Sparse Isotropic Regularization for Spherical Harmonic Representations of Random Fields on the Sphere

Yu Guang Wang
School of Mathematics and Statistics
University of New South Wales

Joint with Quoc T. Le Gia, Ian Sloan, Rob Womersley (UNSW)

On the Frontiers of High Dimensional Computation
MATRIX 2018
Cosmic Microwave Background

- **Cosmic Microwave Background (CMB)** is a black-body radiation from the **recombination epoch** of the early universe, around 13.7 Byr ago.

- **CMB** is critical to understanding of early universe and evidence of the **Big Bang theory**. In particular, **measurement of CMB fluctuations** is key to verification and estimation of the cosmological properties and parameters.

- **CMB** was discovered by **Arno Penzias** and **Robert Wilson** in 1964 (Nobel Prize in Physics 1978), since then **NASA** and **ESA probes** have been collecting CMB data.
CMB field on $S^2$

$$\widetilde{T}(\hat{p}, p_0, t) = T_0(p_0, t)(1 + T(\hat{p}, p_0, t))$$

- $\hat{p}$ is the direction the CMB signal comes from.
- $p_0$ is the observation location, near the planet earth.
- $t$ is the time when the CMB signal was received, Planck Space Observatory in 2009.
- $T_0(p_0, t)$ is around $2.73K$ at present.
- $T(\hat{p}) := T(\hat{p}, p_0, t)$ is modelled as a random field on $S^2$. 
Random fields on $\mathbb{S}^2$

- Probability measure space $(\Omega, \mathcal{F}, P)$

- An $\mathcal{F} \otimes B(\mathbb{S}^2)$-measurable function $T(\omega, x) : \Omega \times \mathbb{S}^2 \to \mathbb{R}$ is called a *real-valued random field* on the sphere $\mathbb{S}^2$.

- Space of square integrable functions on product space $L_2(\Omega \times \mathbb{S}^2, P \otimes \sigma)$
Isotropy of random field

- **(Strongly) Isotropy** of a random field $T$ means, for any $n$ points $x_1, \ldots, x_n$ and any rotation $\rho \in SO(3)$,

\[
(T(x_1), \ldots, T(x_n)) \sim (T(\rho x_1), \ldots, T(\rho x_n))
\]

in distribution.

- **Isotropy** means the field does not change (its distribution) when the coordinate system changes, which is consistent with the physics intuition and is a fundamental hypothesis for the CMB field.
For each realization, we can use spherical harmonics to represent the random field, in $L_2(\Omega \times S^2)$ sense,

$$T^o(\omega, x) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \left(\hat{T}^o\right)_{\ell m}(\omega) Y_{\ell m}(x), \quad (\omega, x) \in \Omega \times S^2.$$

- The $Y_{\ell m}$ is a spherical harmonic orthonormal basis for $L_2(S^2)$.
- Each $\left(\hat{T}^o\right)_{\ell m}(\omega)$ is a random variable.
- The $L_2(\Omega \times S^2)$ convergence holds for the expansion when the field is isotropic.
Big Data of CMB

- Collected by Planck in 2009
- Data stored at 50,331,648 HEALPix points
- \# Fourier coeff. = (4000 + 1)^2 = 16,008,001
Sparse representation by $\ell_1$ regularization


To process signals from random fields, people want to use as small proportion of coefficients to represent the main information of the field.

$$\min_{a_{\ell,m}} \frac{1}{2} \| T - T^o \|_{L_2(S^2)}^2 + \lambda \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} |a_{\ell,m}|,$$

where $T(x) := \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell,m} Y_{\ell,m}(x), \quad x \in S^2$.

where $T^o$ is an observed field.
$\ell_1$ regularization does not preserve isotropy

$$\frac{1}{2} \|T - T^0\|_{L_2(S^2)}^2 + \lambda \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} |a_{\ell,m}|.$$

- The $\ell_1$ regularization is an effective strategy of reducing the number of coefficients but is not feasible for representing CMB fields as it does not preserve isotropy of the random field.

- The essential reason the $\ell_1$ regularization violates isotropy is because, for a given $\ell \geq 1$, the sum $\sum_{m=-\ell}^{\ell} |a_{\ell,m}(\omega)|$ is not rotationally invariant.
Sparse isotropic regularization

\[
\min_{a_{\ell m}} \frac{1}{2} \| T - T^o \|_{L_2(S^2)}^2 + \lambda \sum_{\ell=0}^{\infty} \beta_\ell \left( \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2 \right)^{1/2}.
\]

- For any given \( \ell \geq 1 \) the sum \( \sum_{m=-\ell}^{\ell} |a_{\ell m}(\omega)|^2 \) is rotationally invariant.

- \( \lambda \) is a real number (regularization parameter) and \( \beta_\ell \) is a real-valued sequence.
\[
\min_{a_{\ell m}} \frac{1}{2} \|T - T^o\|^2_{L_2(S^2)} + \lambda \sum_{\ell=0}^{\infty} \beta_{\ell} \left( \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2 \right)^{1/2}
\]

\[
= \min_{a_{\ell m}} \frac{1}{2} \|a_{\ell m} - a^o_{\ell m}\|^2_{l_2} + \lambda \sum_{\ell=0}^{\infty} \beta_{\ell} \left( \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2 \right)^{1/2}.
\]
The regularized field for a given observed field $T^o$ takes the form

$$T^r(x) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a^r_{\ell,m} Y_{\ell,m}(x), \quad x \in S^2,$$

where

$$a^r_{\ell,m} := \begin{cases} 
(1 - \frac{\lambda \beta_\ell}{A^o_\ell}) a^o_{\ell,m}, & \text{if } A^o_\ell > \lambda \beta_\ell, \\
0, & \text{if } A^o_\ell \leq \lambda \beta_\ell,
\end{cases}$$

and

$$A^o_\ell := \left( \sum_{m=-\ell}^{\ell} |a^o_{\ell,m}|^2 \right)^{1/2}, \quad \ell \geq 0.$$

In any application of the present regularization scheme, the choices of the sequence $(\beta_\ell)_{\ell \geq 0}$ and the parameter $\lambda$ are crucial.
Sparse isotropic regularization preserves isotropy

- The regularized field $T^r$ is strongly isotropic if the observed field $T^o$ is isotropic.
Fig. The CMB map with $N_{\text{side}} = 2048$ as computed by SMICA.
The observed field $A_\ell^o$, empirical angular power spectrum $C_\ell = (A_\ell^o)^2/(2\ell + 1)$.
Parameter selection

Fig. The regularized field $A^r_\ell$ with $\beta_\ell = 1$, and with $\lambda = 1.05 \times 10^{-6}$ (left graph) and $\lambda = 9.75 \times 10^{-7}$ (right graph).
Errors of regularized field, Sparsity 72.1%, $L_2$ error $5.82 \times 10^{-5}$

Errors with $\beta_f = 1, \lambda = 1.05 \times 10^{-6}, \gamma = 1.0953$
Errors of regularized field, Sparsity 9.44\%, $L_2$ error $5.54 \times 10^{-5}$

Errors with $\beta_t = 1$, $\lambda = 9.75 \times 10^{-7}$, $\gamma = 1.0888$