ETH zürich

Universität Basel

Cosmology Group
Optimization of Machine Learning Systems Group

Constraining Cosmology with Deep Learning

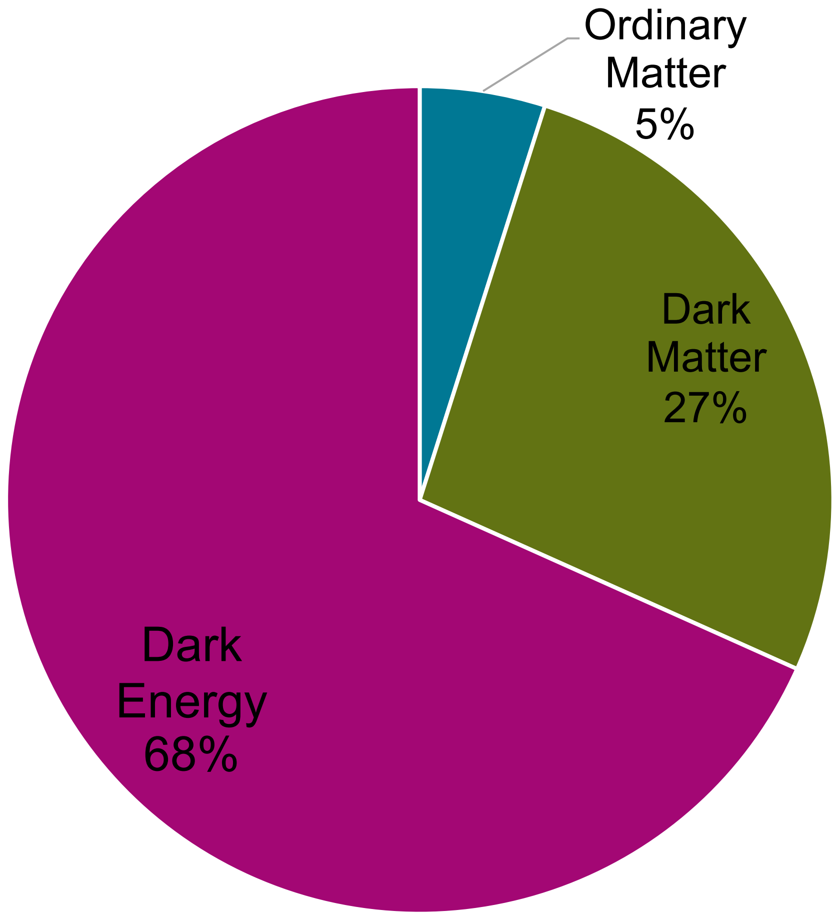
Arne Thomsen

21.03.2023, IPA-ML Workshop

Lambda-CDM

- Refinement of **big bang model**
- Gravity described by **general relativity**
- Major components
  - Ordinary matter
  - Cold dark matter (CDM)
  - Dark energy (Lambda)
- 5 free parameters
  - Mass fraction Ω_M
  - Variance of mass distribution σ_8
  - And $H_0,\ \Omega_b,\ n_s$

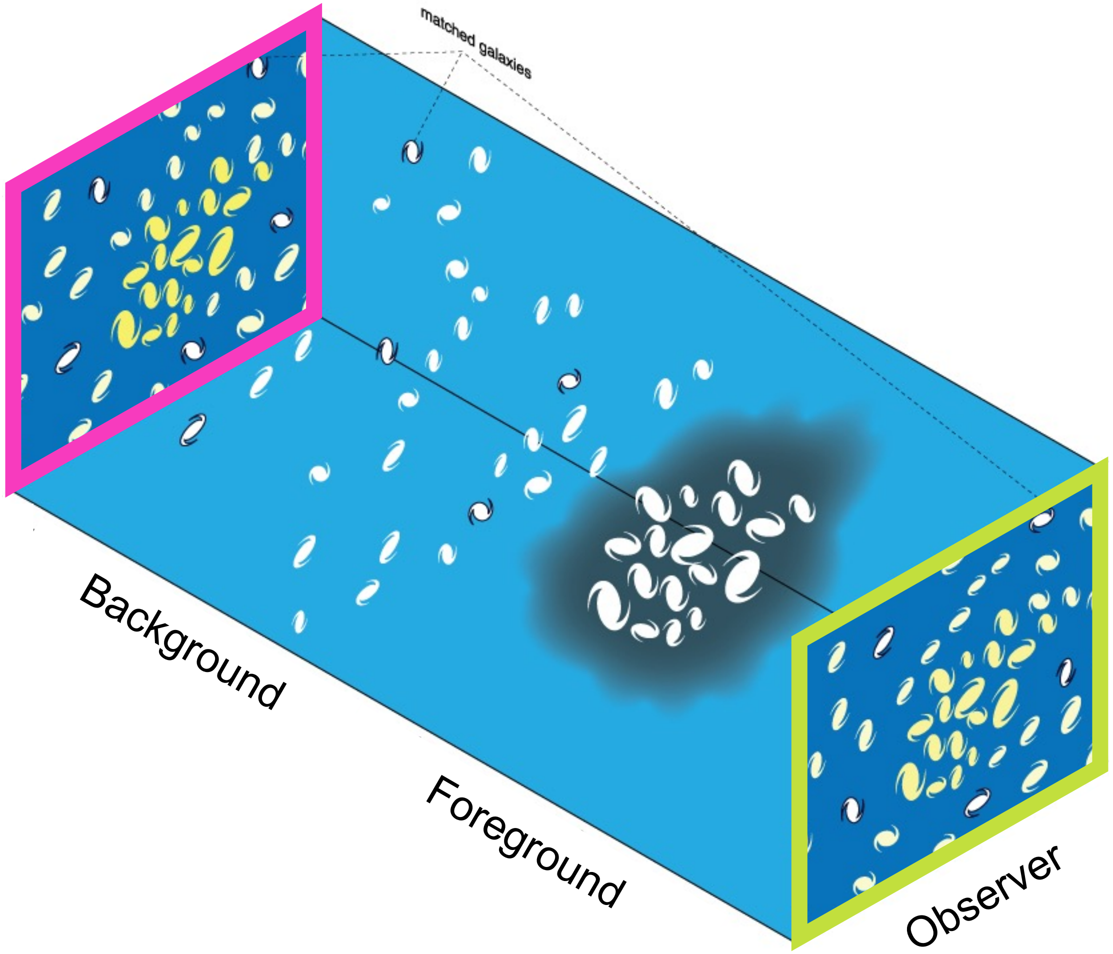
→ Improve constraints



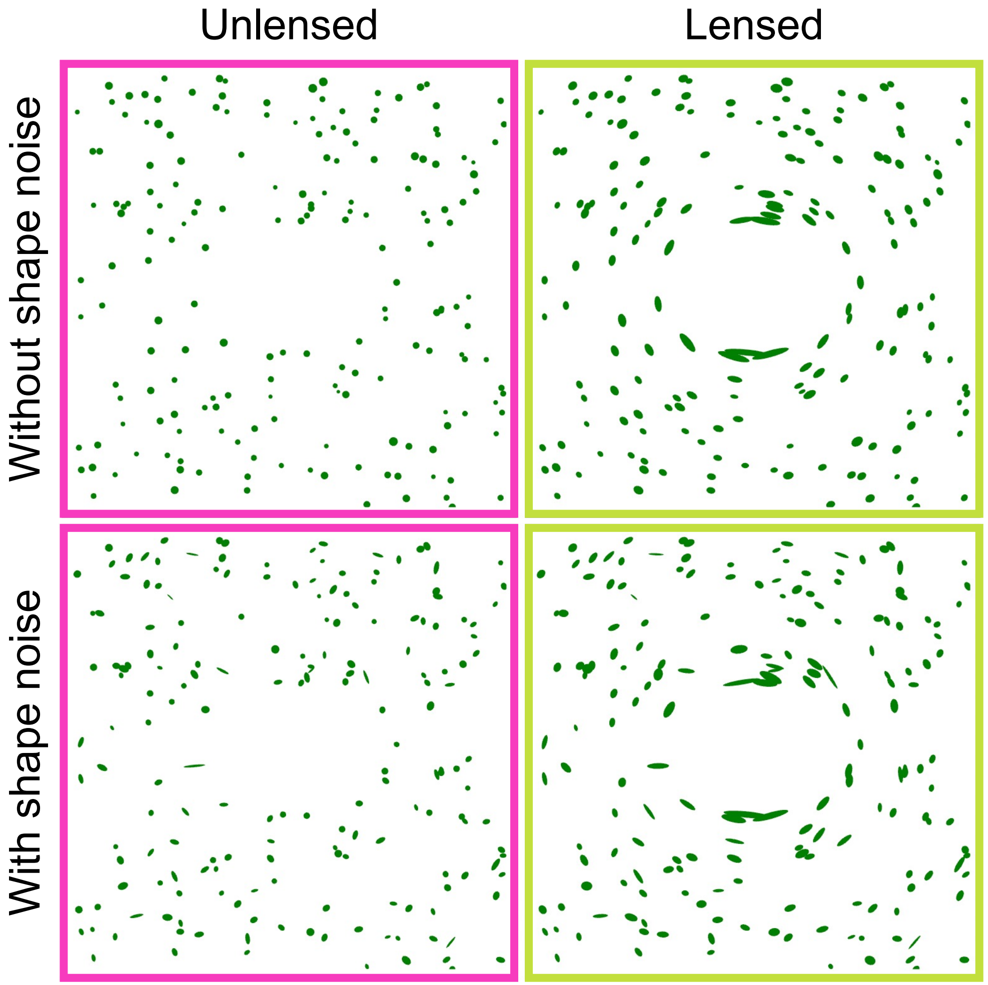
Energy content of the Universe today

Weak (gravitational) lensing

Image credit: [Wikipedia]

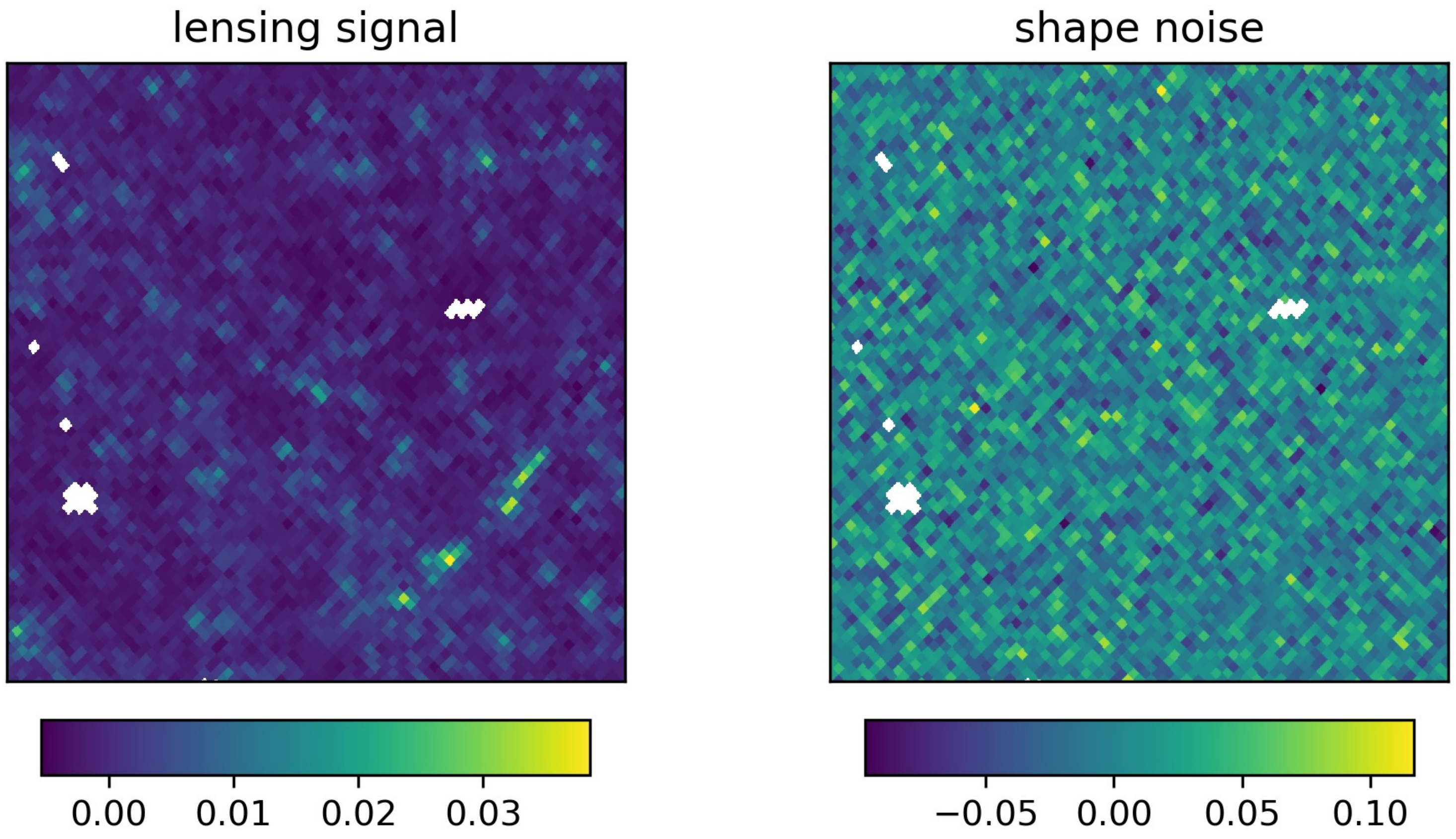


Matter arrangement

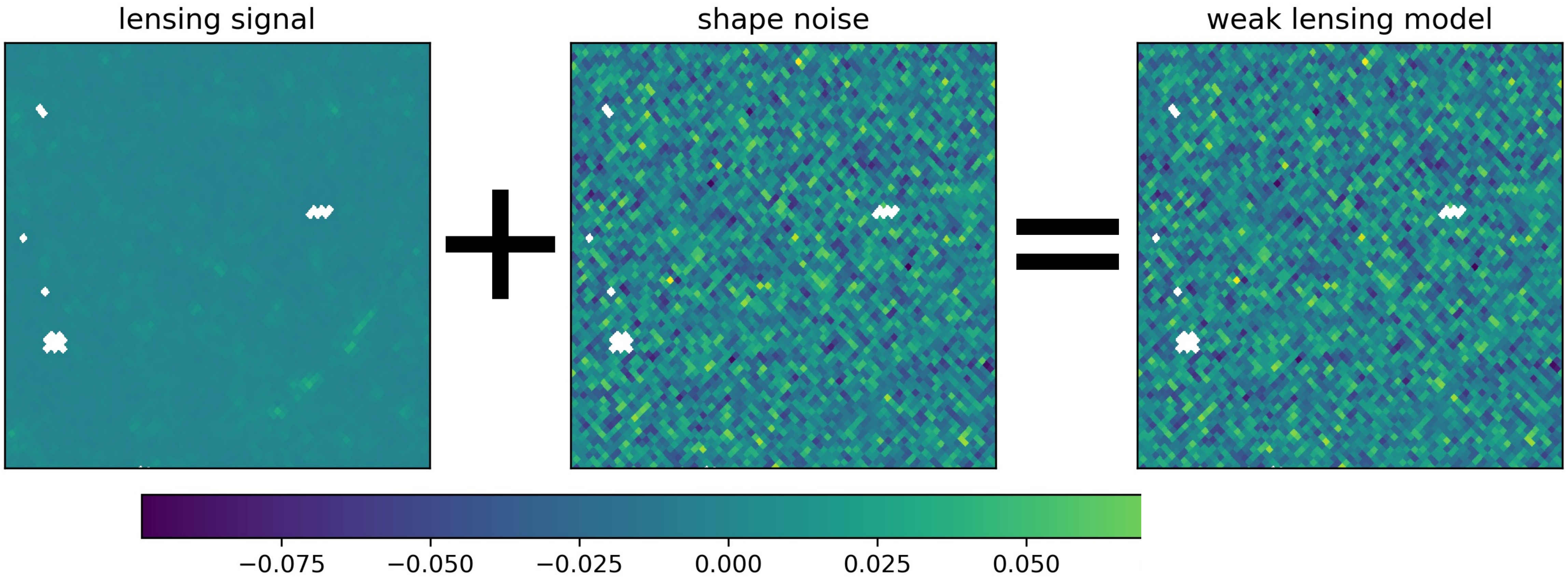


Galaxy shapes

Weak (gravitational) lensing



Weak (gravitational) lensing



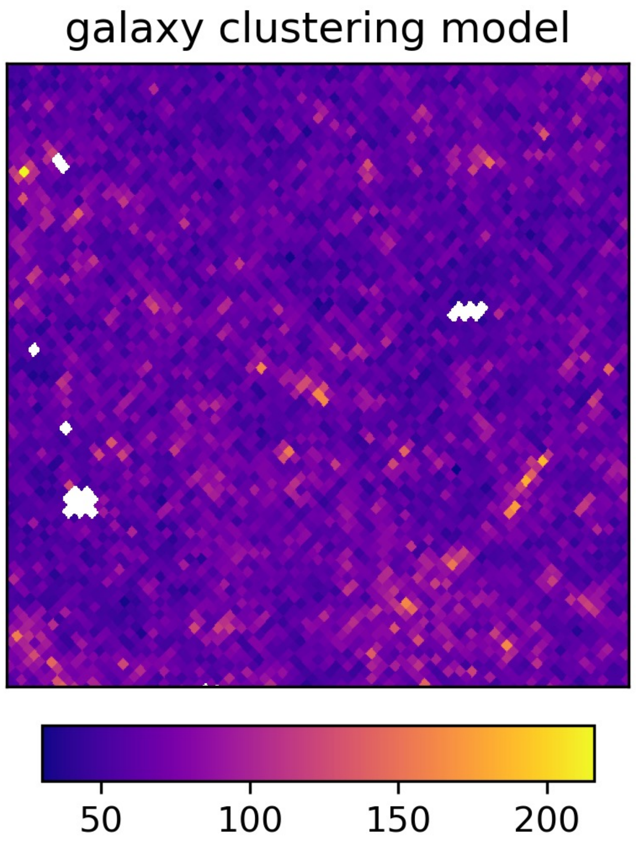
Combined probes of large scale structure

General motivation

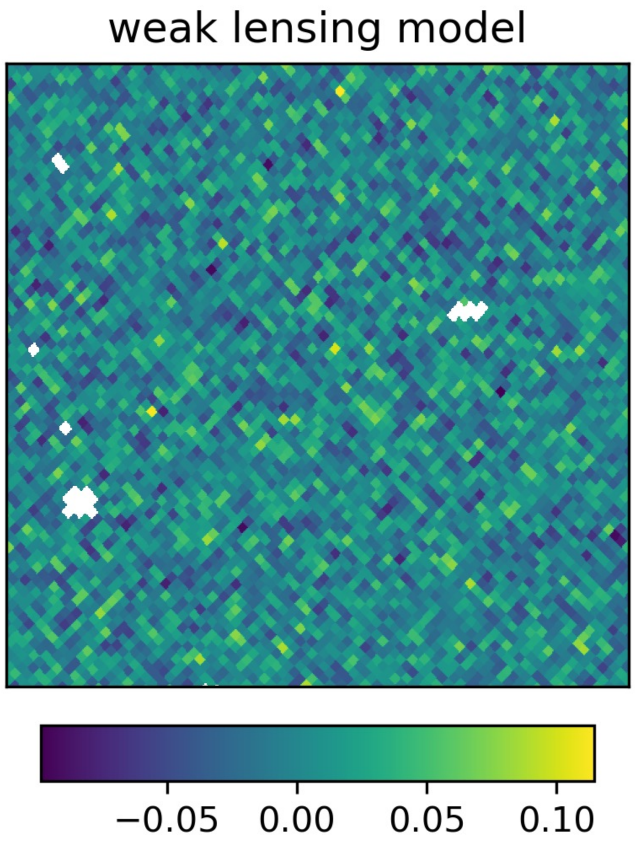
- Breaking of parameter degeneracies
  - Cosmology
  - Astrophysics
  - Measurement

Deep learning specific

- Combining probes in different channels is easy



Number of galaxies per pixel



Convergence value per pixel

Dark Energy Survey (DES)

- International colloboration
  - >400 members
  - 25 institutions
  - USA, Spain, UK, Brazil, Germany, Switzerland, Australia
- 4m telescope in the Chilean Andes
- 570 Megapixel camera with 3 deg field of view
- 758 nights of observations (2013-2019)
- Area of 5000 deg^2 (~1/10 of the sky)
- Half the history of the universe deep

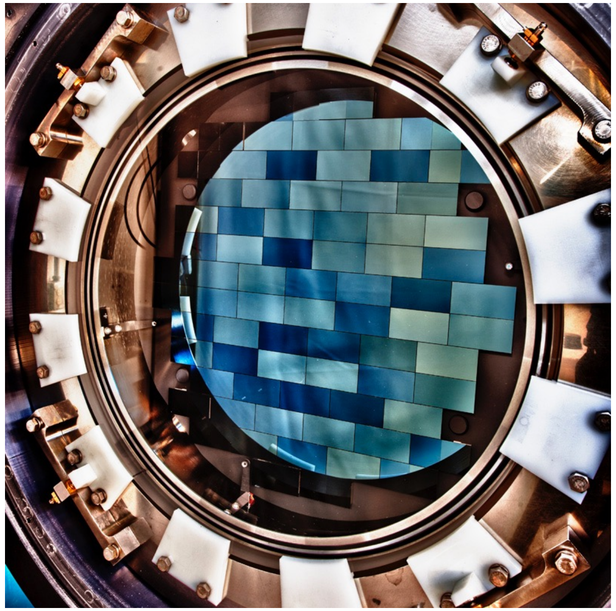
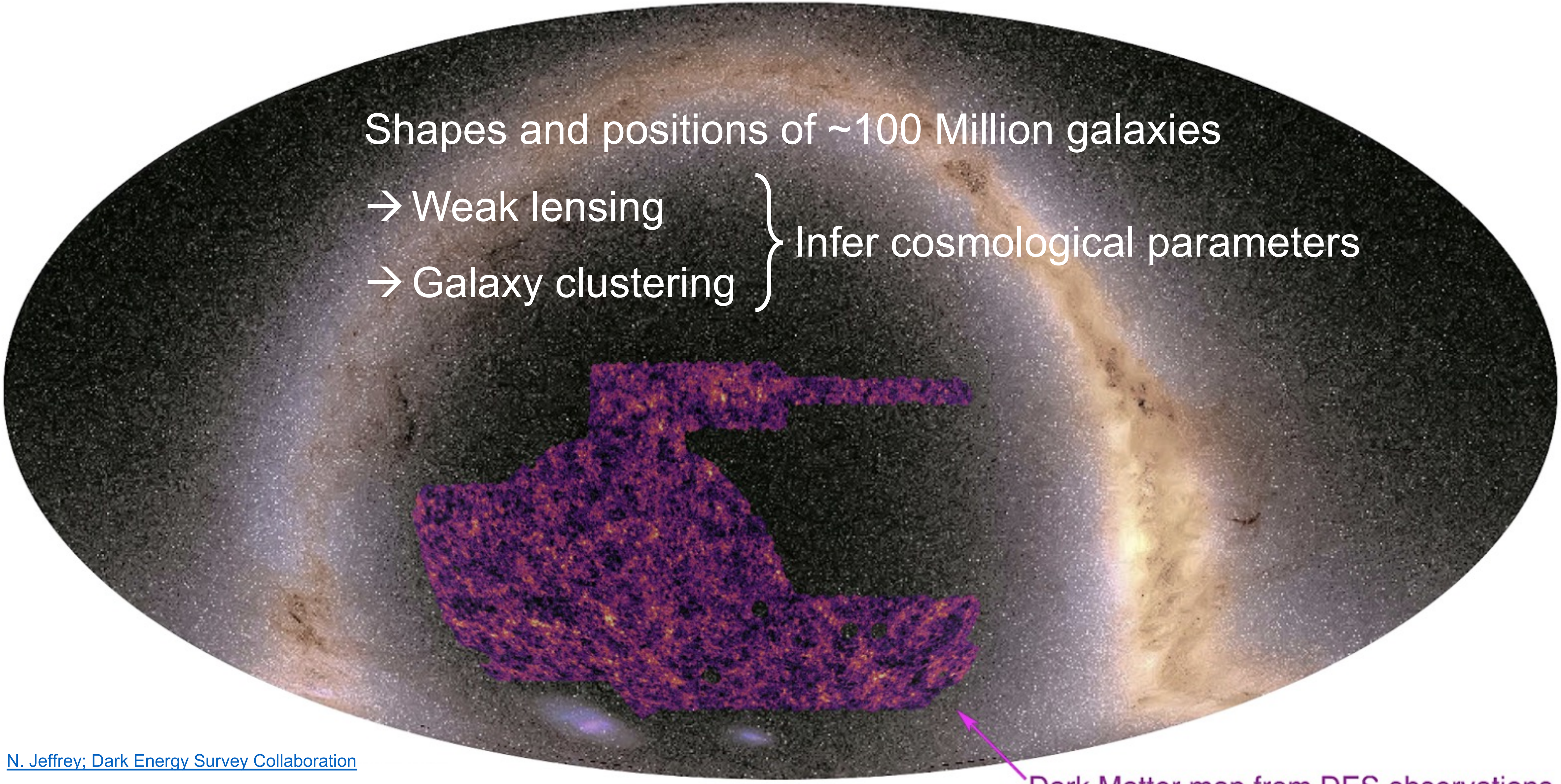


Image credit: Reider Hahn, Fermilab

DES: Survey footprint

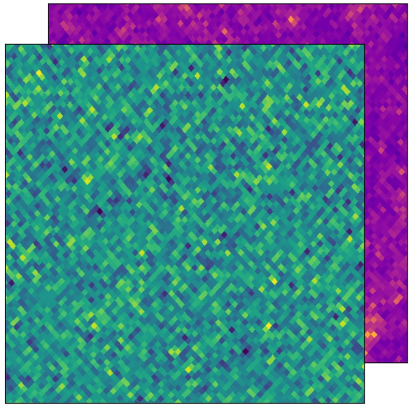


Dark Matter map from DES observations

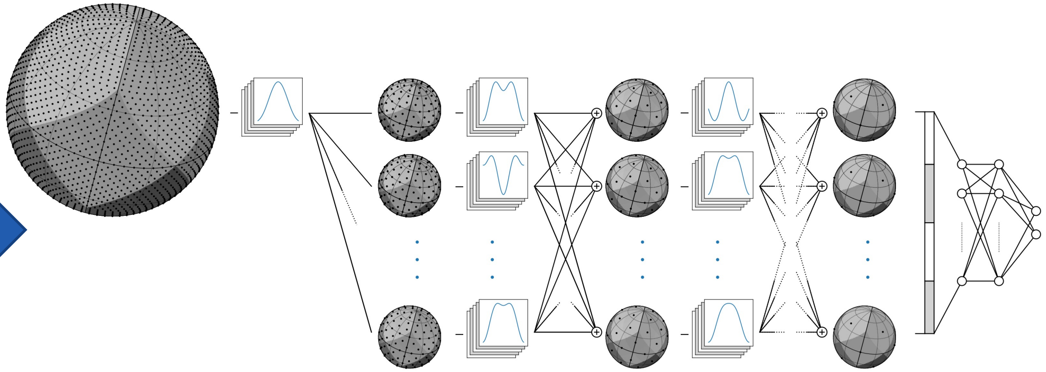
N. Jeffrey; Dark Energy Survey Collaboration

Inference with deep learning

Observation

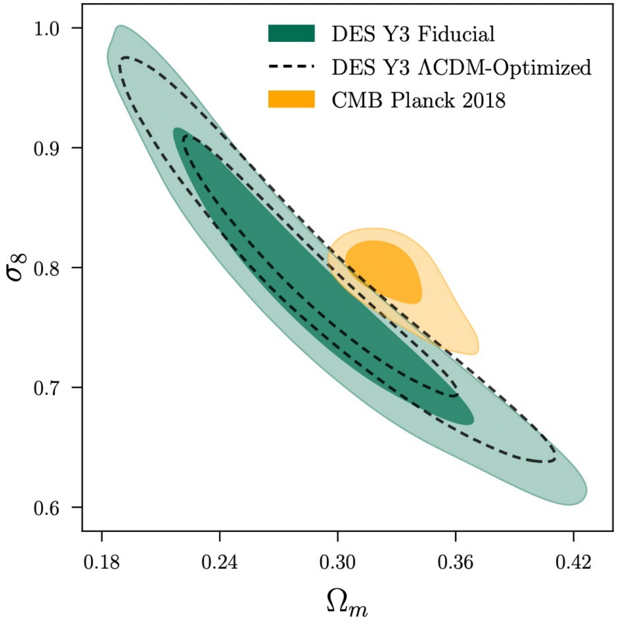


Learned summary statistic



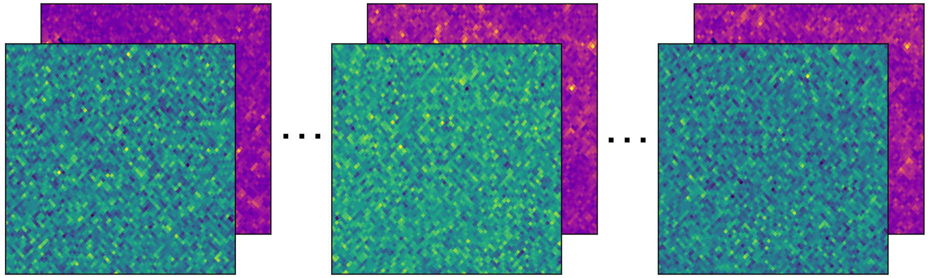
N. Perraudin et al. 2019

Simulation based inference



L. F. Secco et al. 2022

Forward model



Computing resources

- Perlmutter GPU cluster
- National Energy Research Scientific Computing Center (NERSC)
- 8th fastest (TOP500)

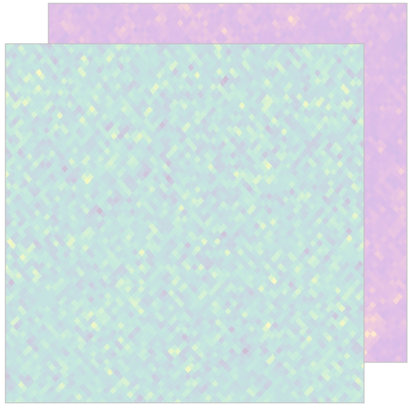
➡ Massive compute resources

➡ Time to solution like standard approach

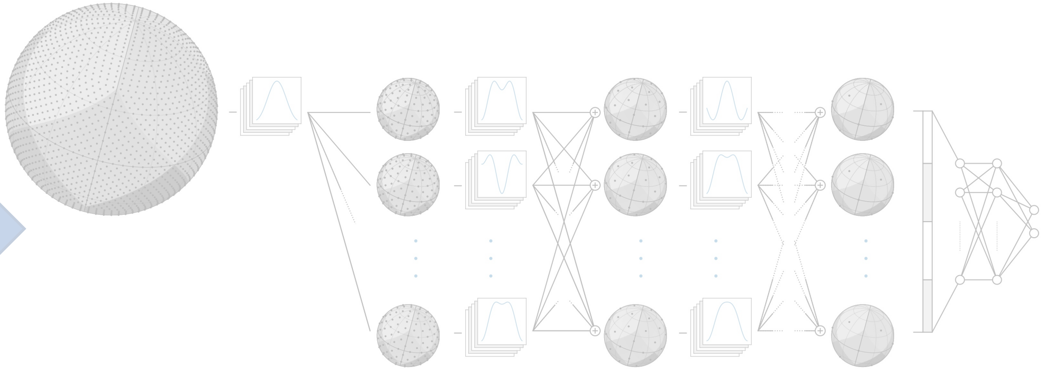
Image credit: NERSC

Inference with deep learning

Observation

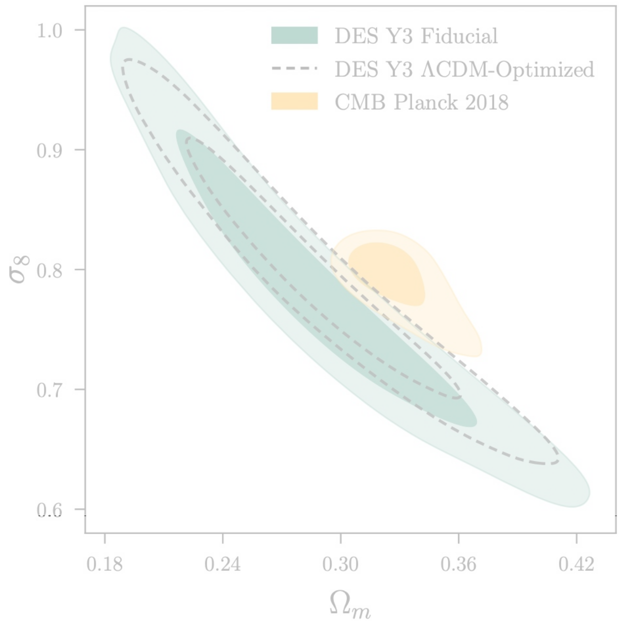


Learned summary statistic



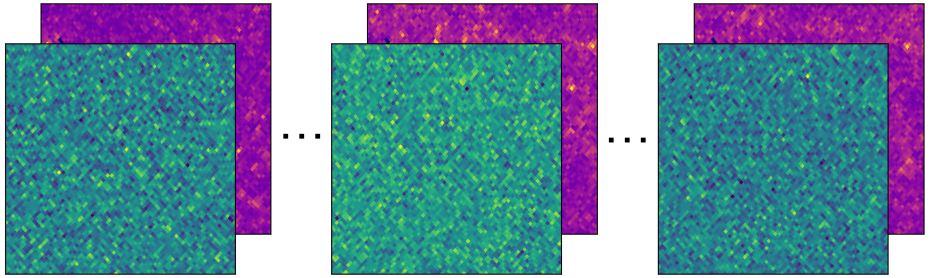
N. Perraudin et al. 2019

Simulation based inference



L. F. Secco et al. 2022

Forward model

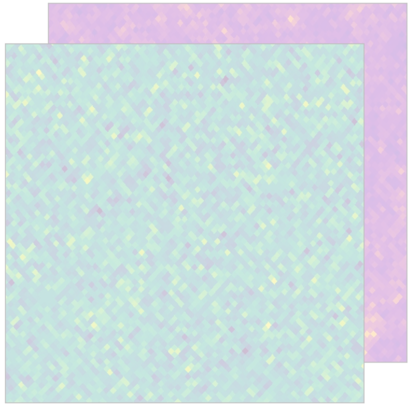


Forward model

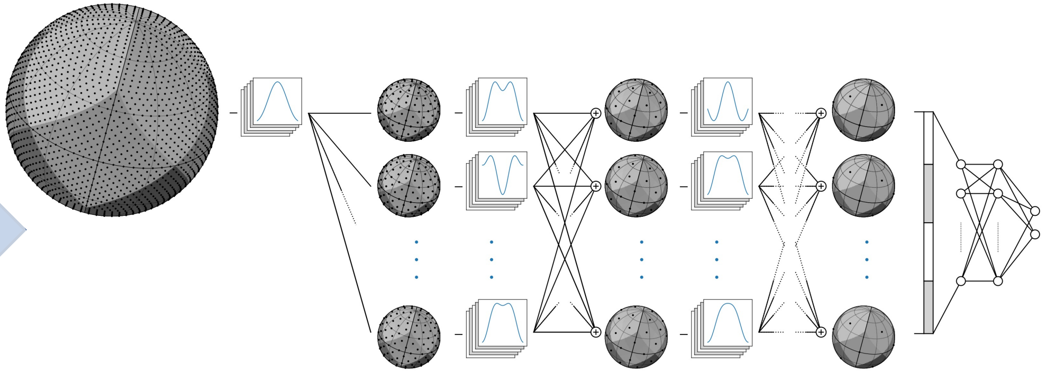
General aspects	Deep learning specific
• Data driven approach → Physics is in the modeling	• Fast pipeline required → Trade-off between accuracy and speed
• No analytical prediction summary statistics needed → Flexible choice	• The networks can learn anything → Extra accurate forward model
	• New map level systematics tests

Inference with deep learning

Observation

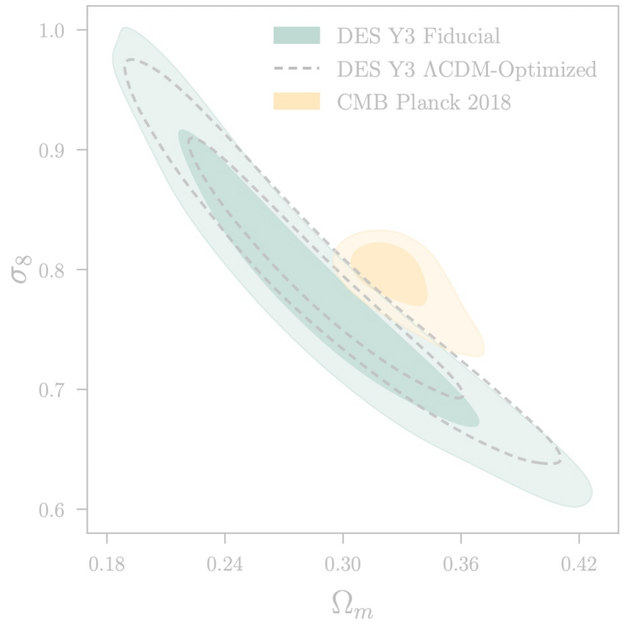


Learned summary statistic



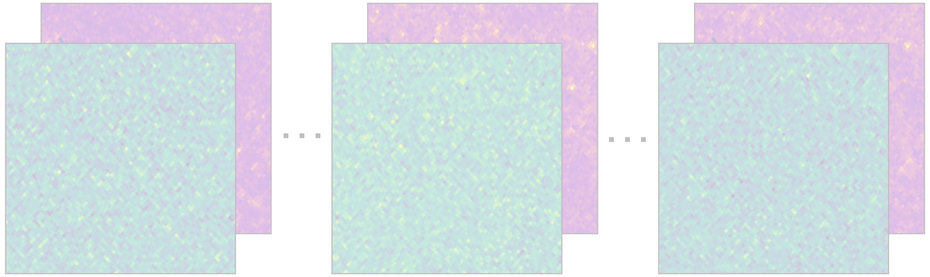
N. Perraudin et al. 2019

Simulation based inference



L. F. Secco et al. 2022

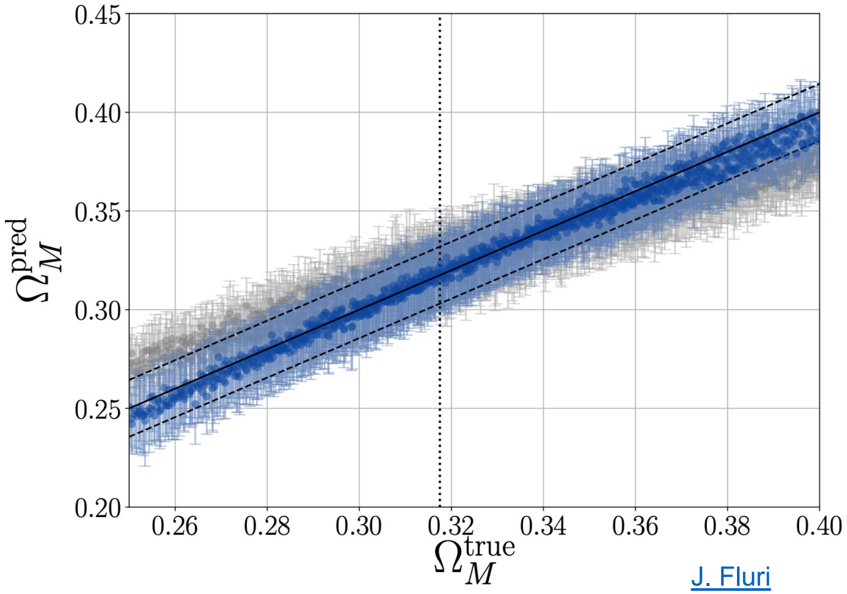
Forward model



Learned summary statistic

General

- Compression $\quad d_\theta \leq d_s \ll d_x$
  - High dimensional input $\quad x \in \mathbb{R}^{d_x}$
  - Low dimensional output $\quad s(x) \in \mathbb{R}^{d_s}$
- Preserve information
  - Ideally Bayesian sufficiency $\quad p(\theta|x) = p(\theta|s(x))$



J. Fluri et al. 2021

Learned

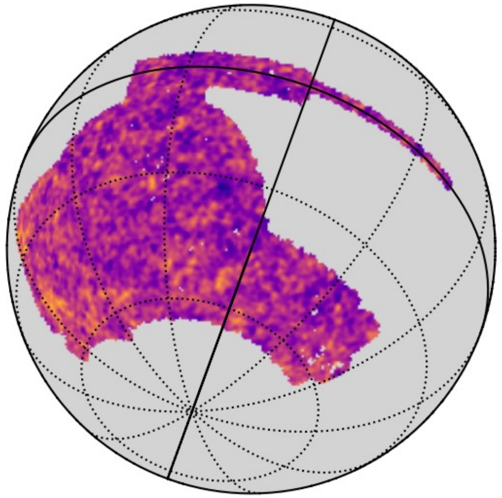
- Loss function
  - Simplest choice: Mean squared error
  - More specialized: Information maximizing loss

$$\underbrace{\mathrm{Cov}_\theta(s) \leq \frac{\partial \Psi_\theta(s)}{\partial \theta} \mathbf{I}(\theta)^{-1} \frac{\partial \Psi_\theta(s)}{\partial \theta}^T}_{\text{Cramér–Rao bound}} \iff -\log\det(\mathbf{I}(\theta)) \leq \overbrace{\log\det\left(\mathrm{Cov}_\theta(s)\right) - 2\log\left|\det\left(\frac{\partial \Psi_\theta(s)}{\partial \theta}\right)\right|}^{\text{Loss function}}$$

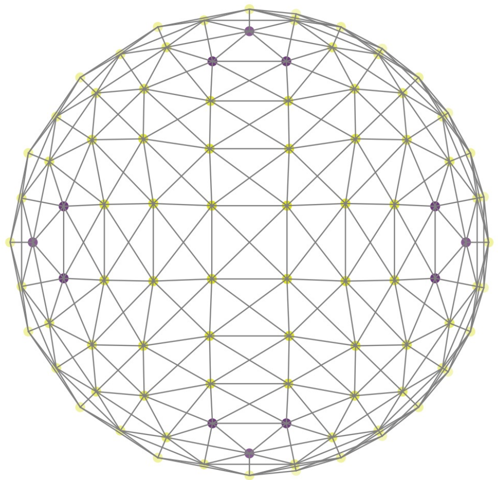
Network architecture

- Large area of the sky → Non-negligible curvature
- Complex survey mask → Graph representation advantageous
- Pixelation of the sphere (HEALPix)

DeepSphere graph convolutional neural networks (https://github.com/deepsphere)

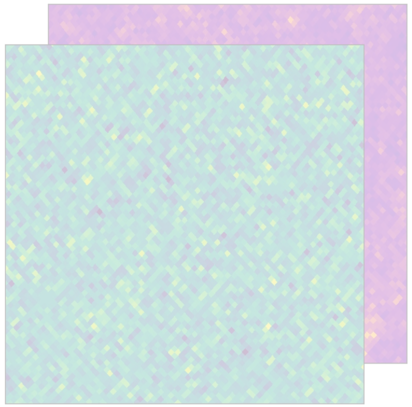


N. Perraudin et al. 2019

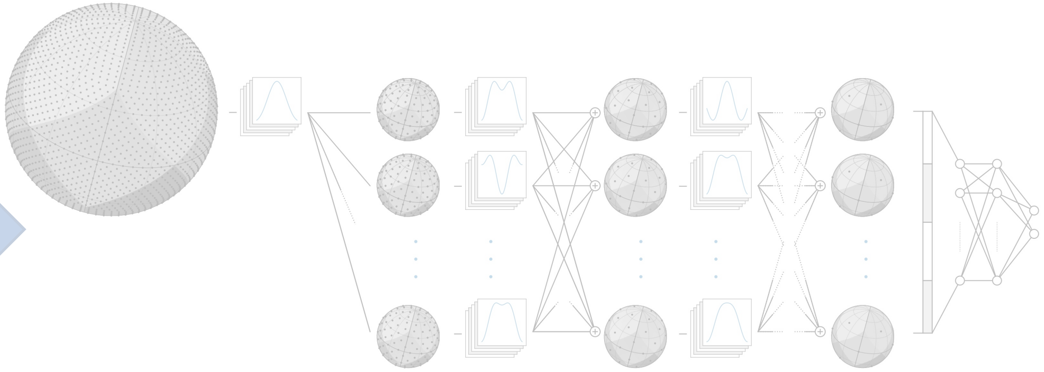


Inference with deep learning

Observation

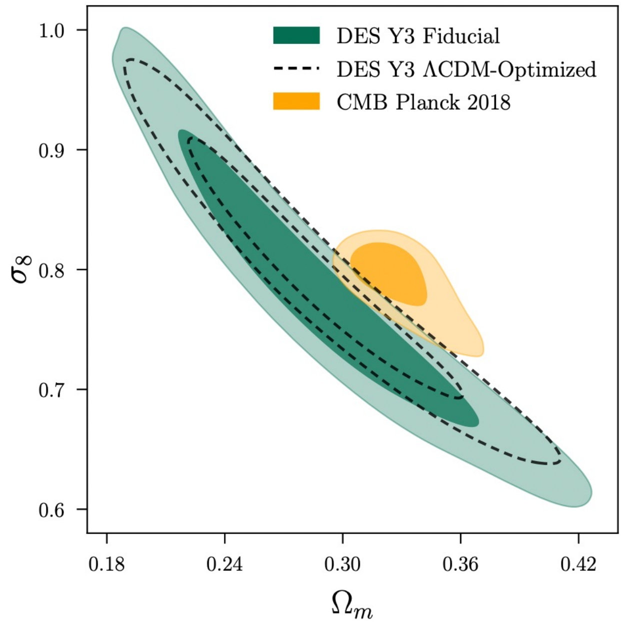


Learned summary statistic



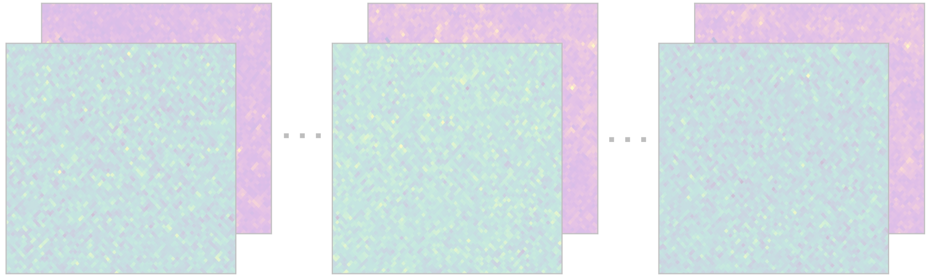
N. Perraudin et al. 2019

Simulation based inference



L. F. Secco et al. 2022

Forward model



Simulation based inference

Goal: Posterior distribution $p(\theta|x) \overset{\text{if sufficient}}{=} p(\theta|s(x))$

- GP ABC J. Fluri et al. 2021
  - Approximate Bayesian (ABC) computation for a kernel K_h

    $$p_{ABC}\left(\theta \mid s_{\text{obs}}\right) \propto \int K_h\left(\left\|s-s_{\text{obs}}\right\|\right) p(s \mid \theta) p(\theta) \mathrm{d} s$$

    - Evaluate on the parameter grid
  - Gaussian process (GP) regression to emulate between grid points
- Neural density estimation
  - Learn $p(s|\theta)$ with a normalizing flow, mixture auf Gaussians, …
  - Get $p(\theta|s_{\text{obs}})$ from Bayes' theorem

Conclusion

Constraining cosmology with deep learning

- Multiple probes of large scale structure
  - Weak lensing
  - Galaxy clustering
- Forward model
  - Contains all physics
- Learned summary statistic
  - Loss function
  - Network architecture
- Parameter inference
  - Approximate Bayesian computation
  - Neural density estimation

Thanks for your attention