



**Fermi National Accelerator Laboratory**

**FERMILAB-Pub-95/375-A**

**Determining Cosmic Microwave Background Anisotropies in the  
Presence of Foregrounds**

S. Dodelson

NASA/Fermilab Astrophysics Center

*Fermi National Accelerator Laboratory*

*P.O. Box 500, Batavia, Illinois 60510*

December 1995

Submitted to *Astrophysical Journal*

## Disclaimer

*This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.*

# Determining cosmic microwave background anisotropies in the presence of foregrounds

Scott Dodelson

*NASA/Fermilab Astrophysics Center  
Fermi National Accelerator Laboratory, Batavia, IL 60510-0500*

## ABSTRACT

Separating foregrounds from the signal is one of the big challenges in cosmic microwave background (CMB) experiments. A simple way to estimate the CMB temperature in a given pixel is to fit for the amplitudes of the CMB and the various foreground components. The variance squared of this estimator is shown to be equal to  $[(\text{FDF})^2 \sigma_{\theta}^{(0)2} + \sigma_{\text{shape}}^2]$ , where  $\sigma_{\theta}^{(0)}$  is the variance in the absence of foregrounds;  $\sigma_{\text{shape}}$  is the variance due to the uncertainty in the shapes of the foreground components; and FDF is the *foreground degradation factor*. This one number, the FDF, gives a good indication of the ability of a given experiment to disentangle the CMB from foreground sources. A variety of applications relating to the planning and analyzing of experiments is presented.

# 1 Introduction

The cosmic microwave background encodes a great deal of information about our universe. In particular the anisotropies – and especially those on small scales – are sensitive to many cosmological parameters and to the initial perturbations which grew into the large structures observed today. Thus a map of the anisotropies in the CMB on small scales can unequivocally answer questions that have plagued cosmologists for decades (or longer). For this reason, a number of groups have set out to make such maps of the sky at varying angular resolution, typically better than half a degree.

There are many complicated experimental issues involved in making such maps. However, even an experiment perfectly designed to minimize atmospheric contamination, sidelobes,  $1/f$  noise, etc. still has to deal with the reality of the sky. And this reality includes not only the “signal” in the form of CMB anisotropies but also “noise” in the form of galactic and extragalactic foregrounds. The most powerful way to extract the CMB signal from foreground contamination is to take measurements at many different frequencies. The CMB anisotropies vary with frequency differently than do the foregrounds. By using the knowledge we have about these different spectral shapes, we can conceivably extract the CMB component from the total signal.

In this paper, I will discuss how to perform this extraction. For a given set of frequencies and given number of foregrounds one wants to eliminate, we can define an estimator,  $\theta^{\text{cmb}}$ , for the true CMB temperature by fitting the amplitudes of a CMB component and various foreground components to the observed temperatures in each channel. On average this estimator will equal the true CMB temperature  $t^{\text{cmb}}$ . The variance of the estimator depends on the instrumental and atmospheric noise of course. But it also depends on the frequency coverage and the foregrounds. In fact the variance can be simply expressed as

$$\sigma_{\theta}^2 = (\text{FDF})^2 \sigma_{\theta}^{(0)2} + \sigma_{\text{shape}}^2, \quad (1)$$

where  $\sigma_{\theta}^{(0)}$  is the variance in the absence of foregrounds (due to instrumental and atmospheric noise) and  $\sigma_{\text{shape}}$  is the contribution to the variance due to the uncertainty in the spectral shapes of the foregrounds. In many cases,  $\sigma_{\text{shape}}$  will be small, so the variance is enhanced over the no-foreground case by the *foreground degradation factor*, FDF. By construction, FDF is greater than or equal to one. Thus the effectiveness of any given set of frequencies can be expressed by this one number. If the FDF for a given frequency set is large, then the contaminating foreground is troubling; if FDF is close to one, then foregrounds may be effectively eliminated.

Another feature of equation 1 deserves mention. The first term is proportional to the instrumental and atmospheric noise; we will see that the second is proportional to the rms amplitude of the foregrounds. While typically the first term dominates, there are situations – e.g. in experiments with very low noise per pixel or experiments in dusty regions of the sky – where the second term is most important. In these cases, the channels should be constructed to minimize  $\sigma_{\text{shape}}$ .

Section 2 presents a prescription for calculating FDF and  $\sigma_{\text{shape}}$  for a given experimental configuration and set of foregrounds. Some of the details are relegated to Appendix A. Section 3 presents a number of applications of this analytic technique; questions which might come up in designing or analyzing an experiment which can be simply addressed with the concept of FDF and  $\sigma_{\text{shape}}$ .

This method of estimating the CMB temperature was carried out by the MSAM team in Cheng et al (1994) when analyzing their data. In several previous papers (Dodelson & Stebbins 1994 and Dodelson & Kosowsky 1995) my collaborators and I analyzed a variety of experiments using an apparently different technique, that of marginalization. In Appendix B, I show that the two techniques are in fact identical.

I should point out that there has already been a good deal of work on the issue of foregrounds. Perhaps the most influential has been the paper by Brandt et al. (1994). Without getting into the details of their work, I simply point out that their basic technique is the Monte Carlo. Here I am more interested in seeing what can be done analytically. In §3.4, I compare this analytic approach with their Monte Carlo methods and find excellent agreement. The work of Toffolatti et al (1994), Danese et al (1995), and Tegmark & Efstathiou (1995) uses information from other maps, such as the IRAS (Neugebauer et al 1984) map of dust. Although the formalism discussed in this paper can probably be extended to include such maps, here I do not attempt to do so. So the conclusions reached here are probably on the conservative side (I assume that less is known about the foregrounds). The fact that these conclusions are still reasonably optimistic is encouraging and offers still more evidence that foregrounds will *not* be an intractable problem for a satellite mission.

## 2 CMB Estimator and Variance

This section is divided into three parts. First there a brief discussion of notation; this provides the information necessary to translate the experimental/foreground information into the vectors used to calculate FDF and  $\sigma_{\text{shape}}$ . The second subsection derives the estimator of the CMB temperature and its variance. One simply performs a best fit to the free parameters: the amplitude of the various components. Calculating the variance of this estimator leads immediately to the concept and definition of FDF and  $\sigma_{\text{shape}}$ . Section 2.3 then presents a simple formula for the FDF in the presence of one and two foregrounds.

### 2.1 Notation

I will label the number of frequency channels with a subscript  $a = 1, \dots, N_{\text{ch}}$ . The observed *antenna* temperature in each channel is denoted  $T_a$ . It will be convenient to group all  $N_{\text{ch}}$  of these numbers into an  $N_{\text{ch}}$ -dimensional vector  $\vec{T}$ . The observed signal is composed of the

CMB component, foreground components, and noise, so

$$\vec{T} = \sum_{i=0}^{N_{\text{fg}}} \vec{T}^i + \vec{N} \quad (2)$$

where the CMB component has superscript 0 and the  $i = 1, \dots, N_{\text{fg}}$  foreground components are appropriately superscripted;  $\vec{N}$  denotes the contribution to the signal from instrumental and atmospheric noise. The noise is assumed Gaussian with

$$\langle \vec{N} \rangle = 0 \quad ; \quad \langle N_a N_b \rangle = C_{ab}. \quad (3)$$

Throughout,  $\vec{T}^i$  will be used to refer to the true temperatures on the sky. These are to be distinguished from the estimators,  $\vec{\Theta}^i$ , which represent our best guess about these temperatures. These estimators will assume that the shape of the foregrounds and CMB are known and take the amplitudes as free parameters. Thus, we set

$$\vec{\Theta}^i = \vec{F}^i \theta^i \quad (4)$$

where  $\theta^i$  is the (unknown) amplitude of the  $i^{\text{th}}$  component and  $\vec{F}^i$  is the (presumed) shape of that component. As a simple example consider the CMB component. We know that it has a blackbody shape, after subtracting off the mean,

$$\hat{F}_a^0 = \frac{1}{2\nu_a^2} \frac{d}{dT} B_{\nu_a}(T) = \frac{x_a^2 e^{x_a}}{(e^{x_a} - 1)^2} \quad (5)$$

where  $x_a \equiv 2\pi\hbar\nu_a/k_B\bar{T}$  and  $\bar{T} = 2.726^\circ\text{K}$ , the average temperature. With this shape vector, the amplitude  $\theta^0$  is the estimate of the *thermodynamic* temperature anisotropy, which of course is frequency independent. Note that in the Rayleigh-Jeans limit ( $x_a \rightarrow 0$ ),  $\hat{F}_a^0 \rightarrow 1$ . The CMB shape vector has a  $\hat{\cdot}$  over it to denote unit vector. That is,  $\hat{F}^0 \cdot \hat{F}^0 = 1$ , where the dot product of any two vectors is defined as

$$\vec{T} \cdot \vec{S} \equiv \sigma_\theta^{(0)2} \sum_{a,b=1}^{N_{\text{ch}}} T_a C^{-1}_{ab} S_b. \quad (6)$$

The prefactor here,  $\sigma_\theta^{(0)2}$ , is the variance in the absence of foregrounds and can be written as

$$\sigma_\theta^{(0)2} \equiv \frac{1}{\sum_{a=1}^{N_{\text{ch}}} \hat{F}_a^0 \hat{F}_b^0 C^{-1}_{ab}} \quad (7)$$

We will see shortly that this is indeed the variance in the no-foreground case, but one can immediately see that this is reasonable by considering the case of equal and uncorrelated noise with variance  $\sigma$  in each channel in an experiment with frequencies in the Rayleigh-Jeans limit. Then  $\sigma_\theta^{(0)} \rightarrow \sigma/\sqrt{N_{\text{ch}}}$ , the correct limit. Finally, it will prove useful to introduce the  $(N_{\text{fg}} + 1) \times (N_{\text{fg}} + 1)$  matrix

$$K_{ij} \equiv \vec{F}^i \cdot \vec{F}^j. \quad (8)$$

## 2.2 Best-Fit Estimator and its variance

To determine the amplitudes  $\theta^i$  of the various components, we can minimize the variance:

$$\frac{\partial}{\partial \theta^i} \left\langle \left( \vec{T} - \sum_{i=0}^{N_{\text{fg}}} \vec{F}^i \theta^i \right)^2 \right\rangle = 0. \quad (9)$$

Appendix A provides the straightforward details of this minimization. The result is that the estimator for the CMB temperature is

$$\theta^0 = \sum_{j=0}^{N_{\text{fg}}} K^{-1}_{0j} \vec{F}^j \cdot \vec{T}. \quad (10)$$

It is important to note that the estimator in equation 10 is linear in the observed temperature  $\vec{T}$ . Therefore, if the noise around  $\vec{T}$  is Gaussian, then the noise around  $\theta$  will also be Gaussian. Had we allowed the foreground shapes to vary as well, the transformation would no longer be linear, and there would be no reason to expect the noise to be Gaussian.

Equation 10 is an estimate for the thermodynamic temperature anisotropy of the CMB. How good an estimator is it? To answer this question, we need to compute

$$\sigma_{\theta}^2 \equiv \langle (\theta^0 - t^0)^2 \rangle \quad (11)$$

where  $t^0$  is the true CMB thermodynamic temperature on the sky. A short calculation (presented in Appendix A) shows that this variance is given by equation 1 with

$$\text{FDF} \equiv \sqrt{K^{-1}_{00}} \quad (12)$$

and:

$$\sigma_{\text{shape}}^2 \equiv \left( \sum_{i=1}^{N_{\text{fg}}} \vec{T}^i \cdot \sum_{j=0}^{N_{\text{fg}}} K^{-1}_{0j} \vec{F}^j \right)^2. \quad (13)$$

One important limit of equation 1 is when no foregrounds are projected out  $N_{\text{fg}} = 0$ . In that case, the matrix  $K$  has only one component, the  $_{00}$  component which is unity. Thus  $\text{FDF} = 1$  and the variance is equal to  $\sigma_{\theta}^{(0)}$ , as defined in equation 7. Another important limit is when the foreground shapes are known. In this limit,  $\sigma_{\text{shape}}$  vanishes. To see this, note that if we have chosen the correct  $\vec{F}^i$  for the foregrounds, then the true foregrounds are proportional to  $\vec{F}^i$ . Then, the dot product  $\vec{T}^i \cdot \vec{F}^j$  in equation 13 is proportional to  $K_{ij}$ . When multiplied by  $K^{-1}_{0j}$  and summed over  $j$ , this gives a delta function,  $\delta_{0i}$  which vanishes for all foregrounds  $i > 0$ .

## 2.3 FDF in the presence of one or two foregrounds

The FDF can be easily calculated via equation 12 once the matrix  $K$  is known.  $K$  in turn depends on the assumed foreground shapes via equation 8. Here I present the results for

FDF in the cases where (i) one foreground is to be projected out and (ii) two foregrounds are to be removed.

If there is one foreground to be removed with shape vector  $\vec{F}^1$ , then the matrix  $K$  depends on only the dot products  $\vec{F}^1 \cdot \vec{F}^1$  and  $\hat{F}^0 \cdot \vec{F}^1$ :

$$K = \begin{pmatrix} 1 & \hat{F}^0 \cdot \vec{F}^1 \\ \hat{F}^0 \cdot \vec{F}^1 & \vec{F}^1 \cdot \vec{F}^1 \end{pmatrix} \quad (14)$$

The inverse of this  $K$  is readily obtained:

$$K^{-1} = \frac{1}{\vec{F}^1 \cdot \vec{F}^1 - (\hat{F}^0 \cdot \vec{F}^1)^2} \begin{pmatrix} \vec{F}^1 \cdot \vec{F}^1 & -\hat{F}^0 \cdot \vec{F}^1 \\ -\hat{F}^0 \cdot \vec{F}^1 & 1 \end{pmatrix} \quad (15)$$

so that

$$\text{FDF} = \left[ \frac{1}{1 - (\hat{F}^0 \cdot \vec{F}^1)^2 / \vec{F}^1 \cdot \vec{F}^1} \right]^{1/2} \quad (16)$$

The limits of equation 16 are interesting. If the foreground component has a much different spectrum than the CMB, then their shape vectors will be much different, and  $\hat{F}^0 \cdot \vec{F}^1 \rightarrow 0$ . In this case, FDF goes to one. That is, a foreground component with a shape much different than that of the CMB does not degrade the sensitivity of the experiment. On the other hand a foreground component with a shape very close to that of the CMB (so that  $(\hat{F}^0 \cdot \vec{F}^1)^2 \rightarrow \vec{F}^1 \cdot \vec{F}^1$ ) produces a very large FDF. To minimize the FDF in a given experiment then, one must measure at frequencies designed to maximize the ‘‘angle’’ between the foreground spectrum and the CMB spectrum.

In the case when there are two foreground sources to project out, we define the three angles:

$$\cos \phi_1 \equiv \hat{F}^{\text{cmb}} \cdot \hat{F}^1 \quad ; \quad \cos \phi_2 \equiv \hat{F}^{\text{cmb}} \cdot \hat{F}^2 \quad ; \quad \cos \phi_{12} \equiv \hat{F}^1 \cdot \hat{F}^2. \quad (17)$$

Then, moving through the same steps as in the one foreground case (but this time with the aid of *Mathematica*), one finds that

$$\text{FDF} = \left[ \frac{\sin^2 \phi_{12}}{\sin^2 \phi_{12} - \cos^2 \phi_1 - \cos^2 \phi_2 + 2 \cos \phi_1 \cos \phi_2 \cos \phi_{12}} \right]^{1/2}. \quad (18)$$

Note again that in the limit that one of the foregrounds is parallel to the CMB ( $\cos \phi_1 = 1$  or  $\cos \phi_2 = 1$ ), the FDF blows up as is expected.

### 3 Applications

I now apply the formalism of the previous section to several practical questions. To set the stage, consider figure 1 which shows the spectra of the three galactic foregrounds of interest

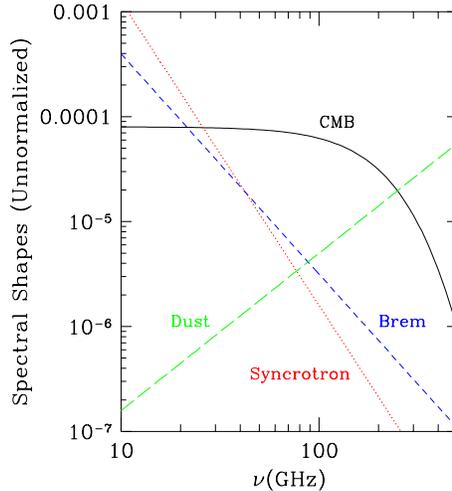


Figure 1: Unnormalized shapes of the different components in the sky.

to us: synchrotron, bremsstrahlung, and dust. The shape of the bremsstrahlung spectrum is pretty well fixed by atomic physics. If we parametrize a given shape by

$$\vec{F}_a^i \propto \nu_a^{p_i} \quad (19)$$

then  $p_{\text{brem}} \simeq -2.1$  with an uncertainty of a few percent. The spectral index of synchrotron is much more uncertain; typical estimates suggest that  $p_{\text{sync}} = -2.9 \pm 0.2$ . Finally the uncertainty in the spectral index of dust is even more pronounced; in fact it is not even clear if a fit along the lines of equation 19 is adequate to represent the complexities of dust. Nonetheless, a rough estimate might give  $p_{\text{dust}} = 1.5 \pm 0.5$ . Figure 1 illustrates these different shapes. Thus dust is expected to dominate at high frequencies and the other components at low frequencies.

### 3.1 One Component: Synchrotron

Let us start with the simplest possible example: one foreground component, synchrotron, with spectral index assumed known. This example, while crude, is not really that unrealistic. At low frequencies dust can be safely ignored, and bremsstrahlung typically comes in lower than synchrotron. Further, as we will see in the next subsection, the uncertainty in the spectral index introduces very little error.

According to equation 1, the uncertainty in our determination of the CMB temperature

has only one piece if the spectral shape of the foreground is known:

$$\sigma_{\theta}^2 = \text{FDF}^2 \sigma_{\theta}^{(0)2} \rightarrow \text{FDF}^2 \frac{\sigma^2}{N_{\text{ch}}} \quad (20)$$

where the limit  $\sigma_{\theta}^{(0)2} \rightarrow \sigma^2/N_{\text{ch}}$  holds in the case of equal and uncorrelated noise in each channel with variance  $\sigma$ , as long as we are safely in the the Rayleigh-Jeans limit. In this simple example with only one foreground to be projected out, we saw in equation 16 that  $\text{FDF} = 1 / \left(1 - (\hat{\mathbf{F}}^0 \cdot \vec{\mathbf{F}}^1)^2 / \vec{\mathbf{F}}^1 \cdot \vec{\mathbf{F}}^1\right)$  The dot products and hence the FDF depends on the shape we assume for synchrotron emission (here I will assume  $p_{\text{sync}} = -2.9$ ) and also on the placement of the frequency channels.

What is the optimal placement of frequency channels? And how many are needed? Let us first consider two frequency channels. For simplicity I will assume that they are centered about  $\nu = 40$  GHz. Figure 2 shows the FDF as a function of the difference  $\nu_{\text{high}} - \nu_{\text{low}}$ . For example  $\nu_{\text{high}} - \nu_{\text{low}} = 10$  GHz indicates two channels placed at 35 and 45 GHz. The FDF in that case is a little over three: the signal to noise is degraded by this factor. For very small frequency differences, it is difficult to disentangle the CMB component from synchrotron; hence the FDF factor is high. If the frequencies can be spread far apart, separating CMB from synchrotron becomes easier and the FDF decreases accordingly. In the limit of very large frequency difference, the FDF asymptotes to:

$$\lim_{\Delta\nu \rightarrow \infty} \text{FDF} = \sqrt{\frac{N_{\text{ch}}}{N_{\text{ch}} - N_{\text{fg}}}} \quad (21)$$

in this case  $\sqrt{2}$ . To see why, note that in the absence of foregrounds, the additional information from all the channels beats down the noise by a factor of  $1/\sqrt{N_{\text{ch}}}$ ; this is the factor explicitly present in equation 20. When a foreground component is present, at least one of the channels must be used to determine the foreground amplitude. Thus even in the ideal case, when the foreground component can be easily distinguished from the CMB, there is still one fewer channel with which to measure the CMB. Hence, the true noise is now down by a factor of  $1/\sqrt{N_{\text{ch}} - 1}$ . And on it goes as more foreground amplitudes must be separated. Note that these arguments are only valid in the Rayleigh-Jeans limit; For higher frequencies, the limits in equation 20 and equation 21 no longer apply.

Now consider three frequency channels. It is clear that it is best to get as large a spread in the frequencies as possible. But where best to place the middle frequency channel? Figure 3 shows the FDF as a function of the frequency of the middle frequency channel when  $\nu_{\text{low}} = 30$  GHz and  $\nu_{\text{high}} = 50$  GHz. Figure 3 shows that the optimal place for the middle frequency channel is at  $\nu = 50$  GHz! At first, this is surprising, but it makes sense upon further reflection: the lowest channel is used to separate out the synchrotron component. The other channels are best placed where they will get the least contaminated by synchrotron; thus all other channels should go as high in frequency as possible. Of course, this example is somewhat artificial: when more than one foreground component is projected out, it becomes

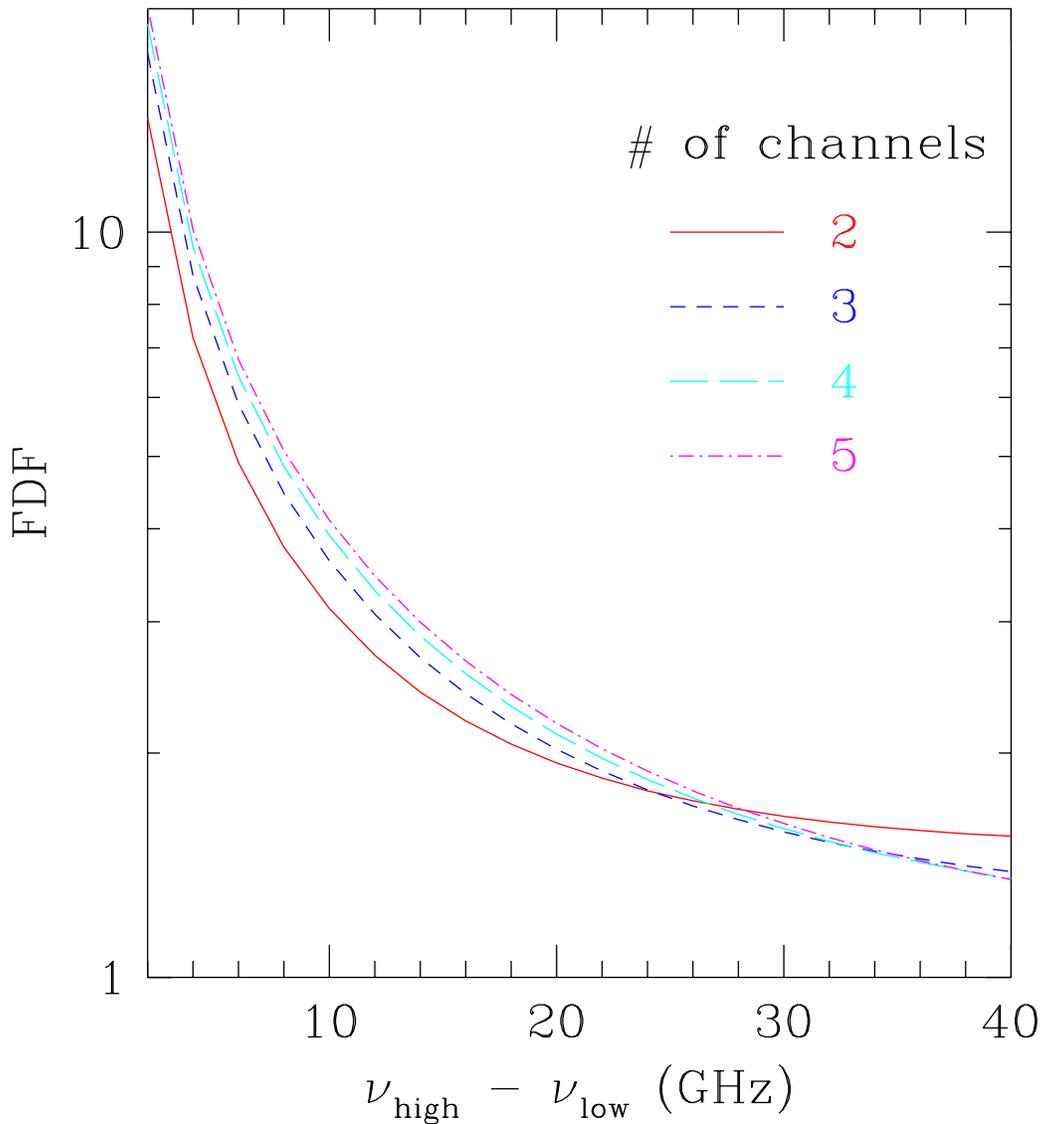


Figure 2: FDF as a function of the difference between the highest and lowest frequencies in an experiment when synchrotron with assumed index  $-2.9$  is fitted for. The extreme channels are centered around 40 GHz (so  $(\nu_{\text{low}} + \nu_{\text{high}})/2 = 40$  GHz). The curves with more than two channels have their frequencies equally spaced between the two extremes. FDF is lowered – hence the experiment has the best discrimination against foregrounds – when the frequency spread is as large as possible.

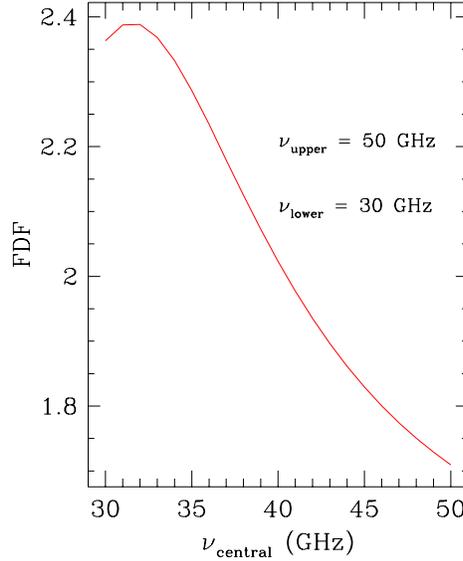


Figure 3: FDF as a function of the placement of the central frequency in an experiment with three channels, the other two at 30 and 50 GHz. Again the foreground component is synchrotron with assumed index  $-2.9$ . The optimal place to locate the third channel is at 50 GHz, where FDF is minimized.

important to space out the channels more evenly. Nonetheless, I hope this simple example alerts experimenters to the possibility that the best signal to noise may be achieved with an unorthodox positioning of the frequency channels. In this simple example, the FDF varies from 2.3 to 1.7, i.e. by roughly 30%, as one varies the placement of the middle channel. So clever positioning of the intermediate channels could be an easy way to increase the final signal to noise.

Figure 2 shows the FDF for this case of three frequency channels as a function of the difference between the highest and lowest channel. In this graph, the middle channel is not placed in the optimal position (usually at the highest frequency possible), but all the channels are evenly spaced. (Thus the point corresponding to  $\nu_{\text{high}} - \nu_{\text{low}} = 10$  GHz and  $N_{\text{ch}} = 3$  has channels at  $\nu = 35, 40, 45$  GHz.) It is interesting that, except for the largest frequency spreads, adding extra channels does not really help in disentangling the foregrounds. (This point was also made by Brandt et al. (1994).) Certainly going beyond  $N_{\text{ch}} = 3$  provides very little gain in this simple case.

### 3.2 Uncertain foreground shape

We can generalize the discussion of the previous subsection by accounting for the fact that the shape of the synchrotron spectrum is not perfectly determined. If we allow for this uncertainty, there arises a new term in the variance of the estimator. Following equation 13, we see that

$$\sigma_{\text{shape}} = \vec{T}^{\text{sync}} \cdot \left[ K^{-1}_{00} \hat{F}^{\text{cmb}} + K^{-1}_{01} \vec{F}^{\text{sync}} \right] \quad (22)$$

where again I emphasize that  $T^{\text{sync}}$  is the *true* synchrotron temperature, with a spectral shape that differs from the assumed one. We will suppose that the true shape of the synchrotron spectrum is still given by equation 19, but with spectral index  $p \neq -2.9$ , the assumed index. Figure 4 shows the error induced by assuming the wrong spectral index. For the kind of

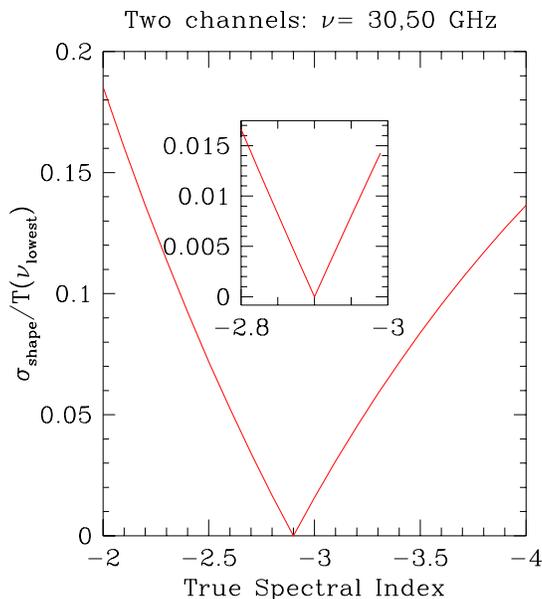


Figure 4:  $\sigma_{\text{shape}}$ : the variance in the determination of the CMB temperature due to uncertainty in the shape of the foreground spectrum being fitted for. The assumed spectral index is  $-2.9$ ; if the true index is equal to this, then  $\sigma_{\text{shape}}$  vanishes.

uncertainty typically measured for synchrotron,  $\Delta p \sim 0.2$ , the error induced is less than a few percent of the synchrotron amplitude. Thus, if the synchrotron amplitude is  $40 \mu\text{K}$ , the uncertainty in the spectral index contributes less than  $1 \mu\text{K}$  to the total error. This error is very small and for reasonable noise values will be much smaller than the (FDF)  $\sigma/\sqrt{N_{\text{ch}}}$  factor discussed above.

It is instructive to understand why the uncertainty in the spectral index leads to very small errors in the CMB temperature determination. By projecting out the  $p = -2.9$  component, we are looking in the space perpendicular to the shape vector defined by  $p = -2.9$ . But, the vector defined by  $p = -2.8$  is almost perfectly parallel to the  $p = -2.9$  shape vector. Thus, it has a very small component in the perpendicular space. Unless the amplitude is extremely large, the perpendicular component is negligible.

Figure 4 shows that, even when we project out only synchrotron emission, we also succeed in eliminating a large fraction of the bremsstrahlung (with  $p = -2.1$ ) as well. In this example, only 15% of the bremsstrahlung amplitude remains after projecting out a  $p = -2.9$  component. So the simple projection of  $p = -2.9$  is sufficient for all but the most sensitive experiments.

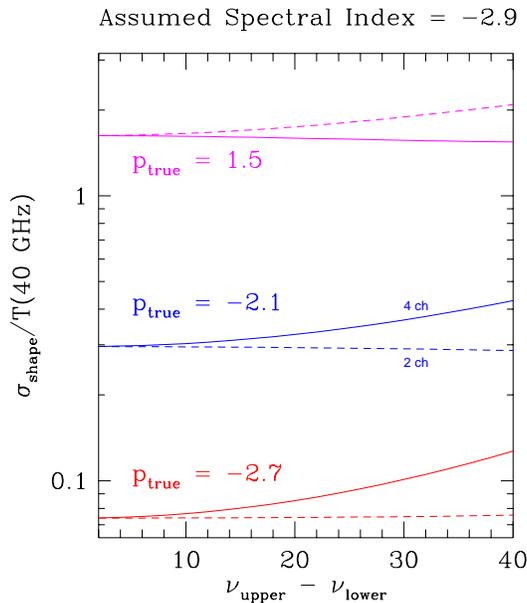


Figure 5:  $\sigma_{\text{shape}}$  for different true foreground shapes as a function of the frequency range of an experiment. Solid lines are for a 4 channel experiment (with equally spaced frequencies); dashed lines for a 2 channel experiment. In all cases, the range is centered around 40 GHz.

Figure 5 shows how  $\sigma_{\text{shape}}$  varies as the frequency coverage changes. In this example, increasing the frequency range always leads to an increase in  $\sigma_{\text{shape}}$ . For some foregrounds, increasing the number of channels also leads to an increase in  $\sigma_{\text{shape}}$ , although this is not true for dust here. I have not been able to figure out any general principles for minimizing  $\sigma_{\text{shape}}$ ; fortunately, it is simple enough to deal with each case individually.

### 3.3 Projecting out two components

This subsection deals with an analysis question. How best to analyze the data? In particular, should one attempt to fit for several components or is it best to fit for fewer components? I will argue that projecting out two components often will lead to a larger variance than if one simply projected out one component as in §3.1.

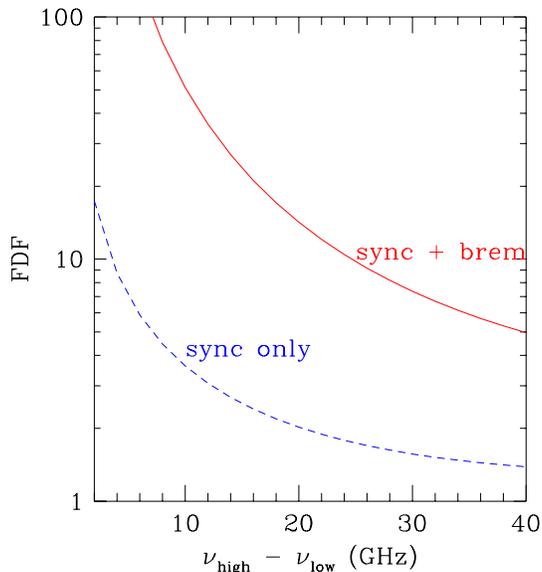


Figure 6: FDF as a function of frequency range in a three channel experiment. Projecting out two components leads to a much larger FDF.

Figure 6 shows the FDF for an experiment with three frequency channels when both synchrotron [with index  $-2.9$ ] and bremsstrahlung are fitted for. For comparison, also shown is the FDF if only synchrotron was fitted for. Again the central channel is at  $\nu = 40$  GHz. In all cases, the FDF is much higher if both components are fitted for. For example, with channels at  $\nu = 30, 40, 50$ , figure 6 shows that the “one-component” FDF = 2 while the “two-component” FDF = 14.

Let me pursue this example further. When is it advantageous to project out two components? The total variance in the two-component analysis is

$$\sigma_{\theta}^2|_{\text{two-component}} = (14)^2 \sigma_{\theta}^{(0)2}. \quad (23)$$

There is also a small uncertainty due to the unknown shape but this is very small so I neglect it. In the one-component analysis, we must include the shape uncertainty since the

bremmstrahlung amplitude is not projected out. Thus

$$\sigma_{\theta}^2|_{\text{one-component}} = (2)^2\sigma_{\theta}^{(0)2} + (0.18)^2\langle T_{\text{brem}}^2(\nu = 30\text{GHz})\rangle. \quad (24)$$

The coefficient 0.18 in equation 24 can be simply read off of figure 4. It becomes useful to analyze the data by fitting for two components only when  $\sigma_{\theta}^2|_{\text{two-component}} < \sigma_{\theta}^2|_{\text{one-component}}$ . Using equations 23 and 24, we find that this occurs when

$$\frac{\langle T_{\text{brem}}^2(\nu = 30\text{GHz})\rangle^{1/2}}{\sigma_{\theta}^{(0)}} > 77. \quad (25)$$

For even the most sensitive experiments, we do not expect foreground amplitudes of this magnitude. So in this example, it would be best to analyze the experiment by fitting for only the synchrotron component.

One must pursue each example on a case by case basis. This simple analytic technique should prove useful in deciding how best to analyze the data. This simple example suggests that fitting for fewer components leads to a smaller variance in the CMB temperature; this agrees with the general point made by Brandt et al. (1994). Now let us turn to a more quantitative comparison with that work.

### 3.4 Comparison with Brandt et al.

The analytic techniques presented here can be compared with the Monte Carlos performed by Brandt et al. (1994). Here I focus on one example of theirs, a seven channel experiment with equally spaced frequencies between 25 and 38 GHz. I will not describe their methodology in detail [please see their paper for a lucid description of what they did]. For the purposes of comparison, note that they were interested in the same quantity I have been focusing on: the total variance in the determination of the CMB temperature. I have called it  $\sigma_{\theta}$ ; they called it  $E_{\text{RMS}}$ . Given the experimental configuration and the average foreground levels, we can plot this variance as a function of instrumental and atmospheric noise per channel,  $\sigma$  [in their notation roughly equivalent to  $\xi$ ]. Figure 7 shows a comparison of the analytic technique and the Monte Carlos. The points are two different techniques that they used to extract the CMB temperature. They allow for free synchrotron amplitude [model *Q2*] and free synchrotron and bremsstrahlung amplitudes [model *P3*]. This corresponds to projecting out  $N_{\text{fg}} = 1, 2$  foregrounds respectively. The curves show  $\sigma_{\theta}$  with these two projections. To get these curves I needed  $\langle T_{\text{sync}}^2 \rangle$  and  $\langle T_{\text{brem}}^2 \rangle$ ; I took the same values they used in their Monte Carlos.

The agreement is excellent and shows clearly that the simple analytic technique adequately describes the situation. The shapes of their curves now becomes obvious: at low noise levels [small  $\sigma$ ],  $\sigma_{\text{shape}}$  dominates over the FDF-enhanced noise. Since  $\sigma_{\text{shape}}$  is independent of noise, the total variance is also independent of noise in this regime. That is, at low  $\sigma$ ,  $\sigma_{\theta}$  is constant. In the opposite limit, FDF-enhanced noise dominates over  $\sigma_{\text{shape}}$ , so the total variance increases linearly with  $\sigma$ .

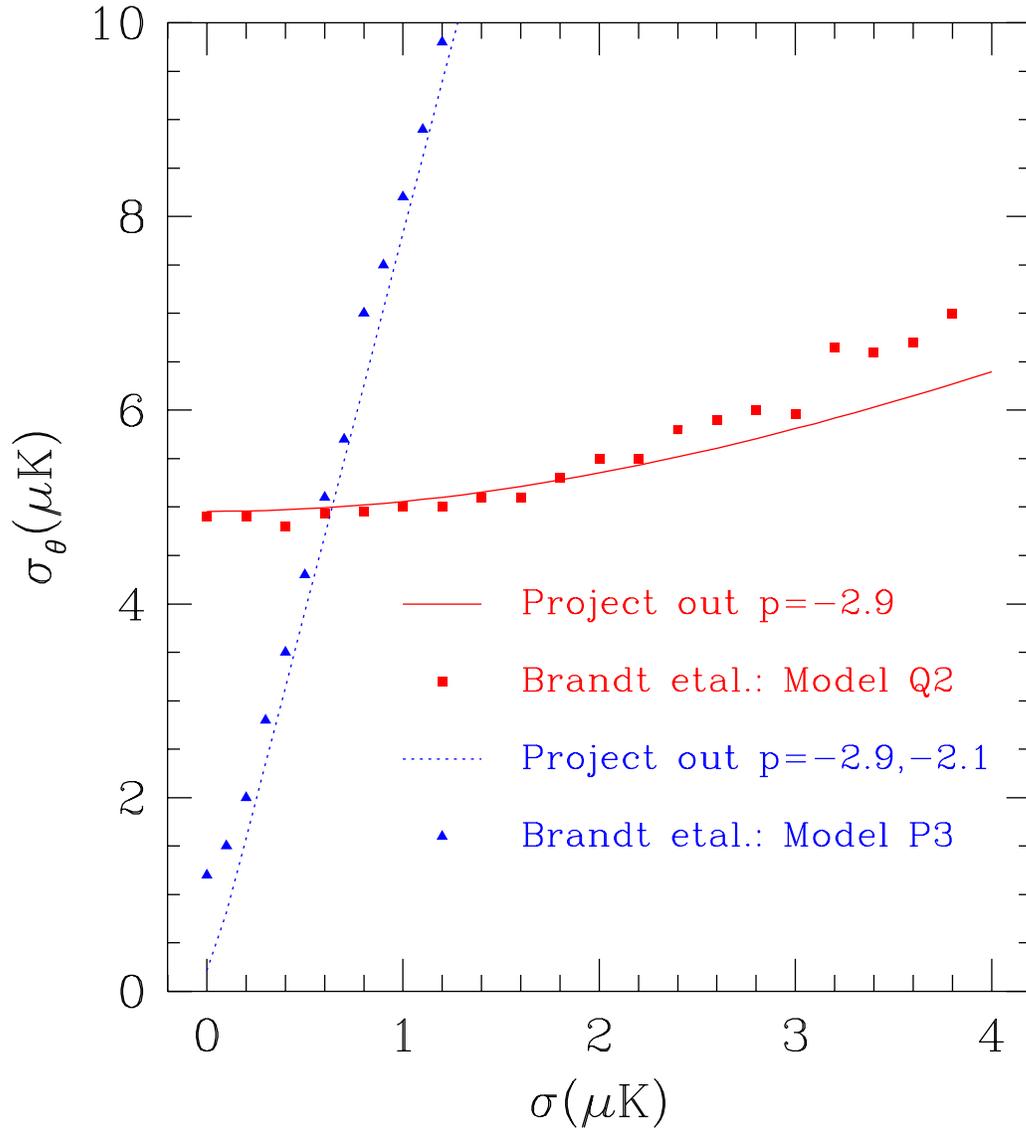


Figure 7: The variance in the determination of the CMB temperature as a function of the noise per channel. The points denote two different methods used by Brandt et al (1994) to extract the CMB temperature. The lines are the variances one gets with the analytic formula of Eq. 1 fitting for one and two foreground components.

One final point about our approaches. They presented many other “model”s for extracting out the CMB temperature. For example, one of their models allowed the synchrotron index to be a free parameter. Within the analytic framework presented here, I cannot allow the shapes to be free parameters. However, the variances using such models are much higher than the variances in the models where only the amplitudes are allowed to vary. So I would argue that the analytic technique cannot do everything but it can do the things that are worth doing.

### 3.5 Breaking up bands and noise correlations

We have seen in the previous sections that adding more frequency channels to an experiment is not necessarily a good thing. For, intermediate channels are not as effective in separating out different spectra; a longer lever arm with good sensitivity at both ends is often preferable. In this section I focus on another possible danger of splitting up bands. Often when a given frequency band is split up, the noise in the new channels is correlated. How does correlated noise impact on the decision to split up bands? Here, I address this question in the context of a simple example.

Consider an experiment with two channels in the Rayleigh- Jeans regime, say  $\nu = 15, 45$  GHz. The noise in each channel is assumed to have variance  $\sigma$ , so in the absence of foregrounds, the variance in the determination of the CMB temperature would be  $\sigma/\sqrt{2}$ . If we wish to project out synchrotron emission (with assumed spectral index  $-3$ ), then this variance is increased by the FDF. In this case a simple computation yields  $\text{FDF} = 1.51$ , so the total variance in the experiment is

$$\sigma_{\theta,2} = 1.07\sigma \tag{26}$$

where the subscript  $_2$  denotes the number of channels. Is it worthwhile to add two new frequency channels at  $\nu = 25, 35$  GHz? I will assume that in so doing, the noise in each channel increases by  $\sqrt{2}$ , so that in the absence of foregrounds, the variance would still be  $(\sqrt{2}\sigma)/\sqrt{4} = \sigma/\sqrt{2}$ . If there were no correlations introduced between the different channels, then we could do a simple calculation and find that  $\text{FDF} = 1.33$ . Thus the total variance in this 4-channel case is only  $0.94\sigma$ , smaller than in the two channel case and perhaps worth the effort.

However, if correlations amongst the different channels are introduced, then the calculation becomes slightly less trivial. Here I carry out the calculation in this correlated case for several reasons. First, this will give us a sense of whether or not it is important to avoid correlations. But more importantly, I hope that this provides yet another example of how useful the formalism of §2 can be when it comes to analyzing specific problems.

For simplicity I will assume that the two lowest channels are correlated as are the two

highest channels, so the new noise correlation matrix is

$$C = 2\sigma^2 \begin{pmatrix} 1 & \epsilon & 0 & 0 \\ \epsilon & 1 & 0 & 0 \\ 0 & 0 & 1 & \epsilon \\ 0 & 0 & \epsilon & 1 \end{pmatrix}. \quad (27)$$

For the calculation we will need the inverse of  $C$ . A short calculation shows that

$$C^{-1} = \frac{1}{2\sigma^2(1-\epsilon^2)} \begin{pmatrix} 1 & -\epsilon & 0 & 0 \\ -\epsilon & 1 & 0 & 0 \\ 0 & 0 & 1 & -\epsilon \\ 0 & 0 & -\epsilon & 1 \end{pmatrix}. \quad (28)$$

We can now immediately calculate the variance in the absence of foregrounds:

$$\begin{aligned} \sigma_{\theta}^{(0)2} &= \frac{1}{\sum_{ab} C^{-1}_{ab}} \\ &= \frac{\sigma^2(1+\epsilon)}{2} \end{aligned} \quad (29)$$

where in the first line I have assumed that we are deep enough into the Rayleigh-Jeans to set  $\hat{F}_a^0 = 1$ . Thus, the variance in such an experiment – in the absence of foregrounds – increases due to correlations by a factor of  $\sqrt{1+\epsilon}$ . This is a very simple way of saying what I and my collaborators tried to illustrate in Dodelson, Kosowsky, and Myers (1995). Now let us include the effects of foregrounds. Our standard formula gives

$$\text{FDF}^2 = \frac{1}{1 - (\vec{F}^1 \cdot \hat{F}^0)^2 / (\vec{F}^1 \cdot \vec{F}^1)} \quad (30)$$

so we need to calculate the two dot products. The only complication here is that we need to account for the non-diagonal structure of  $C$ . Thus,

$$\begin{aligned} \vec{F}^1 \cdot \hat{F}^0 &= \sigma_{\theta}^{(0)2} \sum_{ab} C^{-1}_{ab} F_b^1 \\ &= \sigma_{\theta}^{(0)2} \frac{\sum_a F_a^1}{2\sigma^2(1+\epsilon)} \end{aligned} \quad (31)$$

and

$$\vec{F}^1 \cdot \vec{F}^1 = \sigma_{\theta}^{(0)2} \frac{\sum_a (F_a^1)^2 - 2\epsilon(F_1^1 F_2^1 + F_3^1 F_4^1)}{2\sigma^2(1-\epsilon^2)}. \quad (32)$$

With these expressions for the dot products we can now evaluate the FDF with equation 30.

Figure 8 shows the variance in the determination of the CMB temperature as a function of the correlation between the channels when the synchrotron is fitted for. Apparently, increased correlation does not significantly increase the variance. So, at least in this example, noise correlation should not deter experimenters from adding new channels.

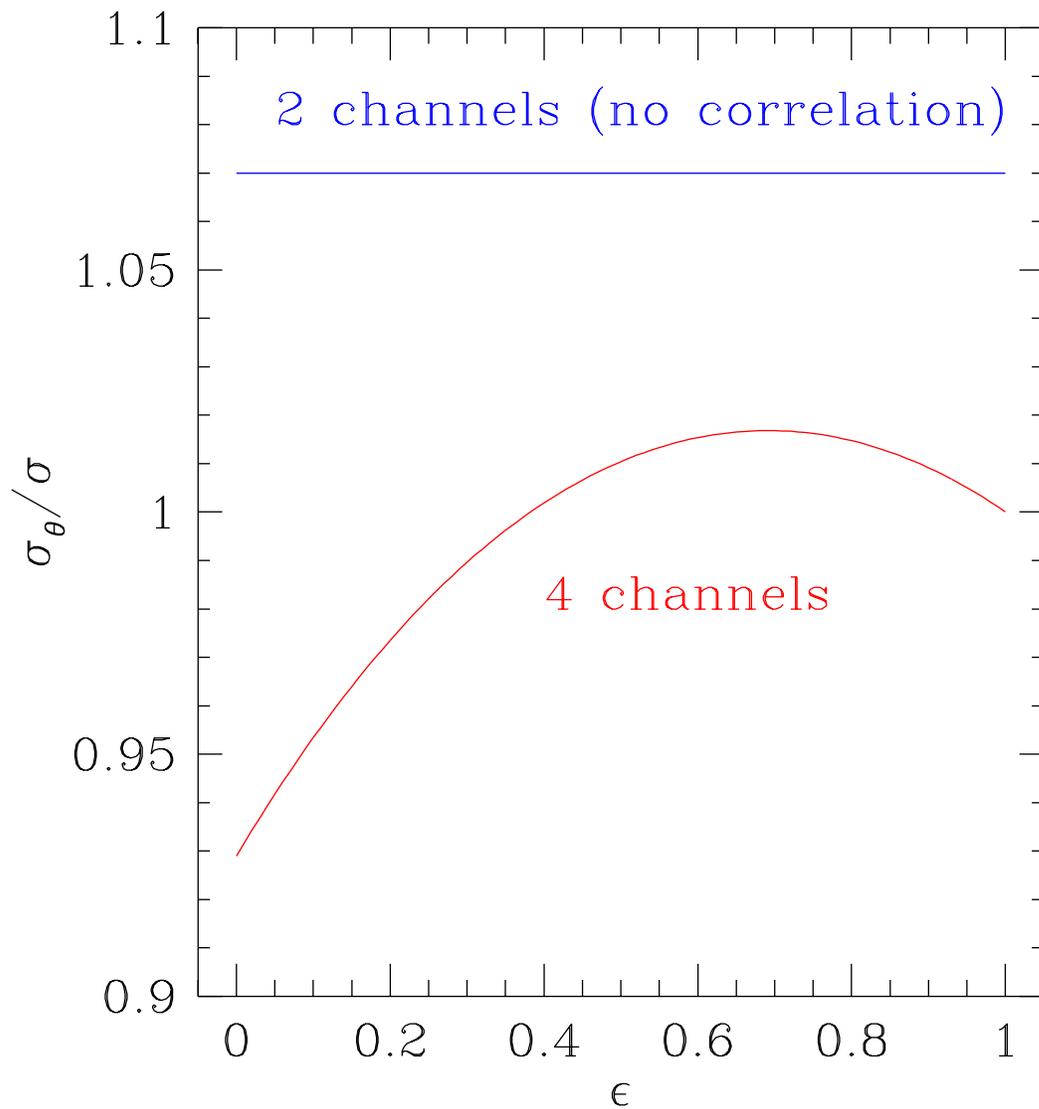


Figure 8: The variance in the determination of the CMB temperature as a function of the correlation amongst the different frequency channels. This is to be compared with the horizontal line, the variance in the two channel case when there is no correlation. Since the four channel curve is lower than the horizontal line, it would always be advantageous to add the extra channels in this case even if correlations were introduced.

### 3.6 Current Experiments

To get a feel for how well current experiments are doing at separating out foregrounds, I compiled Table 1. For each experiment, the FDF is computed for a given spectral index. For example the FDF for COBE fitting for bremsstrahlung is 1.75. Also shown is the uncertainty due to the shape. Again for COBE, if bremsstrahlung is fit for, then a dust component [with index 1.5] contributes an uncertainty  $\sigma_{\text{shape}} = 5.32 \langle T_{\text{lowest}}^2 \rangle^{1/2}$ . For COBE, the lowest frequency channel is at 31 GHz. At this frequency, one expects an rms dust antenna temperature of order a few  $\mu\text{K}$ , so – in the absence of any other maps – the uncertainty due to dust would be of order  $10\mu\text{K}$ . [This is not intended to be a rigorous estimate of the uncertainty due to dust, just a guide to reading the table. COBE has access to – and used much other information to get a handle on dust. See for example Bennett et al. (1994).]

A cursory look at some of the other experiments in Table 1 shows that typical FDF's are of order 1 – 4. Bolometer experiments like MAX and MSAM do very well at projecting out dust. [Note though that MAX3 in particular could not distinguish well between CMB and bremsstrahlung.] The HEMT experiments do not have large frequency coverage, so they discriminate less well than the high frequency experiments. However, the recent modifications to the South Pole and the Saskatoon [additions of higher frequency channels] have significantly reduced their FDF's.

The last line of Table 1 presents the FDF for a hypothetical experiment with equally spaced frequencies between 30 and 120 GHz (this case was also analyzed by Brandt et al. (1994)). Analyzing by projecting out two components leads to a variance squared equal to  $(2.97 \times \sigma/\sqrt{7})^2 + (.169T_{\text{brem}}(30\text{GHz}))^2$ , where  $\sigma$  is the noise in each channel.

Table 1: FDF's for selected experiments.

Experiment	Assumed Index	FDF	Foreground with $p =$ $\rightarrow \sigma_{\text{shape}}/T_{\text{lowest}}$
COBE <sup>a</sup>	-2.1	1.75	1.5 $\rightarrow$ 5.32
FIRS <sup>b</sup>	1.5	1.02	2 $\rightarrow$ 2.45
MAX3 <sup>c</sup>	-2.1	11.5	1.5 $\rightarrow$ 73.7
MAX4 <sup>d</sup>	-2.1	4.09	1.5 $\rightarrow$ 12.4
MAX3	1.5	1.12	-2.1 $\rightarrow$ 2.09
MAX4	1.5	1.06	-2.1 $\rightarrow$ 1.16
MAX3	1.5	1.12	2 $\rightarrow$ 1.34
MAX4	1.5	1.06	2 $\rightarrow$ 1.95
MSAM1 <sup>e</sup>	1.5	1.02	2 $\rightarrow$ 2.48
SK93 <sup>f</sup>	-2.9	4.48	-2.1 $\rightarrow$ .234
SK94 <sup>g</sup>	-2.9	2.35	-2.1 $\rightarrow$ .180
SP91 <sup>h</sup>	-2.9	4.00	-2.1 $\rightarrow$ .225
SP94 <sup>i</sup>	-2.9	2.40	-2.1 $\rightarrow$ .179
Tenerife <sup>j</sup>	-2.9	1.54	-2.1 $\rightarrow$ .094
Satellite	-2.9,1.5	2.97	-2.1 $\rightarrow$ .169

<sup>a</sup> Bennett et al. (1994)

<sup>b</sup> Ganga et al. (1994)

<sup>c</sup> Meinhold et al. (1993)

<sup>d</sup> Clapp et al. (1995)

<sup>e</sup> Cheng et al. (1994)

<sup>f</sup> Wollack et al. (1993)

<sup>g</sup> Netterfield et al. (1994)

<sup>h</sup> Gaier et al. (1992)

<sup>i</sup> Gundersen et al. (1994)

<sup>j</sup> Hancock et al. (1994)

## 4 Conclusions

This paper has introduced an analytic technique that can be used to help design an experiment and to analyze data. The main result coming out of this analytic treatment is that the variance in the determination of the CMB temperature has two components. First, there is a component proportional to noise; due to fitting for foregrounds, noise is amplified by the FDF. Second, there is a component  $\sigma_{\text{shape}}$  proportional to the foreground amplitudes. This component vanishes if the foreground spectra are known, but is non-zero if there is some uncertainty in the shapes. This simple model of CMB extraction was shown in §3.4 to reproduce the Monte Carlo results of Brandt et al (1994) very accurately.

I think the most useful thing to emerge is the technique itself, which is easy to understand and implement. Any given experiment will have its own set of complications, so it is dangerous to make general conclusions about the “best” set of frequency channels. Nonetheless, on the basis of the work in §3, there are several general principles that should be considered in any experimental plan/analysis.

- *A wide range of frequencies does best at minimizing the FDF.* In general this would lead one to go with large frequency ranges. Indeed, I would argue that experiments with bolometers have been more successful to date at extracting the CMB since they allow a larger frequency range. However, one can envision circumstances where increasing the range is not beneficial<sup>1</sup>. In particular, as shown in §3.2, increasing the frequency range often leads to a larger  $\sigma_{\text{shape}}$ . This effect can be even more dramatic if the new frequencies are more sensitive to a different foreground component [e.g. a 120 GHz channel added to a low frequency experiment would be more sensitive to dust].
- Equal spacing of the intermediate channels is not always the optimal way to go. Furthermore, adding more intermediate channels is also not necessarily beneficial. However, it does not appear – at least from the example studied in §3.5 – that noise correlations amongst different channels should be a deterrent in this regard.
- In terms of analysis, in agreement with the results of Brandt et al. (1994), I found in §3.3 that it is best to fit for as few components as possible. A cursory glimpse at present experiments suggests that their signal to noise is degraded due to foregrounds by a factor ranging from one to four, with bolometers at the low end and HEMPTs at the high end. A satellite experiment with frequencies ranging from 30 to 130 GHz would see its signal to noise degraded by about three.

## Acknowledgments

Many people on different experimental groups have been thinking about similar problems to the ones discussed here. I thank them for discussions. Especially helpful were Gary Hinshaw,

---

<sup>1</sup>I am grateful to Steve Levin for emphasizing this point to me.

Lloyd Knox, and Steve Levin. I am also grateful to Arthur Kosowsky and Albert Stebbins, my collaborators on previous aspects of foreground separation. This work was supported in part by the DOE (at Chicago and Fermilab) and the NASA (at Fermilab through grant NAG 5-2788).

## A Derivation of Variances

We want to solve equation 9 for the free parameters, the amplitudes  $\theta^i$ . This minimization requirement is satisfied when

$$\vec{F}^j \cdot \left( \vec{T} - \sum_{i=0}^{N_{\text{fg}}} \vec{F}^i \theta^i \right) = 0. \quad (33)$$

Using the definition of  $K$  in equation 8 leads to

$$\sum_{i=0}^{N_{\text{fg}}} K_{ji} \theta^i = \vec{F}^j \cdot \vec{T}. \quad (34)$$

Multiplying by  $K^{-1}$  and summing over  $j$  leads to

$$\theta^i = \sum_{j=0}^{N_{\text{fg}}} K^{-1}_{ij} \vec{F}^j \cdot \vec{T}. \quad (35)$$

The  $i = 0$  component of this equation is the estimator for the CMB temperature presented in equation 10.

Now we want to calculate the variance of the estimator for the CMB temperature. Start from equation 11 and use 10 for  $\theta^0$  and equation 2 for  $\vec{T}$ . Then,

$$\sigma_{\theta}^2 = \left\langle \left( \sum_{j=0}^{N_{\text{fg}}} K_{0j}^{-1} \vec{F}^j \cdot \left[ \sum_{i=0}^{N_{\text{fg}}} \vec{T}^i + \vec{N} \right] - t^0 \right)^2 \right\rangle. \quad (36)$$

Consider the  $i = 0$  term here. This is

$$\sum_j K_{0j}^{-1} \vec{F}^j \cdot \hat{F}^0 t^0 = \sum_j K_{0j}^{-1} K_{j0} t^0 = t^0 \quad (37)$$

so this part of the sum exactly cancels the  $t^0$  term in equation 36. Thus we are left with

$$\sigma_{\theta}^2 = \left\langle \left( \sum_{j=0}^{N_{\text{fg}}} K_{0j}^{-1} \vec{F}^j \cdot \left[ \sum_{i=1}^{N_{\text{fg}}} \vec{T}^i + \vec{N} \right] \right)^2 \right\rangle. \quad (38)$$

The first term in square brackets is exactly  $\sigma_{\text{shape}}$  in equation 13. The noise term is slightly more complicated. It is

$$\begin{aligned}
\left\langle \left( \sum_{j=0}^{N_{\text{fg}}} K_{0j}^{-1} \vec{F}^j \cdot \vec{N} \right)^2 \right\rangle &= \sum_{j=0}^{N_{\text{fg}}} K_{0j}^{-1} \sigma_{\theta}^{(0)2} \sum_{ab} F_a^j C^{-1}_{ab} \sum_{j'=0}^{N_{\text{fg}}} K_{0j'}^{-1} \sigma_{\theta}^{(0)2} \sum_{a'b'} F_{a'}^{j'} C^{-1}_{a'b'} \langle N_b N_{b'} \rangle \\
&= \sigma_{\theta}^{(0)4} \sum_{jj'} K_{0j}^{-1} K_{0j'}^{-1} \sum_{aa'bb'} F_a^j F_{a'}^{j'} C^{-1}_{ab} C^{-1}_{a'b'} C_{bb'} \\
&= \sigma_{\theta}^{(0)2} \sum_{jj'} K_{0j}^{-1} K_{0j'}^{-1} \left( \sigma_{\theta}^{(0)2} \sum_{aa'} F_a^j C^{-1}_{aa'} F_{a'}^{j'} \right) \tag{39}
\end{aligned}$$

In going from the first to the second line here I have used equation 3. The term in parentheses on the last line in equation 39 is by definition equal to  $\vec{F}^j \cdot \vec{F}^{j'} \equiv K_{jj'}$ . This contracts with one of the  $K^{-1}$ 's to give  $\delta_{0j}$ . Thus all that is left of the sum is  $K^{-1}_{00}$ . This corresponds to the FDF in equation 12.

## B Marginalization

This appendix presents what appears to be another way to extract the CMB signal. For a long time I thought that this estimator was better than the one presented in the body of the paper. I even gave a talk or two explaining that this estimator was preferable to any other. This is not true. Both estimators are identical. In this appendix, I first present the other method and then prove that both estimators, although they look completely different, are in fact identical. Albert Stebbins and I in Dodelson & Stebbins (1994) wrote about marginalization and Dodelson & Kosowsky (1995) describes some more marginalization work. All of this is now shown to be equivalent to the best fit technique used for example by Cheng et al. (1994) to analyse the MSAM data.

The idea is to project out all the foreground sources. The  $N_{\text{fg}}$  foregrounds span an  $N_{\text{fg}}$ -dimensional subspace of  $\mathfrak{R}^{N_{\text{ch}}}$ ; call this subspace  $\mathcal{F}$ . All foreground contributions to the signal live in  $\mathcal{F}$ . The *orthogonal complement* of  $\mathcal{F}$ ,  $\mathcal{F}_{\perp}$ , is the space which contains all vectors orthogonal to the foregrounds. Thus any vector in  $\mathcal{F}_{\perp}$  is completely independent of foregrounds. What we need to do, therefore, is project the observed temperature vector on to  $\mathcal{F}_{\perp}$ . This projected temperature will be independent of any foregrounds and hence will provide an unbiased estimate of the CMB temperature. To project on to  $\mathcal{F}_{\perp}$ , we first need a set of basis vectors in  $\mathcal{F}_{\perp}$ . Let us call these

$$\hat{z}^{(r)} \quad r = 1, \dots, N_{\text{ch}} - N_{\text{fg}}. \tag{40}$$

These basis vectors are chosen to be orthonormal, so they are perpendicular to each other and they have unit norm:

$$\hat{z}^{(r)} \cdot \hat{z}^{(s)} = \delta_{rs}. \tag{41}$$

Of course there are  $N_{\text{ch}} - N_{\text{fg}}$  of these vectors since they span the  $N_{\text{ch}} - N_{\text{fg}}$  dimensional space  $\mathcal{F}_{\perp}$ . Finally, by the definition of  $\mathcal{F}_{\perp}$ , the vectors  $z^{(r)}$  must satisfy

$$\hat{z}^{(r)} \cdot \vec{F}^i = 0 \quad r = 1, \dots, N_{\text{ch}} - N_{\text{fg}} \quad ; \quad i = 1, \dots, N_{\text{fg}}. \quad (42)$$

With the basis vectors  $z^{(r)}$ , we can form the projection operators which project any vector in the full  $N_{\text{ch}}$  dimensional space on to  $\mathcal{F}_{\perp}$ , the space independent of foregrounds. For an arbitrary vector  $\vec{x}$  in the full space,

$$\vec{x}_{\perp} \equiv \sum_{r=1}^{N_{\text{ch}}-N_{\text{fg}}} \hat{z}^{(r)} (\hat{z}^{(r)} \cdot \vec{x}) \quad (43)$$

is the projection on to  $\mathcal{F}_{\perp}$ . Thus  $\vec{x}_{\perp}$  is independent of any foregrounds.

We are now in a position to get the marginalization estimate for the CMB temperature. First we project the observed temperature on to the space independent of foregrounds. Then we find the CMB component of this projected temperature. The estimator is therefore

$$\theta' \equiv \frac{\hat{F}^0 \cdot \vec{T}_{\perp}}{\hat{F}^0 \cdot \hat{F}_{\perp}^0} = \sum_{r=1}^{N_{\text{ch}}-N_{\text{fg}}} (\hat{F}^0 \cdot \hat{z}^{(r)}) (\hat{z}^{(r)} \cdot \vec{T}) / \sum_{r=1}^{N_{\text{ch}}-N_{\text{fg}}} (\hat{F}^0 \cdot \hat{z}^{(r)})^2 \quad (44)$$

The denominator in equation 44 is simply to get the normalization right. [Thus when  $\vec{T} = \hat{F}^0 t^0$ , the estimator  $\theta'$  will give  $t^0$ ].

This estimator looks [to me] completely different from the estimator in equation 10. I now show that the two are equivalent. Let me write  $\theta = \vec{a} \cdot \vec{T}$  and  $\theta' = \vec{a}' \cdot \vec{T}$  so

$$\begin{aligned} \vec{a} &\equiv \sum_{j=0}^{N_{\text{fg}}} K^{-1}_{0j} \vec{F}^j \\ \vec{a}' &\equiv \frac{\hat{F}_{\perp}^0}{\hat{F}^0 \cdot \hat{F}_{\perp}^0} \end{aligned} \quad (45)$$

If I can show that these two vectors are equivalent, then I have shown that the two estimators  $\theta$  and  $\theta'$  are also equivalent. One way to do this is to pick a basis which spans the full  $N_{\text{ch}}$  dimensional space and show that for each basis vector  $\vec{b}$ ,  $\vec{a} \cdot \vec{b} = \vec{a}' \cdot \vec{b}$ . As a basis consider the  $N_{\text{fg}}$  vectors  $\vec{F}^j (j > 0)$  together with the  $N_{\text{ch}} - N_{\text{fg}}$  unit vectors  $\hat{z}^{(r)}$ . It is easy to check that  $\vec{a} \cdot \vec{F}^j = \vec{a}' \cdot \vec{F}^j = 0$  for all  $j$ . Of course this is the way the estimators were constructed, to be independent of foregrounds. Now I will show that  $\vec{a} \cdot \hat{z}^{(r)} = \vec{a}' \cdot \hat{z}^{(r)}$  for all  $r$ .

$$\begin{aligned} \vec{a} \cdot \hat{z}^{(r)} &= (K^{-1})_{00} \hat{F}^0 \cdot \hat{z}^{(r)} \\ \vec{a}' \cdot \hat{z}^{(r)} &= \frac{1}{\hat{F}^0 \cdot \hat{F}_{\perp}^0} \hat{F}^0 \cdot \hat{z}^{(r)} \end{aligned} \quad (46)$$

Thus to show that the two estimators are identical, I need only show that

$$\frac{1}{\hat{F}^0 \cdot \hat{F}_{\perp}^0} = (K^{-1})_{00} \quad (47)$$

To prove equation 47 let us calculate

$$\vec{a} \cdot \vec{a} = \sum_{jj'} (K^{-1})_{0j} (K^{-1})_{0j'} \vec{F}^j \cdot \vec{F}^{j'} = (K^{-1})_{00} \quad (48)$$

since  $\vec{F}^j \cdot \vec{F}^{j'} \equiv K_{jj'}$ . Since  $\vec{a}$  is perpendicular to all the foregrounds it lies in  $\mathcal{F}_\perp$ . Thus it can be written as

$$\vec{a} = \sum_r \hat{z}^{(r)} (\vec{a} \cdot \hat{z}^{(r)}). \quad (49)$$

Therefore, another way of writing the dot product in equation 48 is

$$\begin{aligned} \vec{a} \cdot \vec{a} &= \sum_{rr'} \hat{z}^{(r)} (\vec{a} \cdot \hat{z}^{(r)}) \cdot \hat{z}^{(r')} (\vec{a} \cdot \hat{z}^{(r')}) \\ &= \sum_r (\vec{a} \cdot \hat{z}^{(r)})^2 \\ &= (K_{00}^{-1})^2 \sum_r (\hat{F}^0 \cdot \hat{z}^{(r)})^2 \end{aligned} \quad (50)$$

Using equation 48, we can now equate

$$\frac{1}{\sum_r (\hat{F}^0 \cdot \hat{z}^{(r)})^2} = (K^{-1})_{00}. \quad (51)$$

But the left hand side here is equal to the left hand side of equation 47 by definition. So the identity is proven.

## References

- C. L. Bennett et al. 1994, ApJ, 436, 423
- W.N.Brandt, C.R. Lawrence, A.C.S. Readhead, J.N. Pakianathan, & T.M. Fiola, 1994, ApJ, 424, 1
- E. S. Cheng et al. 1994, ApJ, 422, L37
- A. C. Clapp et al. 1994, astro-ph/9404072
- L. Danese, L. Toffolatti, A. Franceschini, M. Bersanelli, & N. Mandolesi 1995, astro-ph/9501043
- S. Dodelson & A. Stebbins 1994, ApJ, 433, 440
- S. Dodelson & A. Kosowsky 1995, PhysRevLett, 75, 604
- S. Dodelson, A. Kosowsky, & S. T. Myers 1995, ApJ, 440, L37
- T. Gaier, J. Schuster, J. Gundersen, T. Koch, M. Seiffert, P. Meinhold, & P. Lubin 1992, ApJ, 398, L1
- K. Ganga, L. Page, E. Cheng, & S. Meyer 1994, astro-ph/9404009
- J. Gundersen et al. 1994, astro-ph/9401220
- S. Hancock, R. D. Davies, A. N. Lasenby, C. M. Gutie'rrez de la Cruz, R. A. Watson, R. Rebolo, & J. E. Beckman 1994, Nature, 367, 333
- P. Meinhold et al. 1993, ApJ, 409, L1
- C. B. Netterfield, N. C. Jarosik, L. A. Page, D. Wilkinson, & E. J. Wollack 1994, astro-ph/9411035
- Neugebauer et al. 1984, ApJ, 278, L1.
- M. Tegmark & G. Efstathiou 1995, astro-ph/9507009
- L. Toffolatti et al. 1994, astro-ph/ 9410037
- E. J. Wollack, N. C. Jarosik, C. B. Netterfield, L. A. Page, & D. Wilkinson 1993, ApJ, 419, L49