Novel approach to reconstructing signals of isotropy violation from a masked CMB sky

Pavan K. Aluri,[1,*] Nidhi Pant,[1,†] Aditya Rotti,[1,2,‡] and Tarun Souradeep[1,§]

[1]*IUCAA, Pune 411007, India*
[2]*Florida State University, Tallahassee, Florida 32304, USA*

Statistical isotropy (SI) is one of the fundamental assumptions made in cosmological model building. This assumption is now being rigorously tested using the almost full sky measurements of the CMB anisotropies. A major hurdle in any such analysis is to handle the large biases induced due to the process of masking. We have developed a new method of analysis, using the bipolar spherical harmonic basis functions, in which we semianalytically evaluate the modifications to SI violation induced by the mask. The method developed here is generic and can be potentially used to search for any arbitrary form of SI violation. We specifically demonstrate the working of this method by recovering the Doppler boost signal from a set of simulated, masked CMB skies.

I. INTRODUCTION

Cosmic microwave background (CMB) observations have ushered in the precision era in cosmology. This enables us to test the basic premises on which the standard cosmology rests, such as the assumption of statistical isotropy via the *cosmological principle*. Some claims of statistical isotropy (SI) breakdown that appeared in CMB literature were revisited by WMAP and Planck satellite science teams; see Refs. [1–3] and the references therein.

The bipolar spherical harmonic (BipoSH) formalism [4,5] is a powerful tool to identify and study SI violation. This has been one of the methods used to search for deviation from isotropy in the CMB data [1,2,6].

The idea behind searching for SI violation is simple: it involves searching for nonvanishing BipoSH coefficients. This is made nontrivial due to the presence of a mask, as it induces spurious correlations that alias the measurements. In this article, we present a new methodology based on the BipoSH formalism, to estimate signals of SI violation from a masked CMB sky. The algorithm presented here is generic and can be potentially applied to recover any signal of SI violation, such as weak lensing, etc. We have already applied the method described here to Planck 2015 data to recover low-l dipole modulation and Doppler boost signals, and the results are part of Ref. [3]. The current paper elaborates the method employed there in detail.

In this article we demonstrate our procedure by recovering Doppler boost field from a Doppler-boosted CMB sky in the presence of a mask. Doppler boost is a well known and guaranteed isotropy violation in CMB maps at a

challenging level of subtlety. So, we choose to apply our machinery to estimate the Doppler signal, whose faithful recovery would establish the reliability of our method.

The paper is structured as follows. We briefly recap the BipoSH formalism in Sec. II. In Sec. III we first discuss the SI violation induced due to Doppler boosting and masking separately and finally the resultant SI violation when both the effects are considered together. In Sec. IV we discuss the minimum variance estimator that takes into account the effects of masking for an arbitrary signal. The performance of the new estimator is discussed in Sec. V, with the specific example of Doppler boost. Finally our conclusions are presented in Sec. VI.

II. BIPOSH: A BRIEF SUMMARY

A CMB anisotropy map $T(\hat{n})$ defined on a sphere is conventionally decomposed in terms of spherical harmonics (Y_{lm}) as

$$T(\hat{n}) = \sum_{l=1}^{\infty} \sum_{m=-l}^{+l} a_{lm} Y_{lm}(\hat{n}), \qquad (1)$$

where $\hat{n}$ denotes position coordinates on the sphere, $T(\hat{n})$ are the temperature anisotropies observed in the direction $\hat{n}$ and a_{lm} are the spherical harmonic coefficients of expansion.

The two-point correlation function is defined as

$$C(\hat{n}_1, \hat{n}_2) = \langle T(\hat{n}_1) T(\hat{n}_2) \rangle, \qquad (2)$$

where $\langle \cdots \rangle$ denotes an in principle average over an ensemble of statistically independent CMB realizations. If the field T is statistically isotropic, then it can be argued that the correlation function cannot have any explicit directional dependence and that it can only be a function of the separation θ between the two directions $\hat{n}_1$ and $\hat{n}_2$, where $\cos\theta = \hat{n}_1 \cdot \hat{n}_2$. As a result of this simplification, the two-point correlation function is given by

---
[*]aluri@iucaa.in
[†]nidhip@iucaa.in
[‡]adityarotti@gmail.com
[§]tarun@iucaa.in

$$C(\hat{n}_1, \hat{n}_2) \equiv C(\theta) = \frac{1}{4\pi} \sum_{l=1}^{\infty} (2l+1) C_l P_l(\cos\theta), \quad (3)$$

where P_l are the Legendre polynomials and C_l is the well-known *angular power spectrum*. When expressed in terms of the harmonic space covariance matrix

$$\langle a_{lm} a_{l'm'}^* \rangle = C_l \delta_{ll'} \delta_{mm'}, \quad (4)$$

C_l's are the diagonal and the only nonvanishing elements of the matrix. Hence for a statistically isotropic Gaussian random field, C_l completely characterize the statistical properties of the field.

In the absence of statistical isotropy, C_l do not completely describe the CMB sky and the BipoSH basis functions [4] are better suited for this case. BipoSH forms a complete orthonormal basis for an $S^2 \times S^2$ space and hence can be used to describe any two-point correlation function for a field defined on a sphere as

$$C(\hat{n}_1, \hat{n}_2) = \sum_{L,M,l_1,l_2} A_{l_1 l_2}^{LM} \{Y_{l_1}(\hat{n}_1) \otimes Y_{l_2}(\hat{n}_2)\}_{LM}, \quad (5)$$

without making assumptions about statistical isotropy of the field. Here $A_{l_1 l_2}^{LM}$ are the coefficients of expansion and $\{Y_{l_1} \otimes Y_{l_2}\}_{LM} = \sum_{m_1 m_2} C_{l_1 m_1 l_2 m_2}^{LM} Y_{l_1 m_1} Y_{l_2 m_2}$ are the bipolar spherical harmonic basis functions [7]. $C_{l_1 m_1 l_2 m_2}^{LM}$ are the Clebsch-Gordan coefficients whose indices satisfy the properties (a) $|l_1 - l_2| \leq L \leq |l_1 + l_2|$, (b) $m_1 + m_2 = M$, (c) $-l_1 \leq m_1 \leq l_1$, and (d) $-l_2 \leq m_2 \leq l_2$.

The harmonic space covariance matrix for a statistically anisotropic field is fully described in terms of the BipoSH coefficients as

$$\langle a_{l_1 m_2} a_{l_2 m_2} \rangle = \sum_{LM} A_{l_1 l_2}^{LM} C_{l_1 m_1 l_2 m_2}^{LM}, \quad (6)$$

in complete analogy with Eq. (4). Therefore the BipoSH coefficients completely characterize the statistical properties of a Gaussian, statistically anisotropic random field. It can be shown that $A_{ll}^{00} = (-1)^l \sqrt{2l+1} C_l$. Hence the BipoSH coefficients are a generalization of the commonly studied angular power spectrum. Note that Eq. (6) can also be used to define an estimator for the BipoSH coefficients as follows:

$$\hat{A}_{l_1 l_2}^{LM} = \sum_{m_1 m_2} C_{l_1 m_1 l_2 m_2}^{LM} a_{l_1 m_1} a_{l_2 m_2}. \quad (7)$$

It is easy to see that this is an unbiased estimator, since it converges to the true BipoSH coefficients on averaging over an ensemble of CMB skies. These BipoSH coefficients have the following symmetry properties:

(i) exchange symmetry: $C(\hat{n}_1, \hat{n}_2) = C(\hat{n}_2, \hat{n}_1)$

$$A_{l_2 l_1}^{LM} = (-1)^{l_1 + l_2 - L} A_{l_1 l_2}^{LM}; \quad (8)$$

(ii) reality condition: $C^*(\hat{n}_1, \hat{n}_2) = C(\hat{n}_1, \hat{n}_2)$

$$A_{l_1 l_2}^{LM*} = (-1)^M (-1)^{l_1 + l_2 - L} A_{l_1 l_2}^{L,-M}. \quad (9)$$

Alternatively, following the triangularity condition between $\{l_1, l_2, L\}$, it is convenient to define the BipoSH coefficients as $A_{l,l+D}^{LM}$, where we set $l_1 = l$, and $l_2 = l + D$, with the constraint that $|D| \leq L$. Needless to say, following the above-mentioned properties of the BipoSH coefficients, determining the BipoSH coefficients for $0 \leq D \leq L$ and $0 \leq M \leq L$ will amount to computing all the BipoSH coefficients, $A_{l_1 l_2}^{LM}$.

BipoSH has been a powerful tool in exploring isotropy violation in a blind as well as honed manner [8]. It has earlier been used to probe SI violation as induced due to specific models of isotropy violation (phenomenological), cosmic topology (cosmological) and beams (systematic) [1,2,9–11].

III. EFFECT OF ISOTROPY-VIOLATING PHENOMENA

Here we analytically derive the form of BipoSH coefficients as a function of multipole l, induced by Doppler boosting of the CMB sky. Following this, we derive the effect of masking on BipoSH coefficients, in addition to the anisotropy due to Doppler boost. Based on these analytic studies, a new estimator is obtained to recover the Doppler boost vector from a masked CMB sky.

A. BipoSH coefficients due to Doppler boosting of CMB anisotropies

One of the known phenomena that leads to breakdown of statistical isotropy is our relative motion ($\vec{v}$) with respect to the CMB rest frame (i.e. the frame in which an observer does not see a dipole). The strongest anisotropy in an otherwise uniform background temperature of $\mathcal{T}_0 = 2.7255$ K [12,13] is an excess (deficit) temperature in the direction (opposite direction) of this relative motion. The Doppler boost ($\vec{\beta} = \vec{v}/c$) that induces this large scale dipole temperature anisotropy ($= \mathcal{T}_0 \vec{\beta} \cdot \hat{n}$) has a well-measured amplitude and direction given by $|\vec{\beta}| = 1.23 \times 10^{-3}$ and $\hat{\beta} = (264°, 48°)$ in galactic coordinates, respectively [14,15]. In addition to this dipole anisotropy, the velocity boost also modulates and aberrates the CMB temperature fluctuations leading to SI violation [16–21]. These aberration and modulation effects can be used to make an independent measurement of the velocity of our local motion. It has been measured with ESA's Planck satellite data also [3,22,23].

Doppler-boosted CMB temperature anisotropies are given by [22]

$$T(\hat{n}) = \mathcal{T}_0 \vec{\beta} \cdot \hat{n} + \Theta(\hat{n} - \nabla(\vec{\beta} \cdot \hat{n}))(1 + b_\nu \vec{\beta} \cdot \hat{n}), \quad (10)$$

where T and Θ correspond to boosted and unboosted temperature anisotropies, respectively, and $\vec{\beta} = \vec{v}/c$ is the Doppler boost vector. From Eq. (10) we see that, in addition to the dominant dipole component due to the Doppler effect, Doppler boosting also aberrates and modulates the CMB anisotropies. The modulation component has a frequency-dependent factor $b_\nu = (\nu/\nu_0)\coth(\nu/2\nu_0) - 1$, where $\nu_0 = k_B \mathcal{T}_0/h \approx 57$ GHz.

For the rest of the discussion we will not bother with the first term in Eq. (10) since we often work with dipole subtracted maps. Evaluating the two-point correlation function for the temperature anisotropies of a Doppler-boosted CMB sky described in Eq. (10) and ignoring the dipole term, it can be shown that the resultant BipoSH coefficients can be cast in the following form:

$$A_{l_1 l_2}^{LM} = (A_{l_1 l_2}^{LM})_{\text{ub.cmb}} + \beta_{LM} H_{l_1 l_2}^L, \quad (11)$$

up to first order in the isotropy-violating Doppler boost field, $\beta(\hat{n}) = \vec{\beta} \cdot \hat{n}$. Note that Doppler boosting generates only the $L = 1$ BipoSH modes. The term $(A_{l_1 l_2}^{LM})_{\text{ub.cmb}}$ corresponds to BipoSH coefficients due to *unboosted CMB* anisotropies, β_{LM} are the spherical harmonic coefficients of $\beta(\hat{n})$, and $H_{l_1 l_2}^L$ is the shape function corresponding to Doppler boosting given by

$$H_{l_1 l_2}^L = b_\nu (G_{l_1 l_2}^L)_{\text{mod}} - (G_{l_1 l_2}^L)_{\text{abr}}, \quad (12)$$

where

$$(G_{l_1 l_2}^L)_{\text{mod}} = \frac{C_{l_1} + C_{l_2}}{\sqrt{4\pi}} \frac{\Pi_{l_1} \Pi_{l_2}}{\Pi_L} C_{l_1 0 l_2 0}^{L0}, \quad (13a)$$

$$(G_{l_1 l_2}^L)_{\text{abr}} = \frac{[C_{l_1} F(l_1, L, l_2) + C_{l_2} F(l_2, L, l_1)]}{\sqrt{4\pi}}$$
$$\times \frac{\Pi_{l_1} \Pi_{l_2}}{\Pi_L} C_{l_1 0 l_2 0}^{L0}, \quad (13b)$$

$$F(l_1, L, l_2) = \frac{l_1(l_1 + 1) + L(L + 1) - l_2(l_2 + 1)}{2}, \quad (13c)$$

where $\Pi_l = \sqrt{2l + 1}$ and $(G_{l_1 l_2}^L)_{\text{mod}}$ and $(G_{l_1 l_2}^L)_{\text{abr}}$ are the shape functions due to modulation and aberration effects, respectively, and are completely specified by the angular power spectrum C_l.

A full sky minimum variance estimator for the velocity boost can then be defined as a weighted linear combination of the observed BipoSH coefficients, $A_{l_1 l_2}^{LM}$. See Appendix A for details.

B. Effect of masking on BipoSH coefficients

Even though an ideal CMB sky is isotropic, masking the sky to avoid foregrounds makes it highly anisotropic. In this section we derive the BipoSH coefficients arising due to the process of masking. Let a masked CMB sky be defined as

$$\tilde{T}(\hat{n}) = \mathcal{W}(\hat{n}) T(\hat{n}), \quad (14)$$

where $\tilde{T}$ and T denote the masked and unmasked CMB sky, respectively, and $\mathcal{W}(\hat{n})$ denotes the mask used for the analysis. Evaluating the two-point correlation function for the masked CMB sky, it can be shown that the resultant BipoSH coefficients are given by

$$\tilde{A}_{l_1 l_2}^{LM} = \sum_{l_3 l_4} \frac{\Pi_{l_3} \Pi_{l_4}}{\sqrt{4\pi}} \sum_{l_5 l_6} \frac{\Pi_{l_5} \Pi_{l_6}}{\sqrt{4\pi}} C_{l_3 0 l_5 0}^{l_1 0} C_{l_4 0 l_6 0}^{l_2 0}$$
$$\times \sum_{L'M'JK} \begin{Bmatrix} L & l_1 & l_2 \\ L' & l_3 & l_4 \\ J & l_5 & l_6 \end{Bmatrix}$$
$$\times \Pi_{L'} \Pi_J A_{l_3 l_4}^{L'M'} W_{l_5 l_6}^{JK} C_{L'M'JK}^{LM}, \quad (15)$$

where $\tilde{A}_{l_1 l_2}^{LM}$ denotes the BipoSH coefficient for the masked CMB sky, $A_{l_3 l_4}^{L'M'}$ denote the BipoSH coefficients of the full sky map, which is not assumed to be statistically isotropic for generality, $W_{l_5 l_6}^{JK}$ are the BipoSH coefficients of the mask and $\{\}_{3\times3}$ denotes the $9j$ symbol. The mask BipoSH coefficients are defined in the same way as the CMB sky. This equation is a generalization of the MASTER kernel [24] which describes how the angular power spectrum (C_l) is modified by mask. In the special case of $L = 0$ and $L' = 0$, this equation exactly reduces to the MASTER equation relating full sky power spectrum C_l to partial sky power spectrum $\tilde{C}_l$.

C. BipoSH of a masked anisotropic sky

The full sky BipoSH coefficients of an anisotropic Doppler-boosted CMB sky are given by Eq. (11). Thus the BipoSH coefficients of a masked Doppler-boosted CMB sky can be obtained by substituting Eq. (11) in Eq. (15) to get

$$\tilde{A}_{l_1 l_2}^{LM} = (\tilde{A}_{l_1 l_2}^{LM})_{\text{ub.cmb}} + \sum_{L'M'} \beta_{L'M'} K_{LMl_1 l_2}^{L'M'}, \quad (16)$$

where

$$(\tilde{A}_{l_1 l_2}^{LM})_{\text{ub.cmb}} = (-1)^{l_1 + l_2 + L} \sum_{l_3} (-1)^{l_3} C_{l_3} \frac{\Pi_{l_3}^2}{\sqrt{4\pi}}$$
$$\times \sum_{l_5 l_6} \frac{\Pi_{l_5} \Pi_{l_6}}{\sqrt{4\pi}} C_{l_3 0 l_5 0}^{l_1 0} C_{l_3 0 l_6 0}^{l_2 0} W_{l_5 l_6}^{LM}$$
$$\times \begin{Bmatrix} l_5 & l_6 & L \\ l_2 & l_1 & l_3 \end{Bmatrix} \quad (17)$$

denotes the BiPoSH coefficients generated due to masking an unboosted CMB sky characterized by C_l. $\beta_{L'M'}$ are the harmonic coefficients of the Doppler field $\beta(\hat{n}) = \vec{\beta} \cdot \hat{n}$, and the term $\{\}_{2\times3}$ denotes the $6j$ symbol. The modified shape function (MSF) $K^{L'M'}_{LMl_1l_2}$, of the Doppler signal, is the masked analogue of the full sky shape function $H^L_{l_1l_2}$ in Eq. (12) and is given by

$$K^{L'M'}_{LMl_1l_2} = \sum_{l_3l_4} \Pi_{L'} H^{L'}_{l_3l_4} \frac{\Pi_{l_3}\Pi_{l_4}}{\sqrt{4\pi}} \sum_{l_5l_6} \frac{\Pi_{l_5}\Pi_{l_6}}{\sqrt{4\pi}} C^{l_10}_{l_30l_50} C^{l_20}_{l_40l_60}$$

$$\times \sum_{JK} \begin{Bmatrix} L & l_1 & l_2 \\ L' & l_3 & l_4 \\ J & l_5 & l_6 \end{Bmatrix} \Pi_J W^{JK}_{l_5l_6} C^{LM}_{L'M'JK}. \quad (18)$$

The modified shape function $K^{L'M'}_{LMl_1l_2}$ incorporates the mixing of modes $\{JK\}$ due to the mask and the intrinsic anisotropic modes $\{L'M'\}$, giving rise to the observed modes $\{LM\}$. This coupling is captured by the Clebsch-Gordan coefficient $C^{LM}_{L'M'JK}$ and the $9j$ symbol.

Hence, from Eq. (16), it can be seen that masking may leak power from an intrinsic anisotropic mode L' to any observed mode L. Further note that the modified shape function due to masking is now phase (M,M') dependent unlike in Eq. (12). Here we emphasize that the formalism discussed in this section is very generic [8], since we assumed no particular form of isotropy violation.

IV. PARTIAL SKY ESTIMATOR

A. Approximations

In general, masking leads to highly entangled modes as seen from Eqs. (16) and (18). We now specifically consider the case of BiPoSH coefficients generated due to a Doppler-boosted CMB sky. Though in a full sky Doppler-boosted CMB sky the signal is only in $L' = 1$ mode of BiPoSH coefficients, the masked Doppler-boosted sky can have power leaked to $L \neq 1$ modes. The mode coupling of the intrinsic anisotropic signal ($L' = 1$) with the mask indices (J) giving rise to the observed modes (L) is tabulated in Table I.

We now argue that, even though there is coupling between different modes, it is reasonable to assume that only modes with $L' = L = 1$ and $M = M'$, which we refer to as the diagonal approximation, are sufficient to evaluate Eq. (16).

Most masks used in CMB analysis can at first order be approximated as a band along the galactic equator where the astrophysical emission from our own Galaxy is the highest and hence have a significant azimuthal symmetry. It can be argued that, for such a mask, the mixing kernel is primarily diagonal in $\{M, M'\}$, motivated by the fact that for a band mask, which has perfect azimuthal symmetry,

TABLE I. Illustrated here are the observed BiPoSH indices (L) that result from the mixing of an intrinsically anisotropic signal in $L' = 1$ with the mask BiPoSH indices (J). For a mask which is largely azimuthally oriented, the dominant mask modes contributing to the mixing kernel are highlighted in bold.

Intrinsic anisotropic signal, $L' = 1$	
Observed BiPoSH index, L	Mask BiPoSH index, $J = \|L - L'\|$ to $L + L'$
$L = 1$	$J = \mathbf{0}, 1, \mathbf{2}$
$L = 2$	$J = 1, \mathbf{2}, 3$
$L = 3$	$J = \mathbf{2}, 3, \mathbf{4}$

the $\{M, M'\}$ coupling can be shown to be identically zero. For mask BiPoSH coefficients it is seen that the $J = 0$ mode is the most dominant mode. Observing the various couplings in Table I it can be seen that only the $L = 1$ mode couples to $L' = 1$ mode via $J = 0$, while all higher $L > 1$ modes couple to $L' = 1$ via $J > 0$, suggesting that the contribution from terms where $L \neq L'$ will be subdominant.

We follow these heuristic arguments by quantitatively showing that the off-diagonal terms are in fact subdominant by explicitly evaluating the kernel $K^{L'M'}_{LMll'}$. In Fig. 1 we compare the full sky shape function $H^L_{ll'}$ of Doppler boosting with the modified shape function $K^{L'M'}_{LMll'}$,
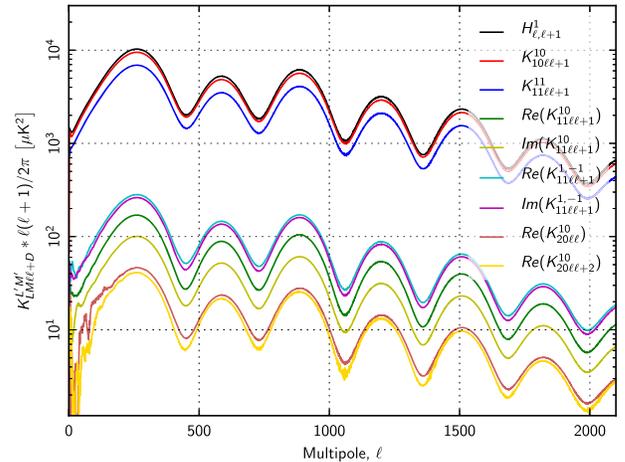

FIG. 1 (color online). In this figure we compare the full sky shape function $H^1_{l,l+1}$ and the modified shape function $K^{L',M'}_{L,M,l,l+1}$, modified due to the mask, for the Doppler-boosted CMB sky. The full sky shape function is independent of phase as it does not depend on M. The most dominant term of the modified shape function is the diagonal denoted by $K^{L,M}_{L,M,l,l+1}$, i.e. when $L = L'$ and $M = M'$. Also plotted are some of the off-diagonal terms of the shape function, i.e. $L \neq L'$ and $M \neq M'$, which are seen to be negligible when compared to the diagonal and hence ignored in the analysis. Further note that, for the diagonal terms, there is an M-dependent suppression of power as compared to $H^1_{l,l+1}$. The varying suppression of power between modes corresponding to different M is determined by the morphology of the mask.
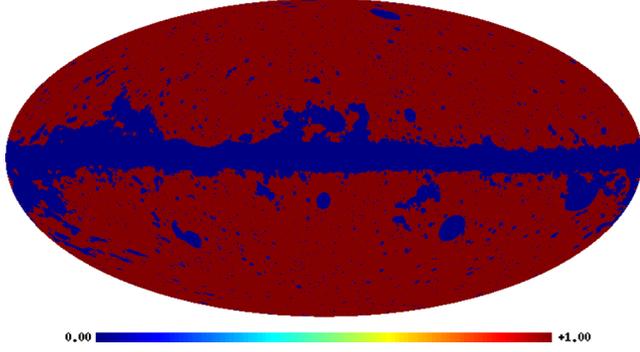
FIG. 2 (color online).　The unapodized galactic mask used in the analysis is shown here. This $\{1, 0\}$ mask has an available sky fraction of $f_{\text{sky}} \approx 0.78$, which remains practically the same on apodization.

evaluated using the apodized version of the mask shown in Fig. 2 (used for all analyses presented in this article). They are plotted using WMAP normalization for even parity BipoSH coefficients [1]. Note that the diagonal terms are nearly 2 orders of magnitude larger than the off-diagonal terms. This justifies the approximations suggested above.

Here it is also important to note that the modified shape function is different for different phase modes ($M = 0, 1$). This distinction is purely determined by the details of the mask. As we will see in the following sections, this difference in MSF corresponding to different phase modes (M) naturally compensates for the differential loss of signal in these modes while recovering the Doppler field harmonics.

Note that in Eq. (16) it is fairly complicated to algebraically solve for the Doppler field harmonics β_{LM}, since it appears inside a summation. By making the approximations discussed in this section, we have simplified Eq. (16) to the form

$$\tilde{A}_{l_1 l_2}^{LM} = (\tilde{A}_{l_1 l_2}^{LM})_{\text{ub.cmb}} + \beta_{LM} K_{LM l_1 l_2}^{LM}. \qquad (19)$$

This is particularly useful, since it can be used to define a partial sky estimator, similar to the full sky estimator.

B. Weighted variance modified shape function estimator (WVMSFE)

In this section we briefly discuss the Doppler estimator. Equation (19) can be inverted to arrive at an estimator for the Doppler field harmonics:

$$\hat{\beta}_{LM} = \frac{\tilde{A}_{l_1 l_2}^{LM} - \langle \tilde{A}_{l_1 l_2}^{LM} \rangle_{\text{ub.cmb}}}{K_{LM l_1 l_2}^{LM}}, \qquad (20)$$

where $\tilde{A}_{l_1 l_2}^{LM}$ denotes the BipoSH coefficients derived from the data maps and $\langle \tilde{A}_{l_1 l_2}^{LM} \rangle_{\text{ub.cmb}}$ is the expected bias due to mask and spatially varying noise in an observed map. The expected bias is estimated from an ensemble of simulations

which are not Doppler boosted. This naive estimator can be optimized by minimizing its variance arising due to cosmic variance and instrument noise. The derivation of the estimator is discussed in Appendix B. The estimator used to reconstruct the Doppler boost vector is given by

$$\hat{\beta}_{LM} = \sum_{l_1 l_2} \hat{w}_{l_1 l_2}^L \frac{\hat{\mathcal{A}}_{l_1 l_2}^{LM}}{K_{LM l_1 l_2}^{LM}}, \qquad (21)$$

where $\hat{\mathcal{A}}_{l_1 l_2}^{LM} = \tilde{A}_{l_1 l_2}^{LM} - \langle \tilde{A}_{l_1 l_2}^{LM} \rangle_{\text{ub.cmb}}$ are the bias-corrected BipoSH coefficients. The weights $\hat{w}_{l_1 l_2}^L$ which minimize the variance are given by the expression

$$\hat{w}_{l_1 l_2}^L = \frac{1}{\sum_M \left(\frac{\hat{\sigma}_{l_1 l_2}^{LM}}{K_{LM l_1 l_2}^{LM}} \right)^2} \left[\sum_{l_1' l_2'} \frac{1}{\sum_M \left(\frac{\hat{\sigma}_{l_1' l_2'}^{LM}}{K_{LM' l_1' l_2'}^{LM}} \right)^2} \right]^{-1}, \qquad (22)$$

where

$$(\hat{\sigma}_{l_1 l_2}^{LM})^2 = \langle |\tilde{A}_{l_1 l_2}^{LM}|^2 \rangle_{\text{ub.cmb}} - |\langle \tilde{A}_{l_1 l_2}^{LM} \rangle_{\text{ub.cmb}}|^2 \qquad (23)$$

is the variance of the unboosted map's BipoSH coefficients.

We see that the effective weights are M dependent owing to the M-dependent shape function in Eq. (21), unlike in the full sky estimator. Since the estimator is weighed by the modified shape function which accounts for loss of power due to the mask, it naturally corrects for the reduced amplitude of the Doppler vector due to masking.

V. DEMONSTRATION OF THE WVMSFE

We demonstrate the working of this newly proposed method for reconstructing the Doppler boost, by evaluating it on an ensemble of Doppler-boosted simulations. We also run a parallel analysis on full skies to allow direct comparison of its efficiency.

A. Generating simulations

We generated a set of 1000 simulations with characteristics of the Planck 217 GHz instrument using the best fit theoretical C_l corresponding to the cosmological parameters from Planck 2013 data [25]. The isotropic simulations were generated using the SYNFAST facility of HEALPix [26,27]. The Doppler-boosted simulations are generated using the *Code for Non-Isotropic Gaussian Sky* (CoNIGS) algorithm [28]. We injected Doppler boost with amplitude $|\vec{\beta}| = 1.23 \times 10^{-3}$, pointing towards the galactic coordinates $(l, b) = (264°, 48°)$. Since these simulations are specific to the 217 GHz channel, we set the frequency-dependent modulation factor to $b_\nu = 3$.

The noise simulations were generated separately and added to the CMB simulations. In our analysis, we first worked with isotropic noise approximation and then with

realistic spatially varying noise, so as to progressively increase the complexity of the data. The isotropic noise simulations were generated using a white Gaussian noise spectrum corresponding to the 217 GHz instrument ($\theta_{fwhm} = 5$ arcmin and $\sigma = 4.8$ μK/K), as quoted in the Planck Bluebook [29]. To simulate the full mission (five surveys) noise levels, we divide the nominal (two surveys) noise standard deviation by $\sqrt{5/2}$.

The realistic, spatially varying noise maps were generated using the noise variance map (no cross-pixel correlations) corresponding to Planck 217 GHz channel available in the public archives [30].

B. Mask and modified shape function

We use the common analysis mask used in the Planck analysis shown in Fig. 2. Sharp $\{0, 1\}$ masks are not ideal for any band-limited analysis as they result in heavy ringing at the edges of the mask. To avoid this unnecessary complication we apodize the mask with a Gaussian beam of $\theta_{fwhm} = 30$ arcmin. The apodized mask has an effective sky fraction of $f_{sky} \approx 0.78$ which is practically the same as the sky fraction available with the unapodized mask.

Given the apodized mask and the fiducial power spectrum used to generate the simulations, we numerically evaluate the modified shape function $K_{LMl_1 l_2}^{LM}$ using the expression in Eq. (18). The $3j$ and $6j$ symbols are calculated using standard routines available in the SLATEC numerical library [31]. The modified shape functions evaluated and used in our analysis are depicted in Fig. 1.

C. Analysis and results

In this section we test the recovery of the Doppler boost vector from simulated data using the new estimator presented in Sec. IV B. We perform the analysis using the full sky simulations in addition to the analysis on masked maps to allow direct comparison.

For the masked analysis, we first evaluate the ensemble-averaged BipoSH coefficients and their variance from the unboosted, noise-added, masked simulations. While the average is used to subtract the biases due to the mask and spatially varying noise, the variance is used in the evaluation of the minimum variance estimator in Eq. (21). The Doppler boost estimator is evaluated on each realization of Doppler-boosted simulations, yielding an ensemble of estimates of the Doppler field $\beta(\hat{n}) = \vec{\beta} \cdot \hat{n}$. While the direction of the Doppler boost is determined using the HEALPix subroutine REMOVE_DIPOLE, the Doppler boost amplitude is recovered by first estimating the power in the reconstructed Doppler field: $|\vec{\beta}| = 1.5\sqrt{\beta_1/\pi}$, where $\beta_1 = \sum_M |\beta_{1M}|^2/3$.

The Doppler power is not expected to vanish when estimated from an unboosted sky, and this bias is termed reconstruction noise. In order to have an unbiased estimate of the Doppler amplitude, the reconstruction noise β_1^N
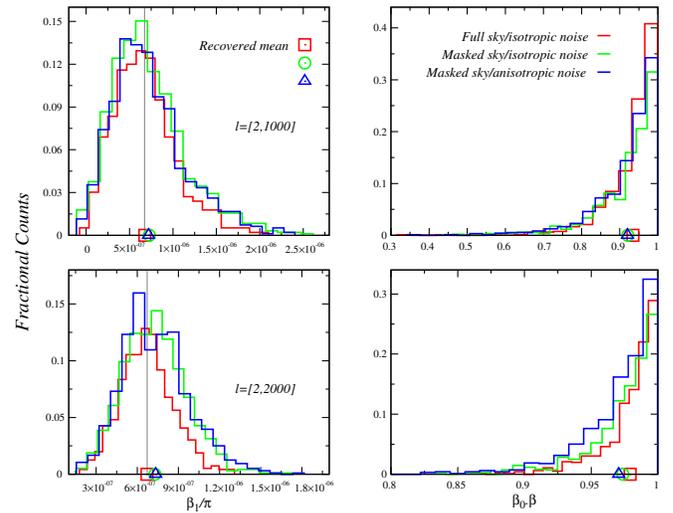

FIG. 3 (color online). This figure depicts the recovered Doppler boost parameters from analysis on simulated CMB skies masked with the apodized Planck common mask. While the left column depicts the histogram of the recovered Doppler power corrected for reconstruction noise bias, the right column depicts the histogram of the cosine of angular separation $\hat{\beta}_0 \cdot \hat{\beta}$, where $\hat{\beta}_0$ and $\hat{\beta}$ denote the injected and recovered direction, respectively. The top and bottom row denote the results from analysis in the multipole bins $l = [2, 1000]$ and $[2, 2000]$, respectively. The injected Doppler power is denoted by a vertical gray line.

needs to be subtracted from the estimated Doppler power. This bias can be dealt with semianalytically in the full sky case (see Appendix A). For the masked analysis we estimate the mean reconstruction noise by applying the Doppler estimator on a set of masked unboosted simulations. The unbiased estimate of the Doppler boost amplitude is given by $|\vec{\beta}| = 1.5\sqrt{(\beta_1 - \beta_1^N)/\pi}$. The full sky analysis follows the same procedure as above, except that we do not have to bias subtract the BipoSH coefficients, since we only run the full sky analysis on simulations with isotropic noise. Note that one still needs to bias subtract the power in the Doppler field to arrive at the correct Doppler boost amplitude.

Finally we run the analysis on two multipole ranges, [2, 1000] and [2, 2000]. At high multipoles $l \gtrsim 1000$, the results can be potentially affected by point source masks. By demonstrating robustness against a range of multipoles used in the analysis, we have shown that our results are not affected by the presence of point source masks.

Results of the reconstruction of Doppler boost vector are summarized in Fig. 3. Table II lists the mean recovered Doppler amplitude and direction from different multipole bins. Note that the injected amplitude and direction are consistently recovered from the masked as well as full sky simulations with isotropic and anisotropic noise. We find that the error on the recovered Doppler amplitude and direction is larger when using smaller multipole range $l \in [2, 1000]$ as compared to estimates from using the larger

TABLE II. The mean recovered Doppler amplitude and direction obtained from 1000 simulations are tabulated here. The boosted simulations were generated with a Doppler boost amplitude $|\vec{\beta}| = 1.23 \times 10^{-3}$ pointing towards the galactic coordinates $(l, b) = (264°, 48°)$. The errors quoted are the standard deviation computed from probability density function inferred from the ensemble of reconstructed Doppler power and direction (l, b).

Mask	Δl	$|\beta|$ ($\times 10^{-3}$)	$b°$	$l°$
Full sky	[2, 1000]	1.23 ± 0.33	46 ± 14	266 ± 26
(Isotropic noise)	[2, 2000]	1.23 ± 0.18	47.7 ± 8.2	266 ± 13
Common mask	[2, 1000]	1.28 ± 0.38	45 ± 15	262 ± 32
(Isotropic noise)	[2, 2000]	1.28 ± 0.20	46.6 ± 9.2	262 ± 14
Common mask	[2, 1000]	1.27 ± 0.38	44 ± 15	263 ± 30
(Anisotropic noise)	[2, 2000]	1.28 ± 0.21	46.2 ± 9.2	262 ± 15

multipole range $l \in [2, 2000]$ as expected. Similarly the error on estimates from the masked sky is larger than the error on estimate from the full sky, as expected due to the reduced sky fraction owing to the mask. Finally we reiterate that we do not have to perform any f_{sky} corrections to the Doppler amplitude as is usually required by other methods.

VI. CONCLUSIONS

For any reliable cosmological analysis, masking the CMB maps is inevitable to avoid biases due to galactic foregrounds. The effects of masking on the angular power spectrum are easy to reverse using the MASTER algorithm [24].

Here we developed a formalism similar to the MASTER algorithm but extended to all BipoSH coefficients which are a generalization of the well-known angular power spectrum. Using this formalism, we derived an expression which describes how masking modifies an arbitrary isotropy-violating signal. Though we write down the mask coupling matrix for the BipoSH coefficients, we do not try to invert the equation as done in the MASTER algorithm. We simplified the equation using the symmetry properties of the mask and properties of Clebsch-Gordan coefficients.

As a very specific example, we studied how the Doppler boost signal is modified due to the presence of the mask. We showed that by making reasonable assumptions it is possible to cast the BipoSH coefficients of a masked Doppler-boosted sky in a form similar to the BipoSH coefficients of Doppler-boosted full CMB sky, where the full sky shape function is replaced by the modified shape function. Eventually we obtained an estimator which recovers the Doppler signal from a masked CMB sky. This estimator naturally accounts for the loss of power due to masking and hence does not require an additional f_{sky} correction as required by other quadratic estimators.

We generated a set of simulations whose statistical properties are tailored to match the 217 GHz Planck maps. We used an apodized form of Planck 2015 common mask for our analysis. Finally we evaluated the newly derived estimator on masked CMB skies to demonstrate its unbiased recovery of the injected signal.

The Doppler amplitude from Planck 2013 data [22] is estimated as projected Doppler amplitude along the known dipole direction to obtain $\beta_{\parallel} = 0.75 \pm 0.19$ (in units of the expected amplitude $= 1.23 \times 10^{-3}$) from the 217 GHz observed map. To allow for a direct comparison with this estimate, we translate our $\vec{\beta}$ estimates from our current simulations to projected Doppler amplitude to find $\beta_{\parallel} = 1.0 \pm 0.18$. While we use the same multipole range $l = [500, 2000]$ as was used in the Planck 2013 analysis and use the Planck noise model, we however use a different mask. Note that, at face value, we get a slightly better error bar. For a robust, direct comparison between the two analysis methods, we would need to run our analysis on the exact same data used for the Planck 2013 analysis.

While it is difficult to draw any strong conclusions with this information, we find that our method is comparable to, if not better than, the previous method used to estimate the Doppler boost signal. Further this method has also been used to estimate the dipole modulation signal from Planck 2015 data [3], demonstrating its versatility.

Finally we note that the method developed here is generic and can be used to reconstruct any arbitrary form of SI violation [32,33].

APPENDIX A: FULL SKY STATISTIC

Many isotropy-violating signals can be written in the generic form [8]

$$\hat{A}^{LM}_{l_1 l_2} = A^{LM}_{l_1 l_2} + \beta_{LM} H^L_{l_1 l_2}, \tag{A1}$$

where $\hat{A}^{LM}_{l_1 l_2}$ denote the BipoSH coefficients measured from the data, $A^{LM}_{l_1 l_2}$ denote the BipoSH coefficients of the statistically isotropic CMB sky, β_{LM} denote the parameters

of the isotropy-violating field, and $H^L_{l_1 l_2}$ denotes the spectral shape function induced by the particular form of isotropy violation.

For a statistically isotropic CMB field, $\langle A^{LM}_{l_1 l_2} \rangle = 0$ for $L > 0$. Following Eq. (A1), we can define a *naive* estimator for the parameters β_{LM} as follows:

$$\hat{\beta}_{LM} = \frac{1}{\sum_{l_1 l_2}} \sum_{l_1 l_2} \frac{\hat{A}^{LM}_{l_1 l_2}}{H^L_{l_1 l_1}}. \quad (A2)$$

However, the above estimator is not optimized to reduce the variance due to noise. It is possible to recover a signal with a better signal to noise ratio by redefining the estimator as

$$\hat{\beta}_{LM} = \sum_{l_1 l_2} \hat{w}^L_{l_1 l_2} \frac{\hat{A}^{LM}_{l_1 l_2}}{H^L_{l_1 l_1}}, \quad (A3)$$

where the weights $\hat{w}^L_{l_1 l_2}$ are to be chosen so as to minimize the quantity

$$\hat{C}^{\beta\beta}_L = \langle \hat{\beta}_{LM} \hat{\beta}^*_{LM} \rangle \quad (A4a)$$

$$= \sum_{l'_1 l'_2} (\hat{w}^L_{l'_1 l'_2})^2 \frac{2 C_{l'_1} C_{l'_2}}{(H^L_{l'_1 l'_2})^2} + C^{\beta\beta}_L \quad (A4b)$$

$$= N_L + C^{\beta\beta}_L, \quad (A4c)$$

where N_L is the reconstruction noise that is minimized by choosing appropriate weights and, $\hat{C}^{\beta\beta}_L$ and $C^{\beta\beta}_L$ denote the reconstructed and true power in the multipole L, respectively, of any isotropy-violating field $\beta(\hat{n})$. Minimizing the reconstruction noise results in down weighting the noisy modes while giving more weightage to the signal-dominated modes. In arriving at Eq. (A4b) we used the covariance of BipoSH coefficients given by [5]

$$\langle A^{LM}_{l_1 l_2} A^{L'M'*}_{l'_1 l'_2} \rangle = C_{l_1} C_{l_2} [\delta_{l_1 l'_1} \delta_{l_2 l'_2} + \delta_{l_1 l'_2} \delta_{l'_1 l_2}] \delta_{LL'} \delta_{MM'}, \quad (A5)$$

which is valid when $L \neq 0$ and $L + l_1 + l_2$ is even.

This minimization of power is evaluated subject to the constraint $\sum_{l_1 l_2} \hat{w}^L_{l_1 l_2} = 1$, resulting in a constrained minimization problem, which is solved using the method of Lagrange multipliers. The weights are determined by minimizing the function

$$\mathcal{L} = \hat{C}^{\beta\beta}_L - \alpha \left[\sum_{l_1 l_2} \hat{w}^L_{l_1 l_2} - 1 \right], \quad (A6)$$

where α is the Lagrange multiplier. Note that the weights for each BipoSH mode are assumed to be independent of each other in the following sense:

$$\frac{\partial w^L_{l_1 l_2}}{\partial w^{L'}_{l'_1 l'_2}} = \delta_{LL'} \delta_{l_1 l'_1} \delta_{l_2 l'_2}. \quad (A7)$$

On setting the derivative of $\mathcal{L}$ with respect to the weights $\hat{w}^{L'}_{l'_1 l'_2}$ to zero and simplifying the resultant equation it can be shown that the weights that minimize the reconstruction noise are given by

$$\hat{w}^L_{l_1 l_2} = \frac{(H^L_{l_1 l_2})^2}{C_{l_1} C_{l_2}} \left[\sum_{l'_1 l'_2} \frac{(H^L_{l'_1 l'_2})^2}{C_{l'_1} C_{l'_2}} \right]^{-1}. \quad (A8)$$

Finally the reconstruction noise of the minimum variance estimator is given by

$$N_L = \left[\sum_{l_1 l_2} \frac{(H^L_{l_1 l_2})^2}{2 C_{l_1} C_{l_2}} \right]^{-1}. \quad (A9)$$

APPENDIX B: PARTIAL SKY STATISTIC

We follow a procedure similar to the one described in Appendix A, to arrive at a minimum variance estimator in the presence of a mask. Following Eq. (19), we can define a minimum variance estimator for any isotropy-violating field $\beta(\hat{n})$ as

$$\hat{\beta}_{LM} = \sum_{l_1 l_2} \hat{w}^L_{l_1 l_2} \frac{\hat{\mathcal{A}}^{LM}_{l_1 l_2}}{K^{LM}_{LM l_1 l_2}}, \quad (B1)$$

where $\hat{\mathcal{A}}^{LM}_{l_1 l_2} = \hat{A}^{LM}_{l_1 l_2} - \langle A^{LM}_{l_1 l_2} \rangle$, $\hat{A}^{LM}_{l_1 l_2}$ are the BipoSH coefficients measured from masked data map and $\langle A^{LM}_{l_1 l_2} \rangle$ denotes the bias due to all known systematic effects in the data like masking, anisotropic noise, noncircular beam, etc., estimated from masked simulated skies which incorporate all these systematic effects, other than the isotropy violating phenomena in question. It is nontrivial to evaluate the covariance of the BipoSH coefficients $\hat{\mathcal{A}}^{LM}_{l_1 l_2}$ analytically, as can be done in the case of ideal full sky CMB. So we approximate the BipoSH covariance to be diagonal, i.e.

$$\langle \hat{\mathcal{A}}^{LM}_{l_1 l_2} \hat{\mathcal{A}}^{LM*}_{l'_1 l'_2} \rangle_{\text{ub.cmb}} \approx (\hat{\sigma}^{LM}_{l_1 l_2})^2 \delta_{l_1 l'_1} \delta_{l_2 l'_2}, \quad (B2)$$

where $(\hat{\sigma}^{LM}_{l_1 l_2})^2$ is the variance of BipoSH coefficients which is estimated from simulations. This approximation has been verified using simulations. We find that the off-diagonals are at a few percent level or smaller compared to the diagonal terms. Note that the covariance of BipoSH coefficients have an explicit M dependence which was absent in the full sky BipoSH covariance given in Eq. (A5). We also assume that the weights for the different BipoSH modes are independent of each other as in Eq. (A7). Using these approximations and following the same procedure discussed in Appendix A, one finds that the weights that minimize the variance of the estimator are given by

$$\hat{w}^L_{l_1 l_2} = \frac{1}{\sum_M \left(\frac{\hat{\sigma}^{LM}_{l_1 l_2}}{K^{LM}_{LM l_1 l_2}}\right)^2} \left[\sum_{l'_1 l'_2} \frac{1}{\sum_M \left(\frac{\hat{\sigma}^{LM}_{l'_1 l'_2}}{K^{LM}_{LM l'_1 l'_2}}\right)^2}\right]^{-1}. \tag{B3}$$

Note that even though the weights themselves do not have an explicit M dependence, the effective weighting of the BipoSH coefficients inferred from data is M dependent owing to the presence of the modified shape function [see Eq. (B1)].

[1] C. Bennett *et al.*, Astrophys. J. Suppl. Ser. **192**, 17 (2011).

[2] P. A. R. Ade *et al.*, Astron. Astrophys. **571**, A23 (2014).

[3] P. A. R. Ade *et al.*, arXiv:1506.07135.

[4] A. Hajian and T. Souradeep, Astrophys. J. **597**, L5 (2003).

[5] N. Joshi, A. Rotti, and T. Souradeep, Phys. Rev. D **85**, 043004 (2012).

[6] A. Hajian, T. Souradeep, and N. Cornish, Astrophys. J. **618**, L63 (2005).

[7] D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskii, *Quantum Theory of Angular Momentum* (World Scientific, Singapore, 1988).

[8] S. Kumar, A. Rotti, M. Aich, N. Pant, S. Mitra, and T. Souradeep, Phys. Rev. D **91**, 043501 (2015).

[9] T. Ghosh, A. Hajian, and T. Souradeep, Phys. Rev. D **75**, 083007 (2007).

[10] N. Joshi, S. Jhingan, T. Souradeep, and A. Hajian, Phys. Rev. D **81**, 083012 (2010).

[11] N. Joshi, S. Das, A. Rotti, S. Mitra, and T. Souradeep, arXiv:1210.7318.

[12] J. C. Mather, D. J. Fixsen, R. A. Shafer, C. Mosier, and D. T. Wilkinson, Astrophys. J. **512**, 511 (1999).

[13] D. J. Fixsen, Astrophys. J. **707**, 916 (2009).

[14] A. Kogut *et al.*, Astrophys. J. **419**, 1 (1993).

[15] D. J. Fixsen, E. S. Cheng, J. M. Gales, J. C. Mather, R. A. Shafer, and E. L. Wright, Astrophys. J. **473**, 576 (1996).

[16] A. Challinor and F. van Leeuwen, Phys. Rev. D **65**, 103001 (2002).

[17] L. Amendola, R. Catena, I. Masina, A. Notari, M. Quartin, and C. Quercellini, J. Cosmol. Astropart. Phys. 07 (2011) 027.

[18] J. Chluba, Mon. Not. R. Astron. Soc. **415**, 3227 (2011).

[19] A. Kosowsky and T. Kahniashvili, Phys. Rev. Lett. **106**, 191301 (2011).

[20] A. Notaria and M. Quartin, J. Cosmol. Astropart. Phys. 02 (2012) 026.

[21] S. Mukherjee, A. De, and T. Souradeep, Phys. Rev. D **89**, 083005 (2014).

[22] N. Aghanim *et al.*, Astron. Astrophys. **571**, A27 (2014).

[23] S. Adhikari, Mon. Not. R. Astron. Soc. **446**, 4232 (2015).

[24] E. Hivon, K. M. Gorski, C. B. Netterfield, B. P. Crill, S. Prunet, and F. Hansen, Astrophys. J. **567**, 2 (2002).

[25] P. A. R. Ade *et al.*, Astron. Astrophys. **571**, A16 (2014).

[26] K. M. Gorski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann, Astrophys. J. **622**, 759 (2005).

[27] http://healpix.sourceforge.net/.

[28] S. Mukherjee and T. Souradeep, Phys. Rev. D **89**, 063013 (2014).

[29] Planck Collaboration, Report No. ESA-SCI(2005)01, 2005.

[30] http://irsa.ipac.caltech.edu/Missions/planck.html.

[31] http://www.netlib.org/slatec/index.html.

[32] Y. Akrami, Y. Fantaye, A. Shafieloo, H. K. Eriksen, F. K. Hansen, A. J. Banday, and K. M. Górski, Astrophys. J. Lett. **784**, L42 (2014).

[33] P. A. R. Ade *et al.*, arXiv:1502.01591.