Decoding the Early Universe: Machine Learning Applications in CMB Analysis

Leonora Kardum

LABORATOIRE ASTROPARTICULE & COSMOLOGIE

for HAMLET – How to Apply Machine Learning to Experimental and Theoretical Physics Conference

August, 2024

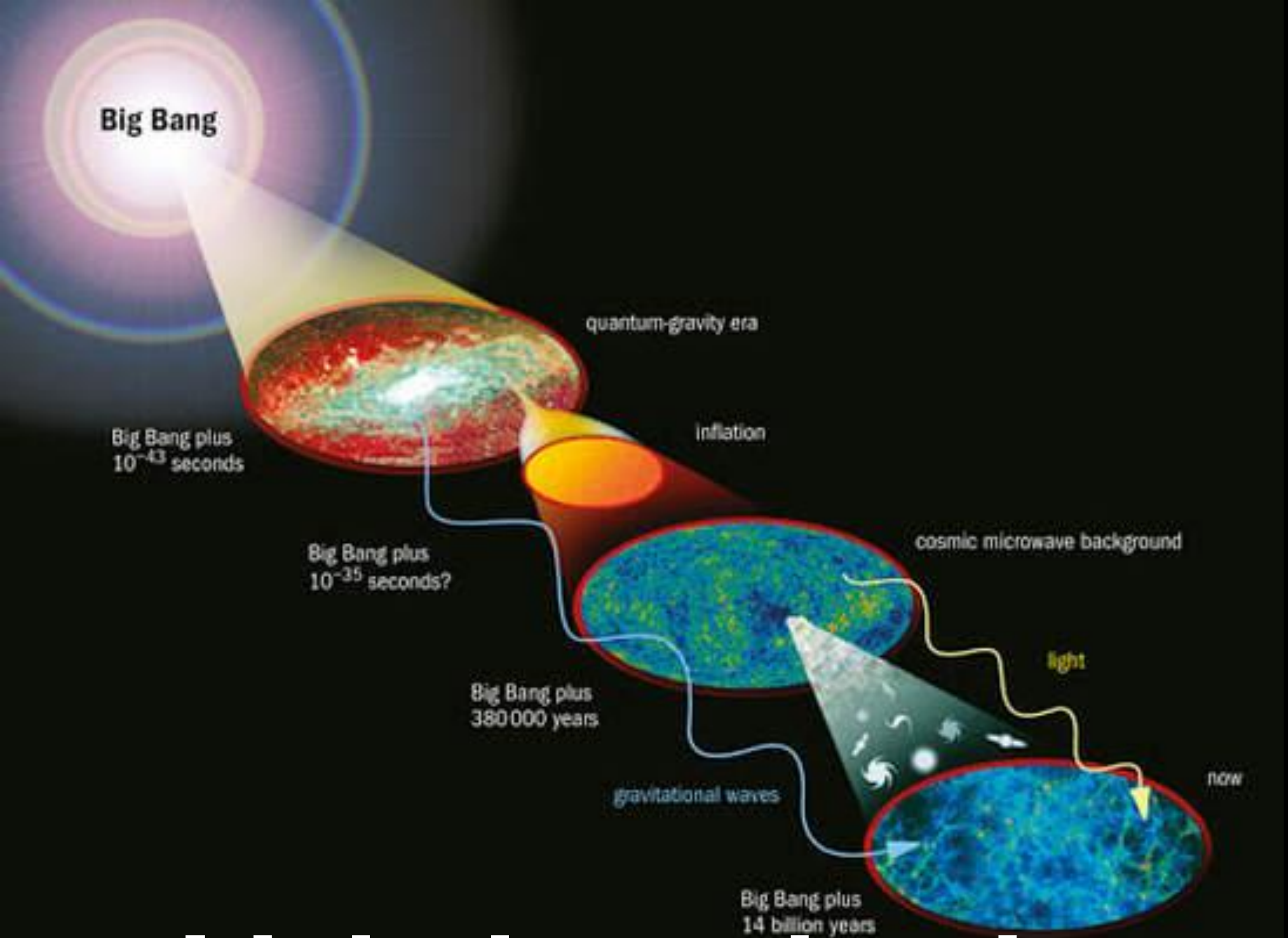
Contents

PART I.

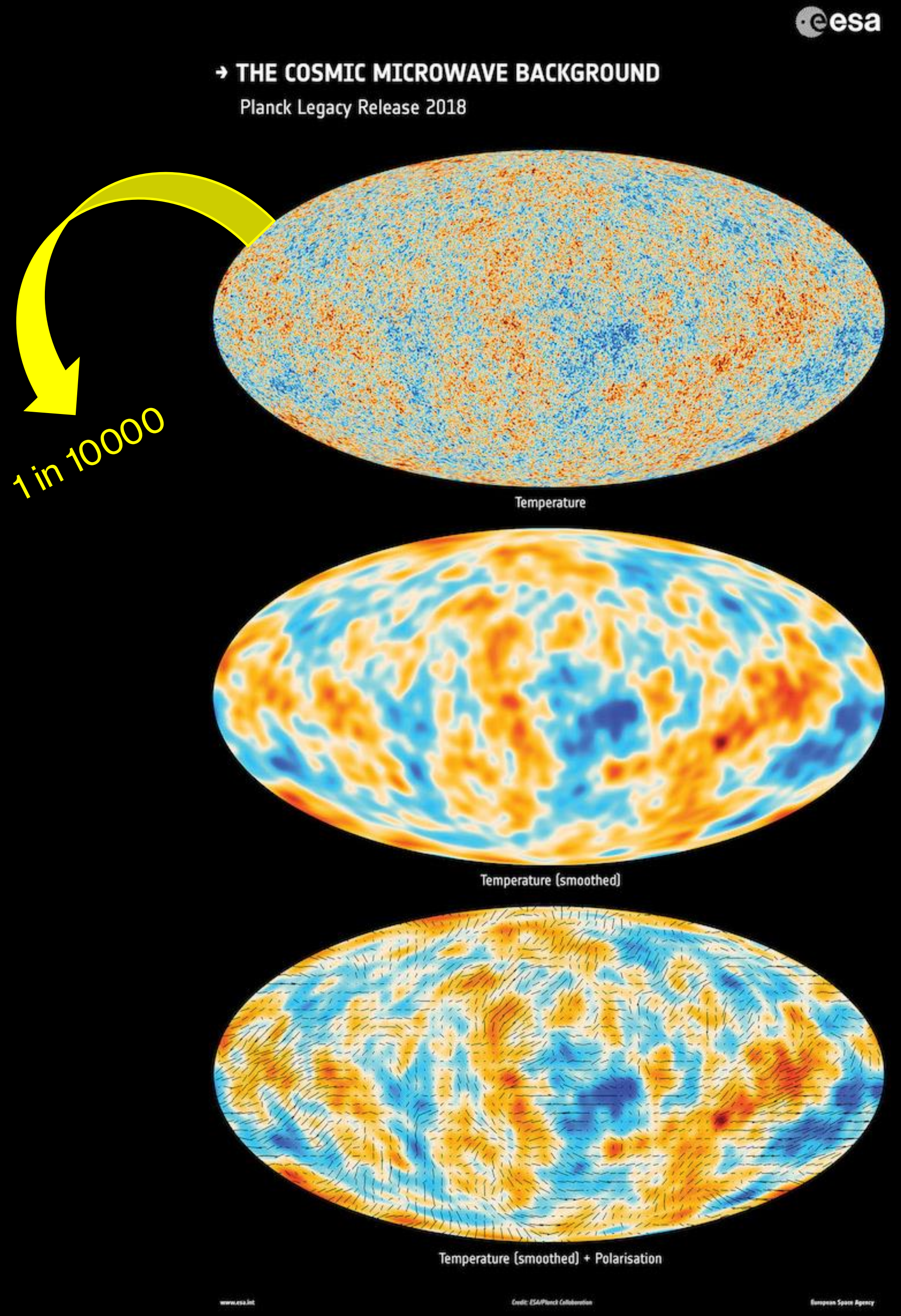
- Light introduction to cosmology and CMB
- Tensor to scalar ratio r
- Challenges in CMB analysis
- QUBIC

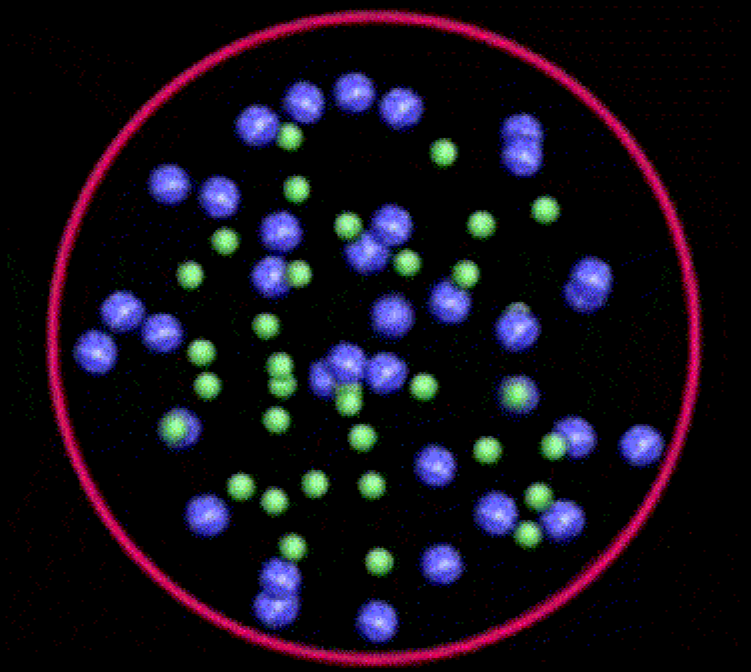
PART II.

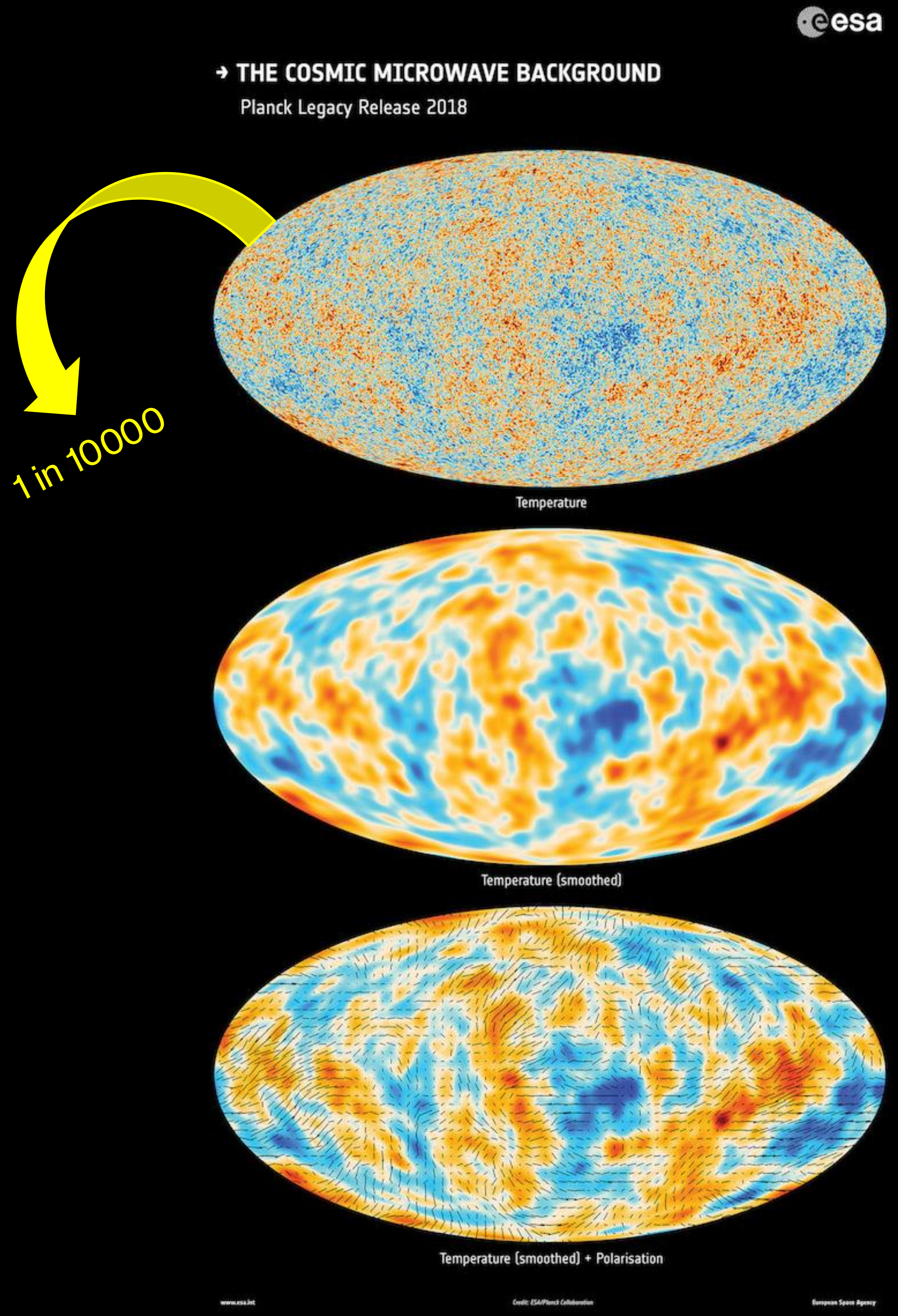
- Neural Networks in CMB analysis
- Adaptation
- Multifeature data
- The framework
- PINNs



Light introduction to cosmology and CMB

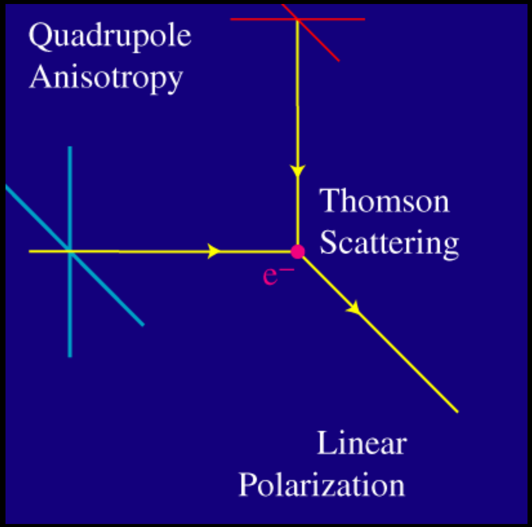


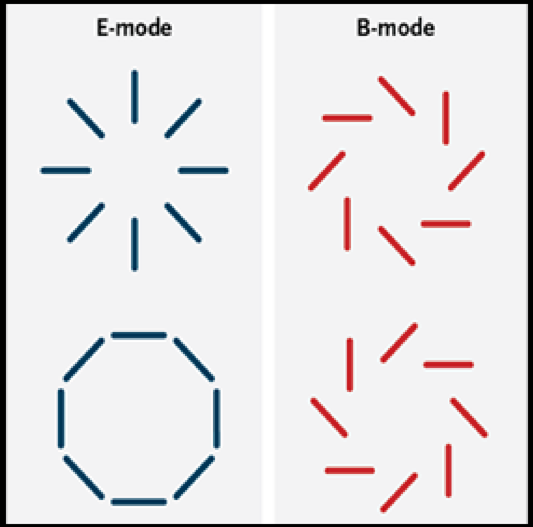




Light introduction to cosmology and CMB

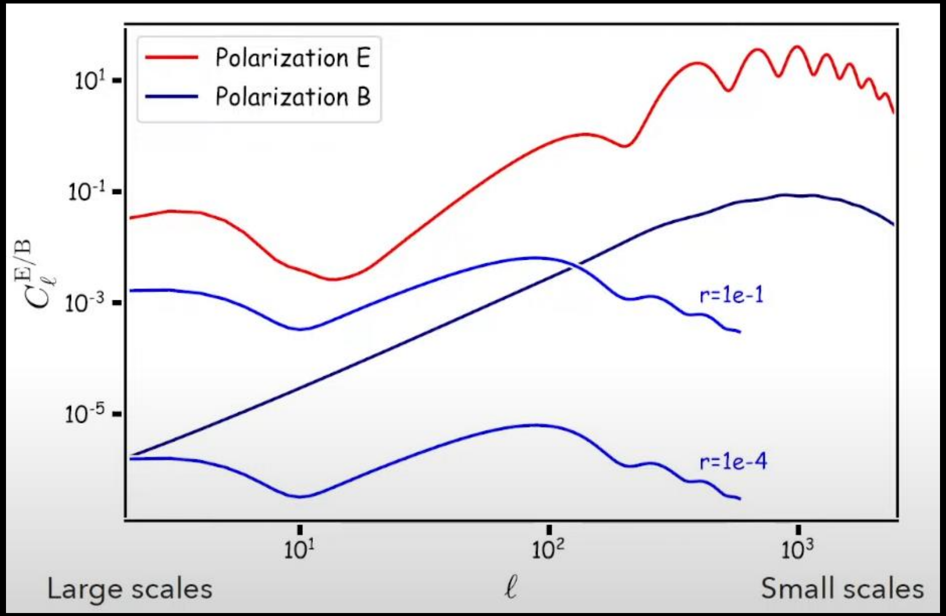
Tensor to scalar ratio r





- A key parameter for understanding the physics of inflation: Represents the relative strength of primordial gravitational waves (tensor perturbations) to density fluctuations (scalar perturbations)

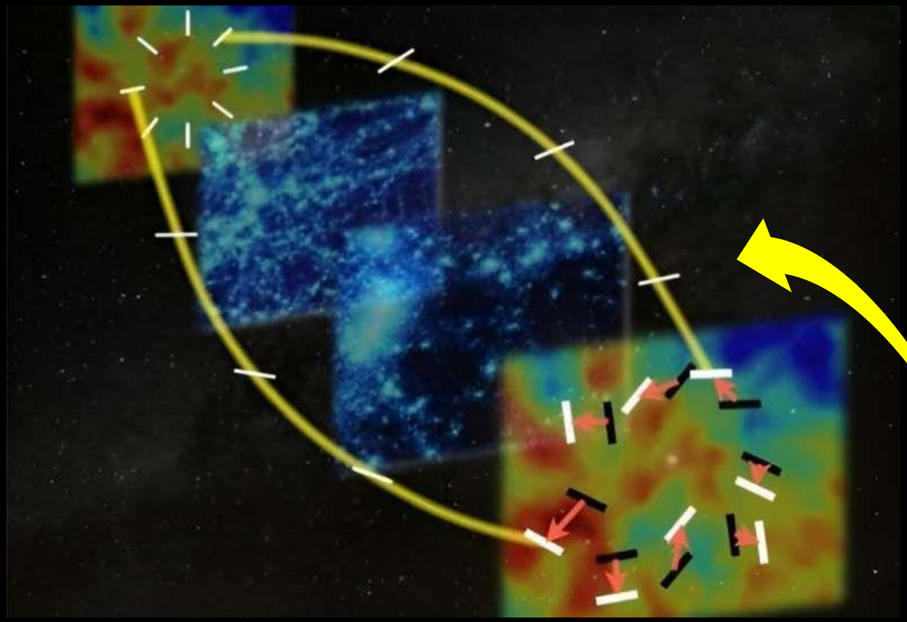
- Provides insights into the energy scale of inflation and the early universe's dynamics



Cosmology with gravitational lensing of the Cosmic Microwave Background, Louis Legrand, ICTP/SAIFR - IFT/Unesp

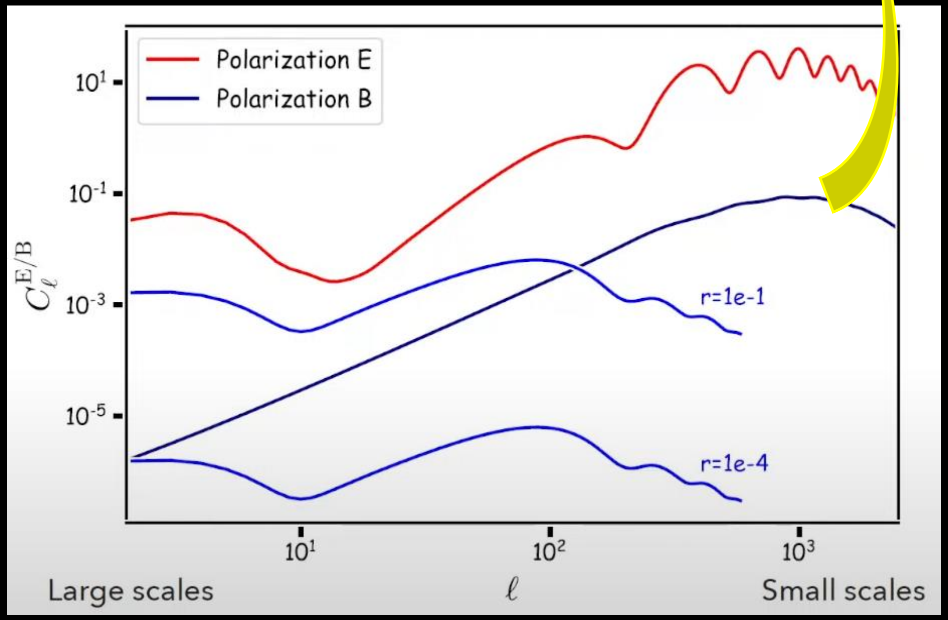
Physical Review D (2023). DOI:
10.1103/PhysRevD.108.063525

Tensor to scalar ratio r



- A key parameter for understanding the physics of inflation: Represents the relative strength of primordial gravitational waves (tensor perturbations) to density fluctuations (scalar perturbations)

- Provides insights into the energy scale of inflation and the early universe's dynamics



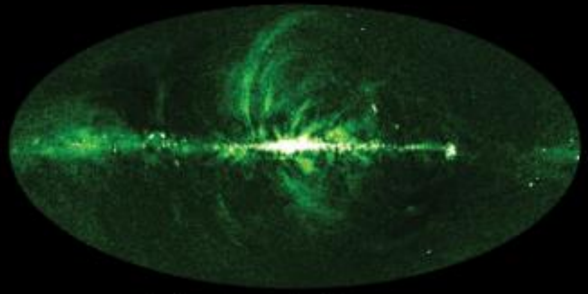
Cosmology with gravitational lensing of the Cosmic Microwave Background, Louis Legrand, ICTP/SAIFR - IFT/Unesp

Challenges in CMB analysis

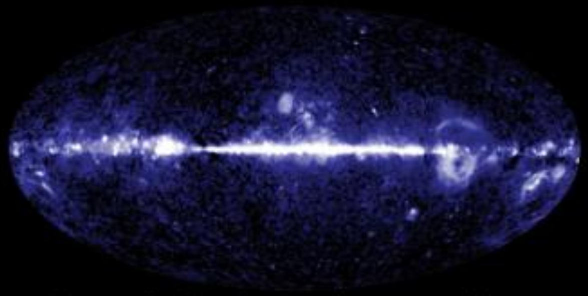
- The **signal is extremely weak**, requiring precise instruments and data analysis techniques

- **Galactic Foregrounds:**
  - **Synchrotron Radiation:** Emitted by relativistic electrons spiraling in the Galactic magnetic field, this radiation is dominant at low frequencies (~10-100 GHz).
  - **Thermal Dust Emission:** Dust grains heated by starlight emit radiation, dominant at high frequencies (~100-1000 GHz). This emission can mimic or overshadow the CMB signal.
  - **Free-Free Emission:** Electrons interacting with ions emit this radiation, adding to the contamination, especially at low frequencies.

Polarised dust emission



Polarised synchrotron emission



Line radiation from carbon monoxide gas

Image credit: ESA and the Planck Collaboration

Challenges in CMB analysis

- **Foreground components vary with frequency**, necessitating observations at multiple frequencies to disentangle the CMB from foregrounds.

- **Instrumental Noise and Systematics:**
  + **Photon Noise:** arising from the quantized nature of light. The noise is inherent to signal due to the random arrival times of photons. Photon noise is especially significant in the low signal regime of the CMB, where the photon count is low, making it a critical factor in the overall noise determination.
  + **Detector Noise:** This noise comes from the instrument's detectors themselves, including thermal noise, electronic noise, and other internal sources.

- **Atmosphere**

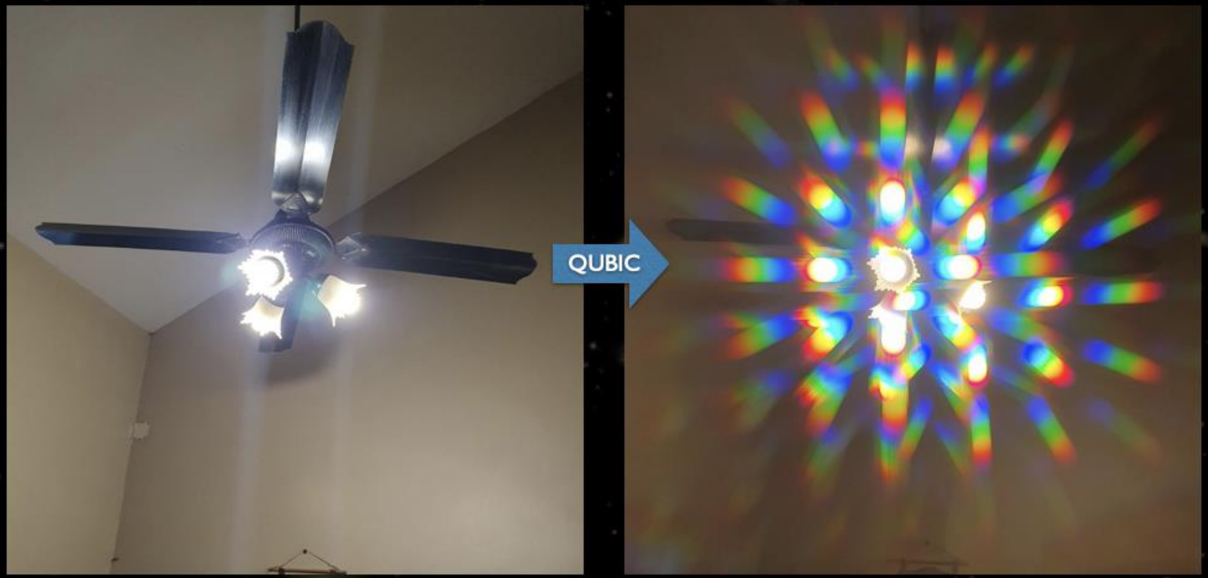
QUBIC

The **Q** & **U Bolometric Interferometer** for Cosmology

QUBIC

- unique instrument combining interferometry and bolometry to observe the CMB, using an array of bolometers to measure temperature and polarization fluctuations

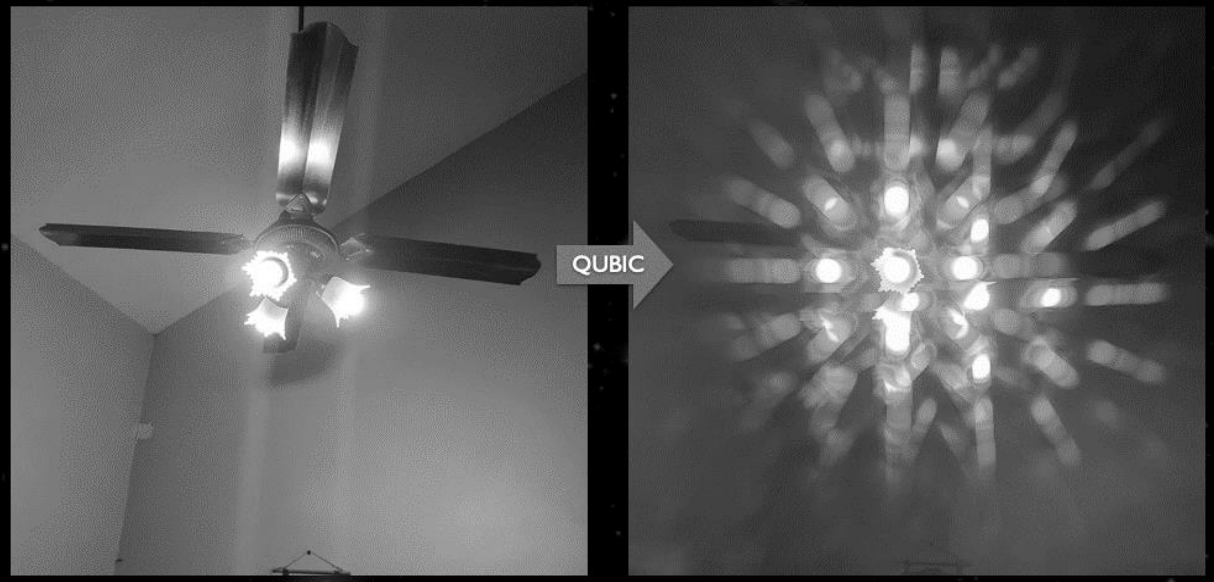
- QUBIC produces multiple peaks due to its interferometric setup: Spectral splitting of light enables the separation of different frequency components, aiding in component separation



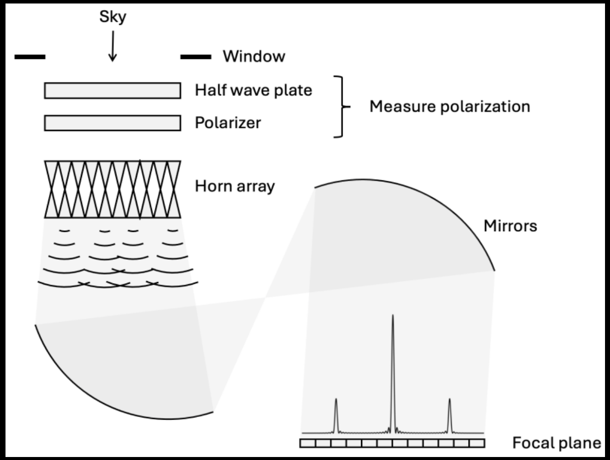
QUBIC

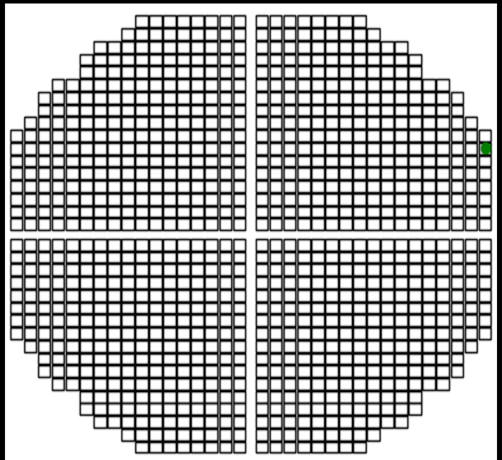
- unique instrument combining interferometry and bolometry to observe the CMB, using an array of bolometers to measure temperature and polarization fluctuations

- QUBIC produces multiple peaks due to its interferometric setup: Spectral splitting of light enables the separation of different frequency components, aiding in component separation

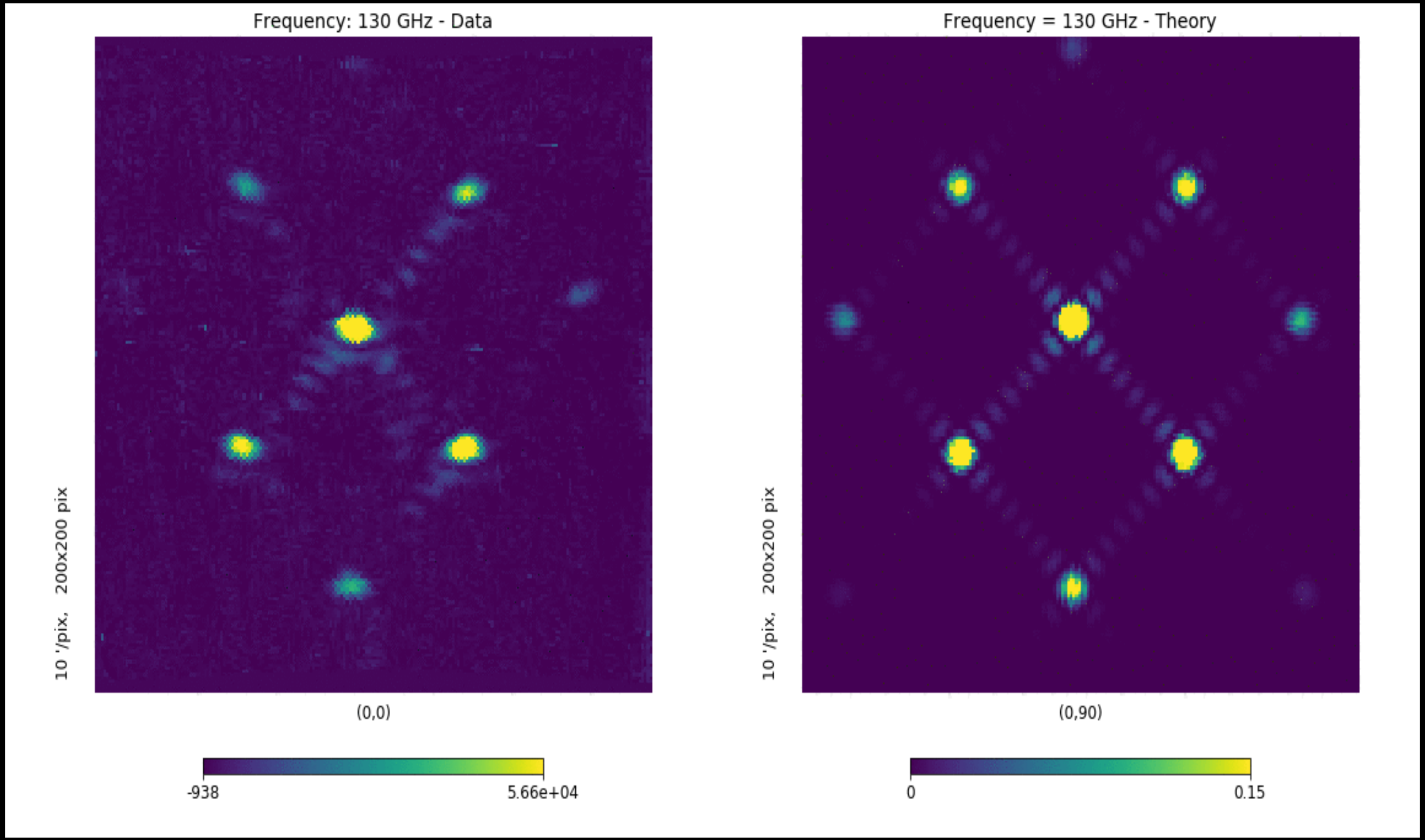


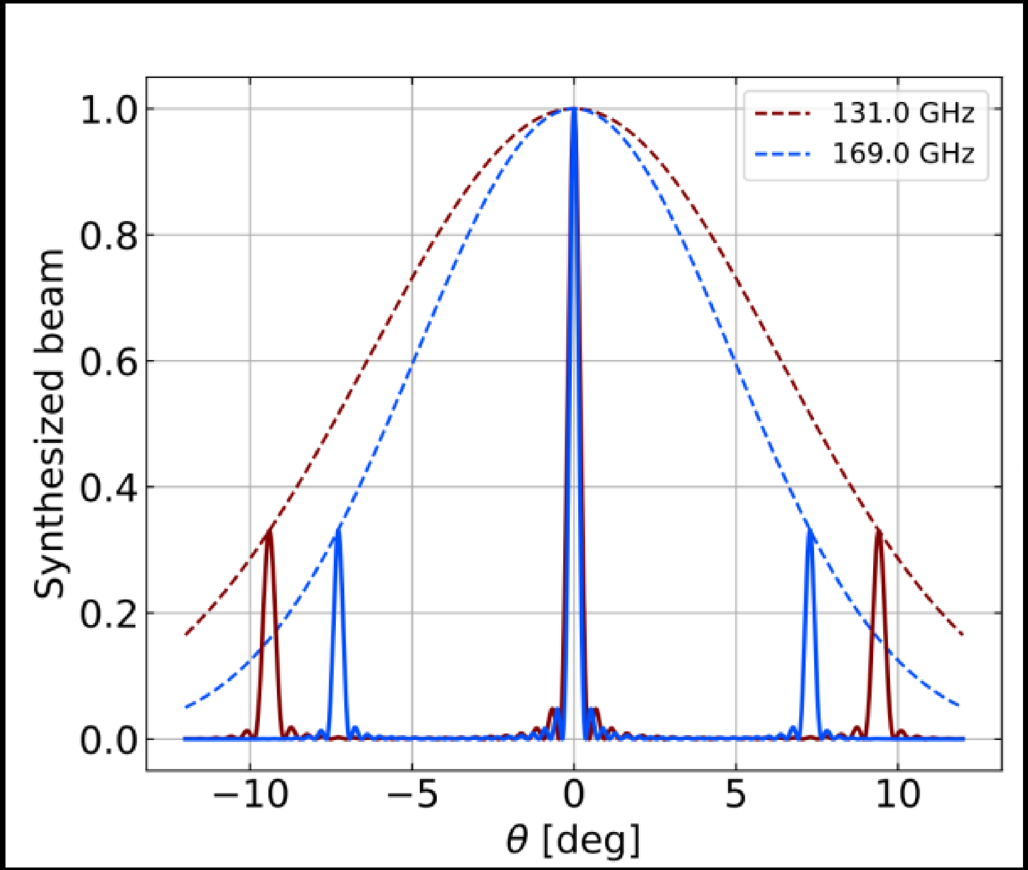
QUBIC





Battistelli, E.S., Ade, P., Alberro, J.G. *et al.* QUBIC: The Q & U Bolometric Interferometer for Cosmology. *J Low Temp Phys* 199, 482–490 (2020). https://doi.org/10.1007/s10909-020-02370-0





Neural Networks in CMB Analysis

- Choosing appropriate layers (e.g., convolutional, recurrent, graph-based) depending on the data structure
- Using layers that incorporate physical knowledge (e.g., known operators)
- Operators like rotation, scaling, and convolution should be well-understood and correctly implemented based on problem specifics and used structures
- Proper architecture design ensures that the network can learn complex relationships without overfitting (due to unneccesary complexity) or underfitting (due to improper introduction of known operators)

Neural Networks in CMB Analysis

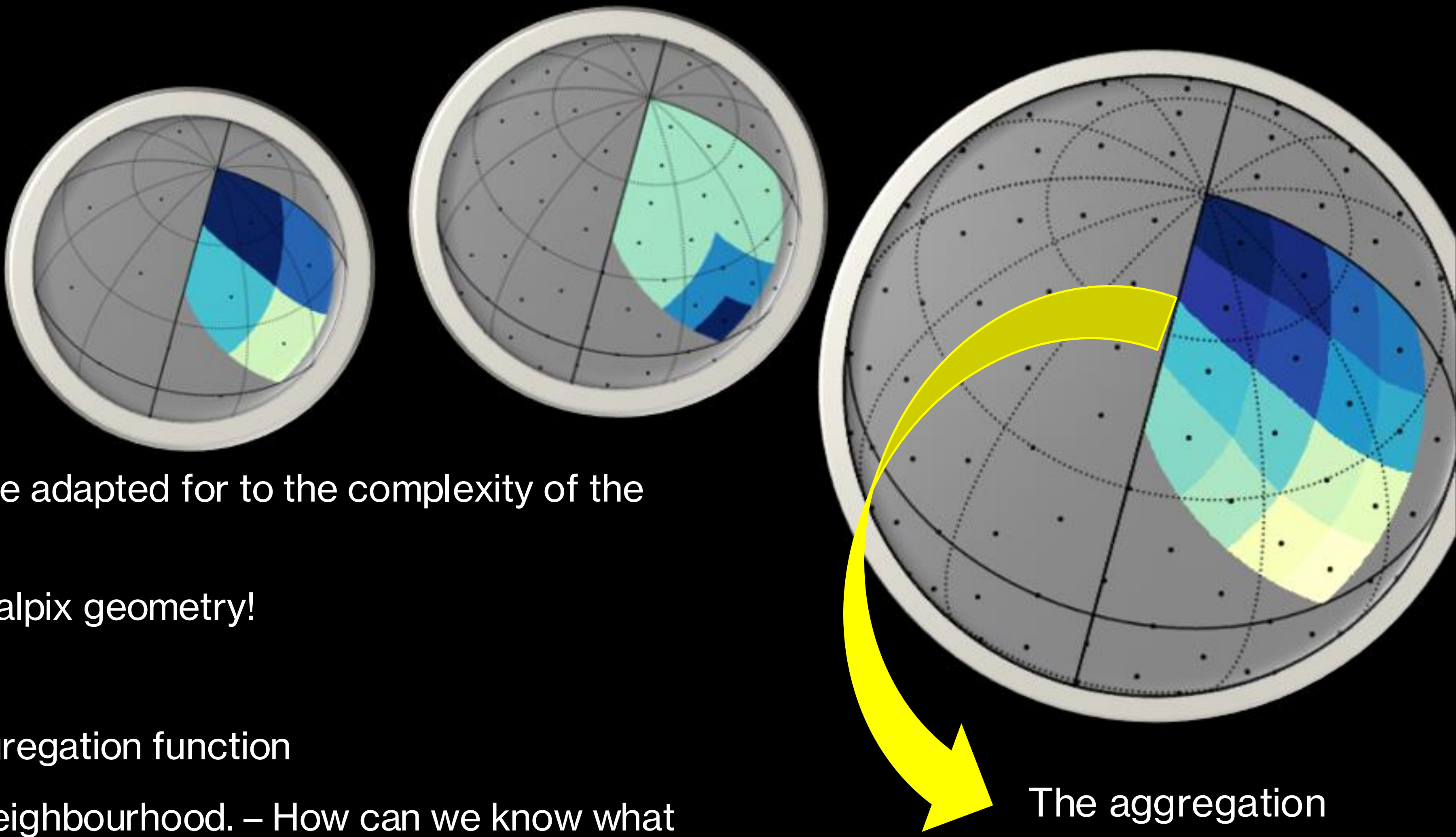
Case-specific issues:

Rotational Invariance: CMB data is inherently spherical, so the analysis must be invariant to rotations on the sphere

Handling Sparsity: Both the data and applied operators can be sparse! CMB observations often cover only parts of the sky, while operators are represented with sparse matrices (e.g. detector beam)

Particularities of the detector: Different instrument process data in different manners. The network has to represent the complex process inside QUBIC.

Adaptation



- Conventional layers need to be adapted for to the complexity of the data to maintain accuracy.
- Layers should respect the Healpix geometry!

The heart of any layer is the aggregation function

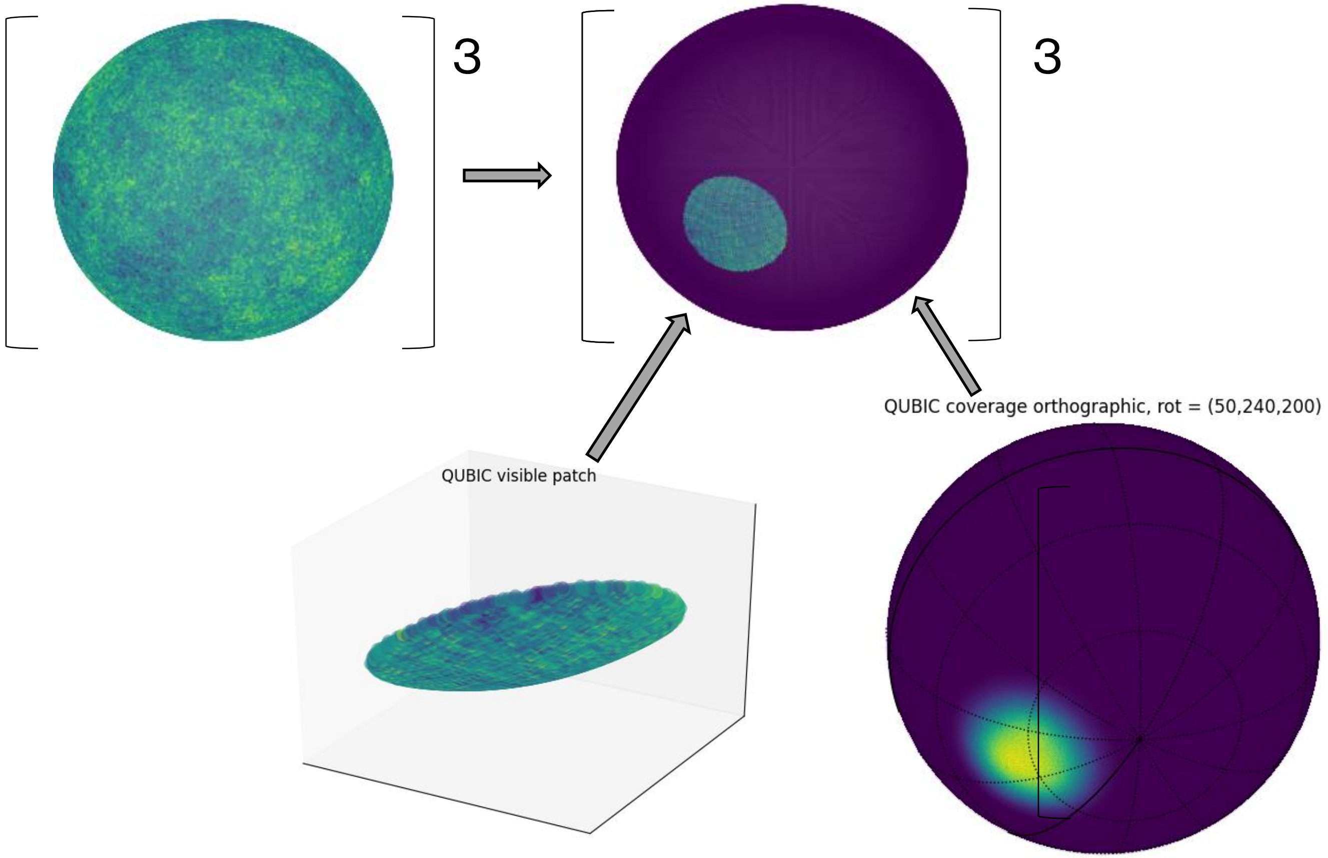
Calls for proper integration of neighbourhood. – How can we know what is proper?

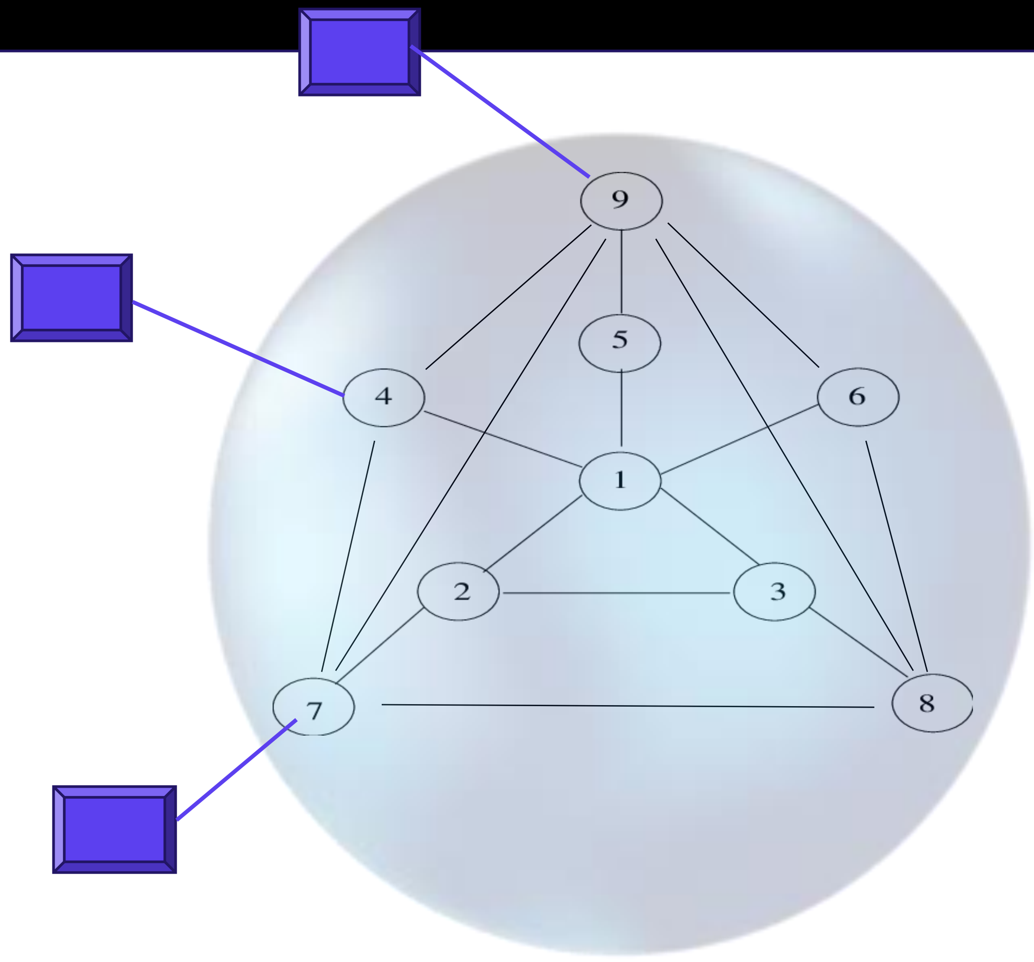
The analysts choose this based on the specifics of the attempted solution!

The aggregation function determines the operation, but will treat patches as Euclidian in default!!

Data from different sources and types can be represented in a unified structure; naive handling of multifeature data can easily lead to discrepancies!

Multifeature data

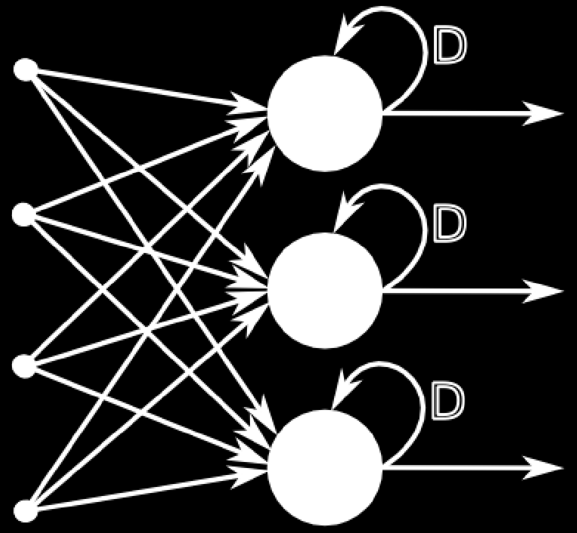


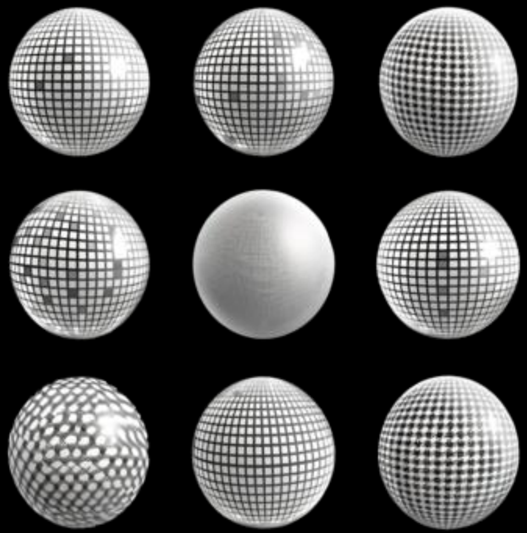


Data from different sources and types can be represented in a unified structure; naive handling of multifeature data can easily lead to discrepancies!

Multifeature data

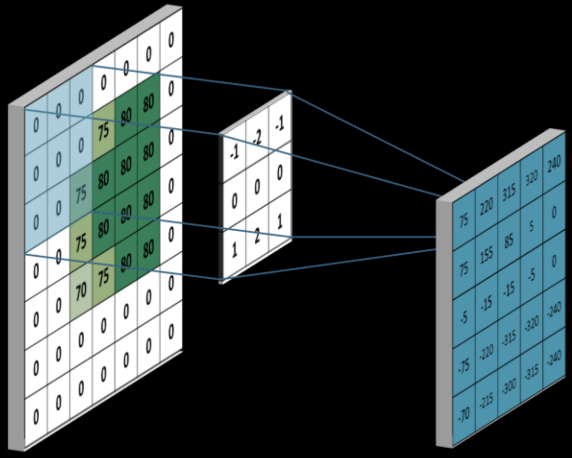
LSTM
Dana type: <u>temporal, sequential</u>
Used for: <u>processing of TOD, frequency dependance</u>

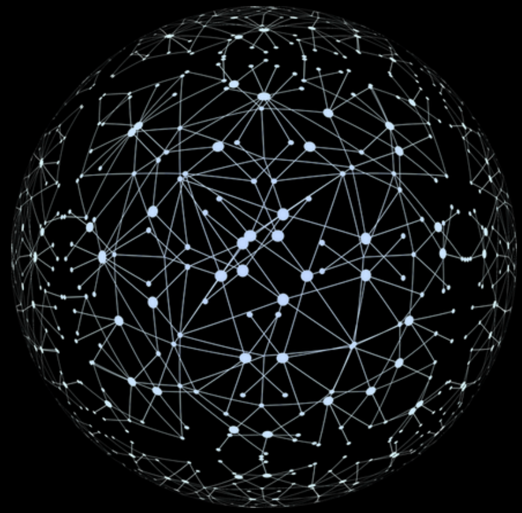




Spherical NN
Dana type: <u>spherical maps</u>
Used for: <u>processing intensity maps</u>

Convolution NN
Dana type: <u>2D projections of maps</u>
Used for: <u>finding structure in small areas, local feature extraction</u>





Graph CN
Dana type: <u>spherical maps with additional features and relationships</u>
Used for: <u>complex and complete studies of maps</u>

The problem

$$d = \mathbf{H}s$$

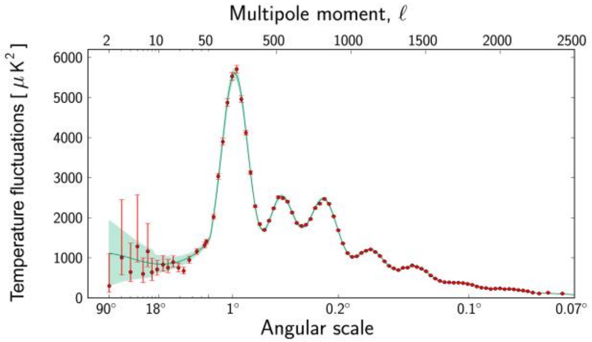
The problem

$$d = \mathbf{H}s$$

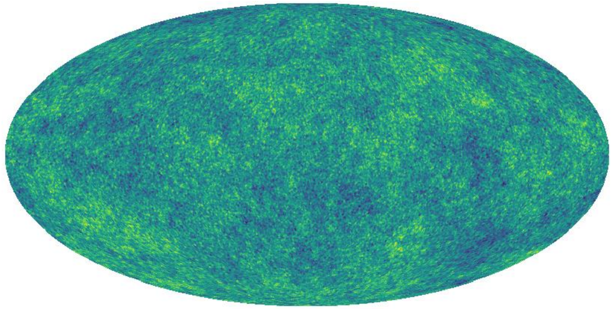
$$d = \mathbf{R}\ \mathbf{T}\ \mathbf{I}\ \mathbf{P}\ (\mathbf{HW}\ \ \mathbf{Proj})\ \mathbf{F}\ \mathbf{A}\ \mathbf{T}\ \mathbf{U}\ s$$

The framework

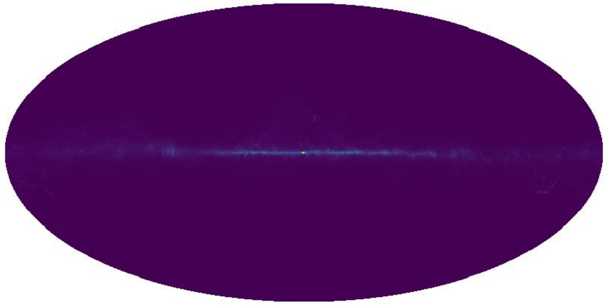
r $\xrightarrow{\text{CAMB}}$

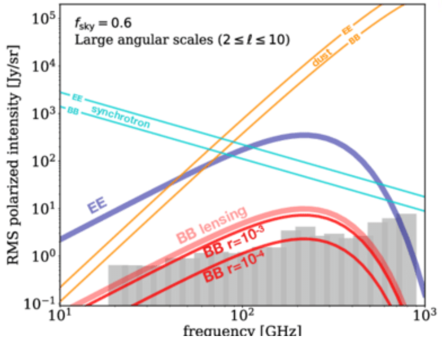


hp

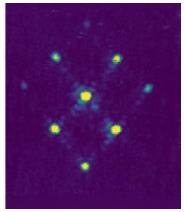


pysm



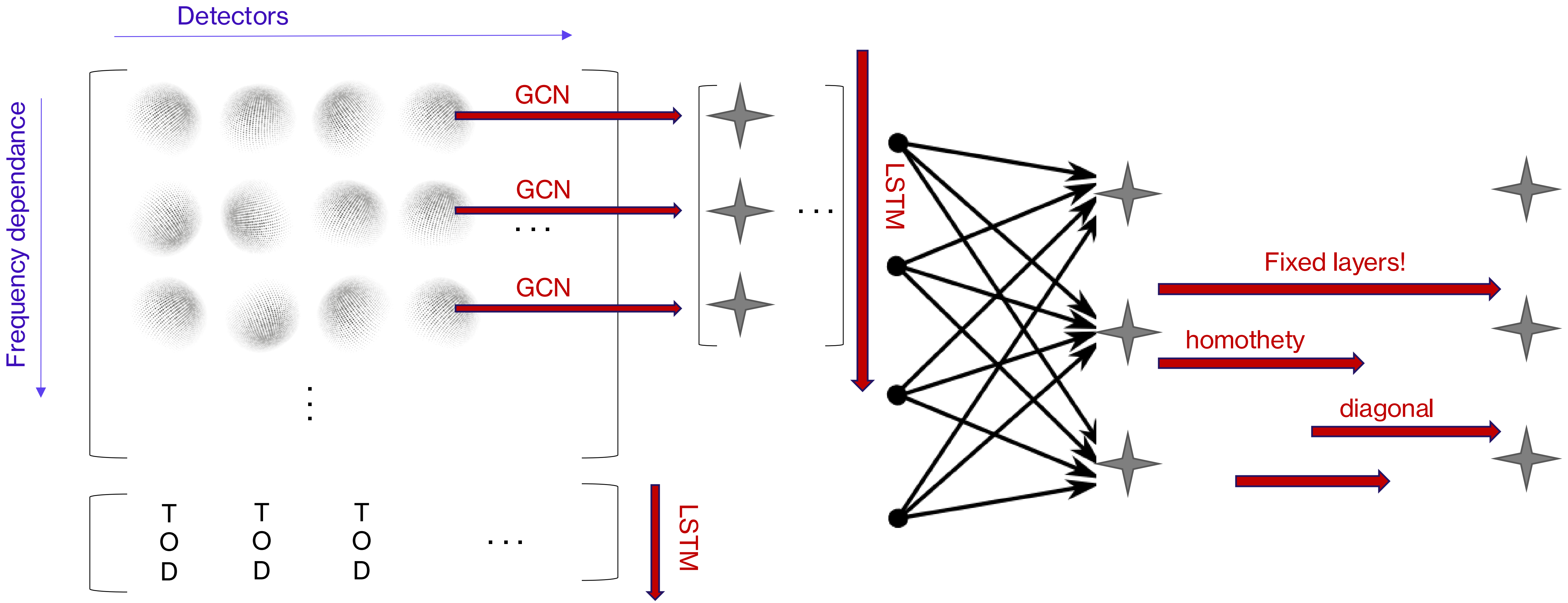


T
O
D

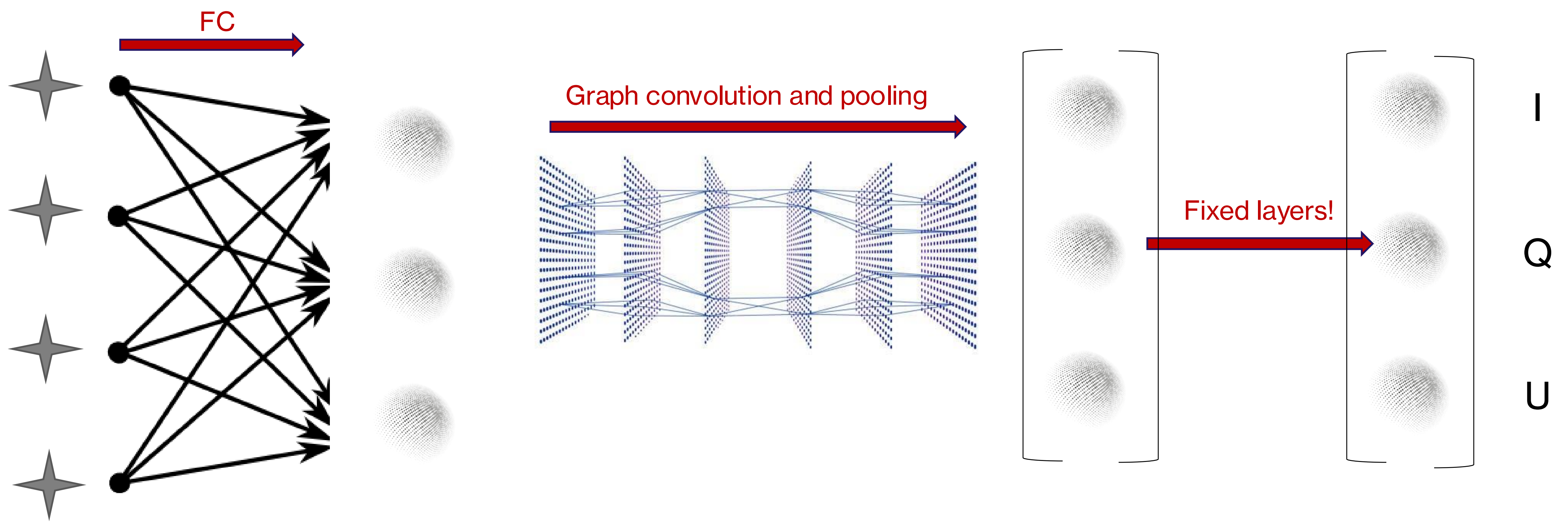


8/19/2024

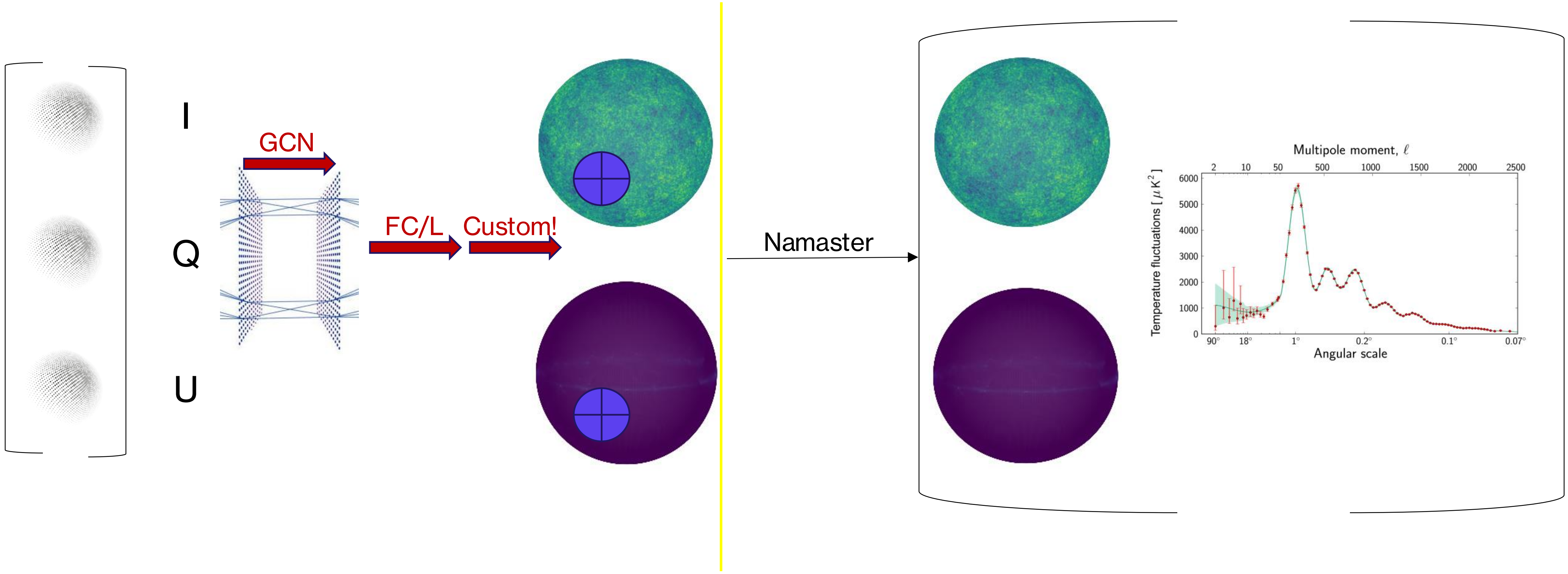
The framework



The framework



The framework



PINNs

- integrating domain knowledge directly into the learning process.

Always ensure that the network's predictions align with known physical principles.

In CMB this can be...

- spectral characteristics of different components (e.g., CMB, dust, synchrotron).
- following the physical properties of polarization (example: Q and U stokes!).
- any response function from the instrument.

PINNs

NETWORK LAYER

- Versatile, modular, captures complex relationships
- Can capture relationships far from ones predicted by the model
- More difficult to interpret

--> Fitting is possible, requires careful consideration of the created kernels and their weights

LOSS FUNCTION

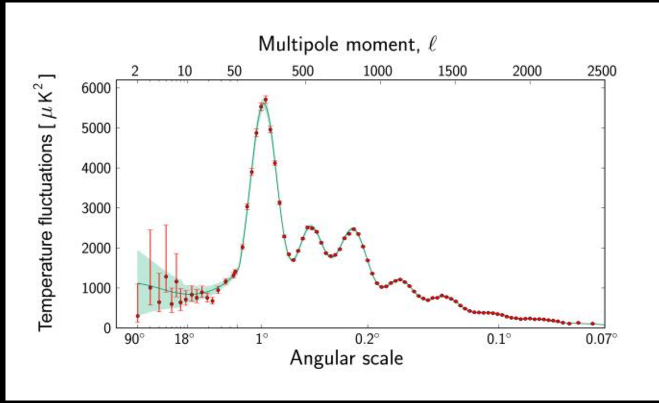
- Best for estimation of parameters from given set of data
- Training data needs to be varied to allow generalization
- The defined loss function should correspond to the physical truth!
- Easier to implement and easier to interpret

--> Fitting is possible, requires handling of possible overfitting to training data

PINNs

Planck 2013 results. XI. All-sky model of thermal dust emission
Planck Collaboration, A&A 571, A11 (2014)
https://doi.org/10.1051/0004-6361/201323195

$$I_\nu = \tau_{\nu_0}\, B_\nu(T_{\rm obs}) \left(\frac{\nu}{\nu_0}\right)^{\beta_{\rm obs}}$$



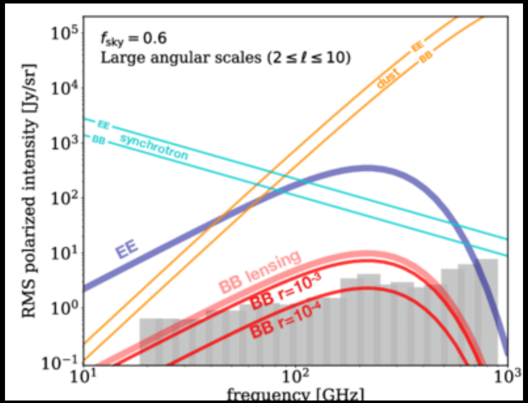
NETWORK LAYER

- Versatile, modular, captures complex relationships
- Can capture relationships far from ones predicted by the model
- More difficult to interpret

--> Fitting is possible, requires careful consideration of the created kernels and their weights

```python
class ModifiedBlackBodyLayer(nn.Module):
    def __init__(self, nu_0, initial_temp=20.0, initial_beta=1.5):
        super(ModifiedBlackBodyLayer, self).__init__()
        self.nu_0 = nu_0
        self.temp = nn.Parameter(torch.tensor(initial_temp, dtype=torch.float32))
        self.beta = nn.Parameter(torch.tensor(initial_beta, dtype=torch.float32))
        self.amplitude = nn.Parameter(torch.tensor(1.0, dtype=torch.float32))


    def forward(self, nu):
        h = 6.626e-34
        k = 1.381e-23
        c = 3.0e8
        exponent = h * nu / (k * (self.temp))
        black_body = (nu**3) / (torch.exp(exponent) - 1)
        power_law = (nu / self.nu_0) ** self.beta
        return self.amplitude * power_law * black_body
```



LOSS FUNCTION

- Best for estimation of parameters from given set of data
- Training data needs to be varied to allow generalization
- The defined loss function should correspond to the physical truth!
- Easier to implement and easier to interpret

--> Fitting is possible, requires handling of possible overfitting to training data

In conclusion...

Leonora Kardum

LABORATOIRE ASTROPARTICULE & COSMOLOGIE

for HAMLET Conference

August, 2024 in Copenhagen