November 2021

The NAT Lecture

Simulating cosmic structures

Simon White

Max Planck Institute for Astrophysics

CMB map from the full *Planck* mission



The current CMB evidence for ΛCDM



Planck Collaboration 2018

Parameter	Combined
$\overline{\frac{\Omega_{\rm b}h^2}{\Omega_{\rm c}h^2}} \dots \dots \dots$	$\begin{array}{c} 0.02233 \pm 0.00015 \\ 0.1198 \pm 0.0012 \end{array}$
$100\theta_{\rm MC}$	$\begin{array}{c} 1.04089 \pm 0.00031 \\ 0.0540 \pm 0.0074 \end{array}$
$\frac{\ln(10^{10}A_{\rm s}) \ldots \ldots}{n_{\rm s} \ldots \ldots}$	3.043 ± 0.014 0.9652 ± 0.0042
$\frac{\overline{\Omega_{\rm m}h^2} \dots \dots}{H_0 [{\rm kms^{-1}Mpc^{-1}}]}$	0.1428 ± 0.0011 67.37 ± 0.54
$\Omega_{\rm m}$	$\begin{array}{c} 0.3147 \pm 0.0074 \\ 13.801 \pm 0.024 \end{array}$
$\sigma_8 \dots \dots \dots$ $S_8 \equiv \sigma_8 (\Omega_m/0.3)^{0.5}$	$\begin{array}{c} 0.8101 \pm 0.0061 \\ 0.830 \pm 0.013 \end{array}$
$z_{\rm re}$ $100\theta_*$	7.64 ± 0.74 1.04108 ± 0.00031 147.18 ± 0.29
drag [101pc] · · · · ·	147.10 ± 0.29

• <u>No</u> local/low-redshift data are used

Measurements of all 6 ΛCDM parameters Cosmic properties, not fitting parameters

• Low-z data needed to specify <u>nature</u> of the DM

The current CMB evidence for ΛCDM



• Low-z data needed to specify <u>nature</u> of the DM

The current CMB evidence for ΛCDM



• Low-z data needed to specify <u>nature</u> of the DM

Lyman α forest spectra compared to ACDM predictions



Cosmology and galaxy formation

- The geometry is flat to better than 0.5%
- Baryon and CDM densities, H_0 and σ_8 are known to ~1%
- Initial P(k) is Λ CDM with $n \sim 0.97$ down to subgalactic scales
- Initial non-gaussianities and Σm_v are both small

Cosmology and galaxy formation

- The geometry is flat to better than 0.5%
- Baryon and CDM densities, H_0 and σ_8 are known to ~1%
- Initial P(k) is Λ CDM with $n \sim 0.97$ down to subgalactic scales
- Initial non-gaussianities and Σm_v are both small
- Late-time expansion history BAO signal in galaxies w(z)
- Late-time growth factor z-space distortions mod.grav., v masses
- Dwarf galaxy core structure / Ly α forest WDM / SIDM / fuzzy DM
- Signatures of DE interactions with DM? with v's? with baryons?

Cosmology and galaxy formation

- The geometry is flat to better than 0.5%
- Baryon and CDM densities, H_0 and σ_8 are known to ~1%
- Initial P(k) is Λ CDM with $n \sim 0.97$ down to subgalactic scales
- Initial non-gaussianities and Σm_v are both small
- Late-time expansion history BAO signal in galaxies w(z)
- Late-time growth factor z-space distortions mod.grav., v masses
- Dwarf galaxy core structure / Ly α forest WDM / SIDM / fuzzy DM
- Signatures of DE interactions with DM? with v's? with baryons?

Does galaxy formation distort or mask these signals at the 1% level?

Making predictions for galaxies (accurately?)

Simulations are *required* to show that nonlinear effects are under control and to represent realistic observational surveys

Making predictions for galaxies (accurately?)

Simulations are <u>required</u> to show that nonlinear effects are under control and to represent realistic observational surveys

- Are the initial conditions well enough represented?
- Is the volume large enough to control cosmic variance?
- Can the code follow growth sufficiently well?
- Is galaxy formation represented at a sufficient level by:

Making predictions for galaxies (accurately?)

Simulations are <u>required</u> to show that nonlinear effects are under control and to represent realistic observational surveys

- Are the initial conditions well enough represented?
- Is the volume large enough to control cosmic variance?
- Can the code follow growth sufficiently well?
- Is galaxy formation represented at a sufficient level by:
 - Halo occupation distribution (HOD) models
 - Subhalo abundance matching (SHAM) models
 - Semianalytic population simulations (SAM)
 - Cosmological hydrodynamical simulations

Modelling galaxies for large-scale structure

Halo Occupation Distributions (HOD)

Input: N-body simulation with halos Fit data: Galaxy abundances and clustering at a given redshift Output: Parameters α in $P_{\alpha} \{L_{cen} \dots | M_{halo} \dots\}, n_{\alpha}(L_{sat}, r | M_{halo} \dots)$

Subhalo Abundance Matching (SHAM)

Input: N-body simulation with halos+subhalos, observed $\Phi(L)$ Fit data: Galaxy clustering at a given redshift Output: Scatter in L – M_{halo} relation, "best" estimator for M_{halo}

Semianalytic/Empirical Models (SAM)

Input: N-body simulation with halos+subhalos+merger tree Fit data: Galaxy abundances (and clustering) at multiple redshifts Output: Parameters of physical/empirical galaxy formation model

Cosmological Hydrodynamical Simulations

Springel et al 2005

125 Mpc/h



Springel et al 2005

z = 0 Galaxy Light

Halo clustering depends on formation history



Gao, Springel & White 2005

The 20% of halos with the <u>latest</u> half-mass assembly redshifts in a 30 Mpc/h thick slice

 $M_{halo} \sim 10^{11} M_{\odot}$

Halo clustering depends on formation history



Gao, Springel & White 2005

The 20% of halos with the <u>earliest</u> half-mass assembly redshifts in a 30 Mpc/h thick slice

 $M_{halo} \sim 10^{11} M_{\odot}$



Halo bias as a function of mass and formation time

Gao, Springel & White 2005

On large scales halo bias increases smoothly with formation redshift

The dependence of bias on formation redshift is strongest at low mass

This behaviour is inconsistent with simple versions of excursion set theory, and of HOD and halo abundance matching models

Halo assembly bias: conclusions

The large-scale bias of halo clustering depends not only on halo mass through $v = \delta_c / D(z) \sigma_o(M)$, but <u>also</u> on

- formation time
- concentration
- substructure content
- spin
- shape
- velocity anisotropy

The dependences on different assembly variables are different and <u>cannot</u> be derived from each other: $b = b(M, \underline{A})$ with \underline{A} multidimensional.

These dependences are likely to be reflected in <u>galaxy</u> bias

Most stars are in galaxies similar in mass to the Milky Way Dark matter is *much* more broadly distributed across halos



Most stars are in galaxies similar in mass to the Milky Way Dark matter is *much* more broadly distributed across halos

► Galaxy to halo mass ratio varies *strongly* with mass



Most stars are in galaxies similar in mass to the Milky Way Dark matter is *much* more broadly distributed across halos

→ Halo to galaxy mass ratio varies *strongly* with mass

Star formation efficiency is reduced at both low *and* high halo mass



Most stars are in galaxies similar in mass to the Milky Way Dark matter is *much* more broadly distributed across halos Halo to galaxy mass ratio varies *strongly* with mass Star formation efficiency is reduced at both low *and* high halo mass $(\Omega_{b} / \Omega_{m}) M_{halo} = M_{hot} + M_{cold} + M_{ejecta} + M_{star} + M_{BH}$ black hole quasar mode accretion radio mode accretion RM feedback cooling cold interstellar **IGM** hot halo gas stripping ▲ISM reheating gas infall SN feedback stellar mass *X* loss winds star formation stars ejected gas

Most stars are in galaxies similar in mass to the Milky Way Dark matter is *much* more broadly distributed across halos

Halo to galaxy mass ratio varies strongly with mass

Star formation efficiency is reduced at both low *and* high halo mass

 $(\Omega_{b} / \Omega_{m}) M_{halo} = M_{hot} + M_{cold} + M_{ejecta} + M_{star} + M_{BH}$



The semi-analytic programme

Follow the DM distribution with high-resolution simulations Identify dark halos/subhalos at all times, building merger trees to describe their growth, internal structure and spatial distribution

Treat baryonic physics within the evolving population of DM objects using simplified physical models for processes such as Gas cooling onto central galaxies Star formation within these central galaxies Central black hole growth Generation of winds through stellar and AGN feedback Production, expulsion and mixing of nucleosynthesis products

Measure the <u>efficiencies</u> of these processes as functions of redshift and galaxy properties by comparing model output directly with observational data

Six parameters fine-tuned to fit three curves

Pl	lanck	k+1	WP

Parameter	Best fit	68% limits
$\Omega_{ m b}h^2$	0.022032	0.02205 ± 0.00028
$\Omega_{ m c}h^2$	0.12038	0.1199 ± 0.0027
$100\theta_{\rm MC}$	1.04119	1.04131 ± 0.00063
au	0.0925	$0.089^{+0.012}_{-0.014}$
$n_{\rm s}$	0.9619	0.9603 ± 0.0073
$\ln(10^{10}A_{\rm s})$	3.0980	$3.089^{+0.024}_{-0.027}$





Calibrating models for (sub)halo occupation



Henriques et al (2015)

The 17 parameters of the SA subhalo occupation model constrained by MF and passive fraction observations over $0 \le z \le 3$ and three orders of magnitude in stellar mass

The MCMC chains show all parameters to be determined to moderate accuracy with no major degeneracies

z=2.0

9

8 1 2

11

Observations used in MCMC

z=3.0

8

This Work - Planck1 Henriques2013a - WMAP7

... Guo2013a - WMAP7

10

 $\log_{10}(M_{\bullet}[h^{-2}M_{\odot}])$









A population simulation prediction for galaxy halos



Central galaxies of a given stellar mass are predicted to have larger halo masses if they are red (passive) than if they are blue (star-forming)

This is because central galaxies stop growing after quenching but their halos do not

This effect is **not** present (by construction) in age+abundance SHAM models

Halo mass dependence on central galaxy colour?



Blue centrals have lower mass halos than red centrals of the same stellar mass according to estimates based on the motions of satellites and on weak gravitational lensing

Semianalytic versus full MHD simulations



Semianalytic versus full MHD simulations



Mass distribution dependence on baryon physics



AGN feedback sufficient to match the stellar mass function of galaxies at high mass affects the power spectrum of the total mass distribution at > 1% for k > 0.3 h/Mpc

This will affect the small-scale lensing power spectrum.

van Daalen et al 2011

In summary...

Precision cosmology with galaxy surveys requires the relation between the galaxy and dark matter distributions to be known <u>precisely</u>

- Halo clustering depends at the 10 to 30% level on many aspects of halo structure and formation history in addition to halo mass
- This complexity carries over to the galaxy population and affects both the spatial and kinematic (peculiar velocity) properties
- Different galaxy types can have BAO features of different shape
- Halo mass depends on both colour and mass of the central galaxy
- Baryon physics can affect the lensing P(k) down to $k \sim 0.3$ h/Mpc

All these effects depend on the <u>details</u> of galaxy formation physics None is easily included in the HOD or SHAM modelling frameworks

Galaxy formation is an insoluble problem

• Galaxies form as gas cools and condenses at the centres of a population of massive halos growing by gravitational amplification of fluctuations in an initially near-uniform distribution of pre-existing dark matter

• Galaxies form as gas cools and condenses at the centres of a population of massive halos growing by gravitational amplification of fluctuations in an initially near-uniform distribution of pre-existing dark matter

Semi-analytic and Subhalo Abundance Matching models <u>assume</u> this and tune a physically based (SAM) or purely statistical (SHAM) relation between galaxy properties and subhalo history to fit observation.

• Galaxies form as gas cools and condenses at the centres of a population of massive halos growing by gravitational amplification of fluctuations in an initially near-uniform distribution of pre-existing dark matter

Subhalo Abundance Matching and Semi-analytic models assume this and tune a more (SAM) or less (SHAM) complicated relation between galaxy properties and subhalo history to fit observation.

Main outstanding issues are:

- I. The dependence of the survival of satellite subhalos on resolution, integration accuracy, and baryon effects the "orphan" problem
- II. The number of properties of subhalo histories needed to predict their galaxy content to the required precision the "assembly bias" problem

- Galaxies form as gas cools and condenses at the centres of a population of massive halos as these grow by gravitational amplification of fluctuations in an initially near-uniform distribution of pre-existing dark matter
- The efficiency of galaxy formation is limited by feedback that is most effective at low and at high halo mass. Different astrophysical processes are required in the two cases.

- Galaxies form as gas cools and condenses at the centres of a population of massive halos as these grow by gravitational amplification of fluctuations in an initially near-uniform distribution of pre-existing dark matter
- The efficiency of galaxy formation is limited by feedback that is most effective at low and at high halo mass. Different astrophysical processes are required in the two cases.

At low mass: Reionization heating; Star-formation-driven winds At high mass: Inefficient cooling; AGN feedback

Galaxy formation is an insoluble problem or G-1-10-1F **problem** Norberg et al. (2002) 10^{-2} • Galaxies f es of a population of massive h: 10^{-3} in an initia ation of fluctuations g dark matter 10^{-4} The efficit $\bigvee_{M_{c}}^{-od}$ • The efficie ack that is most ohysical processes 10⁻⁵ are require Millennium Simulation 10-6 Croton et al (2006) n winds At low ma: Cooling only At high ma 10^{-7} -22 -20 -18 -16 -24 -14 $M_{bJ} = 5 \log_{10} h$

Galaxy formation is an insoluble problem or G-1-10-1 F problem Norberg et al. (2002) 10^{-2} • Galaxies f es of a population of massive ha 10^{-3} in an initia 10^{-3} The efficie M_{E}^{-3} 10^{-4} effective a 10^{-5} tion of fluctuations g dark matter • The efficie ack that is most physical processes 10-5 are require

At low man At high ma 10-6

 10^{-7}

-24

Millennium Simulation Millennium Simulation Croton et al (2006) Cooling+SN winds -22 -20 -18 -16 -14 $M_{bJ} - 5loq_{10}h$

Galaxy formation is an insoluble problem or problem G¹10⁻¹ Norbera et al. (2002) 10^{-2} • Galaxies f es of a population of massive ha in an initia ition of fluctuations g dark matter 10-4 The efficit $\bigvee_{\mathbb{R}^{d}}$ ack that is most • The efficie physical processes 10-5 are require Millennium Simulation 10-6 Croton et al (2006) n winds At low may Cooling+SN winds+AGN At high ma 10^{-7} -22 -20 -18 -16 -14 -24 M_{bJ} - 5loq₁₀h



- Galaxies form as gas cools and condenses at the centres of a population of massive halos as these grow by gravitational amplification of fluctuations in an initially near-uniform distribution of pre-existing dark matter
- The efficiency of galaxy formation is limited by feedback that is most effective at low and at high halo mass. Different astrophysical processes are required in the two cases.

At low mass: Reionization heating; Star-formation-driven winds At high mass: Inefficient cooling; AGN feedback

Main outstanding issues:

I. Mechanical/radiative feedback, B-fields/cosmic rays, ejection/recycling II. Can "subgrid" processes be sufficiently well/uniquely characterised?



Recent cosmological (magneto)hydrodynamical simulations reproduce many aspects of the observed internal structure of galaxies....





Simulating the structure of galaxies

FIRE



- ...but they differ strongly in their treatment of the ISM, of star formation, of feedback, of nuclear BH's...
- They do not include processes known to be significant (cosmic rays/B-fields, binary evolution, dust evolution)
- They make different predictions for properties not used as constraints (gas/bar fractions, CGM/ ISM structure)
- They are not yet checked across the full range of galaxy masses and environments.

Recent cosmological (magneto)hydrodynamical simulations reproduce many aspects of the observed internal structure of galaxies....

• Star formation: IMF as a function of Z, p, ρ , $\langle v^2 \rangle$, B...; GRB, GW precursors

- Star formation: IMF as a function of Z, p, ρ , $\langle v^2 \rangle$, B...; GRB, GW precursors
- Spirals/bars/warps: internal versus external driving, the role of gas

- Star formation: IMF as a function of Z, p, ρ , $\langle v^2 \rangle$, B...; GRB, GW precursors
- Spirals/bars/warps: internal versus external driving, the role of gas
- Phase structure of the ISM/CGM/IGM: role of B-fields, cosmic rays, dust

- Star formation: IMF as a function of Z, p, ρ , $\langle v^2 \rangle$, B...; GRB, GW precursors
- Spirals/bars/warps: internal versus external driving, the role of gas
- Phase structure of the ISM/CGM/IGM: role of B-fields, cosmic rays, dust
- SMBH formation and fuelling

- Star formation: IMF as a function of Z, p, ρ , $\langle v^2 \rangle$, B...; GRB, GW precursors
- Spirals/bars/warps: internal versus external driving, the role of gas
- Phase structure of the ISM/CGM/IGM: role of B-fields, cosmic rays, dust
- SMBH formation and fuelling
- Launching of SF/starburst/AGN winds: the mechanisms of mass loading

- Star formation: IMF as a function of Z, p, ρ , $\langle v^2 \rangle$, B...; GRB, GW precursors
- Spirals/bars/warps: internal versus external driving, the role of gas
- Phase structure of the ISM/CGM/IGM: role of B-fields, cosmic rays, dust
- SMBH formation and fuelling
- Launching of SF/starburst/AGN winds: the mechanisms of mass loading
- Inflow/outflow interactions: galactic fountains, IGM metals

- Star formation: IMF as a function of Z, p, ρ , $\langle v^2 \rangle$, B...; GRB, GW precursors
- Spirals/bars/warps: internal versus external driving, the role of gas
- Phase structure of the ISM/CGM/IGM: role of B-fields, cosmic rays, dust
- SMBH formation and fuelling
- Launching of SF/starburst/AGN winds: the mechanisms of mass loading
- Inflow/outflow interactions: galactic fountains, IGM metals
- Mergers: the genealogy of the 1%, restructuring through major(?) mergers

- Star formation: IMF as a function of Z, p, ρ , $\langle v^2 \rangle$, B...; GRB, GW precursors
- Spirals/bars/warps: internal versus external driving, the role of gas
- Phase structure of the ISM/CGM/IGM: role of B-fields, cosmic rays, dust
- SMBH formation and fuelling
- Launching of SF/starburst/AGN winds: the mechanisms of mass loading
- Inflow/outflow interactions: galactic fountains, IGM metals
- Mergers: the genealogy of the 1%, restructuring through major(?) mergers
- Environment effects: nature vs nurture, stripping/harassment/strangulation

- Star formation: IMF as a function of Z, p, ρ , $\langle v^2 \rangle$, B...; GRB, GW precursors
- Spirals/bars/warps: internal versus external driving, the role of gas
- Phase structure of the ISM/CGM/IGM: role of B-fields, cosmic rays, dust
- SMBH formation and fuelling
- Launching of SF/starburst/AGN winds: the mechanisms of mass loading
- Inflow/outflow interactions: galactic fountains, IGM metals
- Mergers: the genealogy of the 1%, restructuring through major(?) mergers
- Environment effects: nature vs nurture, stripping/harassment/strangulation

(Multiple) phenomenological models have been suggested for all of these Convincing *ab initio* physical models are available for very few Mass and detailed assembly history determine their relative importance

Epistemology for complex systems

(galaxy formation, climate change, ecology, macro-economics, brain function)

- Agreement of the galaxy population in a modern cosmological hydrodynamical simulation with (aspects of) real populations may contribute rather little to our knowledge/understanding of galaxy formation, since
 - part of the agreement is due to calibration/tuning
 - simulations with *different* subgrid models often agree equally well
 - unexamined (but linked) aspects often disagree with observation
 - better resolution or subgrid modelling may ruin the agreement
- It is important to understand *why* simulation and observation agree. Intuition is often helped by models which isolate individual processes
- Stronger conclusions can often be drawn from showing that some aspects of the observations *cannot* be fit, implying e.g. that
 - the integration scheme is insufficiently accurate, or
 - the subgrid models incorrectly represent the astrophysics, or
 - critical processes are not yet included, or
 - $-\Lambda CDM$ is wrong

Summary points?

- ΛCDM is an *a priori* theoretical model with parameters fully specified by CMB measurements
- Of its basic tenets, only the cold nature of the Dark Matter *requires* data from the low-redshift Universe for justification/validation
- In principle, ACDM thus predicts **all** properties of the nonlinear, latetime universe (e.g. all galaxy properties) with no further freedom
- In practice, it can be very hard to calculate these predictions reliably.
- Different (uncertain) treatments of astrophysical processes can lead to very different galaxy properties within the *same* ΛCDM framework

Summary points?

- ΛCDM is an *a priori* theoretical model with parameters fully specified by CMB measurements
- Of its basic tenets, only the cold nature of the Dark Matter *requires* data from the low-redshift Universe for justification/validation
- In principle, ACDM thus predicts **all** properties of the nonlinear, latetime universe (e.g. all galaxy properties) with no further freedom
- In practice, it can be very hard to calculate these predictions reliably.
- Different (uncertain) treatments of astrophysical processes can lead to very different galaxy properties within the *same* ΛCDM framework

It seems very unlikely that the detailed structural properties of galaxies can be used reliably to infer failings of Λ CDM

Summary points?

- ΛCDM is an *a priori* theoretical model with parameters fully specified by CMB measurements
- Of its basic tenets, only the cold nature of the Dark Matter *requires* data from the low-redshift Universe for justification/validation
- In principle, ACDM thus predicts **all** properties of the nonlinear, latetime universe (e.g. all galaxy properties) with no further freedom
- In practice, it can be very hard to calculate these predictions reliably.
- Different (uncertain) treatments of astrophysical processes can lead to very different galaxy properties within the *same* ΛCDM framework

Complex simulations of knowledge?

Limited observations of a more complex reality