

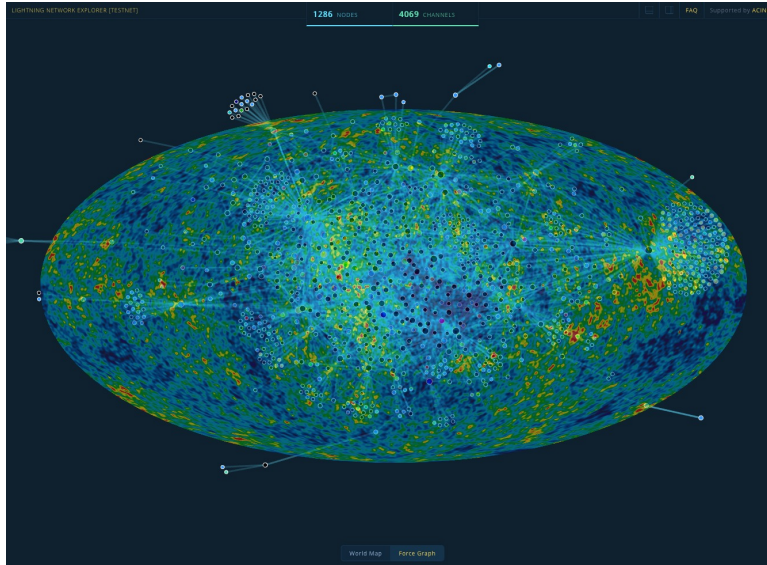
CMBGraph: A new data representation of the Cosmic Microwave Background

J.B.G. Alvey

Theoretical Particle Physics and Cosmology, King's College London

Abstract

The most precise measurement of the Cosmic Microwave Background (CMB) comes from space-bound satellites such as WMAP ([Hinshaw et al., 2013](#)) and the Planck satellite ([Aghanim et al., 2018](#)). This consists of temperature and polarisation data measured against a uniform temperature background of approximately 2.73 K. An important quantity that can be measured from the maps is the angular power spectra. These play a high-precision role in constraining theories of inflation and cosmological evolution. This letter considers a new representation of this data in terms of a planar graph which aims to encode the angular correlation function in a more complete way compared to the power spectrum. We denote this type of structure a *planar correlation graph*. More recently ([Battaglia et al., 2018](#)), there has been interest in a new network architecture for Deep Learning known as a Graph Network. Input to this network is in the form of a planar graph. Whilst we are not advocating the use of a Graph network at this stage for Cosmology, the abstraction of the CMB data to a higher dimensional graph is a natural step forwards. The weighting of the edges of the graph using the power spectrum is then a physically motivated choice for an element of the architecture that is user-defined.



I The Ising Model in 1D

The Ising model is the prototypical example of a statistical system whose correlations and fluctuations lead to non-trivial phenomenology beyond that suggested by the mean field solution. We will use this example as a warm up to the case of the full CMB. We will illustrate how to construct the planar graphs based on the correlation function which is simple to state in this setup. In one dimension, the Ising model is a set of spins, $\{\sigma_i\}$ on a line of sites labelled $i = 1, 2, \dots, N$. Note that we are not necessarily assuming periodic boundary conditions in this case. The Hamiltonian for the system is then given by,

$$\mathcal{H} = -J \sum_{i=1}^N \sigma_i \sigma_{i+1} - H \sum_{i=1}^N \sigma_i, \quad (1.1)$$

from which we can compute the canonical partition function,

$$Z_N = \sum_{\{\sigma_i\}} \exp \left(\beta J \sum_{i=1}^N \sigma_i \sigma_{i+1} + \beta H \sum_{i=1}^N \sigma_i \right). \quad (1.2)$$

Now, for a given system, J and H are typically fixed and characterise the strength of the interactions and the energy of the spin state due to an external field respectively. Furthermore, $\beta = 1/T$ can vary with the temperature.¹ The key physics of the Ising model in our context is the fact that correlations/fluctuations between the spin sites destroy the long range order that would be characteristic of a phase transition at low temperature. An argument by Landau demonstrated that indeed there is no phase transition in the 1D Ising model despite the mean field solution suggesting otherwise. When $H = 0$, it is possible to calculate the correlation function (which is discrete in this case), and one finds,

$$\langle \sigma_i \sigma_j \rangle = \frac{1}{Z_N} \sum_{\{\sigma_i\}} \sigma_i \sigma_j \exp(-\beta \mathcal{H}) = \exp(|i - j| \log(\tanh \beta J)). \quad (1.3)$$

The behaviour of the correlation as a function of the site separation is shown in Figure ?? for $N = 20$ sites. This is now in a suitable form to illustrate the concept of a planar correlation graph. In the case of the Ising model it is constructed as follows,

1. To each lattice site $i = 1, 2, \dots, N$, associate a vertex v_i on the graph, G .
2. At each vertex v_i , store the information (i, σ_i) in a suitable data structure.
3. Between vertices v_i and v_j , draw an edge e_{ij} weighted according to the correlation function $\langle \sigma_i \sigma_j \rangle$ as computed above.

This procedure will lead to a planar graph $G = \{v_i, e_{ij}\}$ along with a set of weights $\{w_{ij} = \langle \sigma_i \sigma_j \rangle\}$, as well as a map $\mathfrak{D} : V \rightarrow \mathcal{M}$. Here \mathfrak{D} should be viewed as a function from the set of vertices V to a suitable data manifold, \mathcal{M} . In the case of the Ising model for example,

$$\mathfrak{D} : V \rightarrow \mathbb{Z} \times \{-1, 1\}, \quad \mathfrak{D}(v) \mapsto (i, \sigma_i). \quad (1.4)$$

For the CMB that we will consider later, this map will instead be of the form, $\mathfrak{D} : V \rightarrow S_2 \times \mathbb{R}$, $\mathfrak{D}(v) \rightarrow (\hat{\mathbf{n}}, \delta T(\hat{\mathbf{n}})/T_{\text{CMB}})$. We formalise this with the following definition;

Definition 1.1. *A planar correlation graph is a planar graph $G = \{V, E\}$, formed of a set of edges $E = \{e_{ij}\}$, and a set of vertices $V = \{v_i\}$ together with;*

1. *A set of weights w_{ij} that characterise the correlation between the data at the vertex v_i and that at v_j .*
2. *A map $\mathfrak{D} : V \rightarrow \mathcal{M}$, where \mathcal{M} is a suitable data manifold. This places the role of ‘storing’ the data at each vertex (e.g. CMB temperature/polarisation or the spin in the Ising model).*

¹ $k_B = 1$ here.

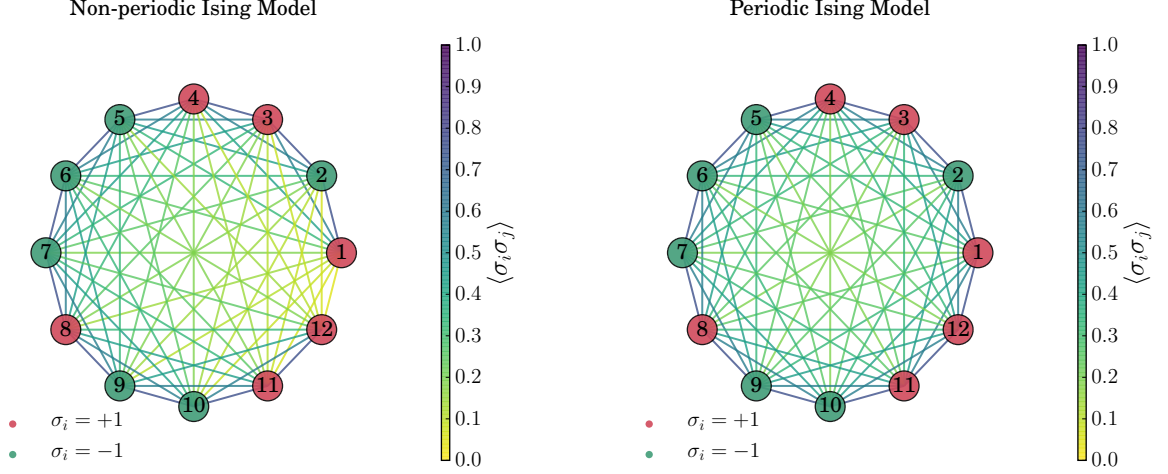


Figure 1.1 The planar correlation graphs in the case of the Ising model for $N = 12$ sites, with $\beta J = 1$. This is shown for non-periodic (Left) and periodic (Right) boundary conditions

We now apply this idea to the explicit example of the Ising model on $N = 12$ lattice sites for both periodic and non-periodic boundary conditions, with $\beta J = 1$. Labelling the nodes as $i \in \{1, 2, \dots, N\}$, we randomly sample an initial spin state $\{\sigma_i\}$ and display the information via the colour of the nodes. Furthermore, we assign the weight $w_{ij} = \langle \sigma_i \sigma_j \rangle$ between site i and site j and indicate this via the colour of the edge on the graph. The results are shown in Figure 1.1. We see that the graphs encode the different correlation graphs illustrate the difference between the two sets of boundary conditions via the weights.

II The Cosmic Microwave Background

We used the Ising model as an example of how to construct a planar correlation graph. We now wish to extend this idea to the CMB. To understand the correlations of the cosmic microwave background, we should first understand the expansion of a general random field, $f(\hat{\mathbf{n}})$, on a sphere. For such a field, we can expand $f(\hat{\mathbf{n}})$ in terms of spherical harmonics, $Y_{lm}(\hat{\mathbf{n}}) = Y_{lm}(\theta, \phi)$ as follows;

$$f(\hat{\mathbf{n}}) = \sum_{l=0}^{\infty} \sum_{m=-l}^l f_{lm} Y_{lm}(\hat{\mathbf{n}}), \quad (2.1)$$

where the set of f_{lm} are the expansion coefficients. The Y_{lm} are orthogonall in the sense that,

$$\int d\hat{\mathbf{n}} Y_{lm}(\hat{\mathbf{n}}) Y_{l'm'}^*(\hat{\mathbf{n}}) = \delta_{ll'} \delta_{mm'}, \quad (2.2)$$

so we can extract the expansion coefficients via,

$$f_{lm} = \int d\hat{\mathbf{n}} f(\hat{\mathbf{n}}) Y_{lm}^*(\hat{\mathbf{n}}). \quad (2.3)$$

Now, in the context described in this paper where we wish to abstract the real space, map representation of the CMB, we are interested in spatial correlations. To compute these, we consider

$$\langle f(\hat{\mathbf{n}}) f^*(\hat{\mathbf{n}}') \rangle = \sum_{lm} \sum_{l'm'} \langle f_{lm} f_{l'm'}^* \rangle Y_{lm}(\hat{\mathbf{n}}) Y_{l'm'}^*(\hat{\mathbf{n}}'). \quad (2.4)$$

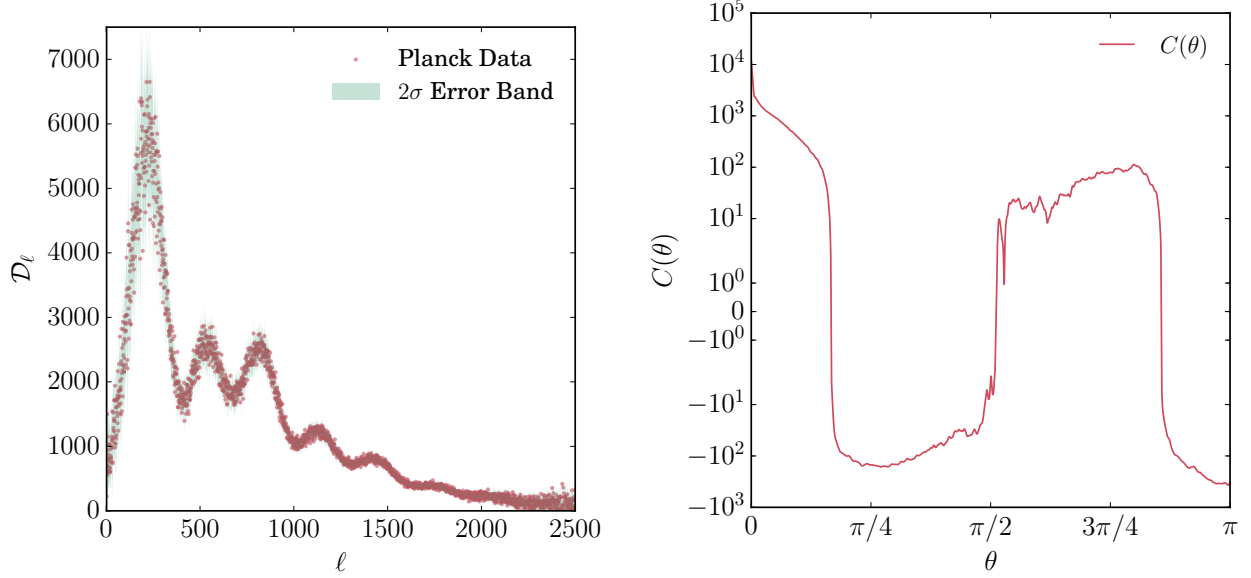


Figure 2.1 Plots of the correlation coefficients, \mathcal{D}_l (Left) and the correlation function, $C(\theta)$ (Right).

Rotational invariance of the CMB (in our case) implies that $\langle f_{lm} f_{l'm'}^* \rangle$ must be of the form

$$\langle f_{lm} f_{l'm'}^* \rangle \equiv C_l \delta_{ll'} \delta_{mm'}. \quad (2.5)$$

So, we can compute the two point correlation function as,

$$\begin{aligned} \langle f(\hat{\mathbf{n}}) f^*(\hat{\mathbf{n}}') \rangle &= \sum_{lm} \sum_{l'm'} \langle f_{lm} f_{l'm'}^* \rangle Y_{lm}(\hat{\mathbf{n}}) Y_{l'm'}^*(\hat{\mathbf{n}}') \\ &= \sum_l C_l \sum_m Y_{lm}(\hat{\mathbf{n}}) Y_{lm}^*(\hat{\mathbf{n}}') \\ &= \sum_l \frac{l(l+1)}{4\pi} C_l P_l(\hat{\mathbf{n}} \cdot \hat{\mathbf{n}}'), \end{aligned}$$

where we have made use of a standard identity relating the Legendre polynomials, $P_l(\hat{\mathbf{n}} \cdot \hat{\mathbf{n}}') = P_l(\cos \theta)$ and the spherical harmonics, $Y_{lm}(\hat{\mathbf{n}})$. We will now work with the CMB data and consider the field $f(\hat{\mathbf{n}}) = \delta T(\hat{\mathbf{n}})/T_{\text{CMB}}$. The Planck analysis (Aghanim et al., 2018) provides the correlation coefficients C_l in their analysis. From these, we can extract the correlation function, $C(\theta)$, which measures the correlation on some angular scale θ .² This is given by,

$$\left\langle \frac{\delta T(\hat{\mathbf{n}})}{T_{\text{CMB}}} \frac{\delta T(\hat{\mathbf{n}}')}{T_{\text{CMB}}} \right\rangle = C(\theta) = \sum_l \frac{2l+1}{4\pi} C_l P_l(\cos \theta). \quad (2.6)$$

Plots of both $\mathcal{D}_l := l(l+1)C_l/2\pi$ as well as $C(\theta)$ are shown in Figure 2.1, illustrating the characteristic shapes. We will be mainly interested in $C(\theta)$ as a proxy by which to weight our planar graph.

III Graph Networks

The computations above, whilst providing perhaps an interesting representation of the data in their own right, were mainly done to provide input to a graph network. Learning based inference

²Here, θ satisfies $\hat{\mathbf{n}} \cdot \hat{\mathbf{n}}' = \cos \theta$.

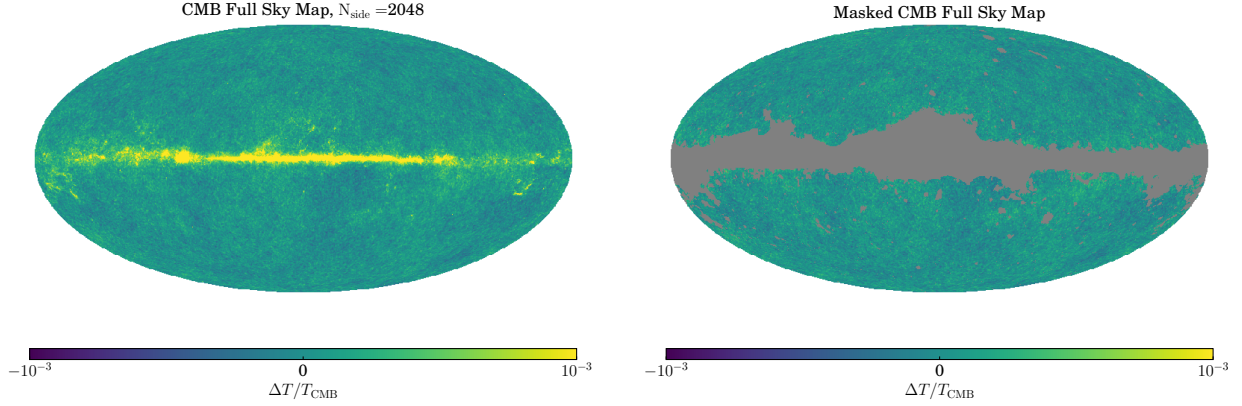


Figure 2.2 The Planck PR3 100GHz Intensity Map of the CMB (Left) and the masked version, using the 2018 Intensity mask (Right).

on such structured input has been recently pioneered as a means to push artificial intelligence forward (Battaglia et al., 2018). The collective reference for this is *combinatorial generalisation*, where the system in question is biased towards certain, known, structural input. In our example, we make use of the natural correlation structure of the Ising model and the CMB pixel data.

Acknowledgements

This work is based on observations obtained with Planck (<http://www.esa.int/Planck>), an ESA science mission with instruments and contributions directly funded by ESA Member States, NASA, and Canada.

References

- G. Hinshaw et al. (WMAP), *Astrophys. J. Suppl.* **208**, 19 (2013), [1212.5226](#).
- N. Aghanim et al. (Planck) (2018), [1807.06209](#).
- P. W. Battaglia, J. B. Hamrick, V. Bapst, A. Sanchez-Gonzalez, V. F. Zambaldi, M. Malinowski, A. Tacchetti, D. Raposo, A. Santoro, R. Faulkner, et al., *CoRR* (2018), [1806.01261](#).