Towards wide-field, field-level simulation-based inference (SBI) for Euclid cosmic shear

Jason McEwen
www.jasonmcewen.org

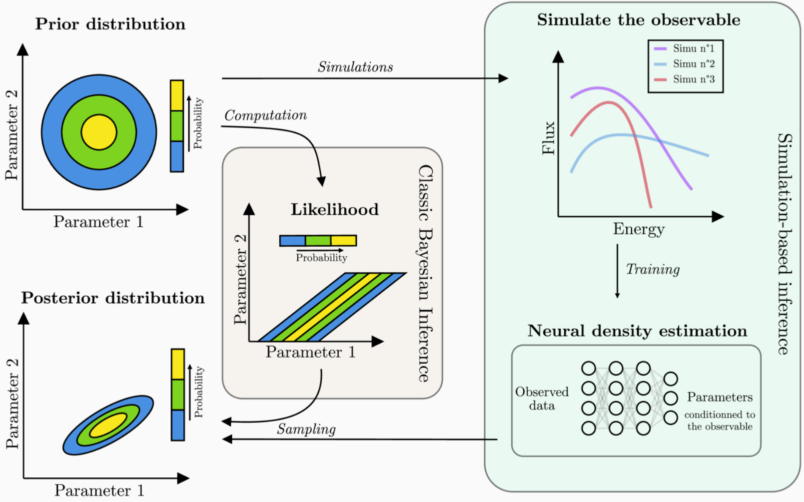
Mullard Space Science Laboratory (MSSL)
University College London (UCL)

National Astronomy Meeting (NAM), July 2024

Outline

SBI: what and why

Classical likelihood-based inference versus SBI



Simulation-based inference (aka. likelihood-free inference) seeks to perform Bayesian inference by **estimating the posterior** $p(\theta \,|\, x_o, M)$ of **parameters** θ for **observed data** x_o using **simulations only**.

Simulation-based inference (aka. likelihood-free inference) seeks to perform Bayesian inference by **estimating the posterior** $p(\theta \mid x_o, M)$ of **parameters** θ for **observed data** x_o using **simulations only**.

Key advantages:

- Forward modelling of complex physics, systematics, observational process.
- No assumptions on the form of the likelihood.

SBI variants

Three variants:

1. **Neural posterior estimation (NPE)**: learn surrogate of posterior (probability distribution over parameters).

2. **Neural likelihood estimation (NLE)**: learn surrogate of likelihood (probability distribution over data).

3. **Neural ratio estimation (NRE)**: learn surrogate of likelihood-to-evidence ratio.

SBI variants

Three variants:

1. **Neural posterior estimation (NPE)**: learn surrogate of posterior (probability distribution over parameters).

2. **Neural likelihood estimation (NLE)**: learn surrogate of likelihood (probability distribution over data).
   - NLE introduced by Papamakarios *et al.* (2019).
   - First applied to cosmology by Alsing *et al.* (2019).
   - First applied to cosmic shear by Taylor *et al.* McEwen (2019).

3. **Neural ratio estimation (NRE)**: learn surrogate of likelihood-to-evidence ratio.

Towards Euclid SBI shear pipeline

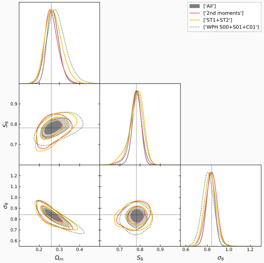
Advantages of a field-level SBI pipeline for Euclid cosmic shear

- Extract informative field-level cosmological information.
- No assumptions regarding likelihood (no need to characterize covariances).
- Capture all uncertainties.
- Accurately model systematic effects at the field-level.

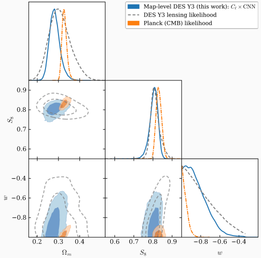
$\Rightarrow$ **More precise** (tighter constraints) and **more accurate** (in right place) Bayesian inference.

Effectiveness of field-level SBI for cosmic shear

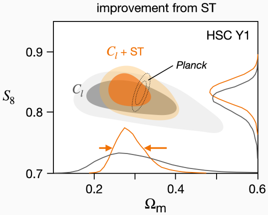
Effectiveness of field-level SBI demonstrated already in **small-field planar setting**.



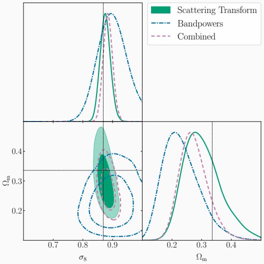
Gatti *et al.* (2023)



Jeffrey *et al.* (2024)



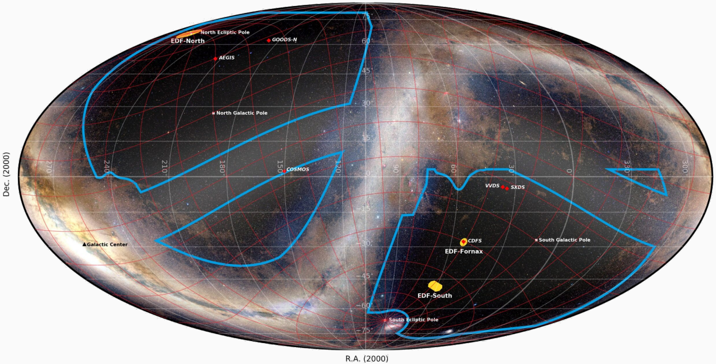
Cheng *et al.* (2024)



Lin, Joachimi, McEwen in prep. (preliminary)

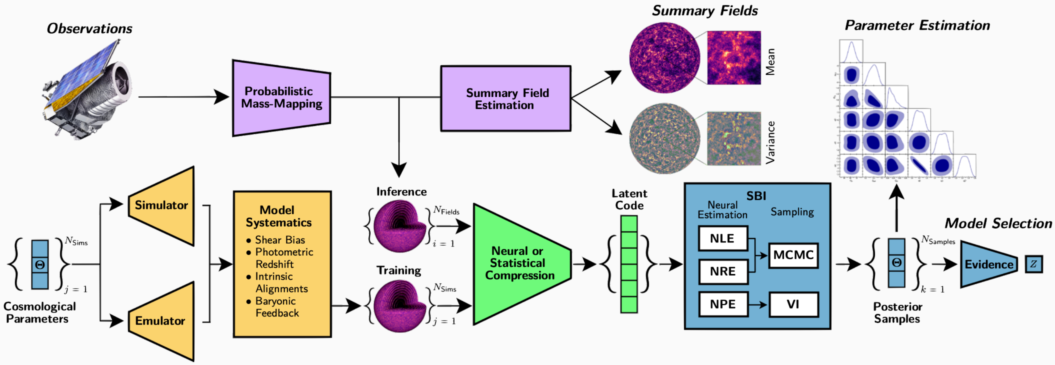
$\Rightarrow$ **Tightest cosmic shear constraints to date** from SBI.

Euclid wide-field survey



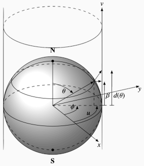
Field-level SBI techniques must be extended to support wide-fields, requiring **spherical methods defined on the curved sky.**

Wide-field, field-level SBI pipeline

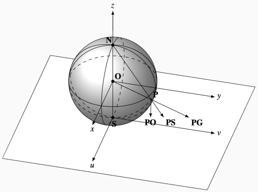


Wide-field mass-mapping

Spherical Kaiser Squires mass-mapping introduced by Wallis *et al.* McEwen (2017) to avoid planar approximations.

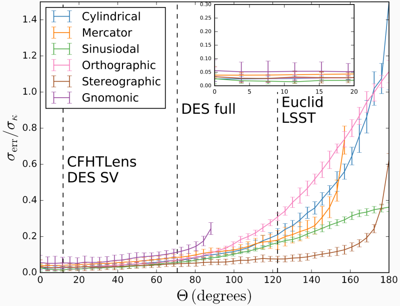


Equitorial



Polar

Projections of sphere to plane.



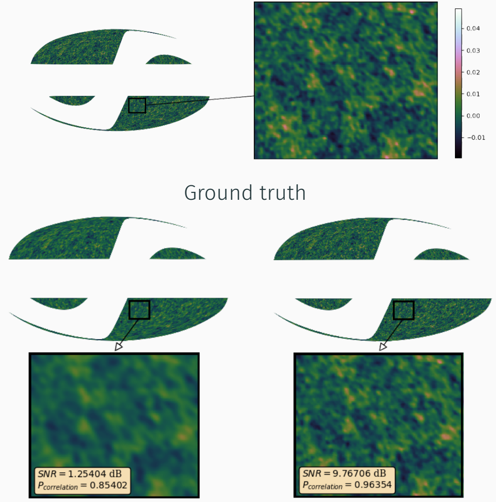
Planar projections introduce **significant error** in mass-mapping.

SciAI UCL
Jason McEwen

Enhanced spherical mass-mapping

Spherical wavelet mass-mapping introduced by Price, McEwen *et al.* 2021.

Spherical AI mass-mapping techniques in preparation...



Ground truth

Spherical Kaiser-Squires

Spherical wavelet regularization

SciAI UCL
Jason McEwen

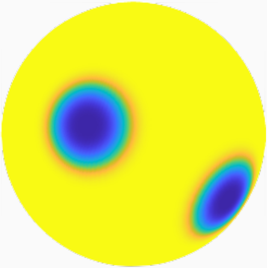
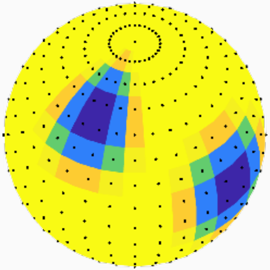
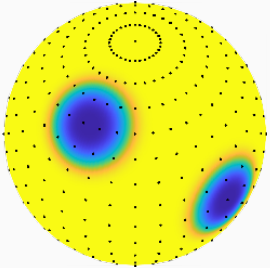
Wide-field compression

Wide-field compression

1. Neural compression
   - CNNs: Convolutional neural networks (*e.g.* Jeffrey *et al.* 2024)
2. Statistical compression
   - Scattering transforms (*e.g.* Cheng *et al.* 2024, Gatti *et al.* 2023)

Require **spherical CNNs** and **spherical scattering transforms** defined on the curved sky.

Categorization of spherical CNN frameworks

Continuous	Discrete	Discrete-Continuous (DISCO)
		
✔ Equivariant ✖ Not Scalable	✖ Not Equivariant ✔ Scalable	✔ Equivariant ✔ Scalable
(Cohen et al. 2018, Esteves et al. 2018, Kondor et al. 2018, Cobb et al. 2021, McEwen et al. 2022, ...)	(Jiang et al. 2019, Zhang et al. 2019, Perraudin et al. 2019, Cohen et al. 2019, ...)	(Ocampo, Price & McEwen 2023)

Scalable and equivariant spherical CNNs

Efficient Generalized Spherical CNNs
(Cobb *et al.* McEwen 2021)

Scalable and Equivariant Spherical CNNs by Discrete-Continuous (DISCO) Convolutions
(Ocampo, Price & McEwen 2023)

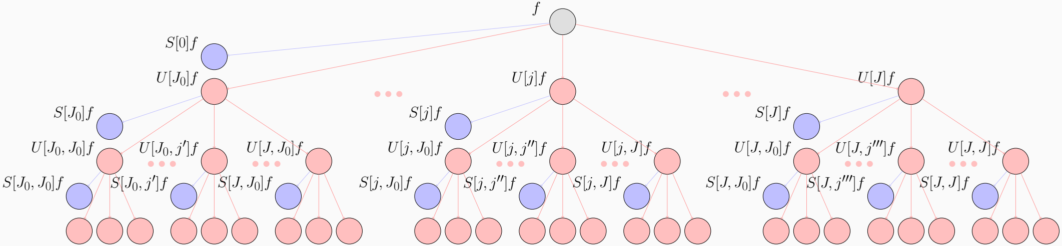
Equivariance $\Rightarrow$ state-of-the-art performance on all problems considered to date.

Scattering networks inspired by CNNs but designed rather than learned filters (Mallat 2012).

Scattering networks on the sphere

(McEwen et al. 2022)

Cascade of **spherical wavelet transforms** (McEwen *et al.* 2018) and non-linearities (modulus).



Spherical scattering covariance (third generation)

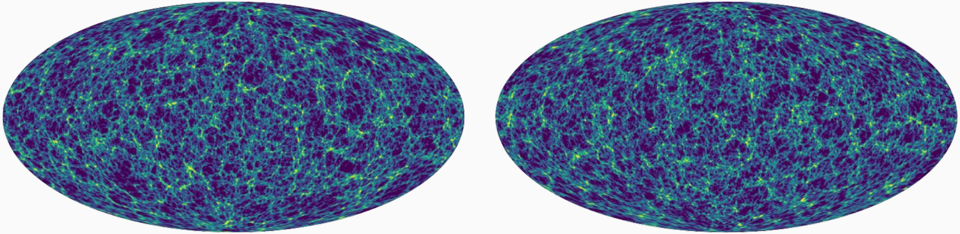
Generative models of astrophysical fields with scattering transforms on the sphere
(Mousset, Allys, Price, *et al.* McEwen 2024)

Scattering covariance statistics considered:

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. $S_4[\lambda_1, \lambda_2, \lambda_3]\, f = \mathrm{Cov}\big[\, |f \star \psi_{\lambda_1}| \star \psi_{\lambda_3}, |f \star \psi_{\lambda_2}| \star \psi_{\lambda_3} \,\big]$.

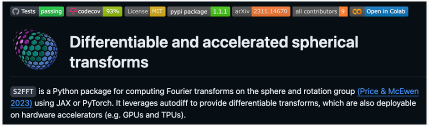
Emulation: Generative modelling with scattering covariances

Which field is emulated and which simulated?



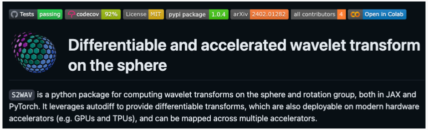
Logarithm (for visualization) of weak lensing field

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



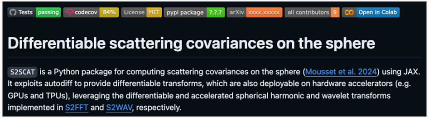
s2fft: Spherical harmonic transforms

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



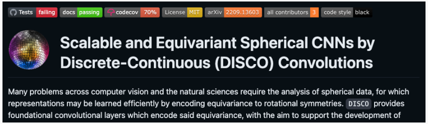
s2wav: Spherical wavelet transforms

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.

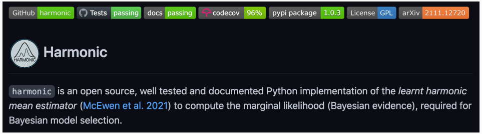
Bayesian model selection

Learned harmonic mean estimation of the Bayesian evidence

Learned harmonic mean estimator (McEwen *et al.* 2021)

$$z^{-1} = \rho = \mathbb{E}_{p(\theta \,|\, x)} \left[\frac{\varphi(\theta)}{\mathcal{L}(\theta)\pi(\theta)} \right] \quad \text{where} \quad \varphi(\theta) \overset{\text{ML}}{\simeq} \varphi^{\text{optimal}}(\theta) = \frac{\mathcal{L}(\theta)\pi(\theta)}{z} .$$

- Requires **posterior samples only**
  - ⇝ Evidence almost for free
- **Agnostic to sampling** technique
  - ⇝ Leverage efficient samplers
  - ⇝ Simulation-based inference (SBI)
  - ⇝ Variational inference
- Scale to **high-dimensions**
  - ⇝ Normalizing flows (Polanska *et al.* McEwen 2024)



`harmonic`: Learned harmonic mean

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

Accelerated Bayesian inference

The future of cosmological (likelihood-based) inference

(Piras, Polanska, Spurio Mancini, Price, McEwen 2024)

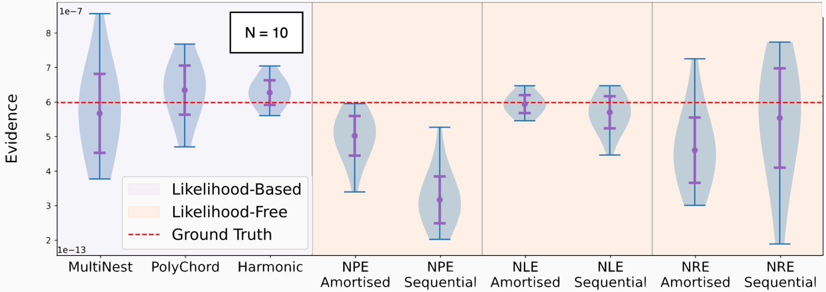
37 parameter cosmic shear analysis of LCDM vs $w_0 w_a$CDM

- CAMB + PolyChord
  $\rightsquigarrow$ 8 months on 48 CPU cores
- CosmoPower-JAX + NumPyro/NUTS + Harmonic
  $\rightsquigarrow$ 2 days on 12 GPUs

157 parameter 3x2pt analysis of LCDM vs $w_0 w_a$CDM

- CAMB + PolyChord
  $\rightsquigarrow$ 12 years on 48 CPUs (projected)
- CosmoPower-JAX + NumPyro/NUTS + Harmonic
  $\rightsquigarrow$ 8 days on 24 GPUs

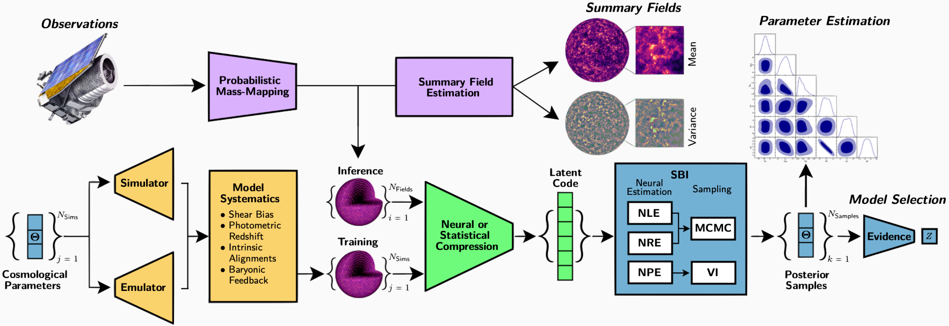
Bayesian model selection for SBI first introduced by Spurio Mancini *et al.* McEwen (2023).



Summary

Summary

- Field-level SBI **highly effective**.
- For Euclid, require spherical methods defined on the curved sky.



Have the methods and codes needed to **develop a highly effective wide-field, field-level SBI pipeline for Euclid cosmic shear.**

Extra slides

Neural likelihood estimation

Construct **training data** $\{(\theta_i, x_i)\}$ where parameter drawn from proposal prior $\theta_i \sim \tilde{p}(\theta \,|\, M)$ and then generate simulation $x_i \sim p(x \,|\, \theta_i) \Rightarrow$ joint distribution $\tilde{p}(\theta, x) = p(x \,|\, \theta, M)\tilde{p}(\theta, M)$.

Learn likelihood

$$\boxed{q_\psi(x \,|\, \theta, M) \simeq p(x \,|\, \theta, M)} \,,$$

where ψ are the parameters of the learned model.

Train by maximum likelihood, *i.e.* by maximising

$$\boxed{\mathbb{E}_{\tilde{p}(\theta,x)}[\log q_\psi(x \,|\, \theta, M)] = -\mathbb{E}_{\tilde{p}(\theta)}[D_{\text{KL}}(p(x \,|\, \theta, M), q_\psi(x \,|\, \theta, M))] + \text{const.}} \,,$$

where D_{KL} is the Kullback-Leibler divergence.

Sample from approximate posterior by **MCMC sampling**.

Conditional normalizing flow as density estimator