Removing Galactic foregrounds on CMB polarization maps using convolutional neural networks

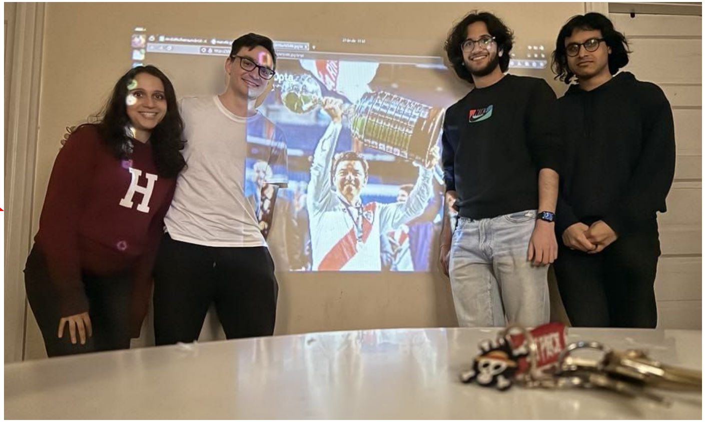
Lic. Luca J. Gómez Bachar

Advisor Dr. Daniel Supanitsky
Co-advisor: Dra. Cora Dvorkin

QUBIC
ITeDA
UNSAM
UNIVERSIDAD NACIONAL DE SAN MARTÍN
CNEA

A brief summary about me

- PhD in physics student at UNSAM, advised by Daniel Supanitsky
- Co-advised by Cora Dvorkin (Harvard University Professor)
- Member of QUBIC collaboration
- Member of CMB-S4 collaboration
- I spent six months working with Cora at Harvard University (from January-July)

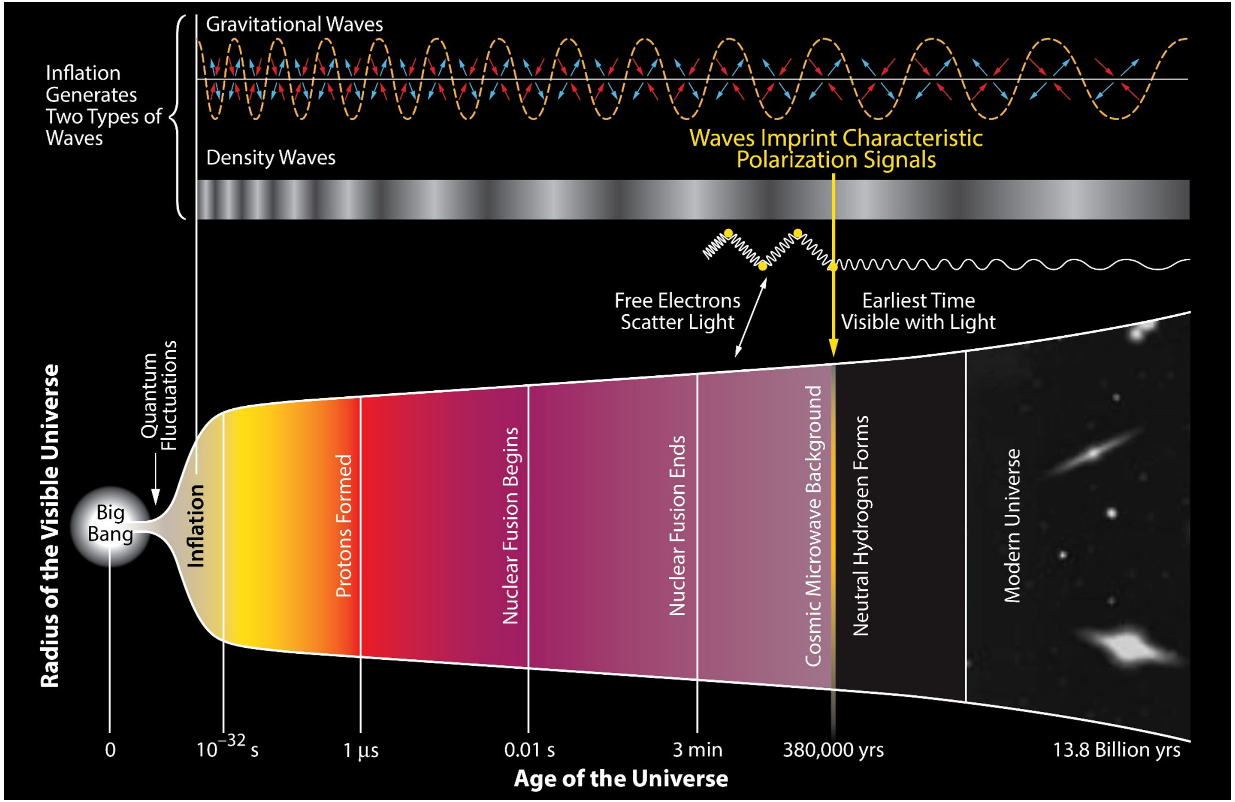


A brief summary about me

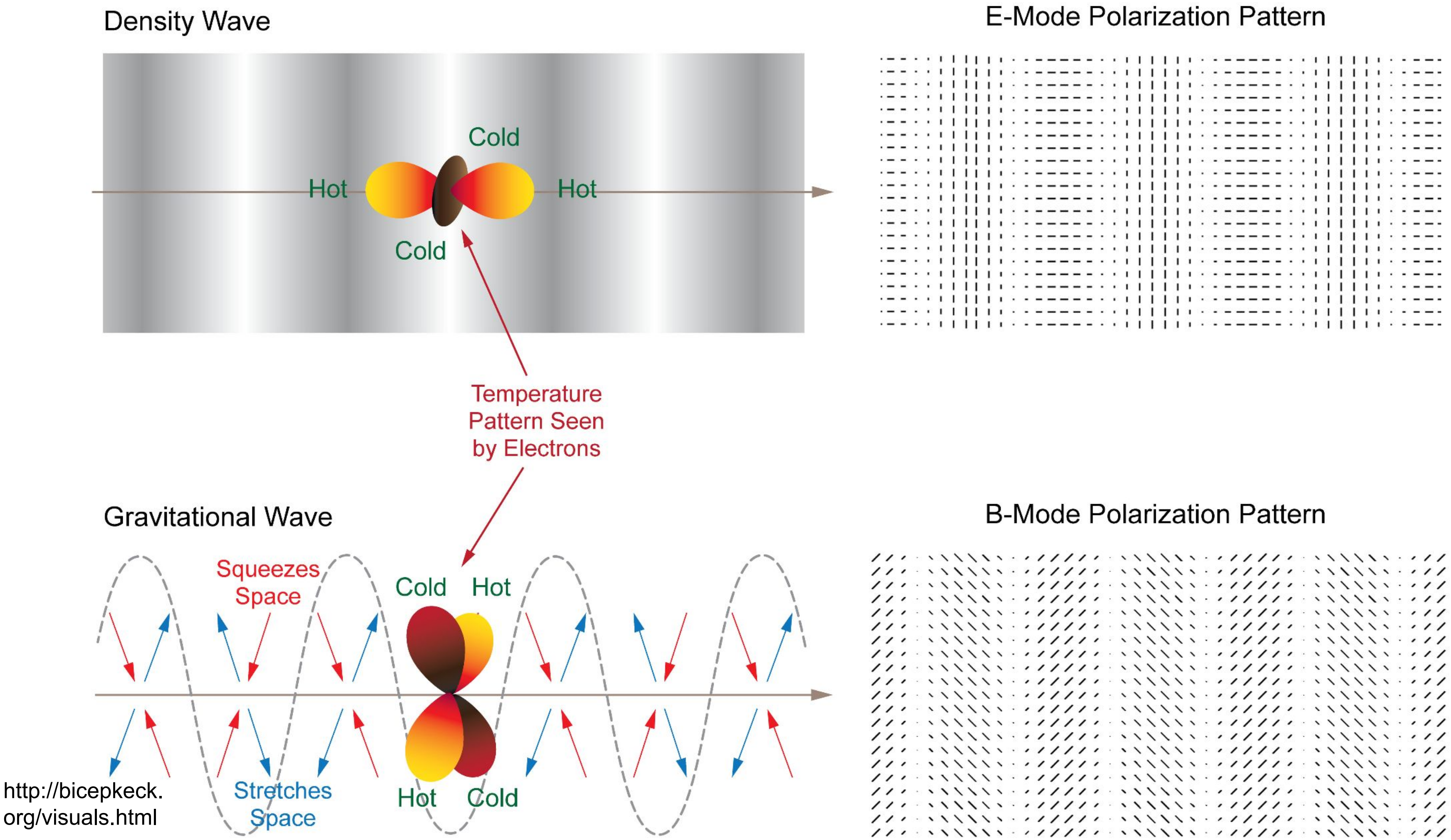
- PhD in physics student at UNSAM, advised by Daniel Supanitsky
- Co-advised by Cora Dvorkin (Harvard University Professor)
- Member of QUBIC collaboration
- Member of CMB-S4 collaboration

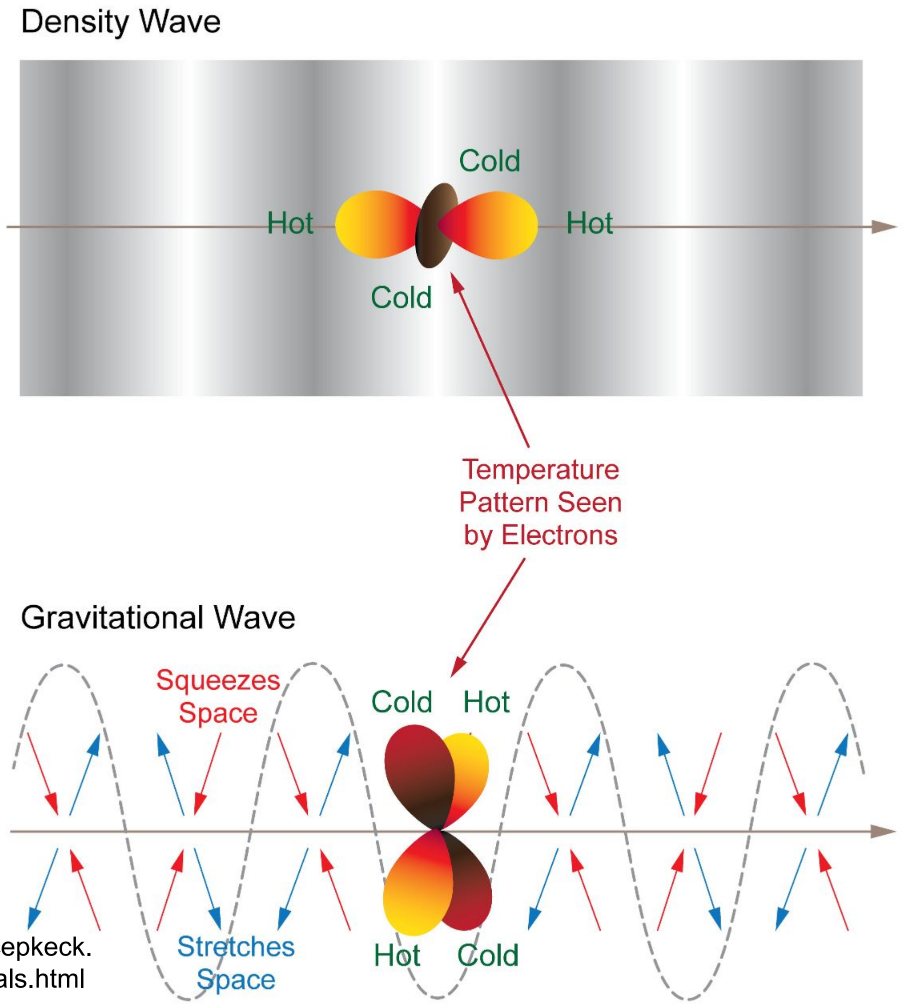
- Field hockey player & coach
- Huge fan of River Plate (futbol)

Primordial waves in the Universe

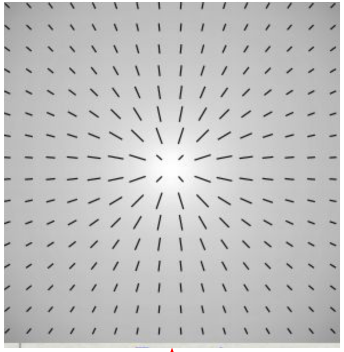


http://bicepkeck.org/visuals.html

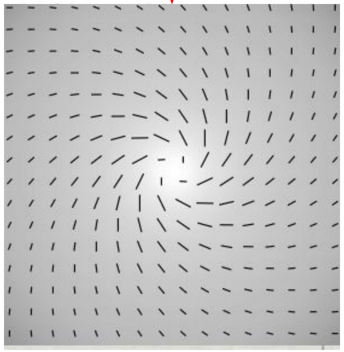

Density Wave
E-Mode Polarization Pattern
Cold
Hot
Hot
Cold
Temperature
Pattern Seen
by Electrons
Gravitational Wave
B-Mode Polarization Pattern
Squeezes
Space
Cold
Hot
Stretches
Space
Hot
Cold
http://bicepkeck.
org/visuals.html



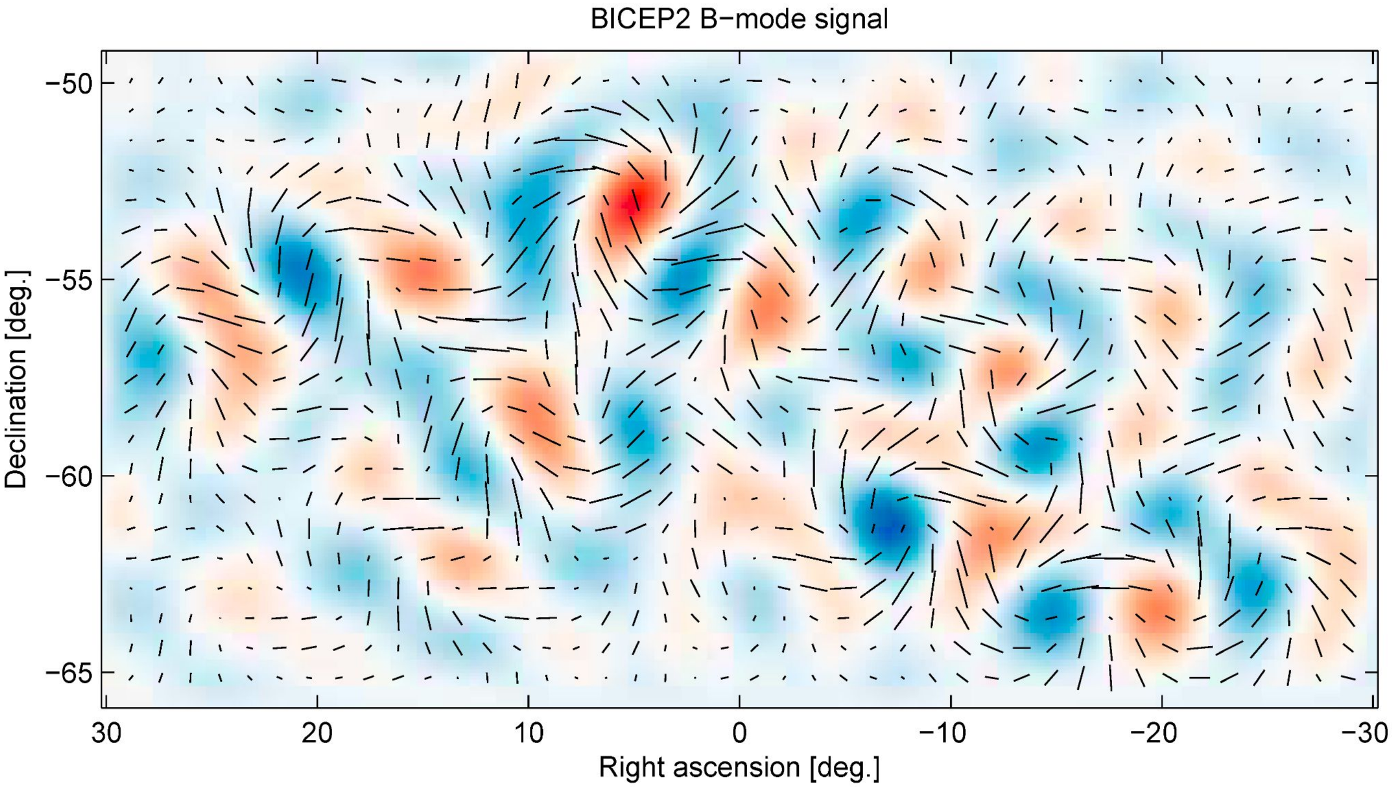
http://bicepkeck.
org/visuals.html



Adding several Fourier modes we obtain sky patterns like these

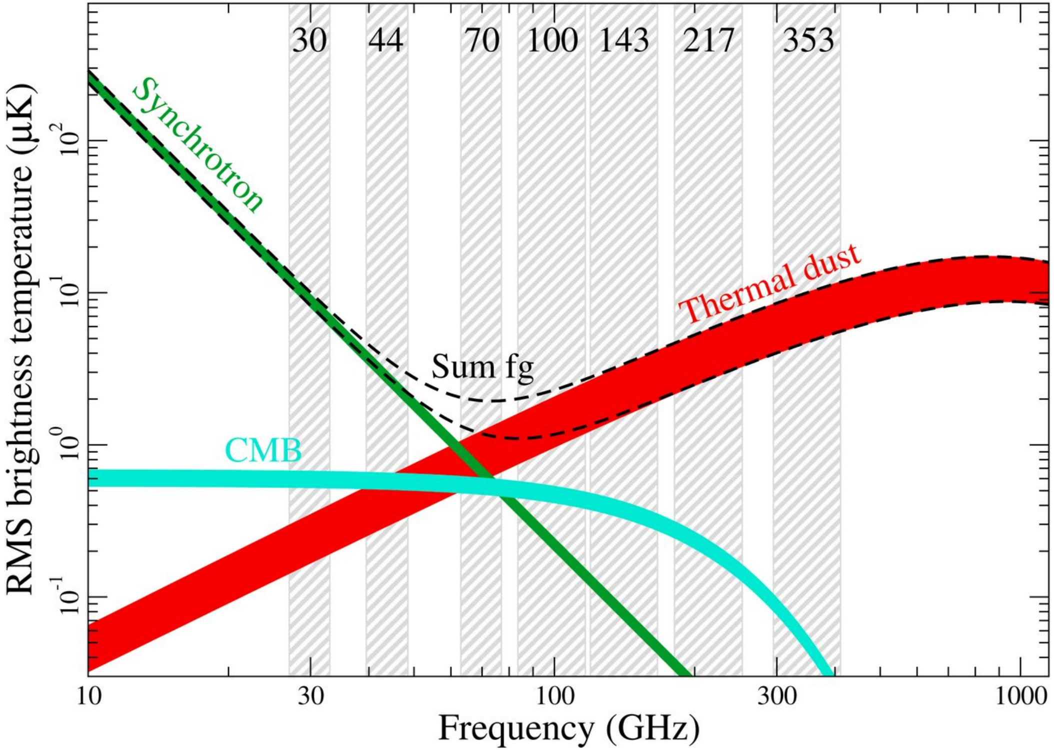


The best measurement so far



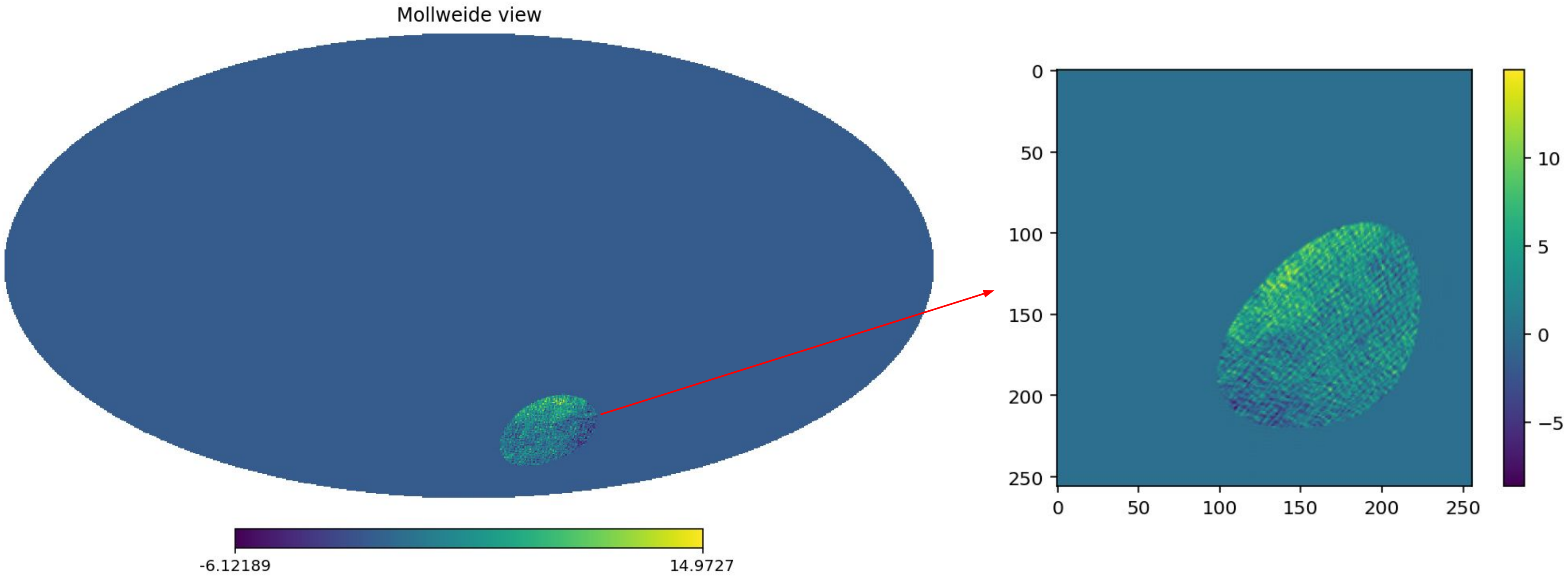
http://bicepkeck.
org/visuals.html

The Galactic foregrounds contamination

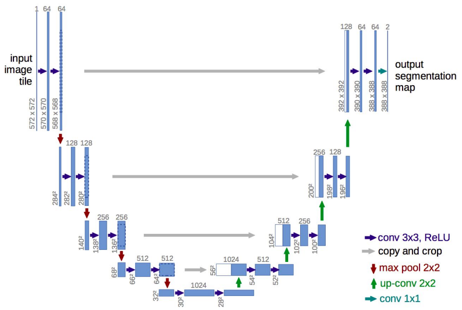


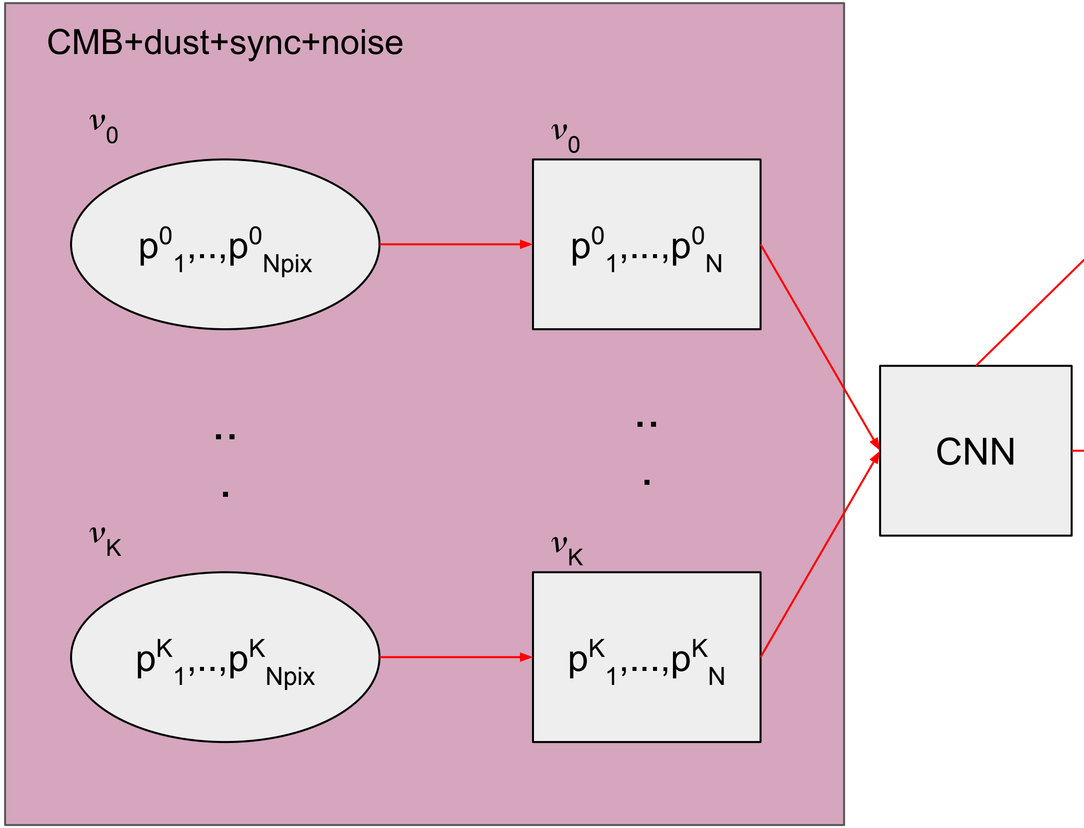
https://www.cosmos.esa.int/documents/387566/425793/2015_FGPolSpectra/b1813da7-ce91-439e-bb3d-cedaef18e20d?t=1423092019867

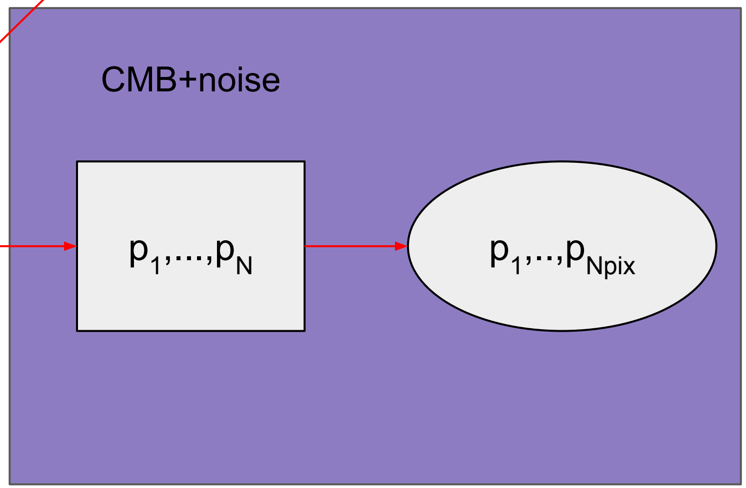
Our current research line



Our current research line



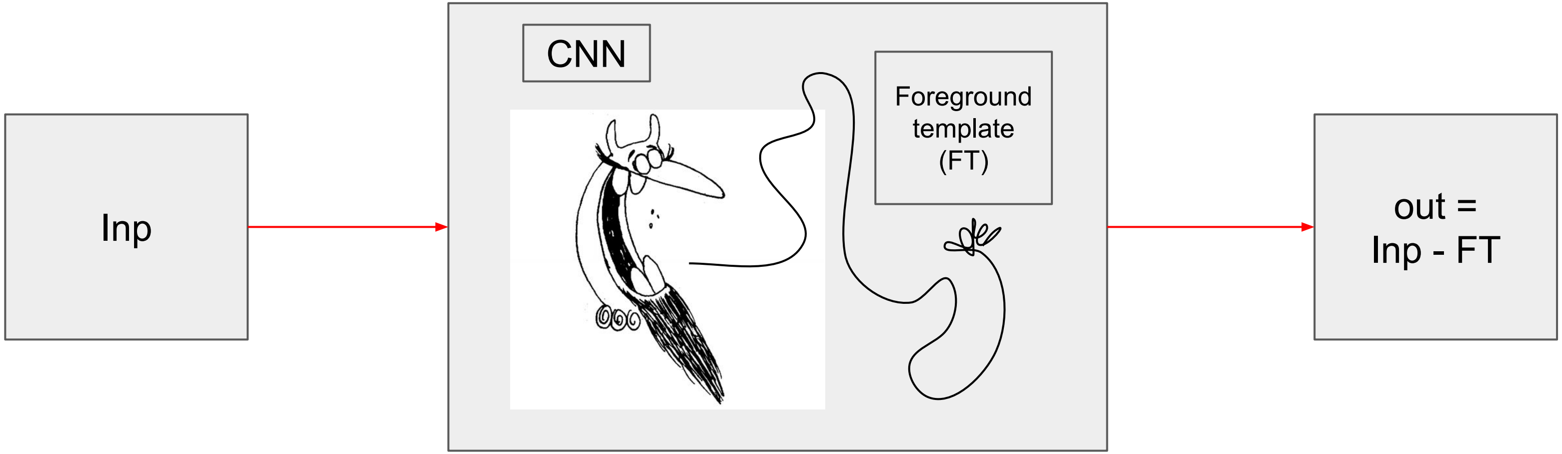




But the principal question of this research line:

How can we be sure that the CNN is not just subtracting a learned template?

The answer of this question relies in the generalization.



Pictorial representation of Maxwell's demon taken from a lecture given by Esteban Calzetta. I drew the elastic hand.

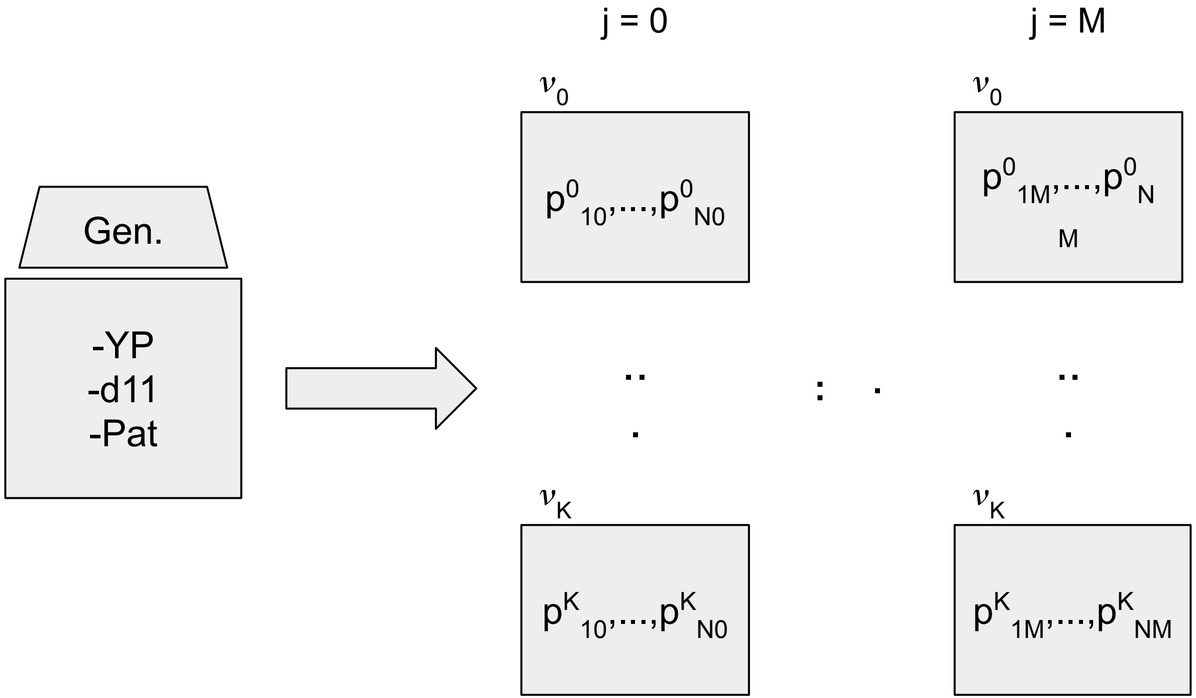
Template based foreground simulation generators

Gen. -YP → Gaussian fluctuations of the foregrounds parameters for each realization

Gen. -d11 → Small scale fluctuations based in a spectra for each realization

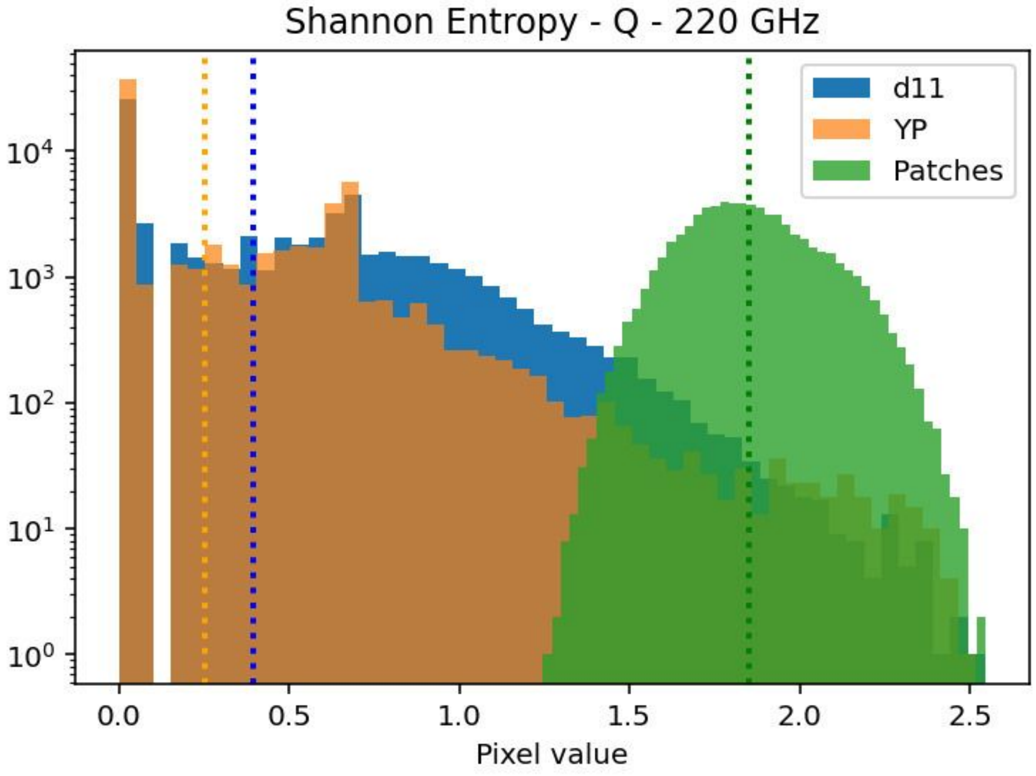
Gen. -Pat → Each realization is a different patch of the sky

Foreground simulation generators



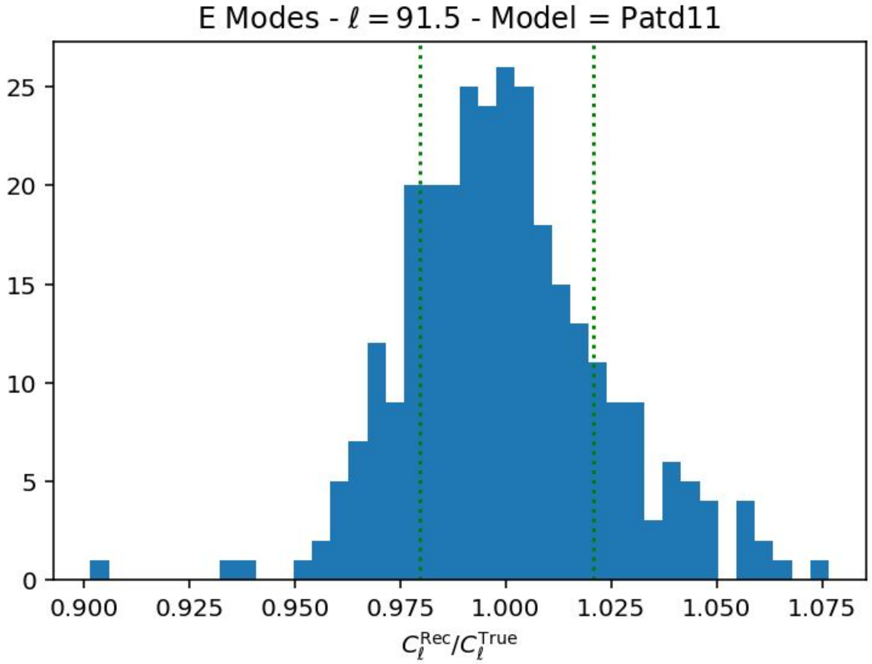
- Variance
- Skewness
- Entropy

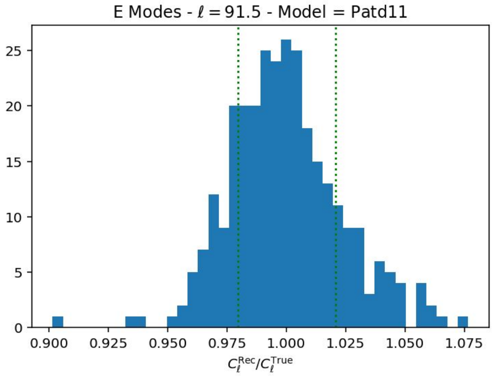
Shannon Entropy for our models



Generalization factor

$$\mathrm{rat}_\ell^{X,i}[\mathrm{M}_1, \mathrm{M}_2] = \frac{C_\ell^{\mathrm{Rec},X,i}[\mathrm{M}_1, \mathrm{M}_2]}{C_\ell^{\mathrm{True},X,i}[\mathrm{M}_2]}$$





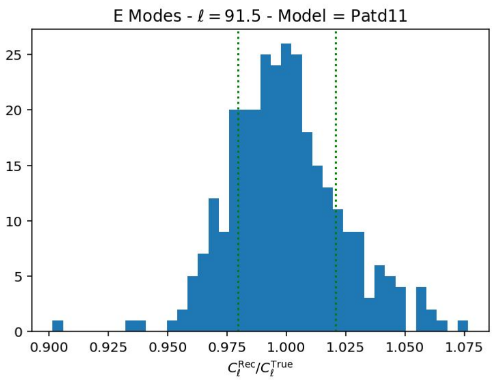
Generalization factor

$$\mathrm{rat}_\ell^{X,i}[\mathrm{M}_1,\mathrm{M}_2] = \frac{C_\ell^{\mathrm{Rec},X,i}[\mathrm{M}_1,\mathrm{M}_2]}{C_\ell^{\mathrm{True},X,i}[\mathrm{M}_2]}$$

$$\mathrm{Med}_\ell^{X}[\mathrm{M}_1,\mathrm{M}_2] = \mathrm{median}\left[\mathrm{rat}_\ell^{X,i}[\mathrm{M}_1,\mathrm{M}_2]\right]_i$$

$$p_\ell^{17,X}[\mathrm{M}_1,\mathrm{M}_2] = \mathrm{percentile}\left[\mathrm{rat}_\ell^{X,i}[\mathrm{M}_1,\mathrm{M}_2], 17\%\right]_i$$

$$p_\ell^{83,X}[\mathrm{M}_1,\mathrm{M}_2] = \mathrm{percentile}\left[\mathrm{rat}_\ell^{X,i}[\mathrm{M}_1,\mathrm{M}_2], 83\%\right]_i$$



Generalization factor

$$\mathrm{rat}_\ell^{X,i}[\mathrm{M}_1,\mathrm{M}_2] = \frac{C_\ell^{\mathrm{Rec},X,i}[\mathrm{M}_1,\mathrm{M}_2]}{C_\ell^{\mathrm{True},X,i}[\mathrm{M}_2]}$$

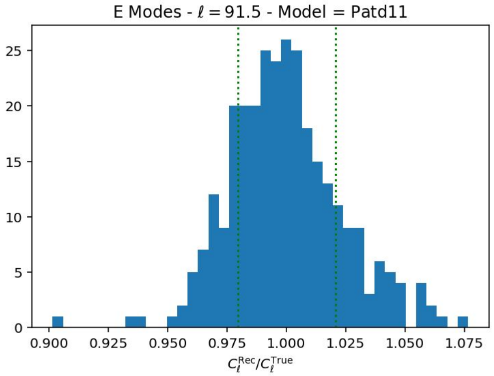
$$\mathrm{Med}_\ell^X[\mathrm{M}_1,\mathrm{M}_2] = \mathrm{median}\left[\mathrm{rat}_\ell^{X,i}[\mathrm{M}_1,\mathrm{M}_2]\right]_i$$

$$p_\ell^{17,X}[\mathrm{M}_1,\mathrm{M}_2] = \mathrm{percentile}\left[\mathrm{rat}_\ell^{X,i}[\mathrm{M}_1,\mathrm{M}_2], 17\%\right]_i$$

$$p_\ell^{83,X}[\mathrm{M}_1,\mathrm{M}_2] = \mathrm{percentile}\left[\mathrm{rat}_\ell^{X,i}[\mathrm{M}_1,\mathrm{M}_2], 83\%\right]_i$$

$$L_\ell^X[\mathrm{M}_1,\mathrm{M}_2] = \mathrm{Med}_\ell^X[\mathrm{M}_1,\mathrm{M}_2] - p_\ell^{17,X}[\mathrm{M}_1,\mathrm{M}_2]$$

$$U_\ell^X[\mathrm{M}_1,\mathrm{M}_2] = \mathrm{Med}_\ell^X[\mathrm{M}_1,\mathrm{M}_2] + p_\ell^{83,X}[\mathrm{M}_1,\mathrm{M}_2]$$



Generalization factor

$$\mathrm{Med}_\ell^X[\mathrm{M}_1,\mathrm{M}_2] = \mathrm{median}\left[\mathrm{rat}_\ell^{X,i}[\mathrm{M}_1,\mathrm{M}_2]\right]_i$$

$$p_\ell^{17,X}[\mathrm{M}_1,\mathrm{M}_2] = \mathrm{percentile}\left[\mathrm{rat}_\ell^{X,i}[\mathrm{M}_1,\mathrm{M}_2], 17\%\right]_i$$

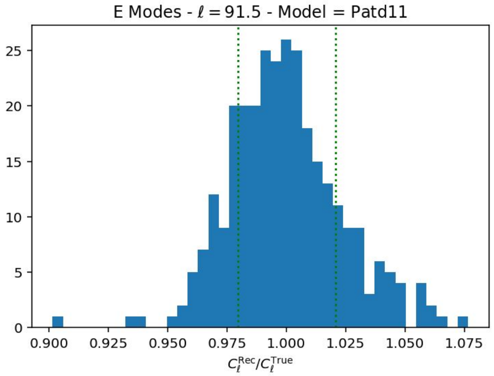
$$\mathrm{rat}_\ell^{X,i}[\mathrm{M}_1,\mathrm{M}_2] = \frac{C_\ell^{\mathrm{Rec},X,i}[\mathrm{M}_1,\mathrm{M}_2]}{C_\ell^{\mathrm{True},X,i}[\mathrm{M}_2]}$$

$$p_\ell^{83,X}[\mathrm{M}_1,\mathrm{M}_2] = \mathrm{percentile}\left[\mathrm{rat}_\ell^{X,i}[\mathrm{M}_1,\mathrm{M}_2], 83\%\right]_i$$

$$\begin{aligned}\Delta_\ell^X[\mathrm{M}_1,\mathrm{M}_2] &= U_\ell^X[\mathrm{M}_1,\mathrm{M}_2] - L_\ell^X[\mathrm{M}_1,\mathrm{M}_2]\\ &= p_\ell^{17,X}[\mathrm{M}_1,\mathrm{M}_2] + p_\ell^{83,X}[\mathrm{M}_1,\mathrm{M}_2]\end{aligned}$$

$$L_\ell^X[\mathrm{M}_1,\mathrm{M}_2] = \mathrm{Med}_\ell^X[\mathrm{M}_1,\mathrm{M}_2] - p_\ell^{17,X}[\mathrm{M}_1,\mathrm{M}_2]$$

$$U_\ell^X[\mathrm{M}_1,\mathrm{M}_2] = \mathrm{Med}_\ell^X[\mathrm{M}_1,\mathrm{M}_2] + p_\ell^{83,X}[\mathrm{M}_1,\mathrm{M}_2]$$



Generalization factor

$$\mathrm{Med}_\ell^X[\mathrm{M}_1,\mathrm{M}_2] = \mathrm{median}\left[\mathrm{rat}_\ell^{X,i}[\mathrm{M}_1,\mathrm{M}_2]\right]_i$$

$$p_\ell^{17,X}[\mathrm{M}_1,\mathrm{M}_2] = \mathrm{percentile}\left[\mathrm{rat}_\ell^{X,i}[\mathrm{M}_1,\mathrm{M}_2], 17\%\right]_i$$

$$p_\ell^{83,X}[\mathrm{M}_1,\mathrm{M}_2] = \mathrm{percentile}\left[\mathrm{rat}_\ell^{X,i}[\mathrm{M}_1,\mathrm{M}_2], 83\%\right]_i$$

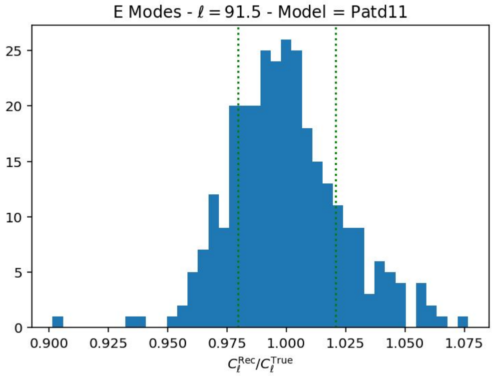
$$\mathrm{rat}_\ell^{X,i}[\mathrm{M}_1,\mathrm{M}_2] = \frac{C_\ell^{\mathrm{Rec},X,i}[\mathrm{M}_1,\mathrm{M}_2]}{C_\ell^{\mathrm{True},X,i}[\mathrm{M}_2]}$$

$$\begin{aligned}\Delta_\ell^X[\mathrm{M}_1,\mathrm{M}_2] &= U_\ell^X[\mathrm{M}_1,\mathrm{M}_2] - L_\ell^X[\mathrm{M}_1,\mathrm{M}_2]\\ &= p_\ell^{17,X}[\mathrm{M}_1,\mathrm{M}_2] + p_\ell^{83,X}[\mathrm{M}_1,\mathrm{M}_2]\end{aligned}$$

$$L_\ell^X[\mathrm{M}_1,\mathrm{M}_2] = \mathrm{Med}_\ell^X[\mathrm{M}_1,\mathrm{M}_2] - p_\ell^{17,X}[\mathrm{M}_1,\mathrm{M}_2]$$

$$U_\ell^X[\mathrm{M}_1,\mathrm{M}_2] = \mathrm{Med}_\ell^X[\mathrm{M}_1,\mathrm{M}_2] + p_\ell^{83,X}[\mathrm{M}_1,\mathrm{M}_2]$$

$$\mathrm{MSE}_\ell^X[\mathrm{M}_1,\mathrm{M}_2] = \left(\mathrm{Med}_\ell^X[\mathrm{M}_1,\mathrm{M}_2]\right)^2 + \left(\Delta_\ell^X[\mathrm{M}_1,\mathrm{M}_2]\right)^2$$



Generalization factor

$$\mathrm{Med}_\ell^X[\mathrm{M}_1,\mathrm{M}_2] = \mathrm{median}\left[\mathrm{rat}_\ell^{X,i}[\mathrm{M}_1,\mathrm{M}_2]\right]_i$$

$$\mathrm{rat}_\ell^{X,i}[\mathrm{M}_1,\mathrm{M}_2] = \frac{C_\ell^{\mathrm{Rec},X,i}[\mathrm{M}_1,\mathrm{M}_2]}{C_\ell^{\mathrm{True},X,i}[\mathrm{M}_2]}$$

$$p_\ell^{17,X}[\mathrm{M}_1,\mathrm{M}_2] = \mathrm{percentile}\left[\mathrm{rat}_\ell^{X,i}[\mathrm{M}_1,\mathrm{M}_2], 17\%\right]_i$$

$$p_\ell^{83,X}[\mathrm{M}_1,\mathrm{M}_2] = \mathrm{percentile}\left[\mathrm{rat}_\ell^{X,i}[\mathrm{M}_1,\mathrm{M}_2], 83\%\right]_i$$

$$\begin{aligned}\Delta_\ell^X[\mathrm{M}_1,\mathrm{M}_2] &= U_\ell^X[\mathrm{M}_1,\mathrm{M}_2] - L_\ell^X[\mathrm{M}_1,\mathrm{M}_2]\\ &= p_\ell^{17,X}[\mathrm{M}_1,\mathrm{M}_2] + p_\ell^{83,X}[\mathrm{M}_1,\mathrm{M}_2]\end{aligned}$$

$$L_\ell^X[\mathrm{M}_1,\mathrm{M}_2] = \mathrm{Med}_\ell^X[\mathrm{M}_1,\mathrm{M}_2] - p_\ell^{17,X}[\mathrm{M}_1,\mathrm{M}_2]$$

$$U_\ell^X[\mathrm{M}_1,\mathrm{M}_2] = \mathrm{Med}_\ell^X[\mathrm{M}_1,\mathrm{M}_2] + p_\ell^{83,X}[\mathrm{M}_1,\mathrm{M}_2]$$

$$\mathrm{MSE}_\ell^X[\mathrm{M}_1,\mathrm{M}_2] = \left(\mathrm{Med}_\ell^X[\mathrm{M}_1,\mathrm{M}_2]\right)^2 + \left(\Delta_\ell^X[\mathrm{M}_1,\mathrm{M}_2]\right)^2$$

$$\boxed{\mathcal{G}_\ell^X[\mathrm{M}_1,\mathrm{M}_2;\mathrm{M}_3] = \frac{\mathrm{MSE}_\ell^X[\mathrm{M}_1,\mathrm{M}_3] - \mathrm{MSE}_\ell^X[\mathrm{M}_2,\mathrm{M}_3]}{\mathrm{MSE}_\ell^X[\mathrm{M}_3,\mathrm{M}_3]}}$$

Generalization factor

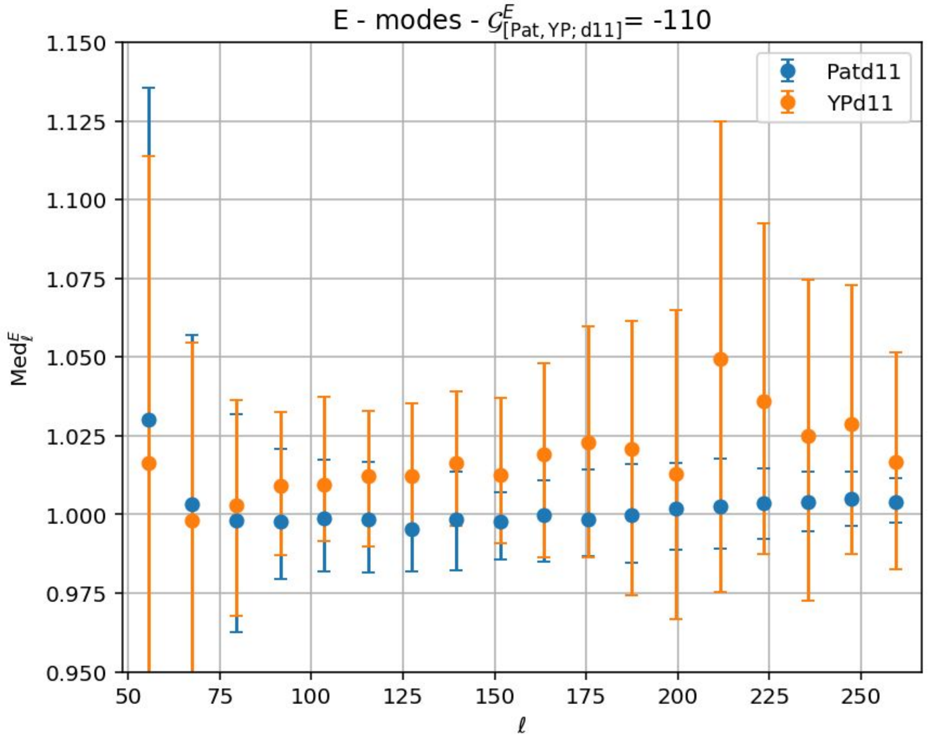
$$\mathcal{G}_\ell^X[\mathrm{M}_1, \mathrm{M}_2; \mathrm{M}_3] = \frac{\mathrm{MSE}_\ell^X[\mathrm{M}_1, \mathrm{M}_3] - \mathrm{MSE}_\ell^X[\mathrm{M}_2, \mathrm{M}_3]}{\mathrm{MSE}_\ell^X[\mathrm{M}_3, \mathrm{M}_3]}$$

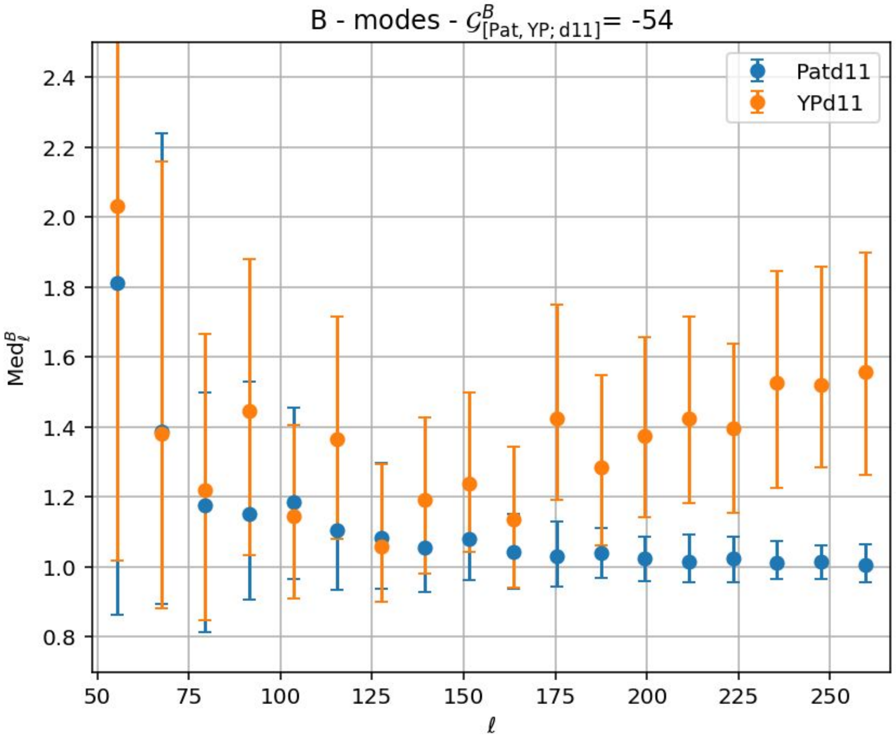
$$\downarrow$$

$$\boxed{\mathcal{G}^{M_1, M_2; M_3} = \left\langle \mathcal{G}_\ell^X[\mathrm{M}_1, \mathrm{M}_2; \mathrm{M}_3] \right\rangle_{\ell_{\min} < \ell < \ell_{\max}}}$$

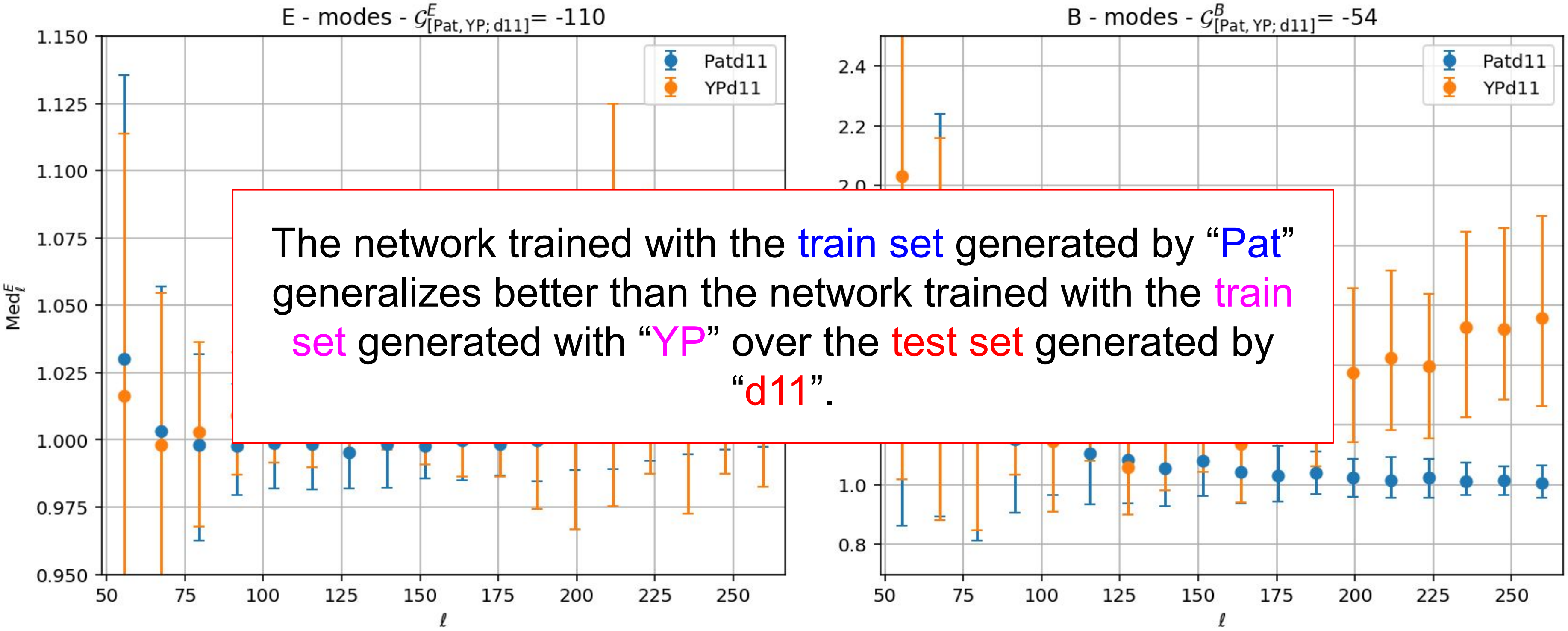
- If $\mathscr{G}^{\mathrm{M1,M2;M3}} < 0$ —> M1 generalizes better than M2 on M3
- If $\mathscr{G}^{\mathrm{M1,M2;M3}} > 0$ —> M2 generalizes better than M1 on M3

Preliminary result: [Pat, YP; d11]





Preliminary result: [Pat, YP; d11]



The network trained with the train set generated by “Pat” generalizes better than the network trained with the train set generated with “YP” over the test set generated by “d11”.

Thank you for the attention!