Forehead component separation with generalized Internal Linear Combination

Mathieu Remazeilles,* Jacques Delabrouille* and Jean-François Cardoso*

APC 10, rue Alice Domon et Léonie Duquet, 75205 Paris Cedex 13, France

Accepted 2011 July 21. Received 2011 July 21; in original form 2011 March 7

ABSTRACT

The ‘Internal Linear Combination’ (ILC) component separation method has been extensively used to extract a single component, the cosmic microwave background (CMB), from the WMAP multifrequency data. We generalize the ILC approach for separating other millimetre astrophysical emissions. We construct in particular a multidimensional ILC filter, which can be used, for instance, to estimate the diffuse emission of a complex component originating from multiple correlated emissions, such as the total emission of the Galactic interstellar medium. The performance of such generalized ILC methods, implemented on a needlet frame, is tested on simulations of Planck mission observations, for which we successfully reconstruct a low-noise estimate of emission from astrophysical foregrounds with vanishing CMB and Sunyaev–Zel’dovich contamination.

Key words: methods: data analysis – ISM: general – cosmic background radiation.

1 INTRODUCTION

The separation of emissions originating from distinct astrophysical components in observations of the millimetre and submillimetre sky is an important step in the scientific exploitation of such observational data. Various methods have been developed to extract the emission of several components out of multifrequency cosmic microwave background (CMB) observations such as those of the WMAP and Planck space missions (see e.g. Delabrouille & Cardoso 2009 for a review).

In many cases, such methods define components through an (explicit or implicit) assumption that the observations are a linear mixture of unknown templates (or sources) scaling rigidly with frequency, that is,

\[ y_i(p) = \sum_j A_{ij} s_j(p) + n_i(p). \]

(1)

Such methods also assume a fixed number of astrophysical emissions (e.g. CMB anisotropies, thermal Sunyaev–Zel’dovich (SZ) effect, thermal dust emission, synchrotron emission, etc.). The rigid scaling of component emission with frequency is imposed by the fact that the mixing coefficients \( A_{ij} \) depend solely on \( i \) and \( j \) (observation channel and component), and not on the pixel \( p \).

Assuming that such a representation holds, blind component separation methods, such as Spectral Matching ICA (Delabrouille, Cardoso & Patanchon 2003; Cardoso et al. 2008), FastICA (Hyvarinen 1999; Maino et al. 2002), IADE (Cardoso 1998), CCA (Bonaldi et al. 2006) or GMCA (Bobin et al. 2008), are designed to solve the problem of recovering the components of interest when their mixing matrix (the matrix of mixing coefficients, which specifies how much each component contributes to a frequency observation) is unknown. By exploiting the assumption of statistical independence between the components, the mixing matrix can be blindly estimated up to permutation and re-scaling of its columns. Once an estimate of the mixing matrix is available, the components can be separated by inverting the linear system, possibly taking into account the presence of instrumental noise. This has been investigated by a number of authors (Tegmark & Efstathiou 1996; Hobson et al. 1998; Bouchet & Gispert 1999; Delabrouille, Patanchon & Audit 2002).

However, in millimetre and submillimetre wave observations, some components cannot be correctly modelled as a single template which would be simply scaled by mixing coefficients (Tegmark 1998). Emissions from the Galactic interstellar medium (ISM) exhibit frequency scaling which depends on local conditions (temperature, chemical composition) at the location of emission, and hence are variable over the celestial sphere.

Some of the blind component separation methods quoted above can take into account the possible variation of the foreground (FG) frequency scaling as a function of the observed pixel. The CCA method, for instance, can use a pixel-localized model of the FG spectral indices. The Spectral Matching ICA method can be (and has been) implemented on wavelet frames. All methods can be applied independently on several regions of the sky, allowing for a different parameter set in each of the selected regions. However, such localization of the model and of the solution is then the result of prior choices. The number of FG components is fixed, and the regions to be masked, or to be analysed separately, are selected a priori. In reality, the number of relevant components is not well...
known before the data are analysed, and also varies in practice both over the sky and over the scales. For instance, most of the galactic FG components become small, and possibly negligible, at small angular scales and at high Galactic latitude. The number of relevant components in any data set, however, is bound to depend on the level of instrumental noise.

The total FG emission can be separated by subtracting a CMB map estimate from the observation maps. This method has been investigated on the WMAP observations by Ghosh et al. (2011) and employed so far by the Planck Collaboration in their early results (Planck Collaboration et al. 2011). However, the CMB-free maps suffer from excessive noise contamination: as the removed CMB map itself is a low-noise estimate obtained from a minimum variance procedure, most of the instrumental noise ends up in the FG maps, which must then be reprocessed (e.g. filtered) after CMB subtraction. The new component separation method investigated in this paper is an extension of the Internal Linear Combination (ILC) method, aimed at the reconstruction of the total FG emission with the intention of both relaxing the prior assumption on the number of FG components and performing a local processing for best suppression of the contamination of the reconstructed FG components by residual instrumental noise.

Classical ILC methods do not assume a particular parametrization of the FG emission. They offer a simple way to extract the map of a single component of interest and have been used by several authors in the analysis of the maps obtained by the WMAP satellite to extract a CMB map (Bennett et al. 2003; Eriksen et al. 2004; Park, Park & Gott 2007; Kim, Naselsky & Christensen 2008; Delabrouille et al. 2009). The traditional ILC, however, can only recover components for which the emission scales rigidly with frequency (hence its use for separating a CMB map). In addition, the ILC method performs satisfactorily only if the component of interest is not correlated with the other emissions.

In a previous publication (Remazeilles, Delabrouille & Cardoso 2011), we have introduced the Constrained ILC method, which extends the ILC method to the case where there is more than one component of interest (e.g. CMB and thermal SZ), and one wishes to cancel out the contamination from one of them into the recovered map of the other. In this paper, we generalize further the ILC and address the blind separation of multidimensional components which cannot be modelled as a single template scaling with frequency according to a single (pixel-independent) emission law.

2 CMB ESTIMATION BY STANDARD ILC

2.1 Model of the measurement

In all of the following, we assume that all available maps \( N_{\text{obs}} \) can be written, for all pixels \( p \) of the observed maps, in the form

\[
y(p) = ax(p) + b\tilde{z}(p) + f(p) + n(p),
\]

where \( s(p) \) is the CMB template map, \( z(p) \) is the thermal SZ effect, \( f(p) \) is the emission of the rest of the FGs as they would be observed by the instrument in absence of anything else and \( n(p) \) is the instrumental noise. Note that \( f(p) \) and \( n(p) \) are represented with \( N_{\text{obs}} \) maps each, while the CMB and the SZ effect are represented by one single map each, scaled across frequency channels using CMB and SZ scaling coefficients, \( a \) and \( b \), which are assumed to be known.

Depending on the objective, any of \( ax(p) \), \( b\tilde{z}(p) \) or \( f(p) \) can be considered as ‘noise’ and implicitly included in the noise term. Similarly, depending on the objectives of the component separation, \( ax(p) \) or \( b\tilde{z}(p) \) can be considered as part of the total ‘FG term’, that is, implicitly included in \( f(p) \).

2.2 Extraction of the CMB

The ILC method provides the estimate \( \hat{s}_{\text{ILC}} \) of the CMB component \( s \) by forming a linear combination of the \( N_{\text{obs}} \) observed maps which has unit response to the component of interest and has minimum variance. Straightforward algebra leads to

\[
\hat{s}_{\text{ILC}} = \frac{a^\intercal \hat{R}^{-1} y}{a^\intercal \hat{R} a},
\]

where \( \hat{R} \) is the empirical covariance matrix of the observations, \( a \) has dimension \( N_{\text{obs}} \times 1 \) and \( y \) is the \( N_{\text{obs}} \times 1 \) vector of observation maps. This standard ILC can be similarly used to recover an estimate \( \hat{z}_{\text{ILC}} \) of the SZ effect (with \( a \) replaced by \( b \) in equation 3). Note that the quality of CMB estimation with an ILC depends on the accuracy with which \( a \) is known. In the presence of errors (for instance, calibration errors), there is no guarantee that the CMB is preserved (Dick, Remazeilles & Delabrouille 2010).

Assuming no correlations between the components, the total covariance matrix \( R \) of the observations \( y \) can be written as

\[
R = aa^\intercal C_{\text{CMB}} + bb^\intercal C_{\text{SZ}} + R_{\text{FG}} + R_y.
\]

2.3 Wavelet-space ILC

In its simplest implementation, the ILC is performed on the complete maps, and one single global matrix \( \hat{R} \) is used for the whole data set. This requires all maps to be at the same resolution. It is possible, however, to decompose the original maps as sums of different data subsets, each covering a different region in pixel space or in harmonic space, to apply independent versions of the ILC to the different data subsets, and then to re-compose a map from all these independent results.

The main interest of such a decomposition is the possibility to adapt the ILC filter to local contamination conditions. Such localization of the filter is useful in pixel space: the Galactic emissions are stronger in the Galactic plane, whereas noise dominates the total error at high Galactic latitudes. It is also useful in harmonic space, because contaminants do not all have the same angular power spectrum and because of the channel dependence of instrumental beams.

Note, however, that some care must be taken when subdividing the original data into small subsets. The ILC indeed relies on the component of interest to be uncorrelated with the contaminants (i.e. \( \langle s(p)n_i(p) \rangle = 0 \), for all channels of observation \( i \)). If this condition does not hold, the ILC introduces a bias in the reconstruction. This has a consequence on the minimum data size on which the ILC should be implemented: too few independent data points result in empirical correlations between the component of interest and the contaminants, which generate a reconstruction bias as described in the appendix of Delabrouille et al. (2009).

In this paper, observations are decomposed using the spherical needlets discussed, in the context of CMB data analysis, by several authors (see e.g. Pietrobon, Balbi & Marinucci 2006; Fay et al. 2008; Marinucci et al. 2008; Guilloux, Fay & Cardoso 2009). This needlet decomposition provides localization of the ILC filters both in pixel and in harmonic space.
We define a set of spectral windows $h^j(\ell)$ such that, over the useful range of $\ell$, we have
\[
\sum_j \left[ h^j(\ell) \right] = 1. \tag{5}
\]

Maps of wavelet (needlet) coefficients are obtained, for each observed map $y(p)$, by inverse spherical harmonic transform (SHT) of the associated-map SHT coefficients $Y_{\ell m}$ filtered by the spectral windows $h^j(\ell)$:
\[
y^{(j)}(p) = \sum_{\ell m} Y_{\ell m} h^j(\ell) Y_{\ell m}(p). \tag{6}
\]

For each scale $j$, for each pixel $p$ of the corresponding needlet coefficient maps $y^{(j)}$ (one such map for each observation $a$), the empirical covariance matrix $\mathbf{R}$ used in equation (3) is computed from an average, in a domain $D_p$ centred at pixel $p$ and including some neighbouring pixels, of the product of needlet coefficients. The $ab$ entry is given by
\[
\mathbf{R}_{ab}(p) = \frac{1}{N_p} \sum_{p' \in D_p} y^{(a)}(p') y^{(b)}(p'). \tag{7}
\]

In Delabrouille et al. (2009), the practical implementation uses, as domains $D_p$, HEALPix ‘superpixels’ obtained by grouping $32 \times 32$ pixels of the needlet coefficient maps $y^{(j)}(p)$ (making use of the hierarchical definition of HEALPix pixels). Here, we use a slightly different prescription: we smooth instead the product map $\gamma^{(a)}(p) \gamma^{(b)}(p)$ with a symmetric, Gaussian window in pixel space. This avoids artificial discontinuities at superpixel edges.

## 3 Foreground estimation by multidimensional ILC

We now address the problem of estimating the set of maps $f(p)$, that is, a ‘catch-all’ FG component comprising the emission of the diffuse Galactic ISM, and of numerous Galactic and extragalactic compact sources. The objective is to construct estimated maps $\hat{f}(p)$ that ‘best match’ what would be observed by the instrument in the absence of CMB, SZ effect and noise (see equation 2).

Astrophysical emission originating from the Galactic ISM and from numerous extragalactic sources is qualitatively different from the CMB and the SZ effect. Each of the latter is somewhat special in the sense that their emission can be modelled, with good accuracy, as a single template scaling in a known way with frequency. The total FG emission $f$ comprises contributions from several different processes. In addition, we cannot even assume a priori that a linear superposition of well-defined templates) does hold.

For extracting such emissions from multifrequency observations, we propose to generalize the ILC method to address the case of such a ‘multidimensional component’.

### 3.1 Multidimensional components

Let $\mathbf{R}_{\text{FG}} = (f f')$ denote the covariance matrix of the observed FGs in $N_{\text{obs}}$ frequency channels. This $N_{\text{obs}} \times N_{\text{obs}}$ matrix $\mathbf{R}_{\text{FG}}$ will be referred to as the FG covariance matrix.

Among astrophysical FGs included in the ‘catch-all’ component $f$, the ISM of our own Galaxy is the main contributor. It emits via the combination of several processes (synchrotron, free–free, thermal dust, ‘anomalous’ dust emission, molecular lines, etc.). In previous work, some of these processes have been individually modelled by a fixed template and an emission law. Bouchet & Gispert (1999), for instance, assume that synchrotron emission scales with frequency proportionally to $v^{-0.9}$, free–free proportionally to $v^{-0.10}$, and dust proportionally to $v^2 B_\nu(T)$ with $T = 18$ K. Since the emission of the ISM in each channel is described as a linear mixture of three templates, such a model predicts that the ISM covariance matrix (which we will denote by $\mathbf{R}_{\text{ISM}}$) is a rank 3 matrix. When the contribution of extragalactic compact sources is neglected (assuming bright point sources are extracted from the maps, and faint ones contribute a negligible amount of emission), the FG covariance matrix itself, $\mathbf{R}_{\text{FG}}$, is equal to $\mathbf{R}_{\text{ISM}}$ (the covariance matrix of the Galactic ISM emission) and is a rank 3 matrix.

Such a model is too crude in the context of the very sensitive measurements performed by Planck and the WMAP: the emission laws of the Galactic emissions vary as a function of the direction on the sky. To make things even more complex, the background of compact sources contributes emission that becomes significant for measurements such as those of the Planck mission (Planck Collaboration et al. 2011), and that cannot be modelled at all as the sum of a few independent components.

The question of the rank of the FG covariance matrix is a crucial one for component separation. This matrix is expected to be, strictly speaking, full rank. In practice, however, the issue is slightly more subtle. Consider its eigendecomposition: $\mathbf{R}_{\text{FG}} = \mathbf{V} \mathbf{D} \mathbf{V}^T$, where $\mathbf{V}$ is an orthonormal matrix and $\mathbf{D}$ is a diagonal matrix with eigenvalues sorted in decreasing order. While the three-component model of Bouchet & Gispert (1999) predicts that only the first three eigenvalues are non-zero, a model with spatially varying spectral indices, and numerous additional emission processes (‘anomalous’ dust emission, molecular lines, extragalactic source background), predicts that all the eigenvalues are non-zero. However, if there is only a small variation in the spectral indices, and if some components are very weak, it is expected, at least in some regions of the sky or at some angular scales, that the smallest eigenvalues are very close to zero so that $\mathbf{R}_{\text{FG}}$ is ‘almost rank-deficient’ (see Section 3.4 below for a more rigorous statement).

In this paper, we propose, as in Cardoso et al. (2008), to model the FG covariance matrix as an $N_{\text{obs}} \times N_{\text{obs}}$ matrix of rank $N_{\text{FG}}$, not necessarily equal to $N_{\text{obs}}$. Loosely speaking, $N_{\text{FG}}$ counts the number of different templates needed to represent most of the emission of the FG in our data set. In other words, we try to capture all FG emission as resulting from $N_{\text{FG}}$ (possibly correlated) templates. The integer $N_{\text{FG}}$ is called the (effective) FG dimension and may vary over the sky with respect to the pixel $p$ or with respect to $\ell$ in harmonic space or with respect to the needlet domain considered for a decomposition of the maps on a needlet frame.

#### 3.2 The foreground subspace

The analysis in this paper is performed on a needlet frame. The temperature map needlet coefficients are indexed by $(j, k)$, where $j$ denotes the scale and $k$ the pixel. In a given needlet domain $D_k^{(j)}$, if the FG covariance matrix $\mathbf{R}_{\text{FG}}$ is an (symmetric, non-negative) $N_{\text{obs}} \times N_{\text{obs}}$ matrix of rank $N_{\text{FG}}$, then FG emission can be represented as a superposition of $N_{\text{FG}}$

1 Note that the methods described throughout this paper do not require a needlet frame in particular and can be implemented in pixel space as well, where domains $D$ should be indexed by pixels $p$, or in harmonic space with the domains indexed by $(\ell, m)$ coefficients.
templates:

\[ f = F g. \]

where the \( N_{\text{obs}} \times N_{\text{FG}} \) matrix \( F \) is called the FG mixing matrix and \( g \) is a vector of dimension \( N_{\text{FG}} \). It follows that the FG covariance matrix is

\[ \mathbf{R}_{\text{FG}} = \langle f f' \rangle = (F g') F' = F \mathbf{G} F', \]

where \( G \) is an \( N_{\text{FG}} \times N_{\text{FG}} \) full-rank covariance matrix.

Note two important points. First, the templates \( g \) are not expected to correspond to physical FGs. They are just a basis of the \( N_{\text{FG}} \)-dimensional subspace spanned by \( f \). We are not interested in recovering \( g \). Our objective with the method discussed here is to recover \( f \) (in addition to \( s \) and \( z \)). Secondly, matrix \( F \) and its number \( N_{\text{FG}} \) of columns may depend on the domain \( D^{(n)} \) considered. For instance, at high Galactic latitude, it is quite possible that our observations contain negligible emission from some of the Galactic FGs, but not so at low Galactic latitude. The needlet implementation allows us to modulate the effective dimension of the multidimensional FG component both in pixel space and in harmonic space, that is, to vary \( N_{\text{FG}} \) across the sky regions and the physical scales.

We also stress that, unlike the case of CMB and SZ reconstruction, where the mixing vectors \( a \) and \( b \) are assumed fully known a priori, we do not assume here that matrix \( F \) is known. We will not resort to prior physical knowledge about the components of the FG emission to determine matrix \( F \). In fact, as the basis templates \( g \) do not correspond to anything physically meaningful, we are not even interested in determining \( F \) itself but, for reconstruction purposes, only the product \( f = F g \). It is only assumed that matrix \( R_{\text{FG}} \) has a given rank \( N_{\text{FG}} \) (which can be estimated from the data, if needed) in the needlet domain.

Matrix \( F \) cannot be determined from the data only, that is, without making use of some prior information or assumption about \( g \). Indeed, let \( T \) be some invertible \( N_{\text{FG}} \times N_{\text{FG}} \) matrix and consider the transformed matrices \( \tilde{F} = F T^{-1} \) and \( \tilde{G} = T G T^{-1} \). These transformed matrices are an alternate, completely equivalent, factorization of the FG covariance matrix, since, by construction, \( \mathbf{F} \mathbf{G} \mathbf{F}' = \tilde{\mathbf{F}} \tilde{\mathbf{G}} \tilde{\mathbf{F}}' \). Physically, it means that the \( N_{\text{FG}} \) underlying templates \( g \) can be replaced by any other linear combination \( Tg \) of them (provided the linear combination is not degenerate, that is, \( T \) is invertible).

However, as we will see in Section 3.3, the implementation of the ILC filter for estimating the total FG emission does not require the full knowledge of \( F \). Indeed, the expression of that filter is strictly unchanged upon the introduction of such an invertible factor \( T \). In Section 3.4, we show how matrix \( F \) can be estimated up to multiplication by a right factor \( T \). It is worth stressing again that this indetermination means that we are only concerned with estimating the column space of matrix \( F \) [denoted by \( \text{Col}(F) \) throughout this paper]. That \( N_{\text{FG}} \)-dimensional space can be called the ‘FG subspace’. Our working assumption that \( R_{\text{FG}} \) has rank \( N_{\text{FG}} \) means that the FG data have a covariance structure which is unknown but is constrained to live in the FG subspace.

Physically, accepting this indetermination amounts to giving up, during the component separation stage discussed here, distinction between processes of emission on the basis of physical criteria such as emission process or physical origin. Obviously, this is not fully satisfactory from an astrophysicist’s point of view, since in the end we would like to know what is the source of the observed emission. This distinction between sources of FG emission, however, can be made at a later stage of the data analysis, that is, we can first separate the CMB and SZ effect from other FGs, and then put physics into the interpretation of the reconstructed multidimensional FG component and interpret it as the sum of emissions from a number of physical emission processes.

### 3.3 Multidimensional ILC filter

Aiming at a direct estimation of the FGs, we generalize the ILC method to address the case of a multidimensional component (here \( N_{\text{FG}} \) dimensional, where \( N_{\text{FG}} \) is the number of components, that is, the rank of the FG covariance matrix). In a given needlet domain, we model the observation maps, collected into the \( N_{\text{obs}} \times 1 \) vector \( y \), as

\[ y = Ax + n, \]

where \( n \) is the \( N_{\text{obs}} \times 1 \) vector of instrumental noise and

\[ A = \begin{bmatrix} a \end{bmatrix}, \quad x = \begin{bmatrix} s \\ g \end{bmatrix}. \]

Note that no rigid scaling with frequency is assumed since all these quantities are needlet-dependent, that is, they depend both on the scale considered and on the pixel. Here the \( (N_{\text{FG}} + 1) \times 1 \) signal vector \( x \) contains the CMB emission \( s \) as the first entry and the \( N_{\text{FG}} \times 1 \) vector \( g \) which collects the emission of the \( N_{\text{FG}} \) components needed to model the total FG emission. The \( N_{\text{obs}} \times (N_{\text{FG}} + 1) \) mixing matrix \( A \) contains, as the first column, the \( N_{\text{obs}} \times 1 \) vector \( a \) giving the frequency scaling of the CMB component. The other columns correspond to the \( N_{\text{obs}} \times N_{\text{FG}} \) FG mixing matrix \( F \), that is, they span the FG subspace. Note that this assumes that \( a \) itself cannot be obtained by linear combinations of the columns of \( F \) (more about this later).

As a refinement, it can be useful to single out both the CMB and the SZ effect, in which case the second column in \( A \) explicitly appears as the frequency scaling vector \( b \) of the SZ component (and the SZ component can be considered as excluded from the rest of the FGs). We get back to this refinement in Sections 3.5 and 4.

Equation (10) assumes that all observations are at the same resolution, which is needed to implement the ILC filter (for practical implementation, maps are put to the same resolution by partial deconvolution in harmonic space). The localization in harmonic space allows dropping out some of the channels at high \( \ell \), if needed, by reason of insufficient resolution.

We consider the estimation of \( f \) by a linear operation

\[ \hat{f} = B y, \]

where, as in standard (one-dimensional) ILC, the \( N_{\text{obs}} \times N_{\text{obs}} \) ILC weight matrix \( B \) is designed to offer unit gain to the FGs while minimizing the total variance of the vector estimate \( \hat{f} \). In other words, matrix \( B \) is the minimizer of \( E(||B y||^2) \) under the constraint \( BF = F \). The weight matrix \( B \) thus solves the following constrained variance minimization problem:

\[ \min_{B, F} \text{Tr}(B B') \text{Tr}(B F F') \]

\[ \text{subject to } BF = F, \]

where \( R \) is the covariance matrix of the observations \( y \) and \( B \) is the matrix trace operator. That problem can be solved by introducing a Lagrange multiplier matrix \( A \) and the Lagrangian

\[ L(B, A) = \text{Tr}(B B') - \text{Tr}(A' (BF - F)). \]

Differentiating equation (14) with respect to \( B \), one finds that

\[ \partial L(B, A)/\partial B = 0 \]

entails

\[ 2BR = A F'. \]
By imposing the constraint $BF = F$ on equation (15), one then finds that $\Lambda = 2F(FR^{-1}F)^{-1}$. Hence, the solution of problem (13) is the FG ILC weight matrix given by

$$B = F(FR^{-1}F)^{-1}FR^{-1}.$$  

(16)

Comparing equation (16) with equation (3), multidimensional ILC appears as a direct generalization of one-dimensional ILC.\(^2\)

One can immediately note that expression (16) for $B$ is invariant if $F$ is changed into $FT$ for any invertible matrix $T$. Hence, as already mentioned in Section 3.2, implementing the FG ILC filter (16) only requires that the FG mixing matrix $F$ be known up to right multiplication by an invertible factor. Again, in other words, the only meaningful and mandatory quantity for implementing a multidimensional ILC is the column space of $F$.

### 3.4 Estimation of the foreground subspace

In this section, we propose a method for estimating the FG subspace locally, that is, in each needlelet domain. We consider only the case where the model accounts for the CMB, an $N_{FG}$-dimensional FG component and noise at a known level:

$$R = C_{CMB}a + FG^f + R_N$$  

(17)

and we want to estimate the FG subspace $\text{Col}(F)$ from an estimate $\hat{R}$ of $R$. Define the $N_{\text{obs}} \times (N_{FG} + 1)$ matrix:

$$L = \left[ a^{1/2}_{CMB} \; \text{FG}^{1/2} \right],$$  

(18)

where the first column $a^{1/2}_{CMB}$ of $L$, containing the FG frequency-scaling vector (which is known), is distinguished from the $N_{FG}$ unknown remaining columns $\text{FG}^{1/2}$ associated to the FGs. Thus, the signal part of the covariance matrix is $LL^t$:

$$\hat{R} = LL^t + R_N.$$

Our procedure for estimating the column space $\text{Col}(F)$ of the FG mixing matrix $F$ consists of two steps. In the first step, we obtain an estimate of $L$ up to right multiplication by a rotation matrix ($\Lambda$) and (an estimate for the dimension $N_{FG}$ of the FG subspace) using the knowledge of the noise covariance matrix. In the second step, we use the fact that the first column of $L$ is known (up to scale) to obtain an estimate of $\text{Col}(F)$. This is described next.

Denote the eigenvalue decomposition of the noise-whitened signal covariance matrix $R_N^{1/2}LL^tR_N^{-1/2}$ as

$$R_N^{1/2}LL^tR_N^{-1/2} = U\Delta U^t,$$

where $U$ is orthonormal, $UU^t = I$, and $\Delta$ is diagonal. Now,

$$R_N^{1/2}LL^tR_N^{-1/2} = R_N^{1/2}(LL^t + R_N)R_N^{-1/2} = R_N^{1/2}LL^tR_N^{-1/2} + I$$

$$= U\Delta U^t + U\Delta + I,$$

showing that $R_N^{1/2}LL^tR_N^{-1/2}$ and $R_N^{1/2}LL^tR_N^{-1/2}$ share the same eigenvectors but that the former has eigenvalues shifted by 1 with respect to the latter. Further, if $L$ has rank $N_{FG} + 1$, then its eigenstructure actually is

$$U\Delta U^t = [U, U_n] \begin{bmatrix} \Delta \Omega & 0 \\ 0 & 0 \end{bmatrix} [U, U_n]^t.$$  

(19)

\(^2\)The ILC estimate of the CMB vector of emission in each frequency channel is obtained by applying the filter $a \left( a^tR^{-1}a \right)^{-1}a^tR^{-1} \left[ \text{i.e. the filter of equation (3) multiplied on the left-hand side by the vector } a \right]$.\(^3\)

where $U$, has $(N_{FG} + 1)$ columns, $U$, has $N_{\text{obs}} - (N_{FG} + 1)$ columns, and $\Delta \Omega$ is a $(N_{FG} + 1) \times (N_{FG} + 1)$ diagonal matrix.

### 3.4.1 Estimation of $N_{FG}$ and $L$

In the needlelet domain considered, given an estimate $\hat{R}$ of $R$, we compute the eigendecomposition:

$$R_N^{1/2}LL^tR_N^{-1/2} = \hat{UD}\hat{U}^t$$  

(20)

and, similar to equation (19), we denote by $\hat{D}$ the sub-block of $\hat{D}$ corresponding to the eigenvalues that are larger than $(1 + \epsilon)$ and by $\hat{U}$ the corresponding subset of columns of $U$. Here $\epsilon$ is a threshold above which the observation is not dominated by instrumental noise (see Section 4 for the choice of the threshold). This threshold condition thus provides an estimate for the dimension $N_{FG}$ of the FG subspace in the needlelet domain, given the dimension $(N_{FG} + 1)$ of the sub-block $\hat{D}$ which fulfills the threshold condition.

In this pre-processing, the dimension of the estimated signal subspace, $(N_{FG} + 1)$, depends on the level of noise in the needlelet domain considered, so the signal subspace is estimated locally, both in space and in scale. This processing thus locally performs a rank reduction of the observations covariance matrix, allowing the reduction of the instrumental noise in the reconstruction.

By construction, the matrix

$$\hat{M} = R_N^{1/2}\hat{U}\hat{S}(\hat{D} - I)^{1/2}$$  

(21)

is such that $R_N^{1/2}\hat{M}\hat{M}^tR_N^{1/2}$ is close to $R_N^{1/2}LL^tR_N^{1/2}$ if $\hat{R}$ is close to $R$. That property, in turn, implies that $\hat{M}O$ is close to $L$ for some (undetermined) rotation matrix $O$, completing the first step of our estimation procedure.

### 3.4.2 Estimation of the foreground column space $\text{Col}(F)$

In the second step, we note that the rotation matrix $O$ should be such that $O \hat{M}$ is close to $L$. However, only the first column of $L$ is known, up to scale. Hence, we partition $O$ as $O = [vl]$, where $v$ is a unit norm vector and $l$ is an $(N_{FG} + 1) \times N_{FG}$ matrix. The only available constraint is thus that $\hat{M}v$ should be close to the first column of $L$. However, we cannot expect to find a $v$ such that $\hat{M}v$ is strictly equal to $a^{1/2}_{CMB}$ because $\hat{M}$ is estimated from the data so that $Col(\hat{M})$ does not necessarily contain $a$ (as would be the case for $\hat{R} = R$). The best we can do is to determine $v$ such that $\hat{M}v$ is equal to the projection of $a^{1/2}_{CMB}$ on $\text{Col}(\hat{M})$. The orthogonal projection matrix is $\hat{M}(\hat{M}\hat{M})^{-1}\hat{M}$ so that the projection is

$$\hat{M}(\hat{M}\hat{M})^{-1}\hat{M}a^{1/2}_{CMB}.$$

Let us then denote the vector $\bar{a}$ by

$$\bar{a} = \left( \hat{M}a^{1/2}_{CMB} \right)^{-1}\hat{M}a.$$  

(22)

The projection of $a^{1/2}_{CMB}$ on to $Col(\hat{M})$ then is $\hat{M}\bar{a} a^{1/2}_{CMB}$ and the vector $v$ is therefore given by $v = \bar{a} a^{1/2}_{CMB}$. Recall that $\bar{v}$ is a unit norm vector, so we must have

$$v = \bar{v}/|\bar{v}| \quad \text{and} \quad C_{CMB} = 1/|\bar{v}|^2.$$  

(23)

Once vector $v$ is determined, the constraint that $O = [v | V]$ is a rotation matrix uniquely determines $V$ up to right multiplication by a rotation factor. However, nothing more is required to determine the FG subspace, as already stressed. Our procedure is therefore complete and can be summarized by the following steps:

---

© 2011 The Authors, MNRAS 418, 467–476
(i) Compute the eigendecomposition (20) of the noise-whitened covariance matrix and obtain an estimate of \(N_{FG}\) from comparing the level of the eigenvalues with the noise level.

(ii) Form matrix \(\mathbf{M}\) by equation (21), compute vector \(\hat{a}\) by equation (22) and get \(\mathbf{v}\) by normalization (23).

(iii) Compute an \((N_{FG} + 1) \times N_{FG}\) matrix \(\mathbf{V}\) such that matrix \([\mathbf{v} \mid \mathbf{V}]\) is a rotation.

(iv) Obtain a basis of the FG subspace as \(\mathbf{F} = \mathbf{M} \mathbf{V}\).

(v) Compute the \(N_{FG}\)-dimensional ILC filter

\[
\mathbf{B} = \mathbf{F} (\mathbf{R}^{-1} \mathbf{F})^{-1} \mathbf{F} \mathbf{R}^{-1}
\]

### 3.5 Projecting foregrounds orthogonally to both thermal SZ and CMB emission

The FG multidimensional ILC filter can be generalized further. Thermal SZ emission can be singled out in the same way as the CMB emission, in which case we may require that there is no thermal SZ residual in the reconstructed FG map. This is doable because the emission law of the SZ component, like that of the CMB, is known. We then write the model of emissions as

\[
\mathbf{A} = \begin{bmatrix} a & b & \mathbf{F} \end{bmatrix}, \quad x = \begin{bmatrix} s \\ z \\ g \end{bmatrix},
\]

where we have explicitly distinguished the thermal SZ emission \(z\) from the other FGs through its frequency scaling vector \(b\) (emission law). We may then generalize the processing developed in Section 3.4. In that spirit, the FG mixing matrix \(\mathbf{F}\) can then be estimated in the needlet domain considered from the set of \(N_{FG}\) columns orthogonal to both the projection of the CMB scaling vector and the projection of the thermal SZ scaling vector on to the estimated \((N_{FG} + 2)\)-dimensional signal subspace. This guarantees that the FG map reconstructed by the multidimensional ILC now contains neither SZ effect nor CMB (with, however, the usual caveat that the statistics used to compute the covariance matrices must be accurate enough). In addition, the rank-reduction procedure [restriction of the observations to the \((N_{FG} + 2)\)-dimensional signal subspace] in each needlet region reduces the level of instrumental noise locally in the reconstructed FGs.

### 3.6 Discussion of special cases

#### 3.6.1 Less channels than foreground dimension

In the discussion above, we assumed that the signal subspace is the direct sum of two subspaces: the CMB subspace which is one dimensional (because of the rigid scaling of the CMB with frequency) and the FG subspace which is \(N_{FG}\) dimensional. The former is not included in the latter if no combination of FG emission has the same scaling as the CMB across available frequencies. Of course, this property requires enough properly chosen frequency channels.

When there are more components than observations, then, unless the FG emissions are either fully correlated or very faint (below noise), we have \(N_{FG} < 1 > N_{obs}\) and the CMB cannot be perfectly separated from the FGs. When there are enough independent observations (i.e. a large number of channels), then \(N_{FG} < N_{obs}\), and in general the CMB subspace is not contained in the (larger) FG subspace. Separation is then possible up to finite-sample-size errors in the determination of the appropriate subspaces.

Note in the passing that the kinetic SZ effect cannot be separated from the CMB on the basis of its emission law. Throughout this paper, the CMB, distinguished solely by its known emission law \(a\), also includes the kinetic SZ effect.

#### 3.6.2 CMB subtraction as an \((N_{obs} - 1)\)-dimensional ILC

A FG estimation has been obtained in Ghosh et al. (2011) by subtracting the CMB ILC estimate from the observation data.

\[
f = y - \hat{f} = y - a \hat{s}.
\]

It is interesting to note that the CMB-subtraction procedure is equivalent to an \((N_{obs} - 1)\)-dimensional ILC filtering, that is, the particular multidimensional ILC filtering where the dimension of the FG subspace is assumed to be constant over the whole sky and the whole range of scales, and equal to \((N_{obs} - 1)\). Indeed, the CMB-subtracted estimate expands as follows:

\[
\hat{f} = y - a \hat{a} \mathbf{R}^{-1} y + a \mathbf{R}^{-1} a.
\]

\[
\hat{f} = \mathbf{W}^{-1} \left( \mathbf{I} - \mathbf{W} \mathbf{a} \left( \mathbf{W} \mathbf{a} \right)^{-1} \left( \mathbf{W} \mathbf{a} \right) \right) \mathbf{W} y,
\]

\[
\mathbf{W}^{-1} \left( \mathbf{I} - \mathbf{P}_1 \right) \mathbf{W} y,
\]

where \(\mathbf{W} = \mathbf{R}^{-1/2}\) denotes the inverse square root of the data covariance matrix.

Matrix \(\mathbf{P}_1 = \mathbf{W} \mathbf{a} \left( \mathbf{W} \mathbf{a} \right)^{-1} \left( \mathbf{W} \mathbf{a} \right)\) is an orthogonal projection (\(\mathbf{P}_1 = \mathbf{P}_1\) and \(\mathbf{P}_1^T = \mathbf{P}_1\) on to the line \(\mathbf{Span} \{\mathbf{W} \mathbf{a}\}\) (one-dimensional ‘whitened’ CMB subspace). It implies that \(\mathbf{I} - \mathbf{P}_1 = \mathbf{P}_H\) is the projection on to the \((N_{obs} - 1)\)-dimensional hyperplane \(\mathbf{H} = \mathbf{Span} \{\mathbf{W} \mathbf{a}\}\), which is orthogonal and complementary to the one-dimensional whitened CMB subspace. Let us denote \(\mathbf{v} = \mathbf{W} \mathbf{a} / \left| \mathbf{W} \mathbf{a} \right|\) and consider an \(N_{obs} \times (N_{obs} - 1)\) matrix \(\mathbf{V}\) such that \([\mathbf{v} \mid \mathbf{V}]\) is a rotation. Then, \(\mathbf{P}_H = \mathbf{V} \left( \mathbf{V}^T \mathbf{V} \right)^{-1} \mathbf{V}^T\) and, by denoting \(\mathbf{F} = \mathbf{W}^{-1} \mathbf{V}\), we get

\[
\hat{f} = \mathbf{W}^{-1} \mathbf{P}_H \mathbf{W} y,
\]

\[
\mathbf{W}^{-1} \mathbf{V} \left( \mathbf{V}^T \mathbf{V} \right)^{-1} \mathbf{V} \mathbf{W} y,
\]

\[
\mathbf{F} \left( \mathbf{F}^T \mathbf{F} \right)^{-1} \mathbf{F} \mathbf{R}^{-1} y,
\]

which completes the proof since \(\mathbf{F}\) is full rank \((N_{obs} - 1)\). Here, it is interesting to note that the \((N_{obs} - 1)\)-dimensional ILC can be obtained without even knowing the mixing matrix \(\mathbf{F}\) since the procedure becomes equivalent to the estimation obtained by subtracting the CMB ILC estimate from the observation data.

This equivalence means that the CMB-subtraction procedure does not take advantage of the fact that the FG mixing matrix can be almost rank-deficient in some regions of the sky or at some scales (for

---

**Figure 1.** The spectral bands used in this work for the definition of the needlets.
instance, at small scale where the instrumental noise is dominant). The \((N_{\text{obs}} - 1)\)-dimensional subspace \(\text{Col}(F)\) reconstructed here thus includes both noise and FG components. Consequently, such a FG reconstruction is noisy. The \(N_{\text{FG}}\)-dimensional ILC procedure described in Section 3.4 performs a cleaner FG reconstruction (in terms of signal-to-noise ratio) because the effective rank \(N_{\text{FG}}\) of the FG subspace and the FG mixing matrix (with reduced rank) are estimated locally in each needlet domain. In effect, this boils

\[\text{FG component separation with generalized ILC}\]
down to performing at the same time both component separation and denoising by thresholding the needlet coefficients.

4 PLANCK SIMULATIONS AND RESULTS

We now turn to illustrating this discussion with examples based on simulated data sets. We apply our multidimensional FG ILC filter on a frame of needlets. The spectral bands used in the definition of the needlets are shown in Fig. 1. For each needlet domain considered, we project the data both on to the ‘full-rank’ FG subspace (equivalent to simple CMB ILC subtraction) and on to a ‘reduced-rank’ FG subspace. For the latter, we reject the eigenvalues of the covariance matrix of the observation needlet coefficients smaller than 1.25 times the noise covariance level, that is, values for which the instrumental noise contributes more than 80 per cent of the total emission. This is a somewhat arbitrary, but reasonable, choice chosen here for illustration. In practice, this threshold can be fixed more rigorously, considering the trade-off between rejecting low-level FG emission and letting in the final map too much instrumental noise.

As a refinement, the number of relevant FG components could be estimated without even imposing any arbitrary threshold, for example, by using the Akaike information criterium (Akaike 1974). This criterium consists in maximizing the likelihood of the observations, given the model, taking into account a particular penalty imposed on the number of free parameters entering in the model (e.g. the number of FG components or, equivalently, the rank of the FG mixing matrix).

Our investigations are carried out on sky temperature simulations generated with the PSM (Planck Sky Model) version 1.6.6. Sky simulations include Gaussian CMB generated assuming a theoretical angular spectrum fitting the WMAP observations, thermal and kinetic SZ effect, four components of Galactic ISM emission including thermal and spinning dust, synchrotron and free–free, and emission from point sources (radio and infrared). The resolution and noise level of the observations correspond to nominal mission parameters as described in the Planck ‘Bluebook’. Nine frequency channels are used in this simulation and correspond to the Planck High Frequency Instrument and Low Frequency Instrument channels. Details about PSM simulations can be found in Leach et al. (2008) and Betoule et al. (2009).

Fig. 2 shows the ‘observed’ 70-GHz map, the input FG map at 70 GHz and the thermal SZ map, all at the resolution of the 70-GHz channel. Our maps are centred on an interesting region which is both at low Galactic latitude, around (l, b) = (72°, −8°), and close to a set of bright galaxy clusters. The 70-GHz reconstructed FGs, recovered by multidimensional ILC filtering, are shown in the same

![Figure 4. FG reconstruction at 70 GHz. Left-hand panels: (N_{obs} − 1)-dimensional needlet ILC output map and needlet error (input−output) maps. Right-hand panels: N_{FG}-dimensional needlet ILC output map and needlet error (input−output) maps. The N_{FG}-dimensional needlet ILC guarantees the reduction of the noise contamination.](https://academic.oup.com/mnras/article-abstract/418/1/467/962882/)

![Figure 5. Power spectrum of the recovered FGs at 70 GHz and of the ILC reconstruction error: (N_{obs} − 1)-dimensional FG ILC (solid red curve) and error (dashed red curve), and N_{FG}-dimensional FG ILC (solid blue curve) and error (dashed blue curve). We clearly see the suppression of the noise at high ℓ for the N_{FG}-dimensional FG ILC reconstruction.](https://academic.oup.com/mnras/article-abstract/418/1/467/962882/)

© 2011 The Authors, MNRAS 418, 467–476

region of the sky in the top panels of Fig. 3. The corresponding reconstruction error (difference between reconstructed output and original input) is displayed in the bottom row of the same figure. The \( (N_{\text{obs}} - 1) \)-dimensional ILC reconstruction (left-hand panels), equivalent to a simple subtraction of the ILC estimate of the CMB map to the observation map, is clearly noisy (as expected from the discussion of Section 3.6.2). The reduction in the noise in the FG reconstruction is achieved by performing an \( N_{\text{FG}} \)-dimensional ILC reconstruction (middle panels), where \( N_{\text{FG}} \) is the local dimension of the FG subspace depending both on the needlet scale and on the pixel. We observe the leakage of a thermal SZ emission in the FG reconstruction in the left-hand and middle panels of Fig. 3. Using the modified ‘reduced-rank’ ILC introduced in Section 3.5, we obtain instead the reconstruction displayed in the right-hand panels of Fig. 3 with no visible contamination by SZ emission.

For completeness, the same results are shown on full sky maps in Fig. 4 and we have plotted the corresponding power spectra in Fig. 5. The suppression of the noise contamination is clearly visible on the spectrum at high \( \ell \) when an \( N_{\text{FG}} \)-dimensional ILC method is employed.

Fig. 6 shows, for the fifth needlet band (scales comprised between \( \ell = 512 \) and 1100), the effective number \( N_{\text{FG}} \) of FG components (i.e. the effective rank of the FG covariance matrix) which has been estimated in each needlet domain from the 80 per cent noise threshold (FG components contributing less than 20 per cent of the total emission in the needlet domain have been neglected). In the right-hand panel of the figure, a 12.5 \( \times \) 12.5 patch of the sky at low Galactic latitude, centred around \( (l, b) = (72^\circ, -8^\circ) \), explicitly shows the effective number of FG components estimated in this region. This number decreases according to the distance from the Galactic plane. This is consistent with the bottom right-hand panel of Fig. 3 showing that the residual noise is locally distributed and decreases according to the distance from the Galactic plane.

5 CONCLUSIONS

In this paper, we have shown how the standard ILC procedure, originally dedicated to the CMB extraction, can be extended for the reconstruction of complex astrophysical emissions, beyond the CMB alone. We have developed generalized ILC filters (multidimensional ILC) to reconstruct the diffuse emission of a complex multidimensional component originating from multiple correlated emissions, such as the total Galactic FG emission. Similar, though pixel based, extensions have also been implemented in a fastICA-based code, altICA, as used in Leach et al. (2008) and are integrated in the Planck LFI Data Processing Center pipeline (C. Baccigalupi and R. Stompor, private communication). Our estimators were implemented on a needlet frame and tested on simulations of Planck observations. This new ILC filtering successfully reconstructs the FG emission, exempt from both the CMB and the SZ emission, and with a reduced level of instrumental noise.

ACKNOWLEDGMENTS

We thank Tuhin Ghosh, Radek Stompor and Carlo Baccigalupi for useful conversations related to this work. Some of the results in this paper have been derived using the HEALPix package (Górski et al. 2005). We also acknowledge the use of psm developed by the Component Separation Working Group (WG2) of the Planck Collaboration.

REFERENCES
