Astrostatistics: The Intersection of Statistics and Outer Space

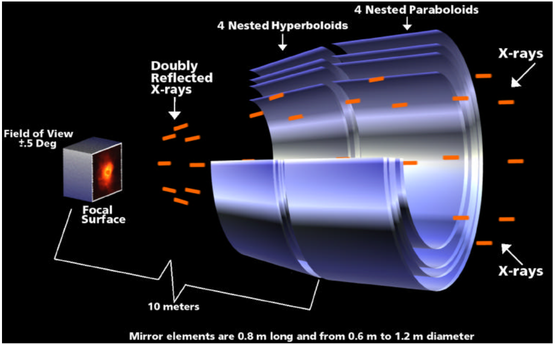
Xiao-Li Meng (HU), Aneta Siemiginowska and Vinay Kashyap (CfA)

Joint work with Y. Chen (Michigan), X. Wang (Two Sigma Inc.), D. van Dyk (Imperial College London), H. Marshall (MIT)

October 29, 2019

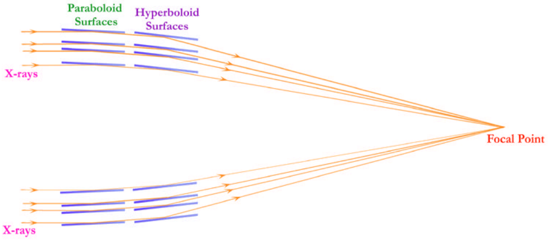
On the nature of High-Energy Astro data: Telescopes

X-rays and γ-rays are easily absorbed by matter, so can't have simple mirror telescopes. High-energy photons are brought to a focus via grazing reflection through a pairs of nested parabolic and hyperbolic surfaces.



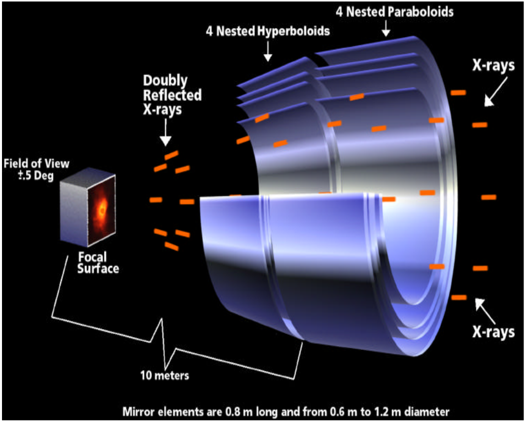
On the nature of High-Energy Astro data: Telescopes

X-rays and γ-rays are easily absorbed by matter, so can't have simple mirror telescopes. High-energy photons are brought to a focus via grazing reflection through a pairs of nested parabolic and hyperbolic surfaces.



On the nature of High-Energy Astro data: Telescopes

X-rays and γ-rays are easily absorbed by matter, so can't have simple mirror telescopes. High-energy photons are brought to a focus via grazing reflection through a pairs of nested parabolic and hyperbolic surfaces.



On the nature of High-Energy Astro data: Photons

- **High-energy photons are rare**
  - One X-ray photon at 10Å is worth 600 green photons at 6000Å
  - Counting rates are **LOW**: often about 1 per kilosecond, must watch for days to build up a decent signal

On the nature of High-Energy Astro data: Photons

- **High-energy photons are rare**
  - One X-ray photon at 10Å is worth 600 green photons at 6000Å
  - Counting rates are **LOW**: often about 1 per kilosecond, must watch for days to build up a decent signal
- Lemonade out of lemons
  - Every photon that comes in is treated as precious

On the nature of High-Energy Astro data: Photons

- **High-energy photons are rare**
  - One X-ray photon at 10Å is worth 600 green photons at 6000Å
  - Counting rates are **LOW**: often about 1 per kilosecond, must watch for days to build up a decent signal
- Lemonade out of lemons
  - Every photon that comes in is treated as precioussss
    (it gets put into a list)

On the nature of High-Energy Astro data: Photons

- **High-energy photons are rare**
  - One X-ray photon at 10Å is worth 600 green photons at 6000Å
  - Counting rates are **LOW**: often about 1 per kilosecond, must watch for days to build up a decent signal
- Lemonade out of lemons
  - Every photon that comes in is treated as precioussss
    (it gets put into a list)
  - What we have then is a *marked* Poisson process

On the nature of High-Energy Astro data: Photons

- **High-energy photons are rare**
  - One X-ray photon at 10Å is worth 600 green photons at 6000Å
  - Counting rates are **LOW**: often about 1 per kilosecond, must watch for days to build up a decent signal
- Lemonade out of lemons
  - Every photon that comes in is treated as precioussss (it gets put into a list)
  - What we have then is a *marked* Poisson process – in addition to time stamps, we also record location on detector and energy deposited in detector

On the nature of High-Energy Astro data: Photons

- **High-energy photons are rare**
  - One X-ray photon at 10Å is worth 600 green photons at 6000Å
  - Counting rates are **LOW**: often about 1 per kilosecond, must watch for days to build up a decent signal
- Lemonade out of lemons
  - Every photon that comes in is treated as precioussss
    (it gets put into a list)
  - What we have then is a *marked* Poisson process – in addition to time stamps, we also record location on detector and energy deposited in detector
  - Histograms made from arrival time information are *light curves*, from deposited-energy information are *spectra*, and from pixel location information are *images*

On the nature of High-Energy Astro data: Missions

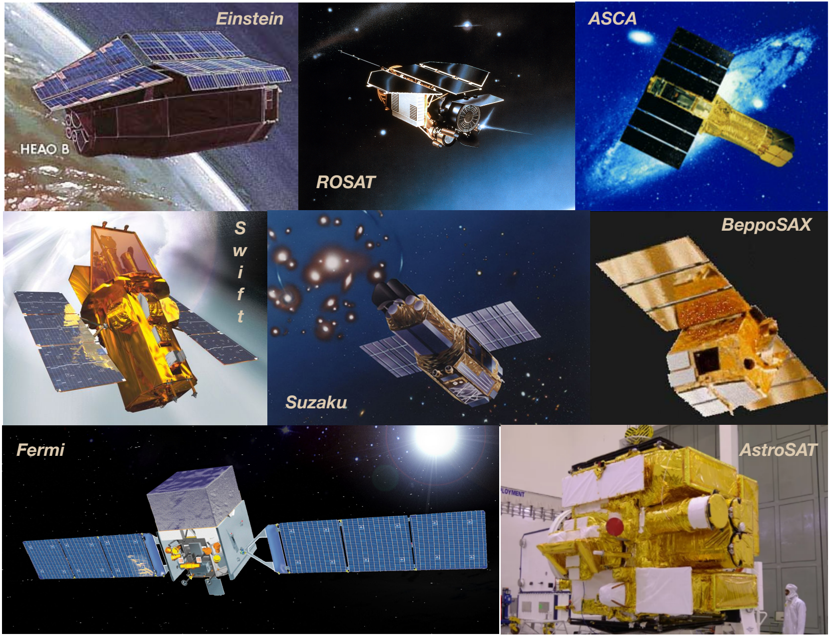
- Our atmosphere does a very good job of protecting us from high-energy radiation. So to view these photons, we must put our telescopes up above all the inconvenient air.

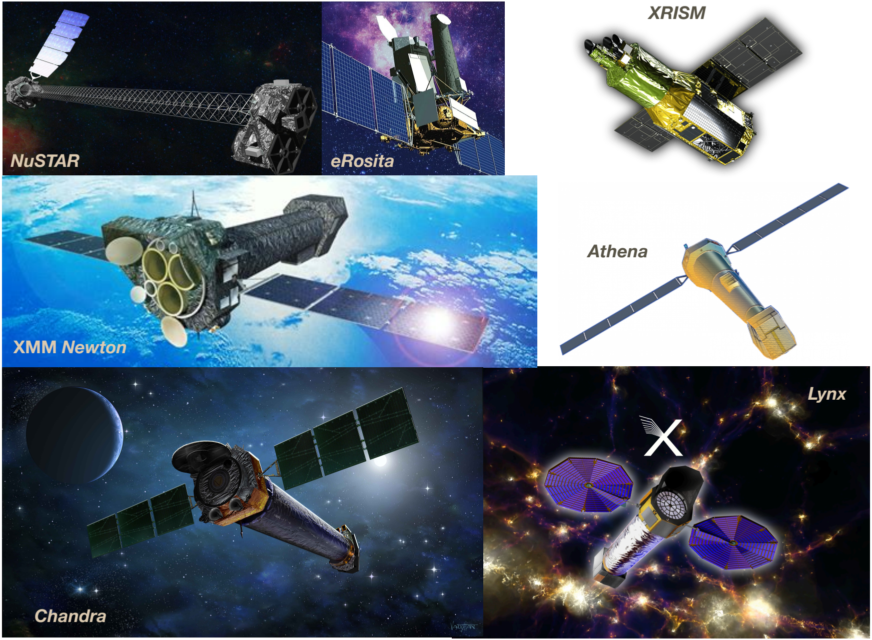
On the nature of High-Energy Astro data: Missions

- Our atmosphere does a very good job of protecting us from high-energy radiation. So to view these photons, we must put our telescopes up above all the inconvenient air.
- X-ray astronomy started in the ′60s with a failed attempt to detect X-rays from the Moon, and got Riccardo Giacconi a Physics Nobel.

On the nature of High-Energy Astro data: Missions

- Our atmosphere does a very good job of protecting us from high-energy radiation. So to view these photons, we must put our telescopes up above all the inconvenient air.
- X-ray astronomy started in the ′60s with a failed attempt to detect X-rays from the Moon, and got Riccardo Giacconi a Physics Nobel.
- Imaging X-ray astronomy started with the *Einstein* X-ray Observatory


Einstein
HEAO B
ROSAT
ASCA
Swift
Suzaku
BeppoSAX
Fermi
AstroSAT


NuSTAR
eRosita
XRISM
XMM Newton
Athena
Chandra
Lynx

On the nature of High-Energy Astro data: Response

- Every mission is unique, and has custom-built mirrors, mirror coatings, and photon detectors, and blocking filters

On the nature of High-Energy Astro data: Response

- Every mission is unique, and has custom-built mirrors, mirror coatings, and photon detectors, and blocking filters

- The efficiency with which a photon of a given energy E is detected varies with E

On the nature of High-Energy Astro data: Response

- Every mission is unique, and has custom-built mirrors, mirror coatings, and photon detectors, and blocking filters
- The efficiency with which a photon of a given energy E is detected varies with E, is different for different instruments

On the nature of High-Energy Astro data: Response

- Every mission is unique, and has custom-built mirrors, mirror coatings, and photon detectors, and blocking filters
- The efficiency with which a photon of a given energy E is detected varies with E, is different for different instruments, and often varies with age for a given instrument

On the nature of High-Energy Astro data: Response

- Every mission is unique, and has custom-built mirrors, mirror coatings, and photon detectors, and blocking filters

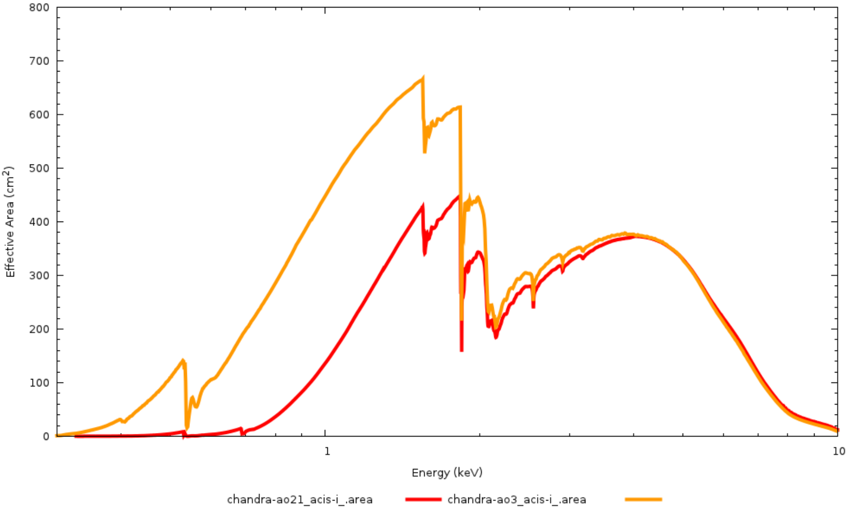
- The efficiency with which a photon of a given energy E is detected varies with E, is different for different instruments, and often varies with age for a given instrument (and sometimes decreases for bright objects)

On the nature of High-Energy Astro data: Response

- Every mission is unique, and has custom-built mirrors, mirror coatings, and photon detectors, and blocking filters
- The efficiency with which a photon of a given energy E is detected varies with E, is different for different instruments, and often varies with age for a given instrument (and sometimes decreases for bright objects)
- This efficiency is encoded and stored as an area [cm^2], called *Effective Area*, A_{eff}
- Maximum theoretical efficiency is 100% of the geometric aperture area, and is *never* achievable in practice.

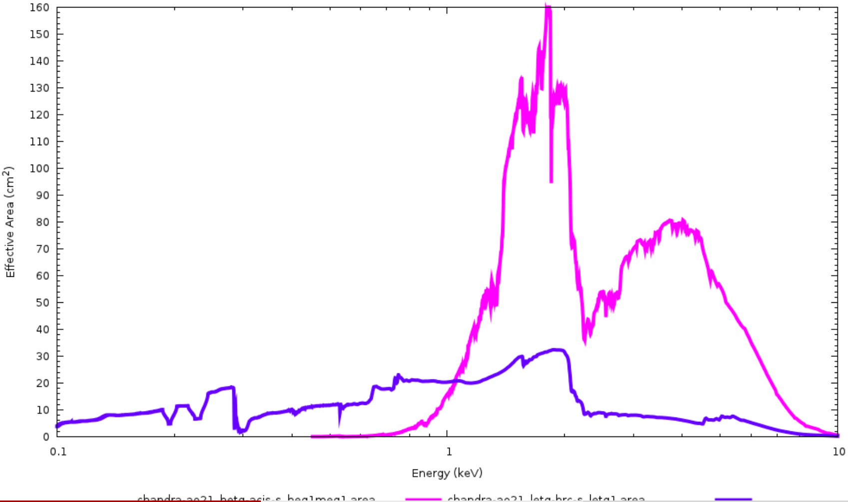
On the nature of High-Energy Astro data: Example A_{eff}

Chandra/ACIS-I: current vs 2001



On the nature of High-Energy Astro data: Example Example A_{eff}

Chandra gratings: LETGS+HRC-S vs HETGS+ACIS-S



On the nature of High-Energy Astro data: A_{eff}

- The correct estimation of the Effective Area is the most critical part of any astronomical mission, because this is what enables us to translate the electrical signals coming out of the instruments into physically meaningful quantities

On the nature of High-Energy Astro data: $A_{\rm eff}$

- The correct estimation of the Effective Area is the most critical part of any astronomical mission, because this is what enables us to translate the electrical signals coming out of the instruments into physically meaningful quantities
- The Effective Area is measured on the ground pre-flight, and monitored during flight

On the nature of High-Energy Astro data: A_{eff}

- The correct estimation of the Effective Area is the most critical part of any astronomical mission, because this is what enables us to translate the electrical signals coming out of the instruments into physically meaningful quantities
- The Effective Area is measured on the ground pre-flight, and monitored during flight
- Our best estimates of A_{eff} are written into a file and shared with all observers, who use it to infer the properties of the astrophysical source they are investigating

On the nature of High-Energy Astro data: FEq

The fundamental equation of observational astronomy

All of astronomy, until very recently, has been about inferring, from the observed photons, the nature of the object that is emitting the photons. Spectra are the biggest hammer in our toolbox.

Spectra are recorded as digitized amplifier signals over a range of *channels*.

The expected counts in digitized detector channels $i : j$,

$$\lambda_{i:j} = \tau \sum_{k=i}^{j} \int dE \ \{A_{\rm eff}(E) \cdot f(E; \mathbf{\Theta})\} \ p(k|E)$$

$>$ $A_{\rm eff}(E)$ [cm^2] is the effective area at photon energy E [keV],
$>$ τ [sec] is the duration of the observation,
$>$ $p(k|E)$ is the probability of registering the photon in detector channel k,
$>$ and $f(E; \mathbf{\Theta})$ [photons cm^{-2} sec^{-1} keV^{-1}] is the astrophysical model spectrum defined at each E
$>$ via astrophysical parameters $\mathbf{\Theta}$

On the nature of High-Energy Astro data: the Dirty Secret

On the nature of High-Energy Astro data: the Dirty Secret

Everybody assumes that $A_{\rm eff}$ is known exactly.

On the nature of High-Energy Astro data: the Dirty Secret

Everybody assumes that A_{eff} is known exactly. It is not.

On the nature of High-Energy Astro data: the Dirty Secret

Everybody assumes that A_{eff} is known exactly. It is not.

Astrostatistics to the rescue!

Xiao-Li will next discuss the consequences of this problem, and show how we are dealing with it.