

23. Inflation

Revised August 2025 by J. Ellis (King’s Coll. London; CERN), V. Vennin (LPENS; CNRS) and D. Wands (Portsmouth U.).

23.1 Motivation and Introduction

The standard Big-Bang model of cosmology provides a successful framework in which to understand the thermal history of our Universe and the growth of cosmic structure, but it is essentially incomplete. As described in Secbigbang:sec:problems, Big-Bang cosmology requires very specific initial conditions. It postulates a uniform cosmological background, described by a spatially-flat, homogeneous and isotropic Robertson-Walker (RW) metric (Eq. (22.1) in “Big-Bang Cosmology” review), with scale factor $R(t)$. Within this setting, it also requires an initial almost scale-invariant distribution of primordial density perturbations as seen, for example, in the cosmic microwave background (CMB) radiation (described in Chap. 29, “Cosmic Microwave Background” review), on scales far larger than the causal horizon at the time the CMB photons last scattered.

The Hubble expansion rate, $H \equiv \dot{R}/R$, in an RW cosmology is given by the Friedmann constraint equation (Eq. (22.8) in “Big-Bang Cosmology” review)

$$H^2 = \frac{8\pi\rho}{3M_{\text{P}}^2} + \frac{\Lambda}{3} - \frac{k}{R^2}, \quad (23.1)$$

where k/R^2 is the intrinsic spatial curvature. We use natural units such that the speed of light $c = 1$ and hence we have the Planck mass¹ $M_{\text{P}} = G_{\text{N}}^{-1/2} \simeq 10^{19}$ GeV (see “Astrophysical Constants and Parameters”). A cosmological constant, Λ , of the magnitude required to accelerate the Universe today (see Chap. 28, “Dark Energy” review) would have been completely negligible in the early Universe where the energy density $\rho \gg M_{\text{P}}^2 \Lambda \sim 10^{-12}(\text{eV})^4$. The standard early Universe cosmology, described in Sec. 22.1.5 in “Big-Bang Cosmology” review, is thus dominated by non-relativistic matter ($p_{\text{m}} \ll \rho_{\text{m}}$) or radiation ($p_{\text{r}} = \rho_{\text{r}}/3$ for an isotropic distribution). This leads to a decelerating expansion with $\ddot{R} < 0$.

The hypothesis of inflation [5,6] postulates a period of accelerated expansion, $\ddot{R} > 0$, in the very early Universe, preceding the standard radiation-dominated era, which offers a physical model for the origin of these initial conditions, as reviewed in Refs. [1–4, 7–9]. Such a period of accelerated expansion (i) drives a curved Robertson-Walker spacetime (with spherical or hyperbolic spatial geometry) towards spatial flatness, and (ii) it also expands the causal horizon beyond the present Hubble length, so as to encompass all the scales relevant to describe the large-scale structure observed in our Universe today, via the following two mechanisms.

1. A spatially-flat Universe with vanishing spatial curvature, $k = 0$, has the dimensionless density parameter $\Omega_{\text{tot}} = 1$, where we define (Eq. (22.13) in “Big-Bang Cosmology” review; see Chap. 25, “Cosmological Parameters” review for more complete definitions)

$$\Omega_{\text{tot}} \equiv \frac{8\pi\rho_{\text{tot}}}{3M_{\text{P}}^2 H^2}, \quad (23.2)$$

with $\rho_{\text{tot}} \equiv \rho + \Lambda M_{\text{P}}^2/8\pi$. If we re-write the Friedmann constraint (Eq. (23.1)) in terms of Ω_{tot} we have

$$1 - \Omega_{\text{tot}} = -\frac{k}{\dot{R}^2}. \quad (23.3)$$

¹We note that another commonly used convention in the literature (*e.g.*, Refs. [1–4]) is to work in terms of the reduced Planck mass, $(8\pi G_{\text{N}})^{-1/2} \simeq 2 \times 10^{18}$ GeV.

Observations require $|1 - \Omega_{\text{tot},0}| < 0.005$ today [10], where the subscript "0" denotes the present-day value. Taking the time derivative of Eq. (23.3) we obtain

$$\frac{d}{dt}(1 - \Omega_{\text{tot}}) = -2\frac{\ddot{R}}{\dot{R}}(1 - \Omega_{\text{tot}}). \quad (23.4)$$

Thus in a decelerating expansion, $\dot{R} > 0$ and $\ddot{R} < 0$, any small initial deviation from spatial flatness grows, $(d/dt)|1 - \Omega_{\text{tot}}| > 0$. A small value such as $|1 - \Omega_{\text{tot},0}| < 0.005$ today requires an even smaller value at earlier times, *e.g.*, $|1 - \Omega_{\text{tot}}| < 10^{-5}$ at the last scattering of the CMB, which appears unlikely, unless for some reason space is exactly flat. However, an extended period of accelerated expansion in the very early Universe, with $\dot{R} > 0$ and $\ddot{R} > 0$ and hence $(d/dt)|1 - \Omega_{\text{tot}}| < 0$, can drive Ω_{tot} sufficiently close to unity, so that $|1 - \Omega_{\text{tot},0}|$ remains unobservably small today, even after the radiation- and matter-dominated eras, for a wide range of initial values of Ω_{tot} .

2. The comoving distance (the present-day proper distance) traversed by light between cosmic time t_1 and t_2 in an expanding Universe can be written, (see Eq. (22.32) in “Big-Bang Cosmology” review), as

$$D_0(t_1, t_2) = R_0 \int_{t_1}^{t_2} \frac{dt}{R(t)} = R_0 \int_{\ln R_1}^{\ln R_2} \frac{d(\ln R)}{\dot{R}}. \quad (23.5)$$

In standard decelerated (radiation- or matter-dominated) cosmology the integrand, $1/\dot{R}$, decreases towards the past, and there is a finite comoving distance traversed by light (a particle horizon) since the Big Bang ($R_1 \rightarrow 0$). For example, the comoving size of the particle horizon at the CMB last-scattering surface ($R_2 = R_{\text{lss}}$) corresponds to $D_0 \simeq 100$ Mpc, or approximately 1° on the CMB sky today (see Sec. 22.2.4 in “Big-Bang Cosmology” review). However, during a period of inflation, $1/\dot{R}$ increases towards the past, and hence the integral (Eq. (23.5)) diverges as $R_1 \rightarrow 0$, allowing an arbitrarily large causal horizon, dependent only upon the duration of the accelerated expansion. Assuming that the Universe inflates with a finite Hubble rate H_* at $t_1 = t_*$, ending with $H_{\text{end}} < H_*$ at $t_2 = t_{\text{end}}$, we have

$$D_0(t_*, t_{\text{end}}) > \left(\frac{R_0}{R_{\text{end}}} \right) H_*^{-1} (e^{N_*} - 1), \quad (23.6)$$

where $N_* \equiv \ln(R_{\text{end}}/R_*)$ describes the duration of inflation, measured in terms of the logarithmic expansion (or “e-folds”) from $t_1 = t_*$ up to the end of inflation at $t_2 = t_{\text{end}}$, and R_0/R_{end} is the subsequent expansion from the end of inflation to the present day. If inflation occurs above the TeV scale, the comoving Hubble scale at the end of inflation, $(R_0/R_{\text{end}})H_{\text{end}}^{-1}$, is less than one astronomical unit ($\sim 10^{11}$ m), and a causally-connected patch can encompass our entire observable Universe today, which has a size $D_0 > 30$ Gpc, if there were more than 40 e-folds of inflation ($N_* > 40$). If inflation occurs at the GUT scale (10^{15} GeV) then we require more than 60 e-folds.

Producing an accelerated expansion in general relativity requires an energy-momentum tensor with negative pressure, $p < -\rho/3$ (see Eq. (22.9) in “Big-Bang Cosmology” review and Chap. 28, “Dark Energy” review), quite different from the hot dense plasma of relativistic particles in the hot Big Bang. However a positive vacuum energy $V > 0$ does exert a negative pressure, $p_V = -\rho_V$. The work done by the cosmological expansion must be negative in this case so that the local vacuum energy density remains constant in an expanding Universe, $\dot{\rho}_V = -3H(\rho_V + p_V) = 0$. Therefore, a false vacuum state can drive an exponential expansion, corresponding to a de Sitter spacetime with a constant Hubble rate $H^2 = 8\pi\rho_V/3M_{\text{P}}^2$ on spatially-flat hypersurfaces.

A constant vacuum energy V , equivalent to a cosmological constant Λ in the Friedmann equation Eq. (23.1), cannot provide a complete description of inflation in the early Universe, since inflation must necessarily have come to an end in order for the standard Big-Bang cosmology to follow. A phase transition to the present true vacuum is required to release the false vacuum energy into the energetic plasma of the hot Big Bang and produce the large total entropy of our observed Universe today. Thus, we must necessarily study dynamical models of inflation, where the time-invariance of the false vacuum state is broken by a time-dependent field. A first-order phase transition would produce a very inhomogeneous Universe [11], unless a time-dependent scalar field leads to a rapidly changing percolation rate [12–14]. However, a second-order phase transition [15, 16], controlled by a slowly-rolling scalar field, can lead to a smooth classical exit from the vacuum-dominated phase.

As a spectacular bonus, quantum fluctuations in that scalar field could provide a source of almost scale-invariant density fluctuations [17, 18], as detected in the CMB (see Chap. 29), which are thought to be the origin of the structures seen in the Universe today.

Accelerated expansion and primordial perturbations can also be produced in some modified gravity theories (*e.g.*, [5, 19]), which introduce additional non-minimally coupled degrees of freedom. Such inflation models can often be conveniently studied by transforming variables to an ‘Einstein frame’ in which Einstein’s equations apply with minimally coupled scalar fields [20–22].

In the following we will review scalar field cosmology in general relativity and the spectra of primordial fluctuations produced during inflation, before studying selected inflation models.

23.2 Scalar Field Cosmology

The energy-momentum tensor for a canonical scalar field ϕ with self-interaction potential $V(\phi)$ is given in Eq. (22.52) in the “Big-Bang Cosmology” review. In a homogeneous background this corresponds to a perfect fluid with density

$$\rho = \frac{1}{2}\dot{\phi}^2 + V(\phi), \quad (23.7)$$

and isotropic pressure

$$p = \frac{1}{2}\dot{\phi}^2 - V(\phi), \quad (23.8)$$

while the 4-velocity is proportional to the gradient of the field, $u^\mu \propto \nabla^\mu \phi$.

A field with vanishing potential energy acts like a stiff fluid with $p = \rho = \dot{\phi}^2/2$, whereas if the time-dependence vanishes we have $p = -\rho = -V$ and the scalar field is uniform in time and space. Thus a classical, potential-dominated scalar-field cosmology, with $p \simeq -\rho$, can naturally drive a quasi-de Sitter expansion; the slow time-evolution of the energy density weakly breaks the exact $O(1, 3)$ symmetry of four-dimensional de Sitter spacetime down to a Robertson-Walker (RW) spacetime, where the scalar field plays the role of the cosmic time coordinate.

In a scalar-field RW cosmology the Friedmann constraint equation (Eq. (23.1)) reduces to

$$H^2 = \frac{8\pi}{3M_{\text{Pl}}^2} \left(\frac{1}{2}\dot{\phi}^2 + V \right) - \frac{k}{R^2}, \quad (23.9)$$

while energy conservation (Eq. (22.10) in “Big-Bang Cosmology” review) for a homogeneous scalar field reduces to the Klein-Gordon equation of motion (Eq. (22.54) in “Big-Bang Cosmology” review)

$$\ddot{\phi} = -3H\dot{\phi} - V'(\phi). \quad (23.10)$$

The evolution of the scalar field is thus driven by the potential gradient $V' = dV/d\phi$, subject to damping by the Hubble expansion $3H\dot{\phi}$.

If we define the Hubble slow-roll parameter

$$\epsilon_H \equiv -\frac{\dot{H}}{H^2}, \quad (23.11)$$

then we see that inflation ($\ddot{R} > 0$ and hence $\dot{H} > -H^2$) requires $\epsilon_H < 1$. In this case the spatial curvature decreases relative to the scalar field energy density as the Universe expands. Hence, in the following we drop the spatial curvature and consider a spatially-flat RW cosmology, assuming that inflation has lasted sufficiently long that our observable Universe is very close to spatial flatness. However, we note that bubble nucleation, leading to a first-order phase transition during inflation, can lead to homogeneous hypersurfaces with a hyperbolic (‘open’) geometry, effectively resetting the spatial curvature inside the bubble [23]. This is the basis of so-called open inflation models [24–26], where inflation inside the bubble has a finite duration, leaving a finite negative spatial curvature.

In a scalar field-dominated cosmology, Eq. (23.11) gives

$$\epsilon_H = \frac{3\dot{\phi}^2}{2V + \dot{\phi}^2}, \quad (23.12)$$

in which case we see that inflation requires a potential-dominated expansion, $\dot{\phi}^2 < V$.

23.2.1 Slow-Roll Inflation

It is commonly assumed that the field acceleration term, $\ddot{\phi}$, in Eq. (23.10) can be neglected, in which case one can give an approximate solution for the inflationary attractor [27]. This slow-roll approximation reduces the second-order Klein-Gordon equation (23.10) to a first-order system, which is over-damped, with the potential gradient being approximately balanced against the Hubble damping:

$$3H\dot{\phi} \simeq -V', \quad (23.13)$$

and at the same time the Hubble expansion (Eq. (23.9)) is dominated by the potential energy

$$H^2 \simeq \frac{8\pi}{3M_{\text{P}}^2} V(\phi), \quad (23.14)$$

corresponding to $\epsilon_H \ll 1$.

A necessary condition for the validity of the slow-roll approximation is that the potential slow-roll parameters

$$\epsilon \equiv \frac{M_{\text{P}}^2}{16\pi} \left(\frac{V'}{V} \right)^2, \quad \eta \equiv \frac{M_{\text{P}}^2}{8\pi} \left(\frac{V''}{V} \right), \quad (23.15)$$

are small, *i.e.*, $\epsilon \ll 1$ and $|\eta| \ll 1$, requiring the potential to be correspondingly flat. If we identify V'' with the effective mass of the field, we see that the slow-roll approximation requires that the mass of the scalar field must be small compared with the Hubble scale. We note that the Hubble slow-roll parameter (Eq. (23.11)) coincides with the potential slow-roll parameter, $\epsilon_H \simeq \epsilon$, to leading order in the slow-roll approximation.

The slow-roll approximation allows one to determine the Hubble expansion rate as a function of the scalar field value, and vice versa. In particular, we can express, in terms of the scalar field value during inflation, the total logarithmic expansion, or number of “e-folds”:

$$\begin{aligned} N_* &\equiv \ln \left(\frac{R_{\text{end}}}{R_*} \right) \\ &= \int_{t_*}^{t_{\text{end}}} H dt \simeq - \int_{\phi_*}^{\phi_{\text{end}}} \sqrt{\frac{4\pi}{\epsilon}} \frac{d\phi}{M_{\text{P}}} \text{ for } V' > 0. \end{aligned} \quad (23.16)$$

Given that the slow-roll parameters are approximately constant during slow-roll inflation, $d\epsilon/dN \simeq 2\epsilon(\eta - 2\epsilon) = \mathcal{O}(\epsilon^2)$, we have

$$N_* \simeq \frac{4}{\sqrt{\epsilon}} \frac{\Delta\phi}{M_{\text{P}}}. \quad (23.17)$$

Since we require $N > 40$ to solve the flatness, horizon, and entropy problems of the standard Big-Bang cosmology, we require either very slow roll, $\epsilon < 0.01$, or a large change in the value of the scalar field relative to the Planck scale, $\Delta\phi > M_{\text{P}}$.

23.2.2 Reheating

Slow-roll inflation can lead to an exponentially large Universe, close to spatial flatness and homogeneity, but the energy density is locked in the potential energy of the scalar field, and needs to be converted to particles and thermalised to recover a hot Big-Bang cosmology at the end of inflation [28, 29]. This process is usually referred to as reheating, although there was not necessarily any preceding thermal era. Reheating can occur when the scalar field evolves towards the minimum of its potential, converting the potential energy first to kinetic energy. This can occur either through the breakdown of the slow-roll condition in single-field models, or due to an instability triggered by the inflaton reaching a critical value, in multi-field models known as hybrid inflation [30].

Close to a simple minimum, the scalar field potential can be described by a quadratic function, $V = m^2\phi^2/2$, where m is the mass of the field. We can obtain slow-roll inflation in such a potential at large field values, $\phi \gg M_{\text{P}}$. However, for $\phi \ll M_{\text{P}}$ the field approaches an oscillatory solution:

$$\phi(t) \simeq \frac{M_{\text{P}}}{\sqrt{3\pi}} \frac{\sin(mt)}{mt}. \quad (23.18)$$

For $|\phi| < M_{\text{P}}$ the Hubble rate drops below the inflaton mass, $H < m$, and the field oscillates many times over a Hubble time. Averaging over several oscillations, $\Delta t \gg m^{-1}$, we find $\langle \dot{\phi}^2/2 \rangle_{\Delta t} \simeq \langle m^2\phi^2/2 \rangle_{\Delta t}$ and hence

$$\langle \rho \rangle_{\Delta t} \simeq \frac{M_{\text{P}}^2}{6\pi t^2}, \quad \langle p \rangle_{\Delta t} \simeq 0. \quad (23.19)$$

This coherent oscillating field corresponds to a condensate of non-relativistic massive inflaton particles, driving a matter-dominated era at the end of inflation, with scale factor $R \propto t^{2/3}$.

The inflaton condensate can lose energy through perturbative decays due to terms in the interaction Lagrangian, such as

$$\mathcal{L}_{\text{int}} \subset -\lambda_i \sigma \phi \chi_i^2 - \lambda_j \phi \bar{\psi}_j \psi_j \quad (23.20)$$

that couple the inflation to scalar fields χ_i or fermions ψ_j , where σ has dimensions of mass and the λ_i are dimensionless couplings. When the mass of the inflaton is much larger than the decay products, the decay rate is given by [31]

$$\Gamma_i = \frac{\lambda_i^2 \sigma^2}{8\pi m}, \quad \Gamma_j = \frac{\lambda_j^2 m}{8\pi}. \quad (23.21)$$

These decay products must in turn thermalise with Standard Model particles before we recover conventional hot Big-Bang cosmology. An upper limit on the reheating temperature after inflation is given by [29]

$$T_{\text{rh}} = 0.2 \left(\frac{100}{g_*} \right)^{1/4} \sqrt{M_{\text{P}} \Gamma_{\text{tot}}}, \quad (23.22)$$

where g_* is the effective number of degrees of freedom and Γ_{tot} is the total decay rate for the inflaton, which is required to be less than m for perturbative decay.

The baryon asymmetry of the Universe must be generated after the main release of entropy during inflation, which is an important constraint on possible models. Also, the fact that the inflaton mass is much larger than the mass scale of the Standard Model opens up the possibility that it may decay into massive stable or metastable particles that could be connected with dark matter, constraining possible models. For example, in the context of supergravity models the reheating temperature is constrained by the requirement that gravitinos are not overproduced, potentially destroying the successes of Big-Bang nucleosynthesis. For a range of gravitino masses one must require $T_{\text{rh}} < 10^9$ GeV [32, 33].

The process of inflaton decay and reheating can be significantly altered by interactions leading to space-time dependences in the effective masses of the fields. In particular, parametric resonance can lead to explosive, non-perturbative decay of the inflaton in some cases, a process often referred to as preheating [28, 34]. For example, an interaction term of the form

$$\mathcal{L}_{\text{int}} \subset -\lambda^2 \phi^2 \chi^2, \quad (23.23)$$

leads to a time-dependent effective mass for the χ field as the inflaton ϕ oscillates. This can lead to non-adiabatic particle production if the bare mass of the χ field is small for large couplings or for rapid changes of the inflaton field. The process of preheating is highly model-dependent, but it highlights the possible role of non-thermal particle production after and even during inflation.

23.3 Primordial Perturbations from Inflation

Although inflation was originally discussed as a solution to the problem of initial conditions required for homogeneous and isotropic hot Big-Bang cosmology, it was soon realised that inflation also offered a mechanism to generate the inhomogeneous initial conditions required for the formation of large-scale structure [17–19, 35].

23.3.1 Metric Perturbations

In a homogeneous classical inflationary cosmology driven by a scalar field, the inflaton field is uniform on constant-time hypersurfaces, $\phi = \phi_0(t)$. However, quantum fluctuations inevitably break the spatial symmetry leading to an inhomogeneous field:

$$\phi(t, x^i) = \phi_0(t) + \delta\phi(t, x^i). \quad (23.24)$$

At the same time, one should consider inhomogeneous perturbations of the RW spacetime metric (see, *e.g.*, Refs. [36–38]):

$$ds^2 = (1 + 2A)dt^2 - 2RB_i dt dx^i - R^2 [(1 + 2C)\delta_{ij} + \partial_i \partial_j E + h_{ij}] dx^i dx^j, \quad (23.25)$$

where A , B , C and E are scalar perturbations while h_{ij} represents transverse and tracefree, tensor metric perturbations. Vector metric perturbations can be eliminated using Einstein constraint equations in a scalar field cosmology.

The tensor perturbations remain invariant under a temporal gauge transformation $t \rightarrow t + \delta t(t, x^i)$, but both the scalar field and the scalar metric perturbations transform. For example, we have

$$\delta\phi \rightarrow \delta\phi - \dot{\phi}_0 \delta t, \quad C \rightarrow C - H \delta t. \quad (23.26)$$

However, there are gauge invariant combinations, such as [39]

$$Q = \delta\phi - \frac{\dot{\phi}_0}{H} C, \quad (23.27)$$

which describes the scalar field perturbations on spatially-flat ($C = 0$) hypersurfaces. This is simply related to the curvature perturbation on uniform-field ($\delta\phi = 0$) hypersurfaces:

$$\mathcal{R} = C - \frac{H}{\dot{\phi}_0} \delta\phi = -\frac{H}{\dot{\phi}_0} Q, \quad (23.28)$$

which coincides in slow-roll inflation, $\rho \simeq \rho(\phi)$, with the curvature perturbation on uniform-density hypersurfaces [18]

$$\zeta = C - \frac{H}{\dot{\rho}_0} \delta\rho. \quad (23.29)$$

Thus scalar field and scalar metric perturbations are coupled by the evolution of the inflaton field.

23.3.2 Gravitational waves from inflation

The tensor metric perturbation, h_{ij} in Eq. (23.25), is gauge-invariant and decoupled from the scalar perturbations at first order. This represents the free excitations of the spacetime, *i.e.*, gravitational waves, which are the simplest metric perturbations to study at linear order.

Each tensor mode, with wavevector $\vec{k}$, has two linearly-independent transverse and trace-free polarization states:

$$h_{ij}(\vec{k}) = h_{\vec{k}} q_{ij} + \bar{h}_{\vec{k}} \bar{q}_{ij}. \quad (23.30)$$

The linearised Einstein equations then yield the same evolution equation for the amplitude as that for a massless field in RW spacetime:

$$\ddot{h}_{\vec{k}} + 3H\dot{h}_{\vec{k}} + \frac{k^2}{R^2} h_{\vec{k}} = 0, \quad (23.31)$$

(and similarly for $\bar{h}_{\vec{k}}$). This can be re-written in terms of the conformal time, $\eta = \int dt/R$, and the conformally rescaled field:

$$u_{\vec{k}} = \frac{M_{\text{P}} R h_{\vec{k}}}{\sqrt{32\pi}}. \quad (23.32)$$

This conformal field then obeys the wave equation for a canonical scalar field in Minkowski spacetime with a time-dependent mass:

$$u_{\vec{k}}'' + \left(k^2 - \frac{R''}{R}\right) u_{\vec{k}} = 0. \quad (23.33)$$

During slow roll

$$\frac{R''}{R} \simeq (2 - \epsilon) R^2 H^2. \quad (23.34)$$

This makes it possible to quantise the linearised metric fluctuations, $u_{\vec{k}} \rightarrow \hat{u}_{\vec{k}}$, on sub-Hubble scales, $k^2/R^2 \gg H^2$, where the background expansion can be neglected.

Crucially, in an inflationary expansion, where $\ddot{R} > 0$, the comoving Hubble length $H^{-1}/R = 1/\dot{R}$ decreases with time. Thus, all modes start inside the Hubble horizon and it is possible to take the initial field fluctuations to be in a vacuum state at early times or on small scales:

$$\langle u_{\vec{k}_1} u_{\vec{k}_2} \rangle = \frac{i}{2} (2\pi)^3 \delta^{(3)}(\vec{k}_1 + \vec{k}_2). \quad (23.35)$$

In terms of the amplitude of the tensor metric perturbations, this corresponds to

$$\langle h_{\vec{k}_1} h_{\vec{k}_2} \rangle = \frac{1}{2} \frac{\mathcal{P}_{\text{t}}(k_1)}{4\pi k_1^3} (2\pi)^3 \delta^{(3)}(\vec{k}_1 + \vec{k}_2), \quad (23.36)$$

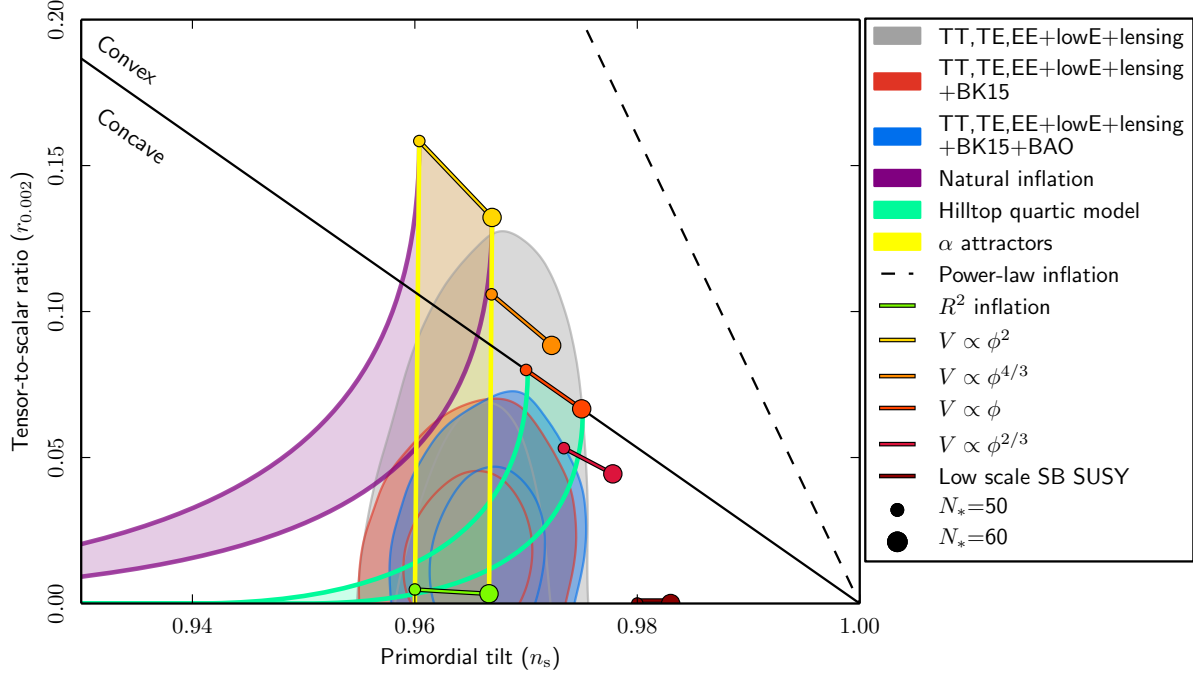


Figure 23.1: The marginalized joint 68 and 95% CL regions for the tilt in the scalar perturbation spectrum, n_s , and the relative magnitude of the tensor perturbations, r , obtained from the *Planck* 2018 and lensing data alone, and their combinations with BICEP2/Keck Array (BK15) and (optionally) BAO data, confronted with the predictions of some of the inflationary models discussed in this review. This figure is taken from Ref. [40].

where the factor $1/2$ appears due to the two polarization states that contribute to the total tensor power spectrum:

$$\mathcal{P}_t(k) = \frac{64\pi}{M_{\text{P}}^2} \left(\frac{k}{2\pi R} \right)^2. \quad (23.37)$$

On super-Hubble scales, $k^2/R^2 \ll H^2$, we have the growing mode solution to Eq. (23.33), $u_{\vec{k}} \propto R$, corresponding to $h_{\vec{k}} \rightarrow \text{constant}$, *i.e.*, tensor modes are frozen-in on super-Hubble scales, both during and after inflation. Thus, connecting the initial vacuum fluctuations on sub-Hubble scales to the late-time power spectrum for tensor modes at Hubble exit during inflation, $k = R_* H_*$, we obtain

$$\mathcal{P}_t(k) \simeq \frac{64\pi}{M_{\text{P}}^2} \left(\frac{H_*}{2\pi} \right)^2. \quad (23.38)$$

In the de Sitter limit, $\epsilon \rightarrow 0$, the Hubble rate becomes time-independent and the tensor spectrum on super-Hubble scales becomes scale-invariant [41]. However slow-roll evolution leads to weak time dependence of H_* and thus a scale-dependent spectrum on large scales, with a spectral tilt

$$n_t \equiv \frac{d \ln \mathcal{P}_t}{d \ln k} \simeq -2\epsilon_*. \quad (23.39)$$

23.3.3 Density Perturbations from single-field inflation

The inflaton field fluctuations on spatially-flat hypersurfaces are coupled to scalar metric perturbations at first order, but these can be eliminated using the Einstein constraint equations to

yield an evolution equation

$$\ddot{Q}_{\vec{k}} + 3H\dot{Q}_{\vec{k}} + \left[\frac{k^2}{R^2} + V'' - \frac{8\pi}{M_{\text{P}}^2 R^3} \frac{d}{dt} \left(\frac{R^3 \dot{\phi}^2}{H} \right) \right] Q_{\vec{k}} = 0. \quad (23.40)$$

Terms proportional to M_{P}^{-2} represent the effect on the field fluctuations of gravity at first order. As can be seen, this vanishes in the limit of a constant background field, and hence is suppressed in the slow-roll limit, but it is of the same order as the effective mass, $V'' = 3\eta H^2$, so must be included if we wish to model deviations from exact de Sitter symmetry.

This wave equation can also be written in the canonical form for a free field in Minkowski spacetime if we define [39]

$$v_{\vec{k}} \equiv RQ_{\vec{k}}, \quad (23.41)$$

to yield

$$v_{\vec{k}}'' + \left(k^2 - \frac{z''}{z} \right) v_{\vec{k}} = 0, \quad (23.42)$$

where we further define

$$z \equiv \frac{R\dot{\phi}}{H}, \quad \frac{z''}{z} \simeq (2 + 5\epsilon - 3\eta)R^2 H^2, \quad (23.43)$$

with the last approximate equality holding to leading order in the slow-roll approximation.

As previously done for gravitational waves, we quantise the linearised field fluctuations $v_{\vec{k}} \rightarrow \hat{v}_{\vec{k}}$ on sub-Hubble scales, $k^2/R^2 \gg H^2$, where the background expansion can be neglected. Thus, we impose

$$\langle v_{\vec{k}_1} v_{\vec{k}_2}' \rangle = \frac{i}{2} \delta^{(3)}(\vec{k}_1 + \vec{k}_2). \quad (23.44)$$

In terms of the field perturbations, this corresponds to

$$\langle Q_{\vec{k}_1} Q_{\vec{k}_2} \rangle = \frac{\mathcal{P}_Q(k_1)}{4\pi k_1^3} (2\pi)^3 \delta^{(3)}(\vec{k}_1 + \vec{k}_2), \quad (23.45)$$

where the power spectrum for vacuum field fluctuations on sub-Hubble scales, $k^2/R^2 \gg H^2$, is simply

$$\mathcal{P}_Q(k) = \left(\frac{k}{2\pi R} \right)^2, \quad (23.46)$$

yielding the classic result for the vacuum fluctuations for a massless field in de Sitter at Hubble exit, $k = R_* H_*$:

$$\mathcal{P}_Q(k) \simeq \left(\frac{H}{2\pi} \right)_*^2. \quad (23.47)$$

In practice there are slow-roll corrections due to the small but finite mass (η) and field evolution (ϵ) [42].

Slow-roll corrections to the field fluctuations are small on sub-Hubble scales, but can become significant as the field and its perturbations evolve over time on super-Hubble scales. This is why it is helpful to work instead with the curvature perturbation, ζ defined in Eq. (23.29), which remains constant on super-Hubble scales for adiabatic density perturbations both during and after inflation [18, 43]. Thus we have an expression for the primordial curvature perturbation on super-Hubble scales produced by single-field slow-roll inflation:

$$\mathcal{P}_\zeta(k) = \left[\left(\frac{H}{\dot{\phi}} \right)^2 \mathcal{P}_Q(k) \right]_* \simeq \frac{4\pi}{M_{\text{P}}^2} \left[\frac{1}{\epsilon} \left(\frac{H}{2\pi} \right)^2 \right]_*. \quad (23.48)$$

Comparing this with the primordial gravitational wave power spectrum (Eq. (23.38)) we obtain the tensor-to-scalar ratio

$$r \equiv \frac{\mathcal{P}_t}{\mathcal{P}_\zeta} \simeq 16\epsilon_*. \quad (23.49)$$

Note that the scalar amplitude is boosted by a factor $1/\epsilon_*$ during slow-roll inflation, because small scalar field fluctuations can lead to relatively large curvature perturbations on hypersurfaces defined with respect to the density if the potential energy is only weakly dependent on the scalar field, as in slow roll. Indeed, the de Sitter limit is singular, since the potential energy becomes independent of the scalar field at first order, $\epsilon \rightarrow 0$, and the curvature perturbation on uniform-density hypersurfaces becomes ill-defined.

We note that in single-field inflation the tensor-to-scalar ratio and the tensor tilt (Eq. (23.39)) at the same scale are both determined by the first slow-roll parameter at Hubble exit, ϵ_* , giving rise to an important consistency test for single-field slow-roll inflation:

$$n_t \simeq -\frac{r}{8}. \quad (23.50)$$

This may be hard to verify if r is small, making any tensor tilt n_t difficult to measure. On the other hand, it does offer a way to rule out single-field slow-roll inflation if either r or n_t is large.

Given the relatively large scalar power spectrum, it has proved easier to measure the scalar tilt, conventionally defined as $n_s - 1$. Slow-roll corrections lead to slow time-dependence of both H_* and ϵ_* , giving a weak scale-dependence of the scalar power spectrum:

$$n_s - 1 \equiv \frac{d \ln \mathcal{P}_\zeta}{d \ln k} \simeq -6\epsilon_* + 2\eta_*, \quad (23.51)$$

and a running of this tilt at second-order in slow roll:

$$\frac{dn_s}{d \ln k} \simeq -8\epsilon_*(3\epsilon_* - 2\eta_*) - 2\xi_*^2, \quad (23.52)$$

where the running introduces a new slow-roll parameter at second order:

$$\xi^2 = \frac{M_{\text{P}}^4}{64\pi^2} \frac{V'V'''}{V^2}. \quad (23.53)$$

23.3.4 Observational Bounds

The observed scale-dependence of the power spectrum makes it necessary to specify the comoving scale, k , at which quantities are constrained and hence the Hubble-exit time, $k = a_* H_*$, when the corresponding theoretical quantities are calculated during inflation. This is usually expressed in terms of the number of e-folds from the end of inflation [44]:

$$\begin{aligned} N_*(k) \simeq & 67 - \ln \left(\frac{k}{a_0 H_0} \right) + \frac{1}{4} \ln \left(\frac{V_*^2}{M_{\text{P}}^4 \rho_{\text{end}}} \right) + \frac{1}{12} \ln \left(\frac{\rho_{\text{rh}}}{\rho_{\text{end}}} \right) \\ & - \frac{1}{12} \ln(g_*), \end{aligned} \quad (23.54)$$

where H_0^{-1}/a_0 is the present comoving Hubble length. Different models of reheating and thus different reheat temperatures and densities, ρ_{rh} in Eq. (23.54), lead to a range of possible values for N_* corresponding to a fixed physical scale, and hence we have a range of observational predictions for a given inflation model; this is illustrated by allowing a fixed range of values for N_* in Fig. 23.1.

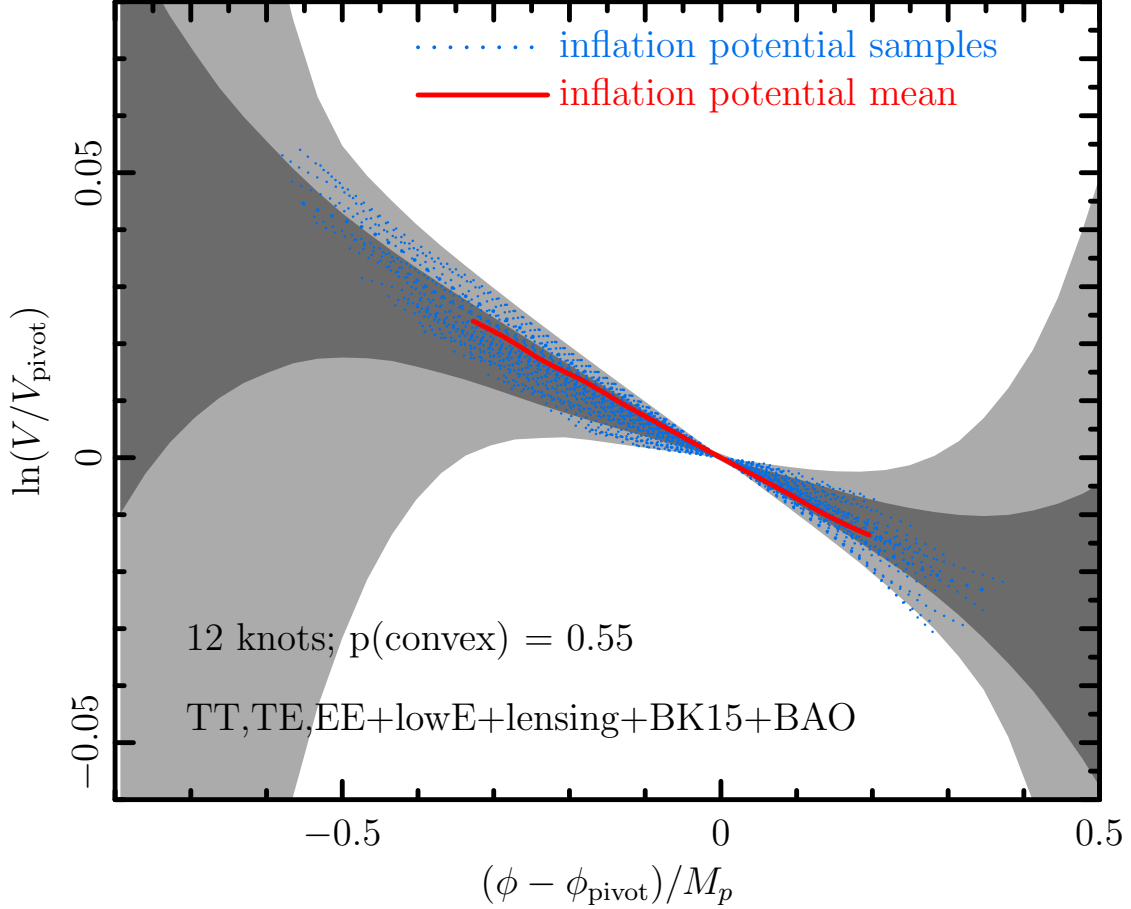


Figure 23.2: The result of reconstructing a single-field inflaton potential using a cubic-spline power-spectrum mode expansion and the full *Planck*, lensing, BK15 and BAO data set. This figure is taken from Refs. [40].

The *Planck* 2018 temperature and polarization data (see Chap. 29, “Cosmic Microwave Background” review) are consistent with a smooth featureless power spectrum over a range of comoving wavenumbers, $0.008 h^{-1} \text{ Mpc}^{-1} \leq k \leq 0.1 h \text{ Mpc}^{-1}$. In the absence of running, the *Planck* data measure the spectral index to be [40]

$$n_s = 0.9649 \pm 0.0042, \quad (23.55)$$

corresponding to a deviation from scale-invariance exceeding the 7σ level. Recent results from the Atacama Cosmology Telescope (ACT) [45] on small angular scales in combination with *Planck* data on larger scales, lensing measurements and DESI DR2 data [46] suggest a somewhat higher value, $n_s = 0.975 \pm 0.003$, although this is not yet a full joint analysis. There are also recent results from the South Pole Telescope [47] that are quite consistent with *Planck*. If running of the spectral tilt is included in the model, this is constrained to be [40]

$$\frac{dn_s}{d \ln k} = -0.0045 \pm 0.0067 \quad (23.56)$$

at the 95% CL, assuming no running of the running. The most recent analysis [48] of the BICEP, Keck Array and *Planck* data places an upper bound on the tensor-to-scalar ratio

$$r < 0.036, \quad (23.57)$$

at the 95% CL.

These observational bounds can be converted into bounds on the slow-roll parameters and hence the potential during slow-roll inflation. Setting higher-order slow-roll parameters (beyond second-order in horizon-flow parameters [49]) to zero, the *Planck* collaboration obtain the following 95% CL bounds when lensing and BK15 data are included [40]

$$\epsilon < 0.0044, \quad (23.58)$$

$$\eta = -0.015 \pm 0.006, \quad (23.59)$$

$$\xi^2 = 0.0029^{+0.0073}_{-0.0069}, \quad (23.60)$$

which can be used to constrain models, as discussed in the next section.

Figure 23.1, which is taken from Ref. [40], compares observational CMB constraints on the tilt, n_s , in the spectrum of scalar perturbations and the ratio, r , between the magnitudes of tensor and scalar perturbations. Important rôles are played by data from the *Planck* satellite and on lensing, the BICEP2/Keck Array (BK15) and measurements of baryon acoustic oscillations (BAO). The reader is referred to Ref. [40] for technical details. These experimental constraints are compared with the predictions of some of the inflationary models discussed in this review. Generally speaking, models with a concave potential are favored over those with a convex potential, and models with power-law inflation are now excluded, as opposed to models with de Sitter-like (quasi-)exponential expansion.

There is no significant evidence for local features within the range of inflaton field values probed by the data [40]. However, the data may be used to partially reconstruct the effective inflationary potential over a range of inflaton field values, assuming that it is suitably smooth. The result of one such exercise by the *Planck* collaboration [40] in the framework of a generic single-field inflaton potential is shown in Fig. 23.2. This reconstruction assumes a cubic-spline power-spectrum mode expansion and employs the full *Planck*, lensing, BK15 and BAO data set. The reader is again referred to [40] for technical details. We see that the effective inflaton potential is relatively well reconstructed over field values ϕ within ± 0.5 of the chosen pivot value, but the potential is only very weakly constrained for larger values of $|\phi - \phi_{\text{pivot}}|$, providing wide scope for inflationary model-builders.

23.4 Models

23.4.1 Pioneering Models

The paradigm of the inflationary Universe was proposed in Ref. [6], where it was pointed out that an early period of (near-)exponential expansion, in addition to resolving the horizon and flatness problems of conventional Big-Bang cosmology as discussed above, would also dilute the prior abundance of any unseen heavy, (meta-)stable particles, as exemplified by monopoles in grand unified theories (GUTs; see Chap. 92, “Grand Unified Theories” review). The possibility of a de Sitter phase in the early history of the Universe was also proposed in the non-minimal gravity model of Ref. [5], with the motivation of avoiding an initial singularity. The original proposal in Ref. [6] was that this inflationary expansion took place while the Universe was in a metastable state (a similar suggestion was made in Refs. [50, 51], where in Ref. [50] it was also pointed out that such a mechanism could address the horizon problem) and was terminated by a first-order transition due to tunneling through a potential barrier. However, it was recognized already in Ref. [6] that this ‘old inflation’ scenario would need modification if the transition to the post-inflationary Universe were to be completed smoothly without generating unacceptable inhomogeneities.

This ‘graceful exit’ problem was addressed in the ‘new inflation’ model of Ref. [15] (see also Ref. [16] and footnote [39] of Ref. [6]), which studied models based on an $SU(5)$ GUT with an

effective potential of the Coleman-Weinberg type (*i.e.*, dominated by radiative corrections), in which inflation could occur during the roll-down from the local maximum of the potential towards a global minimum. However, it was realized that the Universe would evolve to a different minimum from the Standard Model [52], and it was also recognized that density fluctuations would necessarily be too large [17], since they are related to the GUT coupling strength.

These early models of inflation assumed initial conditions enforced by thermal equilibrium in the early Universe. However, this assumption was questionable: indeed, it was not made in the model of Ref. [5], in which a higher-order gravitational curvature term was assumed to arise from quantum corrections, and the assumption of initial thermal equilibrium was jettisoned in the ‘chaotic’ inflationary model of Ref. [53]. These have been the inspirations for much recent inflationary model building, so we now discuss them in more detail, before reviewing contemporary models.

In this section we will work in natural units where we set the reduced Planck mass to unity, *i.e.*, $8\pi/M_{\text{P}}^2 = 1$. All masses are thus relative to the reduced Planck scale.

23.4.2 R^2 Inflation

The first-order Einstein-Hilbert action, $(1/2) \int d^4x \sqrt{-g} R$, where R is the Ricci scalar curvature, is the minimal possible theory consistent with general coordinate invariance. However, it is possible that there might be non-minimal corrections to this action, and the unique second-order possibility is

$$S = \frac{1}{2} \int d^4x \sqrt{-g} \left(R + \frac{R^2}{6M^2} \right). \quad (23.61)$$

It was pointed out in Ref. [5] that an R^2 term could be generated by quantum effects, and that Eq. (23.61) could lead to de Sitter-like expansion of the Universe. Scalar density perturbations in this model were calculated in Ref. [19]. Because the initial phase was (almost) de Sitter, these perturbations were (approximately) scale-invariant, with magnitude $\propto M$. It was also pointed out in Ref. [19] that requiring the scalar density perturbations to lie in the range 10^{-3} to 10^{-5} , consistent with upper limits at that time, would require $M \sim 10^{-3}$ to 10^{-5} in Planck units, and it was further suggested that these perturbations could lead to the observed large-scale structure of the Universe, including the formation of galaxies.

Although the action (23.61) does not contain an explicit scalar field, Ref. [19] reduced the calculation of density perturbations to that of fluctuations in the scalar curvature R , which could be identified (up to a factor) with a scalar field of mass M . The formal equivalence of R^2 gravity (23.61) to a theory of gravity with a massive scalar ϕ had been shown in Ref. [20] (see also Ref. [21]). The effective scalar potential for what we would nowadays call the ‘inflaton’ [54] takes the form

$$S = \frac{1}{2} \int d^4x \sqrt{-g} \left[R + (\partial_\mu \phi)^2 - \frac{3}{2} M^2 (1 - e^{-\sqrt{2/3} \phi})^2 \right] \quad (23.62)$$

when the action is written in the Einstein frame, and the potential is shown as the solid black line in Fig. 23.3. Using Eq. (23.48), one finds that the amplitude of the scalar density perturbations in this model is given by

$$\Delta \mathcal{R} = \frac{3M^2}{8\pi^2} \sinh^4 \left(\frac{\phi}{\sqrt{6}} \right), \quad (23.63)$$

The measured magnitude of the density fluctuations in the CMB requires $M \simeq 1.3 \times 10^{-5}$ in Planck units (assuming $N_* \simeq 55$), so one of the open questions in this model is why M is so small. Obtaining $N_* \simeq 55$ also requires an initial value of $\phi \simeq 5.5$, *i.e.*, a super-Planckian initial condition, and another issue for this and many other models is how the form of the effective potential is

protected and remains valid at such large field values. Using Eq. (23.51) one finds that $n_s \simeq 0.965$ for $N_* \simeq 55$ and using (Eq. (23.49)) one finds that $r \simeq 0.0035$. These predictions are consistent with the present data from *Planck* and other experiments, as seen in Fig. 23.1.

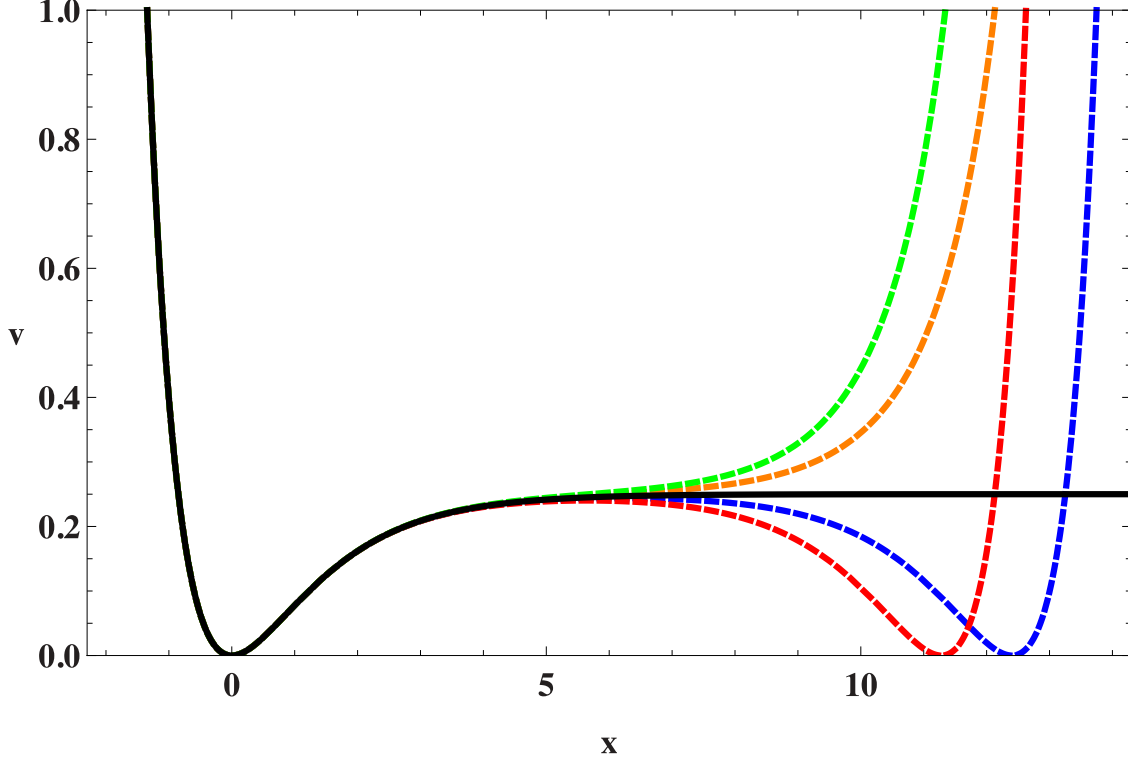


Figure 23.3: Inflationary potential V in the R^2 model (solid black line) compared with its form in various no-scale models discussed in detail in Ref. [55] (dashed coloured lines).

23.4.3 Chaotic Models with Power-Law Potentials

As has already been mentioned, a key innovation in inflationary model-building was the suggestion to abandon the questionable assumption of a thermal initial state, and consider ‘chaotic’ initial conditions with very general forms of potential [53]. (Indeed, the R^2 model discussed above can be regarded as a prototype of this approach.) The chaotic approach was first proposed in the context of a simple power-law potential of the form $\mu^{4-\alpha}\phi^\alpha$, and the specific example of $\lambda\phi^4$ was studied in Ref. [53]. Such models make the following predictions for the slow-roll parameters ϵ and η :

$$\epsilon = \frac{1}{2} \left(\frac{\alpha}{\phi} \right)^2, \quad \eta = \frac{\alpha(\alpha-1)}{\phi^2}, \quad (23.64)$$

leading to the predictions

$$r \approx \frac{4\alpha}{N_*}, \quad n_s - 1 \approx -\frac{\alpha+2}{2N_*}, \quad (23.65)$$

which are shown in Fig. 23.1 for some illustrative values of α . We note that the prediction of the original ϕ^4 model lies out of the frame, with values of r that are too large and values of n_s that are too small. The ϕ^3 model has similar problems, and would in any case require modification in order to have a well-defined minimum. The simplest possibility is ϕ^2 , but this is now also disfavored by the data, at the 95% CL if only the *Planck* data are considered, and more strongly if other data are included, as seen in Fig. 23.1. (For non-minimal models of quadratic inflation that avoid this

problem, see, *e.g.*, Ref. [56].) Indeed, as can be seen in Fig. 23.1, all single-field models with a convex potential (*i.e.*, one curving upwards) are disfavored compared to models with a concave potential.

23.4.4 Hilltop Models

This preference for a concave potential motivates interest in ‘hilltop’ models [57], whose starting-point is a potential of the form

$$V(\phi) = M^4 \left[1 - \left(\frac{\phi}{\mu} \right)^p + \dots \right], \quad (23.66)$$

where the $\dots$ represent extra terms that yield a positive semi-definite potential. To first order in the slow-roll parameters, when $x \equiv \phi/\mu$ is small, one has

$$n_s \simeq 1 - p(p-1)\mu^{-2} \frac{x^{p-2}}{(1-x^p)} - \frac{3}{8}r, \quad r \simeq 8p^2\mu^{-2} \frac{x^{2p-2}}{(1-x^p)^2}. \quad (23.67)$$

As seen in Fig. 23.1, a hilltop model with $p = 4$ can be compatible with the *Planck* and other measurements, if $\mu \gg M_P$.

23.4.5 D-Brane Inflation

Many scenarios for inflation involving extra dimensions have been proposed, *e.g.*, the possibility that observable physics resides on a three-dimensional brane, and that there is an inflationary potential that depends on the distance between our brane and an antibrane, with a potential of the form [58]

$$V(\phi) = M^4 \left[1 - \left(\frac{\mu}{\phi} \right)^p + \dots \right]. \quad (23.68)$$

In this scenario the effective potential vanishes in the limit $\phi \rightarrow \infty$, corresponding to complete separation between our brane and the antibrane. The predictions for n_s and r in this model can be obtained from Eq. (23.67) by exchanging $p \leftrightarrow -p$, and are also consistent with the *Planck* and other data.

23.4.6 Natural Inflation

Also seen in Fig. 23.1 are the predictions of ‘natural inflation’ [59], in which one postulates a non-perturbative shift symmetry that suppresses quantum corrections, so that a hierarchically small scale of inflation, $H \ll M_P$, is technically natural. In the simplest models, there is a periodic potential of the form

$$V(\phi) = M^4 \left[1 + \cos \left(\frac{\phi}{f} \right) \right], \quad (23.69)$$

where f is a dimensional parameter reminiscent of an axion decay constant (see the next subsection) [60], which must have a value $> M_P$. Natural inflation can yield predictions similar to quadratic inflation (which are no longer favored, as already discussed), but can also yield an effective convex potential. Thus, it may lead to values of r that are acceptably small, but for values of n_s that are in tension with the data, as seen in Fig. 23.1.

23.4.7 Axion Monodromy Models

The effective potentials in stringy models [61, 62] motivated by axion monodromy may be of the form

$$V(\phi) = \mu^{4-\alpha}\phi^\alpha + M^4 e^{-C\left(\frac{\phi}{\phi_0}\right)^{p_M}} \cos \left[\gamma + \frac{\phi}{f} \left(\frac{\phi}{\phi_0} \right)^{p_f+1} \right], \quad (23.70)$$

where μ, M, f and ϕ_0 are parameters with the dimension of mass, and C, p, p_M, p_f and γ are dimensionless constants, generalizing the potential [59] in the simplest models of natural inflation.

The oscillations in Eq. (23.70) are associated with the axion field, and powers $p_M, p_f \neq 0$ may arise from ϕ -dependent evolutions of string moduli. Since the exponential prefactor in Eq. (23.70) is due to non-perturbative effects that may be strongly suppressed, the oscillations may be unobservably small. Specific string models having ϕ^α with $\alpha = 4/3, 1$ or $2/3$ have been constructed in Refs. [61, 62], providing some motivation for the low-power models mentioned above.

As seen in Fig. 23.1, the simple axion monodromy models with the power $\alpha = 4/3$ or 1 are no longer compatible with the current CMB data at the 95% CL, while $\alpha = 2/3$ is only marginally compatible at 95% CL. The *Planck* Collaboration has also searched for characteristic effects associated with the second term in Eq. (23.70), such as a possible drift in the modulation amplitude (setting $p_M = C = 0$), and a possible drifting frequency generated by $p_f \neq 0$, without finding any compelling evidence [40].

23.4.8 Higgs Inflation

Since the energy scale during inflation is commonly expected to lie between the Planck and TeV scales, it may serve as a useful bridge with contacts both to string theory or some other quantum theory of gravity, on the one side, and particle physics on the other side. However, as the above discussion shows, much of the activity in building models of inflation has been independent of specific connections with these subjects, though some examples of string-motivated models of inflation were mentioned above.

The most economical scenario for inflation might be to use as the inflaton the only established scalar field, namely the Higgs field (see Chap.11, “Status of Higgs boson physics” review). A specific model assuming a non-minimal coupling of the Higgs field h to gravity was constructed in Ref. [63]. Its starting point is the action

$$S = \int d^4x \sqrt{-g} \left[\frac{M^2 + \xi h^2}{2} R + \frac{1}{2} \partial_\mu h \partial^\mu h - \frac{\lambda}{4} (h^2 - v^2)^2 \right], \quad (23.71)$$

where v is the Higgs vacuum expectation value. The model requires $\xi \gg 1$, in which case it can be rewritten in the Einstein frame as

$$S = \int d^4x \sqrt{-g} \left[\frac{1}{2} R + \frac{1}{2} \partial_\mu \chi \partial^\mu \chi - U(\chi) \right], \quad (23.72)$$

where the effective potential for the canonically-normalized inflaton field χ has the form

$$U(\chi) = \frac{\lambda}{4\xi^2} \left[1 + \exp \left(-\frac{2\chi}{\sqrt{6}M_P} \right) \right]^{-2}, \quad (23.73)$$

which is similar to the effective potential of the R^2 model at large field values. As such, the model inflates successfully if $\xi \simeq 5 \times 10^4 m_h / (\sqrt{2}v)$, where m_h is the Higgs mass, with predictions for n_s and r that are indistinguishable from the predictions of the R^2 model shown in Fig. 23.1.²

This model is appealing because of its connection to particle physics, but must confront several issues. One is to understand the value of ξ , and another is the possibility of unitarity violation. However, a more fundamental issue is whether the effective quartic Higgs coupling is positive at the scale of the Higgs field during inflation. Extrapolations of the effective potential in the Standard Model using the measured values of the masses of the Higgs boson and the top quark indicate that probably $\lambda < 0$ at this scale [64], though there are still significant uncertainties associated with the appropriate input value of the top mass and the extrapolation to high renormalization scales.

²Although Higgs and R^2 inflation share the same inflaton potential, they may reheat the Universe through different mechanisms, resulting in different reheating temperatures and hence different values of N_* that would lead to slightly different predictions; see the discussion around Eq. (23.54).

23.4.9 Supersymmetric Models of Inflation

Supersymmetry [65] is widely considered to be a well-motivated possible extension of the Standard Model that might become apparent at the TeV scale. It is therefore natural to consider supersymmetric models of inflation. These were originally proposed because of the problems of the new inflationary theory [15, 16] based on the one-loop (Coleman-Weinberg) potential for breaking SU(5). Several of these problems are related to the magnitude of the effective potential parameters: in any model of inflation based on an elementary scalar field, some parameter in the effective potential must be small in natural units, *e.g.*, the quartic coupling λ in a chaotic model with a quartic potential, or the mass parameter μ in a model of chaotic quadratic inflation. These parameters are renormalized multiplicatively in a supersymmetric theory, so that the quantum corrections to small values would be under control. Hence, it was suggested that inflation cries out for supersymmetry [66], though non-supersymmetric resolutions of the problems of Coleman-Weinberg inflation are also possible: see, *e.g.*, Ref. [67].

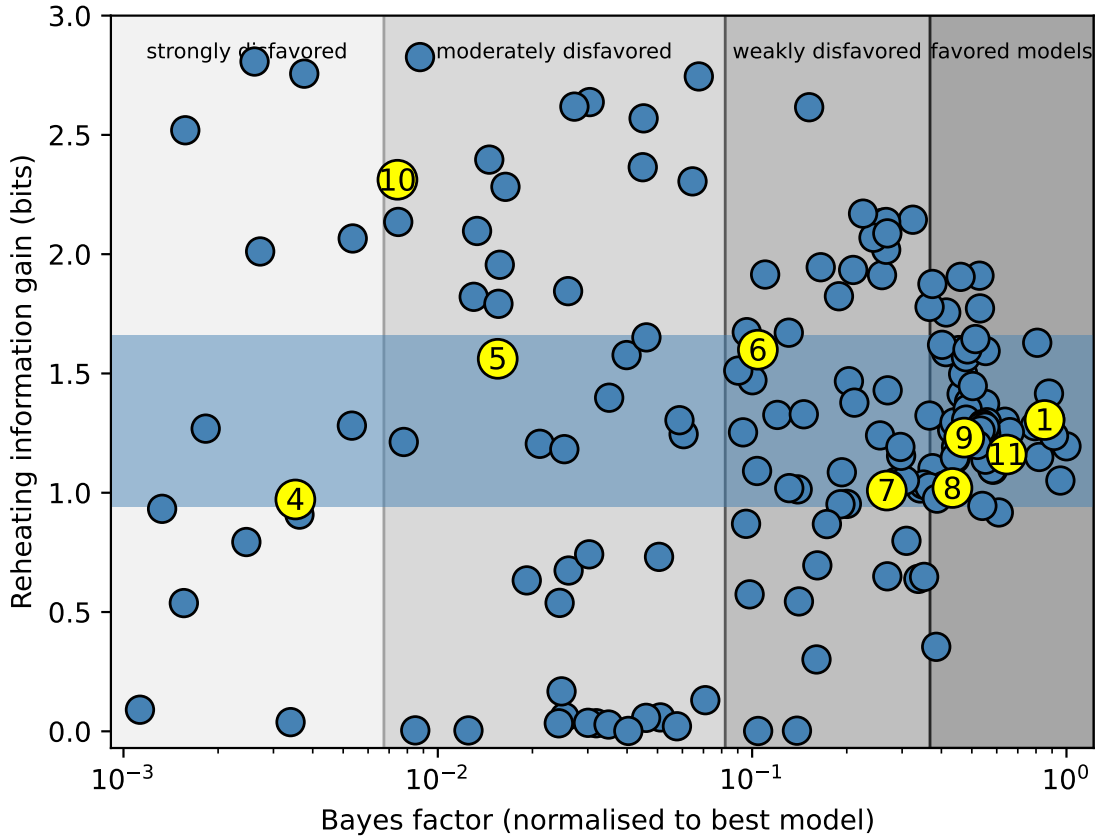


Figure 23.4: Bayes factors calculated in Refs. [68, 69] for 287 single-field inflationary models using *Planck* 2018 data [70] in combination with other CMB and BAO measurements. Those highlighted in yellow are featured in this review, according to the numbers listed in the text (the models labeled "2" and "3" lie outside the plot frame because of their low Bayesian evidence). On the vertical axis, the information gain on the reheating parameter, *i.e.* the k -independent terms in the right-hand side of Eq. (23.54), is displayed in units of bits for the Kullback-Leibler divergence between its prior and posterior distributions [68].

In the Standard Model there is only one scalar field that could be a candidate for the inflaton,

namely the Higgs field discussed above, but even the minimal supersymmetric extension of the Standard Model (MSSM) contains many scalar fields. However, none of these is a promising candidate for the inflaton. The minimal extension of the MSSM that may contain a suitable candidate is the supersymmetric version of the minimal seesaw model of neutrino masses, which contains the three supersymmetric partners of the heavy singlet (right-handed) neutrinos. One of these singlet sneutrinos $\tilde{\nu}$ could be the inflaton [71]: it would have a quadratic potential, the mass coefficient required would be $\sim 10^{13}$ GeV, very much in the expected ballpark for singlet (right-handed) neutrino masses, and sneutrino inflaton decays could also give rise to the cosmological baryon asymmetry via leptogenesis. However, as seen in Fig. 23.1 and already discussed, a purely quadratic inflationary potential is no longer favored by the data. This difficulty could in principle be resolved in models with multiple sneutrinos [72], or by postulating a trilinear sneutrino coupling and hence a superpotential of Wess-Zumino type [73], which can yield successful inflation with predictions intermediate between those of natural inflation and hilltop inflation in Fig. 23.1. Finally, we note that it is also possible to obtain inflation via supersymmetry breaking, as in the model of Ref. [74], whose predictions are illustrated in Fig. 23.1.

23.4.10 Supergravity Models

Any model of early-Universe cosmology, and specifically inflation, must necessarily incorporate gravity. In the context of supersymmetry this requires an embedding in some supergravity theory [75, 76]. An $\mathcal{N} = 1$ supergravity theory is specified by three functions: a Hermitian function of the matter scalar fields ϕ^i , called the Kähler potential K , that describes its geometry, a holomorphic function of the superfields, called the superpotential W , which describes their interactions, and another holomorphic function $f_{\alpha\beta}$, which describes their couplings to gauge fields V_α [77].

The simplest possibility is that the Kähler metric is flat:

$$K = \phi^i \phi_i^*, \quad (23.74)$$

where the sum is over all scalar fields in the theory, and the simplest inflationary model in minimal supergravity has the superpotential [78]

$$W = m^2(1 - \phi)^2, \quad (23.75)$$

where ϕ is the inflaton. However, this model predicts a tilted scalar perturbation spectrum, $n_s = 0.933$, which is now in serious disagreement with the data from *Planck* and other experiments shown in Fig. 23.1.

Moreover, there is a general problem that arises in any supergravity theory coupled to matter, namely that, since its effective scalar potential contains a factor of e^K , scalars typically receive squared masses $\propto H^2 \sim V$, where H is the Hubble parameter [79], an issue called the ‘ η problem’. The theory given by Eq. (23.75) avoids this η problem, but a generic supergravity inflationary model encounters this problem of a large inflaton mass. Moreover, there are additional challenges for supergravity inflation associated with the spontaneous breaking of local supersymmetry [80–82].

Various approaches to the η problem in supergravity have been proposed, including the possibility of a shift symmetry [83], and one possibility that has attracted renewed attention recently is no-scale supergravity [84, 85]. This is a form of supergravity with a Kähler potential that can be written in the form [86]

$$K = -3 \ln \left(T + T^* - \frac{\sum_i |\phi^i|^2}{3} \right), \quad (23.76)$$

which has the special property that it naturally has a flat potential, at the classical level and before specifying a non-trivial superpotential. As such, no-scale supergravity is well-suited for

constructing models of inflation. Adding to its attraction is the feature that compactifications of string theory to supersymmetric four-dimensional models yield effective supergravity theories of the no-scale type [87]. There are many examples of superpotentials that yield effective inflationary potentials for either the T field (which is akin to a modulus field in some string compactification) or a ϕ field (generically representing matter) that are of the same form as the effective potential of the R^2 model Eq. (23.62) when the magnitude of the inflaton field is far above the Planck scale, as required to obtain sufficiently many e-folds of inflation, N_* [88, 89]. This framework also offers the possibility of using a suitable superpotential to construct models with effective potentials that are similar, but not identical, to the R^2 model, as shown by the dashed coloured lines in Fig. 23.3.

23.4.11 Other Exponential Potential Models

This framework also offers the possibility [88] of constructing models in which the asymptotic constant value of the potential at large inflaton field values is approached via a different exponentially-suppressed term:

$$V(\phi) = A \left[1 - \delta e^{-B\phi} + \mathcal{O}(e^{-2B\phi}) \right], \quad (23.77)$$

where the magnitude of the scalar density perturbations fixes A , but δ and B are regarded as free parameters. In the case of R^2 inflation, in Eq. (23.62), $\delta = 2$ and $B = \sqrt{2/3}$. In a general model such as Eq. (23.77), one finds at leading order in the small quantity $e^{-B\phi}$ that

$$\begin{aligned} n_s &= 1 - 2B^2\delta e^{-B\phi}, \\ r &= 8B^2\delta^2 e^{-2B\phi}, \\ N_* &= \frac{1}{B^2\delta} e^{+B\phi}. \end{aligned} \quad (23.78)$$

yielding the relations

$$n_s = 1 - \frac{2}{N_*}, \quad r = \frac{8}{B^2 N_*^2}. \quad (23.79)$$

These relations correspond to the class of models labeled as ‘ α attractors’ [90] in Fig. 23.1. The spectral index in such models is determined by N_* , the number of e-folds Eq. (23.16) when modes exited the Hubble scale before the end of inflation. For example, *Planck* constraints Eq. (23.55) are consistent with $51 < N_* < 65$, placing bounds on the reheating temperature through Eq. (23.54). If the shift of n_s towards higher values suggested by the combination with recent ACT data [45] were to be confirmed, requiring values of $N_* > 60$ at 2σ , the simplest realisations of α -attractor models with conventional reheating would be ruled out [91].

Meanwhile CMB bounds on the tensor-to-scalar ratio translate into an upper bound on the parameter $\alpha = 2/3B^2$, which determines the curvature of the Kähler manifold in supergravity models [92]. There are generalizations of the simplest no-scale model (23.76) with prefactors before the logarithmic factor that are 1 or 2, leading to larger values of $B = \sqrt{2}$ or 1, respectively, and hence smaller values of r than in the R^2 model. CMB constraints on the scalar spectral index can also place a lower bound on the magnitude of α [93], due to the implicit dependence of N_* in Eq. (23.79) upon α [94].

23.5 Model Comparison

Given a particular inflationary model, one can obtain constraints on the model parameters, informed by the likelihood, corresponding to the probability of the data given a particular choice of parameters (see Sec. 40, “Statistics” review). In the light of the detailed constraints on the statistical distribution of primordial perturbations now inferred from high-precision observations of the CMB, it is also possible to make quantitative comparison of the statistical evidence for or

against different inflationary models. This can be done either by comparing the logarithm of the maximum likelihood that can be obtained for the data using each model, *i.e.*, the minimum χ^2 (with some correction for the number of free parameters in each model), or by a Bayesian model comparison [95] (see also Sec. 40.3.4 in “Statistics” review).

In such a Bayesian model comparison one computes [4, 8, 9] the evidence, $\mathcal{E}(\mathcal{D}|\mathcal{M}_A)$ for a model, $\mathcal{M}_A$, given the data $\mathcal{D}$. This corresponds to the likelihood, $\mathcal{L}(\theta_{Aj}) = p(\mathcal{D}|\theta_{Aj}, \mathcal{M}_A)$, integrated over the assumed prior distribution, $\pi(\theta_{Aj}|\mathcal{M}_A)$, for all the model parameters θ_{Aj} :

$$\mathcal{E}(\mathcal{D}|\mathcal{M}_A) = \int \mathcal{L}(\theta_{Aj})\pi(\theta_{Aj}|\mathcal{M}_A)d\theta_{Aj}. \quad (23.80)$$

The posterior probability of the model given the data follows from Bayes’ theorem

$$p(\mathcal{M}_A|\mathcal{D}) = \frac{\mathcal{E}(\mathcal{D}|\mathcal{M}_A)\pi(\mathcal{M}_A)}{p(\mathcal{D})}, \quad (23.81)$$

where the prior probability of the model is given by $\pi(\mathcal{M}_A)$. Assuming that all models are equally likely a priori, $\pi(\mathcal{M}_A) = \pi(\mathcal{M}_B)$, the relative probability of model A relative to a reference model, in the light of the data, is thus given by the Bayes factor

$$B_{A,\text{ref}} = \frac{\mathcal{E}(\mathcal{D}|\mathcal{M}_A)}{\mathcal{E}(\mathcal{D}|\mathcal{M}_{\text{ref}})}. \quad (23.82)$$

Computation of the multi-dimensional integral (Eq. (23.80)) is a challenging numerical task. Even using an efficient sampling algorithm requires hundreds of thousands of likelihood computations for each model, though slow-roll approximations can be used to rapidly calculate the primordial power spectrum using the APSIC numerical library [96] for a large number of single-field, slow-roll inflation models.

The change in χ^2 for selected slow-roll models relative to the Starobinsky R^2 inflationary model, used as a reference, is given in Table 23.1 (values taken from [69]). Table 23.1 also shows the Bayesian evidence as displayed in Fig. 23.4 for all single-field models (see Refs [4, 8, 69] for more

Table 23.1: Observational evidence for and against selected inflation models: $\Delta\chi^2$ and the Bayes factors are calculated relative to the Starobinsky R^2 inflationary model, which is treated as a reference. Results taken from Ref. [69].

Label in Fig. 23.4	Model	$\Delta\chi^2$	$\ln B_{A,\text{ref}}$
1	R^2 inflation	0	0
2	Power-law potential ϕ^2	+20.76	−7.36
3	Power-law potential ϕ^4	+20.81	−7.34
4	Power-law potential $\phi^{2/3}$	+10.82	−5.49
5	Power-law potential ϕ^p	−0.579	−4.01
6	Hilltop potential $1 - \phi^2$	−1.30	−2.10
7	Hilltop potential $1 - \phi^4$	−1.27	−1.16
8	Hilltop potential $1 - \phi^p$	−1.30	−0.67
9	Brane inflation	−1.25	−0.59
10	Natural inflation	+0.55	−4.74
11	α attractor	−1.26	−0.28

complete descriptions of the models and their priors). The Jeffrey’s scale distinguishes favored,

weakly disfavored, moderately disfavored and strongly disfavored models, based on their Bayesian evidence. Disfavored models are those that either produce too large tensor modes, or scalar fluctuations with a spectral tilt that differs from the value preferred by the data, or that require an extreme fine tuning of their parameters. We note that although α -attractor models can provide a good fit to the data, they are slightly disfavored relative to the Starobinsky model due to their larger prior volume. There is now strong evidence ($|\ln B_{A,\text{ref}}| > 5$) against large-field models such as chaotic inflation with a quadratic or a quartic potential. Indeed, over 40% of the slow-roll inflation models considered in Ref. [4, 8, 9, 69] are strongly disfavored by the *Planck* data.

Models discussed in this review are highlighted in yellow in Fig. 23.4, and numbered as follows: (1) R^2 inflation (Sec. 23.4.2) and models with similar predictions, such as Higgs inflation (Sec. 23.4.8) and no-scale supergravity inflation (Sec. 23.4.10); chaotic inflation models (2) with a ϕ^2 potential; (3) with a ϕ^4 potential; (4) with a $\phi^{2/3}$ potential, and (5) with a ϕ^p potential marginalising over $p \in [0.2, 6]$ (Sec. 23.4.3); hilltop inflation models (6) with $p = 2$; (7) with $p = 4$ and (8) marginalising over p (Sec. 23.4.4); (9) brane inflation Sec. 23.4.5); (10) natural inflation (Sec. 23.4.6); and (11) exponential potential models such as α -attractors (Sec. 23.4.11).

In Fig. 23.4, the information gain on reheating is also displayed. It corresponds to the Kullback-Leibler divergence (in units of bits) between the prior and posterior distributions of the so-called reheating parameter, formed by the right-hand side of Eq. (23.54). With recent CMB measurements the best inflationary models have crossed the one-bit threshold [68, 97], which opens up the possibility to constrain the physical processes at play during reheating. This is what is discussed in the next Section.

23.6 Constraints on Reheating

One connection between inflation and particle physics is provided by inflaton decay, whose products are expected to have thermalized subsequently. As seen in Eq. (23.54), the number of e-folds required during inflation depends on details of this reheating process, including the matter density upon reheating, denoted by ρ_{th} , which depends in turn on the inflaton decay rate Γ_ϕ . We see in Fig. 23.1 that, within any specific inflationary model, both n_s and r are sensitive to the value of N_* . In particular, the uncertainty in the experimental measurement of n_s is comparable to the variation in many model predictions for $N_* \in [50, 60]$. This implies that the data start to constrain scenarios for inflaton decay in many models. For example, it is clear from Fig. 23.1 that $N_* = 60$ would be preferred over $N_* = 50$ in a chaotic inflationary model with a quadratic potential.

As a specific example, let us consider α -attractor models that predict small values of r . As seen in Fig. 23.1, within these models the combination of *Planck*, BICEP2/Keck Array and BAO data would require a limited range of n_s , corresponding to a limited range of N_* , as seen by comparing the left and right vertical axes in Fig. 23.5:

$$N_* \gtrsim 52 \quad (68\% \text{ CL}), \quad N_* \gtrsim 44 \quad (95\% \text{ CL}). \quad (23.83)$$

Within any specific model for inflaton decay, these bounds can be translated into constraints on the effective decay coupling. For example, if one postulates a two-body inflaton decay (Eq. (23.21)), the bounds (Eq. (23.83)) can be translated into bounds on $y^2 \equiv 8\pi\Gamma/m$. This is illustrated in Fig. 23.5 [92], where any value of $N_{0.05}$ (the number of e-folds for a scale 0.05 Mpc^{-1} , on the left vertical axis), corresponding to the scalar tilt n_s shown on the right vertical axis, when projected onto the diagonal line representing one of the three illustrated choices of α -attractor models, corresponds to a specific value of y (lower horizontal axis) and T_{reh} (upper horizontal axis). For example, for $\alpha = 1$ one has

$$y \gtrsim 10^{-10} \quad (68\% \text{ CL}), \quad y \gtrsim 10^{-17} \quad (95\% \text{ CL}). \quad (23.84)$$

These bounds are not very constraining, but they can be expected to improve significantly in the coming years and thereby provide significant information on the connections between inflation and particle physics.

23.7 Beyond Single-Field Slow-Roll Inflation

There are numerous possible scenarios beyond the simplest single-field models of slow-roll inflation. These include models with features in the potential leading to deviations from slow roll, as well as theories in which non-canonical fields are considered, such as k-inflation [98] or DBI inflation [99], and multiple-field models, such as the curvaton scenario [100–102]. As well as altering the single-field slow-roll predictions for the primordial curvature power spectrum (Eq. (23.48)) and the tensor-scalar ratio (Eq. (23.49)), they may introduce new quantities that vanish in single-field slow-roll models, such as isocurvature matter perturbations, corresponding to entropy fluctuations in the photon-to-matter ratio, at first order:

$$S_{\text{m}} = \frac{\delta n_{\text{m}}}{n_{\text{m}}} - \frac{\delta n_{\gamma}}{n_{\gamma}} = \frac{\delta \rho_{\text{m}}}{\rho_{\text{m}}} - \frac{3}{4} \frac{\delta \rho_{\gamma}}{\rho_{\gamma}}. \quad (23.85)$$

Another possibility is non-Gaussianity in the distribution of the primordial curvature perturbation (see Chap. 29, “Cosmic Microwave Background” review), encoded in higher-order correlators such as the primordial bispectrum [103]

$$\langle \zeta(\mathbf{k})\zeta(\mathbf{k}')\zeta(\mathbf{k}'') \rangle \equiv (2\pi)^3 \delta(\mathbf{k} + \mathbf{k}' + \mathbf{k}'') B_{\zeta}(k, k', k''), \quad (23.86)$$

which is often expressed in terms of a dimensionless non-linearity parameter

$$f_{\text{NL}} \propto B_{\zeta}(k, k', k'') / P_{\zeta}(k) P_{\zeta}(k').$$

The three-point function (Eq. (23.86)) can be thought of as defined on a triangle whose sides are $\mathbf{k}, \mathbf{k}', \mathbf{k}''$, of which only two are independent, since they sum to zero. Further assuming statistical isotropy ensures that the bispectrum depends only on the magnitudes of the three vectors, k, k' and k'' . The search for f_{NL} and other non-Gaussian effects was a prime objective of the *Planck* data analysis [104, 105].

23.7.1 Enhanced perturbations on small scales

Although the observed primordial power spectrum on CMB scales is constrained to be almost scale invariant, consistent with the simplest slow-roll inflation models, the form of the power spectrum is largely unconstrained on much smaller scales, corresponding to scales that would have left the horizon closer to the end of inflation. Although free-streaming and other damping of perturbations in the primordial plasma are expected to have washed out small-scale fluctuations in the matter sector [106], gravitational relics such as primordial black holes [107, 108] or induced gravitational waves [109–112] produced by sufficiently large primordial density perturbations could have survived and might be detectable in the present day. Rapid growth of the power spectrum on small scales requires violation of slow roll in single-field models [113]. This could arise from features in the scalar field potential, such as an inflection point [114], leading to a transient phase of ultra-slow-roll evolution [115, 116] which can give a k^4 -rise in the scalar power spectrum on small scales [117, 118].

The LIGO-Virgo-KAGRA collaboration has placed an upper limit on such a stochastic gravitational wave background (SGWB), conventionally expressed in terms of the effective density of gravitational waves per logarithmic frequency interval, of $\Omega_{\text{GW}} \leq 5.8 \times 10^{-9}$ for a scale-invariant spectrum in the frequency range 20 – 77 Hz [119]. On the other hand, pulsar timing arrays have found growing evidence for a low-frequency SGWB at nano-Hertz frequencies [120–123]. While

this could be associated with an astrophysical background due to supermassive black hole binary systems, the NANOGrav collaboration report that a cosmological background provides a better fit to their 15-year dataset [124]. This could be generated by large primordial density fluctuations on the corresponding comoving scales, $k \sim 10^6 \text{ Mpc}^{-1}$, but these are constrained by the production of primordial black holes of approximately solar mass [125] and possible one-loop or higher-order corrections to the primordial power spectrum on CMB scales.

23.7.2 Effective Field Theory of Inflation

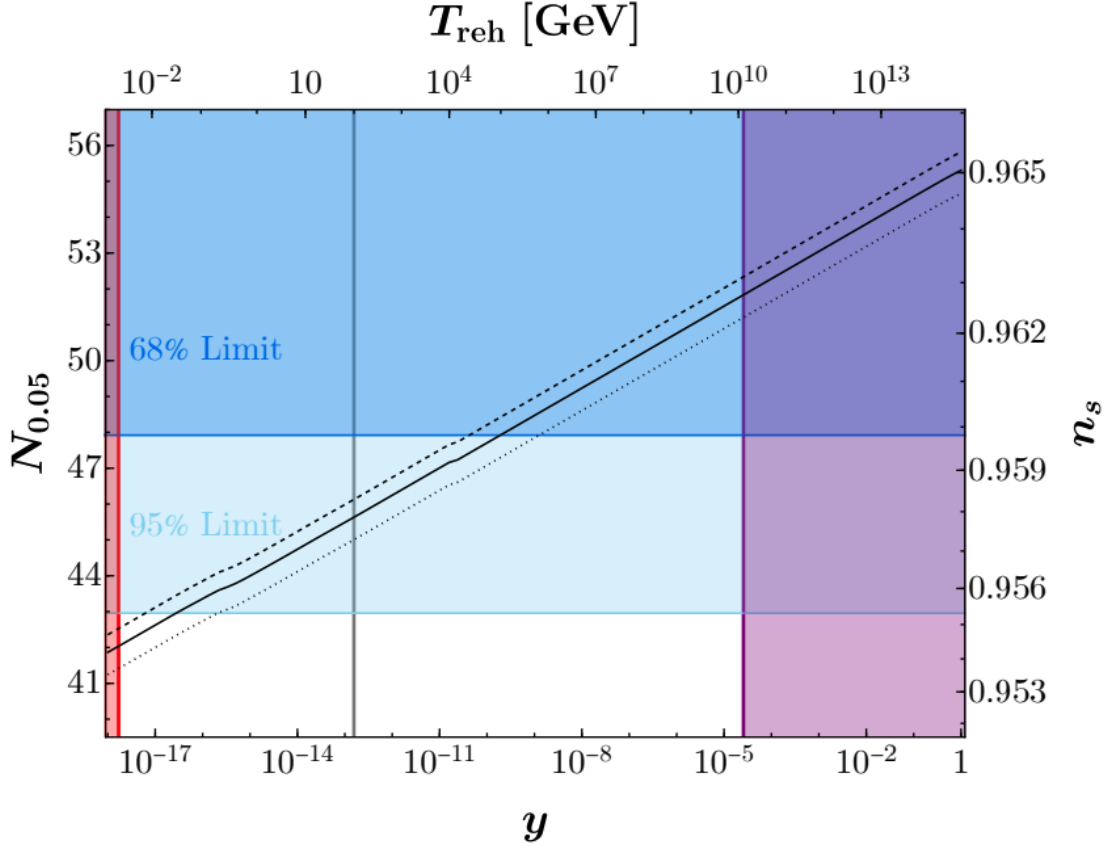


Figure 23.5: Illustration of the impact of the BICEP/Keck [48] and other constraints on the inflaton decay coupling, y , and the number of e-folds at a scale 0.05 Mpc^{-1} , $N_{0.05}$, in α -attractor models of inflation. The horizontal blue lines are 68% and 95% C.L. lower limits on n_s . The left and right axes show the relation between $N_{0.05}$ and n_s and the top and bottom scales show the relation between the inflaton coupling, y , and the reheating temperature, T_{reh} . The diagonal dashed, solid and dotted black lines illustrate the correlations between these quantities for α -attractor models with $\alpha = 0.1, 1$, and 5 . We also include the lower limit on y from Big-Bang nucleosynthesis (red line), the constraint that T_{reh} be no smaller than the electroweak scale (gray line), and a constraint from gravitino production (purple line) for $\alpha = 1$, which strengthens for smaller α . Plot taken from Ref. [92].

Since slow-roll inflation is a phase of accelerated expansion with an almost constant Hubble parameter, one may think of inflation in terms of an effective theory where the de Sitter spacetime symmetry is spontaneously broken down to RW symmetry by the time-evolution of the Hubble rate, $\dot{H} \neq 0$. There is then a Goldstone boson, π , associated with the spontaneous breaking of

time-translation invariance, which can be used to study model-independent properties of inflation. The Goldstone boson describes a spacetime-dependent shift of the time coordinate, corresponding to an adiabatic perturbation of the matter fields:

$$\delta\phi_i(t, \vec{x}) = \phi_i(t + \pi(t, \vec{x})) - \phi_i(t). \quad (23.87)$$

Thus adiabatic field fluctuations can be absorbed into the spatial metric perturbation, $\mathcal{R}$ in Eq. (23.28) at first order, in the comoving gauge:

$$\mathcal{R} = -H\pi, \quad (23.88)$$

where we define π on spatially-flat hypersurfaces. In terms of inflaton field fluctuations, we can identify $\pi \equiv \delta\phi/\dot{\phi}$, but in principle this analysis is not restricted to inflation driven by scalar fields.

The low-energy effective action for π can be obtained by writing down the most general Lorentz-invariant action and expanding in terms of π . The second-order effective action for the free-field wave modes, π_k , to leading order in slow roll, is then

$$S_\pi^{(2)} = - \int d^4x \sqrt{-g} \frac{M_{\text{P}}^2 \dot{H}}{c_s^2} \left[\dot{\pi}_k^2 - \frac{c_s^2}{R^2} (\nabla\pi)^2 \right], \quad (23.89)$$

where ϵ_H is the Hubble slow-roll parameter (Eq. (23.11)). We identify c_s^2 with an effective sound speed, generalising canonical slow-roll inflation, which is recovered in the limit $c_s^2 \rightarrow 1$.

The scalar power spectrum on super-Hubble scales (Eq. (23.48)) is enhanced for a reduced sound speed, leading to a reduced tensor-scalar ratio (Eq. (23.49))

$$\mathcal{P}_\zeta(k) \simeq \frac{4\pi}{M_{\text{P}}^2} \frac{1}{c_s^2 \epsilon} \left(\frac{H}{2\pi} \right)_*^2, \quad r \simeq 16(c_s^2 \epsilon)_*. \quad (23.90)$$

At third perturbative order and to lowest order in derivatives, one obtains [126]

$$S_\pi^{(3)} = \int d^4x \sqrt{-g} \frac{M_{\text{P}}^2 (1 - c_s^2) \dot{H}}{c_s^2} \left[\frac{\dot{\pi} (\nabla\pi)^2}{R^2} - \left(1 + \frac{2\tilde{c}_3}{3c_s^2} \right) \dot{\pi}^3 \right]. \quad (23.91)$$

Note that this expression vanishes for canonical fields with $c_s^2 = 1$. For $c_s^2 \neq 1$ the cubic action is determined by the sound speed and an additional parameter $\tilde{c}_3$. Both terms in the cubic action give rise to primordial bispectra that are well approximated by equilateral bispectra. However, the shapes are not identical, so one can find a linear combination for which the equilateral bispectra of each term cancel, giving rise to a distinctive orthogonal-type bispectrum [126].

Analysis based on *Planck* 2018 temperature and polarization data has placed bounds on several bispectrum shapes including equilateral and orthogonal shapes [105]:

$$f_{\text{NL}}^{\text{equil}} = -26 \pm 47, \quad f_{\text{NL}}^{\text{orthog}} = -38 \pm 24 \quad (68\% \text{ CL}). \quad (23.92)$$

For the simplest case of a constant sound speed, and marginalising over $\tilde{c}_3$, this provides a bound on the inflaton sound speed [105]

$$c_s \geq 0.021 \quad (95\% \text{ CL}). \quad (23.93)$$

For a specific model such as DBI inflation [99], corresponding to $\tilde{c}_3 = 3(1 - c_s^2)/2$, one obtains a tighter bound [105]:

$$c_s^{\text{DBI}} \geq 0.086 \quad (95\% \text{ CL}). \quad (23.94)$$

The *Planck* team analysed a wide range of non-Gaussian templates from different inflation models, including tests for deviations from an initial Bunch-Davies vacuum state, direction-dependent non-Gaussianity, and feature models with oscillatory bispectra [105]. No individual feature or resonance is above the 3σ significance level after accounting for the look-elsewhere effect. These results are consistent with the simplest canonical, slow-roll inflation models, but do not rule out most alternative models; rather, bounds on primordial non-Gaussianity place important constraints on the parameter space for non-canonical models.

23.7.3 Multi-Field Fluctuations

There is a very large literature on two- and multi-field models of inflation, most of which lies beyond the scope of this review [127, 128]. However, two important general topics merit being mentioned here, namely residual isocurvature perturbations and the possibility of non-Gaussian effects in the primordial perturbations.

One might expect that other scalar fields besides the inflaton might have non-negligible values that evolve and fluctuate in parallel with the inflaton, without necessarily making the dominant contribution to the energy density during the inflationary epoch. However, the energy density in such a field might persist beyond the end of inflation before decaying, at which point it might come to dominate (or at least make a non-negligible contribution to) the total energy density. In such a case, its fluctuations could end up generating the density perturbations detected in the CMB. This could occur due to a late-decaying scalar field [100–102] or a field fluctuation that modulates the end of inflation [129] or the inflaton decay [130].

23.7.3.1 Isocurvature Perturbations

Primordial perturbations arising in single-field slow-roll inflation are necessarily adiabatic, *i.e.*, they affect the overall density without changing the ratios of different contributions, such as the photon-matter ratio, $\delta(n_\gamma/n_m)/(n_\gamma/n_m)$. This is because inflaton perturbations represent a local shift of the time, as described in section Sec. 23.7.2:

$$\pi = \frac{\delta n_\gamma}{\dot{n}_\gamma} = \frac{\delta n_m}{\dot{n}_m}. \quad (23.95)$$

However, any light scalar field (*i.e.*, one with effective mass less than the Hubble scale) acquires a spectrum of nearly scale-invariant perturbations during inflation. Fluctuations orthogonal to the inflaton in field space are decoupled from the inflaton at Hubble exit, but can affect the subsequent evolution of the density perturbation. In particular, they can give rise to local variations in the equation of state (non-adiabatic pressure perturbations) that can alter the primordial curvature perturbation ζ on super-Hubble scales. Since these fluctuations are statistically independent of the inflaton perturbations at leading order in slow roll [128], non-adiabatic field fluctuations can only increase the scalar power spectrum with respect to adiabatic perturbations at Hubble exit, while leaving the tensor modes unaffected at first perturbative order. Thus, the single-field result for the tensor-scalar ratio (Eq. (23.49)) becomes an inequality [131]

$$r \leq 16\epsilon_*. \quad (23.96)$$

Hence, an observational upper bound on the tensor-to-scalar ratio does not bound the slow-roll parameter ϵ in multi-field models.

If all the scalar fields present during inflation eventually decay completely into fully thermalized radiation, these field fluctuations are converted fully into adiabatic perturbations in the primordial plasma [132]. On the other hand, non-adiabatic field fluctuations can also leave behind primordial

isocurvature perturbations (Eq. (23.85)) after inflation. In multi-field inflation models it is thus possible for non-adiabatic field fluctuations to generate both curvature and isocurvature perturbations leading to correlated primordial perturbations [133].

The amplitudes of any primordial isocurvature perturbations (Eq. (23.85)) are strongly constrained by the current CMB data, especially on large angular scales. Using temperature and low- ℓ polarization data yields the following bound on the amplitude of cold dark matter isocurvature perturbations at scale $k = 0.002h^{-1}\text{Mpc}^{-1}$ (marginalising over the correlation angle and in the absence of primordial tensor perturbations) [40]:

$$\frac{\mathcal{P}_{S_m}}{\mathcal{P}_\zeta + \mathcal{P}_{S_m}} < 0.025 \text{ (95\% CL)}. \quad (23.97)$$

For fully (anti-)correlated isocurvature perturbations, corresponding to a single isocurvature field providing a source for both the curvature and residual isocurvature perturbations, the bounds become significantly tighter [40]:

$$\frac{\mathcal{P}_{S_m}}{\mathcal{P}_\zeta + \mathcal{P}_{S_m}} < 0.0002 \text{ (95\% CL), correlated}, \quad (23.98)$$

$$\frac{\mathcal{P}_{S_m}}{\mathcal{P}_\zeta + \mathcal{P}_{S_m}} < 0.003 \text{ (95\% CL), anti-correlated} \quad (23.99)$$

23.7.3.2 Local-Type Non-Gaussianity

Since non-adiabatic field fluctuations in multi-field inflation may lead to evolution of the primordial curvature perturbation at all orders, it becomes possible to generate significant non-Gaussianity in the primordial curvature perturbation. Non-linear evolution on super-Hubble scales leads to local-type non-Gaussianity, where the local integrated expansion is a non-linear function of the local field values during inflation, $N(\phi_i)$. While the field fluctuations at Hubble exit, $\delta\phi_{i*}$, are Gaussian in the slow-roll limit, the curvature perturbation, $\zeta = \delta N$, becomes a non-Gaussian distribution [134]:

$$\zeta = \sum_i \frac{\partial N}{\partial \phi_i} \delta\phi_i + \frac{1}{2} \sum_{i,j} \frac{\partial^2 N}{\partial \phi_i \partial \phi_j} \delta\phi_i \delta\phi_j + \dots \quad (23.100)$$

with non-vanishing bispectrum in the squeezed limit ($k_1 \approx k_2 \gg k_3$):

$$B_\zeta(k_1, k_2, k_3) \approx \frac{12}{5} f_{\text{NL}}^{\text{local}} \frac{\mathcal{P}_\zeta(k_1)}{4\pi k_1^3} \frac{\mathcal{P}_\zeta(k_3)}{4\pi k_3^3}, \quad (23.101)$$

where

$$\frac{6}{5} f_{\text{NL}}^{\text{local}} = \frac{\sum_{i,j} \frac{\partial^2 N}{\partial \phi_i \partial \phi_j}}{\left(\sum_i \frac{\partial N}{\partial \phi_i}\right)^2}. \quad (23.102)$$

Both equilateral and orthogonal bispectra, discussed above in the context of generalised single-field inflation, vanish in the squeezed limit, enabling the three types of non-Gaussianity to be distinguished by observations, in principle.

Non-Gaussianity during multi-field inflation is highly model dependent, although $f_{\text{NL}}^{\text{local}}$ can often be smaller than unity in multi-field slow-roll inflation [135]. Scenarios where a second light field plays a role during or after inflation can make distinctive predictions for $f_{\text{NL}}^{\text{local}}$, such as $f_{\text{NL}}^{\text{local}} = -5/4$ in some curvaton scenarios [134, 136] or $f_{\text{NL}}^{\text{local}} = 5$ in simple modulated reheating scenarios [130, 137]. By contrast the constancy of ζ on super-Hubble scales in single-field slow-roll inflation leads to a

very small non-Gaussianity [138, 139], and in the squeezed limit we have the simple result $f_{\text{NL}}^{\text{local}} = 5(1 - n_s)/12$ [140, 141].

A combined analysis of the *Planck* 2018 temperature and polarization data [105] yields the following range for $f_{\text{NL}}^{\text{local}}$ defined in (Eq. (23.102)):

$$f_{\text{NL}}^{\text{local}} = -1 \pm 5 \quad (68\% \text{ CL}). \quad (23.103)$$

This sensitivity is sufficient to rule out parameter regimes giving rise to relatively large non-Gaussianity, but insufficient to probe $f_{\text{NL}}^{\text{local}} = \mathcal{O}(\epsilon)$, as expected in single-field models, or the range $f_{\text{NL}}^{\text{local}} = \mathcal{O}(1)$ found in the simplest two-field models.

Local-type primordial non-Gaussianity can also give rise to a striking scale-dependent bias in the distribution of collapsed dark matter halos and thus the galaxy distribution [142, 143]. Bounds from high-redshift galaxy surveys are not currently competitive with the best CMB constraints. An analysis based on the clustering of quasars in the final data release (DR16) of the extended Baryon acoustic Oscillation Spectroscopic Survey (eBOSS) [144] yields

$$f_{\text{NL}}^{\text{local}} = -12 \pm 21 \quad (68\% \text{ CL}), \quad (23.104)$$

while a recent analysis based on the clustering of luminous red galaxies in imaging surveys from the Dark Energy Spectroscopic Instrument (DESI) [145] finds $-10 < f_{\text{NL}}^{\text{local}} < 58$ at 68% CL.

23.8 Initial Conditions and Fine-tuning

This review is based on the assumption that the inflationary paradigm is valid. However, it remains the object of many criticisms (see, *e.g.*, Ref. [146]), many of them related to the perceived unnaturalness of the required initial conditions.

Most work on inflation is done in the context of RW cosmology, which assumes a high degree of symmetry, or small inhomogeneous perturbations (usually first order) about an RW cosmology. The isotropic RW space-time is an attractor for many homogeneous but anisotropic cosmologies in the presence of a false vacuum energy density [147], or a scalar field with suitable self-interaction potential energy [148, 149]. However it is much harder to establish the range of highly inhomogeneous initial conditions that yield a successful RW Universe, at least in limited studies initially (see, *e.g.*, Refs [150, 151]). A related open question is the general nature of the pre-inflationary state of the inflaton and other fields that could have provided initial conditions suitable for inflation [146]. They would need to have satisfied non-trivial homogeneity and isotropy conditions, and one may ask how these could have arisen, whether they are plausible, and whether there may be some observable signature of the pre-inflationary state. These and other criticisms of inflation were addressed in Ref. [152], which presented studies of the sensitivity of inflation to the initial conditions. Complementing these studies, there have been numerical relativity investigations of highly inhomogeneous initial conditions [153–155]. The general conclusion is that inflation is rather robust with respect to inhomogeneities in the initial conditions in both the scalar field profile and the extrinsic curvature, including large tensor perturbations.

To quantify the fine-tuning of initial conditions requires a measure in the space of possible cosmologies [156]; however, it has been argued that some of the measures historically used to frame this problem are formally invalid [157]. It is sometimes also objected that inflationary models predict the existence of a multiverse, and potentially a loss of predictive power [158], if it undergoes the process termed eternal inflation [159–161]. However, whether this is actually a bug or a feature remains a topic of debate [162, 163]. The existence of the multiverse is a purely philosophical problem, unless it has observable consequences, *e.g.*, in the CMB.

One might expect signatures of any pre-inflationary state to appear at large angular scales, *i.e.*, low multipoles ℓ . Indeed, various anomalies have been noted in the large-scale CMB anisotropies,

as also discussed in Chap. 29, the “Cosmic Microwave Background” review, including a possible suppression of the quadrupole and other very large-scale anisotropies, an apparent feature in the range $\ell \approx 20$ to 30, and a possible hemispheric asymmetry. However, none of these are highly significant statistically, in view of the limitations due to cosmic variance [164]. They cannot yet be regarded as signatures of initial conditions, the multiverse or some pre-inflationary dynamics, such as might emerge from string theory.

A different kind of initial condition problem, called the trans-Planckian problem [165], is that the perturbations now seen in the CMB would have had wavelengths shorter than the Planck length at the onset of inflation. However, under quite general and conservative assumptions the usual inflationary predictions are quite robust [166], although with the possibility of $\mathcal{O}((H/m_{\text{P}})^n)$ corrections that might have interesting signatures in the CMB [167].

When inflation was first proposed [5, 6] there was no evidence for the existence of scalar fields or the accelerated expansion of the Universe. The situation has changed dramatically in recent years with the observational evidence that the cosmic expansion is currently accelerating and with the discovery of a scalar particle, namely the Higgs boson (see Chap. 11, “Status of Higgs boson physics” review). Combined with the lack of any widely accepted alternative model for the origin of cosmic structure, these discoveries have lent support to the idea of a primordial accelerated expansion driven by a scalar field, *i.e.*, cosmological inflation. In parallel, successive CMB experiments have been consistent with generic predictions of inflationary models, although without yet providing irrefutable evidence. It was concluded in Ref. [152] that the inflationary paradigm is not currently in trouble. However, we note that inflation via a formally elementary scalar inflaton should probably only be regarded as an effective field theory valid at energy densities hierarchically smaller than the Planck scale. It should eventually be embedded in a suitable ultraviolet completion, on which inflationary dynamics may be our clearest window.

23.9 Future Probes of Inflation

Prospective future CMB experiments, both ground- and space-based are reviewed in the separate PDG “Cosmic Microwave Background” review, Chap. 29. The main emphasis in CMB experiments in the coming years will be on ground-based experiments providing improved measurements of B -mode polarization and greater sensitivity to the tensor-to-scalar ratio r , and more precise measurements at higher ℓ that will constrain n_s better. As is apparent from Fig. 23.1 and the discussion of models such as R^2 inflation, there is a strong incentive to reach a $5\text{-}\sigma$ sensitivity to $r \sim 3$ to 4×10^{-3} . This could be achieved with a moderately-sized space mission with large sky coverage [168], along with improvements in de-lensing and foreground measurements. The discussion in Sec. 23.3 (see also Fig. 23.5), also brought out the importance of reducing the uncertainty in n_s , as a way to constrain post-inflationary reheating and the connection to particle physics. CMB temperature anisotropies probe primordial density perturbations down to comoving scales of order 50 Mpc, beyond which scale secondary sources of anisotropy dominate. CMB spectral distortions could potentially constrain the amplitude and shape of primordial density perturbations on comoving scales from kpc to Mpc due to distortions caused by the Silk damping of pressure waves in the radiation dominated era, before the last scattering of the CMB photons but after the plasma can be fully thermalised [106].

Improved sensitivity to non-Gaussianities is also a priority. In addition to CMB measurements, future large-scale structure surveys will also have roles to play as probes into models of inflation, for which there are excellent prospects. Current surveys such as DESI may reach $\Delta f_{\text{NL}} \sim 4$ [169] comparable with the *Planck* sensitivity. In the future, radio surveys such as SKA will measure large-scale structure out to redshift $z \sim 3$ [170], initially through mapping the intensity of the neutral hydrogen 21-cm line, and eventually through radio galaxy surveys that will probe local-

type non-Gaussianity to $f_{\text{NL}} \sim 1$.

Galaxy clustering using DESI and *Euclid* satellite data could also constrain the running of the scalar tilt to a precision of $\Delta\alpha_s \approx 0.0028$, a factor of 2 improvement on *Planck* constraints, or a precision of 0.0016 using Rubin-LSST data [169]. This might allow us to confirm, or refute, the prediction from the simplest inflationary models that running should be negative [171].

The satellite mission *SPHEREx* [172] will use measurements of the galaxy power spectrum to target an estimate of the running of the scalar spectral index with a sensitivity $\Delta\alpha_s \sim 10^{-3}$ and local-type primordial non-Gaussianity, $\Delta f_{\text{NL}} \sim 1$. Including information from the galaxy bispectrum one might reduce the measurement error on non-Gaussianity to $\Delta f_{\text{NL}} \sim 0.2$, making it possible to distinguish between single-field slow-roll models and alternatives such as the curvaton scenario for the origin of structure, which generate $f_{\text{NL}} \sim 1$.

Finally we note that future gravitational wave experiments such as the Laser Interferometer Space Antenna (LISA) will be sensitive to a stochastic gravitational wave background, including tensor perturbations produced during inflation or induced by large primordial density perturbations on small scales after inflation [173], which could provide constraints on inflationary perturbations on smaller scales.

Acknowledgements

The work of J.E. was supported in part by the UK STFC research grant ST/T000759/1. The work of D.W. was supported in part by the UK STFC research grant ST/W001225/1.

References

- [1] D. H. Lyth and A. Riotto, *Phys. Rept.* **314**, 1 (1999), [[hep-ph/9807278](#)].
- [2] A.R. Liddle and D.H. Lyth, *Cosmological inflation and large-scale structure* (Cambridge University Press, 2000).
- [3] D. Baumann, in “Physics of the large and the small, TASI 09, proceedings of the Theoretical Advanced Study Institute in Elementary Particle Physics, Boulder, Colorado, USA, 1-26 June 2009,” 523–686 (2011), [[arXiv:0907.5424](#)].
- [4] J. Martin, C. Ringeval and V. Vennin, *Phys. Dark Univ.* **5-6**, 75 (2014), [[arXiv:1303.3787](#)].
- [5] A. A. Starobinsky, *Phys. Lett.* **B91**, 99 (1980), [*Adv. Ser. Astrophys. Cosmol.*3,130(1987)].
- [6] A. H. Guth, *Phys. Rev.* **D23**, 347 (1981), [*Adv. Ser. Astrophys. Cosmol.*3,139(1987)].
- [7] K. A. Olive, *Phys. Rept.* **190**, 307 (1990).
- [8] J. Martin *et al.*, *JCAP* **1403**, 039 (2014), [[arXiv:1312.3529](#)].
- [9] J. Martin, *Astrophys. Space Sci. Proc.* **45**, 41 (2016), [[arXiv:1502.05733](#)].
- [10] P. A. R. Ade *et al.* (Planck), *Astron. Astrophys.* **594**, A13 (2016), [[arXiv:1502.01589](#)].
- [11] A. H. Guth and E. J. Weinberg, *Nucl. Phys.* **B212**, 321 (1983).
- [12] D. La and P. J. Steinhardt, *Phys. Rev. Lett.* **62**, 376 (1989), [Erratum: *Phys. Rev. Lett.*62,1066(1989)].
- [13] A. D. Linde, *Phys. Lett.* **B249**, 18 (1990).
- [14] F. C. Adams and K. Freese, *Phys. Rev.* **D43**, 353 (1991), [[hep-ph/0504135](#)].
- [15] A. D. Linde, *Phys. Lett.* **108B**, 389 (1982), [*Adv. Ser. Astrophys. Cosmol.*3,149(1987)].
- [16] A. Albrecht and P. J. Steinhardt, *Phys. Rev. Lett.* **48**, 1220 (1982), [*Adv. Ser. Astrophys. Cosmol.*3,158(1987)].
- [17] W. H. Press, *Phys. Scripta* **21**, 702 (1980); S. W. Hawking, *Phys. Lett.* **115B**, 295 (1982); A. A. Starobinsky, *Phys. Lett.* **117B**, 175 (1982); A. H. Guth and S. Y. Pi, *Phys. Rev. Lett.* **49**, 1110 (1982).

- [18] J. M. Bardeen, P. J. Steinhardt and M. S. Turner, *Phys. Rev.* **D28**, 679 (1983).
- [19] V. F. Mukhanov and G. V. Chibisov, *JETP Lett.* **33**, 532 (1981), [*Pisma Zh. Eksp. Teor. Fiz.*33,549(1981)].
- [20] K. S. Stelle, *Gen. Rel. Grav.* **9**, 353 (1978).
- [21] B. Whitt, *Phys. Lett.* **145B**, 176 (1984).
- [22] D. Wands, *Class. Quant. Grav.* **11**, 269 (1994), [[arXiv:gr-qc/9307034](#)].
- [23] S. R. Coleman and F. De Luccia, *Phys. Rev.* **D21**, 3305 (1980).
- [24] M. Sasaki *et al.*, *Phys. Lett.* **B317**, 510 (1993).
- [25] M. Bucher, A. S. Goldhaber and N. Turok, *Phys. Rev.* **D52**, 3314 (1995), [[hep-ph/9411206](#)].
- [26] A. D. Linde and A. Mezhlumian, *Phys. Rev.* **D52**, 6789 (1995), [[arXiv:astro-ph/9506017](#)].
- [27] A. R. Liddle, P. Parsons and J. D. Barrow, *Phys. Rev.* **D50**, 7222 (1994), [[arXiv:astro-ph/9408015](#)].
- [28] L. Kofman, A. D. Linde and A. A. Starobinsky, *Phys. Rev.* **D56**, 3258 (1997), [[hep-ph/9704452](#)].
- [29] B. A. Bassett, S. Tsujikawa and D. Wands, *Rev. Mod. Phys.* **78**, 537 (2006), [[arXiv:astro-ph/0507632](#)].
- [30] A. D. Linde, *Phys. Rev.* **D49**, 748 (1994), [[arXiv:astro-ph/9307002](#)].
- [31] A. D. Dolgov and A. D. Linde, *Phys. Lett.* **116B**, 329 (1982).
- [32] J. R. Ellis *et al.*, *Nucl. Phys.* **B238**, 453 (1984), [,223(1983)].
- [33] M. Kawasaki and T. Moroi, *Prog. Theor. Phys.* **93**, 879 (1995), [[hep-ph/9403364](#)].
- [34] J. H. Traschen and R. H. Brandenberger, *Phys. Rev.* **D42**, 2491 (1990).
- [35] G. V. Chibisov and V. F. Mukhanov, *Mon. Not. Roy. Astron. Soc.* **200**, 535 (1982).
- [36] H. Kodama and M. Sasaki, *Prog. Theor. Phys. Suppl.* **78**, 1 (1984).
- [37] V.F. Mukhanov, H.A. Feldman and R.H. Brandenberger, *Phys. Rept.* **215**, 203 (1992).
- [38] K. A. Malik and D. Wands, *Phys. Rept.* **475**, 1 (2009), [[arXiv:0809.4944](#)].
- [39] V. F. Mukhanov, *Sov. Phys. JETP* **67**, 1297 (1988), [*Zh. Eksp. Teor. Fiz.*94N7,1(1988)].
- [40] Y. Akrami *et al.* (Planck), *Astron. Astrophys.* **641**, A10 (2020), [[arXiv:1807.06211](#)].
- [41] A. A. Starobinsky, *JETP Lett.* **30**, 682 (1979), [,767(1979)].
- [42] E. D. Stewart and D. H. Lyth, *Phys. Lett.* **B302**, 171 (1993), [[arXiv:gr-qc/9302019](#)].
- [43] D. Wands *et al.*, *Phys. Rev.* **D62**, 043527 (2000), [[arXiv:astro-ph/0003278](#)].
- [44] A. R. Liddle and S. M. Leach, *Phys. Rev.* **D68**, 103503 (2003), [[arXiv:astro-ph/0305263](#)].
- [45] T. Louis *et al.* (ACT) (2025), [[arXiv:2503.14452](#)].
- [46] M. Abdul Karim *et al.* (DESI) (2025), [[arXiv:2503.14738](#)].
- [47] E. Camphuis *et al.* (SPT-3G) (2025), [[arXiv:2506.20707](#)].
- [48] P. A. R. Ade *et al.* (BICEP/Keck), *Phys. Rev. Lett.* **127**, 15, 151301 (2021), [[arXiv:2110.00483](#)].
- [49] S.M. Leach, A.R. Liddle, J. Martin, and D.J. Schwarz, *Phys. Rev.* **D66**, 23515 (2002).
- [50] D. Kazanas, *Astrophys. J.* **241**, L59 (1980).
- [51] K. Sato, *Mon. Not. Roy. Astron. Soc.* **195**, 467 (1981).
- [52] A. Billoire and K. Tamvakis, *Nucl. Phys.* **B200**, 329 (1982); J. D. Breit, S. Gupta and A. Zaks, *Phys. Rev. Lett.* **51**, 1007 (1983).

- [53] A. D. Linde, *Phys. Lett.* **129B**, 177 (1983).
- [54] D. V. Nanopoulos, K. A. Olive and M. Srednicki, *Phys. Lett.* **127B**, 30 (1983).
- [55] J. Ellis, D. V. Nanopoulos and K. A. Olive, *Phys. Rev. Lett.* **111**, 111301 (2013), [Erratum: *Phys. Rev. Lett.* 111, no. 12, 129902 (2013)], [[arXiv:1305.1247](#)].
- [56] T. S. Koivisto and F. R. Urban, *JCAP* **1503**, 03, 003 (2015), [[arXiv:1407.3445](#)].
- [57] L. Boubekeur and D. H. Lyth, *JCAP* **0507**, 010 (2005), [[hep-ph/0502047](#)].
- [58] G. R. Dvali, Q. Shafi and S. Solganik, in “4th European Meeting From the Planck Scale to the Electroweak Scale (Planck 2001) La Londe les Maures, Toulon, France, May 11-16, 2001,” (2001), [[hep-th/0105203](#)]; J. Garcia-Bellido, R. Rabadan and F. Zamora, *JHEP* **01**, 036 (2002), [[hep-th/0112147](#)]; S. Kachru *et al.*, *JCAP* **0310**, 013 (2003), [[hep-th/0308055](#)].
- [59] F. C. Adams *et al.*, *Phys. Rev.* **D47**, 426 (1993), [[hep-ph/9207245](#)].
- [60] E. Pajer and M. Peloso, *Class. Quant. Grav.* **30**, 214002 (2013), [[arXiv:1305.3557](#)].
- [61] E. Silverstein and A. Westphal, *Phys. Rev.* **D78**, 106003 (2008), [[arXiv:0803.3085](#)].
- [62] L. McAllister, E. Silverstein and A. Westphal, *Phys. Rev.* **D82**, 046003 (2010), [[arXiv:0808.0706](#)].
- [63] F. L. Bezrukov and M. Shaposhnikov, *Phys. Lett.* **B659**, 703 (2008), [[arXiv:0710.3755](#)].
- [64] D. Buttazzo *et al.*, *JHEP* **12**, 089 (2013), [[arXiv:1307.3536](#)].
- [65] H. P. Nilles, *Phys. Rept.* **110**, 1 (1984); H. E. Haber and G. L. Kane, *Phys. Rept.* **117**, 75 (1985).
- [66] J.R. Ellis *et al.*, *Phys. Lett.* **118B**, 335 (1982).
- [67] N. Okada and Q. Shafi (2013), [[arXiv:1311.0921](#)].
- [68] J. Martin, C. Ringeval and V. Vennin, *Phys. Rev.* **D93**, 10, 103532 (2016), [[arXiv:1603.02606](#)].
- [69] J. Martin, C. Ringeval and V. Vennin, *JCAP* **07**, 087 (2024), [[arXiv:2404.10647](#)].
- [70] Y. Akrami *et al.* (Planck), *Astron. Astrophys.* **641**, A10 (2020), [[arXiv:1807.06211](#)].
- [71] H. Murayama *et al.*, *Phys. Rev. Lett.* **70**, 1912 (1993).
- [72] J. Ellis, M. Fairbairn and M. Sueiro, *JCAP* **1402**, 044 (2014), [[arXiv:1312.1353](#)].
- [73] D. Croon, J. Ellis and N. E. Mavromatos, *Phys. Lett.* **B724**, 165 (2013), [[arXiv:1303.6253](#)].
- [74] G. R. Dvali, Q. Shafi and R. K. Schaefer, *Phys. Rev. Lett.* **73**, 1886 (1994), [[hep-ph/9406319](#)].
- [75] D. V. Nanopoulos *et al.*, *Phys. Lett.* **123B**, 41 (1983).
- [76] A. B. Goncharov and A. D. Linde, *Phys. Lett.* **139B**, 27 (1984).
- [77] E. Cremmer *et al.*, *Nucl. Phys.* **B212**, 413 (1983), [[413\(1982\)](#)].
- [78] R. Holman, P. Ramond and G. G. Ross, *Phys. Lett.* **137B**, 343 (1984).
- [79] E. J. Copeland *et al.*, *Phys. Rev.* **D49**, 6410 (1994), [[arXiv:astro-ph/9401011](#)]; E. D. Stewart, *Phys. Rev.* **D51**, 6847 (1995), [[hep-ph/9405389](#)].
- [80] G. D. Coughlan *et al.*, *Phys. Lett.* **131B**, 59 (1983); A. S. Goncharov, A. D. Linde and M. I. Vysotsky, *Phys. Lett.* **147B**, 279 (1984); T. Banks, D. B. Kaplan and A. E. Nelson, *Phys. Rev.* **D49**, 779 (1994), [[hep-ph/9308292](#)]; B. de Carlos *et al.*, *Phys. Lett.* **B318**, 447 (1993), [[hep-ph/9308325](#)]; M. Kawasaki, T. Moroi and T. Yanagida, *Phys. Lett.* **B370**, 52 (1996), [[hep-ph/9509399](#)].
- [81] J. R. Ellis, D. V. Nanopoulos and M. Quiros, *Phys. Lett.* **B174**, 176 (1986).
- [82] T. Moroi, M. Yamaguchi and T. Yanagida, *Phys. Lett.* **B342**, 105 (1995), [[hep-ph/9409367](#)].

- [83] M. Kawasaki, M. Yamaguchi and T. Yanagida, *Phys. Rev. Lett.* **85**, 3572 (2000), [[hep-ph/0004243](#)]; K. Nakayama, F. Takahashi and T. T. Yanagida, *JCAP* **1308**, 038 (2013), [[arXiv:1305.5099](#)].
- [84] E. Cremmer *et al.*, *Phys. Lett.* **133B**, 61 (1983).
- [85] A. S. Goncharov and A. D. Linde, *Class. Quant. Grav.* **1**, L75 (1984); C. Kounnas and M. Quiros, *Phys. Lett.* **151B**, 189 (1985).
- [86] J. R. Ellis, C. Kounnas and D. V. Nanopoulos, *Nucl. Phys.* **B247**, 373 (1984).
- [87] E. Witten, *Phys. Lett.* **155B**, 151 (1985).
- [88] J. Ellis, D. V. Nanopoulos and K. A. Olive, *JCAP* **1310**, 009 (2013), [[arXiv:1307.3537](#)].
- [89] J. Ellis *et al.*, *Class. Quant. Grav.* **33**, 9, 094001 (2016), [[arXiv:1507.02308](#)].
- [90] R. Kallosh, A. Linde and D. Roest, *JHEP* **11**, 198 (2013), [[arXiv:1311.0472](#)].
- [91] M. Drees and Y. Xu, *Phys. Lett. B* **867**, 139612 (2025), [[arXiv:2504.20757](#)].
- [92] J. Ellis *et al.*, *Phys. Rev. D* **105**, 4, 043504 (2022), [[arXiv:2112.04466](#)].
- [93] L. Iacconi *et al.* (2023), [[arXiv:2306.00918](#)].
- [94] J. Ellis *et al.*, *JCAP* **1507**, 07, 050 (2015), [[arXiv:1505.06986](#)].
- [95] A. R. Liddle, *Mon. Not. Roy. Astron. Soc.* **377**, L74 (2007), [[arXiv:astro-ph/0701113](#)].
- [96] J. Martin, C. Ringeval and V. Vennin, ASPIC: Accurate Slow-roll Predictions for Inflationary CosmologyAstrophysics Source Code Library, record ascl:1806.031, <https://github.com/cosmicinflation/aspic> (2018).
- [97] J. Martin, C. Ringeval and V. Vennin, *Phys. Rev. Lett.* **114**, 8, 081303 (2015), [[arXiv:1410.7958](#)].
- [98] C. Armendariz-Picon, T. Damour and V. F. Mukhanov, *Phys. Lett.* **B458**, 209 (1999), [[hep-th/9904075](#)].
- [99] M. Alishahiha, E. Silverstein and D. Tong, *Phys. Rev.* **D70**, 123505 (2004), [[hep-th/0404084](#)].
- [100] K. Enqvist and M. S. Sloth, *Nucl. Phys.* **B626**, 395 (2002), [[hep-ph/0109214](#)].
- [101] D. H. Lyth and D. Wands, *Phys. Lett.* **B524**, 5 (2002), [[hep-ph/0110002](#)].
- [102] T. Moroi and T. Takahashi, *Phys. Lett.* **B522**, 215 (2001), [Erratum: *Phys. Lett.* **B539**, 303 (2002)], [[hep-ph/0110096](#)].
- [103] N. Bartolo *et al.*, *Phys. Rept.* **402**, 103 (2004), [[arXiv:astro-ph/0406398](#)].
- [104] P. A. R. Ade *et al.* (Planck), *Astron. Astrophys.* **594**, A17 (2016), [[arXiv:1502.01592](#)].
- [105] Y. Akrami *et al.* (Planck), *Astron. Astrophys.* **641**, A9 (2020), [[arXiv:1905.05697](#)].
- [106] J. Chluba, J. Hamann and S. P. Patil, *Int. J. Mod. Phys.* **D24**, 10, 1530023 (2015), [[arXiv:1505.01834](#)].
- [107] Y. B. Zel'dovich and I. D. Novikov, *Soviet Astron. AJ (Engl. Transl.)*, **10**, 602 (1967).
- [108] S. W. Hawking, *Nature* **248**, 30 (1974).
- [109] S. Matarrese, O. Pantano and D. Saez, *Phys. Rev. Lett.* **72**, 320 (1994), [[arXiv:astro-ph/9310036](#)].
- [110] S. Matarrese, S. Mollerach and M. Bruni, *Phys. Rev. D* **58**, 043504 (1998), [[arXiv:astro-ph/9707278](#)].
- [111] K. N. Ananda, C. Clarkson and D. Wands, *Phys. Rev. D* **75**, 123518 (2007), [[arXiv:gr-qc/0612013](#)].

- [112] D. Baumann *et al.*, *Phys. Rev. D* **76**, 084019 (2007), [[hep-th/0703290](#)].
- [113] H. Motohashi and W. Hu, *Phys. Rev. D* **96**, 6, 063503 (2017), [[arXiv:1706.06784](#)].
- [114] J. Garcia-Bellido and E. Ruiz Morales, *Phys. Dark Univ.* **18**, 47 (2017), [[arXiv:1702.03901](#)].
- [115] S. Inoue and J. Yokoyama, *Phys. Lett. B* **524**, 15 (2002), [[hep-ph/0104083](#)].
- [116] W. H. Kinney, *Phys. Rev. D* **72**, 023515 (2005), [[arXiv:gr-qc/0503017](#)].
- [117] S. M. Leach *et al.*, *Phys. Rev. D* **64**, 023512 (2001), [[arXiv:astro-ph/0101406](#)].
- [118] C. T. Byrnes, P. S. Cole and S. P. Patil, *JCAP* **06**, 028 (2019), [[arXiv:1811.11158](#)].
- [119] R. Abbott *et al.* (KAGRA, Virgo, LIGO Scientific), *Phys. Rev. D* **104**, 2, 022004 (2021), [[arXiv:2101.12130](#)].
- [120] G. Agazie *et al.* (NANOGrav), *Astrophys. J. Lett.* **951**, 1, L8 (2023), [[arXiv:2306.16213](#)].
- [121] J. Antoniadis *et al.* (EPTA) (2023), [[arXiv:2306.16214](#)].
- [122] D. J. Reardon *et al.*, *Astrophys. J. Lett.* **951**, 1, L6 (2023), [[arXiv:2306.16215](#)].
- [123] H. Xu *et al.*, *Res. Astron. Astrophys.* **23**, 7, 075024 (2023), [[arXiv:2306.16216](#)].
- [124] A. Afzal *et al.* (NANOGrav), *Astrophys. J. Lett.* **951**, 1, L11 (2023), [[arXiv:2306.16219](#)].
- [125] B. Carr *et al.*, *Rept. Prog. Phys.* **84**, 11, 116902 (2021), [[arXiv:2002.12778](#)].
- [126] L. Senatore, K. M. Smith and M. Zaldarriaga, *JCAP* **1001**, 028 (2010), [[arXiv:0905.3746](#)].
- [127] C. Gordon *et al.*, *Phys. Rev.* **D63**, 023506 (2001), [[arXiv:astro-ph/0009131](#)]; R. Easther *et al.*, *Phys. Rev. Lett.* **112**, 161302 (2014), [[arXiv:1312.4035](#)]; J. Ellis *et al.*, *JCAP* **1501**, 010 (2015), [[arXiv:1409.8197](#)]; S. Renaux-Petel and K. Turzynski, *JCAP* **1506**, 06, 010 (2015), [[arXiv:1405.6195](#)].
- [128] C. T. Byrnes and D. Wands, *Phys. Rev.* **D74**, 043529 (2006), [[arXiv:astro-ph/0605679](#)].
- [129] D. H. Lyth, *JCAP* **0511**, 006 (2005), [[arXiv:astro-ph/0510443](#)].
- [130] G. Dvali, A. Gruzinov and M. Zaldarriaga, *Phys. Rev.* **D69**, 023505 (2004), [[arXiv:astro-ph/0303591](#)].
- [131] D. Wands *et al.*, *Phys. Rev.* **D66**, 043520 (2002), [[arXiv:astro-ph/0205253](#)].
- [132] S. Weinberg, *Phys. Rev.* **D67**, 123504 (2003), [[arXiv:astro-ph/0302326](#)].
- [133] D. Langlois, *Phys. Rev.* **D59**, 123512 (1999), [[arXiv:astro-ph/9906080](#)].
- [134] D. H. Lyth and Y. Rodriguez, *Phys. Rev. Lett.* **95**, 121302 (2005), [[arXiv:astro-ph/0504045](#)].
- [135] F. Vernizzi and D. Wands, *JCAP* **0605**, 019 (2006), [[arXiv:astro-ph/0603799](#)].
- [136] M. Sasaki, J. Valiviita and D. Wands, *Phys. Rev.* **D74**, 103003 (2006), [[arXiv:astro-ph/0607627](#)].
- [137] G. Dvali, A. Gruzinov and M. Zaldarriaga, *Phys. Rev.* **D69**, 083505 (2004), [[arXiv:astro-ph/0305548](#)].
- [138] D. S. Salopek and J. R. Bond, *Phys. Rev.* **D43**, 1005 (1991).
- [139] A. Gangui *et al.*, *Astrophys. J.* **430**, 447 (1994), [[arXiv:astro-ph/9312033](#)].
- [140] J. M. Maldacena, *JHEP* **05**, 013 (2003), [[arXiv:astro-ph/0210603](#)].
- [141] V. Acquaviva *et al.*, *Nucl. Phys.* **B667**, 119 (2003), [[arXiv:astro-ph/0209156](#)].
- [142] N. Dalal *et al.*, *Phys. Rev.* **D77**, 123514 (2008), [[arXiv:0710.4560](#)].
- [143] S. Matarrese and L. Verde, *Astrophys. J.* **677**, L77 (2008), [[arXiv:0801.4826](#)].
- [144] E.-M. Mueller *et al.* (2021), [[arXiv:2106.13725](#)].

- [145] M. Rezaie *et al.* (2023), [[arXiv:2307.01753](#)].
- [146] A. Ijjas, P. J. Steinhardt and A. Loeb, *Phys. Lett.* **B723**, 261 (2013), [[arXiv:1304.2785](#)].
- [147] R. M. Wald, *Phys. Rev.* **D28**, 2118 (1983).
- [148] M. Heusler, *Phys. Lett.* **B253**, 33 (1991).
- [149] Y. Kitada and K.-i. Maeda, *Phys. Rev.* **D45**, 1416 (1992).
- [150] D. S. Goldwirth and T. Piran, *Phys. Rept.* **214**, 223 (1992).
- [151] T. Vachaspati and M. Trodden, *Phys. Rev.* **D61**, 023502 (1999), [[arXiv:gr-qc/9811037](#)].
- [152] D. Chowdhury *et al.*, *Phys. Rev. D* **100**, 8, 083537 (2019), [[arXiv:1902.03951](#)].
- [153] W. E. East *et al.*, *JCAP* **1609**, 09, 010 (2016), [[arXiv:1511.05143](#)].
- [154] K. Clough *et al.*, *JCAP* **1709**, 09, 025 (2017), [[arXiv:1608.04408](#)].
- [155] K. Clough, R. Flauger and E. A. Lim, *JCAP* **1805**, 05, 065 (2018), [[arXiv:1712.07352](#)].
- [156] G. W. Gibbons, S. W. Hawking and J. M. Stewart, *Nucl. Phys.* **B281**, 736 (1987).
- [157] J. S. Schiffrin and R. M. Wald, *Phys. Rev.* **D86**, 023521 (2012), [[arXiv:1202.1818](#)].
- [158] A. Ijjas, P. J. Steinhardt and A. Loeb, *Phys. Lett.* **B736**, 142 (2014), [[arXiv:1402.6980](#)].
- [159] A. Vilenkin, *Phys. Rev.* **D27**, 2848 (1983).
- [160] A. D. Linde, *Phys. Lett.* **B175**, 395 (1986).
- [161] A. S. Goncharov, A. D. Linde and V. F. Mukhanov, *Int. J. Mod. Phys.* **A2**, 561 (1987).
- [162] A. H. Guth, D. I. Kaiser and Y. Nomura, *Phys. Lett.* **B733**, 112 (2014), [[arXiv:1312.7619](#)].
- [163] A. Linde, in “Proceedings, 100th Les Houches Summer School: Post-Planck Cosmology: Les Houches, France, July 8 - August 2, 2013,” 231–316 (2015), [[arXiv:1402.0526](#)].
- [164] P. A. R. Ade *et al.* (Planck), *Astron. Astrophys.* **594**, A20 (2016), [[arXiv:1502.02114](#)].
- [165] J. Martin and R. H. Brandenberger, *Phys. Rev.* **D63**, 123501 (2001), [[hep-th/0005209](#)].
- [166] R. H. Brandenberger and J. Martin, *Mod. Phys. Lett.* **A16**, 999 (2001), [[arXiv:astro-ph/0005432](#)].
- [167] J. Martin and R. H. Brandenberger, in “Recent developments in theoretical and experimental general relativity, gravitation and relativistic field theories. Proceedings, 9th Marcel Grossmann Meeting, MG’9, Rome, Italy, July 2-8, 2000. Pts. A-C,” 2001–2002 (2000), [[arXiv:astro-ph/0012031](#)].
- [168] A. Kogut *et al.*, *JCAP* **1107**, 025 (2011), [[arXiv:1105.2044](#)].
- [169] A. Font-Ribera *et al.*, *JCAP* **1405**, 023 (2014), [[arXiv:1308.4164](#)].
- [170] R. Maartens *et al.* (SKA Cosmology SWG), *PoS AASKA14*, 016 (2015), [[arXiv:1501.04076](#)].
- [171] J. Martin, C. Ringeval and V. Vennin, *EPL* **148**, 2, 29002 (2024), [[arXiv:2404.15089](#)].
- [172] O. Doré *et al.* (2014), [[arXiv:1412.4872](#)].
- [173] P. Auclair *et al.* (LISA Cosmology Working Group) (2022), [[arXiv:2204.05434](#)].