## SKA Data Challenge 3 **SERENET: SEgmentation and Recover Network**



Michele Bianco **Tianyue Chen** Shreyam Krishna Arpan Das

**Rohit Sharma** 



EPFL

Sambit Kumar Giri



**Philipp Denzel** 

Hatem Ghorbel Edouard Goffinet Hes.so Massimo De Sanctis



Viraj Nistane **Christopher Finlay** 



Andrei Mesinger David Prelogović



Fri, 08 Sep 2023

SwissSKA Day 2023

#### Cosmology with Radio Interferometric Experiments



# Square Kilometre Array (SKA-Low)



Final layout of the SKA-Low stations and start of the construction in the early months of 2023.

- **512 stations**, each with **128 antennas**
- Maximum baseline ~65 km, station diameter 35 m
  - FoV of 10 deg with 16 arcsec resolution
- Frequency range between 50 to 350 MHz





# Global Evolution of the 21-cm Signal



#### Constrains on the 21-cm Power Spectrum



## Tomographic Imaging of the 21-cm signal

**SKA1-Low** tomographic images of redshifted 21-cm signal challenges:

- Instrumental noise (signal ~ 5 K)
- Foreground emission (signal ~ 1 - 1000 K)
- Antennas gain errors
- Ionospheric refraction effects
- Radio frequency interference
- And more ...



## Tomographic imaging of the 21-cm signal

Probe reionization process by observing the redshifted 21-cm signal

 $\delta T_b \propto (1+\delta) x_{\rm HI}$ 

Dec

RA

Vobs

1,

#### Square Kilometre Array (SKA1-Low): Images sequence of redshifted 21-cm signal

at different observed frequencies.

3D tomographic dataset or a.k.a. 21-cm lightcones

## Tomographic imaging of the 21-cm signal

Probe reionization process by observing the redshifted 21-cm signal

 $\delta T_b \propto (1+\delta) x_{\rm HI}$ 

#### Square Kilometre Array (SKA1-Low):

Images sequence of redshifted 21-cm signal at different observed frequencies.

3D tomographic dataset or a.k.a. 21-cm lightcones



z = 13.2 $v_{obs} = 100 \text{ MHz}$ 

# Goal of the SDC3

Currently we can create EoR mock observation with a combination of numerical models for 21-cm, systematic noise and foregrounds.



#### <u>Goal:</u>

Recover 21-cm differential brightness and the distribition of neutral hydrogen from SKA-Low mock observations.

## SERENEt

## SEgmentation and REgression NEtwork

Combine the predicted binary maps of **SegU-Net** as additional input of **Rec-Unet** training step in order to include prior in the network training.



SKACH 840k hybrid-h allocation projects at Pitz Daint @ CSCS

# Galactic & Extra-galactic Foregrounds

Galactic and extra-galactic **foregrounds** have **frequency smooth flux** compared to the 21-cm signal.

From 2D power spectra, remove k-modes contaminated by foreground as:

- avoidance technique
- model for substraction



#### 



## SERENEt SEgmentation and REgression NEtwork





#### SegU-Net: Tomographic Data & Reionization History

 $\delta T_{b}$  [mK]

-20

0.8

0.4

0.2

0.30

0.25

0.20

0.15

0.10

0.05

0.00



# SegU-Net: HI size distribution

The Island Size Distribution (ISD): statistical distribution of HI regions during EoR (Giri+ 2018)

$$\overline{R}_C(z) = \int_{R_{\min}}^{\infty} R \, \frac{dP}{dR}(z) \, dR$$

SegU-Net results:

$z_c$	pre-process	$\overline{R}_C$ [cMpc]
8.24	Ground Truth	29.54
	all z PCA	$31.37^{+3.09}_{-3.93}$
	PCA	$27.65^{+9.13}_{-6.12}$
	Wedge	$15.20^{+24.13}_{-6.18}$
	GPR	$29.14^{+5.26}_{-4.89}$
	Polynomial	$29.21^{+5.83}_{-5.21}$



#### RecU-Net: Recover 21-cm with U-Net (Bianco+ in prep.)



U-Net architecture with intercepting convolution block to process the binary prior map from SegU-Net

# SERENEt: Recover of 21-cm Signal

Recovered 21-cm signal for EoR for lightcone subvolume centered at redshift z = 8.25 ( $x_{HI} \sim 0.5$ ) and  $\Delta v = 20$  MHz on PCA pre-process images





# SERENEt: Recover of 21-cm Signal

Ground Truth at z = 7.27

**RecU-Net** (no prior)  $R^2 = 64.2 \%$ 

SERENEt(**GT**) +  $x_{H}^{TRUE}$ R<sup>2</sup> = 92.9 %

SERENEt(GT) +  $X_{H}^{PRED}$ R<sup>2</sup> = 82.9 %



# SERENEt: Recover of 21-cm Signal

Coefficient of determination (R<sup>2</sup> score) redshift evolution to quantify the foreground mitigation:

- SERENEt(GT) + X<sub>H</sub>TREU upper limit based on best prior binary map (ground truth)
- SERENEt(GT) + X<sub>H</sub><sup>PRED</sup> next best results when compared to RecU-Net



### SERENEt: Comparison with Current Data



# **Conclusion & Discussion**

- U-Net have the potential to recover 21-cm signal with prior information
  - Need more realistic mock observation
- At SKACH, working on simulation pipeline for realistic mock observation
  - Interferometer systematic
  - DD & DI gain error
  - Update extra-galactic point sources foreground model
  - Machine learning implementation of SERENEt pipeline
- Participation at the SKA Data Challenge (SDC3) with SERENEt pipeline (deadline September 30th)