Dualities in and from Machine Learning

LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

Sven Krippendorf

Corfu, Recent Developments in Strings and Gravity
September 12th 2019

Current ML applications in high energy

Improving sensitivity

Day, SK 1907.07642

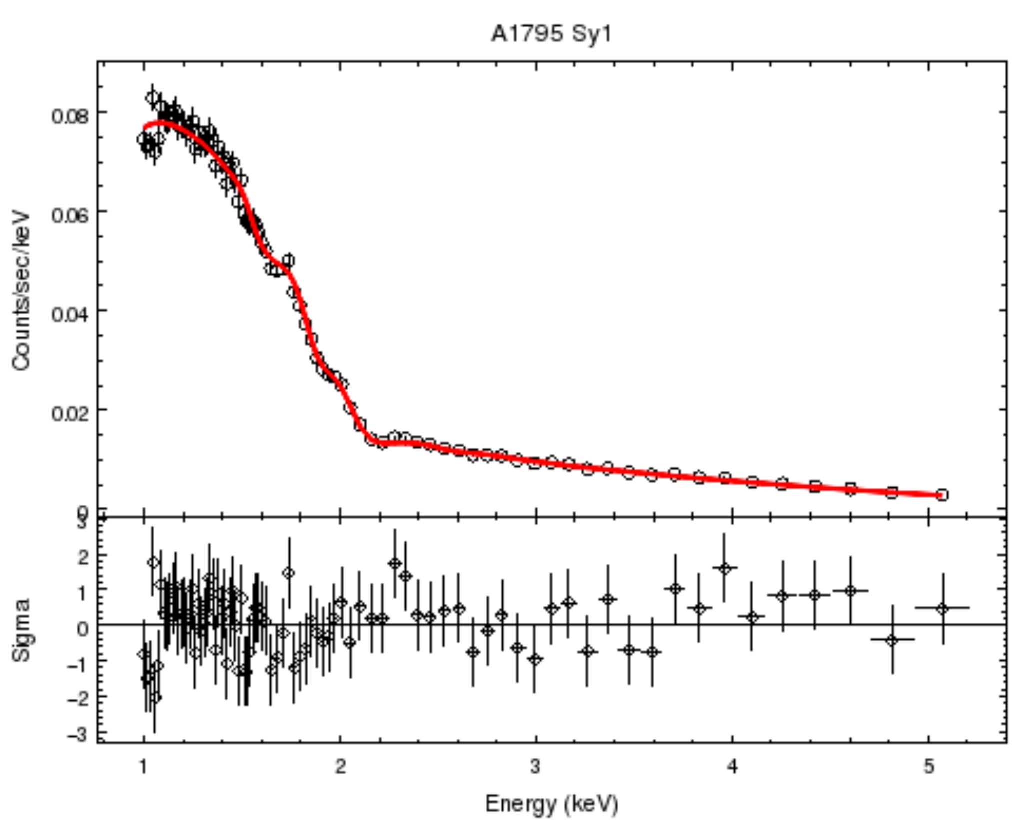
- ML-techniques heavily used in experimental bounds.
- Brief example: Improving sensitivity for ultra-light axion-like particles, compared to previous bounds.
- ML algorithms good at classification. Detecting particles is a classification problem. Our classifiers:

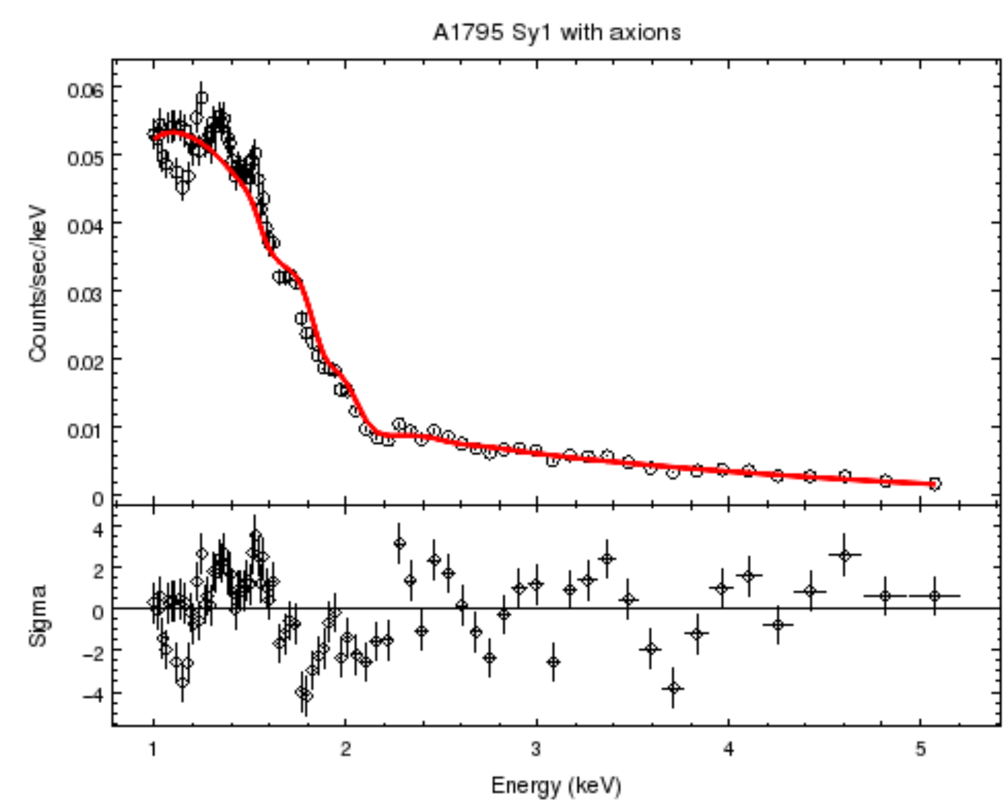
- Training: Simulate data with and without axions for appropriate X-ray sources
- Bounds: Compare fake & real data performance
- Algorithms (sklearn): decision trees, boosted decision trees, random forests, Gaussian Naive Bayes, Gaussian Process classifier, SVM, …

Previous bounds:
NGC1275: 1605.01043, Other sources: 1704.05256, Athena bounds: 1707.00176
with: Conlon, Day, Jennings; Berg, Muia, Powell, Rummel

Improving sensitivity

- Data: Chandra X-ray observations of bright point sources (AGN, Quasar) in or behind galaxy clusters
- Bounds for ALPs with $m<10^{-12}$ eV due to absence of characteristic spectral modulations caused by interconversion between photons and axions in cluster background magnetic field

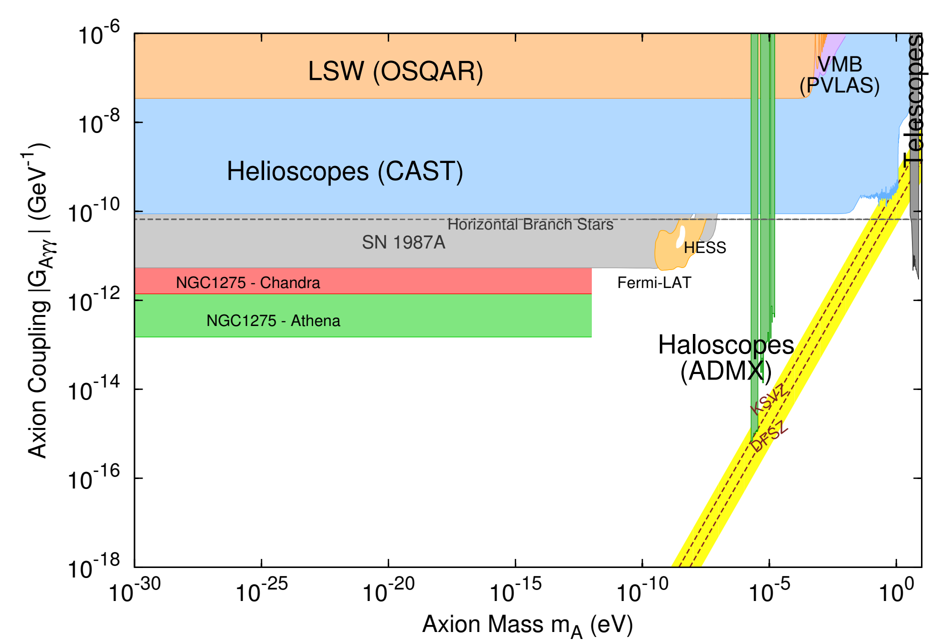




Improving sensitivity

- Data: Chandra X-ray observations of bright point sources (AGN, Quasar) in or behind galaxy clusters
- Bounds for ALPs with $m<10^{-12}$ eV due to absence of characteristic spectral modulations caused by interconversion between photons and axions in cluster background magnetic field

x 10^{-12} GeV^{-1}	AB	DTC	GaussianNB	QDA	RFC	Previous
A1367 (resid.)	1.9	-	-	-	-	2.4
A1367 (up-resid.)	2.0	-	1.9	-	-	2.4
A1795 Quasar (resid.)	-	-	1.7	-	1.4	>10.0
A1795 Quasar (up-resid.)	-	-	-	-	-	>10.0
A1795 Sy1 (resid.)	1.0	0.8	1.2	1.1	0.7	1.5
A1795 Sy1 (up-resid.)	1.1	1.1	1.1	1.0	0.8	1.5



ML for the string landscape?

ML for string landscape

- ML good at following tasks:
  a) analysing & simulating large amounts of data
  b) approximating some unknown functional dependence
- Two examples:
  a) large amounts of string vacua (structure too sophisticated for ordinary MC) **Abel, Rizos; SK, Erbin; Vaudrevange; Altman, Carifio, Halverson, Nelson; Halverson, Nelson, Ruehle; Cole, Schachner, Shiu; …**

  b) performance tests on topological quantities where formulae are known (very flexible fitting algorithm)
  He; Ruehle; Klaewer, Schlechter; Bull, He, Jejjala, Mishra; He, Lee; Brodie, Constantin, Deen, Lukas; …
- What can we expect? development of string landscape analysis tools, functional approximations to many desired phenomenological quantities.

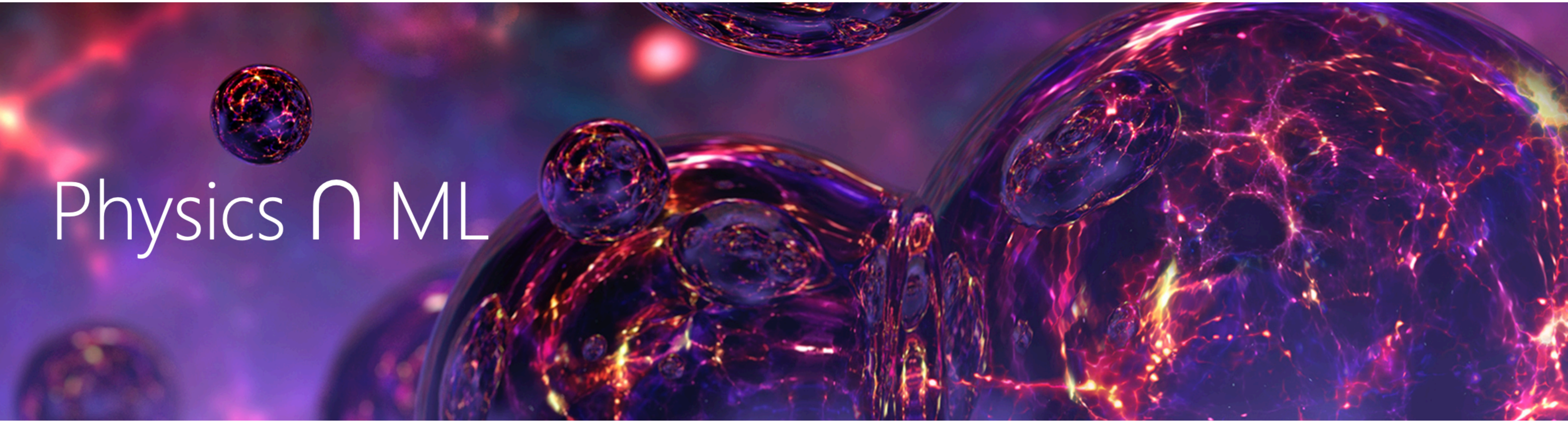
… more in Fabian’s talk

Other avenues?

“Don’t ask what ML can do for you,
ask what you can do for ML.”

– *Gary Shiu*

Physics $\cap$ ML



Overview

Region: North America

Date: April 25, 2019 - April 26, 2019

Location: Microsoft Research Redmond

About Agenda Abstracts Videos

The goal of *Physics ∩ ML* (read 'Physics Meets ML') is to bring together researchers from machine learning and physics to learn from each other and push research forward together. In this inaugural edition, we will especially highlight some amazing progress made in string theory with machine learning and in the understanding of deep learning from a physical angle. Nevertheless, we invite a cast with wide ranging expertise in order to spark new

Physics ∩ ML

Dualities

Betzler, SK: 190x.xxxxx

Why are dualities exciting?

- Multiple EFTs with different DOF describing the same system.
- Present in many dynamical systems: condensed matter physics, AdS/CFT, string dualities
- Allow us to describe dynamics of strongly coupled systems via dual weakly coupled descriptions
- Allow us to get EFT-operators at higher accuracy than normally allowed (theory: large number of diagrams, experiment: large amount of data). Think about Yukawa couplings in heterotic standard embedding
- …

Why are dualities exciting?

- Multiple EFTs with different DOF describing the same system.
- Present in many dynamical systems: condensed matter physics, AdS/CFT, string dualities
- Allow us to describe dynamics of strongly coupled systems via dual weakly coupled descriptions
- Allow us to get EFT-operators at higher accuracy than normally allowed (theory: large number of diagrams, experiment: large amount of data). Think about Yukawa couplings in heterotic standard embedding
- …

What’s your favourite application?

Dualities - Outline

- How are dualities useful in practice?
- What’s special about dual representations?
- Ways to learn/enforce duality transformation via neural networks

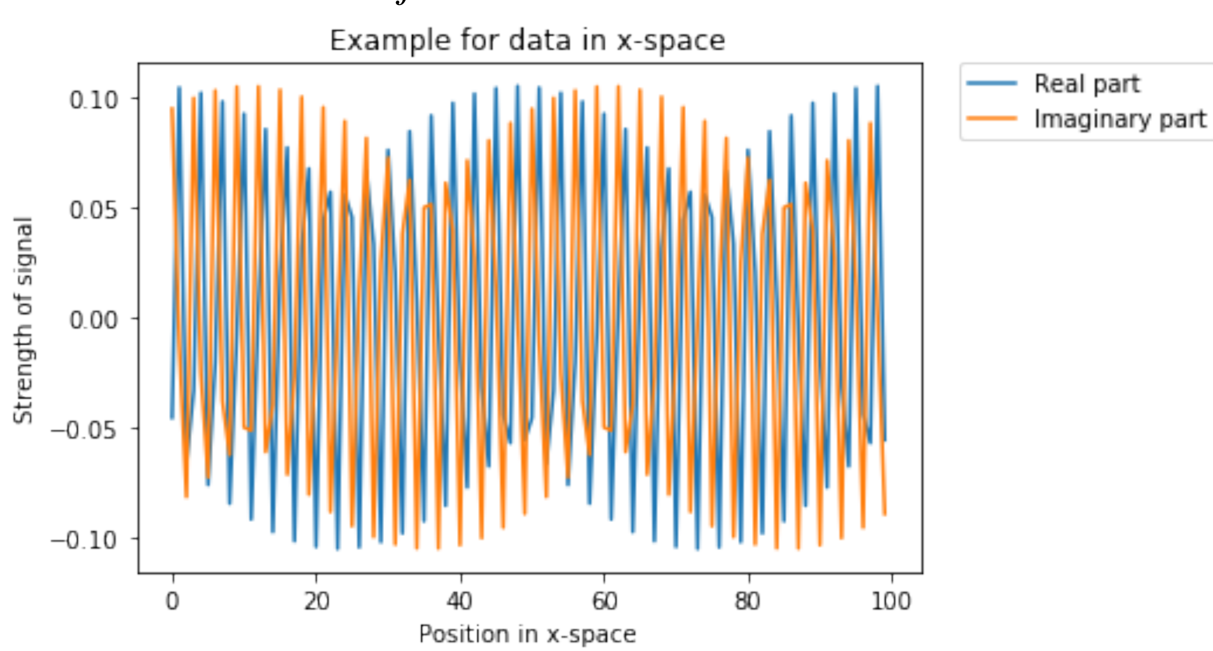
How are dualities useful in practice?

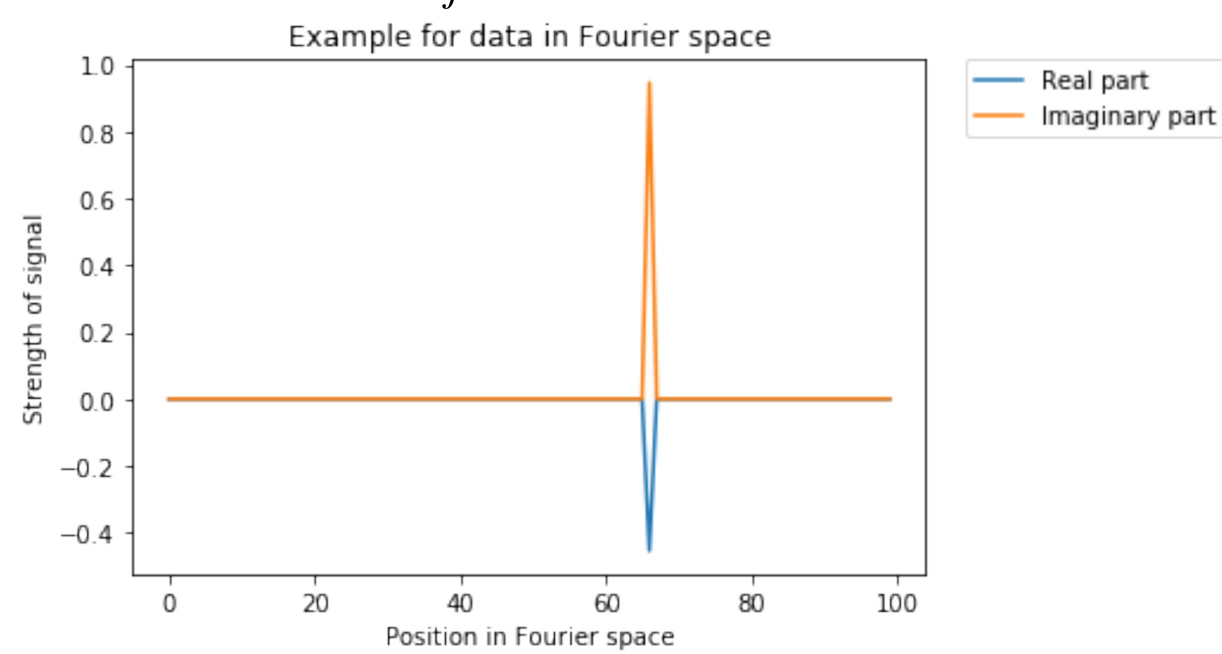
aka connecting physics questions to data questions

Dualities: Practical Use

- (Deep) Neural networks transform data into different representations (if no coarse graining takes place). For instance, these representations can be used for more efficient classification.
- Finding good data representations is hard and often out of the realm of current optimisation strategies.
- Example: Discrete Fourier Transformation (DFT)

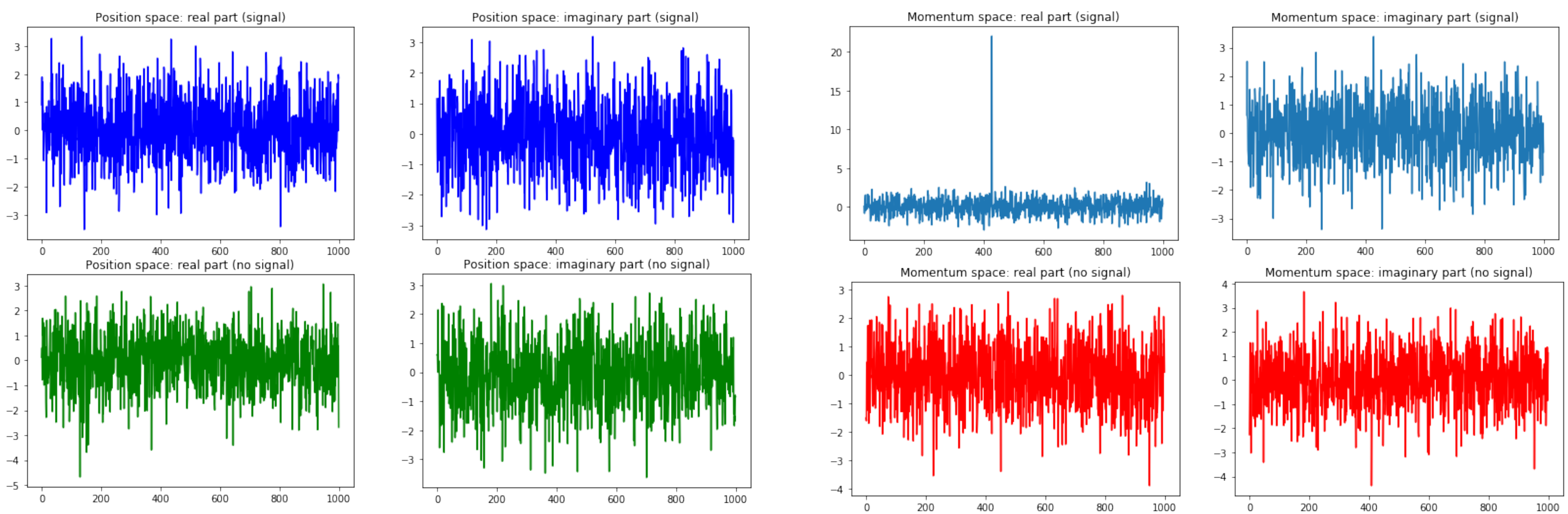
$$x_k = \frac{1}{n}\sum_{j=1}^{n} p_j\, e^{2\pi ijk/n} \qquad\qquad p_k = \sum_{j=1}^{n} x_j\, e^{-2\pi ijk/n}$$





DFT: data question

- Let’s confront this with a data question: Is there a signal in the noise?



- Let’s check the performance on simple networks

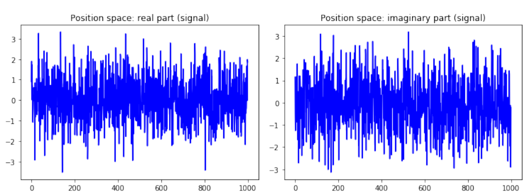
DFT: simple network

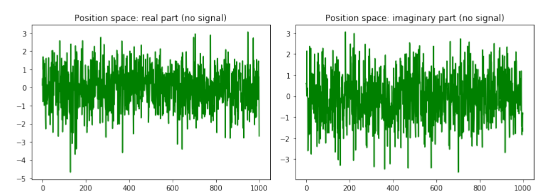
- Supervised learning task (binary classification):

N discrete values

$\{((\mathbf{x}_R, \mathbf{x}_I), y)\}$ $\{((\mathbf{p}_R, \mathbf{p}_I), y)\}$

$y = 1$ **noise + signal**





$y = 0$ **noise**

- Network

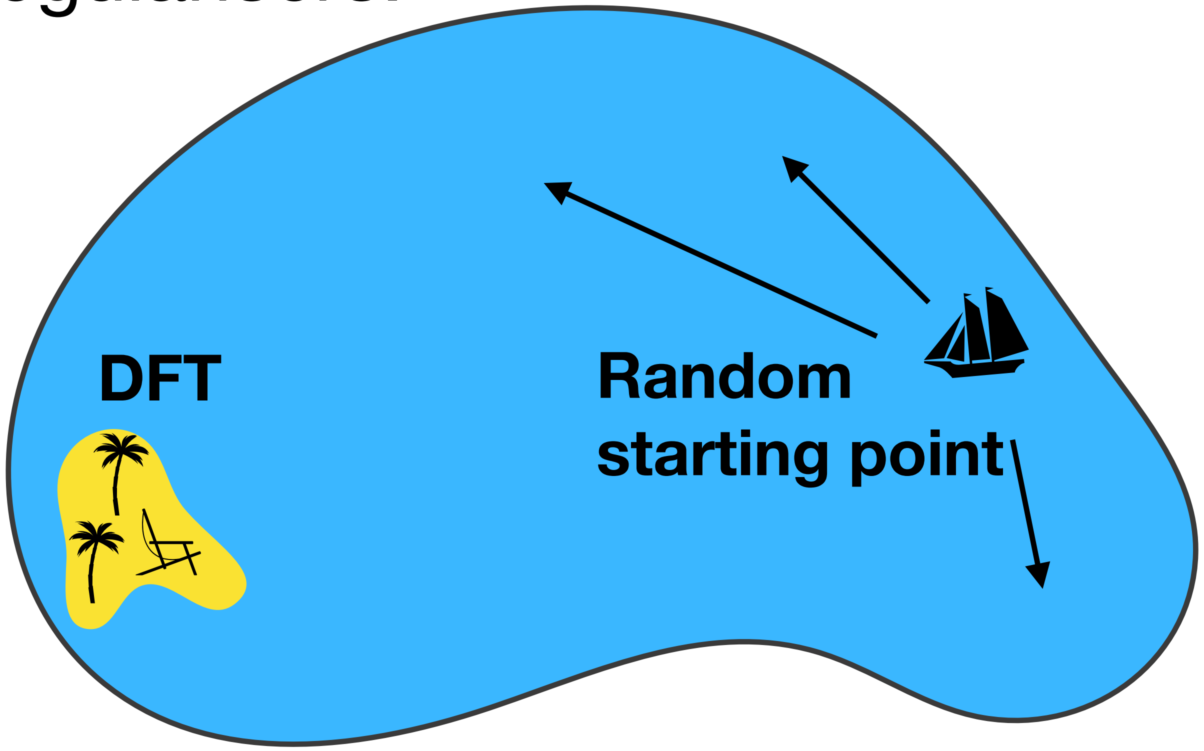
```
# ===================================================
# Define simple 1-layer-CNN.
# ===================================================

model = Sequential()
model.add(Conv1D(2, kernel_size=2,
                 activation='relu',
                 input_shape=(2*number_bins,1)))
model.add(Flatten())
model.add(Dense(1, activation='sigmoid'))
```

For this network classification works in momentum space, but not in position space.

Utilising dual representation

- Goal: improve performance on position space.
- Deeper network? Can do the job in principle
  [DFT can be implemented with a single dense layer]
- However finding it dynamically is `impossible' with standard optimisers, initialisations, and regularisers.

Layer	Shape	Parameters
Dense	(2000,2)	16000000
Conv1D	(2000,2)	4
Activation	(2000,2)	-
Dense	1	4001
Activation	1	-



DFT from modified loss

- Adapt loss function to achieve feature separation, i.e. separating the two classes of data (inspired by triplet loss) [towards generating dual representations dynamically]

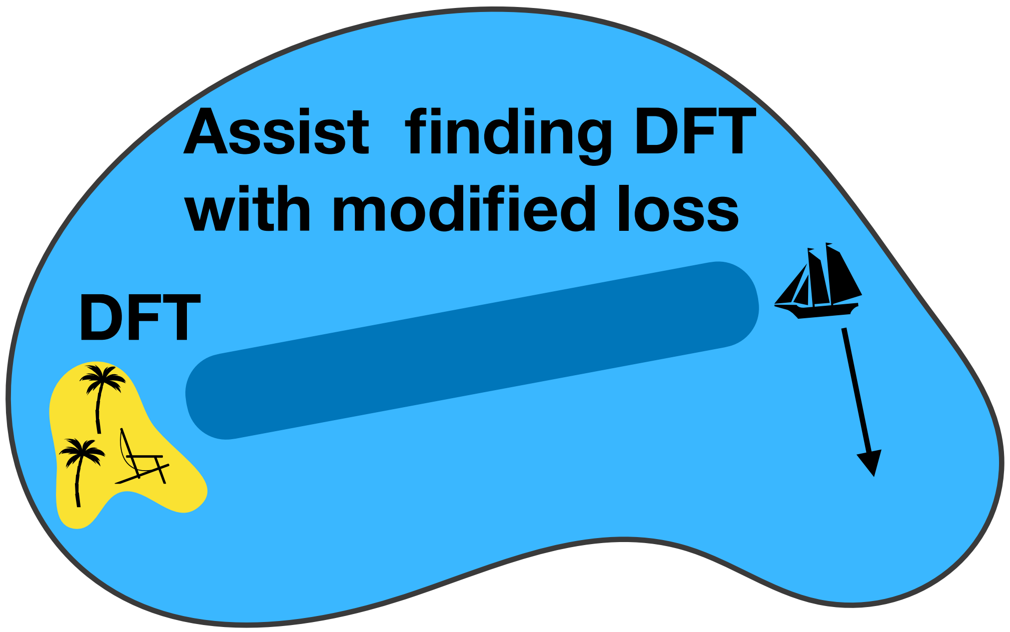
1503.03832

Separation

$$\text{Loss} = |y_{\text{noise}}|^2 - |y_{\text{signal}}|^2 + \alpha$$

Decorrelation via weight regularisation

$$\text{Loss} = \max\left\{0,\ \beta - \sum |w|^2\right\} + \sum_{i \neq j} \max\{0, (w_i \cdot w_j)\}$$



- Note: different data question (signal injected in position and momentum space) can lead to multiple minima in the loss landscape, i.e. using momentum space and position space.

DFT from modified loss

- Adapt loss function to achieve feature separation, i.e. separating the two classes of data (inspired by triplet loss) [towards generating dual representations dynamically]

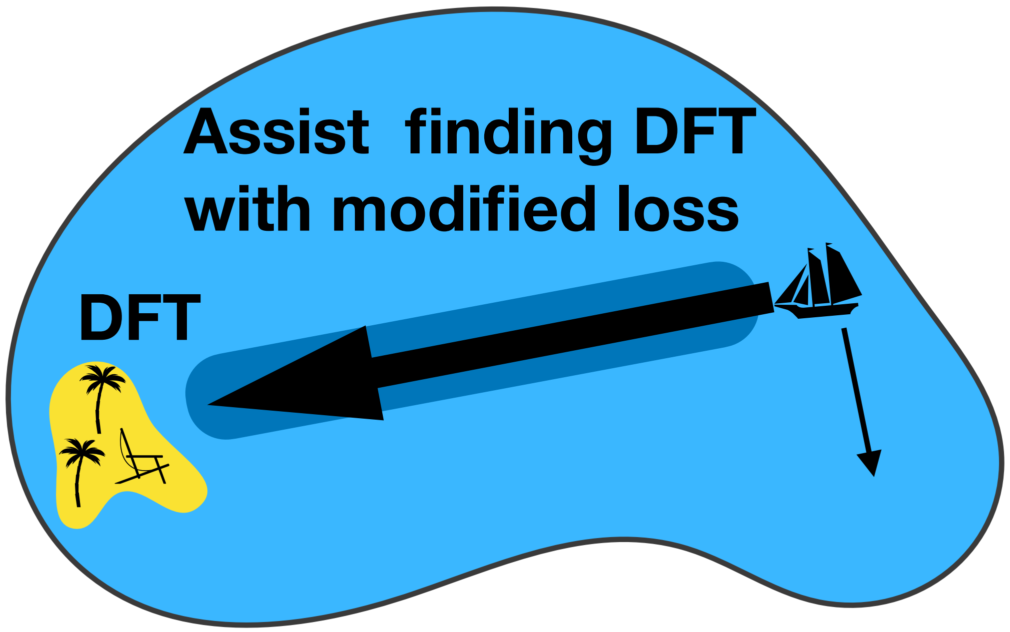
1503.03832

Separation

$$\text{Loss} = |y_{\text{noise}}|^2 - |y_{\text{signal}}|^2 + \alpha$$

Decorrelation via weight regularisation

$$\text{Loss} = \max\left\{0,\, \beta - \sum |w|^2\right\} + \sum_{i \neq j} \max\{0, (w_i \cdot w_j)\}$$



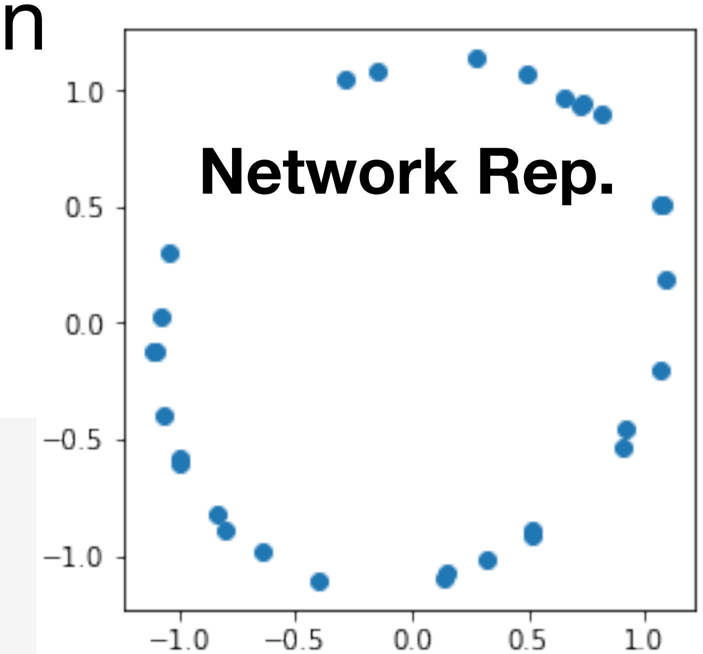
- Note: different data question (signal injected in position and momentum space) can lead to multiple minima in the loss landscape, i.e. using momentum space and position space.

DFT from feature separation

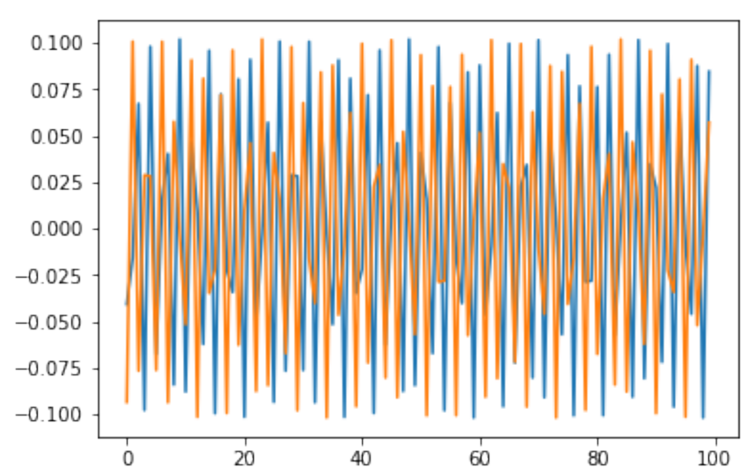
- The final result is a simple network with few parameters utilising `dual' representation. (First layer in deeper network)

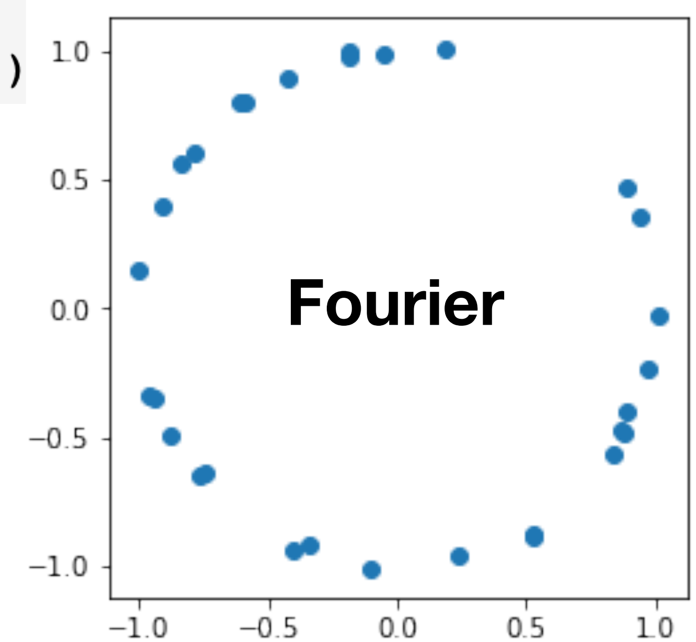
```
ourmodel=Sequential()
ourmodel.add(Dense(2,use_bias='False',activity_regularizer=customReg,
                   input_shape=(int(2*bins),)))
```

```
def customLoss(yTrue,yPred):
    margin = K.constant(1.)
    return K.maximum(K.constant(0),margin-K.sum(K.square(yPred)))

def customReg(weight_matrix):
    margin = K.constant(0.5)
    return K.maximum(K.constant(0.),margin-K.sum(K.square(weight_matrix)))
```



Input sample:



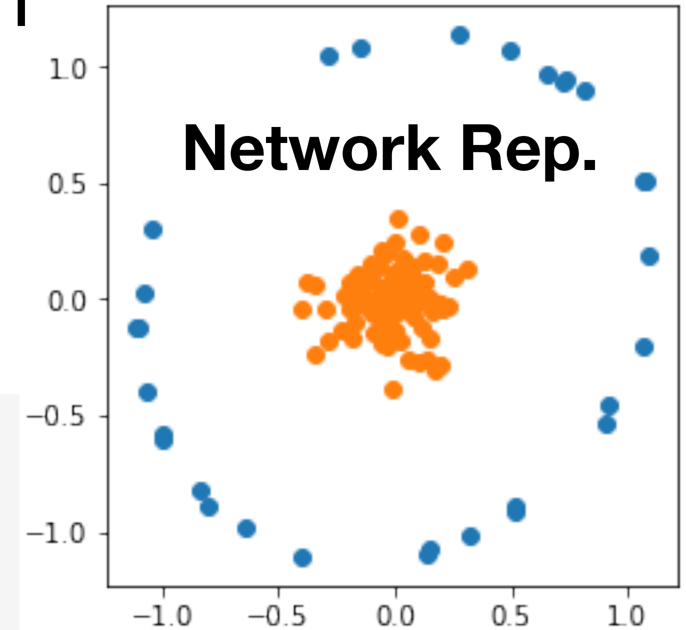


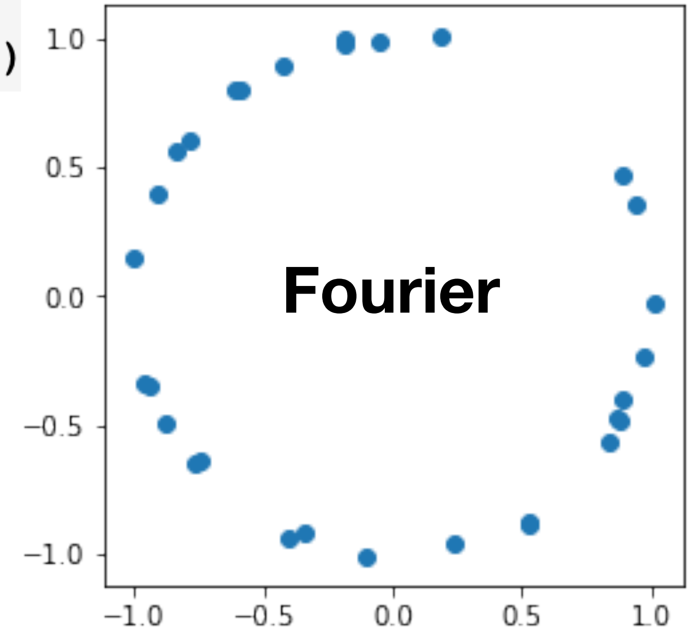
DFT from feature separation

- The final result is a simple network with few parameters utilising `dual' representation. (First layer in deeper network)

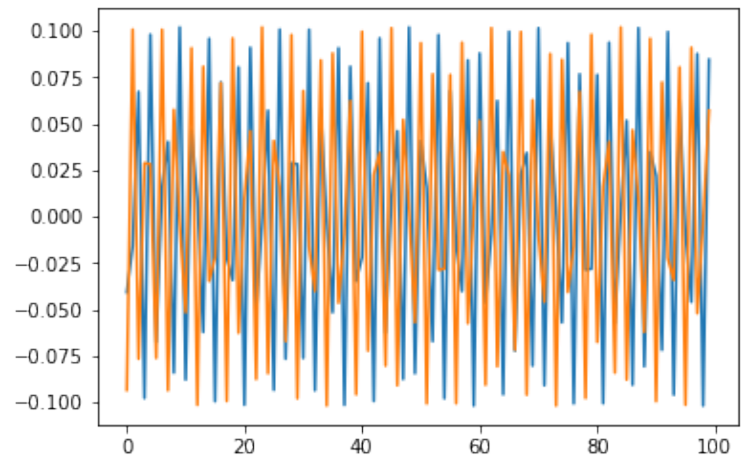
```
ourmodel=Sequential()
ourmodel.add(Dense(2,use_bias='False',activity_regularizer=customReg,
                  input_shape=(int(2*bins),)))
```

```
def customLoss(yTrue,yPred):
    margin = K.constant(1.)
    return K.maximum(K.constant(0),margin-K.sum(K.square(yPred)))

def customReg(weight_matrix):
    margin = K.constant(0.5)
    return K.maximum(K.constant(0.),margin-K.sum(K.square(weight_matrix)))
```





Input sample:



Dualities and Fourier transformation

- Review on dualities and global symmetries (Quevedo: hep-th/9706210), several dualities can be seen as Fourier transformation:

> "Therefore we can see that the dualities we have been dealing with for antisymmetric tensors are only particular cases of Fourier transforms and finding the dual action reduces to finding Fourier transforms. "

- For instance duality between massive antisymmetric tensors.

$$e^{\tilde{F}} \overset{\text{FT}}{\leftrightarrow} e^{F}$$

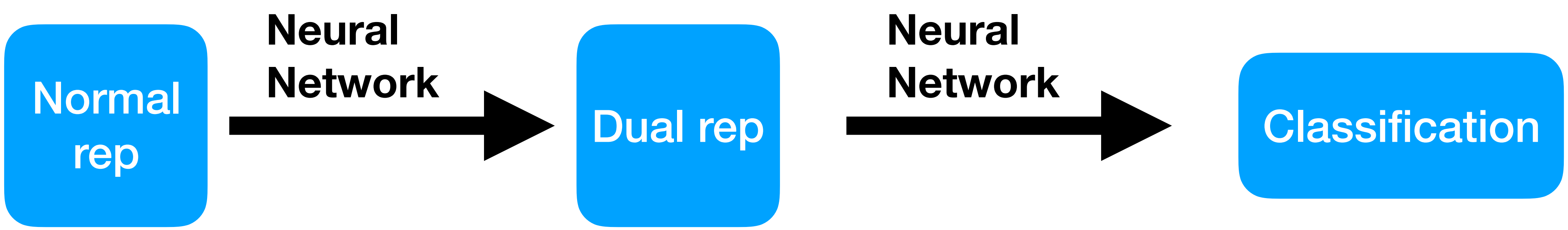
$$\begin{tikzcd}
 & Z = \int \mathscr{D}\tilde{B}_{d-h}\mathscr{D}H_h \; e^{\int d^Dx(H\cdot d\tilde{B}_{d-h}+G(H_h)+\tilde{F}(\tilde{B}_{d-h}))} \arrow[dl, "\tilde{B}_{d-h}"'] \arrow[dr, "H_h"] & \\
S = \int d^Dx \left(F(\partial H_h) + G(H_h)\right) & & \tilde{S} = \int d^Dx \left(\tilde{F}(\tilde{B}_{d-h}) + \tilde{G}(\partial \tilde{B}_{d-h})\right)
\end{tikzcd}$$

Dualities and Fourier transformation

- Fourier transformation ~ Duality transformation

> “Therefore we can see that the dualities we have been dealing with for antisymmetric tensors are only particular cases of Fourier transforms and **finding the dual action reduces to finding Fourier transforms.** “

- Here: network adapts dual representation by demanding feature separation.
- Can we use this?



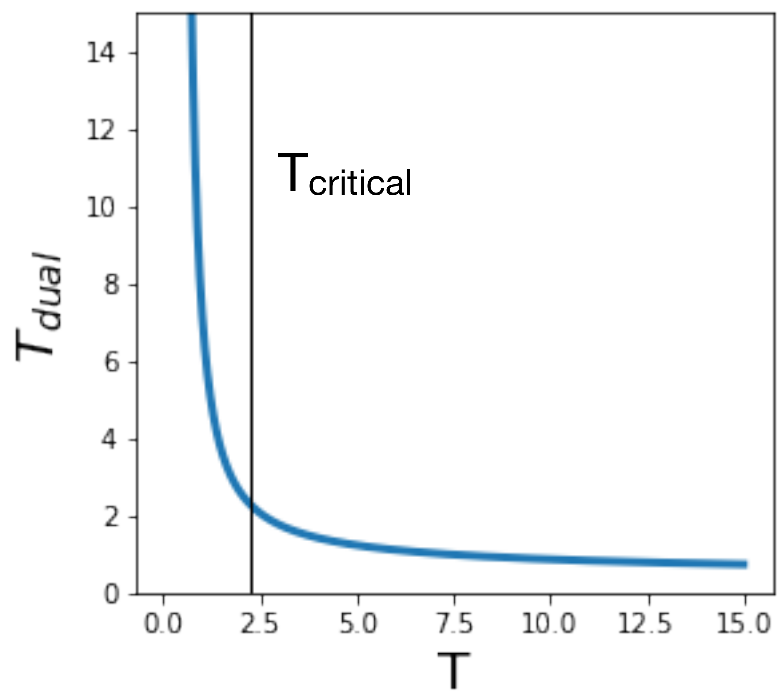
Let’s look for Physics examples

aim: identify what’s special about dual representation

2D Ising Model

Duality in 2D Ising model

- High - low temperature self-duality



Ordered rep. ↔ Disordered rep.

Original

$$H = -J \sum_{\langle i,j \rangle} s_i s_j$$

$$Z = \sum e^{-\beta H(s)}$$

$$\beta = \frac{1}{k_{\mathrm{B}} T}$$

Dual

$$H = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j$$

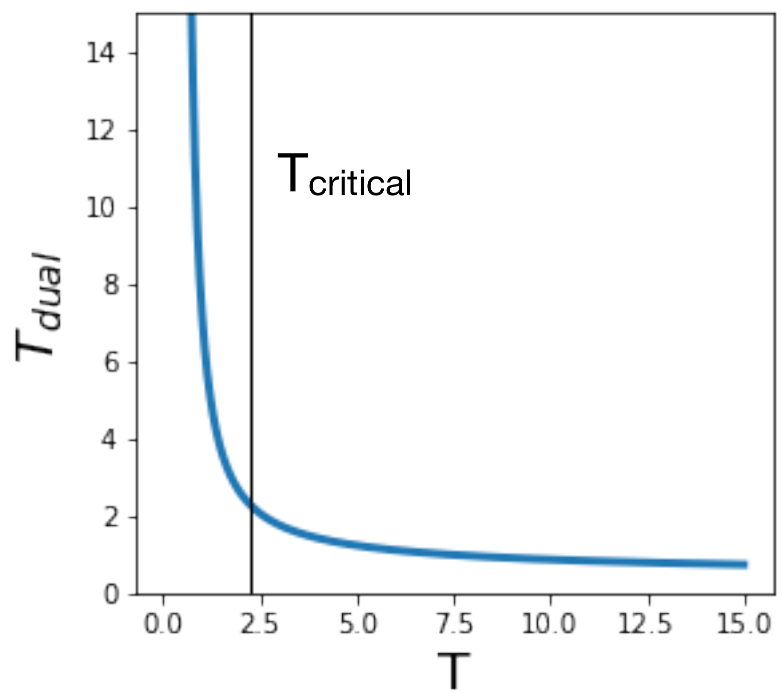
$$Z = \sum e^{-\tilde{\beta} H(\sigma)}$$

$$\tilde{\beta} = -\frac{1}{2} \log \tanh \beta$$

Krammers, Wannier 1941; Onsager 1943; review: Savit 1980

Duality in 2D Ising model

- High - low temperature self-duality



Ordered rep. ↔ Disordered rep.

Position space?

Momentum space?

Original

$$H = -J \sum_{\langle i,j \rangle} s_i s_j$$

$$Z = \sum e^{-\beta H(s)}$$

$$\beta = \frac{1}{k_B T}$$

Dual

$$H = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j$$

$$Z = \sum e^{-\tilde{\beta} H(\sigma)}$$

$$\tilde{\beta} = -\frac{1}{2} \log \tanh \beta$$

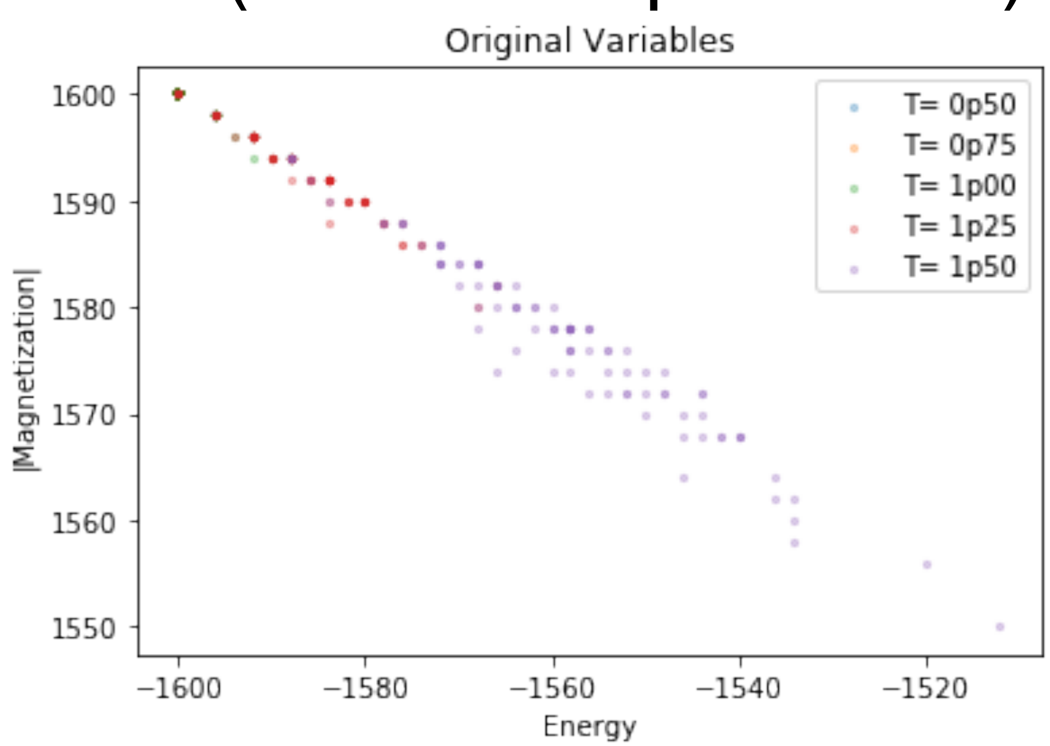
Krammers, Wannier 1941; Onsager 1943; review: Savit 1980

Which data problem?

- Some correlation function which is easier evaluated on dual variables.

$$\langle \sigma_i \sigma_j \rangle, \ \langle E(\sigma) \rangle, \ \langle M(\sigma) \rangle$$

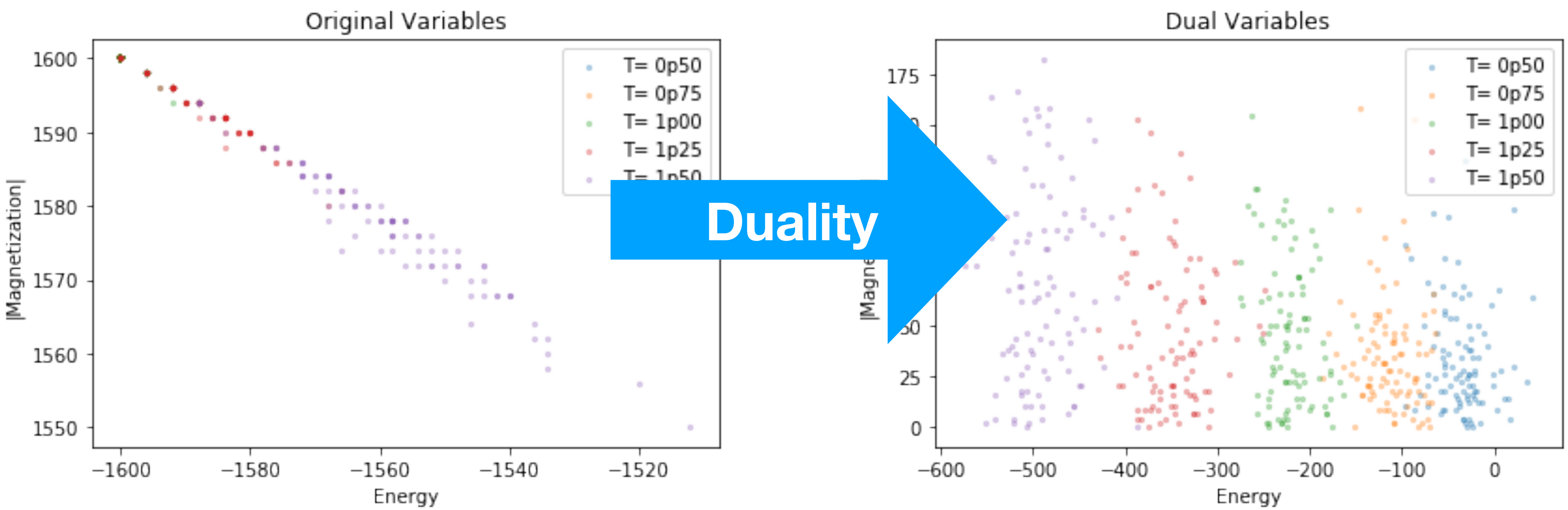
- Can we classify the temperature for low-temperature configurations? Which temperature is a sample drawn from (at low temperatures)?



They look rather similar. How about in the dual rep.?

Data question on Ising

- But at the dual temperatures, our data takes a different shape:

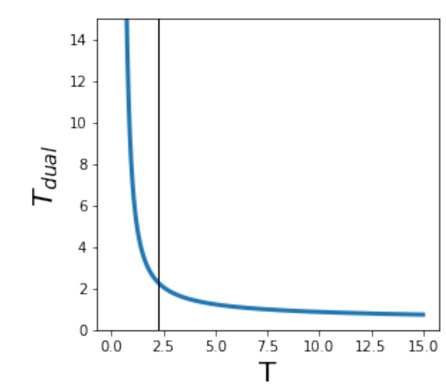


- It is easier to classify temperature of a low-temperature configuration in the dual representation …

- How come?

$$P(\mathbf{s}) = \frac{e^{E/T}}{Z}\,, \qquad P(\sigma) = \frac{e^{\tilde{E}/\tilde{T}}}{Z}$$

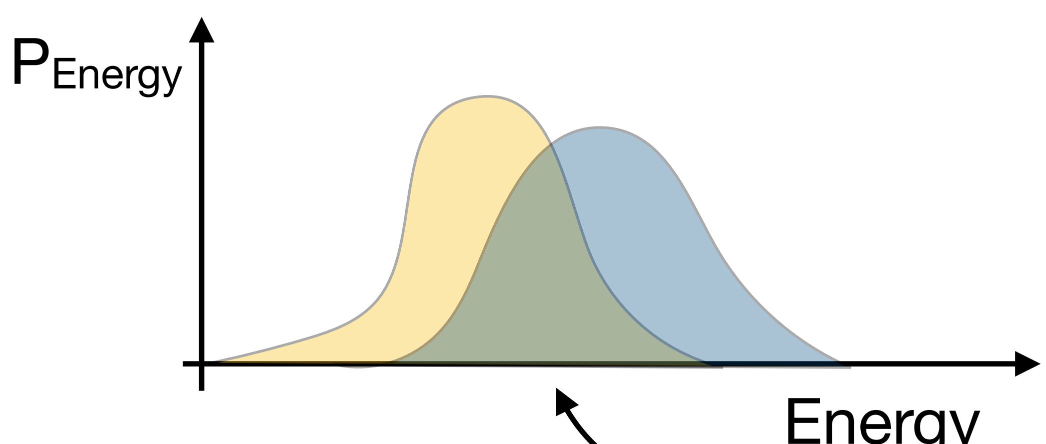
$$\langle \Delta E \rangle \ll \langle \Delta \tilde{E} \rangle \qquad \Delta T \ll \Delta \tilde{T}$$



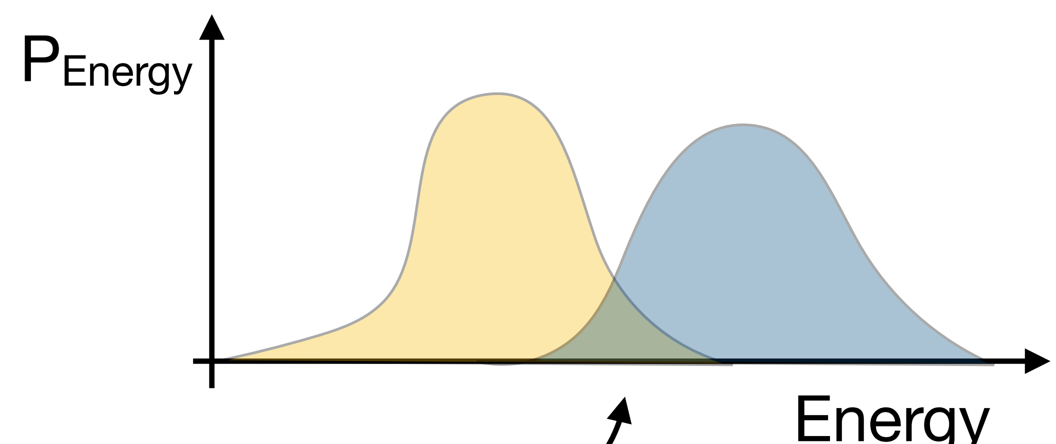
Data question on Ising

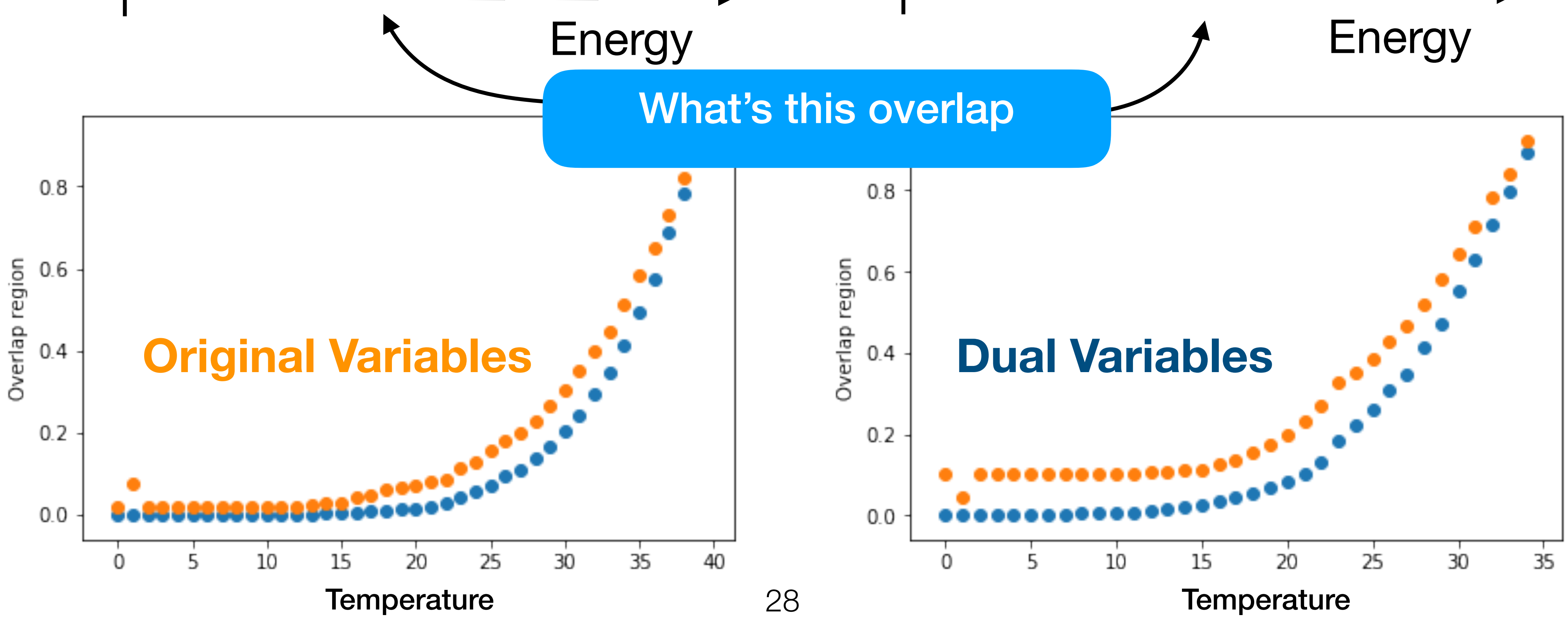
- Let's look at the overlap of energy distributions in finite size samples

Original variables:



Dual variables:





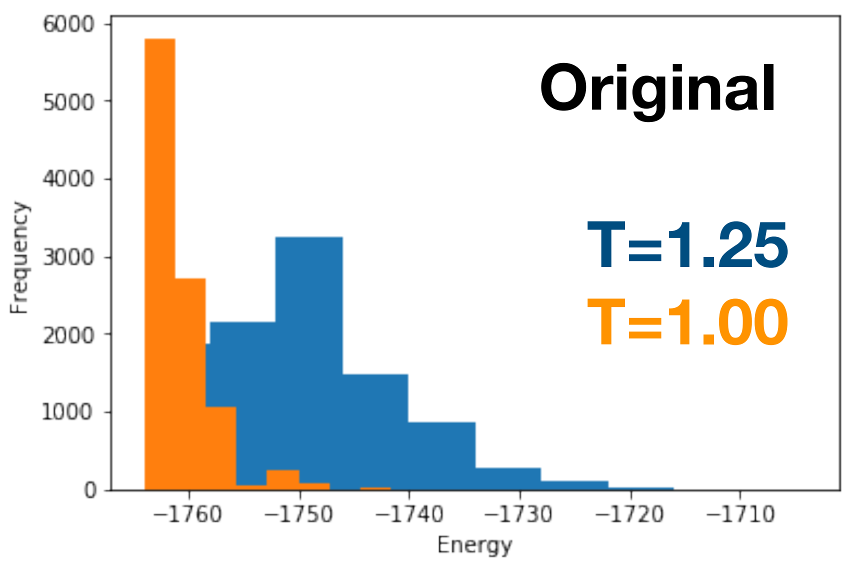
Ising: simple network

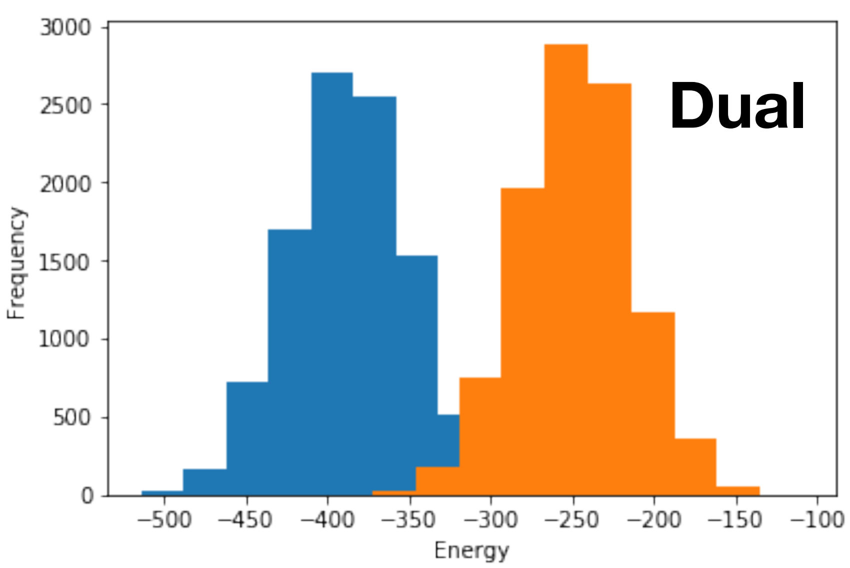
- Let’s confirm this at simple networks:

```
model = Sequential()
model.add(Conv2D(4, kernel_size=3,
                 activation='relu',
                 input_shape=(42,42,1)))
model.add(Flatten())
model.add(Dense(1, activation='sigmoid'))

opt = Nadam(lr=0.002)

model.compile(loss=binary_crossentropy,
              optimizer=opt, metrics=['accuracy'])
```

Original data: <83% val. acc.

Dual data: ~96% val. acc.



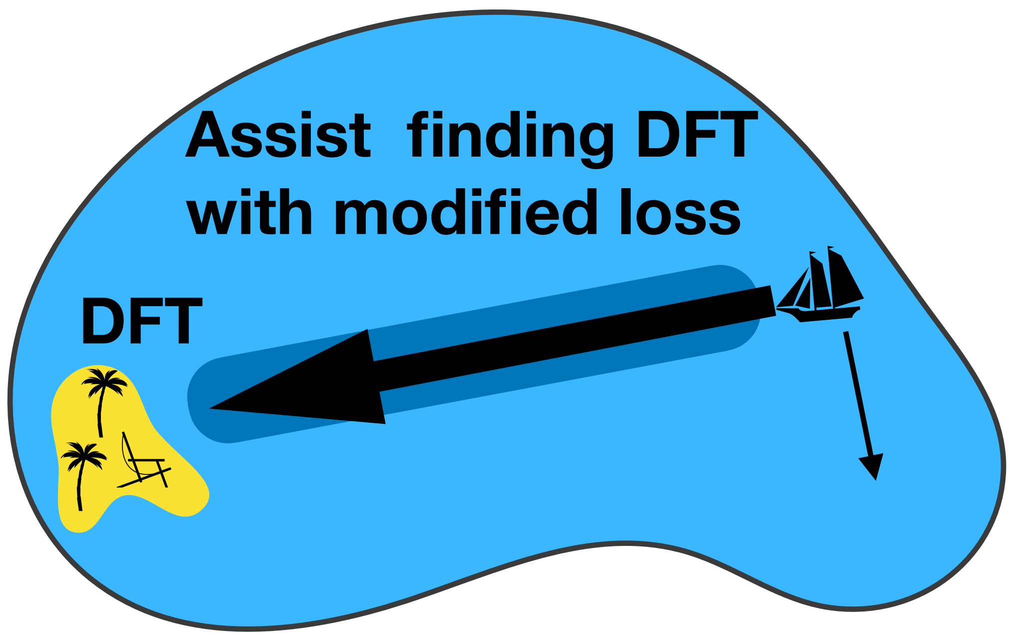


Side remark: way outperforming standard sklearn classifiers

How do we utilise this data representation?

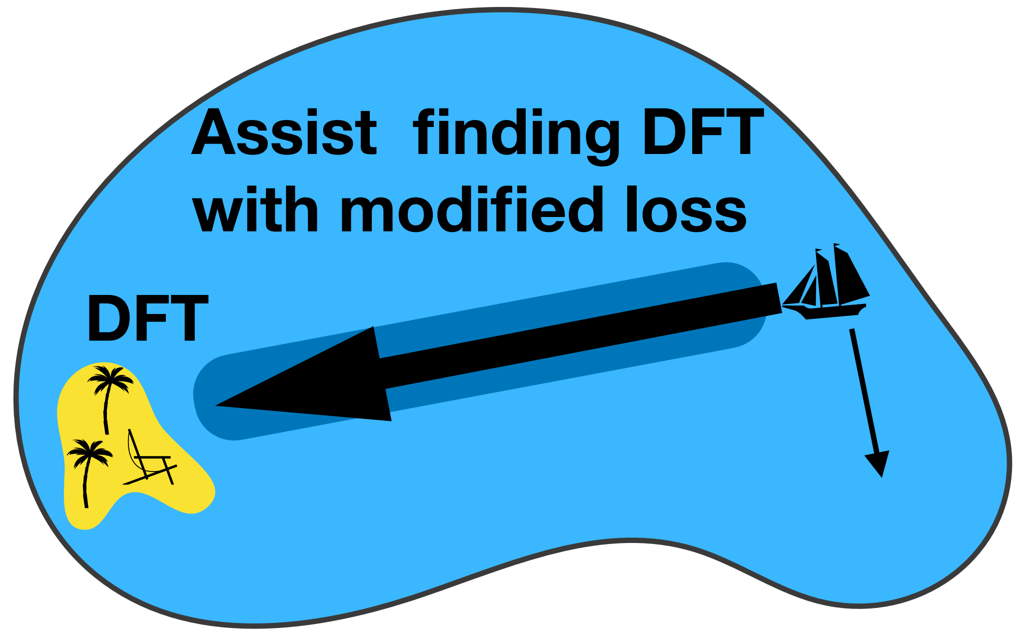
Need duality transformation

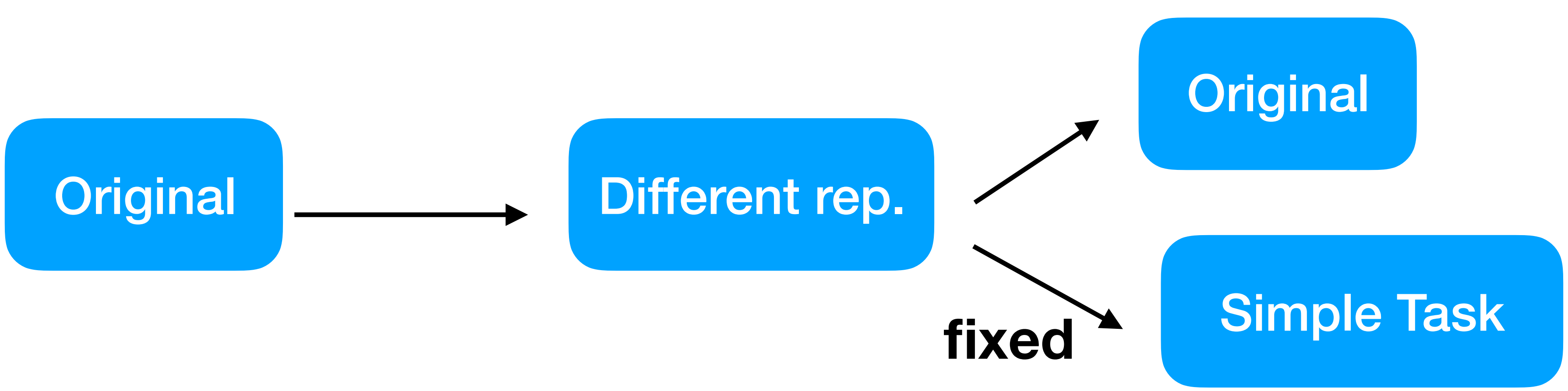
Enforcing dual representation



Enforcing dual representation

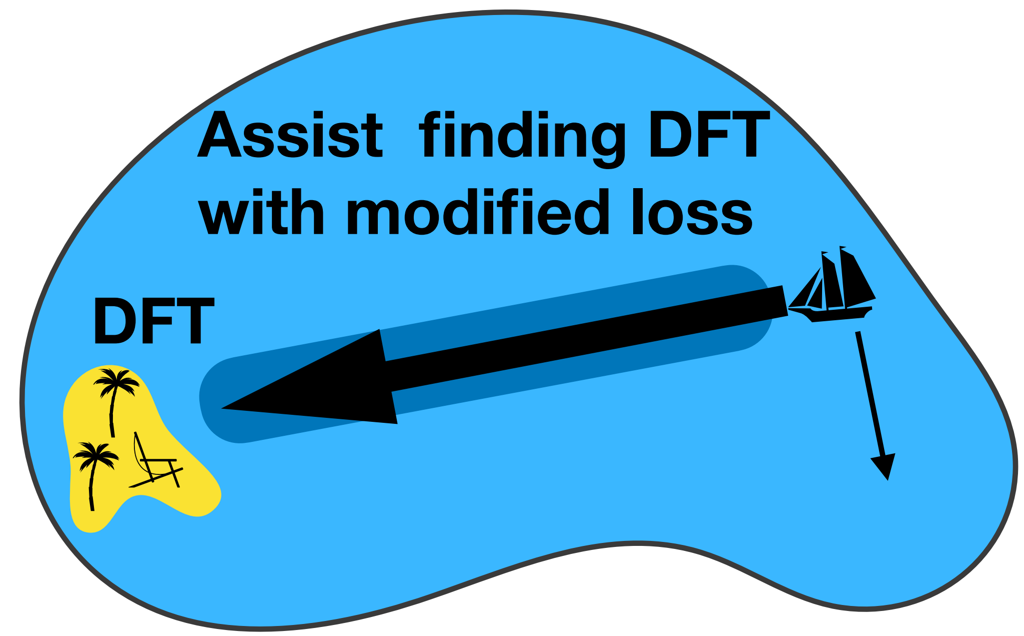
- Observation: deep networks do not find these good representation by themselves
- Which ingredient is missing on the neural network side?
- Work around when task which is accessible in both frames
- Step 1: Find better representation

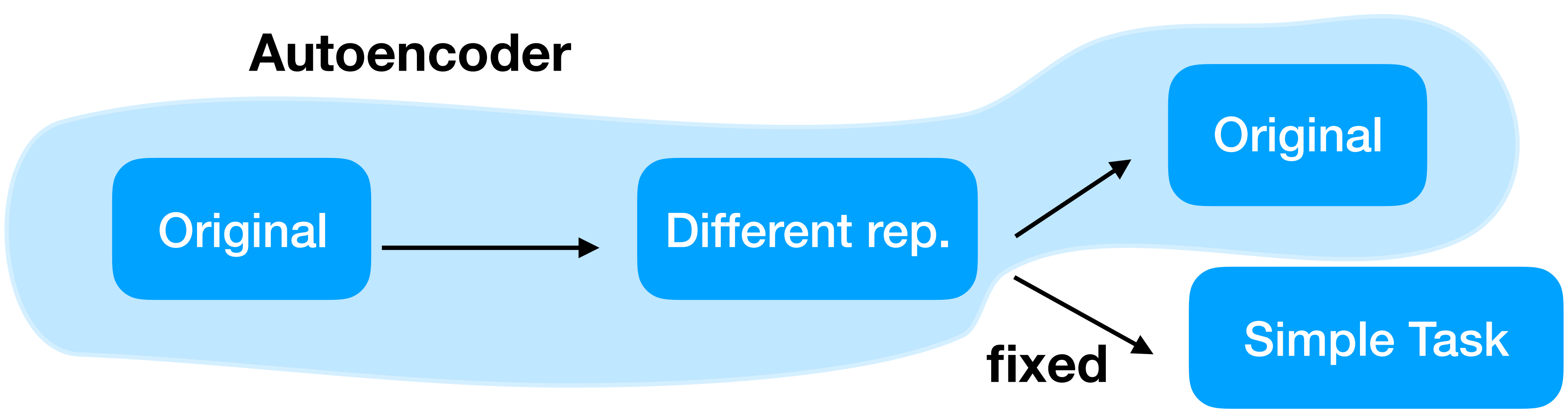




Enforcing dual representation

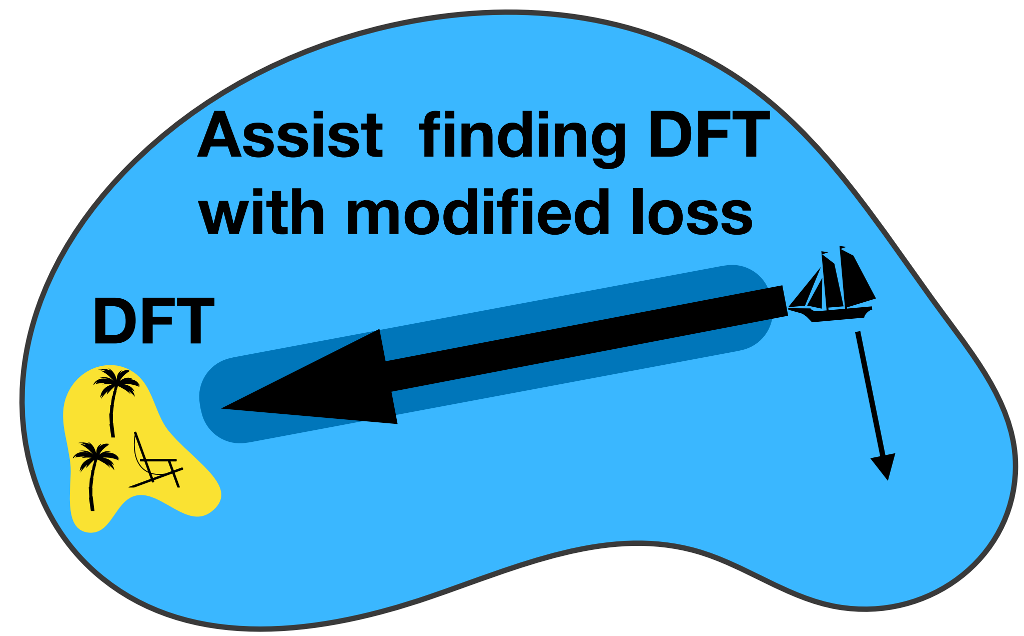
- Observation: deep networks do not find these good representation by themselves
- Which ingredient is missing on the neural network side?
- Work around when task which is accessible in both frames
- Step 1: Find better representation

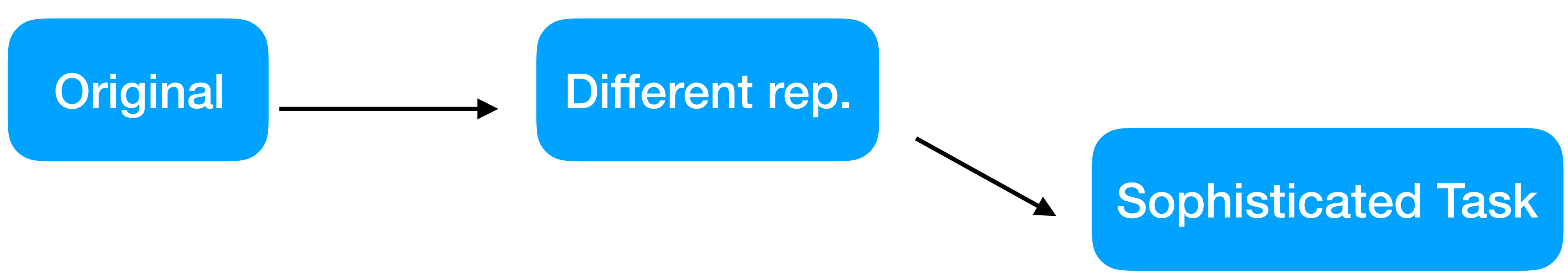




Enforcing dual representation

- Observation: deep networks do not find these good representation by themselves
- Which ingredient is missing on the neural network side?
- Work around when task which is accessible in both frames
- Step 1: Find better representation
- Step 2: Use this representation on previously inaccessible task





Example: 1D Ising model

- Multi-spin interaction model (careful about boundary):

$$H = \sum s_i s_{i+1} \dots s_{i+n} \qquad H_{\text{dual}} = \sum \sigma_i$$

- Simple task: Energy classification (linear in dual representation)
- Hard task: Identify metastable configurations (multiple calculations of σ_i necessary)

-	-	-	-	-	-	-	-	-	-	+	+	-	-	+	-	+	-

- Strategy from previous page shows improvement on performance:
  Normal frame (dense 1024): 92%
  Dual frame (dense 64): 99%
  Autoencoder [64, 18] representation (dense 64): 98%
- Learned representation performs well on dual weights but is not dual representation (*we found a different island*).

2D Ising: finding duality transformation

- In analogy to DFT, we want to go to dual representation.

$$\text{Input: s-configuration} \xrightarrow{\text{NN}} \text{dual: } \sigma\text{-configuration}$$

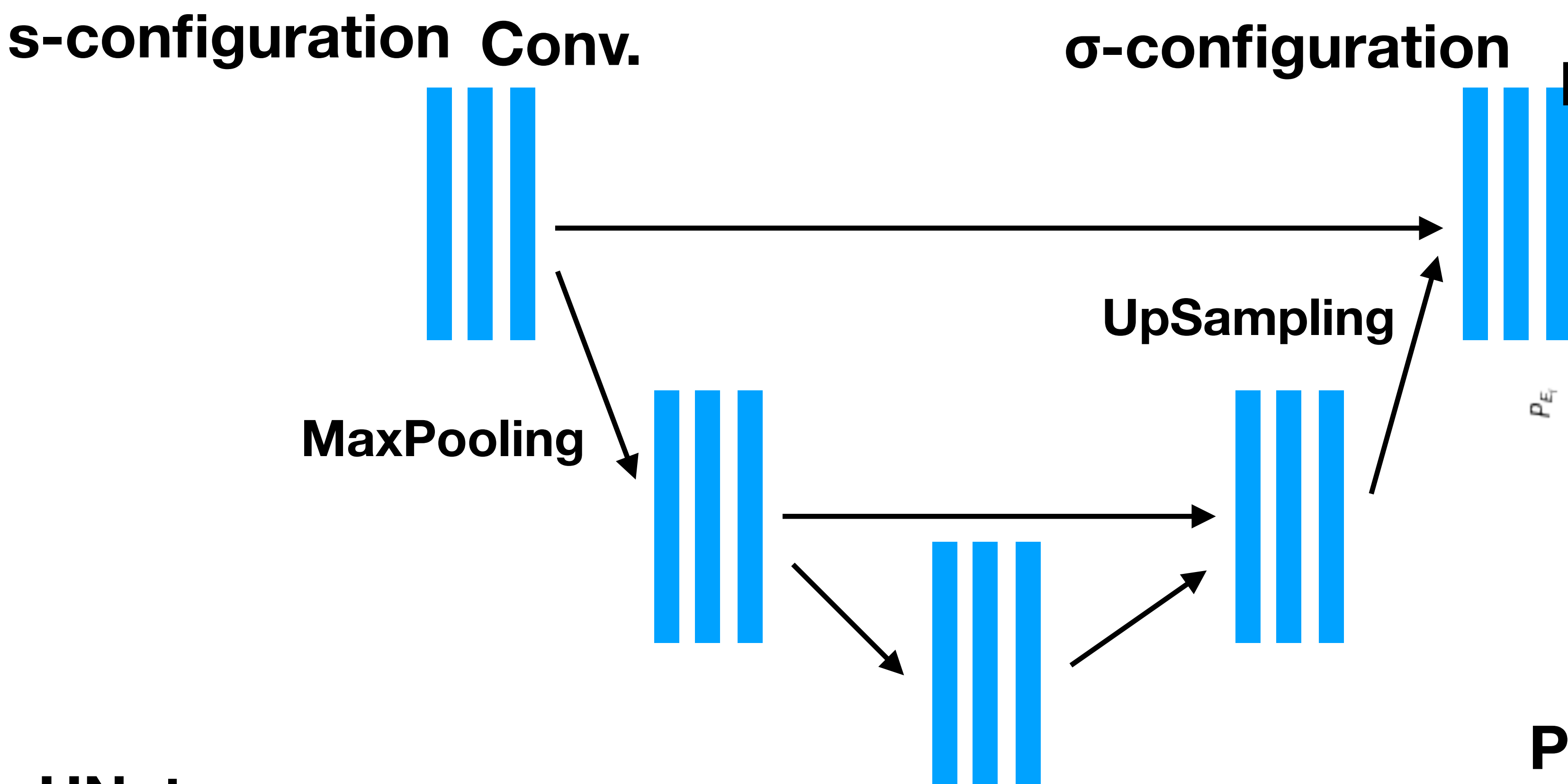
- Which loss (different samples but no 1-1 map from simple MC simulations)?
- Here: KL-divergence (or categorical-crossentropy) between NN distribution and dual temperature distribution; based on p_{Ei} (sparse samples)

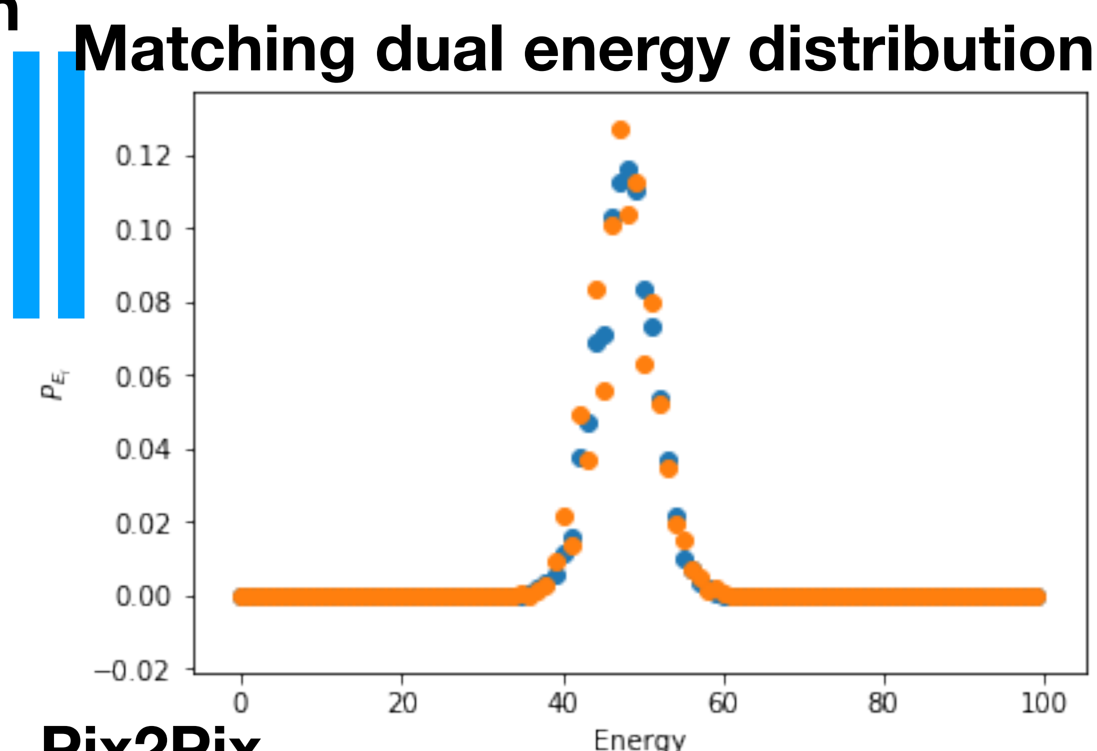
$$D_{KL}(P(E(f(s_i,\beta)))\,||\,P(E(\sigma_i,\beta))) = \sum_j P(E_j(f(s_i,\beta))\log\left(\frac{P(E_j(f(s_i,\beta)))}{P(E_j(\sigma_i,\tilde{\beta}))}\right)$$

a “sophisticated” anchor point

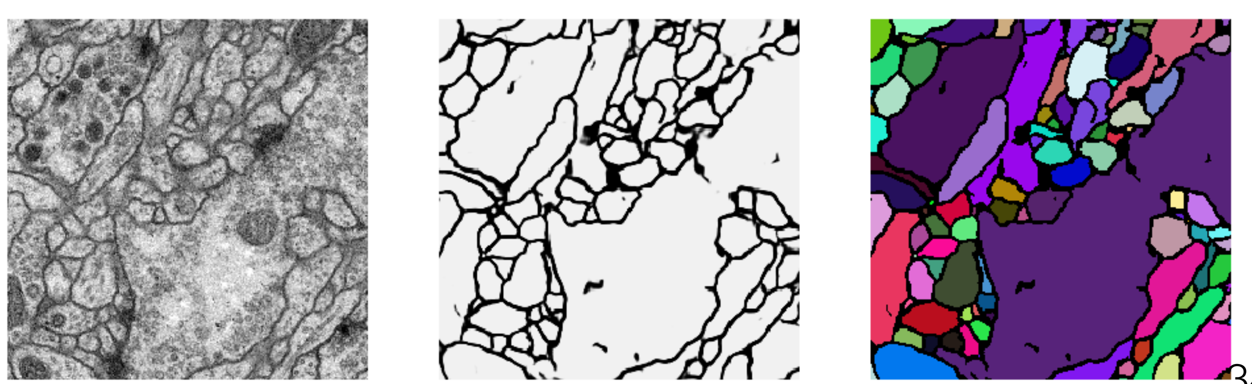
Feature separation in Ising

- Which neural network architecture? Several architectures, so far most promising: U-Net (1505.04597)

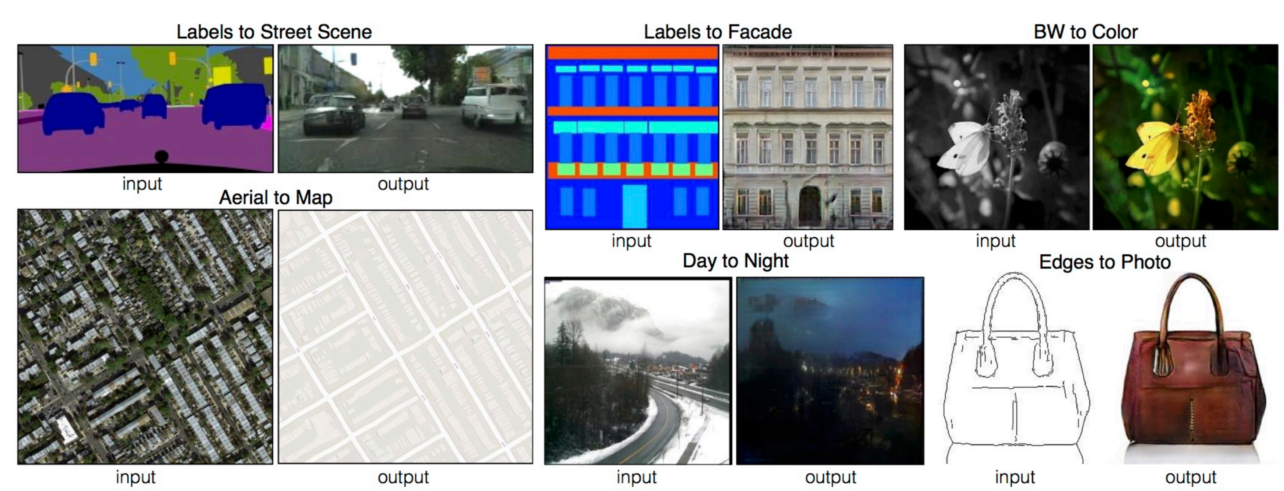




UNet



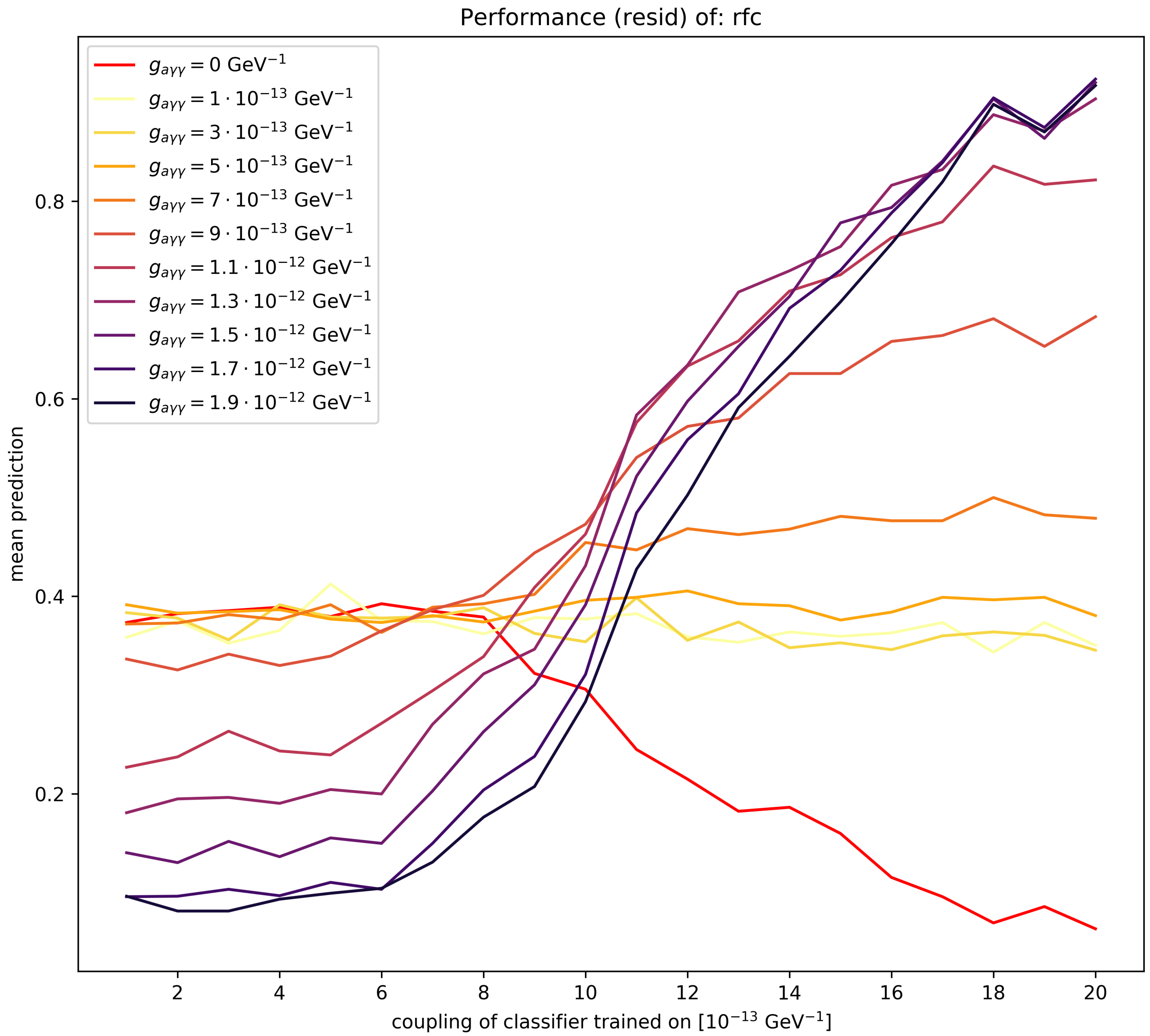
Pix2Pix



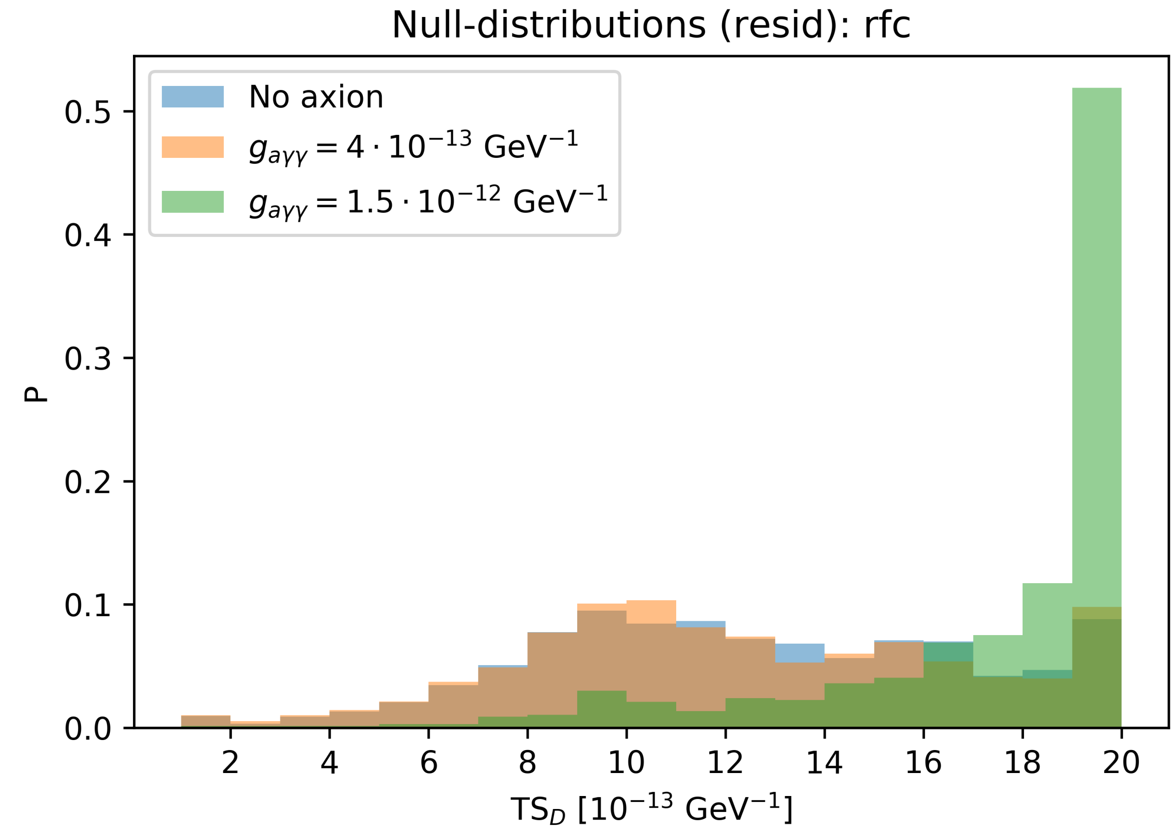
Conclusions

- Dualities tightly connected to data representations.
- Finding good data representations is at the heart of ML and often `good' representations are not found.
- `Good’ representations can be enforced when known to exist.
- This enables classification for previously `inaccessible’ tasks. (1D Ising metastable configurations)
- Interesting representations can be found from feature separation (e.g. DFT)
- Interesting aspect: reduced complexity of networks
- Dualities in Physics motivate multiple minima in a different landscape, those of the cost functions of neural networks.
- Lots of different kinds of dualities in Physics as playground for efficient neural network architectures.

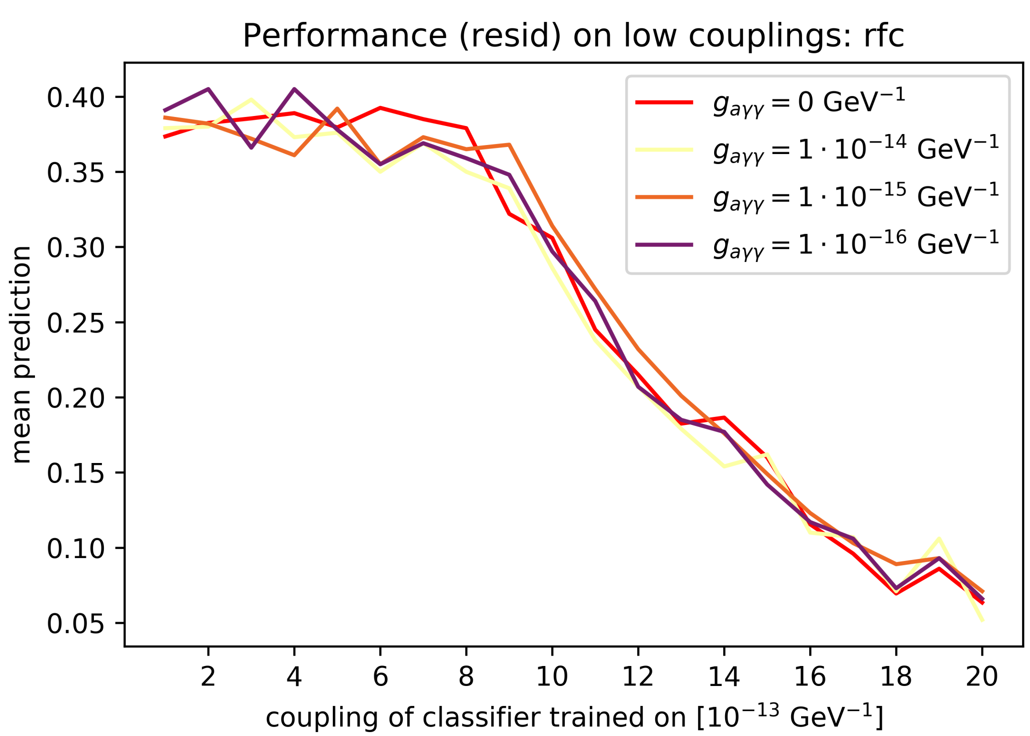
Thank you!



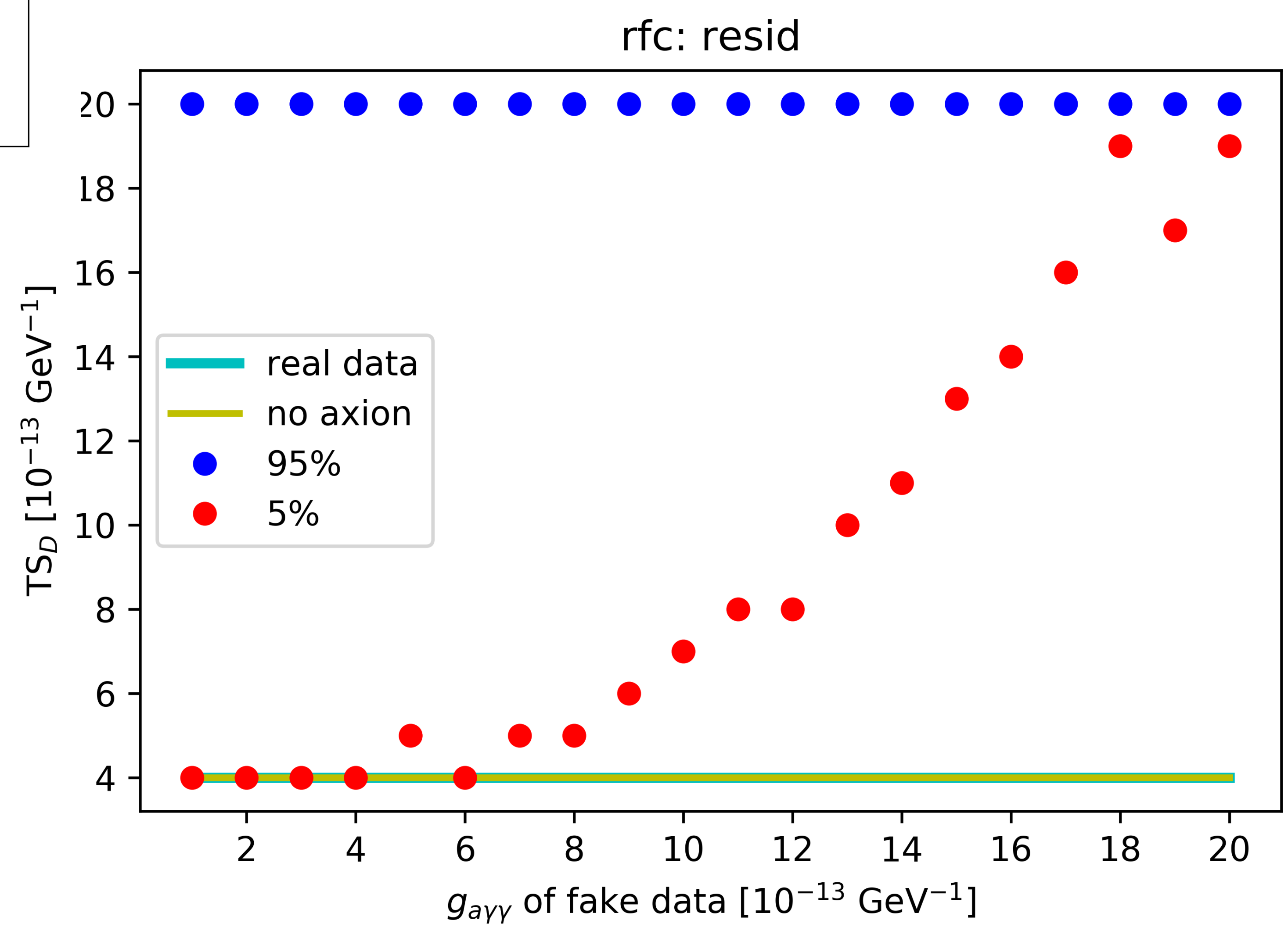
Performance (resid) of: rfc
$g_{a\gamma\gamma} = 0$ GeV^{-1}
$g_{a\gamma\gamma} = 1 \cdot 10^{-13}$ GeV^{-1}
$g_{a\gamma\gamma} = 3 \cdot 10^{-13}$ GeV^{-1}
$g_{a\gamma\gamma} = 5 \cdot 10^{-13}$ GeV^{-1}
$g_{a\gamma\gamma} = 7 \cdot 10^{-13}$ GeV^{-1}
$g_{a\gamma\gamma} = 9 \cdot 10^{-13}$ GeV^{-1}
$g_{a\gamma\gamma} = 1.1 \cdot 10^{-12}$ GeV^{-1}
$g_{a\gamma\gamma} = 1.3 \cdot 10^{-12}$ GeV^{-1}
$g_{a\gamma\gamma} = 1.5 \cdot 10^{-12}$ GeV^{-1}
$g_{a\gamma\gamma} = 1.7 \cdot 10^{-12}$ GeV^{-1}
$g_{a\gamma\gamma} = 1.9 \cdot 10^{-12}$ GeV^{-1}
mean prediction
coupling of classifier trained on [10^{-13} GeV^{-1}]



Null-distributions (resid): rfc
No axion
$g_{a\gamma\gamma} = 4 \cdot 10^{-13}$ GeV^{-1}
$g_{a\gamma\gamma} = 1.5 \cdot 10^{-12}$ GeV^{-1}
P
TS$_D$ [10^{-13} GeV^{-1}]



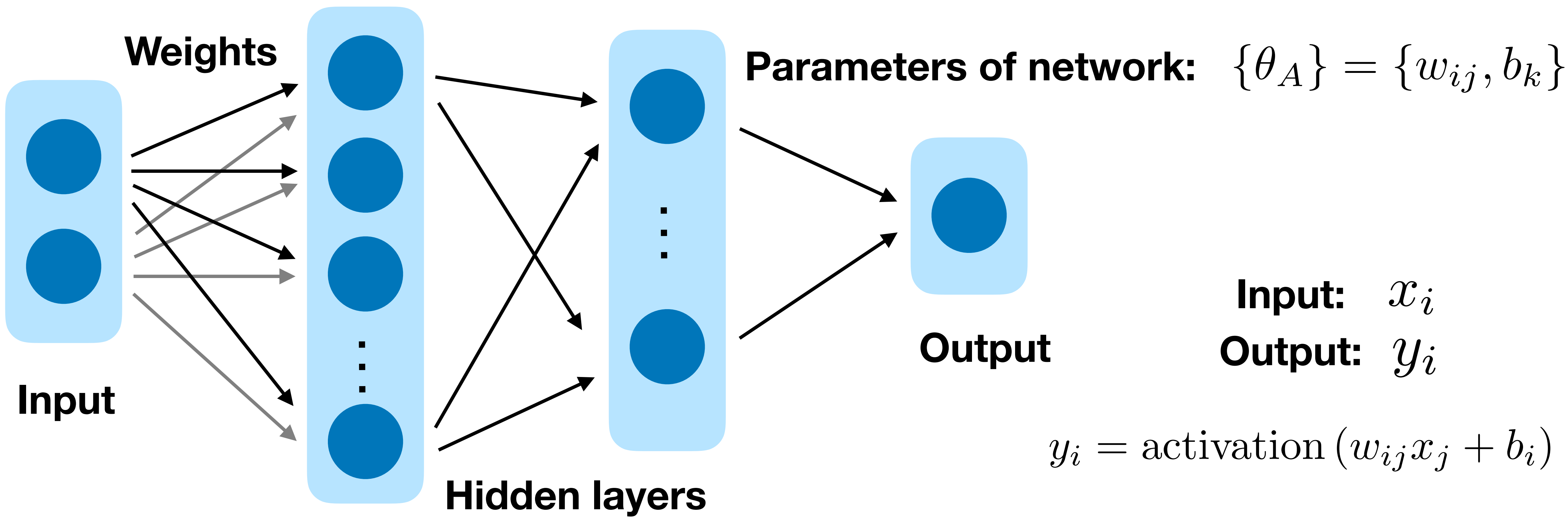
Performance (resid) on low couplings: rfc
$g_{a\gamma\gamma} = 0$ GeV^{-1}
$g_{a\gamma\gamma} = 1 \cdot 10^{-14}$ GeV^{-1}
$g_{a\gamma\gamma} = 1 \cdot 10^{-15}$ GeV^{-1}
$g_{a\gamma\gamma} = 1 \cdot 10^{-16}$ GeV^{-1}
mean prediction
coupling of classifier trained on [10^{-13} GeV^{-1}]



rfc: resid
real data
no axion
95%
5%
TS$_D$ [10^{-13} GeV^{-1}]
$g_{a\gamma\gamma}$ of fake data [10^{-13} GeV^{-1}]

Neural Networks

- Layout of neural nets:



- Cost function depending on parameters of network

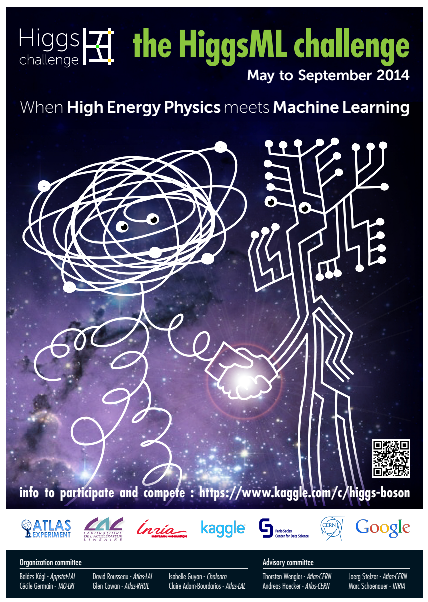
$$\text{cost}(\theta_A) = \sum_{\text{training set}} |y_{\text{desired}} - y_{\text{predicted}}(\theta_A)|$$

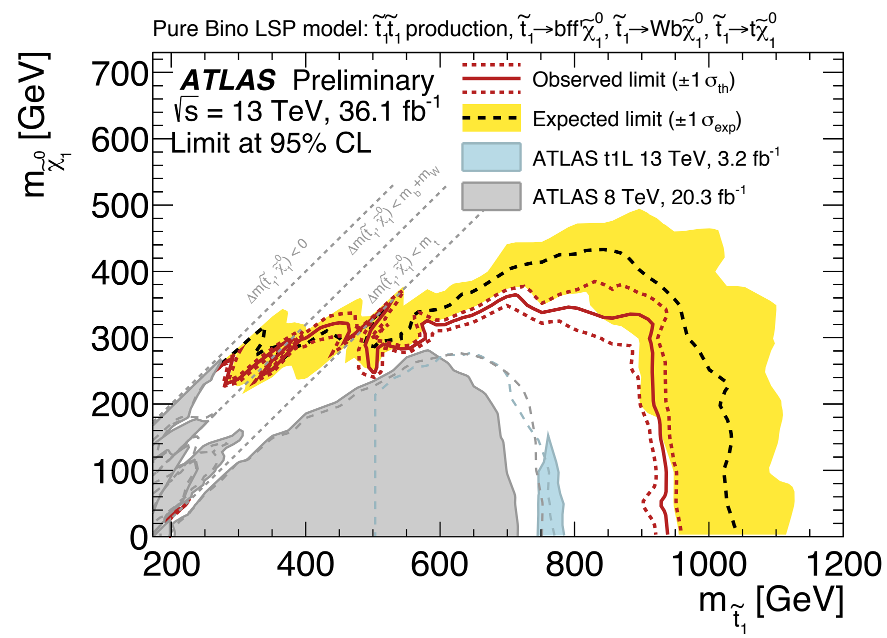
- Parameters of network updated using gradient descent

$$\theta_A \to \theta_A - \eta \, \nabla_A \text{cost}(\theta_A)$$

Improving physics searches with ML

- ML is heavily used in particle physics (since decades) and ML techniques are used to improve bounds on our favourite BSM models





Atlas: 1711.11520

- Similarly in astrophysics (Hoyle: 1504.07255) for distance measurements of galaxies

➡Bottom line: feed the entire data, rather than preprocessed human-designed features

- Let’s see how this works with limits on light axion-like particles

Constraining ALPs

NGC1275: 1605.01043
Other sources: 1704.05256
Athena bounds: 1707.00176
with: Conlon, Day, Jennings, Rummel; Berg, Muia, Powell

- Photon-axion interconversion in background magnetic fields:

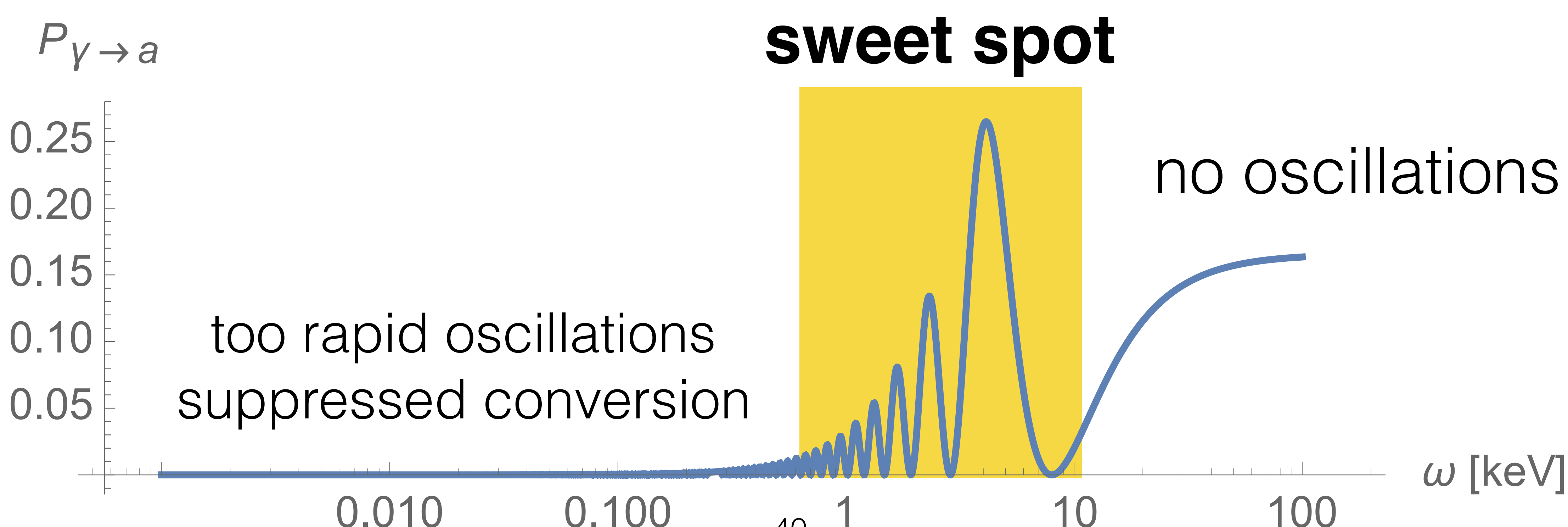
$$\mathscr{L} \supset -\frac{g_{a\gamma\gamma}}{4} a F \tilde{F} = g_{a\gamma\gamma} \mathbf{E} \cdot \mathbf{B}$$

- One interesting parameter region can be obtained for photons from sources in and behind galaxy cluster magnetic fields.

$$\Theta = 0.28 \left(\frac{B_\perp}{1\ \mu\mathrm{G}}\right)\left(\frac{\omega}{1\ \mathrm{keV}}\right)\left(\frac{10^{-3}\mathrm{cm}^{-3}}{n_e}\right)\left(\frac{10^{11}\mathrm{GeV}}{M}\right)$$

$$P_{\gamma\to a} = \frac{1}{2}\frac{\Theta^2}{1+\Theta^2}\sin^2\left(\Delta\sqrt{1+\Theta^2}\right)$$

$$\Delta = 0.54 \left(\frac{n_e}{10^{-3}\mathrm{cm}^{-3}}\right)\left(\frac{L}{10\mathrm{kpc}}\right)\left(\frac{1\mathrm{keV}}{\omega}\right)$$



Constraining ALPs

NGC1275: 1605.01043
Other sources: 1704.05256
Athena bounds: 1707.00176
with: Conlon, Day, Jennings, Rummel; Berg, Muia, Powell

- Photon-axion interconversion in background magnetic fields:

$$\mathscr{L} \supset -\frac{g_{a\gamma\gamma}}{4} a F\tilde{F} = g_{a\gamma\gamma}\mathbf{E}\cdot\mathbf{B}$$

- One interesting parameter region can be obtained for photons from sources in and behind galaxy cluster magnetic fields.

$$\Theta = 0.28\left(\frac{B_\perp}{1\ \mu\mathrm{G}}\right)\left(\frac{\omega}{1\ \mathrm{keV}}\right)\left(\frac{10^{-3}\mathrm{cm}^{-3}}{n_e}\right)\left(\frac{10^{11}\mathrm{GeV}}{M}\right)$$

$$P_{\gamma\to a} = \frac{1}{2}\frac{\Theta^2}{1+\Theta^2}\sin^2\left(\Delta\sqrt{1+\Theta^2}\right)$$

$$\Delta = 0.54\left(\frac{n_e}{10^{-3}\mathrm{cm}^{-3}}\right)\left(\frac{L}{10\mathrm{kpc}}\right)\left(\frac{1\mathrm{keV}}{\omega}\right)$$

Constraining ALPs

NGC1275: 1605.01043
Other sources: 1704.05256
Athena bounds: 1707.00176
with: Conlon, Day, Jennings, Rummel; Berg, Muia, Powell

- Photon-axion interconversion in background magnetic fields:

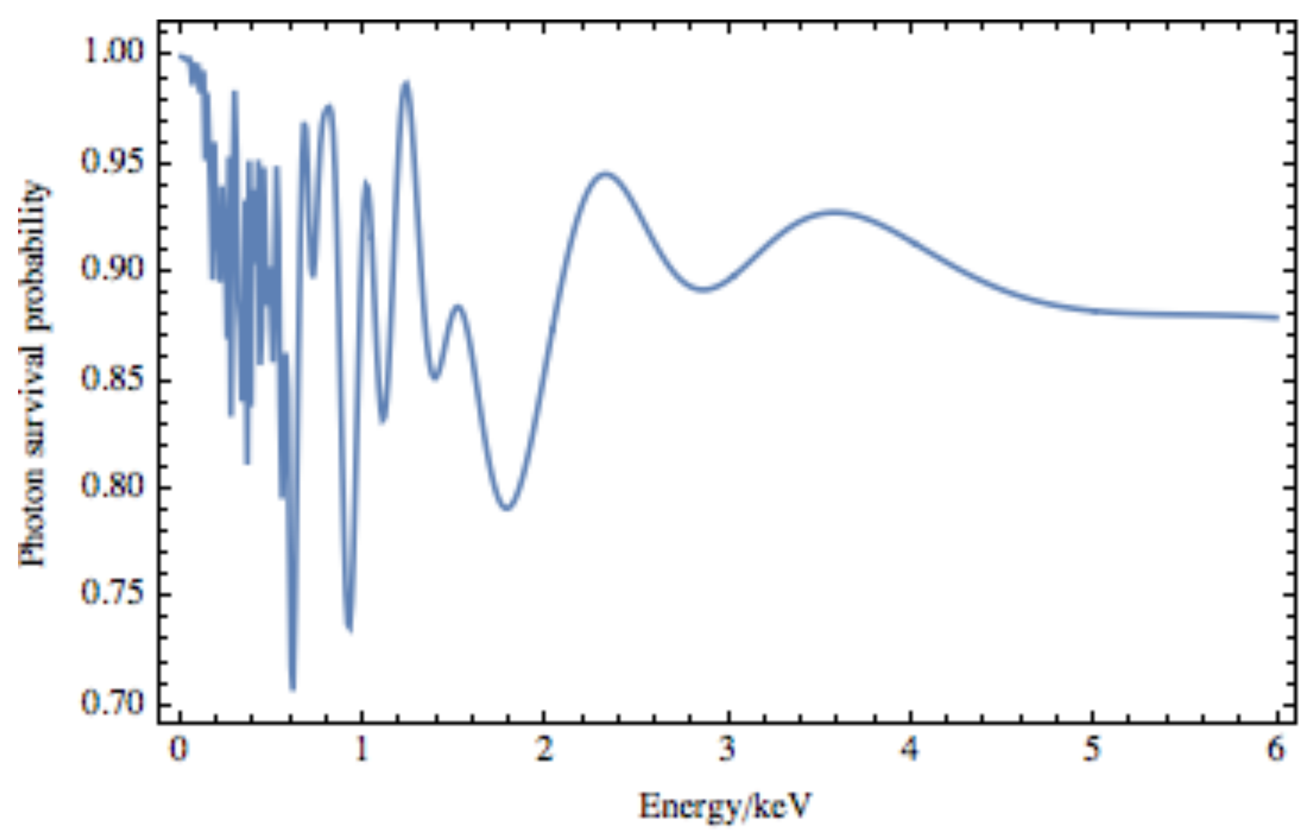
$$\mathscr{L} \supset -\frac{g_{a\gamma\gamma}}{4} a F \tilde{F} = g_{a\gamma\gamma} \mathbf{E} \cdot \mathbf{B}$$

- One interesting parameter region can be obtained for photons from sources in and behind galaxy cluster magnetic fields.

$$\Theta = 0.28 \left(\frac{B_\perp}{1\ \mu\text{G}}\right) \left(\frac{\omega}{1\ \text{keV}}\right) \left(\frac{10^{-3}\text{cm}^{-3}}{n_e}\right) \left(\frac{10^{11}\text{GeV}}{M}\right)$$

$$P_{\gamma \to a} = \frac{1}{2} \frac{\Theta^2}{1+\Theta^2} \sin^2 \left(\Delta \sqrt{1+\Theta^2}\right)$$

$$\Delta = 0.54 \left(\frac{n_e}{10^{-3}\text{cm}^{-3}}\right) \left(\frac{L}{10\text{kpc}}\right) \left(\frac{1\text{keV}}{\omega}\right)$$



Constraining ALPs

NGC1275: 1605.01043
Other sources: 1704.05256
Athena bounds: 1707.00176
with: Conlon, Day, Jennings, Rummel; Berg, Muia, Powell

- Photon-axion interconversion in background magnetic fields:

$$\mathscr{L} \supset -\frac{g_{a\gamma\gamma}}{4} a F\tilde{F} = g_{a\gamma\gamma}\mathbf{E}\cdot\mathbf{B}$$

- One interesting parameter region can be obtained for photons from sources in and behind galaxy cluster magnetic fields.

$$\Theta = 0.28 \left(\frac{B_\perp}{1\ \mu\mathrm{G}}\right)\left(\frac{\omega}{1\ \mathrm{keV}}\right)\left(\frac{10^{-3}\mathrm{cm}^{-3}}{n_e}\right)\left(\frac{10^{11}\mathrm{GeV}}{M}\right)$$

$$P_{\gamma\to a} = \frac{1}{2}\frac{\Theta^2}{1+\Theta^2}\sin^2\left(\Delta\sqrt{1+\Theta^2}\right)$$

$$\Delta = 0.54\left(\frac{n_e}{10^{-3}\mathrm{cm}^{-3}}\right)\left(\frac{L}{10\mathrm{kpc}}\right)\left(\frac{1\mathrm{keV}}{\omega}\right)$$

Constraining ALPs

NGC1275: 1605.01043
Other sources: 1704.05256
Athena bounds: 1707.00176
with: Conlon, Day, Jennings, Rummel; Berg, Muia, Powell

- Photon-axion interconversion in background magnetic fields:

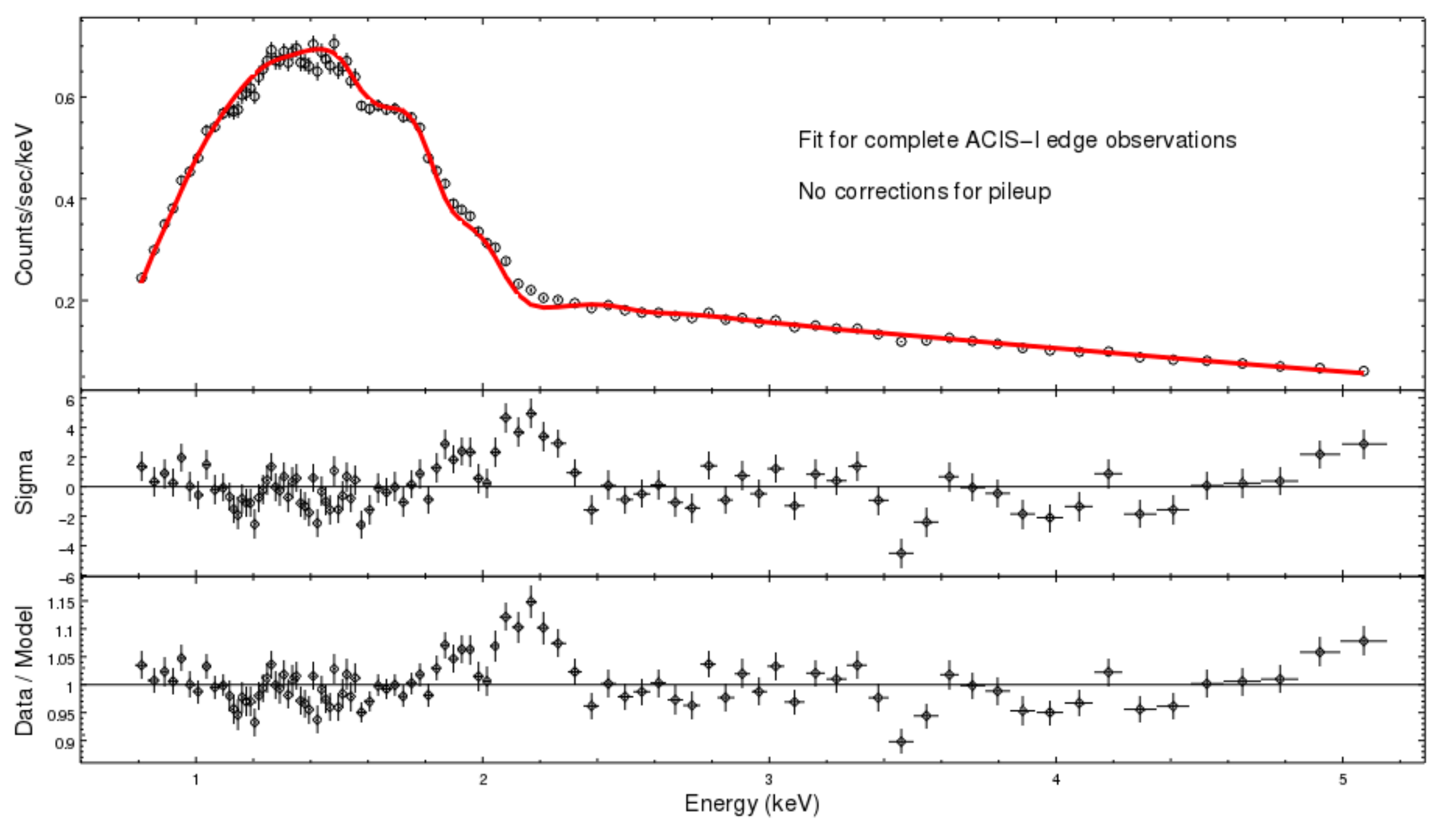
$$\mathscr{L} \supset -\frac{g_{a\gamma\gamma}}{4} a F\tilde{F} = g_{a\gamma\gamma}\mathbf{E}\cdot\mathbf{B}$$

- One interesting parameter region can be obtained for photons from sources in and behind galaxy cluster magnetic fields.

$$\Theta = 0.28\left(\frac{B_\perp}{1\ \mu\mathrm{G}}\right)\left(\frac{\omega}{1\ \mathrm{keV}}\right)\left(\frac{10^{-3}\mathrm{cm}^{-3}}{n_e}\right)\left(\frac{10^{11}\mathrm{GeV}}{M}\right)$$

$$\Delta = 0.54\left(\frac{n_e}{10^{-3}\mathrm{cm}^{-3}}\right)\left(\frac{L}{10\mathrm{kpc}}\right)\left(\frac{1\mathrm{keV}}{\omega}\right)$$

$$P_{\gamma\to a} = \frac{1}{2}\frac{\Theta^2}{1+\Theta^2}\sin^2\left(\Delta\sqrt{1+\Theta^2}\right)$$



Standard bounds

NGC1275: 1605.01043
Other sources: 1704.05256
Athena bounds: 1707.00176
with: Conlon, Day, Jennings;
Berg, Muia, Powell, Rummel

- Spectral distortions (but unknown where & strength as dependent on magnetic field)
- Poisson noise vs. signal
- Essentially: fluctuations larger than noise
- Human-made: Fourier bounds no real improvement

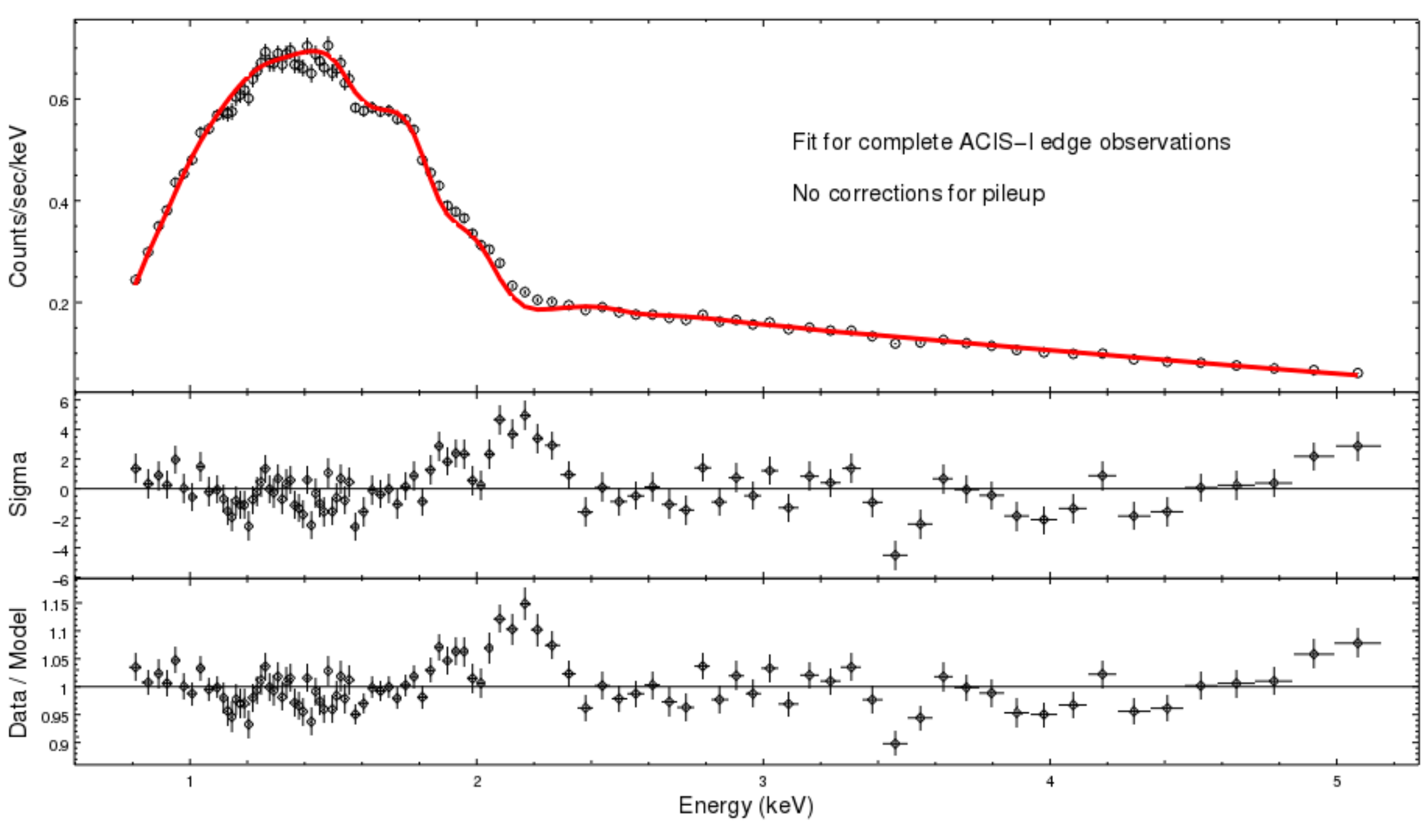
Conlon, Rummel 1808.05916

Standard bounds

NGC1275: 1605.01043
Other sources: 1704.05256
Athena bounds: 1707.00176
with: Conlon, Day, Jennings;
Berg, Muia, Powell, Rummel

- Spectral distortions (but unknown where & strength as dependent on magnetic field)
- Poisson noise vs. signal
- Essentially: fluctuations larger than noise
- Human-made: Fourier bounds no real improvement

Conlon, Rummel 1808.05916

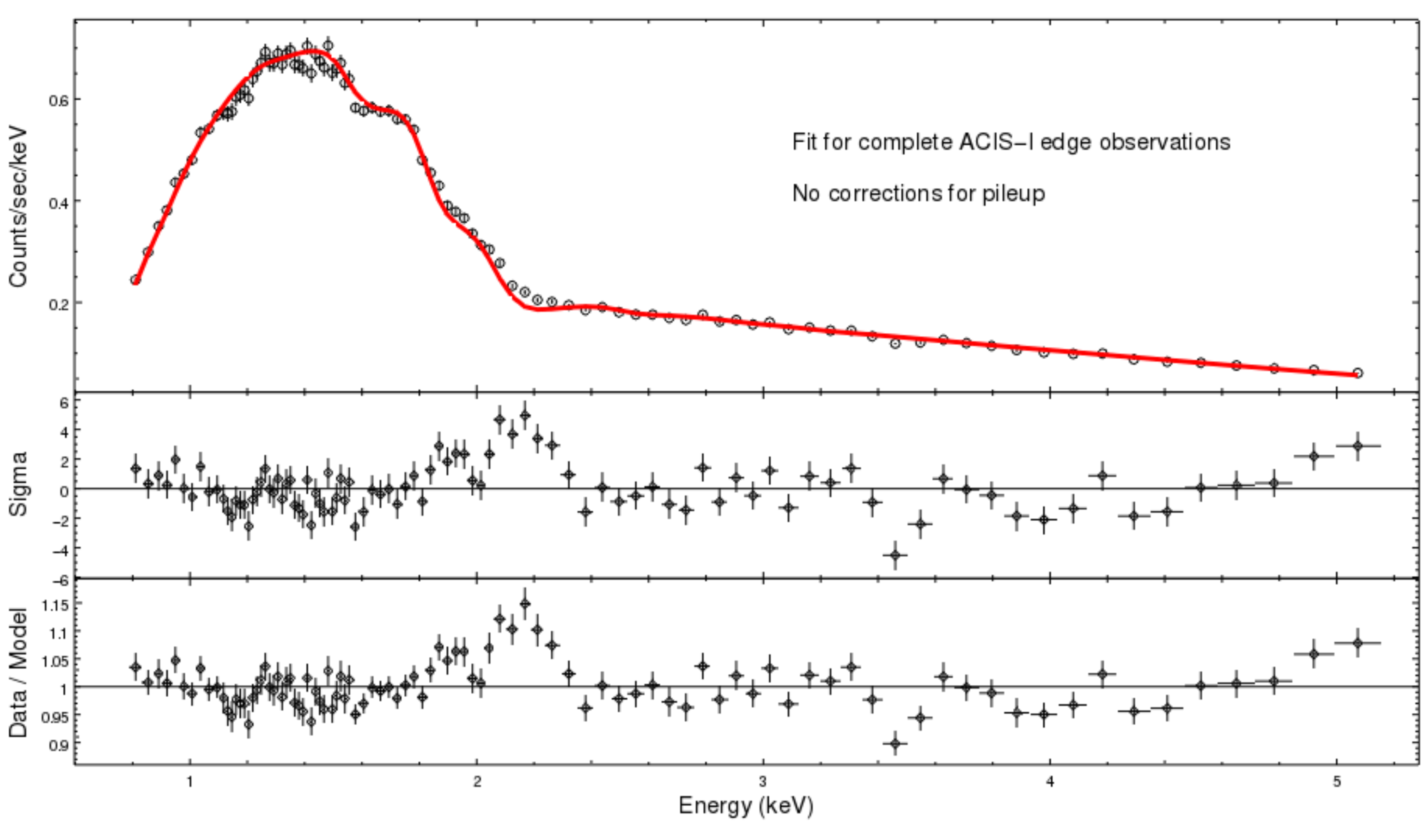


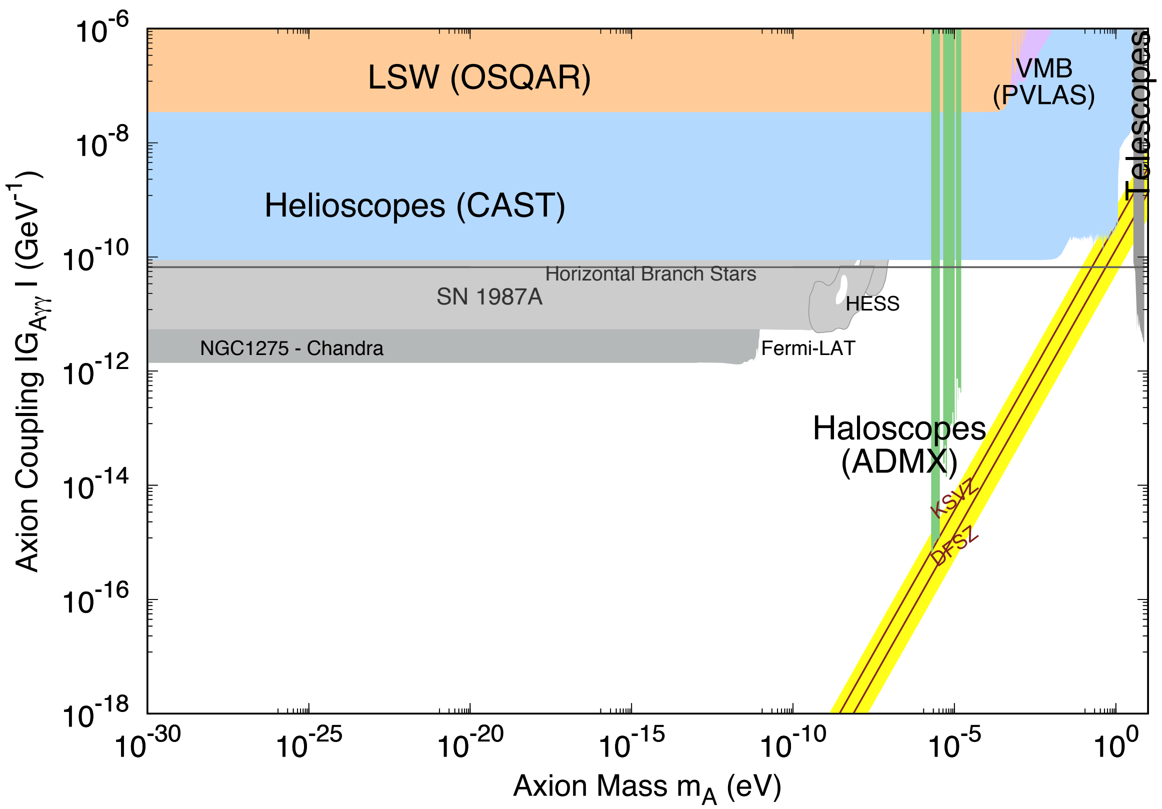
Standard bounds

NGC1275: 1605.01043
Other sources: 1704.05256
Athena bounds: 1707.00176
with: Conlon, Day, Jennings;
Berg, Muia, Powell, Rummel

- Spectral distortions (but unknown where & strength as dependent on magnetic field)
- Poisson noise vs. signal
- Essentially: fluctuations larger than noise
- Human-made: Fourier bounds no real improvement

Conlon, Rummel 1808.05916



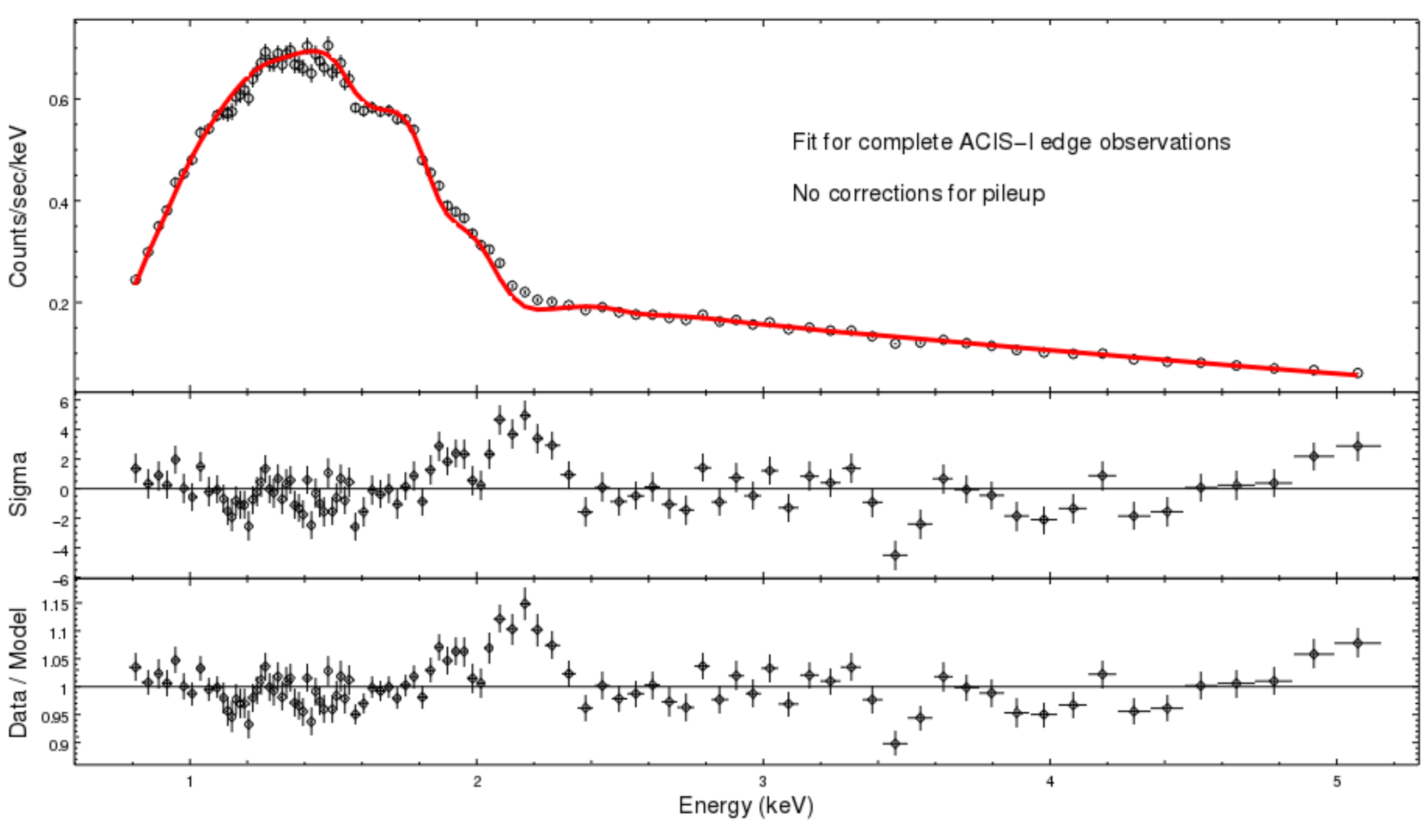


Standard bounds

NGC1275: 1605.01043
Other sources: 1704.05256
Athena bounds: 1707.00176
with: Conlon, Day, Jennings;
Berg, Muia, Powell, Rummel

- Spectral distortions (but unknown where & strength as dependent on magnetic field)
- Poisson noise vs. signal
- Essentially: fluctuations larger than noise
- Human-made: Fourier bounds no real improvement

Conlon, Rummel 1808.05916



Standard bounds

NGC1275: 1605.01043
Other sources: 1704.05256
Athena bounds: 1707.00176
with: Conlon, Day, Jennings;
Berg, Muia, Powell, Rummel

- Spectral distortions (but unknown where & strength as dependent on magnetic field)
- Poisson noise vs. signal
- Essentially: fluctuations larger than noise
- Human-made: Fourier bounds no real improvement

Conlon, Rummel 1808.05916