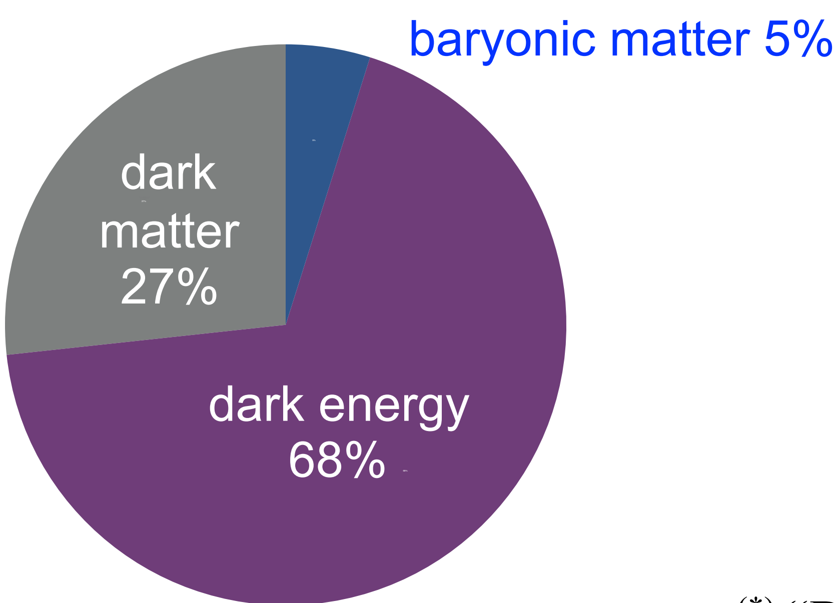
Observational cosmology

Kendrick Smith
TRISEP 2018

Part 1: the standard cosmological model

All current data can be fit by a 6-parameter cosmological model!

$\rho_\Lambda = (2.56 \pm 0.04) \times 10^{-47}\ \mathrm{GeV}^4$	Dark energy density (c.c.)
$\Omega_b = 0.0486 \pm 0.0007$	Baryonic(*) matter abundance
$\Omega_c = 0.267 \pm 0.009$	Cold dark matter abundance
$\Delta_\zeta^2 = (2.11 \pm 0.05) \times 10^{-9}$	Initial power spectrum amplitude
$n_s = 0.967 \pm 0.004$	Spectral index
$\tau = 0.058 \pm 0.012$	CMB optical depth



(*) "Baryons" = protons + neutrons + electrons(!)

Ingredients in the standard cosmological model:

- Background metric is FRW
- Expansion history is ΛCDM
- Initial perturbations are Gaussian random
- Initial perturbations are scalar adiabatic
- Power spectrum of initial perturbations is a power law: $(k^3/2\pi^2)P(k) = \Delta_\zeta^2 (k/k_0)^{n_s-1}$

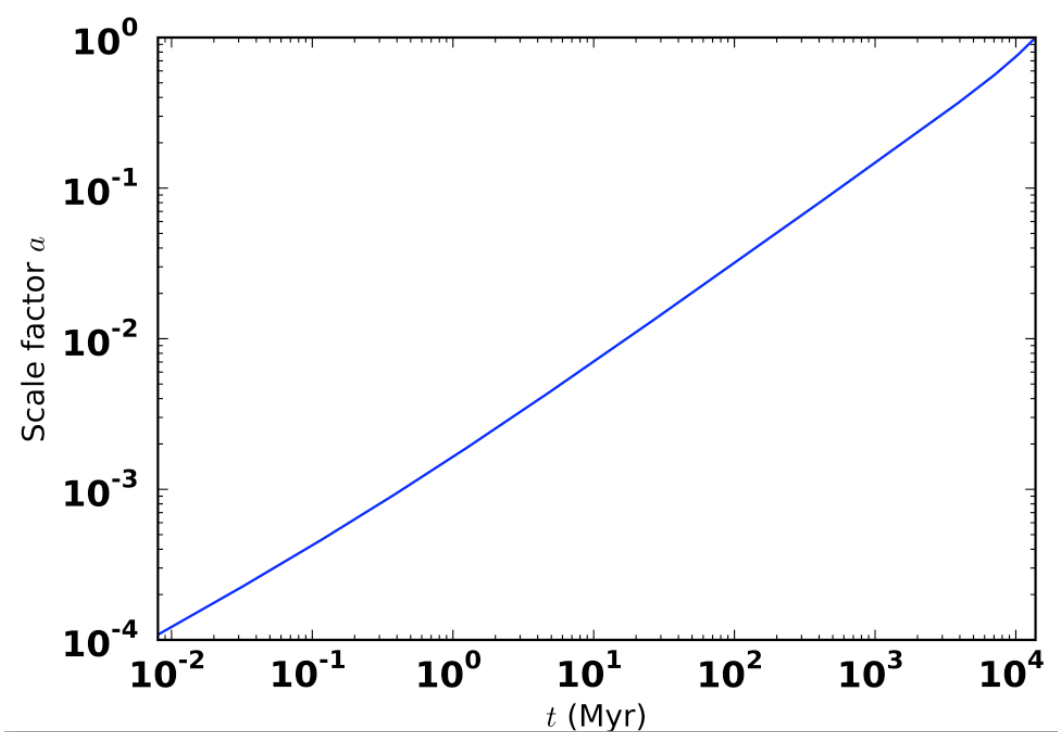
In the next few slides, we'll describe these ingredients at an informal level, just to set the stage. (Focus of these lectures is data analysis and statistics, theory lectures are next week!)

“Background metric is FRW”

The expansion of the universe is described by a function a(t), such that a=0 at the big bang, and a=1 today. (a = “scale factor”)

Formal meaning: metric is $ds^2 = -dt^2 + a(t)^2\, dx^2$

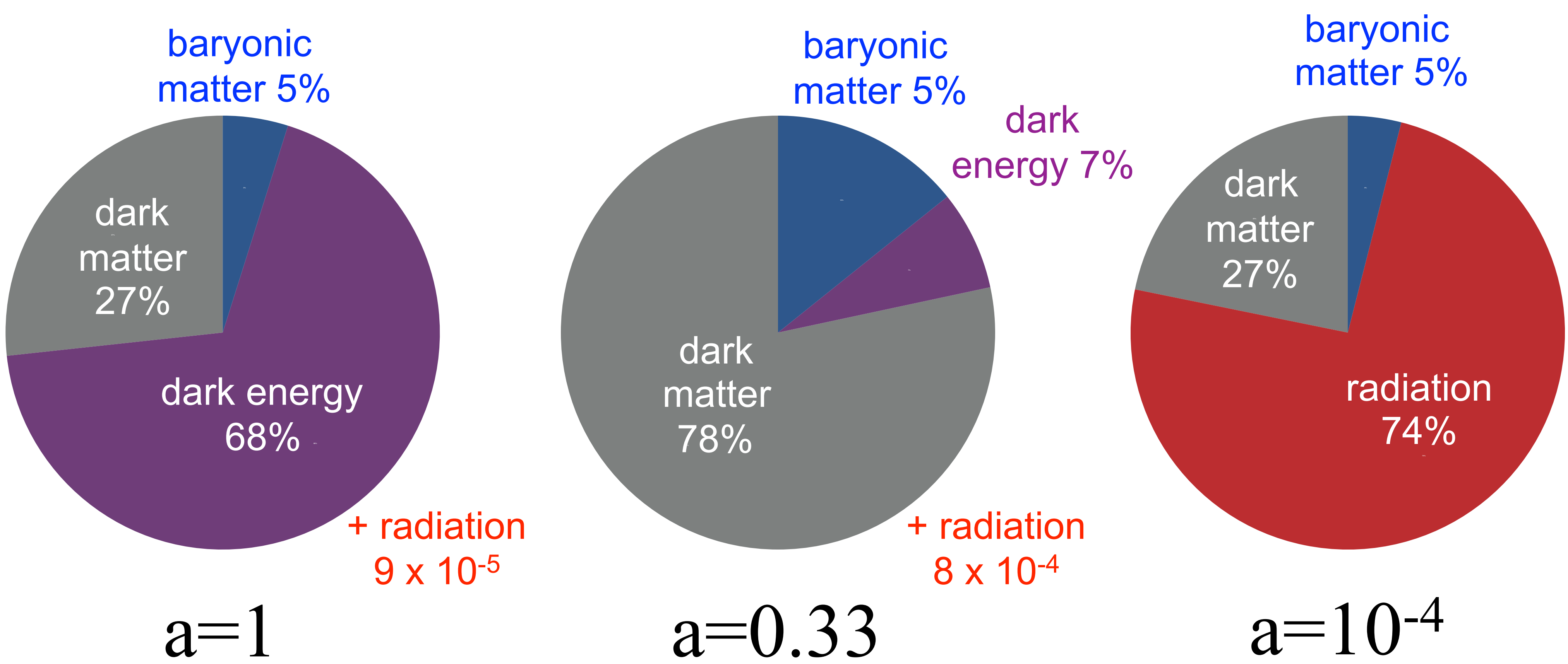
Intuitive meaning: if points x, x’ are separated by distance D today, then their separation at time t is a(t)D.



“Expansion history is ΛCDM”

Energy densities evolve with scale factor a(t):

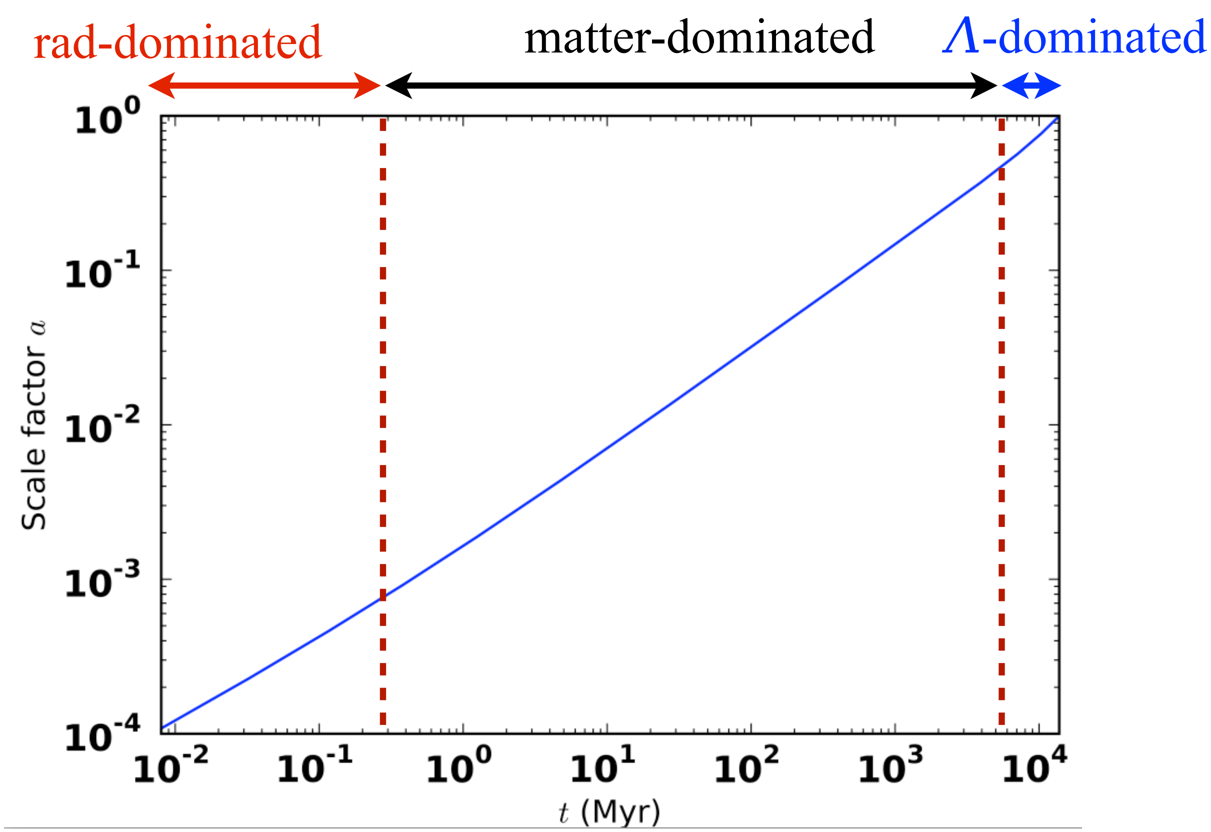
$\rho_{\mathrm{de}} = \text{constant}$	dark energy (assuming it is a c.c.!)
$\rho_{\mathrm{m}} \propto a(t)^{-3}$	nonrelativistic matter (dark + baryonic)
$\rho_{\mathrm{rad}} \propto a(t)^{-4}$	relativistic particles (photons, neutrinos)



"Expansion history is ΛCDM"

Scale factor a(t) evolves via Friedmann equation

$$\frac{d\log a}{dt} = \left(\frac{8\pi G}{3}\rho_{\rm tot}\right)^{1/2} = \left(\frac{8\pi G}{3}\left(\rho_{\rm de} + \rho_{\rm m}(a) + \rho_{\rm rad}(a)\right)\right)^{1/2}$$



The expansion history is parameterized by the first three parameters in the standard model (ρ_Λ, Ω_b, Ω_c).

$\rho_\Lambda = (2.56 \pm 0.04) \times 10^{-47}$ GeV^4	Dark energy density (c.c.)
$\Omega_b = 0.0486 \pm 0.0007$	Baryonic matter abundance
$\Omega_c = 0.267 \pm 0.009$	Cold dark matter abundance
$\Delta\zeta^2 = (2.11 \pm 0.05) \times 10^{-9}$	Initial power spectrum amplitude
$n_s = 0.967 \pm 0.004$	Spectral index
$\tau = 0.058 \pm 0.012$	CMB optical depth

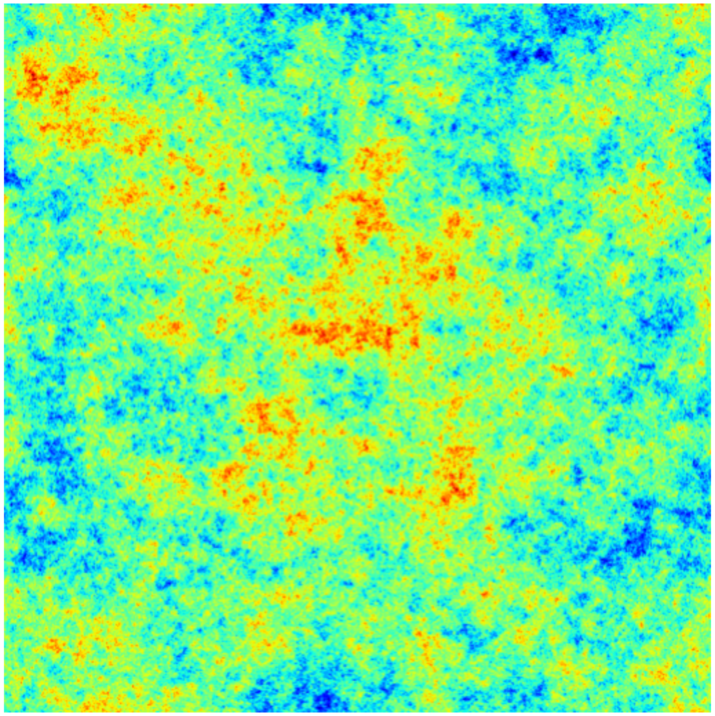
So far, we have not talked about perturbations. The next two parameters ($\Delta\zeta^2$, n_s) specify the *initial* perturbations.

Initial conditions: at early times, the FRW metric has small perturbations.

$$ds^2 = -dt^2 + a(t)^2 e^{2\zeta(x)} dx^2$$

The field $\zeta(x)$ is called the "adiabatic curvature" or the "initial curvature". This is a random field whose statistics can be described informally by the following statements:

- Initial perturbations are self-similar (no preferred scale)
- Almost scale-invariant, small trend toward more power on large scales.
- Characteristic size of fluctuations is $\Delta_\zeta \sim (5 \times 10^{-5})$



Initial conditions: at early times, the FRW metric has small perturbations.

$$ds^2 = -dt^2 + a(t)^2 e^{2\zeta(x)} dx^2$$

More formally, $\zeta(x)$ is a Gaussian random field with the following power spectrum $P_\zeta(k)$. (This will be defined precisely later!)

$$\frac{k^3}{2\pi^2} P_\zeta(k) = \Delta_\zeta^2 \left(\frac{k}{0.05\ h\ \mathrm{Mpc}^{-1}} \right)^{n_s - 1}$$

with free parameters

$\Delta_\zeta^2 = (2.11 \pm 0.05) \times 10^{-9}$	Initial power spectrum amplitude
$n_s = 0.967 \pm 0.004$	Spectral index

"Initial perturbations are scalar adiabatic".

- "Scalar" means that there are no gravity wave perturbations in the initial metric. (Some models of inflation predict this, but so far it has not been observed.)

$$ds^2 = -dt^2 + a(t)^2 e^{2\zeta(x)}(\delta_{ij} + \cancel{h_{ij}(x)})$$

absent

- "Adiabatic" is more technical. It means that the ζ field also completely determines the perturbations in the stress-energy tensor, by a universal set of rules which will be explained later!

$$\rho(\mathbf{x}, t) = \bar{\rho}(t)\left(1 + \frac{4}{7}\zeta(\mathbf{x})\right)$$

$$\ldots$$

The statistics of the initial perturbations are parameterized by parameters (Δ_ζ^2, n_s) below.

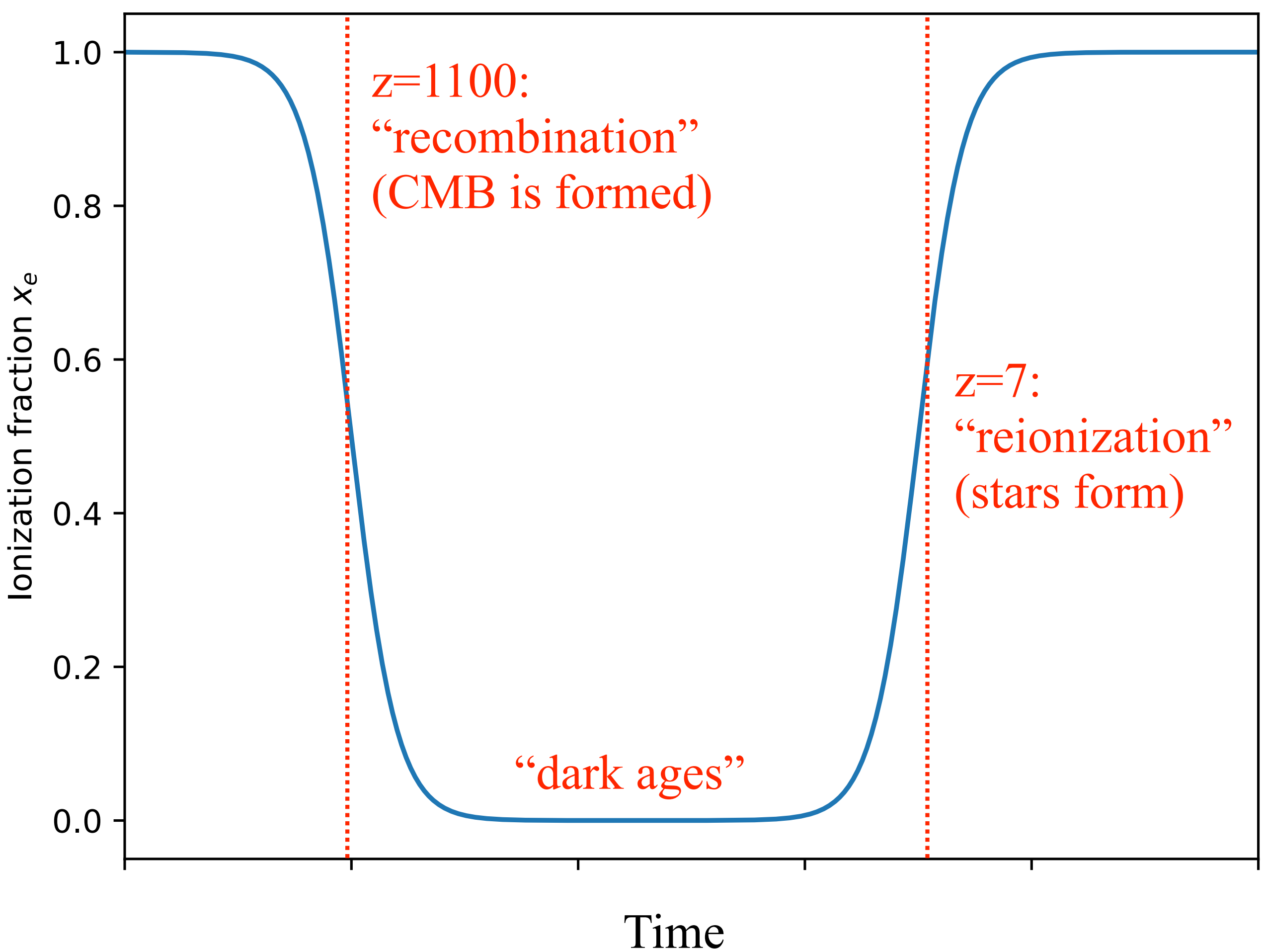
$\rho_\Lambda = (2.56 \pm 0.04) \times 10^{-47}$ GeV^4	Dark energy density (c.c.)
$\Omega_b = 0.0486 \pm 0.0007$	Baryonic matter abundance
$\Omega_c = 0.267 \pm 0.009$	Cold dark matter abundance
$\Delta_\zeta^2 = (2.11 \pm 0.05) \times 10^{-9}$	Initial power spectrum amplitude
$n_s = 0.967 \pm 0.004$	Spectral index
$\tau = 0.058 \pm 0.012$	CMB optical depth

The final parameter τ is an astrophysical nuisance parameter which we define for completeness.

Ionization history of the universe

$x_e(t)$ = electron ionization fraction

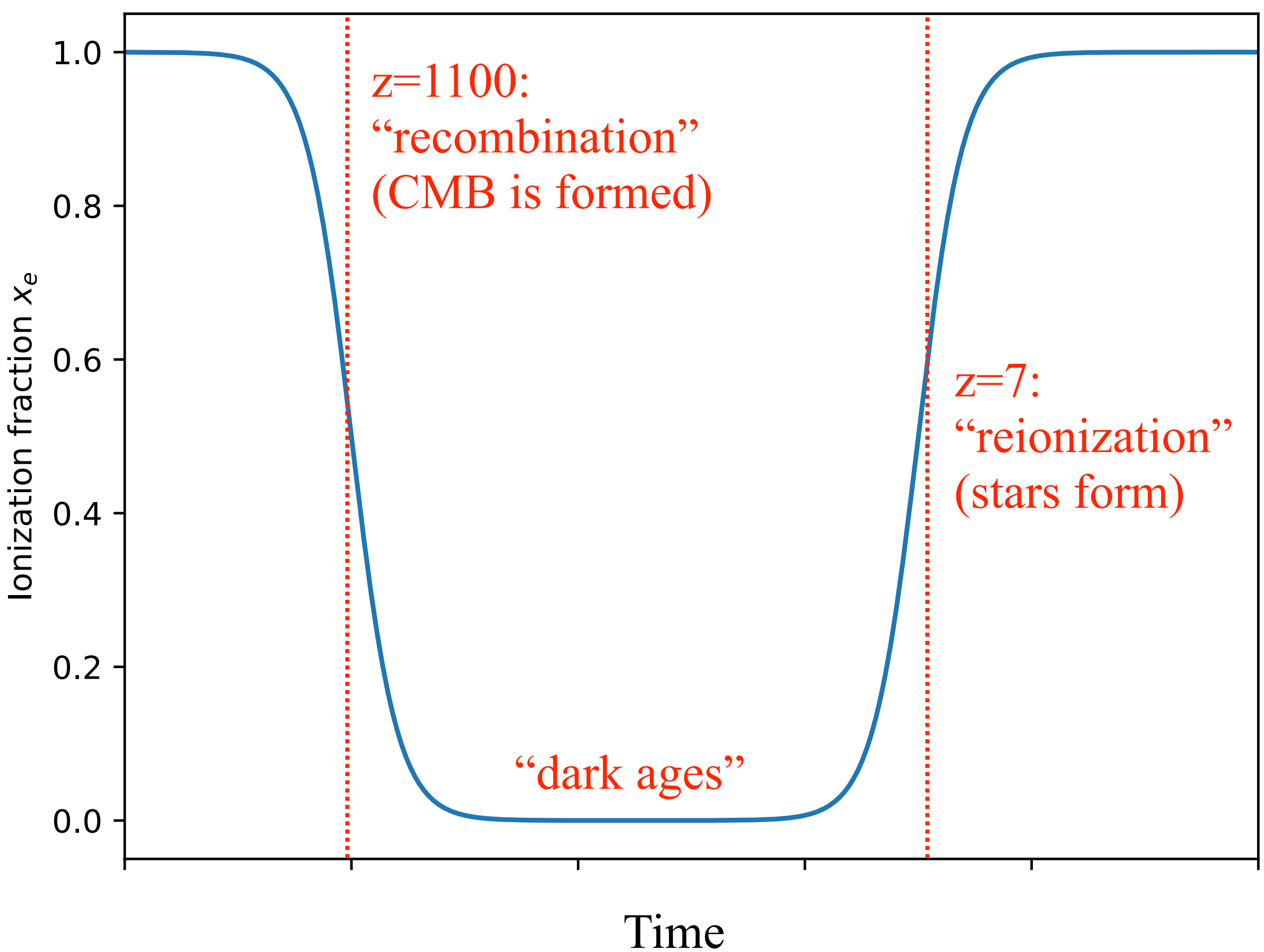
= probability that a random electron in the universe is ionized (rather than being part of an atom)



Ionization history of the universe

τ = CMB optical depth

= probability that a CMB photon emitted at z~1100 scatters from an electron at low redshift, before being observed at z=0.



Ionization history of the universe

τ = CMB optical depth

= probability that a CMB photon emitted at $z\sim1100$ scatters from an electron at low redshift, before being observed at $z=0$.

Astrophysical nuisance parameter: τ affects the CMB power spectrum.

When fitting cosmological parameters from the CMB, we need to include τ in the fit, and account for uncertainty in τ when assigning errors to other parameters.

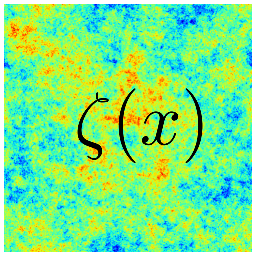
Standard model of cosmology:

- Background metric is FRW
- Expansion history is ΛCDM
- Initial perturbations are Gaussian random
- Initial perturbations are scalar adiabatic
- Power spectrum of initial perturbations is a power law: $(k^3/2\pi^2)P(k) = \Delta_\zeta^2(k/k_0)^{n_s-1}$

Six parameters:

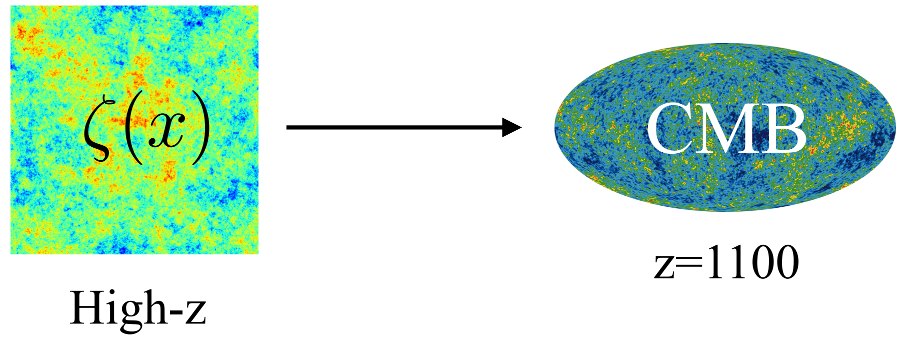
$\rho_\Lambda = (2.56 \pm 0.04) \times 10^{-47}$ GeV^4	Dark energy density (c.c.)
$\Omega_b = 0.0486 \pm 0.0007$	Baryonic(*) matter abundance
$\Omega_c = 0.267 \pm 0.009$	Cold dark matter abundance
$\Delta_\zeta^2 = (2.11 \pm 0.05) \times 10^{-9}$	Initial power spectrum amplitude
$n_s = 0.967 \pm 0.004$	Spectral index
$\tau = 0.058 \pm 0.012$	CMB optical depth

(*) "Baryons" = protons + neutrons + electrons(!)



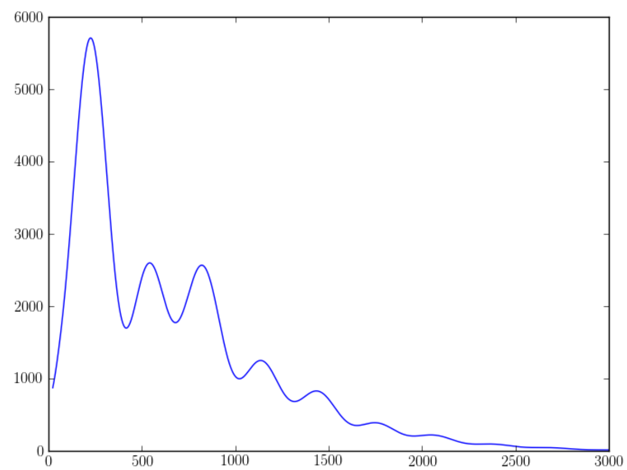
High-z

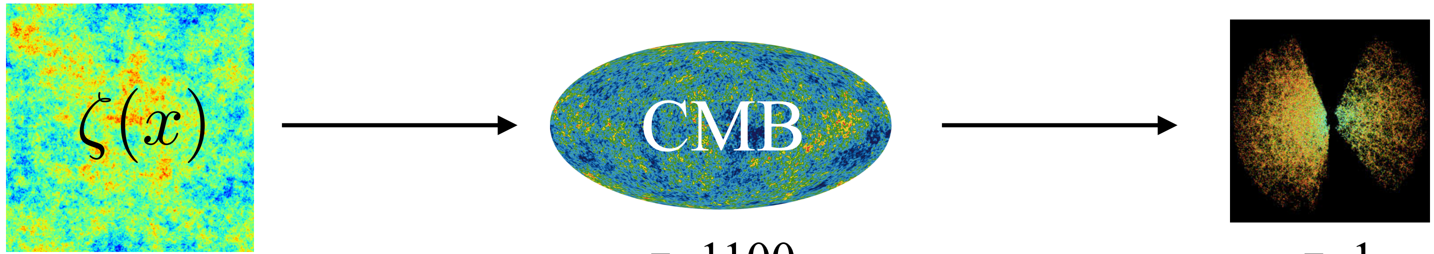
The standard cosmological model specifies the perturbations at very early times (high-z). They are fairly simple, and parameterized by a Gaussian random field $\zeta(x)$ with a featureless power spectrum.



The standard cosmological model specifies the perturbations at very early times (high-z). They are fairly simple, and parameterized by a Gaussian random field $\zeta(x)$ with a featureless power spectrum.

As time evolves, the perturbations become more complex. By the time the CMB is formed (z=1100), a lot of physics has been "imprinted" on the power spectrum.



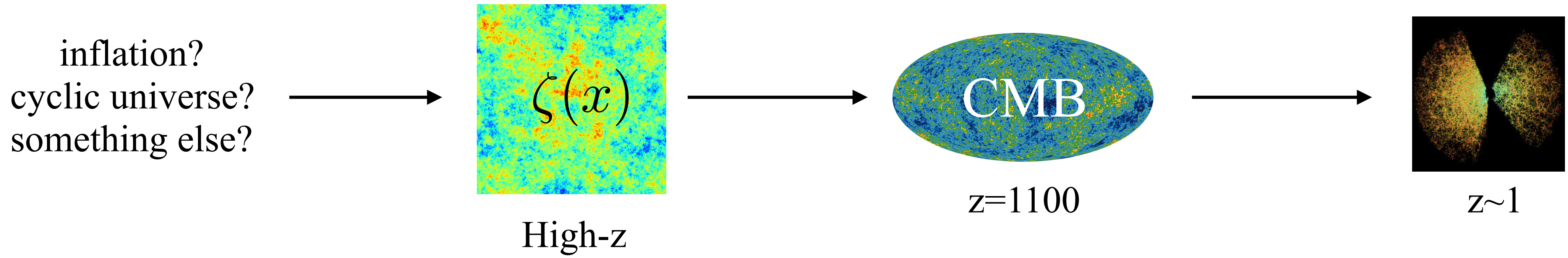


High-z z=1100 z~1

The standard cosmological model specifies the perturbations at very early times (high-z). They are fairly simple, and parameterized by a Gaussian random field $\zeta(x)$ with a featureless power spectrum.

As time evolves, the perturbations become more complex. By the time the CMB is formed (z=1100), a lot of physics has been "imprinted" on the power spectrum.

At late times (z~1), nonlinear effects are important and the perturbations are very non-Gaussian.

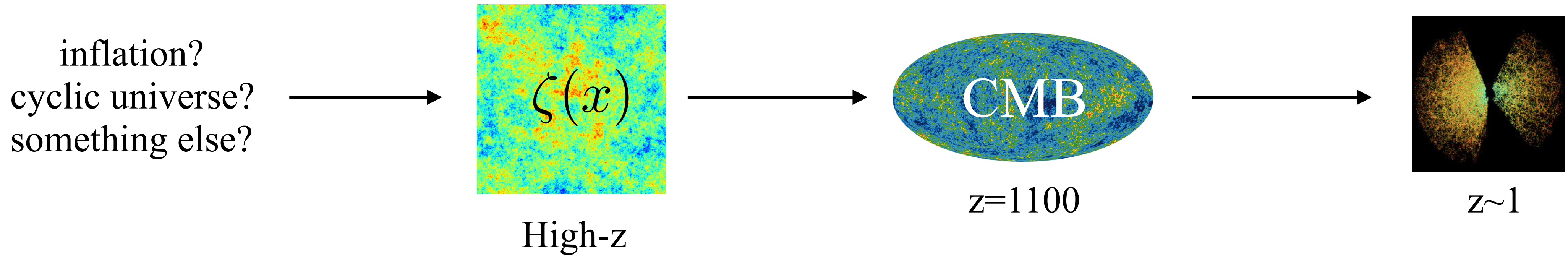


The standard cosmological model specifies the perturbations at very early times (high-z). They are fairly simple, and parameterized by a Gaussian random field ζ(x) with a featureless power spectrum.

As time evolves, the perturbations become more complex. By the time the CMB is formed (z=1100), a lot of physics has been "imprinted" on the power spectrum.

At late times (z~1), nonlinear effects are important and the perturbations are very non-Gaussian.

There are also models for the "early universe", a hypothetical phase preceding the radiation-dominated part of the expansion, which try to explain where the Gaussian field ζ came from.



In each of these three stages, different physics is important:

- Early universe: Quantum mechanics in expanding spacetime generates Gaussian perturbations from vacuum
- Formation of the CMB: Linear perturbation theory in a plasma with multiple components (dark matter, baryons, photons, neutrinos) + metric degrees of freedom
- Late times: Gravitational N-body physics. Messy astrophysics! (galaxy formation, star formation, …)

Fundamental physics →

Cosmological constant ρ_Λ
Baryon abundance Ω_b
Dark matter abundance Ω_c
Initial amplitude Δ_ζ^2
Spectral index n_s
CMB optical depth τ

→ Data analysis

Challenge for observers: which model fits the data?

~1930: Expanding universe
1965: Big bang (discovery of CMB)
~1970: Dark matter
1992: Gaussian, nearly scale-invariant perturbations (COBE)
1998: Cosmological constant
2006: Deviation from scale invariance ($n_s < 1$)

Fundamental physics →

Cosmological constant ρ_Λ
Baryon abundance Ω_b
Dark matter abundance Ω_c
Initial amplitude Δ_ζ^2
Spectral index n_s
CMB optical depth τ

→ Data analysis

Fundamental physics →

Cosmological constant ρ_Λ

Baryon abundance Ω_b

Dark matter abundance Ω_c

Initial amplitude Δ_ζ^2

Spectral index n_s

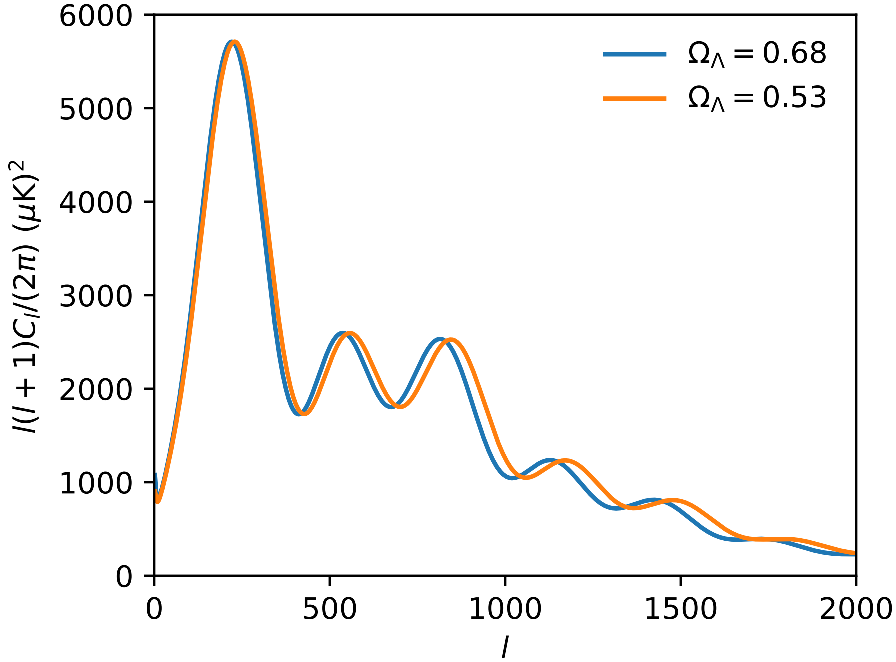
CMB optical depth τ

→ Data analysis

Challenge for theorists: explain this model at a fundamental level

- What is dark matter?
- Why is the cosmological constant so fine-tuned?
  (if late-time accelerated expansion is indeed a c.c.!)
- What physics is responsible for generating the initial Gaussian, nearly scale invariant fluctuations?

Cosmological observables (such as the CMB power spectrum) are sensitive to cosmological parameters, and can jointly constrain multiple parameters.



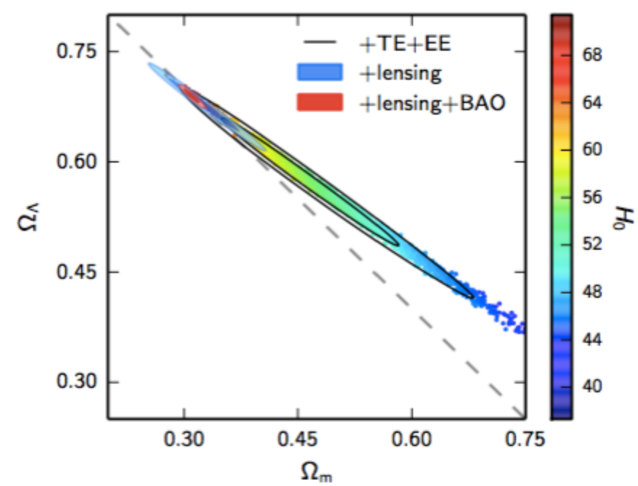
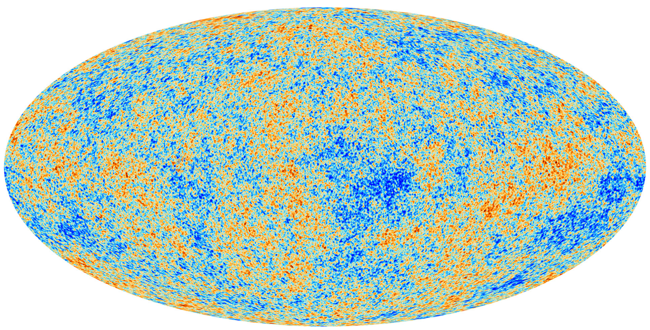


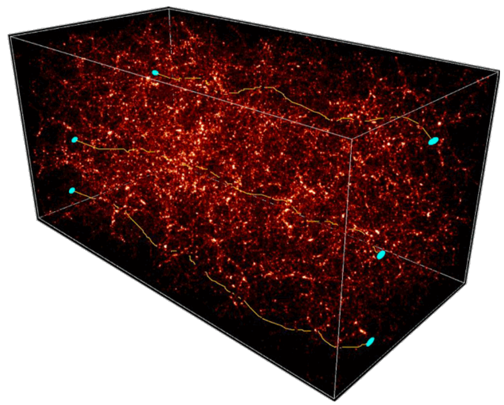
Fig. 26. Constraints in the Ω_m–Ω_Λ plane from the *Planck* TT+lowP data (samples; colour-coded by the value of H_0) and *Planck* TT,TE,EE+lowP (solid contours). The geometric degeneracy between Ω_m and Ω_Λ is partially broken because of the effect of lensing on the temperature and polarization power spectra. These limits are improved significantly by the inclusion of the *Planck* lensing reconstruction (blue contours) and BAO (solid red contours). The red contours tightly constrain the geometry of our Universe to be nearly flat.

Planck 2015

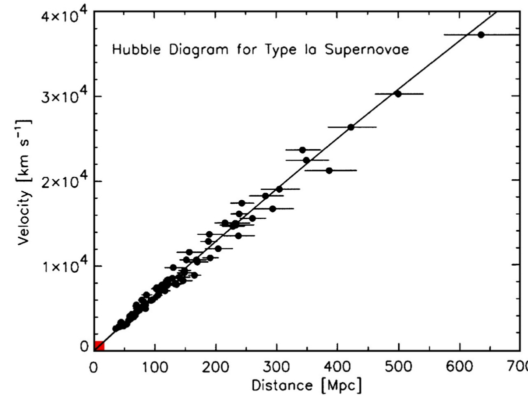
Variety of datasets, field is rapidly evolving:



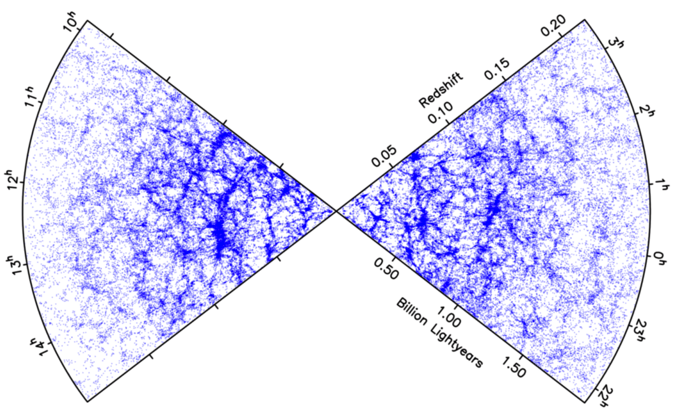
Cosmic microwave background (CMB)



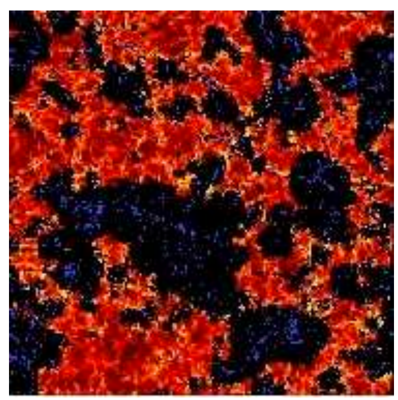
Gravitational lensing



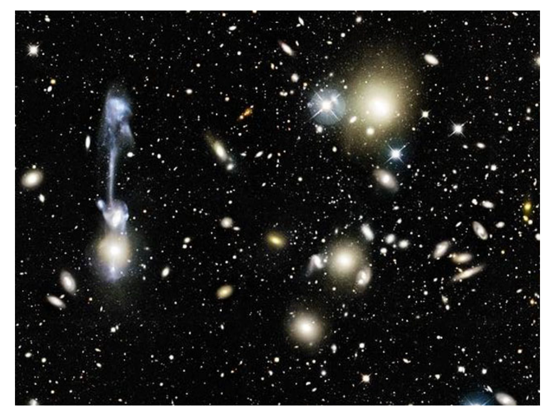
Type Ia supernovae



Galaxy clustering



21-cm intensity mapping



Galaxy cluster abundance

Cosmology is largely concerned with looking for extensions of the 6-parameter standard model.

- Non-Gaussian initial conditions
- Non-minimal neutrino mass
- Extra neutrino species or other light relics
- Interacting dark matter
- Nonzero spatial curvature
- Cosmological gravity waves

\+ many others!

The standard model includes ingredients which were originally surprises (dark matter, cosmological constant, quantum mechanically generated perturbations).

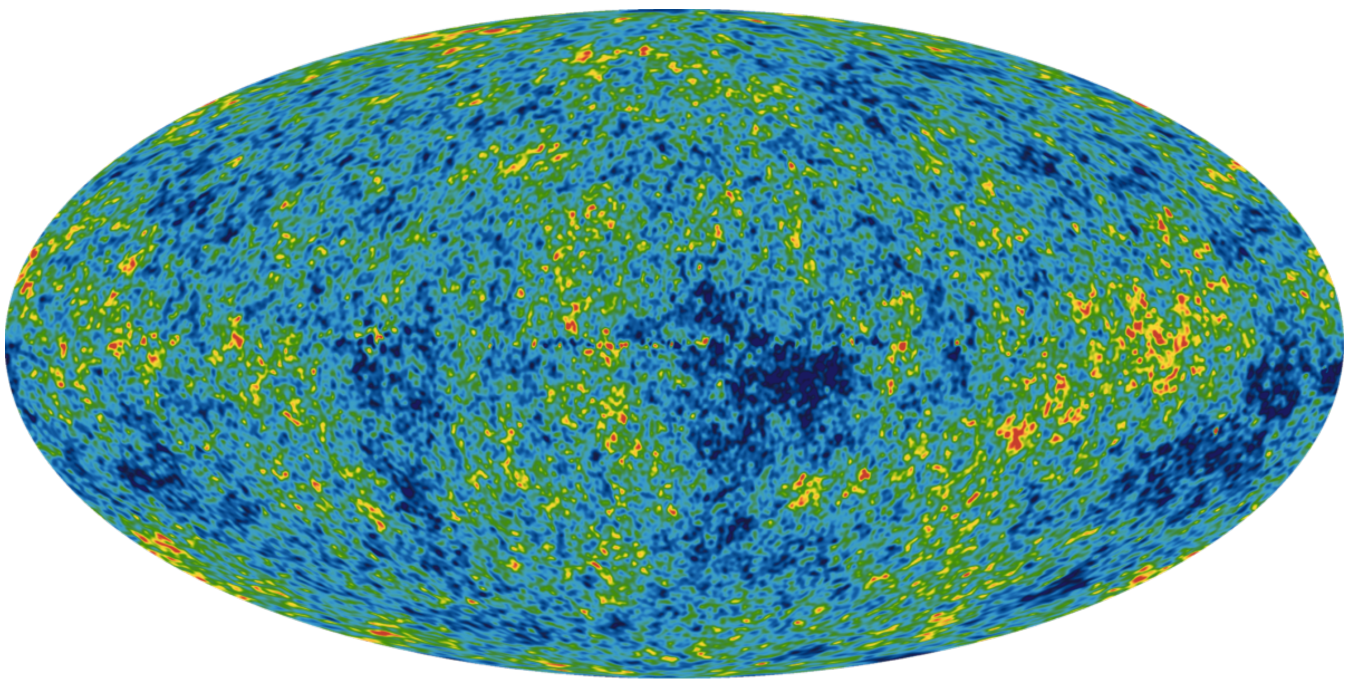
Will we find new surprises?

Part 2: random variables and fields

The standard model of cosmology is a probabilistic model.

For example, it can predict the probability of a given CMB realization occurring, but not the specific realization.

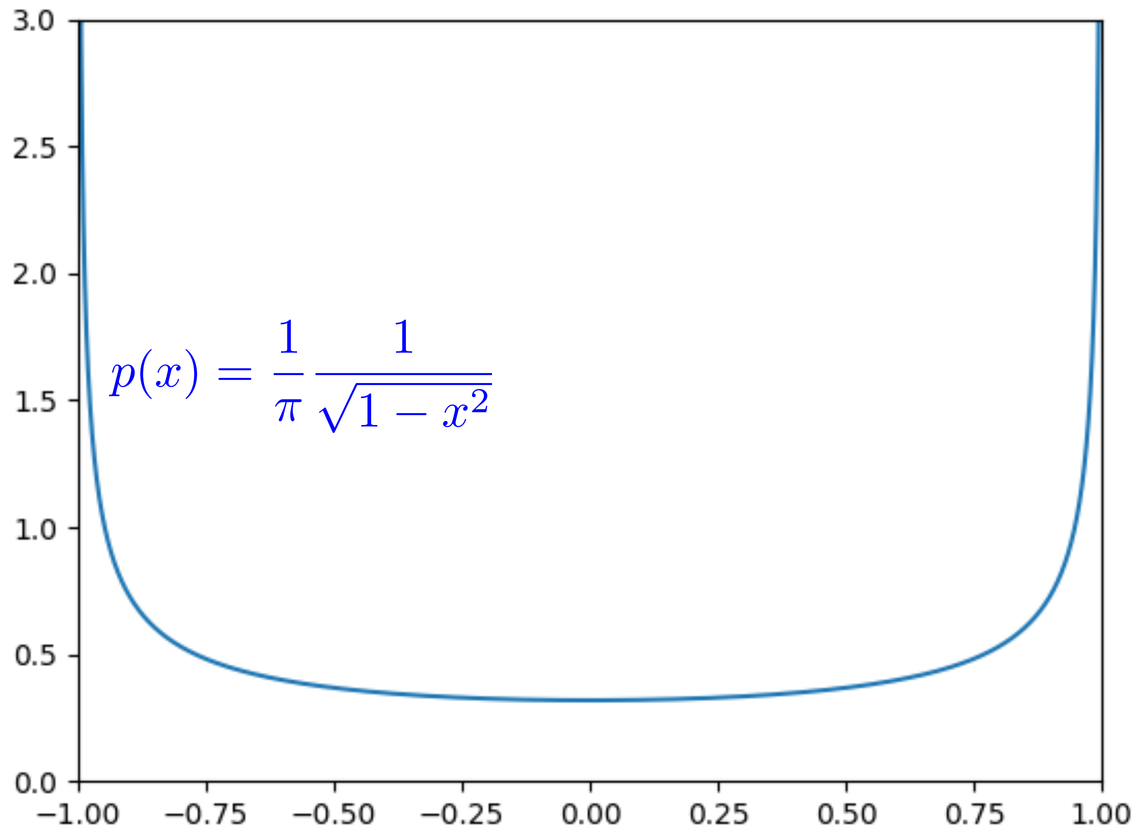
In this part of the lectures, we'll build up some machinery for working with random variables and fields.



Physicist's definition of a one-dimensional random variable X: anything with a probability distribution function (PDF) p(x).

The meaning of p(x) is "probability per unit x".

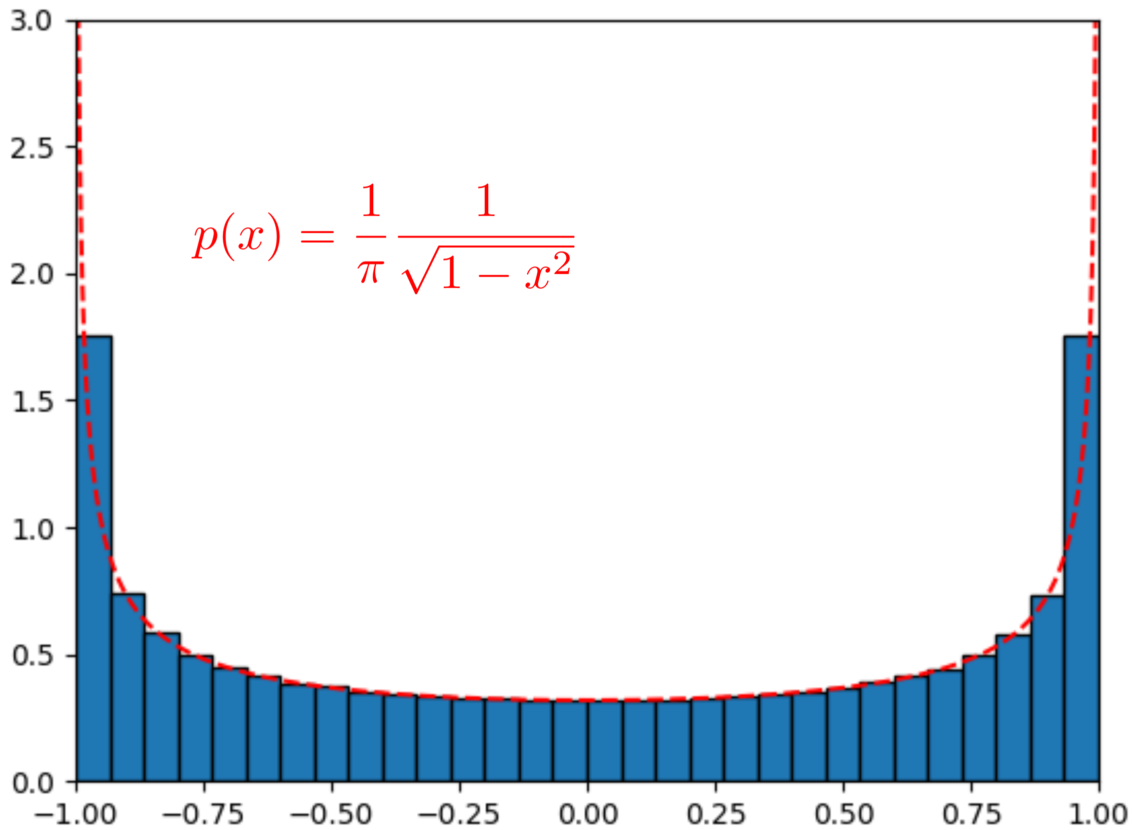
Here is an arbitrarily chosen example.



Histogram of 10^6 random samples in 30 bins, compared to the continuous PDF. The probability for the random variable X to be in bin [a,b] is:

$$\mathrm{Prob}\Big(a < X < b\Big) = \int_a^b dx\, p(x)$$

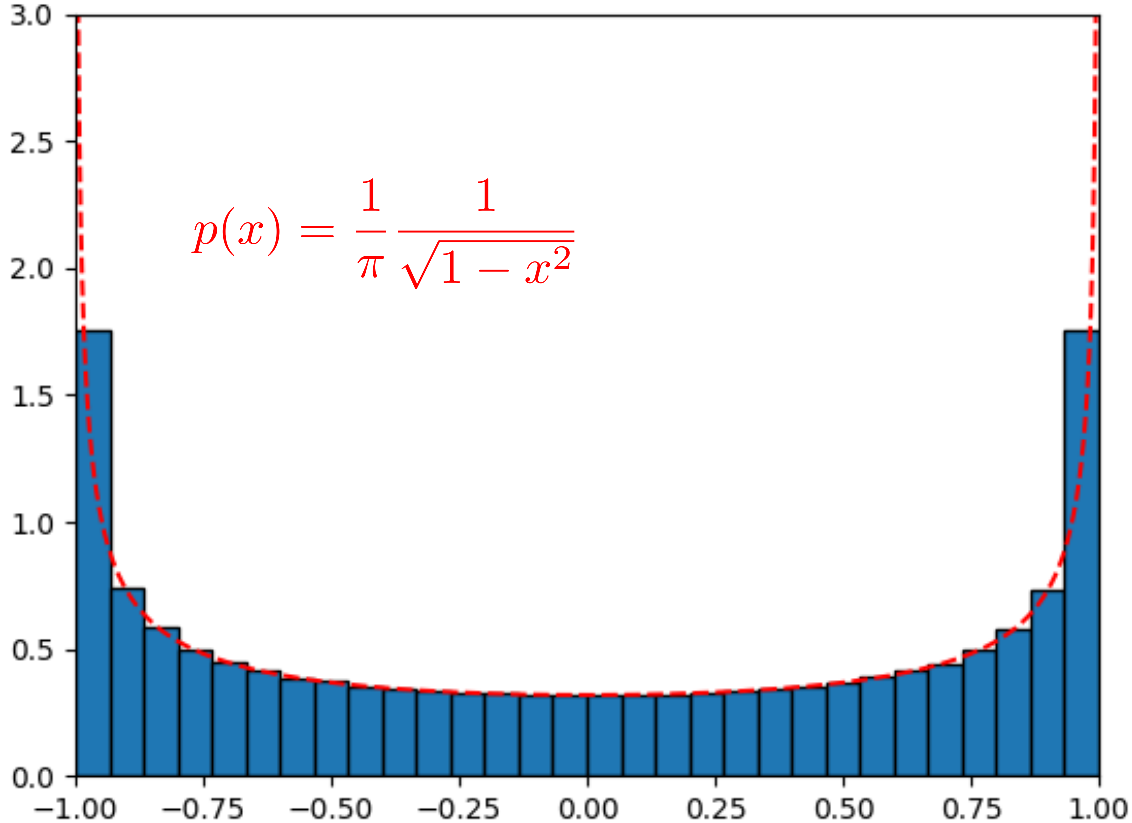
Note that the PDF must satisfy $\int_{-\infty}^{\infty} dx\, p(x) = 1$



The notation < . > denotes an expectation value over realizations of the random variable X. For example:

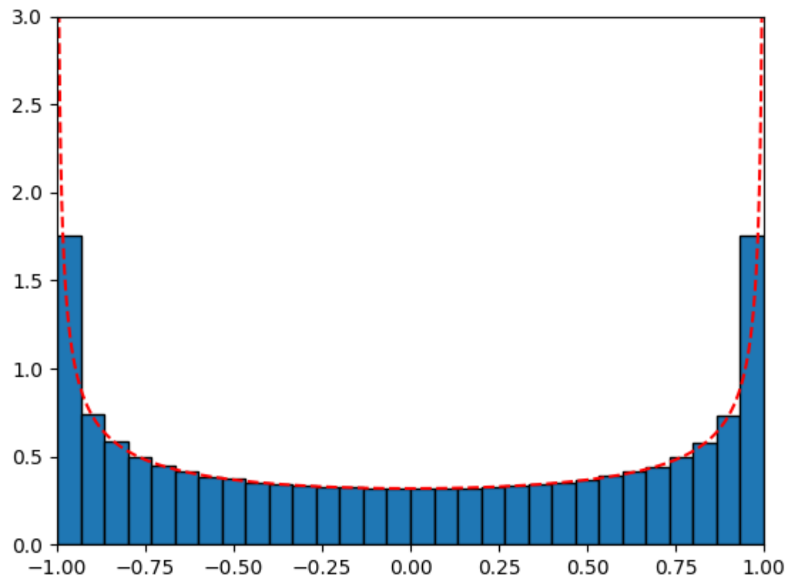
$$\langle X \rangle = \int_{-1}^{1} dx\, x\, p(x) = 0$$

$$\langle X^2 \rangle = \int_{-1}^{1} dx\, x^2\, p(x) = \frac{1}{2}$$



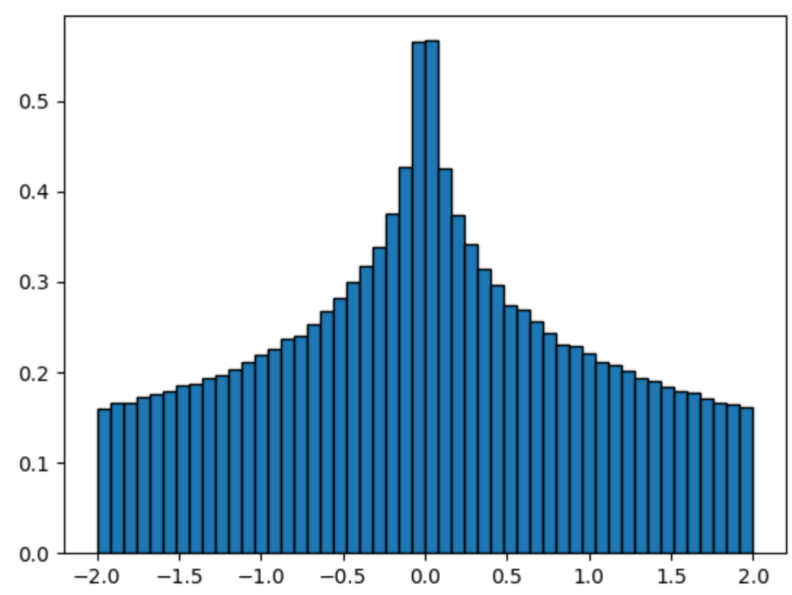
New random variables can be constructed from old ones.

For example, define $Y = (X_1 + X_2)$, where X_1, X_2 are independent random variables with the same PDF as before, $p(x) = \frac{1}{\pi\sqrt{1-x^2}}$



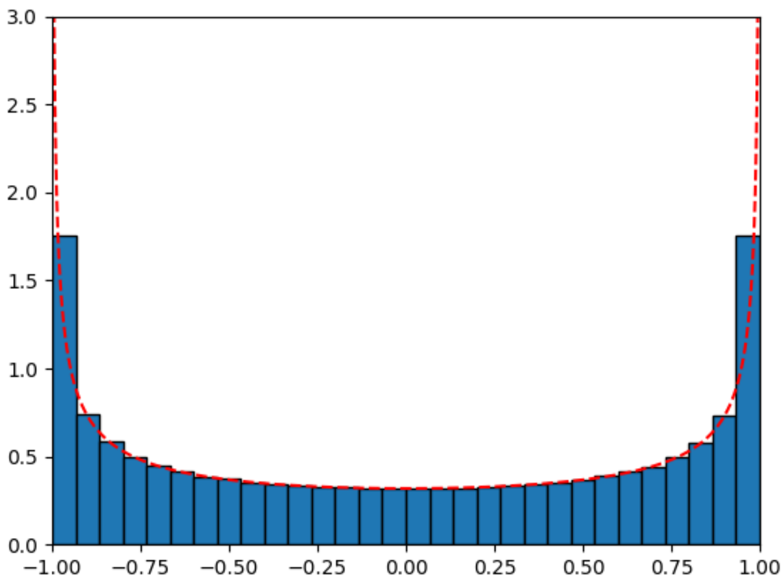
X

$\Rightarrow$



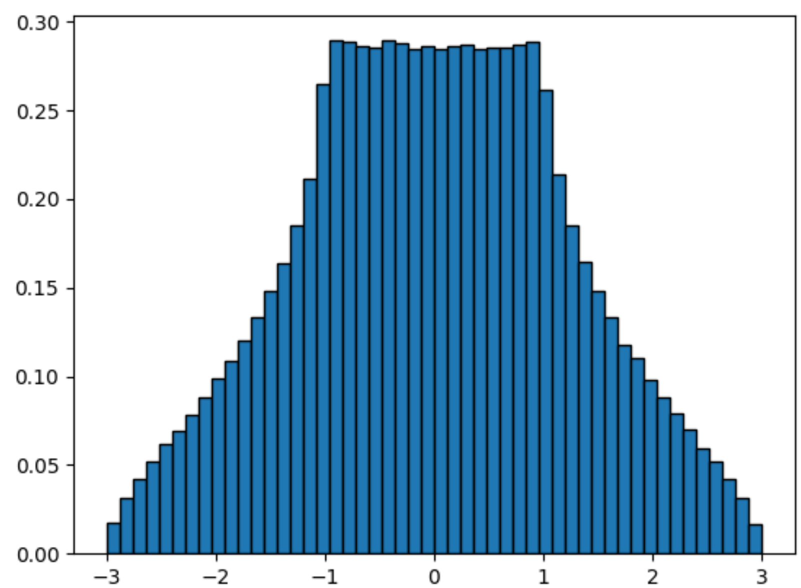
$Y = X_1 + X_2$

Three X's added together: $Y = X_1 + X_2 + X_3$



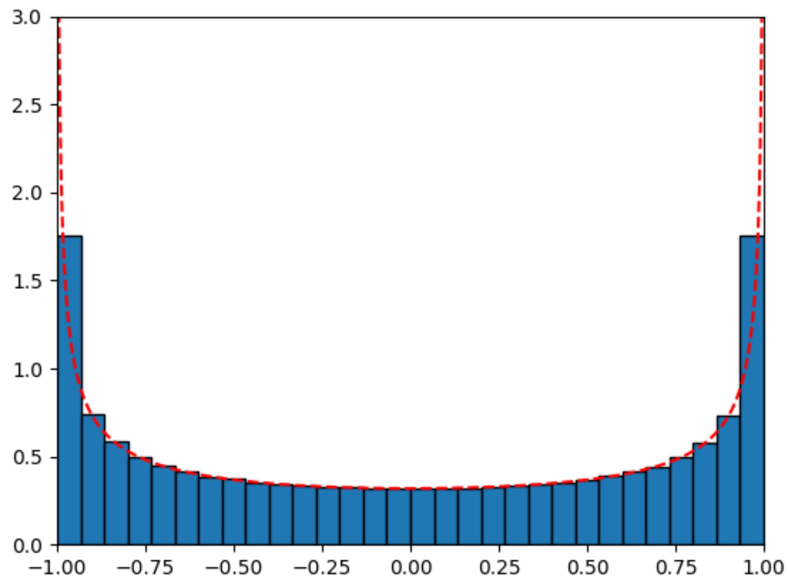
X

$\Rightarrow$



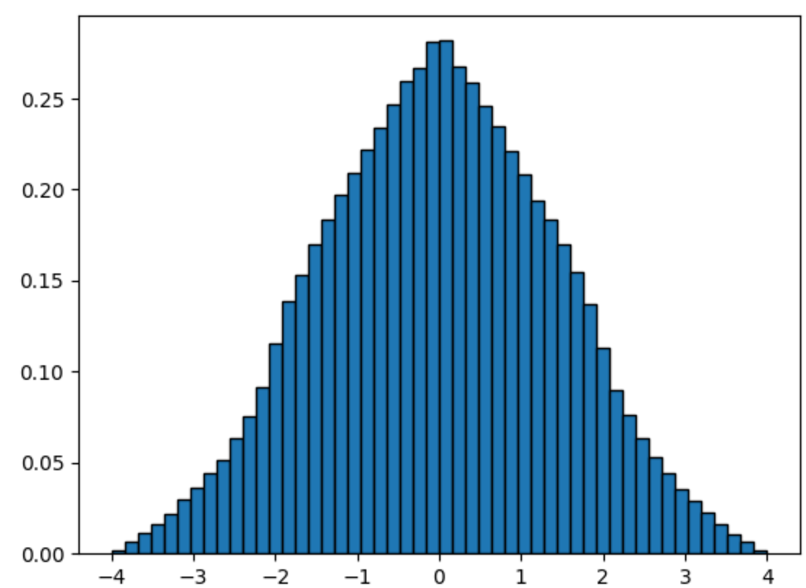
$Y = X_1 + X_2 + X_3$

Four X's added together: $Y = X_1 + X_2 + X_3 + X_4$



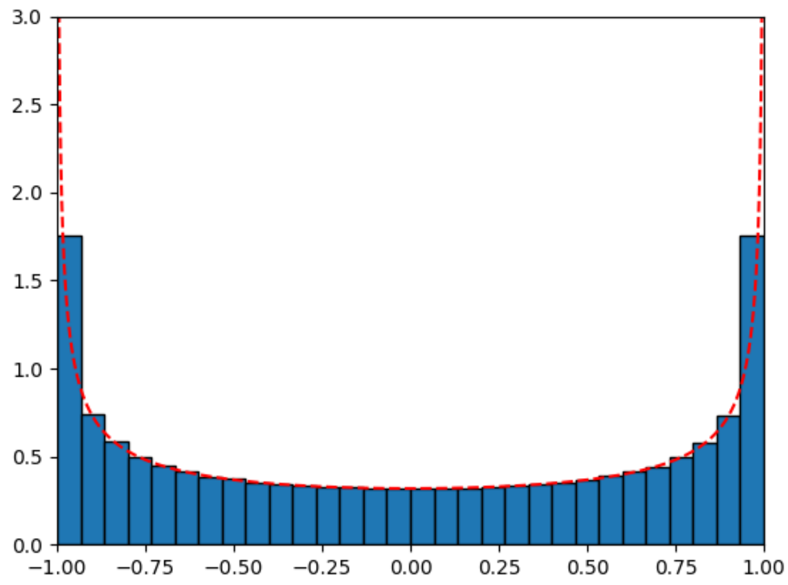
X

$\Rightarrow$



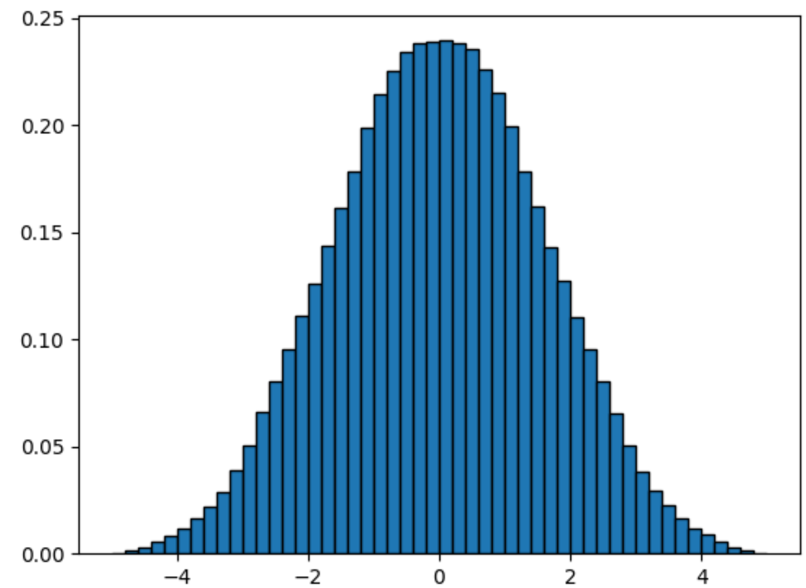
$Y = X_1 + X_2 + X_3 + X_4$

Five X's added together: $Y = X_1 + X_2 + X_3 + X_4 + X_5$



X

$\Rightarrow$

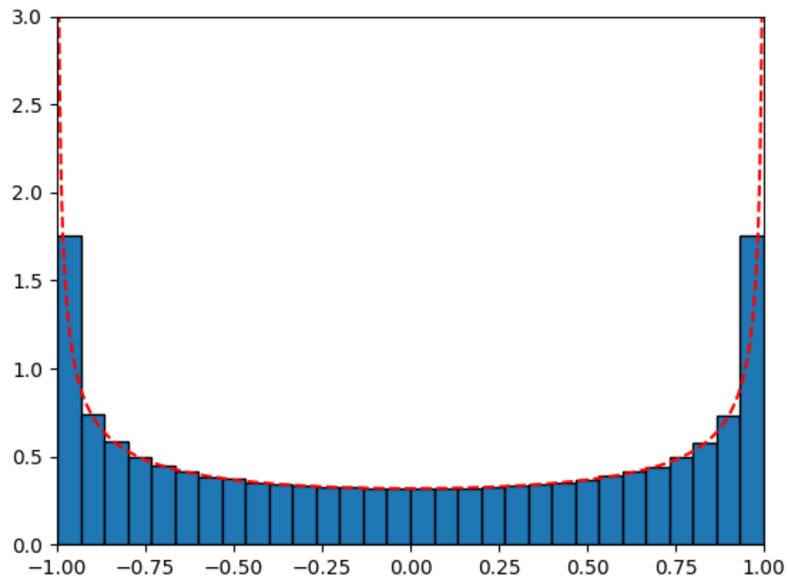


$Y = X_1 + X_2 + X_3 + X_4 + X_5$

Twenty X's added together: $Y = \sum_{i=1}^{20} X_i$

In the next few slides, we'll explain where the limiting PDF $p(x) = \frac{1}{\sqrt{20\pi}} e^{-x^2/20}$ comes from (including factors of 20, π).

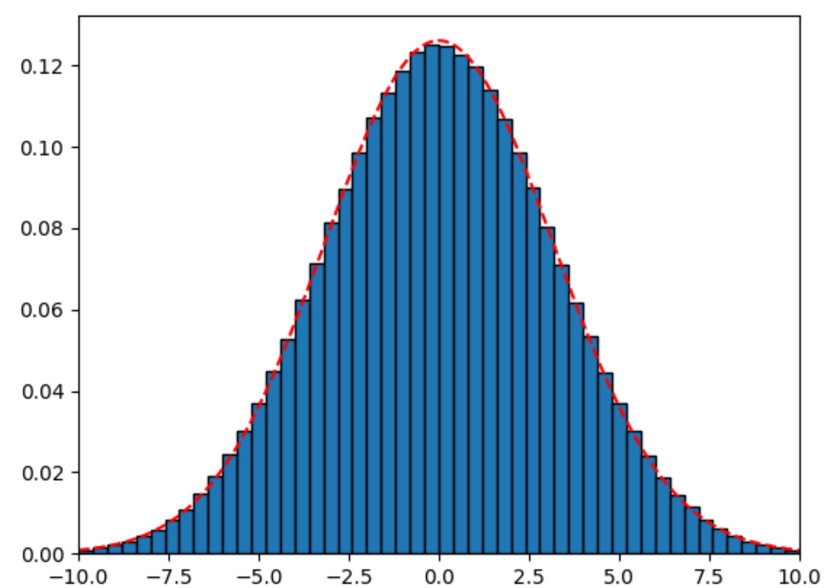
$$p(x) = \frac{1}{\pi} \frac{1}{\sqrt{1-x^2}}$$



X

$\Rightarrow$

$$p(x) = \frac{1}{\sqrt{20\pi}} e^{-x^2/20}$$



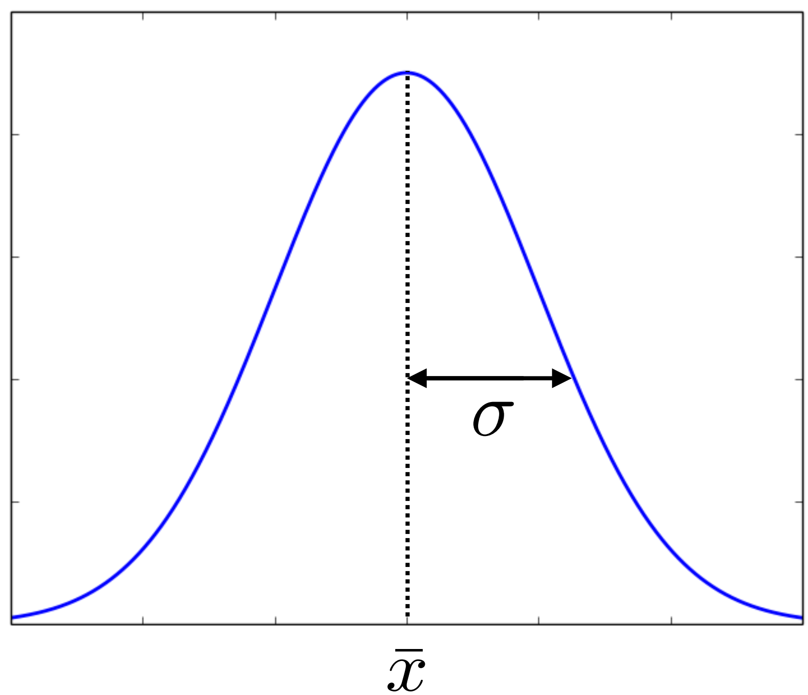
$$Y = \sum_{i=1}^{20} X_i$$

Central limit theorem: the sum of a large number of independent, identically distributed random variables has a PDF which is approximately Gaussian. (Proof omitted!)

The Gaussian PDF is defined by:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right)$$

and has two parameters: a mean $\bar{x}$ and a width σ.



Some definitions: the mean and variance of a random variable X are defined by:

$$\bar{X} = \langle X \rangle$$

[mean]

$$\begin{aligned} \mathrm{Var}(X) &= \langle X^2 \rangle - \langle X \rangle^2 \\ &= \langle (X - \bar{X})^2 \rangle \end{aligned}$$

[variance]

$\sqrt{\mathrm{Var}(X)}$ can be interpreted as the "typical" size of fluctuations around the mean.

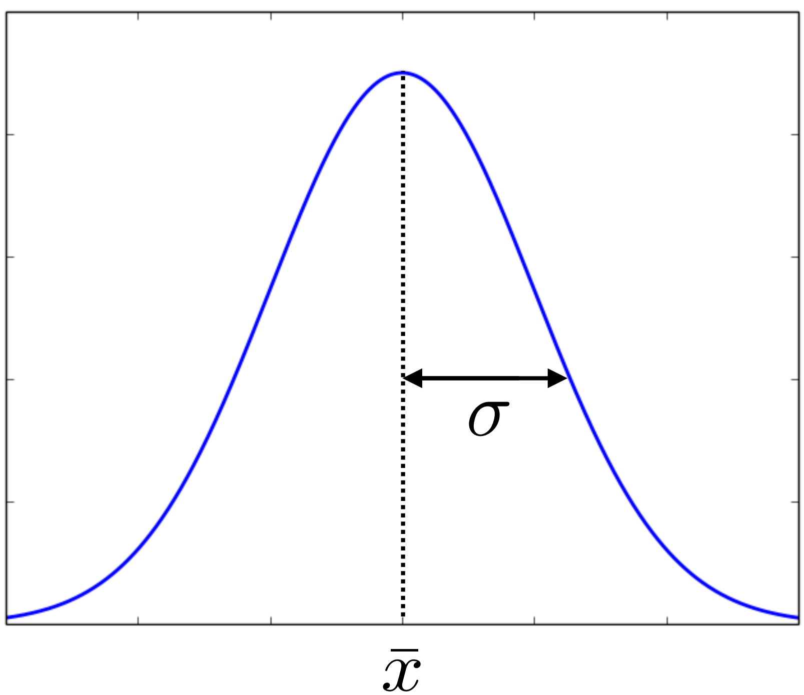
Example: For the Gaussian

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right)$$

a short calculation shows:

$$\text{Mean} = \int_{-\infty}^{\infty} dx\, p(x)\, x = \bar{x}$$

$$\text{Variance} = \int_{-\infty}^{\infty} dx\, p(x)\, (x^2 - \bar{x}^2) = \sigma^2$$

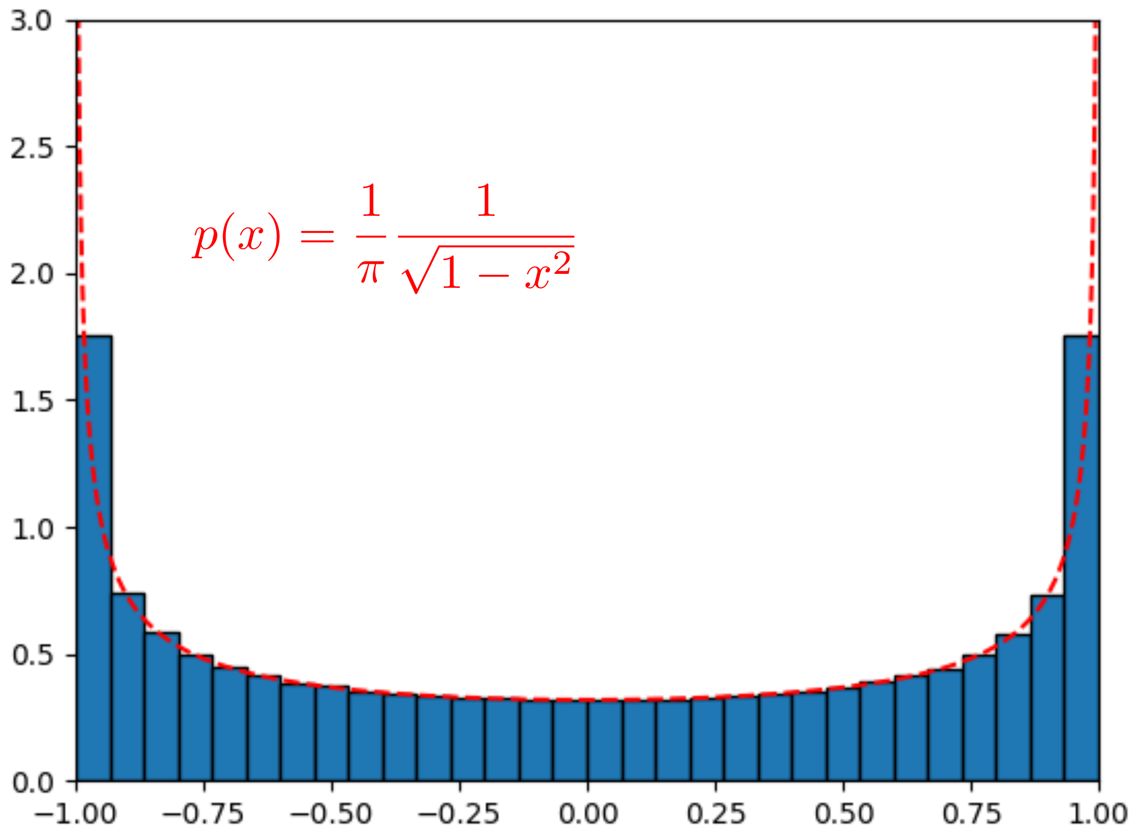


Example 2: for the PDF $p(x) = \frac{1}{\pi}\frac{1}{\sqrt{1-x^2}}$ considered previously,

$$\bar{X} = 0$$

$$\mathrm{Var}(X) = \langle (X - \bar{X})^2 \rangle = \frac{1}{2}$$

Next let's calculate mean and variance of $Y = \sum_{i=1}^{N} X_i$, where the X's are assumed to be independent samples.



Properties of expectation values:

$$\langle X \pm X' \rangle = \langle X \rangle \pm \langle X' \rangle$$

$$\langle cX \rangle = c\langle X \rangle$$ if c is a constant (not a random variable)

$$\langle X X' \rangle = \langle X \rangle \, \langle X' \rangle$$ if X, X' are independent random variables (not true in general!)

Properties of expectation values:

$$\langle X \pm X' \rangle = \langle X \rangle \pm \langle X' \rangle$$

$\langle cX \rangle = c\langle X \rangle$ if c is a constant (not a random variable)

$\langle XX' \rangle = \langle X \rangle \langle X' \rangle$ if X, X' are independent random variables (not true in general!)

Now we can calculate mean and variance of $Y = \sum_{i=1}^{N} X_i$

$$\bar{Y} = \sum_{i=1}^{N} \bar{X}_i = 0$$

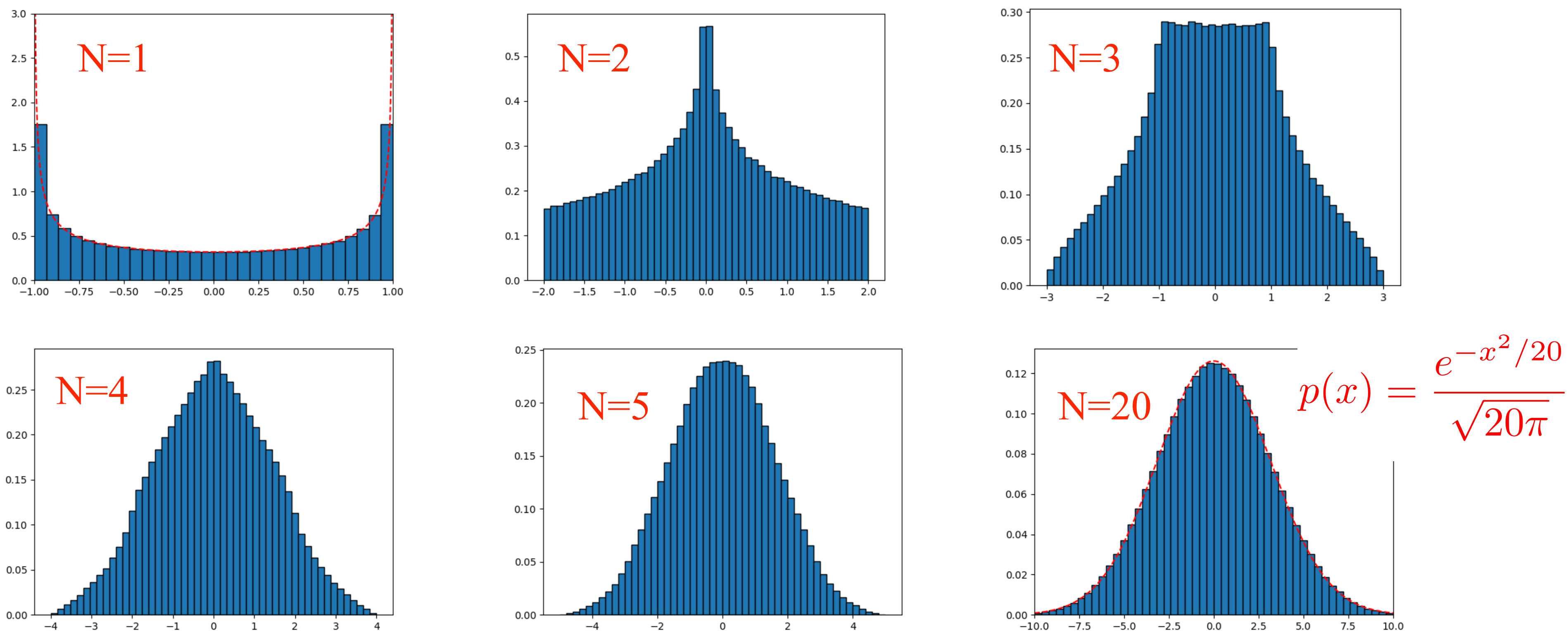
$$\begin{aligned} \mathrm{Var}(Y) &= \langle (Y - \bar{Y})^2 \rangle \\ &= \langle (\textstyle\sum_i X_i)^2 \rangle \\ &= \langle \textstyle\sum_i X_i^2 + \sum_{i \neq j} X_i X_j \rangle \\ &= \textstyle\sum_i \langle X_i^2 \rangle + \sum_{i \neq j} \langle X_i \rangle \langle X_j \rangle \\ &= N\left(\tfrac{1}{2}\right) \end{aligned}$$

This calculation gives the mean and variance of $Y = \sum_{i=1}^{N} X_i$:

$$\bar{Y} = 0 \qquad \mathrm{Var}(Y) = N/2 \qquad\qquad \text{(for all N)}$$

In general, the mean and variance do not determine the PDF p(x). However, for a Gaussian PDF they do!

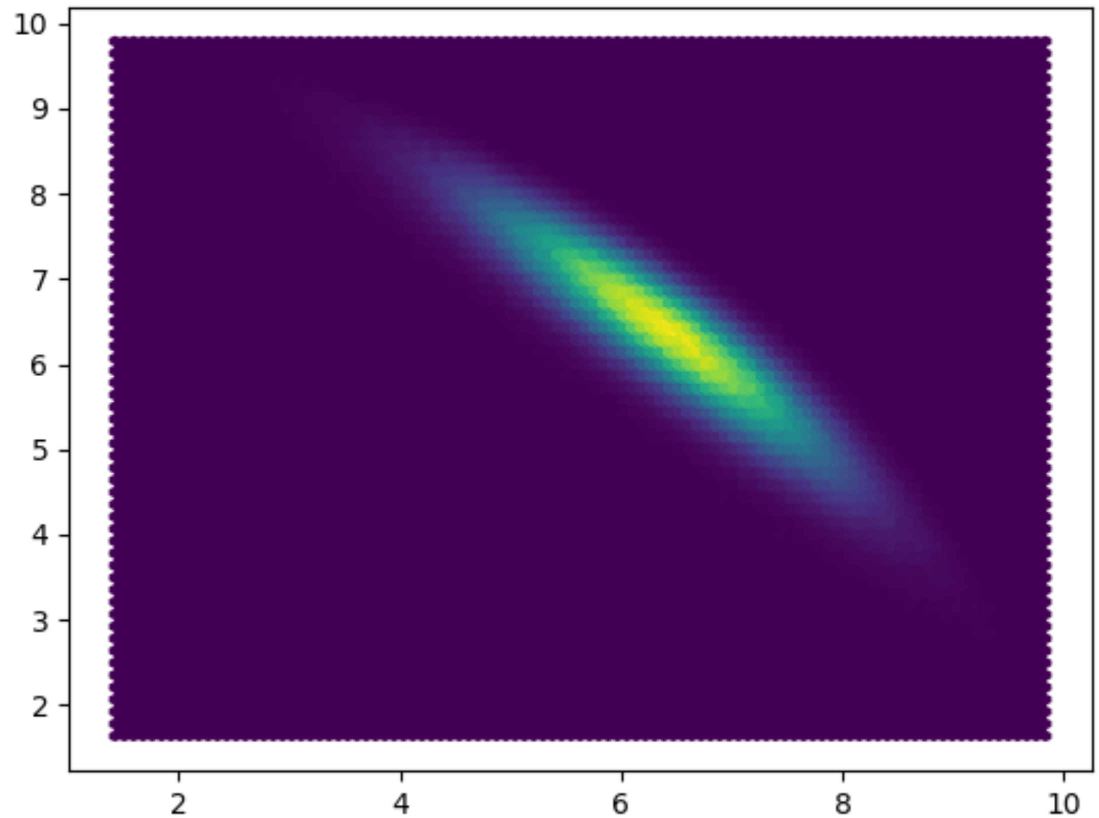
$$p(x) \approx \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\bar{x})^2/2\sigma^2} = \frac{1}{\sqrt{\pi N}} e^{-x^2/N} \qquad \text{(for N >> 1)}$$



Multivariate random variables: let's generalize to the case of N random variables (X_1, …, X_N) which are not assumed independent.

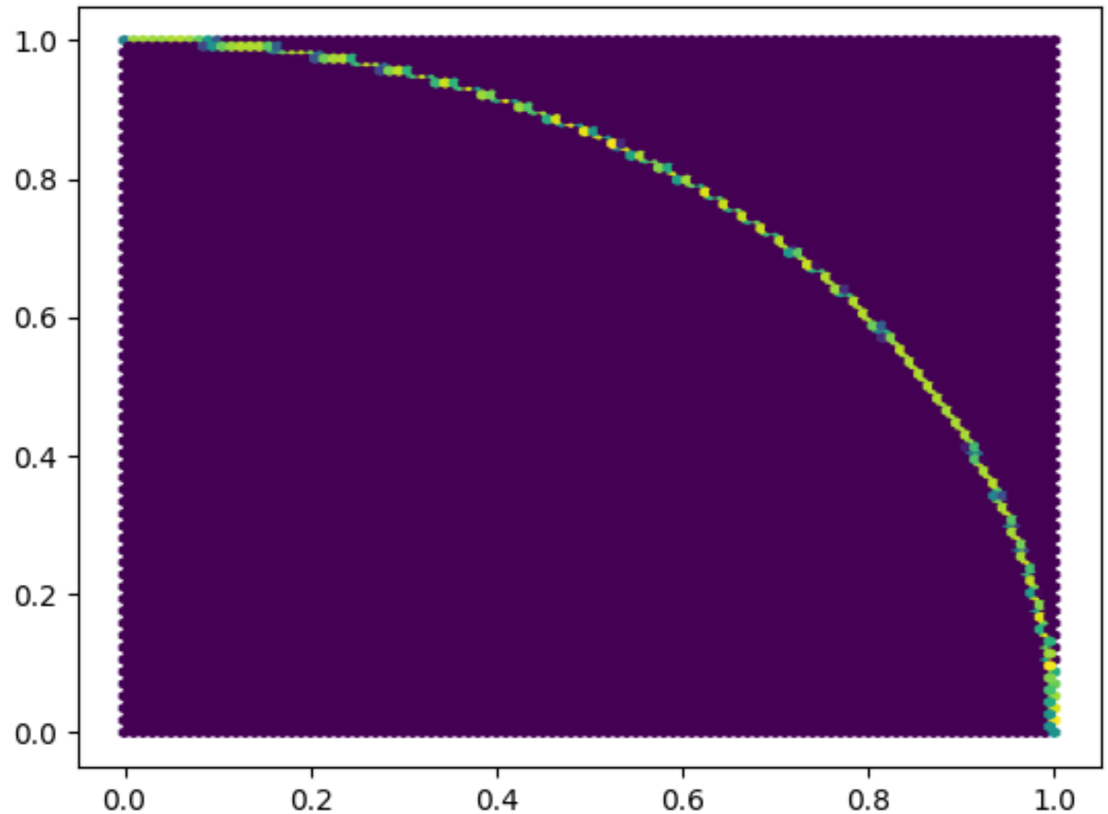
The PDF becomes a function of N variables $p(x_1,\ldots,x_N)$, and represents "probability per unit N-volume".

Example: a multivariate Gaussian (X_1, X_2) with a correlation between X_1 and X_2. (To be defined precisely in a few slides!)



Example: $p(x_1, x_2) = \begin{cases} \frac{2}{\pi}\delta(\sqrt{x_1^2 + x_2^2} - 1) & \text{if } x_1, x_2 \geq 0 \\ 0 & \text{otherwise} \end{cases}$

Just to show an extreme case where the variables x_1, x_2 are very non-independent!

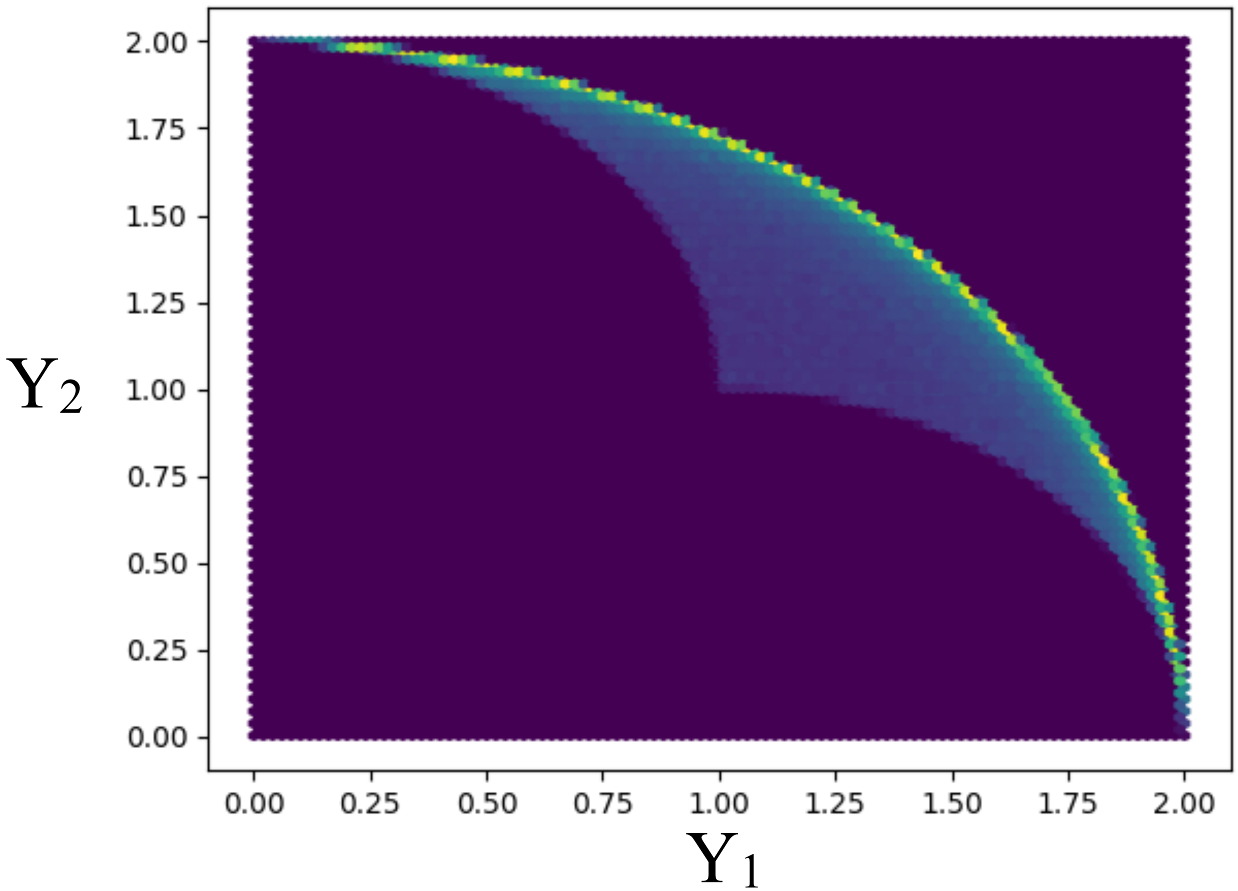


X_2

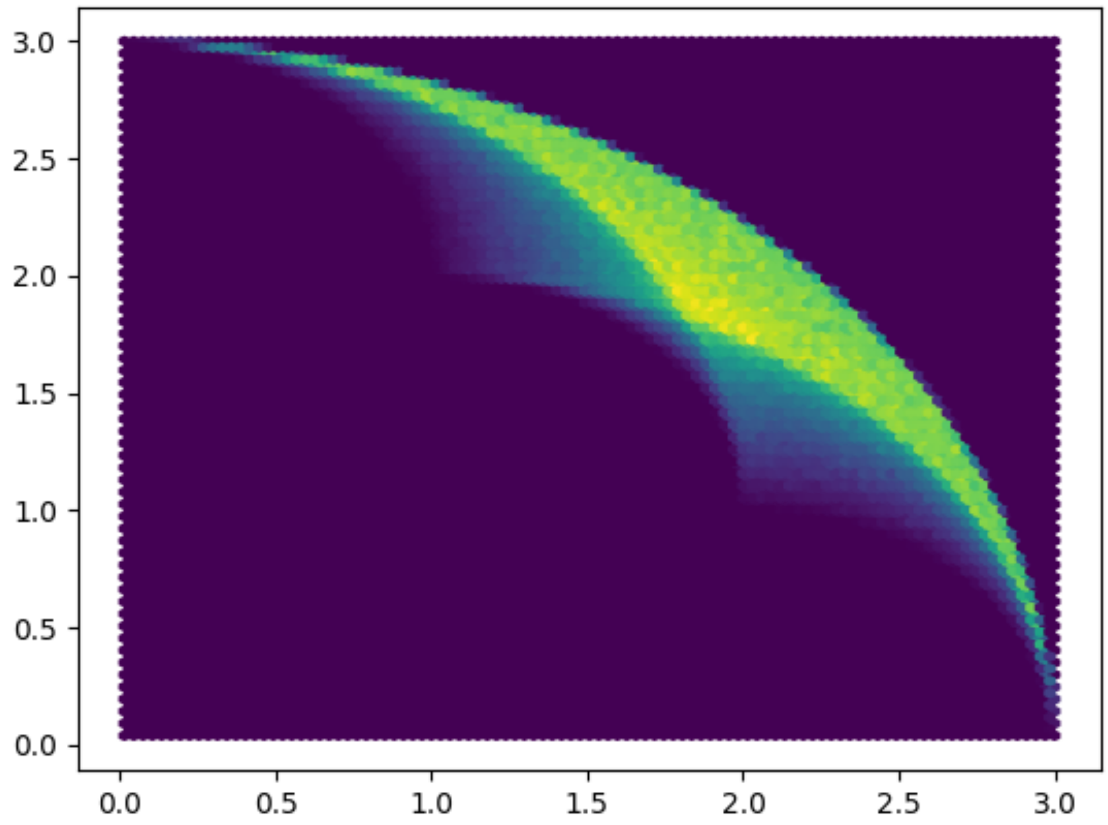
X_1

Does the central limit theorem still hold when the random variable is a vector X_i? (In this case, a two-component vector)

Two X's: $Y_i = X_i^{(1)} + X_i^{(2)}$



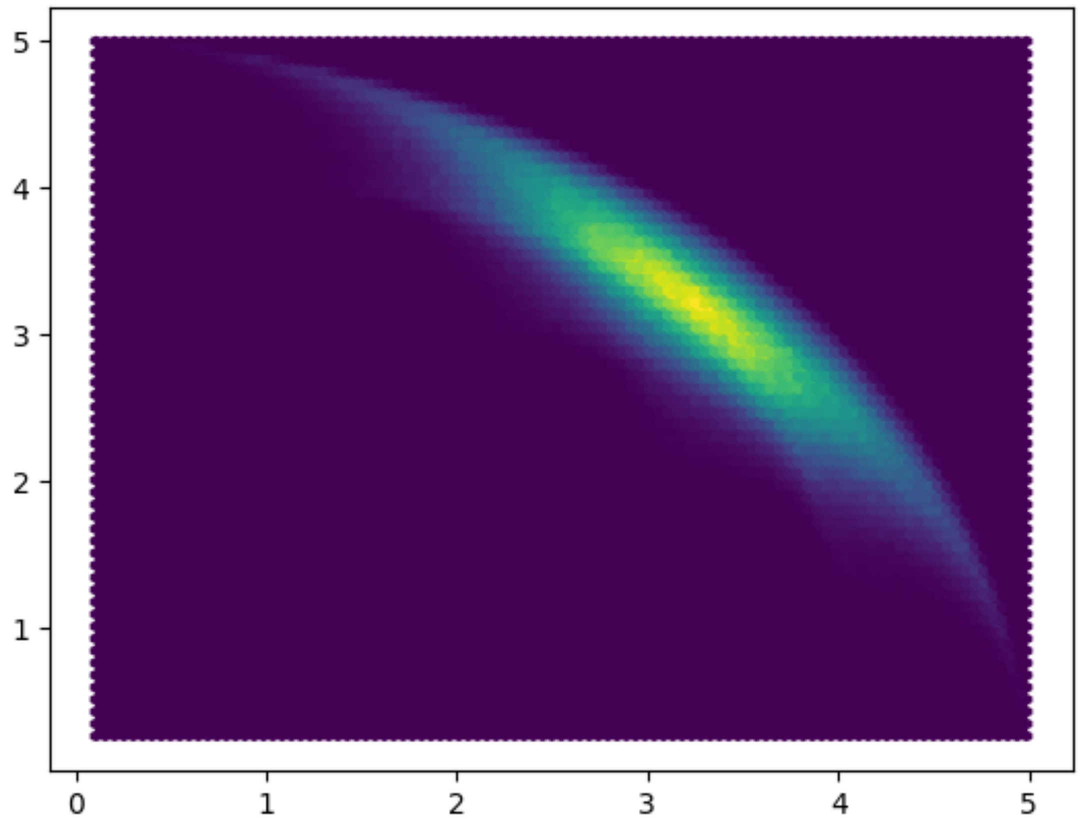
Three X's: $Y_i = X_i^{(1)} + X_i^{(2)} + X_i^{(3)}$



Y_2

Y_1

Five X's: $Y_i = \sum_{j=1}^{5} X_i^{(j)}$



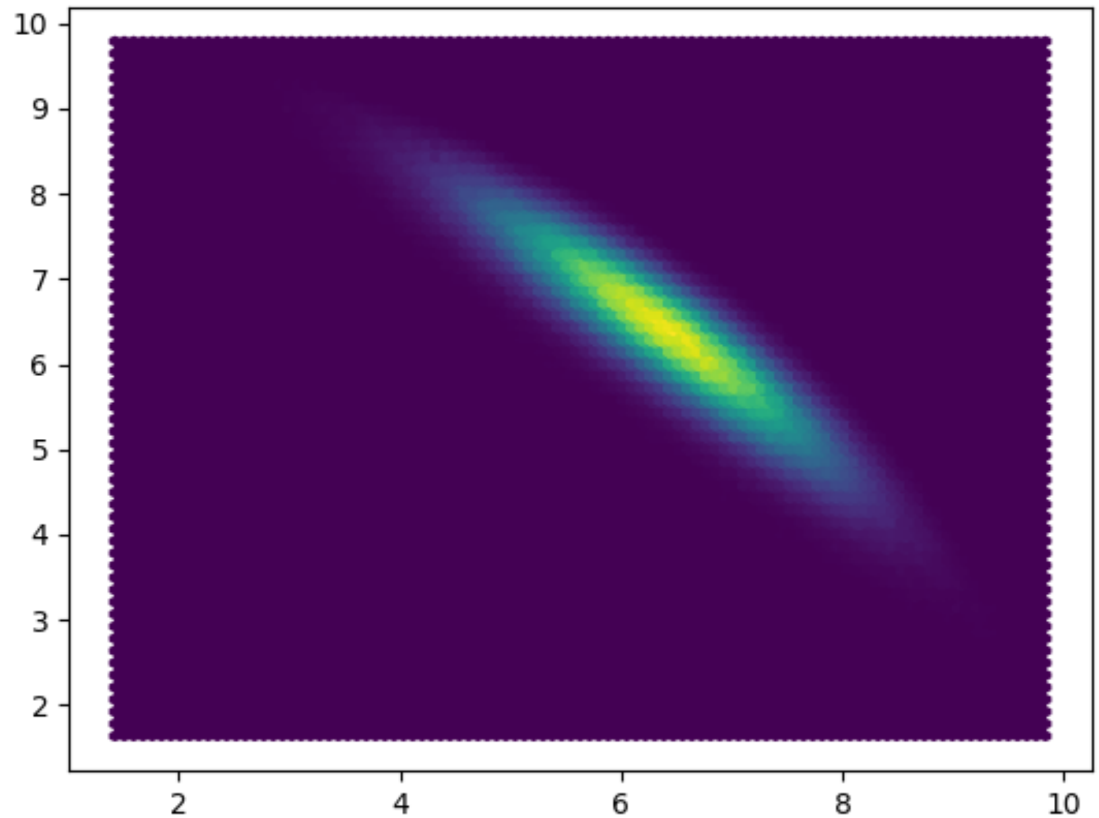
Y_2

Y_1

Ten X's: $Y_i = \sum_{j=1}^{10} X_i^{(j)}$

The distribution has become a multivariate Gaussian.

In two variables, the multivariate Gaussian has five parameters: two "means", and three parameters describing the size and orientation.



Y_2

Y_1

In N variables, the mean becomes an N-component vector

$$\bar{X}_i = \langle X_i \rangle$$

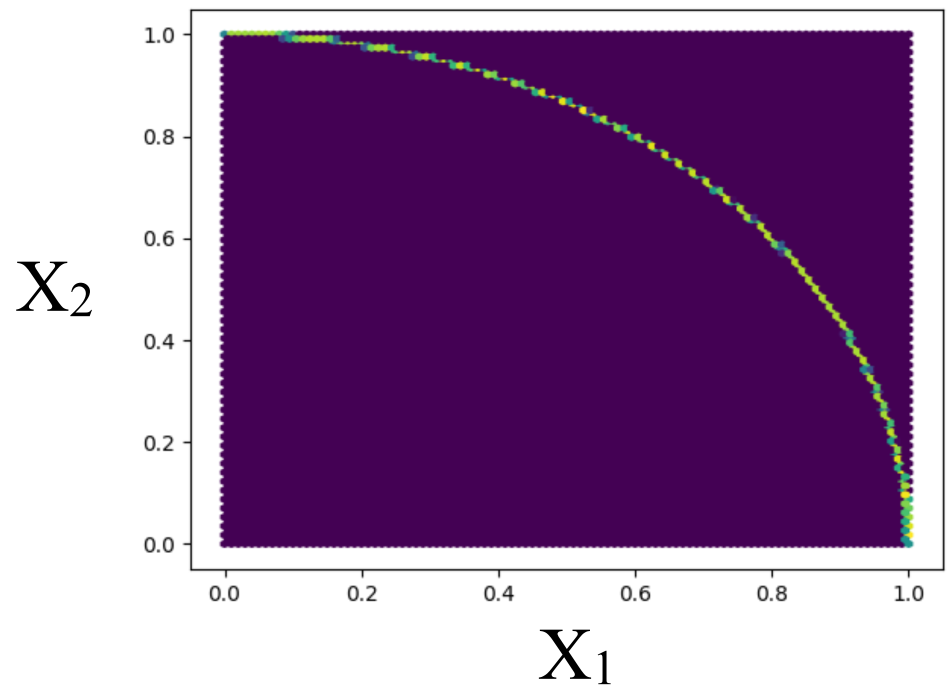
The variance generalizes to an N-by-N covariance matrix:

$$\begin{aligned}\mathrm{Cov}(X_i, X_j) &= \langle X_i X_j \rangle - \langle X_i \rangle \langle X_j \rangle \\ &= \langle (X_i - \bar{X}_i)(X_j - \bar{X}_j) \rangle\end{aligned}$$

In our example, a short calculation gives the mean and covariance:

$$\begin{pmatrix} \bar{X}_1 \\ \bar{X}_2 \end{pmatrix} = \begin{pmatrix} 0.64 \\ 0.64 \end{pmatrix}$$

$$\begin{pmatrix} C_{11} & C_{12} \\ C_{12} & C_{22} \end{pmatrix} = \begin{pmatrix} 0.095 & -0.087 \\ -0.087 & 0.095 \end{pmatrix}$$



Now we can give the definition of a multivariate Gaussian PDF:

$$p(x_1, \cdots, x_N) = \frac{1}{\mathrm{Det}(2\pi C)^{1/2}} \exp\left(-\frac{1}{2}(x_i - \bar{x}_i) C^{-1}_{ij} (x_j - \bar{x}_j) \right)$$

The PDF of a multivariate Gaussian random variable is determined by its mean $\bar{X}_i$ and covariance matrix $C_{ij} = \mathrm{Cov}(X_i, X_j)$

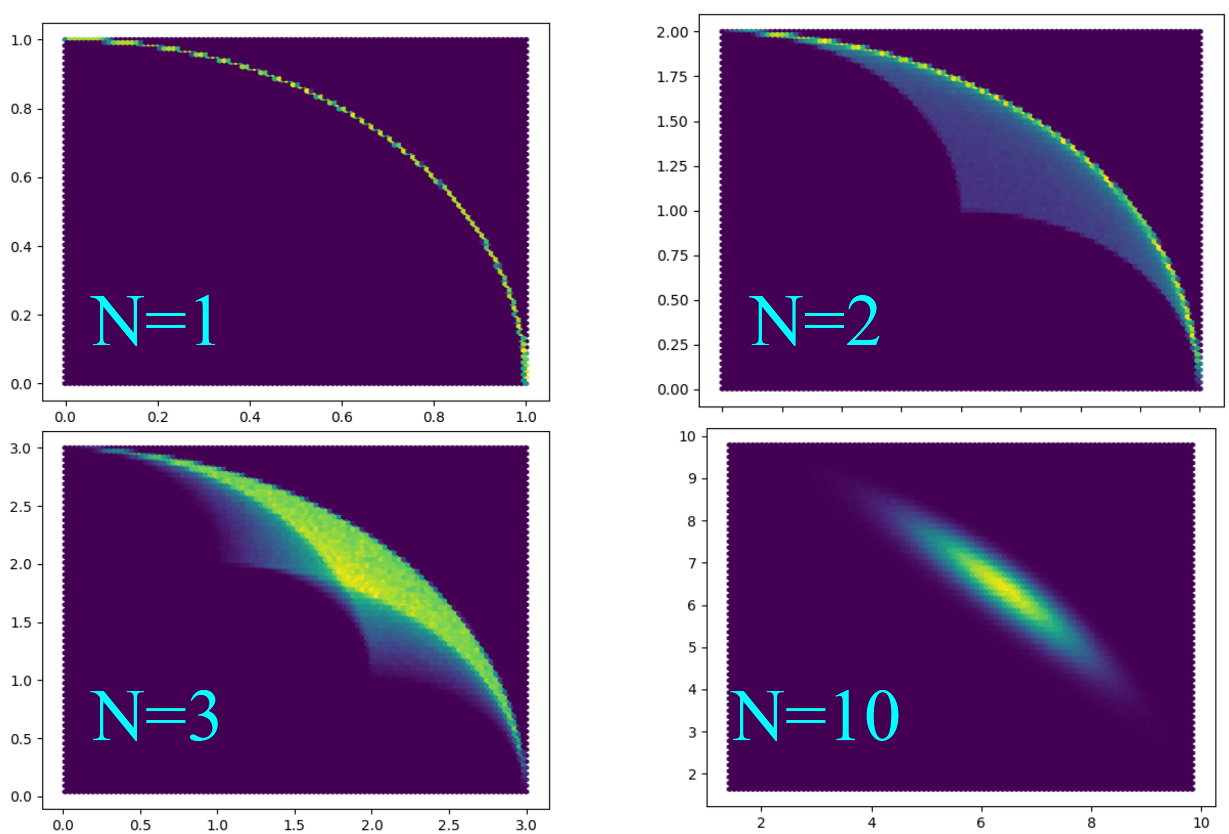
In cosmology, we are usually interested in Gaussian random variables. Therefore, it suffices to keep track of the mean (a vector) and the covariance (a matrix).

In this example:

$$\begin{pmatrix} \bar{X}_1 \\ \bar{X}_2 \end{pmatrix} = N \begin{pmatrix} 0.64 \\ 0.64 \end{pmatrix} \qquad \begin{pmatrix} C_{11} & C_{12} \\ C_{12} & C_{22} \end{pmatrix} = N \begin{pmatrix} 0.095 & -0.087 \\ -0.087 & 0.095 \end{pmatrix}$$

In the large-N limit, these determine the PDF (central limit theorem):

$$p(x_1, x_2) \approx \frac{1}{\mathrm{Det}(2\pi C)^{1/2}} \exp\left(-\frac{1}{2}(x_i - \bar{x}_i) C_{ij}^{-1} (x_j - \bar{x}_j) \right) \qquad (\mathrm{N} >> 1)$$



$$
\begin{aligned}
C_{ij} &= \langle X_i X_j \rangle - \langle X_i \rangle \langle X_j \rangle \\
&= \langle (X_i - \bar{X}_i)(X_j - \bar{X}_j) \rangle
\end{aligned}
$$

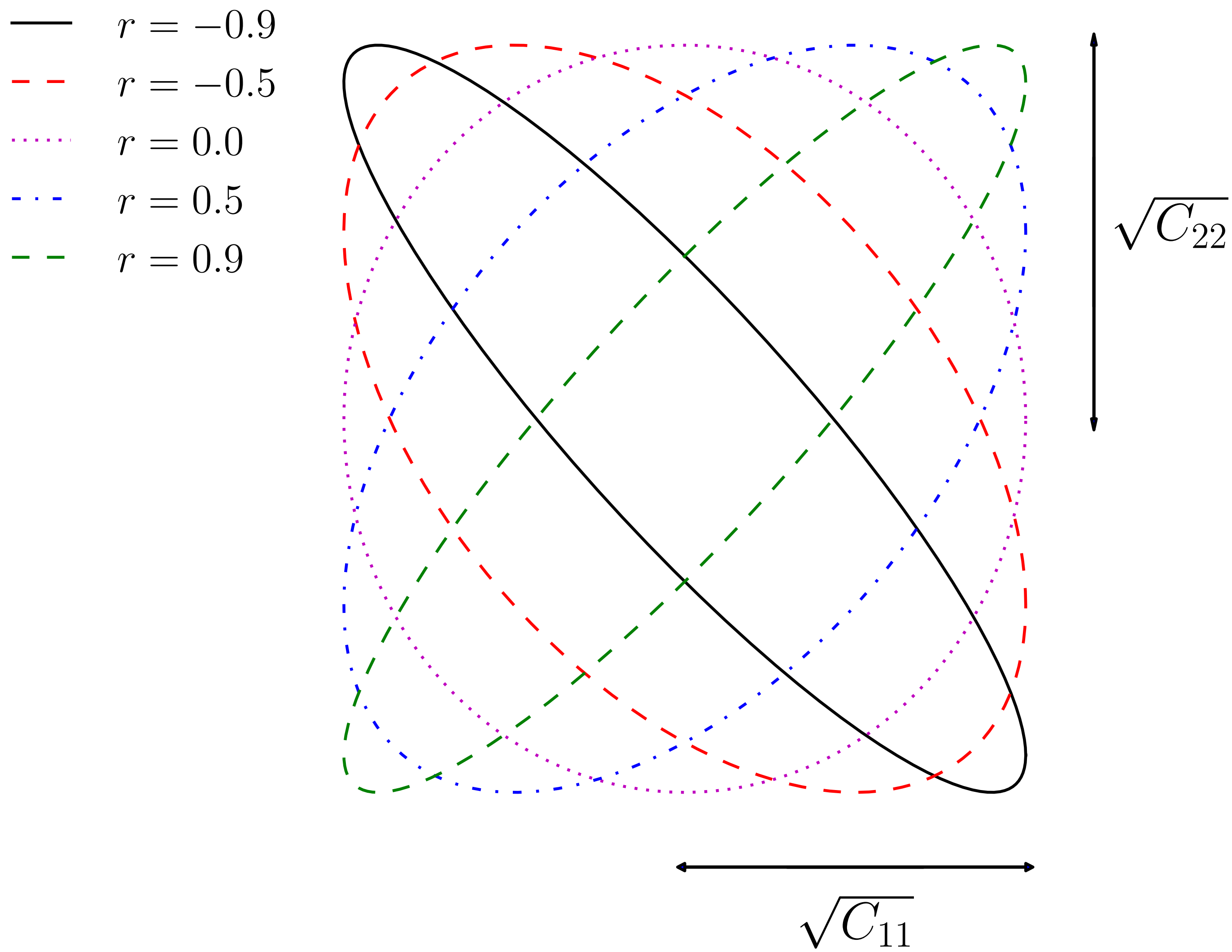
Diagonal elements C_{ii} of the covariance matrix are variances.
$C_{ii}^{1/2} \sim$ characteristic size of fluctuations in X_i around its mean.

Off-diagonals C_{ij} quantify the level of correlation between random variables X_i, X_j. The correlation coefficient

$$
\mathrm{Corr}(X_i, X_j) = \frac{C_{ij}}{\sqrt{C_{ii} C_{jj}}}
$$

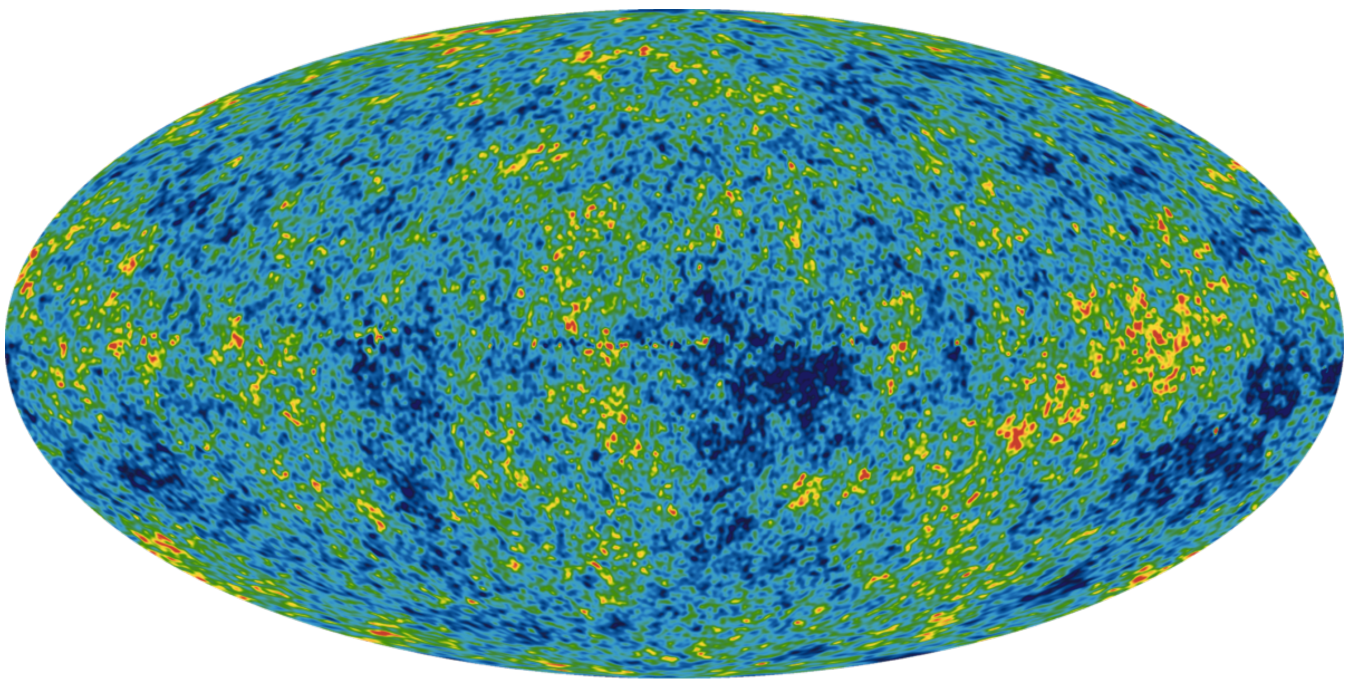
is always between -1 and 1.

Visual representation of covariance matrix (where $r = \dfrac{C_{12}}{\sqrt{C_{11} C_{22}}}$)



The CMB is a multivariate Gaussian random variable!

If the map below is represented with $N=10^7$ pixels, then the statistics are described perfectly (as far as we know) by a multivariate Gaussian, whose N-by-N covariance matrix can be calculated numerically in the standard model.



Behavior of mean and covariance under linear transformations.

Let X_i be an N-component random variable, and define an M-component random variable Y_a by:

$$Y_a = A_{ai} X_i \qquad (A_{ai} \text{ is an M-by-N matrix})$$

Then the mean (a vector) and covariance matrix transform as:

$$\bar{Y}_a = \langle A_{ai} X_i \rangle = A_{ai} \bar{X}_i$$

$$\begin{aligned}
\mathrm{Cov}(Y_a, Y_b) &= \langle (Y_a - \bar{Y}_a)(Y_b - \bar{Y}_b) \rangle \\
&= \langle (A_{ai}(X_i - \bar{X}_i))(A_{bj}(X_j - \bar{X}_j)) \rangle \\
&= A_{ai} A_{bj} \langle (X_i - \bar{X}_i)(X_j - \bar{X}_j) \rangle \\
&= A_{ai} A_{bj} \mathrm{Cov}(X_i, X_j)
\end{aligned}$$

Or in index-free notation:

$$\bar{Y} = A\bar{X} \qquad C_Y = A C_X A^T$$

For arbitrary random variables $Y_a = A_{ai} X_i$, the mean and covariance transform as:

$$\bar{Y} = A\bar{X} \qquad C_Y = AC_X A^T$$

Theorem (proof omitted): if X_i is Gaussian, then Y_a is also Gaussian. In this case, the mean and covariance completely determine the statistics.

In particular, the question of whether a random variable X_i is Gaussian does not depend on the choice of basis. (Changing basis $X_i \rightarrow X'_i$ is the special case where A is invertible.)

Sometimes, problems involving random variables are linear algebra problems in disguise.

Example: how to simulate (on the computer) a Gaussian random variable X_i with specified covariance matrix C_{ij}?

Sometimes, problems involving random variables are linear algebra problems in disguise.

Example: how to simulate (on the computer) a Gaussian random variable X_i with specified covariance matrix C_{ij}?

Answer: diagonalize C

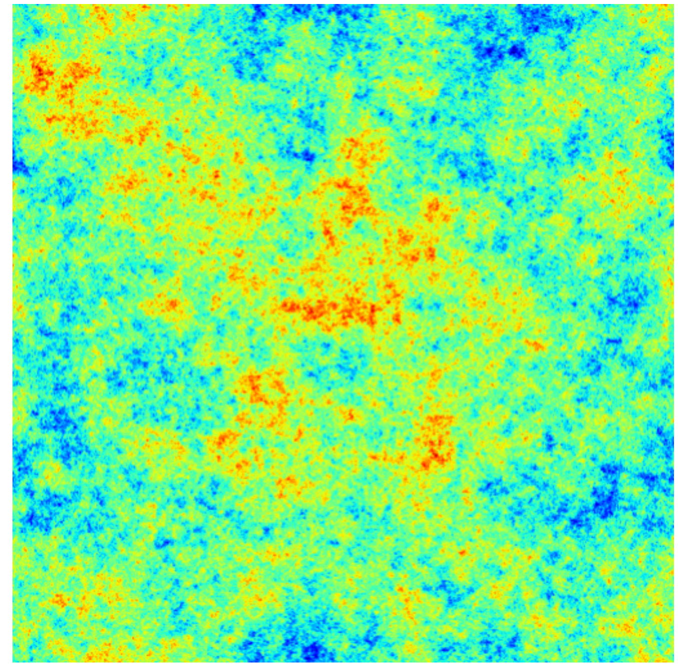
$$C = R\Lambda R^{-1} \qquad \text{where } \Lambda = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \cdots & \\ & & & \lambda_N \end{pmatrix} \text{ and } R^{-1} = R^T$$

Now simulate a Gaussian random variable Y_i with covariance matrix Λ (straightforward, since Λ is diagonal).

Define $X = RY$. This is a Gaussian random variable with covariance matrix $C_X = R\, C_Y\, R^T = R\, \Lambda\, R^{-1} = C$, as desired.

Random fields.

Consider an image f_p with 256^2 (say) pixels. (where $p=1, \ldots, 256^2$).



If f_p is a random variable, then its covariance $C_{pp'}$ is a 256^2-by-256^2 matrix. (Assume mean $\bar{f}_p = 0$ for simplicity.)

Now take the continuum limit:

pixelized image f_p	$\rightarrow$	continuous function $f(\mathbf{x})$
covariance matrix $C_{pq} = \langle f_p f_{p'} \rangle$	$\rightarrow$	two-point correlation function $\langle f(\mathbf{x})\, f(\mathbf{x}') \rangle$

Unless stated otherwise, we will be interested in random fields which are translation and rotation invariant, so that the two-point function < f(**x**) f(**x'**)> depends only on the *scalar* separation |**x** - **x'**|.

$$\langle f(\mathbf{x}) f(\mathbf{x}')\rangle = \zeta(|\mathbf{x} - \mathbf{x}'|)$$

ζ is called the "correlation function".

Now let's compute the two-point function in Fourier space.

$$
\begin{aligned}
\langle f(\mathbf{k}) f(\mathbf{k}')^* \rangle &= \left\langle \left(\int d^n\mathbf{x}\ f(x) e^{-i\mathbf{k}\cdot\mathbf{x}} \right) \left(\int d^n\mathbf{x}'\ f(x) e^{i\mathbf{k}'\cdot\mathbf{x}'} \right) \right\rangle \\
&= \int d^n\mathbf{x}\, d^n\mathbf{x}' \left\langle f(\mathbf{x}) f(\mathbf{x}') \right\rangle e^{-i\mathbf{k}\cdot\mathbf{x}+i\mathbf{k}'\cdot\mathbf{x}'} \\
&= \int d^n\mathbf{x}\, d^n\mathbf{x}'\, \zeta(|\mathbf{x}-\mathbf{x}'|) e^{-i\mathbf{k}\cdot\mathbf{x}+i\mathbf{k}'\cdot\mathbf{x}'} \\
&= \int d^n\mathbf{x}\, d^n\mathbf{r}\, \zeta(|\mathbf{r}|) e^{-i\mathbf{k}\cdot\mathbf{x}+i\mathbf{k}'\cdot(\mathbf{x}-\mathbf{r})} \qquad\qquad (\mathbf{r} = \mathbf{x}-\mathbf{x}') \\
&= \left[\int d^n\mathbf{r}\, \zeta(|\mathbf{r}|) e^{-i\mathbf{k}\cdot\mathbf{r}} \right] (2\pi)^n \delta^n(\mathbf{k}-\mathbf{k}')
\end{aligned}
$$

The quantity in brackets is called the power spectrum P(k).

We have now shown that the two-point statistics of a random field are given equivalently by:

$$\langle f(\mathbf{x}) f(\mathbf{x}') \rangle = \zeta(|\mathbf{x} - \mathbf{x}'|) \qquad \text{in real space}$$

$$\left\langle f(\mathbf{k}) f(\mathbf{k}')^* \right\rangle = P(|\mathbf{k}|)\,(2\pi)^n \delta^n(\mathbf{k} - \mathbf{k}') \qquad \text{in Fourier space}$$

and the correlation function ζ(r) and power spectrum P(k) are related to each other by Fourier transforms (“Weiner-Khinchin theorem”):

$$P(k) = \int d^n\mathbf{r}\,\zeta(|\mathbf{r}|)\,e^{-i\mathbf{k}\cdot\mathbf{r}}$$

$$\zeta(r) = \int \frac{d^n\mathbf{k}}{(2\pi)^n}\,P(|\mathbf{k}|)\,e^{i\mathbf{k}\cdot\mathbf{r}}$$

A random field is Gaussian if its real-space values f(**x**) are a multivariate Gaussian random variable in the usual sense. In this case, the statistics are completely determined by the two-point function (either ζ(r) or P(k)).

Gaussian random fields are easy to think about in Fourier space, since the covariance is always diagonal:

$$\left\langle f(\mathbf{k}) f(\mathbf{k}')^* \right\rangle = P(|\mathbf{k}|)\,(2\pi)^n \delta^n(\mathbf{k}-\mathbf{k}') \qquad (*)$$

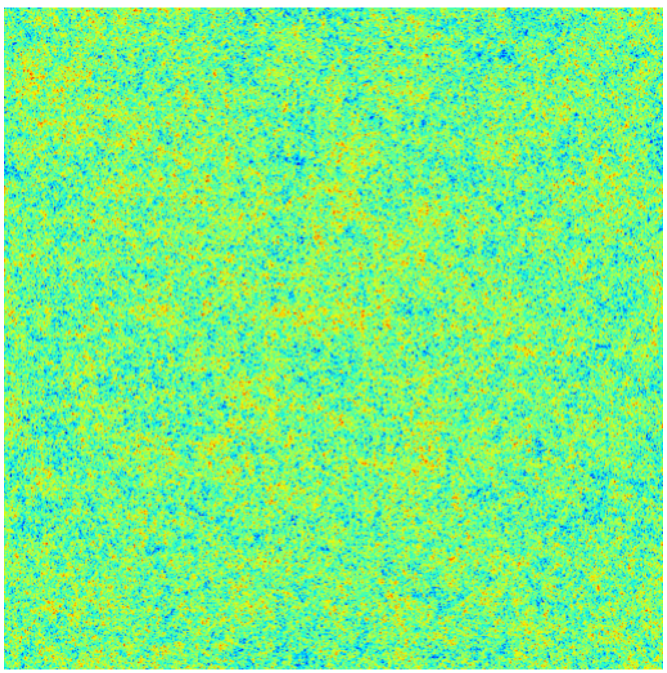
In Fourier space, a Gaussian random field is just a collection of *independent* Gaussian random variables f(**k**).

The delta function on the RHS of (*) can also be understood from translation invariance. Under a translation x => x + a, the two-point function transforms as:

$$\left\langle f(\mathbf{k}) f(\mathbf{k})^* \right\rangle \to e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{a}} \left\langle f(\mathbf{k}) f(\mathbf{k})^* \right\rangle$$

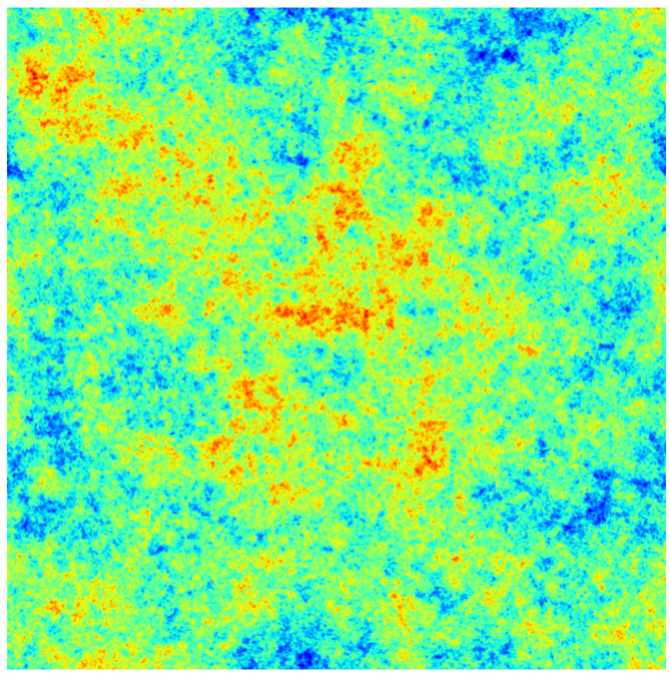
Example: Two-dimensional Gaussian random fields with power-law spectra $P(l) \propto l^{\alpha}$

(Note: cosmologists are hardwired to denote wavenumbers by k in 3D, by l in 2D, and by ω in 1D.)



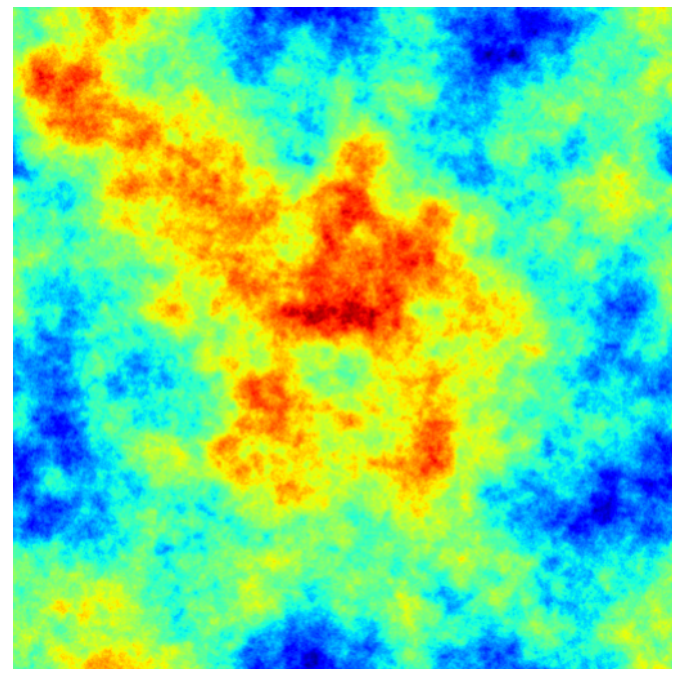
$\alpha = -1$

“blue” spectrum



$\alpha = -2$

scale invariant

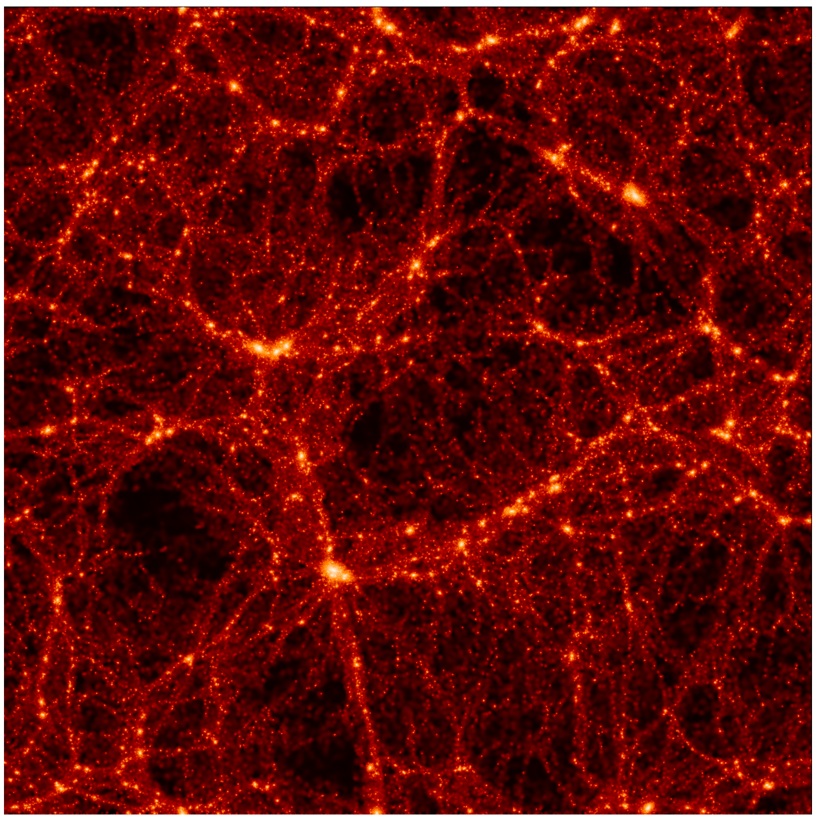


$\alpha = -3$

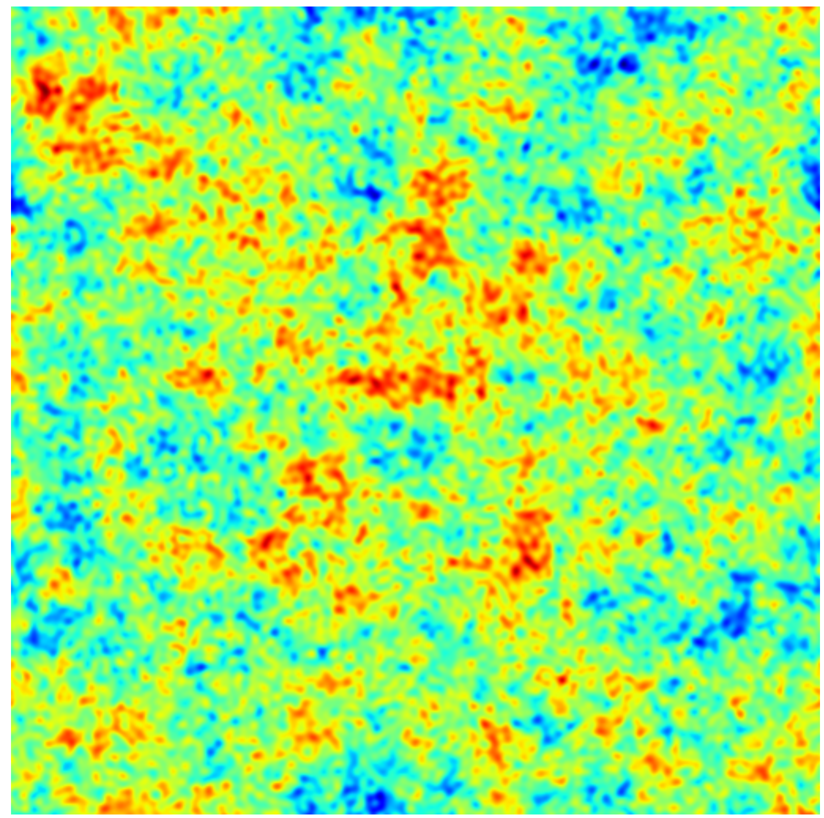
“red” spectrum

Reminder: for a Gaussian field, the statistics of the field are completely determined by the power spectrum P(k).

For a non-Gaussian field, this is not true!



cosmological density field at z=0

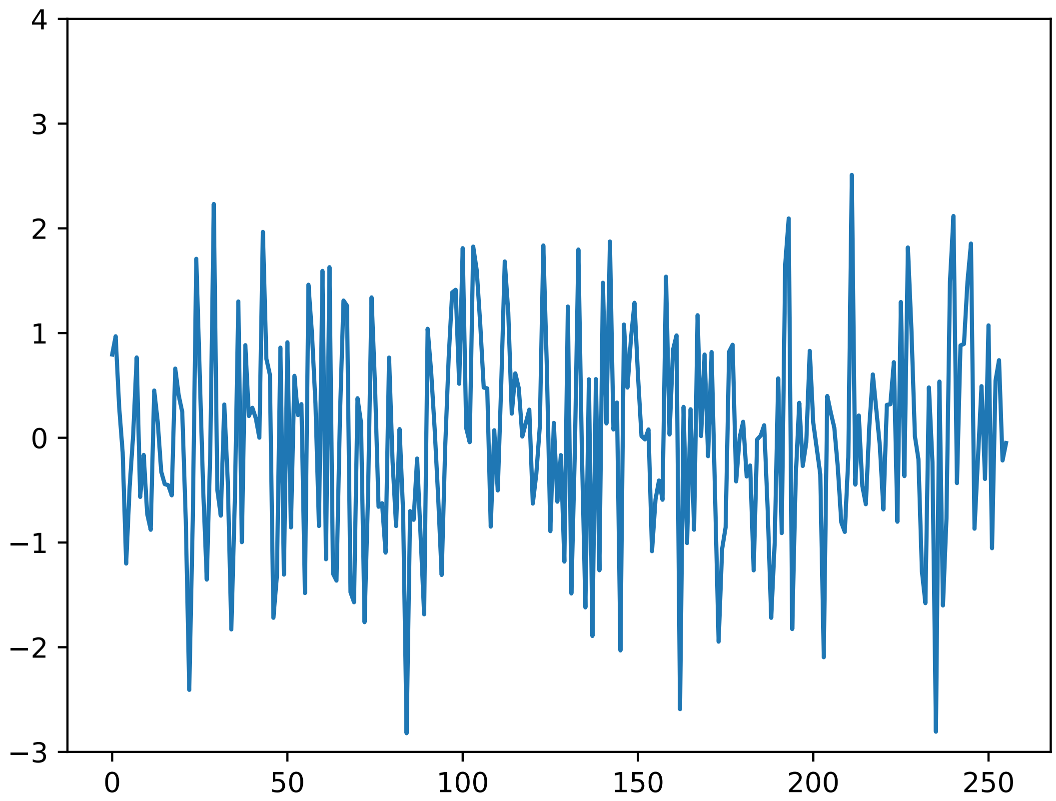


Gaussian field with same power spectrum

Gaussian white noise: simplest example of a Gaussian random field. The correlation function is a delta function.

$$\zeta(\mathbf{r}) = A\,\delta^n(\mathbf{r})$$

Each pixel value is an independent Gaussian random variable. (Covariance matrix is diagonal in real space and Fourier space!)



Gaussian white noise: simplest example of a Gaussian random field. The correlation function is a delta function.

$$\zeta(\mathbf{r}) = A\,\delta^n(\mathbf{r})$$

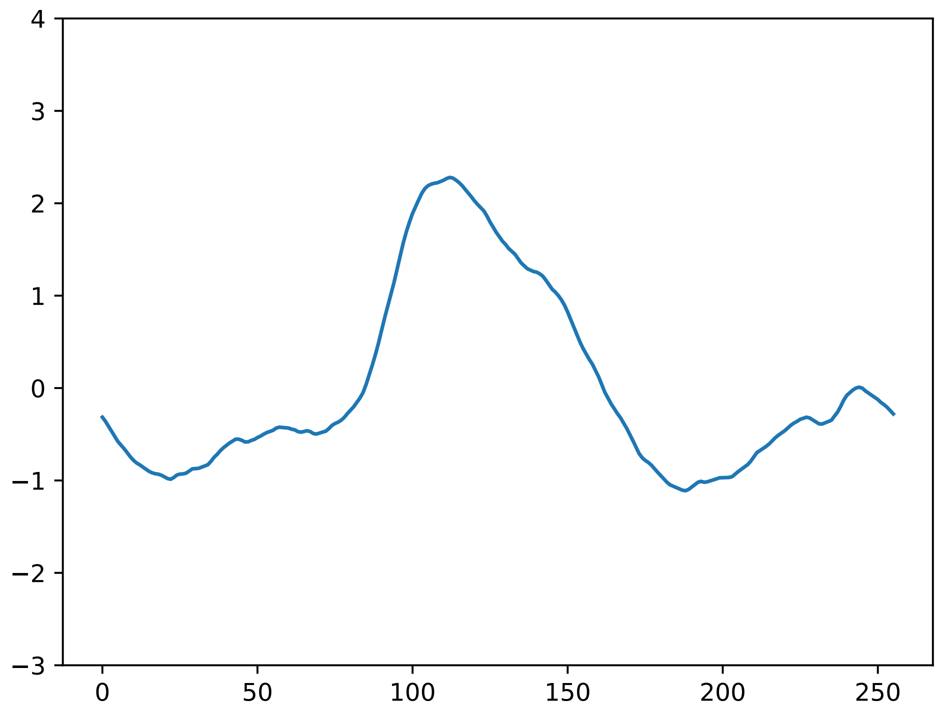
Each pixel value is an independent Gaussian random variable. (Covariance matrix is diagonal in real space and Fourier space!)

Power spectrum P(k) is constant in k. This follows from the Wiener-Khinchin theorem:

$$\begin{aligned} P(k) &= \int d^n\mathbf{r}\,\zeta(\mathbf{r})e^{-i\mathbf{k}\cdot\mathbf{r}} \\ &= \int d^n\mathbf{r}\,(A\delta^n(\mathbf{r}))e^{-i\mathbf{k}\cdot\mathbf{r}} \\ &= A \end{aligned}$$

A linear operator applied to a GRF (Gaussian random field) gives another GRF. (This follows from the general statement that linear combinations of Gaussians are Gaussian.)

Example: what is the power spectrum of a one-dimensional Gaussian random walk? (Obtained by adding an independent Gaussian random number at each timestep.)



A linear operator applied to a GRF (Gaussian random field) gives another GRF. (This follows from the general statement that linear combinations of Gaussians are Gaussian.)

Example: what is the power spectrum of a one-dimensional Gaussian random walk? (Obtained by adding an independent Gaussian random number at each timestep.)

To answer this, we note that a random walk is the integral of white noise. Therefore:

$$
\begin{aligned}
f_{\mathrm{RW}}(\omega) &= \frac{1}{i\omega}\, f_{\mathrm{WN}}(\omega) \\
P_{\mathrm{RW}}(\omega) &= \frac{1}{\omega^2}\, P_{\mathrm{WN}}(\omega) \\
&= \frac{A}{\omega^2}
\end{aligned}
$$

Another example which is more representative of the CMB.
Let f(t,x) be a field which evolves via the wave equation:

$$\left(\frac{\partial^2}{\partial t^2} - c_s^2 \frac{\partial^2}{\partial x^2}\right) f(t,x) = 0$$

c_s = "sound speed"

with the following initial conditions at t=0:

- f(x) is a Gaussian random field with power spectrum $P_0(k)$
  $\partial f / \partial t = 0$

Question: what is the power spectrum $P_T(k)$ of a spatial "snapshot" f(T,x) at time t=T?

We take a spatial Fourier transform x $\rightarrow$ k (but not t$\rightarrow$w).
Then the wave equation $(\partial_t^2 - c_s^2 \partial_x^2) f = 0$ becomes:

$$\left(\frac{\partial^2}{\partial t^2} + c_s^2 k^2 \right) f(t, k) = 0$$

and the solution is (using $\partial f / \partial t = 0$)

$$f(t, k) = \cos(c_s k t) f(0, k)$$

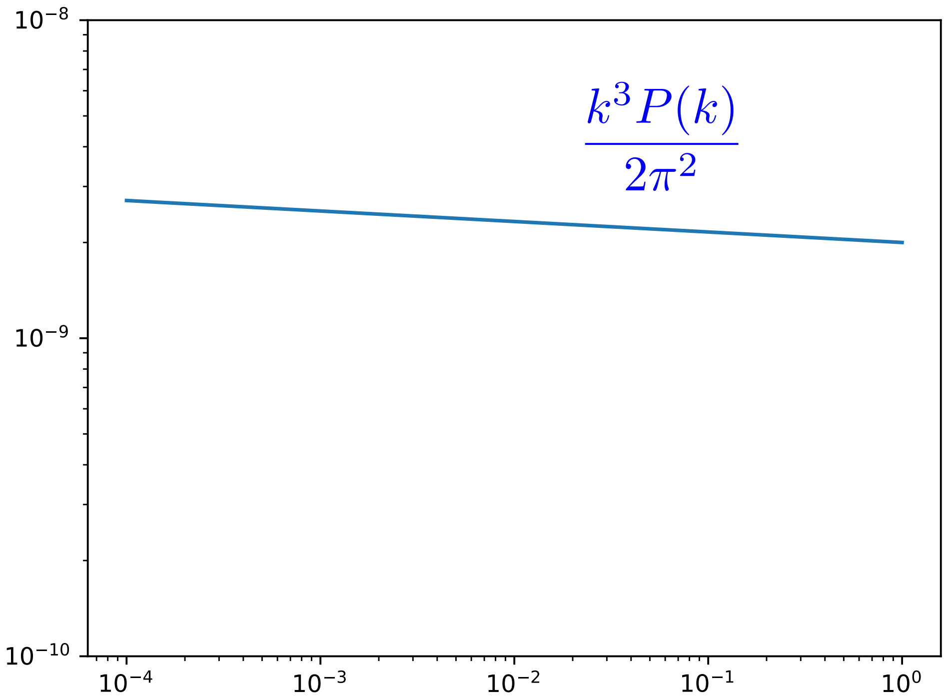
The spatial power spectrum $P_T(k)$ at time t=T is:

$$P_T(k) = \cos(c_s k T)^2 P_0(k)$$

i.e. time evolution imprints peaks on the power spectrum.

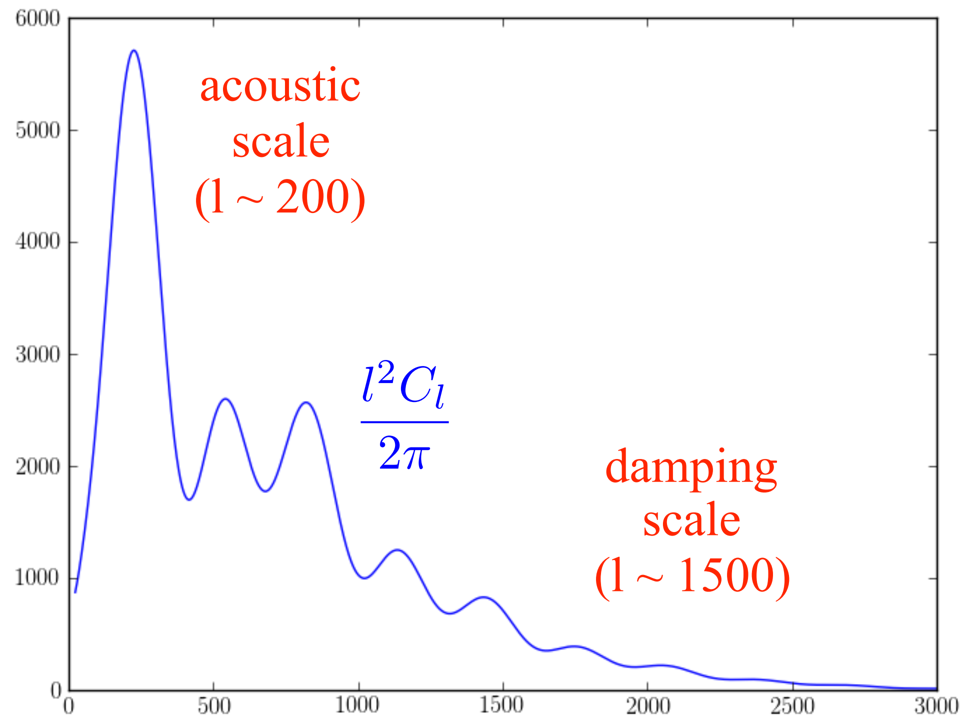
Analogously, time evolution “imprints” features on cosmological power spectra, starting from a featureless initial power spectrum.

3D power spectrum
of initial conditions
(adiabatic curvature)



wavenumber k (Mpc^{-1})

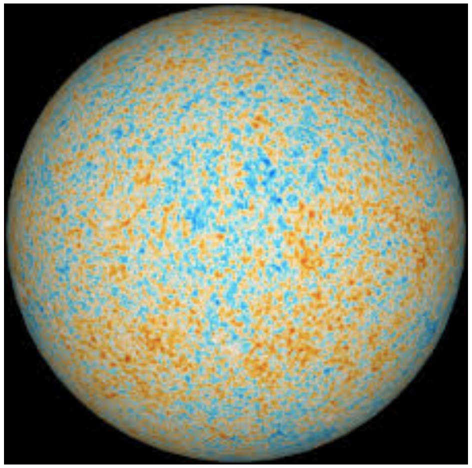
2D CMB power spectrum



angular multipole l

Curved sky.

So far, our fields have been defined on Euclidean space, but some fields are defined on the unit sphere, e.g. CMB temperature $T(\theta,\phi)$.



In Euclidean space, any field $f(x)$ can be represented as a linear combination of plane waves e^{ikx} (Fourier transform).

Analogous statement on the sphere: any field $f(\theta,\phi)$ is a linear combination of spherical harmonics $Y_{lm}(\theta,\phi)$.

The spherical harmonic $Y_{lm}(\theta,\phi)$ is a special function defined for integers $\ell = 0,1,2,\ldots$ and $m = -\ell, (-\ell+1), \ldots, \ell$.

Spherical analogue of a plane wave e^{ikx}. The wavenumber ℓ is quantized (an integer), and there are $(2\ell+1)$ harmonics for each ℓ.

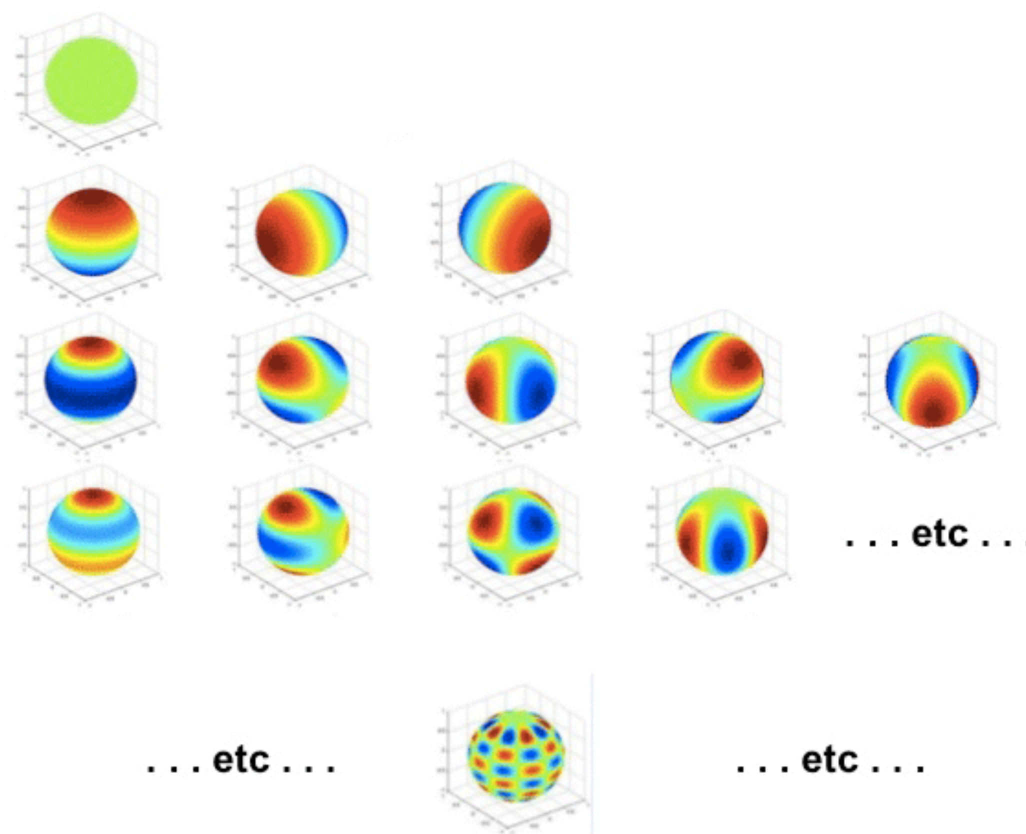
Any function $f(\theta,\phi)$ is representable as $f(\theta,\phi) = \sum_{lm} a_{lm} Y_{lm}(\theta,\phi)$

$\ell=0$ (monopole)

$\ell=1$ (dipole)

$\ell=2$ (quadrupole)

$\ell=3$ (octopole)



	Euclidean field	Spherical field
real-space representation	$f(\mathbf{x})$	$f(\theta, \phi)$
harmonic-space representation	$\tilde{f}(\mathbf{k})$	a_{lm}
harmonic transform	$f(\mathbf{x}) = \int \frac{d^n\mathbf{k}}{(2\pi)^n} \tilde{f}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}}$	$f(\theta, \phi) = \sum_{lm} a_{lm} Y_{lm}(\theta, \phi)$
inverse transform	$\tilde{f}(\mathbf{k}) = \int d^n\mathbf{x}\, f(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}}$	$a_{lm} = \int d(\cos\theta)\, d\phi\, f(\theta, \phi) Y^*_{lm}(\theta, \phi)$
power spectrum	$\langle f(\mathbf{k}) f(\mathbf{k}')^* \rangle = P(k)\, (2\pi)^n \delta^n(\mathbf{k} - \mathbf{k}')$	$\langle a_{lm} a^*_{l'm'} \rangle = C_l \delta_{ll'} \delta_{mm'}$

Part 3: forecasting and the Fisher matrix

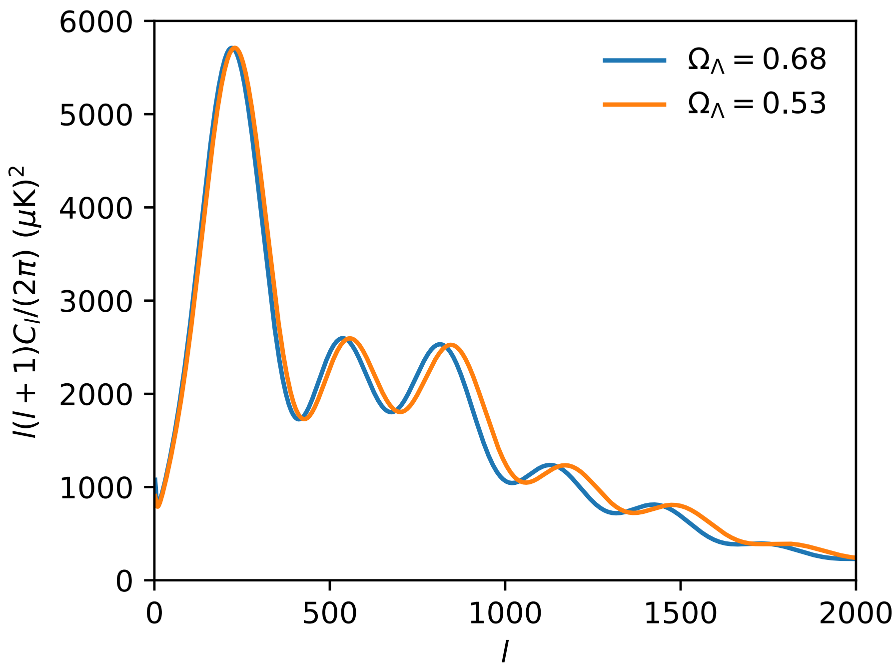
By my estimation, there are ~100 cosmology papers per year which are mainly concerned with computing a "Fisher matrix" F_{ab}.

What is the Fisher matrix and why is it so widespread?

By my estimation, there are ~100 cosmology papers per year which are mainly concerned with computing a "Fisher matrix" F_{ab}.

What is the Fisher matrix and why is it so widespread?

Motivating example: forecasting parameter sensitivity of the CMB



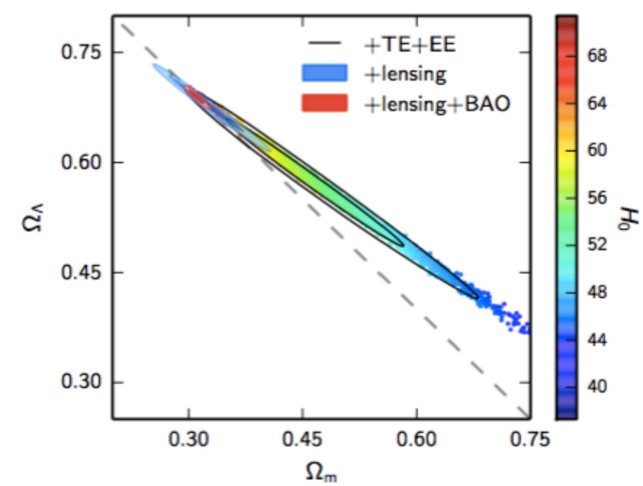


Fig. 26. Constraints in the Ω_m–Ω_Λ plane from the *Planck* TT+lowP data (samples; colour-coded by the value of H_0) and *Planck* TT,TE,EE+lowP (solid contours). The geometric degeneracy between Ω_m and Ω_Λ is partially broken because of the effect of lensing on the temperature and polarization power spectra. These limits are improved significantly by the inclusion of the *Planck* lensing reconstruction (blue contours) and BAO (solid red contours). The red contours tightly constrain the geometry of our Universe to be nearly flat.

Planck 2015

Abstract setup:

"model parameters" $\{\theta^a\}$	$\xrightarrow{p(d\|\theta)}$	"data" d
e.g. ρ_Λ, Ω_b, Ω_c, Δ_ζ^2, n_s, τ		e.g. CMB multipoles a_{lm}

The data d is a random variable whose probability distribution depends on the model parameters θ^a.

$p(d|\theta)$ = conditional probability distribution of data d, given model parameters θ

“model parameters” $\{\theta^a\}$

e.g. ρ_Λ, Ω_b, Ω_c, Δ_ζ^2, n_s, τ

$\xrightarrow{p(d|\theta)}$

“data” d

e.g. CMB multipoles a_{lm}

We might be interested in:

1. Simulation. Given a model θ^a, how do we simulate a random data realization d? (i.e. sample the conditional PDF $p(d|\theta)$)

2. Analysis. Given a data realization d, what are the constraints (say at 95% CL) on the model space θ^a?

3. Forecasting. Given a rough fiducial guess θ_{fid} for the true model, what constraints on the model space do we expect to obtain, for a “typical” realization of the data?

The Fisher matrix is a tool for forecasting (#3).

$$\boxed{\begin{array}{c}\text{“model parameters” } \{\theta^a\} \\ \text{e.g. } \rho_\Lambda, \Omega_b, \Omega_c, \Delta_\zeta^2, n_s, \tau\end{array}} \xrightarrow{p(d|\theta)} \boxed{\begin{array}{c}\text{“data” } d \\ \text{e.g. CMB multipoles } a_{lm}\end{array}}$$

General definition of the Fisher matrix (to be motivated later!)

$$F_{ab} = -\left\langle \frac{\partial^2 \log p(d|\theta)}{\partial\theta^a\,\partial\theta^b} \right\rangle_d$$

where the expectation value is taken over random realizations of the data d, for a preferred fiducial choice of model parameters θ_{fid}

The Fisher matrix depends on the fiducial model θ_{fid}, but does not require a preferred realization of the data d (just the probability distribution $p(d|\theta)$).

"model parameters" $\{\theta^a\}$

e.g. ρ_Λ, Ω_b, Ω_c, Δ_ζ^2, n_s, τ

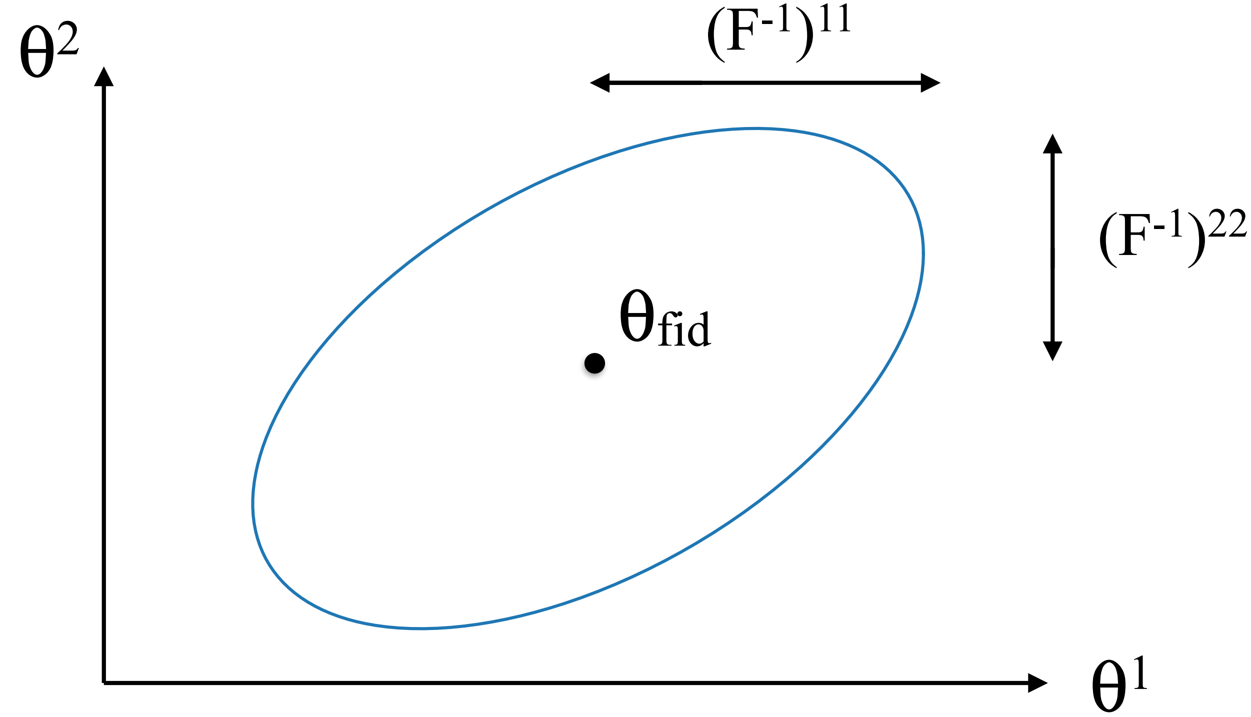
$$\xrightarrow{p(d|\theta)}$$

"data" d

e.g. CMB multipoles a_{lm}

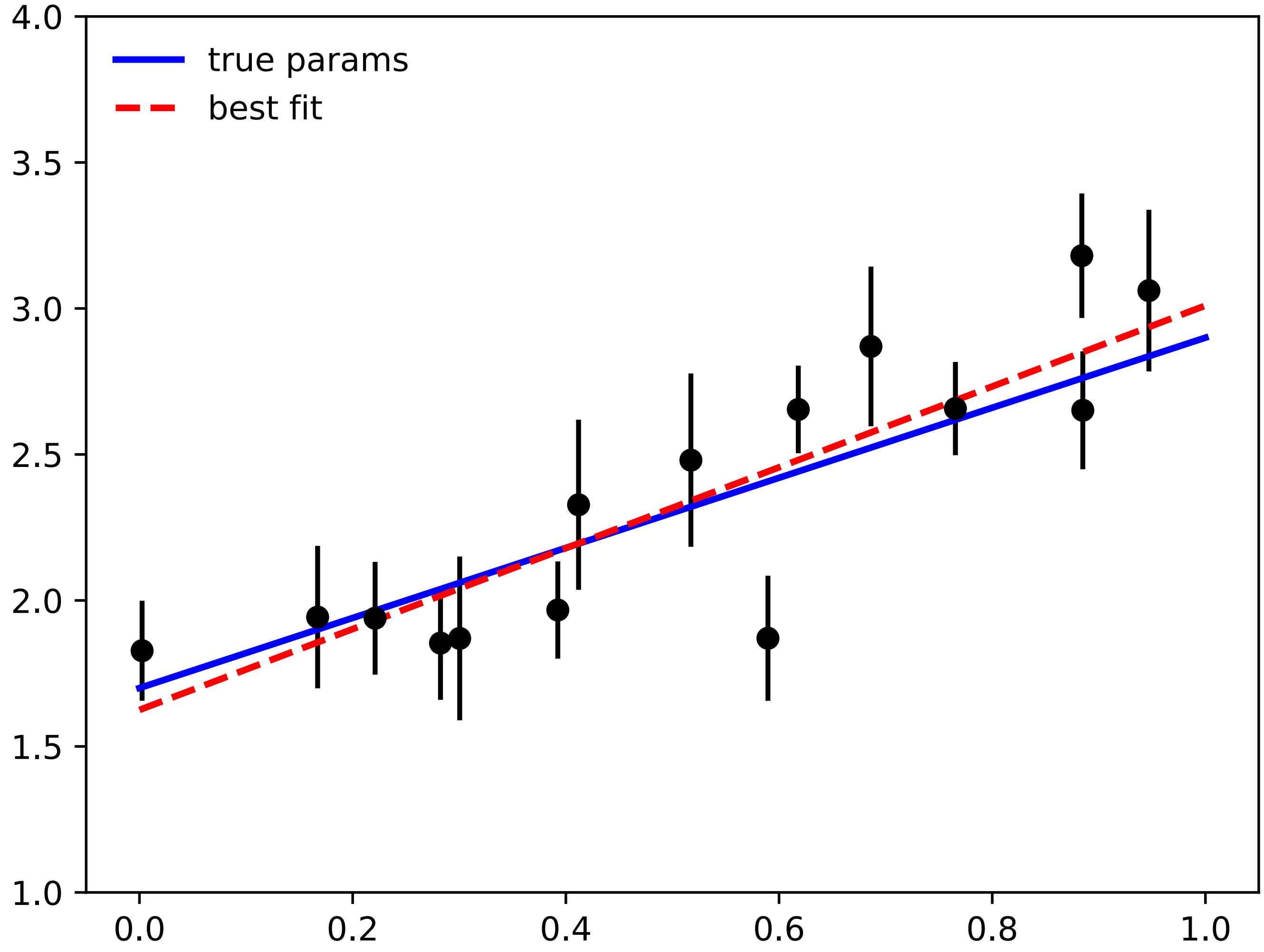
Interpretation: the Fisher matrix F_{ab} is the forecasted *inverse* covariance matrix of the model constraints obtained from a "typical" realization of the data.

$$F_{ab} = -\left\langle \frac{\partial^2 \log p(d|\theta)}{\partial\theta^a\,\partial\theta^b} \right\rangle_d$$

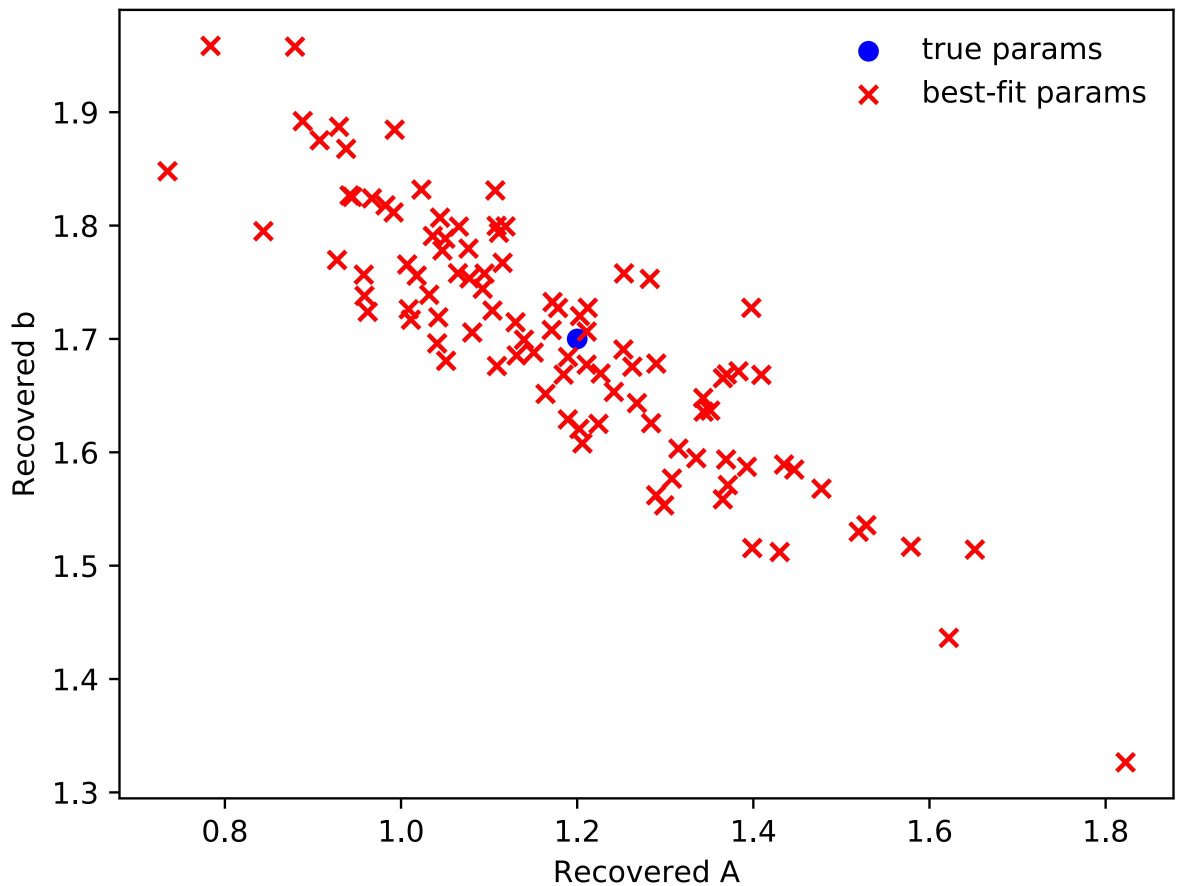


Toy example: linear regression

Fitting a line y=Ax+b through points (x_i,y_i) with error bars σ_i.



The following scatterplot shows the result of repeating the linear regression 100 times. (Detail: the y-values were randomized, but x_i and σ_i were held fixed throughout.)



Let's compute the Fisher matrix for the model parameters $(\theta^1,\theta^2) = (A,b)$, and compare with this plot.

$$
\boxed{\begin{array}{c}\text{“model parameters” } \{\theta^a\} \\ A,b\end{array}} \xrightarrow{p(d|\theta)} \boxed{\begin{array}{c}\text{“data” } d \\ \text{y-values } y_i\end{array}}
$$

First we need to write down the conditional likelihood $p(d|\theta)$. Given $(\theta^1,\theta^2) = (A,b)$, the PDF of an *individual* data value y_i is a Gaussian with mean (Ax_i+b) and variance σ_i^2:

$$
\begin{aligned}
p(y_i|\theta) &= \frac{1}{(2\pi\sigma_i^2)^{1/2}} \exp\left(-\frac{(y_i - Ax_i - b)^2}{2\sigma_i^2}\right) \\
&= \frac{1}{(2\pi\sigma_i^2)^{1/2}} \exp\left(-\frac{(y_i - \theta^a f_a(x_i))^2}{2\sigma_i^2}\right)
\end{aligned}
$$

introducing the two-vector notation $f_a(x) = \begin{pmatrix} x \\ 1 \end{pmatrix}$

“model parameters” $\{\theta^a\}$

A,b

$$\xrightarrow{p(d|\theta)}$$

“data” d

y-values y_i

The conditional likelihood $p(d|\theta)$ is obtained by multiplying the PDF’s for each individual data value y_i.

$$p(d|\theta) = \prod_i \frac{1}{(2\pi\sigma_i^2)^{1/2}} \exp\left(-\frac{(y_i - \theta^a f_a(x_i))^2}{2\sigma_i^2}\right)$$

$$f_a(x) = \begin{pmatrix} x \\ 1 \end{pmatrix}$$

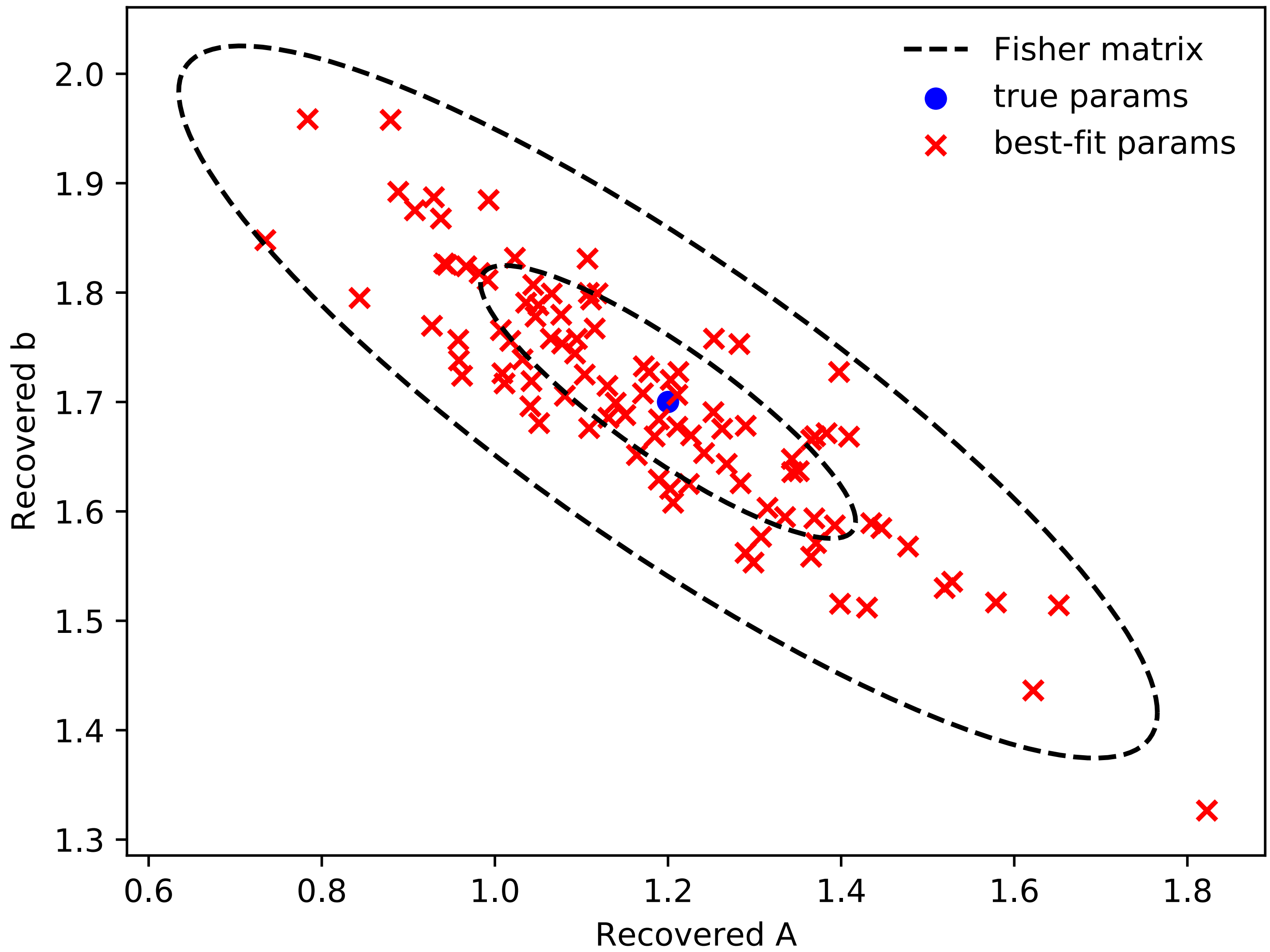
Now we can compute the Fisher matrix $F_{ab} = -\left\langle \frac{\partial^2 \log p(d|\theta)}{\partial\theta^a\,\partial\theta^b} \right\rangle_d$

$$p(d|\theta) = \prod_i \frac{1}{(2\pi\sigma_i^2)^{1/2}} \exp\left(-\frac{(y_i - \theta^a f_a(x_i))^2}{2\sigma_i^2}\right)$$

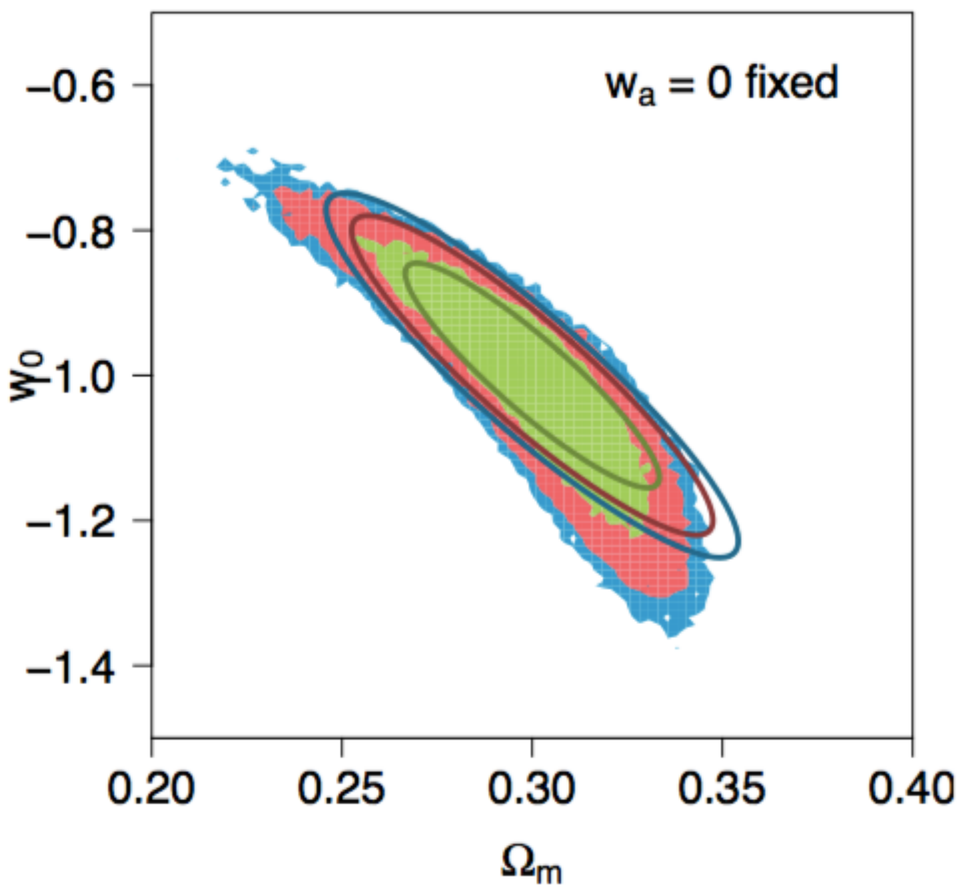
$$\begin{aligned}
\Rightarrow F_{ab} &= -\left\langle \frac{\partial^2 \log p(d|\theta)}{\partial\theta^a\,\partial\theta^b} \right\rangle_d \\
&= -\left\langle \frac{\partial^2}{\partial\theta^a\,\partial\theta^b} \sum_i \left(-\frac{1}{2}\log(2\pi\sigma_i^2) - \frac{1}{2\sigma_i^2}(y_i - \theta^c f_c(x_i))^2\right) \right\rangle_y \\
&= -\left\langle \frac{\partial^2}{\partial\theta^a\,\partial\theta^b} \sum_i \left(-\frac{1}{2\sigma_i^2}(\theta^c f_c(x_i))(\theta^d f_d(x_i))\right) \right\rangle_y \\
&= \left\langle \sum_i \left(\frac{f_a(x_i) f_b(x_i)}{\sigma_i^2}\right) \right\rangle_y \\
&= \sum_i \frac{f_a(x_i) f_b(x_i)}{\sigma_i^2}
\end{aligned}$$

In this example, the expectation value $<>_y$ is trivial, and the Fisher matrix does not depend on a choice of fiducial model $\theta_{fid} = (A_{fid}, b_{fid})$.

Comparison between Fisher matrix and Monte Carlo scatterplot.



A cosmological example. In general, the Fisher matrix is not expected to agree precisely with the Monte Carlo scatterplot. In particular, the contours of the scatterplot need not be ellipses. However, the Fisher matrix is usually easy to compute, and is usually a good approximation.



Example: supernova (not CMB) Fisher matrix from Wolz et al 1205.3984

Switched to blackboard here. For the blackboard part of the lecture (about 90 minutes total), see:

http://pirsa.org/18070002/ 40:00 - 80:00

http://pirsa.org/18070003/ 0:00 - 45:00

Part 4: observational cosmology, past, present and future

In the last few decades, observational cosmology has made amazing progress:

- Sandage (1970): "Cosmology is a search for two numbers" [$H = \dot{a}/a$ and $q = -\ddot{a}a/\dot{a}^2$]

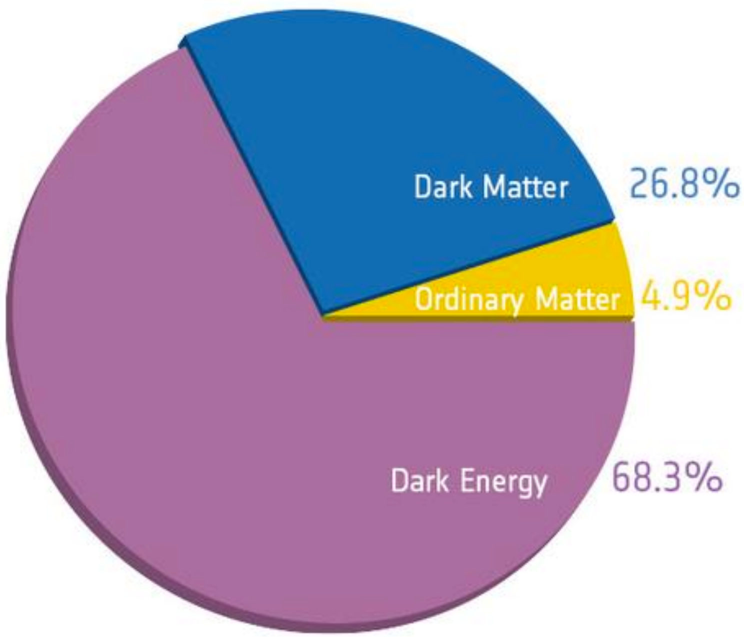
- Peebles: "I did not continue (with computation of CMB anisotropy), in part because I had trouble imagining that such tiny disturbances could be observed" [1992 (COBE)]

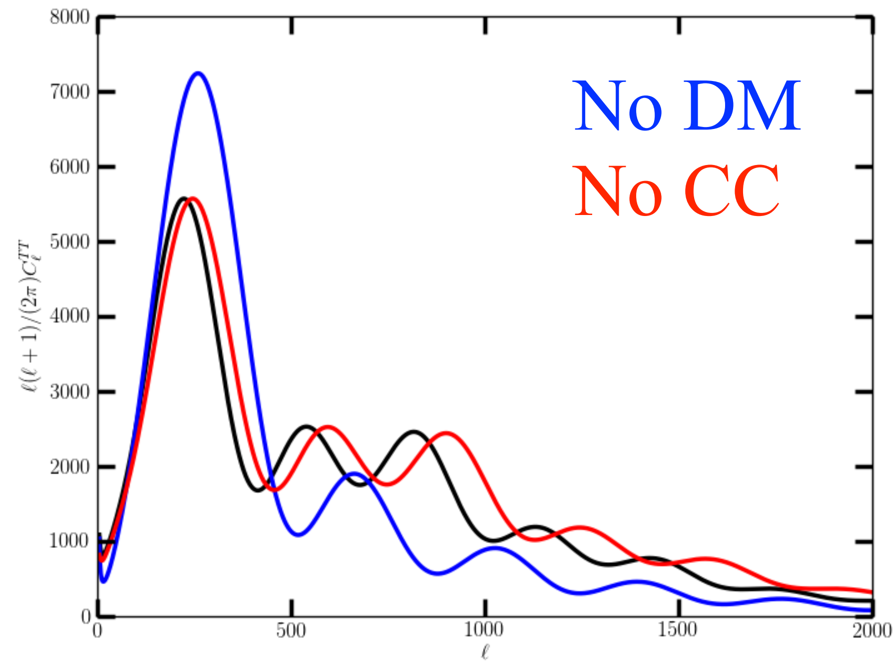
- Sunyaev: "I did not think that the acoustic oscillation would ever be observed" [2000 (multiple experiments)]

- Mukhanov: "I thought it would take 1000 years to detect the logarithmic dependence of the power spectrum" [2006 (WMAP)]

(some quotes taken from a talk by Enrico Pajer)

Good: cosmology provides a real glimpse of physics beyond the (particle physics!) standard model





Planck: Cold dark matter detected at 80σ
Cosmological constant detected at 75σ

Bad: only measure a small number of parameters, hard to test competing hypotheses, or narrow down to a specific model.

Six-parameter standard model:

$\rho_\Lambda = (2.56 \pm 0.04) \times 10^{-47}\ \mathrm{GeV}^4$	Dark energy density (c.c.)
$\Omega_b = 0.0486 \pm 0.0007$	Baryonic matter abundance
$\Omega_c = 0.267 \pm 0.009$	Cold dark matter abundance
$\Delta_\zeta^2 = (2.11 \pm 0.05) \times 10^{-9}$	Initial power spectrum amplitude
$n_s = 0.967 \pm 0.004$	Spectral index
$\tau = 0.058 \pm 0.012$	CMB optical depth

Extensions:

- Non-Gaussian initial conditions ($f_{NL} = 2.5 \pm 5.7$)
- Non-minimal neutrino mass ($m_\nu < 0.23$ eV at 95% CL)
- Extra neutrino species or other light relics ($N_{eff} = 3.04 \pm 0.18$)
- Nonzero spatial curvature ($\Omega_K = 0.000 \pm 0.005$)
- Time-dependent dark energy density ($w = -1.02 \pm 0.08$)
- Cosmological gravity waves ($r < 0.12$ at 95% CL)

+ many more!

Forecasts suggest that there is still a lot of room to shrink error bars in future experiments.

Should we take futuristic forecasts seriously? Historical exercise: Tegmark (1999) forecasts for Planck, very futuristic at the time.

Parameter	Forecasted uncertainty (1999)	Reported uncertainty (2015)
ρ_b	0.94%	0.72%
$(\rho_b + \rho_c)$	1.6%	0.9%
n_s	0.0076	0.0048
$\Lambda/\rho_{\rm tot}$	0.022	0.0087

Planck somewhat outperformed its forecasts.

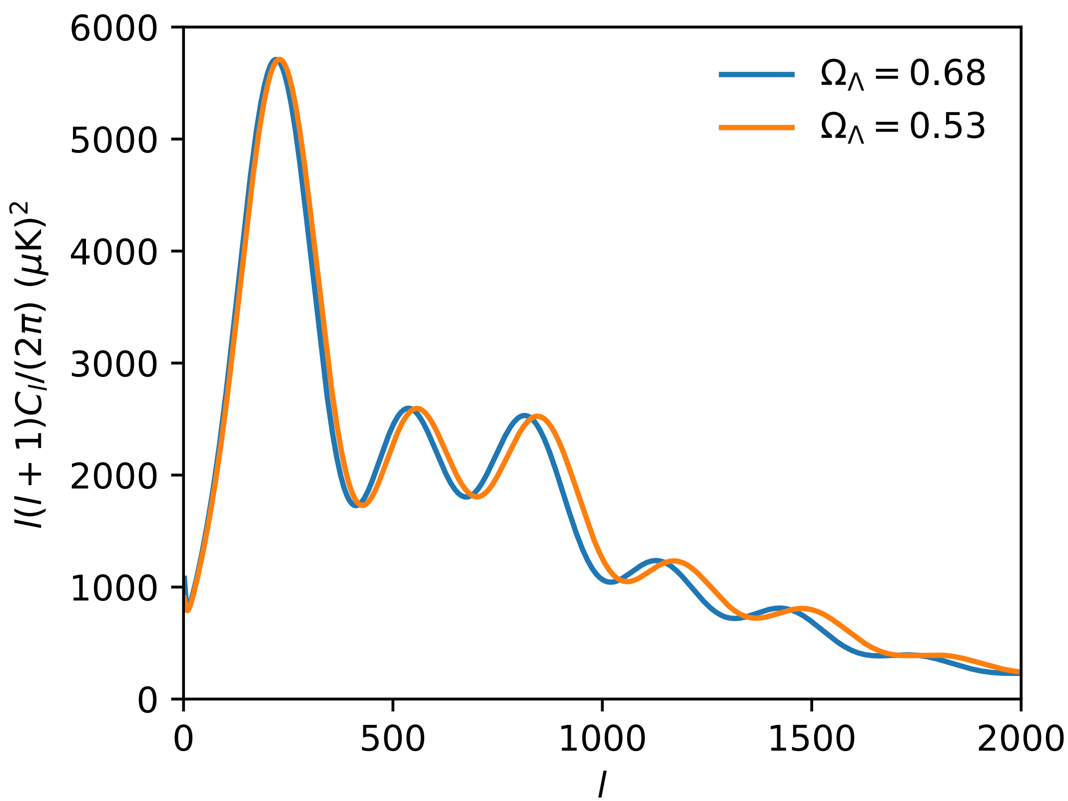
CMB currently dominates constraints on 6-parameter model space

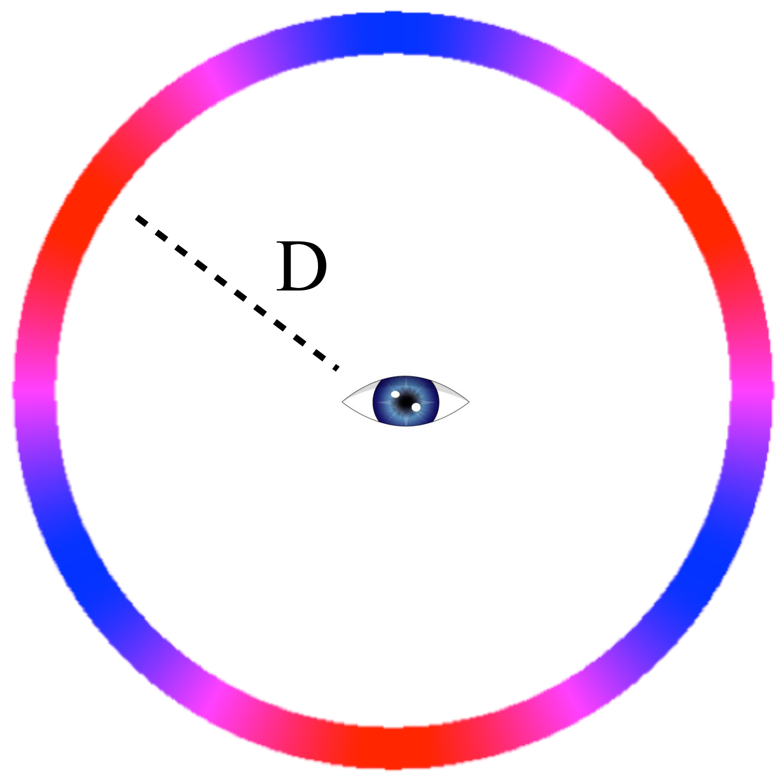
	CMB alone	CMB + baryon acoustic oscillations + type IA supernovae + direct H_0 measurements
DM density $\rho_{c,0}$	0.2618 ± 0.0087	0.2589 ± 0.0063 $(\times \rho_{tot})$
Bary. density $\rho_{b,0}$	0.04884 ± 0.00085	0.04860 ± 0.00070 $(\times \rho_{tot})$
Cosm. constant Λ	2.543 ± 0.071	2.567 ± 0.051 $(\times 10^{-47}\ \mathrm{GeV}^4)$
Amplitude A_ζ	2.130 ± 0.053	2.142 ± 0.049
Spectral index n_s	0.9653 ± 0.0048	0.9667 ± 0.0040
Optical depth τ	0.063 ± 0.014	0.066 ± 0.012

Not true in extensions of standard model! Non-CMB datasets are important when more parameters are added.

CMB temperature observations can constrain the cosmological constant Λ, even though Λ is negligible when the CMB is formed!

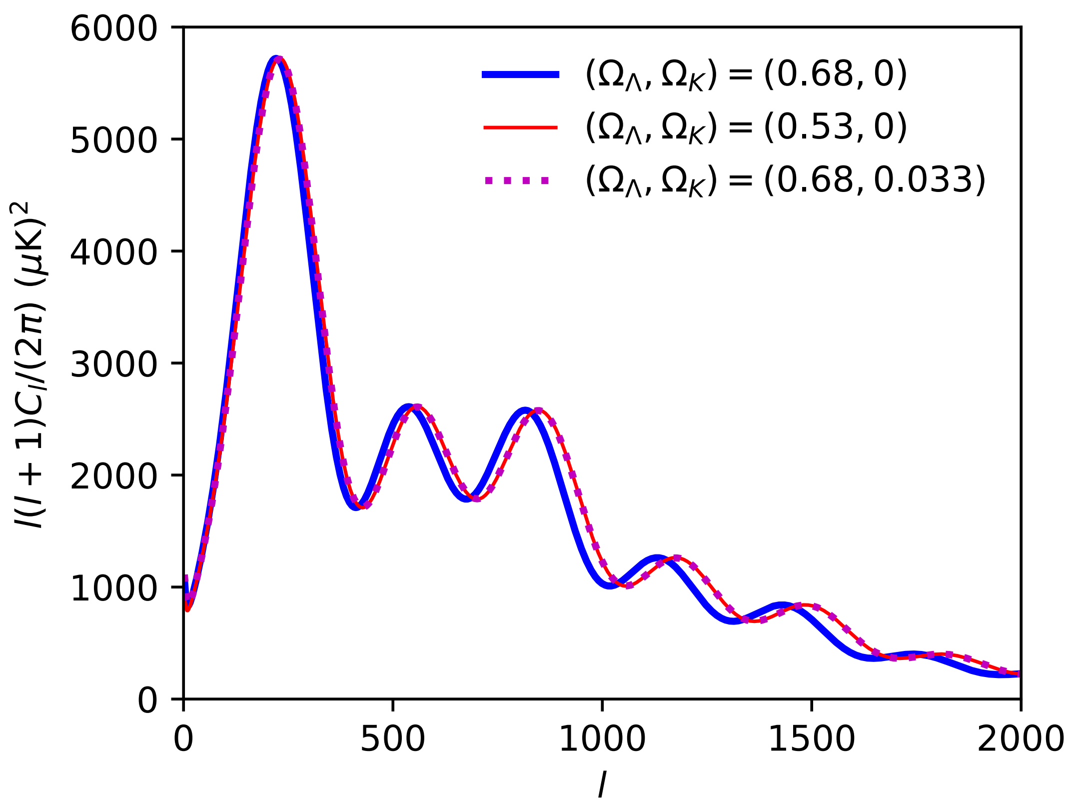
How? By changing the distance D at which the CMB surface is observed.

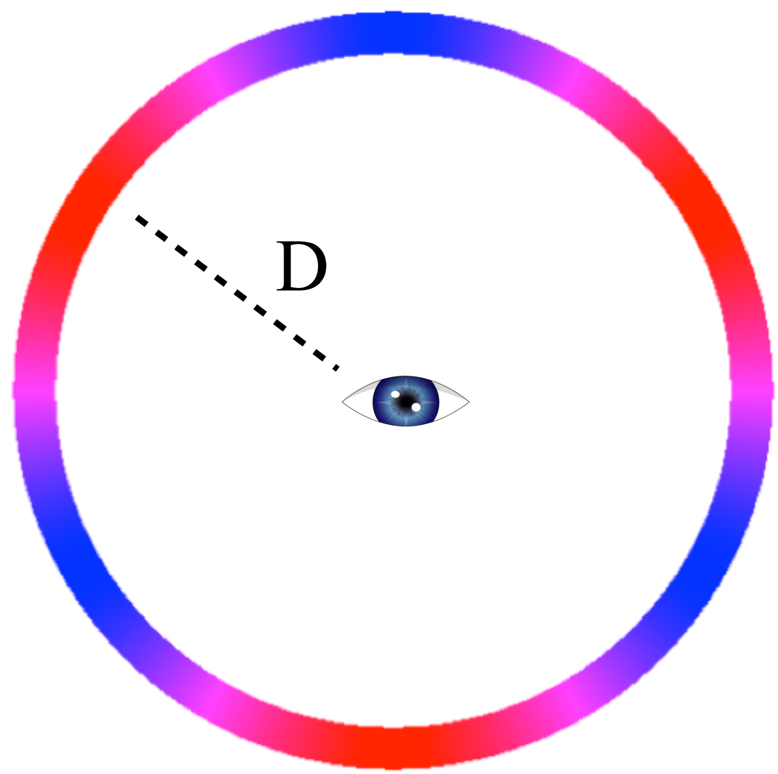




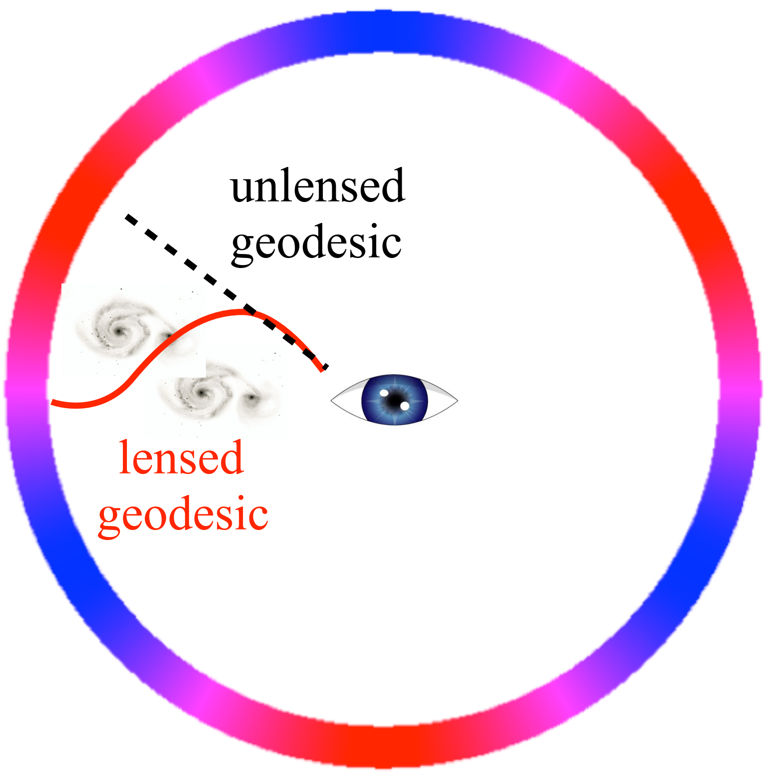
CMB distance degeneracy: in a parameter space with N “late universe” parameters, there is an (N-1)-fold degeneracy.

For example, Ω_Λ and spatial curvature (Ω_K). Before Planck (2013), the CMB power spectrum was consistent with either possibility!





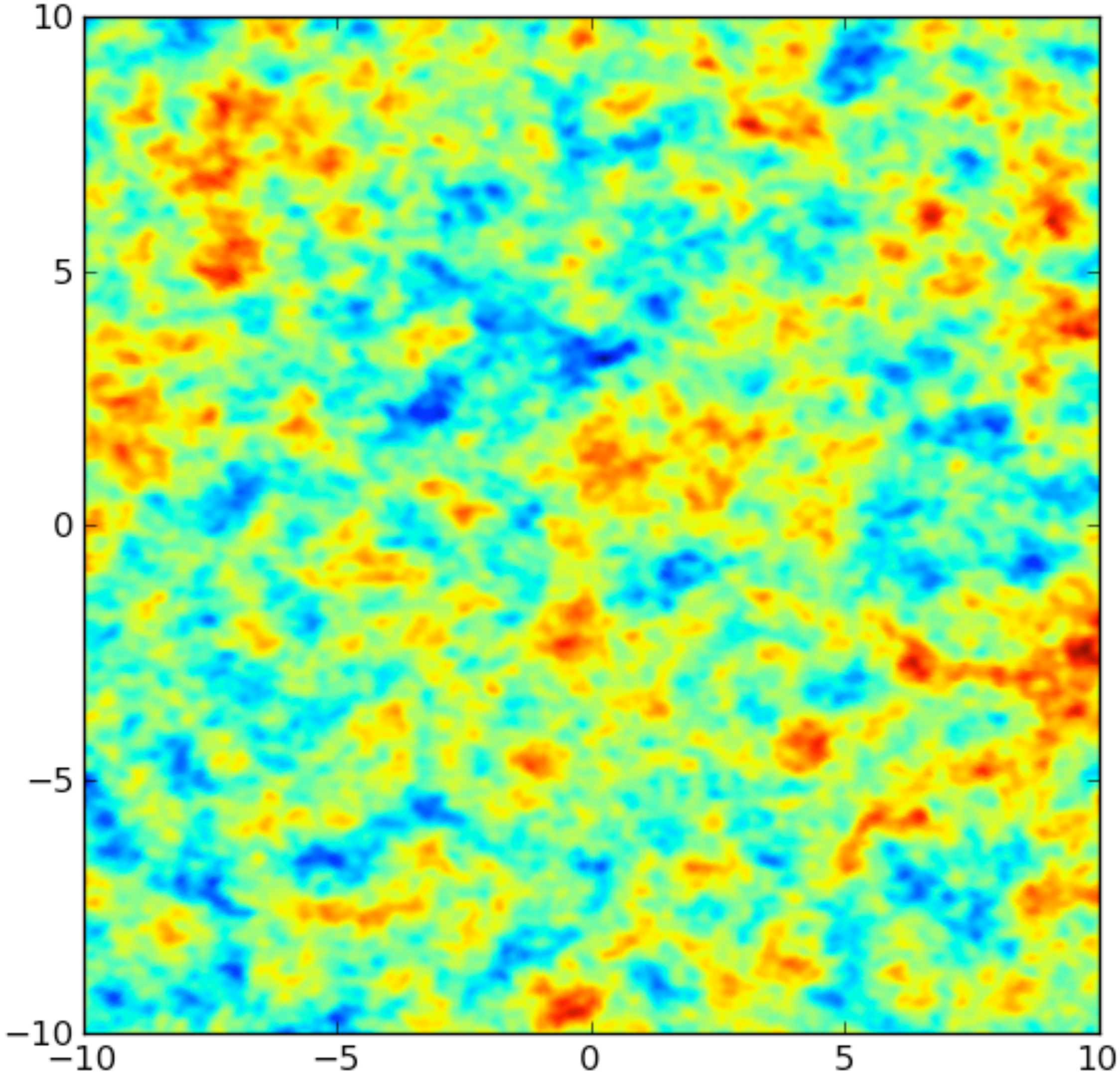
Planck can tell the difference between Λ and spatial curvature. The distance degeneracy is broken by CMB lensing, which I'll describe in the next few slides.



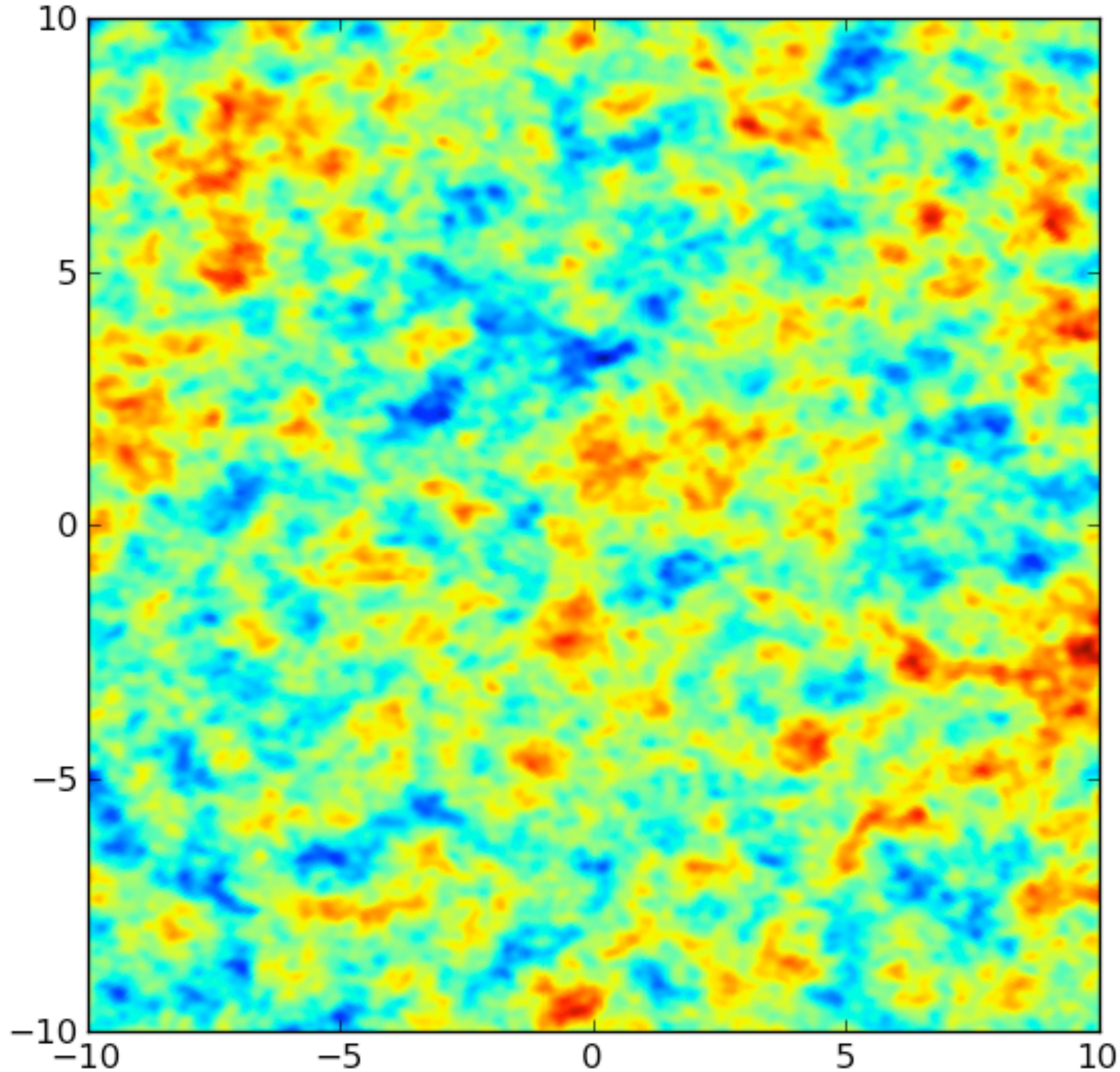
Lensing moves existing temperature fluctuations around, but does not generate new anisotropy (lensing conserves surface brightness)

Shown exaggerated here! Actual lensing deflections are a few arcminutes.

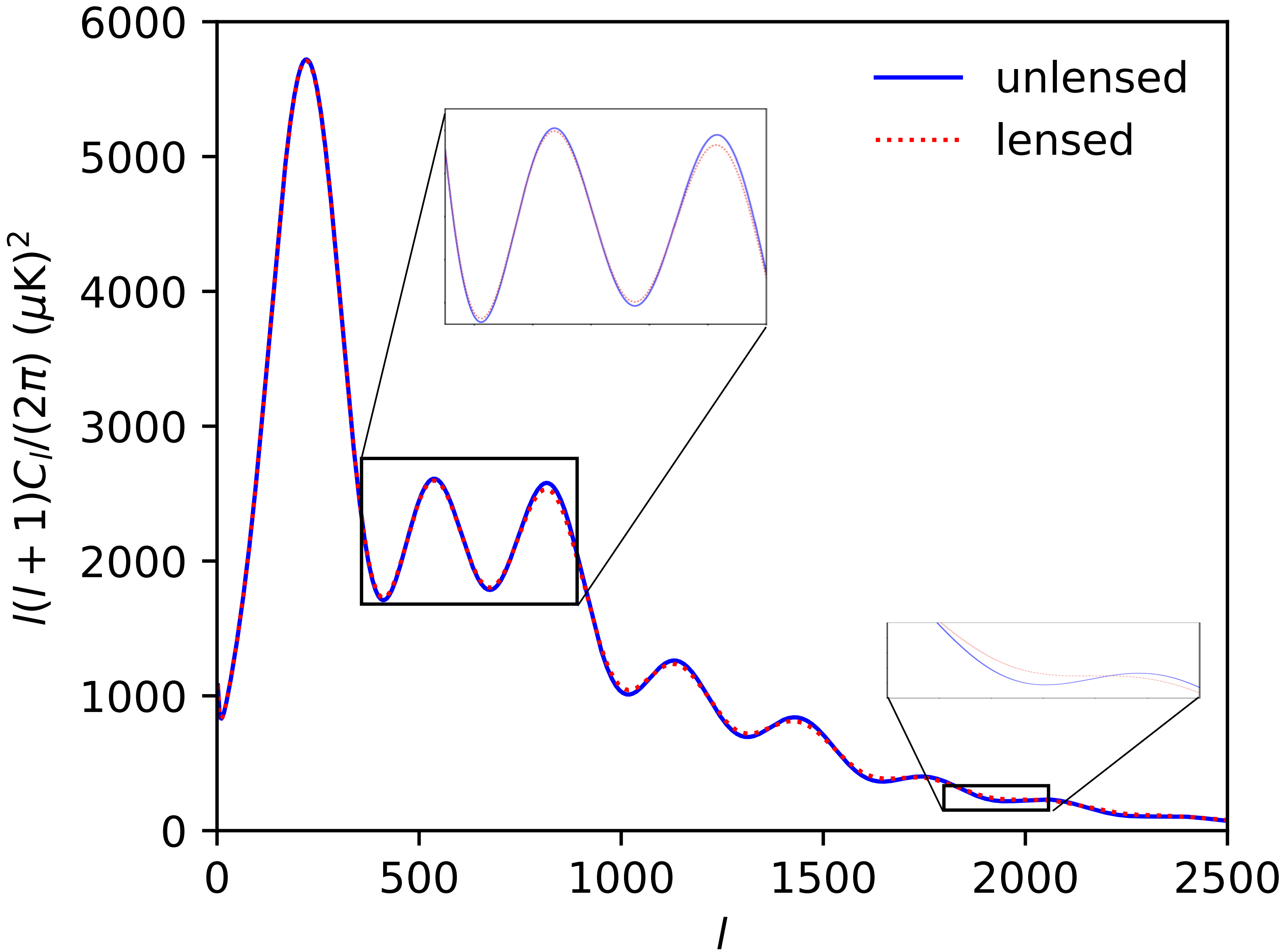
Unlensed vs lensed CMB



Unlensed vs lensed CMB

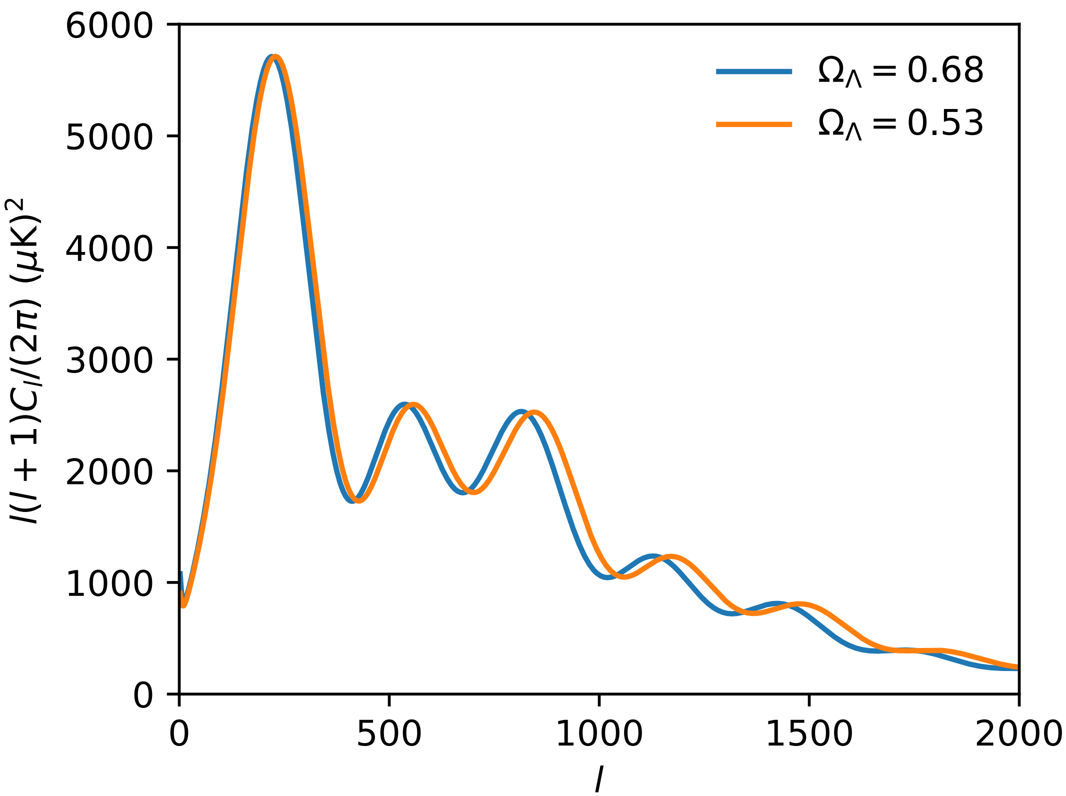


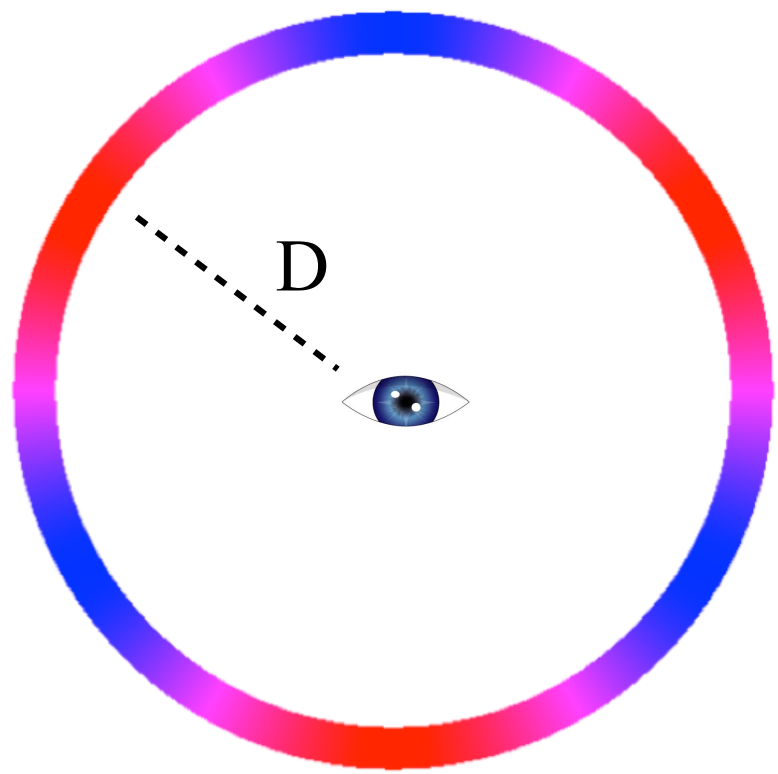
Effect of CMB lensing on the temperature power spectrum.
Lensing smooths peaks and adds power on small scales (high l).
Breaks the distance degeneracy (plot forthcoming in a few slides).



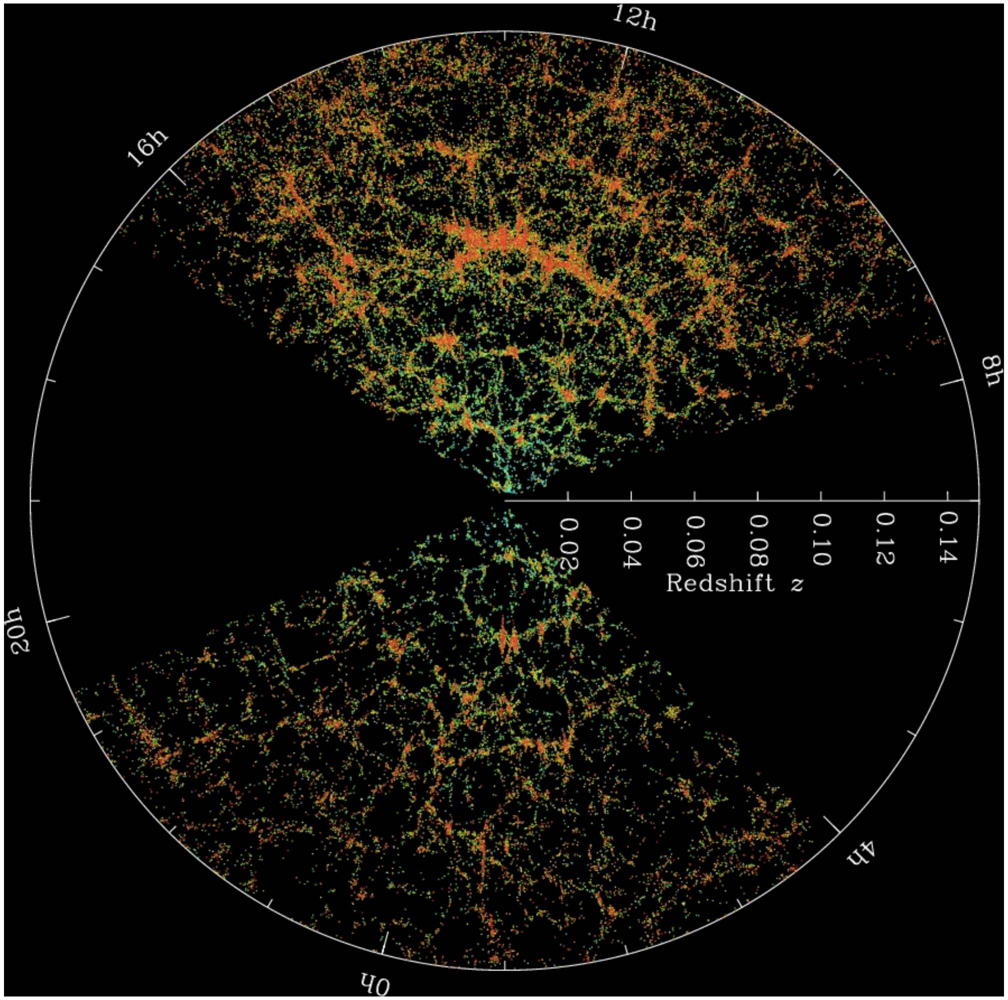
Another way to break the CMB distance degeneracy is by using non-CMB datasets, especially BAO (baryon acoustic oscillations) measurements in galaxy surveys.

First, some terminology: the CMB acoustic peak scale $\ell_a \sim 200$ can be interpreted as a "standard ruler".



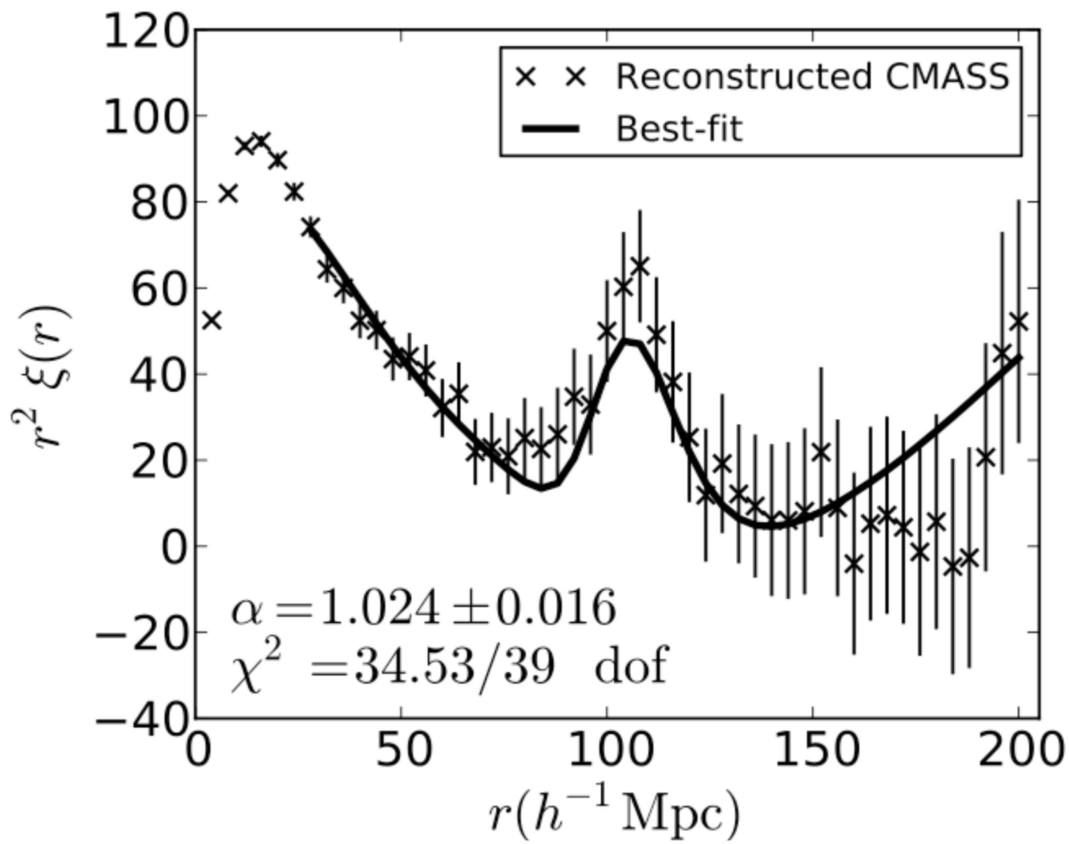


A galaxy survey measures the number density of galaxies throughout the universe. (A 3D field, since redshifts and angular locations are measured.)



The 3D power spectrum of the galaxy density field can be used to constrain cosmological parameters. (In galaxy surveys, the correlation function ζ(r) is usually used instead of the power spectrum P(k)).

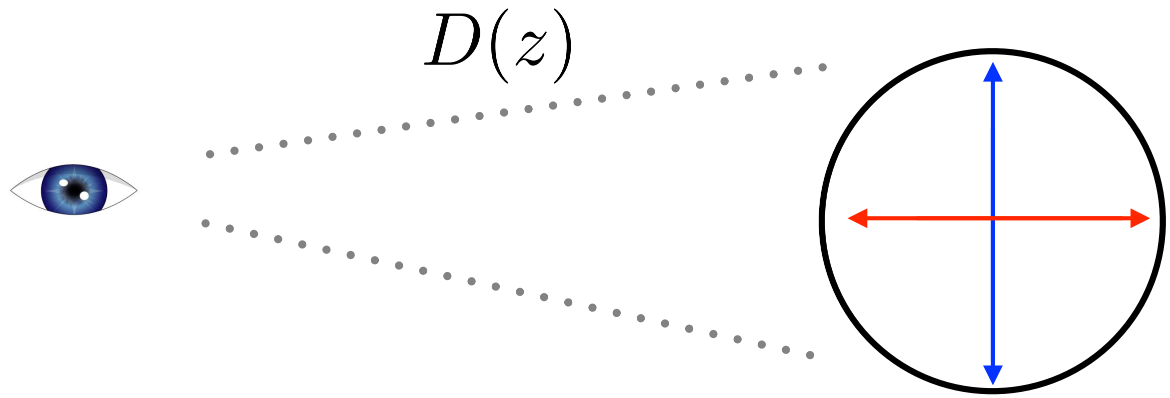
Like the CMB, the correlation function ζ(r) contains an acoustic feature which can be used as a "standard ruler".



SDSS (2012)

More precisely, there are two standard rulers:

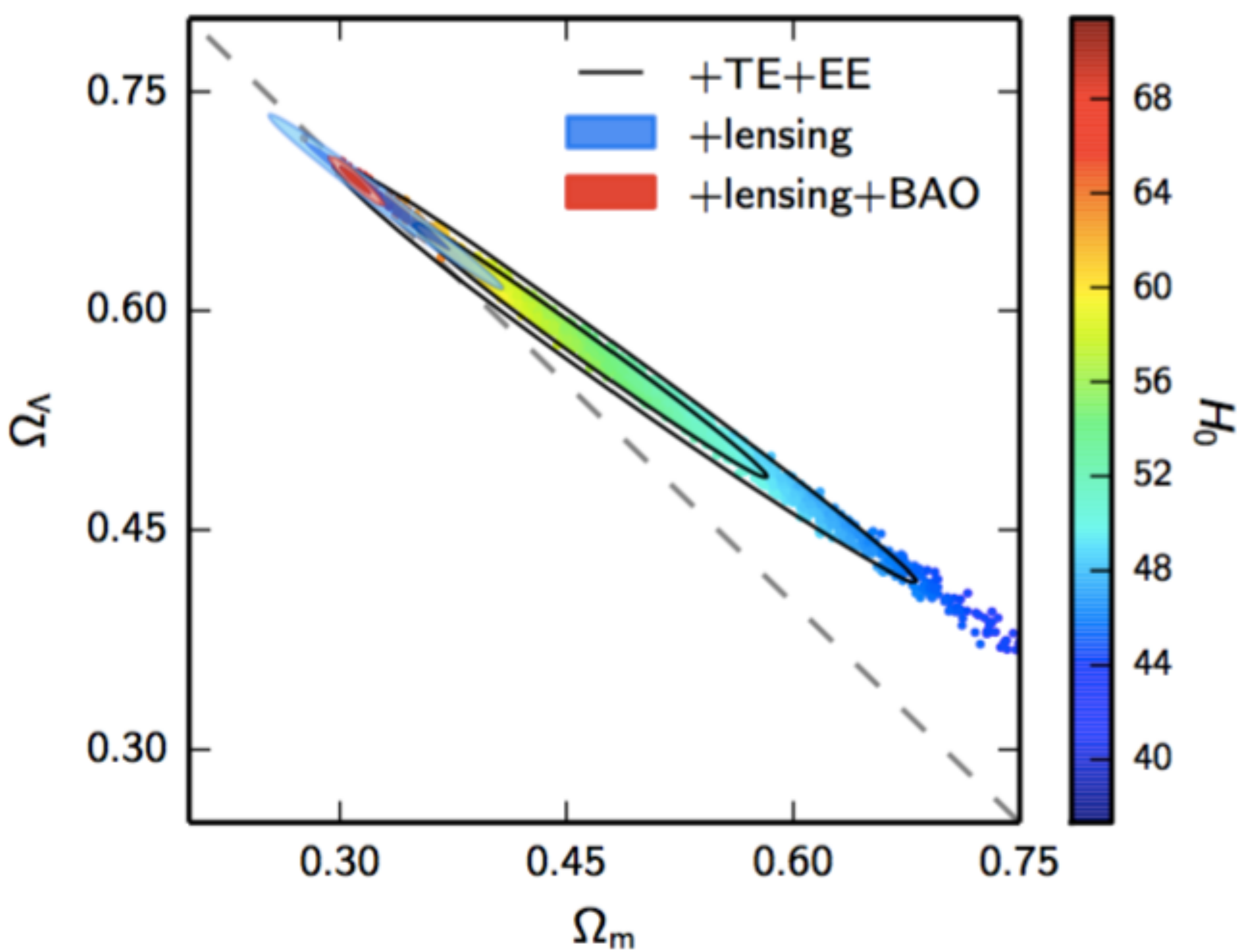
- "Transverse" observation constrains D(z)
- "Radial" observation constrains H(z)



Furthermore, can measure D(z) and H(z) as functions of z!
(Unlike the CMB, where there is only one source redshift.)

Galaxy BAO measurements are very powerful for constraining expansion history and breaking degeneracies between parameters.

Black ellipses: unlensed CMB
Blue ellipses: lensed CMB
Red ellipses: lensed CMB + BAO



Planck 2015

Current constraints on expansion history:

$\Omega_\Lambda = 0.691 \pm 0.006$ cosmological constant

$\Omega_K = 0.000 \pm 0.005$ spatial curvature

$w = -1.02 \pm 0.08$ dark energy equation of state

The last parameter (w) parameterizes the time dependence of dark energy density as $\rho \sim a^{-3(1+w)}$, i.e. $w = -1$ corresponds to a cosmological constant.

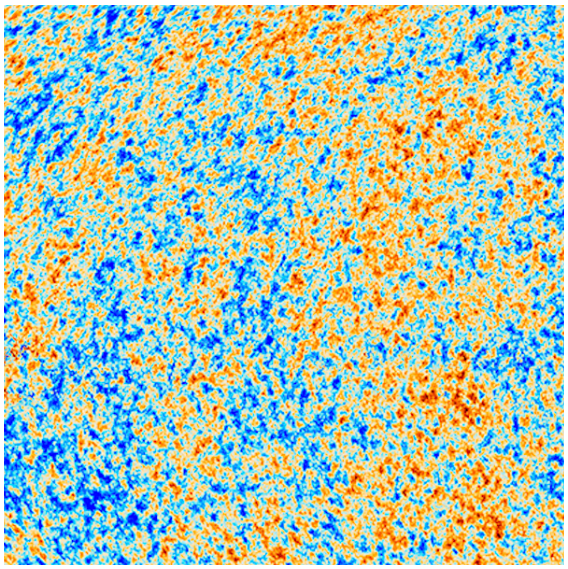
i.e. the DE energy density at z=1 is uncertain by 24% (at 1σ)!

Good news: these constraints will improve by an order of magnitude in the not-too-distant future (mainly from better BAO data).

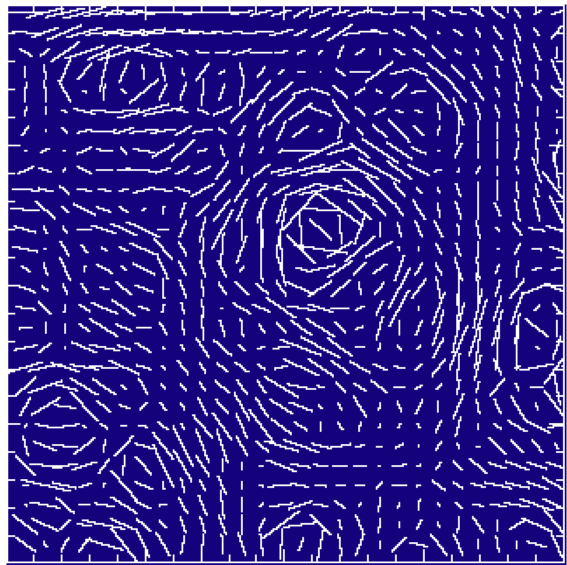
Bad news: there is no natural threshold for these parameters, so this is a fishing expedition (as far as I know!)

CMB polarization

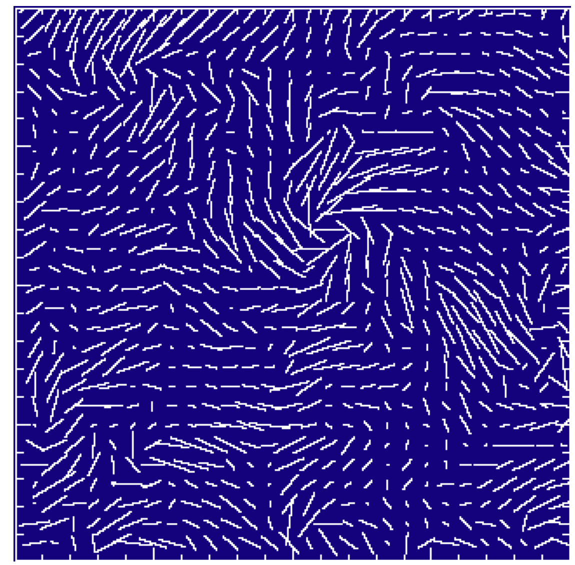
Temperature



E-mode linear polarization



B-mode linear polarization



E/B decomposition of linear polarization (traceless symmetric tensor) is similar to gradient/curl decomposition of vector field.

B-modes are different

Theorem: (scalar sources) + (linear perturbation theory)
$\Rightarrow$ (no B-modes are generated)

Rephrased: B-modes only arise from

- Primordial gravity waves (non-scalar sources)
- Second-order effects (largest by far is CMB lensing)

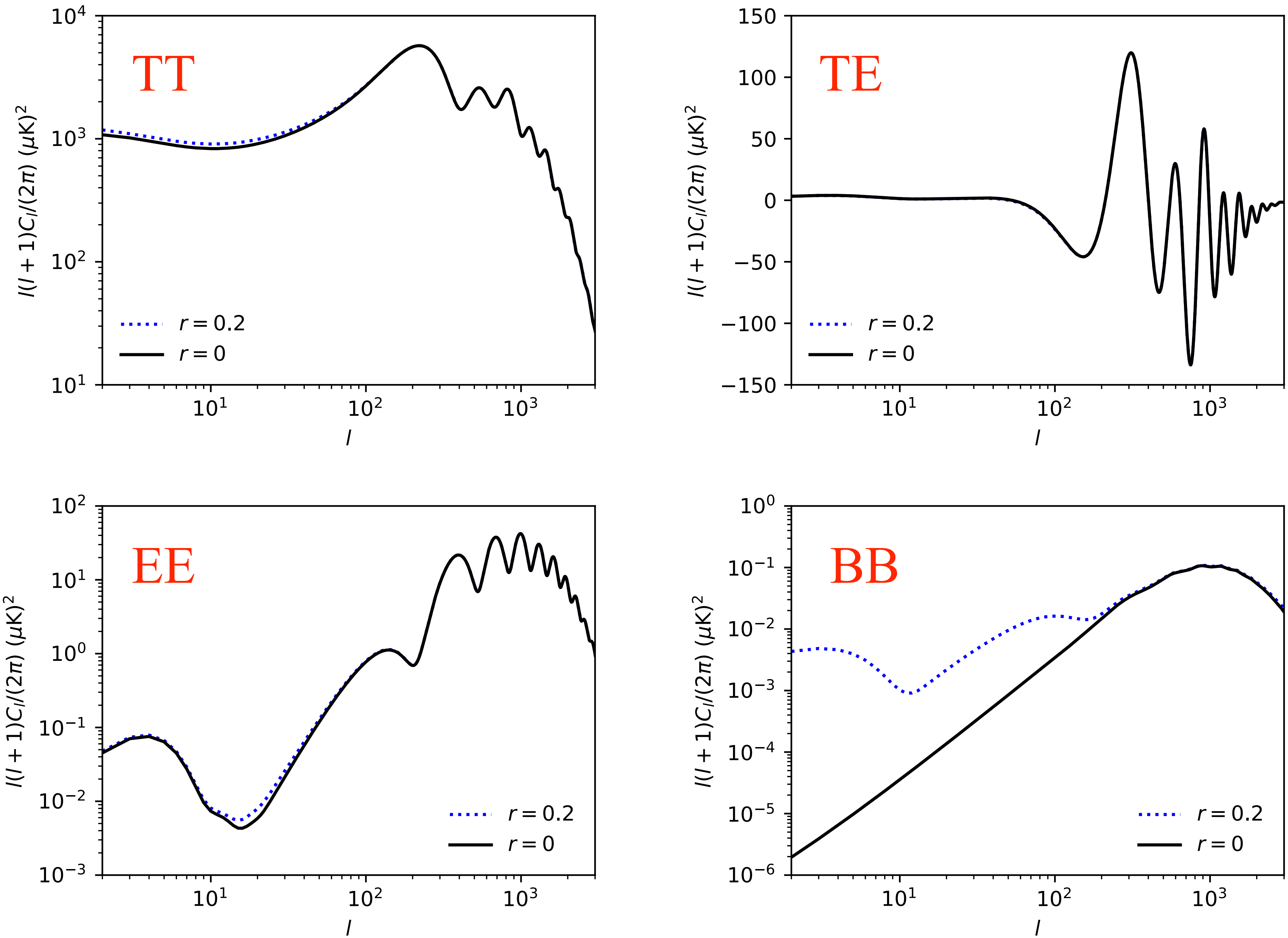
Some models of inflation predict primordial gravity waves.

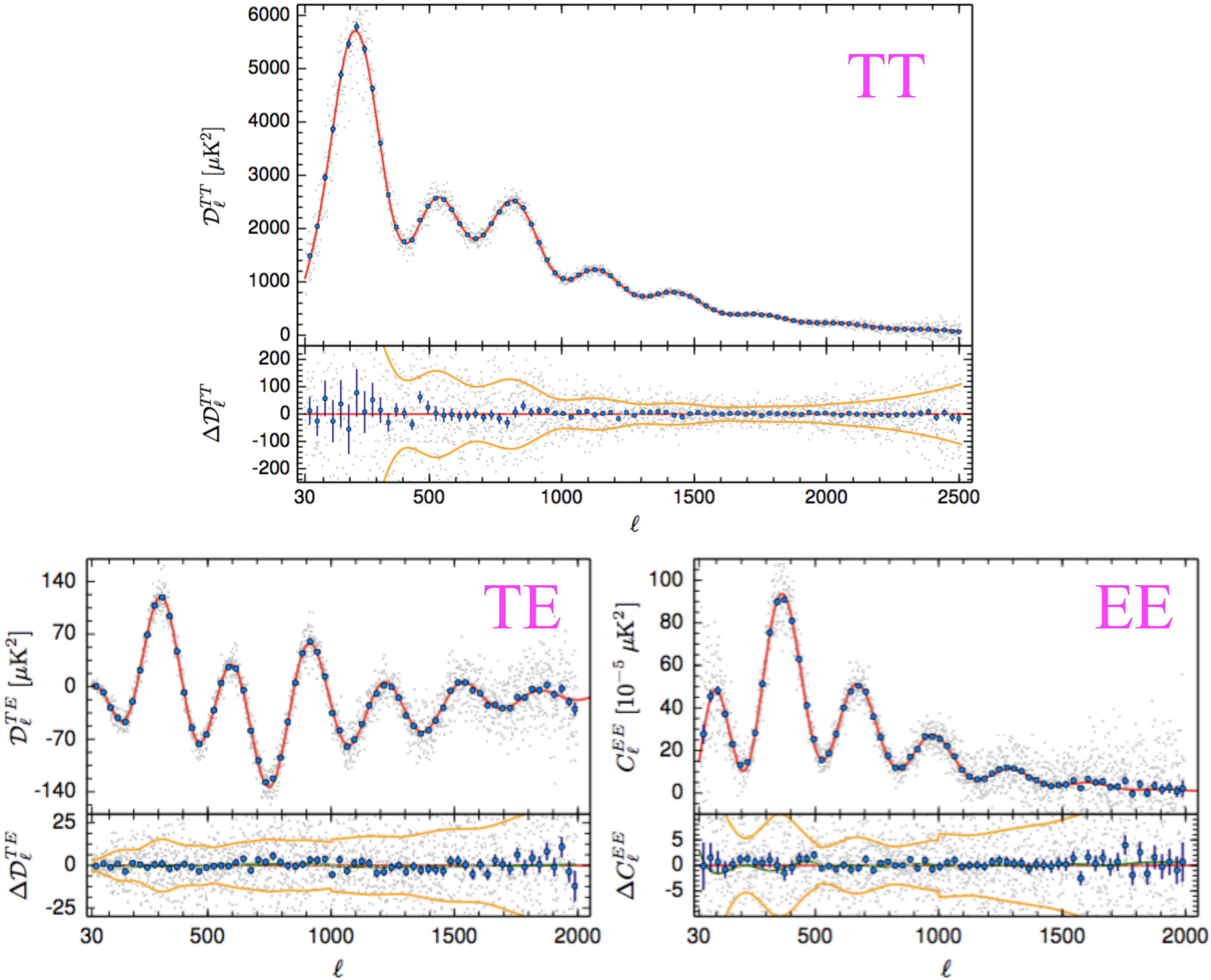
$$ds^2 = -dt^2 + a(t)^2 e^{2\zeta(x)}(\delta_{ij} + h_{ij}(x))$$

Parameterized by "tensor-to-scalar" ratio $r = \frac{P_h(k)}{P_\zeta(k)}$

Current upper limit: r < 0.12 (95% CL), mainly from CMB temperature.

CMB temperature and polarization power spectra



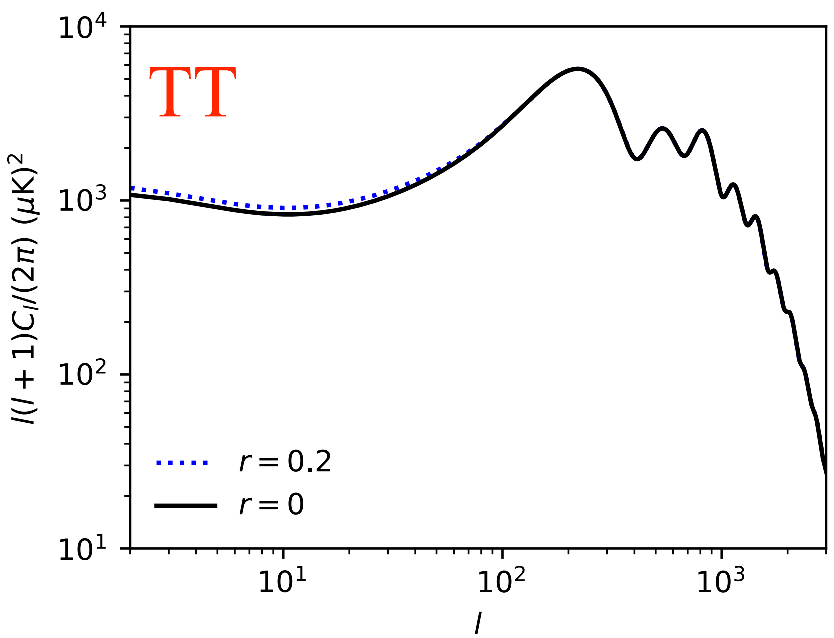


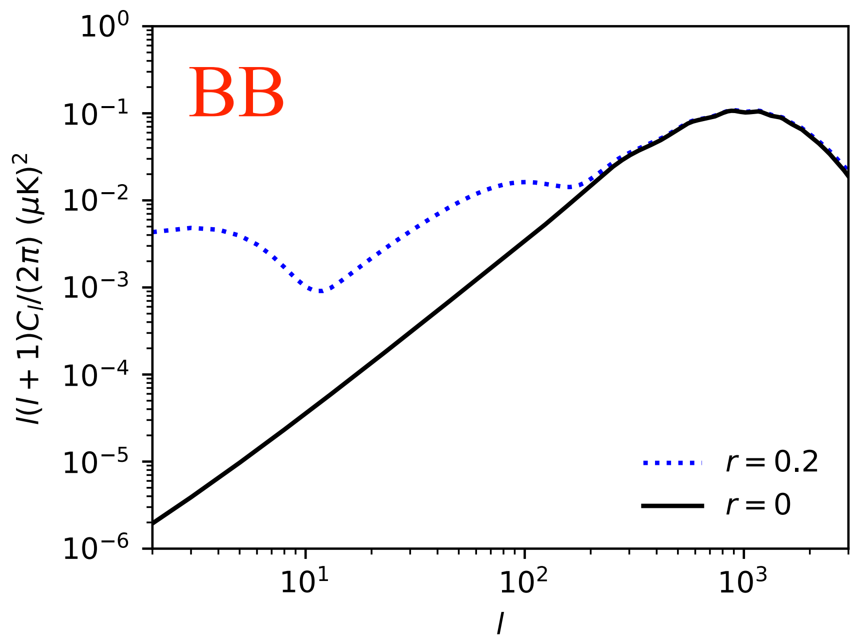
TT
$\mathcal{D}_\ell^{TT}$ [μK^2]
$\Delta\mathcal{D}_\ell^{TT}$
6000
5000
4000
3000
2000
1000
0
200
100
0
-100
-200
30
500
1000
1500
2000
2500
ℓ
TE
$\mathcal{D}_\ell^{TE}$ [μK^2]
$\Delta\mathcal{D}_\ell^{TE}$
140
70
0
-70
-140
25
0
-25
30
500
1000
1500
2000
ℓ
EE
C_ℓ^{EE} [10^{-5} μK^2]
ΔC_ℓ^{EE}
100
80
60
40
20
0
5
0
-5
30
500
1000
1500
2000
ℓ

Current gravity wave constraint: $r < 0.12$ (95% CL)

I predict that:

- In the next few years, BB will be measured much better, and the limit will improve to $r \sim 10^{-2}$
- On a ~20 year timescale, the limit will be $r \sim 10^{-3}$ or 10^{-4}.





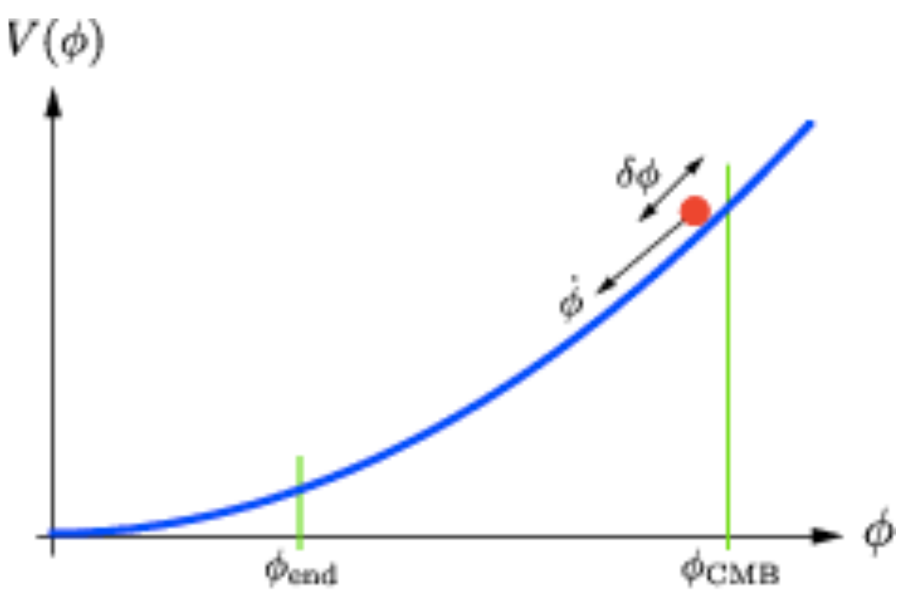
A detection (or non-detection) of r at the $\sim 10^{-2}$ level would be very informative!

E.g. consider single-field slow-roll inflation on potential V(ϕ).

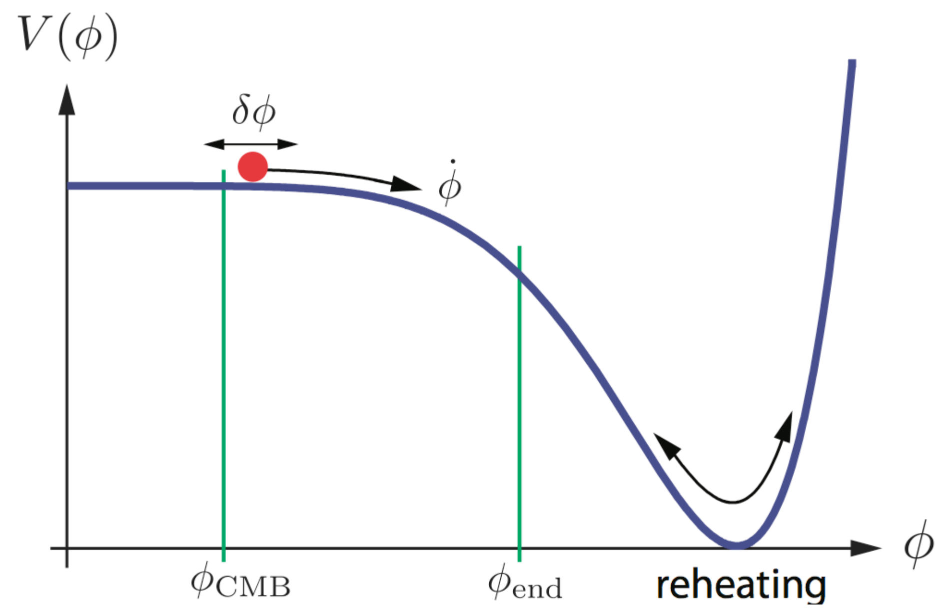
$$S = \int d^4x \sqrt{-g} \left(\frac{M_{\rm Pl}^2}{2} R - \frac{1}{2} (\nabla_\mu \phi)^2 - V(\phi) \right)$$

In the next few slides, I'll explain the following picture:

Case 1: r is detectable



Case 2: r is very small



Single-field slow-roll inflation: 1-slide theory review

For inflation to occur, the “slow-roll” parameters:

$$\epsilon = \frac{M_{\rm Pl}^2}{2}\left(\frac{V'(\phi)}{V(\phi)}\right)^2 \qquad \eta = M_{\rm Pl}^2\left(\frac{V''(\phi)}{V(\phi)}\right)$$

must both be << 1. Some key results:

$$r = 16\epsilon$$

$$n_s - 1 = -6\epsilon + 2\eta$$

$$V^{1/4} = r^{1/4}\left(\frac{3\pi^2\Delta_\zeta^2}{2}\right)^{1/4} M_{\rm Pl}$$ [energy scale of inflation]

Plug in $n_s \sim 0.97$ and $\Delta\zeta^2 \sim (2 \times 10^{-9})$ to the last two equations:

$$-6\epsilon + 2\eta = -0.03$$

$$V^{1/4} = r^{1/4} \times (3.23 \times 10^{16}\ \mathrm{GeV})$$

$$r = 16\epsilon$$

$$-6\epsilon + 2\eta = -0.03$$

$$V^{1/4} = r^{1/4} \times (3.23 \times 10^{16}\ \mathrm{GeV})$$

$$\epsilon = \frac{M_{\mathrm{Pl}}^2}{2} \left(\frac{V'(\phi)}{V(\phi)} \right)^2$$

$$\eta = M_{\mathrm{Pl}}^2 \left(\frac{V''(\phi)}{V(\phi)} \right)$$

$$r = 16\epsilon$$

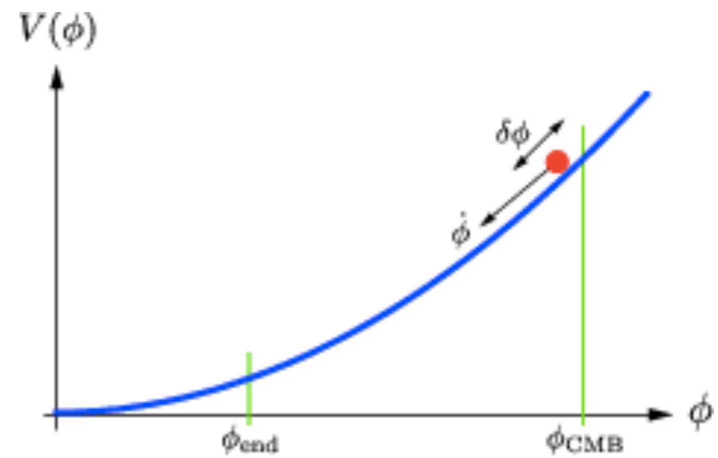
$$-6\epsilon + 2\eta = -0.03$$

$$V^{1/4} = r^{1/4} \times (3.23 \times 10^{16}\ \mathrm{GeV})$$

$$\epsilon = \frac{M_{\mathrm{Pl}}^2}{2}\left(\frac{V'(\phi)}{V(\phi)}\right)^2$$

$$\eta = M_{\mathrm{Pl}}^2\left(\frac{V''(\phi)}{V(\phi)}\right)$$

Scenario 1: ε and η are both of order ~10^{-3}-10^{-2}. Energy scale of inflation must be in a narrow range near 10^{16} GeV.



$$r = 16\epsilon$$

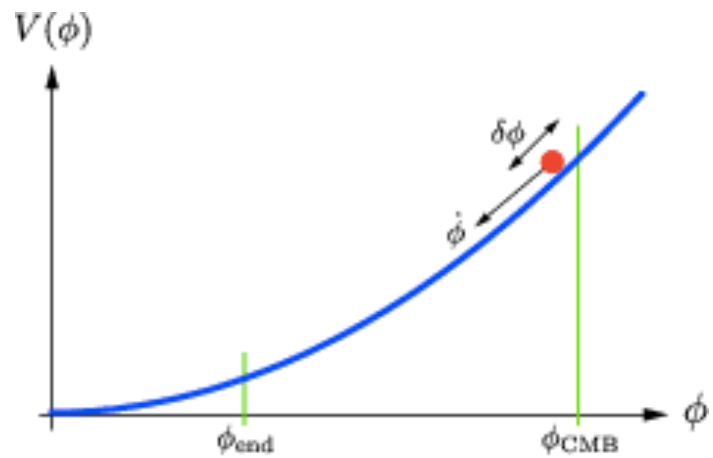
$$-6\epsilon + 2\eta = -0.03$$

$$V^{1/4} = r^{1/4} \times (3.23 \times 10^{16}\ \text{GeV})$$

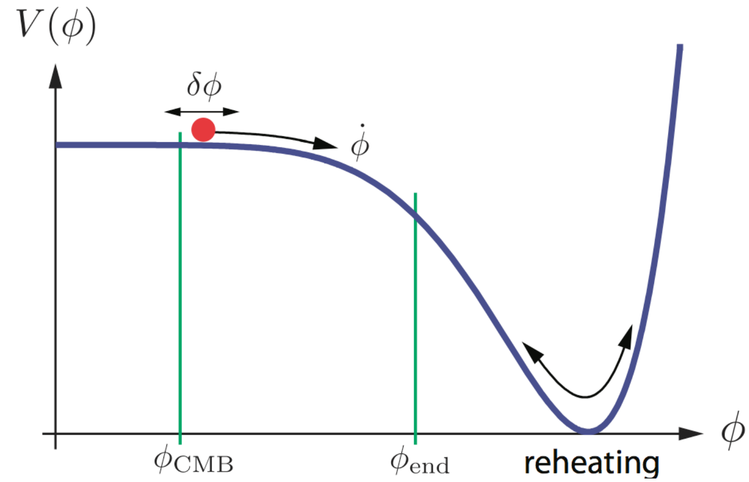
$$\epsilon = \frac{M_{\rm Pl}^2}{2}\left(\frac{V'(\phi)}{V(\phi)}\right)^2$$

$$\eta = M_{\rm Pl}^2\left(\frac{V''(\phi)}{V(\phi)}\right)$$

Scenario 1: ε and η are both of order ~10^{-3}-10^{-2}. Energy scale of inflation must be in a narrow range near 10^{16} GeV.



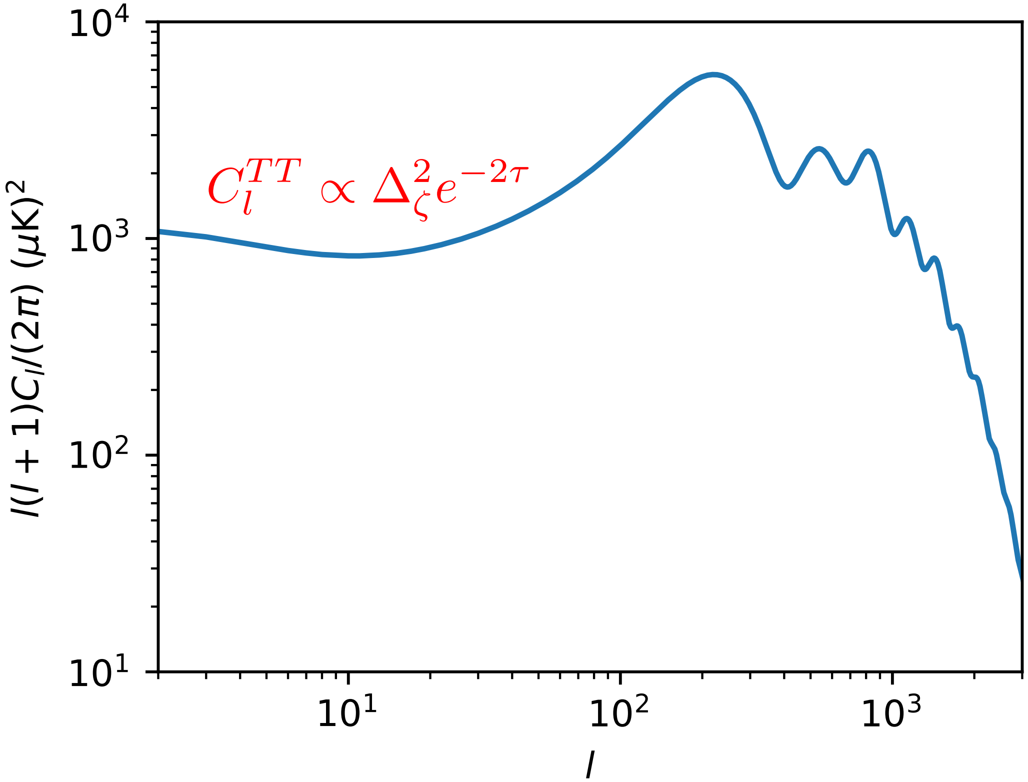
Scenario 2: ε is much smaller than η. (First derivative of potential is very small.) Energy scale of inflation is anywhere between ~10 TeV and ~10^{16} GeV.



The optical depth degeneracy in the CMB.

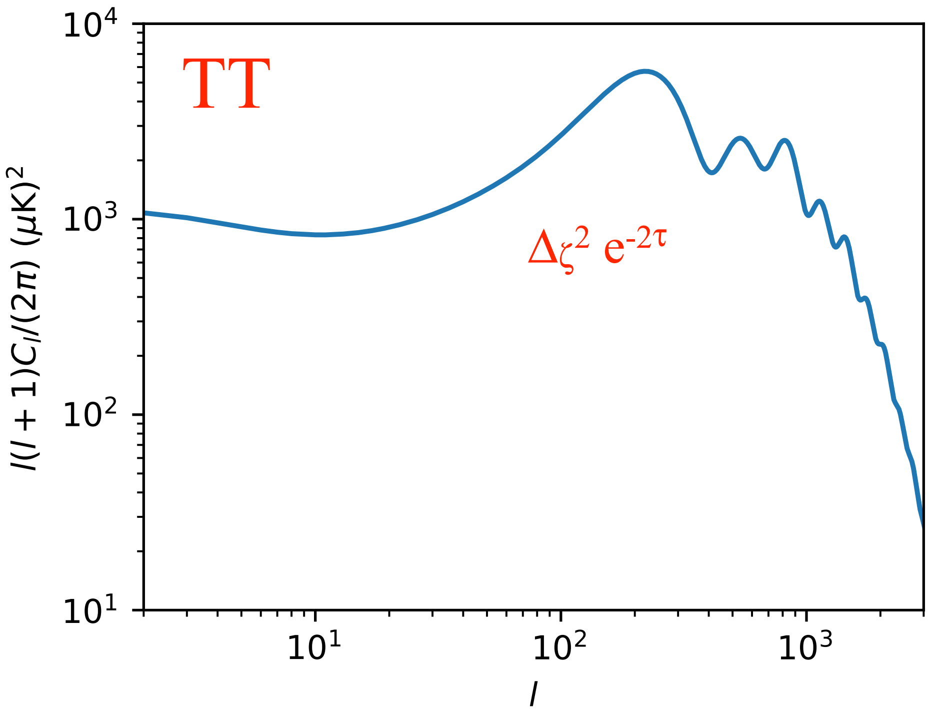
Consider the unlensed TT power spectrum.

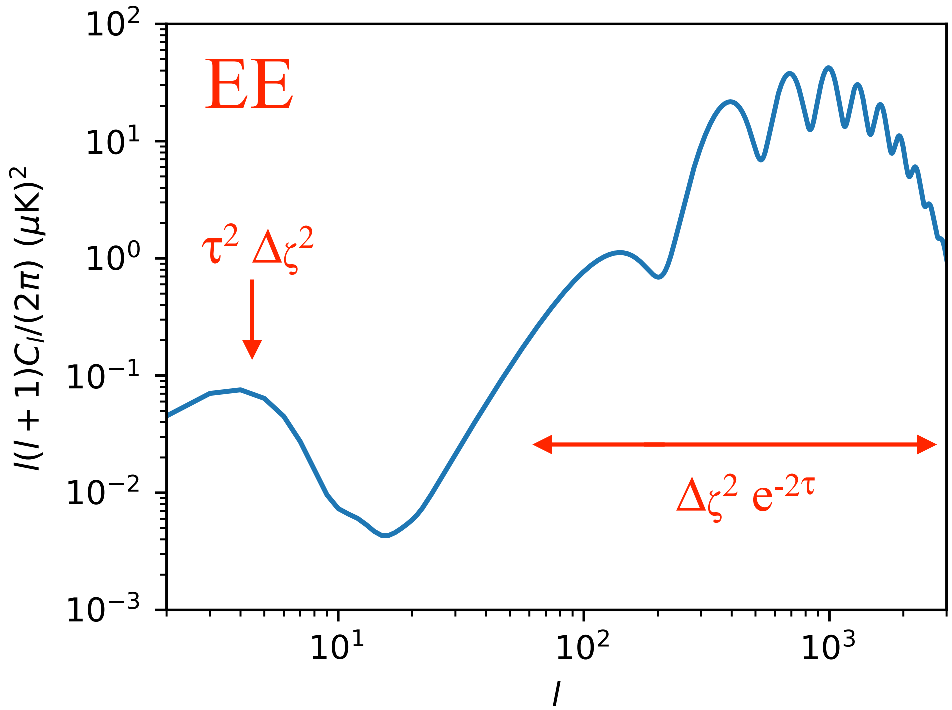
The parameters Δ_ζ^2 and τ are degenerate: they only affect the CMB via the combination ($\Delta_\zeta^2 \, e^{-2\tau}$), which sets the height of the peaks.



The CMB optical depth degeneracy can be broken through either:

- Gravitational lensing, since the amount of lensing in the late universe depends on Δ_ζ^2 but not τ.
- EE at very low l, which is proportional to ($\tau^2 \Delta_\zeta^2$).





Cosmology and neutrino mass

Neutrino oscillation experiments measure Δm_ν^2 between species

$$\Delta m_{31}^2 = (0.049 \pm 0.0012 \text{ eV})^2$$
$$\Delta m_{21}^2 = (0.0087 \pm 0.00013 \text{ eV})^2$$

Cosmology is complementary: mainly sensitive to $\sum_\nu m_\nu$

Neutrino mass suppresses gravitational lensing (and other probes of large-scale structure) in the late universe.

Very schematically:

(Amount of lensing) ~ Δ_ζ^2 * (complicated function of m_ν)

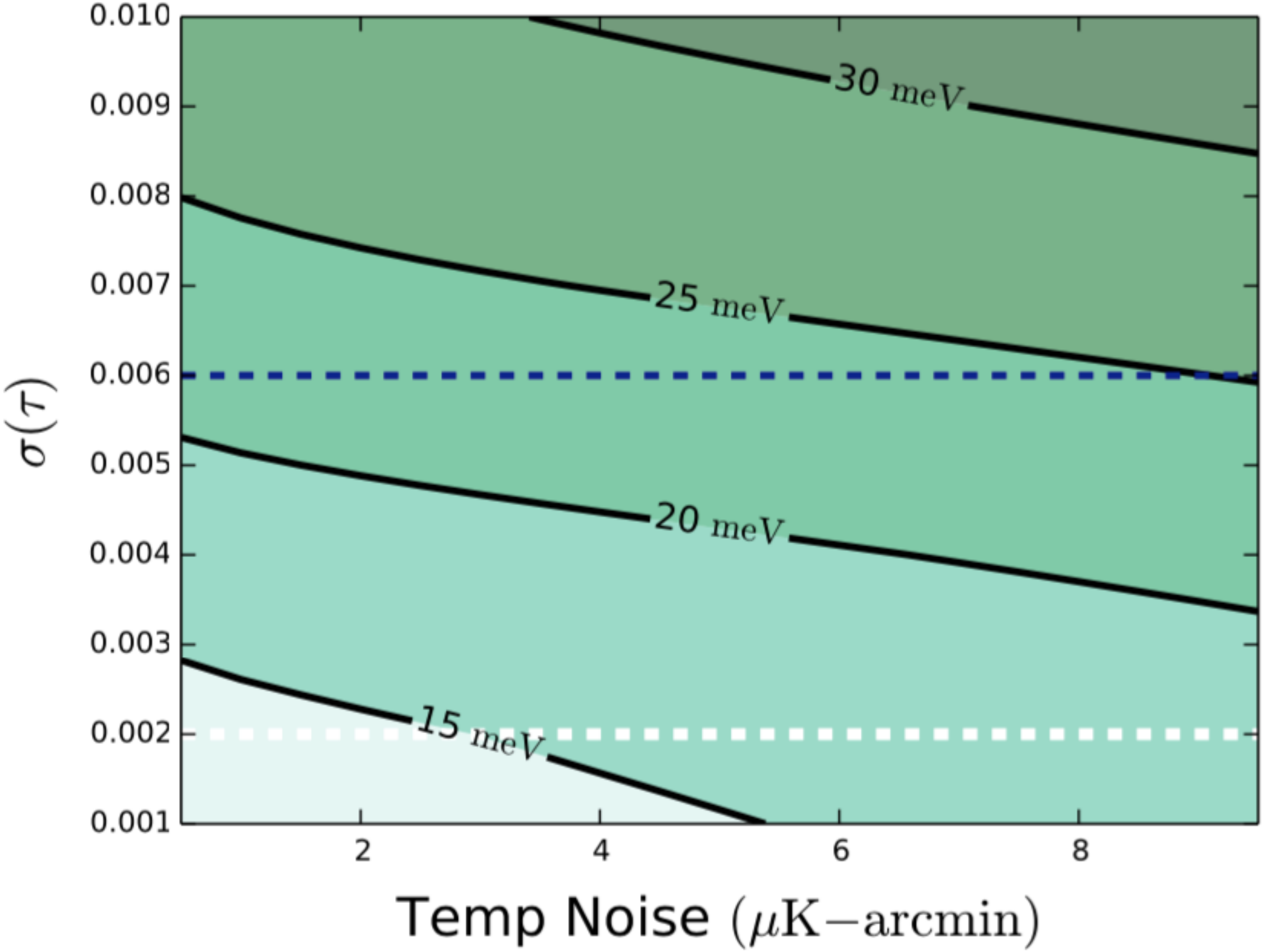
Cosmology and neutrino mass

Consequence: to constrain neutrino mass, we need both lensing and low-l EE to break the CMB optical depth degeneracy.

(Peak heights) ~ ($\Delta_\zeta^2 e^{-2\tau}$)
(Low-l EE amplitude) ~ ($\tau^2 \Delta_\zeta^2$)
(Amount of lensing) ~ Δ_ζ^2 * (complicated function of m_ν)

This may be a serious problem, since current Planck EE measurements may become a limiting factor, and it is not clear how to improve them.

(Difficult to measure low-l EE from ground-based telescopes, and there is not a post-Planck CMB satellite on the horizon.)



CMB-S4 Science Book

Extra light degrees of freedom

The CMB is sensitive to the total energy density of radiation (=relativistic species) during the epoch of CMB formation.

This is usually quoted as an "effective" number of neutrino species N_{eff}, but any relativistic species (neutrino or otherwise) will contribute!

The contribution depends on:

- spin of particle
- boson vs fermion
- time of decoupling from thermal plasma (species which decouple early contribute less)

Goldstone Bosons as Fractional Cosmic Neutrinos

Steven Weinberg*
Theory Group, Department of Physics, University of Texas
Austin, TX, 78712

Abstract

It is suggested that Goldstone bosons may be masquerading as fractional cosmic neutrinos, contributing about 0.39 to what is reported as the effective number of neutrino types in the era before recombination. The broken symmetry associated with these Goldstone bosons is further speculated to be the conservation of the particles of dark matter.

arXiv:1305.1971

Current constraint: $N_{eff} = 3.04 \pm 0.18$

Futuristic CMB constraint: $\Delta N_{eff} \sim 0.02$

Could discover new particles with the CMB!

Thanks!