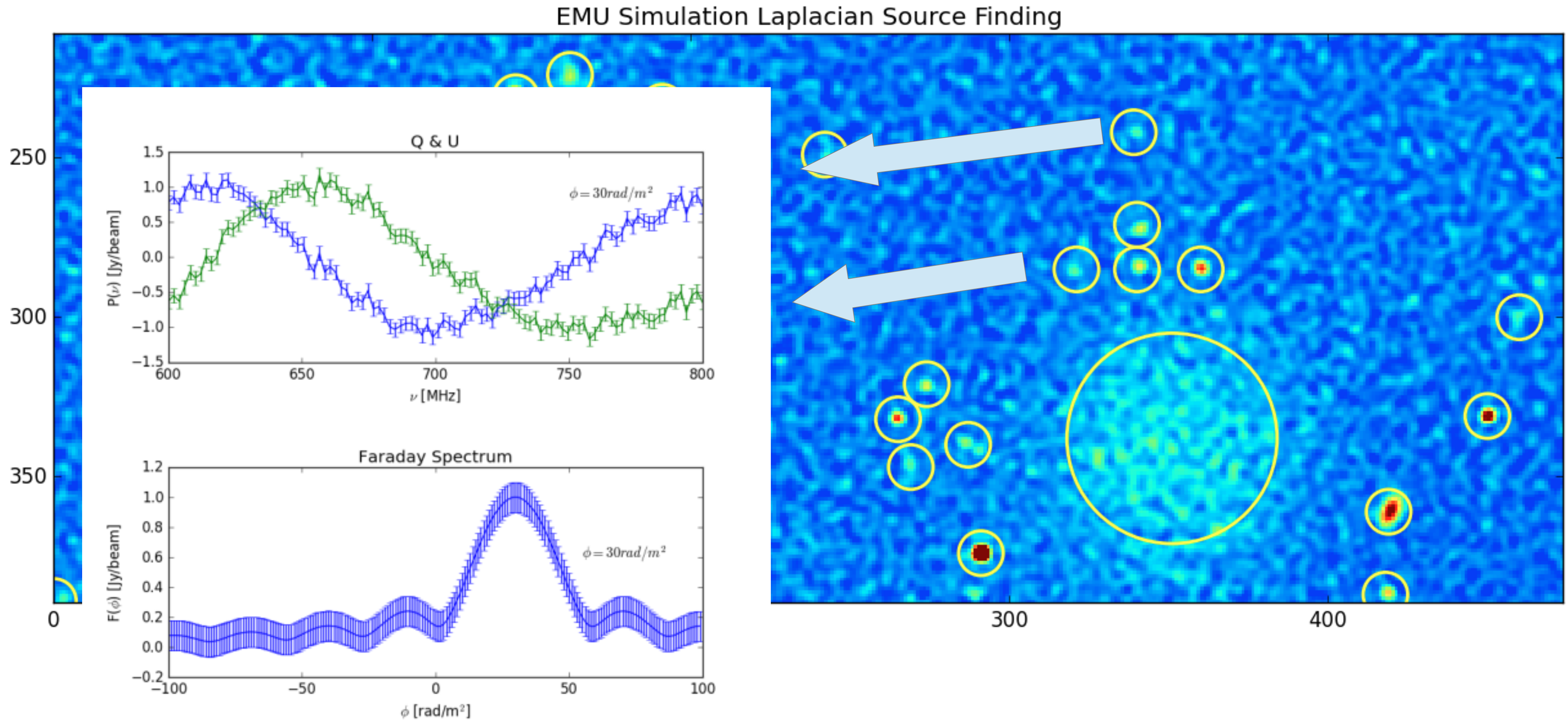
Using Deep Learning to Detect Complex Faraday Rotation Measures

Shea Brown
The University of Iowa
SPARCS IX – Lisboa, Portugal

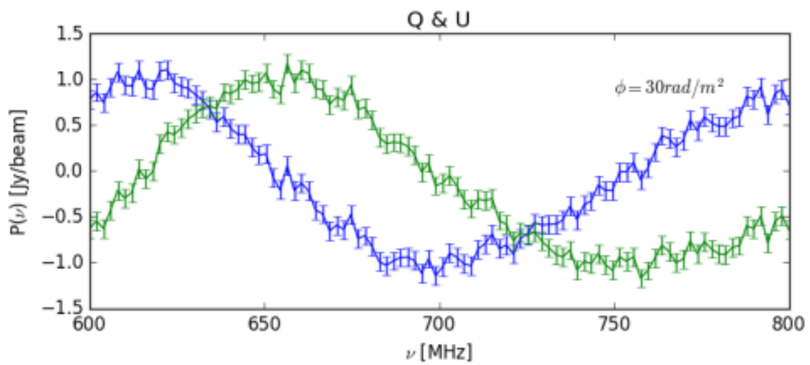
Australian SKA Pathfinder

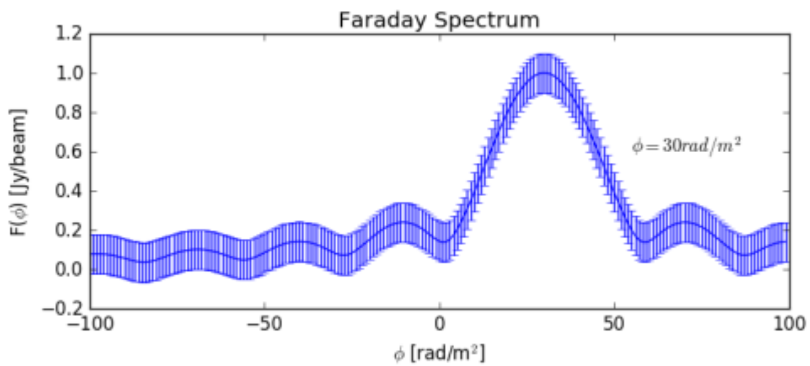
POSSUM: Polarisation Sky Survey of the Universe's Magnetism

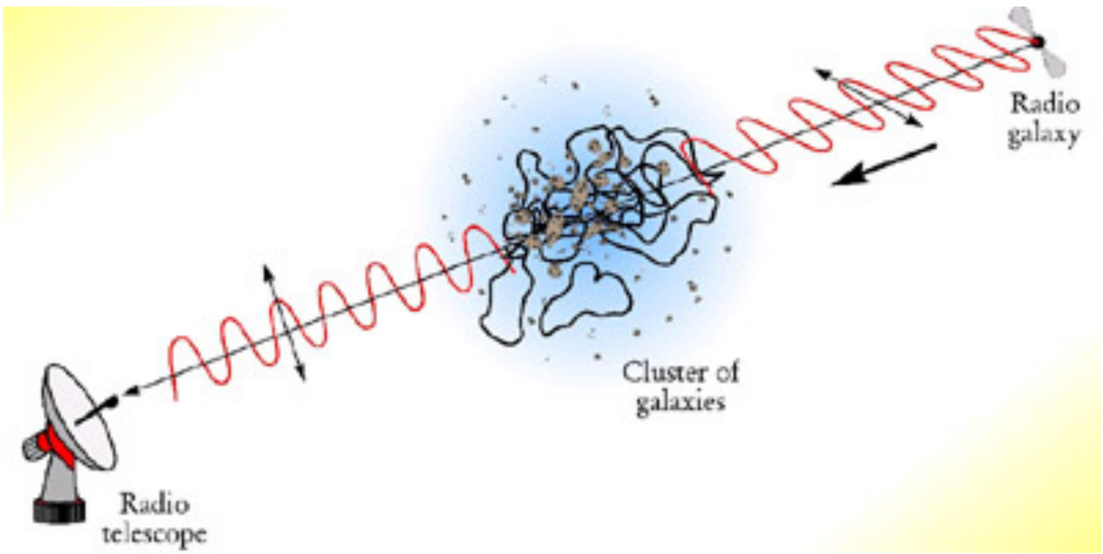


2. Faraday Rotation

A magnetized plasma with an ordered magnetic field will rotate the angle of linearly polarized light



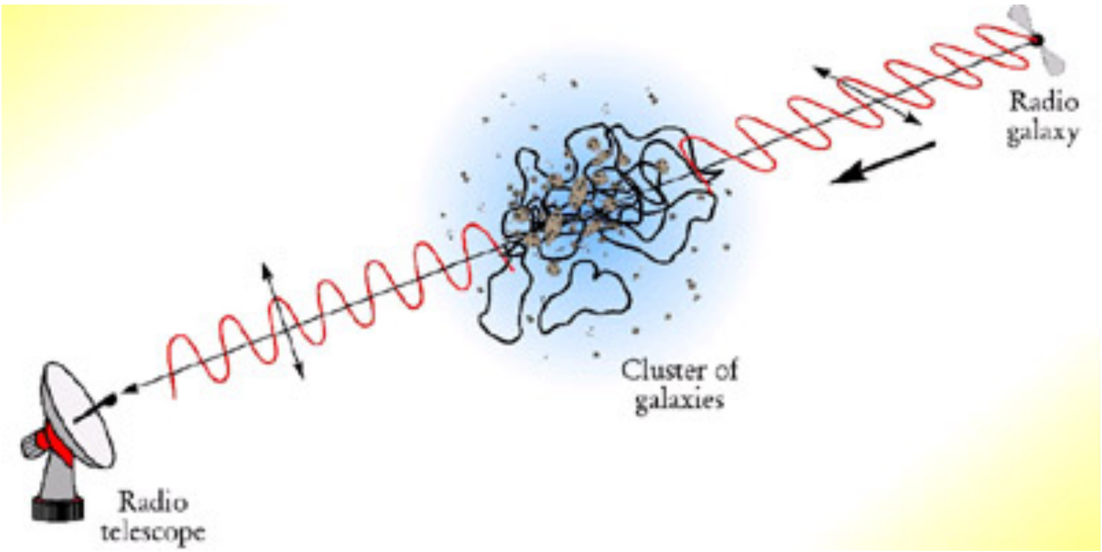


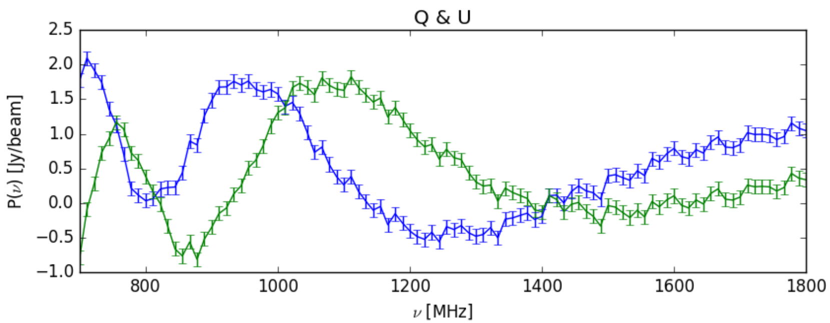


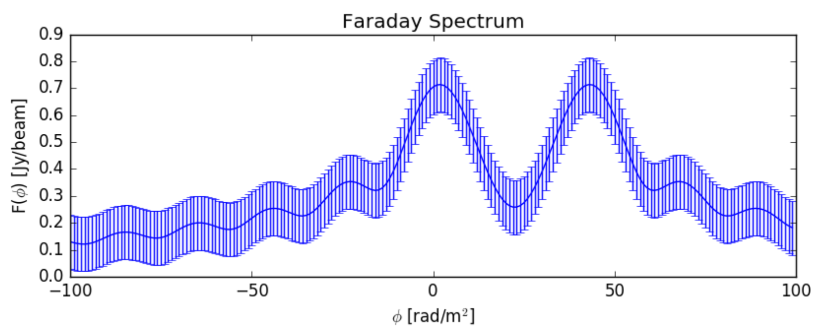
Simple Source

Faraday Rotation

A magnetized plasma with an ordered magnetic field will rotate the angle of linearly polarized light



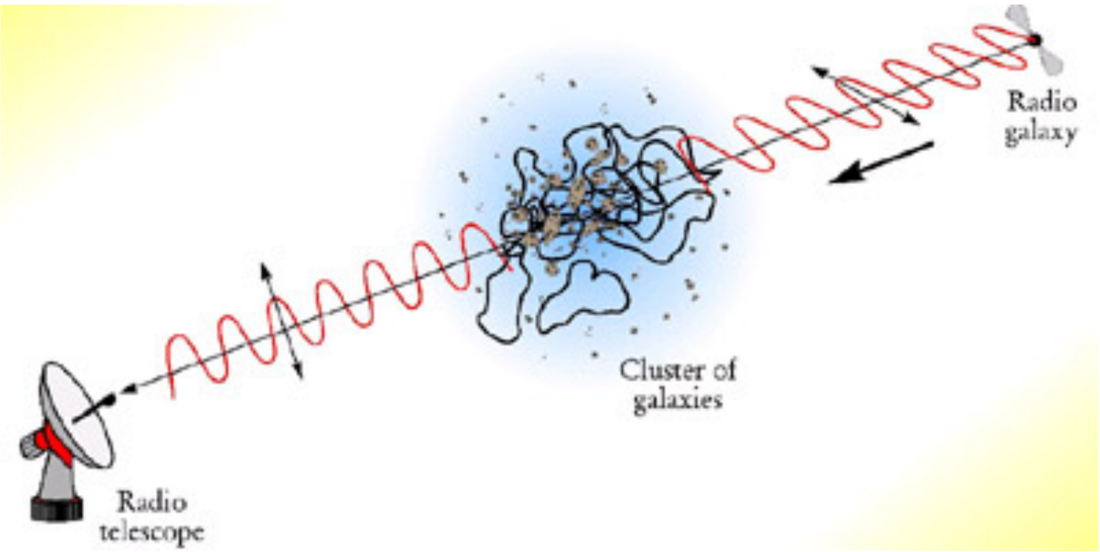


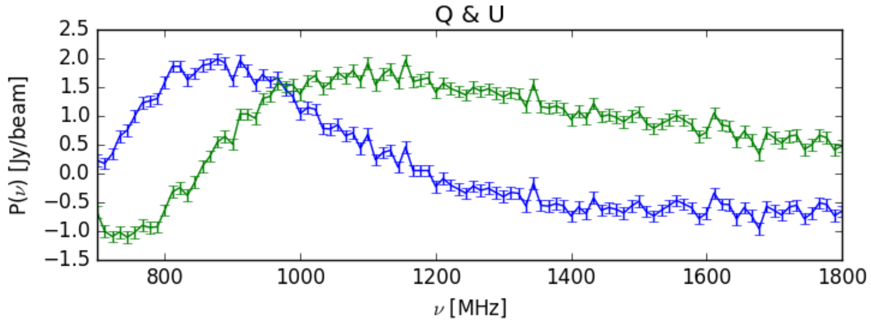


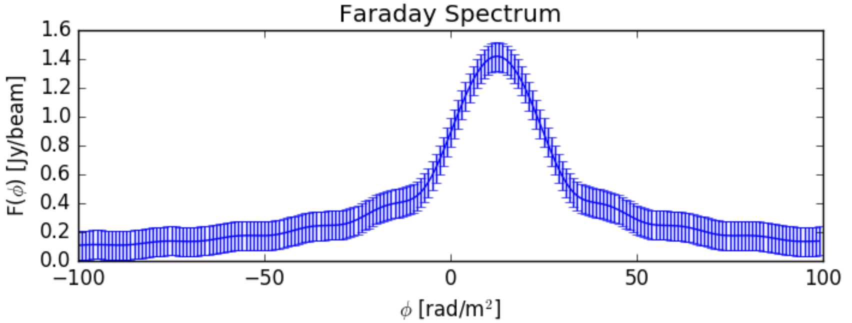
Two Simple Sources -> Complex

Faraday Rotation

A magnetized plasma with an ordered magnetic field will rotate the angle of linearly polarized light



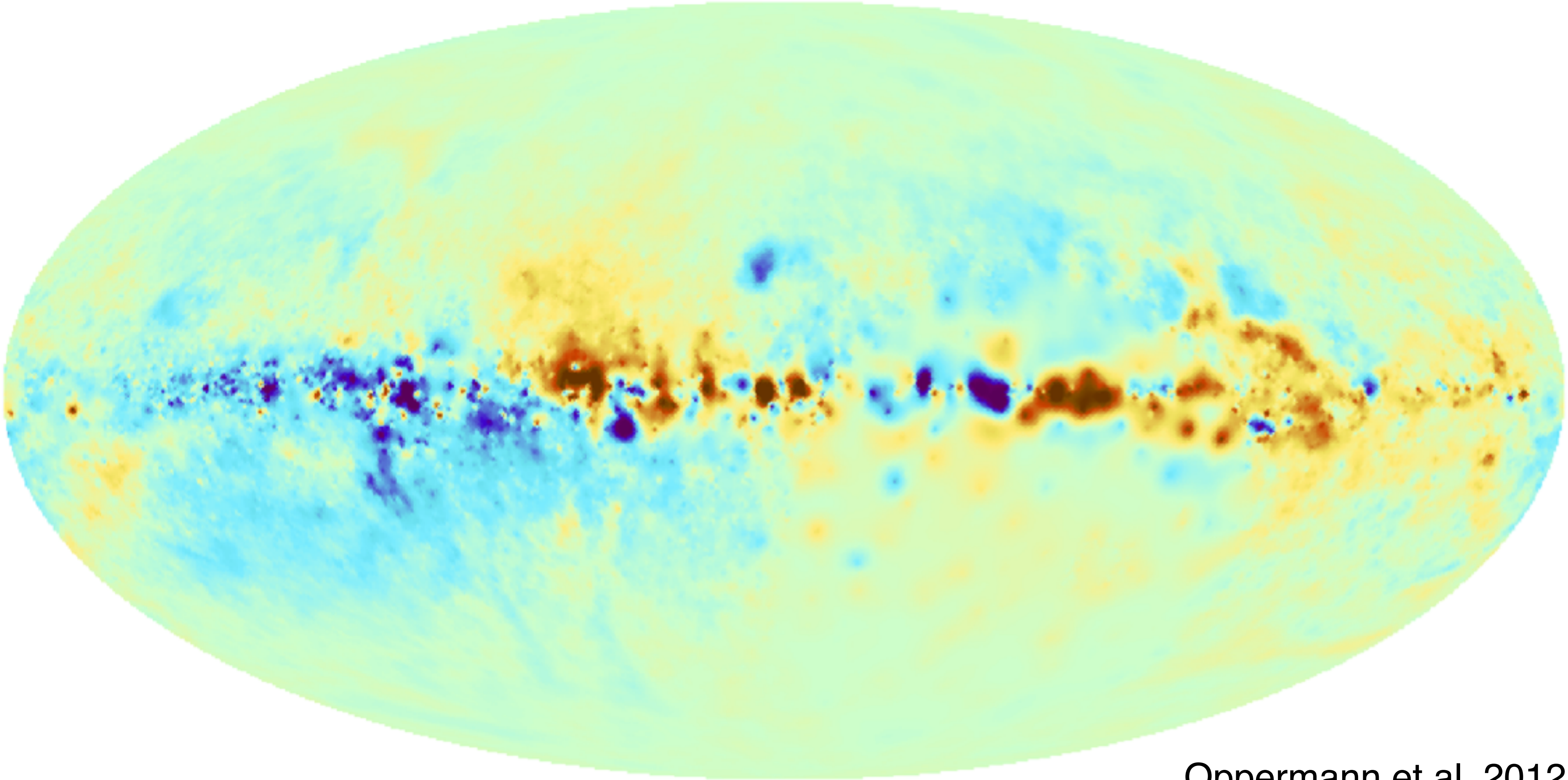




Faraday Thick Source

Faraday Rotation

Bayesian Reconstruction of Rotation Measure Sky

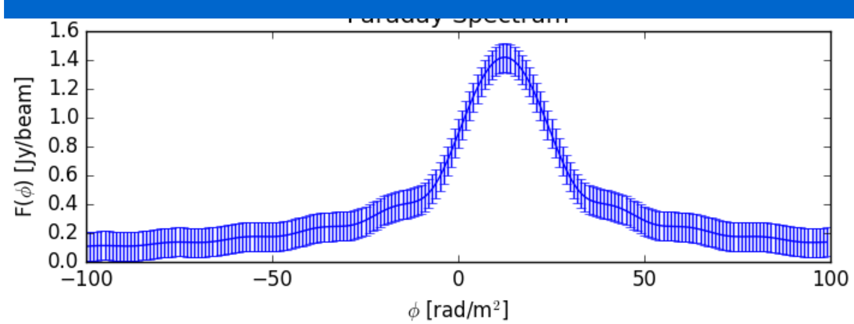


Oppermann et al. 2012

Faraday Rotation

A magnetized plasma with
an ordered magnetic field
w
li

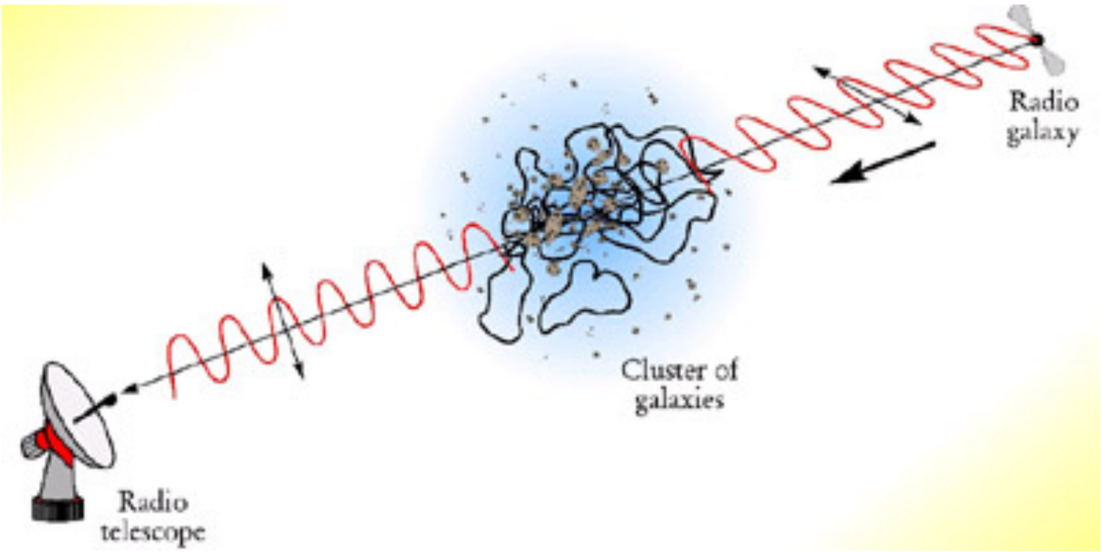
How do we automatically tell the difference between simple and complex sources?

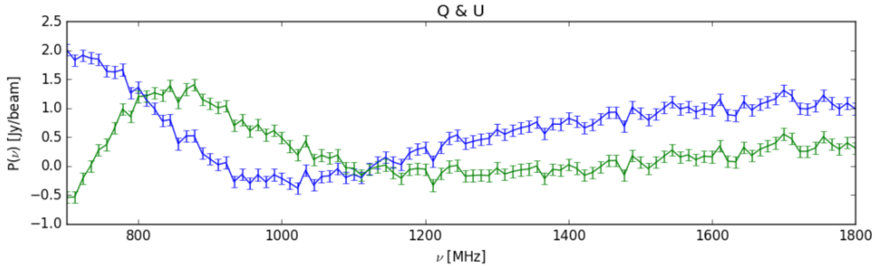


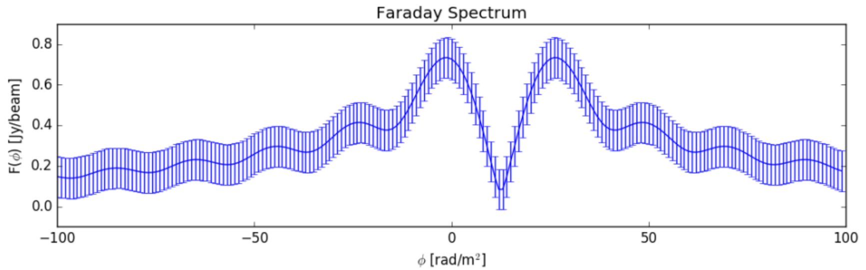
Faraday Thick Source

Faraday Rotation

- We understand the physics and know the important parameters
- We understand the observation process
- We can generate synthetic data to train a classifier to distinguish between simple and complex sources



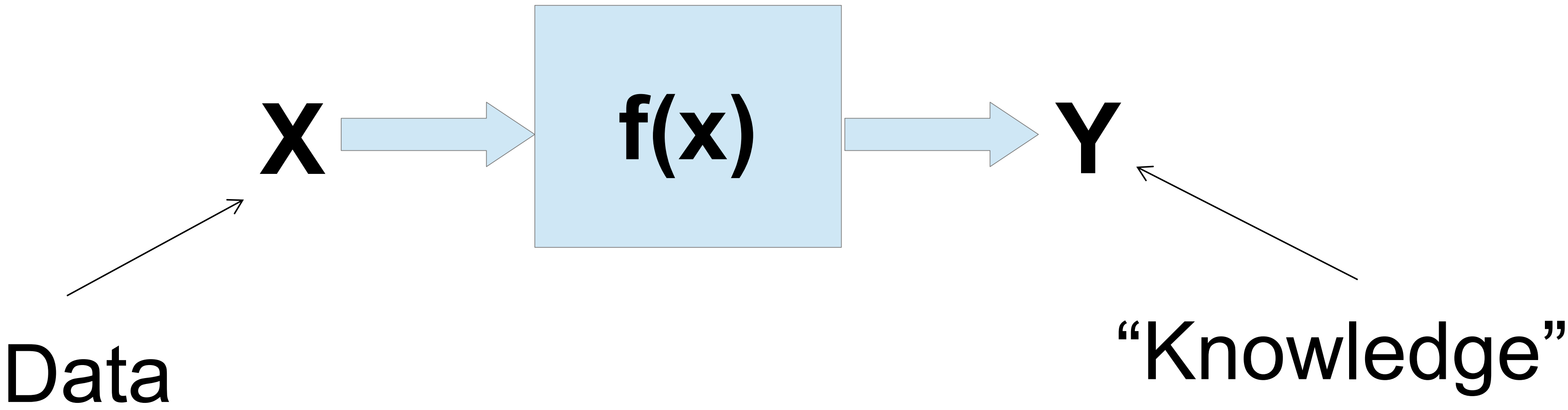




Machine Learning

What is machine-learning?

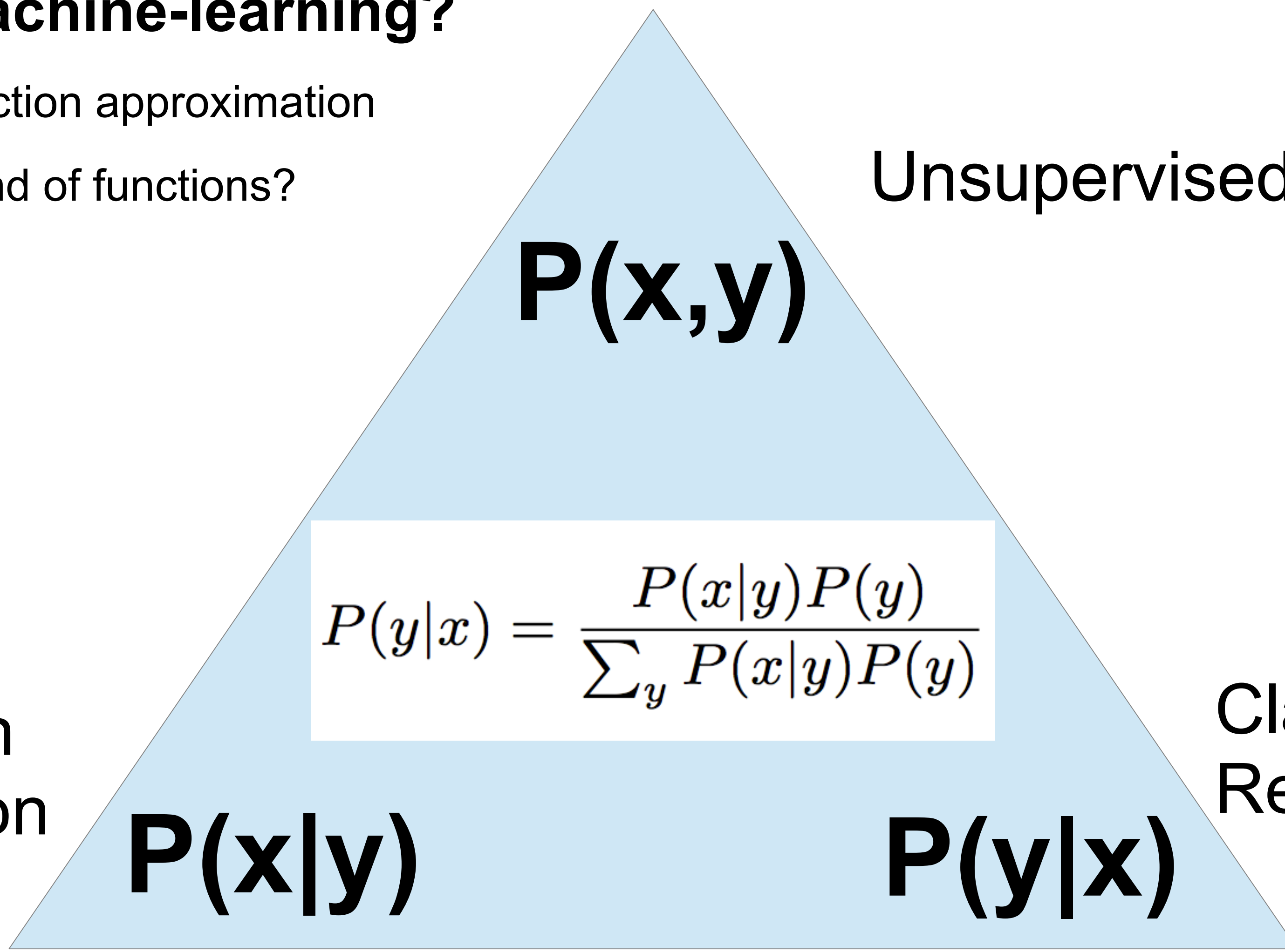
– Just function approximation



Machine Learning

What is machine-learning?

- Just function approximation
- What kind of functions?

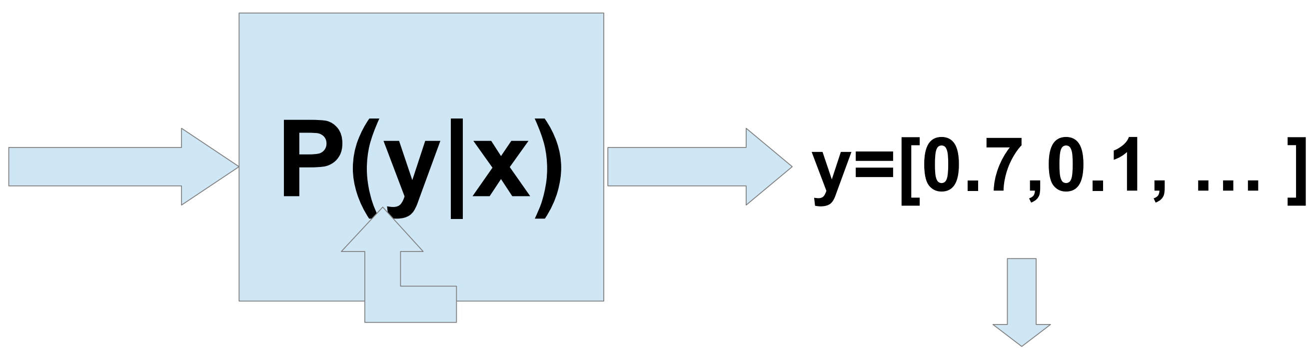


Adapted from Lin & Tegmark 2016

Machine Learning

- **What is machine-learning?**
  - Just function approximation
  - What kind of functions?
- **Classification**
  - **y** is discrete set of classes
    - y = [cat, dog, rock, tree, …]
  - **x** might be an image

P(y|x)

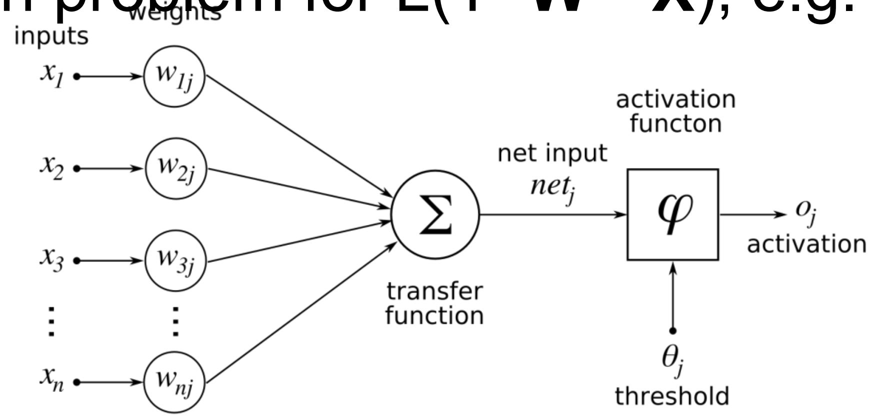


Train with gradient descent

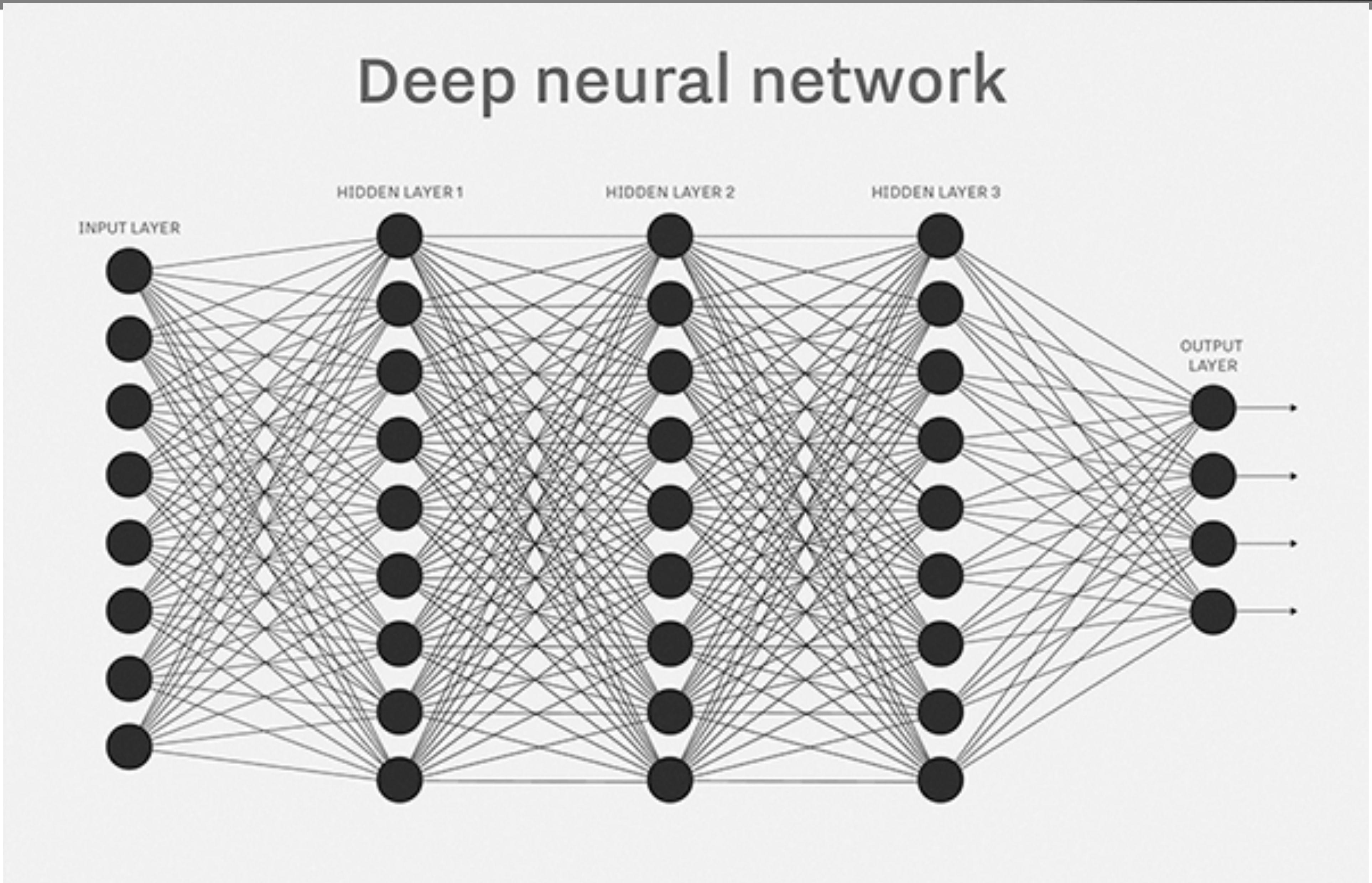
Machine Learning

e.g., binary linear model example

- **X** is a vector of “features” of an person
  - [height, weight, age, miles run per week, time of Facebook,..]
- P(**X**) is the probability that this person is diabetic
- Simplest map from a vector to a scalar: P(**X**) = **W** · **X**
- Training: Initialize **W** with random weights. Get **X** for thousand of people, along with the associated Y (0 or 1).
- Minimization problem for L(Y- **W** · **X**), e.g. gradient decent



Deep Learning



Machine Learning

Information "Hamiltonian"

$$P(y|x) = \frac{P(x|y)P(y)}{\sum_y P(x|y)P(y)}$$

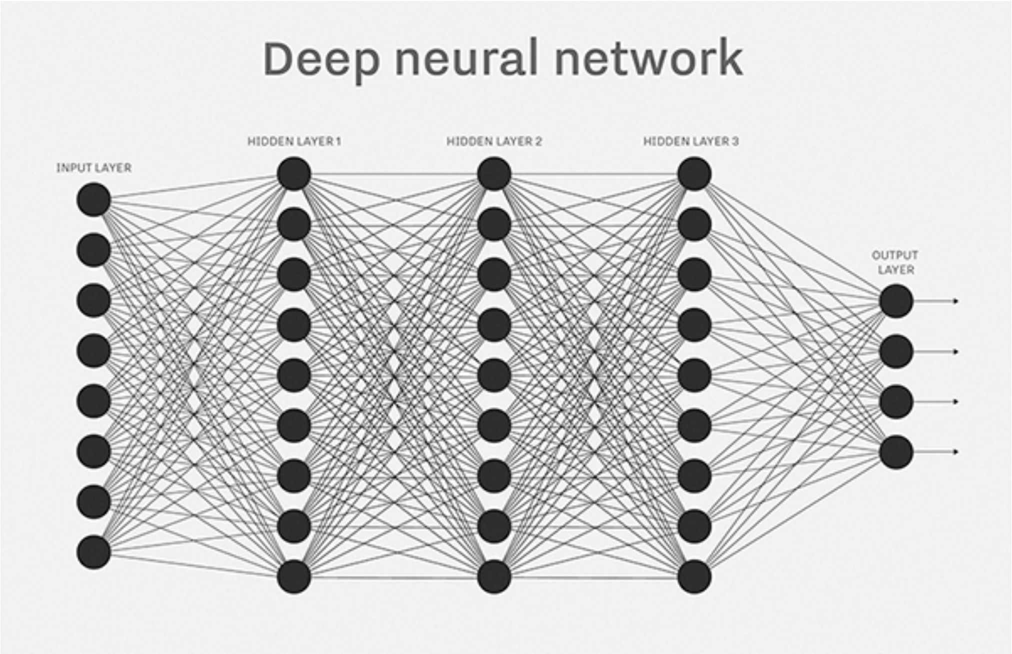
$$\mathbf{p}(\mathbf{x}) = \frac{e^{-\mathbf{H}(\mathbf{x})}}{\sum_y e^{-\mathbf{H}(\mathbf{x})}}$$

$$H_y(\mathbf{x}) = -\ln\left(P(\mathbf{x}|y)P(y)\right)$$

Deep Learning

Why does it work?

- Mimics the natural world
- Generative hierarchy



$$\mathbf{f}(\mathbf{x}) = \boldsymbol{\sigma}_n \mathbf{A}_n \cdots \boldsymbol{\sigma}_2 \mathbf{A}_2 \boldsymbol{\sigma}_1 \mathbf{A}_1 \mathbf{x}$$

$$\mathbf{A}_i \mathbf{x} = \mathbf{W}_i \mathbf{x} + \mathbf{b}_i$$

$$\sigma(x) = max(0, x)$$

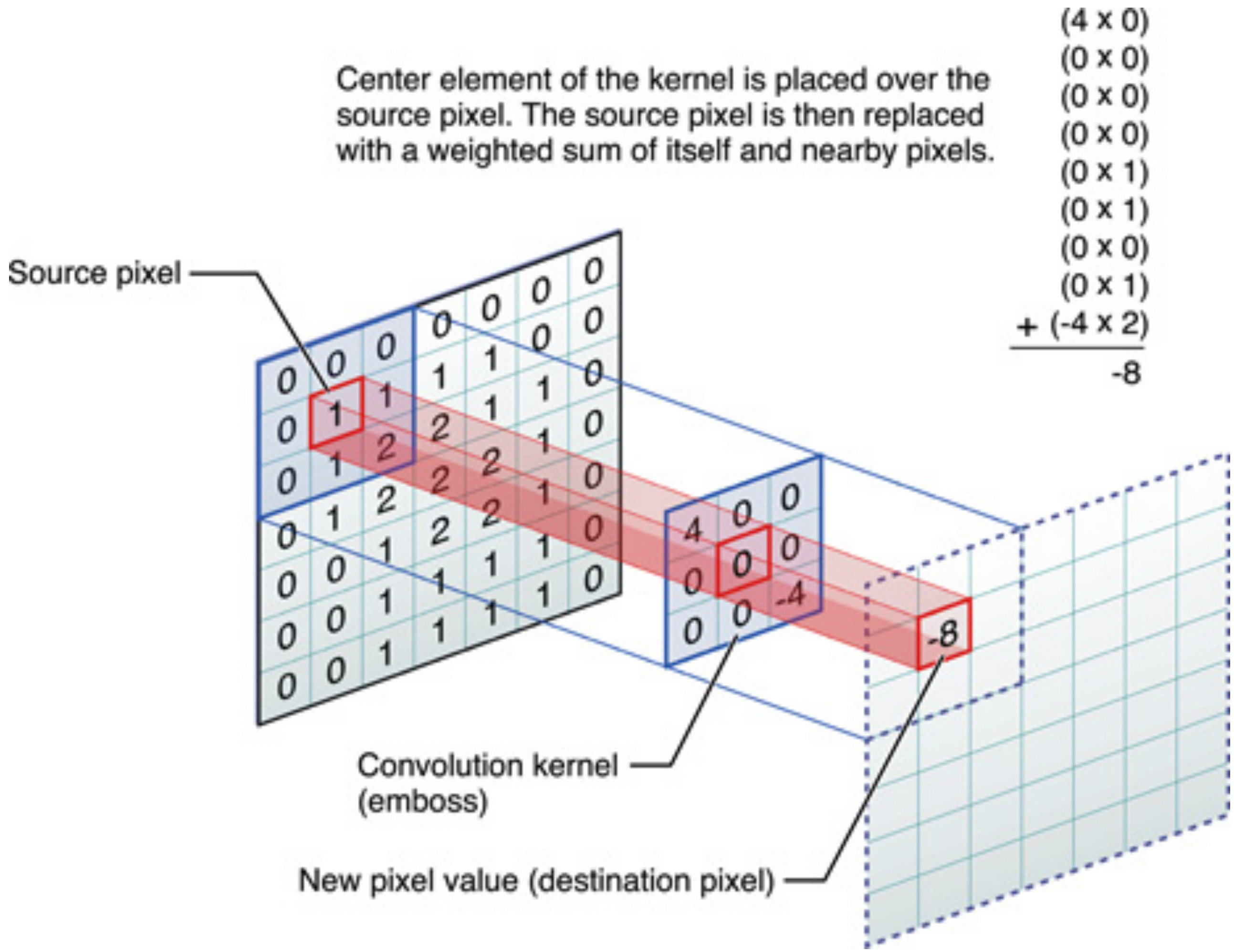
$$\boldsymbol{\sigma}(\mathbf{x}) \equiv \frac{e^{\mathbf{x}}}{\sum_i e^{y_i}}$$

$$\mathbf{p}(\mathbf{x}) = \sigma(-\mathbf{H}(\mathbf{x}))$$

Deep Learning

Convolutional Neural Networks

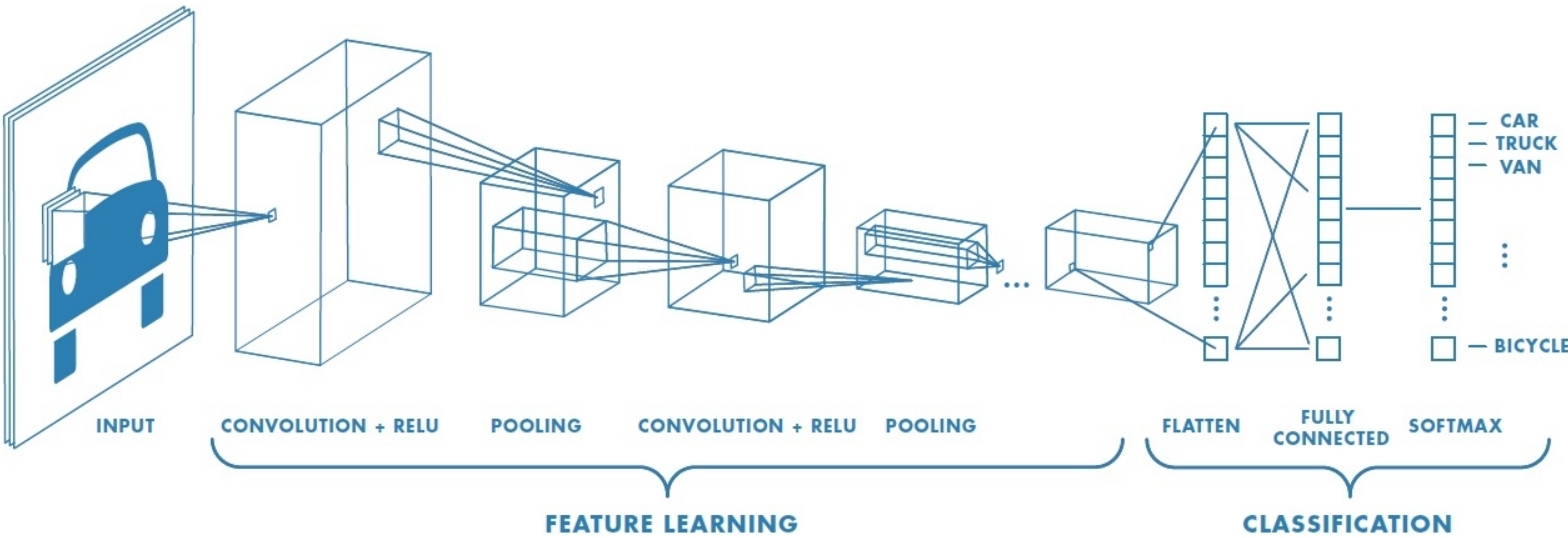
- Reduce number of weights and insert translational invariance



Deep Learning

Convolutional Neural Networks

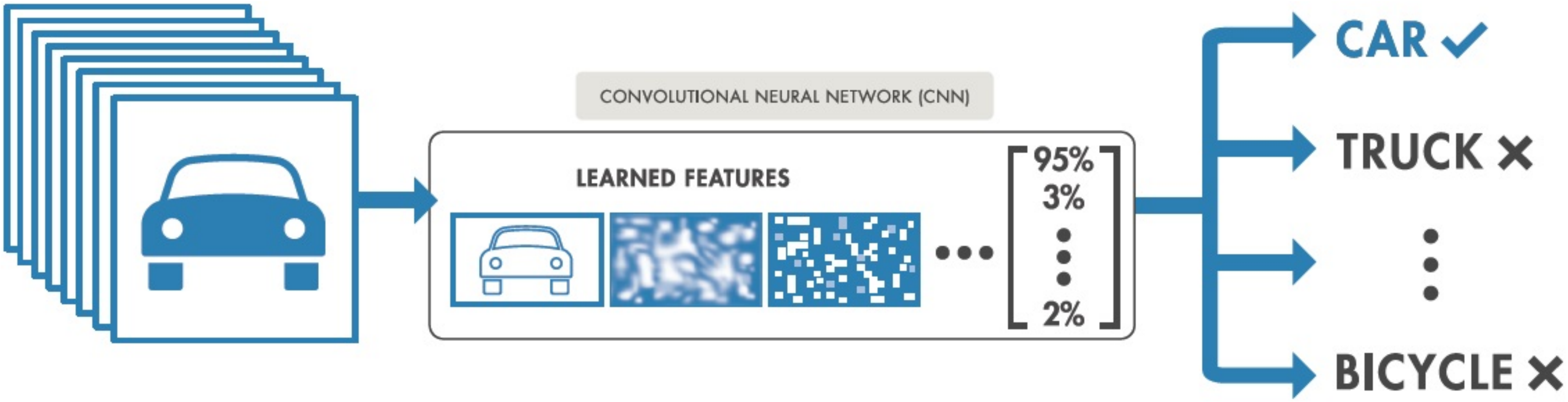
– Reduce number of weights and insert translational invariance



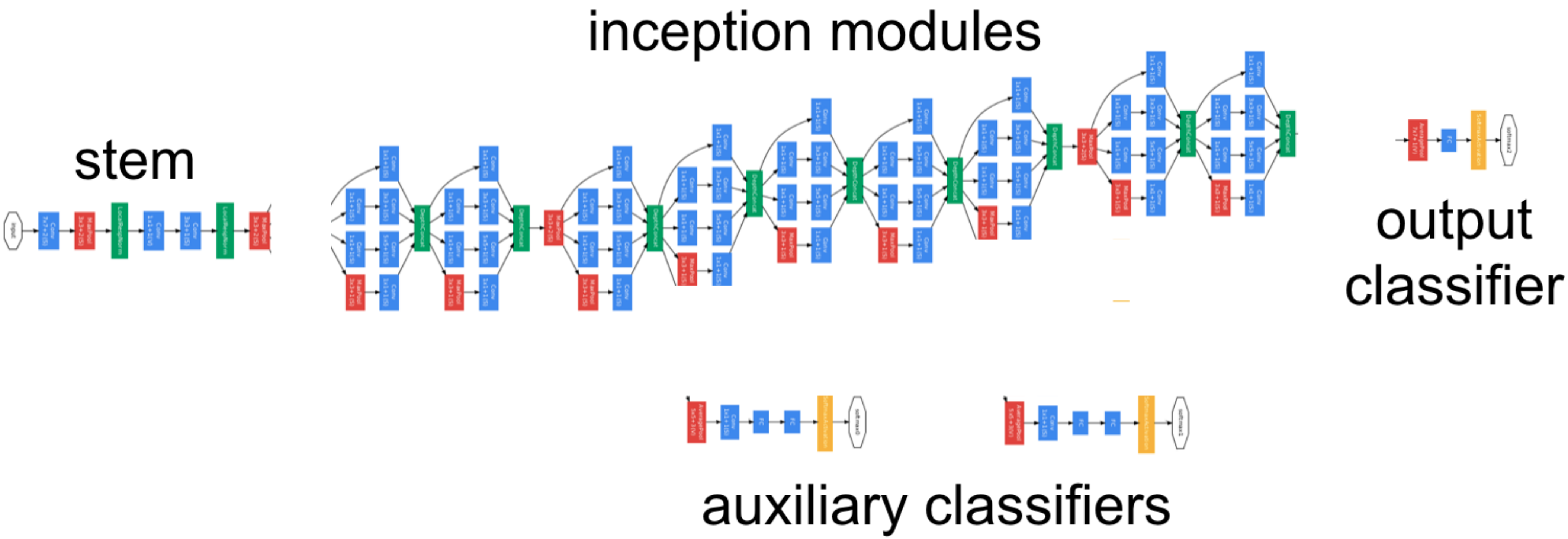
Deep Learning

Convolutional Neural Networks

- Reduce number of weights and insert translational invariance



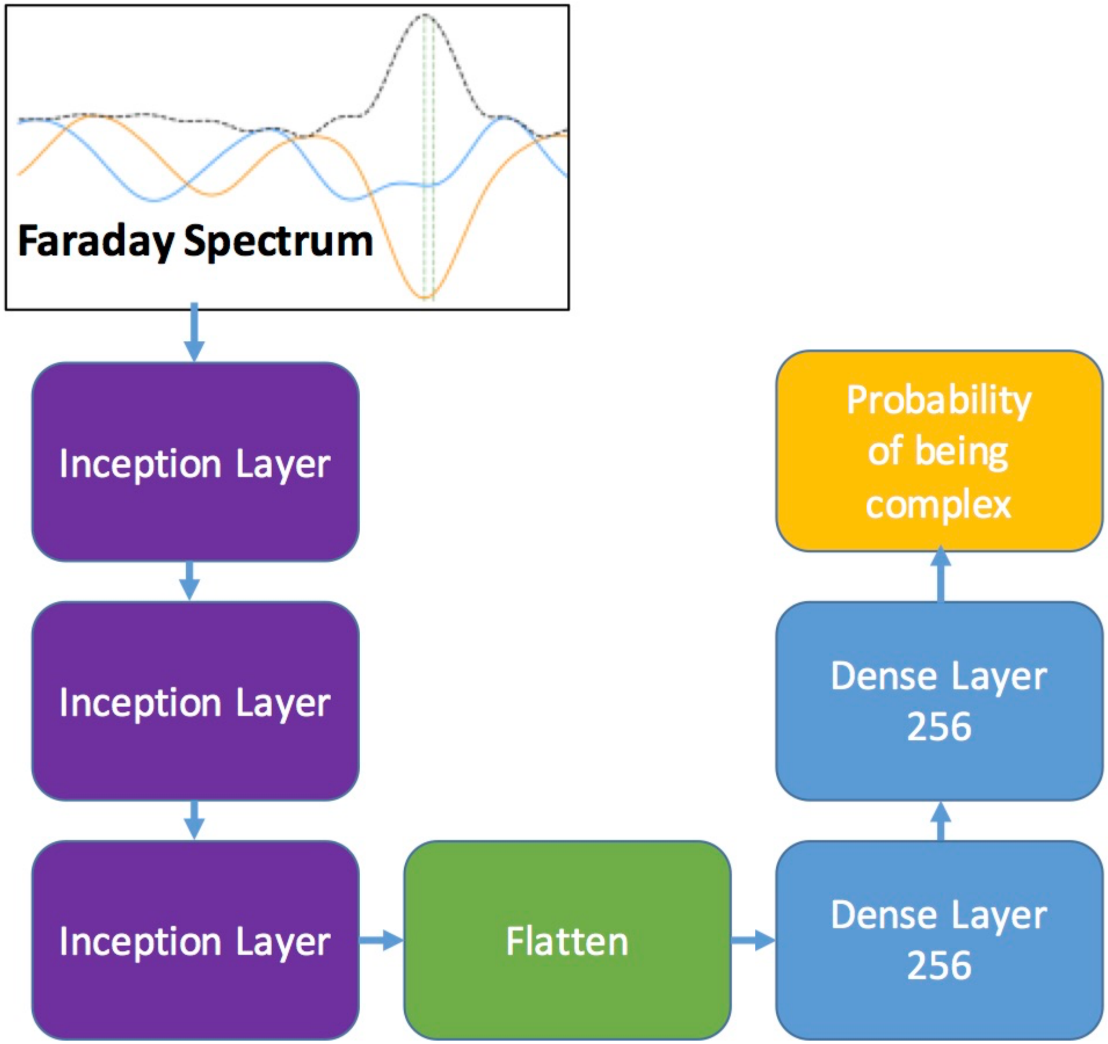
Inception Networks



GoogLeNet team (Szegedy et al. 2014)

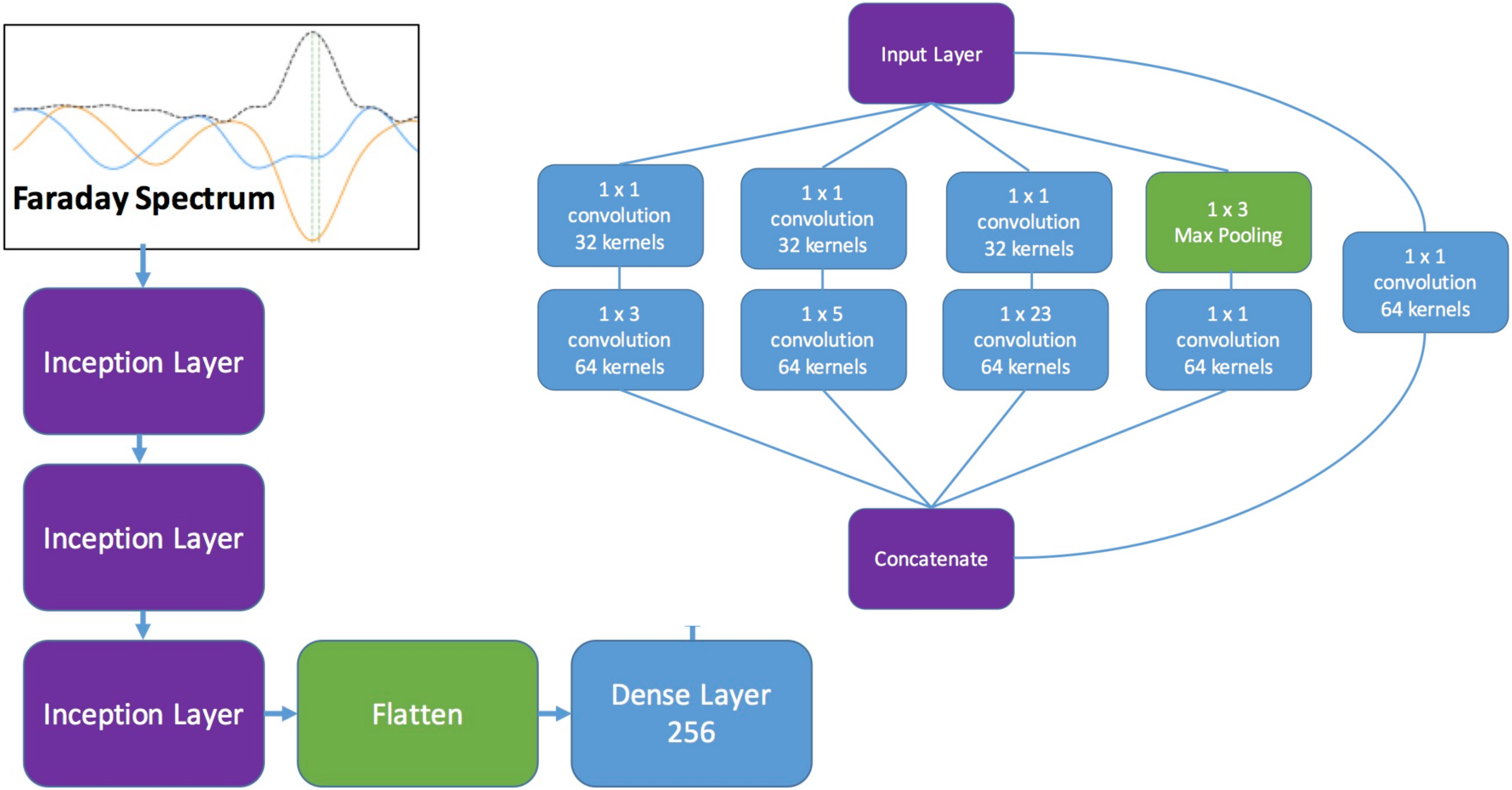
TensorFlow™

Our Network



Faraday Spectrum
Inception Layer
Inception Layer
Inception Layer
Flatten
Dense Layer 256
Dense Layer 256
Probability of being complex

Our Network


Faraday Spectrum
Inception Layer
Inception Layer
Inception Layer
Flatten
Dense Layer
256
Input Layer
1 x 1
convolution
32 kernels
1 x 1
convolution
32 kernels
1 x 1
convolution
32 kernels
1 x 3
Max Pooling
1 x 1
convolution
64 kernels
1 x 3
convolution
64 kernels
1 x 5
convolution
64 kernels
1 x 23
convolution
64 kernels
1 x 1
convolution
64 kernels
Concatenate

Faraday Rotation

MNRAS **000**, 000–000 (0000) Preprint 23 October 2017 Compiled using MNRAS LaTeX style file v3.0

Classifying Complex Faraday Spectra with Convolutional Neural Networks

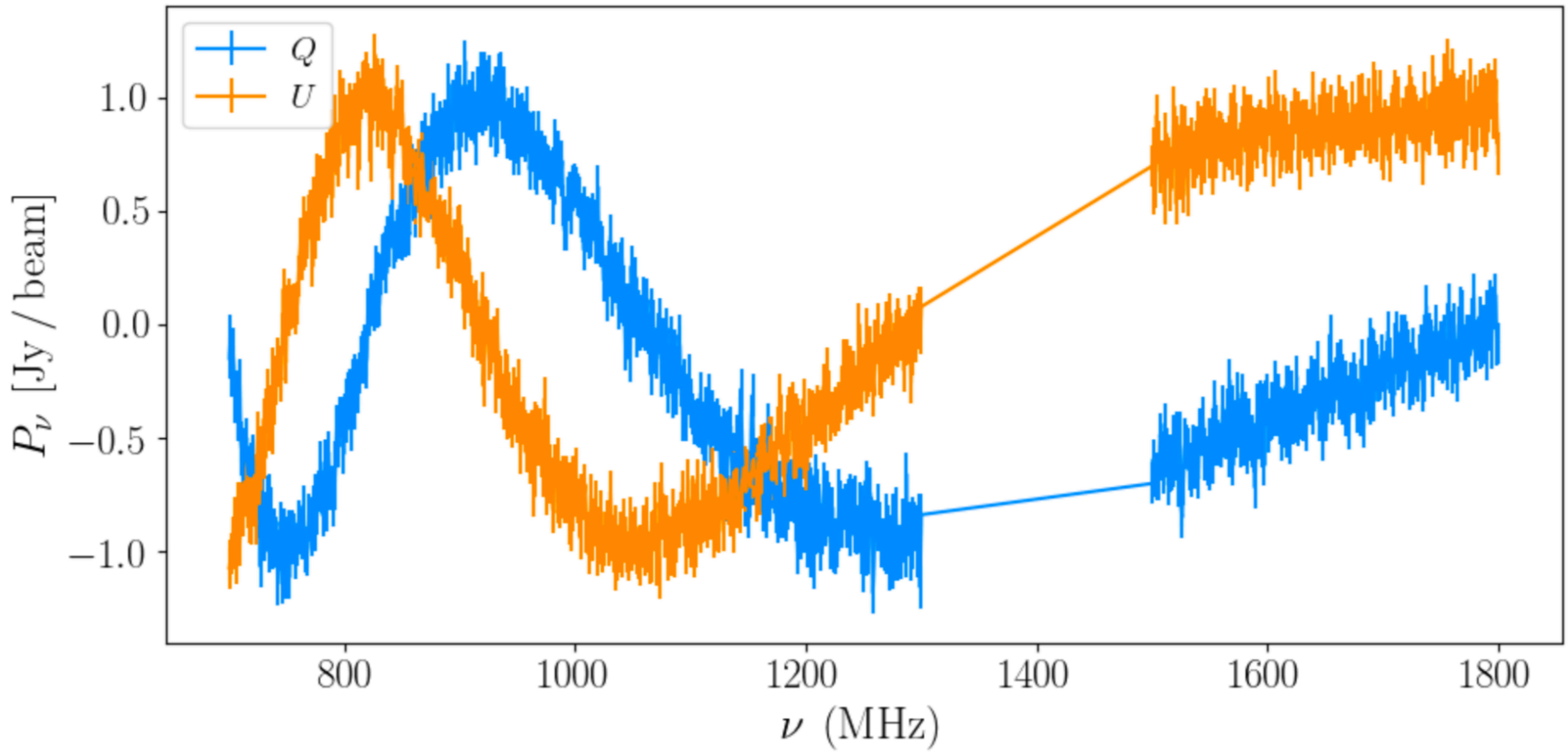
Shea Brown[1*], Brandon Bergerud[1], Allison Costa[1], Bryan Gaensler[2], Jacob Isbell[1], Daniel LaRocca[1], Ray Norris[3], Cormac Purcell[4], Lawrence Rudnick[5], Xiaohui Sun[6]

[1] *Department of Physics & Astronomy, The University of Iowa, Iowa City, IA, 52245*
[2] *Dunlap Institute, University of Toronto*
[3] *Western Sydney University*
[4] *Research Centre for Astronomy, Astrophysics, and Astrophotonics, Macquarie University, NSW 2109, Australia*
[5] *University of Minnesota*
[6] *Department of Astronomy, Yunnan University, and Key Laboratory of Astroparticle Physics of Yunnan Province, Kunming, 650091, China*

ABSTRACT

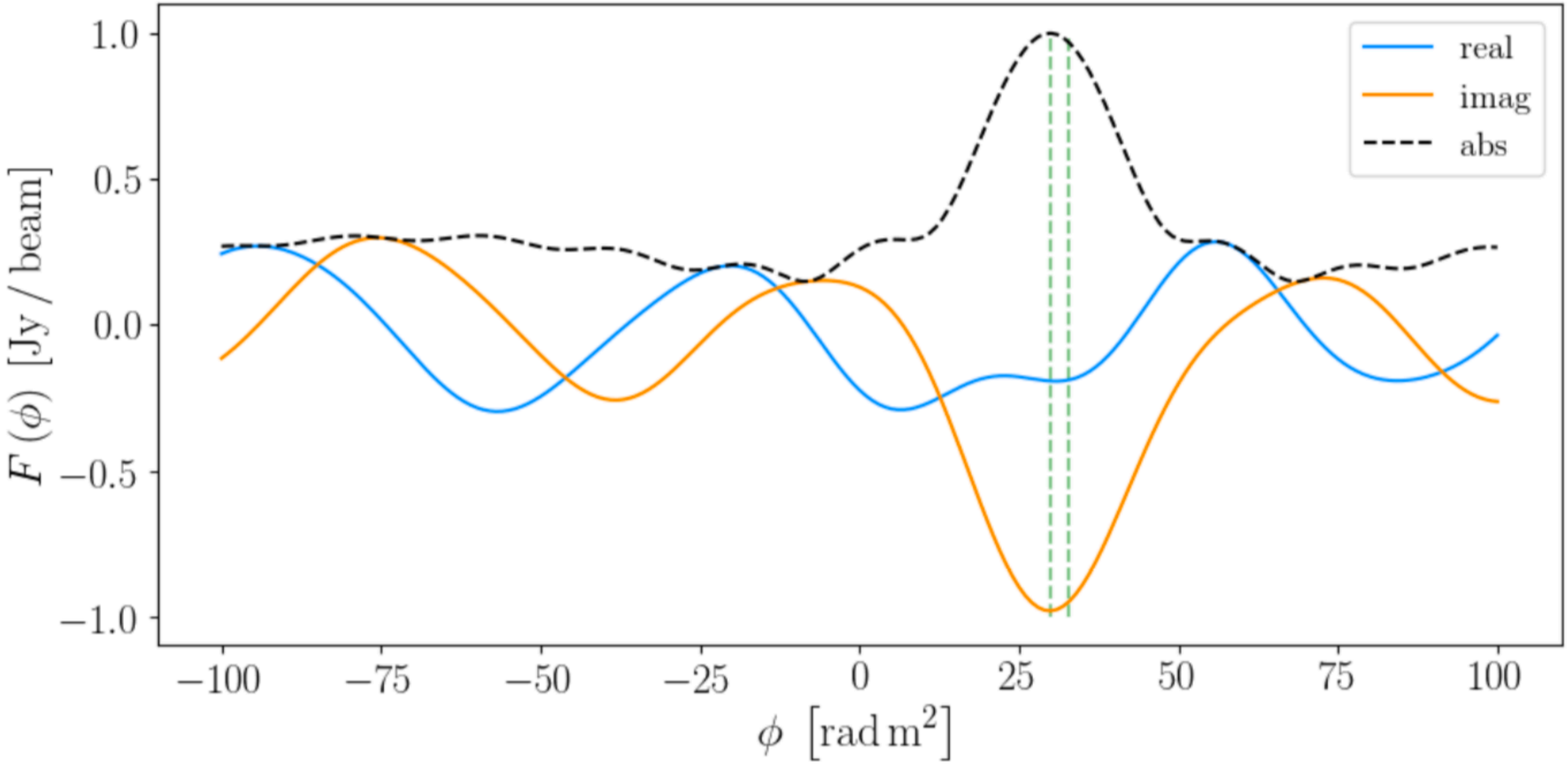
Advancements in radio spectro-polarimetry offer the possibility to disentangle complex regions where relativistic and thermal plasmas mix in the interstellar and intergalactic media. Recent work has shown that even apparently simple Faraday Rotation Mea-

Faraday Rotation



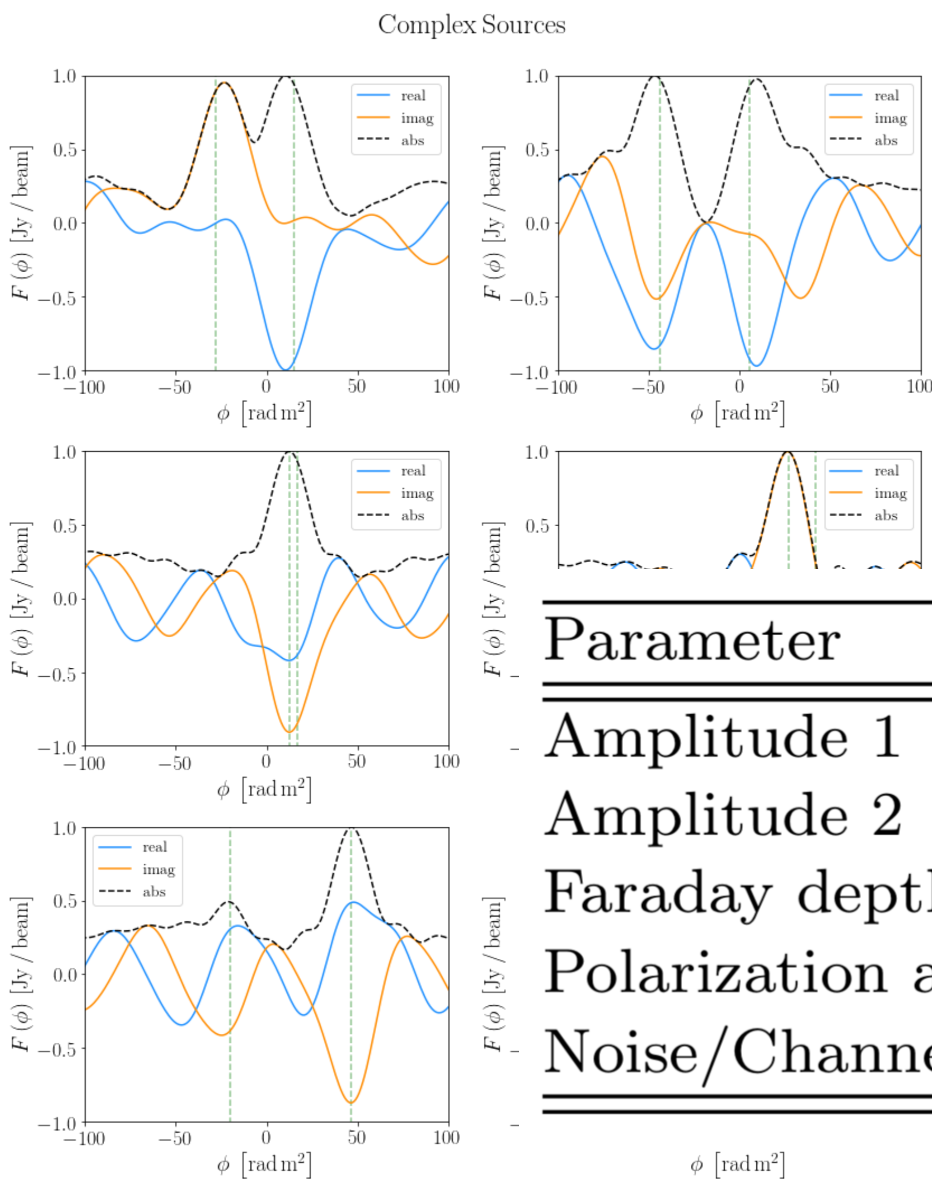
$$P(\lambda^2) = P_1 e^{\left[2i(\chi_1 + \phi_1 \lambda^2)\right]} + P_2 e^{\left[2i(\chi_2 + \phi_2 \lambda^2)\right]}$$

Faraday Rotation



$$F(\phi) \approx K \sum_{i=1}^{N} P_i \mathrm{e}^{-2\mathrm{i}\phi(\lambda_i^2 - \lambda_0^2)}$$

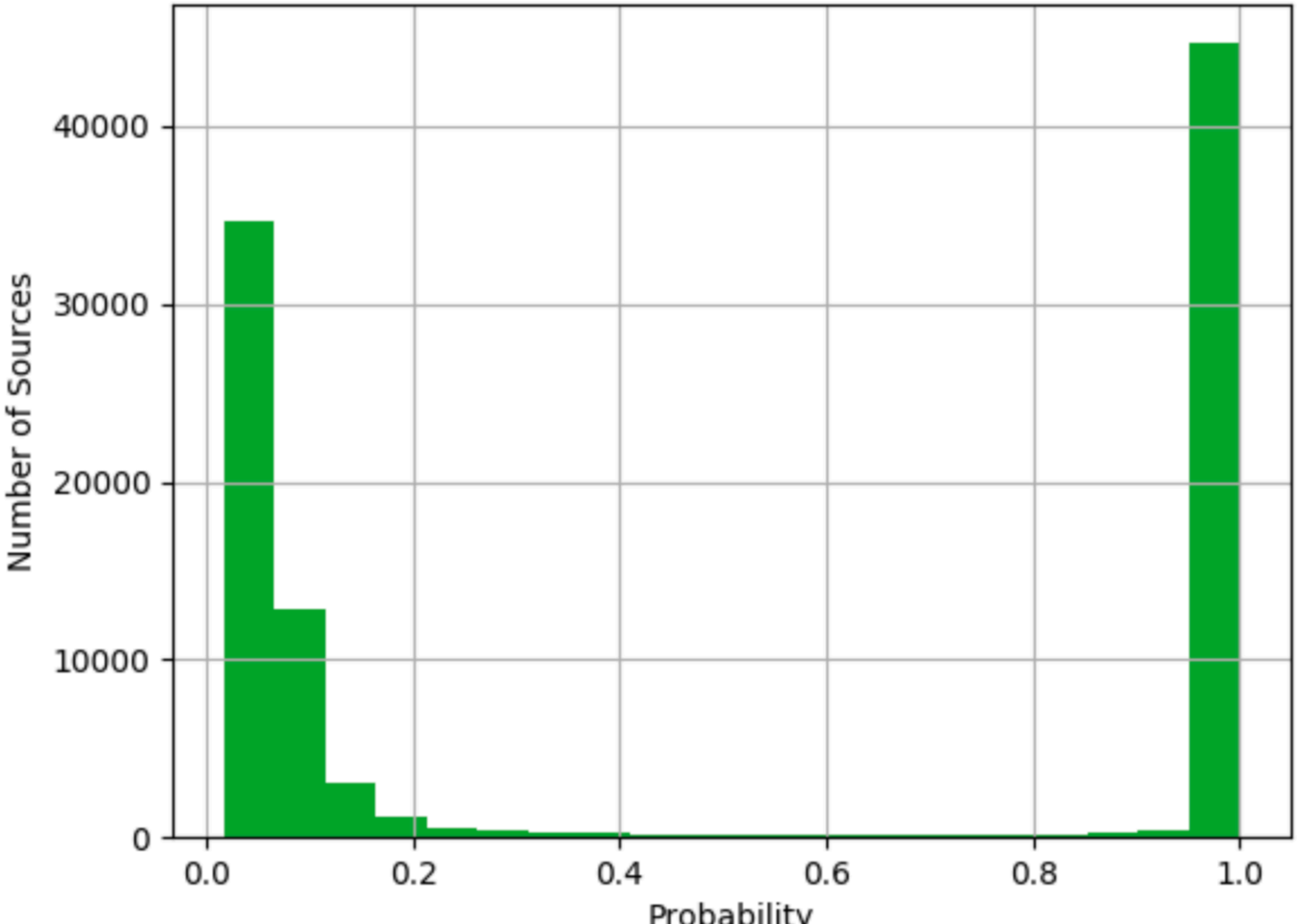
Faraday Rotation



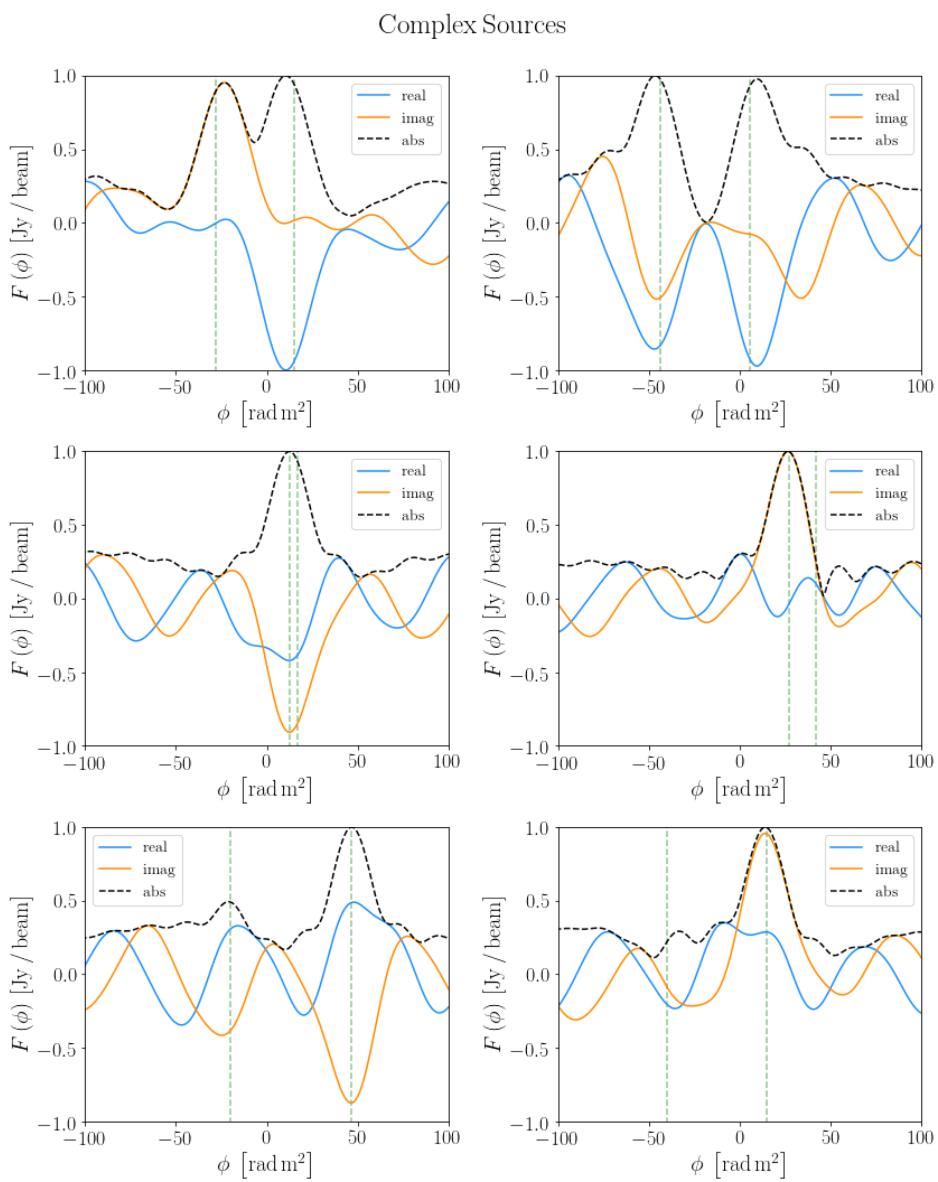
- Simulate simple and complex (two component) sources over relevant parameter space
- Train CNN on simulated sources
- Test CNN on new sim. sources

Parameter	Symbol	Range
Amplitude 1	P_1	1
Amplitude 2	P_2	[0, 1]
Faraday depth {1, 2}	$\phi_{\{1,2\}}$	[-50, +50]
Polarization angle {1, 2}	$\chi_{\{1,2\}}$	[0, $+\pi$]
Noise/Channel	σ	[0, 0.333]

Faraday Rotation



Faraday Rotation



Simulate simple and complex (two component) sources over relevant parameter space

Train CNN on simulated sources

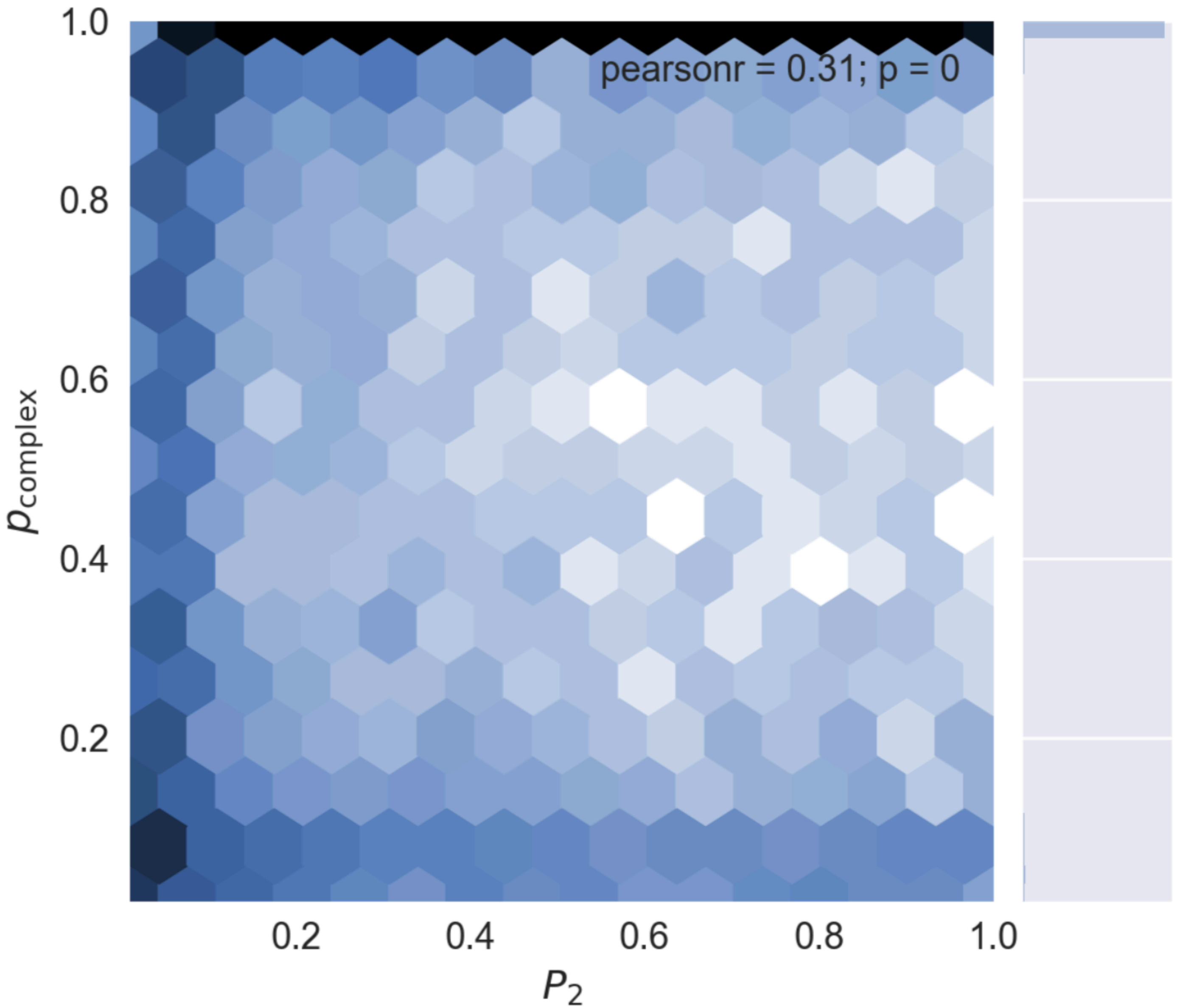
Test CNN on new sim. Sources

~3.0% false positives

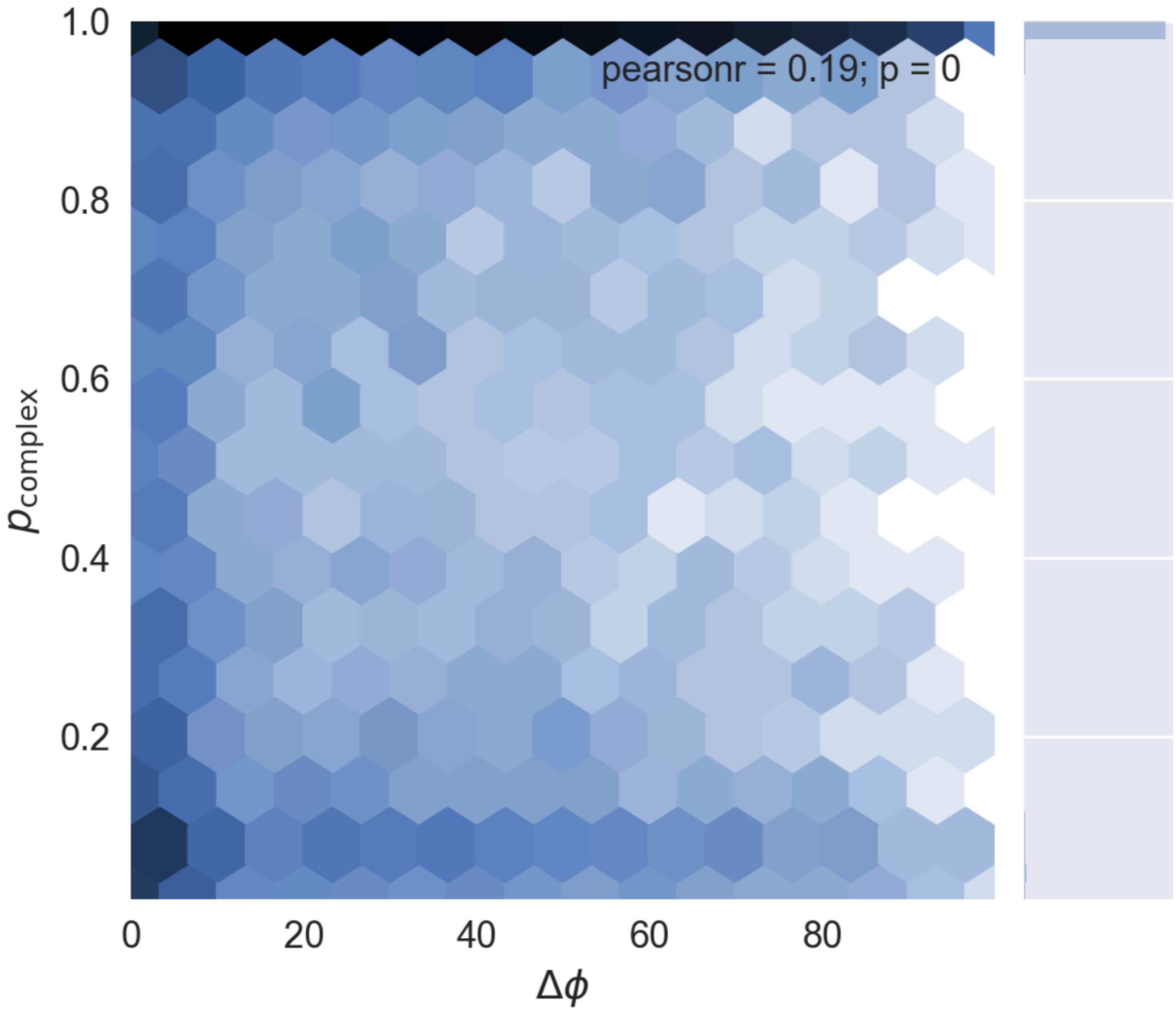
~7.2% false negatives

Predicted −>	Simple	Complex
True Simple	48,318	1481
True Complex	3,618	46,583

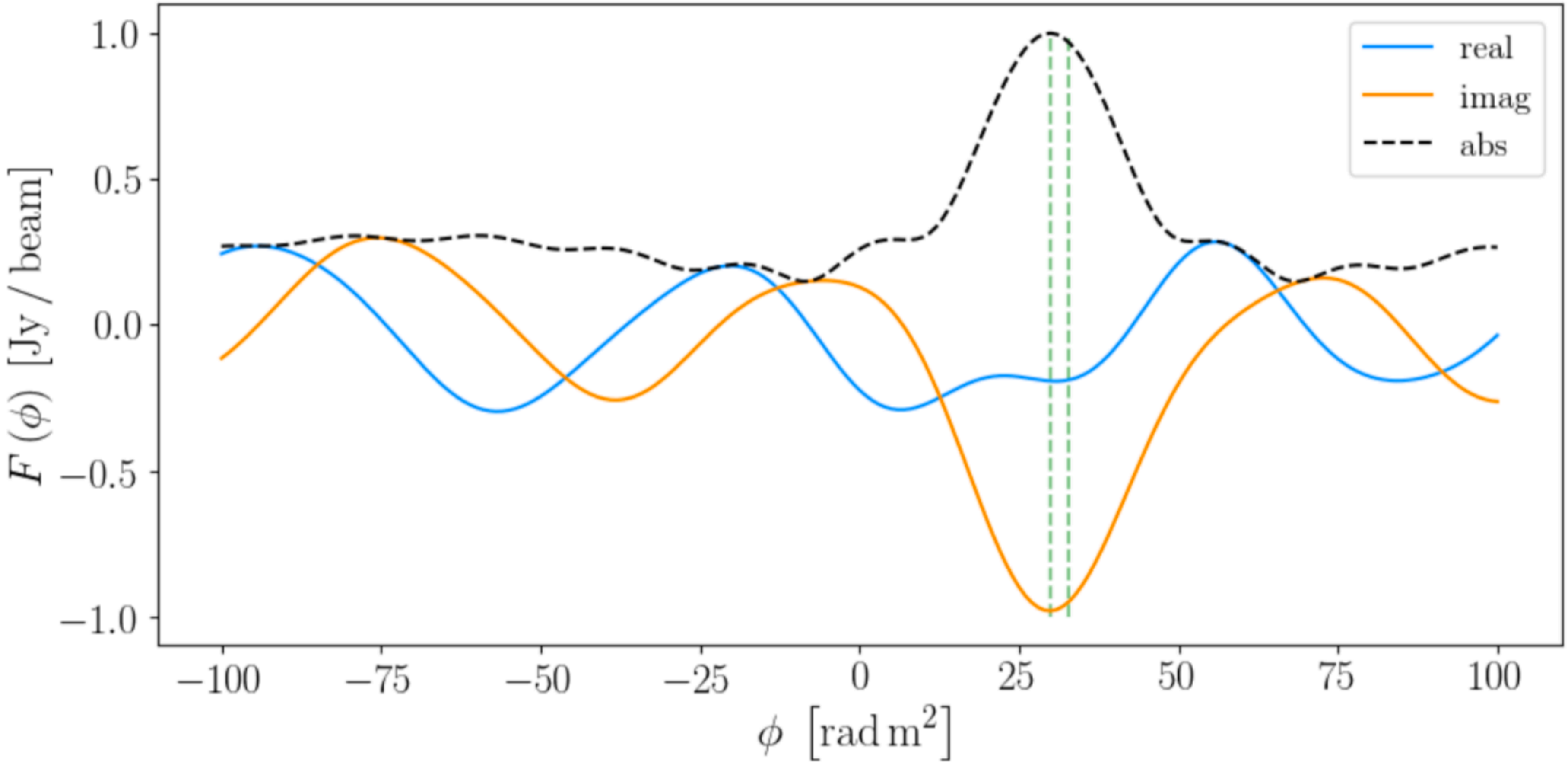
Faraday Rotation



Faraday Rotation

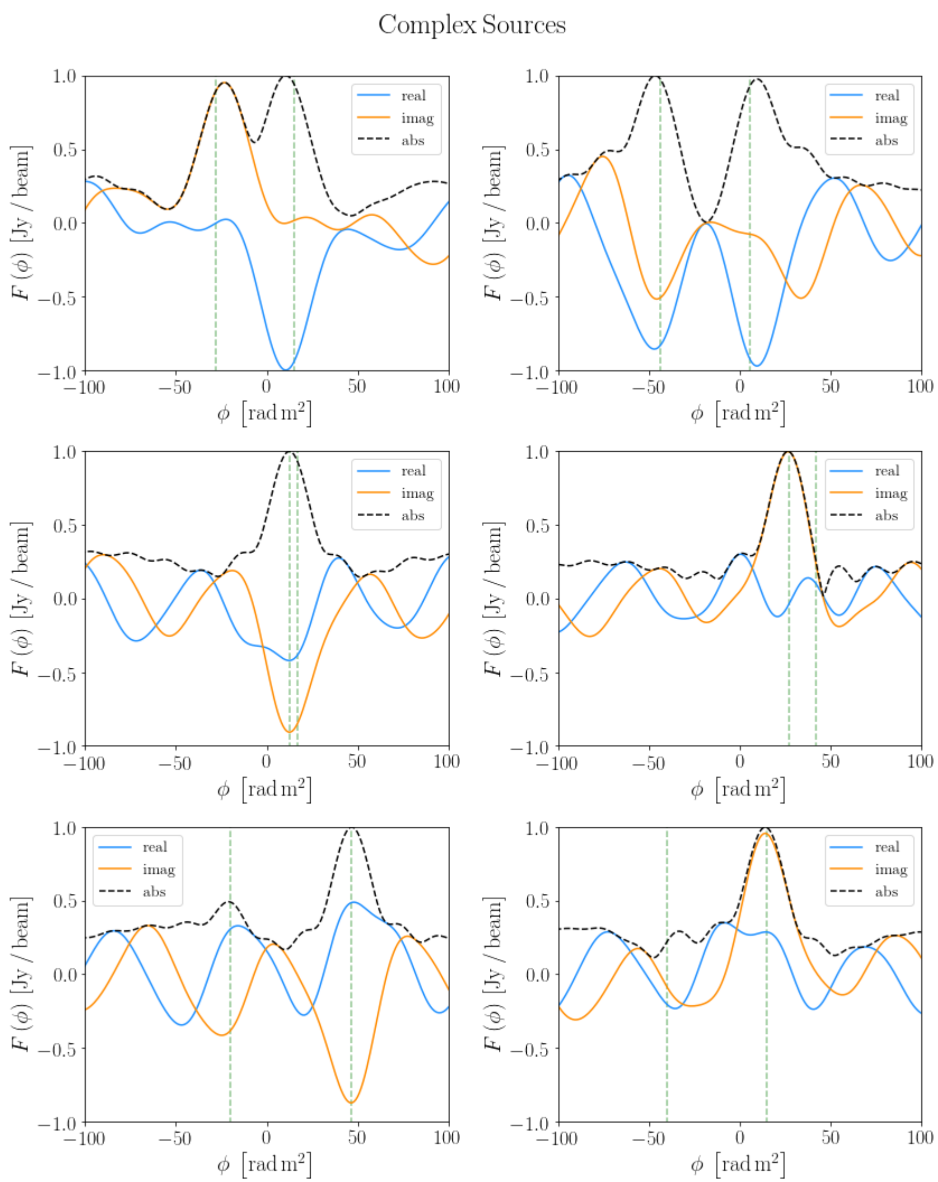


Faraday Rotation



$$F(\phi) \approx K \sum_{i=1}^{N} P_i \mathrm{e}^{-2\mathrm{i}\phi(\lambda_i^2 - \lambda_0^2)}$$

Faraday Rotation



Sub-space that might make it into the catalog:

Delta phi < 10% FWHM

S/N of P_1>5, & P_2/P_1 > 0.1

<0.3% false positives

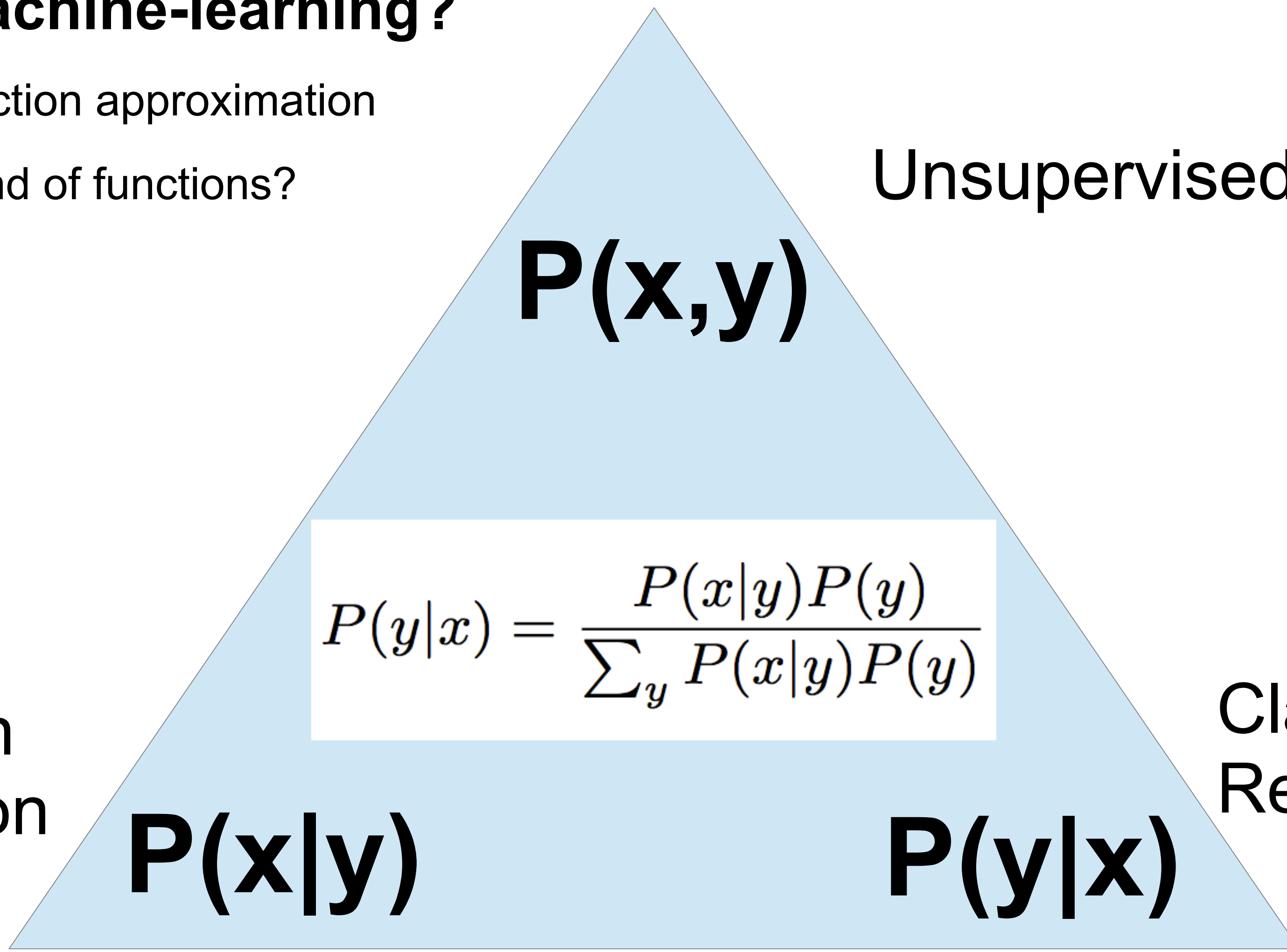
<1.0% false negatives

Predicted – >	Simple	Complex
True Simple	29,281	69
True Complex	247	25,337

Machine Learning

What is machine-learning?

- Just function approximation
- What kind of functions?



Adapted from Lin & Tegmark 2016

Thank You!