Component separation for CMB polarization data using neural networks

Ixaka Labadie García – End of Master Project
University of Cantabria
Master in Particle Physics and the Cosmos
23 July 2021

Contents

Introduction

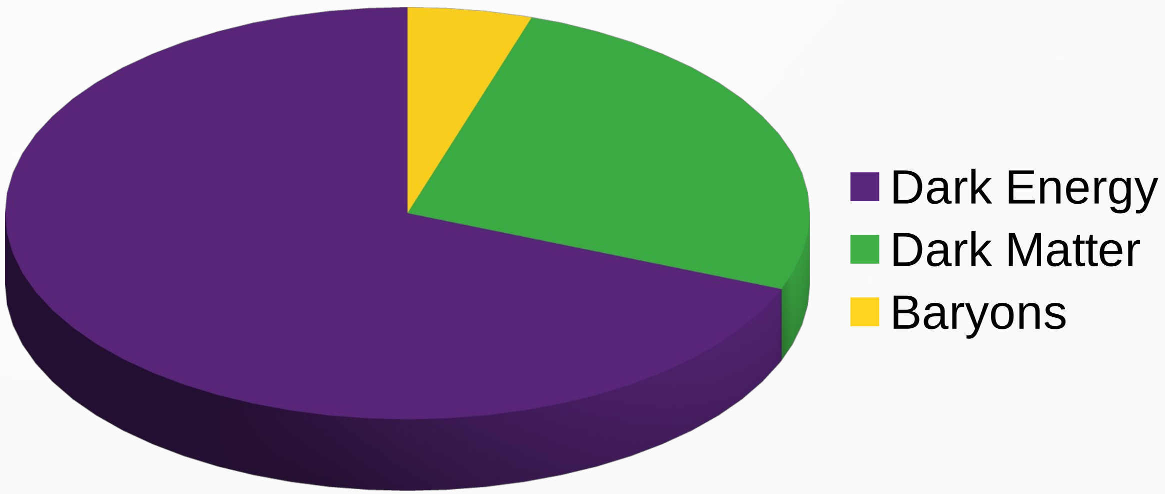
Standard cosmological model (ΛCDM): Big Bang with a short period of inflation

Based on General Relativity, flat geometry and accelerated expansion

Composition: Dark energy (69%), CDM (26%), baryonic (5%)

Homogeneous and **isotropic** Universe at large scales

Cosmological observables: Cosmic microwave background (**CMB**), large scale structure, etc.



Introduction: CMB

Started propagating 380,000 yrs after Big Bang, last scattering surface

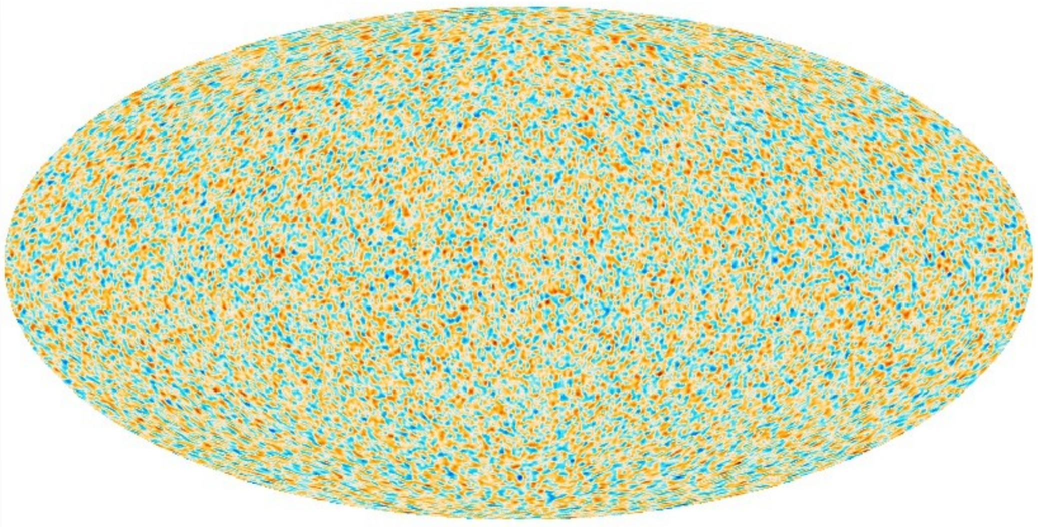
Black body radiation with mean $\mathrm{T}_0 = (2.72548 \pm 0.00057)\,\mathrm{K}$

Fluctuations of the order of $\Delta T / T_0 \approx 10^{-5}$

Primary: during or before last scattering, due to inflation

- **Scalar:** already detected and well known
- **Tensor**: Primordial gravitational waves, very weak. Could be detected in the B-mode

Secondary: after last scattering. Lensing, Sunyaev–Zel'dovich



Introduction: CMB

Temperature fluctuations expanded to **spherical harmonics**

$$\Delta T(\hat{n}) \equiv T(\hat{n}) - T_0 = \sum_{\ell m} a_{\ell m} Y_\ell^m(\hat{n})$$

The CMB radiation is **polarized** due to Thomson scattering

$$Q(n) \pm iU(\hat{n}) = \sum_{\ell m} a_{\pm 2,\ell m} \, {}_{\pm 2}Y_\ell^m(\hat{n})$$

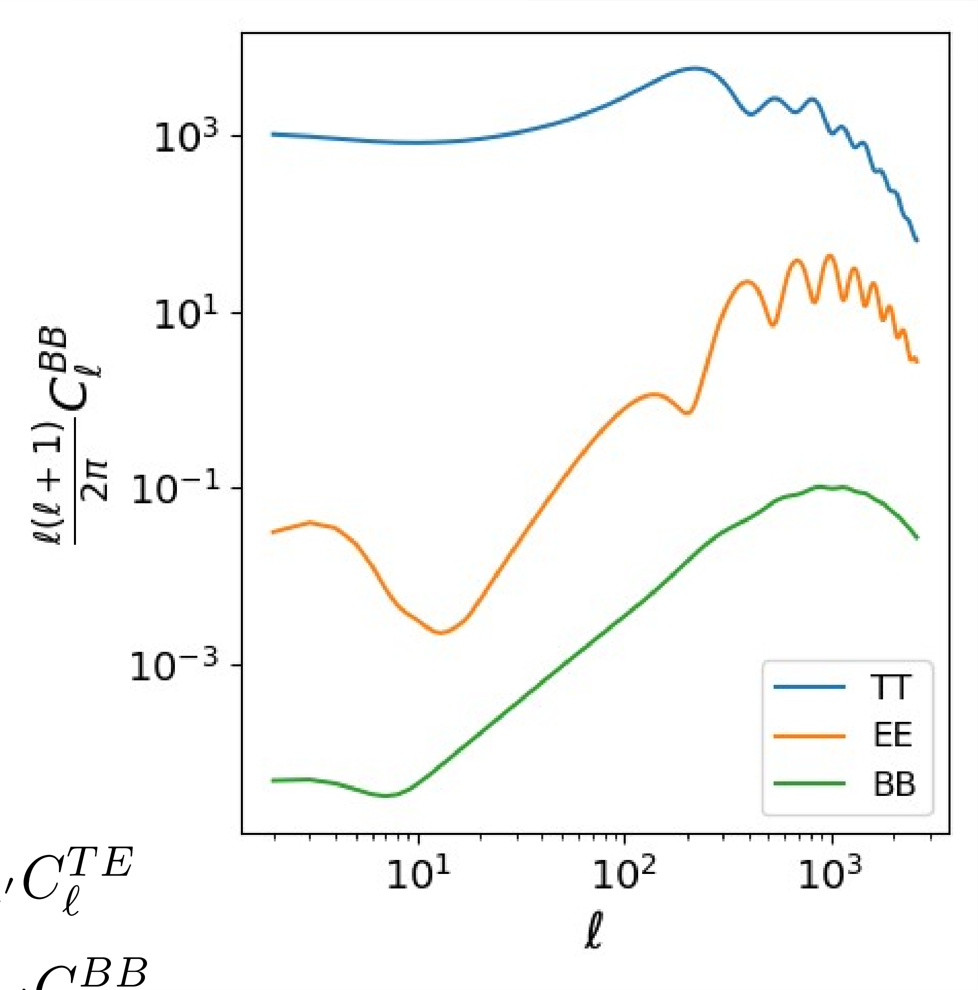
$$a_{E,\ell m} = -(a_{2,\ell m} + a_{-2,\ell m})/2$$

$$a_{B,\ell m} = i(a_{2,\ell m} - a_{-2,\ell m})/2$$

The **multipole coefficients** C_ℓ constitute the power spectrum

$$\langle a^*_{T,\ell m} a_{T,\ell' m'} \rangle = \delta_{\ell\ell'} \delta_{mm'} C_\ell^{TT} \quad \langle a^*_{T,\ell m} a_{E,\ell' m'} \rangle = \delta_{\ell\ell'} \delta_{mm'} C_\ell^{TE}$$

$$\langle a^*_{E,\ell m} a_{E,\ell' m'} \rangle = \delta_{\ell\ell'} \delta_{mm'} C_\ell^{EE} \quad \langle a^*_{B,\ell m} a_{B,\ell' m'} \rangle = \delta_{\ell\ell'} \delta_{mm'} C_\ell^{BB}$$



Introduction: CMB

Inflation: period of exponentially accelerated expansion

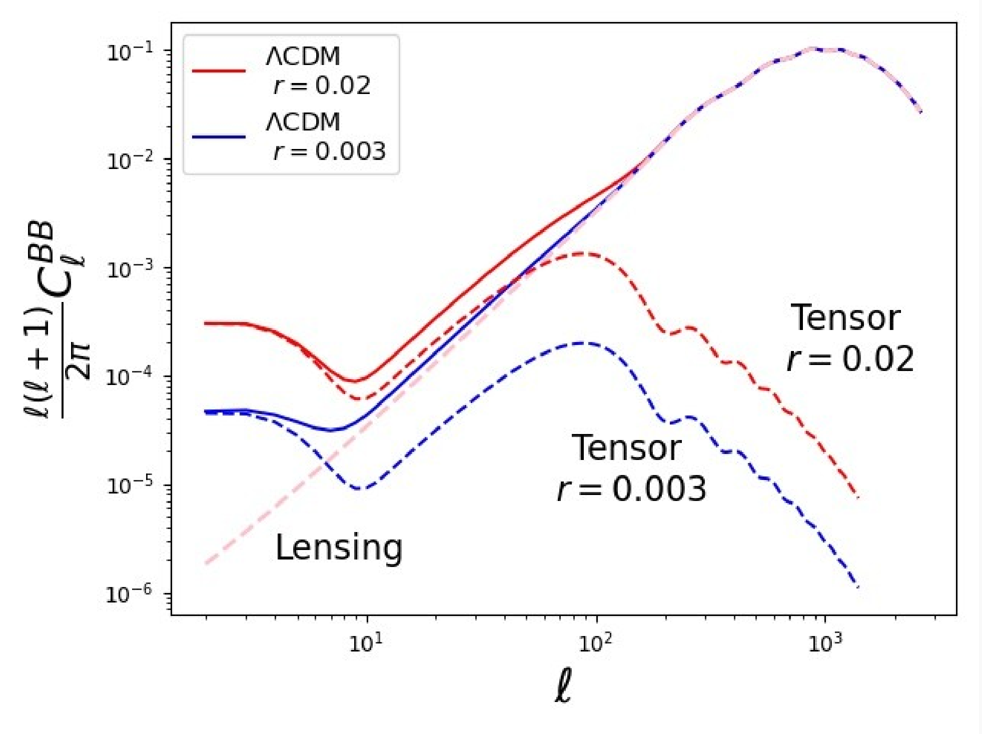
Predicts tensor perturbations in the metric

Thomson scattering with tensor perturbations produces both E- and B-modes, unlike with scalars that produces only E

Strength of GWs given by **tensor-to-scalar ratio** $r = A_t / A_s$

Relevant **models predict r =0.003** others from $r \sim 0.1$ to $r \sim 10^{-52}$

Most recent **constraint r < 0.044** by *Planck* and BICEP/Keck



Introduction: Foregrounds

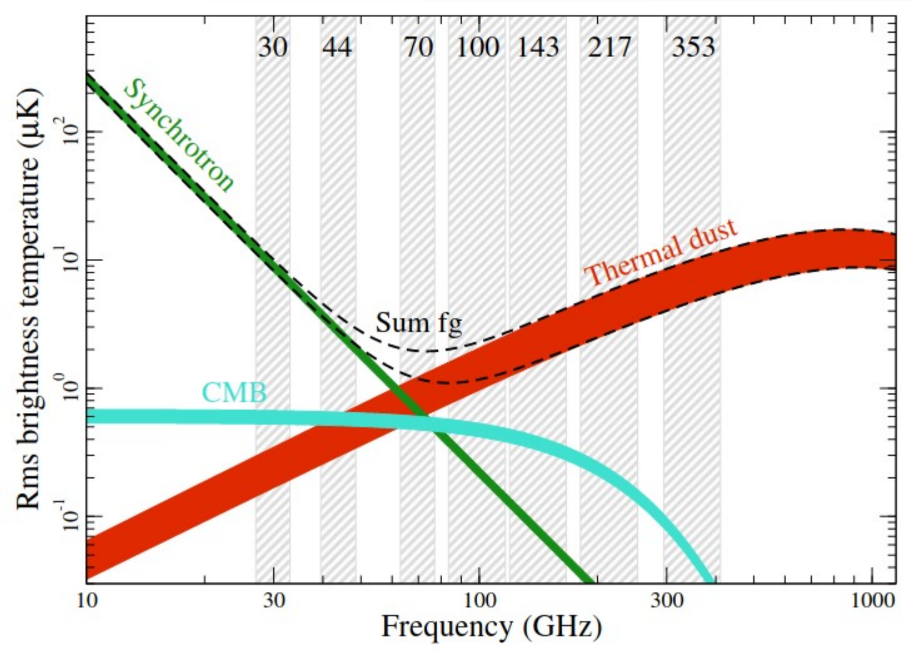
Thermal dust:
Dust particles with $T_d \approx 20$ K
Modified black body

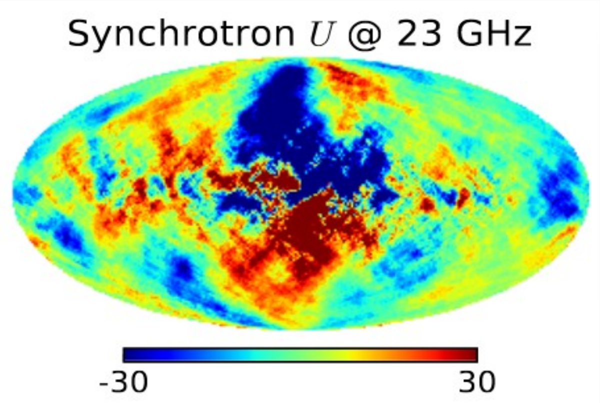
Synchrotron:
Charged particles accelerated in magnetic field
Approximated by power law

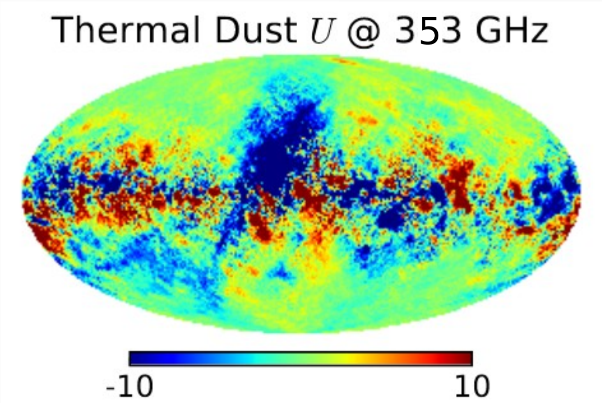
Component separation methods to remove foregrounds:
Commander, SMICA, SEVEM, NILC
(we use ILC for comparison)

Our goal is to **create a new method using Neural Networks**

(Planck Collaboration 2016)





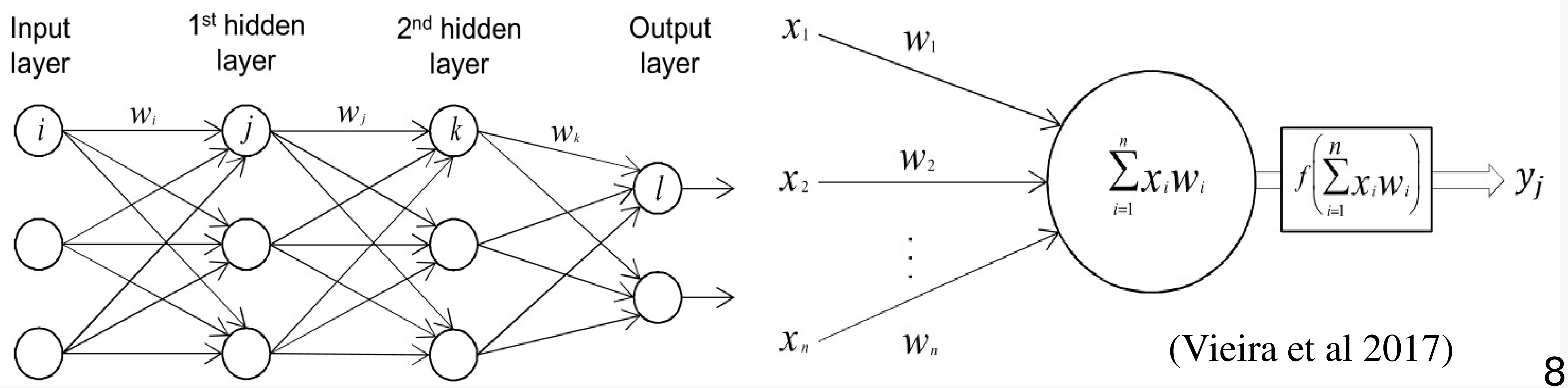


Introduction: Neural Networks

Function that **makes predictions** given input data

During training tries to guess ***labels*** from ***features***

- Number of **layers** and **nodes**: complexity of function
- **Activation** function (ReLU, Softmax): introduces non-linearity
- **Loss** function (MAE, SCCE): function to minimize
- **Optimizer** (Adam): minimization, stochastic gradient descent



(Vieira et al 2017)

Simulations

Freq. (GHz)	$\omega_P^{-1/2}$ (μK arcmin)	θ_{FWHM} (arcmin)
40	37.5	69
50	24.0	56
60	19.9	48
68	16.2	43
78	13.5	39
89	11.7	35
100	9.2	29
119	7.6	25
140	5.9	23
166	6.5	21
195	5.8	20
235	7.7	19
280	13.2	24
337	19.5	20
402	37.5	17

Simulations of Japanese mission **LiteBIRD**

Best resolution equivalent to N_{side}= 512

Total sensitivity of **2.5 μK arcmin over 3 years**

Being made to measure the B-modes **between ℓ=2 and ℓ=200**

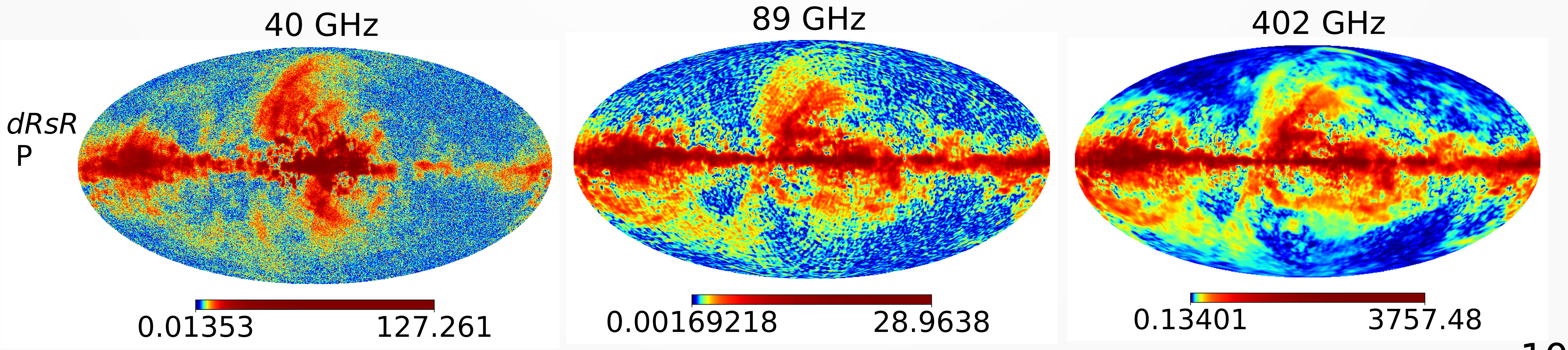
If detected the expected **uncertainty δr = 0.001**

If not, expected **upper limit r < 0.002** at 95% C.L.

Simulations

Foreground maps simulated with PySM

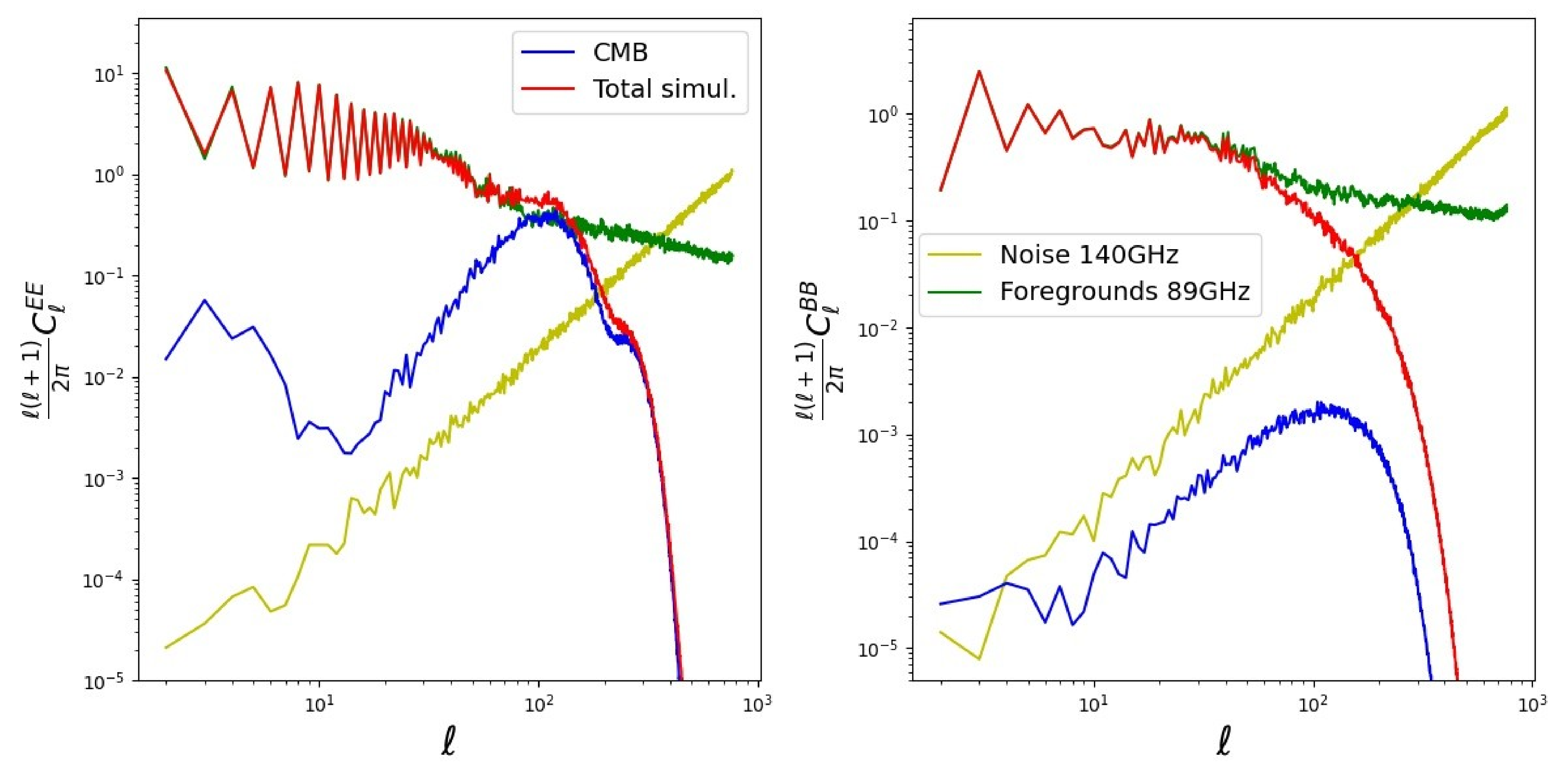
- ***d1:*** Templates from *Planck* using Commander
- ***s1:*** Templates from WMAP 9-year
  Both with spatially varying spectral indexes (also T_d)
- ***dRsR:*** Change spectral index of *d1* and *s1* for random values from Gaussian distribution
  Dust $\beta_d = (1.55 \pm 0.05)$; Synchrotron $\beta_s = (-3.1 \pm 0.3)$



Simulations

Angular power spectrum of the CMB calculated using CAMB

Best-fit parameters from *Planck* 2018 with $r = 0.003$

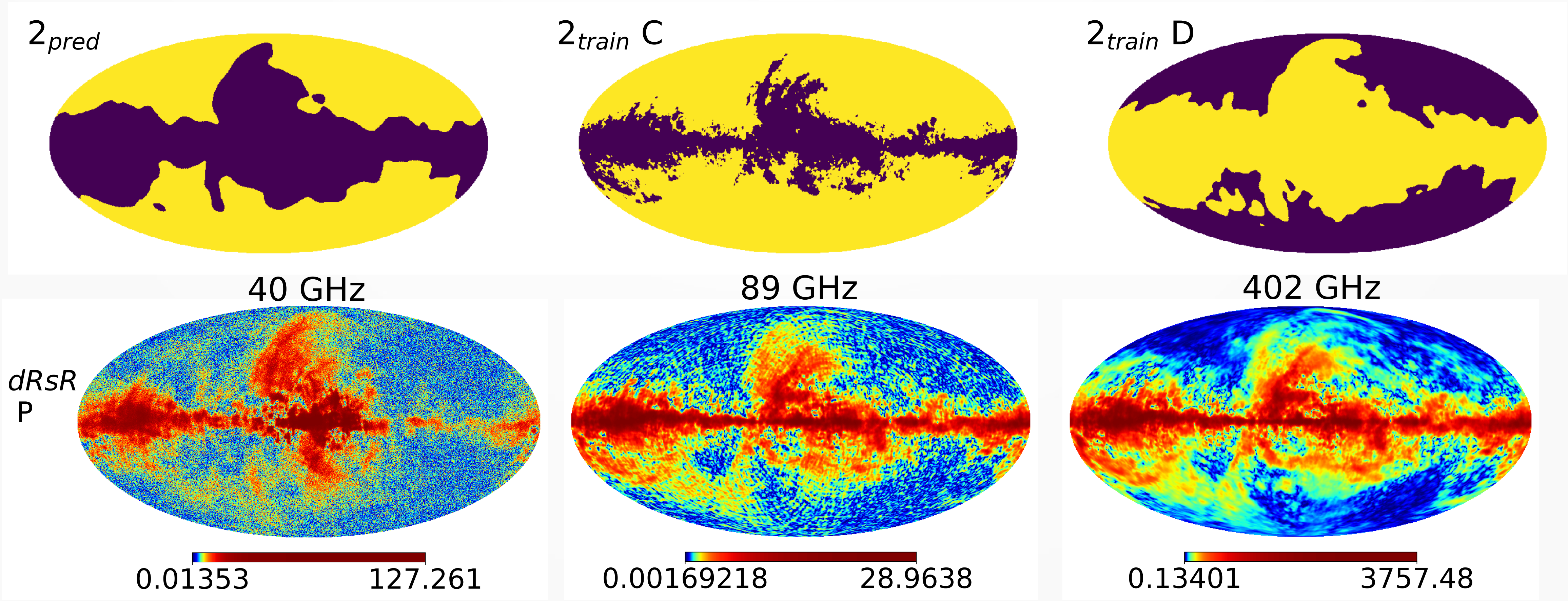


Foregrounds and noise without Gaussian filter

Simulations

Masks for training and predictions based on $\mathrm{P} = \sqrt{Q^2 + U^2}$

Threshold in foreground intensity and Gaussian smoothing



Methodology

Data at hand: values of pixels at 15 frequencies (*features*)

values of CMB at the same pixels (*labels*)

Create training (70%), validation (15%) and test (15%) sets

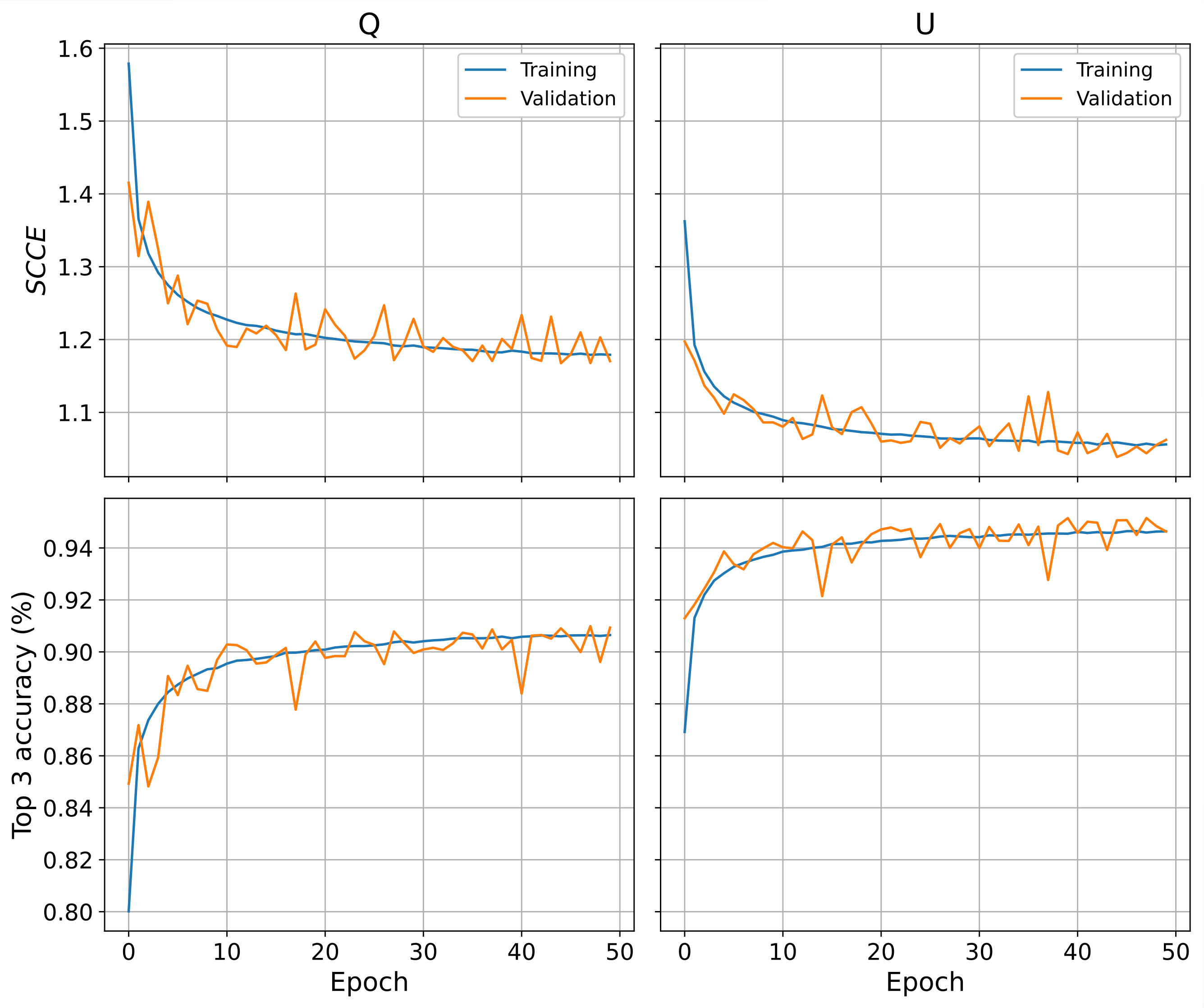
Normalization of features: $(\mathrm{x}_\nu - \bar{x}_\nu)/\sigma_{x_\nu}$

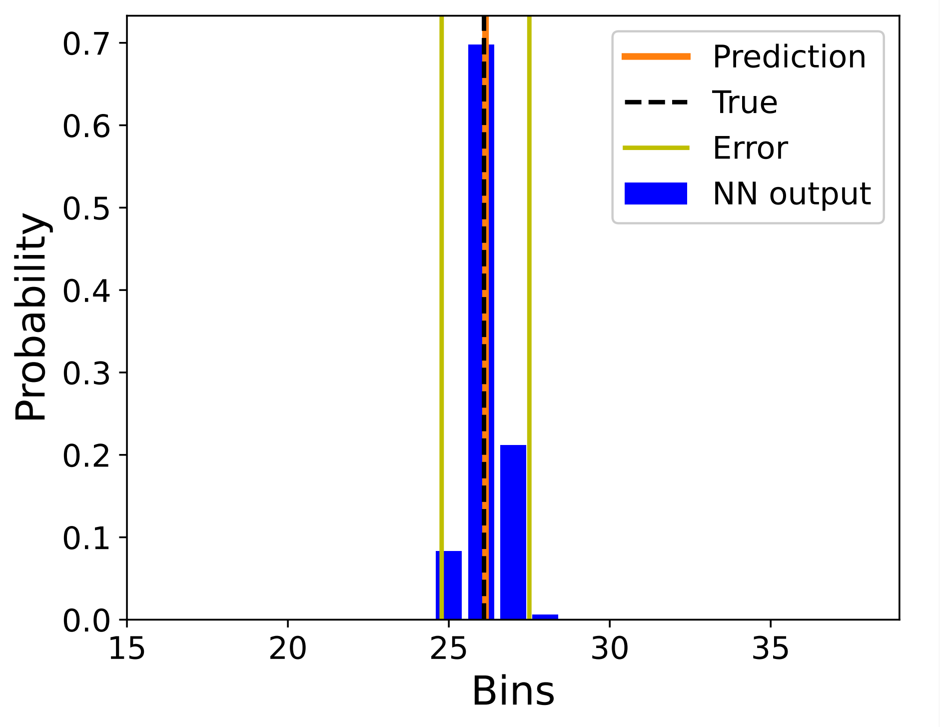
Regression: Predict the true value of the CMB at that pixel

Classification: Give probabilities of true value of CMB at pixel corresponding to temperature bins (0.1 μK)

Using N_{SIDE} = 256 HEALPix maps, N_{PIX} = 786,432

Methodology





Classification gives an estimation of the error

Not correct but good approximation

Results and Analysis

Relative error of 10 maps

Average and dispersion

Various configurations tested, only best shown

Regression and classification similar

No improvement in training:

- Separating Q and U
- Using mask
- With double features (*d1s1*)

	Q_C (%)	U_C (%)	Q_G (%)	U_G (%)
Reg. d1s1	12.11 ±0.19	11.7 ±0.2	28.4 ±0.3	20.5 ±0.5
Clas. d1s1	12.14 ±0.17	11.8 ±0.3	28.4 ±0.3	21.1 ±0.4
d1s1 w/ dRsR	12.08 ±0.17	11.9 ±0.2	28.0 ±0.5	21.1 ±0.5
ILC d1s1	14.8 ±1.6	15.2 ±1.9	36.7 ±0.5	25.2 ±0.7
Reg. dRsR	11.1 ±0.3	10.9 ±0.5	26.5 ±0.5	18.8 ±0.8
Clas. dRsR	11.3 ±0.4	11.2 ±0.4	26.9 ±0.7	19.2 ±1.1
ILC dRsR	14 ± 3	13.7 ±1.7	23 ± 7	19 ± 4

Results and Analysis

dRsR models equal or better with *d1s1* simulations

U better than Q in Galactic region

ILC clearly worse

dRsR 1% less error than *d1s1*

Classification needs CMB x 1.2

	Q_C (%)	U_C (%)	Q_G (%)	U_G (%)
Reg. d1s1	12.11 ±0.19	11.7 ±0.2	28.4 ±0.3	20.5 ±0.5
Clas. d1s1	12.14 ±0.17	11.8 ±0.3	28.4 ±0.3	21.1 ±0.4
d1s1 w/ dRsR	12.08 ±0.17	11.9 ±0.2	28.0 ±0.5	21.1 ±0.5
ILC d1s1	14.8 ±1.6	15.2 ±1.9	36.7 ±0.5	25.2 ±0.7
Reg. dRsR	11.1 ±0.3	10.9 ±0.5	26.5 ±0.5	18.8 ±0.8
Clas. dRsR	11.3 ±0.4	11.2 ±0.4	26.9 ±0.7	19.2 ±1.1
ILC dRsR	14 ± 3	13.7 ±1.7	23 ± 7	19 ± 4

Results and Analysis

dRsR models equal or better with *d1s1* simulations

U better than Q in Galactic region

ILC clearly worse

dRsR 1% less error than *d1s1*

Classification needs CMB x 1.2

	Q_C (%)	U_C (%)	Q_G (%)	U_G (%)
Reg. d1s1	12.11 ±0.19	11.7 ±0.2	28.4 ±0.3	20.5 ±0.5
Clas. d1s1	12.14 ±0.17	11.8 ±0.3	28.4 ±0.3	21.1 ±0.4
d1s1 w/ dRsR	12.08 ±0.17	11.9 ±0.2	28.0 ±0.5	21.1 ±0.5
ILC d1s1	14.8 ±1.6	15.2 ±1.9	36.7 ±0.5	25.2 ±0.7
Reg. dRsR	11.1 ±0.3	10.9 ±0.5	26.5 ±0.5	18.8 ±0.8
Clas. dRsR	11.3 ±0.4	11.2 ±0.4	26.9 ±0.7	19.2 ±1.1
ILC dRsR	14 ± 3	13.7 ±1.7	23 ± 7	19 ± 4

Results and Analysis

dRsR models equal or better with *d1s1* simulations

U better than Q in Galactic region

ILC clearly worse

dRsR 1% less error than *d1s1*

Classification needs CMB x 1.2

	Q_C (%)	U_C (%)	Q_G (%)	U_G (%)
Reg. d1s1	12.11 ±0.19	11.7 ±0.2	28.4 ±0.3	20.5 ±0.5
Clas. d1s1	12.14 ±0.17	11.8 ±0.3	28.4 ±0.3	21.1 ±0.4
d1s1 w/ dRsR	12.08 ±0.17	11.9 ±0.2	28.0 ±0.5	21.1 ±0.5
ILC d1s1	14.8 ±1.6	15.2 ±1.9	36.7 ±0.5	25.2 ±0.7
Reg. dRsR	11.1 ±0.3	10.9 ±0.5	26.5 ±0.5	18.8 ±0.8
Clas. dRsR	11.3 ±0.4	11.2 ±0.4	26.9 ±0.7	19.2 ±1.1
ILC dRsR	14 ± 3	13.7 ±1.7	23 ± 7	19 ± 4

Results and Analysis

dRsR models equal or better with *d1s1* simulations

U better than Q in Galactic region

ILC clearly worse

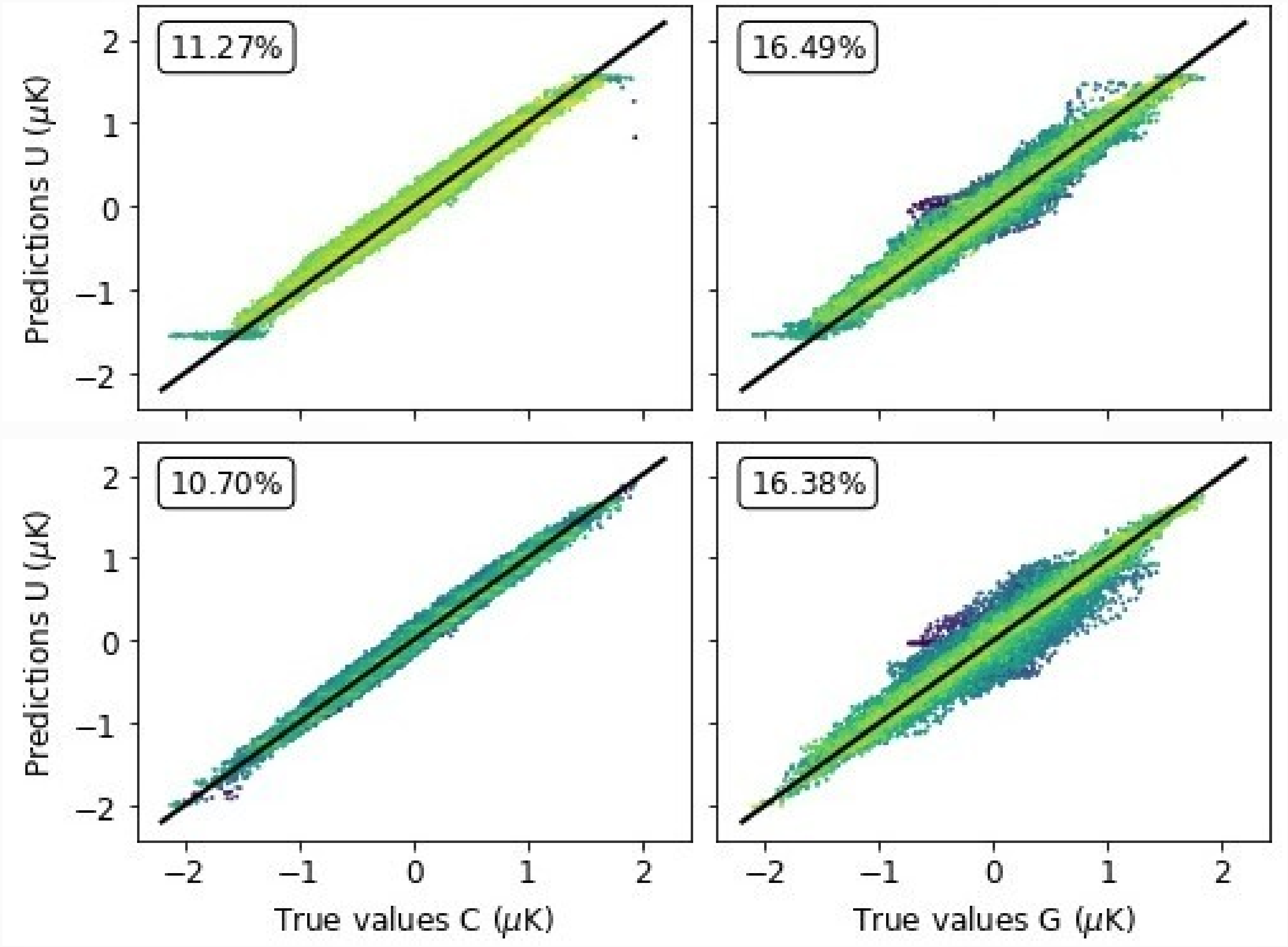
dRsR 1% less error than *d1s1*

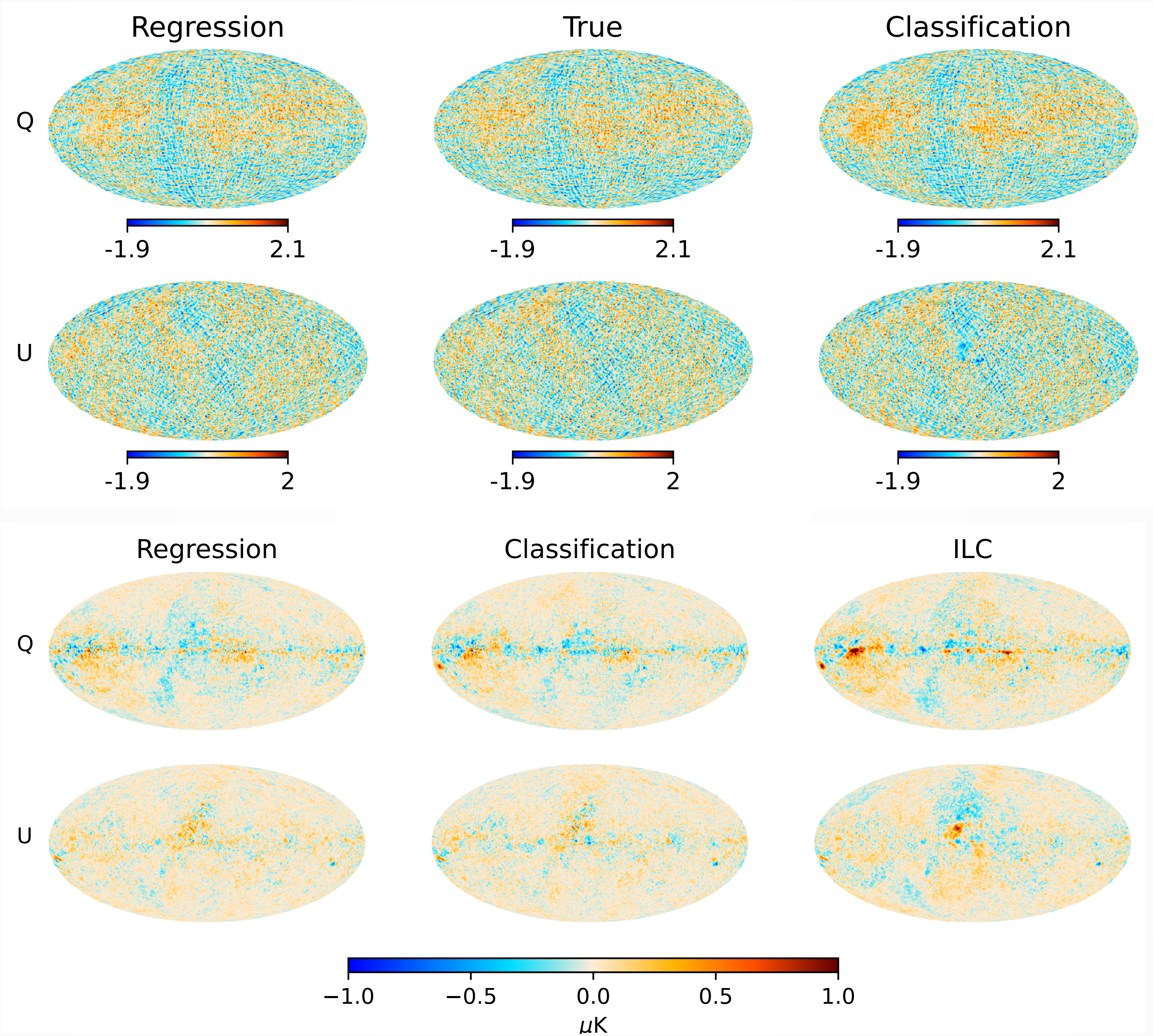
Classification needs CMB x 1.2

	Q_C (%)	U_C (%)	Q_G (%)	U_G (%)
Reg. d1s1	12.11 ±0.19	11.7 ±0.2	28.4 ±0.3	20.5 ±0.5
Clas. d1s1	12.14 ±0.17	11.8 ±0.3	28.4 ±0.3	21.1 ±0.4
d1s1 w/ dRsR	12.08 ±0.17	11.9 ±0.2	28.0 ±0.5	21.1 ±0.5
ILC d1s1	14.8 ±1.6	15.2 ±1.9	36.7 ±0.5	25.2 ±0.7
Reg. dRsR	11.1 ±0.3	10.9 ±0.5	26.5 ±0.5	18.8 ±0.8
Clas. dRsR	11.3 ±0.4	11.2 ±0.4	26.9 ±0.7	19.2 ±1.1
ILC dRsR	14 ± 3	13.7 ±1.7	23 ± 7	19 ± 4

Top: without multiply by 1.2
Bottom: multiply by 1.2

Light: low uncertainty
Dark: high uncertainty



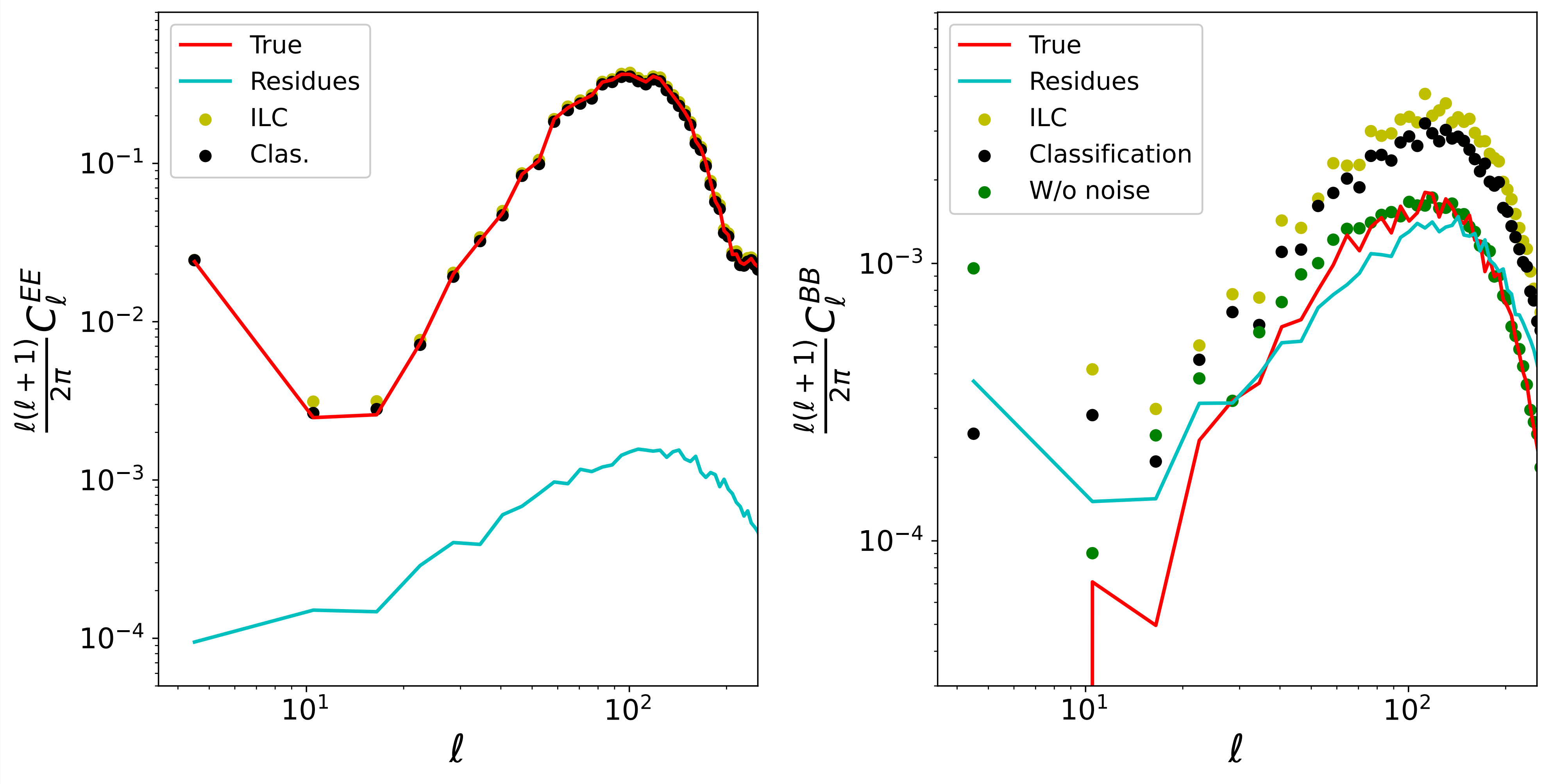

Regression
True
Classification
Q
U
-1.9
2.1
-1.9
2
Regression
Classification
ILC
Q
U
−1.0
−0.5
0.0
0.5
1.0
μK

Results and Analysis

Estimated with f_{SKY} = 0.55

Regression ≈ Classification

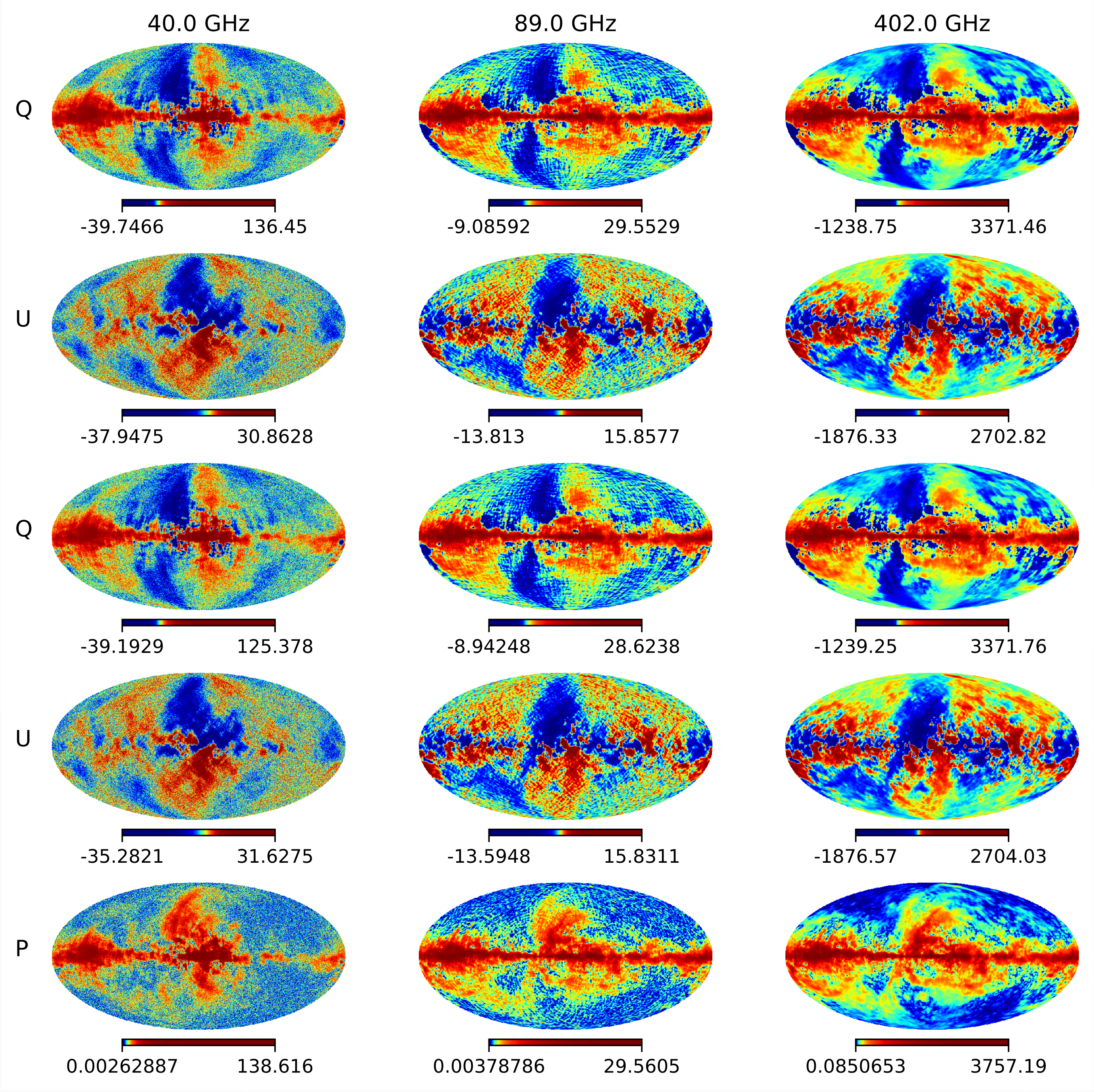
dRsR



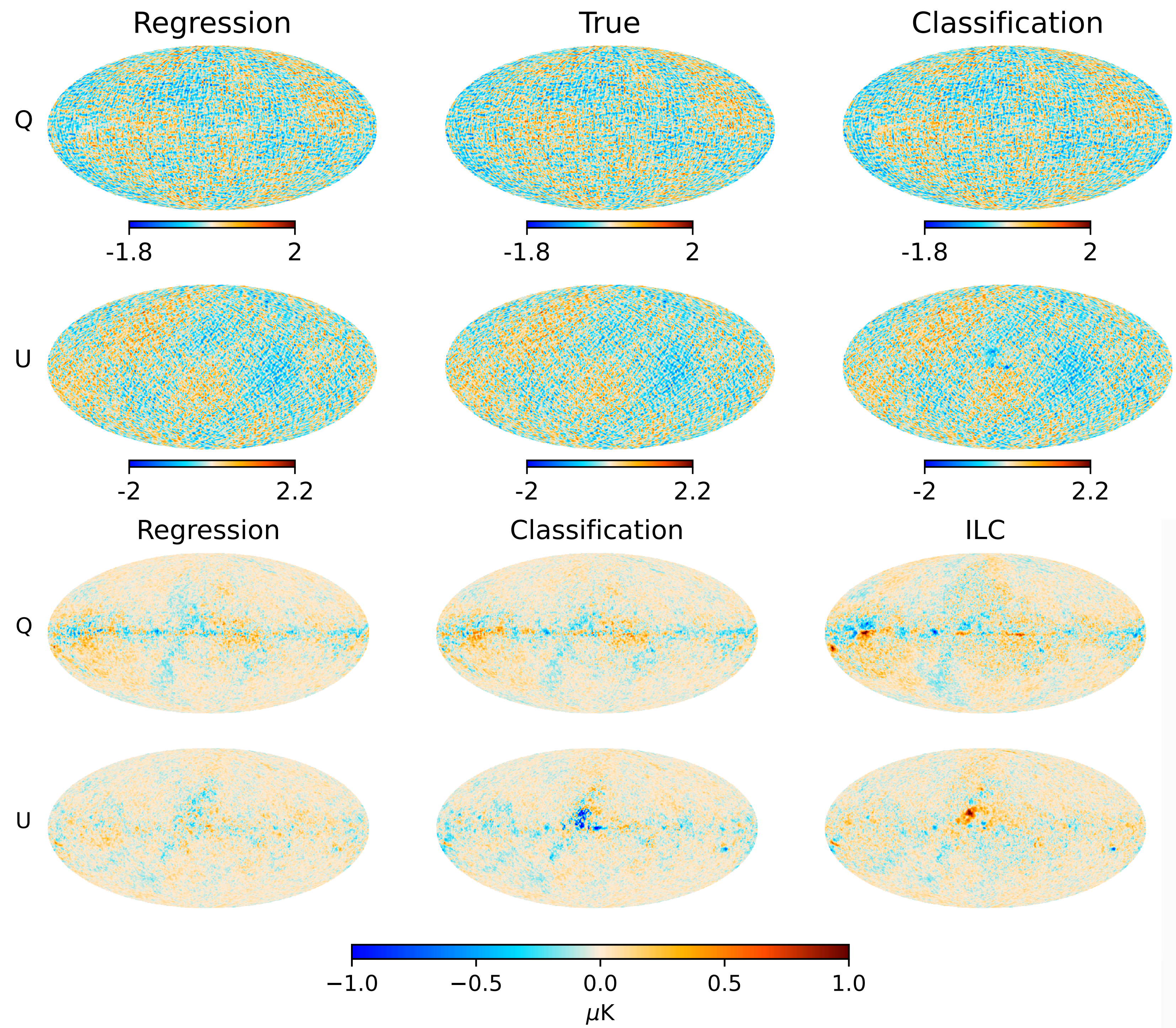
Conclusions

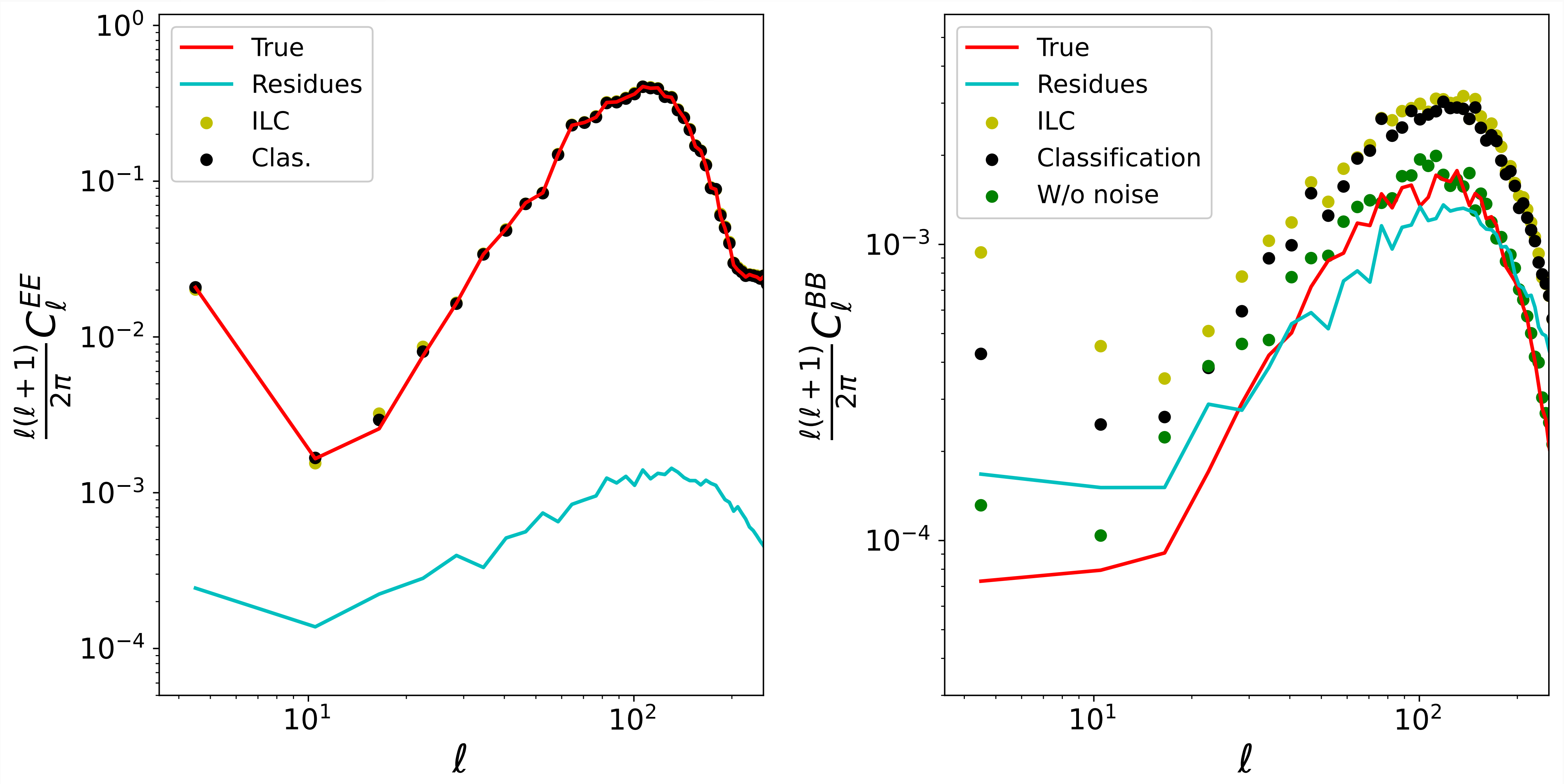
- **Medium-low complexity** networks work better
- Regression and classification **similar** results
- Classification can give an **approximation of the errors**
- Best results: *d1s1* **12%** and >21% ; *dRsR* **11%** and >18%
- Real-space ILC errors around 14%, **3% worse than NNs**
- NNs are a good prospect for a component separation method
- **More analysis needed** regarding classification errors, training with more features, etc.

EXTRA SLIDES

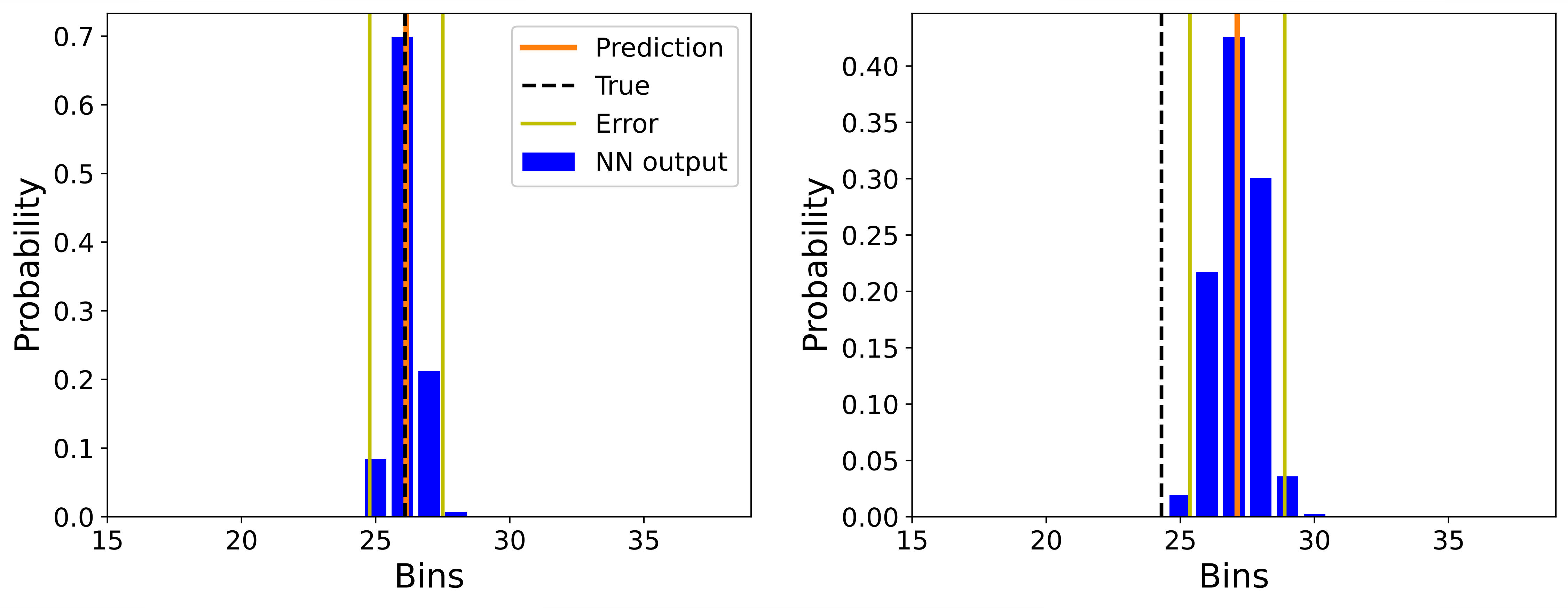


40.0 GHz
89.0 GHz
402.0 GHz
Q
-39.7466
136.45
-9.08592
29.5529
-1238.75
3371.46
U
-37.9475
30.8628
-13.813
15.8577
-1876.33
2702.82
Q
-39.1929
125.378
-8.94248
28.6238
-1239.25
3371.76
U
-35.2821
31.6275
-13.5948
15.8311
-1876.57
2704.03
P
0.00262887
138.616
0.00378786
29.5605
0.0850653
3757.19


Regression
True
Classification
Q
-1.8
2
-1.8
2
-1.8
2
U
-2
2.2
-2
2.2
-2
2.2
Regression
Classification
ILC
Q
U
−1.0
−0.5
0.0
0.5
1.0
μK



True
Residues
ILC
Clas.
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
$\frac{\ell(\ell+1)}{2\pi}C_\ell^{EE}$
10^1
10^2
ℓ
True
Residues
ILC
Classification
W/o noise
10^{-3}
10^{-4}
$\frac{\ell(\ell+1)}{2\pi}C_\ell^{BB}$
10^1
10^2
ℓ



Prediction
True
Error
NN output
Probability
Bins
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
15
20
25
30
35