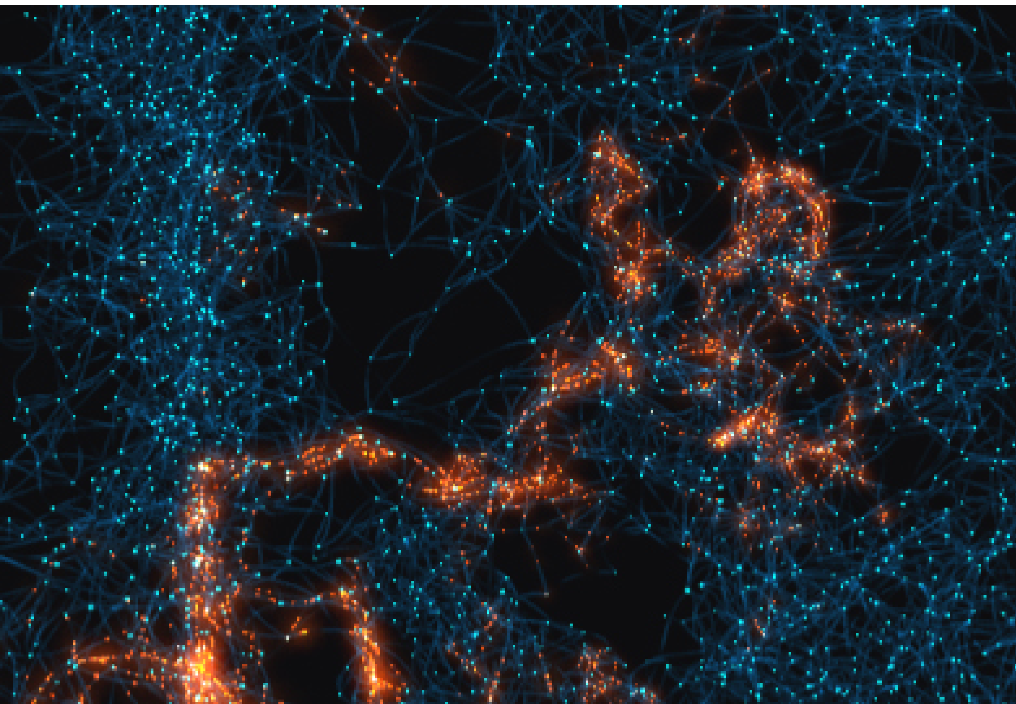
Deep Learning in Astrophysics



Artist representation of a neural network

Large Scale structure of our universe

Marcus Wirtz, Martin Erdmann,
Jonas Glombitza, David Walz

RWTH Aachen, III. Physikalisches Institut A

Astroteilchenschule - October 9th, 2017

Cat or dog?

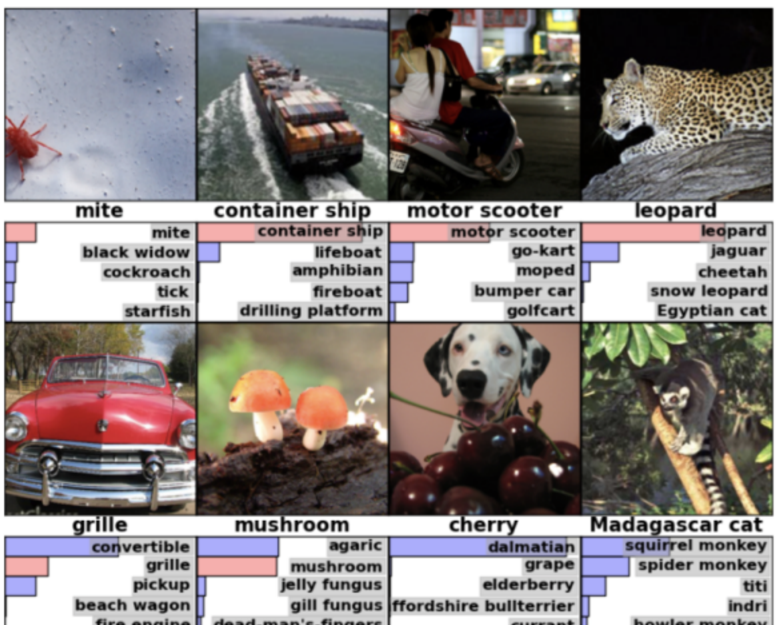
- ✗ Very easy for humans, very hard for machines
- ✗ **Challenges**: High dimensional input (10^4 - 10^7) & Variations in image

Marcus Wirtz – mwirtz@physik.rwth-aachen.de

Image recognition

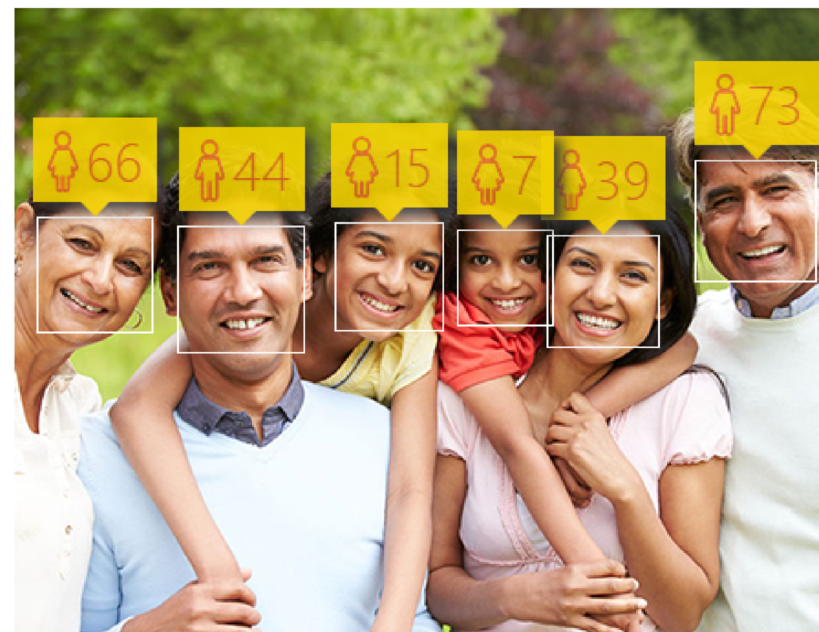
IMAGENET

- 20,000 object classes (categories), 14 Mio images
- 3% error rate (human: 5%)



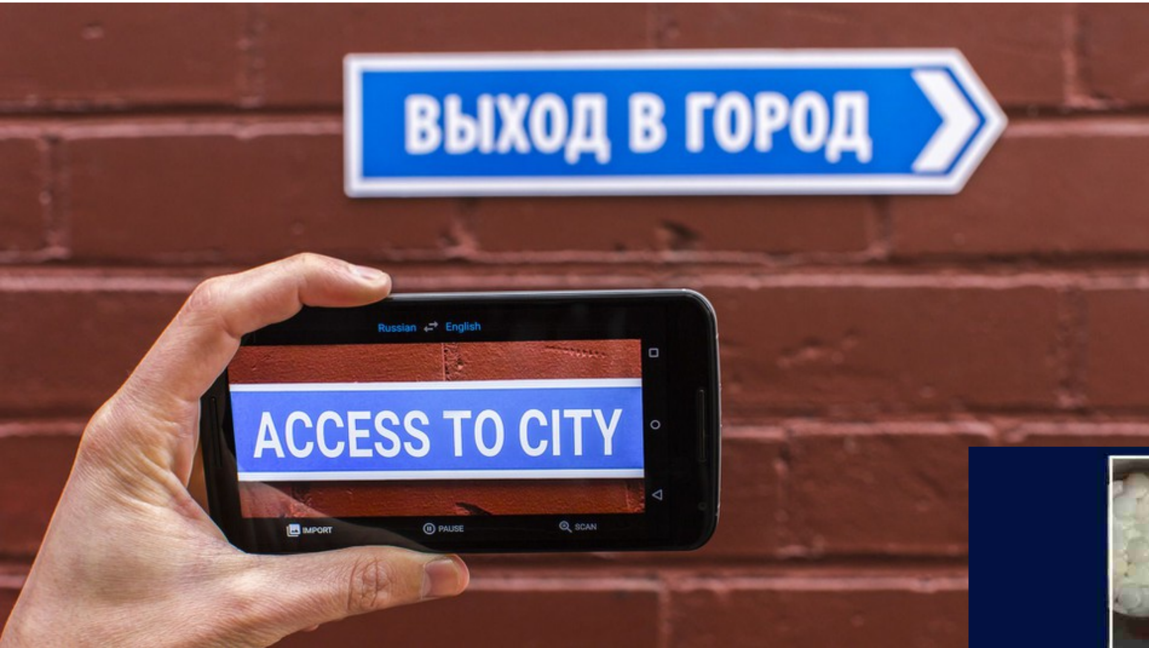
facebook

- Face recognition
- e.g. AmazonGo: Supermarket without cashiers

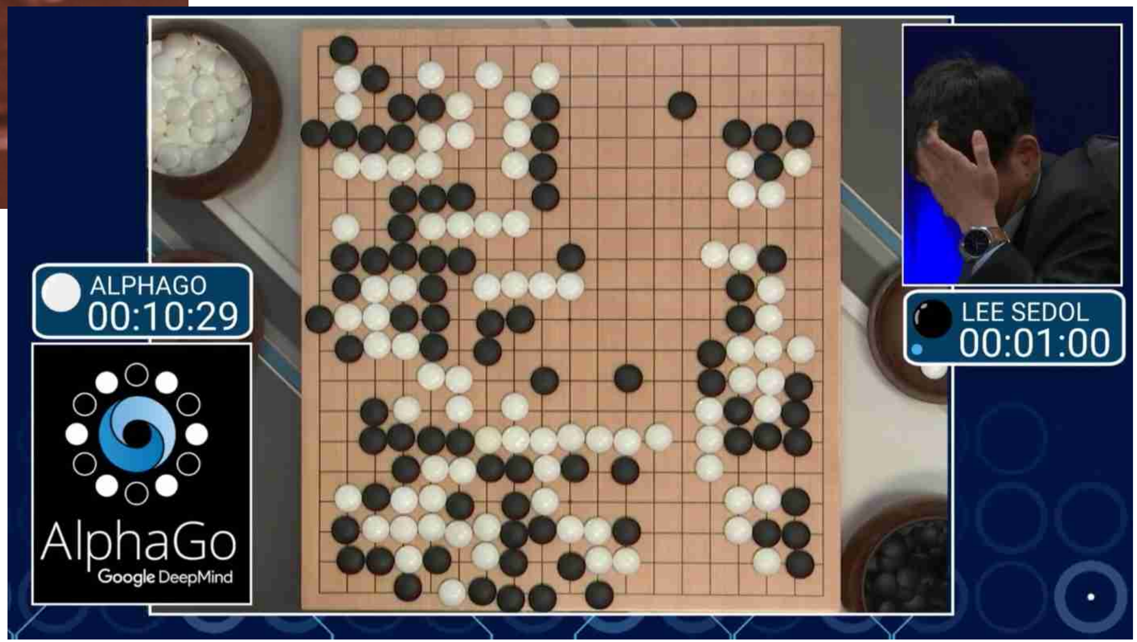


Marcus Wirtz – mwirtz@physik.rwth-aachen.de

Not only image recognition...



- ✗ **Unsupervised learning**
  - – learns by “try and error”
  - – autonomous cars
  - – AlphaGo: neural network beats world champion in game Go

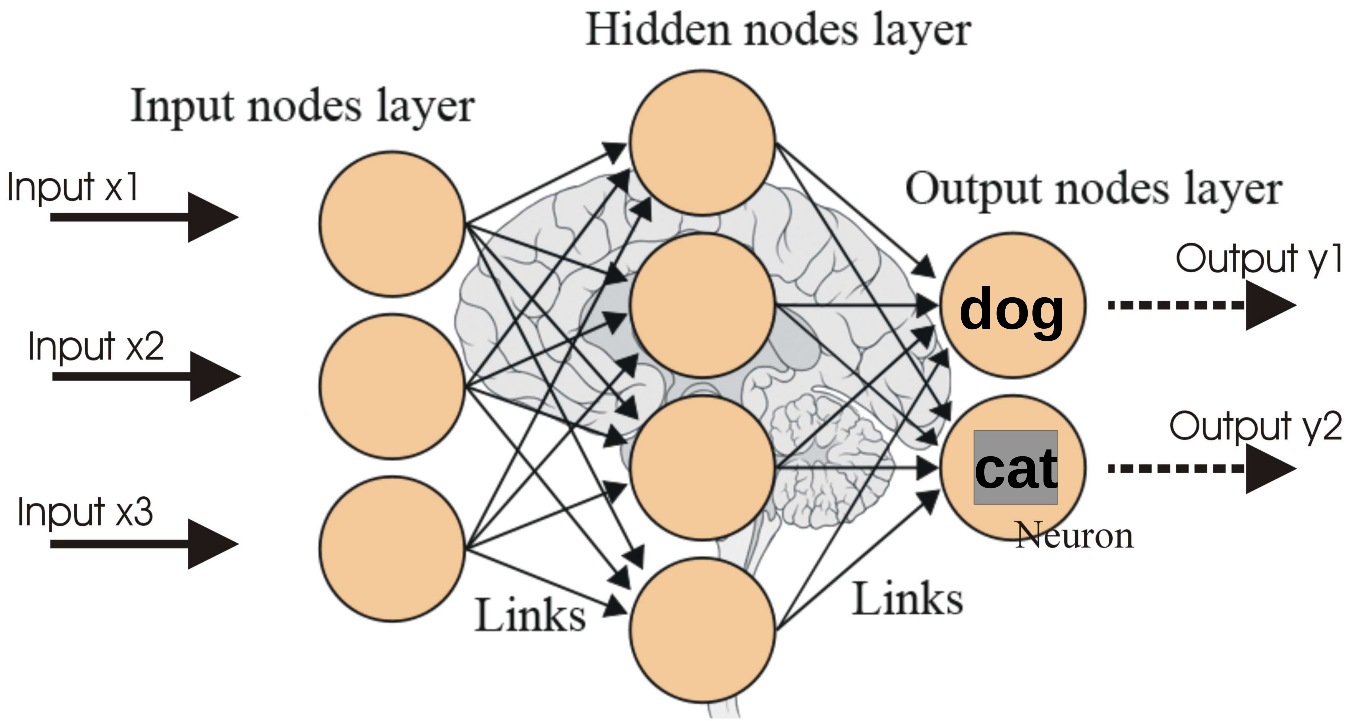


- ✗ **Recurrent neural networks**
  - – Sequence of data (history)
  - – e.g. translation, text-, audioprocessing

Marcus Wirtz – mwirtz@physik.rwth-aachen.de

Neural networks

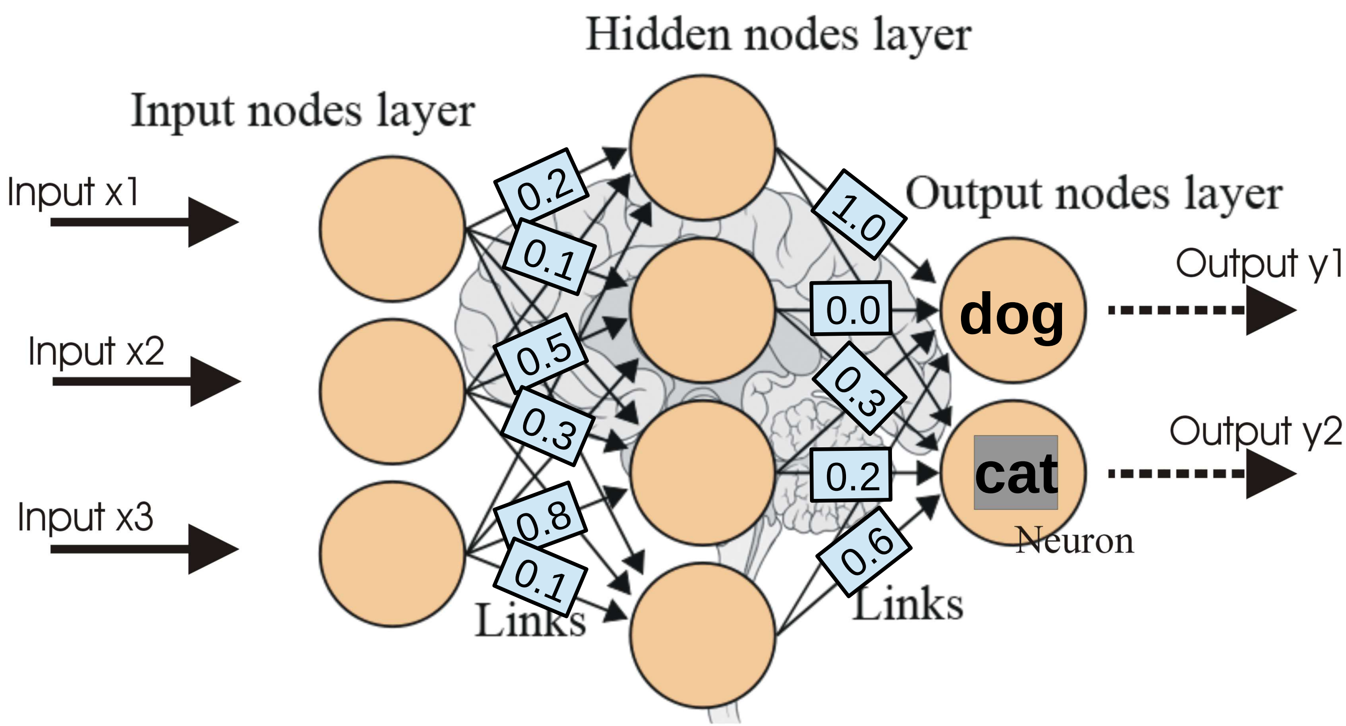
$$\text{Output} = f(\text{Input})$$



Marcus Wirtz – mwirtz@physik.rwth-aachen.de

Neural networks

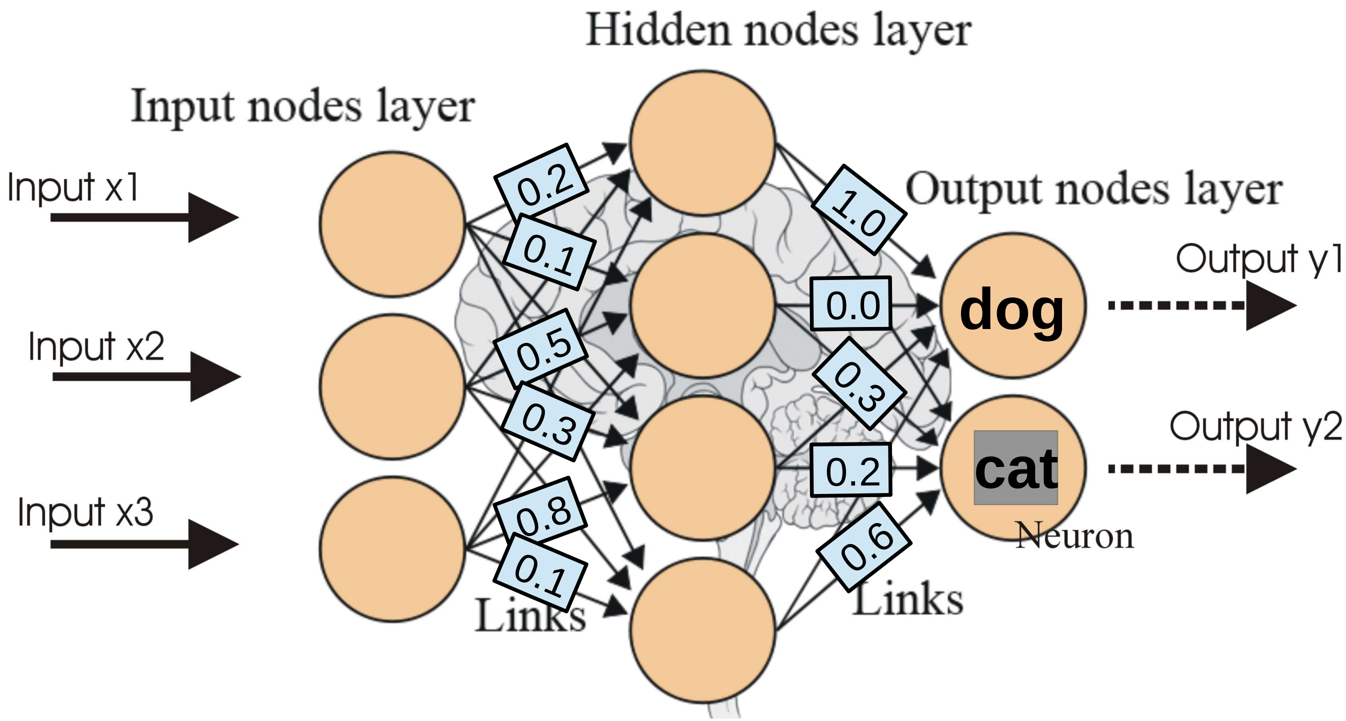
Random initialization...



Marcus Wirtz – mwirtz@physik.rwth-aachen.de

Neural networks

Let's try a cat...



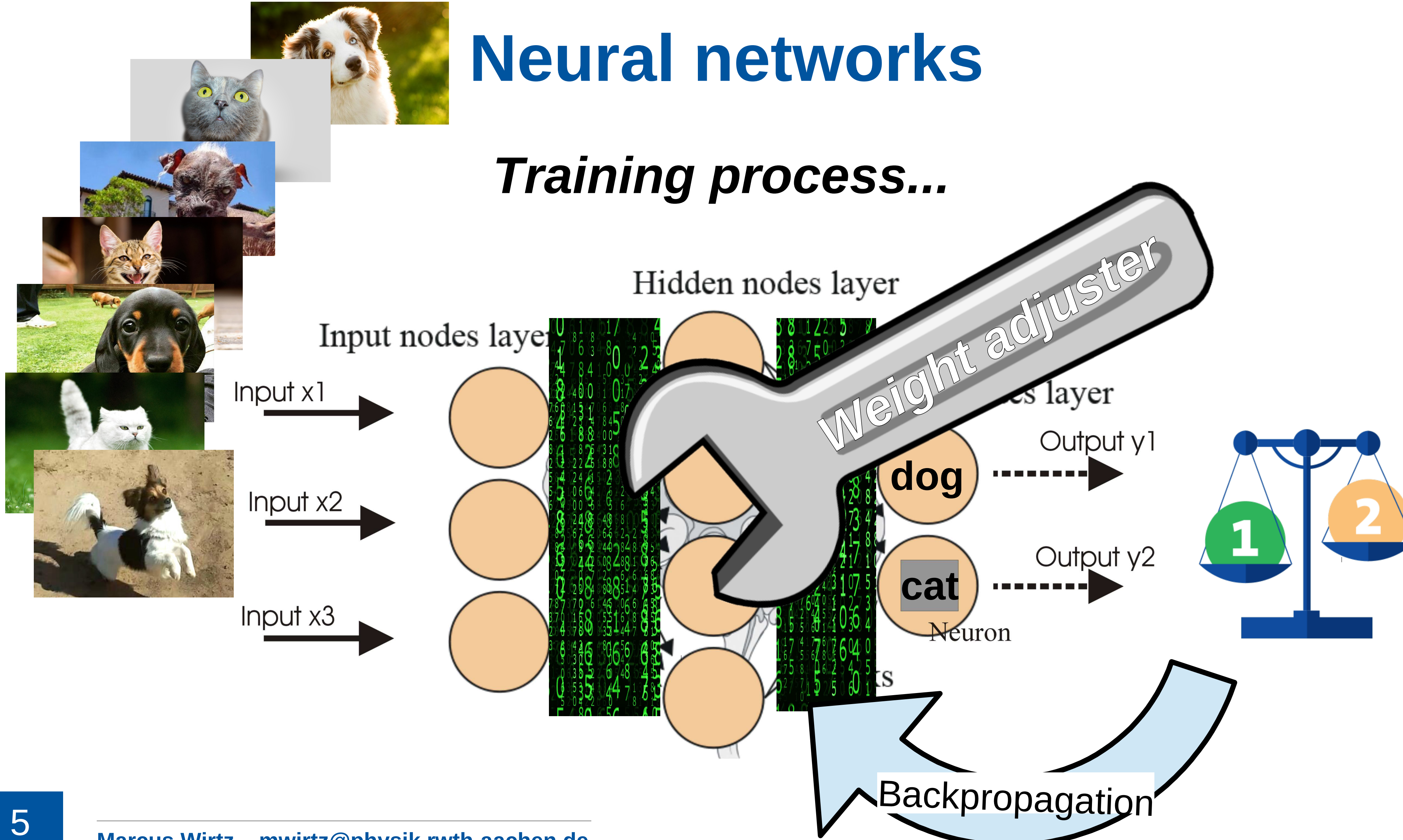
52 %

48 %

Marcus Wirtz – mwirtz@physik.rwth-aachen.de

Neural networks

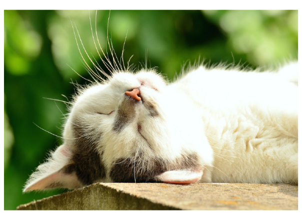
Training process...

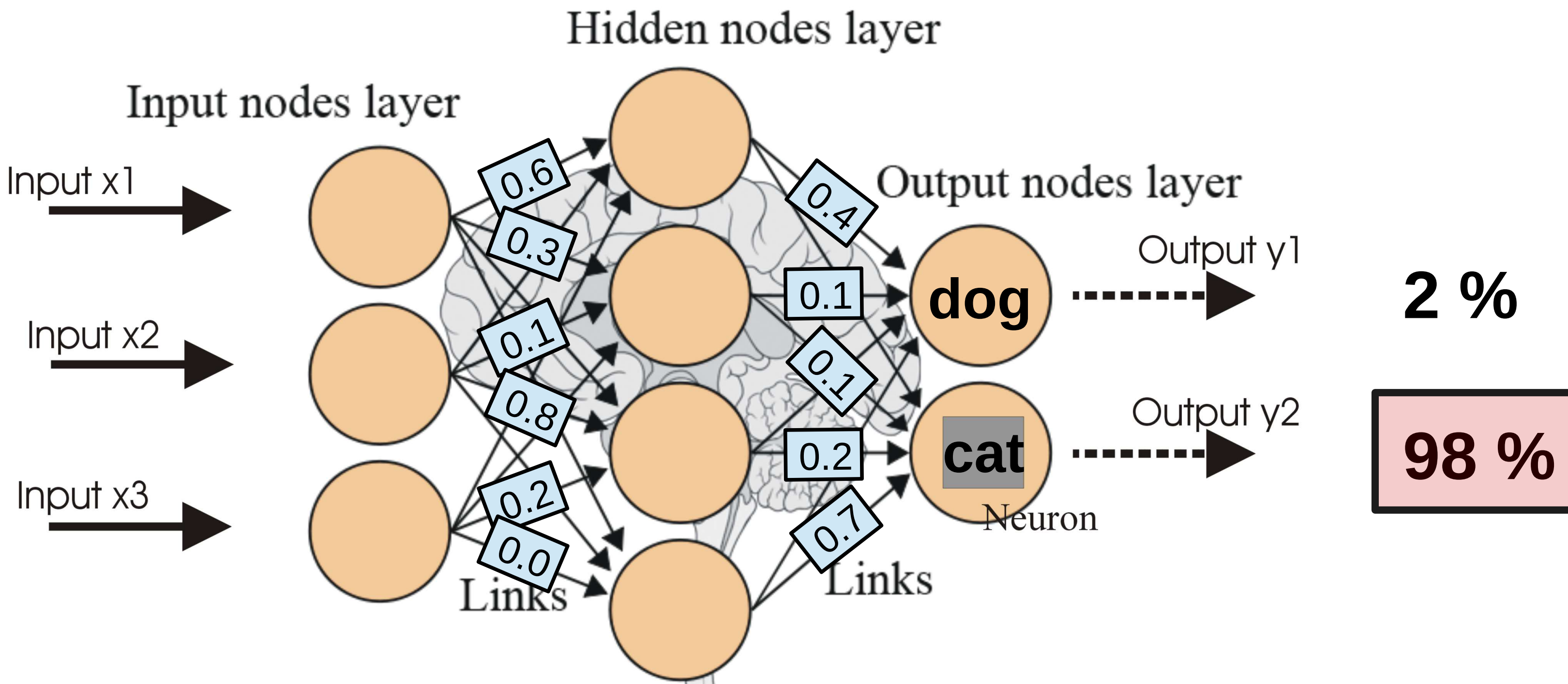


Marcus Wirtz – mwirtz@physik.rwth-aachen.de

Neural networks

Try again...

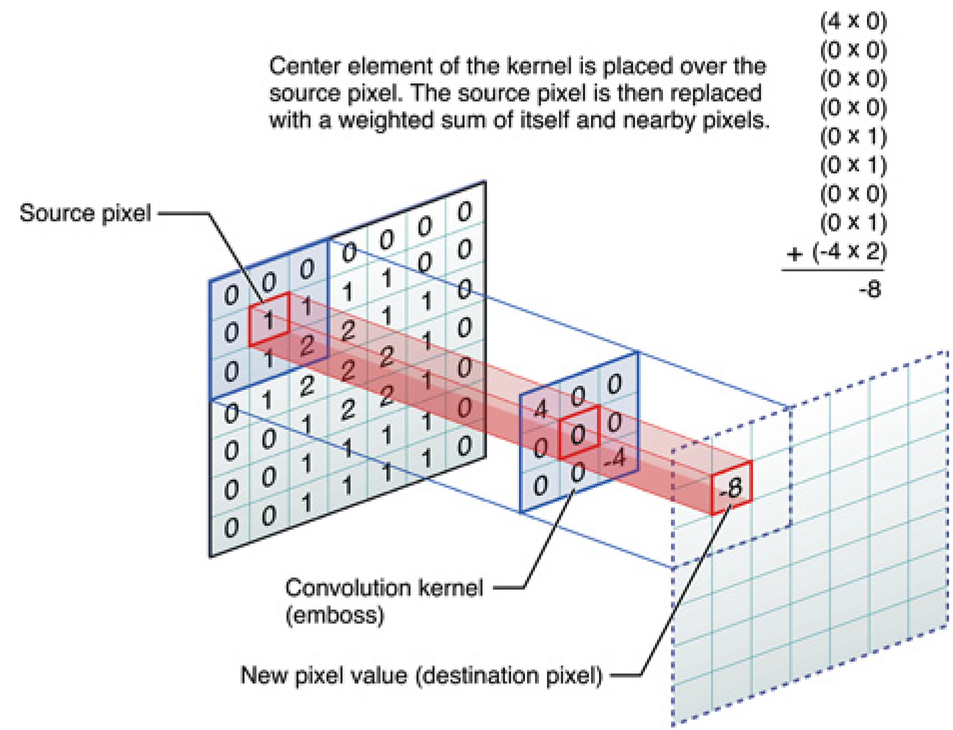


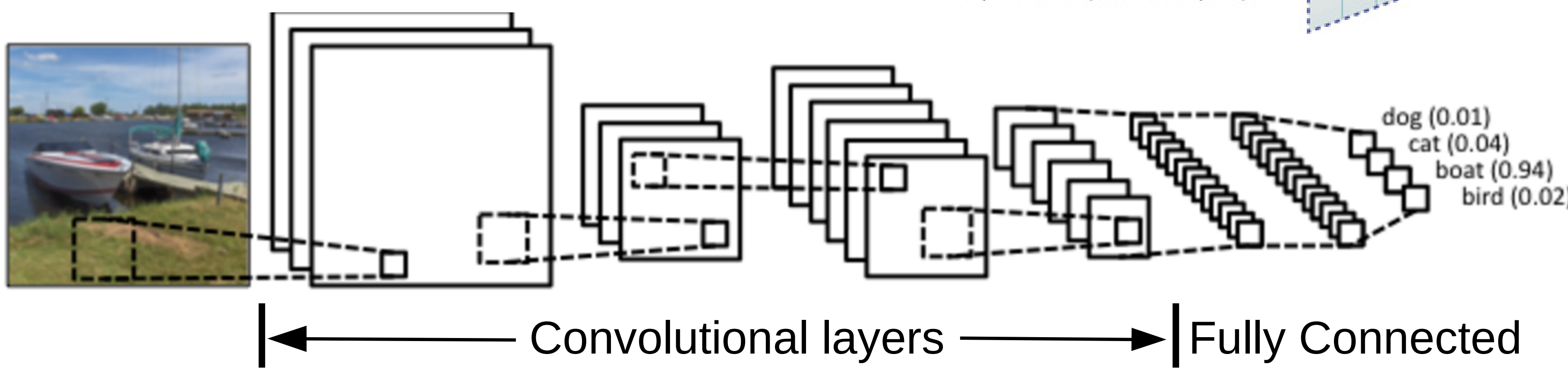


Marcus Wirtz – mwirtz@physik.rwth-aachen.de

Convolutional neural network

- ✗ Used in **image recognition**
- ✗ Looking for local correlations (substructures in image)
- ✗ Kernel consists of trainable parameters
- ✗ Fully connected network at end: classifies extracted features

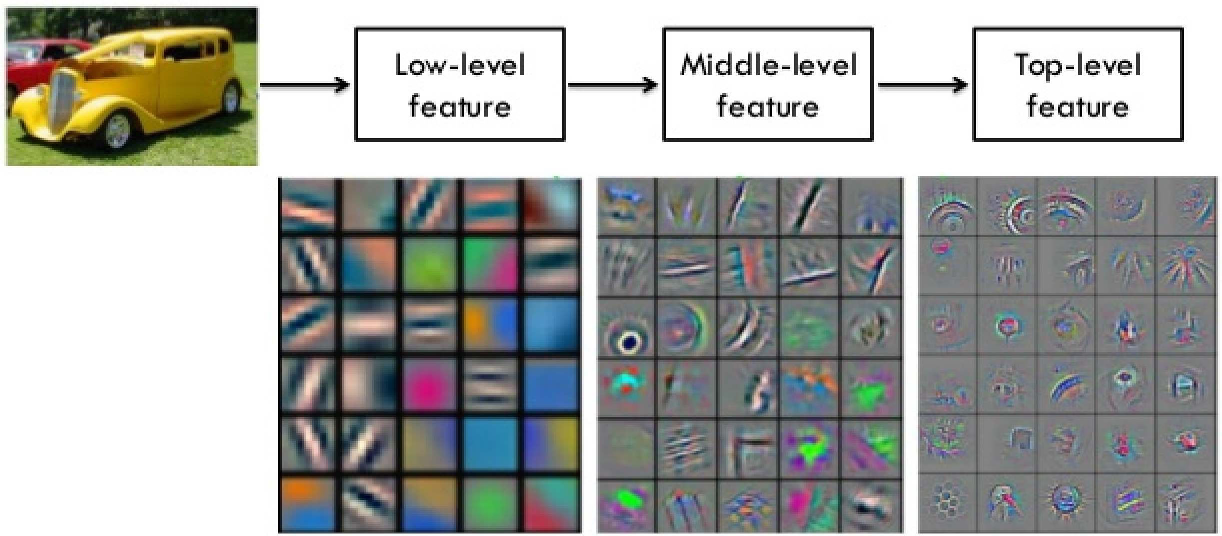




Convolutional neural network

- ✗ Visualize feature maps for image
- ✗ Stacking several convolutions: Extract Features of different hierachie
- ✗ Low-level feature: Edges, gradients
- ✗ Top-level feature: image specific structures visible

Hierarchy of trained representations

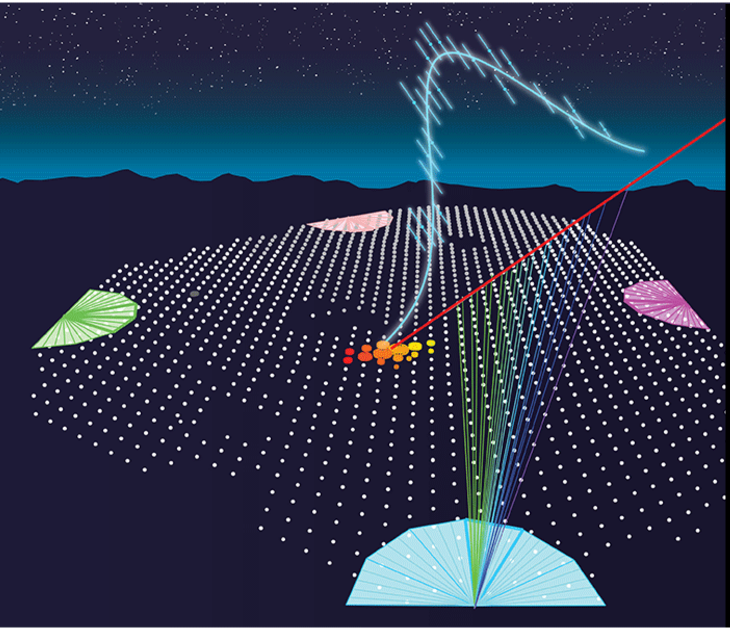


Feature visualization of convolutional net trained on ImageNet from [Zeiler & Fergus 2013]

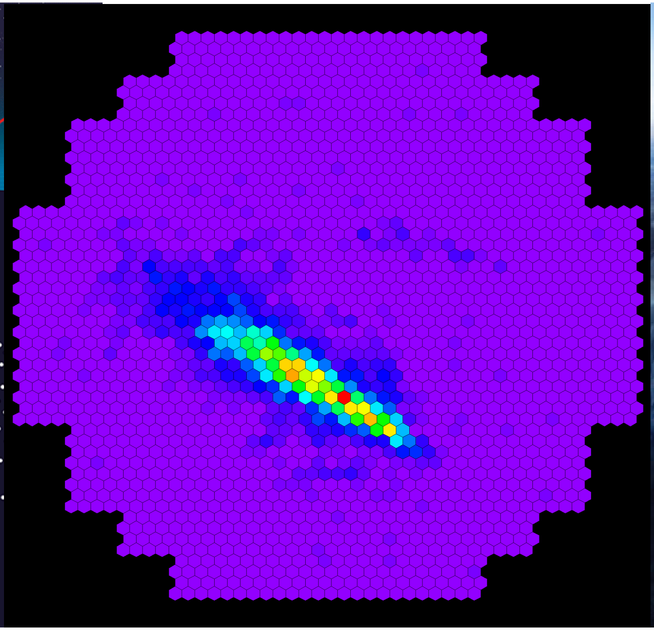
Marcus Wirtz – mwirtz@physik.rwth-aachen.de

Deep Learning in Astrophysics

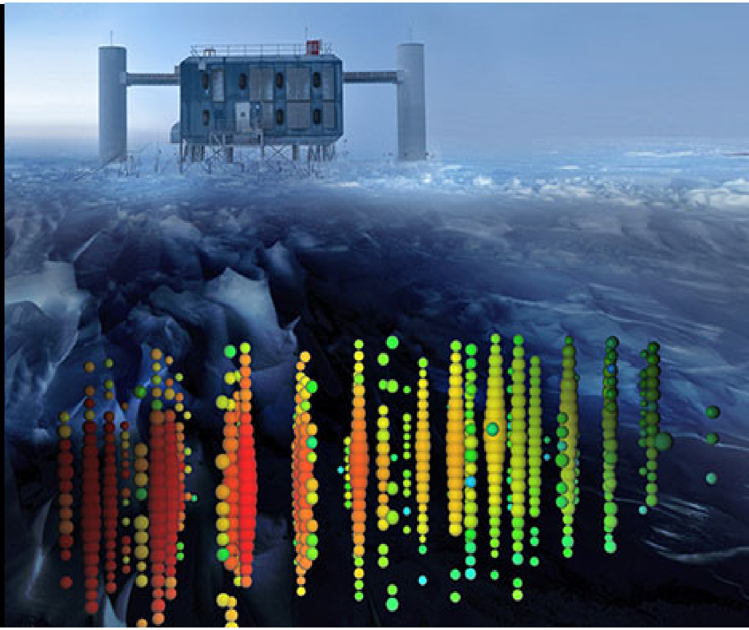
Pierre Auger Observatory



H.E.S.S.



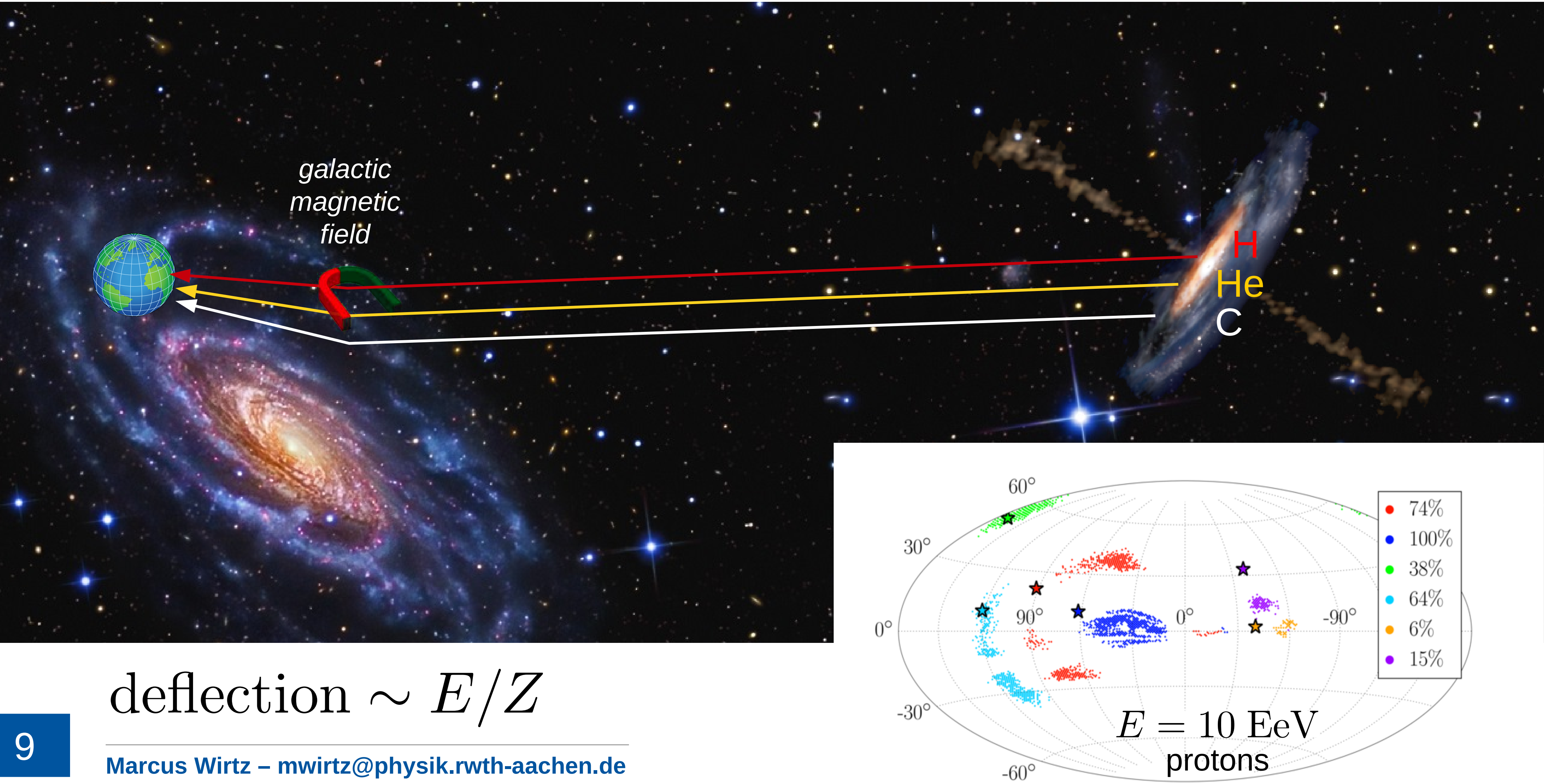
IceCube



Mainly used in the task of event reconstruction (direction, energy, type...)

Marcus Wirtz – mwirtz@physik.rwth-aachen.de

Patterns in arrival directions of UHECRS

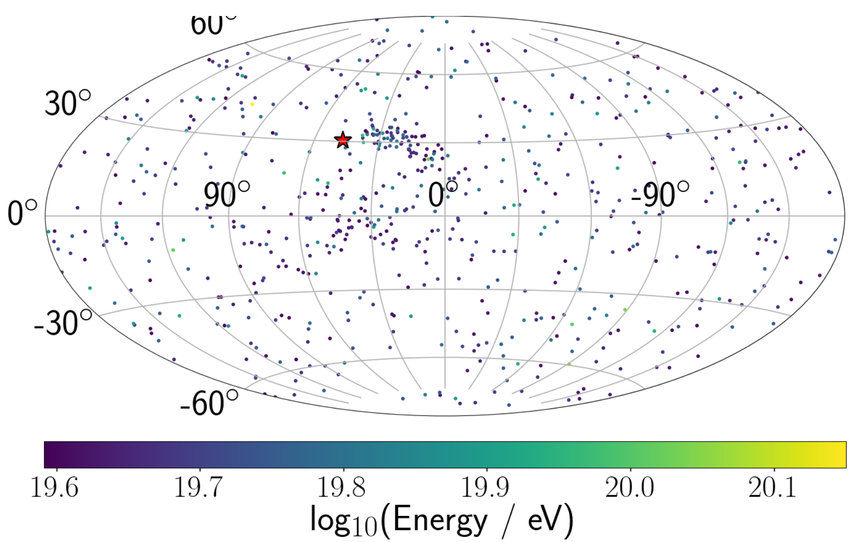


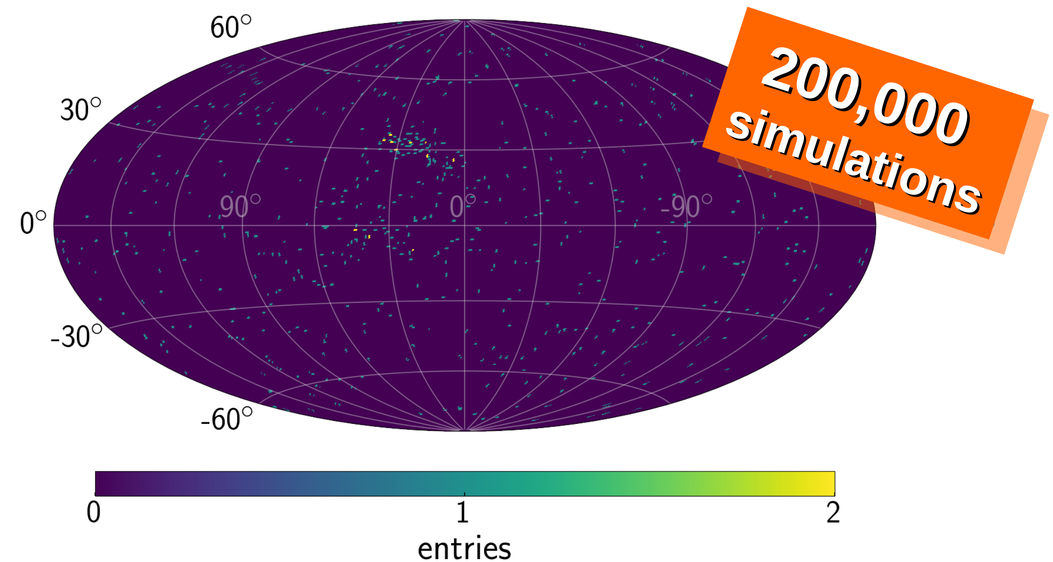
$$\text{deflection} \sim E/Z$$

Marcus Wirtz – mwirtz@physik.rwth-aachen.de

Simulation

- **Supervised learning** with simulated data
- Ultra-high energy cosmic ray arrival skymap with CRPropa3 framework
- Choosing energy cut of E_{min}=39EeV – > 651 UHECRs (AUGER)
- Including variations: source position, signal fraction, mass, fields...

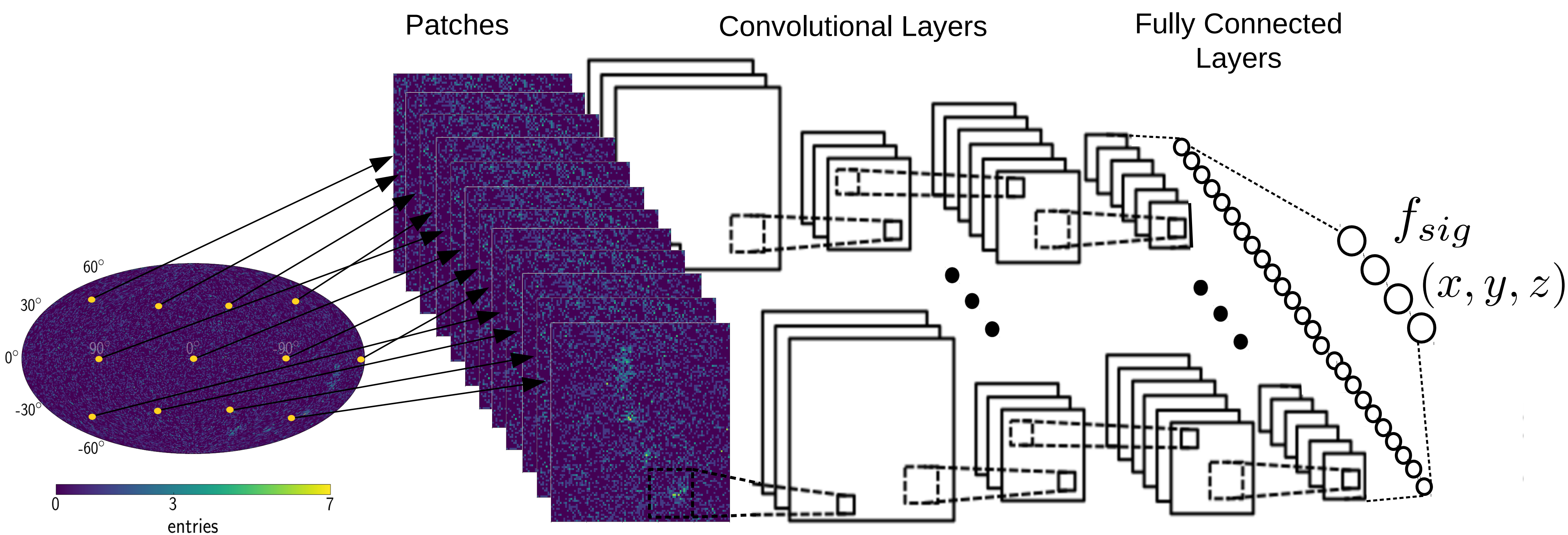




Marcus Wirtz – mwirtz@physik.rwth-aachen.de

Design of the neural network

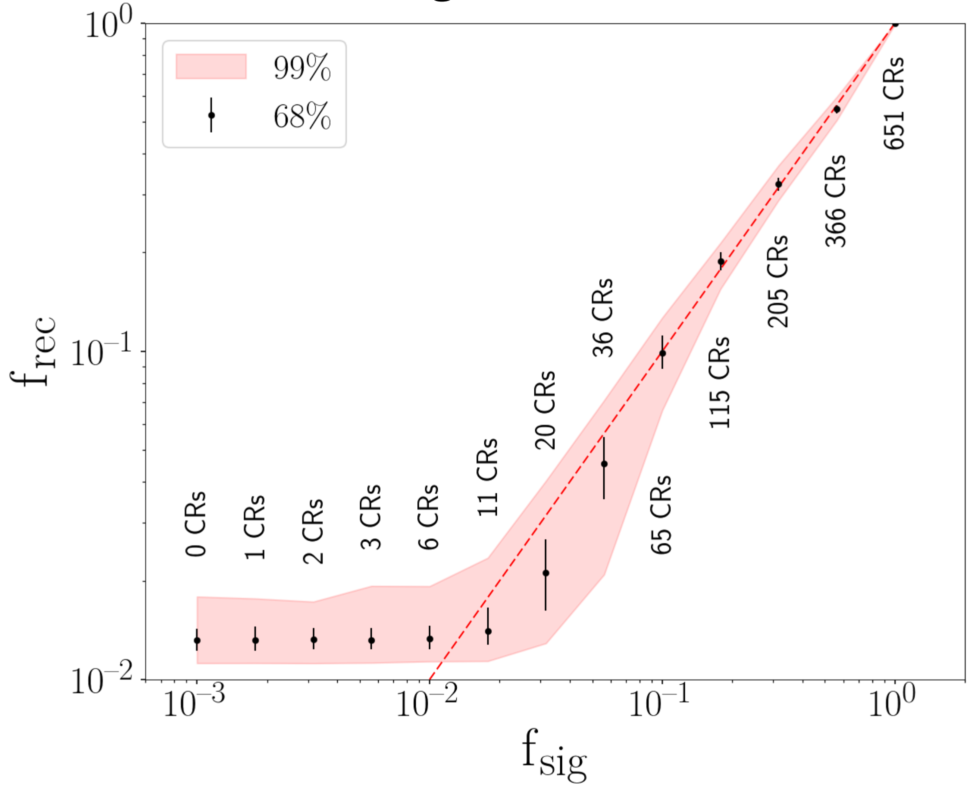
- ✗ Image like input data – > Convolutional network
- ✗ Project spherical surface on 12 planar patches



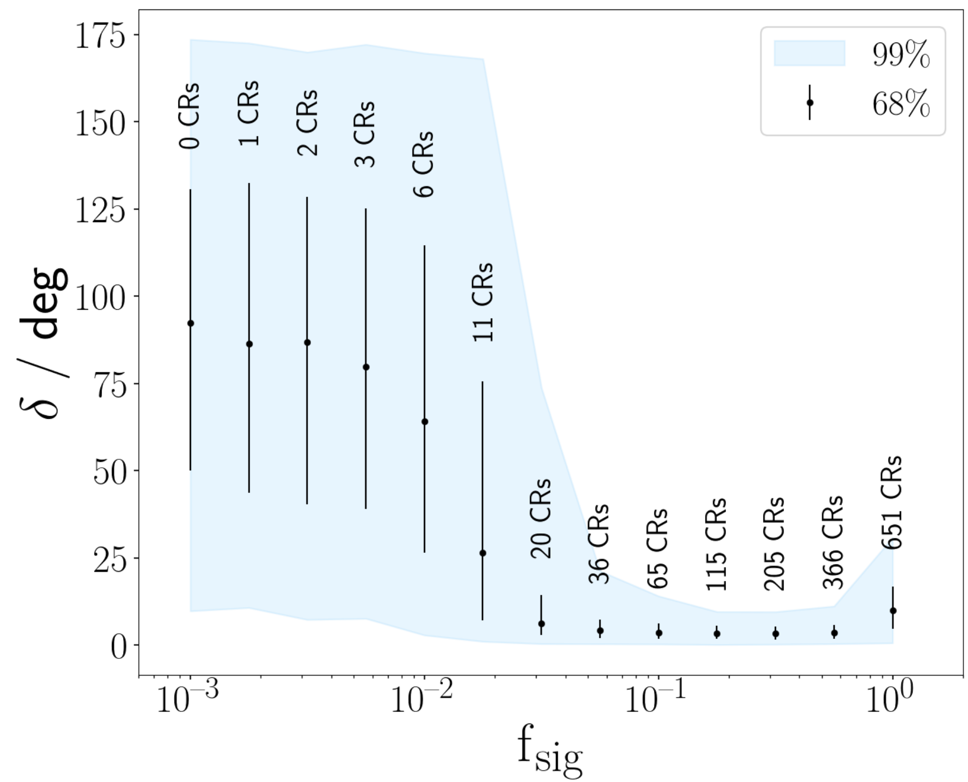
Marcus Wirtz – mwirtz@physik.rwth-aachen.de

Reconstruction on test data

signal fraction



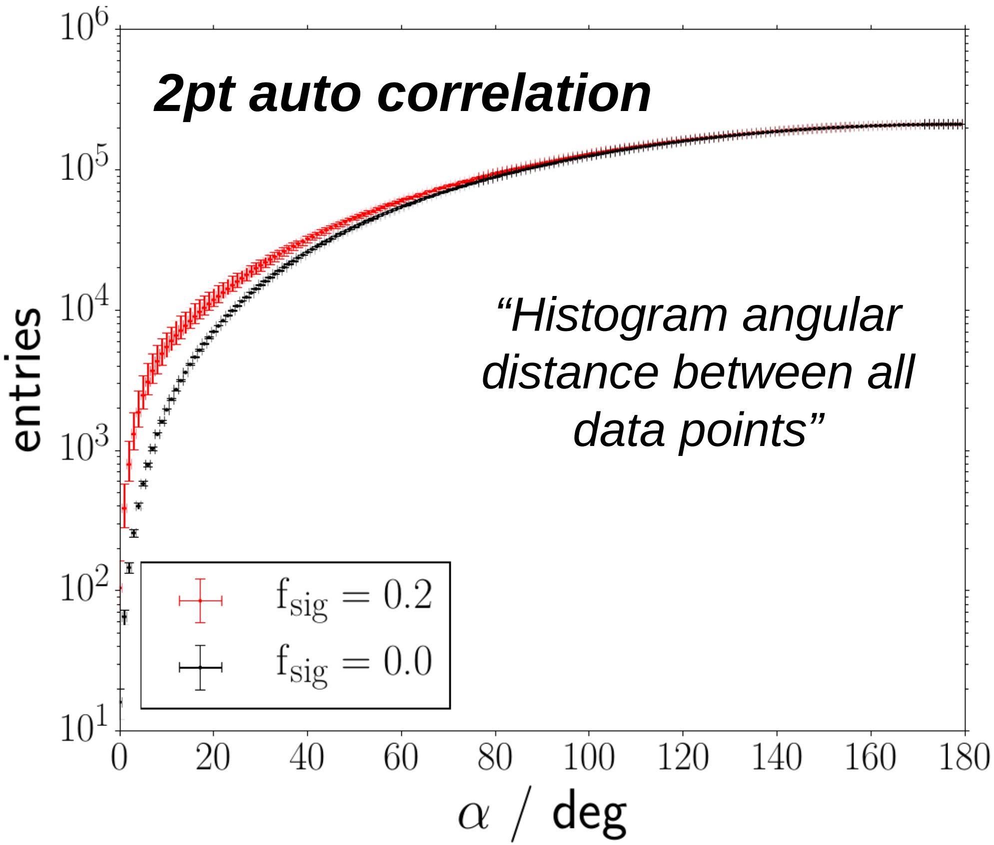
angle to true source

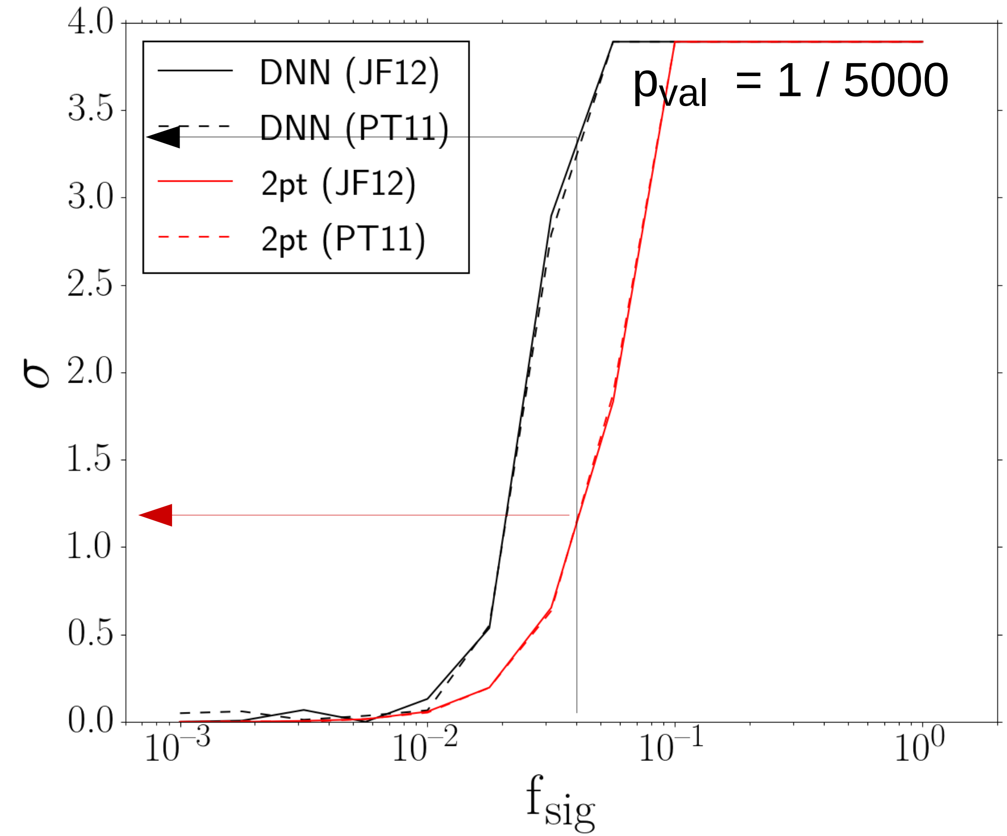


✗ Significance by comparing reconstructed signal fraction to isotropic maps

Marcus Wirtz – mwirtz@physik.rwth-aachen.de

Comparison to conventional analysis

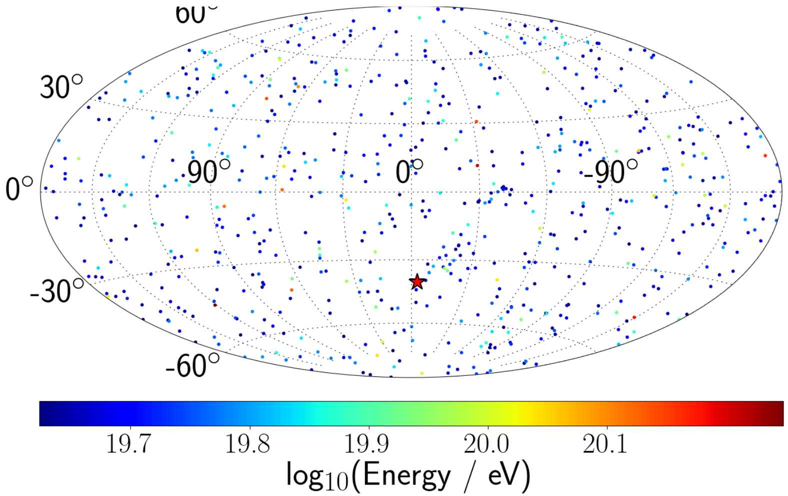


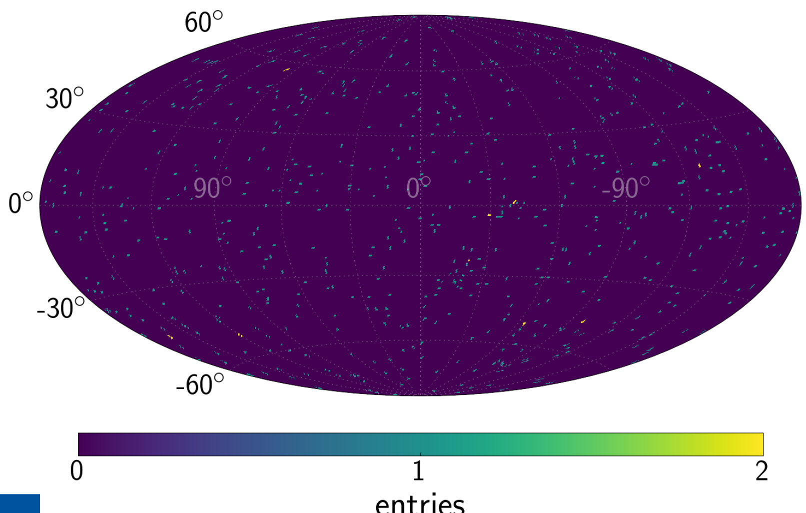


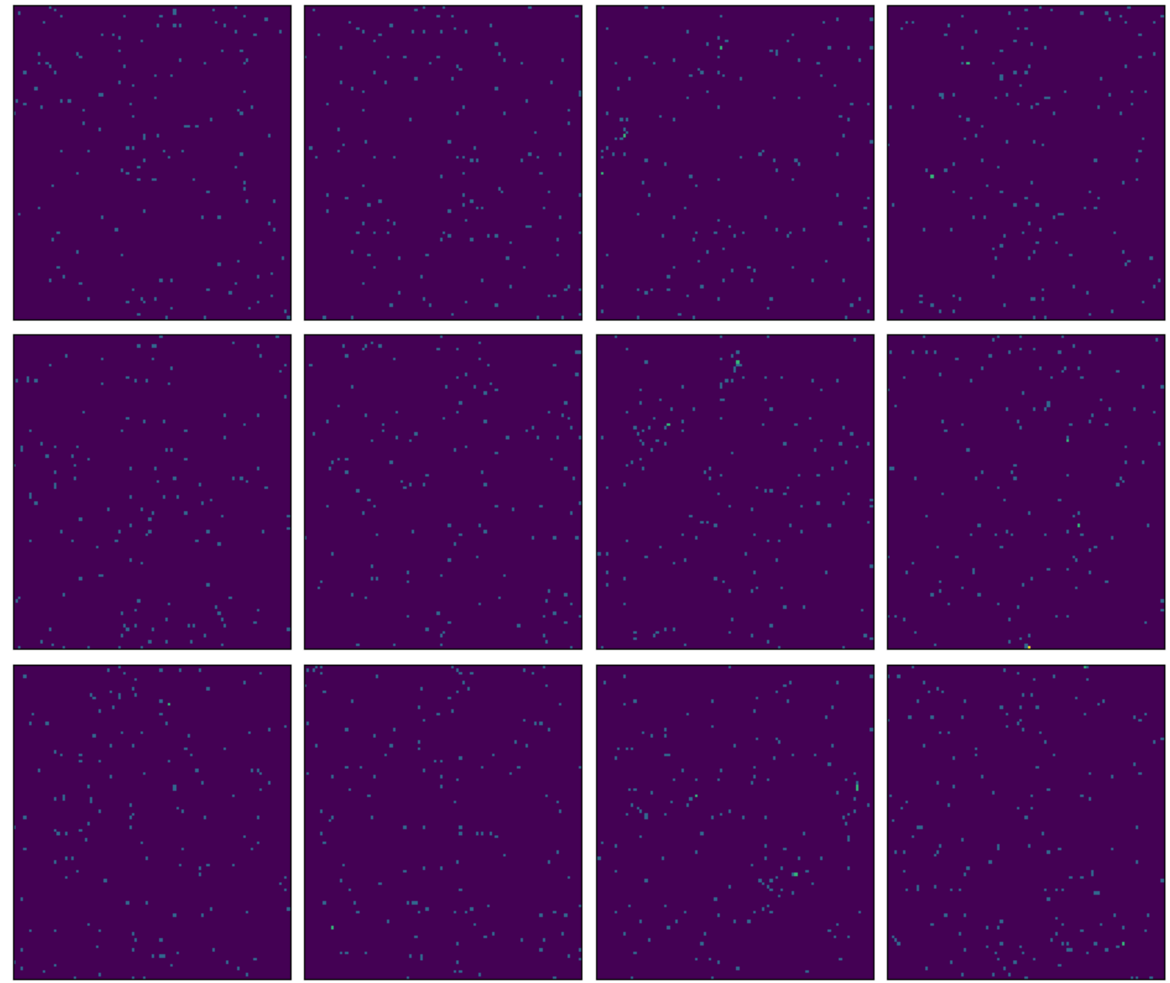
- ✗ 2pt auto: Take angular bin with highest significance to exclude isotropy
- ✗ At 4% signal fraction: 2pt-auto ($1.2\ \sigma$) / DNN ($3.4\ \sigma$)

Marcus Wirtz – mwirtz@physik.rwth-aachen.de

How does 4% signal from 651 cosmic rays look like?





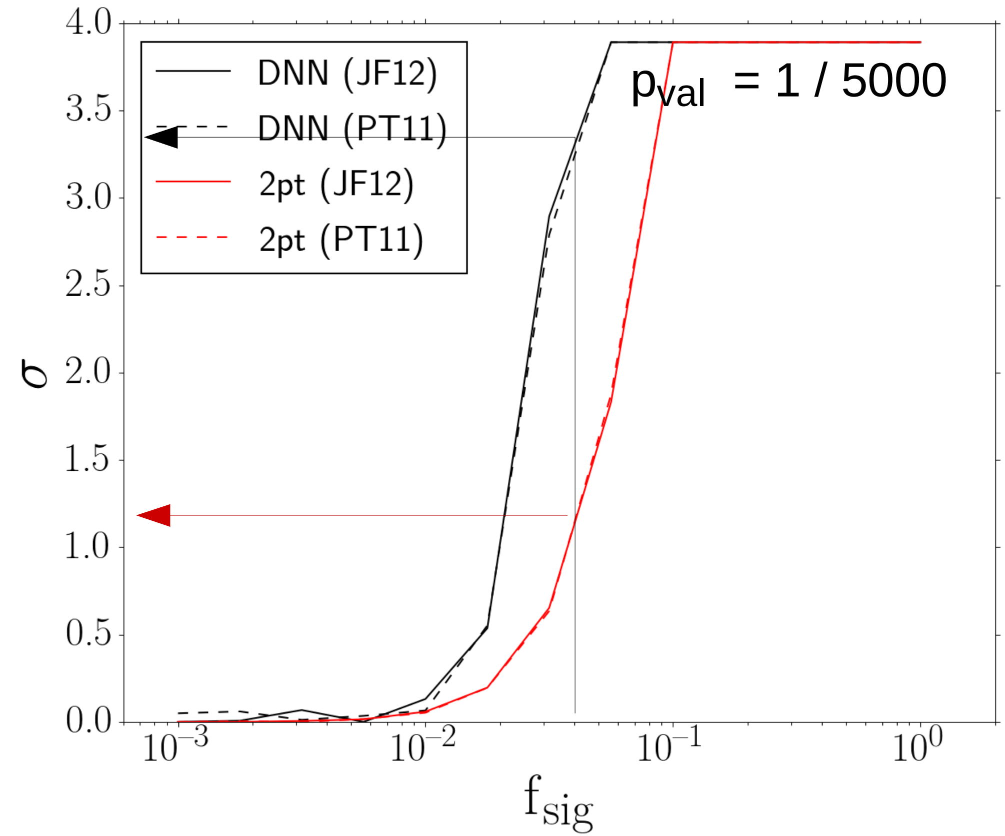


DNN able to detect at 3.5 sigma CL

Marcus Wirtz – mwirtz@physik.rwth-aachen.de

Summary

- Deep Learning is a powerful tool also for data analysis in physics
- Until now especially used for event reconstruction
- Simulated point source search for UHECRs : Deep Learning techniques perform better than a common used analysis method

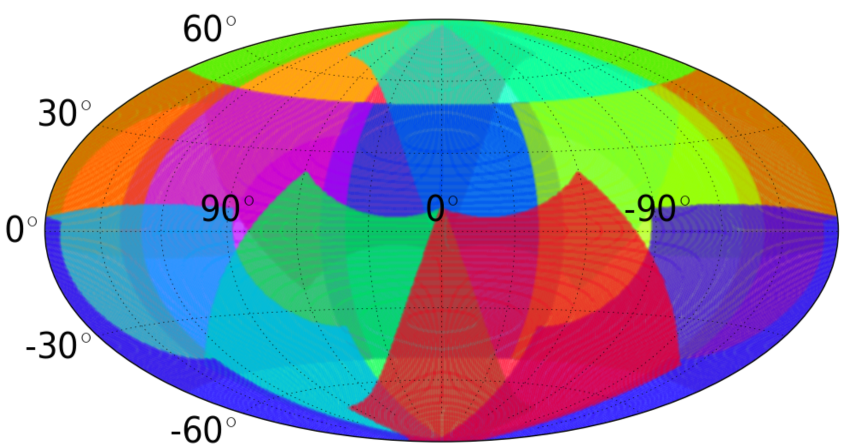


Marcus Wirtz – mwirtz@physik.rwth-aachen.de

Backup

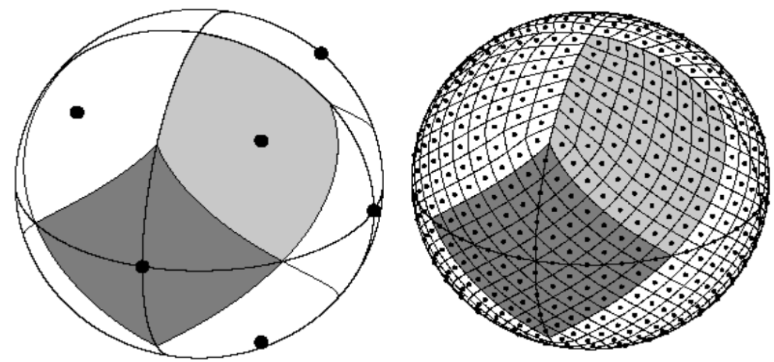
Image recognition on a sphere

- Healpy scheme
- Divide sphere into 12 patches (~planar)
  → size: 100° x 100°



Patch size: 100°

Patch resolution (npix = 12) | Source regions (npix = 768)

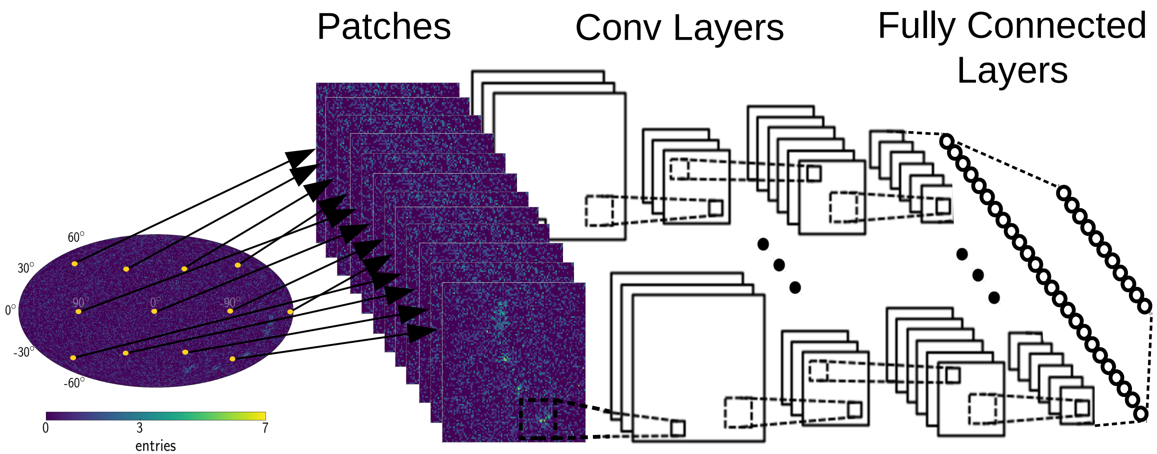


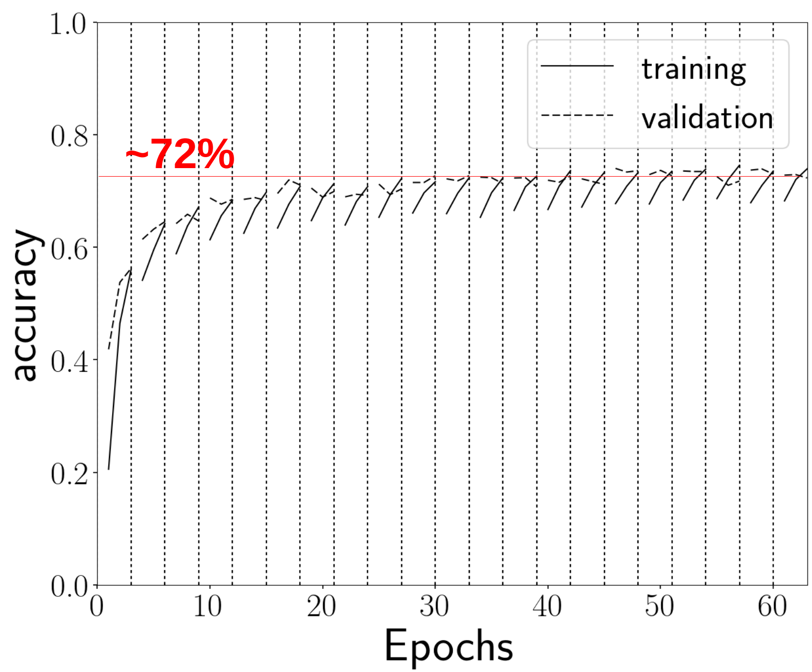
- Run one 2D-CNN on each patch

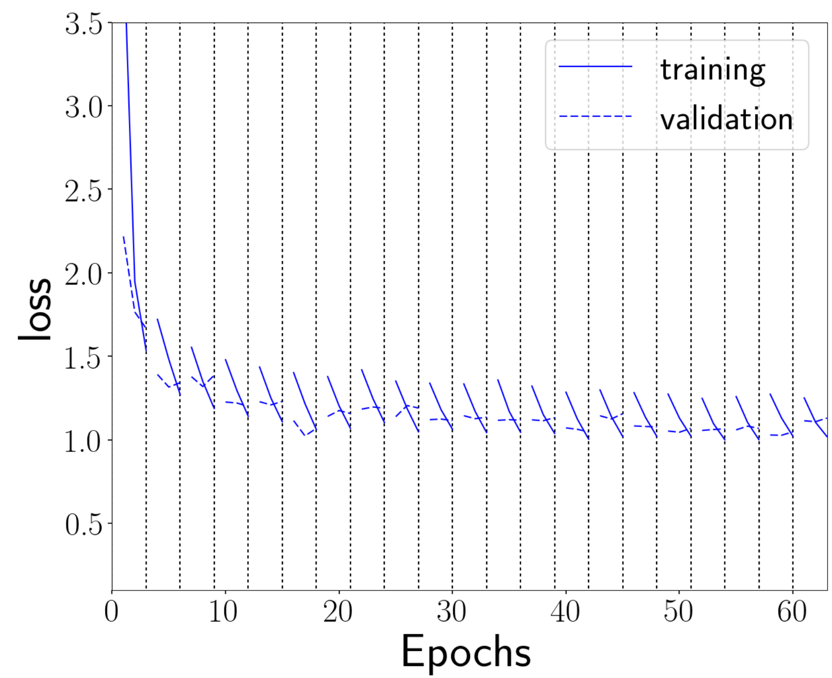
Marcus Wirtz – mwirtz@physik.rwth-aachen.de

Training

- ✗ Training data is expensive: 1,000 skymaps ~ 1 GB
- ✗ Maximum of 20,000 at once



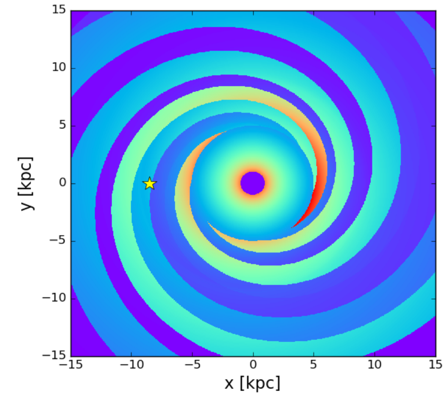




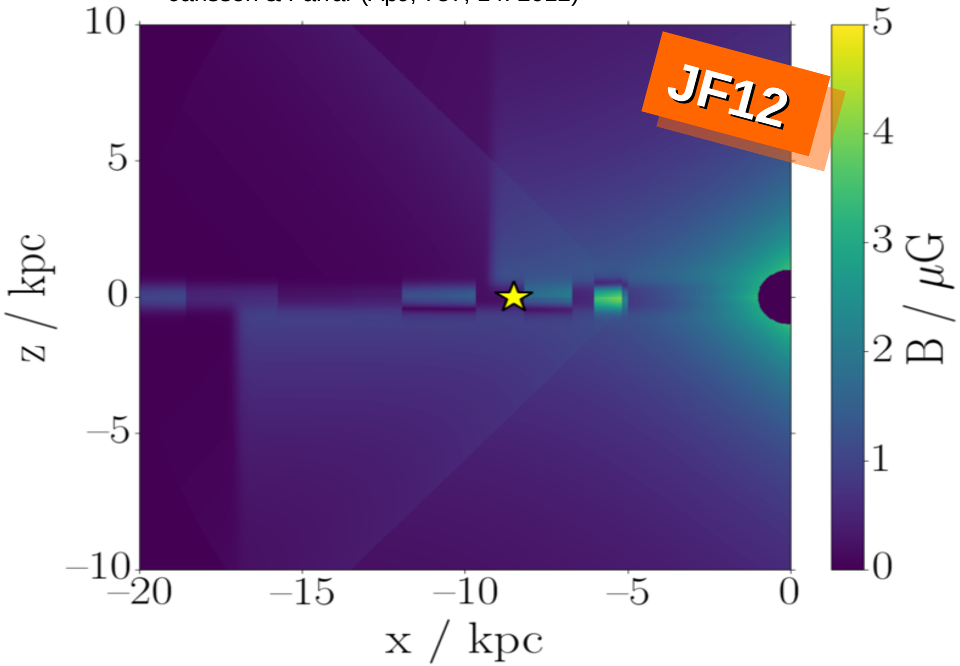
Marcus Wirtz – mwirtz@physik.rwth-aachen.de

Galactic magnetic field - parametrizations

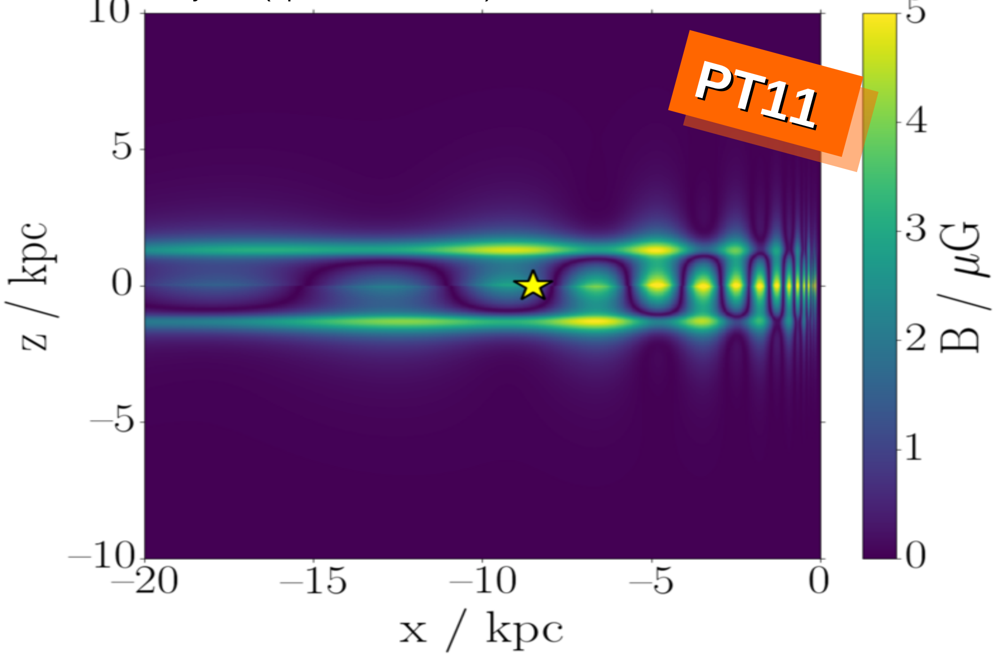
✗ Models tuned to measurements
(e.g. rotation measurements, synchrotron radiation)

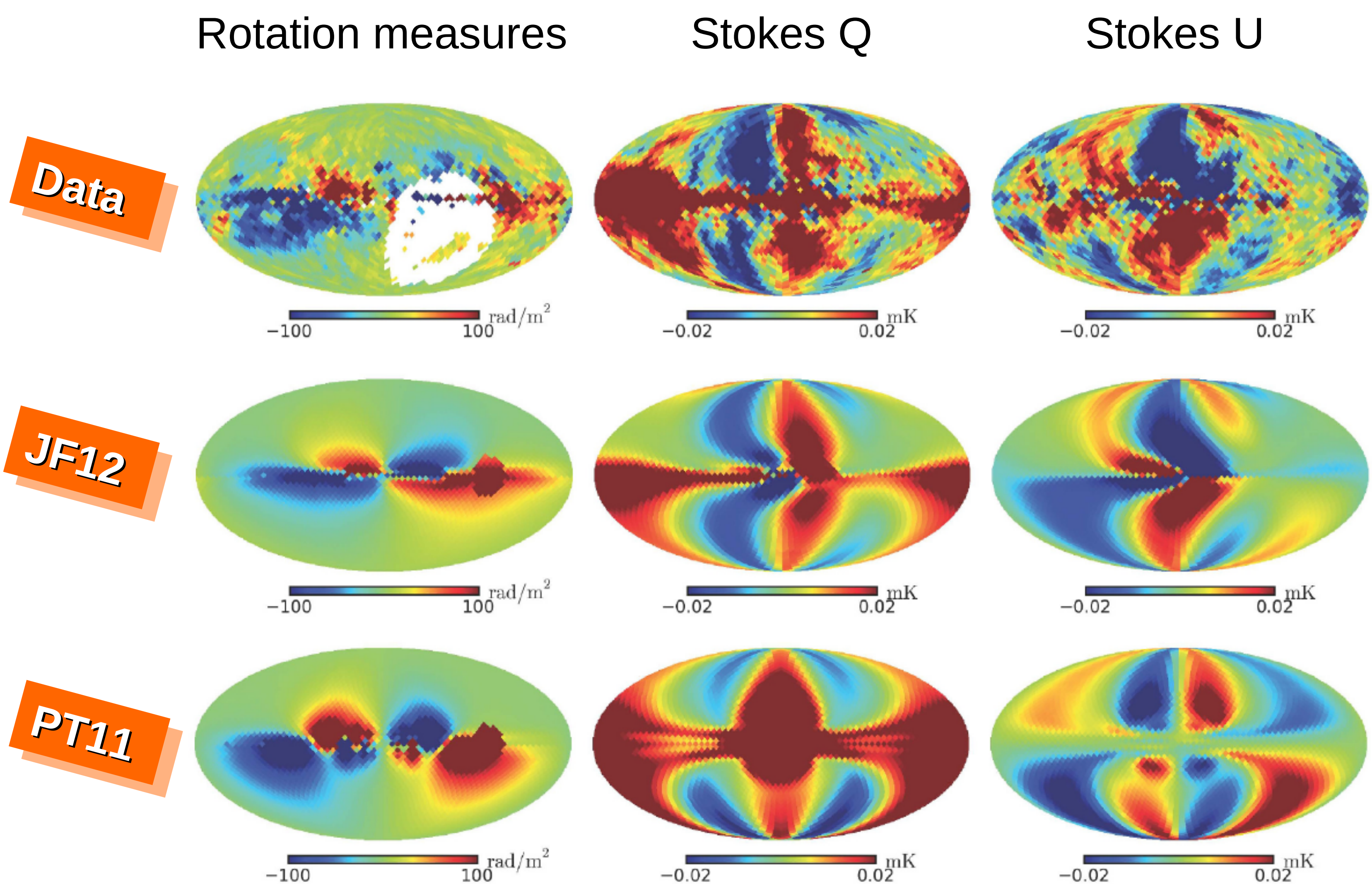


Jansson & Farrar (ApJ, 757, 14: 2012)

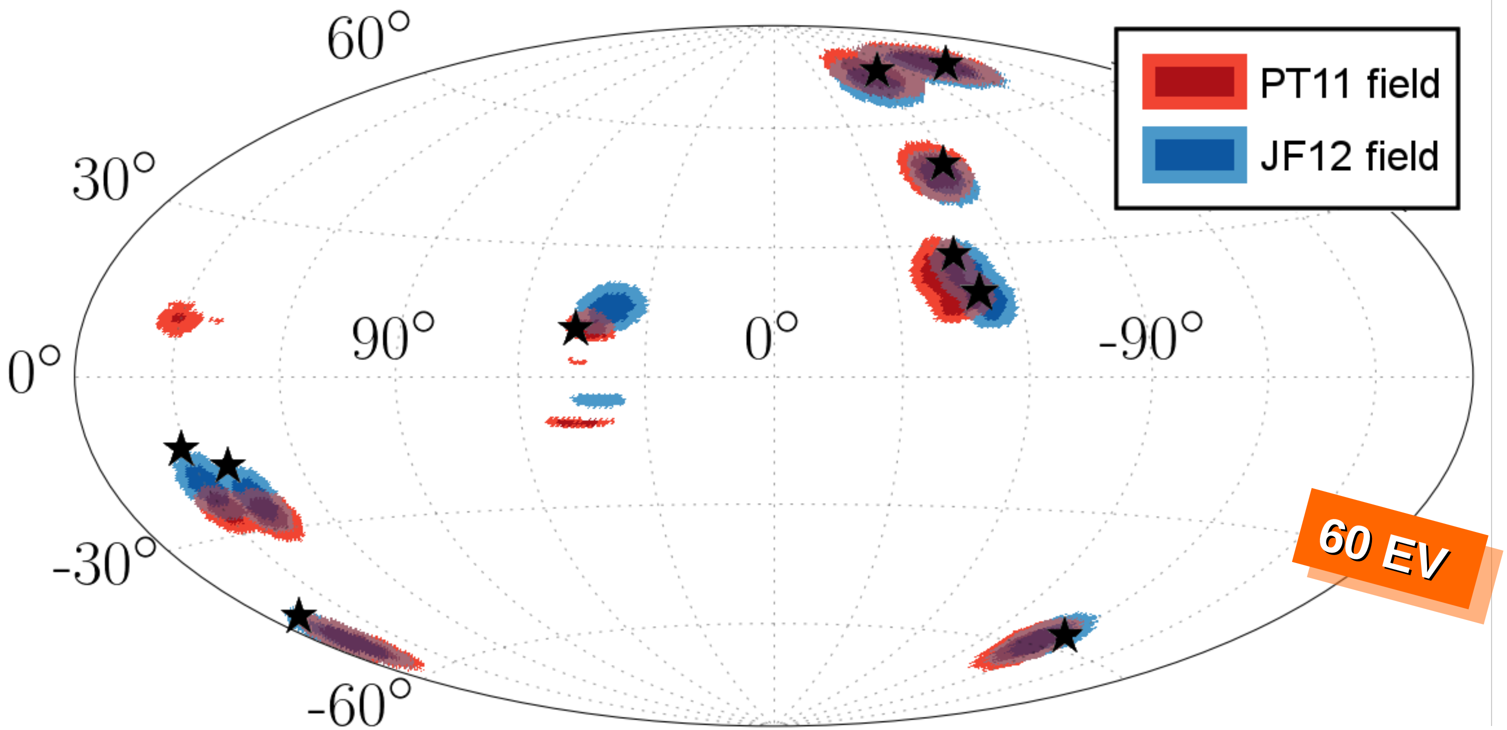


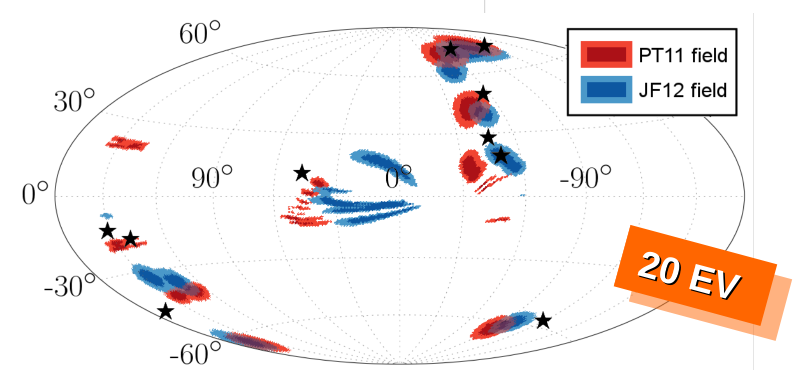
Pshirkov & Tinyakov (ApJ, 738, 192: 2011)




Rotation measures
Stokes Q
Stokes U
Data
JF12
PT11
−100
100
rad/m²
−0.02
0.02
mK

GMF arrival distributions



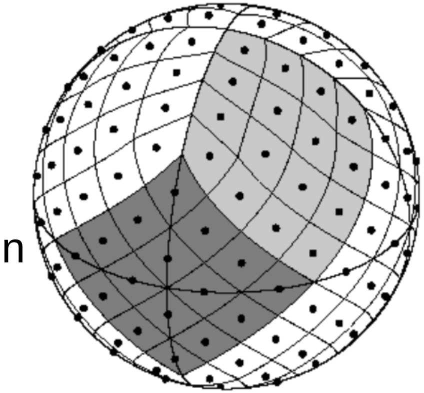


- ✗ Displacements of JF12 and PT11 match mostly
- ✗ Large deflections → point to similar directions

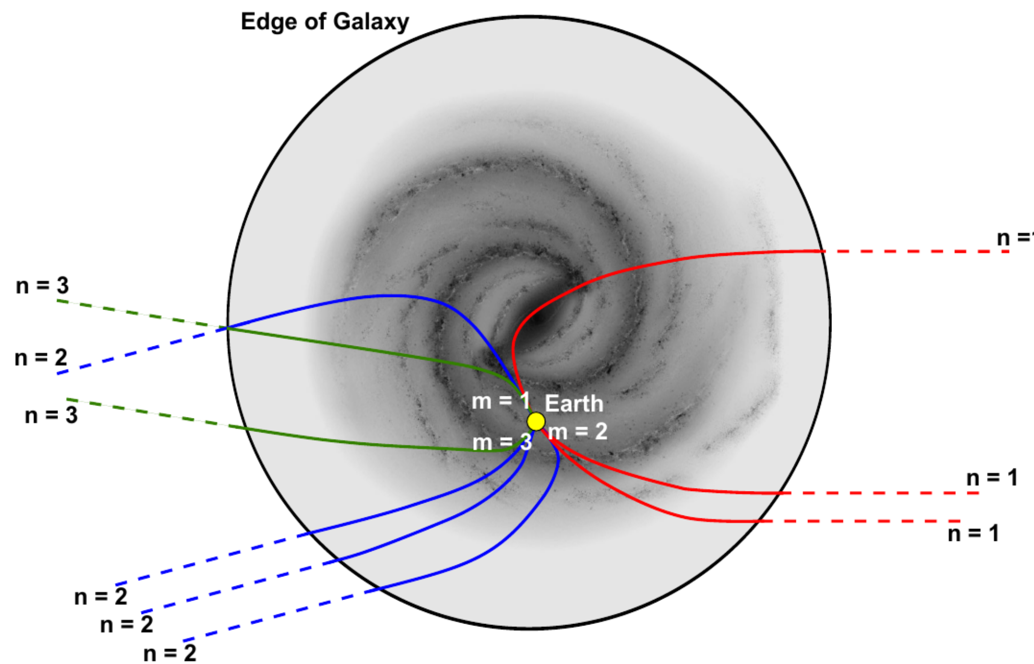
Galactic magnetic field lenses

Matrices for each rigidity R = E / Z mapping
extragalactic directions to **observed arrival directions**

Healpy
Pixelization



- ✗ Based on Healpy framework
  (divide sphere into 49,152 cells)
- ✗ CRPropa simulation:
  Backtrack 5 million particles
  (inverting charge) per rigidity-bin
  to the edge of the galaxy
- ✗ 175 rigidity bins in the range from
  10^{17}eV to $10^{20.5}$eV
- ✗ Matrices are normalized to the
  highest arrival probability
  → rigidity and direction dependent
  transparency



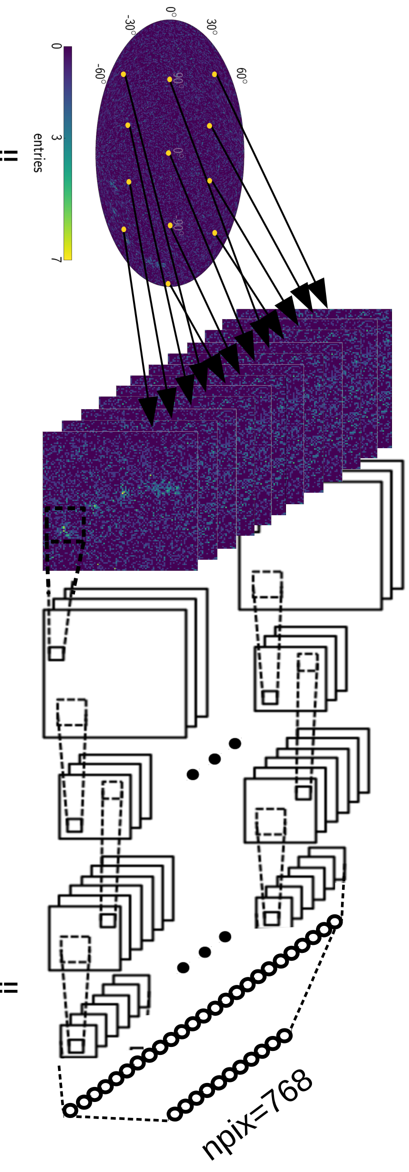
https://web.physik.rwth-aachen.de/Auger_MagneticFields/PARSEC/downloads.php

Marcus Wirtz – mwirtz@physik.rwth-aachen.de

Model.summary()

Layer (type)	Output Shape	Param #	Connected to
convolution3d_1 (Convolution3D)	(None, 12, 98, 98, 32	320	convolution3d_input_1[0][0]
maxpooling3d_1 (MaxPooling3D)	(None, 12, 49, 49, 32	0	convolution3d_1[0][0]
convolution3d_2 (Convolution3D)	(None, 12, 47, 47, 32	9248	maxpooling3d_1[0][0]
maxpooling3d_2 (MaxPooling3D)	(None, 12, 23, 23, 32	0	convolution3d_2[0][0]
convolution3d_3 (Convolution3D)	(None, 12, 21, 21, 32	9248	maxpooling3d_2[0][0]
maxpooling3d_3 (MaxPooling3D)	(None, 12, 10, 10, 32	0	convolution3d_3[0][0]
convolution3d_4 (Convolution3D)	(None, 12, 8, 8, 32)	9248	maxpooling3d_3[0][0]
maxpooling3d_4 (MaxPooling3D)	(None, 12, 4, 4, 32)	0	convolution3d_4[0][0]
flatten_1 (Flatten)	(None, 6144)	0	maxpooling3d_4[0][0]
dropout_1 (Dropout)	(None, 6144)	0	flatten_1[0][0]
dense_1 (Dense)	(None, 768)	4719360	dropout_1[0][0]
dropout_2 (Dropout)	(None, 768)	0	dense_1[0][0]
dense_2 (Dense)	(None, 768)	590592	dropout_2[0][0]

Total params: 5,338,016
Trainable params: 5,338,016
Non-trainable params: 0



Marcus Wirtz – mwirtz@physik.rwth-aachen.de