review chapter, we shall briefly introduce these bremsstrahlung-only models and suggest some background literature that discusses some of the limitations and controversies therein.

Some of the first quantitative models postulated and discussed are those of [Takakura & Kai \(1966\)](#) and [Holt & Ramaty \(1969\)](#). Since then, both thermal and non-thermal models have been considered to explain the observations. To explain the power-law tail in a typical spectrum using a thermal model requires you to consider a multi-thermal approach, where you have several isothermal distributions superimposed on each other to produce a power-law-looking tail (e.g. [Chubb 1972](#); [Brown 1974](#); [Crannell et al. 1978](#); [Maetzler et al. 1978](#); [Brown et al. 1979](#); [Vlahos & Papadopoulos 1979](#); [Krucker & Lin 2008](#)). This has been shown to be a viable way of arguing that thermal models alone are capable of explaining observations ([Emslie & Brown 1980](#); [Smith & Lilliequist 1979](#)), contrary to some other assertions such as those by [Kahler \(1971a,b, 1975\)](#). Moreover, quite often these models seem to be more efficient in producing HXRs and so do not require the sometimes large energy budgets needed for some non-thermal models ([Hoyng, van Beek, & Brown 1976](#); [Brown & Melrose 1977](#); [Emslie 1980](#)).

However, in spite of the above studies, ever since the seminal paper of [Brown \(1971\)](#), the non-thermal model has been the more popular path to take and explain what observations seem to suggest as a non-thermal power-law distribution of electrons above SXR energies. Analytically, this is a much more aesthetic and neat solution than a thermal-only model and it gives us some mathematical leverage to understand the conundrum of X-ray solar physics and weave mathematical and numerical models with this non-thermal model as the basis. It is also very simple in its idea since we assume an isothermal plus a non-thermal distribution of electrons. The isothermal distribution is the product of the plasma temperature. And assuming a power-law form for the non-thermal distribution is consistent with various different situations in nature where particles seem to form a power-law distribution, e.g. asteroid sizes. The non-thermal interpretation has two particular limiting model cases: ‘thick-target’ and ‘thin-target’. A thick target source is one where the accelerated electrons lose all their energy collisionally in the source itself and in doing so produce the corresponding emission during each of these interactions. A thin target source is one where the majority of the elec-

trons pass through it largely unchanged in energy, whereas those that do encounter collisions within the target produce the corresponding radiation. So a thick-target source may be those associated with dense chromospheric footpoints, where we may observe the electron loses all its energy collisionally whereas a thin-target source could be one in the corona where the atmosphere is much more tenuous. Of course, the density itself does not determine whether a source target is thin or thick; the volume and geometry are of importance as well. For example, a low-density region could be a thick-target source if it is large enough and the electron spends sufficient time in it to lose its energy collisionally. Hence thin and thick target models have different equations to describe the physics related to them and produce spectra that have particular diagnostic differences. However, all the previous studies (e.g. [Brown 1971, 1972, 1975](#); [Lin & Hudson 1976](#); [Brown & MacKinnon 1985](#); [Aschwanden & Schwartz 1996](#); [Prato et al. 2006](#), etc.) have assumed non-thermal bremsstrahlung as the only contributor to flare HXRs. In [Chapters 2 \(Brown & Mallik 2008\)](#) and [3 \(Brown & Mallik 2009\)](#), however, we have incorporated both the thin and thick target interpretations of the non-thermal model and included non-thermal recombination as a possible source of HXRs in addition to bremsstrahlung. Some of the problems interpreted successfully through thermal-only models - e.g. efficiency and energy budget requirements as explained earlier in this section - may also be answered if instead we include non-thermal recombination as a possible and sometimes significant source of HXRs in a non-thermal model, which we explain in the subsequent chapters.

## 1.2.2 Particle acceleration

It is clear and universally accepted that the source of the energy that accelerates the particles resides in the magnetic field of the Sun in solar active regions. There is clearly not enough gravitational or thermal energy in a reasonable volume to release  $10^{32}$  ergs typical of a large flare. But studies have shown that small perturbations in the magnetic field over an active region can easily account for this amount of energy ([de Jager 1969](#); [Svestka 1976](#)). However, how this energy is released and the process by which it actually accelerates charged particles are topics of hot debate. The main issue is that our observations imply that particles are accelerated to relativistic energies in a matter of seconds ([Marschhauser et al. 1991](#); [Chupp et al. 1993](#)). The basic idea common to most of these magnetic models is that field lines near a solar active region get stressed due to various dynamical changes, like the differential rotation of the Sun ([Kichatinov 1991](#)), and then break from their ‘original’ topology before reconnecting

into a more stable one. This reconnection process would release large amounts of energy and create large electric fields that should explain the acceleration of particles (and hence the radiation they produce) as well as the mass motion of solar material and also the heating of the solar atmosphere. Many different models, with different advantages and limitations have emerged over the last fifty years to describe what type of magnetic reconnection is plausible.

With improved data from instruments such as the Magnetic Doppler Imager (MDI) (Scherrer et al. 1995) on SoHO, it has been possible to move away from earlier potential (current-free) models. These were postulated for their mathematical simplicity and the lack of high resolution data to determine the magnetic structure in the solar corona - the region where reconnection and hence at least a part of the particle acceleration is believed to occur. In such a model,  $\nabla \times \mathbf{B} = \mathbf{0}$ ;  $\mathbf{B} = \nabla\phi$ , where  $\phi$  is the scalar potential. Since this is already the lowest energy state, a condition is included whereby a current *is* permitted to flow but only parallel to  $\mathbf{B}$ . This is the force-free model, which can lead to a large number of possible twisted flux tube  $\mathbf{B}$  configurations (Alfvén & Carlqvist 1967; Tanaka & Nakagawa 1973; Spicer 1976). Another feature considered is the “neutral sheet” configuration (Dungey 1953; Sturrock 1968; Priest & Heyvaerts 1974), which is produced by bipolar sunspot group interaction (Sweet 1958) or by new magnetic flux emergence (Heyvaerts, Priest, & Rust 1977). However, the energy release in these mechanisms - e.g. Sweet-Parker reconnection (Sweet 1958; Parker 1963) - does not seem quick enough and so a new model was suggested by Petschek (1964), which introduces shock-wave dissipation and thus more complex geometries that increase the energy dissipation rate by up to three orders of magnitude.

Both Sweet-Parker and Petschek reconnection are considered steady mechanisms, i.e. they are time independent. However, several models of unsteady or time-dependent reconnection models, which reflect the dynamic and burst-like nature of solar flares, have also been postulated (Dungey 1953; Craig & McClymont 1991) and in some cases the geometry was made more complex and realistic by looking at three-dimensional models (Schumacher et al. 2000; Priest & Forbes 2000; Birn et al. 2000).

As far as particle acceleration is concerned, Alfvén & Carlqvist (1967) drew the analogy between the current ( $\sim \nabla \times \mathbf{B}$ ) created in a flaring loop and an inductance circuit. The ambient plasma can only support a maximum current and that above this a decrease in the local density creates an increase in the corresponding local voltage, thereby driving the current through this higher impedance region. Therefore a large amount of current is released explosively at this instability region, creating a strong electric field and accelerating particles. Stochastic processes of particle acceleration

have also been suggested, such as those induced by plasma turbulence (e.g. [Sturrock 1968](#)) or MHD shock-waves ([de Jager 1969](#); [Hoyng 1977](#)).

In spite of all these models, what is clear is that none of them can explain all the features of a flare: i.e. the desired total energy rate or the inferred electron fluxes, especially when you consider the features simultaneously. Nevertheless, it is also clear that the models have improved with time and some day will be on the right track to explain these problems as our numerical capabilities improve and models become more sophisticated. However, the physics therein will always have to be based on previous models, essentially on the premise that magnetic energy is in some way converted to electric energy in a very short time-frame, thereby able to accelerate particles to the energies that we observe them directly and indirectly. A more recent paper by [Fletcher & Hudson \(2008\)](#) may pave the way as it explains how Alfvén-wave and turbulence-induced electron acceleration may be the answer to several perplexing flare problems. For detailed reading on energy release models, their features and limitations, please also see [Svestka \(1976\)](#); [Sturrock \(1979\)](#); [Priest & Forbes \(2002\)](#).

## 1.3 Review of Neutron Astronomy in the Inner Heliosphere

### 1.3.1 Motivation

As has been discussed earlier in this chapter, the motivations for pursuing neutron astronomy are clear. Neutrons emitted in solar flares were predicted in the middle of the last century ([Biermann et al. 1951](#)), their production mechanism reviewed some twenty years later ([Chupp 1971](#)) and then finally detected a further decade later ([Chupp et al. 1982](#)). Since then, from detecting highly relativistic neutrons on Earth and in space, our ambitions have now reached the next big step that neutron astronomy needs: observing neutrons in the inner heliosphere. This requires new technology and new instruments at the forefront of neutron detector development, which will be discussed later in this chapter. Being neutrally charged and relatively massive particles, neutrons are not only detected unaffected by strong magnetic and electric fields prevalent during flares, but they are also an excellent complement to  $\gamma$ -ray observations. They are produced in the same nuclear reactions that produce  $\gamma$ -rays, but are a powerful diagnostic tool to probe the acceleration of heavy nuclei, which do not produce narrow  $\gamma$ -ray lines. Hence studying neutrons gives us a fuller picture of particle acceleration occurring at the Sun. The University of New Hampshire (UNH) - where I undertook my undergraduate degree

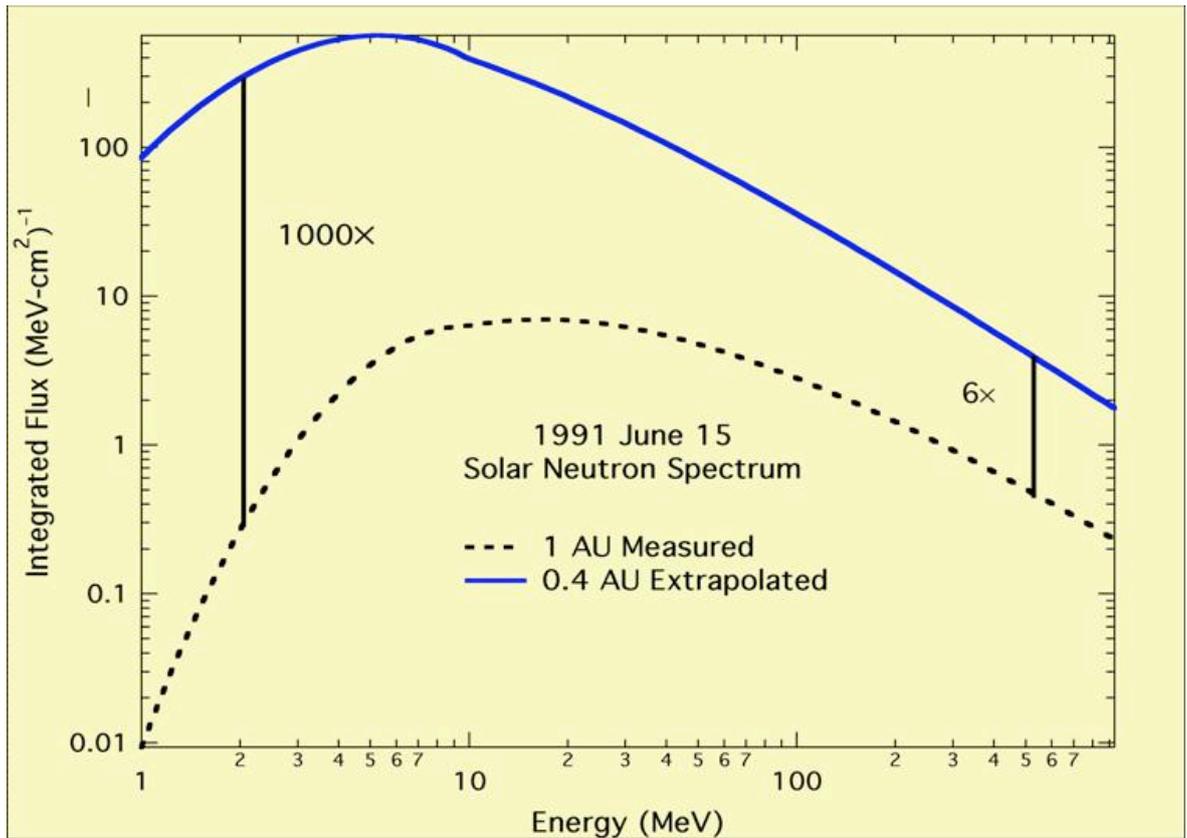


Figure 1.9: Proximity to the Sun is essential to get a good signal of incident neutrons in the 1-20 MeV range. Courtesy UNH

- in Durham, NH, USA has been at the forefront of neutron astronomy for the last forty or more years.

Free neutrons are unstable particles that spontaneously  $\beta$ -decay ( $n \rightarrow p + e^- + \bar{\nu}_e$ ) and have a lifetime of 886 seconds (Amsler et al. 2008), or just under 15 minutes. Hence only the most energetic and relativistic neutrons reach near-Earth orbits and have been observed on Earth by neutron monitors, on balloon-flight instruments as well as near-Earth orbit space instruments like COMPTEL (Ryan et al. 2000). To study neutrons in the sub-relativistic regime (below 250 MeV), one has to deploy instruments on inner heliospheric space missions. This is even more crucial for the energy range of neutrons we are interested in: 1-20 MeV. We are most interested in neutrons of this energy range because they are the most numerous among escaping neutrons. But they almost all decay near the Sun. They will not usually reach beyond 0.2 to 0.4 AU. The way fast electrons - discussed in earlier chapters - have most of their energy in the low end of the scale (1-20 keV) because there are so many of them, similarly most free neutrons and their energy are in the low-energy range of neutrons (1-20 MeV). As

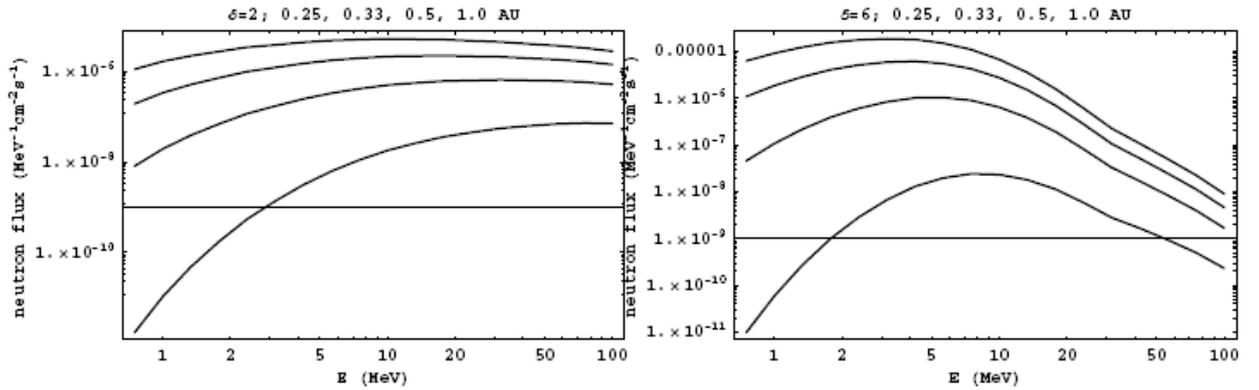


Figure 1.10: Upper-limit Quiet Sun neutron distributions assuming initially isotropic fast ion distributions at various distances from the Sun (MacKinnon & Ryan 2009). The plots are for a power-law primary ion distribution with  $\delta = 2$  (left) and  $\delta = 6$  (right). The four curves for both graphs are for four different distances from the Sun: 0.25, 0.33, 0.5 and 1.0 AU from top to bottom.

can be seen in Figure 1.9, the neutron flux increases by a factor of 1000 for neutrons in the 1-10 MeV range, by moving to an orbit of 0.4 AU, whereas the photon flux would increase by a factor of only  $R^2$ , i.e. 6.25. All these reasons make the necessity compelling for a neutron detector to be deployed on inner heliospheric space missions. MacKinnon & Ryan (2009) elaborate on the importance of studying not only flare, but quiet Sun (QS) neutrons in the inner heliosphere. The premise of QS neutrons is that there are energetic enough nuclear reactions occurring at the Sun even when it is not active, but that these reactions are likely to produce neutrons in the low-end of the energy scale, i.e.  $< 20$  MeV. Unlike during flares, when the most energetic neutrons above 20 MeV may actually reach the Earth, QS neutrons are invariably likely to be in the 1-20 MeV range and hence only detectable in the inner heliosphere, where they are the most numerous (Figure 1.10). Solar neutron spectra in general are likely to be of power-law form, like fast electrons, if detected before they start showing signs of decaying.

One of the great benefits of conducting solar particle measurements in the inner heliosphere is our ability to detect and measure solar neutrons below 10 MeV, an excellent indicator and measure of protons and ions in the low corona (Moser et al. 2005b). Neutrons uniquely sample a wide range of the proton energy spectrum (50-300 MeV), and their numbers reflect the composition of high-energy ions (Debrunner et al. 1997). When complemented with  $\gamma$ -ray data, one has indicators and measures of the proton and ion spectra from a few MeV to GeV energies. Thus, unlike  $\gamma$ -rays, one of the advantages is that neutrons provide us with a continuous energy spectrum

throughout the MeV energy range.

A challenge for deployment on a deep space mission of an instrument of this design is one of limited telemetry. To achieve the lowest possible background and to perform spectroscopy on neutrons of different energies and different arrival times, event-by-event processing and analysis is necessary. For low-Earth orbit and balloon deployment of such instruments, this event-by-event analysis was performed on the ground long after the data were obtained. These event-by-event data were obtained using telemetry with much greater bandwidths than what we will encounter on deep space missions. Therefore, much work remains to automate, as much as possible, the processing and analysis on board the spacecraft. This work breaks down into steps involving, careful instrument characterisation, instrument modelling, algorithm development, and data compression. We have been conducting research into these issues, which will enable a low-background neutron spectroscopy in the inner heliosphere. The initial form of the data is multidimensional with energy deposit information from different detectors and timing measurements between the different detectors. The final form of the data is intended to be in the form of a spectrum as a function of time at the Sun. What takes place on-board and what takes place on the ground is not clear, but with a limited telemetry budget, considerable on-board processing will be required. We will define the magnitude of the problem and develop methods to perform spectroscopy measurements with minimal dead time from limited telemetry. Some of these issues and many more have been addressed in Chapter 6.

In Chapter 5 we shall introduce the types of neutron detectors that are used and tested in labs and then describe in detail the particular instrument that has been developed at UNH. We shall present detection techniques, laboratory and simulated data analysis as well as other uses of such a neutron detector. In Chapter 6, we shall address some of the other problems mentioned above regarding data acquisition and how the data can be used to produce a neutron spectrum.

### 1.3.2 Introduction to neutron detectors

Neutron detection has many limitations and the major challenges include background noise, neutron neutrality and a variable behaviour of detectors with neutron energy. High-energy photons are the main component of background noise when detecting neutrons. Charged particles, like electrons or  $\alpha$ -particles can be screened by using a shield, but photons cannot be thus eliminated and so are a source of noise in neutron detectors. Both neutrons and light deposit similar energies in the detector and so it is hard to

distinguish them. However, by using coincidence or time-of-flight (ToF) techniques, it is possible to isolate neutron events. An example of this is shown in Figure 1.11. Since they are electrically neutral, neutrons must be detected using indirect means (Macri et al. 2007; Bravar et al. 2007; Woolf et al. 2008). They cannot be detected using gas ionisation detectors for example, unlike charged particles. One of the most common techniques is to detect the  $\gamma$ -ray emission resulting from slow (thermal or epithermal) neutron capture on a target material. In both cases the detector is surrounded by a moderator, typically polyethylene or paraffin. Multiple collisions between fast neutrons and protons in the moderator quickly “thermalise” incident neutrons, greatly increasing their detectability. Since the neutrons are detected following many interactions with a moderator the high detection efficiency comes at a price: any information about their arrival direction or primary energy is lost. Neutron Monitors (NMs) are such detectors, adding layers of lead to the paraffin, which both helps to thermalise and undergoes a neutron ‘cascade’, amplifying the initial signal. NMs usually therefore detect these secondary cascaded neutrons, most of which are cosmic-ray generated, but there have been a handful of energetic flare events where NMs have detected the cascaded effect of actual solar neutrons (e.g. Watanabe et al. 2003). Although specialised detectors, such as Bonner spheres (Bramblett et al. 1960), have been employed for neutron spectrometry applications, they rely on statistical inferences from detection of moderated neutrons and still cannot provide any information about incident direction or energy. Another limitation is that different types of detectors are sensitive to different neutron energy ranges. Detectors that rely on neutron absorption are sensitive to low-energy thermal neutrons whereas scintillation detectors are sensitive to fast neutrons. But for astronomical or solar applications, our interest is in fast neutrons (above 1 MeV) and we have been able to overcome most of these disadvantages.

The key to addressing these capability gaps lies in the ‘direct’ measurement of fast neutrons since those measurements are necessary to construct a solar energetic proton/ion spectrum as a function of time within a flare. The instrument concept described here is designed specifically for these purposes and is based on fast neutrons scattering from ambient protons (n-p scattering) that obeys hard-sphere scattering physics. From simple kinematics (see Figure 1.12), the energy of the incident neutron in terms of the energy of the recoil proton is

$$E_n = \frac{E_{p1}}{\sin^2 \theta_n}, \quad (1.1)$$

where  $\theta_n$  is the scatter angle. Organic plastic scintillator is a good choice for a scattering

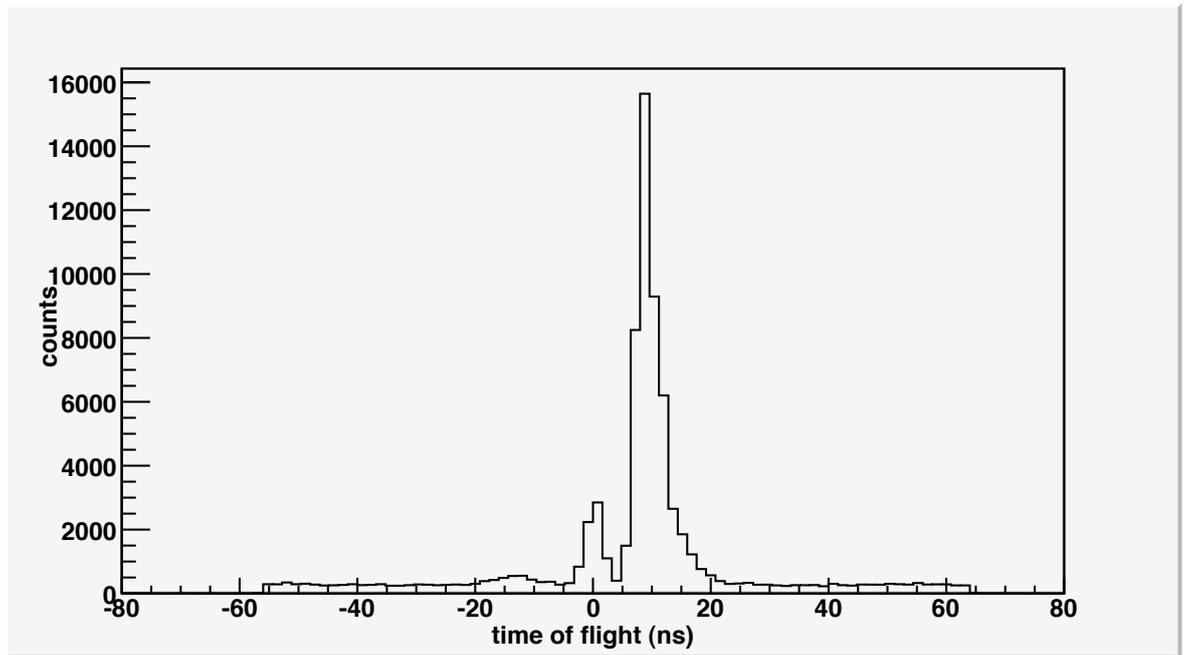


Figure 1.11: ToF spectrum showing the visible distinction between  $\gamma$ -rays (peaking around 0 ns) and neutrons ( $\sim 10$  ns); back-scattered neutrons have a ‘negative’ ToF

medium, because in addition to providing the proton target (H nuclei), it is also an efficient detector of charged particles. The choice of the best scintillator to use is based on the ratio of H to C nuclei within the plastic. The higher the ratio, the higher the efficiency of the instrument and the smaller the chance of unwanted scatters off C nuclei. However, there has to be a balance since it is the C nuclei that hold the plastic together and prevent liquid scintillators from being too volatile.

The detector is configured to locate n-p scatter sites within the detector by using the scintillation light generated due to the n-p interactions. As a neutron scatters off a proton, it transfers some of its momentum to the proton. The proton moves through the plastic medium for a few nano-seconds before coming to a stop due to the continuous deceleration produced by the ambient electrons in the plastic. This loss of energy by the proton results in the emission of light due to the atomic excitations in the material caused by the decelerating proton. The intensity of the light determines the amount of energy deposited in that n-p scatter. Multiple scattering provides important kinematic information for measuring the direction and energy of the incident neutron and for rejecting background. A telescope based on the general principles outlined here will:

- detect fast neutrons directly, without moderation, thereby preserving information about incident direction (to minimise background) and energy (to construct a

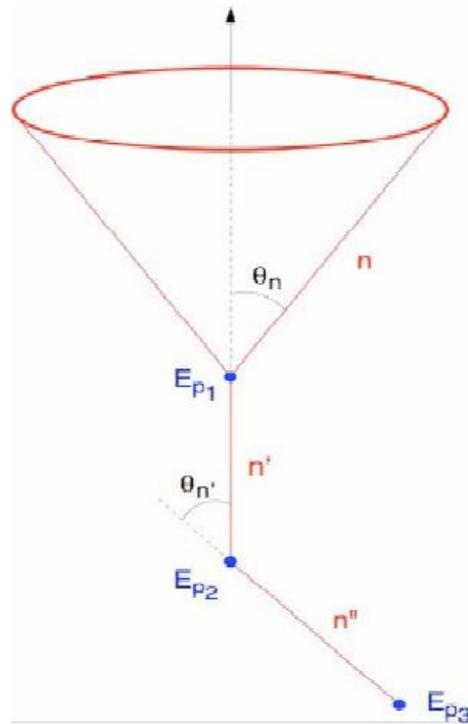


Figure 1.12: Schematic of a double scatter with the event cone and event circle/disc clearly represented. Courtesy UNH

spectrum and remove velocity dispersion effects),

- image the Sun using neutrons by reconstructing the direction of neutrons on an event-by-event basis,
- reduce the effect of  $\gamma$ -ray and isotropic neutron backgrounds using imaging and fast coincidence timing techniques,
- measure the source neutron spectrum and
- measure the arrival time of individual neutrons.

Consider the case of a double n-p scatter shown in Figure 1.12. Here a neutron, whose incident direction is unknown, undergoes two elastic scatters. For neither scatter do we measure the direction of the recoil H nucleus. The recoil neutron velocity direction is defined by the position of the two scatter points that are measured along with the pulse height (i.e. signal strength) of each scatter. If the recoil-neutron energy is fully measured, then the incident neutron direction must then lie on a cone about the recoil neutron velocity vector. The projection of this cone on the image plane or the celestial

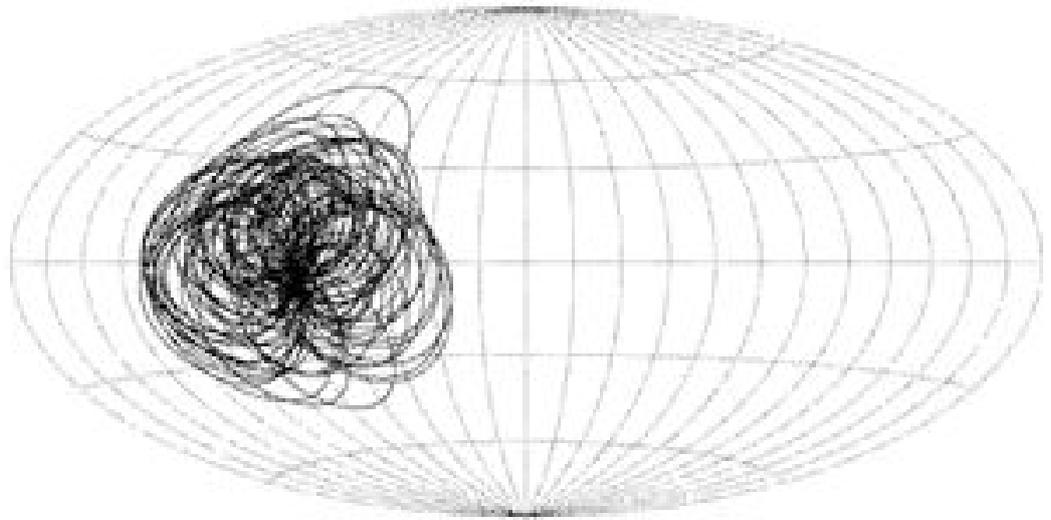


Figure 1.13: Source location identification using event circles for a GRB detected by COMPTEL. Courtesy UNH

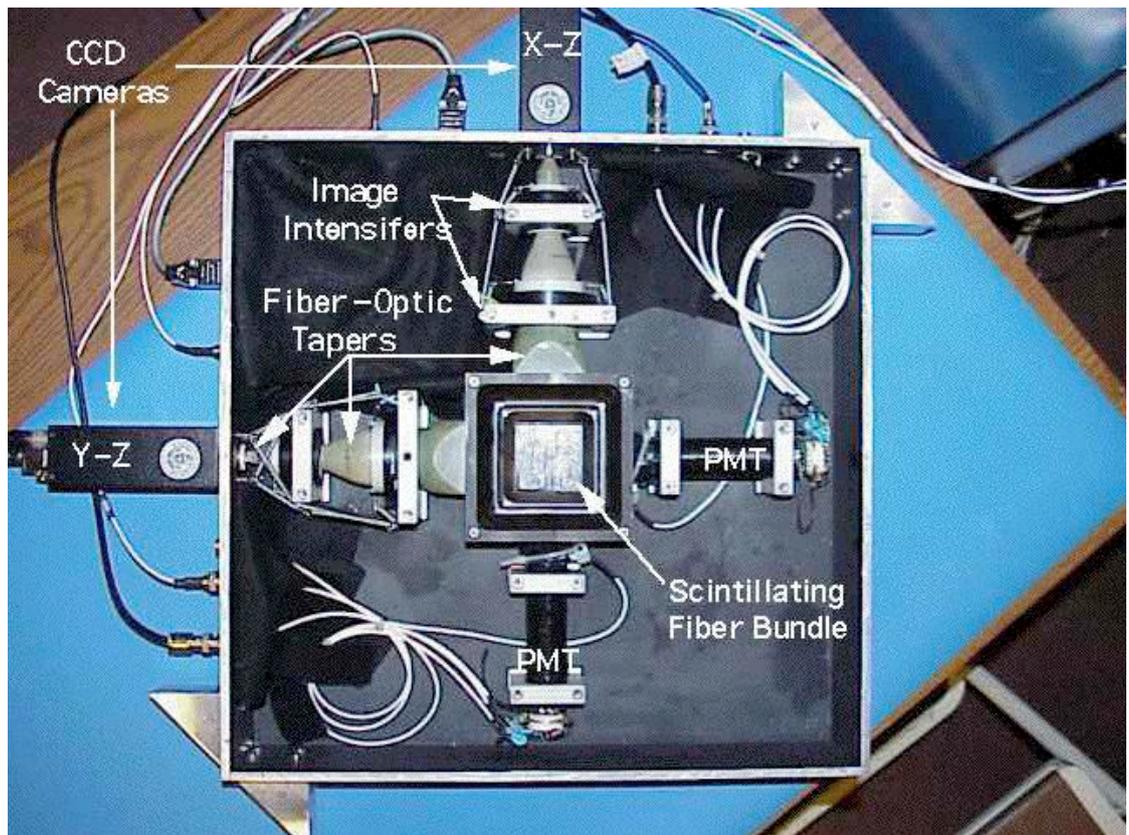


Figure 1.14: The SONTRAC instrument prototype. Courtesy UNH

sphere is an event circle. The superposition of many event circles provides the statistical information necessary to locate an unknown source. Event circles from a point source intersect, while unrelated (background) event circles do not. This procedure is the same as one used successfully on the COMPTEL experiment to image MeV  $\gamma$ -ray and neutron sources (e.g. Ryan et al. 1992, 1993; de Boer et al. 1992). An example of a COMPTEL neutron image of a solar flare was shown in Figure 1.2(i). An example of COMPTEL  $\gamma$ -ray event circles projected on the celestial sphere is given in Figure 1.13. In general, all fast neutron detectors use this principle of n-p scatters in a scintillating material, like plastic, to detect neutrons. An instrument developed by UNH and the University of Alabama-Huntsville, USA called the Solar Neutron Tracking Telescope (SONTRAC) (Figure 1.14) uses a bundle of plastic fibres, which themselves detect the neutrons by scintillating and the light emitted is measured. SONTRAC is designed to measure neutrons in the 20-250 MeV range. However, the instrument that we have been working with is the Fast Neutron Imaging Telescope (FNIT) neutron telescope, also developed at UNH, designed to detect neutrons in the 1-20 MeV range and thus optimised for the energy range of most interest in the inner heliosphere. Single scatters are useful, but triple scatters carry the most information. FNIT will be able to utilise all such events (single, double, triple) thereby providing maximum flexibility. These categories of events are discussed in detail in §5.2.

# Chapter 2

## Recombination as a Neglected Source of Hard X-rays

In Chapters 2 and 3, both of which closely follow joint-author publications, my contributions include, among other things, the following: deducing ionic fraction versus temperature relationships; introducing appropriate values for  $Z_{eff}$ ; running all the data analysis and code to produce all the plots and figures as well as verifying the mathematics and derivations in the theory.

### 2.1 Introduction

This chapter largely follows our published paper [Brown & Mallik \(2008\)](#), with amended results that have been obtained since the paper's publication. The paper made an error regarding the interpretation of some atomic physics that has been corrected here.

Ever since their first detection ([Arnoldy et al. 1968](#); [Kane & Anderson 1970](#)), flare hard X-ray (HXR) bursts (photon energies  $\epsilon > 10$  keV or so) have been recognised as an important diagnostic of electron acceleration and propagation (e.g. [Brown 1971](#); [Lin & Schwartz 1987](#); [Johns & Lin 1992](#)). The large electron flux and power imply they play a substantial role in flare energy budgets and pose challenges for electron acceleration mechanisms (see recent reviews by, e.g. [Vilmer et al. 2003](#); [Brown 2005](#); [MacKinnon 2006](#)). Recent copious high resolution HXR spectral data from the RHESSI mission ([Lin et al. 2002](#)) have created the possibility of detailed reconstruction of source electron spectra (following [Brown 1971](#)) offering important constraints on the electron energy budget and acceleration processes ([Piana et al. 2003](#); [Conway et al. 2003](#); [Massone et al. 2004](#); [Kontar et al. 2004, 2005](#); [Brown et al. 2006](#)).

In inferring electron flux spectra  $F(E)$ , the HXR radiation mechanism has always been taken to be f-f collisional bremsstrahlung of fast electron impacts with atoms and

ions, gyrosynchrotron and inverse Compton radiation being negligible at these energies for solar magnetic and radiation fields (Korchak 1967, 1971). In this thesis, we have also shown that there may be those candidate events and scenarios where inverse Compton radiation is not negligible either (see Chapter 4; MacKinnon & Mallik 2009). Though included for thermal electrons in hot (a few keV) plasma, f-b recombination radiation from non-thermal electrons seems to have been assumed negligible other than in a preliminary study by Landini, Monsignori Fossi, & Pallavicini (1973). In view of the importance of details in the photon spectrum  $J(\epsilon)$  (photons  $\text{sec}^{-1}\text{keV}^{-1}$ ) for accurate reconstruction of  $F(E)$ , we re-examine this assumption, and conclude (cf Mallik & Brown 2007) that it is not valid under some conditions, which quite commonly exist in some flare HXR source regions.

It is not the intention in this chapter to analyse precisely the theoretical recombination radiation spectrum from fast electrons under conditions (e.g. ionisation structure) for specific flares, which are typically both inhomogeneous and time dependent. Rather we give approximate theoretical estimates of how important it may be relative to bremsstrahlung under various limiting conditions. Specifically, we compare the two in the simplest, Kramers, cross-section approximations, for limiting cases of plasma ionisation. The recombination emission rate per electron is very sensitive to the ionic charge, being  $\propto Z^4 A_Z$  (Kramers 1923) per plasma proton for hydrogenic ions of charge  $Ze$  and number abundance  $A_Z$ . Thus the emitted f-b flux and spectrum depend strongly on the ionisation state, hence the temperature, of the plasma where the fast electrons recombine. In practice this will involve several ionisation stages of several target plasma species (since  $Z^4 A_Z$  may be large even for small abundance  $A_Z$ ), which will vary along the paths of the electrons and be time dependent.

This chapter is organised as follows. In §2.2 we briefly discuss relevant processes and the cross-section approximations we use, and obtain expressions for the total continuum photon spectral contributions  $j(\epsilon)$  expected from an electron flux spectrum  $F(E)$  from f-f and from f-b processes. In §2.3, we compare these for a power-law  $F(E)$  with low cut off at  $E < E_c$  and for a shifted power-law, and discuss implications for flare electron spectra and energy budgets under several limiting plasma ionisation assumptions. In §2.4 we discuss the total emission spectra from extended volumes for thin target, collisional thick target and thermal cases. In §2.5 we look at thermal and non-thermal components to show how the relative importance of each contribution depends on conditions in the flare by varying parameters around those for a specific real event. §2.6 discusses the effect of including the f-b contribution on inverse problem inference of  $F(E)$  from  $j(\epsilon)$  while §2.7 talks about the efficiency of non-thermal recombination

and introduces different possible forms of  $F(E)$ . §2.8 summarises our conclusions and suggests directions for future work.

Note that we use the terms “bremsstrahlung” and “free-free (f-f)” as well as “recombination” and “free-bound (f-b)” interchangeably throughout the thesis.

## 2.2 Bremsstrahlung and recombination HXR spectra

### 2.2.1 General considerations

In this section, we discuss only local emissivities  $j(\epsilon)$  (photons  $\text{cm}^{-3} \text{sec}^{-1}$  per unit  $\epsilon$ ). Relativistic and directivity effects are disregarded ( $E, \epsilon \ll m_e c^2$ ) since the f-b/f-f ratio is largest at low  $E$ . Then, if target atom/ion type  $t$  has density  $n_t$  and the fast electron flux spectrum is  $F(E)$  (electrons  $\text{sec}^{-1} \text{cm}^{-2}$  per unit  $E$ ),  $j(\epsilon)$  for a collisional radiation process is

$$j(\epsilon) = \sum_t j_t(\epsilon) = \sum_t n_t \int_{E_{tmin}(\epsilon)}^{\infty} F(E) \frac{dQ_t}{d\epsilon}(\epsilon, E) dE, \quad (2.1)$$

where  $dQ_t/d\epsilon(\epsilon, E)$  is the relevant cross-section per unit  $\epsilon$  for target species  $t$  and the integral is over the range of electron energies relevant to species  $t$ .

### 2.2.2 Bremsstrahlung

In the case of f-f (bremsstrahlung),  $dQ_t/d\epsilon(\epsilon, E)$  is essentially the same for any state of ionisation of an atomic species  $Z$  (Koch & Motz 1959) - since the velocities of the relevant keV electrons are high enough such that their encounters with ions are so close to the nucleus that the ionisation state of the ion is irrelevant - and the  $t$  summation in Equation 2.1 need only be carried out over elements  $Z$  to give, for element abundances  $A_Z$  (by number relative to hydrogen), and total proton (p+H) density  $n_p$ ,

$$j_B(\epsilon) = n_p \sum_Z A_Z \int_{\epsilon}^{\infty} F(E) \frac{dQ_{BZ}}{d\epsilon}(\epsilon, E) dE, \quad (2.2)$$

where  $dQ_{BZ}/d\epsilon(\epsilon, E)$  is the bremsstrahlung cross-section for element  $Z$  and  $E_{min} = \epsilon$  since any free-free transition can only yield a maximum  $\epsilon = E$ . The bremsstrahlung cross-section per nucleus  $Z$  scales as  $Z^2$  and can be written

$$\frac{dQ_{BZ}}{d\epsilon} = \frac{8\alpha r_e^2 Z^2 m_e c^2}{3 \epsilon E} q(\epsilon, E), \quad \epsilon \leq E \quad (2.3)$$

(and zero for  $\epsilon > E$ ). Here  $\alpha = e^2/\hbar c$  is the fine structure constant and  $r_e = e^2/m_e c^2$  the classical electron radius, while  $q(\epsilon, E)$  is the ratio of the actual cross section to the Kramers cross section (Kramers 1923), which is the factor in front of  $q$ . While this is only a first approximation, not suitable for accurate absolute spectral inversion/reconstruction algorithms (Brown 2005), it will be adequate for the present purpose of comparing f-f with f-b emission, which we also treat in the Kramer's approximation. Then Equations 2.2 and 2.3 give, for bremsstrahlung,

$$j_B(\epsilon) = \frac{8\alpha r_e^2 m_e c^2}{3} \frac{m_e c^2}{\epsilon} \zeta_B n_p \int_{\epsilon}^{\infty} \frac{F(E)}{E} dE, \quad (2.4)$$

where

$$\zeta_B = \sum_Z \zeta_{BZ} = \sum_Z A_Z Z^2 \quad (2.5)$$

is the heavy element correction for bremsstrahlung, with  $\zeta_B \approx 1.6$  for the solar coronal abundances we use - see later.

### 2.2.3 Recombination Radiation

The situation here is more complicated. Firstly, 2-body radiative recombination (we neglect 3-body recombination) of a free electron of energy  $E$  to a bound level  $n$  (the principal quantum number) of energy  $-V(Z, i, n)$  in ionic stage  $i$  yields a photon energy  $\epsilon$ , which, apart from quantum uncertainty, is unique, namely:

$$\epsilon = E + V(Z, i, n). \quad (2.6)$$

That is, when a fast electron does recombine, all of its kinetic energy  $E$  plus  $V$  goes into a photon of that energy, in contrast to bremsstrahlung where photons of all energies  $\epsilon \leq E$  are emitted.

Furthermore, for each element  $Z$ , there is a range of  $Z + 1$  distinct ion stages  $i$  each with its own distinct set of energy levels ( $n$ ) and a set of  $Z, i, n$ -dependent recombination cross-sections. Thus recombination collisions of a mono-energetic beam with a multi-species plasma gives rise to a set of delta-function-like spectral features at all energies (2.6) corresponding to elements  $Z$ , ionic stages  $i$  and levels  $n$ . For a continuous electron spectrum, this yields a continuum photon spectrum that is a sum of an infinite series of energy-shifted electron flux contributions. In contrast to bremsstrahlung it does not involve an integral over a continuum of electron energies.

For a general plasma the basic particle type “ $t$ ” onto which recombination occurs is level  $n$  of ion stage  $i$  of element  $Z$  with recombination cross-section differential in  $\epsilon$  for that  $t$ :

$$\frac{dQ_{Rt}}{d\epsilon}(\epsilon) = Q_{Rt}\delta(E - \epsilon + V_t), \quad (2.7)$$

where  $Q_{Rt}$  is the total radiative recombination cross-section for species  $t$  and  $\delta(E')$  is the delta-function in energy such that  $\int_{-\infty}^{\infty} \delta(E')dE' = 1$ . Then the total recombination emission spectrum for electron flux spectrum  $F(E)$  is

$$\begin{aligned} j_R(\epsilon) &= n_p \sum_t A_t \int_{E_{min}(\epsilon,t)}^{\infty} Q_{Rt}(\epsilon, E)\delta(E - \epsilon + V_t)F(E)dE \\ &= \sum_t A_t n_p Q_{Rt}(\epsilon, \epsilon - V_t)F(\epsilon - V_t), \end{aligned} \quad (2.8)$$

where  $A_t$  is the numerical abundance of species  $t$  relative to  $n_p$ . The forms for  $Q_{Rt}$ , for general  $t$ , are complicated and have to be calculated numerically, as do the values of  $A_t$  when individual ionisation states are considered. However, in the Kramers approximation (with unit Gaunt factors) there is an analytic expression for hydrogenic ions, which we will use to estimate  $dj_R/d\epsilon$  compared with  $dj_B/d\epsilon$ , namely, for recombination onto level  $n$  of the hydrogenic ion of element  $Z$  (Kramers 1923; Andersen et al. 1992; Hahn 1997)

$$Q_R = \frac{32\pi}{3\sqrt{3}\alpha} r_e^2 \frac{Z^4 \chi^2}{n^3 \epsilon E}, \quad (2.9)$$

where  $\chi = m_e e^4 / 2\hbar^2$  is the hydrogen ionisation potential.

For an element in its highest purely hydrogenic ion state the emissivity spectrum would then be

$$j_{RZ}(\epsilon) = \frac{32\pi}{3\sqrt{3}\alpha} \frac{r_e^2 \chi^2 Z^4 n_z}{\epsilon} \sum_n \frac{1}{n^3} \frac{F(\epsilon - Z^2 \chi / n^2)}{\epsilon - Z^2 \chi / n^2} \quad (2.10)$$

with the  $n$  summation over  $n \geq Z(\chi/\epsilon)^{1/2}$ , since recombination to level  $n$  yields only photons of  $\epsilon \geq Z^2 \chi / n^2$ . If the source were so hot that all atoms were almost fully ionised the total for all  $Z$  would be, in this approximation,

$$j_R(\epsilon) = \frac{32\pi}{3\sqrt{3}\alpha} \frac{r_e^2 \chi^2}{\epsilon} n_P \sum_Z Z^4 A_Z \sum_m \frac{1}{n^3} \frac{F(\epsilon - Z^2 \chi / n^2)}{\epsilon - Z^2 \chi / n^2} \quad (2.11)$$

for element abundances  $A_Z$ , with the same  $n$  summation limits.

Table 2.1: Elements with their coronal abundances and ionisation potentials at  $T \gg 10^8$  K

Element	Z	$A_z$	$A_z Z^2$	$A_z Z^4$	$V_z = Z^2 \chi$ (keV)
H	1	1	1	1	0.0136
He	2	0.096	0.384	1.536	0.0544
C	6	$3.57 \times 10^{-4}$	0.013	0.463	0.490
O	8	$8.57 \times 10^{-4}$	0.055	3.511	0.870
Ne	10	$1.07 \times 10^{-4}$	0.011	1.071	1.360
Mg	12	$1.33 \times 10^{-4}$	0.019	2.755	1.958
Si	14	$1.27 \times 10^{-4}$	0.025	4.871	2.666
S	16	$1.61 \times 10^{-5}$	0.0041	1.053	3.482
Ca	20	$8.50 \times 10^{-6}$	0.0034	1.360	5.440
Fe	26	$8.61 \times 10^{-5}$	0.058	39.336	9.914
Ni	28	$6.95 \times 10^{-6}$	0.0054	4.27	10.662
			$\Sigma = 1.58$	$\Sigma = 61.2$	

Table 2.2: Ionic species of iron at 20 MK

Element	$Z - z$	$Z_{eff}$	$\Phi_{Z_{eff}}$	$A_z$	$A_z Z_{eff}^4$	$V_z = Z_{eff}^2 \chi$ (keV)
Fe XXII	21	21.98	0.05	$0.43 \times 10^{-5}$	1.004	6.57
Fe XXIII	22	22.61	0.14	$1.21 \times 10^{-5}$	3.152	6.95
Fe XXIV	23	23.20	0.25	$2.15 \times 10^{-5}$	6.232	7.32
Fe XXV	24	23.77	0.56	$4.82 \times 10^{-5}$	15.381	7.68

In reality, even super-hot coronal flare temperatures are not high enough to equal the ultra-hot  $T \gg 10^8$  K needed to almost fully ionise all elements into their hydrogenic states, especially Fe, which is crucial in having by far the highest value of  $A_Z Z^4$  - see Table 2.1. Consequently, to deal accurately with  $j_R$  for real flare data, we would have to take into account the actual ionisation state of the flare plasma, which varies with time and location (being radically different in loop tops from loop footpoints), and actual forms of  $Q_R(Z)$ ,  $V_Z$  for non-hydrogenic ion stages.

For our purpose of making first estimates we make the following simplifying approximations:

- We treat *all* ions using hydrogenic Equations 2.9 - 2.11 but with suitably chosen  $Z_{eff}$  so that

$$V_Z = Z_{eff}^2 \chi \quad ; \quad Q_{RZ} = \frac{32\pi}{3\sqrt{3}\alpha} r_e^2 \frac{Z_{eff}^4 \chi^2}{n^3 \epsilon E}, \quad (2.12)$$

where  $Z_{eff}$  makes allowance for screening and other non-hydrogenic effects. While

this will be a rough estimate for some ions, such approximations are often quite satisfactory for suitable  $Z_{eff}$  (e.g. [Hahn & Krstic 1994](#); [Erdas, Mezzorani, & Quarati 1993](#)). Here we adopt  $Z_{eff}$  such that hydrogenic Equation 2.12 gives the correct value of  $Q_{RZ}$  as given by exact calculations such as those of [Arnaud & Raymond \(1992\)](#) for Fe, which is the most important ion in our analysis. Typically, for an element of atomic number  $Z$  in an ionic state with  $z$  bound electrons left,  $Z_{eff}$  is between  $Z - z$  and  $Z - z + 1$

- Noting that  $Q_R \propto 1/n^3$  we include here only recombination to  $n = n_{\min}$  (in the sense of the lowest empty level of the ion - hydrogenic with  $Z = Z_{eff}$  - not of the atom). Higher  $n$  contributions are weaker, being  $\propto 1/n^3$  though extending to lower energies with edges at  $Z_{eff}^2 \chi / n^2$ . These should be included in quantitative data fitting.
- The Kramers formula assumes, and is applicable to, recombination into an empty shell. For highly-charged ions, reasonable account is taken of initially partially filled  $n$ -shells by applying a ‘vacancy factor’  $p_n$  to the usual Kramers formula. The simplest choice for  $p_n$  is  $N_v / N_n$  where  $N_n = 2n^2$  is the total electron occupation number of an  $n$ -shell and  $N_v$  the number unoccupied. For recombination of a H-like ion to  $n = 1$  of a He-like ion,  $p_1 = 1/2$ . For recombination into a partially filled  $n = 2$  shell,  $N_2 = 8$  would follow. However, a more accurate result is obtained on recognising that recombination into  $n = 2$  in the Kramers formula is dominated by the 6  $p$ -states, i.e., little of the Kramers  $n = 2$  result arises from recombination into  $s$ -states, at least at the electron energies of interest here. Thus, we take  $N_2 = 6$  and  $N_v$  the number of unoccupied  $2p$  states, i.e.  $p_2 = 1$  for Li and Be-like initial ions and  $p_2 = 5/6, 4/6, \dots 1/6$  for B- through F-like initial ions. Comparisons of such modified Kramers cross sections have been made with the results of detailed calculations using the AUTOSTRUCTURE code (cf. [Badnell 2006](#)) for initial H-like through to F-like Fe ions and agreement to within 20% is obtained at the electron energies of interest here.
- We focus on situations where the emitting region is near isothermal and either quite cool, so that only low  $V_Z$  element recombination matters, or very hot so that high  $V_Z$  elements (mainly Fe) are dominant. The former are typically loop chromospheric footpoints (thick target) and the latter very hot coronal loops which are either at the limb with their footpoints occulted, or are so dense as to be coronal thick targets ([Veronig & Brown 2004](#)).

Under these conditions, Equation 2.11 becomes

$$j_R(\epsilon) = \frac{32\pi}{3\sqrt{3}\alpha} \frac{r_e^2 \chi^2}{\epsilon} n_p \sum_{Z_{eff}} \sum_{n \geq n_{min}} p_n \frac{1}{n^3} Z_{eff}^4 A_{Z_{eff}} \frac{F(\epsilon - Z_{eff}^2 \chi / n^2)}{\epsilon - Z_{eff}^2 \chi / n^2}, \quad (2.13)$$

where  $A_{Z_{eff}} = A_Z \Phi_{Z_{eff}}$  with  $\Phi_{Z_{eff}}$  the fraction of atoms of element  $Z$  in ionic state  $Z_{eff}$ .

Note that, since there is no integration over  $E$  here, if  $F(E)$  contains a sharp feature at an electron energy  $E_*$ , such as a low or high  $E$  cut-off, this will appear in the recombination contribution to the photon spectrum  $j(\epsilon)$  as a series of sharp features at photon energies  $\epsilon(n, Z, E_*) = E_* + Z_{eff}^2 \chi / n^2$ ;  $n = (n_{min}, \infty)$  for every ion  $Z$  present. The same is true for broad features like smooth bumps or dips. This is in contrast with the bremsstrahlung contribution, in which such features are smoothed out by integration over  $E$ . Thus, even if  $j_R \ll j_B$ , it may have an important effect in inferring  $F(E)$  from  $j(\epsilon)$  since this essentially involves differentiating  $j(\epsilon)$ .

#### 2.2.4 Element parameters and flare plasma ionisation

The heavy element correction for bremsstrahlung,  $\zeta_B$ , is almost independent of ionisation state (since the bremsstrahlung cross sections for atoms and ions of the same  $Z$  are essentially the same), being  $\zeta_B \approx 1.6$  for solar abundances. On the other hand  $\zeta_{RZ_{eff}} = Z_{eff}^4 A_{Z_{eff}}$  depends on the number of empty ion levels available for recombination. The importance of fast electron recombination radiation thus depends on the state of ionisation of the plasma in which the fast electrons are moving, which is primarily a function of plasma temperature  $T$ . A particular temperature determines the ionisation state of the elements. This abundance of ions of the relevant elements is assumed to remain constant for a given temperature. This assumption is used when considering f-b thermal interactions and so we too have extrapolated the same assumption when considering f-b non-thermal interactions. Questions may be raised on the presence of a highly ionised plasma and non-thermals at the same time, but through co-spatial and co-temporal observations of soft and hard X-rays in flares, we know that there are very hot ( $> 10$  MK) regions where non-thermals exist and so conclude that f-b interactions between non-thermal electrons and highly ionised species of elements such as Fe are bound to take place. That is the basis and the essence of non-thermal recombination.

In Table 2.1 we show the values of  $Z$ ,  $Z^2 A_Z = \zeta_{BZ}$ ,  $Z^4 A_Z = \zeta_{RZ}$ ,  $V_Z$  for various elements/ions whose  $\zeta_{RZ} = Z^4 A_Z$  is large enough to be significant, if the element is

sufficiently ionised. With  $\zeta_{RZ} \approx 40$  for FeXXVII, Fe is by far the most important if conditions are such that it is highly ionised. The  $kT$  where maximum ionisation of an ion stage is reached is typically of the order  $0.1Z_{eff}^2\chi$  to  $Z_{eff}^2\chi$ . In Table 2.2 we show more detailed values for several stages of ionisation of Fe (XXII-XXV, i.e. 21+ to 24+) with the appropriate  $A_{Z_{eff}} = A_{Z_{eff}}\Phi_{Z_{eff}}$  for each of these Fe ionic states for the typical coronal flare case of  $T = 2 \times 10^7$  K. These are taken from [Arnaud & Raymond \(1992\)](#) as are the actual ionisation fractions we adopt later (§2.5) for the temperatures of the real flare we consider.

The radiative recombination coefficients give  $Z_{eff}$ , which differ slightly from the  $Z$  values, as mentioned in §2.2.3. For the 2002 April 14 event, to which we return later, the peak flare temperature was 19.6 MK,  $\sim 5\%$  of the iron appearing as Fe XXII ( $Fe^{21+}$ ),  $\sim 14\%$  in the Fe XXIII ( $Fe^{22+}$ ) state,  $\sim 25\%$  appearing as Fe XXIV and  $\sim 56\%$  as Fe XXV. The respective  $Z_{eff}$  values are 21.98, 22.61, 23.20 and 23.77.

Broadly speaking in typical flare/micro-flare conditions we can consider the following  $T$  regimes:

- At  $T \leq 10^4$  K ('cold') even H and other low  $V_{Z_{eff}}$  ions are neutral so  $\zeta_{RZ} \approx 0$  for all  $Z$ . This would be typical of very dense cool chromospheric thick target footpoints relevant to deeply penetrating electrons.
- For  $10^5 \leq T \leq 10^6$  K ('cool') the predominant elements ionised are H, O, Mg, Si giving  $\sum_Z \zeta_{RZ} \approx 15$ . This is most relevant to upper chromospheric dense warm plasma reached by moderate energy thick target electrons.
- At  $T \geq 10^7$  K ('hot') Fe is well ionised up to about Fe XXV giving  $\sum_Z \zeta_{RZ} \approx 50$ . This is relevant to the hot 'coronal' loop regime, hence either to (i) typical upper (SXR) flare loops of moderate density (thin target) whose HXR emission is seen in isolation either by HXR spectroscopic imaging or volume integrated but with the cool footpoints occulted because they are over the solar limb; or (ii) cases of coronal thick target loops ([Veronig & Brown 2004](#)) where the upper loop density suffices to stop the fast electrons collisionally.

## 2.3 Local thin-target approximation for non-thermal electron distributions

### 2.3.1 Basic expressions for $j_B, j_R$

To estimate how the fast electron recombination  $j_R(\epsilon)$  compares with bremsstrahlung  $j_B(\epsilon)$ , we first consider the commonly studied case of a power-law with a low energy cut-off

$$F(E) = (\delta - 1) \frac{F_c}{E_c} \left( \frac{E}{E_c} \right)^{-\delta} ; \quad E \geq E_c, \quad (2.14)$$

where  $F_c$  is the total electron flux at  $E \geq E_c$ . Then, from Equations 2.4 and 2.14, we obtain for f-f emission

$$\begin{aligned} j_B(\epsilon) &= \frac{\delta - 1}{\delta} \frac{8\alpha\zeta_B}{3} \frac{m_e c^2 r_e^2}{\epsilon} \frac{n_p F_c}{E_c} \\ &\quad \times \left[ \frac{\epsilon}{E_c} \right]^{-\delta} ; \quad \epsilon \geq E_c \\ &\quad \times 1 ; \quad \epsilon < E_c, \end{aligned} \quad (2.15)$$

while for f-b emission from an ion of effective charge  $Z_{eff}$ ,

$$\begin{aligned} j_{RZ_{eff}}(\epsilon) &= (\delta - 1) \frac{32\pi\zeta_{RZ_{eff}}}{3^{3/2}\alpha} \frac{r_e^2 \chi^2}{\epsilon} \frac{n_p F_c}{E_c} \sum_{n \geq n_{\min}} p_n \frac{1}{n^3} \\ &\quad \times \left[ \frac{\epsilon - Z_{eff}^2 \chi / n^2}{E_c} \right]^{-\delta-1} ; \quad \epsilon \geq E_c + Z_{eff}^2 \chi / n^2 \\ &\quad \times 0 ; \quad \epsilon < E_c + Z_{eff}^2 \chi / n^2, \end{aligned} \quad (2.16)$$

where

$$\zeta_{RZ_{eff}} = A_{Z_{eff}} Z_{eff}^4. \quad (2.17)$$

So the total for all relevant  $V_{Z_{eff}}$  is

$$j_R(\epsilon) = \sum_{Z_{eff} \geq n[(\epsilon - E_c)/\chi]^{1/2}} j_{RZ_{eff}}(\epsilon). \quad (2.18)$$

### 2.3.2 Ratio of $j_R$ to $j_B$

For this truncated power-law case, the ratio of f-b to f-f emissivity is

$$\begin{aligned} \Psi &= \frac{j_R(\epsilon)}{j_B(\epsilon)} = \frac{2\pi\delta}{\sqrt{3}} \frac{\chi}{\epsilon} \sum_{Z_{eff}^2 > n^2(\epsilon - E_c)/\chi} p_n \frac{1}{n^3} \frac{\zeta_{RZ_{eff}}}{\zeta_B} \left[ 1 - \frac{Z_{eff}^2 \chi / n^2}{\epsilon} \right]^{-\delta-1} \\ &\approx \frac{0.25(\delta/5)}{\epsilon(\text{keV})} \sum_{Z_{eff}^2 > n^2(\epsilon - E_c)/\chi} p_n \frac{1}{n^3} \frac{\zeta_{RZ_{eff}}}{\zeta_B} \left[ 1 - \frac{Z_{eff}^2 \chi / n^2}{\epsilon} \right]^{-\delta-1}, \end{aligned} \quad (2.19)$$

where each term in the summation is zero at  $\epsilon < E_c + Z_{eff}^2 \chi / n^2$ .

For  $\epsilon \gg E_c$ ,  $\Psi \rightarrow 0.25 \sum_{Z_{eff}} A_{Z_{eff}} Z_{eff}^4 / \epsilon(\text{keV})$ . In pure ionised H ( $\sum_Z \zeta_{RZ} = 1$ ) this is only 2.5% at 10 keV. This rather small value of  $\Psi$  must be the origin of the conventional wisdom that f-b can be ignored compared to f-f emission at HXR energies. However, this notion neglects several crucial facts:

- At high coronal flare temperatures, where all elements are highly ionised, in plasmas of cosmic chemical abundances, heavy elements are the main contributors to the  $A_Z Z^4$  sum. For the extreme ultra-hot case of near-total ionisation of all  $Z$ , and for modern solar coronal abundances the  $\sum_Z$  factor is  $\approx 61.2$ , mainly due to Fe as discussed in §2.2.4 - see Tables 2.1 and 2.2. Note that Fe coronal abundance, for example, has been assumed to be 2.9 times photospheric Fe abundance (Feldman 1992). Even higher factors of 4-10 have been suggested (White et al. 2000; Feldman et al. 2004, Dennis, personal communication).
- At lower  $\epsilon$  the contribution from each  $Z_{eff}$  rises steeply to a sharp recombination edge at  $\epsilon = E_c + V_Z/n^2$ , where the flux can be large, especially if  $E_c$  is small and  $\delta$  large.
- At the edge, the [ ] factor in Equation 2.19 goes to  $[1 + Z_{eff}^2 \chi / n^2 E_c]^{\delta+1}$ . This is because the flux of electrons emitting recombination photons of energy  $\epsilon$  is not the flux of those at  $E \geq \epsilon$ , as for bremsstrahlung, but of those at  $E = \epsilon - Z_{eff}^2 \chi / n^2$ . Consequently  $\Psi$  is not negligible even at  $\epsilon \gg E_c$ . For fully ionised Fe alone, this factor is  $\approx [1 + 10/E_c(\text{keV})]^{\delta+1}$ , which, for  $\delta = 5$  and at  $\epsilon = 10$  keV, is 64, 11.4, 5.5 for  $E_c = 10, 20, 30$  keV respectively. Even for lower stage Fe ions (e.g. XXV), common in flare coronal loops, evidently recombination must be a significant contributor to the HXR emission in those parts of the flare.

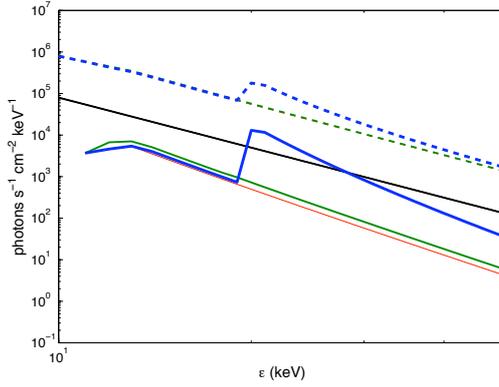
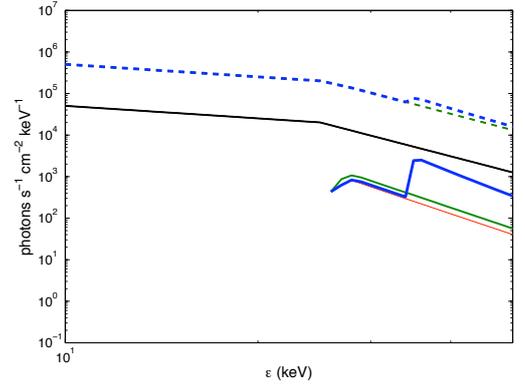
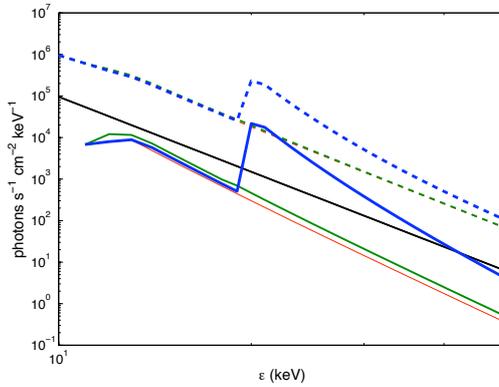
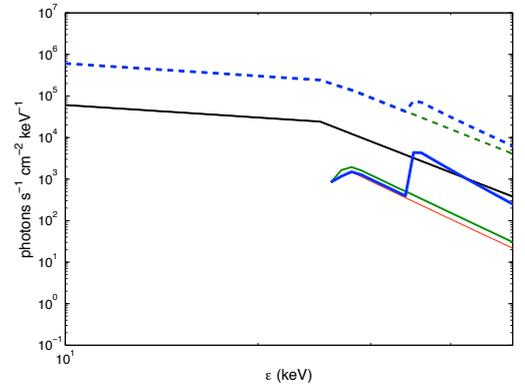
(a)  $E_c = 10$  keV;  $\delta = 3$ (b)  $E_c = 25$  keV;  $\delta = 3$ (c)  $E_c = 10$  keV;  $\delta = 5$ (d)  $E_c = 25$  keV;  $\delta = 5$ 

Figure 2.1: Actual shapes of non-thermal f-b and f-f spectra for different temperature regimes and non-thermal electron parameters. The solid curves represent non-thermal cool (red), hot (green) and ultra-hot (blue) f-b spectra as well as the f-f (black) spectrum. The dashed curves represent the total spectrum multiplied by 10.

### 2.3.3 Typical results in limiting regimes

N.B. All spectrum figures in this chapter (except Figures 2.3, 2.8 and 2.9) have been plotted for a bin-width of 1 keV to match RHESSI's spectral resolution. However, in the figures mentioned above, we use 0.1 keV resolution so as to compare them with the other plots and appreciate the sharpness of the edge features to see how f-b edges would look if they were observed at a higher resolution. The 1 keV binning smears out a lot of the edges of different elements that are clearly noticeable in, for example, Figure 2.8. Hence in Figures 2.4 and 2.7, the edges are not 'infinitely' steep as they

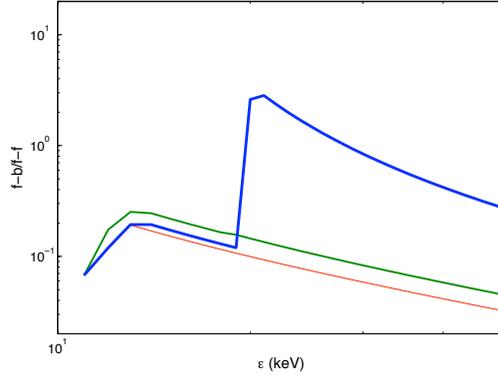
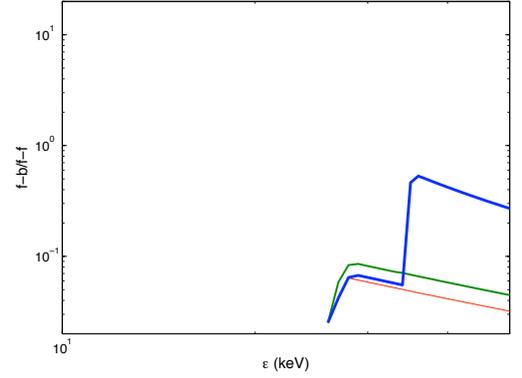
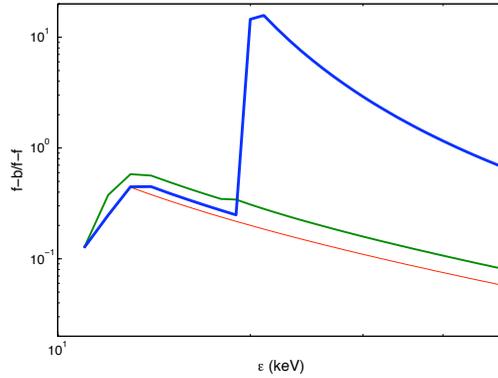
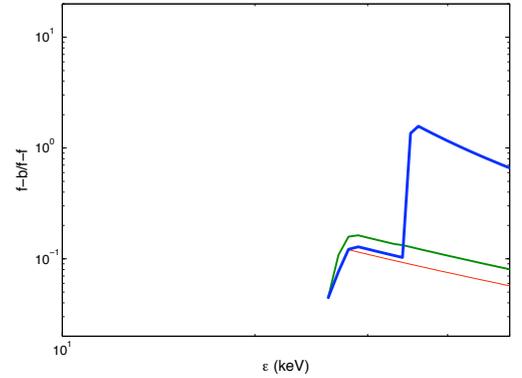
(a)  $E_c = 10$  keV;  $\delta = 3$ (b)  $E_c = 25$  keV;  $\delta = 3$ (c)  $E_c = 10$  keV;  $\delta = 5$ (d)  $E_c = 25$  keV;  $\delta = 5$ 

Figure 2.2: Photon flux ratio of non-thermal f-b to f-f emission for different temperature regimes and parameters. Line styles have the same meaning as in Figure 2.1.

should be; this is evident in Figure 2.8 where they do look far steeper due to the finer resolution. In theory, if the resolution was infinitesimally fine, then the edges would be infinitely steep. Also important to note is that the features seen in Figures 2.4, 2.7, 2.8 etc. are recombination edges and not spectral lines. None of the figures in this thesis include spectral lines - leaving them out shows more clearly where f-b edges exist in the HXR continuum.

In Figure 2.1 we show for  $\delta = 3, 5$  the actual spectral shapes for  $E_c = 10, 25$  keV respectively in plasmas of normal solar coronal abundances, which are: ultra-hot ( $T \gg 10^8$  K; Fe is nearly fully ionised), hot ( $T = 2 \times 10^7$  K; Fe well ionised up to Fe XXV) and cool ( $T = 10^6$  K; elements up to Si are almost fully ionised). In Figure

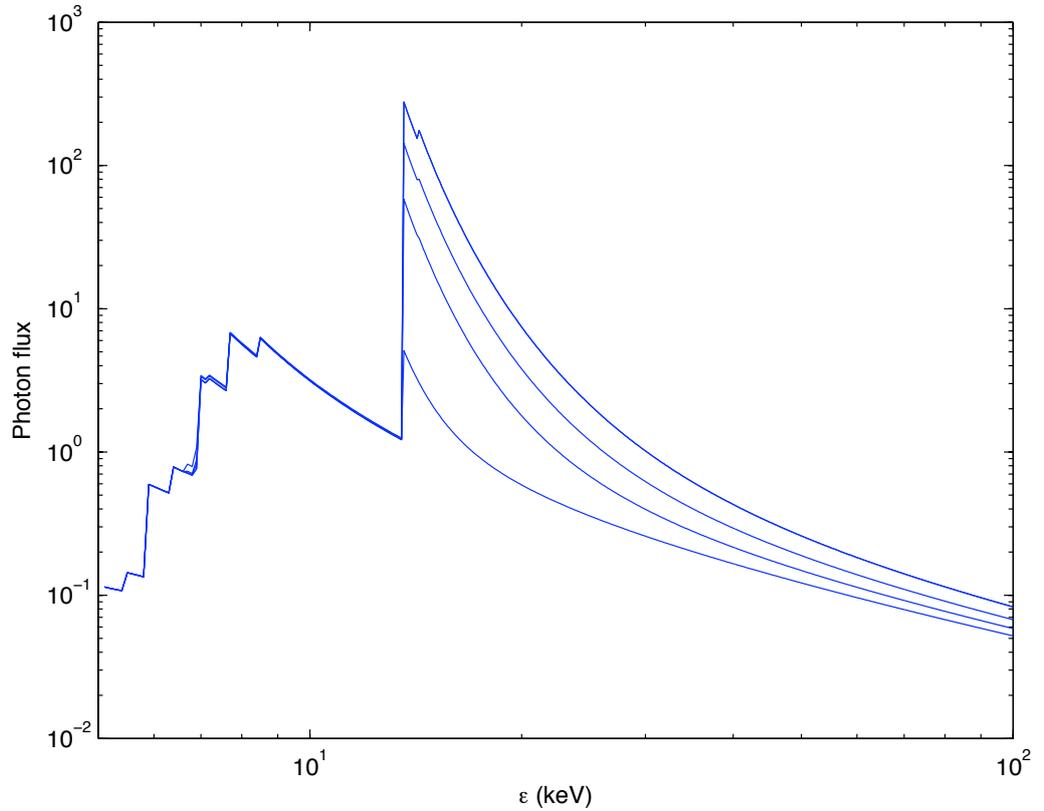


Figure 2.3: Thin-target non-thermal f-b:f-f ratio for  $E_c = 5$  keV and  $\delta = 7$  for 20, 30, 40 and 50 MK. Each incremental increase in the temperature produces a larger f-b:f-f ratio from the Fe edge at about 14 keV ( $= E_c + V_{Fe25+,26+}$ )

**2.2** we show the ratios  $\Psi(\epsilon)$  for the ultra-hot, hot and ‘cool’ cases, respectively. The following key features of the hot thin target situation are apparent from these Figures:

- The maximum non-thermal f-b contribution, in each hot or cool case shown, adds about 30% to the usual f-f one and in some cases ( $E_c = 10$  keV, ultra-hot) is up to 10 times greater (1000% increase). In §2.7 we evaluate the efficiency with which f-b yields HXRs compared to f-f, and also derive the ratio  $\Psi$  for the case of a smooth  $F(E)$  with no cut-off. This proves, that in a hot enough plasma, far fewer electrons and less power are needed than is found when only f-f is included and that, for smooth  $F(E)$ ,  $\Psi$  is largest for large  $\delta$  and low  $E$  spectral roll-over.
- In the ‘cool’ case ( $T \approx 10^6$  K) of elements up to Si almost fully ionised, the f-b contribution is significantly smaller but not in general negligible. For example, in the bottom left panel of Figure 2.1 ( $\delta = 5, E_c = 10$  keV), f-b is about 20 %

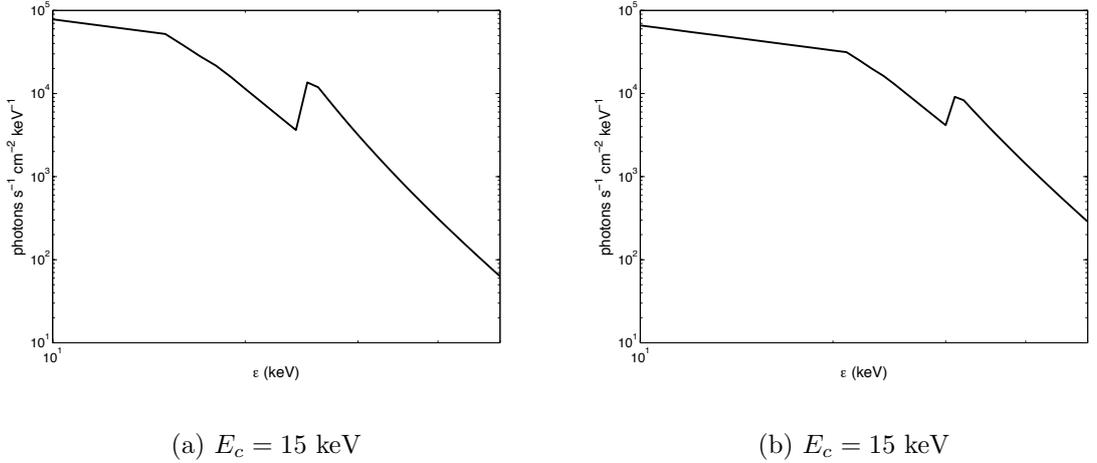


Figure 2.4: Spatially localised spectra from a loop with fully ionised plasma for two values of  $E_c$ . The left plot shows a very distinct iron edge at  $\approx 24$  keV ( $= E_c + V_{Fe26+}$ ) and a break at 15 keV ( $= E_c$ ), whereas the second plot shows an iron edge ( $\approx 28$  keV) and a break at  $E_c = 21$  keV. This shows the value of recombination as an  $E_c$  diagnostic. The ‘edges’ appear to be of finite slope because of the finite (1 keV) resolution used.

of f-f at 15 keV energies. This is amply large enough to have a major impact on inferring  $F(E)$  by inversion or by forward fitting (§2.6).

- In hot plasmas, Fe is by far the most important contributor of recombination radiation.
- The peak ratio of f-b to f-f increases as  $\delta$  is increased and/or  $E_c$  is decreased. This is because f-b photons of energy  $\epsilon$  are emitted by electrons of energy  $E - V$  which have flux  $F(E - V) \propto (E - V)^{-\delta}$  which is greatest when the minimum  $E = E_c$  is smallest,  $V$  is largest and the steepness  $\delta$  greatest.
- Recombination edges are apparent for the elements with the highest values of  $A_{Z_{eff}} Z_{eff}^4$  - Fe, Si, Mg and O and at energies  $\epsilon = E_c + Z_{eff}^2 \chi / n^2$ , thereby creating the possibility of finding the location of a low energy cut-off  $E_c$  should one exist.
- The harder asymptotic  $\gamma = \delta + 1$  for f-f compared with  $\gamma = \delta + 2$  for f-b (Equations 2.15 and 2.16) results in a gradual bend upwards, or ‘ankle’, in the total spectrum clearly visible in Figure 2.1 for  $E_c = 10$  keV but also present for higher  $E_c$  outside the  $\epsilon$  range of the Figure. This could be an important signature in data of a substantial f-b contribution.

While the edge locations and the spectral shape trends will be roughly right, our use

of the hydrogenic and  $Z_{eff}$  approximations, and adoption of unit Gaunt factors, mean that these curves/analytic forms can only be used for approximate quantitative fitting of real data. As far as we are aware (Kaastra, personal communication) the Gaunt factors, rates etc. have only ever been systematically evaluated for Maxwellian  $F(E)$  and sometimes for forms which can be written as sums of these (such as pure power-laws with no cut-off), and some occasional consideration of specific non-thermal spectra (e.g. Landini, Monsignori Fossi, & Pallavicini 1973). Comparison of our Maxwellian results, in the unit Gaunt factor Kramers approximation, with those of Culhane for the same parameters shows the necessary corrections in the Maxwellian case to be significant for quantitative comparison with real data. In addition, in real cases the non-thermal emission will always be superposed on thermal contributions (especially important for the very hot plasmas of special interest here) and also in many cases on a thick target non-thermal contribution (unless this is from occulted footpoints), from the flare volume as a whole. In the next section we derive the generalisation of the above equations to the various cases involved in real flares, viz. finite volume thin targets, Maxwellian plasmas and thick targets for use in §2.5, where we evaluate the sum of all these contributions for a specific case.

## 2.4 Whole flare thin-target, thermal and thick-target expressions

Here we extend the above results on local emissivities  $j(\epsilon)$  to estimate total spectral emission rate  $J(\epsilon)$  (photons  $\text{sec}^{-1}$  per unit  $\epsilon$ ) from extended flare volumes as required for real flare data.

### 2.4.1 Thin Target Coronal Loop

A thin target is one in which  $F(E)$  is not significantly modified by energy losses or gains over the volume. For a loop of half length  $L$ , transverse area  $A$ , volume  $2AL$  and density  $n_p$ , the total emission rate spectra contributions  $J_{thin}(\epsilon)$  are for a power law  $F(E)$  with a low energy cut-off, by Equation 2.15,

$$J_{B_{thin}}(\epsilon) = \frac{\delta - 1}{\delta} \frac{8\alpha\zeta_B}{3} \frac{m_e c^2 r_e^2}{\epsilon} \frac{2n_p A L F_c}{E_c} \times \begin{cases} \left[ \frac{\epsilon}{E_c} \right]^{-\delta} & ; \epsilon \geq E_c \\ 1 & ; \epsilon < E_c \end{cases} \quad (2.20)$$

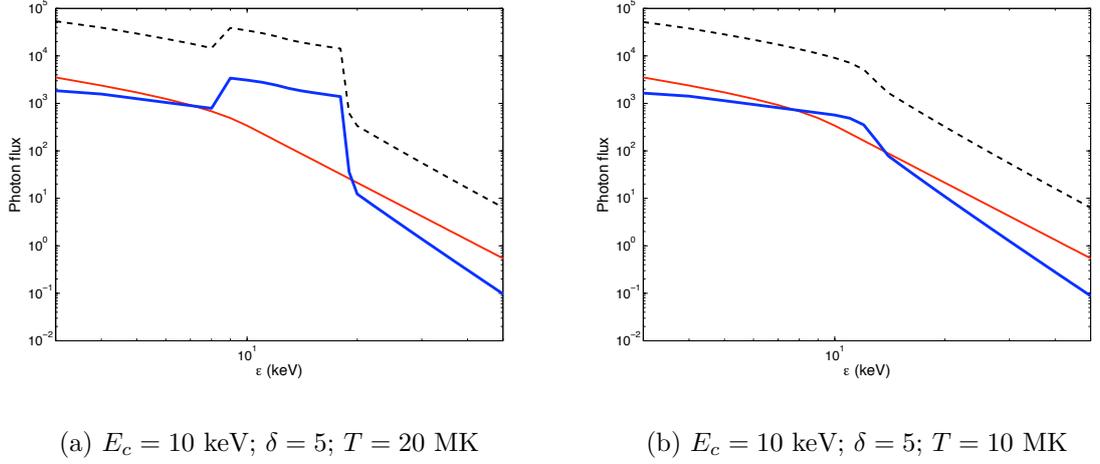


Figure 2.5: Non-thermal f-f and f-b spectra for the thick target case (Equations 2.29 and 2.30) shown for 2 different temperatures: 20 MK that is pertinent to events such as the 2002 April 14 event and 10 MK, which is more in the range of ‘microflare’ temperatures. The red curve represents the non-thermal f-f spectrum and the blue is for non-thermal f-b. The dashed black curve is the total non-thermal emission multiplied by 10. It is interesting to note the three distinct energy regimes for the f-b spectrum, namely:  $\epsilon < V_{Fe}$ ;  $V_{Fe} \leq \epsilon \leq V_{Fe} + E_c/n^2$ ;  $\epsilon > V_{Fe} + E_c/n^2$ . Clearly f-b is very important in the 10-50 keV range, precisely where albedo issues are also important.

and by Equation 2.16,

$$\begin{aligned}
 J_{R_{thin}}(\epsilon) = & (\delta - 1) \frac{32\pi\zeta_{RZ_{eff}} r_e^2 \chi^2 2n_p ALF_c}{3^{1/2}\alpha \epsilon E_c^2} \times \sum_{Z_{eff}} \sum_{n \geq n_{min}} p_n \frac{1}{n^3} \\
 & \times \left[ \frac{\epsilon - Z_{eff}^2 \chi / n^2}{E_c} \right]^{-\delta-1} ; \epsilon \geq E_c + Z_{eff}^2 \chi / n^2 \\
 & \times 0 ; \epsilon < E_c + Z_{eff}^2 \chi / n^2,
 \end{aligned} \tag{2.21}$$

where the summation is over all  $Z_{eff} \leq n(\epsilon - E_c)^{1/2}$ . These spectral shapes  $J(\epsilon)$  are of course just the same as the thin target  $j$  forms, scaled by the plasma volume.

### 2.4.2 Hot Coronal Loop Thermal Emission (in the Kramers approximation)

Both f-f and f-b emissions are included in the standard analyses (e.g. Mewe et al. 1986; Dere et al. 1997) of isothermal hot plasma contributions to flare spectra, using full cross sections and ionisation balance expressions. It is therefore surprising that f-b is

omitted from calculations of non-thermal emission, especially at low  $\epsilon$ , where electrons of comparable energy are present in both thermal and non-thermal populations. In applying our study of the non-thermal f-b to real data we wish to include thermal emission as it is important at energies under about 20 keV and so dilutes the visibility of non-thermal contributions. In order to treat the thermal and non-thermal  $j$  consistently and allow meaningful comparisons we use the expressions for the thermal  $j$  relevant to the Kramers cross sections just as in the non-thermal case - but see remarks previously and below concerning Gaunt factors and absolute accuracy of our results.

For an isothermal plasma the local Maxwellian electron flux spectrum is

$$F_{therm}(E) = \left[ \frac{8}{\pi m_e} \right]^{1/2} \frac{E}{(kT)^{3/2}} n_p \exp(-E/kT), \quad (2.22)$$

which, by Equation 2.4, gives for the thermal bremsstrahlung emission from a uniform loop

$$J_{Btherm}(\epsilon) = \frac{16\alpha r_e^2}{3} \zeta_B m_e c^2 \times \left[ \frac{8}{\pi m_e} \right]^{1/2} \frac{2n_p^2 AL e^{-\epsilon/kT}}{\epsilon(kT)^{1/2}} \quad (2.23)$$

and for the recombination

$$J_{Rtherm}(\epsilon) = \sqrt{\frac{2\pi}{27m_e}} \frac{64r_e^2 \chi^2}{\alpha} \frac{2n_p^2 AL}{\epsilon(kT)^{3/2}} \sum_{Z_{eff}} \zeta_{RZ_{eff}} \sum_{n \geq n_{min}} p_n \frac{1}{n^3} \exp\left(\frac{Z_{eff}^2 \chi / n^2 - \epsilon}{kT}\right).$$

In Figure 2.6, we plot our thermal model f-f and f-b spectra along with that predicted by the atomic database, Chianti (e.g Dere et al. 2009), following on from Landi (2007). The agreement is good for both the magnitude and location of all the edges, although some discrepancies exist since Chianti uses more accurate measured cross-sections as opposed to Kramers' that we have used and Chianti also includes recombination on to higher  $n$  levels rather than just  $n_{min}$ .

These results can be compared with those of Culhane (1969); Culhane & Acton (1970) who were among the first to explicitly address the X-ray spectrum from hot coronal plasmas. Using the Kramers cross sections is essentially equivalent to setting to unity all Gaunt factors in their expressions. When we do so, the  $\epsilon, T$  dependences of our  $J_{Rtherm}, J_{Btherm}$  are identical to theirs - e.g.  $J_{Rtherm}/J_{Btherm}$  is independent of  $\epsilon$ , the only difference being that our  $J_{Rtherm}$  is larger (in absolute value) than theirs, mainly because they used the very much lower value of  $A_Z$  for Fe believed at that time. With modern instruments and improved quality of data, it has been estimated that abundance of low first ionisation potential (FIP) elements such as Fe are much higher

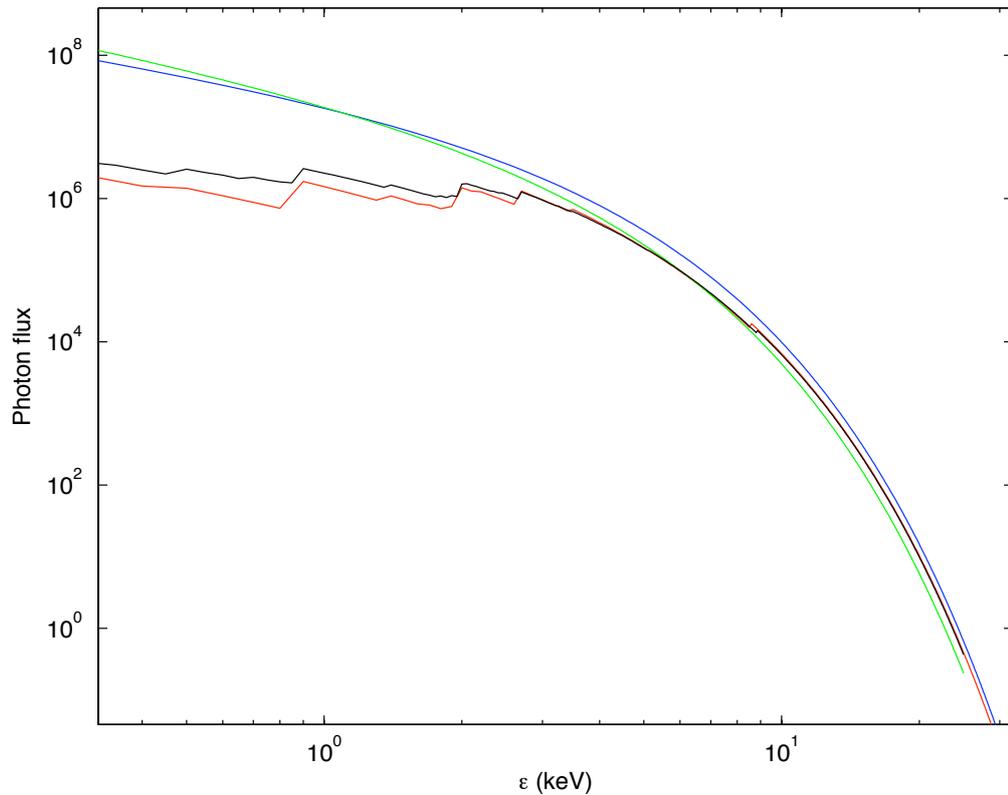


Figure 2.6: The thermal model spectra (blue: thermal f-f; red: thermal f-b) compared with Chianti's (green for f-f and black for f-b). Courtesy E. Landi for the Chianti spectra.

in the corona than in the photosphere and hence much higher than estimates used in the past. Examination of the  $\epsilon, T$  dependences of Culhane's Gaunt factors shows that they affect quite significantly both the f-f and the f-b spectra from a Maxwellian  $F(E)$  and we should expect the same to be true for non-thermal  $F(E)$  like power-laws. Thus, any accurate absolute comparison of predictions with data will require incorporation of appropriate  $g, G$ . However, these do not affect the absolute orders of magnitude of  $J_{Rtherm}, J_{Btherm}$  nor the dependences on  $n_p, V, F_c$  etc., nor the locations of edges. So, for the present purpose of demonstrating the importance of f-b, the Kramers expressions will suffice.

It is also important to note that in addition to interpreting and observing possible non-thermal recombination edges, which has been the focus in this chapter, our model shows that it is also possible that thermal recombination edges be observed. As can be seen in Figure 2.8, thermal edges should occur at set ionisation potential energies

and their presence could tell us something about both the plasma temperature and the relative abundance of elements. The thermal edges we predict closely resemble those discussed by Landi (2007). So although the magnitude of these edges is likely to be small, with higher resolution instruments, some day these edges should be detectable paving a whole new way to diagnose flare parameters.

### 2.4.3 Thick target (dense loop or footpoint) f-f and f-b emission spectra

In the thick target case,  $j$  evolves in space along with the energy losses of the electrons. To find  $j$  locally one uses the continuity equation (Brown 1972) and then integrates over volume to get  $J$ . However, to get the whole volume  $J$ , it is actually simpler (Brown 1971) to start with the electron injection rate spectrum  $\mathcal{F}_o(E_o)$  electrons/sec per unit injection energy  $E_o$  and use the expression

$$J_{thick}(\epsilon) = \int_{E_o} \mathcal{F}_o(E_o) \eta(\epsilon, E_o) dE_o, \quad (2.24)$$

where  $\eta(\epsilon, E_o)$  is the total number of photons per unit  $\epsilon$  emitted by an electron of energy  $E_o$  as it decays in energy. For purely collisional losses  $dE/dN = -K/E$  with  $K = 2\pi e^4 \Lambda$ ,  $e$  being the electronic charge and  $\Lambda$  the Coulomb Logarithm. Then

$$\eta(\epsilon, E_o) = \frac{1}{K} \int_E E \frac{dQ}{d\epsilon} dE \quad (2.25)$$

for the relevant radiation cross section  $dQ/d\epsilon$ . Note that this assumes H to be uniformly and fully ionised along the electron path. For partially ionised H the energy loss constant  $K$  is reduced but this situation is not relevant to our hot source situations.

For our Kramers  $dQ/d\epsilon$  f-f and f-b expressions 2.3, 2.7 and 2.9, the resulting expressions, in the case where  $A_{Z_{eff}}$  are uniform along the path, Equation 2.25 gives

$$\eta_B(\epsilon, E_o) = \frac{8\alpha\zeta_B r_e^2 m_e c^2}{3} \frac{1}{K} \times \begin{cases} \left[ \frac{E_o}{\epsilon} - 1 \right]; & \epsilon \leq E_o \\ 0; & \epsilon > E_o \end{cases} \quad (2.26)$$

and

$$\eta_{RZ}(\epsilon, E_o) = \frac{32\pi A_{Z_{eff}} Z_{eff}^4 r_e^2 \chi^2}{3^{3/2} \alpha} \frac{1}{K \epsilon} \sum_{n \geq n_{\min}} p_n \frac{1}{n^3} \times$$

$$1; E_o \geq \epsilon + Z_{eff}^2 \chi / n^2$$

$$0; E_o < \epsilon + Z_{eff}^2 \chi / n^2. \quad (2.27)$$

For a power-law injection rate spectrum of spectral index  $\delta_o$ , viz

$$\mathcal{F}_o(E_o) = (\delta_o - 1) \frac{\mathcal{F}_{oc}}{E_{oc}} \left[ \frac{E_o}{E_{oc}} \right]^{-\delta_o}; E_o \geq E_{oc}, \quad (2.28)$$

where  $\mathcal{F}_{oc}$  is the total rate above low energy cut-off  $E_{oc}$ , the expressions for the non-thermal emission spectra are then by Equation 2.24

$$J_{Bthick}(\epsilon) = \frac{8\alpha r_e^2}{3} \frac{\zeta_B m_e c^2 \mathcal{F}_{oc}}{(\delta_o - 1)(\delta_o - 2)K} \times$$

$$\left( \frac{\epsilon}{E_c} \right)^{-\delta_o + 1}; \epsilon \geq E_c$$

$$\left[ (\delta_o - 1) \frac{E_c}{\epsilon} - (\delta_o - 2) \right]; \epsilon < E_c \quad (2.29)$$

and, for ion  $Z_{eff}$ ,

$$J_{RZ_{eff}thick}(\epsilon) = \frac{32\pi r_e^2}{3^{3/2} \alpha} \zeta_{RZ_{eff}} \frac{\chi^2}{K \epsilon} \mathcal{F}_{oc} \times \sum_{n \geq n_{\min}} p_n \frac{1}{n^3} \times$$

$$\left[ \frac{\epsilon - Z_{eff}^2 \chi / n^2}{E_{oc}} \right]^{-\delta_o + 1}; \epsilon \geq E_{oc} + Z_{eff}^2 \chi / n^2$$

$$\left[ \frac{E_{oc} - Z_{eff}^2 \chi / n^2}{E_{oc}} \right]^{-\delta_o + 1}; Z_{eff}^2 \chi / n^2 < \epsilon < E_{oc} + Z_{eff}^2 \chi / n^2$$

$$0; \epsilon < Z_{eff}^2 \chi / n^2. \quad (2.30)$$

For the case of a cold thick target footpoint the total  $\zeta_R$  can be almost as small as 1 if only hydrogen and some low  $\zeta_R$  elements are ionised and even zero if  $T < 8000$  K or so (there being almost no charged ions present). In these sources the f-b contribution is negligible or at most a very small correction. For a collisionally thick hot loop  $\zeta_R$  is, however, very much higher.

The main distinction of these hot thick target spectra compared to hot thin targets is that the decay of all electrons to zero energy means that the signature of the cut off  $E_{oc}$  in the injection spectrum appears not as a discontinuity in  $J(\epsilon)$  but only in its gradient  $J'(\epsilon)$ . This gradient break is very noticeable in Figure 2.5 at energy  $\epsilon = E_c + V_{Fe}/4$ . So, even in the thick target case, spectral diagnosis of any  $E_{oc}$  present is possible. The recombination edges themselves appear corresponding to the relevant ionisation energies  $\epsilon = V_{eff}$  (edges appearing at  $\epsilon = V_{eff}/n^2$ ), these being from thick target electrons decelerated to zero  $E$ . These non-thermal recombination spectral edges are then down in the energy regime, well below 10 keV, which is complicated by Fe lines etc., making the interpretation of  $\mathcal{F}_o$  there, and of the lines, more difficult.

## 2.5 Practical case study results derived from a real flare

We saw above that the most favourable conditions for a substantial recombination contribution are when the maximum possible amount of the observable HXR source is a hot plasma (e.g. loop) at SXR temperatures. High density maximises the emission measure but may make the source/loop collisionally thick and smear recombination edge spectral signatures of low energy cut offs. So an optimal case could be a loop which is just tenuous enough to be collisionally thin and for which the cool dense thick target footpoints are occulted. (Footpoint removal by imaging is limited by RHESSI's dynamic range). Such sources will have a strong HXR source in the coronal loop. One such event was adopted as a basis for a case study, starting from the real event parameters. This was the 2002 April 14 event, which Veronig & Brown (2004) showed to be a hot, dense, collisionally thick loop with a strong coronal HXR source and no footpoints up to at least 60 keV. Thus the hot coronal source of non-thermal f-b emission was not diluted by cold footpoint thick target f-f emission though the f-b edges were smeared because the hot loop itself slowed the fast electrons to rest. In Figure 2.4 we show the theoretical spectrum from a hypothetical resolved part of the coronal loop for two  $E_c$  values. We have evaluated the theoretical thermal, non-thermal and the whole volume hypothetical total  $J_B(\epsilon)$ ,  $J_R(\epsilon)$  (from §2.2, 2.3 and 2.4) for such a loop, based on our approximate Kramers expressions, in two loop parameter regimes - Figure 2.7:

- Plot A: With the actual hot thick target loop parameters found by Veronig & Brown (2004), namely  $\delta = 6.7$ ;  $T = 19.6$  MK;  $L = 45 \times 10^8$  cm;  $A = 19.1 \times 10^{16}$

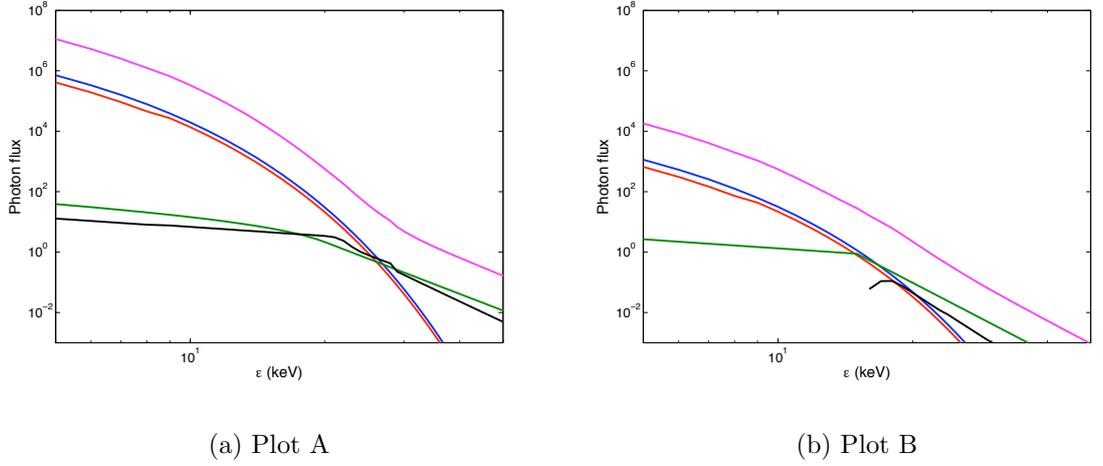


Figure 2.7: The spectral components for 2 different hypothetical situations. We show these spectra by varying the parameters around the results in the [Veronig & Brown \(2004\)](#) paper that analyses the coronal thick target 2002 April 14 event. In all cases we keep the same values of  $\delta = 6.7$ ,  $E_c = 15$  keV and  $T = 19.6$  MK. The curves represent the non-thermal f-f (green), non-thermal f-b (black), thermal f-f (blue) and thermal f-b (red) spectra. The magenta curve shows the total spectrum multiplied by 10. Plot A is for the thick-target coronal case with the actual event parameters  $n_p, \mathcal{F}_{oc}$  according to Veronig and Brown. Plot B was obtained for the same event parameters but with  $n_p$  reduced 25 times to make the loop collisionally thin above 10 keV and with footpoint emission occulted. The injection rate is the same as Plot A so the density fraction of fast electrons is 25 times higher. The non-thermal emission is down by 25 times while the thermal is down by a factor of 625.

$\text{cm}^2$ ;  $n_p = 10^{11} \text{ cm}^{-3}$ ;  $N = 4.9 \times 10^{20} \text{ cm}^{-2}$ ;  $\mathcal{F}_1 = 5 \times 10^{35} \text{ sec}^{-1}$  above  $E_1 = 25$  keV. The total  $J$  is dominated by thermal f-b and f-f at low  $\epsilon$  but thick-target f-b at medium  $\epsilon$  and thick-target f-f at high  $\epsilon$ . Locally within the loop volume, if this were spatially resolved, the spectrum  $j$  would be like those in [Figure 2.4](#), where edges are clearly visible in positions corresponding to cut-off energies of 15 and 21 keV. At a higher resolution, these edges would look similar to the edges shown in [Figure 2.8](#). Should such edges be found in data, they can diagnose the all-important  $E_c$  parameter.

- Plot B: With the actual parameters found by Veronig and Brown except with  $n_p$  reduced by a factor of 25 so that the loop is collisionally thin above about 10 keV but with the footpoints hidden (limb occulted) so there is no cold thick target contribution. In this case the thermal emission is also much reduced because  $EM = 2n_p^2 AL$  is down by a factor of 625. Somewhere between this and the first

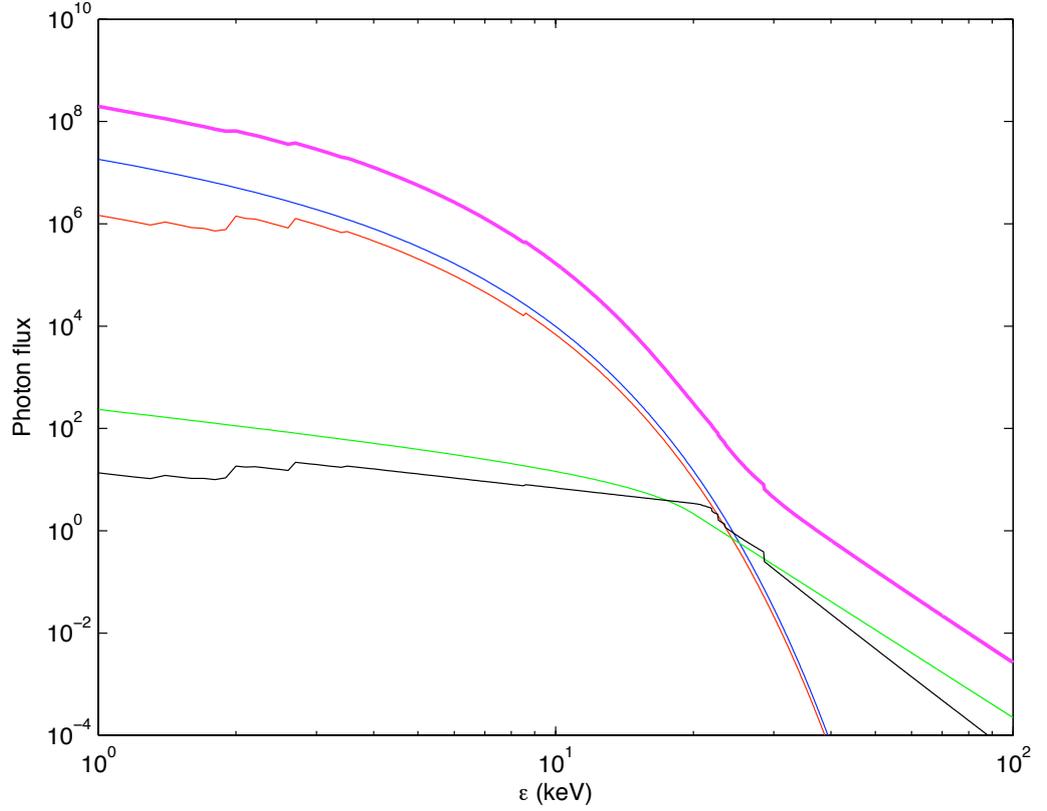


Figure 2.8: Spectral components for a resolution of 0.1 keV. The f-b edges of all elements involved are clearly noticeable. The parameters are  $T = 19.6$  MK,  $E_c = 15$  keV and  $\delta = 6.7$ . This plot can be compared to Plot A of Figure 2.7, which has the same parameters but for 1 keV binning resolution.

Table 2.3: Significance of Fe 25+ and 26+ at 4 different temperatures

$T$ (MK)	$A_Z$ (Fe 25+)	$A_Z$ (Fe 26+)	$\zeta_{RZ}$ (Fe 25+)	$\zeta_{RZ}$ (Fe 26+)	$R_{TR}$ (Fe 25+)	$R_{TR}$ (Fe 26+)
20	$2.4 \times 10^{-7}$	$2.5 \times 10^{-10}$	$4.7 \times 10^{-2}$	$1.1 \times 10^{-4}$	$2.5 \times 10^{-2}$	$8.7 \times 10^{-5}$
30	$3.5 \times 10^{-6}$	$4.5 \times 10^{-8}$	0.68	$2.1 \times 10^{-2}$	0.37	$1.6 \times 10^{-2}$
40	$6.7 \times 10^{-6}$	$2.8 \times 10^{-7}$	1.3	0.13	0.70	$9.8 \times 10^{-2}$
50	$1.7 \times 10^{-5}$	$1.3 \times 10^{-6}$	3.3	0.59	1.8	0.45

case should be the optimum condition for seeing maximum f-b contribution.

- If we include emission from cold thick-target footpoints, the non-thermal f-f emission dominates the spectrum and swamps any or all f-b features.

In Figure 2.1 we see that at no point is NTR greater than NTB for the hot case, but still can be a significant contributor to the total non-thermal flux in and around

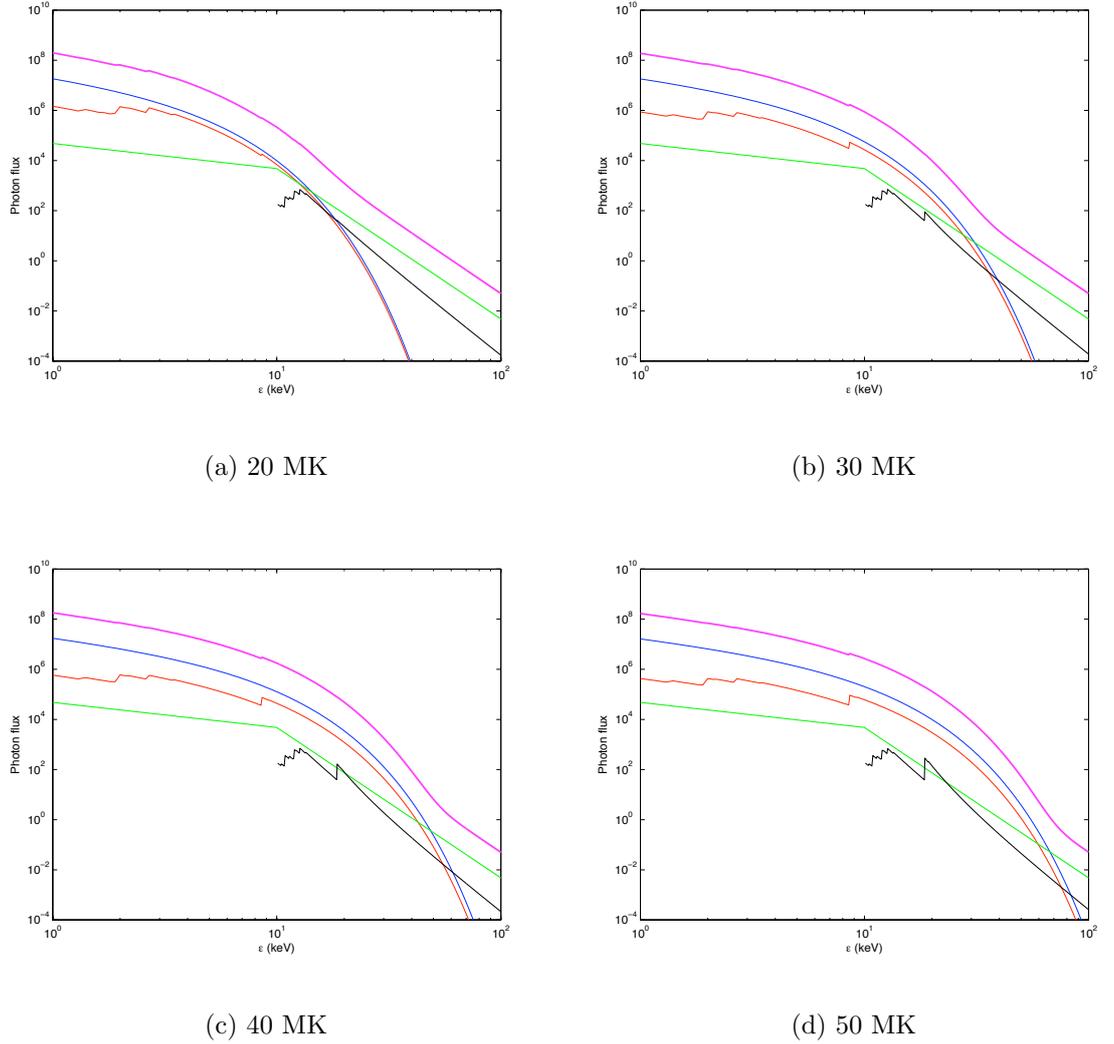


Figure 2.9: The thin-target spectra for 4 different temperatures with  $f_c = 0.01$ . The colour-coding is the same as in Figure 2.7

the cut-off energy. However, the thermal emission at these energies is quite likely to swamp the detectability of the non-thermal edges for these temperatures. Nevertheless, it is important to note that all the major edges seem to occur about 2-4 keV above the cut-off and that is because the  $Z^2\chi/n^2$  value for the important species of Fe, namely Fe 23+ and Fe 24+ in this case, is about 2-3 keV. So these important Fe edges coexist with edges from other elements like Si and O that have significant values of  $\zeta_{RZ}$ , and are almost fully ionised, hence  $n_{\min} = 1$  for them. Also added is Figure 2.3, which shows the ratio for an extreme case, i.e.  $\delta = 7$  and  $E_c = 5$  keV for temperatures of 20, 30, 40 and 50 MK, with each incremental increase in  $T$  producing an increase in the non-thermal f-b:f-f ratio at the Fe 25+ and 26+ edge, which occurs at  $E_c + V \approx 14$

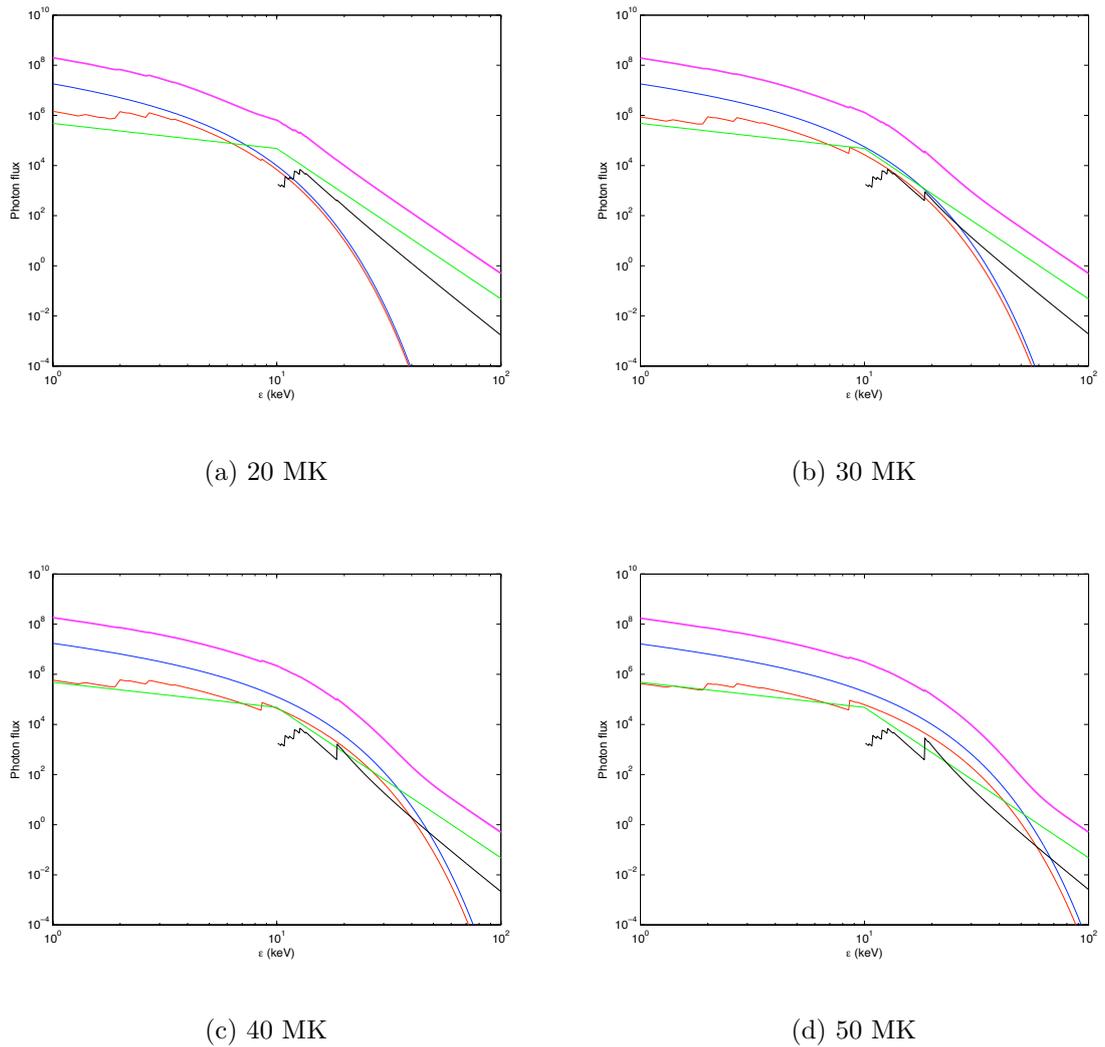


Figure 2.10: The revised thin-target spectra for 4 different temperatures with  $f_c = 0.1$ . The colour-coding is the same as in Figure 2.7

keV.

This is significant since Fe 25+ and Fe 26+ start becoming important only for temperatures above 30 MK (Table 2.3). Hence we decided to look at our model spectra for each of the four different temperatures while keeping  $E_c = 10$  keV and  $\delta = 5$  fixed. And then varying  $f_c = 0.01, 0.1$ , we produced the plots in Figures 2.9 and 2.10 for the thin-target case. However, although it is true that the Fe 25+/26+ edge becomes more prominent for the higher temperatures and non-thermal f-b can be greater than f-f at these edges, higher temperatures also mean an increased thermal emission (where thermal f-b becomes less and less important than thermal f-f), which invariably swamp these non-thermal features. Hence, it is quite obvious that if non-thermal f-b were to

be observed, the optimum temperature would be about 20-30 MK. The same is true for thermal f-b edges (Landi 2007).

The ‘ankle’ apparent in Figure 2.7 A,B at around 40 keV due to the transition from a f-b to a f-f dominated spectrum (cf. §2.3 and Figure 2.4) is rarely seen in data but may be present in some events (Conway et al. 2003). A statistical survey of a large sample of events should shed light on conditions where non-thermal f-b is important. Also note that an ‘ankle’ is present at the transition from a thermal- to a non-thermal-dominated spectrum. The position of this ankle depends on the plasma temperature and may interfere with the f-b to f-f ‘ankle’, which depends mainly on the  $E_c$  parameter. Hence, although for certain parametric conditions one may be able to notice two separate ‘ankles’, if  $E_c$  is low and  $T$  is high, the ‘ankles’ may occur at similar  $\epsilon$  and may not be distinguishable in real data. But it is important to distinguish between those events where edges may not be present at all and those where they may be present but not detectable because present-day instruments are not sensitive enough.

## 2.6 Inference of electron spectra by inversion

We note again that, since even the thin target  $j_B$  involves an integral over  $E$  while  $j_R$  does not, any sharp features in  $F(E)$  would be smoothed out in the bremsstrahlung contribution to the photon spectrum but not in the recombination contribution. Consequently, an important way to study the effect of including f-b on the required properties of  $F(E)$  is to consider it as an inverse problem (Craig & Brown 1986) to infer  $F(E)$  from observed  $j(\epsilon)$ . Here we consider the following experiment for the thin target case (thick target and thermal cases always involve even greater error magnification - Brown & Emslie (1988)): generate the total  $j(\epsilon)$  including f-b as well as f-f from a specified  $F_1(E)$  and evaluate the  $F_2(E)$  which would be erroneously inferred by solving the inverse problem ignoring the presence of the f-b term, as is currently done in all HXR data analysis, whether by inversion or forward fitting.

By Equations 2.4 and 2.11 the total f-f + f-b emission spectrum  $dJ/d\epsilon$  from a homogeneous volume  $V$  can be written

$$H(\epsilon) = \int_{\epsilon}^{\infty} G(E)dE + D \sum_{Z_{eff} \leq n\sqrt{\epsilon/\chi}} \sum_{n \geq n_{\min}} p_n \frac{1}{n^3} Z_{eff}^4 A_{Z_{eff}} G(\epsilon - V_{Z_{eff}}/n^2), \quad (2.31)$$

where

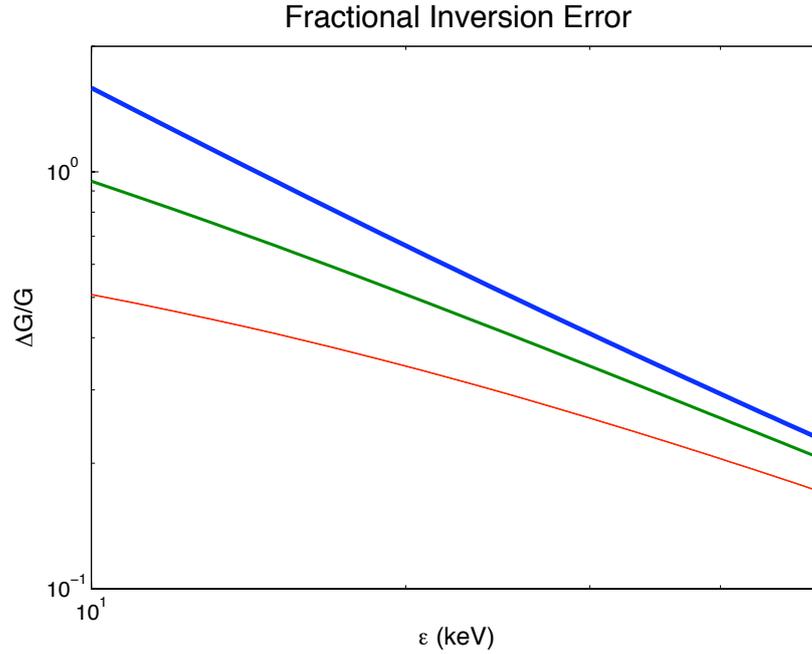


Figure 2.11: Fractional error ( $\Delta G/G$ ) in  $G$  (Equation 2.36) as discussed in §2.6 for  $E_* = 5$  (blue), 10 (green) and 20 (red) keV for a shifted power-law due to inference of  $G$  from  $H$  ignoring the presence of recombination.

$$H(\epsilon) = \frac{3}{8\alpha r_e^2} \frac{1}{\zeta_B m_e c^2 n_p V} \epsilon \frac{dJ}{d\epsilon}; \quad G(E) = F(E)/E \quad (2.32)$$

and  $D$  is as given in Equation 2.39. If we ignore the second (recombination) term in Equation 2.31, as has always been done in the past, for the Kramers f-f term, the inverse is just (Brown & Emslie 1988)

$$G(\epsilon) = -H'(E). \quad (2.33)$$

The neglect of the second term can be thought of as an ‘error’  $\Delta H$  in our data and if we apply inversion formula (2.33) to this ‘data’, ignoring the recombination ‘error’ we get a resulting error  $\Delta G$  in the inferred  $G$  given by

$$\begin{aligned} \Delta G(E) &= \frac{F_2(E) - F_1(E)}{E} \\ &= -D \sum_{Z_{eff} \leq n\sqrt{\epsilon/\chi}} \sum_{n \geq n_{min}} p_n \frac{1}{n^3} Z_{eff}^4 A_{Z_{eff}} G'(E - V_{Z_{eff}}/n^2). \end{aligned} \quad (2.34)$$

It is at once clear that any sharp change in  $j(\epsilon)$  i.e. in  $H(E)$ , such as the presence of f-b edges, however small, can have a very large effect on the inferred  $F_2(E)$ . (If the

inverse problem is addressed for more realistic smoother forms of f-f cross section than Kramers, the ‘error magnification’ is in general even larger - [Brown & Emslie \(1988\)](#); [Piana et al. \(1999\)](#)). For a power law  $F$  with cut off around say 20 keV, analytically speaking this expression gives infinite negatives in  $\Delta G(E)$  at the spectral edges around 30 keV (for Fe). However when smoothed over a few keV and added to the f-f term the result would be a ‘wobble’ in the  $F(E)$  solution in the 30-40 keV range. This is just where enigmatic features have been reported in some RHESSI spectra and variously attributed to the effects of photospheric albedo ([Kontar et al. 2006](#)), possibly pulse pile up [Piana et al. \(2003\)](#), or a high value of  $E_c$  ([Zhang & Huang 2004](#)).

Another case providing insight is that of a smooth shifted power-law  $G(E) = A(E + E_*)^{-\delta-1}$ , which has no edges though the corresponding  $F(E)$  has a smooth peak at  $E = E_*/\delta$ . In this case the fractional error in  $G$  due to applying Equation 2.33 ignoring the recombination term can be expressed as

$$\frac{\Delta G(E)}{G(E)} = (\delta + 1) \frac{D}{E + E_*} \sum_{Z_{eff}} \sum_{n \geq n_{min}} p_n \frac{1}{n^3} Z_{eff}^4 A_{Z_{eff}} \left[ \frac{1}{1 - V_{Z_{eff}}/n^2(E + E_*)} \right]^{\delta+2}, \quad (2.35)$$

where each term in the  $Z_{eff}$  sum is zero for  $E < V_{Z_{eff}}/n^2 = Z_{eff}^2 \chi/n^2$ .

In the case of recombination onto Fe XXV alone (hot plasma), this gives for  $\delta = 5$ ,

$$\frac{\Delta G}{G} \approx \frac{10 \text{ keV}}{E + E_*} [1 - 7 \text{ keV}/4(E + E_*)]^{-7}, \quad (2.36)$$

which is shown in Figure 2.11 for  $E_* = 5, 10, 20$  keV. Evidently errors due to neglect of recombination can be large at low  $E$ . The reason is that the  $Z_{eff}$  recombination contribution to the bremsstrahlung solution for  $G(E)$  at  $E$  comes from the slope of  $G$ , and not just  $G$  itself and at  $E - V_{Z_{eff}}/n^2$  not at  $E$ . Figure 2.11 is similar to Figure 2.13 because  $F_2/F_1 = G_2/G_1 = 1 + \Delta G/G_1$ .

This error has serious consequences for past analyses of HXR flare spectra, at least in cases where a significant hot dense coronal loop is involved. For example, the f-b emission spectrum is most important at lower energies (5-30 keV or so), depending on the plasma temperature  $T$  and low energy electron cut-off or roll-over  $E_c, E_*$  and is steeper than the free-free. This will offset some of the spectral flattening caused around such energies by photospheric albedo ([Kontar et al. 2005](#)) resulting in underestimation of the albedo contribution and hence of the downward beaming of the fast electrons. This fact would weaken the finding of [Kontar & Brown \(2006\)](#) that the electrons are near isotropic, in contradiction to the usual thick target description, but for the fact

that the flares they used had rather hard spectra and substantial footpoint emission - conditions where the f-b correction should be rather small. Nevertheless it illustrates that care is needed to ensure f-b emission is properly considered.

Finally, recognising the presence of the f-b contribution, one can in fact convert integral (2.31) into a differential/functional equation for  $F(E)$  by differentiation, namely

$$G(E) - D \sum_{Z_{eff} \geq n(E/\chi)^{1/2}} \sum_{n \geq n_{\min}} p_n \frac{1}{n^3} A_{Z_{eff}} Z_{eff}^4 G'(E - Z_{eff}^2 \chi / n^2) = -H'(E), \quad (2.37)$$

which is a wholly new class of functional equation in need of exploration.

## 2.7 Recombination efficiency and different forms of F(E)

The question of efficiency has been handled in more detail in the following chapter, but in this section we will briefly introduce how the problem was originally tackled by us and this section can be thought of as a prelude to the efficiency implications we discuss in the following chapter.

### 2.7.1 Comparison of the efficiency of f-b versus f-f HXR yield

In §2.3, we predicted the  $j_B, j_R$  from a power-law  $F(E)$  and found that the  $j_R$  contribution could sometimes be more important than  $j_B$ . It is of interest therefore, to consider the following question. If one observes a power-law  $j(\epsilon) \propto \epsilon^{-\gamma}$  above some  $\epsilon \geq V_Z$ , what electron flux  $F_R(E)$  would be needed to generate it in a plasma of solar abundances *purely* by non-thermal electron recombination on ion  $Z_{eff}$  as compared with the  $F_B(E)$  required to do so *purely* by f-f bremsstrahlung? If we write, from Equation 2.15, the latter as  $F_B(E) = CE^{-\gamma+1}$  then the former has to be, by Equation 2.16,

$$F_R(E) = C(E + V_{Z_{eff}}/n^2)^{-\gamma} / D Z_{eff}^4 A_{Z_{eff}}, \quad (2.38)$$

where

$$D = \frac{2\pi\chi}{\sqrt{3}\zeta_B} \approx 0.04keV \quad (2.39)$$

and the ratio measuring recombination efficiency relative to bremsstrahlung, which is the inversed of the fluxes required, is

$$\frac{F_B(E)}{F_R(E)} = \gamma Z_{eff}^4 A_{Z_{eff}} \frac{D}{E} [1 + V_{Z_{eff}}/n^2 E]^\gamma, \quad (2.40)$$

which we show in Figure 2.12 for  $\gamma = 5$  in terms of each of the dominant f-b contributions from fully ionised O, Mg, Si and Fe respectively while the f-f is for all elements. Evidently, non-thermal recombination could be dominant over bremsstrahlung up to a few 10s of keV as the most efficient HXR source if the electrons are emitted entirely in a plasma hot enough ( $T \approx 25\text{MK}$ ) for elements up to Fe 25+ to be ionised but is not as significant at lower temperatures.

In terms of the total required electron fluxes  $F_{R1}, F_{B1}$  above energy  $E_1$ , the ratio is

$$\begin{aligned} \Psi &= \frac{\gamma - 1}{\gamma - 2} Z_{eff}^4 A_{Z_{eff}} \frac{D}{E_1} [1 + V_{Z_{eff}}/n^2 E_1]^{\gamma-1} \\ &\approx 0.02 Z_{eff}^4 A_{Z_{eff}} \frac{10 \text{ keV}}{E_1} [1 + V_{Z_{eff}}/n^2 E_1]^{\gamma-1}, \end{aligned} \quad (2.41)$$

which is about 10 for Fe, 0.25 for Si and 0.1 for Mg and O at  $E_1 = 10 \text{ keV}$ .

At higher electron energies ( $E \geq \approx 17 \text{ keV}$ ), O becomes more efficient than Mg, as can be seen in Figure 2.12, because of the combined effects of the  $A_Z Z^4$  factor and the term containing  $V_Z$ .

### 2.7.2 Ratio of $j_R$ to $j_B$ for an example of a smooth $F(E)$ with no cut-off

All of the above results are for  $F(E)$  with a sharp cut off  $E_c$ . To illustrate how the appearance of  $j(\epsilon)$  is modified by inclusion of f-b as well as f-f for a smooth  $F(E)$ , a simple case to evaluate is  $F(E) \propto E(E + E_*)^{-\delta-1}$ , which behaves as  $E^{-\delta}$  at  $E \gg E_*$  but has a smooth roll-over at  $E_*/\delta$ . It is simple to show that the resulting  $j_B(\epsilon) \propto (E + E_*)^{-\delta}/\delta$  for f-f alone and that the ratio of f-b to f-f in this case is, for ion  $Z_{eff}$  alone,

$$\Psi_{smooth} = \frac{D \zeta_{Z_{eff}}}{\epsilon + E_*} \left[ 1 - \frac{Z_{eff}^2}{n^2(\epsilon + E_*)} \right]^{-\delta-1}, \quad (2.42)$$

which is shown in Figure 2.13 for  $\delta = 5$ ,  $Z_{eff} = 23.77$  and  $E_* = 5, 10, 20 \text{ keV}$ . We see again that  $\Psi_{smooth}$  is largest for large  $\delta$  and for small  $E_*$ .

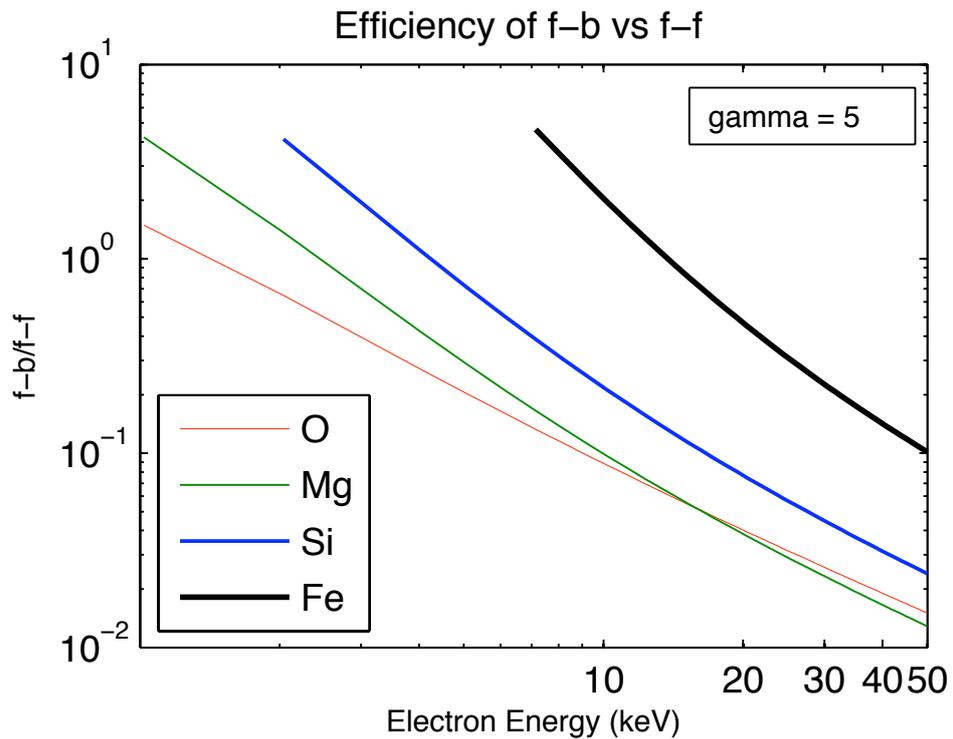


Figure 2.12: The f-b electron efficiency compared to f-f for the 4 elements discussed in §2.7 (Equation 2.41) assuming full ionisation. It is evident from the graph that, if present, highly ionised Fe is the most efficient source of f-b HXR in terms of the  $F(E)$  needed followed by Si, O and Mg. Note that the curves for each element begin at their ionisation potential since f-b emission below the ionisation potential is zero.

## 2.8 Discussion and conclusions

It is clear from our findings that ignoring non-thermal f-b contribution as negligible, as has been done in the past, is erroneous. Even if we ignore coronal enhancement of element abundances, and use photospheric abundances, f-b contribution can be significant. In certain flaring regions, especially in dense-hot coronal sources or occulted loop-top events, fast electron recombination can be of importance in analysing data properly and in inferring electron spectra and energy budgets. It can have a significant influence on inferred electron spectra both as an inverse problem and also in forward fitting parameters, including the important potential to find and evaluate low-energy electron cut-offs, which are vital to flare energy budgets. While incorporating f-b into spectral fitting procedures will make it considerably more complicated, an advantage is that the f-b, unlike the f-f, contribution retains its  $J(\epsilon)$  signatures of any sharp features in  $F(E)$ . As an example, in Figure 2.14 we show the inferred electron spectra with and without including f-b emission for the photon flux shown in Figure 2.8. It

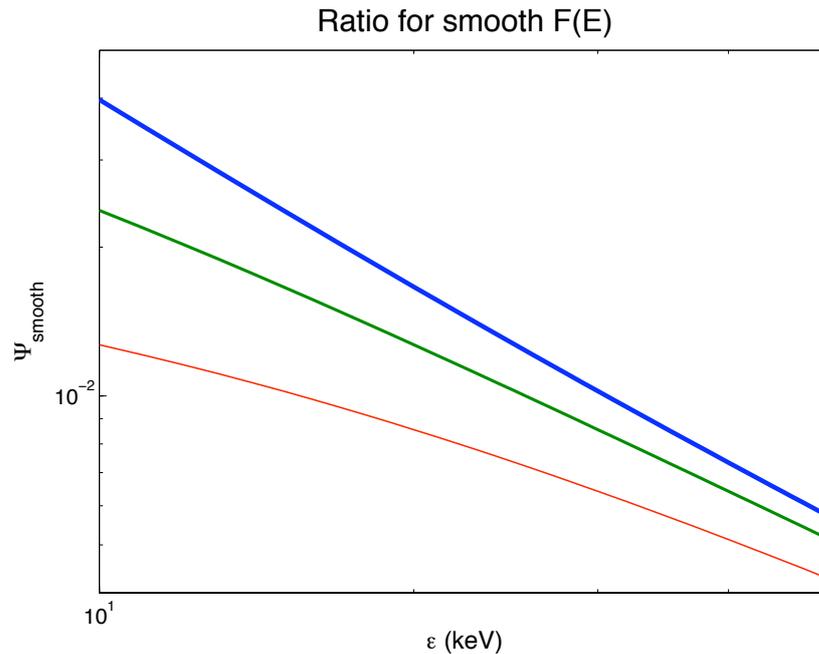


Figure 2.13: The  $\Psi_{smooth}$  as discussed in §2.7 and Equation 2.41. It is the ratio of  $j_R$  to  $j_B$  for the smooth  $F(E) \propto E(E + E_*)^{-\delta-1}$  for  $E_* = 5$  (blue), 10 (green) and 20 (red) keV.

is clear that excluding the f-b component produces prominent ‘bumps’ in the electron spectrum (green curve), whereas the true injected electron spectrum looks similar to the black curve, which is the inferred spectrum if f-b emission is included. This shows that ignoring even the small edges seen in the photon spectrum produces much larger features in the inferred electron spectrum.

A possible major consequence of the low energy f-b contribution is that, to fit an actual photon spectrum, less electrons are needed, than in f-f only modelling, at the low  $E$  end, which is where most of the power in  $F(E)$  lies. For example, if we consider the case  $\delta = 5$ ,  $E_c = 10$  keV and ionisation up to Fe XXVI and above, then we see from Figures 2.1 and 2.2 that inclusion of f-b increases  $j$  by a factor of 2-10 in the 15-20 keV range for  $\delta = 3-5$ . Thus, to get a prescribed  $j$  in that range we need only 10 – 50% as many electrons as inferred from f-f emission only.

We also note that the importance of non-thermal f-b emission is greatest when non-thermal electrons are present at low  $E$  and with large  $\delta$  such as in microflares with ‘hard’ XRs in the few to ten keV range (Krucker et al. 2002). Such low energy electrons have short collisional mean free paths and so are more likely to emit mainly in hot coronal regions, if accelerated there. Microflares are possibly therefore important cases for inclusion of f-b.

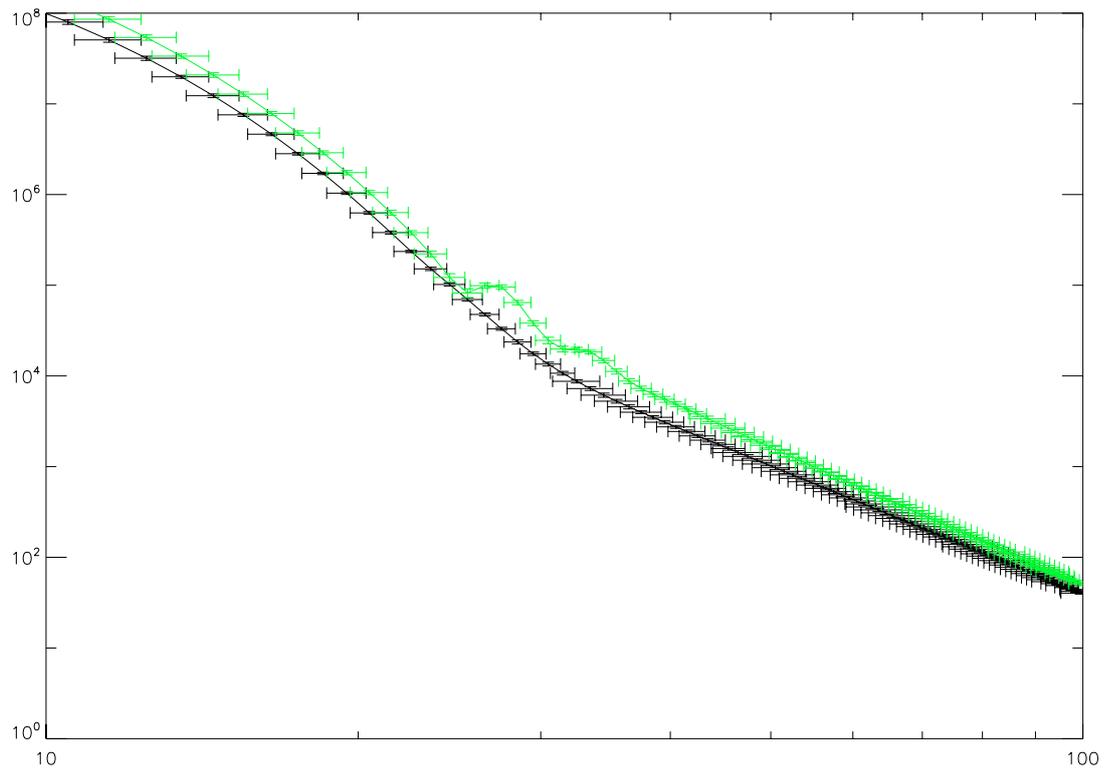


Figure 2.14: Inferred electron spectra using the photon flux as shown in Figure 2.8 with no noise included. The  $x$ -axis is electron energy in keV and the  $y$ -axis is electron flux. The black curve denotes the inferred spectrum if f-b emission is included whereas the green curve shows the inferred electron spectrum and the effects were f-b emission to be ignored.

Before we conduct any precise fitting of  $F(E)$ , involving the f-b contribution, to real data (e.g. from RHESSI) and include it in software packages it will be important to include, for both f-b and f-f, more accurate cross-sections with Gaunt factors etc. and ionisation fractions as functions of plasma temperature. By doing this, it will be possible to show, for certain events, how important recombination is and to improve our understanding of electron spectra and their roles in flares. However, our Kramers results already bring out the fact that recombination should not be ignored in the future, and that it may be invaluable in some cases as a diagnostic of the presence or otherwise of electron spectral features.

# Chapter 3

## Major Contribution of Non-thermal Recombination to Flare Hard X-rays

### 3.1 Introduction

In the previous chapter, we introduced the significance of non-thermal recombination (NTR) and deemed it an important source of HXR for certain flare conditions. In this chapter, we go a bit further to actually determine which type of flares may actually be dominated by NTR. The contents of this chapter closely follow our paper [Brown & Mallik \(2009\)](#), with amended results that have been obtained since the paper's publication. The paper made an error regarding the interpretation of some atomic physics that has been corrected here.

Just to summarise the contents of the last chapter, it was shown that, in the hottest regions of flare plasma, non-thermal HXR emission in the few deka-keV range from non-thermal electrons by recombination onto heavy ions (especially Fe) can be a substantial fraction of bremsstrahlung (NTB), contrary to earlier assumptions. Though non-negligible even at temperatures  $T$  down to  $10^6$  K, importance of such NTR radiation needs  $T > 10$  MK in order for  $\text{Fe}^{22+}$  ions and above to be plentiful. Furthermore, even for an accelerated fraction of only 0.01, the total hot plasma thermal emission begins to exceed NTR only for  $T > 25$  MK. The relative NTR contribution is greatest when the electron flux spectrum is steep and extends to low energies. Thus, in proper modelling of hot HXR sources, inclusion of NTR as well as NTB can be important as it may reduce HXR electron number requirements, by a few times in some cases. This alleviates problems of electron acceleration efficiency.

Even some chromospheric footpoint HXR sources may have substantial NTR con-

tribution if the hot Soft X-Ray (SXR) footpoint plasma there contains fast electrons. Only a small fraction of the plasma emission measure observed in SXR footpoints need be in the form of non-thermals to provide the necessary HXR emission measure. Compared with the Standard Cold Thick Target (bremsstrahlung) Model (CTTM), such a scenario would give fast electrons a lesser role in the flare energy budget, and help solve various problems with CTTM.

We showed that, contrary to previous wisdom (based on e.g. [Elwert & Haug 1970, 1971](#); [Korchak 1967, 1971](#)), flare deka-keV HXRs from non-thermal electrons are not always solely produced by bremsstrahlung. For hot enough ( $> 10$  MK) plasma regions, few-deka-keV HXR emission can have a large contribution from NTR onto highly ionised heavy elements, especially iron (Fe), since : (a) recombination cross sections increase with ion charge  $Z$  like  $Z^4$  ( $\approx 4 \times 10^4$  for fully ionised Fe); (b) while, for hydrogen alone,  $A_Z Z^4 = 1$  ( $A_Z$  is the abundance of ionic species  $Z$ ), the solar (especially coronal) Fe  $A_Z$  is so high ([Feldman 1992](#)) that for Fe alone  $A_Z Z^4 \approx 26$  for iron ionised up to FeXXV and  $\approx 6$  for up to FeXXIV, and  $\approx 2$  even for up to Fe XXIII; (c) bremsstrahlung photons of energy  $\epsilon$  are emitted only by electrons of energy  $E \geq \epsilon$  while photons from recombination with ionisation potential  $V$  are emitted by electrons of  $E = \epsilon - V/n^2$ . The corresponding flux  $F(E - \epsilon)$  of emitting electrons in a power law spectrum,  $E^{-\delta}$ , is thus larger for NTR than the flux  $F(E \geq \epsilon)$  for NTB by a factor exceeding  $(1 - V/n^2\epsilon)^{-\delta}$ , which can be very large at low  $\epsilon$  ( $\rightarrow V$ ), especially for large  $\delta$ . Consequently, for hot sources, including the NTR contribution has major implications for the magnitude and form of the inferred  $F(E)$  and hence for fast electron number and energy budgets, and for acceleration mechanisms. Here we examine under what circumstances NTR may dominate deka-keV HXR emission and by how much this reduces the fast electron number and power requirements. Our results are directly relevant to hot coronal HXR sources ([Veronig & Brown 2004](#); [Krucker et al. 2007, 2008b](#)). In §3.4, we show that they are also relevant to footpoint HXR sources at chromosphere altitudes ([Mrozek & Tomczak 2004](#)), if fast electrons emit NTR in the hot chromospheric plasma component of the SXR footpoint emission. Such a scenario would require fewer fast electrons than the standard CTTM (NTB only model) and could solve some of the problems it poses.

## 3.2 Contribution to deka-keV emission

Since this chapter uses the main ideas from Chapter 2, we will not repeat the details of all the equations used there. Instead, we will abbreviate some of the equations

used there to suit the needs of this chapter. So for relevant background and the basic equations used, please see the previous chapter.

Using Kramers' (Kramers 1923) hydrogenic ion cross section approximations, with effective values  $Z = Z_{eff}$  (which allow for screening and other non-hydrogenic effects), and for a power-law mean source electron flux energy spectrum of the form

$$\bar{F}(E) = (\delta - 1) \frac{F_c}{E_c} \left[ \frac{E}{E_c} \right]^{-\delta}; \quad E \geq E_c \quad (3.1)$$

cut off below  $E = E_c$ , we found the total thin target non-thermal emission rate (NTB + NTR) spectrum (photons per sec per unit  $\epsilon$ ) to be, with  $v_c = (2E_c/m_e)^{1/2}$ ,

$$J(\epsilon) = \frac{\delta - 1}{\delta} \frac{C}{\epsilon} \frac{v_c}{E_c} \zeta_B EM f_c \left( \frac{\epsilon}{E_c} \right)^{-\delta} \left[ 1 + \frac{2\pi\delta}{\sqrt{3}\zeta_B} \frac{\chi}{\epsilon} \sum_{Z_{eff}} \sum_{n \geq n_{min}} p_n \frac{1}{n^3} \zeta_{RZ_{eff}} \left( 1 - \frac{V_{Z_{eff}}/n^2}{\epsilon} \right)^{-\delta-1} \right] \quad (3.2)$$

where the first and second terms are  $J_{NTB}$  and  $J_{NTR}$  respectively.  $C = \alpha m_e c^2 r_e^2 / 3$  with  $m_e, r_e, c, \alpha$  the electron mass and radius, speed of light, and fine structure constant respectively.  $EM = \int_{\mathcal{V}} n_p^2 d\mathcal{V}$  is the total plasma emission measure in volume  $\mathcal{V}$  and  $EM_c = f_c EM$  is the total non-thermal emission measure at  $E > E_c$ .  $n_p$  is the total hydrogen density and we have expressed the total flux  $F_c$  of fast electrons at  $E > E_c$  in terms of their fractional density

$$f_c = n_c/n_p; \quad n_c = \int_0^\infty \frac{F(E)}{v(E)} dE. \quad (3.3)$$

$\chi$  and  $V_{Z_{eff}} = Z_{eff}^2 \chi$  are the ionisation potentials for H and for ions of charge  $Z_{eff}e$ . For element and ion abundances  $A_Z$  and  $A_{Z_{eff}}$  (Feldman 1992; Arnaud & Raymond 1992; Brown & Mallik 2008, (Chapter 3), the heavy element correction factor for bremsstrahlung is  $\zeta_B = \sum_Z A_Z Z^2 \approx 1.6$ , which is independent of the ionisation state while the recombination corrections for each ion  $\zeta_{RZ_{eff}} = A_{Z_{eff}} Z_{eff}^4$  are sensitive to it, and hence to temperature. For typical hot flare plasma temperatures,  $T \approx 10 - 30$  MK, ions up to Fe XXV are important giving  $\zeta_R = \sum_{Z_{eff}} \zeta_{RZ_{eff}} \approx 50$  (for all elements and ions up to Fe XXV) and making the second term in Equation 3.2 important.

Note that, in Equation 3.2 and subsequent discussion, for simplicity we restrict ourselves to considering only  $\epsilon > E_c + V_{Z_{eff}}/n^2$  for all  $Z_{eff}$  so that no recombination edges (Chapter 2) are present. Even for Fe XXVI,  $V_{Z_{eff}} \approx 9$  keV so that if, for example  $E_c \approx 10$  keV, edge effects are absent at  $\epsilon \geq 19$  keV.

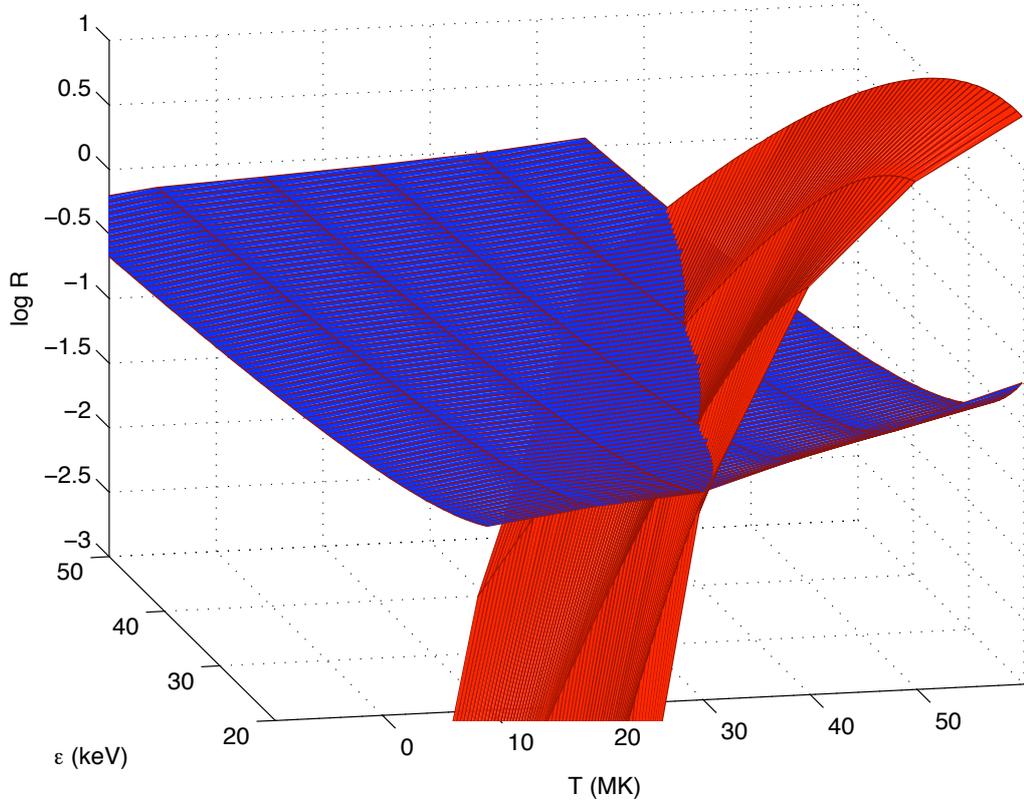


Figure 3.1: The ratios  $R_{NTR}(\epsilon, T)$  (blue) and  $R_{Th}(\epsilon, T)$  (red) versus  $\epsilon$  and  $T$  for  $\delta = 5$  and  $f_c = 0.01$

At very high  $T$ , of course, bremsstrahlung and recombination from *thermal* electrons become important sources of deka-keV HXR so we wish to know in what parameter regimes NTR may dominate NTB while not swamped by the *total* thermal emissions  $J_{Th}(\epsilon) = J_{ThB} + J_{ThR}$ , i.e. where the ratio  $R_{NTR}(\epsilon, T) = J_{NTR}/J_{NTB} > 1$  while the ratio  $R_{Th}(\epsilon, T) = J_{Th}/J_{NTB} < 1$ . The bremsstrahlung and recombination contributions to thermal emission  $J_{Th}(\epsilon)$  are given by Equations 2.20 and 2.21. The surfaces  $R_{NTR}(\epsilon, T)$ ,  $R_{Th}(\epsilon, T)$  are shown in Figure 3.1 in blue and red respectively for  $\delta = 5$  and  $f_c = 0.01$ , with Fe ionisation fractions from Arnaud & Raymond (1992). Even for this quite small  $f_c$ , over a large range of temperatures (4–25 MK) and photon energies ( $> 20$  keV),  $R_{NTR}$  dominates  $R_{Th}$ , so NTR is well above thermal emission so long as  $T < 30$  MK or so. Also,  $J_{NTR}$  exceeds 10% of  $J_{NTB}$  over most of the  $\epsilon, T$  domain plotted, 20% in the 20–30 keV range and by 30–50% for the higher  $T$  end.

Regimes of relative NTR importance in the  $\epsilon, T$  plane are shown in Figures 3.2, 3.3 and 3.4 as distinct coloured areas where  $R_{NTR} \geq 0.05, 0.1, 0.5$  etc., i.e. where NTR

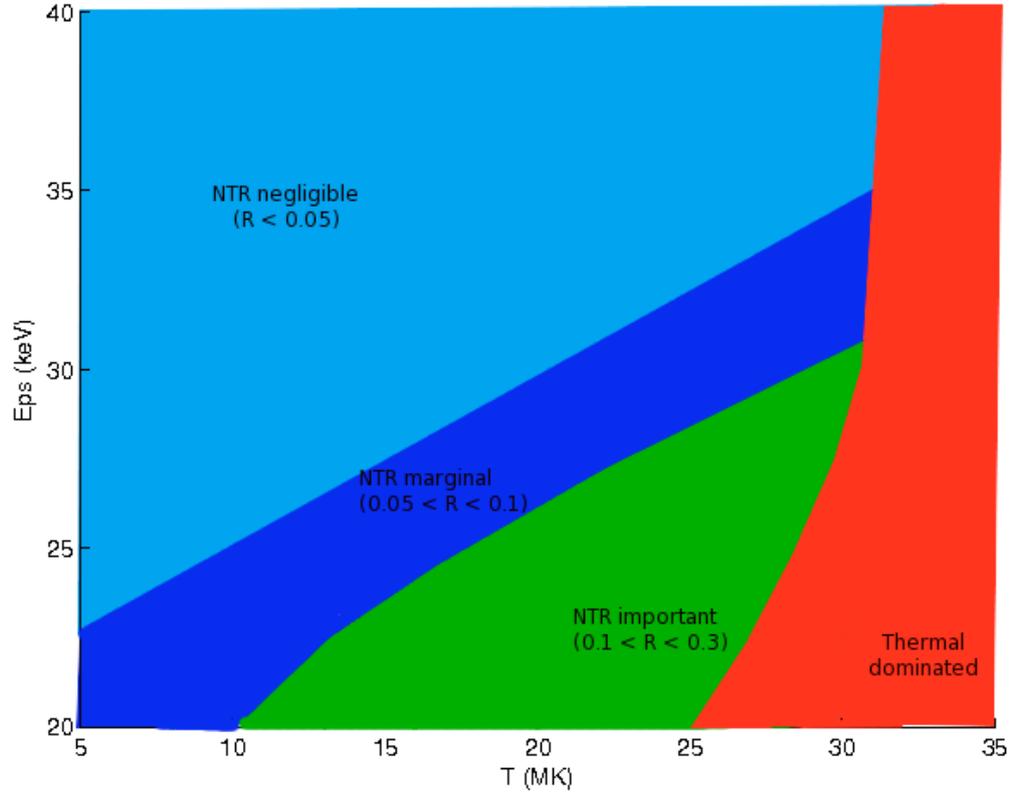


Figure 3.2: Regime plot showing the relevant areas of importance in the  $(\epsilon, T)$  domain for  $E_c = 10$  keV,  $\delta = 5$  and  $f_c = 0.01$

increases progressively from negligible to dominant, and also where thermal emission takes over ( $J_{NTR}(\epsilon, T)/J_{Th}(\epsilon, T) > 1$ ) at high  $T$  for  $f_c = 0.01, 0.1, 1$  respectively.

From Figures 3.2 and 3.3, it is evident that NTR is never dominant but still can be several 10s of percent of the NTB flux. These plots more or less summarise the general importance of NTR. Increasing  $f_c$  from 0.01 (Figure 3.2) to 0.1 (Figure 3.3) pushes the thermal-dominated line by about 5 MK to the right, but of course never changes the NTR:NTB ratio, as seen in Figure 3.3. However, in the  $(\epsilon, T)$  range of 20-30 keV and 20-30 MK respectively, NTR can still be significant enough whereby its inclusion in analysis packages becomes justifiable, even more so where  $f_c$  is larger as that increases the region of importance. As an extreme case, we also consider taking  $f_c = 1$  (as suggested by Krucker et al. (2008c)) and  $\delta = 7$  to see the NTR:NTB ratio in the  $(\epsilon, T)$  domain and this is shown in Figure 3.4. Krucker et al. (2008c) show that for the 2007 December 31 flare, which was highly occulted, non-thermal HXR emissions up to 100 keV are observed to originate entirely in the corona. Using RHESSI data,

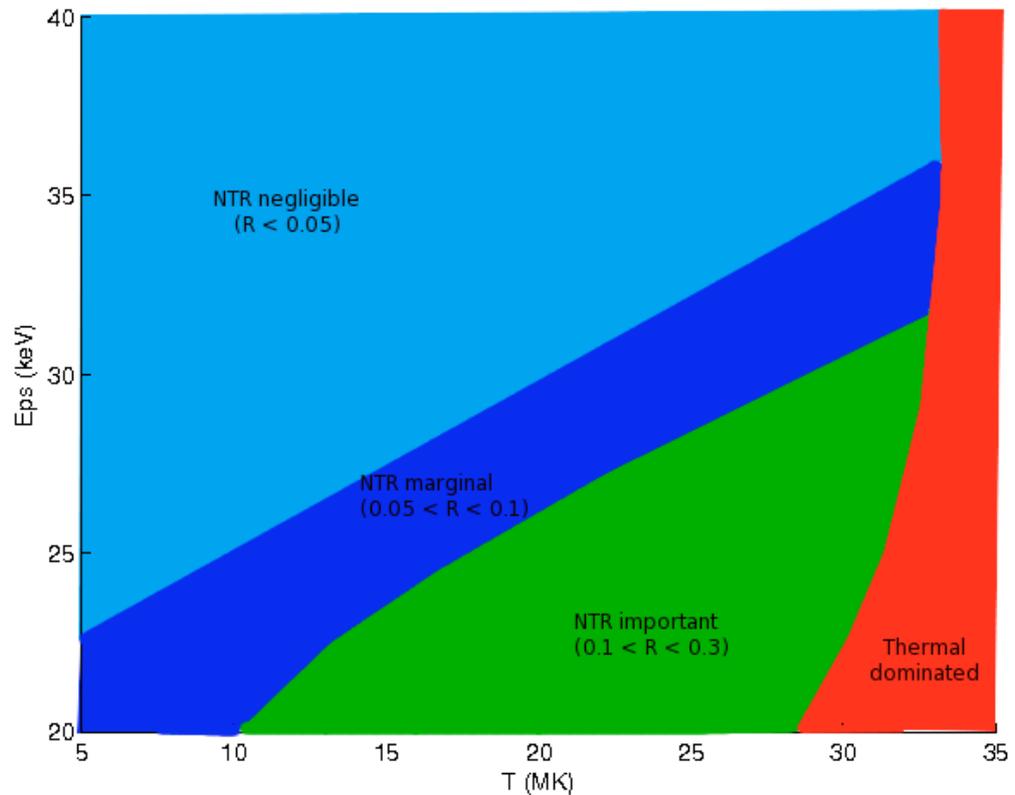


Figure 3.3: Regime plot showing the relevant areas of importance in the  $(\epsilon, T)$  domain for  $E_c = 10$  keV,  $\delta = 5$  and  $f_c = 0.1$ . Note that the area of importance is larger as we go to bigger values of  $f_c$ , as can be seen by comparing this plot to Figure 3.2

the authors were able to show that nearly *all* the electrons need to be accelerated to produce the observed emission; in other words,  $f_c \approx 1$ . As is evident, when  $f_c$  is large, you still get NTR dominated regions, albeit at rather high temperatures and a small  $\epsilon$  range. Nonetheless,  $R$  is substantial ( $> 0.5$ ) for a large  $(\epsilon, T)$  regime of 20-35 keV and 15-40 MK. This suggests that NTR may still play a more major role in efficiency and the electron number problem, but  $f_c$ ,  $T$  and  $\delta$  would have to be rather high for NTR to be so dominant. For a majority of flares, NTR is likely to have a 20-40% effect at the most.

Note that all the other parameters for all these plots have been maintained throughout as the following: density  $n = 1.5 \times 10^{11} \text{ cm}^{-3}$ , flare loop cross-section area  $A = 2 \times 10^{17} \text{ cm}^2$  and loop half-length  $L = 2.5 \times 10^9 \text{ cm}$ , mirroring those observed for the 2002 April 14 event (Veronig & Brown 2004).

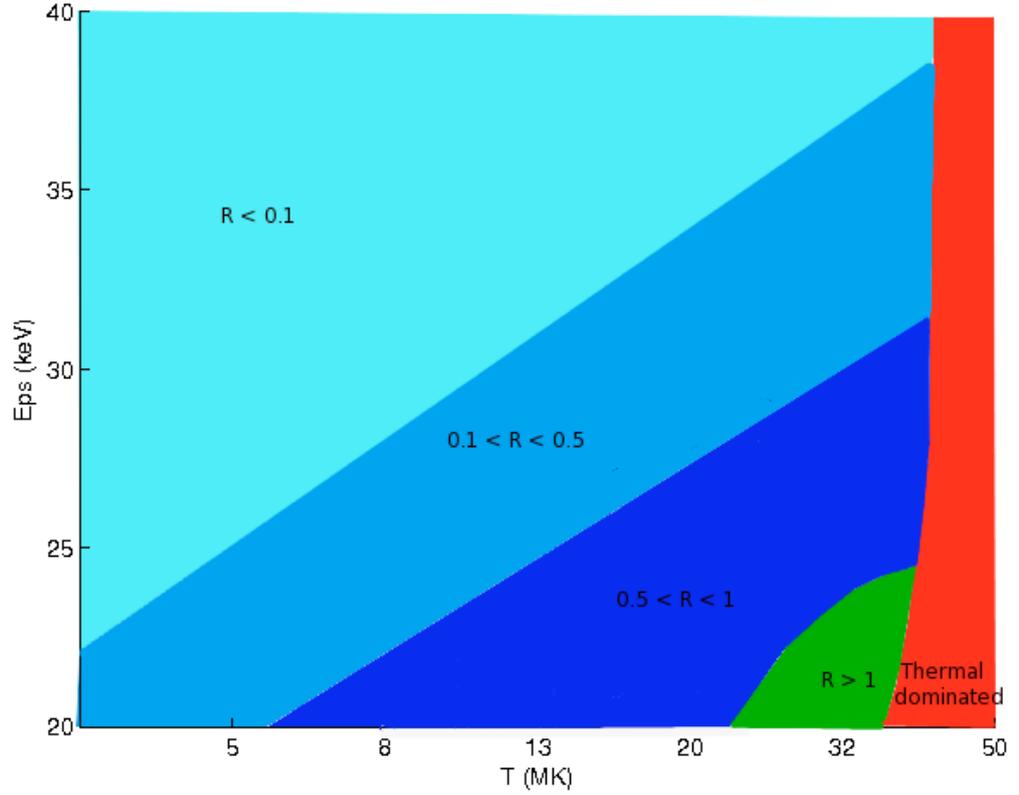


Figure 3.4: Regime plot showing the relevant areas of importance in the  $(\epsilon, T)$  domain for  $E_c = 10$  keV,  $\delta = 7$  and  $f_c = 1$ ; an extreme case. This shows that NTR may still be a dominant source of HXRs in some cases.

### 3.3 Implications of non-thermal recombination dominance

Both  $J_{NTB}(\epsilon)$  and  $J_{NTR}(\epsilon)$  will always contribute to the observed  $J(\epsilon)$ , but with  $J_{NTB}(\epsilon)$  dominating at large  $\epsilon$  due to the  $1/\epsilon$  factor in the ratio  $R_{NTR}$ . However, most of the fast electrons and their energy reside at low  $E$  and the following question is important. Suppose a specific non-thermal HXR event with roughly power-law  $J(\epsilon) \propto \epsilon^{-\gamma}$ , when interpreted (as is usually done) solely as NTB from electrons, requires a total non-thermal emission measure  $EM_{cB} = f_{cB} n^2 \mathcal{V}$  of non-thermal electrons at  $E > E_c$  with spectral index  $\delta_B = \gamma - 1$ . Then, by how much is the necessary  $EM_c$  reduced when we include NTR? (Note that, to produce the same photon  $\gamma$  as NTB at high  $\epsilon$ , NTR requires the electrons to have  $\delta_R = \gamma - 2 = \delta_B - 1$ . To remain consistent in our equations and figures, henceforth  $\delta$  refers to  $\delta_B$ .)

The total photon emission rate integral  $\int_{\epsilon_1}^{\infty} J(\epsilon) d\epsilon$  cannot be expressed analytically

for expression 3.2 so instead we consider the source *luminosity* above  $\epsilon_1$  viz,

$$L(\epsilon_1) = \int_{\epsilon_1}^{\infty} \epsilon J(\epsilon) d\epsilon, \quad (3.4)$$

which gives the following equations (recall our assumption that  $\epsilon \geq E_c + V_{Z_{eff}}/n^2, \forall Z_{eff}$ ):

$$L_{NTB}(\epsilon_1) = \frac{C_1}{\delta - 1} f_1 \epsilon_1^2; \quad ; \quad L_{NTR} = \frac{C_2}{\delta} \chi f_1 \epsilon_1 \sum_{Z_{eff}} \sum_{n \geq n_{\min}} p_n \frac{1}{n^3} \zeta_{RZ_{eff}} \left(1 - \frac{V_{Z_{eff}}}{n^2 \epsilon_1}\right)^{-\delta};$$

$$L_{Th}(\epsilon_1) = C_3 e^{-\epsilon_1/kT} (\chi^3 kT)^{1/2} \left[ \Gamma \frac{\chi}{kT} \sum_{Z_{eff}} \sum_{n \geq n_{\min}} p_n \frac{1}{n^3} \zeta_{RZ_{eff}} e^{V_{Z_{eff}}/n^2 kT} + 1 \right] \quad (3.5)$$

where

$$C_1 = \frac{\delta - 1/2}{\delta} \frac{8\alpha}{3} \zeta_B m_e c^2 r_e^2 \frac{v_1}{\epsilon_1^2}; \quad C_2 = (\delta - 1/2) \frac{32\pi}{\alpha\sqrt{3}} r_e^2 \frac{v_1 \chi}{\epsilon_1^2};$$

$$C_3 = \sqrt{\frac{8}{\pi m_e \chi}} \frac{16\alpha r_e^2}{3} \zeta_B \frac{m_e c^2}{\chi}; \quad \Gamma = \frac{2\pi}{\sqrt{3}\alpha^2 \zeta_B} \frac{\chi}{m_e c^2}. \quad (3.6)$$

We define the luminosity ratios as follows:

$$S_{NTR} = L_{NTR}/L_{NTB}; \quad S_{Th} = L_{NTR}/L_{Th}. \quad (3.7)$$

The quantity,

$$\Phi(T) = 1/(1 + S_{NTR}), \quad (3.8)$$

plotted in Figure 3.5, is roughly the factor by which the  $EM_c$  required to get a specific observed  $L_{HXR}(\epsilon \geq 20 \text{ keV})$  is reduced when both NTR and NTB are included, compared to NTB alone (for  $\delta = 5$ ). This clearly shows that even for temperatures as low as 5 – 10 MK, including NTR can increase HXR production efficiency by about a third, and at 25 MK, by a factor of about 2.

## 3.4 Applications: coronal and footpoint sources

### 3.4.1 Coronal HXR Sources

The parameters ( $EM, T, E_c, \delta$ ) used above were chosen to be similar to those in some typical coronal HXR events as reviewed and classified by Krucker et al. (2008b), e.g.

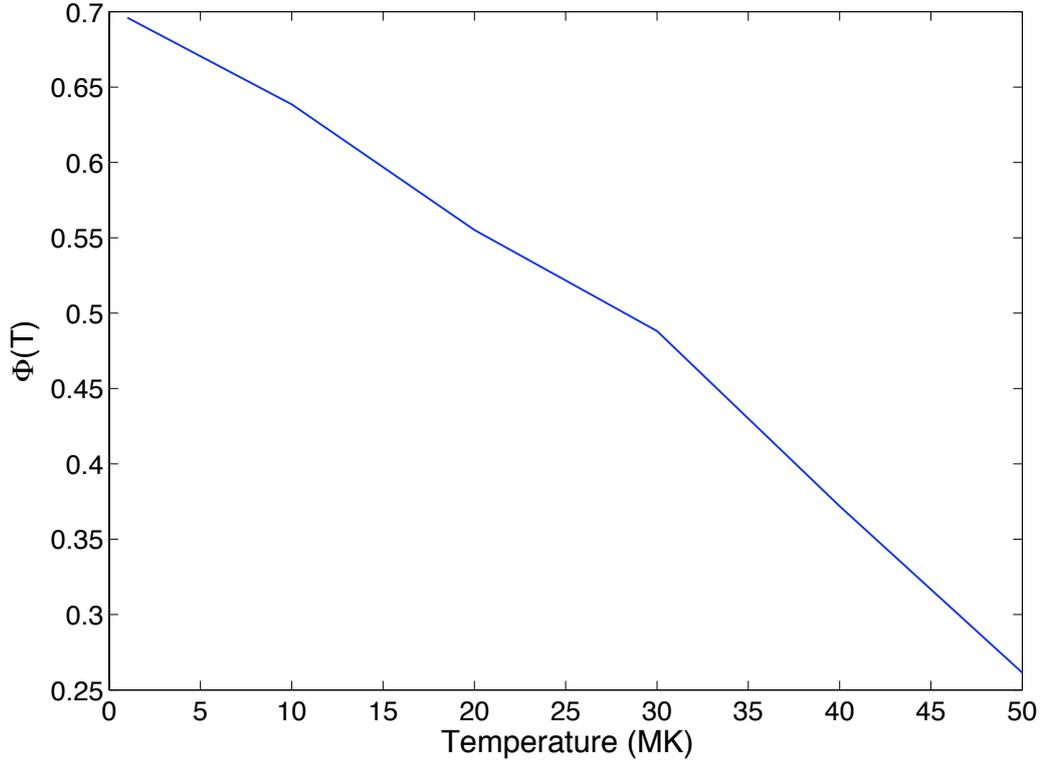


Figure 3.5: The factor  $\Phi(T)$  by which  $EM_c$  is reduced when NTR and NTB are included compared to NTB alone for  $\delta = 5$

2002 April 14 (Veronig & Brown 2004), 2002 October 27 (Krucker et al. 2007) and 2007 December 31 (Krucker et al. 2008c). (Their inferred parameters are based on  $J_{NTB}$ ,  $J_{Th}$  alone). In Chapter 2, we derived plasma parameters for the 2002 April 14 flare to illustrate the significance of NTR for a compact very dense hot coronal loop, which is one class of coronal HXR sources. For this event, the (NTB,  $T_h$  only) parameter estimates were  $EM \approx 4 \times 10^{49} \text{cm}^{-3}$  and, for  $E_c = 20 \text{keV}$ ,  $EM_c \approx 3 \times 10^{46} \text{cm}^{-3}$  implying only  $f_c \approx 7 \times 10^{-4}$  of electrons accelerated. Based on  $T \approx 18 \text{MK}$  for that event, including NTR reduces these to  $EM_c^* \approx 1.5 \times 10^{46} \text{cm}^{-3}$  and  $f_c^* \approx 3.5 \times 10^{-4}$ .

Quite distinct from such dense low-lying loop sources are the more tenuous high altitude HXR sources. 2002 October 27 was one such event for which, neglecting recombination, Krucker et al. (2007) inferred a peak thermal  $EM \approx 5 \times 10^{46} \text{cm}^{-3}$ ,  $T = 12 \text{MK}$  and a non-thermal emission measure  $EM_c \approx 5 \times 10^{45} \text{cm}^{-3}$  for  $E_c = 10 \text{keV}$ , requiring a large acceleration fraction  $f_c = n_c/n = EM_c/EM \approx 0.1$ . More challenging still is the larger 2007 December 31 event for which Krucker et al. (2008c) found flare peak parameters of  $T \sim 19 \text{MK}$ ,  $EM \sim 2 \times 10^{48} \text{cm}^{-3}$ ,  $EM_c \gg 2 \times 10^{47}$

$\text{cm}^{-3}$  and  $\gamma \approx 4.2$  giving  $f_c \gg 0.1$  for  $E_c = 10$  keV, concluding that such a large  $f_c$  implies the usual model of HXR production by NTB alone in ‘cool’ plasma is not viable. However, for these event temperatures, Figure 3.5 shows that including NTR reduces the necessary accelerated fractions  $f_c$  to about 0.06 for the October 27 event and to  $< 0.3$  for December 31, which are physically possible though still demanding.

One has to consider requirements for coronal containment of such electrons for the duration of the events. For the purposes of argument, if we assume the absence of a magnetic field, free motion of deka-keV electrons would result in expansion of coronal HXR source volume  $\mathcal{V}$  on a timescale of order  $L^{1/3}/v_c \approx 0.1\mathcal{V}_{27}^{1/3}$  sec, much shorter than event durations  $t_o$ . The magnetic field pressure required to prevent such expansion due to both thermal and nonthermal particle pressures is of order

$$B > B_{\text{contain}} = [4\pi nkT(2 + f_c E_c/kT)]^{1/2} = 10 \text{ G} \times [n_{10} T_7 (2 + [f_c E_c (\text{keV})/10]/T_7)]^{1/2}, \quad (3.9)$$

which is a modest requirement. However the collisional energy loss lifetime of the energetic electrons is only  $t_{\text{coll}} \approx [E_c (\text{keV})/20]^{3/2}/n_{10}$  sec. To offset this the total energy  $> f_c n E_c$  of the fast electrons has to be replenished roughly  $t_o/t_{\text{coll}}$  times during event duration  $t_o$ . Supplying this by local release of free magnetic energy requires

$$B > B_{\text{energy}} \text{ G} \approx 10^3 n_{10} \left[ \frac{f_c}{[E_c (\text{keV})/20]^{1/2}} \frac{t_o}{100} \right]^{1/2}, \quad (3.10)$$

which is a demanding requirement for the above  $f_c$  values.

By including NTR, the  $f_c$  requirements are of course less demanding for every event, but to varying degrees based on  $T$ ,  $\delta$  and  $E_c$ . Our key result is that, for  $T > 10$  MK and  $\delta \approx 5$ , including NTR reduces the demand for non-thermal electrons by up to 50%.

### 3.4.2 Chromospheric Footpoint HXR Sources

Most discussion of flare structure is in terms of the lower atmosphere being cool (‘chromospheric’). If that were so, highly ionised heavy elements should be rare there and NTR emission unimportant in chromospheric HXR footpoints, for example. However, the observational fact (Hudson et al. 1994; Mrozek & Tomczak 2004) is that, as well as optical and UV footpoints, the impulsive phase commonly also exhibits SXR ‘footpoints’ with  $T$  in the range  $\sim 10$  MK and substantial emission measure. This is easily shown to be ample hot plasma emission measure for a large NTR contribution to be present, as well as NTB, if a substantial fraction of the fast electrons is in the hot SXR

part of the footpoints. To get a rough estimate, we note that for footpoints to be bright in SXR requires  $EM \approx 10^{49} \text{ cm}^{-3}$ ,  $T > 10 \text{ MK}$  and in HXR requires  $EM_c \lesssim 10^{47} \text{ cm}^{-3}$  in the NTB-only interpretation. This requires an accelerated fraction of  $f_c \lesssim 0.01$  in the NTB-only interpretation and hence, for  $\delta = 5$ , a fraction of only  $f_c^* \lesssim 3 \times 10^{-3}$  when NTR is included. To attain a total SXR  $EM = 10^{49} \text{ cm}^{-3}$  in an atmosphere of scale height  $H = 10^8 \text{ cm}$  and where the density  $n = 10^{12} n_{12} \text{ cm}^{-3}$ , requires an SXR source area  $\mathcal{A}_{SXR} = 10^{17}/n_{12}^2$  or around 4 arcsec<sup>2</sup>, which is a reasonable value.

To be more definitive, we can consider the specific data of [Mrozek & Tomczak \(2004\)](#), who were able to make a comparison between impulsive SXR and HXR footpoints for 37 events. The basic parameters involved are  $EM$ ,  $T$ ,  $EM_c$ ,  $\gamma$ ,  $E_c$  and  $\mathcal{V}$ . For every event, [Mrozek & Tomczak \(2004\)](#) give individual values for  $E_c$  and  $\gamma$  and for the total thermal energy,

$$\mathcal{E}_{Th} = 3nkT\mathcal{V} = 3(EM\mathcal{V})^{1/2}kT, \quad (3.11)$$

and nonthermal energy,

$$\mathcal{E}_{NTh} \simeq n_c E_c \mathcal{V}. \quad (3.12)$$

The typical values being  $\gamma = 5 - 6$ ,  $E_c = 10 - 20 \text{ keV}$  and  $\mathcal{E}_{Th} \sim \mathcal{E}_{NTh} \simeq 10^{29} \text{ ergs}$ . For  $T$  and  $n$ , only histograms of distributions were given. The distributions have typical mean  $T \simeq 10 \text{ MK}$  and  $n \simeq 10^{11} \text{ cm}^{-3}$ , corresponding to a source  $\mathcal{V} \simeq 10^{27} \text{ cm}^3$  for the typical observed  $EM \simeq 10^{49} \text{ cm}^{-3}$ , just as we used above. Note that [Mrozek & Tomczak \(2004\)](#) find  $\mathcal{E}_{Th}/\mathcal{E}_{NTh} \simeq 1$ , which implies  $f_c = n_c/n \sim kT/E_c \lesssim 0.1$ . However, their values of  $EM_c$  and hence of  $\mathcal{E}_{NTh}$  are based on NTB emission only. As in the coronal case, including NTR reduces  $EM_c$  by factors of up to 2, hence reducing  $f_c$  to less than 0.05, comparable to our rough estimate above. Comparing the data from [Mrozek & Tomczak \(2004\)](#) with our results thus confirms that NTR can be important in chromospheric footpoints as well.

## 3.5 Conclusions

Through our study we have shown that including NTR can play a significant role in discussing some of the intriguing questions of flare energy budgets and particle acceleration. Using various examples of events both in the tenuous corona and dense chromosphere, it is evident that the conditions are suitable for NTR to be significant. Moreover, in some cases the NTB-only explanation demands a sometimes unphysically high fraction ( $> 10\%$ ) of non-thermal electrons whereas if NTR is included, the fraction

becomes substantially less and quite viable. In a typical flare plasma of  $T = 20$  MK,  $\delta = 5$  and  $E_c = 10$  keV, inclusion of NTR reduces the requirement for accelerated electron numbers by a factor of as much as 2 or 3. For steeper spectra and lower  $E_c$ , NTR would become even more significant. Hence it is important that future studies look to include NTR.

The upshot of our model is that although NTR is seldom dominant over NTB *and* the total thermal emission, it still can be substantial enough whereby its inclusion is necessary. Small thermal and non-thermal recombination edges can clearly be observed in our model spectra (Figure 2.8) and these would have a significant impact on spectral inversion inference as talked about in §2.6. This is so because even a small change in  $J(\epsilon)$ , especially a sharp edge, has a huge impact on inferred  $F(E)$  since inversion invariably magnifies any changes - and sharp features in particular - to the photon spectra. (These aspects of data inversion have been dealt with in detail in Chapter 6, albeit in the context of neutrons.) In extreme parameter regimes ( $\delta > 5$ ,  $E_c < 10$  keV,  $f_c > 0.5$ ,  $25 < T < 35$  MK), NTR is also likely to be a dominant or substantial source of HXRs and so a reconsideration in the number and efficiency problems, may still be necessary for flares with such parameters. However, for a majority of flares, NTR is unlikely to be dominant but still significant enough to demand consideration. Furthermore, the electron energies where NTR is important (10-30 keV) is also the region where most of the electron energy lies, and so this will have an impact on deducing flare energy budgets. Also important to note is that there may well be other stars or astronomical circumstances in which the more extreme parameters are common-place and so NTR dominance may well be typically prevalent in those locations. Finding those sorts of astronomical atmospheres is perhaps a joy for someone else to discover!

Another enhancement to the NTR contribution is the possible enrichment of Fe abundance in the corona by a factor of  $\sim 10$  (Feldman et al. 2004) compared to photospheric levels. We have used an enrichment of about 3 (Feldman 1992). This would increase our NTR predictions by another factor of 3, in which case even for ‘normal’ flare parameters, NTR will start pushing the ‘dominant’ boundary. Clearly, a lot of work needs to be done to see if the First Ionisation Potential (FIP) effect really does enrich low FIP elements, such as Fe, to the levels mentioned by Feldman et al. (2004).

Further improvements to our model can be made by using known measured recombination cross-sections instead of Kramers’. But our research has revealed that Kramers’ cross-section is not a bad approximation and is within a few percent of measured cross-sections. Another refinement in our model would be to include recombination onto 2 or 3 higher levels apart from the lowest available level of the ion. This too would make our

model more accurate by a few more percent. After making these minor adjustments, it would be beneficial to include NTR onto analysis packages such as SolarSoft, such that its diagnostic importance can be realised and that astronomers around the world are made aware of it and benefit from it.

Another concern that has been voiced to us is the question of ionic equilibria when considering recombination. Since we require Fe ions, for example, for recombination to occur, what happens once such an ion has already been recombined in to? This is a much more fundamental problem and deals with the very concept of ionic equilibrium based on the plasma temperature. However, these concerns have never been raised for TR and are hence even less relevant for NTR since TR deals with much larger numbers of electrons. As a concept, it is assumed that a particular temperature corresponds to a particular relative abundance of ions. Another concern is also the availability of electrons, since they too are taken out of circulation once they recombine onto an ion. However, these are valid concerns for recombination as a whole and it would be good to compare recombination and re-ionisation rates to see if there are always a sufficient number of ions and electrons present for recombination to be possible and sustainable. But all these issues will perhaps be addressed in someone else's thesis! We have however included a couple of appendices on dielectronic and "inverse" recombination in Appendix [A](#) and [B](#) respectively.

# Chapter 4

## Inverse Compton Scattering

This chapter closely follows a joint-author publication, where my contributions include, among other things, the following: doing the calculation and calculus to make ICS appropriate to the solar context by tailoring it to the solar radiation field geometry; implementing the result in code to produce all the plots and figures; and analysing particular flare events to interpret them through ICS.

### 4.1 Introduction to ICS

Our motivation to look at inverse Compton scattering (ICS) has not been the same as that had been for non-thermal recombination. For recombination, our intention was to show that it is possibly quite often an important source of HXR's and that it should be incorporated into our present understanding of non-thermal models. For ICS, our interest came from a completely different angle. If we are able to find even a select few events where ICS may be important, then we can make preliminary estimates of particles that cause the ICS, particularly relativistic positrons. This would allow us to get a handle on neutron distributions near the Sun and further our understanding of solar neutron astronomy - the focus of the next half of this thesis - since positrons are produced in the same nuclear reactions that produce neutrons, e.g.  $p + p \rightarrow p + n + e^+ + \nu_e$ , i.e. proton-proton (p-p) interactions. In this chapter, we intend to show how this may be achieved with this preliminary study ([MacKinnon & Mallik 2009](#)).

In solar flares, ICS of photospheric photons might give rise to detectable hard X-ray photon fluxes from the corona where ambient densities are too low for significant bremsstrahlung or recombination.  $\gamma$ -ray lines and continuum in some large flares imply the presence of the necessary  $\sim 100$  MeV electrons and positrons, the latter as by-

products of GeV energy ions. Recent observations of coronal hard X-ray sources in particular prompt us to reconsider here the possible contribution of ICS. We aim to evaluate the ICS X-ray fluxes to be expected from prescribed populations of relativistic electrons and positrons in the solar corona. The ultimate aim is to determine if ICS coronal X-ray sources might offer a new diagnostic window on relativistic electrons and ions in flares. We use the complete formalism of ICS to calculate X-ray fluxes from possible populations of flare primary electrons and secondary positrons, paying attention to the incident photon angular distribution near the solar surface and thus improving on the assumption of isotropy made in previous solar discussions. We know that in regions of typical flare electron density, say  $10^{10} \text{ cm}^{-3}$ , ICS radiation would be minimal compared to bremsstrahlung or recombination, but in more tenuous regions of the corona, where the density may be say  $10^8 \text{ cm}^{-3}$ , ICS will get more important. In the highly tenuous interstellar or intergalactic medium, ICS is known to be an important source of cosmic X-rays and  $\gamma$ -rays produced in galaxies (Strong 1996) or generated near active galactic nuclei (Rogers 1991).

As we have enunciated earlier, Korchak (1967, 1971) considered three possible radiation mechanisms via which solar flare energetic electrons might produce HXR: synchrotron, bremsstrahlung and ICS. He established that fluxes from electron-ion bremsstrahlung would dominate those from the other two mechanisms under normal solar atmosphere conditions and thus laid one of the foundations of the interpretation of flare X-rays. Left open, however, was the possibility that ICS HXR fluxes from low-density regions might exceed those from bremsstrahlung (or, indeed, recombination - Chapters 2 and 3 (Brown & Mallik 2008, 2009)). Recent years have seen increasingly detailed observations of coronal HXR sources (Hudson et al. 2001; Krucker et al. 2008a,b; Tomczak 2009), sometimes from surprisingly tenuous regions. Reconsideration of the possible role of ICS in HXR production thus seems timely (Krucker et al. 2008a).

The basics of ICS are well understood (e.g. Blumenthal & Gould 1970; Rybicki & Lightman 1986). Suppose that electrons of (total) energy  $\gamma m_e c^2$  scatter photons of initial energy  $\epsilon_1$ . Optical photons of photospheric origin, for instance, would have  $\epsilon_1$  typically of order 2 eV. The maximum possible scattered photon energy results from a head-on collision of electron and photon and can be derived as follows (Blumenthal & Gould 1970):

Consider a relativistic electron moving through a gas of photons in the  $x$ -direction of the laboratory system (LS). This electron will scatter off photons, which are moving at various angles,  $\theta$ , with respect to the  $x$ -axis. In the electron's rest system (ERS),

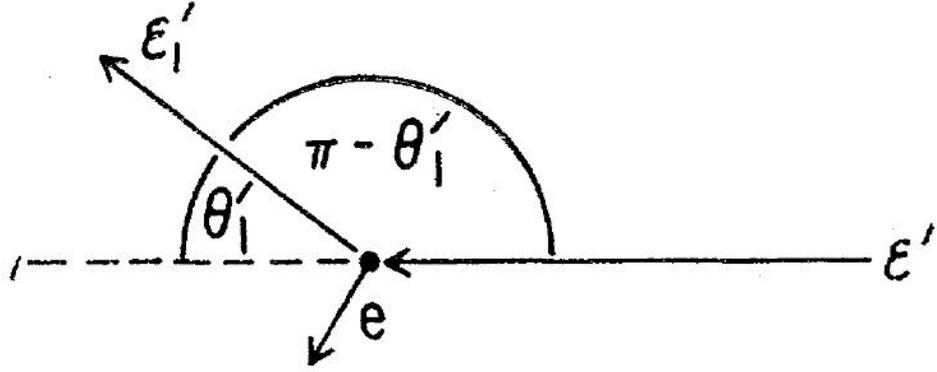


Figure 4.1: Scattering angle in the ERS as given by [Blumenthal & Gould \(1970\)](#). Note that we use  $\epsilon'_1$  and  $\epsilon'_2$  instead of  $\epsilon'$  and  $\epsilon'_1$  for the incident and scattered photon energy respectively.

the corresponding angle,  $\theta'$ , is given by

$$\tan \theta' = \frac{\sin \theta}{\gamma(\cos \theta - \beta)}, \quad (4.1)$$

where  $\gamma$  is the Lorentz factor and  $\beta c$  is the velocity of the electron in the LS. Of course, for the highly relativistic case,  $\gamma \gg 1$  and  $\beta$  tends to  $1 - 0.5\gamma^{-2} \rightarrow 1$ . So only for photons moving virtually along the  $x$ -axis is  $\theta'$  not too small. Hence, Equation 4.1 becomes

$$\tan \theta' = -\frac{\cot(\theta/2)}{\gamma}. \quad (4.2)$$

Hence, in the ERS, the photons are incident in a narrow cone  $1/\gamma$  wide in the direction of the minus  $x'$ -axis. This relativistic effect is alluded to later in this chapter when we apply ICS to the solar context. The photon energy in the ERS can be expressed as

$$\epsilon'_1 = \gamma\epsilon_1(1 - \cos \theta), \quad (4.3)$$

which can range from a value of  $\epsilon'_{1 \min} = \epsilon_1/2\gamma$  for  $\theta = 0$  to  $\epsilon'_{1 \max} = 2\gamma\epsilon_1$  for  $\theta = \pi$ . The scattered photon goes off at an energy of  $\epsilon'_2$  and scattering angle  $\theta'$  in the ERS (Figure 4.1). The energy of the scattered photon is given by the famous relation

$$\epsilon'_2 = \frac{\epsilon'_1}{1 + (\epsilon'_1/m_e c^2)(1 - \cos \theta')}. \quad (4.4)$$

In the LS, therefore, this energy is

$$\epsilon_2 = \gamma\epsilon'_2[1 + \cos(\pi - \theta')] = \gamma\epsilon'_2(1 - \cos \theta'), \quad (4.5)$$

which yields a maximum up-scattered photon energy  $\epsilon_{2 \max} = 2\gamma\epsilon'_2$ . In the Thomson Limit, where  $\epsilon'_1 \ll m_e c^2$ , by Equation 4.4,  $\epsilon'_2 = \epsilon'_1$ , and the maximum up-scattered photon energy has a value of  $\epsilon_{2 \max} \simeq 2\gamma\epsilon'_{2 \max} \simeq 4\gamma^2\epsilon_1$ . In our case, however, we are not in the Thomson Limit (where only a small fraction of the electron energy is imparted to the photon) and so the expression for the maximal up-scattered photon energy differs for the extreme case (where a large fraction of the electron energy is transferred to the photon) as we show in §4.3. Nonetheless, to produce HXR photons via ICS of optical photons needs electrons in the 10s to 100s of MeV energy range.

There is good evidence that electrons attain such energies in flares.  $\gamma$ -ray continuum in this energy range has been observed from some large flares (e.g. [Forrest & Chupp 1983](#); [Kanbach et al. 1993](#); [Talon et al. 1993](#)). This may be due to either or both of: electron-ion bremsstrahlung from primary accelerated electrons; bremsstrahlung from secondary electrons and positrons in the 100 MeV energy range, produced in reactions of accelerated ions in the energy range  $> 0.3$  GeV (e.g. [Murphy et al. 1987](#)). In the latter case positrons are dominant in number since they result from collisions between positively charged particles, e.g. p-p interactions. Continuum in this case is unavoidably accompanied by the flat spectral feature around 70 MeV produced by  $\pi^0$  decay. High-energy continuum can occur both with and without this feature at different times during a single event (e.g. [Vilmer et al. 2003](#)), indicating that both primary accelerated electrons and secondary positrons may be present in the 100 MeV energy range, as needed for ICS HXR production. [Akimov et al. \(1994\)](#) give evidence that the flare of 1991 March 26 accelerated electrons to energies of 300 MeV. The energy distributions of the scattering particles in these two cases will be very different and we consider them separately. Below we sometimes use the word ‘electrons’ as shorthand for both electrons and positrons, however.

In the presence of the solar magnetic field, these high energy electrons would also produce synchrotron emission but at millimetre ([Ramaty et al. 1994](#)), radio and sub-mm wavelengths ([Silva et al. 2007](#)). Synchrotron losses would be important, possibly dominant for these electrons in the corona, but the associated radiation is not in the HXR range. To produce X-rays by synchrotron emission would require electrons of unrealistically high energy, for which there is no evidence.

The ICS estimates of [Korchak \(1967, 1971\)](#) and in [Krucker et al. \(2008a\)](#) employ standard results based on assuming isotropic electron and photon distributions. Electron distributions in the corona may well be isotropic because of pitch-angle scattering by MHD turbulence (e.g. [Miller & Ramaty 1989](#)) but the photon distribution will be isotropic only in the outward hemisphere. As already mentioned, the most energetic

photons result from head-on collisions of photon and electron, which result in the up-scattered photon travelling along the direction of the incident electron (Jones 1968). These most favourable collisions clearly cannot occur, even with an assumed isotropic coronal electron distribution, so a more involved calculation is essential to evaluate likely ICS fluxes, spectra etc.

ICS is certainly important in other areas of astrophysics: of cosmic microwave background photons by hot gas in clusters of galaxies (Sunyaev-Zeldovich effect - Sunyaev & Zeldovich 1970); of solar visible photons by cosmic ray electrons (Orlando & Strong 2008), which has been observed as  $\gamma$ -radiation by the EGRET mission (Kanbach et al. 1989) and corroborated by Orlando & Strong (2008) in their theoretical model. This  $\gamma$ -ray emission is predicted to be an extended and confusing foreground for the diffuse extragalactic background.

The formalism for calculating ICS radiation with arbitrary photon angular distributions has been given most recently by Moskalenko & Strong (2000). Here we adapt their work to the source geometry near the solar surface. We use typically observed power-law distributions of electrons and protons (which produce secondary positrons) to illustrate our study. We elucidate the consequences for observability of this ICS flux and note the difference between the spectra produced by electrons and by secondary positrons, as well as the variation across the solar disc. Our findings reveal that although the ICS intensities are likely to be low, the spectrum is hard and unmistakable. If detected by modern instruments, this would be a new window on extremes of electron and ion acceleration at the Sun, in the case of ions complementing information available from gamma-ray lines and free neutrons detected in space.

In this chapter, we use the units  $\hbar = c = m_e = 1$ .

## 4.2 ICS versus NTB

Let us first look at the following analogy given in (Krucker et al. 2008a). They illustrate how ICS can be a significant source of HXRs and even exceed the bremsstrahlung emission for certain flare conditions. Take for instance the following expression for ICS intensity (photons  $\text{s}^{-1} \text{keV}^{-1}$ ) produced by isotropic relativistic electrons interacting with isotropic mono-energetic photons of energy  $\epsilon_i$  (Blumenthal & Gould 1970; Tucker 1975):

$$I_{ICS}(\epsilon) = \frac{8\pi r_e^2 c}{\epsilon} n_\gamma N_e (\delta - 1) \left( \frac{E_c}{m_e c^2} \right)^{\delta-1} Q(\delta) \left( \frac{\epsilon}{4\epsilon_i} \right)^{(1-\delta)/2}, \quad (4.6)$$

where  $n_\gamma$  is the photon number density,  $N_e$  is the total number of electrons in a homogeneous source above  $E_c$ ,  $r_e$  is the classical electron radius and the electrons have a power-law energy distribution function  $\approx E^{-\delta}$ .  $Q(\delta)$  is a function of  $\delta$  of order unity. To compare the relative strengths of ICS with NTB, using Equations 4.6 and 2.15 duly modified, we get

$$\frac{j_{ICS}(\epsilon)}{j_B(\epsilon)} = \frac{3n_\gamma}{2\alpha n_p} \delta Q(\delta) \left(\frac{\epsilon}{4\epsilon_i}\right)^{(1-\delta)/2} \left(\frac{\epsilon}{m_e c^2}\right)^{\delta-1/2}. \quad (4.7)$$

Since ICS has a harder spectrum, it will eventually dominate over NTB at high energies. However, even at lower energies, if we take tenuous regions of the corona, where say the ambient density is  $10^9 \text{ cm}^{-3}$  and the photospheric photon density is  $10^{12} \text{ cm}^{-3}$ , then  $n_\gamma/n_p = 1000$  and the ratio above can be close to or exceed 1 in the 10-100 keV photon energy range if the electron distribution is hard, e.g.  $\delta = 2$ .

This was a simplistic example to put forward the potential importance of ICS for certain flare conditions. The rest of the chapter looks into the problem in greater detail taking into account the radiation field geometry at the solar surface and even considering a black-body distribution of photons, although the latter has little effect. We also make comparisons with recent observations to suggest that ICS can well be the source of some coronal HXRs.

### 4.3 In the solar context and geometry

In this section we calculate ICS HXR fluxes from relativistic electron and positron populations in the corona, following [Moskalenko & Strong \(2000\)](#).

The rate of photon-particle interactions is given in full generality by ([Weaver 1976](#)):

$$R = n_e n_\gamma \int d\mathbf{p}_\gamma \int d\mathbf{p}_e f_e(\mathbf{p}_e) f_\gamma(\mathbf{p}_\gamma) \frac{p'_\gamma}{\gamma p_\gamma} \sigma(p'_\gamma), \quad (4.8)$$

where  $n_e$ ,  $n_\gamma$  are the electron and photon number densities;  $\mathbf{p}_e$ ,  $\mathbf{p}_\gamma$  are the momenta;  $f_e(\mathbf{p}_e)$ ,  $f_\gamma(\mathbf{p}_\gamma)$  are the respective distribution functions in the laboratory system (LS), normalised to unity;  $\gamma$  is the electron Lorentz factor;  $\sigma$  the cross-section; and the primes signify the electron rest system (ERS) variables. For relativistic electrons, the incoming photons are seen as a narrow beam  $\sim 1/\gamma$  wide in the ERS. We follow [Moskalenko & Strong \(2000\)](#) in using the [Jones \(1968\)](#) approximation that the incident photons are seen as a unidirectional beam in the ERS. This significantly simplifies the calculation of the ICS fluxes while introducing negligible error ([Jones 1968](#)). Since we are interested in relativistic electrons, making this approximation is sufficient and

computationally validated since relativistic effects determine the beamed electrons. Hence all our calculations are based on the relativistic limit rather than the Thomson Limit - see [Blumenthal & Gould \(1970\)](#) for details.

To calculate fluxes from Equation 4.8 we need to specify the electron and photon momentum distributions and the cross-section. Since we deal with highly relativistic particles and situations where the photon may carry away a large fraction of the electron energy, we must use the Klein-Nishina cross-section e.g. as given by [Jauch & Rohrlich \(1976\)](#):

$$\frac{d\sigma}{d\epsilon'_2 d\cos\eta'} = \pi r_e^2 \left(\frac{\epsilon'_2}{\epsilon'_1}\right)^2 \left(\frac{\epsilon'_2}{\epsilon'_1} + \frac{\epsilon'_1}{\epsilon'_2} - \sin^2\eta'\right) \delta\left[\epsilon'_2 - \frac{\epsilon'_1}{1 + \epsilon'_1(1 - \cos\eta')}\right], \quad (4.9)$$

where  $r_e$  is the classical electron radius,  $\epsilon'_1$  and  $\epsilon'_2$  are the ERS energies of the incident and up-scattered photons,  $\eta'$  is the scattering angle in the ERS and  $\delta(x)$  denotes the Dirac delta function.

Appropriately to the general galactic cosmic ray population, [Moskalenko & Strong \(2000\)](#) assume isotropic electrons; this assumption will also be appropriate in the corona as a result of MHD scattering ([Miller & Ramaty 1989](#); [Mandzhavidze & Ramaty 1992](#)). This assumption of isotropy greatly simplifies the calculation and even if we were to consider more ‘realistic’ non-isotropic electrons, the marginal gains in accuracy would not justify the increased complexity of the problem. Note that since the radiation is so highly beamed, some aspects of these results remain useful even if the electron distribution is not isotropic, as we mention later. However, things like the variation across the solar disc would change. However, with the above assumptions, the up-scattered photon distribution over the LS energy,  $\epsilon_2$ , as obtained from Equation 4.8 is ([Moskalenko & Strong 2000](#))

$$\frac{dR}{d\epsilon_2} = \int d\cos\eta' \int d\epsilon_1 d\Omega_\gamma \int d\gamma d\Omega_e f_e(\gamma, \Omega_e) f_\gamma(\epsilon_1, \Omega_\gamma) \epsilon_1^2 \gamma^2 \frac{\epsilon'_1}{\gamma\epsilon_1} \frac{\epsilon'_2}{\epsilon_2} \frac{d\sigma}{d\epsilon'_2 d\cos\eta'}, \quad (4.10)$$

where  $\Omega_\gamma$  and  $\Omega_e$  refer to photon and electron directions respectively.

At this point we depart from [Moskalenko & Strong \(2000\)](#), tailoring our calculation to the radiation field geometry above the solar surface (Figure 4.2). To see how we derive the [Moskalenko & Strong \(2000\)](#) equations in order to get the coefficients and relevant expressions, please see Appendix B. We introduce two spherical polar angular coordinates  $\theta$  and  $\phi$  to label photon direction. Let  $\hat{\mathbf{n}}$  be a unit vector pointing radially outward from the local solar surface, and  $\hat{\mathbf{l}}$  be a unit vector pointing along the line of

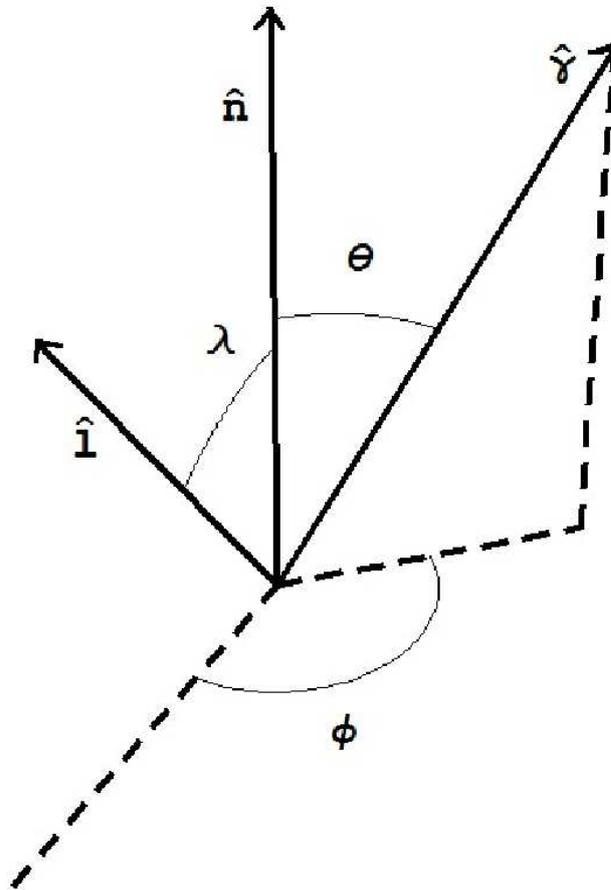


Figure 4.2: Schematic diagram showing the geometry used to describe the radiation field at the solar surface with the relevant angles and vectors.  $\phi$  lies in the solar surface plane.

sight to the observer. Then we have  $\hat{\mathbf{l}} \cdot \hat{\mathbf{n}} = \sin \lambda$  where  $\lambda$  is the heliocentric angle of the source location. Let  $\hat{\mathbf{p}}_\gamma$  be a unit vector in the direction of the photon. The polar angle  $\theta$  measures the angle between  $\hat{\mathbf{n}}$  and  $\hat{\mathbf{p}}_\gamma$ , i.e.  $\hat{\mathbf{n}} \cdot \hat{\mathbf{p}}_\gamma = \cos \theta$ . The photon azimuthal angle  $\phi$  lies in the plane of the solar surface and is measured anticlockwise from the plane defined by  $\hat{\mathbf{n}}$  and  $\hat{\mathbf{l}}$ .

The photon distribution is isotropic in the optically thick photosphere but only includes outward-flowing photons immediately above. It will be close to isotropic, in the hemisphere  $\theta < \pi/2$ , as long as we consider coronal locations below  $\sim R_\odot$  above

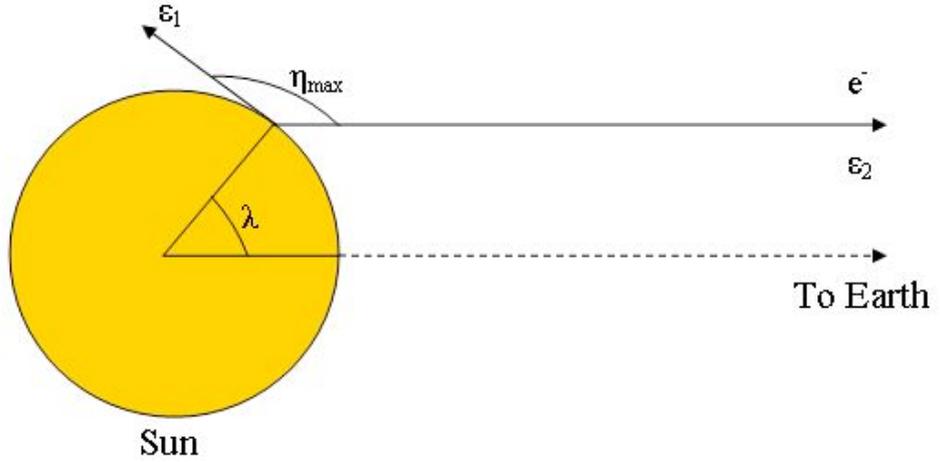


Figure 4.3: Top view of the Sun-Earth ICS geometry system. Note that the Earth is effectively at  $\infty$  to the right and hence electrons/photons from any point on the Sun are parallel to the horizontal axis as viewed from Earth. As  $\lambda$ , or solar longitude, increases from  $0^\circ$  to  $90^\circ$ ,  $\eta_{\max}$  increases from  $90^\circ$  to  $180^\circ$ , which is the most energetic geometric configuration.

the surface. Thus the photon angular distribution takes the simple form

$$f_\gamma(\epsilon_1, \theta, \phi) = \frac{1}{2\pi} H\left(\frac{\pi}{2} - \theta\right) g_\gamma(\epsilon_1), \quad (4.11)$$

where  $H$  is the Heaviside step function.

In the first instance we calculate the ICS flux from monoenergetic electrons with a single energy  $\gamma$ , averaging straightforwardly over more general energy distributions as needed. We also consider monoenergetic primary photon distributions,  $g_\gamma(x) = \delta(x - \epsilon_1)$ . Using Equations 4.9, 4.10 and 4.11, we hence found the total up-scattered photon distribution, per electron, over the LS energy,  $\epsilon_2$ , to be:

$$\begin{aligned} \frac{dR}{d\epsilon_2} = & \left(2 - \frac{4\epsilon_2}{\gamma} + \frac{3\epsilon_2^2}{\gamma^2} - \frac{\epsilon_2^3}{\gamma^3}\right) \int_0^{2\pi} \int_{\theta_{\min}}^{\theta_{\max}} \sin \theta \, d\theta \, d\phi \\ & - \frac{1}{\epsilon_1 \gamma} \left(\frac{2\epsilon_2^2}{\gamma^2} - \frac{2\epsilon_2}{\gamma}\right) \int_0^{2\pi} \int_{\theta_{\min}}^{\theta_{\max}} \frac{d \cos \theta}{1 + \cos \theta} d\phi \\ & - \frac{\epsilon_2^2}{\epsilon_1^2 \gamma^4} \int_0^{2\pi} \int_{\theta_{\min}}^{\theta_{\max}} \frac{d \cos \theta}{(1 + \cos \theta)^2} d\phi. \end{aligned} \quad (4.12)$$

The lower limit of the  $\theta$  integral is given by kinematics:

$$\theta_{\min} = \arccos\left(1 - \frac{\epsilon_2}{2\epsilon_1 \gamma (\gamma - \epsilon_2)}\right) \quad (4.13)$$

and the upper limit by source geometry:

$$\theta_{max} = \arccos(\sin \lambda \cos \phi). \quad (4.14)$$

Performing the integral over polar angle we get

$$\begin{aligned} \frac{dR}{d\epsilon_2} = \frac{r_e^2}{2\epsilon_1(\gamma - \epsilon_2)^2} \times \\ \int_0^{2\pi} \left( \frac{\epsilon_2^3}{\gamma^3} - \frac{3\epsilon_2^2}{\gamma^2} + \frac{4\epsilon_2}{\gamma} - 2 \right) \left( \sin \lambda \cos \phi - \left( 1 - \frac{\epsilon_2}{2\epsilon_1\gamma(\gamma - \epsilon_2)} \right) \right) \\ + \frac{2\epsilon_2}{\epsilon_1\gamma^2} \left( 1 - \frac{\epsilon_2}{\gamma} \right) \left( \ln(1 + \sin \lambda \cos \phi) - \ln \left( 2 - \frac{\epsilon_2}{2\epsilon_1\gamma(\gamma - \epsilon_2)} \right) \right) \\ + \frac{\epsilon_2^2}{\epsilon_1^2\gamma^4} \left( \frac{1}{1 + \sin \lambda \cos \phi} - \frac{1}{2 - \epsilon_2/(2\epsilon_1\gamma(\gamma - \epsilon_2))} \right) d\phi, \end{aligned} \quad (4.15)$$

which is the ICS flux of photons per unit energy per unit time per electron. The following kinematic results (Moskalenko & Strong 2000) are also of importance and differ from the Thomson Limit approximation we derived in the introduction. The detailed derivation given in Jones (1968) produces the following results for the extreme-case approximation:

$$\epsilon'_2 = \epsilon_2/[\gamma(1 - \cos \eta')], \quad \epsilon_2 \leq 2\gamma\epsilon'_1/(1 + 2\epsilon'_1), \quad \epsilon'_1 = \epsilon_1\gamma(1 + \cos \lambda). \quad (4.16)$$

The maximum energy of the up-scattered photon is

$$\epsilon_{2 \max} = 4\epsilon_1\gamma^2/(1 + 4\epsilon_1\gamma). \quad (4.17)$$

Note that the second and third terms in Equation 4.15 have to be evaluated numerically. This was done using MATLAB and the results are portrayed in the following section.

## 4.4 ICS from relativistic flare electrons and positrons

### 4.4.1 ICS from fast electrons

To calculate ICS spectra produced by relativistic electrons, we assumed power-law primary electron kinetic energy distributions extending into the 10s of MeV range,  $\sim (\gamma - 1)^{-\delta}$ . The incident photon population was assumed to have a monoenergetic energy distribution at  $\epsilon_1 = 2$  eV (except in a few cases where we consider the consequences of

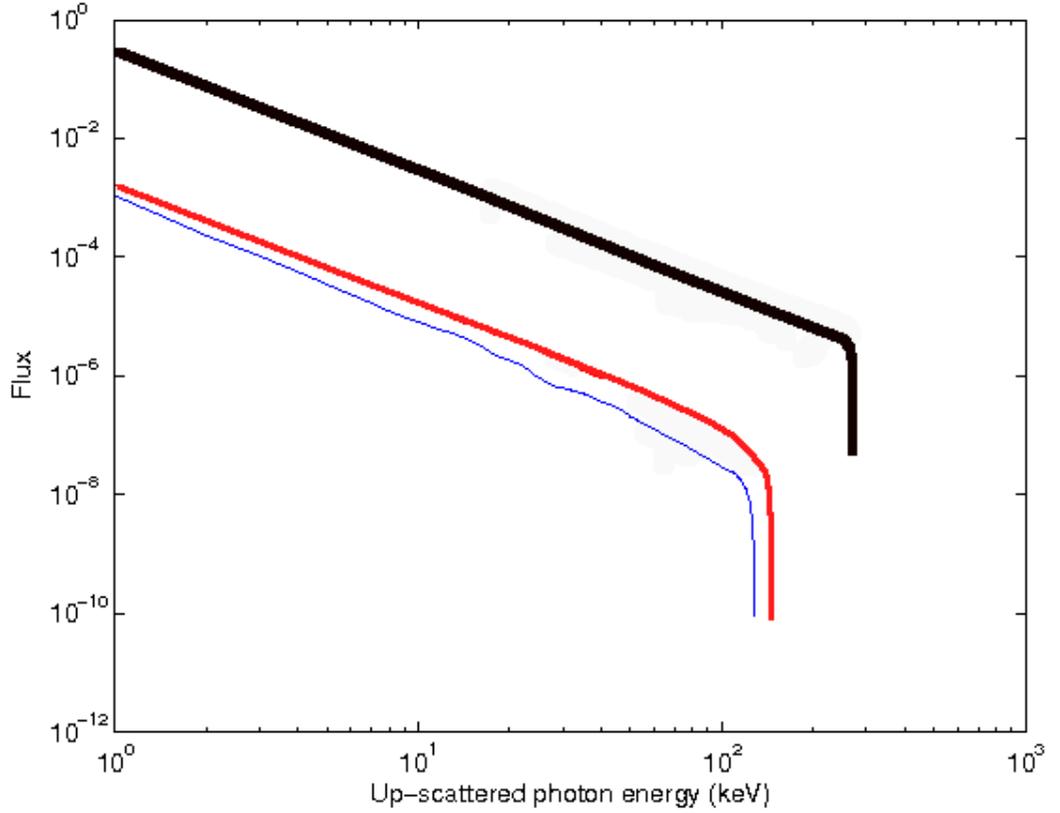


Figure 4.4: Photon spectra at the Sun (photons per keV per second per source electron) from fast electrons with a power-law energy distribution  $E^{-3}$ , where the thin-blue curve is the flux from disc centre ( $\sin \lambda = 0$ ), medium-red for  $\sin \lambda = 0.5$  and thick-black for the solar limb ( $\sin \lambda = 1$ ). Fluxes are normalised to one electron above 0.5 MeV and for an incident photon of energy 2 eV.

more energetic primary photons,  $\epsilon_1 = 200$  eV) so that the solar luminosity implies a photon density  $n_\gamma = 10^{12}$  cm $^{-3}$ . Note that imposing a black-body distribution for the incident photon energy does not change the results significantly enough to justify its inclusion and increase the complexity of the problem - see below. Hence a mean solar black-body photon energy of 2 eV was chosen.

In Figure 4.4, we show the ICS spectra from electrons with energy spectral index  $\delta = 3$ , calculated by weighting the emissivity (Equation 4.15) by this distribution and integrating over electron energy. Fluxes are normalised to one electron (i.e. flux per electron) above 0.5 MeV and we assume an upper cutoff energy of 100 MeV. The three separate curves signify the ICS spectrum as seen from three different viewing angles  $\lambda$ . As expected on geometrical grounds, the most energetic photons come from limb events (since head-on collisions, the most energetic geometric configuration, can occur

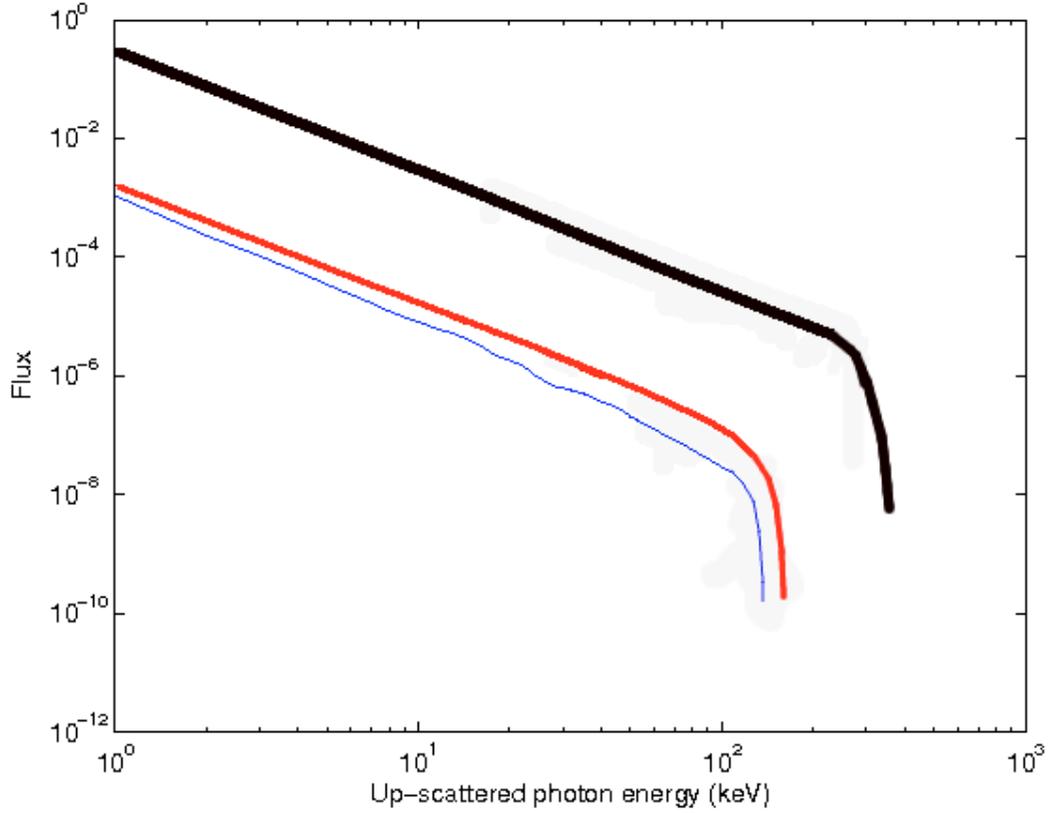


Figure 4.5: Photon spectra at the Sun (photons per keV per second per source electron) from fast electrons with a power-law energy distribution  $E^{-3}$ , where the thin-blue curve is the flux from disc centre ( $\sin \lambda = 0$ ), medium-red for  $\sin \lambda = 0.5$  and thick-black for the solar limb ( $\sin \lambda = 1$ ). Fluxes are normalised to one electron above 0.5 MeV and for an incident solar black-body photon spectrum.

between the beamed electrons coming towards us and a solar photon going away from us only at the limb - as pictured in Figure 4.3), whose flux too can be up to two orders of magnitude greater than that from the disc centre. It is clear that the spectra produced are very hard and would be unmistakable if detected by modern instruments. Over most of the photon range they are described by the expected (e.g. Blumenthal & Gould 1970) ICS power-law  $\sim \epsilon_2^{-(1+\delta)/2}$ , plummeting rapidly close to an upper cutoff determined by the 100 MeV electron upper cutoff, the viewing angle and, as we see below, the energy spectral index  $\delta$ . In § 4.5 we see that observed coronal source photon fluxes imply plausible electron numbers.

For comparison's sake, we have also shown the expected ICS flux if we inject a solar black-body spectrum of photons (Figure 4.5). As you can see, the difference is barely noticeable when this figure is compared to Figure 4.4. The only diagnostic

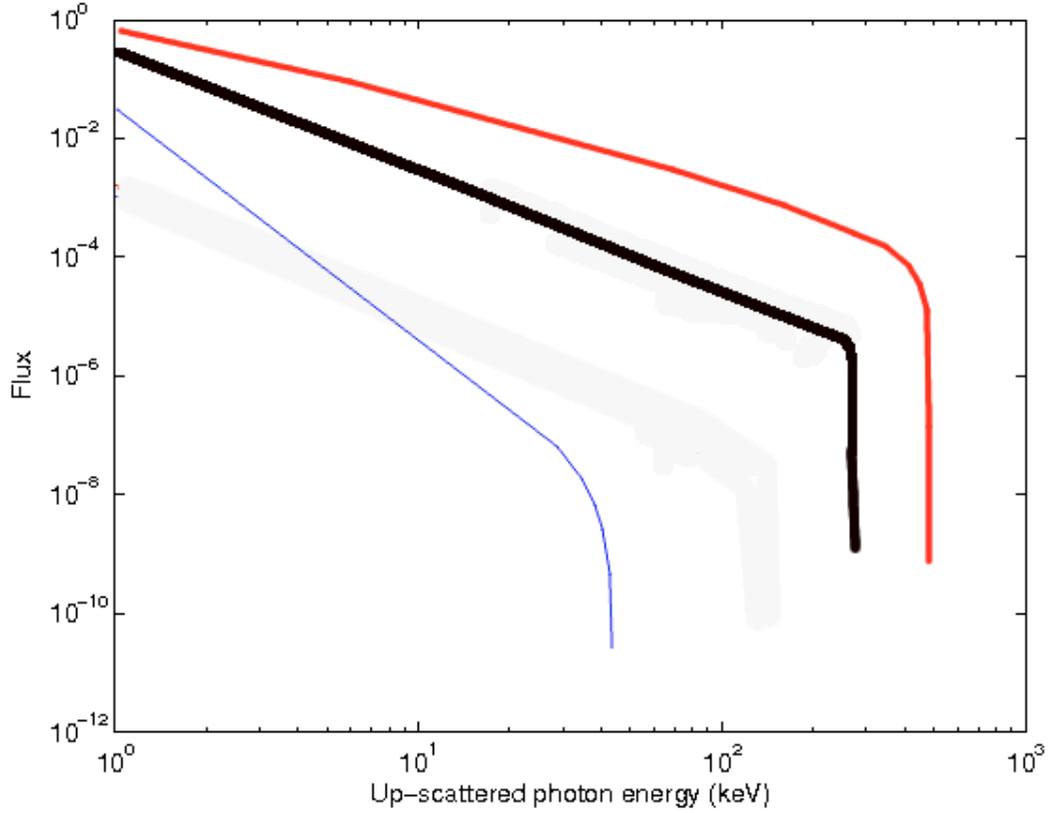


Figure 4.6: Photon spectra at the Sun (photons per keV per second per source electron) from limb fast electrons with different power-law distributions: thin-blue is for  $\delta = 5$ , medium-red for  $\delta = 2$  and thick-black for  $\delta = 3$  with an incident photon energy of 2 eV.

difference is that the flux does not fall off at an upper cut-off as sharply and this is because the black-body spectrum has photons of energy higher than the mean 2 or 2.5 eV. This produces a slightly longer, barely perceivable ‘tail’, but has little impact on the intensity or other diagnostic behaviour. Injecting a black-body spectrum involved adding another integral to the numerical code and the negligible effect it had did not warrant its inclusion to generate all our spectra and increase the complexity of our problem. As you will see later, with secondary positrons you anyway get a ‘tail’ rather than a sharp drop-off and so including the solar black-body-induced ICS spectra for positrons is not shown since there is virtually no difference between the figures.

In Figure 4.6, we show the ICS spectra for an event viewed at the limb but now for different values of  $\delta$ . Again we find power-law behaviour with the expected photon spectral index; we also see that the effective upper cutoff decreases for larger  $\delta$ . Our spectra in Figures 4.4 and 4.6 are based on an incident photospheric photon of energy

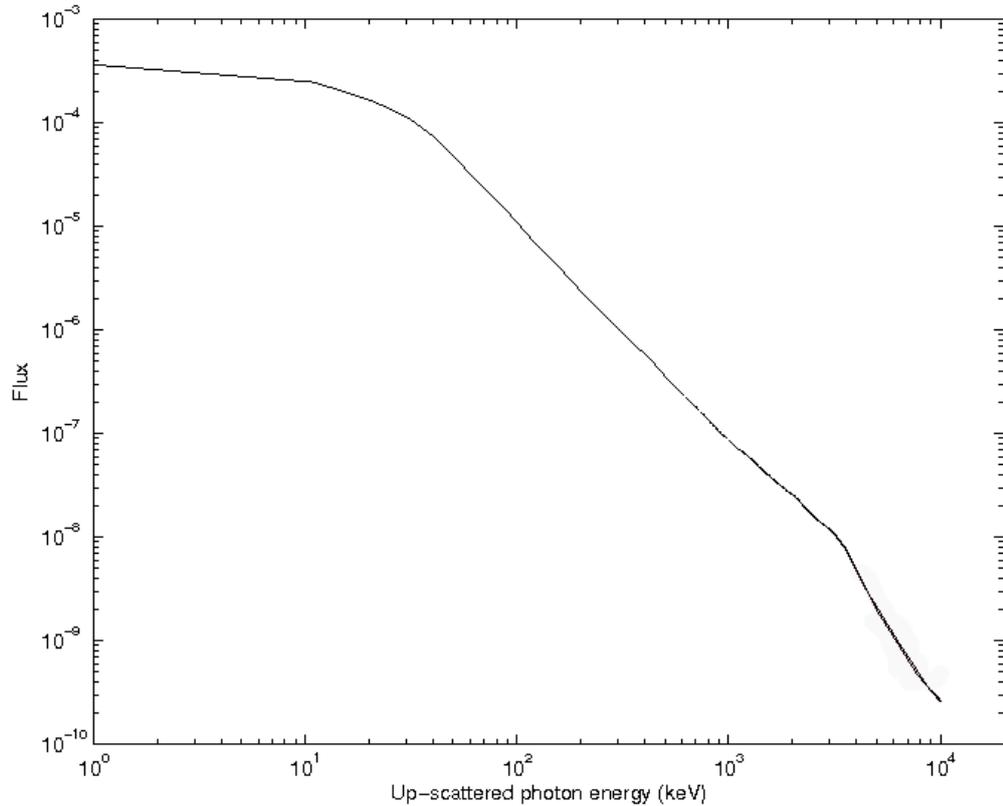


Figure 4.7: Photon spectra at the Sun (photons per keV per second per source electron) from fast electrons with a power-law energy distribution  $E^{-3}$  for an incident photon energy of 200 eV.

2 eV. We can clearly see that these visible photons can quite easily be up-scattered to 10s of keV, even though the actual fluxes and spectra depart from those expected on the basis of an isotropic photon distribution.

Still higher photon energies, but lower fluxes at given  $\epsilon_2$  as we show later, will result from primary photons of higher energy. For illustration we show in Figure 4.7 the spectrum resulting from ICS of primary EUV photons of energy 200 eV, from a flare at disc centre. For easy comparison with the results for optical photons we have adopted the same photon density,  $n_\gamma = 10^{12} \text{ cm}^{-3}$ , although the true EUV density will be many orders of magnitude smaller - see below.

#### 4.4.2 ICS from relativistic positrons

As noted in §4.1, positrons will be produced as secondaries from fast ion reactions. We use the positron energy distributions calculated by [Vilmer et al. \(2003\)](#), in turn

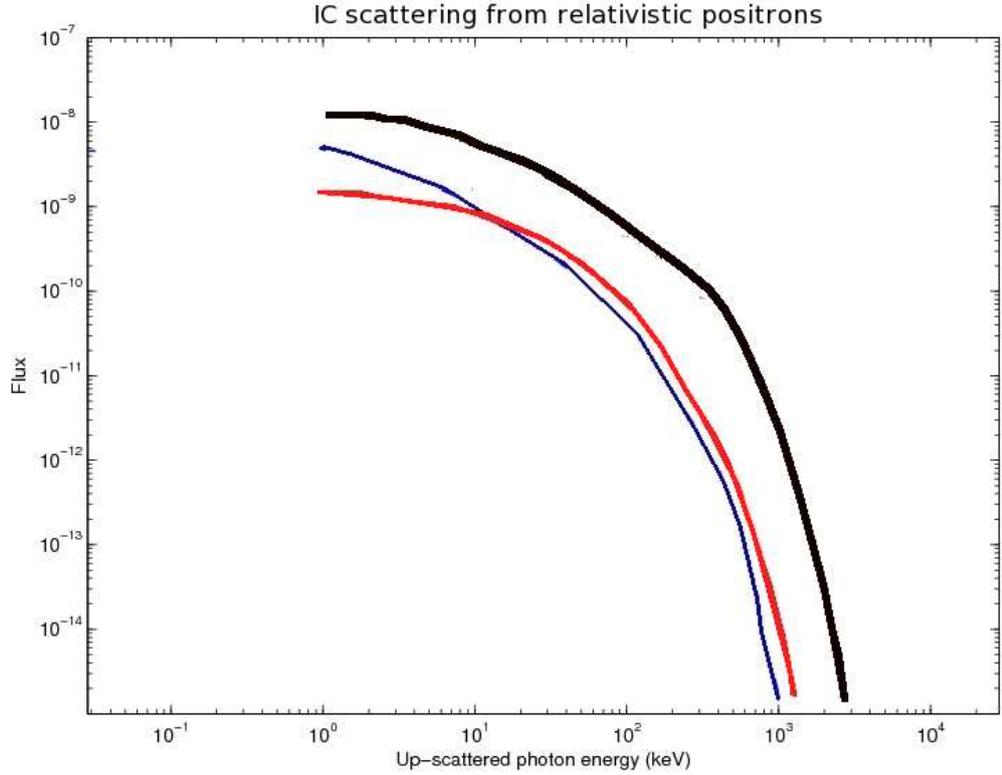


Figure 4.8: Photon spectra at the Sun (photons per keV per second per proton) from relativistic positrons produced by protons with a power-law energy distribution  $E^{-3}$ , where the thin-blue curve is the flux from disc centre ( $\sin \lambda = 0$ ), medium-red for  $\sin \lambda = 0.5$  and thick-black for the solar limb flux ( $\sin \lambda = 1$ ). Fluxes are normalised to one proton above 1 MeV and for an incident photon energy of 2 eV.

calculated following [Dermer \(1986a,b\)](#), assuming they are produced via pion decay following reactions of fast protons with ambient H and He nuclei. ICS spectra from the resulting positrons are shown in [Figure 4.8](#), assuming a power-law proton energy distribution with  $\delta = 3$  extending to an upper cut-off energy of 3 GeV and, again,  $\epsilon_1 = 2$  eV and  $n_\gamma = 10^{12} \text{ cm}^{-3}$ . Electrons and positrons with the same energy distribution would of course produce identical ICS spectra, but the positron energy distribution from p-p collisions, and hence the ICS photon spectrum, are quite different from the power-law electron case considered in [§4.4.1](#). Secondary positron distributions have a maximum at about 300 MeV and a form that is dominated by the nuclear physics of pion formation and decay until primary proton energies significantly exceed the threshold for pion production ([Murphy et al. 1987](#)). Thus the detailed photon spectra depend only rather weakly on proton power-law energy spectral index. However, certain features persist, i.e. the spectrum remains very hard and the most energetic photons

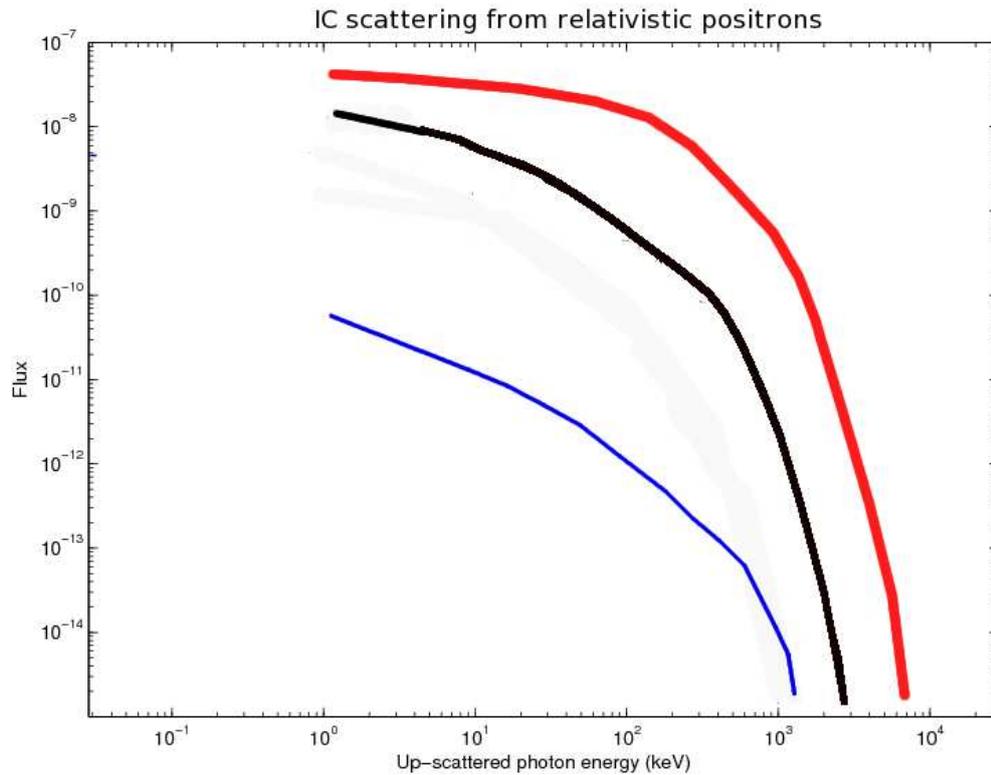


Figure 4.9: Photon spectra at the Sun (photons per keV per second per proton) from limb relativistic positrons produced by protons with different power-law distributions: thin-blue is for  $\delta = 5$ , medium-black for  $\delta = 3$  and thick-red for  $\delta = 2$  with an incident photon energy of 2 eV.

will once again come from the solar limb. The three separate curves are for three different values of  $\lambda$ . And in Figure 4.9, we show the ICS spectra for a range of proton energy distribution  $\delta$  values. The natural energy range of secondary positrons naturally results in up-scattering to the MeV photon energy range.

As in the electron case, we would expect a more energetic ICS flux if we consider incident EUV photons, shown in Figure 4.10 for 200 eV incident photons. However, the relative gain in flux for positrons seems much greater than that for electrons when considering 200 eV incident photons. This is because the positrons extend to more relativistic energies than the electrons and ICS is more effective the more relativistic you get. This can be seen while analysing the kinematic limits for  $\theta$  as shown in Equation 4.13. By increasing  $\epsilon_1$  and  $\gamma$  (which is intrinsically higher for the positrons),  $\theta_{min}$  tends closer to 0, thereby increasing the geometric range of possible ICS scatters and correspondingly increasing the ICS yield. With the photon density held fixed, as for Figure 4.7, the IC flux can be as much as four orders of magnitude greater (for

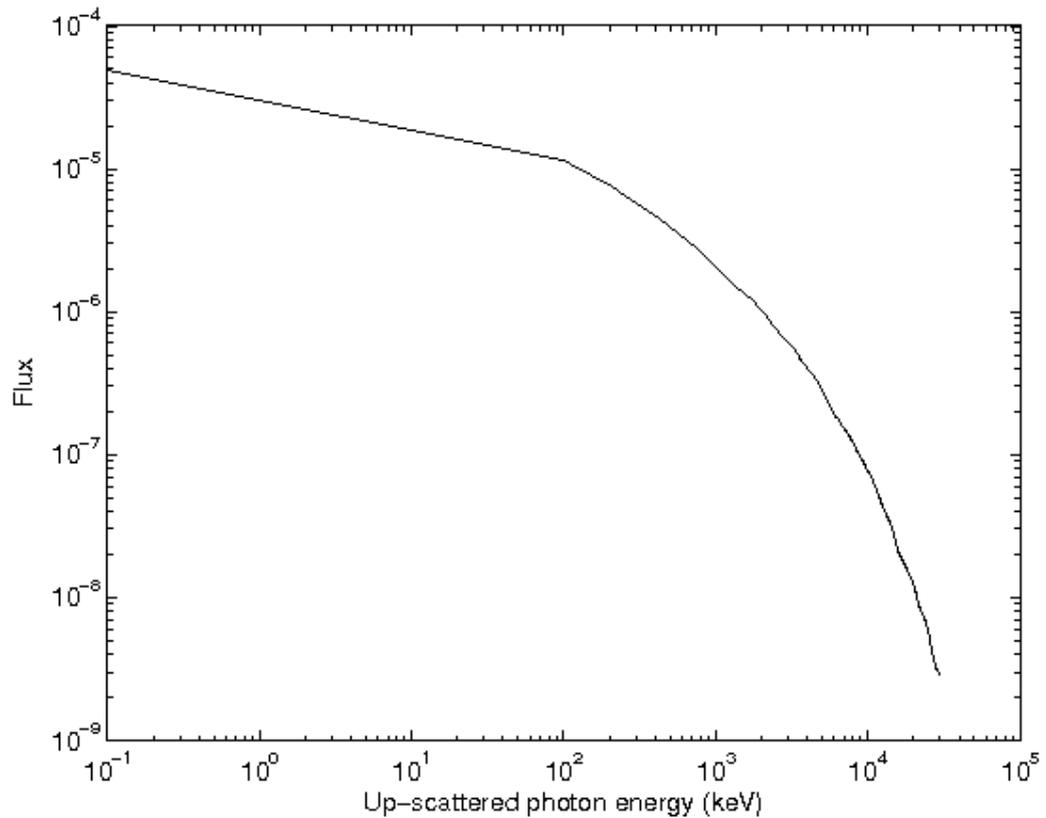


Figure 4.10: Photon spectra at the Sun (photons per keV per second per proton) from relativistic positrons produced by protons with a power-law energy distribution  $E^{-3}$  for an incident photon energy of 200 eV.

electrons the gain was more in the range of two orders of magnitude) for  $\epsilon_1 = 200$  eV than for  $\epsilon_1 = 2$  eV, at the same time extending to higher energies. So we would need an EUV photon density  $\sim 10^{-4}$  times that of visible photons to produce an equally intense ICS flux. A rough estimate of EUV photon density in a large flare suggests this will be  $\sim 10^3$  cm $^{-3}$  (based on a simple calculation of observed luminosities and the black-body spectrum), however, so low that even the greater fluxes obtained with more energetic incident photons will not be observable.

## 4.5 A comment on observations

As explained in §4.1, ICS could be dominant in producing HXRs in low-density regions of the solar atmosphere, which mainly implies the high corona. Consider the coronal X/ $\gamma$ -ray source in the 2005 January 20 flare, described by [Krucker et al. \(2008b\)](#). Could it be due to ICS of photospheric photons?

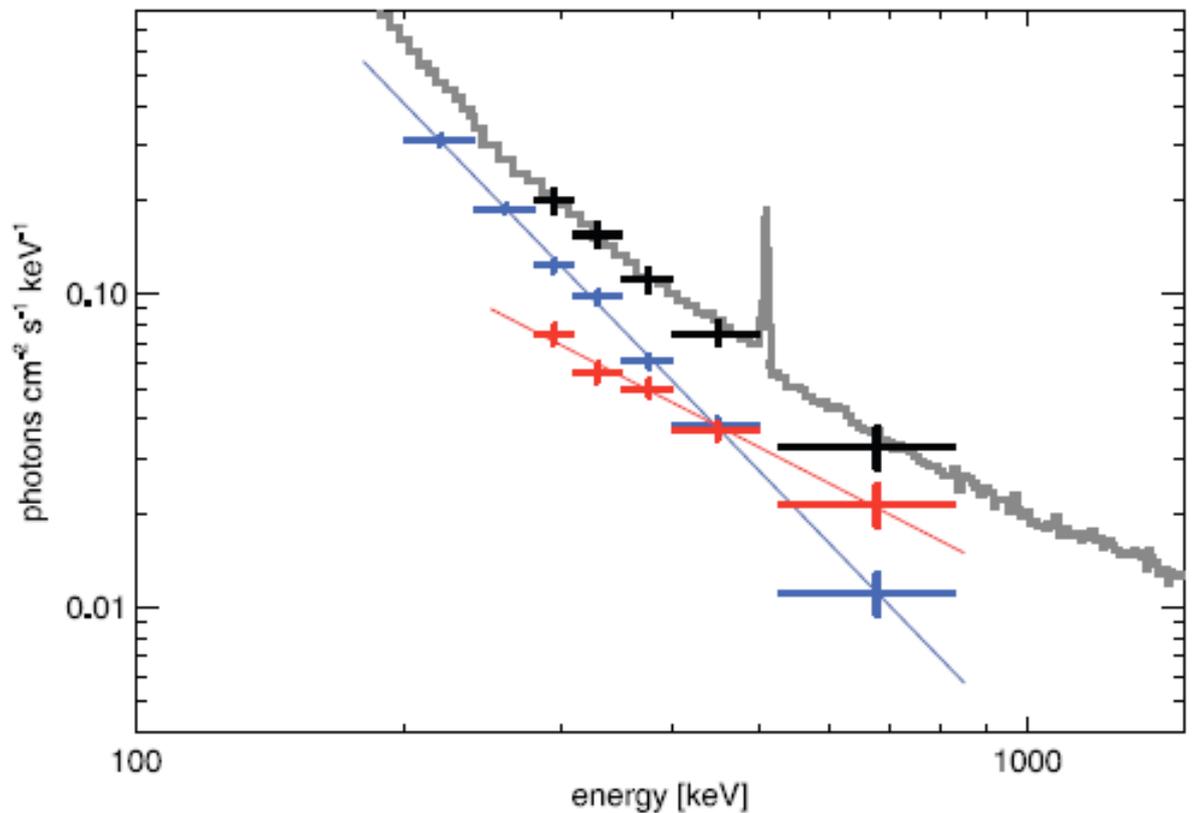


Figure 4.11: Figure taken from [Krucker et al. \(2008b\)](#) showing the imaging spectroscopy of the 2005 January 20 flare. The spectrum of the coronal source is given in red and that of the footpoint source is given in blue; their sum is shown in black with the spatially integrated high resolution spectrum shown in grey. As you can see, the spectral index of the coronal source is hard,  $\approx 1.5$

Continuum  $\gamma$ -radiation in the 100 MeV energy range was observed from this flare by the SONG instrument on CORONAS-F. There is evidence for a pion decay contribution to the observed spectrum ([Kuznetsov et al. 2005](#)), which would also indicate the presence of  $\sim 100$  MeV positrons. The flare was located towards the limb ( $N14^\circ W61^\circ$ ;  $\sin \lambda = 0.88$ ), maximising the likelihood of observable ICS photons. Moreover, the location of the coronal X-ray source is high enough that ICS could be the dominant source of HXRs, given sufficient energetic electrons. The coronal source has a very hard spectrum, photon spectral index  $\approx 1.5$ , consistent with the spectra found in the previous section. A photon spectral index of 1.5 would imply a relativistic electron spectral index of about 2. Continuation of this photon spectrum to at least 700-800 keV implies an electron distribution continuing in this power-law form to at least 120 MeV. To account for the observed coronal source fluxes shown in Figure 4.11 ([Krucker et al. 2008b](#)), we would need  $\sim 10^{31}$  electrons instantaneously present above 0.5 MeV.

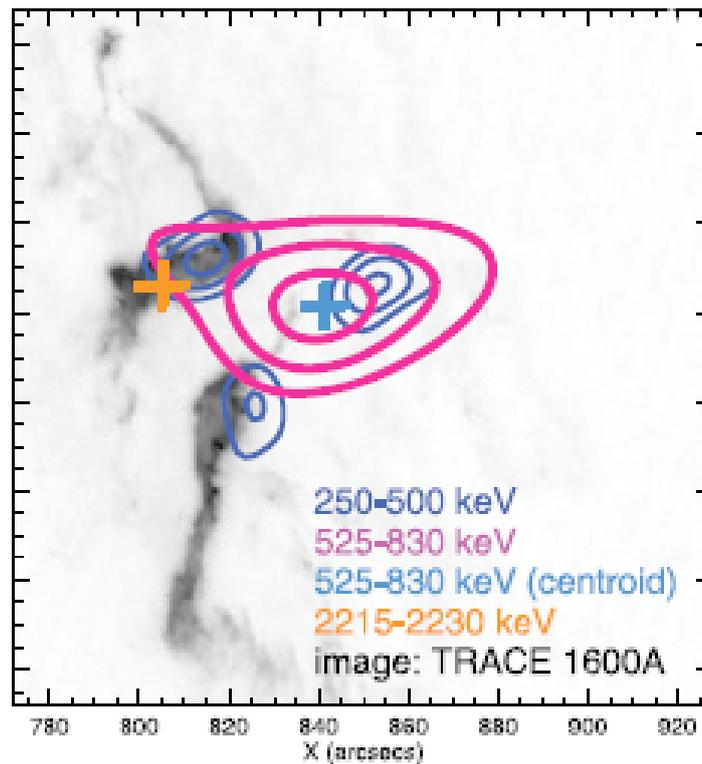


Figure 4.12: TRACE image of the 2005 January 20 flare taken from [Krucker et al. \(2008b\)](#) overlaid with RHESSI contours (50%, 70% and 90%) showing the regions of non-thermal emission. A strong HXR coronal source is clearly observed

The  $\sim 500$  keV source represented by the 50% contour of Figure 4.12 ([Krucker et al. 2008b](#)), is about  $40 \times 80$  arc seconds. Assuming a similar length scale along the line of sight we estimate its total volume as  $5 \times 10^{28}$  cm<sup>3</sup>. Taking for illustration an ambient electron density  $10^8$  cm<sup>-3</sup> we see that the relativistic electrons necessary to account for this source via ICS would represent just  $2 \times 10^{-6}$  of all particles in the volume, which is a perfectly plausible and realistic fraction in view of current theoretical acceleration models. We also estimate that this is  $\sim 10^{-3}$  or less of the electrons  $> 0.5$  MeV implied by a typical, large X-ray burst, assuming that the HXR deduced distribution extends into the relativistic energy range. The minimum energy of 0.5 MeV is of course quite arbitrary; only electrons in the 10s of MeV range and above are demanded by an ICS interpretation of this coronal HXR source.

Close to the limb, the most favourable head-on collisions of electrons with primary photons may occur. The flux and spectrum are very close to those given traditionally for power-law electron distributions and isotropic primary photons (e.g. [Blumenthal & Gould 1970](#); [Krucker et al. 2008a](#)), with modifications resulting primarily from the presence of an upper electron cut-off energy. The number and energy distribution of

electrons found above are close to those that would be found using the traditional results; but this would not be the case for an event further from the limb.

An interpretation in terms of positrons - produced as secondary particles as explained in §4.1 - is also possible. The spectra shown in Figure 4.8 would all give approximately the necessary, hard spectrum in the several hundred keV energy range (although, as discussed above, none has precisely power-law form). For a power-law primary proton energy distribution with energy spectral index = 2, about  $10^{32}$  protons would be needed above 1 MeV. Masson et al. (2009) found  $2.3 \times 10^{31}$  protons above 30 MeV for this event with proton spectral index = 3, i.e.  $2 \times 10^{34}$  protons above 1 MeV. Just a few percent of the secondary positrons produced by this ion population would be enough to account for the coronal source in this flare. Pion decay radiation, which will be accompanied by the production of positrons, often extends over long periods during and after the flare decay phase as viewed in other wavelengths (e.g. Ryan et al. 2000). The persistence into the decay phase of the 2005 January 20 flare coronal HXR source would be consistent with this behaviour.

## 4.6 Conclusions and discussion

ICS needs extreme source parameters if it is to account on its own for the bulk of flare hard X-rays (Korchak 1971; McClements & Brown 1986), particularly when ‘foot-point’ source morphology points to an origin in the dense chromosphere. Our work does not revise this view, just points out that ICS might be important for understanding sources in the tenuous corona. We have seen that very modest numbers of electrons or positrons at relativistic energies could account for already observed coronal HXR sources. Electrons would need to be accelerated into the 100 MeV energy range; positrons are automatically produced with the necessary energies as long as there are  $\sim 300$  MeV protons to produce them in the first place. The electron distribution needed to account for HXR bursts, extended into the 100 MeV energy range, would include enough relativistic electrons that only a small fraction of them would need to be found in the corona to account for at least one, observed coronal HXR source. Moreover, electrons might be accelerated to relativistic energies via a process distinct from the main flare energy release, as appears to occur in the Earth’s magnetosphere (Baker et al. 2001).

Do we need to contain electrons in the corona to produce such sources? The calculations above assume that an isotropic population of electrons is instantaneously present in the source region. The overwhelming contribution to observed ICS, how-

ever, comes from electrons moving instantaneously towards the observer. Electrons could pass freely through the corona, following the field lines and emitting observable ICS HXRs as they pass through the line of sight towards the observer. They would not need to be contained in the corona, and an isolated coronal source might be more naturally explained in this way, as a consequence of relativistic beaming and source magnetic geometry. Instantaneous numbers of electrons needed would be comparable to the numbers found above. A more detailed treatment of electron and positron transport, not given here, would be needed to assess this possibility properly.

The numbers of particles needed to produce ICS coronal HXR sources are not unreasonable, compared with those already known by other means to be present in flares. Such sources may already have been observed, though possibly difficult to distinguish from bremsstrahlung HXRs from electrons with 10s - 100s of keV energies. If definitively recognised in flares, they would open a new window on acceleration and transport of electrons and ions in the 0.1 - 1 GeV energy range.

How might we distinguish these sources from conventional bremsstrahlung HXRs? First of all, they may be expected from locations where the ambient density seems too low for a conventional, bremsstrahlung interpretation. As we have seen, coronal ICS sources should be brightest near the solar limb. If many sources like those described in [Krucker et al. \(2008b\)](#) can be detected, an ICS interpretation would imply a strong centre-to-limb variation. Simultaneous observations from two, widely separated spacecraft (e.g. [Krucker et al. 2008b](#)) would reveal quite different fluxes and spectra. The spectra will always be very hard, possibly also extending to soft X-ray and EUV ranges in a continuous way difficult to account for by other means. Observations of co-spatial radio radiation would have very different spectral properties in the bremsstrahlung and ICS cases.

# Chapter 5

## FNIT - the Fast Neutron Imaging Telescope

This chapter describes the principles of operation of multiple scatter detectors. As part of a team involving the University of Glasgow, the University of New Hampshire (UNH) and the University of Bern, I participated in calibration of a test model of such a detector at the Crocker Nuclear Lab (CNL), at the University of California, Davis. I also developed software to reconstruct incident neutron arrival directions from calibration run data, and thus to the overall evaluation of the prototype detector. This chapter describes the work carried out in detector development in this collaboration, highlighting my own contributions as they appear. As well as describing work in which we participated, this chapter forms essential background to the work of Chapter 6, which in turn aims to develop methods for maximal exploitation of neutron data. So a lot of the work entailed in this chapter and the next involves a collaborative project with UNH and the University of Bern in Switzerland.

### 5.1 Introduction

Development of FNIT has been undertaken at UNH over the last decade. No neutron detector has even been flown on a deep space mission and all neutron measurements in space so far have been taken by *gamma*-ray detectors. Some neutron detectors, like SONTRAC, have been test-flown on balloons but none have been deployed any further than a few km above the Earth's surface. Neutrons from space have of course been detected on Earth by ground-based neutron monitors as elucidated in Chapter 1.

The success of  $\gamma$ -ray instruments such as COMPTEL - which already achieved multiple scatter imaging and spectroscopy of both  $\gamma$ -rays and neutrons - and scintillation detectors like GRS, as well as the additional requirement to study neutrons closer to

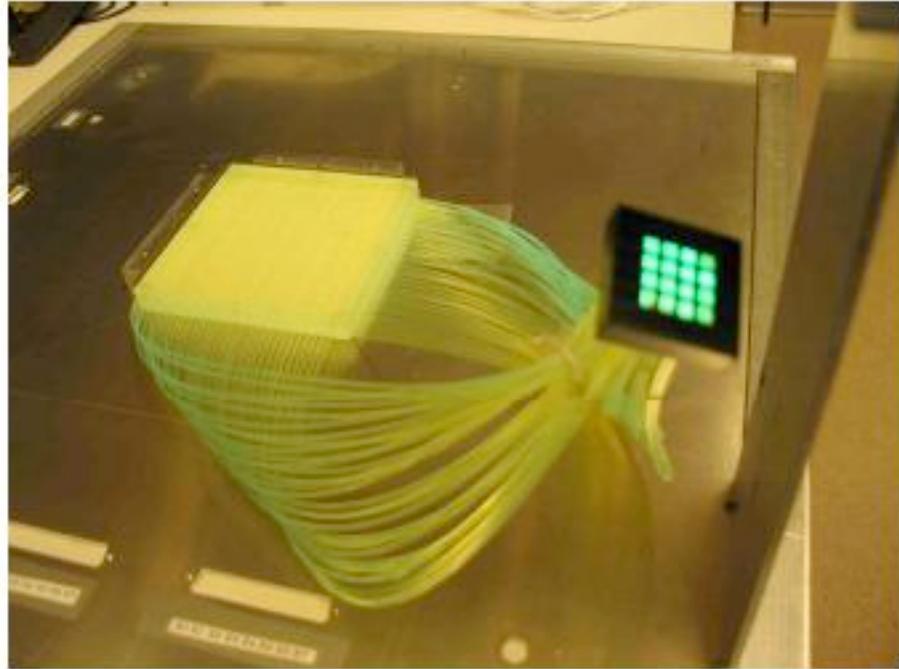


Figure 5.1: Picture of the first FNIT prototype with a plastic slab engraved with a grid of optic fibres. Courtesy UNH

the Sun, led to FNIT's conception. It was first decided that FNIT will be an instrument made of a stack of solid transparent plastic scintillator slabs, each about 10 cm long and broad and about 1cm thick (Figure 5.1) (Moser et al. 2005a; Bravar et al. 2005). Each of these slabs would be adorned with a grid of optic fibres separated by 1-2 cm. Each of these fibres would be connected to a photomultiplier tube (PMT), which in turn would produce the required signal that would be collected by the electronics of the instrument. A schematic diagram of the instrument is shown in Figure 5.2, where 8-10 of these slabs are stacked on top of each other. The concept here was that an n-p scatter in a slab would result in a scintillation signal within it. The location of the scatter would be determined from the signal strength in each optic fibre of that slab. Then a subsequent n-p scatter in another slab would allow us to track the neutron as it travelled down the stack and thereby reveal its incident energy and direction. Several tests were undertaken with this instrument by varying the design parameters to obtain an optimally performing instrument. However, although the concept was sound, it seemed as though we were not able to optimise the instrument's performance to our liking. For instance, improving the spatial resolution involved putting more optic fibres in the grid and this would result in decreasing the signal because more of

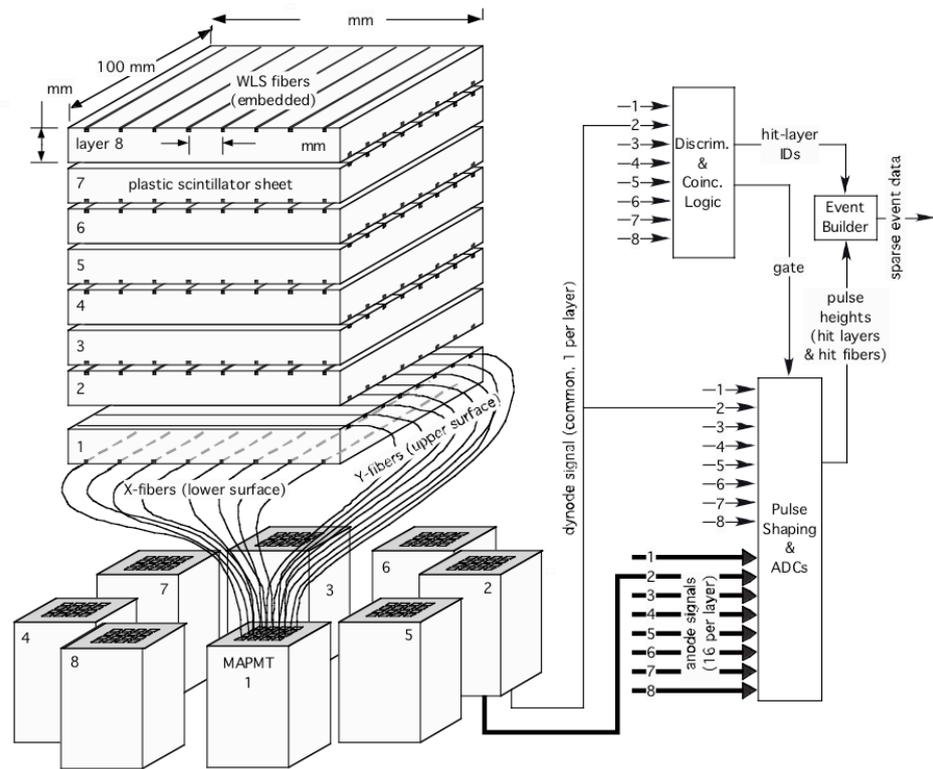


Figure 5.2: Schematic showing the principle of the first FNIT instrument with a stack of slabs. Courtesy UNH

the plastic would be covered by the fibres. On the other hand, widely spaced fibres meant compromising the spatial resolution. These were the kind of problems that we were faced with. Hence, it was decided to instead use organic liquid plastic scintillator instead of solid plastic, especially due to the added advantage of providing pulse shape discrimination (PSD) between neutrons and  $\gamma$ -rays. PSD, as seen in Figure 5.9, produces different signatures for neutrons and  $\gamma$ -rays because of their different decay times. Liquid scintillators had always been on the radar because their actual efficiency in detecting neutrons is higher than for solid scintillators. This is because liquid scintillators have a much higher proportion of hydrogen atoms in them and so n-p interactions become more likely, thereby reducing the likelihood of unwanted scatters off C nuclei and so on. However, using liquid brought with it other technical challenges like the containment of the fluid, which becomes even more tricky on a deep space mission, and so we had decided to test the solid scintillator prototype first. However, since the solid scintillator prototype did not perform to the levels that we required, it was decided to overcome the technical glitches and introduce liquid plastic scintillators. Using liquid obviously also required a complete change in the design of the instrument. In the rest



Figure 5.3: The new FNIT prototype with 3 rods, each filled with liquid plastic scintillator

of this section, it is this new prototype of the instrument that shall be discussed (Ryan et al. 2007, 2008; Macri et al. 2007; Bravar et al. 2007; Woolf et al. 2008, 2009; Pirard et al. 2009).

Liquid scintillator is put in hollow aluminium tubes, about 15 cm tall and 1.5 cm in diameter and these are arranged in a cylindrical formation, each tube about  $15^\circ$  apart with the diameter of the entire instrument being about 15 cm. This cylindrical structure is not only compact and sleek, which enables it to be a candidate instrument for future Solar Sentinel or Solar Orbiter missions, but also maintains its symmetry when put on a rotating spacecraft - Sentinels are envisaged to be rotating spacecraft for thermal control - and gives a full  $360^\circ$  field of view. Both the ends of each tube is connected to a PMT, which then connects with the instrument circuitry. The present prototype built at UNH that is being used and tested is shown in Figure 5.3 and a

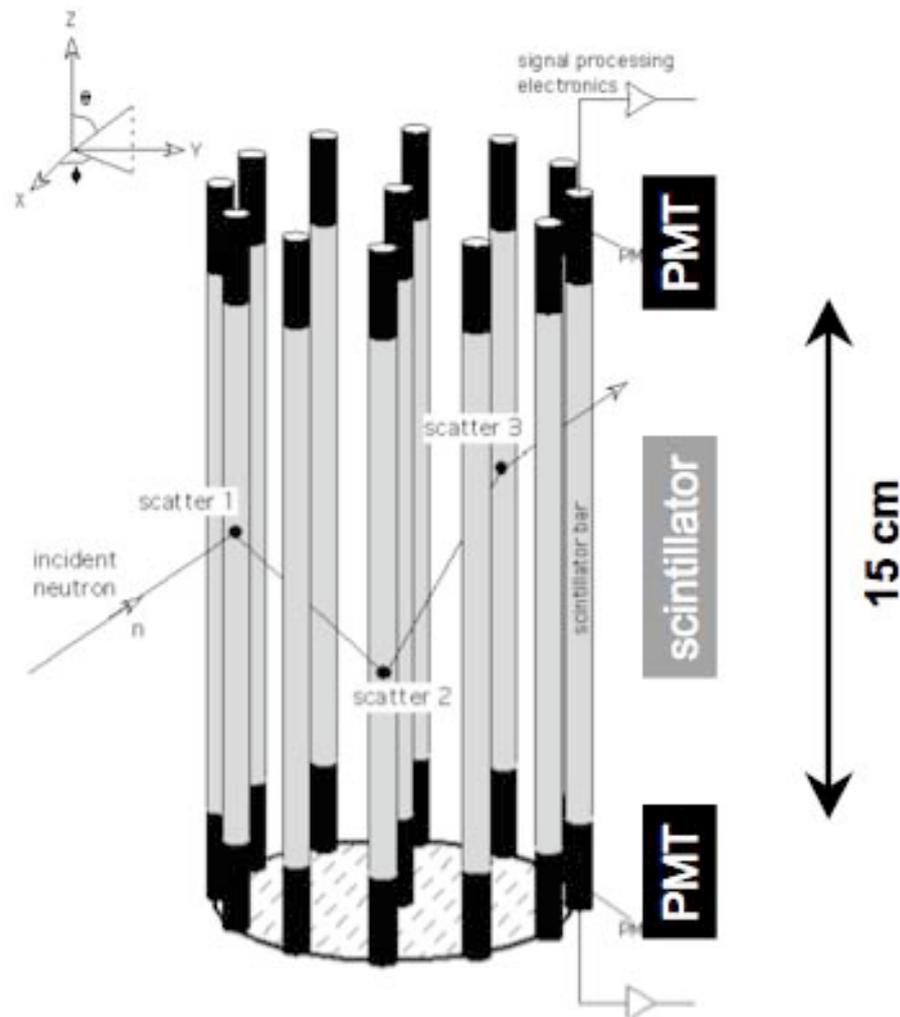


Figure 5.4: A schematic of the new FNIT prototype. Courtesy UNH

schematic of the full instrument is shown in Figure 5.4. To those technically minded, the organic liquid scintillator of choice is either NE-213A, BC-501 or BC-519. The numbers in the names of these plastics are just commercial catalogue numbers with the letters referring to the manufacturer: e.g. French-based St. Gobain uses the code ‘BC’ for all its plastics whereas Massachusetts-based Northeast Plastics uses ‘NE’. In the current prototype (Figure 5.3), we have used different scintillators in different tubes to monitor the performance. A graph of the detection efficiency for BC-519 is given in Figure 5.6. The inner wall of each tube is coated with a diffuse Teflon reflector to reduce the loss of the scintillation light.

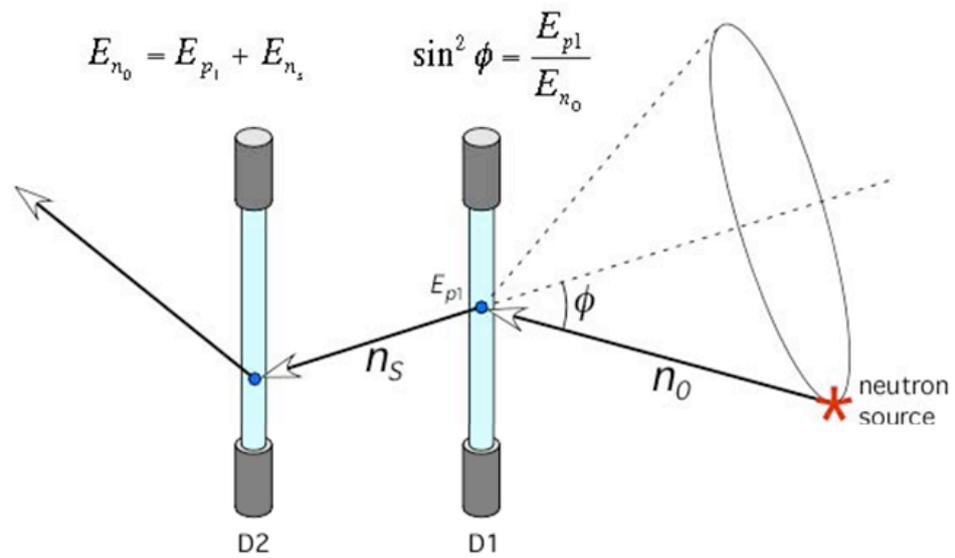


Figure 5.5: The double n-p scatter principle for the FNIT prototype with the neutron source location known. Courtesy UNH

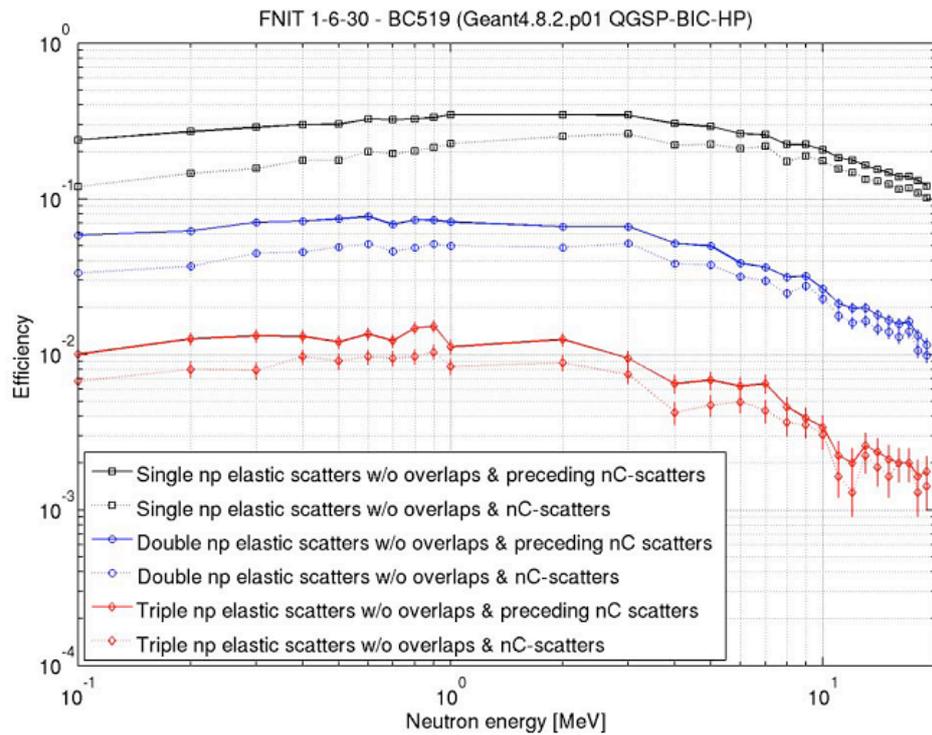


Figure 5.6: The efficiency of BC-519 as established by GEANT4 simulations. Courtesy UNH

## 5.2 Detection technique

In principle, the detection technique is simple. An incident neutron scatters in a tube and this results in a signal. Our choice of the scintillator ensures that most of these scatters are elastic n-p scatters. However, there are those unwanted scatters off heavier nuclei, both elastic and inelastic, like carbon. As discussed below, secondary  $\alpha$  particles and recoil carbon nuclei from these scatters are usually below threshold but they disrupt the trajectory of the neutron destroying the critical spatial information necessary to apply hard-sphere scattering kinematics.

The ratio of pulse heights from the two PMTs of a given rod is used to find the position of interactions along the length of the rod. For example, if the pulse height on both the bottom and top PMTs is equal, then you know that the interaction has occurred close to the middle of the rod. So using the principle that the signal falls off exponentially as you move away from the PMTs, it is possible to build an algorithm using the natural logarithms of the pulse heights to determine the position of the interaction along the rod.

Ideally, neutrons scatter in at least two distinct cylinders (e.g. Figure 5.5) and we call these ‘double’ scatters. But ‘single’ scatter, ‘double’ scatter and ‘triple’ scatter interactions all carry some information about the incident neutrons:

Single scatters: If the neutron flux from a solar flare is large, the rate of single scatters could easily be above the background rate. In such a case, a single scatter event could be assumed to be of solar origin. The measured pulse height provides a lower bound on the neutron energy. Such events could be used for crude spectroscopy if the rate is high, possibly refined via regularisation techniques, as described later.

Double scatters: Two successive n-p scatters are not sufficient to localise an unknown source, as in the case of gamma-ray astronomy. The second scatter provides no guarantee that the full neutron energy has been measured. However, if the full energy has been measured, an event circle can be computed. Additionally, with time of flight measurements, one can estimate the energy of the scattered neutron and thereby draw event circles for even double-scatter events if the location of the source is known, as explained later. However, without making this assumption and having a full energy measure, the radius of the event circle will be too large, implying that if the Sun was the source of the incident neutron then the Sun must lie within the interior of the expanded event circle. In this case the response signature of a neutron is an event disc on the image plane, the circumference of which is the computed event circle. The energy measure is, of course, a lower limit. Double scatter events in which the Sun lies out-

side the disc or event circle would not be candidate solar events. However, undetected carbon reactions between the first and second scatters will randomise the location of the event disc, constituting a background. If the event rate of double scatters is large so that one can assume with some confidence that the event has a solar origin, then the scattering geometry is known and the full energy can be computed if there are no interfering carbon reactions.

Triple scatters: Three n-p scatters allow a full measure of the neutron energy. With the geometry of the second scatter defined by the locations of the three interaction sites, the energy deposition of the second scatter fully determines the recoil neutron energy after the first scatter. Working backward, with the measured recoil proton energy in the first scatter and the kinematic energy measure of the neutron after the first scatter one has a full energy measure of the incident neutron. The response signature in this case is the ideal event circle. The circle's interior is not an allowable location for the source, i.e., the Sun. These circles constitute much less solid angle for the candidate source than the event discs do, thereby improving the signal-to-noise ratio. If the event is believed to be solar then the kinematics is over-determined and the data can be used to verify the solar origin assumption. The kinematics described here have been introduced in §1.3.2 and also discussed in detail in various papers on FNIT, e.g. [Moser et al. \(2005a\)](#).

Carbon scatters that occur after the first n-p scatter will produce subsequent scatter locations that are incompatible with the pulse height of the second scatter according to hard-sphere scattering. One can then reject these events. Carbon scatters that occur before the first scatter randomise the incident neutron direction because the scattered neutron will have an isotropic distribution. Carbon scatters that occur after the third scatter will not likely trigger the electronics and are thus irrelevant. For example, the location of the  $\gamma$ -ray burst (GRB) in Figure 1.13 was produced by the superposition of gamma-ray event circles, many of which have a radius that is too large because of incomplete energy absorption within the COMPTEL instrument. However, we see that the GRB is still located with some precision. Triple scatters within FNIT would intersect in the same manner in the direction of the Sun. Double-scatter events would do similarly but with a broader area of intersection.

So assuming we get an n-p scatter in at least two of the tubes, then by virtue of the signal strengths in each of the PMTs, we can determine where in each tube the collision occurred. Moreover, by knowing the exact relative positions of the tubes, we are able to draw a path between the two scatters and the instrument also notes the time of each scatter. Hence one can establish the time of flight (ToF) of the

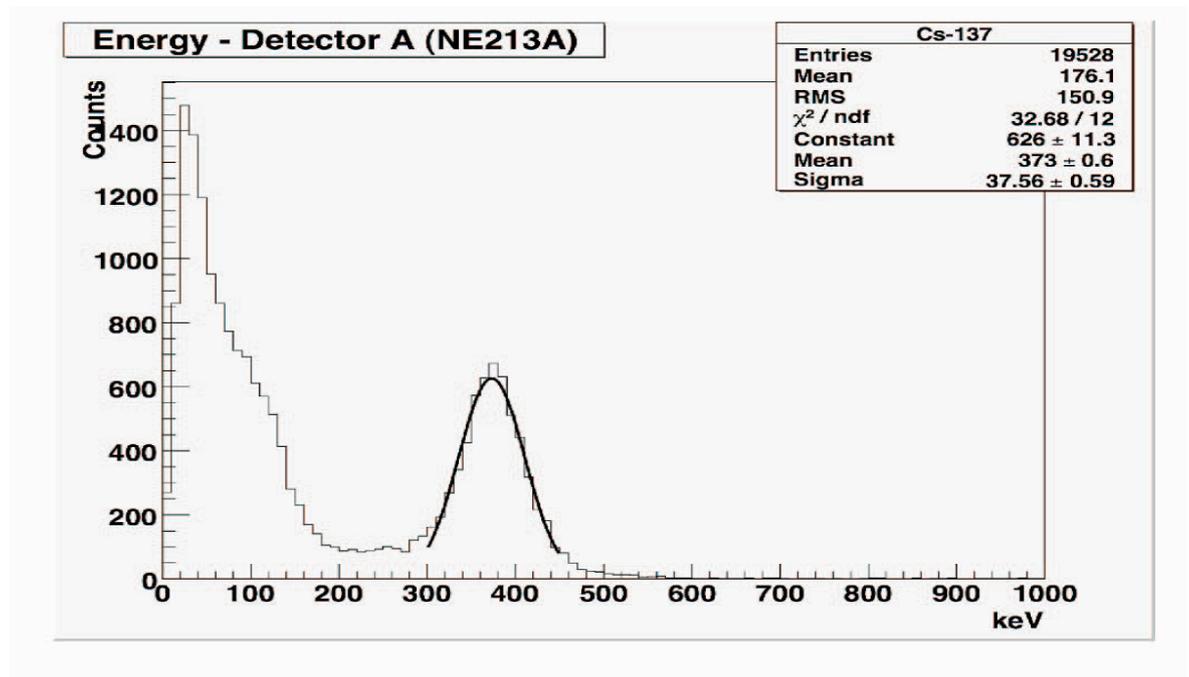


Figure 5.7: Energy resolution of the FNIT scintillator bar was obtained from  $^{137}\text{Cs}$   $\gamma$ -rays scattering in the detector at  $90^\circ$ . The central energy is 374 keV and the resolution was found to be about 10%. Courtesy UNH

scattered neutron, the distance between the scatters and the energy deposited in each tube. All this information easily allows us to calculate the energy (using velocity, i.e. the distance travelled divided by the ToF) and direction of the incident neutron by simple non-relativistic kinematic calculations (Equation 1.1). It also allows us to draw an event circle/disc for each such double (Figure 5.5) or triple scatter. Hence, if we collate many multiple scatter events, then we are able to draw multiple event circles, their intersection telling us the location of the source, and their energy allowing us to produce a neutron spectrum.

### 5.3 Laboratory and simulation data and results

FNIT has been tested in the laboratory at UNH with various  $\beta$  and  $\gamma$  sources like  $^{137}\text{Cs}$ ,  $^{60}\text{Co}$ ,  $^{90}\text{Sr}$  and so on by our colleagues there. This author also tested the first prototype FNIT instrument while studying at UNH as an undergraduate as well as contributing to the testing of the present prototype before and during the beam calibration run. FNIT was exposed to a neutron beam, produced by the bombardment of a  $^7\text{Li}$  target, at the Crocker Nuclear Laboratory (CNL) in the University of California, Davis, CA, USA in July 2007, where this author was present. Here, the instrument was subject

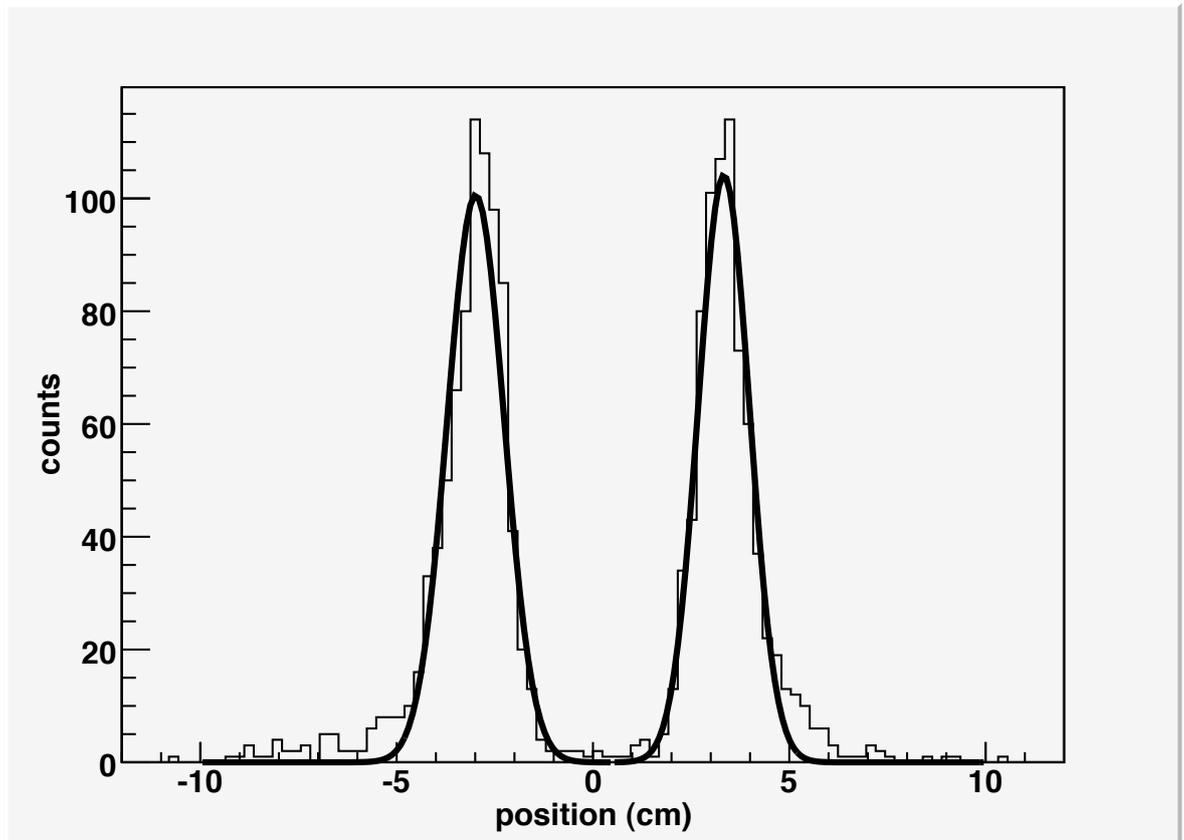


Figure 5.8: Collimated  $^{137}\text{Cs}$  photons located at 3.75 cm away from the centre of the bar (0 cm) on either side. Resolution was found to be 0.85 cm at these positions. Courtesy UNH

to repeated runs in a neutron beam of varying energies from 1-10 MeV. Immediately thereafter, FNIT has also been exposed to a small amount of weapons grade plutonium at the Pacific Northwest National Laboratory (PNNL) in Richland, WA, USA, where a couple of our UNH colleagues got permission to undertake the tests. Since then, it has been under continued testing at the lab in UNH with the strong neutron source of  $^{252}\text{Cf}$ . All this has given us a plethora of laboratory data, but that is not enough. In parallel, our colleagues at both UNH and the University of Bern have been using GEANT4 Monte Carlo simulations to model FNIT. GEANT4 is a toolkit for the simulation of the movement of particles through a given medium. It is used in various branches of physics, especially high energy, nuclear and accelerator physics as well as medical and space sciences. (For a general review of this simulation package, see [Agostinelli et al. \(2003\)](#)). Hence it is the perfect tool to use here as we track the passage of the neutron from the source to the detector and then through the detector.

It is the combination of all these data, namely  $\gamma$ -ray and neutron data from the

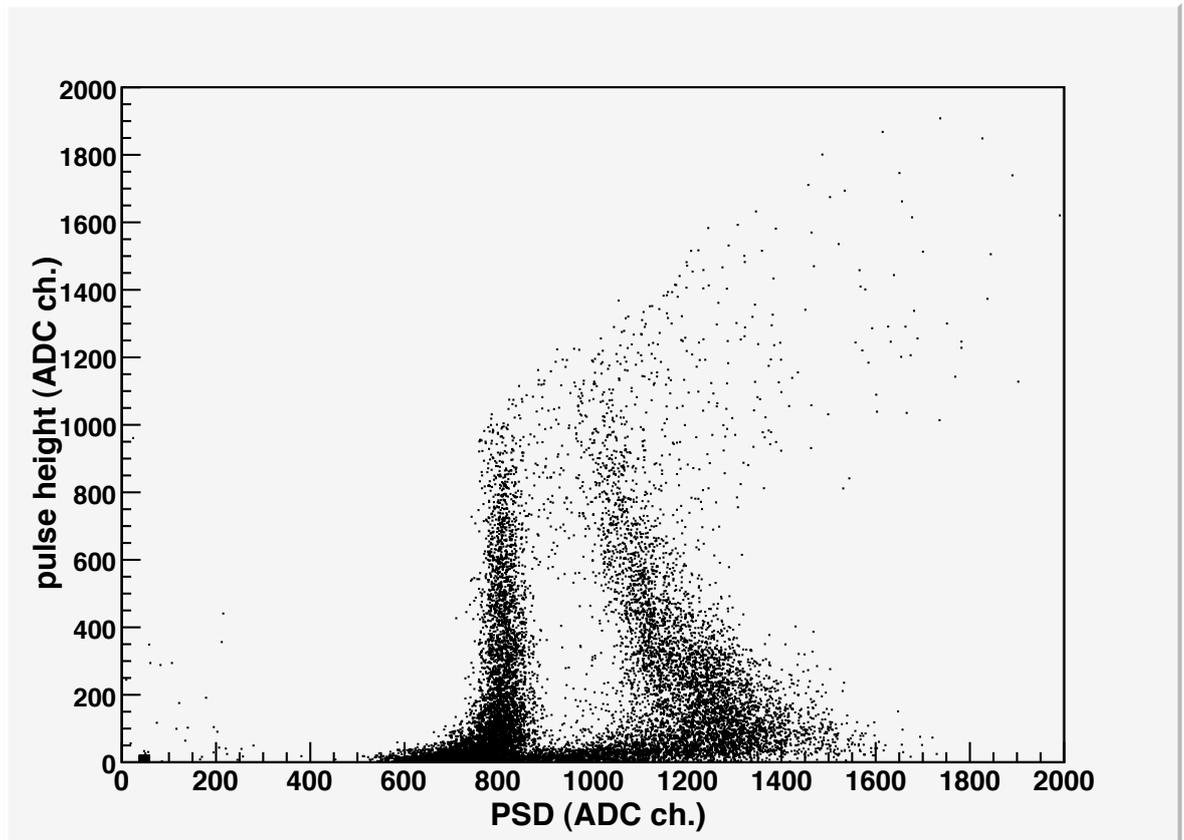


Figure 5.9: Scatter plot of pulse height vs. pulse shape clearly shows the separation between  $\gamma$ -ray-induced and neutron-induced events. Courtesy UNH

lab as well as GEANT4 simulation data, that will allow us to determine an optimum performance of the instrument. Consequently, this will allow us to produce a response matrix for the instrument and perform regularisation and deconvolution tests for it. With the present data we have, we are already able to make some of these preliminary calculations and tests, which will be summarised in the next chapter.

A combination of lab and simulated data have allowed us to make various instrument parameter measurements such as energy, position, angular and timing resolution (derived from multiple scatters) as well as pulse shape discrimination (PSD), which is a property of individual scintillations. These tests have been carried out at UNH over the years. Energy calibration was conducted by exposing FNIT to a  $\gamma$ -source like  $^{137}\text{Cs}$ . Figure 5.7 shows a central energy of 374 keV and an average energy resolution of  $\delta E/E \approx 10\%$  was achieved. Using a collimated source of 662 keV photons impinging on various locations of the rod, an average position resolution of 0.9 cm was also achieved. The performance was best close to the centre of the rod, when both PMTs contribute equally. Figure 5.8 shows one such position resolution measurement

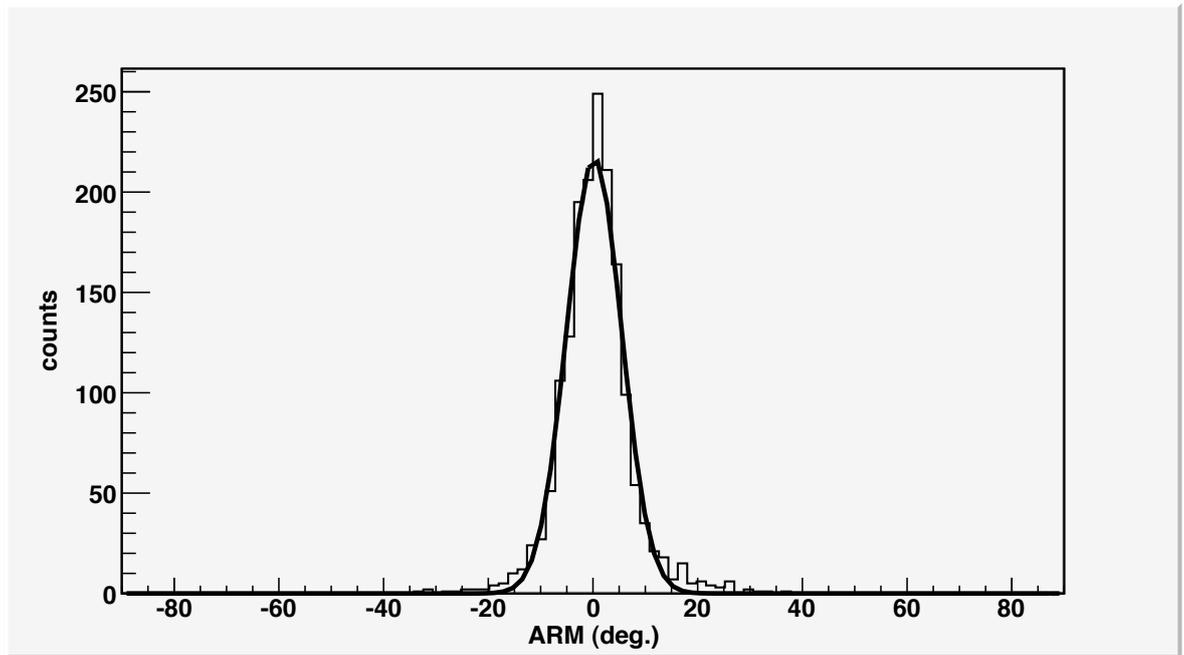


Figure 5.10: Angular resolution measure of 2 MeV neutrons produced with beam data from CNL. The ARM was found to be  $5.3^\circ$  at this energy. Courtesy UNH

3.75 cm from the centre. ToF measurements were made using simultaneously emitted photons from  $^{60}\text{Co}$ , a  $\gamma$ -source, which is placed equidistant between two detectors of the instrument. ToF is the time taken for a particle to travel between each scatter. Since the instrument records the time of each scatter, it is trivial to obtain the ToF. Hence, using the ToF and the distance between the scatters, which too can be deduced from the data by knowing the location of the scatters, it is possible to determine the energy of the scattered neutron. This yielded a timing resolution of 1.3 ns. This also shows why ToF can be used effectively only if the time between scatters of a neutron is greater than the timing resolution. Hence for high energy neutrons and/or closely spaced scatters, ToF becomes less effective. ToF capabilities are shown in Figure 1.11, where photons register a near-zero ToF, neutrons are measured several nano-seconds later whereas the negative ToFs map the back-scattered neutrons.

The scintillation light outputs of neutrons and  $\gamma$ -rays have different decay-time behaviour, so Pulse Shape Discrimination (PSD) may be used to discriminate these species and thus to enable reliable interpretation of the total light output. PSD can be produced using the charge-ratio (Channel Number in Figure 5.9) of the fast to slow components of the scintillation light output (Bryan et al. 2003). This ratio is different for neutron and  $\gamma$  interactions in the liquid scintillator. This can be seen in Figure 5.9. Since ToF is not as effective a tool to distinguish neutrons from  $\gamma$ s at higher energies

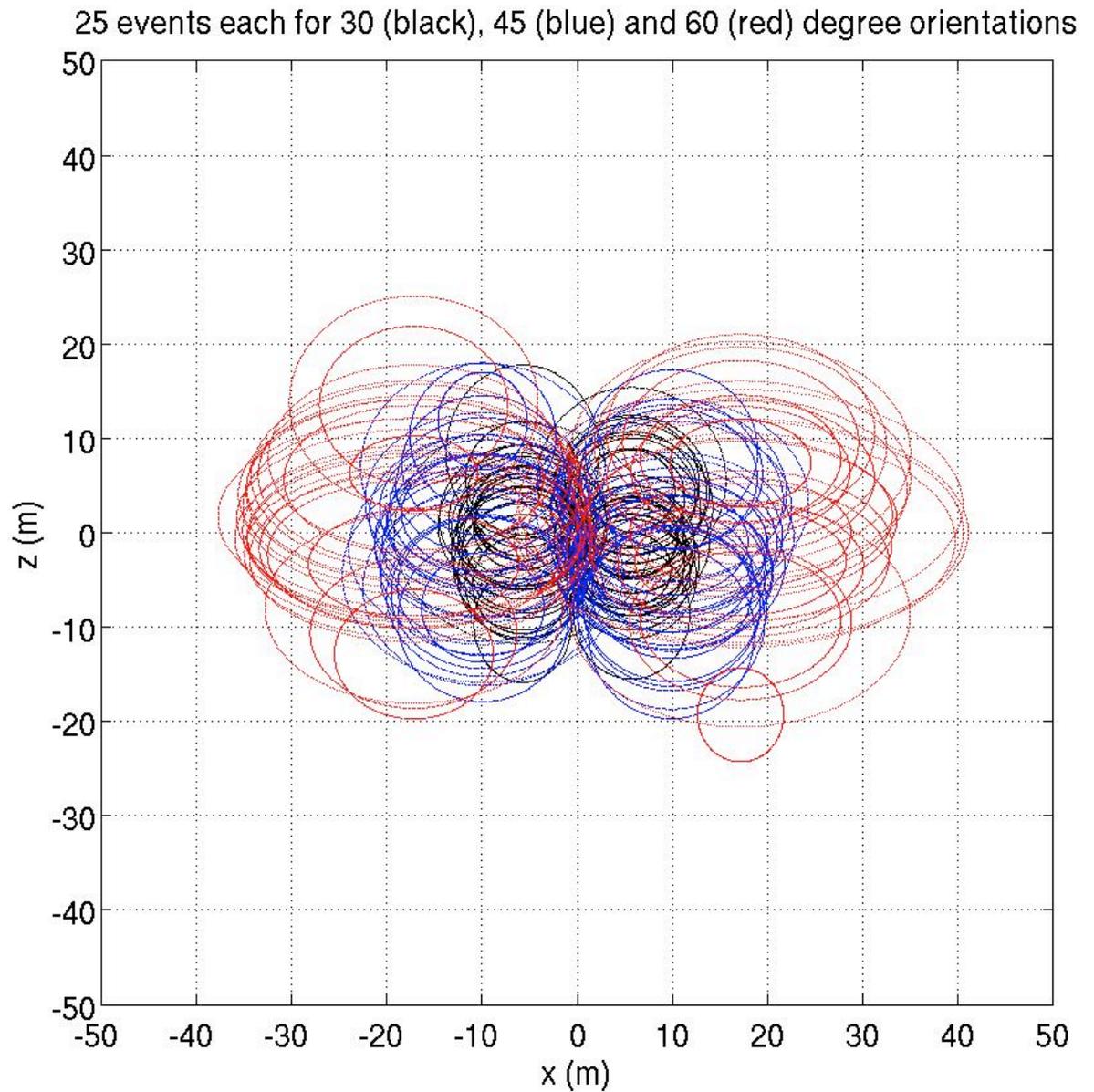


Figure 5.11: Event circles and a composite image of the 2 MeV neutron beam as compiled from measurements at a range of detector rotations

and for as compact an instrument as FNIT, PSD is used to select events at the higher energy end. Conversely, at low energies, ToF is a more effective tool. This is because at higher energies, the neutrons deposit more energy, thereby producing a better PSD.

The same characteristic instrument parameters can be measured using beam data from the CNL runs. At CNL, the energy calibration was conducted using a neutron beam source of 1.8-2.9 MeV and then further data was collected using more energetic neutron beams. These neutrons are produced by the bombardment of a  ${}^7\text{Li}$  target

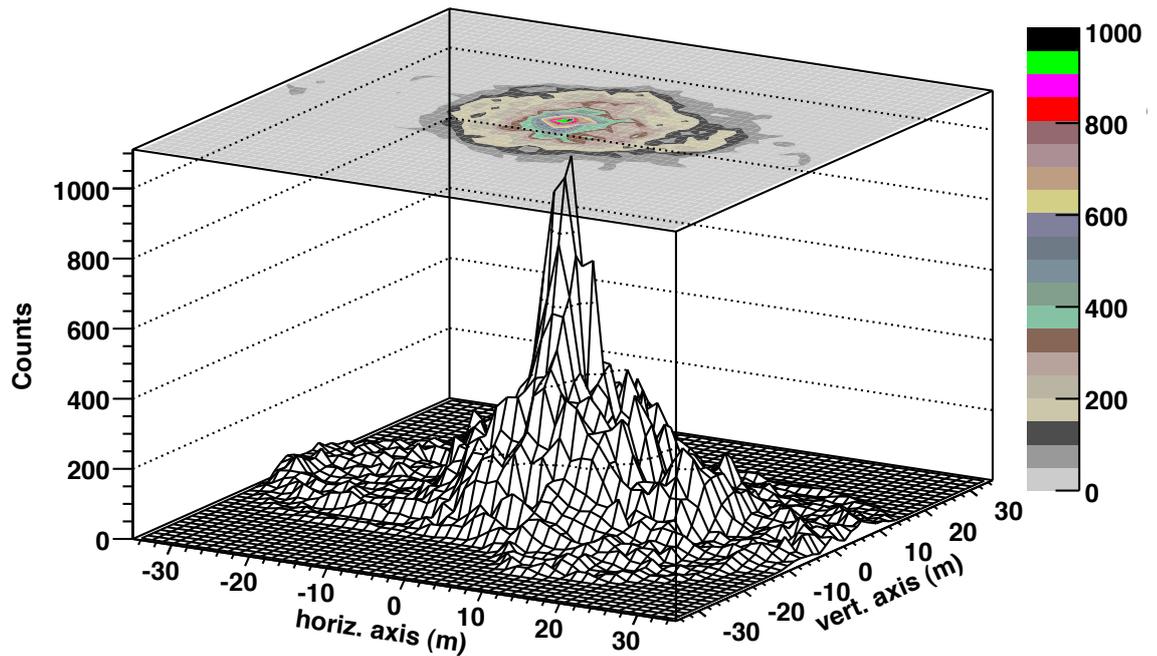


Figure 5.12: Neutron source imaging using FNIT. This image was obtained by reconstructing the event circles (Figure 5.11) of each neutron and projecting them on to an image plane. The height of each bin in the histogram represents that number of circles that intersect at that location on the image plane. Courtesy UNH

by energetic mono-energetic protons produced in a cyclotron. The neutron beam has a distribution of energies based on the thickness of the lithium target. A thicker target produces more neutrons but a broader distribution, so it was a tough task to determine the optimum thickness to get adequate neutron yield without compromising on the instrument calibration too much. The energy resolution was found to be  $\sim 20\%$  at 2 MeV and  $\sim 17\%$  at 10 MeV. Both the inherent detector properties as well as the spread of the incident neutrons produce this spread.

The angular resolution measure (ARM) is defined as the difference between the measured angle and the geometric (i.e. ‘true’) angle. The measured scattering angle, or kinematic angle, is obtained by using the equation in Figure 5.5 (Equation 1.1) and as is evident, its accuracy is dependent on the quality of spatial, energy and timing resolution of the instrument, since it is based on measuring the energy of the recoil proton (signal in D1) and the energy of the scattered neutron (obtained by determining the scatter locations in D1 and D2 and the ToF). However, in the case of neutron beam or laboratory runs, the relative location of the source and the detector is known and so we can use that information to calculate the geometric scatter angle. For the 2 MeV beam calibration run at CNL, an ARM of  $5.3^\circ$  was obtained (Figure 5.10).

The beam calibration runs also allowed us to image the neutron source. By knowing the distance of the detector from the source and drawing event circles for each double scatter, we were able to project these event circles on the source plane using Euler angles in the  $x$  and  $z$  directions. Euler angles are a means of representing the spatial orientation of any co-ordinate system as a composition of rotations from a reference co-ordinate system. Hence they were necessary to use here to ‘project’ the event circles on to a reference frame, in this case the source plane. Since these ‘circles’ are being projected onto a flat surface, their projection is usually an ellipse - as can be seen from Figure 5.11. In the lab at CNL, we kept the detector at a distance of 10m from the source for all our runs and oriented it at six different rotation angles around the  $z$ -axis. A look at Figures 5.3 and 5.4 illustrate the geometry in this context. FNIT is placed on a platform on the  $x$ - $y$  plane and is free to rotate around the  $z$ -axis. Meanwhile, the neutron source is a point on the  $x$ - $z$  plane some 10m away. The different rotation orientations meant that the three detector rods subtended different scattering angles for each of these runs. Event circles (or ellipses) were drawn and projected for all double n-p scatters detected for each of these orientations: 30 degrees and -30 degrees (black), 45 degrees and -45 degrees (blue) and 60 degrees and -60 degrees (red) as shown in Figure 5.11. It is evident that most of these circles intersect at the point (0,0), which we know to be the source location. However, there are clearly also some circles that seem to be completely ‘random’ and these must surely be unwanted scatters off carbon nuclei and so on. Also, the resolution is better in the  $x$ -axis than in the  $z$ -axis because the techniques used to determine the respective ordinate is different. Whereas the  $z$ -location is made purely by analysing the signal strength in the upper and lower PMT of the particular rod, the  $x$ -location is determined by the geometry and orientation of the instrument. The entire event-circle analysis and algorithm was undertaken by this author. A three-dimensional histogram produced by our UNH colleagues representing Figure 5.11 is shown in Figure 5.12, where the height of each bin represents the number of circles that intersect at a particular image-plane co-ordinate. Both these figures give a pictorial representation of FNIT’s point spread function for one incident neutron energy.

These results clearly show that using multiple-scatter n-p events, the FNIT instrument can determine quite accurately the energy and source of the incident neutron. However, to improve its performance, to understand its behaviour better so that we can fully exploit the data and to make it suitable for a space mission, a lot of other work has also been undertaken. The neutron and  $\gamma$  calibrations we have done are at particular neutron energies, which have already made clear that the instrument response

is not perfectly diagonal. We also might be able to extend FNIT's sensitivity into the 10s of MeV range by counting all the inelastic scatter events, at the price of a more non-diagonal response, which would need more sophisticated methods to analyse. In the next chapter, we shall talk about the work we have done on obtaining a response matrix and the deconvolution methods that we have used, but before that we shall briefly introduce the role of simulations that have been undertaken by our colleagues at UNH and Bern.

## 5.4 Simulations

An extended programme of GEANT simulations has demonstrated its validity, both reproducing the results of the laboratory calibrations or measurements made with several neutron detectors (FNIT prototypes, SONTRAC, UNH neutron telescope etc.) and possessing the capacity to provide well-defined response matrices over the past decade or more. A detailed Monte Carlo model of the FNIT instrument based on the CERN GEANT4 libraries ([Agostinelli et al. 2003](#)) has been developed at the University of Bern ([Moser et al. 2005a](#)). It allows one to simulate the interaction of radiation, in particular neutrons and  $\gamma$ -rays, with any given detector configuration under laboratory and space conditions. In addition to the standard GEANT4 elastic neutron scatter models, a new model based on Evaluated Nuclear Data File (ENDF) cross sections for hydrogen and carbon interactions above 20 MeV was implemented, which was originally developed and successfully tested for SONTRAC Monte Carlo simulation purposes ([Desorgher et al. 2003](#)).

In order to validate the Monte Carlo code, the configuration of a previous atmospheric neutron telescope, also developed at UNH ([Saxena 1990](#)), was implemented. By unfolding measurements in the 10-170 MeV regime taken in a campaign during solar minimum in 1987 with the simulated instrumental response matrix, the mid-latitude cosmic ray neutron energy spectrum and angular distribution at sea level could be determined in excellent agreement with data and predictions by other authors ([Moser et al. 2005c](#)).

The Monte Carlo results allow one to determine the characteristics of the neutron interaction processes for stacked layer and cylindrical symmetric configurations with different dimensions, spacings, number of scintillators and scintillator materials. Furthermore, it is possible to simulate scintillation photon propagation through the material until they are absorbed and, shifted in wavelength, re-emitted within the read-out fibres. According to laboratory performance measurements, raw simulation data are

artificially broadened and by applying detection thresholds the instrumental response matrix is computed. From this matrix, important information such as energy and angular resolution are obtained, methods to identify and reject background neutrons and photons are tested, and for the development of deconvolution algorithms it serves as a reference. Together with experimental data, Monte Carlo results help optimising the set of design parameters where one has to find a tradeoff between complexity and detector performance.

On a spacecraft, the instrument is steadily exposed to galactic cosmic rays (GCR) and their secondary particles produced in interactions with the surrounding spacecraft material. During a solar flare the background signal may well exceed the GCR level due to the arrival of solar cosmic rays (SCR) and their induced secondary products. In order to estimate the background particle flux entering the detector volume and to determine the signal-to-noise ratio, it is necessary to model the temporal evolution of the SCR spectrum in the inner heliosphere. This can be done either based on theoretical propagation models or on available spacecraft data. However, the latter have to be scaled in intensity and time to the planned spacecraft orbit. Whereas the first ones may only represent ideal cases. The subsequent convolution of these time dependent spectra with the spacecraft and the instrument response will lead to the counting characteristics during a solar flare. Based on this result it is then possible to estimate the relevance of an external anti-coincidence shield and even the necessity of an additional surrounding low-energy particle absorber. Furthermore, these simulations will then allow one to develop and validate highly efficient methods to discriminate background neutrons from solar flare neutrons, e.g., by using the directional information obtained from multiple scatter events.

[Pirard et al. \(2009\)](#) conducted GEANT4 simulations for FNIT's current prototype instrument with measurement from our beam calibration run at CNL. Simulation results, taking into account a limited number of empirical parameters (ToF, position and energy resolutions, and energy threshold and conversion laws), showed very good consistency with experimental data in terms of double scatter event selection and differential distributions. Efficiency levels could be reproduced reliably with simulations. Above all, variations of double scatter efficiencies with the neutron beam energy and orientation are fully understood: the maximum efficiency has been observed when the two-bar system is oriented with an angle of 30 degrees with respect to the beam, and this is the case for all energies. The finite pulse-height threshold in each single scatter detection represents the main limitation to observe np double scatters for small and large angles at a given energy, and also drives the overall lower energy limit of the

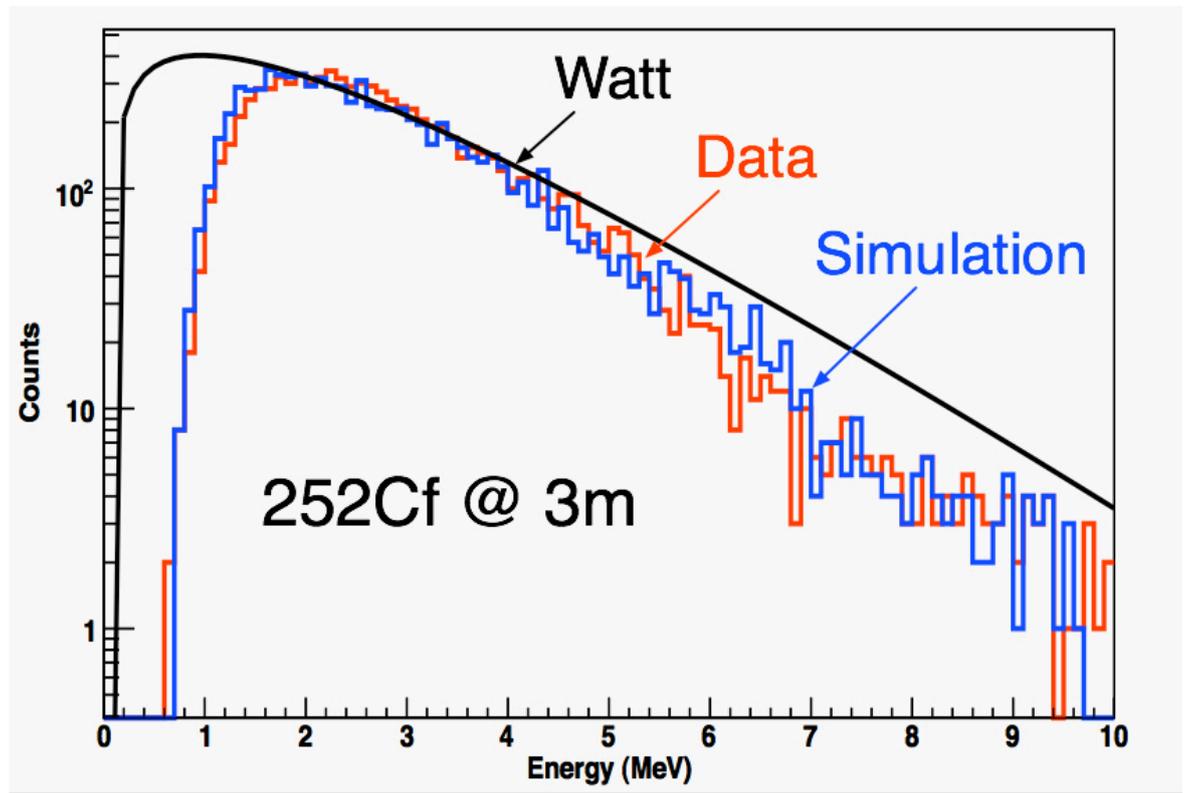


Figure 5.13:  $^{252}\text{Cf}$  spectra: input theoretical distribution (black), uncorrected measured (red) and simulated (blue). Courtesy UNH

detector, evaluated at 0.7 MeV.

Simulations also allow you to evaluate the instrumental response to truly monoenergetic neutrons, and showed us that the resolution (r.m.s.) of the reconstructed energy ranges from 10% to 27%, which can be lowered to 15% with appropriate double scatter event selection. Similarly, simulations and calibrations show that the ARM resolution increases monotonically from  $4^\circ$  to  $11^\circ$ . Without any sophisticated image reconstruction method, nor with strong selection of the double scatter events, the instrument exhibits an overall spatial resolution not greater than  $11^\circ$  (r.m.s.). This was obtained using the simple event-circle superposition to reconstruct the instrument point spread function for the 2 MeV datasets shown in Figure 5.11.

Finally, this study validates the numerical approach to predict the response and to optimise the design of more complex fast neutron telescopes. The simulation tool could also be used in the future not only to support experimental data analysis but also to provide synthetic data sets for imaging algorithm development.

An instrument like FNIT can have several applications. Apart from its use on space missions for solar neutron astronomy, it has various applications on Earth as

well. Since radioactive materials emit neutrons in this energy range (1-10 MeV), FNIT can be used for environmental, safety, security and defence purposes to detect special nuclear materials (SNM). To this end, FNIT was tested at PNNL, where it was kept  $\sim 1\text{m}$  away from 98g of weapons-grade plutonium, i.e. 5.6% Pu-240. Like any fissile material, the neutron distribution emitted by the  $^{239}\text{Pu}$  nucleus follows the Watt energy distribution (Watt 1952):

$$N(E) = K(A, B)e^{-AE} \sinh \sqrt{BE}, \quad (5.1)$$

where  $A$  and  $B$  are the nucleon and temperature-dependent Watt parameters as given by e.g. Ethvignot et al. (2003). The median energy for fission neutrons from this  $^{239}\text{Pu}$  source is 1.6 MeV and the mean energy is 2.1 MeV. Hence FNIT is an ideal instrument to detect dangerous substances such as these and other SNM that may have similar energy distributions. In an era of heightened security and safety concerns and the threat of nuclear materials being smuggled, FNIT has vast commercial application potential.

To understand better the performance of FNIT when exposed to fissile material, it has been exposed to  $^{252}\text{Cf}$ , a strong neutron source, in the lab at UNH. Unlike at CNL, where FNIT was exposed to a quasi-mono-energetic neutron beam,  $^{252}\text{Cf}$  produces a Watt energy distribution (Equation 5.1; Figure 5.13). It has a half-life of a little over 30 months and is a strong neutron emitter:  $1\ \mu\text{g}$  produces 2.3 million neutrons per second, which reflects its 3.1% probability to undergo spontaneous fission. 96.9% of the time, it is an  $\alpha$ -emitter.  $^{252}\text{Cf}$  has a most probable emitted neutron energy of 0.7 MeV and an average energy of 2.1 MeV. Laboratory data as well as GEANT4 simulation data have been assimilated for FNIT when exposed to  $^{252}\text{Cf}$  and a plot of this is shown in Figure 5.13. The results are encouraging in that both the lab and simulation data not only agree with each other closely but also loosely agree with the theoretical Watt distribution. The discrepancy between the theoretical curve and the simulated and data curves is due to the following: at low energies the disagreement is due to the threshold energy of the instrument which requires neutrons above a particular energy to register a scatter; and at high energies there are statistically just not enough high energy neutrons produced by  $\text{Cf}^{252}$  in the lab to validate the Watt distribution. Also, the performance of FNIT starts to deteriorate as neutron energies approach 10 MeV. Hence, all in all, at FNIT's optimum neutron detection energy of 2-5 MeV, all three curves agree the closest.

## 5.5 Characterising detector response

Using laboratory and simulated data, it is possible to obtain a response matrix or point spread function for instruments such as FNIT. We have used both types of data to obtain the instrument response. As explained earlier, [Pirard et al. \(2009\)](#) used the event circles we generated from the CNL calibration run to obtain a point-spread function. However, using simulated data from the same setup, we have been able to obtain a preliminary instrument response for FNIT. It is important to distinguish here between simulated data and data that we would expect from space. In the simulated data, we know the entire detail of the detected neutron, i.e. its initial energy, energy deposited in each detector, location of scatters, type of scatters (e.g. n-p elastic or n-C inelastic etc.), ToF and so on. However, from real-time observations, the raw data will only include scintillation signals from each detector and the timing information. From those data, we will have to determine the nature of the interaction, energies deposited and so on. However, using simulated data, we can select the information such that it reflects the data we would get from the instrument, analyse the data, and see how it compares with what we already know to be the complete information from the simulations.

All the plots herein - Figures [5.14](#) to [5.18](#) - have been produced using GEANT4 simulations. UNH and Bern conducted the simulations and then shared the data produced, which have been analysed by them and by us independently. The resulting plots produced by UNH and by us were compared and verified before being reproduced here. The analyses too have been done in parallel.

The aim is to produce an energy response matrix for the instrument, whose elements represent the probability of a neutron with a particular incident energy giving a signal in the instrument appropriate to another, possibly non-identical energy. It is this ambiguity that makes the instrument response non-diagonal and the reasons for this are many: e.g. multiple scatters in either or both detector rods, inelastic scatters in either or both rods, finite ToF resolution, finite position resolution along the rods and so on. Through the simulation data, we are able to construct such a response matrix for FNIT, as shall be elaborated in [Chapter 6](#).

Since we know the incident neutron energy from the simulations ('energy in'), we can see how that compares with what the instrument tells us is the detected neutron energy ('energy out'). As explained earlier, this is determined using the energy deposited in the first scatter followed by ToF measurements and the scatter geometry to determine the energy of the recoil neutron. The sum of these energies gives us the total neutron

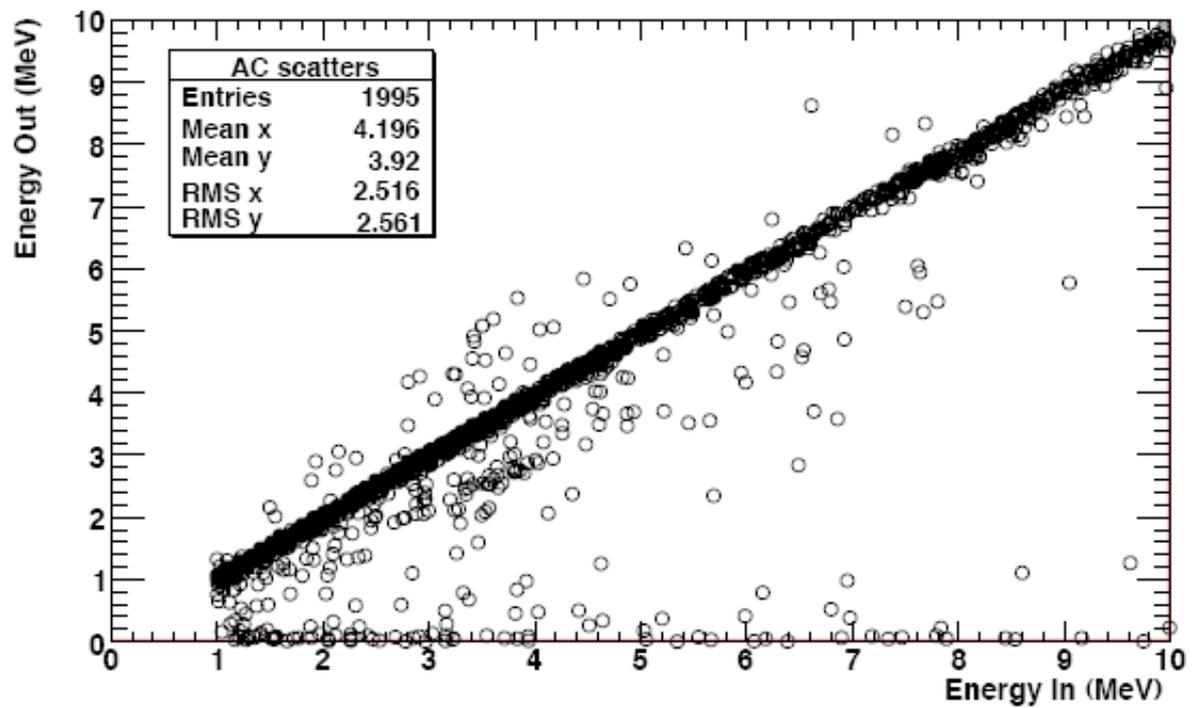


Figure 5.14: Scatter plot representing FNIT response for all double scatter events. Courtesy UNH

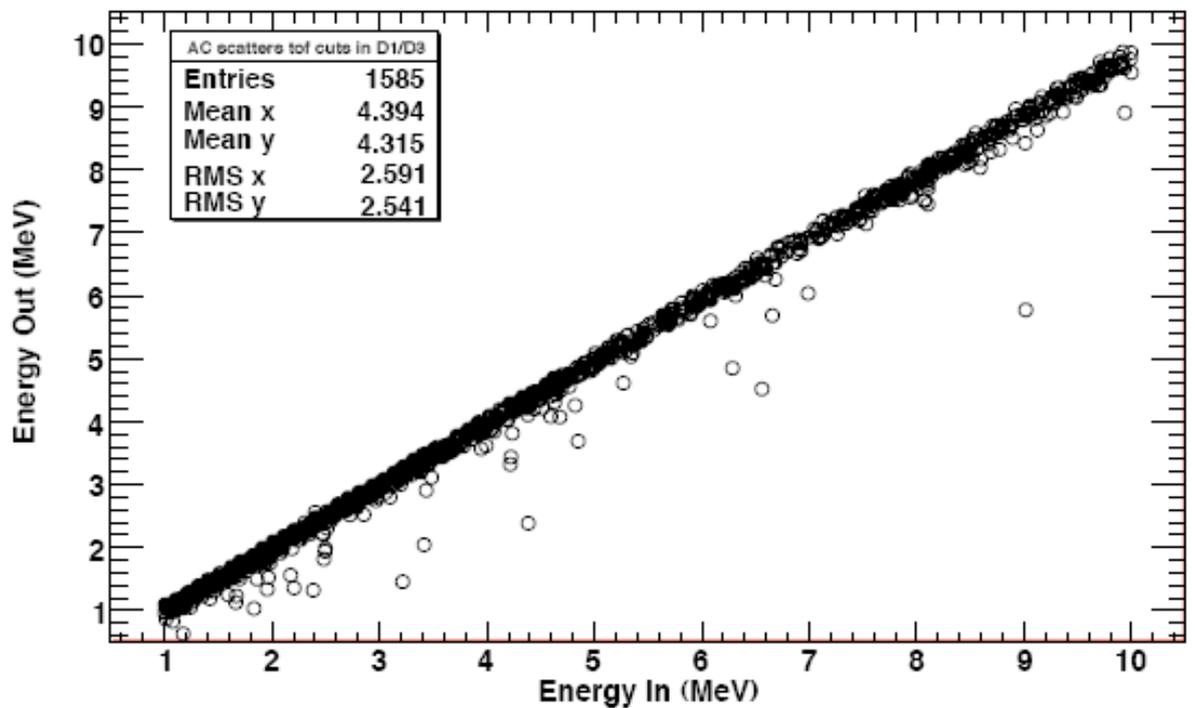


Figure 5.15: Scatter plot representing FNIT response with ToF cuts imposed. Courtesy UNH

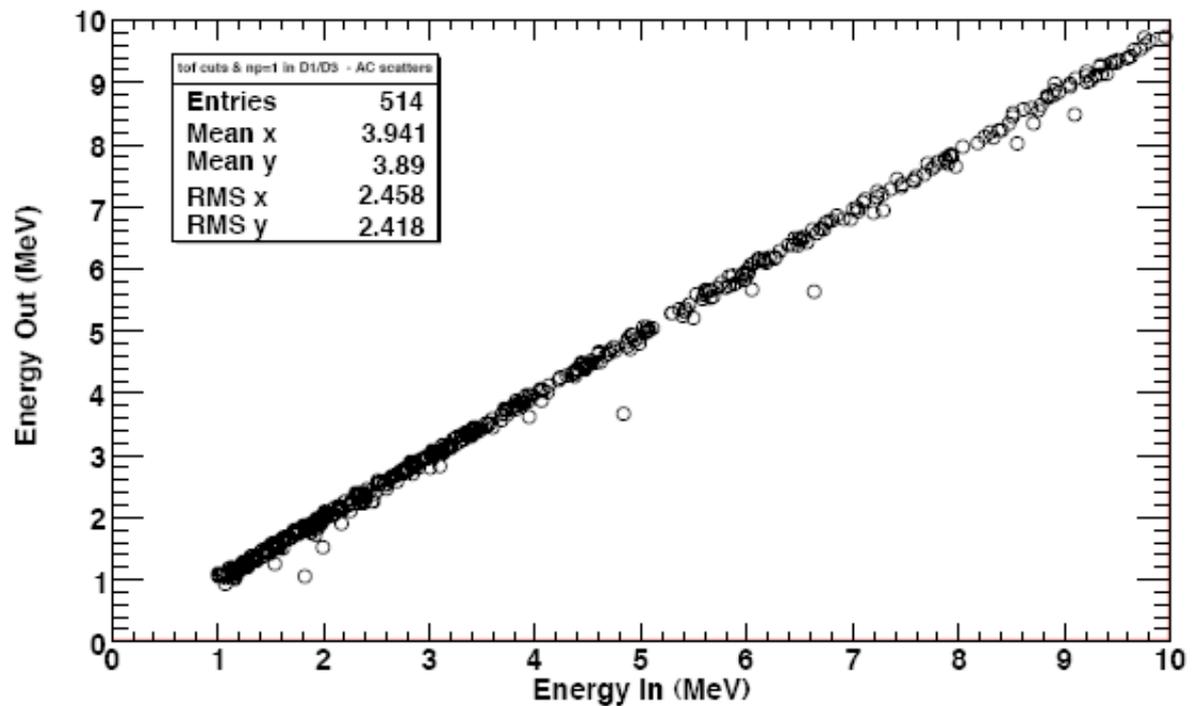


Figure 5.16: Scatter plot representing FNIT response for double scatters with ToF cuts and one elastic n-p scatter in each detector imposed. Courtesy UNH

energy, or energy out, as interpreted by the instrument. Using this, we can simply make the scatter-plots,  $E_{in}$  (*known* from simulations) vs.  $E_{out}$  (*calculated* from simulations), that represent the instrument response. In the Figures 5.14, 5.15, 5.16 and 5.17, we can see these plots made for various “cuts” for a simulation run conducted recently at UNH using a flat energy spectrum of 1-10 MeV incident neutrons. The original simulation has 40 million incident neutrons. Of these, around 2000 are detected as double-scatter events (Figure 5.14). It is clear that most events lie along the  $E_{in} = E_{out}$  diagonal, but there are a fair few events that occur off-diagonal. At first glance, this is encouraging, since it is evident that the instrument detects most neutrons with energy close to its actual energy. However, there are several off-diagonal events that would make the response matrix less ‘diagonal’, and that affects our inversion and regularisation capabilities of the instrument response as will be explained in the next chapter. However, it is possible to make some selections to reduce the number of unwanted non-diagonal events. The first step is to make ToF cuts. We know that neutrons within a certain energy range will travel between two detector rods in a certain timing range. So, in this case, we decided to select events which had ToFs between 6 ns (well above FNIT’s timing resolution of 1.3 ns) and 20 ns. This is something we would be able to do even with ‘real’ data since ToF can be ascertained there. Making this

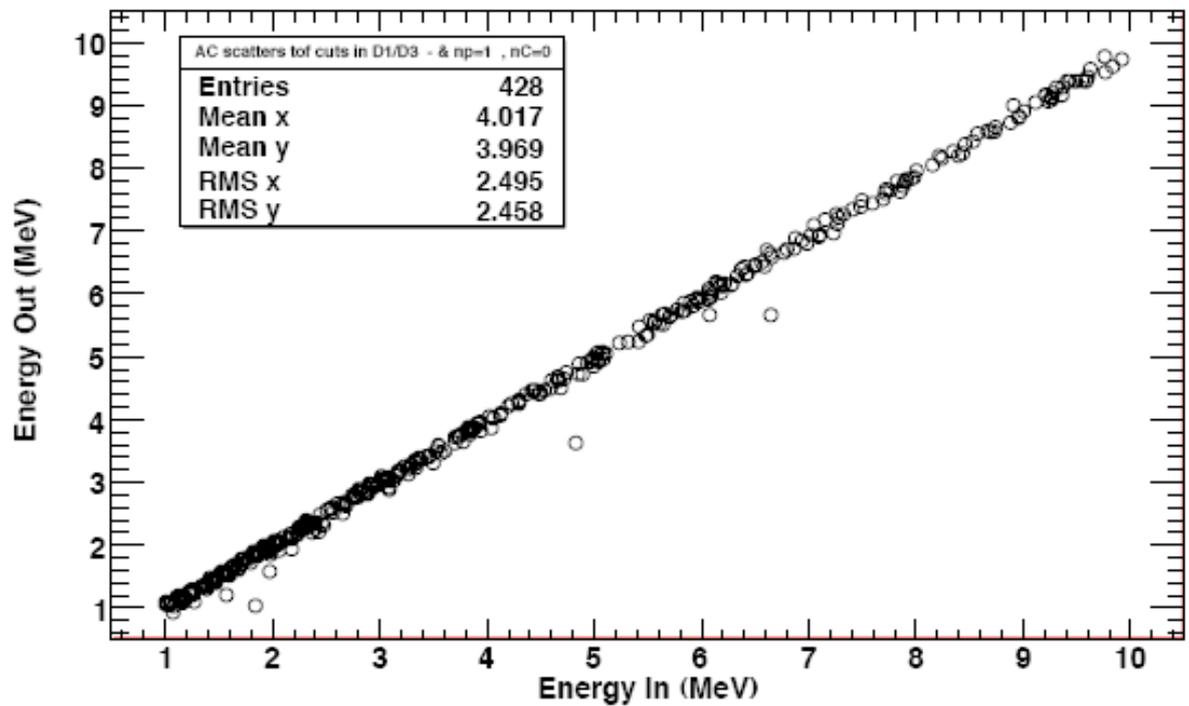


Figure 5.17: Scatter plot representing FNIT response for double scatters with ToF cuts, one elastic n-p scatter in each detector and zero total n-C scatters imposed. Courtesy UNH

simple selection automatically makes our response much more diagonal as can be seen in Figure 5.15 by eliminating more than 300 unwanted events. Even here, about 3% of the events are non-diagonal, but improvement from Figure 5.14 is significant. With the simulated data, we can make even further cuts, e.g. by imposing an n-p scatter in each detector (Figure 5.16) and rejecting all n-C scatter events as well (Figure 5.17). These improve the diagonality even further - about 1% of the events are off-diagonal - but have two immediate problems: number of events comes down dramatically (to less than a third of the initial double scatters detected) thereby affecting instrument efficiency *and* in ‘real’ data, we would not have the information to make these selections offhand. Nevertheless, these results show that FNIT will produce a rather diagonal response matrix and that its performance in space can be relied upon on Earth! It is important to note that all these plots still do not impose instrumental limitations, which would hinder the accuracy when acquiring real data. To simulate ‘real’ data, it has to be ‘smoothed’ and threshold effects of the instrument have to be taken into account. This produces a broadened response as seen in Figure 5.18, which has been obtained by applying filters on threshold as a function of position along the rod and ToF; and smoothed in energy, ToF and position of interaction. It is this figure that will be used

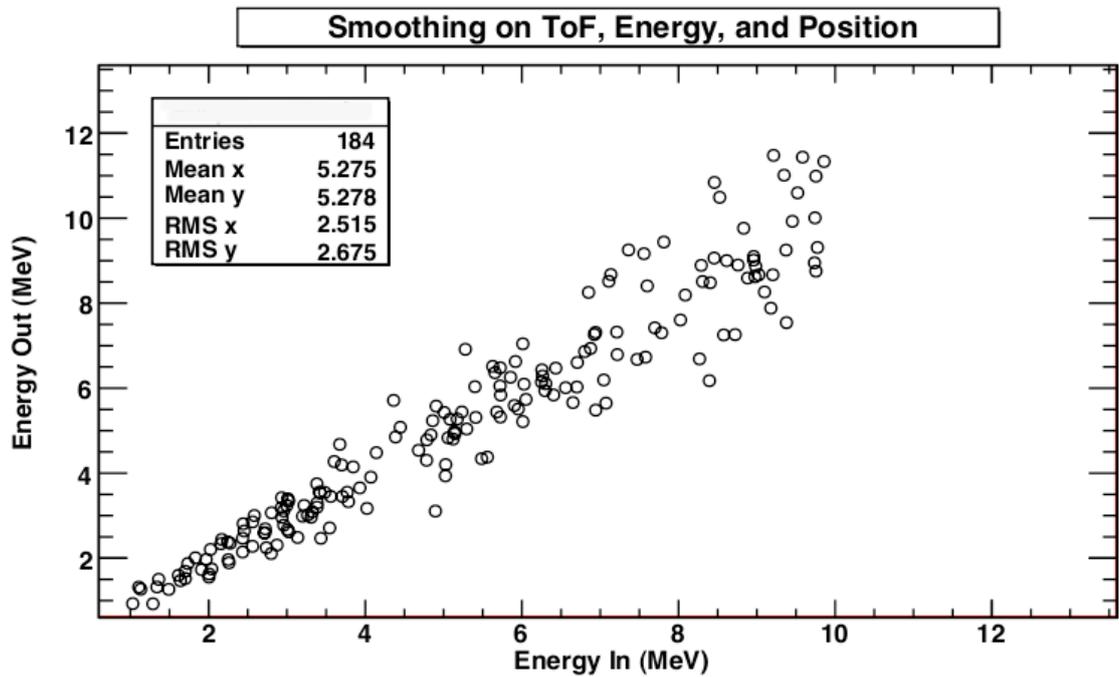


Figure 5.18: Scatter plot representing FNIT response for ‘smoothed’ double scatters with ToF cuts imposed. Courtesy UNH

to obtain a response matrix that we will use to test our regularisation methods in the next chapter. For some cases, we also tested the response matrix generated from un-broadened data, i.e. from Figure 5.15.

In the next chapter, we shall delve into how the above representations can be used to create a response matrix with bins and numbers, and how this detector response matrix can be manipulated to produce the desired quality of neutron spectra and images from real-time data produced by instruments such as FNIT.

# Chapter 6

## Data Inversion Techniques for Solar Neutron Detectors

### 6.1 Relevance, importance and significance of inversion

Inverse problems occur throughout science but are particularly relevant in astronomy. This is because a lot of astronomical data is acquired “remotely” and physical properties of the source are to be determined using these data ([Thompson & Craig 1992](#)). Take the example of studying solar flare X-rays: here we observe the X-rays to deduce the properties of the particles that produce them, namely electrons, and this is an inverse problem. Ideally, measuring the electrons themselves would give us the most direct information about them, but since astronomy is predominantly observational and we are limited by the distance to the objects we study, we mainly rely on data that is possible to acquire. In the case above, X-ray spectra have to be inverted to provide us with what we believe to be the electron spectra at the Sun. For our study on neutrons, the reason that it is an inverse problem is slightly different. For one, we are detecting the particles that we actually want to measure. So, if the instrument we were using behaved exactly as it was supposed to, i.e. the energy of the recoil proton plus the energy of the scattered neutron is the total energy of the incident neutron, then determining the neutron spectra would not be an inverse problem. But in reality, the instrument we use has instrumental limitations as well as the prevalence of experimental noise, making inversion necessary. In a nutshell, the reason for requiring inversion to ascertain the original spectral distribution of the particles of interest is this: if the acquired data does not have a one-to-one relation to the property of the detected particle, then inversion becomes necessary. In the case of X-ray spectra, electrons of

an entire range of energies can produce a photon of a particular energy and so inversion becomes necessary. In the case of neutrons, one can get different signals from neutrons of the same energy coming from the same source based purely on the uncertainty in how the neutron scatters in each detector and what signal that detector produces. Hence, inversion becomes necessary.

Before an instrument can be used in the field, one must demonstrate that the design produces a clear signature of the desired particle. One must then calibrate the instrument with known stimuli, in this case neutrons. Since calibrations cannot be used to measure all possible energies, one must perform a part of this task by way of computer simulations. Finally, one must have the scientific and mathematical tools to take data from an unknown source and from them reconstruct the spectrum of neutrons that produced those data. The aim of FNIT is the deduction of the neutron energy distribution produced at the Sun, which is straightforwardly obtained from the distribution of solar neutrons incident on the instrument. Double and multiple scatter events in FNIT allow the determination of neutron energies and arrival directions. For a variety of reasons, described in Chapter 5, some proportion of events will unavoidably be misrepresented, but GEANT4 simulations can determine this proportion and its dependence on energy and arrival direction. Counting neutrons in (energy and angle) bins, we represent these proportions as a response matrix  $\mathbf{M}$  whose  $(i, j)^{th}$  element  $m_{ij}$  gives the probability of a neutron which belongs in the  $j^{th}$  bin actually being counted in the  $i^{th}$  bin. Introducing the data vector  $\mathbf{x}$ , whose  $i^{th}$  element is the number of neutrons detected in the  $i^{th}$  bin, and the vector  $\mathbf{s}$  whose  $j^{th}$  element is the true number of neutrons incident in the  $j^{th}$  bin, we have

$$\mathbf{x} = \mathbf{M}\mathbf{s} + \epsilon, \quad (6.1)$$

where  $\epsilon$  is the ‘noise’ vector, i.e. its  $i^{th}$  element is the noise on the  $i^{th}$  measurement. If  $\epsilon$  were identically equal to zero, then simple inversion  $\mathbf{s} = \mathbf{M}^{-1}\mathbf{x}$  would be adequate. However, this generally amplifies any data noise and gives unreliable, possibly unphysical results (e.g. [Craig & Brown 1986](#)). This simplified expression of course assumes that the matrix  $\mathbf{M}$  is square and has the same length as vector  $\mathbf{s}$ , but in general this is not necessarily the case. If one feels the data is over-sampled, for example, and wants to recover fewer points but still be reliable, a non-square response matrix may be generated. In this case, both sides have to be multiplied by  $\mathbf{M}^T$  to obtain a square matrix, which may be inverted. Of course, the problems of data noise will still have to be dealt with. However, for simplicity sake, let us henceforth assume  $\mathbf{M}$  is a square

matrix, especially since the  $\mathbf{M}$  we have generated for FNIT has been chosen to be square - see later.

We must use regularisation methods that simultaneously deconvolve and smooth the reconstruction. Various methods of finding  $\mathbf{s}$  have been proposed and the most appropriate seem to be those postulated by [Tikhonov \(1963\)](#) and later implemented successfully by e.g. [Piana \(1994\)](#); [Kontar et al. \(2004\)](#); [Prato et al. \(2006\)](#). The solution can be found by minimising

$$|\mathbf{s} - \mathbf{x}|^2 - \lambda\Phi(\mathbf{s}), \quad (6.2)$$

where the function  $\Phi$  is a measure of ‘smoothness’ and the ‘smoothing parameter’  $\lambda$  expresses the degree of compromise between exact data fit and smoothness. Several forms of  $\Phi$  have been employed, each with its own pros and cons.

## 6.2 Inversion techniques and response matrix

Take, for example, [Toner et al. \(2001\)](#), who implemented a Maximum Entropy (ME) deconvolution of COMPTEL solar neutron data from the flare of 1991 June 15, using the entropy function  $\Phi = -\sum_{s_i} \ln s_i$ . [Figure 6.1](#) compares results from this procedure with those previously obtained using just the diagonal elements of the response matrix ([Kocharov et al. 1998](#)). Using only the diagonal elements of the response matrix underestimates the overall sensitivity to neutrons; but including the off-diagonal elements demands a regularisation procedure, as used here.

The need for such an approach, to allow fullest exploitation of the data, is clear, as is their scientific interest. Specifically, [Kocharov et al. \(1998\)](#) had already found smaller numbers of neutrons than expected on the basis of the observed pion decay continuum, suggesting an under-abundance of heavier species in the flare-accelerated ion population. The more complete analysis of [Toner et al. \(2001\)](#) places still more stringent limits on fast, heavy ions.

Like ME, there are several possible methods and forms of  $\Phi$  that can be used in astronomical inverse problem cases. [Thompson & Craig \(1992\)](#) give a wonderful overview of four such methods and compare their performances. They build on the studies by [Turchin \(1967\)](#); [Craig & Brown \(1986\)](#); [Gull \(1988\)](#); [Thompson et al. \(1991\)](#) and compare several automated methods of selecting the smoothing parameter (Equivalent Degrees of Freedom, [Craig & Brown \(1986\)](#) and Bayesian methods) as well as the more basic  $\chi^2$  method. It is evident from their results that each method has its

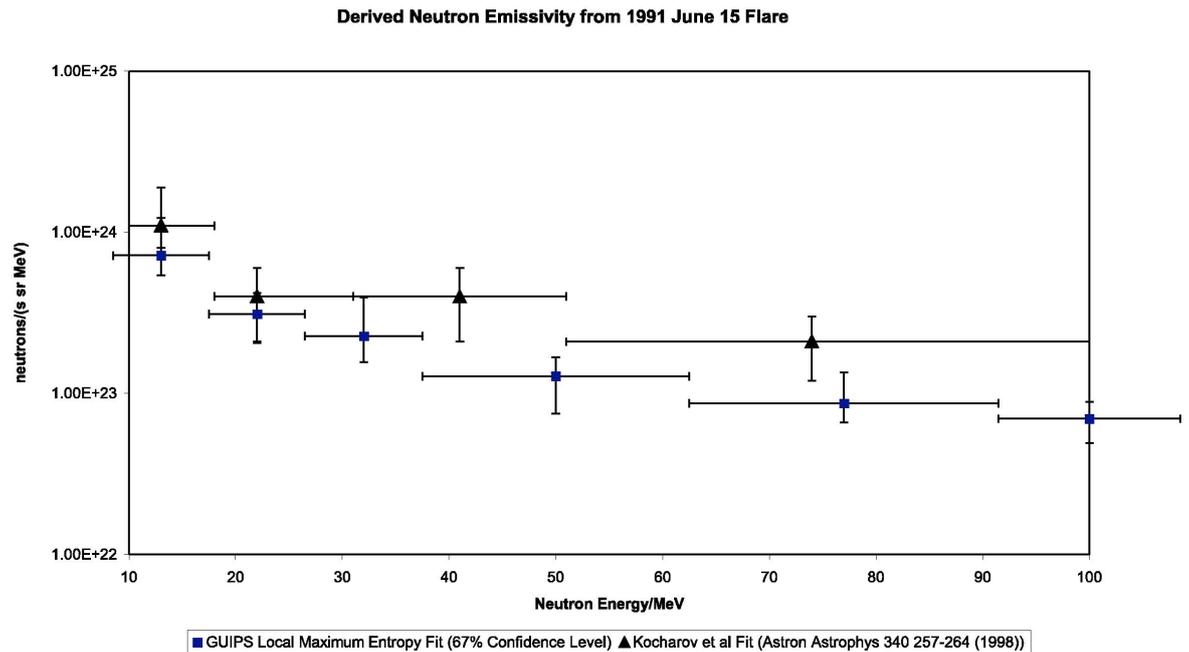


Figure 6.1: Neutron emissivity of the solar flare of 1991 June 15 as a function of neutron energy, derived from COMPTEL data using a Maximum Entropy deconvolution of the instrument response, and compared with the earlier results of [Kocharov et al. \(1998\)](#), obtained using just the diagonal elements of the response matrix. From [Toner et al. \(2001\)](#)

range or scope of optimal performance and only a detailed study such as this would determine which method(s) is (are) appropriate to which situation. In the following sections, we employ three separate methods of regularisation (or deconvolution) for neutron detectors and apply these to FNIT providing us with some interesting results.

Before that, let us briefly explain how we obtain a response matrix from the data and plots described and shown in Chapter 5. Quite simply, this involved dividing the chosen plot into a grid and counting the number of events in each grid. We did this for Figures 5.15, which has a large enough number of events for good statistics with ToF cuts only; and 5.18, which is likely to be the most realistic data we can aspire to get remotely from space since we have imposed only ToF cuts to double n-p scatter events and smoothed that data to match instrumental effects. How one determines the size of each grid, or bin, is one of the primary considerations since that is likely to reflect how effective a particular regularisation method is and would also emphasise how the instrument behaves when collecting real data. In our analysis, we used two different binning strategies. Firstly, we divided  $E_{in}$  and  $E_{out}$  into an equal number of divisions. This is not essential, but is not unrealistic and allows us to obtain a square

response matrix  $\mathbf{M}$ . We then divided the energy bins using either a linear spacing, e.g. each bin 0.3 MeV in size, both in  $E_{in}$  and  $E_{out}$  space; *or* a logarithmic spacing, where each bin was a percentage size of the central energy of that bin. The former method is clearly simpler, but the latter is more appropriate since the energy resolution of FNIT depends on the energy, i.e.  $\delta E/E = 10\%$ . So at 10 MeV, for example, the size of the bin would be 1 MeV whereas at 2 MeV, the bin size would be 0.2 MeV for logarithmic binning. The results for both these binning strategies for the separate regularisation methods we have used are depicted in §6.4 and 6.6.

### 6.3 Choice of regularisation method

The regularisation method we choose to test the FNIT response is Tikhonov regularisation (Tikhonov 1963). Tikhonov regularisation generally leads to results similar to those obtained with e.g. ME (Thompson & Craig 1992), with lesser computational effort. It uses a smoothing function of the form  $\Phi = \|\mathbf{L}\mathbf{s}\|^2$ , with the matrix  $\mathbf{L}$  chosen to make  $\Phi$  give the norm of  $\mathbf{s}$  (zeroth-order regularisation) or an estimate of the  $n^{th}$  differences of the elements of  $\mathbf{s}$  ( $n^{th}$  order regularisation). Especially via introduction of the (Generalised) Singular Value Decomposition - see below - it results in a robust, simply implemented algorithm. Implementation may thus be sufficiently robust and automatically carried out, if necessary, on board the spacecraft. For example, in Figure 6.2, zeroth-order Tikhonov regularisation - described in detail later - is applied in MATLAB to obtain the neutron energy distribution using simulated data for the UNH neutron telescope response matrix, as determined by Moser et al. (2005b). The UNH neutron telescope has enough qualitative similarities to an FNIT type instrument to give a deconvolution problem of a similar character. It is a linear algorithm and therefore easy to implement. It has been used successfully on other instruments of a similar nature. Prior to working with FNIT, we applied zeroth-order Tikhonov regularisation (e.g. Bertero et al. 1988) to the analysis of data recorded by the UNH double-scatter neutron telescope (Saxena 1990), using a response matrix obtained from GEANT4 simulations (Moser et al. 2005b). This neutron telescope was of a similar type to FNIT, but of course was a ground-based detector and so used to detect cosmic ray secondary neutrons. So its applications were different in that the cosmic ray source is extended in angle, unlike the Sun, from which we would expect neutrons coming from a ‘point’ source. In Tikhonov regularisation,  $\Phi$  is a quadratic form in  $\mathbf{s}$  and the minimisation problem (6.2) has an analytic solution for fixed  $\lambda$  in terms of the (Generalised) Singular Value Decomposition - (G)SVD - of  $\mathbf{M}$ . Efficient algorithms exist for

finding the (G)SVD of a matrix, and the solution written in this form avoids numerical instabilities.

To test these methods, one has to choose some sort of ‘fake’ or simulated data to test them on. One could potentially use neutron data produced by say the Watt distribution (Equation 5.1), but this does not have any distinct sharp features with which we can test the robustness of our results. Hence, for these tests, we chose to generate this data ourselves. In fact, the data we chose was initially used to test the UNH neutron telescope (Figure 6.2), and we have chosen to use it here as well because it has particular features that would test the robustness of the regularisation methods. As you can see from Figures 6.2 to 6.10, the data has defined peaks and troughs and the entire amplitude drops off as a power-law. Although this dampening wave-like form was initially chosen to match the energy and angular behaviour of the UNH neutron telescope, we decided to use it to test FNIT as well because of the distinct features, even though this does not reflect ‘real’ neutron energy data we are likely to get in the field. To this energy and angle-dependent data, we added a maximum random Poisson noise of 10%, noise at a similar level to that experienced in COMPTEL neutron measurements from flares, hopefully pessimistic for FNIT. All this was done using MATLAB. It is important to note that the data values are arbitrary numbers and can be used as ‘fake’ neutron energy data. In all the reconstruction plots, the  $x$ -axis is the energy bin number and the  $y$ -axis is the normalised data value. Originally, each energy bin had three separate angular bins associated with it, which we have folded out to all represent an energy bin in this case, providing us with the jagged features.

## 6.4 Tikhonov regularisation and results

Tikhonov regularisation uses Equation 6.2, with particular choices of  $\Phi$ , motivated by simple ideas:

- minimise the norm of  $\mathbf{s}$  with respect to the data fit: zeroth-order Tikhonov
- minimise the derivative so that the sharp features are ‘smoothed’ further: first-order
- minimise the second derivative which is a more refined measure of sharpness, and will leave behind more of the structure: second-order.

We have tested all these three orders of Tikhonov regularisation with our data convolved through the FNIT response.

To understand how zeroth-order Tikhonov regularisation is implemented, consider any matrix,  $\mathbf{M}$ .  $\mathbf{M}$  can be decomposed into three matrices, i.e.  $\mathbf{M} = \mathbf{U}\mathbf{S}\mathbf{V}'$ , where  $\mathbf{U}$  and  $\mathbf{V}'$  are orthogonal matrices and  $\mathbf{S}$  is a diagonal matrix that has the singular values along its primary diagonal whereas all its other matrix values are zero. An orthogonal matrix is one whose transpose is equal to its inverse (e.g. [Golub & van Loan 1996](#)). This decomposition of the matrix  $\mathbf{M}$  is known as its Singular Value Decomposition (SVD), (e.g. [Golub & van Loan 1996](#)). It allows us to find an analytical solution to the regularisation problem (6.2) for the special case  $\mathbf{L} = \mathbf{I}$ , in a compact, numerically stable form. Generalised SVD, which will be explained later, is a more general form of SVD where two matrices can be decomposed simultaneously and this can be used for higher order Tikhonov regularisation. Minimising Equation 6.2 in turn with respect to each of the elements  $s_\lambda$  and using the fact that the singular values  $\sigma_k$  are the eigenvalues of  $\mathbf{M}^T\mathbf{M}$ , it can be shown that ([Piana 1994](#)):

$$s_\lambda = \sum_{k=1}^N \frac{\sigma_k}{\sigma_k^2 + \lambda} (\mathbf{x}, \mathbf{v}_k) u_k, \quad (6.3)$$

where  $\sigma_k$  is the  $k^{th}$  singular value and  $\mathbf{v}_k$  and  $u_k$  are the  $k^{th}$  vectors of the matrices  $\mathbf{V}$  and  $\mathbf{U}$  respectively. With the solution (6.3),  $\mathbf{s}$  is now a function of  $\lambda$ . As already noted,  $\lambda$  expresses the degree of compromise between exact data fit and smoothing. Several strategies may be adopted to choose the best value of  $\lambda$ . In the first instance, we choose  $\lambda$  so that the following equation

$$\|\mathbf{x} - \mathbf{M}\mathbf{s}\|^2 - \sum_1^N \mathbf{x} = 0. \quad (6.4)$$

It is important to automate the choosing of  $\lambda$  and the criterion above ([Turchin 1967](#)) has been used in past studies (e.g. [Kontar et al. 2004](#); [Prato et al. 2006](#)).

Furthermore, we also tested the GSVD method for first- and second-order Tikhonov regularisation ([Golub & van Loan 1996](#)). Following on from the zeroth-order method, the first- or second-order (or  $n^{th}$  order for that matter!) minimisation problem is given by ([Kontar et al. 2004](#)):

$$\|\mathbf{M}\mathbf{s} - \mathbf{x}\|^2 + \lambda\|\mathbf{L}\mathbf{s}\|^2 = \min, \quad (6.5)$$

where  $\mathbf{M}$  is an  $N \times N$  matrix and  $\mathbf{L}$  is a  $P \times N$  matrix. The solution of the equation above given by [Hansen \(1989\)](#) and used by e.g. [Kontar et al. \(2004\)](#) is:

$$\mathbf{s}_\lambda = \sum_{k=1}^P \left( \frac{\sigma_k^2}{\sigma_k^2 + \lambda} \frac{(\mathbf{x} \cdot \mathbf{u}_k)}{\sigma_k^M} \right) \mathbf{w}_k + \sum_{k=P+1}^N (\mathbf{x} \cdot \mathbf{u}_k) \mathbf{w}_k, \quad (6.6)$$

and  $\sigma_k = \sigma_k^M / \sigma_k^L$ , where  $\sigma_k^M$  and  $\sigma_k^L$  are the  $k^{\text{th}}$  singular values of the matrices  $\mathbf{M}$  and  $\mathbf{L}$  respectively as represented in the matrices  $\mathbf{C}$  and  $\mathbf{S}$  as explained below.

Like in SVD, the response matrix  $\mathbf{M}$  can be decomposed such that  $\mathbf{M} = \mathbf{UCW}^{-1}$ , but you also have a complementary matrix  $\mathbf{L}$  that is defined by the minimisation problem in Equation 6.5 (e.g. Hansen 1989, 1990), which is derived from Equation 6.2. The choice of  $\Phi$  is implicit in the choice of  $\mathbf{L}$ . It is the choice of  $\mathbf{L}$  that determines what order Tikhonov regularisation you are implementing. Hence, for the case of first-order regularisation,  $\mathbf{L}$  has one column less than  $\mathbf{M}$  and has 1 along its primary diagonal and -1 along its secondary diagonal. Likewise, for second-order regularisation,  $\mathbf{L}$  has -1 along its primary diagonal, 2 along its secondary diagonal and -1 again as the values of its tertiary diagonal. Both these are explicitly given in Hansen (1989) as well. GSVD also decomposes  $\mathbf{L}$ , such that  $\mathbf{L} = \mathbf{VSW}^{-1}$  and the minimisation solution given by Equation 6.6 can be implemented. Of course, if we choose  $\mathbf{L}$  to be the identity matrix, then we replicate zeroth-order Tikhonov and Equation 6.6 reduces to Equation 6.3 identically.

The (G)SVD solution has several useful consequences. The two unitary matrices for each decomposition ( $\mathbf{U}$  and  $\mathbf{V}$ , or  $\mathbf{U}$  and  $\mathbf{W}$ , or  $\mathbf{V}$  and  $\mathbf{W}$ ) are composed of sets of orthonormal column vectors, which can each be thought of as pertinent to particular data bases and reconstruction spaces. They are associated with the singular values, in decreasing order of magnitude. This property makes the solution in terms of (G)SVD extremely useful for inverse problem applications, because you can see which parts of the solution (corresponding to a particular column vector and singular value) are most sensitive to noise. Additionally, smaller the singular value, smaller the role of its corresponding column vector and this can also play a role in data compression as we may have the opportunity to confidently eliminate redundant parts of the data. All in all, Tikhonov regularisation seems appropriate for an instrument such as FNIT and the following section has the necessary results illustrated.

## 6.5 Testing Tikhonov regularisation

The zeroth-order Tikhonov regularisation method to test data was implemented using the MATLAB SVD routine. Both the UNH neutron detector and FNIT were tested with ‘fake’ simulated data generated on MATLAB to determine the accuracy of this

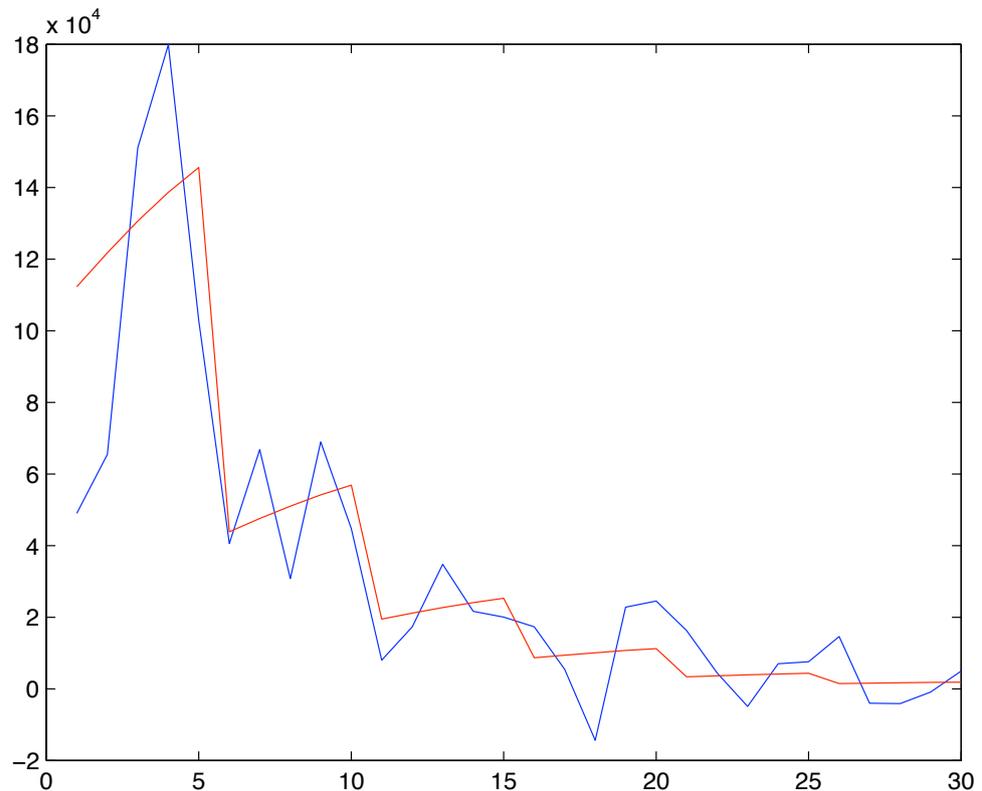


Figure 6.2: SVD regularisation of the inverted solution using the UNH neutron telescope response matrix. The red curve is the original data generated on MATLAB and the blue curve is the reconstructed spectrum. A random Poisson noise of 10% was added to the generated data before the reconstruction. The  $x$ -axis denotes the energy bin number of  $E_{in}$  and the  $y$ -axis is the data/energy value in arbitrary units.

Tikhonov SVD method. The response matrix already assumes a point source in the correct direction. Energy and angle joint deconvolution is beyond the scope of this thesis and is left for others to do. Testing this sample data allows us to see how the various regularisation methods deal with the effects of noise as well as the real structure of the data. These results using ‘fake’ data can be seen in Figure 6.2 for the UNH neutron telescope. As is evident, the reconstructed data makes a pretty good fit but has definite “edge” effects (these are due to the two ends of the data, where the discrete binning produces end edge effects) and is quite under-smoothed. Getting even better reconstructions could involve going higher than just zeroth order regularisation and perhaps modifying our binning strategy as discussed later on.

To reiterate, the ‘real’ solar neutron spectrum is unlikely to have such sharp features as we have used in our ‘fake’ data. Real neutron spectra are likely to be smooth power-

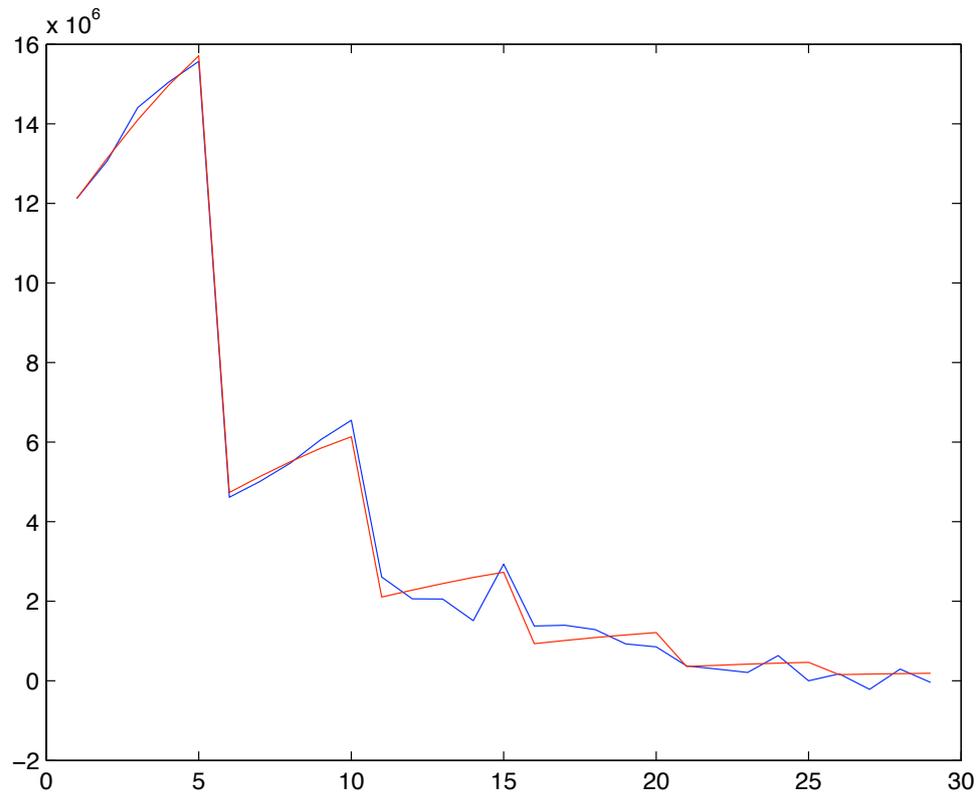


Figure 6.3: SVD regularisation of the inverted solution using the FNIT response matrix as obtained from the ‘smoothed’ data (Figure 5.18) using linear binning. The red curve is the original data generated on MATLAB and the blue curve is the reconstructed spectrum. A random Poisson noise of 10% was added to the generated data before the reconstruction. The  $x$ -axis denotes the energy bin number of  $E_{in}$  and the  $y$ -axis is the data/energy value in arbitrary units.

law spectra as seen in the tail of the Watt distribution or in the example shown in Figure 6.8.

Note that in *all* the following regularisation reconstruction plots (Figure 6.2 to Figure 6.6), the red curve is the ‘fake’ data generated on MATLAB and is identical for all the plots whereas the blue curve is the reconstructed spectrum. The  $x$ -axis denotes the energy bin number of  $E_{in}$  and the  $y$ -axis is the data/energy value in arbitrary units. Figures 6.7 and 6.8 are slightly different and will be described in detail later. For the plots with logarithmic binning, the  $x$ -axis value can approximately be expressed as  $2.5 \log E_{in}$ , whereas for the linear binning, the  $x$ -axis value is simply  $E_{in}/3$ . In our research, we also experimented with the bin-size of  $E_{in}$  and  $E_{out}$  to see what impact that has on the results. In all the plots in this chapter, the linear-binning uses 0.3 MeV

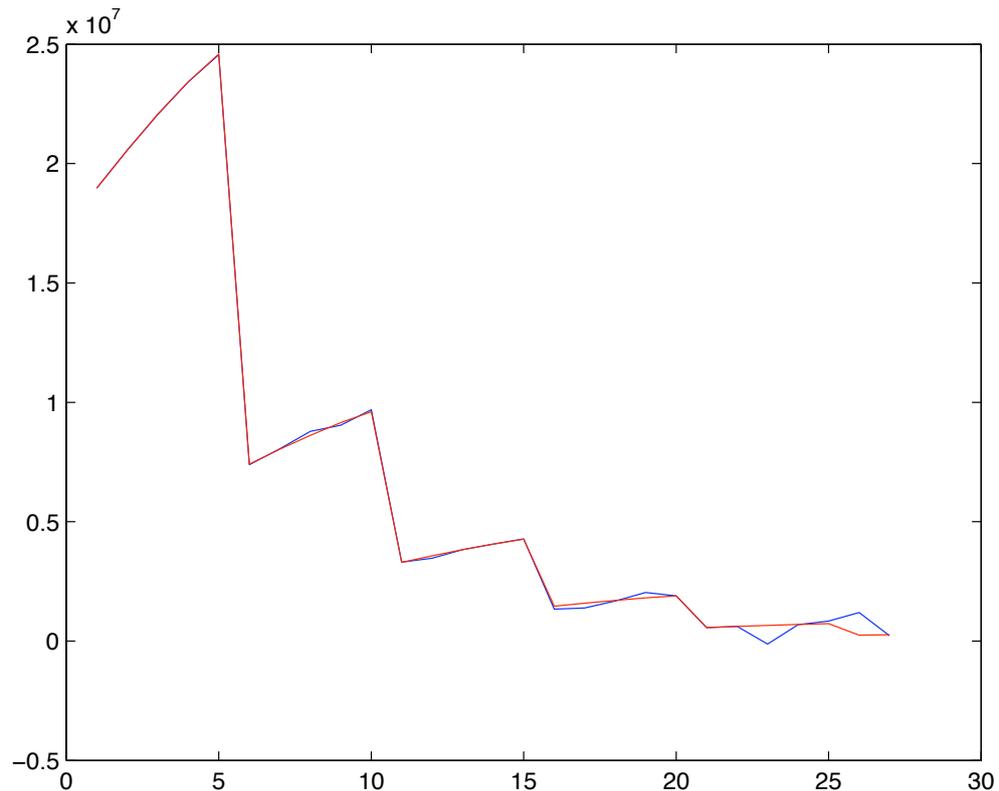


Figure 6.4: SVD regularisation of the inverted solution using the FNIT response matrix as obtained from the ‘smoothed’ data (Figure 5.18) using logarithmic binning. The red curve is the original data generated on MATLAB and the blue curve is the reconstructed spectrum. A random Poisson noise of 10% was added to the generated data before the reconstruction. The  $x$ -axis denotes the energy bin number of  $E_{in}$  and the  $y$ -axis is the data/energy value in arbitrary units.

bin sizes throughout the 1-10 MeV range. So this produces 27 energy bins on each axis. For the logarithmic binning, we used a bin size of  $\sim 10\%$  of the central energy of each bin, so at 1 MeV, the bin size was 0.1 MeV and at 10 MeV, the bin size was 1 MeV. This too produces 27 energy bins. Using larger bin sizes obviously produces fewer bins and the performance of the regularisation methods drop off accordingly. On the other hand, using a larger number of bins does improve the performance, but that is unrealistic since the energy resolution of FNIT is about 10%. Hence the log binning we have used is perhaps the most realistic binning strategy to adopt, although decreasing and increasing bin sizes may be done to judge the performance of each inversion technique and/or if the resolution of FNIT improves with further instrument development. Choosing the 0.3 MeV bins for the linear bins was not totally arbitrary.

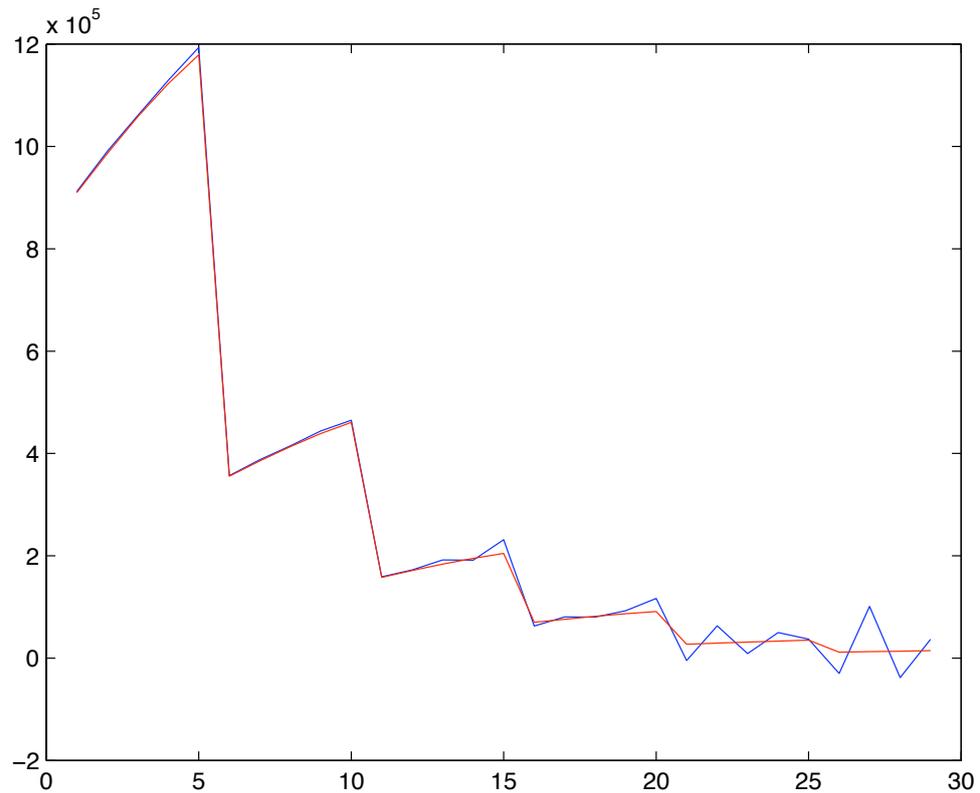


Figure 6.5: SVD regularisation of the inverted solution using the FNIT response matrix as obtained from the ToF cuts only data (Figure 5.15) using linear binning. The red curve is the original data generated on MATLAB and the blue curve is the reconstructed spectrum. A random Poisson noise of 10% was added to the generated data before the reconstruction. The  $x$ -axis denotes the energy bin number of  $E_{in}$  and the  $y$ -axis is the data/energy value in arbitrary units.

First of all, it produces the same number of bins as our log binning did and so that makes it a suitable bin-size to choose. Secondly, 0.3 MeV bins are perhaps too large in size for the lower energy neutrons (1-2 MeV), which would be more numerous in real data (although for our simulations we used a flat spectrum in the 1-10 MeV range), but too small for the higher energy neutrons (4-10 MeV), but these are expected to be smaller in number. Hence, 0.3 MeV sized bins seemed to be a suitable ‘average’ bin size to choose for the linear bins. It is important to note, though, that although implementing linear binning is simpler, log binning is more realistic. So most of our results have been produced for log binning, whereas the plots with linear binning have been shown here so that the two binning strategies can be compared qualitatively.

For FNIT, Figures 6.3 and 6.4 have been produced using linear and logarithmic

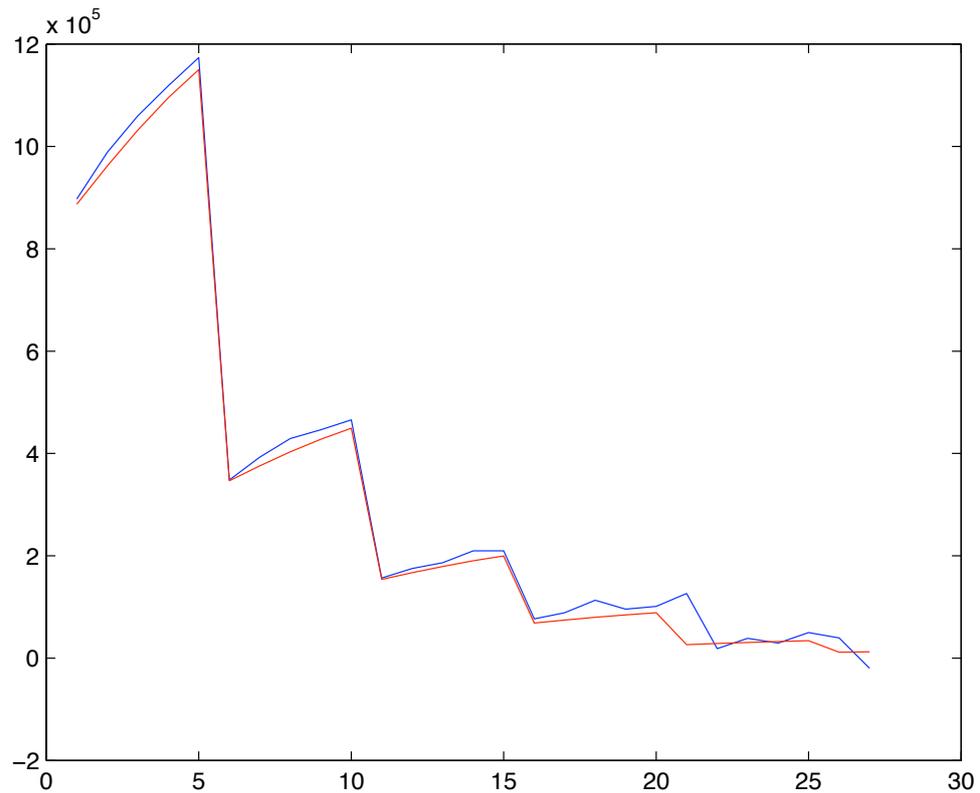


Figure 6.6: SVD regularisation of the inverted solution using the FNIT response matrix as obtained from the ToF cuts only data (Figure 5.15) using log binning. The red curve is the original data generated on MATLAB and the blue curve is the reconstructed spectrum. A random Poisson noise of 10% was added to the generated data before the reconstruction. The  $x$ -axis denotes the energy bin number of  $E_{in}$  and the  $y$ -axis is the data/energy value in arbitrary units.

binning respectively. The FNIT plots have been generated using the response matrix determined from Figure 5.18. However, for comparison, also see Figures 6.5 and 6.6, which have been generated using the response matrix obtained from Figure 5.15. Quite noticeably, SVD works pretty well for FNIT for either use of  $\mathbf{M}$ . The reconstruction represents the original data quite closely, and the results are not as under-smoothed as in the case of the UNH neutron telescope. This is primarily because the FNIT response is more diagonal than that of the UNH neutron telescope. The results of SVD on FNIT are particularly encouraging as it suggests that SVD zeroth-order Tikhonov regularisation is probably a satisfactory tool to use on FNIT. However, other methods also need to be tested to see if and when they may perform adequately as well. We see from Figures 6.2, 6.3 and 6.4 that these initial efforts are highly encouraging,

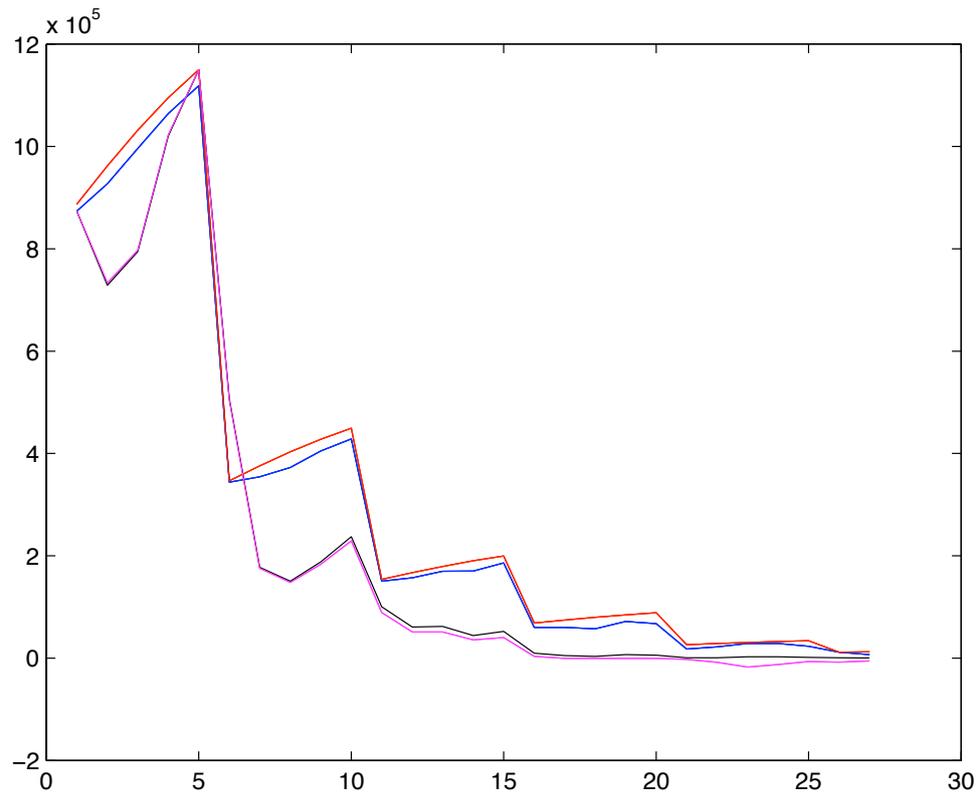


Figure 6.7: GSVd Tikhonov regularisation of the inverted solution using the FNIT response matrix as obtained from the ToF cuts only data (Figure 5.15) using log binning. The red curve is the original data generated on MATLAB and the blue, magenta and black curves are the zeroth, first and second order reconstructed spectra respectively. A random Poisson noise of 10% was added to the generated data before the reconstruction. The  $x$ -axis denotes the energy bin number of  $E_{in}$  and the  $y$ -axis is the data/energy value in arbitrary units.

consistent with other treatments of the same data and with other determinations of the same quantity.

Figures 6.7 and 6.8 have been produced by applying the first-order and second-order  $\mathbf{L}$  regularisation methods - in addition to zeroth order - to the response matrix generated from the ToF cuts only data (Figure 5.15). The red and blue curves, like before, represent the data and the zeroth order regularisation respectively. The magenta curve is the first-order reconstruction and the black curve is the second-order reconstruction. The difference between the plots is that in Figure 6.7 we have used the ‘fake’ data that we have tested throughout the chapter, whereas for Figure 6.8, we have used a smooth quadratic curve that resembles the energy distributions predicted

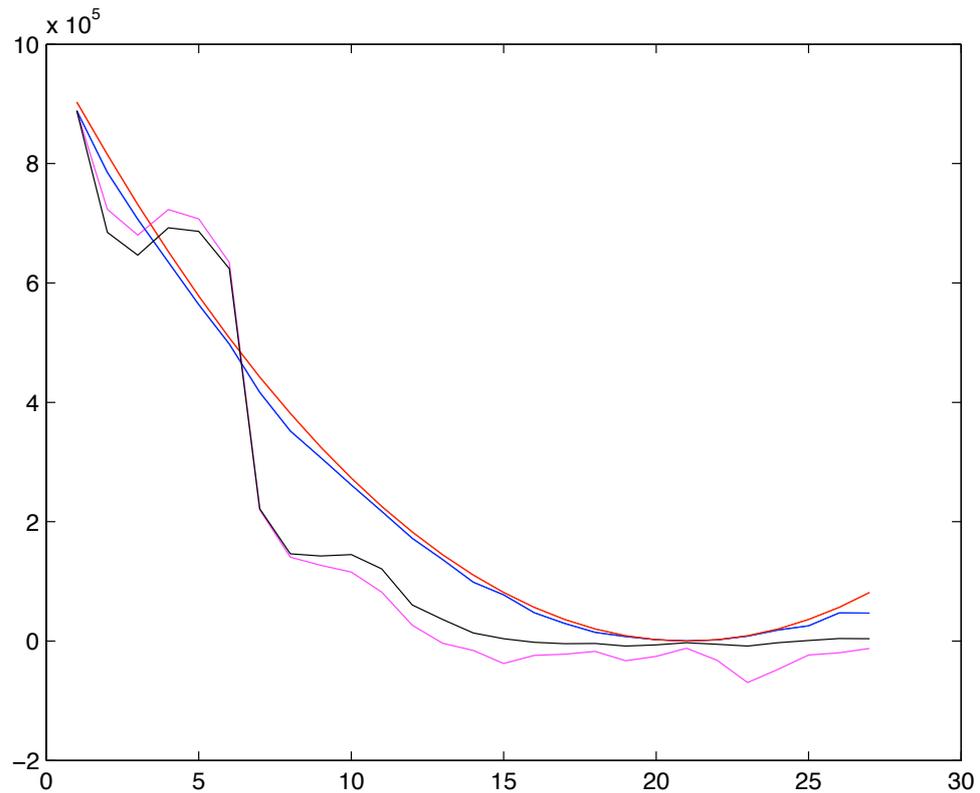


Figure 6.8: GSVD Tikhonov regularisation of the inverted solution using the FNIT response matrix as obtained from the ToF cuts only data (Figure 5.15) using log binning. The red curve is the smooth quadratic test data and the blue, magenta and black curves are the zeroth, first and second order reconstructed spectra respectively. A random Poisson noise of 10% was added to the generated data before the reconstruction. The  $x$ -axis denotes the energy bin number of  $E_{in}$  and the  $y$ -axis is the data/energy value in arbitrary units.

for near the Sun (as in Figure 1.10), with the same 10% maximum Poisson noise imposed. We did this to see whether the performance of any of the three methods varies according to the smoothness of the data, and also because ‘real’ neutron data is likely to be of some sort of smooth shape rather than the jagged data we have been testing throughout. As is evident, neither first nor second order regularisation seem to work as well as zeroth order for either data-set. This seems to result from the fact that no automatic method of choosing  $\lambda$  can work for the higher order methods - see below - and so we manually selected arbitrary values of  $\lambda$ . This suggests to us that perhaps it is the response matrix that limits which method of regularisation is optimum rather than the type of test data used in the case of FNIT.

Since the results of the higher order regularisation techniques are not as promising as those of the zeroth-order regularisation, we attempted to unearth why that might be the case. Our findings show that in the case of GSVD, the code seems unable to automatically generate the smoothing parameter  $\lambda$ , which it was able to do for SVD. This seems to be because the criterion for choosing  $\lambda$  in the SVD method (Equation 6.4), is not satisfied for GSVD. We also discovered that the Picard Condition (Hansen 1990),

$$\sum_{k=1}^{\infty} |(\mathbf{x} \cdot \mathbf{u}_k) / \sigma_k|^2 < \infty ; \quad \sigma_k \neq 0, \quad (6.7)$$

which requires  $\mathbf{x} \cdot \mathbf{u}$  to decrease monotonically for larger values of  $k$  and implies  $\mathbf{x} \cdot \mathbf{u}$  must decay to zero faster than  $\sigma_k$ , is not met for the higher order regularisations. The Picard Condition is a test of convergence, which measures the product of the data vector with the vector column  $\mathbf{u}_k$ . This product needs to converge to 0 for monotonically increasing values of  $k$  faster than the singular values  $\sigma_k$  converge to 0 for the Picard condition to be met. If this convergence criterion is unsatisfied, which happens to be the case for our higher order regularisations, then choosing  $\lambda$  automatically is impossible. This makes the higher order methods harder to implement. However, in spite of this issue, the reconstruction by manually imposing  $\lambda$  is not altogether discouraging. It is evident that particular features of the data are dealt with more appropriately with 1st and 2nd-order regularisation than with zeroth order. This is especially the case with 2nd order regularisation, where you can see the solution is more ‘smoothed’ than in the zeroth order solution and we already state earlier that zeroth order regularisation tends to ‘under-smooth’ the reconstruction. So looking at higher orders of Tikhonov regularisation may be progress in the correct direction, but perhaps both inadequate and unnecessary for the purposes of implementation in FNIT.

Since the Picard condition is not met, even trying different criteria for the selection of  $\lambda$ , like minimising the sum of the residuals (Kontar et al. 2004), did not yield any positive results. Even by varying the values of  $\lambda$  by hand did not yield a satisfactory solution. In fact, the criteria for choosing  $\lambda$  themselves were never satisfied either. Hence we conclude that higher order Tikhonov methods are unavoidably plagued by major convergence problems, crystallised in non-satisfaction of the Picard Condition (Equation 6.7). Altering binning strategy does not appear to improve the situation either. Therefore it is safe to say that our initial results and preliminary tests show that zeroth order Tikhonov regularisation is not only sufficient but better than higher order regularisation. We also tested our GSVD algorithm by setting  $\mathbf{L}$  to be the identity

matrix and  $P = N$ , so that Equation 6.6 reduces identically to the SVD solution in Equation 6.3. Doing this, we got the exact same results as we did for SVD. This led us to believe that our implementation of GSVD is correct, but there is clearly some inherent property in the FNIT response that makes the implementation of first- and second-order Tikhonov regularisation more problematic than zeroth-order. Hence, zeroth-order Tikhonov regularisation seems to be the most appropriate deconvolution strategy to pursue for FNIT. The main advantage of this is that the algorithm for zeroth order regularisation is much simpler and requires less computing power than the higher orders and so analysing FNIT data would be made easier for everyone, especially non-experts in the field who may have to use the instrument.

## 6.6 A preliminary study of blind deconvolution and results

In a more robust capacity to regularise the neutron data, we consider the [Golub, Hansen, & O’Leary \(1999\)](#) approach for a regularised total least squares implementation of Tikhonov SVD regularisation. Tikhonov itself is a least-squares algorithm, but in a total least squares approach, it is not only the right-hand-side of Equation 6.2 (the choice of  $\Phi$  or  $\lambda$ ) that has to be varied to obtain a regularised solution - as in Tikhonov - but it is also assumed the response matrix itself is known only approximately. This method may also be called “Blind Deconvolution” (BD), which too has efficient numerical algorithms available. BD has the capacity, in principle, to adapt to an instrument evolving in time - as space-borne detectors tend to do. As a preliminary experiment, we employed the MATLAB BD routine as a ‘black box’, without having access to the details of its algorithm, to produce the results shown in this section. Figures 6.9 and 6.10 show the reconstruction with the same “fake” simulated data as used in §6.4. These figures can be compared to Figures 6.5 and 6.6 respectively to evaluate the performances of Tikhonov SVD and BD. However, since the BD routine in MATLAB is primarily for image processing, we are not totally confident of its regularisation capabilities and are certain that it is different from the total least squares SVD method that [Golub, Hansen, & O’Leary \(1999\)](#) postulated. However, that does not mean that it is not a suitable method to use; only that there needs to be further study using this method, understanding the algorithm hidden in the code and applying the technique appropriately. Our treatment of it has in all respects been a preliminary test, but the interesting result demanded a mention in this thesis. It is interesting to note that BD was particularly sensitive to the number of energy bins we used; hence its performance

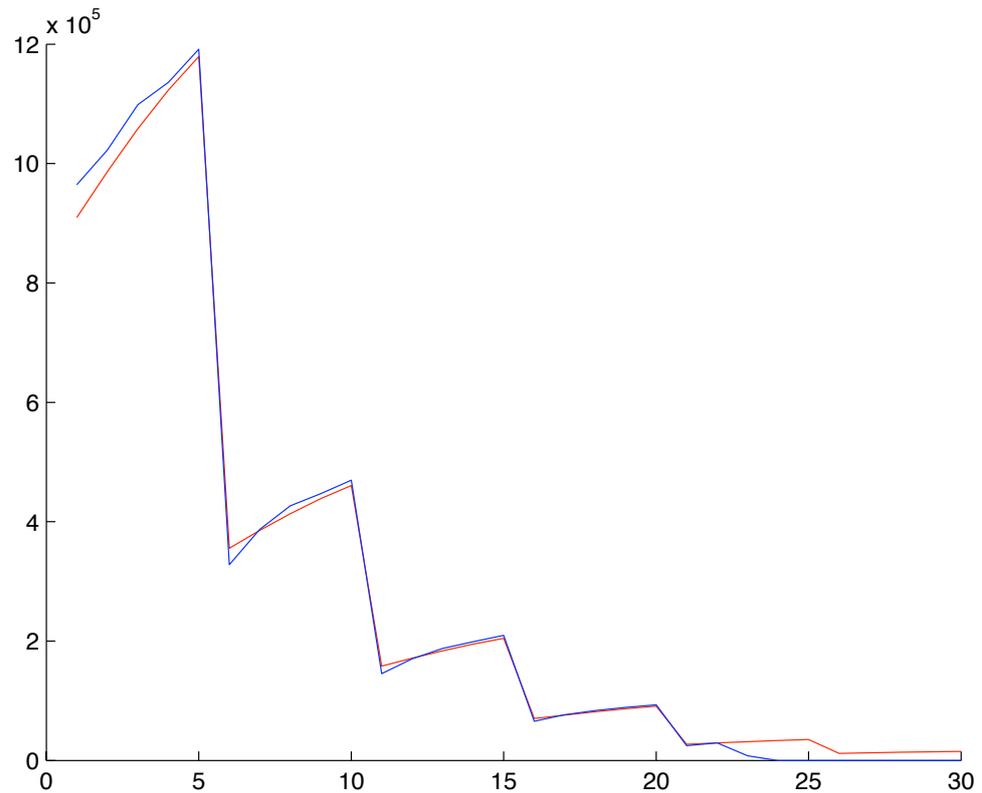


Figure 6.9: Blind Deconvolution of the inverted solution using the FNIT response matrix as obtained from the ToF cuts only data (Figure 5.15) using linear binning. The red curve is the original data generated on MATLAB and the blue curve is the reconstructed spectrum. A random Poisson noise of 10% was added to the generated data before the reconstruction. The  $x$ -axis denotes the energy bin number of  $E_{in}$  and the  $y$ -axis is the data/energy value in arbitrary units.

was far better for linear binning than for log binning. We encountered various other limitations with BD: it did not perform optimally when the first column of the reconstructed point-spread function was equal to the zero-vector; it seemed to perform better when we used other normalisation coefficients to test it and so on. Hence at this point, we cannot say for sure whether this particular BD algorithm is suitable for our purposes. It would also be worthwhile to test the total least squares approach and use the Regularization Tools Package in MATLAB, as mentioned by [Golub, Hansen, & O’Leary \(1999\)](#).

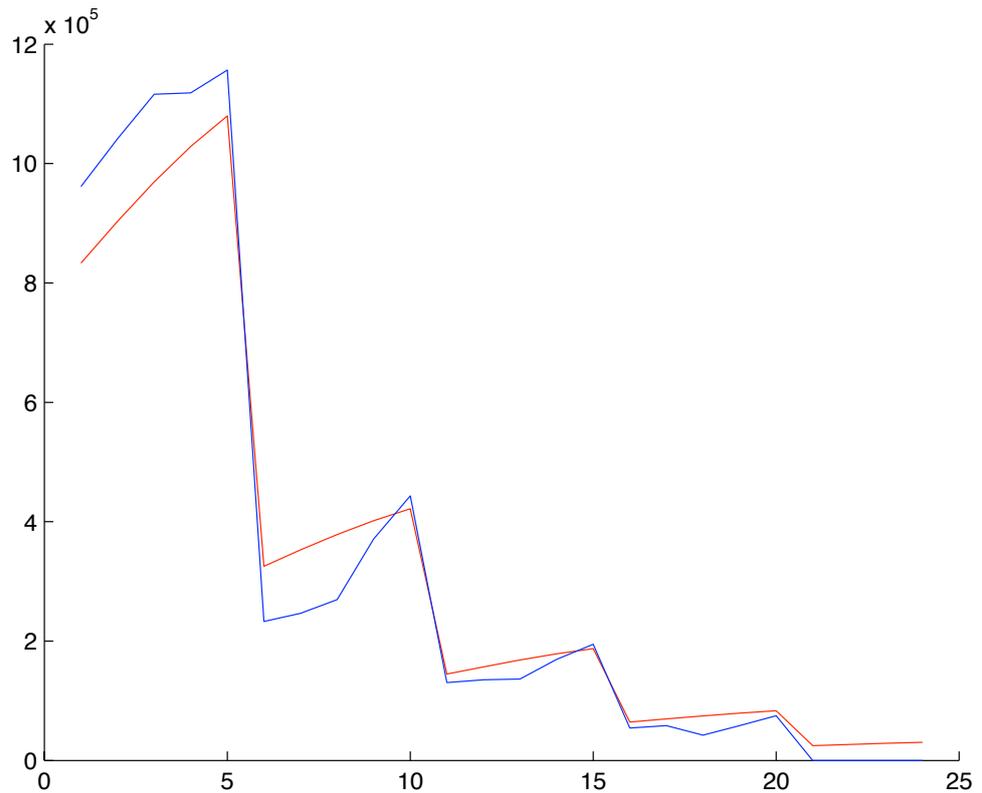


Figure 6.10: Blind Deconvolution of the inverted solution using the FNIT response matrix as obtained from the ToF cuts only data (Figure 5.15) using log binning. The red curve is the original data generated on MATLAB and the blue curve is the reconstructed spectrum. A random Poisson noise of 10% was added to the generated data before the reconstruction. The  $x$ -axis denotes the energy bin number of  $E_{in}$  and the  $y$ -axis is the data/energy value in arbitrary units.

## 6.7 Conclusions

Due to telemetry constraints on a space mission, data compression is one of the main considerations when implementing a regularisation technique. It may be possible, and desirable, to carry out some, if not all, of the deconvolution process onboard the mission. Since a lot of the regularisation is automated and does not require huge computing power, at least partial onboard processing surely seems a viable option. However, one has to first be convinced that the tools used are robust, reliable and compatible for onboard execution. Much testing of our methods needs to be undertaken first before this may become an option. Deduction of the incident neutron distribution via Tikhonov regularisation and the (G)SVD of  $\mathbf{M}$  or blind deconvolution open up

possibilities for a degree of compression, however, should transmission of data to Earth in summary form prove desirable.

SVD and the matrix  $\mathbf{M}$  define bases for the data and reconstruction spaces. Each vector of these bases is associated with one of the so-called singular values of  $\mathbf{M}$ . In the regularised solution for  $\mathbf{s}$ , the vectors associated with smallest singular values play the least significant role and are most easily overcome by noise. Thus the regularised solution for  $\mathbf{s}$  is almost unaffected if we transmit, not the numbers in all data bins, but a smaller number of linear combinations of them that will be used in the Tikhonov solution: we are able to discard some of the data in a way that respects what we are going to use them for. Similar compression can be achieved in blind deconvolution where the final form of the point-spread function has many less bins and entries compared to the original  $\mathbf{M}$ . In the case of SVD and GSVD, one can then consider a compressed decomposition and several such options have been studied in detail by e.g. Hansen (1989, 1990), where instead of the SVD and GSVD of the response matrix, they use the *truncated* SVD (TSVD) and TGSVD. They conclude that TSVD and TGSVD are as robust as their corresponding methodologies, but with the added advantage of needing and using less information, thereby achieving a certain degree of data compression. If such methods are automated, they could be carried out onboard and alleviate concerns of limited telemetry.

Binning of neutrons, at first glance a trivial process, is key to the use of regularisation techniques. The binning strategy that maximises the information content of regularised solutions needs to be determined. Initial results (above) with zeroth, first and second-order Tikhonov regularisation and blind deconvolution are encouraging, but we need a good basis for the best choice between competing regularisation strategies. We must determine the best method of dealing with telemetry constraints: do we transmit all binned data, some binned data, some or all details of some or all neutron events? These decisions have to be made in the light of the instrument response and best strategies for deconvolving it.

The optimal data binning strategy will be determined via a mixture of analytical discussion, based on the SVD, GSVD and blind deconvolution of the response matrix, and testing on simulated data. We feel that the log binning strategy where the bin-size reflects the energy resolution of the instrument is the most suitable strategy to use. However, all methods and bin sizes need to be tested to decide the ultimate approach. This testing will be part of a broader programme aimed at determining the best optimisation strategy to use, and ensuring that implementations are efficient and robust enough for automatic use onboard. Both simulated data and laboratory

---

calibration data will be used. The consequences of data compression will also be investigated using simulated data, with a view particularly to devising an optimal strategy for retaining the full details of a restricted subset of all detected neutrons. In this chapter we have shown how regularisation strategies may be deployed and tested for detectors like FNIT, both in broad principle and in details like choice of binning strategy. If necessary these may be combined with such truncation procedures to meet bandwidth constraints on any particular future space mission. Of course, an instrument like FNIT has other non-astronomical applications as alluded to in Chapter 5, but its deployment as a space instrument to produce solar neutron data is its primary motivation: to be able to produce something like Figure 6.1 for FNIT, but with many more data points and better resolution. It is only by optimising these regularisation, data compression and binning strategies that such an instrument will be suitable for a deep space mission and be able to produce much better data than what we have had so far in this field. In turn this will enable us to gain the fullest possible picture of flare ion acceleration, via the exciting observational frontier of inner heliosphere neutron astronomy.

# Chapter 7

## Conclusions and Future Work

The main motivation of this thesis was to study and better understand some key energetic processes occurring at the Sun by examining the electromagnetic radiation they produce and the particles they accelerate. In Chapters 2 and 3 we delved into the hitherto unexplored world of non-thermal radiative recombination to see if it is a significant contributor to the hard X-ray flux in solar flares. Chapter 4 saw us exploring the wonders of inverse Compton scattering in the solar context to estimate the distribution of energetic particles in solar flares. In Chapter 5 we studied and analysed data from a new-age neutron detector built by our colleagues at the University of New Hampshire and in Chapter 6 we discussed methods of post-processing these data through regularisation and deconvolution techniques, whereby making such an instrument suitable for an inner heliosphere space mission.

Although each chapter provided the necessary conclusions and the possible future work in each area, we shall take the opportunity to summarise those thoughts in this chapter.

### 7.1 Recombination

Studying non-thermal recombination (NTR) took us through a fascinating journey of discovery, doubt and reconciliation. In Chapters 2 and 3 we show that NTR is an X-radiation process that has wrongly been neglected and overlooked over the last several decades. Not only does it have implications for inferred electron spectral analysis, but it could also be an answer to several questions regarding the efficiency of producing hard X-rays (HXR) in solar flares. It also has distinct diagnostic capabilities due to the spectral ‘edges’, which we predict should be visible in solar X-ray spectra. Observing such features would shed light on highly debated concepts like the low-energy electron

cut-off.

In conclusion, it is fair to say that NTR can be a significant source of HXRs in typical flare conditions. If theories of even higher Fe abundance in the corona are true, then NTR can even be dominant. Either way, it has serious implications when analysing data. Firstly, even when the edges are 20-30% of the free-free emission, they may be observable and would surely have an impact on inferred electron spectra. Spectral inversion of any sharp feature is invariably magnified. Moreover, the location of these edges can not only tell us about the presence of a low energy electron cut-off and its value but also about the temperature in a flare. For example, for a very high temperature of 40 MK, you are likely to see a distinct edge caused by Fe 25+ and Fe 26+, but at 20 MK, these edges would not be noticeable. Also, for some extreme parameters, NTR can still be the dominant source of HXRs and that would mean a rethink of fast electron numbers as well as the efficiency to produce HXRs. All in all, NTR has a bright future and significant potential.

Future work on NTR would involve making the minor alterations alluded to earlier in the thesis, e.g. using accurate measured recombination cross-sections, including the recombination effect to higher levels of the ion and so on. Once that is done, it would be possible to introduce an NTR code into present-day solar analysis packages such as SolarSoft. We feel that introducing such an option on a scientific platform, where astronomers from around the world will be able to use it, will pave NTR's way into the mainstream. Once people have the option of using it and then start using it, they are bound to benefit from it. And who knows, in the next several years may be NTR will produce some exciting and interesting results and alter our interpretation of flares and the radiation they produce.

## 7.2 Neutrons and ICS

The second half of this thesis has revolved around the study of solar neutrons in the inner heliosphere. For this endeavour of studying neutrons close to the Sun (0.2-0.4 AU) to be possible, apart from getting the adequate funding, one also needs to have the required instruments and satisfactory analysis tools to deal with such an instrument and its imminent flight into deep space. One also needs to have a basic idea of the kind of neutron flux such an instrument is likely to encounter in the inner heliosphere.

To estimate this, as a corollary, we theorised on the likely inverse Compton scattered (ICS) flux that might be observed in solar flares (Chapter 4). Our main premise was to see if relativistic electrons and positrons in the tenuous regions of the corona

could, during a flare, up-scatter photospheric photons to HXR energies. Since neutrons are produced in the same nuclear reactions that produce positrons, inferring positron distributions through our ICS analysis would further pave the way to determine likely neutron fluxes in the inner heliosphere. To carry out our study, we had to first tailor past ICS studies in other astronomical contexts to our particular one. That involved invoking the radiation field geometry of the Sun and reworking through some of the established ICS expressions to suit our situation. Our results were very interesting: although ICS fluxes are likely to be low, the ICS spectrum should be characteristically hard, and if observable, should be unmistakable. However, since the fluxes are not particularly large, it is also possible that they may not even be observable by today's instruments. Looking further, we considered a particular high coronal source, which has so far been interpreted solely through bremsstrahlung emission, and found that if we interpreted it as an ICS source, the numbers are completely viable and more than enough protons are accelerated to sufficient energies to produce such an ICS flux. Moreover, the conditions satisfy all our criteria for an ICS source, i.e. low-density, energetic flare and hard spectrum.

A lot more work can be done in pursuing this study. Firstly, it will be prudent to look for more candidate events and see if they can be interpreted in a similar fashion. Next, we should also be able to use  $\gamma$ -ray data from new instruments such as the Fermi Gamma-ray Space Telescope (FGST, formerly Gamma Ray Large Area Telescope (GLAST)) (Ritz et al. 2007) to see if they complement our ICS interpretation of energetic particle populations. And finally, we should ideally be able to deduce the relativistic positron flux in such flares, thereby giving us a handle on what neutron fluxes we may be likely to encounter on inner heliosphere space missions.

In Chapter 5, we illustrate the importance and necessity of neutron astronomy in the inner heliosphere. We introduce state-of-the-art neutron detectors being developed today, and in particular the Fast Neutron Imaging Telescope (FNIT) that has been built by our colleagues at the University of New Hampshire. FNIT is designed to detect neutrons in the 1-20 MeV range and that means its maximum benefit can only be realised in the inner heliosphere. Detecting neutrons at these energies is of particular interest because they are the most numerous and carry most of the neutron energy in this range. We discuss all aspects relevant to FNIT and, through simulated and laboratory data, show that it is an accurate and compact instrument making it an ideal candidate to go on an inner heliospheric space mission. The instrument also has a pretty diagonal response and several in-built characteristics that would allow the rejection of unwanted data.

However, like any instrument or experiment, FNIT has its limitations. To deal with these, one has to have robust data inversion methods to analyse the neutron data that FNIT would produce, once in space. Chapter 6 introduces the concept of data regularisation and suggests various algorithms to deconvolve and compress FNIT data. This involves zeroth, first and second order Tikhonov regularisation using the (Generalised) Singular Value Decomposition ((G)SVD) of the FNIT response matrix. The response matrix we create is generated from Monte Carlo simulated data, which was produced using code from the GEANT4 libraries by our colleagues at the University of Bern and New Hampshire. We show that zeroth order Tikhonov regularisation seems to work the best for FNIT. Even though the reconstruction is slightly under-smoothed, it seems to perform better than the higher order Tikhonov regularisations. We conclude that this is because for the higher order regularisations, the Picard condition is not satisfied and so choosing the smoothing parameter becomes problematic as the reconstructions seem to diverge. We also implement MATLAB's Blind Deconvolution method to establish its performance and find that it produces encouraging results. However, not being too sure of how the algorithm works, we conclude that zeroth order Tikhonov regularisation is probably the best method to implement for the deconvolution of FNIT data.

Much work is still needed to be done to make FNIT a viable instrument for a future space flight to detect solar neutrons as well as a commercial detector that can be used on Earth to detect nuclear material. The basic instrument concept is sound and all the data it produces in the lab and simulations are very encouraging. Analysing these data further and carrying out the regularisation tests we have initiated will be vital in making FNIT an instrument it deserves to be. For example, the next step would be to look at simulated data from different rotation angles of FNIT, produce a response matrix for each of those orientations, and hence create a more accurate all-round response for the instrument. In this thesis, we have looked at data only for a single orientation to illustrate the performance of regularisation methods. However, both lab and simulated data exist for a variety of instrument orientations and only by analysing all of them can a total instrument response be ascertained. Once these final steps have been undertaken and everyone involved is satisfied, then FNIT becomes a path-breaking venture that will one day give us images of the Sun we have never seen or imagined before.

# Bibliography

- Agostinelli, S., Allison, J., Amako, K., et al. 2003, Nuclear Instruments and Methods in Physics Research A, 506, 250
- Akimov, V. V., Leikov, N. G., Kurt, V. G., & Chertok, I. M. 1994, in American Institute of Physics Conference Series, Vol. 294, High-Energy Solar Phenomena - a New Era of Spacecraft Measurements, ed. J. Ryan & W. T. Vestrand, 130–133
- Alfvén, H. & Carlqvist, P. 1967, Solar Physics, 1, 220
- Amsler, C., Doser, M., Antonelli, M., Asner, D., & more than a 150 other authors. 2008, Phys. Lett., B667, 1
- Andersen, L. H., Pan, G.-Y., Schmidt, H. T., Pindzola, M. S., & Badnell, N. R. 1992, Phys. Rev. A, 45, 6332
- Ansari, S. M. R. & Alam, B. 1975, Solar Physics, 41, 97
- Arnaud, M. & Raymond, J. 1992, Astrophysical Journal, 398, 394
- Arnoldy, R. L., Kane, S. R., & Winckler, J. R. 1968, Astrophysical Journal, 151, 711
- Aschwanden, M. J. 2004, Physics of the Solar Corona. An Introduction, ed. M. J. Aschwanden (Praxis Publishing Ltd)
- Aschwanden, M. J. & Schwartz, R. A. 1996, Astrophysical Journal, 464, 974
- Badnell, N. R. 2006, Astrophysical Journal, Supplement, 167, 334
- Baker, D. N., Kanekal, S. G., Blake, J. B., & Pulkkinen, T. I. 2001, Journal of Geophysics Research, 106, 19169
- Bastian, T. S., Benz, A. O., & Gary, D. E. 1998, Annual Review of Astronomy and Astrophysics, 36, 131

- Benz, A. O. 2008, *Living Reviews in Solar Physics*, 5, 1
- Bertero, M., DeMol, C., & Pike, E. R. 1988, *Inverse Problems*, 4, 573
- Biermann, L., Haxel, O., & Schlüter, A. 1951, *Zeitschrift Naturforschung Teil A*, 6, 47
- Birn, J., Gosling, J. T., Hesse, M., Forbes, T. G., & Priest, E. R. 2000, *Astrophysical Journal*, 541, 1078
- Blumenthal, G. R. & Gould, R. J. 1970, *Reviews of Modern Physics*, 42, 237
- Bramblett, R. L., Ewing, R. I., & Bonner, T. W. 1960, *Nuclear Instruments and Methods*, 9, 1
- Bravar, U., Bruillard, P., Fluckiger, E., et al. 2007, in *Nuclear Science Symposium Conference Record, 2007. NSS '07. IEEE*, Vol. 1, 264–269
- Bravar, U., Bruillard, P. J., Flueckiger, E. O., et al. 2005, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 5901, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, ed. S. Fineschi & R. A. Viereck, 141–150
- Bray, R. J., Cram, L. E., Durrant, C., & Loughhead, R. E. 1991, *Plasma Loops in the Solar Corona*, ed. R. J. Bray, L. E. Cram, C. Durrant, & R. E. Loughhead
- Brown, D. S. & Priest, E. R. 2001, *Astronomy and Astrophysics*, 367, 339
- Brown, J. C. 1971, *Solar Physics*, 18, 489
- Brown, J. C. 1972, *Solar Physics*, 26, 441
- Brown, J. C. 1974, in *IAU Symposium*, Vol. 57, *Coronal Disturbances*, ed. G. A. Newkirk, 395–+
- Brown, J. C. 1975, in *IAU Symposium*, Vol. 68, *Solar Gamma-, X-, and EUV Radiation*, ed. S. R. Kane, 245–282
- Brown, J. C. 2005, in *Astrophysics and Space Science Library*, Vol. 320, *Astrophysics and Space Science Library*, ed. A. Hanslmeier, A. Veronig, & M. Messerotti, 87–114
- Brown, J. C. & Emslie, A. G. 1988, *Astrophysical Journal*, 331, 554

- Brown, J. C., Emslie, A. G., Holman, G. D., et al. 2006, *Astrophysical Journal*, 643, 523
- Brown, J. C. & MacKinnon, A. L. 1985, *Astrophysical Journal, Letters*, 292, L31
- Brown, J. C. & Mallik, P. C. V. 2008, *Astronomy and Astrophysics*, 481, 507
- Brown, J. C. & Mallik, P. C. V. 2009, *Astrophysical Journal, Letters*, 697, L6
- Brown, J. C. & Melrose, D. B. 1977, *Solar Physics*, 52, 117
- Brown, J. C., Spicer, D. S., & Melrose, D. B. 1979, *Astrophysical Journal*, 228, 592
- Bryan, W., Britton, C., Mihalczko, J., et al. 2003, in *Nuclear Science Symposium Conference Record, 2003 IEEE*, Vol. 2, 1192–1195 Vol.2
- Burgess, A. 1965, in *IAU Symposium, Vol. 23, Astronomical Observations from Space Vehicles*, ed. J.-L. Steinberg, 95–+
- Cane, H. V., Erickson, W. C., & Prestage, N. P. 2002, *Journal of Geophysical Research (Space Physics)*, 107, 1315
- Carrington, R. C. 1859, *Monthly Notices of the Royal Astronomical Society*, 20, 13
- Chambon, G., Hurley, K., Niel, M., et al. 1978, A hard X-ray and gamma ray observation of the 22 November 1977 solar flare, Tech. rep., IKI, Moscow
- Chandra, S. & Dularay, R. 1980, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 23, 585
- Chubb, T. A. 1972, in *Astrophysics and Space Science Library, Vol. 29, The Sun. Part 1 of Solar-Terrestrial Physics/1970*, ed. C. de Jager & E. R. Dyer, 99–+
- Chubb, T. A., Friedman, H., Kreplin, R. W., & Kupperian, J. E. 1957, *Nature*, 179, 861
- Chupp, E. L. 1971, *Space Science Reviews*, 12, 486
- Chupp, E. L. 1984, *Annual Review of Astronomy and Astrophysics*, 22, 359
- Chupp, E. L., Forrest, D. J., Higbie, P. R., et al. 1973, *Nature*, 241, 333
- Chupp, E. L., Forrest, D. J., Ryan, J. M., et al. 1982, *Astrophysical Journal, Letters*, 263, L95

- Chupp, E. L., Trotter, G., Marschhauser, H., et al. 1993, *Astronomy and Astrophysics*, 275, 602
- Clark, G. W. 1973, in *IAU Symposium, Vol. 55, X- and Gamma-Ray Astronomy*, ed. H. Bradt & R. Giacconi, 29–+
- Cliver, E. W. & Ling, A. G. 2007, *Astrophysical Journal*, 658, 1349
- Conway, A. J., Brown, J. C., Eves, B. A. C., & Kontar, E. 2003, *Astronomy and Astrophysics*, 407, 725
- Craig, I. J. D. & Brown, J. C. 1986, *Inverse problems in astronomy: A guide to inversion strategies for remotely sensed data*, ed. I. J. D. Craig & J. C. Brown
- Craig, I. J. D. & McClymont, A. N. 1991, *Astrophysical Journal, Letters*, 371, L41
- Crannell, C. J., Frost, K. J., Saba, J. L., Maetzler, C., & Ohki, K. 1978, *Astrophysical Journal*, 223, 620
- Culhane, J. L. 1969, *Monthly Notices of the Royal Astronomical Society*, 144, 375
- Culhane, J. L. & Acton, L. W. 1970, *Monthly Notices of the Royal Astronomical Society*, 151, 141
- de Boer, H., Bennett, K., Bloemen, H., et al. 1992, in *Data Analysis in Astronomy*, ed. V. di Gesu, L. Scarsi, R. Buccheri, & P. Crane, 241–249
- de Jager, C. 1969, in *Solar Flares and Space Research*, ed. C. de Jager & Z. Svestka, 1–+
- Debrunner, H., Lockwood, J. A., Barat, C., et al. 1997, *Astrophysical Journal*, 479, 997
- Delaboudinière, J., Artzner, G. E., Brunaud, J., et al. 1995, *Solar Physics*, 162, 291
- Dere, K. P., Landi, E., Mason, H. E., Monsignori Fossi, B. C., & Young, P. R. 1997, *Astronomy and Astrophysics, Supplement*, 125, 149
- Dere, K. P., Landi, E., Young, P. R., et al. 2009, *Astronomy and Astrophysics*, 498, 915
- Dermer, C. D. 1986a, *Astrophysical Journal*, 307, 47

- Dermer, C. D. 1986b, *Astronomy and Astrophysics*, 157, 223
- Desorgher, L., Flueckiger, E. O., Moser, M. R., et al. 2003, in *International Cosmic Ray Conference*, Vol. 6, *International Cosmic Ray Conference*, 3383–+
- Dulk, G. A. 1985, *Annual Review of Astronomy and Astrophysics*, 23, 169
- Dungey, J. W. 1953, *Phil. Mag.*, 44, 725
- Elwert, G. & Haug, E. 1970, *Solar Physics*, 15, 234
- Elwert, G. & Haug, E. 1971, *Solar Physics*, 20, 413
- Emslie, A. G. 1978, *Astrophysical Journal*, 224, 241
- Emslie, A. G. 1980, *Astrophysical Journal*, 235, 1055
- Emslie, A. G. & Brown, J. C. 1980, *Astrophysical Journal*, 237, 1015
- Emslie, A. G. & Brown, J. C. 1985, *Astrophysical Journal*, 295, 648
- Erdas, A., Mezzorani, G., & Quarati, P. 1993, *PHYSICAL REVIEW. A*, 48, 452
- Ethvignot, T., Devlin, M., Drog, R., et al. 2003, *Physics Letters B*, 575, 221
- Feldman, U. 1992, *Physica Scripta*, 46, 202
- Feldman, U., Dammasch, I., Landi, E., & Doschek, G. A. 2004, *Astrophysical Journal*, 609, 439
- Fishman, G. J., Meegan, C. A., Parnell, T. A., et al. 1985, in *International Cosmic Ray Conference*, Vol. 3, *International Cosmic Ray Conference*, ed. F. C. Jones, 343–346
- Fletcher, L. & Hudson, H. S. 2008, *Astrophysical Journal*, 675, 1645
- Forrest, D. J. & Chupp, E. L. 1983, *Nature*, 305, 291
- Forrest, D. J., Chupp, E. L., Ryan, J. M., et al. 1980, *Solar Physics*, 65, 15
- Friedman, H., Lichtman, S. W., & Byram, E. T. 1951, *Physical Review*, 83, 1025
- Ginzburg, V. L. & Syrovatskii, S. I. 1969, *Annual Review of Astronomy and Astrophysics*, 7, 375

- Golub, G. H., Hansen, P. C., & O'Leary, D. P. 1999, *SIAM J. Matrix Anal. Appl.*, 21, 185
- Golub, G. H. & van Loan, C. F. 1996, *Matrix Computations*, 3rd edn. (Johns Hopkins University Press)
- Gull, S. F. 1988, *Bayesian inductive inference and maximum entropy* No. 53-74 (Dordrecht: Kluwer)
- Hahn, Y. 1997, *Reports on Progress in Physics*, 60, 691
- Hahn, Y. & Krstic, P. 1994, *Journal of Physics*, B 27, L509
- Handy, B. N., Acton, L. W., Kankelborg, C. C., et al. 1999, *Solar Physics*, 187, 229
- Hansen, P. C. 1989, *BIT*, 29, 491
- Hansen, P. C. 1990, *BIT*, 30, 658
- Hey, J. S., Parsons, S. J., & Phillips, J. W. 1948, *Monthly Notices of the Royal Astronomical Society*, 108, 354
- Heyvaerts, J., Priest, E. R., & Rust, D. M. 1977, *Astrophysical Journal*, 216, 123
- Hodgson, R. 1859, *Monthly Notices of the Royal Astronomical Society*, 20, 15
- Holt, S. S. & Ramaty, R. 1969, *Solar Physics*, 8, 119
- Hoyng, P. 1977, *Astronomy and Astrophysics*, 55, 31
- Hoyng, P., Brown, J. C., Stevens, G., & van Beek, H. F. 1975, in *IAU Symposium*, Vol. 68, *Solar Gamma-, X-, and EUV Radiation*, ed. S. R. Kane, 233–235
- Hoyng, P., van Beek, H. F., & Brown, J. C. 1976, *Solar Physics*, 48, 197
- Hudson, H. & Ryan, J. 1995, *Annual Review of Astronomy and Astrophysics*, 33, 239
- Hudson, H. S., Bai, T., Gruber, D. E., et al. 1980, *Astrophysical Journal*, Letters, 236, L91
- Hudson, H. S., Kosugi, T., Nitta, N. V., & Shimojo, M. 2001, *Astrophysical Journal*, Letters, 561, L211

- Hudson, H. S., Strong, K. T., Dennis, B. R., et al. 1994, *Astrophysical Journal, Letters*, 422, L25
- Hurford, G. J., Schwartz, R. A., Krucker, S., et al. 2003, *Astrophysical Journal, Letters*, 595, L77
- Jauch, J. M. & Rohrlich, F. 1976, *The theory of photons and electrons. The relativistic quantum field theory of charged particles with spin one-half*, ed. J. M. Jauch & F. Rohrlich
- Johns, C. M. & Lin, R. P. 1992, *Solar Physics*, 137, 121
- Johnson, W. N., Kurfess, J. D., Purcell, W. R., et al. 1993, *Astronomy and Astrophysics, Supplement*, 97, 21
- Jones, F. C. 1968, *Physical Review*, 167, 1159
- Kahler, S. 1971a, *Astrophysical Journal*, 168, 319
- Kahler, S. 1971b, *Astrophysical Journal*, 164, 365
- Kahler, S. 1975, in *IAU Symposium, Vol. 68, Solar Gamma-, X-, and EUV Radiation*, ed. S. R. Kane, 211–231
- Kanbach, G., Bertsch, D. L., Favale, A., et al. 1989, *Space Science Reviews*, 49, 69
- Kanbach, G., Bertsch, D. L., Fichtel, C. E., et al. 1993, *Astronomy and Astrophysics, Supplement*, 97, 349
- Kane, S. R. 1974, in *IAU Symposium, Vol. 57, Coronal Disturbances*, ed. G. A. Newkirk, 105–141
- Kane, S. R. & Anderson, K. A. 1970, *Astrophysical Journal*, 162, 1003
- Kane, S. R., Crannell, C. J., Datlowe, D., et al. 1980, in *Skylab Solar Workshop II*, ed. P. A. Sturrock, 187–229
- Kazachevskaya, T. V. & Ivanov-Kholodnyi, G. S. 1959, *Soviet Astronomy*, 3, 937
- Kichatinov, L. L. 1991, *Solar Physics*, 133, 177
- Koch, H. W. & Motz, J. W. 1959, *Rev. Mod. Phys.*, 31, 920

- Kocharov, L., Debrunner, H., Kovaltsov, G., et al. 1998, *Astronomy and Astrophysics*, 340, 257
- Kontar, E. P. & Brown, J. C. 2006, *Advances in Space Research*, 38, 945
- Kontar, E. P., Emslie, A. G., Piana, M., Massone, A. M., & Brown, J. C. 2005, *Solar Physics*, 226, 317
- Kontar, E. P., MacKinnon, A. L., Schwartz, R. A., & Brown, J. C. 2006, *Astronomy and Astrophysics*, 446, 1157
- Kontar, E. P., Piana, M., Massone, A. M., Emslie, A. G., & Brown, J. C. 2004, *Solar Physics*, 225, 293
- Korchak, A. A. 1967, *Soviet Astronomy*, 11, 258
- Korchak, A. A. 1971, *Solar Physics*, 18, 284
- Kosugi, T., Masuda, S., Makishima, K., et al. 1991, *Solar Physics*, 136, 17
- Kosugi, T., Matsuzaki, K., Sakao, T., et al. 2007, *Solar Physics*, 243, 3
- Kramers, H. A. 1923, *Phil. Mag.*, 46, 836
- Krucker, S., Battaglia, M., Cargill, P. J., et al. 2008a, *Astronomy and Astrophysics Reviews*, 16, 155
- Krucker, S., Christe, S., Lin, R. P., Hurford, G. J., & Schwartz, R. A. 2002, *Solar Physics*, 210, 445
- Krucker, S., Hurford, G. J., MacKinnon, A. L., Shih, A. Y., & Lin, R. P. 2008b, *Astrophysical Journal, Letters*, 678, L63
- Krucker, S. & Lin, R. P. 2008, *Astrophysical Journal*, 673, 1181
- Krucker, S., White, S. M., & Lin, R. P. 2007, *Astrophysical Journal, Letters*, 669, L49
- Krucker, S., Wuelser, J.-P., Vourlidas, A., et al. 2008c, 12th European Solar Physics Meeting, Freiburg, Germany, held September, 8-12, 2008. Online at <http://espm.kis.uni-freiburg.de/>, p.2.84, 12, 2
- Kuznetsov, S. N., Kurt, V. G., Yushkov, B. Y., et al. 2005, in *International Cosmic Ray Conference*, Vol. 1, International Cosmic Ray Conference, 49–+

- Landi, E. 2007, *Astronomy and Astrophysics*, 476, 675
- Landini, M., Monsignori Fossi, B. C., & Pallavicini, R. 1973, *Solar Physics*, 29, 93
- Lin, R. P. 1974, *Space Science Reviews*, 16, 189
- Lin, R. P. 1998, *Space Science Reviews*, 86, 61
- Lin, R. P., Dennis, B. R., Hurford, G. J., et al. 2002, *Solar Physics*, 210, 3
- Lin, R. P. & Hudson, H. S. 1976, *Solar Physics*, 50, 153
- Lin, R. P., Levedahl, W. K., Lotko, W., Gurnett, D. A., & Scarf, F. L. 1986, *Astrophysical Journal*, 308, 954
- Lin, R. P., Potter, D. W., Gurnett, D. A., & Scarf, F. L. 1981, *Astrophysical Journal*, 251, 364
- Lin, R. P. & Schwartz, R. A. 1987, *Astrophysical Journal*, 312, 462
- Lingenfelter, R. E. 1969, *Solar Physics*, 8, 341
- MacKinnon, A. L. 2006, Washington DC American Geophysical Union Geophysical Monograph Series, 165, 157
- MacKinnon, A. L. & Mallik, P. C. V. 2009, ArXiv e-prints
- MacKinnon, A. L. & Ryan, J. M. 2009, *Solar Physics*, In preparation
- Macri, J., Bravar, U., Legere, J., et al. 2007, in Nuclear Science Symposium Conference Record, 2007. NSS '07. IEEE, Vol. 1, 173–178
- Maetzler, C., Bai, T., Crannell, C. J., & Frost, K. J. 1978, *Astrophysical Journal*, 223, 1058
- Mallik, P., Brown, J. C., & MacKinnon, A. L. 2009a, in AAS/Solar Physics Division Meeting, Vol. 40, AAS/Solar Physics Division Meeting, Abstract 37.07
- Mallik, P., Brown, J. C., & MacKinnon, A. L. 2009b, in AAS/Solar Physics Division Meeting, Vol. 40, AAS/Solar Physics Division Meeting, Abstract 20.04
- Mallik, P. C. V. & Brown, J. C. 2007, in Royal Astronomical Society NAM, ed. U. of Central Lancashire, Abstract P20.6

- Mandzhavidze, N. & Ramaty, R. 1992, *Astrophysical Journal*, 389, 739
- Mariska, J. T. 1992, *The solar transition region*, ed. J. T. Mariska
- Marschhauser, H., Rieger, E., & Kanbach, G. 1991, in *International Cosmic Ray Conference*, Vol. 3, *International Cosmic Ray Conference*, 61–+
- Masson, S., Klein, K.-L., Bütikofer, R., et al. 2009, *Solar Physics*, 257, 305
- Massone, A. M., Emslie, A. G., Kontar, E. P., et al. 2004, *Astrophysical Journal*, 613, 1233
- McClements, K. G. & Brown, J. C. 1986, *Astronomy and Astrophysics*, 165, 235
- Mewe, R., Lemen, J. R., & van den Oord, G. H. J. 1986, *Astronomy and Astrophysics*, Supplement, 65, 511
- Miller, J. A. & Ramaty, R. 1989, *Astrophysical Journal*, 344, 973
- Moser, M. R., Flückiger, E. O., Ryan, J. M., Macri, J. R., & McConnell, M. L. 2005a, *Advances in Space Research*, 36, 1399
- Moser, M. R., Ryan, J. M., Bravar, U., et al. 2005b, in *International Cosmic Ray Conference*, Vol. 1, *International Cosmic Ray Conference*, 265–+
- Moser, M. R., Ryan, J. M., Desorgher, L., & Flückiger, E. O. 2005c, in *International Cosmic Ray Conference*, Vol. 2, *International Cosmic Ray Conference*, 421–+
- Moses, D., Droege, W., Meyer, P., & Evenson, P. 1989, *Astrophysical Journal*, 346, 523
- Moskalenko, I. V. & Strong, A. W. 2000, *Astrophysical Journal*, 528, 357
- Mrozek, T. & Tomczak, M. 2004, *Astronomy and Astrophysics*, 415, 377
- Muraki, Y., Murakami, K., Miyazaki, M., et al. 1992, *Astrophysical Journal*, Letters, 400, L75
- Murphy, R. J., Dermer, C. D., & Ramaty, R. 1987, *Astrophysical Journal*, Supplement, 63, 721
- Murphy, R. J., Share, G. H., Grove, J. E., et al. 1997, *Astrophysical Journal*, 490, 883
- Nakajima, H., Nishio, M., Enome, S., et al. 1994, *IEEE Proceedings*, 82, 705

- Nelson, G. J. & Melrose, D. B. 1985, Type II bursts, ed. McLean, D. J. & Labrum, N. R., 333–359
- Nishio, M., Nakajima, H., Enome, S., et al. 1994, in Proceedings of Kofu Symposium, 19–33
- Orlando, E. & Strong, A. W. 2008, *Astronomy and Astrophysics*, 480, 847
- Orwig, L. E., Frost, K. J., & Dennis, B. R. 1980, *Solar Physics*, 65, 25
- Parker, E. N. 1963, *Astrophysical Journal, Supplement*, 8, 177
- Petschek, H. E. 1964, *NASA Special Publication*, 50, 425
- Phillips, K. J. H. 1995, *Guide to the Sun*, ed. K. J. H. Phillips
- Piana, M. 1994, *Astronomy and Astrophysics*, 288, 949
- Piana, M., Barrett, R., Brown, J. C., & McIntosh, S. W. 1999, *Inverse Problems*, 15, 1469
- Piana, M., Massone, A. M., Kontar, E. P., et al. 2003, *Astrophysical Journal, Letters*, 595, L127
- Pirard, B., Woolf, R. S., Bravar, U., et al. 2009, *Nuclear Instruments and Methods in Physics Research A*, 603, 406
- Pottasch, S. R. 1967, *Bulletin Astronomical Institute of the Netherlands*, 19, 113
- Prato, M., Piana, M., Brown, J. C., et al. 2006, *Solar Physics*, 237, 61
- Priest, E. & Forbes, T. 2000, *Magnetic Reconnection*, ed. E. Priest & T. Forbes
- Priest, E. R. & Forbes, T. G. 2002, *Astronomy and Astrophysics Reviews*, 10, 313
- Priest, E. R. & Heyvaerts, J. 1974, *Solar Physics*, 36, 433
- Ramaty, R., Paizis, C., Colgate, S. A., et al. 1980, in *Skylab Solar Workshop II*, ed. P. A. Sturrock, 117–185
- Ramaty, R., Schwartz, R. A., Enome, S., & Nakajima, H. 1994, *Astrophysical Journal*, 436, 941
- Reames, D. V. 1999, *Space Science Reviews*, 90, 413

- Ritz, S. M., Michelson, P. F., Meegan, C., Grindlay, J., & GLAST Mission Team. 2007, in *Bulletin of the American Astronomical Society*, Vol. 38, *Bulletin of the American Astronomical Society*, 909–+
- Rogers, R. D. 1991, *Astrophysical Journal*, 383, 550
- Ryan, J., Bennett, K., Debrunner, H., et al. 1993, *Advances in Space Research*, 13, 255
- Ryan, J., Bravar, U., Bruillard, P., et al. 2008, *APS Meeting Abstracts*, 8002
- Ryan, J. M., Aarts, H., Bennett, K., et al. 1992, *NASA Conference Publication*, 3137, 470
- Ryan, J. M., Bravar, U., Macri, J. R., et al. 2007, *AGU Fall Meeting Abstracts*, A846+
- Ryan, J. M., Lockwood, J. A., & Debrunner, H. 2000, *Space Science Reviews*, 93, 35
- Rybicki, G. B. & Lightman, A. P. 1986, *Radiative Processes in Astrophysics*, ed. G. B. Rybicki & A. P. Lightman
- Saint-Hilaire, P. & Benz, A. O. 2002, *Solar Physics*, 210, 287
- Savin, D. W. & Laming, J. M. 2002, *Astrophysical Journal*, 566, 1166
- Saxena, R. 1990, PhD thesis, UNIVERSITY OF NEW HAMPSHIRE.
- Scherrer, P. H., Bogart, R. S., Bush, R. I., et al. 1995, *Solar Physics*, 162, 129
- Schoenfelder, V., Aarts, H., Bennett, K., et al. 1993, *Astrophysical Journal, Supplement*, 86, 657
- Schumacher, J., Kliem, B., & Seehafer, N. 2000, *Physics of Plasmas*, 7, 108
- Silva, A. V. R., Share, G. H., Murphy, R. J., et al. 2007, *Solar Physics*, 245, 311
- Simnett, G. M. 1974, *Space Science Reviews*, 16, 257
- Simnett, G. M. 1995, *Space Science Reviews*, 73, 387
- Simnett, G. M. 2001, *Space Science Reviews*, 97, 231
- Smith, D. F. & Lilliequist, C. G. 1979, *Astrophysical Journal*, 232, 582
- Spicer, D. S. 1976, An unstable arch model of a solar flare, Tech. rep.

- Strong, A. W. 1996, *Space Science Reviews*, 76, 205
- Strong, K., Bruner, M., Tarbell, T., Title, A., & Wolfson, C. J. 1994, *Space Science Reviews*, 70, 119
- Sturrock, P. A. 1968, in *IAU Symposium, Vol. 35, Structure and Development of Solar Active Regions*, ed. K. O. Kiepenheuer, 471–+
- Sturrock, P. A. 1979, *NASA STI/Recon Technical Report N*, 79, 23873
- Sturrock, P. A. & Coppi, B. 1966, *Astrophysical Journal*, 143, 3
- Sunyaev, R. A. & Zeldovich, Y. B. 1970, *Astrophysics and Space Science*, 7, 3
- Suzuki, S. & Dulk, G. A. 1985, *Bursts of Type III and Type V*, ed. McLean, D. J. & Labrum, N. R., 289–332
- Svestka, Z. 1976, *Geophysics and Astrophysics Monographs*, 8
- Sweet, P. A. 1958, in *IAU Symposium, Vol. 6, Electromagnetic Phenomena in Cosmical Physics*, ed. B. Lehnert, 123–+
- Sweet, P. A. 1969, *Annual Review of Astronomy and Astrophysics*, 7, 149
- Sweet, P. A. 1971, in *IAU Symposium, Vol. 43, Solar Magnetic Fields*, ed. R. Howard, 457–+
- Takakara, T. & Kai, K. 1966, *Publications of the ASJ*, 18, 57
- Talon, R., Trottet, G., Vilmer, N., et al. 1993, *Solar Physics*, 147, 137
- Tanaka, K. & Nakagawa, Y. 1973, *Solar Physics*, 33, 187
- Tanaka, Y. 1983, *Solar Physics*, 86, 3
- Thompson, A., Brown, J., Kay, J., & Titterton, D. 1991, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13, 326
- Thompson, A. M. & Craig, I. J. D. 1992, *Astronomy and Astrophysics*, 262, 359
- Tikhonov, A. N. 1963, *Soviet Math. Dokl.*, 4, 1035
- Tomczak, M. 2009, *ArXiv e-prints*

- Toner, M. P., Ryan, J. M., MacKinnon, A. L., et al. 2001, in IAU Symposium, Vol. 203, Recent Insights into the Physics of the Sun and Heliosphere: Highlights from SOHO and Other Space Missions, ed. P. Brekke, B. Fleck, & J. B. Gurman, 573–+
- Tucker, W. H. 1975, Radiation processes in astrophysics, ed. W. H. Tucker
- Turchin, V. F. 1967, USSR Computational Mathematics and Mathematical Physics, 7, 79
- Švestka, Z. & Fritzová-Švestková, L. 1974, Solar Physics, 36, 417
- van Beek, H. F., Hoyng, P., Lafleur, B., & Simnett, G. M. 1980, Solar Physics, 65, 39
- Veronig, A. M. & Brown, J. C. 2004, Astrophysical Journal, Letters, 603, L117
- Vilmer, N. & MacKinnon, A. L. 2003, in Lecture Notes in Physics, Berlin Springer Verlag, Vol. 612, Energy Conversion and Particle Acceleration in the Solar Corona, ed. L. Klein, 127–160
- Vilmer, N., MacKinnon, A. L., Trotter, G., & Barat, C. 2003, Astronomy and Astrophysics, 412, 865
- Vlahos, L. & Papadopoulos, K. 1979, Astrophysical Journal, Letters, 234, L217+
- Watanabe, K., Muraki, Y., Matsubara, Y., et al. 2003, Astrophysical Journal, 592, 590
- Watt, B. E. 1952, Phys. Rev., 87, 1037
- Weaver, T. A. 1976, Physical Review A: General Physics, 13, 1563
- White, S. M., Thomas, R. J., Brosius, J. W., & Kundu, M. R. 2000, Astrophysical Journal, Letters, 534, L203
- Wild, J. P. 1950, Australian Journal of Scientific Research A Physical Sciences, 3, 399
- Wolf, R., Ryan, J., Bloser, P., et al. 2009, APS Meeting Abstracts, C8003+
- Wolf, R. S., Bravar, U., Macri, J. R., et al. 2008, AGU Spring Meeting Abstracts, A21+
- Zhang, J. & Huang, G. L. 2004, Solar Physics, 219, 135

# Appendix A

## Dielectronic and “Inverse” Recombination

During the course of doing the work pertaining to this thesis in Chapters 2 and 3, we have often been asked the question “what about dielectronic recombination?” So in this Appendix we shall address that question briefly and also introduce a new concept, namely “inverse” recombination.

### A.1 Dielectronic Recombination

In all the details we have expounded on in Chapters 2 and 3, we have used the word ‘recombination’ as a short-hand for radiative recombination. In radiative recombination, a free electron binds on to an ion and in the process releases one photon, which has the continuum energy of the electron plus the ionisation potential of the level to which the electron has recombined. We always took this level to be the lowest available principal quantum state of the particular ion because the recombination cross-section drops off as  $1/n^3$  and so higher available states have a more and more negligible probability of being recombined in to. However, while talking about recombination processes, dielectronic recombination can be an equally important physical process (e.g. Burgess 1965; Ansari & Alam 1975). Here, an electron is captured by an ion - as in the radiative process - but occupies an excited state in the ion *and* a portion of the captured electron’s energy goes into exciting a bound electron of that ion into an excited state as well. So you then have an ion with two excited electrons. One of two things can happen thereafter: 1) each electron de-excites to a lower state either directly or by cascading down and thereby photons of corresponding energies are emitted or 2) the process simply reverses, i.e. auto-ionisation occurs whereby the excited electron goes back to its original state and the captured electron is released and the ion is back to

its initial state. Process 1) is dielectronic recombination and this produces photons. Quite often, the cross-section for dielectronic recombination is greater than for radiative recombination for given flare parameters (Mariska 1992), and this prompted us to probe whether we were neglecting something that could turn out to be important for our study.

The importance of dielectronic recombination for astrophysical cases and the solar corona in particular was first explored by Burgess (1965). It has since been elaborated upon by various authors (e.g. Pottasch 1967; Ansari & Alam 1975; Chandra & Dularey 1980) and it is clear from these that it is an important physical process to consider. However, since the initial recombination is radiationless and instead the excess electron energy is used to excite an electron within the ion, the photons eventually emitted are of energies pertaining to the energy levels of the ion and not of the initial electron energy. Hence the energies we are talking about are in the UV and EUV regime. This makes dielectronic recombination a major contributor to UV and EUV line emission in flares and can be used as another diagnostic tool to determine elemental abundances in the solar atmosphere. Various studies looking at more detailed dielectronic recombination rate coefficients and their uncertainties have found that determining the upper solar and stellar atmosphere elemental abundances can be affected by these uncertainties (Savin & Laming 2002). Hence the potential of further study on dielectronic recombination for ionisation balance calculations and abundance diagnostics cannot be overstated. However, the emission this process produces is not in the same energy range as that of the recombining electron and hence does not affect our estimation of the importance of radiative recombination in better understanding the flare X-ray continuum and the production mechanism of such radiation in solar flares.

## A.2 “Inverse” Recombination

Physically, the motion of fast ions at a speed  $v$  through plasma electrons is equivalent to the motion of fast electrons of speed  $v$  through ions, each interaction generating the same radiation apart from a Doppler shift. For example, Emslie & Brown (1985) show HXR production in solar flares produced by high energy proton beams. This process could be called an “inverse” bremsstrahlung process.

Inspired by that, as a corollary to our work on non-thermal recombination in Chapters 2 and 3, we decided to consider the ‘inverse’ process for NTR, viz. recombination of ambient electrons on to a fast ion. It is important to note that the fast ion will be fully ionised by collisions. This scenario is equivalent to recombination of a fast elec-

tron onto an ambient ion for the same speed  $v$ , except that the ion has much higher kinetic energy and range, so we get many more photons per ion. In other words, this is an “inverse” non-thermal recombination process.

Using the example of an Fe ion and assuming an equivalent electron having typical HXR-producing energy of  $E_e = 20$  keV, the Fe ion would have kinetic energy

$$E_{Fe} = 56 \times \frac{m_p}{m_e} \times \frac{1}{2} m_e v^2, \quad (\text{A.1})$$

where  $m_p$  is the mass of a proton. This gives approximately  $10^5 E_e \approx 2$  GeV. For any ion in general of atomic mass  $A$ ,

$$E_i = A \frac{m_p}{m_e} E_e = A m_p v^2 / 2. \quad (\text{A.2})$$

Now, considering the energy loss rate,

$$\frac{dE}{dt} = \frac{-K_i n v}{E}; \quad \text{or} \quad \frac{dE}{dN} = \frac{-K_i}{E}, \quad (\text{A.3})$$

and stopping column density  $N_i = E_i^2 / 2K_i$ , with  $dN = n v dt$ , [Emslie \(1978\)](#) gives

$$\frac{dE}{dt} = \frac{-2\pi Z^2 z^2 e^4 \Lambda}{E} \frac{m}{M} n v, \quad (\text{A.4})$$

where the particle in motion has mass  $m$  and charge  $ze$  and the target particle has mass  $M$  and charge  $Ze$ ;  $\Lambda$  is the Coulomb logarithm. So, in this case, the ion has mass  $m = A m_p$  and  $z = Z_i$  whereas the electron has mass  $M = m_e$  and  $Z = 1$ .

This gives, from Equation [A.4](#),

$$\frac{dE}{dt} = \frac{-2\pi z^2 e^4 \Lambda}{E} A \frac{m_p}{m_e} n v; \quad K_i = A Z^2 \frac{m_p}{m_e} K_e. \quad (\text{A.5})$$

Since  $N = \frac{E^2}{2K}$ , the ratio of stopping column density  $N_i/N_e$  (ranges) of ion to electron is

$$\frac{N_i}{N_e} = \left( \frac{E_i}{E_e} \right)^2 \frac{K_e}{K_i} = \frac{A m_p}{Z^2 m_e}. \quad (\text{A.6})$$

For Fe, this gives a ratio of approximately 200. What this means is that an Fe ion at  $v = 10^{10}$  cms<sup>-1</sup> has about 200 times the range of an electron with the same velocity, so will emit about 200 times as many recombination photons in its lifetime. Therefore, in a thick target, we need an ion injection rate of only 1/200<sup>th</sup> the necessary electron injection rate for the same NTR output. This sounds promising, but it is worth keeping in mind that each Fe ion carries much more energy than the electron of the same  $v$ ,

in fact by a factor  $Am_p/m_e \simeq 10^5$ . So inverse NTR would reduce the necessary beam particle flux  $\sim 200$  times, but would worsen the beam power problem by a factor

$$\frac{Am_p}{m_e} \times \frac{Z^2 m_e}{Am_p} = Z^2, \quad (\text{A.7})$$

which is about 600 for Fe. However, if we consider “inverse” NTR with H ions, the beam power problems do not exist. And the ratio in Equation A.6 is *better* by a factor of 10, i.e.  $N_p/N_e = 2000$ . But the cross-section for H ions is smaller than Fe by a factor of  $Z^4 = 4.6 \times 10^5$ , while the energy per proton is 56 times smaller than Fe. Therefore the net gain for hydrogen compared to an ion is  $1/(Z^2 A^2)$ , which is always less than unity, therefore making hydrogen even more inefficient in producing “inverse” non-thermal recombination emission than ions such as Fe.

# Appendix B

## Derivation of Relevant ICS Integral Formulae

Moskalenko & Strong (2000) do not derive the following equations but simply state that evaluating the integrals are “trivial” and jump from Equation 4.10 to Equation 8 in their paper over two steps. However, we felt the need to derive and solve those integrals in order to understand how to apply them in the solar context. Hence, in this Appendix, you shall find a step-by-step analysis (well, almost!) of how we got from Equation 4.10 to Equation 4.12.

Moskalenko & Strong (2000) give the  $\delta$ -function 3-dimensional electron energy distribution as

$$f_e(\gamma_1, \Omega_e) = \frac{1}{4\pi\gamma_1^2} \delta(\gamma_1 - \gamma) \quad (\text{B.1})$$

and we also use a Heaviside step-function for our photon distribution given by Equation 4.11.

So the ICS photon distribution from Equation 4.10 then becomes

$$\begin{aligned} \frac{dR}{d\epsilon_2} &= \frac{r_e^2}{2\gamma\epsilon_1\epsilon_2} \int d\cos\eta' \int d\Omega_\gamma Q_\gamma(\theta, \phi) \epsilon'_1 \epsilon'_2 \left(\frac{\epsilon'_2}{\epsilon'_1}\right)^2 \\ &\times \left(\frac{\epsilon'_2}{\epsilon'_1} + \frac{\epsilon'_1}{\epsilon'_2} - \sin^2\eta'\right) \delta\left[\epsilon'_2 - \frac{\epsilon'_1}{1 + \epsilon'_1(1 - \cos\eta')}\right]. \end{aligned} \quad (\text{B.2})$$

From Equation 4.16, we can then get the argument of the  $\delta$ -function, which we call  $X$ , as

$$X = \frac{\epsilon_2}{\gamma(1 - \cos\eta')} - \frac{\epsilon'_1}{1 + \epsilon'_1(1 - \cos\eta')}. \quad (\text{B.3})$$

Letting  $\cos\eta' = u$  and solving for  $X = 0$ , we get

$$\frac{\epsilon_2}{\gamma(1-u)} = \frac{\epsilon'_1}{1 + \epsilon'_1(1-u)}. \quad (\text{B.4})$$

Rearranging and solving for  $u$ , we get

$$u = \frac{\epsilon'_1\gamma - \epsilon_2 - \epsilon'_1\epsilon_2}{\epsilon'_1(\gamma - \epsilon_2)} \equiv u_*. \quad (\text{B.5})$$

Also,

$$dX = \left( \frac{\epsilon_2}{\gamma(1-u)^2} - \frac{\epsilon_1'^2}{[1 + \epsilon'_1(1-u)]^2} \right) du \equiv Y(u)du. \quad (\text{B.6})$$

Then the up-scattered photon flux can be written as

$$\begin{aligned} \frac{dR}{d\epsilon_2} &= \frac{r_e^2}{\gamma\epsilon_1\epsilon_2} \int \frac{dX}{Y} \int d\Omega_\gamma Q_\gamma(\theta, \phi) \frac{\epsilon_2'^3}{\epsilon_1'} \left( \frac{\epsilon_2'}{\epsilon_1'} + \frac{\epsilon_1'}{\epsilon_2'} - 1 + u^2 \right) \delta(X) \\ &= \frac{r_e^2}{\gamma\epsilon_1\epsilon_2} \int d\Omega_\gamma Q_\gamma(\theta, \phi) Y^{-1}(X=0) \frac{\epsilon_2'^3}{\epsilon_1'} \left( \frac{\epsilon_2'}{\epsilon_1'} + \frac{\epsilon_1'}{\epsilon_2'} - 1 + u^2|_{X=0} \right). \end{aligned} \quad (\text{B.7})$$

Using this condition and substituting  $u$  with  $u_*$ , we get

$$Y = \frac{(\gamma - \epsilon_2)^2}{\gamma^2\epsilon_2} [\gamma\epsilon_1'^2 - \epsilon_2\epsilon_1'^2]. \quad (\text{B.8})$$

Using Equation 4.16, we can obtain

$$\epsilon_2' = \frac{\epsilon_1'(\gamma - \epsilon_2)}{\gamma}. \quad (\text{B.9})$$

Since  $\sin^2 \eta' = 1 - \cos^2 \eta'$ , we also obtain, after a lot of algebra,

$$\sin^2 \eta' = \frac{2\epsilon_1'^2\gamma^3 - 4\epsilon_1'^2\epsilon_2\gamma^2 - \epsilon_1'^2\epsilon_2^3 + 3\epsilon_1'^2\epsilon_2^2\gamma - 2\epsilon_1'\epsilon_2\gamma^2 + \epsilon_2^2\gamma + 2\epsilon_1'\epsilon_2^2\gamma}{\epsilon_1'^2\gamma(\gamma - \epsilon_2)^2}. \quad (\text{B.10})$$

Then, for the ICS up-scattered photon flux from Equation B.7, we arrive at

$$\begin{aligned} \frac{dR}{d\epsilon_2} &= \frac{r_e^2}{2\epsilon_1\epsilon_2} \int d\Omega_\gamma Q_\gamma(\theta, \phi) \frac{\epsilon_2[\epsilon_1'^2(\gamma - \epsilon_2)^3 + \epsilon_1'^2\gamma^2(\gamma - \epsilon_2) - \epsilon_2\gamma(2\epsilon_1'\gamma - \epsilon_2 - 2\epsilon_1'\epsilon_2)]}{\gamma^2(\gamma - \epsilon_2)(\epsilon_1'^2\gamma - \epsilon_1'^2\epsilon_2)} \\ &= \frac{r_e^2}{2\epsilon_1(\gamma - \epsilon_2)^2} \frac{1}{\gamma} \int d\Omega_\gamma Q_\gamma(\theta, \phi) \left[ \frac{(\gamma - \epsilon_2)^3}{\gamma^2} + (\gamma - \epsilon_2) - \frac{\epsilon_2(2\epsilon_1'\gamma - \epsilon_2 - 2\epsilon_1'\epsilon_2)}{\epsilon_1'^2\gamma} \right] \\ &= \frac{r_e^2}{2\epsilon_1(\gamma - \epsilon_2)^2} \int_{\Omega_\gamma} d\Omega_\gamma Q_\gamma(\theta, \phi) \left[ 2 - \frac{2\epsilon_2}{\gamma} \left( \frac{1}{\epsilon_1'} + 2 \right) + \frac{\epsilon_2^2}{\gamma^2} \left( \frac{1}{\epsilon_1'^2} + \frac{2}{\epsilon_1'} + 3 \right) - \frac{\epsilon_2^3}{\gamma^3} \right] \end{aligned} \quad (\text{B.11})$$

which is the exact expression (accounting for the different photon distribution) that [Moskalenko & Strong \(2000\)](#) get in Equation 8 of their paper. It is from this equation above that we proceeded to integrate over  $d\Omega$  and obtained the expressions (4.12 and 4.15) we used to calculate our estimation of the ICS flux in the solar context. The isotropy in the outward hemisphere of the solar radiation field makes further simplification possible in our case, if we divert from the [Moskalenko & Strong \(2000\)](#) road and make explicit use of this high degree of symmetry.