Title: Probing primordial non-Gaussianity by reconstructing the initial conditions with machine learning

Speakers: Xinyi Chen

Series: Cosmology & Gravitation

Date: November 21, 2023 - 11:00 AM

URL: https://pirsa.org/23110071

Abstract: Inflation remains one of the enigmas in fundamental physics. While it is difficult to distinguish different inflation models, information contained in primordial non-Gaussianity (PNG) offers a route to break the degeneracy. In galaxy surveys, the local type PNG is usually probed by measuring the scale-dependent bias in the galaxy power spectrum on large scales, where cosmic variance and systematics are also large. Other types of PNG need bispectrum, which is computationally challenging and is contaminated by gravity. I will introduce a new approach to measuring PNG by using the reconstructed density field, a density field reversed to the initial conditions from late time. With the reconstructed density field, we can fit a new template at the field level, or compute a near optimal bispectrum estimator, to constrain PNG. By reconstructing the initial conditions, we remove the nonlinearity induced by gravity, which is a source of confusion when measuring PNG. Near optimal bispectrum estimator mitigates computational challenges. This new approach shows strong constraining power, offers an alternative way to the existing method with different systematics, and also follows organically the procedure of baryon acoustic oscillation (BAO) analysis in large galaxy surveys. I will present a reconstruction method using convolutional neural networks that significantly improves the performance of traditional reconstruction algorithms in the matter density field, which is crucial for more effectively probing PNG. This pipeline can enable new observational constraints on PNG from the ongoing Dark Energy Spectroscopic Instrument (DESI) and Euclid surveys, as well as from upcoming surveys, such as that of the Nancy Grace Roman Space Telescope.

Zoom link https://pitp.zoom.us/j/92361466496?pwd=ZlljUGlKaTVlSFZIV21NUHNGY2RRUT09

Probing primordial non-Gaussianity by reconstructing the initial conditions with machine learning

Xinyi Chen
Yale University

w/ Nikhil Padmanabhan, Daniel Eisenstein, Fangzhou (Albert) Zhu, and Sasha Gaines

Cosmology Seminar, Perimeter Institute, 11/21/23

Image: D. Schlegel

Planck, ACT, Simons Observatory, CMB-S4, …

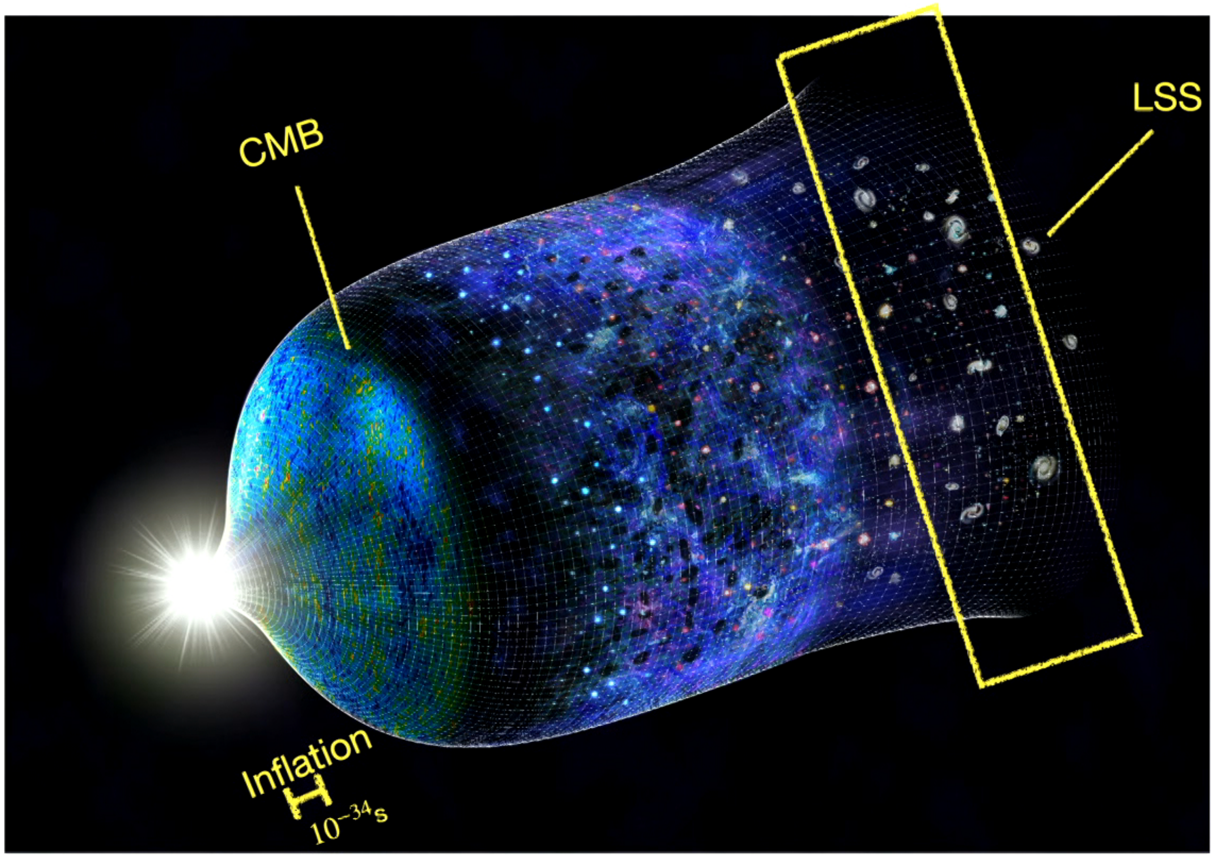


Image: Nicolle R. Fuller, National Science Foundation

DESI, *Euclid*, *Roman*, …

2

Understand the mechanism behind inflation

- Inflation seeded the **density fluctuations** that we can observe today
- **Primordial non-Gaussianities (PNG)**:
  - Deviations from the initial Gaussian density fluctuations. Consequence of many inflation models
  - Robust probe of dynamics during inflation
- The size of PNG — f_{NL}: multiple types —local, equilateral, orthogonal

Local type $f_{\rm NL}^{\rm loc}$

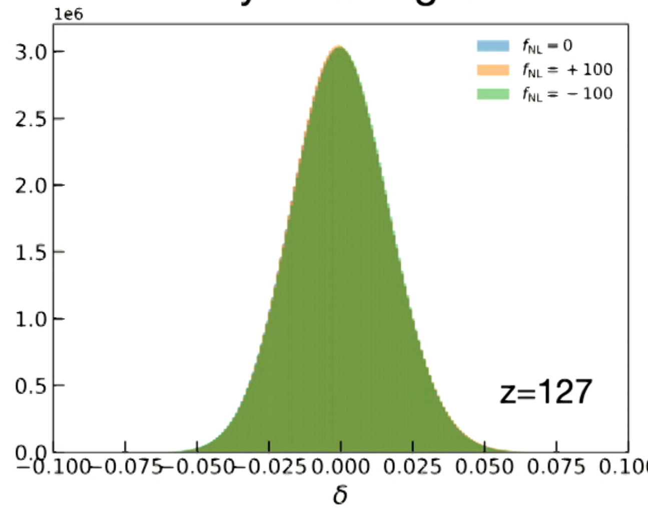
$$\Phi(\boldsymbol{x}) = \phi(\boldsymbol{x}) + f_{\rm NL}^{\rm loc}\{\phi^2(\boldsymbol{x}) - \langle\phi^2(\boldsymbol{x})\rangle\} + \ldots$$

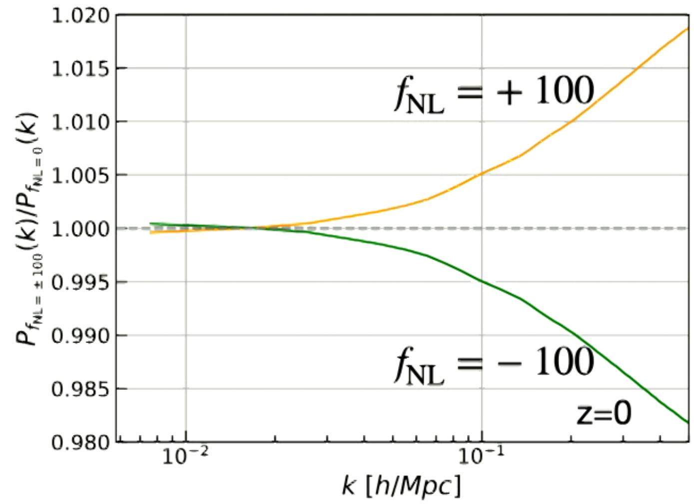
Primordial potential   Gaussian field

- Sensitive probe of **multi-field models**
- Multi-field: $|f_{\rm NL}^{\rm loc}| > 1$, single field $|f_{\rm NL}^{\rm loc}| < 0.01$

A sensitivity of $|f_{\rm NL}^{\rm loc}| < 1$: $\sigma(f_{\rm NL}^{\rm loc}) < 1$

Very small signal

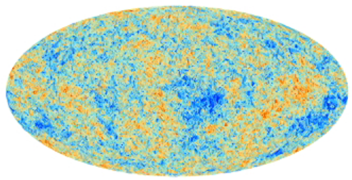




Matter density field

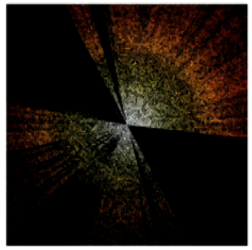
4

Status of CMB



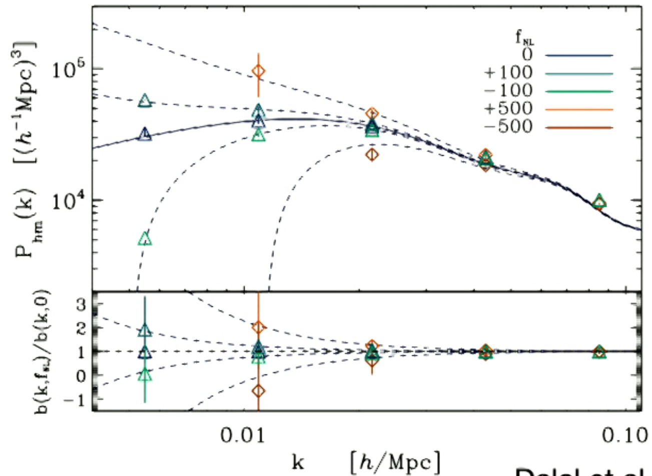
- Current best: 0.9±5.1 (Planck Collaboration 2020)
- Limited by **2D** nature
- Only a factor of 2 improvement in future
- CMB secondary probes x LSS (CMB lensing, kSZ etc)

Status of LSS



- Current best: -12±21 (eBOSS DR16 QSO, Mueller et al. 2022)
- Many **more modes from 3D**
- Usual technique: scale-dependent bias on galaxy power spectrum
  - Systematics
  - **Cosmic variance** on large scales
  - Forecast DESI $\sigma(f_{\rm NL}) \sim 10$ (Sailer et al. 2021)
- **Adding Bispectrum -> tighter constraints**
  - e.g. a factor of ~3 Pk -> Bk, a factor of ~4 Pk -> Pk+Bk (*SPHEREx*, Dore et al. 2014)
  - **Large bispectrum from gravity** Reconstruction
  - **Large data vectors**

Near-optimal 2-pt bispectrum estimator



Dalal et al. 2008

$$\Delta b \propto \frac{f_{\rm NL}}{k^2 T(k)}$$

Transfer function

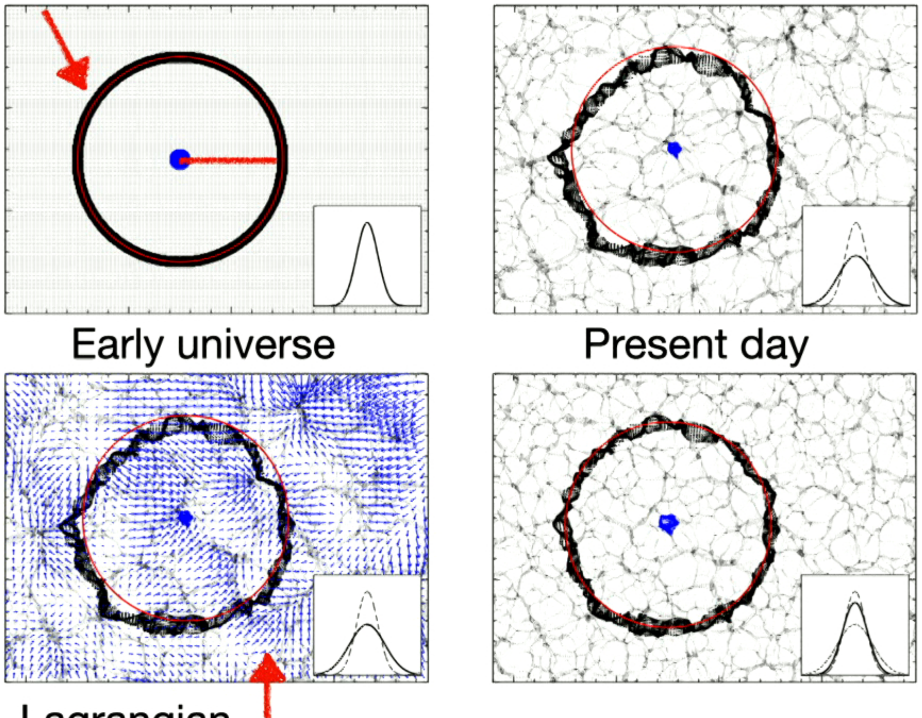
6

New approach to constraining PNG

- Reconstructing the density field
- Fitting templates at field level
- Computing and fitting a near-optimal 2-pt bispectrum estimator

Reconstruction of the initial conditions: reverse a late-time density field back to initial density field

Acoustic feature



Lagrangian displacement field $\nabla \cdot \Psi = -\delta$

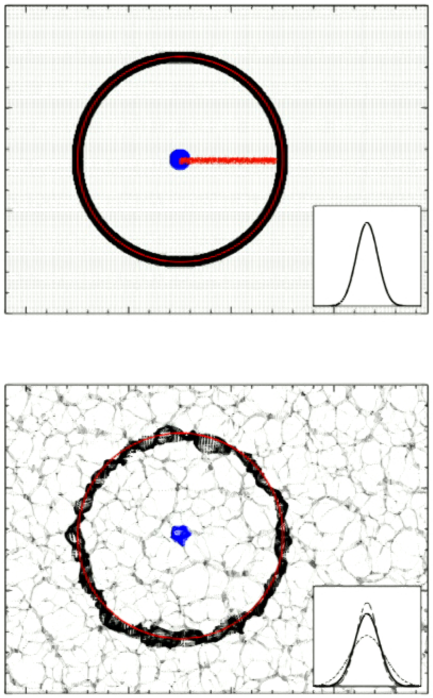
Padmanabhan et al. 2012

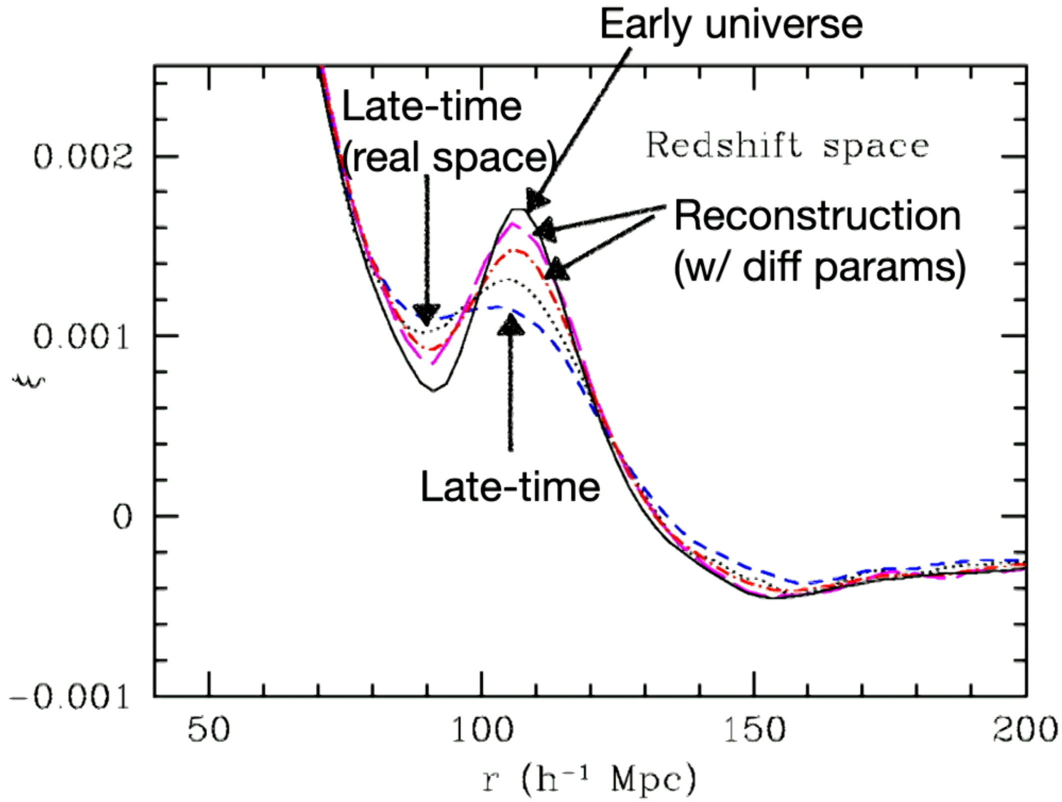
Idea behind standard reconstruction (Eisenstein et al. 2007)

- The initial density field in the early universe is very smooth
- As the universe evolves, the black points spread out which broadens the acoustic feature
- Estimate the displacement field and move the particles back to their initial positions.

- Reduces the distance error in BAO analysis by a factor of ~2

Restoring the BAO feature for BAO analysis

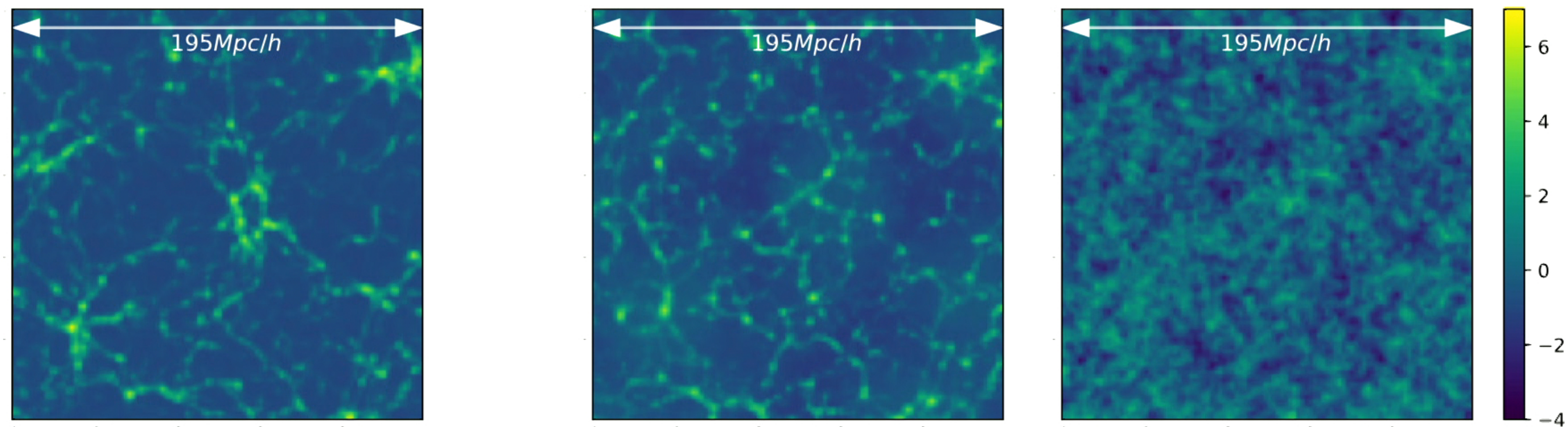




Eisenstein et al. 2007

10

Density field reconstructed by the standard reconstruction algorithm still nonlinear

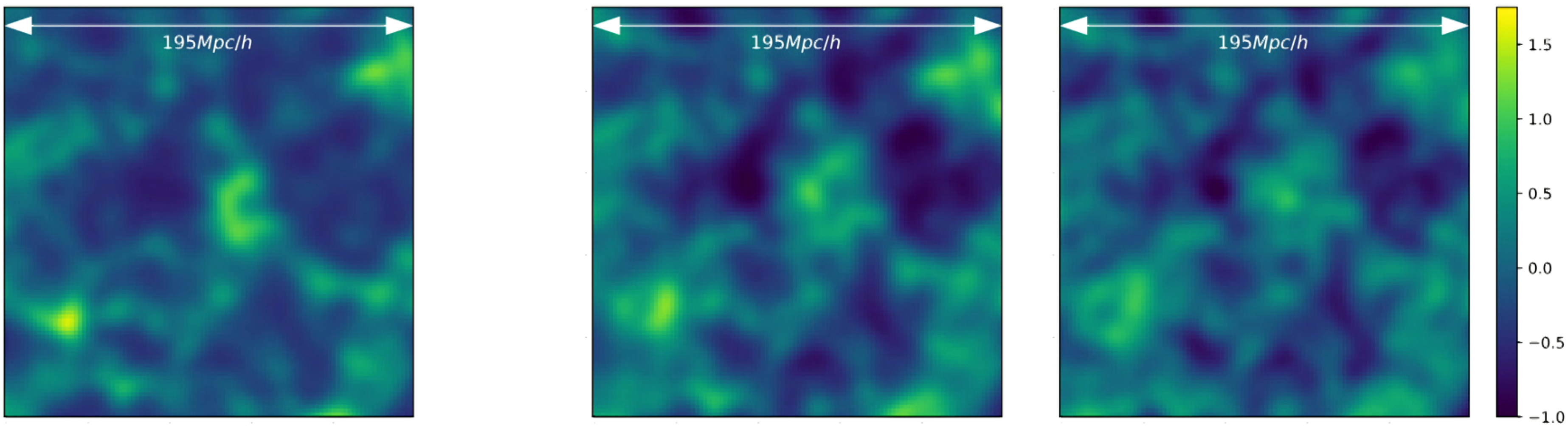


Late-time | Standard reconstruction | Initial

Matter density fields at high resolution (1024^3 particles in 1 Gpc/h box) at z=0, on a 512^3 grid, using Quijote simulations (Villaescusa-Navarro et al. 2020)

Density field reconstructed by the standard reconstruction algorithm still nonlinear

Smoothed at 5 Mpc/h

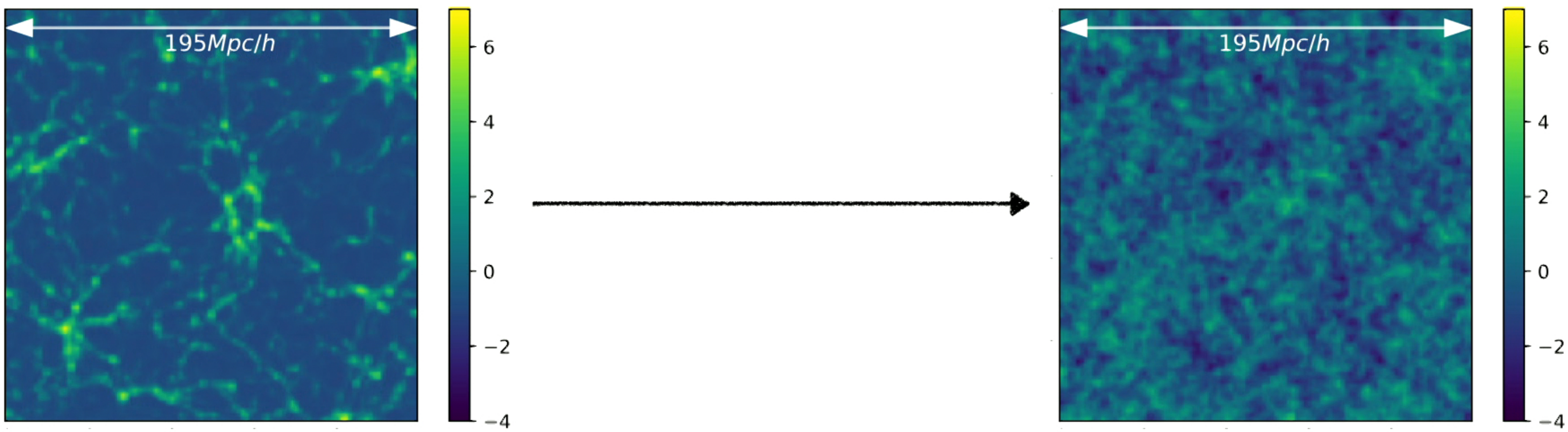


Late-time | Standard reconstruction | Initial

Matter density fields at high resolution (1024^3 particles in 1 Gpc/h box) at z=0, on a 512^3 grid, using Quijote simulations (Villaescusa-Navarro et al. 2020)

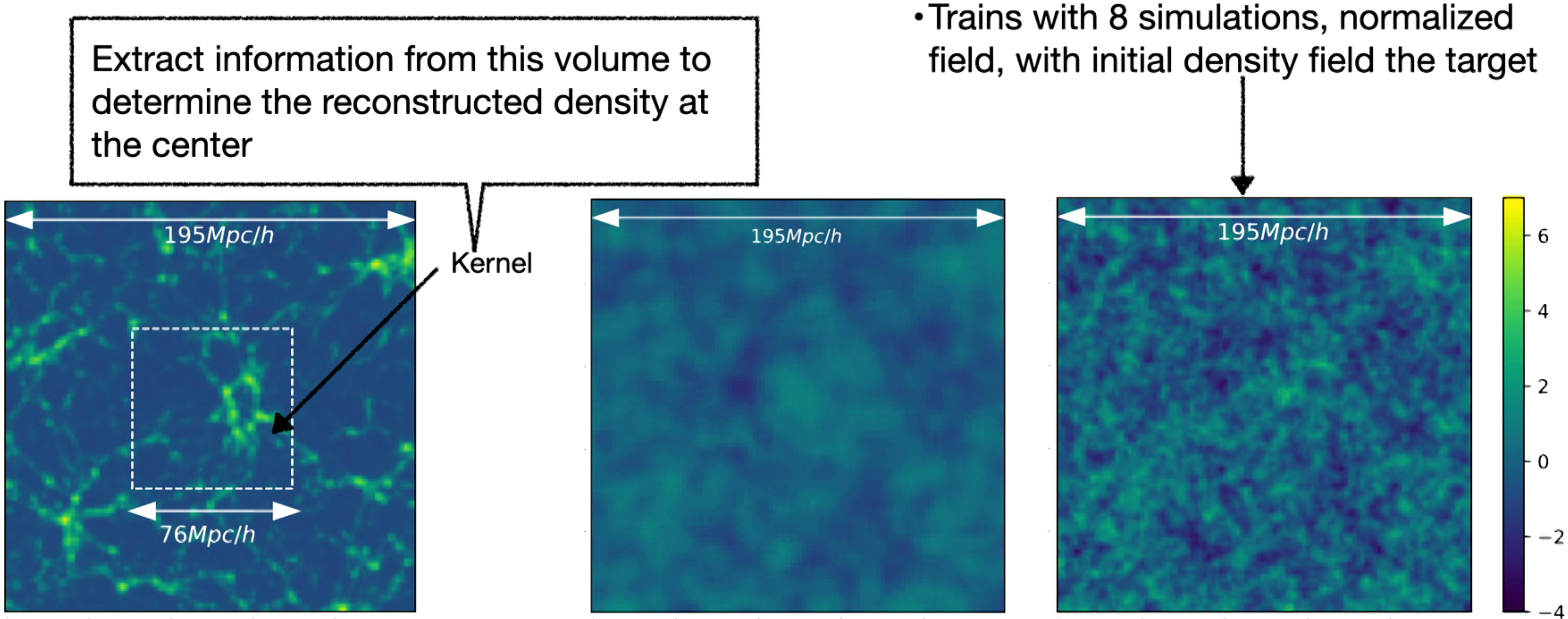
A new reconstruction method

A hybrid method that combines convolutional neural network (CNN) with traditional algorithm based on perturbation theory (Shallue & Eisenstein 2023, **Chen** et al. 2023)



Density field by CNN trained with late-time density field still nonlinear

- Considers neighbor points by grouping them in a batch
- Trains with 8 simulations, normalized field, with initial density field the target

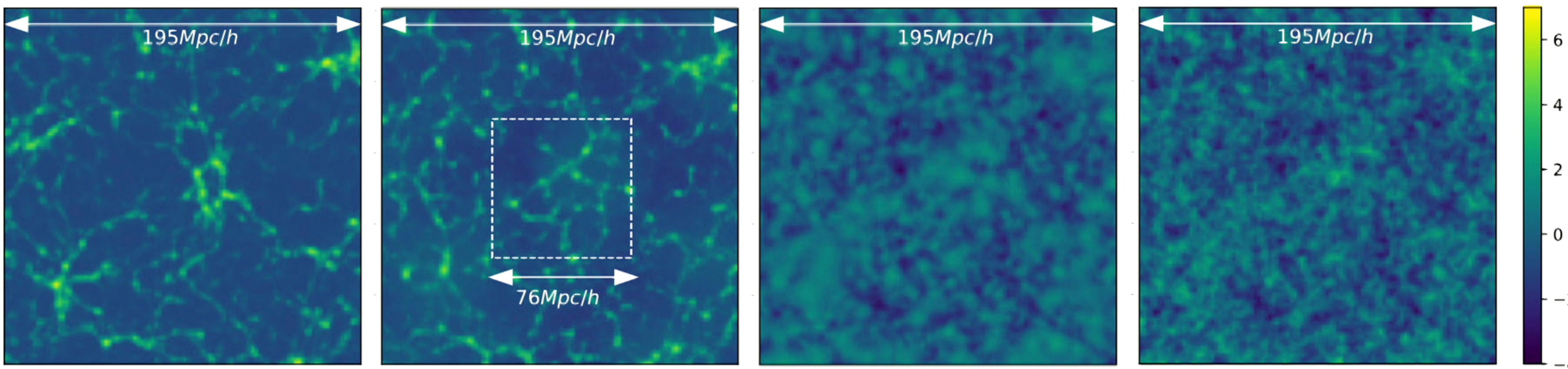


Late-time CNN trained w/ late-time field Initial

14

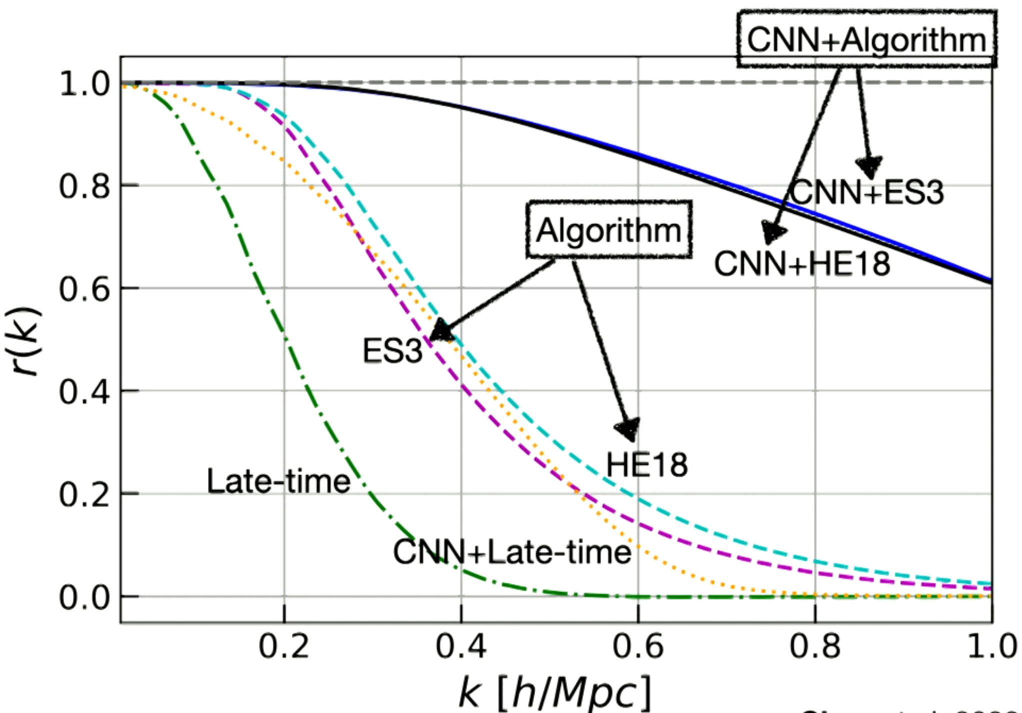
Training with *reconstructed* density field significantly improves performance

CNN is relatively local. Algorithm provides good approximation on large scale (Zel'dovich approximation is only valid for large scales). CNN then reconstructs further on smaller scales.



Late-time | Standard recon | CNN trained w/ standard recon field | Initial

CNN improves cross-correlation



Chen et al. 2023

$$r(k) = \frac{\langle \delta^*(\boldsymbol{k})\delta_{\text{ini}}(\boldsymbol{k})\rangle}{\sqrt{\langle \delta^2(\boldsymbol{k})\rangle\langle \delta^2_{\text{ini}}(\boldsymbol{k})\rangle}}$$

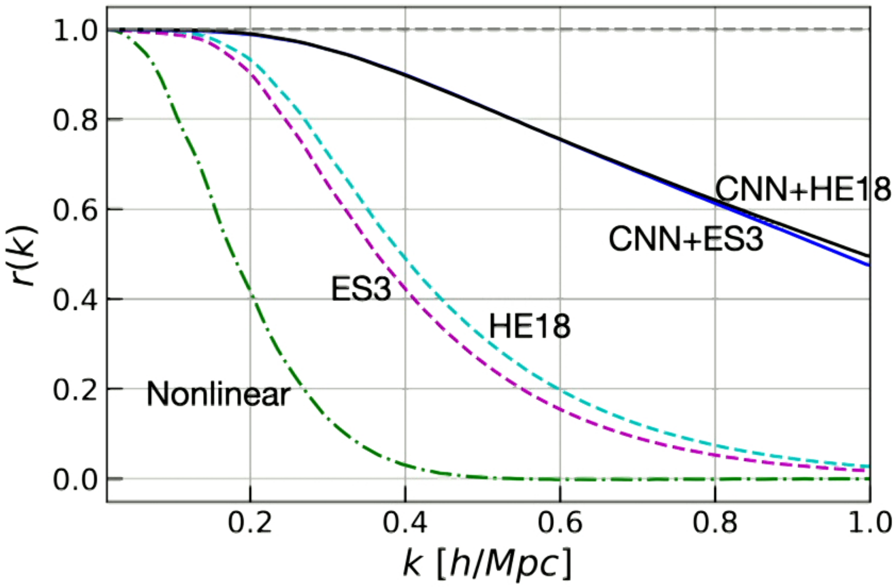
- CNN+Algorithm performs significantly better than algorithms alone and CNN+Late-time density field
- CNN+ES3 and CNN+HE18 are similar

Two reconstruction algorithms:

- Eisenstein et al. 2007, **ES3**, i.e., standard
- Hada & Eisenstein 2018, **HE18**

Real space matter field z=0

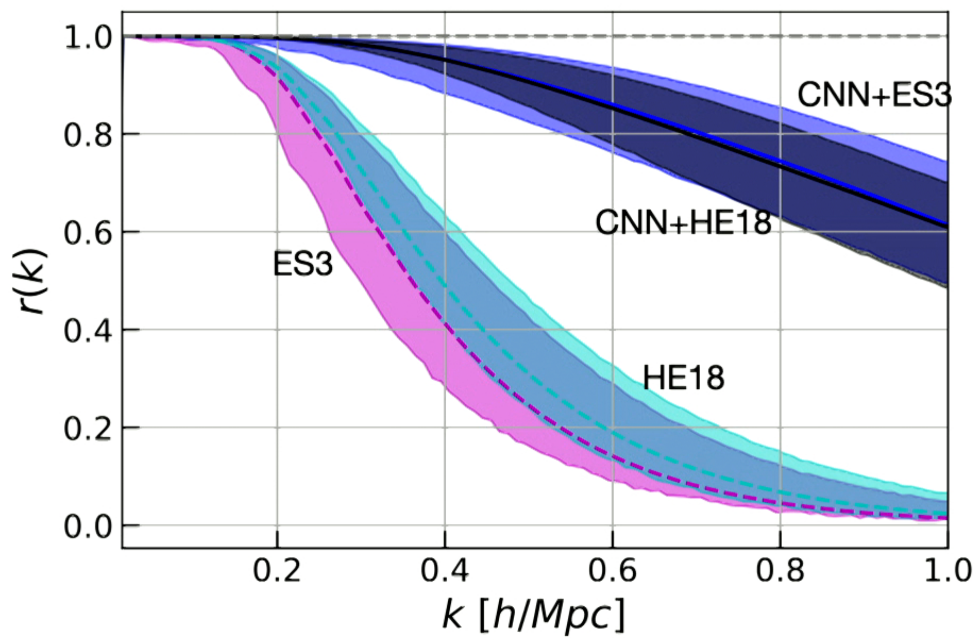
Redshift space cross-correlation also much better



Redshift space

•Very similar trend to real space results

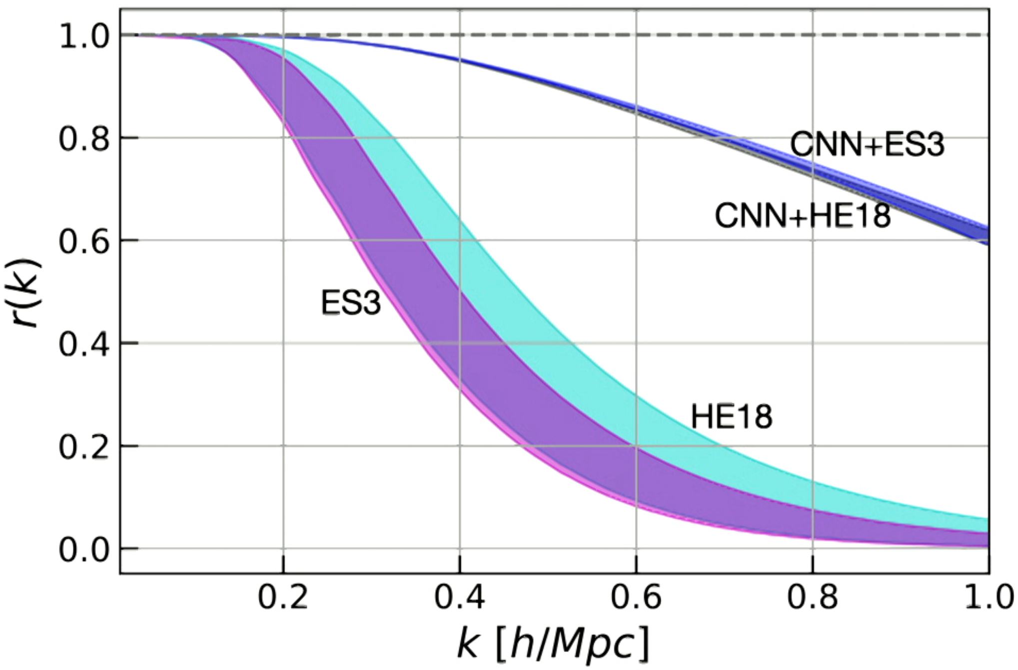
CNN is robust to cosmologies



- Model trained on the fiducial cosmology applied to other cosmologies
- Generalizes well to a wide range of cosmologies

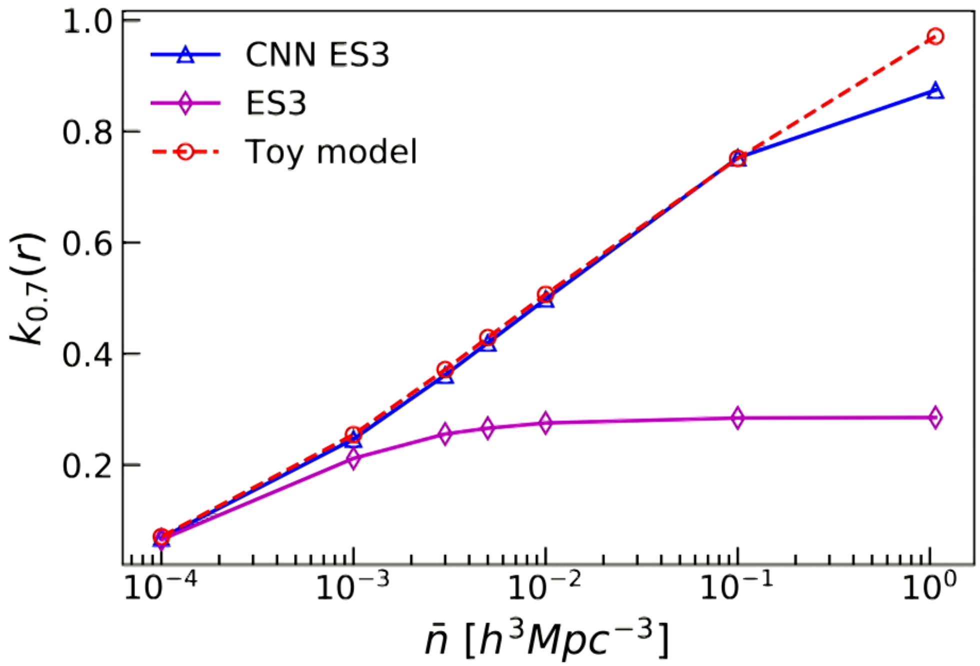
18

CNN minimizes the variance from input reconstruction



- Model removes much of the variance in the input reconstruction (different algorithms, different smoothing scales)

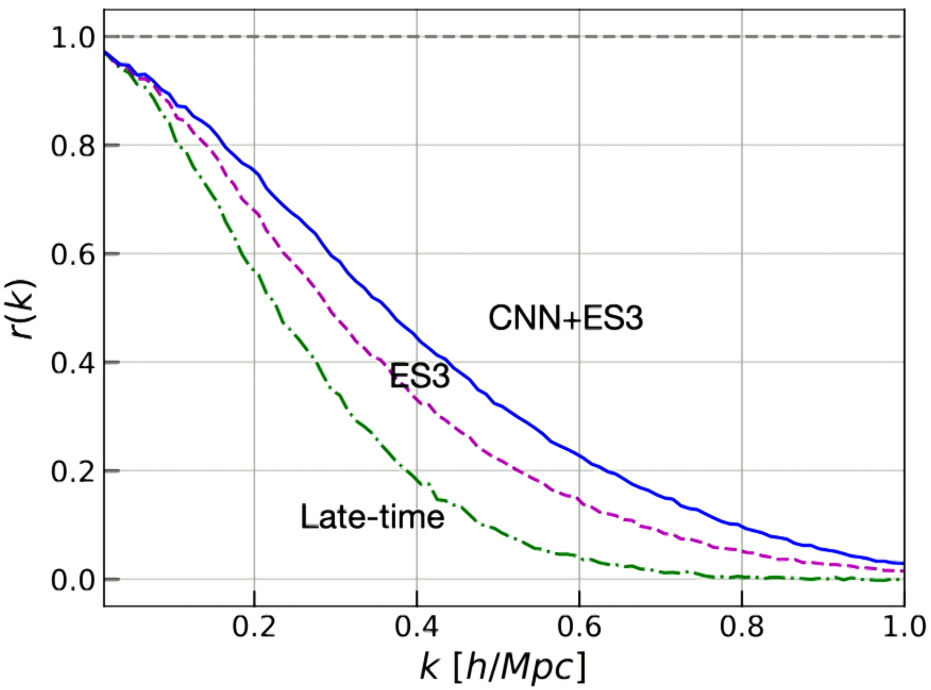
Shot noise is a challenge



- Same architecture retrained with subsampled fields at lower densities
- CNN is sensitive to change in number density
- Toy model captures CNN's behavior
- Advantage with higher number densities, inform design of future surveys

20

Hybrid recon boosts traditional algorithms in halo fields too



z=1

Number density=2.3x10^{-4} h^{3}Mpc^{-3}

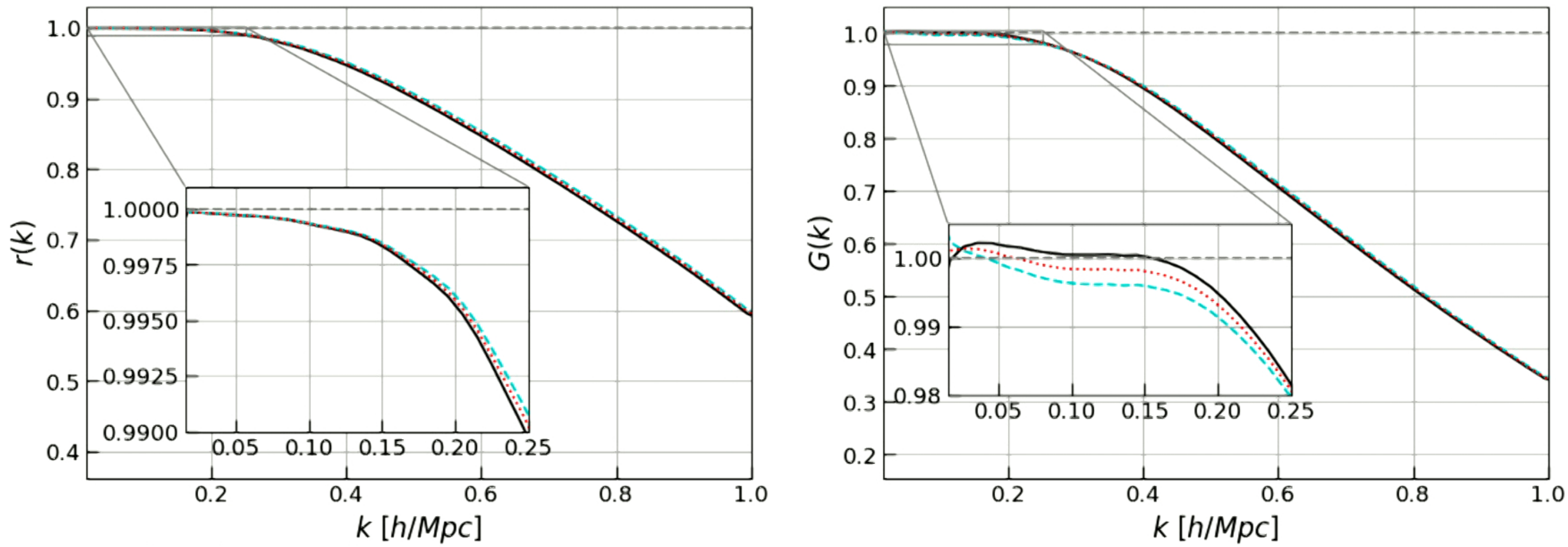
Bias=2.3

Now adding PNG…

Three categories of sims: $f_{\rm NL} = 0, f_{\rm NL} = +\,100, f_{\rm NL} = -\,100$

Model trained with no PNG works for PNG

$$G(k) = \frac{\langle \delta^*(\boldsymbol{k})\delta_{\mathrm{ini}}(\boldsymbol{k})\rangle}{\langle \delta_{\mathrm{ini}}^2(\boldsymbol{k})\rangle}$$



$$r(k) = \frac{\langle \delta^*(\boldsymbol{k})\delta_{\mathrm{ini}}(\boldsymbol{k})\rangle}{\sqrt{\langle \delta^2(\boldsymbol{k})\rangle\langle \delta_{\mathrm{ini}}^2(\boldsymbol{k})\rangle}}$$

— $f_{\mathrm{NL}} = +100$

⋯ $f_{\mathrm{NL}} = 0$

-- $f_{\mathrm{NL}} = -100$

23

CNN+HE18

New approach to constraining PNG

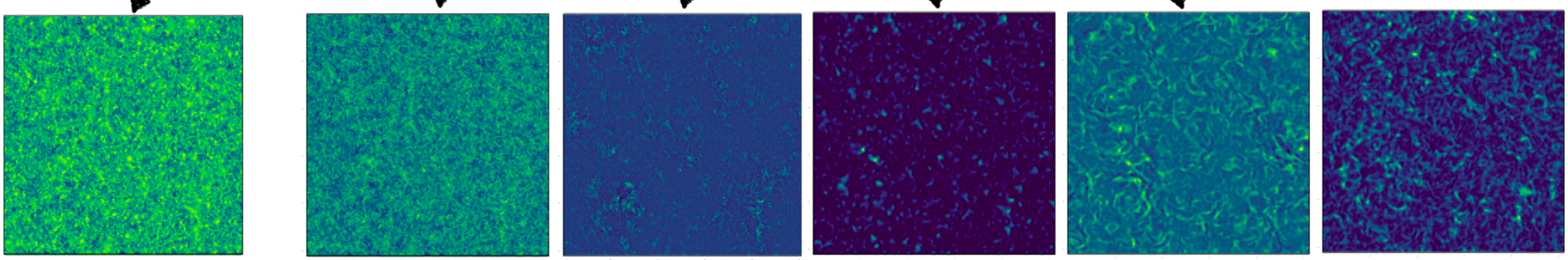
- Reconstructing the density field
- **Fitting templates at field level**
- Computing and fitting a near-optimal 2-pt bispectrum estimator

Templates for fitting f_{NL}

- δ_G=No PNG IC
- $\delta_{f_{\mathrm{NL}}} = \phi_G^2(k) M_\phi(k)$
- δ^2,δ_{∇^2},δ_{s^2} all computed using δ_G

Gaussian　　Growth　　Shift　　Tidal

$$\delta_{\mathrm{CNN}} = b_G \delta_G + f_{\mathrm{NL}} \delta_{f_{\mathrm{NL}}} + b_2 \delta^2 + b_{\nabla^2} \delta_{\nabla^2} + b_{s^2} \delta_{s^2} + \ldots$$



25

Small error but fits are slightly biased

- Errors in 1 Gpc volume, std of 90 sims
- With 3 Mpc/h smoothing for the quadratic fields
- k cut at 0.1 h/Mpc

Fits for $\delta_{\rm CNN}$
z=0 real space

	b_G	$f_{\rm NL}$	b_2	b_{∇^2}	b_{s^2}
$f_{\rm NL}=0$	0.9996±0.0007	-11.8±4.4	0.005±0.001	-0.013±0.001	0.010±0.001
$f_{\rm NL}=+100$	1.0011±0.0007	80.7±4.5	0.004±0.001	-0.013±0.001	0.011±0.001
$f_{\rm NL}=-100$	0.9980±0.0007	-103.7±4.3	0.005±0.001	-0.013±0.001	0.010±0.001

Chen, Padmanabhan & Eisenstein in prep.

Accounting for the shift in the mean at $f_{\rm NL}=0$:

$f_{\rm NL}=+100$: ~+92

$f_{\rm NL}=-100$: ~-92

Slightly biased

For >2 Gpc survey volume (e.g. DESI):

$\sigma(f_{\rm NL})\sim 2$

Much lower error

Small error but fits are slightly biased

- Reconstruction **significantly reduces nonlinearities** at the second order, and still **preserves most PNG and gives tighter constrains**

Fits for $\delta_{\rm CNN}$

z=0 real space

~2x improvement

	b_G	$f_{\rm NL}$	b_2	b_{∇^2}	b_{s^2}
$f_{\rm NL}=0$	0.9996±0.0007	-11.8±4.4	0.005±0.001	-0.013±0.001	0.010±0.001
$f_{\rm NL}=+100$	1.0011±0.0007	80.7±4.5	0.004±0.001	-0.013±0.001	0.011±0.001
$f_{\rm NL}=-100$	0.9980±0.0007	-103.7±4.3	0.005±0.001	-0.013±0.001	0.010±0.001

Chen, Padmanabhan & Eisenstein in prep.

Fits for $\delta_{\rm NL}$

- k cut at 0.05 h/Mpc

	$f_{\rm NL}$	b_2	b_{∇^2}	b_{s^2}
$f_{\rm NL}=0$	-12.8±11.0	0.808±0.006	-1.002±0.008	0.192±0.004
$f_{\rm NL}=+100$	82.3±10.9	0.808±0.006	-1.003±0.009	0.192±0.004
$f_{\rm NL}=-100$	-108.1±11.3	0.807±0.005	-1.002±0.007	0.192±0.004

From F2 kernel: $b_2=\frac{17}{21}\sim 0.81 \quad b_{\nabla^2}=-1 \quad b_{s^2}=\frac{4}{21}\sim 0.19$

Small error but fits are slightly biased

- Errors in 1 Gpc volume, std of 90 sims
- With 3 Mpc/h smoothing for the quadratic fields
- k cut at 0.1 h/Mpc

Fits for δ_{CNN}
z=1 real space

	b_G	f_{NL}	b_2	b_{∇^2}	b_{s^2}
$f_{\mathrm{NL}} = 0$	0.9972±0.0006	0.8±3.1	0.004±0.001	-0.018±0.001	0.017±0.001
$f_{\mathrm{NL}} = +100$	0.9990±0.0006	93.2±3.2	0.004±0.001	-0.018±0.001	0.018±0.001
$f_{\mathrm{NL}} = -100$	0.9955±0.0006	-91.6±3.1	0.004±0.001	-0.017±0.001	0.016±0.001

Chen, Padmanabhan & Eisenstein in prep.

Accounting for the shift in the mean at $f_{\mathrm{NL}} = 0$:

$f_{\mathrm{NL}} = +100$: ~+92

$f_{\mathrm{NL}} = -100$: ~-92

Slightly biased

Small error but fits are slightly biased

- Reconstruction **significantly reduces nonlinearities** at the second order, and still **preserves most PNG and gives tighter constrains**

Fits for δ_{CNN}

z=1 real space

~1.5x improvement

	b_G	f_{NL}	b_2	b_{∇^2}	b_{s^2}
$f_{\mathrm{NL}}=0$	0.9972±0.0006	0.8±3.1	0.004±0.001	-0.018±0.001	0.017±0.001
$f_{\mathrm{NL}}=+100$	0.9990±0.0006	93.2±3.2	0.004±0.001	-0.018±0.001	0.018±0.001
$f_{\mathrm{NL}}=-100$	0.9955±0.0006	-91.6±3.1	0.004±0.001	-0.017±0.001	0.016±0.001

Chen, Padmanabhan & Eisenstein in prep.

Fits for δ_{NL}

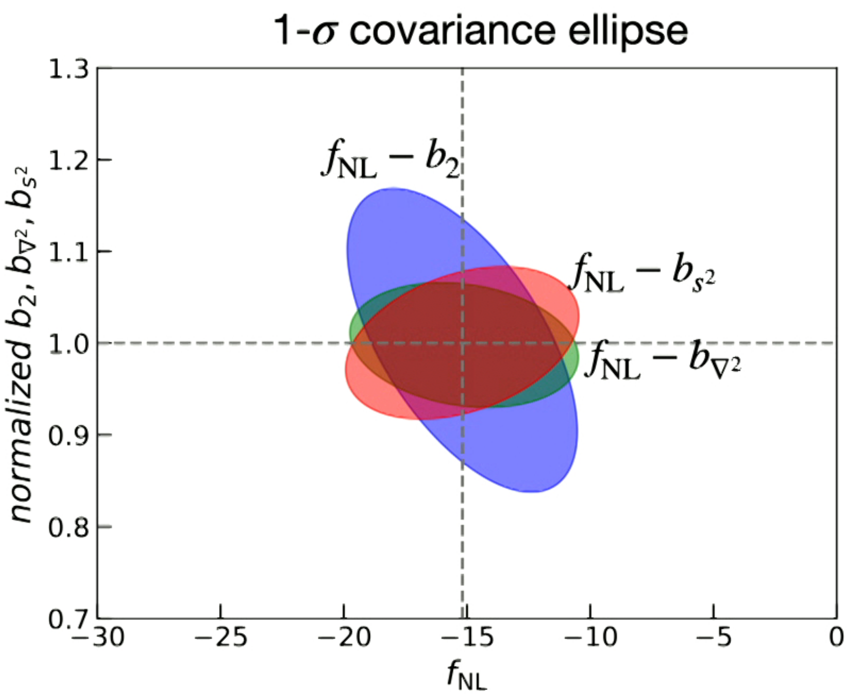
- k cut at 0.05 h/Mpc

	f_{NL}	b_2	b_{∇^2}	b_{s^2}
$f_{\mathrm{NL}}=0$	-7.5±5.3	0.823±0.004	-1.026±0.005	0.201±0.003
$f_{\mathrm{NL}}=+100$	90.6±5.2	0.823±0.005	-1.026±0.006	0.201±0.003
$f_{\mathrm{NL}}=-100$	-105.7±5.5	0.823±0.004	-1.026±0.004	0.201±0.002

From F2 kernel: $b_2=\frac{17}{21}\sim 0.81 \quad b_{\nabla^2}=-1 \quad b_{s^2}=\frac{4}{21}\sim 0.19$

Strong degeneracy between $f_{\rm NL}$ and b_2

z=0



$f_{\rm NL} = 0$

- Cross-correlation coefficient between

$f_{\rm NL} - b_2$: ~-0.6

$f_{\rm NL} - b_{\nabla^2}$: ~-0.2

$f_{\rm NL} - b_{s^2}$: ~0.4

30

Possible sources of bias:

1. Reconstruction distorts PNG
2. Imperfect reconstruction, leftover nonlinearity
3. Degeneracy between $f_{\rm NL}$ and other bias terms
4. Higher order terms not included in the model

New approach to constraining PNG

- Reconstructing the density field
- Fitting templates at field level
- **Computing and fitting a near-optimal 2-pt bispectrum estimator**

Near optimal bispectrum estimator

$$\langle \Phi^2 \delta \rangle$$

Reconstructed/Linear

Primordial potential

Transfer function

$$\Phi(\boldsymbol{k}) = \frac{\delta(\boldsymbol{k})}{M_\Phi(k)} \qquad M_\Phi(k) = \frac{2}{3}\frac{k^2 T(k)}{\Omega_{\mathrm{m},0} H_0^2}$$

$$\Phi^2(\boldsymbol{k}) = \int \mathrm{d}\boldsymbol{x}\, e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} \Phi^2(\boldsymbol{x}) = \frac{1}{(2\pi)^3}\int \mathrm{d}\boldsymbol{k}_1 \Phi(\boldsymbol{k}_1)\Phi(\boldsymbol{k}-\boldsymbol{k}_1)$$

$$\langle \Phi^2(\boldsymbol{k})\delta(-\boldsymbol{k})\rangle = \frac{1}{(2\pi)^3}\int \mathrm{d}\boldsymbol{k}_1 M_\Phi(-\boldsymbol{k}) \underbrace{\boxed{\langle \Phi(\boldsymbol{k})\Phi(\boldsymbol{k}-\boldsymbol{k}_1)\Phi(-\boldsymbol{k})\rangle}}_{\text{Primordial bispectrum}}$$

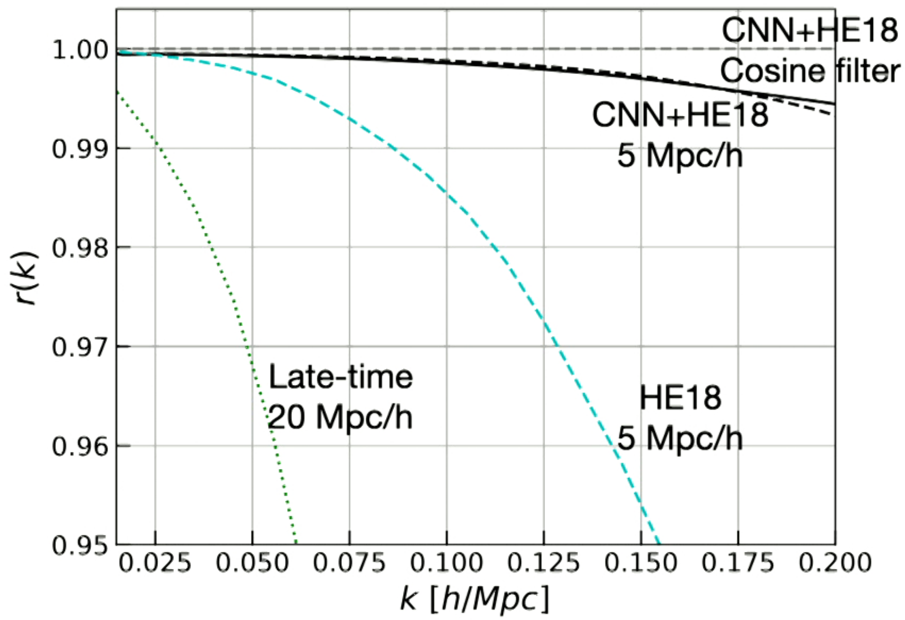
Integral of bispectrum

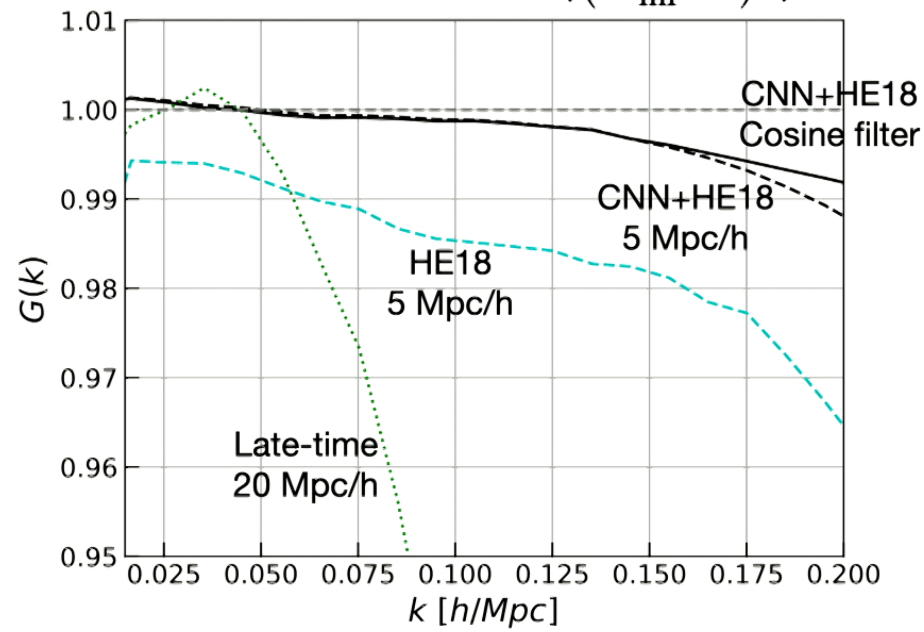
Why near optimal? Maximum likelihood estimation by Schmittfull, Baldauf & Seljak 2015

Reconstructed Φ^2 field cross-correlation close to IC too

$$G(k) = \frac{\langle \Phi^{2*}(\boldsymbol{k})\Phi^2_{\rm ini}(\boldsymbol{k})\rangle}{\langle \left(\Phi^2_{\rm ini}(\boldsymbol{k})\right)^2\rangle}$$

z=0

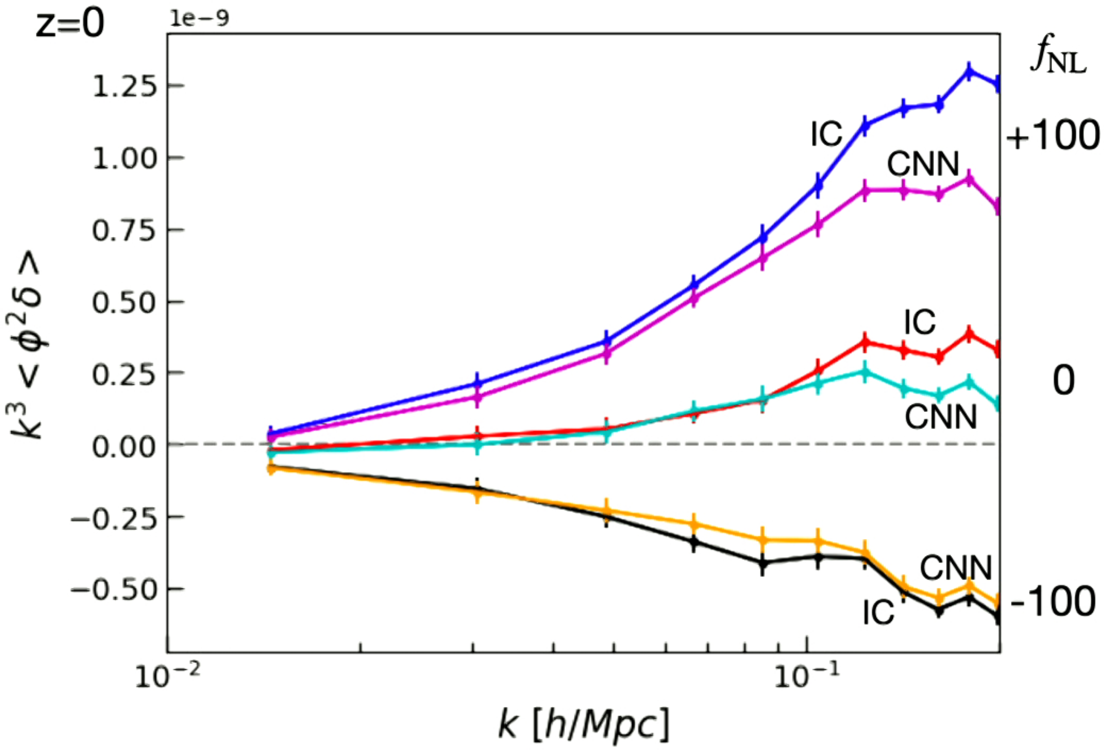




Cosine filter: between k in 0.2-0.25 h/Mpc

$$r(k) = \frac{\langle \Phi^{2*}(\boldsymbol{k})\Phi^2_{\rm ini}(\boldsymbol{k})\rangle}{\sqrt{\langle \left(\Phi^2(\boldsymbol{k})\right)^2\rangle \langle \left(\Phi^2_{\rm ini}(\boldsymbol{k})\right)^2\rangle}}$$

Near optimal bispectrum estimator as a statistic



with cosine filter between
k=0.2-0.25 h/Mpc

- Biased, consistent with template fits
- Single parameter forecast CNN $\sigma(f_{\rm NL})$ ~50, pre-recon $\sigma(f_{\rm NL})$~100 (for k<0.1 h/Mpc) — ~2x improvement
- Optimistic without including other bias terms
- Can compute similar estimator for other fields

35

Plan

- Fitting templates in reality: fitting coefficients together with δ_G, forward model
- Estimate each template term with bispectrum estimator
- High shot noise biased tracer
- Other types of PNG — can be more helpful since can't rely on scale-dependent bias

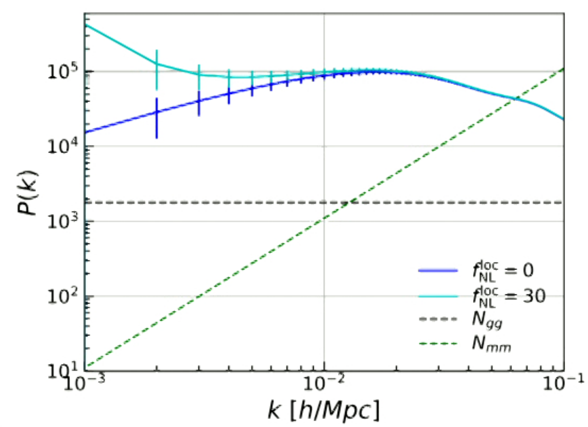
Other near-future plans

Map galaxy field to matter field at same redshift

LSSxCMB kSZ for local type with DESIxACT

- Lower noise on large scales
- Reduce cosmic variance

(Munchmeyer et al. 2018)



38

Summary

- Great opportunities to probe inflationary physics in the current era of cosmology
- Reconstruction with CNN+Algorithm shows promising constraining power for PNG
- Template fits are biased but errors are lower
- Bispectrum estimator consistent with template fitting, ~2x improvement in error with reconstruction

- Plan:
  - Fitting templates in reality: fitting coefficients together with δ_G, forward model
  - Estimate each template term with bispectrum estimator
  - High shot noise biased tracer
  - Other types of PNG — can be more helpful since can't rely on scale-dependent bias

Thank you! 😀