SKA Blind Challenge (Pilot)

Description of SKA layout:

The simulation uses 224 SKA core stations with random positions spread out upto a diameter of ~1 km. Each station has 256 dipole antennas, positioned randomly within a diameter of 38 m.

We assume an SEFD of 2500 Jy at 115 MHz, and compute the SEFD at other frequencies via the relationship SEFD(ν) = SEFD(115 MHz)x(115 MHz/ν)$^{0.55}$. A station has a collecting area of 38 m diameter, with a filling factor of 1, so the effective collecting area is the same as the physical collecting area. Using the equation $SEFD = \frac{2kT_{sys}}{A_{effective}}$, this yields $T_{sys} = 871.5\,K$.

Description of the simulation:

The data covers a 10 deg x 10 deg field of view. It is over the frequency range 115 - 125 MHz. The data thus has a bandwidth of 10 MHz (= $\Delta\nu$). The channel width is 0.5MHz (= $d\nu$).

The simulated data covers an observation period of t_{sim} = 4 hours, with 10 second(= dt)-averaged snapshot data recorded every 30 min (= t). Thus, the data consists of 8 snapshot observations, with the snapshots being spaced 30 min apart. The data is simulated in this manner in order to keep the data size manageable. Since the SKA has excellent instantaneous uv-coverage, this is permissible and the data sets are representative of real observations. We simulate the real observations over longer time periods by appropriately scaling the station SEFD (assumed to be $SEFD_{stn}$ = 2500 Jy) to obtain the effective SEFD:

$$SEFD_{stn,eff} = \frac{SEFD_{stn}}{\sqrt{t/dt}}$$

and the corresponding noise per visibility per baseline as:

$$RMS = \frac{SEFD_{stn,eff}}{\sqrt{2\,dt\,d\nu}}$$

To simulate shallow, medium, and deep surveys with total observation times t_{tot} = 10, 100, 1000 hours, we scaled the noise to the effective noise for these longer observation times:

$$RMS_{eff} = \frac{RMS}{\sqrt{t_{tot}/t_{sim}}}$$

Description of the data:

The data are provided as images (in FITS file format) and visibilities (in measurement set format).
The datasets are Stokes-I only data (no polarization for now). The data are provided for 3 different total observation times - 10, 100, 1000 hours.

Image datasets:

21 images are provided, covering the frequency range 115-125 MHz, with channel width 0.5 MHz. Each single-frequency image is ~366 KB in size, for a total size of ~7.7 MB for the 21 images.

Field of view: 10 deg x 10 deg
Image size: 300 pixels x 300 pixels
Pixel size: 2 arcmin

3 datasets (images) are provided:

- EoR + noise
- EoR + FG
- EoR + noise + FG

as well as the following 2 auxiliary datasets (images):

- Noise
- PSF associated with each image

Visibilities datasets:

The data is also provided as visibilities, in measurement set format. The data covers the frequency range 115-125 MHz, with channel width 0.5 MHz. Each measurement set is ~20 MB in size.

3 datasets are provided:

- EoR + noise
- EoR + FG
- EoR + noise + FG

as well as the following auxiliary dataset:

- Noise

Data files:

The data files are on Dropbox. These include images and visibilities, as well as beam models and telescope layout.

Desired format of submitted results:

The following power spectra of the EoR, FG, and noise signals should be submitted in a text data file:

2D (cylindrically averaged) power spectrum:

1) The file should have 4 columns: $k_{//}$ (in *hc*Mpc^{-1}), $k_{\perp}$ (in *hc*Mpc^{-1}), P (in $K^2 h^{-3} cMpc^3$), P_{err} (in $K^2 h^{-3} cMpc^3$).
2) The values of $k_{//}$ should be 10 equally-spaced points between 0.0459 and 0.459 ((in *hc*Mpc^{-1}, both endpoints inclusive).
3) The values of $k_{\perp}$ should be 25 equally-spaced points between 0.0474 and 0.2288 ((in *hc*Mpc^{-1}, both endpoints inclusive).
4) The bin size is the difference between two consecutive $k_{\perp}$ points / $k_{//}$ points, centered on the values of $k_{//}$ and $k_{\perp}$ specified in list items 2) and 3) above. Thus, the bin size is 0.0459 in $k_{//}$ and 0.00756 in $k_{\perp}$.

3D (spherically averaged) power spectrum:

1) The file should have 3 columns: k (in *hc*Mpc^{-1}), Δ^2 (in mK^2), Δ^2_{err} (in mK^2).
2) The edges of the bins in k are 10 points equidistant in log-space between 0.068 and 0.5 ((in *hc*Mpc^{-1}, both endpoints inclusive).
3) The k value in the first column in 1) above should be the mean of the k values in that bin.

Power spectra should be extracted from each of the 3 datasets (EoR + noise, EoR + FG, EoR + noise + FG), and for each of the values of the total observation time (10 hr, 100 hr, 1000 hr). Errors should include thermal noise and sample variance (separately).