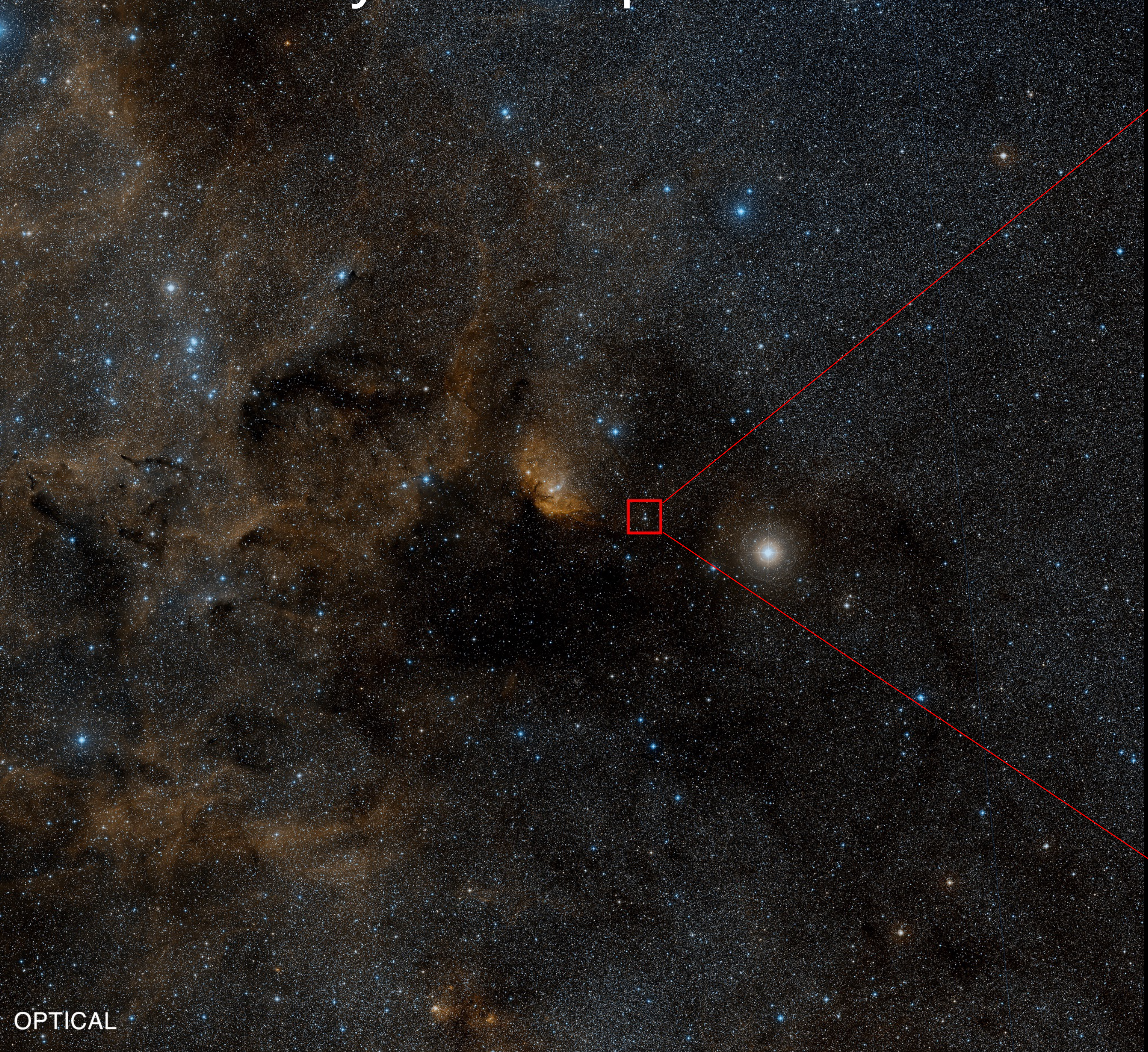
Emulating Photon Pile-up Effects on X-ray Spectra with a Neural Network

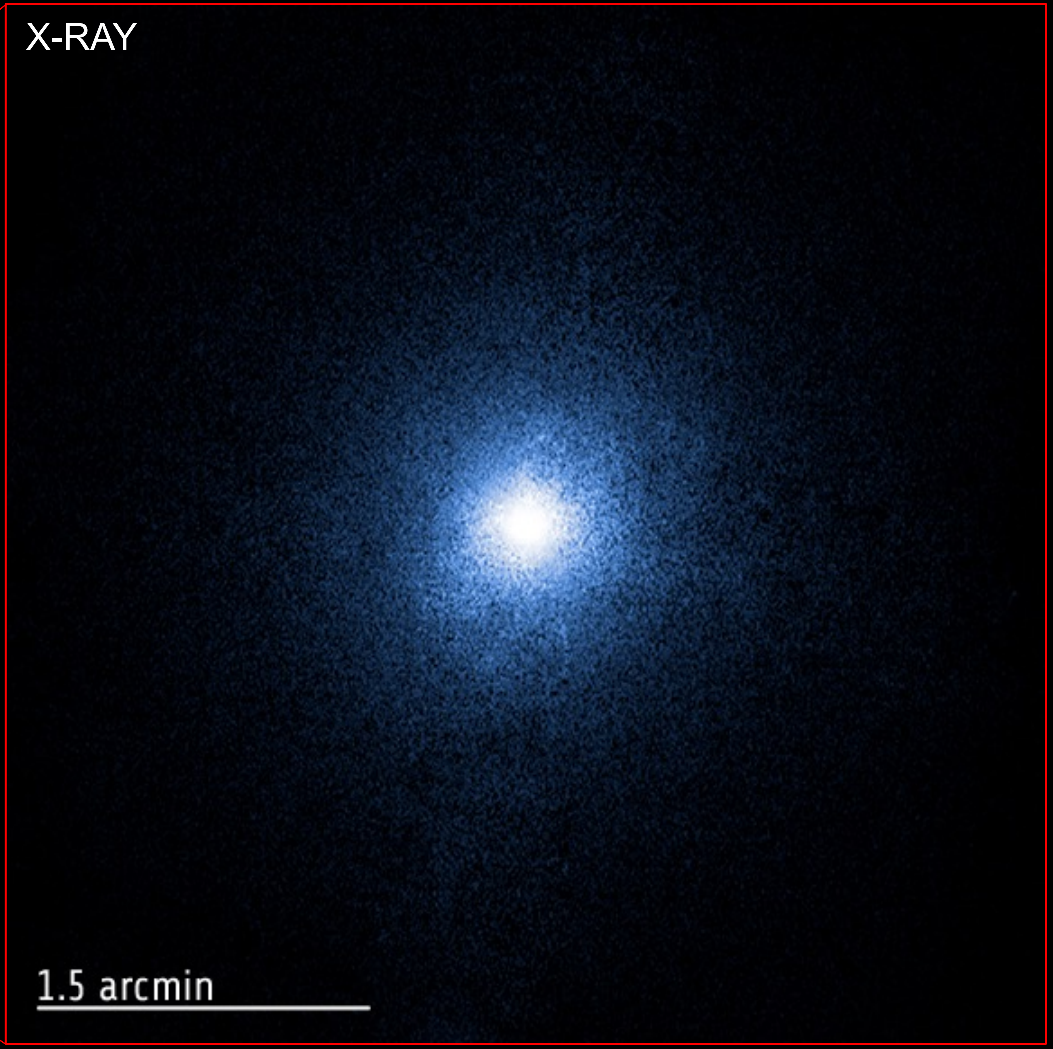
Justina R. Yang
with Daniela Huppenkothen, Rafael Martínez-Galarza, Vinay Kashyap, Douglas Finkbeiner, Aneta Siemiginowska

Wednesday, 6 November 2024

justina.yang@cfa.harvard.edu

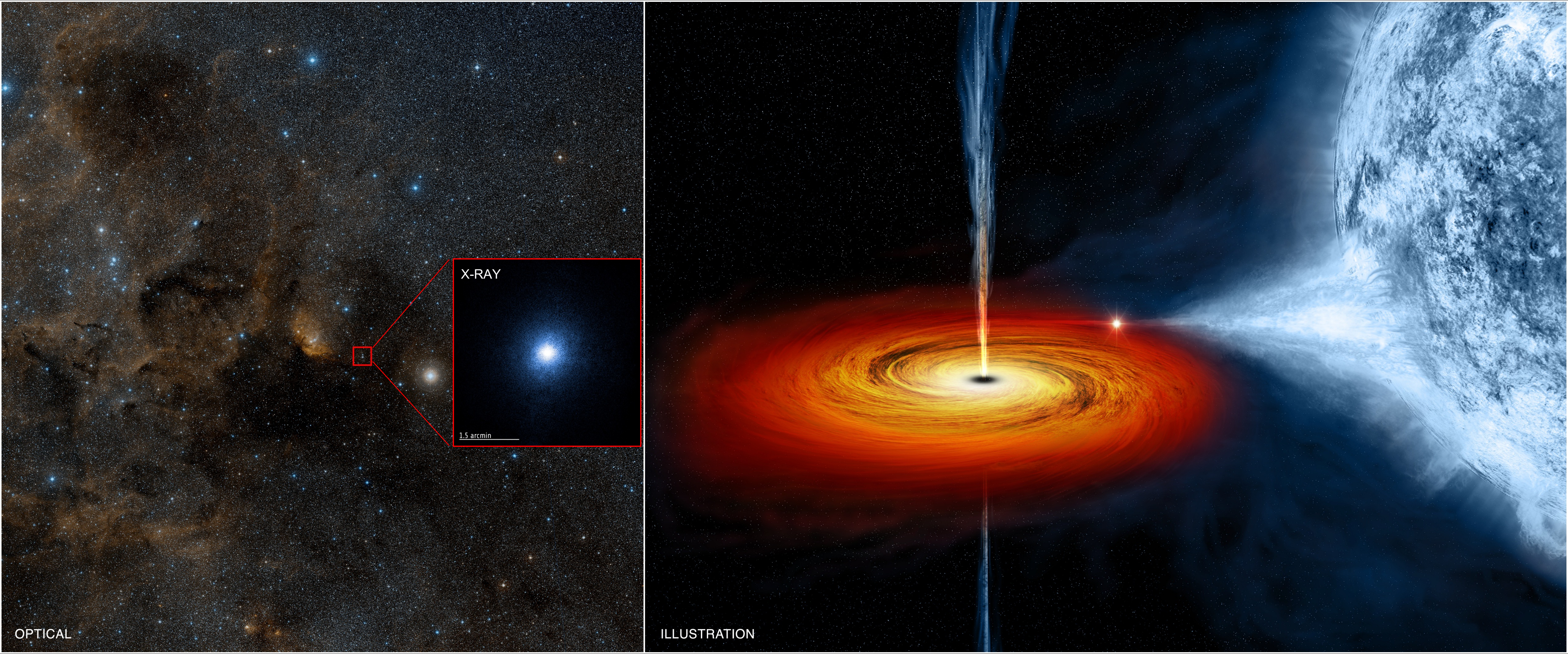
X-ray telescopes enable us to examine distant objects





Cygnus X-1: > 6000 light-years away
(original image: NASA/CXC/M. Weiss)

X-ray telescopes enable us to examine distant objects



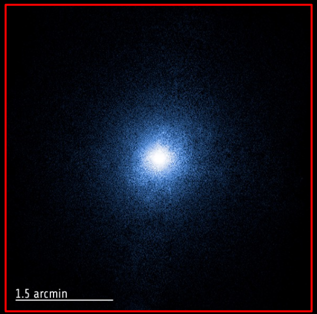
Cygnus X-1 is a black hole X-ray binary (original image: NASA/CXC/M. Weiss)

X-ray data arrives in event lists from which we can extract spectra

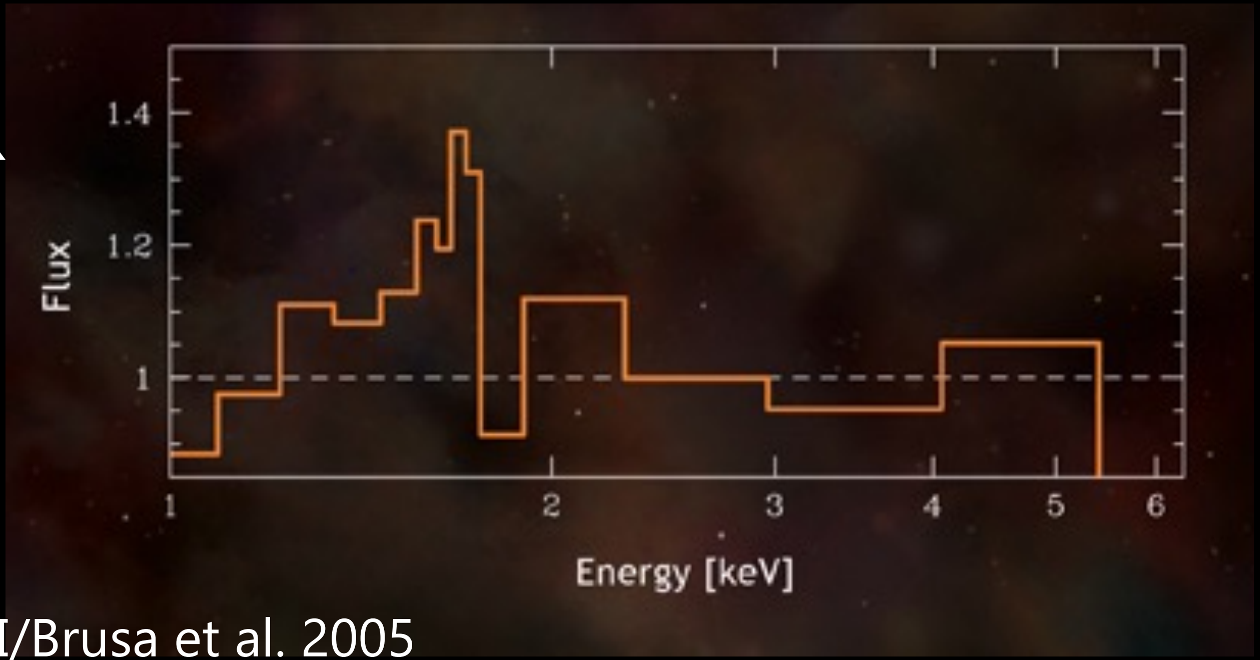
Event list

time	energy	position (x,y)	quality parameter (grade)	…

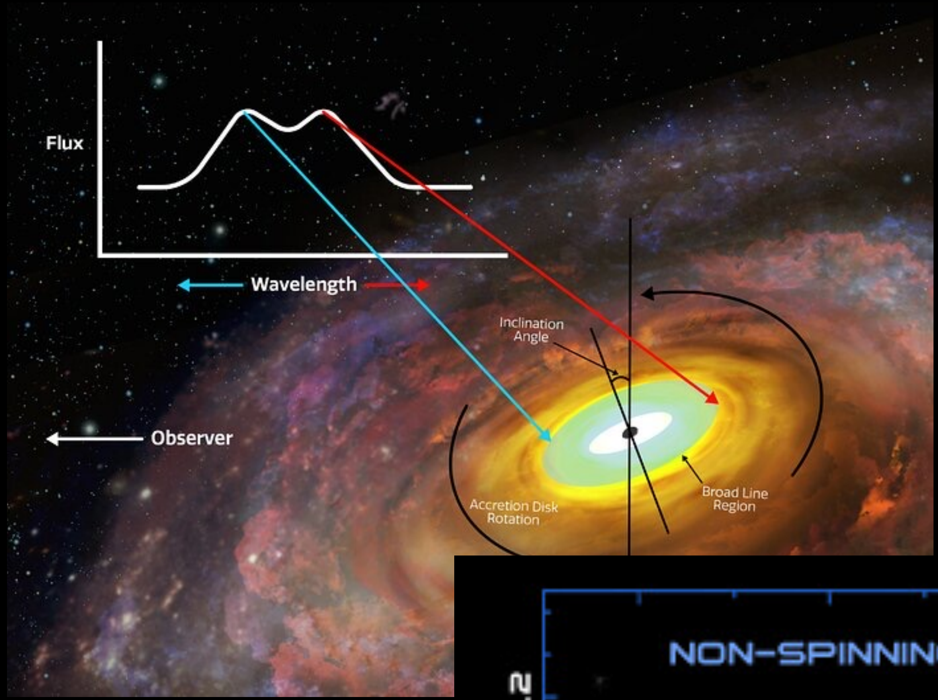
Images



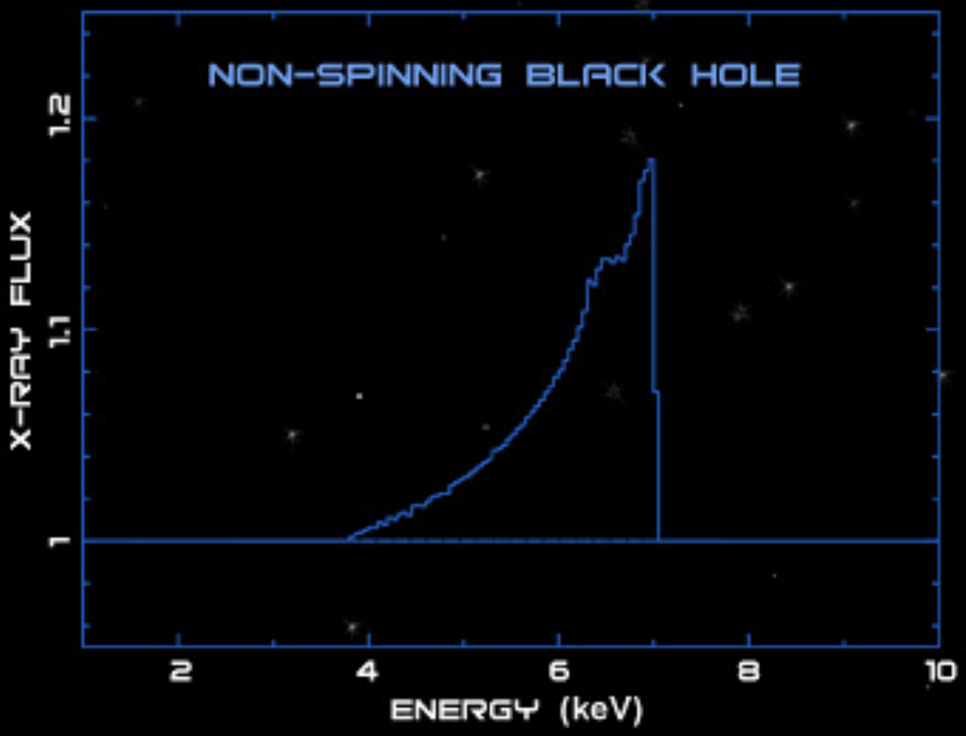
Spectra

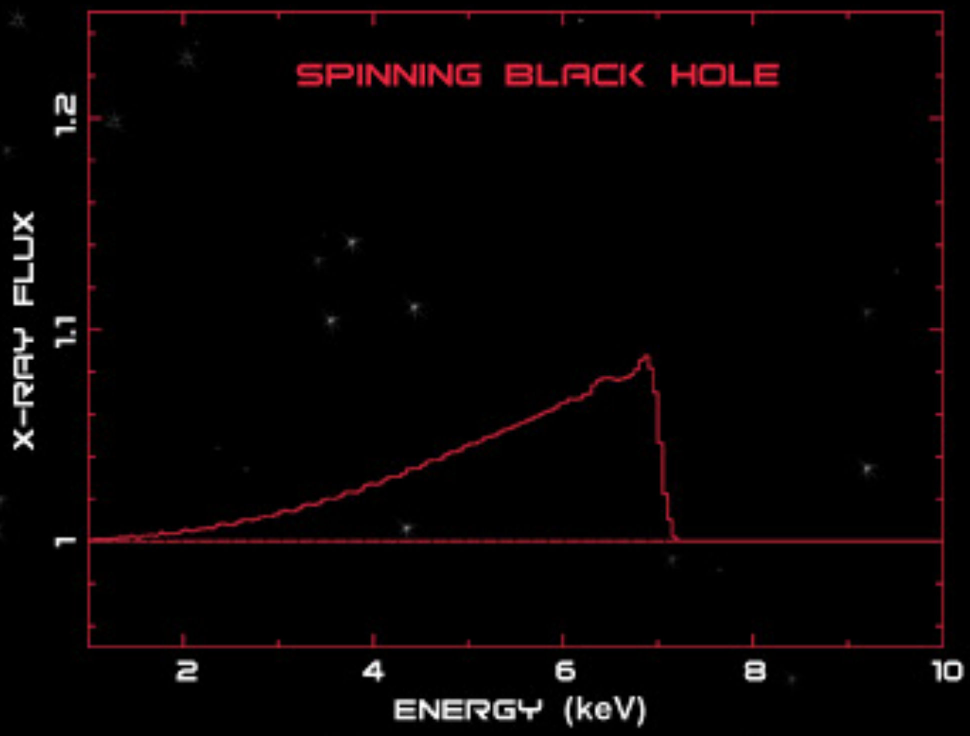


Images: NASA/CXC/M. Weiss, NASA/CXC/MPI/Brusa et al. 2005



X-ray spectra contain information about astrophysical object properties



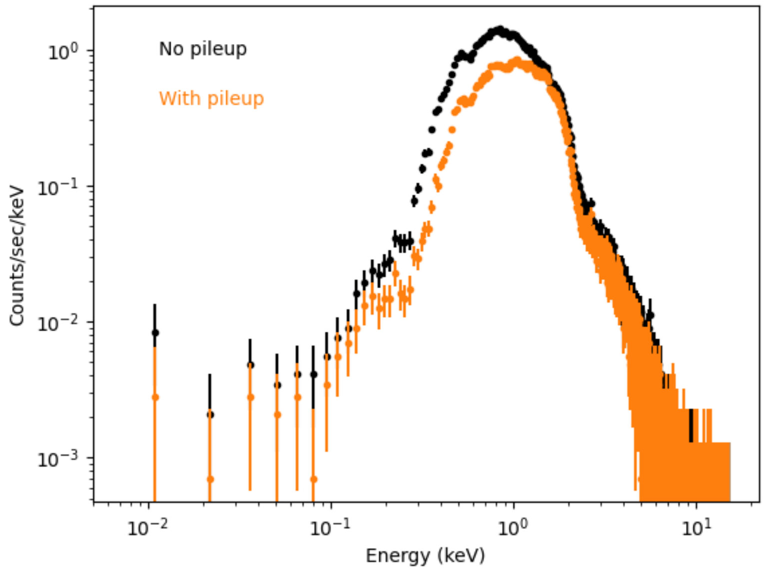


Images: NOIRLab/NSF/AURA/P. Marenfeld, NASA/CXC/M. Weiss, NASA/CXC/SAO/Miller et al. 2002

Photon pile-up

is an instrumental effect related to detector frame rate that is common in X-ray detectors. It affects observed events and distorts spectra.

Accurately understanding pile-up is important for measuring physical properties such as black hole spin.



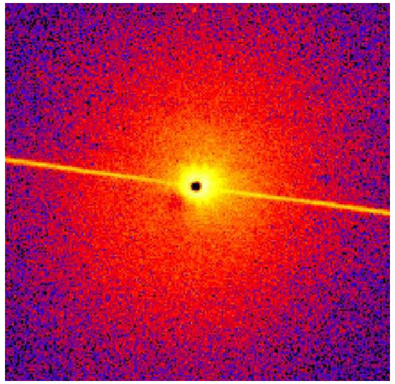


Image: “The *Chandra* ABC Guide to Pileup”

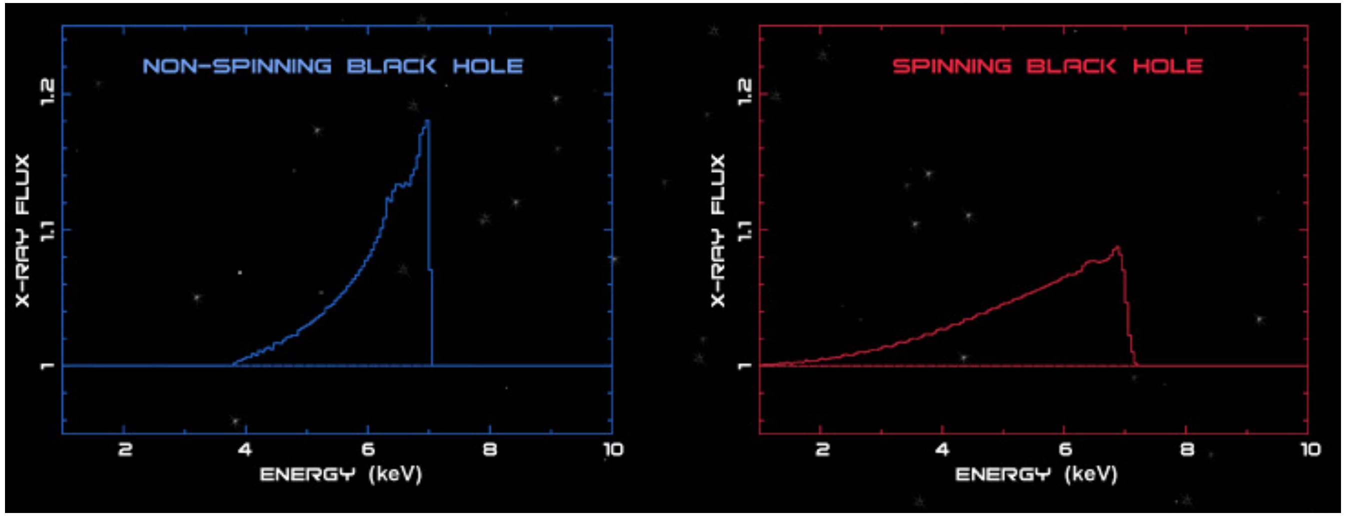
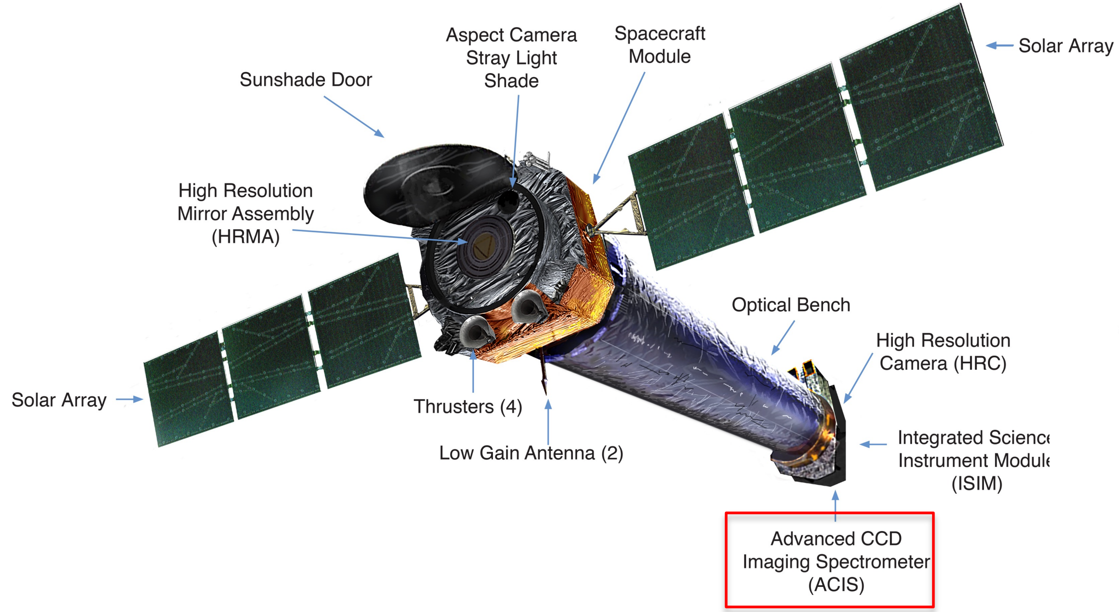
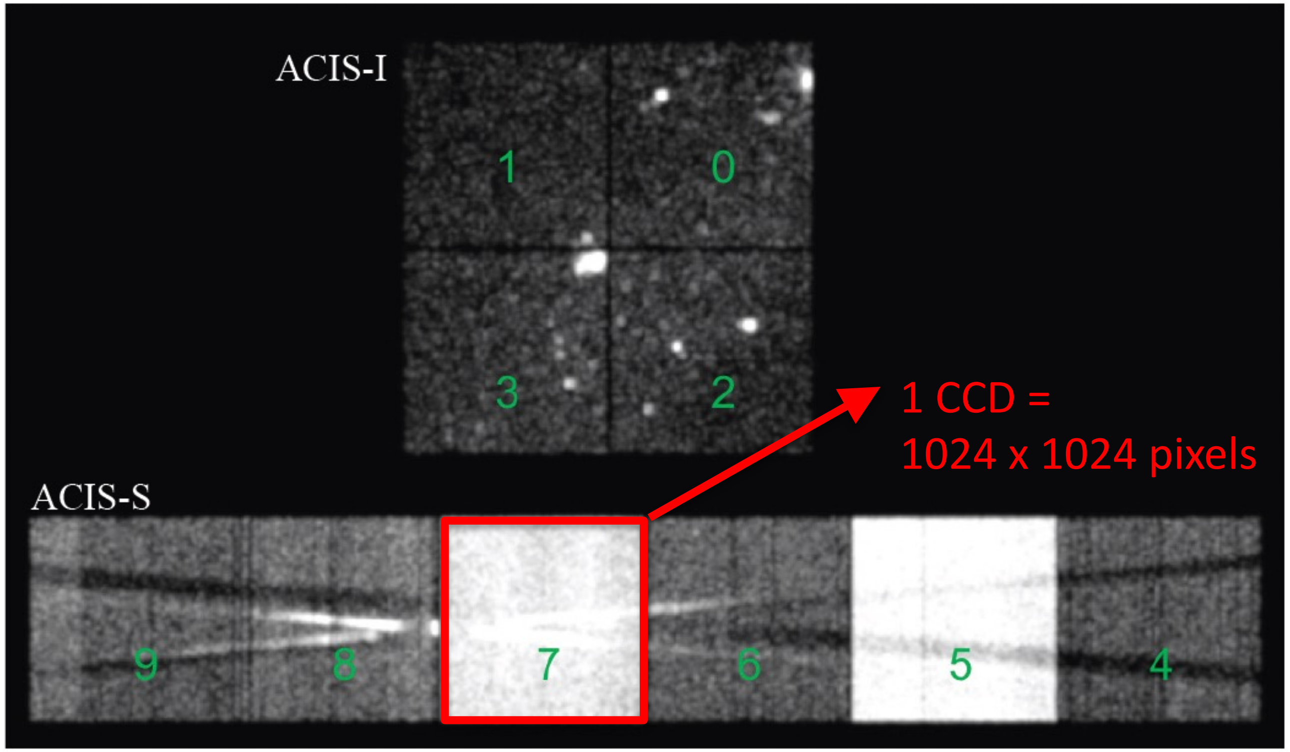


Image: NASA/CXC/M. Weiss, NASA/CXC/SAO/Miller et al. 2002

Chandra X-Ray Observatory, ACIS-S instrument

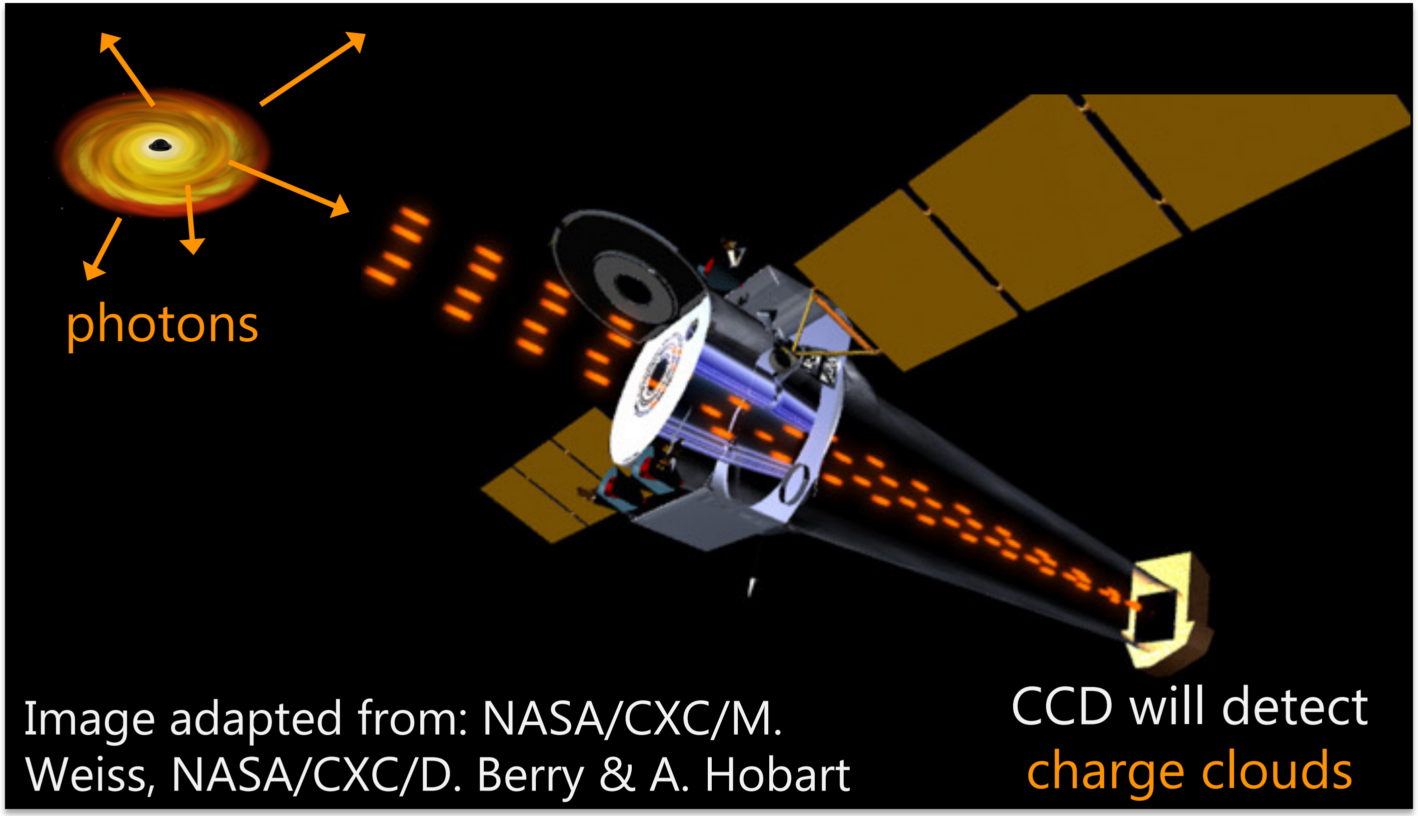


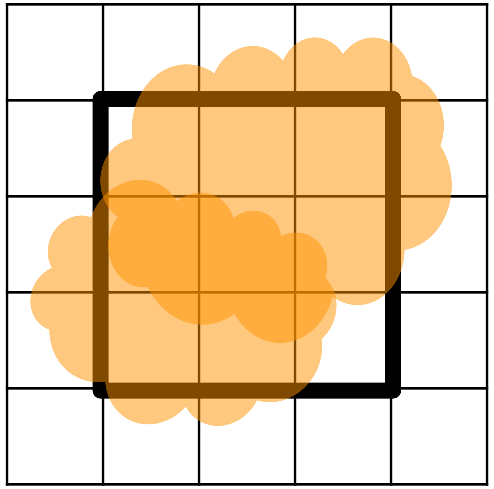


The ACIS focal plane layout in 'sky' coordinates with each CCD labelled with its identification number.

CCD (charge-coupled device):
a set of coupled capacitors that converts an incoming photon to a charge-cloud of electron-hole pairs and produces an electric signal linearly related to photon energy

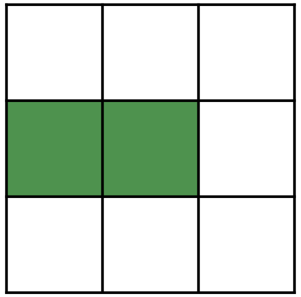
Photon pile-up occurs when a detector registers multiple photons as a single event





2 or more charge clouds in a single frame time and pixel region

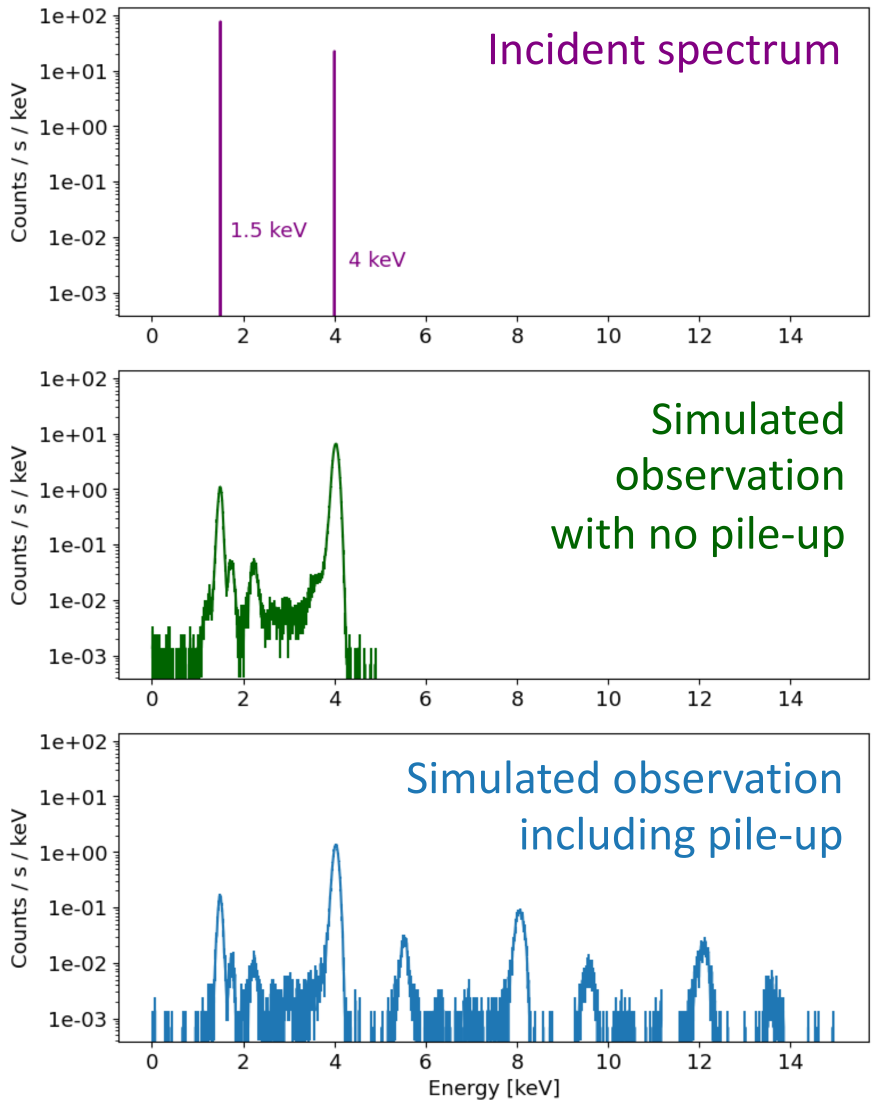
single event (x, y, E, …) with grade



Pile-up affects both events and spectra

Pile-up causes

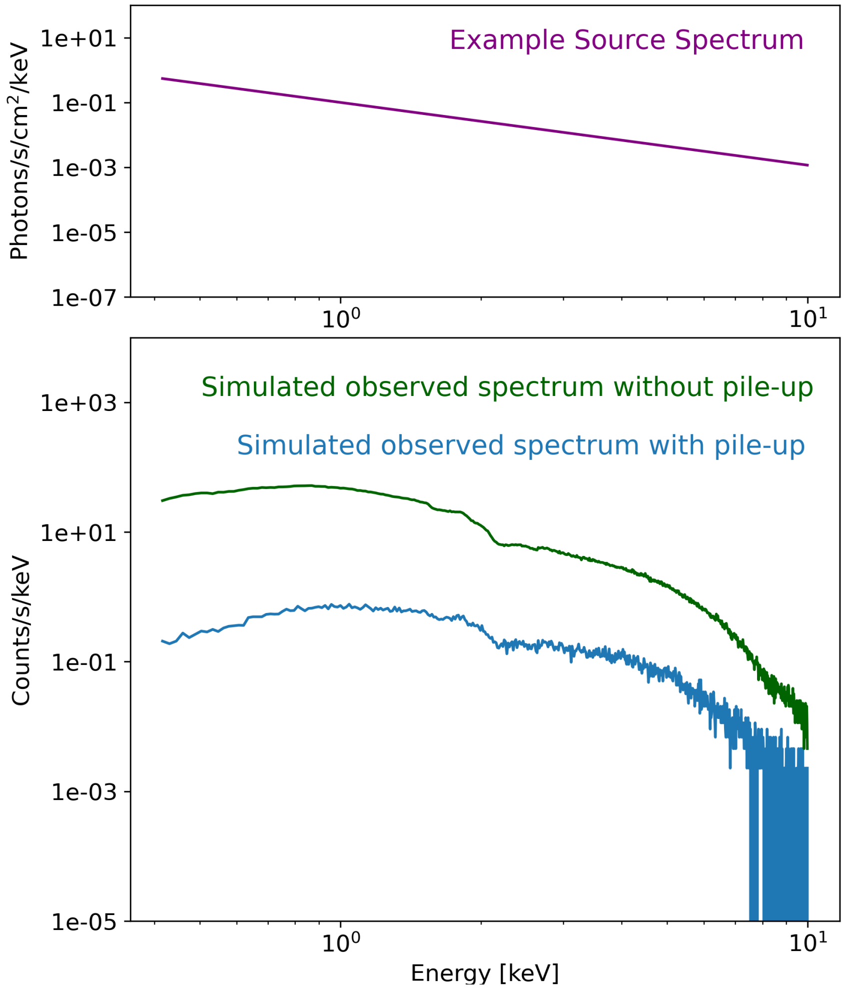
- energy migration (events are associated with higher energies)
- decrease in event count rates
- grade migration (events may be assigned "grades" of poorer quality and may be rejected)



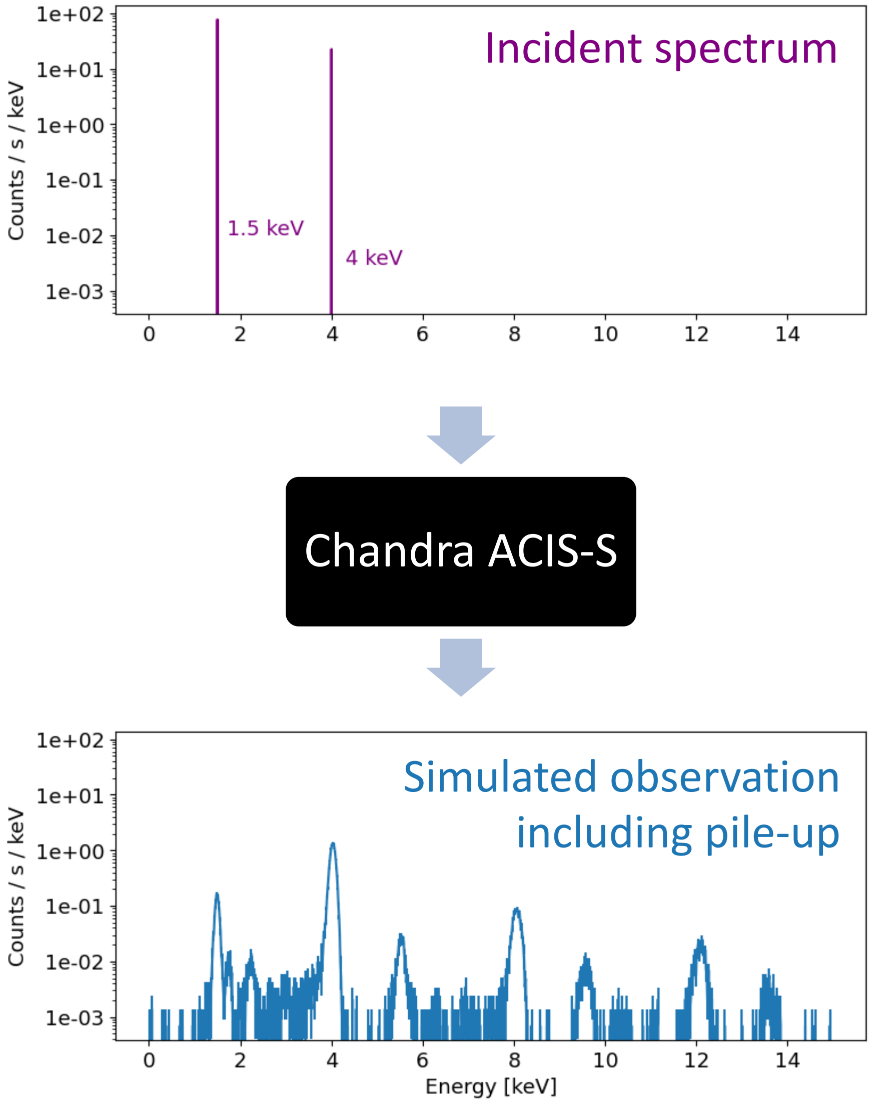
Pile-up affects both events and spectra

Pile-up causes

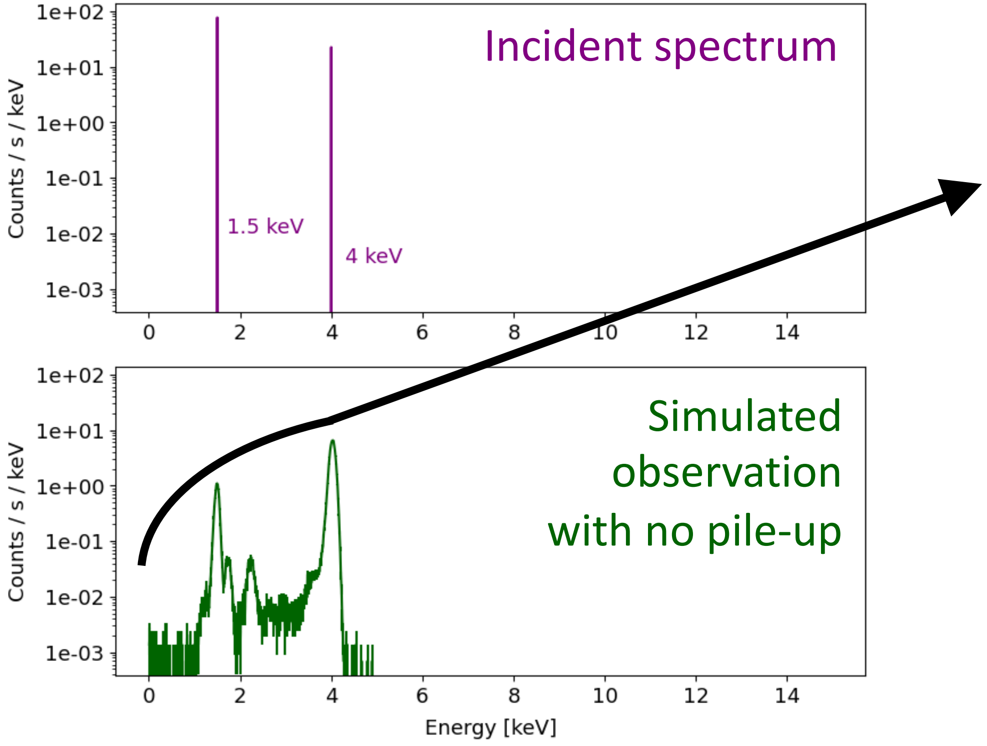
- energy migration (events are associated with higher energies)
- decrease in event count rates
- grade migration (events may be assigned "grades" of poorer quality and may be rejected)



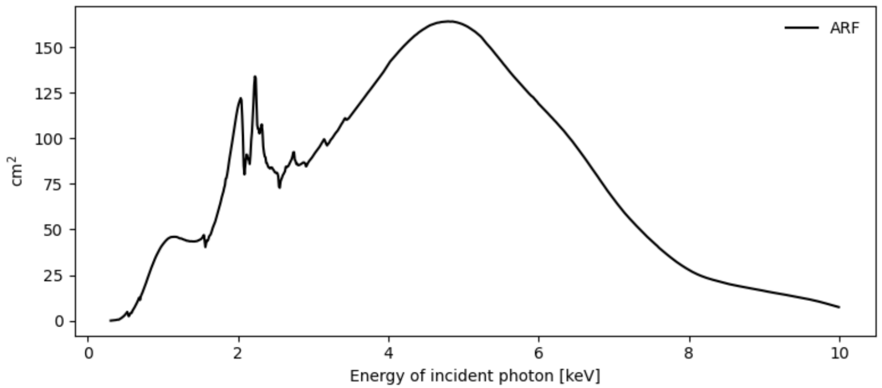
Pile-up is only one part of the instrument response



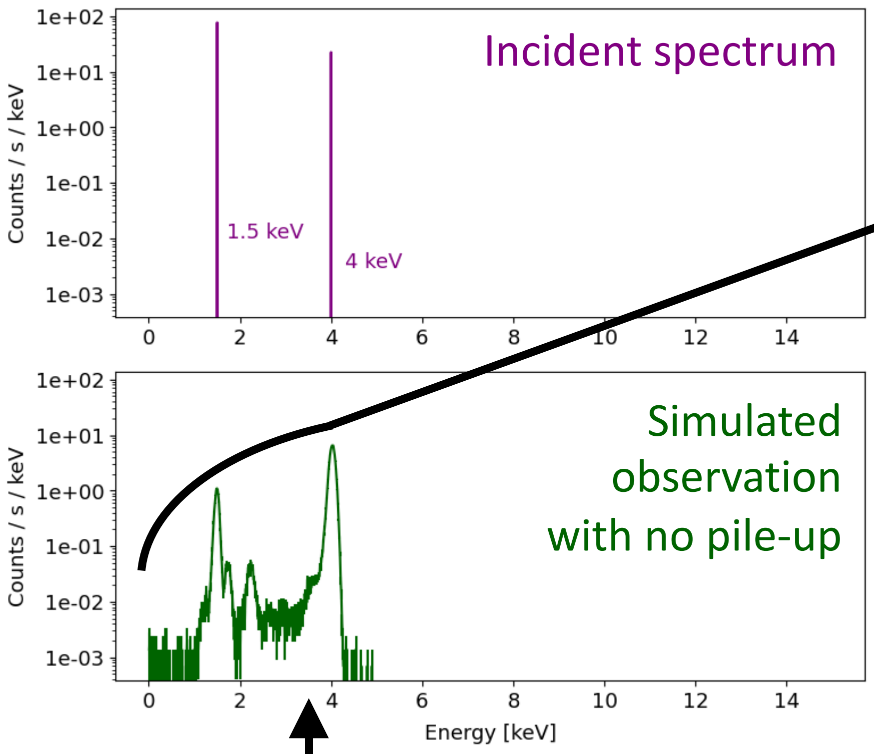
Parts of the instrument response are described by the ARF and RMF, which are well-understood and calibrated



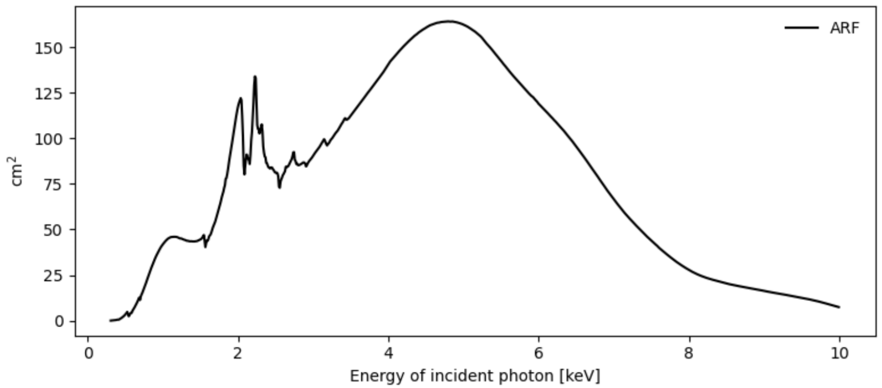
Detector sensitivity is parametrized by the **ARF** (ancillary response file): includes effective area, quantum efficiency



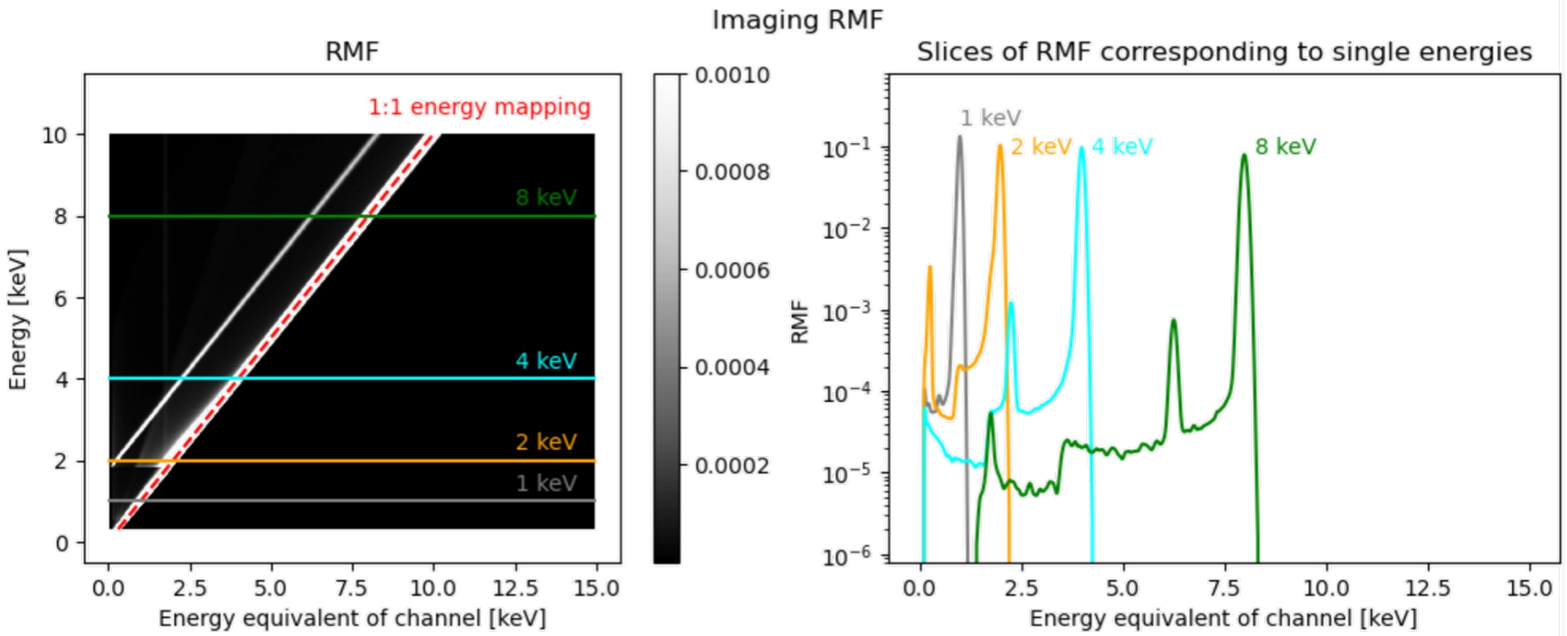
Parts of the instrument response are described by the ARF and RMF, which are well-understood and calibrated



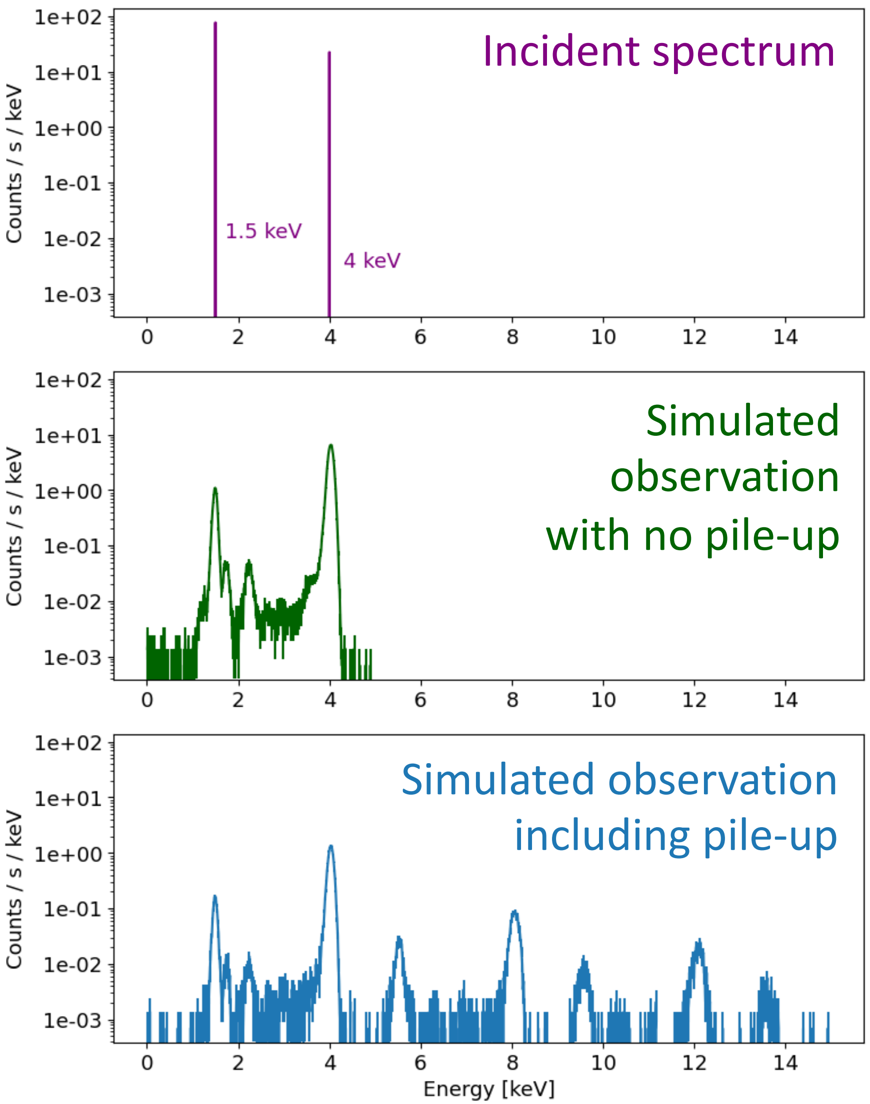
Detector sensitivity is parametrized by the **ARF** (ancillary response file): includes effective area, quantum efficiency

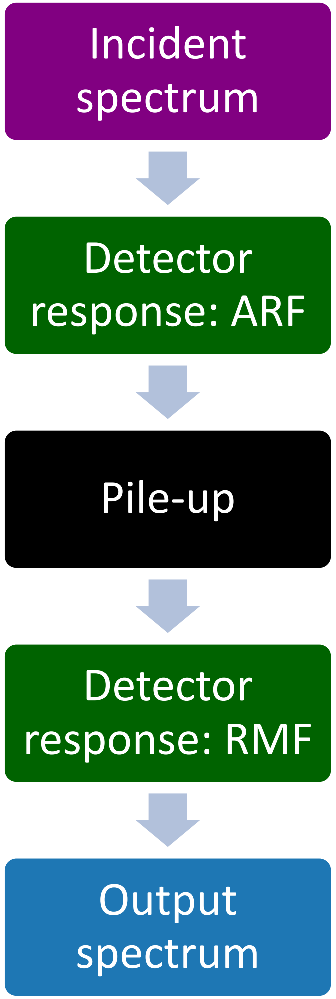


Signal broadening from RMF (redistribution matrix file): maps incident photon energy to signal distributed over channels.

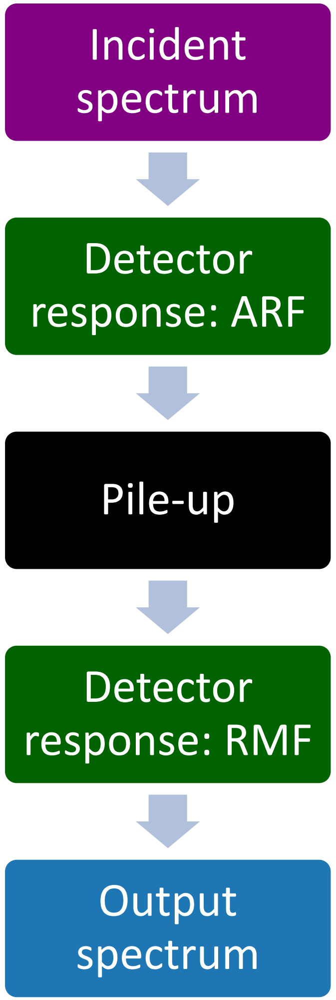


Pile-up is an important effect to model, but it is fundamentally photon-level and is difficult to model analytically





The current standard is the Davis 2001 pile-up model



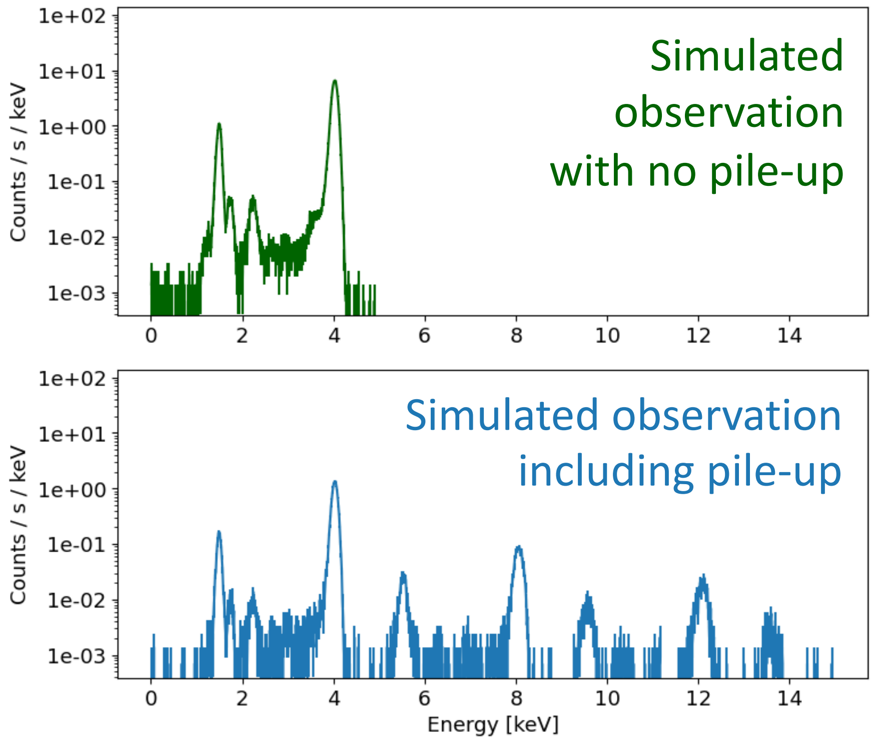
Davis 2001:

Probability (piled-up combination of n photons has a good grade, given that the combination of photons $1, \dots, n-1$ had a good grade) = α

- Often, analysis assumes α = 0.5
- Statistical model does not hold for severe pile-up
- Standard method, e.g. implemented in MARX ("Model of AXAF Response to X-rays"), a ray-trace Monte Carlo simulator of Chandra observations.

Can a neural network emulate a simulated version of pile-up, and eventually the empirical *Chandra*/ACIS pile-up?

- Neural networks are good function approximators and may be able to directly learn the function that distorts the idealized spectrum into an observed spectrum



- Neural networks can be evaluated much faster than MARX simulations

Neural Network Model

parameters
(black hole with accretion disk: gamma, incl, nH, log(ionization), spin)

→

spectral model
(relativistic reflection, relxill, Garcia & Dauser et al. 2014, etc.)

→

incident spectrum

↓ ← ARF, RMF

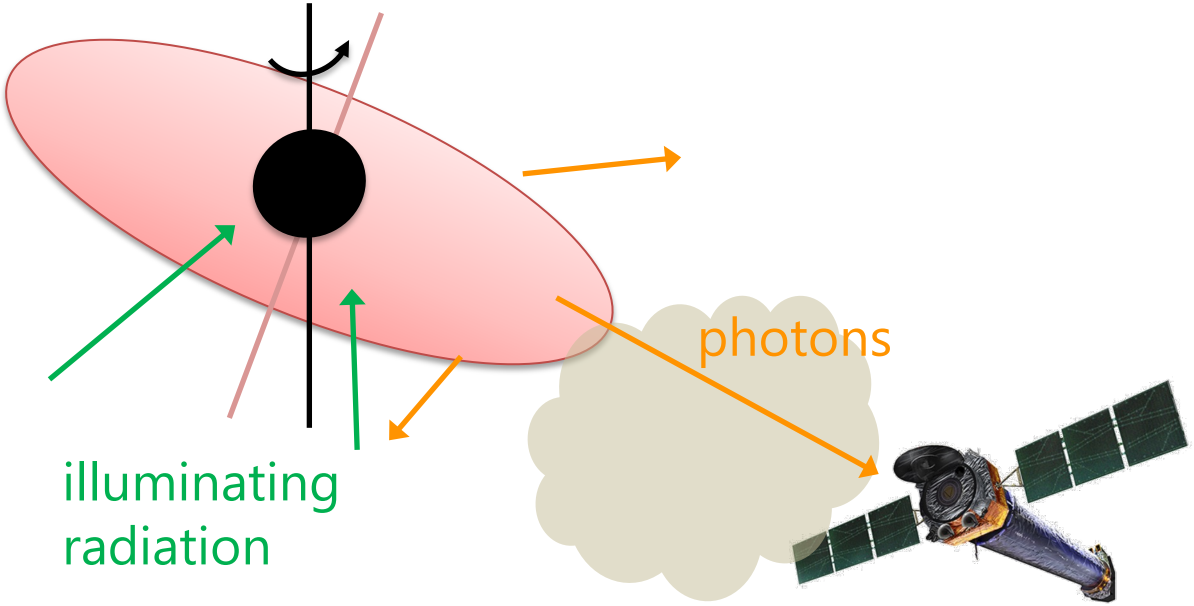
incident spectrum with instrumental response

↓

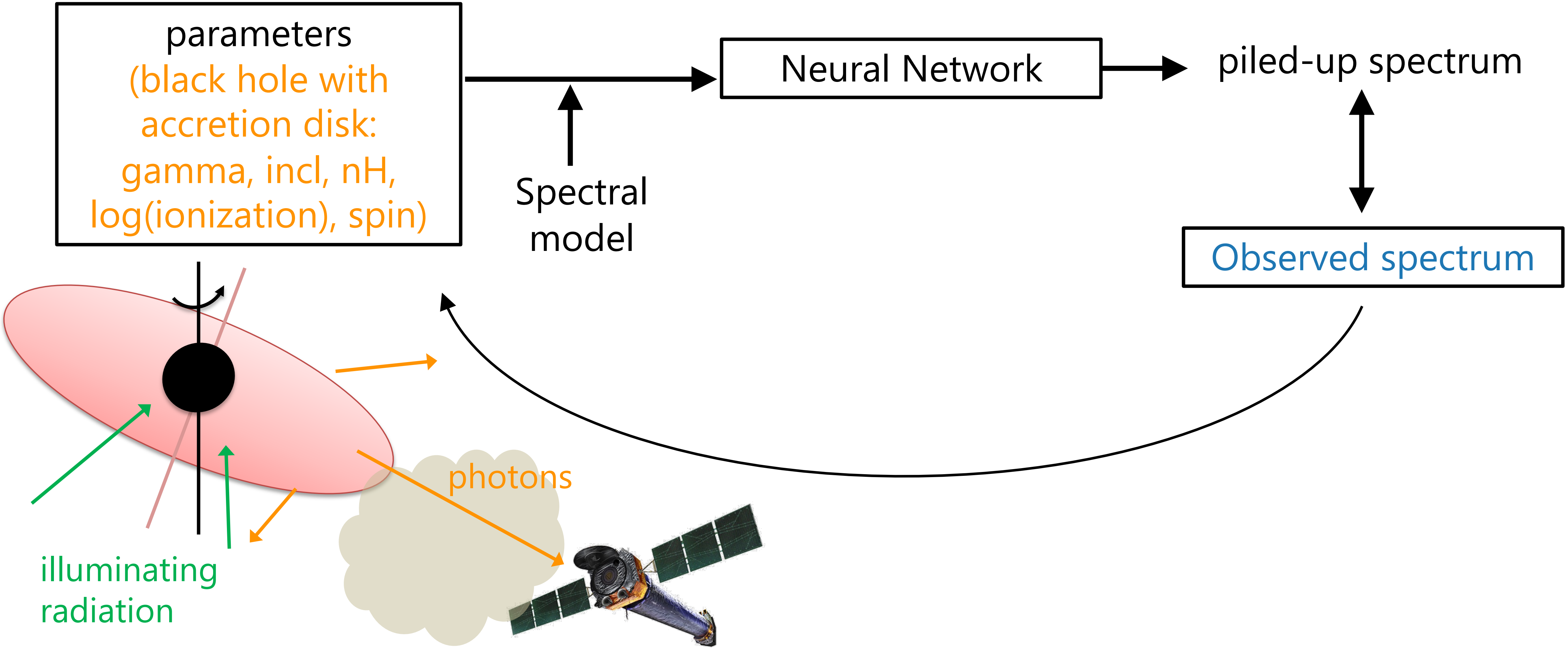
Neural Network

↓

piled-up spectrum

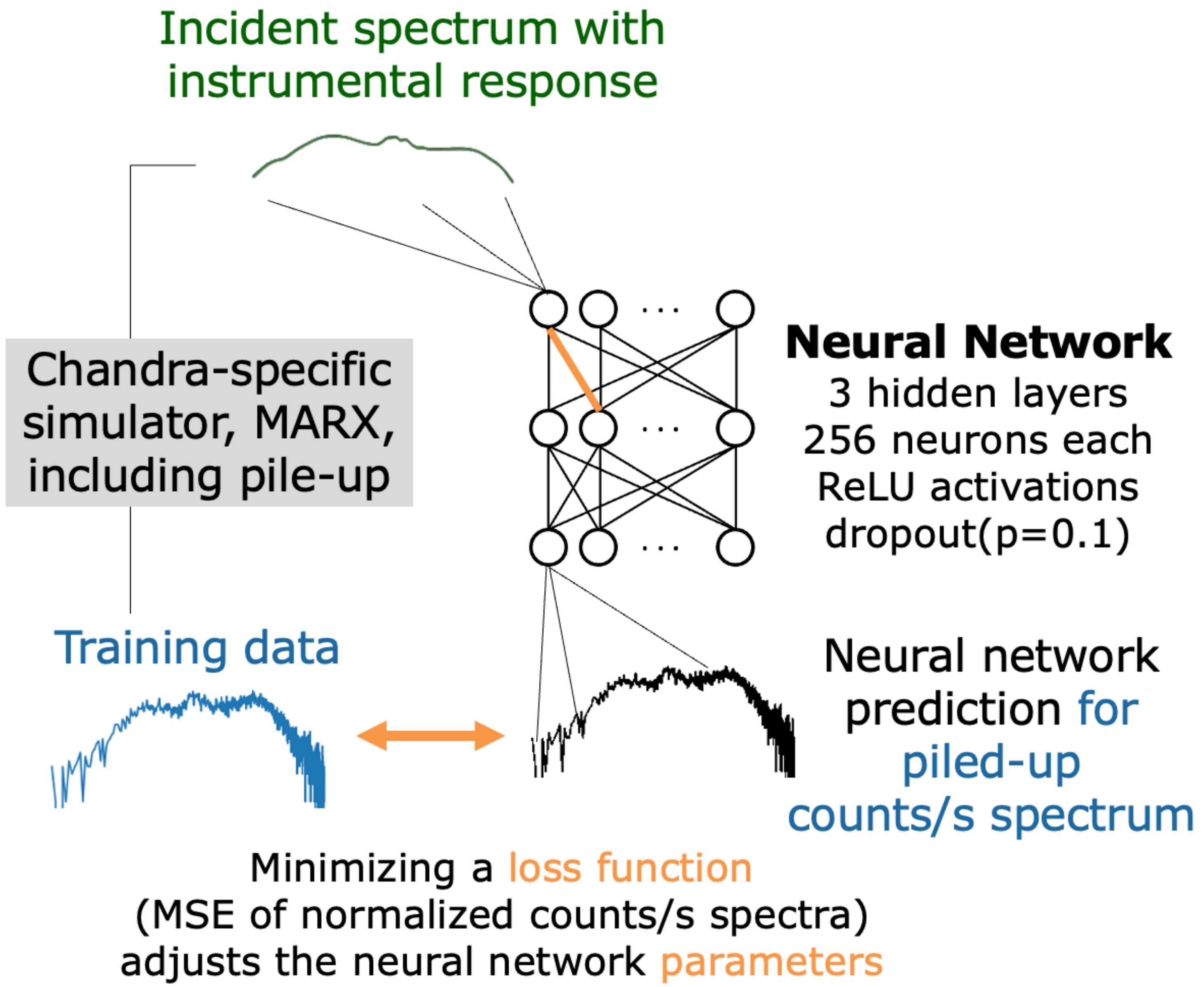


A trained neural network can be incorporated into different parameter inference methods

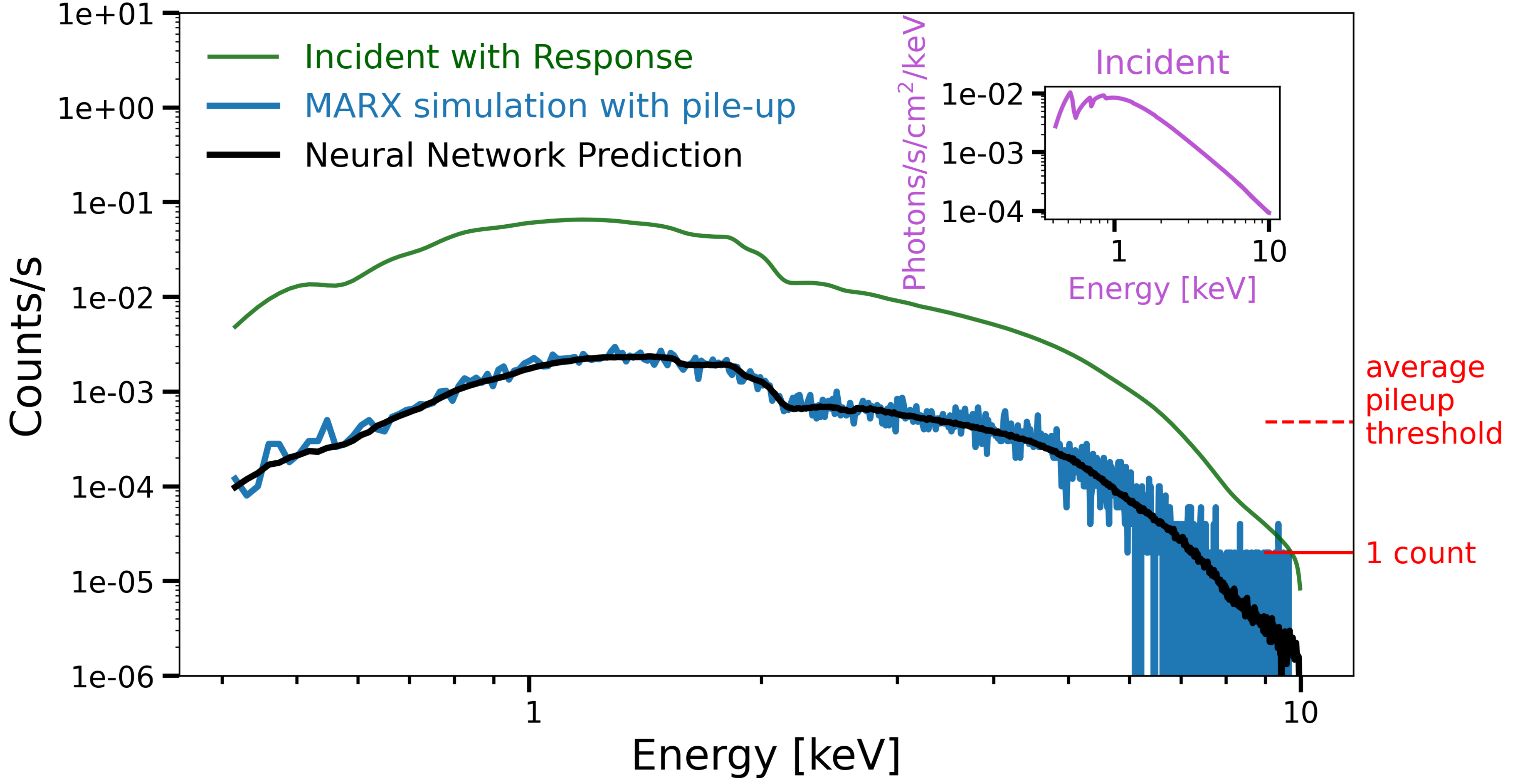


Details of the neural network

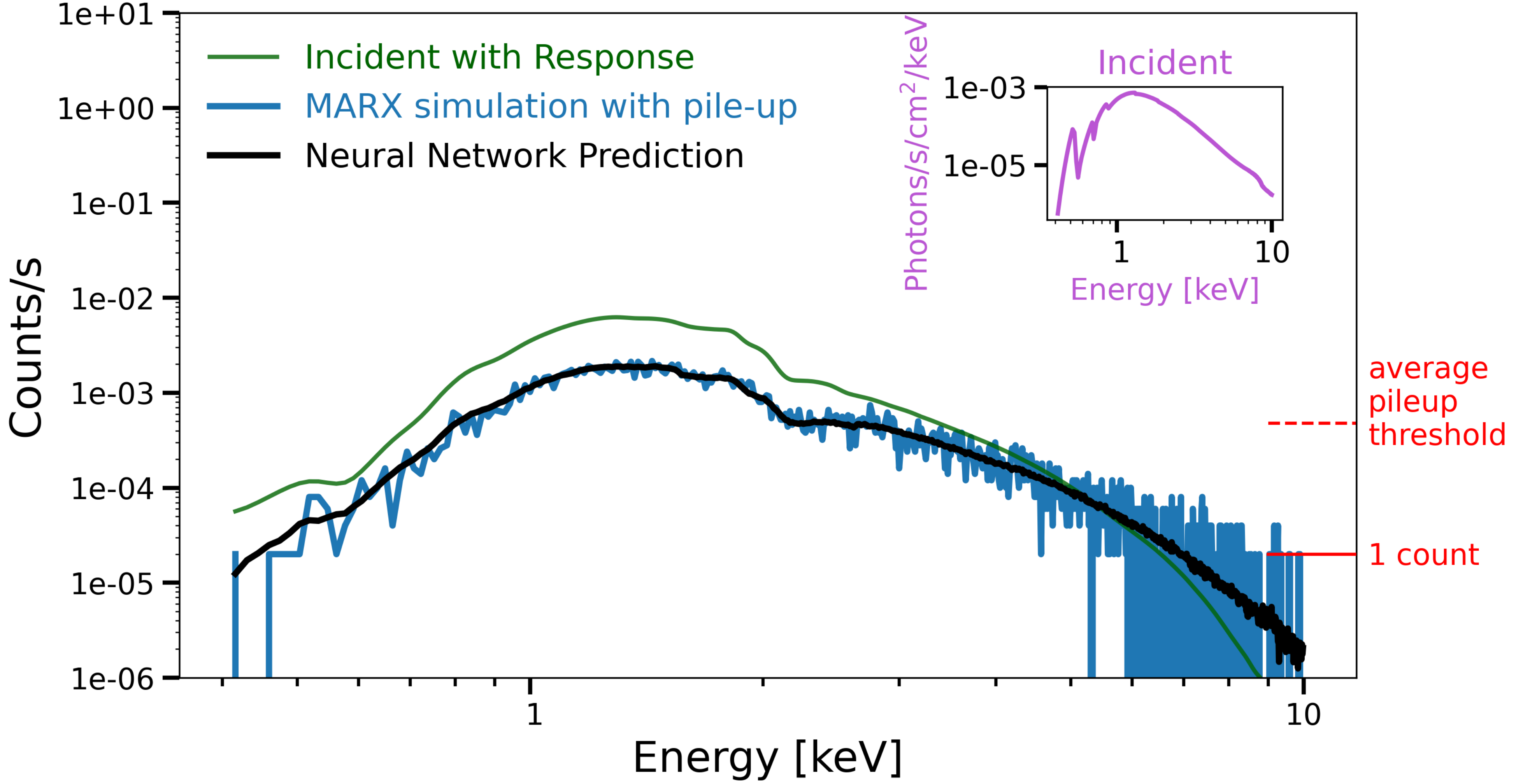
Training data:
>17,000 pairs of (incident spectrum with instrumental response, MARX-simulated spectrum with pile-up and observational noise)



The neural network can closely reproduce a MARX-simulated piled-up spectrum

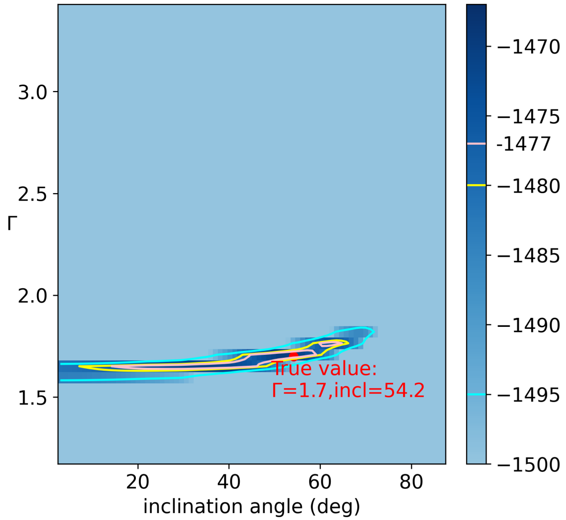


The neural network can closely reproduce a MARX-simulated piled-up spectrum

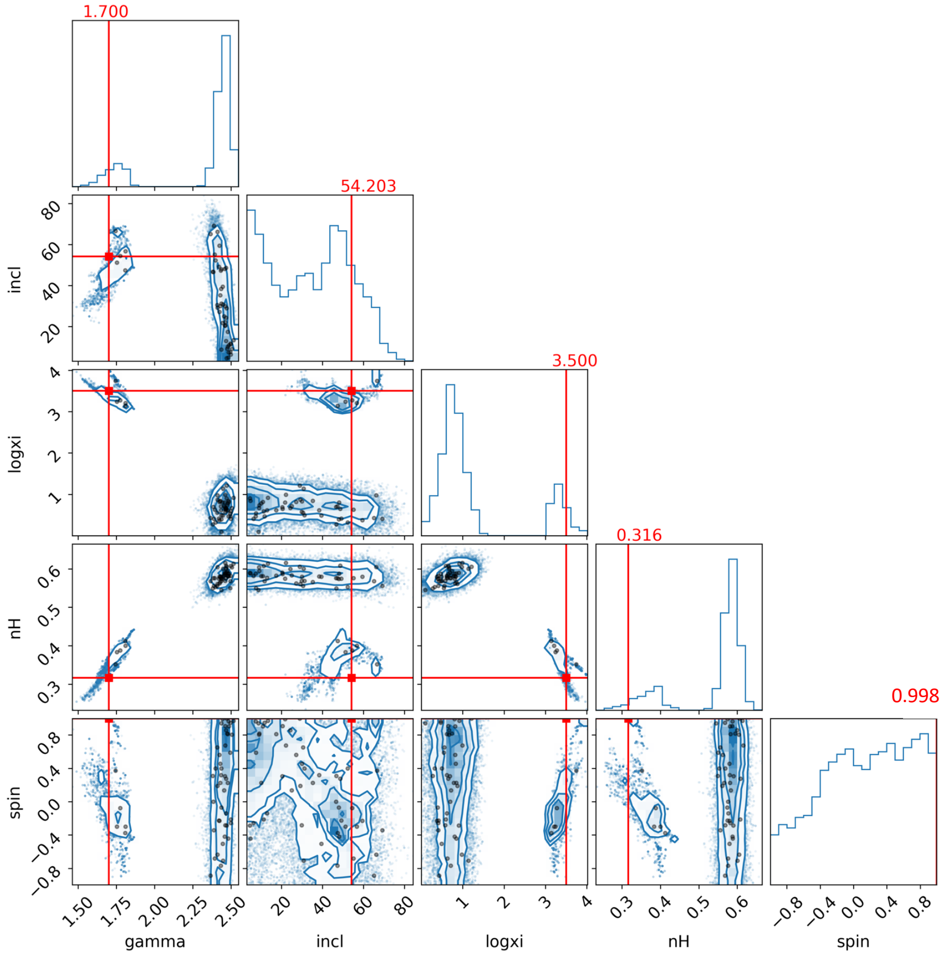


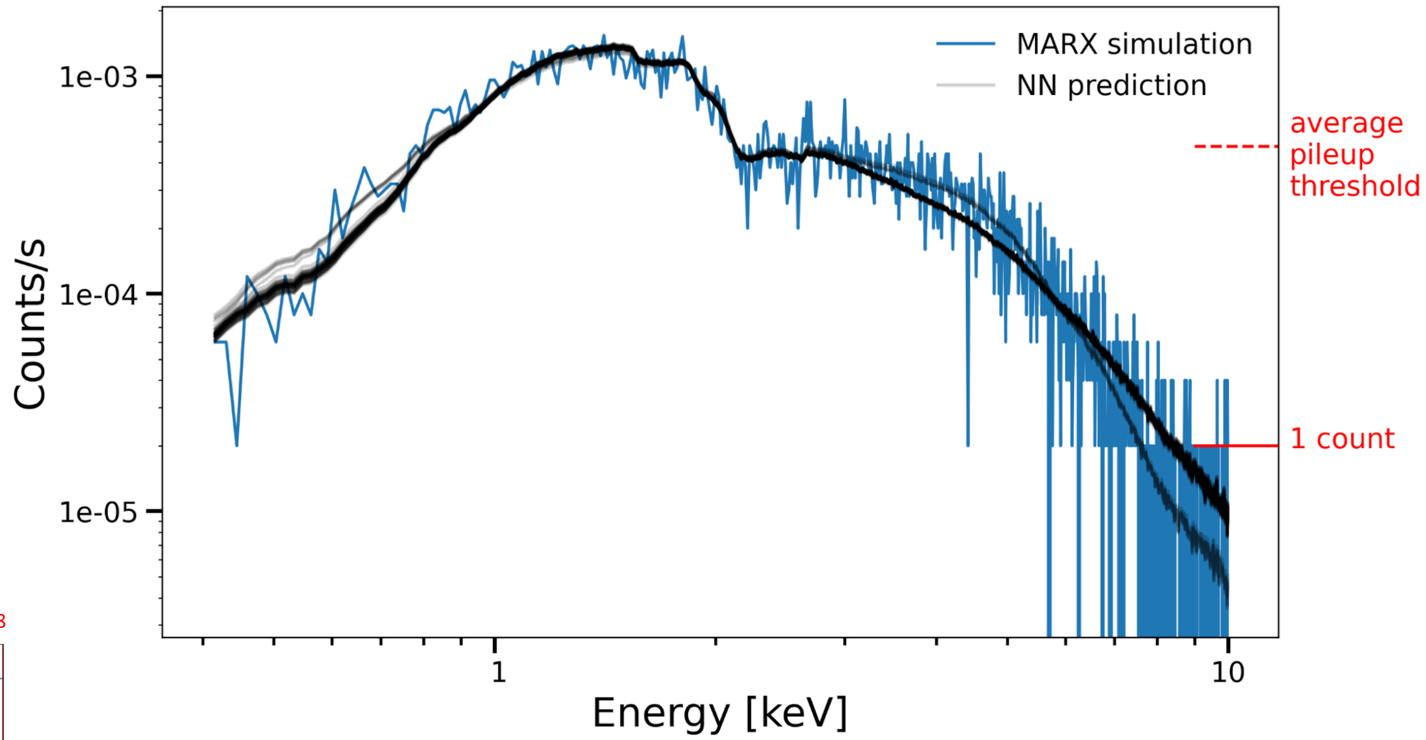
Preliminary work using the neural network in an MCMC is promising

- Parameter inference with **emcee** MCMC code (Foreman-Mackey et al. 2012)
- Optimization quantity = log(P) = log[product over channels k of Probability (MARX counts in channel k, given Poisson mean = NN-predicted counts in channel k)]

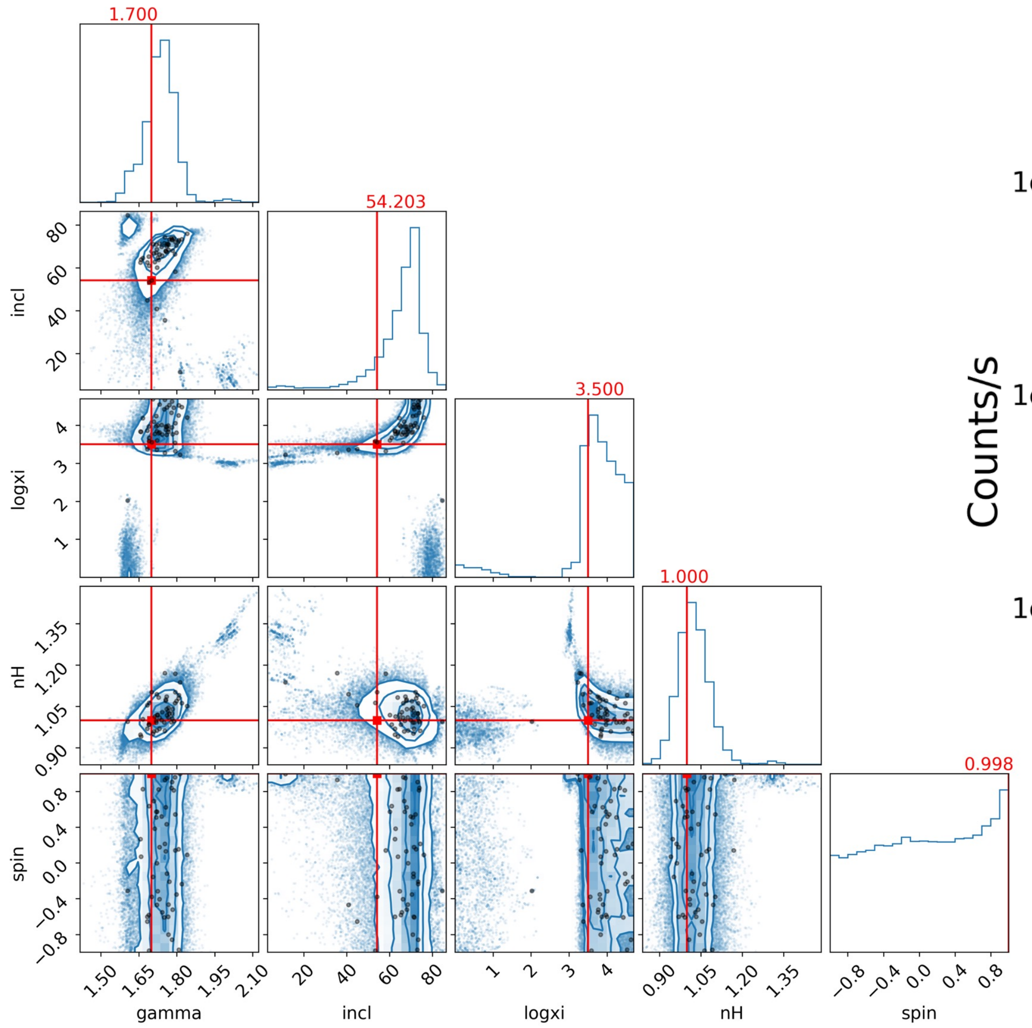


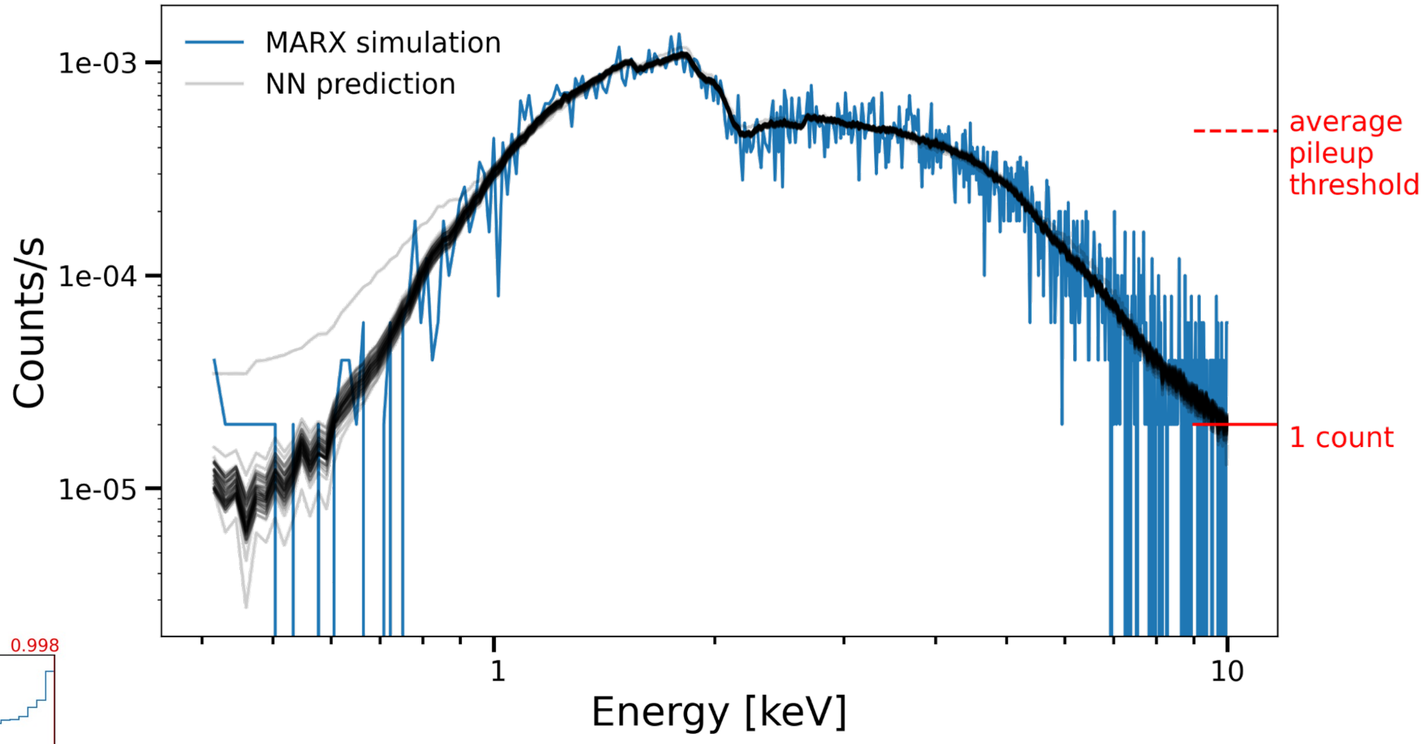
Preliminary work using the neural network in an MCMC is promising





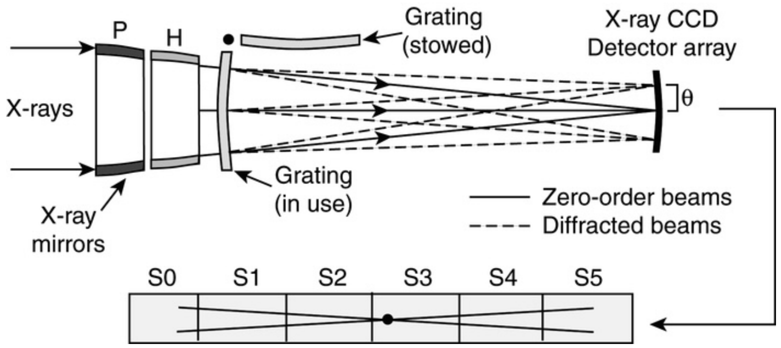
Preliminary work using the neural network in an MCMC is promising

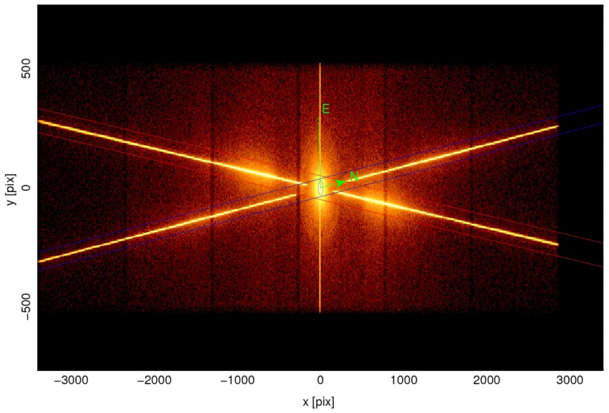




Future work: can the neural network emulate the empirical *Chandra*/ACIS pile-up?

Transfer learning to real Chandra observations with HETG data:



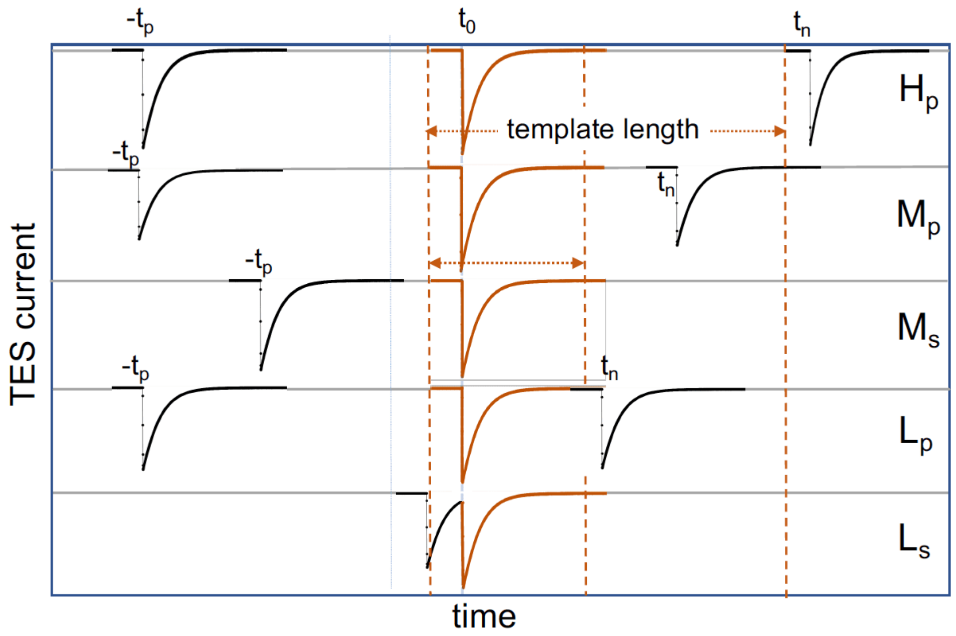


Images: Canizares et al. 2005, TGCat (Huenemoerder et al. 2011)

- Observations taken with *Chandra*/HETG (High-Energy Transmission Grating) distribute photons across more pixels and reduce pile-up in dispersed spectra. We will use 1st-order dispersed spectra as a proxy for unpiled "ground truth" spectra.
- New training data set: ~ 1000 pairs of (unpiled 1st-order HETG spectra, piled-up ACIS-S spectra)

Possible extensions of work

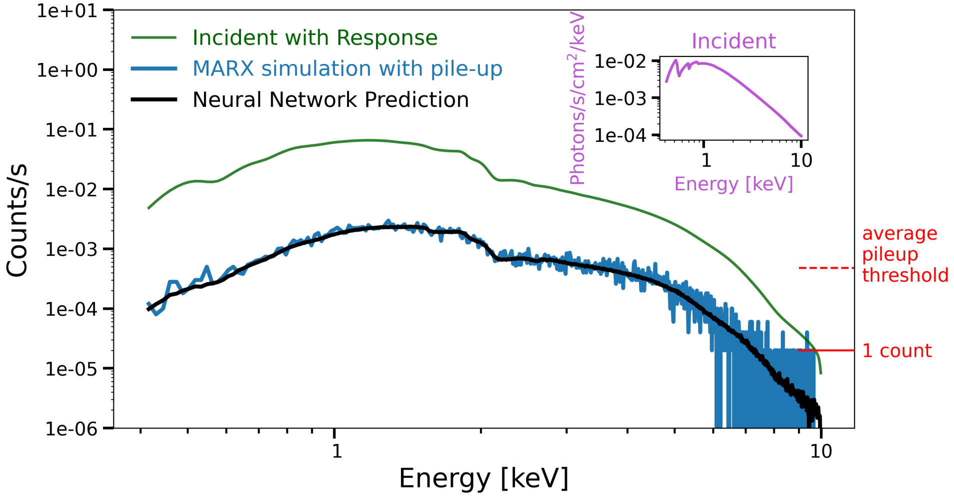
- Relevance for other non-CCD detectors
  - E.g. transition edge sensor (TES) microcalorimeters have a different pile-up effect, related to finite time resolution



Conclusions

- A simple, fully-connected neural network can emulate a simulated version of pile-up.
- We are investigating how to use it for robust and accurate parameter inference from piled-up data.
- Future work will include transfer learning with ~1000 Chandra observations.
- Does your data experience a distortion like the Chandra/ACIS version of pile-up?

Neural Network Prediction



Draws from MCMC