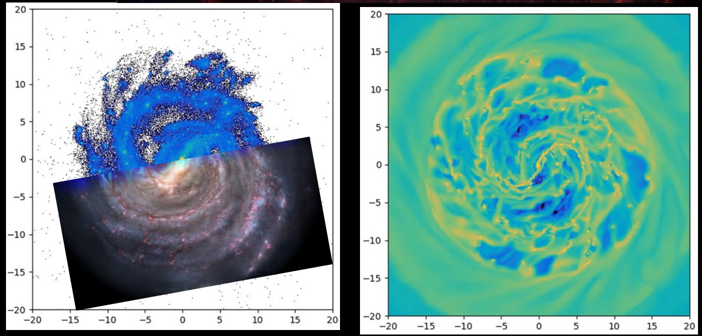
Eric Emsellem — Astronomer / Head of ESO Project Science

Scientific Interests : *constrain physical processes that drive the formation and evolution of galaxies, scale-coupling, galactic dynamics*

Tools and Methods

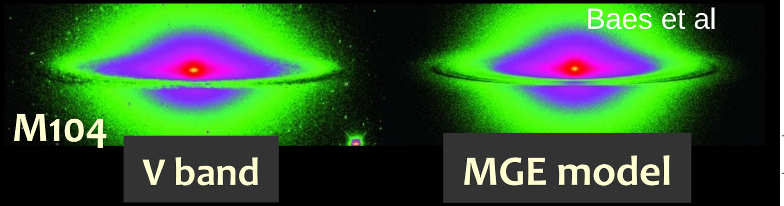
- **Instrumentation** ⇒ Integral-field spectroscopy
- **Observations** ⇒ Nearby galaxies, dynamics, stellar populations
- **Modeling** ⇒ Dynamics / morphology
- **Simulations** ⇒ Hydro-dynamical runs (mostly RAMSES)
- **Codes** ⇒ e.g., MGE, pymusepipe, pPXF, pipelines



M 83

Projects, e.g. :

- **Sauron, Atlas^{3D}** Phangs
- **Condor, Geckos, MAUVE**
- **MUSE, WST**

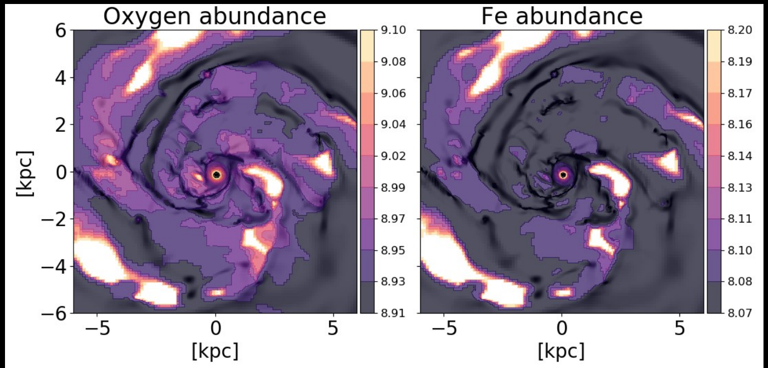


My Path

- **1990** Engineering degree
- **1994** PhD in Lyon
- **1994** ⇒ Post-Docs in Leiden (NL) + Munich (DE)
- **1997 / 2006** ⇒ Assoc. / full Astronomer at CRAL – Lyon
- **2009** ⇒ Head of the ESO Office for Science (Garching, Germany)
- **2020** ⇒ Head of ESO Project Science

Hobbies

Scuba diving, Photography
Music, Science-fiction lit.



NGC0628

Observations & **Simulations**

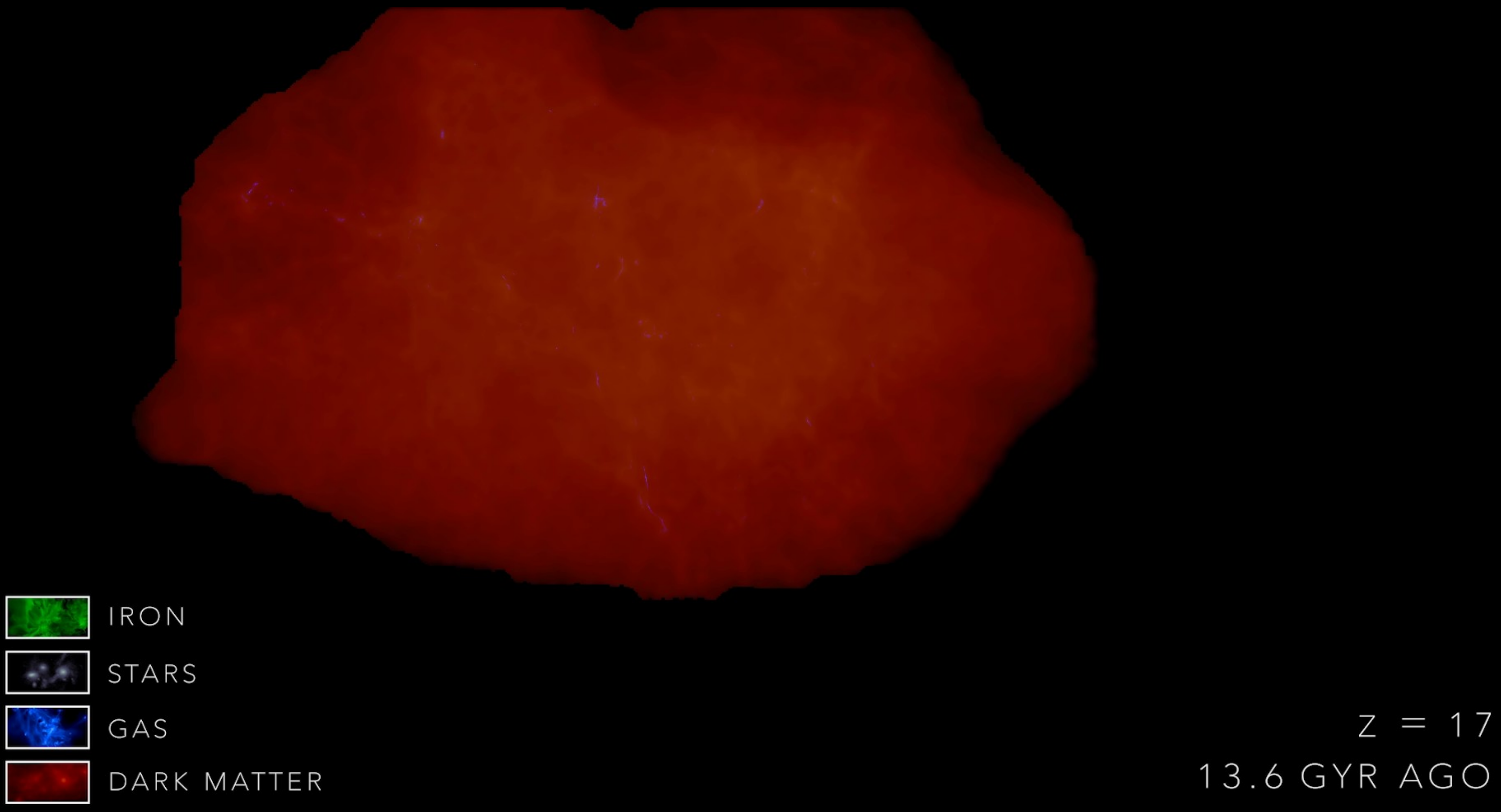
What can we learn ?

With big thanks to:

Jeremy Fensch, Jeremy Blaizot, Florent Renaud

Partly inspired by **Christophe Michel**

Eric Emsellem
ESO



Agertz / Renaud / Segovia Otero [2021, 2022, 2025]

This presentation

Is a way (for me) to emphasise
- The challenges we face with simulations
- The difficulty to compare simulations and observations
- A few biases that emerge from the literature
  → my personal, present, perspective

Is NOT
- A criticism of simulations (or simulators)
  - they are achievements!

→ and in no way, an objective perspective

Numerical simulations = what are they?

Theory?

- Numerical Recipes involves formal knowledge of physics
- Complexity
- Emergence, Predictions

Numerical simulations = what are they?

Theory?

- Numerical Recipes involves formal knowledge of physics
- Complexity
- Emergence, Predictions
- BUT: are not a set of fundamental principles

$\Rightarrow$ Simulations $\neq$ Theory

Numerical simulations = what are they?

Theory?

- Numerical Recipes involves formal knowledge of physics
- Complexity
- Emergence, Predictions
- BUT: are not a set of fundamental principles

⇒ Simulations ≠ Theory

Observations?

- Products include derived quantities, simulations can be "observed"
  BUT
- Some quantities are directly available, some are not
- These are **not** observed data from "our Universe"
  - But from "a" (descriptive) model.

⇒ Simulations ≠ Observations

Numerical simulations = what are they?

Experiments ?

- Involves a setup and rules
- Outcome is not easy to predict
- Several runs may lead to different measurements (noise)
- *Appears* to produce knowledge (?)

  BUT
- **Given a bit of time and patience, …**

Observed Theory or Model ?

- Setup constrained by theory (and the user)
- Products are measurements within that setup
  - Not pure theory, but also

    → implementation-dependent (numerics, recipes, …)

⇒ Observed descriptive Model

How to proceed ?

What does it mean (to run simulations) ?

A Series of operations…

(with some patience and a lot of time: you could do it on paper)

Model

- Representation of the Universe
- Expression of that model with (restricted) recipes and rules

Initial Conditions

- Realisation and setup

Integration time

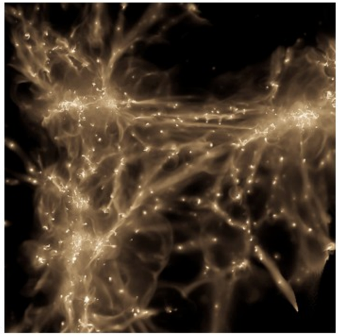
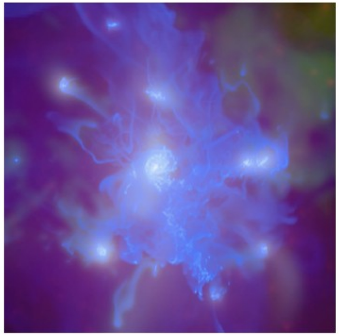
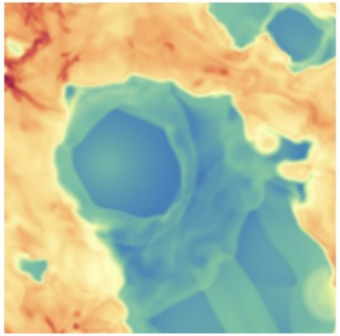
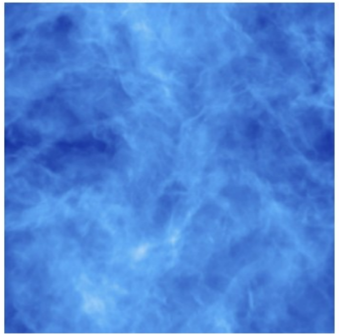
- Hardware, how long

Numerics

- Framework (and coding approach)

→ See Simon Glover's talk!

Several types of galaxy simulations

	New Horizon Dubois et al. (2021)	*Vintergatan* Agertz et al. (2021) Renaud et al. (2021a,b)	Renaud et al. (2021c)	*Tigress* Kim et al. (2017)	Federrath et al. (2008)
					
	cosmological volume	*cosmological zoom-in*	*isolated galaxy*	*galaxy patch*	*ISM box*
	size: > 10 Mpc res.: ~100-500 pc	size: ~1 Mpc res.: ~10-100 pc	size: ~100 kpc res.: ~0.1-10 pc	size: ~0.1-1 kpc res.: ~0.1-10 pc	size: < 100 pc res.: < 0.1 pc
👍	• initial conditions (CMB) • statistics on galaxy pop.	• initial conditions (CMB) • some can capture GMCs	• control on the parameters • can be cheap to run	• easy to setup • relatively cheap to run	• very high resolution • control on the parameters
👎	• poor resolution for describing star formation and feedback • barely resolves galactic disks	• only one galaxy • do not resolve internal GMC physics • very expensive to run	• not realistic environment (mergers and gas accretion missing) • relies on artificial initial conditions • can be very expensive to run	• misses several aspects of disk dynamics • imposed instabilities • not a huge advantage compared to isolated galaxies	• no realistic gas recycling • no effect of galaxy (e.g. potential, turbulence, tides, shear etc.)

Simulations **vs** Observations

Level 1 = direct comparisons

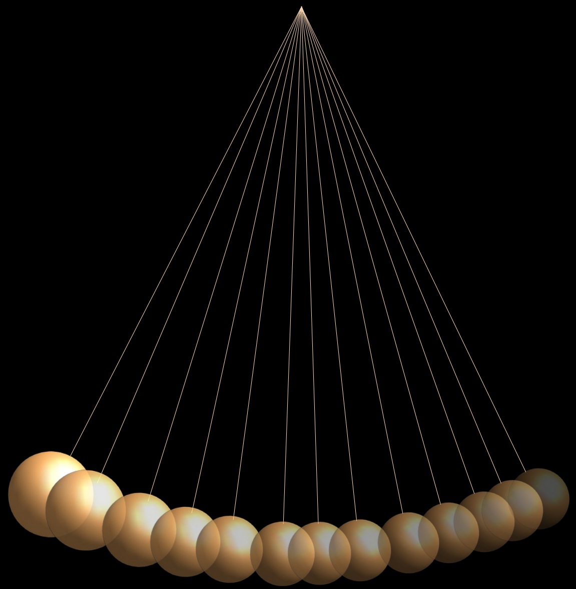
1

Primary

Model → compute → **X, Y, Z**

Initial conditions
Input Physics
Rules & Recipes
Implementation

Densities
Temperatures
Yes/No



Questions

- Do you really understand the physics behind / the rules ?
  - [black box effect]
- How is that quantity derived from Observations ?
  - Is that the same tracer as in the simulation ?

Level 2 = Computed quantities

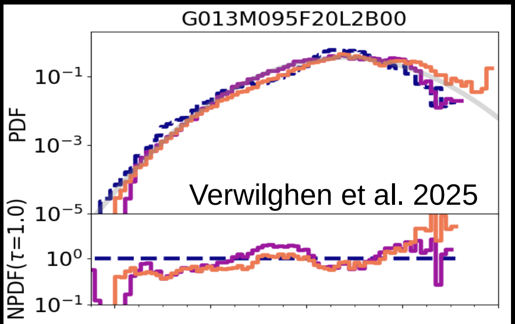
Model → compute → ① **Primary: X, Y, Z** → convert → ② **Secondary: A, B, C**

Initial conditions
Input Physics
Rules & Recipes
Implementation

More physics
More recipes
Implementation

Densities
Temperatures
Yes/No

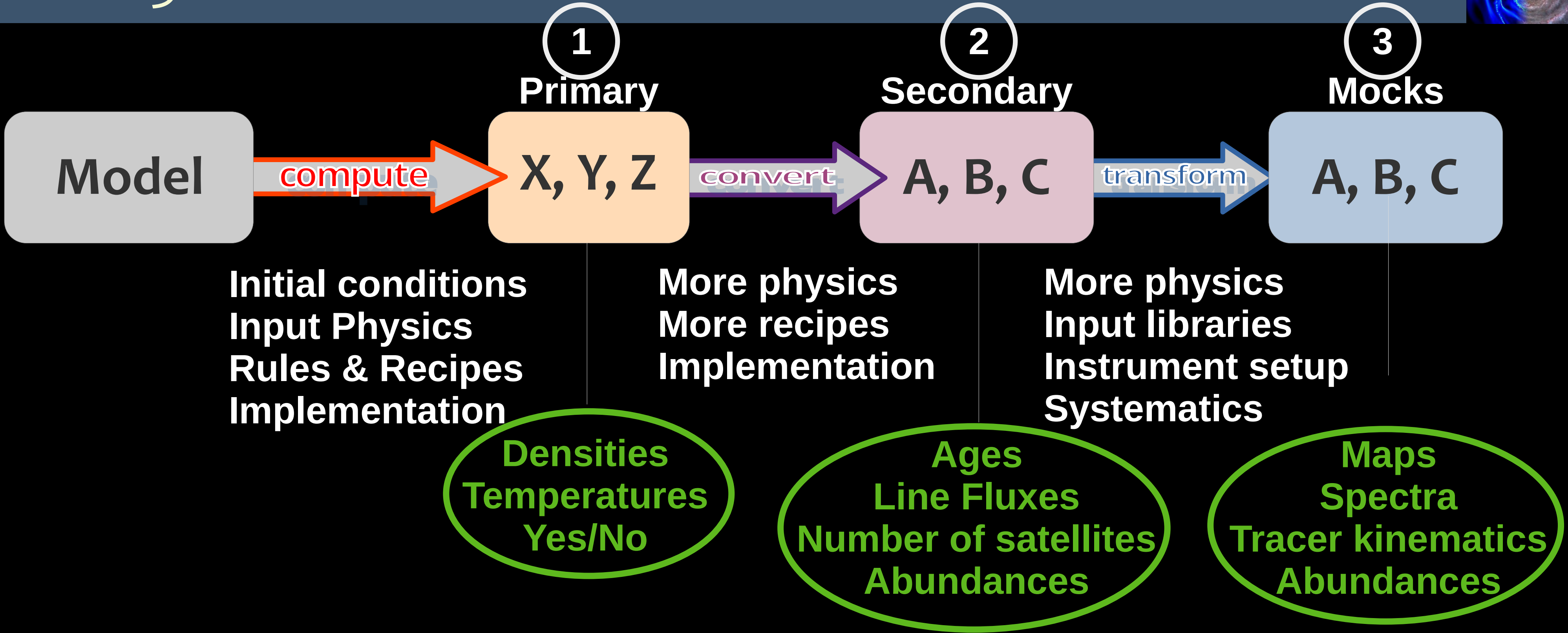
Ages
Line Fluxes
Number of satellites
Abundances



Questions

- Apples with Apples?
- Selection functions

Level 3 = Mocks



Questions

- Adding the “instrument” layer: is that relevant?
- What are the metrics?

Example 1 The SPHINX simulations – Katz et al. 2023

- ΛCDM Cosmological box – 20^3 Mpc^3
- z=10 to 4.64
  - Stars: age, mass, metallicity
    → SED (BPASS), IMF slope, cutoff
  - Emission lines, nebular continuum
  - (Effective) Dust, resonant line RT

→ Many mocks and data have been publicly released

TABLE 2
EMISSION LINES INCLUDED IN THE SPHINX DATA RELEASE. SPECIES ARE DEFINED BY THEIR ELEMENT SYMBOL AND IONIZATION STATE. STATES WITH A SUFFIX OF "R" OR "C" REPRESENT THE RECOMBINATION OR CHARGE EXCHANGE CONTRIBUTION TO A COLLISIONALLY EXCITED LINE. EMISSION LINES SHOWN IN MAGENTA HAVE BEEN PROPAGATED THROUGH DUST AND/OR H I RADIATIVE TRANSFER. WAVELENGTHS ARE THE DEFAULT VALUES IN CLOUDY V17.

Species	State	Wavelength	Species	State	Wavelength
H	1	1215.67 Å	O	3	51.80 μm
H	1	6562.80 Å	O	3	88.33 μm
H	1	4861.32 Å	Ne	3	3868.76 Å
H	1	4340.46 Å	Ne	3	3967.47 Å
H	1	4101.73 Å	C	2	157.64 μm
He	2	1640.41 Å	C	3	1906.68 Å
He	2	4685.68 Å	C	3	1908.73 Å
O	1	6300.30 Å	C	4	1548.19 Å
O	1	6363.78 Å	C	4	1550.78 Å
O	2	3726.03 Å	N	2	5754.61 Å
O	2	3728.81 Å	N	2R	5755.00 Å
O	2R	3726.00 Å	N	2	6548.05 Å
O	2R	3729.00 Å	N	2	6583.45 Å
O	2	7318.92 Å	N	2R	6584.00 Å
O	2	7319.99 Å	N	3	1748.65 Å
O	2	7329.67 Å	N	3	1753.99 Å
O	2	7330.73 Å	N	3	1746.82 Å
O	2R	7332.00 Å	N	3	1752.16 Å
O	2R	7323.00 Å	N	3	1749.67 Å
O	3	1660.81 Å	S	2	6716.44 Å
O	3	1666.15 Å	S	2	6730.82 Å
O	3	4363.21 Å	S	2	4076.35 Å
O	3R	4363.00 Å	S	2	4068.60 Å
O	3C	4363.00 Å	S	3	6312.06 Å
O	3	4958.91 Å	S	3	9068.62 Å
O	3	5006.84 Å	S	3	9530.62 Å

Example 1 The SPHINX simulations – Katz et al. 2023

TABLE 1
STATISTICS OF THE GALAXIES IN THE SPHINX DATA RELEASE. EACH COLUMN SHOWS THE NUMBER OF GALAXIES WITH $\mathrm{SFR} \geq 0.3\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$, THEIR MAXIMUM AND MEDIAN STELLAR MASS, AND THE MAXIMUM AND MEDIAN VIRIAL MASS OF THEIR HOST DARK MATTER HALO, RESPECTIVELY.

Redshift	N_{gal}	$\log M_*$ Max	$\log M_*$ Median	$\log M_{\mathrm{vir}}$ Max	$\log M_{\mathrm{vir}}$ Median
10.0	49	8.56	7.38	10.20	9.30
9.0	66	8.72	7.52	10.48	9.44
8.0	128	9.25	7.62	10.75	9.55
7.0	177	9.65	7.87	10.91	9.69
6.0	276	9.93	8.07	11.12	9.89
5.0	317	10.46	8.27	11.64	10.00
4.6	367	10.63	8.40	11.70	10.07

Example 1 The SPHINX simulations – Katz et al. 2023

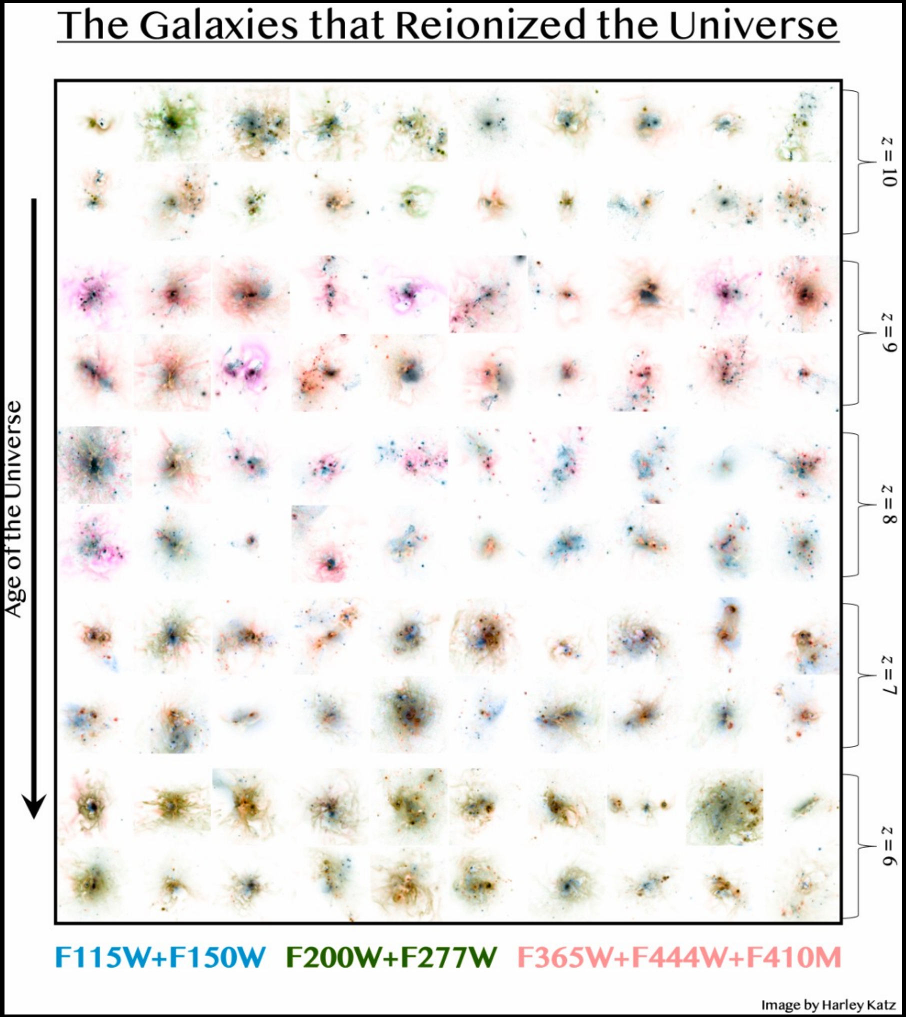
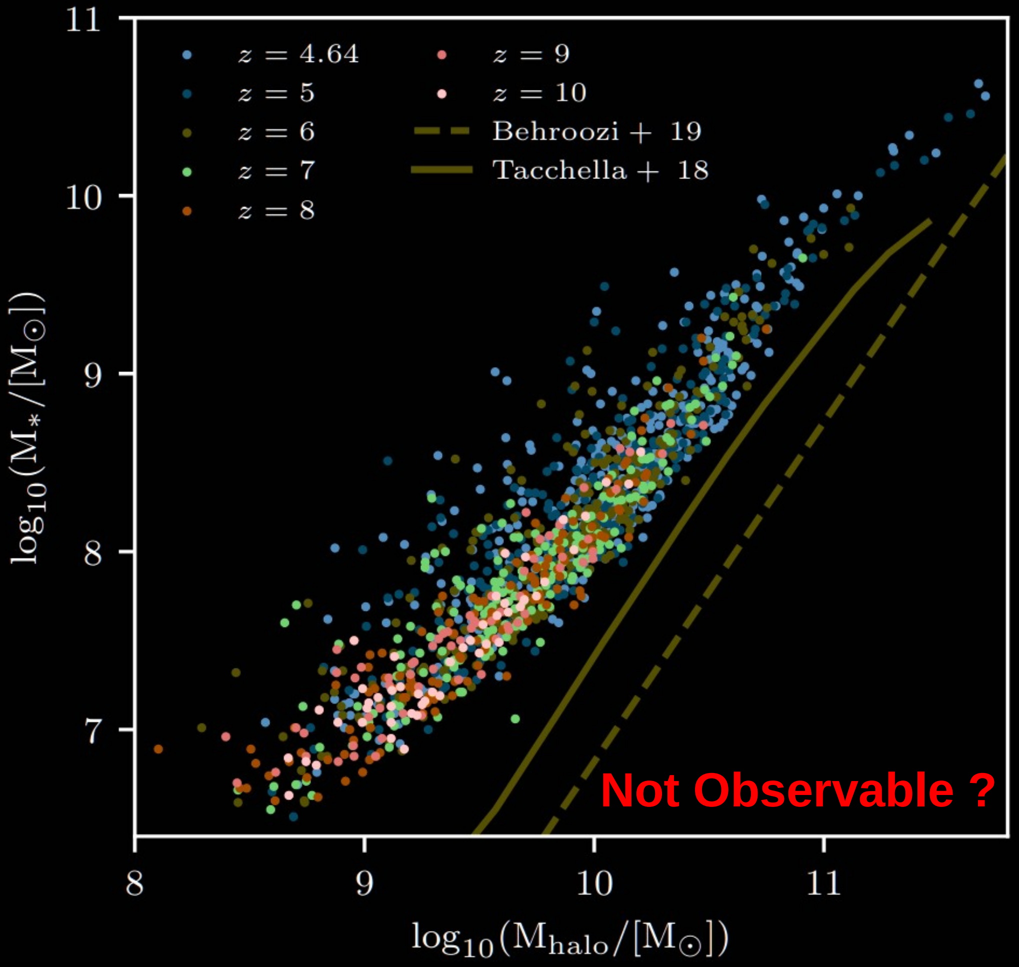
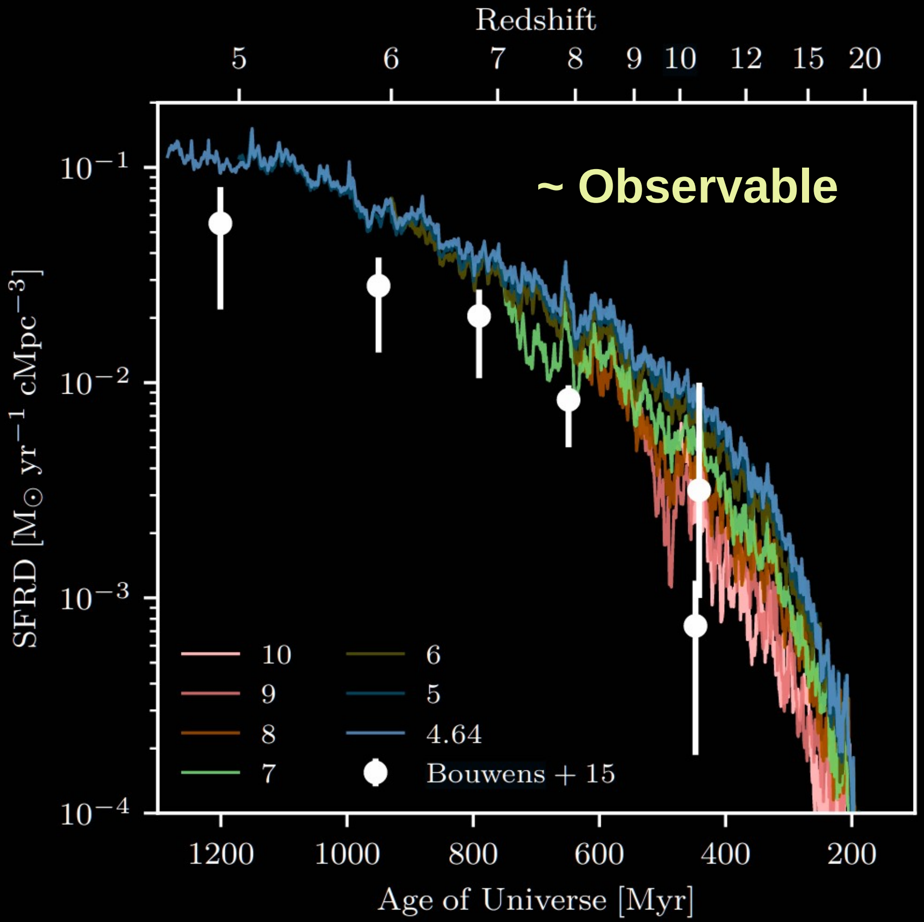


TABLE 3
DETAILS OF THE GALAXY PROPERTIES AVAILABLE AS PART OF THE SPHINX PUBLIC DATA RELEASE.

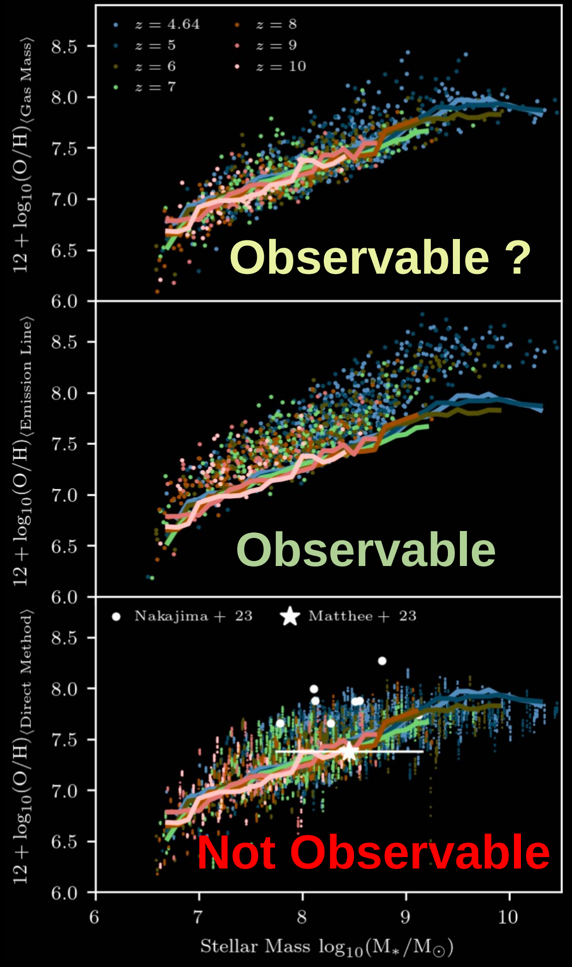
Quantity	Units	Notes
Halo ID		
Redshift		
Halo mass	$\log_{10}(M/M_\odot)$	
Stellar mass	$\log_{10}(M/M_\odot)$	This value is the total stellar mass formed (i.e. the integral of the star formation history) and is not adjusted for mass loss due to stellar feedback
R_{vir}	20 cMpc	
x, y, z position	20 cMpc	3D position of the halo in the simulation volume
Star formation rate	$M_\odot$ yr^{-1}	Provided as an average over 3, 5, 10, and 100 Myr and can be recomputed for any other interval from the star formation history
Star formation history	$M_\odot$ yr^{-1}	Provided for every galaxy on a 1 Myr time cadence
Stellar ages	Myr	Mass-weighted and LyC luminosity-weighted
Stellar metallicity	Absolute	Mass-weighted and LyC luminosity-weighted over all stars
Stellar metallicity history	Absolute	Mass-weighted stellar metallicity of all star particles that formed in bins of 1 Myr
Ionizing luminosity	photons s^{-1}	
LyC escape fraction		Angle-averaged (for all photons with $E > 13.6$ eV) and along ten sight lines (for photons with a wavelength of 900 Å)
ISM gas density	$\log_{10}(n_H/cm^{-3})$	Weighted by intrinsic [O II] $\lambda\lambda$3727 or [C III] $\lambda\lambda$1908
Gas metallicity	$\log_{10}(Z/Z_\odot)$	Mass-weighted as well as [O II] $\lambda\lambda$3727, [O III] λ5007, [N II] λ6583, and Hβ weighted
Emission line luminosities	erg s^{-1}	Intrinsic for all emission lines listed in Table 2, dust attenuated along ten sight lines for Hα, Hβ, Hγ, Hδ, [O II] $\lambda\lambda$3727, [Ne III] λ3869, [O III] λ4363, [O III] λ4959, [O III] λ5007, [N II] λ6583
Stellar continuum luminosities	erg s^{-1} Å^{-1}	Intrinsic & dust attenuated along ten sight lines for 20 wavelengths (1300Å, 1400Å, 1500Å, 1600Å, 1700Å, 1800Å, 1900Å, 2000Å, 2500Å, 3000Å, 3727Å, 3869Å, 4102Å, 4341Å, 4363Å, 4861Å, 4959Å, 5008Å, 6563Å, 6583Å)
Nebular continuum luminosities	erg s^{-1} Å^{-1}	Intrinsic & dust attenuated along ten sight lines for 20 wavelengths (1300Å, 1400Å, 1500Å, 1600Å, 1700Å, 1800Å, 1900Å, 2000Å, 2500Å, 3000Å, 3727Å, 3869Å, 4102Å, 4341Å, 4363Å, 4861Å, 4959Å, 5008Å, 6563Å, 6583Å)
Full SEDs	erg s^{-1} Hz^{-1} cm^{-2}	Intrinsic & dust attenuated along ten sight lines and redshifted to the relevant z. Spectra are computed at 1 Å resolution by interpolating the escape fractions at the 20 continuum wavelengths and for each emission line. SED files provide the total SED as well as the three separate components
E(B−V)		Along ten sight lines. Computed from the Balmer decrement (Hα and Hβ)
Effective radii (R_{eff})	pc	Measured in each of the JWST filters along each line of sight for the largest segment after our image segmentation procedure. We provide the corresponding flux density (nJy) of the segment in addition to its circularized radius
UV continuum slopes (β)		Intrinsic & dust attenuated along ten sight lines. Measured from the full SED (stellar + nebular continuum) as well as only the stellar continuum. Additional values can be measured from the photometry with the inclusion of emission lines
UV magnitudes	AB	Intrinsic & dust attenuated along ten sight lines. Measured at 1500 Å from the stellar and nebular continuum
JWST filter magnitudes	AB	Dust attenuated along ten sight lines. Computed for all NIRCam wide and medium filters (F070W, F090W, F115W, F140M, F150W, F162M, F182M, F200W, F210M, F250M, F277W, F300M, F335M, F356W, F360M, F410M, F430M, F444W, F460M, F480M)
Lyα and Hα spectra	erg s^{-1}	Dust attenuated along ten sight lines. Spectral resolution of 0.1 Å. Values should be divided by the wavelength bins to obtain appropriate units
Lyα and Hα surface brightness profiles	erg s^{-1}	Dust attenuated along ten sight lines. Spatial resolution of $R_{vir}/250$. Values should be divided by the pixel size to obtain surface brightness
Galaxy images	nJy pixel	Dust attenuated along ten sight lines for each JWST filter. Due to data size, these are made available upon request for any emission line or continuum (nebular or stellar) wavelength. Example RGB images combining multiple filters are shown in Figure 1

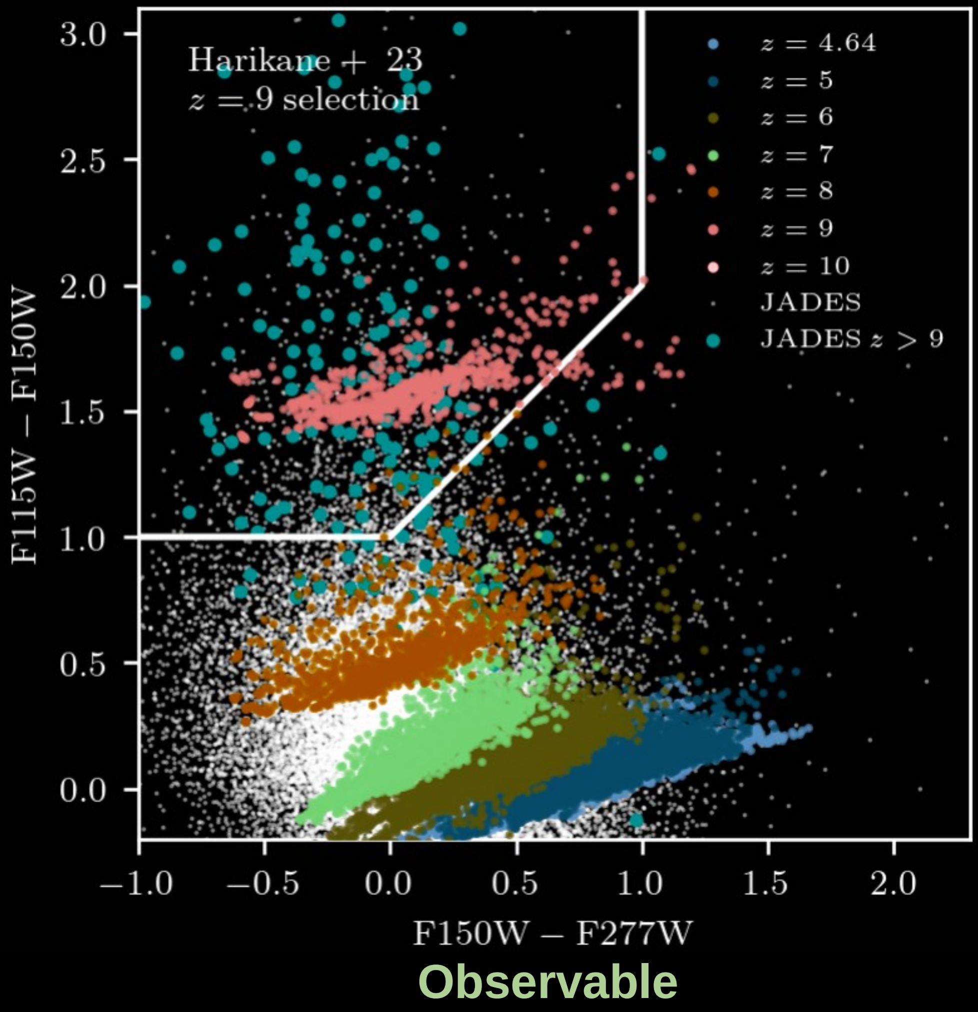
Example 1 The SPHINX simulations – Katz et al. 2023





Example 1 The SPHINX simulations – Katz et al. 2023





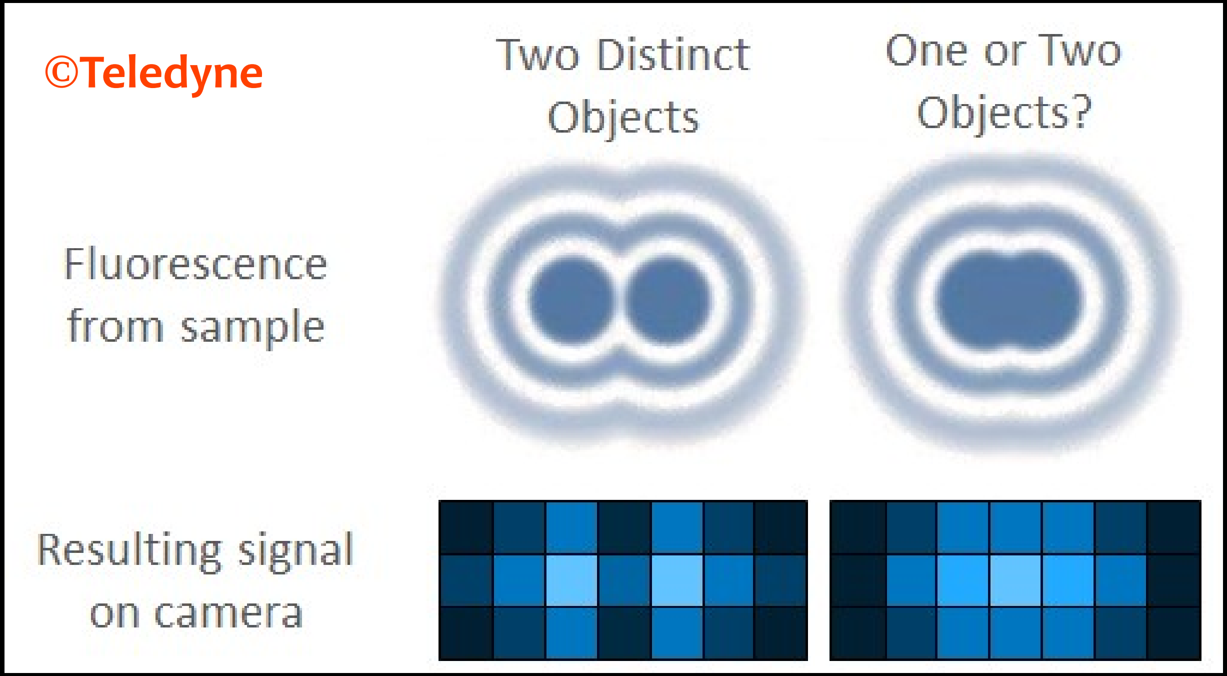
By the way…

What is “resolution” ?

Resolution in Observations

Resolution is about power of separation

- Minimum distance to distinguish 2 objects

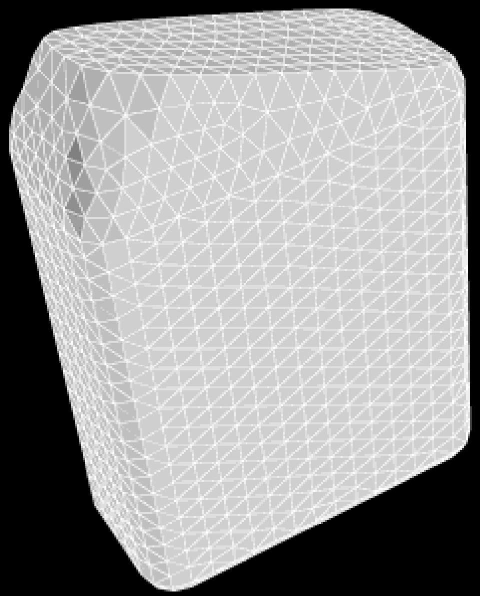


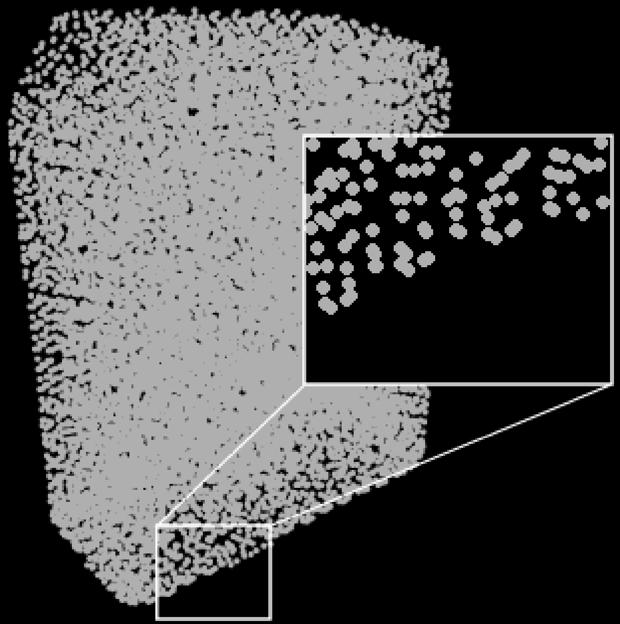
Sampling is about pixel size: signal discretisation

Resolution in Simulations

Resolution in the simulation is about

Where (physical) quantities are computed





Sampling may occur in simulations (AMR)

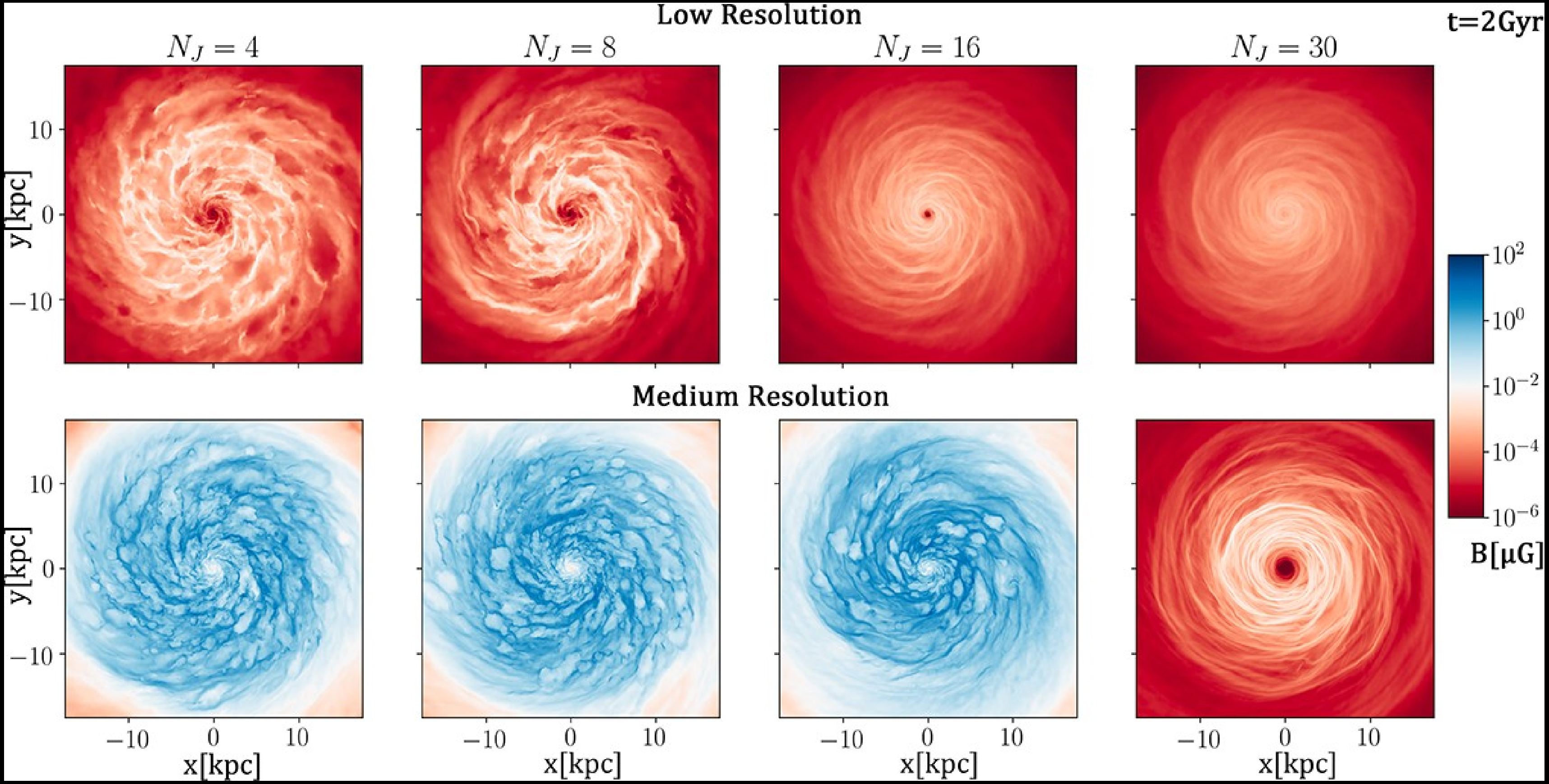
but not always in the same way (i.e., cells, particle+kernel)

Resolution in Observations + Simulations

What "resolution" means in a paper ?

Beware of such a difference when comparing obs / sim

Wissing & Shem 2023



Resolution in Observations + Simulations

What "resolution" means in a paper ?

Beware of such a difference when comparing obs / sim

Strawn et al. 2023 (AGORA – VI)

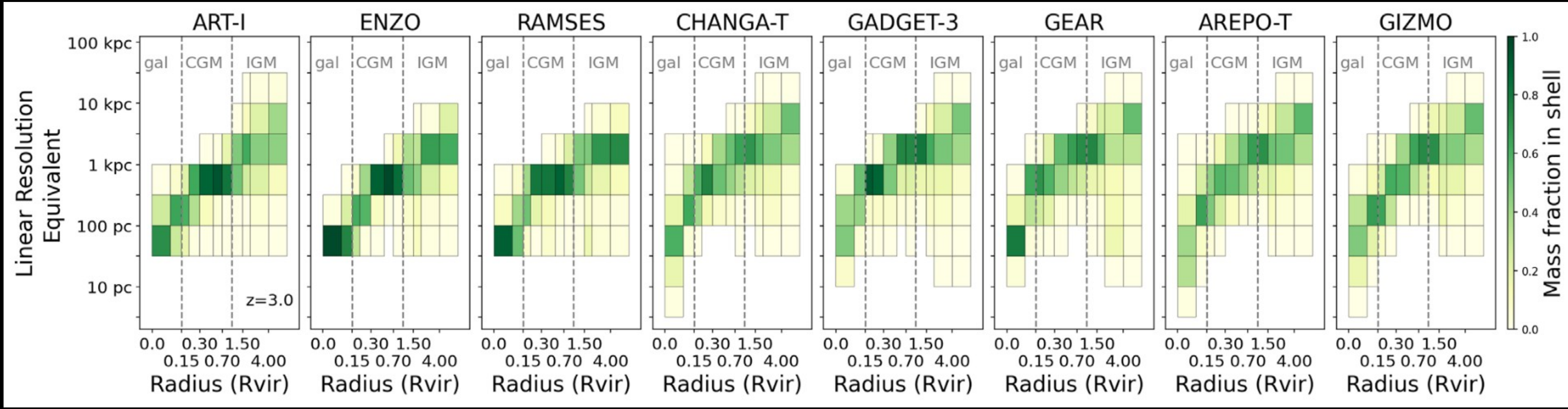


Figure 1. Resolution of all eight AGORA codes at $z = 3$. In each shell of increasing size, color shows the mass fraction contained in "linear resolution equivalent" bins of width 0.5 dex, normalized within columns. For grid and moving mesh codes, "linear resolution equivalent" is defined as cell volume raised to the $1/3$ power. For particle-type codes, it is instead defined as "effective volume" (particle mass divided by particle density) to the $1/3$ power. See Section 2.3 for more details.

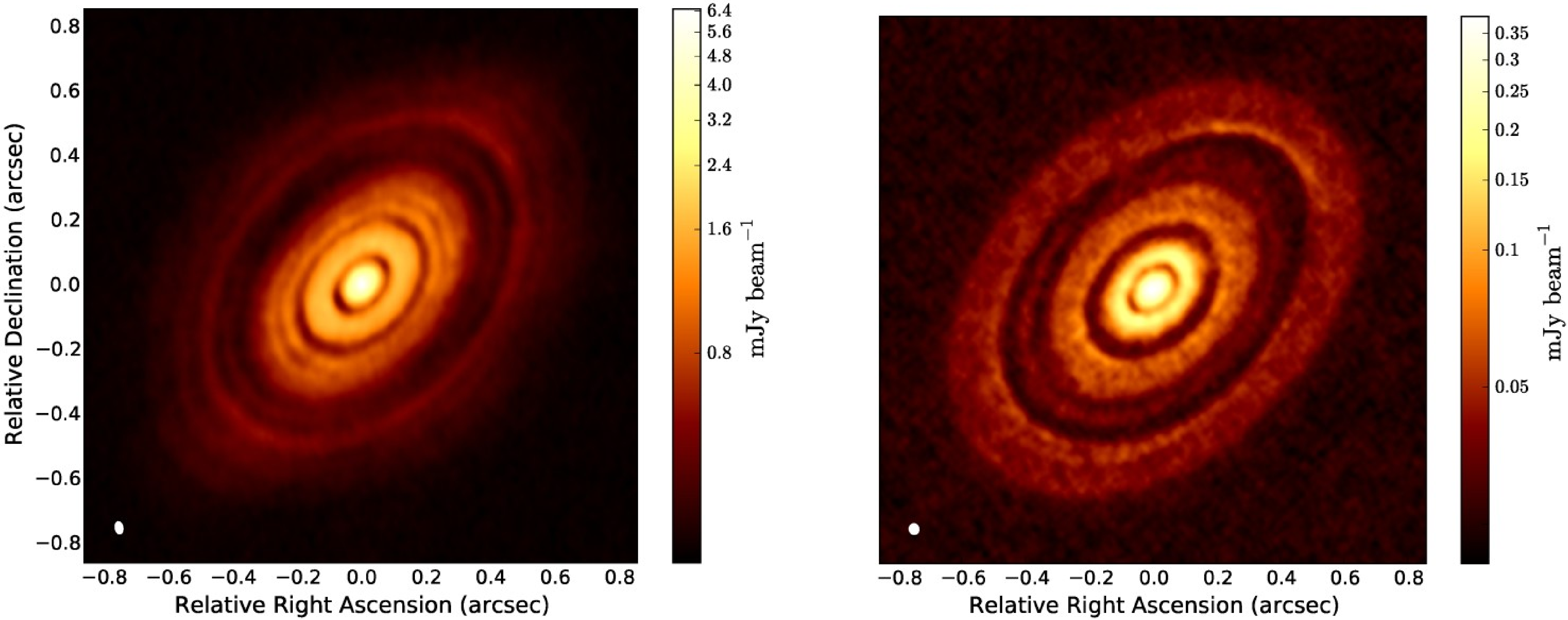
Do we learn anything from running numerical simulations?

Theory/Models versus Experiments – **Usage #1**

A Simulation

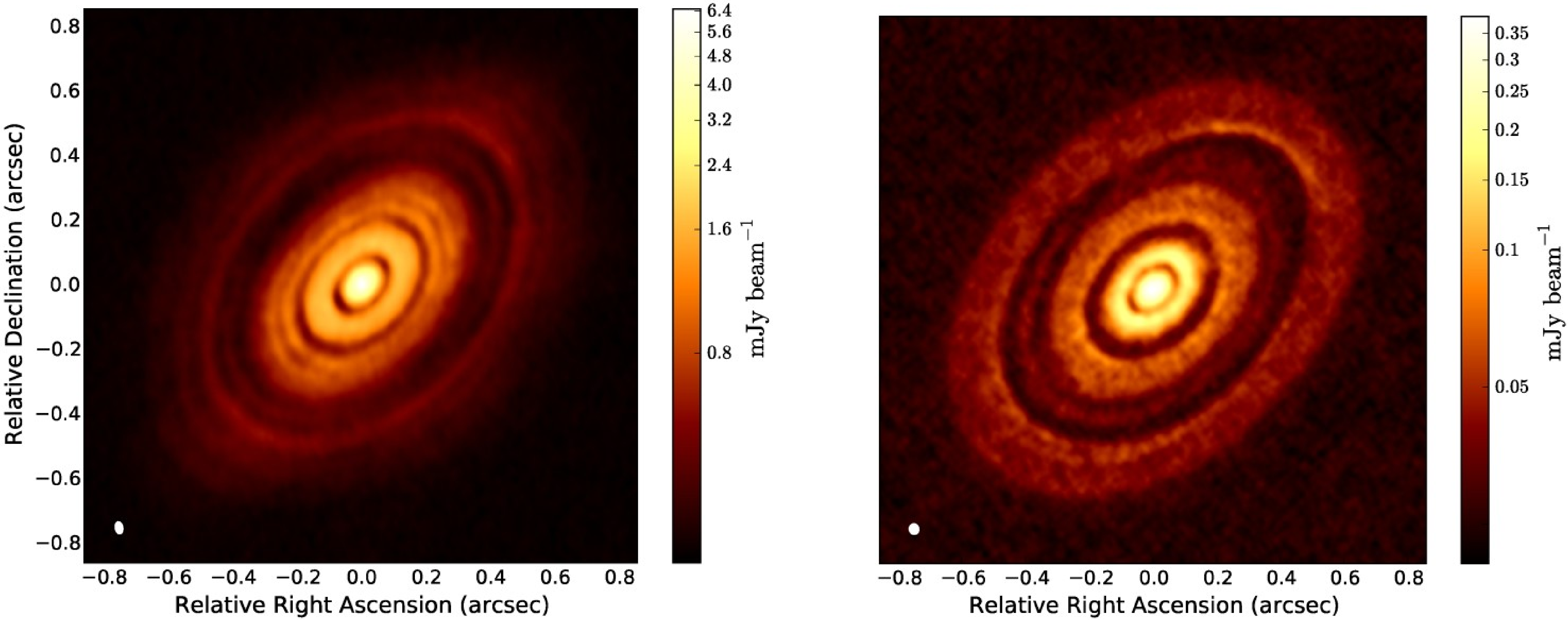
- As a consistency check for a given hypothesis
  → compare with existing data
  → Hypothesis is not inconsistent with the laws of physics "as implemented"
- **Example 2** = DiPierro et al. (2015)
- **Example 3** = Verwilghen et al. (2025)

Theory/Models versus Experiments DiPierro et al. (2015)



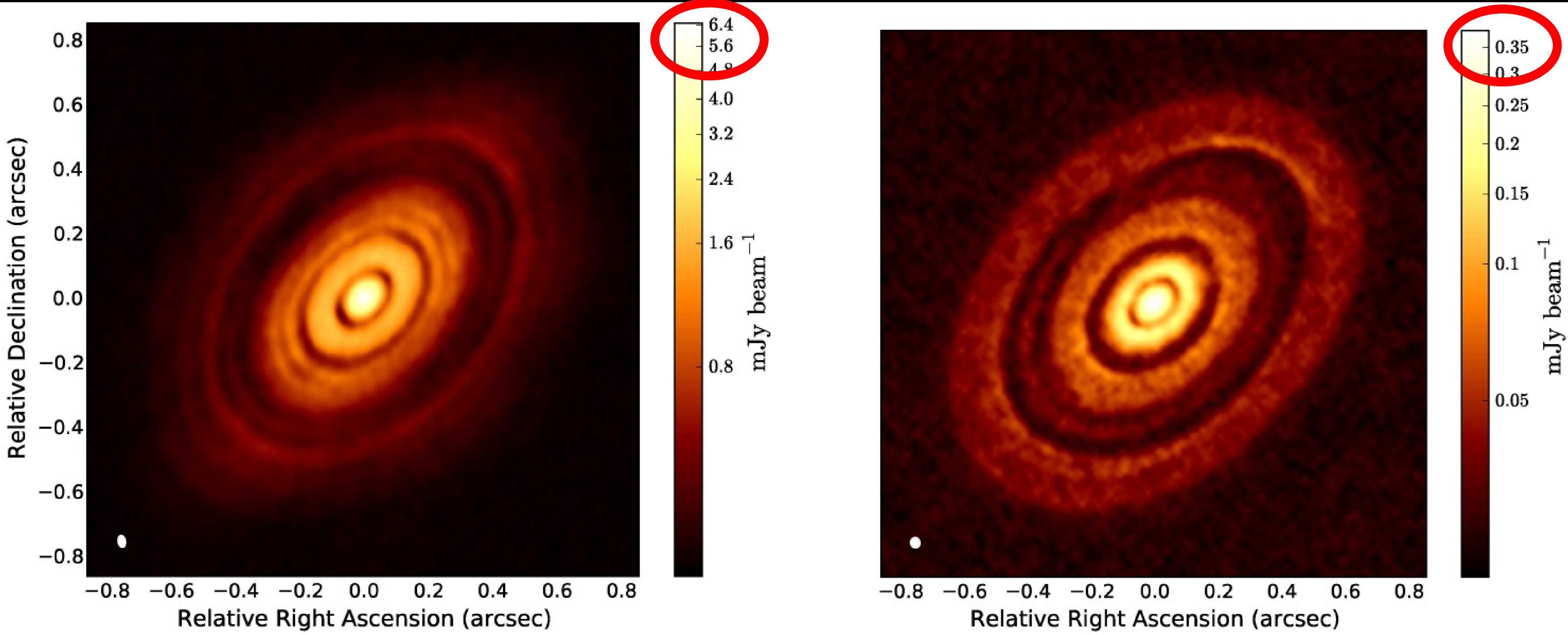
3 Saturns within a disk ?

Theory/Models versus Experiments DiPierro et al. (2015)

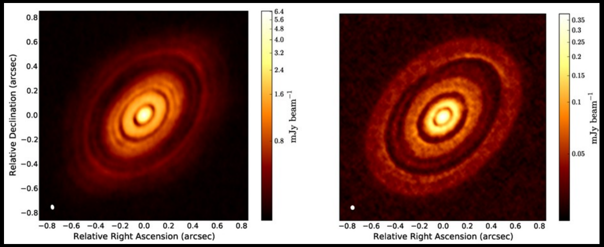


Consistent with the hypothesis that those rings are carved by Saturns

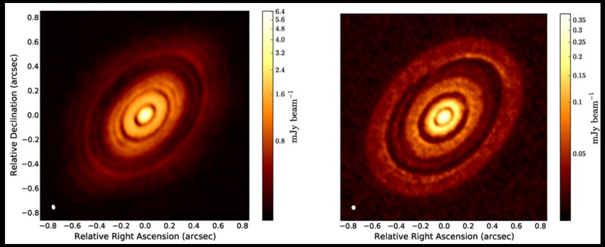
Theory/Models versus Experiments DiPierro et al. (2015)



Consistent, really ?



So: *what did we learn ?*

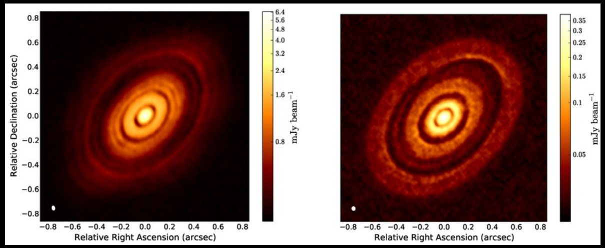


So: *what did we learn ?*

Planets **may** have carved those rings

but not a proof

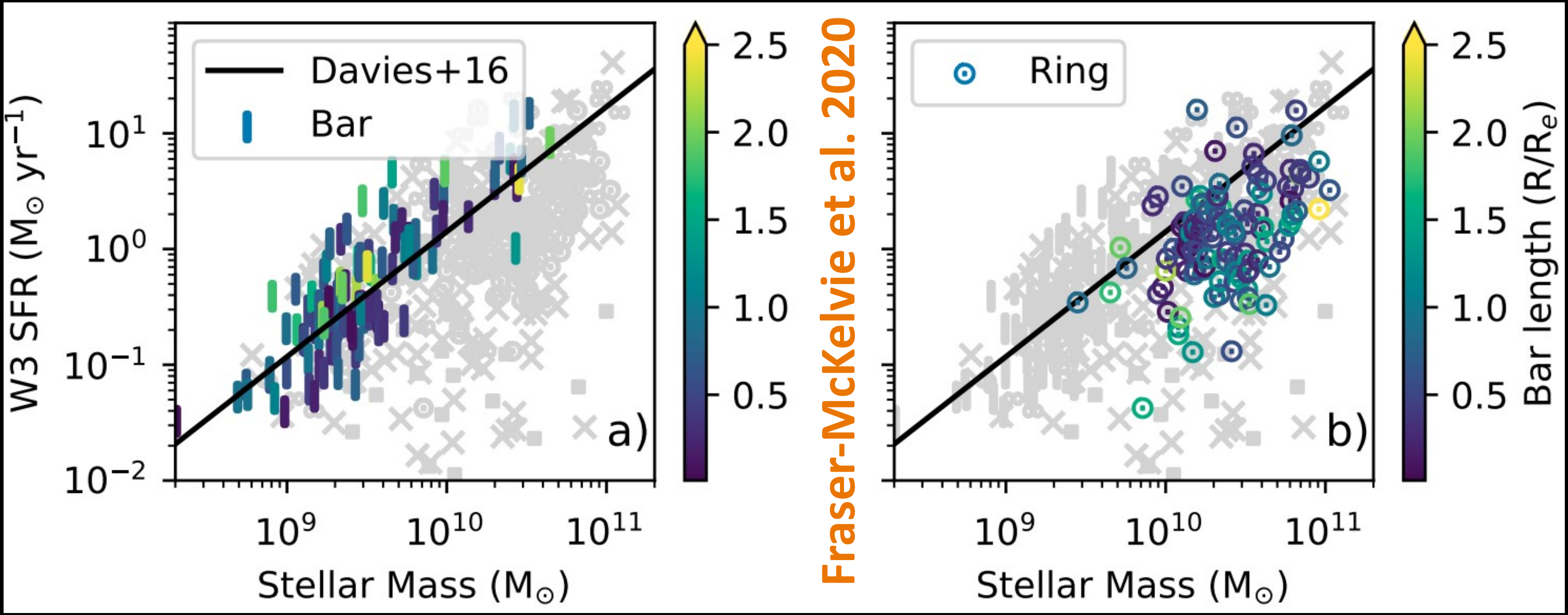
What is missing ?



So: *what did we learn ?*

Planets **may** have carved those rings

Theory/Models versus Experiments Verwilghen et al. (2025)

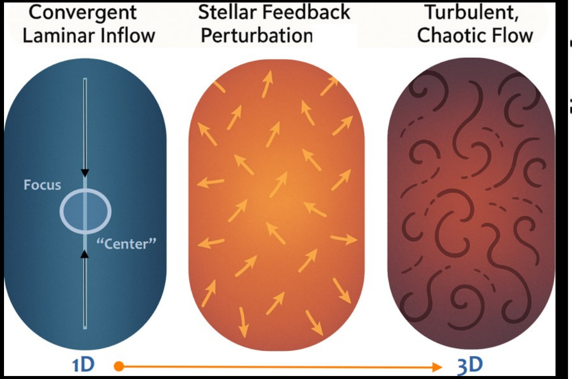


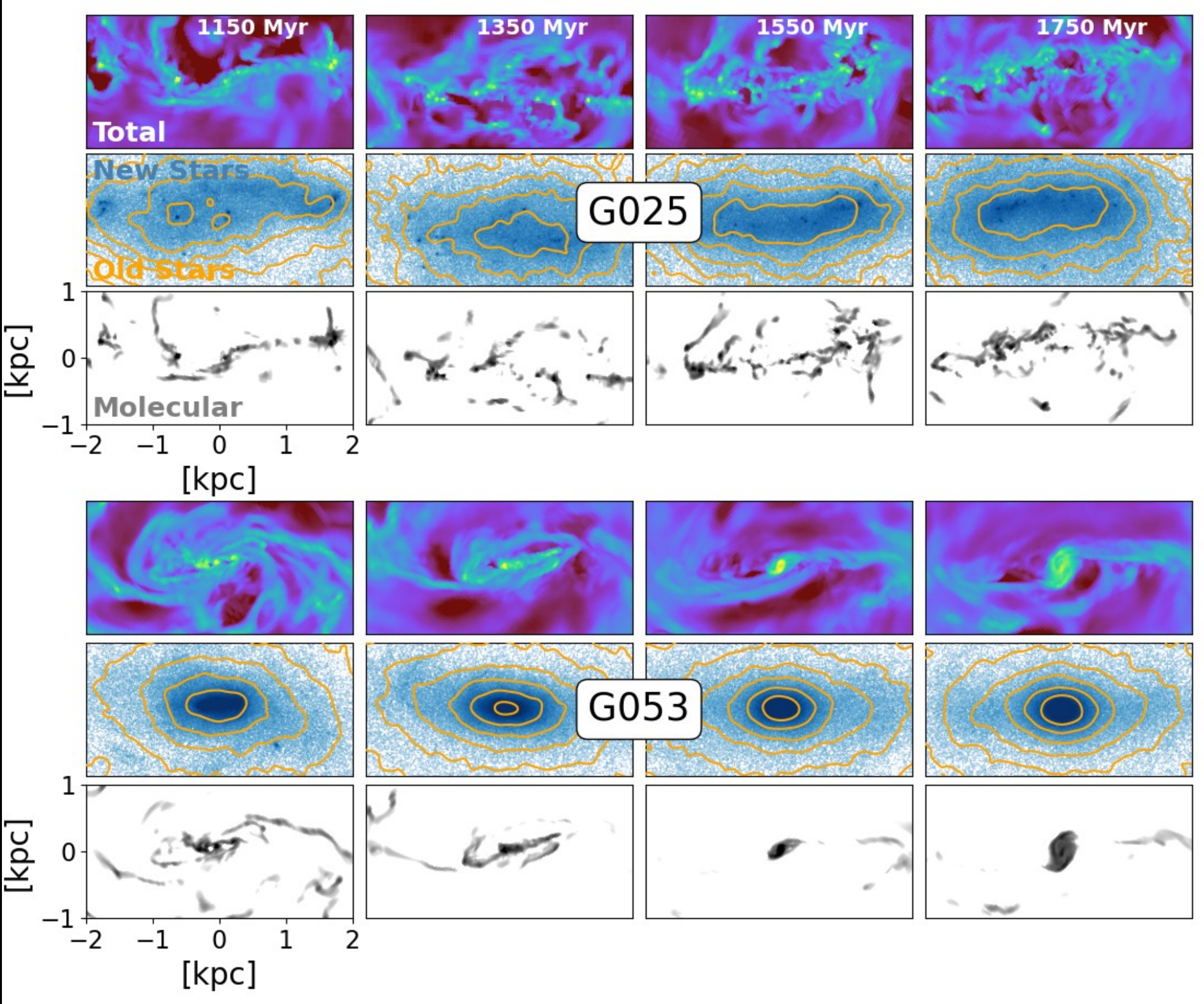
Dependence of SF on stellar mass

⇒ lower-mass = bar

⇒ higher mass = rings / central

Theory/Models versus Experiments **Verwilghen et al. (2025)**





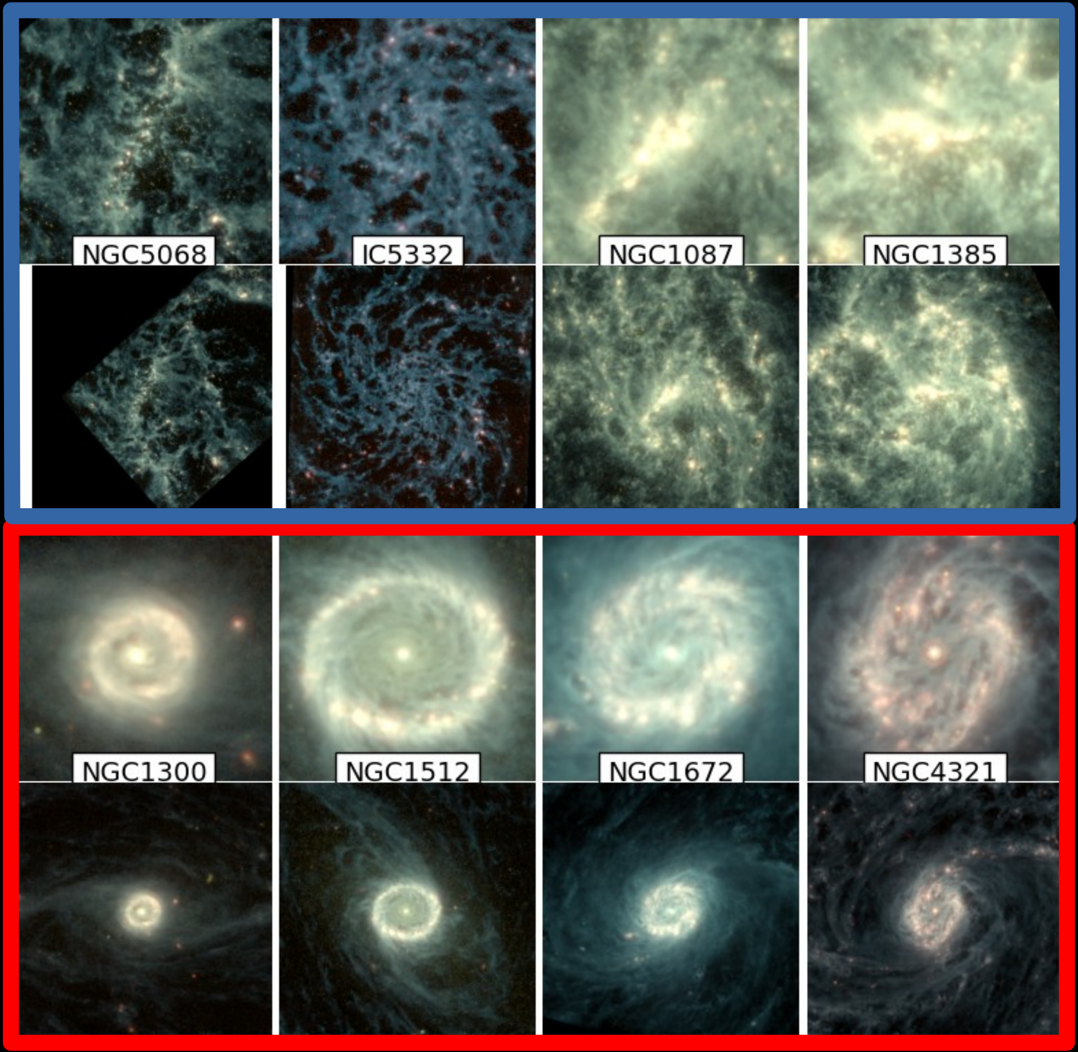
Low $M_\star$

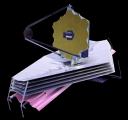
High $M_\star$

Theory/Models versus Experiments

Lower mass

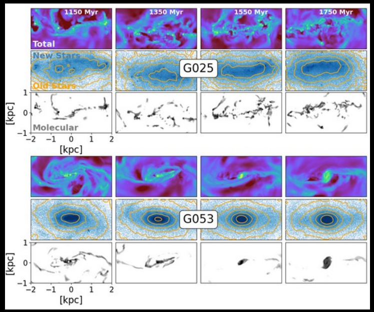
Higher mass



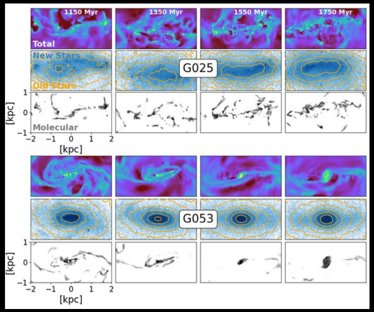


21μm-10.0μm-3μm

Credit: NASA/ESA,CSA;
PHANGS / Chown / Williams / Sutter /Emsellem



So: *what did we learn ?*

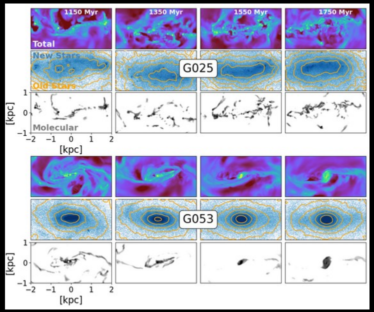


So: *what did we learn ?*

Feedback in the ISM ***may*** be responsible for a (mass-dep.) differential evolution

Not a proof

What is missing ?



So: *what did we learn ?*

Feedback in the ISM ***may*** be responsible for a (mass-dep.) differential evolution

Not a proof

Theory/Models versus Experiments – Usage #2

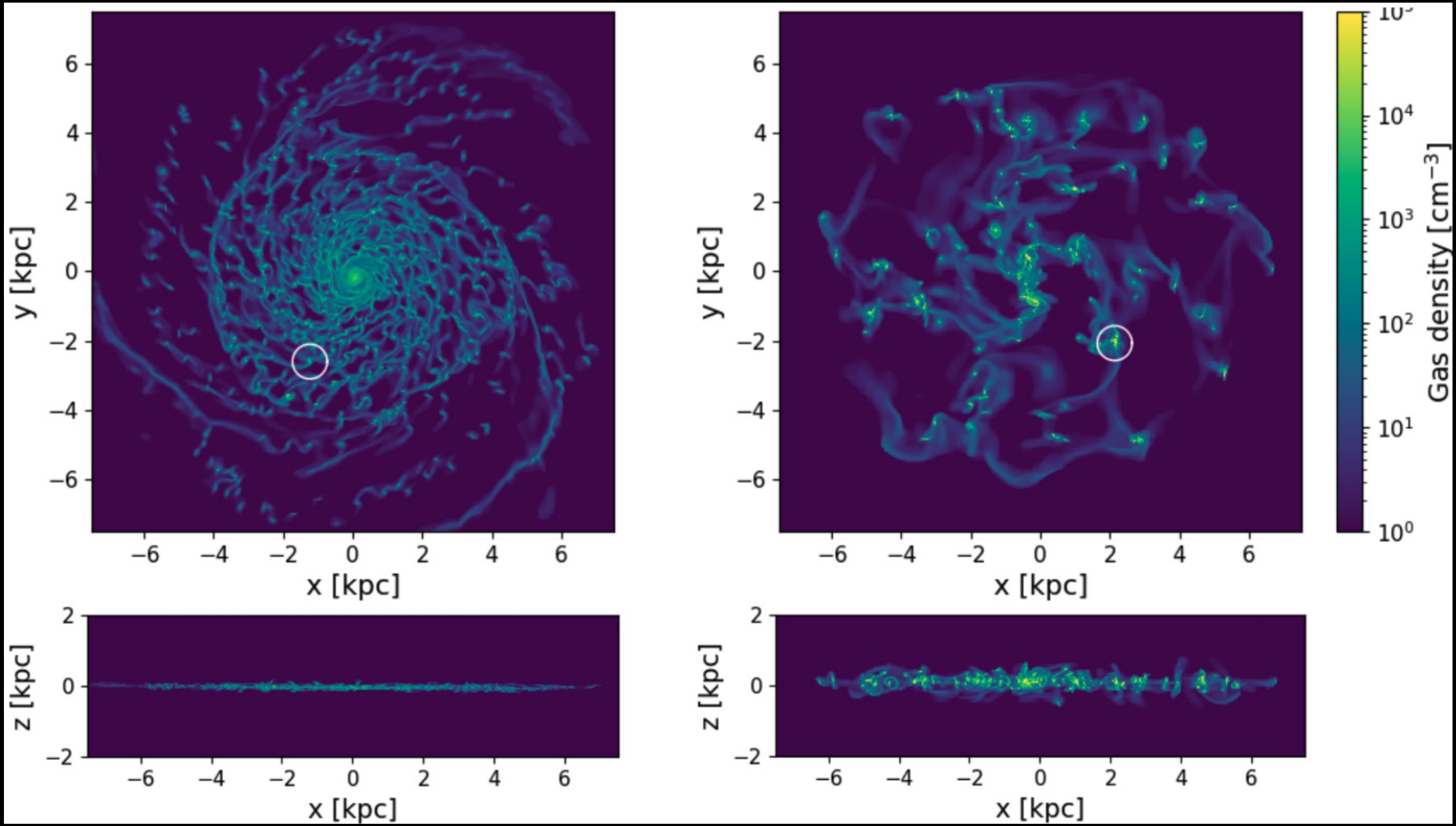
A simulation

- May provide a hint of the emergence of a complex process
- Reminder : biases in Simulations
  → Observation of a model
- May provide new predictions
  → You learn something about your model

- **Example 4** = Fensch et al. 2023

Theory/Models versus Experiments **Fensch et al. 2023**

10%

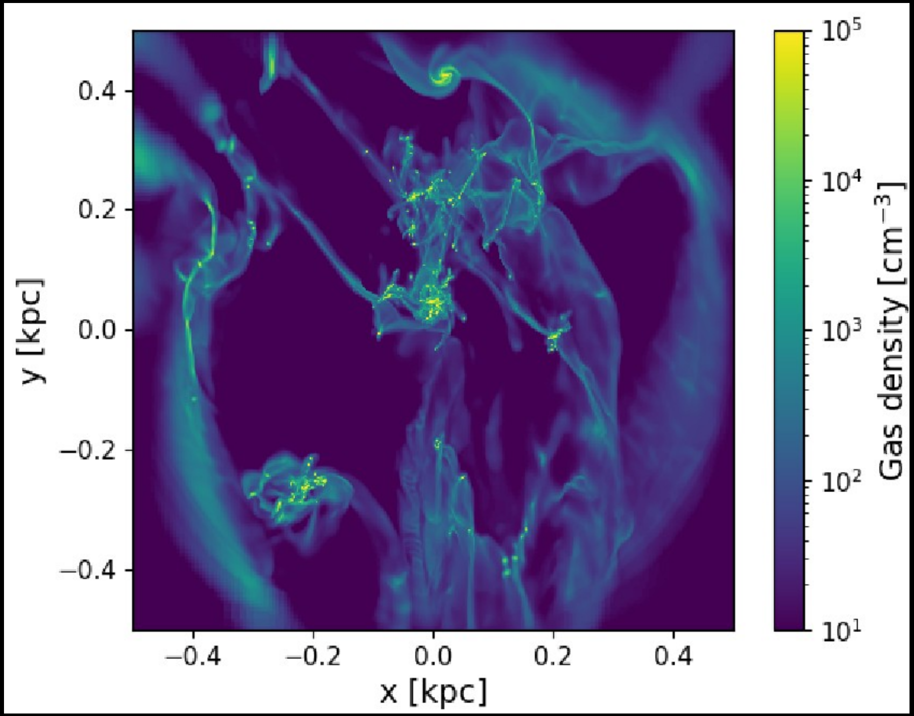


65%

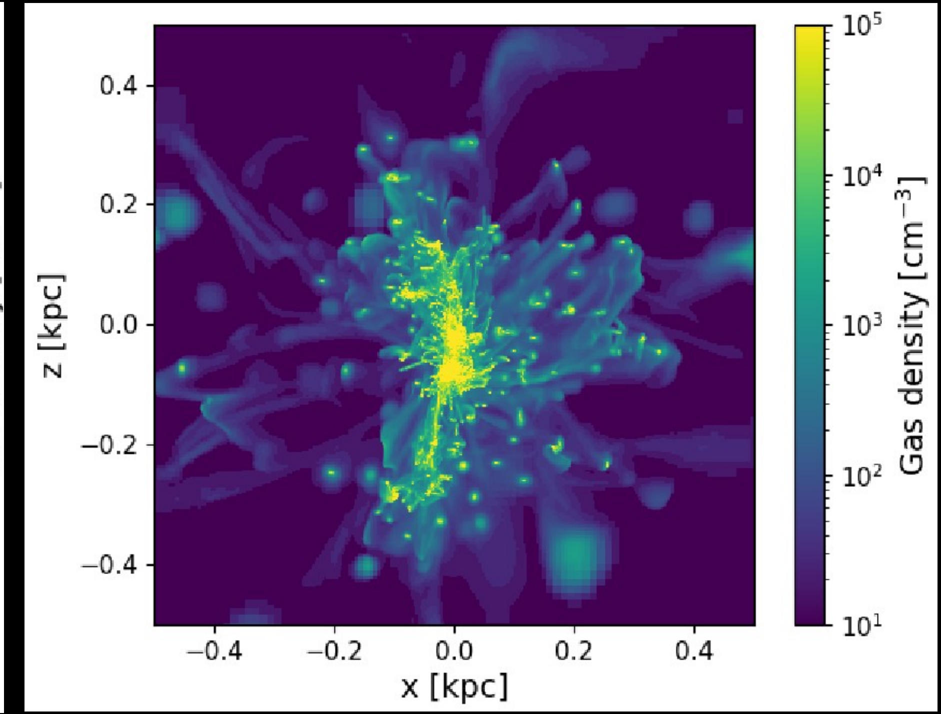
Growth of ISM structures

Theory/Models versus Experiments **Fensch et al. 2023**

10%

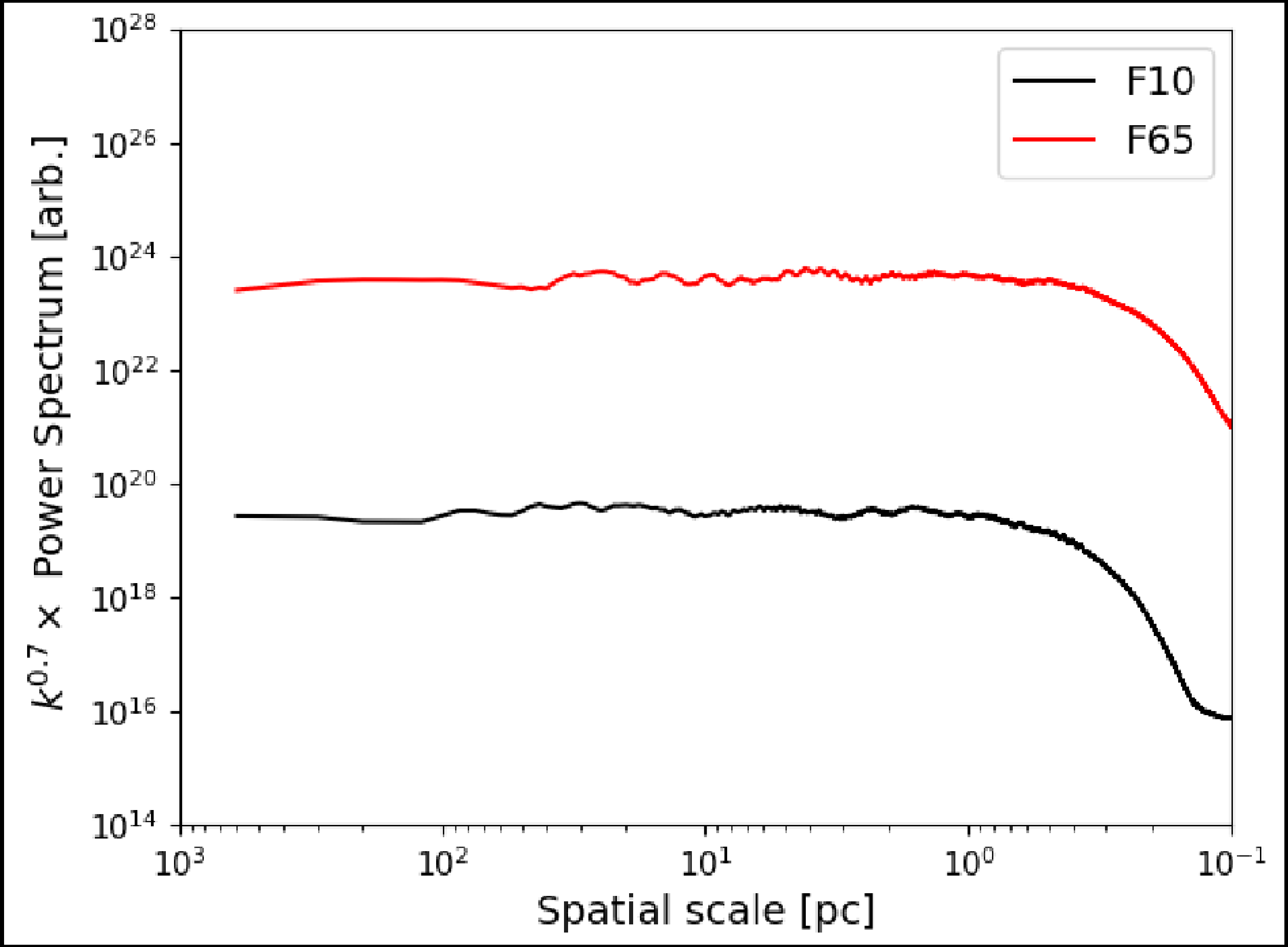


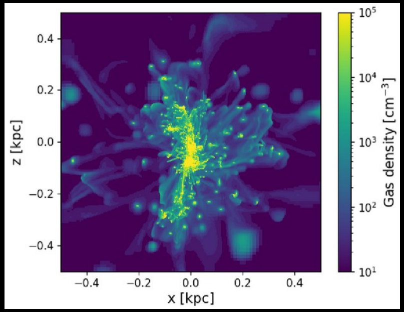
65%



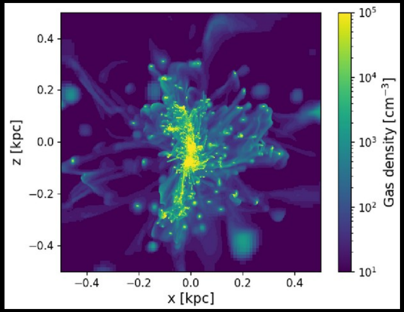
Large contrast of structures

Theory/Models versus Experiments **Fensch et al. 2023**





So: *what did we learn ?*

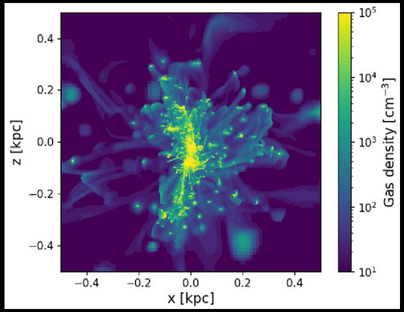


So: *what did we learn ?*

When gravity + Hydrodynamics act alone, turbulence scaling seems to be invariant

Not a proof

What is missing ?



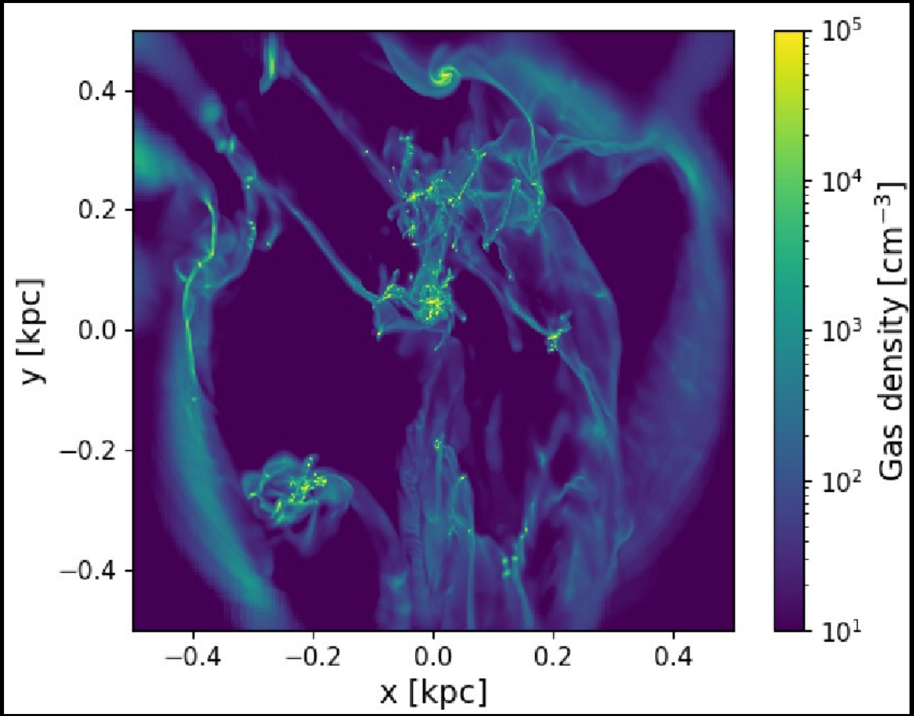
So: *what did we learn ?*

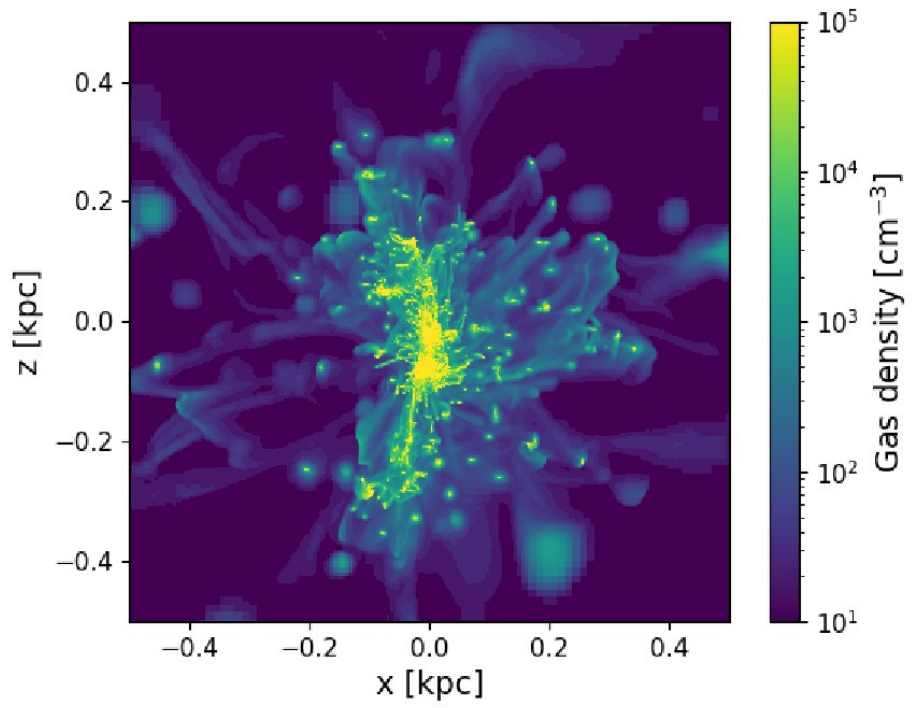
When gravity + Hydrodynamics act alone, turbulence scaling seems to be invariant

Not a proof

Theory/Models versus Experiments Fensch et al. 2023

10% **65%**

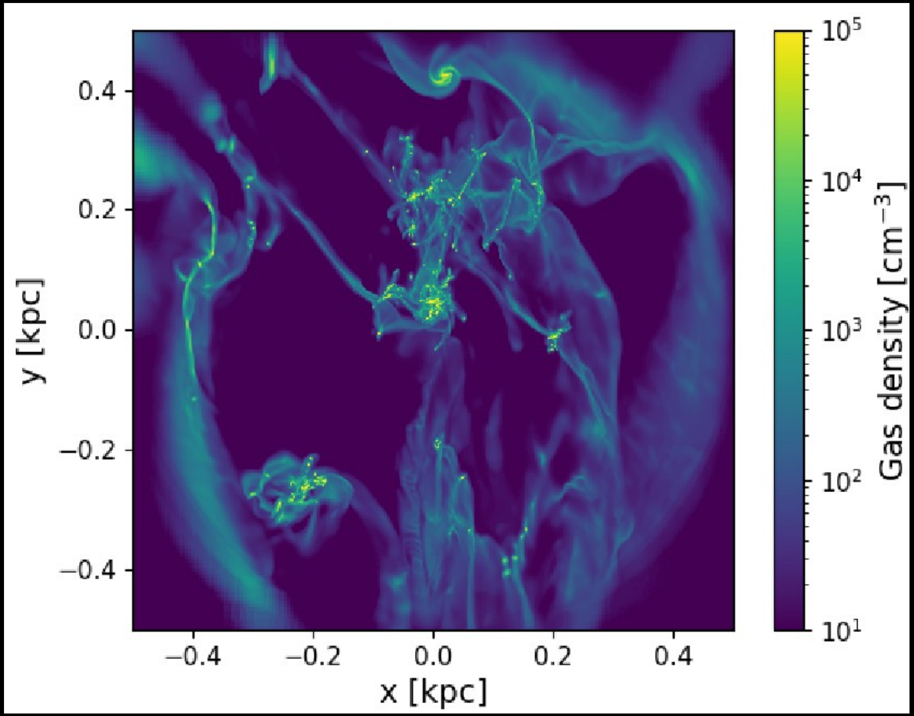




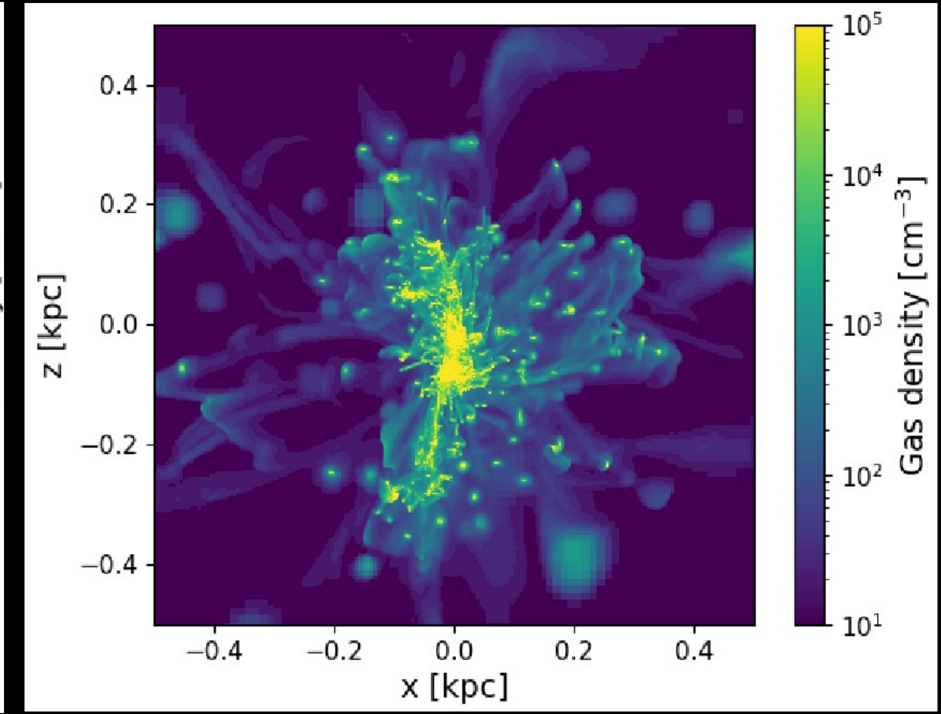
Such a set up will never be realised in nature

Theory/Models versus Experiments **Fensch et al. 2023**

10%



65%



Such a set up will never be realised in nature

But this is true for all simulations ...

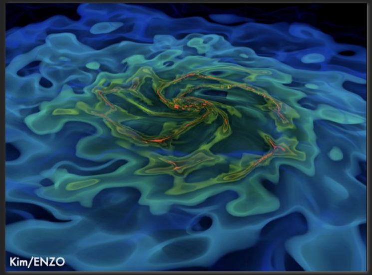
Simulations versus Simulations

AGORA

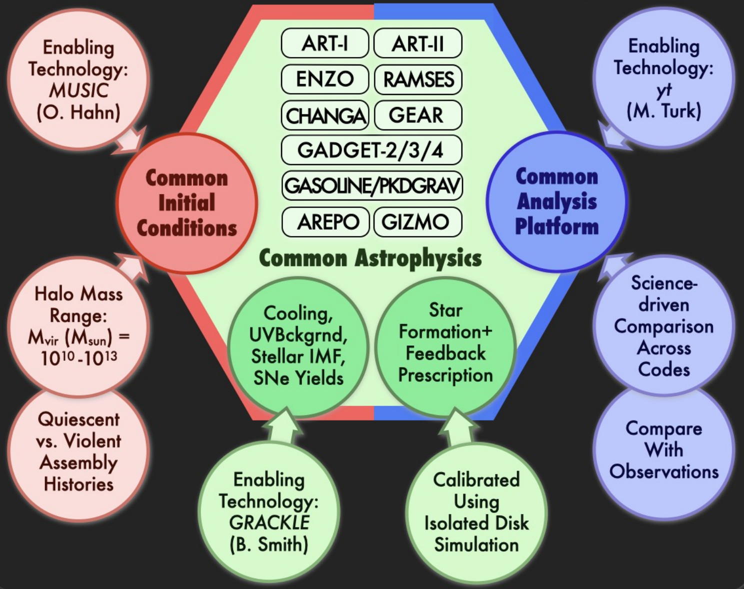
www.AGORAsimulations.org

A High-resolution Galaxy Simulations Comparison Initiative: www.AGORAsimulations.org

High-res Galaxy Simulations



AGORA Comparison Infrastructure



AGORA Goal & Team

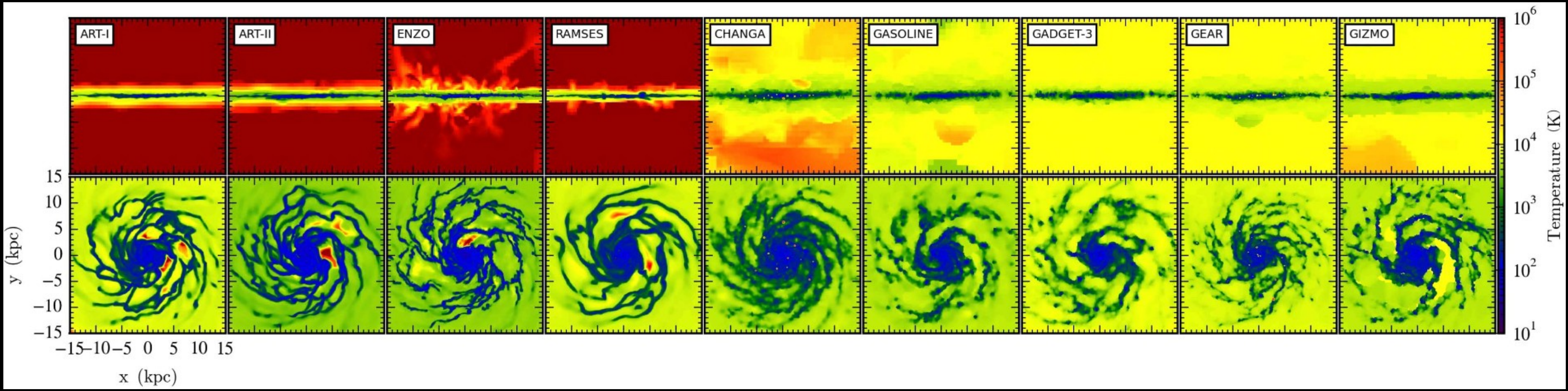
- GOAL: A collaborative, multi-platform study to raise the realism and predictive power of galaxy formation simulations
- TEAM: 160+ participants from 60+ institutions worldwide, representing 10+ codes as of 2024
- DATA SHARING: Simulations outputs and analysis softwares will be shared with the community

- Paper I and II (J. Kim et al. 2014, 2016), Paper III and IV (S. Roca-Fabrega et al. 2021, 2024), Paper V (M. Jung et al. 2024), Paper VI (C. Strawn et al. 2024)

Simulations versus Simulations

www.AGORAsimulations.org

Isolated disks

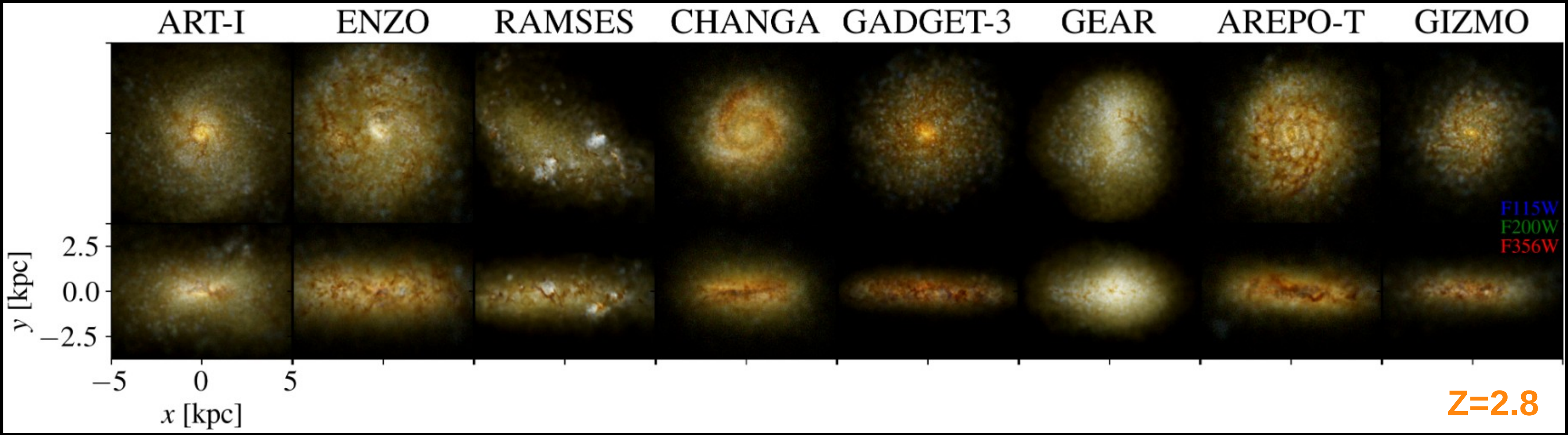


Kim et al. 2016 – AGORA II

Simulations versus Simulations

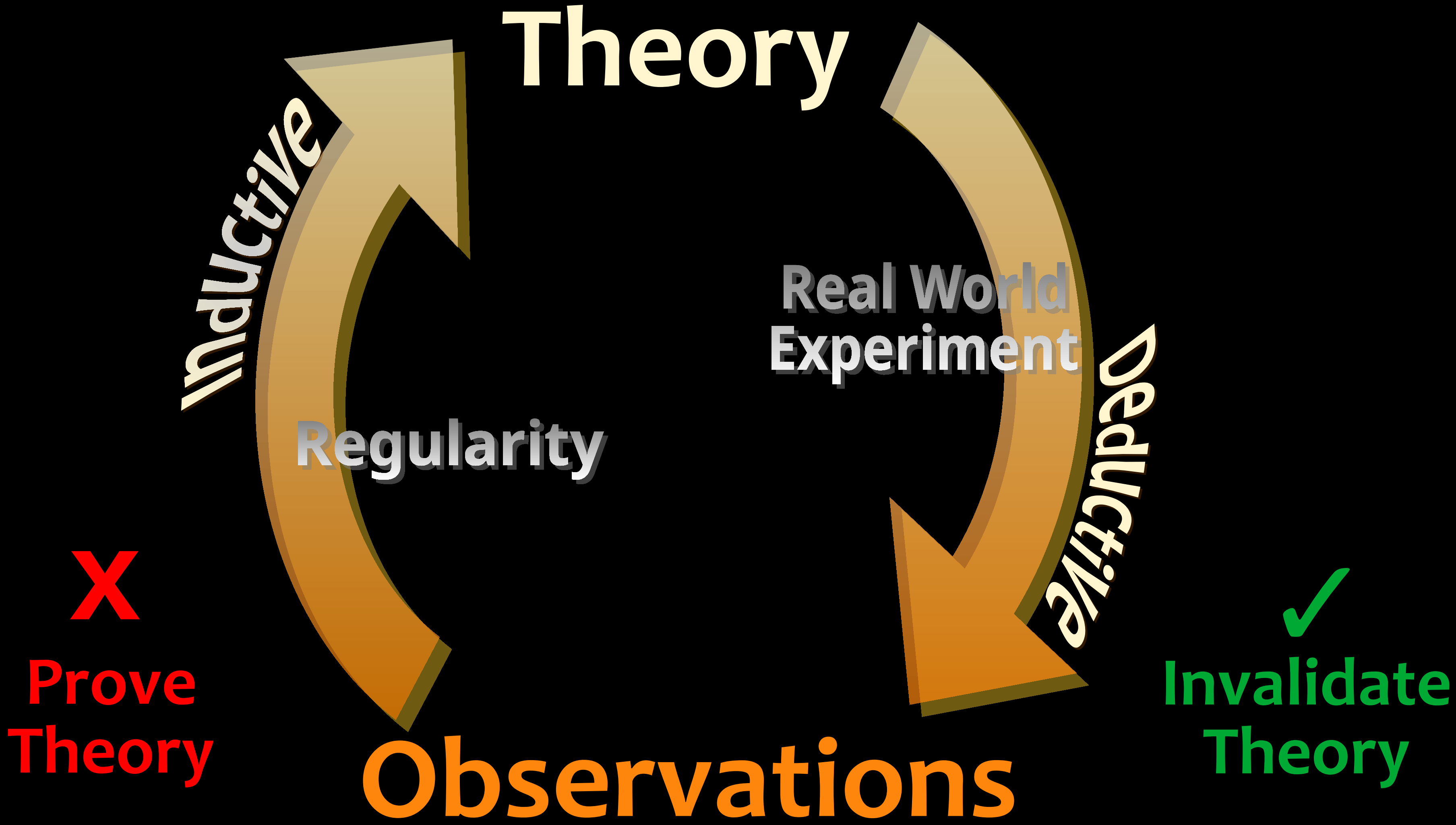
www.AGORAsimulations.org

Cosmological zoom-in / Milky-Way mass progenitors



Jung et al. 2025 – AGORA VIII

Guess why you see such a variety of morphologies, states ?


Theory
Inductive
Regularity
Real World
Experiment
Deductive
X
Prove
Theory
Observations
✓
Invalidate
Theory

Theory/Models versus Experiments

The confirmation bias

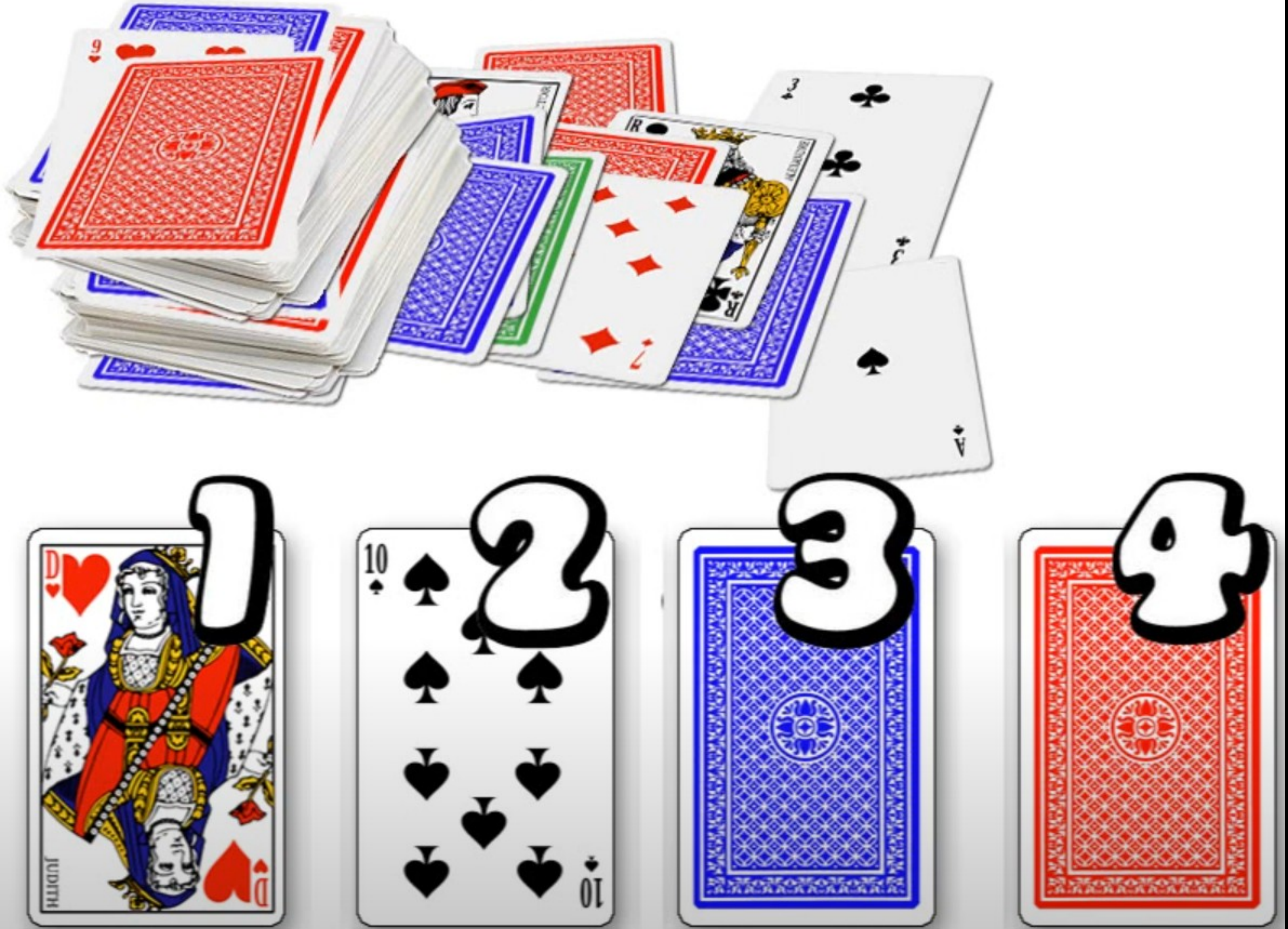
- Evolution may have favoured (wrong) deductive thinking
- We often try to "confirm" something
- We design experiments
  - Often to **confirm** our priors
- And as long as we haven't found a flaw
  - We are happy…

The Universe has disks → Implement state-of-the-art sub-grid recipes → Calibrate on z=0 galaxy populations → Simulations have disks

A small experiment

Theory

If a card is a **Queen,**
The other side is **blue**



Theory

If a card is a **Queen,**
The other side is **blue**

Theory

A : If a card is a **Queen,**
→ X The other side is **blue**

Theory

A : If a card is a **Queen,**
→ X The other side is **blue**

If A → X
And : A
Then : X

Theory

A : If a card is a **Queen,**
→ X The other side is **blue**

If A → X
And : A
Then : X

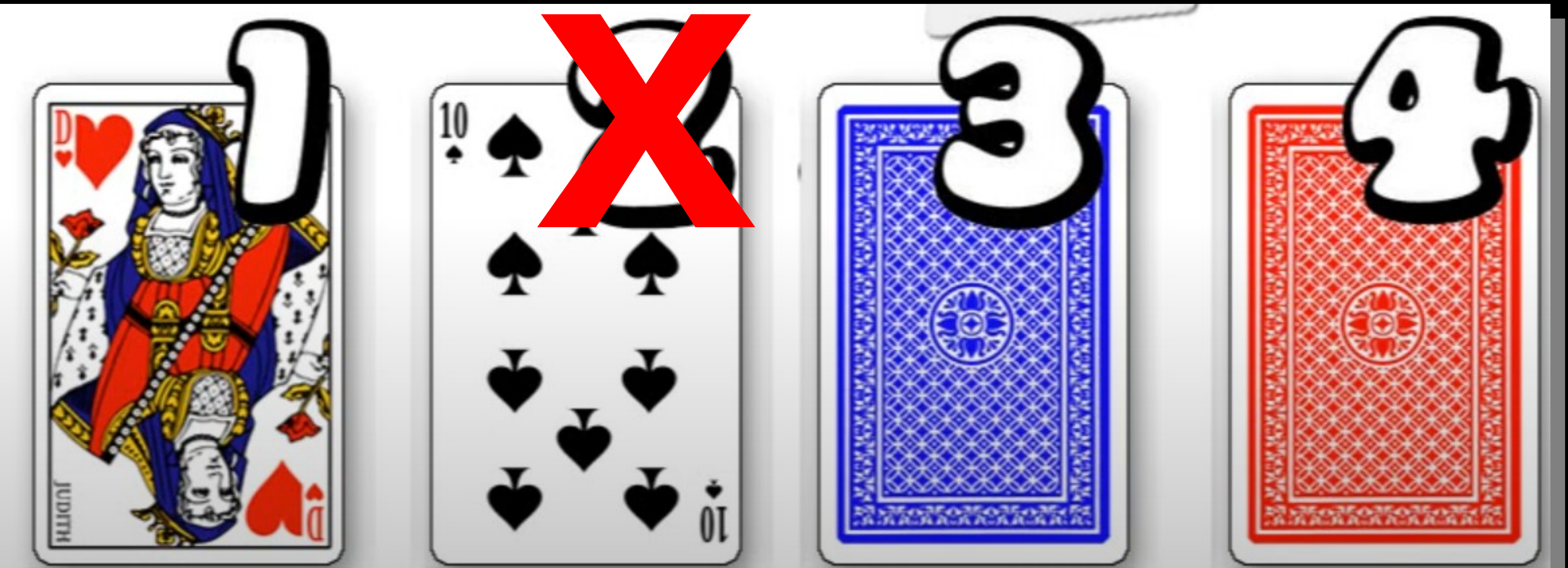
If A → X
And if : no-A
Then : no-X

Theory

A : If a card is a **Queen**,
→ X The other side is **blue**

If A → X
And : A
Then : X

If A → X
And if : no-A
Then : no-X



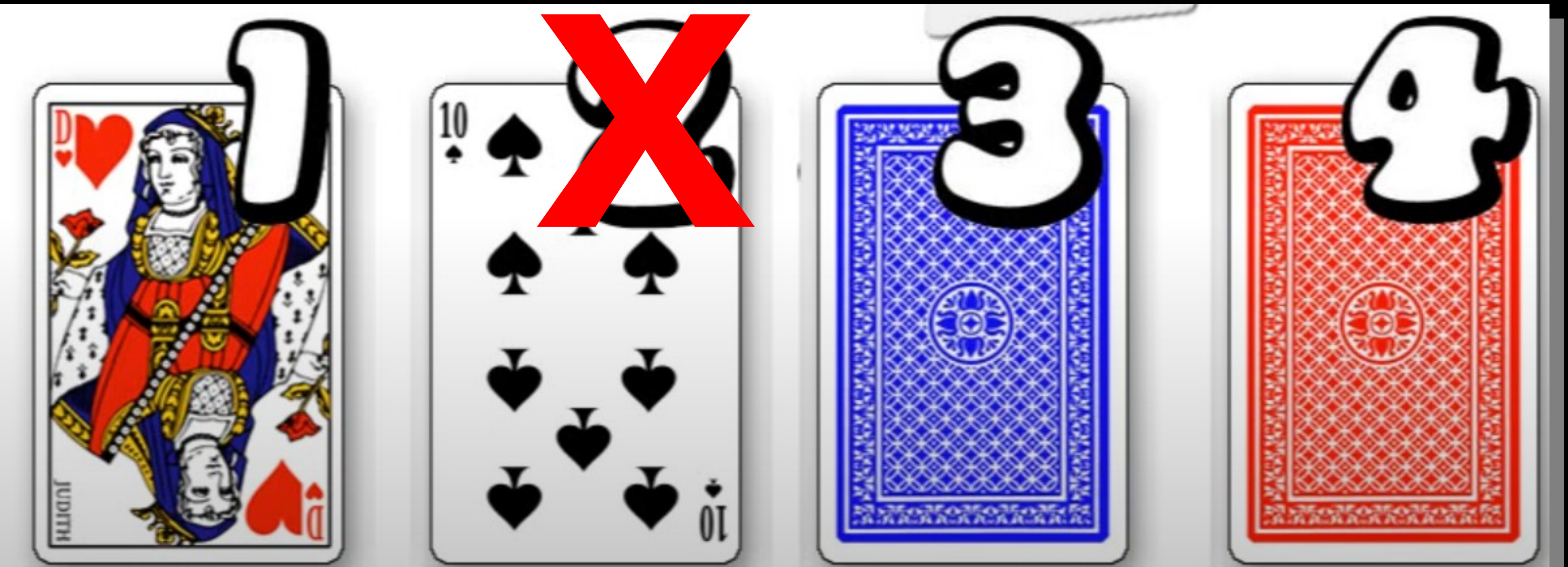
Theory

A : → X — If a card is a **Queen**, The other side is **blue**

If A → X
And : A
Then : X

If A → X
And if : no-A
Then : no-X

If A → X
And if : X
Then : A



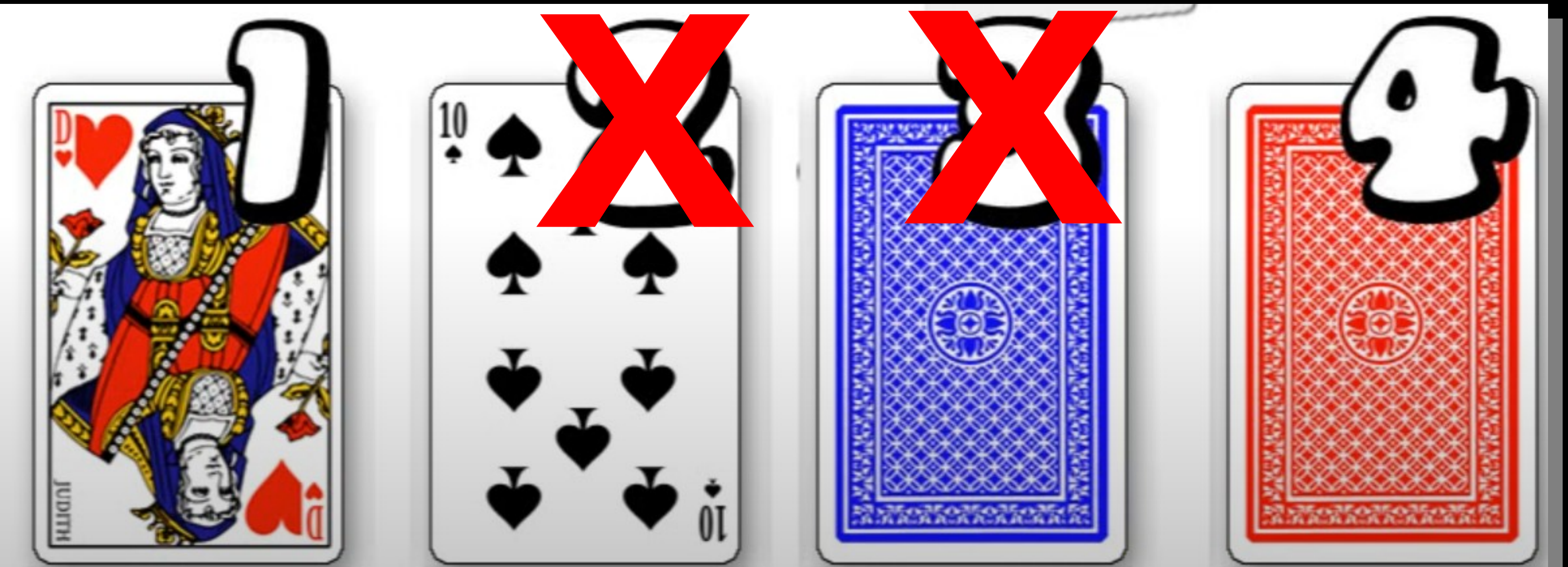
Theory

A : If a card is a **Queen**,
→ X The other side is **blue**

If A → X
And : A
Then : X

If A → X
And if : no-A
Then : no-X

If A → X
And if : X
Then : A



Theory

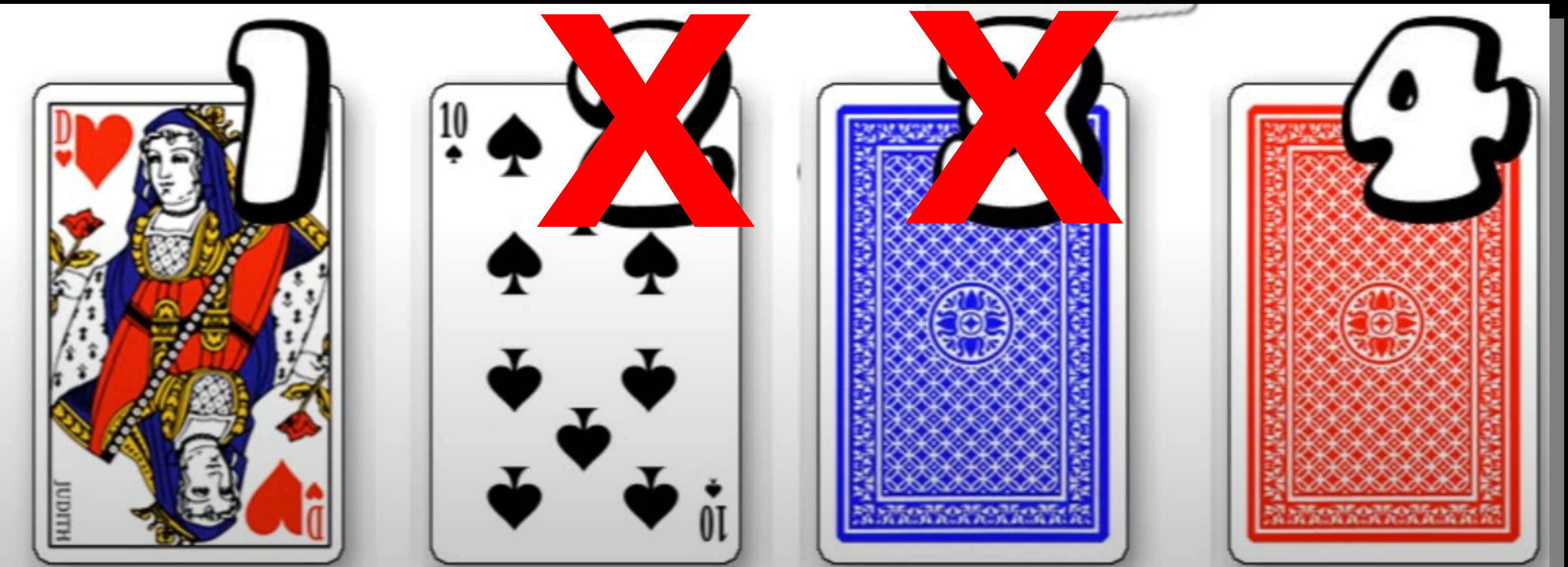
A : If a card is a **Queen**,
→ X The other side is **blue**

If A → X
And : A
Then : X

If A → X
And if : no-A
Then : no-X

If A → X
And if : X
Then : A

If A → X
And if : no-X
Then : no-A



Fallacy #1 and #2

If A → X
And if : no-A
Then : no-X

Fallacy of the inverse

- Star formation leads to turbulence in the ISM
- *If I switch off SF in my simulation, …*

If A → X
And if : X
Then : A

Fallacy of the converse

- I form disks in my simulations
- *My model is a good model of the Universe*

If A → X
And if : no-X
Then : no-A

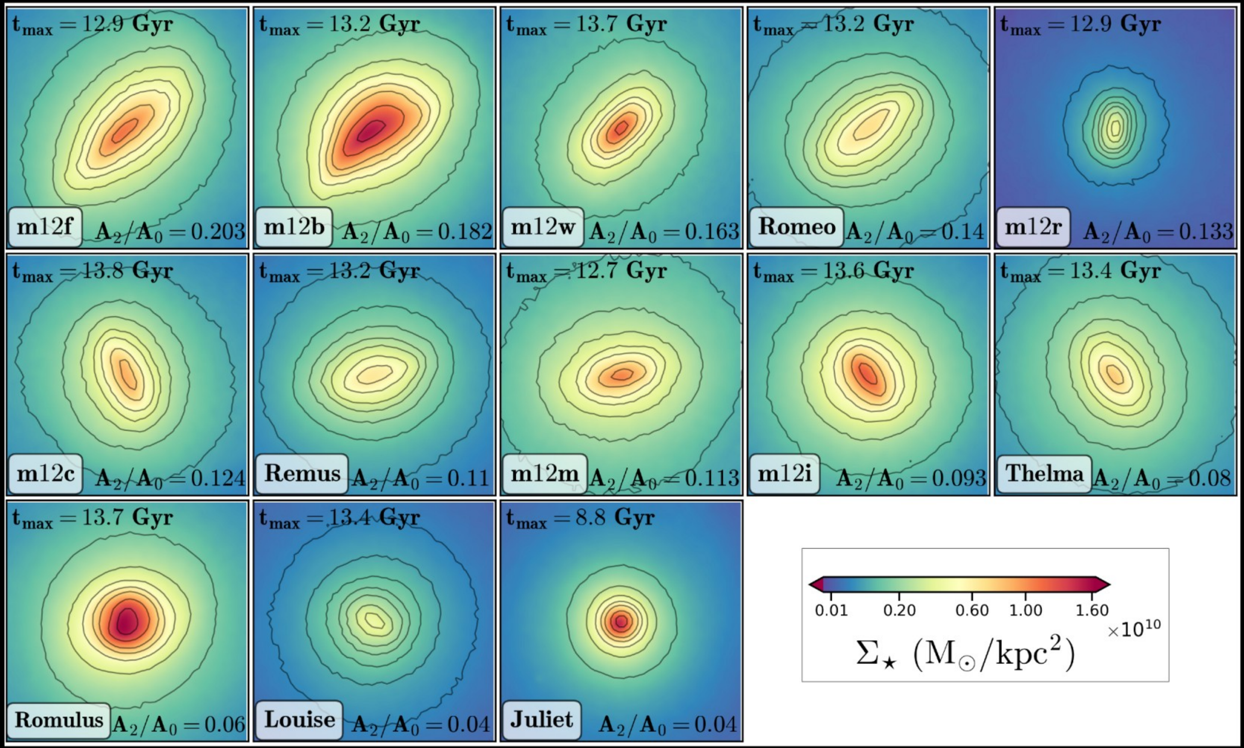
BUT

- If I do not form disks in my simulations
- *My simulation is not a good model of the Universe*

Example 5: Bar formation in cosmological simulations

Ansar et al. 2025

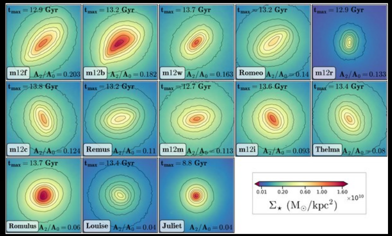
- Cosmological simulations of MWs
- FIRE-2 simulations



Example 5: Bar formation in cosmological simulations

Ansar et al. 2025

- Cosmological simulations of MWs
- FIRE-2 simulations



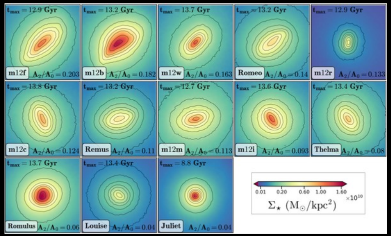
Abstract

The physical mechanisms responsible for bar formation and destruction in galaxies remain a subject of debate. While we have gained valuable insight into how bars form and evolve from isolated idealized simulations, in the cosmological domain, galactic bars evolve in complex environments, with mergers and gas accretion events occurring in the presence of the turbulent interstellar medium with multiple star formation episodes, in addition to coupling with their host galaxies' dark matter halos. We investigate the bar formation in 13 Milky Way–mass galaxies from the Feedback in Realistic Environments (FIRE-2) cosmological zoom-in simulations. 8 of the 13 simulated galaxies form bars at some point during their history: three from tidal interactions and five from internal evolution of the disk. The bars in FIRE-2 are generally shorter than the corotation radius (mean bar radius ~1.53 kpc), have a wide range of pattern speeds (36–97 km s^{-1} kpc^{-1}), and live for a wide range of dynamical times (2–160 bar rotations). We find that the bar formation in FIRE-2 galaxies is influenced by satellite interactions and the stellar-to-dark-matter mass ratio in the inner galaxy, but neither is a sufficient condition for bar formation. Bar formation is more likely to occur, with the bars formed being stronger and longer-lived, if the disks are kinematically cold; galaxies with high central gas fractions and/or vigorous star formation, on the other hand, tend to form weaker bars. In the case of the FIRE-2 galaxies, these properties combine to produce ellipsoidal bars with strengths $A_2/A_0 \sim 0.1$–0.2.

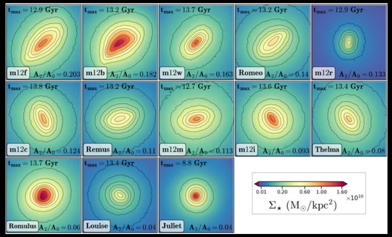
Example 5: Bar formation in cosmological simulations

Ansar et al. 2025

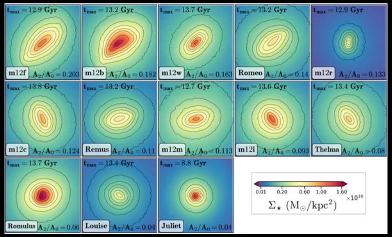
- Cosmological simulations of MWs
- FIRE-2 simulations



et al. 2013) and has number density $n > 1000\ \mathrm{cm}^{-3}$. The FIRE simulations are able to produce disk galaxies with masses, scale radii, and scale heights that are comparable to observed MW-mass galaxies (X. Ma et al. 2017; S. Garrison-Kimmel et al. 2018; R. E. Sanderson et al. 2020; S. Yu et al. 2021; J. Gensior et al. 2023; A. B. Gurvich et al. 2023) and also realistic Giant Molecular Cloud populations (D. Guszejnov et al. 2020; S. M. Benincasa et al. 2020). Importantly for this study, the kinematic "coldness" of the stellar and gas disks of FIRE-2 galaxies has recently been shown by F. McCluskey et al. (2024) to be consistent with observed galaxies, agreeing well with the measurements of M31, M33, and galaxies from the PHANGS survey (J. Sun et al. 2020; I. Pessa et al. 2023), with the MW being somewhat kinematically cold relative to this population. The FIRE-2 model does not include feedback from black hole accretion.



So: *what did we learn ?*



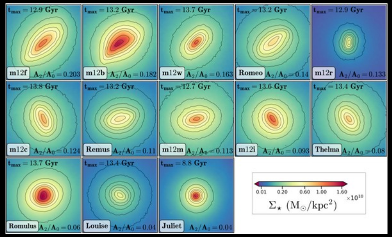
So: *what did we learn ?*

Weak bars form in the FIRE-2 zoom-in setup

Evolution driven by various parameters

Not conclusive for our Universe

What is missing ?



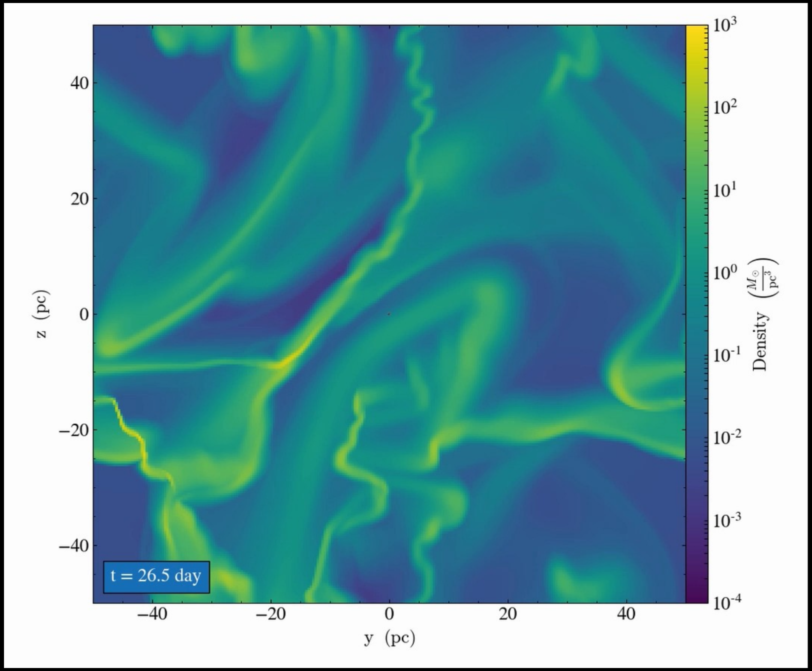
So: *what did we learn ?*

Weak bars form in the FIRE-2 zoom-in setup
Evolution driven by various parameters
Not conclusive for our Universe

Example 6: Supernova explosion

Olhlin, Renaud, Agertz 2019

- SN bubbles in a turbulent medium
- RAMSES code (AMR)
- 100 pc box with 0.4 pc cells
- uniform density of 100 cm^{-3}/10K



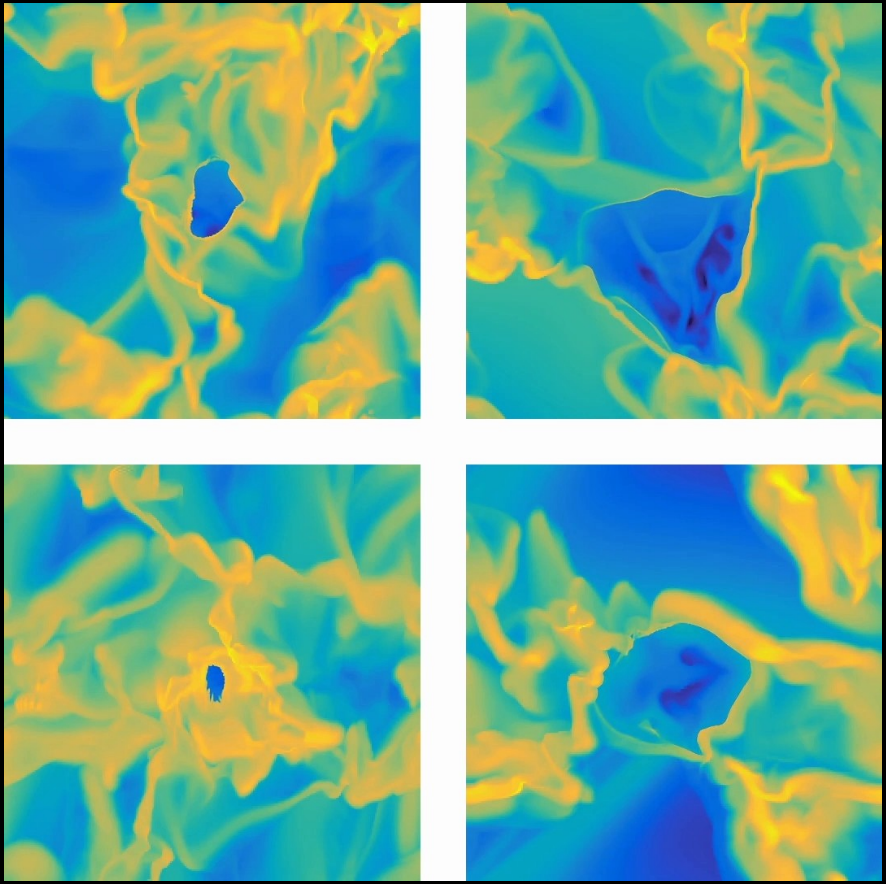
Example 6: Supernova explosion

Olhlin, Renaud, Agertz 2019

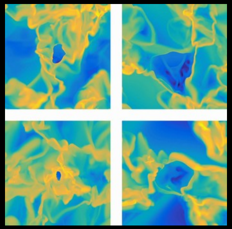
- SN bubbles in a turbulent medium
- RAMSES code (AMR)
- 100 pc box with 0.4 pc cells
- uniform density of 100 cm^{-3}/10K

→ Variance

- Set by the turbulence seed



What is missing ?

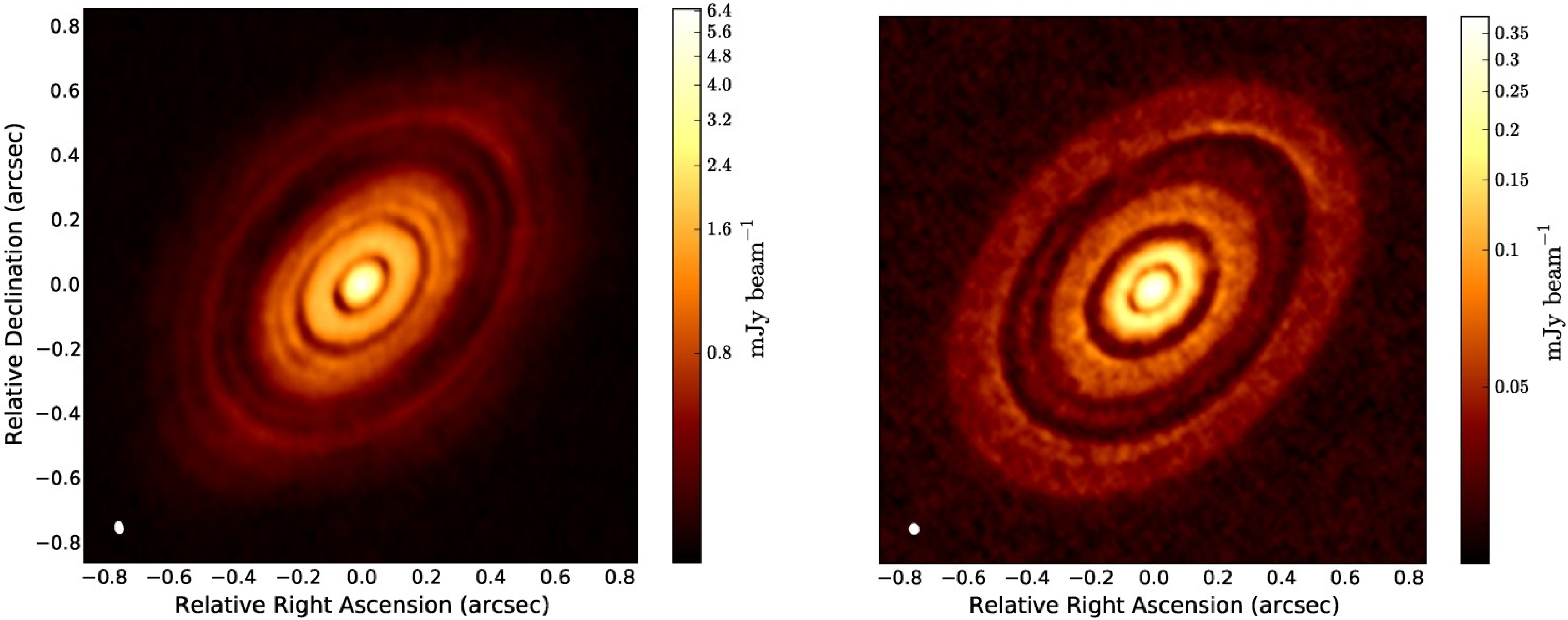


So: *what did we learn ?*

Large variations in the impact of SNe;
Spheres: not a good model
Scary for Sims; No easy extrapolations

The need
to define
(proper) Metrics

Theory/Models versus Experiments DiPierro et al. (2015)



What is good / What is not good?

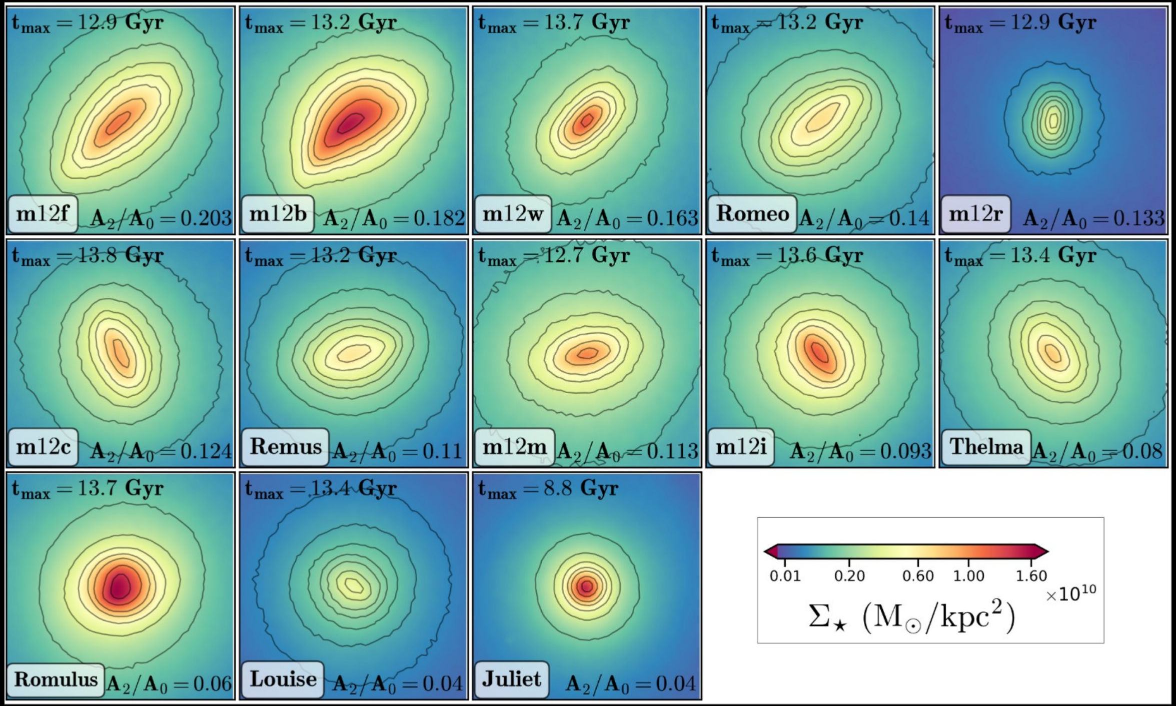
Example 5: Bar formation in cosmological simulations

Ansar et al. 2025

- Bars defined as A2 > 0.1 !
- Are those relevant bars?
- Are those the right population?

Observations

- Bars defined as A2 > 0.2
- Are those relevant metrics?
- Do we know the population?



Fallacy #3

Ambiguity

- Fallacy of the Misplaced concreteness
- Connected with the Authority Bias
- Don’t use the measurement/goals as the new metric

→ see also Goodhart’s and Campbell’s laws
→ Citation impact in journals

I have “bars” (or disks)

- May be beautiful
- May be seeded by complexity
- May be an achievement

→ But not a validation

One more example

A brain switch

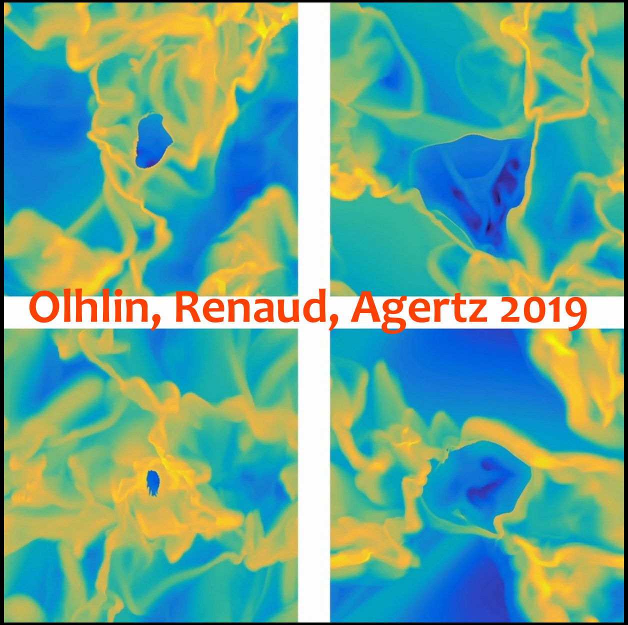

IRON
STARS
GAS
DARK MATTER
Z = 0.8
7.2 GYR AGO
Amazing isn't it ?

This is an AI-generated film (thanks to Renaud)

No it's not...

Of the Importance
of
Coherence

Example 6: Supernova explosion



At low-er resolution

- May be expressed by varying outcomes
- See also, e.g., work by Andersson
  (resolved IMF, overshooting stars)

At higher resolution

- Other parameters may take over

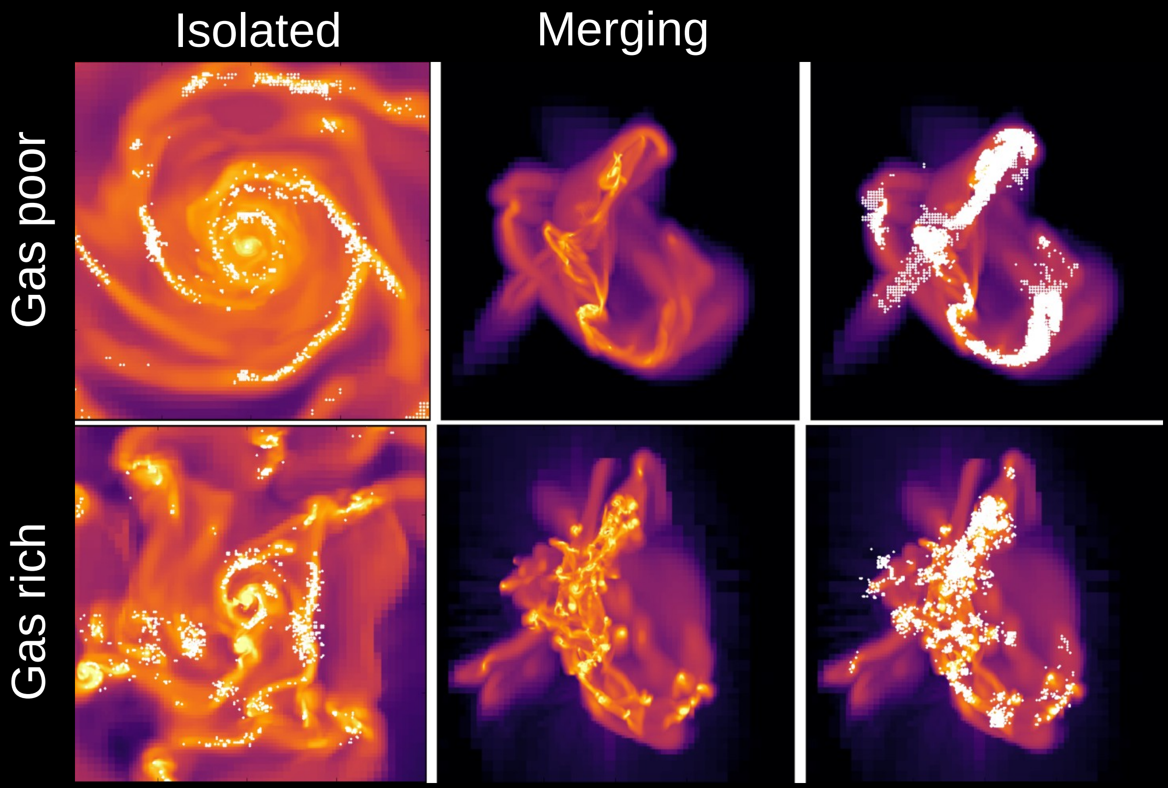
Need to consider

- Sub-grid recipes = e.g., cooling, SF, feedback
  - Are all cross-dependent

  If 1 is changed → may need to adapt all others

- This is also true for resolution

Example 7 : Merging galaxies



Fensch et al. 2016

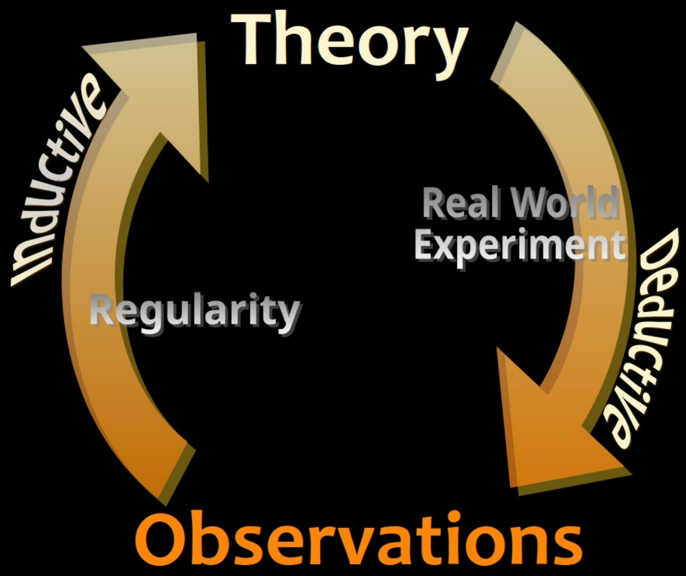
At low resolution
→ central starbursts

Weak SF enhancement
- In gas-rich mergers
  - Tidal compression
  - Pre-processing
    → Saturation

How to adapt ?
- Sub-grid recipes
  - cooling, SF, feedback

Of the Importance
of
Predictions

Predictions ?



Way to refute a theory

→ requires predictive power

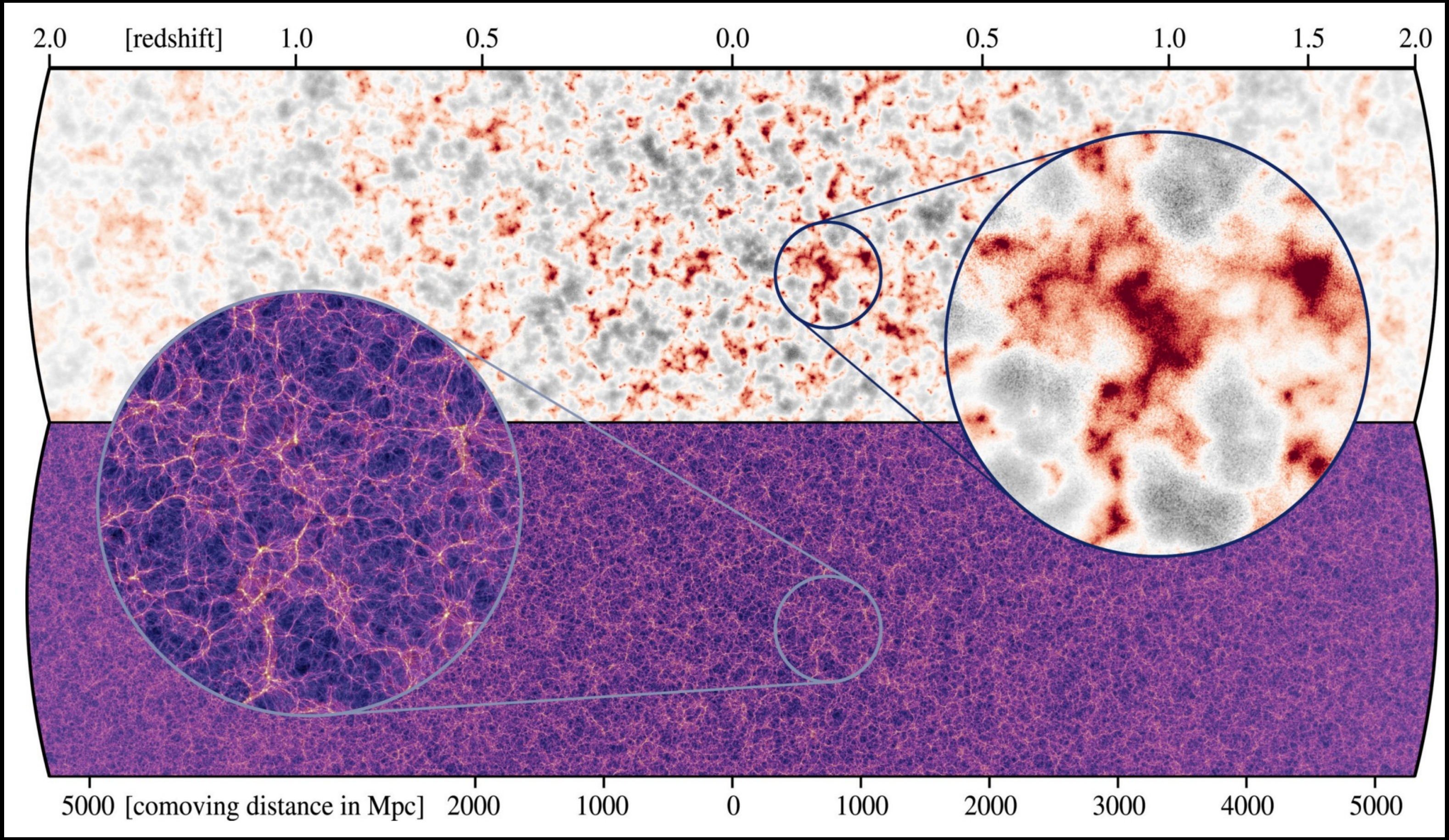
Particularly important

as a simulation : not a theory

→ difficulty of reproducing results

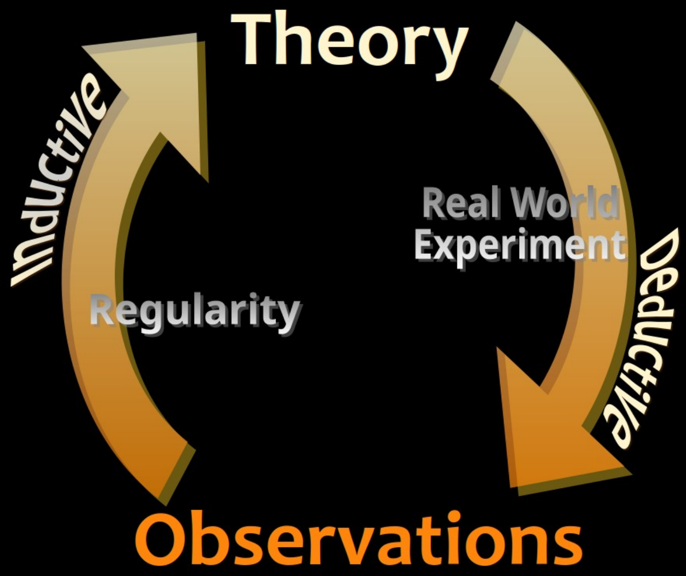
Predictions ? Hernández-Aguayo et al. 2024

Massive neutrinos versus Dark matter



Of the Importance
of
A “science question”

What is the science question ?



Simulations can serve as a seed
→ for further research

But

Without a science question
→ so what?

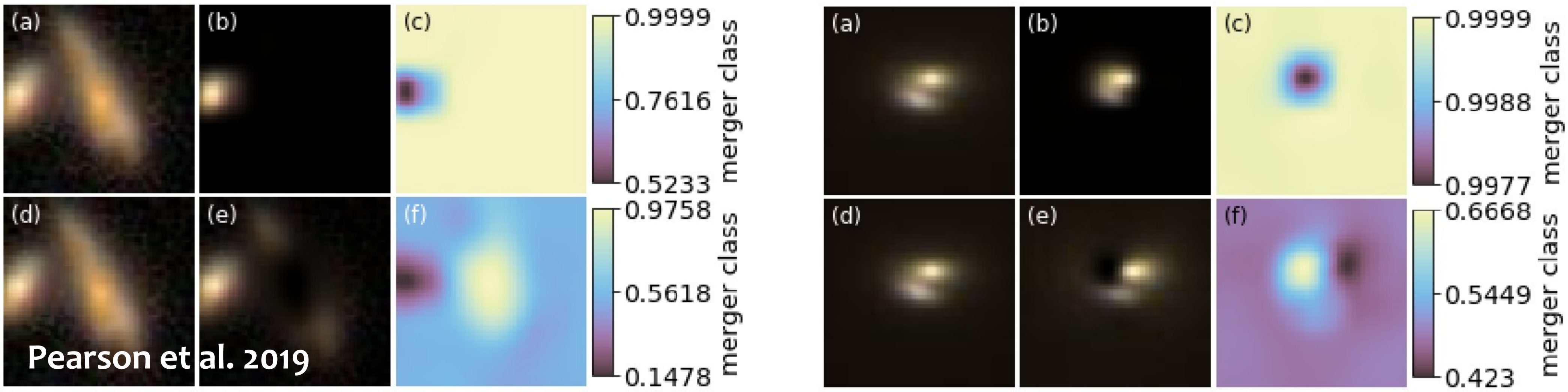
One last piece of Warning

One last piece of warning [© Jeremy Fensch]

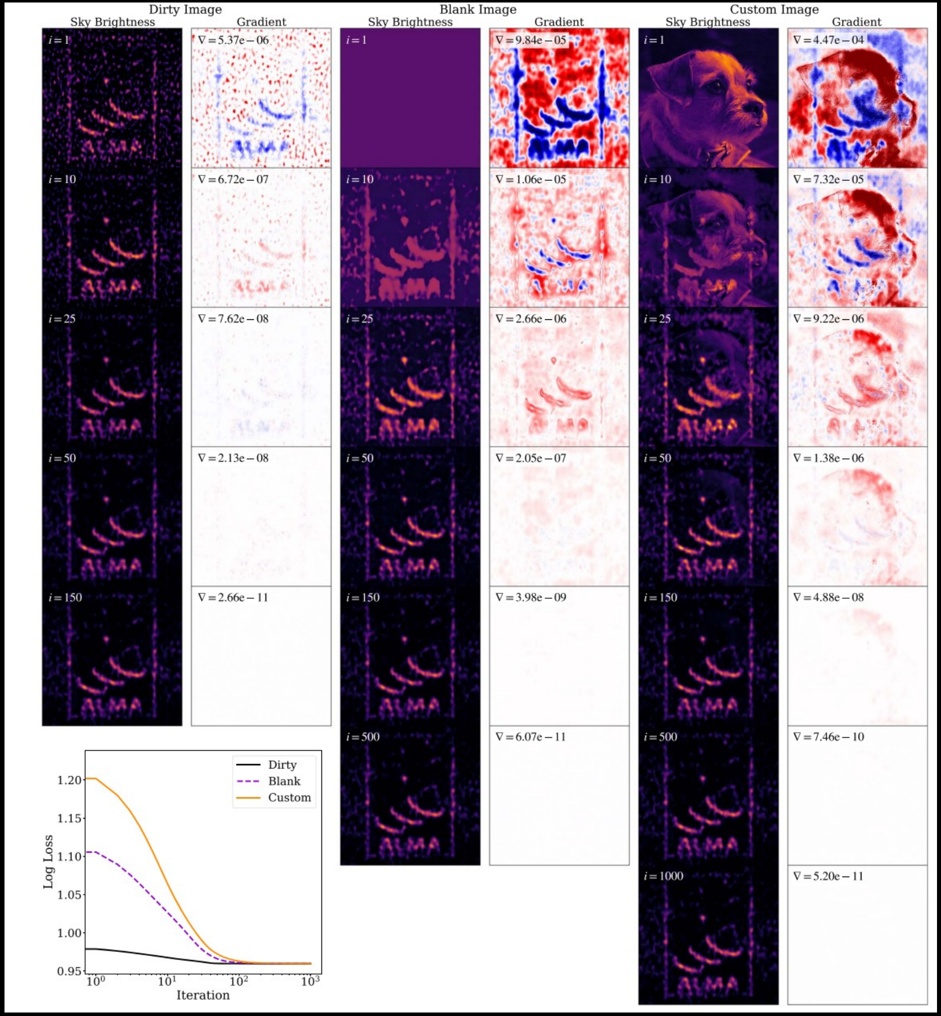
Identifying galaxy mergers in observations and simulations with deep learning – Pearson et al. 2019

Conclusions. The networks trained and tested with the same data perform the best, with observations performing better than simulations, a result of the observational sample being biased towards conspicuous mergers. Classifying SDSS observations with the simulation trained network has proven to work, providing tantalising prospects for using simulation trained networks for galaxy identification in large surveys.

How do we quantify biases?



One last piece of warning [© Jeremy Fensch]



Zawadski et al. 2023

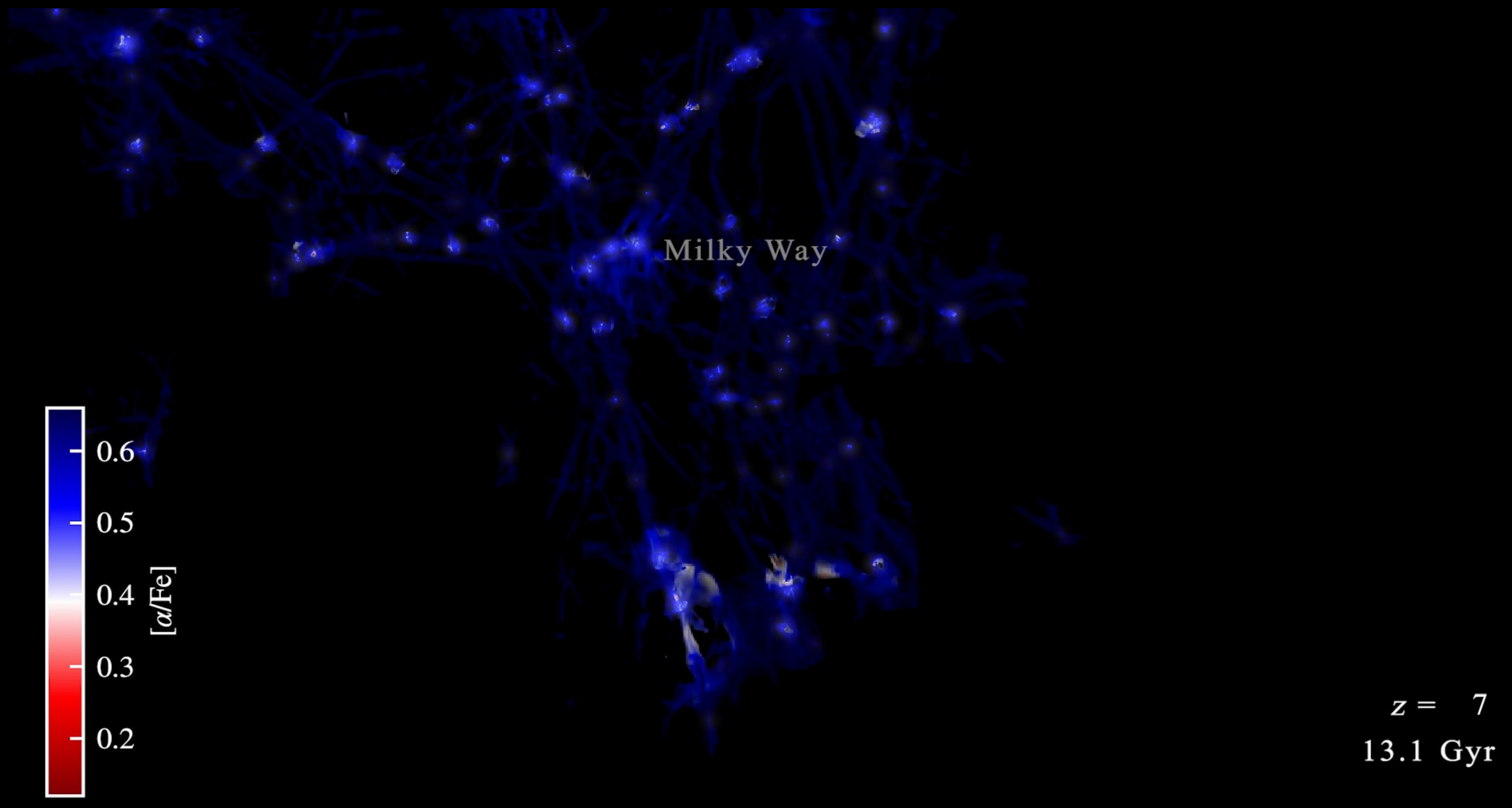
How can we quantify our biases ?

- Regularisation using descent optimisation
- Used on 3 input images
  - ▹ ALMA logo
  - ▹ Blank
  - ▹ Dog

How do we make sure

- We do not erase unknown signal?

Wrapping up



Agertz / Renaud / Segovia Otero [2021, 2022, 2025]

Take home messages

Simulations are ~ Observations of a (restricted) Model

- They are **not** theory per se
- Remember = theories can only be **refuted**

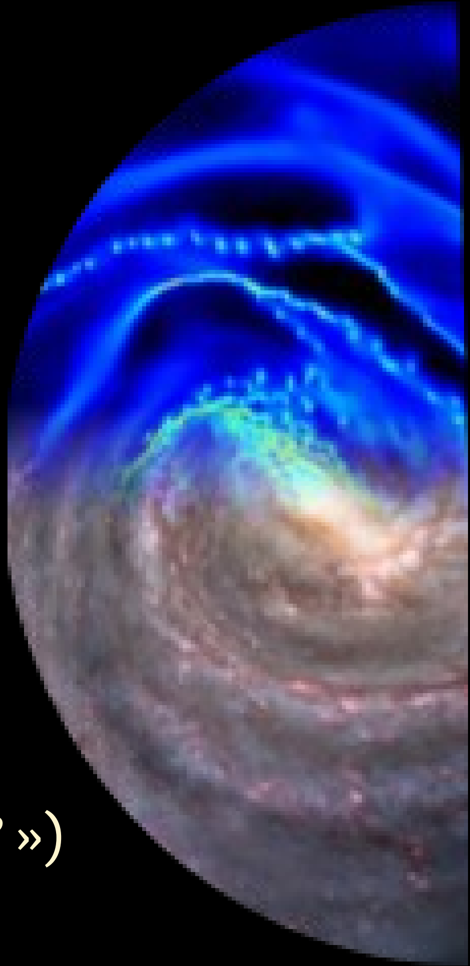
Many types of ISM-related simulations

- Scales, recipes (physics and setup), generic, tuned
  ⇒ each «should» require a science question
  OR at least a scientific context / motivation

Doing simulations is hard. Comparing them to Obs is tough

- Requires many more assumptions
- Requires we understand our observations too
- Let's not get fooled (remember the « bravo but ***so what ?*** »)

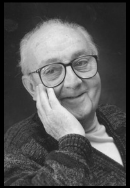
Of the importance of : coherence & predictions



What is your message ?

Take home messages

DavidMCEddy @ en.wikipedia



George Box

All Simulations are wrong

(Many are amazing)

but

Some are useful