Research Article

CMB Map Restoration

J. Bobin,[1] J.-L. Starck,[1] F. Sureau,[1] and J. Fadili[2]

[1] *Laboratoire AIM, IRFU, SEDI-SAP, Service d'Astrophysique, Orme des Merisiers, Bat 709, Piece 282, 91191 Gif-Sur-Yvette, France*
[2] *GREYC CNRS, ENSICAEN, Université de Caen 6, Boulevard du Maréchal Juin, 14050 Caen Cedex, France*

Correspondence should be addressed to J. Bobin, jbobin@cea.fr

Estimating the cosmological microwave background is of utmost importance for cosmology. However, its estimation from full-sky surveys such as WMAP or more recently Planck is challenging: CMB maps are generally estimated via the application of some source separation techniques which never prevent the final map from being contaminated with noise and foreground residuals. These spurious contaminations whether noise or foreground residuals are well known to be a plague for most cosmologically relevant tests or evaluations; this includes CMB lensing reconstruction or non-Gaussian signatures search. Noise reduction is generally performed by applying a simple Wiener filter in spherical harmonics; however, this does not account for the non-stationarity of the noise. Foreground contamination is usually tackled by masking the most intense residuals detected in the map, which makes CMB evaluation harder to perform. In this paper, we introduce a novel noise reduction framework coined LIW-Filtering for Linear Iterative Wavelet Filtering which is able to account for the noise spatial variability thanks to a wavelet-based modeling while keeping the highly desired linearity of the Wiener filter. We further show that the same filtering technique can effectively perform foreground contamination reduction thus providing a globally cleaner CMB map. Numerical results on simulated Planck data are provided.

1. Introduction

In mid-2009, European Spatial Agency (ESA) put in orbit the latest space observatory Planck to investigate the Cosmological Microwave Background (CMB). These data are of particular scientific importance as it will provide more insight into the understanding of the birth of our universe. Most cosmological parameters can be derived from the study of these CMB data. After a series of successful CMB experiments (Archeops, Boomerang, Maxima, COBE, and WMAP [1]), the Planck ESA mission is providing more accurate data which will be the reference full-sky high-resolution CMB data for the next decades. More precisely, recovering useful scientific information requires disentangling the CMB from the contribution of several astrophysical components, namely, Galactic emissions from dust and synchrotron, Sunyaev-Zel'dovich (SZ) clusters, Free-Free emissions, CO emission to only name a few, see [2]. Classically, the CMB map is estimated via the application of some source separation methods. However, the application of very sophisticated source separation methods (see [3–6]) does not prevent the

final estimated map from being contaminated with various spurious components: (i) instrumental noise and (ii) foreground residuals. Dealing with these various sources of contaminations is of paramount importance for most cosmologically relevant tests or evaluations: this includes CMB lensing reconstruction [7] or non-Gaussian signatures search. Tackling the problem of noise and foreground residual reduction is therefore important and is at the heart of this paper.

(a) Instrumental Noise. The very specific scanning pattern in full-sky CMB experiments leads to an instrumental noise closely Gaussian that exhibits significant spatial variation: Planck noise is far from being homogeneous. More precisely, the statistics of the noise vary spatially; in the field of statistics, it is said that the noise is nonstationary. As an illustration, Figure 1 displays a Planck simulated root mean square (RMS) map of instrumental noise at 217 GHz. These data originate from the publicly available simulations described in [3]. Furthermore, while noise is insignificant at large scale, it is largely predominant at small scale (e.g., typically for $\ell > 1500$ for Planck). Most source separation methods
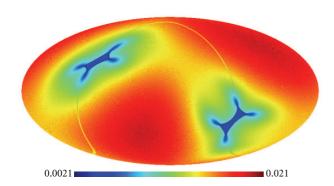
0.0021   0.021

FIGURE 1: *Instrumental noise in Planck*—simulated noise root mean square map at 217 GHz in mK (Antenna temperature). These simulated Planck data are described in [3].

applied so far to CMB data are linear methods: the estimated CMB map can be expressed as a linear combination of all the pixels of the input observations. As such, the propagated instrumental noise in the CMB maps is also typically assumed nonstationary Gaussian.

Noise can be a significant limitation for most cosmological tests or reconstructions. High noise level can hamper non-Gaussianity tests as the noisy map is likely to be "more Gaussian" than the noise-free map. The classical approach in the field of cosmology generally consists in reducing the noise of the CMB via a Wiener filter in spherical harmonics. If C_ℓ stands for the theoretical power spectrum of the CMB and C_ℓ^n the power spectrum of the noise, the Wiener filter is defined as follows:

$$\forall \ell, m > 0; \quad a_{\ell m}^{\hat{x}} = \frac{C_\ell}{C_\ell + C_\ell^n} \, a_{\ell m}^y, \qquad (1)$$

where $\{a_{\ell m}^{\hat{x}}\}_{\ell, m}$ (resp., $\{a_{\ell m}^y\}_{\ell, m}$) stands for the spherical harmonics coefficients of the filtered map $\hat{x}$ (resp, the input map y). While this approach makes profit of some knowledge of the energy distribution of the noise in frequency, it does not account for its spatial variations or nonstationarity. It is very likely that such a nonstationarity may create global non-Gaussian signatures at high frequency when such variations are sharp. Strong nonstationarities are known to have a strong impact on CMB lensing reconstruction (see [8]), creating an undesired reconstruction bias.

Another less traditional but commonly encountered noise reducing technique is the local Wiener filtering. It boils down to apply the Wiener filter in the Fourier space on patches in the pixel domain. The main drawback of this method is that while large patches should be favored to capture the nonstationarity of noise, they are enable to capture correlations at scales larger than the patch size.

(b) Foreground Residual. The first evaluation of already sophisticated CMB-dedicated source separation methods has been performed in [3]. Since, novel very effective methods have been proposed Needlet ILC [4], Generalized ILC [9], SMICA [10] or GMCA [6]. However, due to the very high

complexity of the data, none of these very effective methods is able to provide a CMB map that is guaranteed to be free of foreground contaminations. If the residual contamination is generally low at large scales (e.g., for $\ell < 500$ for Planck), it is generally significant at higher frequencies. It is likely that the remaining contamination mainly originates from point sources or strong features in the galactic plane. Obviously the presence of these residuals highly biases any cosmological tests to be performed on contaminated CMB map. The usual approach consists in masking the known spurious pixels of the CMB prior to any post-processing (e.g., 20 to 40% of the sky are usually masked). However masking raises several issues: (i) it limits the number of samples to which any test can be applied, limiting its statistical relevance and (ii) masking generally makes any post-processing much more complicated to perform properly (e.g., inpainting of the mask is generally needed prior to any CMB lensing reconstruction [7]). Avoiding or at least limiting the extent of the mask to be applied should greatly help improving any cosmological test to be made on the estimated CMB map. Furthermore, masking does not prevent from the presence of residuals of undetected point sources which are likely be be numerous all over the sky. Further reducing the amount of foreground residuals should clearly be helpful.

(c) Contribution of This Paper. The contribution of this paper is twofold:

(i) It introduces a novel noise reduction framework that, in opposition to the classical spherical-harmonics based Wiener filter, accounts for the nonstationarity of the noise. This new method preserves the linearity of the filtering. In fact, linearity is crucial in this field as it makes the study of error propagation via Monte-Carlo simulations much more convenient to carry out. In this framework, we make use of a simple, but efficient, wavelet-based modeling of the noise that allows for the modeling of nonstationarities at different scales. This will be discussed in Section 3.

(ii) We discuss an extension of this method to further estimate an approximate contribution of the foreground residuals per wavelet scales by means of RMS maps. Thereby contamination reduction could be tackled within the same filtering framework.

The noise/contamination reduction problem, recast as a quadratically regularized least square problem, is then effectively solved using recently introduced algorithms based on proximal calculus. Extensive numerical experiments are provided that show the effectiveness of the proposed method in reducing noise in Section 2.3 and foreground residuals in Section 3.3.

2. CMB Map Filtering with Nonstationary Noise

2.1. CMB Map Filtering. In a very general context and more precisely in the context of Planck, it is customary to assume that the input noise data, denoted by y in the following, is

modeled as the linear combination of the sought after noise-free signal x and the noise term n:

$$y = x + n. \tag{2}$$

Decorrelation between x and n is also classically assumed. In what follows, we will assume that each of these signals is sampled on the sphere according to the Healpix [11] sampling grid. The sampled CMB signal x is assumed to be composed of N samples $\{x[k]\}_{k=1,\dots,N}$. Recovering an estimate $\hat{x}$ of x can then be formulated as an inverse problem. According to the Bayesian way of solving inverse problems, some *a priori* knowledge about the signal of interest x has to be expressed. This inference framework is particularly well suited to the CMB denoising issue as strong *prior* assumptions can be formulated. Indeed, since [12], it is well established that the cosmological microwave background can be approximated as an isotropic Gaussian random field with a high accuracy. This particularly means that it is fully characterized by its power spectrum C_ℓ.

More precisely, the CMB x can be defined by its expansion in spherical harmonics:

$$x = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}. \tag{3}$$

Assuming that the CMB is fully Gaussian with power spectrum C_ℓ, its spherical harmonics $\{a_{\ell m}\}_{\ell m}$ are independently distributed following a Gaussian distribution with mean 0 and variance C_ℓ:

$$a_{\ell m} \sim \mathcal{N}(0, C_\ell). \tag{4}$$

The power spectrum of the CMB is generally estimated with the pseudo-power spectrum defined as follows:

$$\hat{C}_\ell = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} a_{\ell m}^{\star} a_{\ell m}, \tag{5}$$

where $a_{\ell m}^{\star}$ stands for the conjugate of $a_{\ell m}$. The *a priori* probability distribution $\mathcal{P}$ of the CMB map x is a normal distribution with covariance matrix $\mathcal{F}^H \mathbf{C} \mathcal{F}$ where $\mathcal{F}$ stands for the spherical harmonics basis and $\mathcal{F}^H$ its adjoint. The matrix $\mathbf{C}$ is diagonal with $\mathbf{C} = \mathrm{diag}(C_\ell)$. This yields the following CMB *prior* probability distribution:

$$\mathcal{P}(x) \propto \exp\left\{-x^T \mathcal{F}^H \mathbf{C}^{-1} \mathcal{F} x\right\}. \tag{6}$$

In this section, we assume that the instrumental noise can be modeled as a nonstationary but uncorrelated Gaussian field in the pixel domain such that

$$\forall k = 1,\dots,N, \quad n[k] \sim \mathcal{N}\left(0, \sigma_k^2\right). \tag{7}$$

Written differently, n is a realization of a multivalued Gaussian variable with mean 0 and covariance matrix $\Sigma = \mathrm{diag}(\sigma_1^2 \cdots \sigma_N^2)$. The case of correlated noise will be discussed later on in this section. From the Gaussian modeling of n, the likelihood function is trivially obtained as follows:

$$\mathcal{L}(y \mid x, \Sigma) \propto \exp\left\{-(y-x)^T \Sigma^{-1}(y-x)\right\}. \tag{8}$$

Following Bayes rule, the posterior distribution for x is the following:

$$P(x \mid y) \propto \mathcal{L}(y \mid x, \Sigma)\mathcal{P}(x) \tag{9}$$

$$\propto \exp-\left[x^T \mathcal{F}^H \mathbf{C}^{-1} \mathcal{F} x + (y-x)^T \Sigma^{-1}(y-x)\right]. \tag{10}$$

The CMB map can be estimated using the maximum *a posteriori* (MAP) estimator with the following analytical solution:

$$\hat{x} = \underset{x}{\mathrm{Argmax}} P(x \mid y) \tag{11}$$

$$= \left(\Sigma^{-1} + \mathcal{F}^H \mathbf{C}^{-1} \mathcal{F}\right)^{-1} \Sigma^{-1} y. \tag{12}$$

The reader would here recognize a classical Wiener filter. Contrary to the classically used Wiener filter in spherical harmonics, the solution of (11) explicitly account for the nonstationarity of the noise. However, the computational evaluation of this Wiener-like solution raises some numerical issues: the noise covariance matrix is diagonal in the pixel domain while the CMB covariance matrix is diagonal in the spherical harmonics domain. The solution filter that appears in (11) is formulated in the pixel domain:

$$\left(\Sigma^{-1} + \mathcal{F}^H \mathbf{C}^{-1} \mathcal{F}\right)^{-1} \Sigma^{-1}. \tag{13}$$

In this expression, the matrix $\mathcal{F}^H \mathbf{C}^{-1} \mathcal{F}$ is the inverse covariance matrix of the CMB in the pixel domain. This matrix is clearly not diagonal. Recalling that N^2 is of the order of $1e13$ for WMAP and $2e15$ for Planck, storing and handling this matrix is way too unrealistic. As an alternative, the most effective approach in this situation would amount to solve the linear system of equations $(\Sigma^{-1} + \mathcal{F}^H \mathbf{C}^{-1} \mathcal{F})\hat{x} = \Sigma^{-1} y$ using a conjugate gradient (CG, [13]) for instance. However, in the next section, we will rather make use of recently introduced proximal algorithms. This optimization framework leads to algorithms which exhibits a lower speed of convergence than CG but with the gain of more flexibility to handle different filtering techniques or better constrain the solution.

2.1.1. Iterative Filtering. As said clearly in the previous paragraph, computing the MAP solution requires inverting a system of equations which turns to be intractable due to the large scale of Planck data. The most straightforward numerical solver is the well-known conjugate gradient. However, while this solver is known to be a fast and accurate numerical method for solving linear systems of equations, it lacks the flexibility of more modern optimization techniques. For this reason we rather opt for the more flexible optimization framework of proximal calculus (see [14] and references therein).

---

(1) **Initialization:** set the number of iterations $I_{\max}$, the step size $\gamma = \min_k \sigma_k^2$, and the starting point x^0.
(2) **Main loop:** apply iteratively the following steps at each iteration k:
    (a) Forward step of the FPS:
      $u = x^k + 2\gamma\Sigma^{-1}(y - x^k)$
    (b) Backward step of the FPS:
      $x^{k+1} = \mathcal{F}^H(I + 2\gamma\mathbf{C}^{-1})^{-1}\mathcal{F}u$
(3) **Stop** when $k = I_{\max}$.

ALGORITHM 1

(a) A Flexible Optimization Framework: Forward-Backward Splitting. The problem in (12) can be written as minimization of the sum of two operators f_1 and f_2:

$$\hat{x} = \underset{x}{\mathrm{Argmin}} f_1(x) + f_2(x) \tag{14}$$

$$= \underset{x}{\mathrm{Argmin}} x^T \mathcal{F}^H \mathbf{C}^{-1} \mathcal{F} x + (y - x)^T \mathbf{\Sigma}^{-1}(y - x). \tag{15}$$

The function $f_1(x) = x^T \mathcal{F}^H \mathbf{C}^{-1} \mathcal{F} x$ is convex and lower semicontinuous and differentiable; the function $f_2(x) = (y - x)^T \Sigma^{-1}(y - x)$ is strictly convex, continuous and differentiable. Such problem as therefore a unique solution (see [14]). If $\nabla f_2(x)$ denotes the gradient of f_2, we will also assumed that this gradient is L-Lipschitz. In the specific setting of interest in this article, $\nabla f_2(x) = -2\Sigma^{-1}(y - x)$ and $L = 2\|\Sigma^{-1}\|_2$. In this context, an effective algorithmic framework is the forward-backward splitting (FPS) technique defined in [14]. In the paradigm of proximal calculus, any convex lower semicontinuous function (see [14]) admits a well-defined proximal operator:

$$\mathrm{prox}_{\gamma f}(z) = \underset{u}{\mathrm{Argmin}} \gamma f(u) + \frac{1}{2}\|z - u\|_{\ell_2}^2. \tag{16}$$

In the FBS framework, the solution to the problem in (14) is computed via an iterative fixed point algorithm such that at iteration k

$$x^{k+1} = \mathrm{prox}_{\gamma f_1}\left(x^k - \gamma\nabla f_2\left(x^k\right)\right) \tag{17}$$

$$= \mathrm{prox}_{\gamma f_1}\left(x^k + 2\gamma\Sigma^{-1}\left(y - x^k\right)\right). \tag{18}$$

As proved in [14], the convergence of the above fixed point algorithm is guaranteed under the condition $\gamma < 2/L$; note that in the setting of CMB map recovery with uncorrelated nonstationary noise, $L = 1/\min_k \sigma_k^2$.

(b) The Proximal Operator of f_1. In the problem of interest in this paper, $f_1(x) = x^T \mathcal{F}^H \mathbf{C}^{-1} \mathcal{F} x$. As this function is continuous and differentiable, it is *a fortiori* lower semicontinuous. Following the definition in (16), the proximal operator of f_1 is defined as follows:

$$\mathrm{prox}_{\gamma f_1}(z) = \underset{u}{\mathrm{Argmin}} \gamma u^T \mathcal{F}^H \mathbf{C}^{-1} \mathcal{F} u + \frac{1}{2}\|z - u\|_{\ell_2}^2. \tag{19}$$

The solution to this problem has a simple closed-form expression:

$$\mathrm{prox}_{\gamma x^T \mathcal{F}^H \mathbf{C}^{-1} \mathcal{F} x} = \left(I + 2\gamma\mathcal{F}^H \mathbf{C}^{-1} \mathcal{F}\right)^{-1} z \tag{20}$$

$$= \mathcal{F}^H\left(I + 2\gamma\mathbf{C}^{-1}\right)^{-1}\mathcal{F}z. \tag{21}$$

The proximal operator of f_1 is a mere Wiener filter in the spherical harmonics space. To sum everything up, the resulting iterative Wiener-like filtering is as we can see in Algorithm 1.

(c) Computational Cost. The computational cost of this iterative Wiener algorithm is particularly low. The first step (2-a) only requires entrywise multiplications and additions of vectors. It is therefore of the order of $\mathcal{O}(N)$. The second step (2-b) is by far the most computationally demanding; it requires entrywise multiplications and additions of vectors of the order of $\mathcal{O}(\ell_{\max})$ (where generally $\ell_{\max} \simeq 3000$) and single forward/backward spherical harmonics transforms. In the numerical experiments of Sections 2.3 and 3.3, the starting point $x^{(0)}$ will be chosen as having entries equal to zero. The algorithm stops when the relative variance—in euclidian norm—between consecutive estimates of the CMB map is lower than a certain threshold: $\|x^{(k+1)} - x^{(k)}\|_{\ell_2}/\|x^{(k)}\|_{\ell_2} \leq 10^{-6}$. The total number of iterations is generally no greater than $I_{\max} = 250$.

(d) Linearity and Bias of the Iterative Wiener Filter. The proposed iterative filtering technique preserves the linearity of the overall CMB map processing. In fact, in every step of the proposed algorithm, each output depends linearly on their inputs. This property is essential as it allows for a simple way to study the propagation of errors via Monte-Carlo simulations; each simulation then has to undergo exactly the same processing step described previously. It is worthwhile to outline that the solution to the proposed iterative Wiener filter is exactly the solution of the Wiener filter described in (11). Therefore, the proposed method exhibits exactly the same bias and variance as the Wiener estimator. Similarly to the global Wiener filter, the bias of this estimator does not vanish and can be shown to depend on the noise statistics and the signal itself. The same holds for the variance as well.

2.2. The Case of Correlated Noise. Assuming a perfect calibration of the data, the nonstationarities of instrumental

noise mainly originate from the nonuniform scanning of the sky. The lower the number of scanning time in one area is, the higher the noise level is and vice versa. Figure 1 shows the RMS map of noise at channel 217 GHz; mathematically this is no more than the square root of the diagonal of Σ. However this is also well known that, contrary to what has been assumed in the previous section, the noise is far from being uncorrelated. There are mainly two origins for the correlation of noise:

(i) The way the sky is scanned by Planck makes the noise correlated along the scanning direction. This means that the noise is correlated along elongated patterns in the pixel domain; for numerical reasons, this can hardly be accounted for in the pixelwise covariance matrix Σ.

(ii) Most modern CMB-dedicated source separation methods, Needlet ILC [4], Generalized ILC [9], SMICA [10] or GMCA [6] rely on a local analysis of the data in the wavelet or needlet domain. Furthermore, the final CMB map is generally obtained as the mixture of observations that do not share the same resolution. This means that the noise that finally contaminates the CMB is, in the pixel domain, correlated.

It then seems natural to model non-stationary-correlated noise: (i) in the wavelet domain to capture its correlation and (ii) locally to capture its nonstationarity.

(1) Wavelets: the Basics. Wavelets are well known to be the tool of choice to analyze nonstationary signals [15]. When considering spherical data, there exist several implementations of wavelets on the sphere. In this paper, we will make use of the isotropic undecimated wavelet transform (IUWT) introduced in [16]. The IUWT can be seen as an extension of the celebrated *à trous* algorithm to spherical data (see the seminal book [15] and references therein).As any wavelet transform, the IUWT is first defined by a scaling function $\phi_{\ell_c}(\theta, \varphi)$ defined by a fixed frequency cut-off ℓ_c. As emphasized in [16], this scaling function is built so that it is azimuth-invariant on the sphere (i.e., it does not de-pend on φ). Classically, each rescaled version of the scaling function $\phi_{\ell_c/2^j}$ is associated to a low-pass filter h_j or equivalently $H_j(\ell)$ in spherical harmonics. Traditionally, the wavelet-based multiscale analysis is performed by computing successive smooth approximations of x. These smooth approximations are obtained by convolving x with dyadically rescaled versions of the scaling function:

$$\forall j = 0, \ldots, J; \quad c_j = h_j \star x, \tag{22}$$

where J is the total number of scales. The so-called wavelet function at scale j, $\psi_{\ell_c/2^j}(\theta, \varphi)$, is defined as the difference of two consecutive rescaled scaling functions:

$$\psi_{\ell_c/2^j}(\theta, \varphi) = \phi_{\ell_c/2^{j-1}}(\theta, \varphi) - \phi_{\ell_c/2^j}(\theta, \varphi). \tag{23}$$

The wavelet function at scale j can be equivalently associated to a high-pass filter g_j or equivalently $G_j(\ell) = 1 - H_j(\ell)$ in

spherical harmonics. The wavelet coefficients at scale j are then computed by convolving the smooth approximation of x at scale j with the j th wavelet function $\psi_{\ell_c/2^j}(\theta, \varphi)$.

$$\forall j = 0, \ldots, J; \quad w_{j+1} = g_j \star c_j. \tag{24}$$

In spherical harmonics, this amounts to a simple multiplication of the signal's spherical harmonic coefficients with the wavelet filter G_j. The decomposition being redundant, several strategies can be designed to reconstruct x from the wavelet coefficients and the coarse resolution c_J. The simplest is the one advocated by the *à trous* algorithm:

$$x = c_J + \sum_{j=1}^{J} w_j. \tag{25}$$

In the following, noise modeling will be performed in each wavelet scale; that is, on each set of coefficients $\{w_j\}_{j=1,\ldots,J}$.

(2) Wavelet-Based Statistical Modeling of Correlated Noise. As emphasized previously, one elegant way to model the instrumental noise is to consider the distribution of its expansion coefficients in the spherical wavelet domain. The basic idea consists in assuming that the wavelet coefficients of noise are approximately decorrelated. This idea takes its roots in the field of multiscale statistical modeling: it has been shown that wavelets exhibit (almost)-decorrelating properties for some classes of nonstationary stochastic processes. To give only one example, this is the case for fractional Brownian motion [17] which has been extensively used to model power-law following stochastic processes. In the following, $w_j^n[k]$ will denote the wavelet coefficient of the noise term n in scale j at pixel k with the convention: $j = 1$ corresponds to the finest scale and $j = J - 1$ to the coarse resolution. From this definition, we will assume that these wavelet coefficients verify:

$$\mathbb{E}\left\{ w_j^n[k] w_j^n[i] \right\} = \sigma_j^n[k]^2 \delta_{k,i}, \tag{26}$$

where $\sigma_j^n[k]^2$ stands for the local variance of the noise at pixel k and $\delta_{k,i}$ for the Kronecker delta function. This entails that the covariance matrix of the noise in wavelet scale j is diagonal and equal to:

$$\Sigma_{n,j} = \text{diag}\left(\sigma_j^n[k]^2 \right). \tag{27}$$

The local standard deviation of the noise, $\Sigma_{n,j}$, has to be estimated from the data beforehand. In the numerical experiments in Section 2.3, these parameters are estimated from a single noise realization. It is commonly assumed that one has access to noise realizations which generally originate from simulations or from the data themselves. Assuming that the local variances $\sigma_j^n[k]^2$ of the noise vary slowly from one pixel to another, it can be approximated at pixel k by the empirical estimate of the variance in a fixed surrounding neighborhood; in the following this neighborhood will be a patch of size $\sqrt{P} \times \sqrt{P}$ centered about the pixel k:

$$\hat{\sigma}_j^n[k]^2 = \frac{1}{P} \sum_{i \in \mathcal{N}[k]} w_j^n[i]^2, \tag{28}$$

where $\mathcal{N}[k]$ stands for the indices of the pixels that compose the patch that surrounds the pixel k. Note that the patch size $\sqrt{P}$ will highly depend on how fast the noise variance varies spatially. This particularly means that $\sqrt{P}$ should be smaller in the finest scale and larger when j increases. Following the natural dyadic scaling of the wavelet transform, the patch size in scale j will scale as follows: $P_j = 4^j P_1$ where P_1 stands for the patch size at the finest scale.

2.3. Numerical Experiments

2.3.1. Experimental Dataset. In the following, numerical experiments will be performed on simulated Planck data described in [3]; to our knowledge, these are the only full-sky simulated data that are publicly available. The input CMB map has been computed by applying a recently introduced source separation technique coined Local-GMCA (L-GMCA, see [6, 18]) to Planck simulated data. These data were already deployed in [3] to evaluate the performances of source separation methods. For more technical details on these data, we refer the interested reader to [3]. To our knowledge this dataset is the only full-sky realistic Planck simulated dataset that is publicly available. It contains contributions as various as synchrotron and free-free emission, anomalous dust, dust, SZ effect and CMB. The instrumental noise on each of these Planck simulated observations was assumed to be uncorrelated but nonstationarity.

The noise that contaminates the raw CMB map at the output of most component separation techniques is nonstationary and correlated. The CMB map estimates being estimated from several observations at different resolutions, the output noise is obviously correlated. Source separation techniques like L-GMCA [6, 18] or Needlet-ILC [4] also analyze the data on small patches in different wavelet or needlet scales. Such local/multiscale process is also at the origin of the correlation of the noise that contaminates the CMB map.

2.4. Experimental Results.

In this section, we particularly focus on the noise reduction aspect of the proposed method. The reference denoising method in the field of astrophysics and more specifically CMB data analysis is the *global* Wiener filter applied in spherical harmonics. In opposition to this classical filtering technique, the proposed iterative filtering approach accounts for the nonstationarity of the noise; thereby, it should particularly provide better solution than global Wiener.

As detailed in the previous section, the modeling of the noise contribution is performed in the wavelet domain: the noise is assumed to be nonstationary but decorrelated at each wavelet scale j. In each wavelet scale, the noise covariance matrix is estimated locally from a single noise realization. The noise covariance per wavelet scale is evaluated locally on patches of size $4^{(j+1)}$ where, by convention, $j = 1$ is defined as the finest wavelet scale. As an illustration, Figure 8 displays the estimated noise RMS map at scale $j = 1$.

The global Wiener solution is computed by filtering the original CMB map y with the linear filter defined in spherical harmonics by (11). The CMB power spectrum C_ℓ is a WMAP7 best-fit estimate; the noise power spectrum C_ℓ^n is estimated via the pseudo-power spectrum of the noise realization.

Iterative Wiener has been applied to the original noisy CMB map y. The single parameter to be tuned is the number of iterations $I_{\max}$; in the following experiments, it has been fixed to 250. Figure 2 shows the noisy map, the LIW-Filtering solution and the noise realization used to estimate the noise local variance in the finest wavelet scale. Figure 3 displays the residual maps of the *global* Wiener and LIW-Filtering solution. As expected the nonstationarity footprint of the noise appears clearly on the residual LIW-Filtering map where the bias of the estimated map is lower.

Figure 4 features the power spectra of the original and filtered maps. We would like to point out that these power spectra have been computed from the raw maps; no masking has been applied. This particularly means that these estimates are likely to be biased by strong foreground from the galactic plane. The first observation is that the global Wiener filter does not fully remove the noise at high ℓ contrary to the iterative filtering technique. This can be explained by the presence of significant nonstationarities at high ℓ which can produce a substantial bias at high frequency. Interestingly, the iterative technique succeed in largely reducing noise at high ℓ thus supporting the significant contribution of the noise nonstationarities in producing a bias on the power spectrum. More technically, it is worth noting that the power spectrum of a stochastic process is properly defined as long as this process is wide sense stationary [19]. Obviously, the nonstationarities of the noise make it depart from this assumption; thereby the power spectrum has no rigorous statistical meaning in this context. Said differently, the power spectrum of the noise C_ℓ^n does not contain all the information that a nonstationary noise does in the pixel domain.

It is well known that the Wiener filter provides a CMB map that has a biased power spectrum. Thereby the Wiener filter—in red in Figure 4—yield a power spectrum that is negatively biased at high ℓ when the noise highly predominates upon the CMB—typically for $\ell > 2500$.

Figure 5 shows the power spectrum of the residual: $x - \hat{x}$. This quantity is particularly well suited to visualize the power discrepancy in frequency between the original CMB map x and its estimate $\hat{x}$. As expected, the *global* Wiener filtering reduces the contribution of noise at high ℓ (see the blue line in Figure 5). Still, the filtered map largely departs from the original map. Interestingly, the proposed iterative filtering technique (in red in Figure 5) largely outperforms the classical Wiener filter from $\ell = 900$ to $\ell = 3000$. More precisely, the residual power is reduced by up to an order of magnitude for $\ell > 2500$.

A more complete measure of discrepancy between the original map x and one of its estimate $\hat{x}$ consists in evaluating the mean squared error (MSE) in the wavelet domain. To that end, Figure 6 displays the MSE of different estimators of the CMB map in 5 wavelet scales in logarithmic scale. The normalized MSE between some CMB map estimate $\hat{x}$ and the "true" CMB map is computed as follows: $\sum_k (\hat{x}[k] - x[k])^2 / \sum_k x[k]^2$. Again, this figure
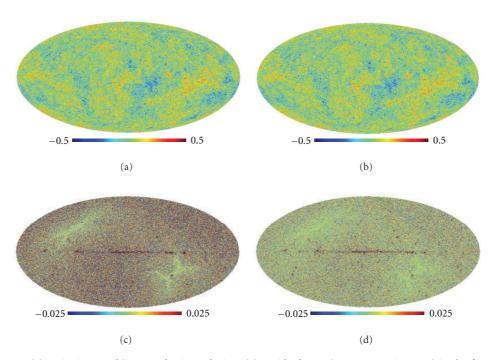
Figure 2: *CMB maps*—(a) Noisy image. (b) LIW-Filtering solution. (c) Residual map (true map-noisy map) in the finest wavelet scale. (d) Residual map (true map-LIW-Filtering solution) in the finest wavelet scale. Units in mK.
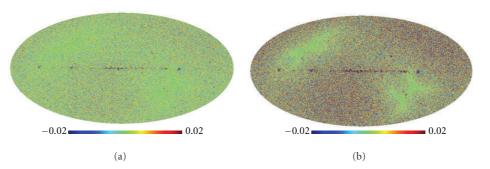

Figure 3: *Removed noise map* (i.e., *noisy data-filtered data*) in the finest wavelet scale—(a) from the *global* Wiener filter. (b) from the LIW-Filtered solution. *These maps correspond to what has been filtered from the noisy map*. The amount of noise removed by LIW-Filtering is more realistic. Units in mK.

shows that the iterative technique which directly account for the nonstationarity of the noise, performs better than the *global* Wiener filter. The impact of the noise tends to vanish for larger scales (*typically* for $\ell < 1500$).

To better illustrate the differences between the *global* Wiener filter and its iterative counterparts, Figure 7 displays the MSE of the different estimators of the CMB map which has been computed in the wavelet domain per bands of latitude of width $10°$. By convention, the galactic plane is centered around the latitude $0°$. Figure 7(a) displays the MSE which has been evaluated from the finest wavelet scale $j = 1$; this corresponds to an analysis window that ranges from $\ell = 1500$ to $\ell = 3000$. Again, taken globally, one can point out the clear improvement between the *global* Wiener filter and its iterative version. As expected the Wiener filter is optimal in the sense of the MSE. More interestingly, the MSE of the

two filtering techniques seems to be rather constant for high latitudes and sharply increases in the range $[-15°, 15°]$. As already emphasized, the impact of any of these linear filtering techniques is the more significant when the signal to noise ratio (SNR) decrease. This is obviously the case at smaller scales. At larger scales, the SNR increases and the impact of the filtering is much less significant. This is exactly what can be observed on the plot on the right of Figure 7(b): the MSE of any of the filtered CMB maps at scale $j = 2$, which grossly correspond to an analysis window ranging from $\ell = 750$ to $\ell = 1500$, are approximately the same.

As a brief conclusion for these experimental results, it appears clearly that the nonstationarity of the noise has an impact on the CMB estimate. The most classical noise reducing technique in the field, that is, the *global* Wiener filter, only relies on the power spectra of the CMB C_ℓ

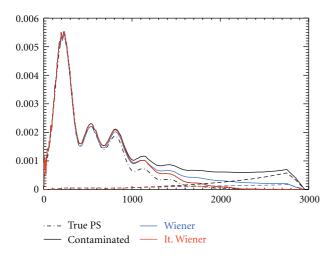FIGURE 4: *Power spectra.* In abscissa: spherical harmonics coefficient. In ordinate: CMB power in mK^2. Dashed line show the power spectra of the noise for the different maps.
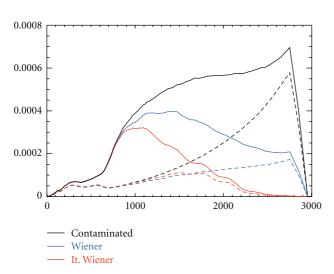

FIGURE 5: *Residual power spectra.* In abscissa: spherical harmonics coefficient. In ordinate: CMB power in mK^2. Dashed line show the power spectra of the noise for the different maps.

and the noise C_ℓ^n. Obviously, the behavior of the noise power spectrum contains almost no information about its spatial behavior. These experiments have shown that the noise nonstationary has a clear impact on the CMB. Further accounting for the spatial variations of the noise variance helps improving the noise reduction process by a dramatic amount.

3. From Noise Reduction to Contamination Reduction

As stated in the introduction in Section 1, one of the purposes of this paper is to tackle the problem of post-separation contamination reduction. In this setting, contamination generally means instrumental noise and residuals of galactic foregrounds. In the case of Planck, CMB maps are generally
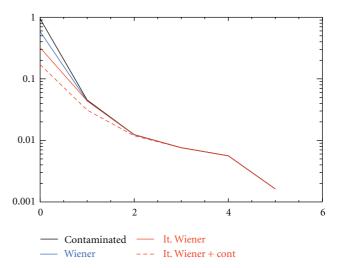

FIGURE 6: *MSE per wavelet scale.* In abscissa: wavelet scale—0 corresponds to $j = 1$. In ordinate: normalized mean squared error.

obtained by applying sources separation methods. Most of these methods provide a solution that can be formulated as a linear combination of the original Planck multichannel observations. Therefore, it is relevant to assume that these contaminants essentially contribute additively to the CMB map:

$$y = x + f + n, \tag{29}$$

where f stands for the foreground contamination and n for the instrumental noise. Furthermore, it is also natural to assume that these three components x, f and n are mutually uncorrelated. The filtering technique introduced in the previous sections mainly relies on the fact that the contaminants are characterized by their covariance matrix Σ. The questions that we aim at tackling in this paragraph is: *How far do we know Σ and how can it be estimated from the data ?*

3.1. Foreground Residuals Modeling. Unlike noise, the foreground residual are generally non-Gaussian, with the exception of the CIB (see [20]). We further restrict the modeling of residuals to their second order statistics thus opting for a Gaussian approximation of their contribution. The contribution of foreground residuals are obviously not known in advance; this means that their covariance matrix has to be estimated from the data directly. From the uncorrelation of the different components, the additive contamination model in (29) then reads at the level of second statistics:

$$\Sigma = \Sigma_x + \Sigma_f + \Sigma_n. \tag{30}$$

In this equation, the correlation of foreground pixels implies that Σ_f is non-diagonal. Similarly to the noise, second order dependencies of foreground pixels are way too complex to be modeled in the pixel domain.

So as to capture the correlation of foreground pixels, a natural and simple strategy boils down to adopting the wavelet-based statistical modeling used to model correlated
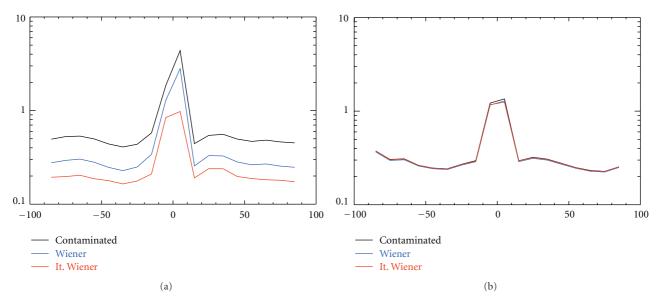
(a)

(b)

FIGURE 7: *MSE per latitude at wavelet scale* (a) $j = 1$ *and* (b) $j = 2$. In abscissa: latitude. The galactic plane corresponds to $0°$. In ordinate: normalized mean squared error.
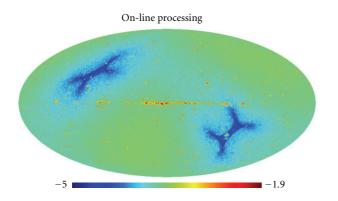

FIGURE 8: *Estimated noise covariance map at wavelet scale* $j = 1$: this covariance map has been estimated from a single realization of the noise that contaminates the CMB estimated by GMCA on simulated Planck data—see [3] for a description of these data. Units in mK^2.

noise in the previous Section 2.2.(2). To this end, $w_j^f[k]$ will denote the wavelet coefficient of f in scale j at pixel k with the convention: $j = 1$ corresponds to the finest scale and $j = J - 1$ to the coarse resolution. From this definition, we will assume that these wavelet coefficients verify:

$$\mathbb{E}\left\{w_j^f[k]w_j^f[i]\right\} = \sigma_j^f[k]^2 \delta_{k,i}, \tag{31}$$

where $\sigma_j^f[k]^2$ stands for the local variance of the foreground at pixel k. The covariance matrix of the foreground residual in wavelet scale j is diagonal and equal to:

$$\Sigma_{f,j} = \text{diag}\left(\sigma_j^f[k]^2\right). \tag{32}$$

In practical situations, $\Sigma_{f,j}$ has to be estimated from the data y themselves. Similarly to the modeling of noise, the local noise variance of the foreground will be approximated

by the empirical estimate of the variance in a surrounding neighborhood of size $\sqrt{P} \times \sqrt{P}$ centered about the pixel k:

$$\hat{\sigma}_j^f[k]^2 = \frac{1}{P} \sum_{i \in \mathcal{N}[k]} w_j^f[i]^2. \tag{33}$$

3.2. Contamination Filtering. In the rest of this paper, we will generally use the term contamination for the contribution of both the foreground residuals and the instrumental noise. In the following, as suggested previously, we also opt for exactly the same model to capture the pixel dependencies of the instrumental noise. Therefore, we will now define locally at pixel k and scale j the variance of contamination (noise and foreground) as $\hat{\sigma}_j^c[k]^2 = \hat{\sigma}_j^f[k]^2 + \hat{\sigma}_j^n[k]^2$. The joint contribution of noise and foreground residuals will be evaluated jointly from the data.

From the estimation of the local variance $\hat{\sigma}_j^y[k]^2$; one has to recall that the contaminants' variance is not directly accessible but is equal to:

$$\hat{\sigma}_j^c[k]^2 = \left[\hat{\sigma}_j^y[k]^2 - \hat{\sigma}_j^2\right]_+, \tag{34}$$

where $\hat{\sigma}_j^2$ stands for the variance of the CMB and $[x]_+ = \max(0, x)$. Let use note that, from the stationarity of the CMB map, $\hat{\sigma}_j^2$ does not explicitly depend on k; this holds as long as the patch size at scale j is large enough compared to the CMB fluctuations (i.e., correlation length) in the jth wavelet scale. The value of the CMB variance at scale jth is directly related to the CMB power within this scale; its value can be derived from the CMB power spectrum C_ℓ and the wavelet analysis filters. To that end, let us denote by $\psi_{j,\ell}$ the spherical harmonics filter from which the wavelet scale j is defined. As emphasized in Section 1 the spherical harmonics coefficients of the CMB map should be mutually uncorrelated. This

particularly entails that the power of the CMB in the jth wavelet scale is exactly:

$$\sigma_j^2 = \frac{1}{4\pi}\sum_\ell (2\ell + 1)\psi_{j,\ell}^2 C_\ell. \tag{35}$$

It follows that the variance of the contamination at pixel k is approximately equal to:

$$\hat{\sigma}_j^c[k]^2 = \left[\hat{\sigma}_j^y[k]^2 - \frac{1}{4\pi}\sum_\ell (2\ell + 1)\psi_{j,\ell}^2 C_\ell\right]_+. \tag{36}$$

Thereby, from the *a priori* knowledge of the CMB power spectrum C_ℓ, an approximate contribution of the residual can be evaluated at each pixel and each wavelet scale. The reader has to keep in mind that the measurement of this contribution is only approximate and has to be considered under the light of the following assumptions: (1) the contribution of the residual is modeled as a nonstationary Gaussian field where the covariance matrix is diagonal in each wavelet scales and (2) its variance being evaluated on small patches of size $\sqrt{P} \times \sqrt{P}$, it is assumed to be approximately stationary in a given analysis patches. However, even with these approximations, our model offers a much more realistic modeling of the complexity of the data than the traditional Wiener method. Improvements of this modeling will be discussed in Section 4.1.

Let $\mathbf{\Sigma}_{f,j} = \mathrm{diag}(\hat{\sigma}_j^c[k]^2)$. Then, in the algorithm introduced in Section 2, the noise matrix $\mathbf{\Sigma}$ is substituted with the following covariance matrix:

$$\mathbf{\Sigma} = \mathbf{W}^\mathbf{T}\begin{bmatrix} \mathbf{\Sigma}_{f,1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{\Sigma}_{f,J} \end{bmatrix}\mathbf{W}, \tag{37}$$

where $\mathbf{W}$ stands for the isotropic wavelet transform and $\mathbf{W}^\mathbf{T}$ its adjoint. Along with the iterative technique introduced in Section 2, the resulting wavelet-based filtering will be coined LIW-Filtering for Linear Iterative Wiener Filtering in the following.

3.3. Contamination Reduction on Planck Simulated Data.
The dataset used in this section is exactly the same that we used to study noise reduction. It is described in Section 2.3.

As emphasized in Section 3, the wavelet-based noise/contamination variance modeling used so far makes it possible to account for the contribution of the foreground contamination in the filtering process. Estimating the contribution of the foreground contamination, as detailed in Section 3, makes the reduction of contamination possible. In the following experiments, the contamination, which includes the contribution of the noise and an approximate contribution of the foreground residuals, has been modeled in the wavelet domain [16] using 6 scales. Within each wavelet scale, with the exception of the coarsest scale to which no filtering is applied, the contribution of the contamination is assumed to be Gaussian and locally stationary on patches
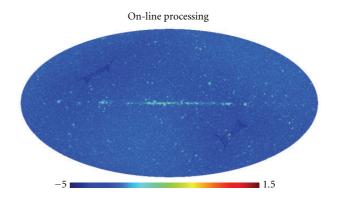
On-line processing

FIGURE 9: *Estimated contamination covariance map at wavelet scale $j = 1$: this covariance map has been estimated from the CMB estimated by GMCA. Units in mK^2.*
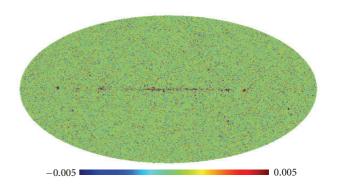

FIGURE 10: *CMB and residual maps* Difference map between the noise-only LIW-Filtering solution and the contamination-aware LIW-Filtering solution in the finest wavelet scale. Units in mK.

of size $\{4^{j+1}\}_{j=1,\dots,5}$. Figure 9 features the estimated variance map in logarithmic scale. As a remark, if the morphology of this contaminant variance map resembles at first glance the noise-only variance map, the dynamic range of the former is clearly larger by 3 orders of magnitudes. These differences are mostly significant on the galactic plane and on local features that are likely point sources or residuals of SZ clusters.

The proposed Wiener-based iterative method has been applied on the aforementioned Planck simulated data. Figure 10 shows the difference map between the noise-only LIW-Filtering solution and the proposed contamination-aware LIW-Filtering solution in the finest wavelet scale. It appears clearly that the difference between the map are mainly concentrated on the galactic plane and on likely point sources at high latitude.

Figure 11 displays the power spectra of the original and filtered CMB maps along with the theoretical power spectrum of the simulated map in black dash-dotted line. At first glance, accounting for the contribution of the foreground residuals, even approximately, yields improvements from $\ell = 750$. Perhaps more interesting, Figure 12 shows the power spectrum of the residual maps $x - \hat{x}$: the contaminant-aware filtering technique clearly outperforms the previous methods
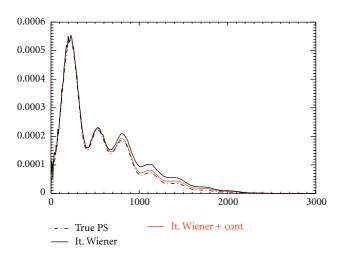
FIGURE 11: *Power spectra*. In abscissa: spherical harmonics coefficient. In ordinate: CMB power in mK². Dashed line show the power spectra of the noise for the different maps.
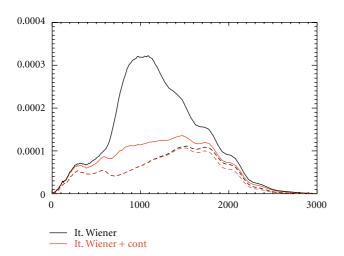

FIGURE 12: *Residual power spectra*. In abscissa: spherical harmonics coefficient. In ordinate: CMB power in mK². Dashed line show the power spectra of the noise for the different maps.

from $\ell = 250$ to $\ell = 3000$ with a dramatic improvement in the range $[750, 1500]$.

Figure 13 displays the wavelet-based MSE we described in the previous paragraph. The results of the contaminant-aware filtering technique have been further added to the plot in Figure 6. Let us recall that the noise reducing techniques described in the previous section mainly help reducing the noise in the noise-dominant regime. This particularly explained why the MSE remains almost unaltered in wavelet scales $j > 1$. Interestingly, accounting for an approximate contribution of the foregrounds helps reducing their impact in larger scales, typically for $j < 3$.

Figure 14 presents the MSE computed at different latitudes in the finest wavelet scale $j = 1$— plot on Figure 14(a)—and $j = 2$ —plot on Figure 14(b). Accounting for the contaminant contribution improves the MSE of the filtered map at all latitudes. Interestingly, the MSE is
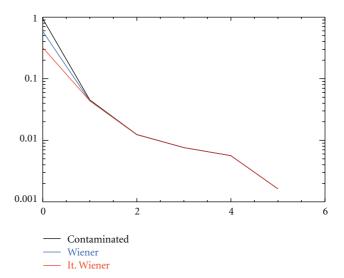

FIGURE 13: *Normalized MSE per wavelet scale*. In abscissa: wavelet scale—0 corresponds to $j = 1$. In ordinate: normalized mean squared error.

decreased by a large amount—a factor of approximately 5— on the galactic (latitude 0°). The same holds in a slighter amount in wavelet scale $j = 2$.

Assuming that the Planck instrumental noise is Gaussian is widely considered as a reasonable assumption. This is obviously not the case for foreground residuals the presence of which is likely to largely distort the search for non-Gaussian features in the CMB. Thereby, reducing the amount of contaminant should help preventing non-Gaussianity test from the non-Gaussian impact of foreground residuals. In the field of CMB non-Gaussianity evaluation, a classical technique boils down to measuring higher-order statistics in the wavelet domain [21]. One of the statistics of choice is the statistics of order 4, denoted κ_4, a.k.a. the kurtosis. The kurtosis of the wavelet coefficients of the estimated CMB map has been evaluated in 5 scales; non-Gaussianity test results are shown in Figure 15. The first observation is that the proposed method helps reducing non-Gaussian contamination even when only the noise is modeled as a spurious component. This suggests that the nonstationary behavior of the noise can generate undesired non-Gaussian signatures. Thereby, accounting for the nonstationary behavior of the noise in the noise reduction process decreases the kurtosis of the filtered map by an order of magnitude in the finest wavelet scale.

The red dashed line in Figure 15 shows a clear reduction of the kurtosis of the filtered map when the contribution of the foreground residuals is modeled. More precisely, the value of κ_4 is grossly reduced by 4 orders of magnitude for $j = 1$, 2 orders of magnitude for $j = 2$ and 1 order of magnitude for $j = 3$. While noise-only filtering yields map that departs from the Gaussian assumption for $J < 2$, the contamination-aware filtered map is now compatible with the Gaussian assumption in all wavelet scales with $\kappa_4 < 1$.

Figure 16 presents more detailed non-Gaussianity tests; the three plots that this figure displays show the evolution of κ_4, in the three finest wavelet scales $j = 1, \ldots, 3$, which
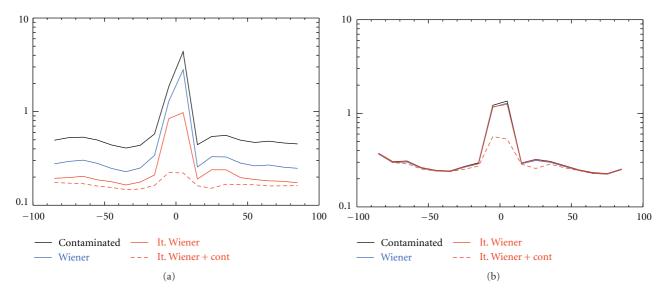
(a)

(b)

FIGURE 14: (a) *Normalized MSE per latitude at wavelet scale $j = 1$ and* (b) $j = 2$. In abscissa: latitude. The galactic plane corresponds to $0°$. In ordinate: normalized mean squared error.
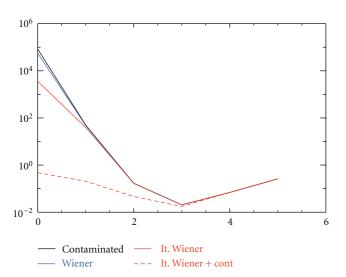

FIGURE 15: *Kurtosis κ_4 per wavelet scale.* In abscissa: wavelet scale—0 corresponds to $j = 1$. In ordinate: value of the kurtosis κ_4.

has been evaluated in bands of latitude of width $10°$. The galactic plane is centered about the point $0°$. The non-Gaussian contamination mainly originates from the presence of galactic foreground residuals which are mainly concentrated on the galactic plane. This explains the very large value of κ_4 in the range of latitude $[-15°, 15°]$. The second well-known source of non-Gaussian signatures at spatial high frequency are the radio and infrared point sources. These foregrounds are roughly uniformly scattered all over the sky. Point sources are likely to be the predominant source of non-Gaussian signatures at high latitude thus explaining the relatively large value of κ_4 in this range of latitudes. As shown in Figure 16(a), further filtering the contamination (see the red dashed line) yields a dramatic reduction of κ_4 up to 5 orders of magnitude on the galactic plane. The filtered CMB map is likely to be compatible with the Gaussian assumption

at every latitude, even on the galactic plane. In the second wavelet scale $j = 2$, the same reduction of non-Gaussian signatures is also visible by a lesser extent, mainly on the galactic plane in the range $[-15°, 15°]$. In the third wavelet scale, the proposed filtering is likely to reduce the non-Gaussian signatures in the galactic plane; in this range of scales (i.e., the third wavelet scale roughly corresponds to an analysis window that spans the range $[375, 750]$), the CMB is largely predominant which entails that: (i) the input map is already quite compatible with the Gaussian assumption and (ii) the impact of the filtering may be less visible when non-Gaussianity is measured.

4. Discussion, Prospects, and Conclusion

4.1. Discussion and Prospects

4.1.1. Filtering the Maps, for What Use? Important cosmological tests and evaluations are performed on estimated CMB maps; this is particularly the case for CMB lensing reconstruction [7] and non-Gaussian signature detection to only name two. These tests are particularly sensitive to spurious components that contaminate the estimated CMB map, whether it is noise or foreground residuals. This is therefore crucial to estimate a CMB map that is the less contaminated possible. The second point we would like to emphasize is that the aforementioned cosmological tests or evaluations are generally performed on the full-sky data; prior to any post-processing the estimated map is masked to get rid of the impact of galactic foregrounds and point sources. This has two major consequences: (i) it limits the size of the samples to which these tests/evaluations are performed thus increasing the statistical variance of these tests and (ii) any of these tests has to deal with this mask (e.g., via an inpainting procedure in [7]) thus making the analysis of the CMB map generally more complex.
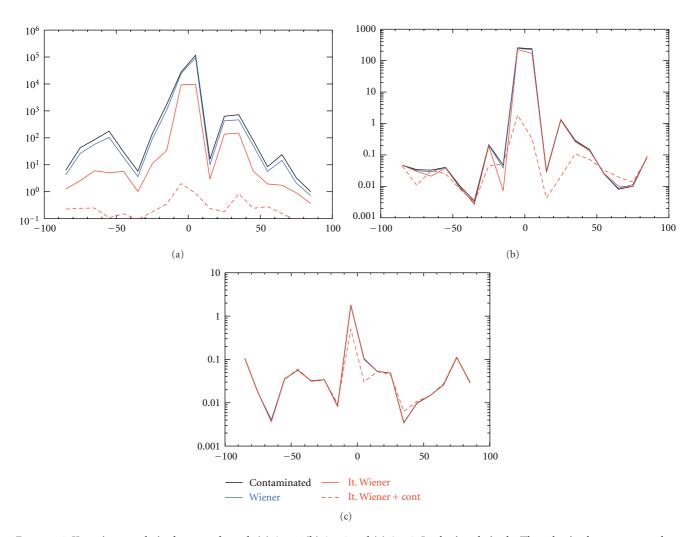
FIGURE 16: *Kurtosis κ_4 per latitude at wavelet scale* (a) $j = 1$, (b) $j = 2$ and (c) $j = 3$. In abscissa: latitude. The galactic plane corresponds to $0°$. In ordinate: value of the kurtosis κ_4.

As shown in the previous experiments, the proposed filtering clearly limits the impact of noise by taking into account its nonstationary behavior. Furthermore, results are presented that clearly show that modeling the contribution of foreground residuals make contamination reduction effective even on the galactic plane. This has two major consequences: (i) it makes it possible the use of a much smaller mask prior to any analysis of the estimated CMB map and (ii) it helps reducing the non-Gaussian features that originate from the presence of foreground residuals.

4.1.2. Beyond the Proposed Methods. Whether noise or foreground residual is at stake, the central assumption that is at the heart of the modeling of contamination is the decorrelation hypothesis we made in Section 3. It is assumed that the contaminants have potentially nonstationary but decorrelated samples in each wavelet scales. The basic idea that supports this assumption is the well-known decorrelating property of wavelets bases. If this assumption almost holds for rather theoretical stochastic processes (e.g., fractional Brownian motion [17] to only name one), this is only

roughly approximate for more general signals such as galactic foreground or complex correlated noise. The basic idea behind this assumption is that the multiscale analysis basis is able to capture the *correlation morphology* of the component to be modeled. In surveys involving raster scanning such as WMAP or Planck, it is very likely that instrumental noise will be highly correlated along the scanning direction. This means that the correlation of the Planck instrumental noise is better modeled as a decorrelated stochastic process in a multiscale signal representation that best represent elongated structures such as the Curvelet tight frame [16]. We can extrapolate the same argument to the modeling of foreground residuals.

The contamination modeling used so far in this paper makes profit of the decorrelation assumption; it particularly helps simplifying the filtering process by only requiring the handling of diagonal matrices (i.e., root mean squared maps). Departing from the decorrelation assumption will largely increase the complexity of the proposed filtering techniques. However, a straightforward way of extending the proposed methods is to choose multiscale signal representations which are better adapted to the morphology or structure of the signal to be modeled.

4.1.3. Potential Impact of Contamination Reduction on CMB Non-Gaussianities. It is important to wonder whether the contamination reduction filtering introduced in Section 3 might affect or reduce potential CMB-related non-Gaussianities. One important point is that the way the local variance of the contamination is estimated in Section 3 leads, in practice, to an estimation of foreground contaminations that largely exceeds the average CMB power in each wavelet scale. Therefore, the filtering should have a much lesser impact on non-Gaussianities whose magnitude is of the order or lower than the magnitude of the CMB. It is worth pointing out that the rule in (36) to estimate the local con-tamination variance can also be chosen differently to only account for non-Gaussianities that are guaranteed to largely exceed the magnitude of the CMB thus avoiding for sure any effect on potential CMB-related non-Gaussianities. Future work will focus on evaluating the potential impact of the proposed filtering technique on CMB lensing reconstruction and non-Gaussianity detection.

4.2. Conclusion. Cosmological microwave background maps estimated from full-sky surveys such as WMAP or more recently Planck generally suffers from various sources of contaminations: (i) instrumental noise is generally nonstationary which may generate non-Gaussian signatures and (ii) foreground residuals generally remain even after the application of state-of-the-art source separation methods. In this context, the most classical denoising technique, aka. *global* Wiener filter; despite its simplicity, it is not able to account for the nonstationarity of the instrumental noise. To that end, we introduce a novel noise reduction technique coined LIW-Filtering for Linear Iterative Wavelet Filtering which combines the linearity of the *global* Wiener filter while accounting for the potential nonstationarity of the noise. In this framework, the noise is modeled as a nonstationary but decorrelated process in the wavelet domain. The denoising problem then takes the very simple form of a quadratically regularized least square where the noise covariance matrix is diagonal in the wavelet domain and the signal covariance matrix is also diagonal in the spherical harmonic domain (i.e., CMB power spectrum). The solution is computed using recently introduced proximal algorithms. When noise reduction is at stake, we showed that the proposed iterative technique succeeds in reducing the mean squared error (MSE) of the filtered solution with respect to the *global* Wiener filter classically applied in the field. Moreover, the modeling/estimation framework introduced in this paper makes the reduction of foreground contamination possible. Similar to the nonstationary instrumental noise, foreground contamination is modeled as a nonstationary but decorrelated process in the wavelet domain. Numerical experiments show that: (i) the MSE of the filtered map is improved; specifically on the galactic plane, and (ii) very interestingly, non-Gaussian signatures are also dramatically reduced. This particularly provides arguments supporting the application of the proposed filtering technique as post-processing step to be applied to the CMB with the crucial aim of reducing noise and perhaps more importantly foreground contamination.

It is also important to point out that, similarly to the *global* Wiener filter, LIW-Filtering is—at this acronym indicates—a linear filtering technique. This means that LIW-Filtering is also relevant when studying errors and their propagation on the estimated CMB map via Monte-Carlo simulations are unavoidable.

Future work will focus on refining the modeling of noise and foreground residuals. Without losing the simplicity of this approach, we will more specifically focus on studying more adapted signal representation to better model the noise/foreground contamination spatial behavior.

The developed IDL code will be released with the next version of ISAP (Interactive Sparse astronomical data Analysis Packages) via the following web site: http://jstarck.free .fr/isap.html.
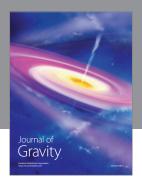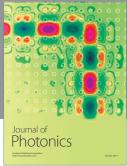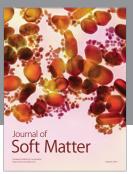
Acknowledgments

References

[1] C. L. Bennett, M. Halpern, G. Hinshaw et al., "First-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: preliminary maps and basic results," *Astrophysical Journal, Supplement Series*, vol. 148, no. 1, pp. 1–27, 2003.

[2] F. R. Bouchet and R. Gispert, "Foregrounds and CMB experiments: I. Semi-analytical estimates of contamination," *New Astronomy*, vol. 4, no. 6, pp. 443–479, 1999.

[3] S. M. Leach, J. F. Cardoso, C. Baccigalupi et al., "Component separation methods for the PLANCK mission," *Astronomy and Astrophysics*, vol. 491, no. 2, pp. 597–615, 2008.

[4] J. Delabrouille, J. F. Cardoso, M. Le Jeune, M. Betoule, G. Fay, and F. Guilloux, "A full sky, low foreground, high resolution CMB map from WMAP," *Astronomy and Astrophysics*, vol. 493, no. 3, pp. 835–837, 2009.

[5] J. Bobin, J. L. Starck, J. Fadili, and Y. Moudden, "Sparsity and morphological diversity in blind source separation," *IEEE Transactions on Image Processing*, vol. 16, no. 11, pp. 2662–2674, 2007.

[6] J. Bobin, F. Sureau, and J.-L. Starck, "Source separation in cosmology, from global to local models," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '11)*, pp. 1297–1300, 2011.

[7] L. Perotto, J. Bobin, S. Plaszczynski, J.-L. Starck, and A. Lavabre, "Reconstruction of the cosmic microwave background lensing for *Planck*," *Astronomy & Astrophysics*, vol. 519, article A4, 14 pages, 2010.

[8] D. Hanson, G. Rocha, and K. Górski, "Lensing reconstruction from Planck sky maps: inhomogeneous noise," *Monthly Notices of the Royal Astronomical Society*, vol. 400, no. 4, pp. 2169–2173, 2009.

[9] M. Remazeilles, J. Delabrouille, and J.-F. Cardoso, "Foreground component separation with generalized Internal Linear Combination," *Monthly Notices of the Royal Astronomical Society*, vol. 418, no. 1, pp. 467–476, 2011.

[10] J. Delabrouille, J. F. Cardoso, and G. Patanchon, "Multidetector multicomponent spectral matching and applications for cosmic microwave background data analysis," *Monthly Notices*
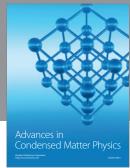
of the Royal Astronomical Society, vol. 346, no. 4, pp. 1089–1102, 2003.

[11] M. Górski, E. Hivon, A. J. Banday et al., "HEALPix: a framework for high-resolution discretization and fast analysis of data distributed on the sphere," *Astrophysical Journal*, vol. 622, no. 2 I, pp. 759–771, 2005.

[12] E. Komatsu, J. Dunkley, M. R. Nolta et al., "Five-year wilkinson microwave anisotropy probe observations: cosmological interpretation," *Astrophysical Journal, Supplement Series*, vol. 180, no. 2, pp. 330–376, 2009.

[13] G. H. Golub and C. F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, Md, USA, 3rd edition, 1996.

[14] P. L. Combettes and V. R. Wajs, "Signal recovery by proximal forward-backward splitting," *Multiscale Modeling and Simulation*, vol. 4, no. 4, pp. 1168–1200, 2005.

[15] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, Calif, USA, 1998.

[16] J. L. Starck, Y. Moudden, P. Abrial, and M. Nguyen, "Wavelets, ridgelets and curvelets on the sphere," *Astronomy and Astrophysics*, vol. 446, no. 3, pp. 1191–1204, 2006.

[17] P. Flandrin, "Wavelet analysis and synthesis of fractional Brownian motion," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 910–917, 1992.

[18] J. Bobin, Y. Moudden, J. L. Starck, J. Fadili, and N. Aghanim, "SZ and CMB reconstruction using generalized morphological component analysis," *Statistical Methodology*, vol. 5, no. 4, pp. 307–317, 2008.

[19] T. Chonavel, *Statistical Signal Processing: Modelling and Estimation*, Springer, New York, NY, USA, 2002.

[20] F. R. Bouchet, R. Gispert, and J. Puget, "Unveiling the cosmic infrared background," in *Proceedings of the IP Conference*, E. Dwek, Ed., pp. 255–258, 1995.

[21] J. L. Starck, N. Aghanim, and O. Forni, "Detection and discrimination of cosmological non-gaussian signatures by multi-scale methods," *Astronomy and Astrophysics*, vol. 416, no. 1, pp. 9–17, 2004.