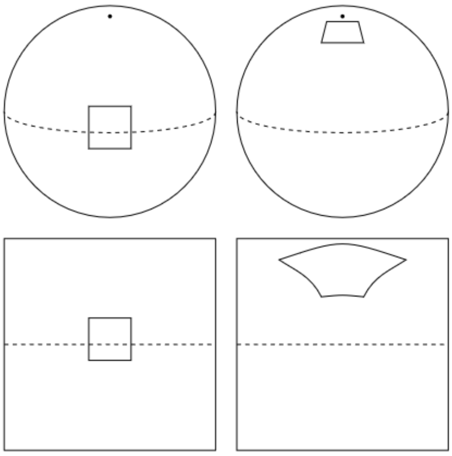
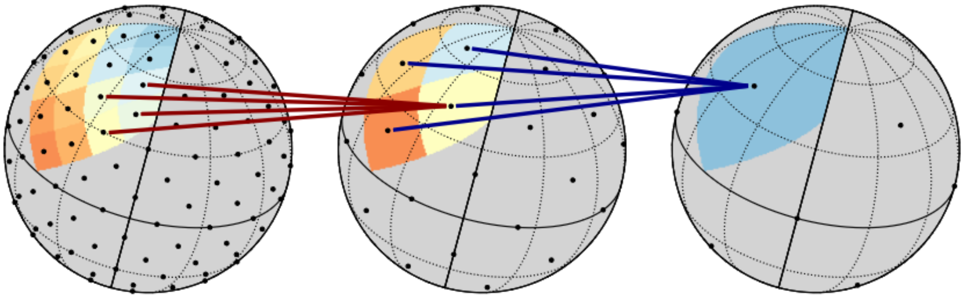
Deep Sphere & Stochastic Weight Averaging for likelihood-free Bayesian Parameter Estimation:

0. Regular CNNs do not learn rotationally invariant filters



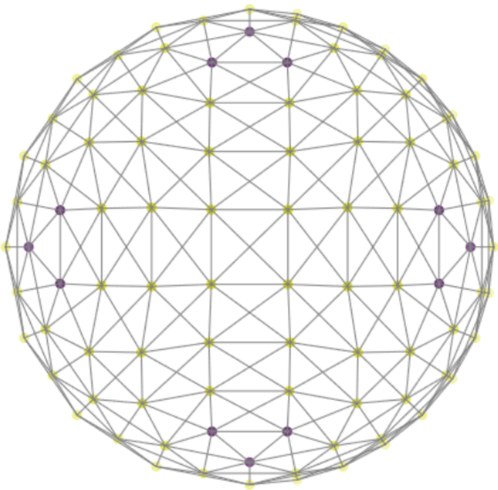
(https://arxiv.org/pdf/1801.10130.pdf)

1. Use Spherical Convolutions to obtain rotational Invariance:



(https://arxiv.org/pdf/1810.12186.pdf)

2. Performing Convolutions after a spherical Harmonic transform is slow so to use this method on large maps we construct neighborhood Graph on the Healpix sphere first and learn filters on the graph. (This also does work for irregular parts in the Graph)



This allows us to incorporate the geometric structure of the data directly into our models. Also (https://arxiv.org/pdf/1810.12186.pdf) note, that the Eigenvectors of the Graph Laplacian approximate the Spherical Harmonics, which is a strong indicator that the model architecture fits our problem.

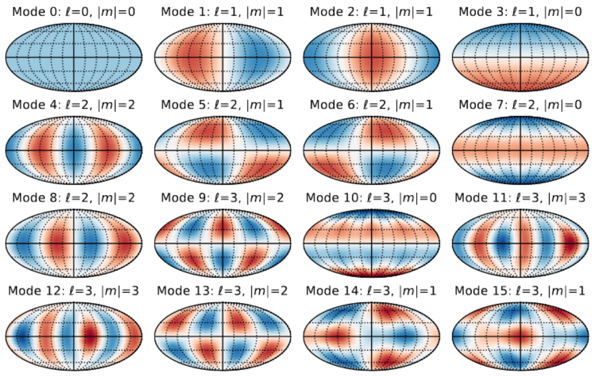


Figure 3: The first 16 eigenvectors of the graph Laplacian, an equivalent of Fourier modes, of a graph constructed from the HEALPix sampling of the sphere ($N_{side} = 16$). Eigenvectors 1–3 could be associated with spherical harmonics of degree $\ell = 1$ and order $|m| = (0, 1)$, eigenvectors 4–8 with degree $\ell = 2$ and order $|m| = (0, 1, 2)$, and eigenvectors 9–15 with degree $\ell = 3$ and order $|m| = (0, 1, 2, 3)$. Nevertheless, graph eigenvectors are only approximating spherical harmonics.

We can use this to directly estimate the cosmological parameters from the maps by training, for example, on a simple L2 objective.
How do we obtain a measure of uncertainty from this?
Stochastic Weight Averaging!
Pretrain a model on lots of data, then run Stochastic Gradient Descent with a constant learning rate schedule to “explore” the parameter space of the model. This leads to better robustness of the model and better generalization capabilities and allows us to obtain posterior credibility intervals for the predictions of the model.
(https://arxiv.org/pdf/1902.02476.pdf)
We also can compare this to variational methods or Dropout for an overview of the obtained intervals.