HITS 2022, 23-27 May 2022

# Painting HI onto the dark matter field for mock catalogs generation

**Francesco Sinigaglia**

francesco.sinigaglia@phd.unipd.it

# Outline

- HI mapping model
- Full pipeline: from the primordial density field to HI field
- Deep learning extension
- Applications
- Work still to do

# The CosmicAtlas project

**GOAL**

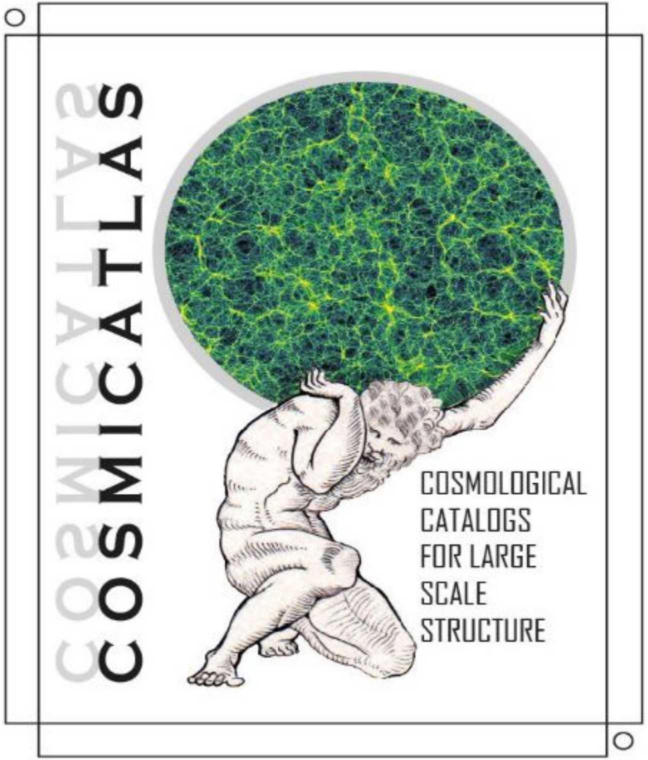
Produce halo/galaxy/Lyman-alpha/HI IM/quasar mock catalogs for forthcoming surveys:
**Bias Assignment Method (BAM)**

**Developers**

Andrés Balaguera-Antolínez (Advanced Severo Ochoa Fellow, IAC, ULL)
Francisco-Shu Kitaura (Ramón y Cajal Fellow, IAC/ULL)
**Myself**

**Collaborators**

Marc Huertas-Company (IAC), Ariel Sánchez (MPE), Chia-Hsun Chuang (Stanford), Shadab Alam (Edinburg) Yu-Yu (Beijin), Cheng Zhao (EPFL), Kentaro Nagamine (Osaka), Metin Ata (Kavli-IPMU)



# The CosmicAtlas project: BAM

**BAM: Bias Assignment Method to generate mock catalogs**

(arXiv:1806.05870)

A. Balaguera-Antolínez★[1,2], Francisco-Shu Kitaura†[1,2], Marcos Pellejero-Ibáñez[1,2], Cheng Zhao[3] and Tom Abel[4]

**One simulation to have them all: performance of the Bias Assignment Method against $N$-body simulations**

(arXiv:1906.06109)

A. Balaguera-Antolínez★[1,2],Francisco-Shu Kitaura†[1,2], M. Pellejero-Ibáñez[3], Martha Lippich[4], Cheng Zhao[5], Ariel G. Sánchez[4], Claudio Dalla Vecchia[1,2], Raúl E. Angulo[3,6] and Martín Crocce[7]

**The bias of dark matter tracers: assessing the accuracy of mapping techniques**

(arXiv:1910.13164)

M. Pellejero-Ibañez ★[1], A. Balaguera-Antolínez†[2,3], Francisco-Shu Kitaura‡[2,3] Raúl E. Angulo[1,4], Gustavo Yepes[5,6], Chia-Hsun Chuang[7], Guillermo Reyes-Peraza[8], Mathieu Autefage[9], Mohammadjavad Vakili[10] & Cheng Zhao[11]



Universidad de La Laguna

OSSERVATORIO ASTRONOMICO DI PADOVA

# The BAM philosophy

BAM learns the **DM-tracers bias relation** from **one** suitable reference simulation

$$\delta'_{\mathrm{tr}} \curvearrowleft \mathcal{B} = P(\delta_{\mathrm{tr}} \,|\, \delta_{\mathrm{dm}}, X, Y, \ldots)$$

# The BAM philosophy

BAM learns the **DM-tracers bias relation** from **one** suitable reference simulation

$$\delta'_{\rm tr} \curvearrowleft \mathcal{B} = P(\delta_{\rm tr} \,|\, \delta_{\rm dm}, X, Y, \ldots)$$

- full non-linear relation (deterministic + stochastic), local & non-local, parameter-free formulation

# The BAM philosophy

BAM learns the **DM-tracers bias relation** from **one** suitable reference simulation

$$\delta'_{\rm tr} \curvearrowleft \mathcal{B} = P(\delta_{\rm tr} \,|\, \delta_{\rm dm}, X, Y, \ldots)$$

- full non-linear relation (deterministic + stochastic), local & non-local, parameter-free formulation
- 1-point PDF reconstructed by construction

# The BAM philosophy

BAM learns the **DM-tracers bias relation** from **one** suitable reference simulation

$$\delta'_{\mathrm{tr}} \curvearrowleft \mathcal{B} = P(\delta_{\mathrm{tr}} \,|\, \delta_{\mathrm{dm}}, X, Y, \ldots)$$

- full non-linear relation (deterministic + stochastic), local & non-local, parameter-free formulation
- 1-point PDF reconstructed by construction
- $P(k)$ constrained through iterative convolutions with an isotropic kernel $K(k)$

# The BAM philosophy

BAM learns the **DM-tracers bias relation** from **one** suitable reference simulation

$$\delta'_{\rm tr} \curvearrowleft \mathcal{B} = P(\delta_{\rm tr} \,|\, \delta_{\rm dm}, X, Y, \ldots)$$

- full non-linear relation (deterministic + stochastic), local & non-local, parameter-free formulation
- 1-point PDF reconstructed by construction
- $P(k)$ constrained through iterative convolutions with an isotropic kernel $K(k)$
- higher order statistics not constrained: bispectrum good proxy for the accuracy of our model

# The BAM philosophy

BAM learns the **DM-tracers bias relation** from **one** suitable reference simulation

$$\delta'_{\mathrm{tr}} \curvearrowleft \mathcal{B} = P(\delta_{\mathrm{tr}} \,|\, \delta_{\mathrm{dm}}, X, Y, \ldots)$$

- full non-linear relation (deterministic + stochastic), local & non-local, parameter-free formulation
- 1-point PDF reconstructed by construction
- $P(k)$ constrained through iterative convolutions with an isotropic kernel $K(k)$
- higher order statistics not constrained: bispectrum good proxy for the accuracy of our model
- Connection to bias PT expansion: $\delta_{\mathrm{tr}} = c_0 + c_1\,\delta_{\mathrm{dm}} + c_2\,\delta_{\mathrm{dm}}^{\,2} + \ldots$ (incl. non-local terms)

(Kitaura *et al.* 2022, incl FS)
**(arXiv:2012.06795)**

# The BAM philosophy

BAM learns the **DM-tracers bias relation** from **one** suitable reference simulation

$$\delta'_{tr} \curvearrowleft \mathcal{B} = P(\delta_{tr} \,|\, \delta_{dm}, X, Y, \ldots)$$

- full non-linear relation (deterministic + stochastic), local & non-local, parameter-free formulation
- 1-point PDF reconstructed by construction
- $P(k)$ constrained through iterative convolutions with an isotropic kernel $K(k)$
- higher order statistics not constrained: bispectrum good proxy for the accuracy of our model
- Connection to bias PT expansion: $\delta_{tr} = c_0 + c_1\,\delta_{dm} + c_2\,\delta_{dm}{}^2 + \ldots$ (incl. non-local terms)

(Kitaura *et al.* 2022, incl FS)
**(arXiv:2012.06795)**

**Strongly-supervised physically-motivated ML:** lack of information in the model compensated by minimizing $|P'_{tr}(k)\text{-}P_{tr}(k)|^2$

# Non-local bias & cosmic web

Different ways to capture non-locality and the CW dependency:

- T-web (Hahn *et al.* 2007): knots, filaments, sheets, voids

  $\mathrm{sgn}(\lambda_1, \lambda_2, \lambda_3)>0$, $\mathrm{sgn}(\lambda_1, \lambda_2, \lambda_3)<0$, $\lambda_i$ eigenvalue of $T_{ij} = \partial_i \partial_j \varphi$

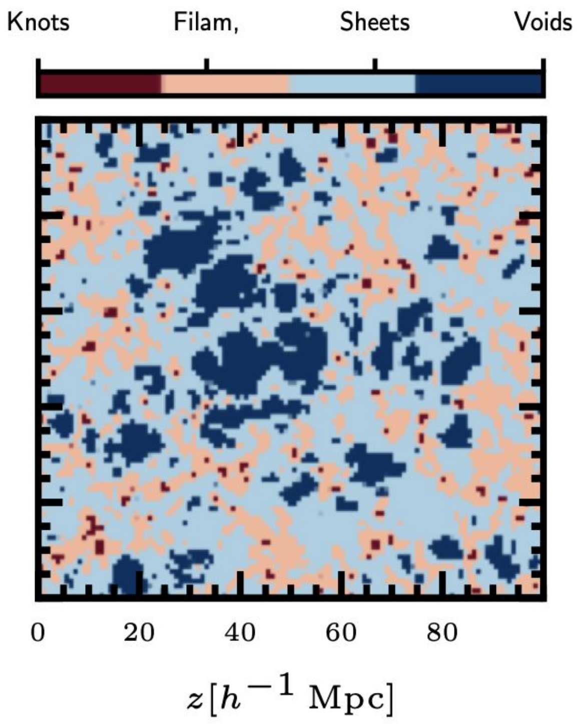
- $\mathrm{I}^{\varphi}$-web (Kitaura *et al.* 2022, incl. FS): invariants of $T_{ij}$

  **(arXiv:2012.06795)**

  $$I_1 = \lambda_1 + \lambda_2 + \lambda_3$$
  $$I_2 = \lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_1\lambda_3$$
  $$I_3 = \lambda_1\lambda_2\lambda_3$$

  Models halo bias PT up to 3rd order + is equivalent to T-web

- $\mathrm{I}^{\delta}$-web (Heavens and Peacock 1986, Sinigaglia *et al.* 2021): Invariants of $\delta_{ij} = \partial_i \partial_j \delta$ **(arXiv:2012.06795)**



# Hydro-BAM: cosmological volumes with baryons

A novel application of the BAM approach

## Goal

Map gas properties in cosmological volumes onto dark matter fields

## State of the art

***Full N-body + Hydro***

Hydro simulations

DM N-body + SAMs

***Predictive***

Fast PM codes (FastPM, COLA, ...)

Approx. gravity solvers (LPT, 2LPT, ALPT) + SAMs

Gaussian random fields + SAMs

**Calibrated**

ML/DL

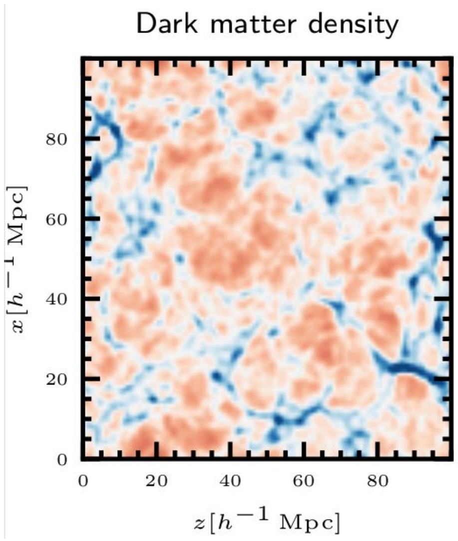
Domain specific methods

**Test case:** $V = (100\ h^{-1}\ \mathrm{Mpc})^3$ , $2 \times 512^3$ particles interpolated on cubic $128^3$ cells mesh with CIC @ $z$=2

→ $l_{\mathrm{cell}} \sim 0.78\ h^{-1}\ \mathrm{Mpc},\ \ k_{\mathrm{nyq}} \sim 4.0\ h\ \mathrm{Mpc}^{-1}$

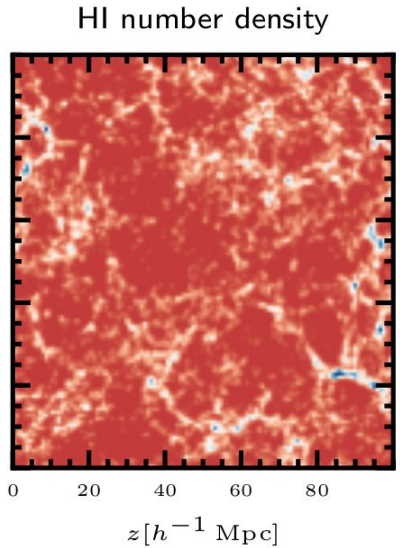
# Hierarchical bias mapping

Sinigaglia et al. (2021), ApJ **(arXiv:2012.06795)**




$P(\delta_{HI} | \delta_{dm})$

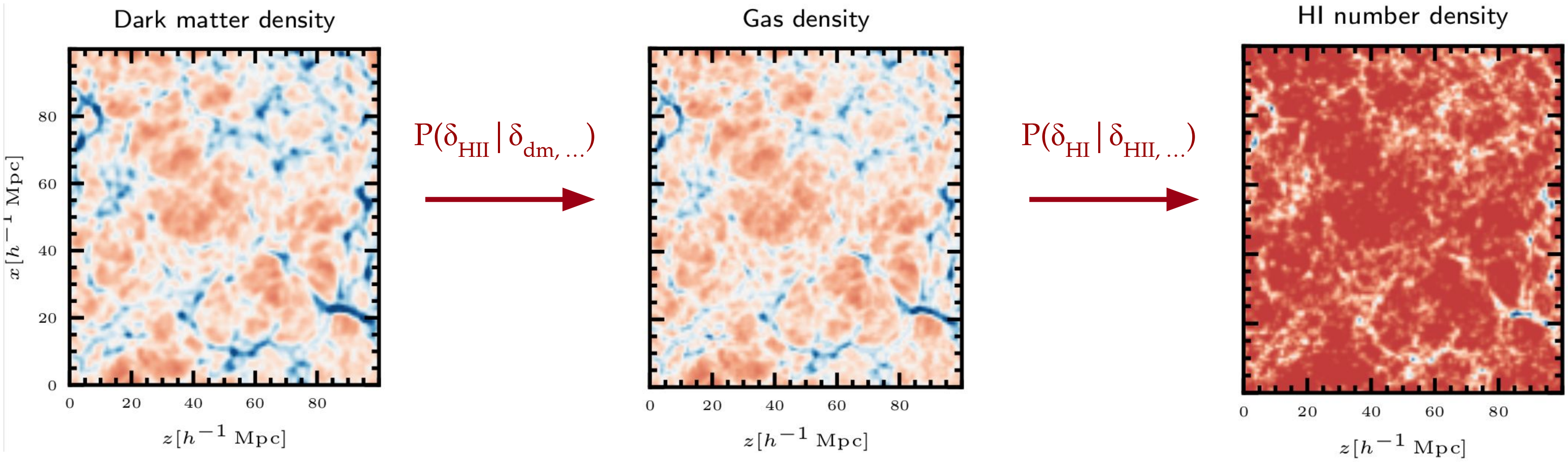
Large scales accurate, small scales not quite...




*Painting HI onto the DM field for mock catalogs generation*

# Hierarchical bias mapping

Sinigaglia et al. (2021), ApJ **(arXiv:2012.06795)**

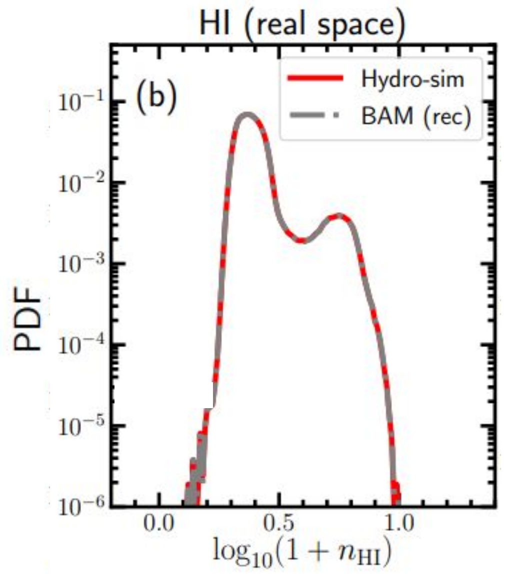



Model HI spatial distribution and clustering with a hierarchical approach: captures both LSS and baryon effects

*Painting HI onto the DM field for mock catalogs generation*

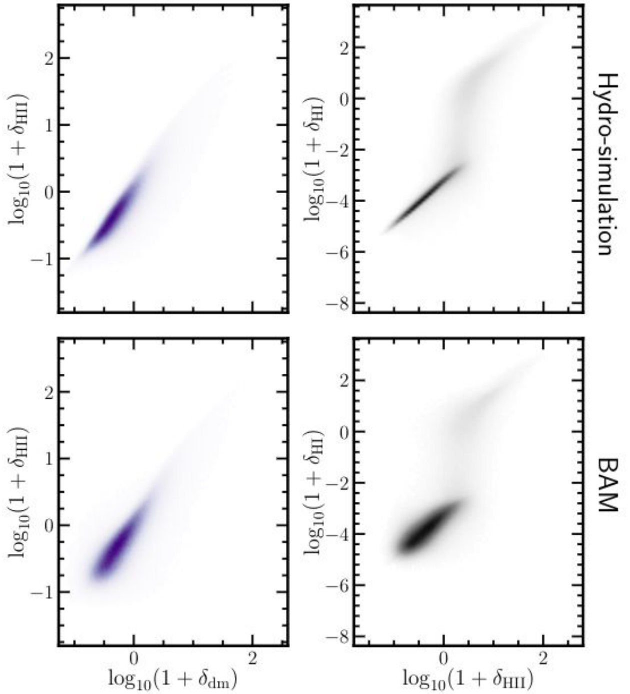
Universidad de La Laguna

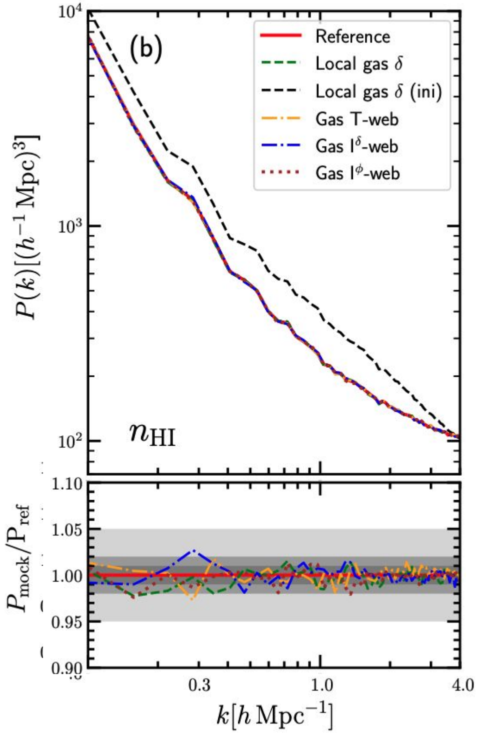
# PDF, gas phases and power spectrum




Sinigaglia et al. (2021), ApJ

**(arXiv:2012.06795)**




Local δ: local density only

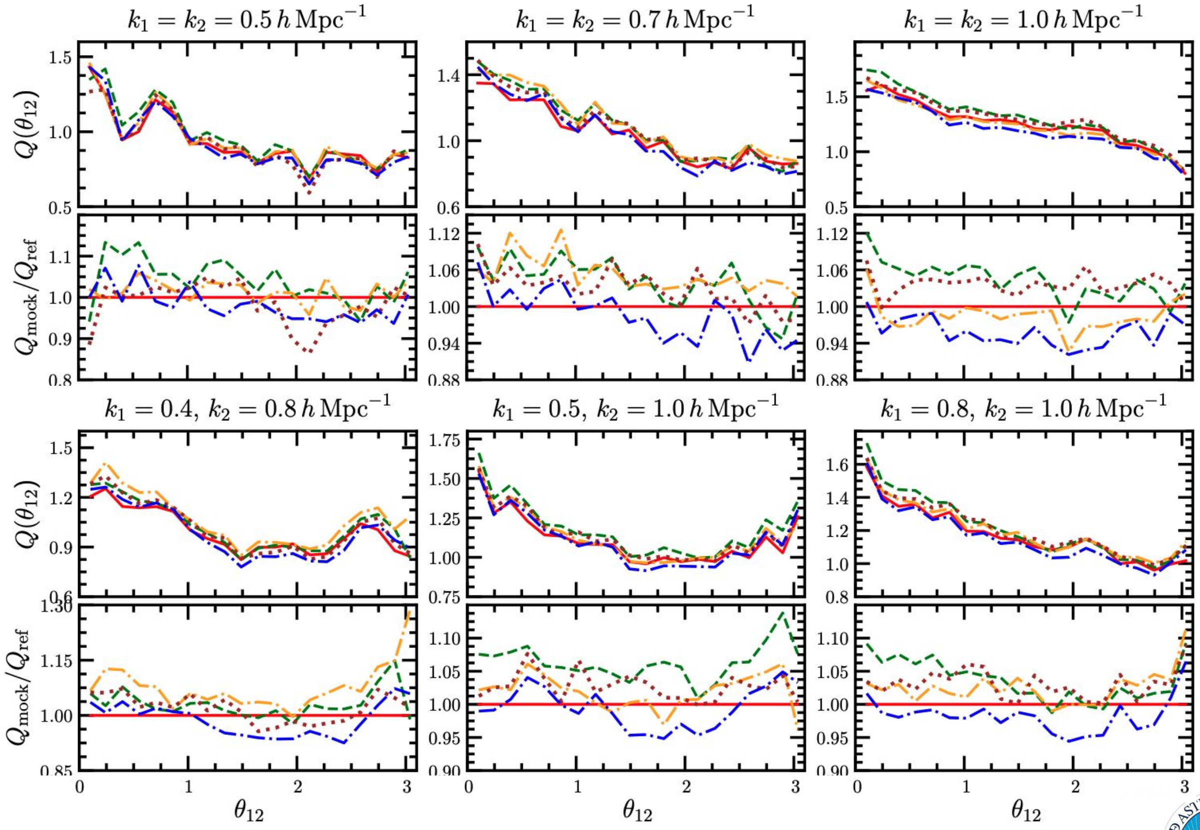
T-web: knots, filaments, sheets, voids

Iδ-web: local curvature tensor $\delta_{ij}=\partial_i\partial_j\delta$

Iφ-web: tidal field tensor $T_{ij}=\partial_i\partial_j\phi$

*Painting HI onto the DM field for mock catalogs generation*

# Bispectrum




Sinigaglia et al. (2021), ApJ **(arXiv:2012.06795)**


Local δ: local density only
T-web: knots, filaments, sheets, voids

Iδ-web: local curvature tensor $\delta_{ij}=\partial_i\partial_j\delta$
Iφ-web: tidal field tensor $T_{ij}=\partial_i\partial_j\varphi$

Reduced bispectrum

$$Q(k_1, k_2, k_3) = \frac{B(k_1, k_2, k_3)}{P(k_1)P(k_2) + P(k_1)P(k_3) + P(k_2)P(k_3)}$$

*Painting HI onto the DM field for mock catalogs generation*

# Hierarchical bias in deep learning

Attempt to transfer the Hydro-BAM hierarchical framework into DL

Conditional Variational Autoencoder (U-net)

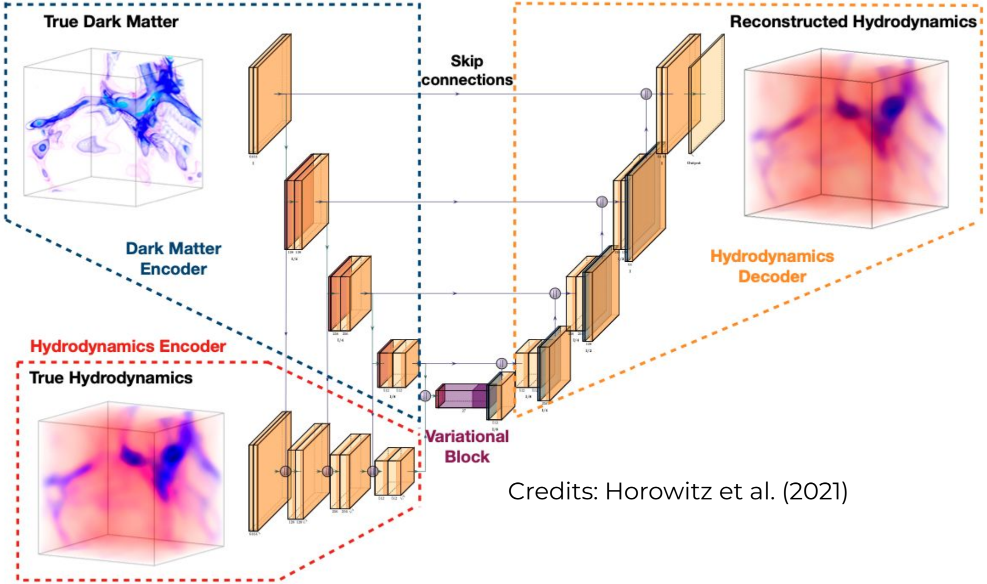
↓

Reconstruct the tracer field giving the
DM density + non-local quantities as condition

↓

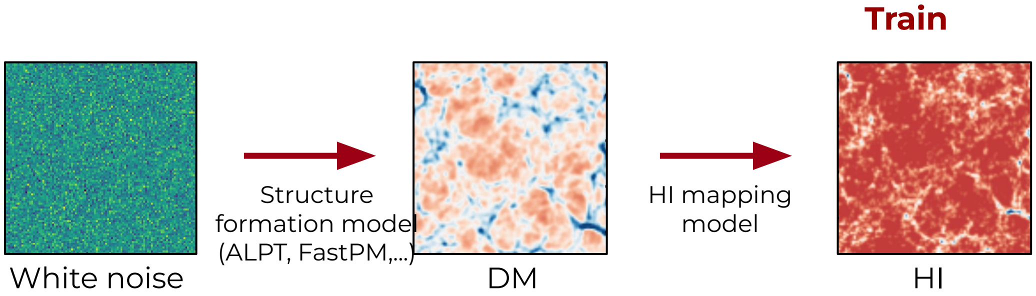
Try to treat DL less as a black box

Sinigaglia et al. (in prep.)

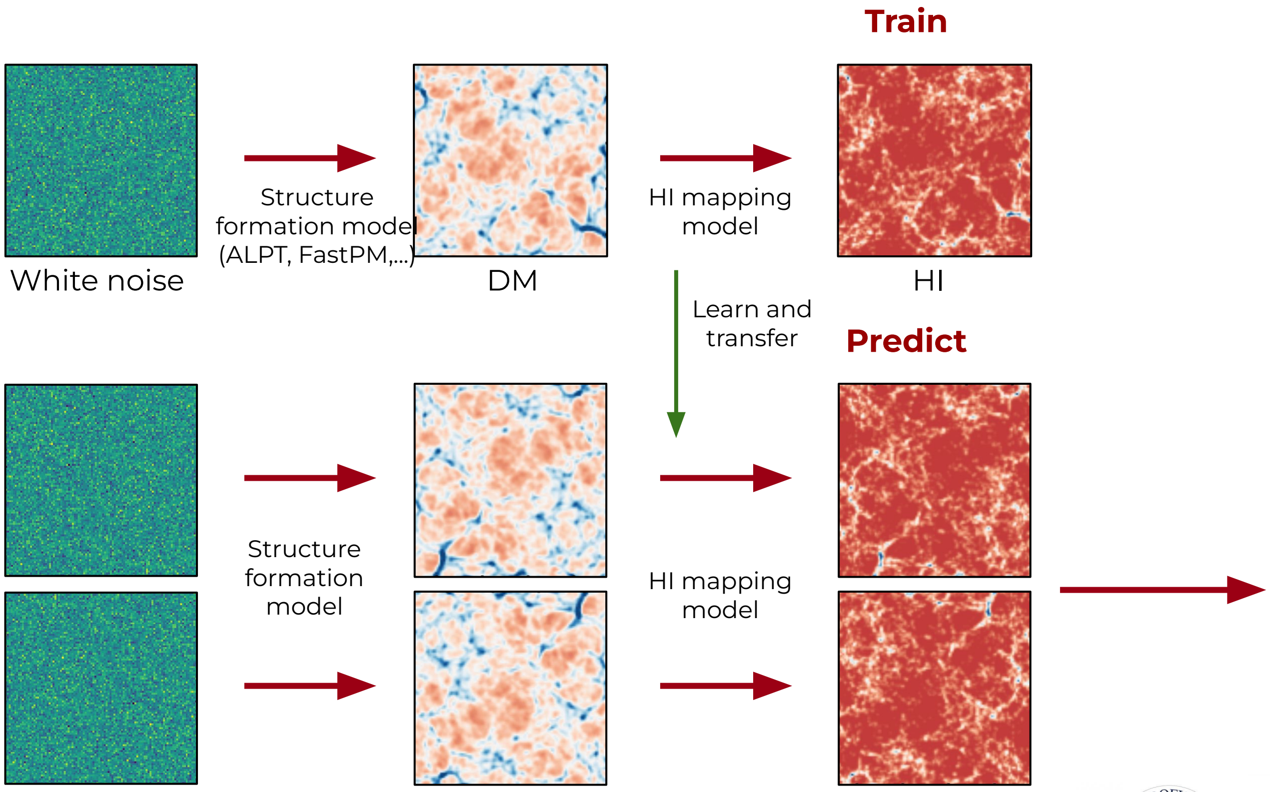



Credits: Horowitz et al. (2021)

# The full pipeline



# The full pipeline




- Repeat at different $z$
- Produce lightcone
- Add instrumental effects

*Painting HI onto the DM field for mock catalogs generation*

# Applications

Use our framework to:

- Study bias relations and scaling relations in general

# Applications

Use our framework to:

- Study bias relations and scaling relations in general

- mocks and covariance matrices: galaxies, HI IM, Lyman-alpha forests, weak lensing

  **Important for multi-tracers cross-correlation!!**

  galaxies at $z<2$, Lyman-alpha forest at $z>2$, …

# Applications

Use our framework to:

- Study bias relations and scaling relations in general

- mocks and covariance matrices: galaxies, HI IM, Lyman-alpha forests, weak lensing

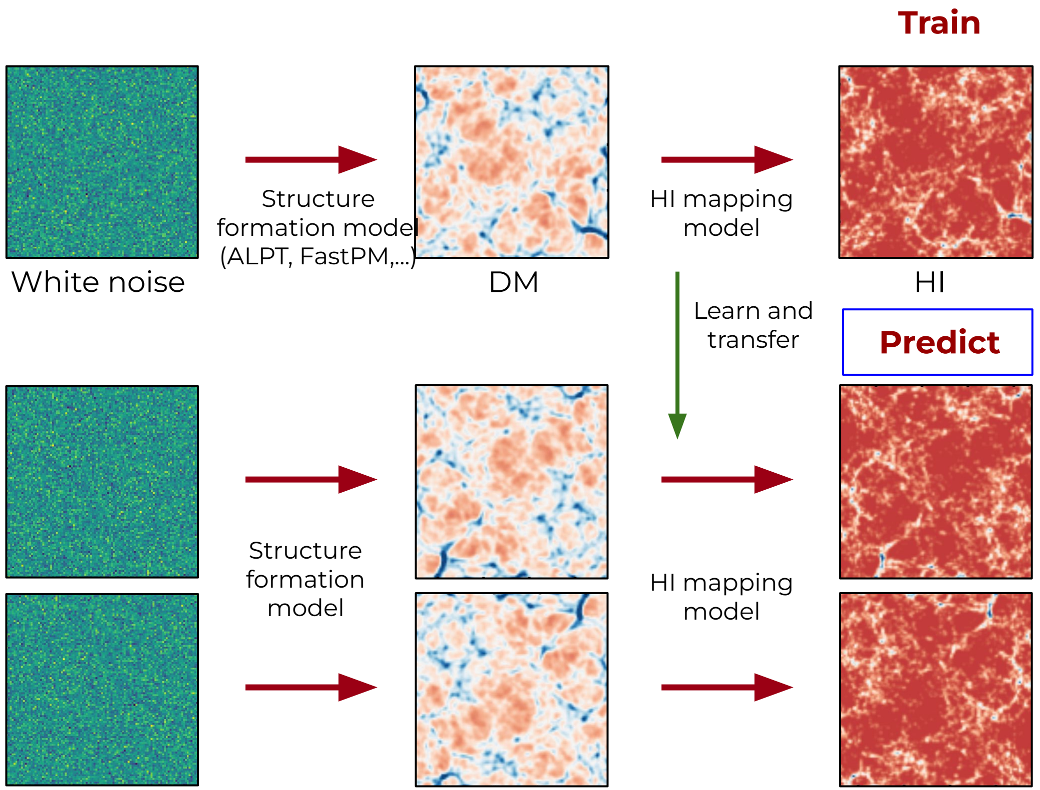
  **Important for multi-tracers cross-correlation!!**

  galaxies at $z<2$, Lyman-alpha forest at $z>2$, ...

- reconstruction of initial density field + constrained simulations

  (Ata *et al.* (2022), *Nature Astro*, reconstruction of COSMOS at cosmic noon ($1.5<z<3.5$) with galaxies)

# Future work




- repeat for different baryon models and cosmologies
- larger volumes $(500\ h^{-1}\ \mathrm{Mpc})^3$
- include relativistic effects

- Repeat at different *z*
- Produce lightcone
- Add instrumental effects

*Painting HI onto the DM field for mock catalogs generation*

# Summary

- Model to paint HI onto DM fields in place
- Stochastic, non-local, non-linear bias formulation
- Use the model to generate mock catalogs
- Mocks of different cosmological tracers in the same box
- Deep learning extension
- Use the forward model for reconstruction

Sinigaglia *et al.* (2021), ApJ

**(arXiv:2012.06795)**

Sinigaglia *et al.* (2022), ApJ

**(arXiv:2107.07917)**



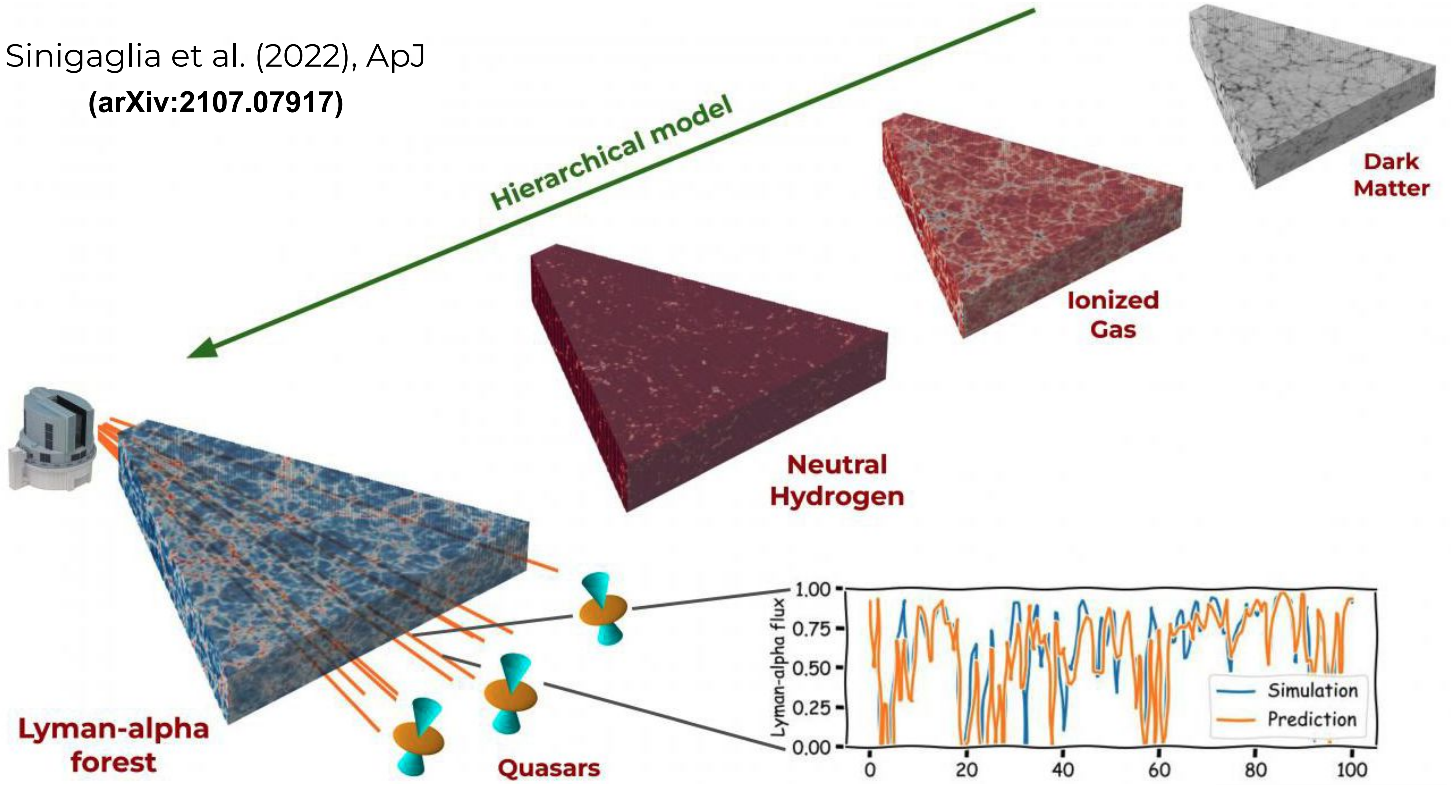
Universidad de La Laguna

OSSERVATORIO ASTRONOMICO DI PADOVA

# Backup slides

# Hierarchical bias mapping

Sinigaglia et al. (2022), ApJ
**(arXiv:2107.07917)**



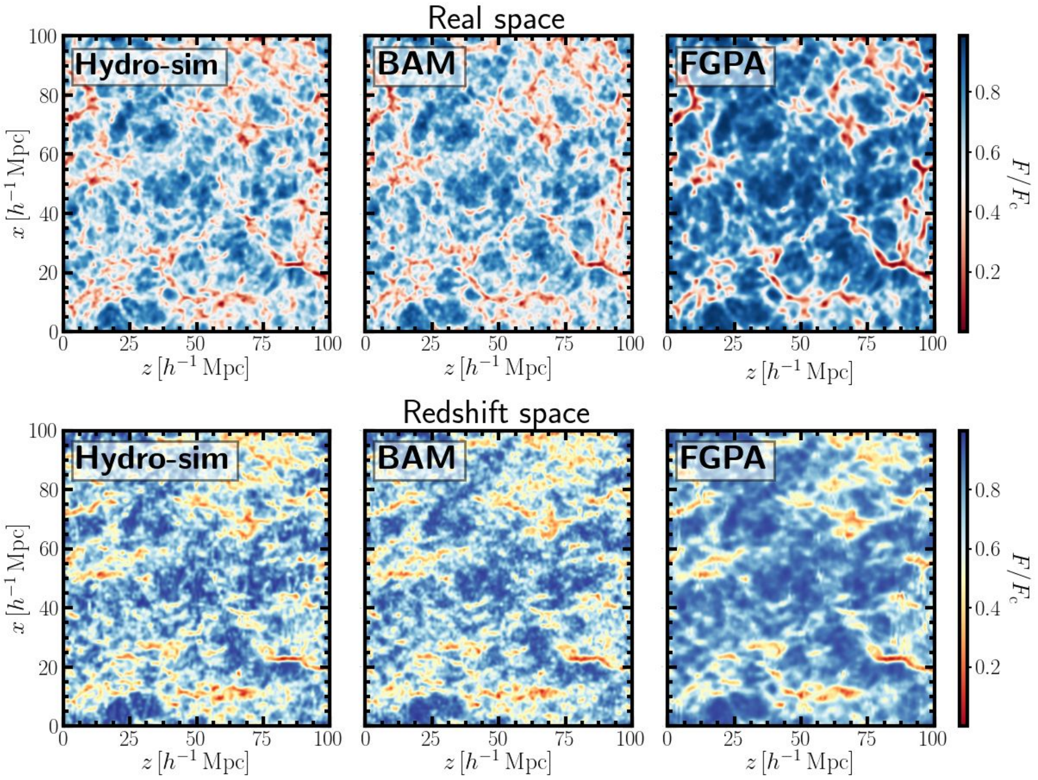

Press release:

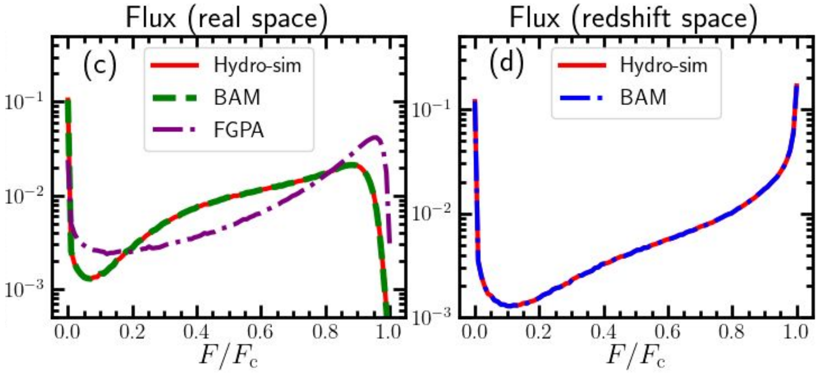
*Painting HI onto the DM field for mock catalogs generation*

Universidad de La Laguna

OSSERVATORIO ASTRONOMICO DI PADOVA

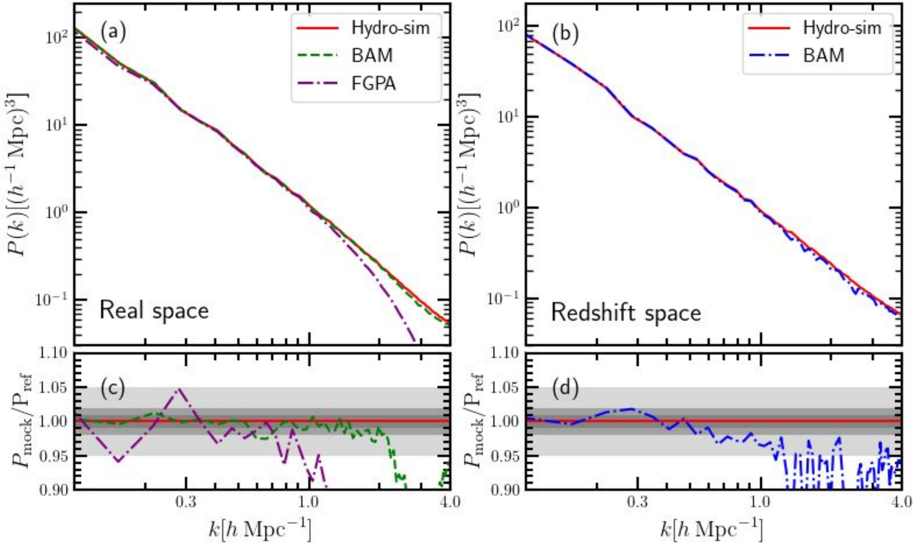
# Lyman-alpha forest




*Painting HI onto the DM field for mock catalogs generation*

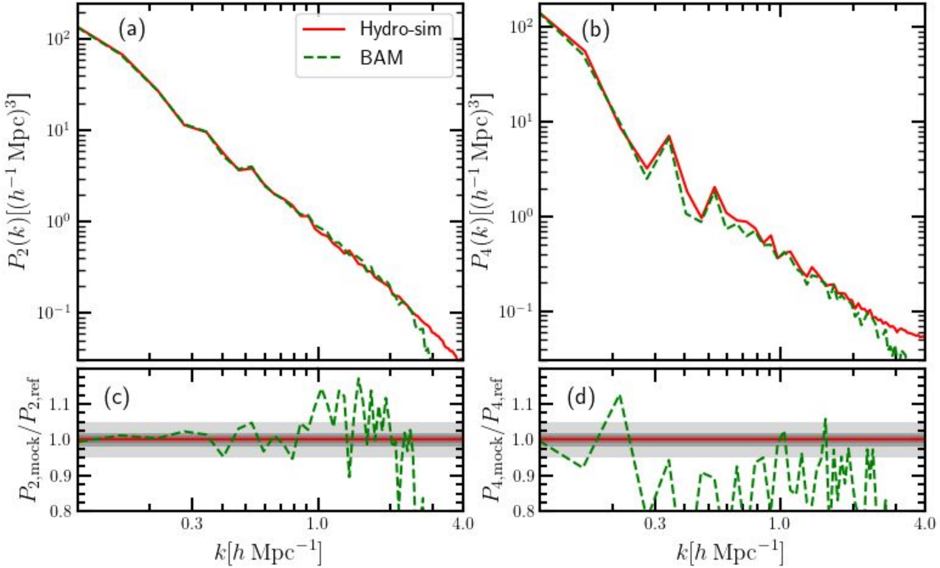
# Lyman-alpha forest



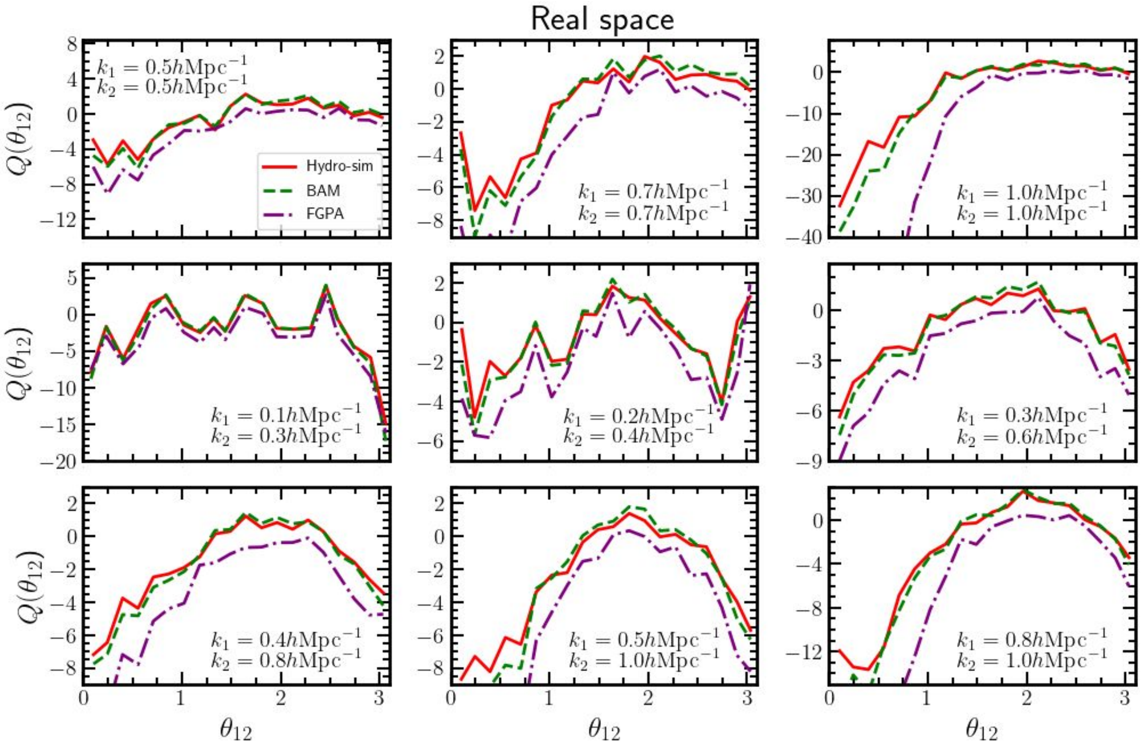

3D power spectrum

Quadrupole

Hexadecapole

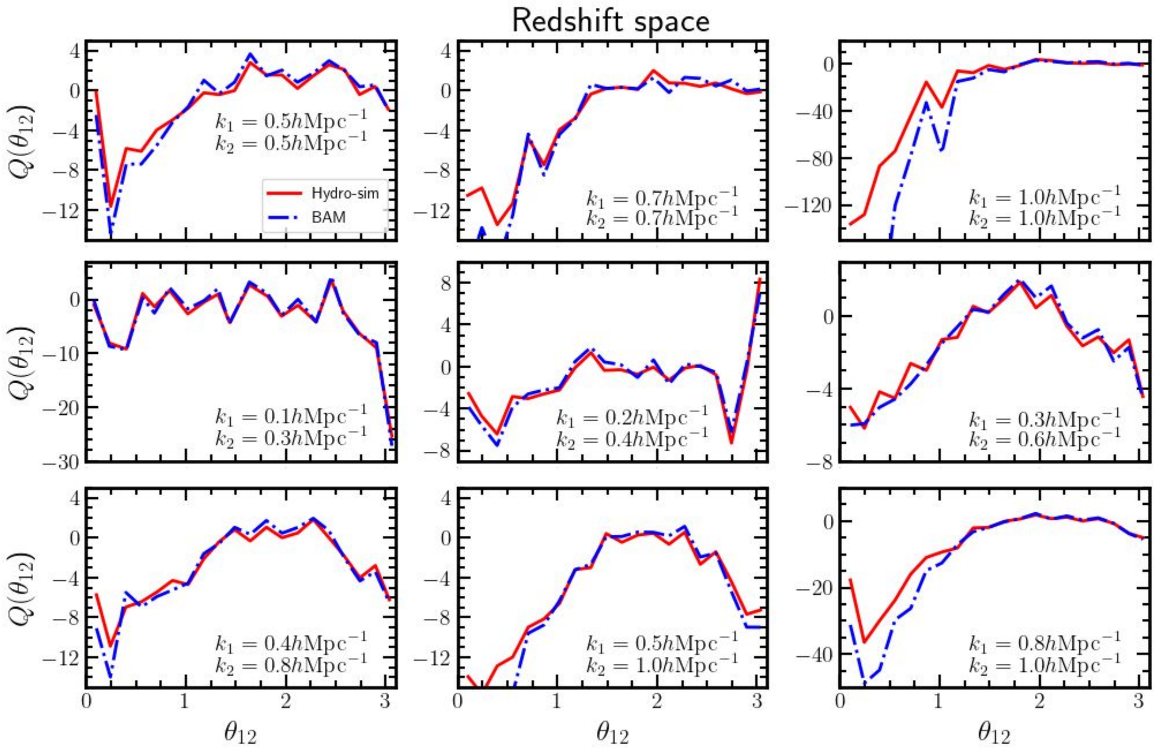


# Lyman-alpha forest




*Painting HI onto the DM field for mock catalogs generation*

# Lyman-alpha forest



# RSD modelling

We want to calibrate the Lyα forest in redshift space: $F(\boldsymbol{s}) \curvearrowleft \mathcal{B} = P(F(\boldsymbol{s}) | \Theta\{\delta_{\rm HII}(\boldsymbol{s}) \otimes \mathcal{K}, \delta_{\rm HI}(\boldsymbol{s})\})$

We need to map $\delta_{\rm HII}$ and $\delta_{\rm HI}$ from real to redshift space on the mesh

- Consider a cell *i* and assign *N* fictitious pseudo-particles with position $\boldsymbol{r}_j$ coincident to the center of the cell
- Displace pseudo particles from real to redshift space following (Kaiser 1987, Hamilton 1998):

  $\mathcal{S}$: $\boldsymbol{s}_j = \boldsymbol{r}_j + [b_{\rm v}(\boldsymbol{v}_{{\rm dm},j} \cdot \boldsymbol{r}_j)\, \boldsymbol{r'}_j] / (aH)$ , $\boldsymbol{r'}_j = \boldsymbol{r}_i / |\boldsymbol{r}_i|$ , $\boldsymbol{v}_{{\rm dm},j} = \boldsymbol{v}_{{\rm dm},j}^{\rm coh} + \boldsymbol{v}_{{\rm dm},j}^{\rm disp}$ ,

  $\boldsymbol{v}_{{\rm dm},j}^{\rm coh} = \boldsymbol{v}_{{\rm dm},i}^{\rm sim}$ = coherent flows, $\boldsymbol{v}_{{\rm dm},j}^{\rm disp} \curvearrowleft \mathcal{N}[0, A(1 + \delta_i)^{\alpha}]$ = quasi-virialized motions , A, α and $b_{\rm v}$ free parameters
- Re-interpolate pseudo-particles on the mesh at coordinates $\boldsymbol{s}_j$ using CIC