Emulating Baryonic fields from cosmic to galaxy scales with Generative Deep Learning

Robert Feldmann, **Mauro Bernardini,** Joachim Stadel, Lucio Mayer, FIRE team

Institute for Computational Science
University of Zurich
robert.feldmann@uzh.ch

References:

Bernardini et al. 2022, arXiv:2110.11970

Feldmann et al. 2023, arXiv:2205.15325

Bernardini et al. in prep

Swiss SKA Days, September 8, 2023

FNSNF SWISS NATIONAL SCIENCE FOUNDATION

University of Zurich[UZH]

Motivation

- Need to understand both formation of (dark matter) structure and the biased tracers accessible to observations

- Most tracers, e.g., galaxies or 21 cm signals, are the product of complex astrophysical processes (“baryonic physics”)

Motivation

- Need to understand both formation of (dark matter) structure and the biased tracers accessible to observations

- Most tracers, e.g., galaxies or 21 cm signals, are the product of complex astrophysical processes (“baryonic physics”)

Marginalize over these (astrophysical) complications

e.g., Effective Field Theory (e.g., Senatore et al. 2015, Schmidt et al. 2019)

- Lose cosmological information from small scales (< few Mpc)
- Lose many of the astrophysical constraints

Motivation

- Need to understand both formation of (dark matter) structure and the biased tracers accessible to observations

- Most tracers, e.g., galaxies or 21 cm signals, are the product of complex astrophysical processes (“baryonic physics”)

Marginalize over these (astrophysical) complications

e.g., Effective Field Theory (e.g., Senatore et al. 2015, Schmidt et al. 2019)

- Lose cosmological information from small scales (< few Mpc)
- Lose many of the astrophysical constraints

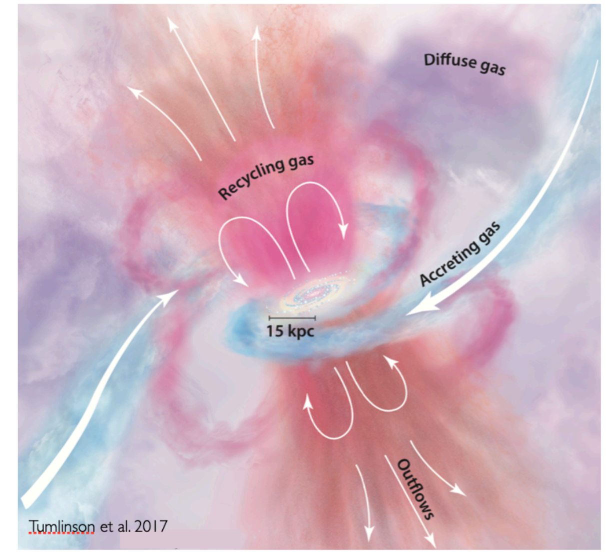
Try to model them

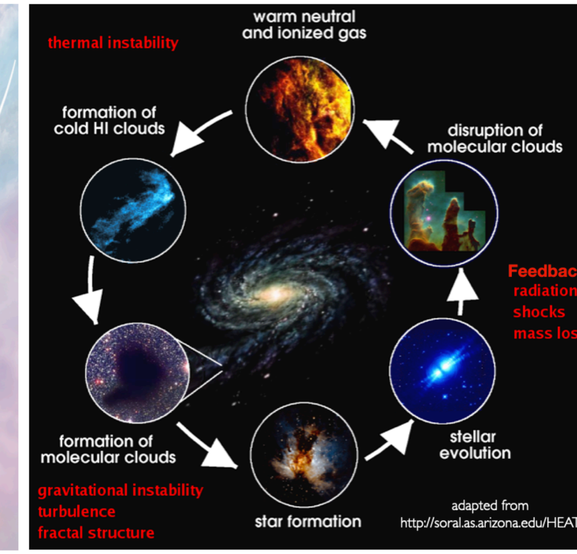
e.g., semi-analytical models, cosmological simulations

- Computationally expensive
- Some of the physics not well understood (e.g., AGN feedback)

Modeling baryons

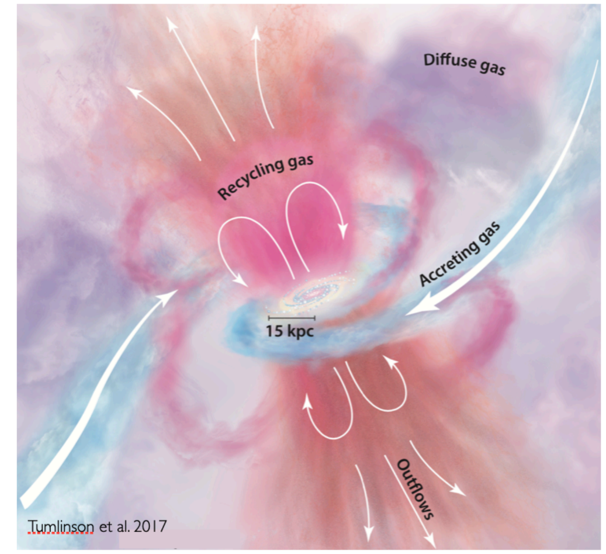
- Rich / complex physics:
  - Multitude of interacting physical processes
  - Wide range of scales
  - Coupling of scales: Stellar and AGN feedback

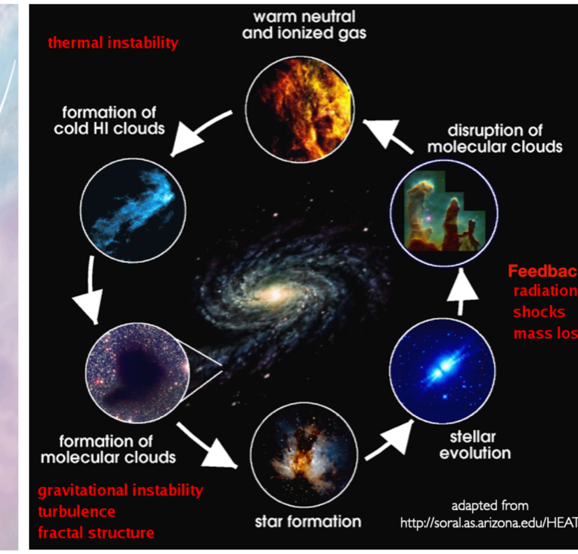




Modeling baryons

- Rich / complex physics:
  - Multitude of interacting physical processes
  - Wide range of scales
  - Coupling of scales: Stellar and AGN feedback



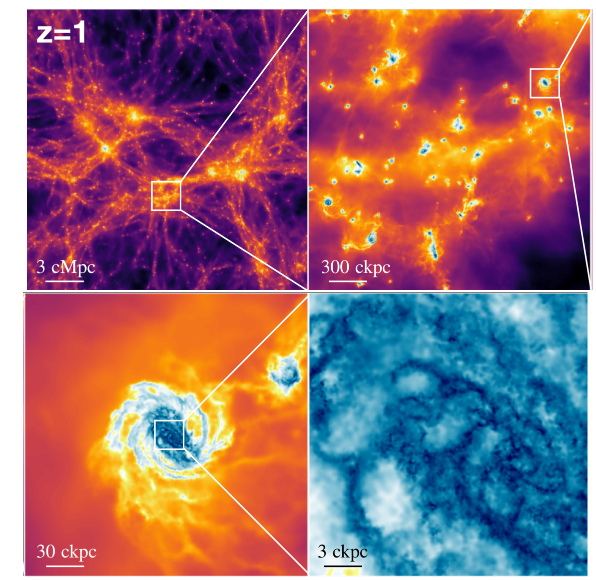


- Hydrodynamical cosmological simulations

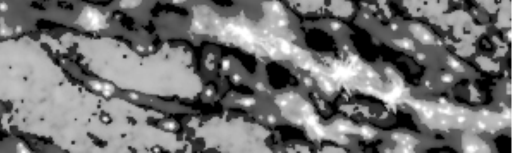
  + predict baryons on all scales and across cosmic history

  + test physical models (e.g., UV background evolution, feedback)

  - high computational cost

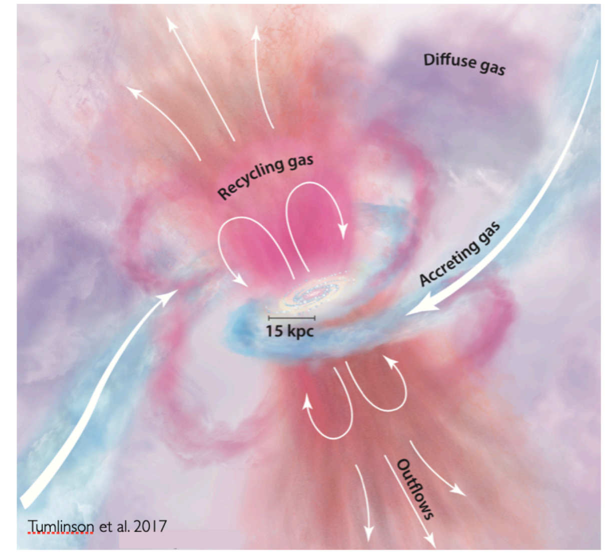


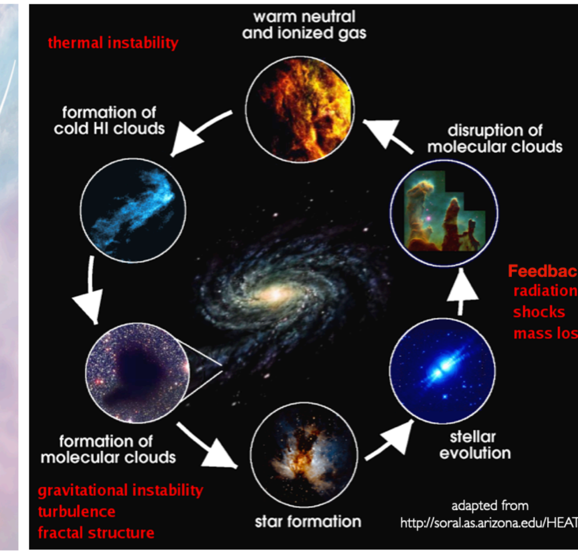
RF et al. 2023



Modeling baryons

- Rich / complex physics:
  - Multitude of interacting physical processes
  - Wide range of scales
  - Coupling of scales: Stellar and AGN feedback



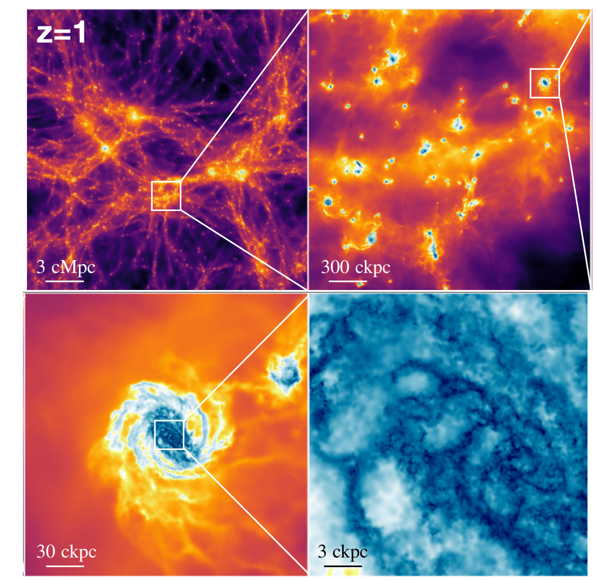


- Hydrodynamical cosmological simulations

  \+ predict baryons on all scales and across cosmic history

  \+ test physical models (e.g., UV background evolution, feedback)

  - high computational cost

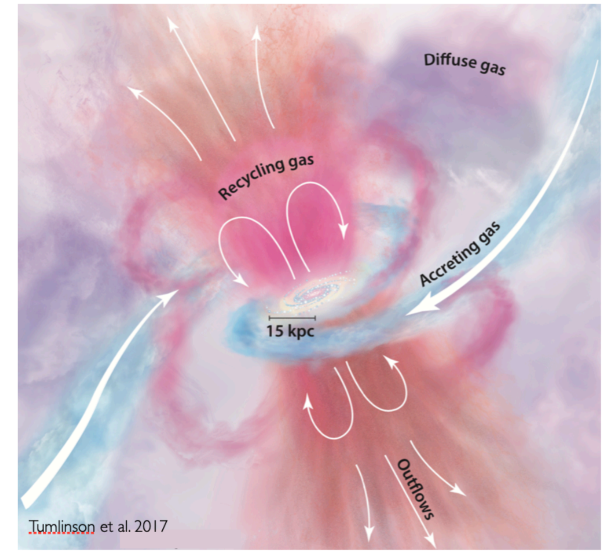


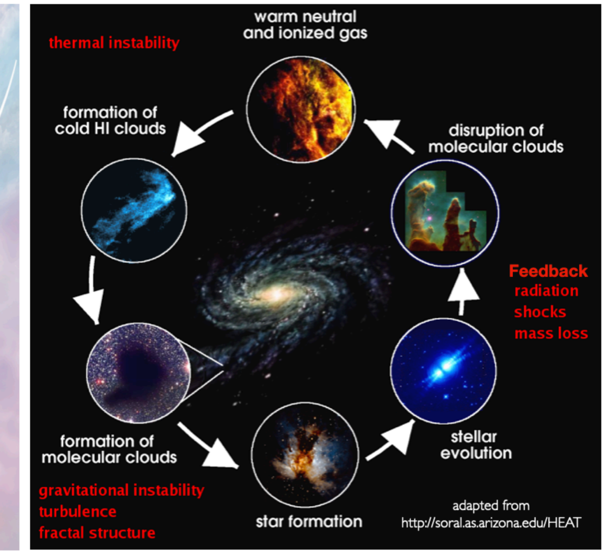
RF et al. 2023

<u>Idea: Encapsulate simulation results by training deep neural network</u>

Modeling baryons

- Rich / complex physics:
  - Multitude of interacting physical processes
  - Wide range of scales
  - Coupling of scales: Stellar and AGN feedback



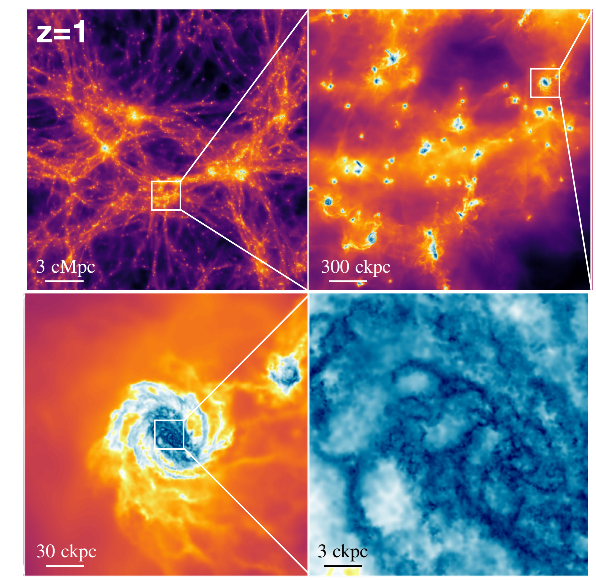


- Hydrodynamical cosmological simulations

  \+ predict baryons on all scales and across cosmic history

  \+ test physical models (e.g., UV background evolution, feedback)

  - high computational cost



RF et al. 2023

<u>Idea: Encapsulate simulation results by training deep neural network</u>

- Once trained, prediction is computationally *inexpensive* => Emulator

  \+ Generation of mock catalogs, light cones, covariance matrix calculation

  \+ Interpolate to new parameter values, redshifts etc. => Likelihood-free inference

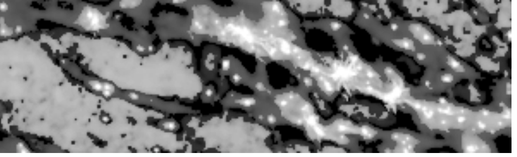
  \+ Increase resolution ("super-resolution")

  \+ Upscale to much larger volumes while preserving high resolution

EMBER - **EM**ulating **B**aryonic En**R**ichment Network

- Deep learning framework to sample “realistic” baryon maps for given DM maps

$$p(\Sigma_{\rm gas} | \Sigma_{\rm DM})$$

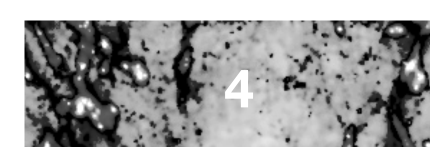


EMBER - **EM**ulating **B**aryonic En**R**ichment Network

- Deep learning framework to sample “realistic” baryon maps for given DM maps

$$p(\Sigma_{\rm gas} \,|\, \Sigma_{\rm DM})$$

- Fiducial architecture based on a Deep Convolutional Neural Network + conditional Generative Adversarial Network (GAN)



EMBER - **EM**ulating **B**aryonic En**R**ichment Network

- Deep learning framework to sample "realistic" baryon maps for given DM maps

$$p(\Sigma_{\rm gas} \,|\, \Sigma_{\rm DM})$$

- Fiducial architecture based on a Deep Convolutional Neural Network + conditional Generative Adversarial Network (GAN)

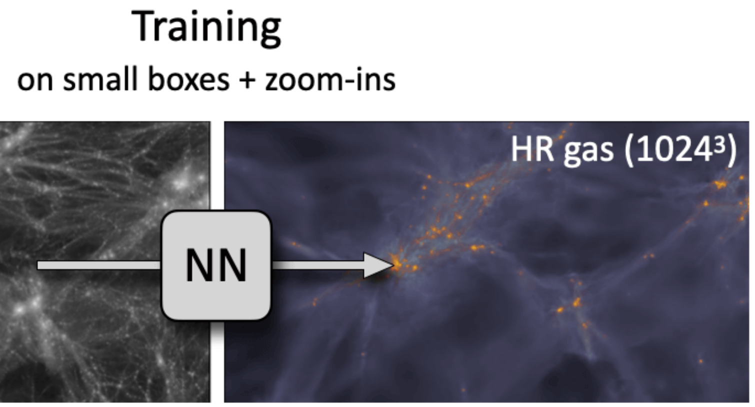
- Trained on cosmological hydrodynamical simulations

EMBER - **EM**ulating **B**aryonic En**R**ichment Network

- Deep learning framework to sample “realistic” baryon maps for given DM maps

$$p(\Sigma_{\rm gas} \,|\, \Sigma_{\rm DM})$$

- Fiducial architecture based on a Deep Convolutional Neural Network + conditional Generative Adversarial Network (GAN)
- Trained on cosmological hydrodynamical simulations



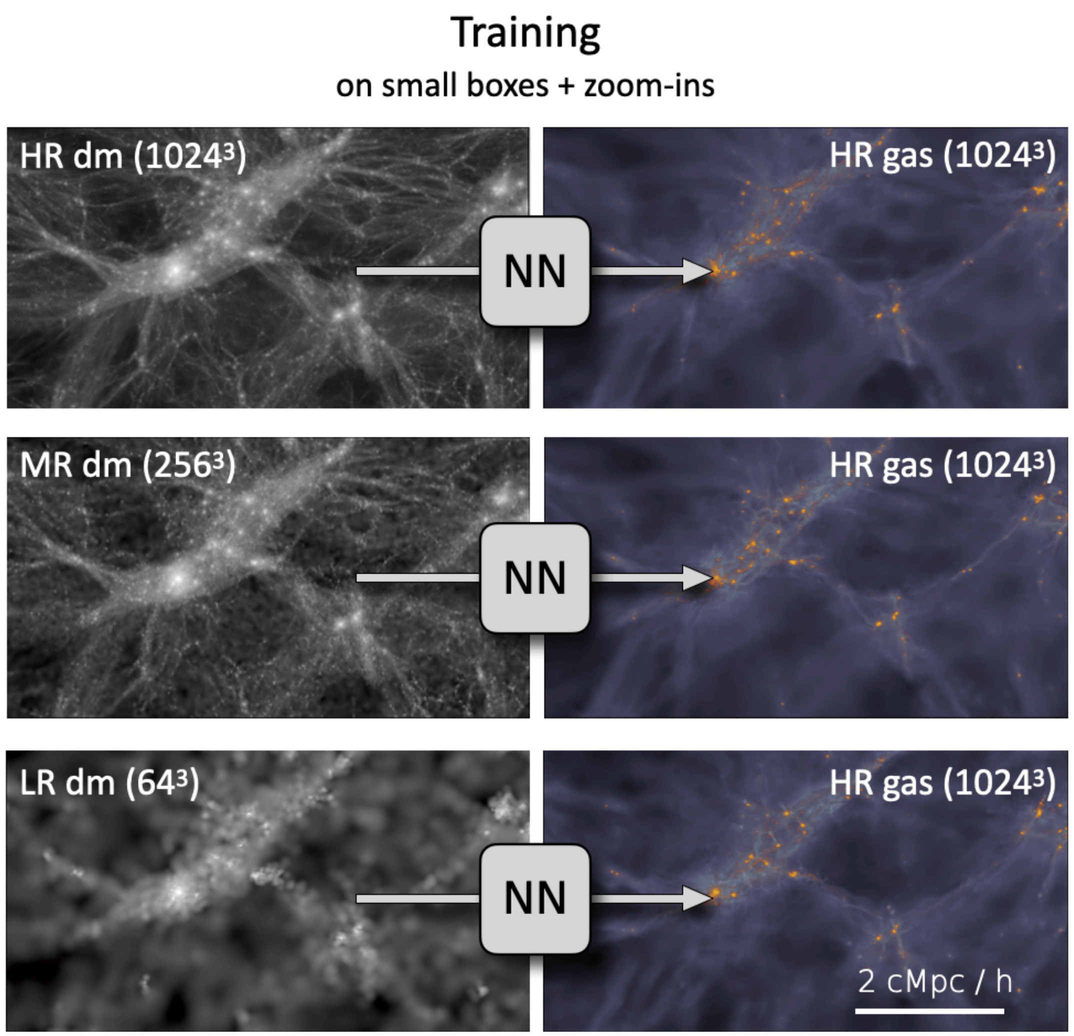
EMBER - **EM**ulating **B**aryonic En**R**ichment Network

- Deep learning framework to sample "realistic" baryon maps for given DM maps

$$p(\Sigma_{\text{gas}} \,|\, \Sigma_{\text{DM}})$$

- Fiducial architecture based on a Deep Convolutional Neural Network + conditional Generative Adversarial Network (GAN)

- Trained on cosmological hydrodynamical simulations

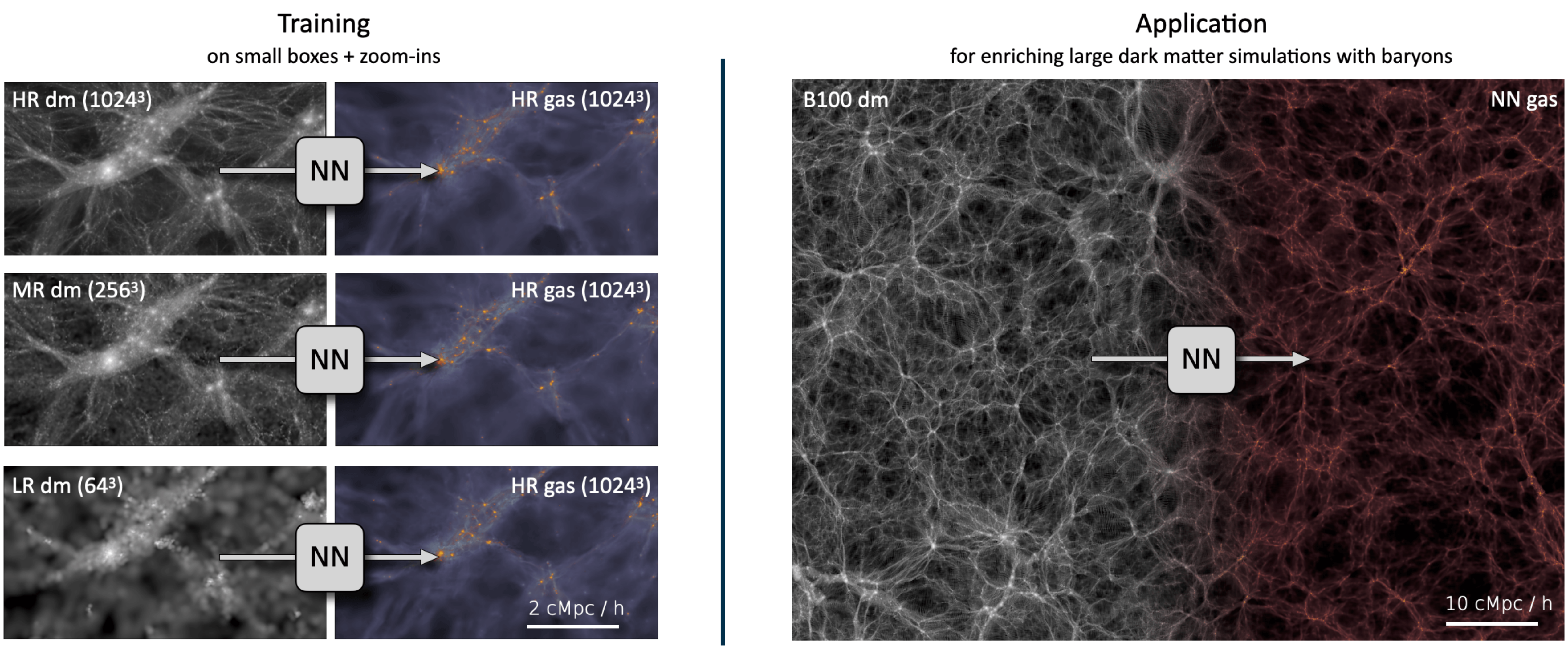


EMBER - **EM**ulating **B**aryonic En**R**ichment Network

- Deep learning framework to sample “realistic” baryon maps for given DM maps

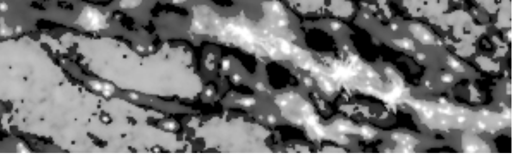
$$p(\Sigma_{\rm gas} \,|\, \Sigma_{\rm DM})$$

- Fiducial architecture based on a Deep Convolutional Neural Network + conditional Generative Adversarial Network (GAN)
- Trained on cosmological hydrodynamical simulations



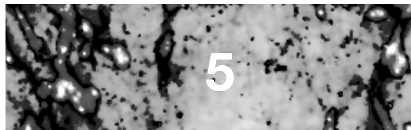
Bernardini, RF, et al. 2022, arXiv:2110.11970

title slide shows part of the B100 map



Emulating observational data on the field level

Input	Output	Posterior	Note	Reference	
Noise	HI field	Yes	3-d, Δx~35 kpc/h	Zamudio-Fernandez et al. 2019	
DM map, Noise	electron pressure map	Yes	2d, Δx~200 kpc/h, tSZ	Tröster et al. 2019	
DM field	elec. pressure, density, momentum fields	No	3d, Δx~100 kpc, tSZ, kSZ	Thiele et al. 2020	
DM field	HI field	No	3d, Δx~140 kpc/h	Wadekar et al. 2021	
DM field	gas density, temp., velocity field	No	3d, Δx~20 kpc/h	Harrington et al. 2021	
DM field, Noise	gas density, temp., velocity field	Yes	3d, Δx~40 kpc/h	Horowitz et al. 2021	
DM maps	HI maps	Yes	2d, Δx~400 kpc/h	Hassan et al. 2022	
DM maps	HI maps, gas maps	Yes	2d, Δx~3.6 kpc/h	Bernardini et al. 2022	**EMBER**
DM maps, l.o.s. velocities	gas maps, temp., velocity, HI maps	Yes	2d, variable resolution	Bernardini et al. in prep	**EMBER-2**



Simulation data

- Simulations run with GIZMO, FIRE-2 physics

FIRE
Feedback In Realistic Environments

fire.northwestern.edu

Simulation data

FIRE
Feedback In Realistic Environments
fire.northwestern.edu

- Simulations run with GIZMO, FIRE-2 physics
- 2 kinds of simulations:

Cosmological volume

Cosmological zoom-in

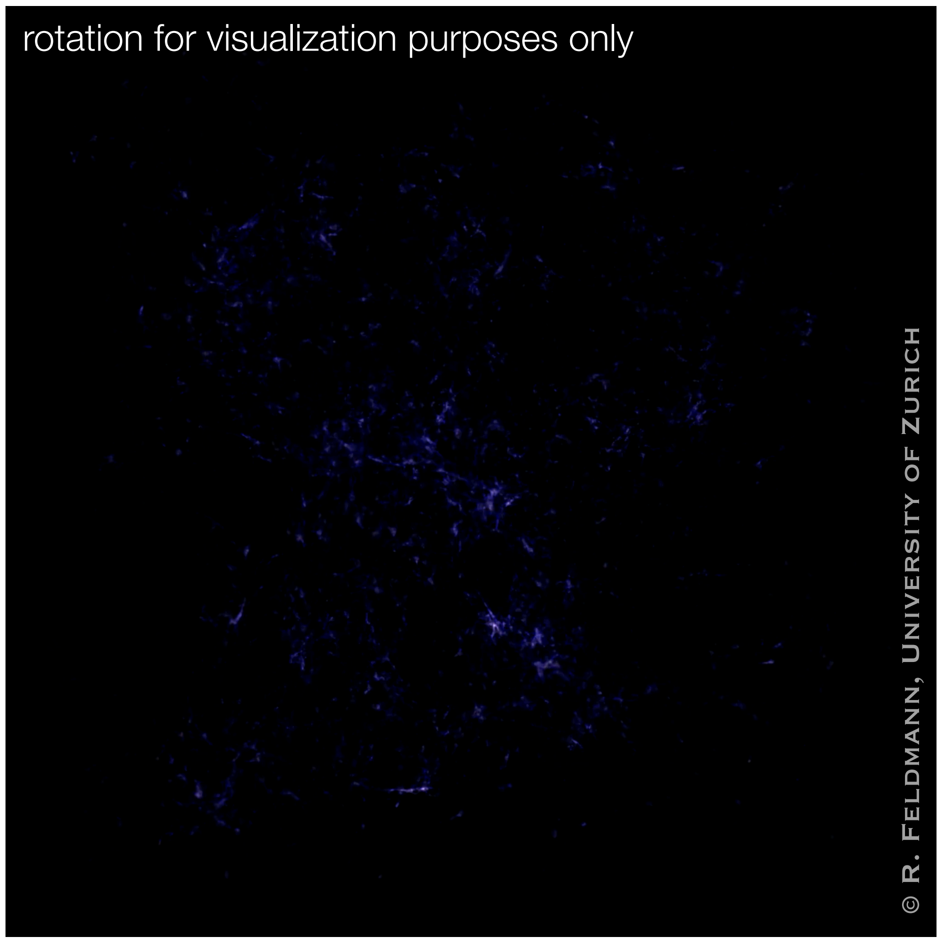
Simulation data

- Simulations run with GIZMO, FIRE-2 physics
- 2 kinds of simulations:

Feedback In Realistic Environments

fire.northwestern.edu

Cosmological volume



FIREbox (RF et al. 2023 arXiv:2205.15325)

- L = 22 cMpc
- 2 billion (gas+DM) particles
- $m_b \sim 6 \times 10^4\, M_\odot$, $\Delta x \sim$ tens of pc

Cosmological zoom-in

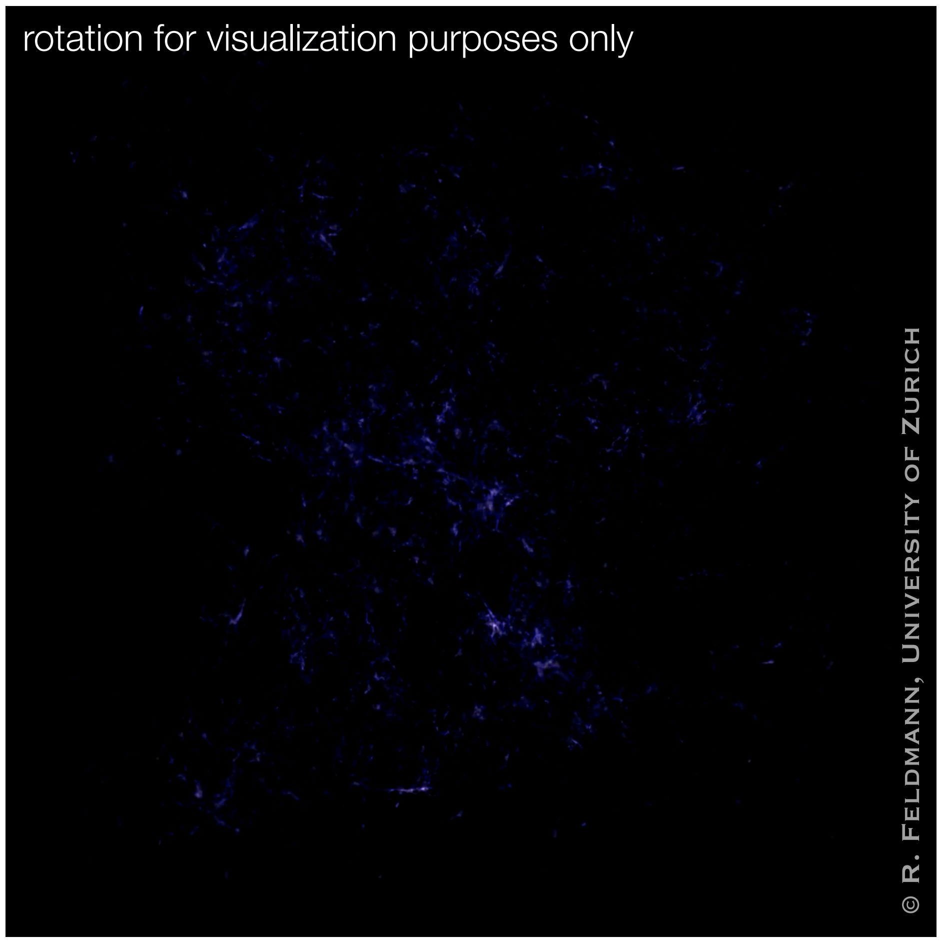
Simulation data

Feedback In Realistic Environments

fire.northwestern.edu

- Simulations run with GIZMO, FIRE-2 physics
- 2 kinds of simulations:

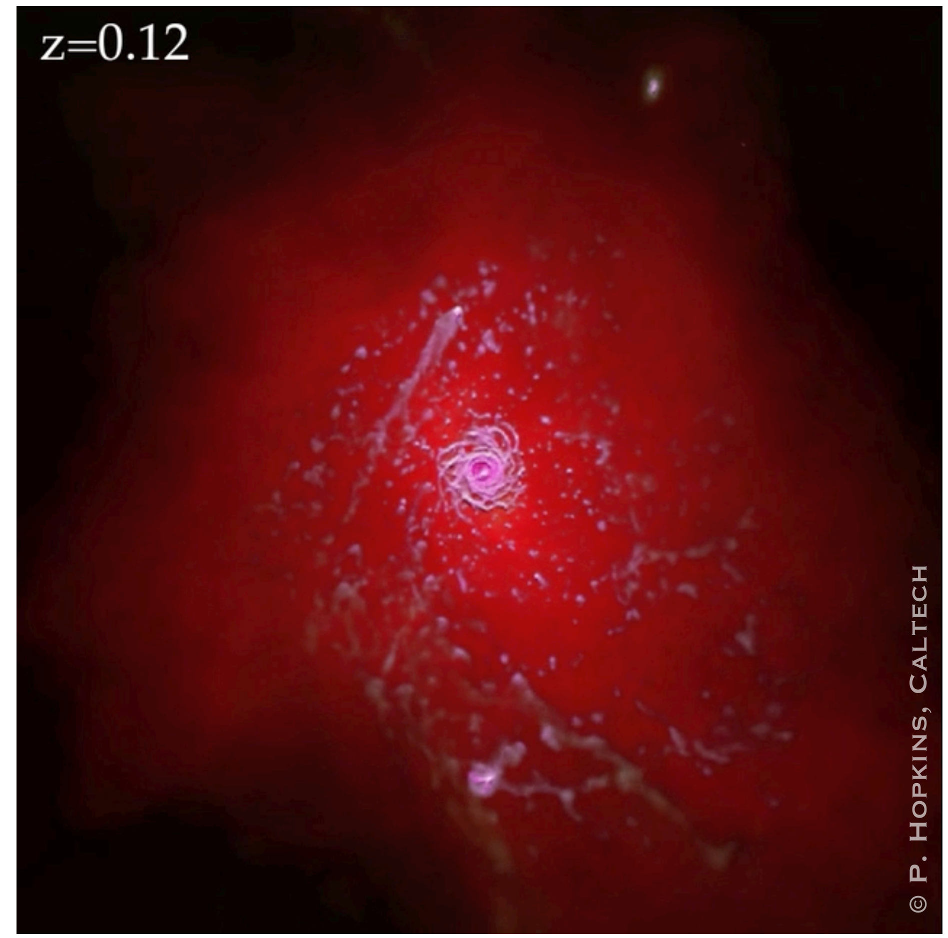
Cosmological volume



FIREbox (RF et al. 2023 arXiv:2205.15325)

- L = 22 cMpc
- 2 billion (gas+DM) particles
- $m_b \sim 6 \times 10^4\, M_\odot$, $\Delta x \sim$ tens of pc

Cosmological zoom-in

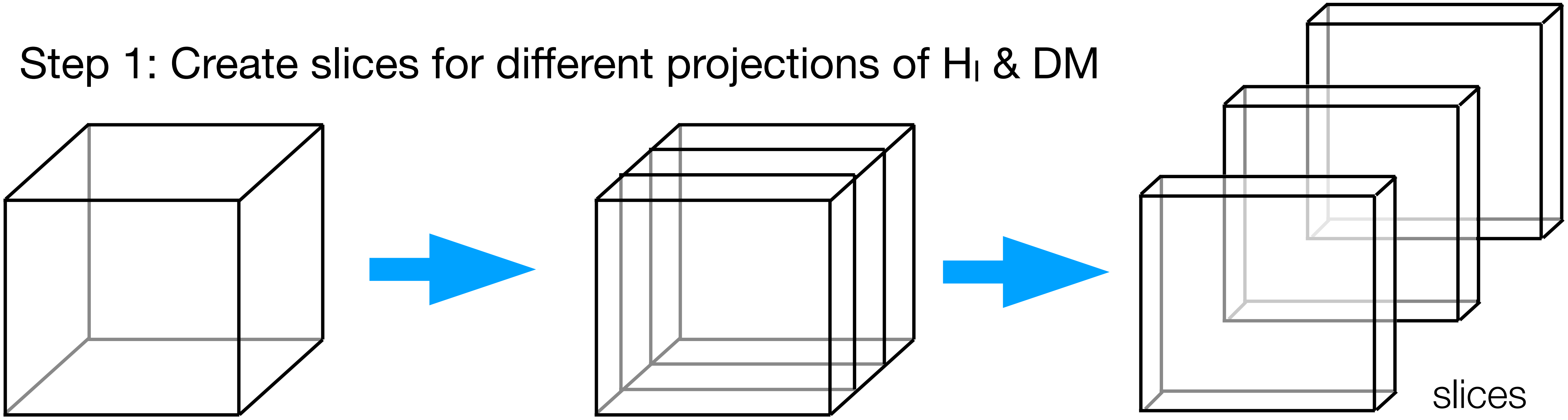


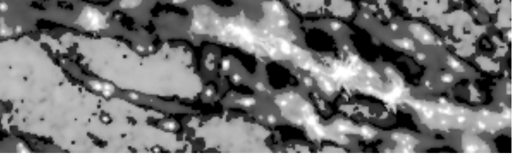
MassiveFIRE (RF+2016, Angles-Alcazar+2017)

- ~few cMpc region
- 4 massive halos
- $m_b \sim 3 \times 10^4\, M_\odot$, $\Delta x \sim$ tens of pc

From FIRE to EMBER

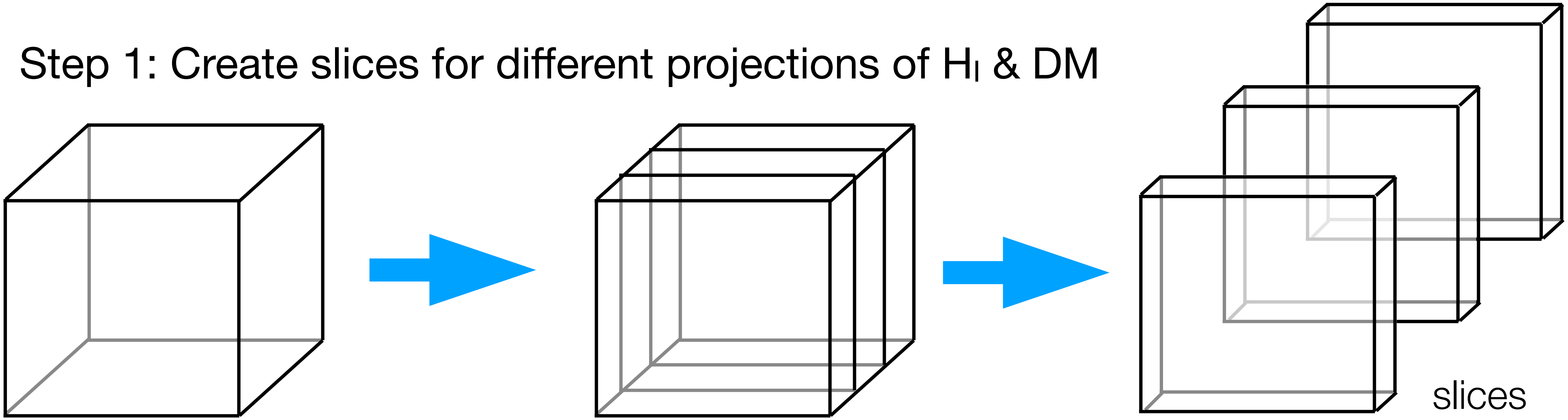
Step 1: Create slices for different projections of H_I & DM



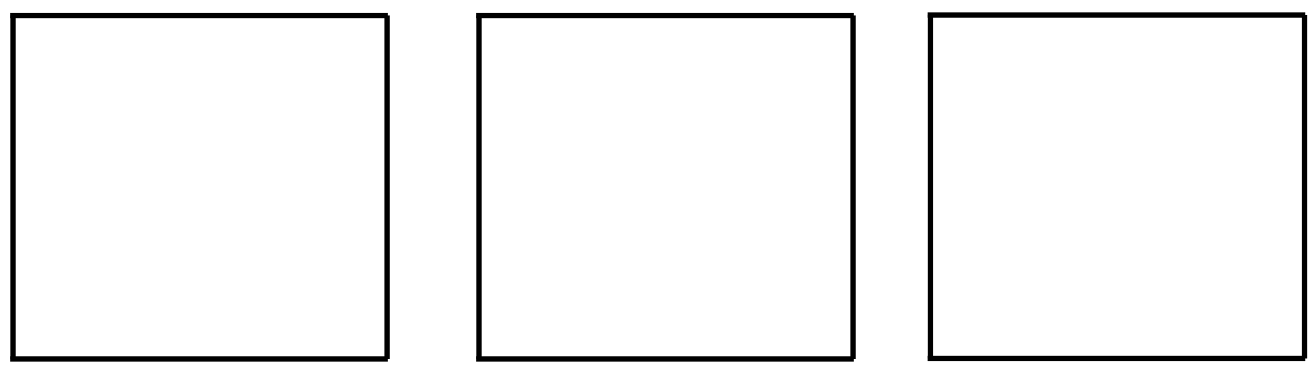


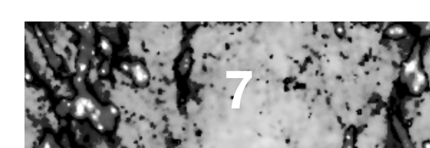
From FIRE to EMBER

Step 1: Create slices for different projections of H_I & DM



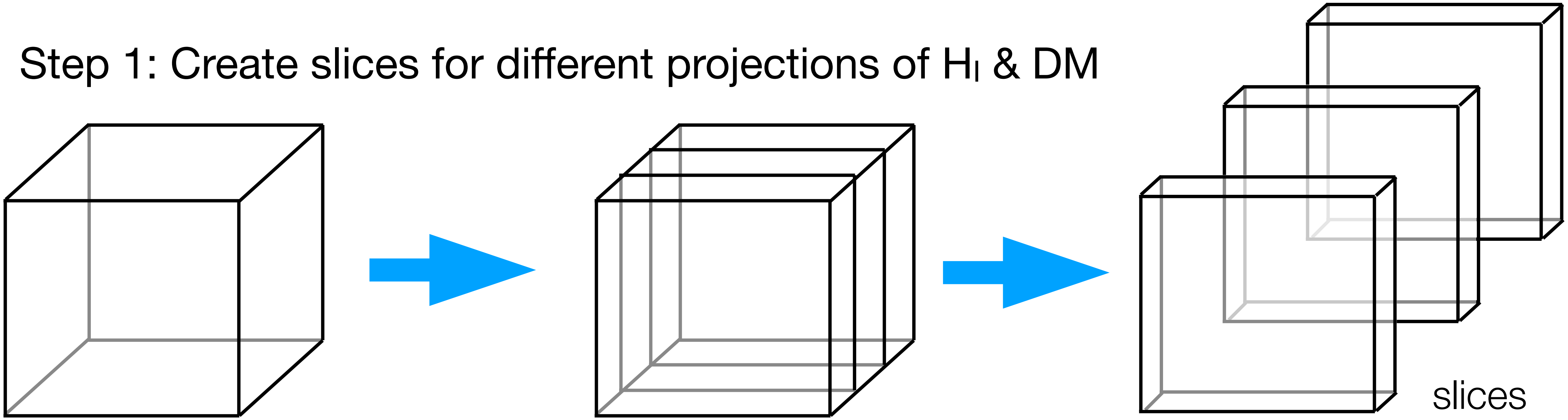
Step 2: Project H_I and DM density fields



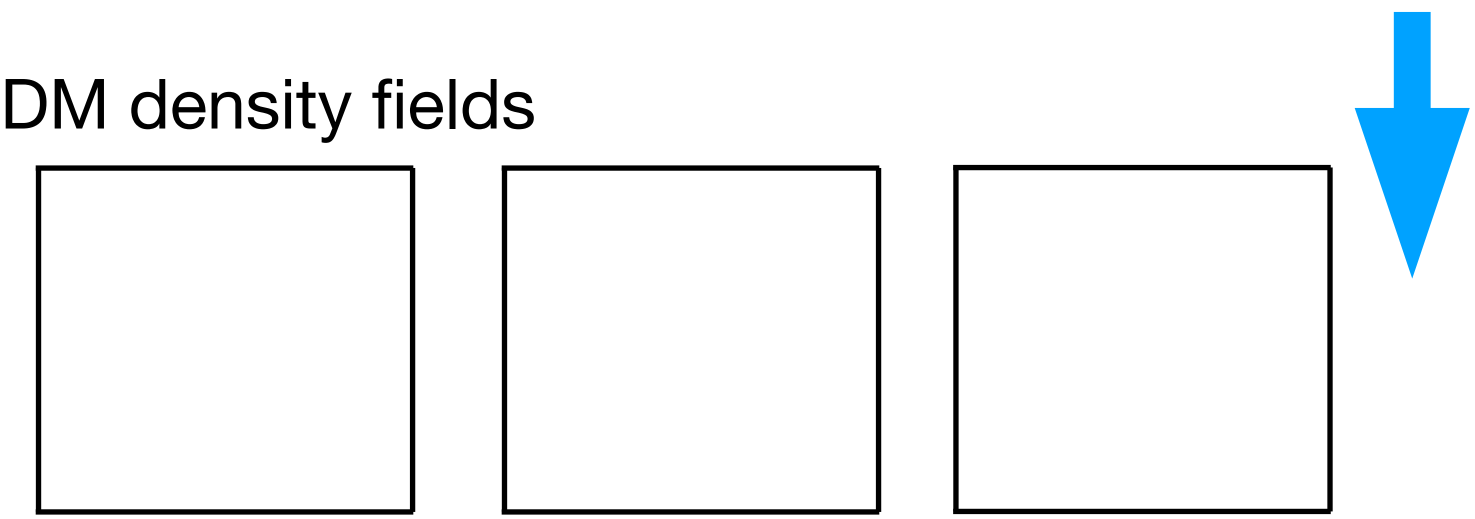


From FIRE to EMBER

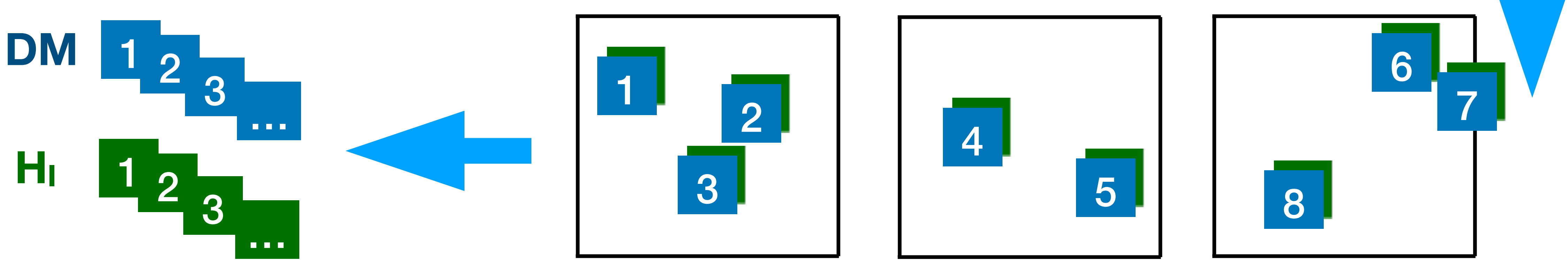
Step 1: Create slices for different projections of H_I & DM



Step 2: Project H_I and DM density fields



Step 3: Extract 512x512 pixel tiles randomly



Basic conditional GAN

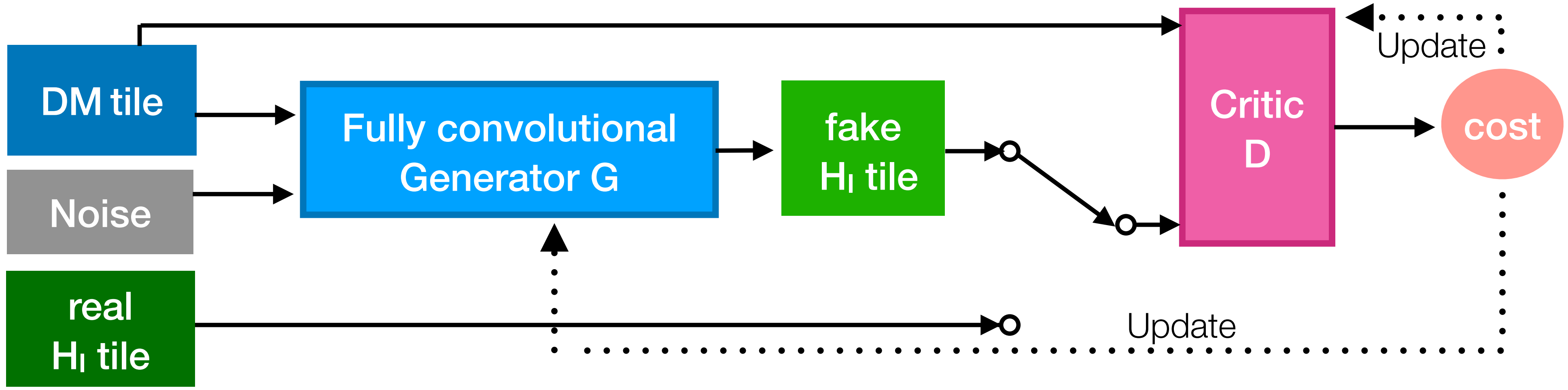
EMBER is a conditional GAN

- Generator G and Discriminator D networks compete in an adversarial game
- G tries to deceive D ⇔ D tries to spot images generated by G

Basic conditional GAN

EMBER is a conditional GAN

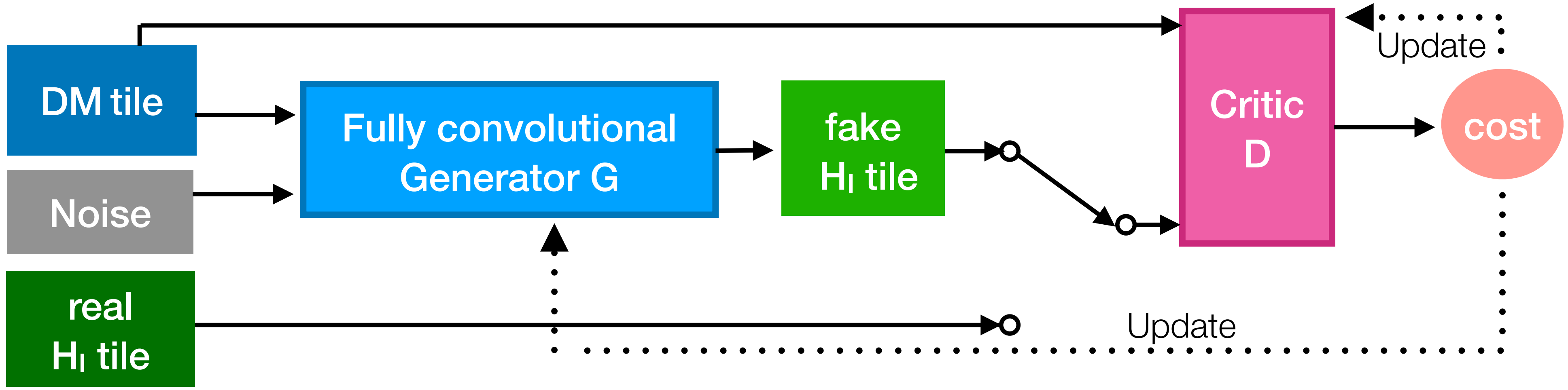
- Generator G and Discriminator D networks compete in an adversarial game
- G tries to deceive D ⇔ D tries to spot images generated by G



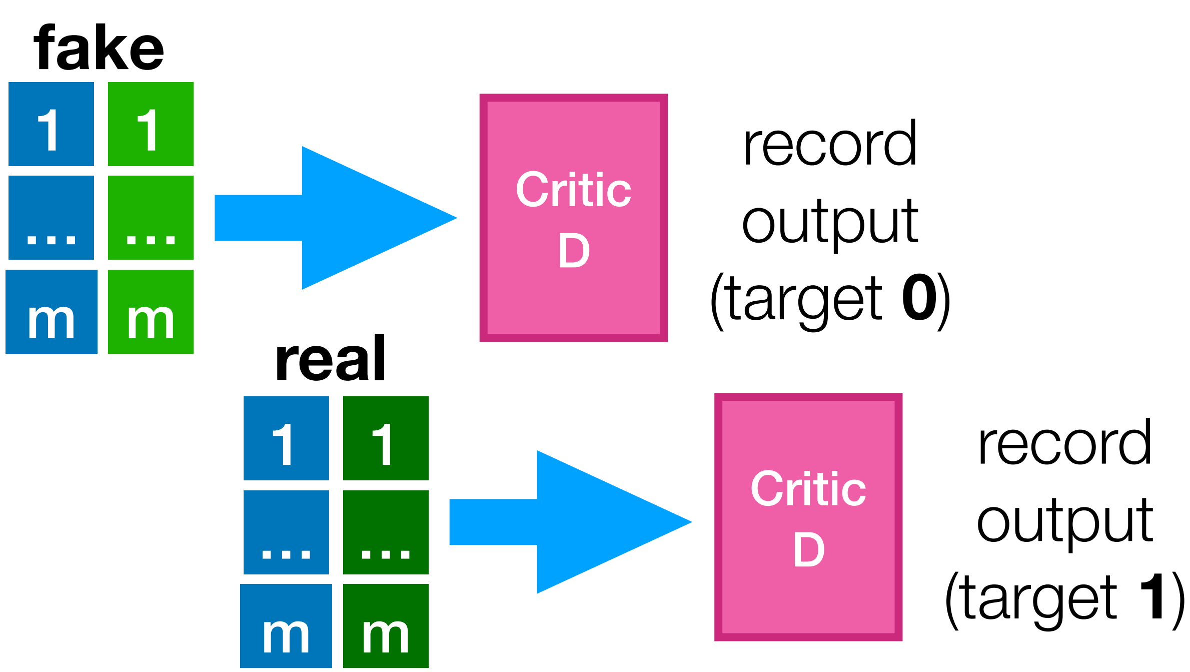
Basic conditional GAN

EMBER is a conditional GAN

- Generator G and Discriminator D networks compete in an adversarial game
- G tries to deceive D ⇔ D tries to spot images generated by G



Step 1: Update Discriminator / Critic D

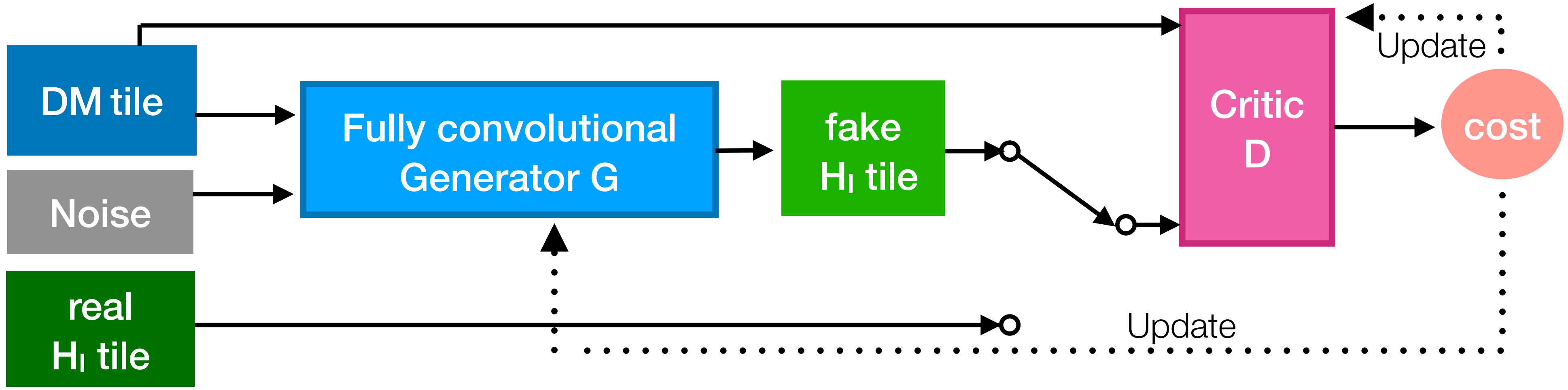


$$\mathrm{Cost}_{\mathrm{D}} = \mathbb{E}[\mathrm{D}(\mathrm{fake})] - \mathbb{E}[\mathrm{D}(\mathrm{real})]$$

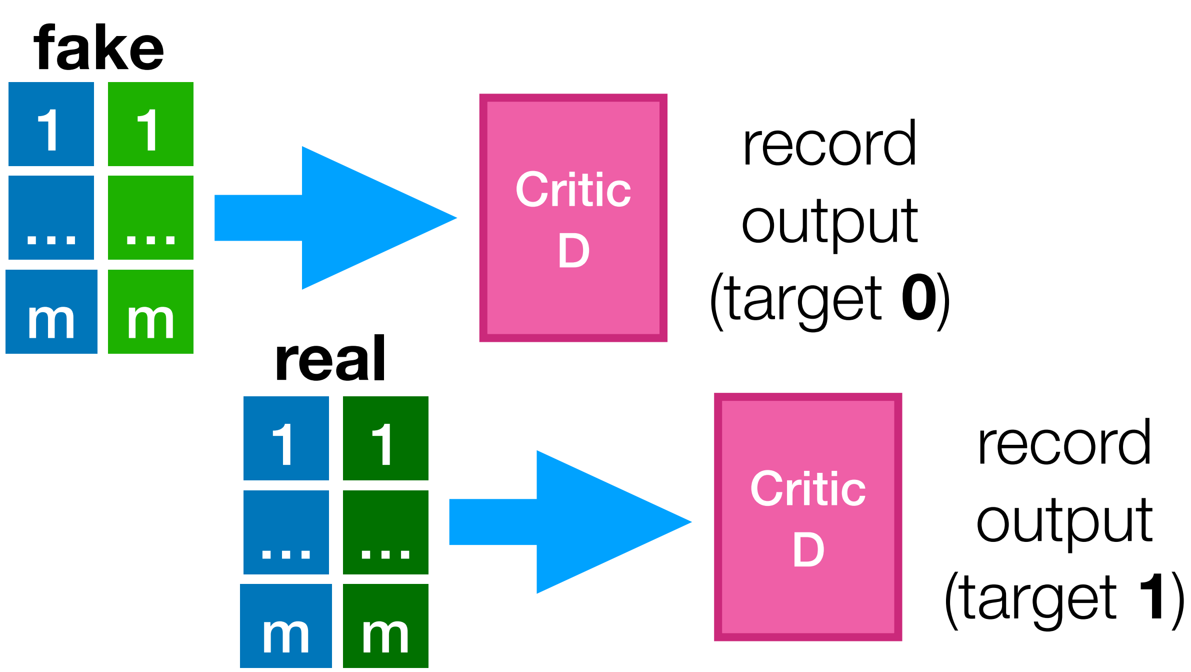
Basic conditional GAN

EMBER is a conditional GAN

- Generator G and Discriminator D networks compete in an adversarial game
- G tries to deceive D ⇔ D tries to spot images generated by G

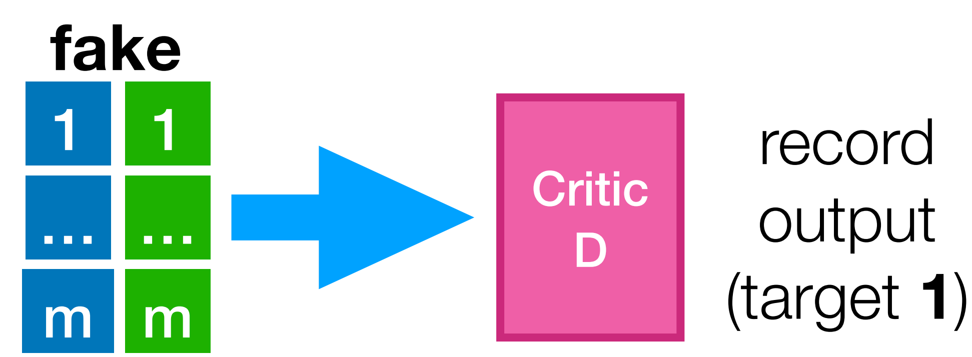


Step 1: Update Discriminator / Critic D

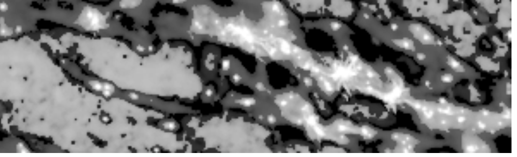


$$\mathrm{Cost}_\mathrm{D} = \mathbb{E}[\mathrm{D}(\mathrm{fake})] - \mathbb{E}[\mathrm{D}(\mathrm{real})]$$

Step 2: Update Generator G



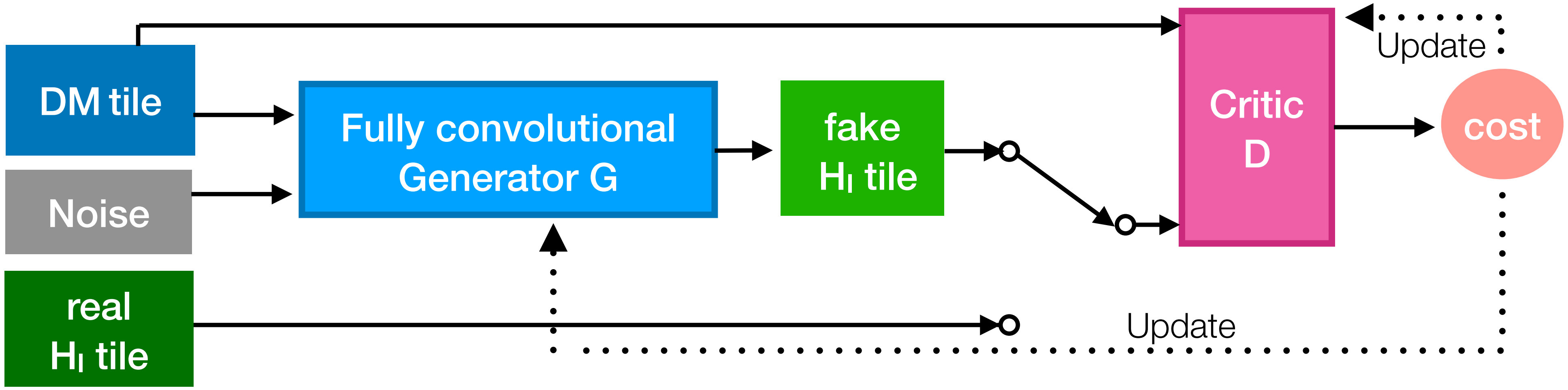
$$\mathrm{Cost}_\mathrm{G} = -\,\mathbb{E}[\mathrm{D}(\mathrm{fake})]$$



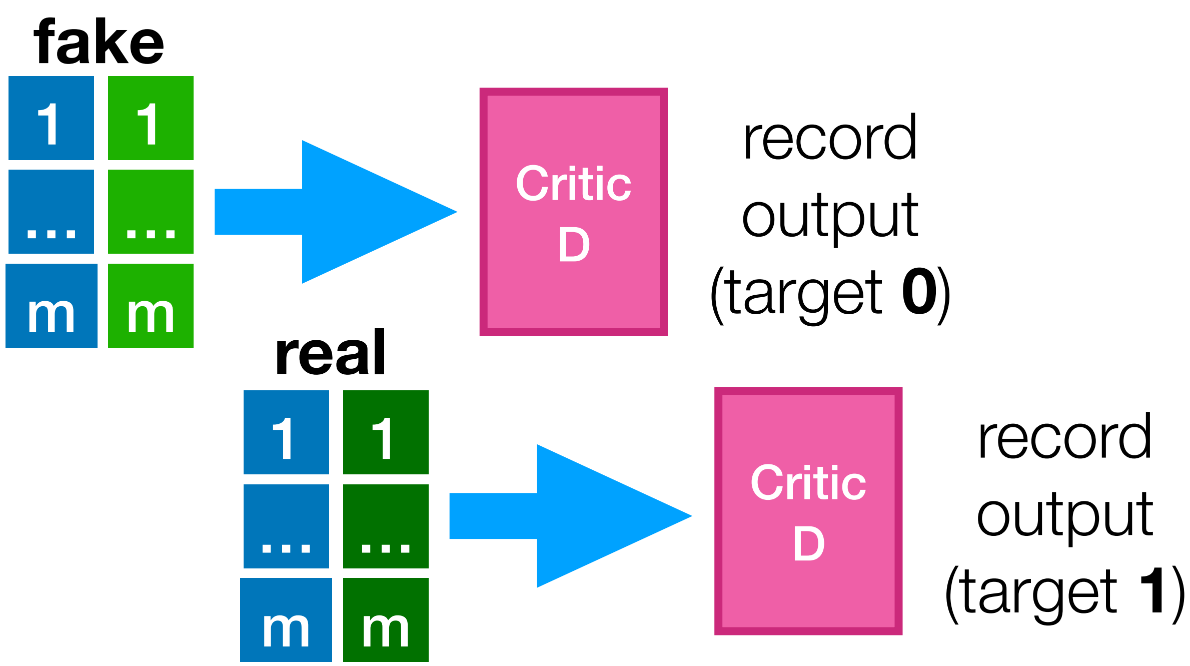
Basic conditional GAN

EMBER is a conditional GAN

- Generator G and Discriminator D networks compete in an adversarial game
- G tries to deceive D ⇔ D tries to spot images generated by G

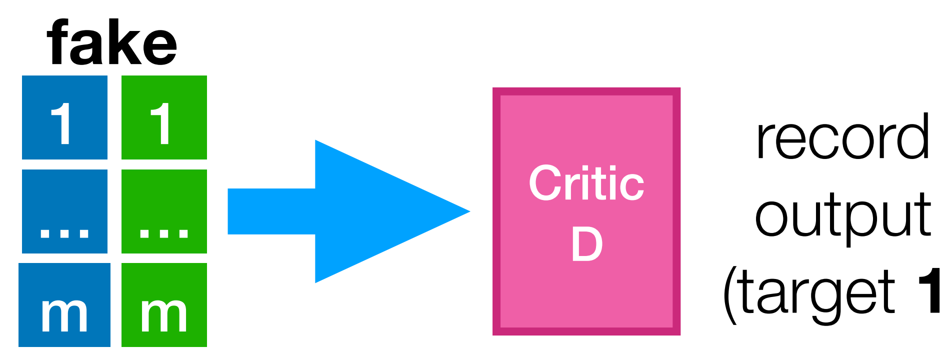


Step 1: Update Discriminator / Critic D



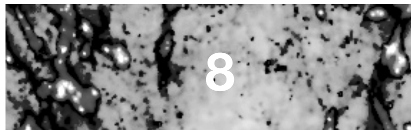
$$\mathrm{Cost}_{\mathrm{D}} = \mathbb{E}[\mathrm{D(fake)}] - \mathbb{E}[\mathrm{D(real)}]$$

Step 2: Update Generator G



$$\mathrm{Cost}_{\mathrm{G}} = -\,\mathbb{E}[\mathrm{D(fake)}]$$

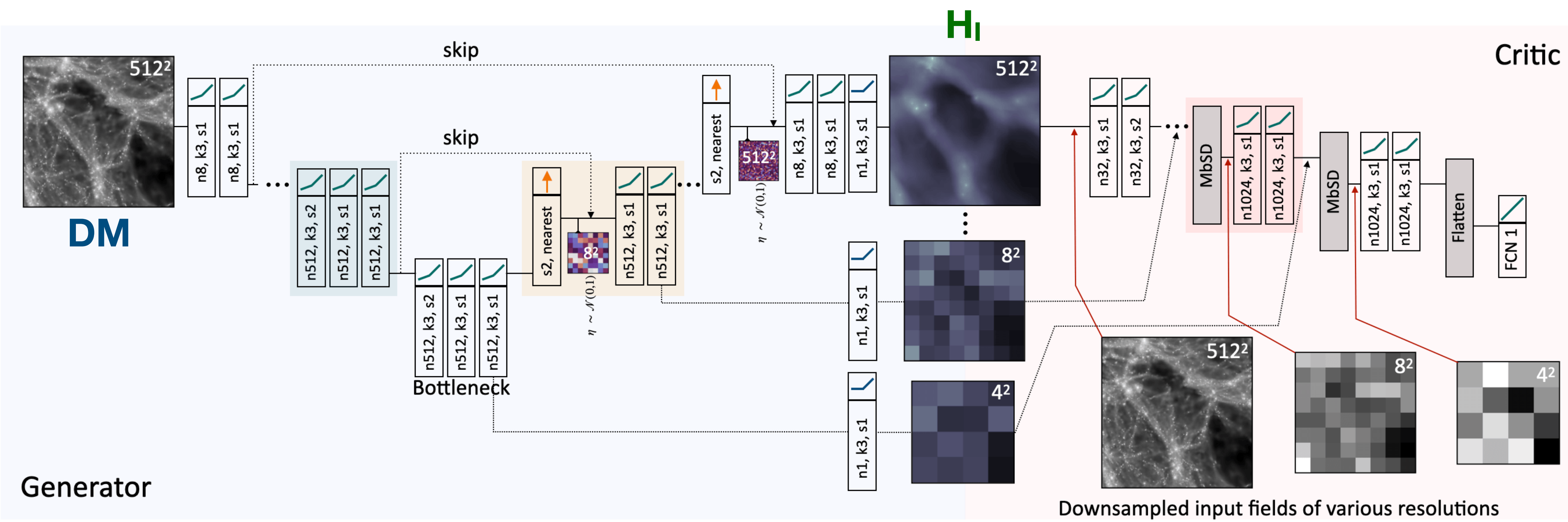
Step 3: Continue with step 1



A detailed view of the EMBER architecture

Fiducial architecture

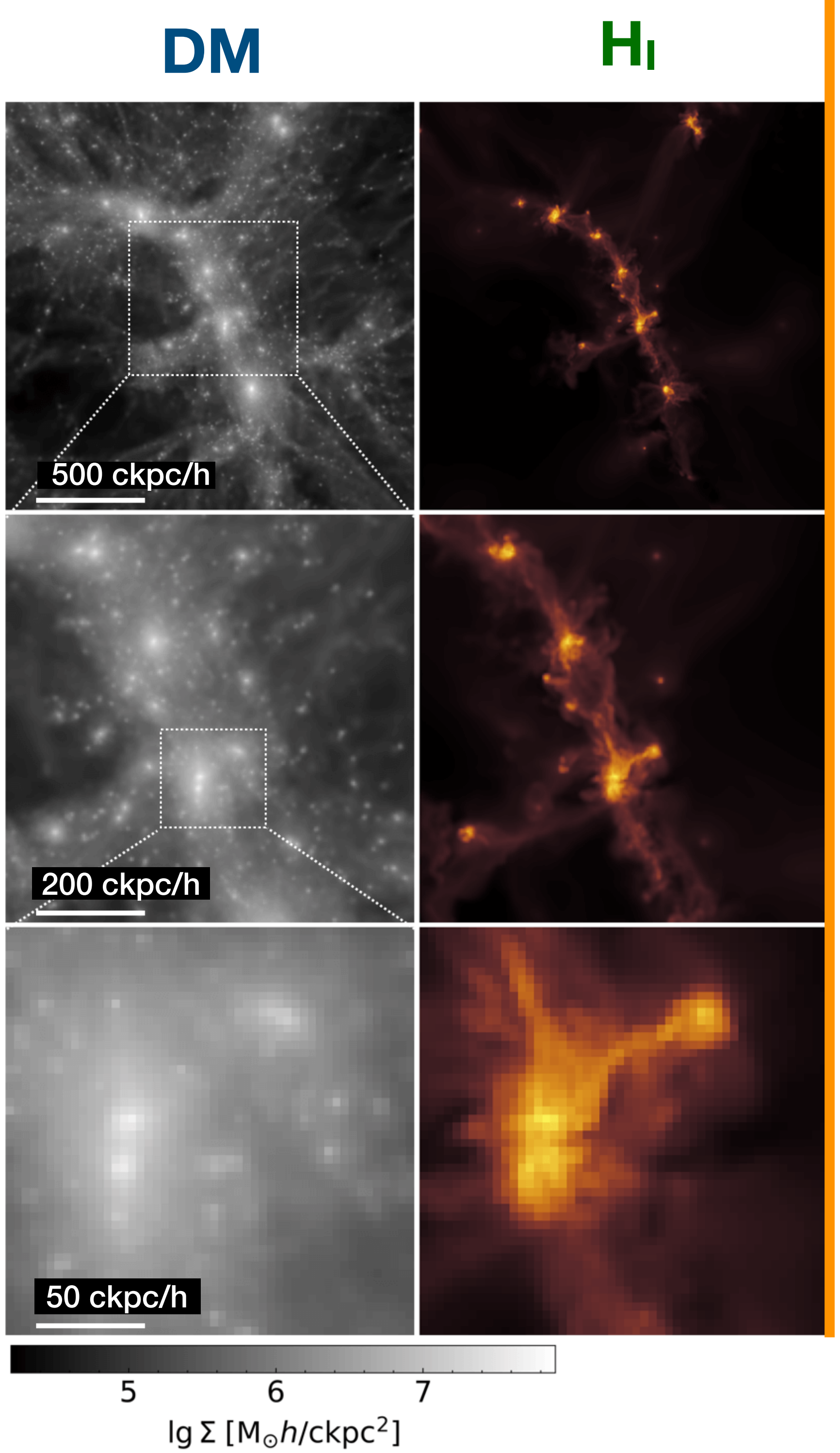
- Fully convolutional Generator = U-Net with noise injection
- conditional GAN using multi-scale mapping
- normalization scheme: $$\tilde{x} = \frac{1}{k} \log \left[\left(\frac{x}{x_0} \right)^q + 1 \right]$$
- Cost function:
  - perceptual loss based on image similarity index & MSLE
  - Wasserstein metric for adversarial loss (WGAN)



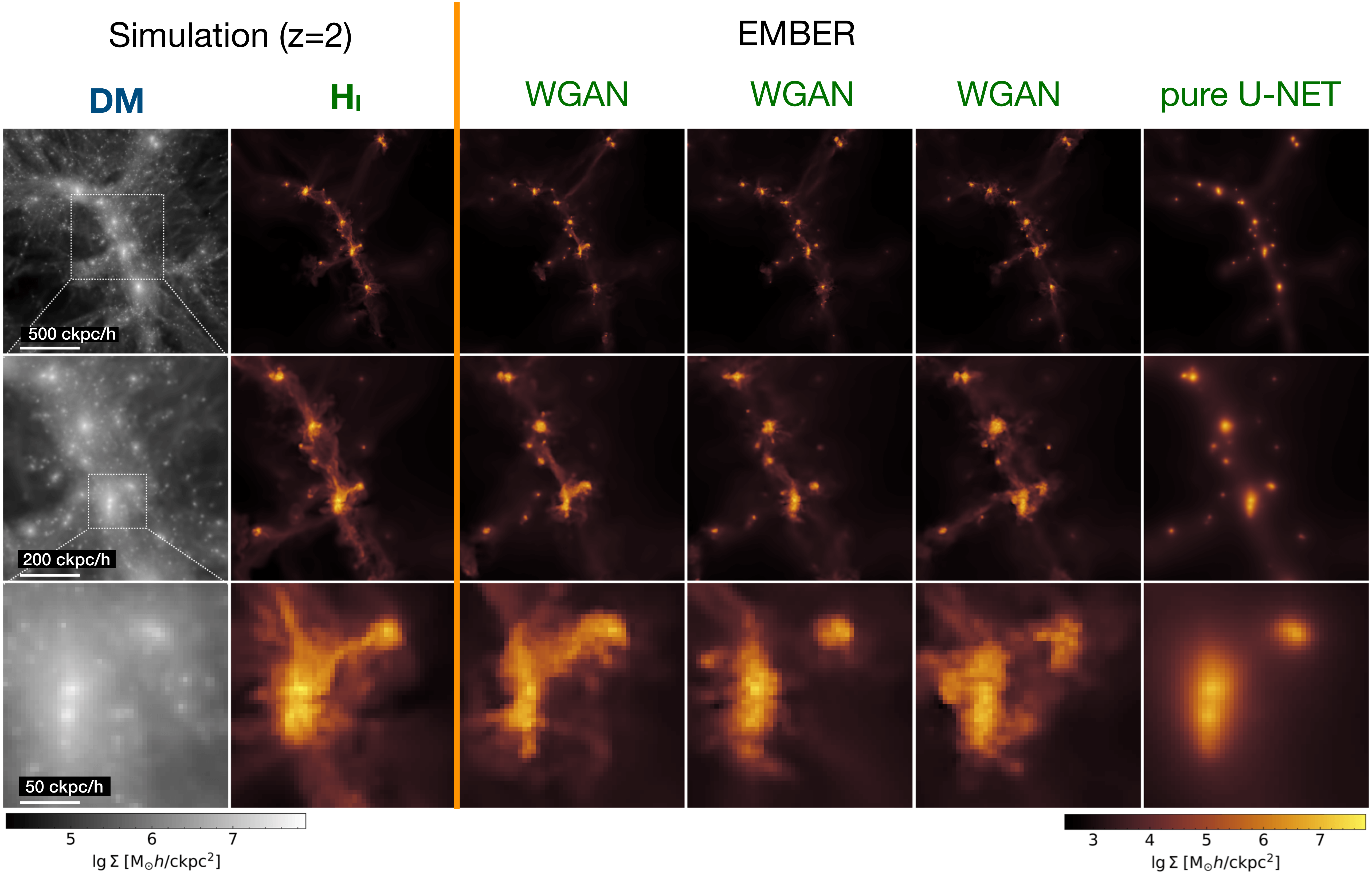
Bernardini, RF, et al. 2022, arXiv:2110.11970

How well does it work?

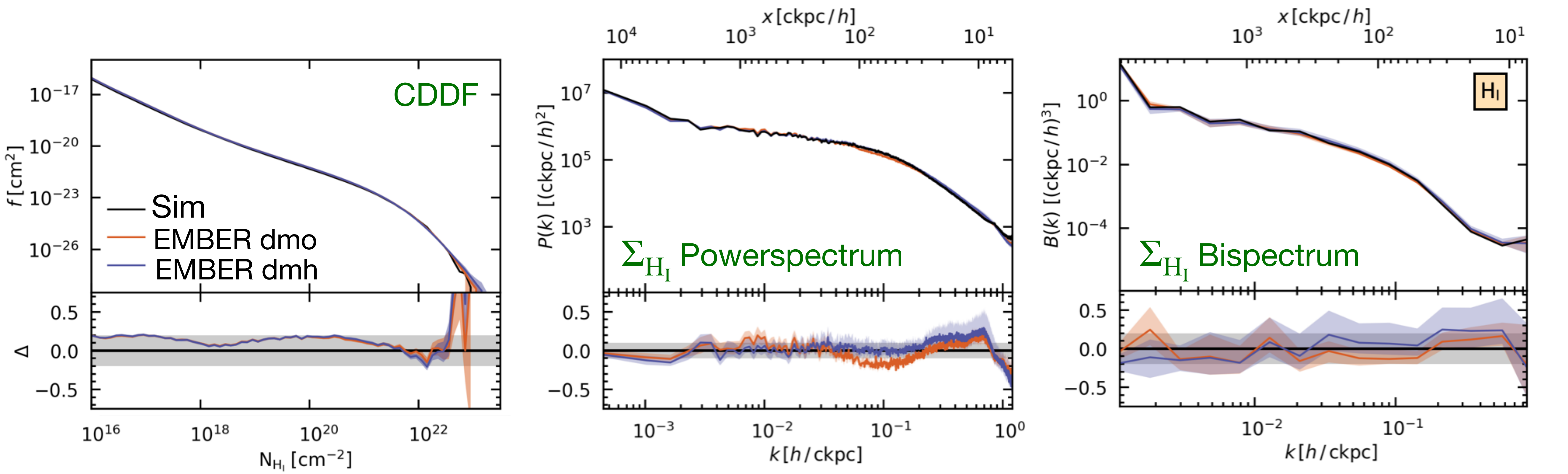
Simulation (z=2)



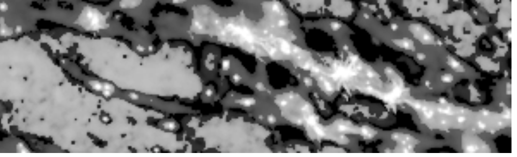
How well does it work?



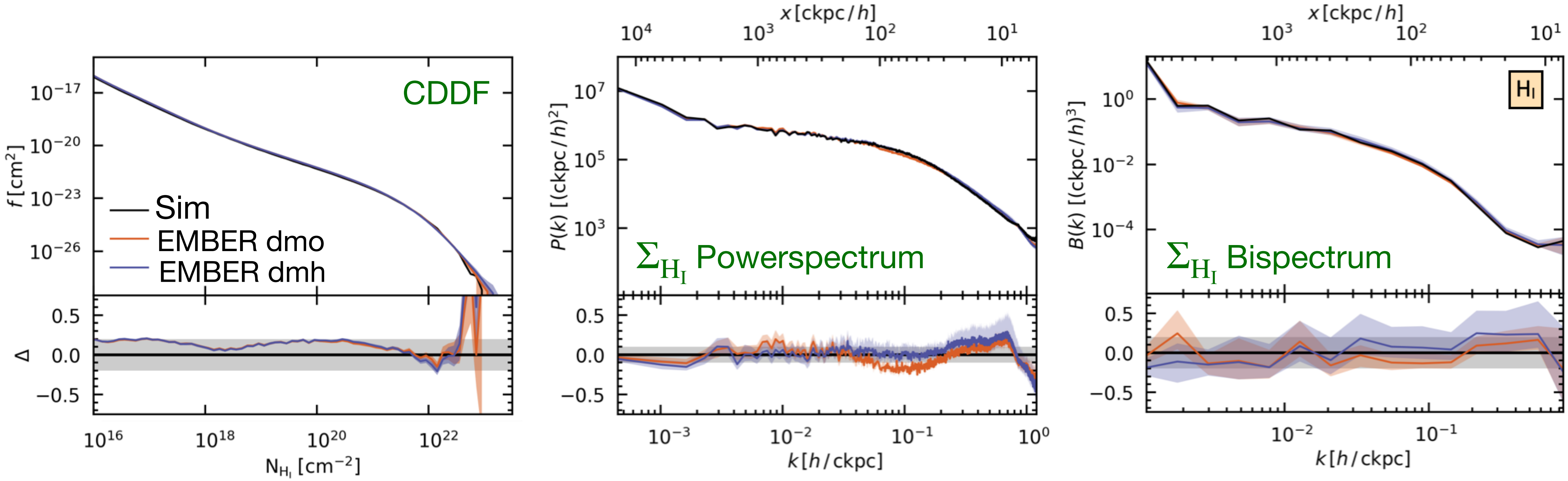
How well does it work?



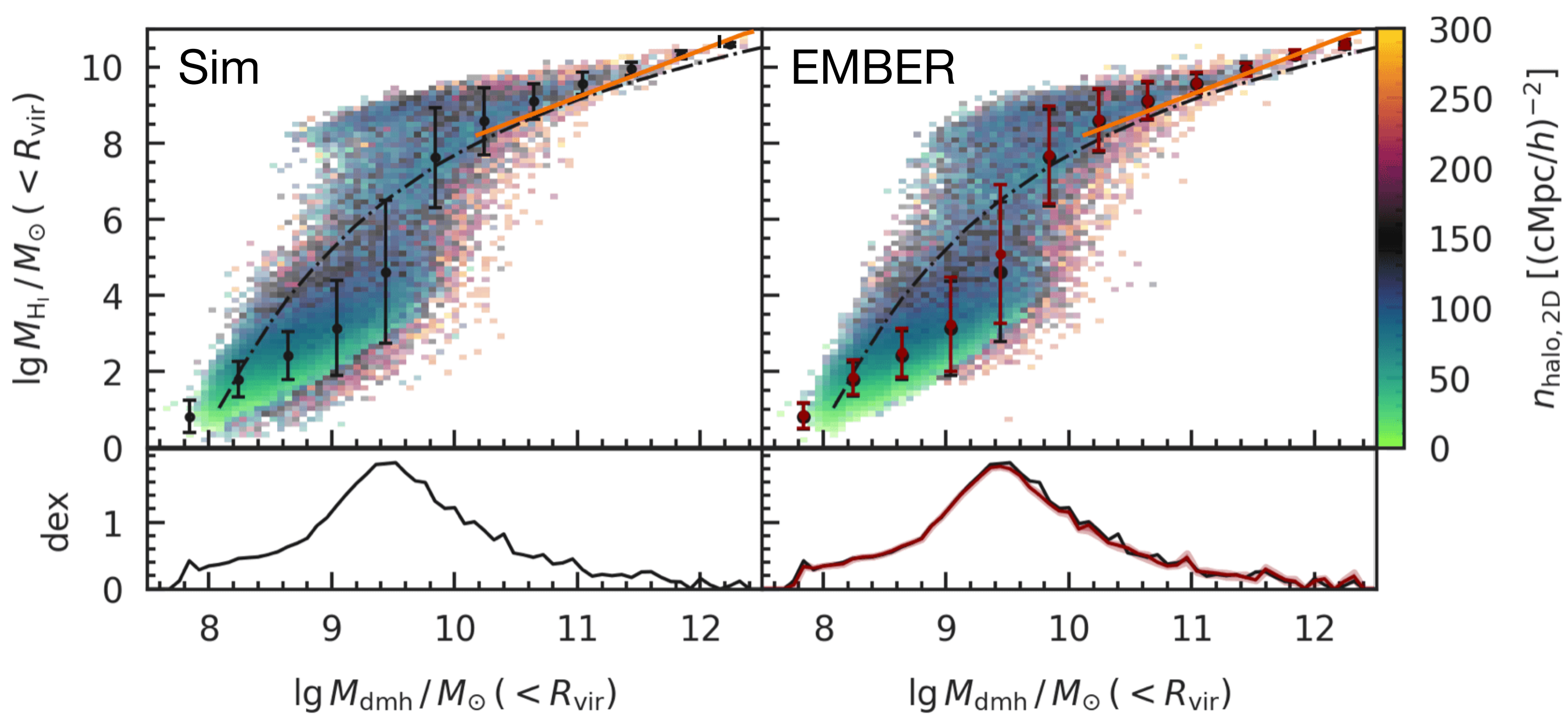
- Agreement to within 10-20% down to galactic scales!



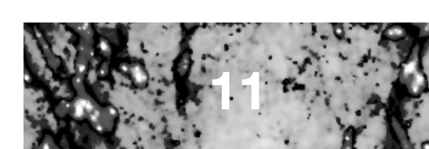
How well does it work?

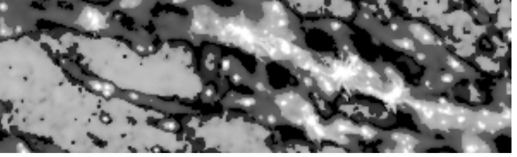


- Agreement to within 10-20% down to galactic scales!



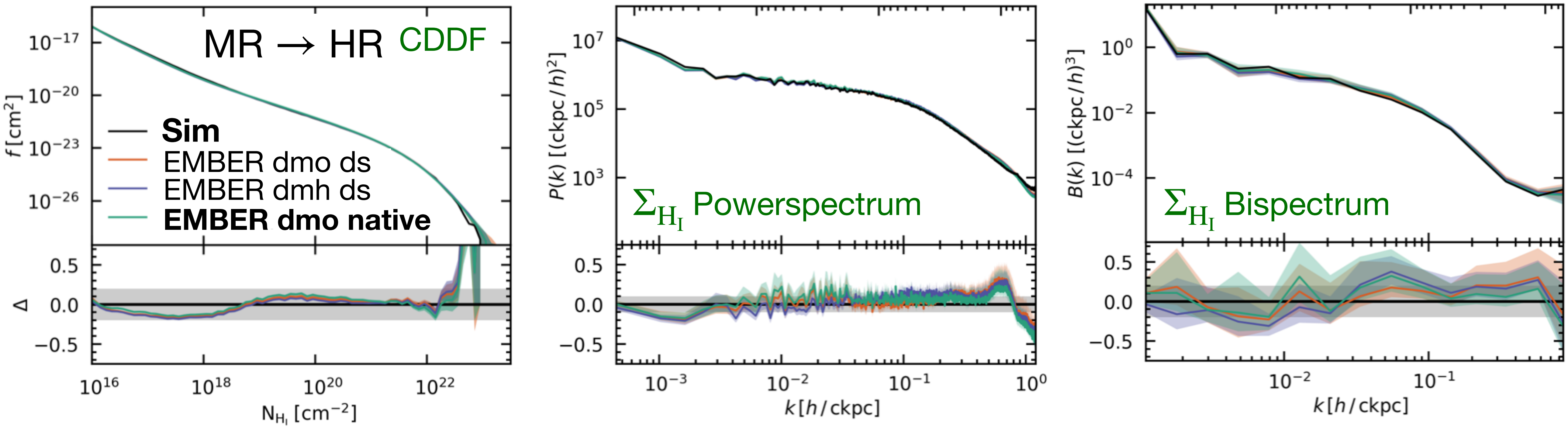
- Reproduces (projected) $\mathrm{H_I}$ masses in halos and scatter!



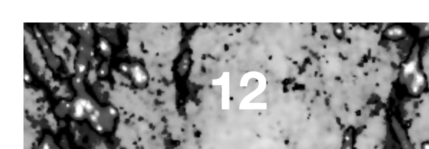


Upscaling

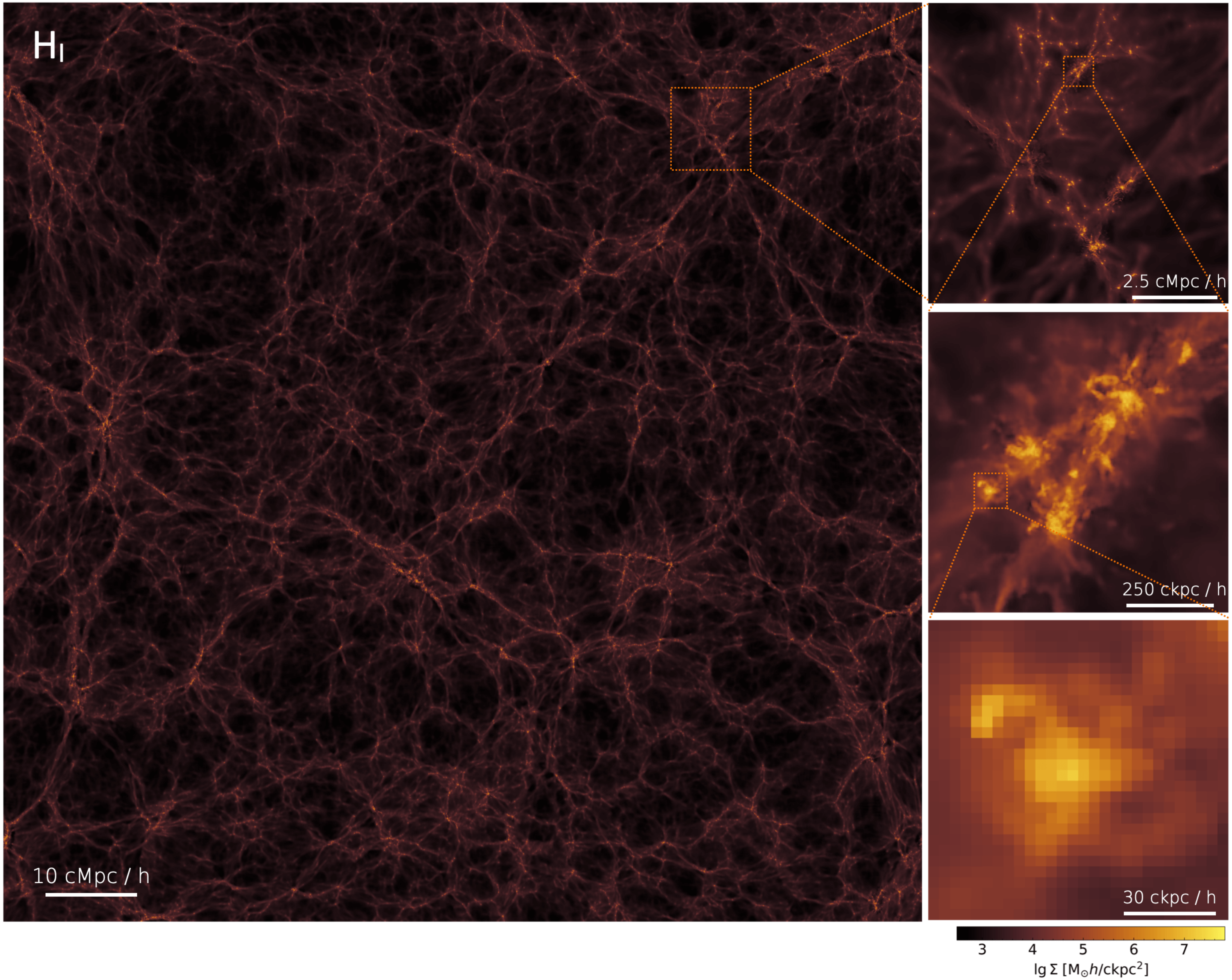
- Goal: create high-resolution HI maps from low resolution DM simulations



- 10-20% accurate when using 64x lower resolution DM maps



Upscaling



H$_{I}$
10 cMpc / h
2.5 cMpc / h
250 ckpc / h
30 ckpc / h
3 4 5 6 7
lg Σ [M$_\odot$$h$/ckpc^2]

Next steps

EMBER

- Training comparably expensive (~1 GPU-week)
- Makes predictions for fixed redshift
- Makes prediction for fixed output resolution
- Only single output (e.g., HI mass map or gas mass map)

Next steps

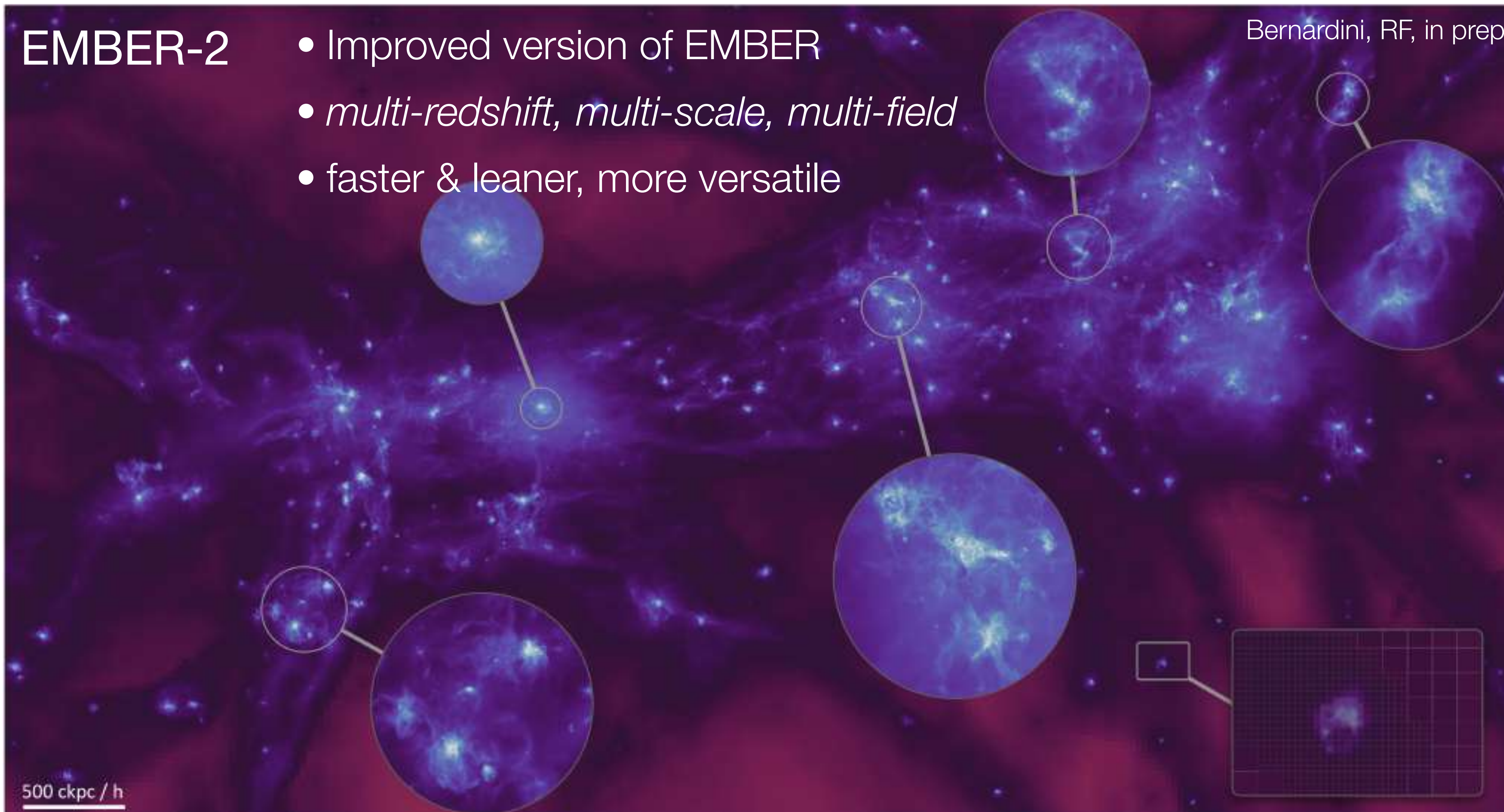
EMBER

- Training comparably expensive (~1 GPU-week)
- Makes predictions for fixed redshift
- Makes prediction for fixed output resolution
- Only single output (e.g., HI mass map or gas mass map)

EMBER-2

- Improved version of EMBER
- *multi-redshift, multi-scale, multi-field*
- faster & leaner, more versatile

Bernardini, RF, in prep



Challenges for EMBER-2

Achieving multi-redshift capability

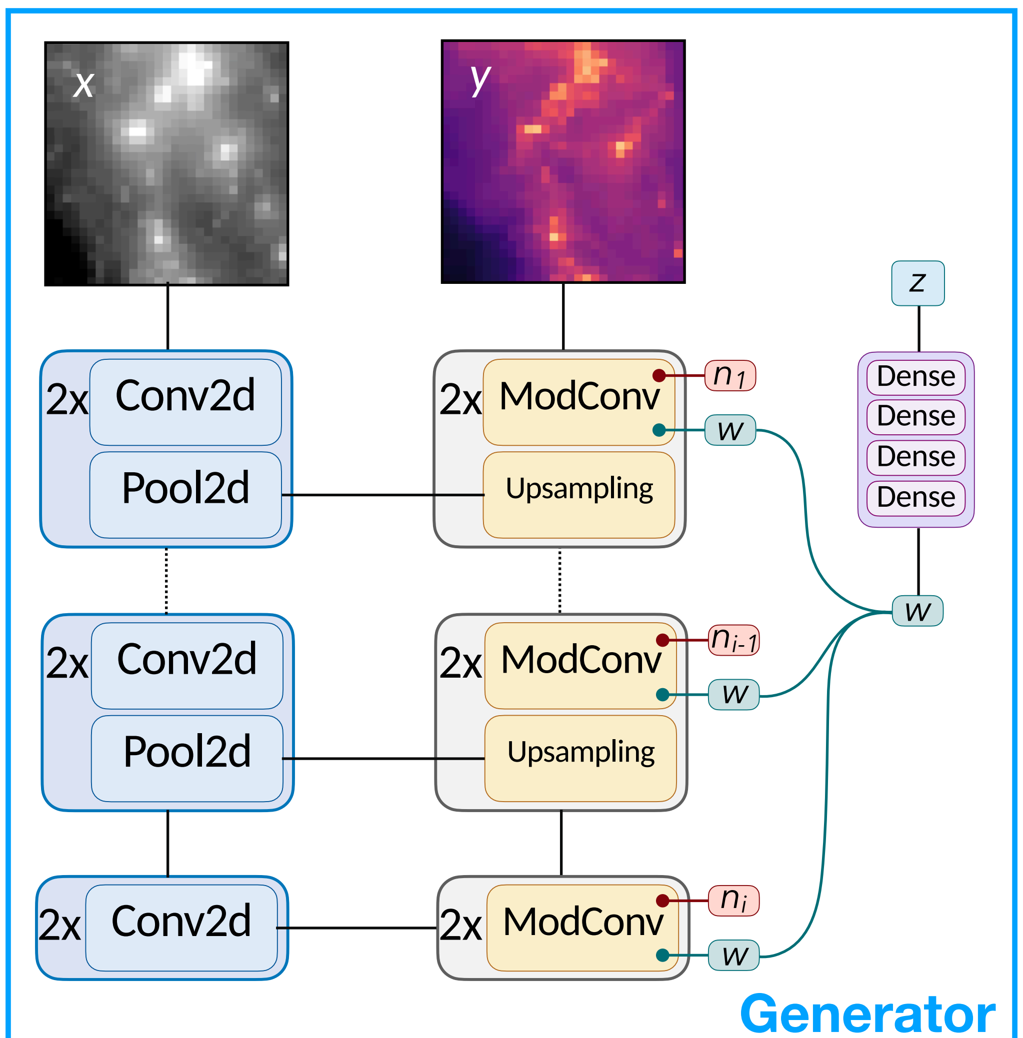
Coupling between DM and HI strongly redshift dependent, network would become too large if using standard convolutions

Achieving multi-redshift capability

Coupling between DM and HI strongly redshift dependent, network would become too large if using standard convolutions

Solution:

Inject the global context information (redshift) directly into the modulation of the convolution kernel (Style2-GAN approach)



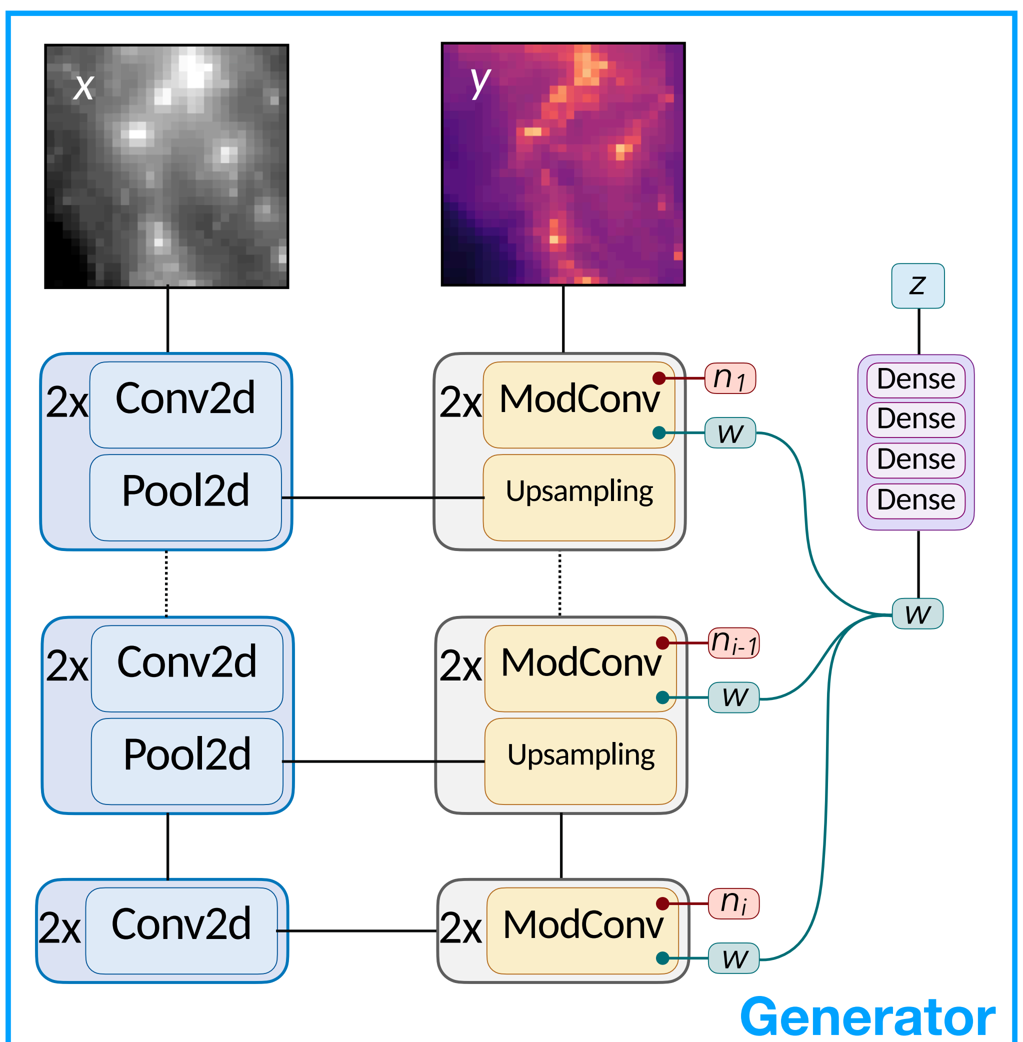
Challenges for EMBER-2

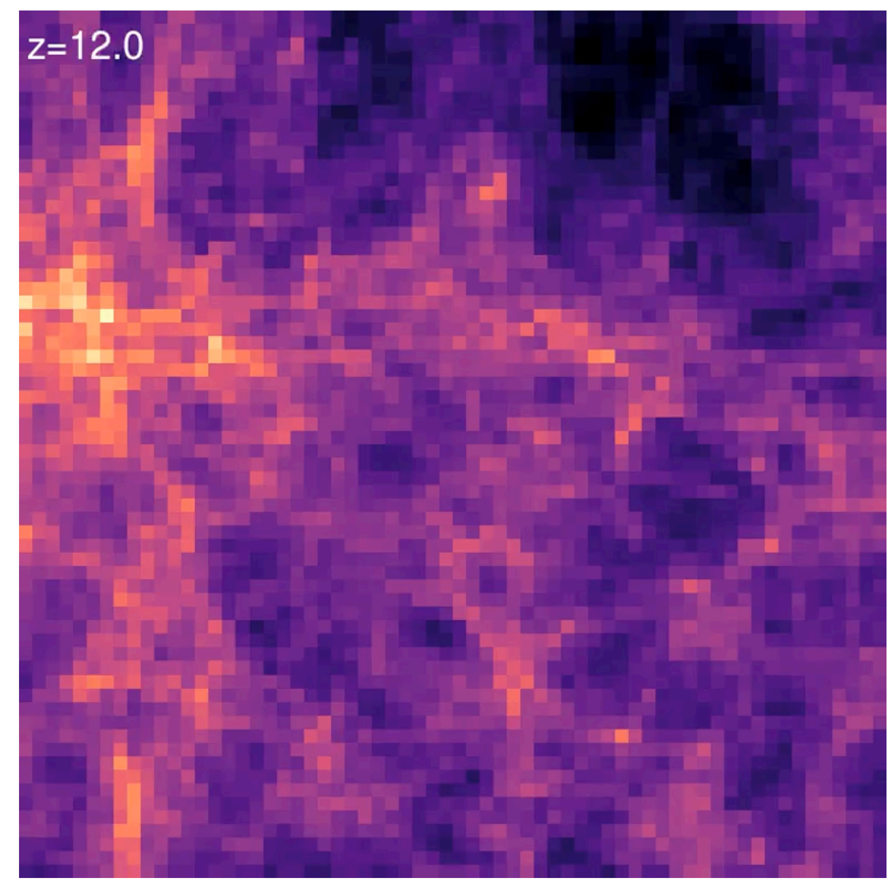
Achieving multi-redshift capability

Coupling between DM and HI strongly redshift dependent, network would become too large if using standard convolutions

Solution:

Inject the global context information (redshift) directly into the modulation of the convolution kernel (Style2-GAN approach)





Changing “redshift” while keeping DM map fixed

Challenges for EMBER-2

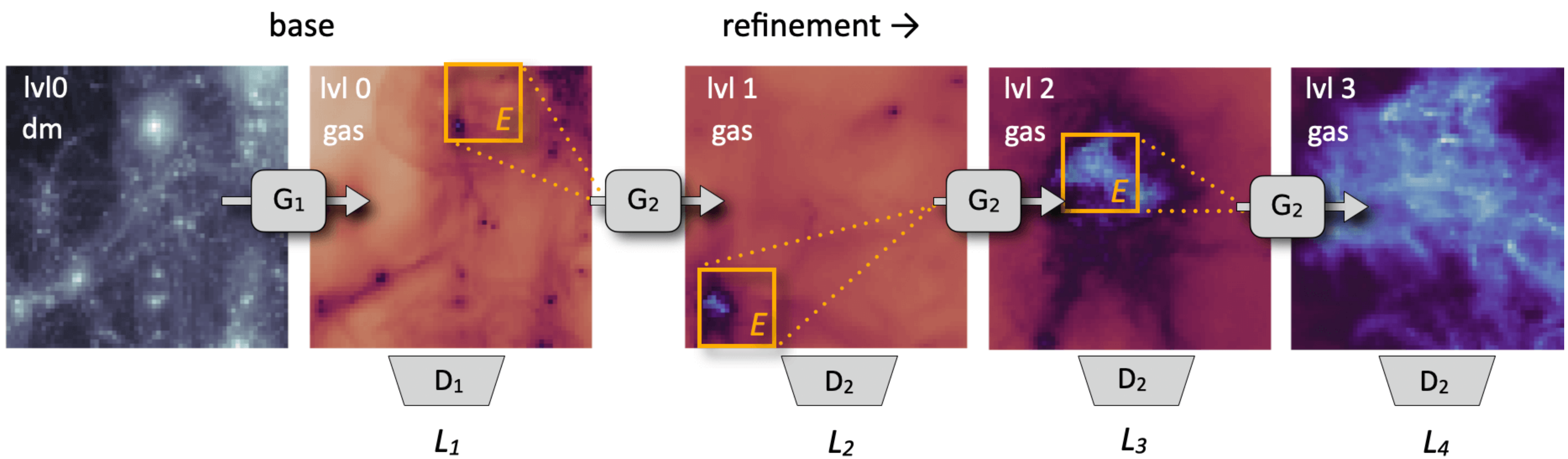
Achieving multi-scale capability

Normally trade-off between learning for high-resolution and including large environment given fixed tile size during training

Achieving multi-scale capability

Normally trade-off between learning for high-resolution and including large environment given fixed tile size during training

Solution:



Combine **base network** which learns large scale features and a **refinement network** which adds small-scale features given large scale features

Resolution level injected as context information (Style2-GAN) directly into the modulation of the convolution kernel of the refinement network

Achieving multi-field capability

- Consistency between output fields, e.g., temperature, density etc.
- Enable additional input fields $(\Sigma_{\rm DM}, v^{\rm rad}_{\rm DM}) \rightarrow (\Sigma_{\rm gas}, v^{\rm rad}_{\rm gas}, T_{\rm gas}, \Sigma_{\rm HI})$

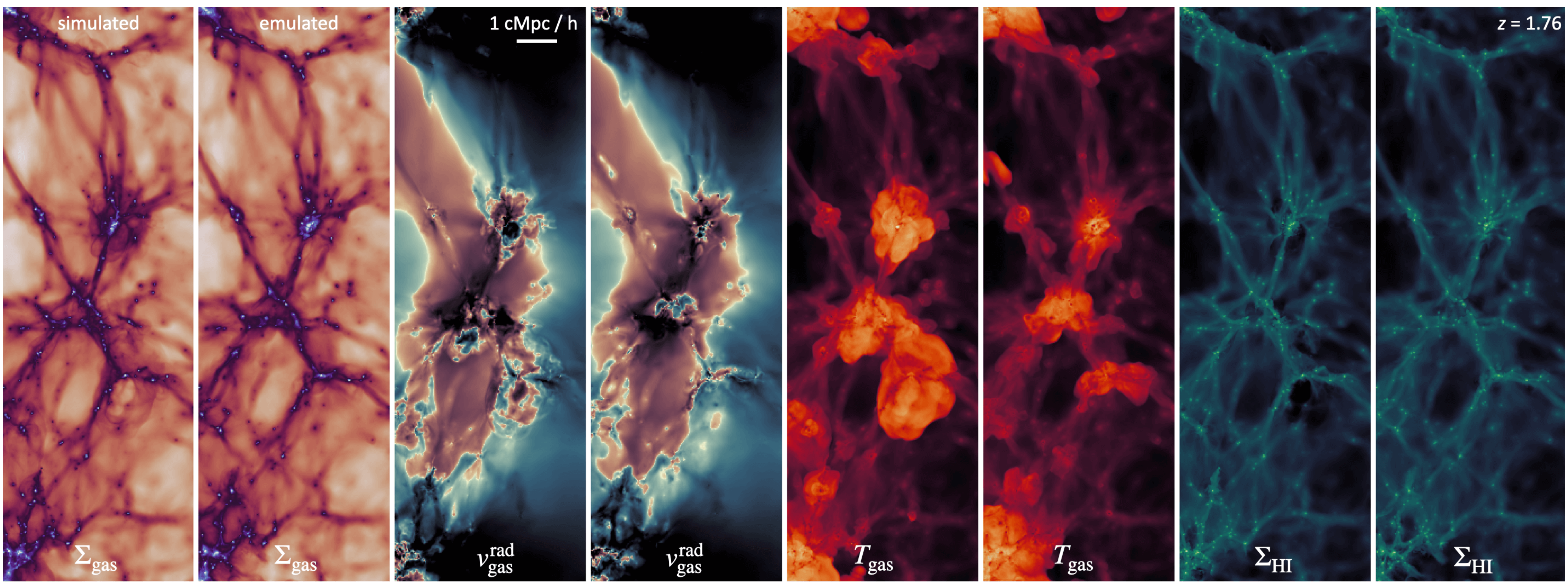
Achieving multi-field capability

- Consistency between output fields, e.g., temperature, density etc.
- Enable additional input fields $(\Sigma_{\rm DM}, v^{\rm rad}_{\rm DM}) \rightarrow (\Sigma_{\rm gas}, v^{\rm rad}_{\rm gas}, T_{\rm gas}, \Sigma_{\rm HI})$

Solution:

Conceptually straightforward: Include extra fields as additional channels

Sim **EMBER-2** **Sim** **EMBER-2** **Sim** **EMBER-2** **Sim** **EMBER-2**



First look at EMBER-2 performance

- EMBER 2.0 significantly leaner, faster

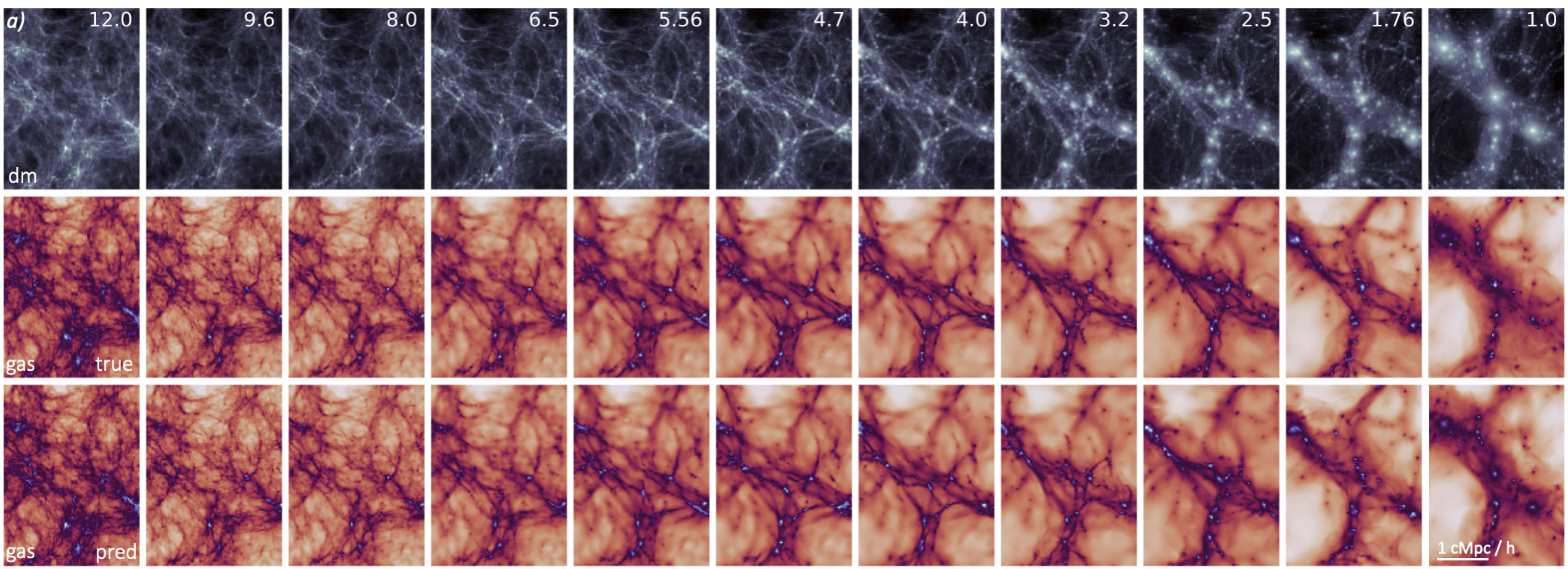
	# params [1e6]	Training time [d]	Inference time	Redshift	Tilesize
EMBER-1	80	7	minutes	z=2 fixed	1.8 cMpc/h
EMBER-2	18	2	seconds	z=12 to 1	1.8 cMpc/h

First look at EMBER-2 performance

- EMBER 2.0 significantly leaner, faster

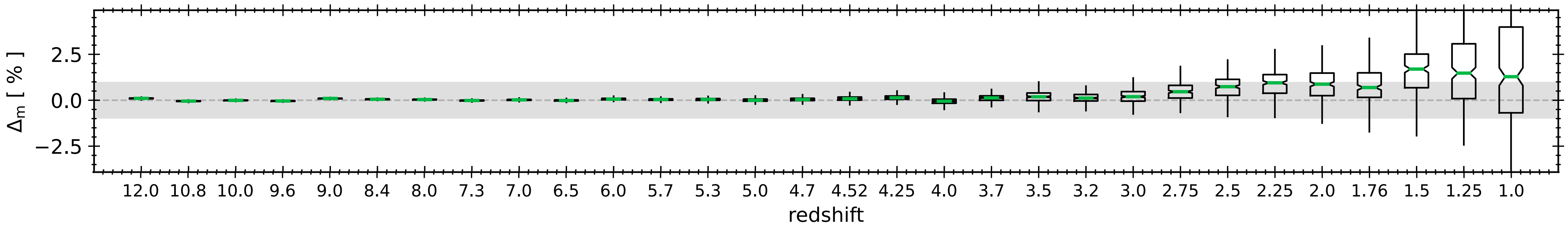
	# params [1e6]	Training time [d]	Inference time	Redshift	Tilesize
EMBER-1	80	7	minutes	z=2 fixed	1.8 cMpc/h
EMBER-2	18	2	seconds	z=12 to 1	1.8 cMpc/h

- Visually accurate predictions for any redshift z=1-12



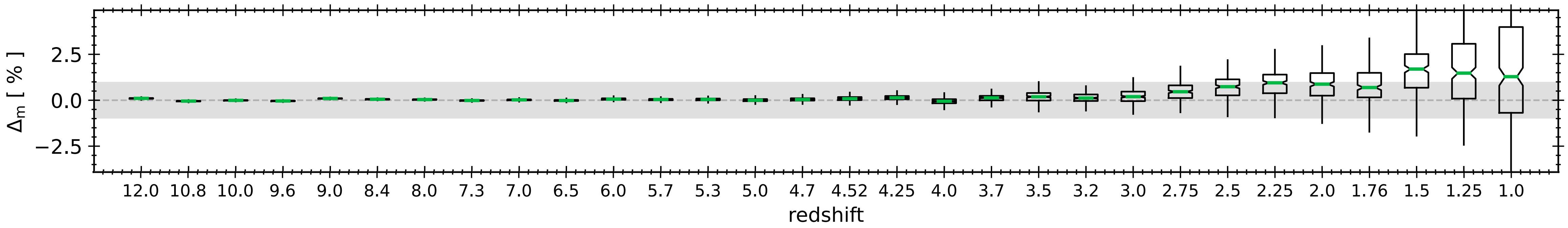
First look at EMBER-2 performance

- Excellent mass conservation for slices (~1%)

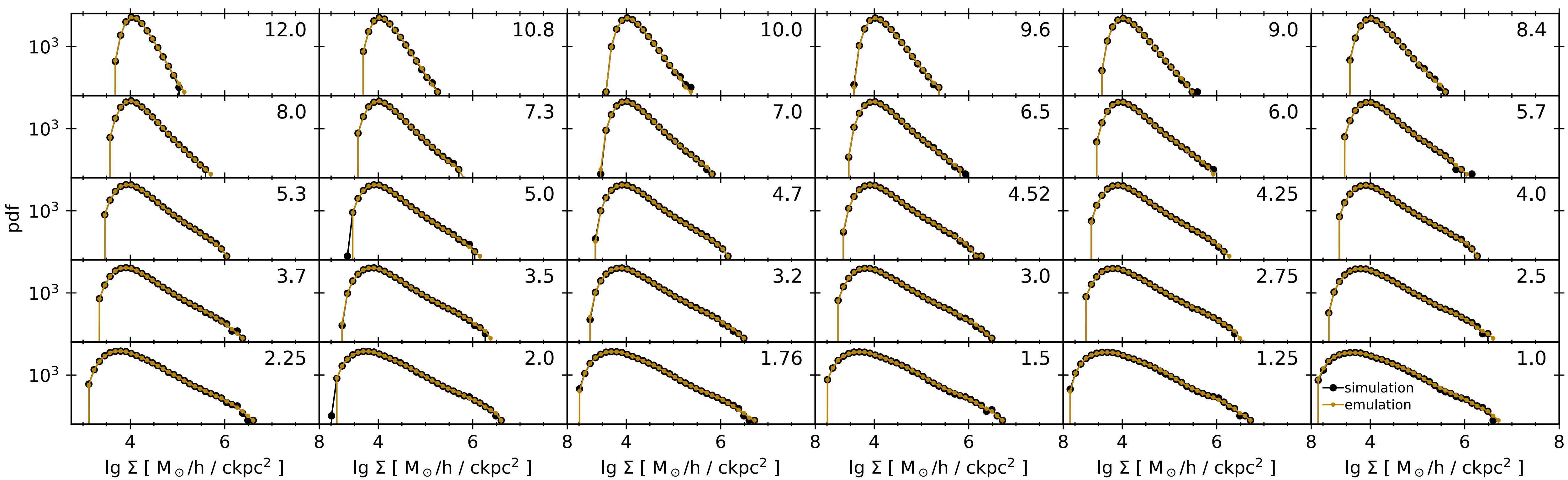


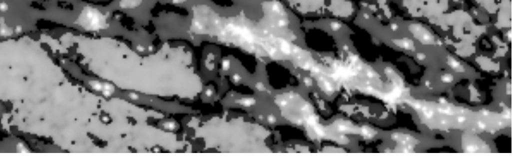
First look at EMBER-2 performance

- Excellent mass conservation for slices (~1%)



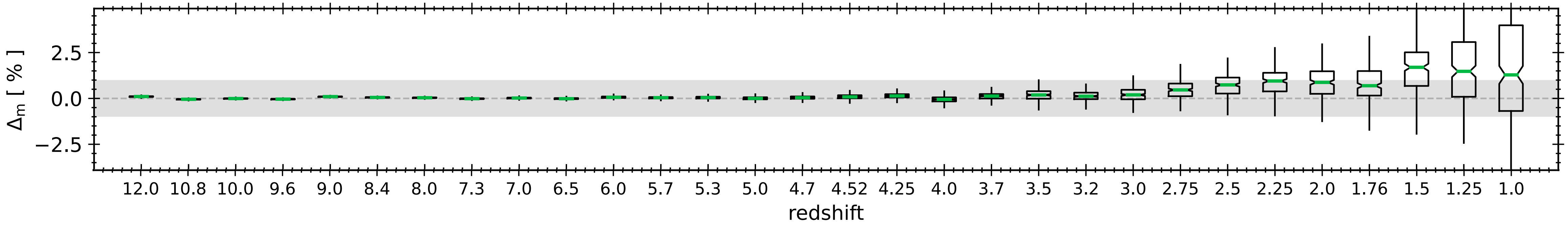
- Excellent agreement for the pixel probability density functions



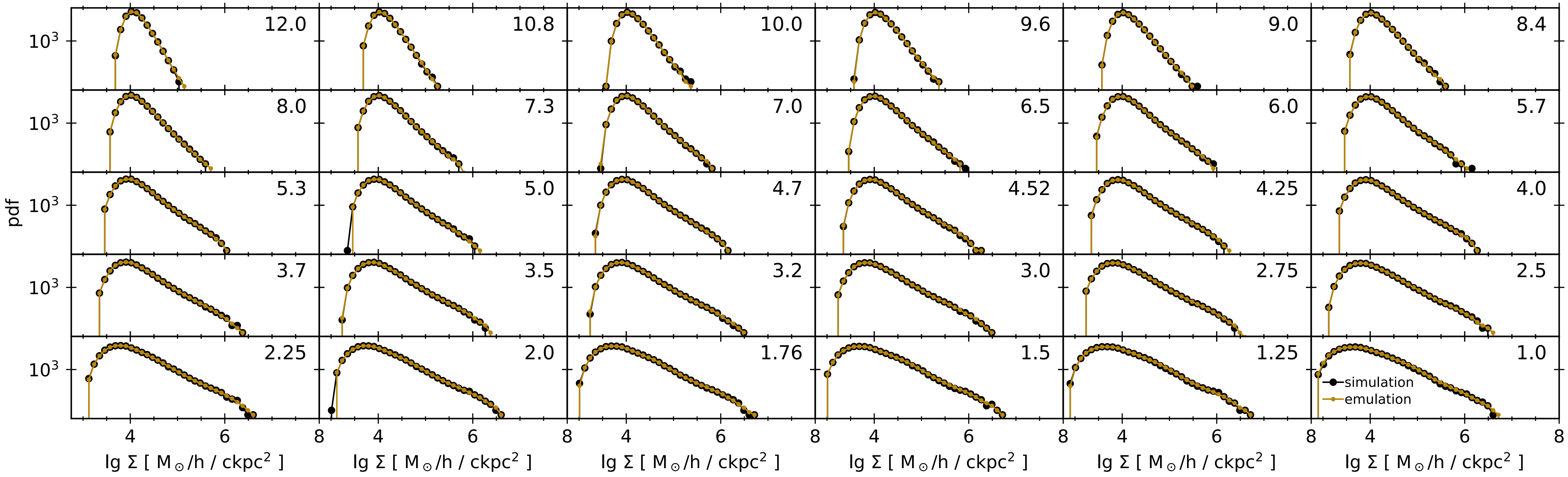


First look at EMBER-2 performance

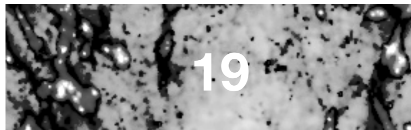
- Excellent mass conservation for slices (~1%)



- Excellent agreement for the pixel probability density functions

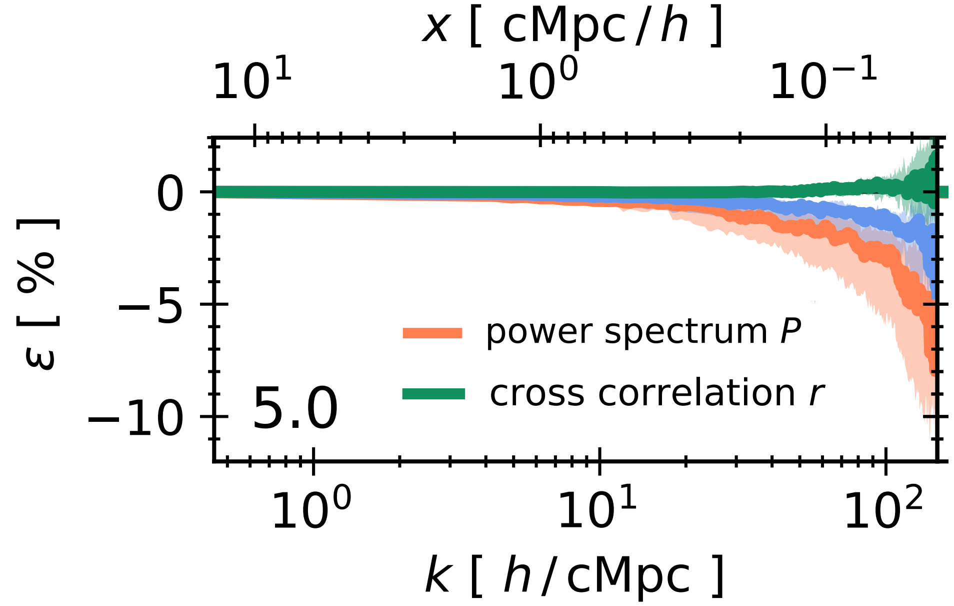


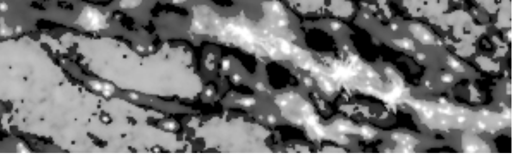
- What about higher order statistics?



First look at EMBER-2 performance

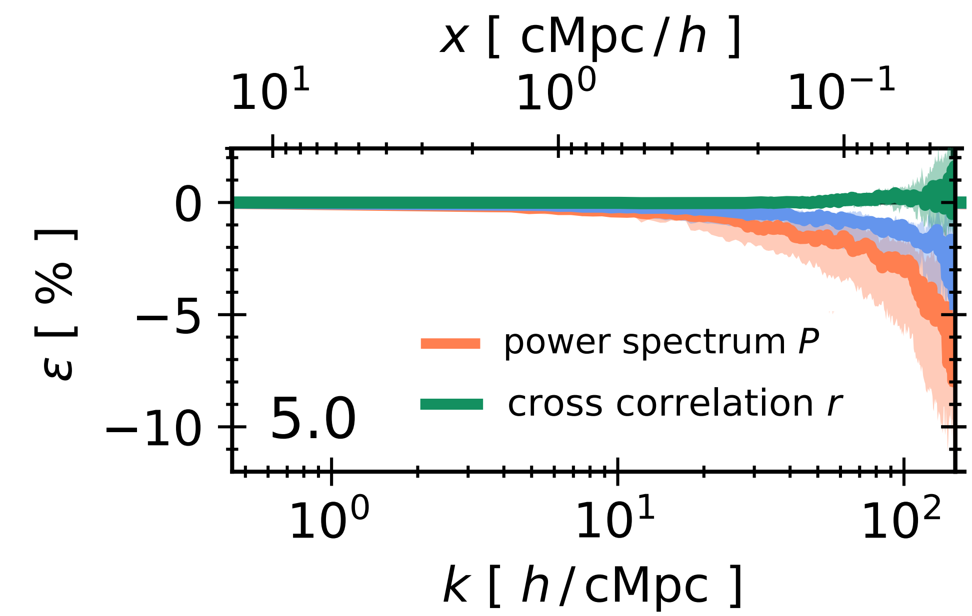
- Often employed metric to measure performance of emulator is power-spectrum
- Small box => high variance
- Test: modify slices by lowering mass in 5 most massive pixels by 10%





First look at EMBER-2 performance

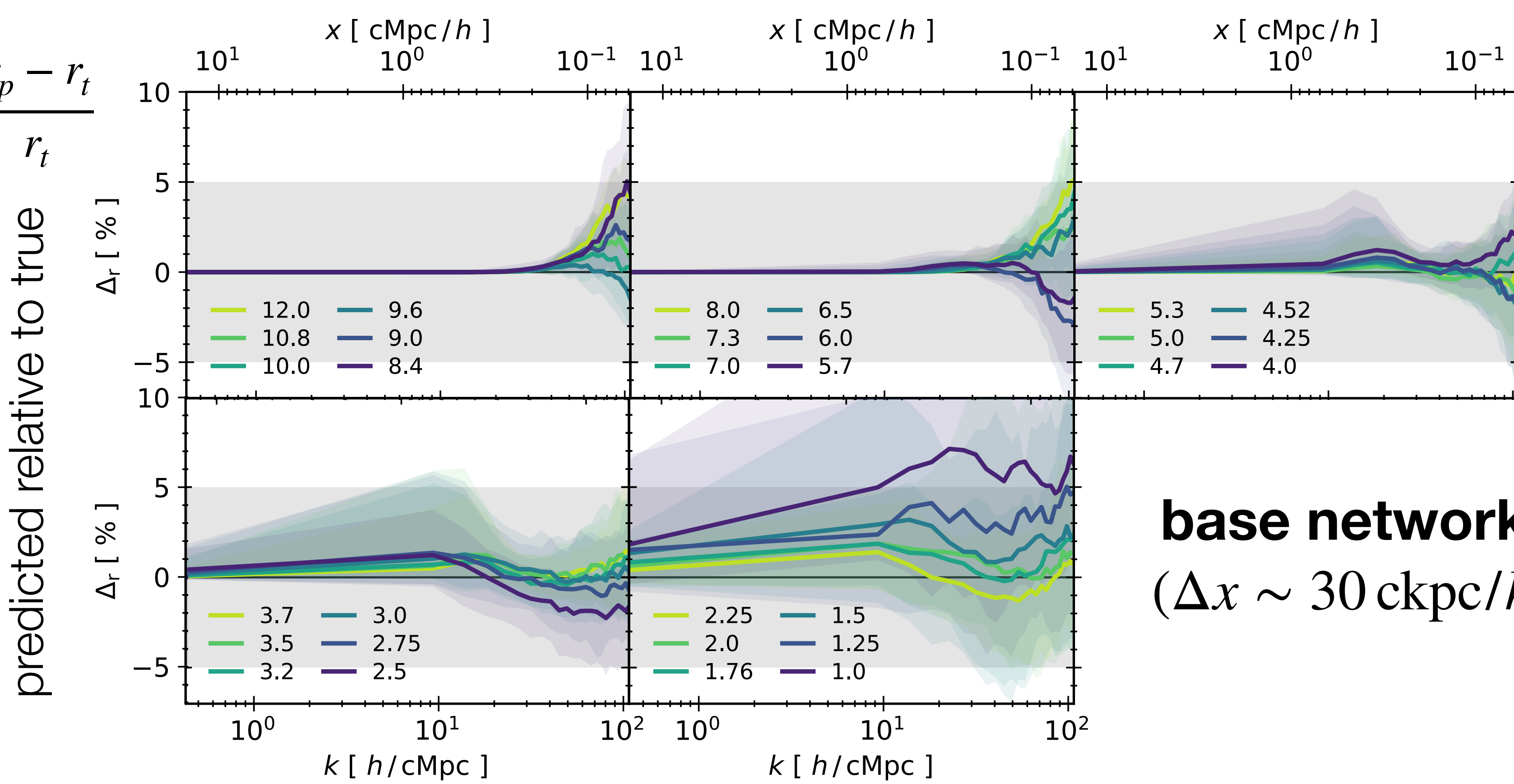
- Often employed metric to measure performance of emulator is power-spectrum
- Small box => high variance
- Test: modify slices by lowering mass in 5 most massive pixels by 10%



- Better performance metric: cross correlation between DM map and gas map:

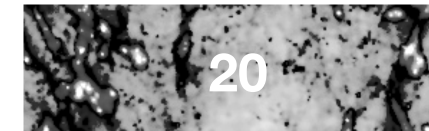
$$r(k) = P^{\times}_{dm,g} / (P_{dm} P_g)^{1/2}$$

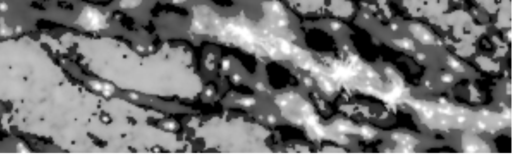
$$\Delta r = \frac{r_p - r_t}{r_t}$$



base network!
$(\Delta x \sim 30\,\mathrm{ckpc}/h)$

➜ Cross-correlation reproduced to better than ~5% at z=1-12 up to ~50 ckpc





Summary

- Deep Learning techniques can reduce computational cost of calculating baryonic fields compared with high-resolution (hydrodynamical) simulations

- EMBER: stochastic, conditional, convolutional GAN that learns baryon maps from cosmological, hydrodynamical simulations; trained on FIRE simulations
- 10-20% accurate mapping down to galactic scales (~tens of kpc) for CDDF, powerspectrum, and bispectrum
- Entirely halo-free method, but can make accurate predictions for halo based properties (including scatter) down to low halo masses; upscaling to large volumes

- EMBER-2: upgraded version of EMBER with new architecture
- Adds *multi-redshift*, *multi-scale*, and *multi-field* capabilities
- Faster to train, faster to run, more versatile