Application of beam deconvolution technique to power spectrum estimation for CMB measurements

E. Keihänen,[1,2]★ K. Kiiveri,[1,2] H. Kurki-Suonio[1,2] and M. Reinecke[3]★

[1]*Department of Physics, Gustaf Hällströmin katu 2, University of Helsinki, FI-00014, Helsinki, Finland*
[2]*Helsinki Institute of Physics, Gustaf Hällströmin katu 2, University of Helsinki, FI-00014, Helsinki, Finland*
[3]*Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, D-85741 Garching, Germany*

ABSTRACT

We present two novel methods for the estimation of the angular power spectrum of cosmic microwave background (CMB) anisotropies. We assume an absolute CMB experiment with arbitrary asymmetric beams and arbitrary sky coverage. The methods differ from the earlier ones in that the power spectrum is estimated directly from the time-ordered data, without first compressing the data into a sky map, and they take into account the effect of asymmetric beams. In particular, they correct the beam-induced leakage from temperature to polarization. The methods are applicable to a case where part of the sky has been masked out to remove foreground contamination, leaving a pure CMB signal, but incomplete sky coverage. The first method (deconvolution quadratic maximum likelihood) is derived as the optimal quadratic estimator, which simultaneously yields an unbiased spectrum estimate and minimizes its variance. We successfully apply it to multipoles up to $\ell = 200$. The second method is derived as a weak-signal approximation from the first one. It yields an unbiased estimate for the full multipole range, but relaxes the requirement of minimal variance. We validate the methods with simulations for the 70 GHz channel of *Planck* surveyor, and demonstrate that we are able to correct the beam effects in the TT, EE, BB and TE spectra up to multipole $\ell = 1500$. Together, the two methods cover the complete multipole range with no gap in between.

Key words: methods: data analysis – methods: numerical – cosmic background radiation.

1 INTRODUCTION

Present-day cosmic microwave background (CMB) experiments require accurate methods for the estimation of the angular power spectrum of the CMB. Several methods have been developed for this purpose. The methods fall roughly into two categories, according to the multipole range that they are applicable to. The quadratic maximum likelihood (QML) estimator (Tegmark 1997; Tegmark & Oliveira-Costa 2001) gives an unbiased minimum-variance estimate for the low-multipole range, when given a CMB sky map as input. A practical implementation of the method is the `BolPol` estimator (Gruppuso et al. 2009), which has been applied to multipoles up to $\ell = 32$.

The high-multipole regime requires different techniques. To mention the ones most relevant for this work, the `Master` (Hivon et al. 2002) method provides an unbiased estimate of the CMB temperature spectrum. The method consists of computing the pseudo-$\mathcal{C}_\ell$ spectrum of the input map through harmonic transform, and correcting it through a kernel matrix that depends on sky coverage. Similar methods have been developed for polarization (Kogut et al. 2003; Chon et al. 2004; Grain, Tristam & Stompor 2009). For a more extensive review of power spectrum estimation (PSE) methods, we refer to Gruppuso et al. (2009).

Beam effects are typically not considered part of the PSE problem, but are taken into account at a later stage in the form of a beam window function. A scalar beam window (for instance, Mitra et al. 2011) captures the average smoothing effect of the beam. A scalar window function does not correct the leakage from temperature to polarization. Methods for correcting the latter include matrix window function formalism (Planck Collaboration IV 2016) and `QuickPol` (Hivon, Mottet & Ponthieu 2016).

On the other hand, methods have been developed for full-beam deconvolution at map level (Armitage & Wandelt 2004; Armitage-Caplan & Wandelt 2009; Harrison, van Leeuwen & Ashdown 2011; Keihänen & Reinecke 2012). Deconvolution map-making yields a sky map with a symmetrized effective beam. This is achieved at the cost of more complicated noise structure, which makes it more difficult to use the deconvolved maps for PSE. The impact of deconvolution on noise properties has been discussed in Keihänen et al. (2016). A low-multipole noise covariance matrix for deconvolved maps has been derived in the same paper.

★ E-mail: elina.keihanen@helsinki.fi (EK); martin@mpa-garching.mpg.de (MR)

In this paper, we aim at combining the best of the two worlds. We present a method that combines full-beam deconvolution with techniques of PSE. The deconvolution technique is based on the `artDeco` deconvolver (Keihänen & Reinecke [2012]). Our goal is to estimate the power spectrum directly from the time-ordered data (TOI), without constructing a sky map along the way, at the same time eliminating the smoothing by an asymmetric beam. In particular, we aim at removing the beam-induced leakage from temperature to polarization.

The situation we have in mind is the following. Assume we have successfully performed component separation for our data. We thus have an estimate of how much foreground emission there is in each sky pixel. This is the starting point in pixel-based PSE methods as well. Instead of proceeding with the foreground-cleaned CMB map, we go back to the TOI. We can scan the foreground map into a TOI, convolving it with the known beam and subtract it from the original TOI. It is unlikely that the cleaning can be done with sufficient accuracy for all of the sky. Therefore, the regions with the strongest foreground emission are masked out. We are thus left with a pure CMB time stream, but with incomplete sky coverage.

We aim at deriving the optimal quadratic method that yields an unbiased estimate of the CMB spectrum, at the same time minimizing its variance, when given a TOI stream as input. The technique is inspired by the work of Tegmark ([1997]), with the distinction that we include beam effects, and, instead of maps, operate at TOI level. The required inputs beside the TOI itself are pointing information and known beam shapes.

We derive yet another method that is applicable to the full multipole range of interest. Again, the method yields an unbiased estimate of the CMB spectrum, however not necessarily with minimal variance. Formally, we derive the high-ell method from the optimal method as a weak-signal approximation. A key element in both methods is a kernel matrix that is applied to a raw spectrum to correct simultaneously for beam effects and for incomplete sky coverage.

We validate both new methods with simulations. As the basic simulation case, we take the 4-yr data set of the *Planck* LFI instrument's 70 GHz channel.

This paper is organized as follows. We start by presenting the fiducial simulation case in Section 2. We review the deconvolution map-making method of `artDeco` in Section 3. To set a reference point, we compare some pixel-based PSE methods in Section 4. The optimal PSE method is presented in Section 5, and the high-multipole version in Section 6. We validate the methods with simulations in Section 7. In Section 8, we present an idea for future development, statistical estimation of the optimal kernel matrix. Finally, we give our conclusions in Section 9.

2 TEST CASE

We selected as test case the *Planck* LFI instrument and its 2015 data release (Planck Collaboration I [2016]). We run simulations on all LFI channels: 30, 44 and 70 GHz. We consider both temperature and polarization. From the cosmology point of view, the 70 GHz channel is the most interesting, and we select it as the fiducial test case. This channel offers a wide multipole range ($\ell = 0$–1500) and detector beams asymmetric enough to demonstrate various beam effects. It is also, computationally, the most demanding one due to the size of the data set and the wide multipole range covered. The FWHM beam widths for the 70 GHz channel vary in the range 12.8–13.5 arcmin, allowing us to estimate the power spectrum up to $\ell_{max} = 1500$. The results shown in this paper are mainly based on

the fiducial 70 GHz data set. In some cases, where an effect under study is more prominent at another channel, or the computation is very heavy, we show results from 30 GHz or 44 GHz.

We used the LEVELS software package (Reinecke et al. [2006]) to generate detector pointings, to convolve the input sky with realistic beams, and to produce time-ordered data from it. The simulations cover 4 yr of mission, during which the sky is scanned eight times. Because we are focusing on CMB PSE, we allowed some simplifications with respect to full *Planck* data analysis. We did not include systematics such as calibration errors and added white noise only. The beams, however, were fully realistic, and included the main and intermediate side-lobe components of real *Planck* beams (Planck Collaboration IV [2016]).

The input sky contained the polarized CMB signal, but no foregrounds. We applied the *Planck* calibration masks (Planck Collaboration V [2016]) to remove the Galactic region from the data set. This mimics a situation where an estimate of the foreground signal has been subtracted from the TOI before passing it to PSE, and the regions of the strongest contamination have been cut out. We are left with a data set that consists of pure CMB, but with incomplete sky coverage. The sky coverage at 70 GHz is 89.67 per cent. As input CMB sky, we used one realization from the FFP8 simulation set (Planck Collaboration XII [2016])

We also performed a series of pure white noise simulations, which we processed through the same deconvolution procedure as the CMB simulations. The noise parameters were taken from Planck Collaboration II ([2016]). Here, we are assuming that the data have been destriped or otherwise cleaned of correlated noise before the PSE phase, in sufficient degree that the residual noise can be considered white. We discuss the effect of non-white residuals briefly in the Results section.

The noise simulations serve several purposes. We use them to evaluate the noise bias, which is subtracted from the estimated spectrum to reveal the actual CMB spectrum and to study the properties of residual noise. We utilize them also in the statistical kernel evaluation method presented in Section 8.

3 DECONVOLUTION MAP-MAKING

3.1 artDeco deconvolver

Our work builds on the beam deconvolution formalism presented in Keihänen & Reinecke ([2012]). Consider a time-ordered data stream of the form

$$t_i = \sum_{s\ell m} A_{i,s\ell m} a_{s\ell m} + n_i. \tag{1}$$

Here, index i labels the samples in the data stream, n_i is noise and harmonic coefficients a_{slm} represent the sky signal. The spin index s takes values 0 and ± 2, 0 representing temperature and ± 2 polarization. There is a linear dependence between the sky $a_{s\ell m}$ and the data t, collected in matrix **A**. The explicit formula for A is given in Keihänen & Reinecke ([2012]) as

$$A_{i,s\ell m} = \sum_k b^*_{s\ell k} D^{\ell*}_{mk}(\omega_i). \tag{2}$$

Here, $b_{s\ell k}$ is the harmonic expansion of the detector beam, D^{ℓ}_{mk} is a Wigner matrix and ω_i is the combination of three pointing angles $\{\theta, \phi, \psi\}$ that define the pointing and orientation of the beam for sample i.

In deconvolution map-making, we perform a linear fit to the TOI to obtain an estimate for the coefficients $a_{s\ell m}$. From those, one can
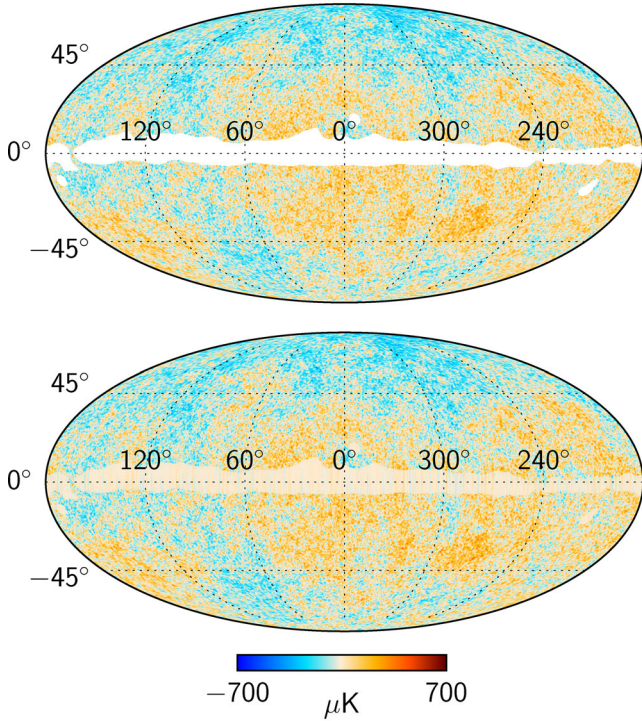
Figure 1. Simulated CMB sky, seen through *Planck* LFI 70 GHz beams. Top: map binned directly from the time-ordered data. Galaxy and strong point sources are masked out, leaving a pure CMB signal. Bottom: beam-deconvolved version of the same map, constructed through harmonic expansion from deconvolved $a_{s\ell m}$ coefficients. Deconvolution restores the CMB signal in the observed region, but creates artefacts in the unobserved region.

Figure 2. TT spectrum as constructed directly from deconvolved $a_{T\ell m}$ co-efficients. Simulation data consist of a pure CMB signal, convolved by *Planck* 70 GHz beams. Here, as throughout the paper, we plot the quantity $\mathcal{D}_\ell = C_\ell \ell(\ell+1)/(2\pi)$, which shows the high-multipole structure more clearly. In the lower panel, we plot the absolute error in the same units. In the ideal case with 100 per cent sky coverage, the deconvolved $a_{s\ell m}$ coefficients directly give a reliable estimate of the CMB spectrum. This is no longer true when deconvolution is applied to a data set where the part of the sky is masked out. Scaling the spectrum by the inverse of sky fraction $f_{\text{sky}} = 0.8967$ provides a rough correction (green), but still leaves a systematic error of the order of 1 per cent.

construct a sky map, where the effective beam is symmetric and Gaussian. The general deconvolution solution can be written in vector notation as

$$\hat{a} = (\mathbf{A}^\dagger \mathbf{N}^{-1} \mathbf{A} + \mathbf{S}^{-1})^{-1} \mathbf{A}^\dagger \mathbf{N}^{-1} t. \tag{3}$$

Here, S represents an optional prior in the form of a first-guess CMB power spectrum. Formally, it is a diagonal matrix, with elements of the angular power spectrum spread along the diagonal. In usual deconvolution map-making, it is omitted, but it has an important role in our PSE method.

3.2 Direct spectrum estimation

The harmonic coefficients obtained from equation (3) are not directly useful for the estimation of the CMB power spectrum if sky coverage is not complete. The deconvolution as implemented in `artDeco` can be performed under incomplete sky coverage as well, and it works correctly in the sense that it recovers the sky map outside the mask. However, the output $a_{s\ell m}$ does not represent the correct CMB spectrum. One can imagine filling the missing sky region by an arbitrary CMB realization. Each possible realization corresponds to a different set of $a_{s\ell m}$, each of which gives the correct sky map in the region outside the mask. From these realizations, `artDeco` arbitrarily picks one, depending on the starting point of the conjugate-gradient iteration. Without additional information, it is not possible to identify the correct realization.

We demonstrate deconvolution map-making in Fig. 1. We show two versions of the temperature map, constructed from a simulated timeline with *Planck* 70 GHz beams. We show first the binned
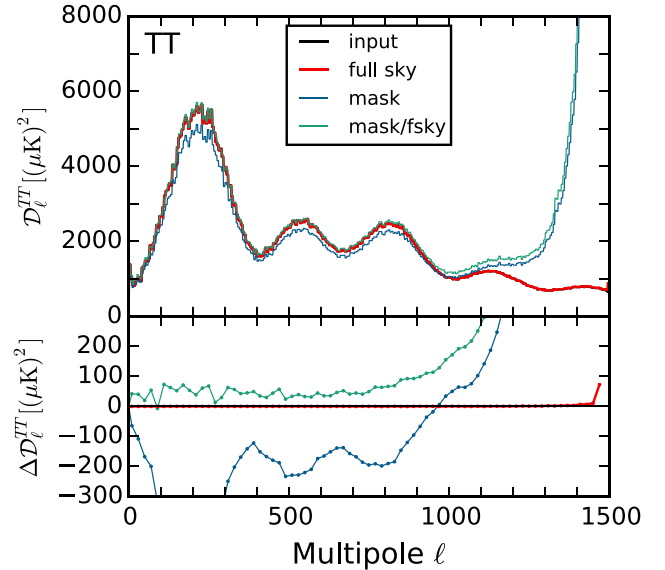
map, which is constructed by coadding the TOI samples into a two-dimensional sky, without attempting to correct for beam shapes. Below it, we show the deconvolved version of the same map. To construct the latter, we run the data through `artDeco` deconvolver to obtain a beam-free $a_{s\ell m}$ expansion of the sky, reconstruct the sky map through harmonic transform and apply Gaussian smoothing with FWHM = 12 arcmin. Both maps make use of HEALPIX pixelization at resolution $N_{\text{side}} = 1024$ (3,5 arcmin). Despite there being no data available from the masked region, deconvolution inserts a signal there as well.

In Fig. 2, we show the raw TT spectrum constructed directly from the deconvolved harmonic coefficients as

$$\hat{C}_\ell^{TT} = \frac{1}{2\ell+1} \sum_m a_{T\ell m}^* a_{T\ell m} \tag{4}$$

for full sky, and in presence of a mask. In the lower panel, we show the absolute error. Throughout the paper, we follow the convention where we compare the recovered spectrum to the spectrum of the single CMB realization that was used as input to the simulation. We plot the quantity $\mathcal{D}_\ell = C_\ell \ell(\ell+1)/(2\pi)$, which shows the high-multipole range more clearly. As a measure of error, we plot the absolute error, calculated as the difference between the recovered spectrum and the input spectrum. This allows us to use the same error measure for the TE cross-spectrum, for which the relative error is an inconvenient measure.

We see in Fig. 2 that in the ideal case where the sky coverage is complete and no noise is included in the simulations, deconvolution recovers the true CMB spectrum almost perfectly. The situation changes when a mask is applied. Direct deconvolution yields a
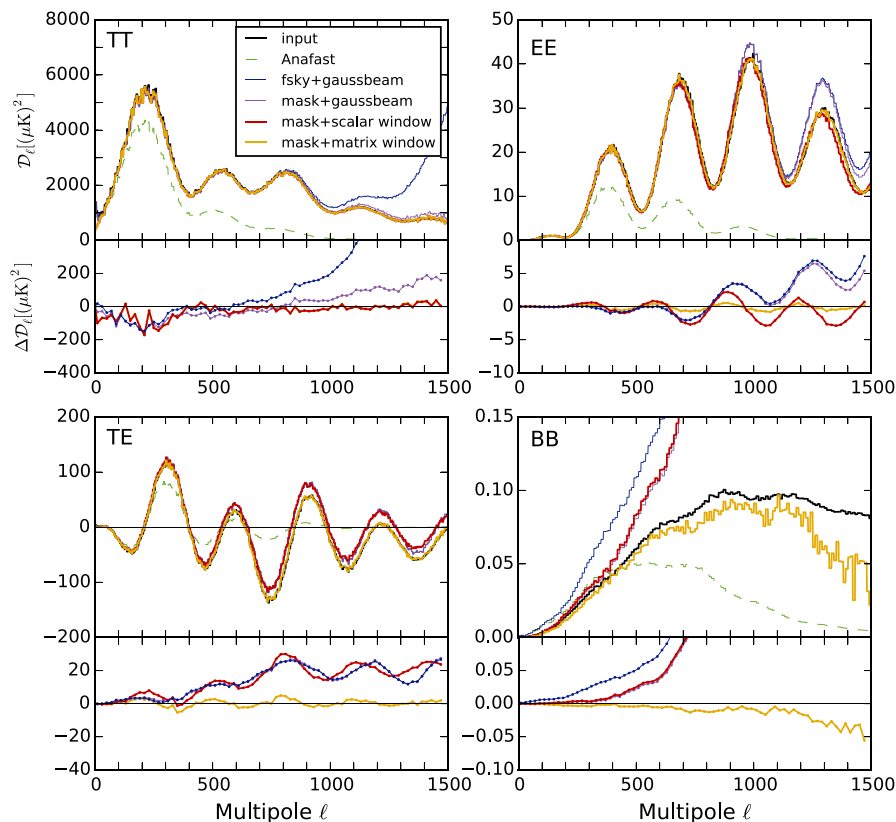
Figure 3. Comparison of selected pixel-based methods for PSE, applied to the fiducial 70 GHz simulation. We show the spectrum of the input CMB realization, the un-corrected Anafast spectrum (dashed) and a series of estimates with increasing complexity: scaling by sky fraction f_{sky} and Gaussian beam window, and PolSpice mask correction combined with Gaussian beam, scalar beam window function or matrix beam window function. The difference between scalar and matrix window functions reveals the level of beam-induced leakage from temperature to polarization. The spectra are averaged over 5 adjacent multipoles to reduce scatter, except for BB, which is averaged over 10 multipoles. The absolute error given in the lower panel is averaged over 20 multipoles.

spectrum that falls below the input at low multipoles, but rises steeply at high multipoles. The effect at low multipoles corresponds roughly to scaling the spectrum by the sky coverage, and can as first approximation be corrected for by inverse scaling. There remains, however, a 1 per cent error, which is unacceptably high for present-day precision cosmology. More accurate methods are thus needed. At high multipoles (above $\ell = 800$), the deconvolved spectrum becomes useless.

Even if we cannot determine the individual harmonic coefficients, we still may be able to estimate their spectrum. The main goal of this paper is to derive a method that extracts the spectral information in an optimal way, given the pre-known beam shapes and known noise level.

4 BEAM CORRECTION AT THE POWER SPECTRUM LEVEL

To put the new methods into context, we first compare some available map-based methods for PSE. We use again the simulated *Planck* 70 GHz data set, with realistic beams. A Galactic mask with $f_0 = 0.8967$ sky coverage is applied. The methods discussed here take as starting point the (I, Q, U) Stokes map triplet binned from time-ordered data. The operation of constructing the map from the TOI utilizes the known direction of polarization sensitivity for

each detector, encoded in parameter ψ_{pol}, but no other information on beam properties.

The pseudo-spectrum constructed from the map is suppressed by both the beam shape and the incomplete sky coverage. We refer to this initial spectrum as Anafast spectrum, according to the standard HEALPIX (Górski et al. 2005) tool. We show in Fig. 3 the Anafast spectrum along with various corrections. The simplest correction consists of dividing the Anafast spectrum by the sky fraction f_{sky}, and by a Gaussian beam window estimate. The FWHM width of the Gaussian window was taken from Planck Collaboration IV (2016) to be 13.315 arcmin. This naive correction works reasonably well at low multipoles, but overestimates all the spectra towards higher multipoles.

Several more sophisticated methods exist, that take into account the exact mask shape and correct the effect of the incomplete sky coverage. We use the publicly available POLSPICE[1] code (Chon et al. 2004) to apply the correction to our data set. we then apply the Gaussian beam window to correct for the beam smoothing. The mask correction improves the spectrum estimate, especially in the BB spectrum, and in TT at high multipoles.

We may further improve the spectrum estimate by replacing the Gaussian beam approximation by a proper beam window computed from the full-beam information. The scalar beam window

---

[1] www2.iap.fr/users/hivon/software/PolSpice/

function W_ℓ^u can be obtained from MC simulations using the known instrument beam as

$$W_\ell^u \equiv \frac{\langle \tilde{\mathcal{C}}_\ell^u \rangle_{\mathrm{MC}}}{\mathcal{C}_\ell^u}, \qquad (5)$$

where $\mathcal{C}_\ell^u$ ($u = $ TT, EE...) is some input spectrum used for the simulations, $\tilde{\mathcal{C}}_\ell^u$ is the output (Anafast) full-sky spectrum and $\langle\rangle_{\mathrm{MC}}$ is an average over many realizations. This method is not suitable for treating leakage between temperature and polarization signals, so to treat just the smoothing effect, we use equation (5) for $u = $ TT, EE, BB with separate **T**-only, **E**-only and **B**-only simulations. For $u = $ TE, TB, EB we then use

$$W_\ell^{\mathrm{TE}} = \sqrt{W_\ell^{\mathrm{TT}} W_\ell^{\mathrm{EE}}}, \quad \text{etc.} \qquad (6)$$

The angular power spectrum estimate $\hat{\mathcal{C}}_\ell^u$ for spectral component u is obtained from the Anafast spectrum by dividing by the corresponding beam window.

For correcting the beam-induced leakage between temperature and polarization signals, a matrix window function method was introduced in Planck Collaboration IV (2016). Here the estimate $\hat{\mathcal{C}}_\ell$ is obtained from the spectrum $\tilde{\mathcal{C}}_\ell$ of the sky map as

$$\hat{\mathcal{C}}_\ell = W_\ell^{-1} \tilde{\mathcal{C}}_\ell, \qquad (7)$$

where the $\mathcal{C}_\ell$ are six-component vectors with components $\mathcal{C}_\ell^u$ and the W_ℓ is a 6×6 matrix with components $W_\ell^{uu'}$. The components $W_\ell^{uu'}$ of the matrix window function are obtained from **T**-only, **E**-only and **B**-only MC simulations as described in (Planck Collaboration IV 2016). For correcting the effect of the incomplete sky coverage prior to the beam correction, we again use the `PolSpice` method.

The difference between the results obtained with scalar beam window and with matrix beam window is an indicator of the level of leakage between spectral components, an effect that cannot be corrected for through a scalar window function. As can be expected, the difference is significant in polarization components, but negligible in temperature.

Of all the correction methods compared here, the last one (`PolSpice` mask correction combined with matrix beam window) provides the best correction. This sets a reference level for the accuracy that we must require from our deconvolution-based spectrum estimation method.

5 OPTIMAL PSE

5.1 Notation

We now set out to derive a PSE method, which corrects simultaneously the effect of the incomplete sky coverage, and that of beam smoothing, including beam-induced leakage effects. We start by introducing some formalism used throughout this work.

Consider first the properties of the CMB $a_{s\ell m}$ elements. We denote by S their covariance,

$$\langle a_{s\ell m} a_{s'\ell'm'}^* \rangle = S_{s\ell m, s'\ell'm'}. \qquad (8)$$

All elements except the ones with $\ell = \ell'$ and $m = m'$ vanish. The non-zero elements are linear combinations of elements of the CMB power spectra $\mathcal{C}_{\mathrm{TT}}$, $\mathcal{C}_{\mathrm{EE}}$, $\mathcal{C}_{\mathrm{BB}}$, $\mathcal{C}_{\mathrm{TE}}$, $\mathcal{C}_{\mathrm{TB}}$ and $\mathcal{C}_{\mathrm{EB}}$. Parity symmetry indicates $\mathcal{C}_{\mathrm{TB}} = \mathcal{C}_{\mathrm{EB}} = 0$. To keep the formalism as general as possible, we retain these components as well. The unwanted elements can be dropped at any later stage.

The spin harmonics are related to the **T**, **E**, **B** harmonic components through

$$a_{0\ell m} = a_{T\ell m}$$
$$a_{2\ell m} = -(a_{E\ell m} + i a_{B\ell m})$$
$$a_{-2\ell m} = -(a_{E\ell m} - i a_{B\ell m}). \qquad (9)$$

From this, we can derive the linear relation between the spectral components,

$$\langle a_{0lm} a_{0l'm'}^* \rangle = \mathcal{C}_\ell^{TT} \delta_{\ell\ell'} \delta_{mm'}$$
$$\langle a_{2\ell m} a_{2\ell'm'}^* \rangle = \left(\mathcal{C}_\ell^{EE} + \mathcal{C}_\ell^{BB} \right) \delta_{\ell\ell'} \delta_{mm'}$$
$$\langle a_{-2\ell m} a_{-2\ell'm'}^* \rangle = \left(\mathcal{C}_\ell^{EE} + \mathcal{C}_\ell^{BB} \right) \delta_{\ell\ell'} \delta_{mm'}$$
$$\langle a_{2\ell m} a_{-2\ell'm'}^* \rangle = \left(\mathcal{C}_\ell^{EE} - \mathcal{C}_\ell^{BB} + i2\mathcal{C}_\ell^{EB} \right) \delta_{\ell\ell'} \delta_{mm'}$$
$$\langle a_{0\ell m} a_{2\ell'm'}^* \rangle = -\left(\mathcal{C}_\ell^{TE} - i\mathcal{C}_\ell^{TB} \right) \delta_{\ell\ell'} \delta_{mm'}$$
$$\langle a_{0lm} a_{-2l'm'}^* \rangle = -\left(\mathcal{C}_\ell^{TE} + i\mathcal{C}_\ell^{TB} \right) \delta_{\ell\ell'} \delta_{mm'}. \qquad (10)$$

We can write this as

$$\langle a_{s\ell m} a_{s'\ell'm'}^* \rangle = \delta_{\ell\ell'} \delta_{mm'} \sum_u g_{ss'}^u \mathcal{C}_{ul}, \qquad (11)$$

where u labels the spectra ($u = $ TT, EE, BB, TE, TB, EB), and g is a $3 \times 3 \times 6$ object whose non-zero elements are

$$g_{00}^{TT} = 1$$
$$g_{0,-2}^{TE} = g_{0,2}^{TE} = g_{-2,0}^{TE} = g_{2,0}^{TE} = -1$$
$$g_{2,2}^{EE} = g_{2,-2}^{EE} = g_{-2,2}^{EE} = g_{-2,-2}^{EE} = 1$$
$$g_{2,2}^{BB} = g_{-2,-2}^{BB} = -g_{2,-2}^{BB} = -g_{-2,2}^{BB} = 1$$
$$g_{0,2}^{TB} = g_{-2,0}^{TB} = -g_{0,-2}^{TB} = -g_{2,0}^{TB} = i$$
$$g_{2,-2}^{EB} = -g_{-2,2}^{EB} = 2i. \qquad (12)$$

We further define object G as

$$G_{s\ell m, s'\ell'm'}^{u\ell''} = \delta_{\ell\ell''} \delta_{\ell\ell'} \delta_{mm'} g_{ss'}^u. \qquad (13)$$

With this definition the relations of equation (10) can be written in the compact form

$$\langle a_{s\ell m} a_{s'\ell'm'}^* \rangle = \sum_{u\ell''} \mathcal{C}_{u\ell''} G_{s\ell m, s'\ell'm'}^{u\ell''}. \qquad (14)$$

We introduce combined indices that allow a still more concise notation. Index n represents the triplet of indices $n = \{s, \ell, m\}$ and index L the combination of indices $L = \{u, \ell\}$. The time-ordered data of equation (1) becomes

$$t_i = \sum_n A_{in} a_n + n_i, \qquad (15)$$

and equation (14) becomes

$$S_{nn'} = \langle a_n a_{n'}^* \rangle = \sum_L \mathcal{C}_L G_{nn'}^L. \qquad (16)$$

It is useful to note that the standard operation of calculating the spectrum of a given harmonic vector a_{slm} is formally given as

$$\hat{\mathcal{C}}_L = \left[\sum_{nn'} G_{nn'}^{*L} G_{nn'}^{L'} \right]^{-1} \sum_{nn'} a_n G_{nn'}^L a_{n'}^*. \qquad (17)$$

When written out, this gives, for instance, for the TT spectrum the usual formula

$$\hat{\mathcal{C}}_\ell^{TT} = \frac{1}{2\ell + 1} \sum_m |a_{0\ell m}|^2. \tag{18}$$

Matrix G represents a quadratic sum of two $a_{s\ell m}$ vectors, and the factor in the brackets takes care of normalization.

5.2 Finding the optimal quadratic method

We aim at finding a quadratic method, represented by matrix E_{ij}^L, which gives an optimal estimate for the power spectrum $\mathcal{C}_L$, given known beams and time-ordered data. The data are assumed to consist of a CMB signal and noise. We follow the example of the quadratic maximum likelihood (QML) analysis of Tegmark (1997), the major difference being that we are dealing with full-time streams instead of sky maps. We refer to the new method as DQML (deconvolution quadratic maximum likelihood).

The optimal estimate must fulfil two requirements.
(i) The solution is unbiased (as ensemble average).
(ii) Under condition (i) the solution minimizes the variance.

Consider the first condition (i). Given a time stream t, we want to find a (symmetric) matrix E_{ij}^L, and constant β_L, such that

$$\hat{\mathcal{C}}_L = \sum_{ij} t_i E_{ij}^L t_j - \beta_L \tag{19}$$

gives an unbiased estimate of the true power spectrum $\mathcal{C}^L$, that is

$$\langle \hat{\mathcal{C}}_L \rangle = \mathcal{C}_L. \tag{20}$$

The brackets denote an ensemble average over noise realizations and realizations of a_n for a given power spectrum.

Inserting equation (15) into (19) and averaging, we obtain, assuming that noise and signal are independent,

$$\langle \hat{\mathcal{C}}_L \rangle = \sum_{ij} E_{ij}^L \left(\sum_{nn'} A_{jn} \langle a_n a_{n'}^* \rangle A_{in'}^* + \langle n_j n_i \rangle \right) - \beta_L \tag{21}$$

$$= \sum_{ij} E_{ij}^L \left(\sum_{nn'} A_{jn} A_{in'}^* \sum_{L'} \mathcal{C}_{L'} G_{nn'}^{L'} + N_{ji} \right) - \beta_L.$$

We have denoted the noise covariance in TOI domain by

$$\langle n_i n_j \rangle = N_{ij}. \tag{22}$$

If noise is white, this is a simple diagonal matrix.

The requirement of equation (20) now splits into two conditions. From the noise part, we get a relation for the noise bias β,

$$\beta_L = \sum_{ij} E_{ij}^L N_{ji}, \tag{23}$$

which we can evaluate once we have found out E. The signal part gives the condition

$$\sum_{ij} E_{ij}^L \sum_{nn'} A_{jn} A_{in'}^* G_{nn'}^{L'} \delta_{LL'}. \tag{24}$$

These can be written in matrix formalism as

$$\beta_L = \mathrm{Tr}(\mathbf{E}^L \mathbf{N}), \tag{25}$$

and

$$\mathrm{Tr}(\mathbf{E}^L \mathbf{A} \mathbf{G}^{L'} \mathbf{A}^\dagger) = \delta_{LL'}. \tag{26}$$

When using matrix formalism we are interpreting E_{ij}^L and N_{ji} as matrices with i, j as row and column index. Each value of index L

thus identifies one matrix. The trace operation applies to indices i, j. Similarly, $G_{nn'}^L$ is interpreted as a matrix with n, n' as column and row index.

The condition of equation (26) does not yet uniquely determine $\mathbf{E}$. Consider then condition (ii). Among all $\mathbf{E}^L$ that satisfy condition (26), we want to find the one that minimizes the variance,

$$V = \langle \hat{\mathcal{C}}_L^2 \rangle - \langle \hat{\mathcal{C}}_L \rangle^2. \tag{27}$$

We can write (β cancels out)

$$V = \left\langle \sum_{ij} t_i E_{ij}^L t_j \sum_{i'j'} t_{i'} E_{i'j'}^L t_{j'} \right\rangle - \left\langle \sum_{ij} t_i E_{ij}^L t_j \right\rangle \left\langle \sum_{i'j'} t_{i'} E_{i'j'}^L t_{j'} \right\rangle$$

$$= \sum_{ii'jj'} E_{ij}^L E_{i'j'}^L (\langle t_i t_j t_{i'} t_{j'} \rangle - \langle t_i t_j \rangle \langle t_{i'} t_{j'} \rangle)$$

$$= \sum_{ii'jj'} E_{ij}^L E_{i'j'}^L (\langle t_i t_{i'} \rangle \langle t_j t_{j'} \rangle + \langle t_i t_{j'} \rangle \langle t_{i'} t_j \rangle)$$

$$= \sum_{ii'jj'} E_{ij}^L E_{i'j'}^L (C_{ii'} C_{jj'} + C_{ij'} C_{i'j}). \tag{28}$$

On the third line, we have used the property of Gaussian random variates:

$$\langle x_1 x_2 x_3 x_4 \rangle = \langle x_1 x_2 \rangle \langle x_3 x_4 \rangle + \langle x_1 x_3 \rangle \langle x_2 x_4 \rangle + \langle x_1 x_4 \rangle \langle x_2 x_3 \rangle. \tag{29}$$

Matrix C on the last line represents the time-domain covariance, including both CMB and noise, and is given by

$$C_{ii'} = \langle t_i t_{i'} \rangle = \sum_{nn'} A_{in} \langle a_n a_{n'}^* \rangle A_{i'n'}^* + \langle n_i n_{i'} \rangle$$

$$= \sum_{nn'} A_{in} S_{nn'} A_{i'n'}^* + N_{ii'}, \tag{30}$$

or, in matrix notation,

$$\mathbf{C} = \mathbf{A} \mathbf{S} \mathbf{A}^\dagger + \mathbf{N}. \tag{31}$$

We want to minimize the variance (28) under condition (26). We use the technique of Lagrange multipliers (see Tegmark's work). The technique is intuitively understood as follows. The constraint of equation (26) picks a surface in the multidimensional space of E_{ij}. At a local minimum of V on this surface, the gradient of V does not have a component along the surface. This is equivalent to saying that the gradient is included in the subspace spanned by $AG^\dagger A^\dagger$ $(AG^\dagger A^\dagger)_{ij}$. The proportionality constants are the Lagrange multipliers, and are adjusted so as to fulfil the constraint (26).

We take the gradient of V with respect to E_{ij}^L, and write

$$2 \sum_{i'j'} (C_{ii'} C_{jj'} + C_{ij'} C_{i'j}) E_{i'j'}^L = 4 \sum_{L'} \lambda_{LL'} \sum_{nn'} A_{in} G_{nn'}^{L'\dagger} A_{jn'}^*. \tag{32}$$

We have arbitrarily scaled the right-hand side by 4 (we can think of it as being absorbed in λ). Rearranging the indices, and remembering that E and C are symmetric, we find that the two terms on the left-hand side are identical. The factor of 4 cancels out, and equation (32) becomes in matrix notation

$$\mathbf{C} \mathbf{E}^L \mathbf{C} = \sum_{L'} \lambda_{LL'} \mathbf{A} \mathbf{G}^{L'\dagger} \mathbf{A}^\dagger. \tag{33}$$

Since C is a covariance matrix, it is invertible. We can thus solve for E as

$$\mathbf{E}^L = \sum_{L'} \lambda_{LL'} \mathbf{C}^{-1} \mathbf{A} \mathbf{G}^{L'\dagger} \mathbf{A}^\dagger \mathbf{C}^{-1}. \tag{34}$$

This gives the optimal quadratic method we are looking for. We yet have to determine the Lagrange coefficients λ. Inserting E to equation (26) we obtain

$$\sum_{L'} \lambda_{LL'} \text{Tr} \left[\mathbf{A}^\dagger \mathbf{C}^{-1} \mathbf{A} \mathbf{G}^{L'\dagger} \mathbf{A}^\dagger \mathbf{C}^{-1} \mathbf{A} \mathbf{G}^{L''} \right] = \delta_{LL''}. \tag{35}$$

From this we can solve the coefficients λ, by inverting the matrix defined by the trace formula. The minimum-variance estimate for the power spectrum is then obtained as

$$\hat{\mathcal{C}}_L = \sum_{L'} \lambda_{LL'} \boldsymbol{t}^T \mathbf{C}^{-1} \mathbf{A} \mathbf{G}^{L'\dagger} \mathbf{A}^\dagger \mathbf{C}^{-1} \boldsymbol{t}. \tag{36}$$

We now have a formal solution for our PSE problem. The solution in its present form is still impractical, since it involves inverting the huge TOI covariance $\mathbf{C}$. We need to work the solution yet into a more practical form. With the help of Sherman–Morrison formula and equation (31), we can write

$$\mathbf{A}^\dagger \mathbf{C}^{-1} \boldsymbol{t} = \mathbf{S}^{-1} (\mathbf{S}^{-1} + \mathbf{A}^\dagger \mathbf{N}^{-1} \mathbf{A})^{-1} \mathbf{A}^\dagger \mathbf{N}^{-1} \boldsymbol{t}. \tag{37}$$

Apart from factor $\mathbf{S}^{-1}$ in front, this is the usual deconvolution equation of equation (3). With a similar technique we get

$$\mathbf{A}^\dagger \mathbf{C}^{-1} \mathbf{A} = \mathbf{S}^{-1} - \mathbf{S}^{-1} (\mathbf{S}^{-1} + \mathbf{A}^\dagger \mathbf{N}^{-1} \mathbf{A})^{-1} \mathbf{S}^{-1}. \tag{38}$$

We have replaced the TOI covariance by the much smaller deconvolution matrix, which has the rank equal to the number of harmonic coefficients $3(\ell_{\max} - 1)^2$.

5.3 DQML algorithm

We are now ready to collect the results into a recipe for optimal PSE.

(i) Run the usual beam deconvolution procedure on the TOI, using some first estimate of the power spectrum as a prior,

$$\hat{\boldsymbol{a}} = (\mathbf{S}^{-1} + \mathbf{A}^\dagger \mathbf{N}^{-1} \mathbf{A})^{-1} \mathbf{A}^\dagger \mathbf{N}^{-1} \boldsymbol{t}. \tag{39}$$

This step is performed by the `artDeco` code.

(ii) Construct the raw spectrum as

$$P_L = \hat{\boldsymbol{a}}^\dagger \mathbf{S}^{-1} \mathbf{G}^{L\dagger} \mathbf{S}^{-1} \hat{\boldsymbol{a}}. \tag{40}$$

Apart from normalization, matrix $\mathbf{G}^L$ represents the familiar squared-sum operation of constructing the spectrum of a harmonic vector $\hat{\boldsymbol{a}}$. Explicit from for $\mathbf{G}$ is given by equation (13). Matrix $\mathbf{S}$ represents scaling by the CMB spectrum used as prior. Formally, it is a diagonal matrix with elements of the prior on the diagonal.

(iii) Construct the kernel matrix

$$\Lambda_{LL'} = \text{Tr} \left[\mathbf{M} \mathbf{G}^L \mathbf{M}^\dagger \mathbf{G}^{\dagger L'} \right], \tag{41}$$

where $\mathbf{M}$ is given by

$$\mathbf{M} = \mathbf{S}^{-1} - \mathbf{S}^{-1} (\mathbf{S}^{-1} + \mathbf{A}^\dagger \mathbf{N}^{-1} \mathbf{A})^{-1} \mathbf{S}^{-1}. \tag{42}$$

This is the computationally heavy step. It involves constructing explicitly the full deconvolution matrix. Routines for that exist in the `artDeco` code.

(iv) Apply the kernel to the raw spectrum

$$\hat{\mathcal{C}}_L = \sum_{L'} \Lambda_{LL'}^{-1} P_{L'}. \tag{43}$$

(v) If noise is present, evaluate the noise bias β_L through Monte Carlo simulations, and subtract the bias from the estimate of step 4.
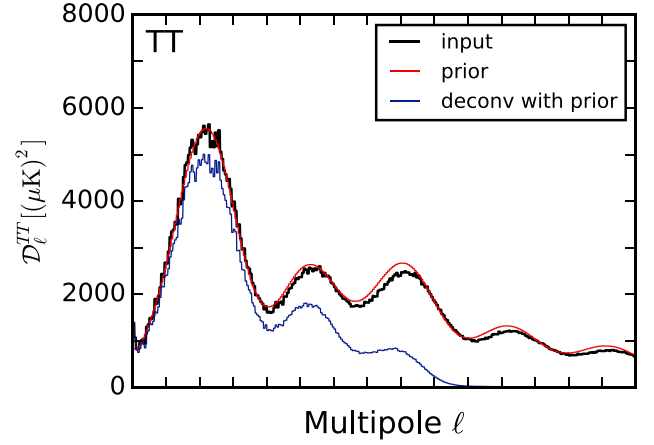
Figure 4. 70 GHz TT spectrum used as input in simulations and the one used as prior. Deconvolution with prior yields a suppressed spectrum estimate. The suppressed spectrum acts as input to the PSE estimation method of Section 5.

The solution contains an estimate of the CMB spectrum $\mathcal{C}_L$ inside the prior $\mathbf{S}$. It is not known, since it is exactly the quantity we are trying to determine. It is important to note, however, that $\mathbf{S}$ only enters in the part of calculation where we minimized the variance. If our estimate of $\mathbf{S}$ is inaccurate, the solution will not be that of minimal variance but the solution is still unbiased.

5.4 Implementation aspects

When implementing the method, we made a couple of further rearrangements. We note that the prior $\mathbf{S}$ is independent of index m. From that, and from the special structure of $\mathbf{G}$, it follows that $\mathbf{S}^{-1}$ can be taken out from equation (40) and transferred inside the kernel of equation (41). Further, we replace the operation $\mathbf{G}$ in equation (40) by the normalized version of equation (17), and apply the name normalization to the left-hand side of the kernel. The modified kernel is dimensionless, but no more symmetric. The raw spectrum of step 2 is replaced by

$$\tilde{P}_L = \left[\text{Tr}(\mathbf{G}^L \mathbf{G}^{L'\dagger}) \right]^{-1} \hat{\boldsymbol{a}}^\dagger \mathbf{G}^{L\dagger} \hat{\boldsymbol{a}}. \tag{44}$$

Equation (44) represents the standard operation of constructing the spectrum of a harmonic vector, and is equivalent to equation (17), only written in matrix notation. The corresponding kernel is

$$\tilde{\Lambda}_{LL'} = \left[\text{Tr} \mathbf{G}^L \mathbf{G}^{L'\dagger} \right]^{-1} \text{Tr} \left[\tilde{\mathbf{M}} \mathbf{G}^L \tilde{\mathbf{M}}^\dagger \mathbf{G}^{\dagger L'} \right], \tag{45}$$

where $\tilde{\mathbf{M}}$ is given by

$$\tilde{\mathbf{M}} = \mathbf{I} - (\mathbf{S}^{-1} + \mathbf{A}^\dagger \mathbf{N}^{-1} \mathbf{A})^{-1} \mathbf{S}^{-1}. \tag{46}$$

The modified algorithm is fully equivalent to the original one, but is a little more intuitive. The right-hand side 'raw' spectrum has the physical interpretation as the spectrum of the deconvolved harmonic coefficients, and can be constructed using standard tools. For us remains the task of constructing the kernel $\tilde{\Lambda}$. The raw TT spectrum in our simulations is depicted in Fig. 4. Because of the prior, it is suppressed with respect to the input spectrum. A dimensionless kernel is then applied to the spectrum to correct for the suppression.

The construction of the kernel in step 3 is the computationally heaviest part in the algorithm. To construct matrix $\mathbf{M}$, we need

to construct and invert the full deconvolution matrix. `ArtDeco` contains routines that allow us to evaluate any element of the matrix. Storing and inverting the matrix, however, soon becomes unfeasible, as the matrix rank increases as proportional to $3(\ell_{max}+1)^2$. We were able to compute the full kernel up to $\ell_{max} = 200$. The construction of the kernel took 4 h on 1440 CPUs.

Constructing the raw spectrum is a lighter task. We run the deconvolution at full resolution, then extract the multipoles up to the ℓ_{max} of the kernel.

5.5 Strong signal limit

It is instructive to look at the optimal solution in the ideal case, where the data cover the full sky, and noise is absent or negligible. As we saw in Section 3, in this case, direct deconvolution recovers the input spectrum to a very high accuracy. If the CMB signal is very strong compared with noise, $\mathbf{S}^{-1}$ becomes negligible compared with $\mathbf{A}^{\dagger}\mathbf{N}^{-1}\mathbf{A}$, and we can make the approximation

$$\tilde{\mathbf{M}} \approx \mathbf{I}. \tag{47}$$

The requirement of complete sky coverage is essential for matrix $\mathbf{A}^{\dagger}\mathbf{N}^{-1}\mathbf{A}$ to be invertible. The kernel reduces into a unity matrix $\tilde{\Lambda}_{LL'} = \delta_{LL'}$, and we see that the spectrum of the deconvolved $\hat{a}_{s\ell m}$ coefficients is indeed the optimal solution.

6 HIGH MULTIPOLES: WEAK SIGNAL APPROXIMATION

6.1 Weak signal limit

The optimal method of Section 5 is only feasible at a limited multipole range. To cover the whole multipole range of interest (in our fiducial simulation up to $\ell_{max}=1500$), we have to find another solution.

In Section 5.5, we considered the extreme case where noise is negligible and sky coverage is complete. Consider now the opposite limiting case, where the signal is weak compared to noise,

$$\mathbf{S}^{-1} \gg \mathbf{A}^{\dagger}\mathbf{N}^{-1}\mathbf{A}. \tag{48}$$

Under this assumption, we can approximate

$$(\mathbf{S}^{-1} + \mathbf{A}^{\dagger}\mathbf{N}^{-1}\mathbf{A})^{-1} \approx \mathbf{S} - \mathbf{S}\mathbf{A}^{\dagger}\mathbf{N}^{-1}\mathbf{A}\mathbf{S}. \tag{49}$$

Matrix M of equation (42) then simplifies into

$$\mathbf{M} \approx \mathbf{A}^{\dagger}\mathbf{N}^{-1}\mathbf{A}. \tag{50}$$

An explicit formula for $\mathbf{A}^{\dagger}\mathbf{N}^{-1}\mathbf{A}$ for one detector is given by Keihänen & Reinecke (2012) as

$$(\mathbf{A}^{\dagger}\mathbf{N}^{-1}\mathbf{A})_{s\ell m, s'\ell'm'}$$

$$= \frac{1}{\sigma^2} \sum_{kk'} (-1)^{m'+k'} b_{s\ell k} b_{s'\ell'k'}^* \sum_{\ell_2} (2\ell_2+1) W_{m'-m,k'-k}^{\ell_2}$$

$$\times \begin{pmatrix} \ell & \ell' & \ell_2 \\ m & -m' & (m'-m) \end{pmatrix} \begin{pmatrix} \ell & \ell' & \ell_2 \\ k & -k' & (k'-k) \end{pmatrix}. \tag{51}$$

Here, $b_{s\ell k}$ is the harmonic beam expansion; σ^2 is the white noise variance in TOI domain, and W is the Wigner transform of the pointing distribution. If several detectors are involved, equation (51) is replaced by the sum over all detectors.

The Wigner transform W is computed as an expansion in Wigner functions of the 3D map data structure, which is the input format assumed by the `artDeco` deconvolver. A 3D map is a

three-dimensional data object constructed from a time-ordered data stream through a binning operation that combines data samples with similar pointing. Two of the dimensions are equivalent to the θ, ϕ angles that define an ordinary HEALPIX map; the third keeps track of the distribution of beam orientations. The Wigner functions offer a natural orthogonal set of base functions for this data structure.

When we insert equation (51) into the kernel formula (41), we obtain a formula with four 3j symbols. We disconnect W from indices m, m' by writing

$$W_{m'-m,k'-k}^{\ell_2} \begin{pmatrix} \ell & \ell' & \ell_2 \\ m & -m' & m'-m \end{pmatrix} =$$

$$\sum_{m_2} W_{m_2,k'-k}^{\ell_2} \begin{pmatrix} \ell & \ell' & \ell_2 \\ m & -m' & m_2 \end{pmatrix}. \tag{52}$$

All dependence on indices m, m' is now in the 3j symbols. The sums over m, m' can be carried out with the help of the relation

$$\sum_{mm'} \begin{pmatrix} \ell & \ell' & \ell_2 \\ m & m' & m_2 \end{pmatrix} \begin{pmatrix} \ell & \ell' & \ell_3 \\ m & m' & m_3 \end{pmatrix} = (2\ell_2+1)^{-1}\delta_{\ell_2\ell_3}\delta_{m_2m_3}. \tag{53}$$

All this put together, we obtain for the kernel matrix

$$\Lambda_{u\ell,u\ell'}$$

$$= \sum_{\alpha\beta} \frac{1}{\sigma_\alpha^2} \frac{1}{\sigma_\beta^2} \sum_{k_1k_2k_3k_4} \sum_{s_1s_2} g_u^{s_1s_2} b_{s_1\ell k_1}^{1*} b_{s_2\ell k_2}^2 \sum_{s_3s_4} g_{u'}^{s_3s_4} b_{s_3\ell'k_3}^1 b_{s_4\ell'k_4}^{2*}$$

$$\times \sum_{\ell_2} (-1)^{k_3+k_4} \begin{pmatrix} \ell & \ell' & \ell_2 \\ k_1 & -k_3 & k_3-k_1 \end{pmatrix} \begin{pmatrix} \ell & \ell' & \ell_2 \\ k_2 & -k_4 & k_4-k_2 \end{pmatrix}$$

$$\times (2\ell_2+1) \sum_{m_2} W_{m_2,k_3-k_1}^{1\ell_2*} W_{m_2,k_4-k_2}^{2\ell_2}, \tag{54}$$

where α, β label the detectors involved in the deconvolution process. This kernel is much cheaper to evaluate than the full kernel that involves inverting the deconvolution matrix. The required inputs are the beam expansion $b_{s\ell k}$ and the pointing Wigner transform W. The latter we can extract from the `artDeco` code.

Consider then the raw spectrum of equation (40). Under the same approximation of equation (49), it simplifies into

$$P_L = \boldsymbol{t}^T \mathbf{N}^{-1}\mathbf{A}^{\dagger}\mathbf{G}^L\mathbf{A}\mathbf{N}^{-1}\boldsymbol{t}. \tag{55}$$

This is equivalent to taking the spectrum of the right-hand side of the deconvolution equation. Note that the prior has disappeared both in the raw spectrum and in the kernel. The solution thus does not depend on a priori information of the CMB spectrum, as long as we can assume that the CMB signal is weak in comparison with noise.

6.2 Uniformization

The obvious problem with the recipe above is that for experiments like *Planck*, the assumption of weak signal does not hold. In fact, an attempt to apply the method presented above as it is leads to a gross misestimation of all the spectra. We can, however, improve the method through preprocessing of the data. A clue is found in Hivon et al. (2002), where instead of the actual hit distribution, a uniform one is assumed when weighting the input map. It turns out that uniformization of the hit count distribution is an essential step in our method too.

The error in a quadratic PSE method comes from two sources. One is the instrument noise and the other is the statistical variation of the harmonic coefficients. In the process of deriving the optimal
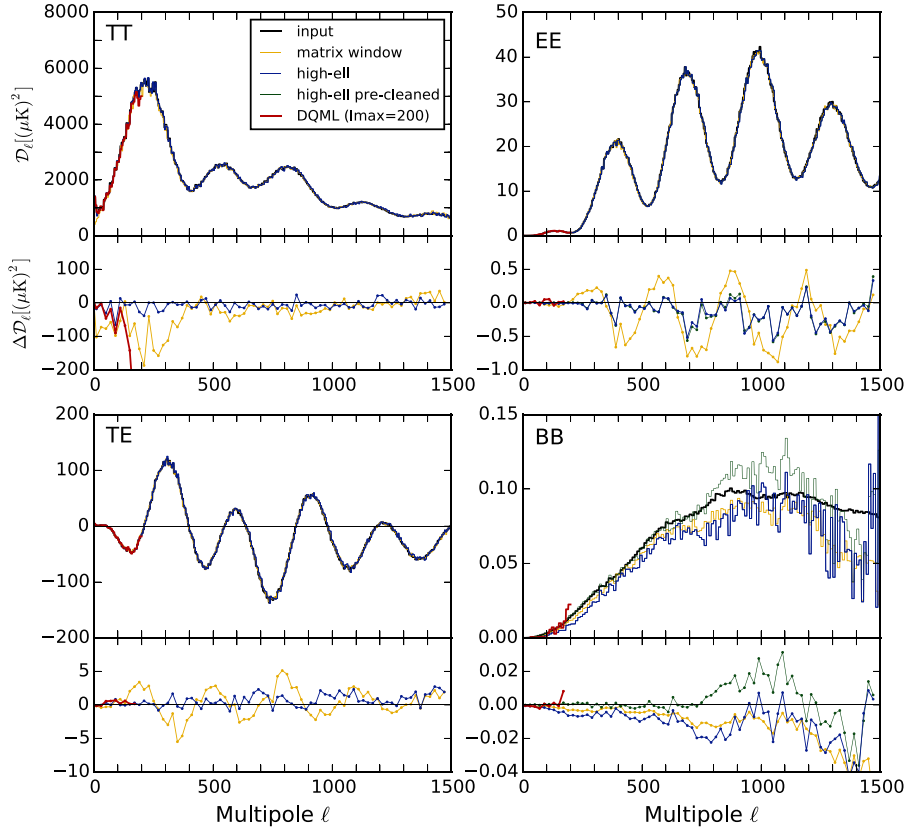
Figure 5. Recovered 70 GHz CMB spectra. We show the spectrum of the input realization (black) along with recovered estimates with the DQML method up to $\ell_{\max} = 200$ and the high-ell deconvolution method with $\ell_{\max} = 1500$. For comparison, we show also the estimate from the matrix window method (best estimate from Fig. 3). The spectra are averaged over 5 adjacent multipoles (10 for BB), the error over 20 multipoles. Note the tighter scale of the error plot as compared to Fig. 3. We show also the effect from pre-cleaning of the temperature signal (only visible in EE and BB).

method, we made the simplification where we replaced the products of the form $a_{s\ell m} a_{s'\ell'm'}$ by their ensemble average. The actual products differ from the ensemble average. This deviation acts as 'noise' in the spectrum estimation. The DQML method finds the optimal balance between the error sources and minimizes their combined effect. The approximation of equation (54), instead, ignores the CMB-induced error component. Thus individual strong T multipoles may leak into the regime of a much weaker signal, destroying the estimate there. To prevent this, we would like to minimize the connection between distant multipoles.

Analysis of the deconvolution matrix reveals that the correlation between distant multipoles is related to the uneven distribution of measurements due to scanning strategy. This motivates the following processing step. We *uniformize* the sample distribution by setting the total hit count for each HEALPIX pixel that is observed at all, to a common value that is taken to be the average hit count. This is done by rescaling of the input 3D map object. In other words, we are reducing the weight of frequently scanned sky pixels. The distribution of beam orientations within a pixel is left intact. Further, when working with *Planck* simulations, we set the assumed noise levels for a detector pair sharing a horn to the same value. This works further into the direction of reducing the leakage between temperature and polarization. We recompute the Wigner transform W with the uniformized distribution as input. This way, as it turns out, we are able to estimate both the temperature and polarization spectra over the full multipole range with very good accuracy.

There is a remarkable similarity between the kernel equation (54) and the kernel involved in the Master method (Hivon et al. 2002). An immediate difference is the presence of $k \neq 0$ beam coefficients and related 3j symbols in our solution. If we assume perfect delta beams, $b_{0\ell k} = \sqrt{2\ell + 1}\delta_{k0}$, the only non-zero terms of equation (54) are those involving W^ℓ_{m0}, which is equivalent to the usual harmonic expansion of the hit count. Under this assumption, our result becomes identical to that of Master.

In the case of polarization, our kernel can be compared to the corresponding one presented by Kogut et al. (2003) and Grain et al. (2009). Despite of obvious similarities, here it is more difficult to see how the methods relate, since the map-based methods involve dividing the polarization signal into Q and U components, something that does not exist in our method. What can be said is that assuming perfect delta beams alone is not sufficient to make the results identical.

7 RESULTS

7.1 Simulations

We have validated both deconvolution methods with simulations. The simulation procedure was described in Section 2. As prior in the DQML algorithm, we use a CMB spectrum generated from slightly different cosmological parameters than those used for the input spectrum. The TT prior depicted in Fig. 4.
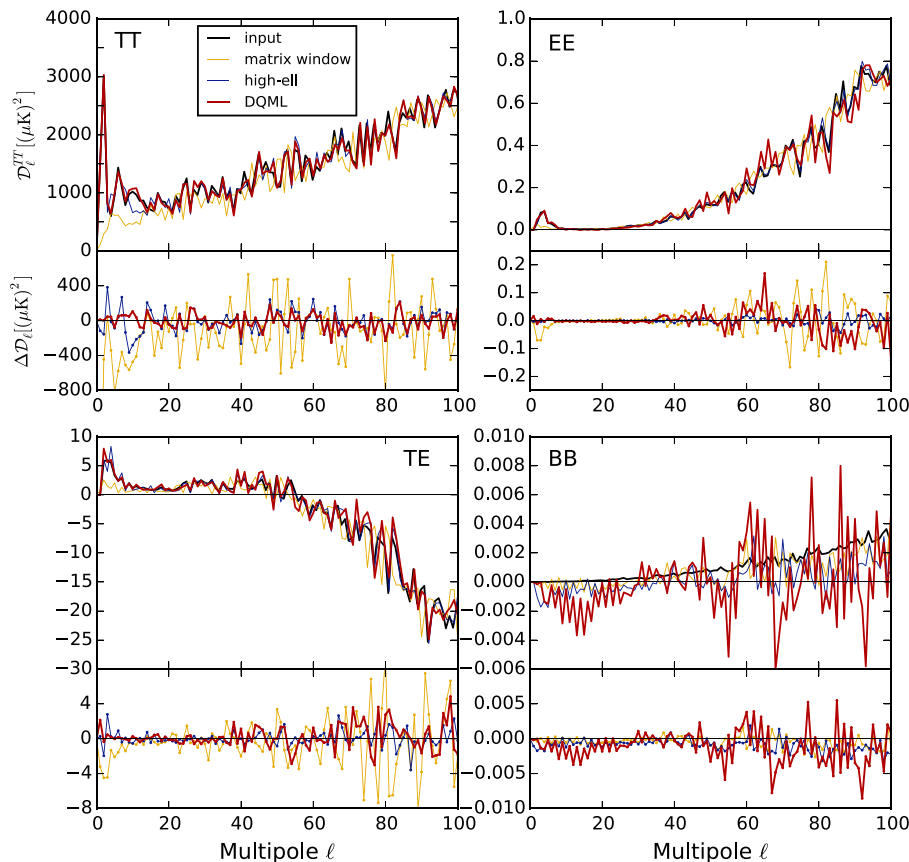
Figure 6. Zoom into the low-multipole region. Data and line types are the same as in Fig. 5. We show all the individual multipoles without binning.

The main results for the fiducial 70 GHz simulation are shown in Fig. 5. We plot in the same figure spectrum estimates from the DQML method of Section 5 and the high-ell method of Section 6. We apply the DQML method to multipoles $\ell = 0$–200, the high-ell method to multipoles $\ell = 0$–1500. For comparison, we show also results from the `PolSpice` mask correction combined with matrix window correction, which was the best of the pixel-based methods of Section 4. We refer to this combined method for brevity as 'matrix window' method.

All three methods recover the CMB spectrum well in their applicable multipole range. An exception is the BB spectrum where the high-multipole methods show a small negative bias. Differences between the methods become visible in the lower panel, where we plot the absolute error (difference between the recovered spectrum and the input spectrum). We observe that in all spectra but BB, the high-ell deconvolution method yields a still more accurate estimate than the matrix window method. The improvement is particularly clear in TT around the first acoustic peak. Also, the matrix window method for EE and TE shows residuals of temperature leakage, which are reduced by the high-ell deconvolution method. In BB, matrix window and high-ell deconvolution give similar results, unless we apply pre-cleaning (see Section 7.2).

The drop in the DQML estimate for TT above $\ell = 100$ is a fringe effect, caused by the limited $\ell_{\max}$ value. Extending the multipole range across the first acoustic peak would probably reduce the effect.

We blow up the low-multipole region in Fig. 6. Here, we show the individual multipoles without binning. Consequently, the amplitude of error is larger than in Fig. 5. The high-ell method is again superior to the matrix window method, and in EE and and in BB, it

is also more accurate than the DQML method. At the very lowest TT multipoles ($\ell < 40$), the DQML method gives the most accurate estimate. The differences, however, are at the same level as the cosmic variance. We see also that the two deconvolution methods give nearly identical results in TT at intermediate multipoles ($20 < \ell < 100$). This means that the two estimates can be joined without discontinuity. The TE spectrum behaves similarly to TT, i.e. DQML is more accurate at lowest multipoles, but the differences are less clear than in TT.

7.2 Pre-cleaning

As discussed in Section 6, the strong low-ell T signal leaking to higher multipoles or to polarization interferes with the estimation at these weaker spectral components. The procedure of uniformizing the hit count distribution was introduced to mitigate the problem. We can do more than that as follows. We run the usual deconvolution procedure (with prior) to obtain an estimate of the $a_{s\ell m}$ coefficients of the sky. We then scan the T components back into a TOI stream, subtract it from the original TOI, and perform the usual PSE procedure on the cleaned data, to obtain a better estimate of the weaker spectrum components. The signal components to be subtracted can be chosen in many ways. In Fig. 5, we show the effect on EE and BB of removing all of the T signal. In EE the effect is negligible, but in BB pre-cleaning step removes the negative bias almost perfectly up to $\ell = 800$.

In the case above, pre-cleaning has little practical meaning, since the small bias in the BB spectrum is well below the noise level of any current CMB experiment. In some cases, however, pre-cleaning
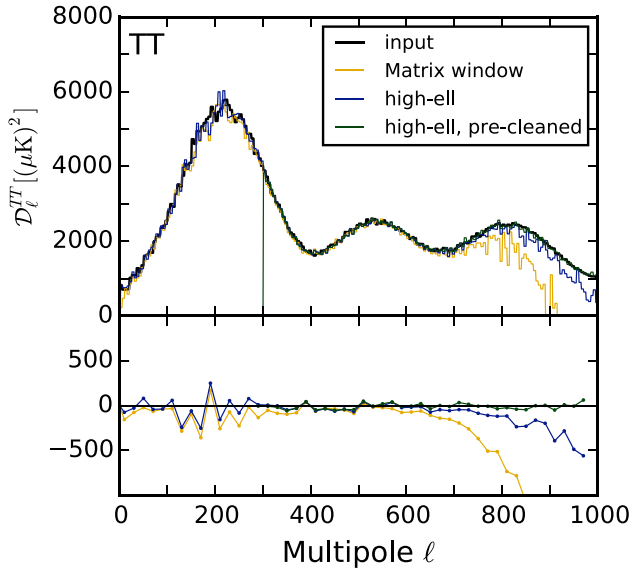
Figure 7. An example of a case where pre-cleaning is important. Shown is the recovered 44 GHz TT spectrum compared against input. The strong quadrupole moment leaks into the high-multipole regime in a way that is not captured by the spectrum estimation methods. Subtracting an estimate of the low-multipole signal up to $\ell = 300$ removes the bias.

may become important. One such case is shown in Fig. 7. We show the recovered TT spectrum from a simulation involving the 44 GHz channel of *Planck*. Beam leakage effects are known to be particularly strong for this channel. Without pre-cleaning, the spectrum is underestimated at high multipoles ($\ell < 700$). Note that this is not specific to deconvolution, but the matrix window method suffers from the same problem. Analysis of individual multipoles revealed that the bias arises from the very lowest temperature multipoles, which leak into the high-multipole regime. We pre-cleaned the 44 GHz data set by subtracting the deconvolved T multipoles up to $\ell_{\mathrm{max}} = 300$. This removed the bias entirely. This particular realization happened to have a very strong quadrupole, which was found to be responsible for most of the leakage. The quadrupole of the actual CMB sky, instead, is known to be small. The simulation shown here probably represents an overly pessimistic case. However, leakage effects depend on the specifics beam shapes and the scanning strategy of a particular experiment, and it is difficult to predict where such effects may arise. It is thus advisable to check the robustness of the results against pre-cleaning parts of the data, or else to validate the results with simulations.

7.3 Instrument noise

At first, it seems surprising that the optimal method appears to be inferior to the high-ell method for all but the lowest TT multipoles, considering that the optimal method was specifically designed to minimize the error. There is an explanation for this. When deriving the optimal method, we required that it minimizes the *combined* error arising from statistical variation in the CMB itself, and from instrument noise. So far, we have applied the method to a pure CMB simulation. It is conceivable that when instrument noise is taken into account, the optimal method will give a smaller total error.

To verify if this is the case, we perform a series of noise simulations. We generate 100 realizations of white noise, at the level of *Planck* 70 GHz channel. The parameters were taken from Planck
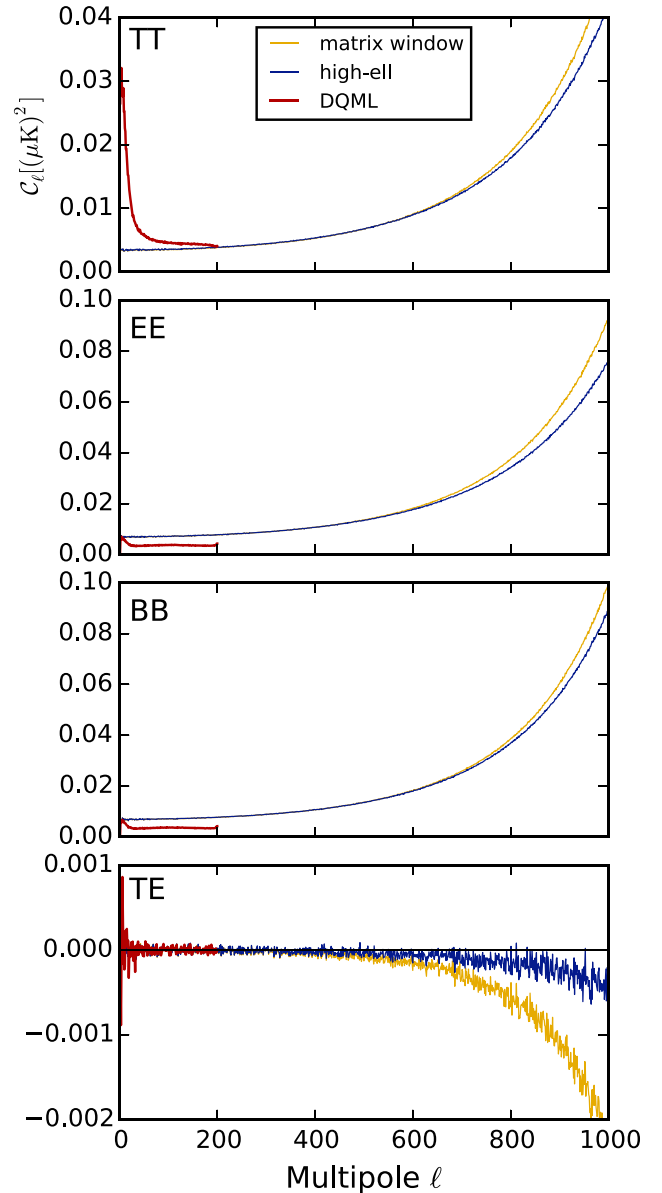

Figure 8. Noise bias, assuming white noise level of *Planck* 70 GHz detectors, as estimated from Monte Carlo simulations. The noise bias is plotted in the $\mathcal{C}_\ell$ convention, i.e. without scaling by $\ell(\ell + 1)/(2\pi)$.

Collaboration II (2016). We processed the noise through the same PSE pipeline as the signal simulations, with the three methods in question. For each method, we obtain 100 noise spectra with realistic properties. We compute the *noise bias* as the average over the 100 noise spectra. The noise bias for the three methods are plotted in Fig. 8. We plot the spectrum without the $\ell(\ell + 1)/(2\pi)$ scaling, to avoid excessive amplification at high multipoles. We cut the plot at $\ell = 1000$ to show the low-multipole region more clearly.

The noise bias for the matrix window method and for the high-ell deconvolution method are very close at low multipoles. At high multipoles ($\ell > 600$), the bias for the matrix window method begins to increase more steeply. The noise bias for the DQML method behaves very differently.

The noise bias itself is not a good measure of residual noise. The noise bias can be estimated through Monte Carlo simulations, and subtracted from the actual spectrum. What is more interesting is the
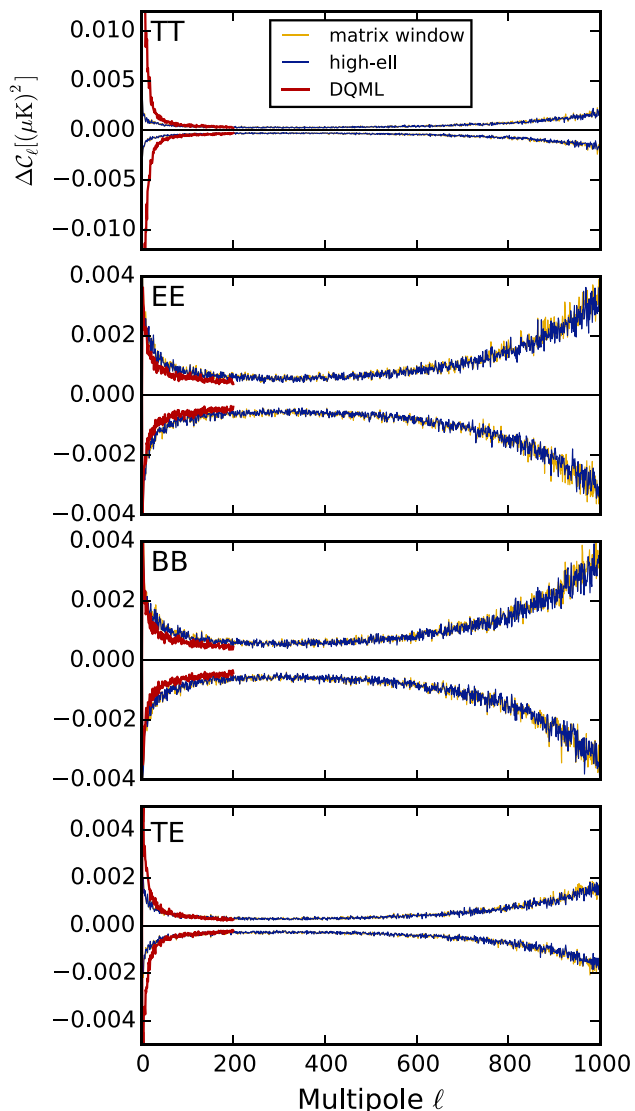
Figure 9. Residual noise level in the recovered spectra, after subtraction of the average noise bias. We assume the white noise level of *Planck* 70 GHz detectors. Shown is the 1σ variation of the residual noise in both directions in $\mathcal{C}_\ell$ units, as estimated from 100 Monte Carlo realizations.

variation of a single realization that remains after the subtraction of the mean noise bias. This is the residual noise in the final spectrum estimate. We estimated the residual noise level from the same MC simulations that we used for the estimation of the noise bias. We subtracted the mean bias from the 100 noise spectra, sorted the 100 values at each individual multipole, and picked the 16th and 85th value. These values correspond to one-sigma variation in both directions around the mean noise bias.

The results are shown in Fig. 9. The DQML method yields a higher residual noise than the high-ell method in the TT spectrum. The noise level is, however, negligible compared with the signal-induced error of Fig. 6, where DQML performed better. In EE and BB components, the DQML method gives a significantly lower residual noise level than the high-ell method. Since noise is the dominant error source in these spectra, this is consistent with the theoretical expectation that the DQML method minimizes the total error.
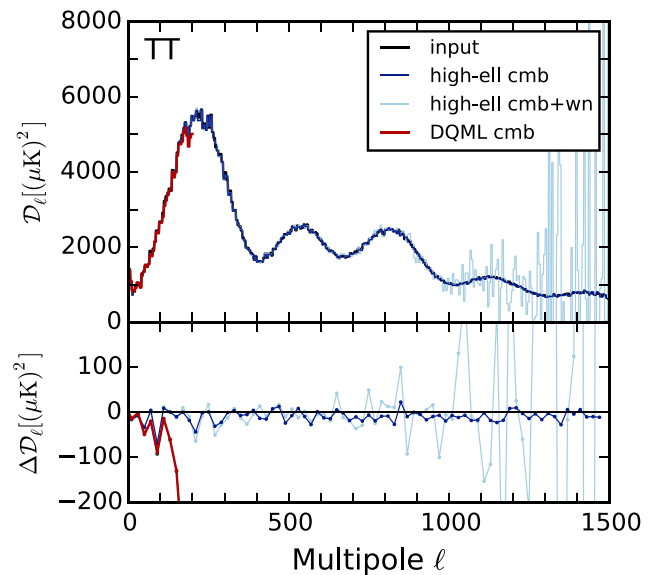

Figure 10. Cross-power spectrum estimation. Shown is the 70 GHz TT cross-spectrum between two 2-yr data sets, for pure CMB (dark blue) and for one realization of white noise at *Planck* level. The fiducial 4-yr data set is split into two halves, which are cross-correlated to give an unbiased spectrum estimate.

7.4 Cross-power spectrum

In this work, we have made the simplifying assumption that the instrument noise is white. This not exactly true for realistic instruments. The TOI of *Planck*, for instance, is contaminated by slowly varying $1/f$ noise. The correlated noise component can partly be removed by destriping techniques, but there remains a residual that contaminates the lowest multipoles. If noise properties are well known, the noise bias can be estimated through MC simulations and subtracted from the spectrum estimate, to yield an unbiased spectrum estimate.

Another possibility is to cross-correlate two independent data sets. If noise is uncorrelated from one set to another, the noise bias vanishes and the estimate is again unbiased. Both deconvolution PSE methods presented in this paper extend trivially to cross-power estimation. Equation (40) is replaced by

$$P_L = \hat{a}_1{}^\dagger \mathbf{S}^{-1} \mathbf{G}^{L\dagger} \mathbf{S}^{-1} \hat{a}_2, \tag{56}$$

where index 1, 2 refers to the two data sets. The kernel of equation (41) takes the form

$$\Lambda_{LL'} = \mathrm{Tr}\left[\mathbf{M}_1 \mathbf{G}^L \mathbf{M}_2^\dagger \mathbf{G}^{\dagger L'} \right]. \tag{57}$$

The high-ell kernel of equation (54) is already written for a detector pair, and requires no modification.

We demonstrate the cross-power estimation in Fig. 10. We split the fiducial 70 GHz simulation into two halves, one consisting of years 1–2 and the other of years 3–4. The cross-power TT estimate is remarkably similar to the auto-spectrum estimate of Fig. 5. The same lows and bumps are present in both error plots, indicating that the error arises from the CMB sky itself.

We plot also the cross-power spectrum from a simulation where we add one realization of white noise at *Planck* level. In the range $\ell < 200$ where the DQML method is applied, noise is negligible. The results from the simulation with noise are practically identical to those from the pure CMB simulation and are not shown separately.

8 STATISTICAL KERNEL EVALUATION

While the light-weight high-ell method works very well in most cases, our simulations have indicated that the optimal DQML method leaves a lower residual noise level. It was also superior in determining the lowest signal-dominated TT multipoles. With future CMB experiments also, the polarization signal may enter the regime of signal domination. It is thus worth the effort to explore ways of extending the method to a wider multipole range. The bottleneck is the current implementation is the construction and inversion of the deconvolution matrix. The rank of the matrix grows as proportional to $\ell_{\max}^2$, and the CPU cost of the inversion as $\ell_{\max}^6$. It is thus unlikely that the multipole limit can be pushed much further even with increased computational resources. A numerical method that allows us to evaluate the kernel of equation (41) without explicit matrix inversion would be highly welcome, but such a method is not in sight.

We explore an alternative way of evaluating the kernel matrix, involving Monte Carlo simulations. To start, we rewrite matrix **M** from equation (42) as

$$\mathbf{M} = \mathbf{A}^\dagger \mathbf{N}^{-1} \mathbf{A} (\mathbf{S}^{-1} + \mathbf{A}^\dagger \mathbf{N}^{-1} \mathbf{A})^{-1} \mathbf{S}^{-1}. \tag{58}$$

Here, $\mathbf{N} = \langle \boldsymbol{n}\boldsymbol{n}^T \rangle$ is the white noise covariance in TOI domain. We rewrite further

$$\mathbf{N}^{-1} = \mathbf{N}^{-1} \langle \boldsymbol{n}\boldsymbol{n}^T \rangle \mathbf{N}^{-1}, \tag{59}$$

where $\boldsymbol{n}$ denotes a white noise realization. We insert this into the first occurrence of N of equation (58) and that further into equation (41). Since M appears there twice, we must introduce two independent noise realizations $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$, so that we can combine the averages as

$$\langle \boldsymbol{n}_1 \boldsymbol{n}_1^T \rangle \langle \boldsymbol{n}_2 \boldsymbol{n}_2^T \rangle = \langle \boldsymbol{n}_1 \boldsymbol{n}_1^T \boldsymbol{n}_2 \boldsymbol{n}_2^T \rangle. \tag{60}$$

Because a permutation operation leaves the trace unchanged, we can rearrange the terms to obtain

$$\Lambda = \Big\langle \mathrm{Tr}[\boldsymbol{n}_1^T \mathbf{N}^{-1} \mathbf{A} \mathbf{G}^L \mathbf{A}^\dagger \mathbf{N}^{-1} \boldsymbol{n}_2 \boldsymbol{n}_2^T \mathbf{N}^{-1} \mathbf{A} (\mathbf{S}^{-1} + \mathbf{A}^\dagger \mathbf{N}^{-1} \mathbf{A})^{-1} \mathbf{S}^{-1}$$

$$\mathbf{G}^{L'\dagger} \mathbf{S}^{-1} (\mathbf{S}^{-1} + \mathbf{A}^\dagger \mathbf{N}^{-1} \mathbf{A})^{-1} \mathbf{A}^\dagger \mathbf{N}^{-1} \boldsymbol{n}_2] \Big\rangle. \tag{61}$$

This lengthy construction actually consists of two scalar terms of the form $\boldsymbol{n}_1^T \cdots \boldsymbol{n}_2$. We can thus drop the trace operation (the trace of a scalar is the scalar itself). The kernel as given by equation (61) can be evaluated through MC simulations as follows. We first generate a sequence of white noise realizations $\boldsymbol{n}_k, k = 1 \ldots n_{\mathrm{mc}}$. For each realization, we evaluate the objects

$$\boldsymbol{a}_k = (\mathbf{S}^{-1} + \mathbf{A}^\dagger \mathbf{N}^{-1} \mathbf{A})^{-1} \mathbf{A}^\dagger \mathbf{N}^{-1} \boldsymbol{n}_k$$

$$\boldsymbol{r}_k = \mathbf{A}^\dagger \mathbf{N}^{-1} \boldsymbol{n}_k. \tag{62}$$

Both objects have the structure of an $a_{s\ell m}$ vector. The first line represents the usual deconvolution operation. The second operation is simply the right-hand side of the deconvolution equation. Both products are standard outputs of `artDeco` code. We thus already have all the necessary machinery in place. We evaluate the spectra

$$P_r^L = \boldsymbol{r}_j^T \mathbf{G}^L \boldsymbol{r}_k$$

$$P_a^L = \boldsymbol{a}_j^T \mathbf{S}^{-1} \mathbf{G}^L \mathbf{S}^{-1} \boldsymbol{a}_k \tag{63}$$
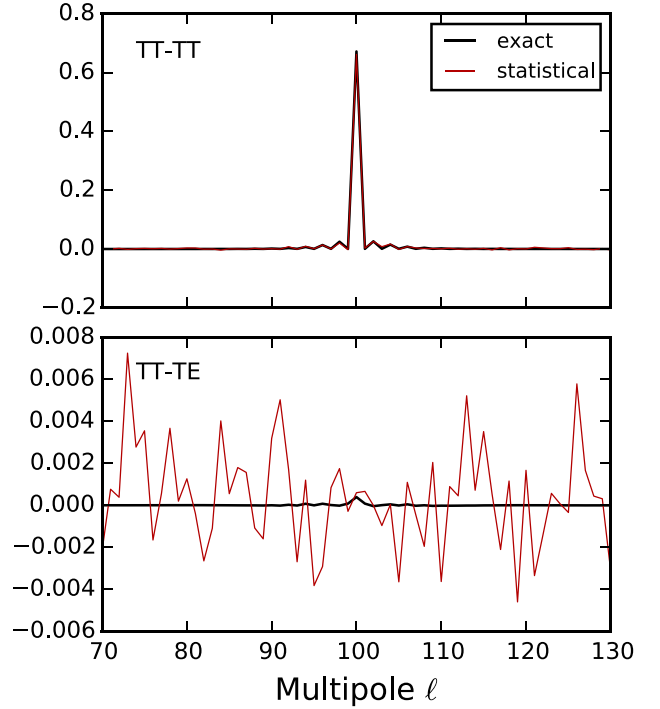
Figure 11. Selected elements of the 30 GHz kernel, evaluated through the exact numerical method of Section 5 and the statistical method of Section 8. Shown are elements along the row corresponding to multipole $\ell = 100$ in TT. The TT–TT elements are recovered well with the statistical method, but the number of realizations (800) is insufficient for the weaker TT–TE elements.

for all realization pairs $j \neq k$. This is equivalent to the familiar operation of evaluating the cross-spectrum between two harmonic vectors, apart from normalization. An estimate for the kernel is obtained as an average over all pairs

$$\hat{\Lambda}_{LL'} = \frac{1}{n_{\mathrm{mc}}(n_{\mathrm{mc}} - 1)} \sum_{j \neq k} \boldsymbol{r}_j^\dagger \mathbf{G}^L \boldsymbol{r}_k \cdot \boldsymbol{a}_k^\dagger \mathbf{S}^{-1} \mathbf{G}^{L'} \mathbf{S}^{-1} \boldsymbol{a}_j, \tag{64}$$

where $n_{\mathrm{mc}}(n_{\mathrm{mc}} - 1)$ counts the pairs $j \neq k$. The estimate converges to the true kernel when $n_{\mathrm{mc}} \to \infty$. Note that the noise simulations here have nothing to do with the actual noise properties of the data. We are simply evaluating the diagonal matrix **N** statistically.

To validate the procedure, we pick the less demanding *Planck* 30 GHz channel, and generate 800 white noise realizations. The procedure took roughly one minute of wall-clock time per realization on 768 cores. We evaluate the kernel matrix with the statistical method up to $\ell_{\max} = 800$, and with exact DQML up to $\ell_{\max} = 200$ and compare the kernel elements in the multipole range covered by both methods. We plot elements of row $\ell = 100$ in Fig. 11. The 800 realizations we have are sufficient to recover the structure of the TT–TT block of the kernel, but the weaker TT–TE cross terms are totally buried under statistical variation. We conclude that a lot more realizations are required, for the method to become useful in practice.

9 CONCLUSIONS

We have studied the possibility of combining beam deconvolution technique with PSE for absolute CMB measurements. We present two new PSE methods, one suitable for low multipoles, the other

for a wide multipole range. The methods yield an estimate for the angular CMB power spectrum in both temperature and polarization. The methods correct simultaneously the effect of incomplete sky coverage and of asymmetric beam shape. In particular, they correct the leakage from temperature to polarization through beam shape mismatch. The required inputs are the time-ordered data, pointing information and known beam shapes.

The first method (DQML) was derived formally as the quadratic estimator that fulfils two requirements: it yields a non-biased estimate of the CMB spectrum and, at the same time, minimizes the residual error in the estimate. The residual error is a combination of instrument noise and the error that arises from the statistical variation of the individual $a_{s\ell m}$ sky coefficients.

We developed the optimal solution into a practical algorithm. The procedure consists of beam-deconvolving the data with a first-guess CMB spectrum as prior, and correcting the spectrum with a kernel that takes into account the exact pointing distribution and beam shapes. The computational burden grows steeply with increasing ℓ_{max}. We successfully applied it to the multipoles in range $\ell = 0$–200.

For the high-multipole regime, we derived another solution, which is suboptimal in the sense that it does not minimize the residual error, but still yields an unbiased spectrum estimate. Formally, it is derived as the weak-signal limit of the optimal DQML method. The high-ell method is computationally light and we apply it to multipoles up to $\ell = 1500$. The core of the method is a kernel matrix (equation 54), which simultaneously corrects for the sky coverage and for beam effects. The kernel resembles the one implemented in the `Master` method, but includes additional terms that take into account the beam.

We validated the methods with simulated data. Our fiducial data set mimics that of the *Planck* LFI 70 GHz channel. We used realistic *Planck* beams and scanning strategy. We applied a Galactic mask with $f_{sky} = 0.8967$ sky coverage. We found the DQML method to be superior when estimating the TT spectrum in the low-multipole range ($\ell < 40$).

The high-ell method appears superior to the DQML method in the polarization components, when we deal with pure CMB simulations. When we add white noise at *Planck* level, however, we observe that DQML produces a lower residual noise level.

The best estimate for the CMB spectrum is obtained by a combination of the two deconvolution methods. For low multipoles, it makes sense to use the DQML method; for high multipoles, the high-ell method. The two estimates overlap cleanly at intermediate multipoles with no apparent discontinuity.

We compared the new methods with selected pixel-based PSE methods. The most accurate of those was one that combines `Pol-Spice` mask correction with a matrix beam window function. The new methods were found to be still more accurate in the studied simulation case. Apart from accuracy, the new methods offer the benefit that they avoid the need for heavy CMB Monte Carlo simulations that are usually needed in the evaluation of a beam window function.

The new methods operate at TOI level, which is the natural domain for handling effects that depend on time or on detector orientation. While this work focuses on asymmetric beams, we can consider the possibility of extending the general idea to other time-dependent effects. For instance, a long detector time constant can be included in the beam model. Further, in the current implementation, it is assumed that residual noise at TOI level is at least approximately white. The formulation is more general, however, and could in principle be extended to incorporate correlated noise residuals. This

is achieved by replacing the diagonal covariance **N** by an appropriate noise filtre.

Beam effects are usually thought to be unimportant at low multipoles. One may therefore ask if the DQML method provides any benefit over the beamless QML method. This is a valid question and a definite answer would require running DQML and `BolPol` or equivalent side by side. There is, however, an important difference between the methods, unrelated to beam shapes. The QML method, as implemented in `BolPol`, takes as input low-resolution CMB maps. The process of downgrading a high-resolution map is not a trivial one, and it can be done in different ways (the problem in the context of *Planck* is discussed in Planck Collaboration V (2016)). The DQML method avoids all these issues, as it always operates at the full resolution of the input 3D map. The transition to low resolution occurs in a natural way in harmonic space, where we simply cut the raw spectrum at the desired ℓ_{max} before applying the kernel. The DQML method is thus not equivalent to the QML method, even if we assume perfect delta beams.

Finally, we studied the possibility of extending the DQML method to higher multipoles. We showed that the kernel matrix can be evaluated statistically through white noise Monte Carlo simulations, but found that 800 noise realizations do not give sufficient accuracy yet. Still, this is a promising idea to study further, since the error variance goes down as inversely proportional to the number of MC realizations. This is to be contrasted with the analytical method, where the computational burden grows as proportional to ℓ_{max}^6 with increasing ℓ_{max}.

REFERENCES

Armitage C., Wandelt B. D., 2004, Phys. Rev. D, 70, 123007
Armitage-Caplan C., Wandelt B. D., 2009, ApJS, 181, 533
Chon G., Challinor A., Prunet S., Hivon E., Szapudi I., 2004, MNRAS, 350, 914
Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759
Grain J., Tristam M., Stompor R., 2009, Phys. Rev. D, 79, 123515
Gruppuso A., De Rosa A., Cabella P., Paci F., Finelli F., Natoli P., de Gasperis G., Mandolesi N., 2009, MNRAS, 400, 463
Harrison D. L., van Leeuwen F., Ashdown M. A. J., 2011, A&A, 532, A55
Hivon E., Górski K., Netterfield C., Crill B., Prunet S., Hansen F., 2002, ApJ, 567, 2
Hivon E., Mottet S., Ponthieu N., 2016, A&A, preprint (arXiv:1608.08833)
Keihänen E., Reinecke M., 2012, A&A, 548, A110
Keihänen E., Kiiveri K., Lindholm V., Reinecke M., Suur-Uski A.-S., 2016, A&A, 587, A27
Kogut A. et al., 2003, ApJS, 148, 161

Mitra S., Rocha G., Górski K. M., Huffenberger K. M., Eriksen H. K., Ashdown M. A. J., Lawrence C. R., 2011, ApJS, 193, 5
Planck Collaboration I, 2016, A&A, 594, A1
Planck Collaboration II, 2016, A&A, 594, A2
Planck Collaboration IV, 2016, A&A, 594, A4
Planck Collaboration V, 2016, A&A, 594, A5
Planck Collaboration XII, 2016, A&A, 594, A12

Reinecke M., Dolag K., Hell R., Bartelmann M., Enßlin T. A., 2006, A&A, 445, 373
Tegmark M., 1997, Phys. Rev. D, 55, 5895
Tegmark M., Oliveira-Costa A., 2001, Phys. Rev. D, 64, 063001

This paper has been typeset from a TEX/LATEX file prepared by the author.