The origin of 26Al in the Galaxy

Jürgen Knödlseder

THE ORIGIN OF ^{26}Al IN THE GALAXY

L'origine de l'^{26}Al galactique

THESE

présentée

à l'UNIVERSITE PAUL SABATIER DE TOULOUSE

pour l'obtention du diplôme de

DOCTORAT DE L'UNIVERSITE PAUL SABATIER

Spécialité : Astrophysique et techniques spatiales

par

Jürgen Knödlseder

Centre d'Etude Spatiale des Rayonnements
9, avenue Colonel-Roche
BP 4346
31028 Toulouse CEDEX
FRANCE

Soutenue le 24-11-1997 à 16.30 heures devant le jury composé de :

G. VEDRENNE	Professeur à l'UPS, Toulouse	Président
M. CASSE	Ingénieur de Recherche, CEA Saclay	Rapporteur
A. MAEDER	Professeur à l'Université de Genève	Rapporteur
R. DIEHL	Directeur de Recherche, MPE Garching	Examinateur
M. GIARD	Chargé de Recherche, CESR Toulouse	Examinateur
N. PRANTZOS	Chargé de Recherche, CNRS/IAP Paris	Examinateur
J.-P. ROQUES	Chargé de Recherche, CESR Toulouse	Examinateur
S. VAUCLAIR	Professeur à l'UPS, Toulouse	Examinatrice
P. VON BALLMOOS	Professeur à l'UPS, Toulouse	Examinateur

The reasearch described in this thesis was supported by the European Community
through grant number ERBFMBICT 950387.

To my parents

To Nathalie

Contents

Abstract

The history of recent galactic nucleosynthesis activity can be studied by measurements of the 1.809 MeV gamma-ray line arising from the decay of radioactive ^{26}Al. The COMPTEL telescope aboard the Compton Gamma-Ray Observatory, launched on April 5, 1991, permits for the first time an extensive investigation of the 1.8 MeV radiation throughout the entire sky. The aim of this thesis is to infer the galactic distribution of ^{26}Al from these measurements and to identify the dominant sources of this radioactive isotope.

The first part of the thesis is dedicated to the reconstruction of the 1.8 MeV intensity distribution from the measured data. It is demonstrated that the use of conventional deconvolution algorithms, like maximum likelihood or maximum entropy inversion, leads to lumpy, noise-dominated intensity distributions. Nevertheless, simulations can help to assess the uncertainties in the reconstructed images, which permits the scientific exploitation of the recovered skymaps. Alternatively, a multiresolution approach is proposed, which largely reduces the uncertainties in the reconstructed 1.8 MeV intensity distribution. In summary, 1.8 MeV emission is mainly concentrated towards the galactic plane, which clearly demonstrates that the bulk of ^{26}Al is of galactic rather than local origin. However, distinct emission features towards Cygnus, Carina, and the Auriga-Camelopardalis-Perseus region are inconsistent with a smooth galactic 1.8 MeV emission profile, pointing towards a massive star origin of ^{26}Al.

The second part of the thesis consists of a multi-wavelength comparison of COMPTEL 1.8 MeV data which aims in the identification of the origin of galactic ^{26}Al. For the comparison, a rigorous Bayesian analysis is applied, which is the only consistent framework that allows inference based on the comparison. It turned out that the 1.8 MeV distribution follows very closely the distribution of free electrons in the Galaxy which is traced by thermal bremsstrahlung, observable in the microwave domain. The similarity of the 1.8 MeV intensity distribution to the thermal bremsstrahlung distribution implies a direct proportionality between the ^{26}Al and the massive star column densities, which strongly supports that massive stars are the origin of galactic ^{26}Al. In particular, ONeMg-novae and AGB stars can be excluded as dominant ^{26}Al sources since their galactic distribution is not expected to correlate with the distribution of free electrons.

The correlation between ^{26}Al and free electrons established, the analysis of 1.8 MeV γ-ray line emission can complement our knowledge about star formation and the distribution of ionized gas throughout the entire Galaxy. While COMPTEL made the first step in providing the first all-sky map in the light of the 1.809 MeV line, INTEGRAL, the next generation γ-ray spectrometer, will allow a detailed study of current star formation in the Galaxy.

Résumé

L'émission de la raie gamma à 1,809 MeV, correspondante à la décroissance du radio-isotope ^{26}Al, témoigne de la synthèse récente d'éléments lourds dans notre Galaxie. Une étude approfondie de cette émission a pu être réalisée grâce au télescope COMPTEL, mis en orbite en avril 1991. En particulier, je présente dans cette thèse une étude de l'origine et de la distribution de l'^{26}Al dans la Galaxie.

La première partie de cette thèse est consacrée à la reconstruction de l'image du ciel dans la lumière de la raie à 1,8 MeV. Les algorithmes conventionnels, comme la méthode de maximum d'entropie ou de vraisemblance maximale, mènent à des images bruitées, dominées par des fluctuations statistique du bruit de fond. Afin d'exploiter scientifiquement celles-ci, il a été nécessaire d'effectuer des simulations en particulier pour déterminer les incertitudes. De plus, une approche multi-résolution permettant une réduction significative des incertitudes est proposée dans la thèse. Il ressort de cette analyse, que l'émission à 1,8 MeV est concentré dans le plan galactique, ce qui montre clairement que l'origine de l'^{26}Al est galactique et non locale. Cependant, l'émission distinct vers la région du Cygne, la Carène, et le complexe du Cocher-Girafe-Persée est incompatible avec des fluctuations statistique, indiquant les étoiles massives comme géniteurs principaux de l'^{26}Al.

La deuxième partie de cette thèse consiste dans une comparaison des données COMPTEL à 1,8 MeV avec d'autres longueurs d'onde avec le but d'identifier l'origine de l'^{26}Al galactique. Pour effectuer cette comparaison, j'ai utilisé une approche Bayesienne qui est l'unique cadre consistent permettant des conclusions basées sur la comparaison. Cette comparaison a montré que la distribution de la raie à 1,8 MeV ressemble fortement à la distribution des électrons libres, tracée par le rayonnement de freinage thermique qui est observé dans la domaine des microondes. Cette ressemblance implique une proportionnalité directe entre la colonne densité de l'^{26}Al et des étoiles massives, supportant fortement que les étoiles massives soient à l'origine de l'^{26}Al galactique. En particulière, les novae O-Ne-Mg et les étoiles AGB peuvent être exclues comme géniteurs important de l'^{26}Al, car ils ne sont pas corrélés à la distribution des électrons libres.

La corrélation entre l'^{26}Al et les électrons libres étant établie, l'analyse de la raie gamma à 1,8 MeV peut compléter nos connaissances sur la formations d'étoiles et la distribution du gaz ionisé dans tout la Galaxie. Si le télescope COMPTEL a produit la première carte du ciel dans la raie gamma à 1,8 MeV, le prochain instrument, INTEGRAL, va permettre une étude approfondie des sites de formation d'étoiles dans notre Galaxie.

Chapter 1

Introduction

1.1 Nuclear Astrophysics and γ-ray Astronomy

If, on a starry night, one turns ones gaze towards the firmament, it is unavoidable that a feeling of rest and eternity comes over the mind. In a life, which is marked by transformations, birth and death, youth and age, the stars form a definite and absolute basis, which seems invariable and endless. For a long time it was a prevailing cultural belief that the heavens represented a perfect and therefore unchangeable creation by God. The only apparent change in the sky, the movement of the planets, was described by perfectly cyclic motions, i.e., motions that have no start and no end. Change was supposed to be permitted only on the imperfect Earth, and when change did in fact occur in the sky the matter was dealt with by ensuring that no written record of it was made. It is therefore not surprising that the phenomenon of novae or supernovae came very lately to the attention of European astronomers, when in 1572 Tycho Brahe discovered a 'new' bright star, visible over 18 months in the constellation of Cassiopeia. With increasing observational effort, more and more life came into the sky: At the transition from the 16th to the 17th century, the first variable star, Mira (o Ceti), was discovered by David Fabricius. In 1718, Edmond Halley recognized that stars are not fixed on a sphere, as had been believed for a long time, but have their proper motions on the sky. Nowadays, a wealth of variable phenomena is observed in the sky, not only at optical wavelengths but throughout the entire electromagnetic spectrum. Among these are: pulsars, which were discovered first in 1967 by their radio emission, quasars, first observed at radio wavelengths in 1963, and the mysterious γ-ray bursts, which were accidentally discovered by satellite-borne detectors meant to monitor violations of the nuclear test ban treaty.

Today it is known that even the stars have only a finite lifetime, which however lasts millions to billions of years, much too long for aging to be directly perceptible to human beings. But physical changes due to the evolution of stars which may be as short as decades or centuries are readily observed (a noticeable example is P Cygni,

which appeared in the 17th century as a reddish, rather variable star, but today is a relatively constant yellowish-white star). In the 1920s the English astrophysicist Sir Arthur Stanley Eddington worked out the internal structure of the stars, and proposed that nuclear energy causes them to shine. In 1938, Hans Bethe and Carl Friedrich von Weizsäcker independently described two nuclear reaction chains (the pp chain and the CNO cycle) which can provide the nuclear energy budget of the stars by transforming hydrogen into helium. Today, the evolution of stars is understood as the story of their nuclear transmutations and the consequent effects on the stellar structure.

During the struggle to understand the energetics of stars in their various evolutionary stages, and to explain the origin and abundance of the chemical elements and their isotopes in the Universe, the field of nuclear astrophysics emerged in the 1940s from the marriage of the old science of astronomy with the young discipline of nuclear physics. Since the epoch-making paper of Burbidge *et al.* (1957), it has been established that the majority of the chemical elements are produced by nuclear reactions in the hot stellar interiors, either during the long, quiescent phases of stellar evolution, or in the violent supernova explosions that mark the death of massive stars. The material ejected from the dying stars is mixed in the interstellar medium from which new stellar generations are formed. The cycle then starts again, progressively enriching the Galaxy with heavy elements.

Stellar nucleosynthesis produces not only stable, but also unstable nuclei with lifetimes ranging from seconds to billions of years. Their radioactive transitions may give rise to γ-ray photons which, under certain conditions, may be detected by sufficiently sensitive instruments. This possibility, first suggested by Clayton and Craddock (1965), would unambiguously allow the detection of isotopes through their characteristic γ-ray signature. Moreover, detection of species with lifetimes considerably shorter than the age of the Galaxy would clearly show that nucleosynthesis is still active today. This thesis deals with the observation of such a nuclear transition line, arising from the radioactive decay of ^{26}Al with a lifetime of $\sim 10^6$ yr.

1.2 Observation of the 1.809 MeV Line

The discovery of γ-ray emission arising from the decay of radioactive ^{26}Al in the interstellar medium (ISM) was a milestone in γ-ray astronomy; it was the first detection of a cosmic radioactive isotope by means of its γ-ray radiation. Using the high-purity germanium (Ge) detectors aboard the HEAO-3 satellite, Mahoney *et al.* (1984) observed a γ-ray line at 1808.49 ± 0.41 keV from the general direction of the galactic center. Due to the high energy resolution of the Ge detectors (3.5 keV FWHM), the observed line could be uniquely identified with the 1808.65 ± 0.07 keV line arising from the radioactive decay of ^{26}Al at rest. The observed intrinsic line width (FWHM) of ≤ 3.0 keV (1σ) was consistent with the Doppler broadening expected from differential galactic rotation and random motions of the ISM. From the observed intensity of $(4.8 \pm 1.0)\,10^{-4}$

ph $cm^{-2}s^{-1}rad^{-1}$ from the vicinity of the galactic center, a total galactic ^{26}Al mass of 3 $M_\odot$ was inferred.

One year later, the γ-ray spectrometer aboard the Solar Maximum Mission (SMM) satellite provided convincing proof that the HEAO-3 discovery was truly the signature of galactic ^{26}Al production (Share *et al.* 1985). Following this promising discovery, a number of balloon borne instruments were set up, aiming to improve the spatial information on galactic 1.809 MeV emissivity (von Ballmoos *et al.* 1987, MacCallum *et al.* 1987, Malet *et al.* 1991, Teegarden *et al.* 1991). Nonetheless, the poor angular resolution of the experiments could only weakly constrain possible distribution models. While all measurements generally agreed about an enhancement of 1.8 MeV emission towards the galactic center, re-analysis of the SMM 1.8 MeV measurements by using the Earth as an occulting disk ruled out a single point source at the galactic center at the 4.7 σ confidence level (Purcell 1989). Additionally, the marginal detection (2 σ) of 1.8 MeV emission from the galactic plane at longitudes $l = 335°$ by GRIS was considered as indication of extended ^{26}Al emission along the galactic plane (Teegarden *et al.* 1991).

The HEAO-3 discovery of 1.8 MeV emission hardly came as a surprise. Indeed, calculations in the late 60s of explosive carbon burning in supernovae (Arnett 1969) had shown that substantial amounts of ^{26}Al may be produced in this environment. On the basis of these estimates it was suggested that the 1.8 MeV line of ^{26}Al would be an interesting target for γ-ray line astronomy (Ramaty and Lingenfelter 1977, Arnett 1977). Due to its relative long lifetime of $\tau_{26} = 1.040\,10^6$ years (Fuller *et al.* 1982), ^{26}Al from several thousand supernovae should accumulate in the Galaxy, giving rise to a diffuse emission in the galactic plane. The flux on Earth was estimated to $\sim 10^{-4}$ ph $cm^{-2}s^{-1}$ from the galactic center direction, and it was pointed out that, because of its long mean life, ^{26}Al would have enough time to be thermalized in the ISM; therefore, it would de-excite essentially at rest and emit a narrow line.

Recent measurements by the balloon borne γ-ray spectrometer GRIS, however, raised some doubt about the thermalization of ^{26}Al in the ISM. Naya *et al.* (1996) reported for a 32 hour balloon flight made from Alice Springs (Australia) the detection of a broadened 1.8 MeV line from the galactic center direction with an intrinsic width of 5.4 ± 1.4 keV. The detection of such a wide feature is suggestive of either high dispersion velocities (> 450 km s^{-1}) or high temperatures ($T > 2\,10^8$ K), for which, however, no standard scenario can account (e.g., Chen *et al.* 1997). Note, however, that the GRIS measurement is inconsistent with the HEAO-3 measurement at the 2σ confidence level.

Since April 5, 1991, the imaging Compton telescope COMPTEL aboard the Compton Gamma-Ray Observatory (CGRO) is in orbit, providing for the first time a complete survey of the γ-ray sky at energies between $0.75 - 30$ MeV. Its unprecedented angular resolution of 3.8° (FWHM) within a wide field of view of about 1 steradian allows a detailed mapping of the 1.8 MeV line throughout the entire sky. After more than 6 years of operation, the point source sensitivity at 1.8 MeV was pushed below

Fig. 1.1: Decay scheme of ^{26}Al from Lederer and Shirley (1978).

$1.5\,10^{-5}$ ph cm^{-2}s^{-1} (3σ) for the combined data, setting a new standard in γ-ray line astronomy. Besides the clear detection of the 1.8 MeV line, new γ-ray lines were discovered with COMPTEL. Iyudin *et al.* (1994) reported the observation of the 1.157 MeV γ-ray line arising from the radioactive decay of ^{44}Ti towards the Cas A supernova remnant, and Morris *et al.* (1995) found evidence for the two principal decay lines of ^{56}Co, at 847 keV and 1238 keV, from the type Ia supernova SN 1991 T. Additional, enhanced γ-ray emission between $3 - 7$ MeV was reported by Bloemen *et al.* (1994b) from the Orion molecular cloud complex, which was tentatively attributed to nuclear deexcitation lines at 4.44 and 6.13 MeV from accelerated ^{12}C and ^{16}O nuclei.

1.3 The Radioactive Isotope ^{26}Al

1.3.1 Nucleosynthesis of ^{26}Al

The ground state of ^{26}Al is unstable to positron emission (β^+ emission; 82%) or to electron capture (EC; 15%) with a mean life of $\tau_{26} = 1.040\,10^6$ years (Fuller *et al.* 1982). Its decay feeds the first excited state of ^{26}Mg at 1.8087 MeV, which de-excites within 0.49 ps by emitting a γ-ray photon of 1.8086 MeV. At 228 keV above the ground state, ^{26}Al has an *isomeric state* (^{26}Al^m) which can directly decay to the ground state of ^{26}Mg by an allowed transition, because both states have the same spin and parity (cf. Fig. 1.1). The half-life of ^{26}Al^m is only 6.36 s. If this state can be thermally excited (at $T \sim 2.6\,10^9$ K), the ^{26}Al decay rate changes by a factor of roughly 10^{12}.

An environment rich in Mg can be found in the carbon and neon shells of massive stars, where ^{26}Al can be produced hydrostatically ($T \sim 10^9$ K) or explosively ($T_p \sim (2-2.5)\,10^9$ K). The ^{25}Mg(p,γ)^{26}Al reaction proceeds rapidely at these high temperatures and leads to large amounts of ^{26}Al, despite the low abundance of protons generated in these sites by reactions such as ^{12}C(^{12}C,p)^{23}Na. In such environments, however, substantial amounts of neutrons may be produced through secondary reactions, and ^{26}Al may be destroyed by ^{26}Al(n,p) and ^{26}Al(n,α) reactions. If neutrons are not present in large quantities and if the timescale is sufficiently long ($> 10^3$ s, i.e. in shell C-burning conditions), ^{26}Al is mainly destroyed by β^+ decay (at such high temperatures the β^+ decay rate is enhanced due to the presence of the isomeric state ^{26}Al^m).

Fig. 1.2: *Principal reactions involved in the synthesis of ^{26}Al in astrophysical sites.* Solid arrows: *(p,γ) reactions;* wavy arrows: *β^+ decays;* dashed arrows *(p,α) reactions.*

To be observable in γ-rays, ^{26}Al has to be ejected in the ISM before destruction. This can easily be achieved in the case of an explosive site, such as a nova or a supernova. On the other hand, objects suffering extensive mass loss and sufficient internal mixing,

like Wolf-Rayet stars (WRs), Luminous Blue Variables (LBVs), or Asymptotic Giant Branch stars (AGBs), could also statisfy this condition. Obviously, the different regimes of temperature, density, timescale and initial composition in these astrophysical sites imply different modes for the production and destruction of ^{26}Al. In the subsequent sections the most important candidate sites that have been proposed up to now are presented.

1.3.2 Massive Stars

Hydrostatic core H burning in massive stars can lead to the production of significant amounts of ^{26}Al if the central temperature of the star is $T_c > (3.5 - 4)\,10^7$ K, i.e., for stars more massive than ~ 15 M$_\odot$ on the main sequence. As the convective stellar core gradually retreats, while the star still burns H on the main sequence, it leaves behind ^{26}Al that has been previously produced and mixed in the core. Whether or not it will appear at the stellar surface depends on internal mixing and the mass loss.

All massive stars suffer mass loss due to a stellar wind which is driven by radiation pressure in absorption lines. Among the stellar objects showing the largest mass loss rates are Wolf-Rayet (WR) stars which are characterized by prominent broad emission lines of nitrogen (WN stars) or carbon and oxygen (WC and WO stars) in their spectra. These lines are interpreted as freshly synthesized products of central hydrogen or helium burning arising at the stellar surface due to heavy mass loss and/or internal mixing (e.g. Langer 1987). Since ^{26}Al is produced during core hydrogen burning, it is expected to show up at the stellar surface together with nitrogen[1], i.e. in WN stars (Prantzos and Cassé 1986). During the subsequent phase of central helium burning ^{26}Al is no longer produced. Neutrons released through ^{13}C(α,n) and ^{22}Ne(α,n) efficiently destroy the remaining ^{26}Al in the stellar core through ^{26}Al(n,α) and ^{26}Al(n,p) reactions. However, ^{26}Al continues to be ejected from the stellar envelope, and disappears only when the He burning products appear at the surface, at their turn, in the WC phase.

Several calculations of ^{26}Al production in WR stars have been performed in the recent past (e.g. Prantzos and Cassé 1986, Langer *et al.* 1995, Meynet *et al.* 1997, and references therein). Despite differences in the stellar models and different physical ingredients (mass loss, treatment of convection, reaction rates) that were used, results agreeing within a factor of ~ 3 are found for stars of a given mass. The calculations show that stars in the mass range $40 < M/\mathrm{M}_\odot < 120$ and with solar metallicity eject a few $10^{-5} - 10^{-3}$ M$_\odot$ of ^{26}Al in the ISM, where higher stellar masses lead to higher ^{26}Al yields. An important feature of the production of ^{26}Al by WR stars is the dependence of the ^{26}Al yield on metallicity (Prantzos and Cassé 1986). Firstly, the production of ^{26}Al is proportional to the available amount of ^{25}Mg which depends on

[1] For massive stars, central hydrogen burning is mainly through the CNO cycle. Due to the slowness of the ^{14}N(p,γ)^{15}O reaction, all C and O is converted to N in hydrostatic equilibrium. Hence, overabundance of N over O and C is a clear sign of the activity of the CNO cycle.

the initial stellar metallicity. Secondly, the initial metallicity determines the mass loss rate on the main sequence since the radiative force on the envelope depends on the amount of metallic ions it contains. Thirdly, higher metallicity reduces, again via the mass loss, the minimum mass required for a star to turn into a WR star. The first two dependencies result in a power law relation $M_{26} \propto (Z/Z_\odot)^n$ (where $n \sim 2-2.2$) between the ^{26}Al yields and the initial metallicity Z (Prantzos and Diehl 1996, Meynet *et al.* 1997). The present day galactic ^{26}Al mass due to WR stars is hence rather sensitive to the galactic metallicity gradient, the precise value of which is still uncertain. Recent estimates amount from $M_{26} = 0.4$ M$_\odot$ (Prantzos and Diehl 1996) to $M_{26} = 0.9 \pm 0.5$ M$_\odot$ (Meynet *et al.* 1997).

The influence of a close companion on the ^{26}Al yields of WR stars has recently been investigated by Braun and Langer (1994). Their calculations show that for massive WR stars the ^{26}Al yield of a binary system may possibly be smaller than that for a single star since ^{26}Al from the primary is accreted onto the secondary. For low mass WR stars ($M \lesssim 40$ M$_\odot$), binarity has the potential to increase the ^{26}Al yield since these stars loose much more mass in an interacting binary system than as single stars. However, the uncertainties in the fraction of WR binaries and in the fraction of the ejected material that leaves the system are too large for any meaningful conclusions.

Evidence of CNO-cycled material at the stellar surface, however, is not only restricted to WR stars. Although most massive main sequence stars have normal He and N abundances (Herrero *et al.* 1992, Maeder 1994, and references therein), there are many exceptions for O and B stars. Among these are O giants and supergiants, in particular when they are fast rotators (Herrero *et al.* 1992, Pauldrach *et al.* 1994), OBN stars, among which are many short-period binaries (Schönberner *et al.* 1988), and B supergiants (Gies and Lambert 1992, Lennon 1994). Maeder (1987) suggested that there may be a bifurcation in stellar evolution: while most stars follow the tracks of inhomogenous evolution, a fraction of about 15%, mainly composed of fast rotators and binaries, may evolve homogenously and become ON blue stragglers. ^{26}Al yield calculations for such homogenously evolving stars have been made by Walter and Maeder (1989) who find that a blue straggler produces at least two times more ^{26}Al than a normal massive star. Recently, Langer *et al.* (1996) presented first ^{26}Al yield calculations for rotating massive stars between $8 - 20$ M$_\odot$, which suggest an ^{26}Al enhancement of factors $3 - 10$ with respect to non-rotating stars. If these results can be confirmed, ^{26}Al production by massive rotating stars could contribute significantly to the total galactic ^{26}Al budget.

1.3.3 Asymptotic Giant Branch Stars

Stars of intermediate or low mass (≤ 8 M$_\odot$) spend the last part of their active life on the Asymptotic Giant Branch (AGB) of the Hertzsprung-Russel diagram. At that phase of evolution a star has terminated He burning in its core and consists of a degenerated and inert carbon-oxygen core, surrounded by a He-burning shell, a H-burning shell, and

a H-envelope. The envelope is completely convective, but the depth of its penetration inside the star is poorly known. The double-shell burning phase is unstable, leading to the development of thermal pulses[2], which progressively expell the stellar envelope on a timescale of a few $10^5 - 10^6$ years (depending on the stellar mass), forming a planetary nebula.

^{26}Al is thought to be produced in two regions of such stars. Firstly, ^{26}Al may be produced in the H-burning shell at temperatures $T \sim 7\,10^7$ K. While in stars of initial masses $M < 5$ M$_\odot$ ^{26}Al is produced by proton capture on pre-existing ^{25}Mg, for more massive stars the H-burning shell temperature is high enough to enhance the ^{26}Al abundance via the ^{24}Mg(p,γ)^{25}Al(β^+)^{25}Mg(p,γ)^{26}Al chain (e.g. Forestini and Charbonnel 1997). However, as the mean H-burning shell temperature increases with time, ^{26}Al begins to be partially destroyed by ^{26}Al(p,γ) reactions late along the AGB evolution. Periodically after the He-flash, when H-shell burning is temporarily extinct, the convective envelope descends down to the H-shell and brings up the H-burning products of that shell, and thus ^{26}Al, to the surface (3$^{\rm rd}$ dredge-up).

Secondly, if the convective envelope reaches the H-burning shell, the temperature at the base becomes sufficiently high for the onset of nuclear burning. In stellar evolution calculations, this hot-bottom burning (HBB) generally appears (for solar metallicity) for stars with initial masses $\gtrsim 5$ M$_\odot$ (e.g. Vassiliadis and Wood 1993, Bloecker 1995, Forestini and Charbonnel 1997). The existence of HBB is supported by observations of enhanced ^{7}Li abundances in AGB stars (Smith and Lambert 1989). Hence ^{26}Al might be directly produced inside the convective envelope by proton capture on either pre-existing ^{25}Mg, or for sufficiently high temperatures which occur at the end of the AGB phase for massive stars, on ^{25}Mg supplied by ^{24}Mg(p,γ) reactions (e.g. Forestini and Charbonnel 1997). Since the mean turn-over time of the convective envelope amounts to only 0.5 yr, ^{26}Al appears almost instantaneously at the surface of the star from which it is expelled into the ISM by stellar winds.

Products of the He-burning shell may also be mixed in the upper layers after a thermal pulse and finally to the surface, after the next dredge-up episode. For sufficiently high temperatures in the base of the He-shell ($T > 2.5\,10^8$ K) ^{25}Mg may be produced by the ^{22}Ne(α,n)^{25}Mg reaction which may be mixed into regions of ^{26}Al production, considerably enhancing the ^{26}Al yield. Observations, however, do not support this idea,

[2]If, after central He exhaustion, a star reaches the AGB, the efficiency of the He-burning shell is drastically increased due to the contraction of the CO-core. This overheating leads to an expansion of the overlaying material which in turn leads to a decrease of the temperature in the H-burning shell and consequently results in extinction of H-shell burning. As the fuel of the He-burning shell is progressively exhausted, the He-burning shell slowly migrates outwards, until it almost reaches the He-H interface. This rises the temperature at the position of the former H-burning shell which results in re-ignition of H-shell burning – the He-burning shell dies down temporarily. In its turn, the H-burning shell migrates outwards, laying down its He-ashes as new fuel for He-burning. When the He-layer becomes sufficiently thick, the He-burning shell re-ignites violently which agains leads to an expansion of the overlaying material and an extinction of the H-burning shell. The cycle of the thermal pulses starts again.

First ^{26}Al yield predictions based on full evolutionary models were recently presented by Forestini and Charbonnel (1997). Due to the complex production mechanism involved, the ^{26}Al yield is not a monotonic function of the stellar mass. Nevertheless, a general trend exists for initially more massive stars to produce higher ^{26}Al yields. In particular, their most massive models show a substantial net destruction of ^{24}Mg together with a net production of ^{25}Mg in the H-burning shell, which in turn leads to an enhancement of the ^{26}Al yield due to the presence of additional seed nuclei. Note, however, that this leads to substantial anomalies in the ^{24}Mg:^{25}Mg:^{26}Mg isotopic abundance ratios, which have not been observed so far (see Lambert 1989, Barbuy *et al.* 1992, and references therein). Forestini and Charbonnel (1997) calculated models for both solar metallicity and $Z = 0.25$ Z$_\odot$. While their 3 M$_\odot$ model roughly scales with Z (which indicates that ^{26}Al was only produced by proton capture on pre-existing ^{25}Mg), the situation becomes more complicated for the more massive models. Indeed, their low metallicity 5 M$_\odot$ model produces roughly ~ 20 times as much ^{26}Al as the solar metallicity model.

Fig. 1.3: *Final stellar mass as function of the initial mass for stars of solar metallicity (from Langer 1995).*

1.3.4 Core Collapse Supernovae

Many stars loose a considerable amount of their initial mass during their evolution by stellar winds (cf. Fig. 1.3). While stars initially less massive than ~ 8 M$_\odot$ terminate their lives as white dwarfs, stars with initial masses above this limit may explode as core collapse supernovae.[3] Because of their important mass loss, the advanced

[3]Core collapse supernovae, also refered to as hydrodynamic supernovae, arise from stars which gradually exhibit all possible phases of central burning (H, He, C, Ne, O, and Si). They end up with a central core made of Fe and Ni, the most stable elements, for which no exothermal nuclear processing is possible. Consequently, the core collapses into its own gravitational potential. Another type of supernova, called 'thermonuclear supernova', arises from the accretion of mass on a white dwarf near

evolution of WR stars (initial mass $\gtrsim 30$ $M_\odot$) is rather uncertain. Those stars may almost completely evaporate during their evolution, leaving a presupernova star of only a few $M_\odot$, which probably explodes as a type Ib/c supernova (e.g., Woosley *et al.* 1993, and references therein).

The evolution of less massive stars in the mass range $8 < M/M_\odot < 25$ is not much effected by mass loss, at least for initial metallicities $Z \leq Z_\odot$, and are expected to explode as type II supernovae. Such stars develop shell H-burning at temperatures $T \sim (7-9)\,10^7$ K leading to a moderate production of ^{26}Al at the base of the shell. Much more important seems to be the production of ^{26}Al in the carbon and oxygen-neon[4] shells, just prior to the supernova explosions. In these shells, ^{26}Al is produced by proton capture on ^{25}Mg at temperatures $T \sim 1.5\,10^9$ K on time scales of a few years. Neutrons, essentially provided by ^{13}C(α,n) and ^{22}Ne(α,n) reactions, are the main destruction agent of ^{26}Al, along with β^+ decay (the β^+ decay rate is substantially enhanced at these high temperatures due to the population of the isomeric state of ^{26}Al). Convection in these shells is of paramount importance since it can simultaneously bring light reactants and seed nuclei into the hot zone to aid in the synthesis, and then remove the fragile product from the high temperature region where it might otherwise be destroyed. Consequently, major uncertainties in the ^{26}Al yield estimates arise from the treatment of time-dependent convection, for which no satisfactory theory is available so far.

When the stellar core collapses, the outgoing shock wave heats the inner zones of the C and O-Ne shells to peak temperatures $T_p \sim (2-3)\,10^9$ K, sufficiently high to ignite explosive nucleosynthesis. Explosive nucleosynthesis of ^{26}Al in the carbon and neon shell have been studied in numerous works which find considerable amounts produced in the O-Ne layer, and smaller amounts in the layers of explosive C-burning. The explosive ^{26}Al production may be further enhanced by $\sim 30\% - 50\%$ due to neutrino-induced nucleosynthesis, also refered to as the ν-process (Woosley *et al.* 1990). In this process, additional protons are liberated by ν-spallation of the abundant preexplosive isotopes (^{20}Ne, ^{16}O, ^{23}Na, ^{24}Mg), which are captured on ^{25}Mg to produce ^{26}Al. The major uncertainty of this mechanism arises from the unknown neutrino spectrum emitted from the core collapse.

^{26}Al production in type II supernovae has been estimated by various groups (e.g. Truran and Cameron 1978, Weaver and Woosley 1993, Thielemann *et al.* 1994, Woosley and Weaver 1995, Thielemann *et al.* 1996). They find that the ^{26}Al yield is not a monotonic function of the stellar mass, depending in a complictated way on the stellar physics. Nevertheless, a general trend exists of increasing ^{26}Al yield with increasing initial stellar mass. Secondly, the ^{26}Al yield depends critically on the treatment of semi-convection and may vary by factors of $3-12$ (Prantzos and Diehl 1996). Finally,

---

the Chandrasekhar limit (type Ia supernova).

[4]The oxygen and neon burning shells frequently combine into a single burning shell during the late stages of stellar evolution (e.g. Timmes *et al.* 1995).

the ^{26}Al production may also depend on the initial metallicity of the progenitor star: Using parametrized calculations Woosley and Weaver (1980) showed that an increase of metallicity from $Z = 1$ $Z_\odot$ to 2.5 $Z_\odot$ leads to a decrease in the ^{26}Al yield of nearly a factor of 2. This can be understood due to an increased neutron flux in the layers of ^{26}Al synthesis, provided by reactions on ^{22}Ne, whose abundance is directly proportional to Z. More recent calculations of Woosley and Weaver (1995) for low-metallicity stars reveal however a complex dependence of ^{26}Al production on metallicity with either decreasing, increasing or even fluctuating relationsships, depending on the initial mass of the progenitor.

Although no ^{26}Al production calculations based on full evolutionary models exist so far for the type Ib/c supernova scenario, the ^{26}Al yield for these objects may be estimated from recent calculations for mass loosing helium stars by Woosley *et al.* (1995). They identified these objects as Wolf-Rayet stars of initial mass $\gtrsim 30$ $M_\odot$, corresponding to helium cores with $M \gtrsim 10$ $M_\odot$. For such cores, Woosley *et al.* (1995) derived ^{26}Al yields between $(4-6)\,10^{-5}$ $M_\odot$. This yield should not depend drastically on the initial mass of the progenitor star since mass dependent mass-loss leads to convergent presupernova masses around ~ 4 $M_\odot$ (cf. Fig. 1.3).

1.3.5 Novae

The outburst of novae are suggested to arise from the accretion of a critical mass of hydrogen-rich material from a companion star filling its Roche lobe on the surface of a white dwarf (WD), leading to a thermonuclear runaway. For the physical conditions that are expected in this environment (temperature $T \sim (2-4)\,10^8$ K, density $\rho \sim 2\,10^4$ g cm^{-3}), thermonuclear burning will proceed by means of hydrogen burning on carbon, nitrogen, and oxygen (CNO) or heavier nuclei present in the shell matter.

Clayton and Hoyle (1974) were the first who suggested that substantial amounts of ^{26}Al may be produced in the explosive hydrogen burning which occurs during nova outbursts. Studies of nucleosynthesis in such conditions show that, depending on the characteristic timescale of the explosion (a few 100 s), not only ^{25}Mg but also the more abundant ^{24}Mg can be transformed into ^{26}Al, by ^{24}Mg(p,γ)^{25}Al(β^+)^{25}Mg(p,γ)^{26}Al, provided there is enough time for the β^+ decay of ^{25}Al to take place ($\tau_{25} \sim 10$ s). While nova events with heavy element composition of solar metallicity lead only to marginal ^{26}Al production, enrichment of the envelope by elements in the range from neon to aluminium greatly increases ^{26}Al yields (Weiss and Truran 1990, Nofar *et al.* 1991, and references therein).

The presence of significant enrichments of CNO or intermediate-mass elements in the ejecta of most classical novae suggests that part of the WD surface is mixed into the accreted envelope material. Recently, considerable attention was directed towards so-called ONeMg-rich novae which show strong emission lines of heavy elements, in particular Ne, in their spectra, and possibly form $10\% - 20\%$ of all galactic nova events

(Livio and Truran 1994, and references therein). The enhancement of Mg in the H-burning envelope then provides an abundant reservoir of seed nuclei for ^{26}Al production.

Two possible scenarios which may give rise to ONeMg-rich novae are currently discussed in the literature. On the one hand, they may arise from thermonuclear explosions on ONeMg WDs (Law and Ritter 1983, Truran and Livio 1986), which possibly evolved from stars of initial masses between $8 < M/\mathrm{M}_\odot < 12$ $\mathrm{M}_\odot$ (e.g. Nomoto 1984). For this scenario, ^{26}Al yields have been estimated recently using 1-dimensional hydrodynamic stellar evolution codes coupled to nuclear reaction networks (Politano *et al.* 1995, José *et al.* 1997). The results depend critically on many, often poorly known, parameters of the models, like the nuclear reaction rates (e.g. Coc *et al.* 1995), the underlying WD composition (e.g. José *et al.* 1997), the treatment of convection, the mass accretion rate, or the amount of enrichment of the nova envelope with heavy elements (e.g. Kolb and Politano 1997). Consequently, yield estimates for different calculations vary by about one magnitude. Additionally, all models fail to accout for the large ejected masses ($10^{-4} - 10^{-3}$ $\mathrm{M}_\odot$) which are observed for ONeMg-rich novae.

This led Shara and Prialnik (1994) to propose a new scenario for ONeMg-rich novae. They suggested that low-mass CO white dwarfs may acquire a relatively thick outer layer of ONeMg-rich material by rapid accretion ($\sim 10^{-6}$ $\mathrm{M}_\odot$ yr^{-1}) of hydrogen-rich matter from the companion, that burns quietly (in thermal equilibrium) into helium, which in turn is burnt explosively into heavier elements in a series of mild flashes. Recently, Prialink and Shara (1995) calculated the first self-consistent ONeMg-rich nova model based on this scenario. Surprisingly, despite the differences between the scenarios, the ejecta composition obtained is rather similar to the ONeMg WD models of Politano *et al.* (1995). In contrast, the ejected masses are about one magnitude larger than for ONeMg WD models, much closer to the observations. José *et al.* (1997) stress, however, that this scenario requires a fine tuning of some parameters relative to the companion, such as the mass accretion rate.

1.3.6 ^{26}Al Yield Estimates

In Fig. 1.4, ^{26}Al yield estimates for the discussed candidate sources are compiled as a function of initial stellar mass. AGB yields were taken from Bazan *et al.* (1993) and Forestini and Charbonnel (1997), where only the latter yields are the result of self-consistent evolutionary models. Bazan *et al.* (1993) derived their estimates using analytic expressions of stellar envelope models assuming that all ^{26}Al is produced during hot-bottom burning. The comparison of both works illustrates the tremendous uncertainties involved in the prediction of ^{26}Al production from AGB stars.

ONeMg nova yields were taken from Prialink and Shara (1995), José *et al.* (1997), and Kolb and Politano (1997). Since novae may undergo numerous outbursts during the lifetime of ^{26}Al, the quoted values are total ^{26}Al yields during 10^6 yr, as derived from $\dot{M}_{26} = \alpha \tau_{26} X_{26} \dot{M}_{acc}$, where α is the ratio of the ejected mass to the accreted

Fig. 1.4: *^{26}Al yield estimates for AGB stars (circles), ONeMg-rich novae (squares), type II supernovae (diamonds), Wolf-Rayet stars (triangles), and type Ib/c supernovae (dashed line).*

mass, τ_{26} is the lifetime of ^{26}Al, X_{26} is the ^{26}Al abundance in the ejecta as predicted by the model, and $\dot{M}_{acc}$ is the mass accretion rate on the WD. For consistency among the different calculations, ^{26}Al yields for 1.25 $M_\odot$ WDs are given, using $\alpha = 1.2$ (as suggested by Kolb and Politano 1997) and $\dot{M}_{acc} = 2\,10^{-10}$ $M_\odot yr^{-1}$.

Type II supernova yields were taken from Woosley and Weaver (1995) and Thielemann *et al.* (1996). Both calculations differ in numerous details like the treatment of convection, nuclear reaction rates, or the explosion mechanism (see Woosley and Weaver 1995). In particular, Woosley and Weaver (1995) followed the persupernova evolution by coupling a large reaction network with time dependent convection, while Thielemann *et al.* (1996) calculated only explosive ^{26}Al production. For type Ib/c supernovae, a mass independent ^{26}Al yield of $6\,10^{-5}$ $M_\odot$ was assumed as suggested by the helium star models of Woosley *et al.* (1995).

For WR stars, ^{26}Al yields were taken from two recent studies of Langer *et al.* (1995) and Meynet *et al.* (1997). Note that except for the lowest masses, both estimates are in fairly good agreement, despite the different physics involved in the calculations (mass loss and treatment of convection). Meynet *et al.* (1997) suggest that this could be due to a subtle balance between the effect of a change in mass loss and the impact of a different mixing scheme. The uncertainties of ^{26}Al production in WR stars may in reality be larger as those suggested by comparison of both works.

The total amount of ^{26}Al produced by each candidate source per ^{26}Al lifetime ($\tau_{26} \sim$

10^6 yr) can be estimated by the expression

$$\dot{M}_{26} = \mathrm{SFR} \int_{M_1}^{M_2} m_{26}(M)\xi(M)dM, \tag{1.1}$$

where SFR is the galactic star formation rate, given here in stars Myr^{-1}, $m_{26}(M)$ is the ^{26}Al yield of one object of a source class (in $\mathrm{M}_\odot$ star^{-1}), which generally is a function of the initial stellar mass M, $\xi(M)$ is the initial mass function (IMF), which specifies the fraction of stars formed with initial stellar mass within the interval $[M, M + dM]$, and M_1 and M_2 are the lower and upper initial stellar mass limits for the considered source class. $\xi(M)$ is normalized here to one star in the mass interval $M_{low} = 0.08$ $\mathrm{M}_\odot$ and $M_{upp} = 100.0$ $\mathrm{M}_\odot$:

$$\int_{M_{low}}^{M_{upp}} \xi(M)dM = 1. \tag{1.2}$$

For the mass spectrum, a power-law IMF

$$\xi(M) \propto M^{\Gamma-1} \tag{1.3}$$

is applied for which slope estimates range from $\Gamma = -1.1$ to $\Gamma = -2.4$ (e.g. Salpeter 1955, Humphreys and McElroy 1984, Kroupa *et al.* 1993, Massey *et al.* 1995). The SFR will be determined by calibration of the IMF on the observed Lyman continuum ($\lambda < 912$ Å) photon production rate Q (e.g. Mezger 1988), using the relation

$$\mathrm{SFR} = \frac{Q}{\int_{M_{low}}^{M_{upp}} \xi(M)Q_0(M)t_{\mathrm{LT}}(M)dM}, \tag{1.4}$$

where $Q_0(M)$ is the number of Lyman continuum (Lyc) photons produced per second by an individual star of initial mass M, and $t_{\mathrm{LT}}(M)$ is the total lifetime of such a star. For $Q_0(M)$ the recent calibration of Vacca *et al.* (1996) was used, lifetimes were taken from the evolutionary calculations of Schaller *et al.* (1992). The galactic Lyc luminosity is estimated to $Q = 3.0\,10^{53}$ ph s^{-1} on basis of radio observations of galactic H II regions (Mezger 1978), and has recently been inferred from FIRAS observations of the N II cooling line at 205 μm to $Q = 3.5\,10^{53}$ ph s^{-1} (Bennett *et al.* 1994b). In this work, the latter value is used for which Bennett *et al.* (1994b) estimate an uncertainty of $\sim 50\%$.

Resulting ^{26}Al yield estimates $\dot{M}_{26}$ are shown for AGB stars, type II and type Ib/c supernovae, and Wolf-Rayet stars in Tabs. 1.1 and 1.2 for two different sets of nucleosynthesis model calculations and for solar metallicity. The tables differ in the assumed mass limit $\mathrm{M_{WR}}$ above which stars exhibit a WR phase and subsequently explode as type Ib/c supernova ($\mathrm{M_{WR}} = 40$ $\mathrm{M}_\odot$ for Tab. 1.1 and $\mathrm{M_{WR}} = 25$ $\mathrm{M}_\odot$ for Tab. 1.2). Stars within the initial mass interval of $8 < M/\mathrm{M}_\odot < \mathrm{M_{WR}}$ were assumed to explode as type II supernovae, while stars within $1 < M/\mathrm{M}_\odot < 8$ were assumed to eject ^{26}Al during the thermal-pulse AGB phase.

$\Gamma = -1.1$, $Z = 1$ $Z_\odot$							
Object	$M_1 - M_2$	w	f	$\langle m_{26}^A \rangle$	$\dot{M}_{26}^A$	$\langle m_{26}^B \rangle$	$\dot{M}_{26}^B$
AGB	$1 - 8$	$5.6\,10^{-2}$	$1.3\,10^5$	$4.4\,10^{-7}$	0.06	$2.8\,10^{-5}$	3.70
SN II	$8 - 40$	$5.2\,10^{-3}$	$1.2\,10^4$	$6.6\,10^{-5}$	0.81	$3.9\,10^{-6}$	0.05
SN Ib/c	$40 - 100$	$6.8\,10^{-4}$	$1.6\,10^3$	$6.0\,10^{-5}$	0.10		
WR	$40 - 100$	$6.8\,10^{-4}$	$1.6\,10^3$	$2.2\,10^{-4}$	0.36	$2.4\,10^{-4}$	0.39
Total	$0.08 - 100$	1.0	$2.3\,10^6$		1.33		4.14 (0.44)
$\Gamma = -1.35$, $Z = 1$ $Z_\odot$							
Object	$M_1 - M_2$	w	f	$\langle m_{26}^A \rangle$	$\dot{M}_{26}^A$	$\langle m_{26}^B \rangle$	$\dot{M}_{26}^B$
AGB	$1 - 8$	$3.1\,10^{-2}$	$2.9\,10^5$	$3.4\,10^{-7}$	0.10	$2.1\,10^{-5}$	6.05
SN II	$8 - 40$	$1.8\,10^{-3}$	$1.7\,10^4$	$5.8\,10^{-5}$	0.97	$3.5\,10^{-6}$	0.06
SN Ib/c	$40 - 100$	$1.6\,10^{-4}$	$1.5\,10^3$	$6.0\,10^{-5}$	0.09		
WR	$40 - 100$	$1.6\,10^{-4}$	$1.5\,10^3$	$2.1\,10^{-4}$	0.32	$2.4\,10^{-4}$	0.36
Total	$0.08 - 100$	1.0	$9.5\,10^6$		1.48		6.47 (0.42)
$\Gamma = -1.7$, $Z = 1$ $Z_\odot$							
Object	$M_1 - M_2$	w	f	$\langle m_{26}^A \rangle$	$\dot{M}_{26}^A$	$\langle m_{26}^B \rangle$	$\dot{M}_{26}^B$
AGB	$1 - 8$	$1.3\,10^{-2}$	$9.2\,10^5$	$2.4\,10^{-7}$	0.22	$1.3\,10^{-5}$	11.77
SN II	$8 - 40$	$3.7\,10^{-4}$	$2.6\,10^4$	$4.8\,10^{-5}$	1.25	$2.9\,10^{-6}$	0.07
SN Ib/c	$40 - 100$	$2.0\,10^{-5}$	$1.4\,10^3$	$6.0\,10^{-5}$	0.08		
WR	$40 - 100$	$2.0\,10^{-5}$	$1.4\,10^3$	$2.0\,10^{-4}$	0.28	$2.2\,10^{-4}$	0.31
Total	$0.08 - 100$	1.0	$6.9\,10^7$		1.83		12.15 (0.38)
$\Gamma = -2.4$, $Z = 1$ $Z_\odot$							
Object	$M_1 - M_2$	w	f	$\langle m_{26}^A \rangle$	$\dot{M}_{26}^A$	$\langle m_{26}^B \rangle$	$\dot{M}_{26}^B$
AGB	$1 - 8$	$2.3\,10^{-3}$	$8.7\,10^6$	$1.4\,10^{-7}$	1.20	$4.7\,10^{-6}$	40.78
SN II	$8 - 40$	$1.6\,10^{-5}$	$5.8\,10^4$	$3.5\,10^{-5}$	2.03	$2.0\,10^{-6}$	0.12
SN Ib/c	$40 - 100$	$3.0\,10^{-7}$	$1.1\,10^3$	$6.0\,10^{-5}$	0.07		
WR	$40 - 100$	$3.0\,10^{-7}$	$1.1\,10^3$	$1.7\,10^{-4}$	0.19	$2.0\,10^{-4}$	0.22
Total	$0.08 - 100$	1.0	$3.8\,10^9$		3.49		41.12 (0.34)

Table 1.1: *Estimated ^{26}Al production rates for solar metallicity using IMF slopes of $\Gamma = -1.1$, -1.35, -1.7, and -2.4. For production rates A, the AGB models of Forestini and Charbonnel (1997), the SN II models of Woosley and Weaver (1995), SN Ib/c estimates of Woosley* et al. *(1995), and the WR models of Meynet* et al. *(1997) were used. Production rates B were calculated using AGB models of Bazan* et al. *(1993), SN II models of Thielemann* et al. *(1996), and WR models of Langer* et al. *(1995). The last line of each table summarizes the total production rate per Myr of stars in the Galaxy (4th column) and the expected ^{26}Al production rate for models A and B (6th and 8th column, respectively). Total ^{26}Al production rates for models B in parentheses exclude the unreasonable high contribution from AGBs. For WR stars and type Ib/c supernovae, a lower mass limit of $M_{\rm WR} = 40$ $M_\odot$ was assumed.*

$\Gamma = -1.1$, $Z = 1$ $Z_\odot$							
Object	$M_1 - M_2$	w	f	$\langle m_{26}^A \rangle$	$\dot{M}_{26}^A$	$\langle m_{26}^B \rangle$	$\dot{M}_{26}^B$
AGB	$1 - 8$	$5.6\,10^{-2}$	$1.3\,10^5$	$4.4\,10^{-7}$	0.06	$2.8\,10^{-5}$	3.70
SN II	$8 - 25$	$4.5\,10^{-3}$	$1.1\,10^4$	$3.3\,10^{-5}$	0.35	$2.4\,10^{-6}$	0.03
SN Ib/c	$25 - 100$	$1.4\,10^{-3}$	$3.3\,10^3$	$6.0\,10^{-5}$	0.20		
WR	$25 - 100$	$1.4\,10^{-3}$	$3.3\,10^3$	$1.2\,10^{-4}$	0.40	$1.3\,10^{-4}$	0.44
Total	$0.08 - 100$	1.0	$2.3\,10^6$		1.01		4.17 (0.47)
$\Gamma = -1.35$, $Z = 1$ $Z_\odot$							
Object	$M_1 - M_2$	w	f	$\langle m_{26}^A \rangle$	$\dot{M}_{26}^A$	$\langle m_{26}^B \rangle$	$\dot{M}_{26}^B$
AGB	$1 - 8$	$3.1\,10^{-2}$	$2.9\,10^5$	$3.4\,10^{-7}$	0.10	$2.1\,10^{-5}$	6.05
SN II	$8 - 25$	$1.6\,10^{-3}$	$1.5\,10^4$	$3.1\,10^{-5}$	0.46	$2.2\,10^{-6}$	0.03
SN Ib/c	$25 - 100$	$3.6\,10^{-4}$	$3.4\,10^3$	$6.0\,10^{-5}$	0.21		
WR	$25 - 100$	$3.6\,10^{-4}$	$3.4\,10^3$	$1.1\,10^{-4}$	0.37	$1.2\,10^{-4}$	0.41
Total	$0.08 - 100$	1.0	$9.5\,10^6$		1.14		6.49 (0.44)
$\Gamma = -1.7$, $Z = 1$ $Z_\odot$							
Object	$M_1 - M_2$	w	f	$\langle m_{26}^A \rangle$	$\dot{M}_{26}^A$	$\langle m_{26}^B \rangle$	$\dot{M}_{26}^B$
AGB	$1 - 8$	$1.3\,10^{-2}$	$9.2\,10^5$	$2.4\,10^{-7}$	0.22	$1.3\,10^{-5}$	11.77
SN II	$8 - 25$	$3.4\,10^{-4}$	$2.4\,10^4$	$2.9\,10^{-5}$	0.67	$1.9\,10^{-6}$	0.05
SN Ib/c	$25 - 100$	$5.2\,10^{-5}$	$3.6\,10^3$	$6.0\,10^{-5}$	0.22		
WR	$25 - 100$	$5.2\,10^{-5}$	$3.6\,10^3$	$9.1\,10^{-5}$	0.33	$1.0\,10^{-4}$	0.37
Total	$0.08 - 100$	1.0	$6.9\,10^7$		1.44		12.19 (0.42)
$\Gamma = -2.4$, $Z = 1$ $Z_\odot$							
Object	$M_1 - M_2$	w	f	$\langle m_{26}^A \rangle$	$\dot{M}_{26}^A$	$\langle m_{26}^B \rangle$	$\dot{M}_{26}^B$
AGB	$1 - 8$	$2.3\,10^{-3}$	$8.7\,10^6$	$1.4\,10^{-7}$	1.20	$4.7\,10^{-6}$	40.78
SN II	$8 - 25$	$1.5\,10^{-5}$	$5.6\,10^4$	$2.4\,10^{-5}$	1.36	$1.5\,10^{-6}$	0.08
SN Ib/c	$25 - 100$	$9.9\,10^{-7}$	$3.7\,10^3$	$6.0\,10^{-5}$	0.22		
WR	$25 - 100$	$9.9\,10^{-7}$	$3.7\,10^3$	$6.7\,10^{-5}$	0.25	$7.6\,10^{-5}$	0.28
Total	$0.08 - 100$	1.0	$3.8\,10^9$		3.03		41.14 (0.36)

Table 1.2: *Similar to Tab. 1.1 but for a lower mass limit for type Ib/c supernovae and WR stars of* $M_{\rm WR} = 25$ $M_\odot$.

The influence of the IMF slope on the resulting ^{26}Al production rates is illustrated by variation of Γ between the extreme boundaries -1.1 and -2.4. Together with the production rates $\dot{M}_{26}$, the number fraction $w = \int_{M_1}^{M_2} \xi(M)dM$, the source frequency $f = \mathrm{SFR} \cdot w$ (in stars per Myr), and the mean ^{26}Al yield

$$\langle m_{26} \rangle = \frac{\int_{M_1}^{M_2} m_{26}(M)\xi(M)dM}{\int_{M_1}^{M_2} \xi(M)dM} \tag{1.5}$$

of each source class is given. Since the star formation rate is calibrated on the Lyc luminosity, which in turn arises mostly from stars initially more massive than $\gtrsim 20$ M$_\odot$, the WR yield is the most stable production rate under variation of Γ. With steepening of the IMF slope, the contribution from less massive stars rises, leading to an increase of the total galactic ^{26}Al production rate. Notice that the AGB stars of Bazan *et al.* (1993) always contribute an unreasonably high amount of ^{26}Al (recall that observations of the 1.8 MeV γ-ray line suggest a total galactic ^{26}Al mass around ~ 3 M$_\odot$). The production rates of Forestini and Charbonnel (1997), which in contrast to Bazan *et al.* (1993) arise from self-consistent evolutionary models, are always lower by factors of $\sim 35 - 60$!

For consistency it is worthwhile to compare some predicted quantities from Tabs. 1.1 and 1.2 with observations. Although rather uncertain, the rate of galactic SN is estimated to 2^{+2}_{-1} per century from which 80% are thought to arise from core collapse events (Weiler and Sramek 1988, and references therein). The frequency of core collapse supernovae in the tables (column 4) varies between $1.4 - 6.0$ per century, where larger values are obtained for steeper IMF slopes. In view of the uncertainty in Q of 50%, all values in the table are uncertain by the same amount, hence the SN fequency must be regarded as consistent with the estimates. More constraining, however, is the relative rate of supernovae of type Ib/c to type II, for which estimates reach from 0.18 to 1.49 (Muller *et al.* 1992, Cappellaro *et al.* 1993, van den Bergh and McClure 1994, Tammann *et al.* 1994, Cappellaro *et al.* 1997). For a WR star mass limit of $\mathrm{M_{WR}} = 40$ M$_\odot$ rather low ratios between $0.02 - 0.13$ are obtained, which are inconsistent with the estimates (cf. Tab. 1.1). Lowering the mass limit towards $\mathrm{M_{WR}} = 25$ M$_\odot$ (cf. Tab. 1.2) increases the number of type Ib/c and lowers those of type II SNe, leading to more consistent ratios between $0.07 - 0.30$. A value of $\mathrm{M_{WR}} = 25$ M$_\odot$ is also supported by correlation studies between the observed distribution of WR stars with that of O-type stars (Doom 1987, van der Hucht *et al.* 1988), modern stellar evolution calculations (e.g. Meynet *et al.* 1997), and the estimated number of galactic WR stars, ranging from 1200 to 7500 (van der Hucht *et al.* 1988, and references therein). The number of galactic WR stars can be inferred from Tabs. 1.1 and 1.2 by multiplying the WR frequencies f with the average WR lifetime of ~ 0.5 Myr. Again, for $\mathrm{M_{WR}} = 40$ the resulting number of galactic WR stars ($550 - 800$) is significantly lower than the estimates. Lowering the mass limit to $\mathrm{M_{WR}} = 25$ rises the number to $1650 - 1850$, being consistent with the predictions.

All in all it becomes clear that $\mathrm{M_{WR}} = 40$ is in disagreement with numerous ob-

servational constraints. $M_{WR} = 25$ is not only favoured by the present analysis but also by correlation studies, stellar evolution calculations, and the observed number of WR stars. Additionally, IMF slopes steeper than $\Gamma \lesssim -1.7$ lead to rather low ratios of type Ib/c to type II supernovae ($\lesssim 0.15$), being inconsistent with estimates ranging from 0.18 to 1.49. Thus, the ^{26}Al production rates presented for $M_{WR} = 25$ and $\Gamma = -1.1$ to -1.7 in Tab. 1.2 should be regarded as the most consistent estimates (but see below). In this case, production rates A, encompassing probably the most reliable[5] calculations, lead to $6 - 15\%$ of the ^{26}Al budget from AGB stars, $35 - 47\%$ from type II SN, $15 - 20\%$ from type Ib/c SN, and $23 - 40\%$ from WR stars. Hence most of the ^{26}Al would be produced by core collapse supernovae. Notice, however, that the nucleosynthesis models are subject to considerable uncertainties (see production rates B and the discussion above), allowing for no firm conclusion on the origin of galactic ^{26}Al.

Prantzos and Cassé (1986) pointed out that metallicity may play an important role in the estimation of the galactic ^{26}Al budget. Since the metallicity increases with $d(\log Z)/dr \sim -0.07$ dex kpc^{-1} towards the galactic center (Shaver *et al.* 1983) and ^{26}Al production from WR stars scales as $M_{26} \propto (Z/Z_\odot)^2$ with metallicity (cf. Section 1.3.2), the WR stars contribution may be considerably increased with respect to the solar metallicity case. Additionally, ^{26}Al yields from type II supernovae may be inversely proportional to metallicity (cf. Section 1.3.4), leading to a decrease of their contribution to the galactic ^{26}Al budget. ^{26}Al production rates for initial stellar metallicities of $Z = 2$ $Z_\odot$, as proposed by Prantzos and Diehl (1996) for the average galactic metallicity, are shown in Tabs. 1.3 and 1.4 for type II and type Ib/c supernovae, and Wolf-Rayet stars (the metallicity dependence for AGB stars seems too complex to allow for a simple extrapolation of the yields of Forestini and Charbonnel (1997), hence no AGB yields are quoted; type Ib/c SN yields were assumed independent on metallicity). The lower mass limit M_{WR} for WR stars was chosen between 25 $M_\odot$ (cf. Tab. 1.3) and 20 $M_\odot$ (cf. Tab. 1.4), where the smaller value is suggested by stellar evolution calculations for high metallicity stars (Meynet *et al.* 1997).

With respect to the solar metallicity case, the use of $Z = 2$ $Z_\odot$ raises the galactic ^{26}Al budget by roughly a factor of 2. For reasonable IMF slopes ($\Gamma \gtrsim -1.7$), more than 70% of the ^{26}Al produced by massive stars originates now from WR stars, less than 30% comes from core collapse explosions. Even if type II SN yields were independent of metallicity (cf. Tab. 1.2), WR stars remain the dominant source of galactic ^{26}Al.

[5]The production rates A (AGB models of Forestini and Charbonnel (1997), SN II models of Woosley and Weaver (1995), SN Ib/c estimates of Woosley *et al.* (1995), and WR models of Meynet *et al.* (1997)) are more reliable in the sense that they represent the most detailed calculations. For AGB stars, only the estimates of Forestini and Charbonnel (1997) are based on self-consistent evolutionary models. For type II SNe, only the models of Woosley and Weaver (1995) include both presupernova and explosive nucleosynthesis, and treat in addition ν-process nucleosynthesis. For WR stars, however, both the calculations of Langer *et al.* (1995) and Meynet *et al.* (1997) are of equivalent detail. The choice of using the latter yields for production rates A is arbitrary, but interchange of both yields would not alter the conclusions, since both calculations lead to similar results.

Due to the galactic metallicity gradient there should also be a gradient in the origin of ^{26}Al: while WR stars should produce most of the ^{26}Al in the inner Galaxy, core collapse supernovae should dominate in the outer Galaxy. However, it should be emphasized again that all nucleosynthesis calculations still suffer from large uncertainties, hence the actual situation may indeed be different. This scenario represents only the current state of knowledge, which has to be verified or disproved by observations.

To estimate the galactic production rate of ^{26}Al by ONeMg novae, a different approach has to be pursued since the progenitors of ONeMg-rich novae are not well known, and additionally, recurrent outbursts may occur on the same white dwarf within the ^{26}Al lifetime of ~ 1 Myr. Using the galactic nova rate R_{novae} (in novae yr^{-1}) and the fraction f_{ONeMg} of ONeMg-rich novae, the production rate is given by

$$\dot{M}_{novae} = 0.1 \left(\frac{R_{novae}}{30\mathrm{yr}^{-1}}\right) \left(\frac{f_{ONeMg}}{33\%}\right) \left(\frac{M_{ej}}{10^{-5}\mathrm{M}_\odot}\right) \left(\frac{X_{26}}{10^{-3}}\right) \mathrm{M}_\odot \ \mathrm{Myr}^{-1}, \qquad (1.6)$$

where M_{ej} is the total amount of mass ejected in a single nova outburst and X_{26} is the mass fraction of ^{26}Al in the nova ejecta. For the nova models of Politano *et al.* (1995), Kolb and Politano (1997) derived a galactic production rate of $\dot{M}_{novae} = 0.15$ M$_\odot$ based on a calibration of nova model parameters using observations. Using improved heavy element abundances and revised rates for the ^{26}Al(p,γ)^{27}Si and ^{27}Al(p,α)^{24}Mg reactions (with respect to Politano *et al.* (1995)), José *et al.* (1997) obtain $X_{26} = 5.4\,10^{-4}$ and $M_{ej} = 1.4\,10^{-5}$ M$_\odot$ for their 1.25 M$_\odot$ ONeMg WD model, resulting in a ^{26}Al production rate of $\dot{M}_{novae} = 0.08$ M$_\odot$. Only the models of Prialink and Shara (1995), which are based on the rapid accretion scenario on CO WDs, yield sufficently high ejection masses ($M_{ej} = (1-10)\,10^{-5}$ M$_\odot$) to provide between 1 and 4 M$_\odot$ of ^{26}Al per Myr.

In summary, all proposed candidate sources could provide a substantial fraction of the observed ^{26}Al. However, the large uncertainties in ^{26}Al yield estimates for AGB stars and ONeMg-rich novae make it difficult to estimate their contribution based on nucleosynthesis model calculations only. Also ^{26}Al yield estimates for core collapse supernovae or Wolf-Rayet stars are too uncertain to allow for firm conclusions on the origin of galactic ^{26}Al. Note, however, that the ^{26}Al production rates for type II SN and $Z = \mathrm{Z}_\odot$ found in this work are considerably lower than those quoted by Timmes *et al.* (1995). On the one hand, this is due to their higher core collapse rate of 3 per century, while normalization on the galactic Lyc luminosity results in somewhat lower rates between $1.4 - 2.7$ for $\Gamma \gtrsim -1.7$ (cf. Tab. 1.1). More important is the lower mass limit M$_{\mathrm{WR}}$ for stars to exhibit a Wolf-Rayet phase. Timmes *et al.* (1995) assume that all stars up to 40 M$_\odot$ explode as type II SNe, enriching the ISM with radioactive ^{26}Al. The above discussion suggests, however, that an upper mass limit of M$_{\mathrm{WR}} = 25$ M$_\odot$ is much more consistent with observations and stellar evolution calculations, reducing the number of stars exploding as type II SNe, hence reducing the ejected amount of ^{26}Al. For M$_{\mathrm{WR}} = 25$ M$_\odot$ (cf. Tab. 1.2), core collapse SN should produce less than 1 M$_\odot$ of ^{26}Al, in constrast to 2.2 ± 1.1 M$_\odot$ as proposed by Timmes *et al.* (1995).

Taking together current nucleosynthesis calculations for all proposed candidate

$\Gamma = -1.1$, $Z = 2$ $Z_\odot$							
Object	$M_1 - M_2$	w	f	$\langle m_{26}^A \rangle$	$\dot{M}_{26}^A$	$\langle m_{26}^B \rangle$	$\dot{M}_{26}^B$
SN II	$8 - 25$	$4.5\,10^{-3}$	$1.1\,10^4$	$1.6\,10^{-5}$	0.17	$1.2\,10^{-6}$	0.01
SN Ib/c	$25 - 100$	$1.4\,10^{-3}$	$3.3\,10^3$	$6.0\,10^{-5}$	0.20		
WR	$25 - 100$	$1.4\,10^{-3}$	$3.3\,10^3$	$4.8\,10^{-4}$	1.59	$5.3\,10^{-4}$	1.75
Total	$0.08 - 100$	1.0	$2.3\,10^6$		1.96		1.76
$\Gamma = -1.35$, $Z = 2$ $Z_\odot$							
Object	$M_1 - M_2$	w	f	$\langle m_{26}^A \rangle$	$\dot{M}_{26}^A$	$\langle m_{26}^B \rangle$	$\dot{M}_{26}^B$
SN II	$8 - 25$	$1.6\,10^{-3}$	$1.5\,10^4$	$1.6\,10^{-5}$	0.23	$1.1\,10^{-6}$	0.02
SN Ib/c	$25 - 100$	$3.6\,10^{-4}$	$3.4\,10^3$	$6.0\,10^{-5}$	0.21		
WR	$25 - 100$	$3.6\,10^{-4}$	$3.4\,10^3$	$4.3\,10^{-4}$	1.47	$4.8\,10^{-4}$	1.63
Total	$0.08 - 100$	1.0	$9.5\,10^6$		1.91		1.65
$\Gamma = -1.7$, $Z = 2$ $Z_\odot$							
Object	$M_1 - M_2$	w	f	$\langle m_{26}^A \rangle$	$\dot{M}_{26}^A$	$\langle m_{26}^B \rangle$	$\dot{M}_{26}^B$
SN II	$8 - 25$	$2.4\,10^{-4}$	$2.4\,10^4$	$1.4\,10^{-5}$	0.34	$9.7\,10^{-7}$	0.02
SN Ib/c	$25 - 100$	$3.7\,10^{-4}$	$3.6\,10^3$	$6.0\,10^{-5}$	0.22		
WR	$25 - 100$	$3.7\,10^{-5}$	$3.6\,10^3$	$3.7\,10^{-4}$	1.31	$4.1\,10^{-4}$	1.47
Total	$0.08 - 100$	1.0	$6.9\,10^7$		1.87		1.49
$\Gamma = -2.4$, $Z = 2$ $Z_\odot$							
Object	$M_1 - M_2$	w	f	$\langle m_{26}^A \rangle$	$\dot{M}_{26}^A$	$\langle m_{26}^B \rangle$	$\dot{M}_{26}^B$
SN II	$8 - 25$	$1.5\,10^{-5}$	$5.6\,10^4$	$1.2\,10^{-5}$	0.68	$7.5\,10^{-7}$	0.04
SN Ib/c	$25 - 100$	$9.9\,10^{-7}$	$3.7\,10^3$	$6.0\,10^{-5}$	0.22		
WR	$25 - 100$	$9.9\,10^{-7}$	$3.7\,10^3$	$2.7\,10^{-4}$	1.00	$3.0\,10^{-4}$	1.14
Total	$0.08 - 100$	1.0	$3.8\,10^9$		1.90		1.18

Table 1.3: *Estimated ^{26}Al production rates for initial metallicities $Z = 2$ $Z_\odot$ using IMF slopes of $\Gamma = -1.1$, -1.35, -1.7, and -2.4. For production rates A, the SN II models of Woosley and Weaver (1995), SN Ib/c estimates of Woosley* et al. *(1995), and the WR models of Meynet* et al. *(1997) were used. Production rates B were calculated using SN II models of Thielemann* et al. *(1996), and WR models of Langer* et al. *(1995). For SN II, a metallicity dependence proportional to Z^{-1} was assumed, while for WRs the relation $m_{26} \propto (Z/Z_\odot)^2$ was applied. For WR stars and type Ib/c supernovae, a lower mass limit of $M_{\rm WR} = 25$ $M_\odot$ was assumed.*

$\Gamma = -1.1$, $Z = 2$ $Z_\odot$							
Object	$M_1 - M_2$	w	f	$\langle m_{26}^A \rangle$	$\dot{M}_{26}^A$	$\langle m_{26}^B \rangle$	$\dot{M}_{26}^B$
SN II	$8 - 20$	$4.0\,10^{-3}$	$9.4\,10^3$	$1.4\,10^{-5}$	0.13	$8.3\,10^{-7}$	0.01
SN Ib/c	$20 - 100$	$1.9\,10^{-3}$	$4.5\,10^3$	$6.0\,10^{-5}$	0.27		
WR	$20 - 100$	$1.9\,10^{-3}$	$4.5\,10^3$	$3.5\,10^{-4}$	1.59	$4.0\,10^{-4}$	1.79
Total	$0.08 - 100$	1.0	$2.3\,10^6$		1.99		1.80
$\Gamma = -1.35$, $Z = 2$ $Z_\odot$							
Object	$M_1 - M_2$	w	f	$\langle m_{26}^A \rangle$	$\dot{M}_{26}^A$	$\langle m_{26}^B \rangle$	$\dot{M}_{26}^B$
SN II	$8 - 20$	$1.4\,10^{-3}$	$1.3\,10^4$	$1.3\,10^{-5}$	0.18	$7.8\,10^{-7}$	0.01
SN Ib/c	$20 - 100$	$5.1\,10^{-4}$	$4.9\,10^3$	$6.0\,10^{-5}$	0.29		
WR	$20 - 100$	$5.1\,10^{-4}$	$4.9\,10^3$	$3.0\,10^{-4}$	1.47	$3.5\,10^{-4}$	1.68
Total	$0.08 - 100$	1.0	$9.5\,10^6$		1.94		1.69
$\Gamma = -1.7$, $Z = 2$ $Z_\odot$							
Object	$M_1 - M_2$	w	f	$\langle m_{26}^A \rangle$	$\dot{M}_{26}^A$	$\langle m_{26}^B \rangle$	$\dot{M}_{26}^B$
SN II	$8 - 20$	$2.2\,10^{-3}$	$2.2\,10^4$	$1.3\,10^{-5}$	0.27	$7.1\,10^{-7}$	0.02
SN Ib/c	$20 - 100$	$5.5\,10^{-4}$	$5.4\,10^3$	$6.0\,10^{-5}$	0.33		
WR	$20 - 100$	$5.5\,10^{-4}$	$5.4\,10^3$	$2.4\,10^{-4}$	1.31	$2.8\,10^{-4}$	1.54
Total	$0.08 - 100$	1.0	$6.9\,10^7$		1.91		1.56
$\Gamma = -2.4$, $Z = 2$ $Z_\odot$							
Object	$M_1 - M_2$	w	f	$\langle m_{26}^A \rangle$	$\dot{M}_{26}^A$	$\langle m_{26}^B \rangle$	$\dot{M}_{26}^B$
SN II	$8 - 20$	$1.4\,10^{-5}$	$5.3\,10^4$	$1.1\,10^{-5}$	0.59	$5.8\,10^{-7}$	0.03
SN Ib/c	$20 - 100$	$1.7\,10^{-6}$	$6.5\,10^3$	$6.0\,10^{-5}$	0.39		
WR	$20 - 100$	$1.7\,10^{-6}$	$6.5\,10^3$	$1.5\,10^{-4}$	1.00	$1.9\,10^{-4}$	1.23
Total	$0.08 - 100$	1.0	$3.8\,10^9$		1.98		1.26

Table 1.4: *Similar to Tab. 1.3 but for a lower mass limit for type Ib/c supernovae and WR stars of* $M_{\rm WR} = 20$ $M_\odot$.

sources, and considering the galactic metallicity gradient leads to a predicted amount of 2 ± 1 $M_\odot$ of ^{26}Al in the Galaxy. The quoted uncertainty comes from the uncertainty in the galactic Lyc luminosity Q (50 %), but uncertainties in the nucleosynthesis calculations may be even larger. The predicted mass of ^{26}Al is in agreement with current estimates of ~ 3 $M_\odot$, based on observations of the 1.8 MeV γ-ray line.

1.3.7 Interaction of Cosmic Rays with the Interstellar Medium

A possibility of non thermonuclear origin for the observed 1.8 MeV line has been suggested by Ramaty *et al.* (1979). The interaction of energetic cosmic-ray particles with the ambient interstellar medium leads on the one hand to direct excitation of nuclear levels of the involved nuclei and on the other hand to the production of excited secondary particles and radioactive species which decay into excited levels of other nuclei. Both processes lead to the emission of nuclear γ-ray lines of energies ranging from tens of keV to about 20 MeV.

The 1.8 MeV γ-ray line may arise from direct excitation of the 1.809 MeV level of ^{26}Mg by proton (^{26}Mg(p,p$'$)^{26}Mg*) and α-particle collisions (^{26}Mg(α,α')^{26}Mg*). It may also emanate from the radioactive decay of ^{26}Al, produced by spallation reactions via ^{26}Mg(p,n)^{26}Al, ^{28}Si(p,ppn)^{26}Al, and ^{27}Al(p,n)^{26}Al. Theoretical calculations of the γ-ray spectrum arising from the interaction of cosmic-ray particles with the ISM predict, however, a wealth of lines, the most important ones being the 4.4 MeV and 6.1 MeV lines from excited ^{12}C and ^{16}O (Ramaty *et al.* 1979). If the 1.8 MeV line would indeed originate from cosmic-ray interactions, these two lines should be much stronger and readily observable by modern γ-ray telescopes. Until now, the ^{12}C and ^{16}O lines were only tentatively identified towards the nearby Orion complex (Bloemen *et al.* 1994b), but no galaxywide emission is detected. Thus, a cosmic-ray origin of the observed 1.8 MeV emission seems unlikely.

1.4 Objective of the Thesis

The objective of this thesis can be summarized in one sentence: The determination of the origin and the distribution of galactic ^{26}Al on basis of recent COMPTEL γ-ray observations. Although the objective appears simple, the realization is rather complicated. γ-ray observations suffer from tremendous instrumental background in conjunction with low source intensities. For example, the effective detection area of COMPTEL amounts to ~ 20 cm^2 (Schönfelder *et al.* 1993) which results in the detection of only one photon per hour for a typical source with a flux of 10^{-5} ph cm^{-2}s^{-1}. During the same time interval, the telescope is bombarded by cosmic-rays, giving rise to about 100 background photons at 1.8 MeV. Additional, the origin of incoming photons can only be restricted to circumferences on the sky, which further decreases the signal-to-noise ratio.

Nevertheless, it was demonstrated that 1.8 MeV gamma-ray line photons are clearly detected by COMPTEL (e.g. Diehl *et al.* 1995b). To understand the analysis and uncertainties of COMPTEL measurements, however, some knowledge about the instrument and the data structure is imperative, hence a short description of COMPTEL is given in chapter 2. In chapter 3, three different image reconstruction techniques are applied to COMPTEL 1.8 MeV all-sky data of observation phases 0.1 - 522.5. It will be demonstrated that COMPTEL image reconstruction suffers from important statistical uncertainties, which, however, may be largely reduced by means of multiresolution algorithms.

While image reconstruction provides valuable information about the distribution of 1.8 MeV photons on the sky, comparison of COMPTEL data with astrophysically motivated intensity distribution models may shed some light on the origin and distribution of ^{26}Al in the Galaxy. In chapter 4 the COMPTEL data is compared with various intensity distributions observed at other wavelengths, from the low frequency radio domain up to the high-energy gamma-ray range. This represents the most unbiased, physically motivated comparison which can be imagined, since 1.8 MeV data are confronted with the entire knowledge about possible emission distributions on the sky. For the comparison, a rigorous Bayesian analysis is applied, which is the only consistent framework that allows inference based on the model comparison. In particular, the Bayesian analysis can be used to reduce the systematical uncertainties of the comparison, and improves the understanding of COMPTEL data. Finally, the results obtained in this thesis are summarized in chapter 5.

Chapter 2

The COMPTEL Telescope

2.1 Mission Concept

On April 5, 1991 the American space shuttle Atlantis launched the Compton Gamma-Ray Observatory (CGRO) into a circular orbit of 450 km altitude and 28.5° inclination. During the first 19 months of operation (mission phase I - from May 1991 to November 1992) the coaligned imaging telescopes, COMPTEL and EGRET, performed the first all-sky survey at γ-ray energies. Given the relatively large field of view of both instruments (about 1 steradian), 44 individual observation periods (periods of constant viewing direction) lasting approximately 14 days each sufficed to complete the survey. The following mission phases were and are devoted to detailed observations of selected objects or regions (e.g., Crab, galactic center, Virgo region). Part of these observations are initiated by guest investigators which allows scientists outside the instrument team to analyze data or even to suggest special observations.

At the low altitude of CGRO the remaining Earth's atmosphere exerts a non negligible friction on the satellite leading to a slow but continuous descent of the instrument. In Summer 1993, a reboost was performed which lifted the instrument from 345 km up to 450 km. A second (and last) reboost was done in summer 1997 which lifted the satellite into an orbit of 550 km altitude. With this reboost the expected lifetime of CGRO should be at least another 5 years (i.e. until 2002).

In this work, data from mission phases I to V (May 1991 to June 1996) were analyzed, corresponding to viewing periods 0.1 - 522.5. All data and corresponding background models were prepared for the analysis by Oberlack (1997), hence for details the reader is referred to his work.

At MeV energies the dominating photon-matter interaction process is Compton scattering, i.e., the inelastic scattering of photons off free electrons at rest (in reality the electrons are bound in atomic shells, but their binding energy is much smaller than the γ-ray energies which allows the consideration of free electrons). From the laws of conservation of energy and momentum, the geometric scattering angle $\varphi_{\rm geo}$ of an incoming photon of energy E_γ is related to the energy loss ΔE of the photon as

$$\varphi_{\rm geo} = \arccos\left\{1 + m_e c^2 \left(\frac{1}{E_\gamma} - \frac{1}{E_\gamma - \Delta E}\right)\right\} \qquad (2.1)$$

($m_e c^2$ being the energy of an electron at rest). For unpolarized photons and electrons the distribution of $\varphi_{\rm geo}$ is given by the energy-dependent Klein-Nishina cross section $d\sigma/d\varphi_{\rm geo}$.

Fig. 2.1: *Relation between incident photon direction and measured scatter angle $\bar{\varphi}$ and scatter direction (χ, ψ).*

COMPTEL measures γ-ray photons by their consecutive interactions in two parallel detector planes.[2] In the ideal case, an incident photon is first Compton scattered in the upper detector plane D_1 and then absorbed in the lower plane D_2. In both layers, the energy deposits E_1 and E_2 and the interaction locations are measured. From the

[1]Notice, however, that recently a γ-ray lens was proposed by von Ballmoos *et al.* (1996) which is based on Laue diffraction of incoming γ-ray photons in slightly inclined Germanium crystals, and operates at energies typically of 200 keV to 1300 keV. A first balloon campaign is planned for 1998.

[2]An *event* is defined by the occurrence of consecutive interactions in both detector layers within a coincidence time window of 40 ns.

energy deposits the total energy

$$E_{\rm tot} = E_1 + E_2 \tag{2.2}$$

and the Compton scattering angle

$$\bar{\varphi} = \arccos\left\{1 + m_e c^2 \left(\frac{1}{E_{\rm tot}} - \frac{1}{E_2}\right)\right\} \tag{2.3}$$

of the photon are determined, the interaction locations in both layers define the scatter direction (χ, ψ). The possible origin of the photon is then restricted to a circumference of radius $\bar{\varphi}$ around the scatter direction (χ, ψ) which is referred to as *event circle* (cf. Fig. 2.1). For photons originating from a point source, all event circles intersect in one point, which corresponds to the position of the source on the sky. In reality, all values suffer from measurement uncertainties, and even more importantly, the scattered photon is often not totally absorbed in D_2. Consequently, the calculated scatter angle $\bar{\varphi}$ differs from the true geometric scatter angle $\varphi_{\rm geo}$ leading to $\bar{\varphi} > \varphi_{\rm geo}$ for incomplete D_2 absorption. For a Compton-scattered photon the difference

$$\mathrm{ARM} = \bar{\varphi} - \varphi_{\rm geo} \tag{2.4}$$

is a measure of the telescope's angular resolution (therefore called Angular Resolution Measure). It is worthwhile to mention that the angular resolution of a Compton telescope depends (partly) on the energy resolution since $\bar{\varphi}$ is determined from two energy measurements.

2.3 Instrument Description

2.3.1 Detector Design

COMPTEL, the first space-borne Compton-telescope, was built by an international collaboration of four institutes: Max-Planck-Institut für extraterrestrische Physik (MPE, Garching, FRG), Laboratory for Space Research (SRON, Utrecht, NL), Space Science Center, University of New Hampshire (UNH, Durham, USA), and Space Science Departement of ESA (SSD, Nordwijk, NL). It consists of two parallel planes of detector modules separated by 157.7 cm (center of detector modules). The upper detector layer has been designed to optimize the chance of a single Compton scattering. It is composed of 7 cylindrical aluminium housings, each with a diameter of 27.6 cm and a thickness of 8.5 cm, filled with the liquid organic scintillator NE213A (mainly made up of carbon and hydrogen with an H/C ratio of 1.333). Each module is viewed through quartz glass windows on the sides of the housings by 8 photomultiplier tubes (PMTs) aiming to convert the weak scintillation light flashes into electrical pulses. From the sum of the absolute PMT pulse heights, the energy deposit E_1 of the ionizing radiation in the detector module is derived. From the relative pulse heights the location

of the interaction within the module is determined with a mean accuracy of 2.3 cm (1σ). The fast response of NE213A allows accurate timing (needed for time-of-flight measurements) and the discrimination between γ-ray and neutron interactions (pulse shape discrimination).

The lower detector layer consists of 14 cylindrical NaI(Tl) scintillator crystals of height 7.5 cm and diameter 28.2 cm enclosed in aluminium housings. Seven PMTs are mounted on the bottom of each module viewing the crystal through quartz glass windows. Similar to D_1, the energy deposit E_2 and the interaction location with a mean accuracy of 1.5 cm (1σ) is inferred from the sum of the absolute pulse heights and the relative pulse heights, respectively. With its high Z, the lower detector is optimized for absorption of MeV gamma-rays.

2.3.2 Instrument Response Description

The response of the instrument is defined as the distribution of measured event parameter values for incident photons of a specific energy and direction. 21 parameters are measured for each event which, on the one hand, are used to determine the fundamental quantities $E_{\rm tot}$, χ, ψ, and $\bar{\varphi}$, and, on the other hand, serve for event selection aimed to suppress background events (cf. Section 2.4). The response of the instrument to γ-ray photons of energy E_γ and incident direction (α, δ) is then given by

$$R = R(E_{\rm tot}, \chi, \psi, \bar{\varphi} | E_\gamma, \alpha, \delta), \tag{2.5}$$

which obviously maps the three photon parameters in a 4-dimensional data space. In principle, COMPTEL data should be analyzed in this 4-dimensional data space, which, however, requires large amounts of computer resources that are not available to date. Consequently, the data are currently analyzed in subspaces which are obtained by integration over some of the parameters. In this work, spectral analysis will be done in a 1-dimensional subspace obtained by integration over scatter angles and directions:

$$R_E(E_{\rm tot} | E_\gamma, \alpha, \delta) = \int_\chi \int_\psi \int_{\bar{\varphi}} R(E_{\rm tot}, \chi, \psi, \bar{\varphi} | E_\gamma, \alpha, \delta)\, d\chi\, d\psi\, d\bar{\varphi}. \tag{2.6}$$

Imaging analysis will be done in a 3-dimensional subspace, spanned by the parameters χ, ψ, and $\bar{\varphi}$ which bear information on the incident direction of the photons. The response in this subspace is obtained by integration over the total energy deposit $E_{\rm tot}$:

$$R_A(\chi, \psi, \bar{\varphi} | E_\gamma, \alpha, \delta) = \int_{E_{\rm tot}} R(E_{\rm tot}, \chi, \psi, \bar{\varphi} | E_\gamma, \alpha, \delta)\, dE_{\rm tot}. \tag{2.7}$$

In the following, the instrumental response function in these subspaces will be presented.

2.3.2.1 Energy Response Function

The energy response function $R_E(E_{\rm tot}|E_\gamma,\alpha,\delta)$ describes the distribution of the measured total energy deposit $E_{\rm tot}$ as function of the energy E_γ and the incident direction (α,δ) of the incoming photons. Figure 2.2 shows this function for 1.809 MeV photons incident from the direction of the telescope z-axis (the axis perpendicular to the detector planes) which was determined by convolution of the energy response of the D_1 detector with that of the D_2 detector (Diehl *et al.* 1992b).

Obviously, the photons are spread over a rather wide domain in $E_{\rm tot}$, ranging from the lower energy threshold at 720 keV to roughly 2 MeV. About 50% of the events are located in a Gaussian peak centered at 1.8 MeV, arising from events which were totally absorbed in D_2, after being Compton scattered in D_1. The width of this peak, which is also referred to as the *photopeak*, has been consistently determined from both preflight calibration data and from Monte Carlo simulations, and can be described by

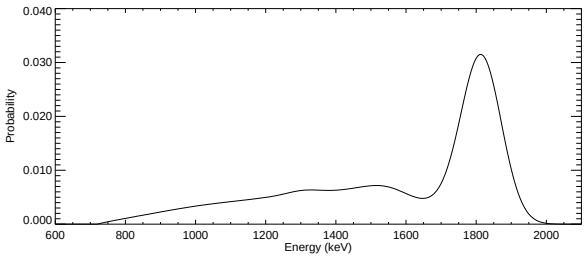

Fig. 2.2: *Energy response function for photons of energy $E_\gamma = 1.809$ MeV.*

$$\sigma(E_\gamma) = 0.01\sqrt{14.61E_\gamma + 2.53E_\gamma^2} \quad \text{MeV}, \tag{2.8}$$

where E_γ is given in MeV (Schönfelder *et al.* 1993). The tail of the response function is composed of events where the photon escapes the D_2 detector after Compton scattering (*Compton tail*) or where one or two annihilation photons created as a result of electron-positron pair production escape the D_2 detector (*Escape peaks*).

2.3.2.2 Angular Response Function

The angular response function $R_A(\chi,\psi,\bar{\varphi}|E_\gamma,\alpha,\delta)$ describes the distribution of the measured scatter directions (χ,ψ) and scatter angles $\bar{\varphi}$ as a function of the energy E_γ and the incident direction (α,δ) of the incoming photons. For an ideal Compton telescope, where all photon interactions are Compton scatters in the upper detector layer followed by total absorption of the scattered photon in one of the lower detector modules, the measured events lie on a cone of opening angle 90° with apex at the source position (α,δ). Figure 2.3 shows the angular response function in the 3-dimensional COMPTEL data space. The left image is a schematic representation of the ideal response function. The right image shows the angular response for 1.809 MeV photons incident from the direction of the telescope z-axis under the assumption of an infinite extent of the D_2 detector layer, and integrated over $E_{\rm tot} = 1.7 - 1.9$ MeV. For illustration purposes the scatter directions (χ,ψ) were integrated over the azimuthal

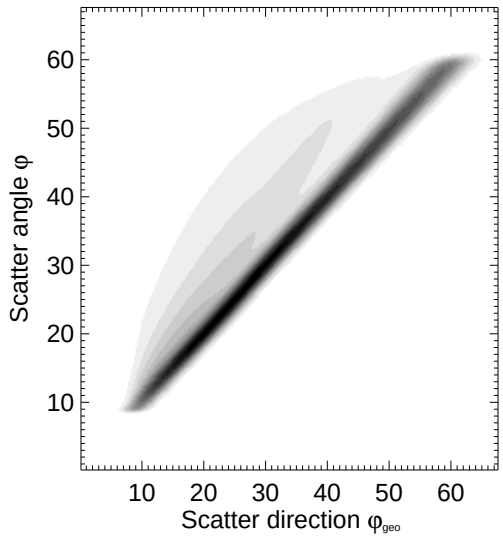

Fig. 2.3: Left: *Schematic diagram of the COMPTEL angular response in the 3-dimensional data space, spanned by the scatter direction* (χ, ψ) *and the scatter angle* $\bar{\varphi}$. Right: *Angular response for 1.809 MeV photons integrated over* $E_{\rm tot} = 1.7 - 1.9$ *MeV and over the azimuth angle of the scatter direction. An infinite extent of the* D_2 *layer was assumed.*

component which leads to an angular response representation in the $\varphi_{\rm geo}$-$\bar{\varphi}$ plane. Indeed, for unpolarized photons the angular response is symmetric with respect to the azimuth angle, hence this representation contains the full information on the angular response.

The angular response function, also referred to as the point spread function (PSF), is bounded at low and high scatter angles due to the energy thresholds of the detector modules (cf. Section 2.4). About half of the registered events lie in a small band along the diagonal (i.e., the cone mantel) and arise from photons which are totally absorbed in D_2. Photons which are not completely absorbed in D_2 lie inside the cone, since according to Eq. (2.3) the measured scatter angle $\bar{\varphi}$ is larger than the true scatter angle $\varphi_{\rm geo}$. Events which lie outside the cone mantel can analogously be attributed to incomplete energy absorption in the D_1 layer.

The 'fringe' around the cone mantel, which would disappear for an ideal Compton telescope, determines the angular resolution of the instrument. It has been consistently determined from both pre-flight calibration data and from Monte Carlo simulations, and can be described by

$$\sigma(E_\gamma) = \frac{1.247}{1 - \exp\left(-0.854 E_\gamma^{0.9396}\right)} \quad \text{degrees}, \tag{2.9}$$

where E_γ is given in MeV (Schönfelder *et al.* 1993).

For 1.8 MeV imaging analysis, events with total energy deposits in the range $E_{\rm tot} =$

$1.7 - 1.9$ MeV are collected into the 3-dimensional data space which is binned in cells of $1^\circ \times 1^\circ$ in (χ, ψ) and 2° in $\bar{\varphi}$. Comparison with the energy response for 1.809 MeV photons shows that the use of the *line-interval* $E_{\rm tot} = 1.7 - 1.9$ MeV encompasses about 50% of all registered 1.809 MeV photons (cf. Fig. 2.2). Extensive description of the treatment of the COMPTEL response in this data space were already given by Diehl *et al.* (1992a), Knödlseder (1994), and van Dijk (1996) and will not be repeated here. In summary, for a fixed telescope pointing the expected number of photons in a data space cell $e(\chi, \psi, \bar{\varphi})$ is to good approximation related to the γ-ray intensity distribution $f(\alpha, \delta)$ by

$$e(\chi, \psi, \bar{\varphi}) = b(\chi, \psi, \bar{\varphi}) + g(\chi, \psi, \bar{\varphi}) \int_\alpha \int_\delta f(\alpha, \delta) A(\alpha, \delta) p(\chi, \psi, \bar{\varphi} | E_\gamma, \alpha, \delta)\, d\alpha\, d\delta \quad (2.10)$$

where $A(\alpha, \delta)$ is the effective exposure of the D_1 layer, $p(\chi, \psi, \bar{\varphi} | E_\gamma, \alpha, \delta)$ is the point spread function defined for an infinite extent of the D_2 layer, $g(\chi, \psi, \bar{\varphi})$ is the probability that a photon scattered in D_1 in direction (χ, ψ) actually encounters a D_2 module, and $b(\chi, \psi, \bar{\varphi})$ is the expected number of events due to instrumental background (the background model). For practical purposes, the angular response function R_A was separated into

$$R_A(\chi, \psi, \bar{\varphi} | E_\gamma, \alpha, \delta) = g(\chi, \psi, \bar{\varphi}) A(\alpha, \delta) p(\chi, \psi, \bar{\varphi} | E_\gamma, \alpha, \delta). \quad (2.11)$$

Observations with different telescope points are added in good approximation using the formulae (e.g., Knödlseder 1994):

$$n(\chi, \psi, \bar{\varphi}) = \sum_{k=1}^{P} n^{(k)}(\chi, \psi, \bar{\varphi}) \quad (2.12)$$

$$b(\chi, \psi, \bar{\varphi}) = \sum_{k=1}^{P} \beta^{(k)} b^{(k)}(\chi, \psi, \bar{\varphi}) \quad (2.13)$$

$$g(\chi, \psi, \bar{\varphi}) = \sum_{k=1}^{P} g^{(k)}(\chi, \psi, \bar{\varphi}) \frac{\hat{A}^{(k)} \tau^{(k)}}{\hat{A} \tau} \quad (2.14)$$

$$\hat{A} = \sum_{k=1}^{P} \hat{A}^{(k)} \frac{\tau^{(k)}}{\tau} \quad (2.15)$$

where P is the number of observations, k is the observation index, $n(\chi, \psi, \bar{\varphi})$ is the number of events observed in data space cell $(\chi, \psi, \bar{\varphi})$, $\hat{A}^{(k)} = \max_{\alpha,\delta} A^{(k)}(\alpha, \delta)$ is the maximum of the effective D_1 detector exposure for observation k, $\tau^{(k)}$ is the duration of observation k, $\tau = \sum_{k=1}^{P} \tau^{(k)}$ is the total observation time, and $\beta^{(k)}$ is the *background scaling factor* of observation k (cf. Section 2.5.2.1). Note, that the effective D_1 detector exposure A becomes independent of the direction (α, δ) of the incident photon in this approximation.

2.4 Instrumental Background: I. Suppression

At an orbiting altitude of $\sim$ 450 km, CGRO is exposed to the continuous bombardment by various types of high-energy cosmic-rays and particles created in the upper atmospheric layers, resulting in, e.g., neutron and proton captures, spallations, and nuclear excitations. Each of these processes may create one or more γ-ray photons, which may subsequently interact in the D_1 and D_2 detectors to mimic valid events. Four basic types of sources of instrumental background events can be discerned (e.g., van Dijk 1996): cosmic-ray particles interacting directly with the spacecraft (mainly $\sim$ 1 GeV protons), neutrons and γ-ray photons produced in the Earth's atmosphere, and protons trapped in the Earth's radiation belts. In order to reduce such unwanted events and to increase the signal-to-noise ratio, several active background suppression techniques are applied, which are described in the following.

Since scintillation detectors register ionizing radiation, they generally respond to charged particles as well as photons. To distinguish both types of interactions, the detector planes are entirely surrounded by 1.5 cm thick plastic scintillator domes (NE110) which are mainly sensitive to charged particles (due to the thinness of the domes, only $\sim$ 8% of the incident 1.8 MeV photons are absorbed). 24 PMTs are mounted on the edge of each dome to convert the scintillation light flashes arising from the passage of charged particles into electrical impulses. Events which occur in coincidence with dome triggers are rejected, since they might originate from the interaction of charged particles (which triggered the dome) and not from Compton scattered γ-ray photons. The use of this anti-coincidence technique reduces the event rate by $\sim$ 99%.

An *event* is defined by the occurrence of consecutive interactions in both detector layers within a coincidence time window of 40 ns. Such a definition, however, requires the accurate determination of the interaction times in both detector layers, which, in particular for NaI scintillators, is difficult to achieve. However, the use of Amplitude and Rise time Compensated (ARC) timing (Knoll 1979) allows to measure the time between the consecutive interactions (the time of flight; TOF) with an accuracy of 1.5 ns (FWHM) (Schönfelder *et al.* 1993), which is largely sufficient for the definition of an event. In particular, this high accuracy permits to discriminate between photons moving from D_1 to D_2 and photons moving the other way, which allows a further reduction of the background noise by about $90\% - 95\%$ (van Dijk 1996).

Besides photons and charged particles, the D_1 liquid scintillator NE213A can also detect neutrons due to its high fraction of hydrogen (H/C ratio = 1.333). In contrast to photons, which interact electromagnetically with the weakly-bound electrons of the atomic shell, neutrons scatter elastically off the hydrogen nuclei of the detector and thereby give up part of their kinetical energy to the protons. The fast electrons resulting from γ-rays mainly excite the singlet state of the organic scintillator, leading to prompt fluorescence light with decay-times of the order of ns. The recoil protons arising from the scatter of neutrons, however, excite more long-lived triplet states along the ionization track, leading to increased decay times of the scintillation signal with

respect to photon interactions. Consequently, the measure of the decay time of the D_1 signal allows a discrimination of γ-ray events from neutron events (pulse shape discrimination – PSD). Vice versa, COMPTEL may be used as a neutron telescope which was successfully demonstrated by the detection of neutrons from the 1991 June 9 solar flare (Ryan *et al.* 1993).

The interaction of cosmic-ray particles with the Earth's atmosphere leads to considerable production of neutrons and γ-ray photons. Due to the relative low orbit of the satellite, the Earth as seen from the telescope occupies one third of the whole sky, hence can almost always 'shine' in the large 'aperture' of the instrument. This troublesome Earth radiation can only be suppressed through adroit 'screening' of the Earth from the field-of-view. For this purpose all events are removed from the data stream for which the event circles come closer than a threshold ζ to the Earth's horizon. Notice that such selective exclusion of events leads to a reduction of the effective exposure towards the direction of the Earth which has to be taken into account for the correct calculation of the instrumental response.

Parameter	Selection Interval
E_1	$70 - 20000$ keV
E_2	$650 - 30000$ keV
$\bar{\varphi}$	$0° - 50°$
TOF	channels $115 - 130$
PSD	channels $0 - 110$
ζ	$> 5°$

Table 2.1: *Selection criteria for the analysis of 1.8 MeV data.*

In order to optimize the signal-to-noise ratio, a set of standard selection criteria has been defined for the measured event parameters. To avoid side-effects due to differences in the lower energy thresholds of individual detector modules, lower boundaries were imposed on the energy deposits. The upper boundaries on the energy deposits reflect the lack of knowledge of the instrumental response above these energies.

2.5 Instrumental Background: II. Modelling

Despite the many selection mechanisms, most of the events ($\gtrsim 95\%$) measured by COMPTEL are made up of instrumental background events. The three contributors to the remaining background are the prompt and delayed γ-ray emission following neutron interactions anywhere in CGRO, prompt and delayed γ-ray emission following proton interactions outside of COMPTEL, and delayed ($\gtrsim 200$ ns) γ-ray interactions due to proton interactions within COMPTEL (van Dijk 1996). However, in lack of sufficient knowledge about the complex processes involved, no physical model for the remaining background distribution exists so far.

Nevertheless, to extract the valuable source events from the data, an accurate description of the distribution of instrumental background events in COMPTEL data is required. For spectral analysis (i.e. the 1-dimensional subspace integrated over the scatter directions and scatter angles), background spectra are usually estimated from

observations off the source (e.g., Lichti *et al.* 1993), but also analytical formulations have proven to give satisfactory results (e.g., Morris *et al.* 1995). For imaging analysis in the 3-dimensional data space, background models are generally derived from the observed events by smoothing or filtering techniques (Bloemen *et al.* 1994a, Knödlseder 1994), or from observations which are assumed to be free of source signal (Strong *et al.* 1993). For γ-ray line analysis a different type of model has proven to be very successful, which is based on measurement at adjacent energy intervals (Knödlseder 1994, Knödlseder *et al.* 1996b). This background model and its uncertainties will be discussed in Section 2.5.2.

2.5.1 Background in Energy Space

2.5.1.1 Background Model

The distribution of the instrumental background in energy space has been extensively studied by van Dijk (1996) and also Oberlack (1997), hence for details the reader is referred to these works. In this work, the background description proposed by Morris *et al.* (1995) is applied, which consists of an analytical fitting function

$$b(E_{\rm tot}) = b_0 \left(1 - {\rm e}^{-a_0(E_{\rm tot}-E_0)}\right) {\rm e}^{-\alpha E_{\rm tot}} \tag{2.16}$$

which is a product of a threshold function and an exponential in total energy deposit. On top of this law, three Gaussians were fitted to account for two prominent background lines at 1.406 MeV, arising from the decay of radioactive ^{40}K in the PMT glass housings of the D_1 modules (van Dijk 1996), and at 2.224 MeV, arising from thermal-neutron capture by the hydrogen nuclei of the D_1 modules. The third Gaussian was included to account for a prominent background feature between $\sim 1.3 - 1.7$ MeV which arises from the activation of ^{22}Na in the telescope structure (Oberlack 1997). Its position was fixed at 1.5 MeV, and the width was restricted to $\sim$ 100 keV. Although such a Gaussian is certainly not a very good description of the ^{22}Na feature, it sufficiently removes the residuals that remain when no such component is fitted.

2.5.1.2 Spectra of Selected Regions

COMPTEL raw count spectra for selected source positions may be obtained using the *software collimation* technique. This method consists in a further restriction of the data stream to events which possibly originated from a specific region on the sky. In other words, only those events are kept for the analysis for which the event circles intersect with a region of interest. For point sources (or circular sky regions), a so called ARM-window $|{\rm ARM}| \leq r$ is applied, which means that only events are accepted for which the observed scatter angle is sufficiently close to the expected scatter angle. r can be interpreted as the *acceptance radius* around the selected sky position which defines a

circular region with which the event circles must intersect. For a point source, r should amount to $1-2$ times the angular resolution ($1.6°$ (1σ) at 1.8 MeV), for extended regions it has to be correspondingly larger.

Examples for COMPTEL raw spectra and background-subtracted spectra are shown in Fig. 2.4. The left panels were obtained from 17 observation periods with pointing directions within 50° of the northern galactic pole. As acceptance region, an ARM-window of $r = 10°$ around $(l, b) = (0°, 90°)$ was chosen, from which no 1.8 MeV emission is expected. Indeed, no line-like feature is present in the residual spectrum. The right panels were derived from the sum of 33 observations periods sufficiently close to the galactic center, with an acceptance circle of $r = 5°$ around $(l, b) = (0°, 0°)$. The 1.8 MeV line, which is already perceptible in the raw spectrum, is the outstanding feature in the residual spectrum.

Further residual spectra are shown in Fig. 2.5 for four selected sky regions. The spectra were obtained by summing the events from numerous observation periods with pointing directions within 50° of the selected positions. The positions of the regions were motivated from imaging analysis which suggests prominent 1.8 MeV emission towards these directions (cf. Section 3.5).

2.5.2 Background in Imaging Data Space

2.5.2.1 The Standard 1.8 MeV Background Model

Based on the observation that the distribution of scatter angles (χ, ψ) is rather independent on the total energy deposit $E_{\rm tot}$ of the events, Knödlseder (1994) proposed an instrumental background model for γ-ray line analysis which makes use of contemporaneously measured events at adjacent energy intervals. The differences in the $\bar{\varphi}$ distributions $P(\bar{\varphi}) = \sum_{\chi,\psi} n(\chi, \psi, \bar{\varphi})$ for different energy deposits can be removed by normalizing the adjacent energy data to the line data, using

$$b(\chi, \psi, \bar{\varphi}) = \langle n_{adj}(\chi, \psi, \bar{\varphi}) \rangle \frac{\sum_{\chi',\psi'} n_{line}(\chi', \psi', \bar{\varphi})}{\sum_{\chi',\psi'} n_{adj}(\chi', \psi', \bar{\varphi})}, \tag{2.17}$$

where n_{adj} and n_{line} are the 3-dimensional event matrices of the adjacent and the line-interval, respectively. To minimize the statistical fluctuations of the background model, adjacent energy intervals of $E_{\rm tot} = 0.85 - 1.7$ MeV and $E_{\rm tot} = 1.9 - 10.0$ MeV are chosen for 1.8 MeV line analysis, which provide ~ 15 times more events than the line interval $E_{\rm tot} = 1.7 - 1.9$ MeV. Remaining statistical fluctuations of the adjacent energy intervals are reduced by applying a slight smoothing in the (χ, ψ) direction, using

$$\langle n_{adj}(\chi, \psi, \bar{\varphi}) \rangle = g(\chi, \psi, \bar{\varphi}) \cos\psi \frac{\sum_{\chi',\psi'} w(\chi', \psi' | \chi, \psi, \bar{\varphi}) n_{adj}(\chi', \psi', \bar{\varphi})}{\sum_{\chi',\psi'} w(\chi', \psi' | \chi, \psi, \bar{\varphi}) g(\chi', \psi', \bar{\varphi}) \cos\psi'}, \tag{2.18}$$

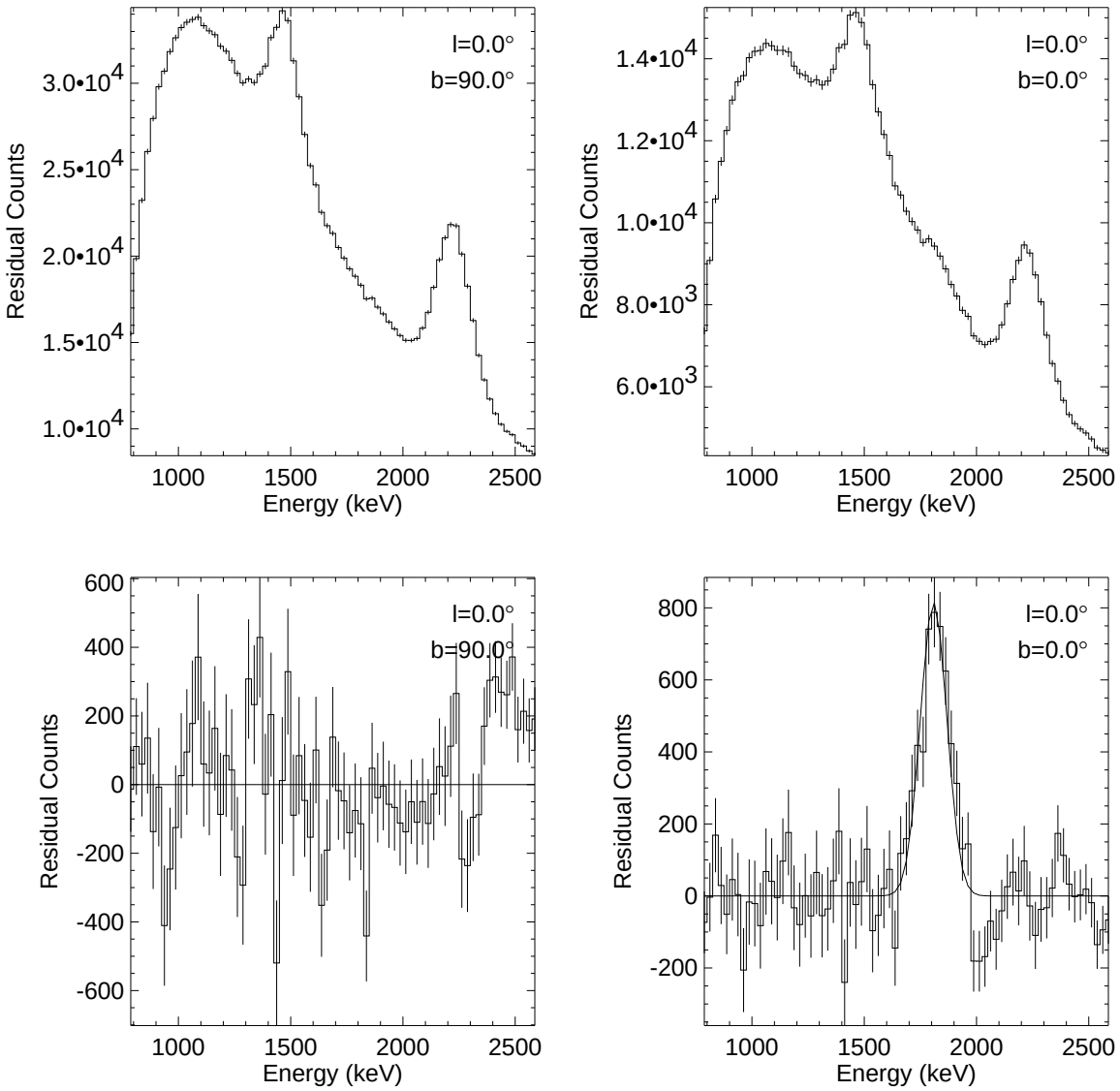

Fig. 2.4: Top-left: *Software-collimated COMPTEL raw spectrum of the northern galactic pole (sum of 17 observation periods using an ARM-window of* $r = 10°$*).* Top-right: *Software-collimated COMPTEL raw spectrum of the galactic center (sum of 33 observation periods using an ARM-window of* $r = 5°$*). Excess emission between* $1.7 - 1.9$ *MeV due to the galactic 1.8 MeV line is already visible in this raw spectrum.* Bottom-left: *Residual spectrum after fitting the background model to the northern galactic pole spectrum.* Bottom-right: *Residual spectrum after fitting the background model to the galactic center spectrum. An additional Gaussian with mean of 1.809 MeV and width of* $\sigma = 58.9$ *keV was fitted to account for the 1.8 MeV line.*

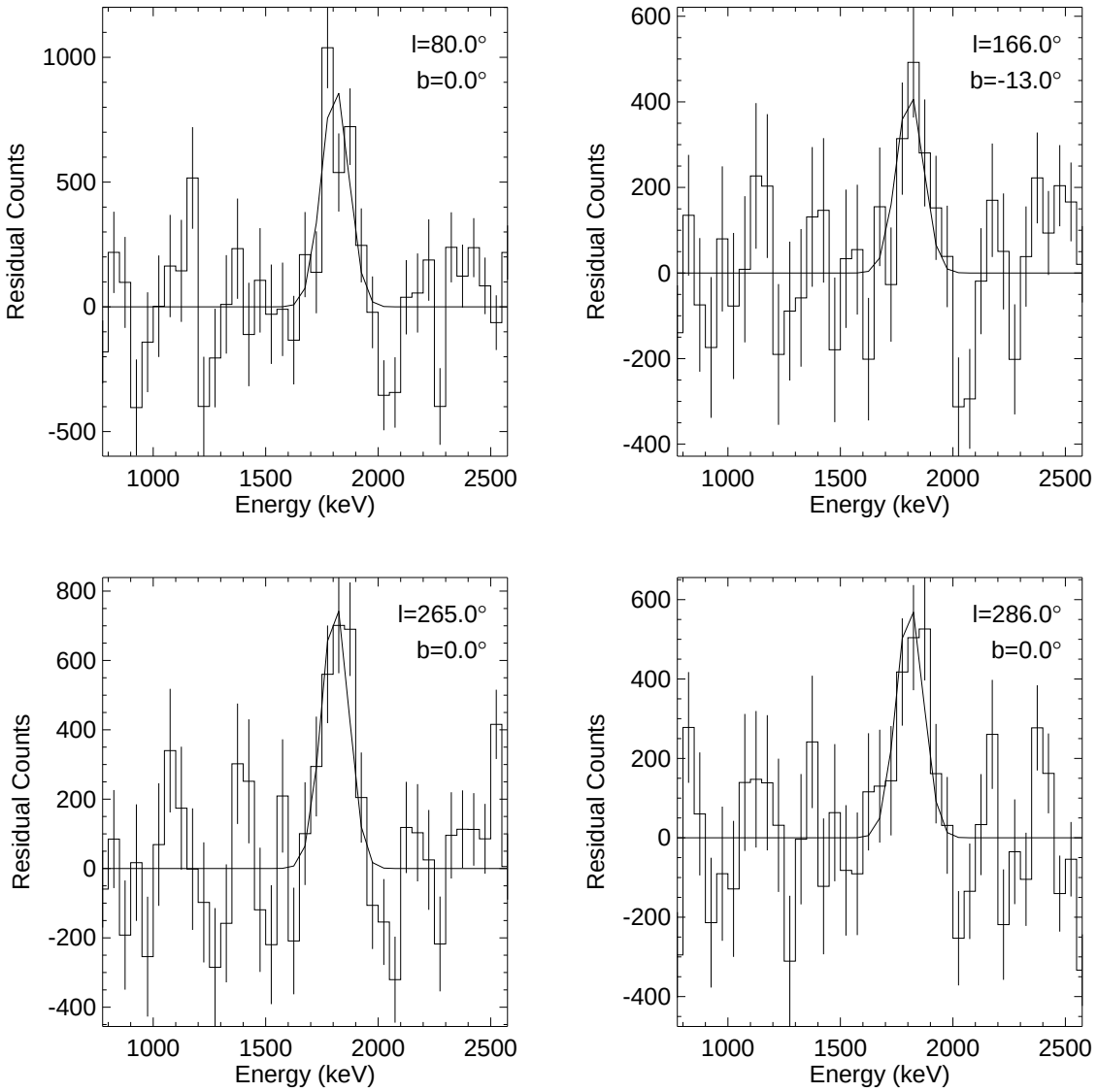

Fig. 2.5: *Background subtracted COMPTEL energy spectra for four selected regions* (top-left: *Cygnus*, $|ARM| \leq 10°$, *25 observation periods;* top-right: *Auriga-Camelopardalis-Persus (ACP) region*, $|ARM| \leq 5°$, *13 observation periods;* bottom-left: *Vela*, $|ARM| \leq 5°$, *17 observation periods;* bottom-right: *Carina*, $|ARM| \leq 4°$, *22 observation periods).*

where

$$w(\chi', \psi' | \chi, \psi, \bar{\varphi}) = \frac{\cos\psi}{2\pi\sigma^2} \exp\left(-\frac{\theta^2}{2\sigma^2}\right) \tag{2.19}$$

is a Gaussian in spherical coordinates with $\cos\theta = \sin\psi \sin\psi' + \cos\psi \cos\psi' \cos\chi - \chi'$ being the cosine of the angular separation of (χ, ψ) and (χ', ψ'). Since the (χ, ψ) distribution of the events roughly follows the geometry function multiplied by the

cosine of the latitude,[3] weighting with this function according to Eq. (2.18) preserves first-order gradients of the (χ, ψ) distribution.

An important feature of this algorithm is that the (χ, ψ) structure of the adjacent energy data space n_{adj} is not modified by Eq. (2.17). Smoothing in the (χ, ψ) direction doesn't affect possible source signatures in the background model if the smoothing width σ is smaller than the instrument's angular resolution. Hence any 'cone' structure which might be present in the adjacent energy data due to continuum source signal will persist, and consequently is included in the background model. Prerequisites for the complete suppression of a continuum source are that the spectral index of the source is the same as that of the instrumental background and that the $\bar{\varphi}$ distribution of the source signal in the background model matches that of the data. Both prerequisites are only roughly fulfilled. From the difference of the spectral indices between the instrumental background and the potential continuum sources, a continuum contribution of 10% at maximum has been estimated for 1.8 MeV analysis (Knödlseder 1994).

Since the normalization procedure Eq. (2.17) forces the total number of predicted background events to be equal to the number of observed events, the instrumental background will be overestimated for observations where 1.8 MeV emission is present. To avoid this overestimation, the background model is scaled by a *background scaling factor* β, which originally was derived by normalization of the total number of expected background events to the number of events in two thin adjacent energy intervals $E_{\rm tot} = 1.6 - 1.7$ MeV and $E_{\rm tot} = 1.9 - 2.0$ MeV. Oberlack (1997), however, found a systematic bias in this normalization, in that it overestimates the instrumental background for observations made later during the mission. As origin of this bias, he identified the build-up of a strong background feature between $1.3 - 1.7$ MeV which is probably due to the radioactive decay of ^{22}Na, created by ^{27}Al(p,x)^{22}Na reactions in the aluminum structure surrounding the D_1 modules. An additional bias may be introduced by the strong ^{24}Na cascade (e.g., van Dijk 1996) which also increased by about a factor 4 in activity after the satellite reboost in 1993. In order to remove this bias, he subtracted the contribution of the ^{22}Na and ^{24}Na background components from the data of each observation period and re-determined individual background scaling factors from the remaining events. He could demonstrate that for these new background scaling factors, no longtime trend exists anymore.

2.5.2.2 Uncertainties

The uncertainty in the background model normalization is twofold: Firstly, the normalization of individual observation periods is uncertain to about 0.5% at maximum (Oberlack, private communication). This results in a symmetric uncertainty in the

[3] Generally, ψ represent the latitude component of the scatter direction. In this work, the selected data space coordinate system is aligned with the galactic coordinate system, hence χ corresponds to galactic longitudes and ψ to galactic latitudes.

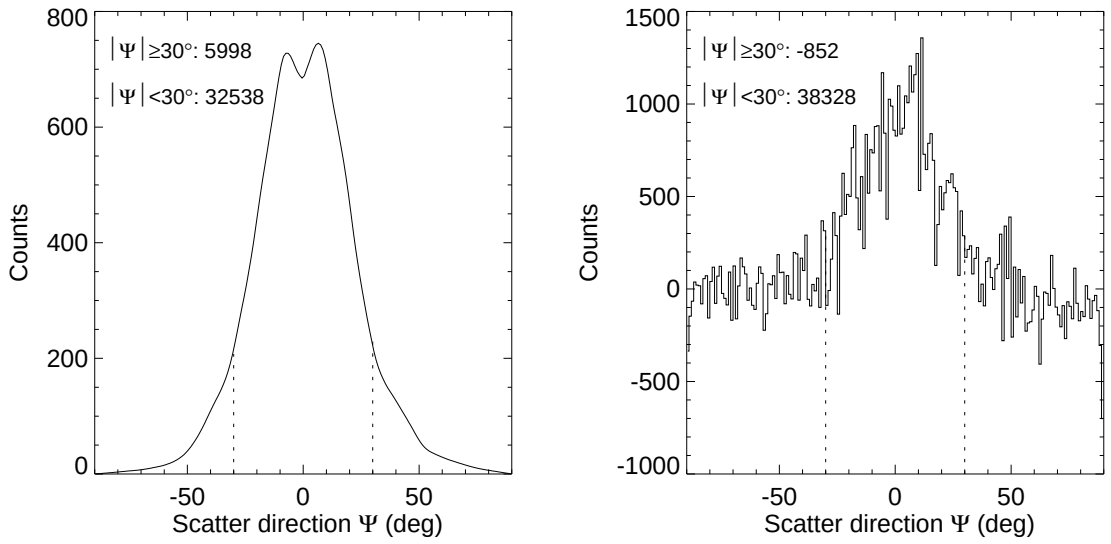

Fig. 2.6: Left: *ψ distribution of an exponential disk model with radial scale length of* $R_0 = 4.5$ kpc, *vertical scale height of* $z_0 = 90$ pc, *and total mass of* $M = 1$ M$_\odot$. *About* 20% *of all events have scatter directions* $|\psi| \geq 30°$. Right: *ψ distribution of observed events minus background model.*

total number of source counts of $37\,476 \pm 19\,648$. Secondly, the zero level was estimated from observations with pointings at galactic latitudes $|b| > 40°$ which implicitly assumes that no 1.8 MeV emission is present in these observations. However, galactic plane emission also contributes events at these high latitudes due to photons with large scatter angles $\bar{\varphi}$. Indeed, with a maximum allowed scatter angle of $\bar{\varphi} \leq 50°$ and typical scatter directions (χ, ψ) $\pm 35°$ around the pointing, only observations with pointings within 5° of the galactic poles are free of galactic plane emission. This is illustrated in Fig. 2.6 which shows in the left panel the ψ distribution of an exponential disk model with a radial scale length of $R_0 = 4.5$ kpc and a vertical scale height of $z_0 = 90$ pc. Although this model, which is a good first order description of COMPTEL 1.8 MeV data (cf. Section 4.4.1), consists only of galactic plane emission, about 20% of the events have scatter directions $|\psi| \geq 30°$. Hence, using high latitude observations to estimate the zero level of the background model leads to an overestimation of the instrumental background. This is illustrated in the right panel of Fig. 2.6 which shows the ψ distribution of the observed events minus the background model. Above $|\psi| \geq 30°$, a negative residual is found which clearly indicates an overestimation of the instrumental background component. To recover the same event distribution as for the exponential disk model (i.e., 20% of the events obey $|\psi| \geq 30°$), the background model has to be scaled by $\beta = 0.9935$, which leads to a rise of the total number of source photons from 37 476 (for $\beta = 1$) to 63 019. Additionally, the ψ distribution of the observed events minus the background model shows an asymmetry with respect to the galactic plane with fewer events at positive ψ than at negative ψ. This could be either due to a normalization

error of the background model or may be a signpost of high-latitude emission in the southern galactic hemisphere (the question of the reality of high-latitude emission will be adressed in Section 4.4.4). Using only positive ψ for the background model normalization leads to a further reduction of the background scaling factor to $\beta = 0.9915$ which corresponds to 70 878 1.8 MeV source photons. Thus, in total, there might be a systematic $\sim 1\%$ uncertainty in the background model normalization towards smaller scaling factors. Indeed, model fits prefer a background scaling factor around $\beta \approx 0.995$ (cf. Section 4.3.2.2), while image reconstruction suggests an even smaller factor of $\beta \approx 0.985$ (cf. Section 3.2.2.4).

Chapter 3

Image Reconstruction

3.1 The Inverse Problem

3.1.1 Bayesian Inference

COMPTEL data are collected in the 3-dimensional image data space, which is related to the sky intensities f_j (the image space) by the set of linear equations (cf. Eq. (2.10))

$$e_i = \sum_{j=1}^{M} R_{ij} f_j + b_i, \qquad 1 \leq i \leq N \tag{3.1}$$

or in more compact matrix notation

$$\mathbf{e} = \mathbf{R}\mathbf{f} + \mathbf{b}, \tag{3.2}$$

where e_i is the number of expected events in a data space cell, R_{ij} is the point spread function (PSF), and b_i is the instrumental background model. N and M are the size of the data and image space, respectively. Such a set of equations is often extremely ill-conditioned, i.e. the solution vector is numerically unstable (Press *et al.* 1992). Additionally, instead of $\mathbf{e}$ – the true sky intensity distribution mapped in the COMPTEL data space – a measurement $\mathbf{n}$ is available which is subject to statistical and systematical errors. In most cases, equations like Eq. (3.2) have very poor error propagation properties[1] (Schwarz 1988). In consequence when Eq. (3.2) is solved using a measurement on the left hand side the solution vector $\mathbf{f}$ generally exhibits high frequency, large amplitude oscillations (see Lucy (1994) for a nice example and Kebede (1994) and Schwarz (1988) for the deeper theoretical understanding of noise amplification). Finally, for the usual choice of data space cell sizes of $1^\circ \times 1^\circ \times 2^\circ$ (in χ, ψ, and $\bar{\varphi}$)

[1] The error propagation properties of a system of linear equations $\mathbf{A}\mathbf{x} = \mathbf{b}$ is determined by the *conditional number* of $\mathbf{A}$ which is defined as the ratio of the largest to the smallest singular value of $\mathbf{A}$ (Schwarz 1988).

and image space pixels of $1^\circ \times 1^\circ$ (in galactic longitude and latitude), the COMPTEL response is not even left-invertible, i.e. there exists no $\mathbf{R}^{-1}$ that fulfills $\mathbf{R}^{-1}\mathbf{R} = \mathbf{1}$. Equation (3.2) is overdetermined and consequently has no (exact) solution.

However, since $\mathbf{n}$ is subject to measurement errors, an exact solution is not even desirable. It is sufficient to find a solution $\mathbf{f}$ for which $\mathbf{e}$ is acceptably close to the measurement $\mathbf{n}$ – acceptably close with respect to the measurement errors. With this relaxation there is suddenly a wealth of 'solutions', especially in the case of an overdetermined problem. To obtain a unique solution, additional constraints are needed.

Finding an acceptably close solution to Eq. (3.2) subject to some constraint(s) is known as regularization. Regularization can be treated in the framework of Bayesian probability theory where it becomes a problem of *Bayesian inference*. In contrast to 'orthodox' statistics, Bayesian probabilities are not frequencies coming from the repetition of an experiment but plausibilities of hypothesis given a measurement and some prior knowledge ('orthodox' methods ignore prior knowledge). Consequently Bayesian probabilities are always conditional probabilities on some collective background information I.

The basic equation of Bayesian inference is[2]

$$P(H|D,I) = P(H|I)\frac{P(D|H,I)}{P(D|I)}, \tag{3.3}$$

where H is the hypothesis to be tested, D is the measured data and I is any prior information available before the data were taken. Equation (3.3) describes a learning process, namely how the plausibility of a hypothesis H is changed when new data D becomes available (on the justification and uniqueness of Eq. (3.3) see Jaynes (1996), chapter 2). The constituents of Eq. (3.3) are interpreted as follows:

$P(H|D,I)$ is called the posterior probability of the hypothesis, given the data and the prior information. It quantifies the knowledge about the hypothesis after the *learning process*, i.e. after consideration of the data.

$P(H|I)$ is the prior probability of the hypothesis (or just *prior*). It quantifies the knowledge about the hypothesis which was available before the measurement.

$P(D|I)$ is the prior probability of the data or *evidence*. It does not depend on the hypothesis to be examined and will (in cases considered in this work) always be a normalization constant.

$P(D|H,I)$ is the direct probability of the data, given the hypothesis and the prior information. The direct probability is called the *sampling distribution* when the hypothesis is held constant and one considers different sets of data, and it is called the *likelihood function* when the data are held constant and one varies the hypothesis. It is this term which will quantify the acceptability of a solution.

[2] also referred to as Bayes' theorem

To complete the framework of Bayesian inference the term *model* has to be introduced. The model links the observable D (i.e. the data) to the hypothesis H which should be tested. In the case of COMPTEL image reconstruction it consists on the one hand of Eq. (3.2), i.e. the way that an intensity distribution is related to the observables. On the other hand, it depends on the pixel grid which was chosen for the intensity distribution, since a particular choice is related to a particular set of hypotheses. If for example a $1^\circ \times 1^\circ$ pixel grid is chosen and no constraint on the correlation of pixels is imposed, the hypotheses under consideration will be "The sky intensities in $1^\circ \times 1^\circ$ pixels will take the mutually independent values f_j". In the case of model fitting (cf. Chapter 4) the second part of the model is more explicitly defined as for example "The sky intensities follow the distribution of molecular hydrogen". While for model fitting the importance of the model is inherent, the importance of the *pixel grid* as part of the model for image reconstruction is often neglected. The problems which may arise due to this ignorance will be discussed in Section 3.4.1.

The Bayesian formulation leads to a posterior *probability bubble* $P(H|D,I)$ and not to any single hypothesis H. For practical purposes it is desirable to get a *best* H – corresponding to a unique solution of Eq. (3.2). A common choice is the most probable hypothesis $\hat{H}$ called the *mode* or the *maximum a posteriori* (MAP) estimate of H. It is the best in the sense of being the single hypothesis one has greatest confidence in. Alternatively one could chose the mean

$$\langle H \rangle = \int H P(H|D,I) dH \tag{3.4}$$

which is probably a better estimate than the mode if H is very broad and flat (Loredo 1990). For image reconstruction only MAP estimates will be considered.

3.1.2 Assigning Probabilities

There is nothing within the Bayesian theory that tells how to assign[3] the probabilities of Eq. (3.3). This must come from outside probability theory and is mainly related to the knowledge the observer has. At present, four fairly general principles for assigning probabilities are known: group invariance, maximum entropy, marginalization, and coding theory (Jaynes 1996, p. 402).

For parameter estimation, the prior $P(H|I)$ and the likelihood $P(D|H,I)$ are both direct probabilities and must be assigned a priori. $P(D|I)$ is independent of H and can be calculated using

$$P(D|I) = \sum_H P(H|I) P(D|H,I), \tag{3.5}$$

[3]In Bayesian inference probabilities are assigned, not measured. Probabilities that are assigned directly are called *direct probabilities*. All other probabilities are derived from direct probabilities using the sum and product rules for probabilities.

where the sum is taken over all possible hypotheses.

For the assignment of direct probabilities the principle of maximum entropy can be used. It is based on Shannon's theorem which states that the only reasonable measure of the amount of ignorance (or uncertainty) in a probability distribution $P(h|I)$ is

$$H \equiv -\sum_h P(h|I) \log P(h|I). \tag{3.6}$$

The principle of maximum entropy then states that if one has some *testable* information[4] I, one can assign a probability distribution to a proposition h such that $P(h|I)$ contains only the information I. This assignment is done by maximizing the entropy H subject to the constraints represented by the information I (Bretthorst 1990). Consequently, probabilities assigned by the maximum entropy principle are maximally uninformative (or least informative), while still incorporating the known information I.

The collection of data by COMPTEL is a Poisson process for which the likelihood function can be derived using the maximum entropy principle (Gull and Skilling 1984):

$$P(D|H,I) \;= \prod_{i=1}^{N} \frac{e_i^{n_i} e^{-e_i}}{n_i!}. \tag{3.7}$$

For the prior $P(H|I)$ two different choices will be discussed: the uniform prior with positivity constraint and the maximum entropy prior. When nothing is known, except that $P(H|I)$ should be normalized,[5] the principle of maximum entropy reduces to the uniform prior (Bretthorst 1990). For image reconstruction the uniform prior with positivity constraint is given by

$$P(H|I) = \begin{cases} \dfrac{1}{f_{max}^M} & \text{if} \quad 0 \le f_j < f_{max} \\ 0 & \text{else} \end{cases} \tag{3.8}$$

where f_{max} is the maximum possible intensity in an image pixel. This maximum is required to make the prior normalizable but the exact value of f_{max} has no influence on the reconstruction. Indeed, Bretthorst (1990) notes that the functional form of the prior makes little difference as long as $P(H|I)$ is slowly varying compared to the likelihood. For the uniform prior it follows that $P(H|D,I) \propto P(D|H,I)$ and finding the MAP estimate becomes a maximum likelihood estimation. An algorithm which extracts the MAP estimate for the uniform prior is presented and discussed in Section 3.2.

If in addition to the normalization constraint the magnitude of the expected intensities is known, the *maximum entropy prior*

$$P(H|I) = \frac{1}{Z_S} \prod_{j=1}^{M} \frac{1}{\sqrt{f_j}} \exp(\alpha S) \tag{3.9}$$

---

[4] Given a probability distribution $P(h|I)$, the information I is called *testable* if it can be determined unambiguously that the distribution is consistent with the information I (Loredo 1990).

[5] This special case of the more general maximum entropy principle is referred to as the *principle of insufficient reason* or the *principle of indifference*.

may be used, where

$$S \equiv \sum_{j=1}^{M} \left[f_j - m_j - f_j \ln \left(\frac{f_j}{m_j} \right) \right] \tag{3.10}$$

is the configurational entropy of the image and

$$Z_S \approx \left(\frac{2\pi}{\alpha} \right)^{\frac{M}{2}} \tag{3.11}$$

is a normalization constant (Gull and Skilling 1985). The parameters of the maximum entropy prior are m_j and α. m_j represents the expected magnitude for the sky intensities f_j, defined on each cell separately. Commonly, in the absence of any reason to favour one pixel over another, m_j is taken to be simply a constant m. α is an inverse measure of the expected spread of the intensities f_j about the expected magnitude m_j.

Gull and Skilling (1985) claim that any positive and additive distribution, and hence sky intensities, can be identified with a probability distribution by removing it's dimensionality $p_j = f_j / \sum f_j$. They consequently define the (negative) information content of the image as $S = -\sum p_j \log p_j$, and conclude that maximizing this entropy automatically produces the most uniform, featureless image that is consistent with the constraints (i.e. the data). Skilling (1988) generalizes the entropy definition to the form of Eq. (3.10). A different viewpoint is taken by Loredo (1990) who argues that the configurational entropy S of a signal is not Shannon's entropy H since a signal is not a probability distribution. Hence identifying the entropy S of a distribution as the uniquely correct measure of its information content, as supported by Skilling (1990), does not apply to signals. Jaynes (1986) notes that, while the maximum entropy prior contains important truth, a fully satisfactory prior for image reconstruction, which expresses all prior information, needs to be able to express the common-sense judgment that correlations vary with the distance between pixels. He nicely illustrates the incomplete incorporation of prior information into the maximum entropy (MAXENT) prior by his statement that "...bare MAXENT is surprised to find isolated stars, but astronomers are not" (Jaynes 1986). Nevertheless, maximum entropy inversion has proven to be successful in many problems.[6] The application of a maximum entropy algorithm to COMPTEL data is presented and discussed in Section 3.3.

[6]The maximum entropy method (MEM) for inverse problems is not identical to the maximum entropy principle for assigning direct probabilities. While the first is a regularization principle (where the generality is called into question; see e.g., Loredo (1990) and Jaynes (1986)) the second is a fundamental principle for assigning least informative priors (Jaynes 1996).

3.2 The Richardson-Lucy Algorithm

3.2.1 The Richardson-Lucy Iteration

Based on Bayes' theorem of conditional probabilities Richardson (1972) proposed an iterative method for the restoration of degraded images. Although he wasn't able to prove the convergence of a solution he noted that in all investigated cases the method did converge. Lucy (1974) rediscussed the iterative scheme and showed that each iteration leads to an increase of the likelihood. He guessed that the iterative scheme converges towards the maximum likelihood (ML) solution and demonstrated that for $M \leq N$ (more data space cells than image pixels) the ML solution is unique. Shepp and Vardi (1982) showed that the Richardson-Lucy (R-L) iterations converge to the ML solution for Poisson statistics in the data when they rediscovered the algorithm for emission tomography.[7]

The R-L iteration step $r \to r+1$ is given by

$$f_j^{r+1} = f_j^r \frac{\sum_{i=1}^{N} \frac{n_i}{e_i^r} R_{ij}}{\sum_{i=1}^{N} R_{ij}} \tag{3.12}$$

where

$$e_i^r = \sum_{j=1}^{M} R_{ij} f_j^r + b_i. \tag{3.13}$$

It is obvious that Eq. (3.12) fulfills the positivity constraint $f_j^{r+1} \geq 0$ if the initial estimate f_j^0 also had this property. Further it is clear that deviations of n_i/e_i^r from unity on a length scale large compared to that of R_{ij} will dominate the correction. Small scale deviations such as pixel-to-pixel noise will, to a large extent, be averaged out when folded with R_{ij} and will result, therefore, in only small corrections to f_j^r. Consequently, it can be anticipated that the iterative scheme will during the first iterations take the significant information into account while later iterations will result only in small corrections that slowly tend to match the successive estimates e_i^r to the statistical fluctuations. This behavior is confirmed in many applications (e.g., Lucy 1974, White 1994, Lucy 1994). A conventional method to avoid fitting of statistical fluctuations is to stop the iterations before *overfit* sets in (an alternative approach of damping the R-L corrections is described by White (1994)). Various stopping criteria were proposed, such as goodness-of-fit tests (Lucy 1974, 1994), cross validation (Perry and Reeves 1994), or the occurrence of a minimum in the curvature of the trajectory in the likelihood-entropy space (Lucy 1994). In this work optimum stopping criteria will be derived from simulations but it will turn out that there is no optimum single stopping criterion. The stopping criterion merely depends on the question which is addressed (cf. Section 3.2.2.2).

[7] Shepp and Vardi (1982) demonstrate that the Richardson-Lucy algorithm is a special case of the expectation-maximization (EM) algorithm (Dempster *et al.* 1977). The expectation step is given by Eq. (3.13), the maximization step by Eq. (3.12).

3.2.2 Application to COMPTEL Data

3.2.2.1 Algorithm Implementation

A major drawback of the Richardson-Lucy iteration is its slow convergence speed. Especially in the COMPTEL case where the estimate e_i is composed of a small signal component and a (fixed) overwhelming background component, correction factors n_i/e_i^r will always lie in a small band around unity, leading to only small intensity corrections. Hence convergence to the ML solution requires some 1000 iterations. Several modifications of the R-L iteration to accelerate convergence were proposed (Fessler and Hero 1994). For COMPTEL data it was found that the ML-LINB-1 algorithm of Kaufman (1987) gives reasonable acceleration without degrading the reconstruction properties. Kaufman (1987) noted that R-L is the special case $\lambda^r \equiv 1$ of the form

$$f_j^{r+1} = f_j^r + \lambda^r \delta f_j^r \tag{3.14}$$

where

$$\delta f_j^r = f_j^r \left(\frac{\sum_{i=1}^N \frac{n_i}{e_i^r} R_{ij}}{\sum_{i=1}^N R_{ij}} - 1 \right) \tag{3.15}$$

is the R-L correction written in additive form. The convergence can be accelerated by using a line-search to find the λ^r which maximize the likelihood subject to the constraint $\lambda^r \delta f_j^r > -f_j^r$ (this constraint assures $f_j^{r+1} > 0$). Application of this scheme to COMPTEL data leads to acceleration factors between 10 and 100.

Since the ML solution for COMPTEL is unique (Lucy 1974), it consequently does not depend on the initial intensity distribution f_j^0. However, to avoid overfitting the R-L iterations are stopped before the ML solution is reached and the obtained reconstruction might depend on the initial estimate since it dictates the way in which in image space the ML solution is reached. In the application of the R-L algorithm to COMPTEL data a uniform grey image $f_j^0 = m$ is used as an initial estimate, corresponding to no prior knowledge about the intensity distribution.[8] Since the total number of source counts is fixed by the normalization of the background model it appears natural to choose

$$m_{\mathrm{nat}} = \frac{\sum_{i=1}^{N} (n_i - b_i)}{\sum_{i=1}^{N} \left(\sum_{j=1}^{M} R_{ij} \right)} \tag{3.16}$$

as the initial intensity. In the following this choice will be called the 'natural initial intensity' m_{nat}.

[8] Shepp and Vardi (1982) note that the choice of the initial estimate is somewhat akin to the choice of a Bayes prior.

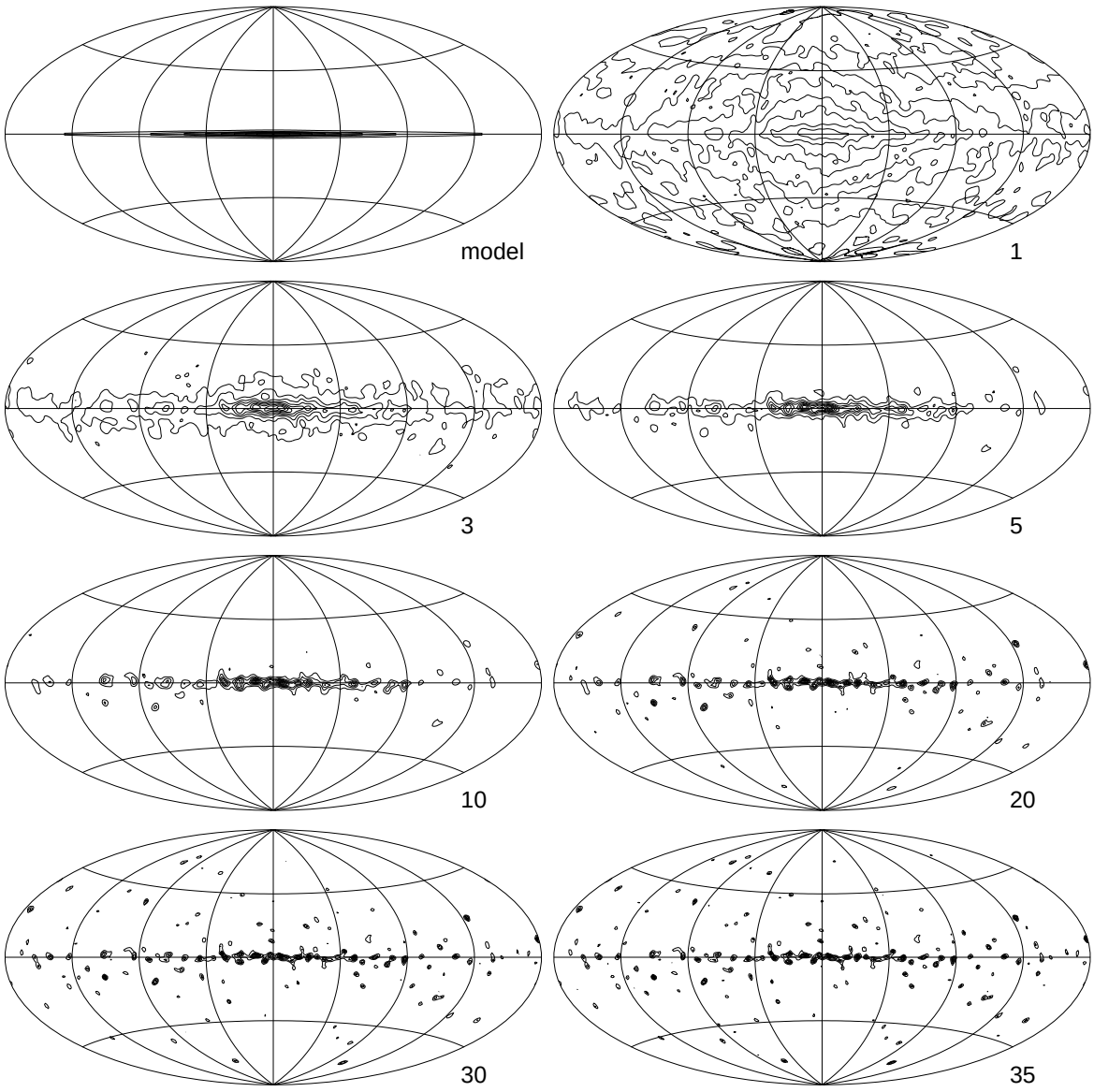

Fig. 3.1: *Richardson-Lucy iterations for the reconstruction of COMPTEL all-sky mock data. The galactic ^{26}Al density distribution was supposed to follow an exponential disk of radial scale length $R_0 = 4.5$ kpc and scale height $z_0 = 90$ pc, the total galactic ^{26}Al mass was fixed to 3 $M_\odot$. The intensity distribution of the exponential disk model is shown in the upper left panel; the other panels show the reconstructed intensity distribution f_j^r after iterations $r = 1, 3, 5, 10, 20, 30,$ and 35. A grey image of natural initial intensity $m_{\rm nat} = 7.8\,10^{-5} ph\ cm^{-2} s^{-1} sr^{-1}$ was used as initial estimate.*

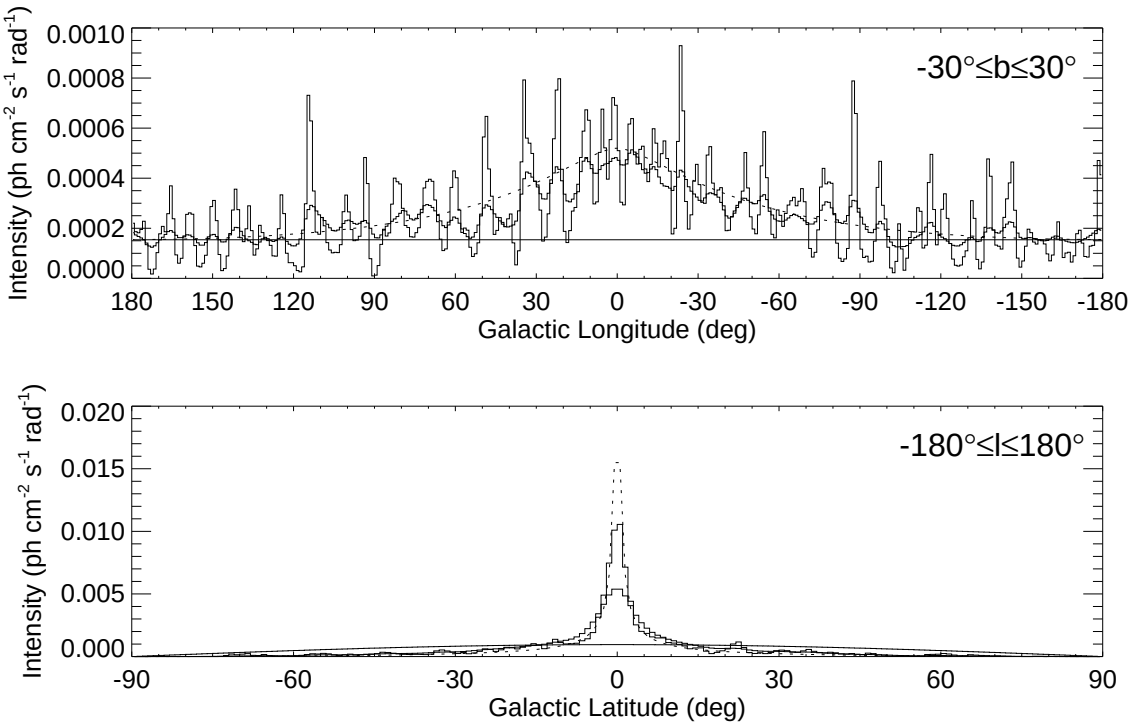

Fig. 3.2: *Longitude (top) and latitude (bottom) profiles of the reconstructed emission after iterations r = 0, 5, and 35. The profile of the model intensity distribution is shown as dotted line, reconstructed profiles are plotted as histograms.*

Examples of R-L reconstructions of COMPTEL all-sky mock data are shown in Fig. 3.1 and Fig. 3.3. All mock data used throughout this work are based on a two-component data space model e_i: the all-sky background model b_i for observation periods 0.1 - 522.5 plus a model intensity distribution h_j convolved through the corresponding COMPTEL PSF R_{ij}. From this two-component model a mock dataset is created by means of a random number generator which returns random deviates n_i, drawn from a Poisson distribution of means e_i. Hence the mock data are possible realizations of COMPTEL measurements throughout observation periods 0.1 - 522.5 as if the 1.8 MeV intensity would follow the selected model intensity distribution. It should be made clear, however, that while such simulations can serve to study the behavior of the reconstruction algorithm under realistic conditions, they have only limited use for the assessment of systematic uncertainties.

For the example shown in Fig. 3.1 an exponential disk of radial scale length $R = 4.5$ kpc and scale height $z = 90$ pc was chosen for the assumed ^{26}Al distribution, the galactic ^{26}Al mass was fixed to 3 $M_\odot$. It will turn out in Section 4.4.1 that such a distribution represents a good first order approximation of the observed 1.8 MeV emission, although the total galactic ^{26}Al mass might be lower than the assumed 3 $M_\odot$. The corresponding model intensity distribution h_j is shown in the upper left panel of Fig. 3.1 in Hammer-

Aitoff projection.[9] The remaining panels show the reconstructed intensity distribution f_j^r after iterations $r = 1$, 3, 5, 10, 20, 30, and 35. For the initial intensity estimate f_j^0, a grey image of natural initial intensity $m_{\rm nat} = 7.8\,10^{-5}$ph cm^{-2}s^{-1}sr^{-1} was chosen.

Figure 3.1 demonstrates the general properties of the R-L algorithm as anticipated in Section 3.2.1. In the first iterations, the large scale structures of the intensity distribution are worked out. Data space deviations n_i/e_i^r on a length scale small compared to the COMPTEL PSF are smeared out by the backprojection (Eq. (3.12)) into the image space, leading to relatively smooth reconstructions for the first iterations. In particular, the latitude distribution of the reconstructed emission is much broader than for the model. Nevertheless, already the first iterations show a considerable amount of structure which is inconsistent with the smooth model intensity distribution. With successive iterations the reconstructions become more and more lumpy and the rather smooth emission breaks up into many artificial 'point sources'. Additionally, spurious 'point sources' become visible throughout the entire sky – a clear sign of fitting statistical fluctuations (i.e., overfit).

The shape of the reconstructed intensity distribution is illustrated in Fig. 3.2 by plotting longitude profiles (intensity distribution integrated over latitudes $-30° < b < 30°$) and latitude profiles (intensity distribution integrated over longitudes $0° \leq l < 360°$) of the R-L iterations. This clearly demonstrates the smoothness of early iterations and the break-up into emission peaks at late iterations. Additionally it can be seen that the algorithm is not capable of reproducing the correct narrow latitude distribution of the model.

The reconstruction of an alternative mock dataset is shown in Fig. 3.3. As a model intensity distribution, 20 point sources each of flux $3\,10^{-5}$ ph cm^{-2}s^{-1} were distributed throughout the sky. Together with the exponential disk simulation the point sources form two extreme examples: while in the first case a totally smooth emission distribution is assumed, the second example treats the case of a pure point source distribution. The reality may lie between these two examples.

Similarly to the exponential disk simulation, the first iterations show a rather smooth emission distribution. The point sources are spread out to extended spots due to the smoothing property of the backprojection. With proceeding iterations the emission spots become more and more concentrated, but spurious sources show up which are indistinguishable from the true point sources (e.g., at $(l, b) \approx (260°, 40°)$ and $(l, b) \approx (290°, -80°)$).

3.2.2.2 The Stopping Criterion

Figure 3.1 and Fig. 3.3 clearly show the effect of overfit in late iterations (cf. Section 3.2.1). For the exponential disk simulation the initially diffuse emission breaks

[9] All-sky maps will be shown throughout this work in Hammer-Aitoff projection with the galactic center at the center of the map and longitudes increasing towards the left.

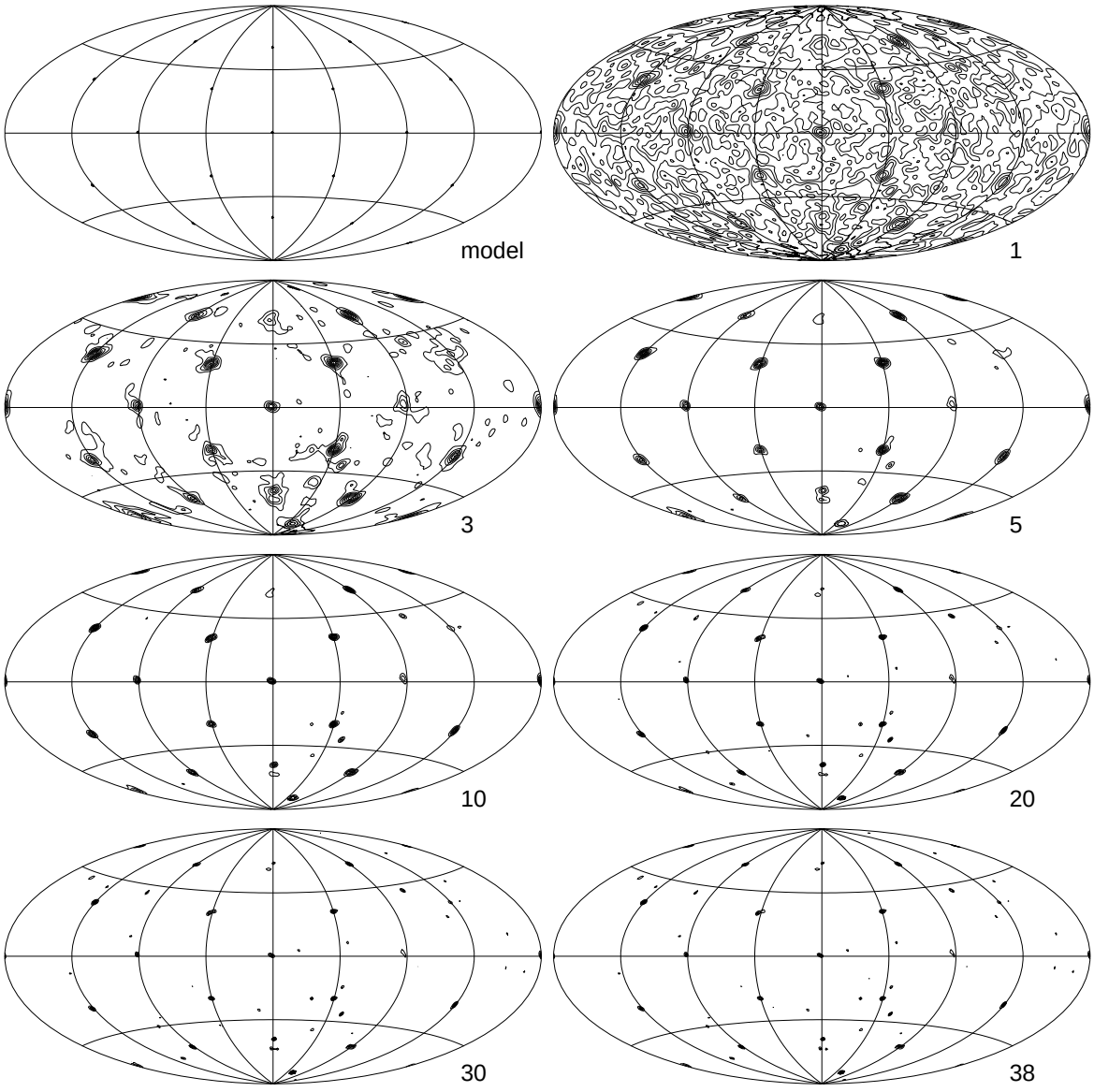

Fig. 3.3: *R-L reconstruction of 20 simulated point sources of flux* $3\,10^{-5}$ *ph* $cm^{-2}s^{-1}$ *at the positions* $(l, b) = (0°, 0°)$, $(90°, 0°)$, $(180°, 0°)$, $(270°, 0°)$, $(45°, \pm 30°)$, $(135°, \pm 30°)$, $(225°, \pm 30°)$, $(315°, \pm 30°)$, $(0°, \pm 60°)$, $(90°, \pm 60°)$, $(180°, \pm 60°)$, *and* $(270°, \pm 60°)$. *The model intensity distribution is shown in the upper left panel, the other panels show the reconstructed intensity distribution* f_j^r *after iterations* $r = 1, 3, 5, 10, 20, 30,$ *and* 38. *A grey image of natural initial intensity* $m_{\rm nat} = 4.0\,10^{-5} ph\ cm^{-2}s^{-1}sr^{-1}$ *was used as initial estimate.*

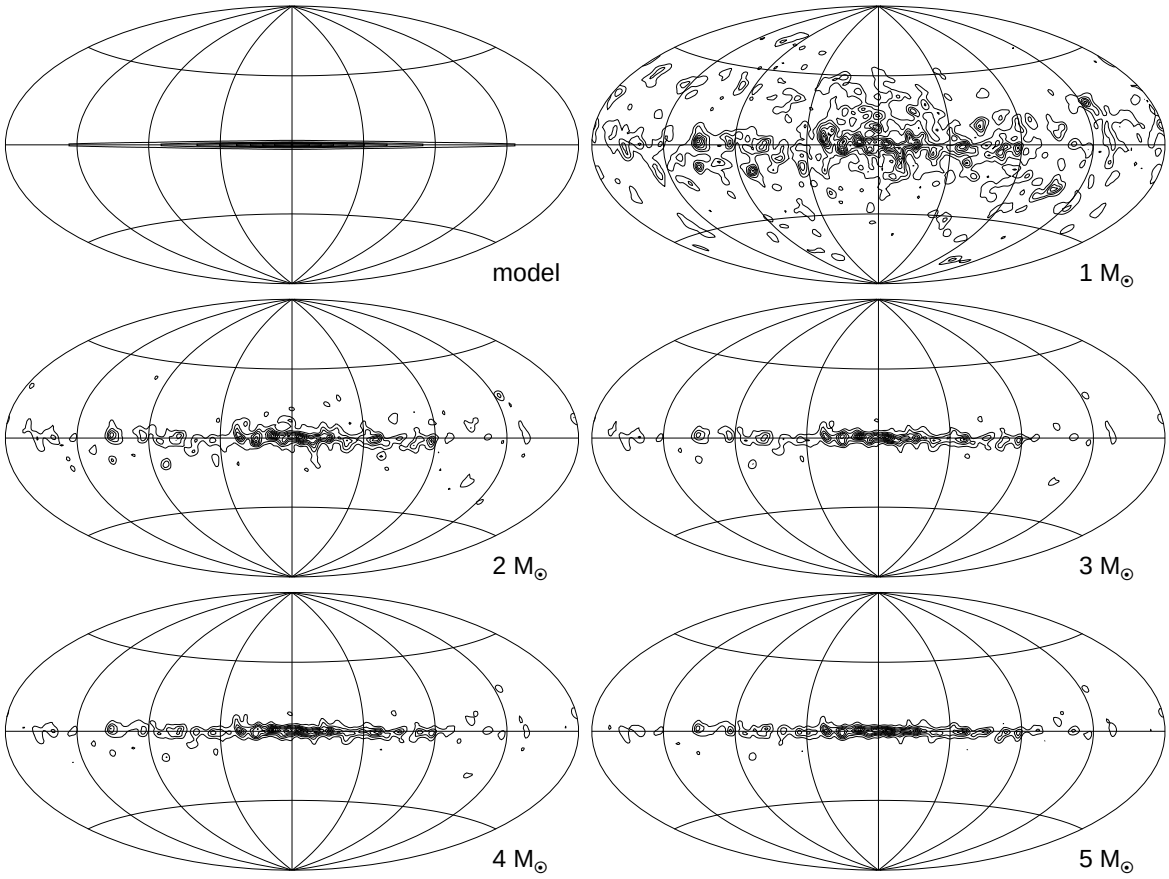

Fig. 3.4: *Optimum exponential disk reconstruction for the parameters* $R_0 = 4.5$ kpc, $z_0 = 90$ pc, *and galactic* ^{26}Al *masses of 1, 2, 3, 4, and 5* $M_\odot$ *as derived from distance measure* A_r. *The optimum iterations are 4, 6, 7, 9, and 11, the best angles* A_r *are 55°, 43°, 37°, 34°, and 30°, respectively (cf. Appendix A.2).*

up into many individual point-like sources situated mainly in regions of high 1.8 MeV intensity. Additional sources at intermediate and high latitudes occur in regions where no signal was present in the model. From the comparison of the late iterations with the model it is clear that these late iterations are a poor description of reality and may be misleading, since they suggest many individual sources where only diffuse emission is present.

The general approach to avoid spurious sources is to stop the iterations before overfit becomes important. A comparison of Fig. 3.1 with Fig. 3.3 shows, however, that the optimum iteration might depend on the underlying intensity distribution: while for diffuse emission overfit sets in very early (about iteration 5) it becomes important for the point source simulation only at late iterations ($r > 20$). Figure 3.2 demonstrates that stopping the R-L reconstruction after few iterations in order to avoid the overfit leads, however, to considerable overestimation of the width of the latitude distribution. Additionally, the flux of emission features is systematically underestimated for early

iterations (cf. Section 3.2.2.3). Hence it becomes clear that the choice of an 'optimum' iteration depends on the question in mind and that there can not be a single optimum iteration.

The question of the optimum iteration for the overall intensity distribution, the longitude, and the latitude profiles is addressed in Appendix A.2 by means of 'distance measures'. It will be shown that for the exponential disk simulation, the overall shape of the reconstructed intensity distribution comes closest to that of the model between iterations 5-8. Comparison with Fig. 3.1 confirms that these iterations provide a rather good compromise between smoothness and latitude extension of the emission distribution. While earlier iterations are smoother, they obey a much wider latitude distribution, hence lie far away from the model. In later iterations the latitude distribution is well represented, but the reconstructions are much more structured than the model. It turned out that the minimum in the distance measures at early iterations is a fundamental property for the reconstruction of diffuse emission. In this sense there exists an optimum iteration for the reconstruction of diffuse emission. Figure 3.4 demonstrates that the optimum iteration slightly increases with galactic ^{26}Al mass if the natural initial intensity is chosen for the reconstruction. For an expected galactic ^{26}Al mass between $2-3$ M$_\odot$ the minimum occurs around iteration $6-7$. Figure 3.4 also shows that with increasing ^{26}Al mass the optimum reconstruction comes closer to the model distribution. This reflects the increase of signal-to-noise ratio with total galactic ^{26}Al mass which reduces the statistical uncertainty in the reconstruction.

3.2.2.3 Flux Calibration

Fluxes Φ of emission features can be derived from the reconstructed intensity map f_j by integrating the intensities over the region of interest using

$$\Phi = \sum_{j \in \text{region}} f_j \Omega_j, \tag{3.17}$$

where

$$\Omega_j = \frac{4\pi}{360} \Delta l \sin\left(\frac{\Delta b}{2}\right) \cos(b) \tag{3.18}$$

is the solid angle of image pixel f_j at galactic latitude b, and Δl and Δb are the pixel sizes in galactic longitude and latitude in degrees, respectively (throughout this work $\Delta l = \Delta b = 1°$). If the reconstructed flux for emission features of the simulations is plotted versus R-L iterations, however, it turns out that for early iterations the flux is systematically underestimated (cf. Fig. 3.5). Depending on the size of the integration region and the amount of flux, a convergent flux level is only reached after $\sim 5 - 15$ iterations. Additionally, the point source simulations show a systematic underestimation of the reconstructed flux with respect to the flux in the intensity model.

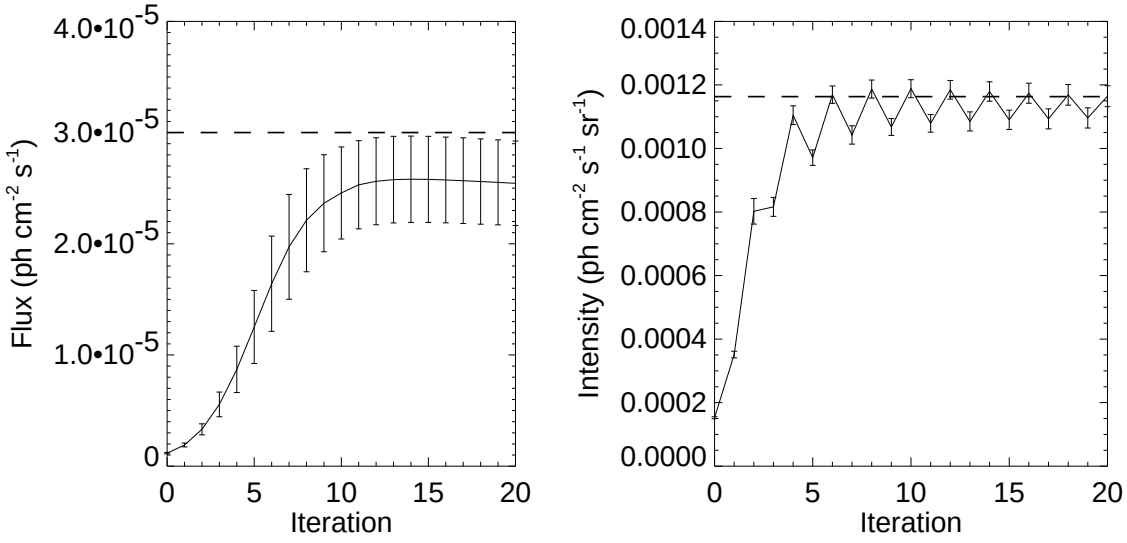

Fig. 3.5: *Reconstructed flux versus iteration for point source (left) and exponential disk simulations (right). For the point source simulation, the reconstructed flux within a radius of 5° around the point source at the galactic center is shown. For the exponential disk simulation, the mean intensity within a box $|l| \leq 30°$ and $|b| \leq 10°$ is given. Error bars were derived from the variance of the reconstructed fluxes for 10 independent mock datasets.*

To allow a reliable flux determination from R-L reconstructions despite these systematic biases, a careful flux calibration has been done by means of numerous point source and exponential disk simulations (cf. Appendix A.3). The picture which arises from this calibration can be summarized as follows: For small integration regions, fluxes are underestimated by a constant amount, independent of the source flux and source position. For the special case of point sources the relation

$$\Phi_{true} = \Phi + 5.6\,10^{-6}\text{ph cm}^{-2}\text{s}^{-1} \tag{3.19}$$

was derived, where Φ is the reconstructed point source flux for late ($\gtrsim 20$) iterations and Φ_{true} is the true flux from the point source. For extended emission features, the underestimation becomes smaller with increasing size of the integration region, falling rapidly below the statistical uncertainties. Reconstructed fluxes can even exceed the true value by a small amount if very extended integration regions are chosen. This is due to the remaining background intensity in empty regions of the sky which, in turn, is due to the positivity constraint of sky intensities. For large integration regions this background intensity can considerably contribute to the emission. To correct for this overestimation, mean background intensities should be estimated from regions free of 1.8 MeV emission (i.e., from high galactic latitudes), which are then subtracted from the reconstructed source fluxes.

3.2.2.4 The 1.8 MeV All-Sky Map

After the study of the properties of the Richardson-Lucy iterations, we are sufficiently prepared to apply the algorithm to the observed 1.8 MeV data. For real data, however, there is an additional parameter which has to be estimated along with the intensity distribution: the background model scaling factor β. Although the background model was carefully normalized on a single observation basis under consideration of time variabilities of major background components, its absolute normalization is uncertain to about $\sim 1\%$ (cf. Section 2.5.2.2). In a full Bayesian analysis of the problem, β would be a 'nuisance parameter', i.e. a parameter which influences the reconstruction but which is not needed to be explicitly known. The full posterior probability distribution for the image reconstruction problem is then given by

$$P(H,\beta|DI) = P(H,\beta|I)\frac{P(D|H,\beta,I)}{P(D|I)}. \tag{3.20}$$

The nuisance parameter is eliminated from the posterior by integration over β[10]

$$P(H|D,I) = \int P(H,\beta|DI)d\beta, \tag{3.21}$$

which leads to the so called *marginal posterior probability density* $P(H|D,I)$. Instead of performing this integration explicitly, which is a rather costly procedure, β may be estimated along with the reconstruction of the intensity distribution. The marginal posterior is then approximated by constraining β to the optimum value $\hat{\beta}$ in the posterior probability:

$$P(H|D,I) \approx P(H,\hat{\beta}|D,I). \tag{3.22}$$

Bretthorst (1988) demonstrates that at least for 'data-dominated' problems, this approximation leads to results which are essentially similar to those obtained by marginalization (i.e., explicit integration).

To estimate the optimum $\hat{\beta}$, R-L reconstructions were performed for a grid of background scaling factors β between 0.96 and 1.00. Figure 3.6 shows the log-likelihood ratios $-2\ln L_0$ which were obtained after R-L iterations $r = 5$, 10, 20, 30, and 40 as a function of β. As anticipated in Section 2.5.2.2, R-L reconstructions prefer a background scaling factor well below unity. For an optimum $\hat{\beta} = 0.985$, as suggested by Fig. 3.6, the total number of 1.8 MeV source photons amounts to 96 420, which is about a factor of ~ 2.5 higher than the number expected from the background model normalization. Note, however, that estimation of the background scaling factor from late R-L iterations might lead to an underestimation of the instrumental background. From the simulations discussed above it is known that the intensity distribution partially fits statistical fluctuations, hence some background counts are assigned to the source component. Consequently, sky intensities may be slightly overestimated, and the subtraction of a mean background intensity for flux determination is mandatory.

[10]This method of the elimination of nuisance parameters is also referred to as marginalization.

Finally, 8 iterations of the R-L reconstruction of the all-sky dataset are shown in Fig. 3.7 for the optimum background scaling factor $\hat{\beta} = 0.985$. Comparison with Fig. 3.1 demonstrates that the reconstruction of real data shows a similar behavior as the exponential disk simulation. The first iterations provide rather smooth intensity distributions where the emission is spread over a considerable range of latitudes. With proceeding iterations the emission becomes more and more concentrated on the plane with a substantial increase of structure along the plane. When the iterations are proceeded until convergence is reached (at about iteration $r \gtrsim 170$), the structure breaks-up into individual emission peaks of pixel size.

Fig. 3.6: *Likelihood ratio as function of background model scaling factor β for R-L iterations $r = 5$, 10, 20, 30, and 40.*

From the simulations, it is known that the lumpiness of the reconstructed images is not necessarily real, but is probably mostly due to statistical fluctuations of the background component. Some features, however, are hard to reconcile with statistical fluctuations only. The most prominent of these is an extended emission feature in Cygnus centered at $l \approx 80°$ with an apparent diameter of about 15°. It is clearly separated from the galactic ridge emission by a region extending from $l = 50 - 70°$ which is relatively free of 1.8 MeV emission. Such segregation is not seen in any of the exponential disk simulations (cf. Fig. 3.4) which strongly suggests that this feature is a real deviation from a smooth emission profile. Another, even more extended feature is seen in the Auriga-Camelopardalis-Perseus (ACP) region ($l = 130 - 180°$) which obeys a latitude extension of $b \gtrsim \pm 30°$. Similar features are seen at high galactic latitudes around $b \sim -60°$. Such extended low intensity features might be due to extended diffuse 1.8 MeV emission, but they could also arise from systematic uncertainties in the background model. The question on the reality of these features will be addressed in Section 4.4.4.

To emphasize the differences between the reconstruction of a smooth exponential disk model and the real 1.8 MeV intensity distribution, longitude and latitude profiles of the R-L reconstructions are shown in Fig. 3.8. The statistical uncertainties in the longitude profile may be inferred from a comparison of the exponential disk reconstruction (dashed line) with the corresponding model (solid line). Firstly, it is obvious that the 1.8 MeV intensity profile is rather similar to the exponential disk reconstruction, within the statistical uncertainties. Secondly, most structures seen in the 1.8 MeV emission profile seem to be consistent with statistical deviations from a

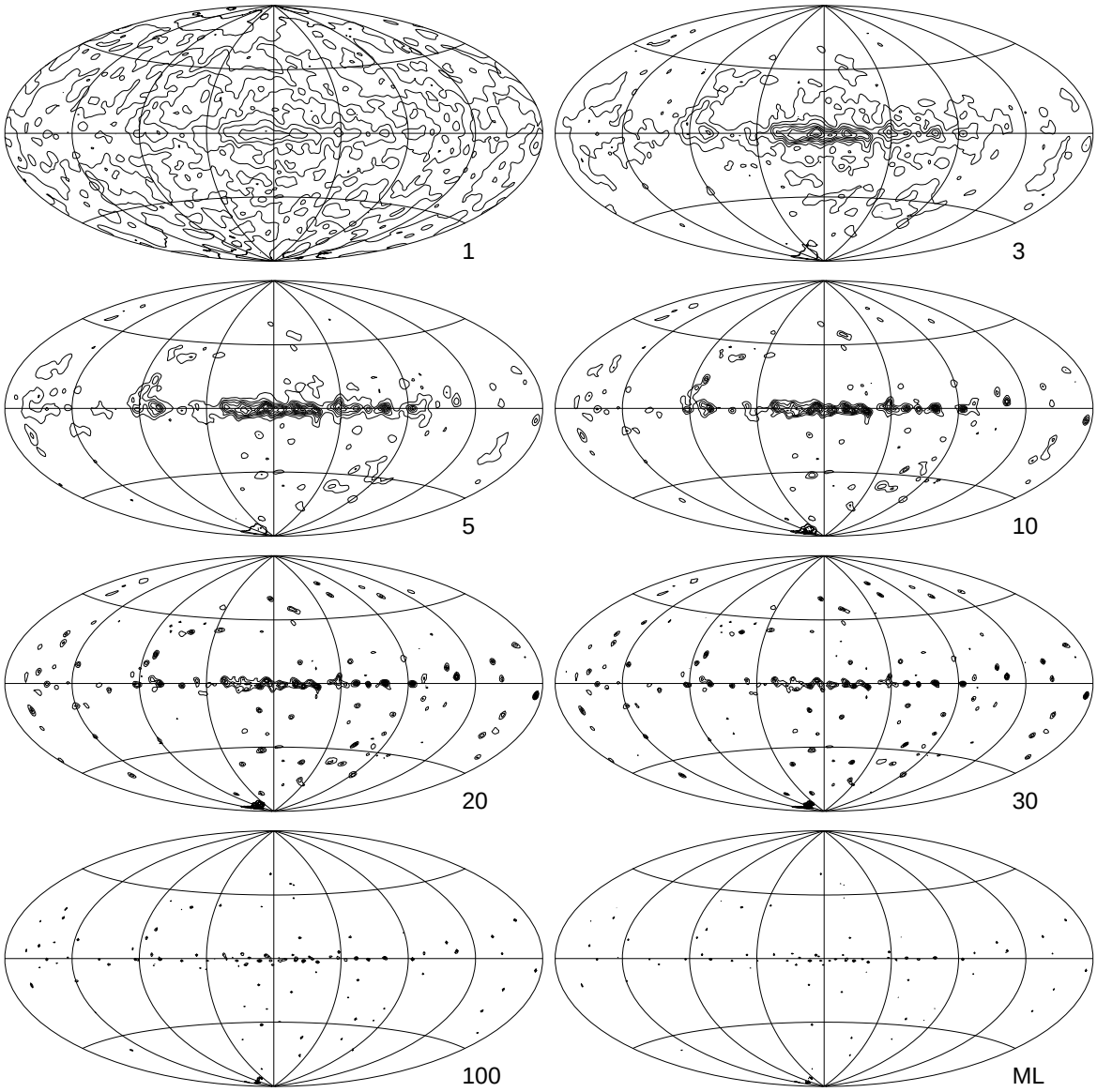

Fig. 3.7: *Iterations r = 1, 3, 5, 10, 20, 30, 100, and ML (170) of the R-L reconstruction of COMPTEL 1.8 MeV all-sky data from observation periods 0.1 - 522.5. A grey image of natural initial intensity $m_{\rm nat} = 9.86\,10^{-5} ph\ cm^{-2}s^{-1}sr^{-1}$ was used as initial estimate. The background scaling factor was fixed to $\beta = 0.985$.*

smooth intensity profile. However, the Cygnus and ACP features are clearly not consistent with a smooth intensity distribution. Additionally, the emission gap at $l \approx 325°$ together with the neighboring peaks at $l \approx 332°$ and $l \approx 317°$ are more pronounced than any fluctuation in the exponential disk simulation. When the view is focussed on emission close to the galactic plane (mid panel of Fig. 3.8) an additional significant deviation appears in Carina ($l \approx 286°$). Furthermore, small deviations from a smooth emission profile may exist near the galactic center ($|l| = \pm 10°$) and in the Vela region

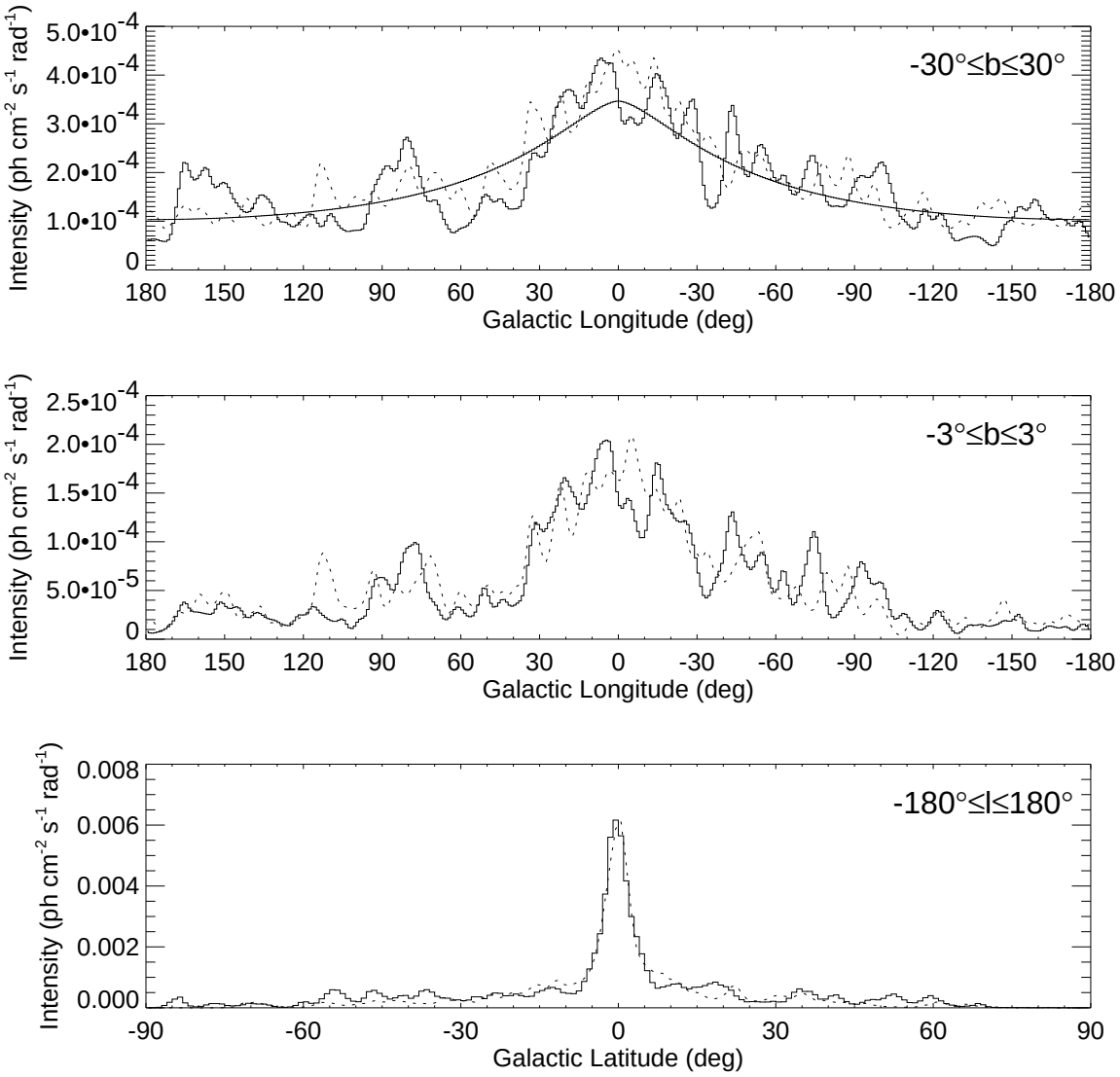

Fig. 3.8: Top: *Longitude profile of R-L iteration* 4 *of the 1.8 MeV all-sky data reconstruction for a latitude integration range of* $|b| \leq 30°$ *(histogram). The solid line indicates the longitude profile of an exponential disk model of* 2 $M_\odot$ *total galactic* ^{26}Al *mass and scale parameters* $R_0 = 4.5$ kpc *and* $z_0 = 90$ pc. *The dashed line shows the corresponding R-L reconstruction for iteration* 4. Middle: *Same as top but for a latitude integration range of* $|b| \leq 3°$. Bottom: *Latitude profile of R-L iteration* 20 *of the 1.8 MeV all-sky data reconstruction (solid). The dashed line indicates the latitude profile of iteration* 20 *of the the exponential disk simulation.*

Name	Position		Radius	All-sky flux	Phase 1 & 2 flux
	l	b		10^{-5} ph cm^{-2}s^{-1}	10^{-5} ph cm^{-2}s^{-1}
Center	2°	-1.5°	5°	4.2 ± 0.6	5.5 ± 1.1
Arm 2 (north)	32°	1°	6°	4.0 ± 0.7	5.1 ± 0.8
Cygnus	80.5°	-3.5°	4°	1.3 ± 0.5	1.8 ± 0.8
Perseus	168.5°	-6.5°	6°	0.8 ± 0.7	3.5 ± 1.1
Puppis	238°	5°	5°	2.0 ± 0.6	4.1 ± 1.2
Vela	265.5°	-1.5°	6°	2.8 ± 0.7	4.6 ± 1.2
Carina	286.5°	0.5°	3°	2.9 ± 0.5	3.3 ± 0.8
Arm 2 (south)	309.5°	-1.5°	4°	1.8 ± 0.5	3.0 ± 1.0
Scorpius	345.5°	-0.5°	6°	5.4 ± 0.7	6.3 ± 1.3

Name	Region				All-sky flux	Phase 1 & 2 flux
	l_{min}	l_{max}	b_{min}	b_{max}	10^{-5} ph cm^{-2}s^{-1}	10^{-5} ph cm^{-2}s^{-1}
Galaxy	-180°	180°	-10°	10°	59.5 ± 4.9	96.5 ± 4.1
Northern Galaxy	0°	180°	-10°	10°	31.2 ± 3.5	44.6 ± 3.7
Southern Galaxy	-180°	0°	-10°	10°	28.3 ± 3.5	51.9 ± 3.6
Inner Galaxy	-35°	38°	-10°	10°	27.4 ± 2.3	35.4 ± 2.6
Arm 2 (south)	302°	325°	-10°	10°	6.2 ± 1.3	8.0 ± 1.4
Central radian	-30°	30°	-10°	10°	24.7 ± 2.1	31.0 ± 2.6
Galactic center	-10°	15°	-10°	10°	10.1 ± 1.4	14.4 ± 1.7
Cygnus	72°	100°	-10°	8°	5.6 ± 1.4	7.4 ± 1.2

Table 3.1: *Fluxes determined from late iterations of the R-L reconstruction. The upper table gives the fluxes for emission spots similar to Tab. 4.2 of Knödlseder (1994). A constant flux of* $6\,10^{-6}$ *ph cm*$^{-2}$*s*$^{-1}$ *was added to account for the flux underestimation for point sources. The lower table gives the fluxes for extended emission regions similar to Tab. 4.3 of Knödlseder (1994). Errors were determined from the error calibration formula.*

($l = 258° - 275°$), their astrophysical reality, however, may be hard to confirm on the base of these profiles, only.

1.8 MeV fluxes of prominent features in the all-sky map are compiled in Tab. 3.1. The integration regions were chosen similar to Knödlseder (1994) to allow for comparison with earlier flux estimates based on observations of mission phases 1 & 2. Knödlseder (1994) derived the fluxes from late iterations of a maximum entropy reconstruction which was calibrated by point source simulations. This calibration resulted in a global flux correction factor of 1.26 which was applied to all flux determinations. Statistical errors were derived using a 'Bootstrap analysis' (e.g., Knödlseder 1994). In this work, relative fluxes are determined from late R-L iterations by subtracting an average background intensity derived for galactic latitudes $|b| \geq 30°$. This 'contrast method' accounts for the probable underestimation of the background scaling factor which may lead to an overestimation of 1.8 MeV intensities. The resulting fluxes, however, are conservative in that they are based on the assumption that no 1.8 MeV emission exists

Fig. 3.9: *Flux error versus integration area as determined from exponential disk simulations. The solid line represent the function* $\Delta\Phi = 3.2\,10^{-5}\sqrt{\Omega} + 0.1\,10^{-5}$ *ph cm*$^{-2}$*s*$^{-1}$ *which was fitted to the data points.*

above $|b| \geq 30°$. To account for the point source flux underestimation, a constant flux of $6\,10^{-6}$ ph cm^{-2}s^{-1} was added to all point sources fluxes.

Statistical flux errors are estimated for all features using the empirical formula

$$\Delta\Phi = 3.3\,10^{-5}\sqrt{\Omega} + 1\,10^{-6} \quad \text{ph cm}^{-2}\text{s}^{-1}, \tag{3.23}$$

where Ω is the solid angle (in sr) of the integration region. This formula is based on the assumptions that, firstly, the flux error is independent of the source flux and source position, and secondly, it scales with the square root of the solid angle of the emission feature. The assumptions are justified by the fact that COMPTEL data are background dominated, hence the statistical uncertainties arise mainly from the background fluctuations and are consequently independent of the source flux. The number of background counts that interfere with the emission feature is, to first order, proportional to the solid angle covered by the feature, hence the error related to the background fluctuations should scale with the square root of the solid angle. The background intensity is, also to first order, a function of the sky exposure which varies by about a factor of ~ 2.4 over the sky and a factor of ~ 1.6 along the galactic plane. Consequently, the variation of the statistical flux error over the sky should only amount to a factor of ~ 1.5, while along the galactic plane it should amount to a factor of ~ 1.3.

The parameters of the functional relationship were determined using 10 independent mock datasets, created for an exponential disk model with radial scale length $R_0 = 4.5$ kpc, scale height $z_0 = 90$ pc, and galactic ^{26}Al mass of 3 $M_\odot$. The variation of the reconstructed fluxes of the samples was determined for rectangular integration boxes centered on $b = 0°$, using a fixed latitude box size of $|b| \leq 3°$, and various longitude box sizes between $\Delta l = 1°$ and $360°$. Possible exposure effects were removed by averaging the flux variation for various box positions along $b = 0°$. The resulting data points are plotted in Fig. 3.9 together with the best fitting relation Eq. (3.23). Obviously, Eq. (3.23) provides a satisfactory fit to the observed statistical uncertainties, at least for $\Omega \gtrsim 5\,10^{-3}$ sr. This solid angle, however, corresponds to the COMPTEL FWHM

beam area, hence it is not surprising that the empirical formula breaks down below this limit. To avoid an underestimation of the statistical error for small integration areas, a lower limit of $\Delta\Phi \geq 5\,10^{-6}$ ph cm^{-2}s^{-1} is applied which was determined from point source simulations and independently from model fitting (cf. Chapter 4).

A comparison of the derived fluxes for the all-sky map with values determined for phase 1 & 2 data shows that the all-sky values lie systematically below the phase 1 & 2 fluxes. For the point sources, this can partially be explained by the fact that the definition of the integration regions were based on the locations of emission spots in the galactic plane skymap of observation phases 1 & 2. Besides the Perseus feature at $(l, b) = (168.5°, -6.5°)$, all emission spots of Tab. 3.1 are still present in the all-sky map, but their position may have shifted by $1° - 2°$, or their shape may have changed, under the influence of the changed statistical 'configuration' of the instrumental background. Examples of such 'morphing' emission are the Vela feature which moved from negative latitudes towards the galactic plane, or the 'Arm 2 (south)' spot which shifted from $l \approx 309°$ to $l \approx 305°$. Such drastic changes between two different datasets may indicate that the underlying 1.8 MeV emission obeys a rather smooth intensity distribution which is very susceptible to the particular 'configuration' of the instrumental background. In contrast, the Carina feature, which was shown to be the most point-like emission in the COMPTEL 1.8 MeV sky (Knödlseder *et al.* 1996a), and the Scorpius feature are the most stable spots, which obey only a marginal flux change of $\sim 10\%$. Probably, these features arise from rather point-like emission regions which are much less affected by the background fluctuations.

The shifts of emission spots, however, can not account for the systematically lower fluxes which are found for extended emission regions. In particular, the estimated 1.8 MeV emission of the entire Galaxy is $\sim 40\%$ lower for the all-sky data than for the phase 1 & 2 galactic plane map. Recall that the use of the contrast method for the flux determination places only a lower limit on the all-sky fluxes. Without background subtraction, the total galactic plane flux ($|b| \leq 10°$) amounts to $(75.9 \pm 4.9)\,10^{-5}$ ph cm^{-2}s^{-1}, which is still about $\sim 20\%$ lower than the phase 1 & 2 value. Another interesting difference between the all-sky and the phase 1 & 2 results is the north-south asymmetry in the emission profile: While phase 1 & 2 data suggested about $\sim 40\%$ more flux from the southern hemisphere than from the northern hemisphere (Diehl *et al.* 1995b), the all-sky skymaps provide about equal fluxes from both hemispheres (indeed, there is marginally more flux from the northern hemisphere).

3.3 The Maximum Entropy Method

Alternatively to the Richardson-Lucy algorithm, COMPTEL intensity distributions are usually reconstructed by means of the maximum entropy method. Historically, this was the first method which was applied to deconvolve COMPTEL data, and summaries about the algorithm implementation and its performance are given by Strong

et al. (1992) and Strong (1995). The application to 1.8 MeV gamma-ray line data of COMPTEL observation phases 1 & 2 was studied in detail by Knödlseder (1994). Recently, Oberlack (1997) extended this study to the all-sky dataset of observation periods 0.1 - 522.5 (the dataset which is also used in this work), hence I will content here with summarizing his results.

3.3.1 MEMSYS2 algorithm

Finding the maximum entropy (ME) MAP consists in finding the maximum of the ME posterior probability distribution. For practical reasons it is convenient to maximize not $P(H|D,I)$ but its logarithm which is proportional to

$$Q(f_j,\alpha) = \ln P(n_i|f_j) + \alpha S(f_j). \tag{3.24}$$

Q can be identified as a Lagrange function with α being the Lagrange multiplier. Another interpretation is that of an image quality function which is composed of a likelihood term $\ln P(n_i|f_j)$ which 'attracts' the image to the data and an entropy term $S(f_j)$ which penalizes this attraction and pulls the image towards the expected magnitudes m_j. The image is reconstructed by solving the M non-linear equations

$$\frac{\partial Q}{\partial f_j} = 0, \qquad 1 \leq j \leq M \tag{3.25}$$

for a given α, which dictates an iterative solution procedure. The solutions $f_j(\alpha)$ define the *trade-off curve* or *maximum entropy trajectory*, starting from $f_j \equiv m_j$ at $\alpha \to \infty$ and ending at the maximum likelihood solution when $\alpha = 0$. Note that even with very small α the entropy still has a regularizing effect since the logarithmic factor imposes positivity.

The software package generally used to reconstruct maximum entropy images from COMPTEL data is MEMSYS2 from Maximum Entropy Data Consultants Ltd of Cambridge, England, which was modified to treat Poisson statistics and to incorporate a background model (see Skilling and Bryan (1984) for a detailed description of the MEMSYS2 algorithm and Appendix B for the modifications). Starting from a grey image ($f_j = m$) with global maximum entropy ($\alpha \to \infty$, $S = 0$), the algorithm iteratively reduces α and determines the maximum entropy image for this particular weighting parameter. Consequently, the images of the first iterations are dominated by the image entropy, hence are close to the expected image, while for late iterations the reconstruction converges to the (positively constrained) maximum likelihood solution.

Similar to the R-L reconstruction there is no criterion where to stop the MEMSYS2 iterations. Originally, Skilling and Bryan (1984) proposed to stop the iterations when the Chi-squared test statistics become identical to the number of image pixels. For COMPTEL data, for which the measurement errors are properly described by the

Poisson statistics, no analog constraint exists, since the formal expectation over possible data of the log-likelihood is not constant, but varies with the reconstructed image. Consequently, similar to the R-L algorithm, the optimum iteration has to be determined by means of simulations.

3.3.2 Properties of MEMSYS2

Figure 3.10 and Fig. 3.11 show some iterations for the reconstruction of the same exponential disk and point source models which were used to study the R-L iterations (cf. Figs. 3.1 and 3.3). Obviously, both the ME and the R-L algorithms lead to essentially identical reconstructions. A detailed inspection of the reconstructed intensity distributions reveals, however, that the maximum entropy images are slightly 'cleaner', showing less noise at high galactic latitudes. As well, the 'speed' with which the iterations proceed is different: while maximum entropy iterates somewhat slower for early iterations (indicated by the flatter reconstruction for iterations 1 and 3), it becomes faster for later iterations.

To determine the optimum ME reconstruction, Oberlack (1997) applied an entropy based distance measure to the exponential disk simulation, and concluded that iterations $6-8$ are the reconstructions which come closest to the model intensity distribution. The use of the distance measures defined in Appendix A.2 gives similar results, with optima between iterations $7-8$. In an earlier study of the application of MEM to the reconstruction of 1.8 MeV galactic plane data from mission phases 1 & 2, Knödlseder (1994) noted a systematic flux underestimation for point sources of about a factor 1.26. Using independent point source mock datasets for various flux levels, Oberlack (1997) showed that point source emission is also systematically underestimated in the ME reconstructions of the all-sky data. From the fit of a scaling factor to the data points, he derived a typical flux underestimation of $\sim 10\%$. The use of a linear regression fit, however, provides a better description of the flux underestimation which is consistent with the relation Eq. (3.19) obtained for the R-L algorithm.

Oberlack (1997) also studied the ME flux recovery for diffuse emission by means of exponential disk mock datasets, derived for scale parameters $R_0 = 4.0$ kpc, $z_0 = 180$ pc, and total mass 3 $M_\odot$. For the relatively small integration regions of $10° \times 10°$ he used, he found a $\sim 30\%$ flux underestimation. Nevertheless, the total flux in the reconstructed skymaps is about $\sim 10\% - 20\%$ higher than in the model, which indicates that the flux underestimation can not be universal and probably depends on the size of the integration region. Indeed, for exponential disk simulations which were performed for this work ($R_0 = 4.5$ kpc, $z_0 = 90$ pc, total mass 3 $M_\odot$), it was found that the flux underestimation disappears for larger integration regions, similar to reconstructions with the R-L algorithm.

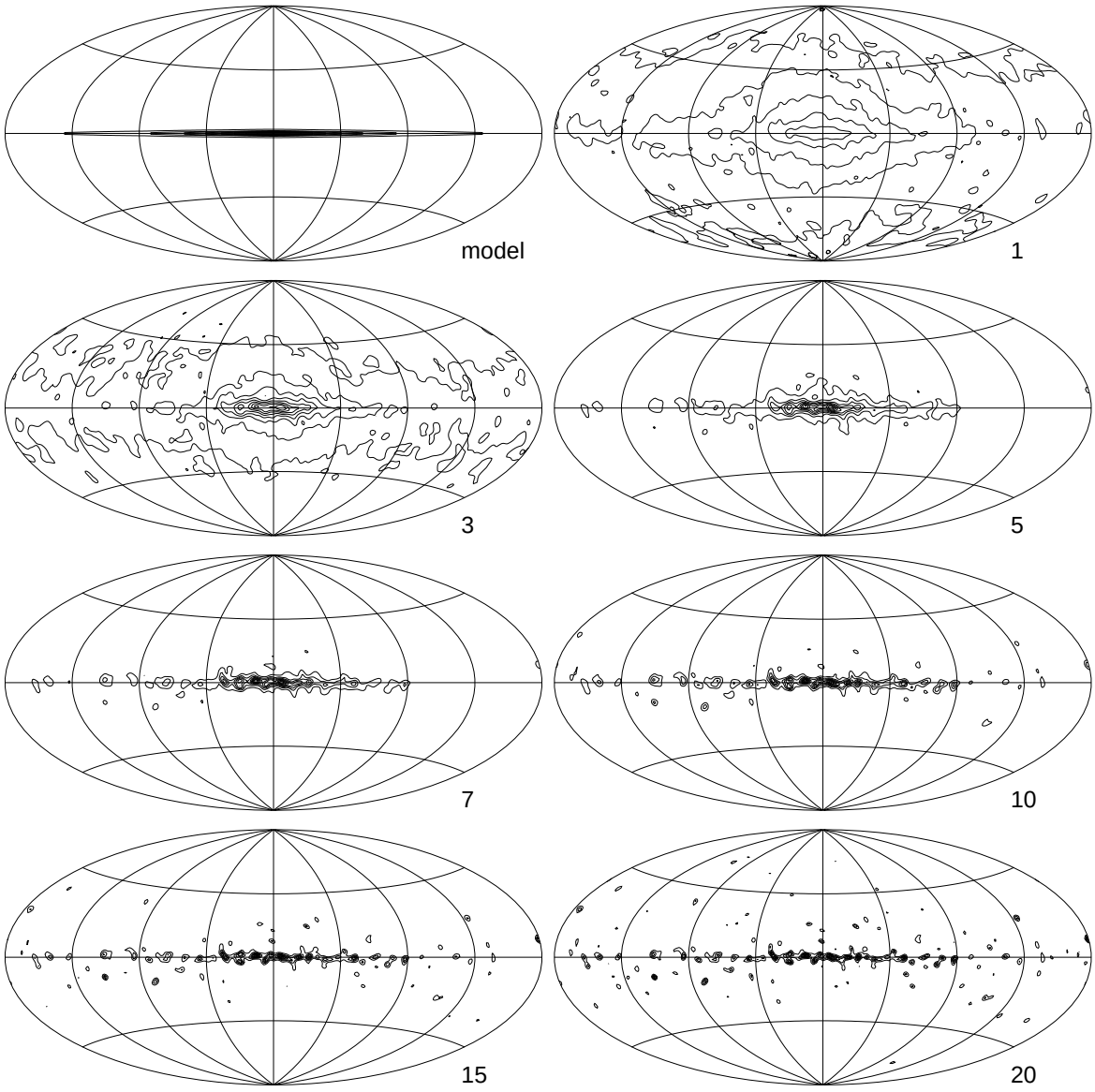

Fig. 3.10: *MEMSYS2 iterations for the reconstruction of simulated exponential disk data, derived for scale parameters* $R_0 = 4.5$ kpc, $z_0 = 90$ pc, *and a total galactic* ^{26}Al *mass of* $3\,M_\odot$. *The intensity distribution of the exponential disk model is shown in the upper left panel; the other panels show the reconstructed intensity distribution* f_j^r *after iterations* $r = 1$, *3*, *5*, *7*, *10*, *15*, and *20*. *A grey image of natural initial intensity* $m_{\rm nat} = 7.8\,10^{-5} ph\ cm^{-2}s^{-1}sr^{-1}$ *was used as initial estimate.*

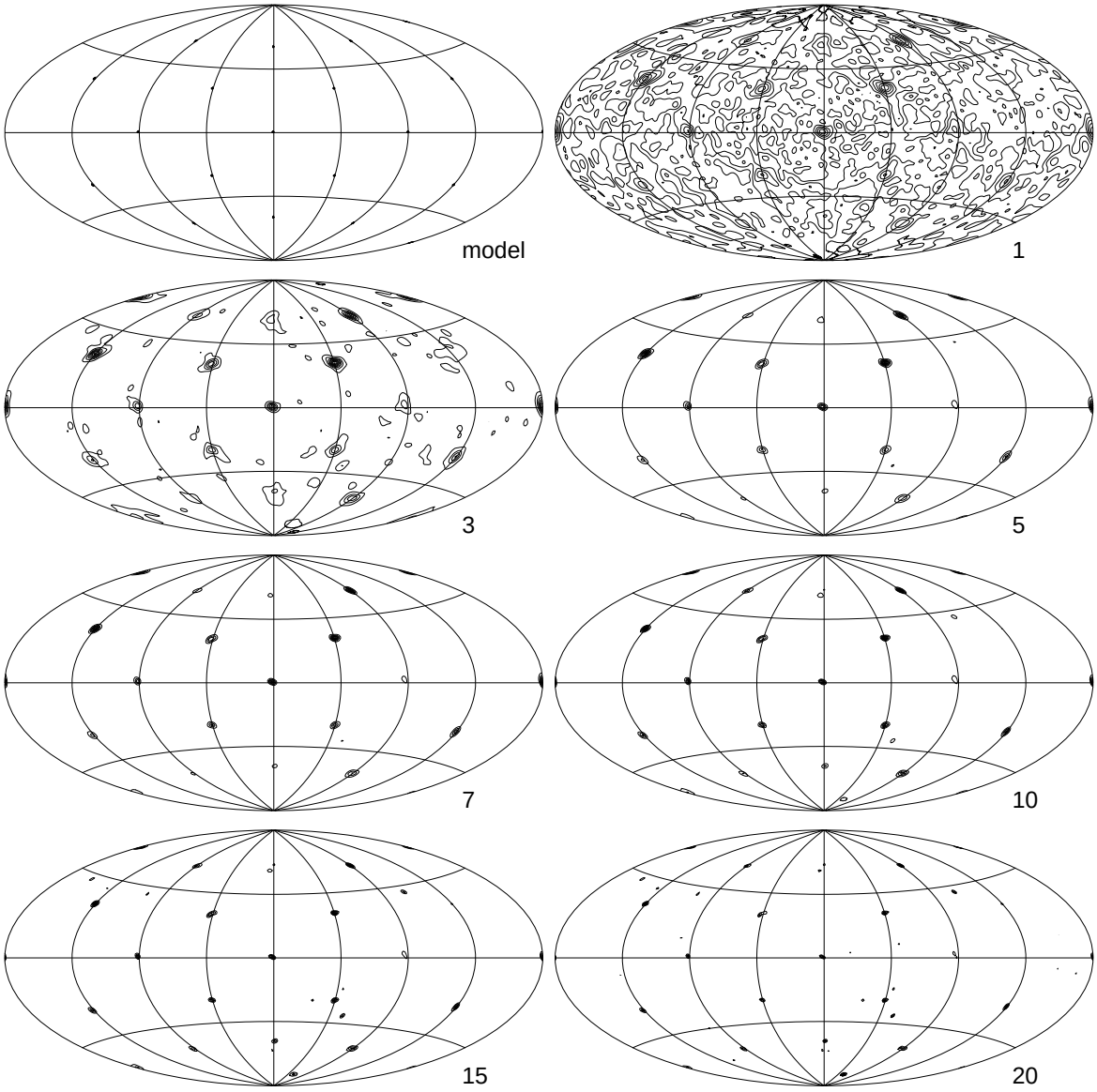

Fig. 3.11: *MEMSYS2 iterations for the reconstruction of 20 simulated point sources. The intensity distribution of the point source model is shown in the upper left panel; the other panels show the reconstructed intensity distribution f_j^r after iterations $r = 1, 3, 5, 7, 10, 15,$ and 20. A grey image of natural initial intensity $m_{\rm nat} = 4.0\,10^{-5} ph\ cm^{-2}s^{-1}sr^{-1}$ was used as initial estimate.*

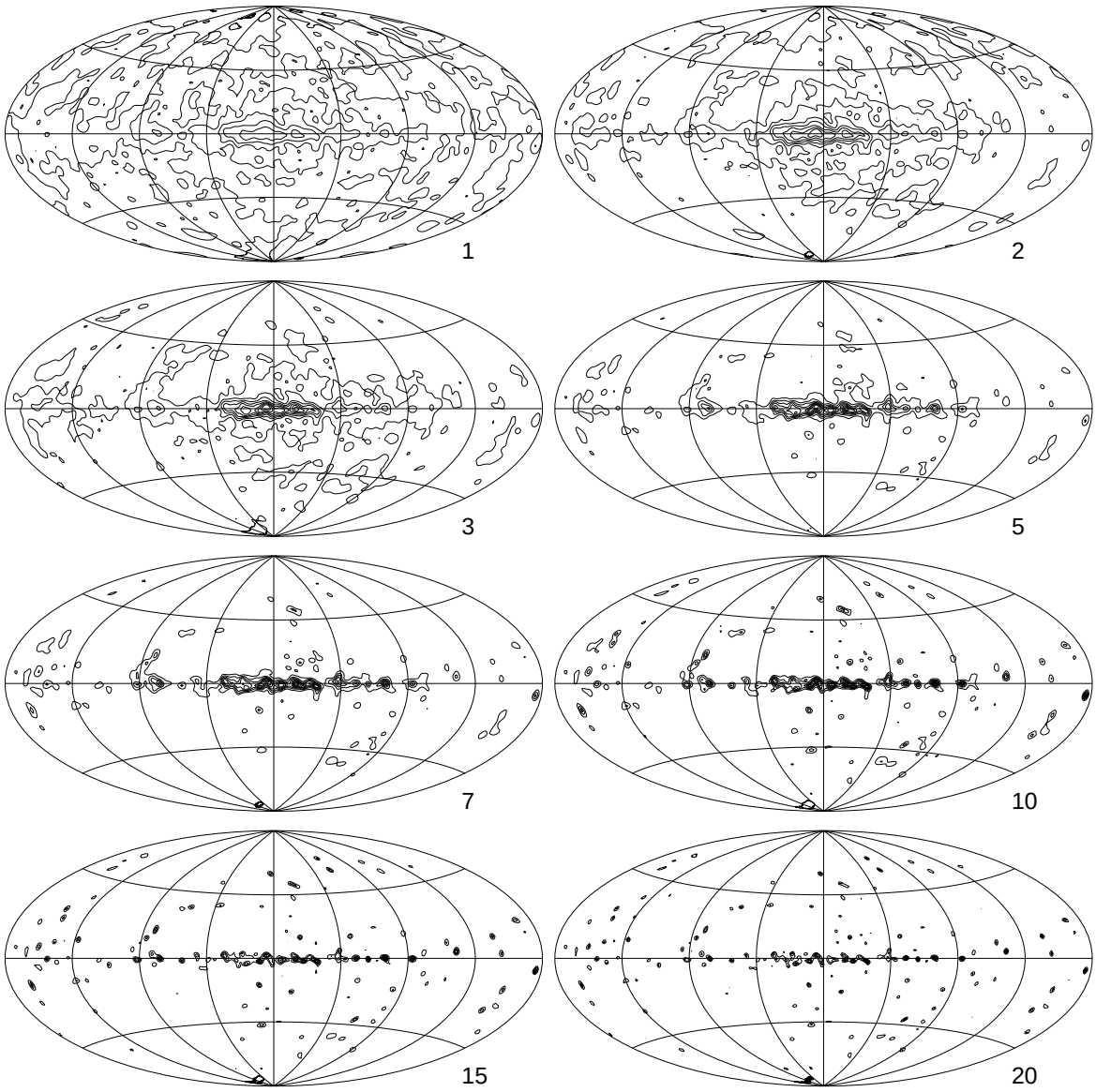

Fig. 3.12: *Iterations $r = 1, 2, 3, 5, 7, 10, 15,$ and 20 of the ME reconstruction of COMPTEL 1.8 MeV all-sky data from observation periods 0.1 - 522.5. A grey image of natural initial intensity $m_{\mathrm{nat}} = 9.86\,10^{-5} ph\ cm^{-2}s^{-1}sr^{-1}$ was used as initial estimate. The background scaling factor was fixed to $\beta = 0.985$.*

3.3.3 The 1.8 MeV All-Sky Map

Figure 3.12 shows 8 iterations of the ME reconstruction of the COMPTEL 1.8 MeV all-sky dataset. Similar to the R-L reconstruction, the background scaling factor was fixed to $\beta = 0.985$ which results in a natural initial intensity of $m_{\rm nat} = 9.86\,10^{-5}$ ph cm^{-2}s^{-1}sr^{-1}. As anticipated by the above simulations, the ME reconstruction is almost identical to the R-L images (cf. Fig. 3.7).

1.8 MeV fluxes of prominent emission features are compiled in Tab. 3.2, which is similar to Tab. 3.1 for the R-L reconstruction. For comparison, two different flux calibrations were applied: For column A, only the point source fluxes were corrected by adding a constant flux of $6\,10^{-6}$ ph cm^{-2}s^{-1} to the values derived from the skymap. For column B, the flux calibration of Oberlack (1997) was applied, which consists in a correction factor of 1.1 for point sources and of 1.3 for extended emission. All fluxes are relative fluxes with respect to an average background intensity, derived from galactic latitudes $|b| \geq 30°$.

Comparison with Tab. 3.1 shows that ME fluxes derived with calibration A (i.e. the same calibration as for the R-L reconstruction) are essentially identical to R-L fluxes. Calibration B, proposed by Oberlack (1997), leads to smaller fluxes for the point sources, while extended emission fluxes are always higher by a factor of 1.3. Nevertheless, these fluxes are also lower than those obtained by Knödlseder (1994) from mission phases 1 & 2 data for the galactic plane.

At first glance it may seem surprising that maximum likelihood and maximum entropy reconstructions lead to almost identical results. This, however, is only true from the practical viewpoint. From the theoretical viewpoint, the maximum likelihood result is fundamentally different from the maximum entropy result. Recall, that the difference in both methods was the choice of the priors Eqs. (3.8) and (3.9). The flat maximum likelihood prior is completely ignorant about the magnitude of the sky intensity, and the solution of the ML problem is found by iterating the R-L algorithm to $r \to \infty$ (which is almost reached after ≈ 170 accelerated iterations for the 1.8 MeV all-sky dataset). Consequently, the maximum likelihood solution to the image reconstruction problem is not one of the early iterations in Fig. 3.7, but the spotty image denoted by ML. In contrast, the maximum entropy prior reflects the prior knowledge about the expected magnitudes m_j in the image pixels, and the expected spread $1/\alpha$ around these magnitudes. Thus, each panel of Fig. 3.12 is indeed a maximum entropy reconstruction, which only differ in the prior knowledge assumed about the spread of the true sky-intensities f_j around the reference magnitudes m_j. Notice, that α is a prior parameter which, from a Bayesian standpoint, should be determined before the reconstruction.[11]

[11] For the 'optimum' ME iterations $6-8$, α lies in the range of $(3.5-6.3)\,10^5$ which corresponds to an entropy prior width of $(1.3-1.7)\,10^{-5}$ ph cm^{-2}s^{-1}sr^{-1} (1σ). For the 'late' iteration 20, $\alpha = 6.7\,10^4$ corresponding to an uncertainty about the intensity magnitude of $\sigma = 3.8\,10^{-5}$ ph cm^{-2}s^{-1}sr^{-1}.

Name	Position		Radius	All-sky flux^A	All-sky flux^B
	l	b		10^{-5} ph cm^{-2}s^{-1}	10^{-5} ph cm^{-2}s^{-1}
Center	2°	-1.5°	5°	4.1 ± 0.6	3.8 ± 0.7
Arm 2 (north)	32°	1°	6°	3.9 ± 0.7	3.7 ± 0.8
Cygnus	80.5°	-3.5°	4°	1.4 ± 0.5	0.6 ± 0.6
Perseus	168.5°	-6.5°	6°	0.8 ± 0.7	0.3 ± 0.8
Puppis	238°	5°	5°	1.8 ± 0.6	1.3 ± 0.7
Vela	265.5°	-1.5°	6°	2.6 ± 0.7	2.2 ± 0.8
Carina	286.5°	0.5°	3°	2.7 ± 0.5	2.3 ± 0.6
Arm 2 (south)	309.5°	-1.5°	4°	1.8 ± 0.5	1.4 ± 0.6
Scorpius	345.5°	-0.5°	6°	5.0 ± 0.7	4.8 ± 0.8

Name	Region				All-sky flux^A	All-sky flux^B
	l_{min}	l_{max}	b_{min}	b_{max}	10^{-5} ph cm^{-2}s^{-1}	10^{-5} ph cm^{-2}s^{-1}
Galaxy	-180°	180°	-10°	10°	58.6 ± 4.9	76.2 ± 6.4
Northern Galaxy	0°	180°	-10°	10°	31.0 ± 3.5	40.3 ± 4.8
Southern Galaxy	-180°	0°	-10°	10°	27.6 ± 3.5	35.8 ± 4.6
Inner Galaxy	-35°	38°	-10°	10°	27.2 ± 2.3	35.4 ± 3.0
Arm 2 (south)	302°	325°	-10°	10°	6.1 ± 1.3	7.9 ± 1.7
Central radian	-30°	30°	-10°	10°	24.5 ± 2.1	31.9 ± 2.7
Galactic center	-10°	15°	-10°	10°	10.1 ± 1.4	13.1 ± 1.8
Cygnus	72°	100°	-10°	8°	5.5 ± 1.4	7.1 ± 1.8

Table 3.2: *Fluxes determined from late iterations of the ME reconstruction. The upper table gives the fluxes for emission spots similar to Tab. 4.2 of Knödlseder (1994). For flux column A, a constant flux of* $6\,10^{-6}$ *ph cm*$^{-2}$*s*$^{-1}$ *was added to account for the flux underestimation as derived in this work. For flux column B, skymap fluxes were multiplied by* 1.1 *according to Oberlack (1997). The lower table gives the fluxes for extended emission regions similar to Tab. 4.3 of Knödlseder (1994). Fluxes in column A are uncorrected, as suggested by this work, fluxes in column B are multiplied by* 1.3 *as proposed by Oberlack (1997). Errors were determined from the error calibration formula, and multiplied by* 1.1 *(point sources) and* 1.3 *(extended emission regions) for flux column B.*

Nevertheless, it is still surprising that the R-L scheme leads to the same reconstructions as the maximum entropy algorithm. On the one hand, this is probably due to the fact that both algorithms are subjected to the positivity constraint of sky intensities, which is indeed already a very powerful constraint for image reconstruction (Lucy 1994). One the other hand, the reconstruction of the COMPTEL 1.8 MeV all-sky map is a so-called data-dominated or large-sample case (e.g., Bontekoe *et al.* 1994, Kass and Wasserman 1994) for which the choice of the prior makes little difference, unless it is smooth in the region of high likelihood. Indeed, differences between MEM and R-L occur at the edges of the image if intensity distributions for single observation periods are reconstructed. In that case, the data only weakly constrain the intensity distribution at the edges, and the ME intensities tend towards the expected magnitudes m_j in

these regions. In contrast, the R-L solution is unconstrained at the edges of the image which leads to prominent artefacts. For the all-sky dataset this problem doesn't occur since it has, by definition, no edges.

3.4 Towards a Multiresolution Approach

3.4.1 Problems of 'Classical' Reconstruction Algorithms

The fundamental problem of the Richardson-Lucy and the maximum entropy reconstructions is overfit of the data, which already sets in during the first iterations, and manifests itself as spurious sources and lumpy emission in the images. The origin of the spurious sources in 'classical' image reconstructions is readily understood (Puetter 1995a): there are too many free parameters in the reconstruction to describe the data. For the 1.8 MeV all-sky map there are 64 800 image pixels which can take mutually independent values, while the total number of 1.8 MeV photons amounts to only $40\,000 - 70\,000$. In particular, the total significance of the 1.8 MeV emission along the galactic plane is $\sim 30\sigma$ which allows only about 100 significant (3σ) image pixels. Consequently, the lumpiness of the images may be reduced by increasing the size of the image pixels and thus decreasing the number of free parameters. Alternatively, a correlation function may be introduced which relates the intensities in neighboring image pixels, and thus effectively reduces the number of independent parameters. Algorithms that deal with such correlations are generally referred to as *multiresolution*, *multiscale*, or *multi-channel* algorithms (e.g., Starck and Murtagh 1994, Bontekoe *et al.* 1994, Puetter 1995b, Murtagh *et al.* 1995, Pantin and Starck 1996, and references therein).

In the Bayesian framework, multiresolution image reconstruction means that the model M is included in the inference procedure (recall that for MEM and R-L the model was a priori fixed to a $1^\circ \times 1^\circ$ pixel grid). Using Bayes' theorem, the multiresolution posterior is given by

$$P(H, M|D, I) = P(M|I)P(H|M, I)\frac{P(D|H, M, I)}{P(D|I)} \tag{3.26}$$

where $P(M|I)$ is the prior probability distribution of the model and $P(H|M, I)$ is the prior probability of the hypothesis given model M (H now simply represents the intensity distribution $\mathbf{f}$). Since for image reconstruction we are not interested in the model, M is a 'nuisance' parameter and the posterior for the intensity distribution may be derived by marginalization

$$P(H|D, I) = \int P(H, M|D, I)dM. \tag{3.27}$$

In the following sections an attempt at multiresolution reconstruction of COMPTEL 1.8 MeV all-sky data is described which is based on noise suppression using the wavelet

transform. This work was recently begun in collaboration with Dave Dixon from the University of California, Riverside (UCR). It is an attempt insofar as not all problems have been solved yet, but first results are both promising and encouraging.

3.4.2 The Wavelet Transform

The fundamental idea behind wavelets is to analyze according to scale, i.e., to decompose a signal into different frequency bands and then study each component with a resolution matched to its scale. This idea is not new, and is most familiar to any physicist as Fourier transformation. The major difference between Fourier analysis and wavelets is that while Fourier basis functions are localized in frequency but not in space, wavelets are local in both frequency (via dilations) and in space (via translations). This localization is an advantage in many cases, since it makes many functions 'sparse' when transformed into the wavelet domain. This sparseness, in turn, results in a number of useful applications such as data compression, detecting features in images, and removing noise. In this work, this last capability will be exploited.

Like the fast Fourier transformation, the discrete wavelet transformation (DWT) is a fast, linear operation, which is invertible and orthogonal. But unlike the Fourier transformation, which is uniquely defined by sines and cosines, there is an infinite set of possible basis functions of wavelets. Roughly, the different sets of wavelets make different trade-offs between how compactly they are localized in space and how smooth they are.

Noise can be removed from a dataset using the *wavelet shrinkage* or *thresholding* methods (Donoho 1993). The wavelet decomposition of a dataset uses filters that act as *averaging* filters and others that produce *details* (Vidaković and Müller 1991). If the details are small, they might be omitted without substantially affecting the main features of the dataset. The idea of *thresholding*, then, is to set to zero all coefficients that are less than a particular threshold. The *thresholded* coefficients are then used in an inverse wavelet transformation to reconstruct the dataset. In contrast to conventional techniques, de-noising is carried out without smoothing out the sharp structures, and the result is a cleaned-up signal that still shows important details (for a nice example see Graps 1995). Note that this de-noising technique has the desired property of introducing pixel-to-pixel correlations in the image of which the correlation length depends on the amount of structure in the data. Regions of the sky with uniform emission will be represented by few large-scale wavelet coefficients, while the small-scale coefficient are essentially zero in these areas. Vice versa, isolated point sources are represented by few small-scale coefficients, while the large-scale coefficients vanish.

One of the major drawbacks of wavelets is their lack of translation invariance, i.e. the wavelet coefficients vary under translations, dilations, and rotations of the input signal (Simoncelli *et al.* 1992). This results in artefacts in the de-noised dataset, in

particular close to discontinuities. To overcome this problem, Coifman and Donoho (1995) proposed the 'cycle spinning' method which averages over shifted dataset, leading to translation-invariant de-noising. In the following, this method will be used to implement noise suppression in the Richardson-Lucy algorithm.[12]

3.4.3 Modification of the Richardson-Lucy Algorithm

The accelerated R-L iteration is given by

$$f_j^{r+1} = f_j^r + \lambda^r \delta f_j^r \tag{3.28}$$

where

$$\delta f_j^r = f_j^r \left(\frac{\sum_{i=1}^N \frac{n_i}{e_i^r} R_{ij}}{\sum_{i=1}^N R_{ij}} - 1 \right) \tag{3.29}$$

is the additive R-L correction. The wavelet transform provides a convenient tool to extract only the significant structures in the R-L correction. This is done by transforming δf_j^r into the wavelet domain, thresholding of the wavelet coefficients, and back-transformation of the thresholded coefficients into the image space. In compact matrix notation, $\delta \mathbf{f}^r$ is replaced by

$$\delta \mathbf{f}^r \rightarrow \mathbf{W}^T \eta_S \mathbf{W} \delta \mathbf{f}^r, \tag{3.30}$$

where $\mathbf{W}$ is the DWT, and η_S is the thresholding operator. Donoho and Johnstone (1994) propose to use the threshold $\lambda = \sigma\sqrt{2 \log n}$ where n is the sample size, i.e. the number of image pixels, and σ is the scale of the noise at a given resolution level. Thresholding may be done using the *hard threshold*

$$\eta_S(w, \lambda) = \begin{cases} 0 & : \quad w < \lambda \\ w & : \quad w \geq \lambda \end{cases} \tag{3.31}$$

or the *soft threshold*

$$\eta_S(w, \lambda) = \text{sign}(w)(|w| - \lambda)_+ \tag{3.32}$$

(Vidaković and Müller 1991). To determine the noise level σ at each resolution, a mock reconstruction is performed in parallel where n_i in Eq. (3.29) is replaced by a mock dataset created from the actual reconstruction e_i^r by means of a random number generator. The resulting R-L correction thus reflects only statistical noise, and the variance of its wavelet transform is used to determine σ. Note, however, that this is only an approximation of the true noise level which may vary throughout the image

[12]The implementation of the 'à trous' wavelet transform in the Richardson-Lucy algorithm is discussed by Starck and Murtagh (1994). Their approach differs from ours in that they de-noise the residuals in the data space, whereas we aim to de-noise directly the R-L corrections in the image space.

due to variation of the column sums $\sum_{i=1}^{N} R_{ij}$. It turns out that image reconstruction in this approach is most stable when a hierarchical scheme is used, which starts with the recovery of large scale structures, and allows, with proceeding iterations, the smaller scales to come in. In practice this is done by zeroing all wavelet coefficients below a specific resolution scale and, once the R-L iterations have converged, successively lowering the minimum permissible scale.[13]

To illustrate the performance of the modified R-L algorithm, it is applied to the usual mock datasets of an exponential disk of scale parameters $R_0 = 4.5$ kpc, $z_0 = 90$ pc, $M = 3$ $M_\odot$ (cf. Fig. 3.13) and for 20 simulated point sources each of flux $3\,10^{-5}$ ph cm^{-2}s^{-1} (cf. Fig. 3.15). What is shown in the panels are the 'converged'[14] R-L iterations for a given limit in resolution levels using soft thresholding (for these and all other examples given in this work, Daubechies wavelets with 12 coefficients were used (Daubechies 1988)). Panel $(2,2)$, for example, means that only the first two resolution levels of the wavelet decomposition were kept and all other coefficients were set to zero. This means that only large-scale R-L corrections are allowed which results in the reconstruction of very smooth and extended emission features.

The exponential disk simulation demonstrates that the modified R-L algorithm provides much better reconstructions than the standard R-L (Fig. 3.1) or the maximum entropy (Fig. 3.10) method. Using only the largest resolution scales $(2,2)$ and $(3,3)$ results in perfectly symmetric intensity distributions with respect to the galactic center and the galactic plane. However, similar to the early R-L/MEM iterations, the latitude distribution of the reconstructed skymaps is much more extended than that of the model. With the permission of smaller resolution scales, i.e. levels $\geq (4,4)$, statistical fluctuations start to propagate into the reconstructions, leading to 'hot spots' of emission along the galactic plane. Nevertheless, the resulting skymaps are much less lumpy compared to the R-L or MEM reconstructions. Notice that for resolution scales $\gtrsim (6,6)$, the reconstructions converge towards a stable image.

To illustrate how closely the reconstructed intensity distribution comes to the exponential disk model, longitude profiles are shown in Fig. 3.14 for the converged modified

[13]As technical sidemark, note that the discrete wavelet transform is only defined on datasets of length $n = 2^s$, where s is the number of resolution levels. The all-sky image grid of 360×180 pixel is thus rebinned internally into a 512×256 pixel grid, resulting in 9×8 resolution levels. While the normal wavelet transform would map the 512×256 pixels to 512×256 wavelet coefficients, the translation-invariant approach results in an overcomplete representation of 9×8 wavelet planes, each of 512×256 wavelet coefficients. Note that coefficients for the first scale ($s = 1$) are not thresholded since they represent the scaling function, i.e. the DC level of the wavelet decompositions.

[14]Indeed, convergence for the modified R-L iterations is rather slow, despite the algorithm acceleration using the line search. The statistical nature of the determination of the thresholding level λ (by means of simulations) can even lead to a decrease in the likelihood of subsequent iterations, while in the long term the likelihood still increases. Therefore, the R-L iterations are considered as 'converged' if the likelihood-changes over ~ 5 iterations are small ($\lesssim 1$), although full convergence may not have been reached yet. Experience suggests, however, that no surprises are expected for the remaining iterations.

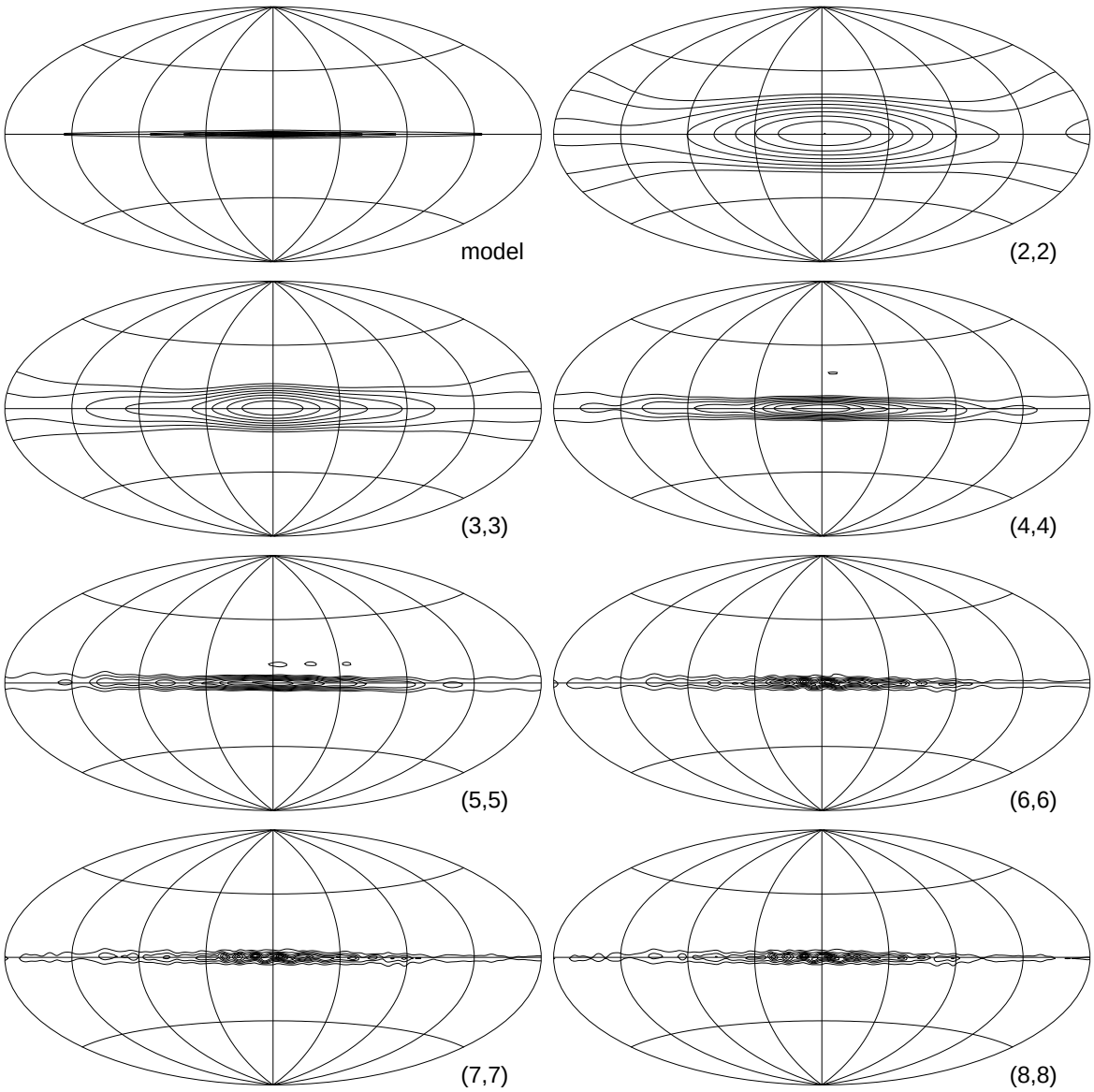

Fig. 3.13: *Multiresolution R-L iterations for the reconstruction of COMPTEL all-sky mock data of an exponential disk model of* $R_0 = 4.5$ kpc, $z_0 = 90$ pc, *and total galactic mass of* 3 $M_\odot$. *The panels show the converged ML solutions for various minimum resolution levels (see text).*

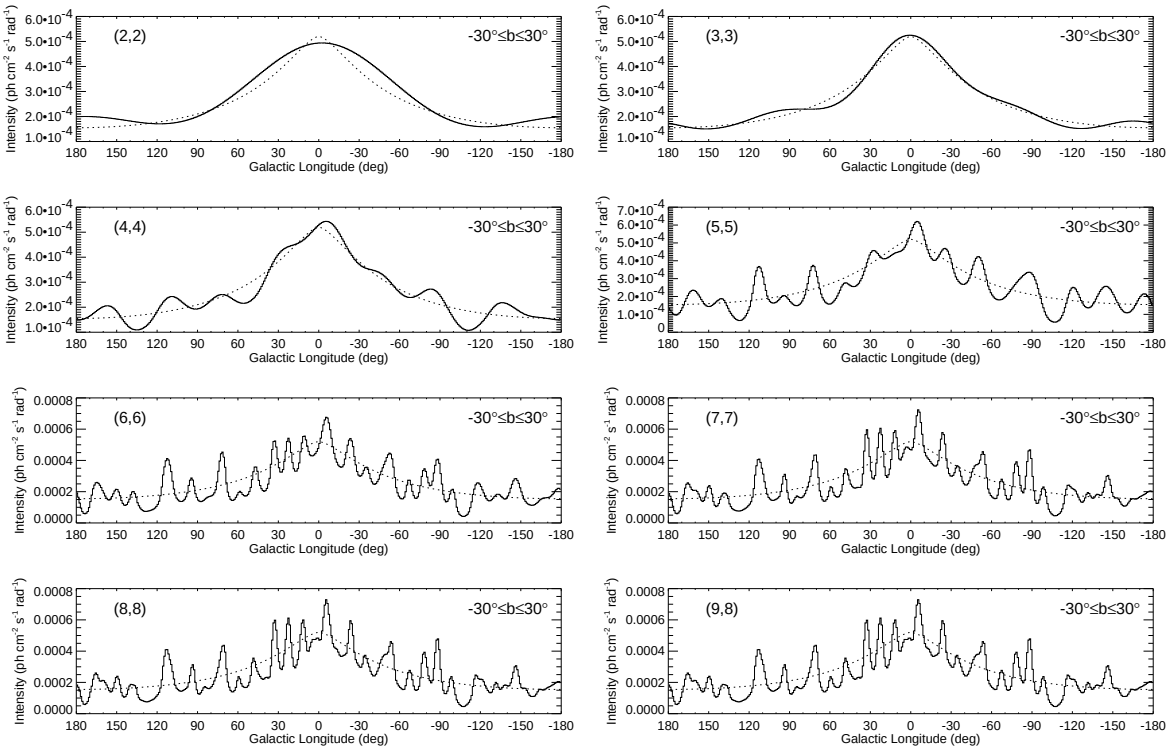

Fig. 3.14: Longitude profiles of the exponential disk reconstructions (cf. Fig. 3.13).

R-L iterations for various resolution levels. These profiles demonstrate another property of the wavelet decomposition: the reconstructed images are composed of 2^{s-1} roughly equidistant 'bumps' in longitude and latitude, where s is the number of resolution levels. For example, the $(2,2)$ profile shows 2 bumps at $l \simeq 0°$ and $l \simeq 180°$, while the $(3,3)$ profile shows 4 bumps at $l \simeq 0°$, 100°, 200°, and 290°.

These 'beating frequencies' can obviously lead to artefacts in the reconstructed intensity distributions. A clear example is the reconstruction of the '20 point sources' mock dataset, shown in Fig. 3.15. Starting from resolution levels $(4,4)$, 'ghost' sources are found half-way between the model sources. Additionally, galactic plane point sources are not recovered by the algorithm. This is due to the fact that for the iterations with the largest resolution scale $(2,2)$, the emission is concentrated in broad bands at galactic latitudes $b \simeq \pm 45°$, leading to a void of emission in the plane (note that the $s = 2$ scale allows only 2 major bumps in latitude). Once the large scale emission structure is established, the following iterations which allow smaller resolution scales are apparently unable to recover the missing point sources. Obviously, the algorithm is trapped in a local minimum of the solution space.

Finally, the modified R-L algorithm is applied to COMPTEL 1.8 MeV all-sky data of observation periods 0.1 - 522.5. Reconstructed intensity distributions are shown for the different resolution levels in Fig. 3.16, corresponding longitude profiles are given in Fig. 3.17. Already from the large resolution scales, some interesting properties of the

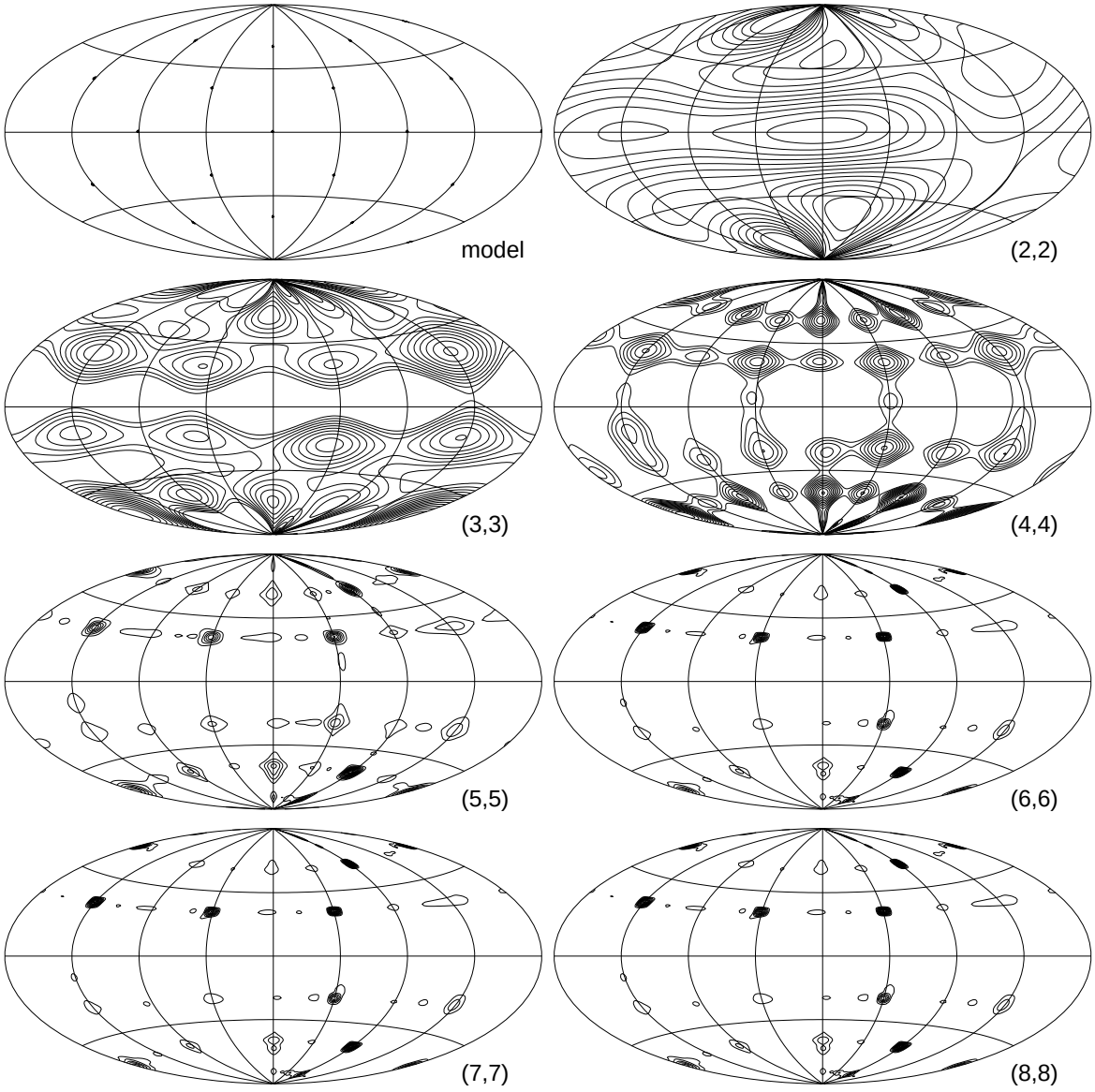

Fig. 3.15: *Multiresolution R-L iterations for the reconstruction of COMPTEL all-sky mock data of 20 point sources of flux* $3\,10^{-5}$ *ph* $cm^{-2}s^{-1}$. *The panels show the converged ML solutions for various minimum resolution levels.*

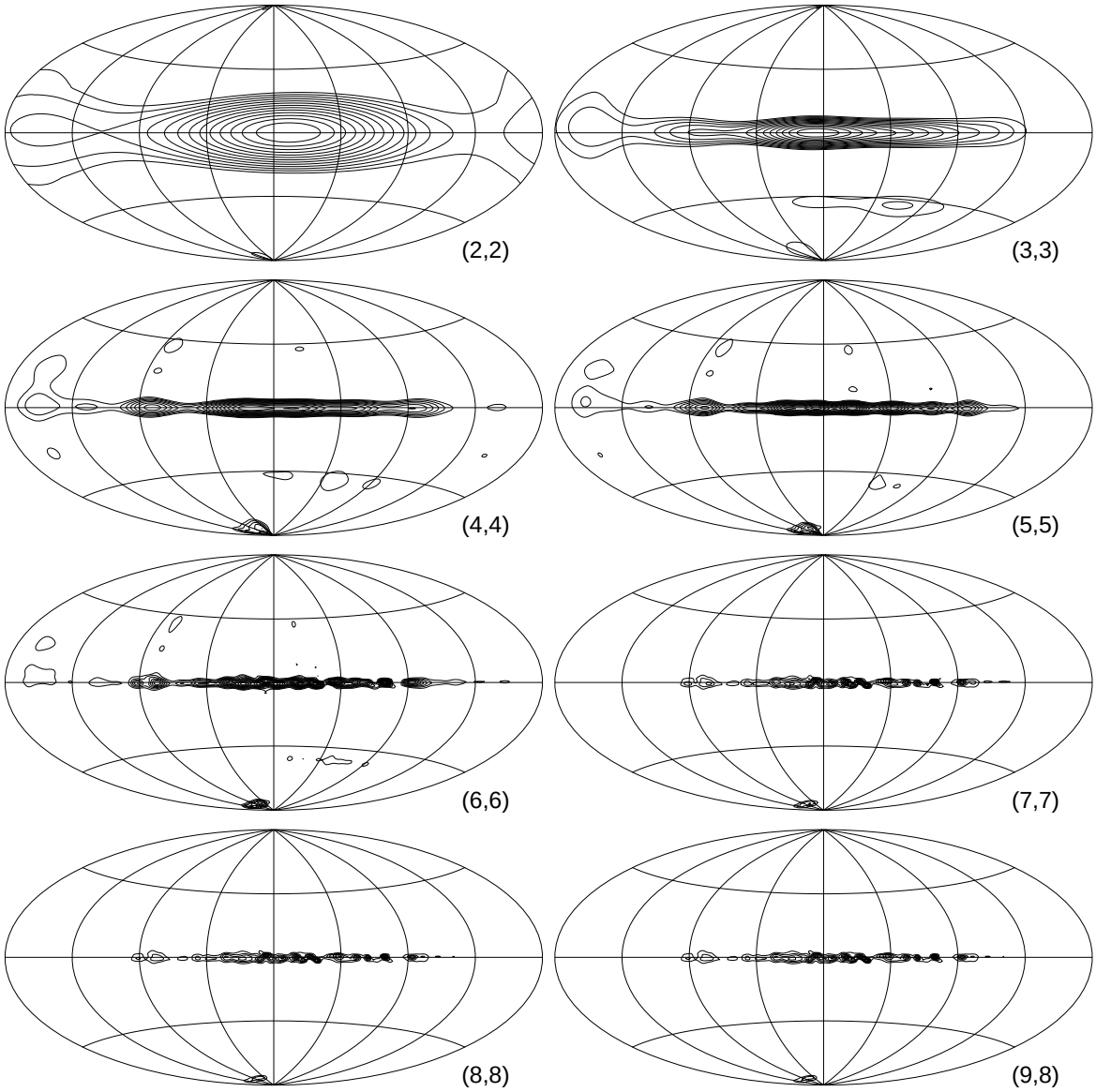

Fig. 3.16: *Multiresolution R-L iterations for the reconstruction of COMPTEL 1.8 MeV all-sky data of observation periods 0.1 - 522.5. In contrast to the R-L or MEM skymaps, a background scaling factor of* $\beta = 1.0$ *was chosen.*

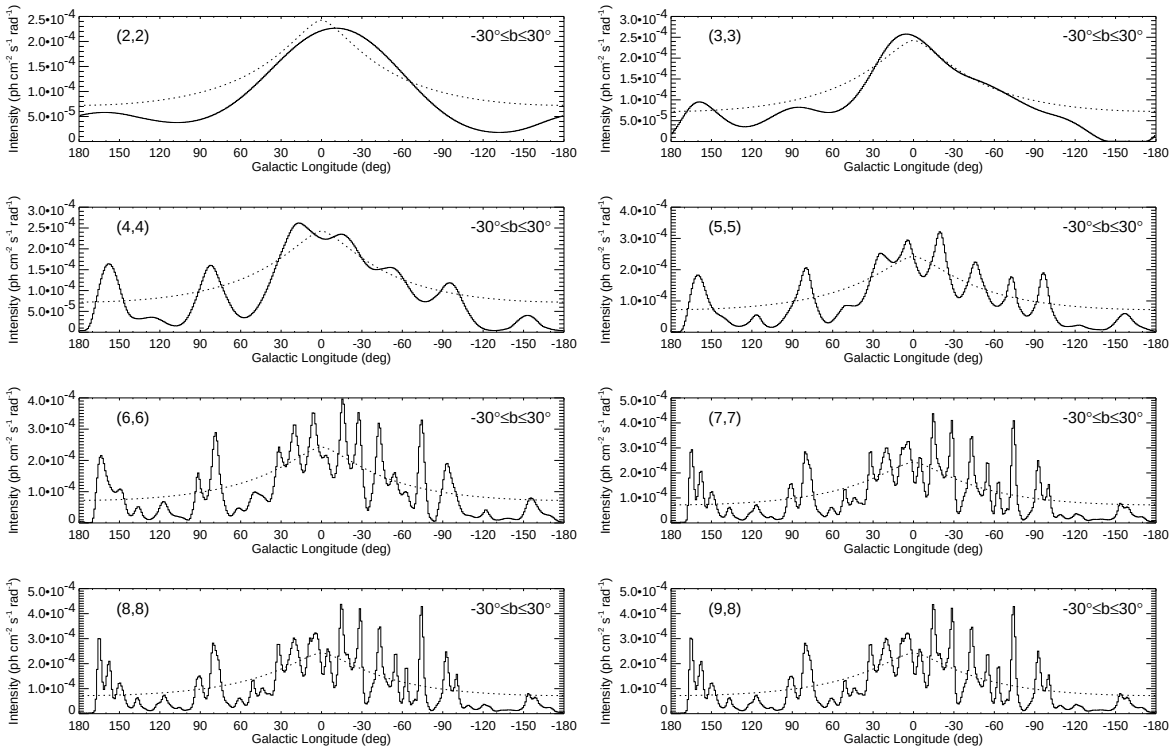

Fig. 3.17: *Longitude profiles of the 1.8 MeV reconstructions (cf. Fig. 3.16). As reference, the longitude profile of an exponential disk model with radial scale length of* $R_0 = 4.5$ kpc *and scale height of* $z_{90} = 90$ pc *is shown as dashed line.*

underlying intensity distribution may be inferred. For example, the $(2,2)$ panel clearly shows an asymmetry of the emission maximum towards negative longitudes. With increasing resolution $(3,3)$, it turns out that the large-scale asymmetry is in reality composed of several small-scale asymmetries: a preference for an offset of the central-radian peak towards positive longitudes, a sharp emission cut at $l \approx 30°$ in contrast to a tail out to $l \approx -120°$, and two emission bumps towards Cygnus and near the anticenter in the ACP region. Notice that the Carina feature at $l \approx 286°$ is invisible until the resolution level $(5,5)$ comes in. Obviously, in level $(4,4)$, which allows for $2^{4-1} = 8$ major bumps, the feature falls exactly in a valley produced by two adjacent bumps, and consequently is suppressed similar to the galactic plane point sources in Fig. 3.15.

The above examples demonstrate that the multiresolution Richardson-Lucy algorithm is a powerful tool for the reconstruction of diffuse emission from COMPTEL γ-ray data. The impact of statistical uncertainties on the reconstructed intensity distribution is largely reduced, leading, within the accuracy of the data, to more reliable representations of the underlying emission profile. Apparently, the current approach has difficulties in reconstructing isolated point sources, but it should be emphasized that we are still in an exploratory phase. The next steps are aimed at consolidating

our understanding of noise propagation in the intensity distributions, and studying the impact of selected wavelet type on the reconstructions.

3.5 Discussion of the 1.8 MeV Sky Maps

The main feature of all 1.8 MeV all-sky maps presented in this chapter is the outstanding emission along the galactic plane – clear evidence that nucleosynthesis is currently active throughout the entire Galaxy. Local ^{26}Al, as proposed by Clayton (1984), Morfill and Hartquist (1985), or Blake and Dearborn (1989), can be clearly excluded as the main origin of the 1.8 MeV emission. Also a single point source at the galactic center, as suggested by von Ballmoos *et al.* (1987) and Hillebrandt *et al.* (1987), is obviously not at the origin of the 1.8 MeV γ-ray line.

Although the simulations of diffuse emission distributions suggest that the lumpiness of the R-L and ME skymaps is mainly due to statistical fluctuations of the instrumental background, several irregularities have proven significant, which are also clearly seen in the multiresolution R-L maps. The most prominent of these is the extended emission feature in Cygnus, centered at $l \approx 80°$ with an apparent diameter of about 15°. The corresponding spectrum shown in Fig. 2.5 clearly demonstrates that this feature is indeed associated with 1.8 MeV line emission. Initiated by the observed 1.8 MeV emission from Cygnus, del Rio *et al.* (1996) studied the population of possible candidate sources in this region. Using core collapse supernova, ^{26}Al yields of Weaver and Woosley (1993), and WR yields of Langer *et al.* (1995), they estimated that the known supernova remnants and Wolf-Rayet stars in the Cygnus region could give rise to a total 1.8 MeV flux of $4.2\,10^{-5}$ ph cm^{-2}s^{-1}, which is consistent with the observed value (cf. Tabs. 3.1 and 3.2). It is worthwhile to note that about $\sim 85\%$ of the 1.8 MeV flux in the model of del Rio *et al.* (1996) is due to ^{26}Al production in Wolf-Rayet stars.

The southern-sky 'counterpart' of the Cygnus feature may be the enhanced emission towards the Vela direction at $l \approx 265°$. This emission could indeed be a physical 'counterpart' since it is suggested that the massive stars found in both regions form a local spiral arm running from the Cygnus region, where it crosses the solar circle, through the Sun (the Sun located near the inner edge of the arm) towards Puppis-Vela (e.g., Sahu 1992, and references therein). Again, Fig. 2.5 demonstrates the presence of a clear 1.8 MeV line signal in this direction. Based on a counterpart search in the Vela region by Oberlack *et al.* (1994), Diehl *et al.* (1995a) tentatively identified the emission feature with the Vela supernova remnant, with a possible contribution from the closest Wolf-Rayet star γ^2 Velorum.

The most concentrated emission feature in the all-sky maps is found at $(l, b) = (286.5°, 0.5°)$ in the direction of Carina (the corresponding spectrum is shown in Fig. 2.5). A counterpart search in the Carina region revealed that the 1.8 MeV feature correlates with the largest concentration of young open clusters known in the Galaxy (Knödlseder *et al.* 1996a). The Carina region also houses NGC 3372, one of the most prominent

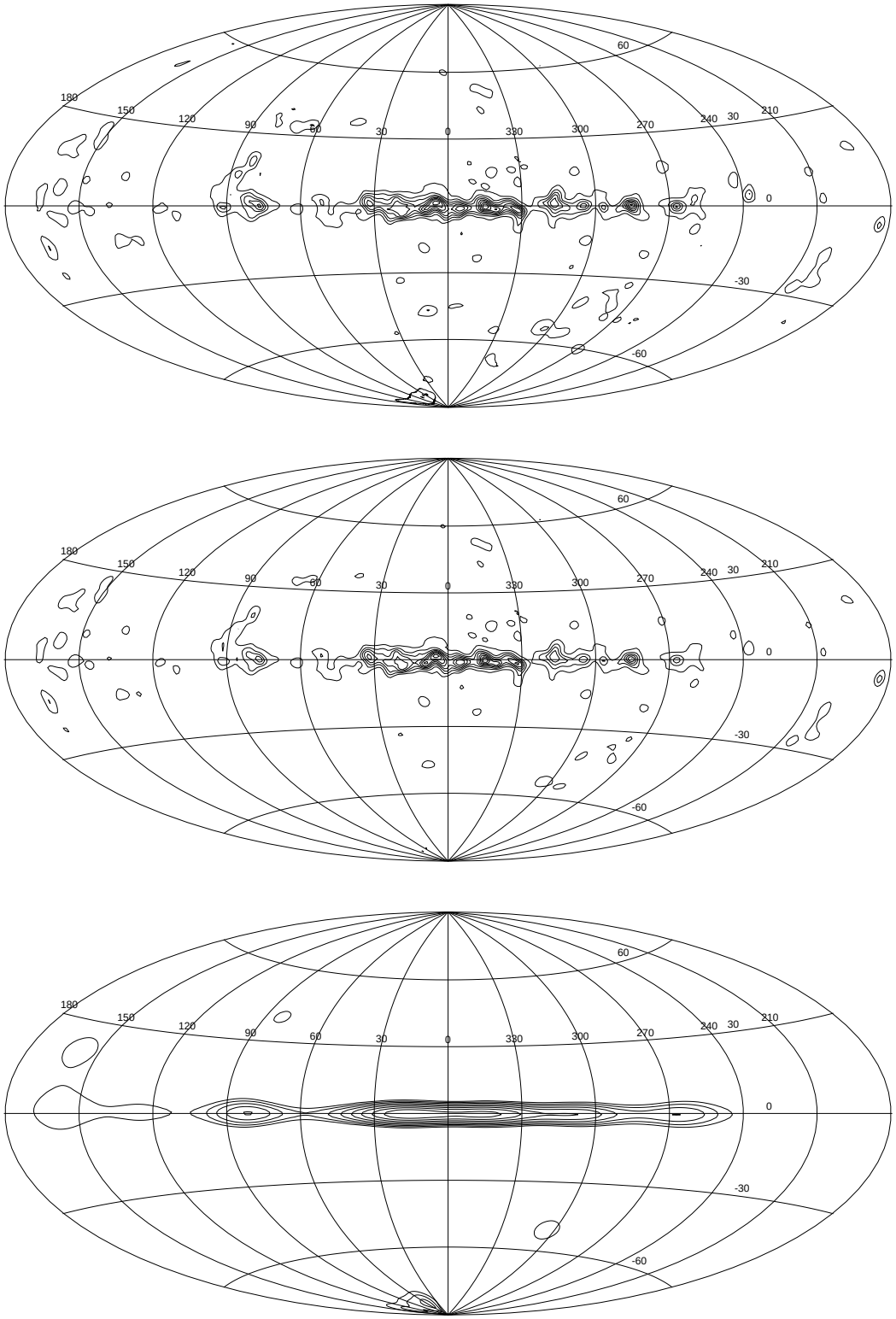

Fig. 3.18: *COMPTEL 1.8 MeV all-sky maps for observation periods 0.1 - 522.5. The panels show (from top) the 'optimum' iteration 6 of the Richardson-Lucy reconstruction, the 'optimum' iteration 6 of the maximum entropy algorithm, and the converged (4, 4) map of the multiresolution R-L algorithm for which no important artefacts due to statistical fluctuations are expected in the image.*

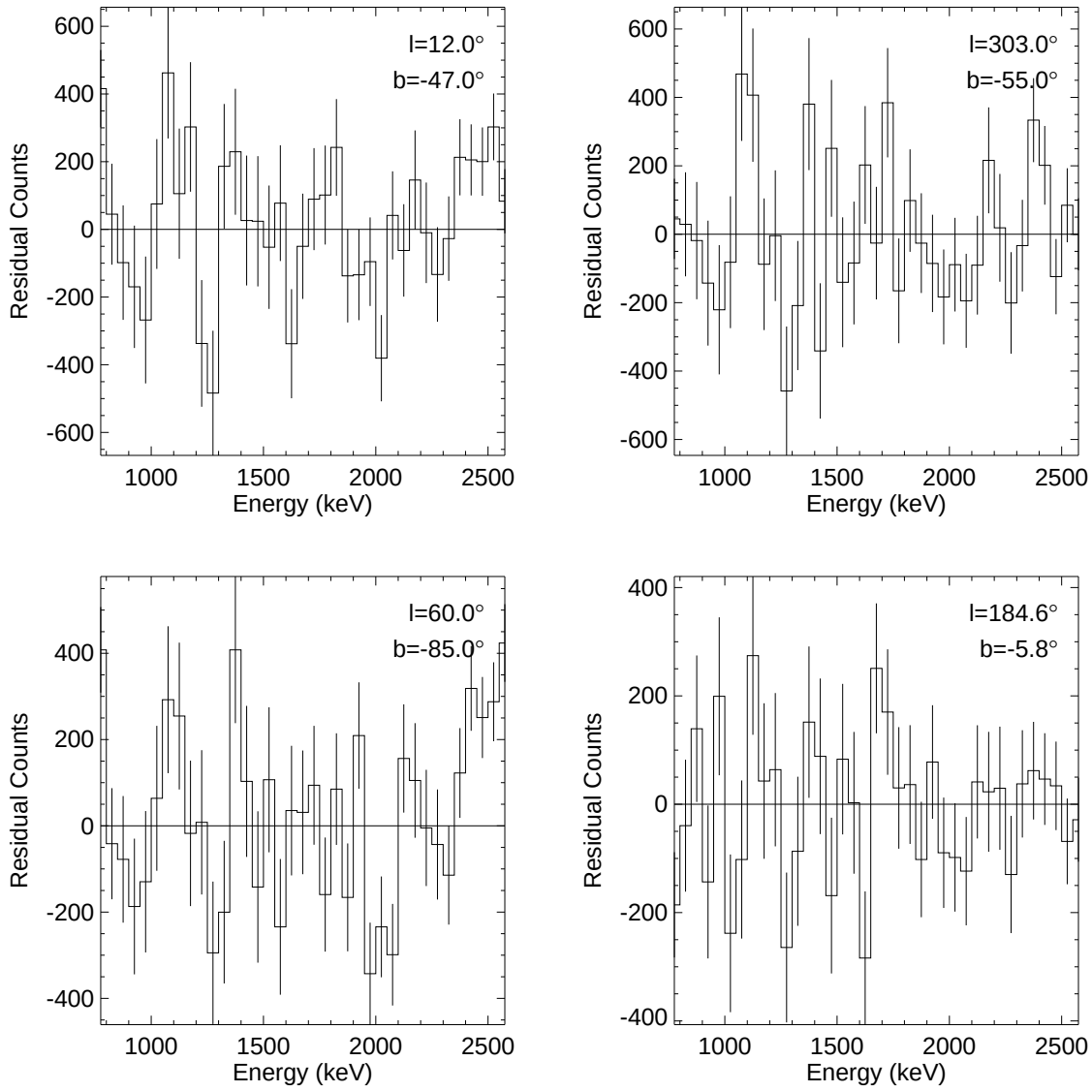

Fig. 3.19: *Background subtracted COMPTEL energy spectra for four emission features in the 1.8 MeV skymaps.*

H II regions of the Galaxy, which is powered by the largest concentration of (six) O3 stars known in the Galaxy. Additionally the most massive star known in the Galaxy is also hosted in this region: η Carinae, a luminous blue variable (LBV). In view of all these massive objects, the most plausible explanation for the 1.8 MeV feature is given by an accumulation of ^{26}Al produced by massive stars along the line of sight, which is supported by the suggested presence of a spiral arm in this direction.

Furthermore, an extended 1.8 MeV emission feature is seen in all skymaps which extends over the constellations Auriga, Camelopardalis, and Perseus (ACP). Again,

Fig. 2.5 confirms that this feature is indeed due to 1.8 MeV γ-ray line emission, at least for latitudes $b \sim -13°$. The large latitude extension of the emission implies that it must be rather nearby: assuming that ^{26}Al resides in a bubble with radius of 100 pc (i.e. a typical wind blown bubble or superbubble), the latitude extension of $\pm 30°$ implies a distance to the bubble center of 200 pc. This is about the distance to a group of B-type stars associated with the Taurus-Auriga molecular cloud complex (Walter and Boyd 1991).

Finally, some features are found in the all-sky maps at high galactic latitudes, in particular around $b \sim -55°$ or near the southern galactic pole. However, spectra derived for these features (cf. Fig. 3.19) show no evidence for 1.8 MeV gamma-ray line emission in these directions, indicating that the features arise from statistical fluctuations or systematic uncertainties of the background model. In this context, it is worthwhile to mention that the weak emission feature in the all-sky R-L/MEM maps at $(l, b) = (184.6°, -5.8°)$ is residual emission from the Crab nebula, the brightest continuum source at MeV energies. Notice that this small residual is a proof of successful inclusion of continuum emission in the background model, since without suppression, the Crab would outshine the entire 1.8 MeV emission.

Chapter 4

Model comparison

4.1 Principles of Model Comparison

4.1.1 Bayesian Parameter Estimation

In the previous chapter it was shown how an intensity distribution is derived from the measured data using the framework of Bayesian inference. The aim was the estimation of gamma-ray intensities in a rectangular $1^\circ \times 1^\circ$ pixel grid under the assumption that the pixels take mutually independent values. This assumption was called the *model*, the pixels being the 64 800 parameters of that model. Alternatively, some idea about the intensity distribution might be present, such as the shape of the emission or the presence of point sources, but their amplitudes and positions are unknown. In particular, model intensity distributions based on assumptions of the ^{26}Al distribution in the Galaxy should be compared directly to the data, with the total galactic ^{26}Al mass being an unknown parameter.

These questions can be addressed by the same principles as discussed for image reconstruction, only the model and the parameters will change. In particular, the number of parameters will be very small ($O(1)$), which suggests different techniques for their estimation. For the moment consider a given model M which contains m parameters denoted by $\theta \equiv \{\theta_1 \ldots \theta_m\}$ (an example is a given sky intensity distribution with unknown total flux as parameter). Given data D and prior information I, the probability distribution for the model parameters is determined by Bayes' theorem

$$P(\theta|D, M, I) = P(\theta|M, I)\frac{P(D|\theta, M, I)}{P(D|M, I)}, \tag{4.1}$$

in which the denominator serves as the normalizing constant

$$P(D|M, I) = \int P(D|\theta, M, I)P(\theta|M, I)d^m\theta. \tag{4.2}$$

When nothing is known about the model parameters a uniform prior probability may be chosen. Nevertheless, most model parameters are continuous positive variables (e.g. the flux of a point source or the total galactic ^{26}Al mass), for which Jeffreys (1939) showed that the proper way to express complete ignorance is to assign uniform prior probability to its logarithm; i.e., the prior probability density is

$$P(\theta_j|M,I) = \frac{1}{\theta_j}. \tag{4.3}$$

Jaynes (1996) notes, however, that almost any prior which is smooth in the region of high likelihood, will lead to substantially the same final conclusions. He remarks that "from a pragmatic standpoint, arguments about which prior probabilities correctly express a state of "complete ignorance" [...] usually amount to quibbling over pretty small peanuts" (Jaynes 1996, p. 631). If in this spirit constant priors are chosen for all parameters, the posterior probability distribution for the parameters becomes directly proportional to the likelihood function $L(\theta) \equiv P(D|\theta, M, I)$:

$$P(\theta|D,M,I) \propto L(\theta). \tag{4.4}$$

Finding the most probable parameters for a given model consists then in maximizing the likelihood function over the parameters – a technique known as *maximum likelihood parameter estimation.*

4.1.2 The Maximum Likelihood Ratio Test

In principle, the degree of reliability which is attached to the parameter estimates can be derived by drawing the entire probability distribution $P(\theta|D, M, I)$ to make a statement of the form, "90 % of the posterior probability is concentrated in the interval $\theta_{min} \leq \theta \leq \theta_{max}$". Calculation of the entire posterior probability distribution, however, can be computationally very demanding. Therefore, an approximation is introduced which allows fast error estimation using the *maximum likelihood ratio test.* In this test, two parametric hypotheses H_0 and H_1 with parameters θ are considered which belong to the same continuous family.[1] While for H_1 all parameters are free, the *null hypothesis* H_0 specifies fixed values $\tilde{\theta}_j$ for some subset $1 \ldots q$ of the m parameters.[2] The *maximum log-likelihood ratio* test statistic Λ is given by

$$\Lambda = -2 \ln \left(\frac{\max\limits_{\theta_j, j=q+1 \ldots m} L(\tilde{\theta}_1 \ldots \tilde{\theta}_q, \theta_{q+1} \ldots \theta_m)}{\max\limits_{\theta_j, j=1 \ldots m} L(\theta_1 \ldots \theta_m)} \right) \tag{4.5}$$

[1] A continuous family of hypotheses is a set of parametric hypotheses for which the whole family is obtained by continuous variation of the parameters.

[2] Such hypotheses are called *nested hypotheses* since the parameter space of H_0 is nested in that of H_1.

where the likelihood functions are maximized with respect to the free parameters. It can be shown (Eadie *et al.* 1982, Cash 1979) that Λ is asymptotically a central χ^2_q distribution under H_0 with q degrees of freedom. In other words, if H_0 imposes q constraints on the m parameters of the model, then Λ is distributed as a χ^2_q under H_0.

Following Strong (1985), this property can be used to assign uncertainties to maximum likelihood parameter estimates. For this purpose, the log-likelihood function $\ln L(\theta)$ is expanded at the point of the maximum likelihood estimate (MLE) $\hat{\theta} = \left\{\hat{\theta}_1 \ldots \hat{\theta}_m\right\}$:

$$\ln L(\theta) = \ln L(\hat{\theta}) + \frac{1}{2}\sum_{j,k=1}^{m}(\theta_j - \hat{\theta}_j)H_{jk}(\theta_k - \hat{\theta}_k) + \ldots \tag{4.6}$$

where $H_{jk} = \partial^2 \ln L/\partial\theta_j\partial\theta_k$ is the Hessian matrix of the log-likelihood function, taken at $\hat{\theta}$. Since H is real, symmetric, and positive definite, it can be diagonalized by $H = Q \cdot \mathrm{diag}(\lambda_j) \cdot Q^T$, where Q is a real orthogonal matrix of eigenvectors of H, and $\mathrm{diag}(\lambda_j)$ is a diagonal matrix composed of the corresponding eigenvalues λ_j. In terms of the transformed variables, the log-likelihood expansion becomes

$$\ln L(\theta) = \ln L(\hat{\theta}) + \frac{1}{2}\sum_{j=1}^{m}\lambda_j x_j^2, \tag{4.7}$$

where

$$x_j = \sum_{k=1}^{m} Q_{kj}(\theta_k - \hat{\theta}_k). \tag{4.8}$$

Strong (1985) showed that the offsets $\Delta\theta_j = \theta_j - \hat{\theta}_j$ can be expressed as a function of the log-likelihood difference $2\Delta \ln L = 2\left(\ln L(\theta) - \ln L(\hat{\theta})\right)$, using

$$\Delta\theta_j = \pm\sqrt{2\Delta \ln L \sum_{k=1}^{m}\frac{Q_{jk}^2}{\lambda_k}}. \tag{4.9}$$

But due to the diagonalization, $2\Delta \ln L$ is identical to the maximum likelihood test statistic Λ for $q = 1$. Consequently, $2\Delta \ln L$ is distributed as χ^2_1, hence 1σ parameter errors are determined from Eq. (4.9) by setting $2\Delta \ln L = 1.0$.

4.1.3 Bayes Factors

The previous sections treated the conduction of inference within the context of a preassigned model, representing some working hypothesis about the phenomenon being observed. In this section the question of how to decide between different models when both seem able to account for the facts is addressed. For this purpose, a set of different models $\{M^{(1)} \ldots M^{(r)}\}$ with parameter vectors $\{\theta^{(1)} \ldots \theta^{(r)}\}$ is considered. Bayes' theorem gives the posterior probability for the j'th model as

$$P(M^{(j)}|D, I) = P(M^{(j)}|I)\frac{P(D|M^{(j)}, I)}{P(D|I)}, \qquad 1 \le j \le r. \tag{4.10}$$

The denominator $P(D|I)$ can be eliminated by considering the *posterior odds* for model $M^{(j)}$ over $M^{(k)}$, given by

$$\frac{P(M^{(j)}|D,I)}{P(M^{(k)}|D,I)} = \frac{P(M^{(j)}|I)}{P(M^{(k)}|I)} \cdot \frac{P(D|M^{(j)},I)}{P(D|M^{(k)},I)} \tag{4.11}$$

where the *prior odds* $P(M^{(j)}|I)/P(M^{(k)}|I)$ quantify the prior favour of model $M^{(j)}$ over $M^{(k)}$. Comparison with Eq. (4.1) shows that the same probability $P(D|M^{(j)},I)$ that appears in the single-model parameter estimation problem only as a normalization constant, now appears as the fundamental quantity determining the status of model $M^{(j)}$ relative to $M^{(k)}$. The second term in Eq. (4.11) is called the *Bayes factor* (Jeffreys 1939)

$$B_{jk} = \frac{P(D|M^{(j)},I)}{P(D|M^{(k)},I)} \tag{4.12}$$

which is defined as the ratio between posterior to prior odds. Note that, unlike in parameter estimation where only a maximum is searched, all probabilities appearing in the definition of $P(D|M^{(j)},I)$ must be correctly normalized when computing B_{jk}.

The basic task for model comparison is the evaluation of the integrals

$$P(D|M,I) = \int L(\theta)P(\theta|M,I)d^m\theta \tag{4.13}$$

which is also called the *marginal probability* or the *marginal likelihood* of the data. Analytical evaluation is, however, often intractable. Numerical methods are generally so inefficient for these integrals that they are of little use. The main reason is that for moderate or large sample sizes the integrand becomes highly peaked around its maximum, and quadrature methods have difficulty finding the region where the integrand mass is accumulating. This is the domain where the Laplace method for integrals becomes successful. It is based on a Taylor series expansion of the function $f(u)$ of the $m-$dimensional vector u, and yields the approximation

$$\int \exp\{f(u)\}du \approx (2\pi)^{m/2} \det(A)^{\frac{1}{2}} \exp\{f(\tilde{u})\}, \tag{4.14}$$

where $\tilde{u}$ is the value of u where f attains its maximum, and A is the negative of the inverse Hessian of f evaluated at $\tilde{u}$. When applied to Eq. (4.13) it yields

$$P(D|M,I) \approx (2\pi)^{m/2} \det(\Sigma)^{\frac{1}{2}} L(\tilde{\theta})P(\tilde{\theta}|M,I), \tag{4.15}$$

where m is the dimension of θ, $\tilde{\theta}$ is the posterior mode of θ, and Σ is the negative of the inverse Hessian of $\ln\{L(\theta)P(\theta|M,I)\}$, evaluated at $\theta = \tilde{\theta}$. Tierney and Kadane (1986) show that in regular statistical models[3] the relative error in Eq. (4.15), and hence in the resulting approximation to B_{jk}, is $O(n^{-1})$[4].

[3] A regular statistical model is one in which the MLE is asymptotically normal with mean at the true value and variance matrix equal to the inverse expected Fisher information matrix (Raftery 1995).

[4] For Poisson data, n is the total number of counts (Raftery 1994). In the case of COMPTEL 1.8 MeV all-sky data, n = 3 967 035.

A second approximation is obtained by replacing the posterior mode by the maximum likelihood estimate $\tilde{\theta} \approx \hat{\theta}$ and taking the inverse Hessian matrix $\Sigma \approx H^{-1}$ at this point, which yields

$$2 \ln B_{jk} \approx \Lambda + (E_j - E_k) \tag{4.16}$$

where $\Lambda = 2(\ln L(\hat{\theta}^{(j)}) - \ln L(\hat{\theta}^{(k)}))$ is the generalized maximum log-likelihood ratio test statistic (cf. Eq. (4.5)), and

$$E_j = -\ln\{\det(H^{(j)})\} + 2 \ln P(\hat{\theta}^{(j)}|M^{(j)}, I) + m^{(j)} \ln(2\pi). \tag{4.17}$$

This approximation is less accurate than Eq. (4.15) since the relative error increases to $O(n^{-\frac{1}{2}})$ (Raftery 1994). Instead of E_j, Jaynes (1996) use the quantity

$$W_j = exp\left\{\frac{1}{2}E_j\right\} = V(\Omega^{(j)})P(\hat{\theta}|M^{(j)}, I), \tag{4.18}$$

where

$$V(\Omega^{(j)}) = (2\pi)^{m^{(j)}/2}\sqrt{\det(H^{(j)})^{-1}} = (2\pi)^{m^{(j)}/2}\left(\prod_{i=1}^{m^{(j)}} \lambda_i^{(j)}\right)^{-\frac{1}{2}} \tag{4.19}$$

is the effective high-likelihood volume[5] of the parameter space ($\lambda_i^{(j)}$ are the eigenvalues of the Hessian matrix $H^{(j)}$ as derived in Section 4.1.2). W_j is also called the Ockham[6] factor, since it is related to the "simplicity" of the model (Jaynes 1996). It can be interpreted as the *amount of prior probability* contained in the high-likelihood region $\Omega^{(j)}$. In this way, W_j quantifies the simplicity of a given model, because the more complicated a set of possible hypotheses (i.e. a parametric model) is, the larger the manifold of conceivable alternatives becomes, and so the smaller must be the prior probability of any particular hypothesis in the set. In this sense complicated models are expected to provide smaller Ockham factors.

---

[5] A high-likelihood region Ω in parameter space is defined by the conditions that Ω is as compact as possible and within Ω the likelihood everywhere exceeds some nominal threshold value L_0. The volume of this region is then

$$V(\Omega) = \int_{L>L_0} d^m\theta \tag{4.20}$$

where L_0 is specified by the condition that the integrated likelihood should be given by $L(\hat{\theta})V(\Omega)$:

$$L(\hat{\theta})V(\Omega) \doteq \int L(\theta)d^m\theta = L(\hat{\theta})(2\pi)^{m/2}\sqrt{det(H^{-1})}. \tag{4.21}$$

Then a rectangular function equal to $L(\hat{\theta})$ on Ω, zero elsewhere, is a crude approximation to the likelihood function and it has, in the present approximation, the same implications for model comparison as does the actual likelihood function.

[6] William of Ockham (lat. Occam), $\sim 1290 - 1349$, Franciscan monk. By the principle "Entities are not multiplied without necessity", later known as "Ockham's Razor", he insisted that what can be done with fewer is done in vain with more.

B_{10}	$2 \ln B_{10}$	Evidence for M_1
< 1	< 0	Negative (supports M_0)
1 to 3	0 to 2.2	Marginal
3 to 12	2.2 to 5.5	Positive
12 to 150	5 to 10	Strong
> 150	> 10	Decisive

Table 4.1: *Decision rules for model comparison with unity prior odds (after Raftery 1994).*

Throughout this work, Bayes factors will be computed using Eq. (4.16). Prior odds were set to unity, corresponding to prior information that is indifferent between competing models. Consequently, the posterior odds are identical to the Bayes factor. The specification of parameter priors $P(\theta|M, I)$ has to be done more carefully. Although Raftery (1994) notes that the exact prior form is not very important, its proper normalization is essential for the Bayes factor analysis. Independent priors were assumed for all parameters, corresponding to the situation where the individual variables are of interest in their own right. Throughout this work, bounded uniform and normal priors were applied. Their motivation is discussed in the applications below. For the interpretation of Bayes factors, Raftery (1994) proposes the decision rules given in Tab. 4.1. It is important to note that the categories given in the table are not issue of a calibration of the Bayes factor but rather a rough descriptive statement about standards of evidence in scientific investigations – probability is a meaningful scale itself!

4.1.4 Bayesian Model Averaging

Parameter estimation as proposed in Section 4.1.1 has a major drawback: the results are only valid in the context of the considered model. In real problems, like those treated in this work, a wealth of possible models exists, each of which generally leads to different parameter estimates. In this case, data analysts typically select the most probable model and then proceed as if the model generated the data. This approach, however, ignores the uncertainty involved in the model selection, which can be a major part of overall uncertainty. Generally, this leads to underestimation of the uncertainty of the quantities of interest, and implies over-confident inferences.

Bayesian model averaging (BMA) provides a coherent mechanism for accounting for the model uncertainty. We wish to confine our attention to a selected set of models $\{M^{(1)} \dots M^{(r)}\}$. Bretthorst (1988) showed that for answering questions within this set one may normalize

$$\sum_{j=1}^{r} P(M^{(j)}|D, I) = 1, \tag{4.22}$$

leading to the expression

$$P(M^{(j)}|D,I) = \frac{P(M^{(j)}|I)P(D|M^{(j)},I)}{\sum_{k=1}^{r} P(M^{(k)}|I)P(D|M^{(k)},I)} = \frac{P(M^{(j)}|I)B_{j1}}{\sum_{k=1}^{r} P(M^{(k)}|I)B_{k1}} \tag{4.23}$$

for the posterior model probabilities (the choice of model 1 as reference model in the Bayes factor is arbitrary). If x is the quantity of interest, its posterior probability, given data D and knowledge I, is determined by

$$P(x|D,I) = \sum_{j=1}^{r} P(x|M^{(j)},D,I)P(M^{(j)}|D,I). \tag{4.24}$$

This is an average of the predictive distributions of x, given model $M^{(j)}$,

$$P(x|M^{(j)},D,I) = \int P(x|D,\theta^{(j)},M^{(j)},I)P(\theta^{(j)}|D,M^{(j)},I)dx \tag{4.25}$$

weighted by the models' posterior probabilities. The posterior mean and standard deviation of x are then calculated from

$$\mathrm{E}[x|D,I] = \sum_{j=1}^{r} \bar{x}_j P(M^{(j)}|D,I), \tag{4.26}$$

$$\mathrm{Var}[x|D,I] = \sum_{j=1}^{r} \left[(\Delta x_j)^2 + (\bar{x}_j)^2\right] P(M^{(j)}|D,I) - \left(\mathrm{E}[x|D,I]\right)^2, \tag{4.27}$$

where $\bar{x}_j \equiv \mathrm{E}[x|D,M^{(j)},I]$ and $(\Delta x_j)^2 \equiv \mathrm{Var}[x|D,M^{(j)},I]$ are the expectation and variance of x under model $M^{(j)}$, respectively. If x is identical to a parameter θ_x which is common to all models the expectation and variance can be approximated by the MLE $\bar{x}_j \approx \hat{\theta}_x^{(j)}$ and the error estimate $(\Delta x_j)^2 \approx (\Delta\theta_x^{(j)})^2$ (cf. Eq. (4.9)).

4.2 Application to COMPTEL data

4.2.1 Basic Algorithm

All methods described above require the maximization of the log-likelihood function

$$\ln L(\theta|M) = \sum_{i=1}^{N} n_i \ln e_i(\theta|M) - e_i(\theta|M) - \ln n_i! \tag{4.28}$$

(cf. Eq. (3.7)) where n_i is the number of observed counts in data space cell i and $e_i(\theta|M)$ is the prediction of model M with parameter vector θ for this cell. Consequently, the base of any model comparison is a maximum likelihood algorithm. For the analysis

of COMPTEL data, the *Levenberg-Marquardt method* was implemented (Press *et al.* 1992) which makes use of the log-likelihood gradient

$$\frac{\partial \ln L}{\partial \theta_j} = \sum_{i=1}^{N} \frac{\partial e_i(\theta|M)}{\partial \theta_j} \left[\frac{n_i}{e_i(\theta|M)} - 1 \right], \tag{4.29}$$

and the Hessian matrix

$$\frac{\partial^2 \ln L}{\partial \theta_j \partial \theta_k} = \sum_{i=1}^{N} \left\{ \frac{\partial^2 e_i(\theta|M)}{\partial \theta_j \partial \theta_k} \left[\frac{n_i}{e_i(\theta|M)} - 1 \right] - \frac{\partial e_i(\theta|M)}{\partial \theta_j} \frac{\partial e_i(\theta|M)}{\partial \theta_k} \frac{n_i}{e_i^2} \right\}. \tag{4.30}$$

Starting from an initial guess of the parameters, this method iteratively improves the log-likelihood until convergence is reached. It turns out that the method is rather independent on the initial parameter values and requires only $2-3$ iterations to achieve an absolute $\ln L$ error of better than 10^{-4} (comparison with Tab. 4.1 shows that decisions are based on log-likelihood differences of $O(1)$, hence the achieved accuracy[7] is largely sufficient). Eigenvectors and corresponding eigenvalues of the Hessian matrix at maximum likelihood are calculated using Jacobi transformations (Press *et al.* 1992). From these, 1σ parameter errors $\Delta\theta_j$ are derived using the method described in Section 4.1.2. The eigenvalues are also used for the calculation of the high-likelihood volume $V(\Omega)$ following Eq. (4.19).

Using the output of the maximization algorithm, namely the maximum log-likelihood $\ln L(\hat{\theta}|M)$ and the high-likelihood volume $V(\Omega)$, Bayes factors can be calculated independently from the maximization (indeed, this is a major advantage of the approximation Eq. (4.16)). Nevertheless, the functional relationship between some of the model parameters and the predicted counts in data space may be so complex that evaluation of the derivatives and the Hessian matrix are impossible, and direct maximization using the Levenberg-Marquardt method is infeasible. In this case, the intractable parameters are fixed, and maximum likelihood with respect to the specified values is estimated by optimizing the remaining free parameters. The global maximum likelihood is determined by repeating this procedure for a grid of parameter values which cover the conceivable parameter space with sufficient density.[8] The derivatives and the Hessian matrix, which are required for parameter error estimation and Bayes factor analysis, are then approximated numerically using the grid points adjacent to the maximum likelihood point.

---

[7] As technical sidemark, notice that all calculations were done using double precision variables. This is necessary since the all-sky data space comprises $N = 1\,360\,800$ valid bins (the $\bar{\varphi}$-layers $1-4$ are empty for 1.8 MeV data), and summing so many quantities leads to considerable roundoff errors.

[8] Sufficient density means that around maximum likelihood the variation of $\ln L$ between adjacent parameter grid points becomes $\lesssim 1$, hence decisions won't be altered by the crude approximation.

4.2.2 An Example: Fitting Exponential Disk Models

To illustrate the procedure and discuss systematic uncertainties, an example is considered: the fit of an exponential disk model. In Section 4.4.1 it will turn out that this model is a good first order parametric description of the observed 1.8 MeV data. A mock dataset was created by means of a Poisson random number generator from the combination of the all-sky background model and the expected source count distribution from an exponential disk model with radial scale length of $R_0 = 4.5$ kpc, scale height of $z_0 = 90$ pc, and total galactic ^{26}Al mass of $M = 3$ M$_\odot$. These data were then compared with the parametric exponential disk model on top of the background model, where R_0, z_0, M, and the background scaling factor β are 'unknown' parameters. Obviously, this example is an ideal case where the background model is known exactly and the parametric source model comprises the 'true' model.

The first step in Bayesian model comparison is the assignment of prior probabilities for all model parameters. For the scale length R_0 and scale height z_0, bounded constant priors were chosen since the prior knowledge about the shape of the Milky Way sets some constraints on these parameters. In particular, the lower boundary of z is well defined. All stars in the Galaxy show some scatter with respect to the galactic plane due to the vertical dispersion of their birth places – the molecular clouds – and, at least for older stars, due to interactions with the interstellar medium (ISM). The most tightly confined population are young massive stars which obey a scale height of $z \sim 50$ pc (Cruz-Gonzalez *et al.* 1974, Garmany *et al.* 1982). Older stars are more widespread because they already encountered significant collisions with interstellar clouds, causing an increase in their velocity dispersion over time. Since ^{26}Al is believed to be produced in stars, it is reasonable to assume that its scale height is at least that of the most tightly confined population (i.e., that of massive stars). Consequently, the lower bound of z_0 is set to 50 pc. The upper bound may be derived from an assumed maximum velocity of ^{26}Al, since the million years lifetime of the isotope limits the mean travel distance to ~ 1 pc (km s^{-1})$^{-1}$. ^{26}Al velocities may in turn be estimated from 1.8 MeV line width measurements using high resolution germanium detectors. While Mahoney *et al.* (1984) could only state that the width of the line is consistent with the galactic rotation velocity ($v \sim 200$ km s^{-1}), recent observation using the *GRIS* balloon experiment indicates a line broadening corresponding to $v \sim 500$ km s^{-1} (Naya *et al.* 1996). Considering the uncertainties in these measurements and the maximum scale height for disk stars of 325 pc (Wainscoat *et al.* 1992), a conservative upper limit of $z_0 < 3$ kpc is assumed.[9] For the radial scale length, the boundaries are more fuzzy. Estimates by numerous workers span an enormous range from $1 - 6$ kpc, depending on the method and investigated wavelength range used for the determination (see Kent *et al.* 1991, Wainscoat *et al.* 1992, for summaries of some of these values). Due to the large

[9] One might argue that a halo population is much more extended than the assumed scale height limit, but the considered model is one of the galactic disk population, not of a halo population – for the latter a radial scale length and vertical scale height are not very meaningful parameters.

uncertainty of the galactic scale length it is advisable to choose a rather uninformative prior, covering the range of plausible scale lengths. Consequently, a bounded uniform prior was applied, reaching from $R_0 = 1 - 10$ kpc.

The prior range of the galactic ^{26}Al mass for exponential disk models can be determined from previous 1.8 MeV observations. Including the measurement uncertainties, all observations provide a flux from the central radian around $(4.5 \pm 2.0)\,10^{-4}$ ph cm^{-2}s^{-1}rad^{-1} (Prantzos and Diehl 1996). For the range of permissible scale parameters of the exponential disk model, corresponding intensity distributions lead to $(1 - 2)\,10^{-4}$ ph cm^{-2} s^{-1} rad^{-1} M$_\odot{}^{-1}$. Hence, prior measurements suggest an exponential disk mass in the range of 3 ± 2 M$_\odot$, and a Gaussian prior with mean of 3 M$_\odot$ and standard deviation 2 M$_\odot$ is chosen for the exponential disk mass to reflect this knowledge.

The all-sky background model, which is based on events observed in adjacent energy bands, was carefully normalized on single observation basis under consideration of time variabilities of major background components (Oberlack 1997, see also Section 2.5.2.2). This normalization makes the background scaling factor β a rather well-defined parameter, where the major uncertainties arise from unidentified time-variable background components, and the zero level estimation from high galactic latitudes $|b| > 40°$. To quantify these uncertainties, a Gaussian prior with mean $\beta = 1$ and standard deviation $\sigma = 0.01$ is used for the background scaling factor.

Now, all prior probabilities are assigned and the exponential disk model can be compared to the mock data. Unfortunately, the relation of scale length R_0 and scale height z_0 to the predicted counts in the COMPTEL data space is intractable. For this reason, exponential disk models were calculated for a grid of scale lengths and scale heights. For each of the models, the maximum likelihood algorithm was applied to determine the optimum galactic ^{26}Al mass M and background scaling factor β. The resulting likelihood 'landscape' is shown in the upper-left plot of Fig. 4.1. Contour lines were chosen to represent significance levels of the maximum likelihood ratio test as derived from the χ^2_3 distribution. The posterior probability distribution is shown in the upper-right plot using identical contour spacing. As anticipated in Section 4.1.1, the posterior probability distribution is essentially similar to the likelihood function. The only net difference is the sharp edge at $z_0 = 50$ pc due to the lower boundary of the scale height prior. The true parameter values $R_0 = 4.5$ kpc, $z_0 = 90$ pc, and $M = 3$ M$_\odot$ are enclosed by the 1σ likelihood contour, demonstrating that the maximization algorithm was set up correctly. At maximum likelihood, the parameter estimates were $M = 3.0 \pm 0.1$ M$_\odot$, $R_0 = 4.5 \pm 0.2$ kpc, $z_0 = 90 \pm 17$ pc, and $\beta = 0.9995 \pm 0.0009$. Note, that optimization of the scale parameters was done on a grid with mesh size 0.5 kpc in R_0 and roughly 40 pc in z_0, hence the coincidence of the best estimates with the exact values is not surprising. From the likelihood contours one can infer that the optimum parameters on a finer grid would have been around $R_0 \sim 4.8$ kpc and $z_0 \sim 110$ pc.

At this point it is worthwhile to compare the results for the ideal case with a similar analysis of measured data. For this purpose, the exponential disk model was fitted

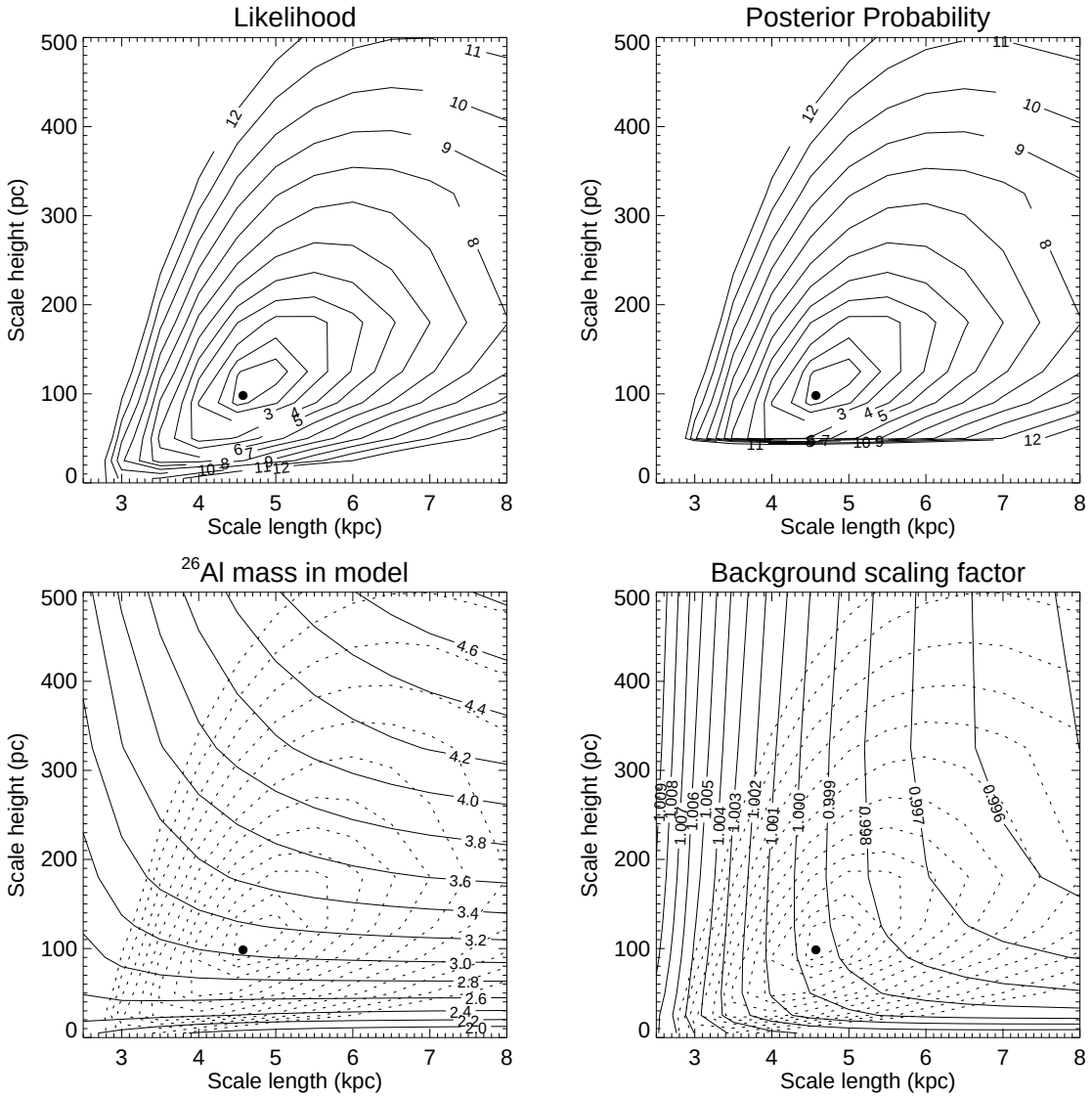

Fig. 4.1: *Fit results for the ideal case: An exponential disk model with parameters scale length R_0, scale height z_0 and total mass M, was fitted along with the 1.8 MeV all-sky background model to mock data generated from an exponential disk of $R_0 = 4.5$ kpc, $z_0 = 90$ pc, and $M = 3$ $M_\odot$ on top of the background model.* Upper left: *Likelihood contours obtained by fitting the model mass M for a grid of scale lengths R_0 and scale heights z_0. The contour levels indicate the maximum likelihood ratio with respect to the global maximum in Gaussian σ, as determined from the χ^2_3 distribution. The black dot marks the parameters of the input model.* Upper right: *Corresponding posterior probability contours.* Lower left: *Assigned galactic ^{26}Al mass M in units of $M_\odot$ as function of the exponential disk parameters. For comparison, the likelihood 'hill' is shown as dotted contours.* Lower right: *Corresponding background scaling factor β.*

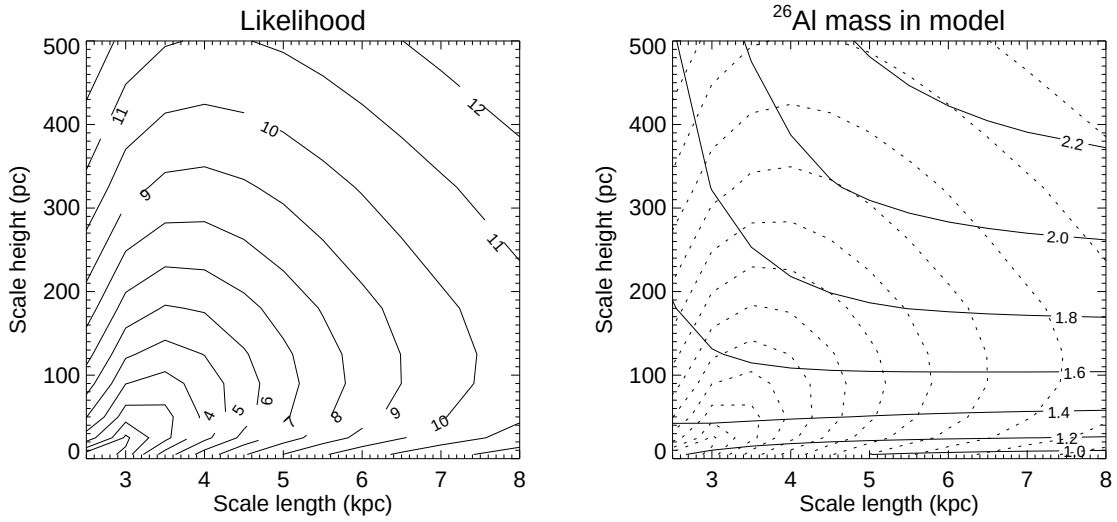

Fig. 4.2: *Likelihood and mass estimate contours for the comparison of the exponential disk model on top of the all-sky background model with the 1.8 MeV all-sky data. Contour line spacing was chosen identical to Fig. 4.1.*

together with the all-sky background model to the observed 1.8 MeV all-sky data for the same grid of scale parameters. The resulting likelihood contours and corresponding mass estimates are shown in Fig. 4.2. Global maximum likelihood is obtained at $R_0 = 2.5$ kpc and $z_0 \lesssim 5$ pc, the corresponding galactic mass estimate and background scaling factors are $M = 1.2$ M$_\odot$ and $\beta = 0.998$, respectively. Surprisingly, the optimum scale height is much lower than those known for galactic stellar populations ($z_0 \sim 50-325$ pc), and consequently it falls outside the prior range. The posterior distribution differs from the likelihood only in that it assigns zero probability for $z_0 < 50$ pc which pushes the maximum a posteriori (MAP) towards $R_0 = 3.2$ kpc and $z_0 = 50$ pc. The corresponding galactic mass estimate and background scaling factor are $M = 1.4$ M$_\odot$ and $\beta = 0.996$, respectively.

However, a maximum likelihood falling outside the prior range should alert the Bayesian analyst. It signals that, (1) either the prior probabilities were badly assigned, or (2) the prior information was wrong, or (3) the model is an insufficient description of the data, and consequently its parameters are not meaningful. For the considered example, the first two possibilities can be excluded because the lower z_0 boundary is probably the most definite prior knowledge which went into the analysis – it is really very hard to imagine how a galactic population with a scale height smaller than $z_0 \sim 50$ pc could form. Thus there remains the last possibility that the model is 'wrong'.

To examine this question, recall that the considered model consists of two components: the parametric exponential disk distribution, convolved into the COMPTEL data space, and the instrumental background model, scaled by some factor β. The more important of these two components is the background model since it has to account for more than 98% of the observed events. Consequently, fit results may depend

crucially on uncertainties of this component. The crudest approximation which is made during the generation of the background model is the normalization of its $\bar{\varphi}$ distribution to that of the data. This approximation becomes exact only, if either no source signal is present, or the source signal obeys a $\bar{\varphi}$ distribution which is similar to that of the instrumental background. In other cases, a deformation of the all-sky background model occurs which, on the one hand, depends on the difference between the source and background $\bar{\varphi}$ distributions, and on the other hand, is proportional to the local source signal strength. Thus, the $\bar{\varphi}$ normalized background model introduces a bias in the fitting procedure in that it exerts an attractive 'force' that favors source models with $\bar{\varphi}$ distributions similar to that of the background model.

To confirm this thesis, $\bar{\varphi}$ distributions of the exponential disk model for various scale parameters are compared to the $\bar{\varphi}$ distribution of the data – which in turn is a good approximation of the $\bar{\varphi}$ distribution of the instrumental background. Two extreme examples are shown in Fig. 4.3: the solid histogram represents the $\bar{\varphi}$ distribution of the data, dotted and dashed histograms those of the models. It becomes obvious that the $\bar{\varphi}$ distributions of the models are rather different from that of the data. However, calculation of the χ^2 distance between model and data shows a clear favour for the $R_0 = 2.5$ kpc, $z_0 = 5$ pc disk, similar to what was found for the maximum likelihood fitting. This model obeys a steeper $\bar{\varphi}$ distribution than the other models which makes it more similar to the background model. Consequently, the $\bar{\varphi}$ discrepancy which occurs when this source model is added to the background model and compared to the data has a minimum. Since COMPTEL data are compared in the 3-dimensional data space, the best fit is always a compromise between proper intensity distribution and minimum $\bar{\varphi}$ discrepancy. Apparently, the $R_0 = 2.5$ kpc, $z_0 = 5$ pc model is such an optimum.

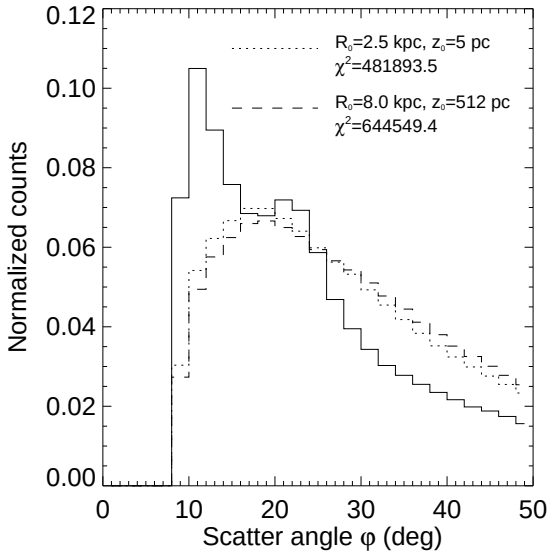

Fig. 4.3: *$\bar{\varphi}$ distributions for the data (solid) and two exponential disk models (dotted and dashed).*

To understand why this model yields a steeper $\bar{\varphi}$ distribution than the other investigated exponential disk models, recall that the instrumental point spread function (PSF) varies in $\bar{\varphi}$ over the field of view due to exposure and selection effects. Consequently, a model's $\bar{\varphi}$ distribution is a weighted average of the position-dependent PSF, being dominated by the slope of the $\bar{\varphi}$ distribution in high model-intensity regions.

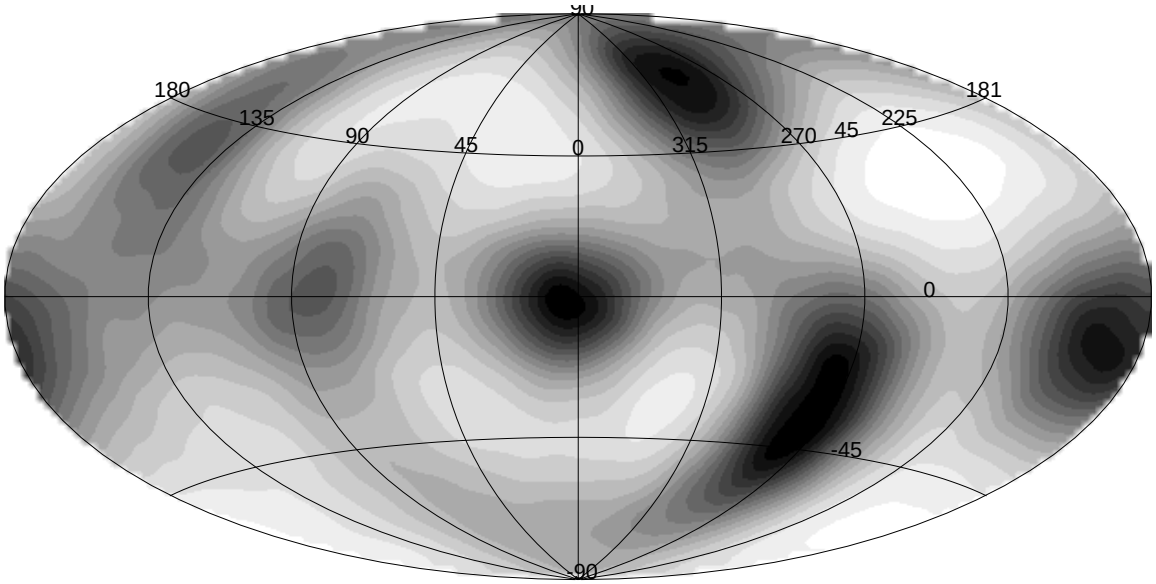

Fig. 4.4: *'Map of attraction' for 1.8 MeV all-sky data. The image shows the ratio between predicted data space counts for $\bar{\varphi} < 14°$ versus predicted counts for $\bar{\varphi} > 44°$. Higher values (darker regions) indicate steeper $\bar{\varphi}$ distribution slopes, hence model intensity distributions which are concentrated towards these regions will be favored by the fit.*

The variation of $\bar{\varphi}$ distribution slope over the sky can be examined by comparing the expected number of source counts for low scatter angles with that for high scatter angles. For this purpose, two exposure maps were created for $\bar{\varphi} < 14°$ and $\bar{\varphi} > 44°$, and the ratio between both was used as a measure of the local $\bar{\varphi}$ distribution slope. Higher ratios indicate steeper slopes, hence intensity distributions which are concentrated towards these regions will be favored by the fit. The $\bar{\varphi}$ distribution slope map, which can also be interpreted as a 'map of attraction' is shown in Fig. 4.4. From this map it becomes evident that the fit favors intensity distributions which are concentrated towards the galactic center. Remember, however, that the best fit is always a compromise between proper intensity and best $\bar{\varphi}$ distribution, hence it is not surprising that e.g. a single point source at the galactic center is discarded: Although it obeys the best $\bar{\varphi}$ distribution, its (χ, ψ) distribution is apparently far away from reality.

As small excursion, it is worthwhile to recognize that the 'attraction map' is also valid to explain some features seen in the COMPTEL all-sky maps. In this case, the same $\bar{\varphi}$ bias occurs, but the intensity model is so flexible that it can often alleviate the bias. It happens, however, that the central 1.8 MeV emission peak in the galactic plane is situated around $l \sim 5°$, exactly where the 'attraction map' has a local maximum. Additionally, a weak spur is seen in the 1.8 MeV all-sky maps around $l \sim 80°$ moving from the galactic plane towards high galactic latitudes, which could be explained by a similar spur in the 'attraction map'. It is obvious, however, that in absence of 1.8 MeV signal, no attraction occurs since in this case, the background model is correct. Consequently, no emission is seen in the skymaps towards Virgo or the LMC, despite the

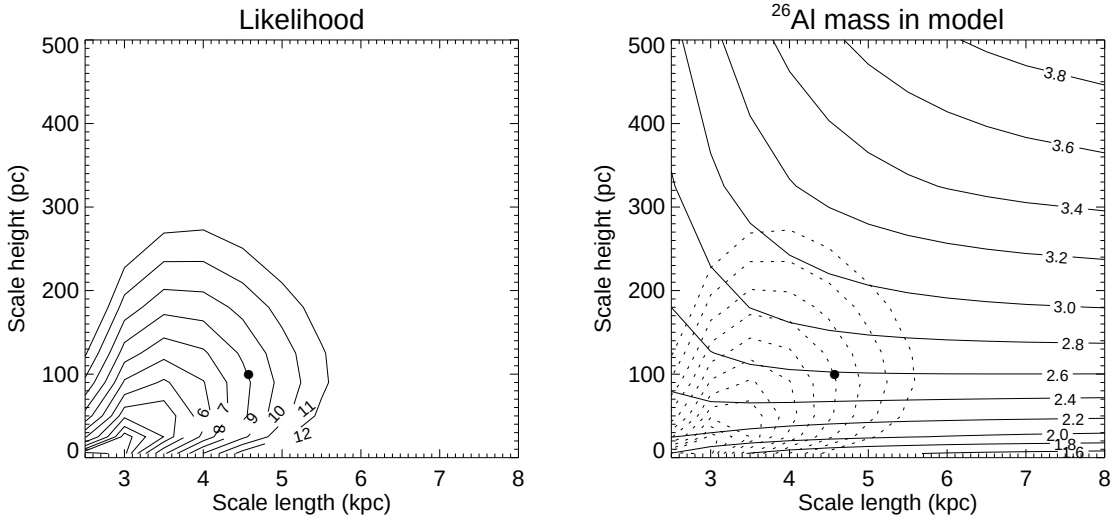

Fig. 4.5: *Fit results for the comparison of the exponential disk model to mock data, using a faked background model which was $\bar{\varphi}$ normalized to the mock data.*

strong attraction points in these regions. For other COMPTEL energy bands, where no satisfactory background model exists so far, the situation is much worse. Comparison with Fig. 1 of Strong *et al.* (1997) shows that most extended features in the $1-3$ MeV continuum map which lie off the galactic plane coincide with maxima of the attraction map. This confirms that the 'attraction map' is a reasonable indicator of systematic bias which is introduced by the uncertainties of the instrumental background models. After this excursion to image reconstruction, let's return to model comparison.

Further evidence, which makes the inaccurate $\bar{\varphi}$ distribution of the background model responsible for the unreasonable fit results comes from a modification of the 'ideal case' study. To simulate the background model generation procedure, a faked background model was created by re-normalizing its $\bar{\varphi}$ distribution on that of the mock data. Although this only approximates the background model generation procedure, it introduces the same type of inaccuracy. The likelihood contours and mass estimates which result from a combined fit of the exponential disk model along with the faked background model are shown in Fig. 4.5. Comparison with the likelihood contours of the real data fit (cf. Fig. 4.2) confirms that in both examples the same bias occurs: small radial scale lengths and unreasonable small scale heights. The steeper likelihood contours for the faked mock data fit probably stem from the higher galactic ^{26}Al mass with respect to the real data, which amplifies the bias. It is also worth noting that the estimated ^{26}Al mass at maximum likelihood is $M \sim 2$ M$_\odot$, one third smaller than the input value. The MAP parameter estimates are $R_0 = 3.3$ kpc $z_0 = 50$ pc, and $M = 2.3$ M$_\odot$, which also fall below the nominal values. Even if the scale parameters of the exponential disk had been known, the resulting mass would have been $M = 2.5 \pm 0.1$ M$_\odot$, significantly below the input mass. Hence, the background inaccuracy not only leads to wrong scale parameters, it also causes a mass underestimation.

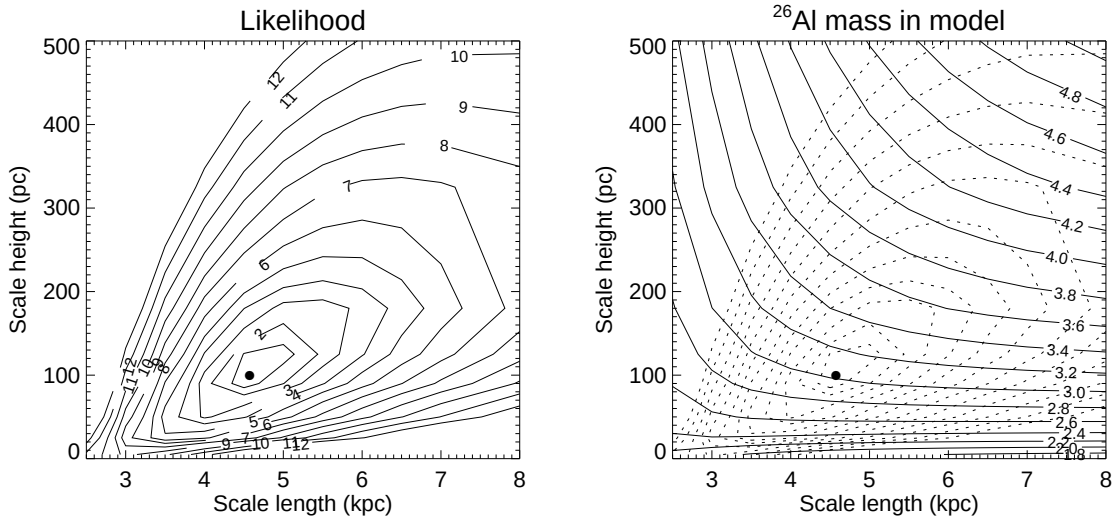

Fig. 4.6: *$\bar{\varphi}$ fitting results for the comparison of the exponential disk model to mock data, using a faked background model which was $\bar{\varphi}$ normalized to the mock data.*

The fact that simple re-normalization of the background model on the $\bar{\varphi}$ distribution of the mock data leads to the same bias as for real data, points towards a solution of the problem. The re-normalization of the background model could be reversed during the fit procedure if the single background scaling factor β would be replaced by individual scaling factors $\beta_{\bar{\varphi}}$ for each $\bar{\varphi}$ layer.[10] In that case, the $\bar{\varphi}$ distribution of the background model would be determined along with the source model parameters, and the influence of uncertainties in the background component should be greatly reduced. Application of this procedure, which will from now on be called '$\bar{\varphi}$ fitting', for fitting the faked background model to mock data leads indeed to results which are essentially similar to the 'ideal case'. For comparison, the corresponding likelihood contours are shown together with the mass estimates in Fig. 4.6.

The introduction of new parameters in the fitting procedure, however, requires the assignment of new prior probability distributions, which needs some thinking about the information that is available for the individual scaling factors. Details about the prior assignment for $\bar{\varphi}$ fitting can be found in Appendix C. In summary, the individual background scaling factors are projected into a manifold spanned by the mean background scaling factor

$$\tilde{\beta} = \frac{\sum\limits_{\bar{\varphi}} \beta_{\bar{\varphi}} B_{\bar{\varphi}}}{\sum\limits_{\bar{\varphi}} B_{\bar{\varphi}}} \tag{4.31}$$

[10]Note that for the selected energy thresholds and $\bar{\varphi}$ binning of 2° there are 21 $\bar{\varphi}$ layers, and hence 21 $\beta_{\bar{\varphi}}$ parameters for 1.8 MeV data.

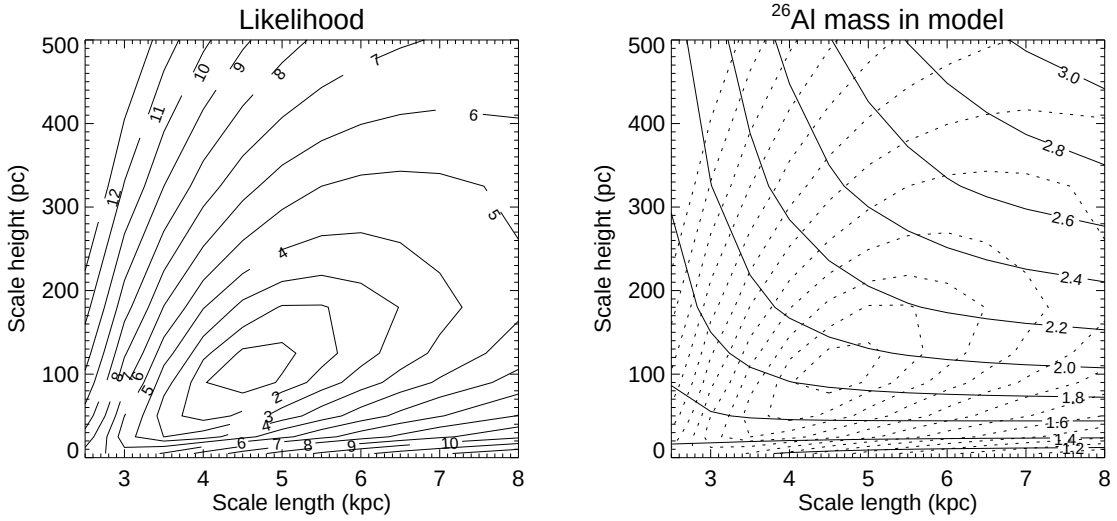

Fig. 4.7: *$\bar{\varphi}$ fitting results for the comparison of the exponential disk model with the 1.8 MeV all-sky data.*

and the normalized individual scaling factors

$$\tilde{\beta}_{\bar{\varphi}} = \frac{\beta_{\bar{\varphi}} - f}{\tilde{\beta} - f}, \tag{4.32}$$

where $B_{\bar{\varphi}}$ is the number of counts in layer $\bar{\varphi}$ of the background model, and $f = 1.00954$ is the ratio of observed counts to the counts in the background model. $\tilde{\beta}$ is identical to the fixed background model scaling factor β, and measures the deviation from the expected number of background events. $\tilde{\beta}_{\bar{\varphi}}$, which are independent of the source or background amplitude, measure the deviation between the $\bar{\varphi}$ distributions of source and background component. The advantage of the projection $\beta_{\bar{\varphi}} \rightarrow \{\tilde{\beta}, \tilde{\beta}_{\bar{\varphi}}\}$ is that prior knowledge is mutually independent for these parameters, which allows the separation of the prior probability distribution into independent Gaussians (cf. Appendix C).

The logical next step is applying the $\bar{\varphi}$ fitting procedure to real data. The results, shown in Fig. 4.7, are surprisingly similar to the 'ideal case'. In particular, the likelihood 'landscape' shows no bias towards small scale parameters, as seen for the 'fixed $\bar{\varphi}$' fits (the fitting procedure where the background model is fitted using a single scaling factor β will from now on be called the 'fixed $\bar{\varphi}$' method). The contour line spacing is somewhat wider than for the 'ideal case' which can be explained by a smaller signal in the real data than in the simulations. Indeed, a MAP mass of $M = 1.8 \pm 0.1$ $M_\odot$ is found for the exponential disk, lower than the 3 $M_\odot$ in the simulation, but significantly higher than for the fixed $\bar{\varphi}$ fit. Also similar to the 'ideal case', maximum likelihood parameter estimates are identical to the MAP estimates, demonstrating that the prior information does not really influence the conclusions for parameter estimation. For completeness, the MAP estimates of the exponential disk scale parameters are $R_0 = 4.5 \pm 0.4$ kpc and $z_0 = 90 \pm 25$ pc. The astrophysical interpretation of this

result will be postponed till Section 4.4.1, where the exponential disk model will be compared to other models for the galactic ^{26}Al distribution.

Although the introduction of a $\bar{\varphi}$ dependent background factor improved the exponential disk fit results, some cases were found where this method obviously fails to provide reliable parameter estimates. An example is the fit of an isotropic intensity distribution to the data. Using the fixed $\bar{\varphi}$ method results in a marginal all-sky intensity of $(1.3 \pm 0.6)\,10^{-6}$ ph cm^{-2}s^{-1}sr^{-1}. In contrast, $\bar{\varphi}$ fitting assigns $\sim 7\%$ of the total number of observed counts to the isotropic component, corresponding to a mean all-sky intensity of $(4.0 \pm 0.3)\,10^{-4}$ ph cm^{-2}s^{-1}sr^{-1}; the mean background scaling factor $\tilde{\beta}$ is reduced to 0.93. Such a strong 1.8 MeV signal, which should be seen in COMPTEL high-latitude raw spectra, is not observed, and consequently should be interpreted as systematic error. Hence, fixed $\bar{\varphi}$ fitting should be preferred in this case.

The question, whether the $\bar{\varphi}$ distribution should be fitted or not can be addressed by the tools of Bayesian model comparison. For the example of the isotropic sky intensities, the small background scaling factors assigned by $\bar{\varphi}$ fitting lead to a huge penalization by the background prior, which in turn leads to a clear favour of the fixed $\bar{\varphi}$ method over $\bar{\varphi}$ fitting. In contrast, for the exponential disk model, the fixed $\bar{\varphi}$ result is punished by the scale height prior, and the Bayes factor analysis gives clear support for $\bar{\varphi}$ fitting.

To illustrate how the prior knowledge may influence the decision, the ideal case is considered one last time. In this case, the background model is known in advance to be correct, hence using $\bar{\varphi}$ fitting to estimate the model parameters should be an introduction of unnecessary parameters. Thus, in spirit of William of Ockham, $\bar{\varphi}$ fitting should be discarded. Figure 4.8 shows the log-Bayes factor of the fixed $\bar{\varphi}$ method against $\bar{\varphi}$ fitting as a function of the width σ of the β prior.[11] Positive log-Bayes factors favour the fixed $\bar{\varphi}$ method, negative factors give preference to $\bar{\varphi}$ fitting.[12] The shape of the curve is readily understood: If σ is smaller than the largest statistical uncertainty on the individual background scaling factors, $\bar{\varphi}$ fitting is penalized by the statistical fluctuations of $\beta_{\bar{\varphi}}$ which leads to the steep increase of the log-Bayes factor for $\sigma_{\bar{\varphi}} < 4.2\,10^{-3}$. The penalization is overlayed by the general trend of an increasing favour for the fixed $\bar{\varphi}$ method with increasing prior width σ. This trend simply reflects the fact that with increasing σ the β prior is spread over larger regions of parameter space which reduces the amount of prior probability contained in the high-likelihood region. Consequently, the Ockham factor for $\bar{\varphi}$ fitting decreases, leading to an increasing favour of the fixed $\bar{\varphi}$ method (cf. Eq. (4.18)). For any choice of σ, however, the Bayes factor analysis leads to a favour for fixed $\bar{\varphi}$ parameter estimation over $\bar{\varphi}$ fitting. This is exactly as expected since the correct background model is the

[11] The prior knowledge for the ideal case differs from realistic applications in that the background model is known exactly. Hence independent Gaussians with means of unity and standard deviations σ were chosen as priors for the individual background scaling factors $\beta_{\bar{\varphi}}$.

[12] In contrast to orthodox statistics, where hypotheses are only rejected but never accepted, Bayesian analysis can favour a hypothesis against others.

fixed $\bar{\varphi}$ model. Apparently, Ockham's razor works perfectly in that it discards the more 'complicated' $\bar{\varphi}$ fitting model.

It is interesting to compare this result to classical statistics. One might be tempted to use the maximum likelihood ratio test (cf. Section 4.1.2) to judge if $\bar{\varphi}$ fitting is superior to the fixed $\bar{\varphi}$ method. The maximum likelihood ratio of $\bar{\varphi}$ fitting against the fixed $\bar{\varphi}$ method is $\Lambda = 34.6$. Since both models belong to the same continuous family of hypotheses, where $q = 20$ parameters are fixed for the fixed $\bar{\varphi}$ model, Λ is asymptotically distributed as χ^2_{20} under the assumption that the fixed $\bar{\varphi}$ background model is the true one. Hence, the maximum likelihood ratio converts to a rejection probability of 97.8% or 2.3 Gaussian σ. Taken literally, if the fixed $\bar{\varphi}$ method is correct, the observed (or larger) likelihood ratios should occur by chance in only 2 of 100 cases. However, it is notorious for classical statistics that the rare things happen so often (Raftery 1995), and most orthodox analysts would interpret 2.3σ not as strong evidence against the fixed $\bar{\varphi}$ background model. However, they are left with the open question which of both methods they should favour.

The last topic addressed in this section is the control of the fit quality. For this purpose, some verification procedures are carried out for each fit, which are illustrated for the 'ideal case' in Fig. 4.9. The upper two panels show the significance and flux of point source residuals on top of the model for galactic latitudes $|b| < 30°$. They were derived by independent fits of point source models along with the now fixed source and background models for a grid of positions with longitude and latitude spacing of 1° (i.e. 21 600 grid points). This provides a 'sky-map' of point source significances (top panel) which were estimated using the maximum likelihood ratio test for $q = 1$ degrees of freedom (de Boer *et al.* 1992). Corresponding flux estimates are shown in the second panel, where positive residuals are indicated by solid, negative residuals by dashed contours. Due to the telescope's angular resolution of 4° FWHM (at 1.8 MeV), neighboring 'sky-map' pixels are not independent and consequently residuals will appear somewhat extended. In particular, diffuse residuals will show up in these maps as connected irregular low significance features.

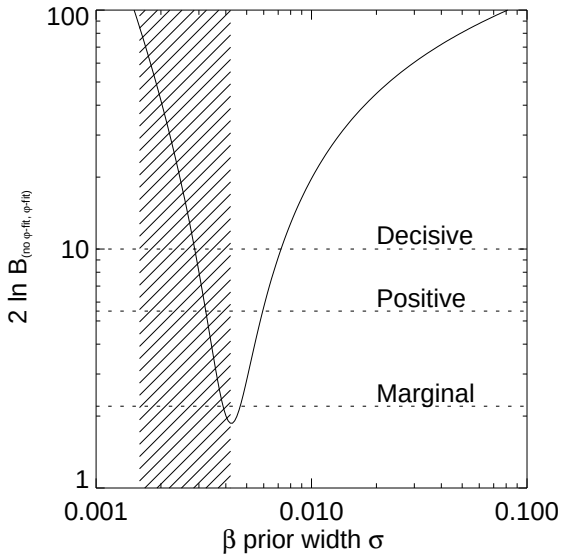

Fig. 4.8: *Log-Bayes factor versus width σ of the β prior. The hatched area indicates the error range on the individual background factors for $\bar{\varphi}$ fitting. Dotted lines represent the decision criteria of Tab. 4.1.*

The two lower panels show a longitude scan along the galactic plane ($b = 0°$) which

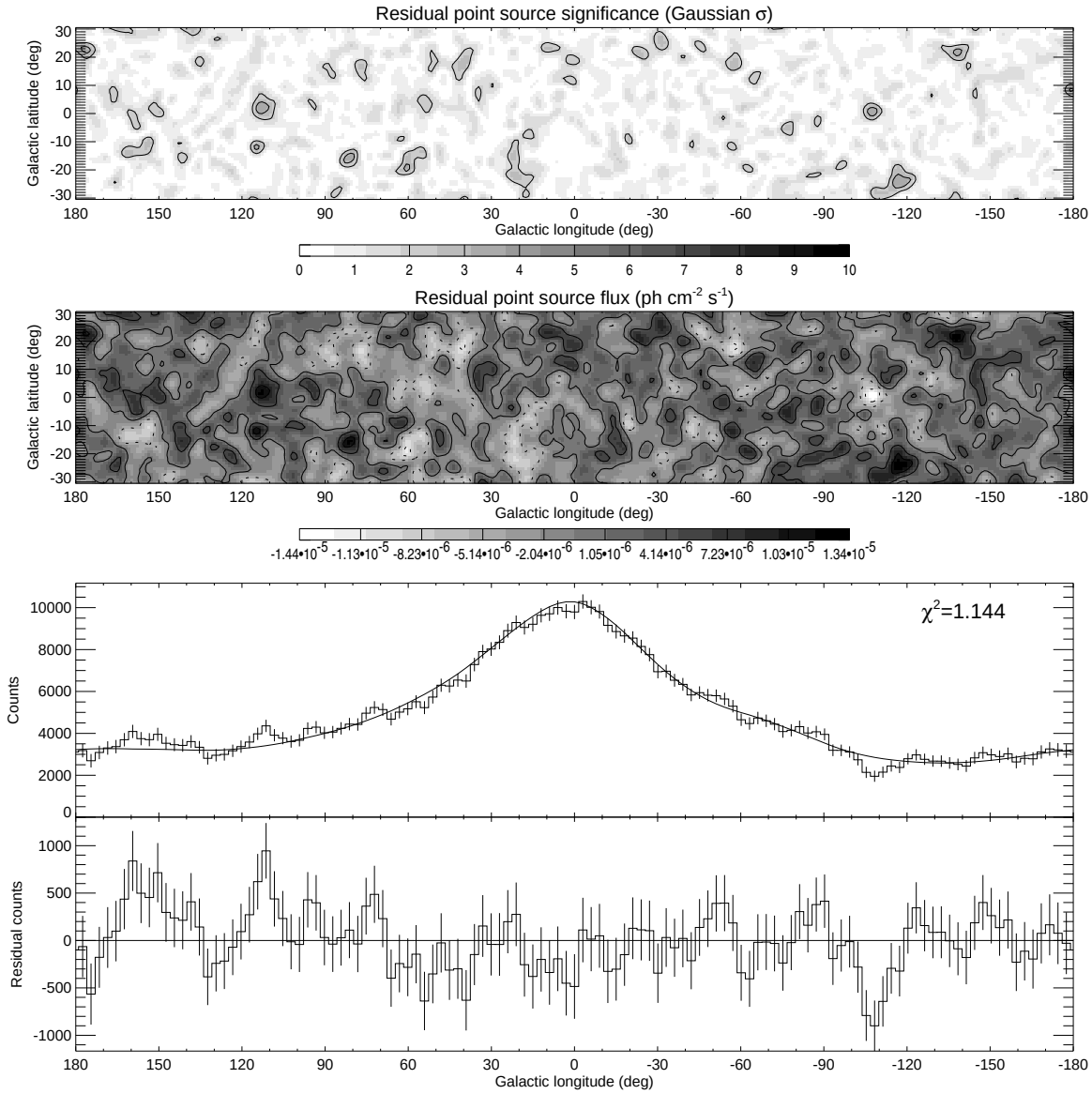

Fig. 4.9: Model fit verification for the 'ideal case' (see text).

was derived by means of the software collimation technique (cf. Section 2.5.1.2). This technique, which is normally used for the generation of COMPTEL count spectra, allows a direct visualization of the fit residuals in the COMPTEL data space. For a given scan position χ_0, ψ_0 on the sky, the events of all data space cells with an angular resolution measure

$$\mathrm{ARM} = \bar{\varphi} - \varphi_{\mathrm{geo}} = \bar{\varphi} - \arccos\left\{\sin\psi \sin\psi_0 + \cos\psi \cos\psi_0 \cos(\chi - \chi_0)\right\} \qquad (4.33)$$

of less than $\pm 3°$ are added, which corresponds to an integration of the observed counts along the PSF cone for a given source position. In parallel, the corresponding number

of predicted events in the source and the background models are determined. The background-subtracted observed counts together with the scan profile of the source model are shown in the first of the two lower panels for a longitude stepsize of 3°. Statistical errors for each scan position were derived from counting statistics, the value of the reduced χ^2 statistics for the longitude scan is given in the upper right corner. The lower plot shows the residual counts after subtraction of both the background and the source models. It is important to note, however, that the software collimation technique has only poor angular resolution of $\sim 10°$ (FWHM), hence neighboring scan positions are not independent and χ^2 is only a formal quantity indicating the agreement of the observed longitude profile with that of the source model.

Since Fig. 4.9 shows the residuals for the 'ideal case', it illustrates what kind of residuals are expected on the basis of statistical uncertainties only. Low significance residual point sources appear in the maximum likelihood map (top panels) which, however, show no spatial correlation. Similar features are seen at corresponding positions in the longitude scan (bottom panels). In particular, the positive residual at $l \sim 110°$ and the negative excess around $l \sim -110°$ correspond to $\sim 3\sigma$ excesses in the maximum likelihood map, demonstrating that both methods provide consistent results.

4.3 A Multi-Wavelength Comparison of COMPTEL 1.8 MeV Data

After the introduction to model comparison techniques and the demonstration of their systematic uncertainties, COMPTEL 1.8 MeV data will now be compared to 'tracers' observed at other wavelengths. The aim of such a study is the identification of similarities between the 1.8 MeV intensity distribution and emission at other wavelengths, which may help to identify the origin of galactic ^{26}Al. In previous works (Diehl *et al.* 1995b, 1997) only specific tracers like CO or the DIRBE 100 μm far-infrared map were considered, which has led to the conclusion that the ^{26}Al emission can not be described by a large-scale galactic emission model, and that multi-component models are required to explain the particularities of the 1.8 MeV distribution (Diehl *et al.* 1996).

In this work, the comparison will be extended to the entire observed wavelength range, starting at the $\lambda \sim 10$ m radio band up to the energetic $E > 100$ MeV γ-ray photons. In this way, it is the most unbiased tracer map comparison which can be imagined since the 1.8 MeV data is confronted with the entire knowledge about possible emission distributions on the sky. It will turn out that the 1.8 MeV distribution follows very closely the distribution of free electrons in the Galaxy which is traced by thermal bremsstrahlung, observable in the microwave domain. In contrast to previous works, it will be shown that no additional components are needed to describe the 'irregularities' found in 1.8 MeV sky-maps. Using the technique of Bayesian model averaging, a total flux of 3.1 ± 0.1 ph cm^{-2}s^{-1} will be derived for the central steradian, which is slightly

higher than flux determinations from reconstructed intensity distributions (2.8 ± 0.3 ph $cm^{-2}s^{-1}$ for both R-L and ME).

4.3.1 Tracer Maps

In the following sections, the different tracer maps which were used for comparison with COMPTEL 1.8 MeV data are presented. To understand their (possible) relation to ^{26}Al, the primary emission mechanisms which underlie the observed radiation will be introduced briefly. In addition to diffuse galactic emission, some tracer maps contain instrumental or extragalactic diffuse background, or strong, mainly extragalactic point sources. When possible, these components were removed from the tracer maps, which led in all cases to an improvement of the fit to the data. The preparation of each map for COMPTEL analysis is discussed. All tracer maps are compiled in Tab. 4.2.

4.3.1.1 The 408 MHz Survey

At radio frequencies, the lowest frequencies which are observed in astrophysics, synchrotron radiation of relativistic cosmic-ray electrons, accelerated by magnetic fields, is the dominant source of galactic continuum emission. The synchrotron emissivity is proportional to the product of the cosmic-ray electron density and approximately the square of the magnetic field intensity B^2. The large scale distribution of B^2 may in turn possibly be related to the distribution of interstellar gas (Haslam *et al.* 1981).

The only all-sky survey available which traces galactic synchrotron radiation at the same frequency and with similar telescope beamsize, is the Bonn 408 MHz survey of Haslam *et al.* (1982). The outstanding feature of the sky map is diffuse extended emission from the plane of the Galaxy, with intensity steps which are probably related to the position of galactic spiral arms (Phillipps *et al.* 1981a, and references therein). In addition, Phillipps *et al.* (1981b) find that a non-spherical halo extending ~ 10 kpc from the plane with an emissivity $\lesssim 10\%$ of that in the plane is required to account for the high latitude brightness temperature. Superimposed on the disk and halo components are ridges of enhanced emission extending to high galactic latitudes, which are generally regarded as local features due to old supernova remnants (designated Loop I-IV by Berkhuijsen *et al.* 1971). For the analysis, the isotropic extragalactic background, which arises from cosmic background radiation and the sum of unresolved distant radio sources, was removed by subtracting 6 K from the observed antenna temperatures (Phillipps *et al.* 1981a). Additionally, the brightest point-like sources in the map, mostly radio galaxies, were removed by estimating the diffuse galactic intensity at the source position from the mean intensity in a ring around the source.

4.3.1.2 Atomic Neutral Hydrogen (H I)

Diffuse gas in the interstellar medium, which is essentially composed of hydrogen, either in atomic or molecular form, is observed to exist in three dominant phases: a cold ($T \gtrsim 50$ K) neutral medium (CNM), a warm ($T \sim 8000$ K) medium, both neutral (WNM) and ionized (WIM), and a hot ($T \sim 10^6$ K) ionized medium (HIM) (McKee and Ostriker 1977). Atomic neutral hydrogen (H I), which occurs in the CNM and WNM phase, is readily observed in emission by its 21 cm hyperfine line at 1420.4058 MHz. Indeed, this is the strongest thermal spectral line in radio astronomy (Dickey and Lockman 1990). In this work, the H I map presented by Dickey and Lockman (1990) is used which is given in units of hydrogen column density $N_{\rm H}$, assuming optically thin emission. The map is dominated by galactic plane emission, which has a FWHM of about 4° on average. Additionally, loops and arching filaments are observed at high galactic latitudes which have been swept up by supernovae.

4.3.1.3 The Sky at Microwave Wavelengths

With increasing frequencies, the galactic continuum emission at microwave wavelengths becomes a mixture of synchrotron radiation of relativistic electrons, free-free emission (thermal bremsstrahlung) from ionized matter in H II regions, and a small contribution of thermal dust emission. For the comparison of the microwave sky with 1.8 MeV data, the four year full sky maps derived by the Differential Microwave Radiometers (DMR) instrument aboard the *Cosmic Background Explorer (COBE)* at 31.5, 53, and 90 GHz were used (Kogut *et al.* 1996a, Bennett *et al.* 1994a). The intensity distribution for the two DMR channels A and B were averaged by weighting with the statistical uncertainties in individual sky map pixels. Since DMR measures only differential temperatures, the zero level in the sky maps is poorly defined. Assuming that no galactic emission is present above galactic latitudes of $|b| > 40°$, the zero level was determined from the mean intensity in the DMR maps above this latitude. Note, that the DMR horn antennas have a full-width at half power of 7°, hence the angular resolution of the sky maps is inferior to that of COMPTEL data.

The different spatial and spectral morphology of the dominant emission processes may be used to separate these components from DMR microwave data. Using a model of galactic synchrotron emission, based on 408 and 1420 MHz radio surveys, and a model of thermal dust emission, based on *COBE* FIRAS data, Bennett *et al.* (1992) created the first all-sky map of free-free emission. This map, which traces the distribution of ionized matter throughout the entire Galaxy, will also be compared to the 1.8 MeV data. The synchrotron and dust models used by Bennett *et al.* (1992) for the component separation will be included in the comparison as reference.

4.3.1.4 Molecular Hydrogen (H_2)

An important fraction of the interstellar gas, probably about half, appears to be in molecular form (Bronfman 1992). Most of the mass ($\sim$ 90%) of the molecular ISM is gathered in giant molecular clouds (GMCs), which are believed to be the sites of most current star formation in the Galaxy (Combes 1991). Since the symmetric H_2 molecule, which is by far the most abundant molecule in the ISM, does not radiate at radio frequencies, the asymmetric CO molecule, the most abundant molecule after H_2, provides the best tracer of molecular clouds in the Galaxy. This is because, in the cold ISM, the CO molecule is excited to the first rotational level ($J = 1 \rightarrow 0$) by collisions with H_2 molecules, giving rise to a microwave spectral line at 115 GHz or $\lambda = 2.6$ mm. The CO emission intensity turns out to be proportional to the H_2 mass, hence it is a valuable tool for studies of the molecular mass distribution in the Galaxy (see Combes 1991, for a discussion of the uncertainties in the H_2/CO conversion ratio). Only in the vicinity of the galactic center, the H_2 column density may be significantly lower than that inferred from CO observations (Sodroski *et al.* 1995, and references therein).

The most complete survey of CO emission is the Columbia / Cerro-Tololo survey of Dame *et al.* (1987) which covers the entire Milky Way with latitude extension from $|b| \leq 10^\circ - 30^\circ$. Due to the uncertainties about the H_2/CO conversion ratio in the galactic center, two different H_2 models were used for comparison: the first consists of the original CO map of Dame *et al.* (1987) which shows a strong intensity peak between galactic longitudes $-4.5^\circ < l < 4.5^\circ$. For the second map, this peak was removed by estimating the underlying diffuse emission from the longitude intervals $4.5^\circ < |l| < 5.5^\circ$ for each latitude between $|b| < 1.5^\circ$ separately.

4.3.1.5 Far-Infrared Continuum Emission

In the far-infrared (FIR) wavelength range from $\lambda \sim 40 - 1000$ μm, for which the Galaxy is optically thin, the galactic large-scale emission arises primarily from the cooling of radiatively heated large dust grains, being in thermal equilibrium with the ambient radiation field. Most of the galactic FIR luminosity (55% – 85%) originates from cool ($T_d \lesssim 22$ K) dust that is primarily heated by the Galaxy's ambient interstellar radiation field (ISRF) (Sodroski *et al.* 1997). Dust that is associated with OB stars that are embedded in or in the vicinity of their progenitor molecular clouds or that are ionizing the galactic extended low-density H II regions contributes 15% – 45% of the total galactic FIR luminosity (Sodroski *et al.* 1997, Kogut *et al.* 1996b,c).

With its 10 photometric bands from 1 to 240 μm, the Diffuse Infrared Background Experiment (DIRBE) aboard the *COBE* satellite provides an almost complete coverage of the infrared (IR) wavelength range with adequate angular resolution of 0.7° (e.g. Boggess *et al.* 1992). Thermal emission of dust associated with the solar system, also known as zodiacal light, contributes 12%, 1%, 0.4%, and 0.2% of the total emission in the inner Galaxy in the FIR continuum bands at 60, 100, 140, and 240 μm, respectively

(Sodroski *et al.* 1997). For the 1.8 MeV comparison, zodiacal light subtracted FIR maps were used, which were kindly provided by Bernard *et al.* (1995).

4.3.1.6 Far-Infrared Lines

Embedded in the FIR continuum are spectral lines arising from transitions of atoms in the gaseous phase of the ISM. The most prominent of these gas cooling lines are the 158 μm ground state transition of C^+, the dominant heat loss mechanism for neutral gas in the Galaxy, and the 205 μm transition of N^+, tracing the large-scale low-density extended ionized component of the Galaxy (Bennett *et al.* 1994b).

Both lines have been successfully mapped by the Far-Infrared Absolute Spectrophotometer (FIRAS) aboard the *COBE* satellite with an effective angular resolution of 7° (Wright *et al.* 1991, Bennett *et al.* 1994b). The C^+ emission follows the FIR continuum emission rather closely because the dominant source of heat for the interstellar gas is provided by photoelectrons ejected from grains by UV photons (Wright *et al.* 1991, Watson 1972). The N^+ intensities are related to C^+ intensities via $I(C^+) \propto I(N^+)^{1.5}$ which is understood as a volume/surface area ratio geometric effect: while the volume of very large internally ionized regions give rise to N^+ emission, these regions are partially surrounded by neutral gas with the interface surface traced by C^+ emission (Bennett *et al.* 1994b). Note, however, that both maps are subject to substantial statistical uncertainties, and $\sim 5\%$ of the sky is not covered by the data (in particular, a bright emission feature in the Cygnus regions suffers from incomplete coverage).

4.3.1.7 Mid- and Near-Infrared Continuum Emission

In the mid-infrared (MIR) wavelength range between 5 μm$< \lambda <$ 40 μm, the infrared sky is dominated by zodiacal light emission. Additionally, extensive dust emission is seen from the Galaxy which is believed to come from small grains (≤ 0.01 μm) briefly heated to high temperatures ($T \sim 100$ K) each time they absorb a single photon (Boulanger and Pérault 1988). As a third component, low-mass AGB stars are the most important point source component of the MIR sky, contributing about $3\% - 4\%$ of the total emission in the 12 and 25 μm bands (Wainscoat *et al.* 1992, Sodroski *et al.* 1997). Due to their mass loss, these stars are to varying degrees surrounded by dust, leading to peak emission in the MIR wavelength range.

In the near-infrared (NIR) domain (1 μm $< \lambda <$ 5 μm), zodiacal and interstellar dust emission declines and stellar emission becomes the dominant source of IR radiation. The stellar populations, mainly late K and M giants, form a superposition of two components: (1) a probably bar-shaped bulge with typical axis ratios of $1 : 0.3 : 0.2$ and scale length between $1 - 2$ kpc for the major axis (Dwek *et al.* 1995, Weiland *et al.* 1994), and (2) an exponential disk with radial scale length between $R_0 = 1 - 3$ kpc and vertical scale height of $z_0 = 90 - 325$ pc (Wainscoat *et al.* 1992, and references therein). With decreasing wavelength, however, extinction in the galactic plane

becomes important, which limits the use of the NIR maps as tracer of galaxywide emission.

For the analysis, DIRBE all-sky maps derived at wavelengths 25, 12, 4.9, 3.5, 2.2, and 1.25 μm were used. The MIR bands at 12 and 25 μm are both heavily contaminated by zodiacal light which is 61% and 68% of the total emission in the inner Galaxy, and 90% and 93% in the outer Galaxy (Sodroski *et al.* 1997). Unfortunately, no zodiacal light corrected maps were available at the moment of writing this thesis, and the uncorrected DIRBE all-sky maps were used for the analysis.

4.3.1.8 Visible Light

The distribution of visible light (400 nm $< \lambda <$ 700 nm) fits continuously to the NIR all-sky maps. Owing to the strong obscuration by interstellar dust, the light is primarily from stars within $\sim$ 1 kpc of the Sun. For the analysis, the Lund all-sky panorama of the Milky Way was used. Note, however, that this panorama is based on a drawing under the supervision of Knut Lundmark, hence cannot be taken as a photometrically correct representation of the sky. Therefore, the results for this tracer map should not be taken too literally.

4.3.1.9 Soft X-rays

The soft X-ray sky (0.15 keV $< E <$ 3 keV) is heavily affected by the presence of the interstellar medium which is only partially transparent in this domain. Photoelectric absorption by interstellar gas reduces the mean free path of $\frac{1}{4}$ keV X-rays to $\sim$ 100 pc within the galactic plane, resulting in mainly local galactic emission at these energies. For 1.5 keV photons, the mean free path increases to $\sim$ 3 kpc in the plane, leading to a mixture of galactic emission from nearby and distant regions (Snowden *et al.* 1997). Amongst the most prolific galactic soft X-ray sources are supernova remnants (SNRs) and X-ray binaries, where X-rays probably arise from thermal plasma radiating at temperatures in the range $10^6 - 10^7$ K. But the zoo of soft X-ray sources is much more extensive, including the moon, comets, stars, cataclysmic variables, white dwarfs, pulsars, neutron stars, globular clusters, galaxies, active galactic nuclei (AGNs), clusters of galaxies, and the X-ray background, until recently one of the long-standing puzzles of modern astrophysics (e.g., Hasinger *et al.* 1993).

For the analysis, the ROSAT all-sky maps of Snowden *et al.* (1993) were used, which cover $\sim$ 98% of the sky in the $\frac{1}{4}$ keV, $\frac{3}{4}$ keV, and 1.5 keV bands. Although discrete X-ray sources have not been removed from the maps, the angular resolution of $\sim 2^\circ$ emphasizes low surface brightness extended features like nearby SNRs and the X-ray background. Due to the opacity of interstellar gas at $\frac{1}{4}$ keV, the X-ray background is heavily absorbed towards the galactic plane, leading to a strong anti-correlation of the $\frac{1}{4}$ keV intensity with the column density of galactic neutral hydrogen

(Snowden *et al.* 1993). Superimposed on this is the emission from two nearby SNRs, the Cygnus Loop, and the Vela SNR, and a diffuse unabsorbed foreground component which probably arises from hot ($T \sim 10^6$ K) gas in the local interstellar medium (Snowden *et al.* 1997). The other soft X-ray bands show a completely different angular structure due to the increased transparency of the Galaxy. Although some residual absorption along the galactic plane remains, point sources, mainly X-ray binaries, become visible. Additionally, diffuse emission is seen in the galactic center hemisphere, which at $\frac{3}{4}$ keV arises from the Loop I complex (Egger 1995), and at 1.5 keV is probably due to a galactic X-ray bulge (Snowden *et al.* 1997). At high galactic latitudes, most of the flux is extragalactic in origin (Hasinger *et al.* 1993).

4.3.1.10 Hard X-rays

At hard X-ray energies (3 keV $< E <$ 400 keV) the entire Galaxy becomes fully transparent to X-ray photons. The most prominent galactic sources seen in the hard X-ray sky are X-ray binaries, i.e. systems where gas is flowing through an accretion disk from a 'normal' star onto a neutron star or black hole. The X-ray emission arises from the hot ($T \sim 10^9$ K) thermal plasma of the accretion disk which is heated by the release of gravitational energy of the infalling matter. Further bright hard X-ray sources comprise SNRs, AGNs, and cluster of galaxies. Additionally, a highly isotropic diffuse X-ray background is detected, which possible arises from a superposition of yet unresolved AGNs (Setti 1995).

For the hard X-ray domain, the all-sky map constructed from the combination of the HED3 and MED detectors of the *HEAO-1* A2 experiment was used (Rothschild *et al.* 1979). The combination of these detectors is sensitive from $2-60$ keV. Conventionally all-sky maps from this combination are analyzed in a number of standard colors defined as weighted sums of counts from certain combinations of detector layers and pulse height channels (discovery scalers; see Allen *et al.* 1994). In this work, the R15 color is used which is the most stable color, since it is a sum of discovery scalers which did not change during the mission. Instrumental background was removed from the maps using the difference between the large and small FOV collimator configurations as described in Allen *et al.* (1994). To remove the isotropic cosmic background, a constant intensity was subtracted from the skymap which was estimated from regions at high galactic latitude which are devoid of prominent point sources. Additionally, some prominent extragalactic point-like sources were removed from the map.

4.3.1.11 High-Energy γ-rays

The outstanding feature of the high-energy gamma-ray sky (E > 100 MeV) is diffuse radiation from the galactic plane. Most of the emission originates from π°-decay following the inelastic collisions between cosmic-ray (CR) nuclei (mainly protons) and the interstellar gas (minor contributions come from inverse Compton scattering of CR

Name	Emission mechanism	Tracer of
408 MHz	synchrotron	cosmic rays / magnetic field
21 cm	H hyperfine transition	neutral hydrogen
†DMR synch.	synchrotron	cosmic rays / magnetic field
DMR 31.5 GHz	free-free / synchrotron	ionized gas / cosmic rays
DMR 53 GHz	free-free / synchrotron	ionized gas / cosmic rays
DMR 90 GHz	free-free / synchrotron	ionized gas / cosmic rays
DMR free-free	free-free	ionized gas
*CO	CO rotational transition	molecular gas
†DMR dust	thermal dust	warm dust
DIRBE 240 μm	thermal dust	warm dust ($T \sim 12$ K)
FIRAS 205 μm	N^+ ($^3P_1 - ^3P_0$)	low-density ionized gas
FIRAS 158 μm	C^+ ($^2P_{3/2} - ^2P_{1/2}$)	neutral gas cooling
DIRBE 140 μm	thermal dust	warm dust ($T \sim 21$ K)
DIRBE 100 μm	thermal dust	warm dust ($T \sim 30$ K)
DIRBE 60 μm	thermal dust	warm dust ($T \sim 50$ K)
DIRBE 25 μm	thermal dust / star light	warm dust ($T \sim 120$ K) / AGBs
DIRBE 12 μm	thermal dust / star light	warm dust ($T \sim 250$ K) / AGBs
DIRBE 4.9 μm	star light	stars (K and M giants)
DIRBE 3.5 μm	star light	stars (K and M giants)
DIRBE 2.2 μm	star light	stars (K and M giants)
DIRBE 1.25 μm	star light	stars (K and M giants)
Visible	star light	stars
ROSAT 1/4 keV	thermal bremsstrahlung	local hot gas
ROSAT 3/4 keV	thermal bremsstrahlung	hot gas
ROSAT 1.5 keV	thermal bremsstrahlung	hot gas / X-ray binaries
HEAO-1	thermal bremsstrahlung	X-ray binaries
EGRET	nuclear interactions	interstellar gas / cosmic rays

Table 4.2: *Tracer maps used for 1.8 MeV model comparison. Maps marked by † are the models of Bennett* et al. *(1992) which were used to extract free-free emission from DMR data. For maps marked by * several modifications of the original data were compared to the 1.8 MeV sky (see text). Specified dust temperatures were calculated from the Wien law.*

electrons off galactic photon fields and CR electron bremsstrahlung). It turned out that the diffuse γ-ray intensity shows a fairly good quantitative correlation with the galactic gas density (H I and H_2) (e.g., Strong *et al.* 1988).

For the analysis, a maximum entropy deconvolved EGRET all-sky map for energies E > 100 MeV was used, which was kindly provided by M. Pohl (MPE). To account for cosmic diffuse background radiation, an isotropic intensity of $1.5\,10^{-5}$ ph $cm^{-2}s^{-1}sr^{-1}$ was subtracted from the map (Kniffen *et al.* 1996). In addition, to provide a map of the galactic diffuse emission, all sources from the Second EGRET Catalog (Thompson

et al. 1995) with a significance greater than 4σ were removed.

4.3.2 Comparison of Tracer Maps

4.3.2.1 Prior Probabilities

Before the tracer maps can be compared to COMPTEL 1.8 MeV data, prior probabilities have to be assigned to all model parameters. Since the intensity distribution for the source models is fixed, the only free parameter for each map is a scaling factor. Prior knowledge about this scaling factor may be determined from previous observations of the 1.8 MeV line. Unfortunately, flux values obtained by other instruments are not quoted in a consistent manner. In general, intensities were estimated assuming a longitude profile for the 1.8 MeV emissivity (mostly the COS-B high-energy γ-ray distribution integrated over an unspecified range of latitudes) which was convolved with the instrumental response and compared to the observed line strength. Results are then given as mean or peak intensity of the longitude profile towards the galactic center, and lie in the range of $(4.5 \pm 2.0)\,10^{-4}$ ph cm^{-2}s^{-1}rad^{-1} (Prantzos and Diehl 1996). Most instruments, however, were large aperture collimators for which the observed intensity depends on the total flux in the field of view. For the present tracer map comparison it is therefore much more convenient to transform the quoted values to model independent mean intensities towards the galactic center. This was done using the EGRET high-energy γ-ray distribution which was integrated over galactic latitudes of $|b| < 30°$ to provide a longitude profile in units of ph cm^{-2}s^{-1}rad^{-1}. The observed intensities of $(4.5 \pm 2.0)\,10^{-4}$ ph cm^{-2}s^{-1}rad^{-1} transform then to $(4.5 \pm 2.0)\,10^{-4}$ ph cm^{-2}s^{-1}sr^{-1}, which will be considered as prior knowledge about the magnitude of the expected 1.8 MeV intensity. Consequently, all tracer maps were scaled to obey the same intensity within the central steradian (i.e. within a circumference of 0.572 rad around the galactic center), and the prior on the scaling factor was chosen to be a Gaussian of mean $4.5\,10^{-4}$ ph cm^{-2}s^{-1}sr^{-1} and standard deviation $2\,10^{-4}$ ph cm^{-2}s^{-1}sr^{-1}. For the background scaling factor, the priors discussed in Section 4.2.2 were used.

4.3.2.2 Systematics

The results of the comparison of all tracer maps to COMPTEL 1.8 MeV data are compiled in Fig. 4.10. As expected, due to the additional number of free parameters, $\bar{\varphi}$ fitting always results in higher log-likelihood ratios than the fixed $\bar{\varphi}$ method. The comparison with the log-Bayes factors shows, however, that the Ockham factor can reverse the likelihood judgment. For the HEAO-1 map the Bayes factor analysis gives a decisive favour for the fixed $\bar{\varphi}$ background model against $\bar{\varphi}$ fitting. To understand why the HEAO-1 model prefers the fixed $\bar{\varphi}$ method, the χ^2 distances between the $\bar{\varphi}$ distributions of the tracer maps and the background model are plotted in Fig. 4.11 versus the log-Bayes factor of the fixed $\bar{\varphi}$ method against $\bar{\varphi}$ fitting. Positive log-Bayes

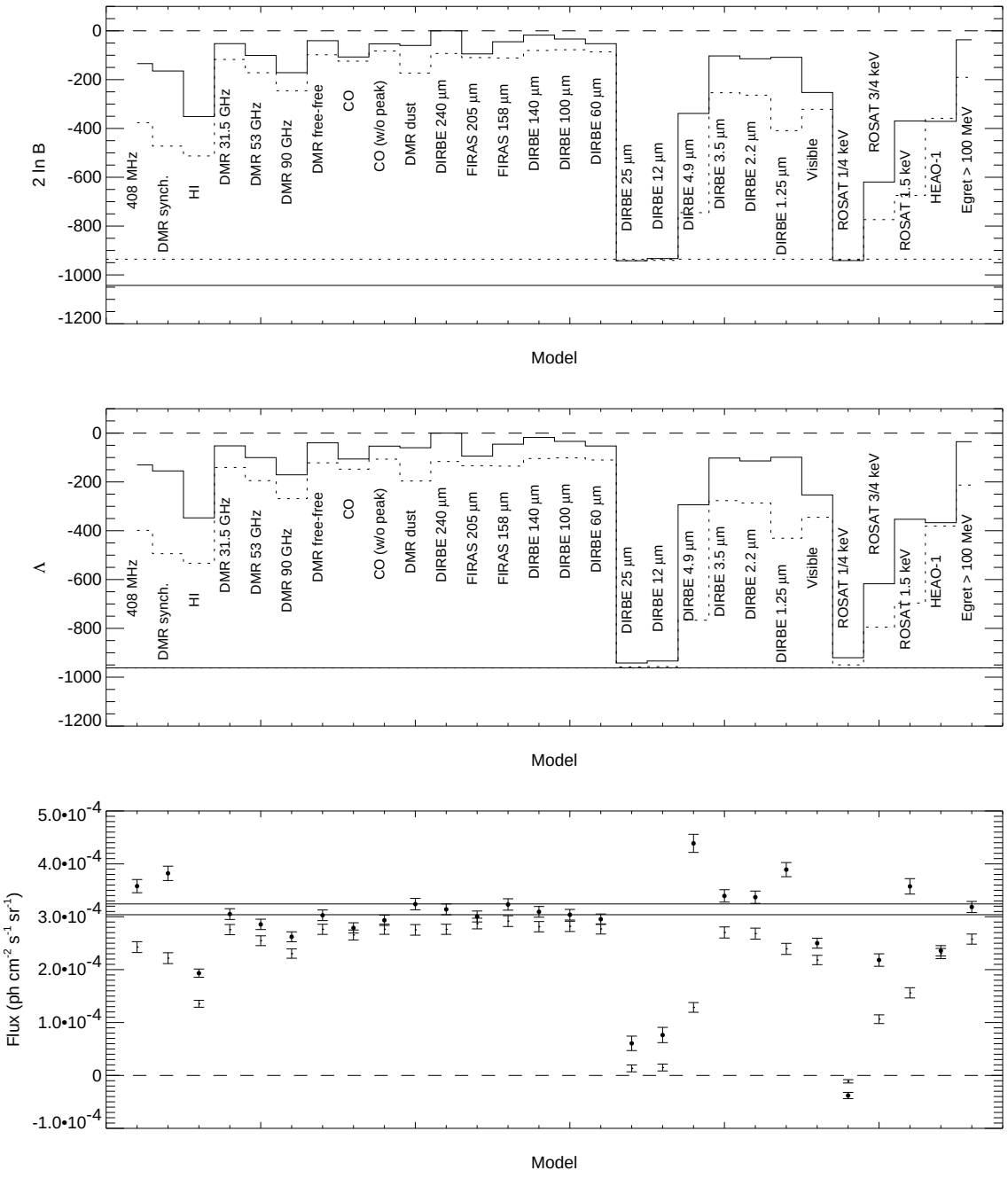

Fig. 4.10: *Comparison of all tracer maps.* Top panel: *Log-Bayes factor of all tracer maps against the best model (DIRBE 240 μm). The solid histogram corresponds to model fits using $\bar{\varphi}$ fitting, the dotted histogram are the results for the fixed $\bar{\varphi}$ method. The solid and dotted horizontal lines indicate the log-Bayes factors of the background model fitted to COMPTEL 1.8 MeV data alone (i.e. without source component). The dashed line indicates the log-Bayes factor for the best model.* Mid panel: *Corresponding log-likelihood ratios (the log-likelihood ratio of the background model fit is almost identical for both fitting methods).* Lower panel: *Assigned flux in the central steradian and the 1σ maximum likelihood ratio error bars. Error bars with dots indicate results obtained using $\bar{\varphi}$ fitting, those without dots give the results for the fixed $\bar{\varphi}$ method. The solid lines show the flux range derived from Bayesian model averaging, the dashed line indicates zero flux.*

factors give favour for the fixed $\bar{\varphi}$ method, negative factors prefer the use of $\bar{\varphi}$ fitting. Obviously, there is a strong correlation between both quantities which results in a decreasing favour for $\bar{\varphi}$ fitting with decreasing χ^2 distance. This is reasonable since the aim of $\bar{\varphi}$ fitting was to alleviate the error in the background model which arises from the normalization of its $\bar{\varphi}$ distribution to that of the data. For decreasing χ^2 distance the error introduced by the $\bar{\varphi}$ normalization becomes smaller, hence $\bar{\varphi}$ fitting leads to a declining improvement of the fit. In fact, the HEAO-1 map obeys the smallest χ^2 distance of all tracer maps and the Bayes factor analysis shows that $\bar{\varphi}$ fitting is only an introduction of unnecessary parameters. Consequently, $\bar{\varphi}$ fitting is rejected by Ockham's razor.

The only models which deviate from the strong correlation are the DIRBE 12 μm and 25 μm, and the ROSAT 1/4 keV maps. A look at Fig. 4.10 shows that these models provide extremely low fluxes for the fixed $\bar{\varphi}$ method. Since the error in the background model approximation depends not only on the $\bar{\varphi}$ distribution discrepancy but also on the source strength, it is obvious that the favour for $\bar{\varphi}$ fitting decreases with decreasing source flux (recall that in the case of zero source flux the background model approximation becomes exact).

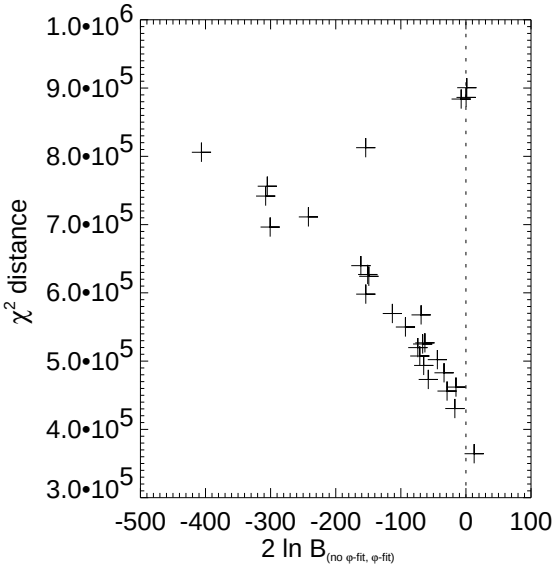

Fig. 4.11: *χ^2 distance between the $\bar{\varphi}$ distributions of all tracer maps and the data versus the log-Bayes factor of the fixed $\bar{\varphi}$ method against $\bar{\varphi}$ fitting.*

Nevertheless, also $\bar{\varphi}$ fitting is only an approximation of the true background model and there might still be some systematic bias in the model comparison. The remaining bias due to the $\bar{\varphi}$ approximation of the background model is visualized in Fig. 4.12, where the χ^2 distance is plotted versus the log-Bayes factors (left) and the log-likelihood ratios (right) against the best model. Obviously, as indicated by the large linear correlation coefficients, there is a clear correlation between χ^2 distance and log-Bayes factor or log-likelihood ratio, which favors models with $\bar{\varphi}$ distributions close to that of the data. The correlation is strongest for the fixed $\bar{\varphi}$ method and becomes weaker (but does not vanish) for $\bar{\varphi}$ fitting. If only 'good' fits are considered (those with log-Bayes factors or log-likelihood ratios which differ by less than 200 or 100 from the best model) the correlation is drastically reduced for $\bar{\varphi}$ fitting. Nevertheless, some small bias seems to remain in the comparison which suggests that log-Bayes factors should probably not be taken too literally. This does not mean that Bayesian analysis fails, but that the uncertainty of the background model introduces

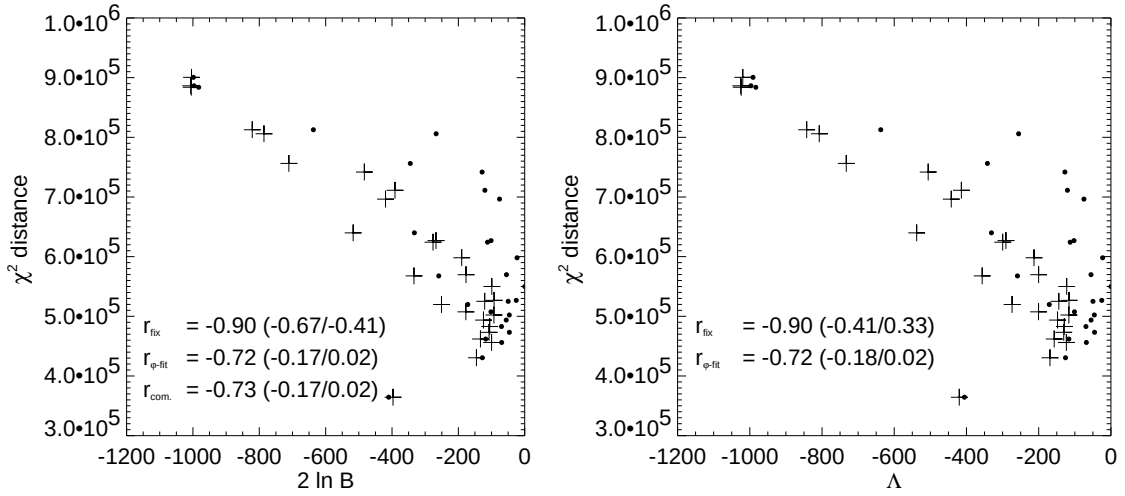

Fig. 4.12: Left: χ^2 *distance between the* $\bar{\varphi}$ *distributions of tracer maps and data plotted versus log-Bayes factor* $2\ln B$*. The fixed* $\bar{\varphi}$ *models are shown as 'plus' symbols,* $\bar{\varphi}$ *fitted models are plotted as dots. The linear correlation coefficient between the* χ^2 *distance and the log-Bayes factors is given for the two background model fitting methods (*$r_{\rm fix}$ *and* $r_{\bar{\varphi}-\rm fit}$*, respectively), and for the combination of the results using the favored fitting method for each model (*$r_{\rm com.}$*). Values in parentheses were obtained by considering only the highest log-Bayes factors with* $2\ln B - (2\ln B)_{max} \geq -200$ *and* -100*, respectively.* Right: χ^2 *distance between the* $\bar{\varphi}$ *distributions of tracer maps and data plotted versus maximum log-likelihood ratio* Λ*.*

some residual bias which may mislead inference (see also Section 4.3.2.3).

The influence of prior knowledge on the model comparison is illustrated in Fig. 4.13 where the assigned flux and (mean) background scaling factor β (or $\tilde{\beta}$) are shown as a function of the negative log-Ockham factor $-2\ln W$. Recall, that the Ockham factor measures on the one hand the simplicity of a model, and on the other hand the deviation of the model parameters from the prior means. Consequently, since they are 'simpler', the fixed $\bar{\varphi}$ models show systematically larger Ockham factors ($-2\ln W = 11 - 25$) than the $\bar{\varphi}$ fitted models ($-2\ln W = 35 - 56$). Fixed $\bar{\varphi}$ models show a clear trend of increasing Ockham factor with increasing flux which obviously is due to the flux prior. All fluxes lie below the prior mean of $4.5\,10^{-4}$ ph cm^{-2}s^{-1}sr^{-1}, hence an increase in flux leads to an increase of the prior probability. For $\bar{\varphi}$ fitting the situation is different. The mean background scaling factor shows large deviations from unity which are penalized by the background prior. Since small background scaling factors induce high fluxes, a trend of decreasing Ockham factor with increasing flux appears above $\sim 3\,10^{-4}$ ph cm^{-2}s^{-1}sr^{-1}.

It is worthwhile to note that the (mean) background scaling factor almost always falls below unity which suggests that the background model normalization is probably somewhat too high. Bayesian model averaging (BMA, cf. Section 4.1.4) yields a MAP background scaling factor of $\beta = 0.9946 \pm 0.0005$, but it is only one model (DIRBE

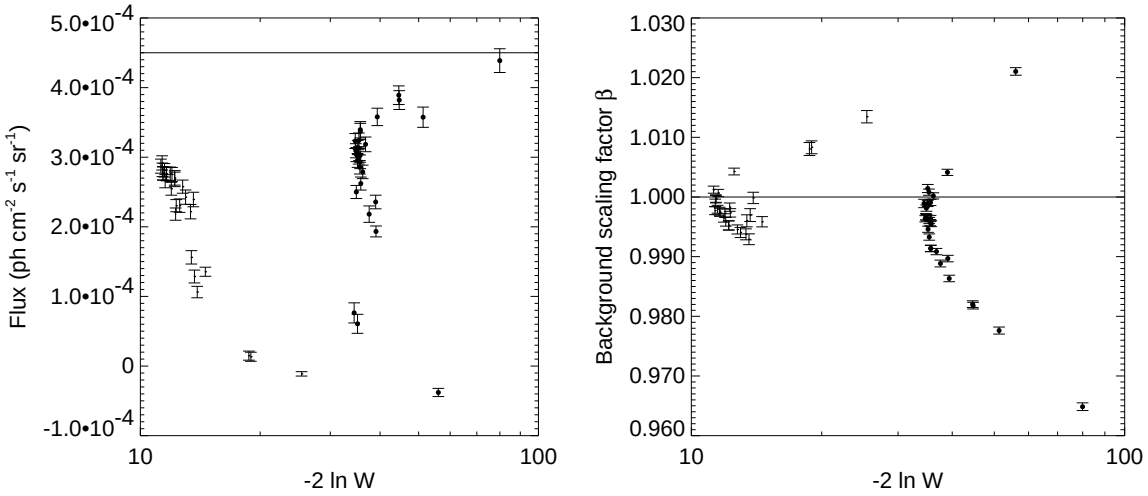

Fig. 4.13: Left: *Assigned model flux within the central steradian versus the negative log-Ockham factor for all tracer maps. The solid line at* $4.5\,10^{-4}$ *ph* $cm^{-2}s^{-1}sr^{-1}$ *indicates the mean of the flux prior.* Right: *Background scaling factor* β *or mean background scaling factor* $\tilde{\beta}$ *versus the negative log-Ockham factor for all tracer maps. The solid line at* $\beta = 1$ *represent the mean of the background prior.*

240 μm) which contributes significantly to this average. Due to the uncertainty of the systematic bias of the analysis, one may alternatively average the background factor with equal weight for all models with sufficiently high Bayes factor (this approach will be called the Ockham window average (OWA)[13]). Using all models with log-Bayes factors above $2\ln B \geq -60$ results in $\beta = 0.9963 \pm 0.0024$ which is consistent with the BMA estimate but significantly below unity. Although the discrepancy from unity may appear small, it raises the number of 1.8 MeV source counts from 37 476 for $\beta = 1$ to $52\,000 \pm 9\,400$ events. Note that the lower background factors agree with the results anticipated from the ψ distribution of the data versus the background model (cf. Section 2.5.2.2).

4.3.2.3 The Best Model

Figure 4.10 shows that the highest Bayes factors are obtained for the DIRBE far-infrared, the DMR free-free emission, and the EGRET high-energy γ-rays maps. Mod-

[13]The Ockham window was introduced by Madigan and Raftery (1994) to reduce the number of models for Bayesian Model Averaging to a manageable number by selecting only data-supported models. In this work a 'numerical' Ockham window will be applied which is simply related to the fact that the posterior $P(M|D,I) \sim B = \exp(\ln B)$ is of the order of the exponential of the log-Bayes factor, hence becomes smaller that a computer's floating point accuracy for $\ln B \lesssim -100$. For Ockham window averaging, all models with $2\ln B$ larger than a given limit are used with equal weight to account for unknown systematic uncertainties. If both fixed $\bar{\varphi}$ and fitted $\bar{\varphi}$ models lie above the limit, the BMA average of both models is used.

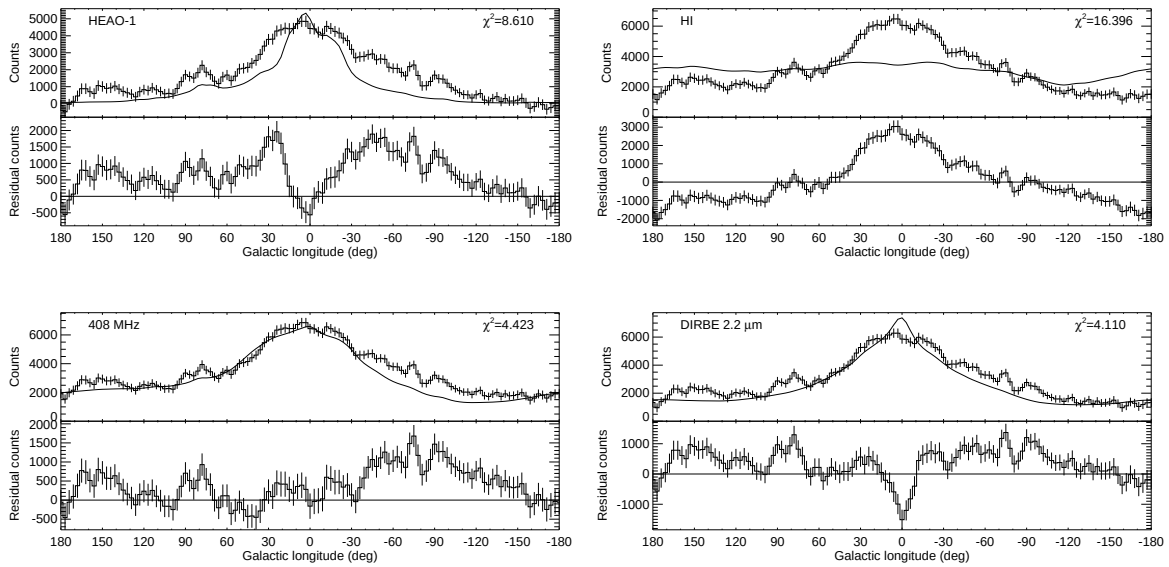

Fig. 4.14: *Software collimator scans for the $\bar{\varphi}$ fit of the HEAO-1 (upper-left), H I (upper-right), 408 MHz (lower-left), and DIRBE 2.2 μm tracer maps. The histogram shows observed 1.8 MeV counts minus the background model, the solid line is the convolved tracer map model. The χ^2 statistics for the longitude scans is given in the upper right corner of the plots.*

els with an essential high galactic latitude emission, like the DIRBE 4.9 μm - 25 μm and ROSAT 1/4 keV - 1.5 keV maps, are clearly rejected by the data. In particular, although two strong sources in the Cygnus and Vela region dominate the 1/4 keV ROSAT soft X-ray band, the fit assigns a negative flux to this tracer map. This reflects the global anticorrelation between soft X-rays and interstellar gas which results in the Galaxy seen in absorption at 1/4 keV. Since the Galaxy is the dominant feature in the 1.8 MeV data, the best fit is obtained for an inverted 1/4 keV map.

The HEAO-1 hard X-ray map, which is mainly composed of point sources that are concentrated towards the galactic center, provides only a poor fit to the data. This is illustrated in Fig. 4.14 where the software collimator longitude scan is shown for the corresponding $\bar{\varphi}$ fit. Obviously, hard X-ray emission is much more concentrated towards the galactic center than the 1.8 MeV emission, leading to huge residual wings in the data. Also the H I map which, in contrast to HEAO-1, obeys a flat longitude profile, is rejected by the data. Considerably better, although not satisfactory fits are obtained for the synchrotron models (408 MHz and DMR synch.) and the DIRBE near-infrared maps. While synchrotron emission provides a reasonable fit for the central galactic radian, it fails to explain the 1.8 MeV emission in the southern hemisphere ($l < 0°$) and is also unable to account for the emission in Cygnus ($l \sim 80°$) and in the Auriga-Camelopardalis-Perseus (ACP) region ($l \sim 130° - 180°$). The near-infrared models, which are well described by a small scale length exponential disk plus a stellar bulge component, obviously provide a profile that is sharper than the 1.8 MeV

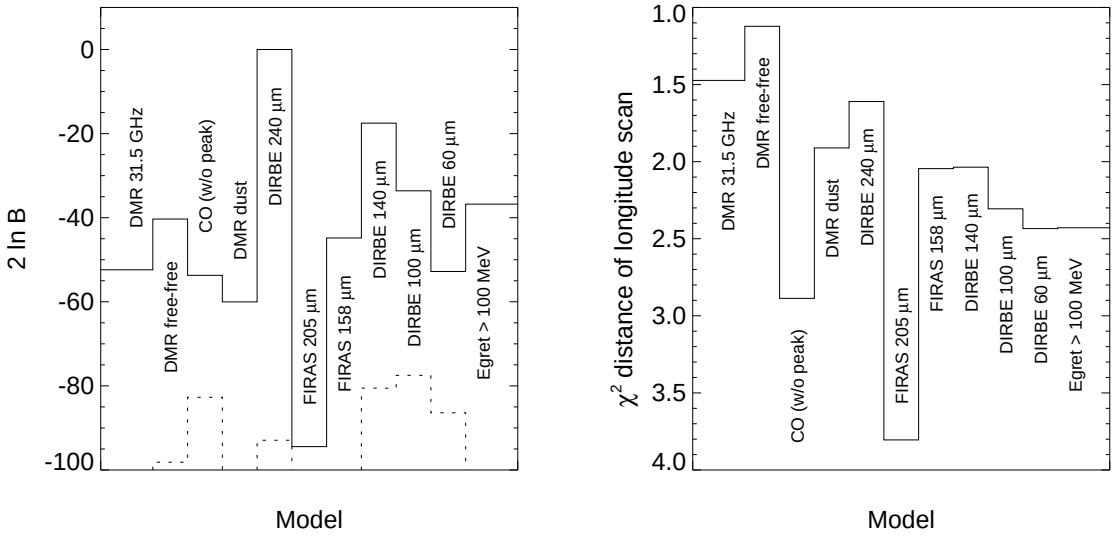

Fig. 4.15: Left: *Log-Bayes factor for models within the Ockham window* $2\ln B \geq -100$. Right: *Corresponding* χ^2 *distance of the longitude profiles determined using the software collimation technique.*

distribution with a pronounced peak towards the galactic center. Consequently, the 1.8 MeV emission in the center is overestimated by the models, while significant residuals remain far off the galactic center.

The log-Bayes factors for the best models are magnified in the left panel of Fig. 4.15, which shows all models that fall in the Ockham window $2\ln B \geq -100$. The software collimation longitude scans for these models are shown together with maximum likelihood ratio residual maps in Appendix D. Apparently, although the DIRBE 240 μm map obeys the highest Bayes factor, it shows larger residuals than the DMR free-free emission and DMR 31.5 GHz maps. While the residual maps for the latter two models show no significant emission in the galactic plane, there are significant (4σ) features in the DIRBE 240 μm residual map situated in Carina and Vela. The longitude scan for the DIRBE 240 μm model (Fig. 4.16) shows that it systematically underestimates the 1.8 MeV intensity in the southern hemisphere ($l < 0°$). The steep slope of the 1.8 MeV intensity decrease between $l = 30° - 60°$ is also not exactly followed by the DIRBE 240 μm profile, leading to negative residuals in this area. In contrast, the DMR free-free map provides a satisfactory explanation of the 1.8 MeV data. It not only reproduces the overall 1.8 MeV longitude profile rather closely, but also accounts quite well for the particularly pronounced emission in Cygnus. Only marginal residuals are seen in the galactic center, Carina, Vela, and the ACP region, but comparison with the residuals of the 'ideal case' (cf. Fig. 4.9) shows that such residuals are expected from statistical fluctuations.

So since the DMR free-free model provides a better explanation of the observed

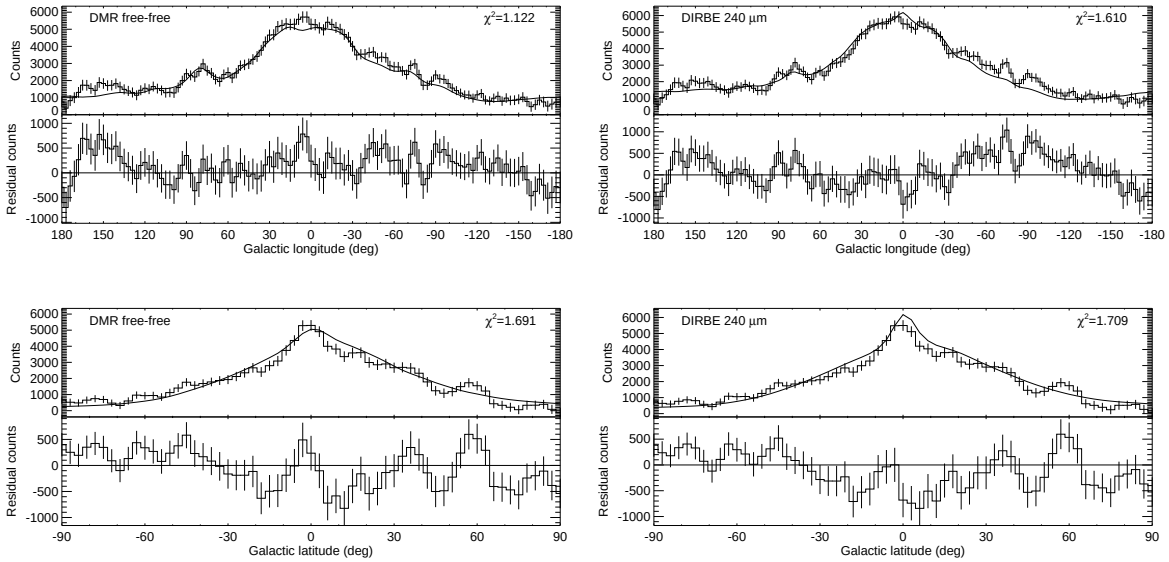

Fig. 4.16: *Software collimator scans for the $\bar{\varphi}$ fit of the DMR free-free emission (left) and the DIRBE 240 μm (right) tracer maps. The upper panels show a longitude scan for $b = 0°$, lower panel show a latitude scan for $l = 0°$. The χ^2 statistics for the scans is given in the upper right corner of the plots.*

1.8 MeV longitude profile, why does the Bayes factor analysis prefer the DIRBE 240 μm map? Obviously, the preference is also not due to a better description of the latitude profile of the data. The lower panels of Fig. 4.16 show software collimator scans made in galactic latitude for $l = 0°$. Both models provide similar, although not perfect descriptions of the 1.8 MeV data. The number of excess counts at positive compared to negative galactic latitudes is due to lack of exposure in the southern galactic hemisphere. Nevertheless there are some excess residual counts at negative galactic latitudes which are systematically found for all tracer maps.

Since neither the longitude nor the latitude distribution of the DIRBE 240 μm map seems to be the origin of the high Bayes factor, the only remaining explanation of the favour is the $\bar{\varphi}$ distribution. Indeed, Fig. 4.17 shows that for the DIRBE 240 μm map the $\bar{\varphi}$ distribution residuals are smaller than for the DMR free-free emission map. Although the differences, indicated by the χ^2 distance, are only small, their impact on the analysis is important: A change in the global $\bar{\varphi}$ distribution affects so many pixels at a time that any inaccuracy in the $\bar{\varphi}$ distribution of the background or source model will disturb the fit. From the discussion in Section 4.2.2 it is known that the $\bar{\varphi}$ distribution of the background model is only an approximation. Although $\bar{\varphi}$ fitting can alleviate this uncertainty some residual bias remains which gives an unjustified strong favour to one of the best models.

The bias of the analysis by the uncertainty in the $\bar{\varphi}$ distribution of the background model can be estimated from Fig. 4.18 where the χ^2 distance of the longitude scan

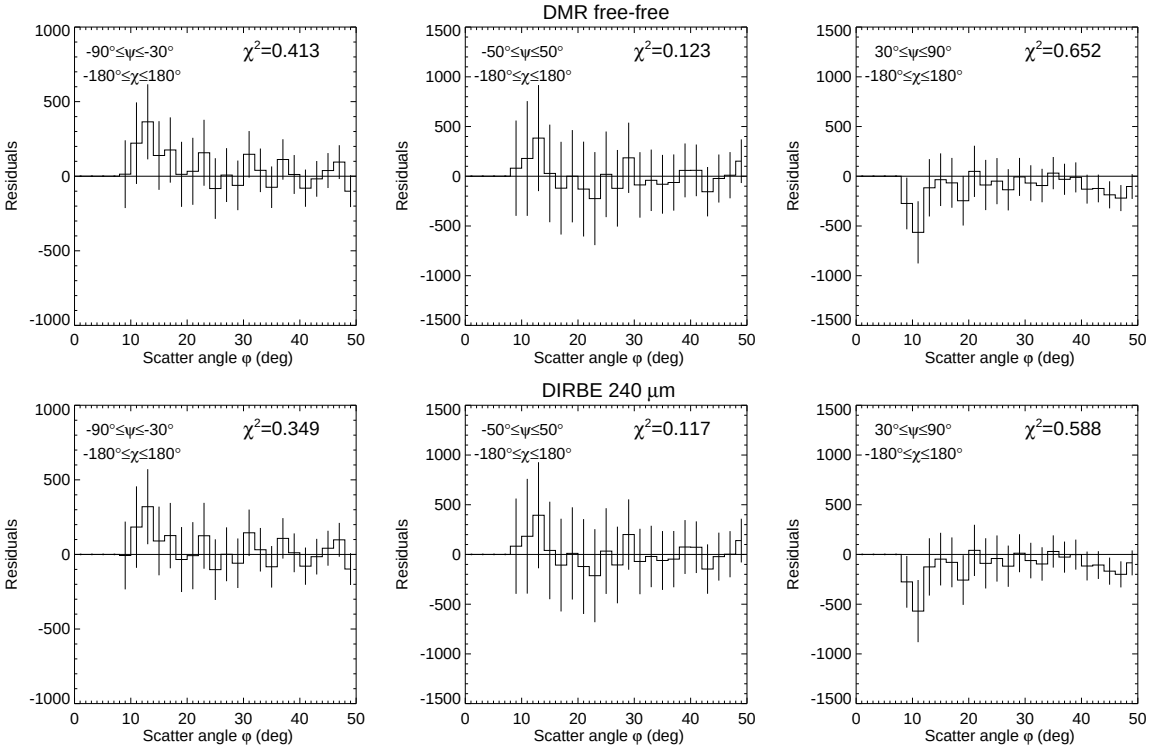

Fig. 4.17: Top: $\bar{\varphi}$ *distribution residuals for the DMR free-free emission model for three data space regions. The distance between the data and the model is again measured by the* χ^2 *statistics.* Bottom: *Same for DIRBE 240* μ*m map.*

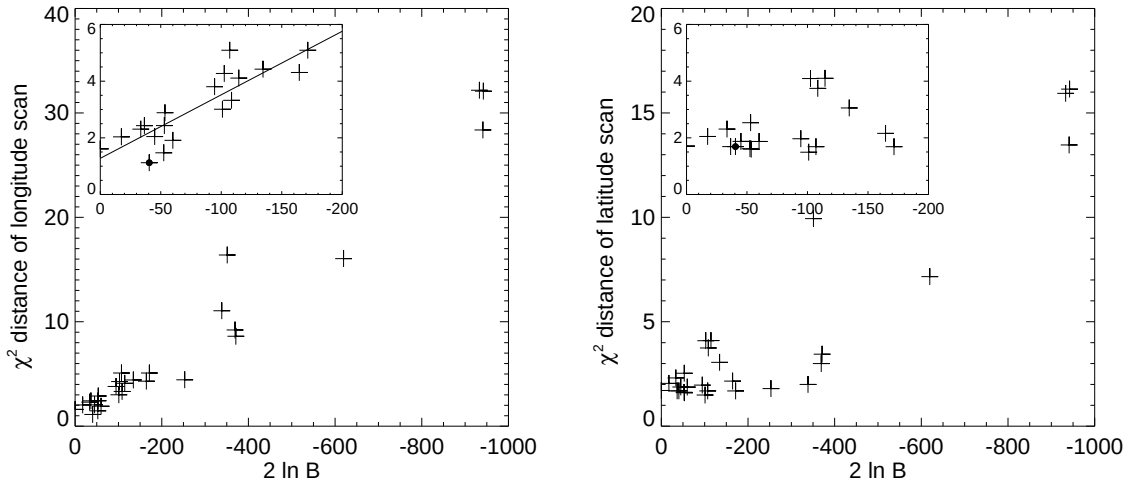

Fig. 4.18: Left: χ^2 *distance of the software collimator longitude scan along* $b = 0°$ *versus log-Bayes factor. The inset represents a zoom on the smallest* χ^2 *distances and log-Bayes factors. The solid line is a linear regression fit to the data points above* $2 \ln B > -200$*. The dot indicates the DMR free-free model.* Right: *Same for a software collimator latitude scan along* $l = 0°$*.*

is plotted for all models versus the log-Bayes factor. Although a clear correlation of increasing χ^2 distance with decreasing Bayes factor exists, it is not strictly monotonic. The DMR free-free model, for example, which provides the best description of the longitude profile, has a log-Bayes factor of $2\ln B \approx -40$ with respect to the DIRBE 240 μm model. From a linear regression to the data points above $2\ln B > -200$, a standard deviation of $\Delta 2\ln B = 28$ was found, which may be taken as systematic uncertainty in the log-Bayes factor. This finding justifies a posteriori the choice of an Ockham window of $2\ln B \geq -60$ for the Ockham window average (cf. Section 4.3.2.2).

Taken together, the uncertainty in the Bayes factor and the results from the residual study, the DMR free-free emission map can be considered as the best model. To illustrate the similarity between this tracer map and the 1.8 MeV data, both the COMPTEL all-sky map of 1.8 MeV emission and the DMR free-free emission map are shown in Fig. 4.19. Obviously, essentially all significant structures in the 1.8 MeV map find their counterpart in the DMR map: the prominent emission in the Cygnus region ($l \approx 80°$), the sharp intensity edge at $l \approx 30°$, and the galactic ridge-emission peaks at $l \approx 265°$ (Vela), $l \approx 287°$ (Carina), $l \approx 310° - 320°$ (Arm 2 south), and $l \approx 345°$ (Scorpius). The only apparent discrepancies are the ACP feature in the 1.8 MeV map around $l \approx 150° - 170°$ which has no obvious counterpart in the DMR map, and possibly the galactic center spot around $l \approx 5°$ which coincides with a gap in the free-free emission map. Notice, however, that the DMR free-free map is limited by statistical noise, hence the diffuse low intensity feature in the ACP region might be missed. On the other hand, the ACP feature and the galactic center spot in the 1.8 MeV map may be influenced by the background model uncertainties, which leads to an 'attraction' of 1.8 MeV emission towards these directions (cf. Section 4.2.2).

Using the χ^2 distance of the longitude profile as a measure of the fit quality, the right panel of Fig. 4.15 illustrates the superiority of the DMR free-free emission model. After the DMR free-free map, good models are the DMR 31.5 GHz map and the far-infrared dust tracers. Both, Bayesian Model Averaging (BMA) and Ockham Window Averaging (OWA), result in a total flux of $(3.1 \pm 0.1)\,10^{-4}$ ph cm^{-2}s^{-1} from the central steradian, which is slightly higher than values obtained from imaging analysis (2.8 ± 0.3 ph cm^{-2}s^{-1} for both R-L and ME), but lower than the prior estimate of $(4.5 \pm 2.0)\,10^{-4}$ ph cm^{-2}s^{-1}.

Similar to Tabs. 3.1 and 3.2 presented in Chapter 3 for the reconstructed intensity maps, fluxes for prominent 1.8 MeV emission features were derived from the tracer maps using BMA, OWA, and the DMR free-free emission map (cf. Tab. 4.3). For comparison, fluxes derived from reconstructed sky intensity distributions are given in columns R-L and MEM. R-L fluxes are from the Richardson-Lucy reconstruction of the all-sky dataset from observation periods 0.1 - 522.5 (cf. Tab. 3.1). MEM fluxes are from galactic plane data of mission phases 1 & 2 as quoted by Knödlseder (1994). Notice that the BMA fluxes are dominated by the DIRBE 240 μm map, hence the quoted values correspond to fluxes in that map multiplied by the scaling factor as derived

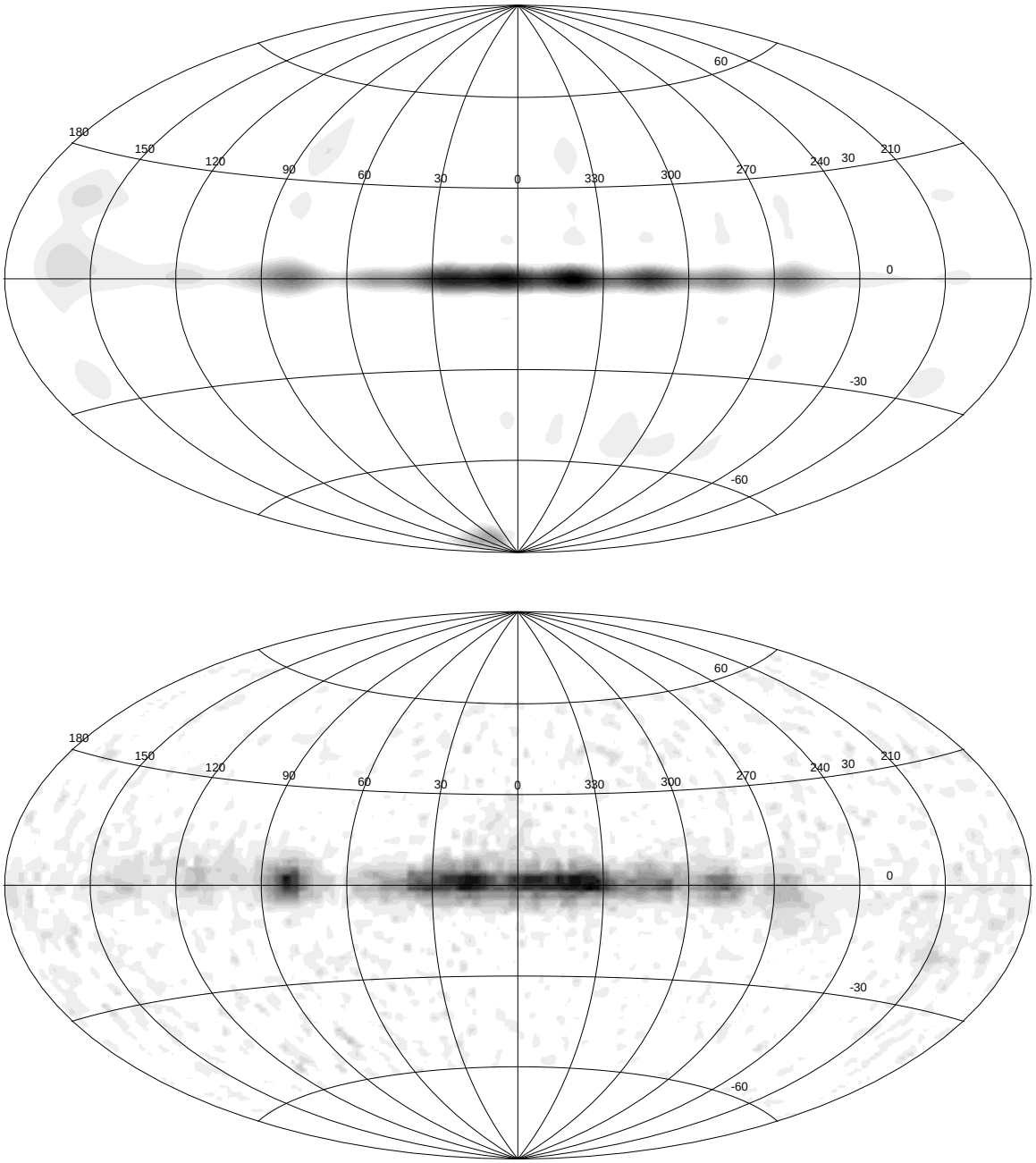

Fig. 4.19: *Comparison of the COMPTEL 1.8 MeV all-sky map as derived by the modified Richardson-Lucy algorithm for resolution levels* $(5, 5)$ *(top) with the DMR free-free emission map of Bennett* et al. *(1992) (bottom). Notice that the DMR map has only an angular resolution of* $7°$ *(FWHM).*

Position		Radius	BMA	OWA	DMR	R-L	MEM
l	b						P1+2
2°	-1.5°	5°	5.3 ± 0.2	4.7 ± 1.5	3.0 ± 0.1	4.2 ± 0.6	5.5 ± 1.1
32°	1°	6°	5.0 ± 0.2	4.5 ± 0.6	4.1 ± 0.2	4.0 ± 0.7	5.1 ± 0.8
80.5°	-3.5°	4°	0.8 ± 0.1	0.9 ± 0.4	1.6 ± 0.1	1.3 ± 0.5	1.8 ± 0.8
168.5°	-6.5°	6°	0.5 ± 0.1	0.5 ± 0.2	0.5 ± 0.1	0.8 ± 0.7	3.5 ± 1.1
238°	5°	5°	0.2 ± 0.1	0.2 ± 0.1	0.2 ± 0.1	2.0 ± 0.6	4.1 ± 1.2
265.5°	-1.5°	6°	1.6 ± 0.1	1.5 ± 0.3	1.9 ± 0.1	2.8 ± 0.7	4.6 ± 1.2
286.5°	0.5°	3°	0.9 ± 0.1	1.0 ± 0.4	1.0 ± 0.1	2.9 ± 0.5	3.3 ± 0.8
309.5°	-1.5°	4°	1.9 ± 0.1	1.6 ± 0.2	1.5 ± 0.1	1.8 ± 0.5	3.0 ± 1.0
345.5°	-0.5°	6°	6.3 ± 0.3	6.2 ± 0.6	6.2 ± 0.3	5.4 ± 0.7	6.3 ± 1.3

Region				BMA	OWA	DMR	R-L	MEM
l_{min}	l_{max}	b_{min}	b_{max}					P1+2
-180°	180°	-10°	10°	84.4 ± 3.8	80.5 ± 8.4	89.8 ± 4.1	59.5 ± 4.9	96.5 ± 4.1
0°	180°	-10°	10°	44.4 ± 2.0	41.7 ± 5.2	45.2 ± 2.0	31.2 ± 3.5	44.6 ± 3.7
-180°	0°	-10°	10°	40.0 ± 1.8	38.9 ± 4.0	44.6 ± 2.0	28.3 ± 3.5	51.9 ± 3.6
-35°	38°	-10°	10°	38.6 ± 1.7	38.6 ± 2.4	40.0 ± 1.8	27.4 ± 2.3	35.4 ± 2.6
302°	325°	-10°	10°	7.6 ± 0.3	7.0 ± 0.7	7.8 ± 0.4	6.2 ± 1.3	8.0 ± 1.4
-30°	30°	-10°	10°	32.5 ± 1.5	33.0 ± 2.1	34.3 ± 1.6	24.7 ± 2.1	31.0 ± 2.6
-10°	15°	-10°	10°	14.1 ± 0.6	14.0 ± 1.2	13.7 ± 0.6	10.1 ± 1.4	14.4 ± 1.7
72°	100°	-10°	8°	6.3 ± 0.3	6.3 ± 1.6	8.6 ± 0.4	5.6 ± 1.4	7.4 ± 1.2

Table 4.3: *1.8 MeV fluxes in units of* 10^{-5} *ph* $cm^{-2}s^{-1}$ *as determined from the tracer map comparison, using Bayesian Model Averaging (BMA), Ockham Window Averaging (OWA), and the DMR free-free emission model (DMR). For comparison, fluxes derived from reconstructed sky intensity distributions are shown in columns R-L and MEM. R-L fluxes are from the Richardson-Lucy reconstruction of the all-sky dataset from observation periods 0.1 - 522.5 (cf. Tab. 3.1). MEM fluxes are from galactic plane data of mission phases 1 & 2 as quoted by Knödlseder (1994).*

from the fitting procedure. A more reliable estimate of the uncertainty introduced by the choice of the model intensity distributions is given by the OWA fluxes and the corresponding flux errors.

It is remarkable that the OWA point source fluxes (upper table) are fairly consistent with flux estimates from the R-L skymap. This demonstrates that the best tracers maps provide a quite good and detailed description of the COMPTEL 1.8 MeV data. The only exceptions to this good agreement are the Puppis feature at $(l, b) = (238°, 5°)$, which obviously has no significant counterpart in the tracer maps, and the Carina feature at $(l, b) = (286.5°, 0.5°)$. For the Carina feature, however, a small integration radius of only 3° was chosen since it appears as a very concentrated emission feature in the 1.8 MeV all-sky maps. Raising this radius to 6° results in tracer-map fluxes which are consistent with the corresponding R-L flux of $(2.8 \pm 0.7)\,10^{-5}$ ph $cm^{-2}s^{-1}$.

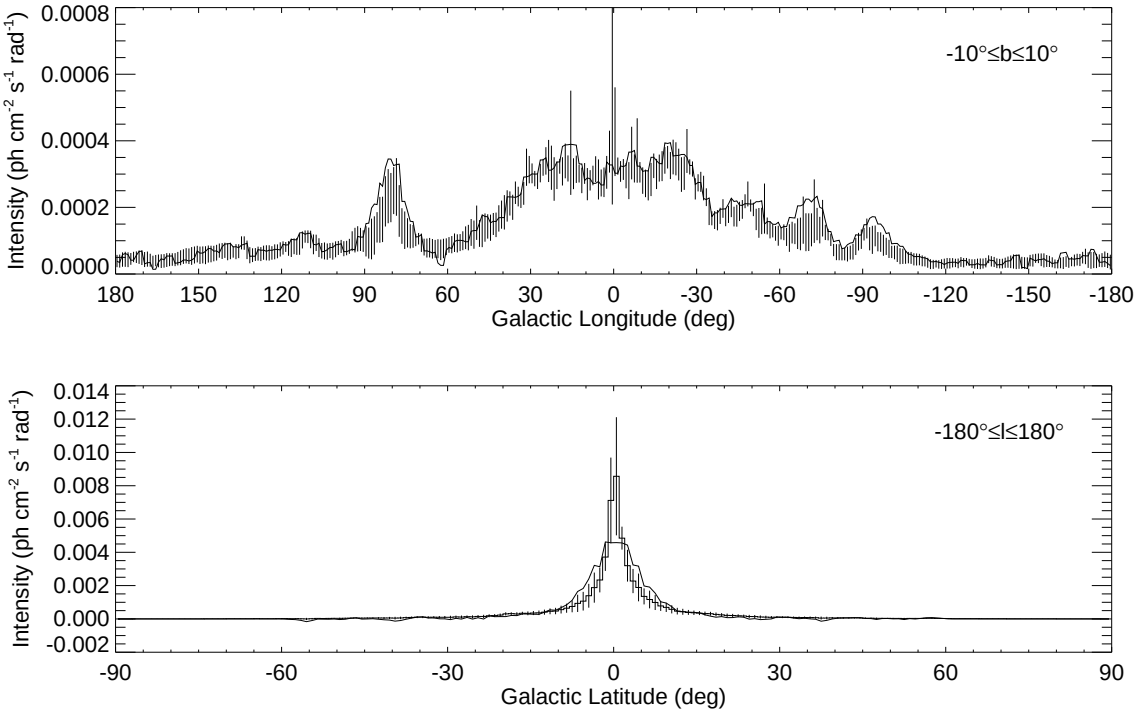

Fig. 4.20: *Ockham Window Average longitude (top) and latitude (bottom) profiles. The error bars indicate the OWA intensities and errors as determined from the* $2\ln B \geq -60$ *Ockham window. The solid line shows the corresponding profile of the DMR free-free emission map.*

The tracer map fluxes for diffuse emission (lower table), however, are considerably higher than corresponding R-L fluxes. On the other hand they are rather close to flux values derived from 1.8 MeV galactic plane data from mission phases 1 & 2 (MEM, P1+2), which points to a systematic flux underestimation in the reconstructed all-sky maps. Obviously, the underestimation can not be due the reconstruction algorithms only, since the careful flux calibrations of Section 3.2.2.3 and Section 3.3.2 showed no hint for flux underestimation for extended emission features. Additionally, it can not be due to the uncertainties in the background model only, since for both image reconstruction and model fitting the same background model was used. Consequently, it must be due to the combination of an uncertain background model with an all-sky image reconstruction, but the reason for this discrepancy remains unclear.

Finally, longitude and latitude profiles of the 1.8 MeV emission are shown in Fig. 4.20 which were derived from the tracer maps using Ockham Window Averaging.

4.3.3 Interpretation of the Results

4.3.3.1 The Ionized Interstellar Medium

The finding that only the DMR free-free emission map provides a satisfactory fit of COMPTEL 1.8 MeV data has important implications on the origin of galactic ^{26}Al. Free-free radiation, or thermal bremsstrahlung, arises from the acceleration of free electrons in the electrostatic field of ionized atoms (mainly hydrogen), hence it traces the distribution of ionized gas in the interstellar medium (ISM). Ionized gas in the Galaxy is composed of two components: On the one hand there are 'normal' H II regions with electron densities n_e of at least 10 cm^{-3}, and which appear as Strömgren spheres that are photoionized by the Lyman continuum (Lyc) photons ($\lambda < 912$ Å) of the embedded OB star(s). One the other hand there is widespread diffuse ionized gas (DIG) lying outside the boundaries of traditional H II regions, occupying more than 20% of the interstellar volume and constituting $\sim 90\%$ of the ionized hydrogen mass in the Galaxy (see review by Reynolds 1991a, and references therein). Pulsar dispersion measures and interstellar emission lines indicate that the DIG (also referred to as warm ionized medium, WIM) is clumped into regions which have an average electron density $n_e \sim 0.1$ cm^{-3}, an electron temperature around $T_e \sim 8000$ K, and which form a thick H II layer (referred to as 'Reynolds Layer') with a scale height of ~ 1 kpc (Reynolds 1991b). Although the source of ionization of the DIG is still under debate, most authors agree that only Lyc photons from OB stars can easily provide the tremendous power requirement of 10^{-4} erg s^{-1} cm^{-2} to keep the DIG ionized (e.g., Ferguson *et al.* 1996a, Dove and Shull 1994, Miller and Cox 1993, Mathis 1986).

Observations of the DIG in other spiral galaxies provide crucial information that is difficult to obtain for the galactic DIG, in particular the overall spatial distribution across and perpendicular to galactic disks. From an analysis of deep Hα images obtained for two nearby face-on spirals, Ferguson *et al.* (1996a) find that the DIG is highly correlated with bright H II regions over both small and large scales, supporting the hypothesis that the DIG is photoionized by Lyc photons which leak out of traditional H II regions. Similar results were obtained by Ferguson *et al.* (1996b), Hoopes *et al.* (1996), and Walterbos and Braun (1994) for inclined spirals which all find that DIG contributes 30% to 50% to the total Hα luminosity in these galaxies. It appears that the contribution of DIG to the total Hα luminosity (the diffuse fraction) seems to be a constant fraction of the H II region luminosity, both within and among different galaxies (Hoopes *et al.* 1996, and references therein). This and the noted correlation with star-forming regions implies a close correlation of massive stars with the DIG.

4.3.3.2 Free-Free Emission

In our Galaxy, ionized gas can be observed by optical line emission arising from the recombination of ionized species, pulsar dispersion measures, far-infrared cooling lines,

or microwave continuum free-free emission. Most of these tracers, however, suffer either from low signal-to-noise ratio (e.g., N II emission; Bennett *et al.* 1994b) or from undersampling and selection biases (e.g., Hα emission and pulsar dispersion measure; Reynolds 1992, 1984). Only the microwave emission presents an unbiased all-sky tracer of the ionized gas phase of the ISM (Kogut *et al.* 1996b).

The volume emissivity (in erg cm^{-3} s^{-1} sr^{-1} Hz^{-1}) for free-free emission is given by

$$\varepsilon_\nu = 6.507\,10^{-38}\,Z^2 n_e n_i T_e^{-0.35} \nu^{-0.1} \exp\left(-\frac{h\nu}{kT_e}\right) \tag{4.34}$$

where Z is the charge number of the ions, n_e and n_i are the electron and ion volume densities (in cm^{-3}), T_e is the electron temperature (in K), and ν is the observation frequency (in Hz) (e.g., Brown 1987). For microwaves ($\nu \sim 10^{11}$ Hz) and typical electron temperatures of $T_e \sim 10^4$ K, $h\nu \ll kT_e$, hence the exponential factor can be omitted. Assuming that the abundance of He^{++} is much less than the abundance of H^+, one can take $Z = 1$ and $n_e = n_i$ (e.g., Bennett *et al.* 1992). Assuming further that the electron temperature is constant along the line of sight, the microwave free-free intensity $I_{\rm ff}(l,b)$ (in erg cm^{-2} s^{-1} sr^{-1} Hz^{-1}) towards a direction (l,b) on the sky is given by[14]

$$I_{\rm ff}(l,b) = 2.00\,10^{-19}\,T_e^{-0.35}\nu^{-0.1}{\rm EM}. \tag{4.36}$$

EM $= \int n_e^2 ds$ (in cm^{-6} pc) is the emission measure resulting from the integration of n_e^2 along the line of sight. In ionization equilibrium, the radiative recombination rate equals the Lyc luminosity, hence

$$\rho_{Lyc} = n_e^2 \alpha_B(T) \tag{4.37}$$

where ρ_{Lyc} is the Lyc emissivity (in cm^{-3} s^{-1}), and $\alpha_B(T)$ is the Case B recombination coefficient[15] for hydrogen (in cm^3 s^{-1}). For electron temperatures characteristic of ionized regions ($T_e \sim 10^4$ K), the recombination coefficient is well approximated by

$$\alpha_B(T) \approx 4.10\,10^{-10}\,T^{-0.8} \tag{4.38}$$

(Brown 1987) which leads to the relation

$$I_{\rm ff}(l,b) = 1.59\,10^{-28}\,T_e^{0.45}\nu^{-0.1}\int \rho_{Lyc} ds. \tag{4.39}$$

---

[14]The DMR free-free emission map is given in units of antenna temperature which is related to $I_{\rm ff}$ by

$$T_A = \frac{I_{\rm ff} c^2}{2\nu^2 k} = 8.442\,10^{-7}\left(\frac{\nu}{53\,{\rm GHz}}\right)^{-2.1}\left(\frac{T_e}{8000\,{\rm K}}\right)^{-0.35}{\rm EM} \tag{4.35}$$

where EM is in units of cm^{-6} pc.

[15]Case B recombination assumes that the ionized region is thick to Lyman lines (e.g., Péquignot *et al.* 1991).

Equation (4.39) demonstrates that the microwave free-free intensity is directly proportional to the Lyc emissivity integrated along the line of sight (the Lyc column density). The major source of Lyc photons in the Galaxy are massive stars (OB and Wolf-Rayet stars), minor ($\lesssim 1\%$) contributions may come from planetary nuclei, white dwarfs, and supernovae (Reynolds 1984, Abbott 1982). Therefore, the galactic Lyc emissivity may be written as the product of the present-day mass function (PDMF), which characterizes the density of stars in a region as a function of stellar mass M, and the stellar Lyc luminosity function $Q_0(M)$, which is the number of H-ionizing photons produced per second by an individual star of mass M. In principle, $Q_0(M)$ depends not only on the stellar mass of a star but also on its evolutionary age and metallicity. For simplicity one generally assumes that all stars are Zero-Age Main Sequence (ZAMS) stars and consequently the PDMF is replaced by the Initial Mass Function (IMF), $\xi(M_i)$, which characterizes the density of stars in a region as a function of initial stellar mass M_i (Vacca 1994). The Lyc emissivity is then given by

$$\rho_{Lyc} = \int_{M_{low}}^{M_{upp}} \xi(M_i) Q_0(M_i) dM_i, \tag{4.40}$$

where M_{low} and M_{upp} specify the lower and upper mass limits, respectively, of stars contributing to the ionizing flux. The star density integrated over this mass range is given by

$$n_* = \int_{M_{low}}^{M_{upp}} \xi(M_i) dM_i. \tag{4.41}$$

The dependence of the Lyc emissivity on stellar mass is illustrated in Fig. 4.21, which is based on the $Q_0(M_i)$ calibration of Vacca *et al.* (1996) for solar metallicities. For the initial mass function, a power-law IMF

$$\xi(M_i) = \frac{dn_*}{dM_i} = n_0 M_i^{\Gamma-1} \tag{4.42}$$

was chosen. The canonical Salpeter mass function has a slope of $\Gamma = -1.35$, but also the extreme values $\Gamma = -0.5$ and -2.5 were considered which span the observational and theoretical range found for massive stars in various galaxies (Vacca 1994, and references therein). Obviously, only the most massive stars provide a significant amount of Lyc photons. Regardless of the IMF slope, O stars ($M \gtrsim 20$ M$_\odot$) contribute more than 95% of the galactic ionizing photons, while the remaining 5% is provided by stars above $M \gtrsim 10$ M$_\odot$. For convenience, the Lyc emissivity may be expressed as the density of 'equivalent O7 V' stars using

$$\rho_{Lyc} = n_*^{\rm O7V} Q_0^{\rm O7V}, \tag{4.43}$$

where $n_*^{\rm O7V}$ is the 'equivalent O7 V' star density, assuming that all ionizing stars are O7 V stars, and $Q_0^{\rm O7V}$ is the ionizing luminosity of an individual O7 V star of solar metallicity. The total density of O V stars, $n_*^{\rm OV}$, is then determined by

$$n_*^{\rm OV} = n_*^{\rm O7V}/\eta_0, \tag{4.44}$$

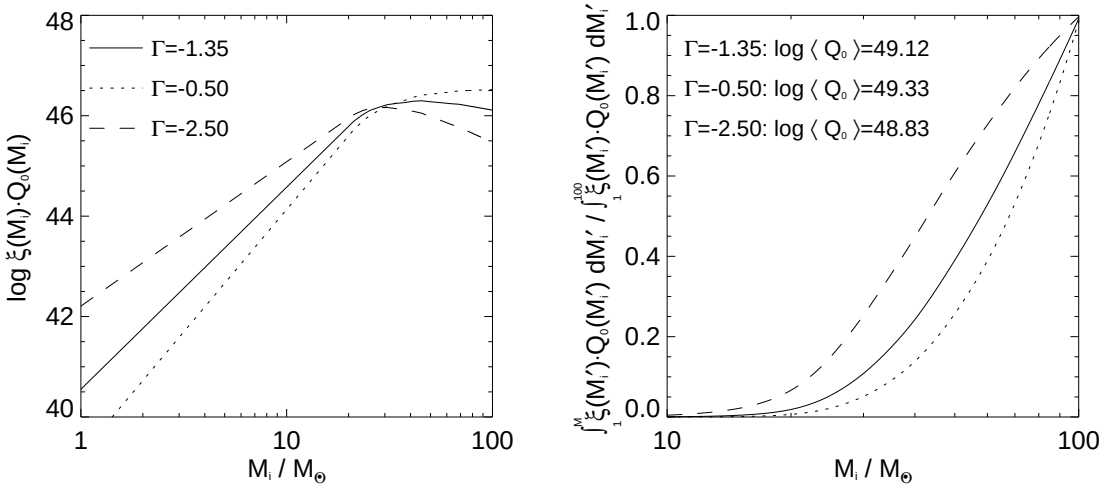

Fig. 4.21: Left: *Lyman continuum luminosities weighted with power-law IMFs of slopes* $\Gamma = -1.35$ *(solid),* -0.5 *(dotted), and* -2.5 *(dashed) and normalized to one star within the mass interval* $17.3 - 100$ $M_\odot$. Right: *Fractional Lyc contribution of stars within the mass interval* $1 - M$ $M_\odot$ *for the same IMF slopes. The quoted values are the average Lyc luminosities over the mass interval* $17.3 - 100$ $M_\odot$ *for* $\Gamma = -1.35, -0.5$, *and* -2.5, *corresponds to mean stellar masses of 38, 45, and 31* $M_\odot$, *respectively.*

where

$$\eta_0 = \frac{\int_{M_{low}}^{M_{upp}} \xi(M_i) Q_0(M_i) dM_i}{Q_0^{\rm O7V} \int_{M_{\rm OV}}^{M_{upp}} \xi(M_i) dM_i} \tag{4.45}$$

is tabulated by Vacca (1994) for various mass ranges and stellar metallicities. Using $\log Q_0^{\rm O7V} = 49.12$, as suggested by the most recent $Q_0(M_i)$ calibration of Vacca *et al.* (1996), the microwave free-free intensity (in erg cm^{-2} s^{-1} sr^{-1} Hz^{-1}) may be rewritten as

$$I_{\rm ff}(l,b) = 1.06\,10^{-15} \left(\frac{T_e}{8000\,{\rm K}}\right)^{0.45} \left(\frac{\nu}{53\,{\rm GHz}}\right)^{-0.1} \int n_*^{\rm O7V} ds, \tag{4.46}$$

with $n_*^{\rm O7V}$ given in units of 'equivalent O7 V' stars pc^{-3}. This corresponds to an antenna temperature T_A (in K) of

$$T_A(l,b) = 1.23 \left(\frac{T_e}{8000\,{\rm K}}\right)^{0.45} \left(\frac{\nu}{53\,{\rm GHz}}\right)^{-2.1} \int n_*^{\rm O7V} ds. \tag{4.47}$$

Equations (4.46) and (4.47) demonstrate that the observed microwave free-free intensity is directly proportional to the ionizing star density integrated along the line of sight. Since the Galaxy is transparent at these frequencies, the line integral has to be taken over the entire galactic disk.

4.3.3.3 Equivalent O7 V star ^{26}Al yield

Similar relations hold for the 1.8 MeV intensity: since the Galaxy is also transparent to MeV γ-rays, the observed 1.8 MeV intensity (in ph cm^{-2} s^{-1} sr^{-1}) is proportional to the 1.8 MeV emissivity $\varepsilon_{1.8}$ (in ph cm^{-3} s^{-1}) integrated along the line of sight over the entire galactic disk:

$$I_{1.8}(l,b) = \frac{1}{4\pi} \int \varepsilon_{1.8} ds. \tag{4.48}$$

1.8 MeV photons arise from the radioactive decay of ^{26}Al with a production rate of $Q = 1.404\,10^{42} (M/\mathrm{M}_\odot)$ ph s^{-1}. Hence the 1.8 MeV emissivity can also be written as ^{26}Al mass density ρ_{26} (in $\mathrm{M}_\odot$ pc^{-3}) leading to

$$I_{1.8}(l,b) = 1.178\,10^{4} \int \rho_{26} ds. \tag{4.49}$$

Since the DMR free-free emission map ($\nu = 53$ GHz) provides a satisfactory fit to COMPTEL 1.8 MeV data, one can equate $I_{1.8}(l,b) = X_{1.8} T_A(l,b)$, where $X_{1.8}$ is the scaling factor obtained by the fit. Using Eqs. (4.47) and (4.49), the ^{26}Al column density then is rewritten as

$$\int \rho_{26} ds = Y_{26}^{\mathrm{O7V}} \int n_*^{\mathrm{O7V}} ds, \tag{4.50}$$

where

$$\boxed{Y_{26}^{\mathrm{O7V}} = 1.046\,10^{-4} X_{1.8} \left(\frac{T_e}{8000\mathrm{K}} \right)^{0.45}} \tag{4.51}$$

is the **'equivalent O7 V star ^{26}Al yield'**. Assuming a typical electron temperature of $T_e = 8000$ K, and using the MAP scaling factor of $X_{1.8} = 0.86 \pm 0.03$ ph cm^{-2} s^{-1} sr^{-1} K^{-1} as assigned by the fit, the equivalent O7 V star ^{26}Al yield as derived by COMPTEL 1.8 MeV all-sky data from observation periods 0.1 - 522.5 is given by

$$\boxed{Y_{26}^{\mathrm{O7V}} = (9.0 \pm 0.3)\,10^{-5} \mathrm{M}_\odot \text{ per 'equivalent O7 V' star}} \tag{4.52}$$

4.3.3.4 Justification of the Assumptions

The above discussion demonstrated that the satisfactory fit of the DMR free-free emission map to COMPTEL 1.8 MeV data implies a direct proportionality between the ^{26}Al and the massive star column densities. To derive this proportionality, however, it was assumed that, firstly, the galactic H II regions and the DIG are in ionization equilibrium, and, secondly, that the spatial Lyc distribution is identical to the spatial distribution of massive stars. In the following both assumptions will be discussed.

The timescale for initially neutral gas to reach ionization equilibrium once the gas is exposed to an ionizing radiation is $\tau_{\mathrm{eq}} \approx 10$ yr (Dove and Shull 1994), much shorter

than the lifetimes of massive stars. When the ionizing flux stops, because the massive stars vanished in supernova explosions, the ionized gas recombines with a timescale of

$$\tau_{\rm r} = (n_e \alpha_B(T))^{-1} = 1.03\,10^6 \left(\frac{T_e}{8000{\rm K}}\right)^{0.8} \left(\frac{0.1{\rm cm}^{-3}}{n_e}\right) {\rm yr}. \qquad (4.53)$$

Hence, H II regions with electron densities above 10 cm^{-3} should disappear after less than 10 000 yr, while DIG remains ionized for about 1 Myr. Nevertheless, $\tau_{\rm r}$ is still shorter than the evolutionary time scales of massive stars, hence it is valid to assume that the ionized gas is in ionization equilibrium.

Lyc photons are heavily absorbed by the interstellar gas through photoelectric absorption, leading to mean free path lengths of some 100 pc and consequently to a concentration of the ionized gas around the O stars (e.g., Miller and Cox 1993). About half of the free-free luminosity arises from traditional H II regions, which form compact Strömgren spheres of radii

$$R_S = 3.16 \left(\frac{Q_0}{10^{49}\ {\rm s}^{-1}}\right)^{1/3} \left(\frac{n_e}{100\ {\rm cm}^{-3}}\right)^{-2/3} {\rm pc} \qquad (4.54)$$

around the ionizing stars. The other half originates from DIG, which is probably ionized by Lyc photons leaking out of traditional H II regions. Consequently, most of the DIG should be found in halos around H II regions, which is indeed confirmed by observations of nearby spiral galaxies (e.g., Ferguson *et al.* 1996a, Hoopes *et al.* 1996). Nevertheless, a fraction of Lyc photons may escape the galactic disk since their mean free path length is of the order of the vertical scale height of the gas layer. In particular the so called 'chimneys', which could be the conduits into the halo for hot gas around disk star-forming regions, could also provide low-density paths for the passage of Lyc photons from the disk into the halo. Thus, while within the galactic plane the spatial correlation of ionized gas with massive stars should be good to within some 100 pc, the vertical extent of the DIG is much higher ($z_0 \sim 1000$ pc) than that of massive stars. Notice, however, that the diffusion of Lyc photons should be no serious problem for the current COMPTEL 1.8 MeV – DMR free-free emission map comparison, since (1) the angular resolution of the DMR (7° FWHM) and COMPTEL (4° FWHM) data can provide no information about small scale correlations anyway, (2) only a fraction of the Lyc photons diffuse out of the plane and form the extended 'Reynolds layer', and (3) ^{26}Al may also travel substantial distances within its lifetime of 1 Myr. In this context it is interesting to recall that recent 1.8 MeV γ-ray line observations using the *GRIS* balloon experiment indicate a line broadening corresponding to ^{26}Al Doppler velocities of $\sim$ 500 km s^{-1}, which would allow ^{26}Al to travel about 500 pc during its lifetime (Naya *et al.* 1996). Consequently, not only the Lyc photons, but also ^{26}Al could form halos around massive star associations.

4.3.3.5 The Origin of Galactic ^{26}Al

The finding that only the DMR free-free emission map provides a satisfactory fit of COMPTEL 1.8 MeV data together with the direct proportionality of free-free emissivity and O star density leads unavoidably to the conclusion that ^{26}Al is mainly produced by O stars ($M \gtrsim 20$ M$_\odot$) in the Galaxy. Although less massive stars may occur in association with ionized gas (e.g., because they were formed together with the ionizing stars in the same associations) they are likewise expected in non-ionized regions. Thus if ^{26}Al would be produced by e.g. ONeMg novae or AGB stars, significant violations of the correlation between 1.8 MeV and free-free emission would be expected. I want to stress that even massive AGB stars can be excluded as prolific ^{26}Al sources. While the lifetime of ionizing stars ($\gtrsim 20$ M$_\odot$) is 9 Myr at maximum (Schaller *et al.* 1992), the typical lifetime for massive AGB progenitors (initial mass $M \approx 8$ M$_\odot$) is ~ 40 Myr, hence they outlive the ionizing stellar population by about a factor $\gtrsim 4$. Additionally, ^{26}Al is supposed to be ejected during the thermal pulse AGB phase at the end of the lifetime of the stars, hence no correlation between ionizing flux and AGB ^{26}Al production is expected from a coevally formed stellar association.

Another strong argument in favour of the massive star origin comes from a Bayesian viewpoint. First, the DMR free-free map represents the simplest model which could be imagined to explain the 1.8 MeV emission. Its only parameter is a single scaling factor, hence Ockham's reasoning discards more complex multicomponent models as proposed by, e.g., Chen *et al.* (1995), Diehl *et al.* (1996), and Diehl *et al.* (1997). Additionally, the scaling factor $X_{1.8}$ between the 1.8 MeV emission and the free-free emission model is consistent with current theories of ^{26}Al production by massive stars, hence the DMR free-free model has also a rather high prior probability. To see this, recall that the galactic Lyc photon production rate has been estimated to $Q = 3.5\,10^{53}$ ph s^{-1} on basis of observations of the N II cooling line by FIRAS (Bennett *et al.* 1994b). Using an O7 V star Lyc luminosity of $\log Q_0^{\rm O7V} = 49.12$ (Vacca *et al.* 1996) results in 26 550 'equivalent O7 V' stars in the Galaxy. With the determined equivalent O7 V yield of $Y_{26}^{\rm O7V} = (9.0 \pm 0.3)\,10^{-5}$ M$_\odot$, a total galactic ^{26}Al mass of 2.4 ± 0.1 M$_\odot$ is obtained, which is comparable to ^{26}Al yields obtained for $Z = 2$ Z$_\odot$ (cf. Tabs. 1.3 and 1.4 in Chapter 1). Note that the uncertainty in Q (possibly $\sim 50\%$; Bennett *et al.* 1994b) does not alter this consistency, since both the yield estimates and the galactic ^{26}Al mass were derived using the same value of Q. Uncertainties in Q will simply cancel out! The total galactic ^{26}Al mass, both as determined from the 1.8 MeV data and as estimated from nucleosynthesis calculations, however, is uncertain to about 50% due to the uncertainty in Q.

There are two very interesting and also very important points which can additionally be inferred from the comparison of the determined galactic ^{26}Al mass with ^{26}Al yield estimates. Firstly, for reasonable IMF slopes, ^{26}Al yields estimated using solar metallicity (cf. Tabs. 1.1 and 1.2) are significantly below the determined galactic ^{26}Al mass of 2.4 ± 0.3 M$_\odot$ (AGB and nova yields should be excluded in the comparison

since it already was discussed that these sources cannot contribute significantly to the galactic ^{26}Al budget). Thus if the ^{26}Al yield predictions are taken seriously, the 1.8 MeV observations exclude an average galactic metallicity of $Z = 1$ Z$_\odot$. Secondly, assuming an average galactic metallicity of $Z = 2$ Z$_\odot$, as indicated by the galactic metallicity gradient (e.g., Prantzos and Diehl 1996), makes the Wolf-Rayet stars the dominant source of galactic ^{26}Al, irrespectively of the assumed IMF slope. For reasonable IMF slopes ($\Gamma \gtrsim -1.7$), more than 70% of the predicted galactic ^{26}Al originates from WR stars, less than 30% comes from core collapse supernovae. This finding is in contradiction to estimates by other workers, which, however, suffer from inconsistencies in the yield estimates and/or the neglection of metallicity effects (e.g., Timmes *et al.* 1995, Prantzos and Diehl 1996). Recall, however, that ^{26}Al yields predictions by nucleosynthesis calculations are still rather uncertain, and the actual situation may indeed be different.

In summary, the following picture of the origin of galactic ^{26}Al arises: The tracer map comparison suggests that ^{26}Al is produced by massive stars, significant contributions from ONeMg-novae and AGB stars can be excluded. The scaling factor of the best fitting DMR free-free emission map is consistent with theoretical ^{26}Al yield estimates for WR stars and core collapse supernovae if an average galactic metallicity of $Z = 2$ Z$_\odot$ is assumed. For $Z = 2$ Z$_\odot$, nucleosynthesis calculations suggest that $2/3 - 3/4$ of the galactic ^{26}Al budget is provided by Wolf-Rayet stars, while core collapse supernovae account only for $1/4 - 1/3$ of the mass. Due to the increase of metallicity towards the galactic center, the inner Galaxy ^{26}Al production is probably dominated by WR stars, while core collapse supernovae provide most of the ^{26}Al in the outer, metallicity poor parts of the Galaxy. The total galactic ^{26}Al mass is 2.4 ± 0.1 M$_\odot$ with an additional uncertainty of 50% due to the uncertainty in the total Lyc luminosity Q of the Galaxy. In the following, it will be attempted to constrain the galactic ^{26}Al mass more tightly using 3-dimensional ^{26}Al distribution models.

4.4 Parameters of the ^{26}Al Distribution

Parameters of the ^{26}Al distribution, such as the galactic ^{26}Al mass or the scale length or scale height of the ^{26}Al density distribution, may be inferred from the comparison of 3-dimensional models of galactic ^{26}Al distribution with the COMPTEL 1.8 MeV data. To perform this comparison, the expected 1.8 MeV intensity distribution $I_{1.8}(l, b)$ (in ph cm^{-2}s^{-1}sr^{-1}) arising from the ^{26}Al distribution is derived by integration of the ^{26}Al mass density ρ_{26} (in M$_\odot$ pc^{-3}) along the lines of sight:

$$I_{1.8}(l, b) = 1.178\,10^4 \int \rho_{26}(l, b, s) ds. \tag{4.55}$$

s is the heliocentric distance along the line of sight which relates to the galactocentric distance R by

$$R^2 = R_\odot^2 + s^2 \cos^2 b - 2 s R_\odot \cos l \cos b, \tag{4.56}$$

where $R_\odot = 8.5$ kpc is the assumed distance of the Sun from the galactic center. The total galactic ^{26}Al mass is derived by integration of ρ_{26} over the galactic volume using

$$M = \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} \int_0^{R_{max}} \rho(l,b,s) s^2 \cos b \, db \, dl \, ds, \tag{4.57}$$

where $R_{max} = 15$ kpc is assumed as the outer radius of the ^{26}Al producing Galaxy. In the following, model intensity maps were derived for various ^{26}Al density distributions, normalized to a total galactic ^{26}Al mass of 1 $M_\odot$. The intensity maps were convolved into the COMPTEL data space where they are directly compared to the observed 1.8 MeV data. Free parameters of these fits are the scaling factor of the model (which corresponds directly to the total galactic ^{26}Al mass), the background scaling factor(s), and possible scale height or scale length of the intensity distribution.

4.4.1 The Radial ^{26}Al Distribution

Observations of edge-on galaxies suggest that the stars follow a double-exponential density profile

$$\rho(R,z) \propto \exp \frac{R}{R_0} \exp \frac{z}{z_0} \tag{4.58}$$

with parameters scale length R_0 and z_0 (e.g., Freeman 1992). Since ^{26}Al is believed to be produced in stars it is reasonable to assume that its distribution in the Galaxy can at least on average be approximated by such a profile. The fit of exponential disk models to COMPTEL 1.8 MeV data, which was already discussed in Section 4.2.2, results in MAP estimates of $M = 1.8 \pm 0.1$ $M_\odot$, $R_0 = 4.5 \pm 0.4$ kpc, and $z_0 = 90 \pm 25$ pc for the galactic ^{26}Al mass, the radial scale length, and the vertical scale height of the disk, respectively.

The Bayes factor analysis gives $2 \ln B = -16.2$ against the DIRBE 240 μm tracer map which places the exponential disk model among the best fitting models. Indeed, the residual analysis (Fig. D.14 in Appendix D) demonstrates that the exponential disk profile provides a reasonable first order description of the 1.8 MeV data. The χ^2 distance of the longitude profile is 1.656 which falls between the DMR dust and the DIRBE 240 μm model. Nevertheless, the exponential disk can not compete with the DMR free-free emission model since it is unable to explain the asymmetry between the northern and southern hemisphere and since it cannot account for the prominent emission feature seen in the Cygnus region.

Estimates of the exponential scale length of the Galaxy have been derived by numerous workers and span an enormous range from 1 to 6 kpc (Kent *et al.* 1991). The ^{26}Al scale length of 4.5 ± 0.4 kpc lies within this range, but is slightly above the often referenced standard value of 3.5 kpc (e.g., Bahcall and Soneira 1980). However, it is noteworthy that the determinations based on near-infrared data (mainly late K and M giants; cf. Section 4.3.1.7) all suggest small scale lengths between 1 and 3 kpc (Kent *et al.* 1991, Wainscoat *et al.* 1992, and references therein).

A more detailed image of the radial distribution of ^{26}Al in the Galaxy may be obtained by splitting a flat disk model into galactocentric rings and simultaneously fitting all rings to the 1.8 MeV data. This is a common technique for the determination of galactocentric density distributions which has been applied to far-infrared data by e.g. Bloemen *et al.* (1990) and Sodroski *et al.* (1997), or to high-energy γ-ray data by e.g. Strong *et al.* (1988).

To determine the radial ^{26}Al distribution, a flat disk with exponential scale height was splitted into 5 galactocentric rings. It turned out that if the scale height z_0 of the flat disk is left as a free parameter of the fit, an optimum value of 25 ± 45 pc is found, again smaller than the prior lower limit of 50 pc. From the experience of Section 4.2.2 it is clear that such a finding is probably related to the systematic uncertainties of the background model, hence a reasonable scale height of $z_0 = 90$ pc was chosen for the analysis (anyway, the choice of the scale height has negligible influence on the resulting radial distribution). The resulting ^{26}Al mass surface density is shown in Fig. 4.22. The corresponding mass estimate is $M = 1.6 \pm 0.2$ M$_\odot$ which is somewhat smaller but still consistent with the exponential disk mass. Notice that despite the coarse radial binning, Fig. 4.22 shows the typical radial profile of a massive star tracer with the strong density enhancement below ~ 6 kpc due to the 'molecular ring' (e.g., Bronfman 1992), and a local enhancement around 8.5 kpc due to the local spiral arm (e.g., Comerón and Torra 1996).

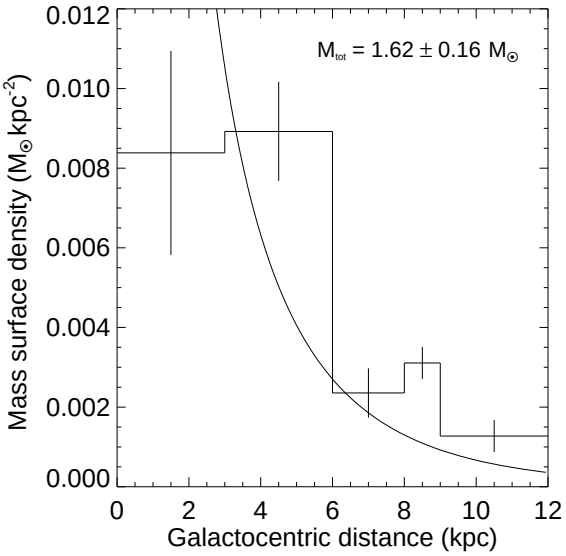

Fig. 4.22: *^{26}Al mass surface density as derived from the fit of a flat disk splitted in galactocentric rings. The solid line indicates the corresponding mass surface density of the best exponential disk disk model.*

4.4.2 The Scale Height of Radioactive ^{26}Al

The scale height of the galactic ^{26}Al distribution may in principle be determined from fits of parametric models where z_0 is a free parameter. Notice, however, that with an angular resolution of 3.8° (FWHM) at 1.8 MeV, COMPTEL is not very sensitive to the exact scale height of the distribution, at least for small scale heights as expected for a young stellar population. In an earlier work, Knödlseder *et al.* (1996c) determined the galactic ^{26}Al scale height to $z_0 = 180^{+240}_{-130}$ pc from fits of exponential disk models to galactic plane data of mission phases 1 & 2. Recently, Diehl *et al.* (1997) used the

all-sky dataset to re-determine the exponential disk scale height to $z_0 = 130^{+100}_{-75}$ pc (3σ errors), a value slightly different from the MAP estimate of $z_0 = 90 \pm 25$ pc presented in this work (Diehl *et al.* (1997) obtain a slightly higher scale height since they assumed a radial scale length of 5 kpc for their models; cf. Fig. 4.7).

However, the fit of the split disk model illustrates again that the scale height determination is easily biased by the uncertainties of the background model. Due to these systematic uncertainties, the quoted values should probably not be taken too literally. Nevertheless, recently Oberlack (1997) estimated the scale height of the 1.8 MeV emission to $z_0 = 170 \pm 50$ pc on the basis of maximum entropy reconstructions. He derived this value by comparing the latitude profile of the 1.8 MeV all-sky map with latitude profiles from reconstructions of various simulated mock datasets. Notice that the width of the reconstructed latitude profile of mock data is rather sensitive to the assumed galactic ^{26}Al mass M, leading to wider profiles for smaller mass (cf. Fig. 3.4). Since Oberlack (1997) assumed $M = 3$ M$_\odot$ for his exponential disk simulations, but the galactic ^{26}Al mass probably falls below this value, his scale height of 170 ± 50 pc should be taken as an upper limit. Thus, it can be concluded that the ^{26}Al distribution has a scale height of about $50 - 220$ pc, which corresponds to typical values for the young or intermediate age stellar populations. On the other hand, it constrains the typical dispersion velocity of ^{26}Al to $v \lesssim 250$ km s^{-1}, which is inconsistent with the 1.8 MeV line broadening as observed by *GRIS* (Naya *et al.* 1996) if the broadening is interpreted as a Doppler broadening.

4.4.3 The Galactic ^{26}Al Mass

The total galactic ^{26}Al mass may be derived from the scaling factors which were assigned to the models by the fitting procedure. It is important to notice, however, that the galactic ^{26}Al mass is rather sensitive to the scale height of the model distribution, which in turn is subject to substantial uncertainties (see above). The relation of galactic ^{26}Al mass to scale height is illustrated in Fig. 4.7 where the fit result for M is plotted versus the scale length and scale height of the exponential disk model. If the scale height uncertainty is taken to be $z_0 = 50 - 220$ pc, the galactic ^{26}Al mass for the exponential disk models varies between $M = 1.6 - 2.2$ M$_\odot$. Recall that the MAP estimate was $M = 1.8 \pm 0.1$ M$_\odot$, hence the formal statistical error is possibly underestimated by about a factor of ~ 3.

The following 3-dimensional models were compared to the 1.8 MeV data in order to estimate the galactic ^{26}Al mass: (1) the exponential disk model with radial scale length R_0 as free parameter, (2) the flat disk, split in 5 galactocentric rings, (3) a spiral arm and molecular ring model, and (4) a spiral arm + molecular ring model on top of an exponential disk with radial scale length R_0 as free parameter. The spiral arm model is that of Taylor and Cordes (1993) (TC) who adjusted a four-arm spiral pattern based on giant H II regions and radio-survey tangent points to pulsar dispersion measures. In addition to the four spiral arms, the model includes a molecular ring component

Model	$2 \ln B$	M_{26} (in $M_\odot$)	χ^2
Exponential disk ($R_0 = 4.5$ kpc, $z_0 = 90$ pc)	-16.2	1.8 ± 0.1	1.656
Galactocentric rings ($z_0 = 90$ pc)	-4.8	1.6 ± 0.2	1.421
Spiral arms + molecular ring ($z_0 = 90$ pc)	-59.3	1.4 ± 0.1	3.042
Exp. disk + arms + ring ($R_0 = 7.0$ kpc, $z_0 = 90$ pc)	13.8	1.8 ± 0.2	1.342

Table 4.4: *Galactic ^{26}Al masses as obtained by fits of 3-dimensional ^{26}Al distribution models. Log-Bayes factors $2 \ln B$ are given with respect to the DIRBE 240 μm tracer map. χ^2 is the usual measure of the longitude profile fit quality (for the DMR free-free map, $\chi^2 = 1.122$).*

with ring radius of 4 kpc and radial exponential density profile of scale length 1.8 kpc. For consistency with the other models, an exponential profile has been adopted for the vertical density distribution. Notice that Taylor and Cordes (1993) proposed their model for the quantitative description of the distribution of free electrons in the Galaxy. In addition to the spiral arms and the molecular ring components used in this work, their model contains the Gum nebula and a thick disk component, to account for the Reynolds layer, as explicit components. Notice, however, that their model doesn't account for the important number of free electrons seen towards the Cygnus region. Further, Heiles *et al.* (1996) pointed out that TC's 'molecular ring' component is not associated with strong H II regions, and neither the diffuse ionized gas nor H II regions are distributed in rings. Thus it is clear that it cannot be expected that the TC free-electron model gives a completely satisfactory description of the 1.8 MeV data.

The results of the model comparison are compiled in Tab. 4.4. For consistency and due to the systematic uncertainties associated with the scale height of the ^{26}Al distributions, the scale height of all models was fixed to $z_0 = 90$ pc. The first column indicates the model and the best fitting radial scale length (if it was a parameter of the model), the second column gives the log-Bayes factors against the DIRBE 240 μm tracer map (i.e. the tracer map with the highest Bayes factor), the third column quotes the derived galactic ^{26}Al mass, and the forth column shows the χ^2 distance between the software-collimation longitude scan of model and data.

The χ^2 distance shows that none of the 3-dimensional ^{26}Al distribution models can explain the 1.8 MeV longitude profile as well as the DMR free-free emission map. Indeed, best results are obtained for the spiral arm + molecular ring model on top of an exponential disk, and for a flat disk split into 5 galactocentric rings. While the first of these models can account for some of the asymmetries in the 1.8 MeV data due to the presence of galactic spiral arms, it is too stiff to describe prominent 1.8 MeV features such as Cygnus. In contrast, the split disk has enough free parameters to account (at least partly) for deviations from a smooth emission profile, but its axisymmetry hinders a satisfactory explanation of 1.8 MeV asymmetries.

Table 4.4 shows that the best fitting models all result in a total ^{26}Al mass around $M \approx 1.7$ $M_\odot$. Only the 'spiral arms + molecular ring' model assigns a marginal lower galactic ^{26}Al mass, but the high χ^2 distance of 3.042 indicates that this model

	$\Gamma = -1.35$		$\Gamma = -1.7$	
M_{26} ($M_\odot$)	1.5	2.2	1.5	2.2
WR star rate (Myr^{-1})	2130 ± 90	3120 ± 130	2250 ± 90	3300 ± 140
SN II rate ($century^{-1}$)	0.9 ± 0.1	1.4 ± 0.1	1.5 ± 0.1	2.2 ± 0.1
WR ^{26}Al yield ($M_\odot$ Myr^{-1})	1.0 ± 0.1	1.4 ± 0.1	0.9 ± 0.1	1.3 ± 0.1
SN II ^{26}Al yield ($M_\odot$ Myr^{-1})	0.14 ± 0.01	0.21 ± 0.01	0.21 ± 0.01	0.31 ± 0.01
SN Ib/c ^{26}Al yield ($M_\odot$ Myr^{-1})	0.13 ± 0.01	0.20 ± 0.01	0.14 ± 0.01	0.20 ± 0.01
Star formation rate ($M_\odot$ yr^{-1})	1.7 ± 0.1	2.5 ± 0.1	8.7 ± 0.4	12.8 ± 0.5
Q (10^{53} ph s^{-1})	2.2 ± 0.1	3.2 ± 0.1	2.2 ± 0.1	3.2 ± 0.1

Table 4.5: *Some galactic parameters as derived from the analysis of COMPTEL 1.8 MeV all-sky data. An average galactic metallicity of $Z = 2$ $Z_\odot$ was assumed. The star formation rate was derived for the mass interval $0.08 - 100$ $M_\odot$. Q is the total Lyc luminosity of the Galaxy.*

describes the 1.8 MeV data poorly. Typical statistical uncertainties are 0.2 $M_\odot$, but the uncertainty due to the poorly known vertical scale height of the ^{26}Al distribution may be larger. Taking together the statistical and systematical uncertainties the total galactic ^{26}Al mass probably lies in the interval $M_{26} = 1.5 - 2.2$ $M_\odot$.

Using this mass estimate, various galactic parameters can be determined as follows: From the DMR free-free emission map, a total galactic ^{26}Al mass of 2.4 ± 0.1 $M_\odot$ was derived which is only based on the relation Eq. (4.50) between the ^{26}Al mass density and the equivalent O7 V star density, and a total Lyc luminosity of $Q = 3.5\,10^{53}$ ph s^{-1} (cf. Section 4.3.3.5). Galactic parameters were derived in Chapter 1 using the same galactic Lyc luminosity, hence the COMPTEL ^{26}Al mass estimation of $M = 1.5 - 2.2$ $M_\odot$ can be used to 're-calibrate' these parameters. It is important to recognize that the results are independent of theoretical ^{26}Al nucleosynthesis calculations, in contrast to estimates proposed by Timmes *et al.* (1997) which rely on the assumption that only SN II produce the galactic ^{26}Al.

The resulting galactic parameters are shown in Tab. 4.5 for a total galactic ^{26}Al mass of 1.5 $M_\odot$ and 2.2 $M_\odot$, and a lower mass limit of $M_{WR} = 25$ $M_\odot$ for stars to exhibit a Wolf-Rayet phase. The quoted errors are statistical uncertainties arising from the fit of the DMR free-free map to COMPTEL 1.8 MeV data. Values are shown for two 'reasonable' IMF slopes of $\Gamma = -1.35$ and -1.7; notice however that the slope of the IMF may flatten for low-mass stars (e.g., Miller and Scalo 1979), hence the star formation rate may be overestimated, in particular for $\Gamma = -1.7$. Wolf-Rayet star ^{26}Al yields were estimated by averaging the yields from Langer *et al.* (1995) and Meynet *et al.* (1997), SN II yields were taken from Woosley and Weaver (1995), and SN Ib/c estimates come from Woosley *et al.* (1995). Notice that these yields are based on theoretical nucleosynthesis calculations and stand in no relation to the observed galactic ^{26}Al mass. The fact that the predicted WR and the core collapse SN yields sum-up to the observed ^{26}Al mass M_{26} simply means that the estimates agree within

the statistical uncertainties!

4.4.4 A Local ^{26}Al Component?

In this last section, a question is addressed which was already raised several times in this work: Is there a local ^{26}Al component? A local ^{26}Al origin was proposed by Clayton (1984), Morfill and Hartquist (1985), and Blake and Dearborn (1989) at a time when nothing was known about the distribution of the 1.8 MeV emission on the sky and additionally, the proposed ^{26}Al candidate sources failed to explain the observed 1.8 MeV luminosity (e.g., Prantzos and Diehl 1996). The present analysis, however, clearly demonstrates that radioactive ^{26}Al is a galaxywide phenomenon and a local origin can be clearly excluded. Nevertheless, if ^{26}Al is produced galaxywide it is also produced locally, and nearby candidate sources could give rise to spectacular 1.8 MeV features on the sky.

To search for such features, the 1.8 MeV all-sky residual map as derived by the maximum likelihood ratio test after subtraction of the DMR free-free emission model is shown in Fig. 4.23. Remember that this map was obtained by independent fits of point source models on top of the DMR free-free and the background models for a $1^\circ \times 1^\circ$ grid of assumed source positions. Due to the telescope's angular resolution of 4° (FWHM), neighboring skymap pixels are not independent and consequently 1.8 MeV residuals will appear somewhat extended. In particular, residuals arising from diffuse extended emission will show up as connected irregular low significance features in this map.

Two promising candidate regions for local 1.8 MeV emission can be identified from Fig. 4.23: Firstly, there is an important feature mainly located within galactic longitudes $240^\circ < l < 30^\circ$ below galactic latitude $b < -30^\circ$ in the map which exceeds the 2σ significance level over substantial areas, and even obeys two emission peaks of 4σ. However, the search for line signatures in software-collimated spectra for this region gave no indication for 1.8 MeV line emission, neither for spectra located on the most prominent emission spots (cf. Fig. 3.19) nor for integrated spectra of the region. Secondly, the ACP feature between galactic longitudes $l \approx 140^\circ - 180^\circ$ and latitude $b \pm 30^\circ$ is still present in the residual map, although the intensity maximum shifted somewhat to positive latitudes. However, the search for line signatures towards the maximum of the residual (around $(l, b) \approx (160^\circ, 20^\circ)$) were also not fruitful (cf. Fig. 4.24). It appears that only in a rather small area around $(l, b) \approx (166^\circ, -13^\circ)$ 1.8 MeV line emission is detected, as illustrated by the nice spectrum given in Fig. 2.5. Hence, the ACP feature probably is a combination of real 1.8 MeV emission and systematic background uncertainties which gives rise to the prominent feature in the all-sky maps. Notice that an 'attraction' feature is present in this region of the sky (cf. Fig. 4.4) which could be at the origin of the appearance of the ACP emission.[16]

[16] Remember that 'attraction' features only attract if 1.8 MeV emission is present in their vicinity.

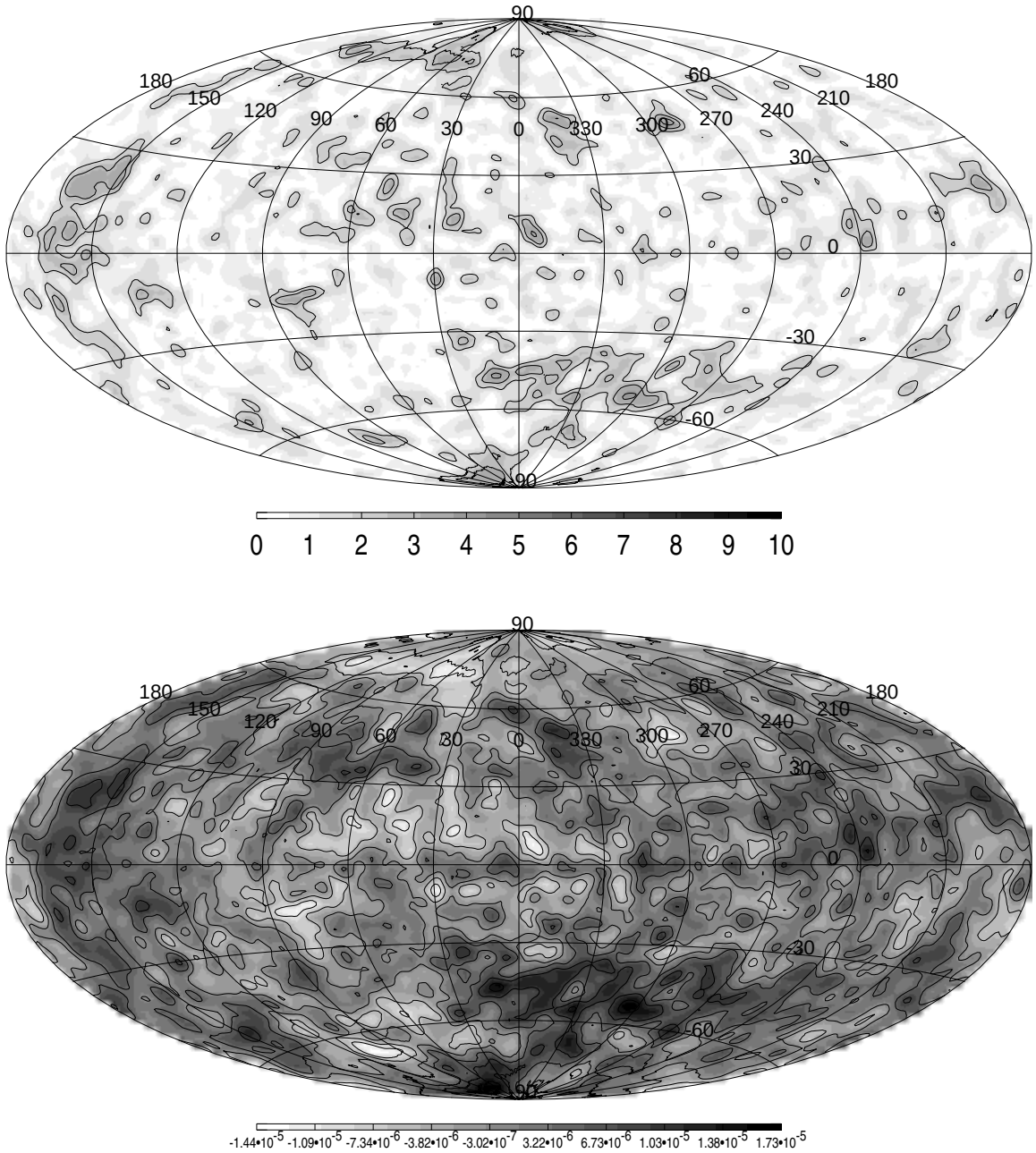

Fig. 4.23: *Residuals in the COMPTEL 1.8 MeV sky after subtraction of the DMR free-free emission model.* Top: *Point source significances in Gaussian σ as derived from the maximum likelihood ratio test for $q = 1$ degrees of freedom (de Boer* et al. *1992).* Bottom: *Corresponding point source fluxes in units of ph $cm^{-2}s^{-1}$.*

Software-collimated energy spectra were also investigated towards all other suspicious residual features, again with no positive results. Two examples are shown in Fig. 4.24 for the features near Virgo at $(l, b) = (342°, 53°)$ and Puppis at $(l, b) = (238°, 5°)$.

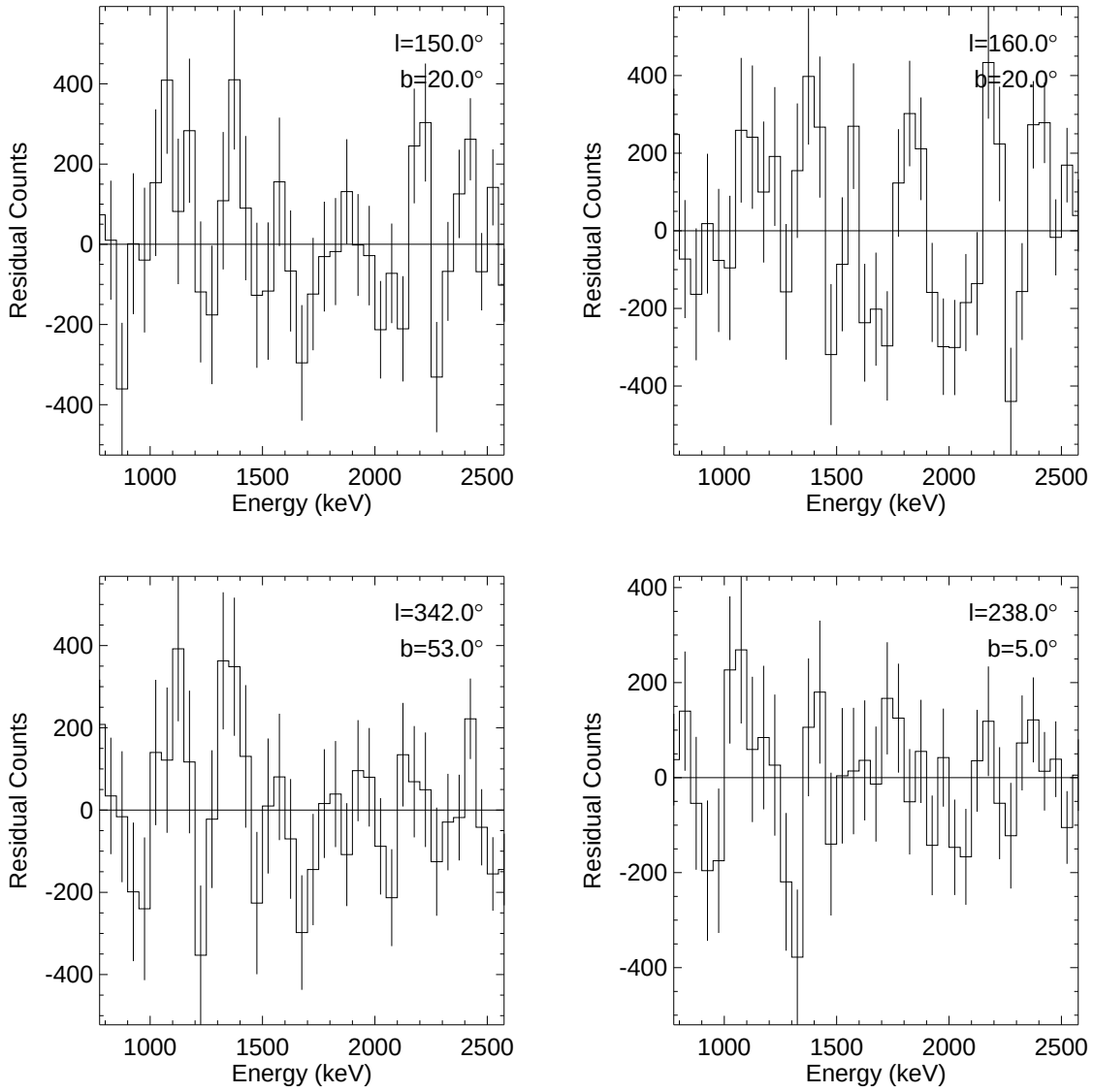

Fig. 4.24: Top: *Background subtracted COMPTEL energy spectra taken towards the maximum of the ACP residual.* Bottom: *Examples of spectra towards two features in the residual map.*

In conclusion, no evidence for a local ^{26}Al component which can not be accounted for by the DMR free-free emission model is found in COMPTEL 1.8 MeV data for observation periods 0.1 - 522.5. The residuals, in particular the extended excess in the

southern galactic hemisphere, must be interpreted as systematic error, which probably arises from shortcomings of the instrumental background model. Remember that already in the ψ distribution of the background-subtracted events, a clear asymmetry between positive and negative latitudes was found (cf. Section 2.5.2.2).

Chapter 5

A New Global View of Galactic ^{26}Al

5.1 The 1.8 MeV Sky after COMPTEL

On 1984 November 15, Mahoney *et al.* (1984) informed the scientific community about the discovery of ^{26}Al in the interstellar medium by means of its distinctive fingerprint – a nuclear gamma-ray line at 1.809 MeV. This discovery, however, hardly came as a surprise since nucleosynthesis calculations in the late 60s had shown that substantial amounts of ^{26}Al may be produced during explosive carbon burning in supernovae (Arnett 1969). However, the observed 1.8 MeV luminosity from the galactic center direction was even more intense than originally estimated, which pushed theoreticians to look for other ^{26}Al sources. As a consequence, many candidate ^{26}Al sources and/or ^{26}Al production scenarios were proposed in order to explain the observed 1.8 MeV line. To note only two extreme examples, Hillebrandt *et al.* (1987) suggested the explosion of a single high-metallicity supermassive star near the galactic center, while Morfill and Hartquist (1985) proposed a recent local supernova origin. The result was that at the end of the 80s, the origin and distribution of ^{26}Al in the Galaxy was more uncertain than ever.

Meanwhile, the community of observers had also been very active, aiming in improving the sensitivity and angular resolution of their instruments. Starting the development in 1971 at the Max-Planck-Institut für extraterrestrische Physik (MPE, Garching, FRG), Schönfelder *et al.* (1973) proposed a new concept for the observation of MeV γ-rays using a 'double-Compton' telescope. In a series of balloon flights, the concept was proven very successful, which lead to the selection of a Compton telescope (COMPTEL) as one of the four instruments on NASA's second 'Great Observatory' CGRO. Launched on 1991 April 5 by the Space Shuttle Atlantis, COMPTEL is now in orbit since more than 6 years, and provided (and still provides) an unprecedented view of the γ-ray sky at MeV energies (for a review see Schönfelder *et al.* 1996).

One of the highlights of the COMPTEL mission is unquestionably the observation of the 1.8 MeV γ-ray line. With the first mapping of the Galaxy in the light of 1.8 MeV

photons (Knödlseder 1994, Diehl *et al.* 1995b, Oberlack *et al.* 1996) the discussion of the origin and distribution of galactic ^{26}Al got new updraft (for a review see Prantzos and Diehl 1996). The outstanding feature of the COMPTEL 1.8 MeV map is emission along the galactic plane which clearly excludes single point source or local ^{26}Al origin scenarios – ^{26}Al production appeared to be a galaxywide phenomenon. The lumpiness of the skymap (cf. Fig. 3.12) has been often emphasized (e.g., Diehl *et al.* 1995b, Schönfelder *et al.* 1996), which lead to the unsuitable notion of 'hot spots' in the 1.8 MeV sky. It was clearly established that distinct emission features in the Cygnus, Carina, and Vela regions were incompatible with a smooth emission distribution, suggesting that massive stars are at the origin of ^{26}Al production. The remaining 'hot spots', however, were most easily explained by statistical fluctuations only (Knödlseder 1994, Oberlack *et al.* 1996, Knödlseder *et al.* 1996b). Nevertheless, the 'hot spots' have been interpreted as signs of nearby localized ^{26}Al sources which possibly could lower the galactic ^{26}Al mass below ≤ 1 M$_\odot$ (Diehl *et al.* 1995b), placing the Sun in a privileged region of the Galaxy. Others tried to explain the 'hot spots' by tuning galaxywide ^{26}Al distribution models (e.g., Chen *et al.* 1996).

A more reliable image of the distribution of 1.8 MeV emission is presented in this work using a multiresolution reconstruction based on wavelets (cf. Fig. 3.18). It is more reliable in the sense that it does not overemphasis irregularities in the skymap arising from statistical uncertainties in the data. Notice that it cannot be excluded that the 1.8 MeV sky is composed of individual point sources, but with the sensitivity and angular resolution of COMPTEL it can not be decided whether this is the case or not. In this sense, the wavelet reconstruction presented in this work can be considered as more objective than the 'classical' maximum entropy or maximum likelihood reconstructions, since it does not show features which possibly do not exist. Indeed, application of this method to COMPTEL 1.8 MeV data leads to a rather smooth skymap with distinct emission features in the Cygnus, Carina, Vela and the anticenter regions. Most of the 'hot spots' along the galactic plane have disappeared.

The most important discovery presented in this work, however, is the close correlation of the 1.8 MeV emission with free-free emission arising from the ionized ISM. This finding was only possible thanks to (1) an unbiased tracer map comparison established by using all-sky maps from the entire explored electromagnetic spectrum (from radio to high-energy γ-rays), (2) a deeper understanding of the systematic uncertainties in COMPTEL model comparison, mainly arising from uncertainties of the background model, and, of course, (3) the availability of an all-sky free-free emission map which has been derived only 5 years ago by Bennett *et al.* (1992). The beauty in this close correlation is that the entire 1.8 MeV distribution can be explained by a simple, one-parametric model, dismissing the earlier claims for the need of multicomponent models to explain the 1.8 MeV sky. In particular, there is no need to place the Sun in a privileged region of the Galaxy.

Additionally, the single free parameter in this correlation, the scaling factor of the free-free emission model, is not even free if one trusts current nucleosynthesis calcu-

lations. Since the free-free intensity is directly proportional to the number of O stars along the line of sight, the expected 1.8 MeV intensity can be calculated under the assumption that massive stars are at the origin of galactic ^{26}Al. For a mean average galactic metallicity of two times solar, as suggested by observations of the galactic metallicity gradient, the predicted 1.8 MeV map is in perfect agreement with the observations, both in distribution and in intensity level. Hence it is very suggestive that indeed massive stars are at the origin of galactic ^{26}Al.

Massive stars in this context means stars initially more massive than $M \gtrsim 20$ M$_\odot$. Only these stars lead to significant ionization of the ISM, giving rise to the observed free-free emission. In particular, ONeMg-novae and AGB stars (even massive ones) can be clearly excluded as major ^{26}Al producers, because their distribution should not be correlated with free-free emission. Obviously, one could speculate about a blend of contributions from different source types together with specific assumptions about their correlation to the massive star population. Yet there is no need for a composite model, neither from the 1.8 MeV data, nor from nucleosynthesis calculations.

Although the massive star origin seems certain, it is difficult to judge whether galactic ^{26}Al is mainly produced by core collapse supernovae or massive stars during the Wolf-Rayet phase. Current nucleosynthesis estimates, however, favour WR stars as primary origin, accounting for roughly $2/3 - 3/4$ of the galactic ^{26}Al mass, while SN II should only contribute $1/4 - 1/3$ of the mass.

5.2 The 1.8 MeV Sky before INTEGRAL

The identification of massive stars as the origin of galactic ^{26}Al is far from being the end of the interest in the 1.8 MeV line. On the contrary! It is the beginning of a serious astrophysical exploitation of nuclear gamma-ray data. In Section 4.4.3 it has been shown how observations of the 1.8 MeV gamma-ray line may be used to derive fundamental parameters of the Galaxy, even independently of theoretical nucleosynthesis calculations. Among these are the galactic star formation rate, the number of Wolf-Rayet stars, the core-collapse supernova rate, or the galactic Lyc luminosity. Usually these parameters are derived from observations of local source populations and extrapolation to the entire Galaxy, from comparison with other Galaxies, or from observations of the ionized ISM, either locally or galaxywide. It should be pointed out that the study of 1.8 MeV data not only complements these investigations, but also enriches the available data by qualitatively new information.

In particular, galactic rotation should lead to distinct γ-ray line profiles due to Doppler shifts, implying important information about the distribution of ^{26}Al, hence massive star formation, along the lines of sight (Skibo and Ramaty 1991, Gehrels and Chen 1996). Although 1.8 MeV line profiles are not measurable by COMPTEL due to the insufficient energy resolution, future telescopes employing cooled germanium (Ge) detectors are capable of resolving nuclear lines. Indeed, with the International

Gamma-Ray Astrophysics Laboratory (INTEGRAL) such an instrument is currently being developed. The INTEGRAL observatory, that is scheduled for launch in 2001 (e.g., Winkler 1996) consists of two main gamma-ray instruments, each of them with both spectral and angular resolution, but differently optimized in order to complement each other. The high-resolution γ-ray spectrometer SPI will achieve a sensitivity of $7\,10^{-6}$ ph cm^{-2}s^{-1} (3σ; T$_{obs} = 10^6$ s), an energy resolution of 3 keV (FWHM), and an angular resolution of 2° (FWHM) at 1.8 MeV. With these performances it should be possible to establish a coarse 3-dimensional mapping of star formation activity in our Galaxy. This mapping will significantly enhance our understanding of the spiral arm structure of the Galaxy.

Yet many other questions related to galactic radioactivity wait for an answer. Among those are the existence of the galactic center 'hot spot', which was suggested to arise from a galactic 'bar' (Hartmann 1994, Chen *et al.* 1995), or the origin of the 1.8 MeV emission in Vela, tentatively identified with the Vela supernova remnant (Oberlack *et al.* 1994). The identification of an individual ^{26}Al source would be of great interest for the direct verification, or even calibration, of theoretical nucleosynthesis calculations. In this context, the Cygnus region may play a key role in future investigations of the 1.8 MeV line. This region is generally interpreted as the local spiral arm seen tangentially over a length of at least 3 kpc. The proximity of the region together with the presence of eight young stellar associations, numerous WR stars, prominent H II regions, and several supernova remnants make it a primary candidate for the study of the interplay of ^{26}Al production, massive stars, ionization, mixing phenomena with the ISM, supernova explosions, and star formation in general.

Appendix A

Performance of the Richardson-Lucy Algorithm

A.1 Algorithm Parameters

To illustrate the R-L reconstruction process the following parameters are shown in Fig. A.1 for the exponential disk simulation versus iteration r:

The log-likelihood ratio $-2\ln L_0 = -2\ln\left(P(f_j^r)/P(\mathbf{b})\right)$ where $P(f_j^r)$ is the likelihood of the reconstructed intensity distribution at iteration r and $P(\mathbf{b})$ is the likelihood when solely the background model is fitted to the data. For the R-L reconstruction $-2\ln L_0$ should monotonically increase with proceeding iterations since it is a convex function of the likelihood $P(f_j^r)$. Following the theory of maximum likelihood ratio testing, $-2\ln\mathcal{L}_0$ is a measure for the significance of the detected emission (Cash 1979).

The log-likelihood ratio $-2\ln L_\Delta = -2\ln\left(P(f_j^r)/P(f_j^{r-1})\right)$ as a measure of the likelihood increment between successive iterations. This parameter visualizes the convergence properties of the algorithm; for convergence $-2\ln L_\Delta \to 0$.

The R-L acceleration factor λ^r which measures how much faster convergence is reached in contrast to the unaccelerated algorithm.

The total flux in the reconstructed image (solid line) together with the total flux in the model intensity distribution (dashed line). While the original R-L algorithm always predicts as many events as measured, the inclusion of a fixed background model in the measurement prediction destroys this property. Consequently it can not be expected that the reconstructed total flux is identical to that of the emission model. The question on flux bias will be adressed in more detail in Section A.3.

Fig. A.2: *Some parameters of the R-L reconstruction of 20 simulated point sources versus iteration r (see text).*

Figure A.1 shows that about 75% of the maximum log-likelihood ratio $-2 \ln L_0$ is obtained after only 5 iterations. A comparison with Fig. 3.1 reveals that during these iterations the large scale structure of the emission is recovered. The remaining iterations increase the likelihood only slowly while the diffuse emission breaks up into individual spots. From the statistical point of view, the maximum log-likelihood ratio which may be obtained without overfit can be determined by fitting the exponential disk model to the mock data (cf. Chapter 4). The result, shown as dotted line in the upper left panel of Fig. A.1, confirms that the data become overfitted at least after iteration 5 (overfit might even set in earlier since overfit in some regions might be compensated by misfit in other regions). The R-L acceleration factor shows an oscillating behaviour which indicates that the optimum λ^r was never heavily influenced by the positivity constraint. The oscillations propagate into the total reconstructed flux where they become damped with proceeding iterations. In average, the total flux in the reconstructed image is always overestimated, rising from about 5% at early iterations towards 10% at the end of the reconstruction.

For the point source simulation, the evolution of the log-likelihood ratio is similar to the exponential disk reconstruction: about 65% of the maximum likelihood ratio is obtained after 5 iterations which is also the point where overfits sets in (for comparison the log-likelihood ratio determined from the model fit is shown as dotted line). The R-L acceleration factor was always constrained by the positivity of the intensities which avoided the oscillations seen for the exponential disk reconstruction. Only at late iterations, acceleration becomes chaotic which is probably due to roundoff errors in the lowest intensities. Nevertheless, the total flux seems to behave rather stable, converging to an overestimated level of about 25%.

A.2 The Stopping Criterion

Figures 3.1 and 3.3 clearly show the effect of overfit in late iterations (cf. Section 3.2.1). For the exponential disk simulation the initially diffuse emission breaks up into many individual point-like sources situated mainly in regions of high 1.8 MeV intensity. Additional sources at intermediate and high latitudes occur in regions where no signal was present in the model. From the comparison of the late iterations with the model it is clear that these late iterations are a poor description of the reality and may be misleading since they suggest many individual sources where only diffuse emission is present.

The general approach to avoid spurious sources is stopping the iterations before overfit becomes important. A comparison of Fig. 3.1 with Fig. 3.3 shows, however, that the optimum iteration might depend on the underlying intensity distribution: while for diffuse emission overfit sets in very early (about iteration 5) it becomes important for the point source simulation only at late iterations ($r > 20$). Figure 3.2 also demonstrates that stopping the R-L reconstruction after only a few iterations

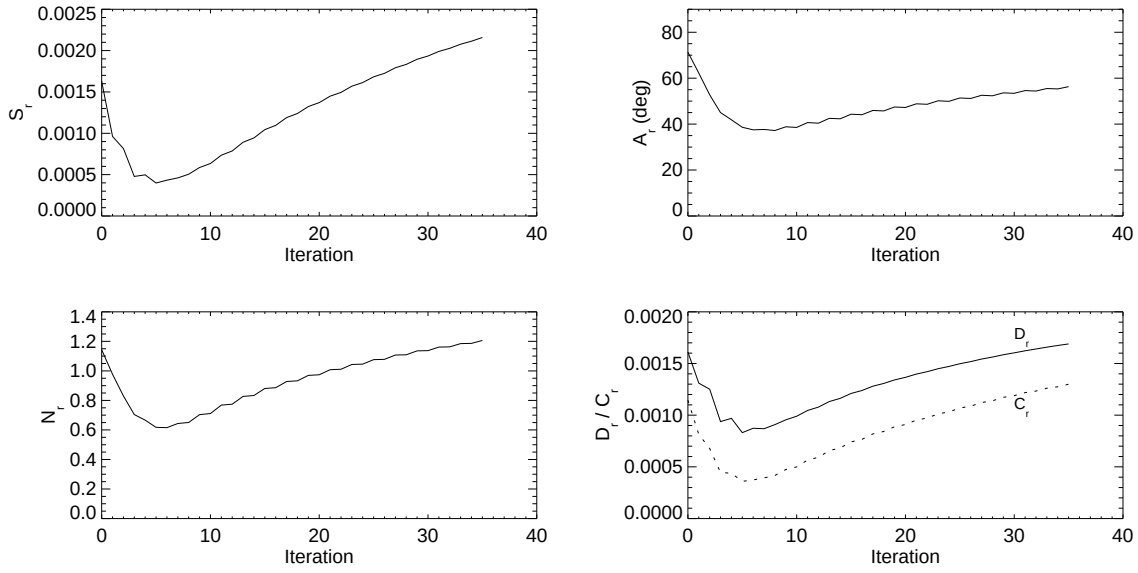

Fig. A.3: Distance measures for the reconstruction of the exponential disk data.

leads to an overestimation of the width of the latitude distribution. Additionally, the flux of emission features is systematically underestimated for early iterations (cf. Section A.3). Hence it becomes clear that the choice of an 'optimum' iteration depends on the question in mind and that there could not be a single optimum iteration.

To find the iteration which comes closest to the model intensity distribution, the 'distance measures'

$$\begin{aligned} S_r &= \sum_{j=1}^{M} f_j^r - h_j - f_j^r \ln\left(\frac{f_j^r}{h_j}\right) \\ A_r &= \arccos\left\{\frac{\sum_{j=1}^{M} f_j^r h_j}{\sqrt{\sum (f_j^r)^2}\sqrt{\sum h_j^2}}\right\} \\ N_r &= \sum_{j=1}^{M}\left|\frac{f_j^r}{\sum f_j^r} - \frac{h_j}{\sum h_j}\right| \\ D_r &= \sum_{j=1}^{M} |f_j^r - h_j| \\ C_r &= \sum_{j=1}^{M} \frac{(f_j^r - h_j)^2}{f_j^r + h_j} \end{aligned} \tag{A.1}$$

are calculated for each iteration, where f_j^r is the reconstruction after iteration r and h_j is the model intensity distribution. The measures differ in the way how 'closest'

is defined. In light of section 3.1.2 the entropy S_r has probably most justification as a measure of spurious sources in the reconstruction but it is also sensitive on the total flux in the reconstructed images. A_r and N_r measure how close the shape of the reconstructed image comes to the model regardless of the total flux. A_r is the angle between the reconstructed image vector $\mathbf{f}^r$ and the model vector $\mathbf{h}$ using the euclidian metric. If both vectors are parallel ($A_r = 0$) the shape of the reconstructed image is identical to that of the model. Finally, D_r measures the absolute deviation of the reconstruction and C_r is the Chi-square statistics. Like the entropy both measures depend on the total reconstructed flux and the shape of the intensity distribution.

Figure A.3 shows the distance measures for the reconstruction of the exponential disk mock data. All measures consistently show a clear minimum between iterations $5-8$. The jagged shape of S_r, D_r, and C_r during early iterations reflects the fluctuations of the total flux between subsequent iterations (cf. Fig. A.1). Consequently, the minimum occurs somewhat earlier for these quantities due to a minimum in total flux at iteration 5. The reconstructed emission shape, as measured by A_r and N_r, comes closest to the model between iterations $5-8$. Comparison with Fig. 3.1 confirms that these iterations show a rather good compromise between smoothness and latitude extension of the emission distribution. While earlier iterations are smoother they obey a much wider latitude distribution, hence lie far away from the model. In later iterations the latitude distribution is well represented but the reconstructions are much more structured than the model. The minimum in the distance measures for diffuse emission is a fundamental property of the R-L algorithm. In this sense there exists an optimum iteration for the reconstruction of diffuse emission. Figure 3.4 demonstrates that the optimum iteration slightly increases with galactic ^{26}Al mass if the natural initial intensity is chosen for the reconstruction. For an expected galactic ^{26}Al mass between $2-3$ $M_\odot$ the minimum occurs around iteration $6-7$. Figure 3.4 also shows that with increasing ^{26}Al mass the optimum reconstruction comes closer to the model distribution. This reflects the increase of signal-to-noise ratio with total galactic ^{26}Al mass which reduces the statistical uncertainty in the reconstruction.

For the point source simulation the situation is different (Fig. A.4). In this case the model intensity distributions is basically zero except for the 20 point source pixels. Zero model intensity pixels have to be omitted in the entropy calculation, hence S_r is only composed of 20 pixels. Consequently, S_r can't be considered as measure of spurious sources in the image but merely as a measure of flux concentration at the point source locations. Except for D_r and C_r, all measures decrease monotonically with proceeding iterations, indicating that the ML solution is the optimum reconstruction. D_r and C_r show a maximum around iteration 5 which reflects the rise in total flux during the first iterations (cf. Fig. A.2). From the maximum both measures decrease monotonically which confirms the superiority of late iterations. Nevertheless, even for the ML reconstruction the angle between the reconstructed image vector and the model remains large ($\sim 60°$) as a result of spurious point sources in the reconstruction.

The R-L reconstructions may also be used to derive longitude and latitude profiles

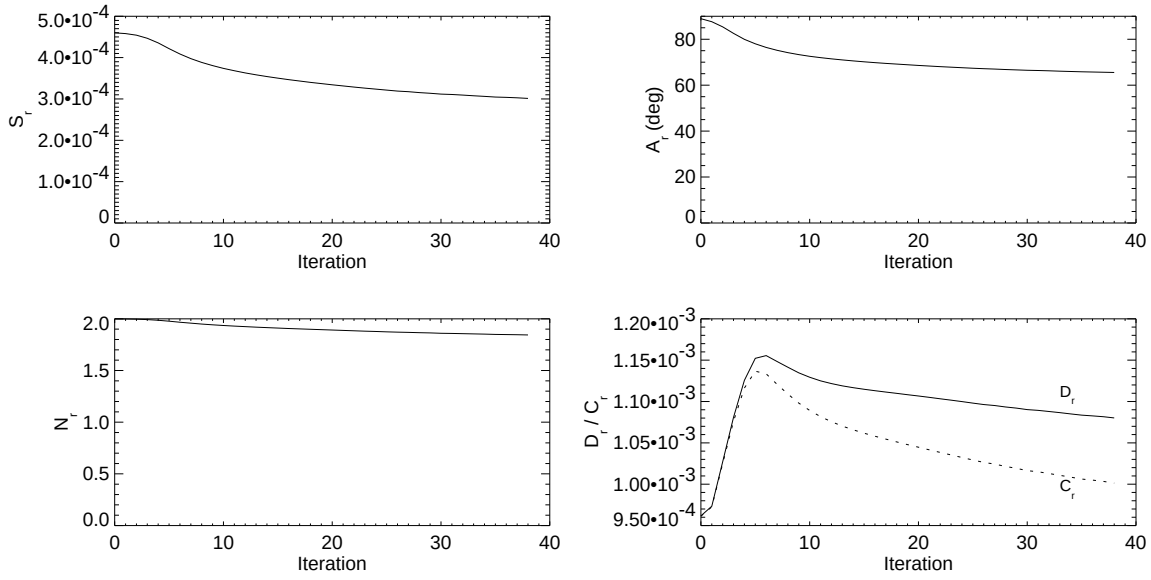

Fig. A.4: *Distance measures for the reconstruction of 20 point sources of flux* $3\,10^{-5}$ *ph* $cm^{-2}s^{-1}$.

of the 1.8 MeV emission. Optimum iterations which provide the closets profile recovery may be determined by application of the distance measures to the profiles themselves. The results for the longitude profiles of the exponential disk mock data of Fig. 3.2 are shown in Fig. A.5. All measures provide a clear minimum after only 2 ± 1 iterations. The jagged shape of S_r, D_r and C_r for the first iterations again reflect oscillations in the total reconstructed flux. After this early minimum the distances rise with proceeding iterations, becoming even worse than the initial flat profile for $r \gtrsim 25$. This demonstrates that overfit sets in very early in longitude profiles. It turned out that the distance minimum at iteration ~ 3 is universal for exponential disk simulations with total galactic ^{26}Al mass in the range $1 - 5$ $M_\odot$.

The distance measures for the corresponding latitude profile are shown in Fig. A.6. While A_r monotonically decreases with proceeding iterations the other measures indicate a weak minimum between iterations $10 - 30$, pointing towards the occurenence of spurious high latitude emission in later iterations. For exponential disk mock data with less galactic ^{26}Al mass, the minimum becomes more pronounced, confirming the onset of high latitude overfit around iteration 10. Nevertheless, Fig. 3.2 shows that the reconstructed profiles remain slightly broader than the narrow exponential disk profile. Consequently, the closets profile is obtained for the ML solution although spurious high latitude sources perturb the reconstruction. Taking together the results for the longitude and latitude profiles, the pixel-to-pixel distance as shown in Fig. A.3 can be understood as compromise between the early optimum longitude and the late optimum latitude iteration.

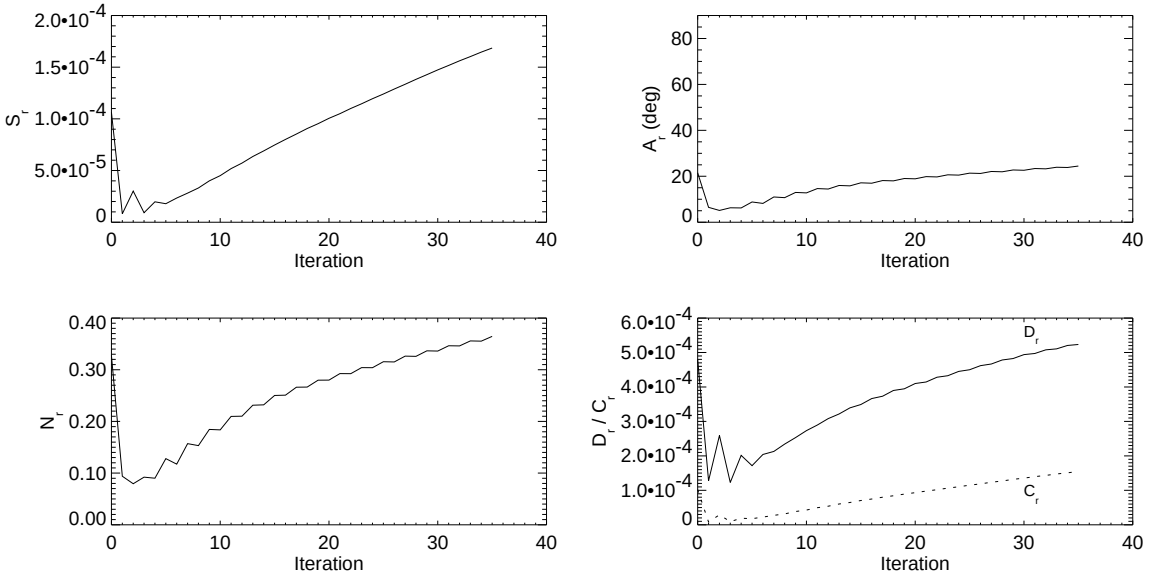

Fig. A.5: *Distance measures for the recovery of the longitude profile in the exponential disk simulation versus iteration r.*

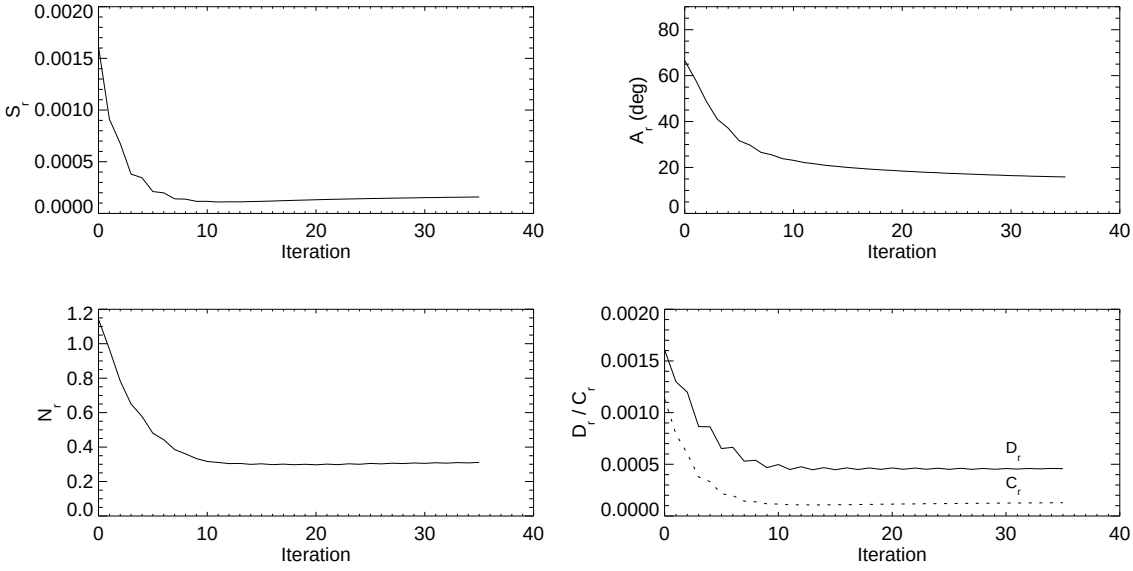

Fig. A.6: *Distance measures for the recovery of the latitude profile in the exponential disk simulation versus iteration r.*

A.3 Flux Calibration

In principle, fluxes Φ of emission features can be derived from the reconstructed intensity map f_j by integrating the intensities over the region of interest using

$$\Phi = \sum_{j \in \text{region}} f_j \Omega_j \tag{A.2}$$

where

$$\Omega_j = \frac{4\pi}{360} \Delta l \sin\left(\frac{\Delta b}{2}\right) \cos(b) \tag{A.3}$$

is the solid angle of image pixel f_j at galactic latitude b, and Δl and Δb are the pixel size in galactic longitude and latitude in degrees, respectively (throughout this work $\Delta l = \Delta b = 1°$). The reconstructed flux for 4 simulated point sources of flux $3\,10^{-5}$ ph cm^{-2}s^{-1} is shown in Fig. A.7 for circular integration regions centered on the point source positions. The plots show 'growth curves' of determined flux Φ as function of integration radius for different iterations. Since sky intensities are positive quantities these curves are strictly monotonic. As expected for an isotropic intensity distribution the flux for the initial estimates (lower solid curve) increase quadratically with integration radius. With proceeding iterations (dotted curves) the growth curves converge from the quadratic towards the ML growth curve (upper solid curve). For each example there exists a critical integration radius below which the flux monotonically increases towards the ML value. Above this radius, the flux 'overshoots' the ML growth curve by some small amount from which it decreases monotonically to the ML value. This demonstrates the concentration of the point source intensity within a distinct region with proceeding iterations: flux is 'diffusing' from outside the critical radius inwards, the flux at the critical radius remains constant.

Above the critical radius the growth curves increase due to background fluctuations in the intensity distribution. The different shape of the curves in the four examples can be explained by the statistical nature of the fluctuations which lead to different background intensity patterns around the sources. This is demonstrated in Fig. A.8 where the average growth curves are shown for the same sources for 10 independent simulations. Due to the averaging, the background intensity becomes rather isotropic, leading to a nearly quadratically increase of the growth curve above the critical radius. Remaining wiggles can be explained by the limited sample of only 10 independent simulations.

The statistical fluctuations largely supressed, it can been seen that the critical radius lies for all sources around $4° \pm 1°$. It is now obvious that the fluxes at the critical integration radius are systematically underestimated by about 20%. To eliminate the background contribution a mean background intensity is estimated from a ring around the point source position and subtracted from the observed fluxes. The resulting growth curves are shown in Fig. A.9. Now a clear convergent flux level becomes visible which justifies the derivation of point source fluxes by integration over circular areas of radius

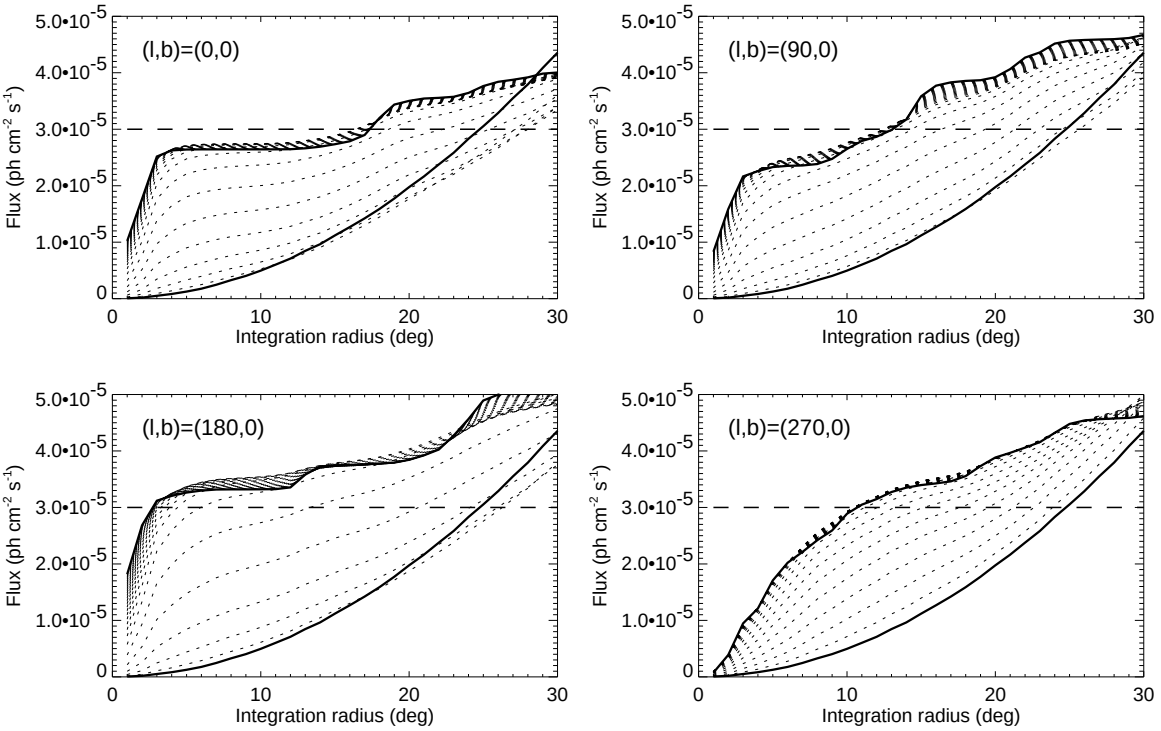

Fig. A.7: *Reconstructed point source flux for 4 simulated point sources of flux* $3\,10^{-5}$ *ph* $cm^{-2}s^{-1}$ *versus integration radius.*

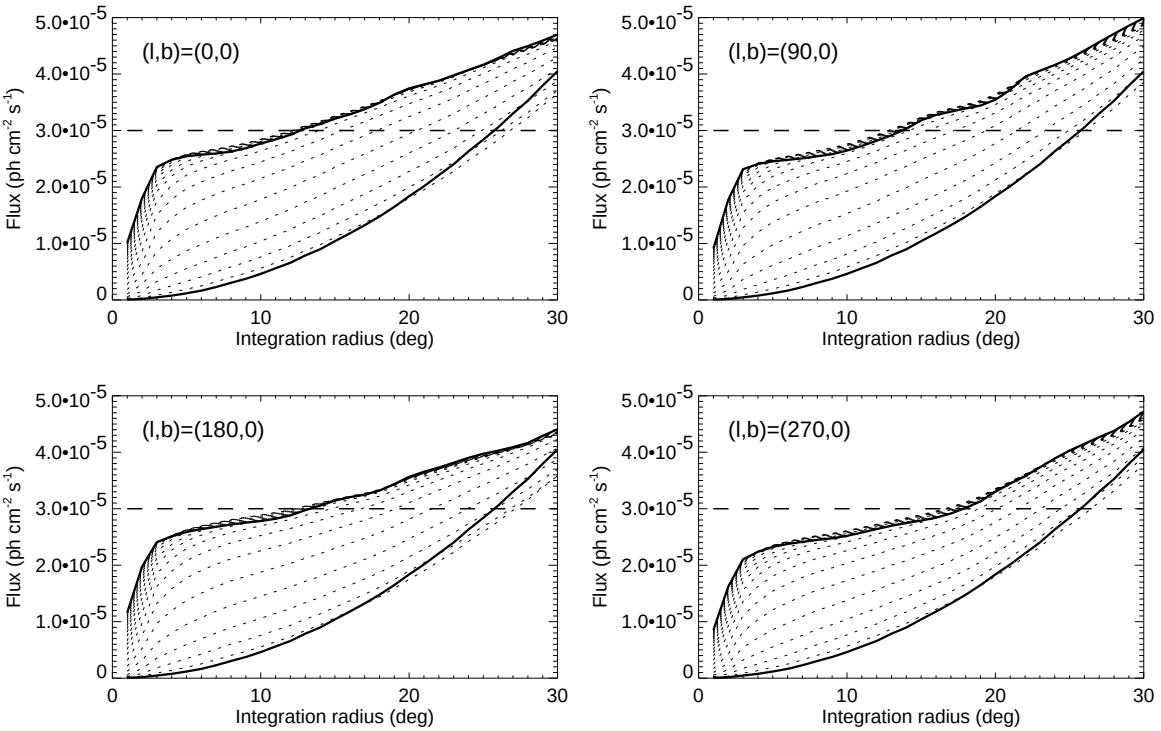

Fig. A.8: *Mean fluxes for 4 simulated point sources of flux* $3\,10^{-5}$ *ph* $cm^{-2}s^{-1}$ *from 10 independent simulations.*

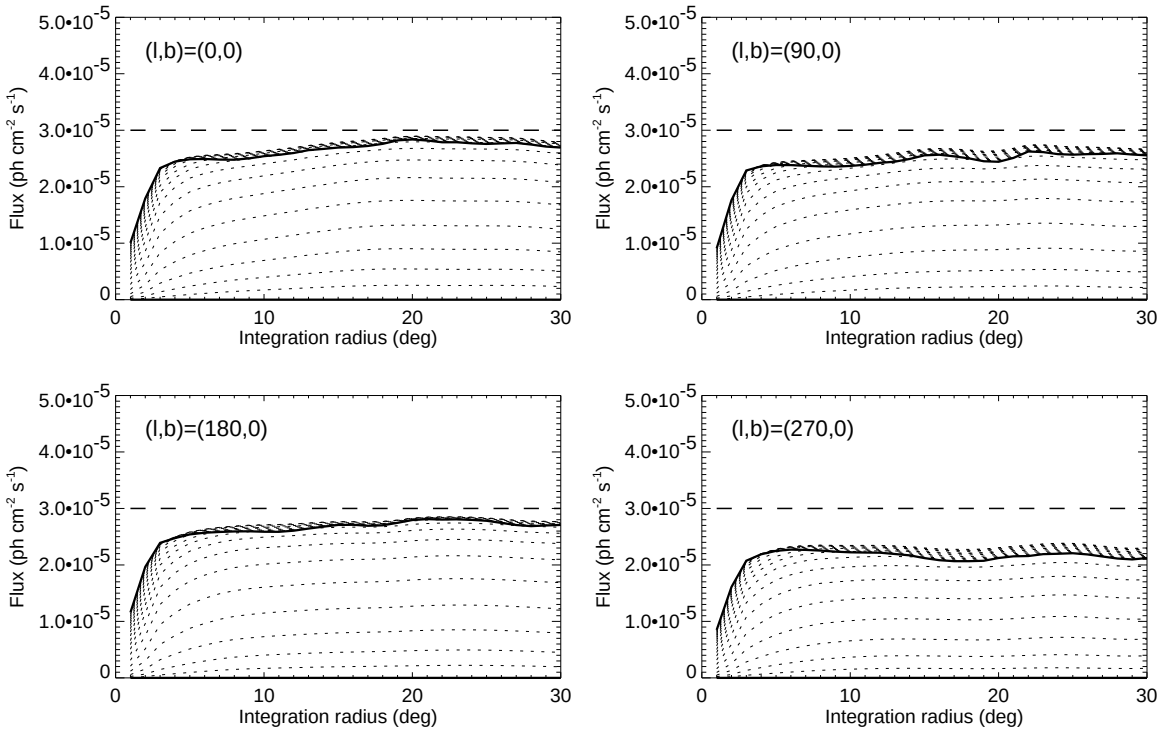

Fig. A.9: *Mean fluxes for 4 simulated point sources of flux* $3\,10^{-5}$ *ph cm*$^{-2}$*s*$^{-1}$ *from 10 independent simulations as derived by the contrast method. A background intensity was derived from a ring around the source position with inner and outer radii* 15° *and* 30°*, respectively.*

5°. From the number of dotted growth curves below the ML curve it can be deduced that only about $10 - 15$ iterations are required until the ML flux level is reached.

To derive meaningful point source fluxes from the R-L intensity maps, the flux underestimation has to be quantified. For this purpose, independent mock datasets of point source simulations with source fluxes between $1\,10^{-5}$ ph cm^{-2}s^{-1} and $5\,10^{-5}$ ph cm^{-2}s^{-1} are analyzed. The distribution of the point sources is always similar to that shown in Fig. 3.3. First, the dependency of reconstructed flux with exposure is studied. For this purpose the mean fluxes at integration radius 5° are plotted for different iterations as function of the sky exposure. As example, the average from 10 independent simulation of 20 point sources of flux $3\,10^{-5}$ ph cm^{-2}s^{-1} is shown in Fig. A.10. The error bars express the statistical uncertainty in the reconstructed fluxes, which is determined from the standard deviations of the samples. Linear regression curves are fitted to the data points, the significance of the presence of a dependency is given in Gaussian σ[1]. This example demonstrates that reconstructed point source

[1] Significance levels in astrophysics are often quoted in Gaussian σ which means that the probability of rejecting a truth hypothesis is identical to the area under the (normalized) Gaussian curve integrated from $-s\sigma$ to $+s\sigma$ (s is the significance in Gaussian σ). In that way significance levels of 1σ, 2σ, and

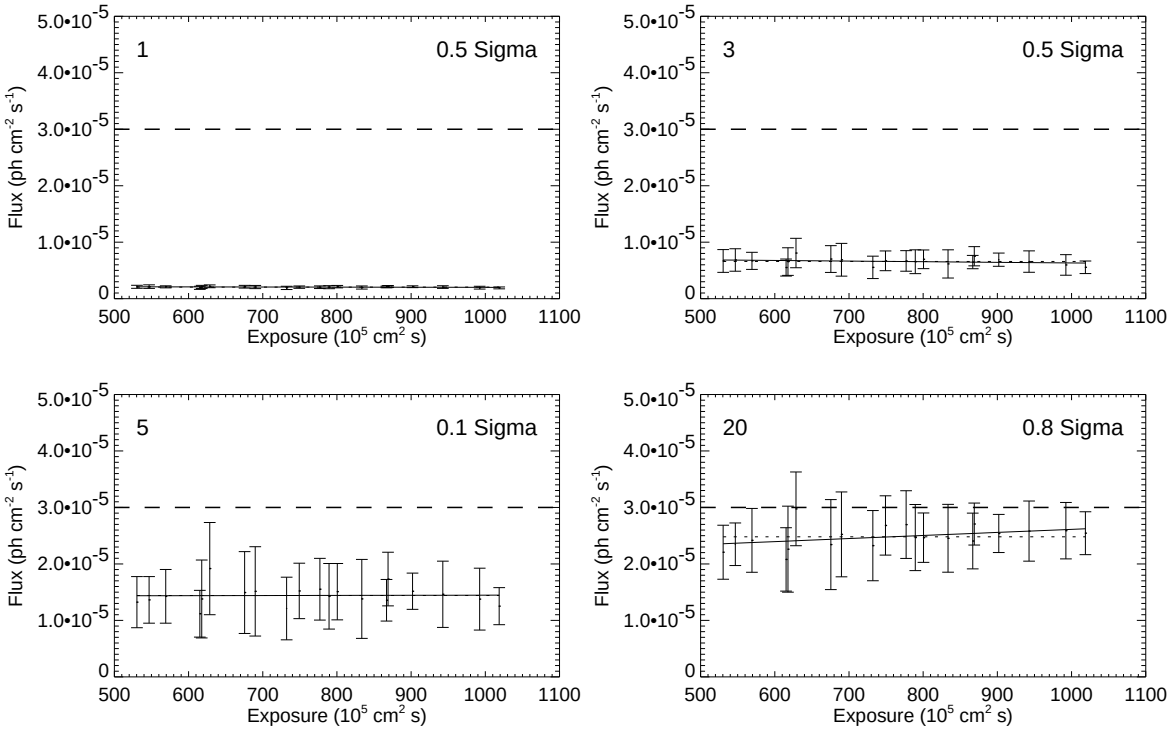

Fig. A.10: *Mean fluxes from 10 independent samples of simulated point sources (flux* $3\,10^{-5}$ *ph* $cm^{-2}s^{-1}$*) as function of sky exposure. Reconstructed fluxes are shown for iteration 1, 3, 5, and the ML image. Solid lines show the best fitting linear trend, dotted lines the best fitting constant.*

fluxes are rather independent of exposure for any R-L iteration.

Second, the dependency of reconstructed flux with galactic latitude is investigated. This is done by plotting the averaged fluxes for point sources at galactic latitude $|b| = 0°$, $30°$, and $60°$ as function of absolute galactic latitude. An example for the same data as above is shown in Fig. A.11 for different iterations. As before a constant and a linear dependency is fitted to the data points and the significance of the fit discrepancy is given in Gaussian σ. It turned out that there is no significant latitude dependence in the reconstruction of point source fluxes for any iteration.

The above investigations demonstrated that the point source flux underestimation is independent of source location. As third step the relationship between the true flux (i.e. the model flux) and the reconstructed flux is explored. For this purpose, the reconstructed fluxes from all point sources of all independent samples are averaged and plotted versus the model flux (Fig. A.12). The solid diagonal marks the line of equality between reconstructed and model flux. All data points lie below this line, demonstrating the systematic underestimation of the reconstructed flux. The func-

---

3σ correspond to rejection probabilities of 68.3%, 95.4%, and 99.7%, respectively.

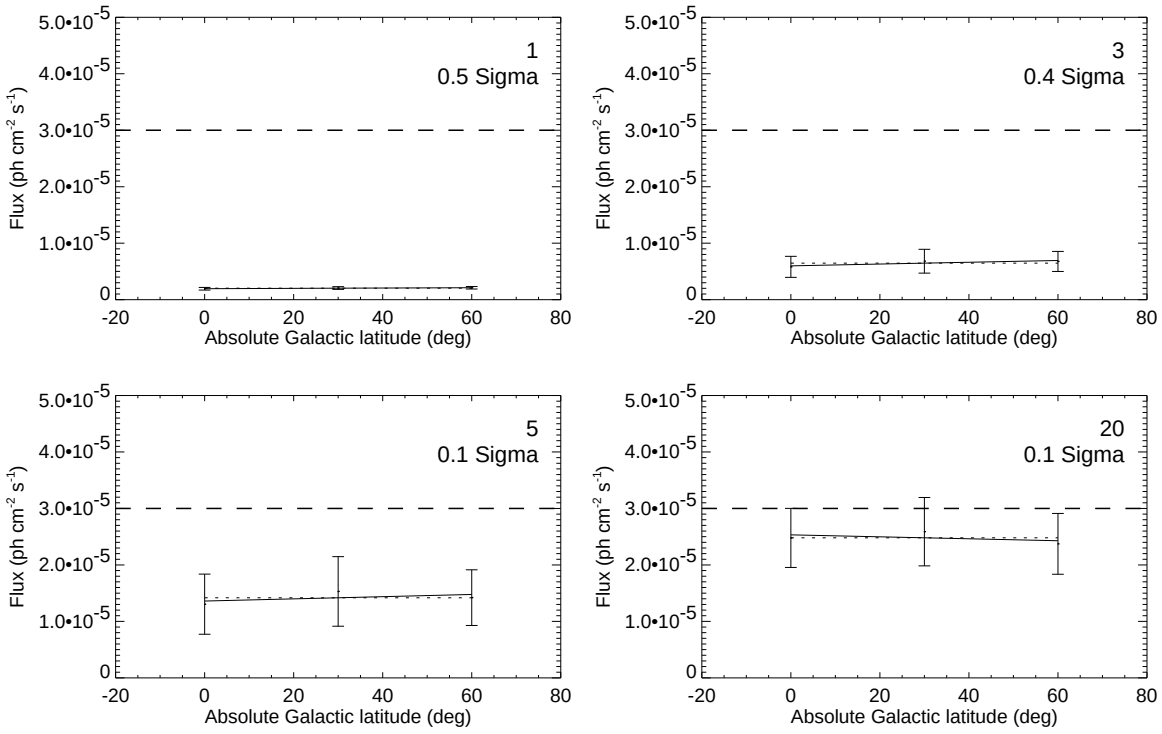

Fig. A.11: *Mean fluxes from 10 independent samples of simulated point sources (flux* $3\,10^{-5}$ *ph* $cm^{-2}s^{-1}$*) as function of absolute galactic latitude of the source position. Reconstructed fluxes are shown for iteration 1, 3, 5, and the ML image. Solid lines show the best fitting linear trend, dotted lines the best fitting constant.*

tional relationship between reconstructed point source flux Φ and model flux Φ_0 is well described by the linear relation

$$\Phi = a\Phi_0 + b \tag{A.4}$$

where $a = 1.01 \pm 0.16$ and $b = (-5.6 \pm 4.9)\,10^{-6}$ ph cm^{-2}s^{-1} were determined by a least square fit (the best fitting relation $\Phi = a'\Phi_0$ is shown as dotted curve where $a' = 0.86 \pm 0.07$). The reconstructed point source fluxes are thus consistent with a systematic underestimated of $(6 \pm 5)\,10^{-6}$ ph cm^{-2}s^{-1}, independent of model flux. It is important to note, however, that the above formula only holds for the explored flux range of $1\,10^{-5} - 5\,10^{-5}$ ph cm^{-2}s^{-1} (corresponding to reconstructed fluxes between roughly $5\,10^{-6} - 5\,10^{-5}$ ph cm^{-2}s^{-1}). Equation (A.4) always predicts positive fluxes even in the absence of reconstructed flux hence it cannot be valid for very low fluxes. Nevertheless, flux measurements around $1\,10^{-5}$ ph cm^{-2}s^{-1} are heavily limited by statistical fluctuations and are therefore not reliable at all. Under these limitations, point source fluxes can be determined from R-L reconstructions by integration of the intensities of late iterations ($r > 10$) within about 5° around the assumed point source position. The true source flux Φ_0 is then estimated from the measured flux Φ by $\Phi_0 = (\Phi - b)/a$.

Finally, flux recovery is investigated for diffuse emission. For this purpose, 10 independent mock datasets, created for an exponential disk model with radial scale length $R_0 = 4.5$ kpc, scale height $z_0 = 90$ pc, and galactic ^{26}Al mass of 3 $M_\odot$ are analyzed. Reconstructed intensities are determined within rectangular boxes, centered on galactic latitude $b = 0°$, for various box positions along galactic latitude $l = 0° \ldots 360°$. In Figs. A.13 and A.14, the flux recovery as function of RL iteration and model flux is illustrated. Two examples for different box size are given to study the case of small (Fig. A.13: $|b| < 5°$ and $\Delta l = 4°$) and large (Fig. A.14: $|b| < 20°$ and $\Delta l = 20°$) integration regions.

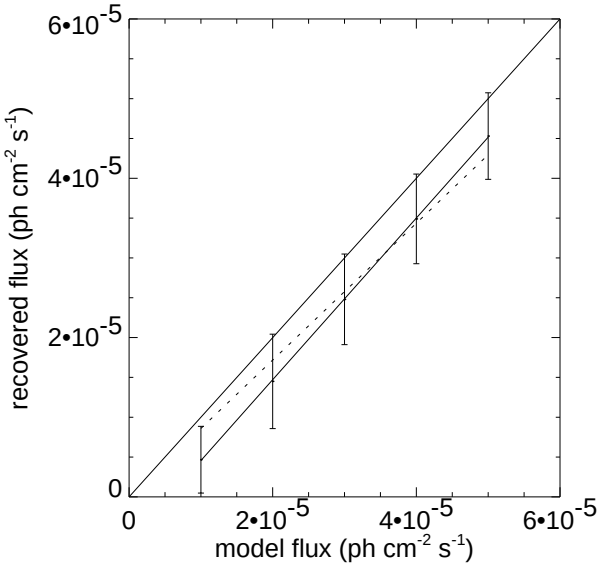

Fig. A.12: *Reconstructed point source flux as function of input flux.*

The left plots show the reconstructed intensity within three boxes of equal size, placed at longitudes with low, medium and high model intensity. Like for point sources, the reconstructed intensities converge to a constant level with proceeding iterations, which is generally reached earlier for high intensities and large integration regions (the oscillating behaviour, which was already noticed for the total flux, originates in the unconstrained RL acceleration (cf. Section 3.2.2.1)). In both examples, the convergent intensity levels are rather close to the model intensities (dashed lines). In particular, there seems to be no systematic trend towards low fluxes, as found in the point source simulations.

The intensity recovery as function of model intensity is studied by moving the integration boxes along galactic longitude in steps of 13°. To supress flux oscillations between successive iterations, convergent intensities are determined by averaging the reconstructed fluxes over the last two iterations. Reconstructed convergent intensities as function of model intensities are shown in the plots on the right of Figs. A.13 and A.14. Intensities found at earlier iterations are shown as dots; the diagonal of exact intensity recovery is indicated as dashed line. To determine the functional relationship between reconstructed and model intensity, a linear relation Eq. (A.4), which was already used for the point source flux calibration, is fitted to the data points. Resulting slope a and offset b with 1σ statistical errors are given in the upper left corner of the plots, the corresponding relation is shown as solid line. For the small integration regions (Fig. A.13), a marginal intensity underestimation is detected, while for large boxes (Fig. A.14) the recovery is perfect.

This trend is quantified in Fig. A.15 where the slope a and offset b of the linear

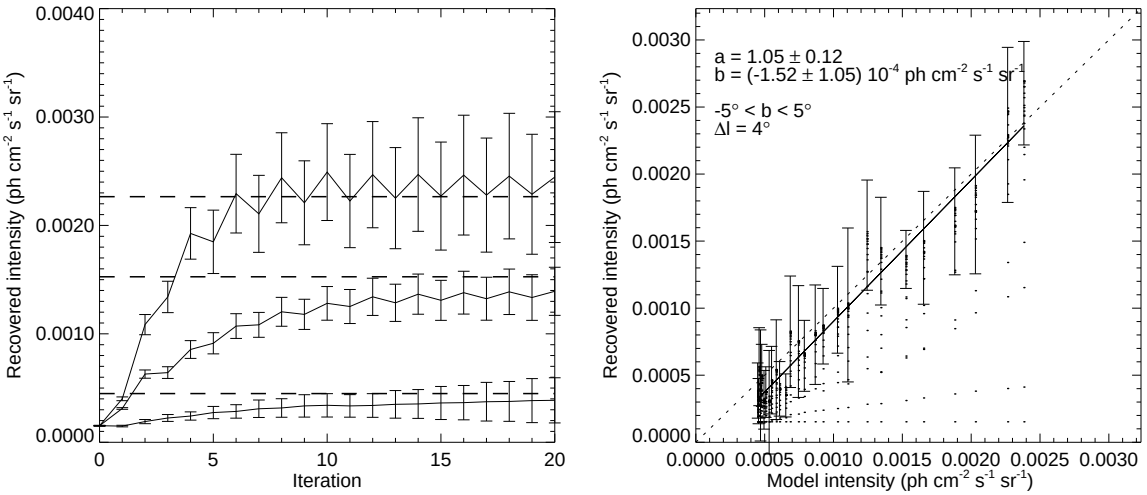

Fig. A.13: *Mean reconstructed intensity versus iteration (left), and versus model intensity (right) for 10 independent simulated exponential disk reconstructions. The intensities were derived for (small) rectangular boxes of size $\Delta b = 10°$ and $\Delta l = 4°$, centered on $b = 0°$. Error bars were determined from the standard deviation of the independent reconstructed intensities. On the left, reconstructed intensities for 3 different boxes placed at $l = 0°$, $80°$, and $180°$ are plotted versus iteration, dashed lines indicate the corresponding model intensities. On the right, reconstructed intensities are plotted versus model intensities for various integration boxes placed along the galactic plane. Error bars indicate the convergent intensities as derived by averaging over the last 2 iterations, dots show reconstructed intensities at earlier iterations. The solid line is a linear regression fit to the data points, the dashed line indicates the regime of exact intensity recovery.*

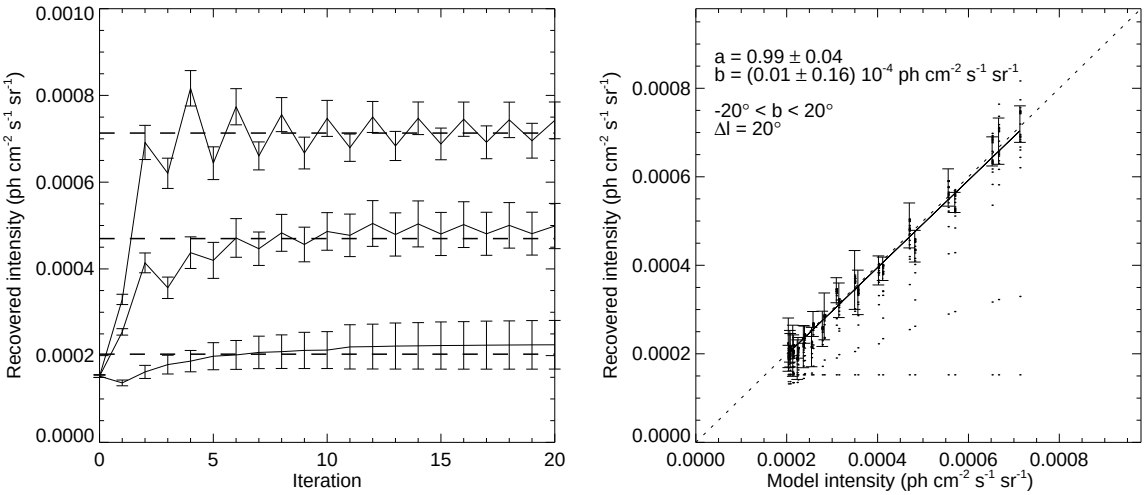

Fig. A.14: *Same as Fig. A.13, but for large integration boxes with $\Delta b = 40°$ and $\Delta l = 20°$.*

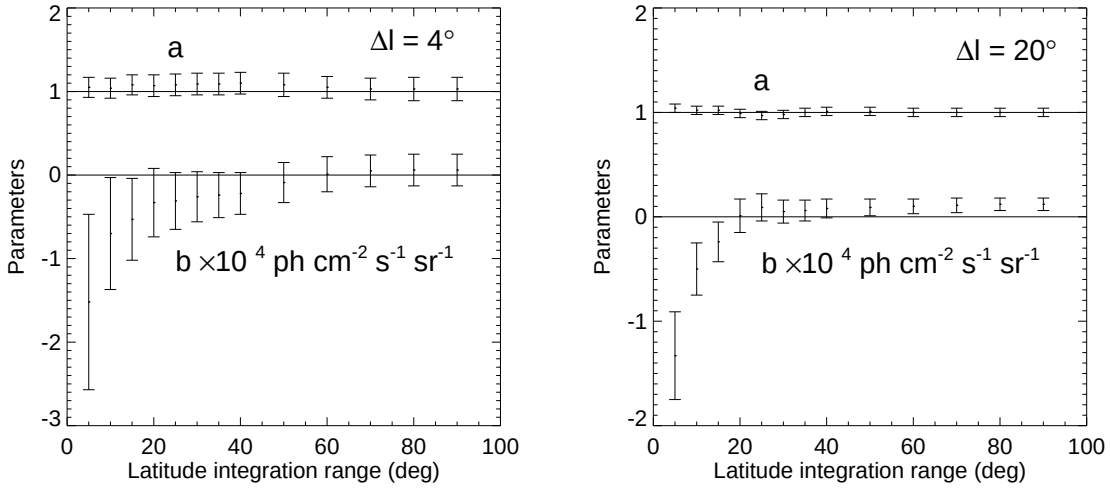

Fig. A.15: *Slope a and offset b of the linear relation between reconstructed and model intensity as function of integration box size.*

relation is shown as function of the latitude extension of the integration box for longitude sizes $\Delta l = 4°$ (left) and $\Delta l = 20°$ (right). Remarkably, the slopes a are consistent with unity over the entire range of integration regions, similar to the slope which was derived from the point source simulations. For small integration boxes (i.e. small latitude integration range) significantly negative offsets b are found in both examples, indicating flux underestimation. With increasing integration region the offsets cross the zero level, becoming slightly positive if integration is performed to high latitudes. This is consistent with the findings of Section 3.2.2.1 where a slight overestimation of the total flux was noticed.

Together with the point source calibration, the following picture arises: For small integration regions, fluxes are underestimated by a constant amount, indepedent of the source flux and source position. For the special case of point sources a linear correction relation was derived. With increasing size of the integration region, the underestimation becomes smaller than the statistical uncertainty. Reconstructed fluxes can even exceed the true value by a small amount if extended integration regions are chosen. This is due to the remaining background intensity in empty regions of the sky which is probably due to the positivity constraint of sky intensities. Subtracting mean intensities derived from high galactic latitudes corrects for this overestimation. For large integration regions this background intensity can considerably contribute to the emission.

Appendix B

The MEMSYS2 Algorithm

The software package generally used to reconstruct maximum entropy images from COMPTEL data is MEMSYS2 from Maximum Entropy Data Consultants Ltd of Cambridge, England, which was modified to treat Poisson statistics and to incorporate a background model. In this section the MEMSYS2 algorithm and the modifications are described.

MEMSYS2 maximizes the entropy S subject to $C \leq C_{aim}$ where C is a statistics measuring the misfit between the data and the image and C_{aim} is the upper bound to the values which C can plausibly take if the image fits the data.[1] As usual for such problems, a Lagrangian function $Q = S - \alpha C$ is set up, and the solution will lie at an extrema of Q for some value of the Lagrangian multiplier α. The non-linear nature of the problem forces an iterative solution.

In MEMSYS2 the extrema is repeatedly sought not along a single search direction, but in a three-dimensional subspace, spanned by vectors that are calculated anew at each landing point. The subspace basic vectors are chosen in such a way as to avoid directions leading to negative intensities f_j. The basis vectors of the three-dimensional subspace are given by

$$\begin{aligned} e_j^{(1)} &= f_j [\nabla C]_j \\ e_j^{(2)} &= f_j [\nabla S]_j \\ e_j^{(3)} &= \|\nabla S\|^{-1} f_j \sum_{k=1}^{M} [\nabla\nabla C]_{jk} f_k [\nabla S]_k - \|\nabla C\|^{-1} f_j \sum_{k=1}^{M} [\nabla\nabla C]_{jk} f_k [\nabla C]_k \end{aligned} \quad \text{(B.1)}$$

where lengths are defined by the entropy metric

$$\|x\| = \sqrt{\sum_{j=1}^{M} f_j x_j^2}. \quad \text{(B.2)}$$

[1] Skilling (1990) calls this method the 'historic' maximum entropy reconstruction.

Within the three-dimensional subspace a quadratic model

$$\tilde{Q} = Q_0 + \sum_{\mu=1}^{3} Q_\mu x_\mu + \frac{1}{2}\sum_{\nu=1}^{3}\sum_{\mu=1}^{3} H_{\mu\nu} x_\mu x_\nu \tag{B.3}$$

is constructed with components

$$\begin{aligned} Q_\mu &= \sum_{j=1}^{M} e_j^{(\mu)} [\nabla Q]_j \\ H_{\mu\nu} &= \sum_{j=1}^{M}\sum_{k=1}^{M} e_j^{(\mu)} e_j^{(\nu)} [\nabla\nabla Q]_{jk} \end{aligned} \tag{B.4}$$

chosen to agree with the local gradient and curvature components of Q itself. Then within the subspace, $\tilde{Q}$ is maximized at

$$x_\mu = -\sum_{\nu=1}^{3} (H_{\mu\nu})^{-1} Q_\nu \tag{B.5}$$

The image is then iteratively updated using

$$f_j^{r+1} = f_j^r + \sum_{\mu=1}^{3} x_\mu e_j^{(\mu)}. \tag{B.6}$$

In MEMSYS2 the Chi-square statistics

$$C = \frac{1}{2}\sum_{i=1}^{N} (F_i - D_i)^2 E_i \tag{B.7}$$

is used to evaluate the model-data misfit. For the evaluation of the basis vectors of the three-dimensional subspace and minimization in the subspace only derivates of C and S are required. Therefore the MEMSYS2 algorithm can be modified for Poisson statistics if the derivates of C are identical to the derivates of the Poisson log-likelihood function

$$\mathcal{L} = -2\sum_{i=1}^{N} n_i \ln e_i - e_i - n_i! \tag{B.8}$$

where

$$e_i = \sum_{j=1}^{M} f_j R_{ij} + b_i. \tag{B.9}$$

It can be shown that this is almost the case for the following assignment:

$$\begin{aligned} F_i &= e_i \\ D_i &= \begin{cases} e_i - \dfrac{e_i\,(e_i - n_i)}{n_i} & \text{for} \quad n_i > 0 \\ e_i - 2 & \text{for} \quad n_i = 0 \end{cases} \\ E_i &= \begin{cases} 2n_i/e_i^2 & \text{for} \quad n_i > 0 \\ 1 & \text{for} \quad n_i = 0 \end{cases} \end{aligned} \tag{B.10}$$

The gradient and the Hessian matrix of C are then given by

$$\begin{aligned}
[\nabla C]_j = \sum_{i=1}^{N}(F_i - D_i)E_i R_{ij} &= 2\sum_{n_i\neq 0}\frac{e_i - n_i}{e_i}R_{ij} + 2\sum_{n_i=0} R_{ij} \\
[\nabla\nabla C]_{jk} = \sum_{i=1}^{N} E_i R_{ij} R_{ik} &= 2\sum_{n_i\neq 0}\frac{n_i}{e_i^2}R_{ij}R_{ik} + 2\sum_{n_i=0} R_{ij}R_{ik}
\end{aligned} \qquad \text{(B.11)}$$

which compares to the gradient and Hessian matrix of the Poisson log-likelihood $\mathcal{L}$

$$\begin{aligned}
[\nabla\mathcal{L}]_j &= 2\sum_{i=1}^{N}\frac{e_i - n_i}{e_i}R_{ij} \\
[\nabla\nabla\mathcal{L}]_{jk} &= 2\sum_{i=1}^{N}\frac{n_i}{e_i^2}R_{ij}R_{ik}
\end{aligned} \qquad \text{(B.12)}$$

Equation (B.11) is identical to Eq. (B.12) except for $n_i = 0$ where $[\nabla\nabla C]_{jk}$ has the additional term $2\sum_{n_i=0} R_{ij}R_{ik}$. Since R_{ij} is positive, $[\nabla\nabla C]_{jk} \geq [\nabla\nabla\mathcal{L}]_{jk}$ with the equality reached if no empty data pixels $n_i = 0$ occur. For the all-sky data of observation periods 0.1 - 522.5 about 35% of the data space pixels are empty.

Due to the assignment Eq. (B.10) C is no longer identical to the Chi-square statistics and the constraint $C \leq C_{aim}$ looses its' meaning as stopping criterion. For Poisson data, there exists no analogue constraint since the formal expectation over possible data of the log-likelihood is not constant, but varies with the image. Hence the historical stopping criterion is not applicable for COMPTEL data. The optimum image is merely determined by means of simulations.

Appendix C

Background Scaling Factor Prior Assignment

For the fixed $\bar{\varphi}$ method, the instrumental background model is scaled by a global background scaling factor β which accounts for the uncertainty about the total number of 1.8 MeV events in the COMPTEL all-sky data set. The background model, which is based on events observed at adjacent energy bands, was carefully normalized on single observation basis under consideration of time variabilities of major background components (Oberlack 1997, see also Section 2.5.2.1). This normalization makes β a rather well defined parameter, where the major uncertainties arise from unidentified time variable background components 0.5% (Oberlack, private communication)), and the zero level estimation from high galactic latitudes $|b| > 40°$ (relative error $\sim 1\%$ if a signal similar to that below $|b| < 40°$ is present at high latitudes). To quantify these uncertainties, a Gaussian prior

$$P(\tilde{\beta}|M,I) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\frac{(\tilde{\beta}-\hat{\beta})^2}{\sigma^2}\right) \tag{C.1}$$

with mean $\beta = 1$ and standard deviation $\sigma = 5\,10^{-3}$ is used for the background scaling factor (the statistical error on β is $< 1\,10^{-3}$).

For $\bar{\varphi}$ fitting, the global background scaling factor β is replaced by individual scaling factors $\beta_{\bar{\varphi}}$ for each $\bar{\varphi}$ layer. The prior knowledge about the magnitude of these scaling factors is twofold. First, the total number of counts in the background model is known to some accuracy. This leads to an integral property of $\beta_{\bar{\varphi}}$ which can be expressed by

$$\tilde{\beta} = \frac{\sum\limits_{\bar{\varphi}} \beta_{\bar{\varphi}} B_{\bar{\varphi}}}{\sum\limits_{\bar{\varphi}} B_{\bar{\varphi}}} \tag{C.2}$$

where $\tilde{\beta}$ is a mean background scaling factor which has the same meaning as the single global background scaling factor β for fixed $\bar{\varphi}$ fitting ($B_{\bar{\varphi}}$ is the number of counts in layer

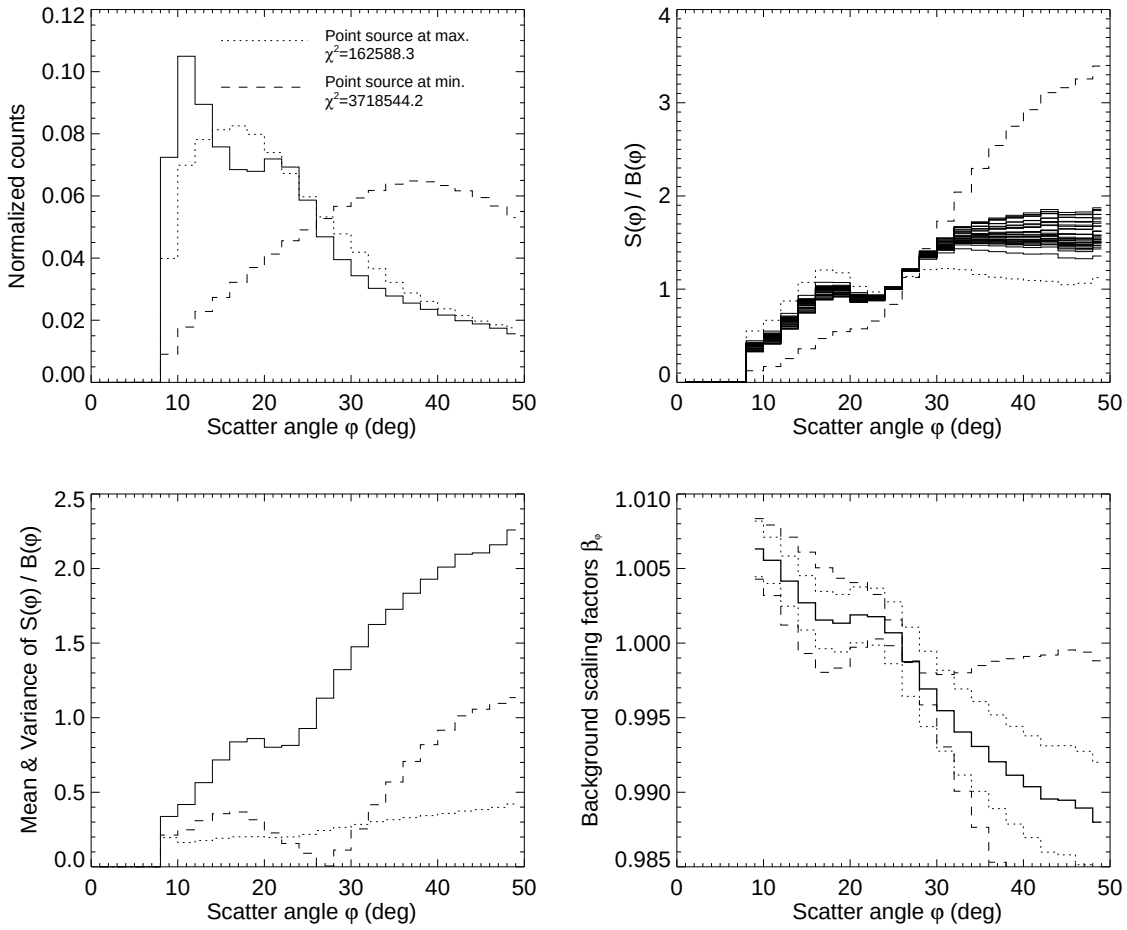

Fig. C.1: Upper-left: *Normalized $\bar{\varphi}$ distributions for the background model ($\hat{B}_{\bar{\varphi}}$, solid) and two extreme source models ($\hat{S}_{\bar{\varphi}}$, dotted).* Upper-right: *Ratio $\hat{S}_{\bar{\varphi}}/\hat{B}_{\bar{\varphi}}$ for various source models, including all tracer maps, extreme point source models, and ^{26}Al distribution models.* Lower-left: *Mean (solid) and variance (dashed) of $\hat{S}_{\bar{\varphi}}/\hat{B}_{\bar{\varphi}}$. The dotted histogram shows the uncertainty about $\hat{S}_{\bar{\varphi}}/\hat{B}_{\bar{\varphi}}$ due to statistical fluctuations.* Lower-right: *Individual background scaling factors $\beta_{\bar{\varphi}}$ derived for the mean $\hat{S}_{\bar{\varphi}}/\hat{B}_{\bar{\varphi}}$ ratio and mean background scaling factor $\beta = 1$ (solid). The dashed histrograms show the expected variations due to different source models, the dotted histograms represent the variations due to statistical uncertainties.*

$\bar{\varphi}$ of the background model). Consequently, the same prior probability distribution will be assigned to this quantity.

Second, the distribution of individual background scaling factors $\beta_{\bar{\varphi}}$ depends on the magnitude of the source signal and the discrepancy between the $\bar{\varphi}$ distribution of the source signal and the background model. To investigate these dependencies the basic relation

$$N_{\bar{\varphi}} = S_{\bar{\varphi}} + \beta_{\bar{\varphi}} B_{\bar{\varphi}} \tag{C.3}$$

is considered which connects the $\bar{\varphi}$ distribution $N_{\bar{\varphi}}$ of the observed events with those

of the source signal ($S_{\bar{\varphi}}$) and the background model ($B_{\bar{\varphi}}$). Due to the $\bar{\varphi}$ normalization of the background model on the data, $N_{\bar{\varphi}}$ and $B_{\bar{\varphi}}$ obey the same $\bar{\varphi}$ distribution, hence

$$N_{\bar{\varphi}} = f B_{\bar{\varphi}} \tag{C.4}$$

with $f = 1.00954$ for the 1.8 MeV COMPTEL all-sky data set. Putting Eq. (C.4) in Eq. (C.3) and summing over all $\bar{\varphi}$ layers provides a relation between $\tilde{\beta}$ and the total number of source counts:

$$\sum_{\bar{\varphi}} S_{\bar{\varphi}} = (f - \tilde{\beta}) \sum_{\bar{\varphi}} B_{\bar{\varphi}}. \tag{C.5}$$

On the other hand, using Eqs. (C.3) and (C.4), $\beta_{\bar{\varphi}}$ can be expressed by

$$\beta_{\bar{\varphi}} = f - \frac{S_{\bar{\varphi}}}{B_{\bar{\varphi}}} \tag{C.6}$$

which may be rewritten as

$$\beta_{\bar{\varphi}} = f + (\tilde{\beta} - f)\tilde{\beta}_{\bar{\varphi}} \tag{C.7}$$

where

$$\tilde{\beta}_{\bar{\varphi}} \equiv \frac{\hat{S}_{\bar{\varphi}}}{\hat{B}_{\bar{\varphi}}} = \frac{S_{\bar{\varphi}} / \sum_{\bar{\varphi}'} S_{\bar{\varphi}'}}{B_{\bar{\varphi}} / \sum_{\bar{\varphi}'} B_{\bar{\varphi}'}} \tag{C.8}$$

is the ratio between the normalized $\bar{\varphi}$ distributions of the source signal and the background model.

With Eq. (C.7), the individual background scaling factors $\beta_{\bar{\varphi}}$ are now separated into a source strength dependent mean background scaling factor $\tilde{\beta}$, and source strength independent parameters $\tilde{\beta}_{\bar{\varphi}}$ which quantify the difference in $\bar{\varphi}$ shape between the source and the background models. Prior knowledge about $\tilde{\beta}_{\bar{\varphi}}$ can be obtained from a comparison of the $\bar{\varphi}$ distribution of the background model with that of possible source models. For this purpose, $\bar{\varphi}$ distributions of various models were compared to that of the background model. Two most extreme examples which were found are shown in the upper-left panel of Fig. C.1: As anticipated in Section 4.2.2, the steepest $\bar{\varphi}$ distribution was found for a point source situated at $(l, b) = (5°, -2°)$, i.e. at a maximum of the 'attraction map' (Fig. 4.4). Correspondingly, the 'flattest' $\bar{\varphi}$ distribution is obtained for a point source at $(l, b) = (230°, 30°)$ where the 'attraction map' obeys a minimum. The resulting $\tilde{\beta}_{\bar{\varphi}}$ distributions are shown with those for various models in the upper-right panel of Fig. C.1.

With Eqs. (C.2) and (C.7) the individual background scaling factors $\beta_{\bar{\varphi}}$ are mapped into a manifold spanned by the parameters $\{\tilde{\beta}, \tilde{\beta}_{\bar{\varphi}}\}$ in which the prior information becomes mutually independent (note that for the original scaling factors $\beta_{\bar{\varphi}}$ the prior information is not mutually independent). Due to this independence, the background prior can be separated into

$$P(\tilde{\beta}, \tilde{\beta}_{\bar{\varphi}}|M, I) = \alpha P(\tilde{\beta}|M, I) \prod_{\bar{\varphi}} P(\tilde{\beta}_{\bar{\varphi}}|M, I) \tag{C.9}$$

where α is a yet undetermined normalization constant. The prior knowledge about $\{\tilde{\beta}, \tilde{\beta}_{\bar{\varphi}}\}$ is in the form of expected values with some variance, hence Gaussians are appropriate to quantify the uncertainties of the parameters: For the mean background factor we write

$$P(\tilde{\beta}|M,I) = \exp\left(-\frac{1}{2}\frac{(\tilde{\beta}-\hat{\beta})^2}{\sigma^2}\right) \tag{C.10}$$

with $\hat{\beta} = 1$ and $\sigma = 5\,10^{-3}$, similar to Eq. (C.1). The prior knowledge about the possible $\bar{\varphi}$ distribution is quantified for each $\tilde{\beta}_{\bar{\varphi}}$ by a Gaussian prior

$$P(\tilde{\beta}_{\bar{\varphi}}|M,I) = \exp\left(-\frac{1}{2}\frac{(\tilde{\beta}_{\bar{\varphi}}-\hat{\beta}_{\bar{\varphi}})^2}{\sigma_{\bar{\varphi}}^2}\right) \tag{C.11}$$

where the means $\hat{\beta}_{\bar{\varphi}}$ were derived from the average and the standard deviations $\sigma_{\bar{\varphi}}$ from the discrepancy of $\tilde{\beta}_{\bar{\varphi}}$ for the most extreme examples. Only for $\bar{\varphi}$ layers where the $\tilde{\beta}_{\bar{\varphi}}$ discrepancy falls below the statistical uncertainty, $\sigma_{\bar{\varphi}}$ was replaced by this uncertainty.

The normalization constant α is determined by the normalization constraint of the prior probability distribution

$$\int_{-\infty}^{\infty} P(\beta_{\bar{\varphi}}|M,I)d\beta_{\bar{\varphi}} = 1. \tag{C.12}$$

Since the prior is defined on a manifold with one additional dimension, the integration has to be done on the hypersurface $dS(\tilde{\beta}, \tilde{\beta}_{\bar{\varphi}})$ which is accessible by variation of the individual background scaling factors $\beta_{\bar{\varphi}}$. The normalization constraint writes then

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} P(\beta_{\bar{\varphi}}|M,I)d\beta_{\bar{\varphi}} = \\ &= \int_S P(\tilde{\beta}, \tilde{\beta}_{\bar{\varphi}}|M,I)dS(\tilde{\beta}, \tilde{\beta}_{\bar{\varphi}}) = \\ &= \int_{-\infty}^{\infty} P(\Phi(\beta_{\bar{\varphi}})|M,I)\sqrt{g(\beta_{\bar{\varphi}})}d\beta_{\bar{\varphi}} \end{aligned} \tag{C.13}$$

where $\Phi(\beta_{\bar{\varphi}})$ is the mapping defined by Eqs. (C.2) and (C.7) from the background scaling factor space into the prior manifold, and $g(\beta_{\bar{\varphi}}) \equiv \det(g_{\bar{\varphi}'\bar{\varphi}''})$ is the Gram determinant of the metrik tensor

$$g_{\bar{\varphi}'\bar{\varphi}''} = \frac{\partial\tilde{\beta}}{\partial\beta_{\bar{\varphi}'}}\frac{\partial\tilde{\beta}}{\partial\beta_{\bar{\varphi}''}} + \sum_{\bar{\varphi}}\frac{\partial\tilde{\beta}_{\bar{\varphi}}}{\partial\beta_{\bar{\varphi}'}}\frac{\partial\tilde{\beta}_{\bar{\varphi}}}{\partial\beta_{\bar{\varphi}''}} \tag{C.14}$$

which is associated to the mapping. For completness, the functional matrix of the mapping is given by

$$\frac{\partial(\tilde{\beta}, \tilde{\beta}_1, \cdots, \tilde{\beta}_n)}{\partial(\beta_1, \cdots, \beta_n)} = \begin{pmatrix} \frac{B_1}{B} & \frac{B_2}{B} & \cdots & \frac{B_n}{B} \\ \frac{1}{\tilde{\beta}-f} - \frac{\beta_1-f}{(\tilde{\beta}-f)^2}\frac{B_1}{B} & -\frac{\beta_1-f}{(\tilde{\beta}-f)^2}\frac{B_2}{B} & \cdots & -\frac{\beta_1-f}{(\tilde{\beta}-f)^2}\frac{B_n}{B} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{\beta_n-f}{(\tilde{\beta}-f)^2}\frac{B_1}{B} & -\frac{\beta_n-f}{(\tilde{\beta}-f)^2}\frac{B_2}{B} & \cdots & \frac{1}{\tilde{\beta}-f} - \frac{\beta_n-f}{(\tilde{\beta}-f)^2}\frac{B_n}{B} \end{pmatrix} \tag{C.15}$$

where n is the number of $\bar{\varphi}$ layers ($n = 21$ for COMPTEL 1.8 MeV analysis) and $B = \sum_{\bar{\varphi}} B_{\bar{\varphi}}$ is the total number of counts in the background model. Analytical integration over the hypersurface, however, is impossible since the mapping between the background scaling factors and the prior manifold is nonlinear. For this reason the integral is approximated using Laplace's method which is based on a Taylor series expansion of the function $y(x)$ of the $m-$dimensional vector x, and yields the approximation

$$\int \exp\{y(x)\}dx \approx (2\pi)^{m/2}\sqrt{\det(A)}\exp\{y(x^{max})\}, \tag{C.16}$$

where x^{max} is the value of x where y attains its maximum, and A is the negative of the inverse Hessian of y evaluated at x^{max}. The metrik tensor varies only slowly over the prior manifold hence it can be assumed to be constant for the high-prior region. The normalization constraint then becomes

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} P(\Phi(\beta_{\bar{\varphi}})|M,I)\sqrt{g(\beta_{\bar{\varphi}})}d\beta_{\bar{\varphi}} = \\ &= \int_{-\infty}^{\infty} \exp\left\{\ln\alpha - \frac{1}{2}\left[\frac{(\tilde{\beta}-\hat{\beta})^2}{\sigma^2} + \sum_{\bar{\varphi}}\frac{(\tilde{\beta}_{\bar{\varphi}}-\hat{\beta}_{\bar{\varphi}})^2}{\sigma_{\bar{\varphi}}^2}\right] + \ln\sqrt{g(\beta_{\bar{\varphi}})}\right\} d\beta_{\bar{\varphi}} \approx \\ &\approx \alpha(2\pi)^{(n+1)/2}\sqrt{\det(A)}\sqrt{g(\beta_{\bar{\varphi}}^{max})} \end{aligned} \tag{C.17}$$

where $\beta_{\bar{\varphi}}^{max} = f + (\hat{\beta} - f)\hat{\beta}_{\bar{\varphi}}$ are the background scaling factors which correspond to the maximum of the prior probability distribution. The normalization constant α is then given by

$$\alpha = \left((2\pi)^{(n+1)/2}\sqrt{\det(A)}\sqrt{g(\beta_{\bar{\varphi}}^{max})}\right)^{-1} \tag{C.18}$$

with

$$A^{-1}_{\bar{\varphi}'\bar{\varphi}''} = \frac{1}{\sigma^2}\frac{B_{\bar{\varphi}'}}{B}\frac{B_{\bar{\varphi}''}}{B} + \sum_{\bar{\varphi}}\frac{1}{\sigma_{\bar{\varphi}}^2}\left(\frac{\delta_{\bar{\varphi}\bar{\varphi}'}}{\hat{\beta}-f} - \frac{\hat{\beta}_{\bar{\varphi}}-f}{(\hat{\beta}-f)^2}\frac{B_{\bar{\varphi}'}}{B}\right)\left(\frac{\delta_{\bar{\varphi}\bar{\varphi}''}}{\hat{\beta}-f} - \frac{\hat{\beta}_{\bar{\varphi}}-f}{(\hat{\beta}-f)^2}\frac{B_{\bar{\varphi}''}}{B}\right). \tag{C.19}$$

Appendix D

Fit Quality

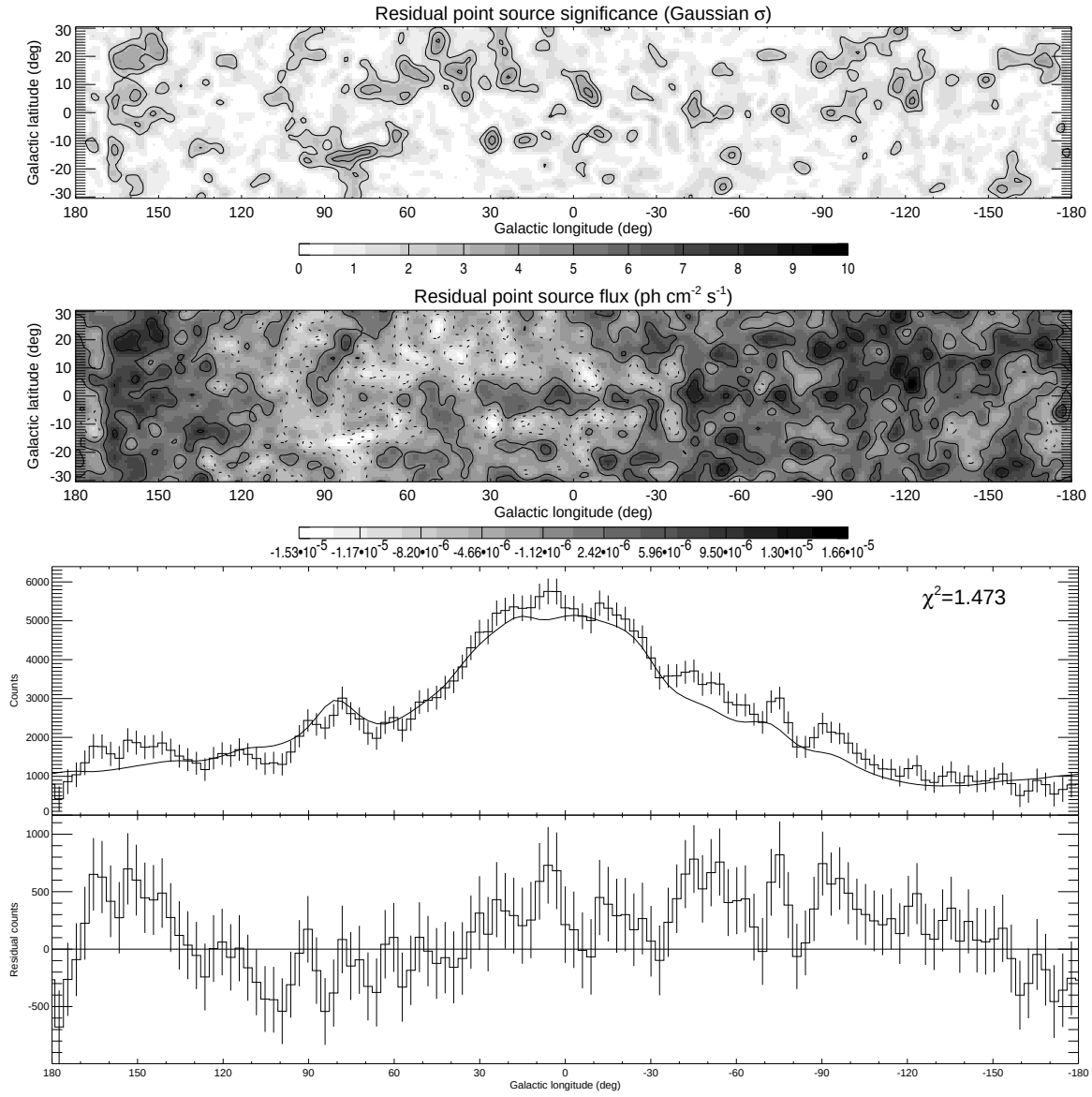

Fig. D.1: DMR 31.5 GHz model.

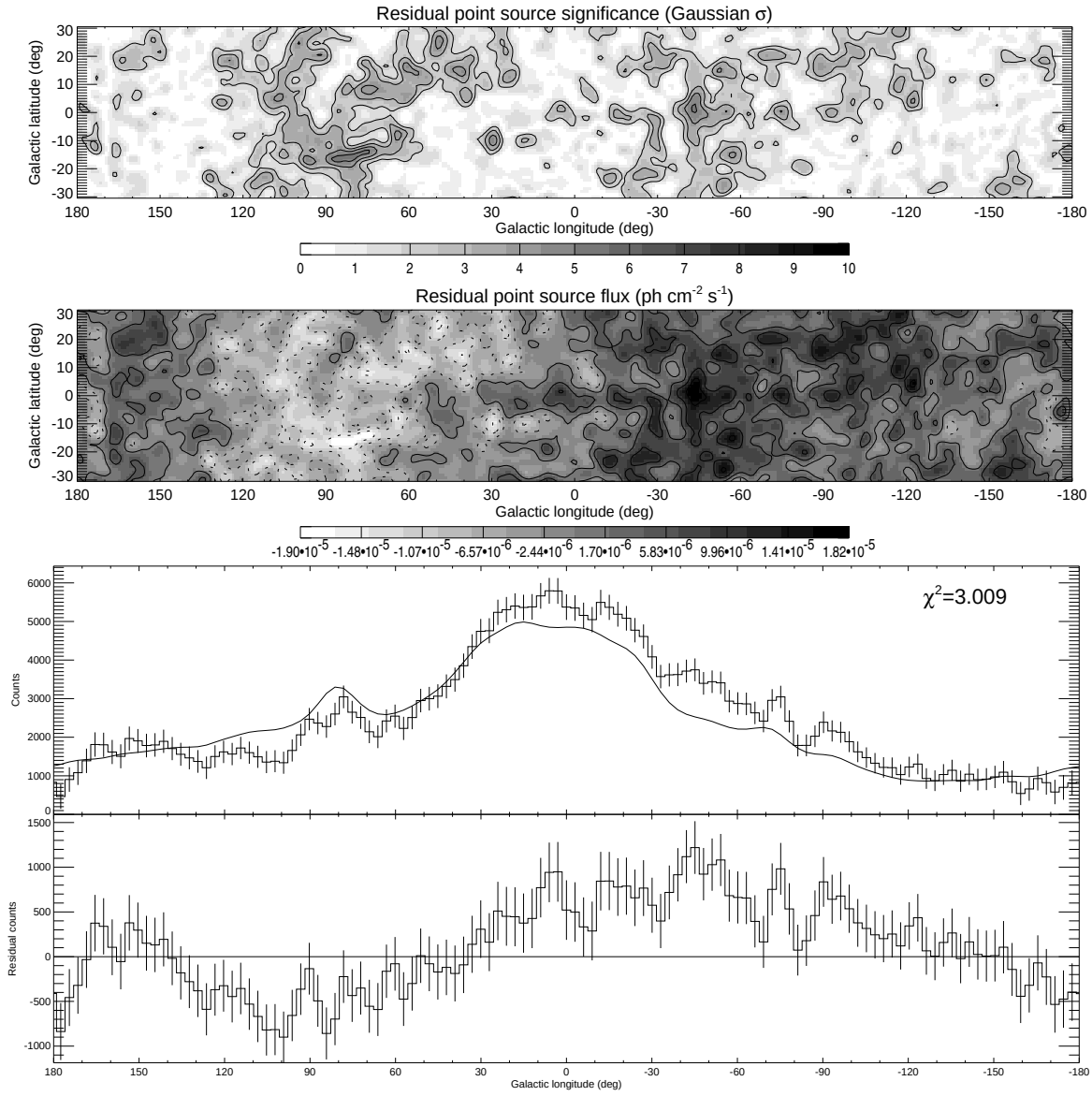

Fig. D.2: DMR 53 GHz model.

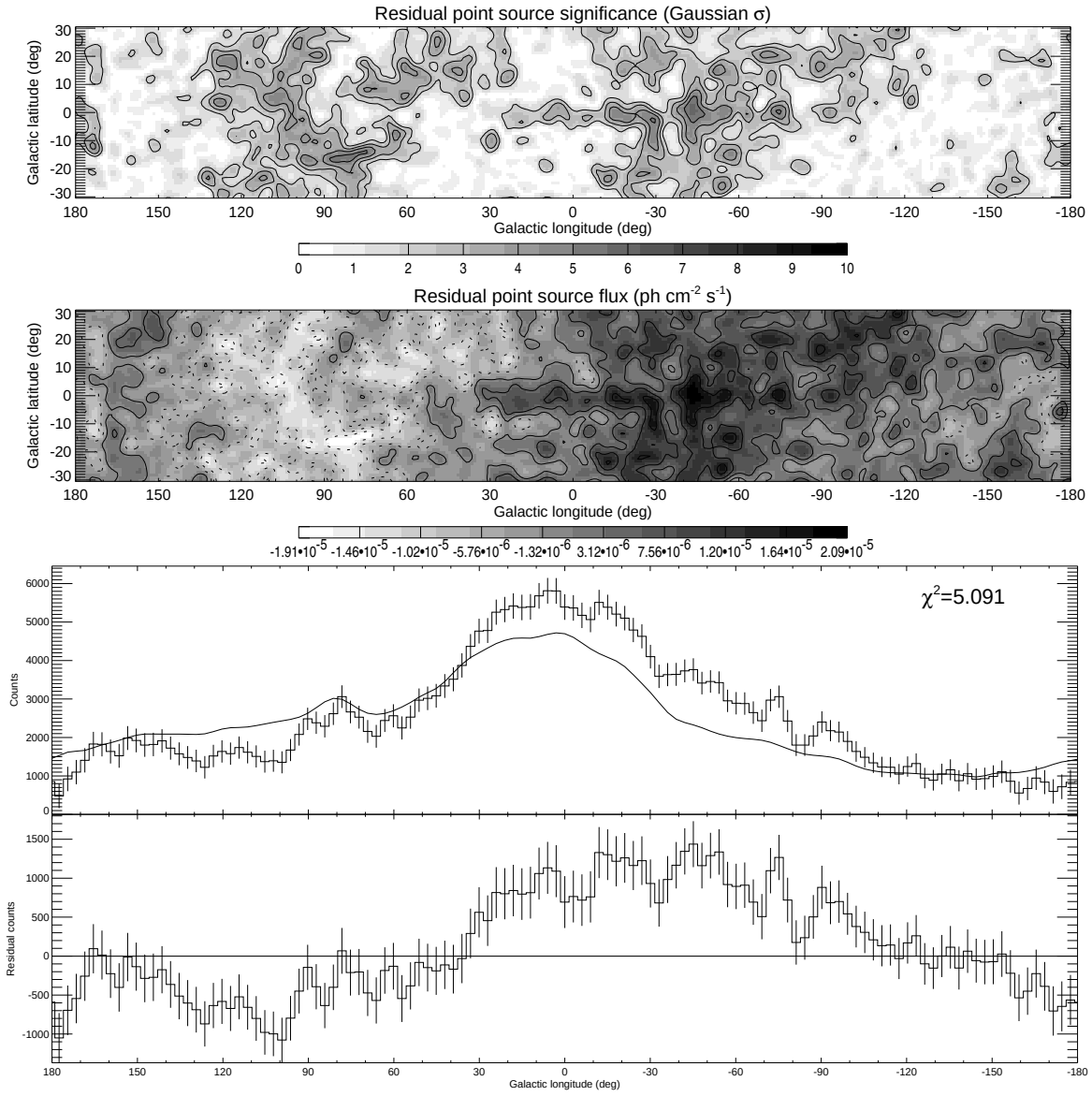

Fig. D.3: DMR 90 GHz model.

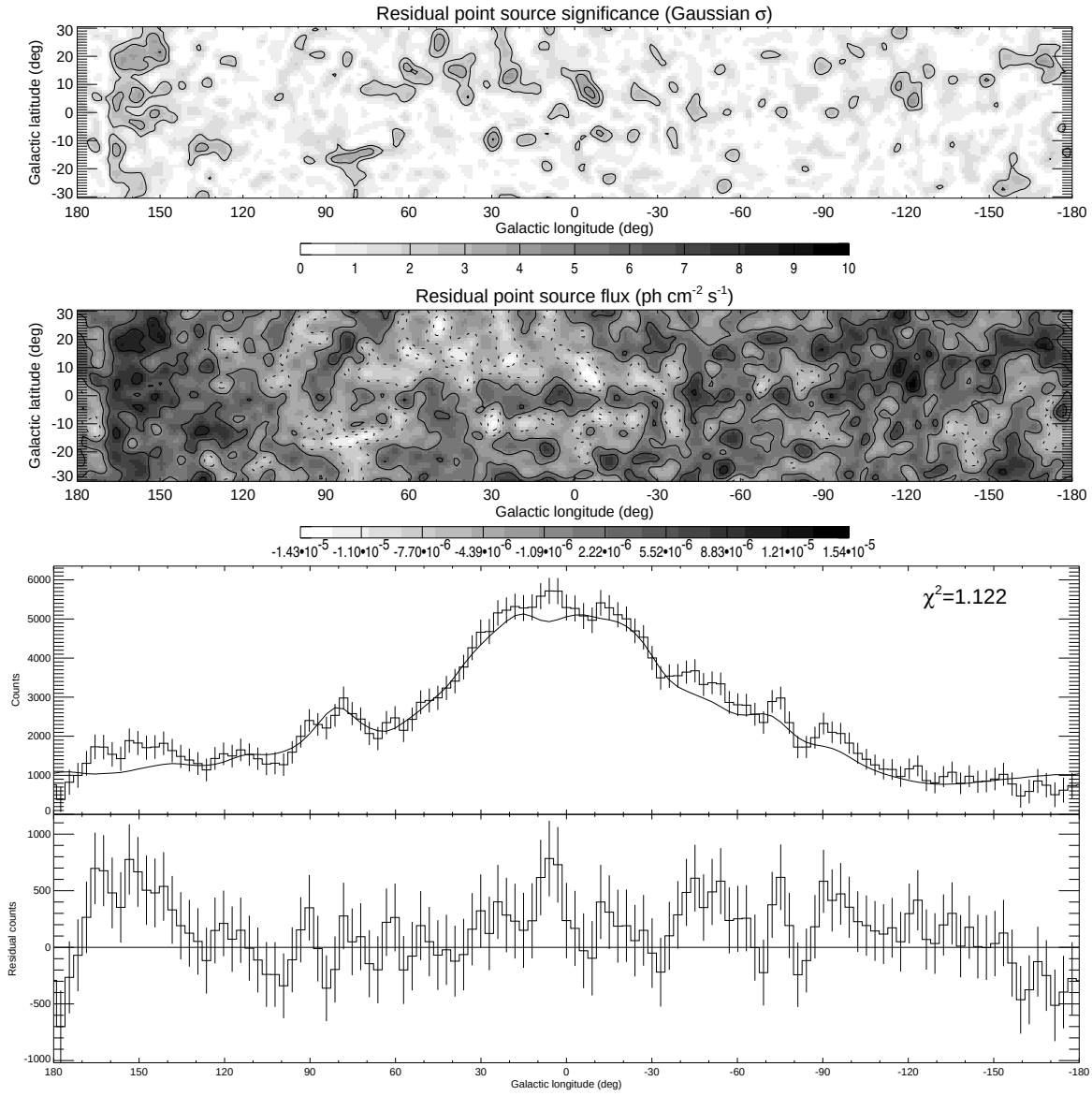

Fig. D.4: DMR free-free emission model.

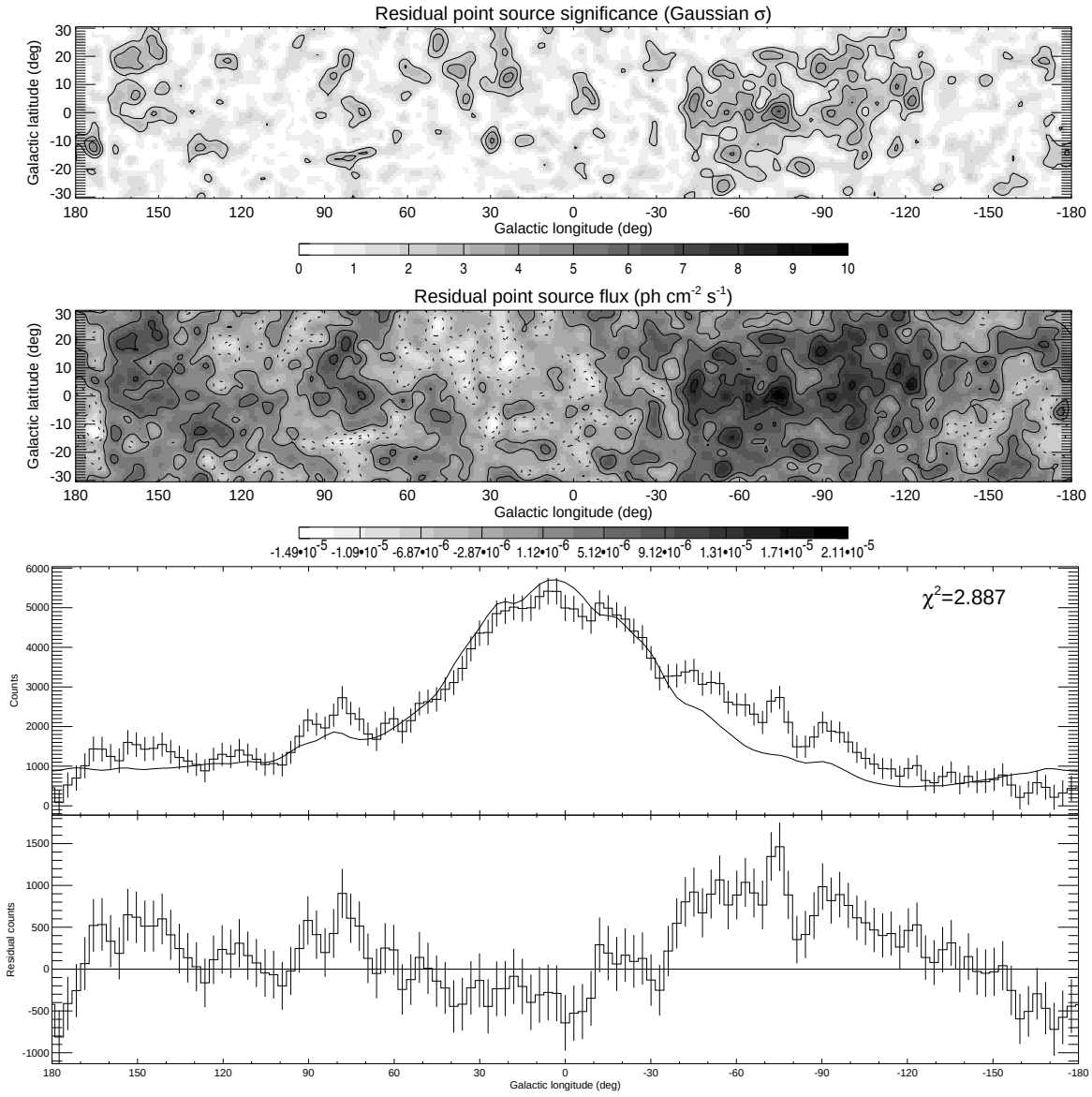

Fig. D.5: CO model.

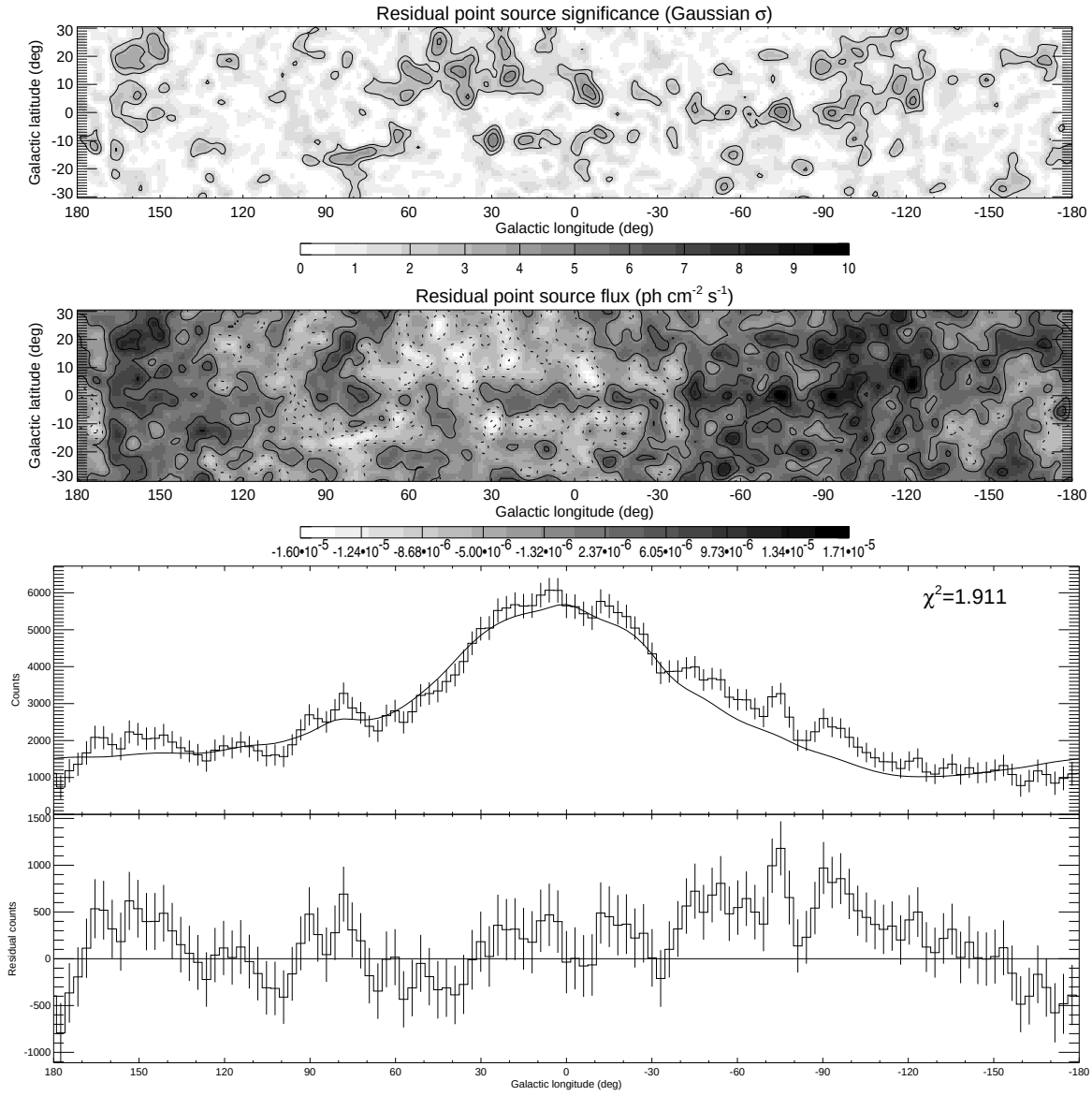

Fig. D.6: DMR dust model.

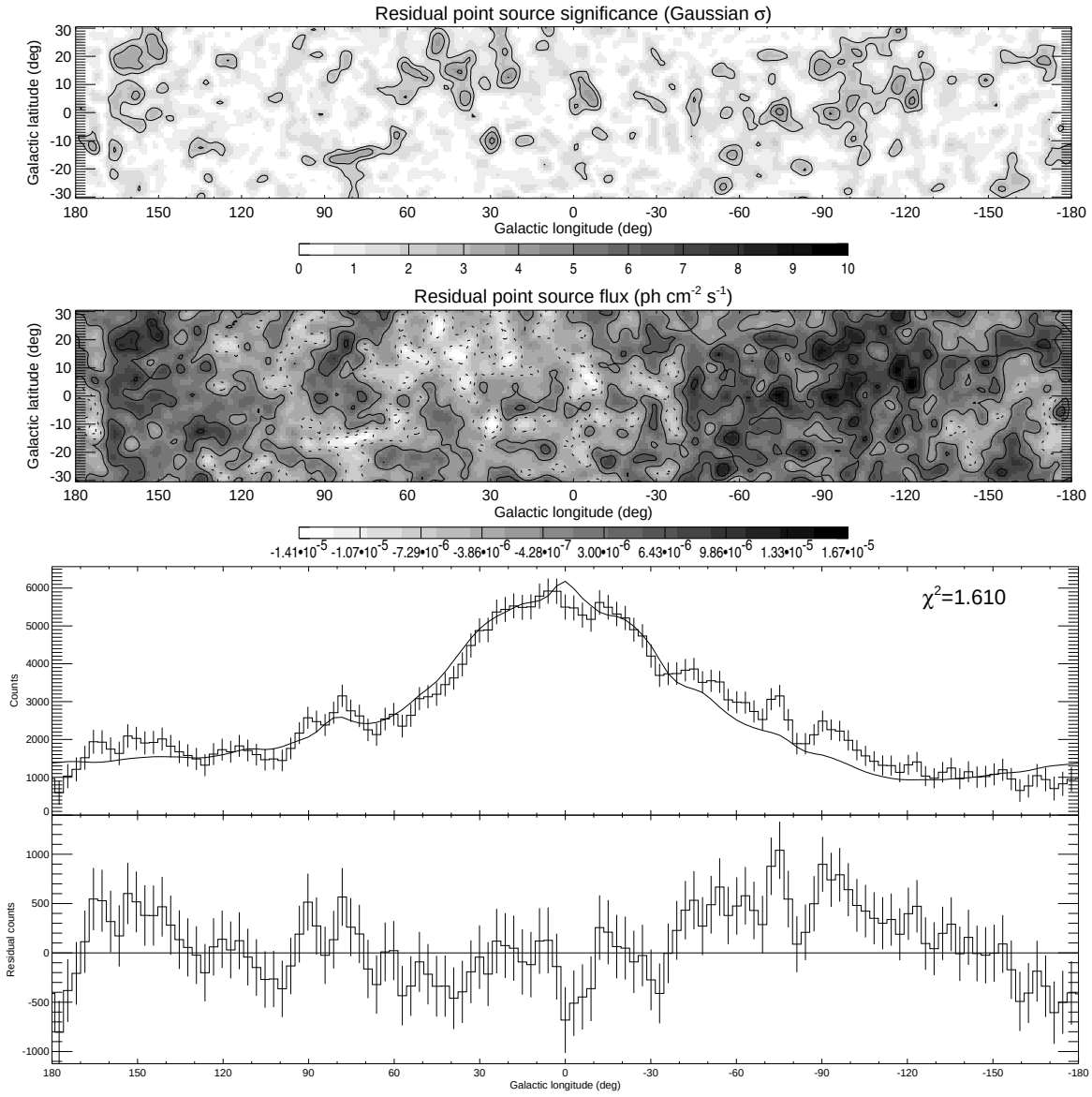

Fig. D.7: DIRBE 240 μm model.

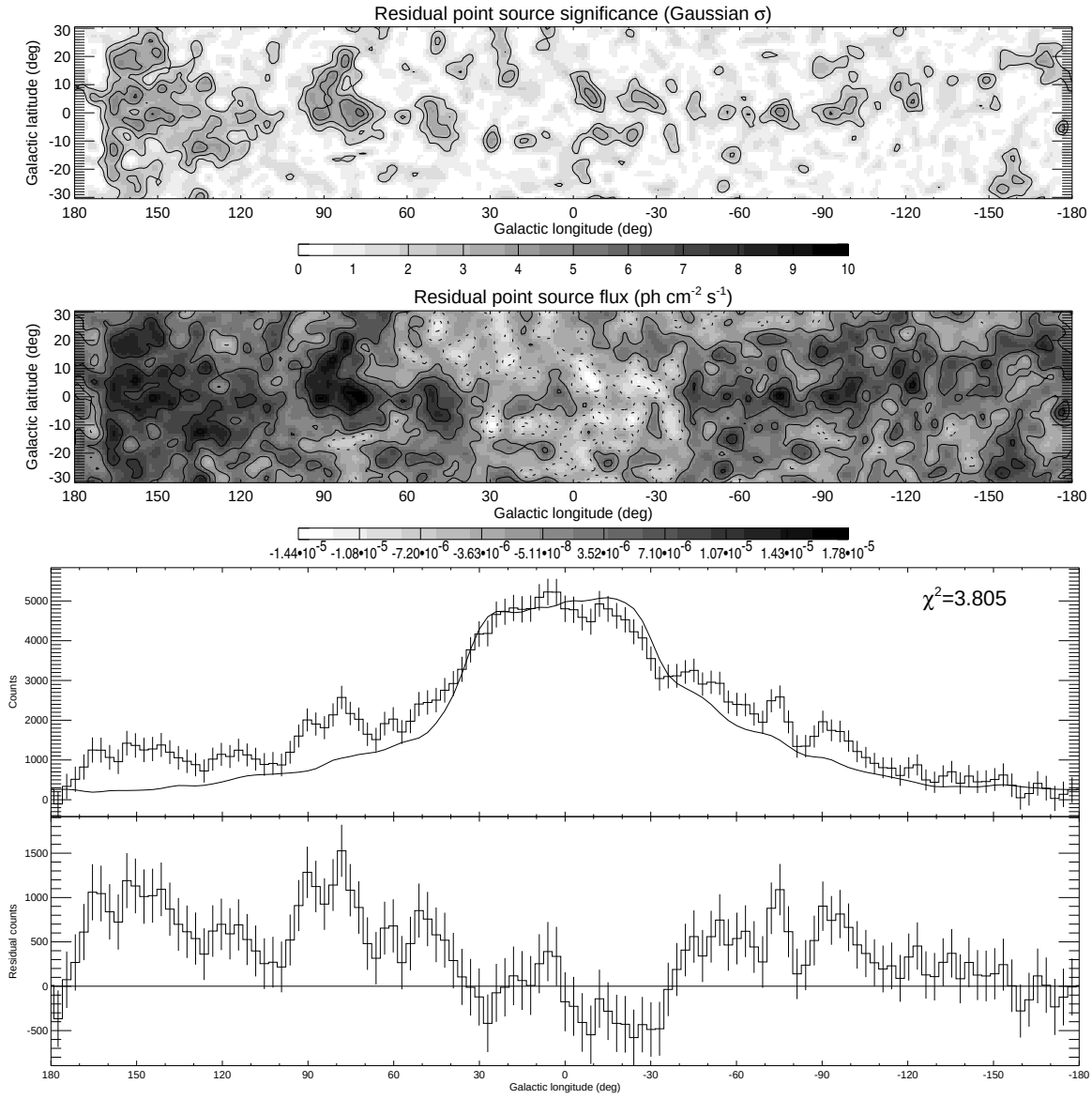

Fig. D.8: FIRAS 205 μm model.

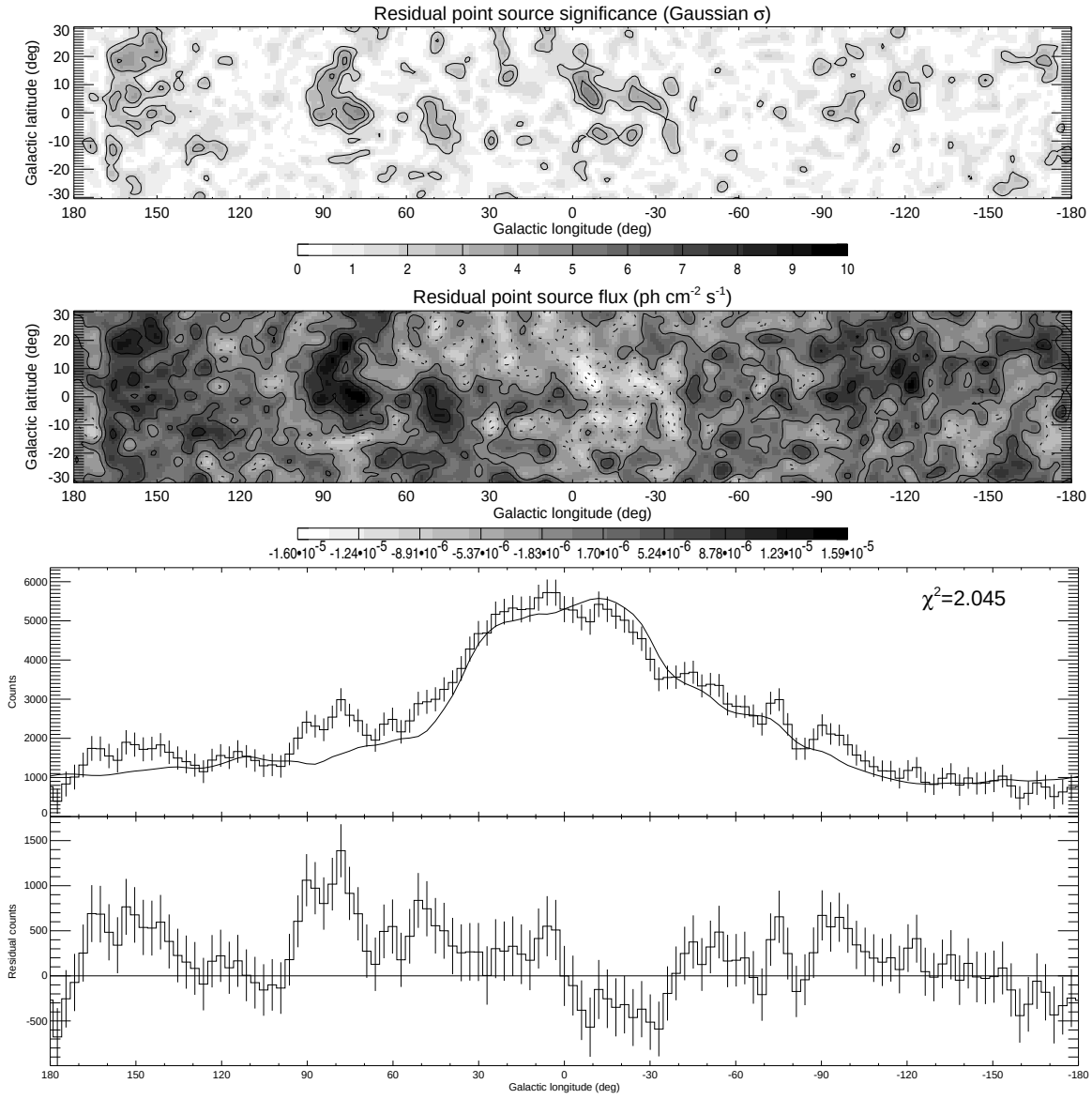

Fig. D.9: FIRAS 158 μm model.

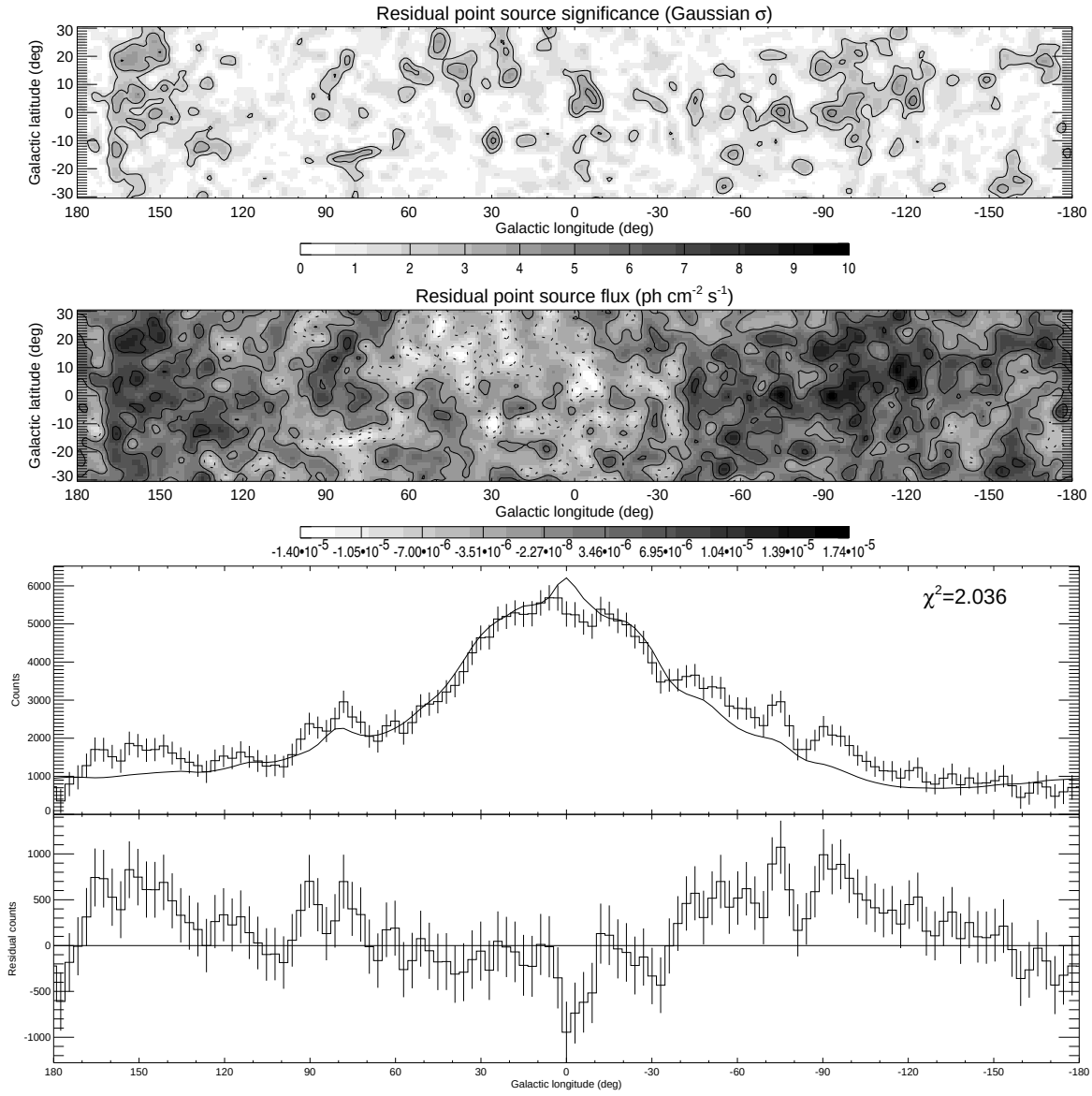

Fig. D.10: DIRBE 140 μm model.

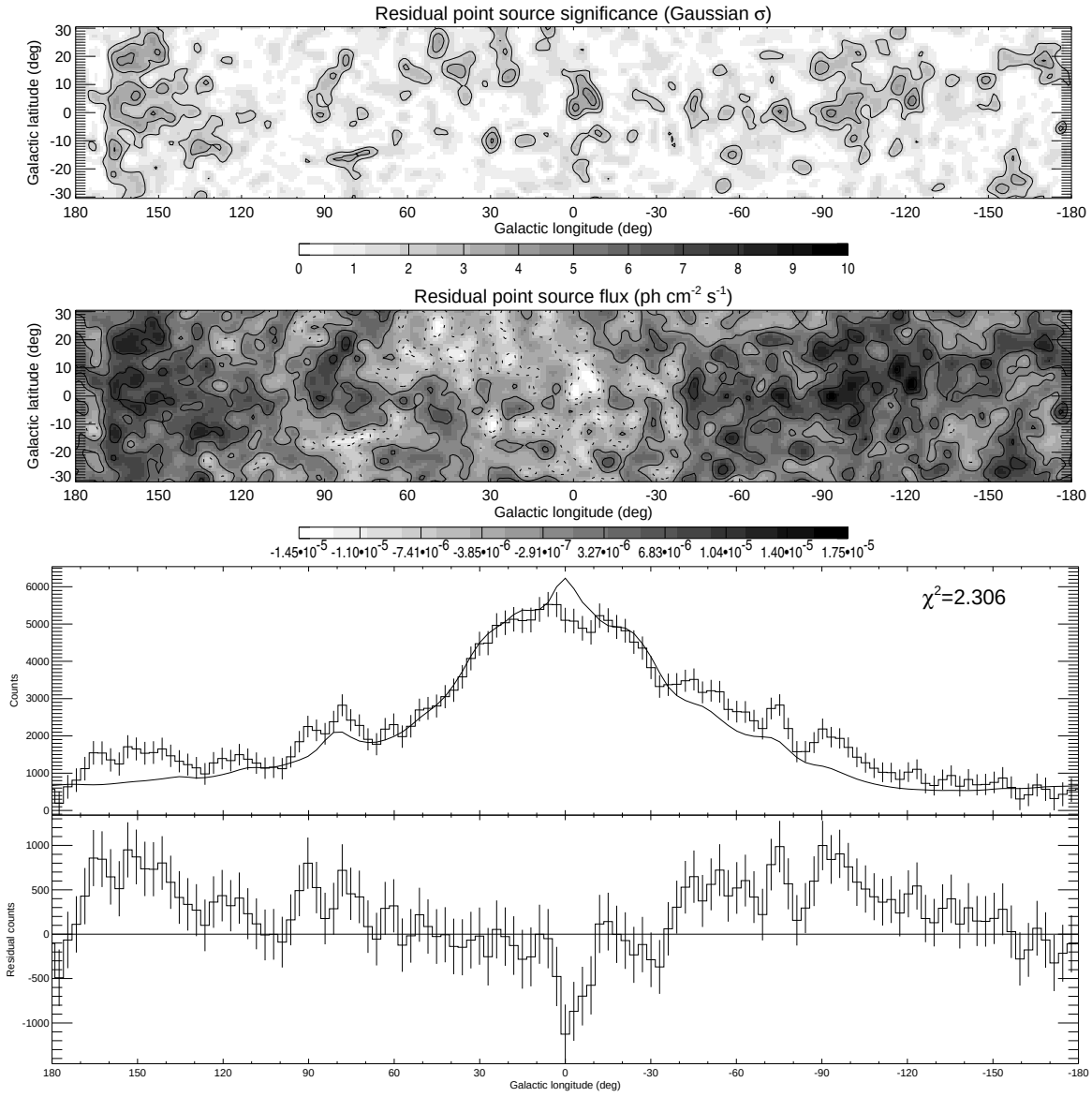

Fig. D.11: DIRBE 100 μm model.

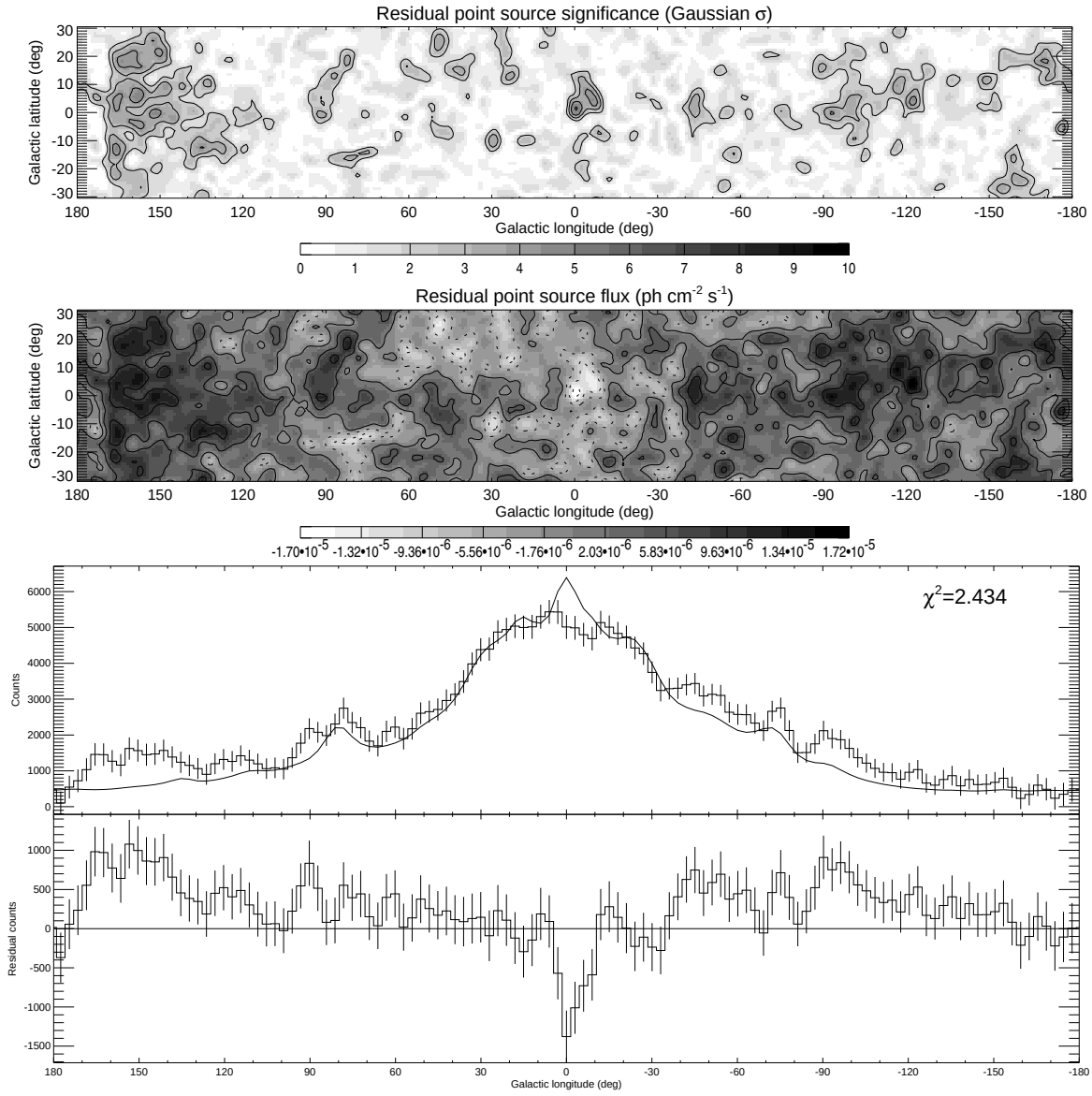

Fig. D.12: DIRBE 60 μm model.

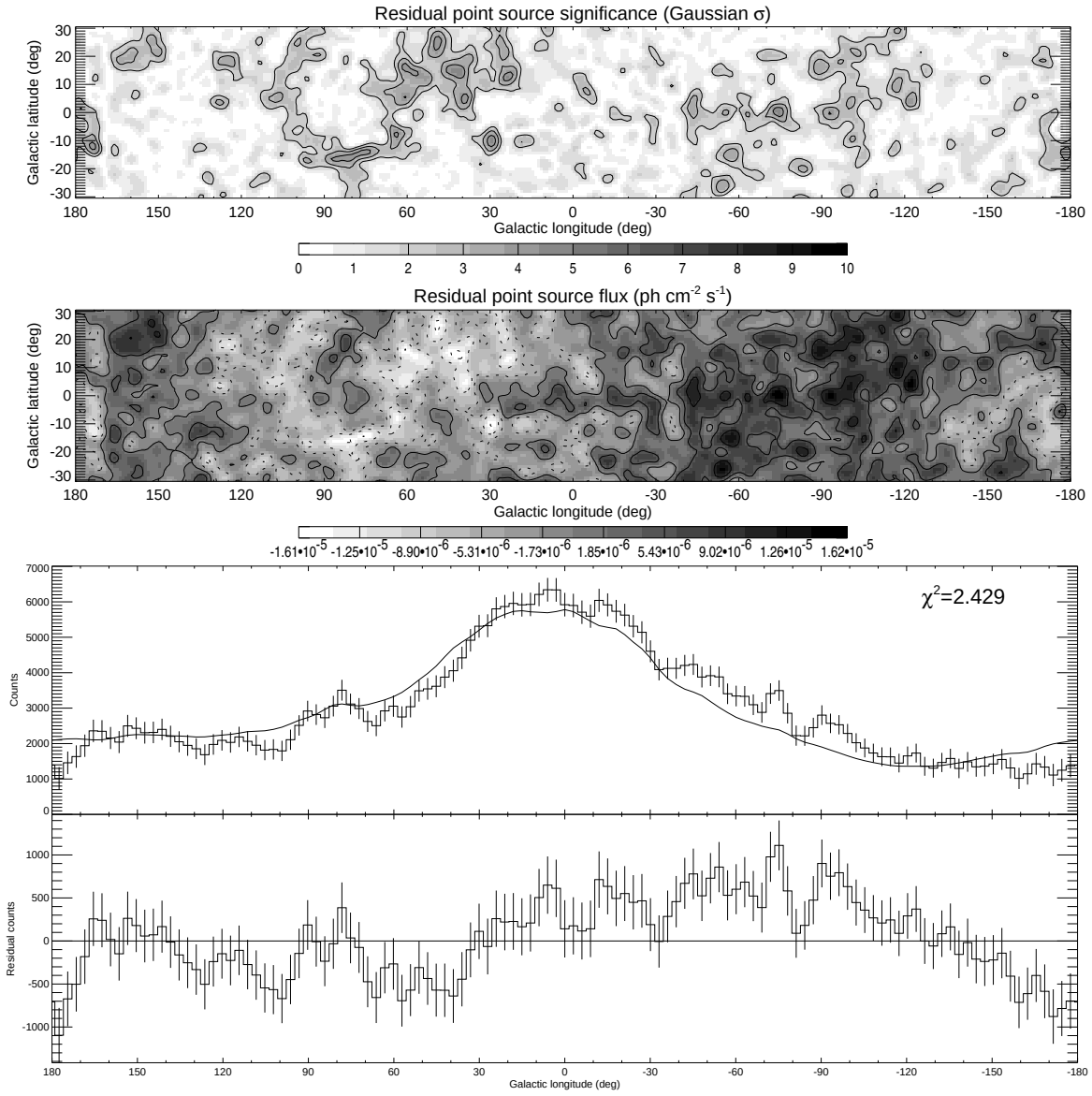

Fig. D.13: EGRET > 100 MeV model.

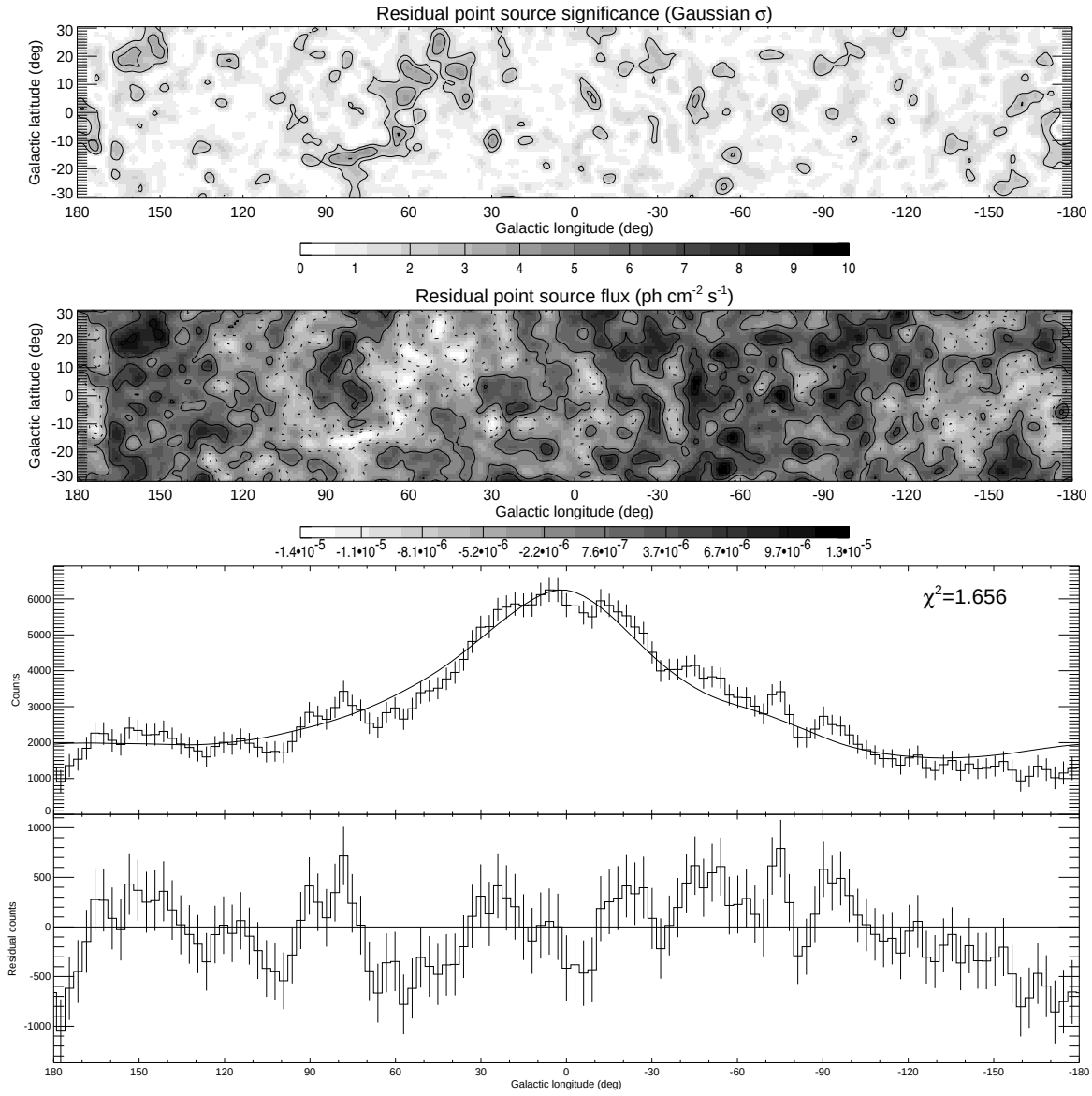

Fig. D.14: Exponential disk of scale length $R_0 = 4.5$ kpc and scale height $z_0 = 90$ pc.

Index

Bibliography

Abbott, D. C. 1982, Astrophys. J. 263, 723

Allen, J. S., K. Jahoda, L. A. Whitlock 1994, Legacy 5 5, 27

Arnett, D. W. 1969, Astrophys. J. 157, 1369

Arnett, W. D. 1977, Ann. N.Y. Acad. Sci. 302, 90

Bahcall, J. N., R. M. Soneira 1980, Astrophys. J., Suppl. Ser. 44, 73

Barbuy, B., *et al.* 1992, Astron. Astrophys. 262, 216

Bazan, G., *et al.* 1993, Rev. Mex. Astron. Astrofis. 27, 87

Bennett, C. L., *et al.* 1992, Astrophys. J. 396, L7

Bennett, C. L., *et al.* 1994a, Astrophys. J. 436, 423

Bennett, C. L., *et al.* 1994b, Astrophys. J. 434, 587

Berkhuijsen, E. M., C. G. T. Haslam, C. J. Salter 1971, Astron. Astrophys. 14, 252

Bernard, J. P., *et al.* 1995, in *Unveiling the Cosmic Infrared Background*, ed. E. Dwek, p. 105–111 (AIP Conference Proceedings)

Blake, J. B., D. S. P. Dearborn 1989, Astrophys. J. 338, L17

Bloecker, T. 1995, Astron. Astrophys. 297, 727

Bloemen, H., *et al.* 1994a, Astrophys. J., Suppl. Ser. 92, 419

Bloemen, H., *et al.* 1994b, Astron. Astrophys. 281, L5

Bloemen, J. B. G. M., E. Deul, E. Thaddeus 1990, Astron. Astrophys. 233, 437

Boggess, N. W., *et al.* 1992, Astrophys. J. 397, 420

Bontekoe, T. R., E. Koper, D. J. M. Kester 1994, Astron. Astrophys. 284, 1037

Boulanger, F., M. Pérault 1988, Astrophys. J. 330, 964

Braun, H., N. Langer 1994, in *WR stars: binaries, colliding winds, evolution*, eds. K. van der Hucht, P. M. Williams, Vol. 163, p. in press (IAU)

Bretthorst, G. 1990, in *Maximum Entropy and Bayesian Methods*, ed. P. F. Fougère, Vol. 39 of *The Fundamental Theories of Physics: Their Clarification, Development and Application*, p. 53–79 (Kluwer Academic Publishers)

Bretthorst, G. L. 1988, *Bayesian Spectrum Analysis and Parameter Estimation* (Springer-Verlag, Berlin), 1st Edition

Bronfman, L. 1992, in *The Center, Bulge, and Disk of the Milky Way*, ed. L. Blitz, p. 131–154 (Kluwer Academic Publishers)

Brown, R. L. 1987, in *Spectroscopy of astrphysical plasmas*, eds. A. Dalgarno, D. Layzer, Cambridge Astrphysics Series, p. 35–58 (Cambridge University Press, Cambridge)

Burbidge, M., *et al.* 1957, Rev. Mod. Phys. 29, 547

Cappellaro, E., *et al.* 1993, Astron. Astrophys. 273, 383

Cappellaro, E., *et al.* 1997, Astron. Astrophys. 322, 431

Cash, W. 1979, Astrophys. J. 228, 939

Chen, W., *et al.* 1996, Astron. Astrophys. Suppl. Ser. 120C, 315

Chen, W., *et al.* 1997, in *Proceedings 2nd INTEGRAL Workshop 'The Trasparent Univers'*, eds. C. Winkler, T. J.-L. Courvoisier, P. Durouchoux, p. 105–108 (ESA Publications Division)

Chen, W., N. Gehrels, R. Diehl 1995, Astrophys. J. 440, L57

Clayton, D. D. 1984, Astrophys. J. 280, 144

Clayton, D. D., W. Craddock 1965, Astrophys. J. 142, 189

Clayton, D. D., F. Hoyle 1974, Astrophys. J. 187, L101

Coc, A., *et al.* 1995, Astron. Astrophys. 299, 479

Coifman, R. R., D. L. Donoho 1995, *Translation-Invariant De-Noising*, tech. rep., Stanford University, Department of Statistics, To Appear, Wavelets and Statistics, Anestis Antoniadis, ed. Springer-Verlag Lecture Notes

Combes, F. 1991, Annu. Rev. Astron. Astrophys. 29, 195

Comerón, F., J. Torra 1996, Astron. Astrophys. 314, 776

Cruz-Gonzalez, C., *et al.* 1974, Rev. Mex. Astron. Astrofis. 1, 211

Dame, T. M., *et al.* 1987, Astrophys. J. 322, 706

Daubechies, I. 1988, Communications on Pure and Applied Mathematics 41, 909

de Boer, H., *et al.* 1992, in *DATA ANALYSIS IN ASTRONOMY - IV*, ed. D. Gesù, p. 241 (Plenum Press, New York)

del Rio, E., *et al.* 1996, Astron. Astrophys. 315, 237

Dempster, A., N. Laird, D. Rubin 1977, J. Royal Stat. Soc. Ser. B 39, 1

Dickey, J. M., F. J. Lockman 1990, Annu. Rev. Astron. Astrophys. 28, 215

Diehl, R., *et al.* 1992a, in *The Compton Observatory Science Workshop*, p. 95–101

Diehl, R., *et al.* 1992b, in *DATA ANALYSIS IN ASTRONOMY - IV*, p. 201

Diehl, R., *et al.* 1995a, Astron. Astrophys. 298, L25

Diehl, R., *et al.* 1995b, Astron. Astrophys. 298, 445

Diehl, R., *et al.* 1996, Astron. Astrophys. Suppl. Ser. 120C, 321

Diehl, R., *et al.* 1997, in press

Donoho, D. L. 1993, in *Proceedings of Symposia in Applied Mathematics*, Vol. 47, p. 173–205, American Mathematical Society

Donoho, D. L., I. M. Johnstone 1994, Biometrika 81(3), 425

Doom, C. 1987, Astron. Astrophys. 182, L43

Dove, J. B., J. M. Shull 1994, Astrophys. J. 430, 222

Dwek, E., *et al.* 1995, Astrophys. J. 445, 716

Eadie, W., *et al.* 1982, *Statistical Methods in Experimental Physics* (North-Holland Publishing Company, Amsterdam), 1st Edition

Egger, R. 1995, in *The physics of the interstellar medium and the inter-galactic medium*, eds. A. Ferrara, C. Heiles, C. McKee, P. Shapiro, Vol. 80 of *PASP*, p. 45

Ferguson, A. M. N., *et al.* 1996a, Astron. J. 111, 2265

Ferguson, A. M. N., *et al.* 1996b, Astron. J. 112, 2567

Fessler, J. A., A. O. Hero 1994, *Space-alternating generalized EM algorithms for penalized maximum-likelihood image reconstruction*, tech. rep., Comm. and Sign. Proc. Lab., Dept. of EECS, Univ. of Michigan, Ann Arbor, MI, 48109-2122

Forestini, M., C. Charbonnel 1997, Astron. Astrophys. Suppl. Ser. 123, 241

Freeman, K. C. 1992, in *The Stellar Populations of Galaxies*, eds. B. Barbuy, A. Renzini, Vol. 149 of *IAU*, p. 65–73 (IAU)

Fuller, G. M., W. A. Fowler, M. J. Newman 1982, Astrophys. J., Suppl. Ser. 48, 279

Garmany, C. D., P. S. Conti, C. Chiosi 1982, Astrophys. J. 263, 777

Gehrels, N., W. Chen 1996, Astron. Astrophys. Suppl. Ser. 120C, 331

Gies, D. R., D. L. Lambert 1992, Astrophys. J. 387, 673

Graps, A. 1995, IEEE Computational Science and Engineering 2, 50

Gull, S., J. Skilling 1984, in *IEE Proc.*, Vol. 131 (F), p. 646–659

Gull, S., J. Skilling 1985, in *Maximum-Entropy and Bayesian Methods in Inverse Problems*, eds. C. R. Smith, J. W.T. Grandy, Vol. 14 of *The Fundamental Theories of Physics: Their Clarification, Development and Application*, p. 287–301 (D.Reidel Publishing Company)

Hartmann, D. 1994, in *AIP*, eds. C. E. Fichtel, N. Gehrels, J. P. Norris, Vol. 304, p. 176

Hasinger, G., *et al.* 1993, Astron. Astrophys. 275, 1

Haslam, C. G. T., *et al.* 1981, Nature 289, 470

Haslam, C. G. T., *et al.* 1982, Astron. Astrophys. Suppl. Ser. 47, 1

Heiles, C., W. T. Reach, B.-C. Koo 1996, Astrophys. J. 466, 191

Herrero, A., *et al.* 1992, Astron. Astrophys. 261, 209

Hillebrandt, W., F.-K. Thielemann, N. Langer 1987, Astrophys. J. 321, 761

Hoopes, C. G., R. A. M. Walterbos, B. E. Greenwalt 1996, Astron. J. 112, 1429

Humphreys, R. M., D. B. McElroy 1984, Astrophys. J. 284, 565

Iyudin, A. F., *et al.* 1994, Astron. Astrophys. 284, L1

Jaynes, E. 1986, in *Maximum Entropy and Bayesian Methods in Applied Statistics*, ed. J. H. Justice (Camdridge University Press, Cambridge)

Jaynes, E. 1996, *Probability Theory: The Logic of Science* (Camdridge University Press, Cambridge), fragmentary Edition

Jeffreys, H. 1939, *Theory of Probability* (Clarendon Press, Oxford), 1st Edition

José, J., M. Hernanz, A. Coc 1997, Astrophys. J. 479, L55

Kass, R. E., L. Wasserman 1994, *Formal Rules for Selecting Prior Distributions: A Review and Annotated Bibliography*, tech. rep., Carnegie Mellon University, Carnegie Mellon University

Kaufman, L. 1987, IEEE Trans. Med. Img. 6, 37

Kebede, L. 1994, Astrophys. J. 423, 878

Kent, S. M., T. M. Dame, G. Fazio 1991, Astrophys. J. 378, 131

Kniffen, D. A., *et al.* 1996, Astron. Astrophys. Suppl. Ser. 120C, 615

Knödlseder, J. 1994, *Untersuchung spezieller Regionen mit 1.8 MeV Linienemission von radioaktivem ^{26}Al*, master thesis, Technische Universität München

Knödlseder, J., *et al.* 1996a, Astron. Astrophys. Suppl. Ser. 120C, 327

Knödlseder, J., *et al.* 1996b, Proc. SPIE 2806, 386

Knödlseder, J., *et al.* 1996c, Astron. Astrophys. Suppl. Ser. 120, 335

Knoll, G. F. 1979, *Radiation Detection and Measurements* (John Wiley & Sons New York)

Kogut, A., *et al.* 1996a, Astrophys. J. 470, 653

Kogut, A., *et al.* 1996b, Astrophys. J. 460, 1

Kogut, A., *et al.* 1996c, Astrophys. J. 464, L5

Kolb, U., M. Politano 1997, Astron. Astrophys. 319, 909

Kroupa, P., C. A. Tout, G. Gilmore 1993, Mon. Not. R. Astron. Soc. 262, 545

Lambert, D. 1989, in *Evolution of Peculiar Red Giants*, eds. H. Johnson, B. Zuckerman, p. 101– (IAU Colloqu. 106)

Langer, N. 1987, in *Nuclear Astrophysics*, ed. W. Hillebrandt, p. 180–186

Langer, N. 1995, *Leben und Sterben der Sterne*, C.H. Beck Wissen (C.H. Beck'sche Verlagsbuchhandlung, München), 1st Edition

Langer, N., H. Braun, J. Fliegner 1995, Astrophys. Space. Sci. 224, 275

Langer, N., J. Fliegner, A. Heger 1996, in *8th Workshop on Nuclear Astrophysics*, eds. W. Hillebrandt, E. Müller, p. 5–9

Law, W. Y., H. Ritter 1983, Astron. Astrophys. 123, 33

Lederer, C. M., V. S. Shirley 1978, *Table of Isotopes* (John Wiley & Sons, Inc.), 7th Edition

Lennon, D. J. 1994, Space Sci. Rev. 66, 127

Lichti, G. G., *et al.* 1993, Astron. Astrophys. Suppl. Ser. 97, 215

Livio, M., J. W. Truran 1994, Astrophys. J. 425, 797

Loredo, T. 1990, in *Maximum Entropy and Bayesian Methods*, ed. P. F. Fougère, Vol. 39 of *The Fundamental Theories of Physics: Their Clarification, Development and Application*, p. 81–142 (Kluwer Academic Publishers)

Lucy, L. 1974, Astron. J. 79, 745

Lucy, L. 1994, Rev. Mod. Astron. 7, 31

Lucy, L. B. 1994, Astron. Astrophys. 289, 983

MacCallum, C. J., *et al.* 1987, Astrophys. J. 317, 877

Madigan, D., A. E. Raftery 1994, Journal of the American Statistical Association 89, 1335

Maeder, A. 1987, Astron. Astrophys. 178, 159

Maeder, A. 1994, Physical processes in astrophysics in press

Mahoney, W. A., *et al.* 1984, Astrophys. J. 286, 578

Malet, I., *et al.* 1991, in *Gamma-ray line astrophysics*, eds. P. Durouchoux, N. Prantzos, Vol. 232 of *AIP Conference Proceedings*, p. 123–128 (American Institute of Physics)

Massey, P., K. E. Johnson, K. Degioia-Eastwood 1995, Astrophys. J. 454, 151

Mathis, J. S. 1986, Astrophys. J. 301, 423

McKee, C. F., J. P. Ostriker 1977, Astrophys. J. 218, 148

Meynet, G., *et al.* 1997, Astron. Astrophys. 320, 460

Mezger, P. G. 1978, Astron. Astrophys. 70, 565

Mezger, P. G. 1988, in *Galactic and Extragalactic Star Formation*, eds. R. E. Pudritz, M. Fich, p. 227–250 (Kluwer Academic Publishers)

Miller, G. E., J. M. Scalo 1979, Astrophys. J., Suppl. Ser. 41, 513

Miller, W. W., D. P. Cox 1993, Astrophys. J. 417, 579

Morfill, G. E., T. W. Hartquist 1985, Astrophys. J. 297, 194

Morris, D. J., *et al.* 1995, in *Seventeenth Texas Symposium on Relativistic Astrophysics and Cosmology*, eds. H. Böhringer, G. E. Morfill, J. E. Trümper, Vol. 759 of *Annals of the New York Academy of Sciences*, p. 397–400 (The New York Academy of Sciences)

Muller, R. A., *et al.* 1992, Astrophys. J. 384, L9

Murtagh, F., J. L. Starck, A. Bijaoui 1995, Astron. Astrophys. Suppl. Ser. 112, 179

Naya, J. E., *et al.* 1996, Nature 384, 44

Nofar, I., G. Shaviv, S. Starrfield 1991, Astrophys. J. 369, 440

Nomoto, K. 1984, Astrophys. J. 277, 791

Oberlack, U. 1997, *Über die Natur der galaktischen ^{26}Al-Quellen. Untersuchung des 1,8-MeV-Himmels mit COMPTEL*, PhD thesis, Technische Universität München

Oberlack, U., *et al.* 1994, Astrophys. J., Suppl. Ser. 92, 433

Oberlack, U., *et al.* 1996, Astron. Astrophys. Suppl. Ser. 120C, 311

Pantin, E., J. L. Starck 1996, Astron. Astrophys. Suppl. Ser. 118, 575

Pauldrach, A. W. A., *et al.* 1994, Astron. Astrophys. 283, 525

Péquignot, D., C. Boisson, P. Petitjean 1991, Astron. Astrophys. 251, 680

Perry, K., S. Reeves 1994, in *The Restoration of HST Images and Spectra II*, eds. R. Hanisch, R. White, p. 97–103 (Space Telescope Science Institute)

Phillipps, S., *et al.* 1981a, Astron. Astrophys. 98, 286

Phillipps, S., *et al.* 1981b, Astron. Astrophys. 103, 405

Politano, M., *et al.* 1995, Astrophys. J. 448, 807

Prantzos, N., M. Cassé 1986, Astrophys. J. 307, 324

Prantzos, N., R. Diehl 1996, Phys. Rep. 267, 1

Press, W., *et al.* 1992, *Numerical Recipes in FORTRAN* (Camdridge University Press, Cambridge), 2nd Edition

Prialink, D., M. M. Shara 1995, Astron. J. 109, 1735

Puetter, R. C. 1995a, in *Proc. of the 11th Florida Workshop on Non-linear Astronomy and Physics*, p. in press (University of Florida)

Puetter, R. C. 1995b, Int. J. Image Sys. & Tech. 6, 314

Purcell, W. R. 1989, *Gamma-ray observations and the distribution of interstellar aluminum-26*, PhD thesis, Northwestern Univ., Evanston, IL.

Raftery, A. E. 1994, *Approximate Bayes Factors and Accounting for Model Uncertainty in Generalized Linear Models*, tech. rep., University of Washington, University of Washington

Raftery, A. E. 1995, in *Social Methodology 1995*, ed. P. Marsden (Blackwells, Cambridge)

Ramaty, R., B. Kozlovsky, R. E. Lingenfelter 1979, Astrophys. J., Suppl. Ser. 40, 487

Ramaty, R., R. E. Lingenfelter 1977, Astrophys. J. 213, L5

Reynolds, R. J. 1984, Astrophys. J. 282, 191

Reynolds, R. J. 1991a, in *The Disk-Halo Connection in Galaxies*, ed. H. Bloemen (Kluwer, Dordrecht)

Reynolds, R. J. 1991b, Astrophys. J. 372, L17

Reynolds, R. J. 1992, Astrophys. J. 392, L35

Richardson, W. 1972, J. Opt. Soc. Am. 62, 55

Rothschild, R., *et al.* 1979, Sp. Sci. Instr. 4, 269

Ryan, J., *et al.* 1993, Adv. Space Res. 13, 255

Sahu, M. S. 1992, *A study of the ISM in Puppis-Vela including the Gum nebula*, PhD thesis, Rijksuniversiteit Groningen

Salpeter, E. E. 1955, Astrophys. J. 121, 161

Schaller, G., *et al.* 1992, Astron. Astrophys. Suppl. Ser. 96, 269

Schönberner, D., *et al.* 1988, Astron. Astrophys. 197, 209

Schönfelder, V., *et al.* 1993, Astrophys. J., Suppl. Ser. 86, 657

Schönfelder, V., *et al.* 1996, Astron. Astrophys. Suppl. Ser. 120C, 13

Schönfelder, V., A. Hirner, K. Schneider 1973, Nucl. Instr. Meth. 107, 385

Schwarz, H. 1988, *Numerische Mathematik* (B.G. Teubner Stuttgart, Stuttgart), 2nd Edition

Setti, G. 1995, in *Seventeenth Texas Symposium on relavistic astrophysics and cosmology*, eds. H. Böhringer, G. Morfill, J. Trümper, Vol. 759 of *Annals of the New York Academy of Sciences*, p. 110–126 (The New York Academy of Sciences)

Shara, M. M., D. Prialnik 1994, Astron. J. 107, 1542

Share, G. H., *et al.* 1985, Astrophys. J. 292, L61

Shaver, P. A., *et al.* 1983, Mon. Not. R. Astron. Soc. 204, 53

Shepp, L., Y. Vardi 1982, IEEE Trans. Med. Img. 1, 113

Simoncelli, E. P., *et al.* 1992, IEEE Transactions on Information Theory 38(2), 587

Skibo, J., R. Ramaty 1991, in *Gamma-Ral Line Astrophysics*, eds. P. Durouchoux, N. Prantzos, Vol. 232 of *AIP Conference Proceedings*, p. 168–170 (AIP)

Skilling, J. 1988, in *Maximum-Entropy and Bayesian Methods in Science and Engineering*, eds. G. Erickson, C. Smith, Vol. 1 of *The Fundamental Theories of Physics: Their Clarification, Development and Application*, p. 173–187 (Kluwer Academic Publishers)

Skilling, J. 1990, in *Maximum Entropy and Bayesian Methods*, ed. P. F. Fougère, Vol. 39 of *The Fundamental Theories of Physics: Their Clarification, Development and Application*, p. 341–350 (Kluwer Academic Publishers)

Skilling, J., R. K. Bryan 1984, Mon. Not. R. Astron. Soc. 211, 111

Smith, V. V., D. L. Lambert 1989, Astrophys. J. 345, L75

Snowden, S. L., *et al.* 1993, Astrophys. J. 454, 643

Snowden, S. L., *et al.* 1997, Astrophys. J. 485, 125

Sodroski, T. J., *et al.* 1995, Astrophys. J. 452, 262

Sodroski, T. J., *et al.* 1997, Astrophys. J. 480, 173

Starck, J. L., F. Murtagh 1994, Astron. Astrophys. 288, 342

Strong, A. 1985, Astron. Astrophys. 150, 273

Strong, A., *et al.* 1988, Astron. Astrophys. 207, 1

Strong, A. W. 1995, Exp. Astron. 6, 97

Strong, A. W., *et al.* 1992, in *DATA ANALYSIS IN ASTRONOMY - IV*, ed. D. Gesù, p. 251 (Plenum Press, New York)

Strong, A. W., *et al.* 1993, in *Compton Gamma Ray Observatory*, Vol. 280, p. 70– (AIP Conf. Proc.)

Strong, A. W., *et al.* 1997, in *Proceedings 2nd INTEGRAL Workshop 'The Transparent Universe'*, eds. C. Winkler, T. J.-L. Courvoisier, P. Durouchoux, Vol. 382, p. 533–536 (ESA Publications Division, ESTEC, Nordwijk, The Netherlands)

Tammann, G. A., W. Loeffler, A. Schroeder 1994, Astrophys. J., Suppl. Ser. 92, 487

Taylor, J. H., J. M. Cordes 1993, Astrophys. J. 411, 674

Teegarden, B. J., *et al.* 1991, in *Gamma-ray line astrophysics*, eds. P. Durouchoux, N. Prantzos, Vol. 232 of *AIP Conference Proceedings*, p. 116–122 (American Institute of Physics)

Thielemann, F.-K., M.-A. Hashimoto, K. Nomoto 1994, in *Les Houches Ecole d'Ete de Physique Theorique*, p. 629

Thielemann, F.-K., K. Nomoto, M.-A. Hashimoto 1996, Astrophys. J. 460, 408

Thompson, D. J., *et al.* 1995, Astrophys. J., Suppl. Ser. 101, 259

Tierney, L., J. B. Kadane 1986, J. Amer. Statist. Ass. 81, 82

Timmes, F. X., *et al.* 1995, Astrophys. J. 449, 204

Timmes, F. X., R. Diehl, D. H. Hartmann 1997, Astrophys. J. 479, 760

Truran, J. W., A. G. W. Cameron 1978, Astrophys. J. 219, 226

Truran, J. W., M. Livio 1986, Astrophys. J. 308, 721

Vacca, W. D. 1994, Astrophys. J. 421, 140

Vacca, W. D., C. D. Garmany, J. M. Shull 1996, Astrophys. J. 460, 914

van den Bergh, S., R. D. McClure 1994, Astrophys. J. 425, 205

van der Hucht, K. A., *et al.* 1988, Astron. Astrophys. 199, 217

van Dijk, R. 1996, *Gamma-ray observations of X-ray binaries with COMPTEL*, PhD thesis, Sterrenkundig Instituut "Anton Pannekoek", Universiteit van Amsterdam

Vassiliadis, E., P. R. Wood 1993, Astrophys. J. 413, 641

Vidaković, B., P. Müller 1991, *Wavelets for kids: A tutorial introduction*, tech. rep., Duke University

von Ballmoos, P., *et al.* 1996, Proc. SPIE 2806, 372

von Ballmoos, P., R. Diehl, V. Schoenfelder 1987, Astrophys. J. 318, 654

Wainscoat, R. J., *et al.* 1992, Astrophys. J., Suppl. Ser. 83, 111

Walter, F. M., W. T. Boyd 1991, Astrophys. J. 370, 318

Walter, R., A. Maeder 1989, Astron. Astrophys. 218, 123

Walterbos, R. A. M., R. Braun 1994, Astrophys. J. 431, 156

Wasserburg, G. J., *et al.* 1994, Astrophys. J. 424, 412

Watson, W. D. 1972, Astrophys. J. 176, 103

Weaver, T. A., S. E. Woosley 1993, Phys. Rep. 227, 65

Weiland, J. L., *et al.* 1994, Astrophys. J. 425, L81

Weiler, K. W., R. A. Sramek 1988, Annu. Rev. Astron. Astrophys. 26, 295

Weiss, A., J. W. Truran 1990, Astron. Astrophys. 238, 178

White, R. 1994, in *Astronomical Data Analysis Software and Systems III*, eds. D. Crabtree, R. Hanisch, J. Barnes, Vol. 61 of *ASP Conference Series*, p. 292–299

Winkler, C. 1996, Astron. Astrophys. Suppl. Ser. 120C, 637

Woosley, S. E., *et al.* 1990, Astrophys. J. 356, 272

Woosley, S. E., N. Langer, T. A. Weaver 1993, Astrophys. J. 411, 823

Woosley, S. E., N. Langer, T. A. Weaver 1995, Astrophys. J. 448, 315

Woosley, S. E., T. A. Weaver 1980, Astrophys. J. 238, 1017

Woosley, S. E., T. A. Weaver 1995, Astrophys. J., Suppl. Ser. 101, 181

Wright, E. L., *et al.* 1991, Astrophys. J. 381, 200