SLAC-PUB-14861

Spectral components analysis of diffuse emission processes

Dmitry Malyshev[1]

Kavli Institute for Particle Astrophysics and Cosmology,
Department of Physics and SLAC National Accelerator Laboratory,
Stanford University, Stanford, CA 94305, USA

ABSTRACT

We develop a novel method to separate the components of a diffuse emission process based on an association with the energy spectra. Most of the existing methods use some information about the spatial distribution of components, e.g., closeness to an external template, independence of components etc., in order to separate them. In this paper we propose a method where one puts conditions on the spectra only. The advantages of our method are: 1) it is *internal*: the maps of the components are constructed as combinations of data in different energy bins, 2) the *components may be correlated* among each other, 3) the method is *semi-blind*: in many cases, it is sufficient to assume a functional form of the spectra and determine the parameters from a maximization of a likelihood function. As an example, we derive the CMB map and the foreground maps for seven yeas of *WMAP* data. In an Appendix, we present a generalization of the method, where one can also add a number of external templates.

Keywords: methods: data analysis; methods: statistical; gamma rays, infrared, submillimeter, X-rays: diffuse background; cosmology: cosmic background radiation.

[1]malyshev at stanford.edu
On leave of absence from ITEP, Moscow, Russia, B. Cheremushkinskaya 25

Submitted to Astrophysical Journal

Work supported in part by US Department of Energy contract DE-AC02-76SF00515.

1. Introduction

In many observations, understanding the diffuse emission components plays a significant role in the interpretation of results. For instance, diffuse background modeling is crucial for the observation of the CMB fluctuations in radio frequencies and for constraining DM annihilation in gamma-rays.

Diffuse emission in a pixel i at an energy E_α can be described as a sum over emission components μ

$$d_\alpha^i = \sum_{\mu=1}^{m} u_\alpha^{i\mu} s_\mu^i + r_\alpha^i, \tag{1}$$

where s_μ^i is the line of sight density of component μ at pixel i, $u_\alpha^{i\mu}$ is the corresponding energy spectrum at energy E_α (the matrices u are also called mixing matrices). r_α^i denotes random noise (instrumental or physical).

In various applications it is important to find the energy spectrum and/or the spatial distribution of the components. In many cases, one can neglect the dependence of the spectrum on pixels within the region of interest, i.e. $u_\alpha^{i\mu} = u_\alpha^\mu$. Let us also for the moment neglect the noise term r_α^i. Then the problem is to find u_α^μ and s_μ^i from the system of equations

$$d_\alpha^i = \sum_{\mu=1}^{m} u_\alpha^\mu s_\mu^i. \tag{2}$$

Suppose that the total number of energy bins (frequency bands) k is larger than the total number of significant components of the emission m. We also assume that the number of pixels N is larger than both k and m. Then, in general, this system is both overdetermined and redundant. It is overdetermined because the rank of the matrix d_α^i is k while the rank of the matrices u_α^μ and s_μ^i is m. It is redundant because there is a symmetry of multiplying the matrix u from the right with a matrix A and multiplying the matrix s from the left with the inverse of A

$$u \longrightarrow u \cdot A, \qquad\qquad s \longrightarrow A^{-1} \cdot s. \tag{3}$$

Different methods can be classified according to the usage of the overdetermination of the system to most efficiently extract the components of the emission and how they deal with the redundancy of the system.

One of the most simple cases is when either the spectra u or the spatial distributions s of the components are known. If u is known then s can be found from Equation (2), and vice versa,

$$\begin{aligned} s &= (u^T u)^{-1} u^T d, \\ u &= d s^T (s s^T)^{-1}. \end{aligned} \tag{4}$$

Denote $f = (u^T u)^{-1} u^T$. The matrix f satisfies $f \cdot u = I$, it is the Moore-Penrose pseudo-inverse matrix for the rectangular matrix u. Analogously, $h = s^T (s s^T)^{-1}$ is the pseudo-inverse for s, $s \cdot h = I$.

If the uncertainty does not depend on both the energy and the position, $\sigma_\alpha^i = \sigma$, then Equations (4) can be also derived by minimizing

$$\chi^2 = \sum_{\alpha,i} \frac{(d_\alpha^i - u_\alpha^\mu s_\mu^i)^2}{\sigma^2}. \tag{5}$$

There is a summation over the repeated index μ. In the following we will assume a summation over the repeated indices unless otherwise stated.

In general, the methods to separate the diffuse emission components have the following steps:

1. Choose a data analysis principle.
2. Find a functional such that the principle from p.1 is realized at the minimum (maximum).
3. Break the redundancy which is not broken by the functional.

In many methods the separating principle is the maximum likelihood. In the case of Gaussian statistics, the minimum of the χ^2 is the maximum likelihood point. The χ^2 has the change of the basis degeneracy (Equation (9)). Some examples of the ways to break the degeneracy are as follows (a review of the methods can be found in Leach et al. (2008))

1. **Independent components analysis** (ICA), Hyvarinen (1999); Maino et al. (2002). There is an additional term in the functional that depends on the mutual information (the minimum corresponds to maximal independence).

2. **Maximal entropy method** (MEM), Hinshaw et al. (2007); Hobson et al. (1998); Stolyarov et al. (2002). There is an additional term that quantifies a closeness to some prior (usually an external template).

3. **Generalized morphological component analysis** (GMCA), Bobin et al. (2007). There is an additional term that quantifies the separation of components in a certain dictionary.

Other methods have a different physical principle to start with. For example,

1. **Principal components analysis** (PCA), e.g., Francis & Wills (1999); Steiner et al. (2009). The functional is the variance in the space of data vectors. The first PCA component corresponds to the direction of the largest variance, the second component corresponds to the next largest variance direction etc. The degeneracy is broken by choosing the components in the order of decreasing variance.

2. **Internal linear combination** (ILC), Eriksen et al. (2004); Hinshaw et al. (2007). One uses the fact that the expected variance of a sum of two components is larger than the variance of either of the components. The CMB map is modeled as a linear combination of data vectors at different *WMAP* frequencies with the constraint on the coefficients ensuring that the CMB signal is preserved in the linear combination. The CMB map is determined by minimizing the variance across the sky of the linear combination map.

3. **Correlated component analysis** (CCA), Bedini et al. (2005). The basic observables are not the fluxes but the two-point correlators. One uses some parameterization of the source correlator matrices and the energy spectra of the sources. The parameters are estimated from the least squared difference between the observed correlators and the model correlators. The redundancy is broken by choosing a particular functional form of the spectra and/or assuming that certain components are uncorrelated.

On general grounds, the degeneracy can be broken by either putting constraints on the energy spectra or on the spatial distribution of the components. Most of the models (apart from the CCA) assume some properties of the component distributions. Usually this means either using external templates or assuming some independence of the components. In CCA one uses the functional form of the spectra to break the degeneracy. This allows one, in principle, to separate correlated components internally.

In this paper, we describe a novel method, which we call the spectral components analysis (SCA). In this method, we use the maximum likelihood for the fluxes and assume a functional form of the energy spectra to break the redundancy. In comparison with the methods that put constraints on the spatial distribution of the components, the new method allows one to study correlated components using only internal data. In contrast with the CCA method, one formulates a model directly for the fluxes rather than for the correlators: the decomposition is simpler and the calculations are easier in this case.

2. Spectral components analysis

In this section we formulate the SCA method to separate the components of a diffuse emission process. The components of the emission are modeled by linear combinations of the data in different energy bins. The degeneracy between the components is broken by assuming a functional form of the energy spectra. The linear combination parameters and the parameters of the spectra are found from a χ^2 minimization.

To start, we introduce vectors with components labeled by pixels, e.g., the data in a bin with the central energy E_α is written as $\vec{d_\alpha} = \{d^i_\alpha\}$ where index i labels the pixels. In vector notations, Equation (1) takes the form

$$\vec{d_\alpha} = \sum_\mu u^\mu_\alpha \vec{s}_\mu + \vec{r}_\alpha, \tag{6}$$

where we assume that the energy spectra u^μ_α do not depend on pixels. We assume that the number of the emission components is smaller than the number of energy bins and model the spatial distribution of the emission components as linear combinations of data vectors

$$\vec{s}_\mu = f^\beta_\mu \vec{d}_\beta. \tag{7}$$

Following the *WMAP* terminology (Hinshaw et al. 2007), we will sometimes call the emission components determined as linear combinations of the data vectors by internal linear combination (ILC) vectors. The spectra u and the linear decomposition coefficients f are found by minimizing the χ^2

$$\chi^2 = \sum_{\alpha,i} \frac{(d^i_\alpha - u^\mu_\alpha f^\beta_\mu d^i_\beta)^2}{{\sigma^i_\alpha}^2}. \tag{8}$$

The matrices u and f can be determined only up to the change of basis degeneracy, which corresponds to multiplying u from the right by a matrix A and multiplying f from the left by the inverse matrix A^{-1}

$$u \cdot f = uA \cdot A^{-1} f. \tag{9}$$

We break the degeneracy by assuming a functional form of the spectra parameterized by a set of parameters q

$$u^\mu_\alpha(q) = u^\mu(E_\alpha; q). \tag{10}$$

Before we formulate a general algorithm, let us consider the simplest case. We fix the parameters of the spectra q and assume homogeneous independent of energy uncertainty, σ. Then according to the first equation in (4)

$$f = (u^T u)^{-1} u^T. \tag{11}$$

The product of the linear combinations matrix defined by this equation and the matrix of spectra equals the unit matrix

$$\sum_\alpha f^\alpha_\mu u^\nu_\alpha = \delta^\nu_\mu, \tag{12}$$

which means that a component with spectrum u^ν contributes only to the linear combination f_ν and no other components. In general, due to different properties of noise in different energy bins, the conditions in Equation (12) will be slightly violated.

Notice, that χ^2 depends quadratically on f and non-linearly on q

$$\chi^2(f,q) = \sum_{\alpha,i} \frac{(d^i_\alpha - u^\mu_\alpha(q) f^\beta_\mu d^i_\beta)^2}{{\sigma^i_\alpha}^2}. \tag{13}$$

The algorithm has two steps:

- Find f from quadratic fitting.
- Minimize χ^2 with respect to q by a non-linear fitting.

At each step of the non-linear fitting procedure we choose a set of parameters q that describe the spectra and find the best fit f by differentiating χ^2 with respect to f and by solving the corresponding linear in f equations

$$\sum_\alpha u^\mu_\alpha u^\nu_\alpha (\vec{d}_\beta \cdot \vec{d}_\gamma)_{g_\alpha} f^\gamma_\nu = \sum_\alpha u^\mu_\alpha (\vec{d}_\alpha \cdot \vec{d}_\beta)_{g_\alpha}, \tag{14}$$

where by $(\vec{d}_\beta \cdot \vec{d}_\gamma)_{g_\alpha} = g_{\alpha\, ij} d^i_\beta d^j_\gamma$ we denote the scalar product in the space with metric g_α. In our case, it is the inverse of the standard deviation squared

$$g_{\alpha\, ij} = \frac{1}{{\sigma^i_\alpha}^2} \delta_{ij}. \tag{15}$$

In general, $g_{\alpha\, ij}$ is the inverse of the covariance matrix.

Equation (14) is a usual linear equation, the non-trivial part is that the unknowns f^γ_ν have two indices instead of one. We can solve this equation, if we choose a single index corresponding to all possible pairs of indices. The mathematical structure behind this operation is the tensor product, i.e., we represent the coefficients f^β_μ as vectors in a linear space F which is a tensor product of two linear spaces: the space D spanned by vectors $\vec{d}_\beta$ and the space S spanned by vectors $\vec{s}_\mu$, $F = D \otimes S$. Suppose that the dimensions of spaces D and S are k and m respectively, then the dimension of F is $k \cdot m$. The basis vectors in F are denoted as

$$\vec{f}_\kappa = \vec{d}_\beta \otimes \vec{s}_\mu. \tag{16}$$

Loosely speaking the operation of the tensor product is a substitution of a pair of indices $\{\beta, \mu\}$ by a single index κ. Denote $\kappa = \{\beta, \mu\}$ and $\lambda = \{\gamma, \nu\}$, then the coefficients f^γ_ν can be represented as a vector with a single index $\tilde{f}_\lambda = f^\gamma_\nu$. The right hand side of Equation (14) is a vector

$$V^\kappa = \sum_\alpha u^\mu_\alpha (\vec{d}_\alpha \cdot \vec{d}_\beta)_{g_\alpha}. \tag{17}$$

The sum on the left hand side is a matrix

$$M^{\kappa\lambda} = \sum_\alpha u^\mu_\alpha u^\nu_\alpha (\vec{d}_\beta \cdot \vec{d}_\gamma)_{g_\alpha}. \tag{18}$$

In the tensor product notations, Equation (14) takes the form

$$\sum_{\lambda} M^{\kappa\lambda}\tilde{f}_{\lambda} = V^{\kappa}. \tag{19}$$

The solution is

$$\tilde{f}_{\lambda} = \sum_{\kappa} {M^{-1}}_{\lambda\kappa} V^{\kappa}. \tag{20}$$

This equation gives the minimal value of χ^2 for a given set of parameters q. If we substitute $f^{\gamma}_{\nu}(q) = \tilde{f}_{\lambda}(q)$ in Equation (13), we get $\chi^2(q)$ as a function of q only. Now we can use a non-linear minimization procedure to find the best fit parameters q_*.

We note that the spectral components analysis allows one to find both the spectra and the spatial distributions of components with only two rather mild assumptions:

- The spatial distribution of components does not depend on the energy. The only dependence comes in the form of overall normalizations, i.e., energy spectra.
- These energy spectra have some functional form.

3. Example: *WMAP* data

In this section we use the SCA method described in the previous section to study the *WMAP* data (Jarosik et al. 2011). The *WMAP* experiment measures the fluctuations of temperature in five frequency bands with the center frequencies (in GHz)

K	Ka	Q	V	W
22.8	33.0	40.7	60.8	93.5

(21)

The data are naturally split into five vectors corresponding to the five frequency bands. In the analysis we use seven years of data (Jarosik et al. 2011) with $\sim 7'$ pixel size corresponding to the HEALPix parameter nside = 512 (Górski et al. 2005). We average the data to get $\sim 1^{\circ}$ pixel size resolution (nside = 64). We mask the Galactic plane within $\pm 30^{\circ}$, detected point sources (Gold et al. 2011), and the one hundred brightest pixels in the K band which we consider to be outliers. The remaining number of pixels is 24,228.

In the analysis, instead of using the data itself in the formulation of the models in Equation (7) we use the data vectors randomized with the detector noise. In this case, the expected χ^2 does not depend on the number of components. Without randomization, the expected χ^2 decreases as the number of components increases. For example, in the extreme case of a five-component model for the five frequency bands the χ^2 is zero.

We start the analysis by assuming only one component of the emission. The index of the energy spectrum in this case is $n = -0.13$. The reduced χ^2 is $\chi^2/dof \approx 14$, which is relatively large. We conclude that one component is not sufficient to describe the *WMAP* data.

For two components, we get the following indices of the energy spectra $n_1 = 0.04$ and $n_2 = -3.1$. The fit in this case is sufficiently good, $\chi^2/dof \approx 1.16$. Based on the spectra, we conclude that most of the contribution to these components comes from the CMB and the synchrotron emission respectively. These are the two strongest components at high latitudes.

Note, that in this approach we don't need to assume that the CMB signal exists. Instead, we derive from the data that there is a component with an index $n \approx 0$ which can be interpreted as the CMB fluctuations.

The two-component model is already very good. If we add a third component and allow the indices to be unconstrained, then the fit converges to some collinear vectors. In this case, one may choose to use some of the external knowledge. In this analysis, we assume that the third significant component is the dust emission and fix the index of the third component $n_{\rm dust} = 2$, which is a usual assumption for the spectrum of the dust emission (e.g., Hinshaw et al. 2007). Although it is true only approximately (Finkbeiner et al. 1999), the exact value is not important since this component is subdominant to the CMB and the synchrotron emissions at the *WMAP* frequencies. We also fix the index of the CMB emission $n_{\rm CMB} = 0$. The reduced χ^2 is $\chi^2/dof \approx 0.97$. The best-fit index of the second component is $n_2 = -2.6$. We interpret this component as a combination of synchrotron and free-free emission.

The components are modeled by taking linear combinations of the *WMAP* temperature maps. The best fit linear combinations are:

$$\begin{array}{rcccccc} \vec{f}_{\rm CMB} & = & (-0.37 & 0.46 & 0.67 & 0.34 & -0.11\ ) \\ \vec{f}_{\rm syn+ff} & = & (\ 0.84 & -0.18 & -0.42 & -0.28 & 0.04\ ) \\ \vec{f}_{\rm dust} & = & (\ 0.31 & -0.44 & -0.52 & -0.01 & 0.66\ ) \end{array} \quad (22)$$

Note, that the sum of the components for the CMB vector is approximately equal to 1

$$\sum_\beta f^\beta_{\rm CMB} = 0.99. \quad (23)$$

In the ILC method (Hinshaw et al. 2007) this is a constraint on the linear combinations describing the CMB map. In our approach this is a property of the linear combination of data vectors corresponding to a component with $n = 0$ spectrum. The sum of components for the foreground models is less than 0.01, i.e., the CMB signal is canceled in these models.

The *WMAP* temperature maps and the residuals after the three-component model subtraction are shown in Figures 1 and 2. The residuals are sufficiently small with little large scale structures. The SCA model of the CMB map is presented in Figure 3. It is also compared to the *WMAP* internal linear combination (ILC) model of the CMB map (Gold et al. 2011). In Figure 4 we present the map of the component corresponding to the spectral index $n = -2.6$. We compare it to the sum of the synchrotron and the free-free emission *WMAP* Markov chain Monte Carlo (MCMC) models (Gold et al. 2011). In Figure 5 we compare the SCA dust model to the *WMAP* MCMC dust model (Gold et al. 2011). Although the SCA models and the WMAP models of the emission components are derived using completely different methods, there is a fairly good agreement among the models.

In Figure 6 we compare the angular power spectra of the SCA models to the corresponding angular power spectra of the *WMAP* MCMC models. The CMB angular power spectra are very similar to each other, while the spectra for the SCA models of the Galactic emission have more power at small ℓ relative to the *WMAP* MCMC models. In Figure 7 we present the root mean squared (RMS) of the temperature maps for the three SCA components at the *WMAP* frequencies.

The SCA models generally reproduce the *WMAP* emission models. The SCA models have a higher level of random noise, which is expected for the models derived from internal data in comparison with the models that also use external data. The only assumptions that were used to derive the SCA maps of the components are the power-law energy spectra with fixed indices for the CMB and the dust emission, while the index for the third component is derived by minimizing the χ^2.

4. Conclusions

In the paper we propose a novel method to separate the components of a diffuse emission process based on an association with energy spectra. The new method is

- internal (the components are modeled as combinations of data at different energies);
- semi-blind (we assume the functional form of the spectra and fit for the parameters);
- can separate correlated components.

Using only internal data we avoid systematic uncertainties related to the usage of external data. For instance, in various diffuse emission models one needs to know the distributions of atomic, molecular, and ionized gas. The distribution of the atomic hydrogen HI (H_1) is traced by the hyperfine splitting emission at 21 cm (Kalberla et al. 2005; Kalberla & Kerp 2009). The corresponding observations have uncertainties related to the determination of the radial velocity of the gas and the spin temperature T_S (Dickey et al. 2009; Johannesson et al. 2010). The distribution of the molecular H_2

gas is traced by the emission from the CO molecules (Dame et al. 2001), some uncertainties may arise from a space dependent conversion coefficient X_{CO} (e.g., Strong et al. 2004). The distribution of ionized hydrogen HII can be modeled by studying the H-alpha emission (e.g., Finkbeiner 2003; Gaensler et al. 2008). In addition to uncertainties in the distribution of the interstellar gas, the models of cosmic ray (CR) propagation in the Galaxy (Ginzburg & Syrovatskii 1964; Strong et al. 2007; Vladimirov et al. 2011; Cholis et al. 2011) have uncertainties in the distribution of the CR sources, in the propagation parameters, and in the distribution of the interstellar radiation (e.g., Moskalenko et al. 2007). In general, the uncertainties arise from the fact that one either needs to make some assumptions when using a certain map as a template for an emission process or there is some processing involved in constructing an emission component from external data.

The advantage of internal methods is based on the following simple observation: if a physical process contributes significantly to the data, then one should be able to use the data to describe the spatial distribution and the energy spectrum of this emission process. Otherwise, if the process is insignificant for a particular observation, then one does not need to include it in the consideration.

In the SCA method, one does not need to assume a priori either the number of the emission components, their spectra, or the spatial distributions. All this information is obtained by fitting the model to the data. As a result, the SCA method is a unique tool to search for new components of emission with spectra sufficiently different from the astrophysical foregrounds. In template fitting and in galactic propagation frameworks, a new component can be observed as a residual. The spectrum and the significance of a "residual" component are usually biased. An advantage of the SCA method is that the new component and the galactic components are fitted to the data simultaneously, i.e. all components in this approach are on equal footing and the there is no reason to expect a bias for any of the components.

Another advantage of the SCA method is the possibility to separate correlated components. This property distinguishes the SCA method from other internal data analysis methods such as internal linear combinations, principal components analysis, or independent components analysis. One of the disadvantages of the SCA method is that we need to assume a homogeneous scaling of the components, i.e., that the spectra do not depend on the position in the sky. Although it is not true in general, it is usually a good first order approximation, especially if the region of interest is small.

In the paper, we use the SCA method to study seven years of the *WMAP* data. We find that at high latitudes a three component model describes the data with a sufficiently good accuracy. We interpret the three components as the CMB, the thermal dust, and a combination of the synchrotron and the free-free emission. The derived spatial distributions of the components are generally con-

sistent with the corresponding spatial distributions provided by the *WMAP* collaboration.

The SCA method is sufficiently universal. It can be applied to other types of diffuse emission data, such as infrared, optical, x-ray, and gamma-ray. One can also use any representation allowing linear space interpretation, e.g., coordinates space, spherical harmonics, wavelets. The method has a straightforward generalization to the case where some of the components are modeled by external templates (Appendix A).

Acknowledgments.

The author is thankful to Elliott Bloom, Anna Franckowiak, Daniel Grin, David W. Hogg, Igor Moskalenko, Dmitry Prokhorov, Kendrick Smith, David Spergel for stimulating discussions. This work was supported in part by the US Department of Energy contract to SLAC no. DE-AC02-76SF0051. The data analysis have been done using the HEALPix package (Górski et al. 2005).

A. General case

In this appendix we present an algorithm that includes some number of external templates together with components modeled as internal linear combinations of data vectors (ILC vectors). In the most general form, we will not put any constraints on the spectra for some of the ILC vectors. As we have discussed in the introduction, if there is more than one ILC vector for which we do not put constraints on the spectra, then one needs to break the change of the basis degeneracy (Equation (9)). In the algorithm described below we break this degeneracy by choosing some sub-matrix of the linear combinations matrix f (Equation (7)) to be the unit matrix. Components for which we do not put constraints on the spectra have an independent scale factor in every energy bin. The process of subtracting these components is called marginalization.

In the general SCA algorithm we can consider the following models for the components together with the notations (bold font letters denote vectors, while usual font letters denote numbers):

1. Components with known spatial distributions (templates) that we marginalize over, $\tilde{w}_\alpha^\nu \tilde{\mathbf{e}}_\nu$ (no summation).

2. Components with known spatial distribution for which we assume a parametric form of the spectra, $w(p)_\alpha^\nu \mathbf{e}_\nu$ (no summation).

3. Components with unknown spatial distributions (modeled as ILC vectors) that we marginalize over, $\tilde{u}_\alpha^\mu \tilde{f}_\mu^\beta \mathbf{v}_\beta$ (no summation over μ).

4. Components with unknown spatial distributions (modeled as ILC vectors) for which we assume a parametric form of the spectra, $u^\mu_\alpha(q) f^\beta_\mu \mathbf{v}_\beta$ (no summation over μ).

The general χ^2 has the form

$$\chi^2 = \sum_\alpha \left| \mathbf{d}_\alpha - u^\mu_\alpha(q) f^\beta_\mu \mathbf{v}_\beta - \tilde{u}^\mu_\alpha \tilde{f}^\beta_\mu \mathbf{v}_\beta - w(p)^\nu_\alpha \mathbf{e}_\nu - \tilde{w}^\nu_\alpha \tilde{\mathbf{e}}_\nu \right|^2_{g_\alpha}, \tag{A1}$$

where $|\mathbf{v}|^2_{g_\alpha} \equiv (\mathbf{v}, \mathbf{v})_{g_\alpha}$ is the norm in space with metric g_α. The metric is given by the inverse of the standard deviation squared, or, in general, by the inverse of the covariance matrix.

The algorithm has the following steps:

1. Marginalize over $\tilde{w}^\nu_\alpha$. This is equivalent to projecting the vectors $\mathbf{d}_\alpha$, $\mathbf{v}_\beta$ and $\mathbf{e}_\nu$ onto the space perpendicular to the space spanned by $\tilde{\mathbf{e}}_\nu$. The residual χ^2 has the form

$$\chi^2 = \sum_\alpha \left| \mathbf{d}^\perp_\alpha - u^\mu_\alpha(q) f^\beta_\mu \mathbf{v}^\perp_\beta - \tilde{u}^\mu_\alpha \tilde{f}^\beta_\mu \mathbf{v}^\perp_\beta - w(p)^\nu_\alpha \mathbf{e}^\perp_\nu \right|^2_{g_\alpha}. \tag{A2}$$

2. Choose parameters p. Subtract $w(p)^\nu_\alpha \mathbf{e}^\perp_\nu$ from $\mathbf{d}^\perp_\alpha$ and $\mathbf{v}^\perp_\alpha$.

3. Choose $\tilde{f}^\beta_\mu$ and marginalize over the parameters $\tilde{u}^\mu_\alpha$ by projecting $\mathbf{d}^\perp_\alpha$ on $\tilde{f}^\beta_\mu \mathbf{v}^\perp_\beta$.

4. Choose parameters q and marginalize over f^β_μ (Equations (13) to (20)).

5. Repeat steps 2, 3, 4, (and 5) to find the best fit non-linear parameters p_*, q_*, and $\tilde{f}_*$.

Given the best fit parameters p_*, q_*, and $\tilde{f}_*$, the maps of the components are reconstructed as follows:

1. Components with templates and constrained energy spectra are given by $w(p_*)^\nu_\alpha \mathbf{e}_\nu$ (no summation). In order to get the other maps, we subtract these components from the data and define $\mathbf{d}'_\alpha = \mathbf{d}_\alpha - w(p_*)^\nu_\alpha \mathbf{e}_\nu$.

2. The maps for the templates with marginal spectra are obtained by marginalization of $\mathbf{d}'_\alpha$ with respect to the templates $\tilde{\mathbf{e}}_\nu$. The scaling coefficients are obtained from minimizing the χ^2

$$\tilde{w}^\nu_\alpha = \sum_\mu (\mathbf{d}'_\alpha, \tilde{\mathbf{e}}_\mu)_{g_\alpha} E^{-1\,\mu\nu}, \tag{A3}$$

where $E^{-1\,\mu\nu}$ is the inverse of the matrix of scalar products $E_{\mu\nu} = (\tilde{\mathbf{e}}_\mu, \tilde{\mathbf{e}}_\nu)_{g_\alpha}$. The maps are given by $\tilde{w}^\nu_\alpha \tilde{\mathbf{e}}_\nu$ (no summation).

3. The ILC components are obtained by linear combinations of model vectors $\mathbf{v}'_\alpha$ derived from the data vectors $\mathbf{d}'_\alpha$. If $\mathbf{v}_\alpha = \mathbf{d}_\alpha$, then $\mathbf{v}'_\alpha = \mathbf{v}_\alpha - w(p_*)^\nu_\alpha \mathbf{e}_\nu$. We also marginalize over the templates $\tilde{\mathbf{e}}_\nu$ with unconstrained spectra by projecting $\mathbf{d}'_\alpha$ and $\mathbf{v}'_\alpha$ on $\tilde{\mathbf{e}}_\nu$: $\mathbf{d}'_\alpha \rightarrow {\mathbf{d}'_\alpha}^\perp$ and $\mathbf{v}'_\alpha \rightarrow {\mathbf{v}'_\alpha}^\perp$. The maps for the ILC components with marginal spectra are given by $\tilde{u}^\mu_\alpha \tilde{f}^\beta_{*\mu} {\mathbf{v}'}^\perp_\beta$ (no summation over μ). The overall coefficients $\tilde{u}^\mu_\alpha$ are found similarly to $\tilde{w}^\nu_\alpha$ by minimizing the residual χ^2 that depends on ${\mathbf{d}'_\alpha}^\perp$

$$\tilde{u}^\nu_\alpha = \sum_\mu ({\mathbf{d}'}^\perp_\alpha, \tilde{\mathbf{f}}_\mu)_{g_\alpha} F^{-1\,\mu\nu}, \tag{A4}$$

where $\tilde{\mathbf{f}}_\mu = \tilde{f}^\beta_{*\mu} {\mathbf{v}'}^\perp_\beta$ are the ILC combinations and $F^{-1\,\mu\nu}$ is the inverse of the matrix of scalar products $F_{\mu\nu} = (\tilde{\mathbf{f}}_\mu, \tilde{\mathbf{f}}_\nu)_{g_\alpha}$.

4. The maps for the ILC components with functional forms of the spectra are equal to $u^\mu_\alpha(q_*) f^\beta_\mu {\mathbf{v}'}^\perp_\beta$ (no summation over μ). The coefficients f^β_μ are found from Equations (14) to (20), where we substitute $u^\mu_\alpha \rightarrow u^\mu_\alpha(q_*)$ and $\mathbf{d}_\alpha \rightarrow {\mathbf{d}'}^\perp_\alpha$.

The general SCA method gives simultaneously the spectra and the maps of all four types of models for the emission components: templates with and without assuming a functional form of the spectra and ILC combinations with and without assuming a functional form of the spectra.

We note that in some cases instead of a non-linear fitting procedure one can use a convergent iterative process (e.g. Bobin et al. 2007; Tsalmantza & Hogg 2012) to define the spectra and the spatial distributions of the emission components.

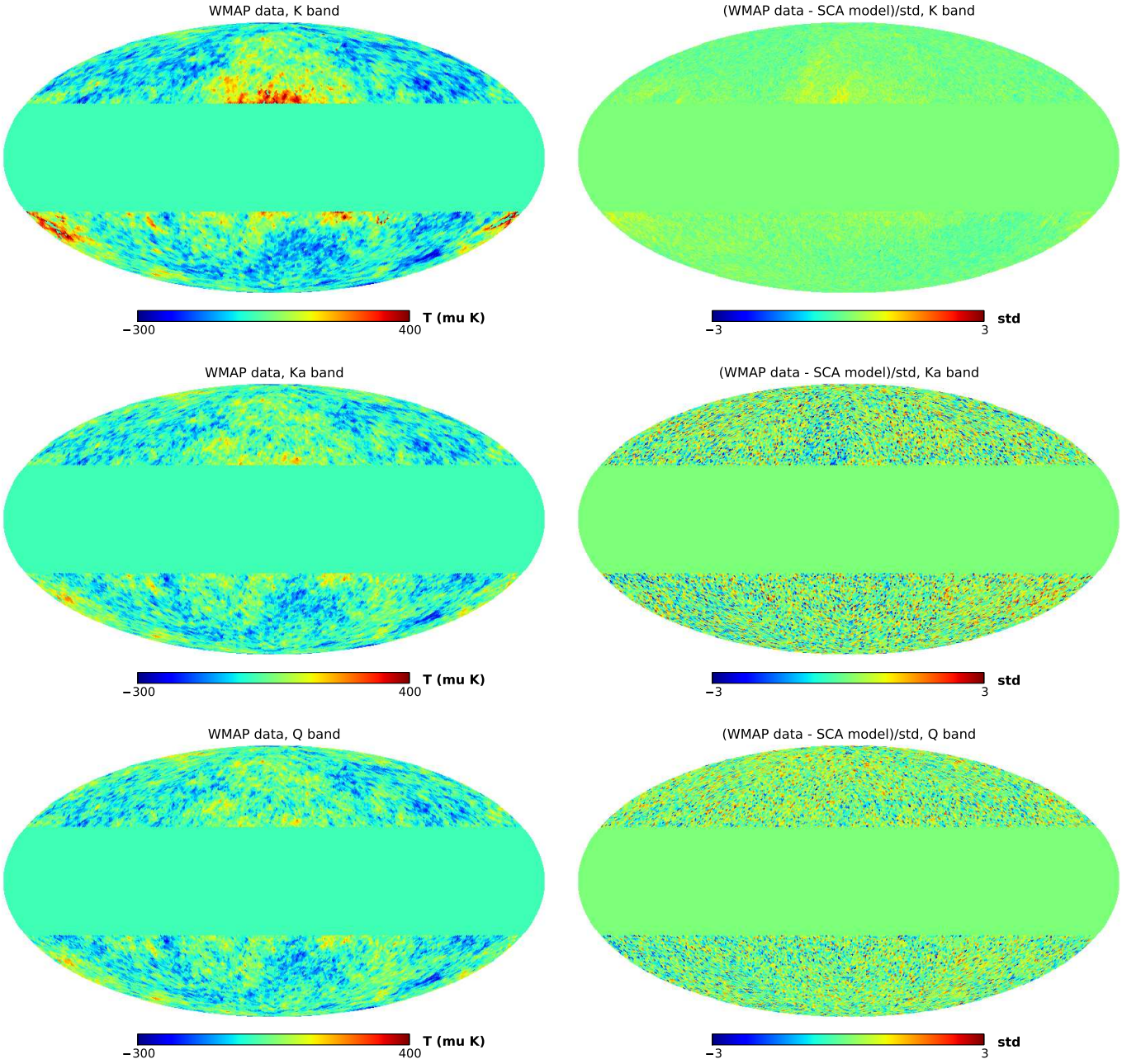

Fig. 1.— Left: maps of seven years of *WMAP* data in K, Ka, and Q frequency bands (Jarosik et al. 2011). Right: residuals after subtracting the SCA models (Figures 3, 4, and 5) divided by the instrumental noise.

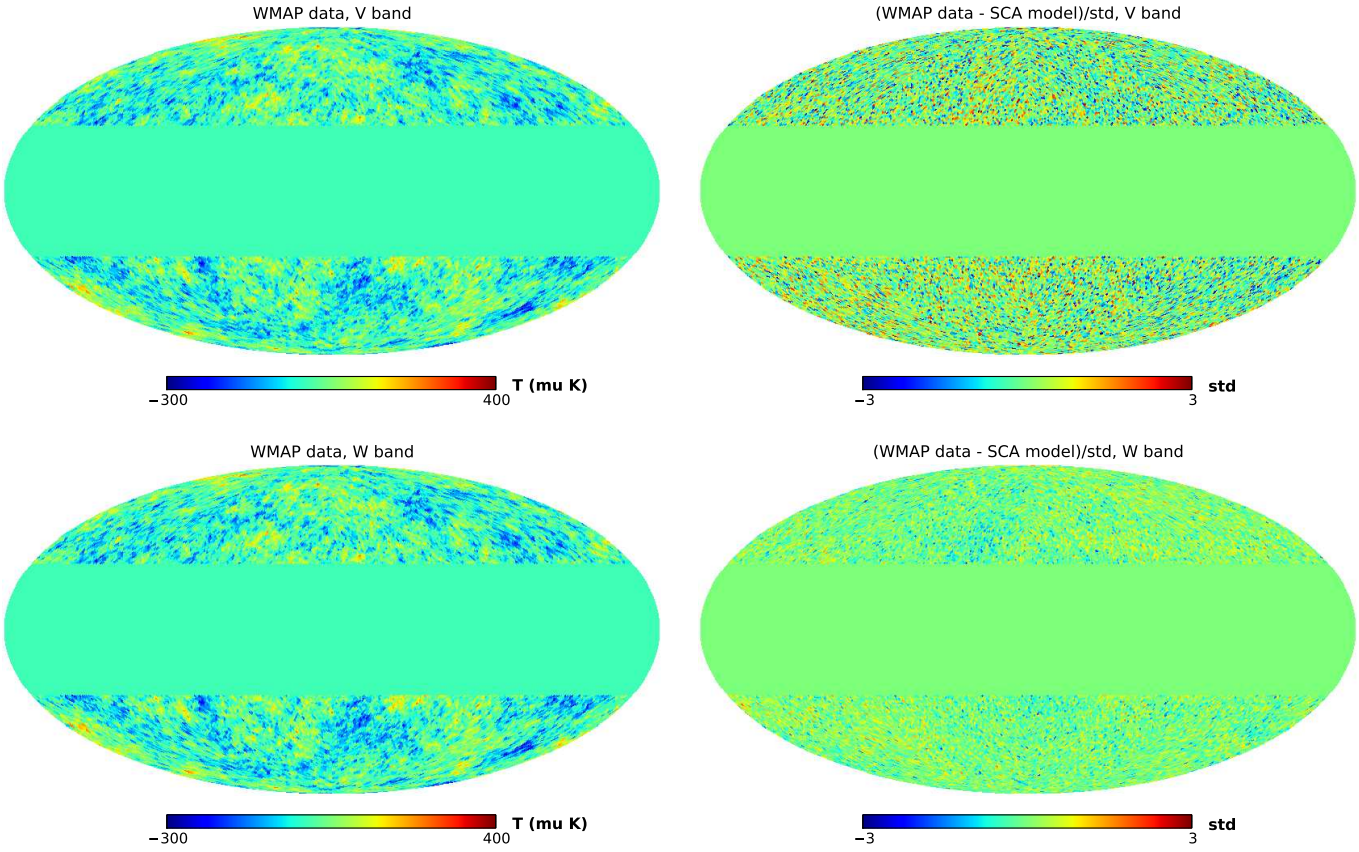

Fig. 2.— Same as in Figure 1 for V and W frequency bands.

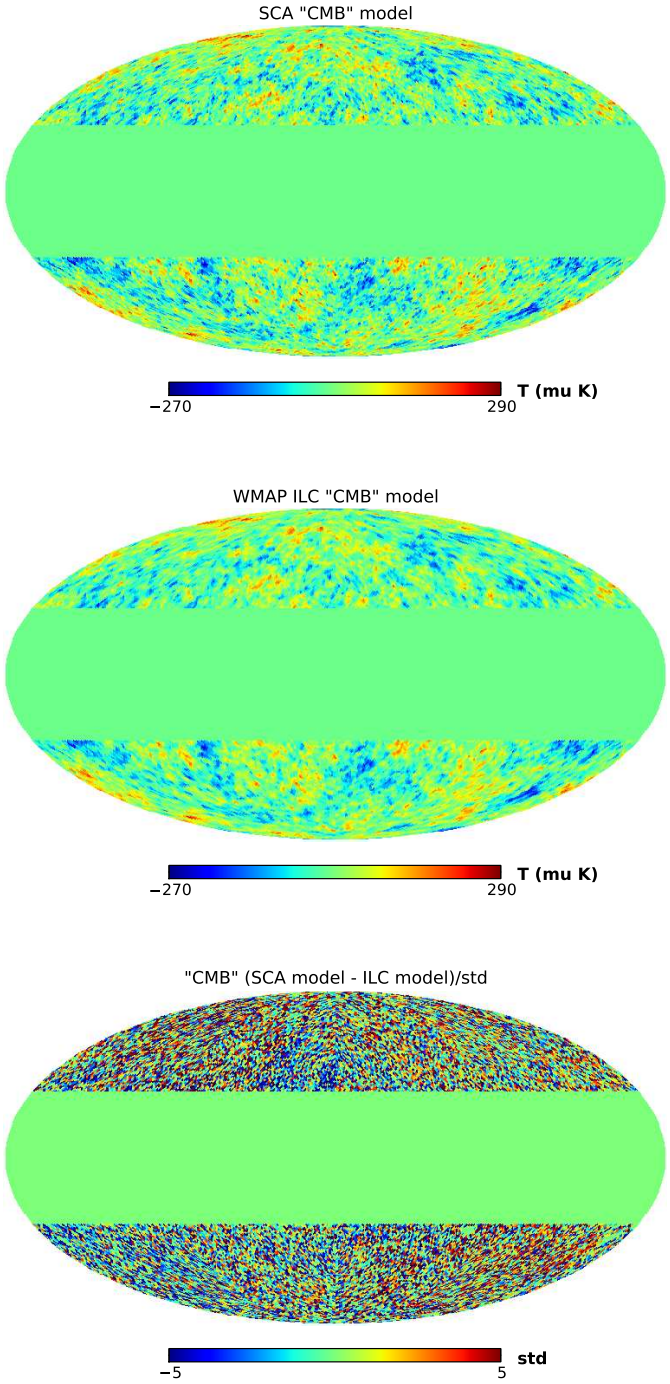

Fig. 3.— Top: the SCA model of the CMB fluctuations. Middle: the *WMAP* ILC model of the CMB (Gold et al. 2011). Bottom: the difference between the SCA model and the *WMAP* ILC model of the CMB divided by the instrumental noise. The SCA CMB model is very similar to the *WMAP* CMB model, but has a little higher random noise level.

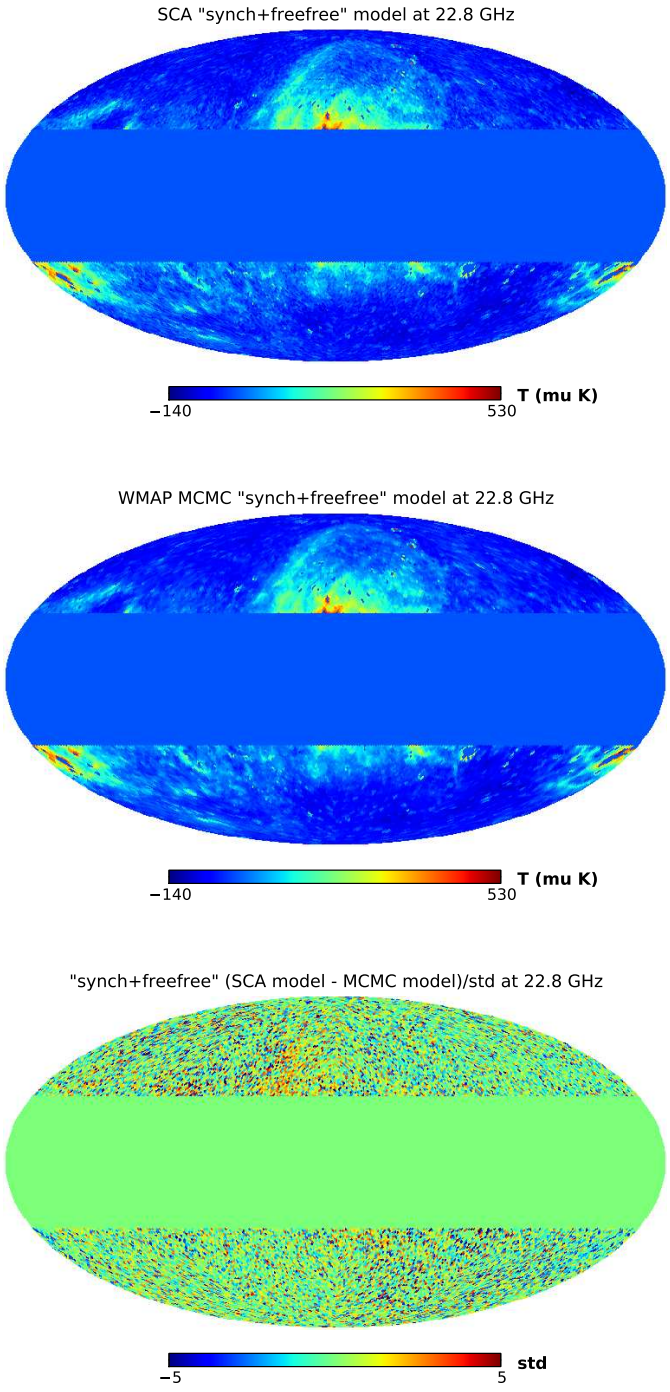

Fig. 4.— Top: the SCA model of the synchrotron+free-free emission. Middle: the *WMAP* MCMC model of the synchrotron+free-free emission (Gold et al. 2011). Bottom: the difference between the SCA model and the *WMAP* MCMC model divided by the instrumental noise.

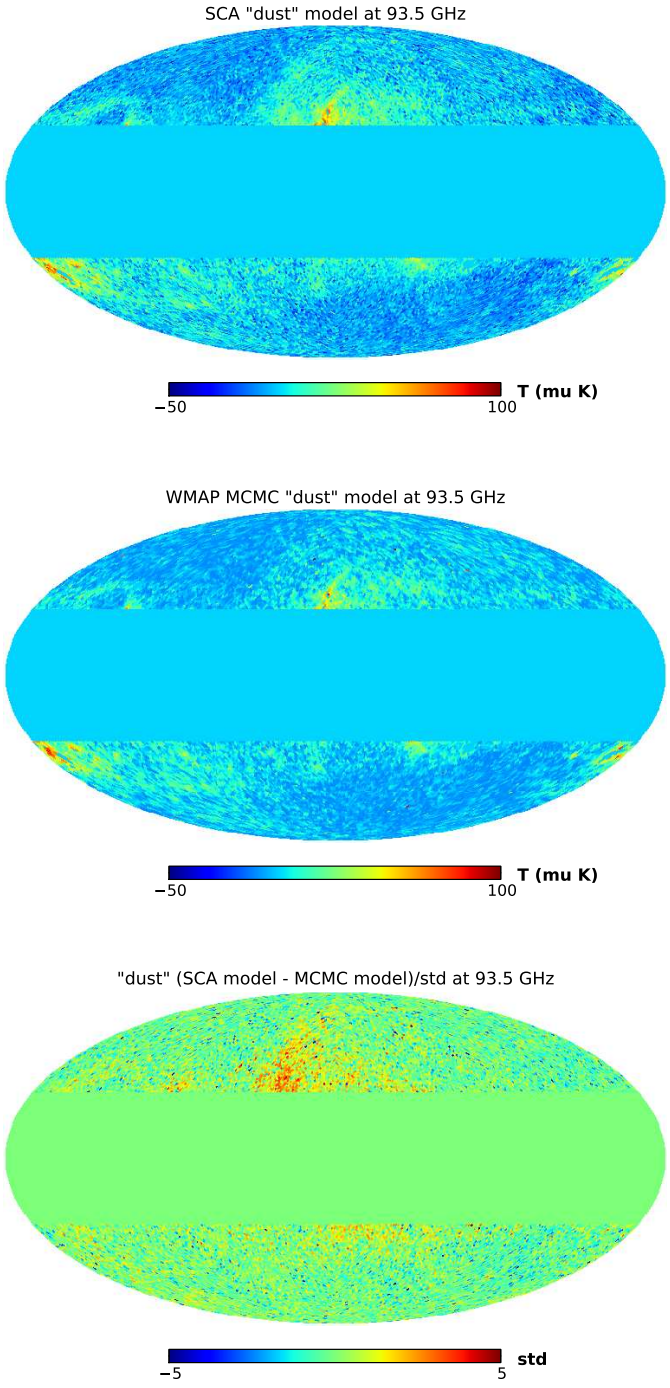

Fig. 5.— Top: the SCA model of the thermal dust emission. Middle: the *WMAP* MCMC model of the thermal dust emission (Gold et al. 2011). Bottom: the difference between the SCA model and the *WMAP* MCMC model divided by the instrumental noise.

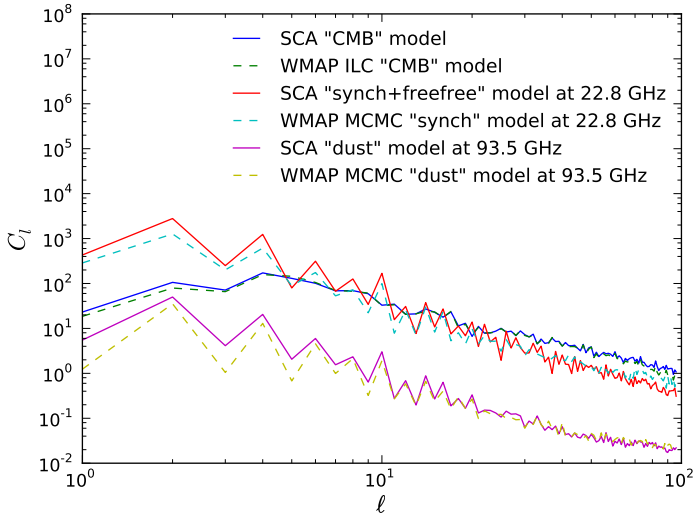

Fig. 6.— Angular power spectra for the SCA models and the *WMAP* models presented in Figures 3, 4, and 5. The angular power spectra of the CMB are very similar to each other. The SCA models of the synchrotron+free-free and the dust emission have a little more power at small ℓ than the *WMAP* MCMC models. Large up and down fluctuations of the foreground emission C_ℓ's is due to the symmetry properties of the mask. In calculating the C_ℓ's, we divide by the fraction of the sky that is unmasked.

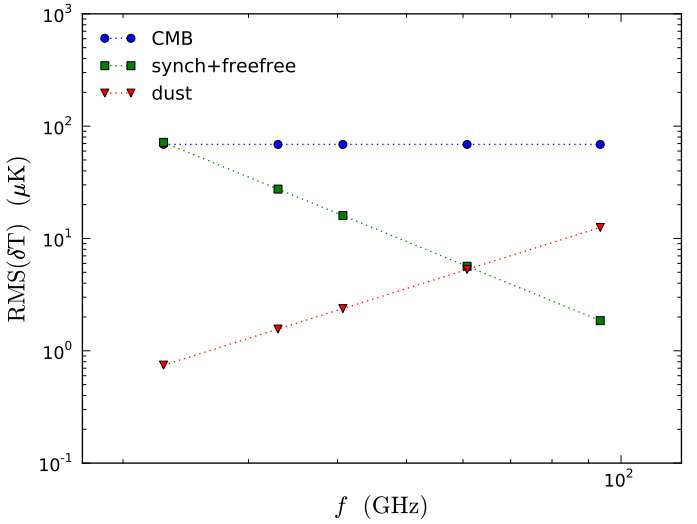

Fig. 7.— Root mean squared of the temperature maps for the SCA models (Figures 3, 4, and 5) at different frequencies. We assume power-law energy spectra. The indices for the CMB and the thermal dust emission are fixed, $n_{\rm CMB} = 0$ and $n_{\rm dust} = 2$. The index of the third component and the normalizations are found from fitting the models to the data in Section 3. The index for the third components is found to be $n_{\rm s+ff} = -2.6$. This component is interpreted as a combination of the synchrotron and the free-free emissions.

REFERENCES

Bedini, L., Herranz, D., Salerno, E., Baccigalupi, C., Kuruouglu, E. E., & Tonazzini, A. 2005, EURASIP Journal on Applied Signal Processing, 2005, 2400

Bobin, J., Starck, J.-L., Fadili, J., & Moudden, Y. 2007, IEEE Transactions on Image Processing, 16, 2662

Cholis, I., Tavakoli, M., Evoli, C., Maccione, L., & Ullio, P. 2011, ArXiv:1106.5073

Dame, T. M., Hartmann, D., & Thaddeus, P. 2001, ApJ, 547, 792

Dickey, J. M., Strasser, S., Gaensler, B. M., Haverkorn, M., Kavars, D., McClure-Griffiths, N. M., Stil, J., & Taylor, A. R. 2009, ApJ, 693, 1250

Eriksen, H. K., Banday, A. J., Górski, K. M., & Lilje, P. B. 2004, ApJ, 612, 633

Finkbeiner, D. P. 2003, ApJS, 146, 407

Finkbeiner, D. P., Davis, M., & Schlegel, D. J. 1999, ApJ, 524, 867

Francis, P. J., & Wills, B. J. 1999, in Astronomical Society of the Pacific Conference Series, Vol. 162, Quasars and Cosmology, ed. G. Ferland & J. Baldwin, 363

Gaensler, B. M., Madsen, G. J., Chatterjee, S., & Mao, S. A. 2008, PASA, 25, 184

Ginzburg, V. L., & Syrovatskii, S. I. 1964, The Origin of Cosmic Rays (Pergamon, Oxford)

Gold, B., et al. 2011, ApJS, 192, 15

Górski, K. M., Hivon, E., Banday, A. J., Wandelt, B. D., Hansen, F. K., Reinecke, M., & Bartelmann, M. 2005, ApJ, 622, 759

Hinshaw, G., et al. 2007, ApJS, 170, 288

Hobson, M. P., Jones, A. W., Lasenby, A. N., & Bouchet, F. R. 1998, MNRAS, 300, 1

Hyvarinen, A. 1999, IEEE Signal Processing Letters, 6, 145

Jarosik, N., et al. 2011, ApJS, 192, 14

Johannesson, G., Moskalenko, I., Digel, S., & for the Fermi LAT Collaboration. 2010, ArXiv:1002.0081

Kalberla, P. M. W., Burton, W. B., Hartmann, D., Arnal, E. M., Bajaja, E., Morras, R., & Pöppel, W. G. L. 2005, A&A, 440, 775

Kalberla, P. M. W., & Kerp, J. 2009, ARA&A, 47, 27

Leach, S. M., et al. 2008, A&A, 491, 597

Maino, D., et al. 2002, MNRAS, 334, 53

Moskalenko, I. V., Digel, S. W., Porter, T. A., Reimer, O., & Strong, A. W. 2007, Nuclear Physics B Proceedings Supplements, 173, 44

Steiner, J. E., Menezes, R. B., Ricci, T. V., & Oliveira, A. S. 2009, MNRAS, 395, 64

Stolyarov, V., Hobson, M. P., Ashdown, M. A. J., & Lasenby, A. N. 2002, MNRAS, 336, 97

Strong, A. W., Moskalenko, I. V., & Ptuskin, V. S. 2007, Annual Review of Nuclear and Particle Science, 57, 285

Strong, A. W., Moskalenko, I. V., Reimer, O., Digel, S., & Diehl, R. 2004, A&A, 422, L47

Tsalmantza, P., & Hogg, D. W. 2012, ArXiv:1201.3370

Vladimirov, A. E., et al. 2011, Computer Physics Communications, 182, 1156