Foreground removal from *Planck Sky Model* temperature maps using an MLP neural network

H. U. Nørgaard-Nielsen[1,⋆] and **K. Hebert**[2]

[1] National Space Institute (DTU Space),
Technical University of Denmark
Juliane Mariesvej 30, DK-2100 Copenhagen, Denmark

[2] Department of Computer Science,
Copenhagen University
Universitetsparken 1, DK-2100 Copenhagen, Denmark

The dates of receipt and acceptance should be inserted later

Key words Cosmology: Cosmic Microwave Background – method: data analysis

Unfortunately, the Cosmic Microwave Background (CMB) radiation is contaminated by emission originating in the Milky Way (synchrotron, free - free and dust emission). Since the cosmological information is statistically in nature, it is essential to remove this foreground emission and leave the CMB with no systematic errors. To demonstrate the feasibility of a simple multilayer perceptron (MLP) neural network for extracting the CMB temperature signal, we have analyzed a specific data set, namely the Planck Sky Model maps, developed for evaluation of different component separation methods before including them in the Planck data analysis pipeline. It is found that a MLP neural network can provide a CMB map of about 80 per cent of the sky to a very high degree uncorrelated with the foreground components. Also the derived power spectrum shows little evidence for systematic errors.

© 0000 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

1 Introduction

The Cosmic Microwave Background (CMB) is one of the key observations behind all contemporary cosmology research. Since its discovery in 1965 by Penzias and Wilson (1965), tremendous efforts to improve the accuracy of the CMB maps have been performed. The latest major step forward is the ongoing *Wilkinson Microwave Anisotropy Probe* (WMAP Bennet et al. 2003a), and with the launch of the ESA Planck mission in Spring 2009, substantial improvements in the angular resolution and sensitivity of the CMB maps, in both temperature and polarization, are expected.

From the CMB temperature and polarization angular power spectra unique information about the earliest phases of the evolution of the Universe can be extracted. Unfortunately, it is not possible to observe a pure CMB signal. It will always be mixed up with unrelated, and in the cosmological context, unwanted, signals such as emission from the Milky Way, and signals of local origin, e.g. from the detector systems. To ensure that the cosmological information is unbiased, it is essential to removeemission from all other sources before the CMB maps are extracted.

A lot of effort has been put into developing algorithms for foreground removal from CMB data in order to ensure minimum systematic residuals in the derived angular power spectra. A comprehensive review has been given by Delabrouille and Cardoso (2007). Of course, a method guaranteeing both a systematic-free CMB map and power spectrum is most desirable. For Planck this is a requirement, since one of the key scientific objectives will be the search for non-Gaussian cosmological features in the CMB map.

In deriving the power spectrum of the CMB map a lot of averaging of the signals of individual sky pixels is performed, implying that to obtain unbiased cosmological information, the crucial issue is not so much to minimize the random errors per sky pixel, but to minimize the systematic errors in the CMB maps.

With the COBE satellite it was clearly demonstrated that the CMB radiation has a virtually perfect black body spectrum, Mather et al. 1999. Fortunately, all known non - cosmological signals have a clear non-thermal spectral behaviour. Depending on the observational accuracy and frequency coverage, it should thus be possible to split the raw microwave signals into their different components.

As an important part of the preparation for the Planck mission, a detailed comparison of 8 different methods has been performed using the same simulated maps, called the Planck Sky Model (PSM). Planck Working Group 2, led by Drs. J. Delabrouille and G. de Zotti, has been responsible for producing these maps and has performed the comparison (called WG2 Challenge-2). Detailed discussion of the results have been published by Leach et al., (2008).

The Internal Linear Combination method (ILC) for extracting the CMB was introduced by Bennett el al. (2003b)

⋆ Corresponding author: e-mail: hunn@space.dtu.dk

© 0000 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

in their analysis of the WMAP 1 yr data. It is expected that a simple linear combination will never give a systematic-free CMB map. The foreground spectra in the WMAP frequency range are known to be power laws, or significant modifications to power laws, with details depending on the position on the sky. Furthermore, the relative contribution of the different foregrounds also vary over the sky. To compensate for these basic problems the linear solutions are derived in small areas on the sky, and these submaps are then combined to make a full sky map. The WMAP team has warned against using the ILC map for cosmological investigations.

By exploiting the non-linear feature of neural networks (NN), Nørgaard-Nielsen and Jørgensen (2008, hereafter NNJ), have shown that with observational errors as expected from the Planck satellite, and with spectral behaviour as described in the literature for the diffuse galactic components (synchrotron, free-free, thermal dust), it is possible to extract the CMB signal with negligible systematic errors in the derived CMB temperatures.

In order to extend this conclusion from a sky pixel level to a sky map level, in this paper we will extend the analysis in NNJ to investigate a specific data set, namely the PSM temperature maps, the same data used in the WG2 Challenge-2. Since we are using basically the same method as in NNJ, the basic description and neural network references can be found there.

2 The ESA CMB mission: Planck

One of the main drivers of the design of the Planck payload has been to ensure proper removal of non-CMB footprints from the sky maps by observing over a wide range of frequencies, much wider than the peak of the CMB signal.

Planck contains 2 detector systems: the Low Frequency Instrument (LFI) based on HEMT technology (Principal Investigator N. Mandolesi), and the High Frequency Instrument (HFI) based on bolometers (Principal Investigator J. L. Puget). The reflector system is provided by ESA and a Danish consortium (Principal Investigator H. U. Nørgaard - Nielsen). The Planck frequency bands are 30GHz, 44GHz and 70GHz (LFI) and 100GHz, 143GHz, 217GHz, 353GHz, 545GHz and 857 GHz (HFI), much wider than the CMB peak at around 200 GHz. All LFI detectors are polarization sensitive, while HFI has polarization-sensitive bolometers for all frequencies less than 545 GHz. The sensitivities of the Planck detector systems used in PSM have been taken from the Planck 'Blue Book' (ESA-SCI(2005)1).

3 The Planck WG2 Challenge-2 Sky Maps

The main purpose of the Planck WG2 Challenge - 2 was to test a number of component separation methods with a data set as expected from the Planck satellite (Leach et al. 2008). The data used in this paper has been taken from the Planck WG2 Challlenge-2 ftp area[1] For each Planck frequency separate files are given for each diffuse component (CMB, synchrotron, free-free, thermal dust, spinning dust). In the WG2 Challenge-2 the spinning dust was not included. Since it has been shown (e.g. Davies et al., 2006) that spinning dust will give a major contribution to the Planck signal at the low frequencies, we have included this component in the current investigation.

The point spread function of each frequency is assumed to be Gaussian, with a FWHM as defined in the Planck 'Blue Book'.

For each frequency a noise map is also given. It has been assumed that the noise is Gaussian and scaled to the expected number of hits per sky pixel. Due to the Planck scanning strategy and the arrangement of the detectors in the focal plane, these noise maps for the different frequencies are very similar, except for a scale factor.

So, for each frequency we construct a map by adding the PSM maps: CMB, synchrotron, free-free, thermal dust, spinning dust and noise, keeping the original resolution (Healpix nside = 2048).

3.1 The PSM modeling of the diffuse foregrounds

In contrast to the real world, it is possible for the WG2 Challenge-2 to examine noise-free maps for each diffuse component and for each Planck frequency, thereby independently establishing the spectral behaviour as a function of frequency.

3.1.1 The PSM synchrotron, free-free and dust spectral modeling

For each component spectral slopes has been calculated, as defined as (for the frequency range f1 to f2)

$$\mathrm{slope\ f1-f2} = \log(\mathrm{flux(f1)/flux(f2)})/\log(\mathrm{f1/f2}) \quad (1)$$

In Fig. 1 the PSM synchrotron slopes in the bands 44 GHz - 70 GHz, 70 GHz - 100 GHz etc., are plotted as a function of the 30 GHz - 44 GHz slope. It is seen that the slopes are quite similar.

In Fig. 2 The PSM free-free slopes in the bands 44 GHz - 70 GHz, 70 GHz - 100 GHz etc., are plotted as a function of the 30 GHz - 44 GHz slope. It is seen that the slopes flatten out to a constant, -0.14, at high frequencies. It is difficult to judge how representative of the real sky this spectral behaviour is, but the important point for this paper is that we have used exactly the same data used in the WG2 Challenge-2, so the results can be directly compared with the results given in Leach et al., (2008).

In Fig. 3 the PSM thermal dust slopes in the bands 100 GHz - 143 GHz, 143 GHz - 217 GHz, 217 GHz - 353 GHz, 353 GHz - 545 GHz, are plotted as a function of the 545 GHz - 857 GHz slope. The slopes 30 GHz - 44 GHz and 44 GHz - 70 GHz are constant at 3.31 and 3.15 respectively.

[1] ftp://planck-wg2.planck.fr/Challenge-2/PSM-maps_v0/

© 0000 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Fig. 1 The PSM synchrotron slopes (44 GHz–70 GHz, 70 GHz–100 GHz, 100 GHz–143 GHz etc) as a function of the 30 GHz–44 GHz slope. It is seen that the slopes are close to being independent of frequency

Fig. 2 The PSM free-free slopes as a function of the 30 GHz–44 GHz slope. At frequencies higher than 217 GHz the slope is constant at -0.14

In our modeling for the neural network we have used the linear relations given in Figs. 1, 2 and 3.

Concerning the spectral behaviour of the spinning dust, the PSM slopes are plotted in Fig. 4 as a function of the 30 GHz–44 GHz slope. For frequencies higher than 143 GHz the spinning dust flux is negligible. It can be seen that the spectral behaviour is rather complex, and in our neural network modeling we have interpolated from tables derived from a simple mean curve for each frequency band.

From the large scatter in Figs. 1, 2 and 3 it is evident that our mathematical models of the diffuse components are much simpler than what has been used in setting up the PSM. But, as shown in Section 7, despite this simplification the neural network is able to derived CMB temperatures with negligible systematic errors.

3.1.2 Flux ranges at 100 GHz

In Fig. 5 the sky fraction of the PSM component maps at 100 GHz is given as a function of the flux limit. Following

Fig. 3 The PSM thermal dust slopes as a function of the 545 GHz–857 GHz slope

Fig. 4 The PSM spinning dust slopes as a function of the 30 GHz–44 GHz slope. For frequencies larger than 143 GHz the flux is negligible

NNJ, it is expected that the neural network will not provide a systematic-free CMB map for the full sky. As seen below, the neural network solution covers about 80 per cent of the sky, with flux limits for the mask of 3, 3 and 15 for synchrotron, free-free and thermal dust, respectively, (units: 10^{-20} erg/s/cm^2/Hz/sr).

As emphasized by Eriksen et al. (2006), and others, there is a basic limitation in how sophisticated the spectral models of the diffuse components can be, due simply to the small number of observed frequencies. This is true both for Planck and all the CMB missions planned for the future. Therefore, the number of independent variables in the models of the diffuse components should be minimized as much as possible.

In Fig. 6 the thermal dust flux at 100 GHz is plotted against the synchrotron flux. It is evident that they are well correlated (correlation coefficient = 0.81). Therefore the synchrotron 100 GHz flux is calculated in our neural network as

$$\mathrm{f}_{100\,\mathrm{GHz}}(\mathrm{synch}) = (0.115 + 0.075 * \mathrm{r}_{\mathrm{syn}}) * \mathrm{f}_{100\,\mathrm{GHz}}(\mathrm{dust}) \quad (2)$$

where r_{syn} is randomly chosen from 0.00 to 1.00.

© 0000 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Fig. 5 The PSM sky fraction as a function of flux at 100 GHz (units: 10^{-20} erg/s/cm^2/Hz/sr)for the synchrotron, free-free and thermal dust emission

Fig. 6 The correlation between the PSM thermal dust and the PSM synchrotron emission at 100 GHz

Since the discovery of what is now called 'spinning dust', it is clear that this emission is well correlated to the thermal dust emission e.g Leitch et al. (1997). In Fig. 7 the PSM thermal dust flux at 70 GHz is plotted against the PSM spinning dust 70 GHz flux. It is seen that the correlation is tight (correlation coefficient = 0.91).

Therefore, the spinning dust 70 GHz flux is calculated in the neural network modeling as

$$\mathrm{f}_{70\,\mathrm{GHz}}(\mathrm{spindust}) = (0.2 + 0.3 * \mathrm{r}_{\mathrm{spin}}) * \mathrm{f}_{70\,\mathrm{GHz}}(\mathrm{dust}) \quad (3)$$

where r_{spin} is randomly chosen from 0.00 to 1.00.

Altogether, the spectral models used in the neural networks have the following independent parameters:

- the synchrotron 30 GHz–44 GHz slope
- the free-free flux at 100 GHz
- the free-free 30 GHz–44 GHz slope
- the thermal dust flux at 100 GHz
- the thermal dust 545 GHz–857 GHz slope
- the spinning dust 30 GHz–44 GHz slope
- the CMB temperature

Fig. 7 The correlation between the PSM thermal and spinning dust emission at 70 GHz

Furthermore, the synchrotron 100 GHz flux and the spinning dust flux at 70 GHz are not independent parameters but calculated from the r_{syn} and r_{spin} parameters given in Eqs. 2 and 3.

Noise was created for each Planck frequency by the IDL *randomn* function, random numbers with a zero mean and standard deviation = 1.0, and scaled to cover the dynamic ranges of the PSM noise maps.

4 The Internal Linear Combination (ILC) method

By a linear combination of the 5 WMAP frequency maps, the ILC method (Bennett et al. 2003b), tries to eliminate foreground signals and extract the cosmological signal. With the known complex spectral shapes of the foreground signals, this assumption only makes sense on small submaps of the sky. The WMAP team divided the sky into 12 regions, of which 11 cover the galactic plane (see Hinshaw et al. 2007 Fig. 8). Due to the basic limitations of the method, the WMAP team warned against using the ILC map for cosmological investigations.

Eriksen et al. (2004) have re-analysed the WMAP ILC map and improved the variance of the map significantly by introducing a Lagrange multiplier fitting algorithm. Using detailed Monte Carlo simulations, they investigated the limitations of the method and emphasized that great care needs to be taken both in its implementation and in understanding the influence of the residual foregrounds on the cosmological results.

5 Brief description of the neural network concept

NNJ gives an introduction to the concept of neural networks, especially the Multi Layer Perceptron (MLP) network, which is one of the simplest networks, and also the most used. Basic references to the concept of neural networks can be

© 0000 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Fig. 8 A schematic of a **M**ulti**L**ayer **P**erceptron network with 1 hidden layer

found in NNJ. An excellent introduction to neural networks can be found in Bishop (1995).

Briefly, an MLP consists of a network of units called neurons, as illustrated in Fig. 8. Each unit is shown as a circle and the lines connecting them are known as weights. The network can be understood as an analytical mapping between a set of input variables x_m $(m = 1, ..., M)$ and a set of output variables y_n $(n = 1, ..., N)$. The input variables are applied to the M input units on the left of the figure: M = 4 and N = 2 in the example shown. These variables are multiplied by a matrix of parameters w_{lm} $(l = 1, ..., L)$, $(m = 1, ..., M)$ corresponding to the first layer of weighted links. Here L is the number of neurons in the middle (hidden) layer: L = 3 in the shown example. This results in a vector of inputs to the units in the hidden layer. Each component of this vector is then transformed by a non-linear function F, so we have

$$z_l \;=\; F\left(\sum_{m=1}^{M} w_{lm}x_m \;+\Theta_l\right) \quad (l = 1, ..., L) \tag{4}$$

where Θ_l is an offset, or threshold.

We used the Neural Network Toolbox in the MATLAB software environment (`www.mathworks.com`) and have used the $tansig$ function as the non-linear function:

$$\mathrm{tansig(x)} \;=\; \frac{2}{1 \;+\; \exp(-2\,\mathrm{x})} \;-1 \tag{5}$$

Function $tansig$ has an S-like shape, with values falling within the interval -1 to $+1$.

From the hidden layer to the output neurons a linear transformation with weights $\widehat{w}_{nl}$ $(n = 1, ..., N; l = 1, ..., L)$ and offsets $\widehat{\Theta}_n$ are applied:

$$y_n \;=\; \sum_{l=1}^{L} \widehat{w}_{nl} z_l \;+\; \widehat{\Theta}_n \qquad (n = 1, ..., N) \tag{6}$$

By combining Eqs. 1 and 2, it is seen that the entire network transforms the inputs x_m to the outputs y_n by the following analytical function:

$$y_n(x_1, ..., x_M) = \sum_{l=1}^{L} \widehat{w}_{nl}\, F\left(\sum_{m=1}^{M} w_{lm}x_m + \Theta_l\right) + \widehat{\Theta}_n \tag{7}$$

Clearly, such an MLP can easily be generalized to more than one hidden layer.

Given a set of P sample input and output vector pairs, $\{x_m^p\; y_n^p\}$ $p = 1, ..., P$, for a specific mapping, a technique known as error back propagation, can derive estimates of the parameters w_{lm}, Θ_m and $\widehat{w}_{nl}$, $\widehat{\Theta}_n$, so that the network function (7) will approximate the required mapping. The training algorithm minimizes the error function

$$E_{NET} \;=\; \sum_{p=1}^{P}\sum_{n=1}^{N}[y_n(x^p) \;-\; y_n^p]^2 \tag{8}$$

The general procedure for setting up a neural network is first to provide a test data set. This is traditionally split into three data sets: one set used directly to train the network; a validation data set used in the iteration scheme, not directly in the training, but in the evaluation of the improvement of the network; and a test set which is only used at the end of the training to get an independent estimate of the accuracy of the derived network.

Fig. 9 The residual (T(CMB)−T(NN calc)) map of the MLP neural network, reduced to nside = 64 (resolution 55′, unit: μK). The RMS of the map outside the mask is 2.5 μK, so the colour scale is $\pm 3\sigma$. The mask covers about 20 per cent of the sky.

6 The applied neural network

As symmetrical Gaussian beams and white noise are assumed in WG2 Challenge-2, (e.g. no $1/f$ noise contribution), the noise of the individual sky pixels is independent, and it is possible to treat each sky pixel separately.

Under these conditions, and assuming that the 4 foreground components can be treated as discussed in Section 3, it is straight forward to set up a suitable neural network.

We have set up networks to establish an algorithm for deriving the CMB temperature anisotropies from simulated

© 0000 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Fig. 10 The PSM map of number of observations per sky pixel for the 100 GHz channel (scale: 0–10000 hits). Due to the Planck scanning strategy the hit maps for the other frequencies are quite similar, except for a scale factor.

spectra, using all 7 foreground parameters, including noise, and covering the 9 Planck frequencies such that $[obs_i, i = 1, 9] \rightarrow \delta T$.

To ensure the feasibility of the network, it is, of course, important to see that the expected parameter ranges are reasonably covered, (see Figs. 1–7), and that the different combinations of parameters are also covered to a reasonable extent. Experience has shown that a training data set of about 10,000 spectra is sufficient.

To set up a neural network, the number of hidden layers, the number of neurons for each layer and the number of data sets in the training set, N_{TRAIN}, need to be specified. For a running validation of the development of the network iteration process, a data set for this purpose alone is also calculated with N_{VAL} spectra; normally 25 percent of N_{TRAIN}.

The final network is then tested on a data set with N_{TEST} spectra derived completely independently of the 2 data sets used to train the network.

With these parameters fixed, the following procedure is used:

1. Draw 7 parameter values, uniformly distributed, scaled to the parameter ranges given in Figs 1–5.
2. Draw 2 parameter values, uniformly distributed, for r_{syn} and r_{spin}.
3. Calculate the combined fluxes of the 4 foreground components and CMB at the 9 Planck frequencies.
4. For each frequency add a Gaussian randomly distributed number scaled to the noise level of each PSM noise map.
5. Repeat 1–3 until the desired number, N_{TRAIN}, of spectra in the training set have been obtained.
6. Repeat 1–3 until the desired number, N_{VAL}, of spectra in the validation set have been obtained.
7. Train the neural network to find the weights describing the mapping between the input spectra and the true temperature anisotropy (known for the simulated input).
8. Obtain independent test samples of spectra by repeating steps 1–3 $N_{(\text{TEST})}$ times.
9. Run N_{TEST} spectra through the network to get an independent estimate of the reliability of the network.

The basic neural network used in this paper is an MLP with 1 hidden layer of 2 neurons. The training set contained 10,000 spectra. To set up this kind of network the requirements on computer power are quite modest. The training of the reference network took a total of about 2.5 CPU hours on a Sun Fire V40z (8 AMD Optron @ 2.2 GHz, 16 GB RAM).

7 Results of the neural network

In order to evaluate the neural network method in relation to the PSM simulations, we have examined both the residual (T(CMB)−T(NN calc)) map and the power spectrum of the T(NN calc) map for systematic effects.

The residual maps of the 8 component separation methods of the WG2 Challenge-2 (Leach et al. 2008, Fig. 5), have been smoothed to a resolution of 45′ and are shown with a colour scale $\pm 30\mu$K. It is evident that all maps show systematic errors outside the Galactic Plane, except for the GMCA method (Bobin et al. 2007), which has a significantly larger RMS than the other methods.

In Fig. 9 the T(CMB)−T(NN calc) map is given with nside = 64 (resolution 55′), the RMS of the maps outside the Milky Way mask is 2.5 μK and the colour scale of the figure is $\pm 3\sigma$. The main structures seen in the figure are the areas with colour variations smaller than the average local RMS. By comparing with Fig. 10 it is evident that these areas correspond closely to the areas with a high number of observations, which are, of course, expected to have better accuracy.

Altogether, the neural network method provides a significantly better CMB map for 80 per cent of the sky outside the mask, compared to the methods tested in the WG2 Challenge-2.

The procedure for correcting the power spectrum of the observed T(NN calc) map is shown in Figs. 11–15.

Fig. 11 shows the power spectrum of the residual (T(CMB) − T(NN calc)) map (Fig. 9), together with the mean power spectrum of all 9 noise maps, each scaled to the observed power of the T(NN calc) map, between $l = 3000$ and $l = 3500$. For clarity the mean noise spectrum has been displaced 0.0002 mK^2. It is seen that the 2 curves follow each other closely, as expected for pure random noise in the residual map.

Fig. 12 shows the power spectrum of the found T(NN calc) map together with the mean PSM noise spectrum (see Fig. 11). In Fig. 13, the difference of these spectra is shown together with the power spectrum, including the mask, of the CMB map used in the WG2 Challenge-2. Since it is a single realization of the theoretical CMB model, the cosmic variance can be seen.

The PSM maps have been smoothed with Gaussian beams assuming FWHMs in the range 5′–33′), depending on frequency. In our neural network analysis, we have not corrected for these differences in resolution. Therefore, it is no surprise that the derived window function, calculated as the

© 0000 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Fig. 11 The power spectrum of the residual (T(CMB)−T(NN calc)) map presented in Fig. 9, (units: mK^2), and the mean of the power spectrum of the 9 noise maps, each scaled to the power spectrum of the T(NN calc) map, for $l = 3000$ to $l = 3500$. For clarity the mean noise spectrum has been displaced by 0.0002 mK^2.

Fig. 12 The power spectrum of the T(NN calc) map together with the mean power spectrum of the 9 PSM noise maps, each scaled to the T(NN calc) power spectrum, for $l = 3000$ to $l = 3500$.

power spectrum of the observed T(NN calc) map divided by the power spectrum of the applied CMB map, (both including the mask), shown in Fig. 14 does not follow a simple 'Gaussian' window function, as seen in Fig. 14). The error bars are derived from the RMS of the T(NN calc) power spectrum within each l interval.

Leach et al. (2008, Fig. 7) give only the power spectrum, incl. error bars, derived from the SMICA method (Delabrouille et al. 2003), but no information about how the correction for the window function has been performed. In order to be able to compare our errors with the SMICA errors we have divided the observed power spectrum with the mean window function shown in Fig. 14.

In Fig. 15 the final T(NN calc) power spectrum is shown. The error bars are derived from the RMS of the spectrum within each l interval. The SMICA power spectrum is also plotted. The point to notice in this plot is that our error bars are smaller than the SMICA.

Fig. 13 The T(NN calc) power spectrum with the mean noise spectrum subtracted (the average of the 9 PSM noise spectra, each scaled to the T(NN calc) power spectrum), for $l = 3000$ to $l = 3500$. The upper curve is the power spectrum of the CMB map used in Challenge-2. Therefore, the spectrum is more noisy than the theoretical spectrum used to produce the map.

Fig. 14 The window function, calculated as the power spectrum of the T(NN calc) map divided by the power spectrum of the CMB map. The window functions for Gaussian beams, FWHM = 3′, 4′, 5′ are also plotted. The error bars are determined from the RMS of the T(NN calc) power spectrum within each l interval.

8 Conclusions

In this paper we have extended the neural network method by NNJ to extract the CMB signal from the Planck WG2 Challenge-2 sky maps. As for NNJ we have shown that this method can remove the intervening foreground signals to a high degree and provide a cleaned CMB map over 80 per cent of the sky with very small systematic errors. Similarly, the method can also provide a CMB power spectrum with high accuracy and negligible systematic errors.

The analysis of the PSM maps is an ideal case, where we are able to investigate the spectral behaviour of the different galactic components from simulated, noise-free maps.

In a forthcoming paper, we will analyse the WMAP 5 yr maps to test the method with a much more realistic, noisy data set, in preparation for the Planck mission.

© 0000 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Fig. 15 The power spectrum of the T(NN calc) map corrected for the window function in Fig. 14. The error bars are determined from the RMS of the power spectrum within each l interval. The blue curve is the power spectrum obtained with the SMICA method as given by Leach et al. (2008). The noisy curve is the power spectrum of the CMB map actually used in Challenge-2. The point to notice is that our error bars are smaller than the SMICA

Acknowledgements. We would like to thank Drs. S. Leach for providing us with the SMICA PSM power spectrum, H. E. Jørgensen and A. Hornstrup for constructive comments and C. A. Oxborrow for significant improvements to the language and typesetting of this paper. We acknowlegde the use of the Planck Sky Model developed by the Component Separation Working Group of the Planck Collaboration

References

Bennett et al., the WMAP Team 2003a, ApJS, 148, 1
Bennett et al., the WMAP Team 2003b, ApJS, 148, 97
Bishop, C. M., 1995, Neural Networks for Pattern Recognition, Oxford Univ. Press
Bobin, J. Starck, J.-L.,Fadili, J., 2007, IEEE Transactions on Image Processing, in press
Delabrouille, J., Cardoso, J. F., & Patanchon, G., 2003, MNRAS, 346, 1089
Delabrouille, J., & Cardoso, J-F., 2007,lecture paper for the international summer school 'Data Analysis in Cosmology', 6–10 september 2005, Valencia, Spain
Eriksen, H. K., Banday, A. J., Gorski, K. M., & Lilje, P. B., 2004, ApJ, 612, 633
Eriksen, H. K., Dickenson, C., Lawrence, C. R., Baccigalupi, C., Banday, A. J., Gorski, K. M., Hansen, F. K., Lilje, P. B., Pierpaoli, E., Seiffert, D., Smith, K. M., & Vanderlinde, K., 2006, ApJ, 641, 665
Hinshaw et al., the WMAP Team 2007, ApJS, 170, 288
Leitch, E. M., Readhead, A. C. S., Pearson, T. J., Myers, S. T., 1997, ApJ, 486, L23
Leach et al., asto-ph 0805.0269
Mather, J. C., Fixsen, D. J., Shafer, R. A., Mosier, C., & Wilkinson, D. T., 1999, ApJ, 512, 511
Nørgaard-Nielsen, H. U., Jørgensen, H. E., 2008, Ap&SS , 318, 195
Penzias, A. A. & Wilson, R. W., 1965, ApJ, 142, 419

© 0000 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim