MillimeterDL

Deep Learning Simulations of the Microwave Sky

Dongwon 'DW' Han

Neelima Sehgal, Francisco Villaescusa-Navarro

arXiv:2105.11444

mmDL Simulations are now public at

[NERSC] https://portal.nersc.gov/project/cmb/data/generic/mmDL/
[Lambda] https://lambda.gsfc.nasa.gov/simulation/tb_sim_ov.cfm
[PySM3] https://github.com/galsci/pysm (coming soon!)

Stony Brook University

LSS x CMB Data Products & Experiments

LSS
Galaxies (photometric / spectroscopic), LRGs,
Quasars, AGNs
Galaxy groups / clusters
Weak lensing
Lyman-alpha forest
HI maps
Other line-intensity maps

×

CMB
Temperature
Polarization
Lensing
tSZ
kSZ

A Slide from Chihway's Talk from CMB-S4 Summer 2020 Meeting

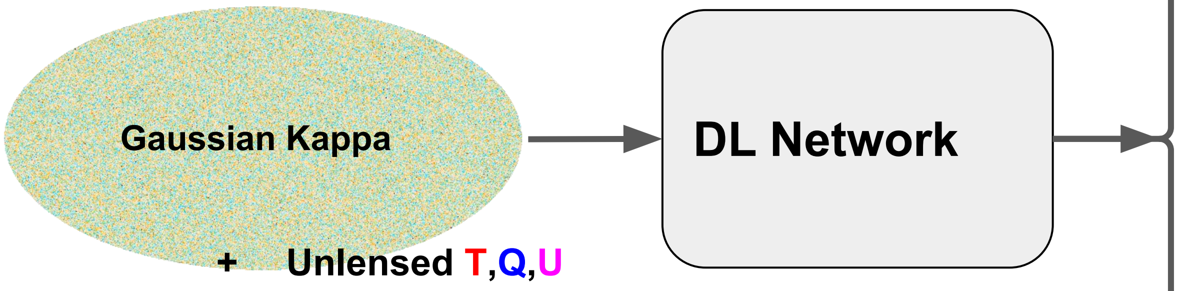
Requirements for Correlated Simulations

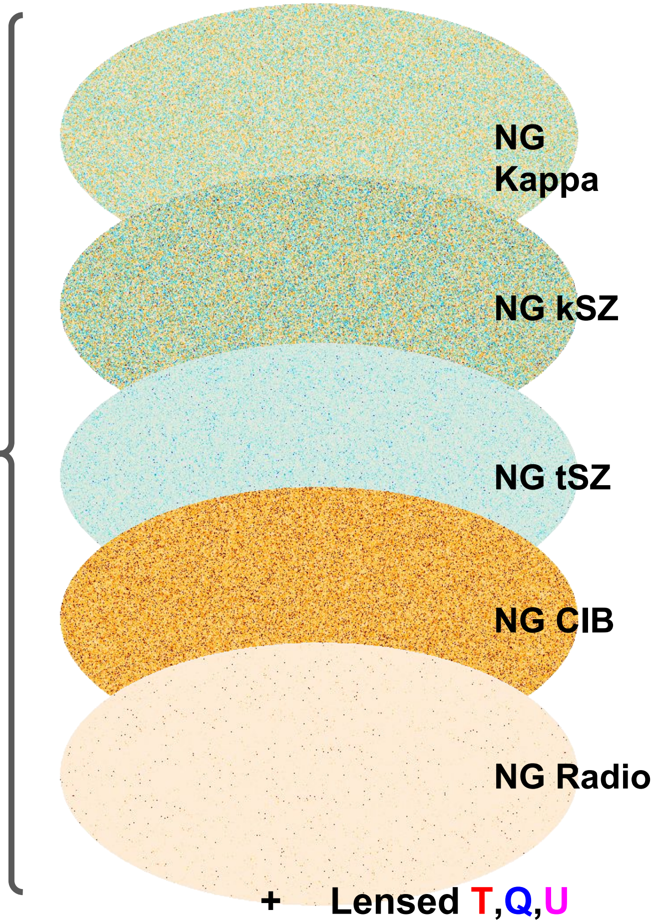
- Require many realizations (500 simulations or more) for statistical analysis
- Full-sky, high-resolution (half-arcminute), multi-frequency
- Contains non-Gaussian Information and correlations among different components.

A natural solution is to generate N-body simulations and post-process them (CrowCanyon, DC2, MDPL2, S10, Websky). Unfortunately, this approach is computationally expensive.

However, once you have just a single realization of such simulation, you can use a Deep Learning (DL) Generative Model to get around this restriction.

Rough Schematic

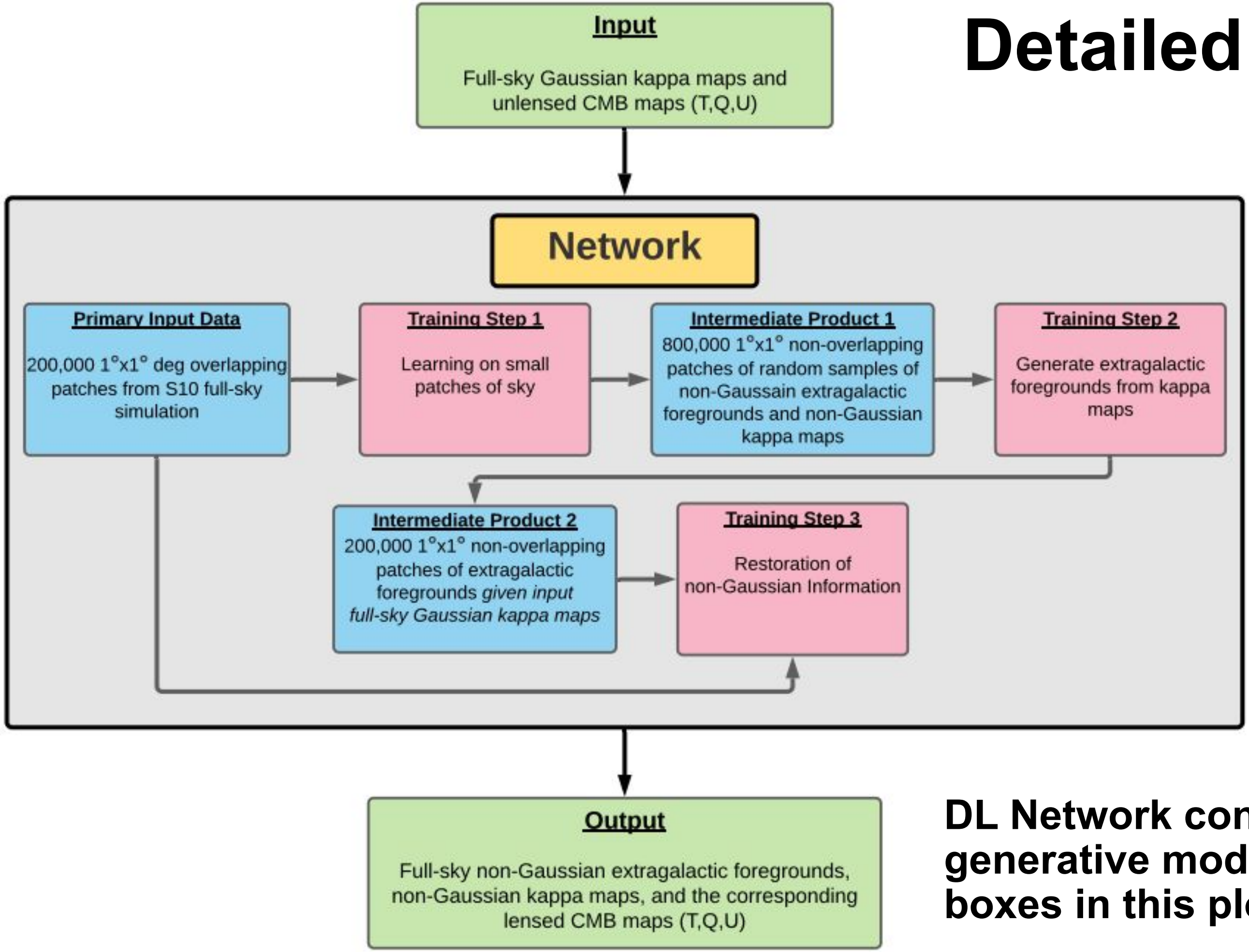




Using a Gaussian kappa map and unlensed T,Q,U as an input, predict (1) Non-Gaussian (NG) kappa, (2) NG kSZ, (3) NG tSZ, (4) NG CIB, and (5) NG Radio, (6) Lensed T,Q,U

Han et al 2021, arXiv:2105.11444

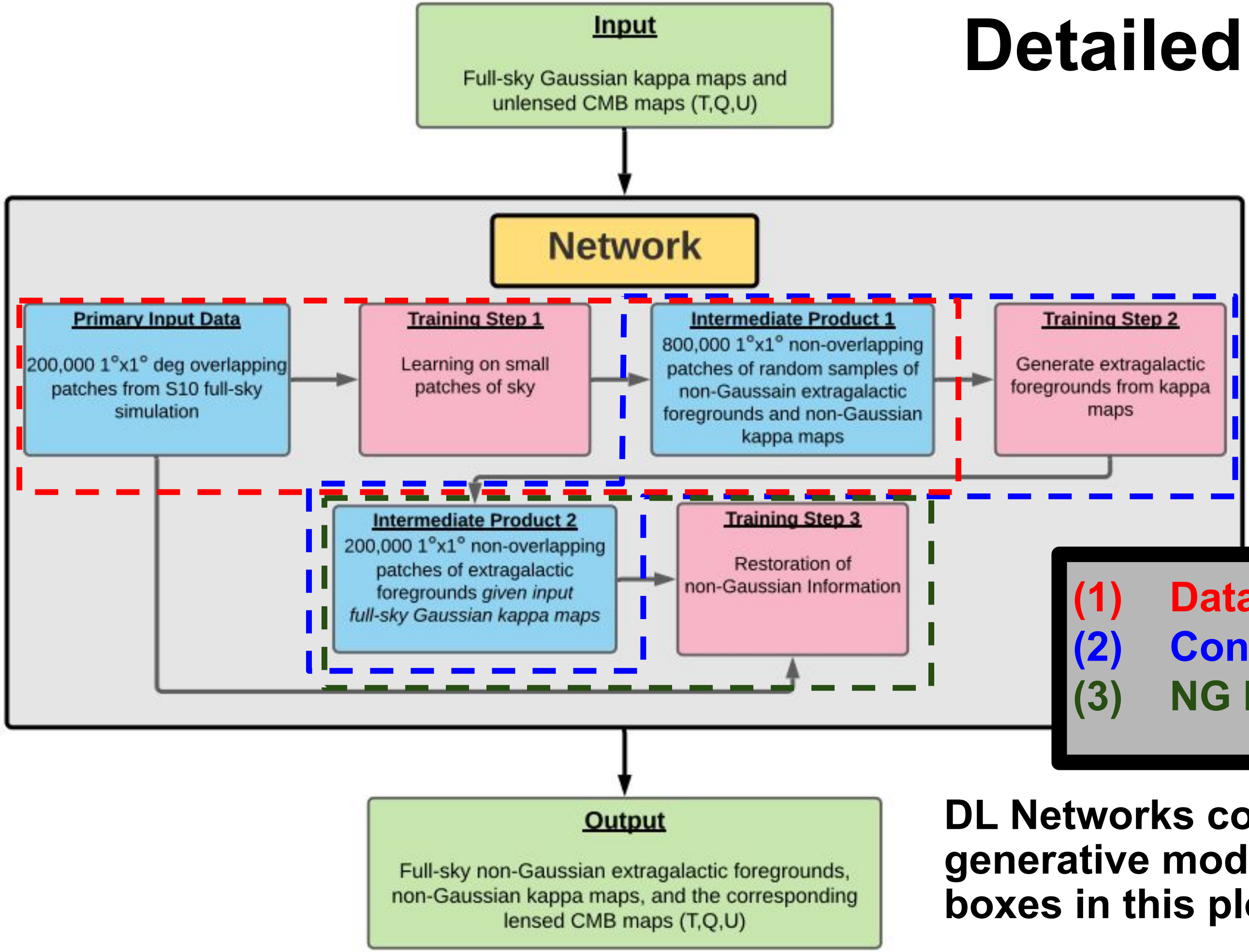
Detailed Schematic



DL Network consists of three layers of generative models represented by the pink boxes in this plot

Han et al 2021, arXiv:2105.11444

Detailed Schematic

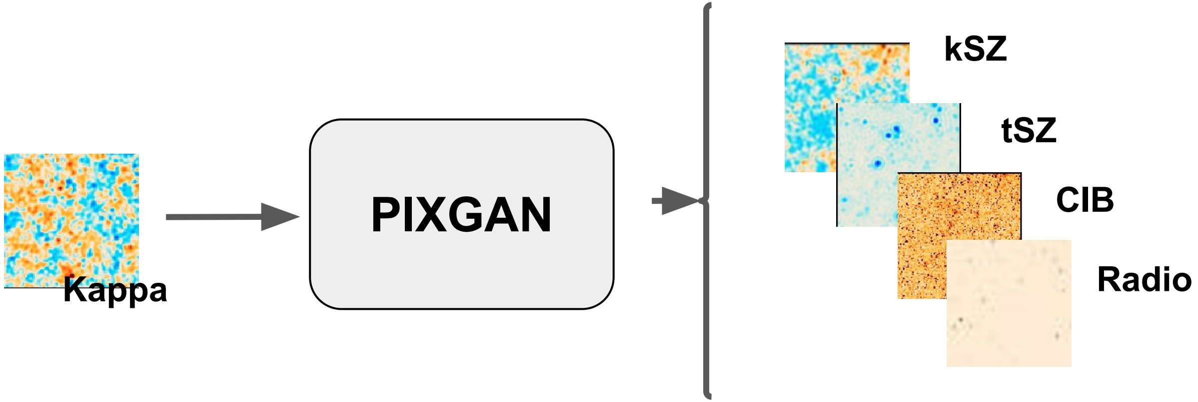


(1) Data Augmentation Step
(2) Conversion (Prediction) Step
(3) NG Restoration Step

DL Networks consists of three layers of generative models represented by the pink boxes in this plot

Han et al 2021, arXiv:2105.11444

Prediction Step (Step#2 on the schematic)

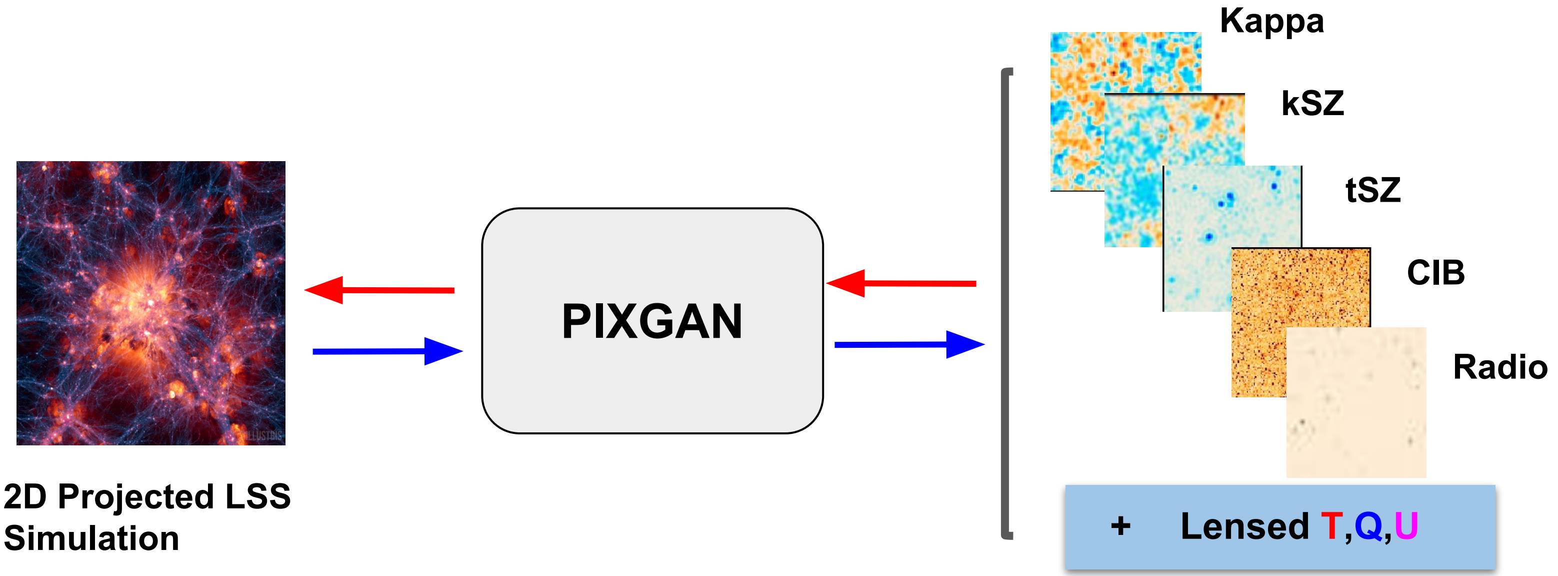


We train a conditional GAN (PIXGAN) to predict multiple fields (kSZ, tSZ, CIB, Radio maps) from kappa maps.

In principle, you can predict any field(s) from any input field(s) using this network.

Han et al 2021, arXiv:2105.11444

For Example: LSS Simulation to Correlated CMB Sims
(Or the other way around)

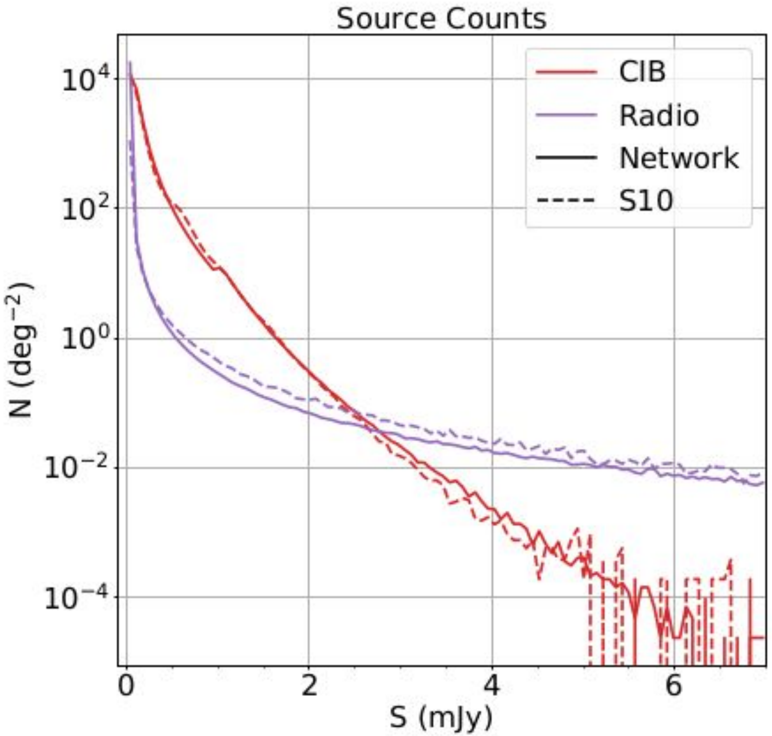


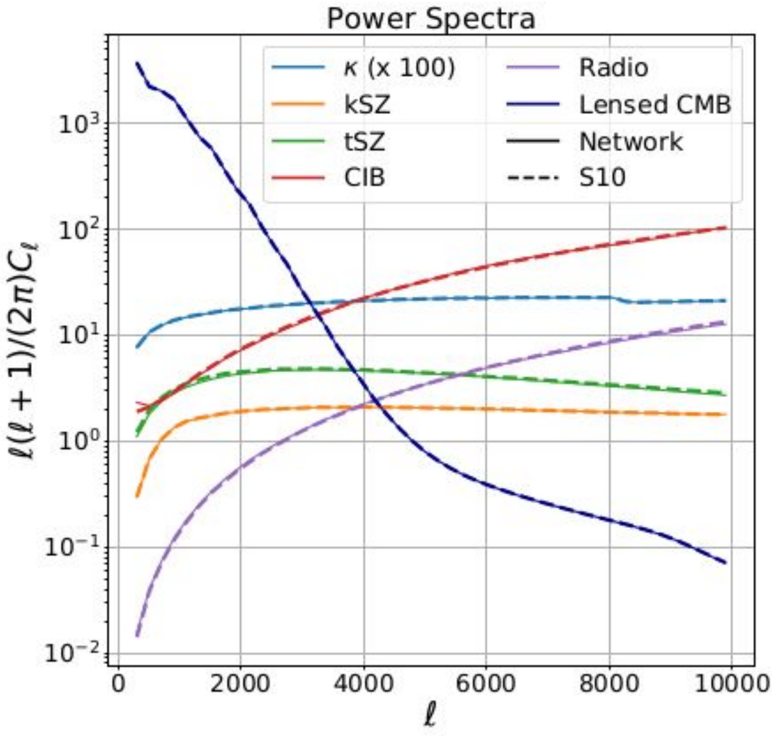
You can predict CMB fields from any 2D projected LSS simulation, and vice versa.

Han et al 2021, arXiv:2105.11444

LSS Image Credit : Illustris Collaboration

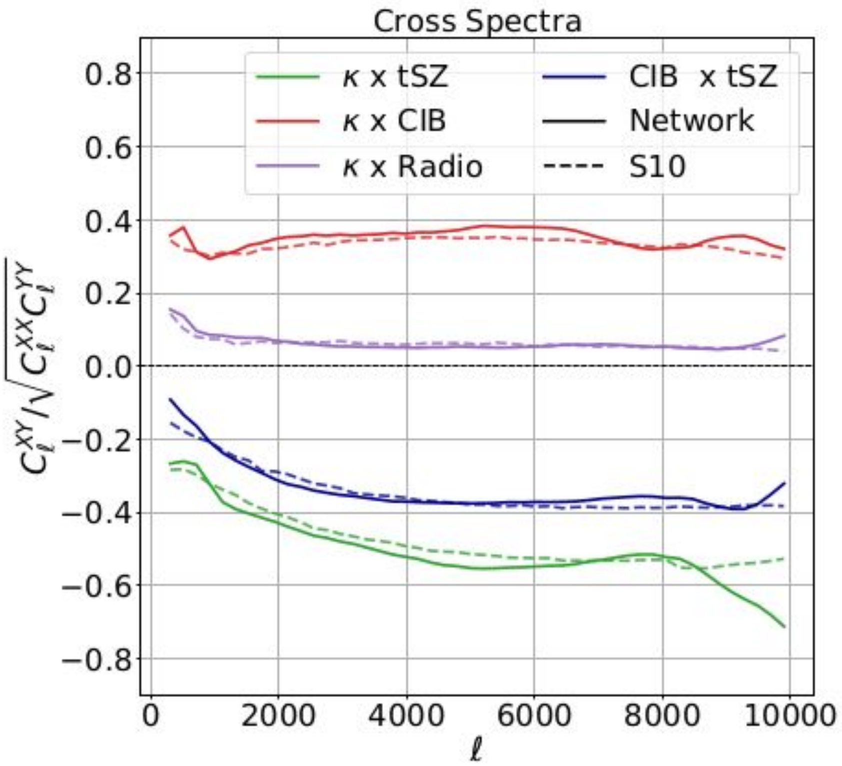
mmDL Results: Source Counts & Power Spectra

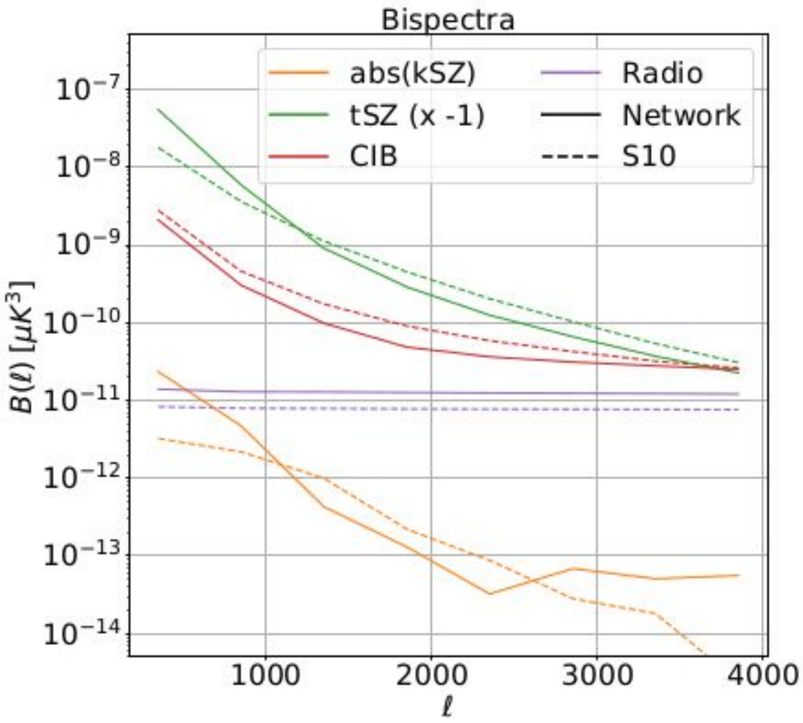




Han et al 2021, arXiv:2105.11444

mmDL Results: Cross-spectra and 3pt functions

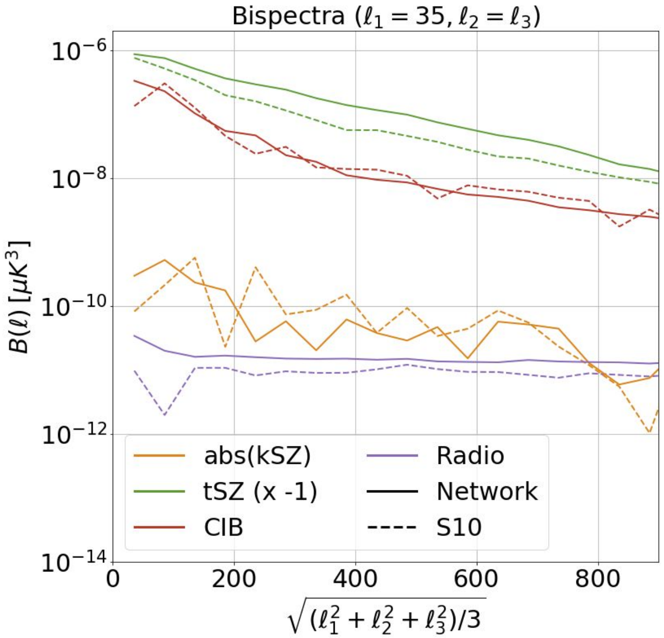


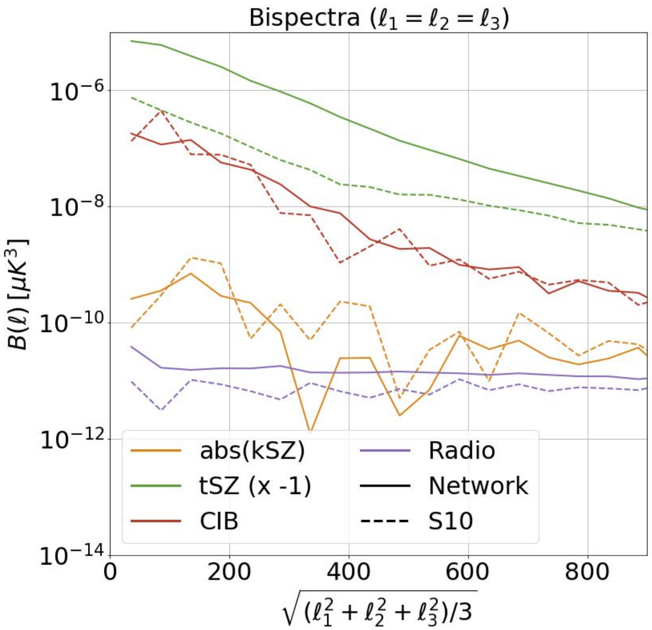


Han et al 2021, arXiv:2105.11444

mmDL Results: 3pt functions

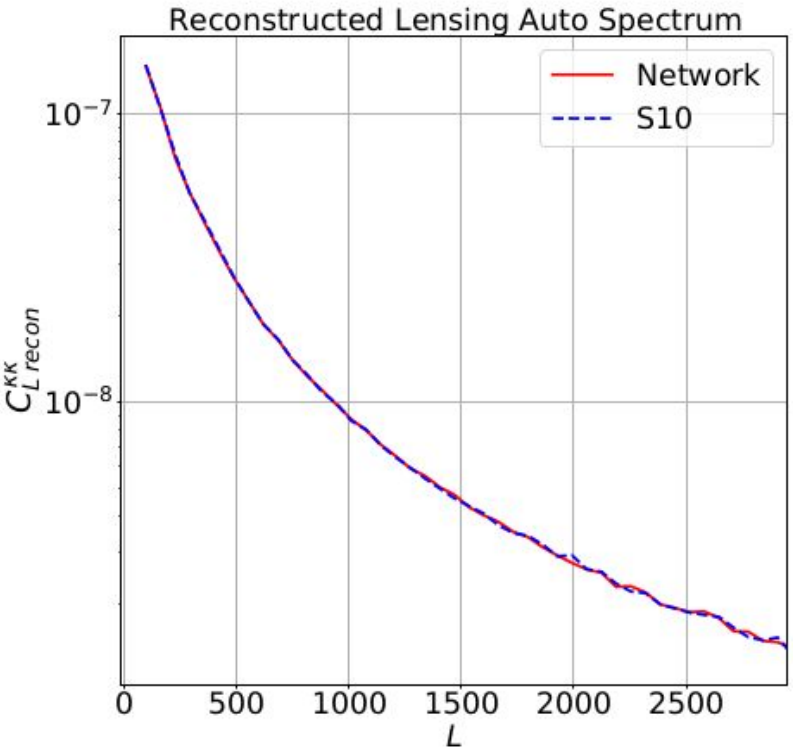
Han et al 2021, arXiv:2105.11444

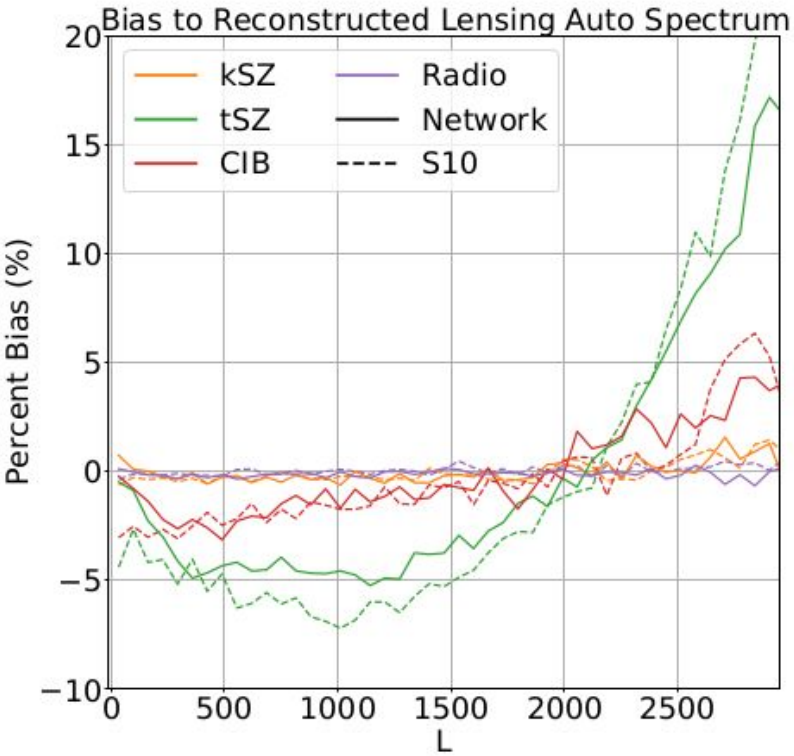




Since we "stitch" up 1x1 sq deg tiles to full-sky maps, you don't necessarily expect the network to preserve the correlations between large super-sample modes (ell < 200) and small scale modes (ell > 200). But our network can recover them.

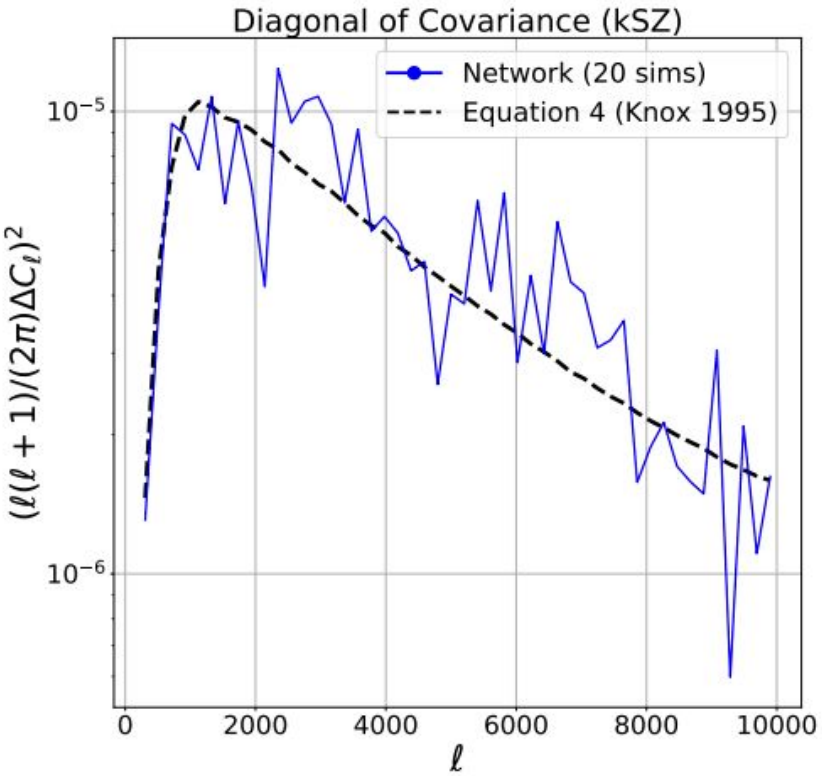
mmDL Results: 4pt functions

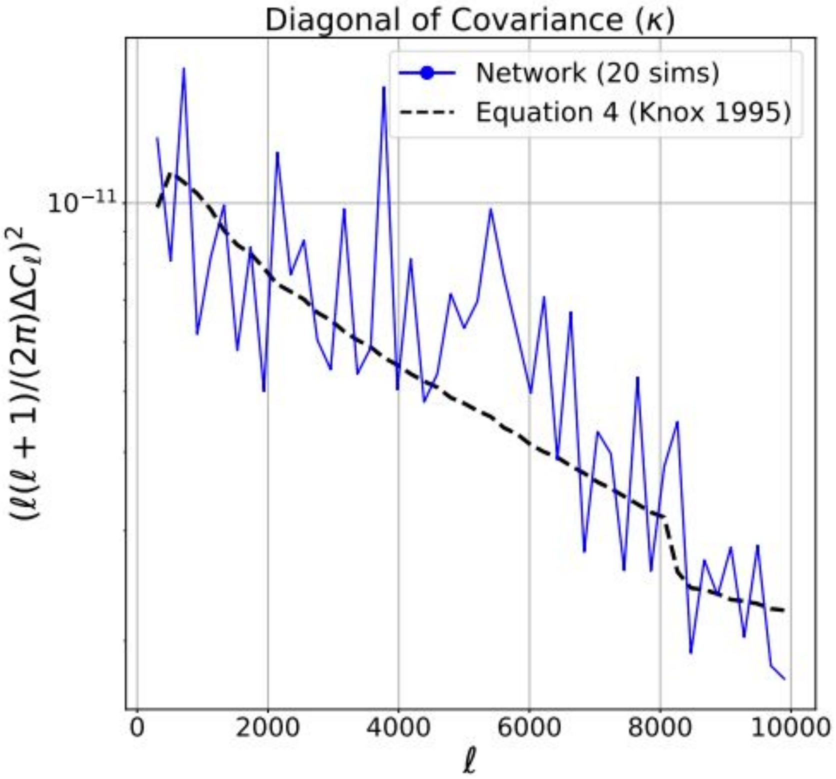




Han et al 2021, arXiv:2105.11444

mmDL Results: Variance of mmDL sims





mmDL Data Product Release

We have generated 500 mmDL realizations using our network. These are fullsky simulations at half-arcminute resolution at six different frequencies (30, 90, 148, 219, 277, and 350 GHz),

which include:

- the lensing convergence map (κ),
- the kinetic Sunyaev-Zelďovich effect (kSZ),
- the thermal Sunyaev-Zelďovich effect (tSZ),
- the Cosmic Infrared Background (CIB)
- the radio galaxies (Radio), and
- the lensed CMB (T,Q,U)

The simulations are available publicly at:

- **[NERSC]** https://portal.nersc.gov/project/cmb/data/generic/mmDL/
- **[Lambda]** https://lambda.gsfc.nasa.gov/simulation/tb_sim_ov.cfm
- **[PySM3]** https://github.com/galsci/pysm (coming soon!)

Han et al 2021, arXiv:2105.11444

Takeaways from mmDL project

- MillimeterDL (mmDL) can reproduce a wide range of non-Gaussian summary statistics.
- We can mass produce independent full-sky realizations from a single expensive full-sky simulation.
- We can “stitch” small patches up to make a full-sky realization that reproduces all the NG statistics.
- **Using our methods, we can take a full-sky lensing convergence map from any large-scale structure (LSS) simulation and generate the corresponding lensed CMB and correlated foreground components at millimeter wavelengths.**
- We can also adopt our procedure to quickly generate FG maps for forward-modelling.

Han et al 2021, arXiv:2105.11444

Deep Learning for Correlated Simulations (Cont.)

- Network is trained at a particular fiducial cosmology and with a particular baryonic model.

- Missing associated catalogues.

- In order to generate the simulations, the network must have learned all the relevant statistics from the fields (one-point, two-points, three-points, ….). Can we use this information to make “optimal” summary statistics?

Deep Learning for Correlated Simulations (Cont.)

- Network is trained at a particular fiducial cosmology and with a particular baryonic model.
  - **Given few simulations at different cosmology and byronic models (i.e. CAMELS simulations), one might be able to train a network to generate simulations at new cosmology.**
- Missing associated catalogues.
  - **A variation of our model can generate an associated catalogues along with fields.**
- In order to generate the simulations, the network must have learned all the relevant statistics from the fields (one-point, two-points, three-points, ….). Can we use this information to make "optimal" summary statistics?
  - **Bayesian NN for rescue? (Villaescusa-Navarro 2021, astro-ph/2011.05992)**