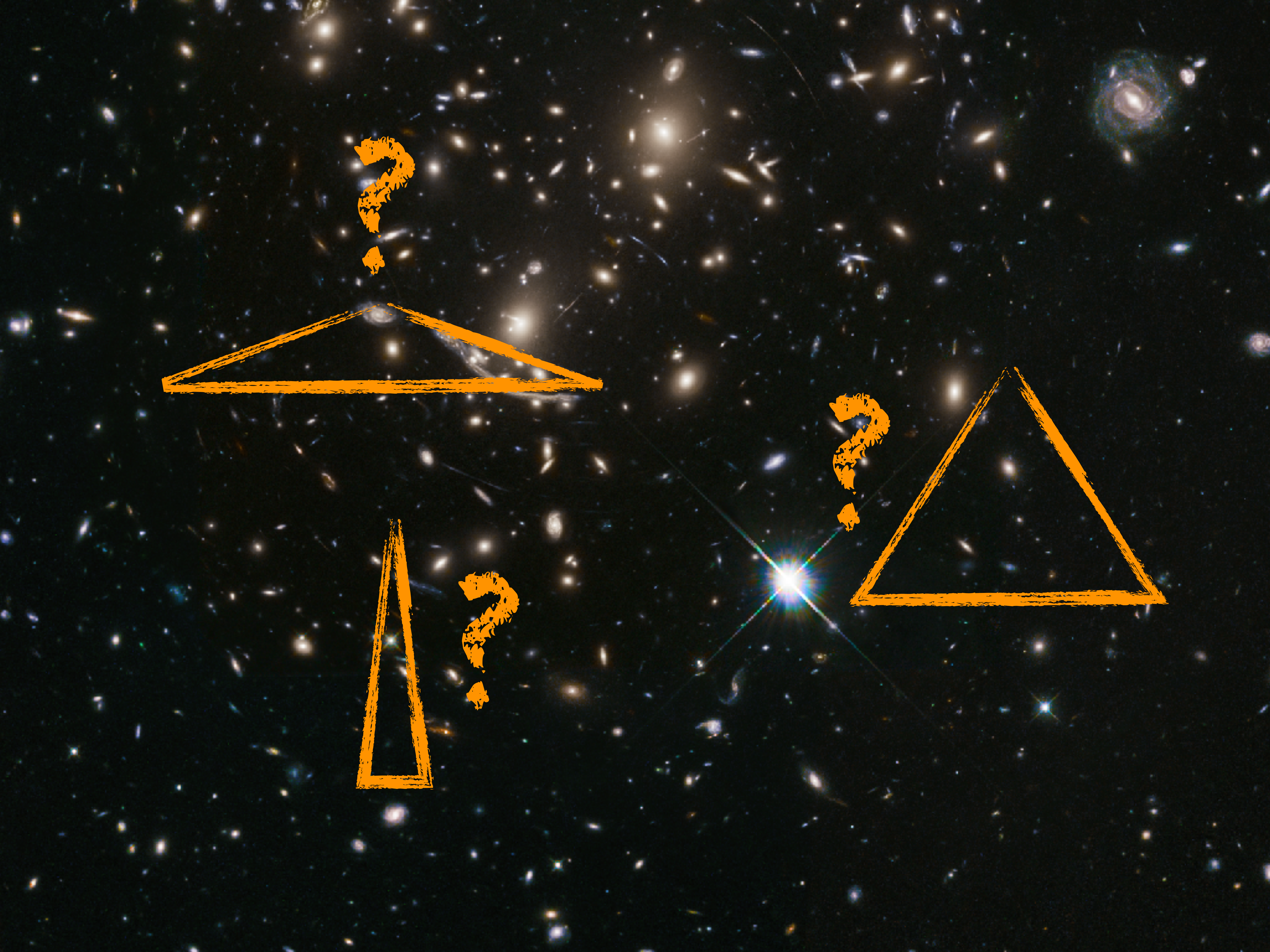
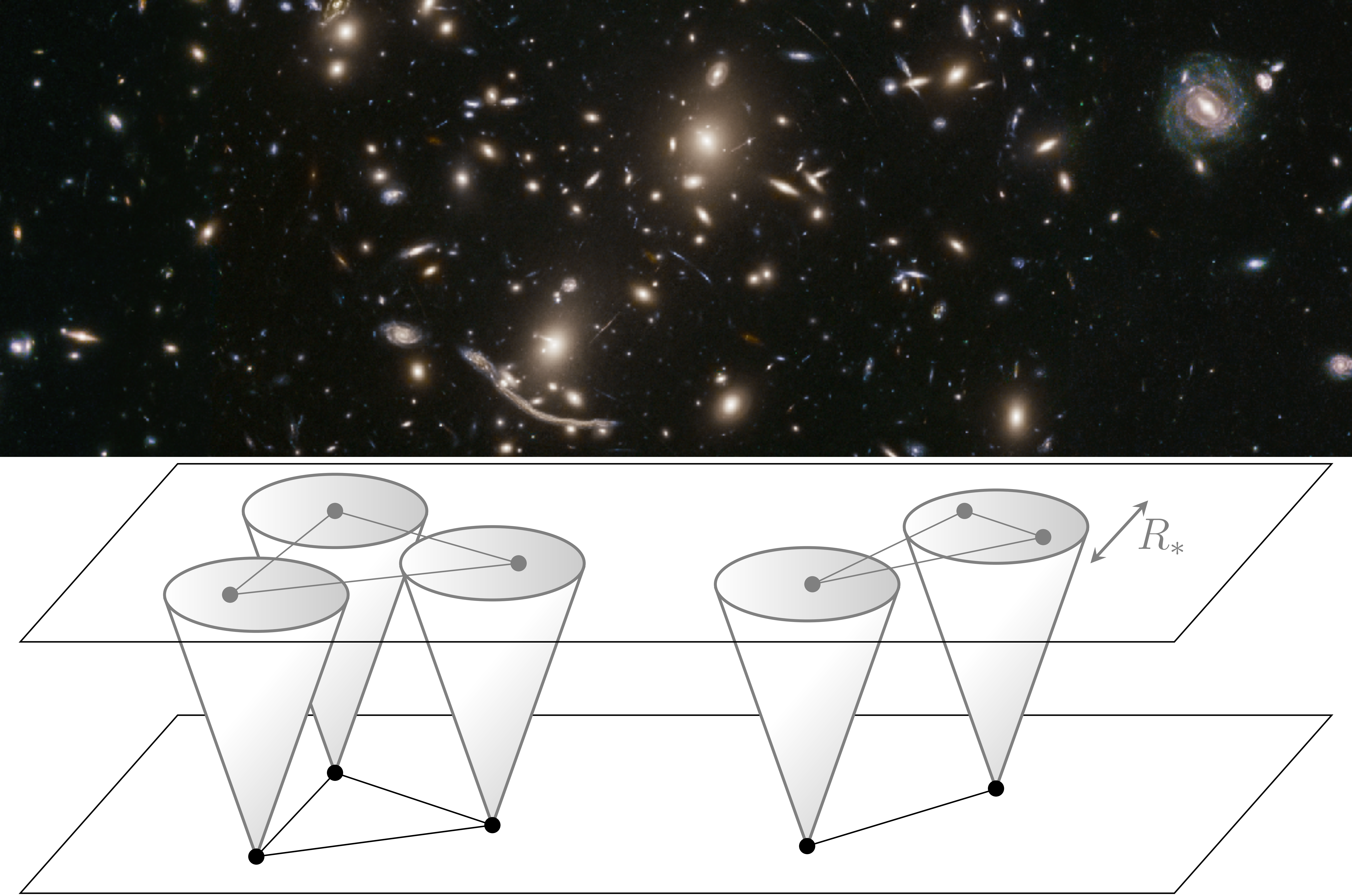
Non-Gaussianity in Maps of Galaxies

Daniel Green
UC San Diego

arXiv: 2112.14645 with Daniel Baumann




R_*

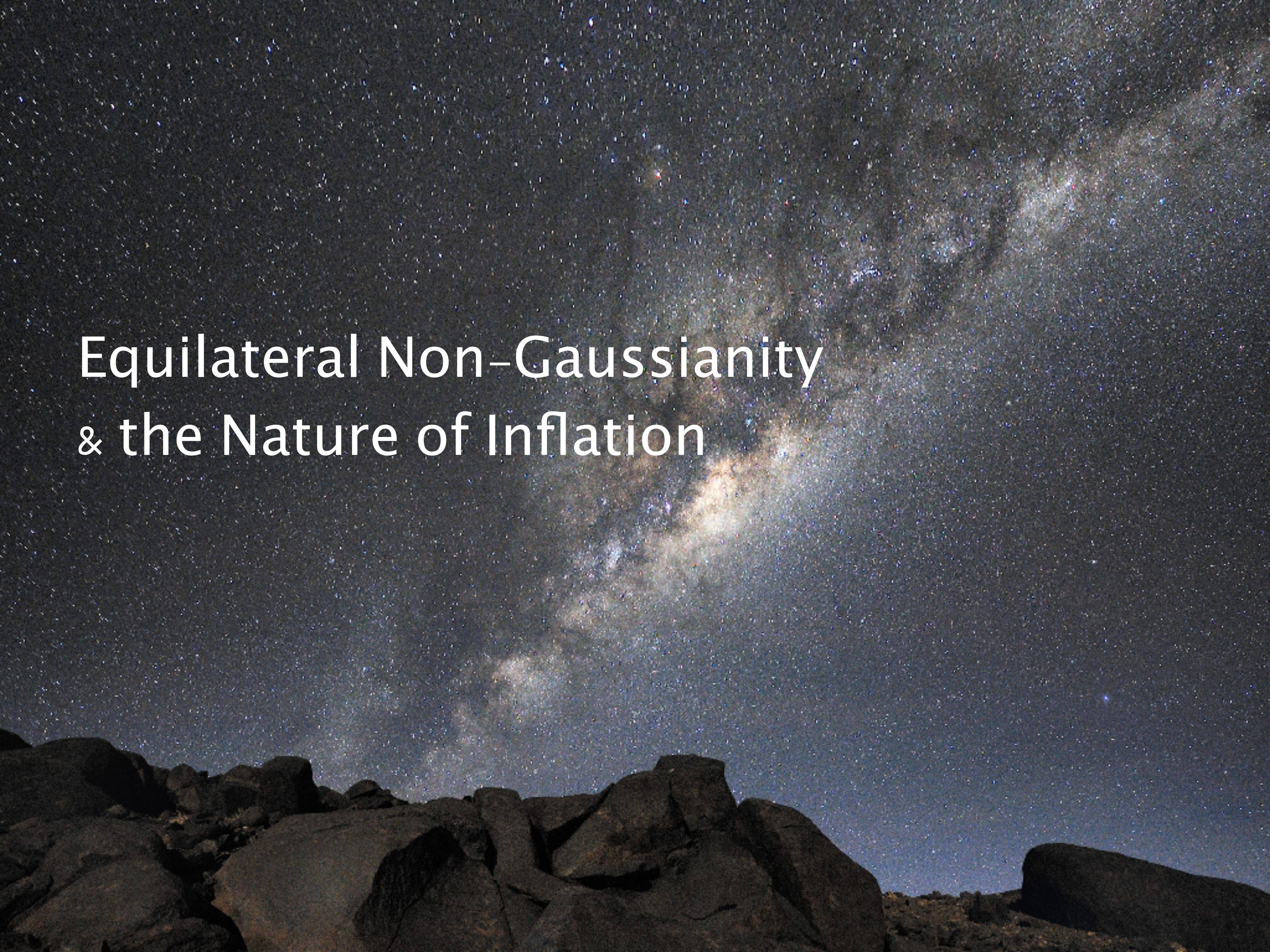
Outline

Equilateral Non-Gaussianity

Challenge of Nonlinearity

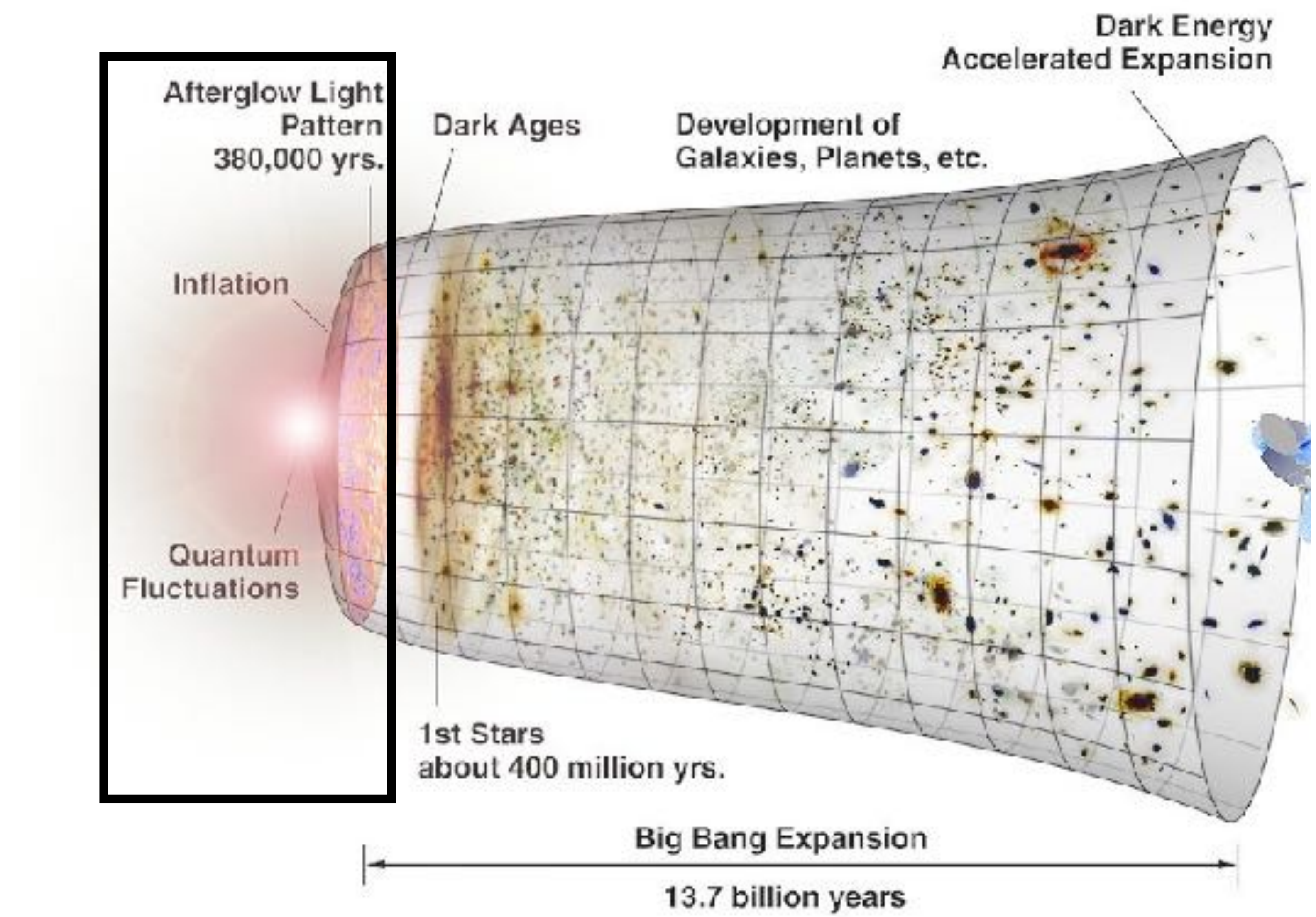
Power of Locality

Future Prospects


Equilateral Non-Gaussianity
& the Nature of Inflation

The Nature of Inflation

The story of inflation is often told in one way

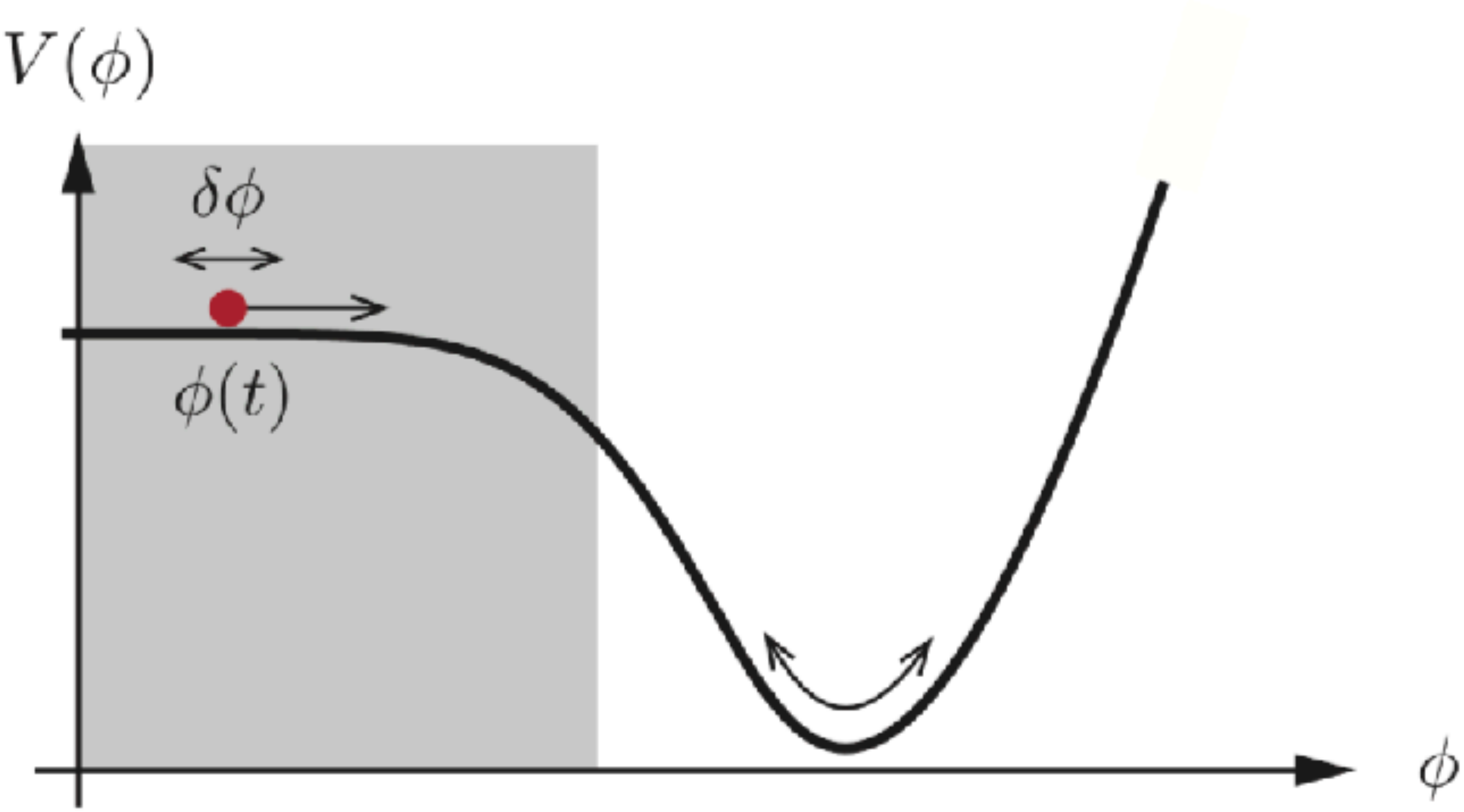


Period of exponential expansion

From WMAP

The Nature of Inflation

The story of inflation is often told in one way

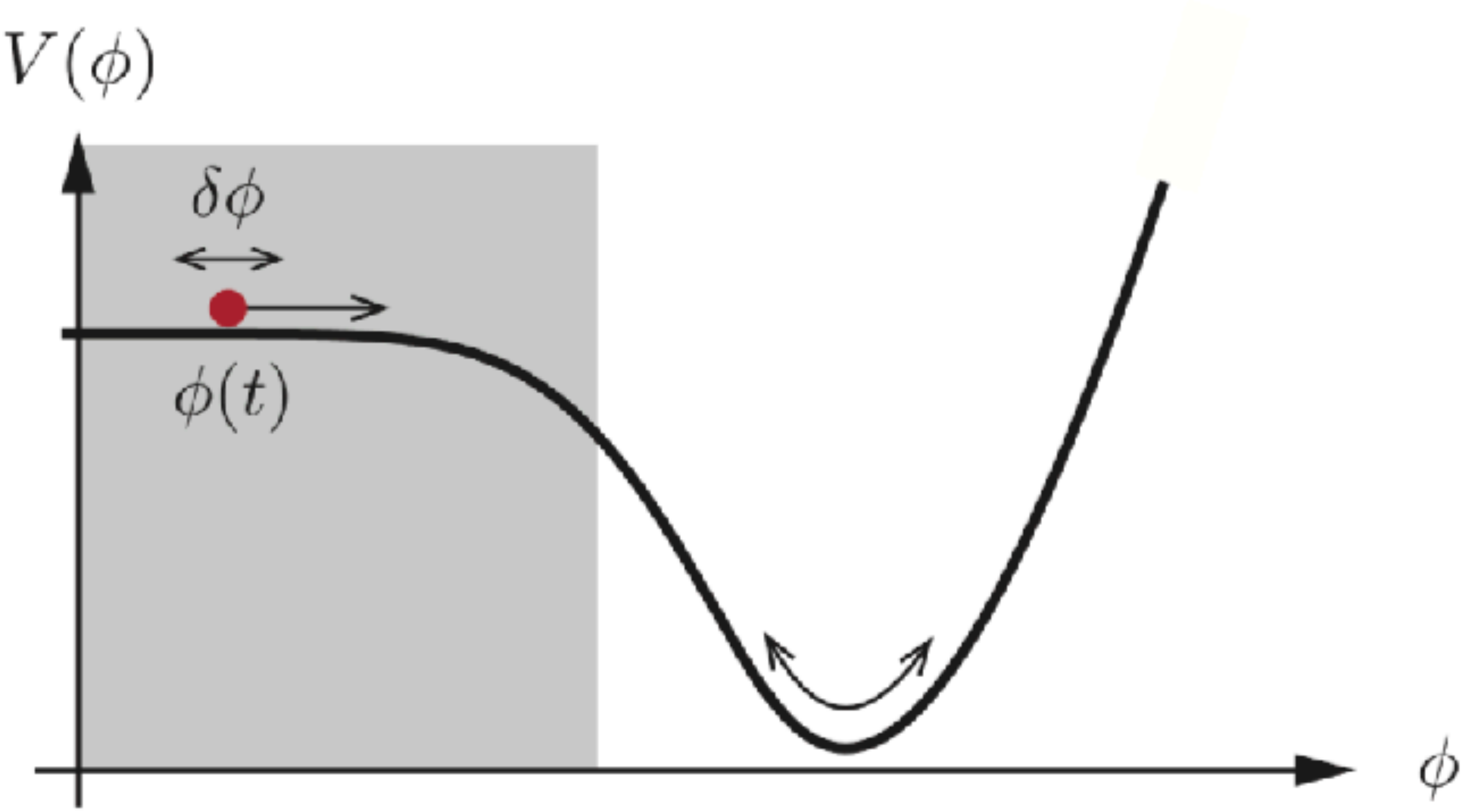


Driven by a slowly rolling scalar field

From Baumann & McAllister

The Nature of Inflation

The story of inflation is often told in one way

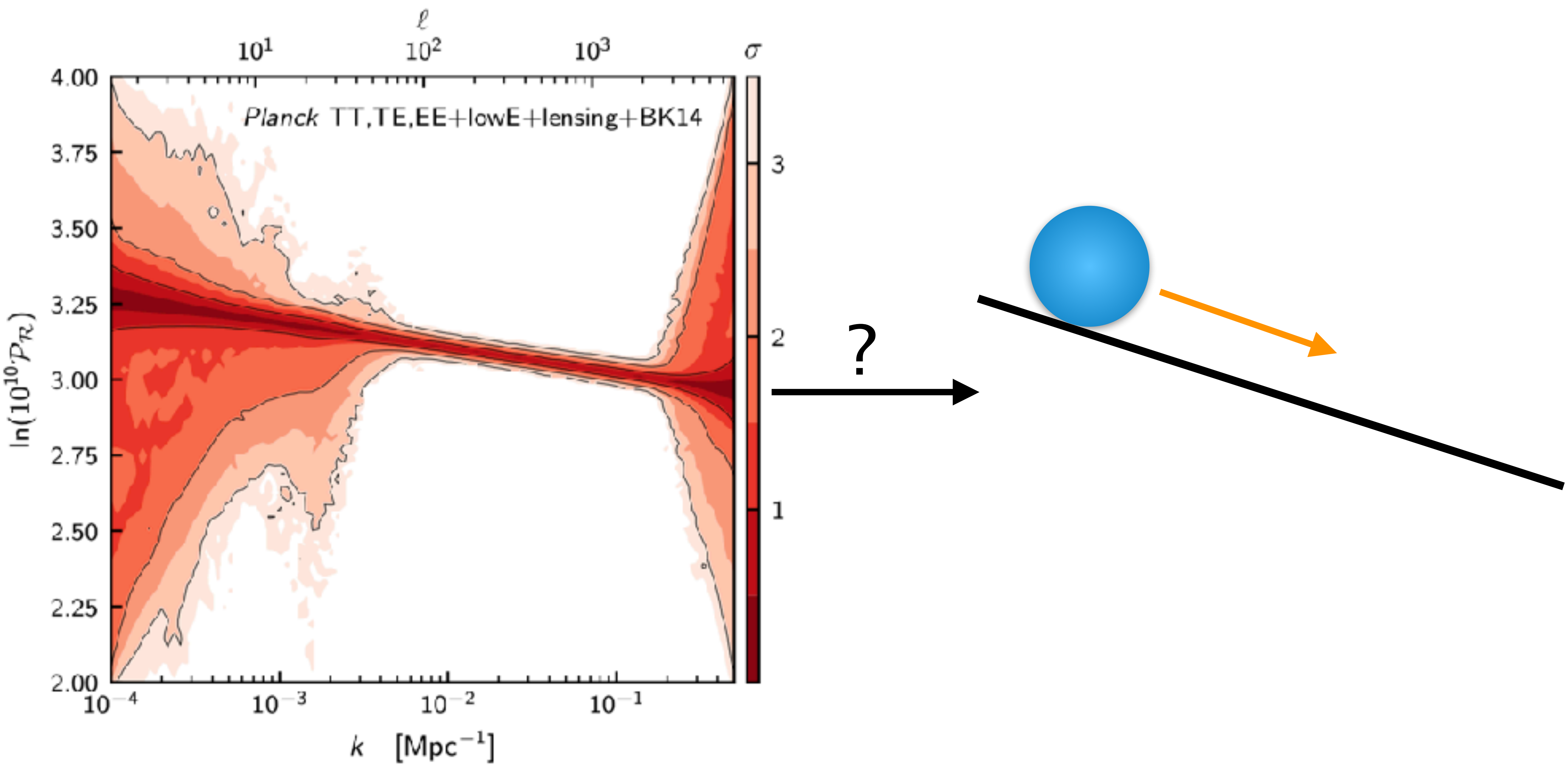


Quantum fluctuations of this field = initial conditions

From Baumann & McAllister

The Nature of Inflation

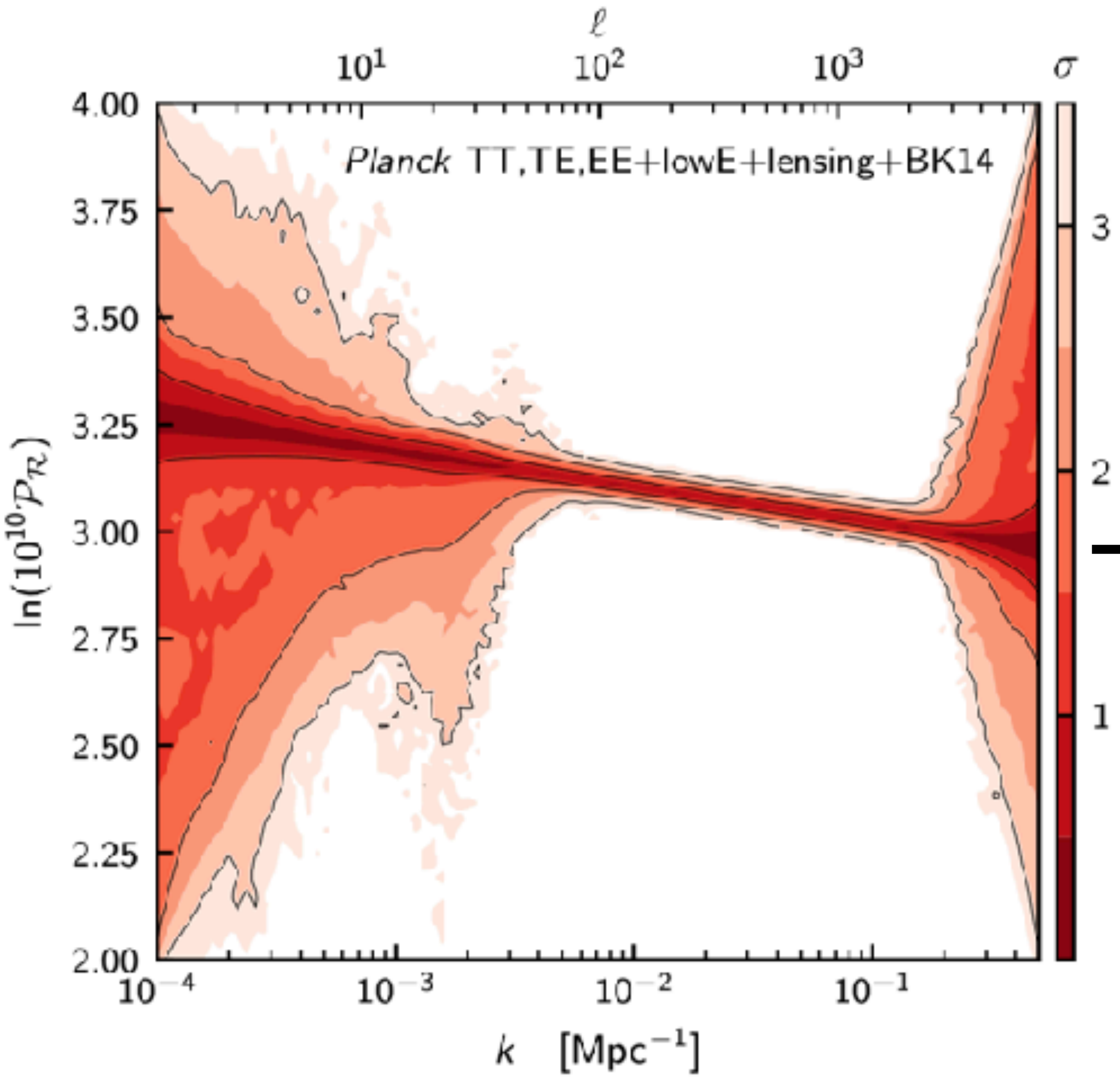
This pictures is consistent with observations



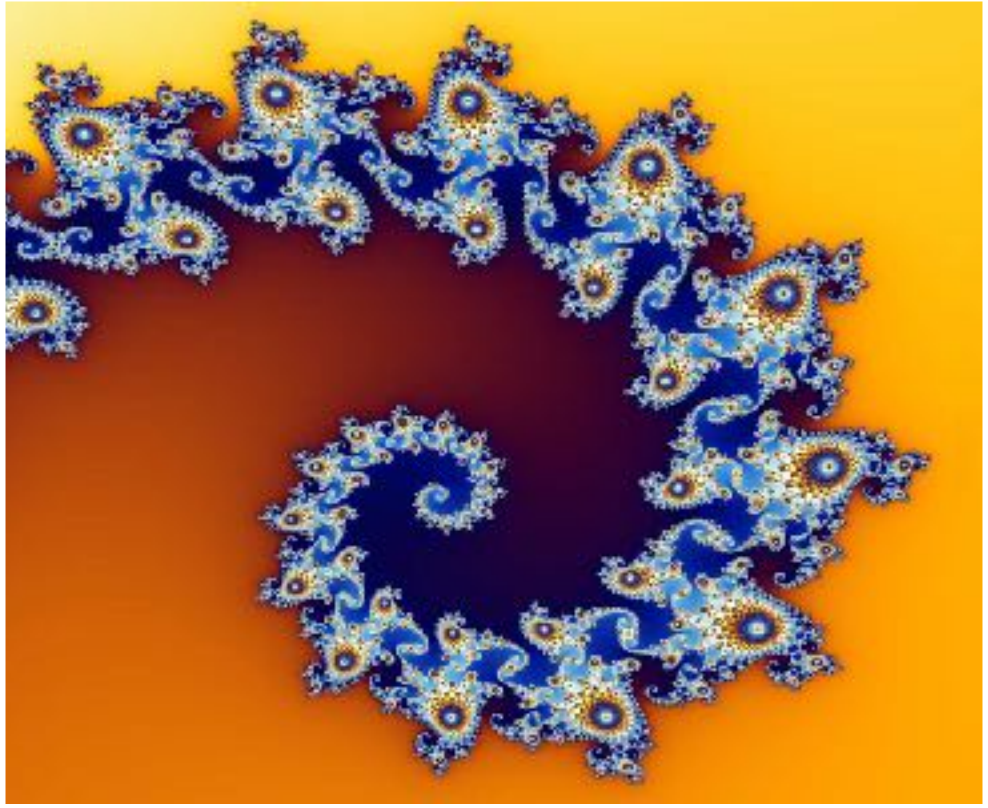
Planck 2018

The Nature of Inflation

But is it necessary ?



?



Planck 2018

The Nature of Inflation

Inflation: A definition

(1) A period of quasi-de Sitter expansion

$$H \equiv \frac{\dot{a}}{a} \qquad \dot{H}(t) \ll H^2 \qquad a(t) \approx e^{Ht}$$

(2) Inflation ends: requires a physical clock

In slow roll inflation – we set our clocks to $\phi(t) \approx \dot{\phi}\, t$

The Nature of Inflation

What is inflation?

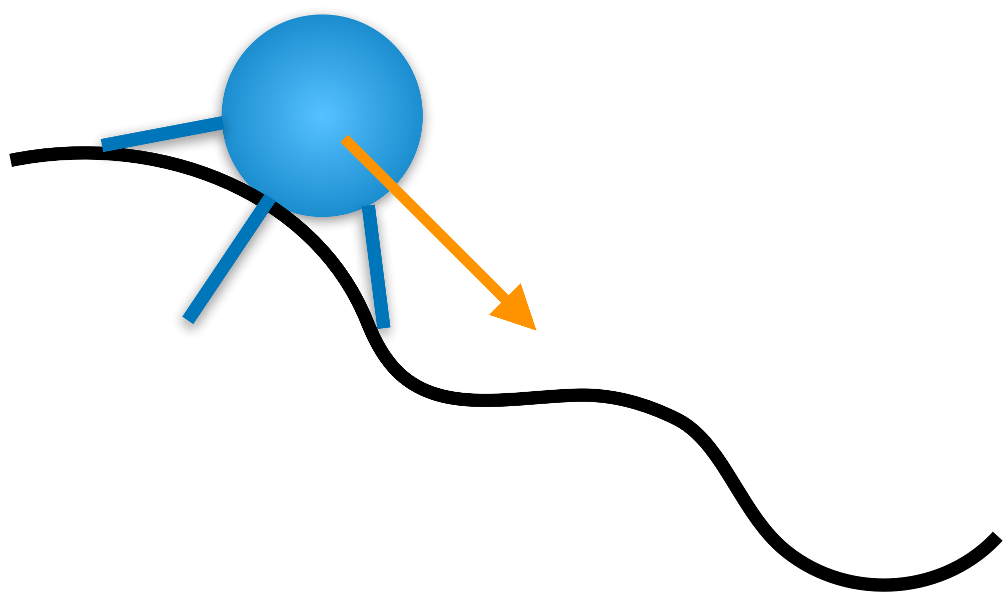
- There are lots of ways to make clock
- Could it be some exotic state of matter?
- Is inflation a weakly or strongly coupled phenomenon?
- How can we tell the different?

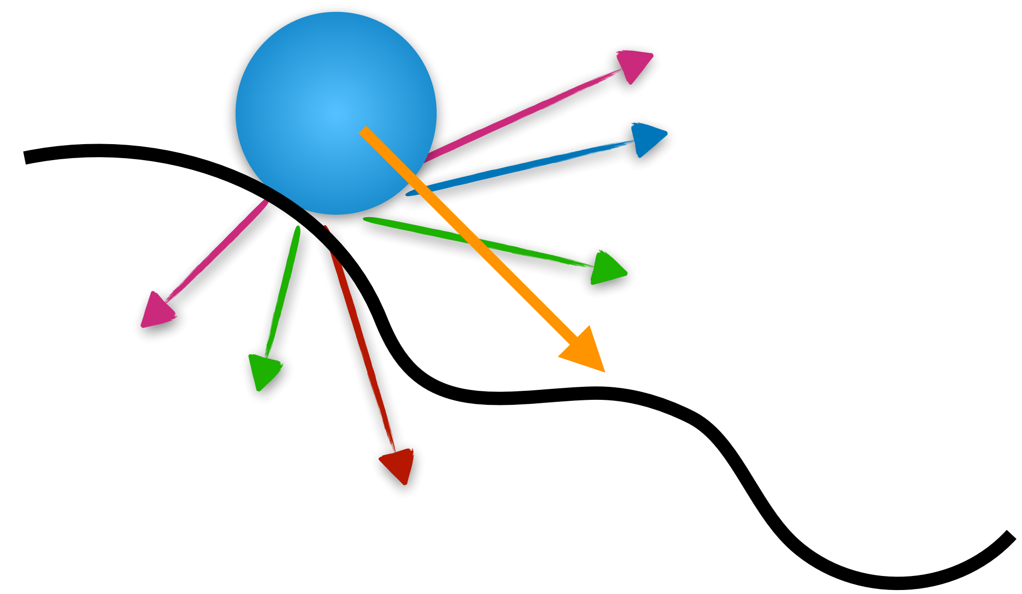
We would like data to answer these kinds of questions

The Nature of Inflation

Might expect dynamics = non-Gaussian

Seen in specific examples





e.g. self-interactions or particle-production

Both lead to large non-Gaussian correlations

The Nature of Inflation

EFT: Inflation is spontaneous time-translation* breaking

Cheung et al. (2007)

$$\langle \mathcal{O} \rangle \propto t \qquad\qquad U \equiv t + \pi$$

The field $\pi(\vec{x}, t)$ are the fluctuations of the clock

We can write the most general possible action

$$S \supset \int d^4x \sqrt{-g}\, F(U, \nabla_\mu)$$

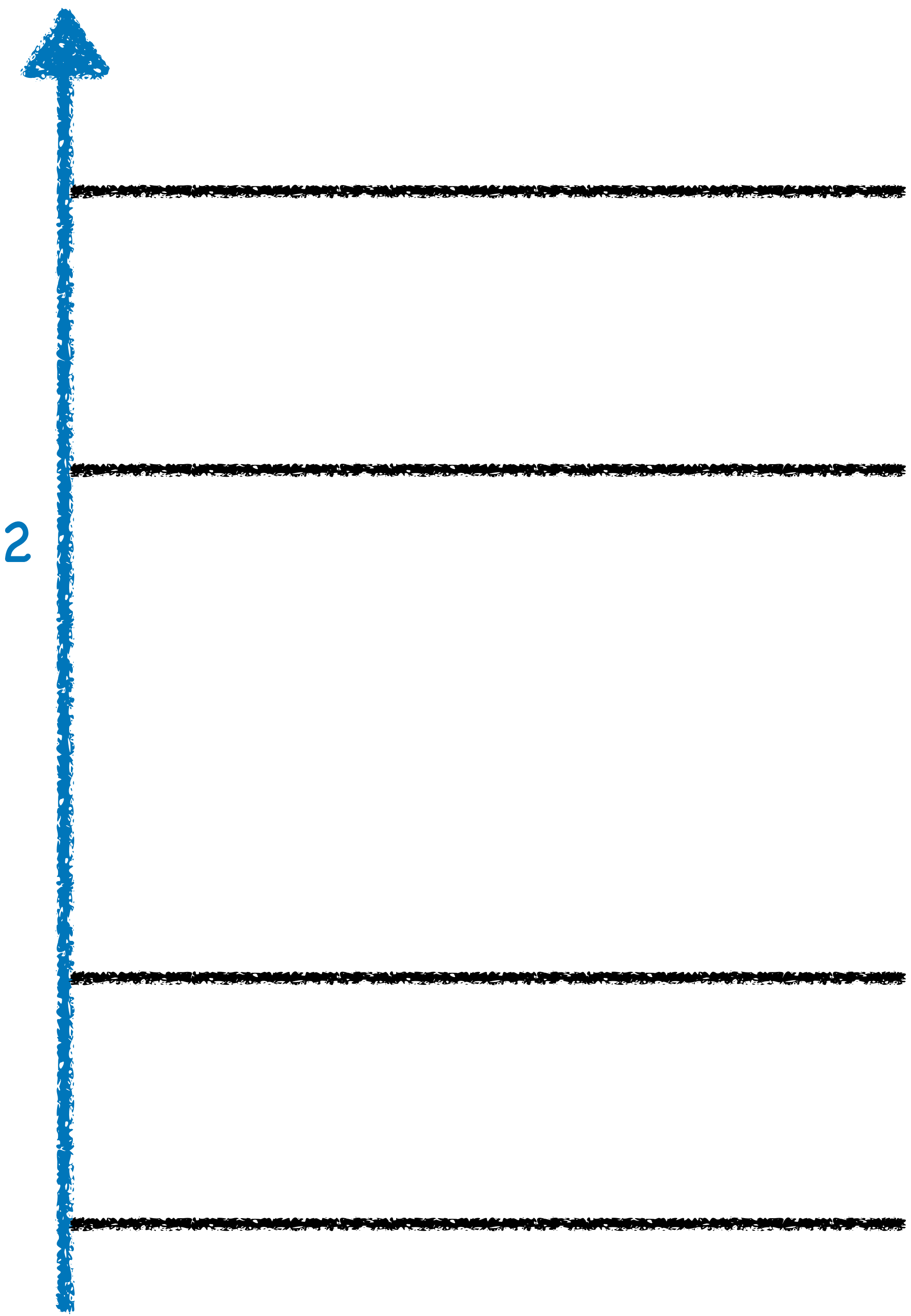
Nothing about this requires a fundamental scalar

*time-translations are eventually gauged / goldstone is eaten by the metric

The Nature of Inflation

What do we know from data?

Energy^2



$M_{\rm pl}^2$ Gravity

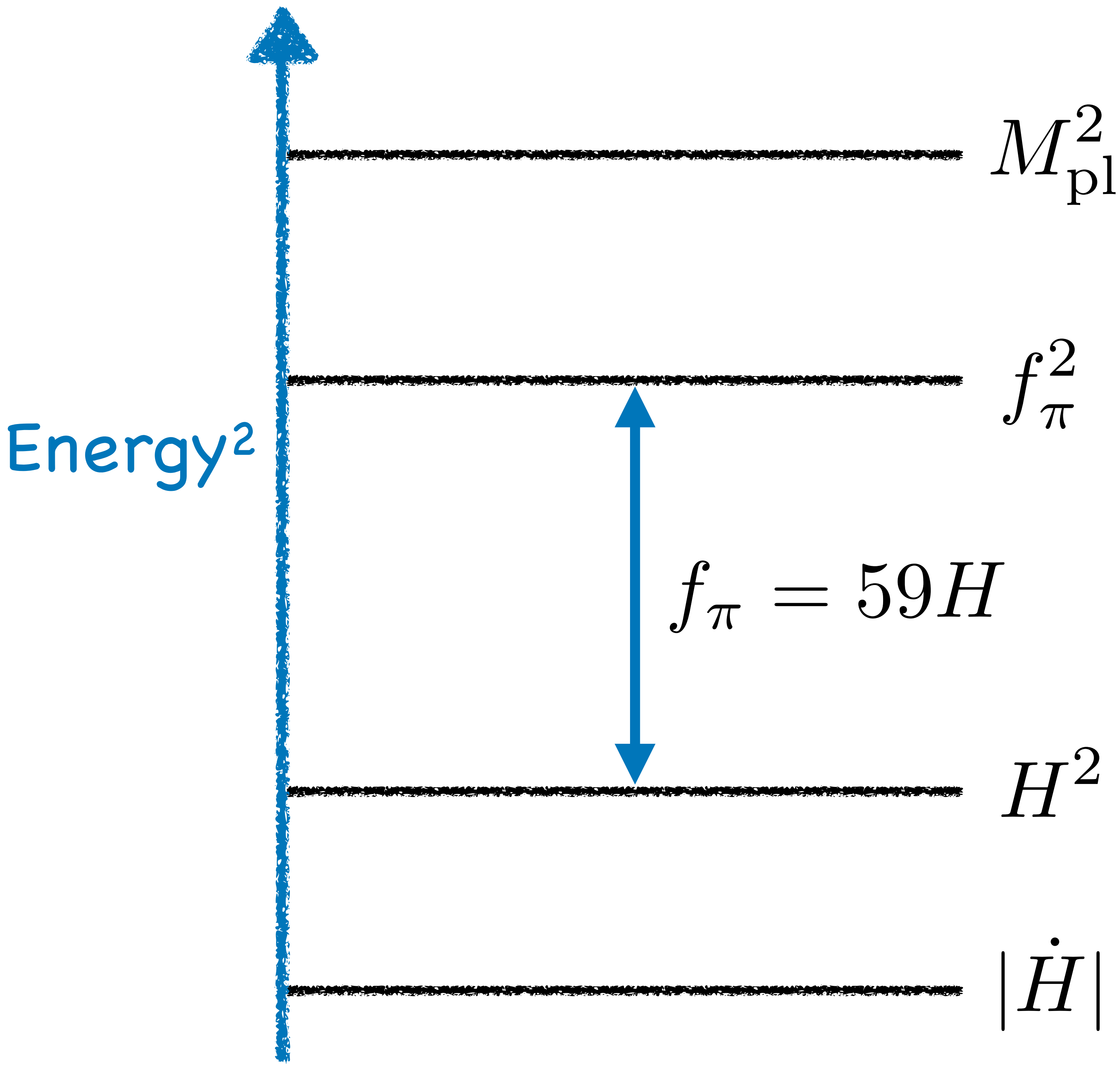
f_π^2 Symmetry breaking: time-translations, boosts

H^2 Freeze-out

$|\dot{H}|$ Explicit breaking of (diagonal) time-translations

The Nature of Inflation

What do we know from data?



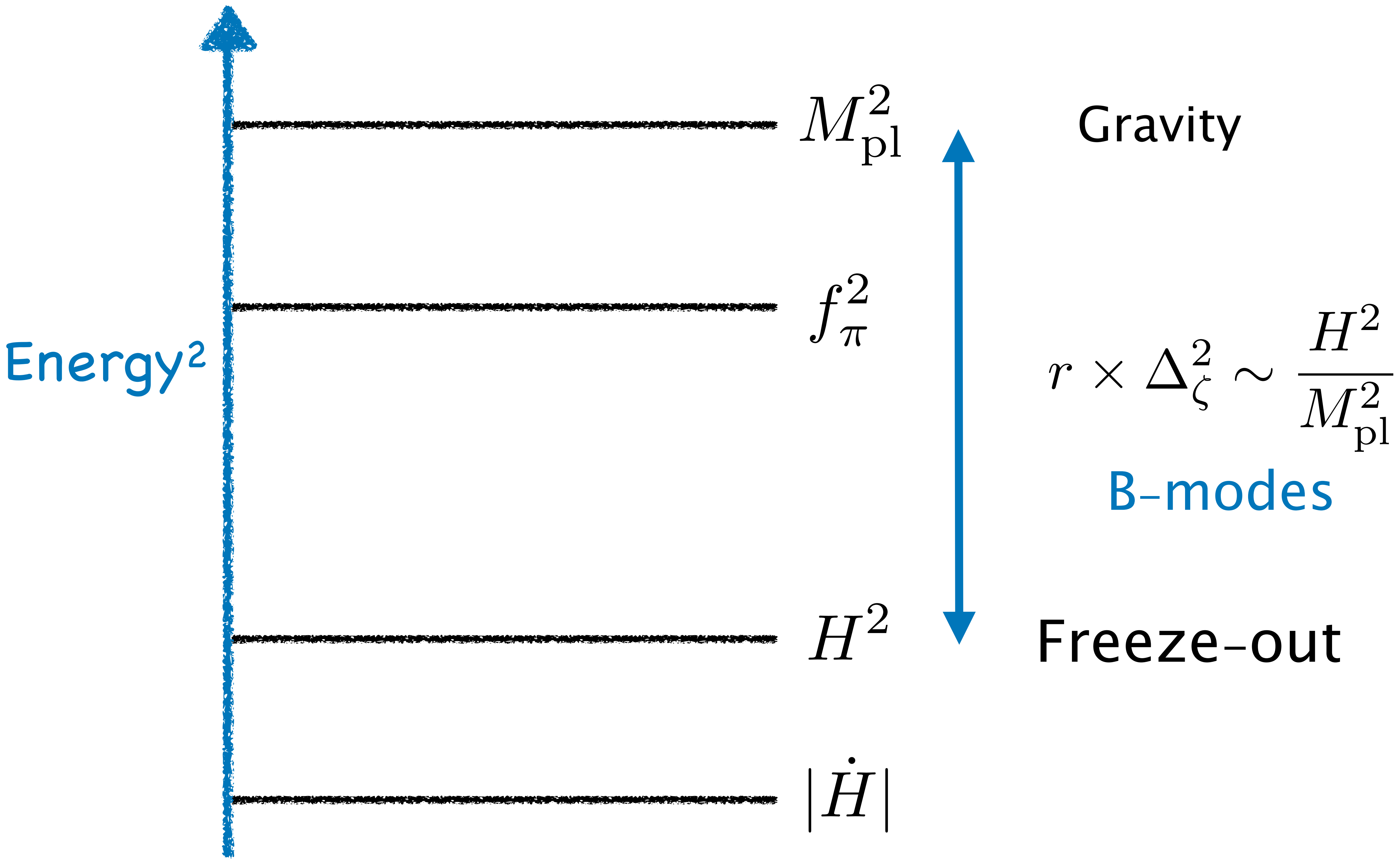
Symmetry breaking

$$2\pi\Delta_\zeta \sim \left(\frac{H}{f_\pi}\right)^2$$

Scale of observations

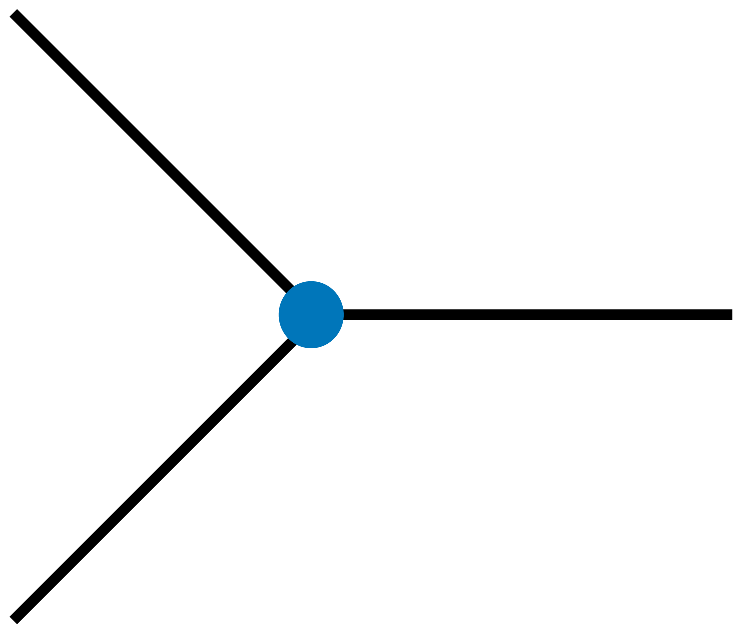
The Nature of Inflation

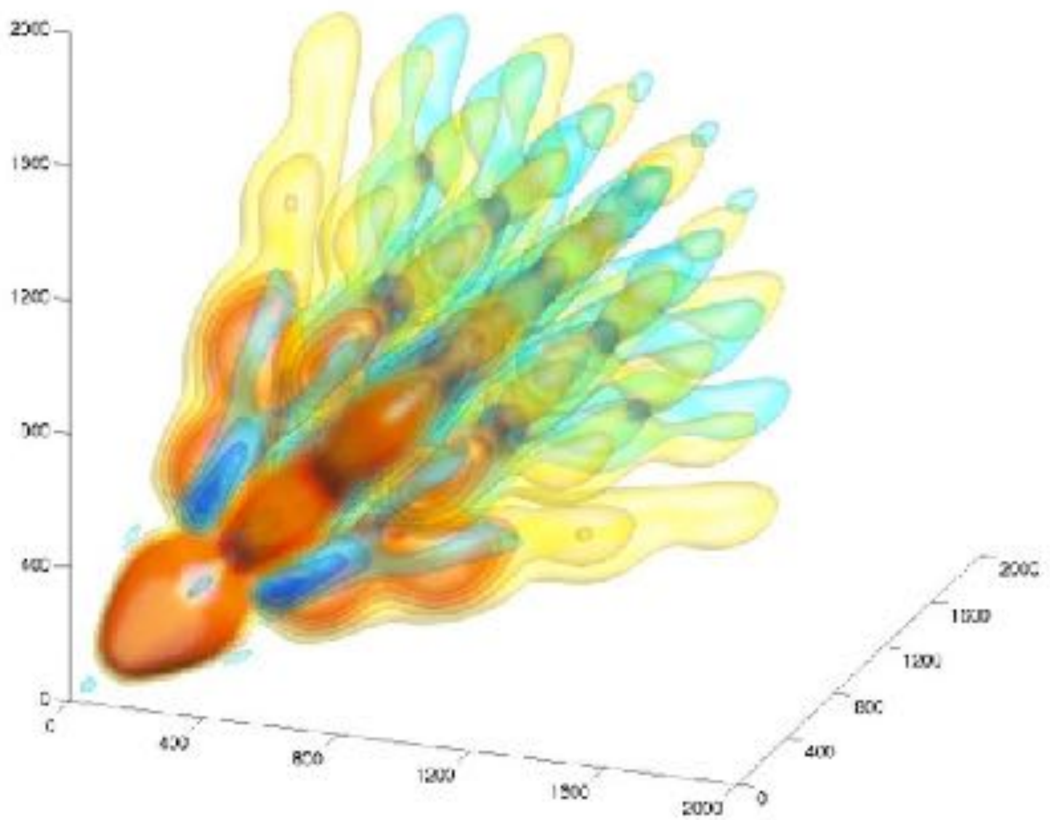
What do we know from data?



The Nature of Inflation

What do we know from data?



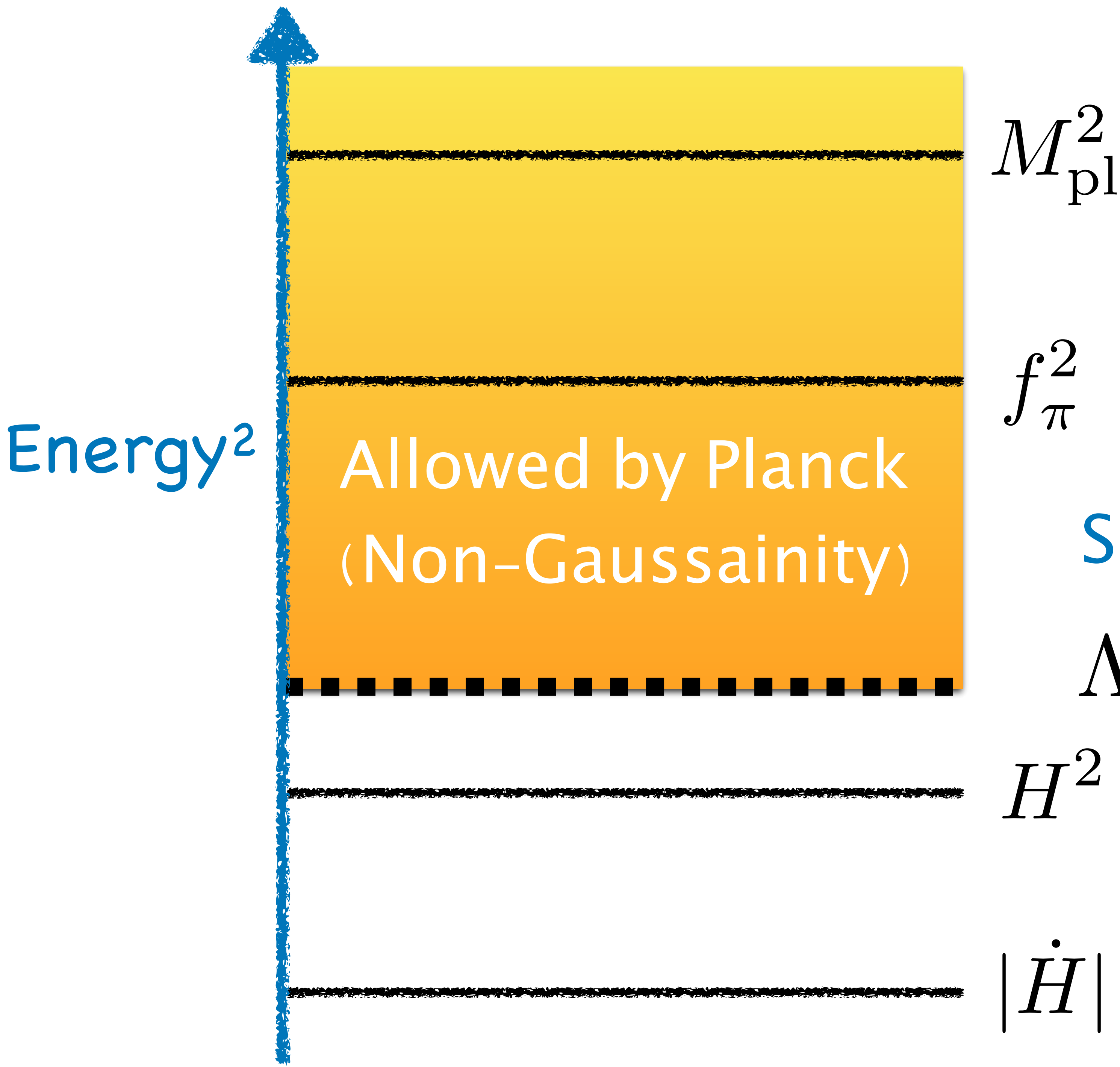


$$\mathcal{L}_{\rm int} \supset \frac{1}{\Lambda^2}\dot{\pi}_c \nabla_\mu \pi \nabla^\mu \pi \qquad \Delta_\zeta^{-1}\frac{H^2}{\Lambda^2} \approx f_{\rm NL}^{\rm eq} = -26 \pm 94\,(95\%)$$

Courtesy of Fergusson & Shellard

The Nature of Inflation

What do we know from data?



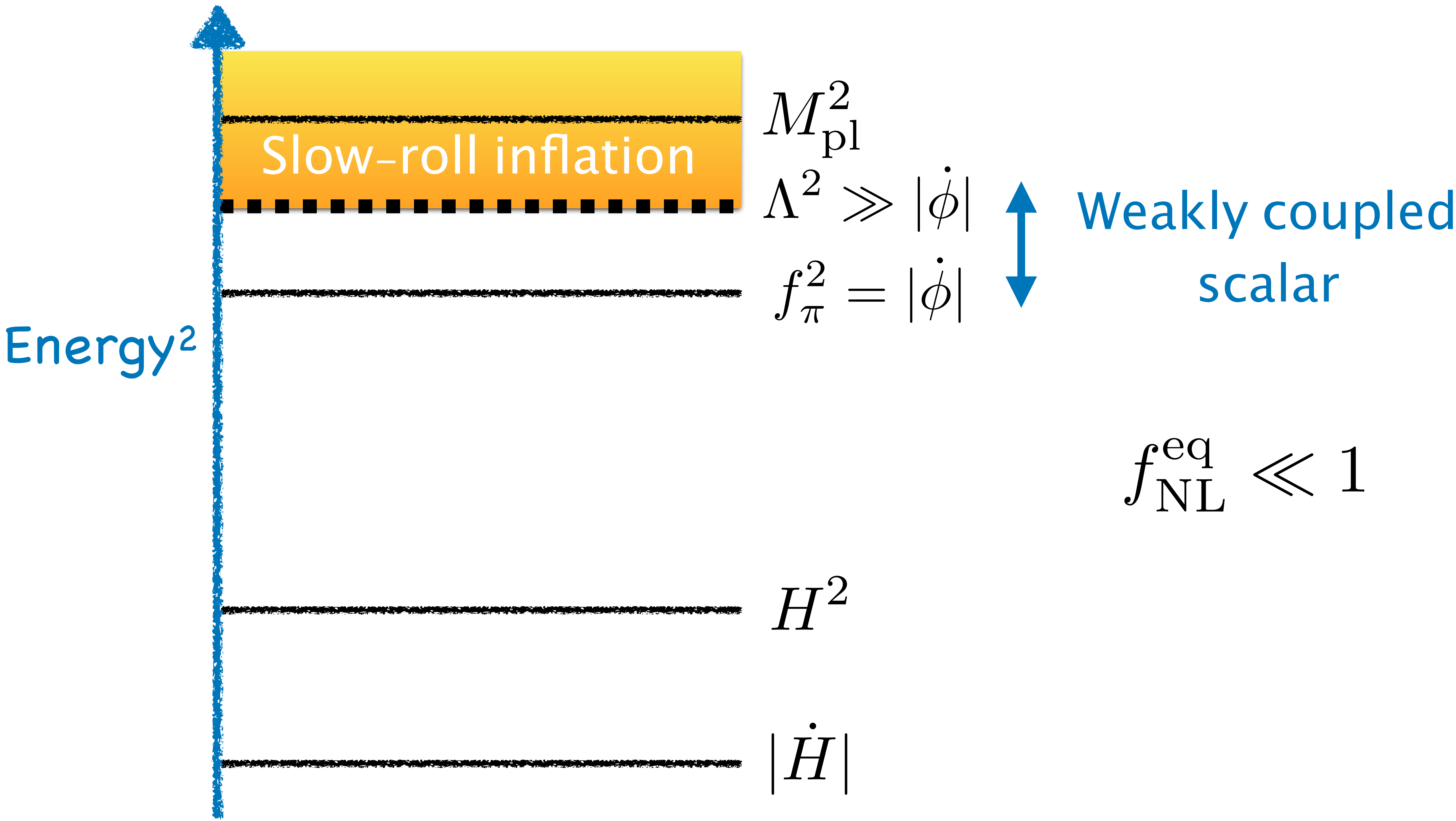
Scale of Cubic-Interactions

$$\Lambda \geq (5-10)\, H \quad (95\%)$$

The Nature of Inflation

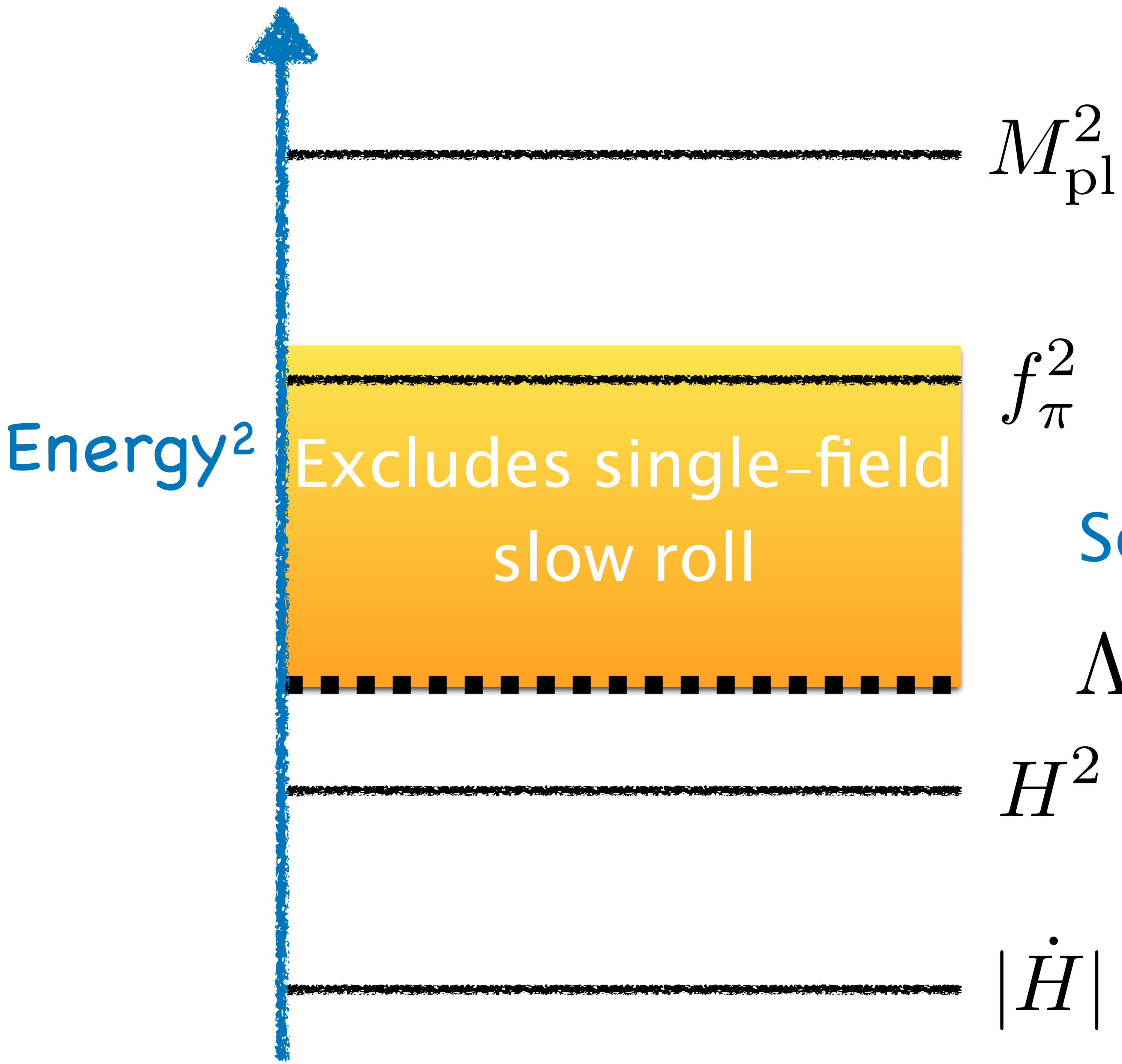
Slow-roll inflation predicts:

Creminelli (2003)



The Nature of Inflation

What do we know from data?

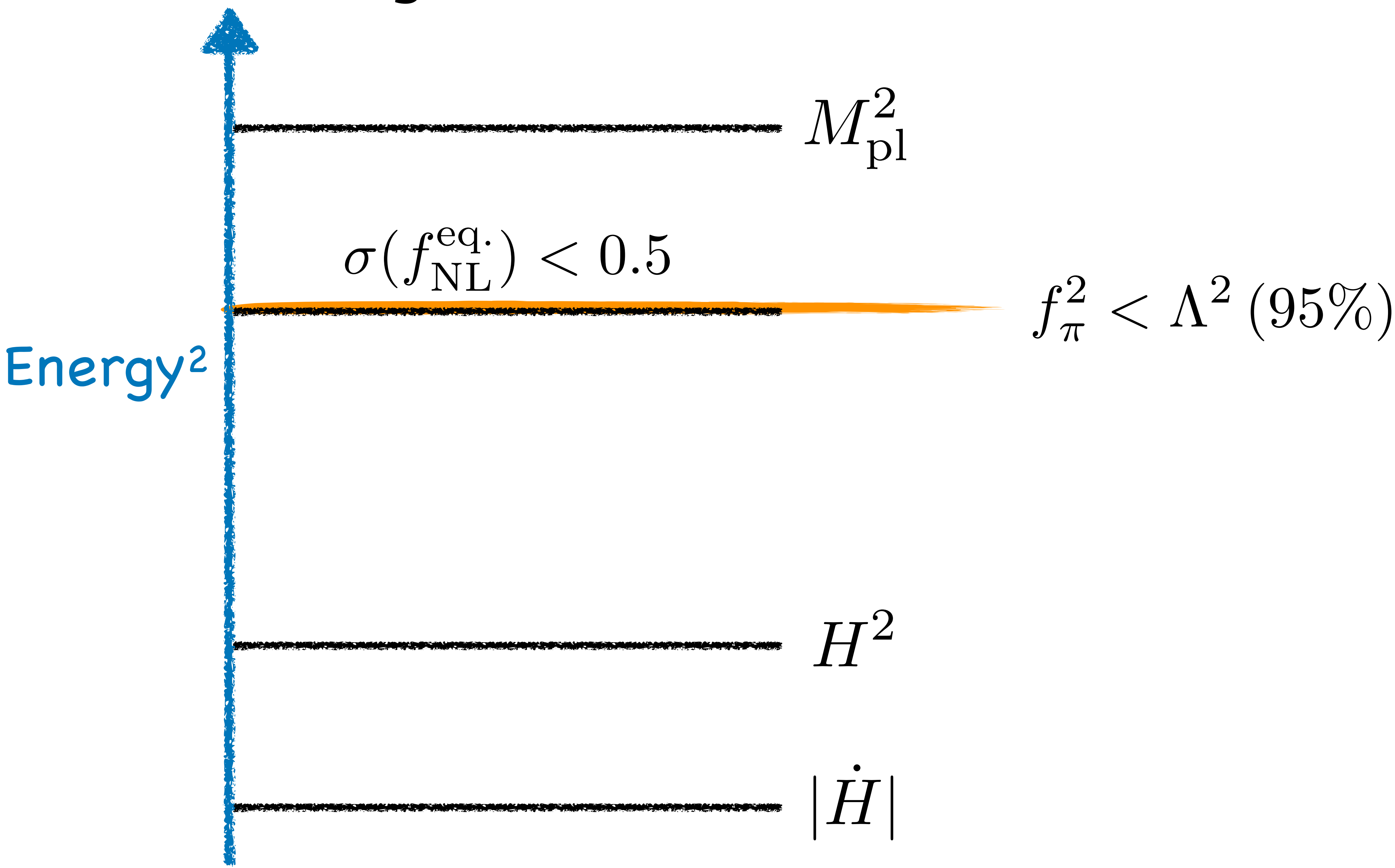


Scale of Cubic-Interactions

$$\Lambda \geq (5-10)\, H \quad (95\%)$$

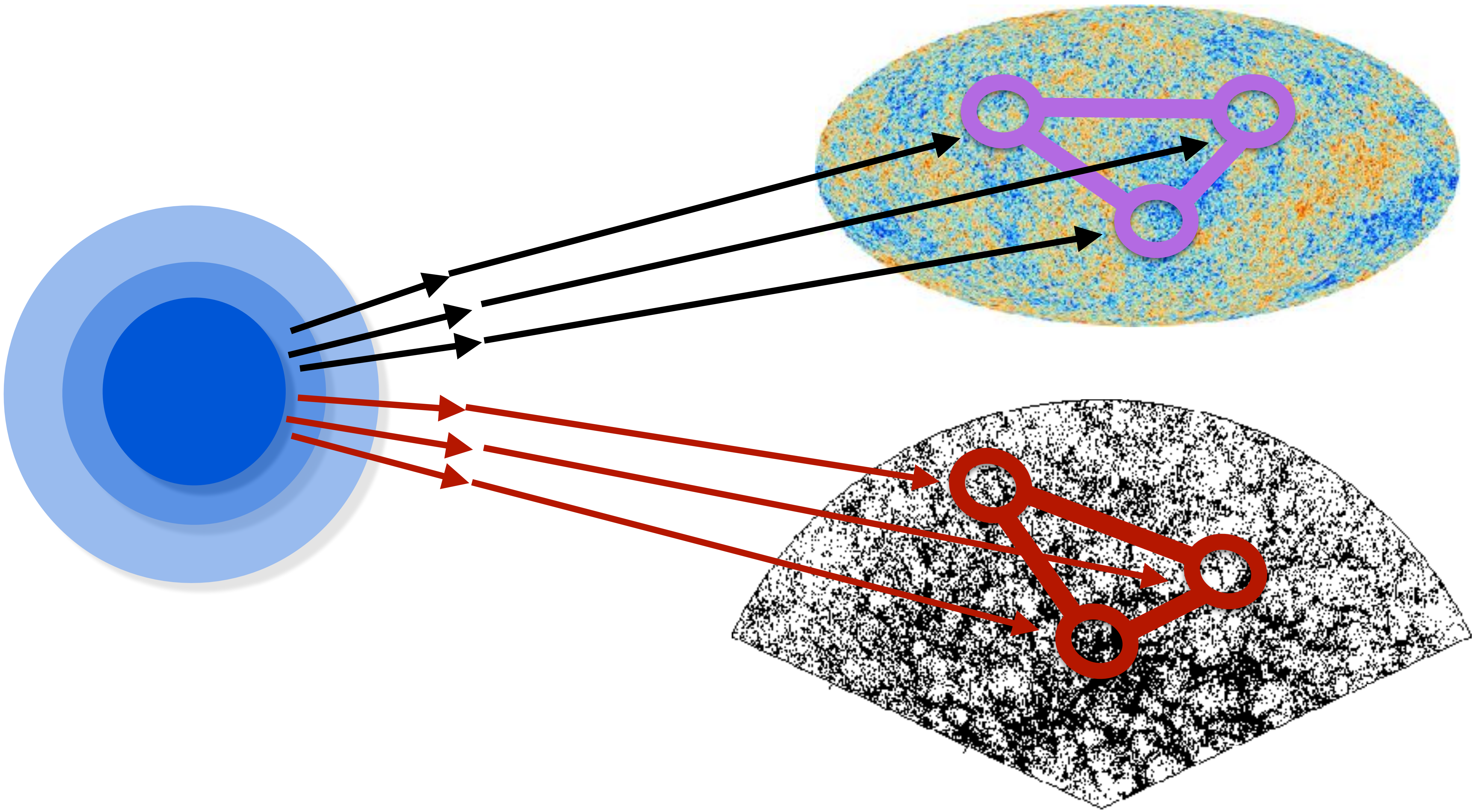
Theoretical Target

Observational goal: reach the slow-roll threshold




The Challenge

Primordial Non-Gaussianity

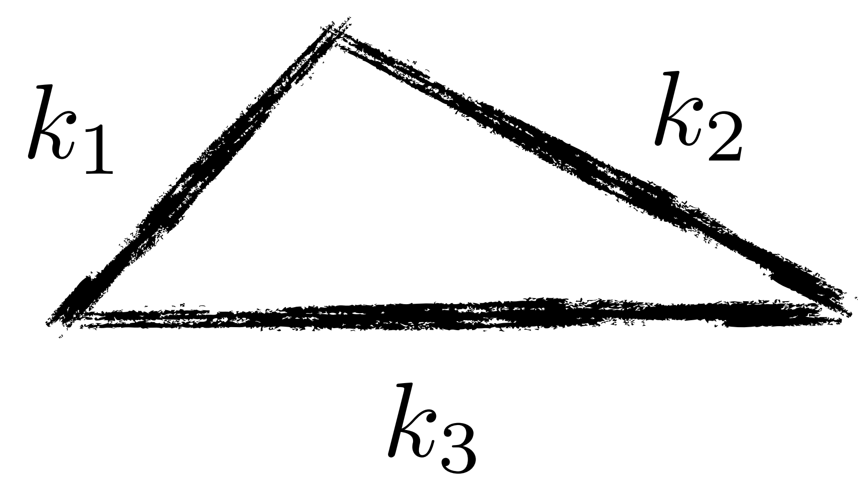


Primordial Non-Gaussianity

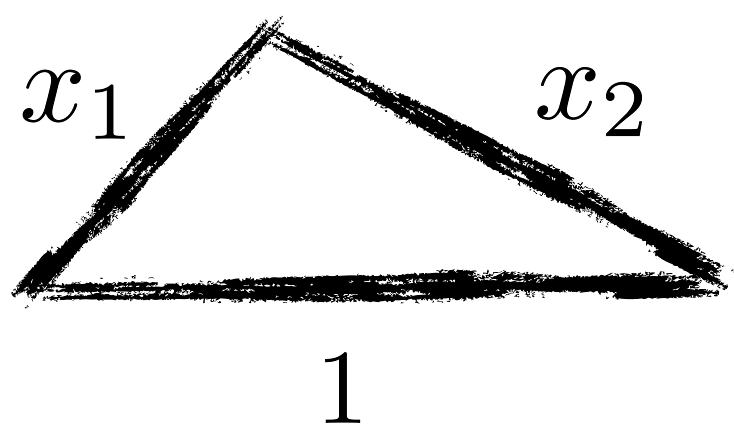
On general grounds, bispectra take the form

$$\langle \Phi(\vec{k}_1)\Phi(\vec{k}_2)\Phi(\vec{k}_3)\rangle = f_{\rm NL} B_\Phi(k_1, k_2, k_3)(2\pi)^3\delta(\vec{k}_1+\vec{k}_2+\vec{k}_3)$$

Momentum conservation:



Scale invariance:



$$\times \frac{1}{k_3^6}$$

Primordial Non-Gaussianity

Defined by amplitude and "shape"

$$f_{\rm NL} B_\Phi(k_1, k_2, k_3) = f_{\rm NL} \frac{6\,\Delta_\Phi^4}{k_3^6 x_1^2 x_2^2} S(x_1, x_2)$$

The shapes live in a basis of orthogonal functions

$$\int dx_1 dx_2 S_1(x_1, x_2) S_2(x_1, x_2) = S_1 \cdot S_2 = \cos_{12}$$

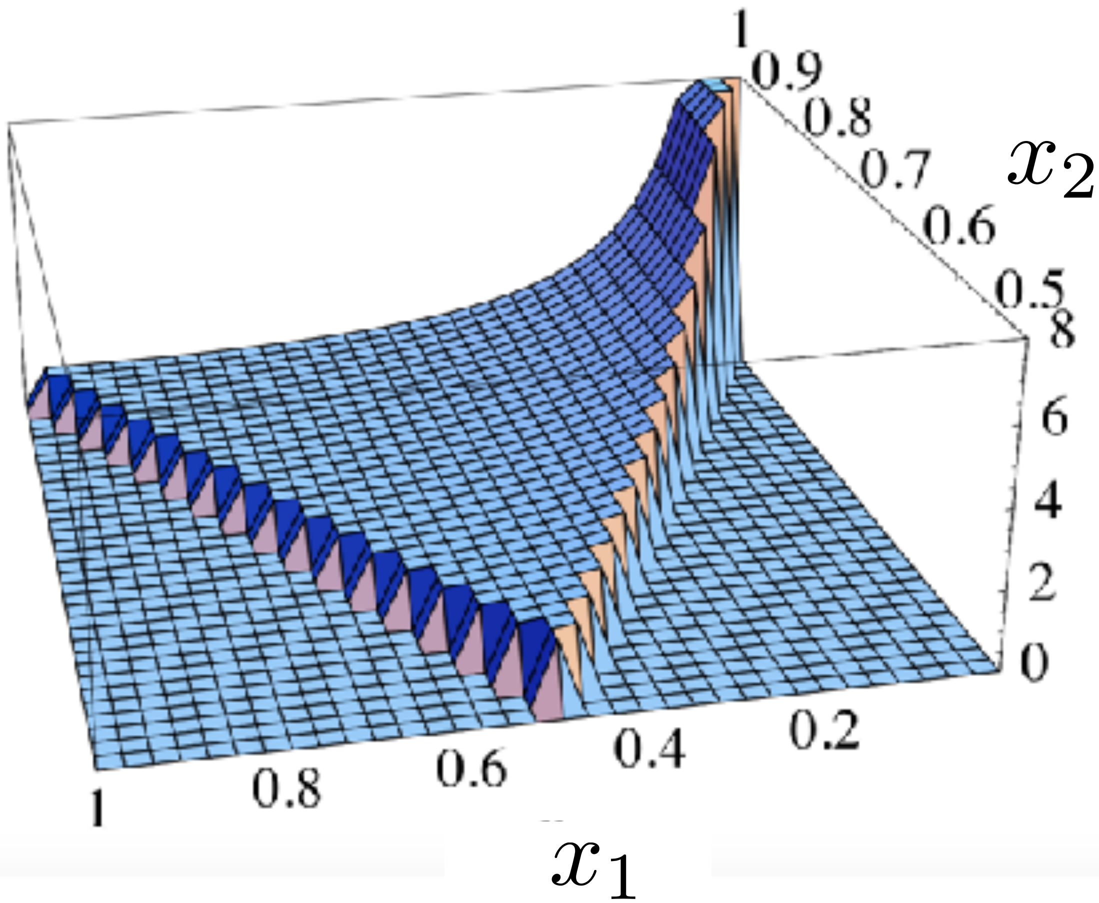
Cosine is how easily they are distinguish (in 3pt)

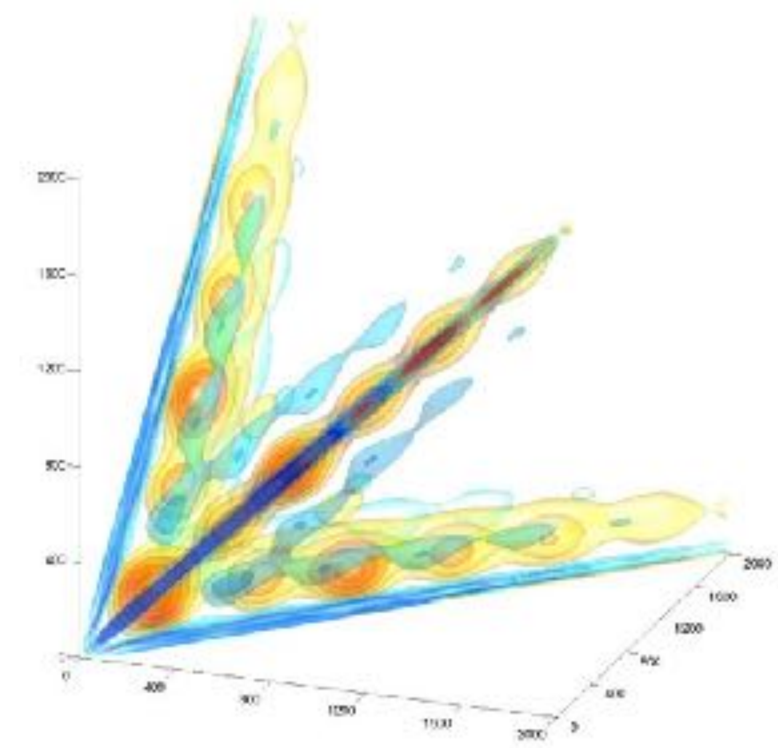
Babich, Creminelli, Zaldarriaga (2004)

The "Local Shape"

$$f_{\rm NL}^{\rm local} = -0.9 \pm 5.1$$

Planck 2018





Courtesy of Fergusson & Shellard

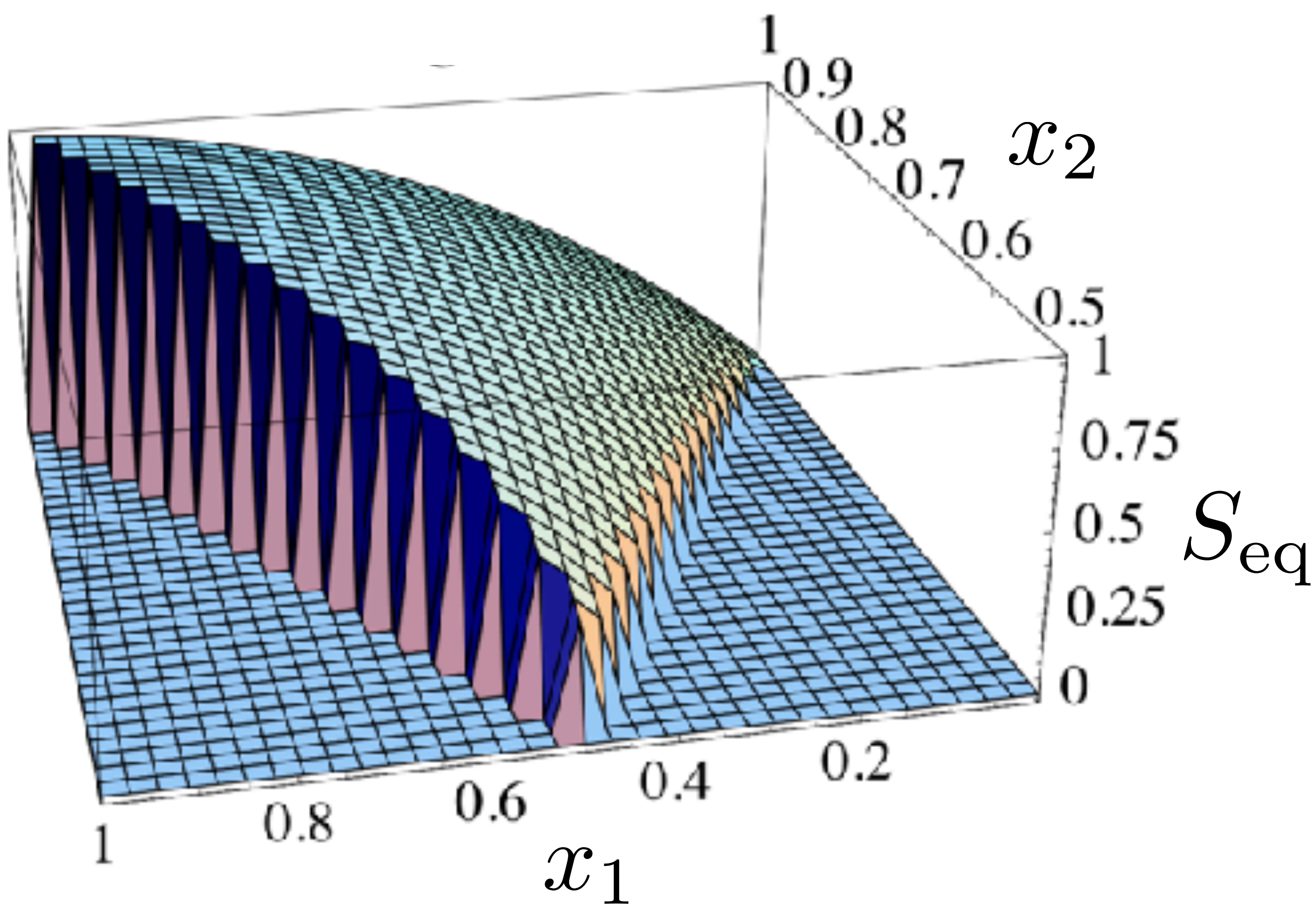
$$\Phi(\vec{x}) = \phi_g(\vec{x}) - f_{\rm NL}^{\rm local}\phi_g^2(\vec{x}) + \dots$$

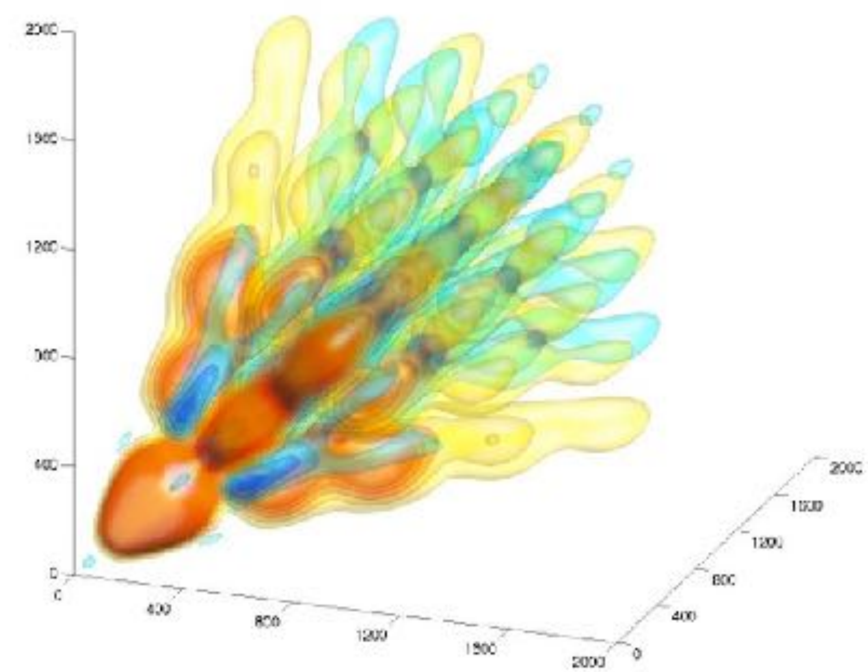
Current Limits

The "Equilateral Shape"

$$f_{\rm NL}^{\rm equil} = -26 \pm 47$$

Planck 2018



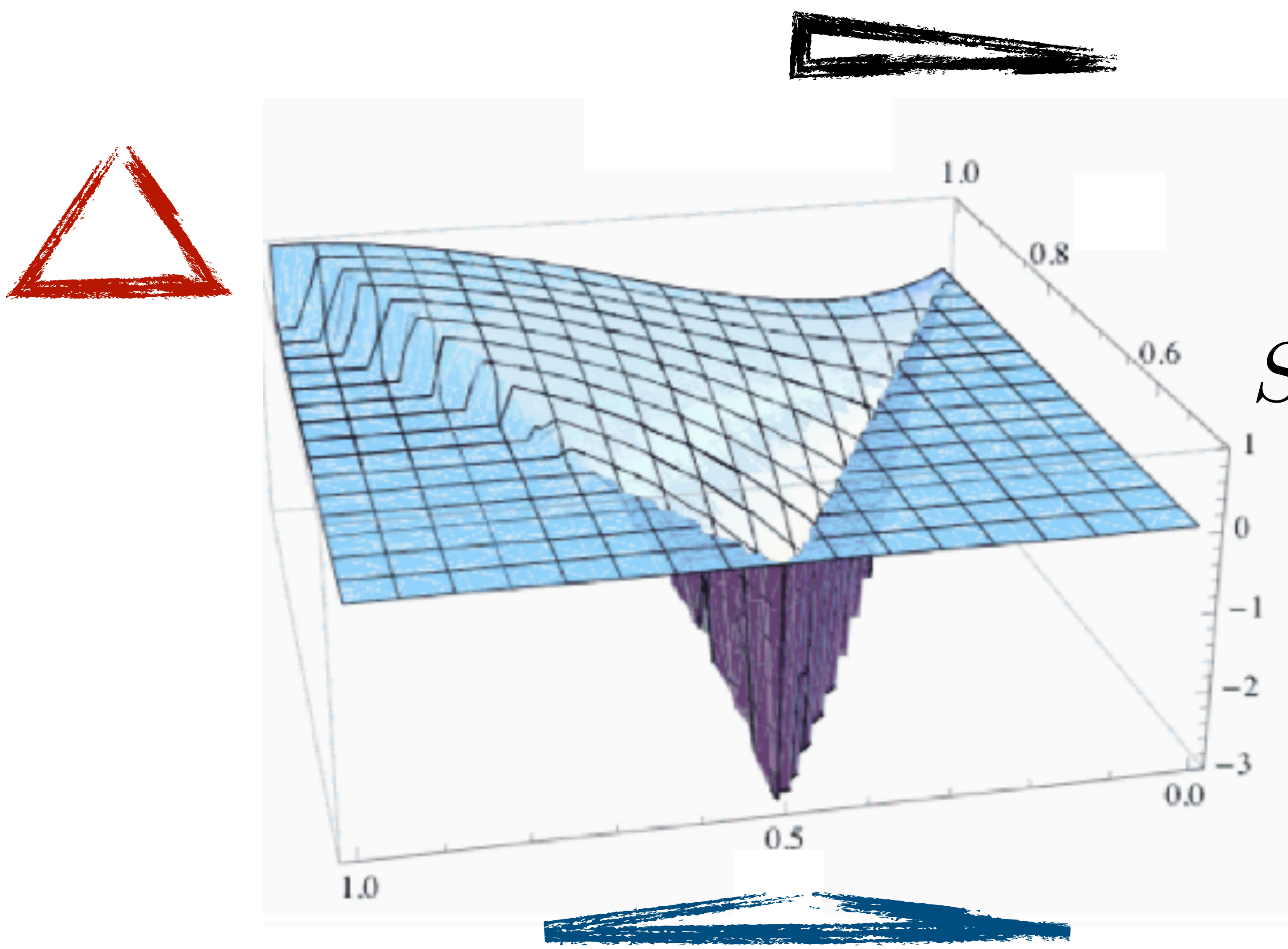


Courtesy of Fergusson & Shellard

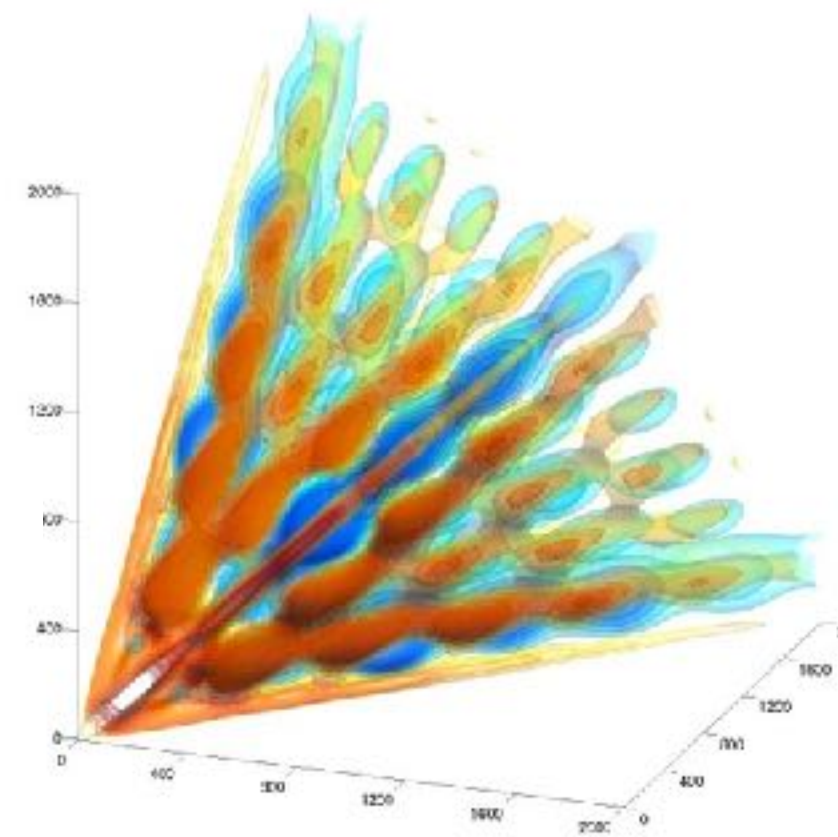
The "Orthogonal Shape"

$$f_{\mathrm{NL}}^{\mathrm{ortho}} = -38 \pm 24$$

Planck 2018



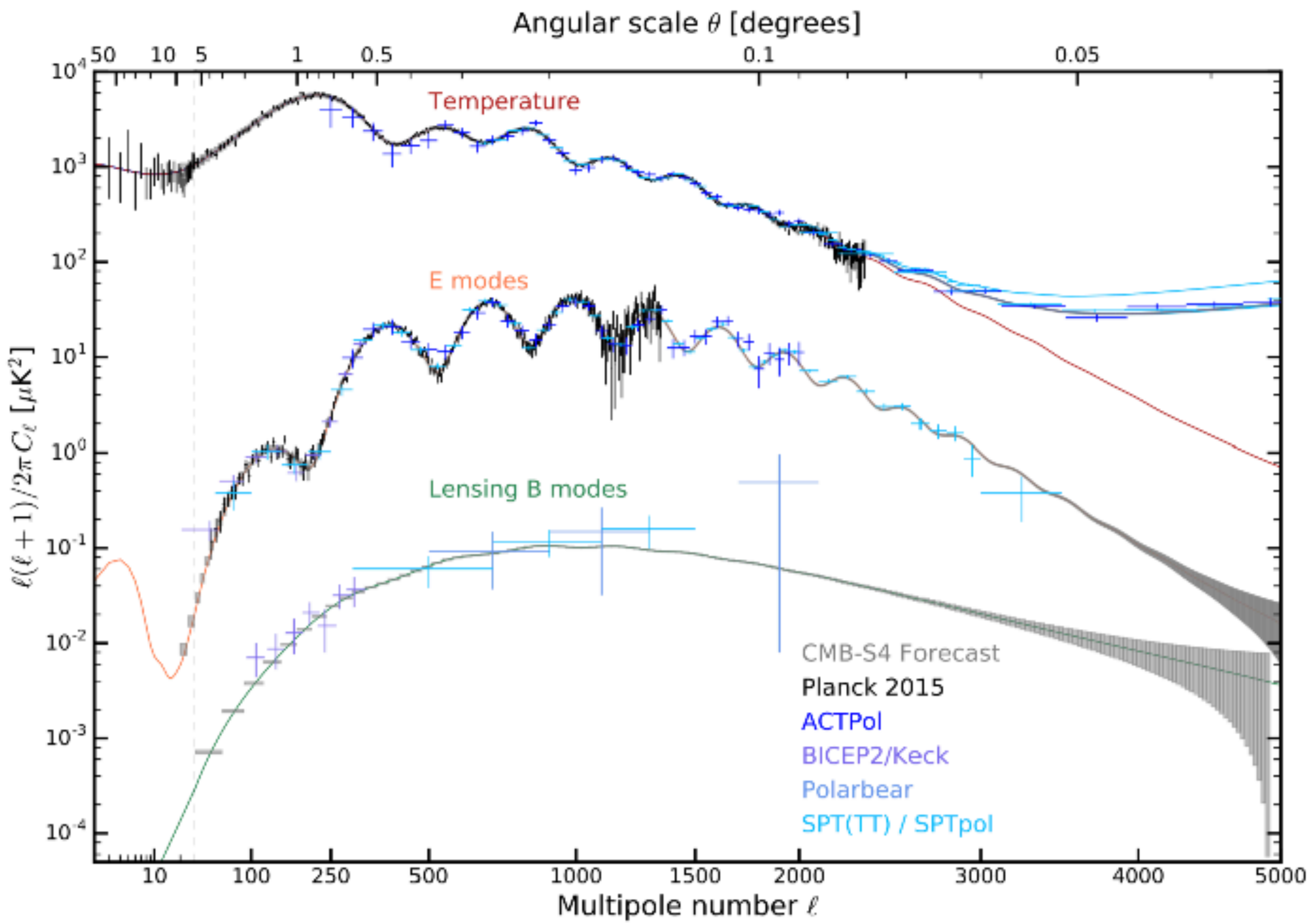
S_{ortho}



Courtesy of Fergusson & Shellard

Cosmic Microwave Background

Not enough modes left in the CMB



Cosmic Microwave Background

Not enough modes left in the CMB

Type	*Planck* actual (forecast)	CMB-S4	CMB-S4 + low-ℓ *Planck*
Local	$\sigma(f_{\rm NL}) = 5\,(4.5)$	$\sigma(f_{\rm NL}) = 2.6$	$\sigma(f_{\rm NL}) = 1.8$
Equilateral	$\sigma(f_{\rm NL}) = 43\,(45.2)$	$\sigma(f_{\rm NL}) = 21.2$	$\sigma(f_{\rm NL}) = 21.2$
Orthogonal	$\sigma(f_{\rm NL}) = 21\,(21.9)$	$\sigma(f_{\rm NL}) = 9.2$	$\sigma(f_{\rm NL}) = 9.1$

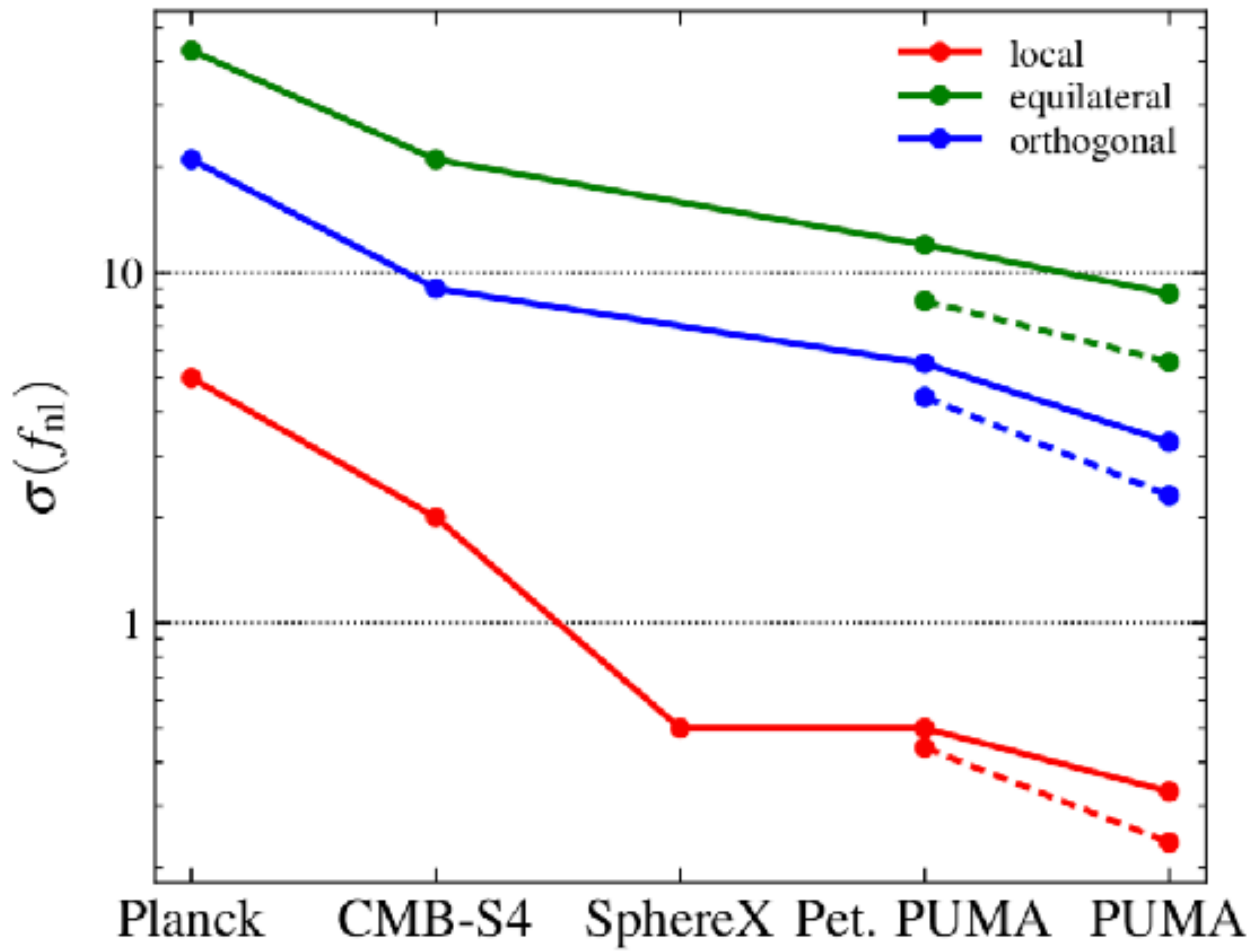
Naive mode counting tells us that $\sigma(f_{\rm NL}^{\rm eq}) \propto \ell_{\rm max}^{-1}$

In detail, we only get the scaling $\sigma(f_{\rm NL}^{\rm eq}) \propto \ell_{\rm max}^{-0.55}$

Kalaja et al. (2020)

Lose information from projection from 3d to 2d

Improvements expected for local non-Gaussianity



PUMA collaboration (2019)

Large Scale Structure

Local NG benefits from scale-dependent bias

$$\delta_g(\vec{k}) \propto f_{\rm NL}^{\rm eq} \Phi(\vec{k}) \propto \frac{f_{\rm NL}}{k^2}\delta(\vec{k})$$

Dalal et al. (2007)

This is apparent violation of equivalence principle

This can only arise in multi-field inflation

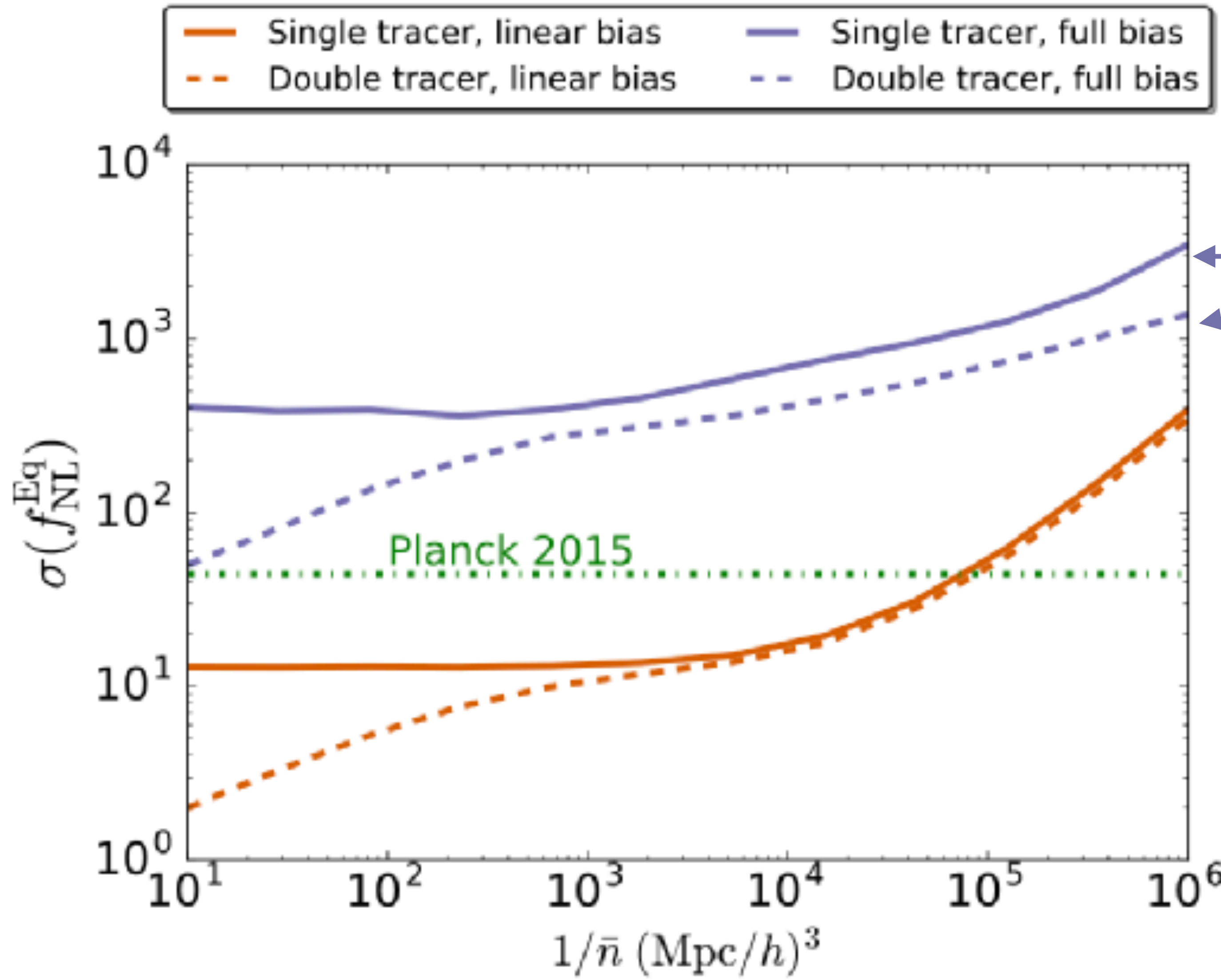
Creminelli & Zaldarriaga (2004)

Other shapes also introduce scale dependence

$$\delta_g(\vec{k}) \approx \frac{1}{k^\alpha}\delta(\vec{k})$$

Large Scale Structure

Not competitive for equilateral



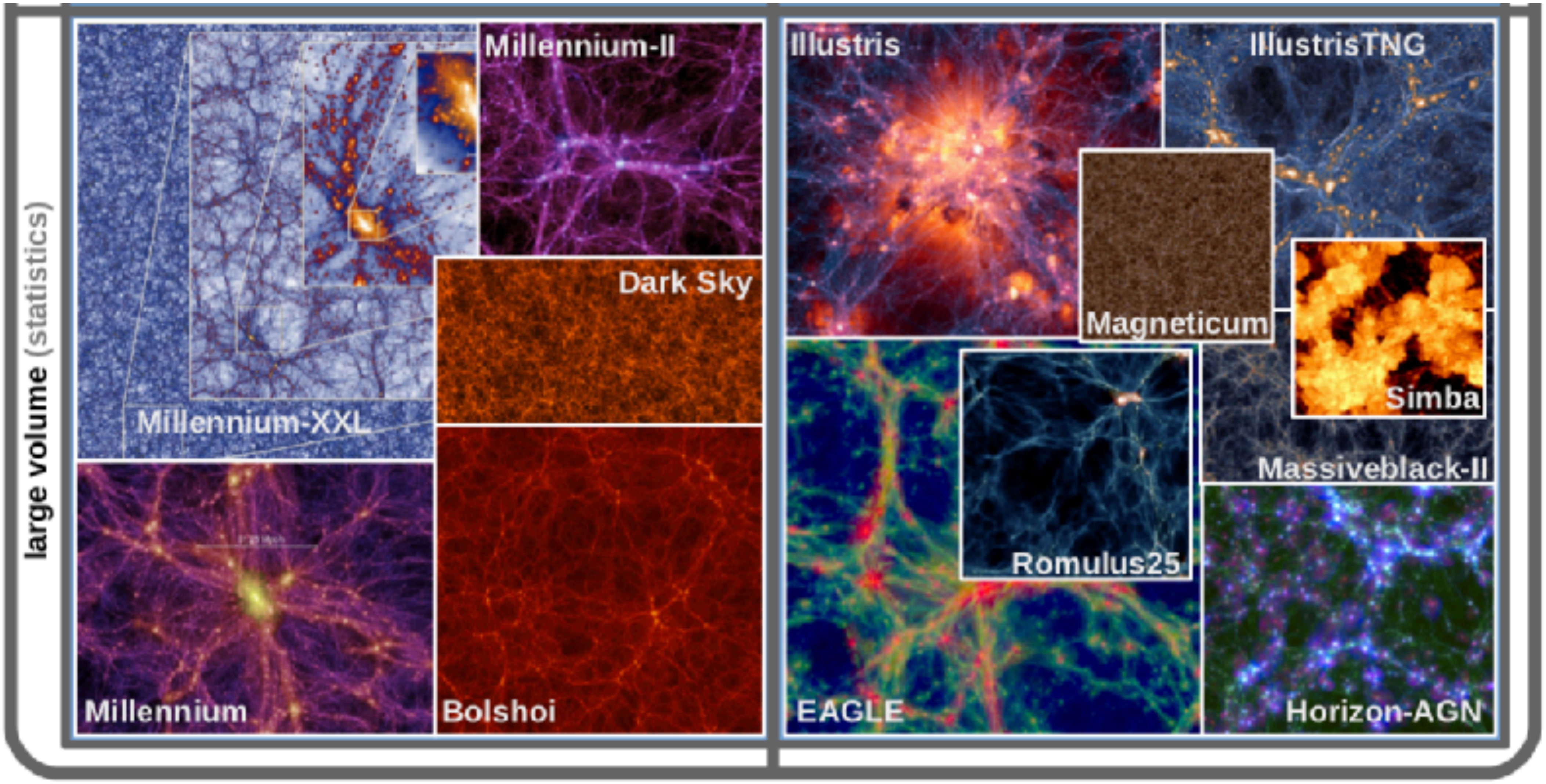
Gleyzes et al. (2016)

Need to measure non-Gaussianity directly

Large Scale Structure

Problem: low redshift universe is hard to model

DM-only DM + Baryons



Vogelsberger et al. (2019)

Large Scale Structure

Galaxies trace the underlying matter distribution

$$\delta_g(\vec{x}) = \sum b_{\mathcal{O}} \mathcal{O}$$
$$= b_1 \delta + b_2 \delta^2 + b_{\mathcal{G}_2} \left((\partial_j \partial_j \Phi)^2 - (\partial^2 \Phi)^2 \right) + b_3 \delta^3 + \cdots$$

We will expand in linear density contrast

Bias parameters are (largely) unknown / fit to data

This works very well in Lagrangian space

Schmittfull et al. (2018)

Late-time Non-Gaussianity

The problem is that this also produces a bispectrum

$$\langle \delta_g(\vec{k}_1)\delta_g(\vec{k}_2)\delta_g(\vec{k}_3)\rangle = b_2 b_1^2 \left(\langle [\delta^2](\vec{k}_1)\delta(\vec{k}_2)\delta(\vec{k}_3)\rangle + \text{perms}\right) + b_2^3 \langle [\delta^2](\vec{k}_1)[\delta^2](\vec{k}_2)[\delta^2](\vec{k}_3)\rangle + b_3 b_2 b_1 \left(\langle [\delta^3](\vec{k}_1)[\delta^2](\vec{k}_2)\delta(\vec{k}_3)\rangle + \text{perms}\right) + \cdots$$

Each new biasing term contributes

We must distinguish equilateral from all these terms

Late-time Non-Gaussianity

Quantify the problem with the inner-product

Bispectra are defined by $B\left(k_1, k_2, k_3\right) \equiv \left\langle \delta\left(\vec{k}_1\right) \delta\left(\vec{k}_2\right) \delta\left(\vec{k}_3\right) \right\rangle'$

$$B_2 = 2\left[P(k_1)P(k_2) + P(k_1)P(k_3) + P(k_2)P(k_3)\right],$$

$$B_{\mathcal{G}_2} = 2\left[\left(\left(\frac{\vec{k}_1 \cdot \vec{k}_2}{k_1 k_2}\right)^2 - 1\right) P(k_1)P(k_2) + \text{perms}\right],$$

$$B_{\text{eq}} = 162\, f_{\text{NL}}^{\text{eq}} \, \frac{\mathcal{T}(k_1)\mathcal{T}(k_2)\mathcal{T}(k_3)\Delta_\Phi^2}{k_1 k_2 k_3 (k_1 + k_2 + k_3)^3}$$

Late-time Non-Gaussianity

Quantify the problem with the inner-product

$$B_i \cdot B_j \equiv V \int \frac{\mathrm{d}^3 k_1 \ \mathrm{d}^3 k_2 \ \mathrm{d}^3 k_3}{(2\pi)^9} \frac{B_i\,(k_1, k_2, k_3)\, B_j\,(k_1, k_2, k_3)}{P\,(k_1)\, P\,(k_2)\, P\,(k_3)} (2\pi)^3 \delta_{\mathrm{D}} \left(\sum_i \vec{k}_i \right)$$

$$\cos\left(B_i, B_j\right) \equiv \frac{B_i \cdot B_j}{\sqrt{B_i \cdot B_i B_j \cdot B_j}}$$

With this we can quantify how similar the shapes are

$$\cos\left(B_2, B_{\mathrm{eq}}\right) \approx 0.96,$$
$$\cos\left(B_3, B_{\mathrm{eq}}\right) \approx 0.87,$$
$$\cos\left(B_4, B_{\mathrm{eq}}\right) \approx 0.80.$$

Theoretical Error

One can account for this challenge with theoretical error

Baldauf et al. (2016)

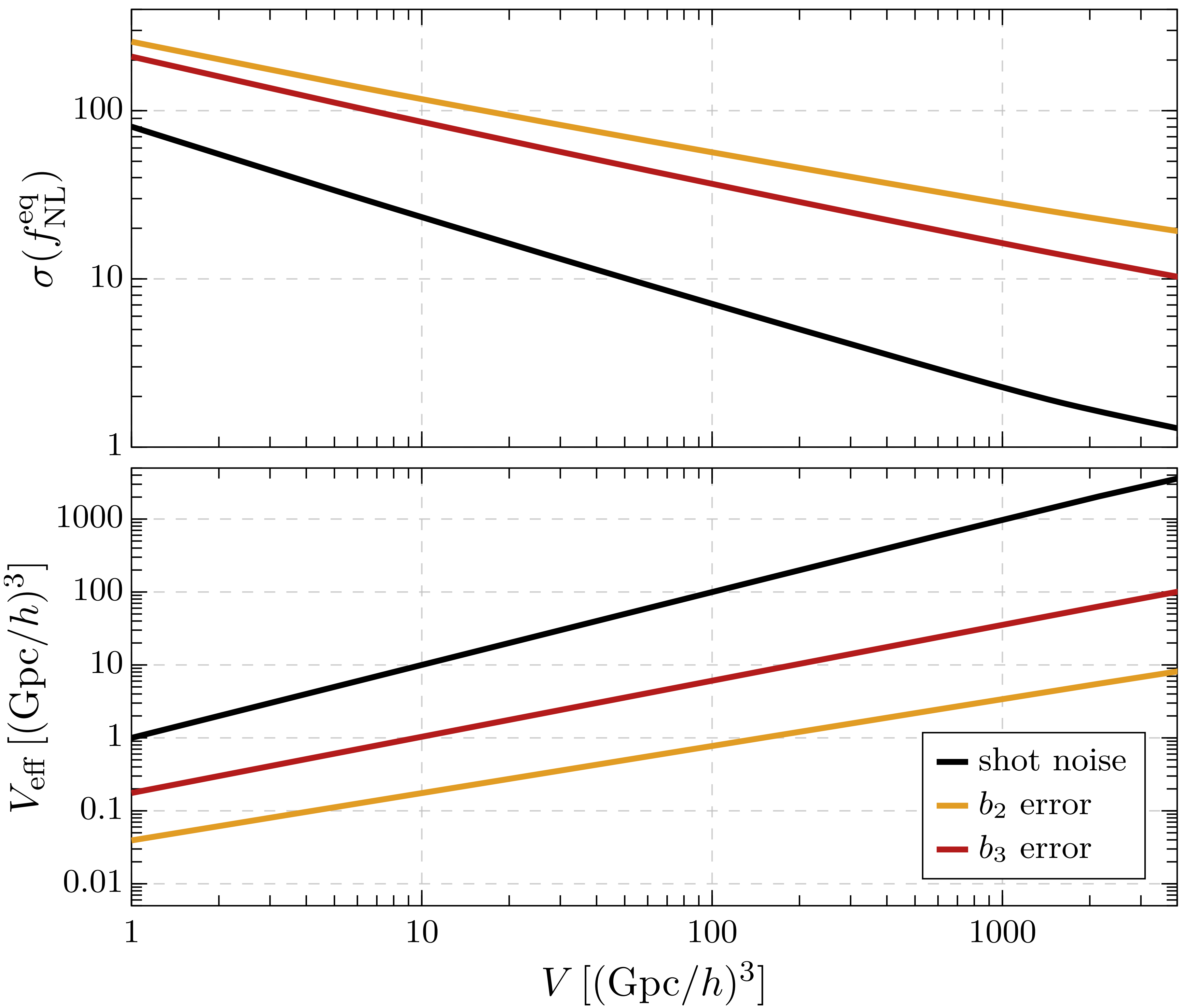
Basic idea: add next term in bias expansion to the noise

$$C^{-1}_{kk'} = \frac{(2\pi)^3}{V} \frac{1}{2\pi k^2 \text{k}} \left(P_g(k) + \frac{1}{\bar{n}} \right)^2 \delta_{kk'} + (C_e)_{kk'} \, ,$$

Simplifying assumption, local in k: $(C_e)_{kk'} \equiv C_e^2(k)\, \delta_{kk'}$

$$C_e(k) \approx \begin{cases} b_2^2 \langle [\delta^2](\vec{k})[\delta^2](-\vec{k}) \rangle' \approx P(k)(\hat{k}/0.31)^{1.8} & b_2 \text{ error} \\ b_3^2 \langle [\delta^3](\vec{k})[\delta^3](-\vec{k}) \rangle' \approx P(k)(\hat{k}/0.23)^{3.3} & b_3 \text{ error} \end{cases}$$

Theoretical Error



Summary of the Challenge

The low redshift universe is nonlinear:

- First principles modeling essentially impossible
- Uncertainty are absorbed into free parameters
- Resulting bispectra degenerate with NG
- Potential for biased measurement of $f_{\rm NL}^{\rm eq}$
- We can remove bias by introducing theoretical error
- Sensitivity to pNG decreases significantly

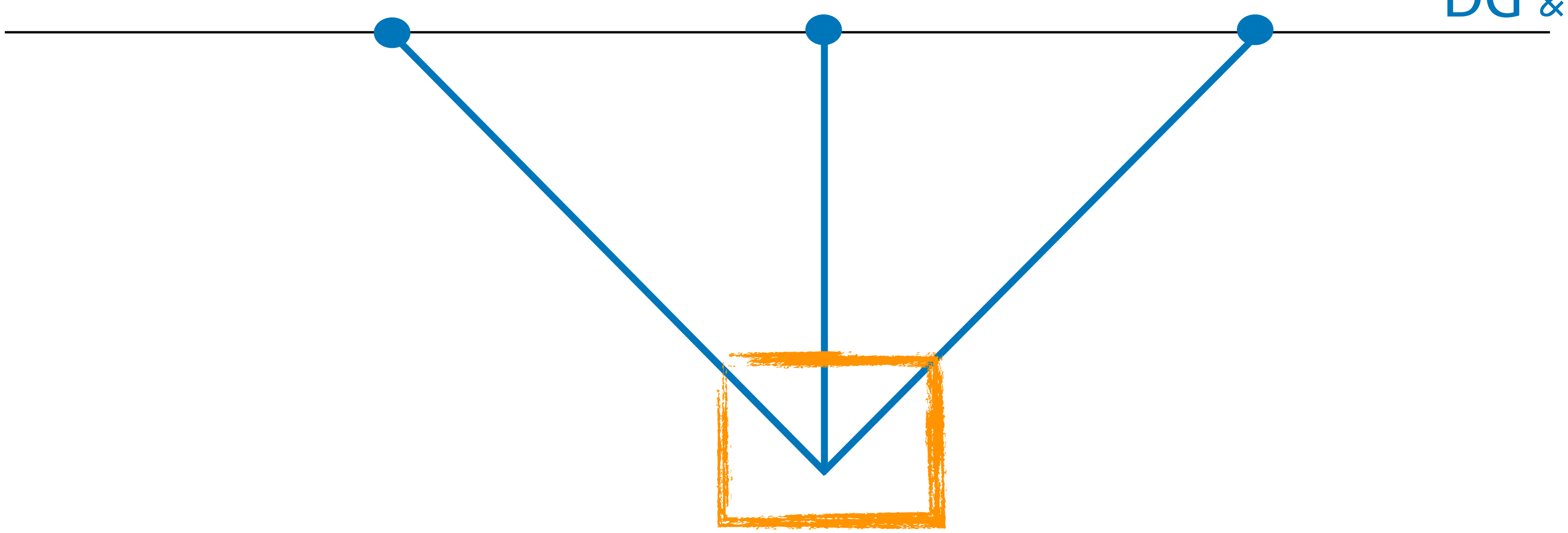
The Power of Locality

Locality

The inflationary signal is nonlocal in space

Created at the past intersection of the light cones

DG & Porto (2020)

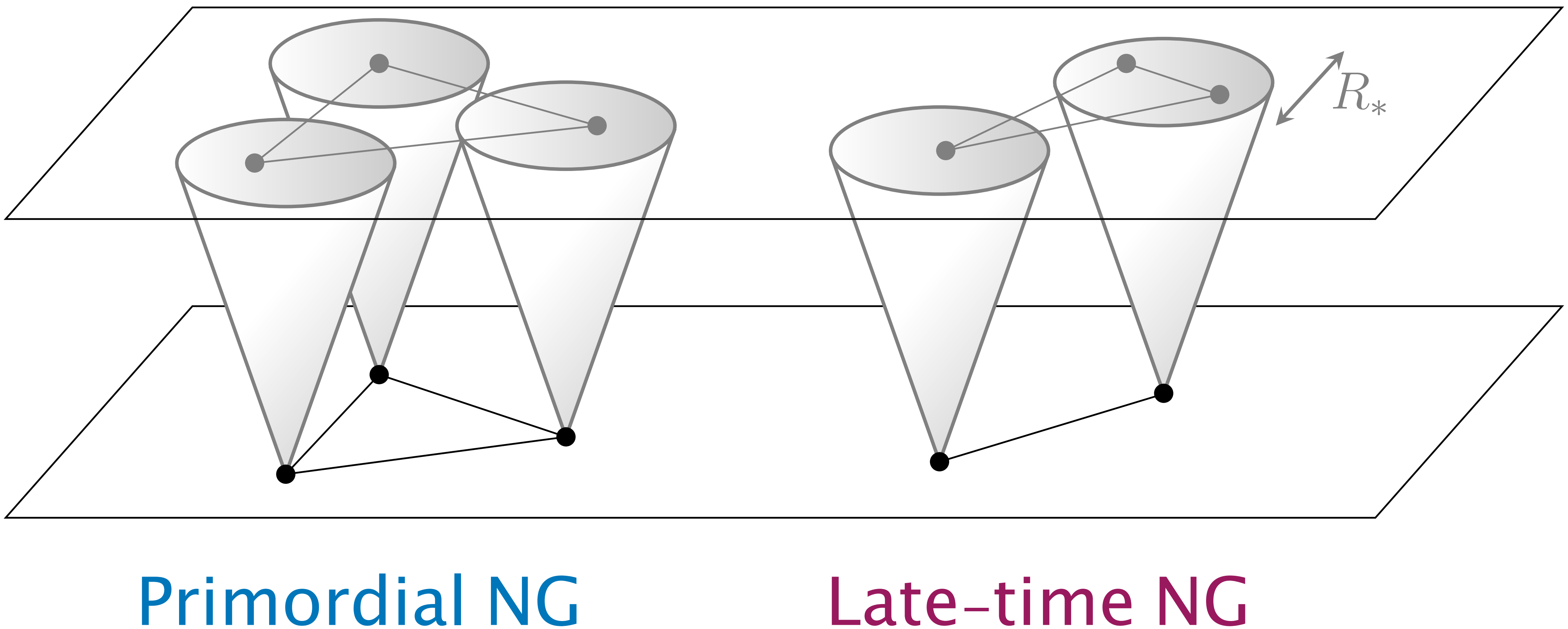


$$B_{\rm eq} = 162\, f_{\rm NL}^{\rm eq}\, \frac{\mathcal{T}(k_1)\mathcal{T}(k_2)\mathcal{T}(k_3)\Delta_\Phi^2}{k_1 k_2 k_3 (k_1+k_2+k_3)^3}$$

Proportional to 3 commutators: uniquely quantum!

Locality

Dark matter is slow: late-time evolution is ultra-local*



Nonlinearity can never completely mimic the signal

* we will neglect bulk-flows to simplify the presentation

Locality

Why didn't we see this in the cosines of the bispectra?

All the information about $f_{\rm NL}^{\rm eq}$ is in the bispectrum

But projecting LSS to the bispectrum misses information

$$\langle \delta(\vec{x}_1) \ldots \delta(\vec{x}_n) \rangle$$

$b_n \delta^n(\vec{x})$

Map

$$\vdots$$

Correlation Functions

$$\langle \delta(\vec{x}_1) \delta(\vec{x}_2) \rangle$$

We should quantity the cosine of the maps

Map-level Information

Work with the likelihood of a given map

$$\mathcal{L} = \mathcal{L}_{\rm CV} \times \mathcal{L}_{\rm obs}$$

$$\mathcal{L}_{\rm CV} = \exp\left(- \int \frac{d^3k}{(2\pi)^3} \frac{|\delta(\vec{k})|^2}{P(k)} \right)$$

$\delta(k)$ is the Gaussian initial conditions for the density

$P(k)$ is the matter power spectrum

This is just cosmic variance:

i.e. the likelihood of starting with this Gaussian field

Map-level Information

Work with the likelihood of a given map

$$\mathcal{L} = \mathcal{L}_{\rm CV} \times \mathcal{L}_{\rm obs}$$

$$\mathcal{L}_{\rm obs} = \exp\left(-\frac{1}{2N^2}\int\frac{d^3k}{(2\pi)^3}\left(\delta_g(\vec{k}) - \delta_g^{\rm obs}(\vec{k})\right)^2\right)$$

Model: $$\delta_g(\vec{x}) = \sum_{n=1}^{\infty}\left(b_n[\delta^n](\vec{x}) + n f_{\rm NL}^{\rm eq} b_n[\bar{\delta}^{n-1}\,\delta_{\rm NG}](\vec{x})\right)$$

$$\delta_{\rm NG}(\vec{k}) = \int\frac{d^3p\,d^3q}{(2\pi)^3}\frac{B_{\rm eq}(\vec{p},\vec{q})}{6\,P_\Phi(p)P_\Phi(q)}\,\delta(\vec{p})\delta(\vec{q})\,\delta_{\rm D}(\vec{p}+\vec{q}-\vec{k})$$

Map-level Information

Work with the likelihood of a given map

$$\mathcal{L} = \mathcal{L}_{\rm CV} \times \mathcal{L}_{\rm obs}$$

$$\mathcal{L}_{\rm obs} = \exp\left(-\frac{1}{2N^2}\int\frac{d^3k}{(2\pi)^3}\left(\delta_g(\vec{k}) - \delta_g^{\rm obs}(\vec{k})\right)^2\right)$$

Data: $\delta_g^{\rm obs}(\vec{x}) = \bar{\delta}_g(\vec{x}) + \epsilon(\vec{x})$ (observed map of galaxies)

Assume stochastic bias dominated by shot noise

$$\langle\epsilon(\vec{x})\epsilon(\vec{x}')\rangle = N^2\delta(\vec{x}-\vec{x}') \qquad N^2 = \frac{1}{\bar{n}_g}$$

Maximum Likelihood Forecasts

In principle, one could just MCMC the full likelihood

Goal: find best fit contours for all the variables

$$\delta(\vec{x}),\ b_n,\ f_{\rm NL}^{\rm eq},\ \ldots$$

Note we are reconstructing the initial map (not $P(k)$)

This is a very high-dimensional optimization problem

Many groups working on this, speed up via ML/AI

We don't need to solve this problem for forecasting

Maximum Likelihood Forecasts

First, we find the maximum likelihood ICs

$$\bar{\delta}^{\rm obs}(\vec{x}) \;=\; \sum_n \tilde{b}_n \star [(\delta_g^{\rm obs})^n](\vec{x}) \;-\; f_{\rm NL}^{\rm eq}\,\bar{\delta}_{\rm NG}[\delta_g^{\rm obs}](\vec{x})$$

$$\tilde{b}_1(\vec{k}) = \frac{b_1 P(k)}{b_1^2 P(k) + N^2}\,,$$

$$\tilde{b}_2(\vec{k}) = -\frac{b_1 P(k)}{b_1^2 P(k) + N^2}\,\frac{b_2}{b_1^2}$$

filter out the noise dominated modes

Basically we are just inverting the map from $\delta \to \delta_g$

Possible in the perturbative regime

Maximum Likelihood Forecasts

Next, calculate the Fisher matrix with fixed map

$$F_{ij} = -\frac{\partial^2}{\partial\theta_i \partial\theta_j} \log \mathcal{L}\Big|_{b_n=\bar{b}_n, f_{\mathrm{NL}}^{\mathrm{eq}}=0}$$

Where $\theta_i \in [\{b_n\}, f_{\mathrm{NL}}^{\mathrm{eq}}]$

For $\theta_i \in \{b_2,\ f_{\mathrm{NL}}^{\mathrm{eq}}\}$, this is identical to the bispectrum

Reason: we use the bispectrum because it is optimal

In fact, the usual proof uses the map-level information

Maximum Likelihood Forecasts

The benefits of the map become apparent for $b_{n>2}$

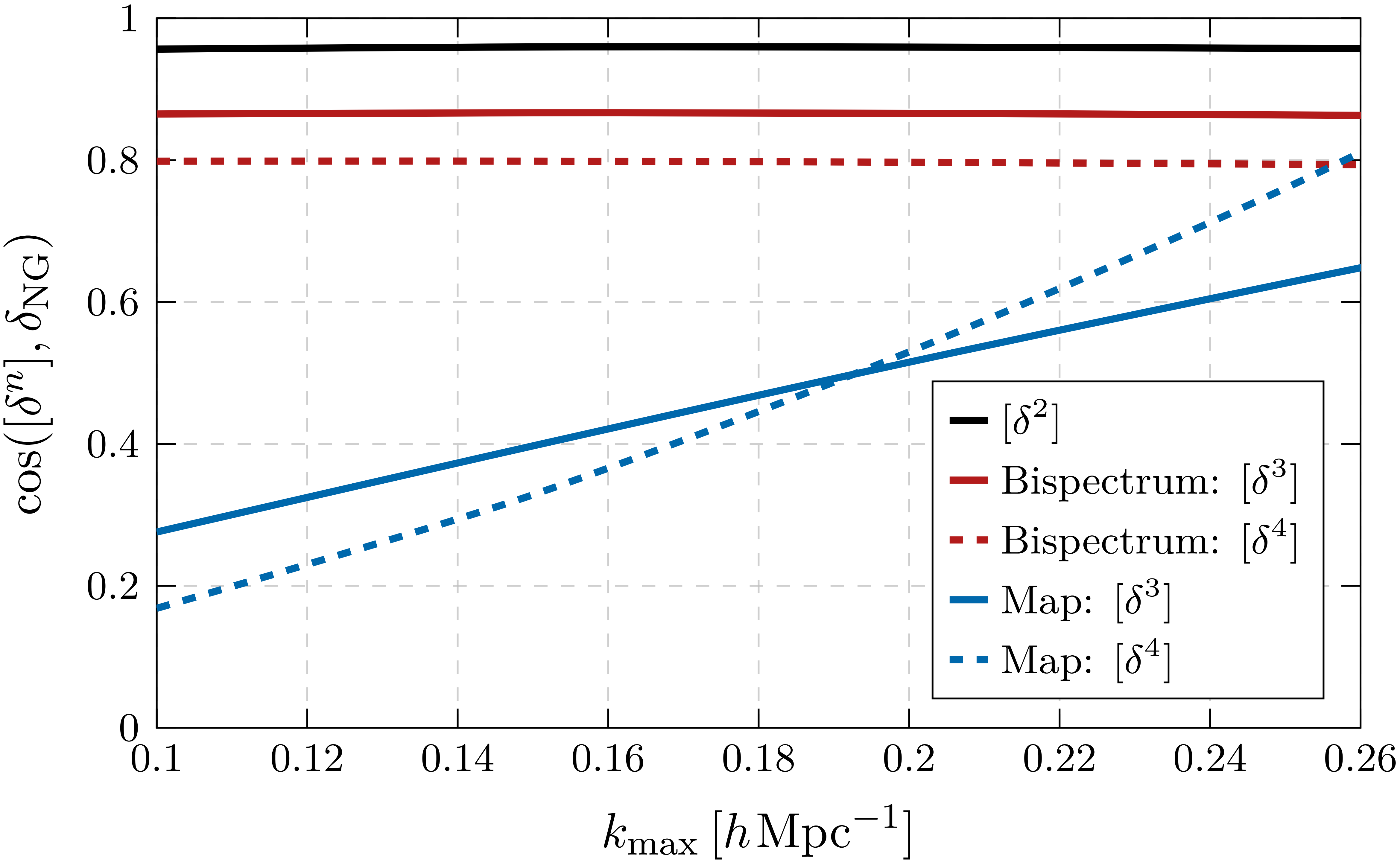
$$F_{n,m} \to V \int \frac{d^3k}{(2\pi)^3} \frac{\langle [\delta^n](\vec{k})[\delta^m](-\vec{k})\rangle' + n\delta_{n,m} n! \sigma^{2n}}{P(k) + N^2} \,,$$

$$F_{n,\mathrm{eq}} \to V \sum_m \int \frac{d^3k}{(2\pi)^3} \, m\bar{b}_m \frac{\langle [\delta^n](\vec{k})[\delta^{m-1}\delta_{\mathrm{NG}}](-\vec{k})\rangle'}{P(k) + N^2}$$

Define a cosine for the maps

$$\cos([\delta^n], \delta_{\mathrm{NG}}) = \frac{F_{n,\mathrm{eq}}}{\sqrt{F_{n,n} F_{\mathrm{eq,eq}}}} \propto \left(\langle \delta^2(\vec{x})\rangle'\right)^{(n-2)/2}$$

Maximum Likelihood Forecasts



Lessons

Changes maps from biasing are distinguishable from NG

In terms of correlation function the dominant info is

$$b_n \delta^n(\vec{x}) \longrightarrow \text{(N+1)-point function}$$

Reflected in two-point functions of composite operators

$$\langle [\delta^n](\vec{x}) [\delta^m](0) \rangle = 0 \qquad m \neq n$$

$$\langle [\delta^n](\vec{x}) \delta_{\rm NG}(0) \rangle = 0 \qquad n > 2$$


Future Prospects

Low-z Forecasts

We will consider a box at z=0 of some fixed size V

This is a harder problem that a realistic survey

Higher-z will be better, test the worst case version

Unless otherwise stated

$$V = 300\,(\mathrm{Gpc}/h)^3 \qquad N_g = 10^{10}$$

$$k_{\max} = 0.2\,h\,\mathrm{Mpc}^{-1}$$

We want to see the limitations of nonlinearity only

Low-z Forecasts

Expand the biasing model to include additional terms:

$$\delta_g(\vec{x}) = b_1\delta + b_2\delta^2 + b_{\mathcal{G}_2}\left((\partial_j\partial_j\Phi)^2 - (\partial^2\Phi)^2\right) + b_{\partial^2}\partial^2\delta + b_3\delta^3 + b_4\delta^4 + \cdots$$

All quadratic terms are captured by bispectrum

Cubic terms and higher we distinguish map & bispectrum

Low-z Forecasts

We will compare to bispectrum with modeling noise

i.e. theoretical error applied equally to all correlators

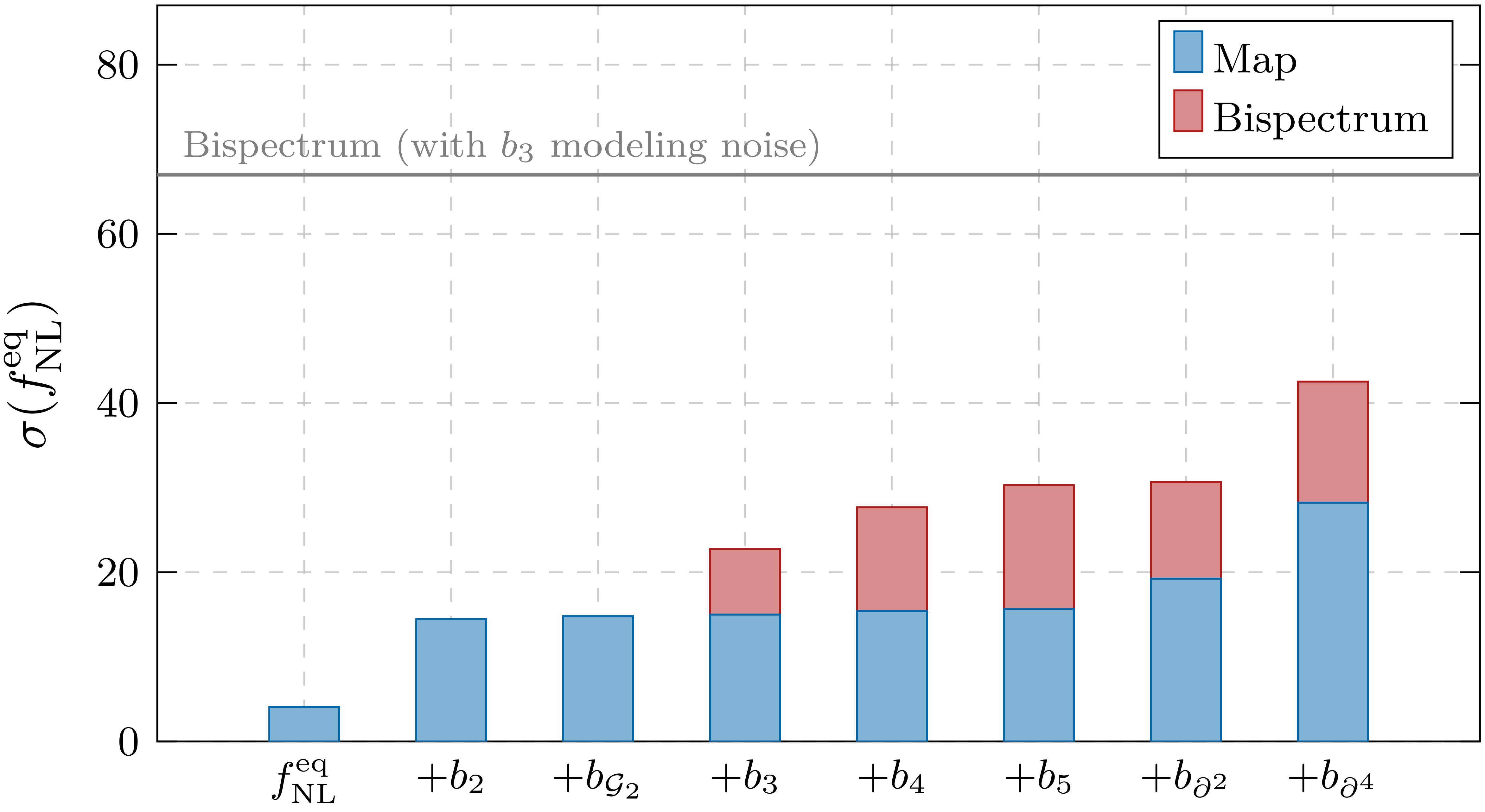
$$P(k) + \frac{1}{\bar{n}} \to P(k) \boxed{+ N^2 + k^2 \Delta k V C_e(k)}$$

Modeling noise = extra contribution to Gaussian noise

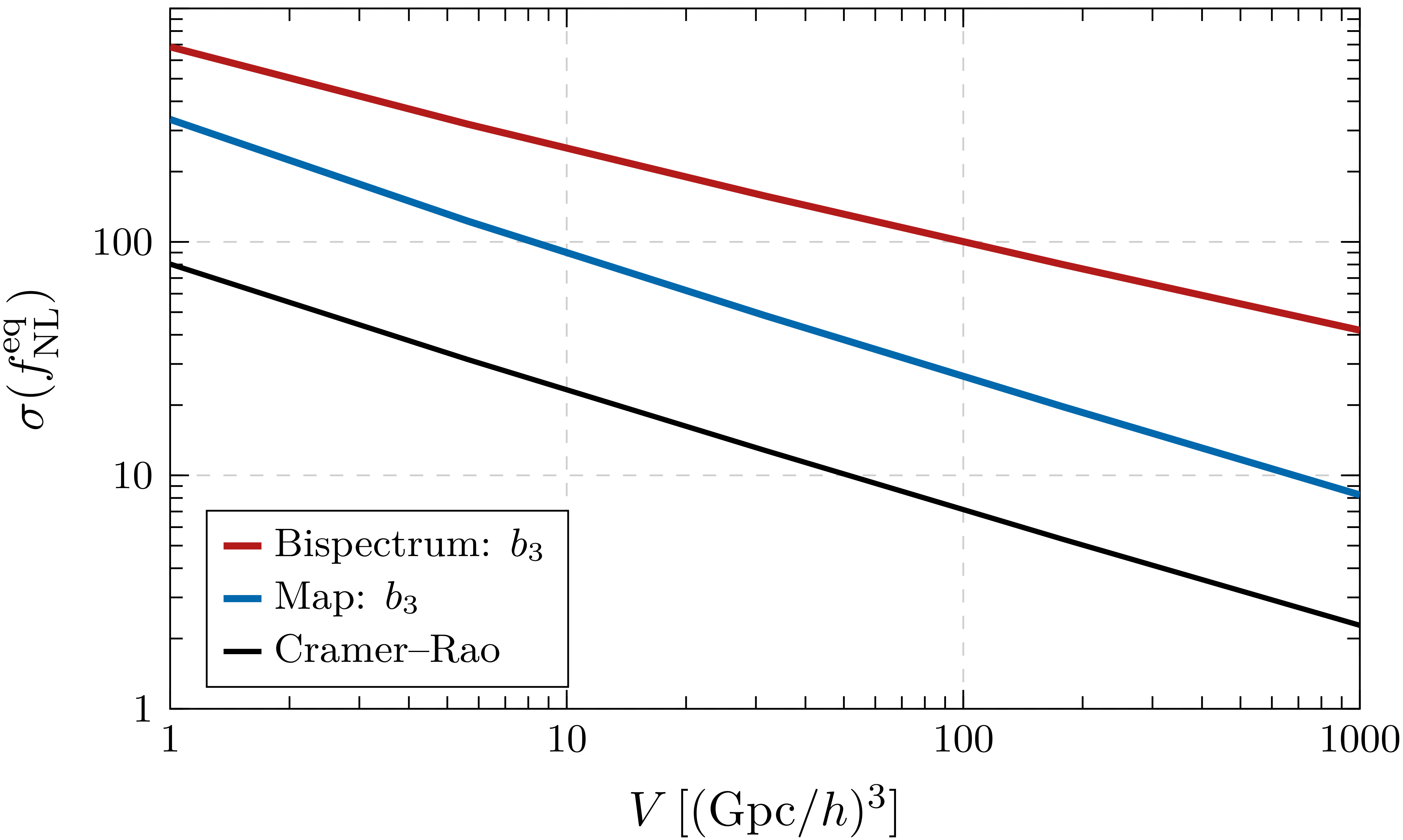
$$C_e(k) \approx b_3^2 \langle [\delta^3](\vec{k})[\delta^3](-\vec{k}) \rangle' \approx P(k)(\hat{k}/0.23)^{3.3} \qquad b_3 \text{ error}$$

We will compare to marginalizing over $b_{n \geq 3}$

Low-z Forecasts



Low-z Forecasts



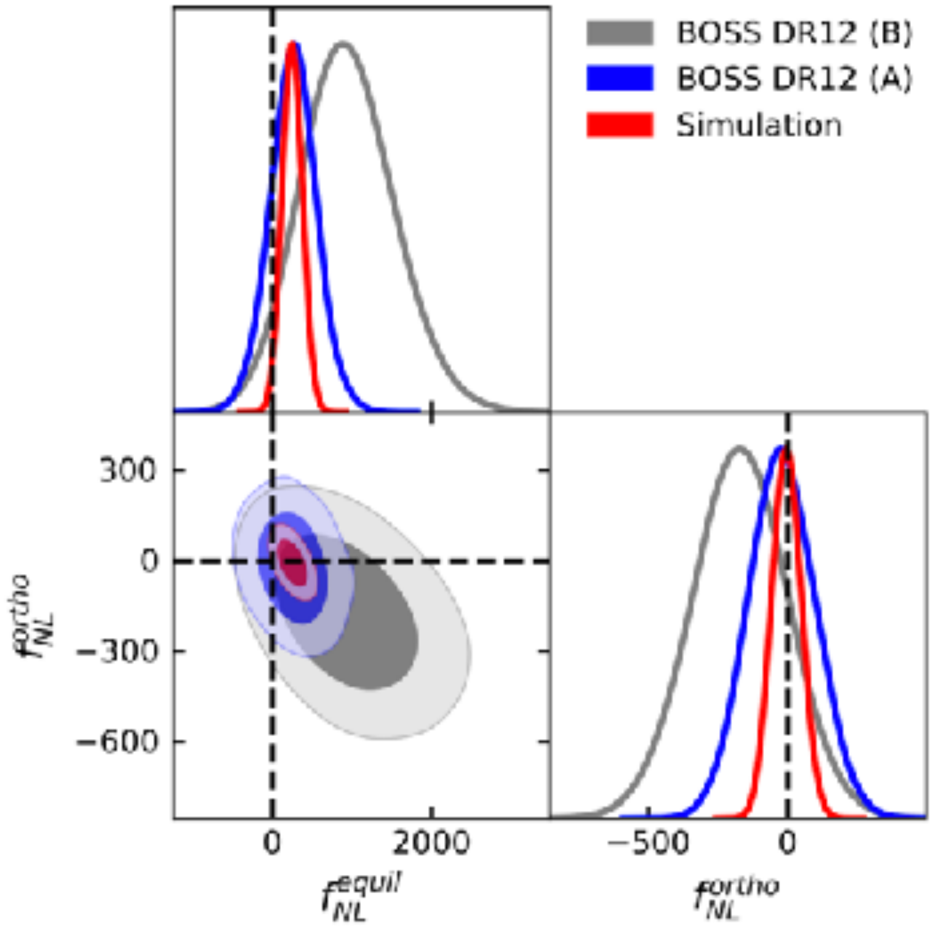
Summary

What we learn from map-level forecasts:

- Map level forecasts ~ bispectrum with $b_2, b_{\mathcal{G}_2}, f_{\rm NL}^{\rm eq.}$
- Theoretical/modeling error is overly conservative
- Bias coefficients capture our uncertainty of modeling
- Higher derivatives terms are dangerous = nonlocality

Status

First ever LSS constraints on pNG from the bispectrum



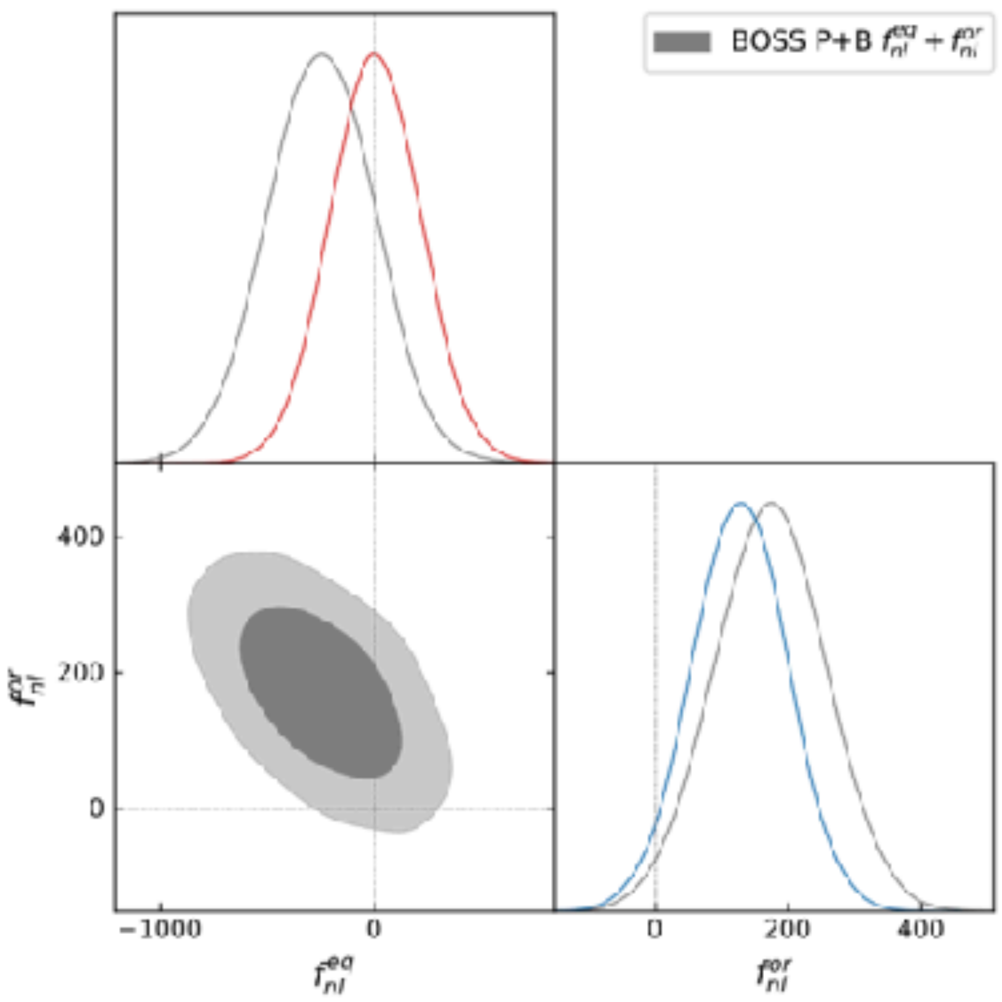
Analysis B $f_{\rm NL}^{\rm eq} = 940^{+570}_{-650}$

Analysis A $f_{\rm NL}^{\rm eq} = 257^{+300}_{-300}$

Cabass, Ivanov, Philcox, et al. (2022)

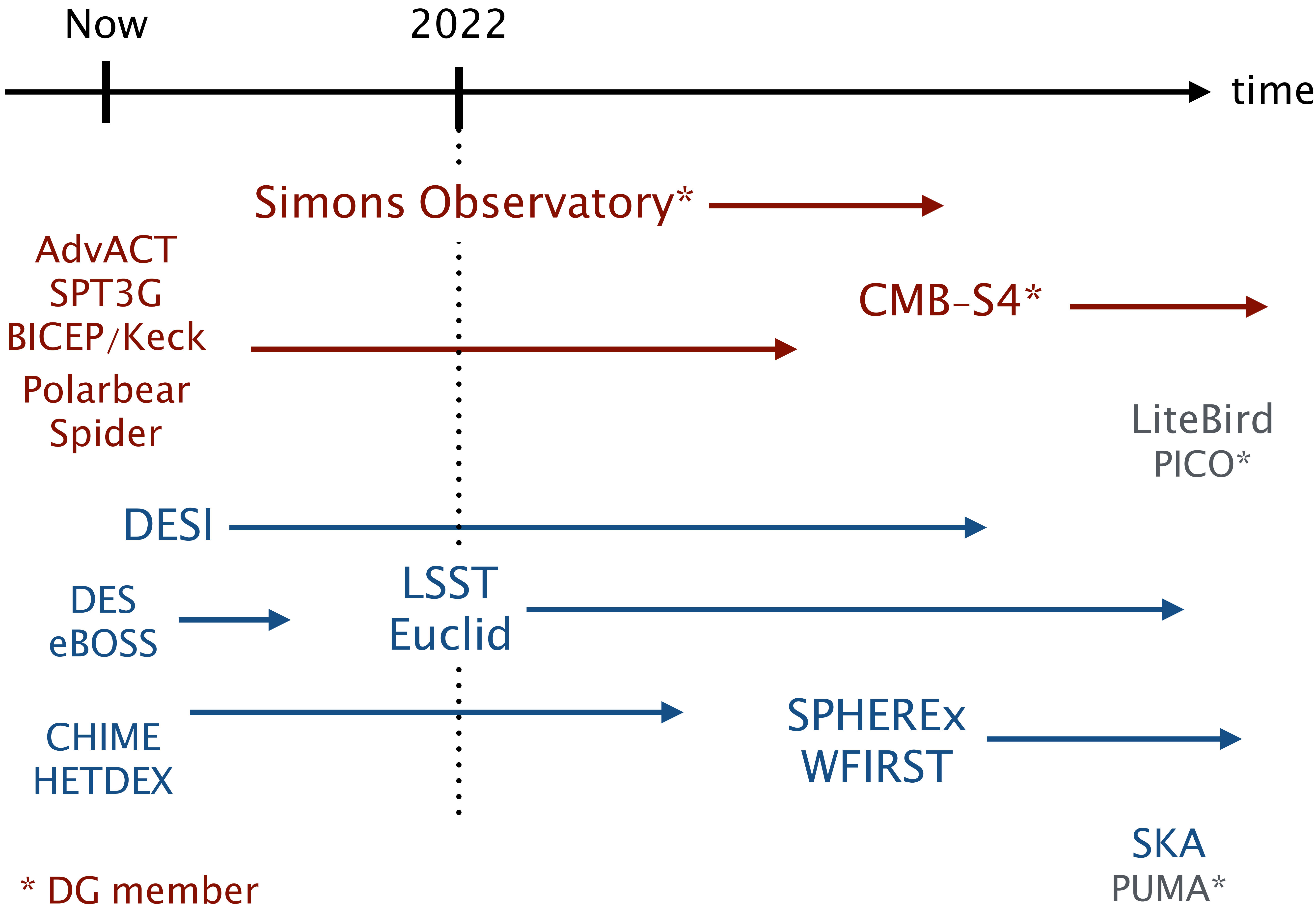
Status

First ever LSS constraints on pNG from the bispectrum



	BOSS
$f_{\rm NL}^{\rm equil.}$	2 ± 212
$f_{\rm NL}^{\rm orth.}$	126 ± 72
$f_{\rm NL}^{\rm loc.}$	-30 ± 29

D'Amico, Lewandowski, et al. (2022)


Now
2022
time
Simons Observatory*
AdvACT
SPT3G
BICEP/Keck
Polarbear
Spider
CMB-S4*
LiteBird
PICO*
DESI
DES
eBOSS
LSST
Euclid
CHIME
HETDEX
SPHEREx
WFIRST
SKA
PUMA*
* DG member

Status

Bispectrum forecasts without T.E = map level forecasts

Euclid: $\sigma(f_{\rm NL}^{\rm eq}) = 7.5$ $k_{\rm max} = 0.15\, h\, {\rm Mpc}^{-1}/D(z)$

Euclid: $\sigma(f_{\rm NL}^{\rm eq}) = 16$ $k_{\rm max} = 0.1\, h\, {\rm Mpc}^{-1}/D(z)$

21cm intensity mapping (e.g. PUMA)

$\sigma(f_{\rm NL}^{\rm eq}) = 4.5$ $k_{\rm max} = 0.1\, h\, {\rm Mpc}^{-1}/D(z)$

with theoretical error

Summary

Summary

The future of pNG lies in LSS surveys

- There has been a lot of pessimism pNG in LSS
- Unknown nonlinear physics is a fundamental concern
- We showed locality protects us from these issues
- Analysis is possible without perfect modeling
- Challenge now is to put this into practice

Future Directions

There are additional challenges in real surveys:

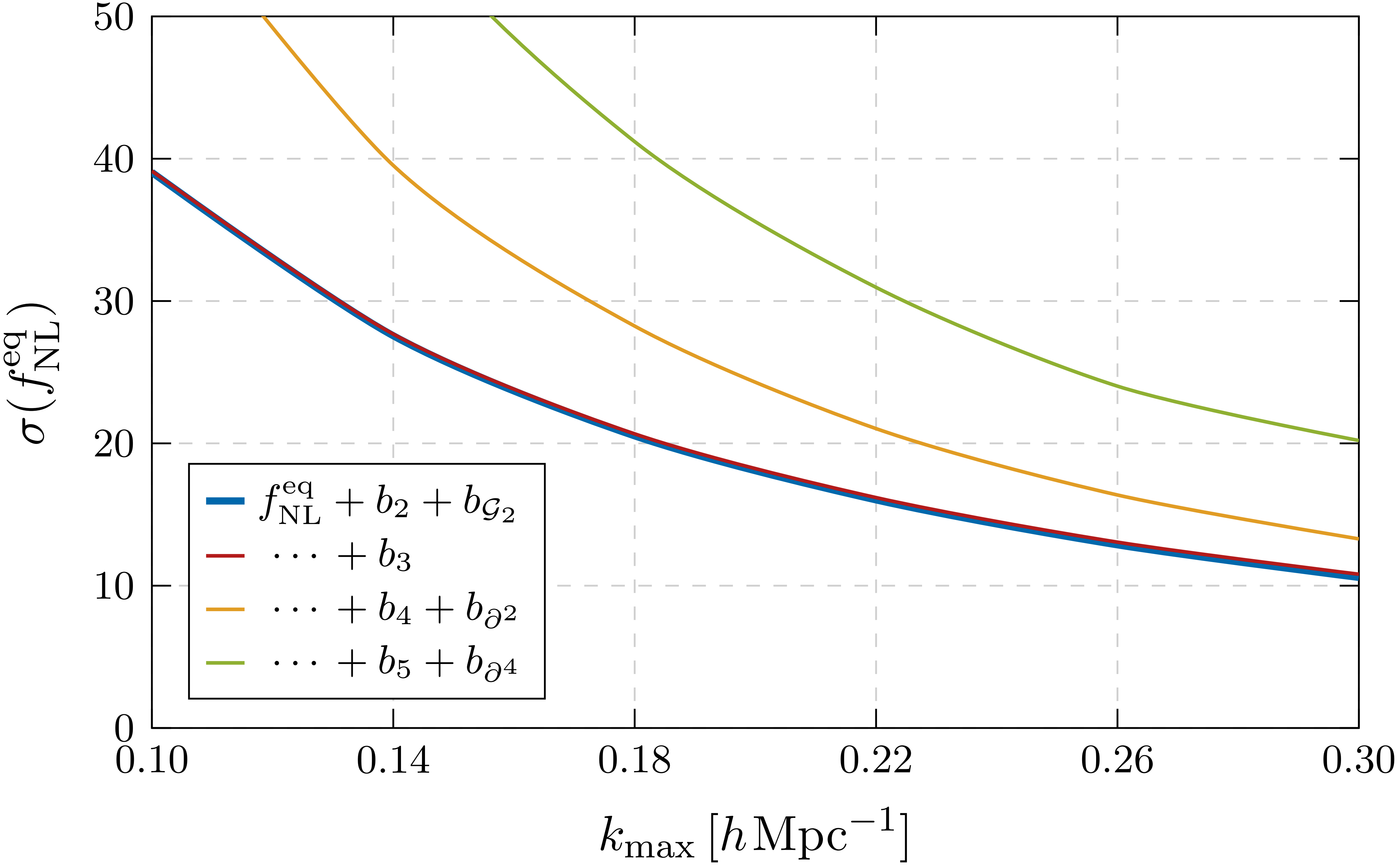
- We have neglected RSD and bulk flow / reconstruction
- Galaxy surveys may be limited by shot noise
- 21cm intensity avoid shot noise
- Currently limited by beams and foregrounds

CHIME collaboration (2022)


Thank you


Extra Slides

Low-z Forecasts



Low-z Forecasts