Astro/Phys 224          Spring 2014

# Origin and Evolution of the Universe

## Week 9

# ΛCDM Small Scale Issues; Cosmic Inflation

## Joel Primack
### University of California, Santa Cruz

## Small-Scale Challenges to ΛCDM

**Many more small halos than observed small galaxies**
    **1) Field galaxies**
    **2) Satellite galaxies**

**Cusp-Core issue at centers of small galaxies**

**"Too Big to Fail" problem for satellite galaxies**

## Evidence Supporting ΛCDM

**Evidence that the large numbers of small subhalos predicted by ΛCDM actually exist:**
    **1) Gaps in cold stellar streams in the Milky Way**
    **2) Gravitational lensing "flux anomalies"**

**Velocity Function**

Discrepancy due to incomplete observations or ΛCDM failure?

theoretical VF with AC

theoretical VF without AC

observed VF (HIPASS + SDSS)

Bolshoi Sub-Halo Abundance Matching

Theory & Observations Agree Pretty Well
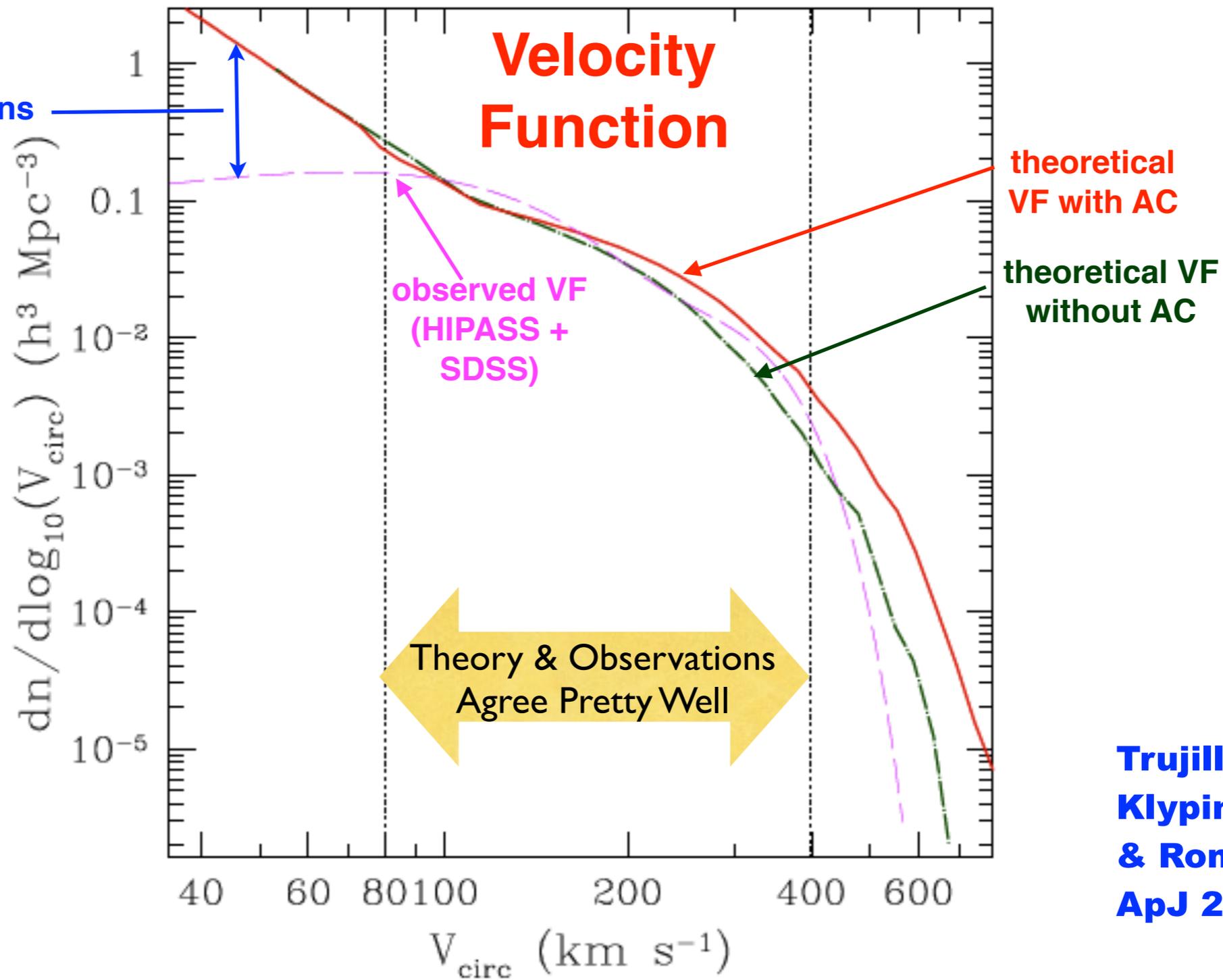
Trujillo-Gomez, Klypin, Primack, & Romanowsky ApJ 2011

Fig. 11.— Comparison of theoretical (dot-dashed and thick solid curves) and observational (dashed curve) circular velocity functions. The dot-dashed line shows the effect of adding the baryons (stellar and cold gas components) to the central region of each DM halo and measuring the circular velocity at 10 kpc. The thick solid line is the distribution obtained when the adiabatic contraction of the DM halos is considered. Because of uncertainties in the AC models, realistic theoretical predictions should lie between the dot-dashed and solid curves. Both the theory and observations are highly uncertain for rare galaxies with $V_{\rm circ} > 400$ km s$^{-1}$. Two vertical dotted lines divide the VF into three domains: $V_{\rm circ} > 400$ km s$^{-1}$ with large observational and theoretical uncertainties; $< 80$ km s$^{-1}$ $< V_{\rm circ} < 400$ km s$^{-1}$ with a reasonable agreement, and $V_{\rm circ} < 80$ km s$^{-1}$, where the theory significantly overpredicts the number of dwarfs.

## Klypin, Karachentsev, Nasonova 2012

**Deeper Local Survey -- better agreement with ΛCDM but still more halos than galaxies below 50 km/s**

## Local Volume: D <10 Mpc

Total sample:  813 galaxies
Within 10 Mpc:   686
    $M_B$<-13  N=304
    $M_B$<-10  N=611

80-90% are spirals or dIrr (T>0)

Errors of distances are 8-10%

80% with D<10Mpc have HI linewidth

Vrot =
    $150 \times 10^{-(20.5+M_B)/8.5}$ km/s



ΛCDM

A factor of two disagreement
at V = 40 km/s

No disagreement
for V > 60 km/s

Distribution of observed line-widths
(similar after correction for inclination)

$dN/d\log_{10}(V)$ $(h^{-1}Mpc)^{-3}$

V (km/s)

# Abundance of field Galaxies

Anatoly Klypin[1]*, Igor Karachenrsev[2], Dmitry Makarov[2], and Olga Nasonova[2]

[1]*New Mexico State University, Las Cruces, NM 88001, USA*
[2]*Special Astrophysical Observatory, Nizhny Arkhyz, Russia*

We present new measurements of the abundance of galaxies with a given circular velocity in the Local Volume: a region centered on the Milky Way Galaxy and extending to distance ~10 Mpc. The sample of ~ 800 mostly dwarf galaxies provides a unique opportunity to study the abundance and properties of galaxies down to absolute magnitudes $M_B \approx -10$, and virial masses $M_{vir} = 10^9 M_\odot$. We find that the standard $\Lambda$CDM model gives remarkably accurate estimates for the velocity function of galaxies with circular velocities $V \geq 60$ km s$^{-1}$ and corresponding virial masses $M_{vir} \geq 3 \times 10^{10} M_\odot$, but it badly fails by over-predicting ~ 5 times the abundance of large dwarfs with velocities $V = 30 - 50$ km s$^{-1}$. The Warm Dark Matter (WDM) models cannot explain the data either, regardless of mass of WDM particle. Though reminiscent to the known overabundance of satellites problem, the overabundance of field galaxies is a much more difficult problem. For the standard $\Lambda$CDM model to survive, in the 10 Mpc radius of the Milky Way there should be 1000 dark galaxies with virial mass $M_{vir} \approx 10^{10} M_\odot$, extremely low surface brightness and no detectable HI gas. So far none of this type of galaxies have been discovered.
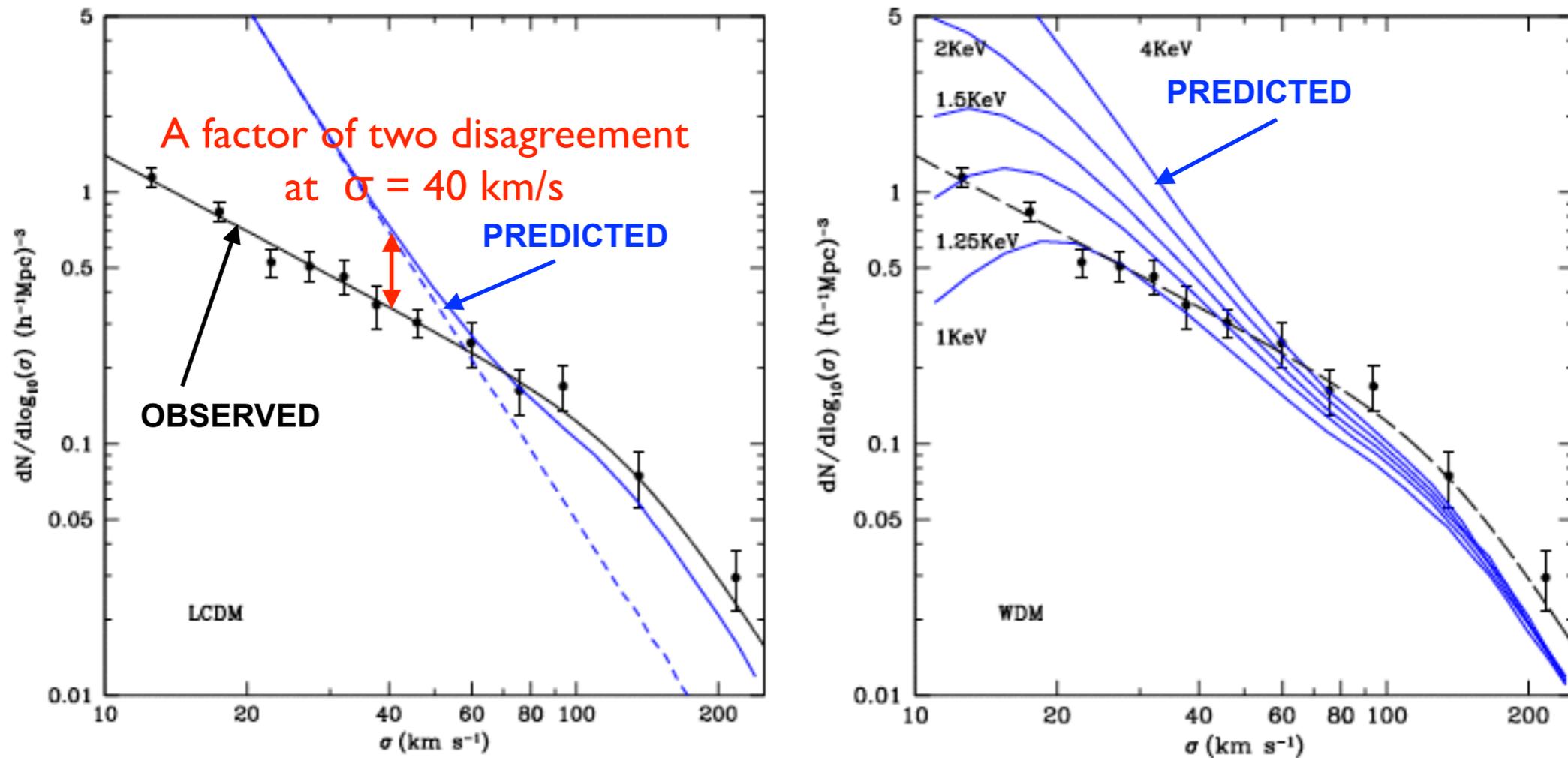
**Figure 6.** Comparison of the distribution function of line-widths $V_{los}$ for galaxies in the Local Volume with theoretical predictions for the LCDM (left panel) and the Warm Dark Matter models (right panel). *Left:* Filled circles and the full curve present velocity function for the 10 Mpc sample. Theoretical predictions for the $\Lambda$CDM model with the Planck cosmological parameters are presented by the upper full curve. The short-dashed curve shows the predictions of the dark matter-only estimates without correction for baryon infall. Enhanced mass of baryons (mostly due to stars) in the central halo regions results in the increase of the circular velocity observed in this plot as the shift from the dashed to the full curve.

# Can the **large** number of small-V<sub>circ</sub> subhalos be reconciled with the **small** number of faint galaxies?
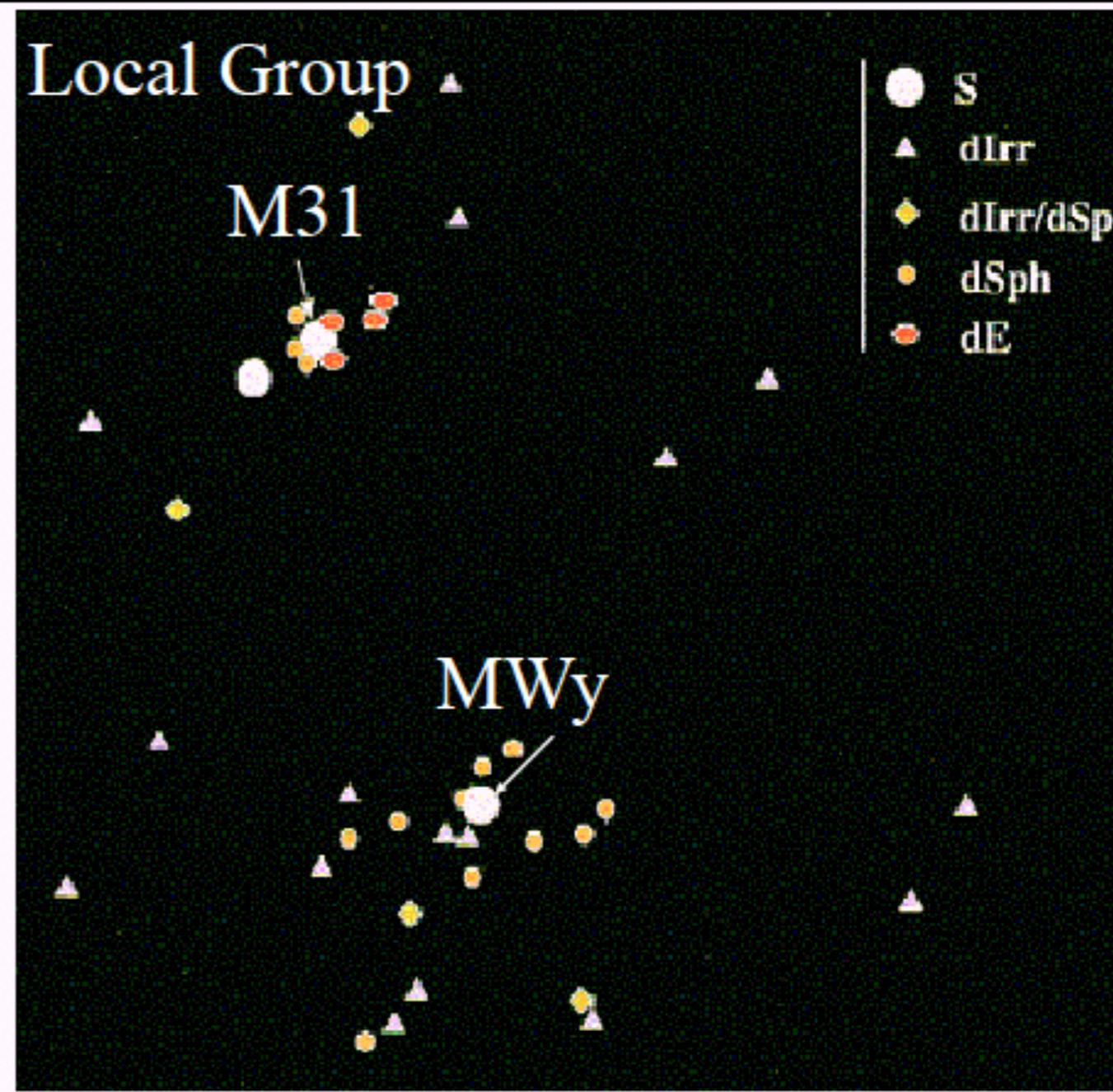
# Modeling of dwarfs

## 1. N-body (DM only) Simulations

- solve equations of gravity for particles of dark matter (& sometimes stars)

## 2. Hydrodynamic Simulations

- solve equations of gravity and hydrodynamics/thermodynamics for particles of dark matter and gas

## 3. Semi-Analytic Models

- treat gravity and "gastrophysics" via analytic approximations based on 1 & 2
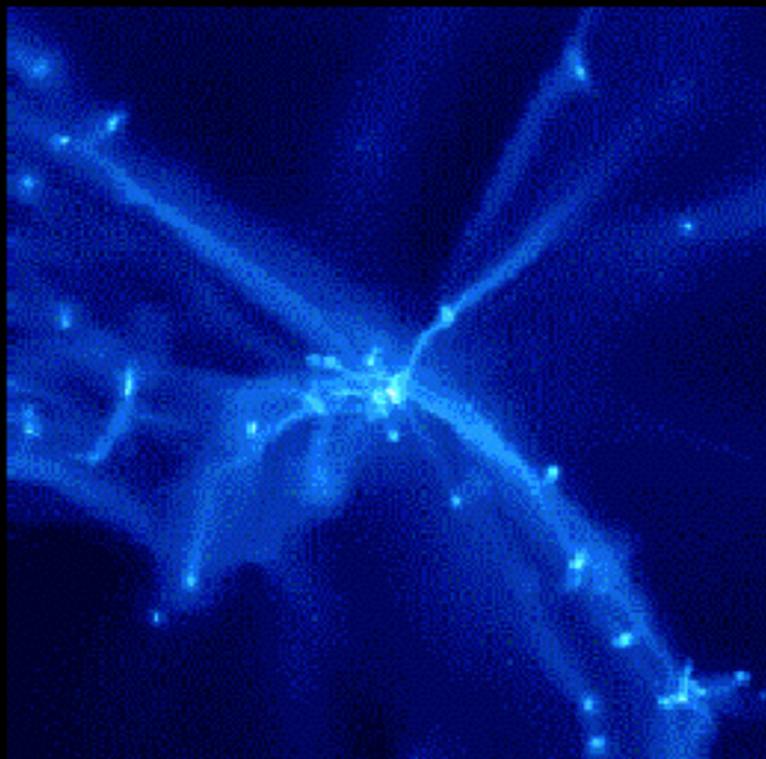
**Flores & Primack94 and Moore94 first pointed out that dark matter simulations have density ρ(r) ~ r$^\alpha$ at small r with α ≈ –1 ("cusp") while observed small spiral galaxies and clusters appeared to have α ≈ 0 ("core").**

**Governato+10,13 and the *Nature* review by Pontzen & Governato14 show that in high-resolution galaxy simulations, baryonic physics softens the central DM cusp to a core as long as enough stars form, M* ≥ 10$^7$ M$_\odot$.** This happens because of repeated episodes when the baryons cool and slowly fall into the galaxy center, and are then expelled rapidly (in less than a dynamical time) by energy released by stars and supernovae.

Observers (e.g., Walker & Peñarrubia11, Amorisco & Evans12) had agreed that the larger dwarf spheroidal Milky Way satellite galaxies such as Fornax (L ≈ 1.7x10$^7$ L$_\odot$) have cores, but recent papers (e.g., Breddels & Helmi13 A&A, Jardel & Gebhardt13, Richardson & Fairbairn14) have questioned this. **Thus the cusp-core question is now observational *and* theoretical.** Adams, Simon+14 find α ≈ - 0.5 for dwarf spirals, in agreement with recent high-resolution simulations with baryons.



**Pontzen & Governato14**
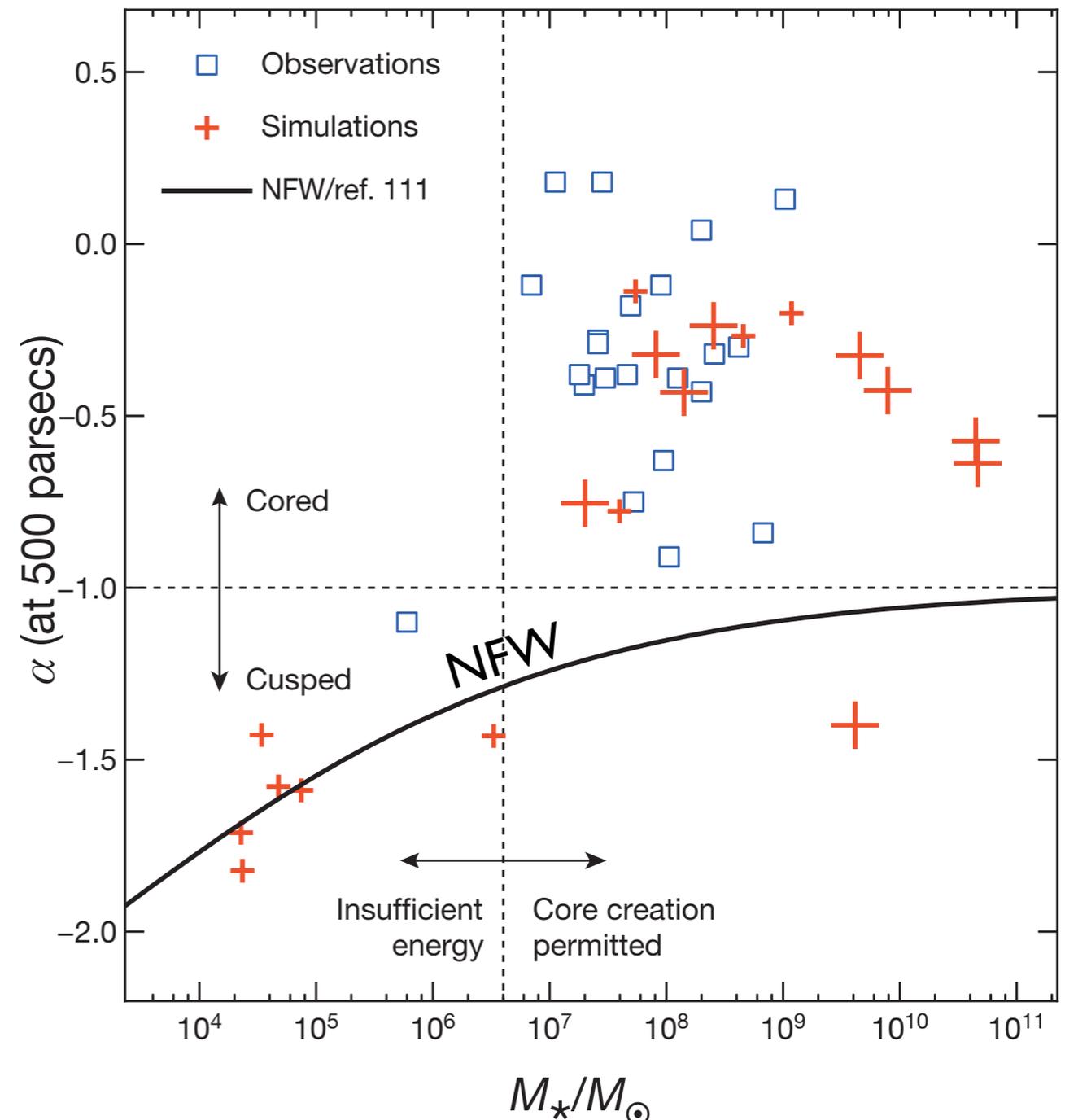
# Challenges: Cusp, Satellite Galaxies

In addition to the Governato... ptov+12, Brooks+13) there are several other important rece... 4, Trujillo-Gomez+14) arguing that baryonic effects convert ...-resolution simulation yet of **a dwarf spiral was describ... e Near Field-Deep Field Connections conference a ... star formation converted the central cusp to a core, red...**

$M_{vir} = 1E10 M_\odot$ at $z = 0$    $M_* = 4 \times 10^6 M_\odot$



**Bursty Star Formation**



Ultrafaint

— Collision-less HR
— Hydro HR

Dwarf

— Collision-less HR
---- Hydro* HR
— Hydro HR

Core

**Repeated episodes when baryons cool and slowly fall into the galaxy center, dominate the mass, and then are expelled rapidly (in less than $t_{dyn}$) by radiation pressure and supernovae, soften the central DM cusp to a core .**
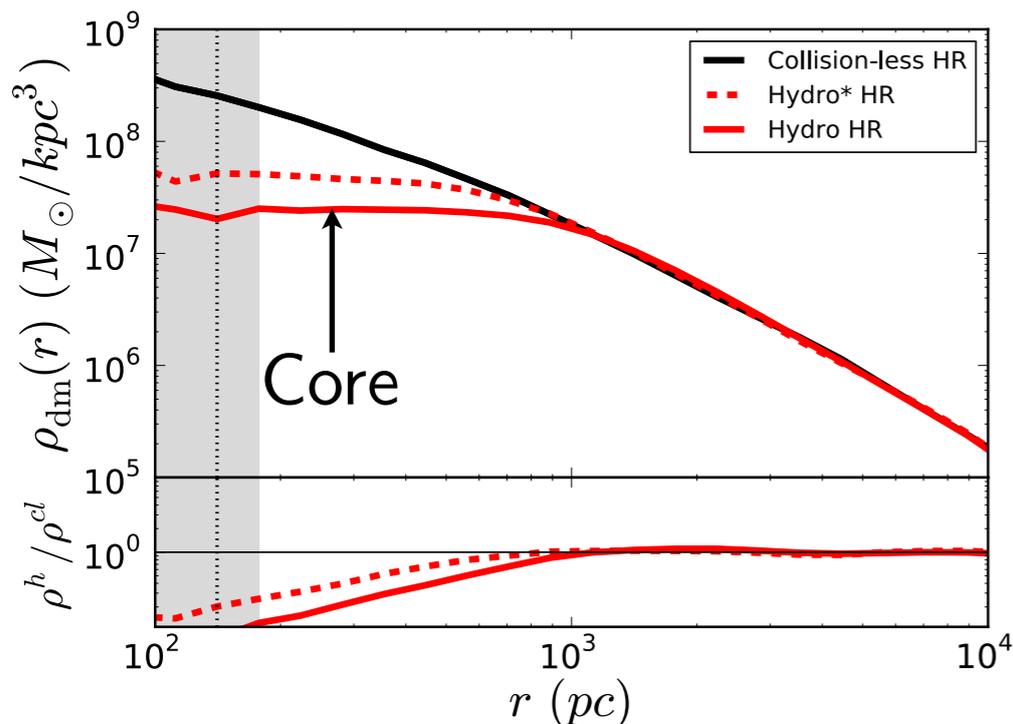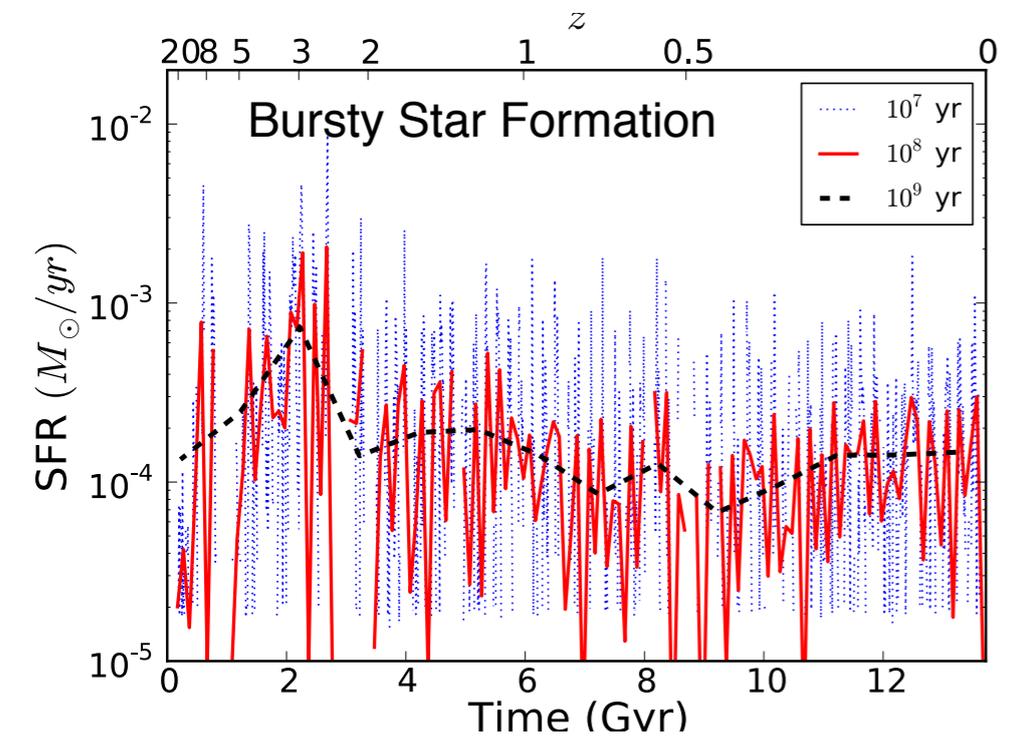
$M_{vir} = 3E9 M_\odot$ at $z = 0$
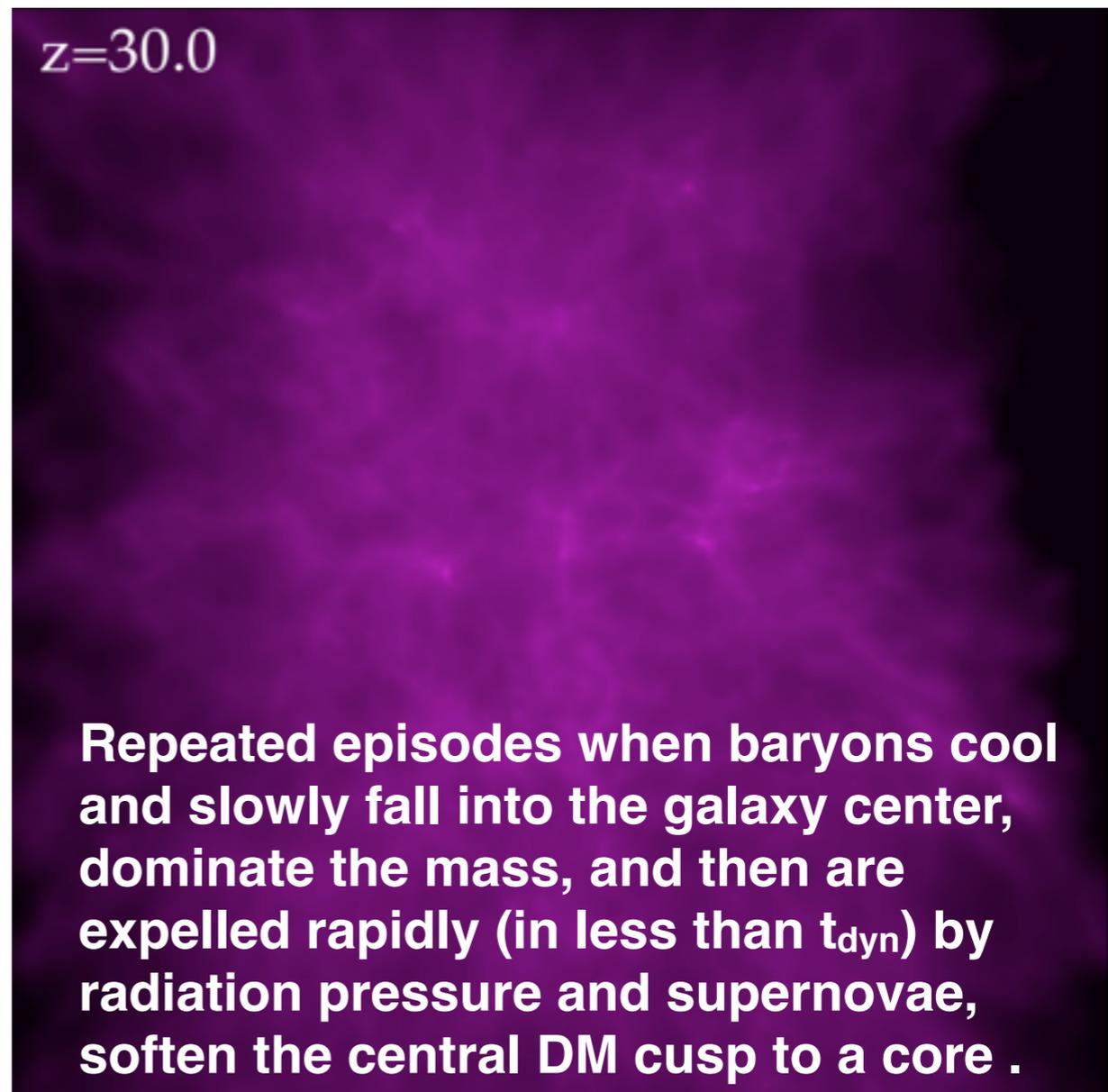
$M_{vir} = 1E10 M_\odot$ at $z = 0$

Enough Star Formation & Realistic Feedback May Solve Cusp-Core and TBTF



$M_* = 2 \times 10^4 M_\odot$
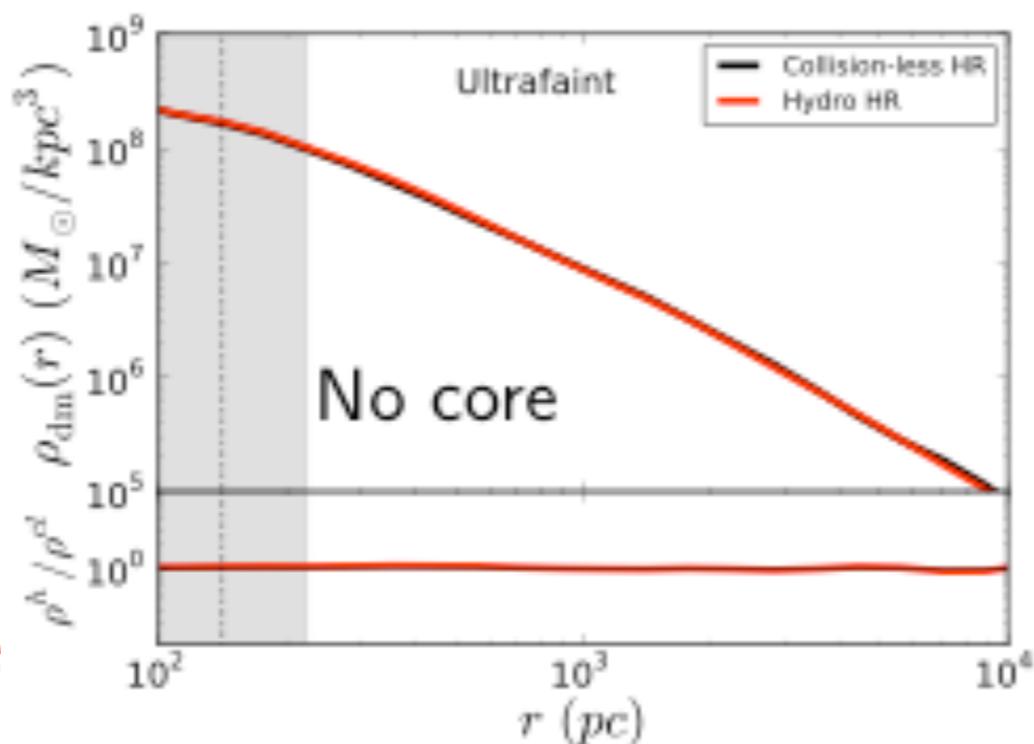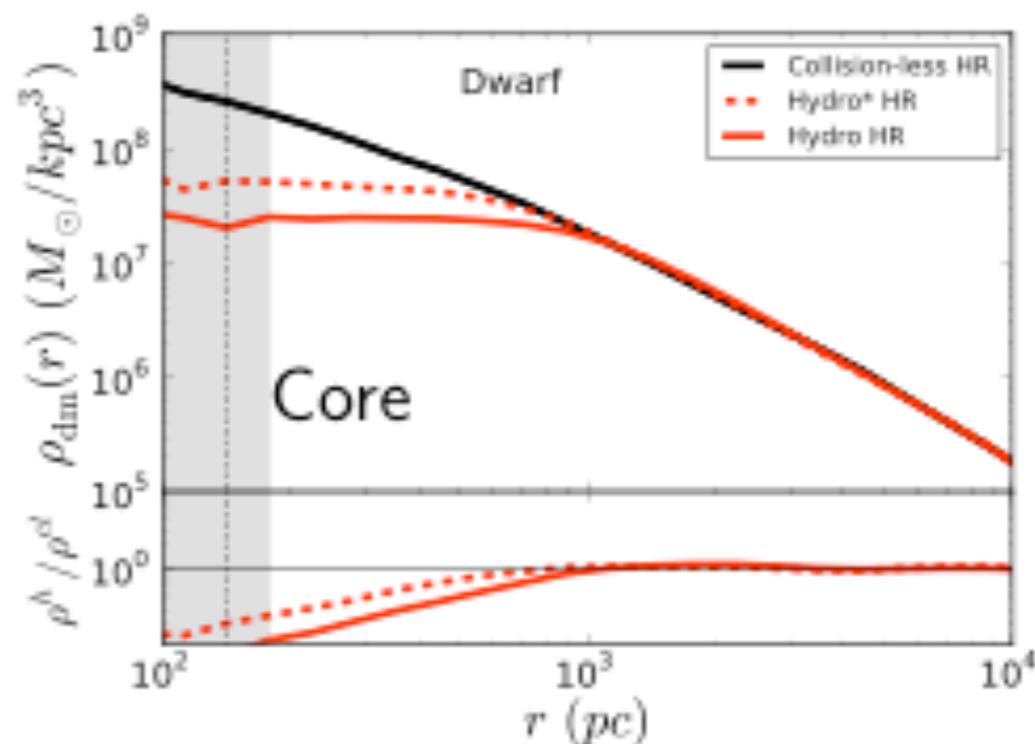
$M_* = 4 \times 10^6 M_\odot$

S. Trujillo-Gomez, A. Klypin, P. Colín, D. Ceverino, K. Arraki, & J. Primack

$M_{vir} = 3 \times 10^{10} M_\odot$ at $z = 0$
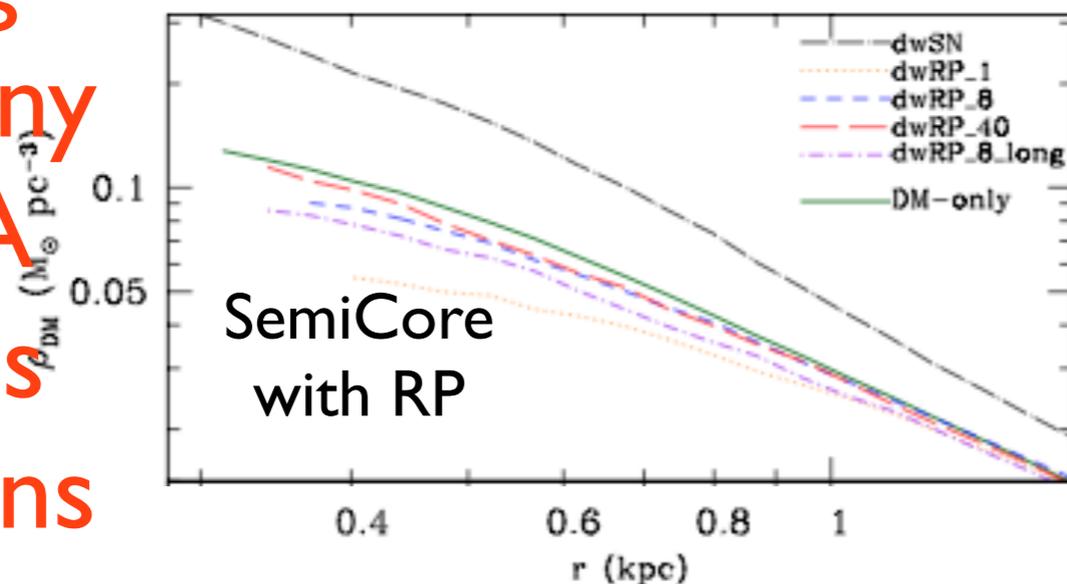
$M_{vir} = 2 \times 10^{11} M_\odot$ at $z = 0$
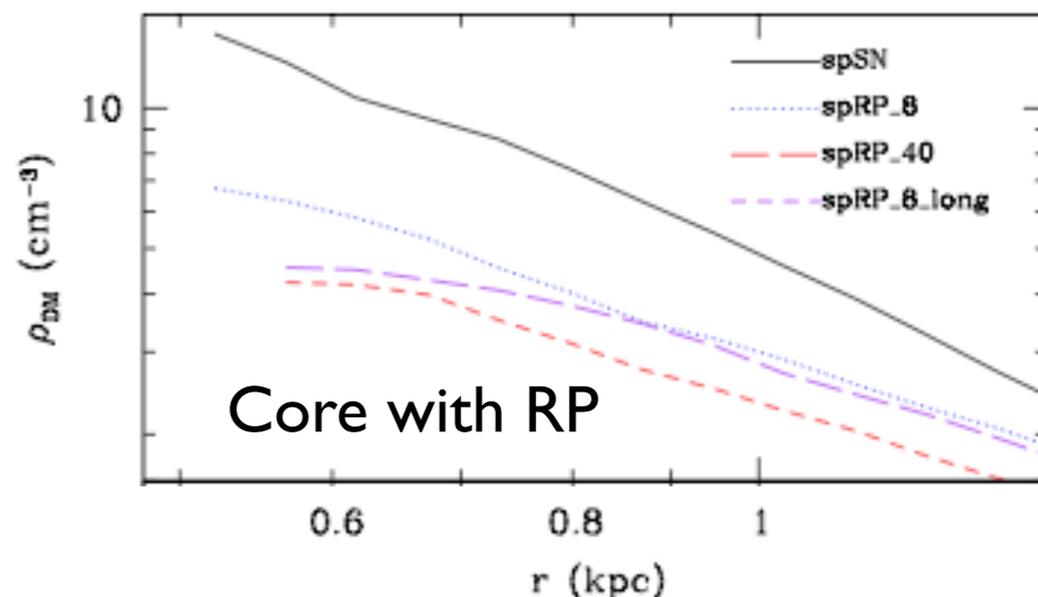
Coming Soon: Many AGORA High-Res Simulations

# The "too big to fail" problem

## ΛCDM subhalos vs. Milky Way satellites

"Missing satellites": Klypin et al. 1999, Moore et al. 1999



**Aquarius Simulation**

**Diameter of visible Milky Way**
**30 kpc = 100,000 light years**

**Diameter of Milky Way Dark Matter Halo**
**1.5 million light years**



dwarf satellites around the Milky Way

(◆ Leo T)

- ● Spiral galaxy
- ● dSph, dE
- ● new dSph
- ■ dIrr

250 kpc sphere

CVn I — Leo II — Leo I — Leo V — CVn II — Leo IV — Her — UMa I — Boö I — Coma — Sextans — Draco — Seg I — UMi — Sagittarius — LMC — Carina — SMC — Sculptor — Fornax

100 / 0 / 100kpc / 0
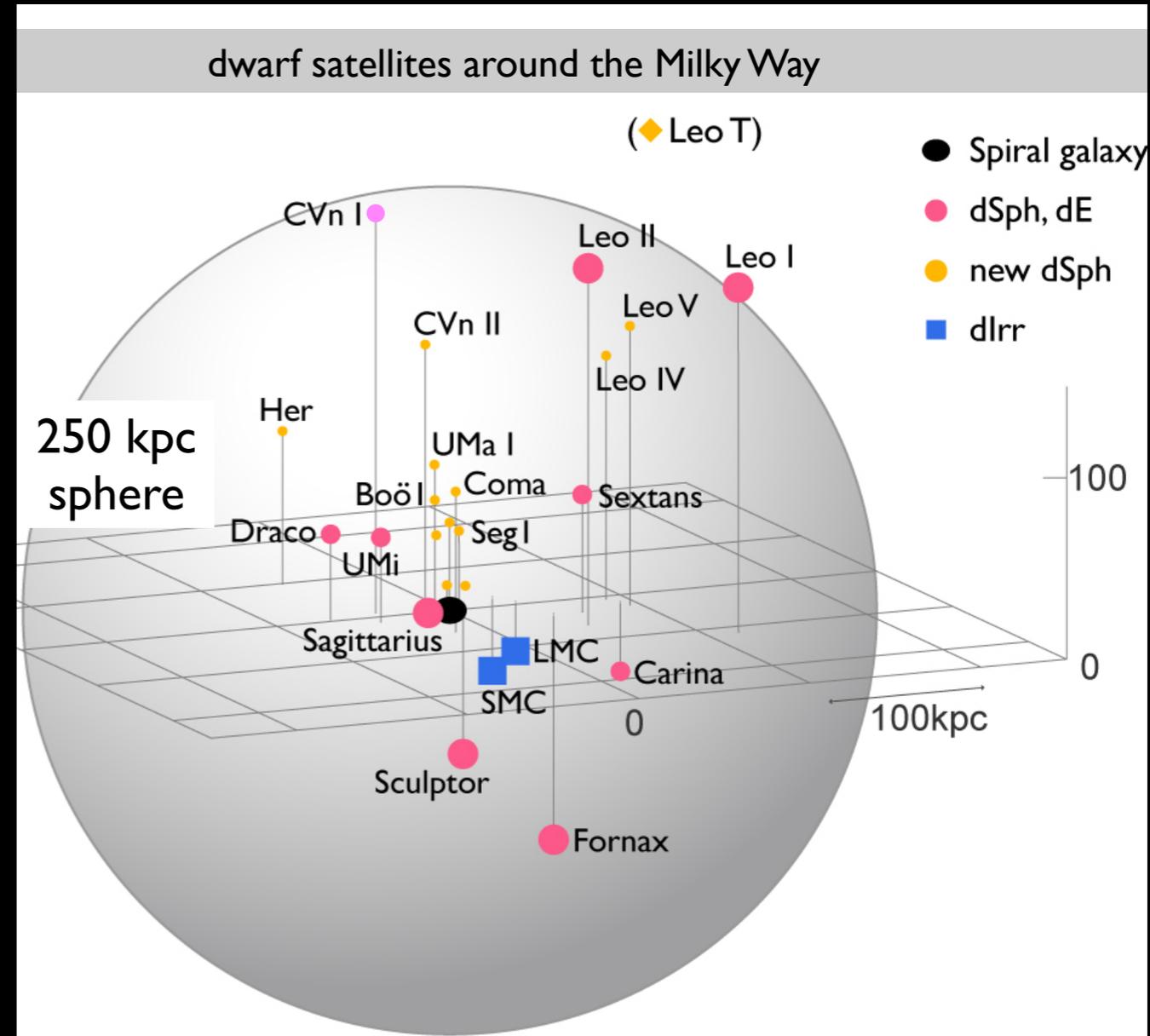
$>10^5$ identified subhalos

12 bright satellites $(L_V > 10^5 L_\odot)$
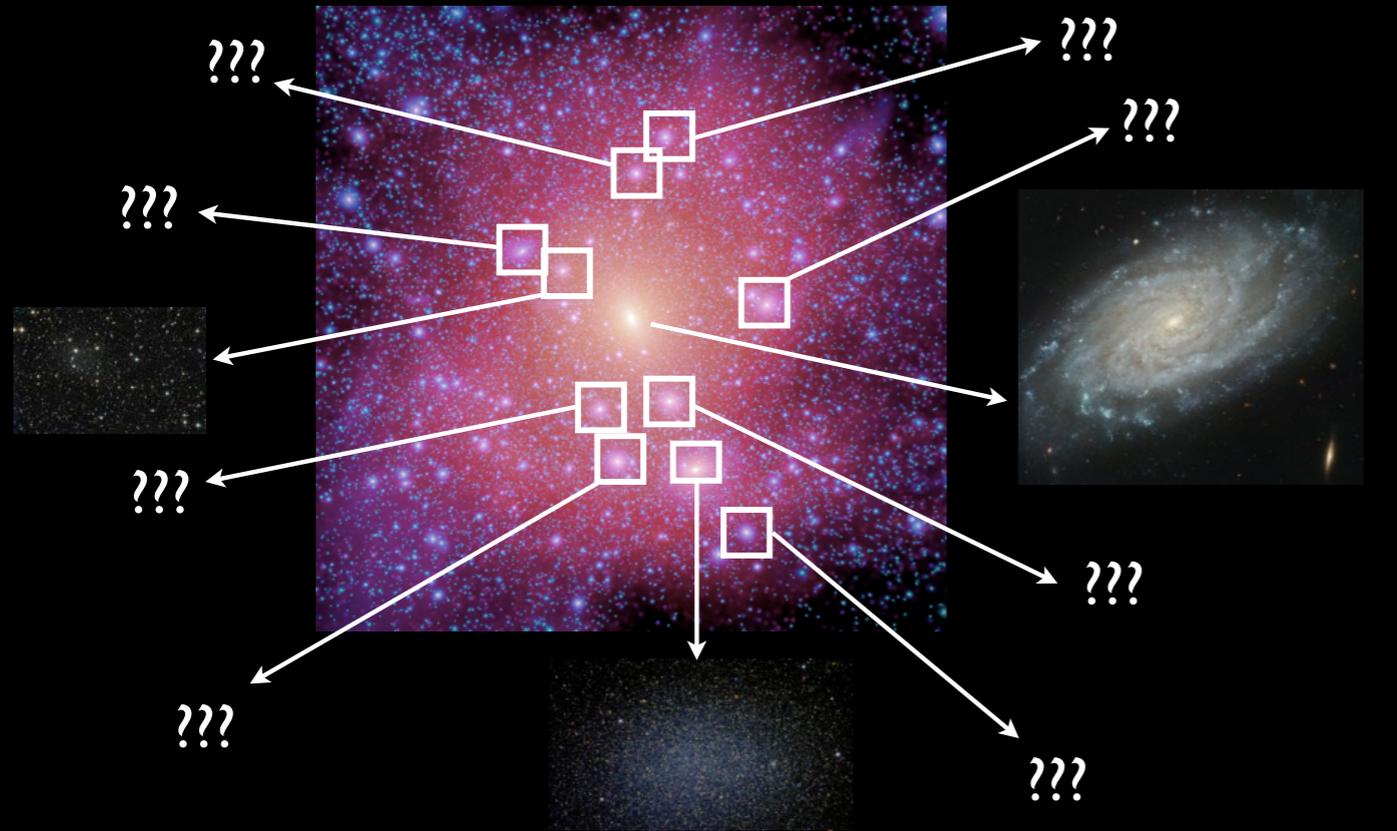
*V. Springel / Virgo Consortium*

*S. Okamoto*

**Of the ~10 biggest subhalos, ~8 cannot host *any* known bright MW satellite**

??? ??? ??? ??? ??? ??? ??? ??? ???

**Observed Milky Way Satellites**

■ LMC

■ SMC

**"massive failures"**: highest resolution LCDM simulations predict ~10 subhalos in this range in the MW, but we don't see *any* such galaxies [except Sagittarius (?)]

**All** of the bright MW dSphs are consistent with $V_{\max} \lesssim 25\,\mathrm{km/s}$ (see also Strigari, Frenk, & White 2010)

No indication that more massive halos host more luminous galaxies

c.f. Strigari et al. 2008

$V_{\max}\ [\mathrm{km/s}]$

$L_V\ [L_\odot]$
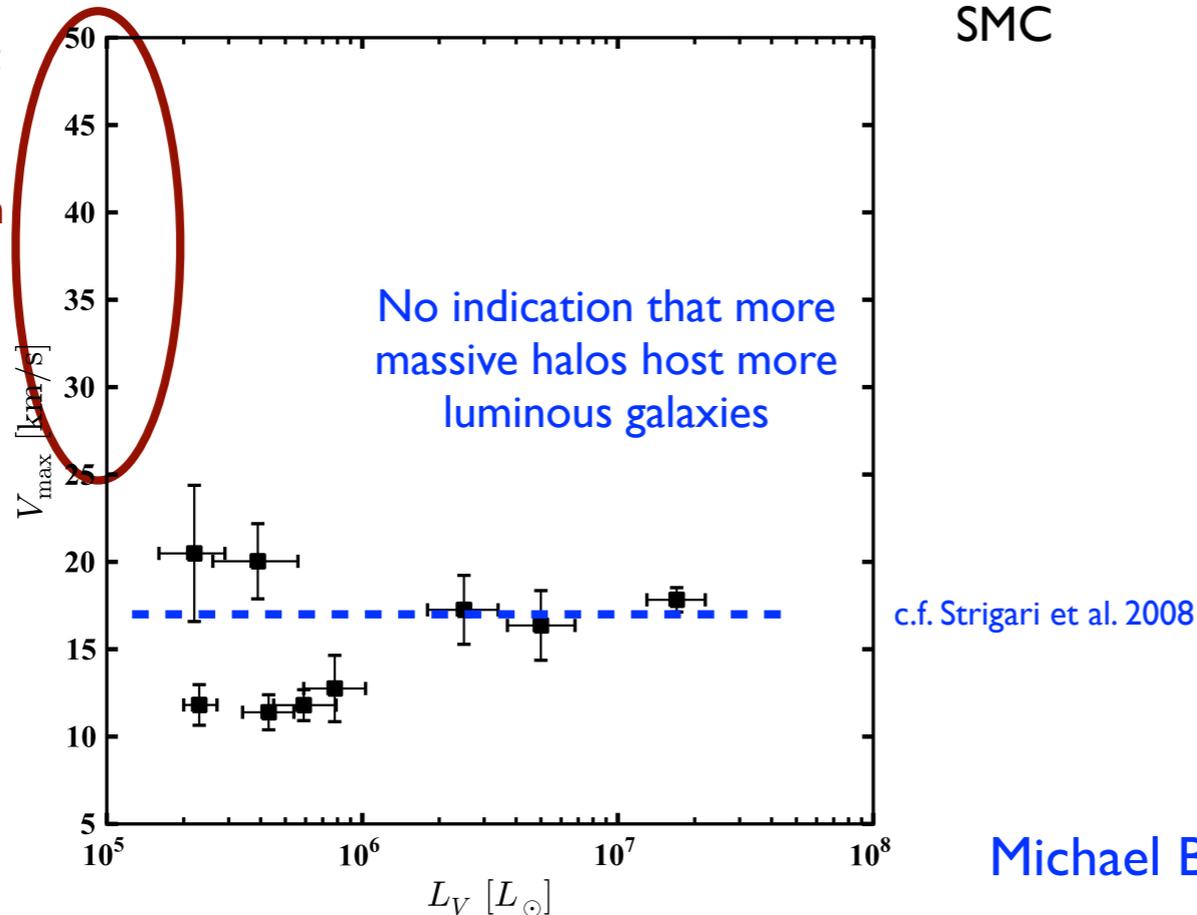
# Possible Solutions to "Too Big to Fail"

The Milky Way is anomalous?

The Milky Way has a low mass dark matter halo?

Galaxy formation is stochastic at low masses?

Dark matter is not just CDM -- maybe WDM (e.g., Lovell+12,13)?

Or even self-interacting DM (Rocha+13, Peter+13, Zavala +14)?

Michael Boylan-Kolchin, Bullock, Kaplinghat 2011, 2012

CDM

WDM
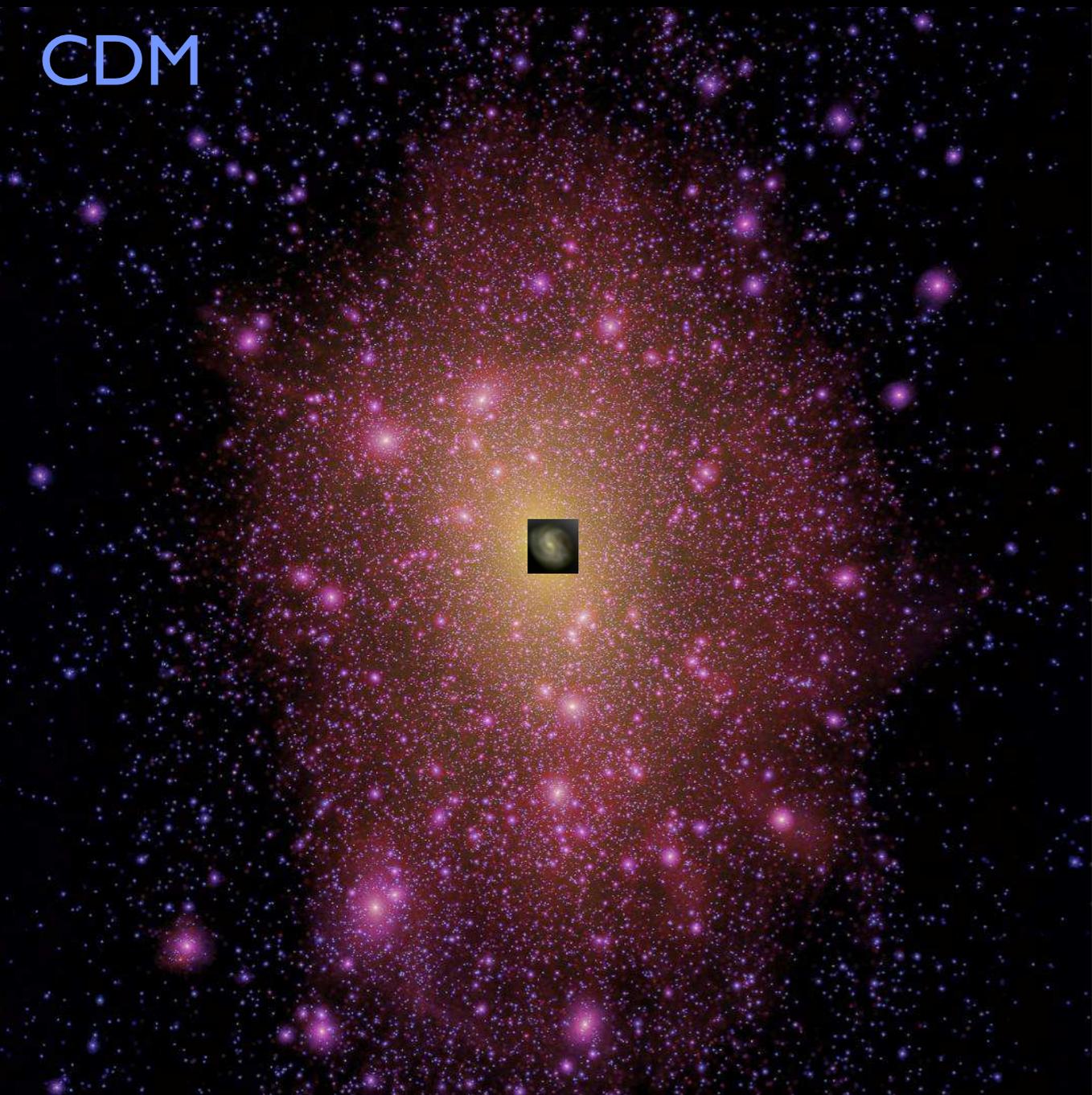
**Aquarius simulation. Springel et al. 2008**      **Lovell, Eke, Frenk, et al. 2012**
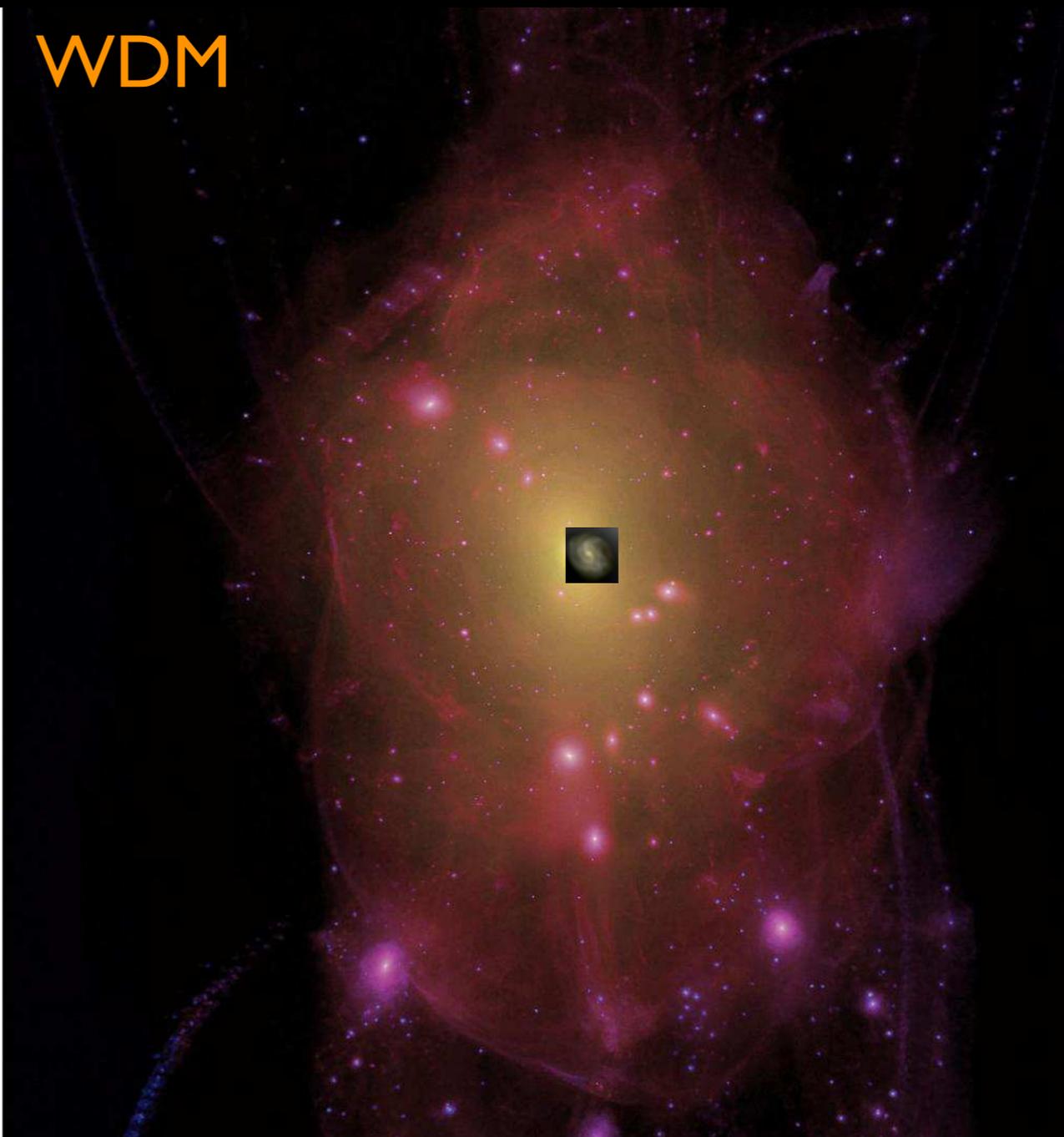
WDM simulation at right has no "too big to fail" subhalos, but it doesn't lead to the right systematics to fit dwarf galaxy properties as Kuzio de Naray+10 showed.  It also won't have the subhalos needed to explain grav lensing flux anomalies and gaps in stellar streams.

CDM

WDM
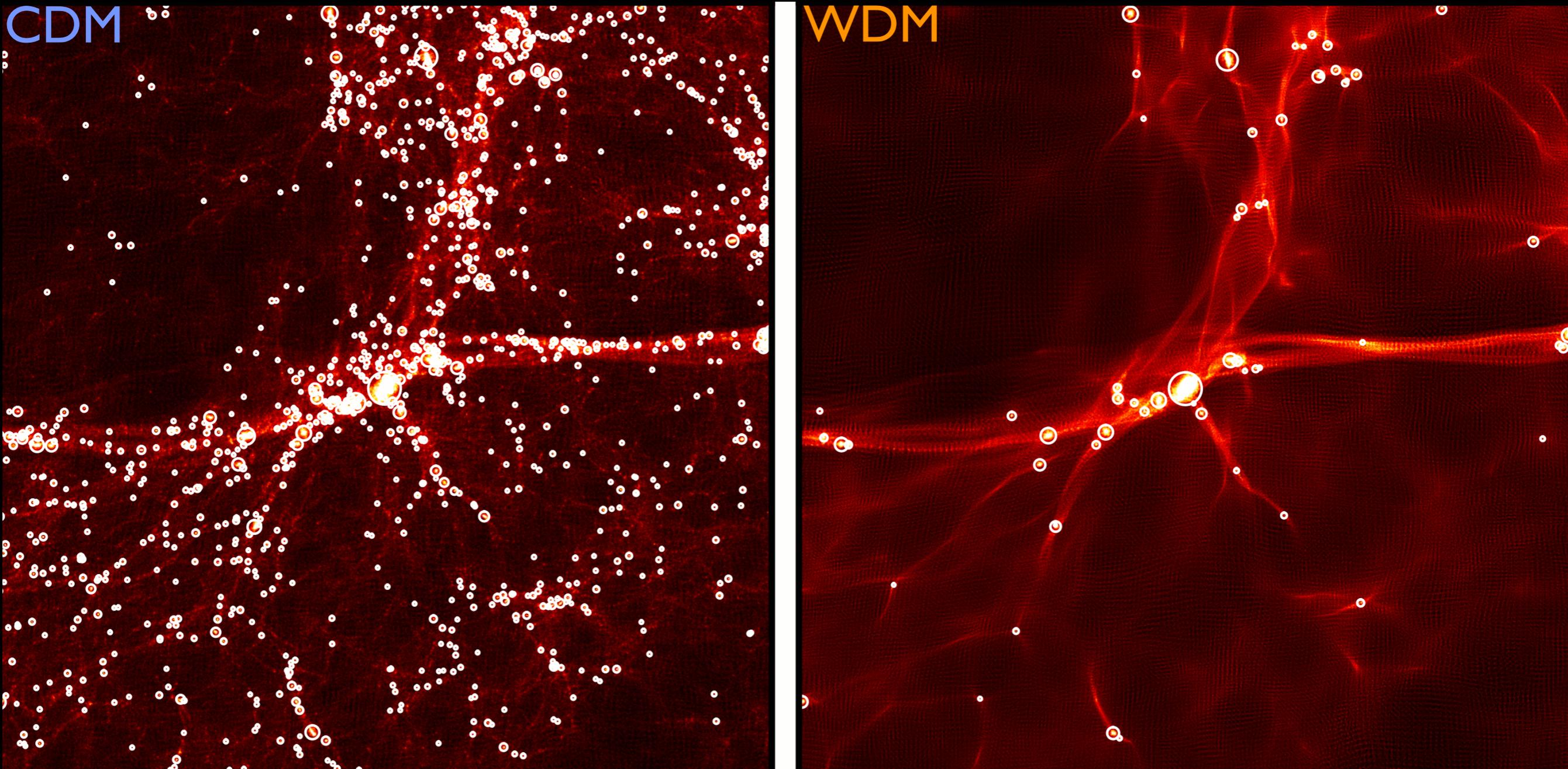
**WDM** simulation at right has no "too big to fail" subhalos, but it is inconsistent at >10σ with Ultra Deep Field galaxy counts. It also won't have the subhalos needed to reionize the universe unless $m_\nu^{thermal} \gtrsim 2.6$ keV (or $m_\nu^{sterile} \gtrsim 15$ keV) assuming an optimistic ionizing radiation escape fraction (Schultz, Onorbe, Abazajian, Bullock14). And the new Ly-α forest analysis (Viel+13) excludes $m_\nu^{thermal} \lesssim 2$ keV at 4σ.

# However there is new evidence for WDM with $m_\nu^{sterile} \approx 7$ keV from detection of 3.5 keV X-rays. Will this be consistent with high-z galaxies, breaks in cold stellar streams and gravitational lensing flux anomalies?

## DETECTION OF AN UNIDENTIFIED EMISSION LINE IN THE STACKED X-RAY SPECTRUM OF GALAXY CLUSTERS arXiv:1402.2301

Esra Bulbul, Maxim Markevitch, Adam Foster, Randall K. Smith, Michael Loewenstein, and Scott W. Randall
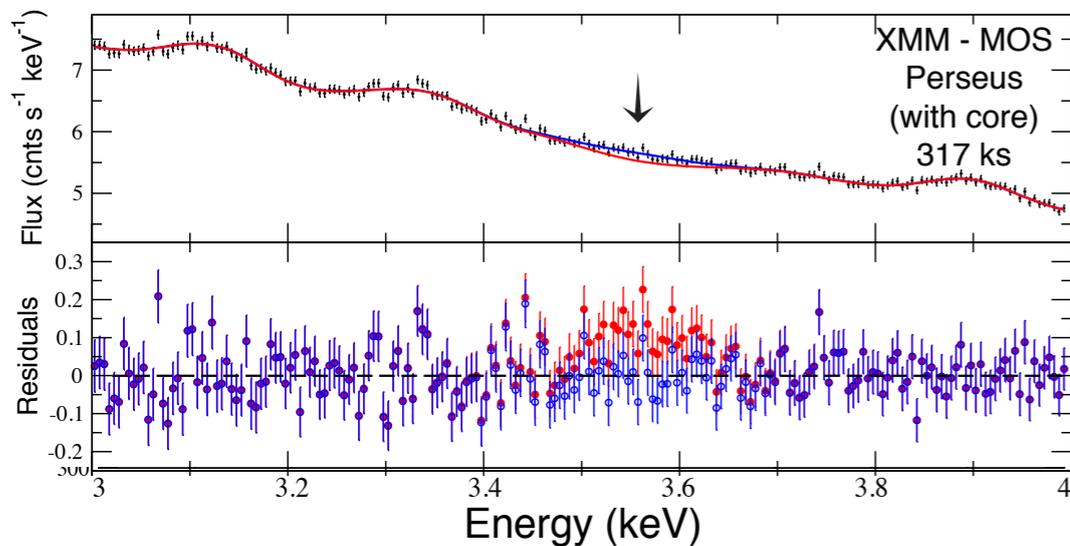
Recent constraints on sterile neutrino dark matter production models (Abazajian+07). Lines in black show theoretical predictions assuming sterile neutrinos are the dark matter with lepton number L = 0, L = 0.003, L = 0.01, L = 0.1. The ★ is consistent with upper limits.
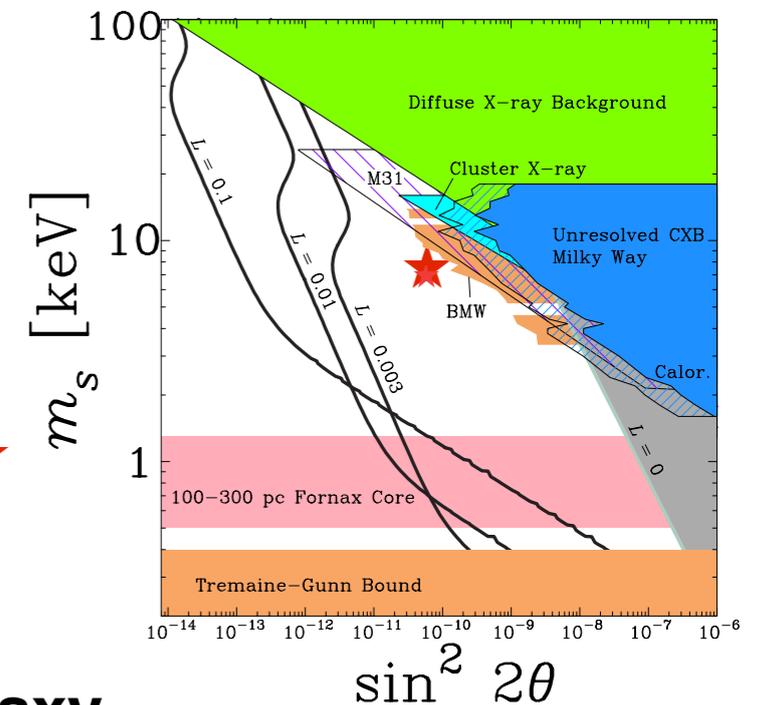


## An unidentified line in X-ray spectra of the Andromeda galaxy and Perseus cluster arXiv:1402.4119 A. Boyarsky, O. Ruchayskiy, D. Iakubovskyi and J. Franse
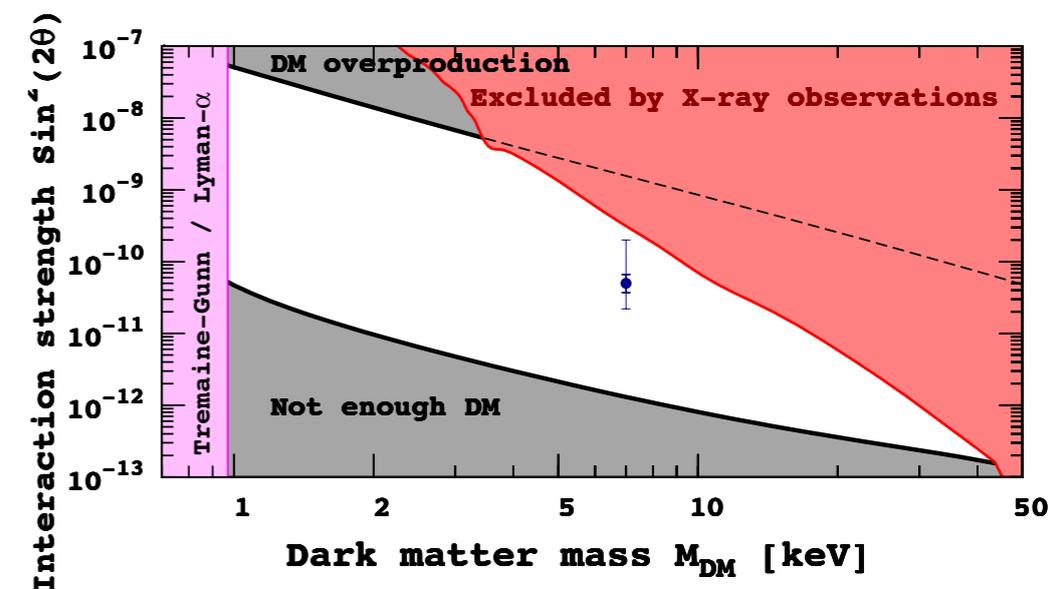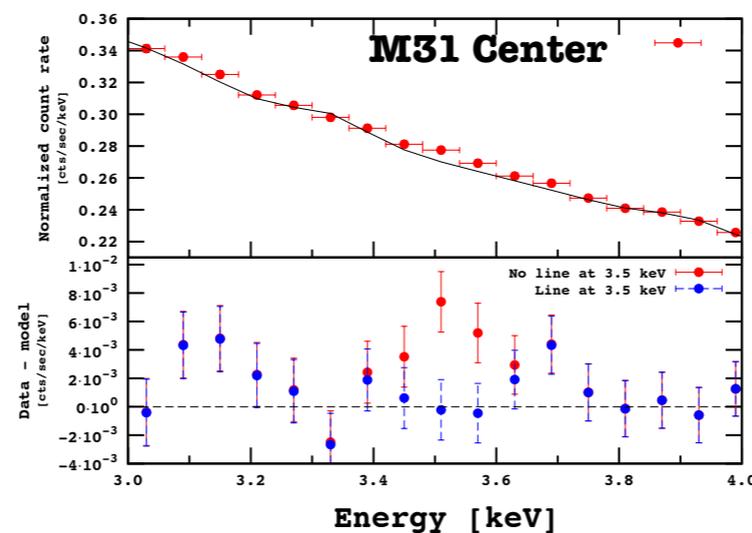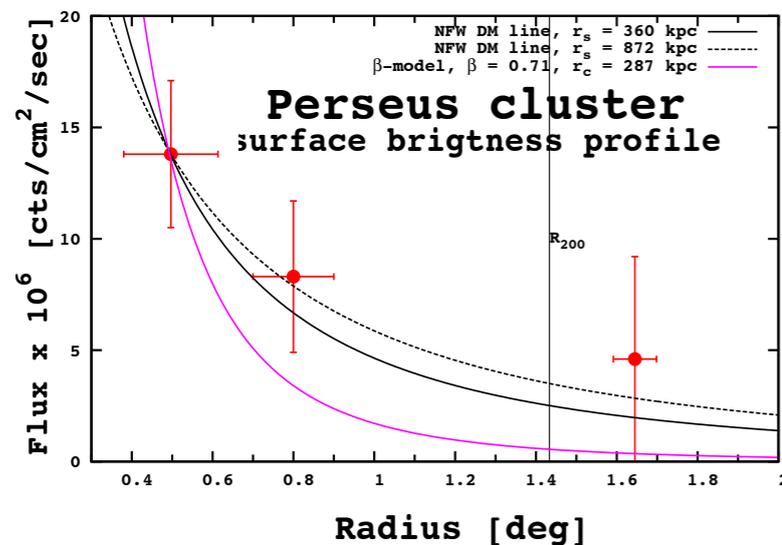
## CLUMPY STREAMS FROM CLUMPY HALOS: DETECTING MISSING SATELLITES WITH COLD STELLAR STRUCTURES

2011 ApJ 731, 58

JOO HEON YOON[1], KATHRYN V. JOHNSTON[1], AND DAVID W. HOGG[2]

Dynamically cold stellar streams are ideal probes of the gravitational field of the Milky Way. This paper re-examines the question of how such streams might be used to test for the presence of "missing satellites" — the many thousands of dark-matter subhalos with masses $10^5 - 10^7 \mathrm{M}_\odot$ which are seen to orbit within Galactic-scale dark-matter halos in simulations of structure formation in $\Lambda$CDM cosmologies. Analytical estimates of the frequency and energy scales of stream encounters indicate that these missing satellites should have a negligible effect on hot debris structures, such as the tails from the Sagittarius dwarf galaxy. However, long cold streams, such as the structure known as GD-1 or those from the globular cluster Palomar 5 (Pal 5) are expected to suffer many tens of direct impacts from missing satellites during their lifetimes. Numerical experiments confirm that these impacts create gaps in the debris' orbital energy distribution, which will evolve into degree- and sub-degree-scale fluctuations in surface density over the age of the debris. Maps of Pal 5's own stream contain surface density fluctuations on these scales. The presence and frequency of these inhomogeneities suggests the existence of a population of missing satellites in numbers predicted in the standard $\Lambda$CDM cosmologies.

See also other recent papers by Carlberg et al.



FIG. 11.— The estimated gap rate vs stream width relation for M31 NW, Pal 5, the EBS and the CDM halo prediction. All data have been normalized to 100 kpc. The width of the theoretical relation is evaluated from the dispersion in the length-height relation of Fig. 8. Predictions for an arbitrary alternative mass functions, $N(M) \propto M^{-1.6}$ normalized to have 33 halos above $10^9 \mathrm{M}_\odot$ is shown with a dotted line.

## DARK MATTER SUB-HALO COUNTS VIA STAR STREAM CROSSINGS

R. G. CARLBERG[1]     2012 ApJ 748, 20

Comparison of the CDM based prediction of the gap rate-width relation with published data for four streams shows generally good agreement within the fairly large measurement errors. **The result is a statistical argument that the vast predicted population of sub-halos is indeed present in the halos of galaxies like M31 and the Milky Way.** The data do tend to be somewhat below the prediction at most points. This could be the result of many factors, such as the total population of sub-halos is expected to vary significantly from galaxy to galaxy, allowing for the stream age would lower the predicted number of gaps for the Orphan stream and possibly others as well, and most importantly these are idealized stream models.

# More evidence for substructure in DM halos: lensing flux anomalies

**Direct Detection of Cold Dark Matter Substructure**

Neal Dalal & Christopher S. Kochanek                                      ApJ 572, 25 (2002)

We devise a method to measure the abundance of satellite halos in gravitational lens galaxies and apply our method to a sample of seven lens systems. After using Monte Carlo simulations to verify the method, we find that substructure comprises $f_{sat}=0.02$ (median, $0.006<f_{sat}<0.07$ at 90% confidence) of the mass of typical lens galaxies, in excellent agreement with predictions of cold dark matter (CDM) simulations.

**Effects of Line-of-Sight Structures on Lensing Flux-ratio Anomalies in a ΛCDM Universe**

D. D. Xu, Shude Mao, Andrew Cooper, Liang Gao, Carlos S. Frenk, Raul Angulo, John Helly        MNRAS (2012)

We conclude that line-of-sight structures can be as important as intrinsic substructures in causing flux-ratio anomalies. ...This alleviates the discrepancy between models and current data, but a larger observational sample is required for a stronger test of the theory.

**Constraints on Small-Scale Structures of Dark Matter from Flux Anomalies in Quasar Gravitational Lenses**

R. Benton Metcalf, Adam Amara                                      MNRAS 419, 3414 (2012)

We investigate the statistics of flux anomalies in gravitationally lensed QSOs as a function of dark matter halo properties such as substructure content and halo ellipticity. ...The constraints that we are able to measure here with current data are roughly consistent with ΛCDM N-body simulations.

**Gravitational detection of a low-mass dark satellite galaxy at cosmological distance,** Simona Vigetti+ 2012 Nature

This group uses galaxy-galaxy lensing to look for the effects of substructure. Our results are consistent with the predictions from cold dark matter simulations at the 95 per cent confidence level, and therefore agree with the view that galaxies formed hierarchically in a Universe composed of cold dark matter.

**Many more small halos than observed small galaxies**
    **1) Field galaxies**
    **2) Satellite galaxies**

**Cusp-Core issue at centers of small galaxies**

**"Too Big to Fail" problem for satellite galaxies**

## Evidence Supporting ΛCDM

**Evidence that the large numbers of small subhalos predicted by ΛCDM actually exist:**
    **1) Gaps in cold stellar streams in the Milky Way**
    **2) Gravitational lensing "flux anomalies"**

**Of the ~10 biggest subhalos, ~8 cannot host *any* known bright MW satellite**
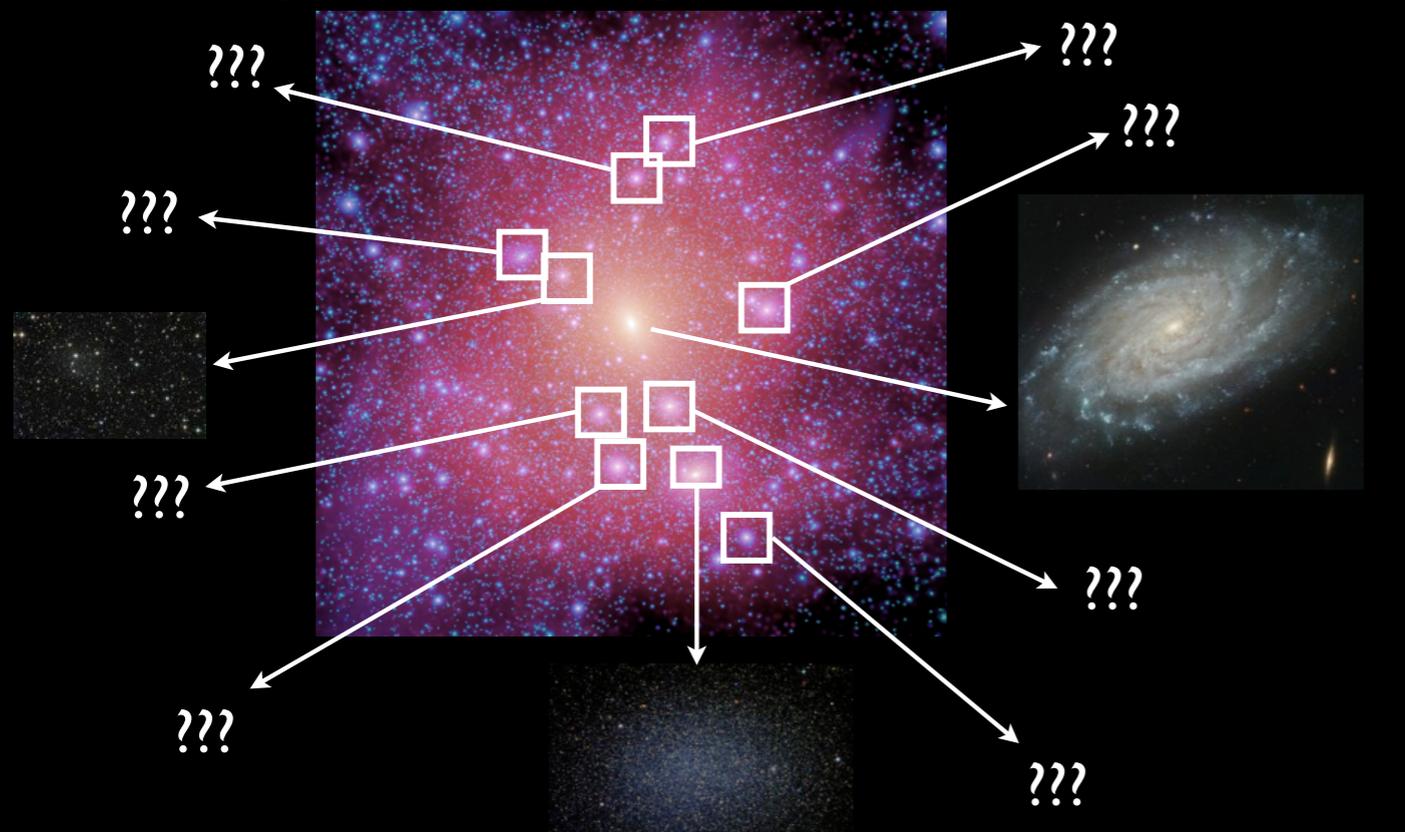


??? ??? ??? ??? ??? ??? ??? ???

## Observed Milky Way Satellites

LMC

SMC

**"massive failures"**: highest resolution LCDM simulations predict ~10 subhalos in this range in the MW, but we don't see *any* such galaxies [except Sagittarius (?)]

**All** of the bright MW dSphs are consistent with $V_{\max} \lesssim 25\,\mathrm{km/s}$ (see also Strigari, Frenk, & White 2010)

No indication that more massive halos host more luminous galaxies

c.f. Strigari et al. 2008



---
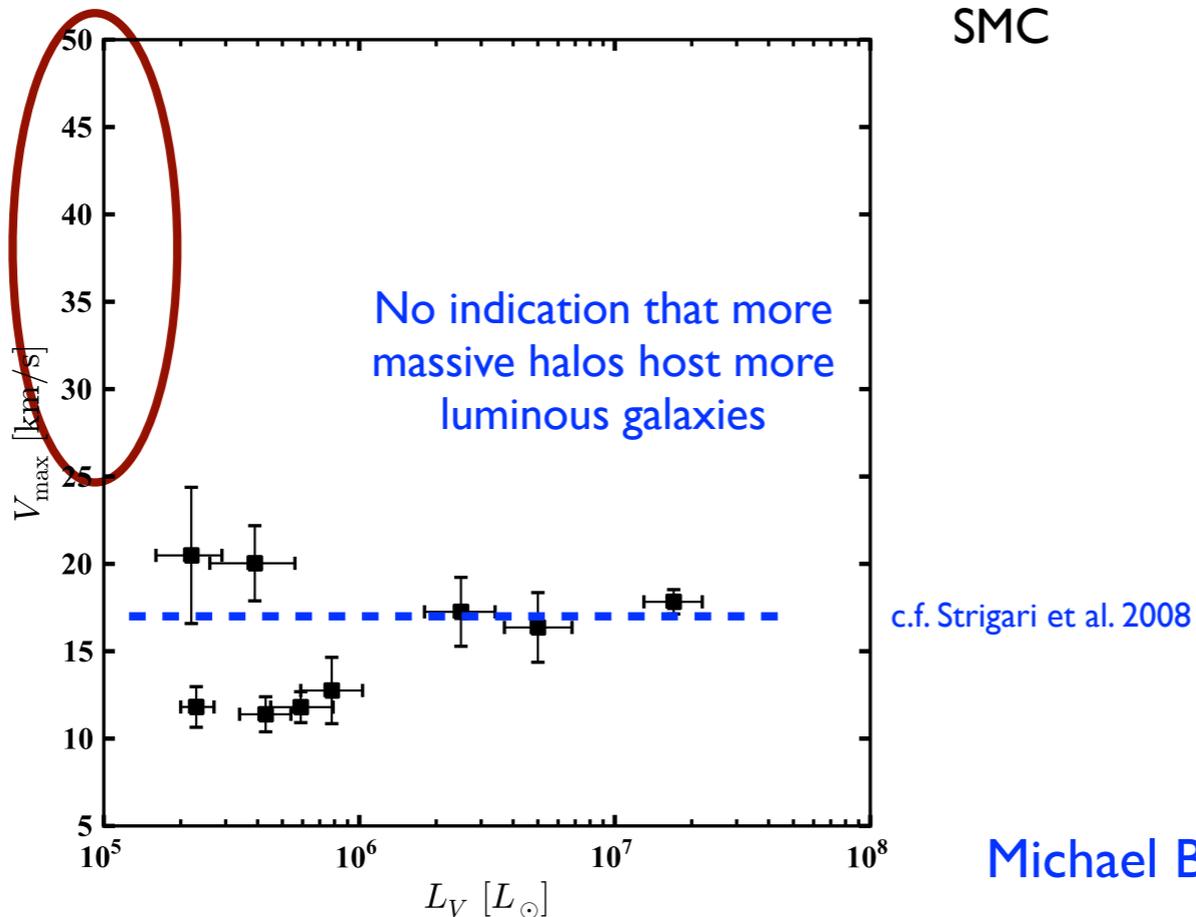
## **Possible Solutions to "Too Big to Fail"**

The Milky Way is anomalous?

The Milky Way has a low mass dark matter halo?

Galaxy formation is stochastic at low masses?

Dark matter is not just CDM -- maybe WDM or even repulsive self-interacting DM?

Or maybe high-resolution CDM-only simulations are being misinterpreted? Stellar feedback can strongly modify the central structure of subhalos, and may resolve the TBTF challenge to ΛCDM.

Michael Boylan-Kolchin, Bullock, Kaplinghat 2011, 2012

# THE BARYON CYCLE OF DWARF GALAXIES: CORED AND NOT TOO-BIG-TO-FAIL

Piero Madau & Sijing Shen

Department of Astronomy and Astrophysics, University of California, Santa Cruz

## ABSTRACT

We present more results from a fully cosmological, "zoom-in", CDM simulation of a group of seven field dwarf galaxies with present-day virial masses in the range $M_{vir}$ = 4. 4 x$10^8$ – 3. 6x$10^{10}$ M$_\odot$. The simulation, run with the TreeSPH code Gasoline at 86 (proper) parsec force resolution, has been previously shown to successfully reproduce the observed stellar mass and cold gas content, resolved star formation histories, and metallicities of field dwarfs in the Local Volume. Here we show that repeated, supernova-driven gas outflows turn dark matter (DM) cusps into kpc-size cores of nearly constant density in all systems having a stellar mass $M_* \gtrsim 10^7$ M$_\odot$, with a "DM removal efficiency" that increases with decreasing host halo mass. DM cores form early, survive during galaxy mergers, and grow secularly over time to kpc scales as the energy input from supernovae exceeds $10^{56}$ ergs. The "scouring" of the core is not as energetically taxing as estimated in some non-cosmological idealized calculations, and the fraction of the energy released by supernovae that is absorbed by the DM is only of order a few percent. The present-day slopes of the inner dark matter mass profiles of the simulated "Bashful" and "Doc" massive dwarfs are similar to those measured in the luminous Fornax and Sculptor dwarf spheroidals. None of the simulated dwarfs has a circular velocity profile exceeding 20 km s$^{-1}$ in the inner 1 kpc, i.e. **supernova feedback can plausibily solve the "too-big-to-fail" problem for Milky Way subhalos.**

## Many Opportunities for Progress Now:
Halo Substructure, Early Galaxies, Galactic Archeology

- **AGORA high-resolution galaxy simulation comparison**
  Will clarify cusp-core and TBTF ΛCDM predictions for
  satellite and dwarf galaxies, and larger galaxies

- **New ways of observing dark matter halo substructure**
  Optical lensing of quasar narrow line regions
  ALMA spectral detection of lensing of dusty galaxies

- **HST Frontier Fields program**
  Uses lensing clusters to get a preview of JWST
  Will clarify nature of high-z galaxies, reionization

- **GAIA will do astrometry on $10^9$ stars in the Milky Way**
  This will allow new probes of dark matter substructure
  Parallel spectroscopic programs on chemical evolution

# ● New ways of observing dark matter halo substructure

## Optical lensing of quasar narrow line regions

**Detection of a substructure with adaptive optics integral field spectroscopy of the gravitational lens B1422+231**

A. M. Nierenberg, T. Treu, S. A. Wright, C. D. Fassnacht, M. W. Auger         arXiv:1402.1496

In this paper we demonstrate for the first time that subhalos can be detected using strongly lensed narrow-line quasar emission, as originally proposed by Moustakas & Metcalf (2003). Many quasars have detectable narrow line emission, so this technique can really measure substructure.

## ALMA spectral detection of lensing of dusty galaxies

**Dark Matter Substructure Detection Using Spatially Resolved Spectroscopy of Lensed Dusty Galaxies**

Yashar Hezaveh, Neal Dalal, Gilbert Holder, Michael Kuhlen, Daniel Marrone, Norman Murray, Joaquin Vieira    ApJ 2013



**Lensed SubMM Galaxies**

apparent brightness
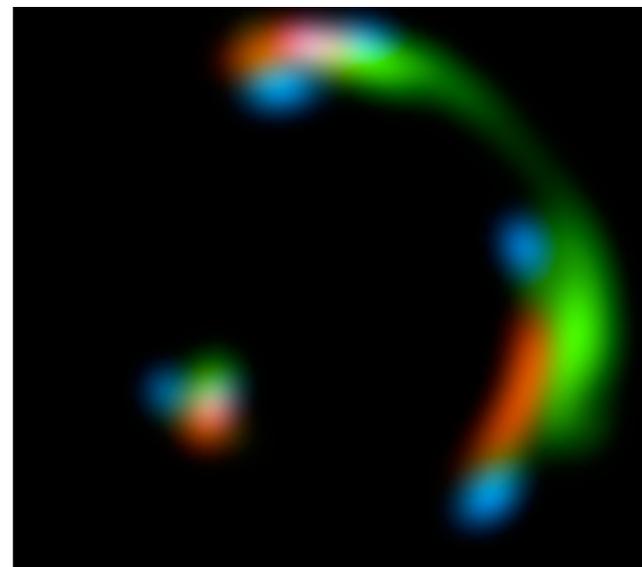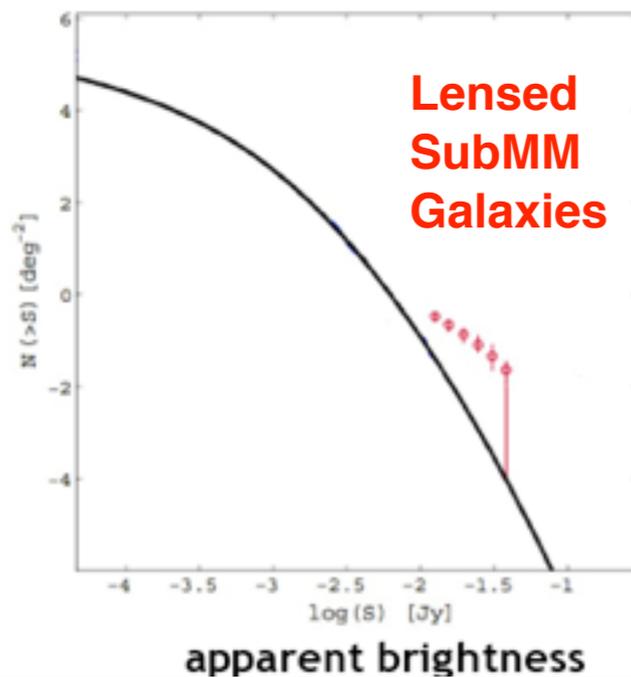
We find that modeling of the full, three-dimensional (angular position and radial velocity) data can significantly facilitate substructure detection, increasing the sensitivity of observables to lower mass subhalos. We find that in typical DSFG lenses, there is a ~55% probability of detecting a substructure with $M > 10^8 M_\odot$ with more than $5\sigma$ detection significance in *each lens*, if the abundance of substructure is consistent with previous lensing results.

- **HST Frontier Fields program**
  Uses lensing clusters to get a preview of JWST
  Will clarify nature of high-z galaxies, reionization
  The program runs 2014-2016, includes Chandra & Spitzer



**Abell 2744 Cluster**
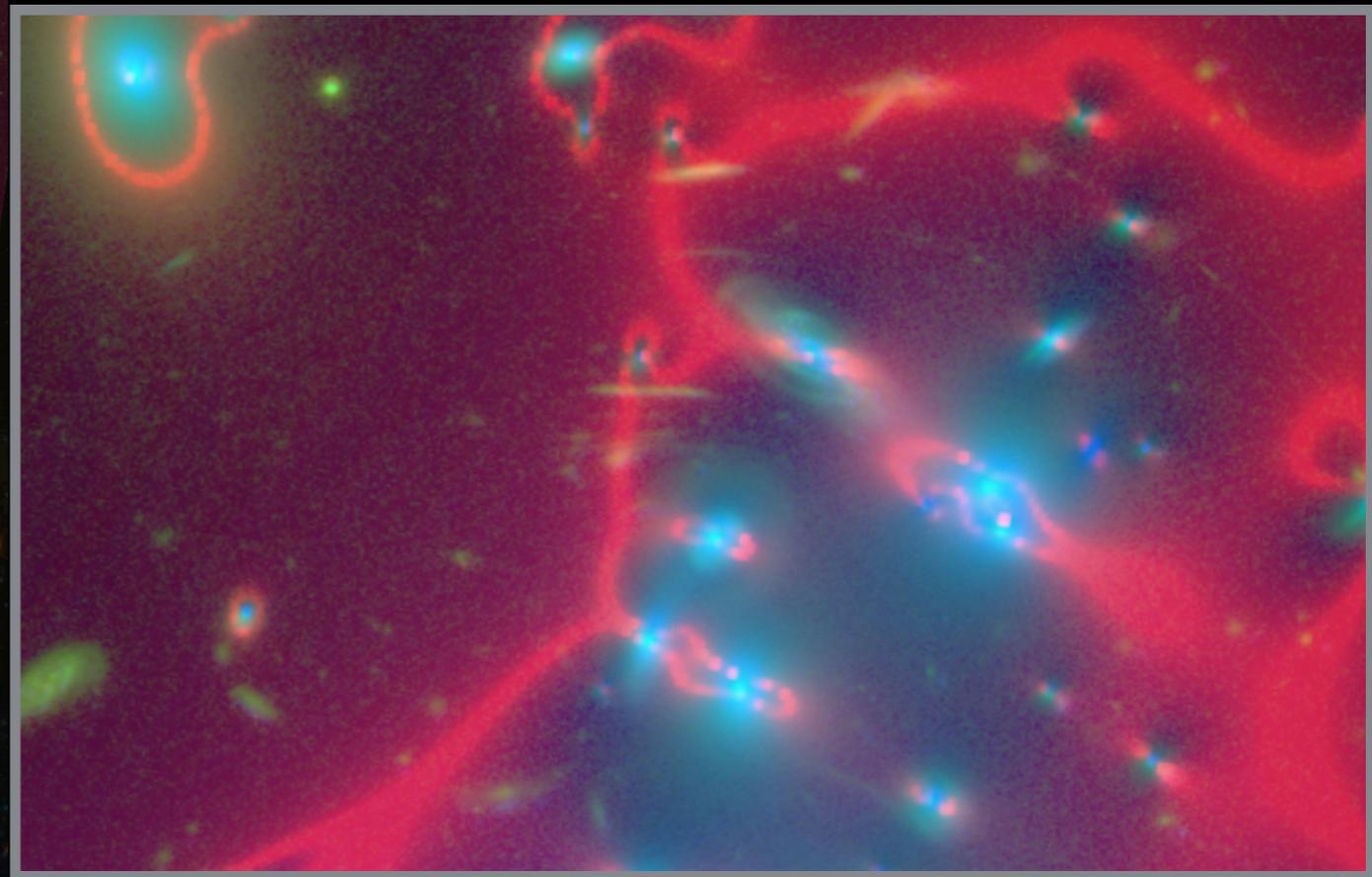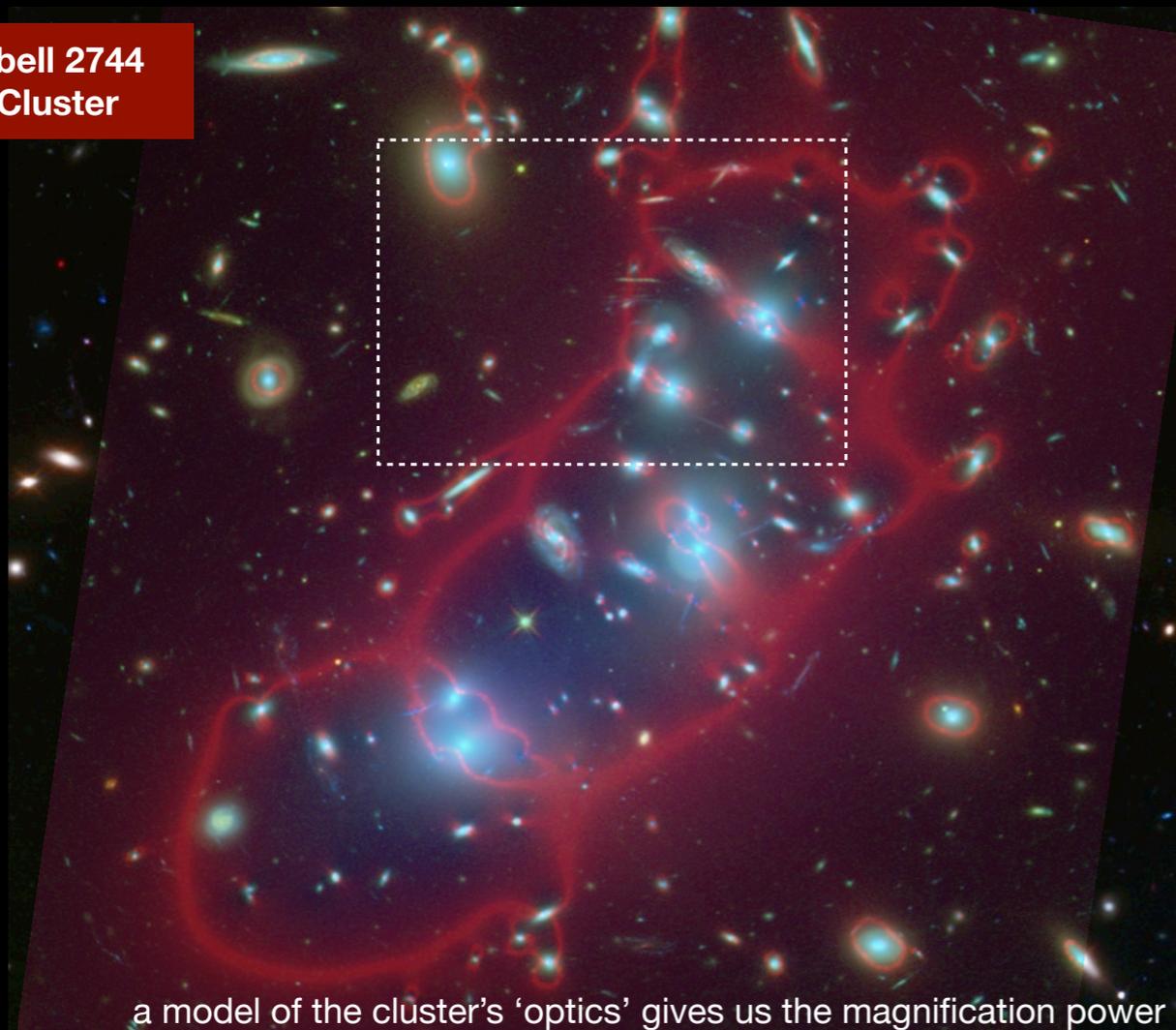
a model of the cluster's 'optics' gives us the magnification power

Spitzer is dedicating >900 hrs of DD time
    deep IRAC imaging at 3.6, 4.5 µm to ~26.5 ABmag depths

archival Chandra data available for all of Frontier Fields;
Chandra FOV encompasses both cluster + parallel fields

- **GAIA will do astrometry on $10^9$ stars in the Milky Way**
  This will allow new probes of dark matter substructure
  Parallel spectroscopic programs on chemical evolution

# Outline

Grand Unification of Forces
Phase Transitions in the Early Universe
Topological Defects: Strings, Monopoles
Cosmic Inflation
  Motivations: Horizon, Flatness, Dragons, Structure
  How much inflation is needed?

# Grand Unification

The basic premise of grand unification is that the known symmetries of the elementary particles result from a larger (and so far unknown) symmetry group G. Whenever a phase transition occurs, part of this symmetry is lost, so the symmetry group changes. This can be represented mathematically as

$$G \rightarrow H \rightarrow ... \rightarrow SU(3) \times SU(2) \times U(1) \rightarrow SU(3) \times U(1).$$

Here, each arrow represents a symmetry breaking phase transition where matter changes form and the groups - G, H, SU(3), etc. - represent the different types of matter, specifically the symmetries that the matter exhibits and they are associated with the different fundamental forces of nature.

The liquid phase of water is rotationally symmetric, that is, it looks the same around each point regardless of the direction in which we look. We could represent this large three-dimensional symmetry by the group G (actually SO(3)). The solid form of frozen water, however, is not uniform in all directions; the ice crystal has preferential lattice directions along which the water molecules align. The group describing these different discrete directions H, say, will be smaller than G. Through the process of freezing, therefore, the original symmetry G is broken down to H.

# Grand Unified Theory

**GUT** refers to a theory in physics that unifies the strong interaction and electroweak interaction. Several such theories have been proposed, but none is currently universally accepted. The (future) theory that will also include gravity is termed theory of everything. Some common GUT models are:

- Georgi-Glashow (1974) model -- SU(5)
- SO(10)
- Flipped SU(5) -- SU(5)×U(1)
- Pati-Salam model -- SU(4)×SU(2)×SU(2)
- $E_6$

There is still no hard evidence that nature is described by a GUT theory. But since grand unification is realized with supersymmetry, i.e. the three forces do come together at about $10^{16}$ GeV, the GUT hypothesis is theoretically attractive.

However, GUT models generically predict the existence of topological defects such as monopoles, cosmic strings, domain walls, and others. None have been observed and explaining their absence is known as the monopole problem in cosmology. Solving it was Alan Guth's motivation for inventing Cosmic Inflation.

# Topological Defects

These arise when some n-component scalar field $\phi_i(\mathbf{x}) = 0$ because of topological trapping that occurs as a result of a phase transition in the early universe (as I will explain shortly).

If the $\phi$ field is complex then n=2, and $\phi_i(\mathbf{x}) = 0$ occurs along a linear locus of points, a **string,** in three dimensional space.  This corresponds to a 2-dimensional world-sheet in the 3+1 dimensions of spacetime.

If the $\phi$ field has three components, then $\phi_i(\mathbf{x}) = 0$ occurs at a point in three dimensional space, a **monopole**.  This corresponds to a 1-dimensional world-line in the 3+1 dimensions of spacetime.

If the $\phi$ field has four components, then $\phi_i(\mathbf{x}) = 0$ occurs at a point in space-time, an **instanton**.  A related concept is **texture**.

Topological defects were once thought to be a possible origin of the fluctuations that lead to galaxy formation.

# Phase transitions

The cosmological significance of symmetry breaking is due to the fact that symmetries are restored at high temperature (just as it is for liquid water when ice melts). For extremely high temperatures in the early universe, we might even achieve a grand unified state G. Viewed from the moment of creation forward, the universe will pass through a sucession of phase transitions at which the strong nuclear force will become differentiated and then the weak nuclear force and electromagnetism.

Phase transitions can have a wide variety of important implications including the formation of topological defects - cosmic strings, domain walls, monopoles and textures, or it may even trigger a period of exponential expansion (inflation).

Phase transitions can be either dramatic - first order, or smooth - second order.

During a first-order phase transition, the matter fields get trapped in a `false vacuum' state from which they can only escape by nucleating bubbles of the new phase, that is, the `true vacuum' state.



**False vacuum**

**True vacuum**

$V(\phi)$

bubble nucleation

$\phi$
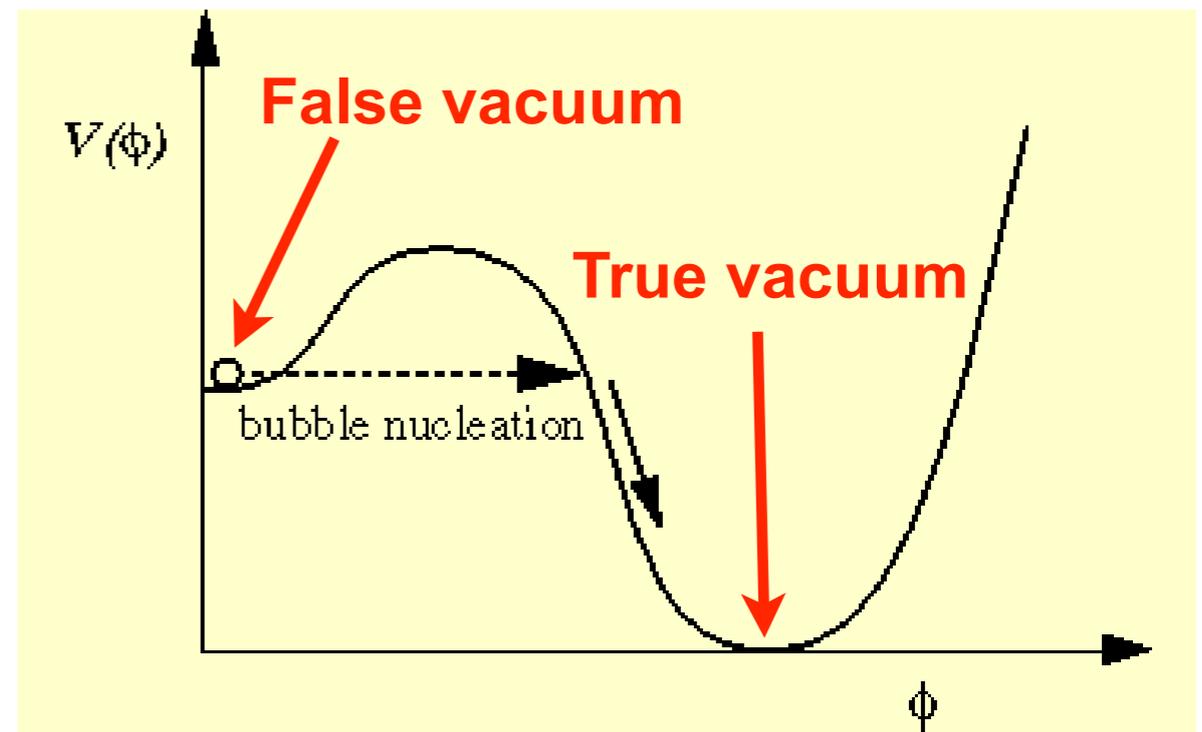
First-order phase transitions (illustrated below) occur through the formation of bubbles of the new phase in the middle of the old phase; these bubbles then expand and collide until the old phase disappears completely and the phase transition is complete.
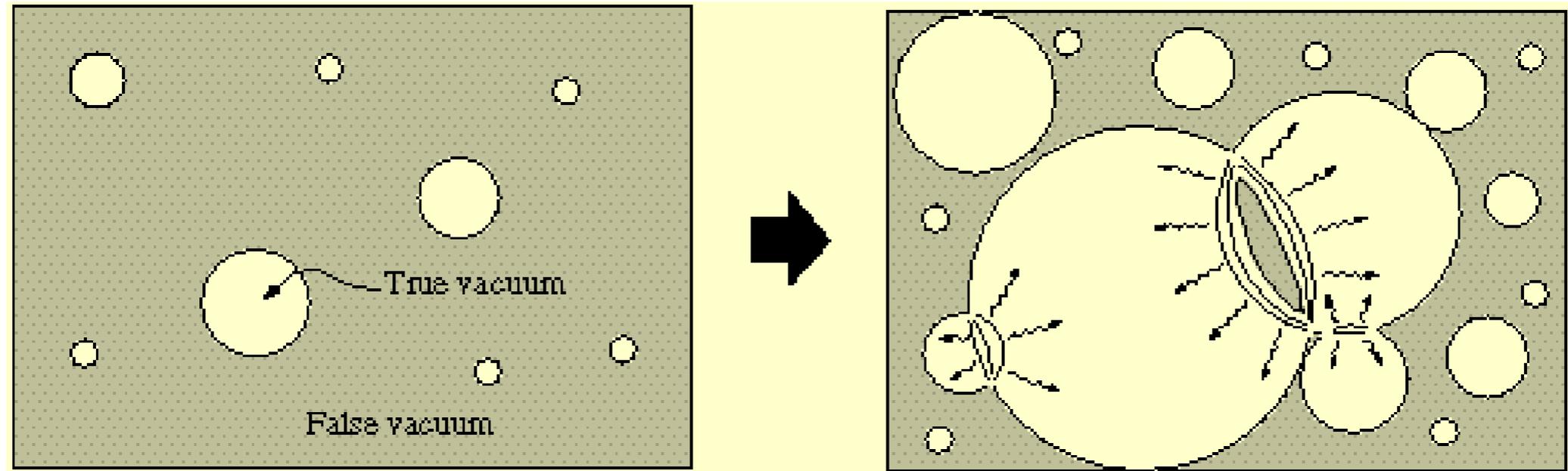


First-order phase transitions proceed by bubble nucleation. A bubble of the new phase (the true vacuum) forms and then expands until the old phase (the false vacuum) disappears. A useful analogue is boiling water in which bubbles of steam form and expand as they rise to the surface.

Second-order phase transitions, on the other hand, proceed smoothly. The old phase transforms itself into the new phase in a continuous manner.  There is energy (specific heat of vaporization, for example) associated with a first order phase transition.

Either type of phase transition can produce stable configurations called "topological defects."

# Cosmic Strings & Other Topological Defects

Topological defects are stable configurations that are in the original, symmetric or old phase, but nevertheless for topological reasons they persist after a phase transition to the asymmetric or new phase is completed - because to unwind them would require a great deal of energy. There are a number of possible types of defects, such as domain walls, cosmic strings, monopoles, and textures. The type of defect is determined by the symmetry properties of the matter and the nature of the phase transition.

**Domain walls:**  These are two-dimensional objects that form when a discrete symmetry is broken at a phase transition. A network of domain walls effectively partitions the universe into various `cells'. Domain walls have some rather peculiar properties. For example, the gravitational field of a domain wall is repulsive rather than attractive.
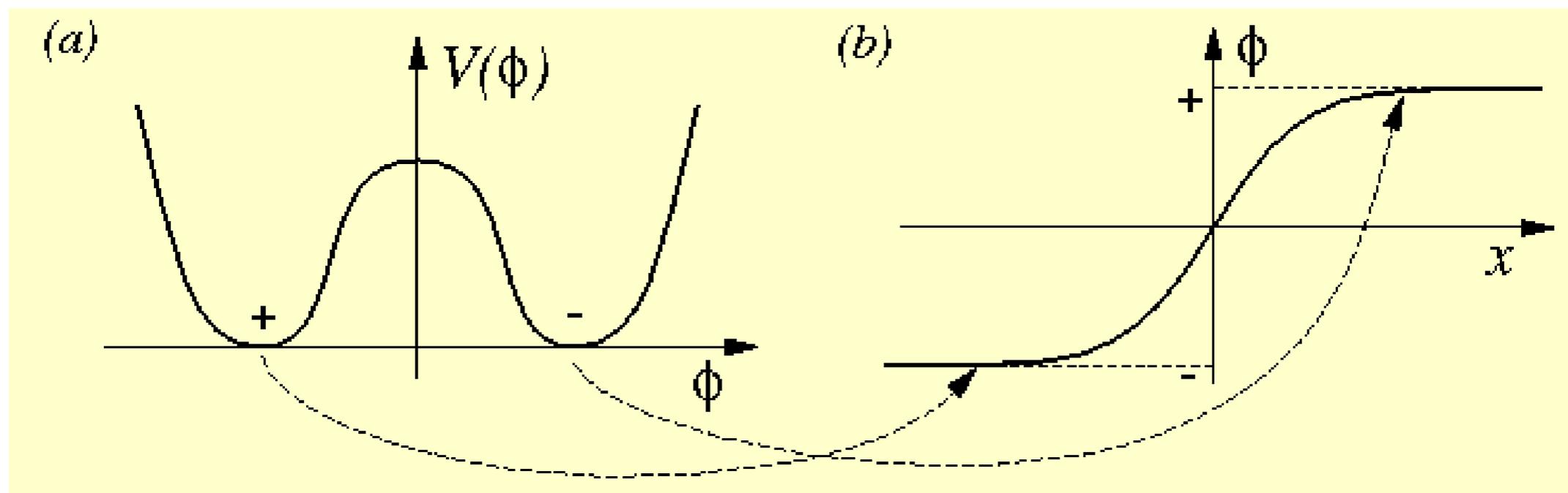
**Cosmic strings:** These are one-dimensional (that is, line-like) objects which form when an axial or cylindrical symmetry is broken. Strings can be associated with grand unified particle physics models, or they can form at the electroweak scale. They are very thin and may stretch across the visible universe. A typical GUT string has a thickness that is less then a trillion times smaller that the radius of a hydrogen atom, but a 10 km length of one such string would weigh as much as the earth itself!
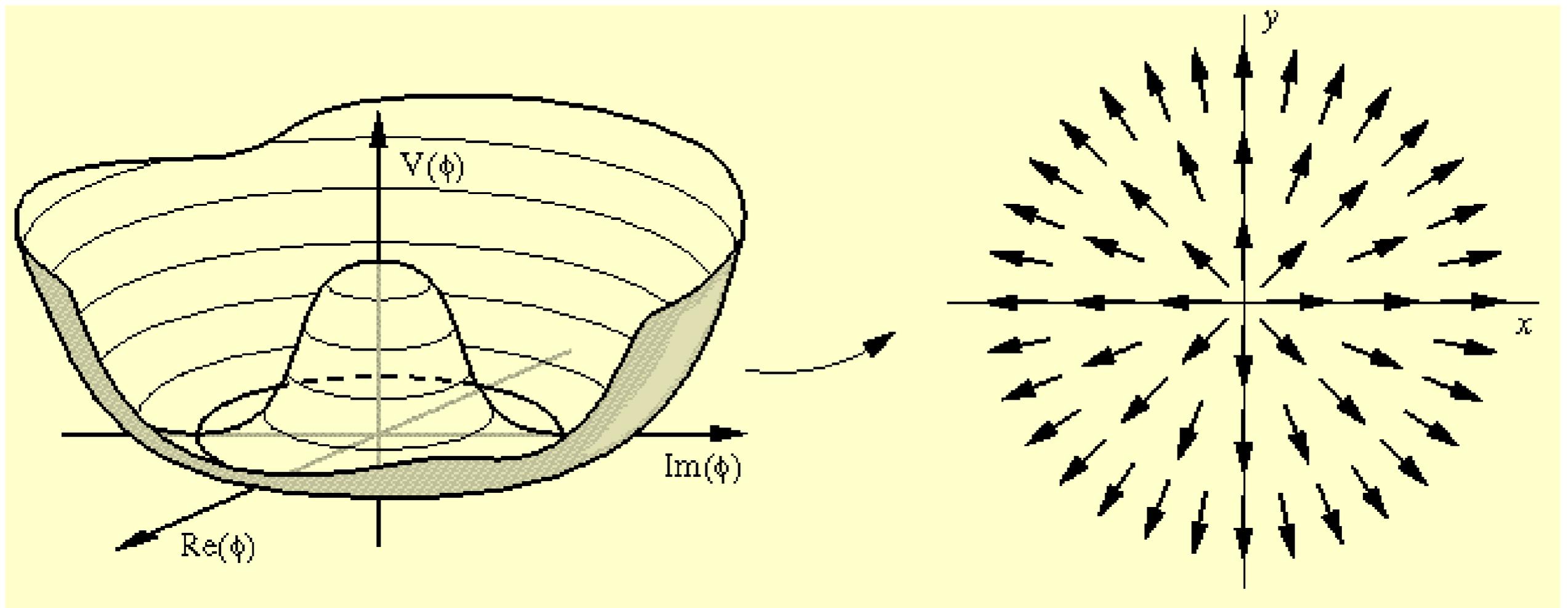


Cosmic strings are associated with models in which the set of minima are not simply-connected, that is, the vacuum manifold has `holes' in it. The minimum energy states on the left form a circle and the string corresponds to a non-trivial winding around this.

**Monopoles:** These are zero-dimensional (point-like) objects which form when a spherical symmetry is broken. Monopoles are predicted to be supermassive and carry magnetic charge. The existence of monopoles is an inevitable prediction of grand unified theories (GUTs - more on this shortly); why the universe isn't filled with them is one of the puzzles of the standard cosmology.



Monopole
(a)

No Monopole
(b)
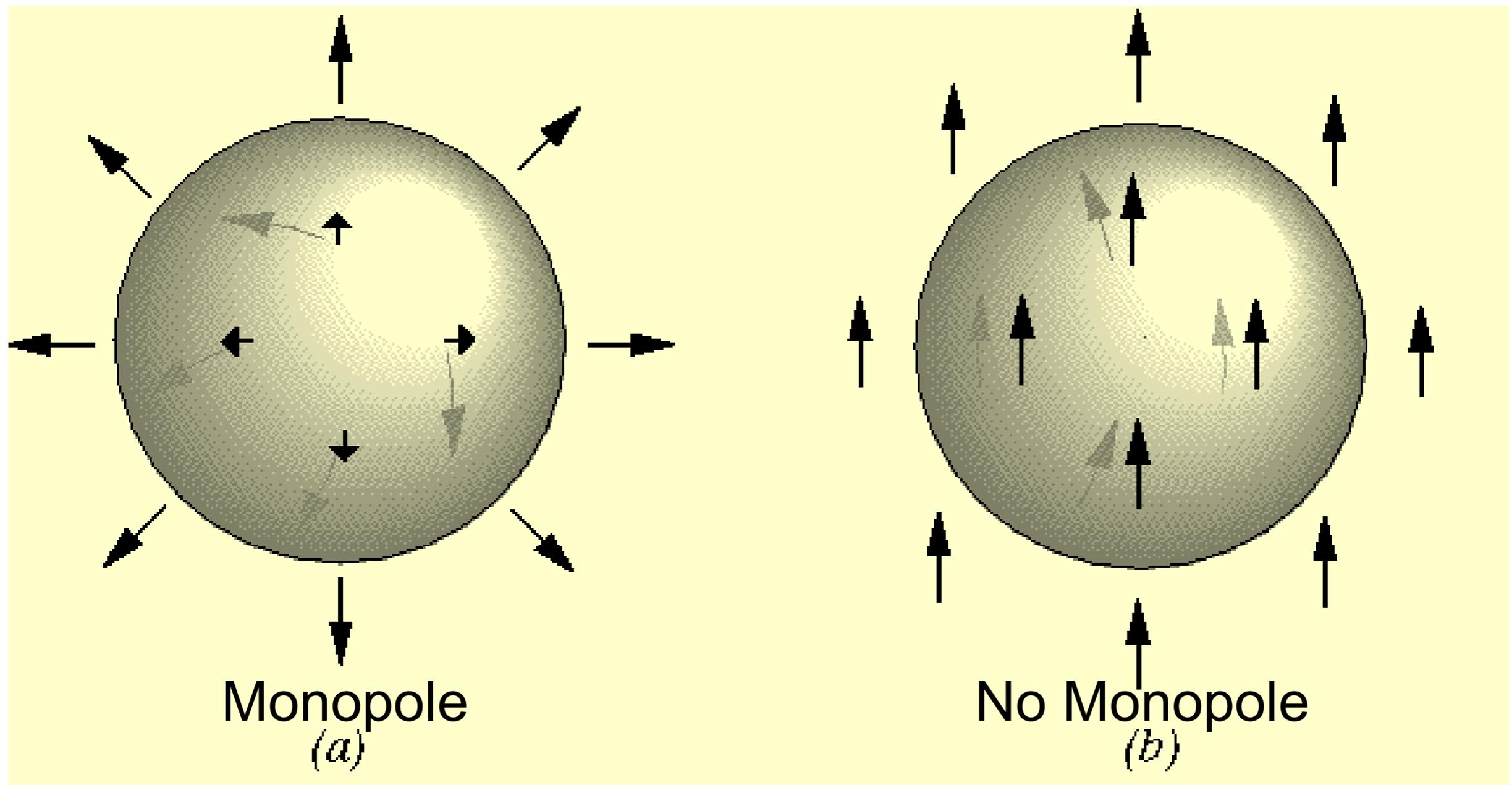
**Textures:** These form when larger, more complicated symmetry groups are completely broken. Textures are delocalized topological defects which are unstable to collapse. A speculation that the largest "cold spot" in the WMAP CMB data was caused by cosmic textures was published by Cruz et al. (2007, Science 318, 1612), but by 2010 this was shown to be unlikely.
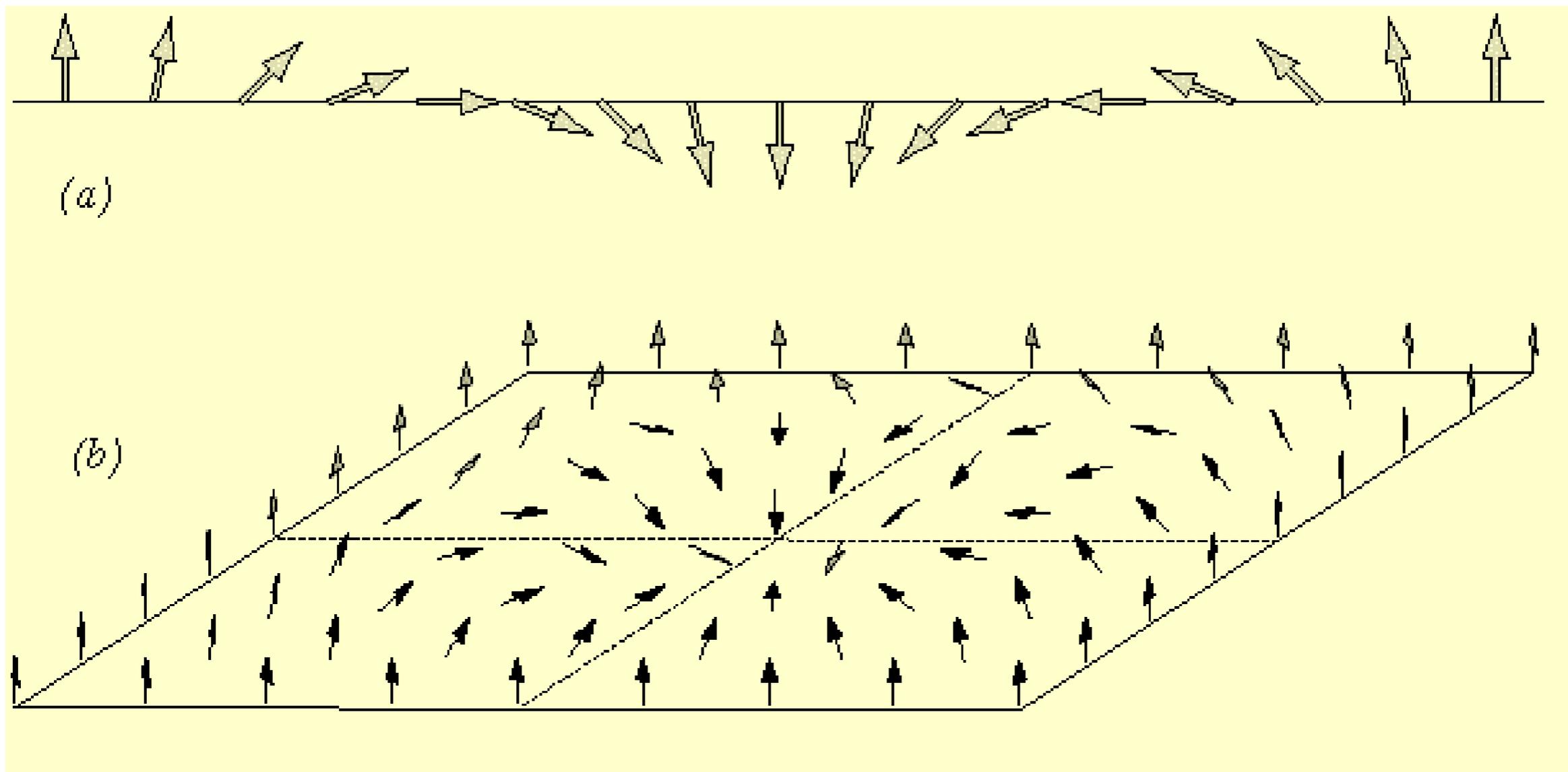


Examples of delocalized texture configurations in one and two dimensions.

# A Cosmic Microwave Background Feature Consistent with a Cosmic Texture

M. Cruz,[1,2]* N. Turok,[3] P. Vielva,[1] E. Martínez-González,[1] M. Hobson[4]

The Cosmic Microwave Background provides our most ancient image of the universe and our best tool for studying its early evolution. Theories of high-energy physics predict the formation of various types of topological defects in the very early universe, including cosmic texture, which would generate hot and cold spots in the Cosmic Microwave Background. We show through a Bayesian statistical analysis that the most prominent 5°-radius cold spot observed in all-sky images, which is otherwise hard to explain, is compatible with having being caused by a texture. From this model, we constrain the fundamental symmetry-breaking energy scale to be $\phi_0 \approx 8.7 \times 10^{15}$ gigaelectron volts. If confirmed, this detection of a cosmic defect will probe physics at energies exceeding any conceivable terrestrial experiment.

## The Axis of Evil revisited

Kate Land, Joao Magueijo, 2007 MNRAS, 378, 153

Abstract: In light of the three-year data release from WMAP we re-examine the evidence for the ``Axis of Evil'' (AOE) [anomalous alignment of CMB multipoles in the direction $l \approx -100$, $b = 60$]. We discover that previous statistics are not robust with respect to the data-sets available and different treatments of the galactic plane. We identify the cause of the instability and implement an alternative ``model selection'' approach. A comparison to Gaussian isotropic simulations find the features significant at the 94-98% level, depending on the particular AOE model. The Bayesian evidence finds lower significance, ranging from ``substantial'' to no evidence for the most general AOE model.
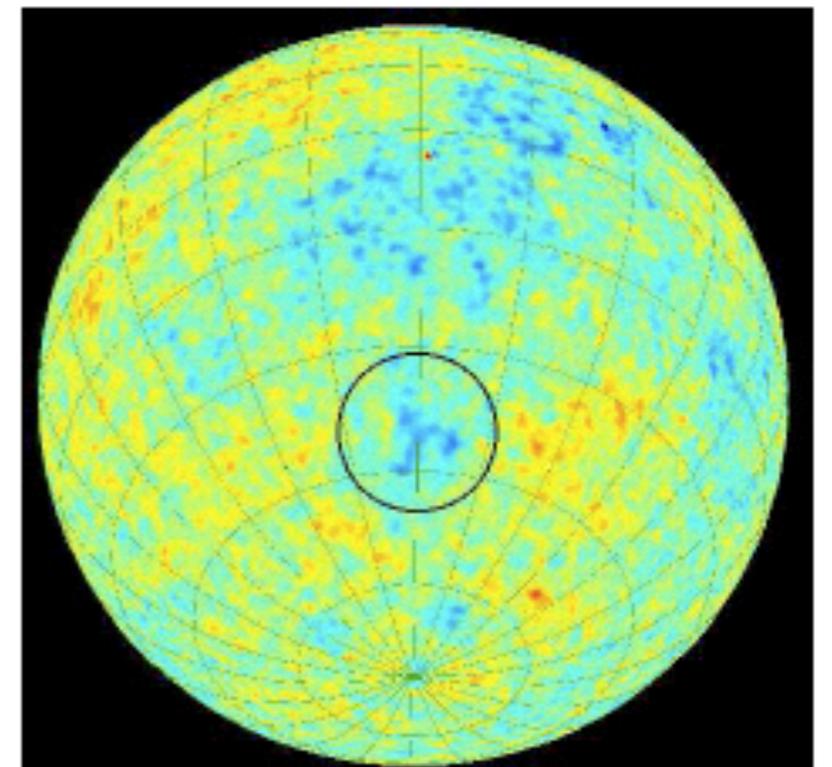
The zone of the CS has been placed at the center of the black circle.
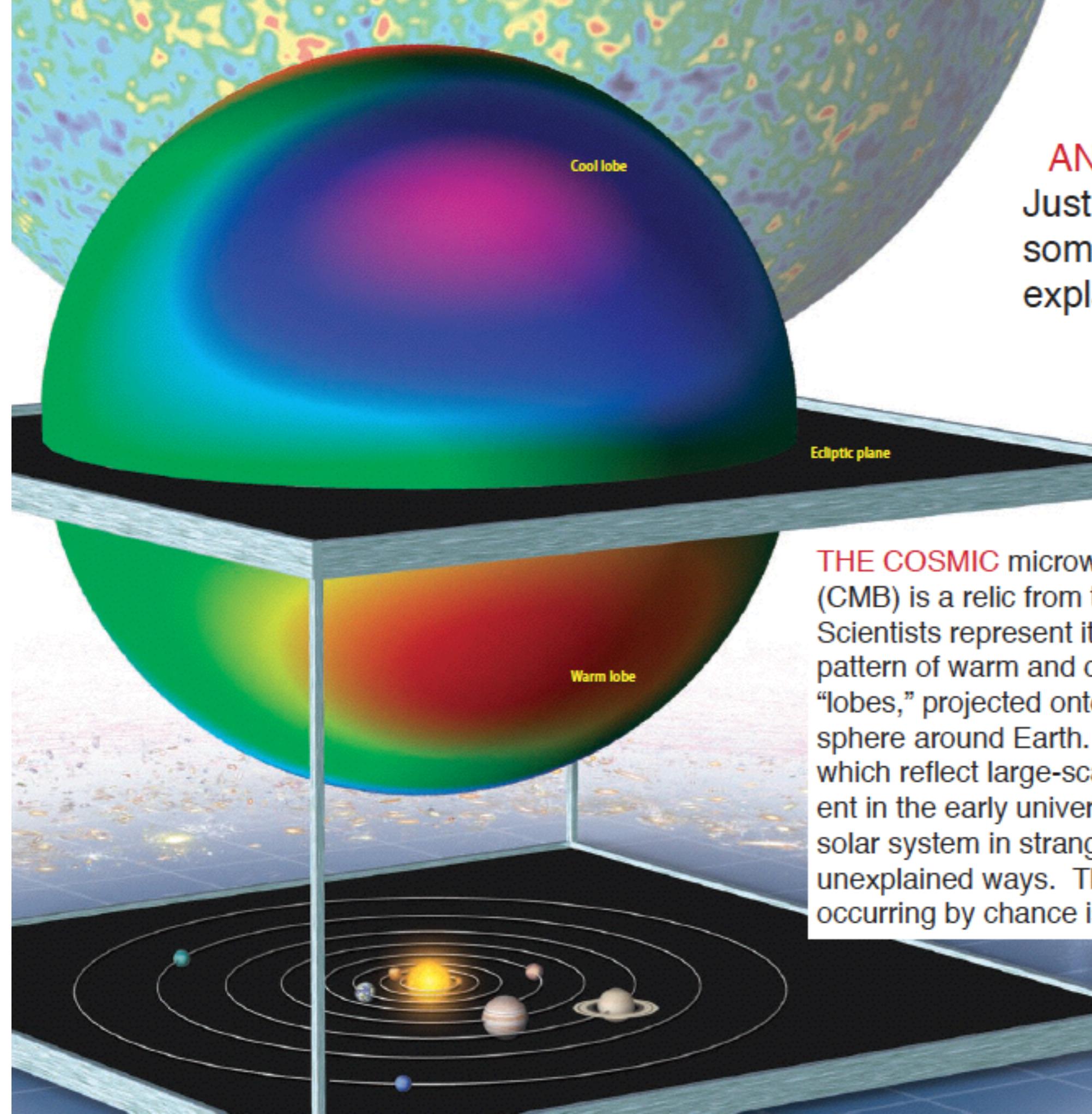
ANOMALIES
Just chance, or something to be explained?

Cool lobe

Ecliptic plane

Warm lobe

THE COSMIC microwave background (CMB) is a relic from the early universe. Scientists represent it as a complex pattern of warm and cool spots, or "lobes," projected onto the celestial sphere around Earth. The patterns, which reflect large-scale structures present in the early universe, line up with the solar system in strange and as-yet-unexplained ways. The likelihood of this occurring by chance is less than 0.1%.

D. Huterer
*Astronomy*
Dec 2007

NASA's Wilkinson Microwave Anisotropy Probe team, who have just released their most detailed map yet of the CMB, used Hawking's initials to draw attention to a serious point. With each new round of WMAP data – the latest is based on seven years of data – apparent anomalies called "anisotropies" in the CMB have puzzled physicists. Such patterns have also been used to justify various exotic theories.

One notorious anomaly is the "axis of evil", an apparent alignment in the hot and cold regions where there should be randomness. Another is the "cold spot", a particularly large void in the CMB, which some have proposed is evidence of another universe nestling next to our own.

The WMAP team point out that if something as apparently unlikely as Hawking's initials can be found in the CMB data, then the chances of finding other apparently improbable patterns may also be quite high.
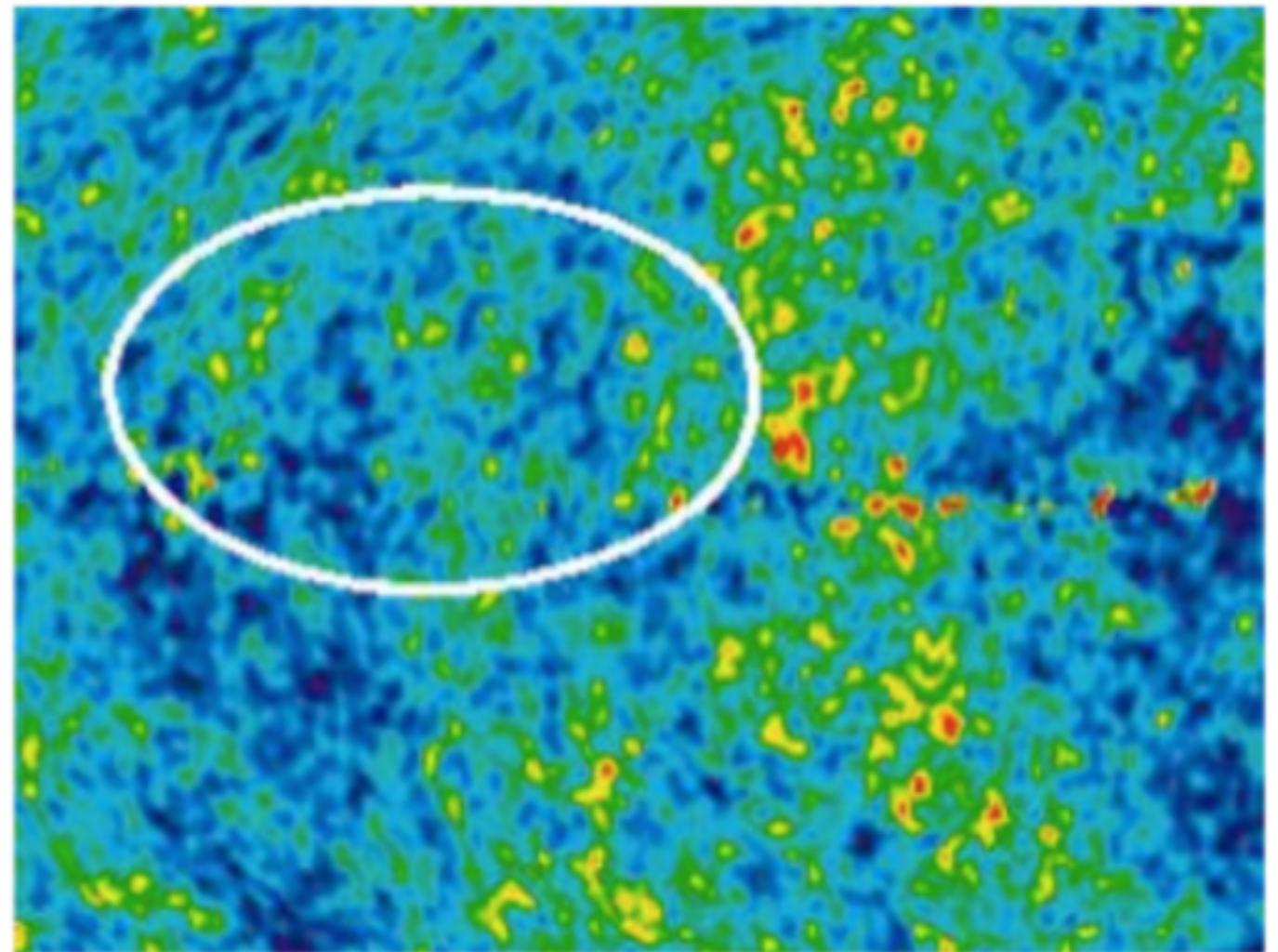


Stephen Hawkings leaves his mark (Image: NASA/WMAP Science Team)
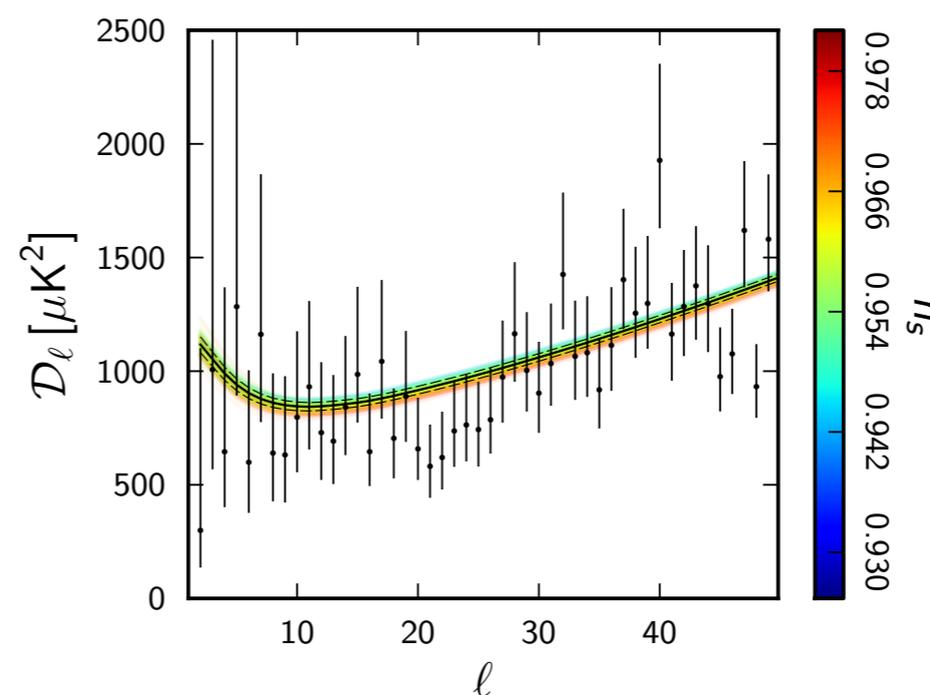


The main Planck anomaly is the low amplitudes at $\ell \approx$ 21-27

# Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Are There Cosmic Microwave Background Anomalies? C. Bennett et al. WMAP7 Jan 2010

A simple six-parameter LCDM model provides a successful fit to WMAP data, both when the data are analyzed alone and in combination with other cosmological data. Even so, it is appropriate to search for any hints of deviations from the now standard model of cosmology, which includes inflation, dark energy, dark matter, baryons, and neutrinos. The cosmological community has subjected the WMAP data to extensive and varied analyses. While there is widespread agreement as to the overall success of the six-parameter LCDM model, various "anomalies" have been reported relative to that model. In this paper we examine potential anomalies and present analyses and assessments of their significance. In most cases we find that claimed anomalies depend on posterior selection of some aspect or subset of the data. Compared with sky simulations based on the best fit model, one can select for low probability features of the WMAP data. Low probability features are expected, but it is not usually straightforward to determine whether any particular low probability feature is the result of the a posteriori selection or of non-standard cosmology. We examine in detail the properties of the power spectrum with respect to the LCDM model. We examine several potential or previously claimed anomalies in the sky maps and power spectra, including cold spots, low quadrupole power, quadropole-octupole alignment, hemispherical or dipole power asymmetry, and quadrupole power asymmetry. We conclude that there is no compelling evidence for deviations from the LCDM model, which is generally an acceptable statistical fit to WMAP and other cosmological data.

# Why do cosmic topological defects form?

If cosmic strings or other topological defects *can* form at a cosmological phase transition, then they *will* form. This was first pointed out by Tom Kibble and, in a cosmological context, the defect formation process is known as the **Kibble mechanism**.

The simple fact is that causal effects in the early universe can only propagate (as at any time) at the speed of light c. This means that at a time t, regions of the universe separated by more than a distance $d=ct$ can know nothing about each other. In a symmetry breaking phase transition, different regions of the universe will choose to fall into different minima in the set of possible states (this set is known to mathematicians as the vacuum manifold). Topological defects are precisely the "boundaries" between these regions with different choices of minima, and their formation is therefore an inevitable consequence of the fact that different regions cannot agree on their choices.

For example, in a theory with two minima, plus + and minus -, then neighboring regions separated by more than ct will tend to fall randomly into the different states (as shown below). Interpolating between these different minima will be a **domain wall**.
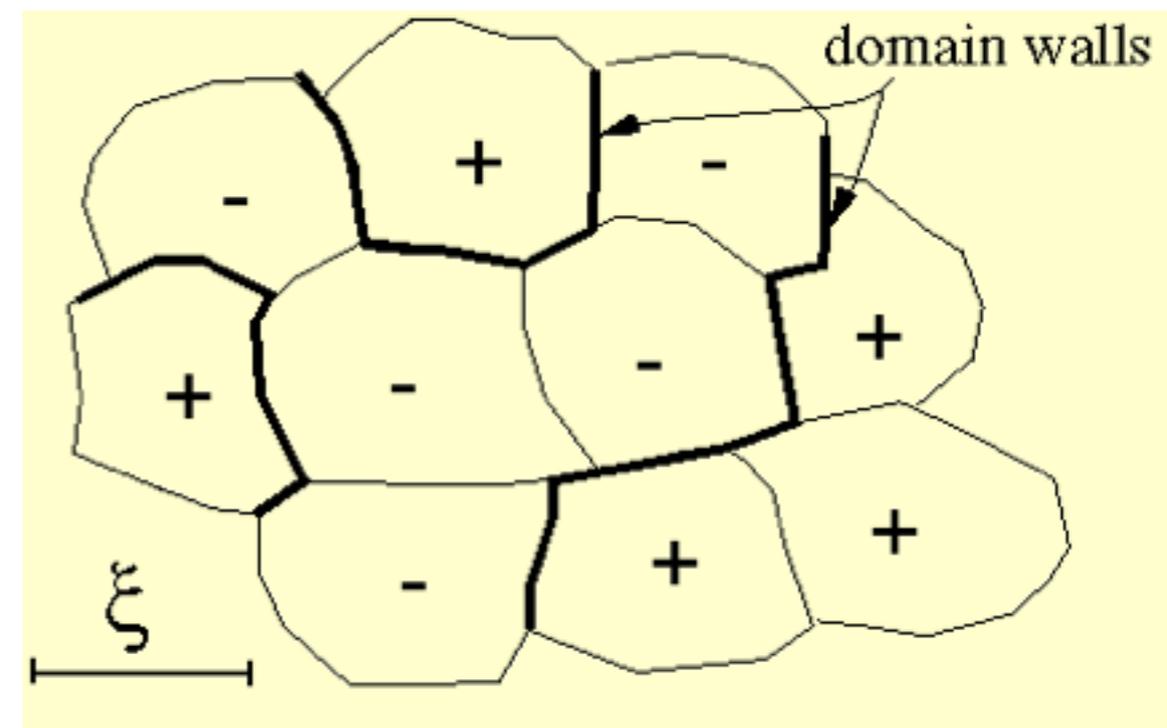
**Cosmic strings** will arise in slightly more complicated theories in which the minimum energy states possess `holes'. The strings will simply correspond to non-trivial `windings' around these holes (as illustrated at right).



*The Kibble mechanism for the formation of cosmic strings.*
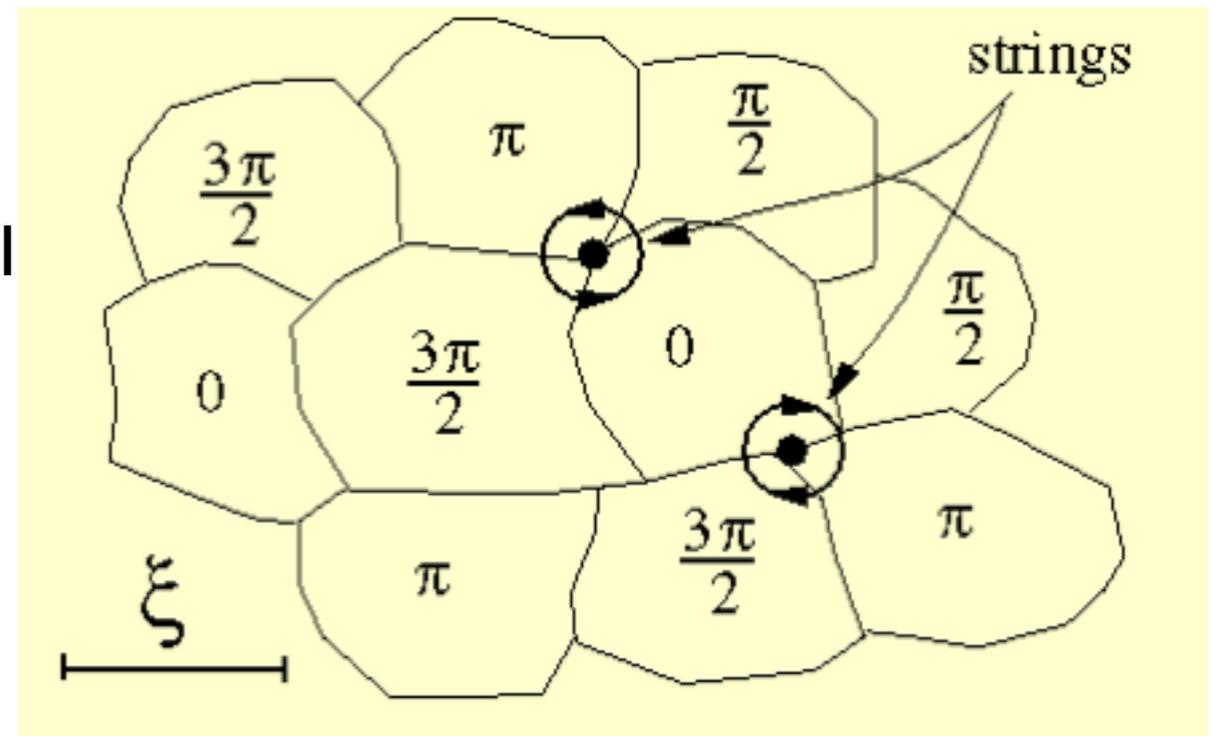
Topological defects can provide a unique link to the physics of the very early universe. Furthermore, they can crucially affect the evolution of the universe, so their study is an unavoidable part of any serious attempt to understand the early universe. The cosmological consequences vary with the type of defect considered. **Domain walls and monopoles are cosmologically catastrophic.** Any cosmological model in which they form will evolve in a way that contradicts the basic observational facts that we know about the universe. Such models must therefore be ruled out! **Cosmic inflation was invented to solve this problem.**

Cosmic strings and textures are (possibly) much more benign. Among other things, they were until recently thought to be a possible source of the fluctuations that led to the formation of the large-scale structures we observe today, as well as the anisotropies in the Cosmic Microwave Background. However, the CMB anisotropies have turned out not to agree with the predictions of this theory.

# Cosmic String Dynamics and Evolution

The evolution of cosmic string network is the relatively complicated result of only three rather simple and fundamental processes: cosmological expansion, intercommuting & loop production, and radiation.

## Cosmological expansion

The overall expansion of the universe will 'stretch' the strings, just like any other object that is not gravitationally bound. You can easily undertand this through the well-known analogy of the expanding balloon. If you draw a line of the surface of the balloon and then blow it up, you will see that the length of your `string' will grow at the same rate as the radius of the balloon.

## Intercommuting & loop production

Whenever two long strings cross each other, they exchange ends, or `intercommute' (case (a) in the figure below). In particular, a long string can intercommute with itself, in which case a loop will be produced (this is case (b) below).
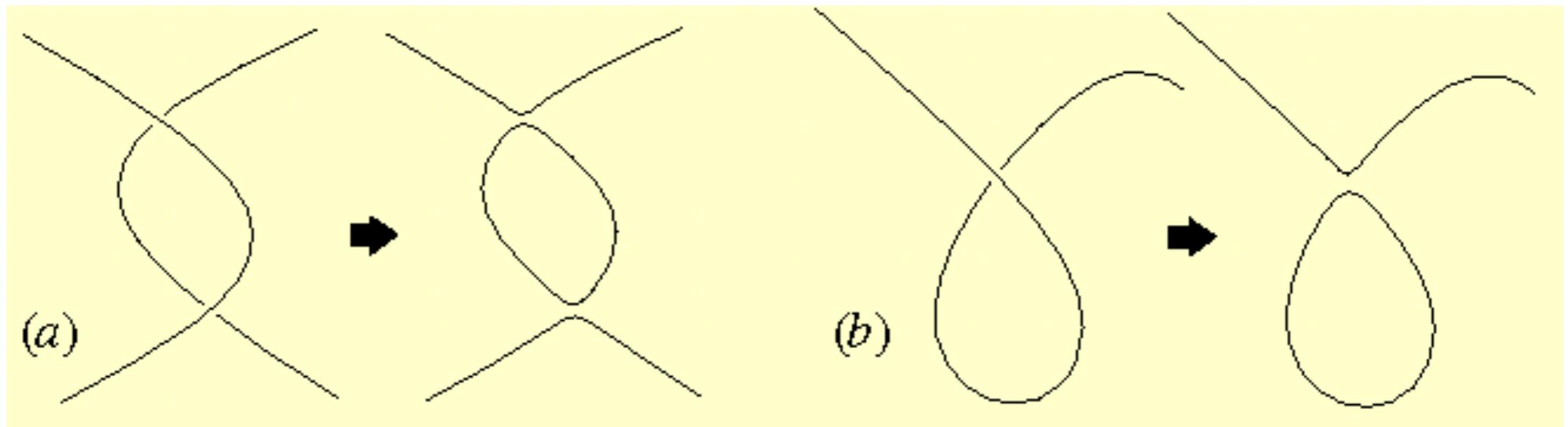
# Radiation from strings

Both long cosmic strings and small loops will emit radiation. In most cosmological scenarios this will be **gravitational radiation**, but electromagnetic radiation or axions can also be emitted in some cases (for some specific phase transitions).

The effect of radiation is much more dramatic for loops, since they lose all their energy this way, and eventually disappear. Here you can see what happens in the case of two interlocked loops. This configuration is unlikely to happen in a cosmological setting, but it is nevertheless quite enlightening. Notice the succession of complicated dynamic processes before the loop finally disappears!
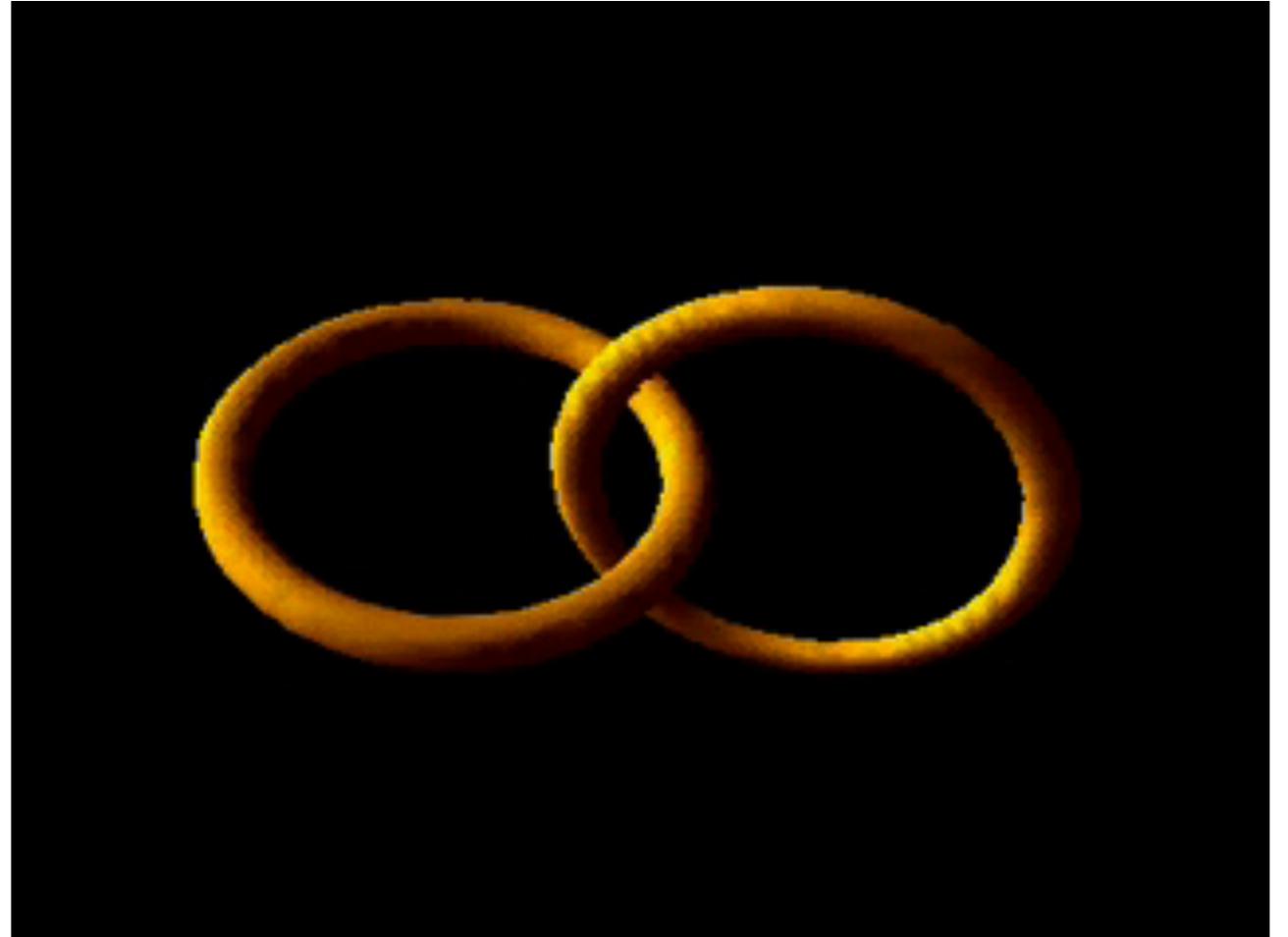


After formation, an initially high density string network begins to chop itself up by producing small loops. These loops oscillate rapidly (relativistically) and decay away into gravitational waves.

The net result is that the strings become more and more dilute with time as the universe expands. From an enormous density at formation, mathematical modelling suggests that today there would only be about 10 long strings stretching across the observed universe, together with about a thousand small loops!

In fact the network dynamics is such that the string density will eventually stabilize at an exactly constant level relative to the rest of the radiation and matter energy density in the universe. Thus the string evolution is described as `**scaling**' or scale-invariant, that is, the properties of the network look the same at any particular time $t$ if they are scaled (or multiplied) by the change in the time. This is schematically represented below:
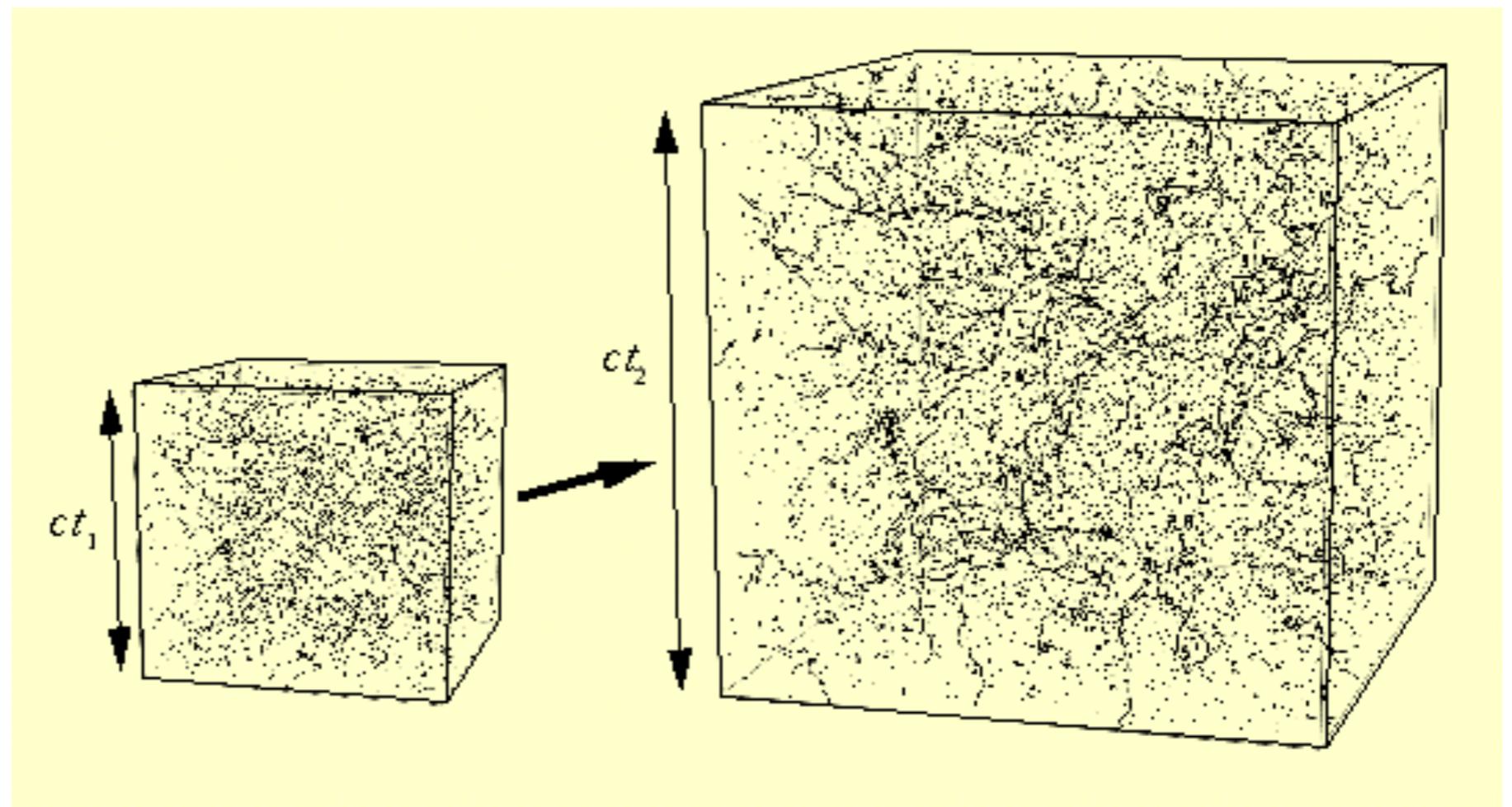
Because strings are extremely complex non-linear objects, the only rigorous way to study their evolution and cosmological consequences is to simulate in on the computer. One of the aims of performing numerical simulations of the evolution of cosmic string networks is to subsequently use the resulting information as an input to build (relatively) simpler semianalytic models that reproduce (in an averaged sense) the crucial properties of these objects. One starts by generating an initial "box of stings" containing a configuration of strings such as one would expect to find after a phase transition in the early universe. Then one evolves this initial box, by using the laws of motion of the strings.
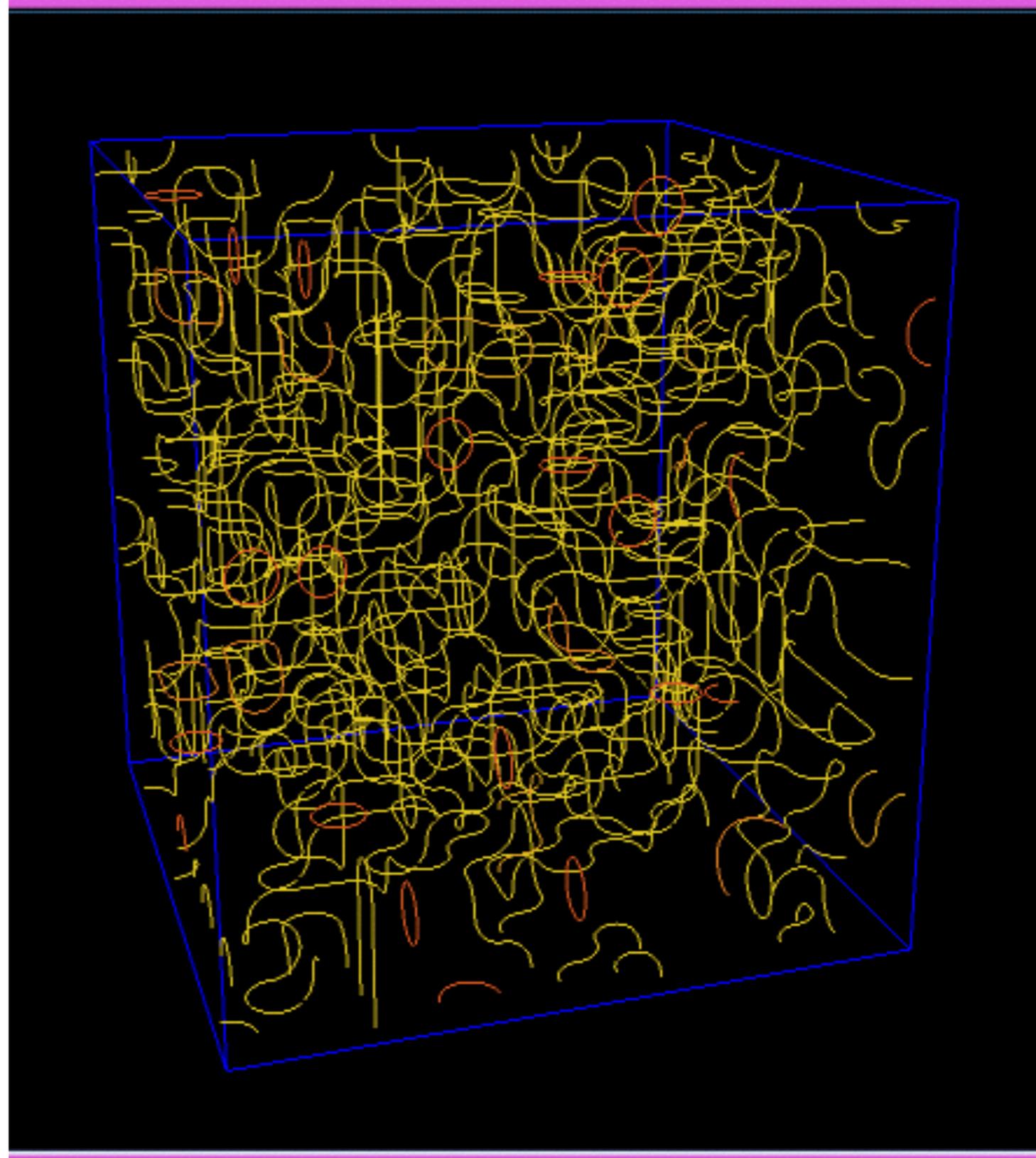


*In this and all other pictures and movies below long strings are shown in yellow, while small loops have a color code going from yellow to red according to their size (red loops being the smallest).*

# Why do the two boxes below look different? Because the rate at which the universe is expanding is different.
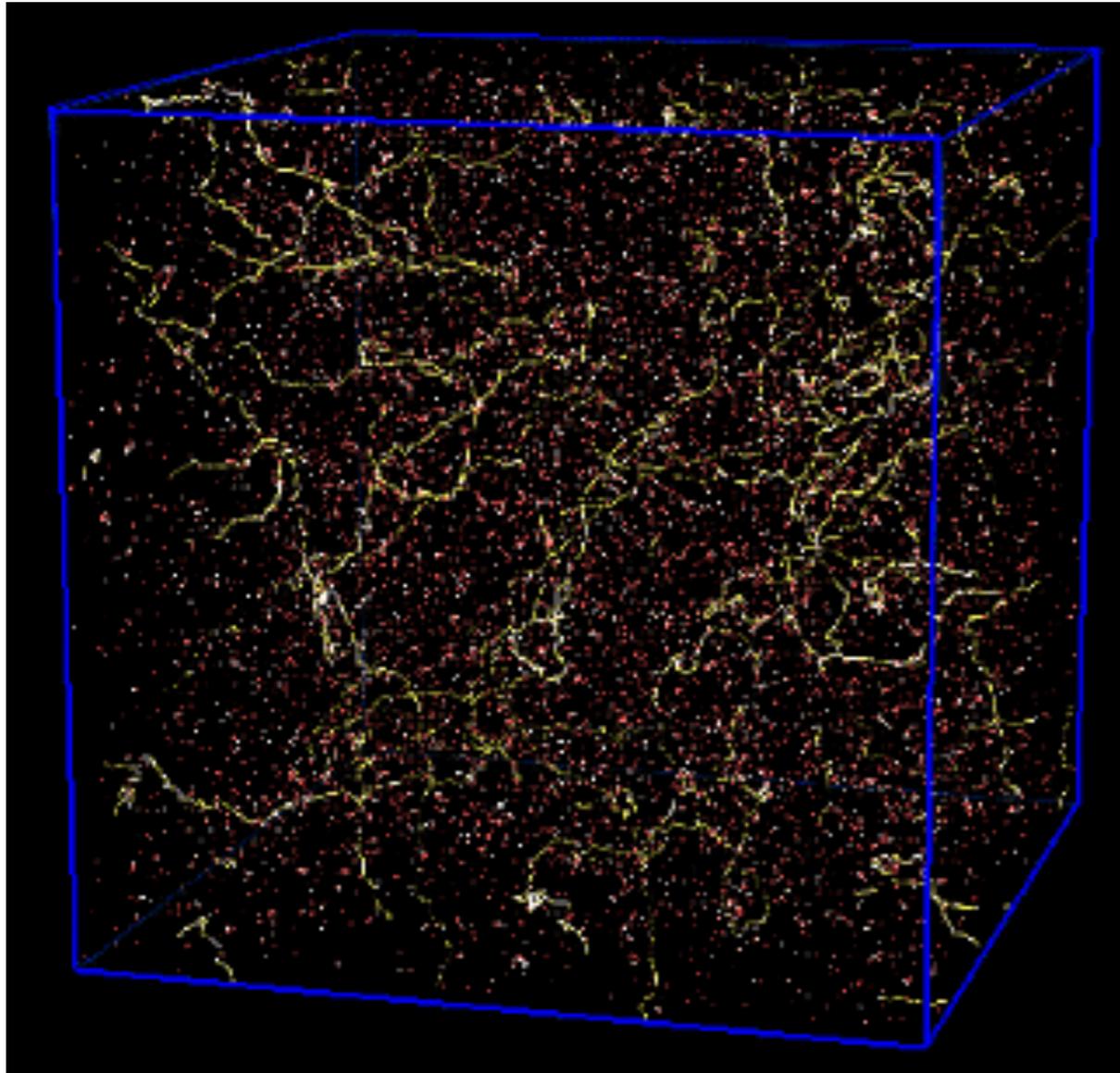


Snapshot of a string network in the **radiation era**. Note the high density of small loops and the `wiggliness' of the long strings in the network. The box size is about 2ct. (B. Allen & E. P. Shellard)
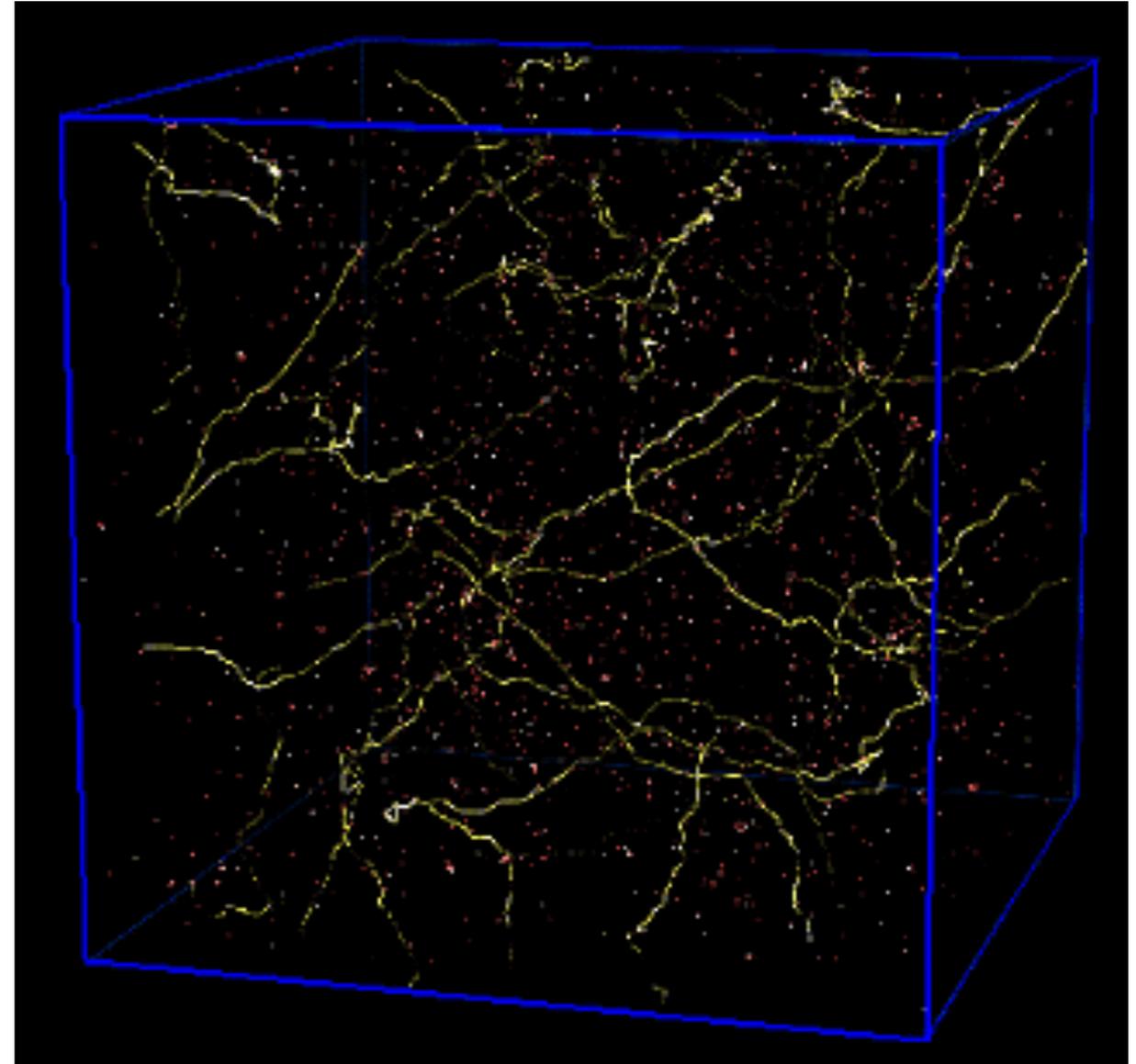
Snapshot of a string network in the **matter era**. Compare with the radiation case at left. Notice the lower density of both long strings and loops, as well as the lower `wiggliness' of the former. The box size is again about 2ct.

**Two movies of the evolution of a cosmic string network in the radiation era. In the movie on the left the box has a fixed size (so you will see fewer and fewer strings as it evolves), while in the one on the right it grows as the comoving horizon. (C. Martins & E. P. Shellard)**



fixed size box

box grows with horizon
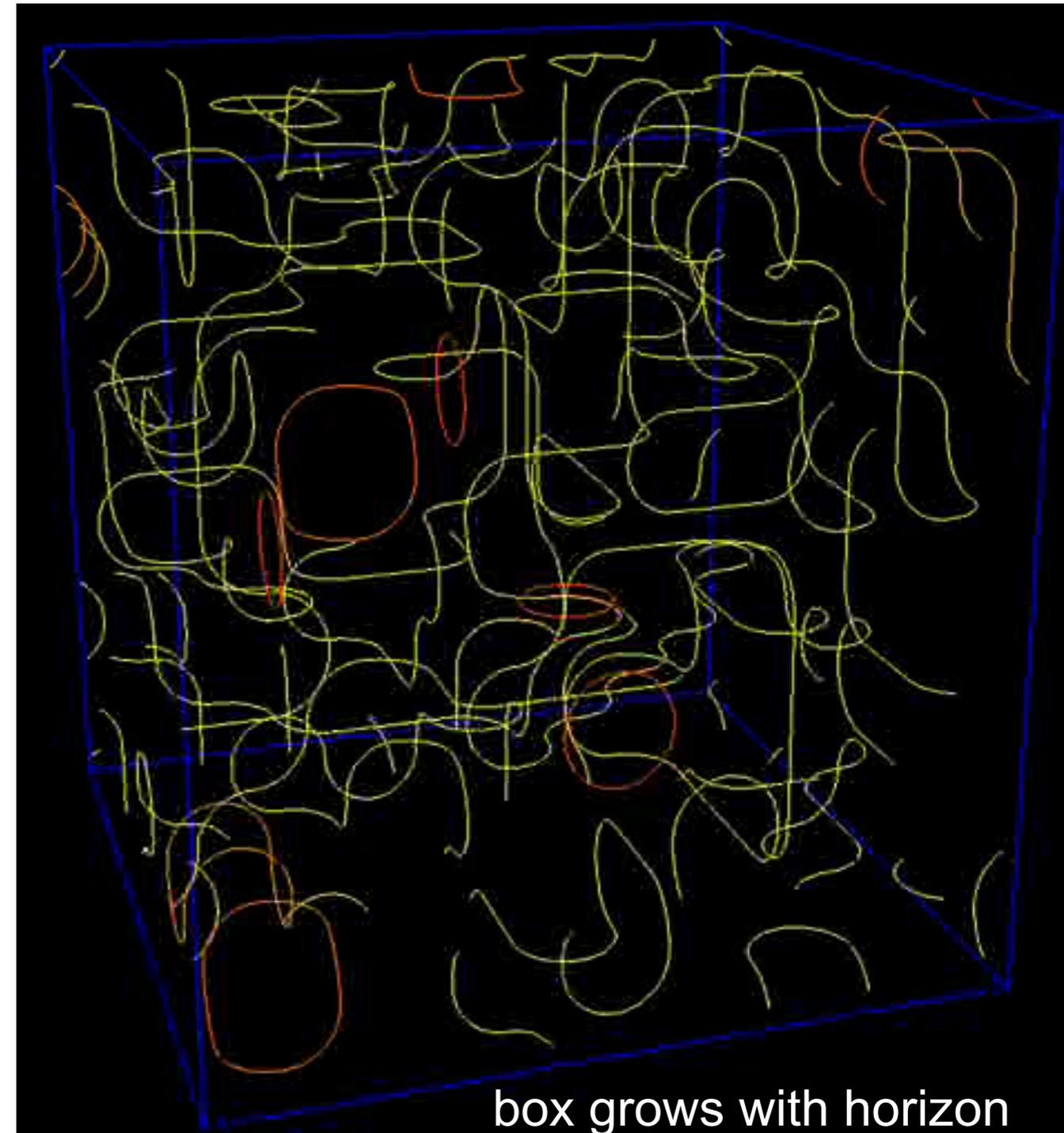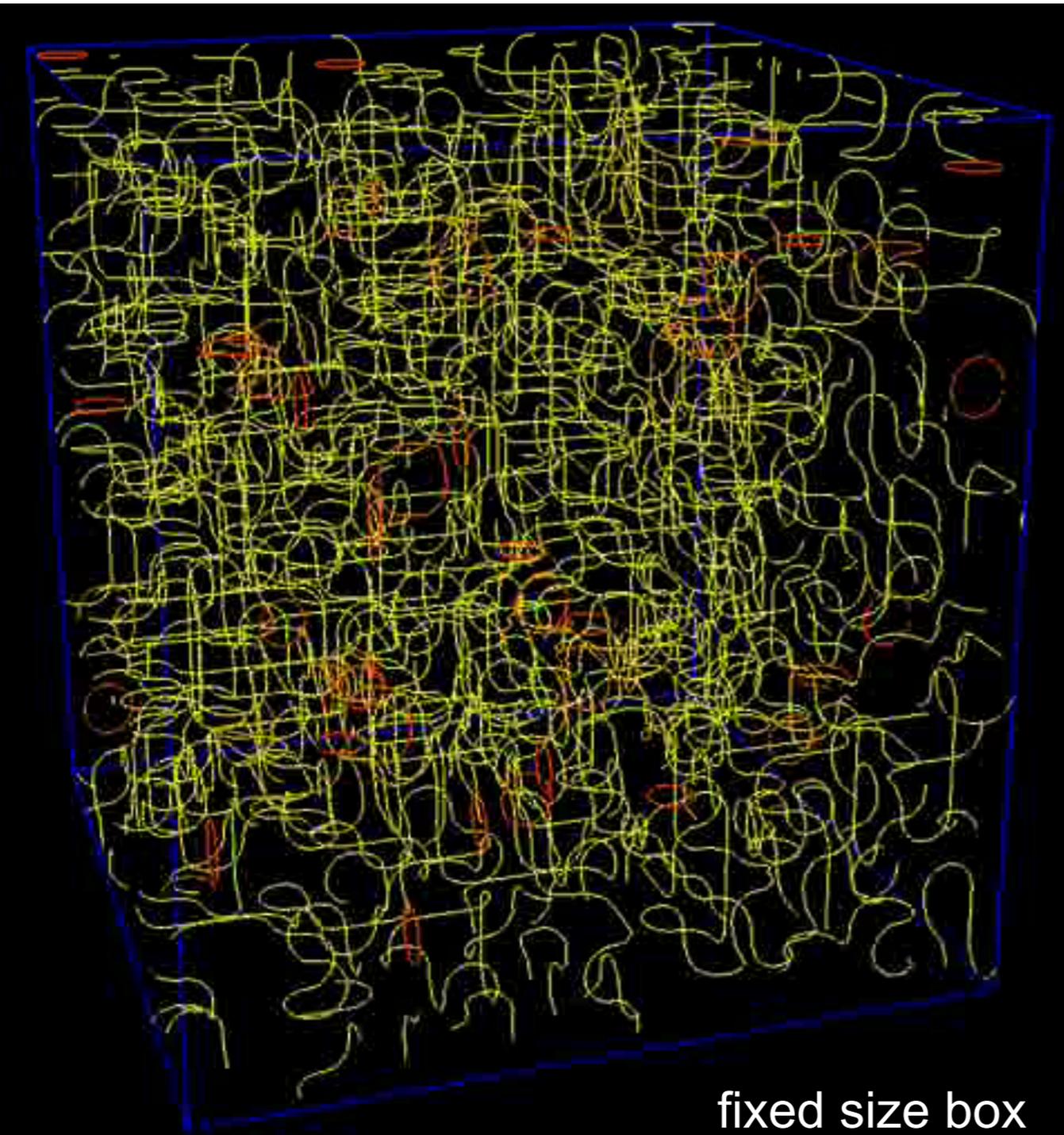
Notice that the number of long strings in the box that grows with the horizon remains roughly constant, in agreement with the scaling hypothesis. This is because the additional length in strings is quickly converted into small loops.

**Two movies of the evolution of a cosmic string network in the matter era.** In the movie on the left the box has a fixed size (so you will see fewer and fewer strings as it evolves), while in the one on the right it grows as the comoving horizon. (C. Martins & E. P. Shellard)



fixed size box

box grows with horizon
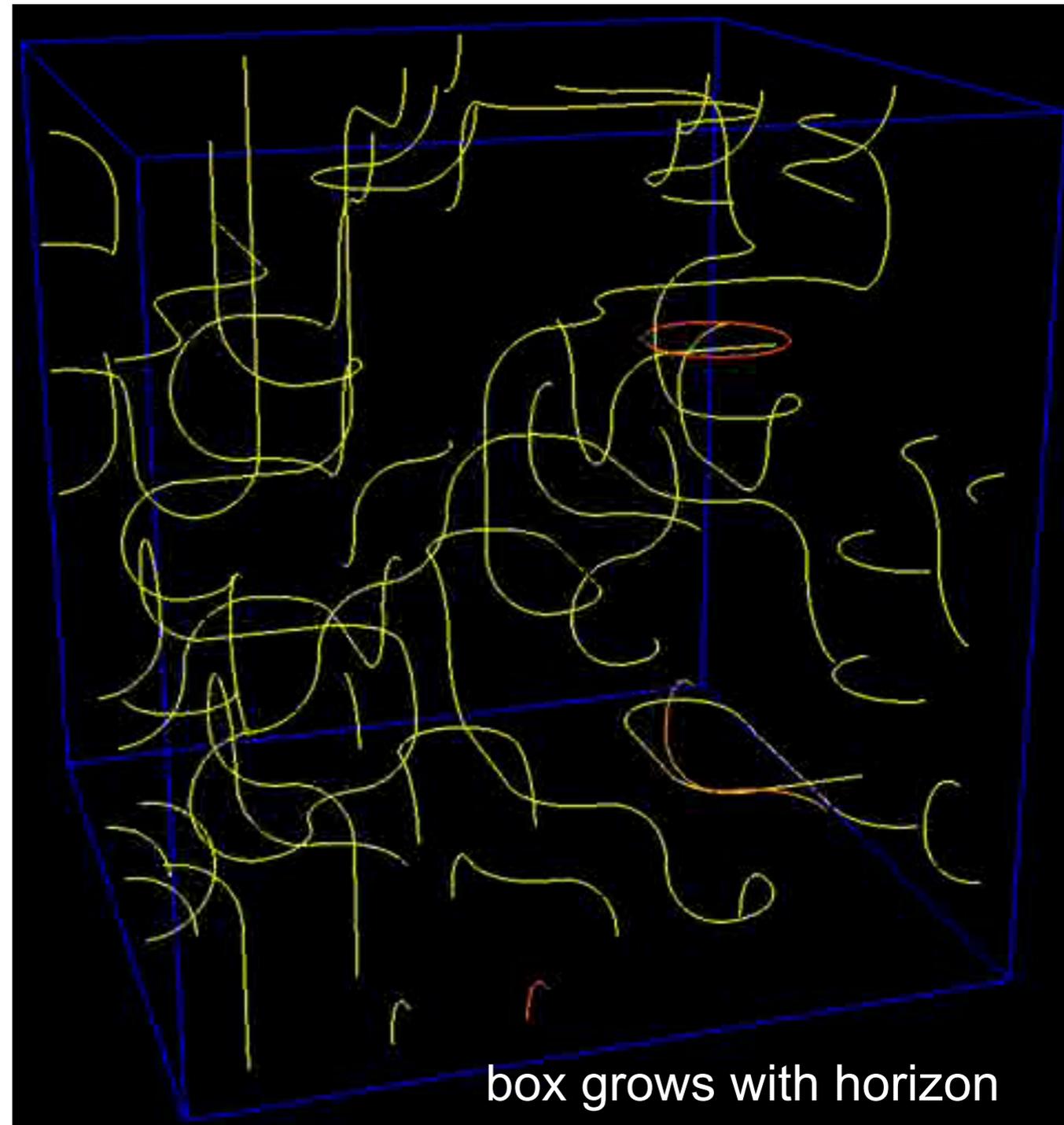
Notice that the number of long strings in the box that grows with the horizon again remains roughly constant, in agreement with the scaling hypothesis. This is because the additional length in strings is quickly converted into small

When strings evolve, scaling from smaller scales to larger ones, they create perturbations in the matter energy density of the universe.  Because of their tension, cosmic strings pull straight as they come inside the horizon.  Although there is no gravitational force from a static string, such moving cosmic strings produce wakes toward which matter falls, thus serving as seeds for structure formation. For a static string along the z axis of mass $\mu$ per unit length, the energy momentum tensor is

$$T^{\mu\nu} = \mu \, \mathrm{diag}(1,0,0,-1)\delta(x)\delta(y)$$

and the metric is

$$ds^2 = dt^2 - dz^2 - dR^2 - (1 - 4G\mu)^2 R^2 d\varphi^2$$

$G\mu \approx (M_{GUT}/M_{Pl})^2 \approx 10^{-6}$ is just the magnitude needed for GUT string structure formation.  There is an **angular defect** of $8\pi G\mu = 5.18''$ ($10^6 \, G\mu$). This implies

that the geodesic path of light is curved towards a string when light passes by it. Two copies of a galaxy near a cosmic string will appear to observers on the other side of the string.
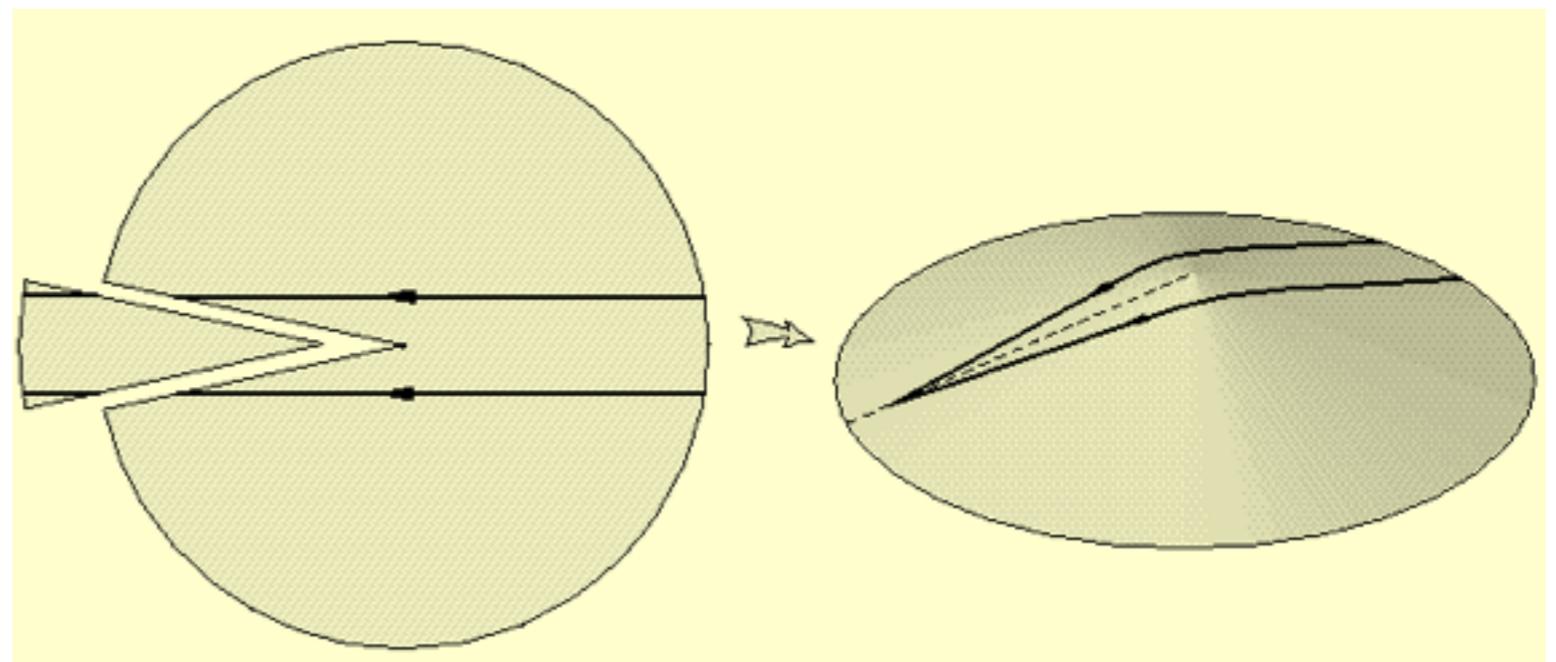


A. Vilenkin, E. P. S. Shellard, **Cosmic Strings and Other Topological Defects** (Cambridge U P, 1994)

# Cosmic Strings Summary

Cosmic strings arise in spontaneously broken (SB) gauge theories

$$\mathcal{L} = -\tfrac{1}{4} F_{\mu\nu} F^{\mu\nu} - |D_\mu \phi|^2 - \lambda \left( |\phi|^2 - \phi_v^2 \right)^2$$

as a consequence of causality in the expanding universe.

As the temperature T falls, a complex scalar field $\phi$ gets a nonzero expectation value

$$\phi(x) = \phi_v \, e^{\, i\theta(x)}$$

The phase $\theta$ will inevitably be different in regions separated by distances greater than

the horizon size when the SB phase transition occurred. If $\theta$ runs over $0 \to 2\pi$ as x goes around a loop in space, the loop encloses a string.



direction represents phase of $\phi$
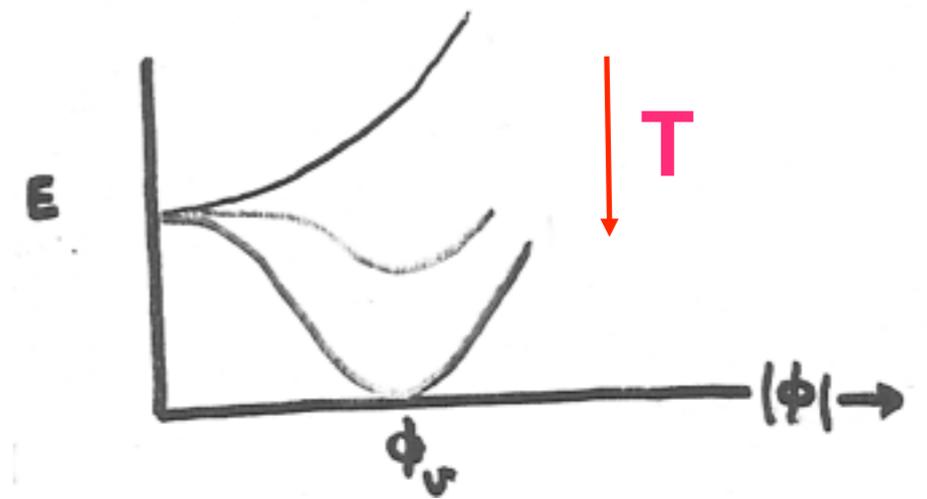
# By 2000, it was clear that **cosmic defects are not the main source of the CMB anisotropies**.
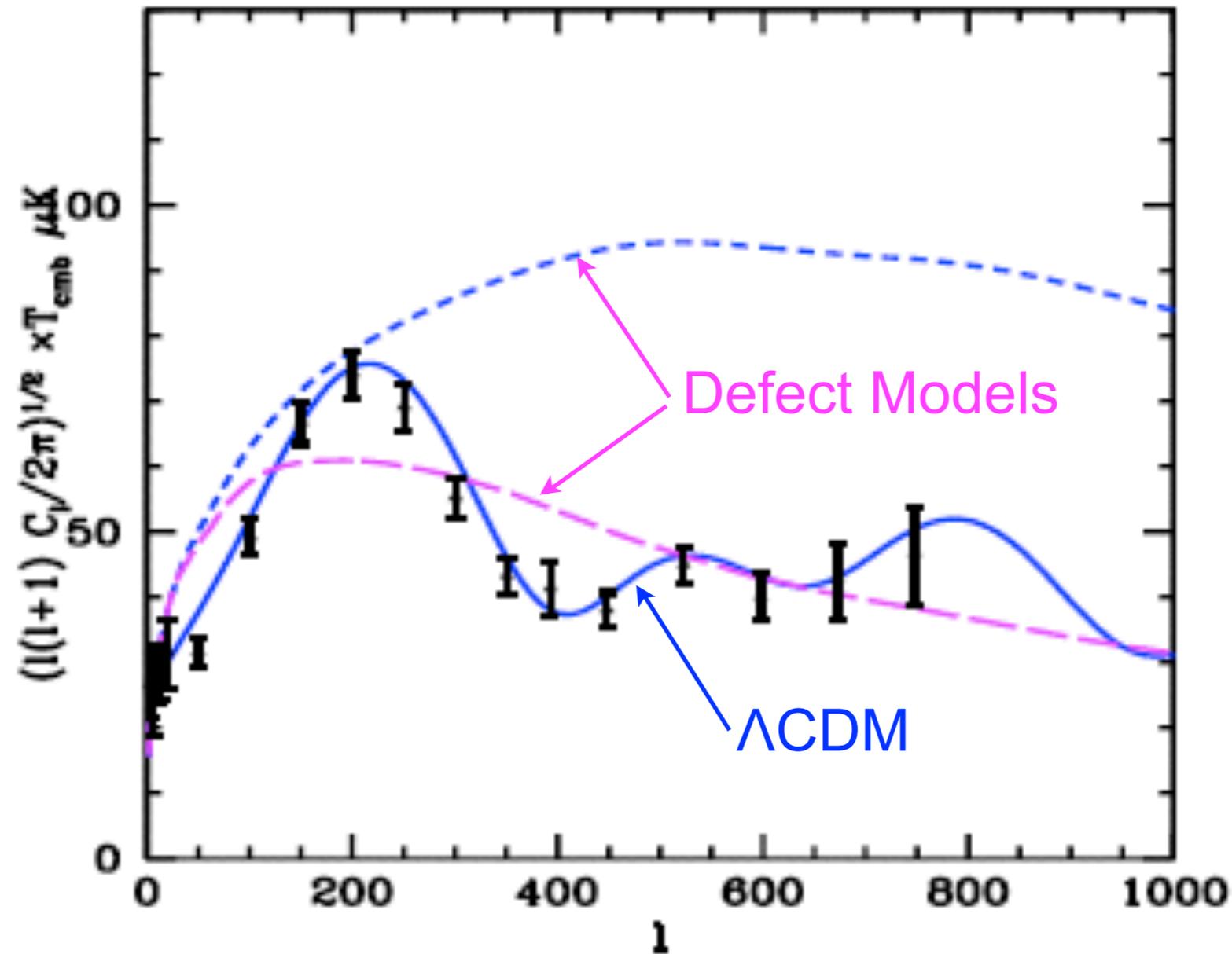


Figure 3: Current data (as complied by Knox[22]) with two defect models (dashed) and an inflation-based model (solid). The upper defect model has a standard ionization history and the lower model has an ionization history specifically designed to produce a sharper, shifted peak.

Andreas Albrecht, Defect models of cosmic structure in light of the new CMB data, XXXVth Rencontres de Moriond ``Energy Densities in the Universe'' (2000).

## Improved limits on short-wavelength gravitational waves from the cosmic microwave background, by Irene Sendra, Tristan L. Smith (arXiv:1203.4232)

We present updated observational bounds for both adiabatic and homogeneous initial conditions using the latest CMB data at small scales from the South Pole Telescope (SPT) in combination with Wilkinson Microwave Anisotropy Probe (WMAP), current measurements of the baryon acoustic oscillations, and the Hubble parameter. With the inclusion of the data from SPT the adiabatic bound on the CGWB density is improved by a factor of 1.7 to $\Omega_{gw} < 8.7 \times 10^{-6}$ at the 95% confidence level (C.L.), with weak evidence in favor of an additional radiation component consistent with previous analyses. The homogeneous bound improves by a factor of 3.5 to $\Omega_{gw} < 1.0 \times 10^{-6}$ at 95% C.L., with no evidence for such a component from current data.



## Constraints on cosmic string tension imposed by the limit on the stochastic gravitational wave background from the European Pulsar Timing Array, by S. Sanidas + (arXiv:1201.2419)
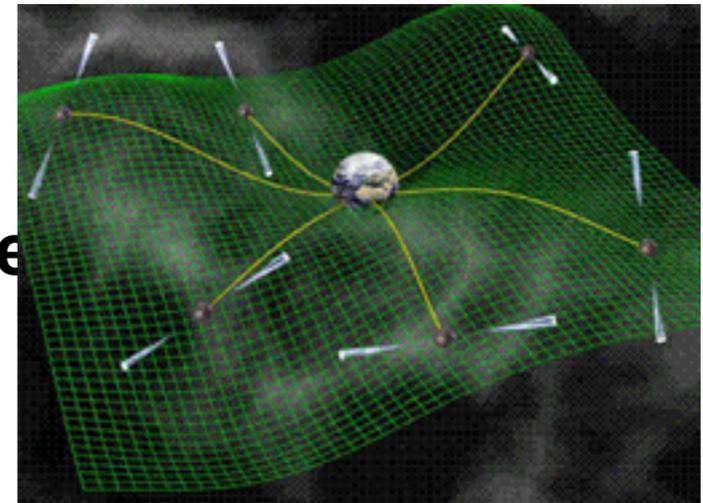
We investigate the constraints that can be placed on the cosmic string tension by using the current Pulsar Timing Array limits on the stochastic gravitational wave background (SGWB). Taking into account all the possible uncertainties in the parameters we find a conservative upper limit of $G\mu c^2 < 5.3 \times 10^{-7}$. We discuss the prospects for lowering this limit by two orders of magnitude, or even a detection of the SGWB, in the very near future in the context of the Large European Array for Pulsars and the Square Kilometre Array.

# GUT Monopoles

A simple SO(3) GUT illustrates how nonsingular monopoles arise. The Lagragian is

$$\mathcal{L} = \frac{1}{2}D_\mu\Phi^a D^\mu\Phi^a - \frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} - \frac{1}{8}\lambda(\Phi^a\Phi^a - \sigma^2)^2,$$

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - e\varepsilon_{abc}A^b_\mu A^c_\nu,$$

$$D_\mu\Phi^a = \partial_\mu\Phi^a - e\varepsilon_{abc}A^b_\mu\Phi^c.$$

The masses of the resulting charged vector and Higgs bosons after spontaneous symmetry breaking are

$$M^2_V = e^2\sigma^2,$$

$$M^2_S = \lambda\sigma^2.$$

If the Higgs field $\Phi^a$ happens to rotate about a sphere in SO(3) space as one moves around a sphere about any particular point in $\mathbf{x}$-space, then it must vanish at that point. Remarkably, if we identify the massless vector field as the photon, this configuration corresponds to a nonsingular magnetic monopole, as was independently discovered by 'tHooft and Polyakov. The monopole has magnetic charge twice the minimum Dirac value, g = 2π/e = (4π/e$^2$)(e/2) ≈ 67.5 e.

The singular magnetic field is cut off at scale σ, and as a result the GUT monopole has mass $M_{monopole} \approx M_V/\alpha \approx M_{GUT}/\alpha \approx 10^{18}$ GeV.

The first accurate calculation of the mass of the 't Hooft - Polyakov non-singular monopole was Bais & Primack (Phys. Rev. D13:819,1976).

# GUT Monopole Problem

The Kibble mechanism produces ~ one GUT monopole per horizon volume when the GUT phase transition occurs. These GUT monopoles have a number density over entropy (using the old $T_{GUT} \sim M_{GUT} \sim 10^{14}$ GeV)

$$n_M/s \sim 10^2 \, (T_{GUT}/M_{Pl})^3 \sim 10^{-13}$$

(compared to $n_B/s \sim 10^{-9}$ for baryons) Their annihilation is inefficient since they are so massive, and as a result they are about as abundant as gold atoms but $10^{16}$ times more massive, so they "overclose" the universe. This catastrophe must be avoided! **This was Alan Guth's initial motivation for inventing cosmic inflation.**

I will summarize the key ideas of inflation theory, following my lectures at the Jerusalem Winter School, published as the first chapter in Avishai Dekel & Jeremiah Ostriker, eds., *Formation of Structure in the Universe* (Cambridge University Press, 1999), and Dierck-Ekkehard Liebscher, *Cosmology* (Springer, 2005) (available online through the UCSC library).

# Motivations for Inflation

| PROBLEM | SOLVED |
|---|---|
| Horizon | Homogeneity, Isotropy, Uniform T |
| Flatness/Age | Expansion and gravity balance |
| "Dragons" | Monopoles, doman walls,...banished |
| Structure | Small fluctuations to evolve into galaxies, clusters, voids |

Cosmological constant $\Lambda > 0 \Rightarrow$ space repels space, so the more space the more repulsion, $\Rightarrow$ de Sitter exponential expansion $a \propto e^{\sqrt{\Lambda} t}$.

Inflation is exponentially accelerating expansion caused by effective cosmological constant ("false vacuum" energy) associated with hypothetical scalar field ("inflaton").

| | FORCES OF NATURE | Spin |
|---|---|---|
| Known { | Gravity | 2 |
| | Strong, weak, and electromagnetic | 1 |
| Goal of LHC | Mass (Higgs Boson) | 0 |
| Early universe | Inflation (Inflaton) | 0 |

Inflation lasting only $\sim 10^{-32}$s suffices to solve all the problems listed above. Universe must then convert to ordinary expansion through conversion of false to true vacuum ("re-"heating).

Joel Primack, in *Formation of Structure in the Universe*, ed. Dekel & Ostriker (Cambridge Univ Press, 1999)

# Inflation Basics

The basic idea of inflation is that before the universe entered the present adiabatically expanding Friedmann era, it underwent a period of de Sitter exponential expansion of the scale factor, termed *inflation* (Guth 1981). Actually, inflation is never precisely de Sitter, and any superluminal (faster-than-light) expansion is now called inflation. Inflation was originally invented to solve the problem of too many GUT monopoles, which, as mentioned in the previous section, would otherwise be disastrous for cosmology.

The de Sitter cosmology corresponds to the solution of Friedmann's equation in an empty universe (i.e., with $\rho = 0$) with vanishing curvature ($k = 0$) and positive cosmological constant ($\Lambda > 0$). The solution is $a = a_o e^{Ht}$, with constant Hubble parameter $H = (\Lambda/3)^{1/2}$. There are analogous solutions for $k = +1$ and $k = -1$ with $a \propto \cosh Ht$ and $a \propto \sinh Ht$ respectively. The scale factor expands exponentially because the positive cosmological constant corresponds effectively to a negative pressure. de Sitter space is discussed in textbooks on general relativity (for example, Rindler 1977, Hawking & Ellis 1973) mainly for its geometrical interest. Until cosmological inflation was considered, the chief significance of the de Sitter solution in cosmology was that it is a limit to which all indefinitely expanding models with $\Lambda > 0$ must tend, since as $a \to \infty$, the cosmological constant term ultimately dominates the right hand side of the Friedmann equation.

Joel Primack, in *Formation of Structure in the Universe*, (Cambridge Univ Press, 1999)

As Guth (1981) emphasized, the de Sitter solution might also have been important in the very early universe because the vacuum energy that plays such an important role in spontaneously broken gauge theories also acts as an effective cosmological constant. A period of de Sitter inflation preceding ordinary radiation-dominated Friedmann expansion could explain several features of the observed universe that otherwise appear to require very special initial conditions: the horizon, flatness/age, monopole, and structure formation problems. (See Table 1.6.)

Let us illustrate how inflation can help with the horizon problem. At recombination $(p^+ + e^- \rightarrow H)$, which occurs at $a/a_o \approx 10^{-3}$, the mass encompassed by the horizon was $M_H \approx 10^{18} M_\odot$, compared to $M_{H,o} \approx 10^{22} M_\odot$ today. Equivalently, the angular size today of the causally connected regions at recombination is only $\Delta\theta \sim 3°$. Yet the fluctuation in temperature of the cosmic background radiation from different regions is very small: $\Delta T/T \sim 10^{-5}$. How could regions far out of causal contact have come to temperatures that are so precisely equal? This is the "horizon problem". With inflation, it is no problem because the entire observable universe initially lay inside a single causally connected region that subsequently inflated to a gigantic scale. Similarly, inflation exponentially dilutes any preceeding density of monopoles or other unwanted relics (a modern version of the "dragons" that decorated the unexplored borders of old maps). Joel Primack, in *Formation of Structure in the Universe*, (Cambridge Univ Press, 1999)

In the first inflationary models, the dynamics of the very early universe was typically controlled by the self-energy of the Higgs field associated with the breaking of a Grand Unified Theory (GUT) into the standard 3-2-1 model: GUT$\rightarrow SU(3)_{color} \otimes [SU(2) \otimes U(1)]_{electroweak}$. This occurs when the cosmological temperature drops to the unification scale $T_{GUT} \sim 10^{14}$ GeV at about $10^{-35}$ s after the Big Bang. Guth (1981) initially considered a scheme in which inflation occurs while the universe is trapped in an unstable state (with the GUT unbroken) on the wrong side of a maximum in the Higgs potential. This turns out not to work: the transition from a de Sitter to a Friedmann universe never finishes (Guth & Weinberg 1981). The solution in the "new inflation" scheme (Linde 1982; Albrecht and Steinhardt 1982) is for inflation to occur *after* barrier penetration (if any). It is necessary that the potential of the scalar field controlling inflation ("*inflaton*") be nearly flat (i.e., decrease very slowly with increasing inflaton field) for the inflationary period to last long enough. This nearly flat part of the potential must then be followed by a very steep minimum, in order that the energy contained in the Higgs potential be rapidly shared with the other degrees of freedom ("reheating"). A more general approach, "chaotic" inflation, has been worked out by Linde (1983, 1990) and others; this works for a wide range of inflationary potentials, including simple power laws such as $\lambda\phi^4$. However, for the amplitude of the fluctuations to be small enough for consistency with observations, it is necessary that the inflaton self-coupling be very small, for example $\lambda \sim 10^{-14}$ for the $\phi^4$ model. This requirement prevents a Higgs field from being the inflaton, since Higgs fields by definition have gauge couplings to the gauge field (which are expected to be of order unity), and these would generate self-couplings of similar magnitude even if none were present.

It turns out to be necessary to inflate by a factor $\gtrsim e^{66}$ in order to solve the flatness problem, i.e., that $\Omega_0 \sim 1$. (With $H^{-1} \sim 10^{-34}$ s during the de Sitter phase, this implies that the inflationary period needs to last for only a relatively small time $\tau \gtrsim 10^{-32}$ s.) The "flatness problem" is essentially the question why the universe did not become curvature dominated long ago. Neglecting the cosmological constant on the assumption that it is unimportant after the inflationary epoch, the Friedmann equation can be written

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\frac{\pi^2}{30}g(T)T^4 - \frac{kT^2}{(aT)^2}$$

where the first term on the right hand side is the contribution of the energy density in relativistic particles and $g(T)$ is the effective number of degrees of freedom. The second term on the right hand side is the curvature term. Since $aT \approx$ constant for adiabatic expansion, it is clear that as the temperature $T$ drops, the curvature term becomes increasingly important. The quantity $K \equiv k/(aT)^2$ is a dimensionless measure of the curvature. Today, $|K| = |\Omega - 1|\,H_o^2/T_o^2 \leq 2 \times 10^{-58}$. Unless the curvature exactly vanishes, the most "natural" value for $K$ is perhaps $K \sim 1$. Since inflation increases $a$ by a tremendous factor $e^{H\tau}$ at essentially constant $T$ (after reheating), it increases $aT$ by the same tremendous factor and thereby decreases the curvature by that factor squared. Setting $e^{-2H\tau} \lesssim 2 \times 10^{-58}$ gives the needed amount of inflation: $H\tau \gtrsim 66$. This much inflation turns out to be enough to take care of the other cosmological problems mentioned above as well.

$e^{66}=$ $4 \times 10^{28}$

$10^{-32}$ seconds of COSMIC INFLATION

13.7 billion years of COSMIC EXPANSION

Cosmic Uroboros

GUT

$10^{30}$ cm
$10^{25}$ cm
$10^{20}$ cm
$10^{15}$ cm
$10^{10}$ cm
$10^{5}$ cm
1 cm

$10^{-25}$ cm
$10^{-20}$ cm
$10^{-15}$ cm
$10^{-10}$ cm
$10^{-5}$ cm

DM?

W, Z
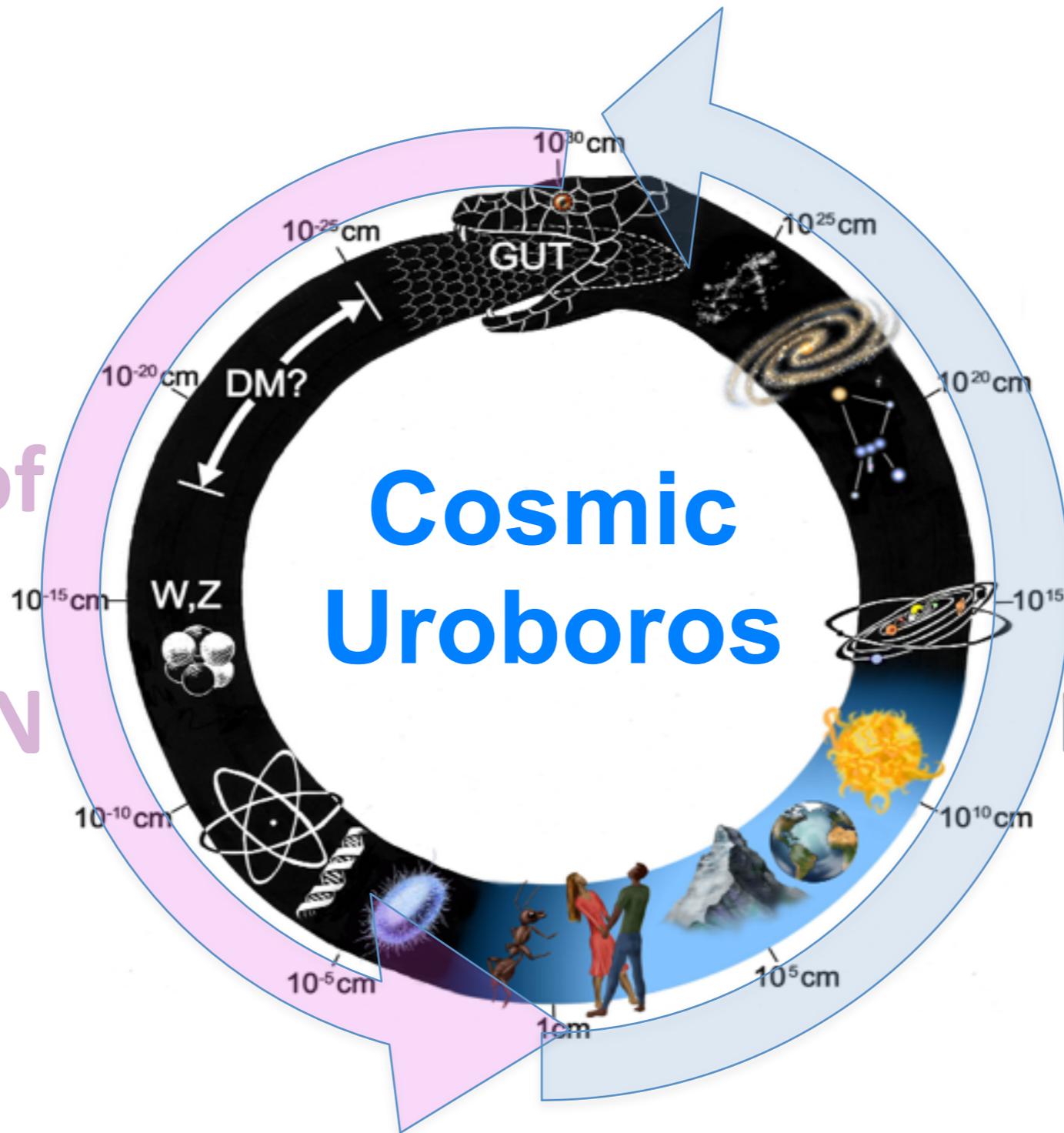
According to Cosmic Inflation theory, the entire visible universe was once about $10^{-29}$ cm in size. Its size then inflated by a factor of about $10^{30}$ so that when Cosmic Inflation ended (after about $10^{-32}$ second) it had reached the size of a newborn baby.

During its entire subsequent evolution, the size of the visible universe has increased by a similar factor of $10^{29}$.

**$10^{-32}$ seconds of COSMIC INFLATION**

**13.7 billion years of COSMIC EXPANSION**



Cosmic Uroboros

# COSMIC TIME



Fast Cosmic Inflation

Milky Andromeda forms

NOW

Slow Cosmic Inflation

Virgo Cluster Disappears Over Cosmic Horizon

0    10    14    20    30    40    50    60    70

time, in billions of years

Slow Expansion Slowing Down

Expansion Speeding Up
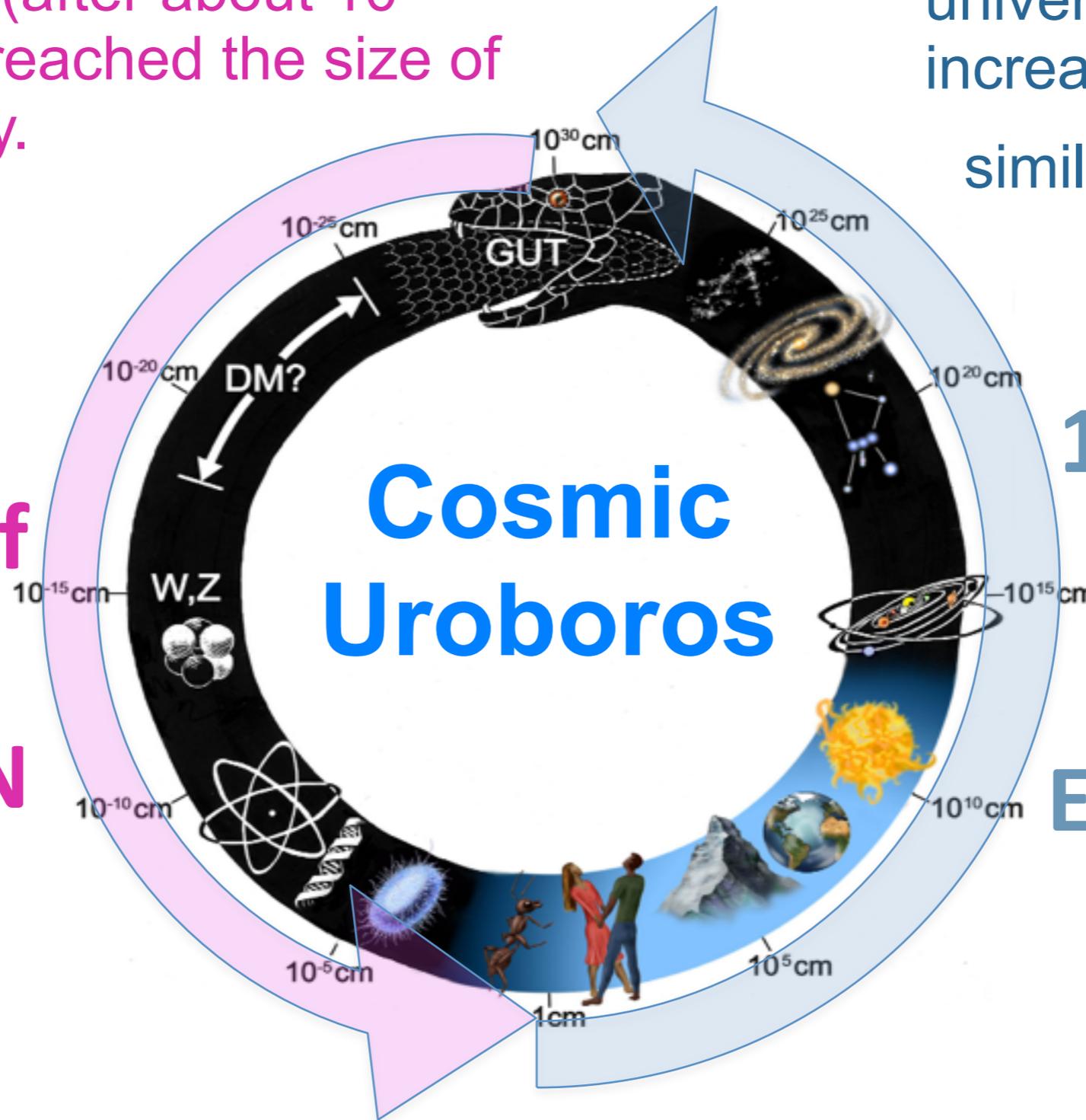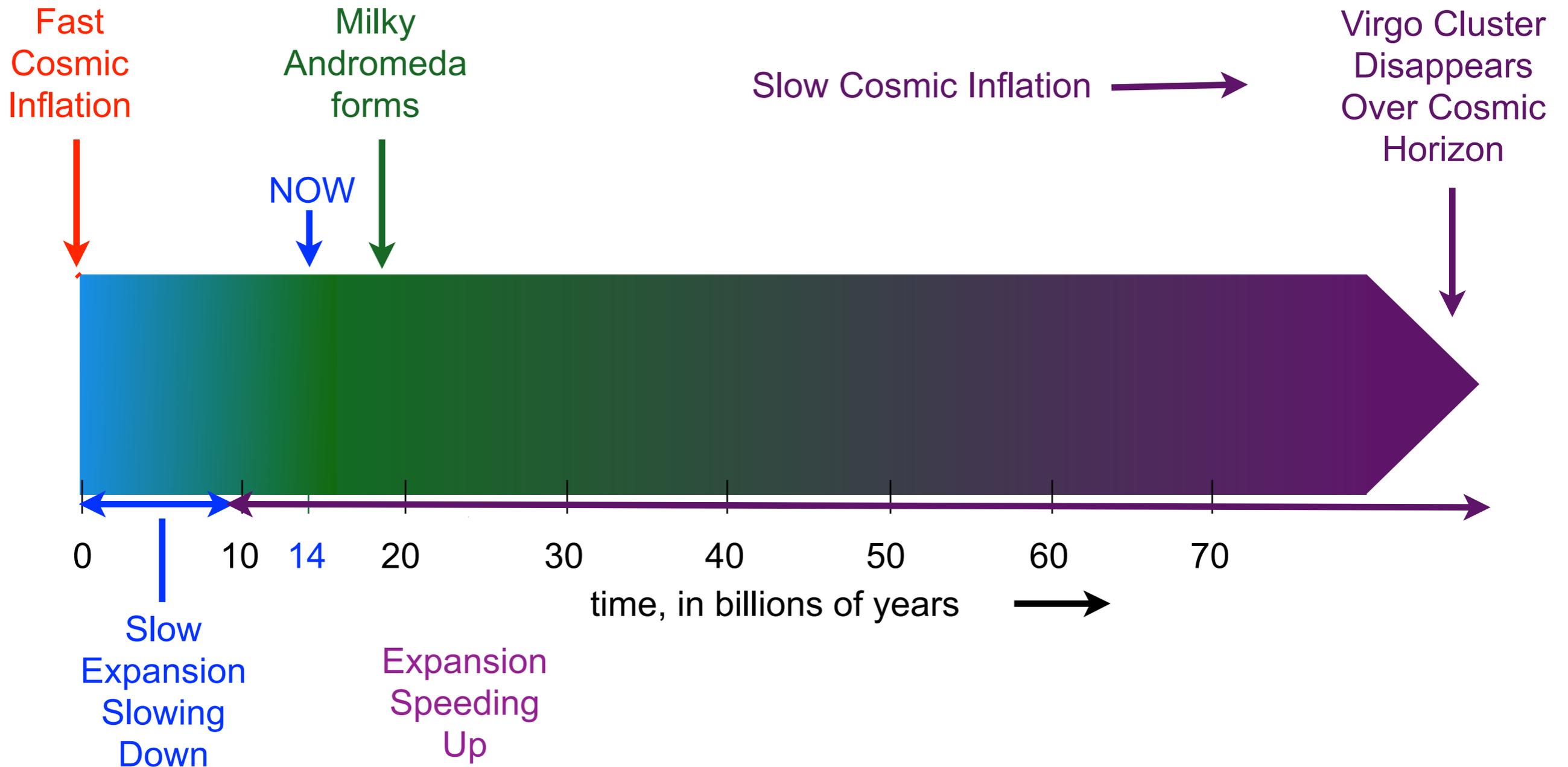
**Inflationary Fluctuations**
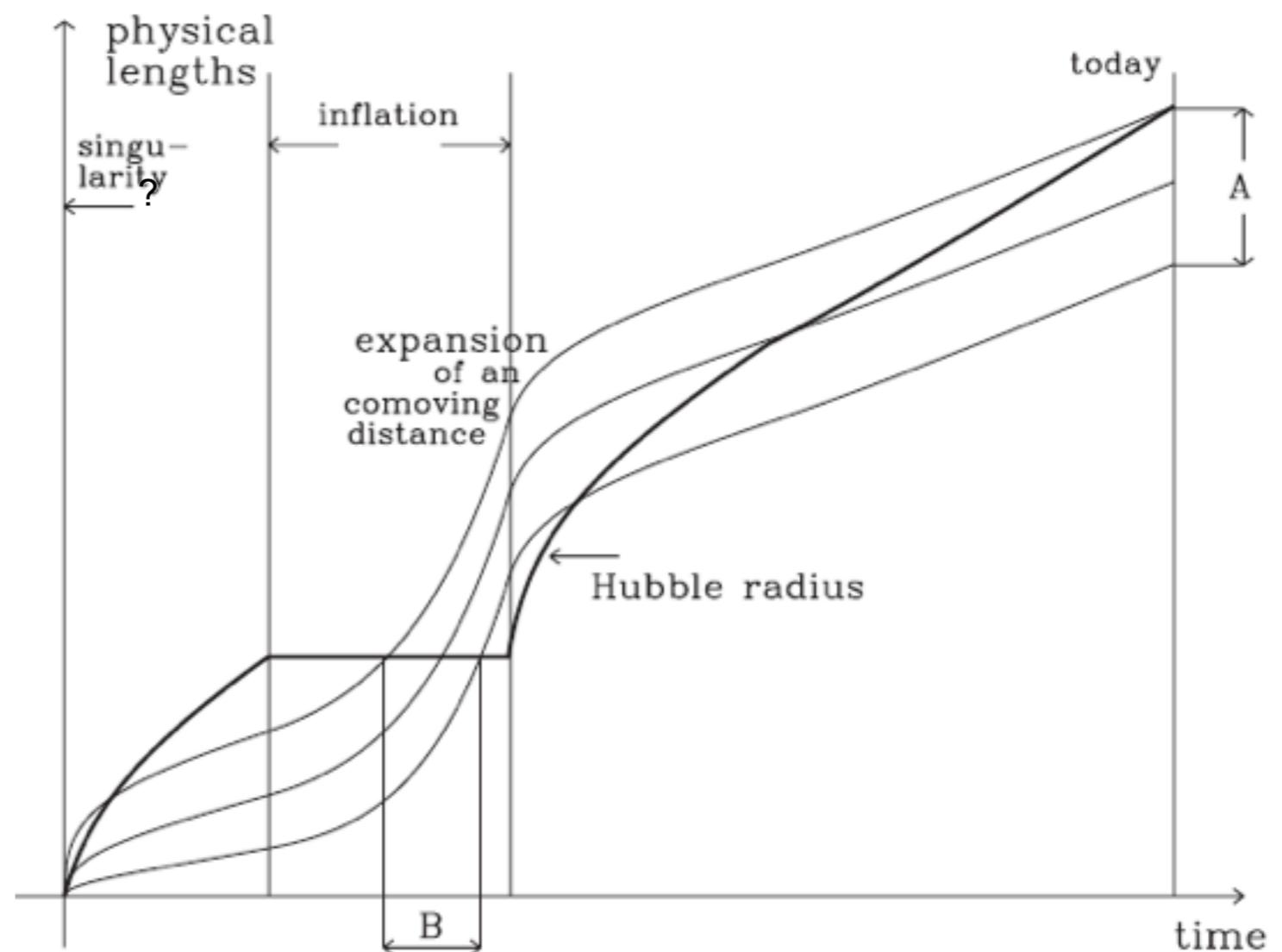
Thus far, it has been sketched how inflation stretches, flattens, and smooths out the universe, thus greatly increasing the domain of initial conditions that could correspond to the universe that we observe today. But inflation also can explain the origin of the fluctuations necessary in the gravitional instability picture of galaxy and cluster formation. Recall that the very existence of these fluctuations is a problem in the standard Big Bang picture, since these fluctuations are much larger than the horizon at early times. How could they have arisen?

The answer in the inflationary universe scenario is that they arise from quantum fluctuations in the inflaton field $\phi$ whose vacuum energy drives inflation. The scalar fluctuations $\delta\phi$ during the de Sitter phase are of the order of the Hawking temperature $H/2\pi$. Because of these fluctuations, there is a time spread $\Delta t \approx \delta\phi/\dot\phi$ during which different regions of the same size complete the transition to the Friedmann phase. The result is that the density fluctuations when a region of a particular size re-enters the horizon are equal to (Guth & Pi 1982; see Linde 1990 for alternative approaches) $\delta_H \equiv (\delta\rho/\rho)_H \sim \Delta t/t_H = H\Delta t$. The time spread $\Delta t$ can be estimated from the equation of motion of $\phi$ (the free Klein-Gordon equation in an expanding universe): $\ddot\phi + 3H\dot\phi = -(\partial V/\partial\phi)$. Neglecting the $\ddot\phi$ term, since the scalar potential $V$ must be very flat in order for enough inflation to occur (this is called the "slow roll" approximation), $\dot\phi \approx -V'/(3H)$, so $\delta_H \sim H^3/V' \sim V^{3/2}/V'$. Unless there is a special feature in the potential $V(\phi)$ as $\phi$ rolls through the scales of importance in cosmology (producing such "designer inflation" features generally requires fine tuning — see e.g. Hodges et al. 1990), $V$ and $V'$ will hardly vary there and hence $\delta_H$ will be essentially constant. These are fluctuations of all the contents of the universe, so they are adiabatic fluctuations.

Joel Primack, in *Formation of Structure in the Universe*, (Cambridge Univ Press, 1999)

Thus *inflationary models typically predict a nearly constant curvature spectrum $\delta_H$ = constant of adiabatic fluctuations.* Some time ago Harrison (1970), Zel'dovich (1972), and others had emphasized that this is the only scale-invariant (i.e., power-law) fluctuation spectrum that avoids trouble at both large and small scales. If $\delta_H \propto M_H^{-\alpha}$, where $M_H$ is the mass inside the horizon, then if $-\alpha$ is too large the universe will be less homogeneous on large than small scales, contrary to observation; and if $\alpha$ is too large, fluctuations on sufficiently small scales will enter the horizon with $\delta_H \gg 1$ and collapse to black holes (see e.g. Carr, Gilbert, & Lidsey 1995, Bullock & Primack 1996); thus $\alpha \approx 0$. The $\alpha = 0$ case has come to be known as the Zel'dovich spectrum.

Inflation predicts more: it allows the calculation of the value of the constant $\delta_H$ in terms of the properties of the scalar potential $V(\phi)$. Indeed, this proved to be embarrassing, at least initially, since the Coleman-Weinberg potential, the first potential studied in the context of the new inflation scenario, results in $\delta_H \sim 10^2$ (Guth & Pi 1982) some six orders of magnitude too large. But this does not seem to be an insurmountable difficulty; as was mentioned above, chaotic inflation works, with a sufficiently small self-coupling. Thus inflation at present appears to be a plausible solution to the problem of providing reasonable cosmological initial conditions (although it sheds no light at all on the fundamental question why the cosmological constant is so small now). Many variations of the basic idea of inflation have been worked out

# Many Inflation Models

following Andrei Linde's classification

## HOW INFLATION BEGINS

Old Inflation — $T_{\text{initial}}$ high, $\phi_{\text{in}} \approx 0$ is false vacuum until phase transition
Ends by bubble creation; Reheat by bubble collisions

New Inflation — Slow roll down $V(\phi)$, no phase transition

Chaotic Inflation — Similar to New Inflation, but $\phi_{\text{in}}$ essentially arbitrary:
any region with $\frac{1}{2}\dot{\phi}^2 + \frac{1}{2}(\partial_i \phi)^2 \lesssim V(\phi)$ inflates

Extended Inflation — Like Old Inflation, but slower (e.g., power $a \propto t^p$),
so phase transition can finish

## POTENTIAL $V(\phi)$ DURING INFLATION

Chaotic typically $V(\phi) = \Lambda \phi^n$, can also use $V = V_0 e^{\alpha\phi}$, etc.
$$\Rightarrow a \propto t^p, \quad p = 16\pi/\alpha^2 \gg 1$$

## HOW INFLATION ENDS

First-order phase transition — e.g., Old or Extended inflation
Faster rolling $\rightarrow$ oscillation — e.g., Chaotic $V(\phi)^2 \Lambda \phi^n$
"Waterfall" — rapid roll of $\sigma$ triggered by slow roll of $\phi$

## (RE)HEATING

Decay of inflatons
"Preheating" by parametric resonance, then decay

## BEFORE INFLATION?

Eternal Inflation? Can be caused by
- Quantum $\delta\phi \sim H/2\pi >$ rolling $\Delta\phi = \phi\Delta t = \phi H^{-1} \approx V'/V$
- Monopoles or other topological defects

# Comparing different inflationary models:

- **Chaotic inflation** can start in the smallest domain of size $10^{-33}$ cm with total mass $\sim M_p$ (less than a milligram) and entropy $O(1)$
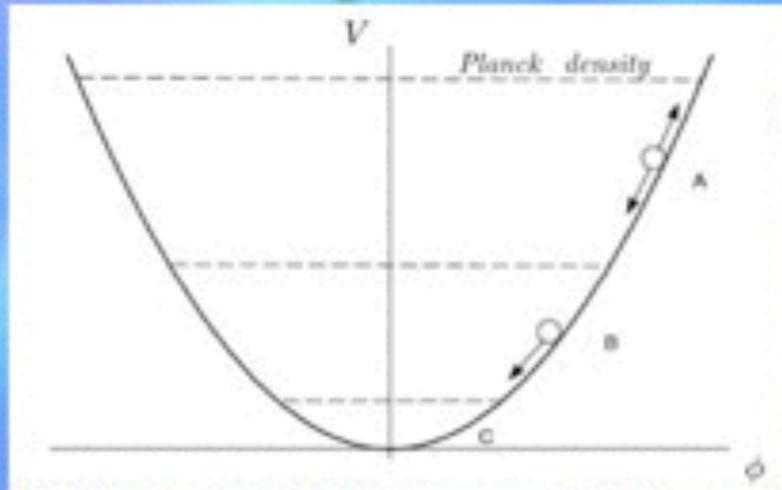
Solves flatness, mass and entropy problem

- **New inflation** can start only in a domain with mass 6 orders of magnitude greater than $M_p$ and entropy greater than $10^9$
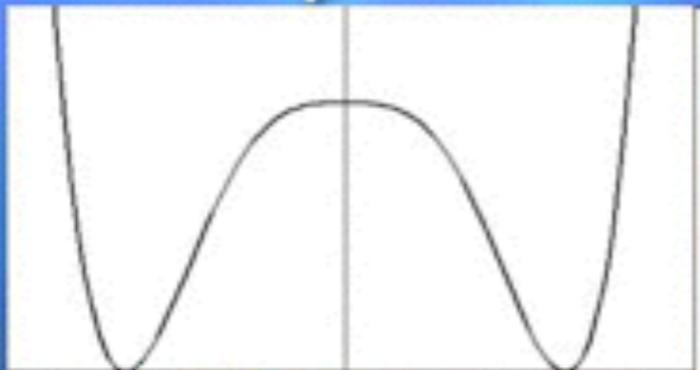
Not very good with solving flatness, mass and entropy problem

- **Cyclic inflation** can occur only in the domain of size greater than the size of the observable part of the universe, with mass $> 10^{55}$ g and entropy $> 10^{87}$
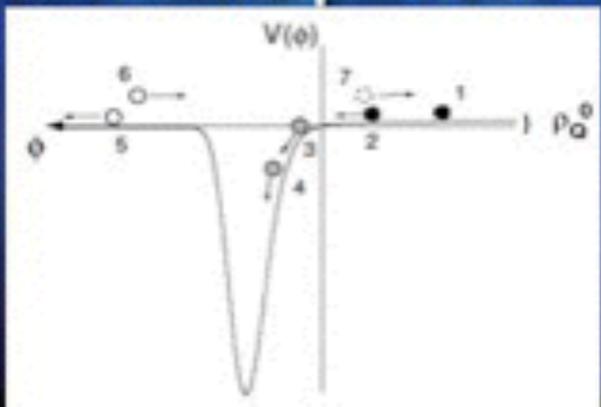
Does not solve flatness, mass and entropy problem

Andrei Linde

# Is the simplest chaotic inflation natural?

Andrei Linde

- Often repeated (but incorrect) argument:

$$V = V_0 + \frac{m^2}{2}\phi^2 + \frac{\lambda}{4}\phi^4 + \sum C_n \frac{\phi^n}{M_p^n}$$

Thus one could expect that the theory is ill-defined at $\phi > M_p$

However, quantum corrections are in fact proportional to

$$\left(\frac{V}{M_p^4}\right)^n \qquad \text{and to} \qquad \left(\frac{m^2(\phi)}{M_p^2}\right)^n$$

These terms are harmless for sub-Planckian masses and densities, even if the scalar field itself is very large.

# Inflaton Theory in More Detail

The action of <span style="color:blue">gravity</span> + scalar inflaton field is

<span style="color:red">Lagrangian for Scalar Field $\phi$</span>

$$S = -\frac{c^3}{16\pi G} \int \mathrm{d}^4 x \sqrt{-\det g_{mn}}\, R$$

$$+ \int \mathrm{d}^4 x \sqrt{-\det g_{mn}}\, \hbar \left( \frac{1}{2} \frac{\partial \phi}{\partial x^i} \frac{\partial \phi}{\partial x^k} g^{ik} - V[\phi] \right)$$

The simplest $V$ is just quadratic $\quad V[\phi] = \frac{1}{2} \frac{m^2 c^2}{\hbar^2} \phi^2$

which just gives the inflaton field a mass $m$.

The Einstein equation $G_{ij} = -(8\pi G/c^4)\, T_{ij}$ follows by varying the metric $g_{ij}$; the first term in the action $S$ leads to the Einstein tensor $G_{ij}$ and the second term leads to the energy-momentum tensor $T_{ij}$.

The energy–momentum tensor is given by

$$T_{ik} = \hbar c \left( \frac{\partial \phi}{\partial x^i} \frac{\partial \phi}{\partial x^k} - g_{ik} \left( \frac{1}{2} g^{lm} \frac{\partial \phi}{\partial x^l} \frac{\partial \phi}{\partial x^m} - V[\phi] \right) \right)$$

which implies that the energy density and pressure are given by

and

$$\varepsilon = \hbar c \left( V + \frac{1}{2c^2} \dot{\phi}^2 + \frac{1}{2} \frac{1}{a^2[t]} (\nabla \phi)^2 \right)$$

$$p = \hbar c \left( -V + \frac{1}{2c^2} \dot{\phi}^2 - \frac{1}{6} \frac{1}{a^2[t]} (\nabla \phi)^2 \right) .$$

Thus a scalar field with a nearly constant potential V corresponds to

$$\varrho c^2 = \varepsilon = -p \ (= \hbar c V[\phi]).$$

Since w = p/ε = -1, this is effectively a cosmological constant.  More generally, a scalar field that is not at the minimum of its potential generates generates "*dark energy*".

The field equation for the inflaton in expanding space is

$$\frac{\partial^2 \phi}{c^2 \partial t^2} - \frac{1}{a^2}\nabla^2\phi + \frac{3\dot{a}\dot{\phi}}{c^2 a} + \frac{dV}{d\phi} \quad (+3H\Gamma\dot{\phi}^2) = 0 \ .$$

This becomes the following equation if the spatial variations of $\phi$ (and the last term, which allows the inflaton to decay into other fields at the end of inflation, thus *reheating* the universe) can be neglected

$$\ddot{\phi} + 3H[t]\dot{\phi} = -c^2 \frac{dV[\phi]}{d\phi} \ .$$

This equation must be solved along with the Einstein equations:

$$H^2 = \frac{8\pi G}{3}\frac{\hbar}{c}\left(V + \frac{1}{2c^2}\dot{\phi}^2\right) \quad \text{and } \dot{H} = -4\pi\frac{\hbar G}{c^3}\dot{\phi}^2$$

With a suitably chosen potential V, the inflaton will quickly reach its ground state and inflation will end.

The last equation leads to

$$H' = \frac{\mathrm{d}H[\phi]}{\mathrm{d}\phi} = -4\pi\frac{\hbar G}{c^3}\dot{\phi}$$

which allows us to write the Friedmann equation as

$$\left(\frac{\mathrm{d}H}{\mathrm{d}\phi}\right)^2 = 12\pi\frac{\hbar G}{c^3}H^2 - 32\pi^2\frac{\hbar^2 G^2}{c^4}V[\phi] .$$

When the inflaton is rolling slowly, the evolution of the inflaton is governed by the "slow roll" equations

$$\dot{\phi} = -\frac{c^2}{3H}\frac{\mathrm{d}V}{\mathrm{d}\phi} , \quad H^2 = \frac{8\pi\hbar G}{3c}V .$$

Then the number N of e-folds of the scale factor a is given by

$$N = \ln\frac{a}{a_1} = \int_{t_1}^{t} H\,\mathrm{d}t = \int_{\phi_1}^{\phi} \mathrm{d}\phi\frac{H}{\dot{\phi}} = 4\pi\frac{\hbar G}{c^3}\int_{\phi}^{\phi_1}\mathrm{d}\phi\frac{H}{H'} \approx 8\pi\frac{\hbar G}{c^3}\int_{\phi}^{\phi_1}\mathrm{d}\phi\frac{V}{V'} .$$

The last approximate equality used the slow roll approximation.

# Inflationary Models in More Detail

**Prototype model for "new inflation"**

Assume that in some region, $\phi \approx 0$. The transition $\phi \to \sigma$ is governed by the Klein-Gordon equation



$$\ddot{\phi} + 3H\dot{\phi} + \Gamma\dot{\phi} = -V'$$
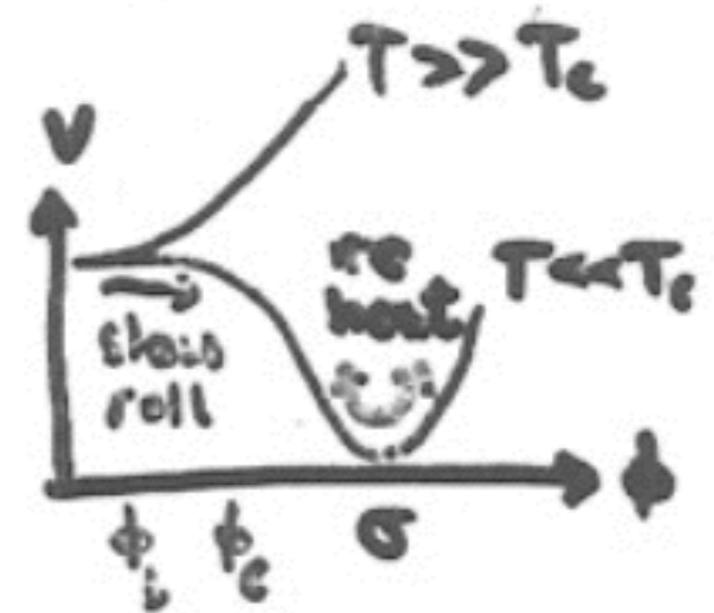
where the Hubble parameter $H^2 = (8\pi/3M_{\mathrm{Pl}}^2)\rho$, energy density $\rho = \dot{\phi}^2/2 + V(\phi) + \rho_{rad}$, pressure $p = \dot{\phi}^2/2 - V(\phi) + \rho_{rad}/3$, $\Gamma$ represents couplings to other fields that will be important at the end of inflation, and we have neglected spatial derivatives of $\phi$ (we just need to start in a region where this is ok).

If $V(\phi)$ is large and flat enough, $V(\phi)$ will be $>> \dot{\phi}^2$ and $\rho_{rad}$, $\ddot{\phi} << 3H\dot{\phi}$, and

$$\dot{R}/R = H \approx (8\pi V/3M_{\mathrm{Pl}}^2)^{1/2}$$

implies that $R = R_0 e^N$, where $N = \int H\, dt$.

For example, take $\sigma = 10^{14}$ GeV, $T_i = \sigma$, $R_i = H^{-1} = M_{\rm Pl}/\sigma^2 = 10^{19} GeV/(10^{14} {\rm GeV})^2$ $= R_i = 10^{-9}$ GeV$^{-1}$ $= 10^{-23}$ cm, which implies that the initial entropy

$$S_i = (R_i T_i)^3 = (M_{\rm Pl}/\sigma^2)^3 \sigma^3 = 10^{15}.$$

Then after inflation ends and the universe is at reheat temperature $T_{RH}$, the final entropy is $S_f = (R_f T_{RH})^3 = (e^N R_i T_{RH})^3$. For example, if $N = 100$ e-folds, $e^{300} = 10^{130}$ and this gives $S_f = 10^{145}$. Requiring this to be at least as big as the present entropy of the observable universe, which is about $10^{88}$,

$$S_f = e^{3N} M_{\rm Pl}/\sigma \geq 10^{88} \quad \text{implies} \quad 3N \geq \ln 10^{88} + \ln(\sigma/M_{\rm Pl}),$$

or $N \geq 58 + \ln(\sigma/10^{15} {\rm GeV})$. This solves the Horizon problem, the Flatness problem ($k/R^2$ decreases by $e^{-2N}$), and the Relics problem (density $\sim R^{-3}$ decreases by $e^{-3N}$).

A is from
$\sim 1 - 10^4$ Mpc
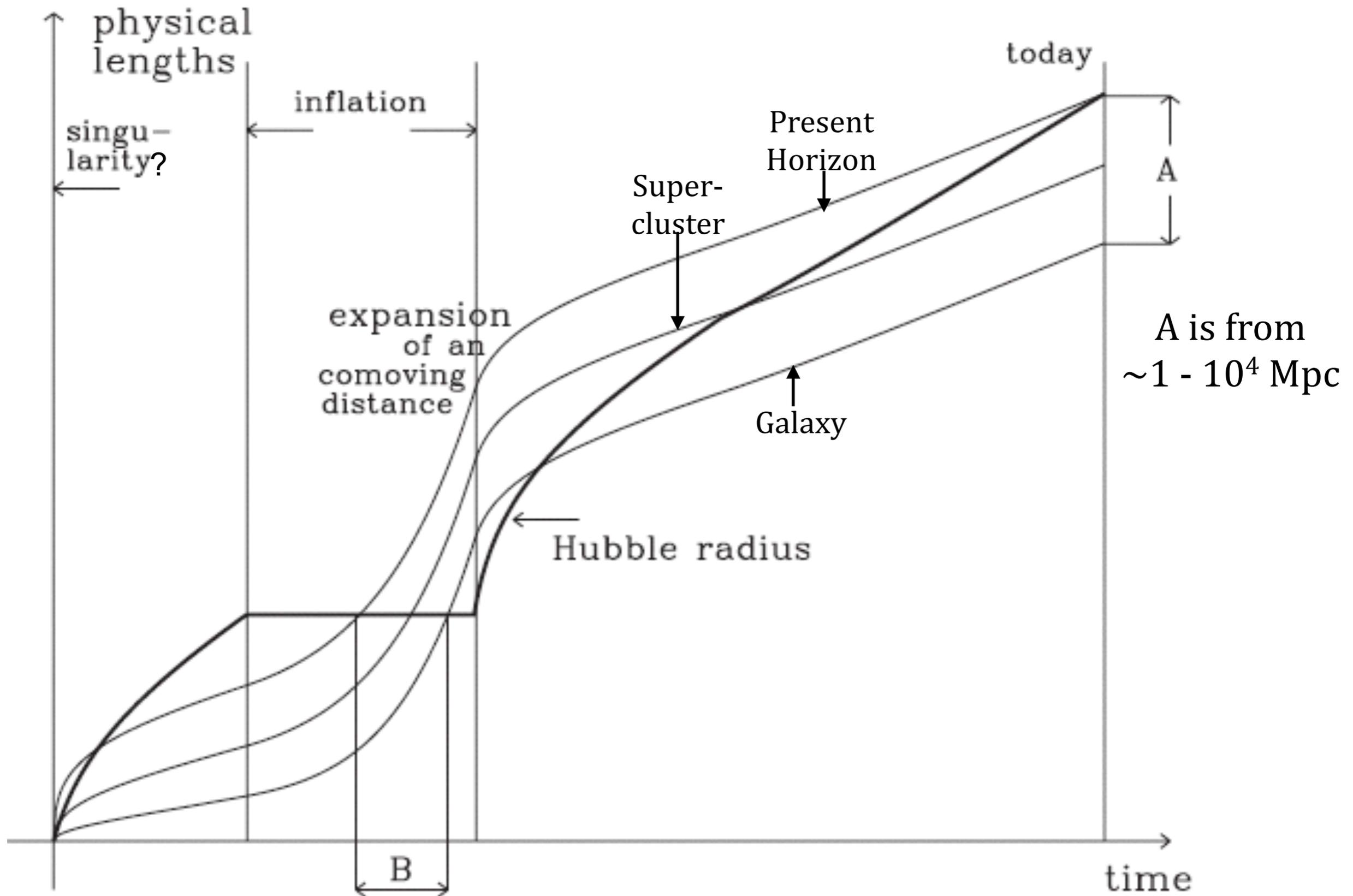
The evolution of the scales of perturbations. The larger scales overtake the Hubble radius at an early time and fall below it again later. They measure the inflation at an earlier time than do the smaller scales, which overtake the Hubble radius during inflation later and fall below it again earlier. The region *A* of scales that are accessible to evaluation today corresponds to a time span *B* of the inflation and related values of the inflaton field; for this time span, we can tell something – at least in principle – about the potential of the inflaton.

# Fluctuations in Inflation



**LOFI**: The last scales to cross outside the horizon in Inflation are the first to cross inside in the subsequent Friedmann-Robertson-Walker phase. If the present cosmic horizon, which encompasses a mass of about $10^{22} M_\odot$, crossed outside the de Sitter event horizon at (say) 60 e-folds before the end of Inflation, galaxies (about $10^{12} M_\odot$ in mass) crossed about 52 e-folds before the end, and any horizon mass $M_H$ crossed at $N = 60 - (1/3)\ln(M_H/10^{22} M_\odot)$.

There are quantum fluctuations in a de Sitter universe corresponding to the Hawking radiation temperature $T_H = H/2\pi$, so $< (\Delta\phi)^2 >= (H/2\pi)^2$. (Guth & Pi 1982 got the same answer using the de Sitter propagator.) This leads to density fluctuations $\delta\rho = V'\Delta\phi = -3H\dot{\phi}\Delta\phi$, where in the last step we used the slow-roll approximation $V' \equiv dV/d\phi = -3H\dot{\phi}$.
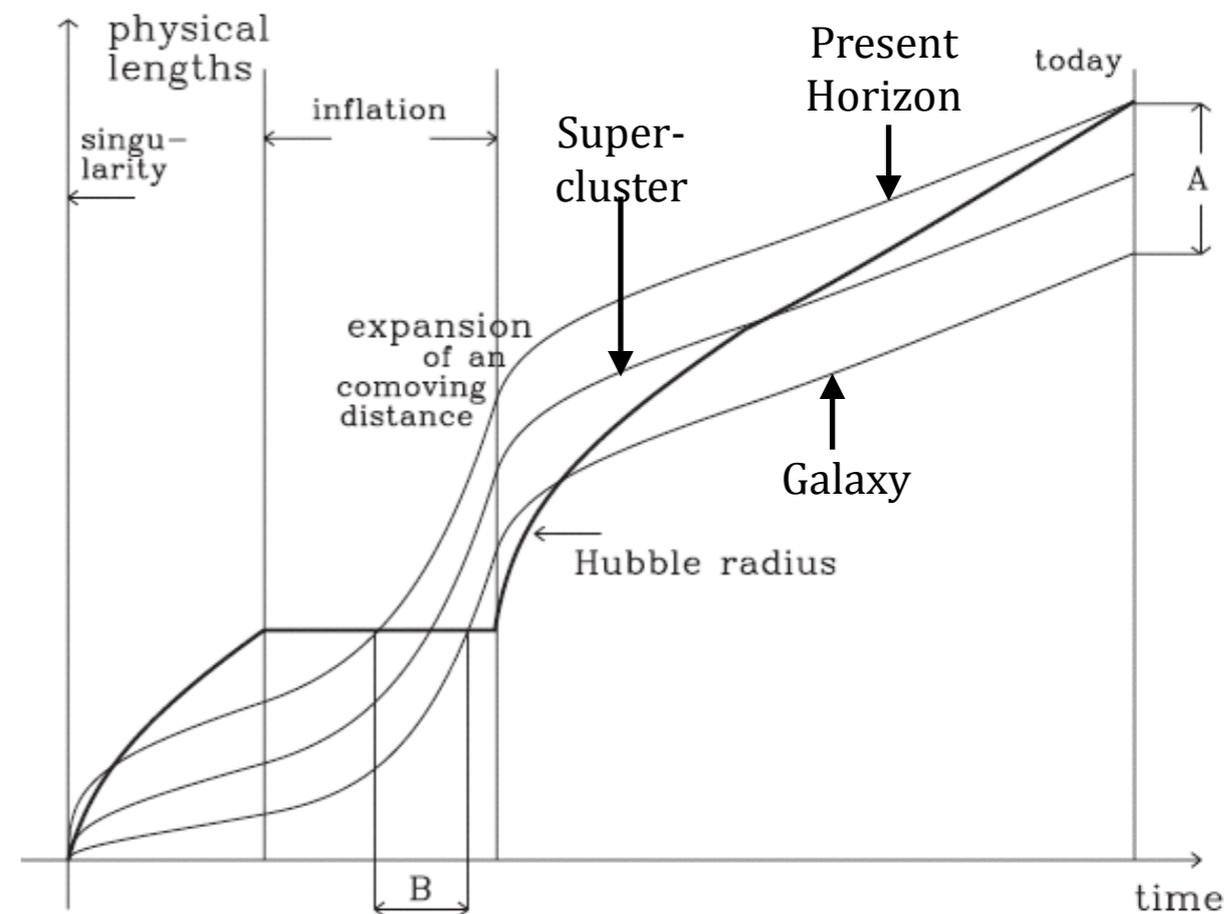
Thus in Inflation, $\rho = \dot{\phi}^2/2 + V(\phi) + \rho_{rad}$ and $p = \dot{\phi}^2/2 - V(\phi) + \rho_{rad}/3$, and we expect that $\rho_{rad}$ will be negligible in Inflation. Note then that $\rho + p = \dot{\phi}^2$.

Bardeen's gauge invariant parameter $\zeta = \delta\rho/(\rho + p)$ is constant outside the horizon, so we equate the value of $\zeta$ as a given scale goes outside the horizon in Inflation with the value when it comes back inside:

$$\text{Inflation}: \quad \zeta = \frac{\delta\rho}{\rho + p} = \frac{\delta\rho}{\dot{\phi}^2} = \frac{V'\Delta\phi}{\dot{\phi}^2} = \frac{-3H\dot{\phi}\Delta\phi}{\dot{\phi}^2} = \frac{-3H^2}{2\pi\dot{\phi}^2}$$

$$\text{FRW}: \quad \zeta = \frac{\delta\rho}{\rho + p} = \frac{\delta\rho}{(4/3)\rho}$$

When the scale comes back inside the horizon in the Friedmann-Robertson-Walker phase, $\dot{\phi}^2$ and $V(\phi)$ are negligible and $\rho + p = (4/3)\rho$. Then

$$\frac{\delta\rho}{\rho} = \frac{4}{3}\zeta = \frac{-2H^2}{\pi\dot{\phi}^2} = \frac{6H^3}{\pi V'(\phi)} = \frac{6(8\pi V/3M_{\rm Pl}^2)^{3/2}}{\pi V'(\phi)},$$

where the next-to-last expression assumes that the slow-roll equation $V' \equiv dV/d\phi = -3H\dot{\phi}$ applies when scale crossed outside the horizon in Inflation.

# Eternal Inflation

Vilenkin (1983) and Linde (1986, 1990) pointed out that if one extrapolates inflation backward to try to imagine what might have preceeded it, in many versions of inflation the answer is "eternal inflation": in most of the volume of the universe inflation is still happening, and our part of the expanding universe (a region encompassing far more than our entire cosmic horizon) arose from a tiny part of such a region. To see how eternal inflation works, consider the simple chaotic model with $V(\phi) = (m^2/2)\phi^2$. During the de Sitter Hubble time $H^{-1}$, where as usual $H^2 = (8\pi G/3)V$, the slow rolling of $\phi$ down the potential will reduce it by

$$\Delta\phi = \dot\phi\Delta t = -\frac{V'}{3H}\Delta t = \frac{m^2_{pl}}{4\pi\phi}. \tag{1.7}$$

Here $m_{Pl}$ is the Planck mass $(m_{Planck} = 1/G^{1/2})$. But there will also be quantum fluctuations that will change $\phi$ up or down by

$$\delta\phi = \frac{H}{2\pi} = \frac{m\phi}{\sqrt{3\pi}\,m_{Pl}} \tag{1.8}$$

These will be equal for $\phi_* = m^{3/2}_{pl}/2m^{1/2}$, $V(\phi_*) = (m/8m_{Pl})m^4_{Pl}$. If $\phi \gtrsim \phi_*$, *positive quantum fluctuations dominate* the evolution: after $\Delta t \sim H^{-1}$, an initial region becomes $\sim e^3$ regions of size $\sim H^{-1}$, in half of which $\phi$ increases to $\phi + \delta\phi$. Since $H \propto \phi$, this drives inflation faster in these regions.

# Inflation as a theory of a harmonic oscillator

$$V(\phi) = \frac{m^2}{2}\phi^2$$



SPACE-TIME FOAM

PLANCK DENSITY

LARGE QUANTUM FLUCTUATIONS

POTENTIAL ENERGY

INFLATION

SMALL QUANTUM FLUCTUATIONS

HEATING OF UNIVERSE

SCALAR FIELD
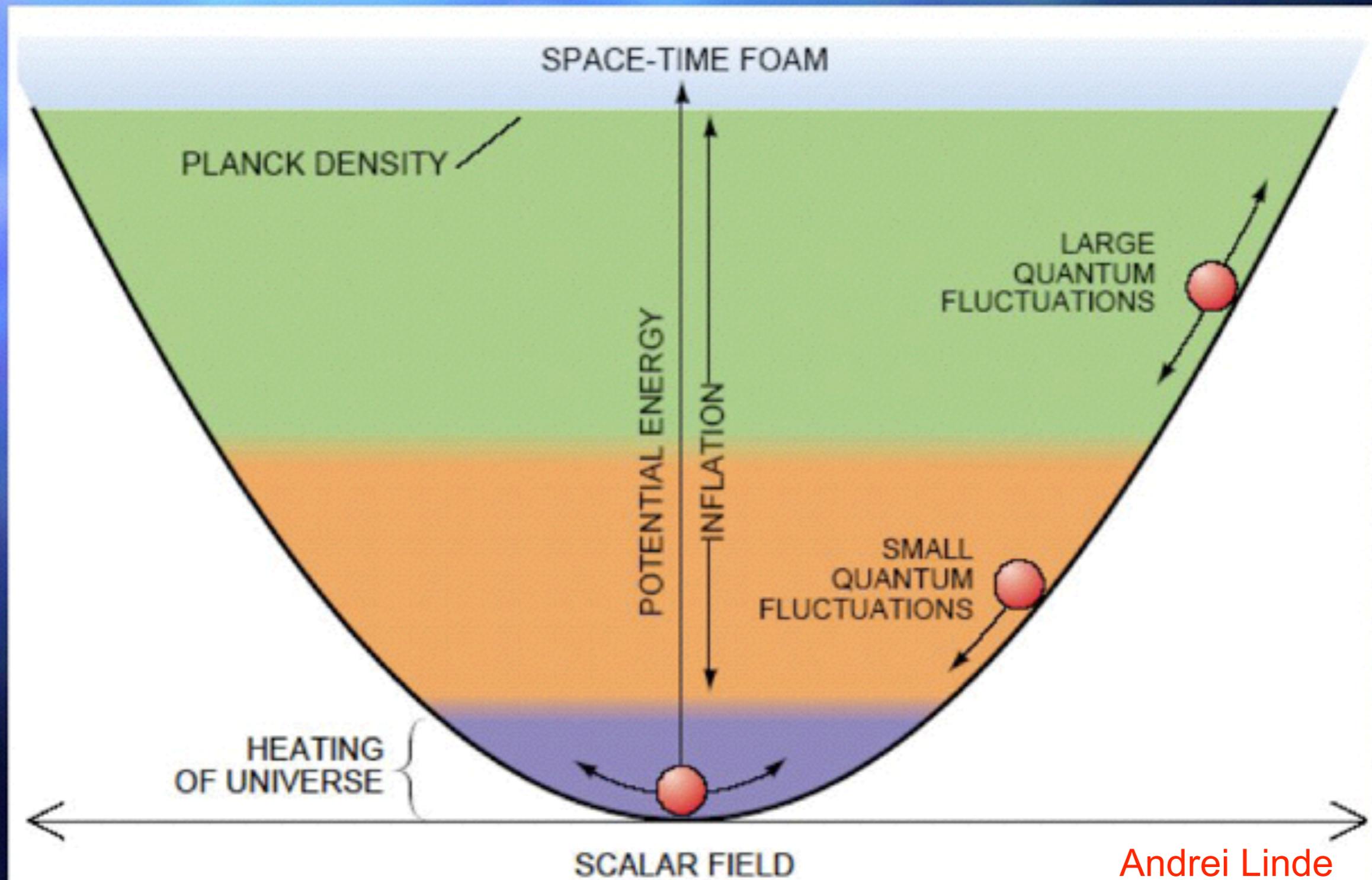
Andrei Linde

# Landscape of eternal inflation



Different colors represent
different vacuua
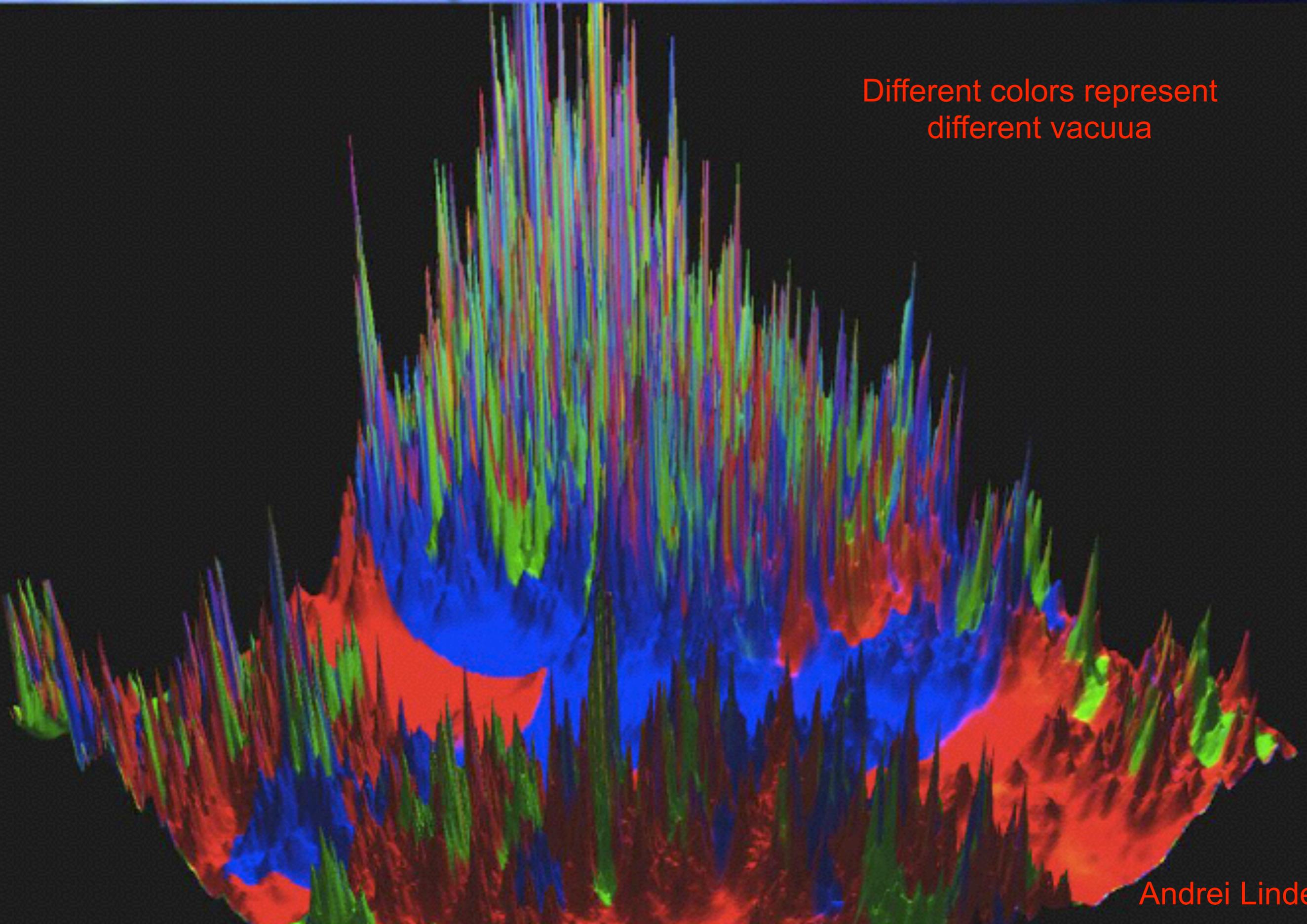
Andrei Linde

# THE COSMIC LAS VEGAS

**Coins constantly flip. Heads, and the coin is twice the size and there are two of them. Tails, and a coin is half the size.**

**Consider a coin that has a run of tails. It becomes so small it can pass through the grating on the floor.**

**At the instant it passes through the floor, it exits eternity.**

**Time begins with a Big Bang, and it becomes a universe and starts evolving.**

The Multiverse

SPACE-TIME FOAM

PLANCK DENSITY

LARGE QUANTUM FLUCTUATIONS

POTENTIAL ENERGY

INFLATION

SMALL QUANTUM FLUCTUATIONS

**"grating"**

HEATING OF UNIVERSE

SCALAR FIELD

Andrei Linde
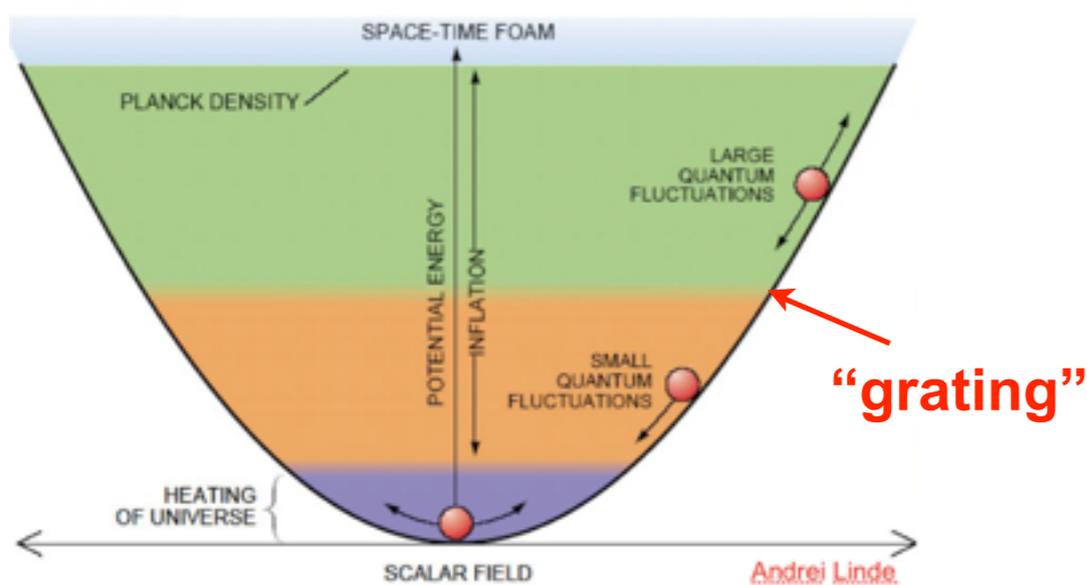
OUR COSMIC BUBBLE IN ETERNAL INFLATION

BUBBLE UNIVERSES IN ETERNAL INFLATION

Center of Coordinates (for orientation)
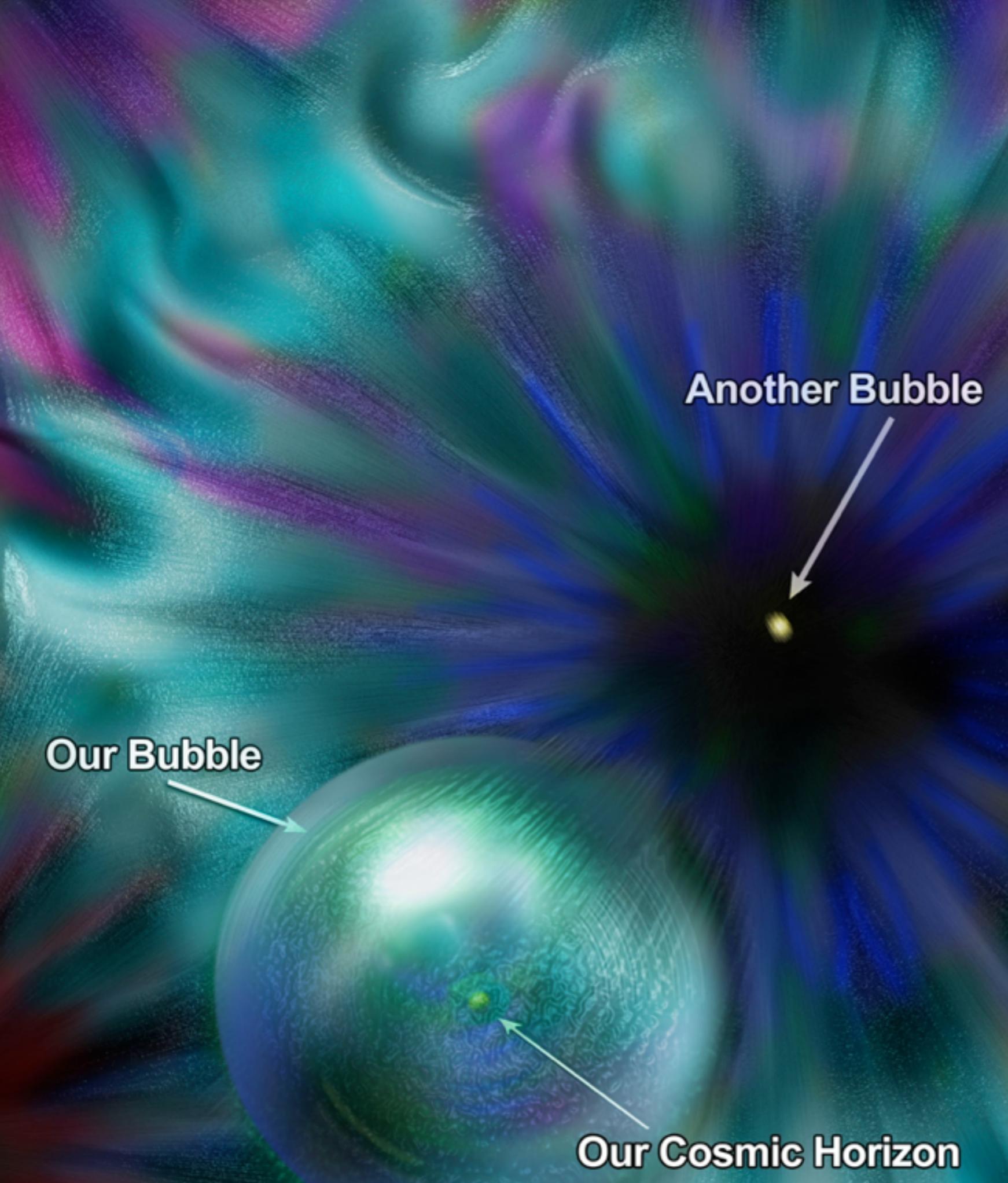
Bubble Collision
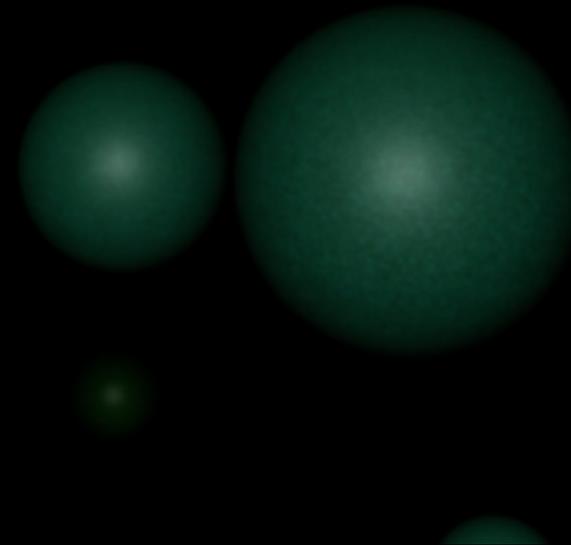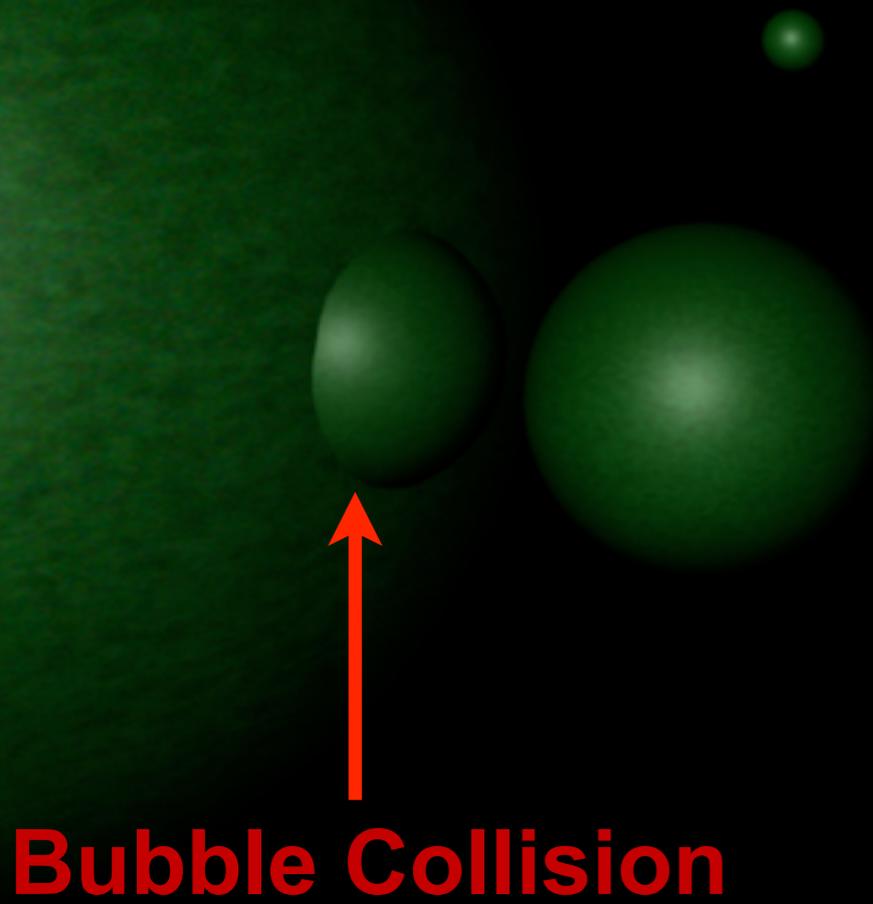
Nancy Abrams
Anthony Aguirre
Nina McCurdy
Joel Primack

**Expanding Bubbles Getting Dimmer Are Receding**

**BUBBLE UNIVERSES IN ETERNAL INFLATION**
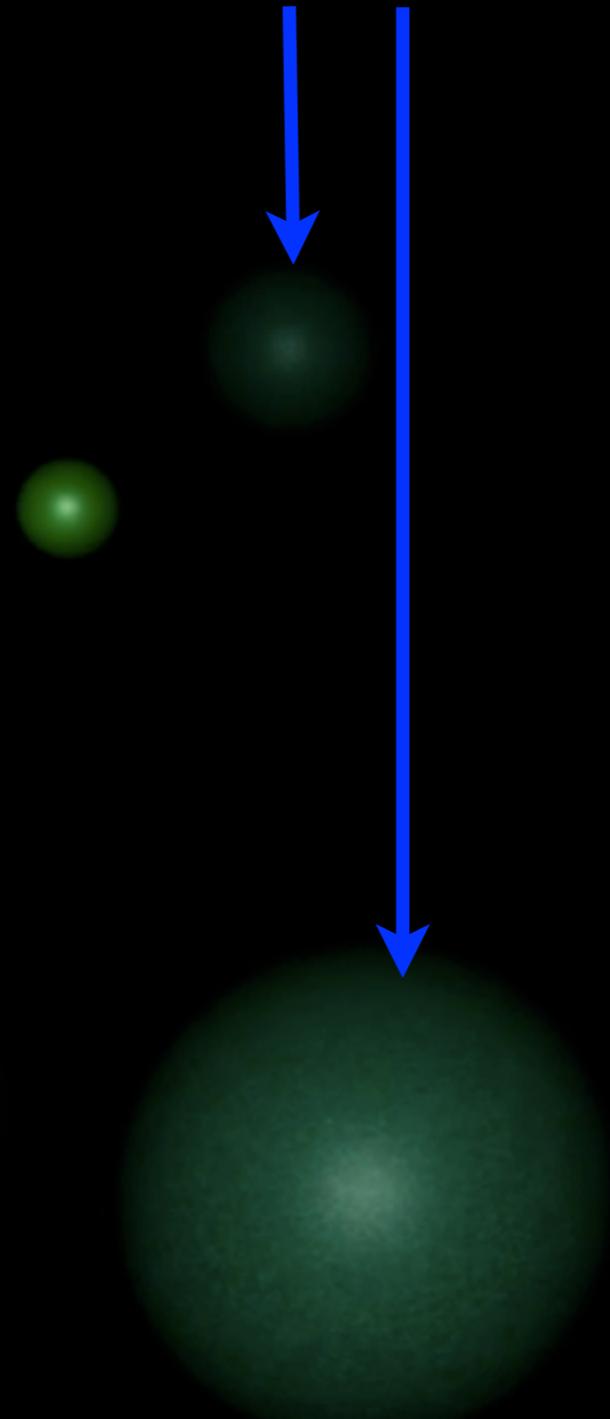
# Supersymmetric Inflation

When Pagels and I (1982) first suggested that the lightest supersymmetric partner particle (LSP), stable because of R-parity, might be the dark matter particle, that particle was the gravitino in the early version of supersymmetry then in fashion. Weinberg (1982) immediately pointed out that if the gravitino were not the LSP, it could be a source of real trouble because of its long lifetime $\sim M_{\mathrm{Pl}}^2/m_{3/2}^3 \sim (m_{3/2}/\mathrm{TeV})^{-3} 10^3$ s, a consequence of its gravitational-strength coupling to other fields. Subsequently, it was realized that supersymmetric theories can naturally solve the gauge hierarchy problem, explaining why the electroweak scale $M_{\mathrm{EW}} \sim 10^2$ GeV is so much smaller than the GUT or Planck scales. In this version of supersymmetry, which has now become the standard one, the gravitino mass will typically be $m_{3/2} \sim$ TeV; and the late decay of even a relatively small number of such massive particles can wreck BBN and/or the thermal spectrum of the CBR. The only way to prevent this is to make sure that the reheating temperature after inflation is sufficiently low: $T_{\mathrm{RH}} \lesssim 2 \times 10^9$ GeV (for $m_{3/2} = \mathrm{TeV}$) (Ellis, Kim, & Nanopoulos 1984, Ellis et al. 1992).

# Inflation Summary

The key features of all inflation scenarios are a period of superluminal expansion, followed by ("re-")heating which converts the energy stored in the inflaton field (for example) into the thermal energy of the hot big bang.

Inflation is *generic*: it fits into many versions of particle physics, and it can even be made rather natural in modern supersymmetric theories as we have seen. The simplest models have inflated away all relics of any pre-inflationary era and result in a flat universe after inflation, i.e., $\Omega = 1$ (or more generally $\Omega_0 + \Omega_\Lambda = 1$). Inflation also produces scalar (density) fluctuations that have a primordial spectrum

$$\left(\frac{\delta\rho}{\rho}\right)^2 \sim \left(\frac{V^{3/2}}{m_{Pl}^3 V'}\right)^2 \propto k^{n_p}, \tag{1.12}$$

where $V$ is the inflaton potential and $n_p$ is the primordial spectral index, which is expected to be near unity (near-Zel'dovich spectrum). Inflation also produces tensor (gravity wave) fluctuations, with spectrum

$$P_t(k) \sim \left(\frac{V}{m_{Pl}}\right)^2 \propto k^{n_t}, \tag{1.13}$$

where the tensor spectral index $n_t \approx (1 - n_p)$ in many models.

The quantity $(1 - n_p)$ is often called the "tilt" of the spectrum; the larger the tilt, the more fluctuations on small spatial scales (corresponding to large $k$) are suppressed compared to those on larger scales. The scalar and tensor waves are generated by independent quantum fluctuations during inflation, and so their contributions to the CMB temperature fluctuations add in quadrature. The ratio of these contributions to the quadrupole anisotropy amplitude $Q$ is often called $T/S \equiv Q_t^2/Q_s^2$; thus the primordial scalar fluctuation power is decreased by the ratio $1/(1+T/S)$ for the same COBE normalization, compared to the situation with no gravity waves ($T = 0$). In power-law inflation, $T/S = 7(1 - n_p)$. This is an approximate equality in other popular inflation models such as chaotic inflation with $V(\phi) = m^2\phi^2$ or $\lambda\phi^4$. But note that the tensor wave amplitude is just the inflaton potential during inflation divided by the Planck mass, so the gravity wave contribution is negligible in theories like the supersymmetric model discussed above in which inflation occurs at an energy scale far below $m_{Pl}$. Because gravity waves just redshift after they come inside the horizon, the tensor contributions to CMB anisotropies corresponding to angular wavenumbers $\ell \gg 20$, which came inside the horizon long ago, are strongly suppressed compared to those of scalar fluctuations.

Joel Primack, in *Formation of Structure in the Universe*, ed. Dekel & Ostriker (Cambridge Univ Press, 1999)

# Basic Predictions of Inflation

1. **Flat universe.** This is perhaps the most fundamental prediction of inflation. Through the Friedmann equation it implies that the total energy density is always equal to the critical energy density; it does not however predict the form (or forms) that the critical density takes on today or at any earlier or later epoch.

2. **Nearly scale-invariant spectrum of Gaussian density perturbations.** These density perturbations (scalar metric perturbations) arise from quantum-mechanical fuctuations in the field that drives inflation; they begin on very tiny scales (of the order of $10^{-23}$ cm, and are stretched to astrophysical size by the tremendous growth of the scale factor during inflation (factor of $e^{60}$ or greater). Scale invariant refers to the fact that the fuctuations in the gravitational potential are independent of length scale; or equivalently that the horizon-crossing amplitudes of the density perturbations are independent of length scale. While the shape of the spectrum of density perturbations is common to all models, the overall amplitude is model dependent. Achieving density perturbations that are consistent with the observed anisotropy of the CBR and large enough to produce the structure seen in the Universe today requires a horizon crossing amplitude of around $2 \times 10^{-5}$.

3. **Nearly scale-invariant spectrum of gravitational waves**, from quantum-mechanical fluctuations in the metric itself .  These can be detected as CMB "B-mode" polarization, or using special gravity wave detectors such as LIGO and LISA.

# Density Fluctuations from Inflation

The relationship between the inflationary potential and the power spectrum of density perturbations today $(P(k) \equiv \langle |\delta_k|^2 \rangle)$ is given by

Transfer function

Power Spectrum $\quad P(k) \;=\; \dfrac{1024\pi^3}{75}\,\dfrac{k}{H_0^4}\,\dfrac{V_*^3}{m_{\mathrm{Pl}}^6 V_*'^2}\left(\dfrac{k}{k_*}\right)^{n-1} T^2(k)$

Tilt $\quad n_{\mathrm{s}} - 1 \;=\; -\dfrac{1}{8\pi}\left(\dfrac{m_{\mathrm{Pl}} V_*'}{V_*}\right)^2 + \dfrac{m_{\mathrm{Pl}}}{4\pi}\left(\dfrac{m_{\mathrm{Pl}} V_*'}{V_*}\right)'$

generally nonzero, $\approx 0.04$ according to WMAP & Planck

Running Tilt $\quad \dfrac{dn}{d\ln k} \;=\; -\dfrac{1}{32\pi^2}\left(\dfrac{m_{\mathrm{Pl}}^3 V_*'''}{V_*}\right)\left(\dfrac{m_{\mathrm{Pl}} V_*'}{V_*}\right)$

$$+\dfrac{1}{8\pi^2}\left(\dfrac{m_{\mathrm{Pl}}^2 V_*''}{V_*}\right)\left(\dfrac{m_{\mathrm{Pl}} V_*'}{V_*}\right)^2 - \dfrac{3}{32\pi^2}\left(m_{\mathrm{Pl}}\dfrac{V_*'}{V_*}\right)^4$$

$$T(q) \;=\; \dfrac{\ln\left(1 + 2.34q\right)/2.34q}{\left[1 + 3.89q + (16.1q)^2 + (5.46q)^3 + (6.71q)^4\right]^{1/4}}, \qquad (4)$$

where $V(\phi)$ is the inflationary potential, prime denotes $d/d\phi$, $V_*$ is the value of the scalar potential when the scale $k_*$ crossed outside the horizon during inflation, $T(k)$ is the transfer function which accounts for the evolution of the mode $k$ from horizon crossing until the present, $q = k/h\Gamma$, and $\Gamma \simeq \Omega_M h$ is the "shape" parameter . The fitting formula (4) isn't accurate enough for precision work; instead, use the website http://camb.info/ .

# Gravity Waves from Inflation

Unlike the scalar perturbations, which must have an amplitude of around $10^{-5}$ to seed structure formation, there is an upper but no lower limit on the amplitude of the tensor perturbations. They can be characterized by their power spectrum today

$$P_T(k) \equiv \langle |h_k|^2 \rangle = \frac{8}{3\pi} \frac{V_*}{m_{\text{Pl}}^4} \left( \frac{k}{k_*} \right)^{n_T - 3} T_T^2(k)$$

$$n_T = -\frac{1}{8\pi} \left( \frac{m_{\text{Pl}} V_*'}{V_*} \right)^2$$

$$\frac{dn_T}{d \ln k} = \frac{1}{32\pi^2} \left( \frac{m_{\text{Pl}}^2 V''}{V} \right) \left( \frac{m_{\text{Pl}} V'}{V} \right)^2 - \frac{1}{32\pi^2} \left( \frac{m_{\text{Pl}} V'}{V} \right)^4 = -n_T[(n-1) - n_T]$$

$$T_T(k) \simeq \left[ 1 + \frac{4}{3} \frac{k}{k_{\text{EQ}}} + \frac{5}{2} \left( \frac{k}{k_{\text{EQ}}} \right)^2 \right]^{1/2}, \tag{11}$$

where $T_T(k)$ is the transfer function for gravity waves and describes the evolution of mode $k$ from horizon crossing until the present, $k_{\text{EQ}} = 6.22 \times 10^{-2}\,\text{Mpc}^{-1}\,(\Omega_M h^2 / \sqrt{g_*/3.36})$ is the scale that crossed the horizon at matter-radiation equality, $\Omega_M$ is the fraction of critical density in matter, and $g_*$ counts the effective number of relativistic degrees of freedom (3.36 for photons and three light neutrino species). The quantity $k^{3/2}|h_k|/\sqrt{2\pi^2}$ corresponds to the dimensionless strain (metric perturbation) on length scale $\lambda = 2\pi/k$.

Root mean square fluctuations in temperature (T) and polarization (E and B modes) of the CMB predicted by inflation.
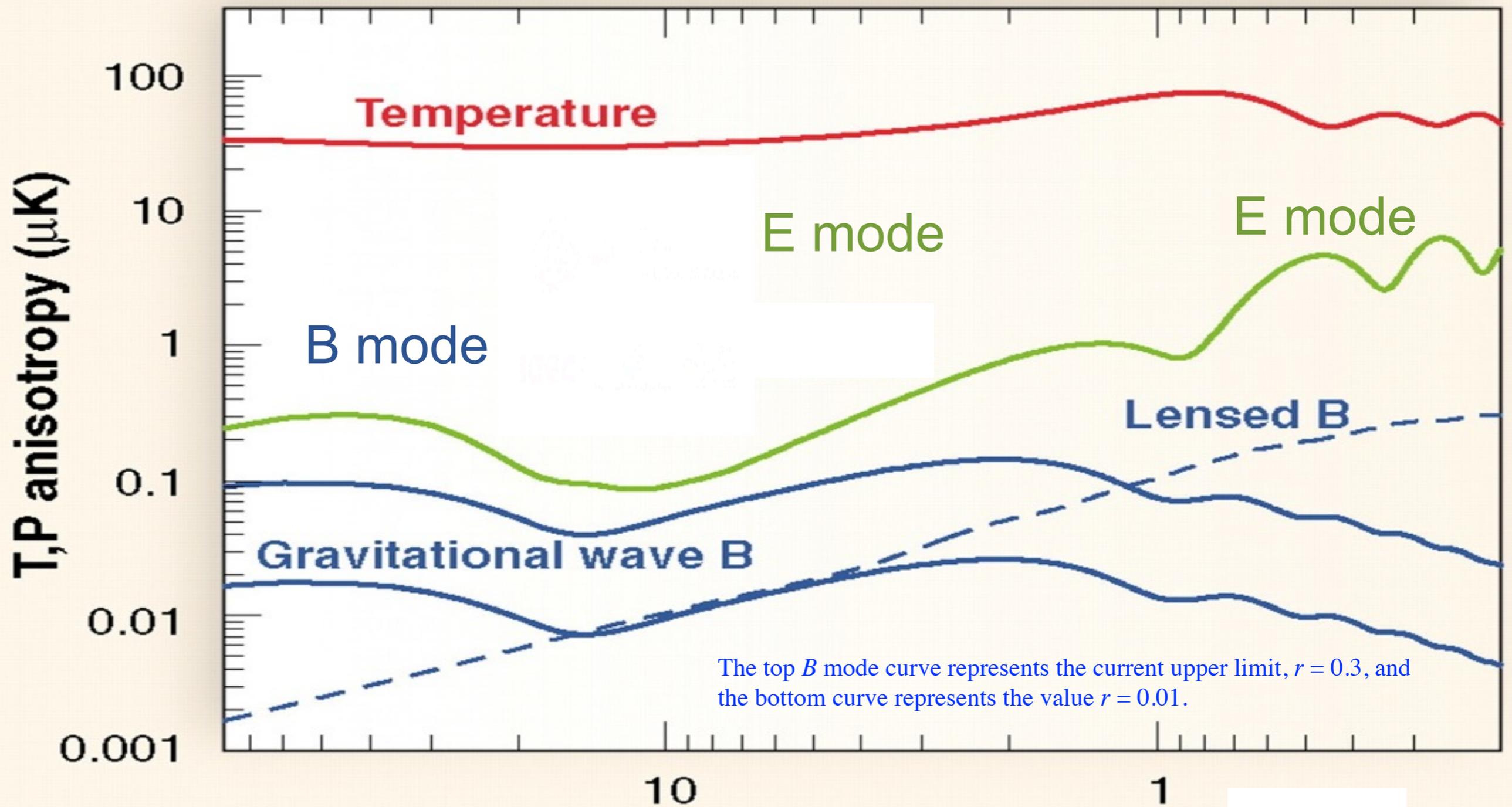
$$V^{1/4} = 1.06 \times 10^{16}\,\text{GeV}\left(\frac{r}{0.01}\right)$$

$$r \equiv \frac{P_t}{P_s}$$

The top B mode curve represents the current upper limit, $r = 0.3$, and the bottom curve represents the value $r = 0.01$.

# **Observational Status of Inflation**

## I. The predictions of inflation are right:

(i) the universe is flat with a critical density

(ii) superhorizon fluctuations

(iii) density perturbation spectrum nearly scale invariant: $P(k) = Ak^n$, $n \approx 1$

(iv) Single slow-roll field models vindicated: Gaussian perturbations, not much running of spectral index


■ If primordial fluctuations are Gaussian distributed, then they are completely characterized by their two-point function $\xi(r)$, or equivalently by the power spectrum. All odd-point functions are zero.

■ If nonGaussian, there is additional info in the higher order correlation functions

■ The lowest order statistic that can differentiate is the 3-point function, or bispectrum in Fourier space: $\langle \Phi(\boldsymbol{k}_1)\Phi(\boldsymbol{k}_2)\Phi(\boldsymbol{k}_3) \rangle = (2\pi)^3 \delta^{(3)}(\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3) B_\Phi(k_1, k_2, k_3)$. Here $B_\Phi(k_1, k_2, k_3) = f_{NL} F(k_1, k_2, k_3)$. The quantity $f_{NL}$ is known as the nonlinearity parameter. Planck data: $f_{NL} = 2.7 \pm 5.8$ -- small!


## II. Data differentiate between models

-- each model makes specific predictions for density perturbations and gravity modes

-- WMAP and Planck rule out many models (see graph on next page, from Planck 2013 Results XXII, http://arxiv.org/abs/1303.5082).

# Observational Status of Inflation
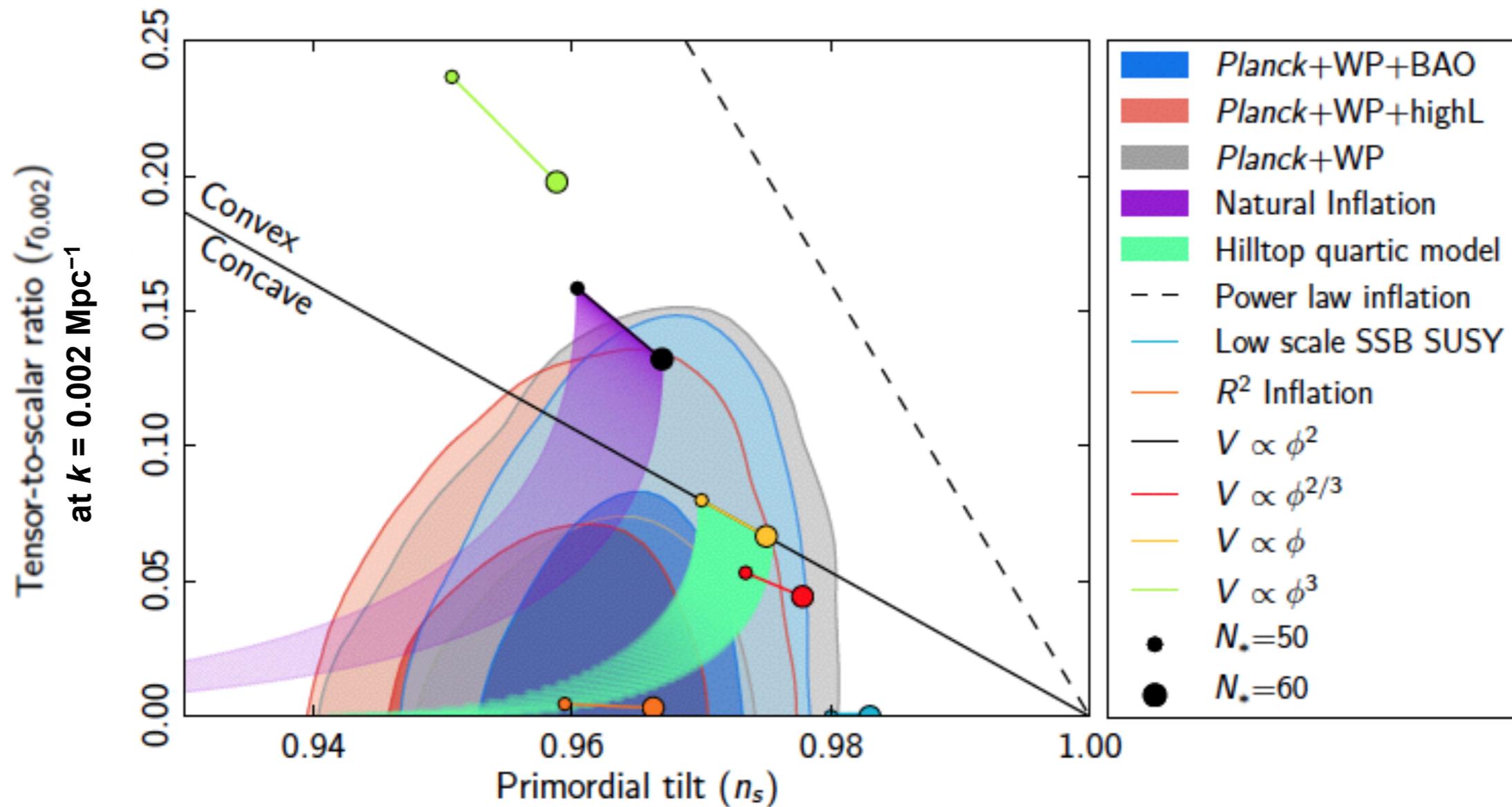
Fig. 1. Marginalized joint 68% and 95% CL regions for $n_s$ and $r_{0.002}$ from *Planck* in combination with other data sets compared to the theoretical predictions of selected inflationary models.

**CONCLUSIONS:** We find that standard slow-roll single field inflation is compatible with the Planck data. Planck in combination with WMAP 9-year large angular scale polarization (WP) yields $\Omega_K$ −0.006 ±0.018 at 95%CL by combining temperature and lensing information (Planck Collaboration XVI, 2013; Planck Collaboration XVII, 2013). The bispectral non-Gaussianity parameter $f_{NL}$ measured by Planck is consistent with zero (Planck Collaboration XXIV, 2013). These results are compatible with zero spatial curvature and a small value of $f_{NL}$, as predicted in the simplest slow-roll inflationary models. Planck+WP data give $n_s$ =0.9603±0.0073 (and $n_s$ = 0.9629±0.0057 when combined with BAO).  The 95% CL bound on the tensor-to-scalar ratio is r 0.002 < 0.12; this implies an upper limit for the inflation energy scale of $1.9 \times 10^{16}$ GeV.

# BICEP2 I: DETECTION OF B-mode POLARIZATION AT DEGREE ANGULAR SCALES
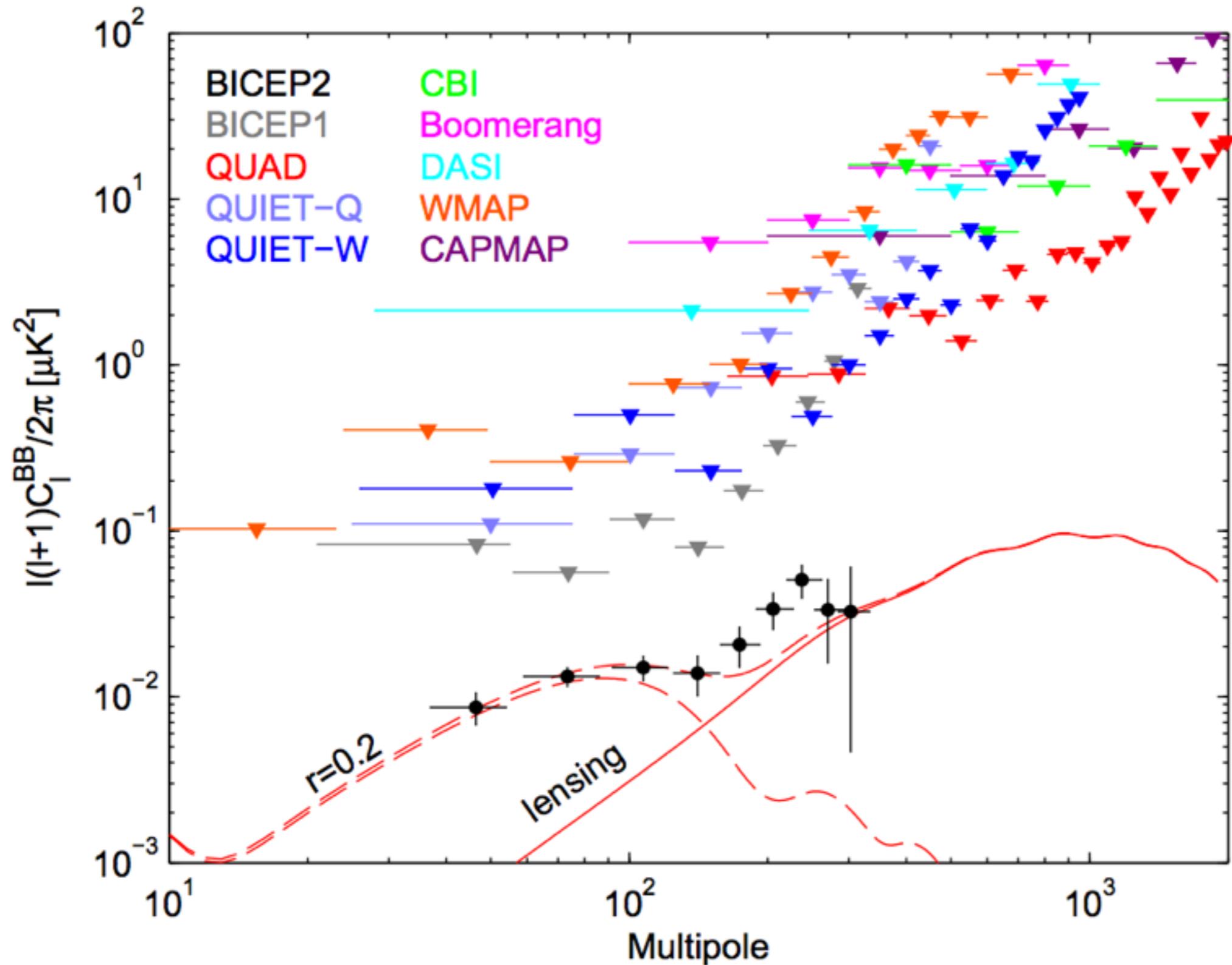


BICEP2 BB auto spectra and 95% upper limits from several previous experiments. The curves show the theory expectations for Tensor/Scalar = r = 0.2 and lensed-ΛCDM.

# BICEP2 I: DETECTION OF B-mode POLARIZATION AT DEGREE ANGULAR SCALES
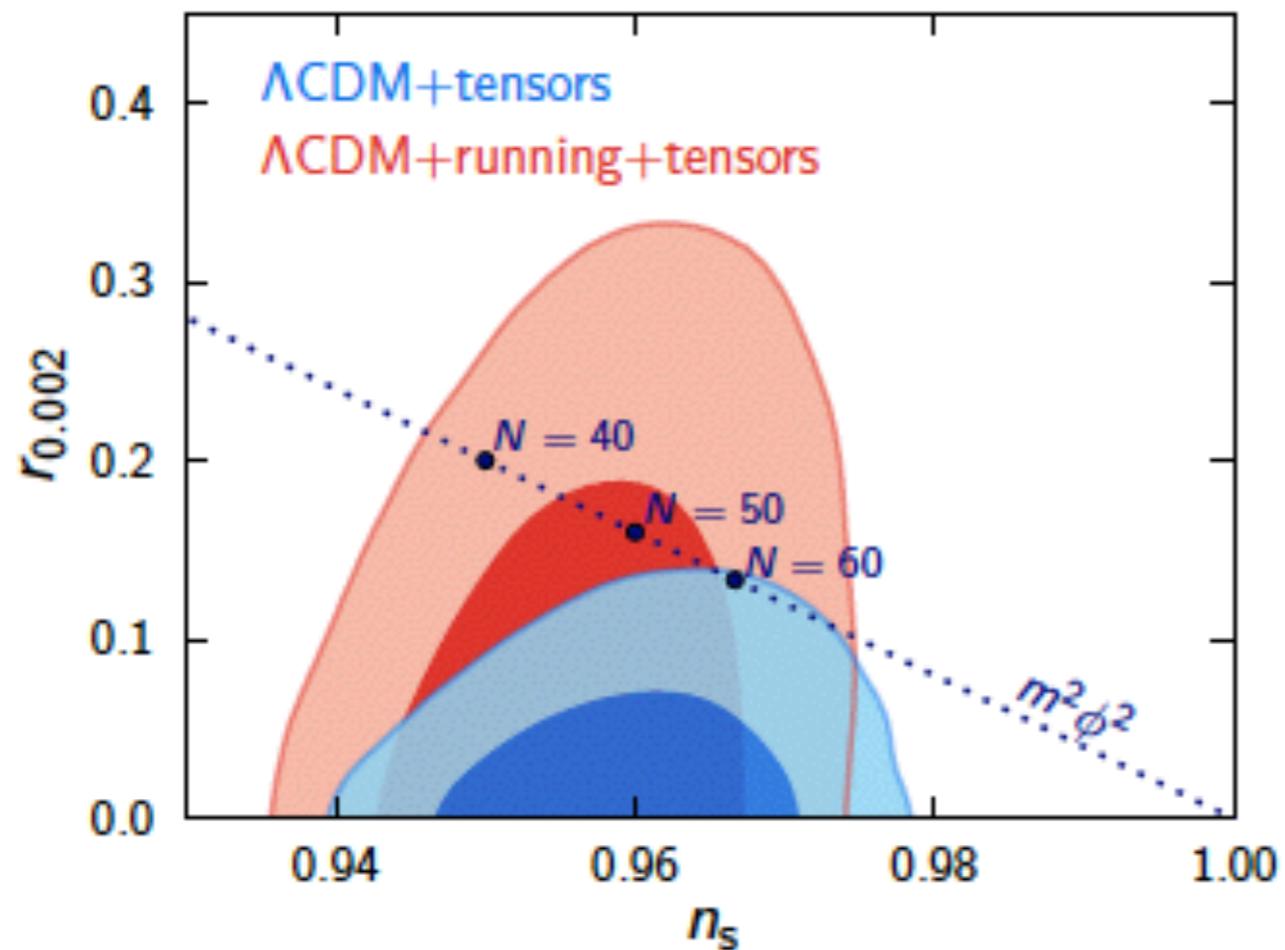


Planck Collaboration XVI (2013) Figure 23: Posterior distribution for $n_s$ for the ΛCDM model with tensors (blue) compared to the posterior when a tensor component and running scalar spectral index are added to the model (red). The dotted line shows the relation between $r$ and $n_s$ for a $V(\varphi)$ inflaton potential where $N$ is the number of inflationary e-foldings.
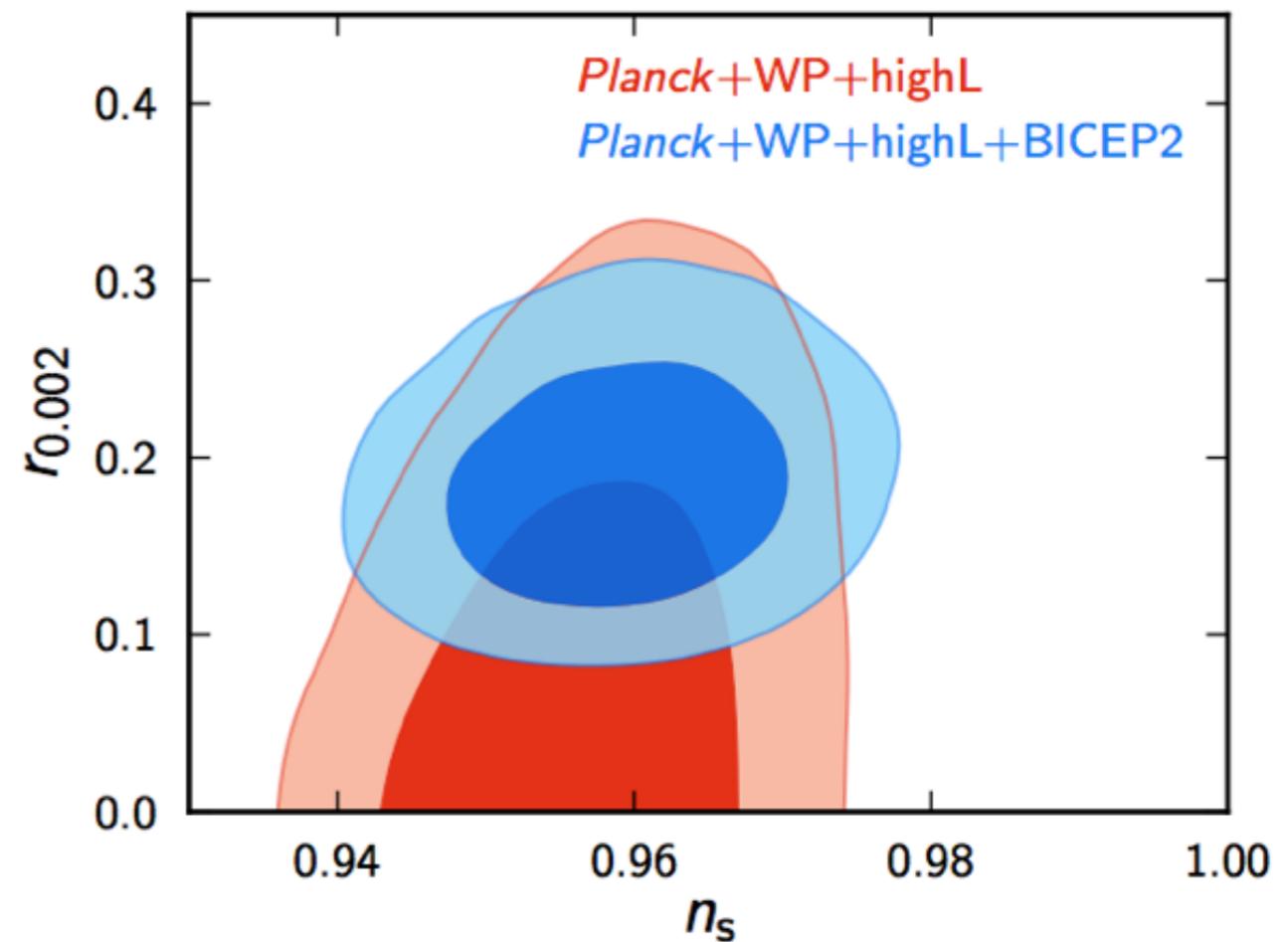
Planck indirect constraints on r from CMB temperature spectrum measurements relax in the context of various model extensions. Shown here, following Planck Collaboration XVI (2013) Figure 23, with tensors and running of the scalar spectral index added to the base ΛCDM model. The contours show the resulting 68% and 95% confidence regions for r and the scalar spectral index $n_s$ when also allowing running.

# Post-Inflation

**Baryogenesis**: generation of excess of baryon (and lepton) number compared to anti-baryon (and anti-lepton) number. In order to create the observed baryon number today

$$\frac{n_B}{n_\gamma} = \left(6.1\, ^{+0.3}_{-0.2}\right) \times 10^{-10}$$

it is only necessary to create an excess of about 1 quark and lepton for every ~$10^9$ quarks+antiquarks and leptons +antileptons.

**Other things that might happen Post-Inflation:**

**Breaking of Pecci-Quinn symmetry** so that the observable universe is composed of many PQ domains.

**Formation of cosmic topological defects** if their amplitude is small enough not to violate cosmological bounds.