Entropy methods for CMB analysis of anisotropy and non-Gaussianity

Momchil Minkov,1‡ Marvin Pinkwart,2§ and Peter Schupp2∗

1Department of Electrical Engineering, Stanford University, 350 Serra Mall, Stanford, CA 94305-9005, USA
2Department of Physics and Earth Sciences, Jacobs University Bremen, Campus Ring 1, 28759 Bremen, Germany

(Dated: September 25, 2018)

In recent years high-resolution CMB measurements have opened up the possibility to explore statistical features of the temperature fluctuations down to very small angular scales. One method that has been used is the Wehrl entropy, which is however extremely costly in terms of computational time. Here, we propose several different pseudo-entropy measures (projection, angular, and quadratic) that agree well with the Wehrl entropy but are significantly faster to compute. All of the presented alternatives are rotationally invariant measures of entanglement after identifying each multipole \( l \) of temperature fluctuations with a spin-\( l \) quantum state, and are very sensitive to non-Gaussianity or anisotropy in the data. We provide a simple proof that the projection pseudo-entropy converges to the Wehrl entropy with increasing dimensionality of the ancilla projection space. Furthermore, for \( l = 2 \), we show that both the Wehrl entropy and the angular pseudo entropy can be expressed as one-dimensional functions of the squared chordal distance of multipole vectors, giving a tight connection between the two measures. We also show that the angular pseudo entropy can clearly distinguish between Gaussian and non-Gaussian temperature fluctuations at large multipoles and henceforth provides a non-brute force method for identifying non-Gaussianities. This allows us to study possible hints of statistical anisotropy and non-Gaussianity in the CMB up to multipole \( l = 1000 \) using Planck 2015, Planck 2018 and WMAP 7-year data. We find that \( l = 5 \) and \( l = 28 \) have a large entropy at \( 3\sigma \) significance and a slight hint towards a connection of this with the Cosmic Dipole. On a wider range of large angular scales we do not find indications of violation of isotropy or Gaussianity. We also find a small scale range, \( l \in [895; 905] \), that is incompatible with the assumptions at \( 3\sigma \)-level, although how much this significance can be reduced by taking into account the selection effect is left as an open question. Furthermore, we find overall similar results in our analysis of the 2015 and the 2018 data. Finally, we also demonstrate how a range of angular momenta can be studied with the range angular pseudo entropy, which measures averages and correlations of different multipoles. We believe that the formalism developed here can underpin future studies of the Gaussianity and isotropy of the CMB and help to identify deviations especially at small angular scales.

Keywords: CMB – data analysis – coherent states – pseudo entropy

I. INTRODUCTION

Since its discovery in 1964 by Penzias and Wilson the Cosmic Microwave Background has served as the main source of information about the current and past Universe. Originating in the process of recombination at about 380,000 years after the Big Bang at redshift of about 1100, the CMB temperature distribution on the celestial sphere displays the energy density distribution. This assumption has been confirmed by the second data release provide the most precise measurements of cosmological parameters[2][3] measured up to order of \( 10^{-5} \) K to \( 10^{-4} \) K and one larger dipole modulation of order \( 10^{-3} \) K which is assumed to be of pure kinematic origin. The anisotropy of the CMB temperature has been measured first by the COBE satellite from 1989 to 1993, followed up by the WMAP mission from 2001 to 2010. The most recent CMB investigation satellite, Planck, was launched in 2009 and shutdown in 2013. Its 2015 results from the second data release provide the most precise values of cosmological parameters[2][3] measured up to this point. Recently, they have been refined in the 2018 data release [1][5].

It is commonly assumed that the tiny temperature fluctuations follow a Gaussian and statistically isotropic distribution. This assumption has been confirmed by the Planck mission to a large extent[6][7], but nevertheless the search for possible non-Gaussianities[8] and statistical anisotropies[9][10][11][12][13][14][15][16][17][18][19] has been rich and certain anomalies have been found, as for example unusual (anti-)correlation of the lowest multipoles with the Cosmic Dipole as well as with each other, a sign of parity asymmetry and a lack of large-angle correlation, see e.g. the review[20].

A common tool in these analyses are multipole vectors which were introduced for cosmological data anal-
y in [21] and whose properties have been elaborated in [10,22,23,24,25]. For the most recent results on possible CMB anomalies using multipoles and an overview over the mathematical approaches see [20,27].

Multipole vectors are closely related to Bloch coherent states, see [10], which were also used in the past to prove special cases of Lieb’s conjecture[28] for the Wehrl entropy.

In this work we develop and compare several rotationally invariant measures of randomness on functions on the two-sphere, namely the angular, projection, and quadratic pseudo entropies. We show that for \( l = 2 \) the Wehrl and angular entropy can be expressed as a function of the squared chordal distance of multipole vectors. We find that all these measures except the quadratic one show the same features, making the quadratic pseudo-entropy the least preferred measure. Because of the shared properties, we then restrict ourselves to the numerically fastest method, and use it to analyze Planck 2015 and 2018 full sky as well as WMAP 7-year ILC maps. The angular pseudo entropy allows for comparing the data to many ensembles of Gaussian and isotropic random maps up to \( l = 1000 \) in short computing time. With a better theoretical understanding of confidence levels, also the Wehrl entropy could be used easily since the computing time for a single map is still reasonable. In general it is especially nice to have a single number for each multipole even in the case that the data would not be Gaussian and isotropic. In this case the CMB would be described by more than one degree of freedom per multipole. Non-Gaussian distributions need higher correlation functions and anisotropic distributions yield an \( m \)-dependent two-point function. In the tradition of thermodynamics, with these pseudo entropies one can approximately reduce a possibly large set of data again to one number for each multipole. Since all considered types of entropies show a similar behavior the information does not depend on the definition of the entropy. Eventually there exists also an extension of the angular entropy to ranges and collections of multipoles, which we call range angular entropy.

This paper is organized as follows: In Sect. II we briefly recapitulate the basic ingredients of CMB spherical harmonic statistics. Afterwards in Sect. III we introduce our methods mathematically, clarify their properties and show the connection to multipole vectors. We also provide a simple proof of the convergence of the projection entropy to the Wehrl entropy up to a term which is independent of the input density matrix. Sect. IV is dedicated to the application of our methods to real data. We compare the different pseudo entropy methods, then we apply the angular pseudo entropy to 2015 Planck and 7-year WMAP full sky foreground-cleaned maps before comparing the 2015 results to those obtained with 2018 data and also applying the range entropy and comparing it to the statistics used before. Eventually in Sect. V we summarize and discuss our findings.

II. CMB STATISTICS

As a function on \( S^2 \) the CMB temperature fluctuations \( \Delta T := \delta T/T_0 \) can be decomposed uniquely according to irreducible representations of \( SO(3) \), i.e. into spherical harmonics

\[
\Delta T(\theta, \phi) = \sum_{l=0}^\infty \sum_{m=-l}^l a_{lm} Y_{lm}(\theta, \phi) \in \mathbb{R},
\]

(1)

where the \( l = 0 \)-summand is omitted because the fluctuations average to zero, and \( \theta, \phi \) denote the usual spherical coordinates. The fact that \( \Delta T \) is real together with the property \( Y_{l,-m} = (-1)^m Y_{lm} \) impose the constraints

\[
a_{l,-m} = (-1)^m a_{lm}^\ast
\]

(2)

on the spherical harmonic coefficients, leaving for each multipole number \( l \) exactly \( 2l + 1 \) real degrees of freedom. The multipole number \( l \) corresponds to angular scales of \( \approx 180/\ell \) deg. The orthonormality of \( \{Y_{lm}\} \) allows to compute the coefficients from the temperature map via

\[
a_{lm} = \int_{S^2} d\Omega \Delta T(\Omega) Y_{lm}^\ast(\Omega).
\]

(3)

Simple inflationary models together with linear perturbation theory predict nearly Gaussian temperature fluctuations and a further common assumption is statistical isotropy. The spherical harmonic coefficients inherit both properties from the temperature map, meaning that

\[
p(\vec{a}_l) = \frac{1}{\mathcal{N}} e^{-\frac{1}{2} \mathbf{a}_l^\dagger \mathbf{D}_l \mathbf{a}_l} \quad \text{(Gaussianity)},
\]

(4)

where \( p \) denotes the joint probability distribution, \( \vec{a}_l = (a_{l0}, \ldots, a_{l\ell})^\dagger \), \( \mathbf{C}_l = (\mathbf{D}_l^{-1})_{mn} = \langle a_{lm}^\ast a_{ln} \rangle \) and \( \mathcal{N} \) denotes a normalization constant, and

\[
\forall \mathbf{R} \in SO(3), \vec{e}_1, \ldots, \vec{e}_n \in S^2, n \in \mathbb{N} : \quad \mathcal{G}_n(\{\mathbf{R}\vec{e}_i\}) = \mathcal{G}_n(\{\vec{e}_i\}) \quad \text{(isotropy)},
\]

(5)

where \( \mathcal{G}_n(\{\vec{e}_i\}) = \{\prod_{i=1}^n \Delta T(\vec{e}_i)\} \) denotes the \( n \)-point function of temperature fluctuations. The isotropy condition is equivalent to rotationally invariance of the joint \( a_{lm} \)-probability distribution. The averaging \( \langle \cdot \rangle \) is meant to be performed over all possible universes, which of course is not possible, wherefore we are left with a natural and inevitable variance in all quantities, called cosmic variance. If we impose both Gaussianity and isotropy then the two-point correlation of spherical harmonic coefficients is diagonal

\[
C_{lm}^{(2)} = C_l \delta_{mn}.
\]

(6)

In practice, the power spectrum \( C_l \) is calculated using the unbiased estimator

\[
\hat{C}_l = \frac{1}{2l+1} \sum_m |a_{lm}|^2 ; \quad \langle \hat{C}_l \rangle = C_l,
\]

(7)

with cosmological variance

\[
\text{var}(\hat{C}_l) = \frac{2}{2l+1} C_l^2.
\]

(8)
III. PSEUDO ENTROPIES AND THEIR PROPERTIES

In this section we shall discuss the mathematical properties and interrelation of various macroscopic entropy measures that can be used as powerful tools to analyze Gaussianity and isotropy of the CMB and can also be useful in other contexts. This section contains a review of the mathematical background as well as new definitions, results and insights. The motivation to look for macroscopic entropy measures is the same as in statistical physics: A microscopic description of a physical system, e.g. the positions and momenta of a fluid, is useful for simulation purposes, but not when comparing to a real fluid. Instead one would resort to the study of macroscopic quantities and parameters like internal energy, temperature, entropy, pressure etc. that are well-defined because of symmetries. In the analysis of the CMB, the $a_{lm}$ coefficients are an analog of the microscopic quantities. For low $l$ these and derived quantities like multipole vectors can be studied individually, but for high $l$ this quickly becomes impractical: The Planck mission data easily comprises several million reliable data points. For the CMB the obvious underlying spacetime symmetry is rotation invariance. The representations of the rotation group decompose into irreducible components labeled by the angular momentum quantum number $l$ (multipole expansion) and we can focus on fixed-$l$ subspaces. A loose analog of internal energy is the angular power spectrum, i.e. the $C_l$ coefficients, see Eqs. 3 and 7. They have proven immensely useful in the analysis of the CMB and its cosmological implications, but when it comes to questions of isotropy and preferred directions, individual $m$ matter and an analog of entropy would be useful. A natural idea is to consider the abstract quantum state $|\psi\rangle := \sum a_{lm}|l,m\rangle$ that can be formally computed from the $a_{lm}$ and associate an entropy $S$ to it. We will usually focus on one $l$ at a time and normalize the states by rescaling the $a_{lm}$ appropriately. Since the states are by construction pure, the von Neumann entropy will be trivially zero, but there are also non-trivial pseudo entropies that can distinguish pure states and turn out to be sensitive to non-Gaussianity and anisotropy. The general strategy is as follows:

$$T(\theta, \phi) \rightarrow a_{lm} \rightarrow |\psi\rangle := \sum a_{lm}|l,m\rangle$$

$$\rightarrow \rho = |\psi\rangle \langle \psi | \rightarrow \rho_{\text{mixed}} \rightarrow S,$$

where $\rho_{\text{mixed}}$ is obtained from the pure state $\rho$ by applying a rotationally symmetric quantum channel $\Phi$, i.e. a completely positive map between Hilbert spaces with possibly different dimensions, or by computing its lower symbol, i.e. its expectation value in spin coherent states. The latter choice leads to the Wehrl entropy $|\psi\rangle := \sum a_{lm}|l,m\rangle$

$$S_W = -(2l + 1) \int \frac{d\Omega}{4\pi} \langle \Omega | \psi \rangle^2 \ln |\langle \Omega | \psi \rangle|^2,$$

where $|\Omega\rangle$ is a spin-$l$ coherent state. Wehrl entropy was first proposed in $[33]$ as a useful tool for CMB analysis. See Fig. 7d for a showing the Wehrl entropy for CMB data. Closely related is the “quadratic entropy” that is obtained by replacing $-x \ln x$ in the formula for the Wehrl entropy by the concave function $x(1-x)$:

$$S_{\text{quad}} = 1 - \frac{2l + 1}{4l + 1} |P_{2l}|^2,$$

where $P_{2l}$ is the projector onto the spin-$2l$ part, i.e. the highest spin component of the tensor product. Other examples using the choice $-x \ln x$ are what we call angular entropy

$$S_{\text{ang}} = \text{Tr} \left[ \phi \left( \sum_{i=1}^{3} \frac{L_i |\psi\rangle \langle \psi | L_i}{l(l+1)} \right) \right],$$

with $\phi(x) := -x \ln x$, where the $L_i$ are angular momentum generators in the spin-$l$ representation, and $j$-projection entropy

$$S_{\text{proj}}^{(j)} = \text{Tr} \left[ \phi \left( \frac{2l+1}{2(l+j)+1} P_{l+j} |\psi\rangle \langle \psi | \otimes \mathbf{1} \right) \right],$$

where $\mathbf{1}$ is the unit operator on a spin-$j$ ancilla $[j]$ and $\phi$ is as in Eq. 13. An overview of these entropies applied to CMB data can be found in Fig. 7. In the following section we shall explain the mathematics in detail, derive relations between the various entropies and point out interesting side results including a fast way to compute multipole vectors. Readers that are mostly interested in results and numerics can skip to the algorithm 50-54 at the end of Sect. IIIA.

A. Coherent states, multipole vectors and entropy

Coherent states were originally introduced by Schrödinger and are well known in the context of the quantum harmonic oscillator, where they can be defined either as eigenstates of the lowering operator or, equivalently, as elements of the orbit of the ground state under the Heisenberg group. Perelomov has generalized the latter notion to orbits of a fiducial vector in some representation of a Lie group under the action of that group. The choice of the fiducial vector is essential for the properties of the resulting coherent states. Spin coherent states – also called Bloch coherent states – in a spin-$l$ irreducible representation $[l] \equiv C^{2l+1}$ of $SU(2)$ with $2l + 1 \in \mathbb{N}$ are defined as orbits of the highest weight vector $|l,l\rangle$. The stability group of that vector is $U(1)$ and spin coherent states can thus be labeled by points $\Omega = (\theta, \phi)$ on the sphere $S_2 \cong SU(2)/U(1)$,

$$|\Omega\rangle = \mathcal{R}(\Omega)|l,l\rangle,$$

where $\mathcal{R}(\Omega)$ denotes a rotation that takes the north pole to the point $\Omega$ and $l$ labels the representation of $SU(2)$. For the CMB data $l$ will be an integer, but everything
we discuss here is also valid for half-integer \( l \). For \( l = \frac{1}{2} \) this gives for example

\[
|\Omega_{\frac{1}{2}}\rangle = e^{-i\phi/2} \cos \frac{\theta}{2} |\Omega_{\frac{1}{2}}, \frac{1}{2}\rangle + e^{+i\phi/2} \sin \frac{\theta}{2} |\Omega_{\frac{1}{2}}, -\frac{1}{2}\rangle .
\]  

(16)

Coherent states inherit nice properties from the underlying fiducial vector. A particular important one is that the tensor product of coherent states is again a coherent state and lies in the highest spin component:

\[
|\Omega_1\rangle \otimes |\Omega_2\rangle = |\Omega_{l+j}\rangle
\]

(17)

Using this property repeatedly yields an explicit formula for any spin from \([10]\):

\[
|\Omega_l\rangle = |\Omega_{\frac{1}{2}}\rangle \otimes \ldots \otimes |\Omega_{\frac{1}{2}}\rangle \\
= \sum_{m=-l}^l (2l) \frac{1}{2} e^{-im\phi/2} \cos^{l+m}(\frac{\theta}{2}) \sin^{l-m}(\frac{\theta}{2}) |l,m\rangle
\]

(18)

Interestingly, such a product representation in terms of spin-\( \frac{1}{2} \) states exists for any state \(|\psi\rangle \in [l] \), but except for coherent states, a projection onto the highest spin component and a renormalization are required \([28]\):

\[
|\psi_l\rangle = \sum_{m=-l}^l a_{lm} |l,m\rangle = cP_l \left( (|\Omega_{\frac{1}{2}}^{(1)}\rangle \otimes \ldots \otimes |\Omega_{\frac{1}{2}}^{(2l)}\rangle \right)
\]

(19)

where \( P_l \) is the projector onto \([l] \) and \( c \) is a normalization constant. The \( \Omega^{(i)} \) point into the direction of the \( 2l \) multipole vectors that characterize the state \(|\psi\rangle \). Contracting \([18]\) with \([19]\) and using the stereographic projection to express points on the sphere in terms of complex numbers \( z = e^{i\phi} \cot (\frac{\theta}{2}) \), leads to a polynomial

\[
\sum_{m=-l}^l \left( \frac{2l}{l+m} \right) ^{1/2} z^{l+m} a_{lm}
\]

(20)

whose \( n \leq 2l \) zeros (roots) correspond to points on the sphere that are antipodal to \( n \) of the \( 2l \) multipole vectors. The remaining \( 2l-n \) multipole vectors point to the south pole of the sphere. For the CMB data \( l \) is an integer, \( \Delta T(\Omega) \) is real and consequently \( a^{*}_{lm} = (-)^{m} a_{lm} \). This implies that the zeros of the polynomial are located at pairs of antipodal points on the sphere and the multipole vectors come in anti-aligned pairs. For details see \([33]\). In \([28]\) this method was introduced to determine explicit formulas for the Wehrl entropy and to prove Lieb’s conjecture. Applying those explicit formulas to the case \( l = 2 \) with two pairs of anti-aligned multipole vectors of length 1/2 gives the following formula for the Wehrl entropy as a function of the squared chordal distance \( \epsilon = \sin^2 (\frac{\theta}{2}) \) between the vectors, where \( \alpha \) is the angle between them:

\[
S_{W}(\epsilon) = c - \ln c + \frac{32}{15} - \ln 6 ,
\]

(21)

FIG. 1: Dependence of the angular pseudo entropy \( S_{ang}(\epsilon) \) (solid) and the Wehrl entropy \( S_{W}(\epsilon) \) (dashed) on the squared chordal distance \( \epsilon \) between multipole vectors on the sphere with radius \( r = \frac{1}{2} \) for \( l = 2 \). The maximum is obtained when both multipole vectors are orthogonal and the minima when they agree.

where

\[
c = c(\epsilon) := \frac{1}{1 - \epsilon (1 - \epsilon)} .
\]

(22)

For the angular entropy a similar computation gives

\[
S_{ang}(\epsilon) = -c \left( (1 - \epsilon)^2 \ln(1 - \epsilon)^2 + \epsilon^2 \ln \epsilon^2 \right) - \ln \frac{c}{2}
\]

(23)

with \( c \) as above. Plots of the two functions look very similar, see Fig. \( 1 \) confirming the observed similarities in behavior of the two entropy measures in the CMB analysis, see Fig. \( 7 \) The polynomial method provides a very fast and convenient way to determine multipole vectors and has been used in \([33]\) and \([25]\) and many other publications to analyze the CMB.

Spin coherent states are complete via Schur’s lemma

\[
(2l + 1) \int \frac{d\Omega}{4\pi} |\Omega_l| |\Omega_l| = P_l ,
\]

(24)

where \( P_l \) is the projector onto \([l] \). They are normalized \(|\Omega_l| |\Omega_l| = 1 \) but not orthogonal

\[
|\langle \Omega_l^l | \Omega_l^j| \rangle|^2 = \cos^{4l}(\alpha(\Omega_l, \Omega_l)) ,
\]

(25)

i.e. they form an overcomplete basis of \([l] \). In the \( l \to \infty \) limit, \( (2l+1)|\langle \Omega_l^l | \Omega_l^j| \rangle|^2 \) becomes a delta function \( \delta(\Omega_l, \Omega_l) \) and in this limit the coherent states form an infinite-dimensional orthonormal basis labeled by points on the sphere.

A striking property of coherent states is that the diagonal matrix elements

\[
A(\Omega_l) = \langle \Omega_l^l | A |\Omega_l^j\rangle \quad \text{(lower symbol)}
\]

(26)
of an operator \( A \) on \([l]\) already determine that operator uniquely: Let \( C = A - B \) with an arbitrary operator \( B \), then \( C(\Omega) = 0 \) for all \( \Omega \) implies \( C = 0 \), i.e. \( A = B \). The proof uses analytic properties of the lower symbol. The lower symbol is thus a faithful representation of an operator. Using Eq. (21), the trace of an operator \( A \) on \([l]\) can be computed as an integral over its lower symbol

\[
\text{Tr}_{[l]}(A) = (2l+1) \int \frac{d\Omega}{4\pi} \langle \Omega_l|A|\Omega_l \rangle = (2l+1) \int \frac{d\Omega}{4\pi} A(\Omega) .
\]

Another interesting property is that any operator \( A \) on \([l]\) can be expanded diagonally in coherent states

\[
A = (2l + 1) \int \frac{d\Omega}{4\pi} h_A(\Omega) \langle \Omega_l | \Omega_l \rangle ,
\]

where \( h_A(\Omega) \) is called an upper symbol of \( A \). These two properties are in fact closely related: Contracting Eq. (28) with an operator \( C \) gives \( \text{Tr}(C^\dagger A) \propto \int \frac{d\Omega}{4\pi} h_A(\Omega) \Omega^*(C) \), i.e. the operators that can be represented by an upper symbol as in Eq. (28), are orthogonal to the operators that are in the kernel of the lower symbol map. Hermitean operators have real lower and upper symbols. Positive semi definite operators and density matrices have unique non-negative lower symbols, but the same is in general not true for upper symbols. Following Wehrl, these properties suggest to interpret the lower symbol of a density matrix \( \rho \), which is by definition positive semi definite and normalized, as a probability density and compute

\[
(2l + 1) \int \frac{d\Omega}{4\pi} \phi(\rho(\Omega)) = (2l+1) \int \frac{d\Omega}{4\pi} \phi(c) .
\]

With \( \phi(x) = x \) we can verify the normalization and get \( \text{Tr}(\rho) = 1 \). With \( \phi = -x \ln x \) we compute the Shannon entropy of the probability density \( \rho(\Omega) \), which is precisely the Wehrl entropy

\[
S_W(\rho) = -(2l + 1) \int \frac{d\Omega}{4\pi} \rho(\Omega) \ln \rho(\Omega) ,
\]

with the special case \([10]\) for a pure state \( \rho = |\psi\rangle\langle\psi| \). The Wehrl entropy was introduced as a semi-classical entropy, which is mathematically better behaved than the Boltzmann entropy in classical statistical mechanics. The Wehrl entropy is always larger than the von Neumann entropy and it is positive even for pure states. Furthermore its definition is rotationally symmetric and it is hence perfectly suited for our purposes. The minimum of the Wehrl entropy is attained for coherent states. This fact is surprisingly difficult to prove. It was first shown for low spin in \([28]\) and then finally in general in \([31] \) and \([32]\). Computational evidence suggested in fact an analogous but much stronger conjecture for any concave function \( \phi(x) \) \([25] \) \([28] \) \([30]\), which has also been settled affirmatively in \([31]\). For our application the maximal value of the Wehrl entropy is more interesting. The exact value is not known, in fact finding the maximizing pure state is another hard problem, but a reasonable upper limit can be obtained very simply from a totally mixed state: \( S_W \leq \ln(2l + 1) \). To summarize: The Wehrl entropy is the Shannon entropy of the probability density obtained from the (faithful) lower symbol representation of a density matrix. It has all the right properties for our purposes. The only drawback is that its computation with suitable precision has a high computational complexity. We will now introduce and discuss several alternatives with similar properties, but better computability.

Instead of \(-x \ln x \), one can consider other concave functions defined on the interval \([0, 1]\). For \( \phi(x) = x(1 - x) \) we obtain the quadratic entropy

\[
S_{\text{quad}}(\rho) = 1 - \frac{2l + 1}{4l + 1} \text{Tr}(P_{2l}(\rho \otimes \rho)) ,
\]

where \( P_{2l} \) is the projector onto spin \( 2l \) and we have used the following trick \([25]\)

\[
(\rho(\Omega))^2 = \langle \Omega_l | \rho | \Omega_l \rangle^2 = \langle \Omega_l \otimes \Omega_l | \rho \otimes \rho | \Omega_l \otimes \Omega_l \rangle = \langle \Omega_{2l} | \rho \otimes \rho | \Omega_{2l} \rangle
\]

and the trace formula \([27]\) adapted to spin-\(2l\). For a pure state \( \rho = |\psi\rangle\langle\psi| \) we obtain formula \([11]\). The quadratic entropy has a large computational advantage as compared to the Wehrl entropy, see Fig. \([9]\) but its qualitative behavior is a bit different as can be seen in Fig. \([7]\). This is not surprising, since many important properties of entropy like additivity and (strong) sub-additivity depend crucially on the choice of the function \( \phi(x) = -x \ln x \). We shall hence not pursue quadratic entropy any further and will try to identify and construct other alternatives of the Wehrl entropy that share its characteristic features but are computationally more accessible. Let us start by reconsidering the main ingredient of Wehrl entropy.

Let \( \rho \) be a density matrix on \([l] = \mathbb{C}^{2l+1} \) and introduce an ancilla Hilbert space \([j] = \mathbb{C}^{2j+1} \). Using the product property \([17]\) and normalization of coherent states, we can rewrite the lower symbol \( \rho(\Omega) \) that enters the formula for the Wehrl entropy as follows:

\[
\langle \Omega_l | \rho | \Omega_l \rangle = \langle \Omega_l \otimes \Omega_j | \rho \otimes 1 | \Omega_l \otimes \Omega_j \rangle = \langle \Omega_{l+j} | \rho \otimes 1 | \Omega_{l+j} \rangle ,
\]

where \( 1 \) is the unit operator on \([j]\). The values of the lower symbol are thus the diagonal elements of a family of infinite-dimensional matrices

\[
\rho_j(\Omega, \Omega') = \langle \Omega_{l+j} | \rho \otimes 1 | \Omega_{l+j}' \rangle .
\]

By an infinite-dimensional compact analog of the Schur-Horn theorem the diagonal elements \( \rho(\Omega) \) are majorized by the eigenvalues of the \( \rho_j(\Omega, \Omega') \) matrices. This implies that any concave function of the values \( \rho(\Omega) \) will be larger or equal to the respective function of the eigenvalues of \( \rho_j(\Omega, \Omega') \). The Wehrl entropy is therefore larger than or equal to the von Neumann entropy of \( \rho_j(\Omega, \Omega') \). For convex functions the inequalities are reversed. See e.g. \([37]\)
for an overview of the mathematical background. In the limit \( j \to \infty \) and in view of Eq. (25) the off-diagonal matrix elements of \( \rho_j(\Omega, \Omega') \) become zero and the inequalities become equalities. Using the property (\ref{eq:property}) on both sides of Eq. (33) we can recover a finite-dimensional matrix

\[ P_{l+j} (\rho \otimes 1) P_{l+j} \]  

from \( \rho_j(\Omega, \Omega') \), where \( P_{l+j} \) is the projector onto the highest spin component \([l+j]\) of the tensor product. The matrix (\ref{eq:matrix}) has the same eigenvalues as \( \rho_j(\Omega, \Omega') \). In fact, if

\[ P_{l+j} (\rho \otimes 1) P_{l+j} |V_\lambda\rangle = \lambda |V_\lambda\rangle \]  

then \( V_\lambda(\Omega) := \langle \Omega_i | V_\lambda \rangle \) satisfies

\[ (2l+j+1) \int \frac{d\Omega'}{4\pi} (\Omega_{l+j} \rho \otimes 1) \Omega_{l+j} V_\lambda(\Omega') = \lambda V_\lambda(\Omega) \]  

(\ref{eq:matrix}) and vice versa if \( V_\lambda(\Omega) \) is a solution of Eq. (\ref{eq:matrix}), then

\[ |V_\lambda\rangle = (2l+j+1) \int \frac{d\Omega'}{4\pi} \Omega_{l+j} V_\lambda(\Omega) \]  

(\ref{eq:matrix}) satisfies Eq. (\ref{eq:matrix}). We have shown that the eigenvalues of the matrix (\ref{eq:matrix}) majorize the values of the lower symbol of \( \rho \) in the sense explained above, namely that inequalities are implied for concave (or convex) functions of these values. It can furthermore be shown that pure states majorize mixed states and that among the pure states, projectors \( |\Omega\rangle \langle \Omega| \) onto coherent states will lead to matrices (\ref{eq:matrix}) that majorize all other choices. Among the concave functionals we are interested in entropy and define an appropriately normalized mixed density matrix

\[ \rho_{\text{proj}}^{(j)} = \frac{2l+1}{2(l+j)+1} P_{l+j} (\rho \otimes 1) P_{l+j} , \]  

(\ref{eq:rho_proj}) whose von Neumann entropy is what we call the “projection entropy”

\[ S_{\text{proj}}^{(j)}(\rho) = \text{Tr} \left[ \phi \left( \frac{2l+1}{2(l+j)+1} P_{l+j} (\rho \otimes 1) P_{l+j} \right) \right] , \]  

(\ref{eq:projection_entropy}) with \( \phi(x) = x \ln(x) \). From the fact that the mixed density matrix (\ref{eq:rho_proj}) has at most \( 2j+1 \) non-zero eigenvalues, we get an upper bound for the projection entropy (\ref{eq:projection_entropy})

\[ S_{\text{proj}}^{(j)}(\rho) \leq \ln(2j+1) . \]  

From the \([l+j]\) perspective the Wehrl entropy should also be computed from Eq. (\ref{eq:projection_entropy}) and we get the aforementioned inequalities. The only differences from the original definition of Wehrl entropy (\ref{eq:wehrl_entropy}) is a rescaling of the density matrix and related renormalization of the integral, which leads to a shift in entropy and the following inequality:

\[ S_W(\rho) \geq S_{\text{proj}}^{(j)}(\rho) + \ln \left( \frac{2l+1}{2(l+j)+1} \right) . \]  

(\ref{eq:wehrl_lower_bound})

In the limit \( j \to \infty \) this inequality becomes an equality, see Fig. 2 for the converge of \( S_{\text{proj}}^{(j)}(\rho) \) to \( S_W(\rho) \) for NILC Planck data and \(^\text{[39]}\) for an alternative proof. The projector

\[ P_{l+j} : |l\rangle \otimes |j\rangle \to |l+j\rangle \]  

can be expressed in terms of Clebsch-Gordan coefficients:

\[ P_{l+j} (|l,m\rangle \otimes |j,M\rangle) = \sqrt{\binom{2l}{l+m} \binom{2j}{j+M} \binom{j+M}{l+j+m+M}} \]  

(\ref{eq:projection_matrix})

For large \( j \) the projection method provides a good way to compute the Wehrl entropy with high precision. For small \( j \) we get an entropy measure with all the nice properties of Wehrl entropy, but a pretty large computational advantage. We will now focus on the case where \( \rho \) is a pure state, i.e. \( \rho = |\psi\rangle \langle \psi| \) with \( |\psi| \) computed from the \( a_m \) of the CMB data with \( l \) fixed. For a pure state \( \rho \) the matrix (\ref{eq:rho_proj}) can be rewritten as the Gram matrix of a set of vectors \( \tilde{V}_M \in [l+j] \) that are labeled by a basis of \([j]\):

\[ P_{l+j} (|\psi\rangle \otimes |1\rangle) P_{l+j} = \sum_{M=-j}^{j} \tilde{V}_M \tilde{V}_M^\dagger \]  

(\ref{eq:projection_matrix})

with \( \tilde{V}_M = P_{l+j} (|\psi\rangle \otimes |j,M\rangle) \) .

The dual Gram matrix

\[ \text{Tr} \tilde{V}_M \tilde{V}_M^\dagger = \tilde{V}_M^\dagger \tilde{V}_M \]  

(\ref{eq:gram_matrix})

has the same non-zero eigenvalues as the original matrix, because for any matrix \( C, \text{CC}^\dagger \) and \( C^\dagger C \) have the same non-zero singular values. We can therefore also use the dual Gram matrix for the computation of the projection
entropy. Appropriately normalized and written in basis-independent notation we have
\[ \hat{\rho}^{(j)}_{\text{proj}} = \frac{2l+1}{2(l+j)+1} \langle \psi_l | P_{l+j} | \psi_l \rangle \]
and
\[ S^{(j)}_{\text{proj}}(\rho) = -\text{Tr} \left( \hat{\rho}^{(j)}_{\text{proj}} \ln \left( \hat{\rho}^{(j)}_{\text{proj}} \right) \right) , \]
where the expectation value is taken in the first tensor slot of $P_{l+j}$. Unlike $\rho^{(j)}_{\text{proj}}$ the new density matrix $\hat{\rho}^{(j)}_{\text{proj}}$ is in general not a faithful representation of the underlying $\rho$ for $j < l$, but the entropy is precisely the same, while its computation involves smaller matrices and is faster. The computational advantage is particularly large for small $j$. The projection entropy computed in this way is an excellent tool for the analysis of the CMB and other spherically distributed data.

Expanding the unit operator on $[j]$ in Eq. (37) in terms of basis states, it can be seen that the map $\rho \rightarrow \rho^{(j)}_{\text{proj}}$ is in fact a trace preserving completely positive map (quantum channel) $[l] \rightarrow [l+j]$ in Kraus form:
\[ \rho^{(j)}_{\text{proj}} = \sum_M A_M \rho A_M^\dagger, \quad \sum_M A_M^\dagger A_M = 1 \]
\[ A_M = \sqrt{\frac{2l+1}{2(l+j)+1}} P_{l+j}[j, M], \]
where the last terms can also be written $P_{l+j}[j, M] = \sum_m [j + l, m + M] [l, m]$. There is a similar formula for the transformation of the density matrix in the dual Gram matrix formulation. In view of the $j \rightarrow \infty$ limit, the lower symbol of a density matrix can also be interpreted as resulting from a completely positive map.

We shall now introduce yet another natural choice of a rotationally invariant quantum channel, leading to what we call “angular entropy”, which shares the nice properties of the aforementioned entropies with the additional advantage of being even faster to compute. Let $L_1, L_2, L_3$ be the standard angular momentum generators in the spin-$l$ representation and define a mixed density matrix and entropy via
\[ \rho_{\text{ang}} = \frac{1}{l(l+1)} \sum_{i=1}^3 L_i \rho L_i^\dagger \]
\[ S_{\text{ang}} = -\text{Tr}(\rho_{\text{ang}} \ln(\rho_{\text{ang}})) . \]
The transformation is obviously of Kraus form and therefore completely positive. It is trace-preserving because $C = \sum_i L_i^\dagger L_i$ is the quadratic casimir and has value $l(l+1)$ in the spin $l$ representation. The formula for angular entropy can be written in a basis-independent way by replacing $\sum L_i \otimes L_i$ by $\frac{1}{2}(\Delta C - C \otimes 1 - 1 \otimes C)$, where $\Delta C$ the coproduct of the casimir. In practice the formula is usually rewritten in terms of $\frac{1}{\sqrt{2}} L_{\pm}$ instead of $L_1$ and $L_2$. Therefore we have included the dagger $\dagger$ in Eq. (49), which is of course not necessary for Hermitian $L_i$. For a pure state $\rho = |\psi\rangle\langle\psi|$, there is also a dual Gram matrix formulation of the angular entropy:
\[ G_{ij} = \langle \psi | C^{-1} L_i^\dagger L_j | \psi \rangle , \quad S_{\text{ang}} = -\text{Tr}(G \ln(G)) , \]
see also [29][30] for application to CMB data. In the way we have written this formula, it is now in fact no longer restricted to individual angular momentum (multipole) numbers $l$. It can also be applied to a range or even a selection of $l$: One simply needs to insert an appropriately normalized state
\[ |\psi\rangle = \sum_{l \in \text{select}} \sum_m a_{lm} |l, m\rangle , \]
leading to the following explicit algorithm: First determine the Hermitian $3 \times 3$ dual Gram matrix $G$
\[ G = \begin{pmatrix} G_{11} & G_{12} & G_{13} \\ G_{21} & G_{22} & G_{23} \\ G_{31} & G_{32} & G_{33} \end{pmatrix} \]
with matrix elements
\[ G_{ij} = \sum_{l \in \text{select}} \frac{1}{2(l+1)} \theta_{ij}(l, \{a_{lm}\}) , \]
with
\[ \theta_{11} = \sum_{m=-l}^l (l + m + 1)(l - m) \cdot |a_{lm}|^2 \\
\theta_{12} = \sum_{m=-l+1}^{l-1} \sqrt{(l^2 - m^2)((l + 1)^2 - m^2)} \cdot a_{l,m-1}^* a_{l,m+1} \\
\theta_{13} = \sqrt{2} \sum_{m=-l}^{l-1} (m+1) \sqrt{(l + m + 1)(l - m)} \cdot a_{lm} a_{l,m+1}^* \\
\theta_{22} = \sum_{m=-l}^l (l - m + 1)(l + m) \cdot |a_{lm}|^2 \\
\theta_{23} = \sqrt{2} \sum_{m=-l+1}^l (m-1) \sqrt{(l - m + 1)(l + m)} \cdot a_{lm} a_{l,m-1}^* \\
\theta_{33} = 2 \sum_{m=-l}^l m^2 \cdot |a_{lm}|^2 , \]
where “select” is a chosen selection of multipole angular momentum quantum numbers $l$. In this paper we typically select a single value at a time, but this can also be a range of values or an even more complex selection.

The angular entropy is then computed in terms of the three eigenvalues $\lambda_i \in [0,1]$ of the normalized mixed angular density matrix $\rho_{\text{ang}} = G / \text{Tr}(G)$, with
\[ \text{Tr}(G) = \sum_{l \in \text{select}} (2l + 1) \hat{C}_l , \]
\[ S_{\text{ang}} = -\text{Tr}(\rho_{\text{ang}} \ln(\rho_{\text{ang}})) = -\sum \lambda_i \ln(\lambda_i) . \]
See [40] for an overview of algorithms for the fast and precise computation of eigenvalues of Hermitian $3 \times 3$
matrices. There exist also two range entropy measures using the Wehrl entropy which were identified in [41].

The computation of the angular entropy involves only $3 \times 3$ matrices and their eigenvalues. It is by far the fastest method and numerical experiments with actual and simulated data show that it has similar behavior as the Wehrl entropy. From a theoretical point of view there are some similarities between the angular and projection entropy and hence also the Wehrl entropy: As we have mentioned, the projector $P_{l+l}$ is related to the Clebsch-Gordon decomposition of $[l] \otimes [j]$. Likewise, the angular momentum operators can be interpreted as Clebsch-Gordon coefficients of the decomposition of $[l] \otimes [l]$, while the projector is onto the highest spin component, the angular momentum generators pick out the adjoint spin-1 representation. For the angular entropy there are similar conjectures as for the Wehrl and projection entropies, which are still open and under current consideration. There are several further generalized pseudo entropies – for example one could choose a (convex) function of the Casimir in the definition of angular entropy, e.g.

$$\frac{1}{l(l+1)} \sum_{i_1, \ldots, i_k} L_{i_1} \cdots L_{i_k} \rho L_{i_1} \cdots L_{i_1}. \quad (55)$$

Below, we focus on the projection and angular entropies that we have defined in this section.

### B. Probability distribution

In this section we want to compare numerically the behavior of the angular entropy for isotropic and Gaussian maps to its behavior for maps that are constructed from multipole vectors which are distributed uniformly on the sphere according to the surface measure. Such maps are statistically isotropic but not Gaussian, hence we investigate for deviations of Gaussianity without violating statistical isotropy. Figs. 3 and 4 show the probability and cumulative distributions at various multipoles for Gaussian as well as uniform multipole vector maps; Fig. 5 shows a direct comparison of both map types at $l = 6$. In the isotropic and Gaussian case the probability density squeezes for increasing values of $l$ while the peak shifts to the right, most probably converging to $\ln(3)$ for $l \to \infty$. In turn, the cumulative distribution shows a simple behavior for $l \geq 5$, namely it can be fitted with a Gaussian bell curve and for $l = 5$ it suits the numerics nicely except for $S_{\text{ang}} \geq 1.0$, while for $l \geq 10$ the fit coincides with the numerical result except for a small range at the lower tail where the deviation is still small. At $l = 3$ the bell curve fit breaks down since the cumulative distribution increases too fast. In contrast to this the probability distribution of uniform multipole vectors shows a more chaotic behavior in dependence on $l$. The curve disperses and the peak moves interchangeably to the left and right.

Apparently, the angular pseudo entropy allows to distinguish clearly between random multipole vectors and isotropic, Gaussian maps and therefore it is a measure that is sensitive to deviations from Gaussianity as a result of the fact that multipole vectors, that one calculates from an isotropic and Gaussian map are, contrary to intuition, not uniformly distributed but feel repulsion. The reason for this is that multipole vectors emerge as zeros of a random polynomial which has Gaussian coefficients if the temperature map is Gaussian. It is widely known that such zeros always experience repulsion. From Fig. 5 the probability density in the isotropic Gaussian case is very low where in the uniform multipole vector case the probability density has its maximum. In turn, the maximum of the isotropic and Gaussian distribution is situated approximately at FWHM of the uniform mul-

![Graph showing probability and cumulative distributions](image-url)
FIG. 4: Distributions of angular pseudo entropy for uniformly distributed multipoles vectors at multipoles \( l = 2, 3, 4, 5, 6, 7 \). When increasing \( l \), the curves move alternately to the right (even \( \rightarrow \) odd) and left (odd \( \rightarrow \) even), depending on the parity of \( l \), but on average the curves move to the left.

Interestingly, for \( l > 3 \) both the coherent state and the uniform multipole vectors result in lower angular entropy than the Gaussian and isotropic maps, although these two cases display the extremal cases. A coherent state corresponds to alignment of all multipole vectors while uniform multipole vectors span widely across the sphere. So the maximum entropy state, which is a maximally mixed state, corresponds to a situation between uniformity and absolute confinement.

It should be noted that an equivalent expression to the angular entropy has already been introduced under the name of power entropy in [29] but without reference to the Wehrl entropy and completely positive maps. Furthermore, that work made the wrong assumption that the maximal entropy value \( \ln(3) \) would be obtained for isotropic maps. The method was applied to Planck and WMAP in [30] but with the main focus on the correlation of multipoles with the quadrupole. There, no large-scale anomalies were observed, but significant correlations with the quadrupole were found on a wider range of scales.
IV. APPLICATION TO CMB DATA

We use Planck 2015 second release data, in particular the four cleaned full sky maps COMMANDER, NILC, SEVEM and SMICA, together with the WMAP 7-year ILC cleaned full sky map. The names stand for different cleaning algorithms applied to the original data. COMMANDER uses astrophysical models in order to fill in masked regions that contain foreground contamination, NILC stands for "Needlet Internal Linear Combination" and represents a refinement of the ILC algorithm using needlets in harmonic space, SEVEM uses template fitting and SMICA fills masked regions by a Metropolis Monte Carlo random process. Later in this section we also compare the 2015 results we obtain with results obtained from recently published 2018 Planck data. We process the data using the Healpy[42] and Numpy packages for Python 2.7. In order to compute confidence levels, a number of ensembles of Gaussian and isotropic random $a_{lm}$ are treated as input data for the various entropies. Depending on the entropy the number of ensembles ranges from 30 to $10^4$.

A. Comparison of pseudo entropies

The considered pseudo entropies differ in computational expense, see Fig. 6. Computing the angular entropy up to $l = 1000$ takes about 90 seconds per run, while the quadratic entropy is slightly more slow. The quadratic entropy is also fast to compute, but it should be used with care since, due to its $-x^2$ instead of the usual $x \log(x)$ behavior it lacks some of the usual entropy properties. Because the projection entropy converges to the Wehrl entropy for $l \to \infty$ up to a term which does not depend on the data, see Eq. (39) in Sect. III, its running time converges as well. Clearly, the Wehrl entropy is the quantity that needs the largest computation time, namely about 3000 seconds up to $l_{\text{max}} = 30$, for system resources see app. A.

Fig. 7 shows that all measures except the quadratic pseudo-entropy exhibit very similar features in the data analysis, which has also been noticed in [43]. In particular, we observe unusually large values at $l = 5, 28$ and conspicuously small values at $l = 6, 16, 17, 30$. On the other hand, the quadratic pseudo entropy singles out other unlikely multipoles, e.g. $l = 14$. This shows again that this measure should be used with care and that the other measures suit our purposes better. It is interesting to see that the most unusual multipoles $l = 5$ and $l = 28$ have an entropy that is far above the expectation value. For non-Gaussian or non-isotropic maps one would in general expect the entropy to be lower than the expectation, as will be shown later in more detail, the confidence levels are constant from intermediate $l$ on, while the Gaussian expectation for the angular entropy shows a very simple functional dependence on $l$ as well. Furthermore, in contrast to unlogarithmic plotting both the upper and lower confidence levels have approximately the same width allowing for a better identification of unusual multipoles. For comparison see fig. 19 in app. B which shows the pure angular pseudo entropy.

Concluding, the agreement of features in the different pseudo entropies suggests considering only the numerically cheapest entropy apart from the quadratic one. Hence, in the following only the angular pseudo entropy will be considered.
a) \((j = 1)\)-projection entropy, 10000 ensembles of random \(a_{lm}\).

b) Angular entropy, 10000 ensembles of random \(a_{lm}\).

c) \((j = 10)\)-projection entropy, 1000 ensembles of random \(a_{lm}\).

d) Wehrl entropy, 30 ensembles of random \(a_{lm}\).

e) \((j = 100)\)-projection entropy, 100 ensembles of random \(a_{lm}\).

f) Quadratic entropy, 1000 ensembles of random \(a_{lm}\).

FIG. 7: Comparison of pseudo entropies from \(l = 1\) to \(l = 30\). The sigma boundaries were determined by a certain number of random maps and have been smoothed with a Gaussian filter in all plots. Note, that concerning the angular entropy the smoothing broadens the confidence levels for \(l = 2, 4\) and straightens them for \(l = 3\). For higher multipole numbers the smoothing does not add or remove any features.
FIG. 8: Expectation, upper and lower confidence levels of \( \log \left( \frac{1}{\log(3)^{-S_{\text{ang}}(l)}} \right) \) with 10000 ensembles of isotropic and Gaussian random \( a_{lm} \) up to \( l = 100 \). The dashed lines represent fits.

B. Results for angular pseudo entropy with 2015 data

From fig. 7b we read off that for the angular entropy in the range \( 1 \leq l \leq 30 \) five NILC data points lie at 2\( \sigma \) or even outside of it (\( l = 5, 16, 17, 28, 30 \)), two of which are even close to 3\( \sigma \) (\( l = 5, 28 \)). One could now argue, that it is expected that some data points lie at low confidence levels, but a quick estimation shows that the deviations observed here are still unlikely. The probability for five out of 30 data points to lie outside 2\( \sigma \) approximately equals the Poisson distribution for five events with a mean rate \( \lambda = 1.2 = 30 \cdot 0.04 \), i.e.

\[
P_\lambda(5) = \frac{\lambda^5}{5!} e^{-\lambda} \approx 0.6%,
\]

implying that the significance of these unlikely data points is above 2\( \sigma \).

Turning to higher multipole numbers it would be beneficial to find a method of calculating confidence levels even faster. In this regard we observe that in the logarithmic reciprocal depiction the Gaussian expectation value and confidence levels of the angular entropy behave in a simple fashion, namely the expectation can be fitted with \( f(x) = a \log(bx + c) \) and the confidence levels with \( g(x) = a(1 - e^{-b(x-1)}) \), see Fig. 8. In particular, it turns out that the confidence levels are constant from about \( l = 30 \) up to \( l = 100 \). In the following we assume that this holds true for \( l > 100 \). This assumption is justified by continuity of the angular pseudo entropy and the isotropic, Gaussian probability distribution of spherical harmonic coefficients, i.e. no sudden jumps should be expected. Tab. III in App. B contains all optimal parameters.

While the fit of the expectation value coincides well with the numerical graph on the whole considered range, the lower confidence fits are not suited for \( l = 2 \) and \( l = 3 \) and the upper confidence fits suit the numerical results from \( l = 3 \) (1\( \sigma \)), \( l = 4 \) (2\( \sigma \)) and \( l = 20 \) (3\( \sigma \)). Hence, using the fits in analysis slightly underestimates the most conspicuous multipoles with values above the expectation value in the range \( 3 \leq l \leq 20 \), but the fits allow for a comparison of the entropy to the expectation from \( l = 1 \) to \( l = 1000 \).

In Fig. 9 we applied the fits for the angular entropy up to \( l = 1000 \), once for the pure cleaned full sky maps and once masked with the SEVEM mask and without refilling the masked region. In the second case WMAP was taken out because of dissimilar NSIDE number of this map and the SEVEM mask. The unmasked shows no obvious deviation of COMMANDER, NILC and SMICA from the expected behavior of a Gaussian map on the whole ob-
served range of multipoles, while single multipoles stick out, as for example NILC at $l = 896$, but the data does not exhibit unusual global deviations from the expectation, i.e. deviations on a large range of angular scales. On the other hand WMAP and SEVEM clearly fall off from $l = 200$ on. While the WMAP data is commonly accepted to be inaccurate on very small angular scales, the large drop of SEVEM surprises at first glance. However, the masked plot shows that the deviation of SEVEM from the other Planck maps can be explained largely by the strong influence of residual foreground pollution in the SEVEM map. Indeed, the masked Planck maps all coincide very well on the whole range, leaving only minor deviations. It can be seen that the pure masking process lowers the entropy for large values of $l$ indicating, as expected, that masking singles out certain directions by removing the galactic plane. This can nicely be seen by taking into account the dashed red line which shows the entropy of a coherent state which represents a map that is confined to a single direction. That masking lowers angular pseudo entropy is not a priori clear since we normalize the $a_l$ before computing pseudo entropies and hence there is no lack of absolute power due to masking.

Taking a closer look to angular scales around $l = 900$, see Fig. 10 and Tab. 1 for the full sky maps reveals that NILC behaves unusually between $l = 895$ and $l = 905$. We measure unusualness of multipoles with the p-value, using the convention

$$p(l) = \begin{cases} \int_{S_{\text{ang}}(l)}^{\log(3)} ds \ p_S(s) & \text{if } S_{\text{ang}}(l) > \langle S_{\text{ang}} \rangle, \\ \int_{S_{\text{ang}}(l)}^{S_{\text{ang}}(l)} ds \ p_S(s) & \text{if } S_{\text{ang}}(l) < \langle S_{\text{ang}} \rangle, \end{cases}$$

where $p_S(s)$ denotes the probability distribution of $S_{\text{ang}}$

![FIG. 10: Angular pseudo entropy between $l = 890$ and $l = 910$, confidence levels and expectation calculated with 1000 random ensembles and smoothed with a Gaussian filter.](image)

| $l$ | pipeline | $p\%$ | $p_{\text{geom}}\%$ |
|-----|----------|-------|---------------------|
| 895 | NILC     | 18.8  | 17.7                |
|     | COMMANDER| 36.6  | 8.1                 |
|     | SMICA    | 8.1   |                     |
| 896 | NILC     | 0.1   | 1.2                 |
|     | COMMANDER| 29.5  | 0.6                 |
|     | SMICA    | 0.6   |                     |
| 897 | NILC     | 46.3  | 23.4                |
|     | COMMANDER| 55.5  | 6.1                 |
|     | SMICA    | 6.1   |                     |
| 898 | NILC     | 2.6   | 6.9                 |
|     | COMMANDER| 13.1  | 9.7                 |
|     | SMICA    | 9.7   |                     |
| 899 | NILC     | 3.1   | 5.9                 |
|     | COMMANDER| 10.5  | 5.9                 |
|     | SMICA    | 6.4   |                     |
| 900 | NILC     | 2.2   | 1.7                 |
|     | COMMANDER| 1.1   | 1.7                 |
|     | SMICA    | 2.1   |                     |
| 901 | NILC     | 4.1   | 4.3                 |
|     | COMMANDER| 5.4   | 3.7                 |
|     | SMICA    | 3.7   |                     |
| 902 | NILC     | 0.4   | 7.7                 |
|     | COMMANDER| 4.0   | 7.7                 |
|     | SMICA    | 25.9  |                     |
| 903 | NILC     | 36.4  | 37.9                |
|     | COMMANDER| 33.2  | 37.9                |
|     | SMICA    | 45.1  |                     |
| 904 | NILC     | 3.0   | 13.9                |
|     | COMMANDER| 31.4  | 13.9                |
|     | SMICA    | 28.3  |                     |
| 905 | NILC     | 45.7  | 32.3                |
|     | COMMANDER| 17.3  | 32.3                |
|     | SMICA    | 42.8  |                     |

TABLE 1: P-values of angular pseudo entropy for $895 \leq l \leq 905$, calculated with 1000 ensembles of Gaussian $a_{lm}$, rounded to one decimal place. In the third column the geometric mean of the p-values of the three maps was chosen because of the multiplicative behavior of probabilities.

for Gaussian and isotropic $a_{lm}$ and $\langle S_{\text{ang}} \rangle$ the respective expectation value. The entropy value at $l = 896$ lies outside the $3\sigma$-region with p-value $\leq 0.1$, where the reason for the inequality is the low number (1000) of random ensembles that have been used to calculate this value, which yields a resolution of 0.1. This means that on average at most one out of 1000 realizations is expected to be larger than the expectation and to be as unusual as or more unusual than the data point. On the other hand one interpretation is that one out of 1000 multipoles is expected to be at least as unusual as the data point. Since $l = 986$ is the only NILC data point on $1 \leq l \leq 1000$ that is outside of $3\sigma$, this multipole is still allowed by statistics. Nevertheless, the NILC values of the angular entropy exhibit small p-values at $l = 896, 898, 899, 900, 901, 902, 904$. SMICA behaves a bit less extreme than NILC on the considered range and COMMANDER stays inside or close to the $1\sigma$-region.
with on average large p-values. While for the multipoles $l = 898$ to $l = 901$ all the pipelines behave similarly, they deviate from each other at the other multipoles between $l = 890$ and $l = 910$.

In order to estimate the significance of this multipole range, we calculate the geometric mean over p-values,

$$\langle p \rangle_{\text{geom}}(\text{map}) = \left( \prod_{l=895}^{905} p^{(\text{map})}(l) \right)^{1/11},$$

and compare it to the distribution of p-values for Gaussian and isotropic random maps, see Fig. 11. For NILC the geometric mean is $\langle p \rangle_{\text{geom}}(\text{NILC}) = 4.4\%$. From 1000 ensembles of random Gaussian and isotropic maps not a single map attains such a small mean p-value, hence we can give an estimate on the upper bound of the likelihood of the NILC data in the given multipole range assuming Gaussianity and isotropy as a null hypothesis

$$L(\text{NILC}; 895 \leq l \leq 905) \lesssim 0.1\%,$$

i.e. a 3σ-significance. For judging all three full sky maps together, we use the geometric mean of the p-value over the three maps and proceed with these mean p-values as with NILC, resulting in a mean p-value of 8.6% on $[895, 905]$ with a likelihood of

$$L(\text{COMMANDER} \cdot \text{NILC} \cdot \text{SMICA}; 895 \leq l \leq 905) \approx 0.8\%,$$

i.e. a more than 2σ-significance.

We conclude that either NILC, and to a lesser extent SMICA, might induce bad characteristics to the data on the mentioned scales or that the COMMANDER algorithm might induce arbitrary isotropy and/or Gaussianity on these scales and therefore distorts the real data. Furthermore, even if one considers all three maps at once by multiplying the p-values and comparing to the expectation, the data is inconsistent with the assumption of isotropy and Gaussianity at a 2σ-level.

It should be noted that this might well be a selection effect due to the particular chosen, non-physically motivated range of scales that was considered. Moreover we do not perform a fully developed and precise statistical analysis, hence the estimated significance might need corrections. The essence here is that the entropy method is capable of highlighting unusual behavior at high multipole numbers.

Now, we return to the large angular range $l \leq 30$, where we computed p-values with 10000 sets of random $a_{lm}$, see App. [13] Fig. 20 for a plot of the p-values. Tab. [11] shows the six most pronounced large scale multipoles revealing again that unmasked SEVEM does not exhibit the same behavior as the other maps, yielding large p-values at these multipoles while the other maps show small p-values. It turns out that $l = 28$ with an average (excluding SEVEM) p-value of about 0.26% sticks out most, followed by $l = 30$ and $l = 5$. Although two of the most conspicuous multipoles display a too large value of the entropy, neither too large nor too small values can directly be identified to be preferred. At these angular scales, the three non-SEVEM Planck maps behave quite similarly and the large discrepancy between COMMANDER and NILC is not yet present. Another feature that can be observed is the slight improvement of Planck compared to WMAP, even for $l < 250$ because on average WMAP yields the smallest p-values at these conspicuous multipoles.

It has been conjectured in [10] that some of the large scale features could be produced by data processing,
namely by a non-linearity in the masking process which mixes the large dipole moment to higher moments when subtracting the dipole. We try to answer this question in a very simplified approach. In Fig. 12 we plot the angular entropy for ten isotropic and Gaussian random full sky maps and in a second step we add the non-relativistic contribution of a dipole to the map

\[ T(\vec{e}) \rightarrow T'(\vec{e}) = T(\vec{e}) + A \hat{\vec{d}} \cdot \vec{e}, \]

where \( A = 3364.5 \mu \text{K} \) denotes the Cosmic Dipole amplitude and \( \hat{\vec{d}} = (x(l,b), y(l,b), z(l,b))^T \) with \( (l,b) = (264.00 \text{ deg}, 48.24 \text{ deg}) \) denotes its direction in the galactic coordinate system, then we mask the map and remove the dipole afterwards again, using this time the build-in Healpy function `remove_dipole`, which returns a map \( \tilde{T} \) that is the closest – in the meaning of a least square fit – map to the original \( T \) among those maps obeying \( \sum_{p \in \mathcal{P}} \tilde{e}_p \tilde{T}_p = 0 \), where \( \mathcal{P} \) denotes the set of all unmasked pixels. Finally, we refill the masked region with the original data \( T \) in order to receive a full sky map, see Fig. 21 in App. B for a depiction of this process by maps in Mollweide view. Since the WMAP and Planck maps behave similarly on angular scales and working with WMAP is computationally cheaper than working with Planck maps – WMAP has NSide = 512 and Planck NSide = 2048 – we use the WMAP intensity mask for masking as well as the WMAP power spectrum up to \( l = 200 \) as the variance of the isotropic and Gaussian \( a_{lm} \). Although a sizable residual effect of the dipole can be seen in the maps, the entropies get modified only slightly. Clearly, at \( l = 1 \) the entropies show the residual part of the dipole, and also at higher \( l \) the curves are distorted a little, but the described procedure does not impose any large anomalies and especially it does not result in conspicuous values at \( l = 5, 16, 17, 28, 30 \).

Thus, we conclude that with our simplified approach no sizable mixing of dipole power to higher multipoles via the masking process can be observed. Finally note that other masking processes with Fourier methods were also applied to the angular and projection pseudo entropies in [44] and more extensively in [45][46].

TABLE II: P-values of angular pseudo entropy for most conspicuous large angle multipoles, calculated with 10000 ensembles of Gaussian \( a_{lm} \), rounded to two decimal places.

| \( l \) | pipeline | \( p \) [%] | \( p_{geom} \) [%] (excluding SEVEM) |
|-------|---------|------------|-----------------------------|
| 5     | NILC    | 1.03       | 0.92                        |
|       | COMMANDER | 1.35       |                             |
|       | SMICA   | 0.99       |                             |
|       | SEVEM   | 16.80      |                             |
|       | WMAP    | 0.53       |                             |
| 16    | NILC    | 1.52       | 1.89                        |
|       | COMMANDER | 2.01       |                             |
|       | SMICA   | 3.58       |                             |
|       | SEVEM   | 32.10      |                             |
|       | WMAP    | 1.16       |                             |
| 17    | NILC    | 1.66       | 2.35                        |
|       | COMMANDER | 1.06       |                             |
|       | SMICA   | 3.82       |                             |
|       | SEVEM   | 19.50      |                             |
|       | WMAP    | 4.36       |                             |
| 28    | NILC    | 0.07       | 0.26                        |
|       | COMMANDER | 0.65       |                             |
|       | SMICA   | 1.13       |                             |
|       | SEVEM   | 4.47       |                             |
|       | WMAP    | 0.09       |                             |
| 30    | NILC    | 1.29       | 0.63                        |
|       | COMMANDER | 0.32       |                             |
|       | SMICA   | 0.45       |                             |
|       | SEVEM   | 1.93       |                             |
|       | WMAP    | 0.84       |                             |
Comparison of 2015 and 2018 data with angular pseudo entropy

In the following we compare the angular pseudo entropy of 2015 Planck data to the newest 2018 data release. Since the 2018 component separation process has been optimized for polarization data, it is expected to come equipped with a few drawbacks in temperature maps, especially for COMMANDER, which carries more residual foreground contamination in the 2018 than in the 2015 temperature map. One should expect to see this feature in the angular pseudo entropy and indeed Fig. 13, which shows the comparison of 2015 and 2018 COMMANDER angular entropy on the two ranges considered in Sect. V B as well as Fig. 17 which shows the relative deviation of 2018 to 2015 data for all Planck foreground cleaned full sky maps, confirm this expectation. While even for large angular scales the deviation of COMMANDER is larger than that of SMICA and NILC, for small angular scales COMMANDER drops even below SEVEM. Since in our work we do not want to mask the maps, but need to work with full sky data, it becomes obvious that for our purposes the 2015 COMMANDER temperature data should be preferred to the 2018 data.

Both NILC and SMICA show only a slight deviation in 2018 compared to 2015, both on small and large angular scales with NILC 2018 entropy being identical to the 2015 entropy with at most 5% deviation, which is reached only at \( l = 28 \), see Figs. 14, 16 and 17. For all other multipoles the NILC deviation is nearly negligible. Since the NILC component separation process has been left nearly unaltered from 2015 to 2018, NILC is most useful for observing the effects of the improved Cosmic Dipole calculation and the removed AD non-linearity. The influence of the former is restricted mainly to a very slight reduction of significance of the both most unlikely multipoles we considered, namely \( l = 28, 896 \).

SEVEM has been clearly enhanced in 2018, as shown in Figs. 15 and 17. In 2015 data the angular entropy of SEVEM was far too low from \( l = 13 \). That behavior came particularly clear at small angular scales. In the preceding sections we argued that this effect stems from the residual contamination of SEVEM data by the galactic plane. In 2018 the entropy is constantly shifted to higher values from \( l = 13 \) on, approaching a nearly constant relative improvement of about 18% at small angular scales. Nevertheless for our purposes the SEVEM map still lacks quality at small angular scales and visible residual foreground pollution is left.

At large angular scales COMMANDER and SMICA exhibit a joint deviation behavior at the most unlikely multipoles. Both large entropy values at \( l = 5 \) and \( l = 28 \) are slightly suppressed, but still outside of 2\( \sigma \), and the small entropy value at \( l = 30 \) is enlarged. While for SMICA all three multipoles still lie outside of 2\( \sigma \), COMMANDER shifts them towards smaller confidence. In contrast to that, the two multipoles \( l = 16, 17 \) get shifted to more unlikely values in both maps. These considerations show that the most conspicuous multipoles at large angular scales could partly be caused by unoptimized component separation, but that AD non-linearity and the Cosmic Dipole identification show only a minor effect, since NILC is nearly unaltered.

At small angular scales we can only use NILC and SMICA since SEVEM is still off and COMMANDER has been degraded due to polarization optimization. The main observation concerning single multipoles, that can be made at this point, is that for SMICA the most unlikely multipole \( l = 896 \) on the considered range is improved but the previously normal multipole \( l = 910 \) is shifted towards 2\( \sigma \).

In Tab. IV in App. B we gather p-values and likelihoods for 2015 and 2018 data. There we also include the range \([890, 910]\) in order to compare it to \([895, 905]\). Since COMMANDER is off at small angular scales in 2018, we
also consider the geometric mean of NILC and SMICA alone. The range \([2, 30]\) is normal in both data releases with 2015 being slightly better than 2018 both for NILC alone and for the geometric mean of COMMANDER, NILC and SMICA. Fig. 13 shows that even though the geometric mean of p-values is smaller than the expectation, the significance for that is too low and hence we can conclude that using the angular entropy method the whole range \([2, 30]\) is compatible with the assumption of isotropic and Gaussian temperature fluctuations, which was also pointed out in [27] for the range \([2, 50]\).

In contrast the range \([895, 910]\) displays unlikely behavior in both releases. We observe a slight enhancement from 4.4% to 6.2% in the geometric mean of p-values for NILC and from 5.3% to 8.1% in the geometric mean of p-values for the geometric mean of NILC and SMICA. These enhancements correspond to changes in likelihood from 0.1% in 2015 to 0.5% in 2018 for the latter and no change of likelihood for the former. None of thousand random isotropic and Gaussian maps admits such a low geometric mean of p-values as NILC in both 2015 and 2018, hence the Likelihood is bounded from above by \(\approx 0.1\%\) in both releases. Enlarging the range a bit from \([895, 905]\) to \([890, 910]\) increases the likelihoods about a factor of 1 to 15, but keeps them under 2.5%. We can conclude that the improvements made in the 2018 data processing improve also the small angular scales, but the features are still at nearly 2\(\sigma\) for the mean of NILC and SMICA at \([890, 910]\) and at or outside of 3\(\sigma\) for NILC at \([895, 905]\). At this point we should clarify again that the question, how likely it is to find a range of such width outside of 3\(\sigma\) is postponed to the future and that here we might fall for the selection effect.
D. Results for range angular pseudo entropy with 2015 data

The range angular pseudo entropy provides an additional measure for quantifying unlikeliness of multipole ranges and also collections of different multipoles which are not necessarily in a row, see Eqs. (51)-(54) in Sect. III. Unfortunately it is a mix of a correlation and an averaging measure, hence one needs to consider both the range angular and the single multipole angular entropies in order to identify effects of correlation of different multipoles, which are usually expressed by small range entropies. On the other hand, if one is solely interested in the mean likelihood of a given range, the geometric mean of p-values of the single multipole angular entropy is surely the better measure. Aside the partial correlation interpretation, the big advantage of the range angular entropy is its pseudo entropy nature and henceforth its interpretation as an entanglement measure. In Tab. V in App. B we gather p-values for the range angular entropy for various ranges on large angular scales and for the small scale range [895, 905], as well as the signed deviation of the entropy for 2015 NILC from the isotropic and Gaussian expectation.

One directly observes that the range entropy of the range [2, 3] is too small at $\approx 1\%$ p-value and $l = 6$ too low at $\approx 7\%$ p-value, the only unlikely collection of
The range \([2, 30]\) calculated with 10000 ensembles.

FIG. 18: Probability density of geometric mean of p-values and respective values for NILC and the geometric mean of maps for 2018 data.

b) Range \([895, 905]\) calculated with 1000 ensembles.

FIG. 18: Probability density of geometric mean of p-values and respective values for NILC and the geometric mean of maps for 2018 data.

The range \([2, 30]\) yields a p-value of around 18% which is compatible with the likelihood of about 15% calculated with the geometric mean of p-values for the single multipole angular entropy. The range entropy lies slightly below the expectation indicating a mixture of a slight averaged preference for a direction on the sky and a slight correlation of multipoles, though being within 1σ.

For the collection \(\{5, 28\}\) the p-value is smaller than one would expect from the average of both multipoles indicating correlation of these two multipoles.

On small angular scales the p-value 0.5% for the range \([895, 905]\) is compatible with the likelihood from the geometric mean, which we gave the approximate upper bound 0.1%. For both ranges \([2, 30]\) and \([895, 905]\) we obtain slightly larger p-values with the range entropy than likelihoods with the geometric mean, which could be caused by reduction of significance due to averaging of large and small values of the angular entropy. The fact that the range entropy lies below the expectation again indicates, that a direction might be preferred in the data and/or different multipoles might be correlated.

V. SUMMARY AND DISCUSSION

Building upon the Wehrl entropy we introduced three types of pseudo entropies which approximate the Wehrl entropy but allow for much faster computation and hence analyzing CMB data up to \(l = 1000\). Those entropies are the \(j\)-projection, the quadratic and the angular entropy. While the quadratic is simple in fashion it should be disregarded, because of its \(x^2\) instead of \(x \log(x)\) behavior and the numerical problems one runs into with it for high \(l\) if not approximating the expression. All pseudo entropies are rotationally invariant measure of quantum randomness or entanglement on spin-\(l\) states and hence on each multipole of CMB temperature fluctuations on the sphere. Contrary to the usual von Neumann entropy these pseudo entropies do not vanish for pure states. In the spirit of thermodynamics, the entropies are useful for reducing \(2l + 1\) degrees of freedom per multipole to a single number per multipole just as one usually does with \(C_l\) but which complements it in the case of anisotropies or non-Gaussianities.

We showed that for \(l = 2\) both the Wehrl and the angular entropy depend only on the squared chordal distance of multipole vectors, yielding another view on this method and a connection to many previous studies of CMB analysis.

Although our focus was on introducing the methods and clarifying their properties, in order to demonstrate the usage of these methods we applied the introduced types of pseudo entropies for analysis of CMB temperature full sky maps and it turned out that they all show similar behavior and the same characteristic features of the maps, except for the quadratic pseudo entropy. Since the angular entropy is the computationally cheapest measure, the rest of the analysis was devoted solely to the angular entropy, which reaches its maxi-
mum $\log(3)$ for maximally mixed states, which cannot be reached by pure temperature maps, and its minimum probably for coherent states; it is mathematically known for sure for $l = 1/2, 1$ only. The physical data from the Planck 2015 maps and WMAP LLC was compared to isotropic and Gaussian maps, and some multipoles with particularly small $p$-values were found, in particular $l = 5, 16, 17, 28, 30$ on large angular scales, and the range $895 \leq l \leq 905$ for NILC, the likelihood of which we approximately bounded from above by 0.1%, but there could be as well further unusual angular scales and we did not take into account the selection effect statistically. On average three out of four Planck maps do not show abnormal global behavior for $l \leq 1000$, that means deviations from isotropy and Gaussianity on a large range of scales, and the abnormality of the fourth map – SEVEM – can be removed by masking the galactic plane. As expected, the Planck maps can be considered as a clear improvement compared to WMAP on small angular scales $l > 200$.

A comparison of isotropic, Gaussian random maps to maps constructed from uniform multipole vectors showed that our method is sensitive to deviations from Gaussianity and, due to its rotationally invariant nature, also isotropy. One should note that uniform multipole vectors are clearly distinguishable from isotropic and Gaussian maps for which the multipole vectors exhibit repulsion. We were not able to identify the data processing as a reason for all or some of the mentioned conspicuous multipoles with a simple masking approach. The fact that these multipoles have low $p$-values in all of the maps except for SEVEM indicates a different reason behind them. Nevertheless, one should not withhold that a statistical fluke cannot be excluded, even if the $p$-value for $l = 28$ lies below half a percent.

Comparing Planck 2015 to 2018 data confirmed the expectation that the COMMANDER full sky 2018 map cannot be used for our methods at large angular scales without masking. SEVEM has been enhanced from 2015 to 2018 but still carries too much foreground pollution in it. SMICA only deviates slightly and NILC is nearly unaltered. The unlikely features we observe are present in both data releases but with slightly less significance in 2018 than in 2015. The fact that NILC is left nearly unchanged suggests that AD non-linearity and unoptimized Cosmic Dipole removal do not account for the observed features. Nevertheless the component separation still might do.

Eventually we considered the angular range entropy as a mixture of a measure of range-or-collection-averaged angular entropy and of correlation between different multipoles. It turned out that the results for the ranges $[2, 30]$ and $[895, 905]$ are consistent with the likelihoods obtained from the geometric mean of $p$-values of the standard angular entropy. We found the anti-correlation of the quadrupole with $l = 5$ and the correlation of $l = 2, 3$ especially interesting. While the latter supports the previously observed quadrupole-octupole correlation, the former hints towards a connection between the high angular entropy value at $l = 5$ and the CMB dipole, which itself has been proposed to be influenced mainly by the Cosmic Dipole in the past.

There are several tasks left for the future. First of all a deeper statistical analysis needs to be done, especially considering wider and smaller angular scale ranges. Our analysis has shown, that there might be something hidden at small angular scales and it would be interesting to see more results on this. So far we have introduced the methods, explained some of their mathematical behavior and performed a perfunctory analysis in order to illustrate their usage. Furthermore one could try to apply some of these methods to polarization data. For the Wehrl entropy it is clear that a direct generalization is possible, but for the angular entropy further analysis is needed. Angular and Wehrl entropy are also useful in quantum information theory as real entanglement measures. Finally, our work is also interesting from a mathematical perspective. The fact that the Wehrl entropy is minimized in general by SU(N)-coherent states is known as the Lieb conjecture and has been proven [35]. The same question is, however, still open for the angular entropy.

ACKNOWLEDGMENTS

We acknowledge financial support by the DFG Research Training Group 1620 Models of Gravity. We thank Valentin Buciumas, Jessica Fintzen, Otgonbaatar Myagmar, and Gregor Mittelstaedt for collaboration at earlier stages of the research leading to this publication. Useful discussions with Elliott Lieb are gratefully acknowledged.

[1] Fixsen, D. (2009) The Temperature of the Cosmic Microwave Background. The Astrophysical Journal, 707, 916–920.
[2] Adam, R. et al. (2016) Planck 2015 results. I. Overview of products and scientific results. Astron. Astrophys., 594, A1.
[3] Ade, P. A. R. et al. (2016) Planck 2015 results. XIII. Cosmological parameters. Astron. Astrophys., 594, A13.
[4] Akrami, Y. et al. (2018) Planck 2018 results. I. Overview and the cosmological legacy of Planck.
[5] Aghanim, N. et al. (2018) Planck 2018 results. VI. Cosmological parameters.
[6] Ade, P. A. R. et al. (2016) Planck 2015 results. XVI. Isotropy and statistics of the CMB. Astron. Astrophys., 594, A16.
[7] Ade, P. A. R. et al. (2016) Planck 2015 results. XVII. Constraints on primordial non-Gaussianity. Astron. Astrophys., 594, A17.
Appendix A: System resources

- Model: OptiPlex 790, Dell Inc., Version 01, 64 bits
- CPU: IntelCore i5-2400, 3.10 GHz, 1 CPU, 4 Cores, 4 Threads
- Cache: L1 – 256 kB, L2 – 1 MB, L3 – 6 MB
- RAM: 4 GB
- GFlops (tested with linpack): From 50.9 up to 76.2

Appendix B: Additional plots and tables

FIG. 19: Angular pseudo entropy plotted unlogarithmicly between $l = 1$ and $l = 30$. Confidence levels have been smoothed.

| quantity | a   | b   | c   |
|----------|-----|-----|-----|
| 1σ_upper | 0.658 | 0.999 | -   |
| 1σ_lower | 0.663 | 0.929 | -   |
| 2σ_upper | 1.672 | 1.742 | -   |
| 2σ_lower | 1.198 | 0.477 | -   |
| 3σ_upper | 2.982 | 68.283 | -   |
| 3σ_lower | 1.654 | 0.345 | -   |
| exp      | 0.948 | 1.545 | 0.976 |

TABLE III: Fitted parameters of angular pseudo entropy rounded to 3 decimals.
TABLE IV: Comparison between 2015 and 2018 of geometric mean of p-values and likelihoods for all three ranges calculated for NILC, the geometric mean of COMMANDER, NILC and SMICA and the geometric mean of only NILC and SMICA. Due to 10000 ensembles used we take two digits for large angular scales and due to 1000 ensembles used we take one digit for small angular scales.
a) Map from isotropic and Gaussian $a_{lm}$ with power spectrum (up to $l = 200$) of WMAP 7-year ILC, smoothed

b) Map after induction of dipole

c) Map after induction of dipole and masking with WMAP intensity mask

d) Map after removing dipole again with Healpy remove_dipole

e) Map after removing dipole and filling masked region with data from a)

FIG. 21: Process of inducing a dipole, masking the map, removing the dipole and filling the masked region.
| Range | p-value [%] | Entropy | Expectation | Above/Below |
|-------|-------------|---------|-------------|-------------|
| [2,3] | 2.0         | 0.820   | 0.960       | -           |
| [2,4] | 44.3        | 1.015   | 1.016       | -           |
| [2,5] | 41.4        | 1.057   | 1.045       | +           |
| [2,6] | 55.9        | 1.055   | 1.054       | +           |
| [3,4] | 27.5        | 1.030   | 0.991       | +           |
| [3,5] | 20.9        | 1.066   | 1.035       | +           |
| [3,6] | 40.2        | 1.060   | 1.047       | +           |
| [4,5] | 11.8        | 1.067   | 1.015       | +           |
| [4,6] | 36.0        | 1.056   | 1.038       | +           |
| [5,6] | 20.3        | 1.051   | 1.007       | +           |
| {2,5} | 3.2         | 1.077   | 0.998       | +           |
| {3,5} | 51.6        | 1.013   | 1.008       | +           |
| {4,6} | 30.6        | 0.990   | 1.007       | -           |
| 5     | 1.03        | 1.080   | 0.988       | +           |
| 6     | 6.95        | 0.831   | 1.004       | -           |
| [2,20] | 17.9       | 1.0847  | 1.0857      | -           |
| [2,30] | 18.4       | 1.0789  | 1.0893      | -           |
| [26,27] | 40.8       | 1.085   | 1.080       | +           |
| [26,28] | 16.6       | 1.093   | 1.086       | +           |
| [26,29] | 24.7       | 1.094   | 1.090       | +           |
| [26,30] | 19.9       | 1.088   | 1.092       | -           |
| [27,28] | 6.4        | 1.094   | 1.081       | +           |
| [27,29] | 12.2       | 1.094   | 1.087       | +           |
| [27,30] | 17.3       | 1.086   | 1.090       | -           |
| [28,29] | 3.9        | 1.095   | 1.082       | +           |
| [28,30] | 23.7       | 1.084   | 1.088       | -           |
| {29,30} | 4.4        | 1.062   | 1.083       | -           |
| {5,28} | 2.3        | 1.084   | 0.986       | +           |
| 28     | 0.07        | 1.097   | 1.074       | +           |
| [895,905] | 0.5       | 1.09827 | 1.09851     | -           |

TABLE V: Range angular pseudo entropy: The second column shows p-values for different multipole ranges and collections, the third and fourth columns show the value of the range angular pseudo entropy for NILC and the expectation using 10000 ensembles of isotropic and Gaussian random maps, the fifth column indicates if the NILC entropy lies below and above the expectation, which is important for the interpretation of the results. The single multipole angular entropy values for $l = 5, 6$ are included for comparison with the ranges that include both multipoles.