A WAVELET–GALERKIN ALGORITHM OF THE E/B DECOMPOSITION OF COSMIC MICROWAVE BACKGROUND POLARIZATION MAPS

LIANG CAO1,2 AND LI-ZHI FANG3

1 Key Laboratory for Research in Galaxies and Cosmology, Shanghai Astronomical Observatory, Chinese Academy of Sciences, 80 Nandan Road, Shanghai 200030, China
2 Graduate University of the Chinese Academy of Science, 19A Yuquan Road, Beijing 100049, China
3 Department of Physics, University of Arizona, Tucson, AZ 85721, USA

Received 2009 January 20; accepted 2009 October 26; published 2009 November 13

The Astrophysical Journal, 706:1545–1555, 2009 December 1
doi:10.1088/0004-637X/706/2/1545
© 2009. The American Astronomical Society. All rights reserved. Printed in the U.S.A.

1. INTRODUCTION

The scalar component of primordial perturbations of the universe can be detected by the maps of temperature fluctuations of cosmic microwave background radiation (CMBR), while the tensor component of the primordial perturbations has to be probed by the maps of the Stokes parameters Q and U of the linear polarization of the CMBR. A tensor field generally contains electric-like E mode and magnetic-like B mode. In the linear regime, vortical mode of primordial perturbations do not grow during the clustering of density field, and therefore, the perturbed field initially has to be curl free. That is, the primordial perturbations can only yield the E mode, but not B mode of the CMBR polarization field. On the other hand, B-mode perturbations can be produced by gravitational waves. Therefore, extracting the B-mode information from CMBR polarization maps is crucial to verify the existence of gravitational wave background produced at the inflationary epoch. Moreover, gravitational lensing of clusters and hot electron scattering of reionization would be able to yield both E and B modes. To study these problems, a sharp decomposition of E and B modes from Q and U maps is required.

If both Q and U maps are available over the whole sky, one can find the whole sky maps of E and B modes with spherical harmonic decomposition, because the relation between the maps of (Q, U) and (E, B) in the space spanned by bases of spin two harmonics is local (Kamionkowski et al. 1997; Zaldarriaga & Seljak 1997). However, the observed maps cannot be global; it is always limited by the contamination of our galaxy and other foreground sources. The relation between the maps of (Q, U) and (E, B) in physical space contains the Laplace operator, and therefore, it is non-local. The E/B decomposition with the spatially limited maps of Q and U will not be unique if we lack information on the polarization and its derivative on the boundary of the maps.

For noiseless samples, the problem of uniqueness would be solved by constructing orthogonal modes with window functions to fit the requirements of boundary conditions (Lewis et al. 2002; Bunn et al. 2003; Smith 2006; Smith & Zaldarriaga 2007). It is, however, similar to the domain (or windowed) Fourier analysis. The result will not be useful to study the structures in physical space (e.g., Chiueh & Ma 2002).

The other challenge caused by the derivative operator is because the Q and U maps are discrete. Mathematically, the derivative operators $\partial_x^i$ or $\partial_x^2$ are continuous linear operators mapping functions defined in Hilbert space, while the observed samples $Q$ and $U$ actually are defined in space spanned by base $v_i$, $i \in I$, which is a set of finite indices. This difference leads to large numerical errors when the discrete maps are noisy.

The last, but not the least, problem is from the smallness of B modes. On the scale of one degree order, the power of B mode caused by gravitational waves at inflationary epoch is less than that of E modes by a factor of at least $10^2$. As the maps of Q and U are random fields, the variance of the random fields will lead to large numerical errors when the discrete maps are noisy. Consequently, E and B modes would be contaminated from each other. Therefore, it is difficult to recover the power of B mode if the powers of E modes are much larger than that of B modes.

In this paper, we develop an algorithm of the E/B decomposition based on the discrete wavelet transform (DWT) analysis, which is a compromise between the decompositions in physical space and scale space. The DWT analysis of the CMBR temperature fluctuation maps has attracted much attention in the last decade (Pando et al. 1998; Sanz et al. 1999; Mukherjee et al. 2000). Besides these points, we especially take the advantage of the so-called wavelet–Galerkin discretization (e.g., Louis et al. 2000).

1545
the effectiveness of the wavelet–Galerkin discretization. Finally, the variance of Gaussian random field is analyzed, especially the tests for samples of Gaussian random field. The effect of the method of the tests the DWT algorithm with samples having known spatial coordinates (which can be approximated as a plane described by Cartesian coordinates ($x$, $y$)). We will study the conditions, under which the information of small $B$ mode can approximately be extracted from noisy maps of $Q$ and $U$.

The paper is organized as follows. Section 2 presents the method of the $E/B$ separation in the DWT space. Section 3 tests the DWT algorithm with samples having known spatial structures. We show that the method effectively suppresses the uncertainties from boundary effect and noise. It is also effective to identify spatial structures of $E$ and $B$ fields. Section 4 presents the tests for samples of Gaussian random field. The effect of the variance of Gaussian random field is analyzed, especially the problem of the mixing of $E$ and $B$ modes. Section 5 addresses the effectiveness of the wavelet–Galerkin discretization. Finally, conclusions are given in Section 6. The DWT representations of derivative operators are given in the Appendix.

2. METHOD

2.1. $E/B$ Separation in DWT Representation

Let us consider polarization samples in a patch of sky, which can be approximated as a plane described by Cartesian coordinates $(x, y)$. In this case, the fields of $E(x, y)$ and $B(x, y)$ are related to the maps of Stokes parameters $Q(x, y)$ and $U(x, y)$ by (Seljak 1997)

$$\nabla^2 E(x, y) = (\partial_x^2 - \partial_y^2)Q(x, y) + 2\partial_x \partial_y U(x, y),$$  
(1)

$$\nabla^2 B(x, y) = 2\partial_x \partial_y Q(x, y) - (\partial_x^2 - \partial_y^2)U(x, y),$$  
(2)

where $\nabla^2$ is two-dimensional Laplace $\partial_x^2 + \partial_y^2$.

We first take a wavelet–Galerkin discretization of Equations (1) and (2) to rewrite these equations in the DWT space. We can assume that the patch is an $L \times L$ square. The size of each pixel is $L/2^J$. Being an integral, one can project the maps into the DWT space by

$$\epsilon^{\Omega}_{i, l,z} = \int Q(x, y)\phi_{J, l,z}(x)\phi_{J, l,z}(y)dx dy,$$  
(3)

$$\epsilon^{U}_{i, l,z} = \int U(x, y)\phi_{J, l,z}(x)\phi_{J, l,z}(y)dx dy,$$  
(4)

where $\phi_{J, l}(x)$ is orthogonal scaling function on scale $J$ (e.g., Fang & Thews 1998). It is non-zero mainly in the cell in $x$-space from $L/2^J$ to $(l + 1)L/2^J$. The index $l$ runs from 0 to $2^J - 1$. It spans the spatial range from 0 to $L$. The variables $\epsilon^{\Omega}_{i, l,z}$ and $\epsilon^{U}_{i, l,z}$ actually are the maps of $Q$ and $U$ on scale $J$. Since the observed maps of $Q$ and $U$ are always pixelized, the projection of Equations (3) and (4) does not lose information if the size $L/2^J$ is the same as that of pixels of observed samples.

One can further take Equations on (1) and (2) as

$$\nabla^2 E(x, y, \phi_{J, l,z}(x)\phi_{J, l,z}(y)) = (\partial_x^2 - \partial_y^2)Q(x, y) + 2\partial_x \partial_y U(x, y, \phi_{J, l,z}(x)\phi_{J, l,z}(y)),$$  
(5)

$$\nabla^2 B(x, y, \phi_{J, l,z}(x)\phi_{J, l,z}(y)) = 2\partial_x \partial_y Q(x, y) - (\partial_x^2 - \partial_y^2)U(x, y, \phi_{J, l,z}(x)\phi_{J, l,z}(y)).$$  
(6)

With the DWT decomposition of $E(x, y)$ and $B(x, y)$

$$E(x, y) = \sum_{i, l,z} \epsilon^{E}_{i, l,z} \phi_{J, l,z}(x)\phi_{J, l,z}(y),$$  
(7)

$$B(x, y) = \sum_{i, l,z} \epsilon^{B}_{i, l,z} \phi_{J, l,z}(x)\phi_{J, l,z}(y),$$  
(8)

Equations (5) and (6) yield matrix equations

$$\sum_{i, l,z} \left[ T^{(2)}_{i, l,z} \delta_{i, l,z} + T^{(2)}_{i, l,z} \delta_{i, l,z} \right] \epsilon_{i, l,z}^{\Omega} = \sum_{i, l,z} \left[ T^{(2)}_{i, l,z} \delta_{i, l,z} - T^{(2)}_{i, l,z} \delta_{i, l,z} \right] \epsilon_{i, l,z}^{Q},$$  
(9)

$$\sum_{i, l,z} \left[ T^{(2)}_{i, l,z} \delta_{i, l,z} + T^{(2)}_{i, l,z} \delta_{i, l,z} \right] \epsilon_{i, l,z}^{U} = \sum_{i, l,z} \left[ 2T^{(1)}_{i, l,z} T^{(1)}_{i, l,z} \epsilon_{i, l,z}^{U} - T^{(2)}_{i, l,z} \delta_{i, l,z} \epsilon_{i, l,z}^{U} \right].$$  
(10)

where the matrix $T^{(n)}_{i, l,z}$ is given by

$$T^{(n)}_{i, l,z} = \int \phi_{J, l}(x)\phi_{J, l}(y)dx.$$  
(11)

Obviously, we can do the projection of Equations (5) and (6) using any bases in two-dimensional space $L \times L$. However, for proper wavelet scaling functions, the integral $\int \phi_{J, l}(x)\phi_{J, l}(y)dx$ are zero for $J \neq J^*$. This point is important for a wavelet–Galerkin discretization (see discussion in Section 4). In this case, all quantities of Equations (9) and (10) are on scale $J$, and Equation (11) gives

$$T^{(n)}_{i, l,z} = \frac{1}{h^2} r^{(n)}_{l-i},$$  
(12)

where $h = 1/2^J$. $r^{(n)}_{l-i}$ is non-zero only in a narrow band $|l-l'| < M$, where $M$ is an integral, depending on wavelet. For Daubechies 6 wavelet, the non-zero coefficients $r^{(n)}_{l-i}$ are $|l-l'| \leq 4$, or $M = 4$. The values of $r^{(n)}_{l-i}$ and $r^{(n)}_{l-i}$ of Daubechies 6 wavelet are listed in Table 1 of Appendix.

Thus, Equations (1) and (2) defined in continuous space $(x, y)$ are reduced to Equations (9) and (10) defined in a space spanned by orthogonal bases $\phi_{J, l}(x)\phi_{J, l}(y)$. The discretized Equations (9) and (10) are an approximation of Equations (1) and (2). Equations (9) and (10) do not contain information on scales less than $L/2^J$. However, this discretization is reasonable in the sense that it does not introduce false correlations, or lose information of discrete data sets $Q$ and $U$. The derivative operator on a function defined in a space spanned by bases $\phi_{J, l}(x)\phi_{J, l}(y)$ will yield a function in the same space. This is required by a wavelet–Galerkin discretization (Section 4). It ensures no signal to be produced on scales less than $L/2^J$. 

Section 3 presents the $E/B$ separation in the DWT space. Section 4 tests the DWT algorithm with samples having known spatial structures. We show that the method effectively suppresses the uncertainties from boundary effect and noise. It is also effective to identify spatial structures of $E$ and $B$ fields. Section 5 addresses the effectiveness of the wavelet–Galerkin discretization. Finally, conclusions are given in Section 6. The DWT representations of derivative operators are given in the Appendix.
The Equations (9) and (10) give a E/B decomposition from observed maps Q and U.

It should be pointed out that not all wavelets yield J-diagonal matrices like Equation (12). For instance, the popular wavelet Daubechies 4 is not suitable for this discretization, as the matrix of derivative operator in space spanned by Daubechies 4 scaling functions is not J-diagonal. More discussion on the wavelet–Galerkin discretization will be given in Section 4. We will first study how to develop the algorithm of the E/B decomposition with Equations (9) and (10).

2.2. Maps of $\mathbb{E}_{l_t,l_z}$ and $\mathbb{B}_{l_t,l_z}$

Equations (9) and (10) can be rewritten as follows:

$$\sum_{l_t',l_z'} M_{l_t,l_z}(l_t',l_z') \varepsilon^F_{l_t',l_z'} = \mathbb{E}_{l_t,l_z},$$

$$\sum_{l_t',l_z'} M_{l_t,l_z}(l_t',l_z') \varepsilon^B_{l_t',l_z'} = \mathbb{B}_{l_t,l_z},$$

where the matrix $M$ is

$$M_{l_t,l_z}(l_t',l_z') = T_{l_t,l_z}^{(2)} \delta_{l_t,l_t'} + T_{l_t,l_z}^{(2)} \delta_{l_z,l_z'}.$$  (15)

If we use Daubechies 6, $\mathbb{E}_{l_t,l_z}$ and $\mathbb{B}_{l_t,l_z}$ are given by

$$\mathbb{E}_{l_t,l_z} = \sum_{m_1}^{4} T_{m_1}^{(2)} \varepsilon_{l_t+m_1,l_z} + \sum_{m_2}^{4} T_{m_2}^{(2)} \varepsilon_{l_t,l_z+m_2} + \sum_{m_1,m_2=-4}^{4} 2 T_{m_1}^{(1)} T_{m_2}^{(1)} \varepsilon^U_{l_t+m_1,l_z+m_2},$$  (16)

$$\mathbb{B}_{l_t,l_z} = \sum_{m_1,m_2=-4}^{4} 2 T_{m_1}^{(2)} T_{m_2}^{(1)} \varepsilon^Q_{l_t+m_1,l_z+m_2} + \sum_{m_1=-4}^{4} T_{m_1}^{(2)} \varepsilon^U_{l_t+m_1,l_z} + \sum_{m_2=-4}^{4} T_{m_2}^{(2)} \varepsilon^U_{l_t,l_z+m_2},$$  (17)

Equations (13) and (14) look like the matrix equations of the DWT variables of E and B fields, $\varepsilon^F_{l_t,l_z}$ and $\varepsilon^B_{l_t,l_z}$. Equations (16) and (17) give the sources $\mathbb{E}_{l_t,l_z}$ and $\mathbb{B}_{l_t,l_z}$ on the right-hand side of Equations (13) and (14), respectively. It seems that one can separate E/B by solving the matrix Equations (13) and (14). However, with the coefficients $T^{(n)}_{l_t,l_z}$ given in Appendix, we can show

$$\sum_{l_t,l_z} M_{l_t,l_z}(l_t',l_z') = 0.$$  (18)

That is, the matrix $M$ is singular. One cannot use a standard linear solver to solve Equations (13) and (14). This problem, of course, is directly related to the non-uniqueness of the solutions $E(x, y)$ and $B(x, y)$ given by the Poisson Equations (1) and (2) without knowledge of boundary conditions. We will not try to solve the singular matrix Equations (13) and (14), but directly use Equations (16) and (17) for the E/B decomposition.

2.3. E/B Decomposition with $\mathbb{E}_{l_t,l_z}$ and $\mathbb{B}_{l_t,l_z}$

The spatial resolution of the source terms $\mathbb{E}_{l_t,l_z}$ and $\mathbb{B}_{l_t,l_z}$ is the same as maps Q and U. It can be used to calculate the DWT power spectrum of $\nabla^2 E$ and $\nabla^2 B$ fields, and other statistics. To do these, we should first find the wavelet function coefficient (WFC) of the maps $\mathbb{E}_{l_t,l_z}$ and $\mathbb{B}_{l_t,l_z}$ by

$$\varepsilon^E_{j,k} = \sum_{l_t} C_{j,k,l_t} \mathbb{E}_{l_t,l_z},$$  (19)

$$\varepsilon^B_{j,k} = \sum_{l_t} C_{j,k,l_t} \mathbb{B}_{l_t,l_z},$$  (20)

where, for simplification, we use two-dimensional vector notation defined by $j = (j_1, j_2)$ and $l_t = (l_1, l_2)$, and $l_1 = 0...2^{j_1} - 1, l_2 = 0...2^{j_2} - 1$; $j_1, j_2$ can be any integral less than J. Index $(j, l_t)$ refers to the cell on scale $j$ and at position $l_t$. The $\mathbf{I} \times \mathbf{I}$ matrix $C_{j,k,l_t}$ is given by

$$\langle \varepsilon^E_{j,k} \varepsilon^B_{j',k'} \rangle = \int \psi_{j_1,l_1}(x) \psi_{j_2,l_2}(y) \bar{\psi}_{j_1',l_1'}(x) \bar{\psi}_{j_2',l_2'}(y) dxdy,$$  (21)

where $\psi_{j,l}(x)$ is one-dimensional wavelet function referring to cell on scale $j$ and at position $l_t$. $C_{j,k,l_t}$ is a banded matrix with respect to $\mathbf{I}$, $\mathbf{I}$. Therefore, the relation between $\varepsilon^E_{j,k}, \varepsilon^B_{j,k}$ and $Q, U$ are spatially quasi-local.

With the WFCs, the DWT power spectrum is given by (Fang & Feng 2000)

$$P^E,B_{j} = \langle \varepsilon^E_{j,k} \rangle^2,$$  (22)

where $\langle \cdots \rangle$ is the average over all cells 1. One can directly use the DWT power spectrum to measure E and B modes. $P^E,B_{j}$ is banded Fourier power spectrum. For a statistically homogeneous random field, the DWT power spectrum is related to the Fourier power spectrum $P^E,B(n_1, n_2)$ of $(\mathbb{E}, \mathbb{B})$ maps by

$$P^E,B_{j} = \frac{1}{L^2} \sum_{n_1, n_2=-\infty}^{\infty} |\hat{\psi}(n_1/2^j) \hat{\psi}(n_2/2^j)|^2 P^E,B(n_1, n_2).$$  (23)

Clearly, $P^E,B_{j}$ is a banded Fourier power spectrum with the window function

$$W^E_{j}(n_1, n_2) = \frac{1}{L^2} |\hat{\psi}(n_1/2^j) \hat{\psi}(n_2/2^j)|^2.$$  (24)

Function $\hat{\psi}(n)$ is the Fourier transform of the basic wavelet. $P_{E}$ contains all valuable quantities of second-order statistics from random samples in a finite area $L \times L$ and pixel $L/2^j$. The window function may cause spurious features and false correlation, such as aliasing effect, in the Fourier power spectrum. With the DWT analysis, the aliasing effects can be effectively suppressed (Fang & Feng 2000).

2.4. Effect of Noise on Power Spectrum

The maps of Q and U are usually noisy and can be given by $Q + \Delta Q$ and $U + \Delta U$. The DWT variables of noisy DWT maps are then $\varepsilon^Q + \Delta \varepsilon^Q$ and $\varepsilon^U + \Delta \varepsilon^U$, where $\Delta \varepsilon^Q$ and $\Delta \varepsilon^U$ are the DWT variables of $\Delta Q$ and $\Delta U$. Assuming that the noise is Gaussian and statistically homogeneous, the DWT variables $\Delta \varepsilon^Q$ and $\Delta \varepsilon^U$ of noise have to satisfy the following statistical properties:

$$(\Delta \varepsilon^Q)^2 = \sigma_Q^2 \delta_{1T}, (\Delta \varepsilon^U)^2 = \sigma_U^2 \delta_{1T}, (\Delta \varepsilon^Q \Delta \varepsilon^U) = 0,$$  (25)
where $\sigma_Q$ and $\sigma_U$ are the variance of the noise of $Q$ and $U$ maps, respectively, and are independent of $l$.

Using $\Delta Q_l$ and $\Delta U_l$ to replace $\epsilon_l^Q$ and $\epsilon_l^U$ in Equations (16) and (17), we can construct the noise maps of $\Delta E_l$ and $\Delta B_l$. First, with Equation (25) we can show

$$\langle (\Delta E_l^2) \rangle = N^{(1)} \sigma_U^2 + N^{(2)} \sigma_Q^2,$$

(27)

$$\langle (\Delta B_l^2) \rangle = N^{(1)} \sigma_Q^2 + N^{(2)} \sigma_Q^2,$$

(28)

where

$$N^{(1)} = \left[ 2 \sum_{m=1}^{4} (T_m^{(1)})^2 \right]^2, N^{(2)} = 2 \sum_{m=1}^{4} [T_m^{(2)}]^2.$$

(29)

$N^{(1)}$ and $N^{(2)}$ are from the terms containing $T_{l-l}^{(1)}$ and $T_{l-l}^{(2)}$, respectively, in Equations (16) and (17). For Daubechies 6 wavelet, $\sqrt{N^{(1)}} \approx 1.2$ and $\sqrt{N^{(2)}} \approx 5$. That is, in the Daubechies 6 DWT algorithm (Equations (16) and (17)), the operator of derivative $\partial$ does not significantly change the level of the noise, while the operator for $\partial^2$ leads to an increase of the variance by a factor of 5 with respect to the variance of $\Delta Q$ and $\Delta U$ maps. This shows that derivative will generally affect the size of noise. However, the matrix $T_{l-l}^{(m)}$ is exactly diagonal with respect to $l$, the derivative operator in the DWT representation does not transfer the noise from one scale to others. In this sense, we have a handle on the noise.

As noise and signal are statistically uncorrelated, the power spectrum of $E_l$ and $B_l$ can be reconstructed by subtracting the power of noise as

$$P_j^E = P_j^E - P_j^{\Delta E},$$

(30)

$$P_j^B = P_j^B - P_j^{\Delta B},$$

(31)

where $P_j^E$ and $P_j^B$ are the DWT power spectrum of maps $E_l$ and $B_l$ given by Equations (16) and (17), respectively, using noisy $Q$ and $U$. $P_j^{\Delta E}$ and $P_j^{\Delta B}$ are the DWT power spectra of noise maps $\Delta E_l$ and $\Delta B_l$. The algorithm of subtracting the noise DWT power spectrum $P_j^{\Delta E}$ and $P_j^{\Delta B}$ scale by scale is similar to the subtraction of shot noise power from the DWT power spectrum of galaxy survey (Fang & Feng 2000).

2.5. The Effect of Boundary

For a sample of finite area, the DWT power spectrum analysis does not need a window function to treat the spatial domain. The effect of boundary can effectively be reduced by dropping the DWT variables related to cells $(j,l)$ located on or near the boundary (Pando & Fang 1998). When derivative operators, $\partial_x$, $\partial_y$, are involved, the boundary effect would be more serious, because the matrix of derivative operators in the DWT representation is not exactly diagonal with respect to the spatial index $l$. Nevertheless, the matrix $T_{l-l}^{(m)}$ (Equation (11)) is narrowly banded, the effect of boundary can still be reduced by dropping boundary modes.

3. Tests with Samples Having Known Spatial Structures

To test the DWT algorithm developed in Section 3, we consider, in this section, samples with given spatial structures, and compare the maps $E_{l,f}$ and $B_{l,f}$ given by Equations (16) and (17) with that directly calculated from $E$ and $B$. This comparison is only to test the discretization of derivative operator, but says nothing about the amount of information loss associated with the algorithm.

3.1. Samples

We use two scalar functions $\psi_E(x, y)$ and $\psi_B(x, y)$ to produce $E$ and $B$ maps in two-dimensional space by the following way:

$$E = -\nabla^2 \psi_E, B = -\nabla^2 \psi_B.$$  

(32)

One can then produce the DWT variables $\epsilon_{l,f}^E$ and $\epsilon_{l,f}^B$ with Equations (7) and (8). With these results, we can further produce the maps of $E_{l,f}$ and $B_{l,f}$ with Equations (13) and (14). Thus, for given scalar functions $\psi_E(x, y)$ and $\psi_B(x, y)$, we have the samples $E_{l,f}$ and $B_{l,f}$, and then, the DWT power spectrum and other statistical properties of maps $E_{l,f}$ and $B_{l,f}$. The ratio between the powers of $E$ and $B$ modes can be adjusted by the ratio between the functions $\psi_E$ and $\psi_B$.

On the other hand, using the function $\psi_E(x, y)$ and $\psi_B(x, y)$, one can produce the samples of the Stokes parameters $Q(x, y)$ and $U(x, y)$ maps by

$$Q(x, y) = (\partial_x \partial_x - \partial_y \partial_y) \psi_E(x, y) - 2 \partial_x \partial_y \psi_B(x, y),$$

(33)

$$U(x, y) = 2 \partial_x \partial_y \psi_E(x, y) + (\partial_x \partial_x - \partial_y \partial_y) \psi_B(x, y).$$

(34)

We add Gaussian white noise in the $Q$ and $U$ maps pixel-by-pixel with signal-to-noise ratio (S/N) equal to 10, 50, and 100. These $Q$ and $U$ maps are used as the simulation of observed samples.

With noisy maps $Q$ and $U$, we can produce the variables $\epsilon_{l,f}^Q$ and $\epsilon_{l,f}^U$ by the projection of Equations (3) and (4). Finally, we have maps $E_{l,f}$ and $B_{l,f}$ using Equations (16) and (17). Thus, we can test the algorithm by comparing the statistics of the maps $E_{l,f}$ and $B_{l,f}$ given by $Q$ and $U$ (Equations (33) and (34)) with that directly calculated from $E$ and $B$ of Equation (32).

3.2. Recovery of Spatial Structures

The scalar functions $\psi_E$ and $\psi_B$ are taken to be sample A. Gaussian function $\psi_{E,B}(x, y) = a_{E,B} \exp(-(x^2+y^2)/(2\sigma^2)$; sample B. The Legendre function $\psi_{E,B}(x, y) = a_{E,B} \exp(x \cdot P_m(x)P_m(y))$. Both samples are in the area $-0.5 \leq x \leq 0.5$ and $-0.5 \leq y \leq 0.5$ and pixels 512 $\times$ 512, i.e., $J = 8$. The coefficients $a_E$ and $a_B$ are used to adjust the ratio of the powers of $E_{l,f}$ and $B_{l,f}$. We use $a_E = 1$ and $a_B = 1/10$. That is, the power of $E$ mode is larger than $B$ mode by a factor $10^2$. The maps of $E_{l,f}$ for samples A and B in the central square 32 $\times$ 32 pixels are shown Figure 1. The maps of $B_{l,f}$ have the same shape of Figure 1, but the intensity is weaker than Figure 1 by a factor 10.

As mentioned in Section 3.1, with $\psi_E$ and $\psi_B$ one can produce the maps of $Q$ and $U$ by Equations (33) and (34). Using noise-added maps of $Q$ and $U$, we can further calculate noisy variables $E_{l,t_0}$ and $B_{l,t_0}$ with Equations (16) and (17). These are the recovered maps of $E_{l,t_0}$ and $B_{l,t_0}$.
Figure 1. DWT maps of $E_l$ of sample A: $\psi_E = \exp\left(-\frac{x^2 + y^2}{2d^2}\right)$, and $2d^2 = 1600$ (left), and sample B: $\psi_E = \psi_B/a_B = P_m(x)P_m(y)$, and $m = 100$ (right).

Figure 2. Maps $E_{l_1,l_2}$ (top panels) and $B_{l_1,l_2}$ (bottom panel) of sample A, which are recovered by Equations (16) and (17) with noisy maps $Q$ and $U$, and the signal-to-noise ratios are taken to be $S/N = 10$ (left), 50 (middle), and 100 (right).

3.3. Recovery of Power Spectrum

We now turn to the recovery of the power spectrum. For wavelet Daubechies 6, the non-zero elements of the matrix $T_{n,l,l'}$ are in a band $|l - l'| \leq 4$, and therefore, cells distant from boundary, larger than $\Delta l = 4$, will be less affected by the boundary. Thus, one may expect that the power spectrum recovery would be reasonable with dropping four boundary cells.

Figures 4 and 5 show the DWT power spectra of $E_{l,l'}$ and $B_{l,l'}$ of both original and recovered samples of sets A and B, respectively. It includes (1) the power spectra of the original maps, i.e., the map directly given by $E$ and $B$ (Equation (32)); (2) the ratio $P_{o,j}/P_{r,j}$, where $P_{o,j}$ is the original power spectra from maps $E_{l_1,l_2}$ and $B_{l_1,l_2}$ of Equation (32), and $P_{r,j}$ is the recovered power spectra of maps $Q$ and $U$ with Equations (16) and (17) without dropping boundary cells; (3) same as (2) but dropping four boundary cells in the recovered power spectra.

In Figures 4 and 5, the scale $j = (j_1, j_2)$ is described by an effective scale defined as $j_{\text{eff}} = (j_1^2 + j_2^2)^{1/2}$. The samples are symmetric with respect to $x = y$. The power of mode $(j_1, j_2)$ should be the same as $(j_2, j_1)$. Thus,
Figure 3. Same as Figure 3, but for sample B.

Figure 4. Left panel: the DWT power spectra $P_j$ of the original maps $E_l$ and $B_l$ of sample A. The power spectrum on the top is for the $E$ mode, and the lower is for the $B$ mode. The 10 data points correspond to $(j_1, j_2) = (4, 4), (4, 5), (4, 6), (4, 7), (5, 5), (5, 6), (5, 7), (6, 6), (6, 7), \text{ and } (7, 7)$ from left to right. The power-ratio $E/B$ is equal to 10. Middle panel: the ratio $P_{o,j}/P_{r,j}$, where $P_{o,j}$ is the original power spectra from maps $E_{l_1,l_2}$ and $B_{l_1,l_2}$ of Equation (32), and $P_{r,j}$ is the recovered power spectra of $Q$ and $U$ with Equations (16) and (17) and without dropping boundary cells. The top is for $E$ mode, and bottom is for $B$ mode. Right panel: same as middle panel, but with four boundary cells dropped.

Figure 5. Same as Figure 4, but for sample B.
for $J = 8$, the available pairs $(j_2, j_1)$ are $(4, 4), (4, 5), (4, 6), (4, 7), (5, 5), (5, 6), (5, 7), (6, 6), (6, 7), (7, 7)$, corresponding to $J_{\text{eff}} = 3.50, 3.84, 3.95, 3.98, 4.50, 4.84, 4.96, 5.50, 5.84, 6.50$. The modes on scales with $j_1$ or $j_2 = 7$ are dropped, as all cells are affected by the boundary effect.

We see from Figures 4 and 5 that the power spectra can indeed be well recovered by dropping four boundary cells. However, the boundary effects of sample A are less serious than sample B. This is because, for sample A, both $\psi_E, B(x, y)$ and $\partial_n \psi_E, B(x, y)$ are very small at boundary. The contribution to power by boundary cells is low. On the other hand, for sample B, the power spectra recovered without dropping boundary cells are significantly different from the original one. The B-mode power spectrum is hugely affected by the boundary. On small scale, the derivative operator in the DWT representation is determined by data at a few discrete points, which leads to large error. This problem is always present in algorithms involving taking derivative on discrete data sets. Nevertheless, the error caused by boundary is decreases rapidly as the scale increases.

To measure how good the recovery of power spectrum is, we use the ratio $P^{j}_j/P^j_1$ to describe the deviation of the recovered power spectrum with noisy maps from the original one. All error bars are the variances calculated from 100 independent noisy maps. The results are plotted in Figures 6 and 7.

First, we see that the effect of Gaussian noise is small at larger scales, because the noise is added on each pixels (finest scale) of the maps $Q$ and $U$, and the uncertainty on large scales is suppressed. This point can also be seen from the fact that the error bars of modes $(4, 7), (5, 7), (6, 7), (7, 7)$ are much larger than others. It is because the Gaussian noise on smallest scales, $j_1$ or $j_2 = 7$, is not suppressed.

Figures 6 and 7 show that other than the modes with $j_1$ or $j_2 = 7$, the power of E mode can be reasonably recovered up to mode $(6, 6)$, or $J_{\text{eff}} = 4.5$, when $S/N = 50$ (sample A) or $S/N = 10$ (sample B).

An interesting point shown in Figures 6 and 7 is that the effects of noise on samples A and B are different. The error bars of sample A generally are larger than that of sample B. This is probably because for sample A, other than the central part, most cells are smooth, and have low local fluctuations. For those cells, the fluctuations of noise will strongly contaminate the power of original field, especially when derivative is involved. On the other hand, for sample B, most cells have relatively stronger local fluctuations, and the effect of noise is relatively low.

4. TESTS WITH SAMPLES OF GAUSSIAN RANDOM FIELDS

4.1. Samples

With the preparation given in the previous section, we can consider the case that $\psi_E(x, y)$ and $\psi_B(x, y)$ as random fields. The sample of $E$ and $B$ can still be generated with the same procedure of Section 3.1, but $\psi_E(x, y)$ and $\psi_B(x, y)$ are taken to be Gaussian random fields with Fourier power spectra $a_E k^{-\alpha}$.
and \( a_B k^{-a} \), respectively, and \( a = 3.6 \). The variable \( k^2 = k_x^2 + k_y^2 \), and \( k_x, k_y \) are the Fourier variables of \( x, y \) space, respectively. The constant factors \( a_E \) and \( a_B \) are used to adjust the ratio of \( E/B \) power. From Equation (32), the power spectra of \( E \) and \( B \) are

\[
P_E = a_E k^{2-a}, \quad P_B = a_B k^{2-a}.
\]

We produce the maps in an area described by coordinate \((x, y)\) in range \(-0.5 \leq x \leq 0.5\) and \(-0.5 \leq y \leq 0.5\) with pixelized into \(512 \times 512\). With maps \( E \) and \( B \), one can find the maps \( E \) and \( B \) by Equations (13) and (14).

The DWT power spectrum of \( E \) and \( B \) is shown in Figure 8. Since \( J = 8 \), the available modes \((j_1, j_2)\) still are \((4, 4), (4, 5), (4, 6), (4, 7), (5, 5), (5, 6), (5, 7), (6, 6), (6, 7),\) and \((7, 7)\). The modes with \( j_1, j_2 \leq 3 \) are dropped, as they have only boundary cells. The ratio of the \( E/B \) power is taken to be 10. The error bars are from the variance of 100 samples. We see from Figure 8 that the powers of modes \((4, 4), (5, 5),\) and \((6, 6)\) are nearly about the same as \((7, 7)\). Similarly, the powers \((4, 6), (5, 6)\) are nearly about the same as \((6, 6)\); the power \((4, 5)\) is nearly about the same as \((5, 5)\). It is because in the case of \( j_1 < j_2 \), the power is dominated by the small-scale \( f_j \).

**4.2. Effects of Random Field**

Unlike the maps in Section 3, all the maps of \( \psi_E(x, y), \psi_B(x, y); Q, U, \) and \( \mathcal{E}, \mathcal{B} \) are random fields. A serious problem caused by random fields is that the power of \( E \) mode may leak to \( B \) mode, and vice versa. That is, even when the original \( B \)-mode power is zero, the recovered \( B \)-mode power would not be zero.

To demonstrate the power leakage, we take \( \psi_E(x, y) \) to be a Gaussian random field with the same Fourier power spectrum as Figure 8, while \( \psi_B(x, y) \) to be 0. That is, the power of \( B \) mode originally is zero. Figure 9 presents the recovered DWT power spectra of \( E \) and \( B \) modes. We see that the recovered \( E \)-mode power spectrum is nearly the same as the original one shown in Figure 8, except that the recovered power spectrum on the finest scale is a little smaller than the original one. However, the recovered \( B \)-mode power spectrum is not zero. It is spurious \( B \) power. It arises from the leaking of \( E \)-mode power to \( B \) mode. On large scales \( j_{\text{eff}} \leq 4 \), the ratio of the \( E/B \) power is about \( 10^6 \), while on small scales \( j_{\text{eff}} \geq 6 \), this ratio is less than \( 10^2 \). This is caused by the variance of random field. Thus, one may conclude that for the \( 512 \times 512 \) sample of a Gaussian random fields of Equation (35), the developed algorithm would be effective only if the ratio \( E/B \) is less than \( 10^2 \) on small scales.

As a comparison, we plot Figure 10, in which the \( \psi_E(x, y) \) is given by samples A and B, while \( \psi_B(x, y) = 0 \), i.e., the power of \( B \) mode originally is also zero. Figure 10 shows that the recovered \( B \)-mode powers are also not zero. However, it generally is less than the original one by at least 3 orders. These powers seem to come from the numerical processes. Therefore, the errors caused by the variance of random fields are serious.

**4.3. Recovery of \( E, B \) Power Spectra**

As in Section 3.3, we measure the soundness of the recovery of power spectrum by the ratio \( P_j^r/P_j^p \), where \( P_j^p \) is the power spectra of original maps \( \psi_{E,j,l} \) and \( \psi_{B,j,l} \) from Equation (32), and \( P_j^r \) is the recovered power spectra from noisy maps of \( Q \) and \( U \). The Gaussian noise added on the maps \( Q \) and \( U \) are on the levels \( S/N = 100, 20, \) and 10. Similar to Section 3.3, four boundary cells are dropped. The results are plotted in Figures 11, 12, and 13, for which the ratio of the powers of \( E \) and \( B \) are equal to 10, 20, and 100, respectively.

Figure 11 shows that the powers of \( E \) mode can be recovered on all scales \( j_{\text{eff}} \leq 6 \) on all noise levels. On the smallest scale, \( j_{\text{eff}} = 6.5 \), the ratio \( P_j^r/P_j^p \) of \( E \) mode is slightly lower than 1. This deviation is almost independent of the level of \( S/N \). Therefore, the errors mostly are not due to the Gaussian noise addition, but from the effect of leakage. This point is consistent with the leaking shown in Figure 9, which also give a little small power on the smallest scale. More interesting, Figure 11 shows...
that the $B$ mode can be perfectly recovered on all scales and all noise levels considered.

Figure 12 presents the case of $E/B = 20$. The results on scales $j_{\text{eff}} \leq 5.5$ are about the same as the case of $E/B = 10$, while the deviations of $P^e_j$ from $P^o_j$ on scales $j_{\text{eff}} > 5.5$ are larger than that of $E/B = 10$. On small scales, the recovered $E$ powers, $P^e_j$, are little lower than the original power $P^o_j$, while the recovered $B$ powers are slightly higher than the original power.

Figure 13 is for the case of $E/B = 100$. It shows that the recovered $E$-mode power spectrum is still good on scales $j_{\text{eff}} \leq 5.5$ for all S/N. However, on scales $j_{\text{eff}} > 5$, the recovered $B$-mode power spectrum generally is higher than the original one. This deviation is expected, as Figure 9 shows that the leaked power from $E$ mode to $B$ mode can be as high as 1% on small scales. Nevertheless, the recovered $B$-mode power spectrum is reasonable on scales $j_{\text{eff}} < 5$, even when the S/N is equal to 10. That is, one can pick up the weak signal of $B$ mode with the DWT algorithm even when the Gaussian noise level of $Q$ and $U$ maps is comparable with the $B$ mode signal.

5. DISCUSSIONS

The relationships between the polarization maps of $(E, B)$ and $(Q, U)$ are differential. For pixelized samples of $Q$ and $U$ in finite area, the algorithm of $E/B$ decomposition should be able to properly handle the derivative operation on a spatially discrete and noisy data set. The derivative operator $\partial_x$ is a continuous linear operator to mapping functions defined in Hilbert space, while the functions $Q$ and $U$ are defined in space spanned by bases $u_i$, in which $i$ is a set of finite index. Therefore, we should approximate the derivative operator from mapping between functions defined in Hilbert space, to a mapping in subspace spanned by $n_i$.

What we need to calculate is

$$O f = g,$$  \hspace{1cm} (36)

where $O$ is a linear continuous operator, like Laplace or derivative, and $f$ and $g$ are function of $x$, $y$ in continuous space $0 < x, y < L$. However, we do not know $f$, but only the discretized $\tilde{f}$, which is given in $N$ pixels (cells). That is, $\tilde{f}$
can be expressed as
\[ \tilde{f}(x, y) = \sum_{k=1}^{N} \alpha_k w_k(x, y), \]  
(37)

where function \( w_k(x, y) \) is the binning function of pixel \( k \), and \( \alpha_k \) is the observed \( f \) at pixel \( k \). The simplest binning function would be the top-hat window function of pixel \( k \).

With \( N \)-dimensional space spanned by bases \([v_1(x, y), \ldots, v_N(x, y)]\), Equations (36) and (37) yield
\[ g_i \equiv \langle g, v_i \rangle = \sum_{k=1}^{N} \langle Ow_k, v_i \rangle \alpha_k. \]  
(38)

The matrix \( O_{ik} \equiv \langle Ow_k, v_i \rangle \) gives a discretization of operator \( O \) from the space \( 0 < x, y < L \) to a finite-dimensional subspace \([v_i]\).

A Galerkin discretization requires the following equation to be hold for all \( v_i \)
\[ \langle (g - Of), v_i \rangle = 0 \]  
(39)

That is, Equation (36) should be hold in the subspace spanned by bases \([v_1(x, y), \ldots, v_N(x, y)]\). In this case, the matrix \( O_{ik} \) is a linear operator to map functions defined in the subspace \( v_i \). If \( f \) is a function of the subspace \( v_i \), \( g = Of \) is also a function of the subspace.

It can be seen from Equations (36)–(38) that the discretization of operator \( O \) actually is inevitable for all algorithms. To treat the data Equation (37), we must use some base \( v_i \) in the spatial domain. It will yield a matrix \( O_{ik} \), regardless of whether Equation (39) is held with the bases \( v_i \). The Galerkin method gives best discretization of Equation (36) or the operator \( O \) (e.g., Louis et al. 1997).

We use \( \phi_{l,l'}(x) \phi_{l,l'}(y), l, l' = 0, 2^J - 1 \) to be the bases to span the subspace with dimension \( N = 2^J \times 2^J \). It can be shown that the conditions of Galerkin discretization, Equation (39), will be satisfied for operator \( O = \partial_x, \partial_y, \partial_x^2 \) and \( \partial_y^2 \). That is, for any function \( f(x, y) \) of the subspace spanned by bases \( \phi_{l,l'}(x) \phi_{l,l'}(y), l, l' = 0, 2^J - 1 \), the result of \( Of \) are also functions of the subspace. This is the wavelet–Galerkin discretization. Matrix \( \langle Ow_k, v_i \rangle \) will be invertible. In this sense, discretization does not lose information, or introduces false data or correlations.

Obviously, the Galerkin discretization is not unique. One can use different wavelets to do the Galerkin discretization. To apply the discretization, the matrix \( O_{ik} \equiv \langle Ow_k, v_i \rangle \) should have the following desirable properties. First, the matrix \( O_{ik} \) has to be sparse, narrowly banded. In this case, one can effectively minimize the information lose due to dropping boundary modes. A narrowly banded matrix can also effectively reduce the spreading of errors among cells with different \( l \). Second, in order that the errors not increase with the size of the matrix \( N \) (number of data), the “width” of the band in which the matrix elements \( O_{ik} \) is non-zero, should be independent on \( N \).

6. CONCLUSIONS
The algorithm developed in this paper can be summarized as follows.

1. From observed noisy and discrete maps \( Q(x, y) \) and \( U(x, y) \) we calculate their DWT maps \( Q_{l_1,l_2} \) and \( U_{l_1,l_2} \) on the finest scale \( j = (J, J) \), which is given by the resolution.
2. Using Equations (16) and (17), we decompose \( Q_{l_1,l_2} \) and \( U_{l_1,l_2} \) into \( E_{l_1,l_2} \) and \( B_{l_1,l_2} \).

3. Using Equations (19) and (20), we calculate WFCs \( \tilde{\xi}_{j,1}^E \) and \( \tilde{\xi}_{j,1}^B \) on scales \( (j_1, j_2) \) and \( j_1, j_2 \leq J \).

4. Using the WFC maps \( \tilde{\xi}_{j,1}^E \) and \( \tilde{\xi}_{j,1}^B \), we calculate the DWT power spectra by Equations (30) and (31).

5. We identify spatial structures with maps of \( E_{l_1,l_2} \) and \( B_{l_1,l_2} \).

With this algorithm, it is possible to recover the power spectrum of B-mode random fields from noisy Stokes parameter maps \( Q \) and \( U \) when the power ratio \( E/B \) is as high as \( 10^2 \), and the S/N is equal to or higher than 10. For samples with given structures, the B-mode structure can also be identified when the power ratio \( E/B \) is equal to \( 10^2 \). Besides power spectrum, the DWT variables of SFCs \( (\epsilon_{j,1}^E, \epsilon_{j,1}^B) \) and WFCs \( (\tilde{\xi}_{j,1}^E, \tilde{\xi}_{j,1}^B) \) can be used for high-order statistics, such as high-order moments, scale-scale correlation, cross correlation between the \( E \) and \( B \) and other maps.

With the DWTs, one can construct orthogonal, divergence-free vector waves. It has been used for a local analysis of the velocity field of incompressible turbulence (Urban 1995; Kishida et al. 1999; Albukrek et al. 2002). The divergence-free B field is similar to a two-dimensional velocity field of turbulence (e.g., Pina 1998). Therefore, it would be valuable to further study the DWT \( E/B \) decomposition with the divergence-free vector waves.

We thank Dr. Priya Jamkhedkar for her help. L.C. acknowledges the support by the Knowledge Innovation Program of the Chinese Academy of Sciences (09090611009) and the CAS Special Grant for Postgraduate Research, Innovation, and Practice (09092921009). This work is partially supported by NSFC (10530303, 10878001), US NSF AST-0507340, and ICRAnet grant 2008.

**APPENDIX**

**DERIVATIVE OPERATOR IN WAVELET REPRESENTATION**

In the DWT space, the operators of derivatives are represented as a matrix

\[
T_{j,l,r}^{(n)} \delta_{j,j'} = \int \phi_{j,l}(x) \partial_r^n \phi_{j',l}(x) dx. \quad (A1)
\]

That is, the matrix is diagonal with respect to \( j, j' \). \( T_{j,l,r}^{(n)} \) is given by (Beylkin 1992; Kwon 1998)

\[
T_{j,l,r}^{(n)} = \frac{1}{h^n} r_{j,l}^{(n)}. \quad (A2)
\]

| Table 1 | Coefficient of \( r_m^{(n)} \) |
|---------|-----------------------|
| \( l-l'=m \) | \( n=1 \) | \( n=2 \) |
| 4       | 1/2920 | 3/560 |
| 3       | 1/16/1095 | 4/35 |
| 2       | 272/365 | 356/105 |
| 1       | 0 | −295/56 |
| 0       | −272/365 | 356/105 |
| −1      | 53/365 | −92/105 |
| −2      | −16/1095 | 4/35 |
| −3      | −1/2920 | 3/560 |

where \( h = 1/2^l \). The matrix elements \( r_{l-l'}^{(n)} \) depend on the type of wavelet. For Daubechies 6 wavelet, the non-zero coefficients \( r_{l-l'}^{(n)} \) are \( |l-l'| \leq 4 \). The values of \( r_{l-l'}^{(n)} \) are listed in Table 1, in which \( m = l - l' \). Therefore, the coefficients of \( T_m^{(n)} \) of Equations (11) and (12) are given by

\[
T_m^{(n)} = \frac{1}{h^n} r_m^{(n)}. \quad (A3)
\]

It is interesting to see that the non-zero band of first and second-order derivative operators \( \partial_r \) and \( \partial_r^2 \) are the same. This is different from the estimation of derivative operator by differential approximation.

**REFERENCES**

Albukrek, C. M., Urban, K., Rempfer, D., & Lumley, J. L. 2002, J. Sci. Comput., 17, 49
Beylkin, G. 1992, SIAM J. Numer. Anal., 29, 1716
Bunn, E. F., Zaldarriaga, M., Tegmark, M., & Oliveira-Costa, A. 2003, Phys. Rev. D, 67, 023501
Chiu, T., & Ma, C. J. 2002, ApJ, 578, 12
Fang, L. Z., & Feng, L. L. 2000, ApJ, 539, 5
Fang, L. Z., & Thews, R. 1998, Wavelet in Physics (Singapore: World Scientific)
Kamionkowski, M., Kosowsky, A., & Stebbins, A. 1997, Phys. Rev. D, 55, 7368
Kishida, K., Araki, K., Kishiba, S, & Suzuki, K. 1999, Phys. Rev. Lett., 83, 5487
Kwon, S. G. 1998, Bull. Korean Math. Soc., 35, 629
Lewis, A., Challinor, A., & Turok, N. 2002, Phys. Rev. D, 65, 3505
Louis, A. K., Maass, P., & Rieder, A. 1997, Wavelets: Theory and Applications (New York: Wiley)
Mukherjee, P., Hobson, M. P., & Lasenby, A. N. 2000, MNRAS, 318, 1157
Pando, J., & Fang, L. Z. 1998, A&A, 340, 325
Pando, J., Valls-Gabaud, D., & Fang, L. Z. 1998, Phys. Rev. Lett., 81, 4568
Pino, E. 1998, J. Plasma Physs., 59, 719
Sanz, J. L., Argeso, F., Cayn, L., Martnez-Gonzlez, E., Barreiro, R. B., & Toffolati, L. 1999, MNRAS, 309, 672
Seljak, U. 1997, ApJ, 482, 6
Smith, K. M., & Zaldarriaga, M. 2007, Phys. Rev. D, 76, 043001
Urban, K. 1995, Adv. Comput. Math., 4, 51
Zaldarriaga, M., & Seljak, U. 1997, Phys. Rev. D, 55, 1830