Peer Review File

Morphometry on the sphere: Cartesian and irreducible Minkowski tensors explained and implemented

Reviewers' comments:

Reviewer #1 (Remarks to the Author):

The report is submitted as a pdf file. Please see attachment.

Reviewer #2 (Remarks to the Author):

In this manuscript the authors discuss a Minkowski tensor-based analysis method to investigate morphological properties of random fields on a sphere. Previous work in this direction was limited to Euclidean space – here the authors generalise the approach to spherical geometry. After demonstrating their method on some basic geometric shapes, the authors use it analyse Planck's cosmic microwave background temperature map and identify two potentially anomalous regions.

Given that the cosmic microwave background is arguably the best probe of the physics of the early Universe, new and innovative approaches to extracting information from it are certainly of relevance. The analysis is sound and the paper is well written and clearly presented. I also appreciate that the authors have released their code to the community in the form of an open software package – this is useful for extending its reach to applications in other disciplines (e.g., earth or climate science) as well.

In summary, I congratulate the authors on this very interesting paper. In my opinion, their work is clearly suitable for publication.

A few detailed comments:

1) In lines 145-146 (page 5), the authors correctly point out that the Healpix pixelisation scheme has vertices where only three edges meet (if I remember correctly this only applies to eight vertices, independent of the resolution), and mention that this "is taken into account accordingly". I am wondering if a bit more detail about how precisely this was handled would be helpful for the reader.

2) The percentile maps in Figure 5 are useful for identifying potentially anomalous regions, but by themselves the percentiles values cannot be interpreted in terms of a "statistical significance" of these features. Note that the authors do not make any claims along those lines, but an inattentive reader might misinterpret the meaning of the maximum percentiles or "sigmas".

While the pixels in the Minkowski map are correlated, there are still a few hundred statistically independent regions, and it's probably not too surprising to find a few pixels in the top 99% after all. A proper significance analysis in the frequentist sense would need to account for the "look-elsewhere-effect", i.e., correct for the fact that the most extreme pixels were selected only after having seen the data. Effectively, it would be nice to know if the Planck Minkowski map as a whole is unusual compared to the simulations.

I won't strongly insist on this point, so treat the following merely as a suggestion: I believe that with their tools and the FFP10 simulation maps at hand, it would not be too difficult for the authors to perform such a significance analysis. A simple example for this could involve, e.g., creating Minkowski percentile maps akin to the ones in Figure 5 for all the FFP10 maps, identifying the most extreme percentile value for each of them, and then evaluating the fraction of FFP10 MM percentile maps that have more extreme maxima than the Planck MM percentile map.

3) If the detected anomalies were to be statistically significant, are there any physical models that could lead to such features?

4) Colour scales should be added to Figures 1 and 2 of the supplementary material (and perhaps the authors might consider using different colour schemes for the temperature map on the left and the percentile deviation maps to avoid confusion).

5) There's a minor typographical error in line 92 (page 3): tensors -> tensor.

Reviewer #3 (Remarks to the Author):

In this paper, the authors have developed an open-source software called litchi, which computes the Minkowski tensors (MT) on the sphere. A novelty of this code is to compute the MT on the sphere with constant curvature using the famous HEALPix grid. This paper describes some details of how and what the code calculates. The code is applied to the Planck data of the Cosmic Microwave Background (CMB) as a test case, and some results are illustrated in this paper.

The MT is a generalization of the Minkowski Functionals (MF), and is recently interested as a new statistical way to analyze cosmological data such as galaxy distributions and the CMB, focusing on morphological structures of cosmological fluctuation fields. Therefore, I think this paper should be timely and helpful for those who are looking for new methods not only for cosmological analysis, but also for general analysis of data on the sphere.

I think the paper is well written and I can recommend this paper for publication in Communications Physics.

Report on COMMSPHYS-24-0165

The work presented in this manuscript consists of the following. First the authors extend the mathematical formalism of Minkowski tensors (MTs) on the sphere S^2 which was earlier introduced by other authors in ref [26]. In particular, their extension focuses on introducing irreducible Minkowski tensors and 4th rank MTs on S^2. Both these extensions were previously introduced for flat space MTs by one or more of the current list of authors, and what is new here is their generalization to S^2. Secondly, the authors implement a marching squares algorithm on the S^2 to identify structures defined by chosen field levels and define the so-called Minkowski maps. Thirdly, they apply their methodology on CMB temperature data from Planck to search for anomalous regions, finding two such regions. They perform the analysis for different datasets at different scales and using different schemes for estimating the significance of the deviation of observed data from simulations, after which one region persists as significantly anomalous.

The methodology developed is interesting and will be useful for applications in cosmology and other areas. Their numerical implementation code "litchi" that they have made publicly available is potentially of good use for analysis of currently available and upcoming large cosmological data sets. The paper is interesting and sufficiently novel, and hence worth considering for publication. However, I have the following questions/comments which can be addressed by the authors first.

Major comments

1. Regarding MTs of rank higher than 2:
   (a) Introduction section, last two paras: the authors mention "arbitrary rank" at a few places. However, the paper describes up to only 4th rank tensors. Please clarify.
   (b) More importantly, a short description of what exactly is the new information gained from rank 4 MT compared to rank 2 is needed.
2. Section on IMT:
   (a) I find the description in this section unclear. A reader has to refer to refs [36-38] to get a clearer understanding. To make the paper self-contained, the authors can either have all definitions and explanation of physical meaning in the same place, or refer to the most relevant ref and point our the specific section and/or figure which conveys the physical meaning clearly.
   (b) "...powerful characterization of spherical anisotropy, as demonstrated below." . Please explain specifically what makes it "powerful". A comparison with some other measures of anisotropy will also be helpful.
3. Section on Calculating Minkowski maps
   (a) para 1: "The body to be described by the tensors is obtained by choosing a brightness threshold, treating everything below as part of the body." I am guessing that the word 'everything' refers to 'all pixels'. The usual convention found in the literature is to define the body as all pixels above the threshold. While this choice does not matter for Gaussian fields (or those with symmetric PDFs), it does matter for fields with skewed PDFs. Please explain the choice.
   (b) para 2: "....only three pixels meet in a triangular shape, which is taken into account accordingly". One such example is visible in Fig. 2 (b) above left of the marching square indicated, and another on the right closer to the boundary of the image. These could also be indicated or mentioned in words. I guess there are only 8 such triangular configurations and the number can be mentioned. Further, the description "taken into account accordingly" does not explain how such shapes are handled. Please clarify.

4. Section MT analysis of Planck data:
   (a) para 2: "All thresholds are combined in each MM...". What is the procedure used to combine thresholds?
   (b) How many simulations are used for Commander pipeline?
   (c) para 7: "...600 simulations". Are these SMICA simulations? If yes, why not 999 mentioned above?
   (d) Even though the authors mention that understanding the origin of the anomalous nature is beyond the scope of the paper, it will be helpful if some possible causes can be mentioned. Can this be related to being broadly in the direction of the ecliptic poles?
5. Methods: Basics of Planck data analysis: According to the description of subroutine ud_grade in Healpix site, downgrading a map is done by averaging map values over associated nested pixels. It is mentioned that in this paper what is followed is not the simple averaging, but convolution with suitable beam size in harmonic space. Since the referred paper ref [20] is quite a detailed paper it will be useful for those who wish to reproduce the results of this paper if brief details of the downgrading process including convolution FWHMs are provided.
6. References: In the introduction section I find some references inappropriately cited, while some important ones are missing. I think the authors will find it beneficial to carry out careful check of what has been studied in the literature.
   (a) I find that ref [26] is where MTs on S^2 were first introduced. However, this ref is incorrectly cited along with [25] as "Rank 2 MT have been used by [25,26], finding consistency with Gaussianity via there approach". Please correct this. Also ref [25] is incorrectly described. It deals with statistical isotropy and not Gaussianity. Some of the authors of ref [26] have also discussed MT on S^2 in S. Appleby et al, ApJ 942, 110 (2023).
   (b) Derivation of analytic formulas for ensemble expectations of Minkowski functionals for non-Gaussian fields has been an important theoretical advancement. In this connection the following refs may be added: T. Matsubara and S. Kuriki, Phys. Rev. D 104, 103522 (2021), and C. Gay, C. Pichon, and D. Pogosyan, PRD 85, 023011 (2012).

Minor comments and suggestions

1. Fig 2:
   (a) It may be clearer to show panel (b) on the left by interchanging with (a).
   (b) Panel (b) is not referenced in the text though Healpix sphere is mentioned.
   (c) In the text it may also be clearer to refer to the panels as (a), (b), (c) rather than left, middle, right.
   (d) I think it is better to use "angular distance" instead of "distance" when referring to α and d.
2. I suggest combining Figs. 3 and 4 so that a reader an visually compare them easily.
3. MT analysis for Planck data, para 3 onward: It is a bit disorienting to go back and forth and remember (l, b) values for the two anomalous spots and the cold spot. Particularly, Fig. 7 gives only the (l, b) values. I suggest mentioning the position values for (l_1, b_1), (l_2, b_2) and the cold spot only once at the beginning, and subsequently refer only as (l_1, b_1), (l_2, b_2) and cold spot. However, in Fig. 7 both names and values can be given for ease of reading.

Response to the referees

Reviewer # 1:

The work presented in this manuscript consists of the following. First the authors extend the mathemati cal formalism of Minkowski tensors (MTs) on the sphere S2 which was earlier introduced by other authors in ref [26]. In particular, their extension focuses on introducing irreducible Minkowski tensors and 4th rank MTs on S2. Both these extensions were previously introduced for flat space MTs by one or more of the current list of authors, and what is new here is their generalization to S2. Secondly, the authors implement a marching squares algorithm on the S2 to identify structures defined by chosen field levels and define the so-called Minkowski maps. Thirdly, they apply their methodology on CMB temperature data from Planck to search for anomalous regions, finding two such regions. They perform the analysis for different datasets at different scales and using different schemes for estimating the significance of the deviation of observed data from simulations, after which one region persists as significantly anomalous. The methodology developed is interesting and will be useful for applications in cosmology and other areas. Their numerical implementation code “litchi” that they have made publicly available is potentially of good use for analysis of currently available and upcoming large cosmological data sets. The paper is interesting and sufficiently novel, and hence worth considering for publication. However, I have the following questions/comments which can be addressed by the authors first.

We thank the referee for the positive assessment of our work and sincerely appreciate the referee’s insightful comments and thoughtful feedback, which enabled us to enhance the clarity and quality of our work. All points are addressed in detail below.

Major comments

1. Regarding MTs of rank higher than 2:

(a) Introduction section, last two paras: the authors mention “arbitrary rank” at a few places. However, the paper describes up to only 4th rank tensors. Please clarify.

Our methods permit the calculation of any rank, and we chose to use ranks 2 and 4 for demonstration purposes where IMT and cartesian MT can be compared. We have clarified this choice in the text. Note that, in the Cartesian case, these are (to our knowledge) the only ranks accessible to anisotropy analysis.

(b) More importantly, a short description of what exactly is the new information gained from rank 4 MT compared to rank 2 is needed.

We thank the referee for this suggestion. A sentence has been added in the rank 4 subsection of the test cases: ”While rank 2 MT encode information about elongation (2-fold symmetry), rank 4 MT contain information about 4-fold symmetry (’square-ness’).” We have, moreover, added specific references in the section on IMTs, as suggested by the referee in the next comment.

2. Section on IMT:

(a) I find the description in this section unclear. A reader has to refer to refs [36-38] to get a clearer understanding. To make the paper self-contained, the authors can either have all definitions and explanation of physical meaning in the same place, or refer to the most relevant ref and point our the specific section and/or figure which conveys the physical meaning clearly.

We agree. This section was indeed a bit intransparent to readers from other fields. We now explicitly point out Figs. 1 and 2 in ref 37 for simple demonstrations of the IMT and an intuitive understanding of the geometrical information that they capture.

(b) "...powerful characterization of spherical anisotropy, as demonstrated below." . Please explain specifically what makes it "powerful". A comparison with some other measures of anisotropy will also be helpful.

MT and MF are sensitive to information not visible with other methods, such as mean intercept length or wavelets; relevant references have been added to the introduction. IMT in particular are more simple to use as compared to Cartesian MT at any rank higher than 2, as explained in the paper.

3. Section on Calculating Minkowski maps
(a) para 1: "The body to be described by the tensors is obtained by choosing a brightness thresh- old, treating everything below as part of the body." I am guessing that the word 'everything' refers to 'all pixels'. The usual convention found in the literature is to define the body as all pixels above the threshold. While this choice does not matter for Gaussian fields (or those with symmetric PDFs), it does matter for fields with skewed PDFs. Please explain the choice.

Our choice has indeed been described wrongly in the text. We have now corrected the corresponding descriptions both in this section and in the caption of Fig. 2. In the code, pixels above the threshold have always been interpreted as part of the body. We thank the referee for catching this mistake.

(b) para 2: "....only three pixels meet in a triangular shape, which is taken into account accordingly". One such example is visible in Fig. 2 (b) above left of the marching square indicated, and another on the right closer to the boundary of the image. These could also be indicated or mentioned in words. I guess there are only 8 such triangular configurations and the number can be mentioned. Further, the description "taken into account accordingly" does not explain how such shapes are handled. Please clarify.

There are indeed only 8 such positions, independent of the resolution. They are treated identically (by interpolating contours between pixels above/below threshold) but need to be distinguished in the code when the exact number of pixels in a window is relevant. This procedure has been clarified in the text.

4. Section MT analysis of Planck data:
(a) para 2: "All thresholds are combined in each MM...". What is the procedure used to combine thresholds?

This technique is mentioned in the last paragraph of the section on Minkowski maps; the tensors calculated for the contours at various thresholds are added up. The text has been adapted slightly in both sections and is hopefully more clear now.

(b) How many simulations are used for Commander pipeline?

We thank the referee for this question. Due to an oversight this number was previously not stated. We used 999 simulations as for SMICA and have added this to the text.

(c) para 7: "...600 simulations". Are these SMICA simulations? If yes, why not 999 mentioned above?

These are SMICA simulations and we used only 600 due to the higher computation time for the technique in the paragraph. This has been clarified in the text.

(d) Even though the authors mention that understanding the origin of the anomalous nature is beyond the scope of the paper, it will be helpful if some possible causes can be mentioned. Can this be related to being broadly in the direction of the ecliptic poles?

There are several models that could lead to anisotropies in general, such as cosmic strings, cosmic topology, or Bianchi models. We have added a corresponding paragraph in the discussion section. Because we only want to point out possible avenues for further research into the anomalies, we do not go into much detail on these models. The ecliptic poles are several window radii away from the regions we discuss, so we do not expect any direct causal connection.

5. Methods: Basics of Planck data analysis: According to the description of subroutine ud grade in Healpix site, downgrading a map is done by averaging map values over associated nested pixels. It is mentioned that in this paper what is followed is not the simple averaging, but convolution with suitable beam size in harmonic space. Since the referred paper ref [20] is quite a detailed paper it will be useful for those who wish to reproduce the results of this paper if brief details of the downgrading process including convolution FWHMs are provided.

Our description of the degrading process has indeed been a bit intransparent. We have now added more details on the degrading process and mention the steps taken in our script. For transforming into spherical harmonics, pixel window functions (provided by the healpy package) are taken into account. FWHMs are now provided for the Nside values used in the analysis.

6. References: In the introduction section I find some references inappropriately cited, while some important ones are missing. I think the authors will find it beneficial to carry out careful check of what has been studied in the literature.
(a) I find that ref [26] is where MTs on S2 were first introduced. However, this ref is incorrectly cited along with [25] as "Rank 2 MT have been used by [25,26], finding consistency with Gaussianity via there approach". Please correct this. Also ref [25] is incorrectly described. It deals with statistical isotropy and not Gaussianity. Some of the authors of ref [26] have also discussed MT on S2 in S. Appleby et al, ApJ 942, 110 (2023).

We thank the referee for this detailed comment. We corrected these statements and added Appleby et al. (2023). These errors may have happened over the course of several iterations of the text in the introduction.

(b) Derivation of analytic formulas for ensemble expectations of Minkowski functionals for non-Gaussian fields has been an important theoretical advancement. In this connection the following refs may be added: T. Matsubara and S. Kuriki, Phys. Rev. D 104, 103522 (2021), and C. Gay, C. Pichon, and D. Pogosyan, PRD 85, 023011 (2012).

We have added those (and more) references about ensemble expectations.

Minor comments and suggestions

1. Fig 2:
(a) It may be clearer to show panel (b) on the left by interchanging with (a).
(b) Panel (b) is not referenced in the text though Healpix sphere is mentioned.
(c) In the text it may also be clearer to refer to the panels as (a), (b), (c) rather than left, middle, right.
(d) I think it is better to use "angular distance" instead of "distance" when referring to α and d.
2. I suggest combining Figs. 3 and 4 so that a reader an visually compare them easily.
3. MT analysis for Planck data, para 3 onward: It is a bit disorienting to go back and forth and remember (l, b) values for the two anomalous spots and the cold spot. Particularly, Fig. 7 gives only the (l, b) values. I suggest mentioning the position values for (l1, b1), (l2, b2) and the cold spot only once at the beginning, and subsequently refer only as (l1, b1), (l2, b2) and cold spot. However, in Fig. 7 both names and values can be given for ease of reading.

We thank the reviewer for this constructive and detailed feedback. We have incorporated all of these suggestions.

Reviewer # 2 :

In this manuscript the authors discuss a Minkowski tensor-based analysis method to investigate morphological properties of random fields on a sphere. Previous work in this direction was limited to Euclidean space – here the authors generalise the approach to spherical geometry. After demonstrating their method on some basic geometric shapes, the authors use it analyse Planck's cosmic microwave background temperature map and identify two potentially anomalous regions.

Given that the cosmic microwave background is arguably the best probe of the physics of the early Universe, new and innovative approaches to extracting information from it are certainly of relevance. The analysis is sound and the paper is well written and clearly presented. I also appreciate that the authors have released their code to the community in the form of an open software package – this is useful for extending its reach to applications in other disciplines (e.g., earth or climate science) as well.

In summary, I congratulate the authors on this very interesting paper. In my opinion, their work is clearly suitable for publication.

We are grateful for this detailed and positive review and the insightful comments. The points raised are addressed individually below.

A few detailed comments:

1) In lines 145-146 (page 5), the authors correctly point out that the Healpix pixelisation scheme has vertices where only three edges meet (if I remember correctly this only applies to eight vertices, independent of the resolution), and mention that this "is taken into account accordingly". I am wondering if a bit more detail about how precisely this was handled would be helpful for the reader.

There are indeed only 8 such vertices. Conceptually these triangular regions are treated the same as those with four corners by interpolating the contours between pixel centers. The number of corners only matters in the implementation of litchi whenever the explicit number of pixels in a window is relevant. This has been clarified in the text.

2) The percentile maps in Figure 5 are useful for identifying potentially anomalous regions, but by themselves the percentiles values cannot be interpreted in terms of a "statistical significance" of these features. Note that the authors do not make any claims along those lines, but an inattentive reader might misinterpret the meaning of the maximum percentiles or "sigmas".

While the pixels in the Minkowski map are correlated, there are still a few hundred statistically independent regions, and it's probably not too surprising to find a few pixels in the top 99% after all. A proper significance analysis in the frequentist sense would need to account for the "look-elsewhere-effect", i.e., correct for the fact that the most extreme pixels were selected only after having seen the data. Effectively, it would be nice to know if the Planck Minkowski map as a whole is unusual compared to the simulations.

I won't strongly insist on this point, so treat the following merely as a suggestion: I believe that with their tools and the FFP10 simulation maps at hand, it would not be too difficult for the authors to perform such a significance analysis. A simple example for this could involve, e.g., creating Minkowski percentile maps akin to the ones in Figure 5 for all the FFP10 maps, identifying the most extreme percentile value for each of them, and then evaluating the fraction of FFP10 MM percentile maps that have more extreme maxima than the Planck MM percentile map.

We thank the referee for this insightful suggestion of a more thorough analysis of the significance of our findings. We have performed a more extended analysis and describe the results in our revised manuscript. For each data MM pixel, we count the number of simulation maps that contain more extreme values. Using this method, (l_1, b_1) *is surpassed by 943 out of the 999 simulation maps and does not seem to be very significant apart from the fact that several neighboring pixels reach percentile values >99% and its location near the Cold Spot. However, one caveat of this method here is that* (l_1, b_1) *does not have the highest pixel value in the data map (that would be* (l_2, b_2)*), but we compare it to the highest values in the simulations. A truly thorough analysis would need to be more complex.* (l_2, b_2) *is only surpassed in 6 simulation maps, making it look a lot more outstanding.*

3) If the detected anomalies were to be statistically significant, are there any physical models that could lead to such features?

There are several models that could lead to anisotropies in general, such as cosmic strings, cosmic topology, or Bianchi models. We have added a paragraph on this in the discussion section. Since we only want to point out possible avenues for further research into the anomalies, we do not go into much detail on any model.

4) Colour scales should be added to Figures 1 and 2 of the supplementary material (and perhaps the authors might consider using different colour schemes for the temperature map on the left and the percentile deviation maps to avoid confusion).

We thank the referee for these helpful suggestions. Colour scales have been added and all subfigures now use different colour schemes.

5) There's a minor typographical error in line 92 (page 3): tensors → tensor.
fixed

Reviewer # 3:

In this paper, the authors have developed an open-source software called litchi, which computes the Minkowski tensors (MT) on the sphere. A novelty of this code is to compute the MT on the sphere with constant curvature using the famous HEALPix grid. This paper describes some details of how and what the code calculates. The code is applied to the Planck data of the Cosmic Microwave Background (CMB) as a test case, and some results are illustrated in this paper.

The MT is a generalization of the Minkowski Functionals (MF), and is recently interested as a new statistical way to analyze cosmological data such as galaxy distributions and the CMB, focusing on morphological structures of cosmological fluctuation fields. Therefore, I think this paper should be timely and helpful for those who are looking for new methods not only for cosmological analysis, but also for general analysis of data on the sphere.

I think the paper is well written and I can recommend this paper for publication in Communications Physics.

We thank the referee for these encouraging comments. We deeply appreciate the constructive feedback and are grateful for the recommendation.

REVIEWERS' COMMENTS:

Reviewer #1 (Remarks to the Author):

The authors have incorporated the comments and suggestions made by the referees. I recommend acceptance for publication.

Reviewer #2 (Remarks to the Author):

In my opinion, the authors have appropriately addressed all of my and the other referee's comments.

I am happy to recommend the manuscript for publication.