Eliminate the effects of polarization leakage with PCA and deep learning

Liyang Gao, Yichao Li, Shulei Ni, and Xin Zhang

Supervisor: Prof. dr. Xin Zhang and Prof. dr. Léon Koopmans

Shenyang, P.R. China

arXiv:2212.08773

Background

21cm Intensity Mapping

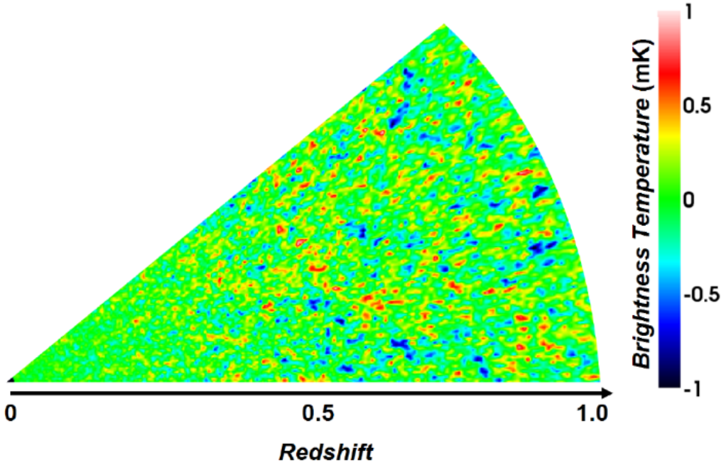
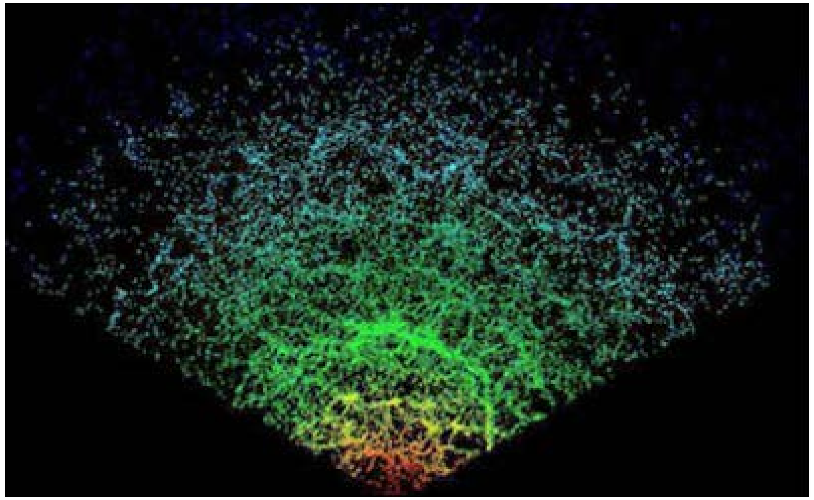


Figure 2: Simulated fluctuations in the brightness temperature of 21cm emission from galaxies in a slice through the universe. The emission is smoothed over 8/h Mpc. The redshift, z, translates to frequency: ν=1.42GHz/(1+z). Red indicates overdensity and blue underdensity.

21cm Intensity Mapping
(from arXiv:0902.3091)



Galactic Survey
(from Sloan Digital Sky Survey)

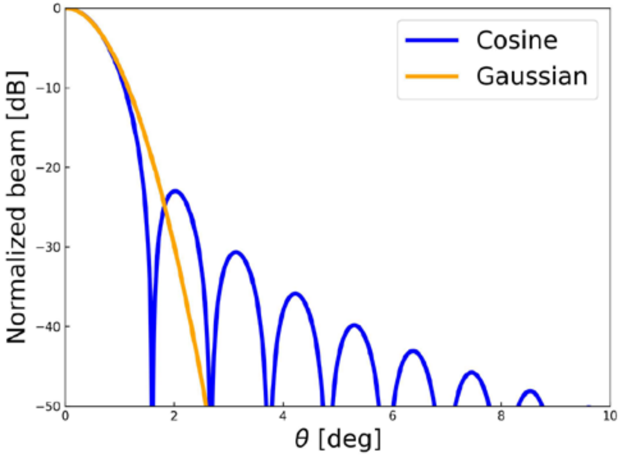
Detects dark ages, cosmic dawns, and reionisation periods;
fast survey speeds and large survey volumes;
can interrelate with optical surveys.

Low resolution and difficulty in extracting 21cm signals.

Challenge:
Beam effect
polarization leakage

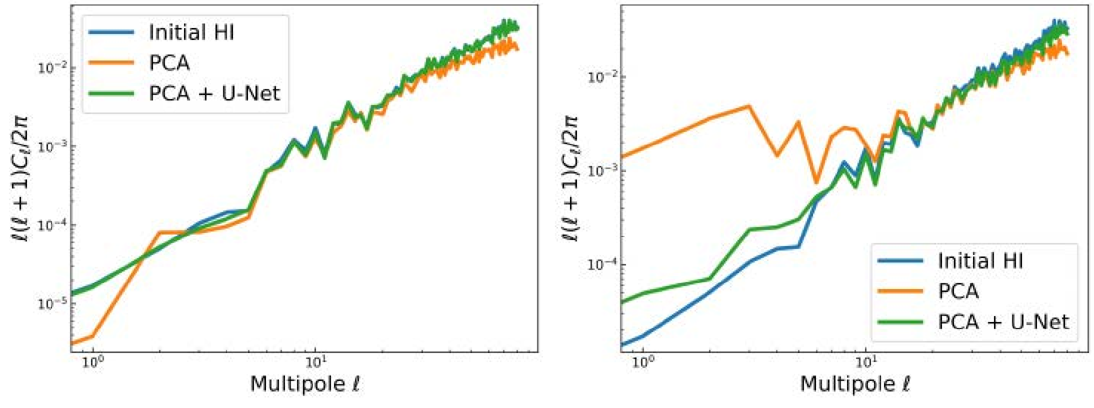
Eliminate beam effects

[ArXiv:2204.02780](https://arxiv.org/abs/2204.02780)



Do convolution with Gaussian beam

Do convolution with Cosine beam



(a) (b)

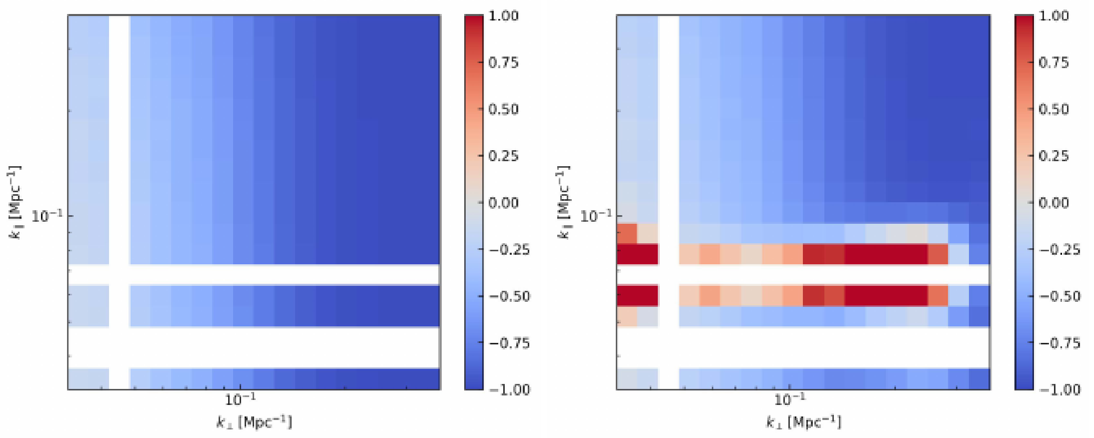
Power spectrum:

$$R(k_\parallel, k_\perp)_{\text{auto}} = \frac{\bar{P}_{\text{cln}}(k_\parallel, k_\perp)}{P_{\text{HI}}(k_\parallel, k_\perp)} - 1$$

- For Gaussian beam : U-Net improved by 27.4% over PCA;
- For Cosine beam : U-Net improved by 144.8 % over PCA.

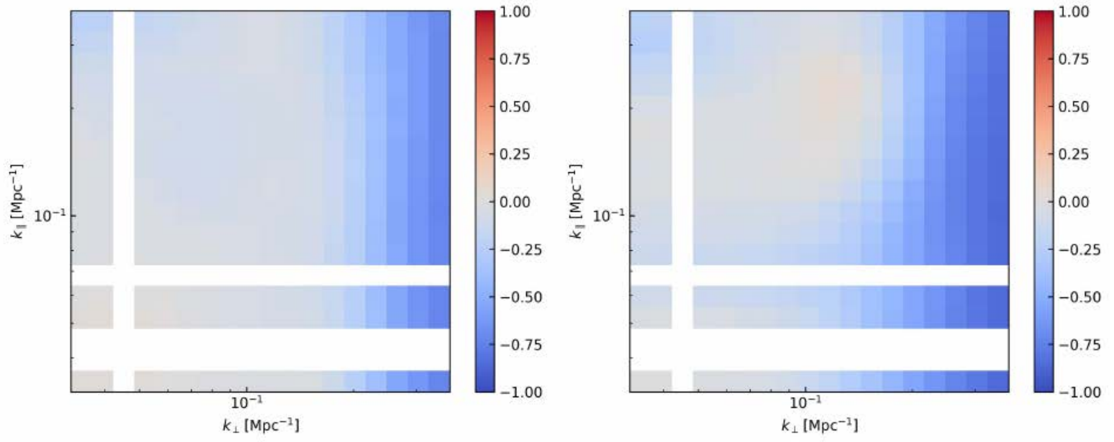
Do convolution with Gaussian beam

Do convolution with Cosine beam



(a) (b)

auto-correlation power spectrum ratio (PCA& initial HI)



(a) (b)

auto-correlation power spectrum ratio (U-Net& initial HI)

Simulation

whole sky map

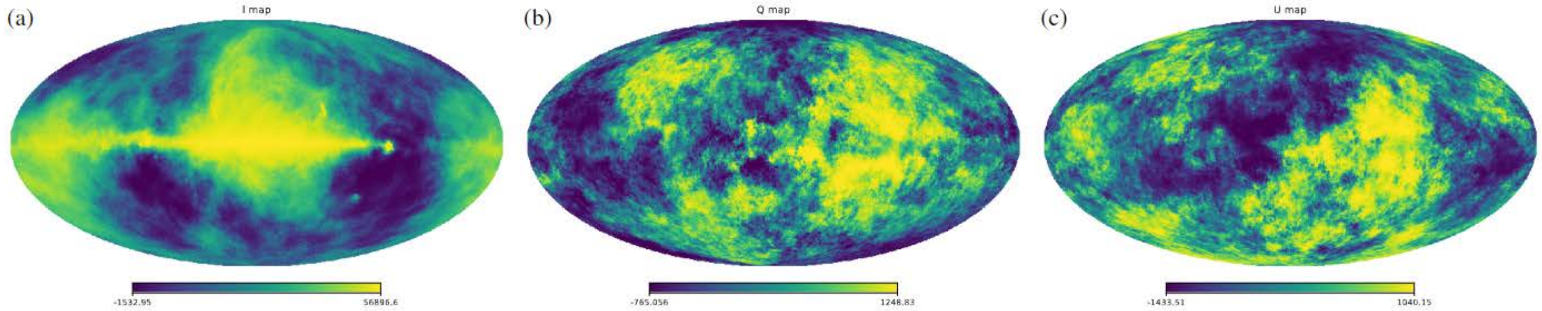


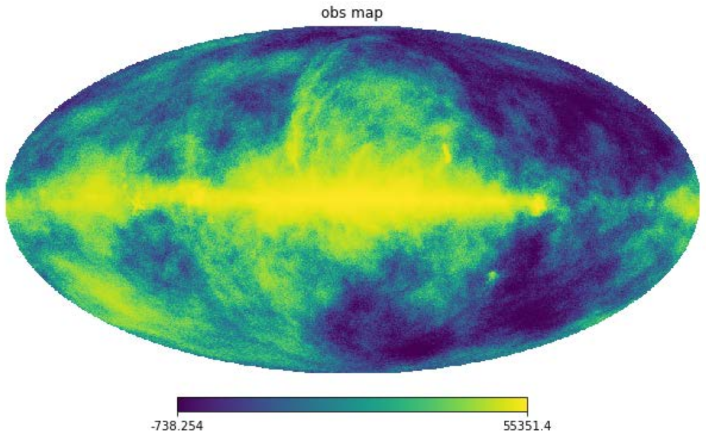
Figure 1. (a), (b), (c) are the I, Q, U sky maps of synchrotron emission at one random frequency, respectively. The unit of each pixel of the sky maps is mK.

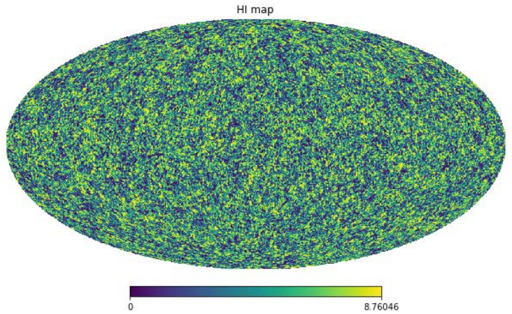
970 MHz ~ 1023 MHz

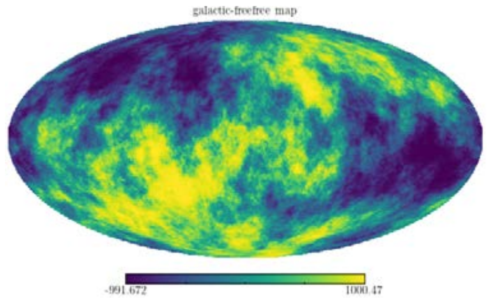
$$T_{\rm leak} = \epsilon_Q T_{\rm FG}^{Q} + \epsilon_U T_{\rm FG}^{U}, \qquad (6)$$

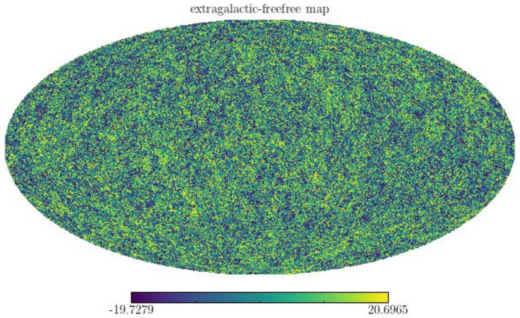
- $(\epsilon_Q, \epsilon_U) = (0.0\%, 0.0\%)$;
- $(\epsilon_Q, \epsilon_U) = (0.5\%, 0.0\%)$;
- $(\epsilon_Q, \epsilon_U) = (1.0\%, 0.0\%)$;
- $(\epsilon_Q, \epsilon_U) = (2.0\%, 0.0\%)$;
- $(\epsilon_Q, \epsilon_U) = (0.0\%, 1.0\%)$;
- $(\epsilon_Q, \epsilon_U) = (0.5\%, 1.0\%)$;
- $(\epsilon_Q, \epsilon_U) = (1.0\%, 1.0\%)$;
- $(\epsilon_Q, \epsilon_U) = (2.0\%, 1.0\%)$.

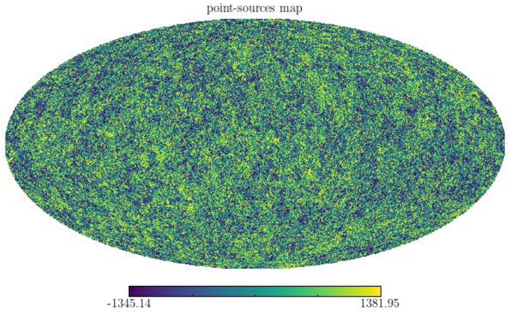
$$T_{\rm sky}(\nu, \hat{\mathbf{n}}) = T_{\rm HI} + T_{\rm unpol} + T_{\rm leak}. \qquad (7)$$





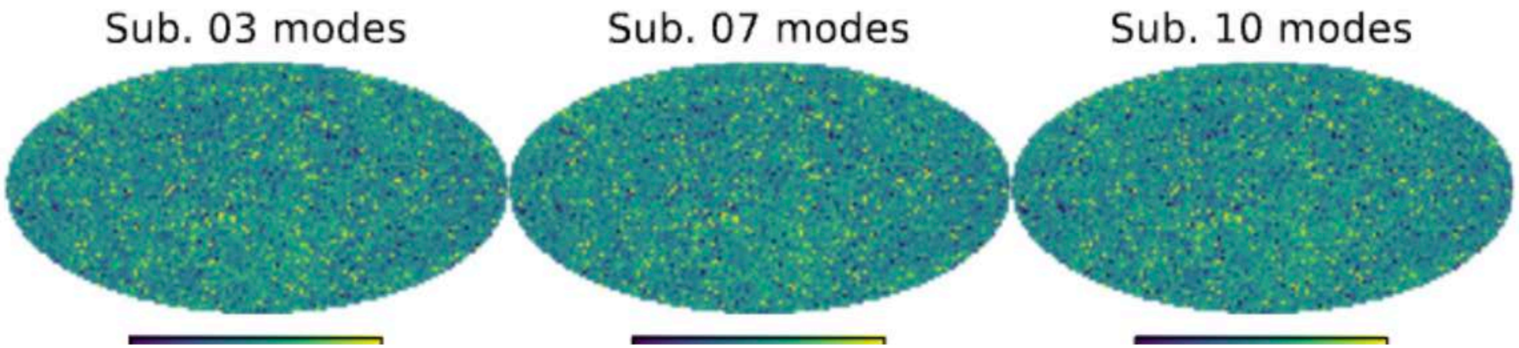


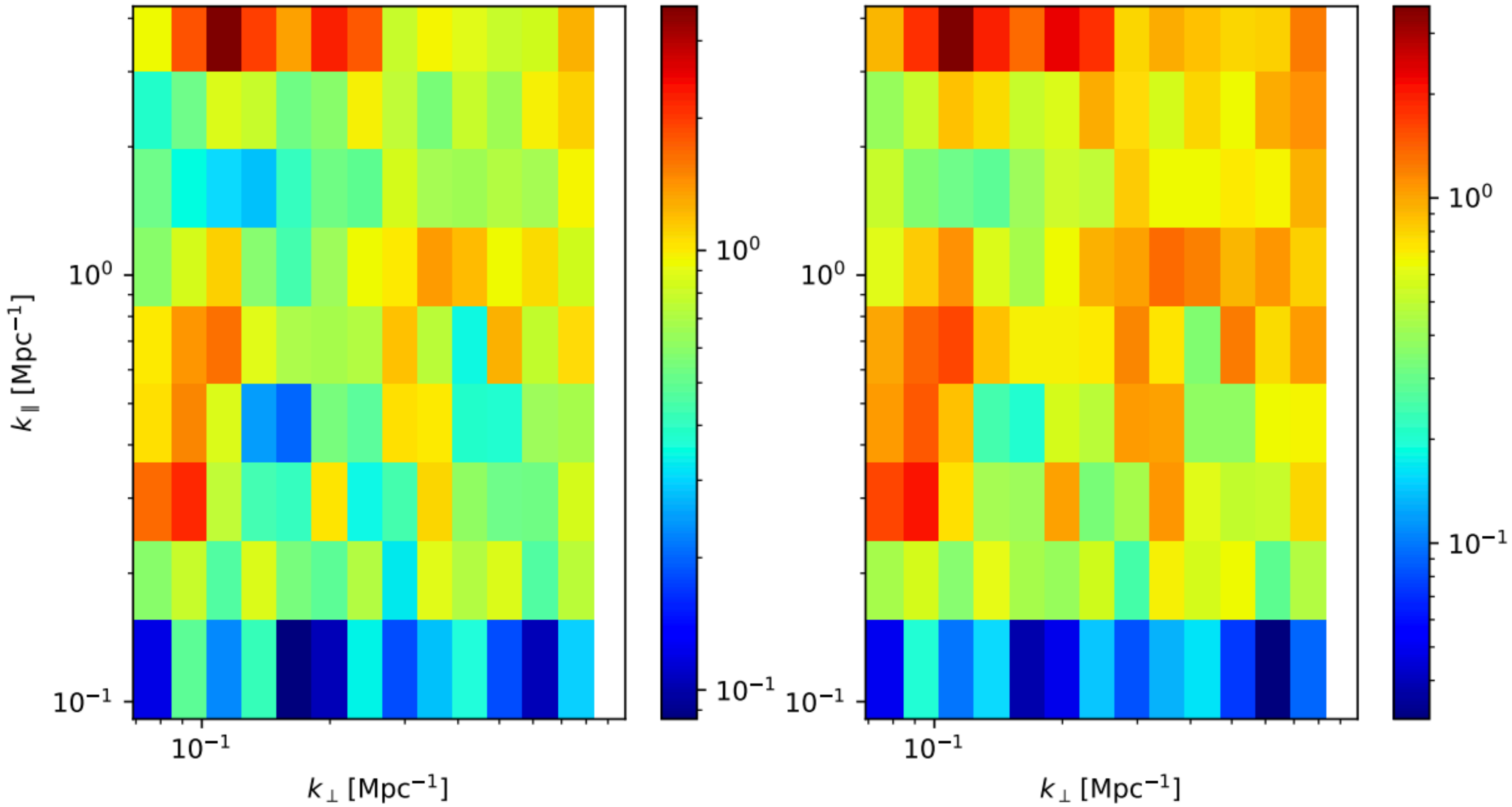


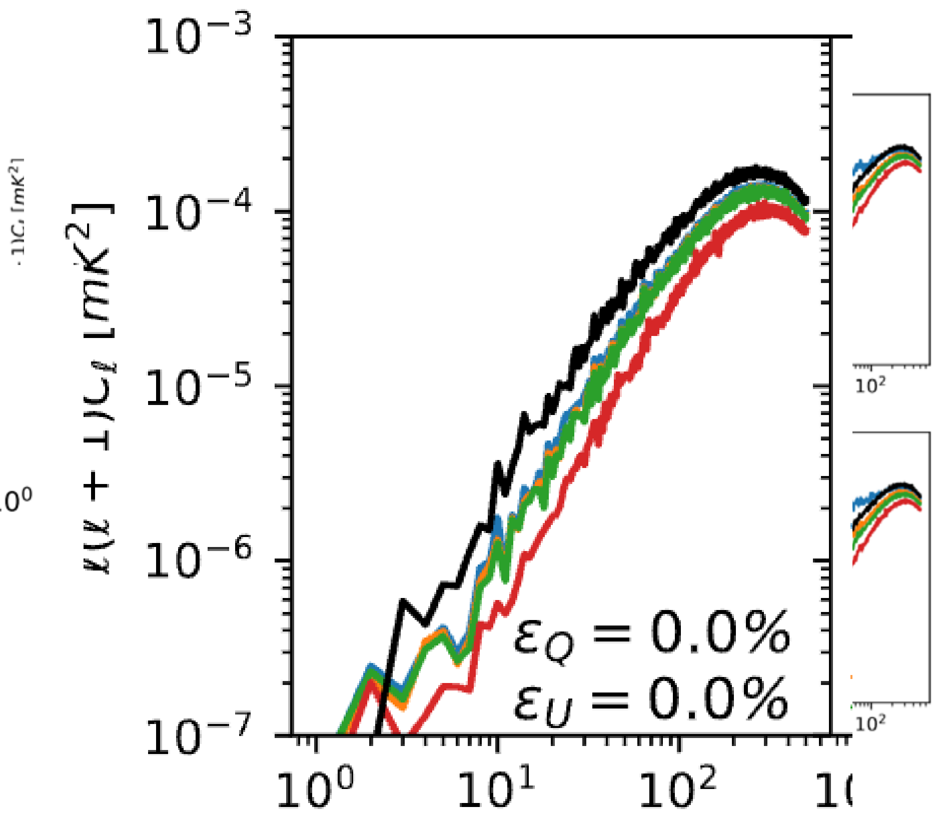


CRIME (http://intensitymapping.physics.ox.ac.uk/CRIME.html, Alonso et al. 2014.)

PCA Results



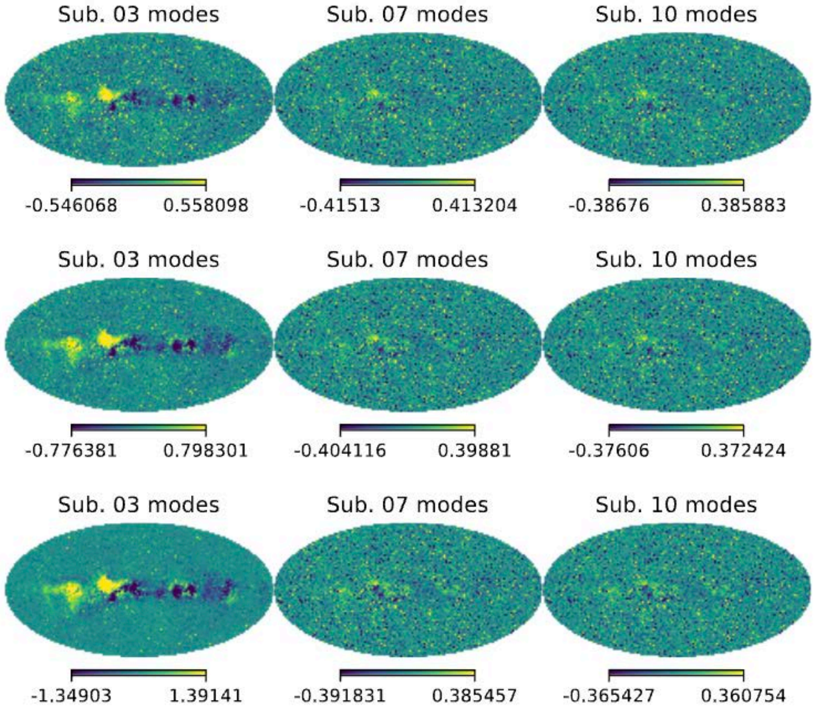


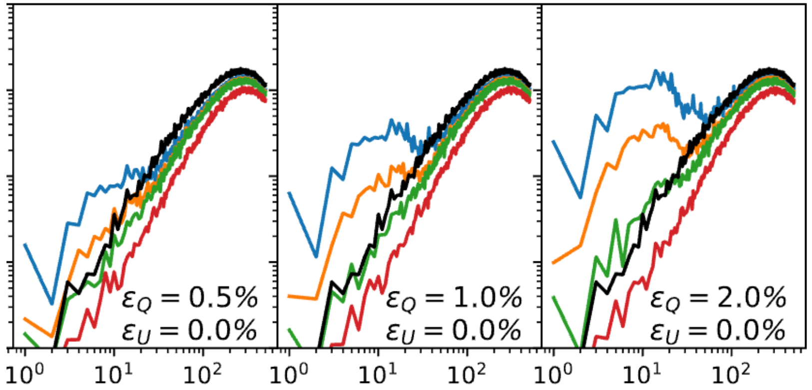


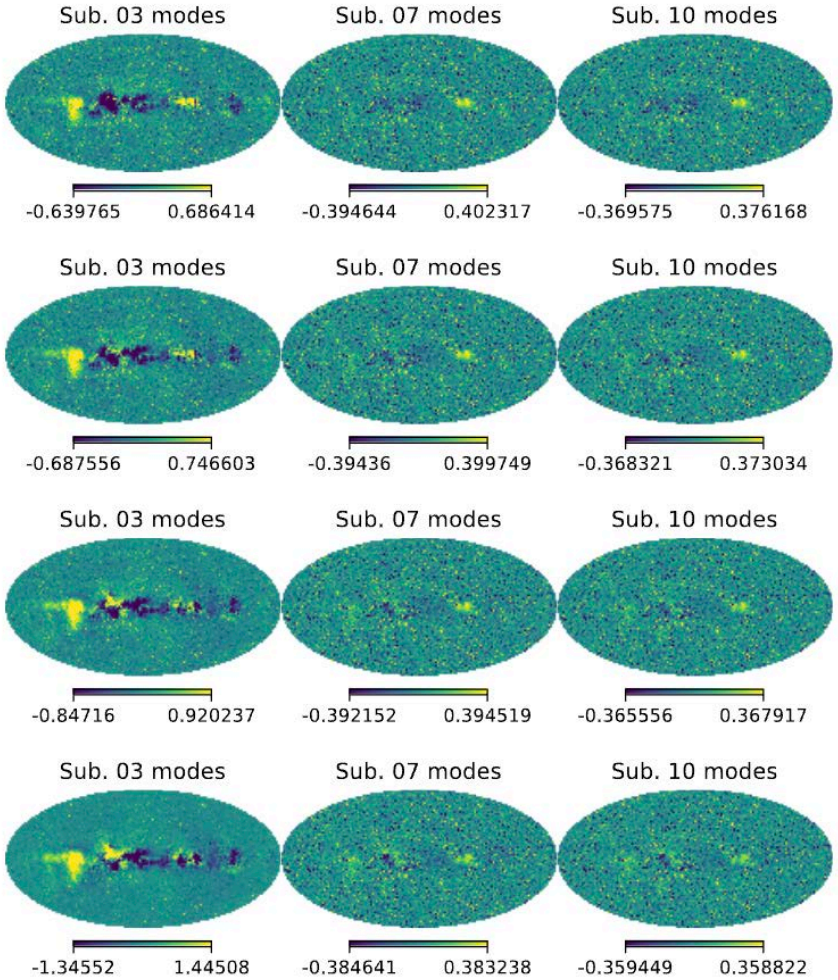
to 4 represent
)5, 0.01, and 0.02 respectively. The first and second rows represent ϵ_U taken as 0 and 0.01 respectively. The black lines represent the sky
l. The blue, yellow, green, red, and purple lines represent the power spectrum after subtracting 3, 4, 5, and 10 modes using PCA, respectively.

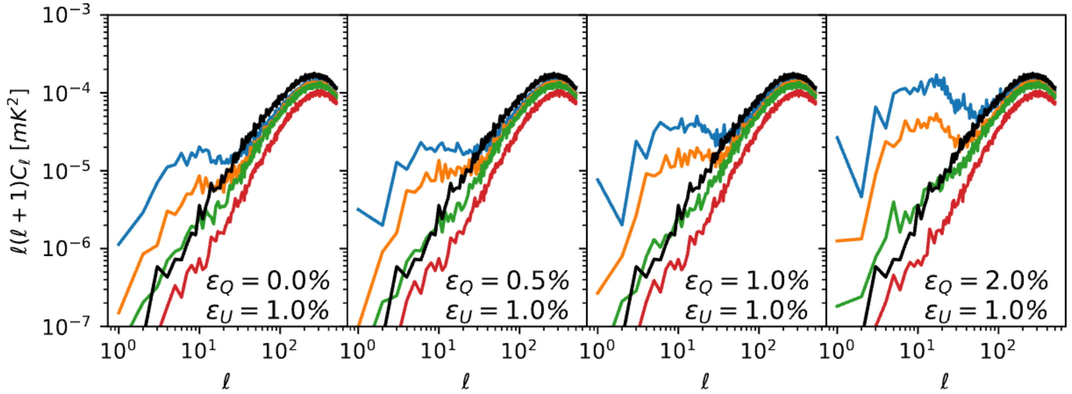
Fig
to t
(0.(
in d
in tl
sky

PCA Results


Sub. 03 modes
Sub. 07 modes
Sub. 10 modes
-0.546068 0.558098
-0.41513 0.413204
-0.38676 0.385883
Sub. 03 modes
Sub. 07 modes
Sub. 10 modes
-0.776381 0.798301
-0.404116 0.39881
-0.37606 0.372424
Sub. 03 modes
Sub. 07 modes
Sub. 10 modes
-1.34903 1.39141
-0.391831 0.385457
-0.365427 0.360754


$\epsilon_Q = 0.5\%$
$\epsilon_U = 0.0\%$
$\epsilon_Q = 1.0\%$
$\epsilon_U = 0.0\%$
$\epsilon_Q = 2.0\%$
$\epsilon_U = 0.0\%$
10^0 10^1 10^2


Sub. 03 modes
Sub. 07 modes
Sub. 10 modes
-0.639765 0.686414
-0.394644 0.402317
-0.369575 0.376168
Sub. 03 modes
Sub. 07 modes
Sub. 10 modes
-0.687556 0.746603
-0.39436 0.399749
-0.368321 0.373034
Sub. 03 modes
Sub. 07 modes
Sub. 10 modes
-0.84716 0.920237
-0.392152 0.394519
-0.365556 0.367917
Sub. 03 modes
Sub. 07 modes
Sub. 10 modes
-1.34552 1.44508
-0.384641 0.383238
-0.359449 0.358822


$\ell(\ell+1)C_\ell\ [mK^2]$
10^{-3} 10^{-4} 10^{-5} 10^{-6} 10^{-7}
$\epsilon_Q = 0.0\%$
$\epsilon_U = 1.0\%$
$\epsilon_Q = 0.5\%$
$\epsilon_U = 1.0\%$
$\epsilon_Q = 1.0\%$
$\epsilon_U = 1.0\%$
$\epsilon_Q = 2.0\%$
$\epsilon_U = 1.0\%$
10^0 10^1 10^2
ℓ

Deep learning and Hyperparameters

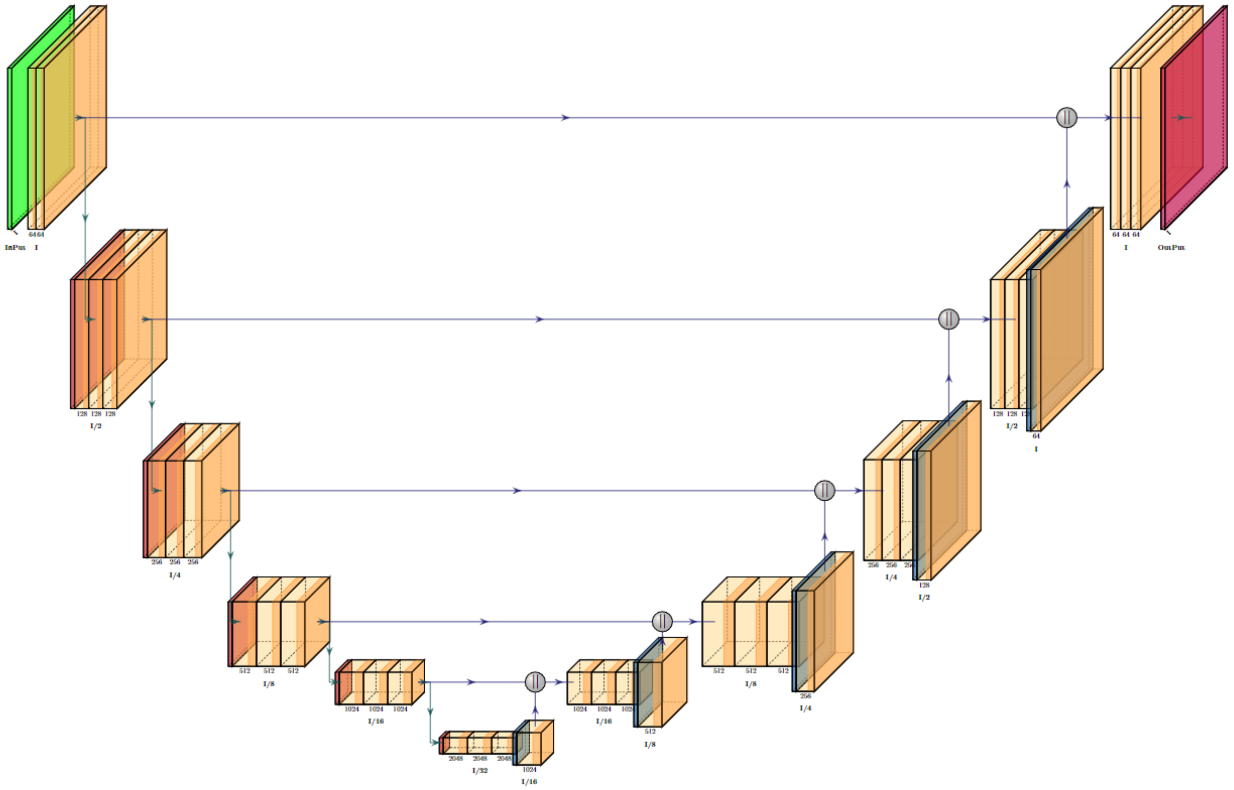


Figure 2. Training process of CNN with U-Net architecture. Each colour represents a structure in the U-Net network, where yellow cubes represent the convolutional layers and ReLU sections, red cubes represent pooling layers in down-sampling, blue cubes represent the transposed convolutional layers, and gray spheres represent connection layers.The green and purple squares at the beginning and end of this figure represent the input and output, respectively.

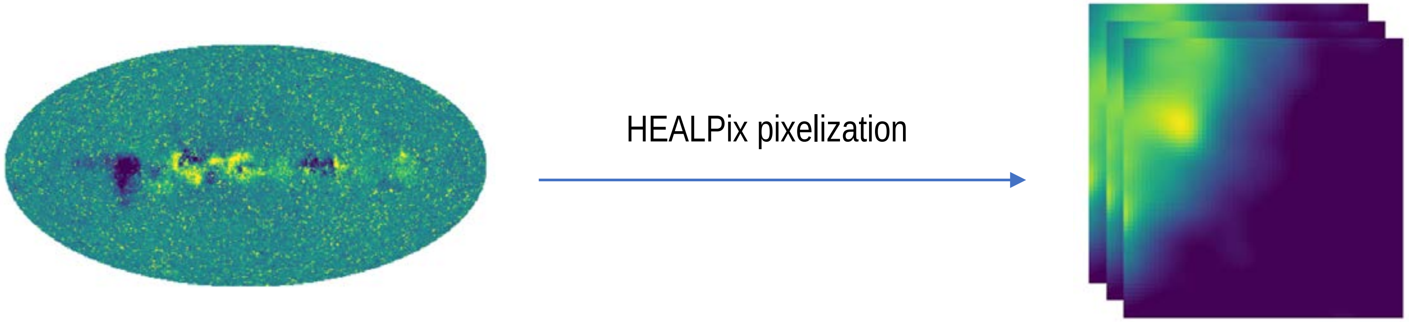


Table 2. Description of the hyperparameters in the U-Net architecture design.

Hyperparameter	Optimum value
n_{block} (number of convolutions for each block)	3
n_{down} (number of down-convolutions)	5
n_{tc} (number of transpose convolutions)	4
batchnorm (batch normalization for given layer)	Ture*
ω (weight decay for optimizer)	10^{-5}
batch size (number of samples per gradient descent step)	16
n_{filter} (initial number of convolution filters)	36
Ω (optimizer for training)	AdamW
η (learning rate for optimizer)	10^{-3}
β_{mom} (batch normalization momentum)	0.02

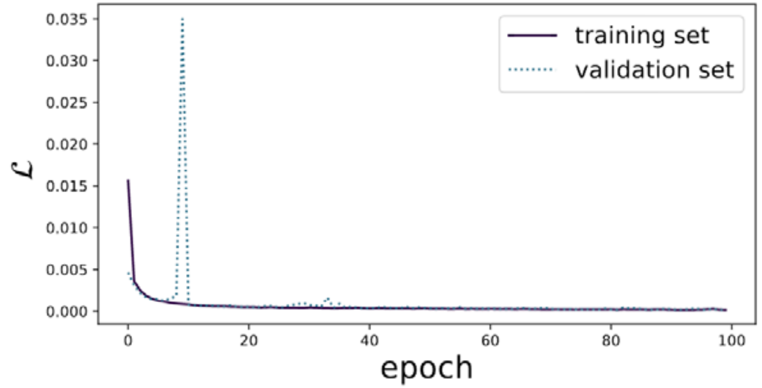


Figure 3. LogCosh loss evolves with the number of epochs. The solid line represents the evolution of the loss in the training set and the dashed line represents the evolution of the loss in the validation set.

PCA+U-Net Results

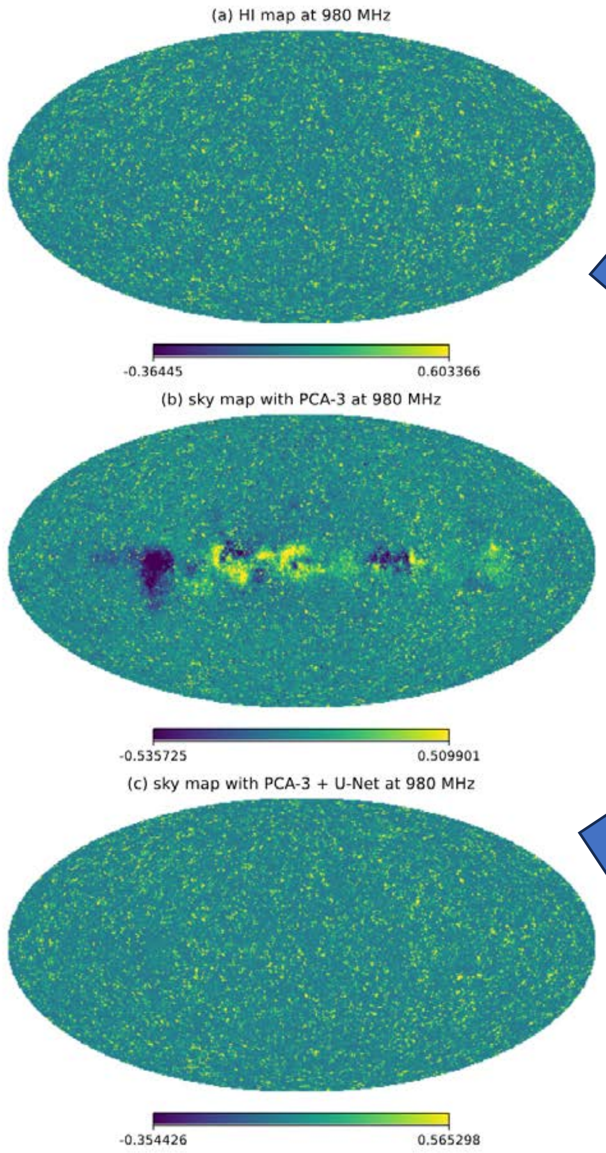


Figure 8. Comparison of the HI signal, PCA-3 result, and sky map processed with U-Net and synthesized by HEALPix **with $(\epsilon_Q, \epsilon_U) = (0.5\%, 1.0\%)$. Complete sky map is processed with U-Net and synthesized by HEALPix.** Panel (a) shows the HI map, panel (b) shows the PCA-3 processed sky map, and panel (c) shows the sky map processed using PCA-3 + U-Net. For convenience, we show only the sky maps at 980 MHz. The unit of each pixel of the sky maps is mK.

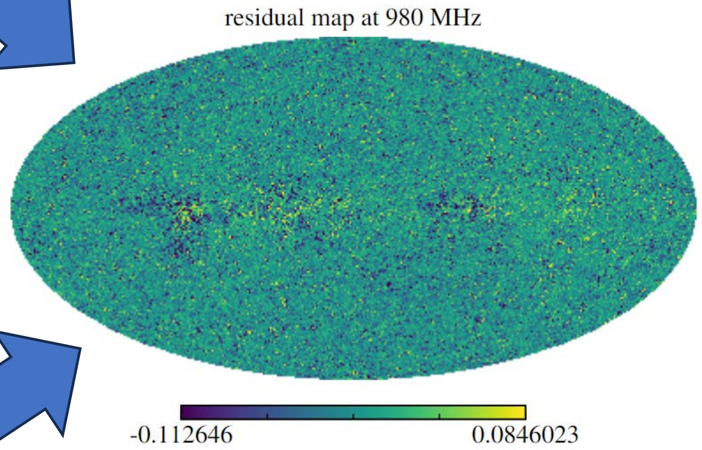


Figure 9. Residual sky map obtained by subtracting the target HI signal sky map from the sky map obtained by PCA-3 + U-Net in $(\epsilon_Q, \epsilon_U) = (0.5\%, 1.0\%)$ case at 980 MHz. The unit of each pixel of the sky maps is mK.

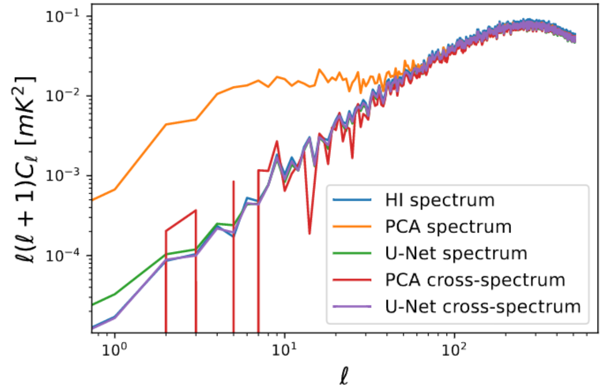
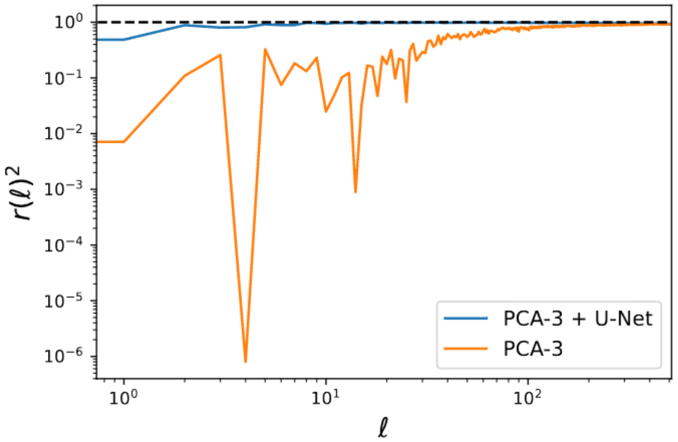


Figure 10. Comparison of the power spectra of HI signal, PCA-3, and U-Net with $[\epsilon_Q, \epsilon_U] = [0.5\%, 1.0\%]$. To eliminate errors arising from individual frequencies, the results are obtained using ten frequency bands averaged. The blue line represents the real HI signal. The yellow and green lines represent the results of PCA-3 and PCA-3 + U-Net, and the red and purple lines represent their cross power spectra with the HI signal.

(a)



$$r(\ell) = \frac{P_{\mathrm{cross}}(\ell)}{\sqrt{P_{\mathrm{HI}}(\ell) P_{\mathrm{p}}(\ell)}},$$

Alleviating Signal Loss

$$T(\ell) = \sqrt{\frac{P_{\mathrm{p}}(\ell)}{P_{\mathrm{HI}}(\ell)}}.$$

(b)

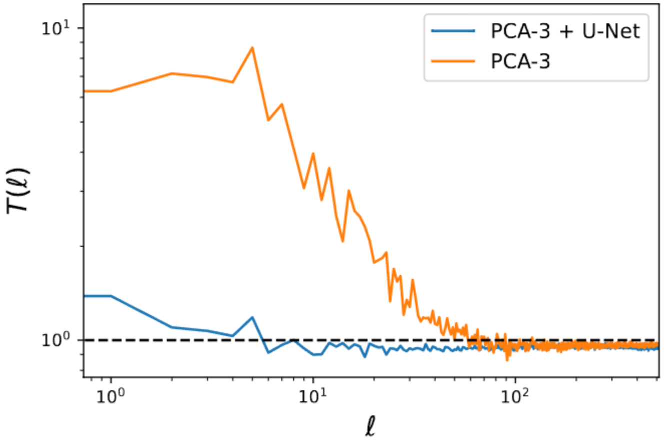
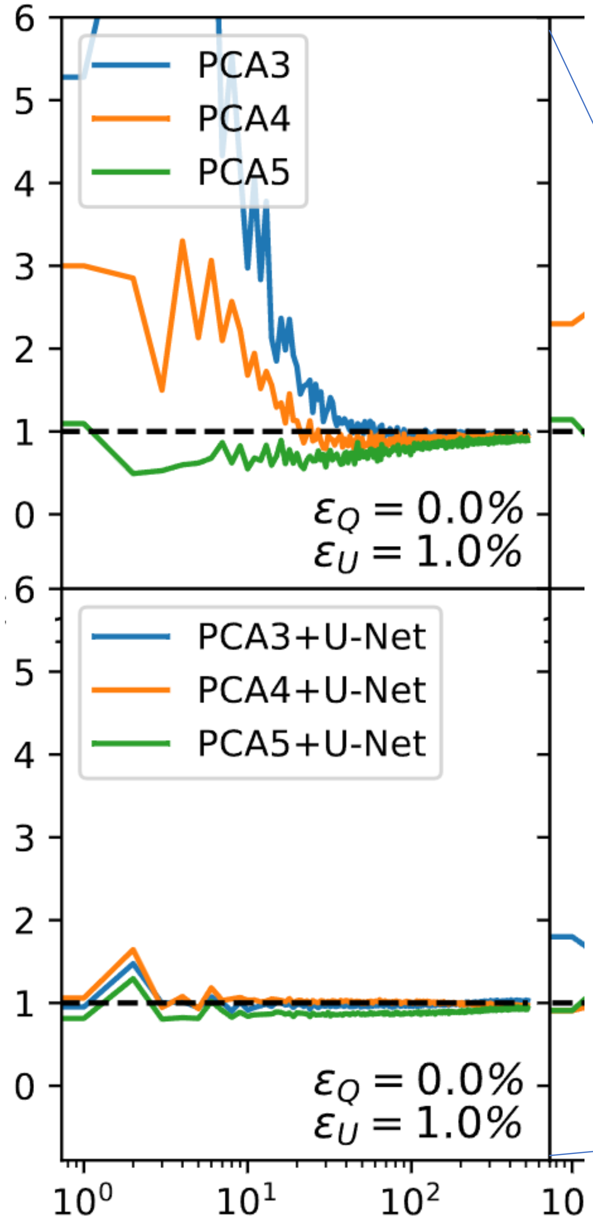


Figure 11. Comparison of the results of U-Net and PCA-3 power spectra. Panels (a) and (b) show the correlation statistic $r(\ell)^2$ and cleaning error $T(\ell)$ for the PCA and U-Net power spectra. Here PCA represents PCA-3 and U-Net represents the use of PCA-3 plus U-Net.



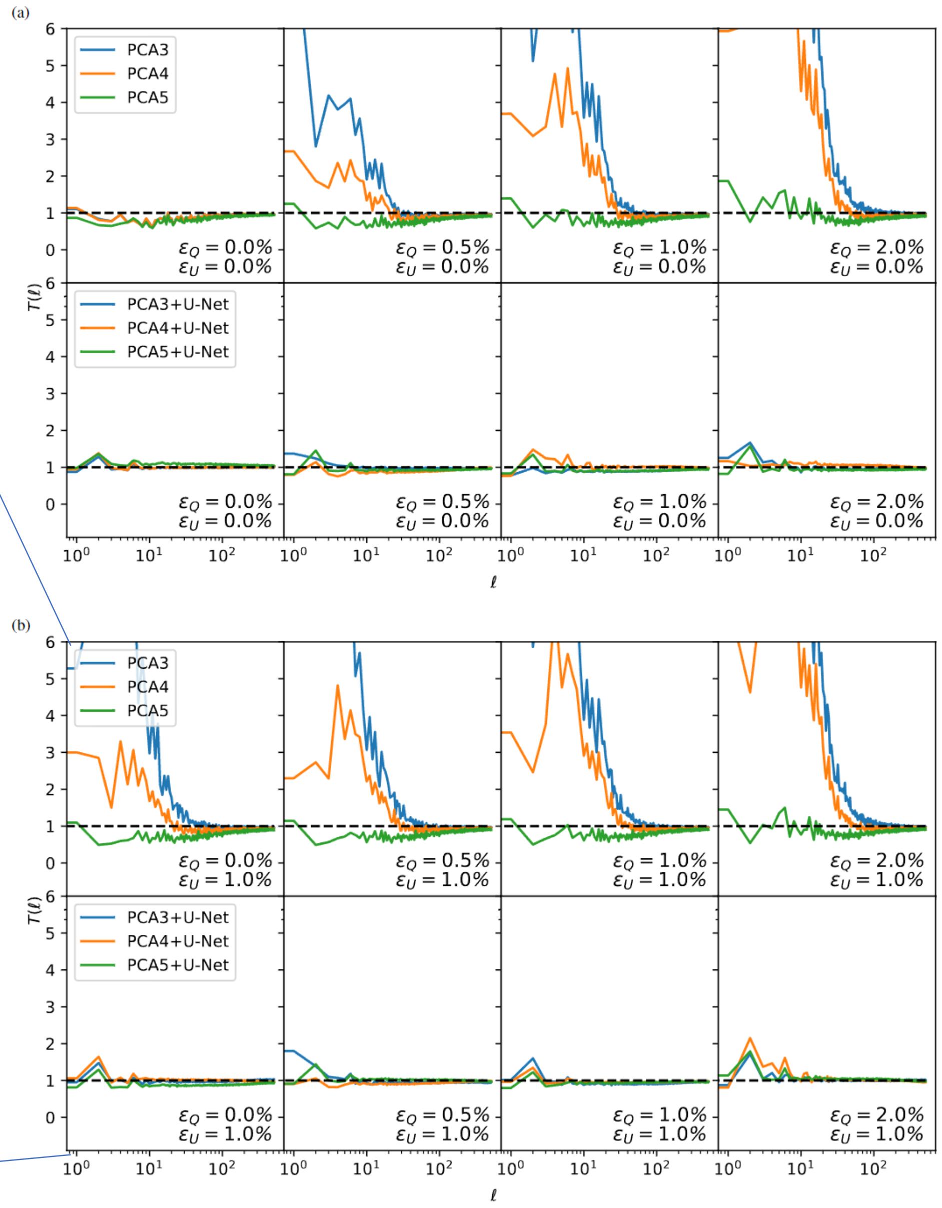


Figure 12. Results of the cleaning errors $T(\ell)$ obtained using U-Net for simulated sky maps of 8 different $[\epsilon_Q, \epsilon_U]$ after PCA-3, PCA-4, PCA-5 processes. The blue, yellow, and green lines represent the input as the result of subtracting 3, 4, 5 modes from the sky maps using PCA, respectively. The top of each of the two plots is the result of PCA and the bottom is the result of PCA+U-Net.

Robustness Analysis

test sets： {0.01, 0.1, 1, 10,100}

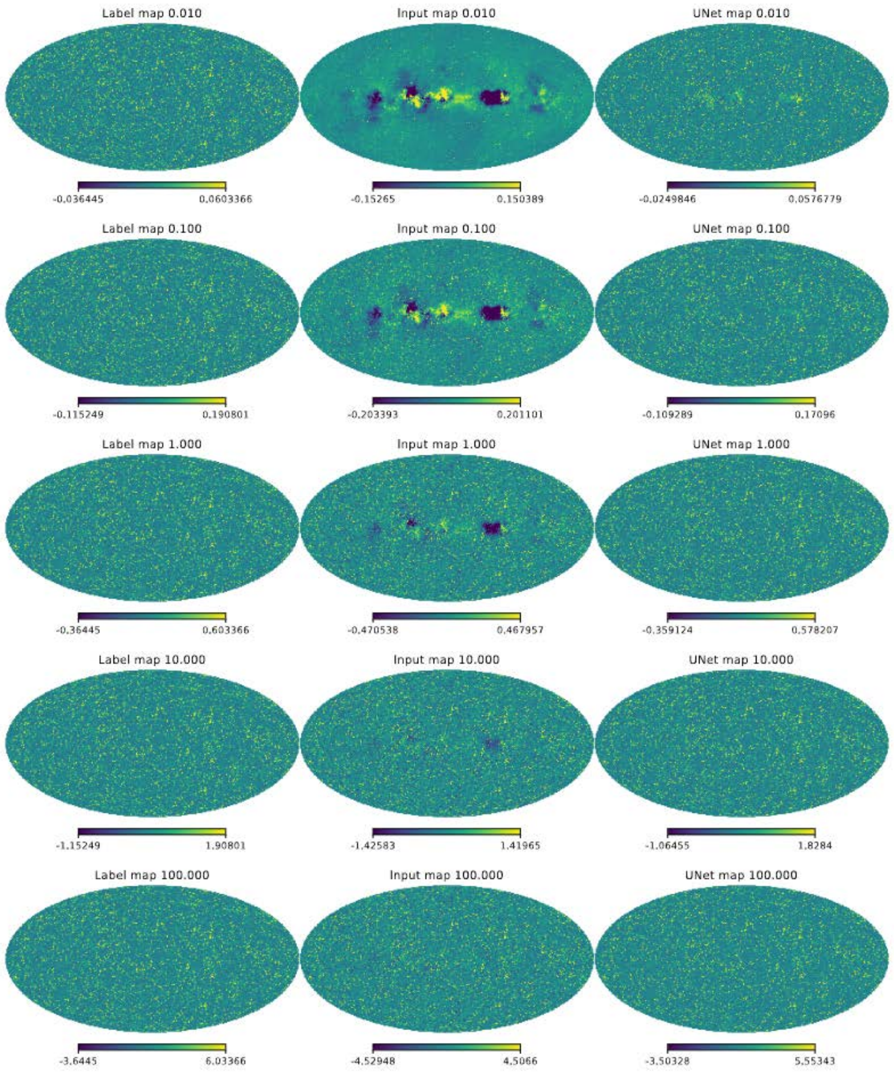


Figure 12. Sky map results of robustness testing. From top to bottom, the 21cm signals of the sky maps used for the test set were multiplied by 0.01, 0.1, 1, 10, and 100. From left to right, they are label maps, input maps, and the results from U-Net. The unit of each pixel of the sky maps is mK.

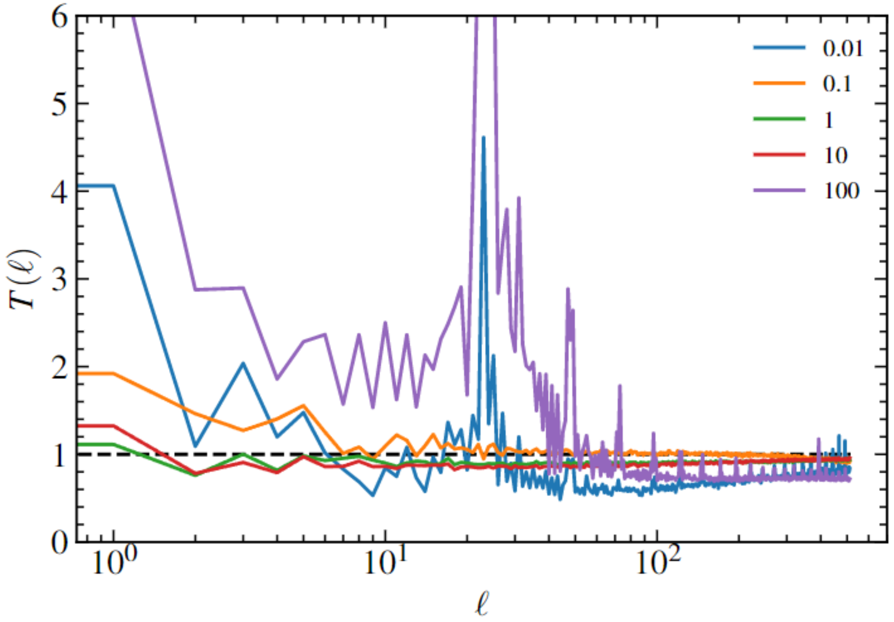


Figure 13. The cleaning error $T(\ell)$ results of PCA-4 + U-Net where $(\epsilon_Q, \epsilon_U) = (0.5\%, 1.0\%)$. The blue, yellow, green, red, and purple lines represent the HI signal of the test sets as 0.01, 0.1, 1, 10, and 100 times the original HI signal, respectively.

Summary

1. PCA foreground subtraction approach produces significant foreground residual.

2. The U-Net architecture can successfully remove the foreground at different polarization leakage levels.
3. U-Net can compensate for the signal loss caused by the aggressive PCA.
4. The U-Net subtraction strategy is reliable when the amplitude of 21cm signal varies within [0.1, 10].