Statistical characterization and generative modelling of cosmological fields

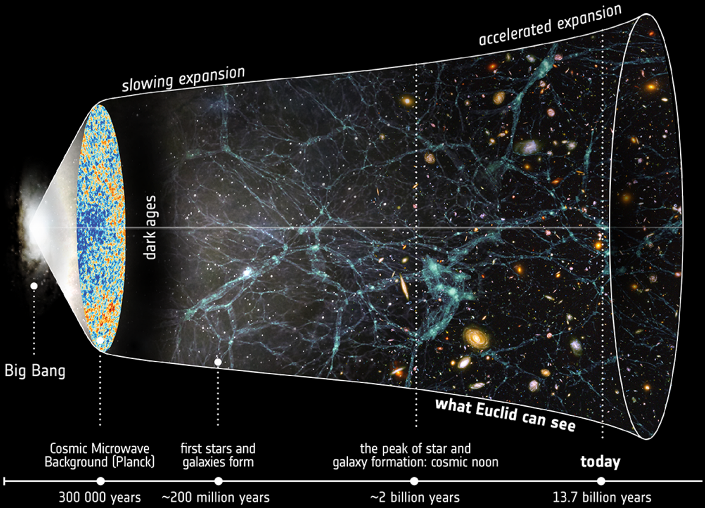
Jason McEwen

www.jasonmcewen.org

Mullard Space Science Laboratory (MSSL)
University College London (UCL)

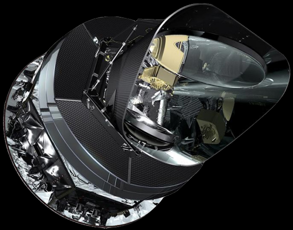
Connecting the Dots: Pattern Recognition in the Physical Sciences
July 2024

Cosmic timeline

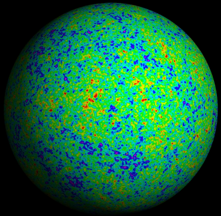


Cosmic microwave background (CMB)

What is the origin of structure in our Universe?



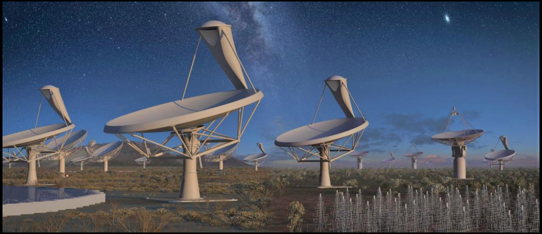
Planck satellite



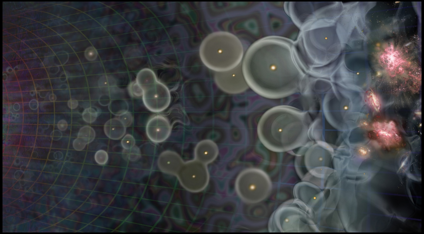
CMB

Epoch of reionisation

How did the first luminous objects in the Universe form?



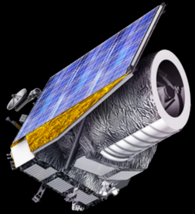
Square Kilometre Array (SKA)



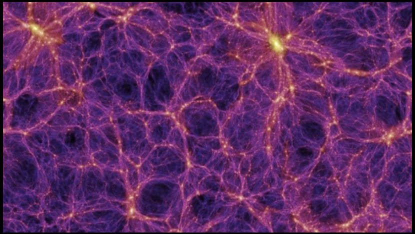
Ionised bubbles in neutral hydrogen

Large-scale structure (LSS) of the Universe

What is the nature of dark energy?

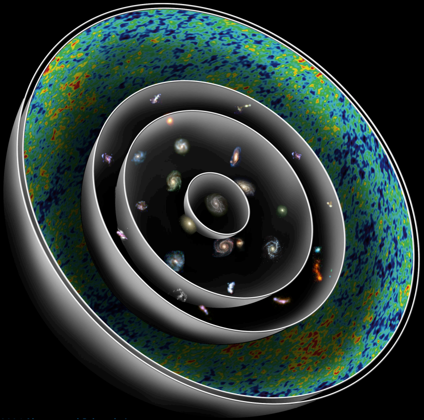


Euclid satellite



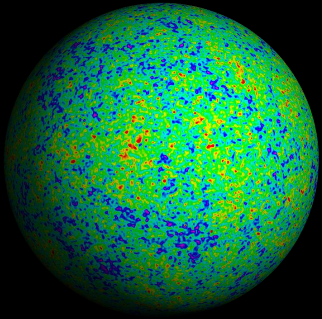
Large-scale structure

Cosmological observations on the celestial sphere

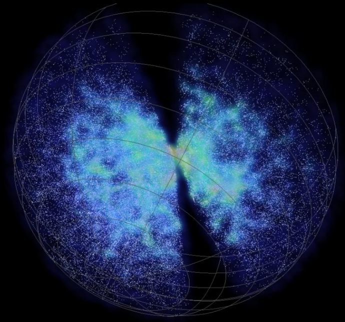


Cosmic textures on the celestial sphere

Characterization and generative modelling of cosmic textures (patterns) on the celestial sphere.

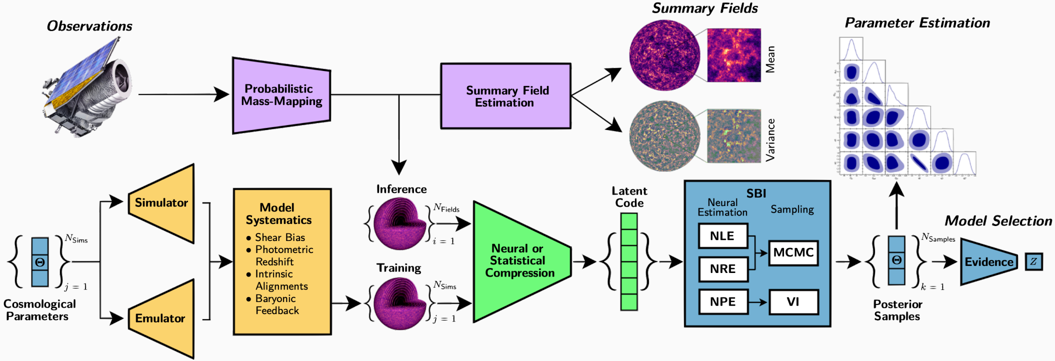


CMB



LSS

For use in simulation-based inference and beyond



SciAI UCL
Jason McEwen

Why not use standard machine learning?

Aim

Characterization and generative modelling of cosmic textures (patterns) on the celestial sphere.

Standard machine learning techniques can be applied but:

- **Requires substantial training data** (which we typically do not have in cosmology).
- **Suffers covariate shift** (*i.e.* change in cosmological model).
- **Fails to capture symmetries** of data (unless encode in model architecture).

Why not use standard machine learning?

Aim

Characterization and generative modelling of cosmic textures (patterns) on the celestial sphere.

Standard machine learning techniques can be applied but:

- **Requires substantial training data** (which we typically do not have in cosmology).
- **Suffers covariate shift** (*i.e.* change in cosmological model).
- **Fails to capture symmetries** of data (unless encode in model architecture).

$\Rightarrow$ Statistical characterization and generative modelling (inspired by CNNs).

Wavelet scattering networks and representations

Wavelet scattering networks and representations inspired by CNNs but designed rather than learned filters (Mallat 2012).

$\Rightarrow$ Scattering networks on the sphere
(McEwen et al. 2022, ICLR, arXiv:2102.02828)

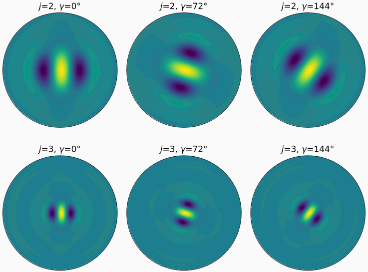
$\Rightarrow$ Generative models of astrophysical fields with scattering transforms on the sphere
(Mousset, Allys, Price, *et al.* McEwen, in prep.)

Wavelets on the sphere

Adopt scale-discretized wavelets on the sphere (e.g. McEwen et al. 2018, McEwen et al. 2015).

Wavelets $\psi_j \in \mathrm{L}^2(\mathbb{S}^2)$ capture spatially-localised, high-frequency signal content at scale j.

Scaling function $\phi \in \mathrm{L}^2(\mathbb{S}^2)$ captures spatially-localised, low-frequency content.



Orthographic plot of spherical wavelets.

Adopt scale-discretized wavelets on the sphere (e.g. McEwen et al. 2018, McEwen et al. 2015).

Wavelets $\psi_j \in \mathrm{L}^2(\mathbb{S}^2)$ capture spatially-localised, high-frequency signal content at scale *j*.

Scaling function $\phi \in \mathrm{L}^2(\mathbb{S}^2)$ captures spatially-localised, low-frequency content.

Spherical wavelet transform given by

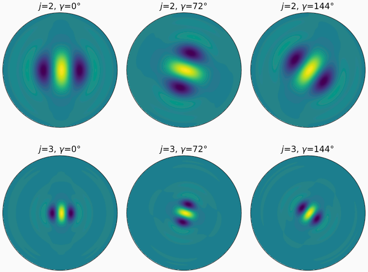
$$W_j(\rho) = (f \star \psi_j)(\rho) = \int_{\mathbb{S}^2} \mathrm{d}\mu(\omega') f(\omega') (R_\rho \psi_j)^*(\omega').$$

Spherical convolution ($f \star \psi_j$)

Rotated wavelet ($R_\rho \psi_j$)

Fast algorithms available

(e.g. McEwen et al. 2007, 2013, 2015).



Orthographic plot of spherical wavelets.

Spherical scattering propagator for scale j:

$$U[j]f = |f \star \psi_j|.$$

Modulus function is adopted for the activation function (since non-expansive and preserves stability of wavelet representation).

Scattering transform on the sphere

Spherical scattering propagator for scale j:

$$U[j]f = |f \star \psi_j|.$$

Modulus function is adopted for the activation function (since non-expansive and preserves stability of wavelet representation).

Spherical cascade of propagators:

$$U[p]f = |||f \star \psi_{j_1}| \star \psi_{j_2}| \dots \star \psi_{j_d}|,$$

for the path $p = (j_1, j_2, \dots, j_d)$ with depth d.

Scattering transform on the sphere

Spherical scattering propagator for scale j:

$$U[j]f = |f \star \psi_j|.$$

Modulus function is adopted for the activation function (since non-expansive and preserves stability of wavelet representation).

Spherical cascade of propagators:

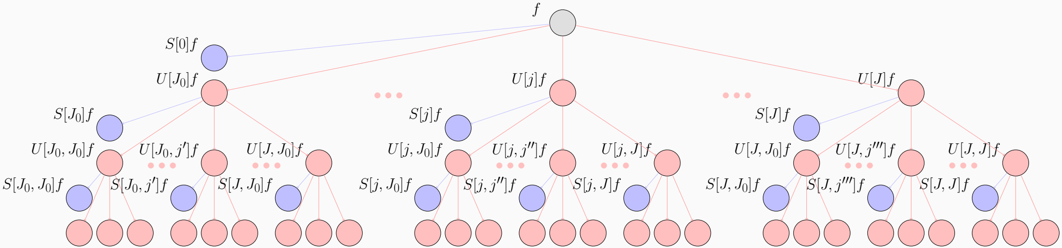
$$U[p]f = |||f \star \psi_{j_1}| \star \psi_{j_2}| \dots \star \psi_{j_d}|,$$

for the path $p = (j_1, j_2, \dots, j_d)$ with depth d.

Scattering coefficients:

$$S[p]f = |||f \star \psi_{j_1}| \star \psi_{j_2}| \dots \star \psi_{j_d}| \star \phi.$$

Spherical scattering network is collection of scattering transforms for a number of paths: $\mathcal{S}_{\mathbb{P}} f = \{S[p]f : p \in \mathbb{P}\}$, where the general path set $\mathbb{P}$ denotes the infinite set of all possible paths $\mathbb{P} = \{p = (j_1, j_2, \dots, j_d) : J_0 \leq j_i \leq J, 1 \leq i \leq d,\ d \in \mathbb{N}_0\}$.



Capture all information content at infinite depth and typically $> 99\%$ for depth $d = 3$.

Properties

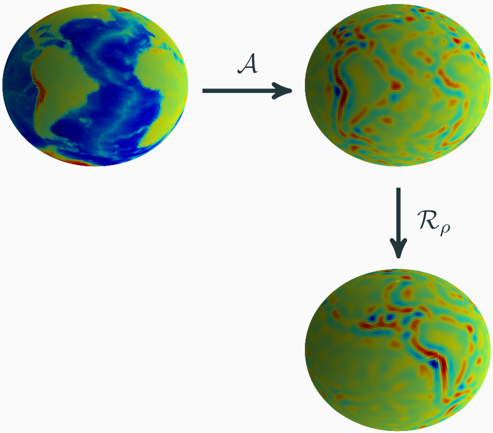
Latent representation is very well-behaved and satisfies a number of important properties:

1. **Rotational equivariance**
2. **Isometric invariance**
3. **Stability to diffeomorphisms**

Rotationally equivariance

Rotational Equivariance

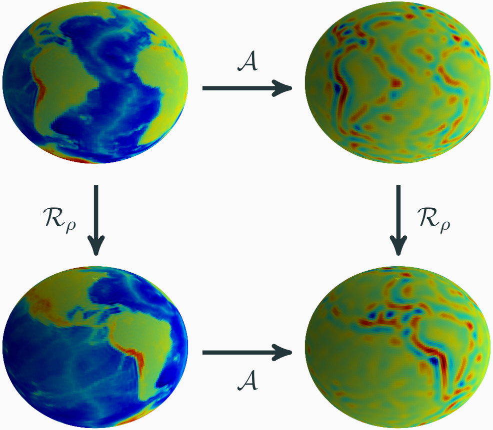
$$((\mathcal{R}_\rho f) \star \psi)(\rho') = (\mathcal{R}_\rho (f \star \psi))(\rho').$$



Rotationally equivariance

Rotational Equivariance

$$((\mathcal{R}_\rho f) \star \psi)(\rho') = (\mathcal{R}_\rho (f \star \psi))(\rho').$$



Isometric Invariance

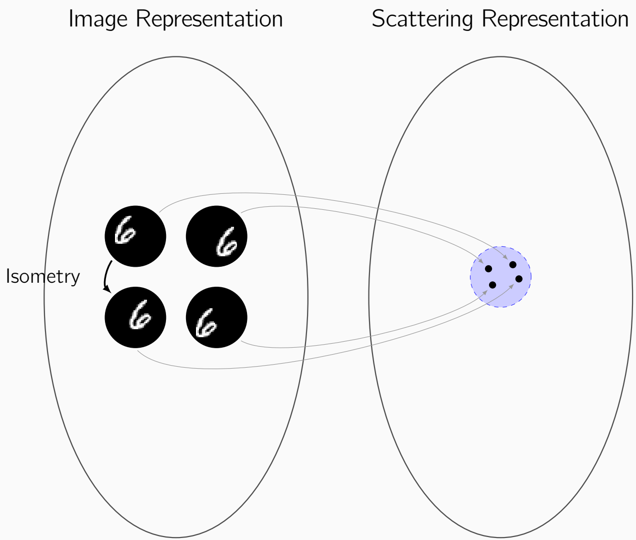
Let $\zeta \in \mathrm{Isom}(\mathbb{S}^2)$, then there exists a constant C such that for all $f \in \mathrm{L}^2(\mathbb{S}^2)$,

$$\| \mathcal{S}_{\mathbb{P}_D} f - \mathcal{S}_{\mathbb{P}_D} V_\zeta f \|_2 \leq C L^{5/2} (D+1)^{1/2} \lambda^{J_0} \|\zeta\|_\infty \|f\|_2 .$$

Difference in representation.

Scattering network representation is invariant to isometries up to a scale.

Isometric invariance



Stability to diffeomorphisms

Stability to Diffeomorphisms

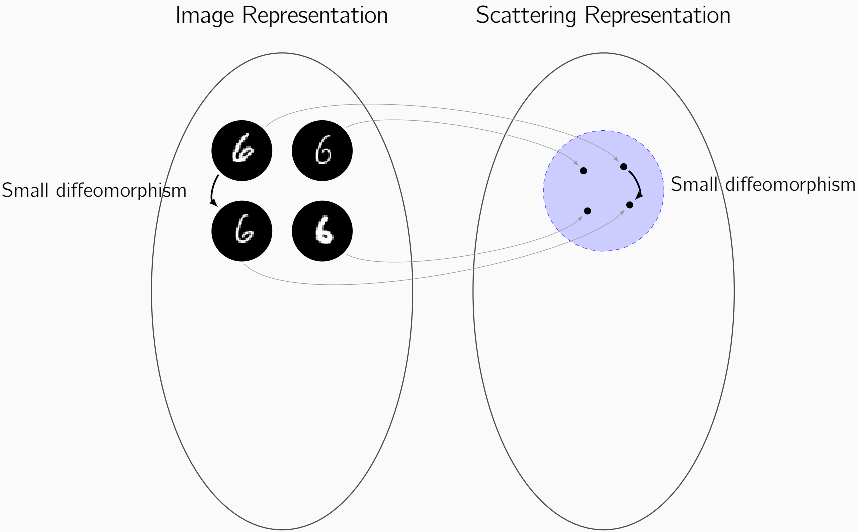
Let $\zeta \in \mathrm{Diff}(\mathbb{S}^2)$. If $\zeta = \zeta_1 \circ \zeta_2$ for some isometry $\zeta_1 \in \mathrm{Isom}(\mathbb{S}^2)$ and diffeomorphism $\zeta_2 \in \mathrm{Diff}(\mathbb{S}^2)$, then there exists a constant C such that for all $f \in \mathrm{L}^2(\mathbb{S}^2)$,

$$\|\mathcal{S}_{\mathbb{P}_D} f - \mathcal{S}_{\mathbb{P}_D} V_\zeta f\|_2 \leq C L^2 \left[L^2 \|\zeta_2\|_\infty + L^{1/2} (D+1)^{1/2} \lambda^{J_0} \|\zeta_1\|_\infty \right] \|f\|_2 .$$

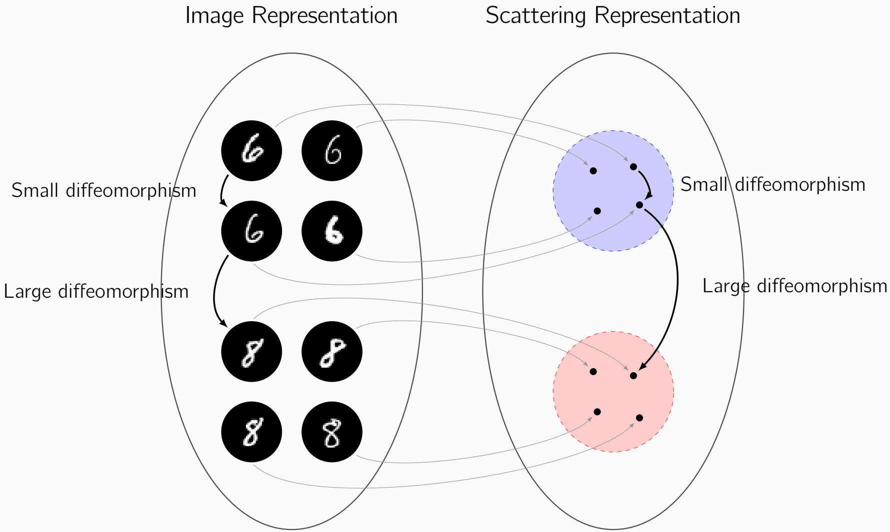
Difference in representation.

Scattering network representation is stable to small diffeomorphisms about isometry.

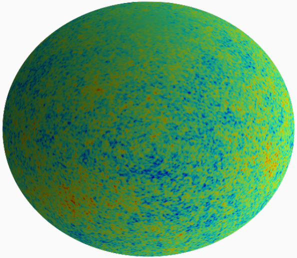
Stability to diffeomorphisms



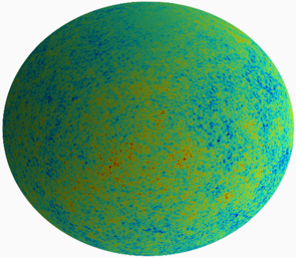
Stability to diffeomorphisms



Toy problem: Gaussianity of the cosmic microwave background (CMB)



Gaussian



Non-Gaussian

53% classification accuracy without scattering versus 95% with scattering network.

Generative models of astrophysical fields with scattering transforms on the sphere

(Mousset, Allys, Price, *et al.* McEwen, in prep.)

Scattering covariance statistics:

1. $S_1[\lambda]\, f = \mathbb{E}\big[\, |f \star \psi_\lambda| \,\big]$.
2. $S_2[\lambda]\, f = \mathbb{E}\big[\, |f \star \psi_\lambda|^2 \,\big]$.
3. $S_3[\lambda_1, \lambda_2]\, f = \mathrm{Cov}\big[\, f \star \psi_{\lambda_2}, |f \star \psi_{\lambda_1}| \star \psi_{\lambda_2} \,\big]$.
4. $S_4[\lambda_1, \lambda_2, \lambda_3]\, f = \mathrm{Cov}\big[\, |f \star \psi_{\lambda_1}| \star \psi_{\lambda_3}, |f \star \psi_{\lambda_2}| \star \psi_{\lambda_3} \,\big]$.

Spherical scattering covariance for generative modelling

Generative models of astrophysical fields with scattering transforms on the sphere

(Mousset, Allys, Price, *et al.* McEwen, in prep.)

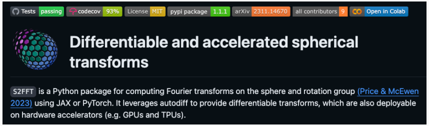
Scattering covariance statistics:

1. $S_1[\lambda]\, f = \mathbb{E}\big[\, |f \star \psi_\lambda| \,\big]$.
2. $S_2[\lambda]\, f = \mathbb{E}\big[\, |f \star \psi_\lambda|^2 \,\big]$.
3. $S_3[\lambda_1, \lambda_2]\, f = \mathrm{Cov}\big[\, f \star \psi_{\lambda_2}, |f \star \psi_{\lambda_1}| \star \psi_{\lambda_2} \,\big]$.
4. $S_4[\lambda_1, \lambda_2, \lambda_3]\, f = \mathrm{Cov}\big[\, |f \star \psi_{\lambda_1}| \star \psi_{\lambda_3}, |f \star \psi_{\lambda_2}| \star \psi_{\lambda_3} \,\big]$.

Generative modelling by **matching set of scattering covariance statistics $\mathcal{S}(f)$** with a (single) target simulation:

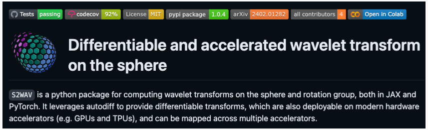
$$\min_f \|\mathcal{S}(f) - \mathcal{S}(f_{\mathrm{target}})\|^2 .$$

Differentiable and GPU-accelerated spherical transform codes (in JAX)



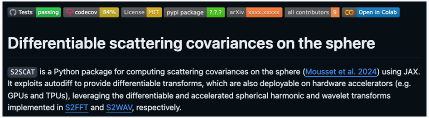
`s2fft`: Spherical harmonic transforms

https://github.com/astro-informatics/s2fft



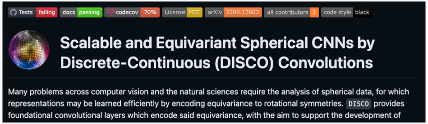
`s2wav`: Spherical wavelet transforms

https://github.com/astro-informatics/s2wav



`s2scat`: Spherical scattering transforms

https://github.com/astro-informatics/s2scat



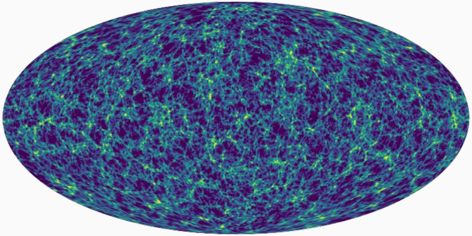
`s2ai`: Spherical AI

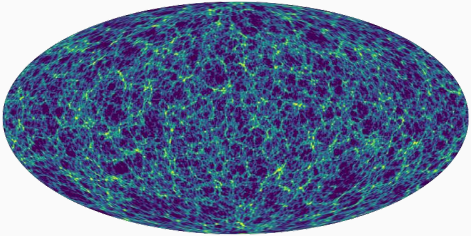
Coming very soon! Contact us for early access.

SciAI UCL

Generative modelling of large scale structure (LSS)

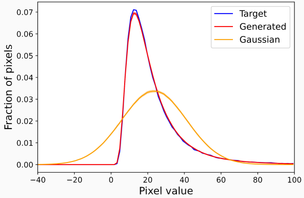
Which field is emulated and which simulated?



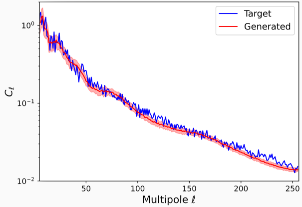


Logarithm (for visualization) of weak lensing field.

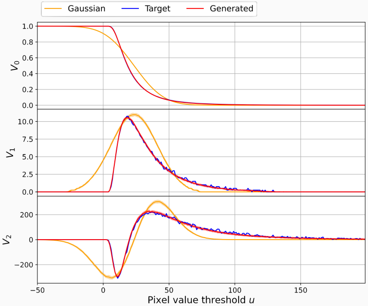
Generative modelling of large scale structure (LSS)



Pixel distribution



Power spectrum



Minkowski functionals

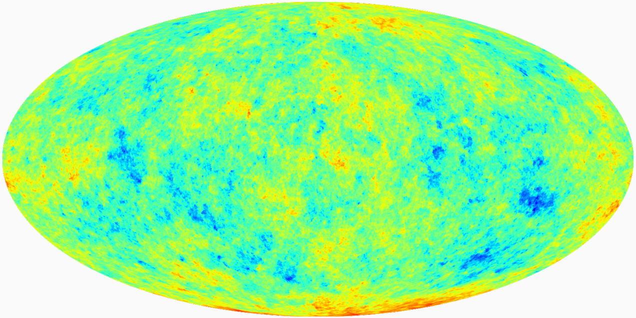
Statistical validation.

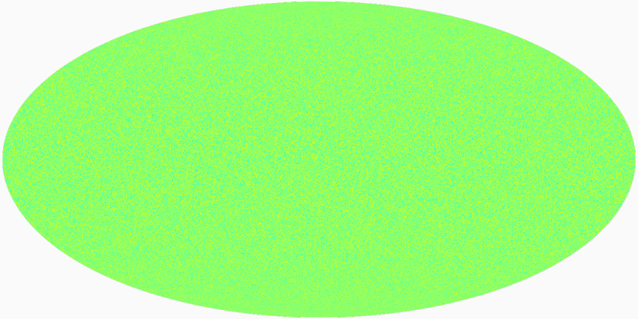
Generative modelling of cosmic strings in the CMB

Need to **simulate full physics**, evolving a network of strings through cosmic time, and then ray-trace CMB photons through the string network (Ringeval et al. 2012).

A single simulation requires **800,000 CPU hours on a supercomputer**.

There are **only three full-sky string maps in existence**.

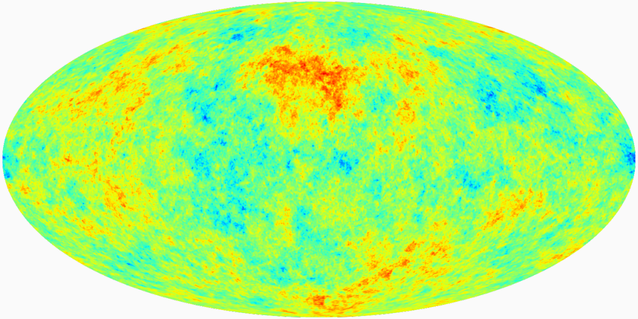




Computation time: 800,000 CPU hours on supercomputer $\rightarrow \mathcal{O}(1)$ hours on A100 GPU.

Still work in progress (statistical validation in progress).

Generative modelling of cosmic strings in the CMB



Computation time: 800,000 CPU hours on supercomputer $\rightarrow$ $\mathcal{O}(1)$ hours on A100 GPU.

Still work in progress (statistical validation in progress).

Summary

Characterization and generative modelling of cosmic textures (patterns) on the celestial sphere with wavelet scattering representations.

Advantages:

- Little to no training data.
- No covariate shift.
- Capture spherical symmetries.

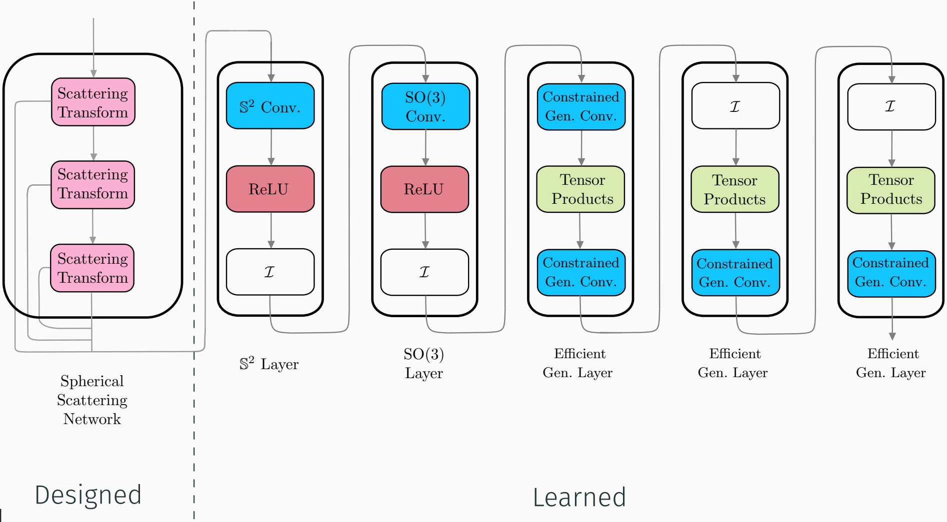
Well-behaved latent representation:

1. Rotational equivariance.
2. Isometric invariance.
3. Stability to diffeomorphisms.

Excellent latent representation to characterize cosmological fields or for generative modelling (**saving of** 10^6 **in computational time, rendering new analyses feasible**).

Extra slides

Scalable and rotationally equivariant spherical CNNs



SciAI UCL
Jason McEwen