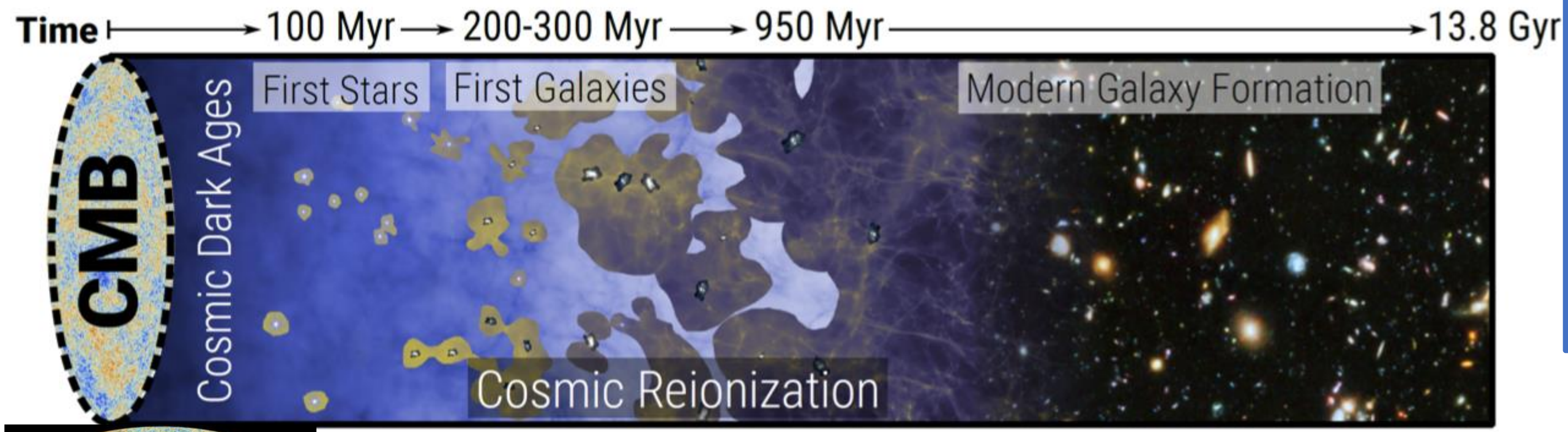
Reconstructing the dark-matter density field and 21-cm intensity field using AI

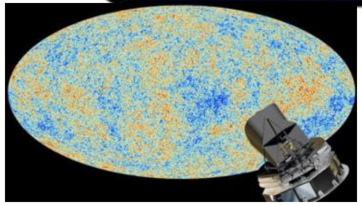
Feng Shi （史峰）

21cm cosmology workshop 2024 @ Hangzhou Dianzi University
July 23, 2024

西安电子科技大学
XIDIAN UNIVERSITY
空间科学与技术学院
School of Aerospace Science And Technology

Background





Filament spin

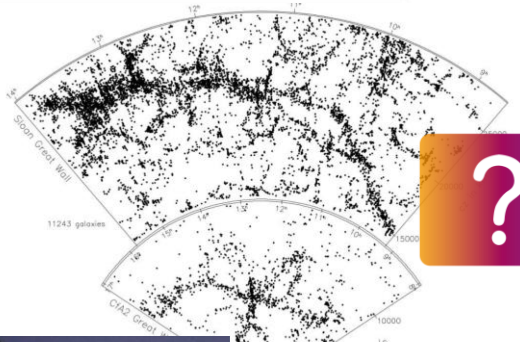
Filament lensing

Supernova distance correction

Constrained simulation

...

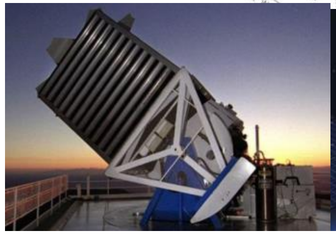
21cm intensity mapping



dark matter density field

Velocity & Tidal field

Initial density field



Fitting the field

Challenge 1: The function form is unknown

Output field

input field

$$\delta_{out} = f(\delta_{in};\theta)$$

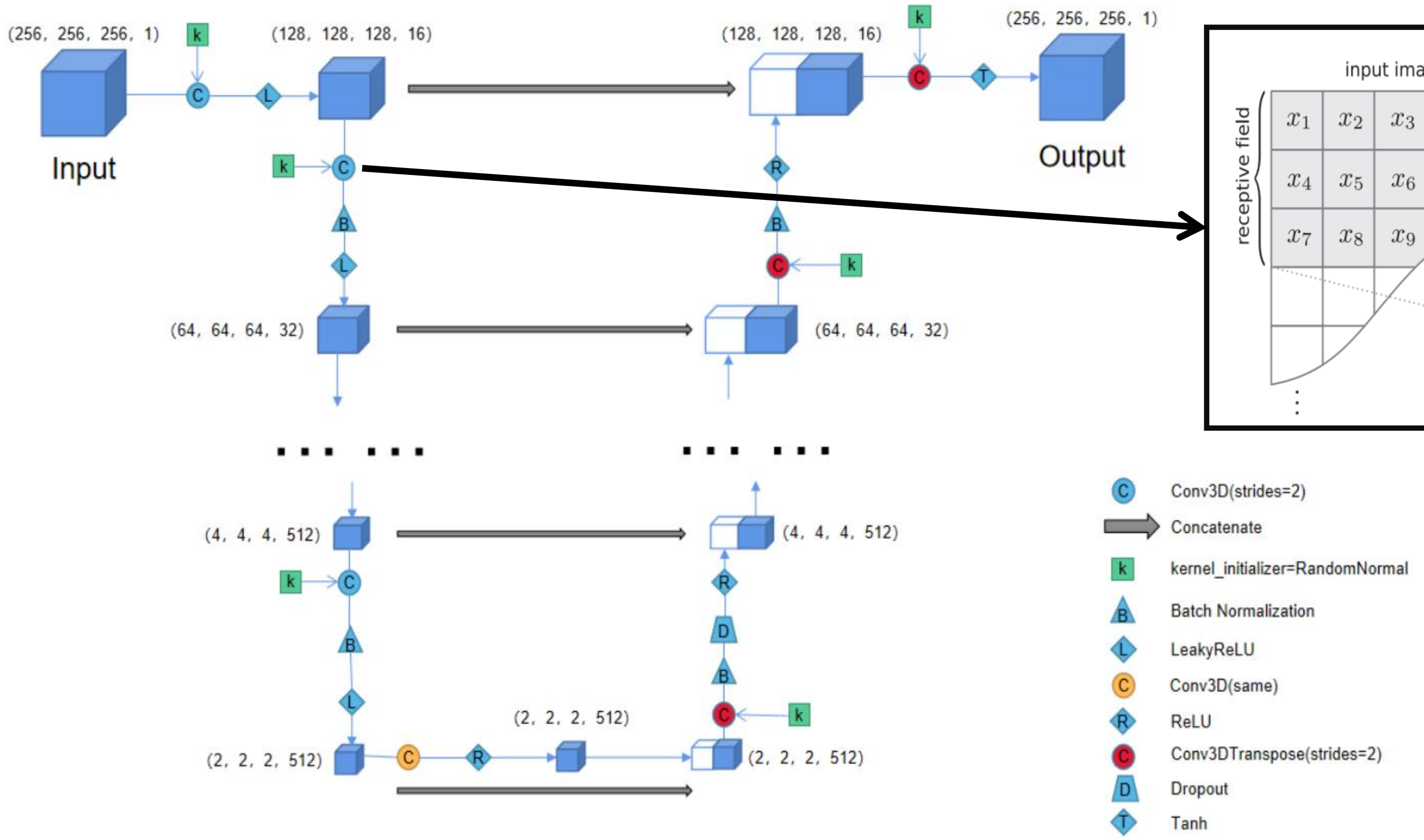
Challenge 2: The fitted data size is huge, such as $N_p=128^3$, 256^3 , 512^3

Likelihood: $$\mathcal{L} = \prod_{i=1}^{N_p} P(\delta_{out}|\delta_{in};\theta)$$

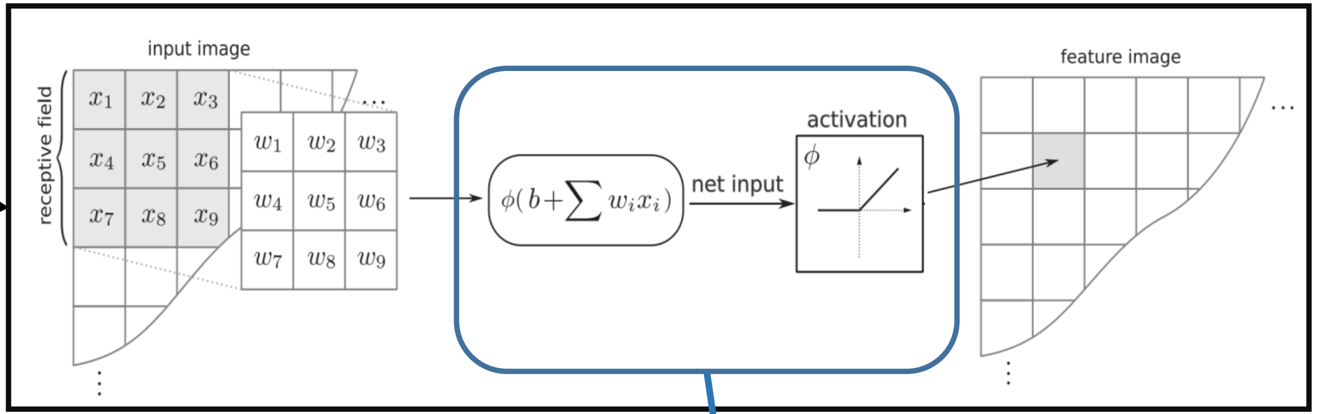
$$\hat{\theta}_{\mathrm{ML}} = arg\ \max_{\theta} \mathcal{L} = arg\ \max_{\theta} \sum_{i} \log P(\delta_{out}|\delta_{in};\theta)$$

Deep learning method: UNet model

U-Net：
Encoder-decoder with skip connections



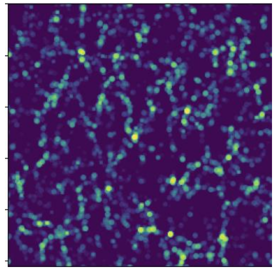
Convolutional neural network (CNN) :

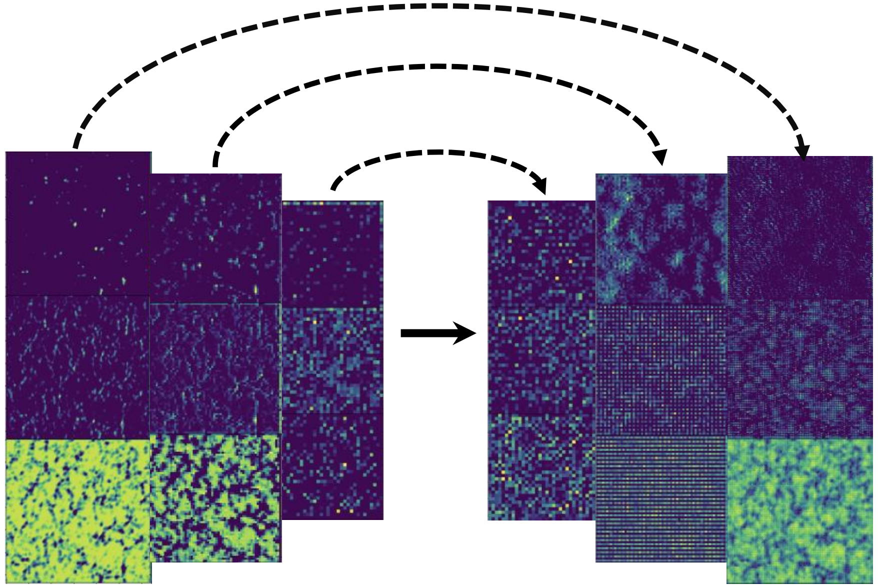


$$\delta_{out} = f(\delta_{in};\theta)$$

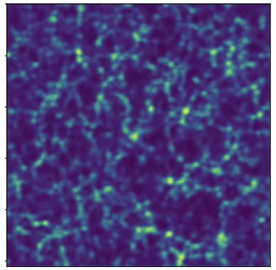
Reconstructing the dark matter density field

Redshift-space halo denstiy field





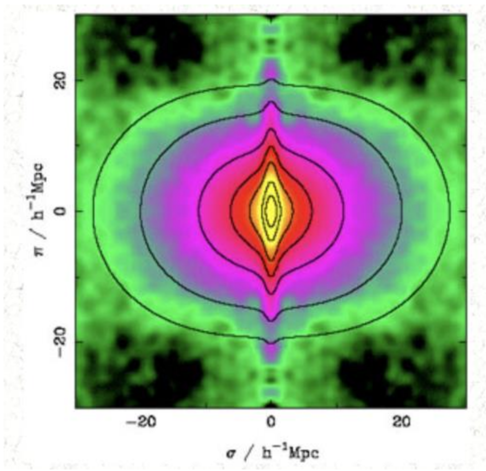
Real-space dark matter field



1) The bias effect

$$\delta_h(\boldsymbol{x}) = b_1\delta(\boldsymbol{x}) + \frac{1}{2}b_2[\delta(\boldsymbol{x})^2 - \sigma_2] + \frac{1}{2}b_{s2}[s(\boldsymbol{x})^2 - \langle s^2\rangle]$$
$$+ \text{higher order terms}\,.$$

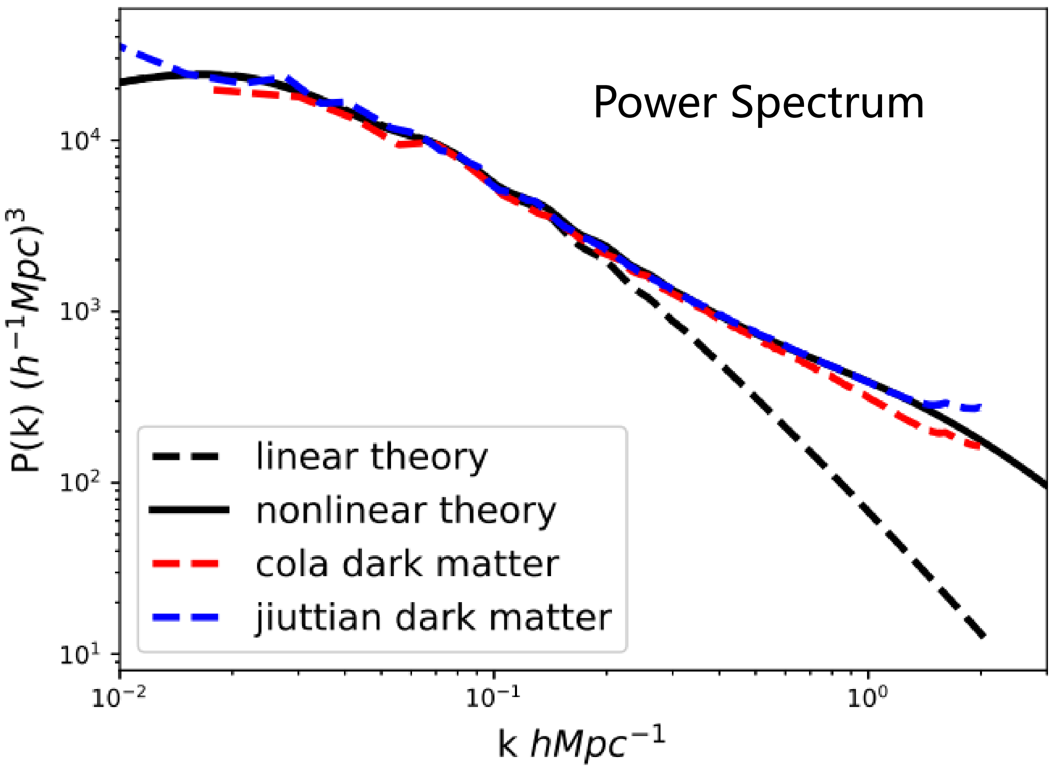
2) Redshift distortions

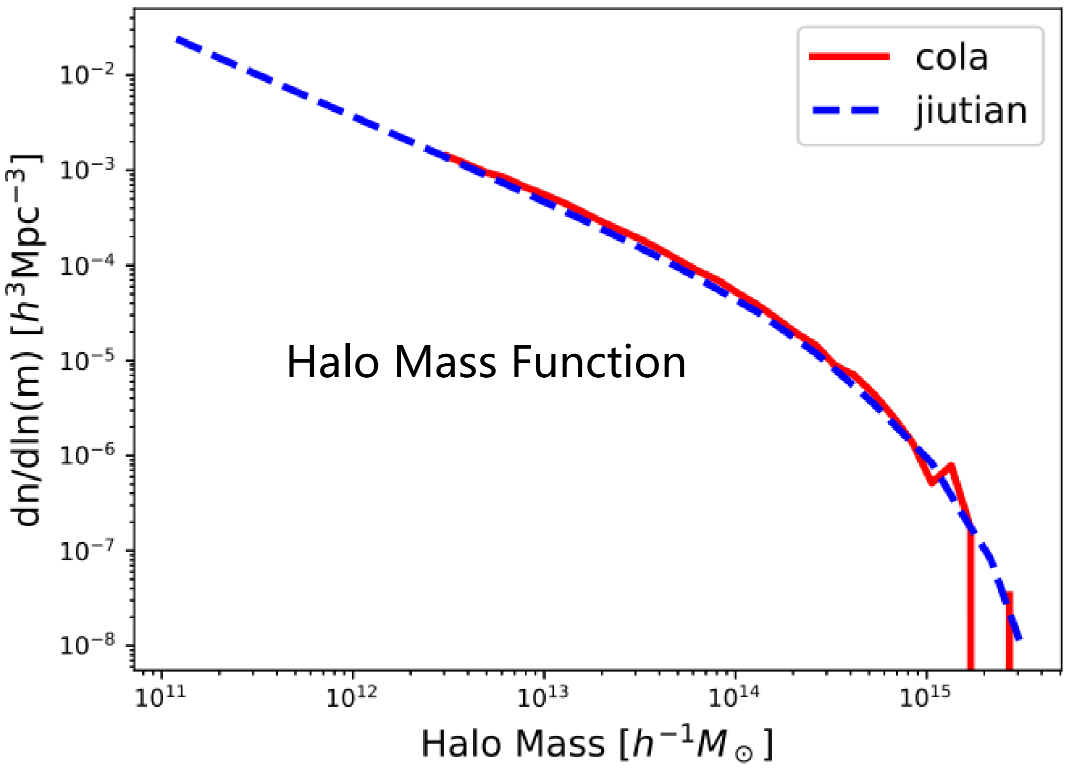


Training strategy

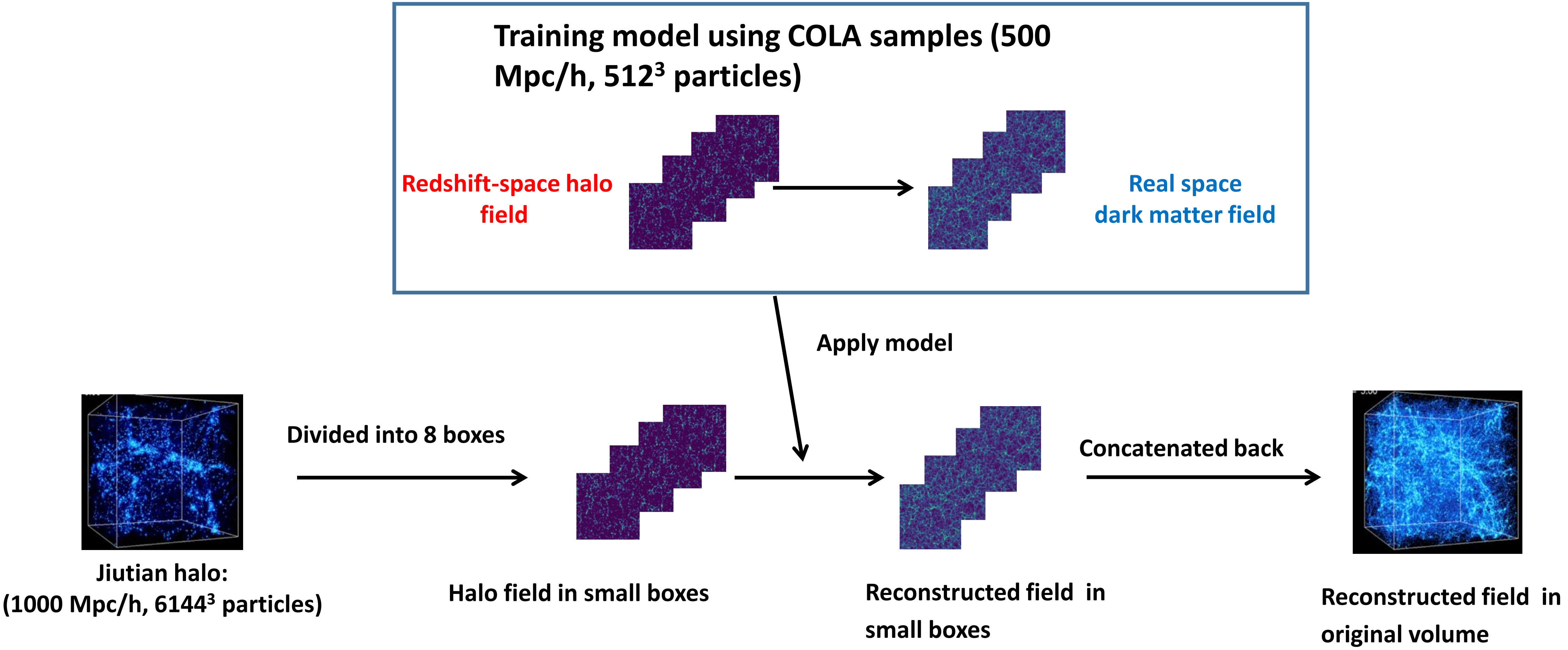
Training data：COLA, Fast simulation
Box size：500Mpc/h
Particle number：512^3
Runing time：28 CPUs，0.5 hours
Memory：<3GB
Storage：10GB

Testing data：Jiutian, Nbody simulation
Box size：1000Mpc/h
Particle number：6144^3
Running time：10^4 CPUs，28 days
Memory：22TB+
Storage：900TB+



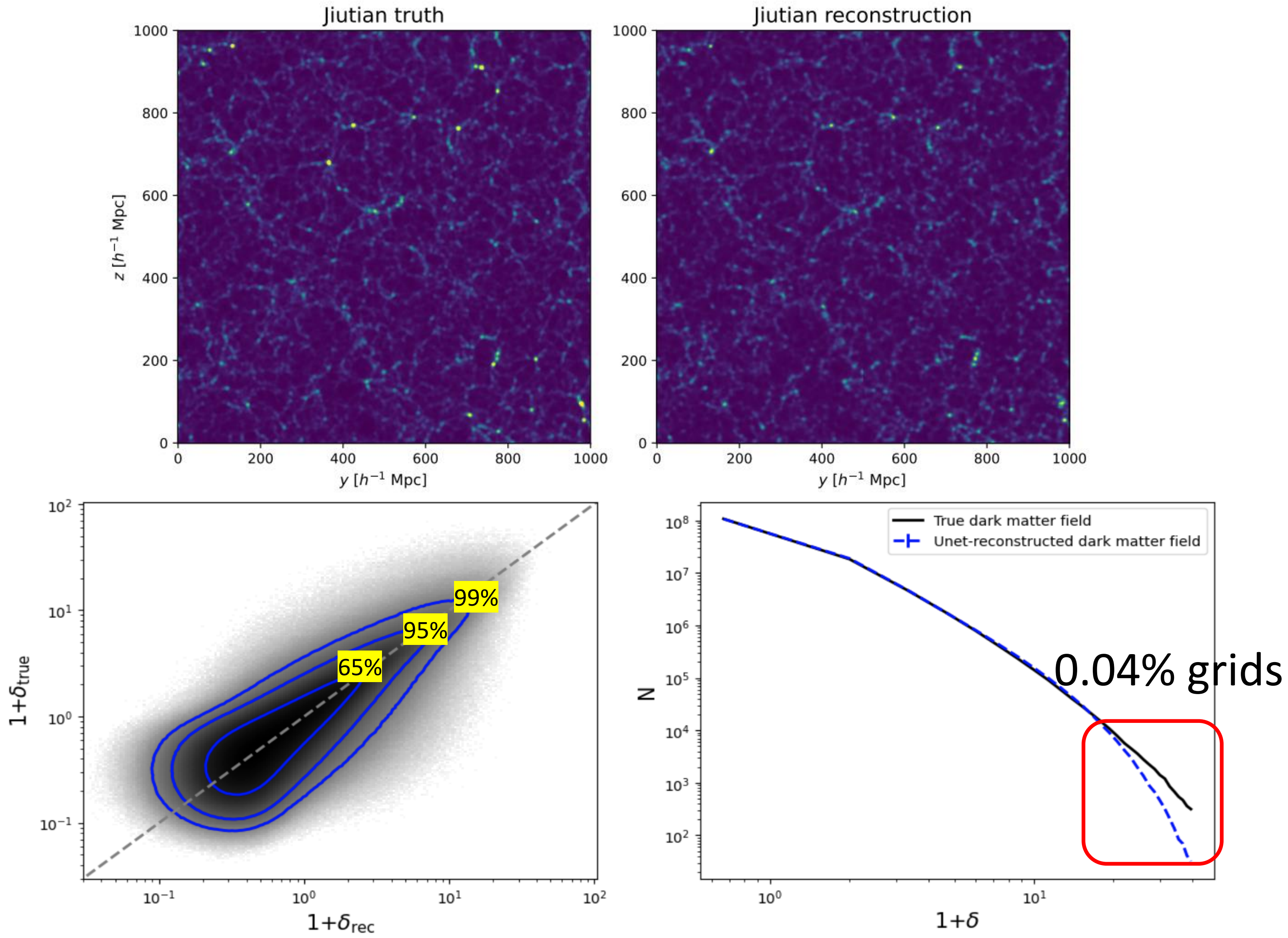


Training and testing

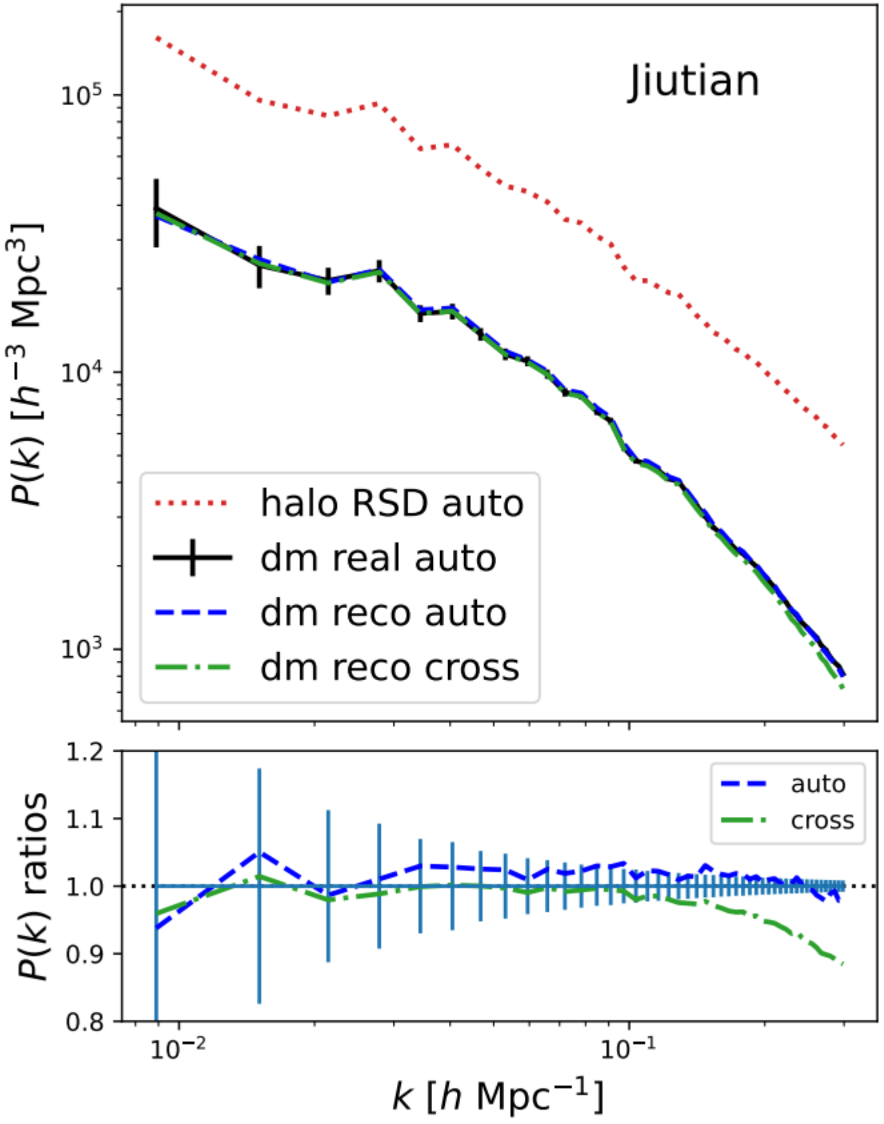


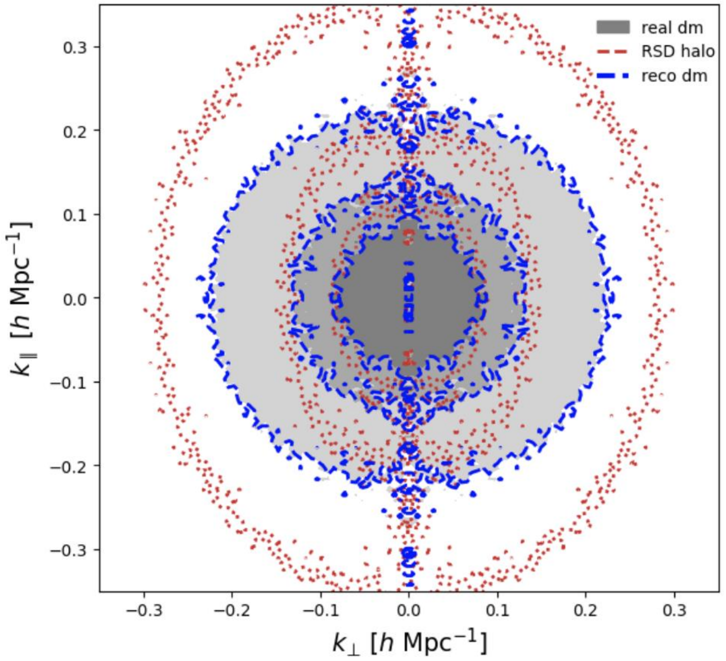
Results: the reconstructed density field for Jiutian

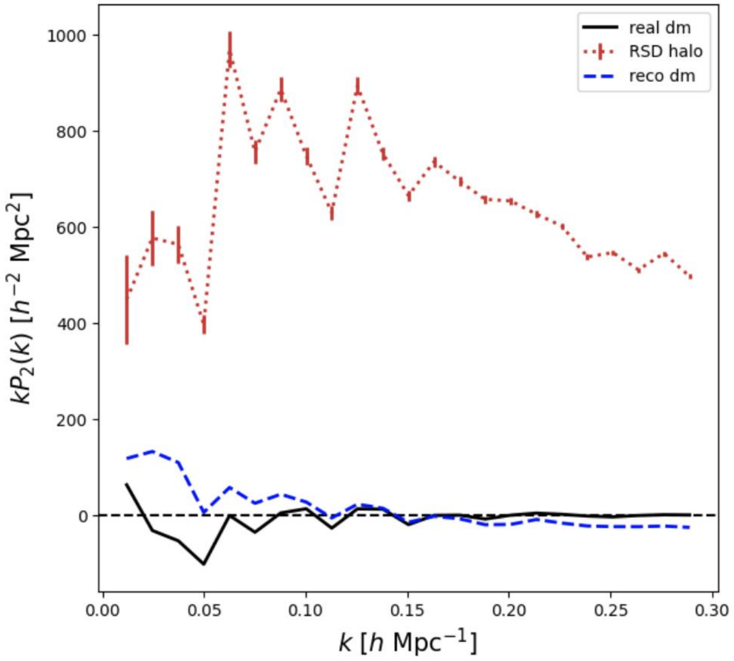
- Reconstructing the dark matter density field based on UNet



Results: the reconstructed density field for Jiutian







The real-space P(k) is also recovered accurately, with only small reduction of the cross-correlation power spectrum at 1% and 10% levels at k = 0.1 and 0.3 h Mpc^{-1}, respectively.

Testing the impact of cosmology

Training in Planck2018 cosmology

- COLA simulations :

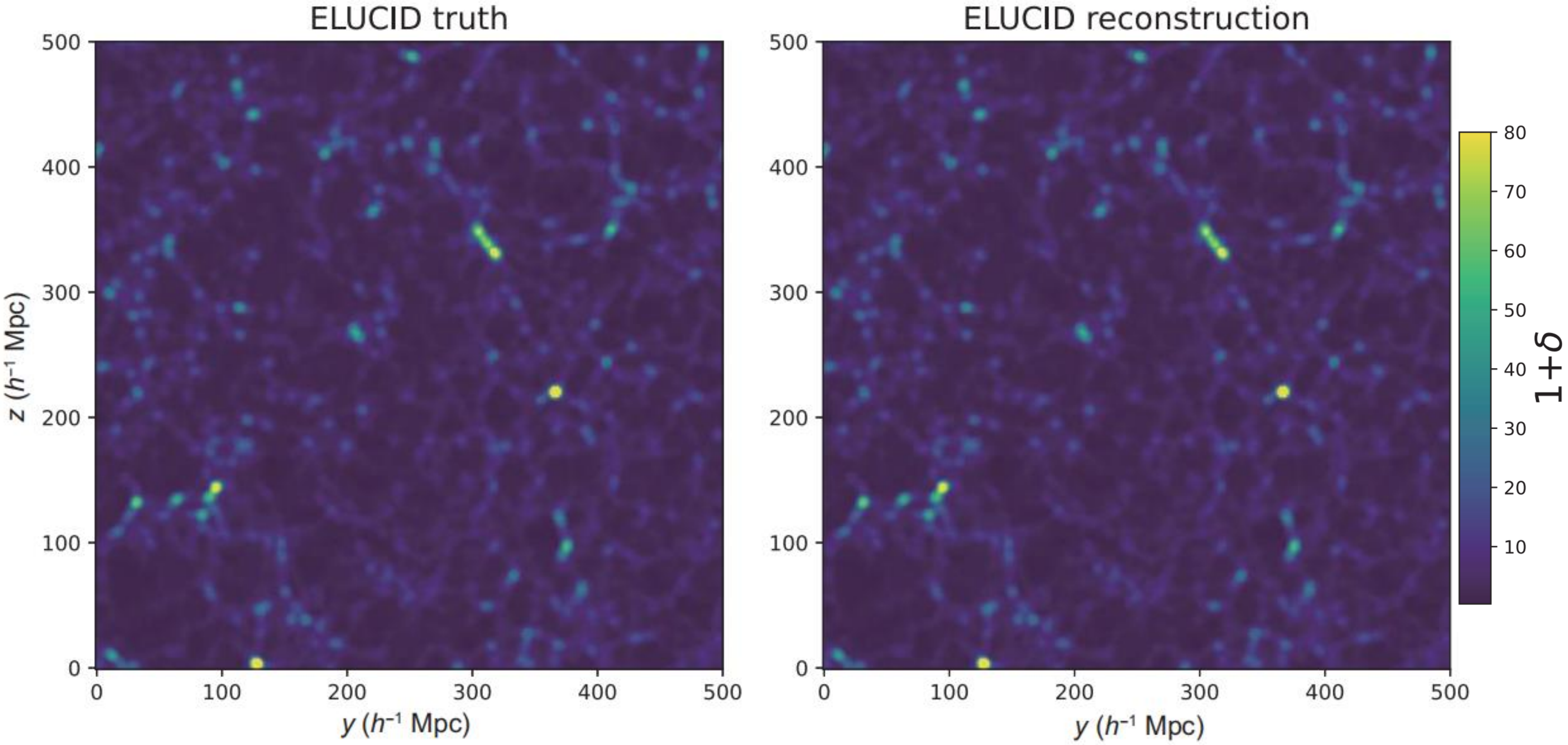
$\Omega_m = 0.3111$, $\Omega_\Lambda = 0.6889$, $h = 0.6766$, $\Omega_b = 0.049$, $\sigma_8 = 0.817$

Applying to WMAP5 cosmology

- ELUCID Nbody simulation (500Mpc/h, 3072^3 particles):

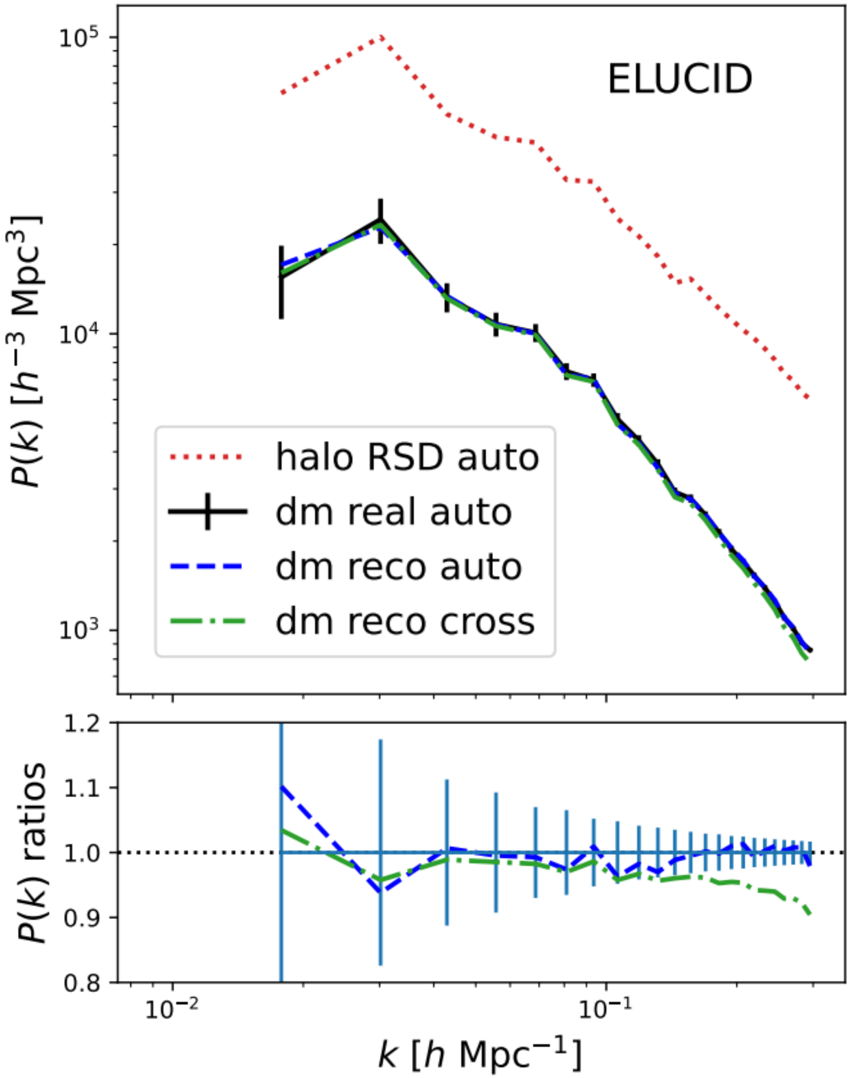
$\Omega_m = 0.258$, $\Omega_\Lambda = 0.742$, $\Omega_b = 0.044$, $h = 0.72$, $\sigma_8 = 0.80$

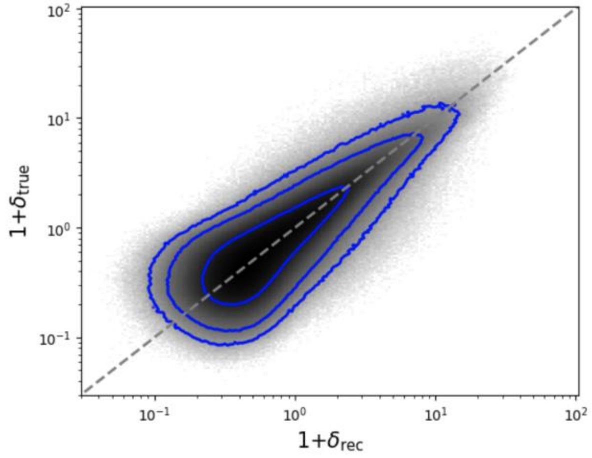
Results: the reconstructed density field for ELUCID

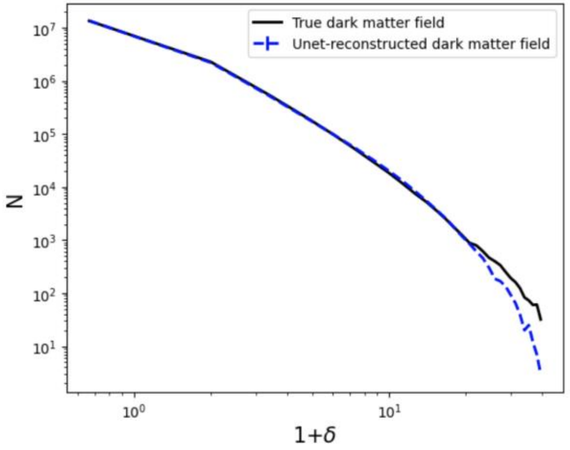


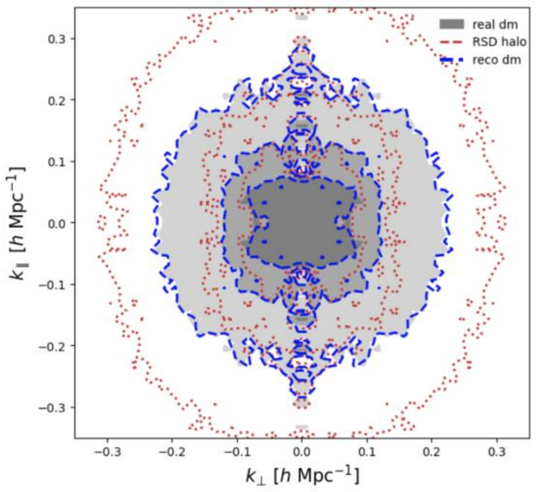
Results: the reconstructed density field for ELUCID

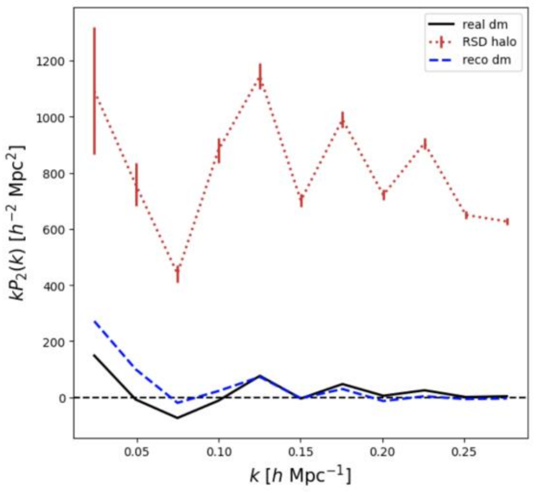
- No large distinction of the results between the WMAP5 and Planck18 cosmology











Testing the reconstruction of velocity field

- Reconstruct velocity field

UNet-reconstructed $\delta(k)$

$$\boldsymbol{v}(\boldsymbol{k}) = H\ a\ f(\Omega)\ \frac{i\boldsymbol{k}}{k^2}\ \delta(\boldsymbol{k})$$

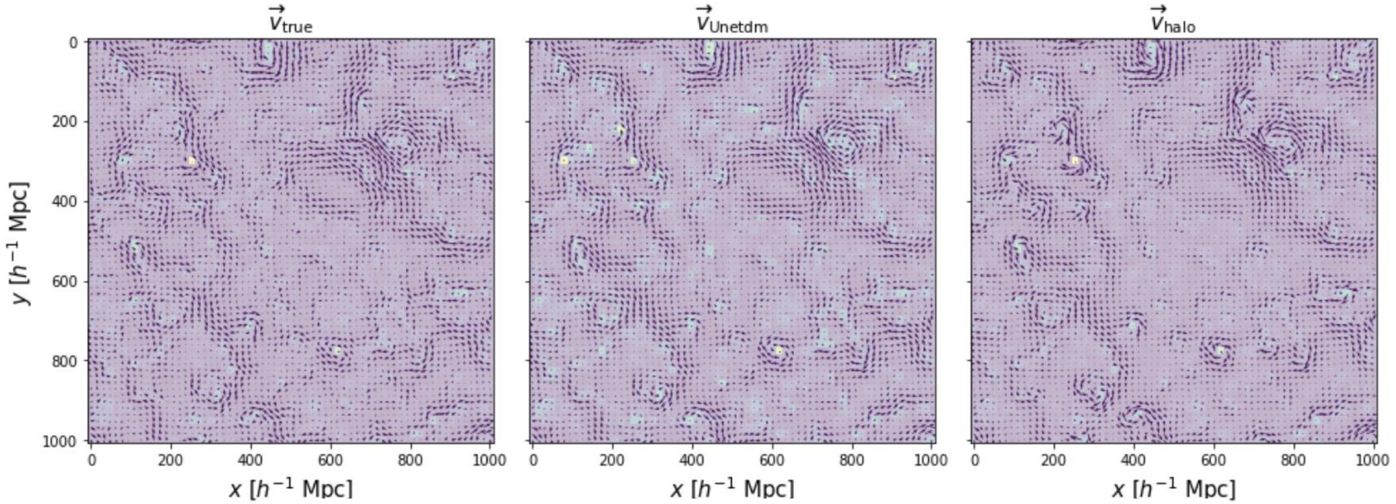
VS.

Halo density field $\delta_h(k)$ with a bias b_{hm}

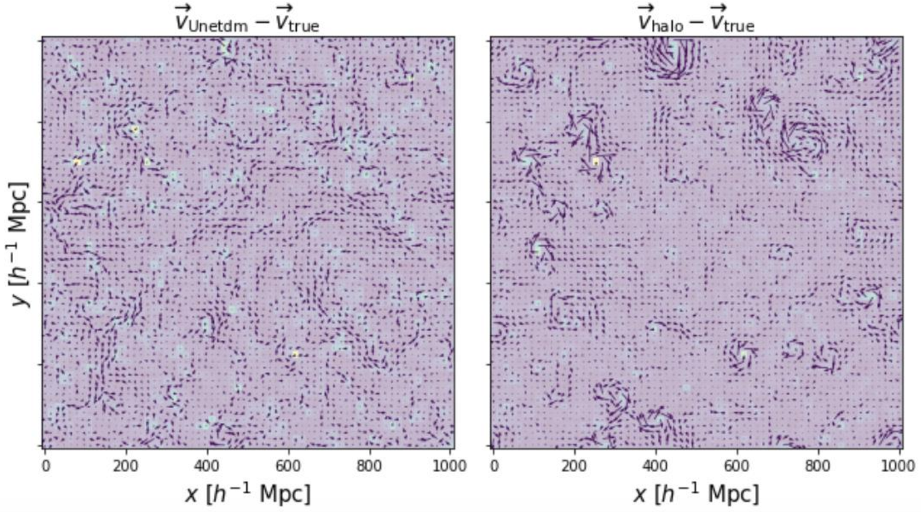
$$\boldsymbol{v}(\boldsymbol{k}) = H\ a\ f(\Omega)\ \frac{i\boldsymbol{k}}{k^2}\ \frac{\delta_{\mathrm{h}}(\boldsymbol{k})}{b_{\mathrm{hm}}}$$

(Wang et al 2012, Shi et al 2016)

Velocity field



Velocity difference field



UNet dark matter

Halo-based with linear bias

Testing the reconstruction of velocity field

- Reconstruct velocity field

UNet-reconstructed δ(k)

$$\boldsymbol{v}(\boldsymbol{k}) = H\ a\ f(\Omega)\ \frac{i\boldsymbol{k}}{k^2}\ \delta(\boldsymbol{k})$$

VS.

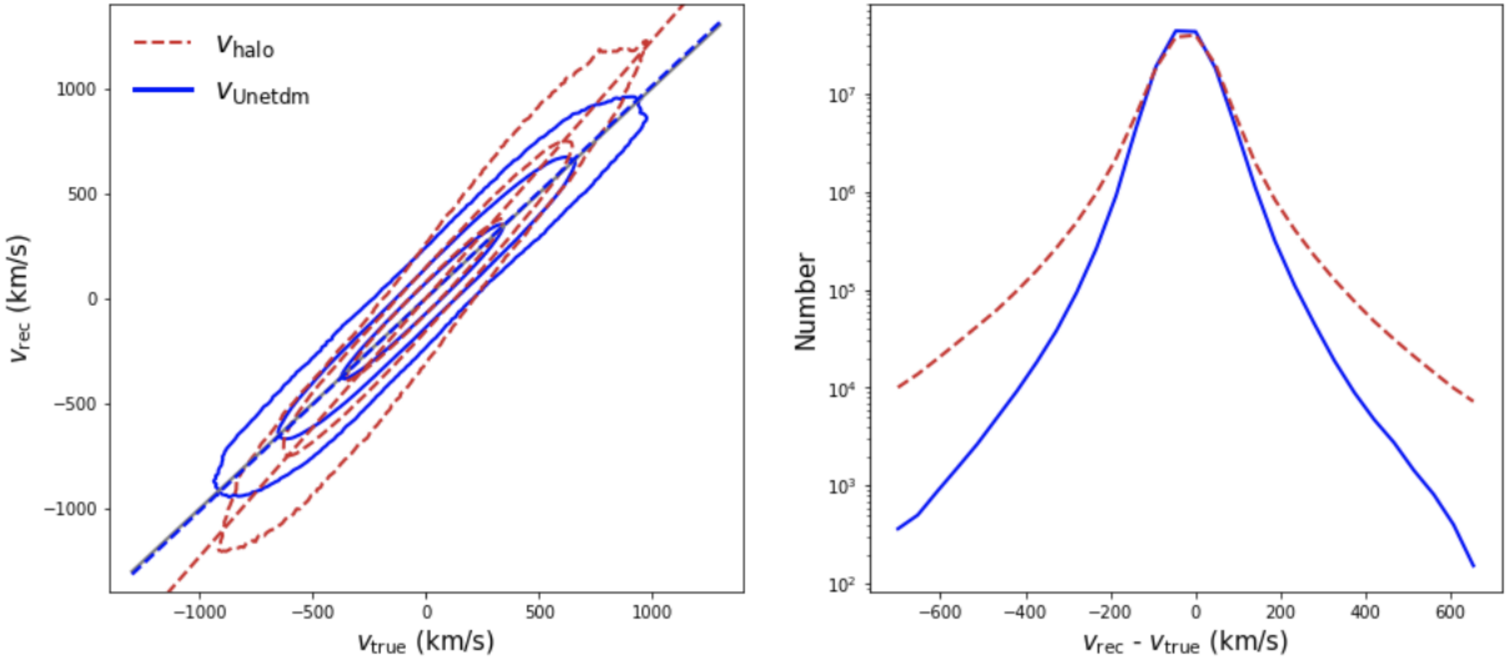
Halo density field $\delta_h(k)$ with a bias b_{hm}

$$\boldsymbol{v}(\boldsymbol{k}) = H\ a\ f(\Omega)\ \frac{i\boldsymbol{k}}{k^2}\ \frac{\delta_{\mathrm{h}}(\boldsymbol{k})}{b_{\mathrm{hm}}}$$

(Wang et al 2012, Shi et al 2016)

	Slope	Scatter
Halo :	1.15	78.2 km/s
UNet :	1.01	57.0 km/s

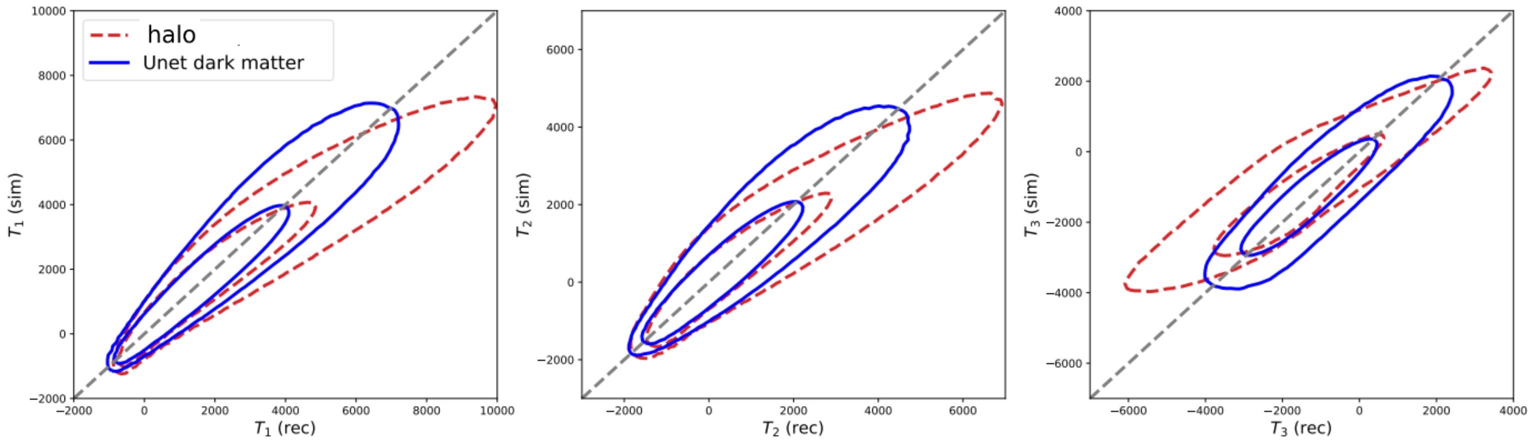
- Unbiased relation
- 21.1% scatter errror reduction



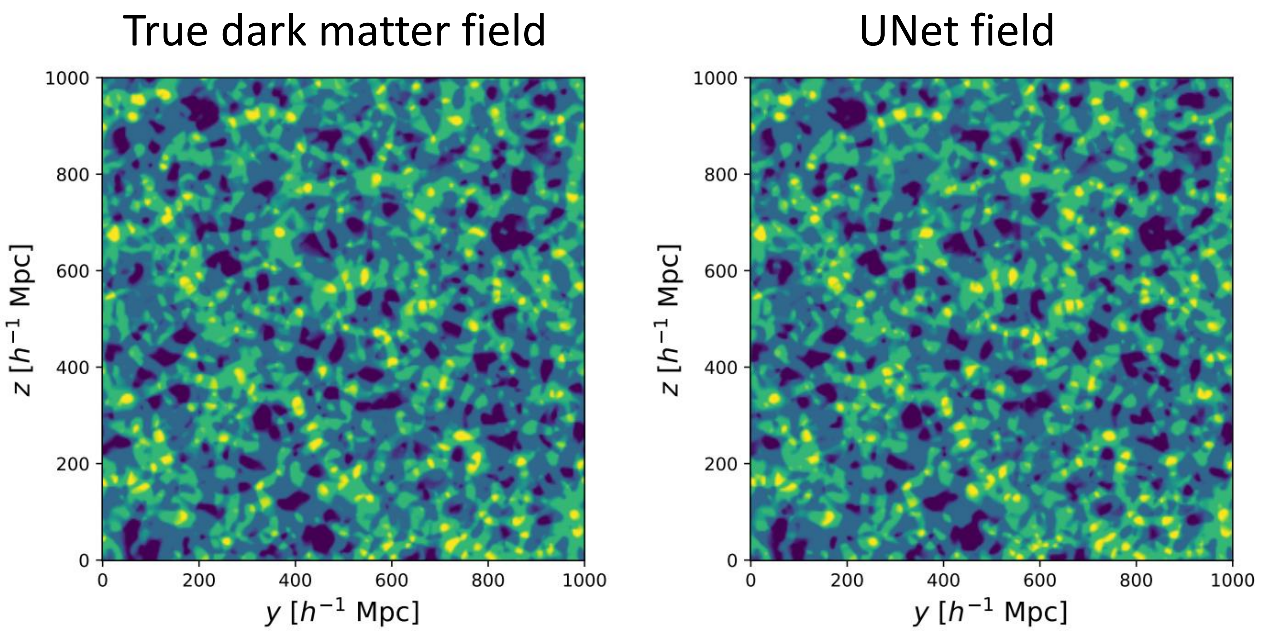
The three contours encompass 67%, 95%, and 99% of the grid cells

Testing the reconstruction of tidal field

- Reconstruct tidal field:



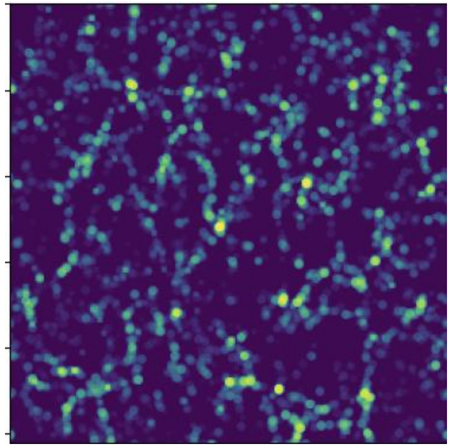
- Classification of the large-scale structure:



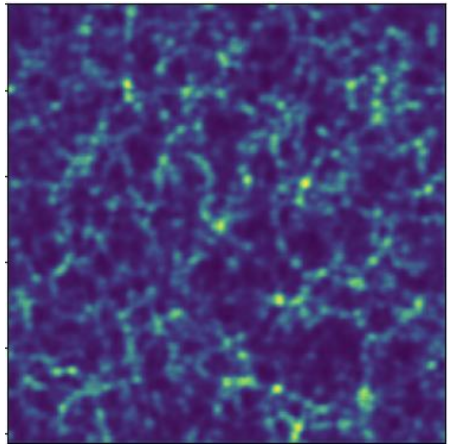
- cluster : yellow
- filament: yellow-green
- sheet: green
- void: black

Deeping learning: UNet model

Redshift-space halo field

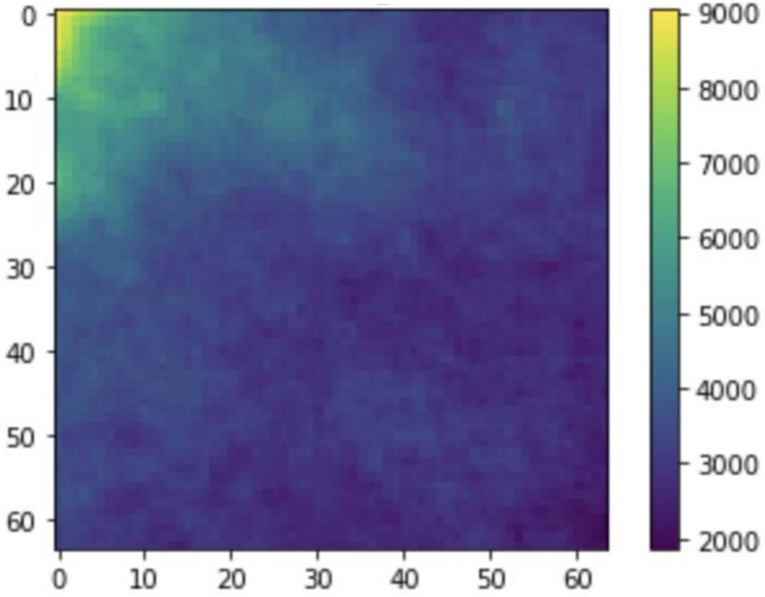


Real-space dark matter field



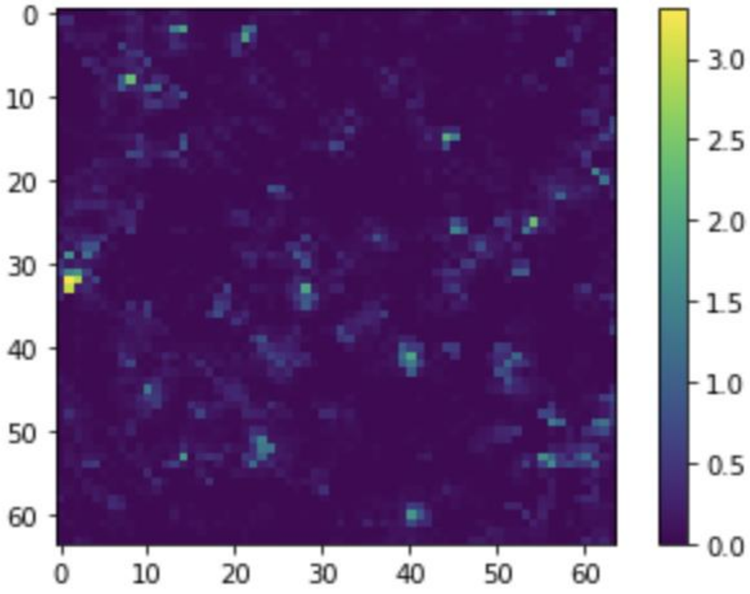
UNet works

21cm + foreground



UNet for removing foregrounds?

21cm



Testing UNet for cleaning different-level foreground

- We run the UNet by inputting $T_{HI} + T_{fg}$, where $T_{fg}=\alpha T_{fg}^{CRIME}$

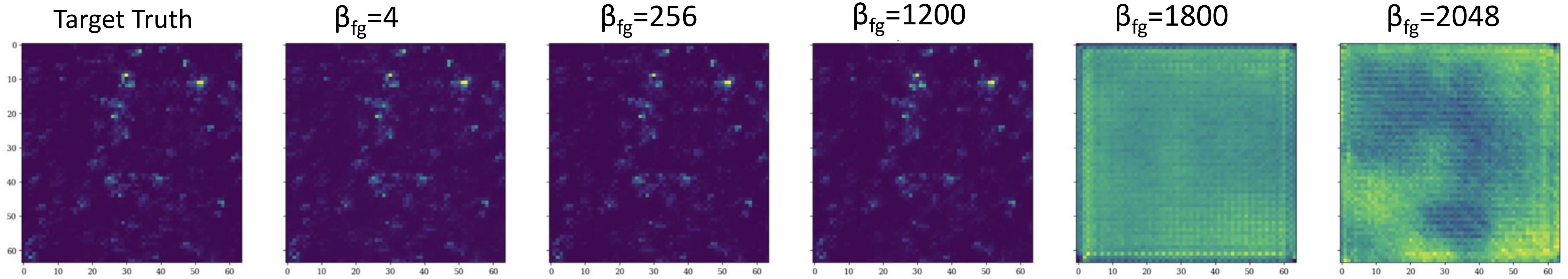
$$\beta_{fg} = \sqrt{\frac{\langle (T_{fg} - \overline{T}_{fg})^2 \rangle}{\langle (T_{HI} - \overline{T}_{HI})^2 \rangle}}$$

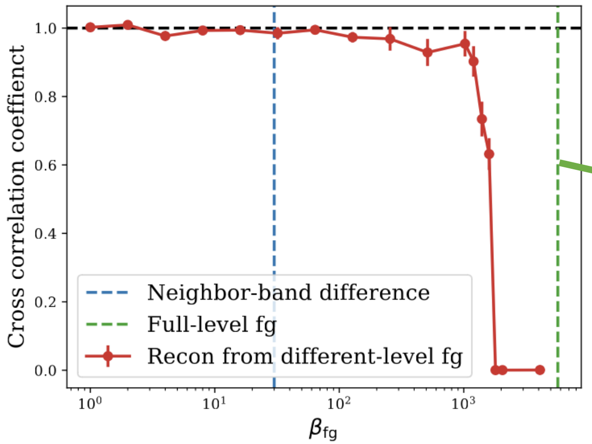
CRIME-generated foreground:

$$T_{syn,0}(\nu, \hat{\boldsymbol{n}}) = T_{Haslam}(\hat{\boldsymbol{n}}) \left(\frac{408\text{MHz}}{\nu} \right)^{\beta(\hat{\boldsymbol{n}})}$$

$$C_\ell(v_1, v_2) = A \left(\frac{\ell_{ref}}{\ell} \right)^{\beta} \left(\frac{v_{ref}^2}{v_1 v_2} \right)^{\alpha} \exp\left(-\frac{\log^2(v_1/v_2)}{2\xi^2} \right)$$

UNet Output:

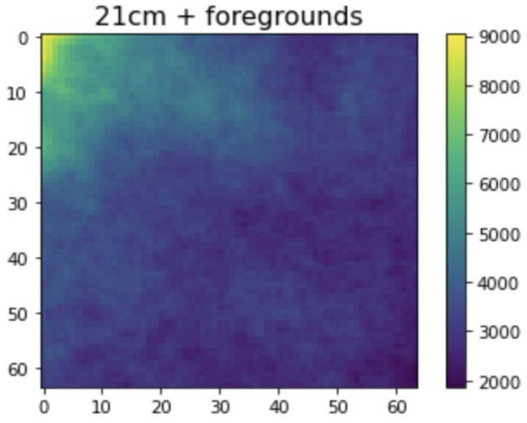




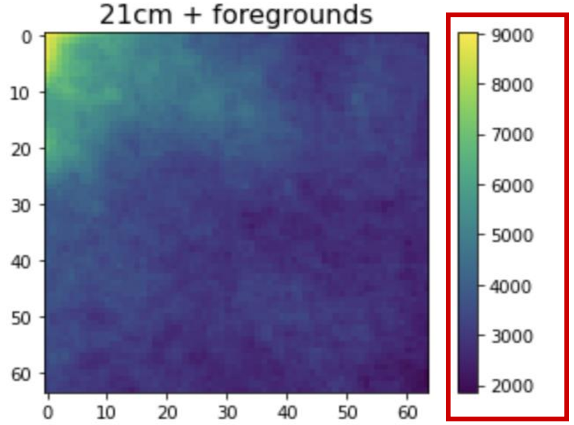
Full-level foregrounds: T_{fg}^{CRIME}

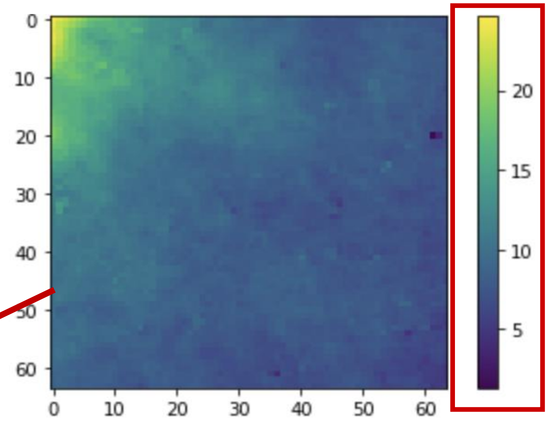
Temperature difference between neighboring frequency bands

Frequency ν_0



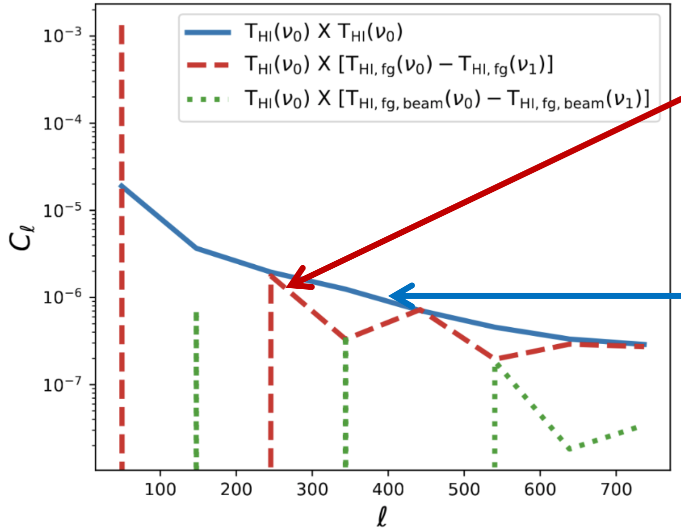
Frequency ν_1

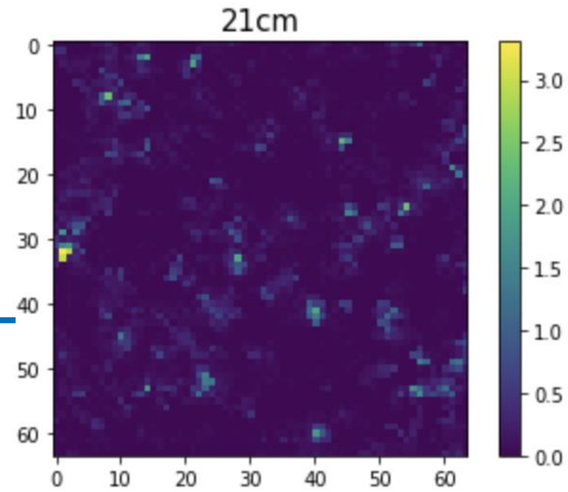




HI : $\nu_0 + \nu_1$

Foregrounds: $\nu_0 - \nu_1$





HI: ν_0

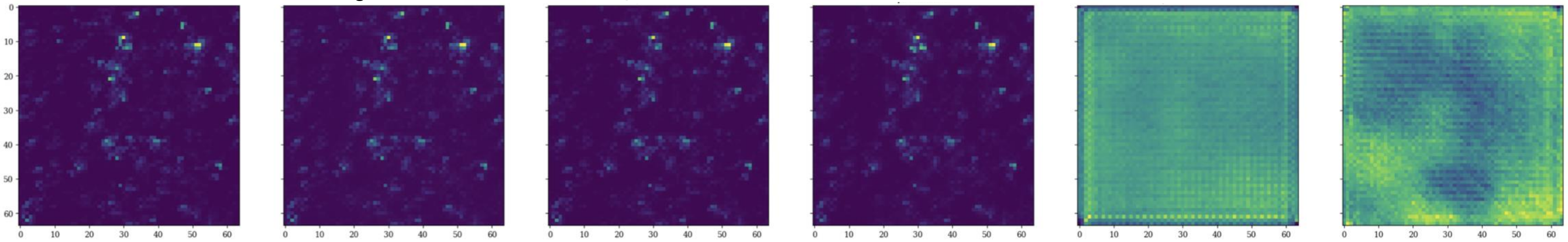
Testing UNet for cleaning different-level foreground

- We run the UNet by inputting $T_{HI} + T_{fg}$, where $T_{fg}=\alpha T_{fg}^{CRIME}$

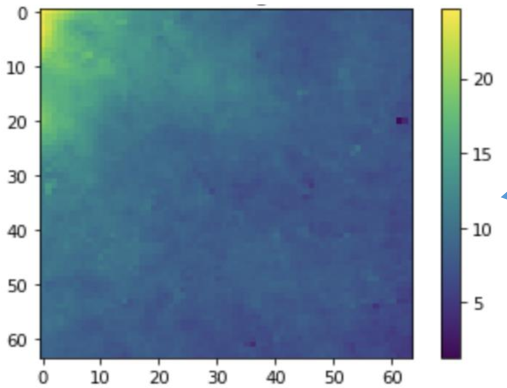
$$\beta_{\mathrm{fg}} = \sqrt{\frac{\langle (T_{\mathrm{fg}} - \overline{T}_{\mathrm{fg}})^2 \rangle}{\langle (T_{\mathrm{HI}} - \overline{T}_{\mathrm{HI}})^2 \rangle}}$$

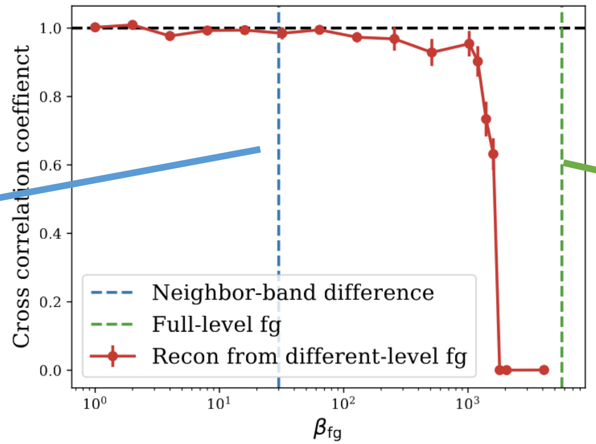
UNet Output：

Target Truth | $\beta_{fg}=4$ | $\beta_{fg}=256$ | $\beta_{fg}=1200$ | $\beta_{fg}=1800$ | $\beta_{fg}=2048$



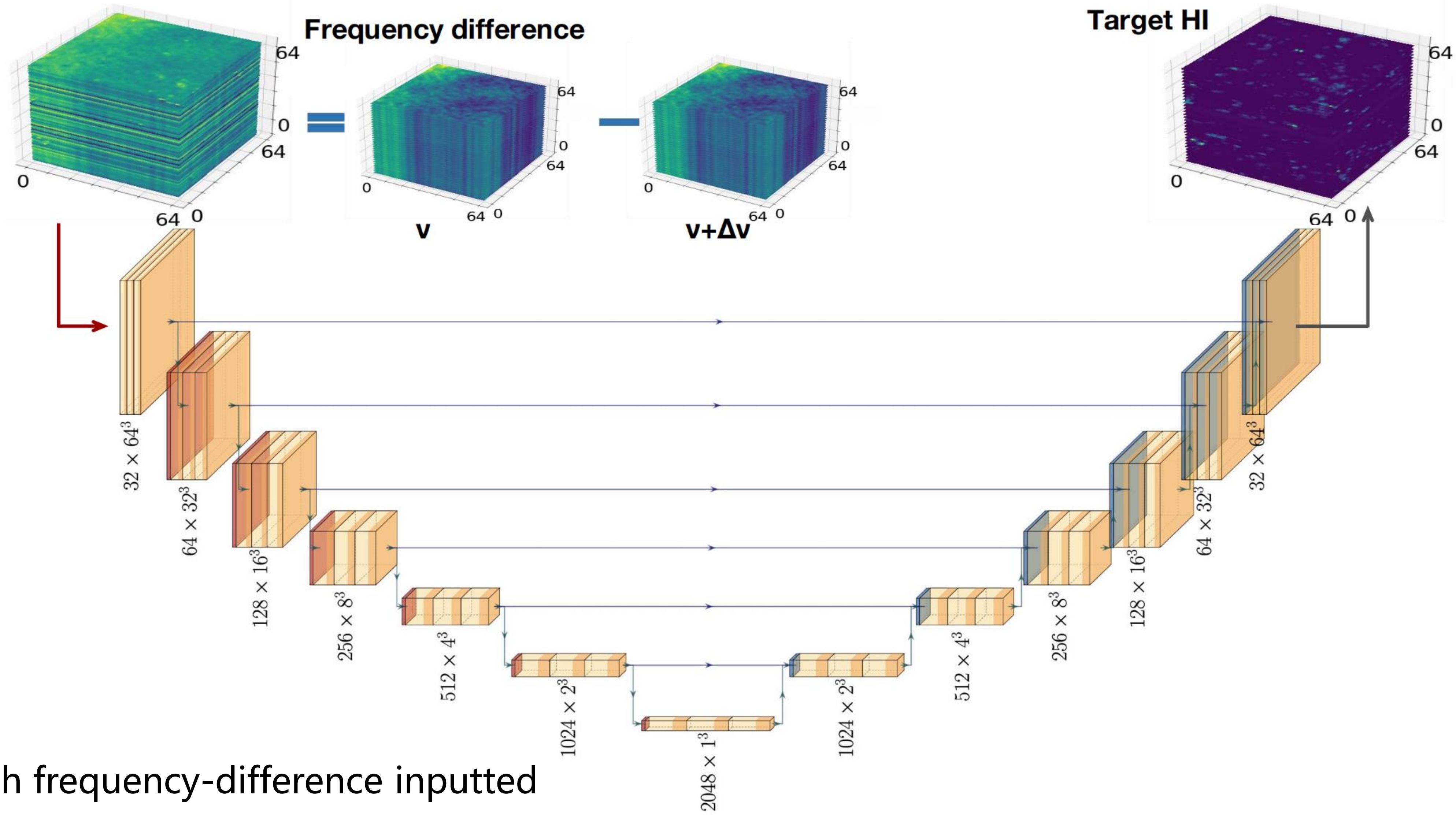
$\nu_0-\nu_1$





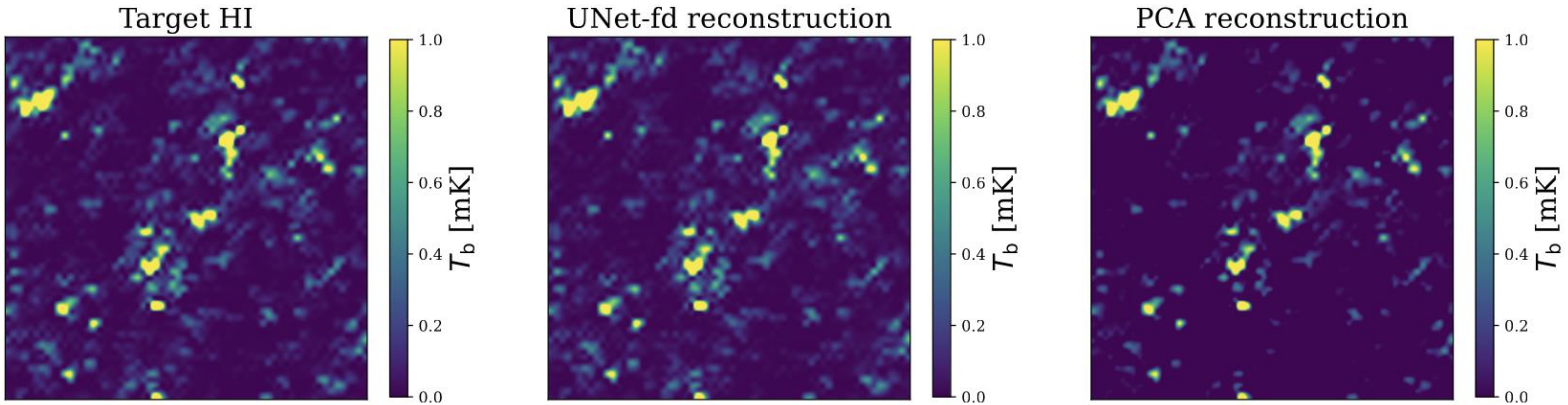
Full-level foregrounds: T_{fg}^{CRIME}

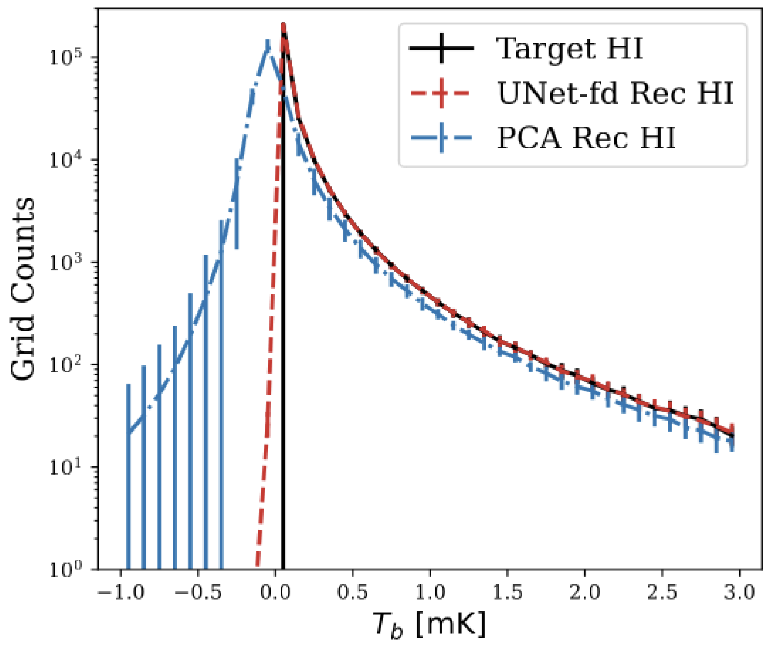
Cleaning model: frequency difference + UNet

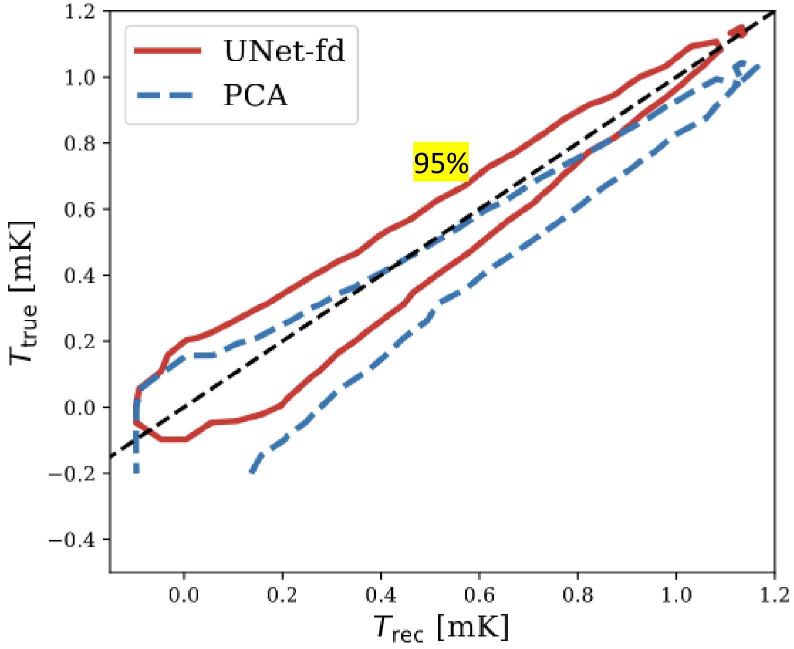


UNet-fd:
UNet with frequency-difference inputted

Test I: 21cm + foreground



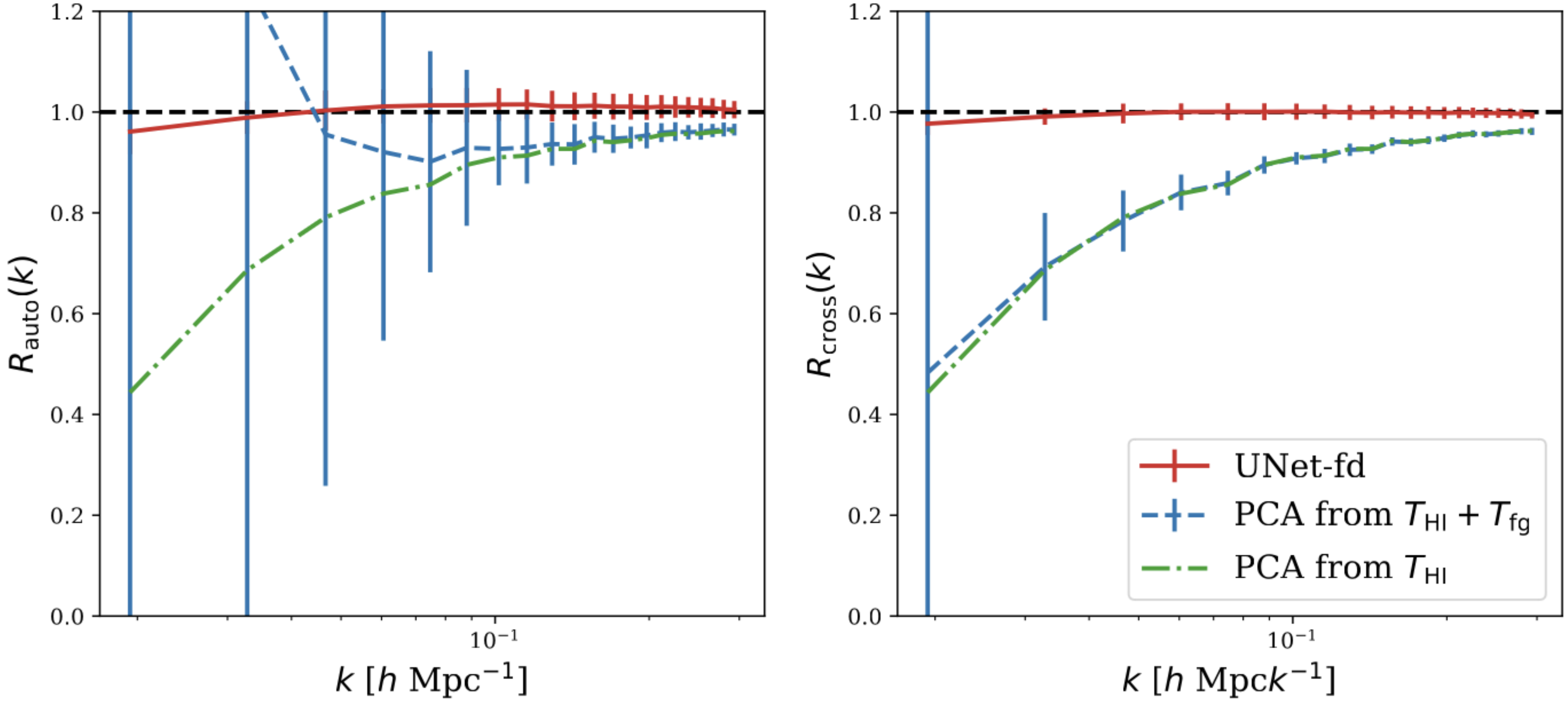




Test I: 21cm + foreground

$$R_{\mathrm{auto}}(k) = \frac{P_{\mathrm{rec,rec}}(k)}{P_{\mathrm{HI,HI}}(k)}$$

$$R_{\mathrm{cross}}(k) = \frac{P_{\mathrm{rec,HI}}(k)}{P_{\mathrm{HI,HI}}(k)}$$

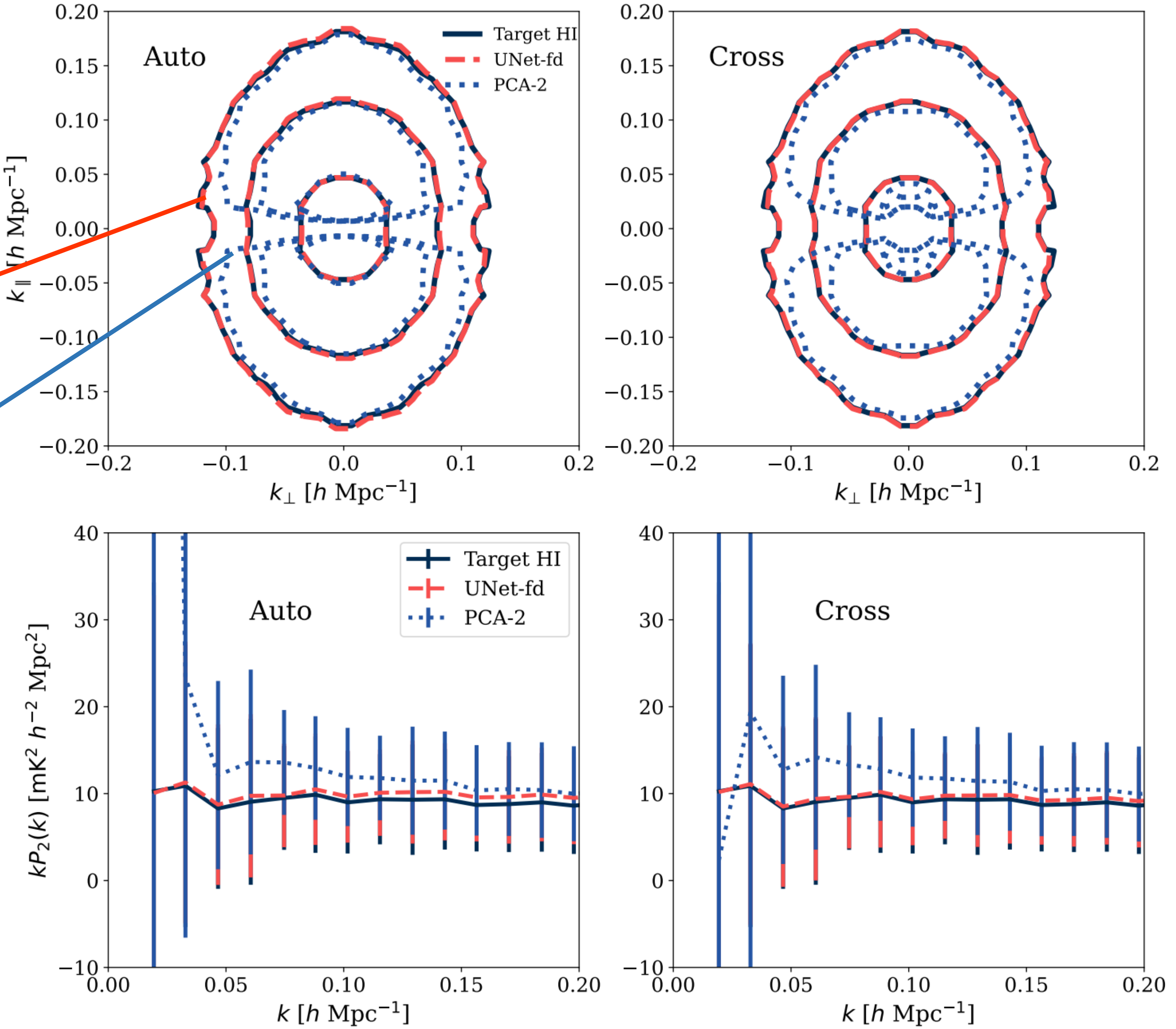


Test I: 21cm + foreground

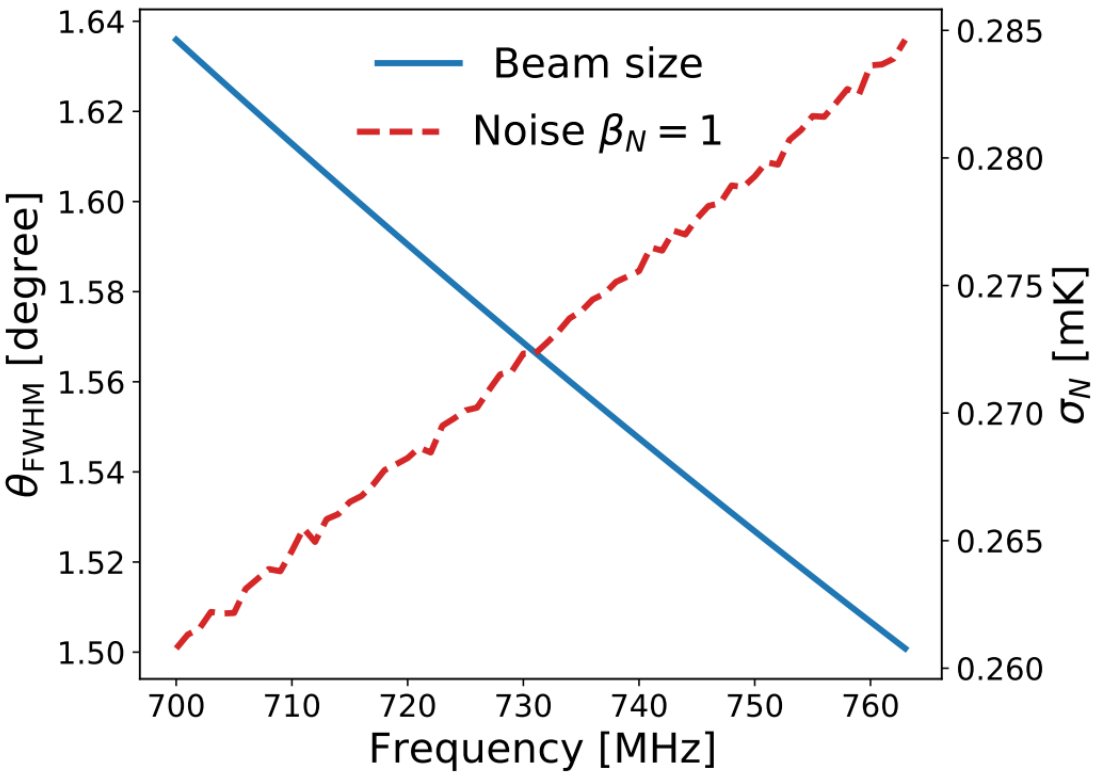
RSD reconstruction

This work

PCA



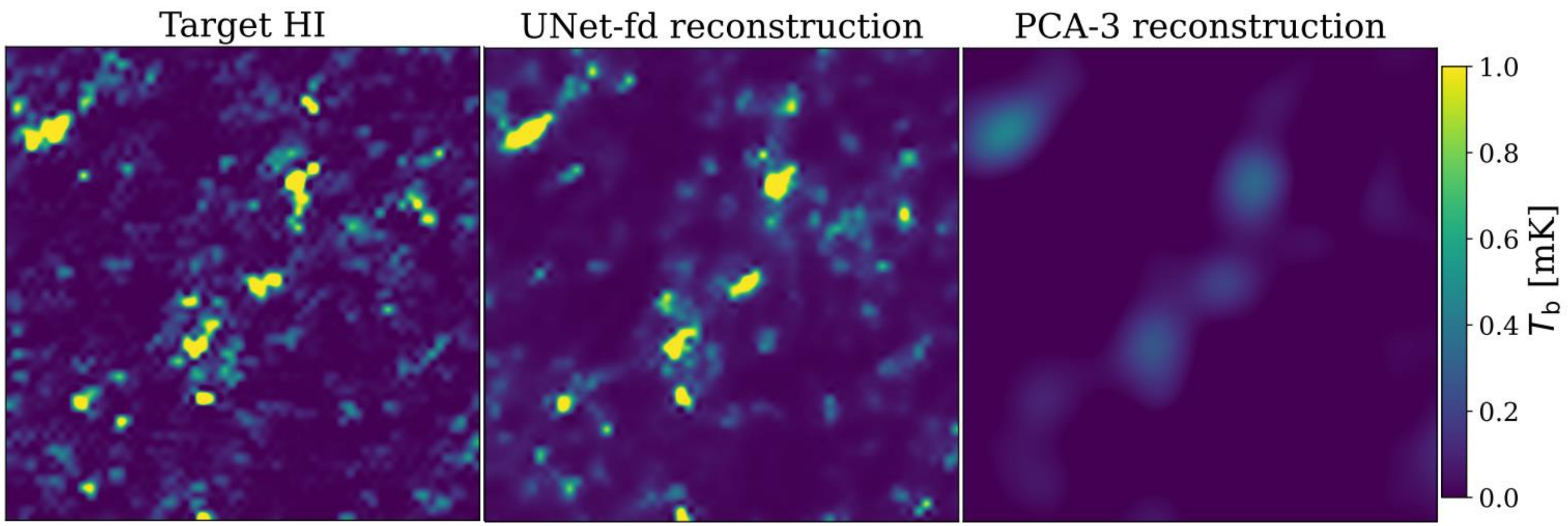
Test II : Adding beam effect

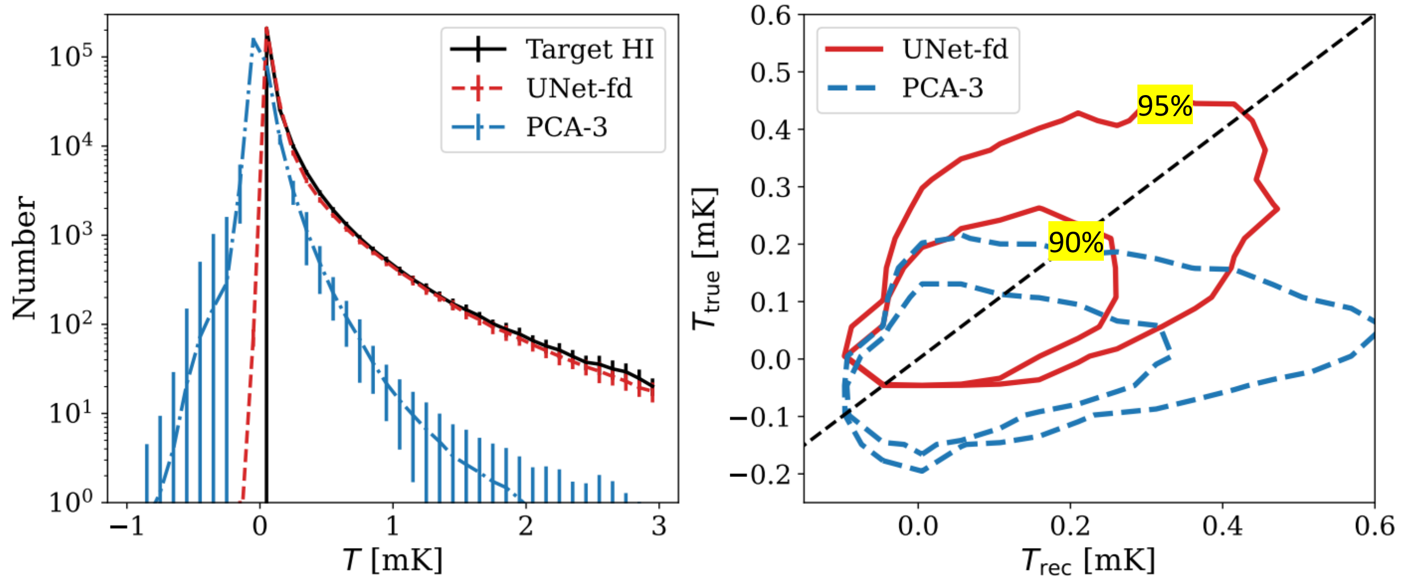


Gaussian beam:

$$B(\nu,\theta) = \exp\left[-4\ln 2\left(\frac{\theta}{\theta_{\mathrm{FWHM}}(\nu)}\right)^2\right]$$

Test II : Adding beam effect

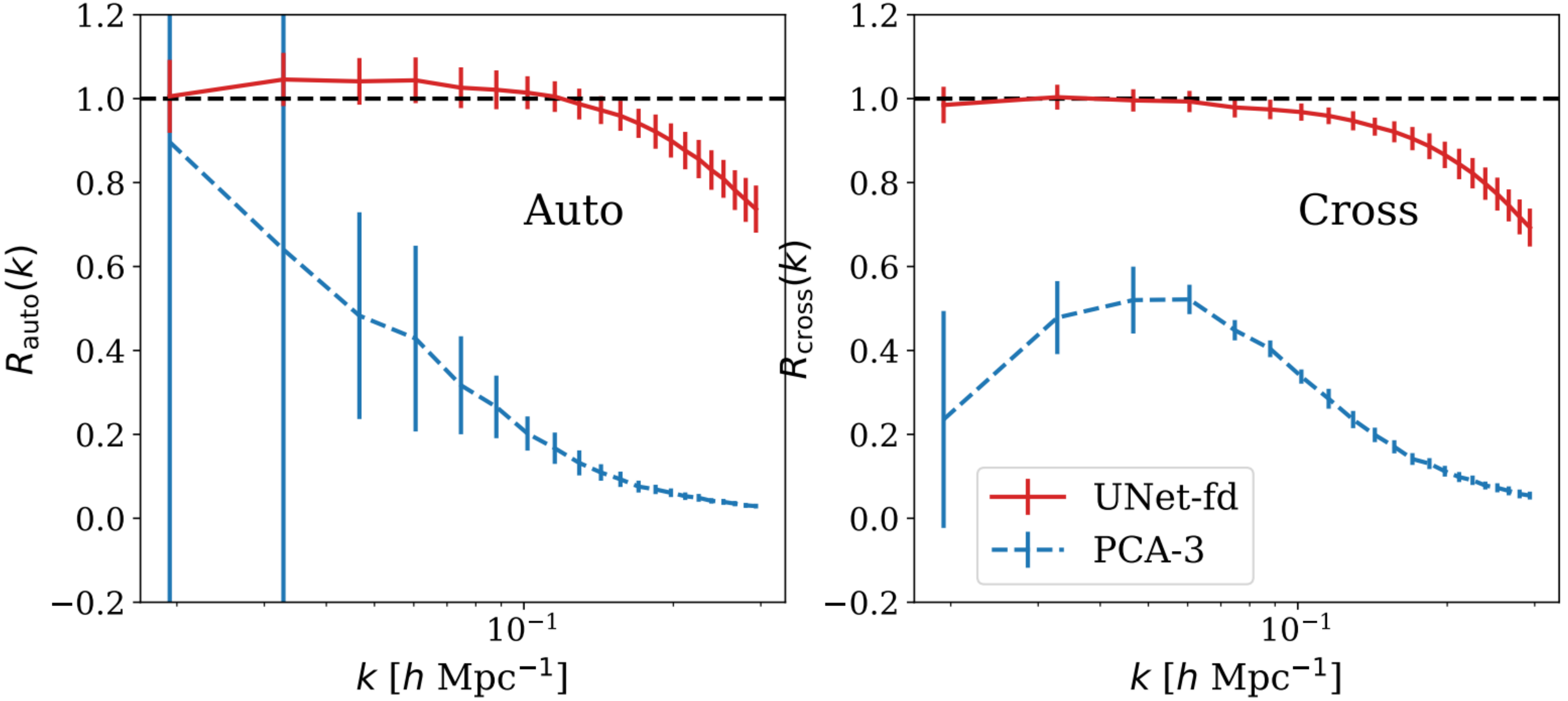




Test II : Adding beam effect

$$R_{\mathrm{auto}}(k) = \frac{P_{\mathrm{rec,rec}}(k)}{P_{\mathrm{HI,HI}}(k)}$$

$$R_{\mathrm{cross}}(k) = \frac{P_{\mathrm{rec,HI}}(k)}{P_{\mathrm{HI,HI}}(k)}$$

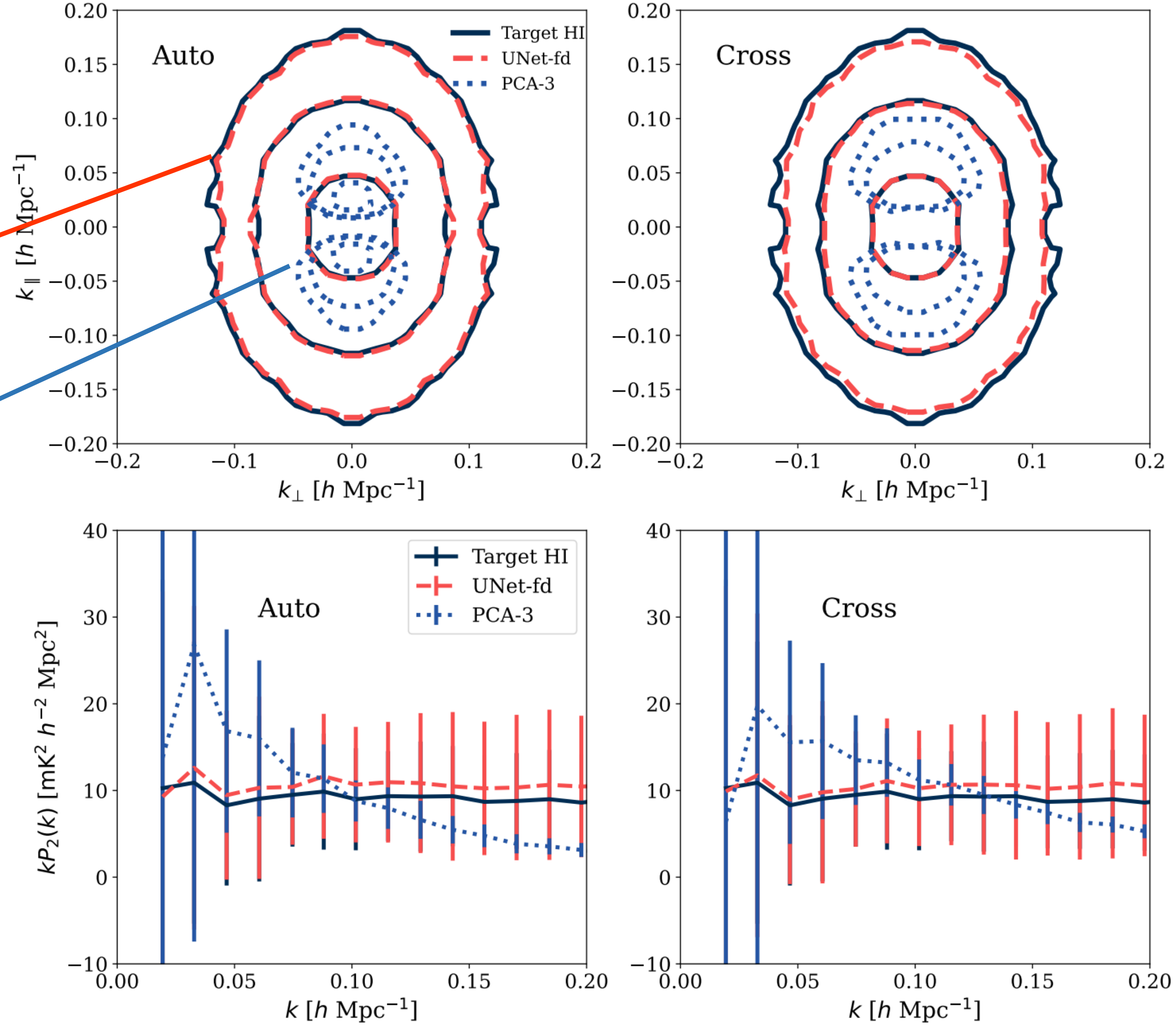


Test II : Adding beam effect

RSD reconstruction

This work

PCA



Summary

(1) Reconstructing the dark matter density field (velocity field, tidal field) using AI

(arXiv:2305.11431)

(2) Reconstructing the 21-cm intensity field using AI and the frequency-difference technique

(arxiv:2310.06518)

https://dark-ai.top/index.html

Science

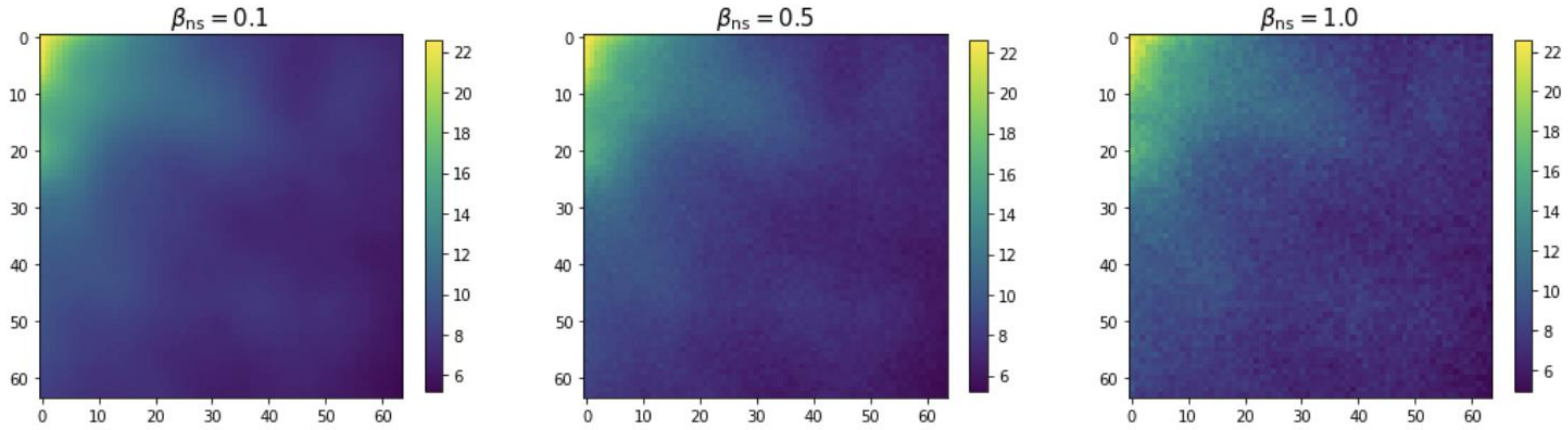
The **DarkAI** project aims to apply state-of-the-art machine learning algorithms to address frontier problems in cosmology. Its scientific goals include exploring the nature of dark energy and dark matter, probing neutrino properties, investigating cosmic expansion and structure growth histories, measuring the Hubble parameter, refining descriptions of the cosmic web, and so on. Our current research focuses on:

- Estimating cosmological parameters at the field level.
- Inferring galaxy velocities.
- Reconstructing the underlying dark matter field.
- Simulating spectroscopic and photometric surveys.
- Integrating machine learning techniques with traditional large-scale structure statistics.
- 21cm foreground removal.

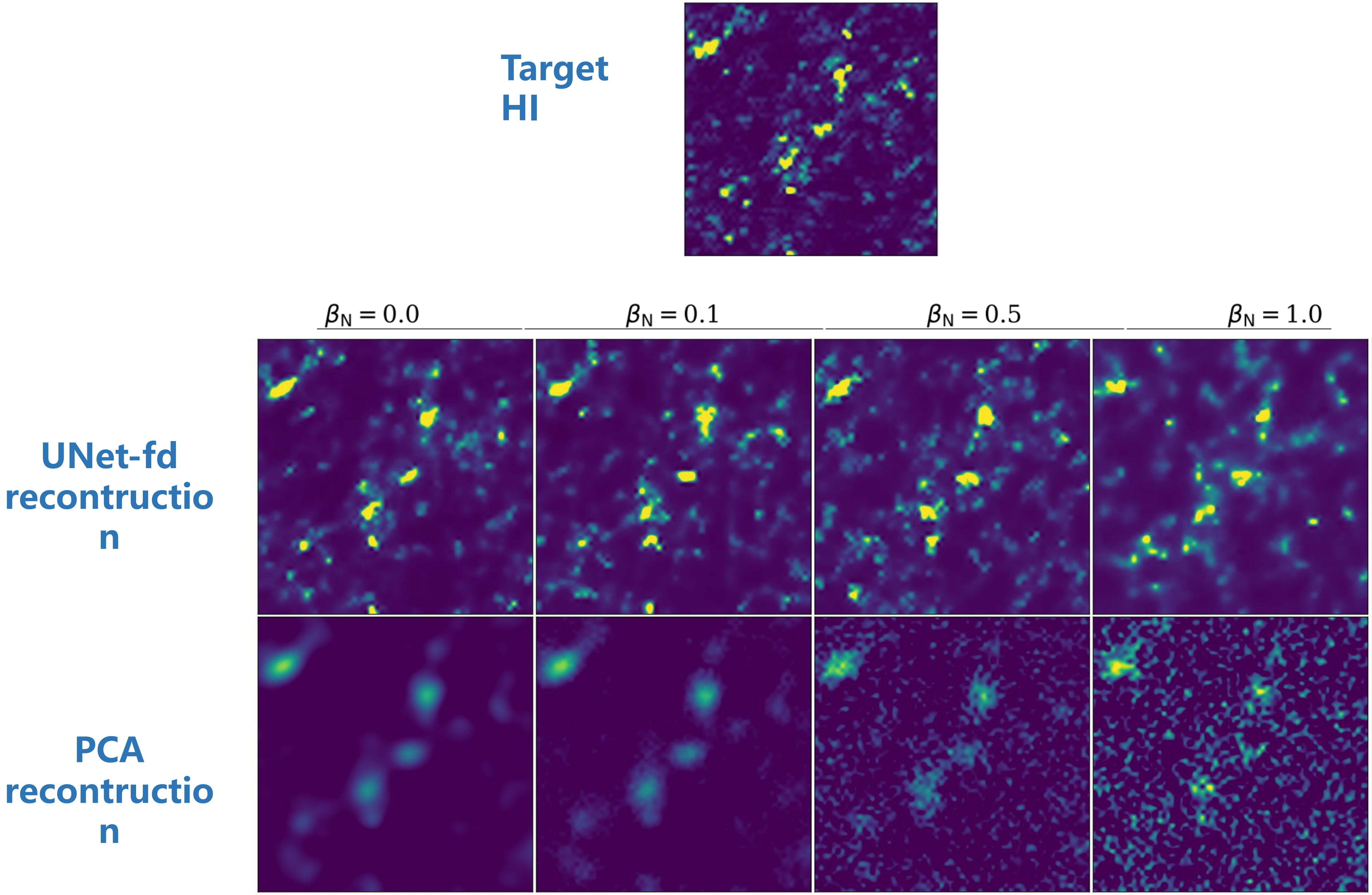
Test III : Varying thermal noise

Input:

$$\sigma_{\rm ns} = T_{\rm sys}\sqrt{\frac{4\pi f_{\rm sky}}{\Omega_{\rm beam} N_{\rm dish} t_{\rm obs} \Delta\nu}} \qquad \beta_{\rm ns} = \frac{\sigma_{\rm ns}}{\sigma_{\rm HI}}$$



Test III : Varying thermal noise



Test III : Varying thermal noise

$$R_{\rm auto}(k) = \frac{P_{\rm rec,rec}(k)}{P_{\rm HI,HI}(k)}$$

$$R_{\rm cross}(k) = \frac{P_{\rm rec,HI}(k)}{P_{\rm HI,HI}(k)}$$

Averaged $R_{\rm cross}(k)$ over $0<k<0.1$ hMpc^{-1}