23 February 2023, *The 1st XRISM Data Analysis Workshop, University of Maryland College Park (USA)*

Analyzing (and proposing) extended sources with XRISM

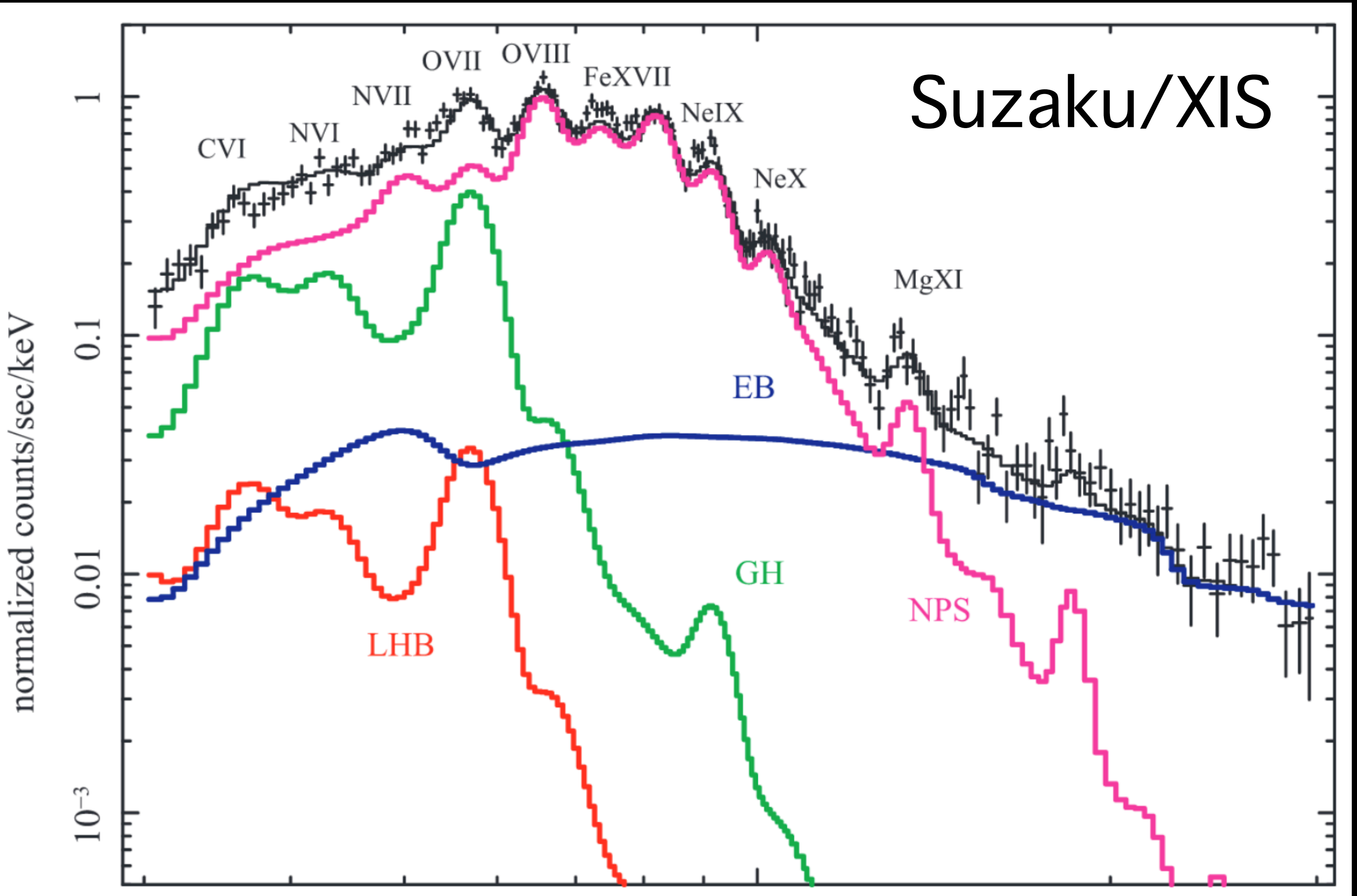
François Mernier

fmernier@umd.edu

The basics of analyzing extended sources

- ✓ **The emission is extended… but (often) faint!**
  - ✓ If not properly accounted for, background may bias the results

- ✓ Background subtraction is good…
- ✓…Background **modelling** is (much) better! (see Suzuki-san's talk)



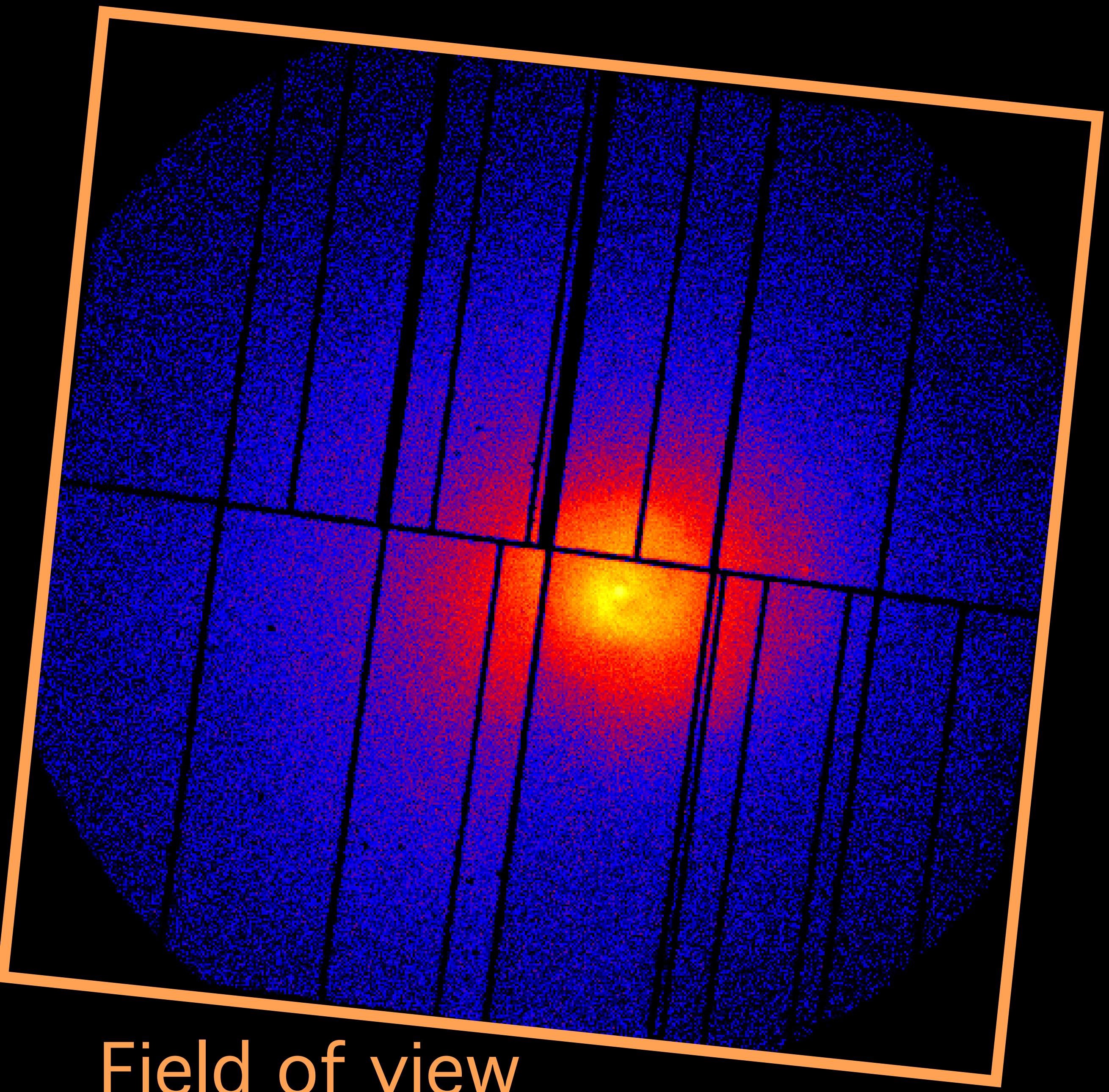
Miller et al. (2007)

- ✓ Your favorite (extended) source
- ✓ Local Hot Bubble (LHB)
- ✓ Galactic Halo (GH)
- ✓ Extragalactic Background (EB)

- ✓ Hard particle background (e.g. XMM-Newton)
- ✓ Soft proton background (e.g. XMM-Newton)

The problem

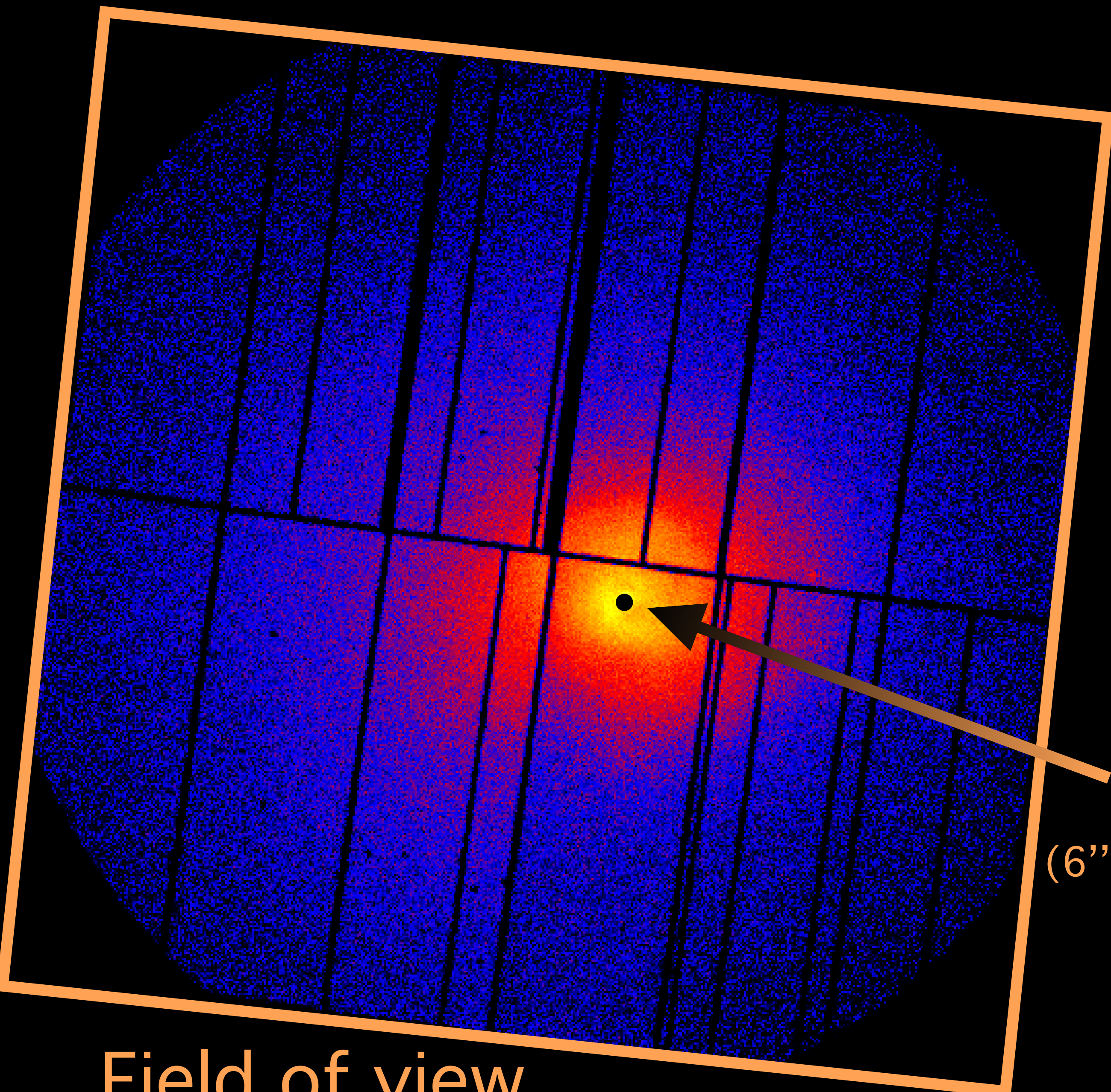
Point spread function (PSF) and field of view



Field of view
(25’ x 25’ total equivalent)

XMM-Newton

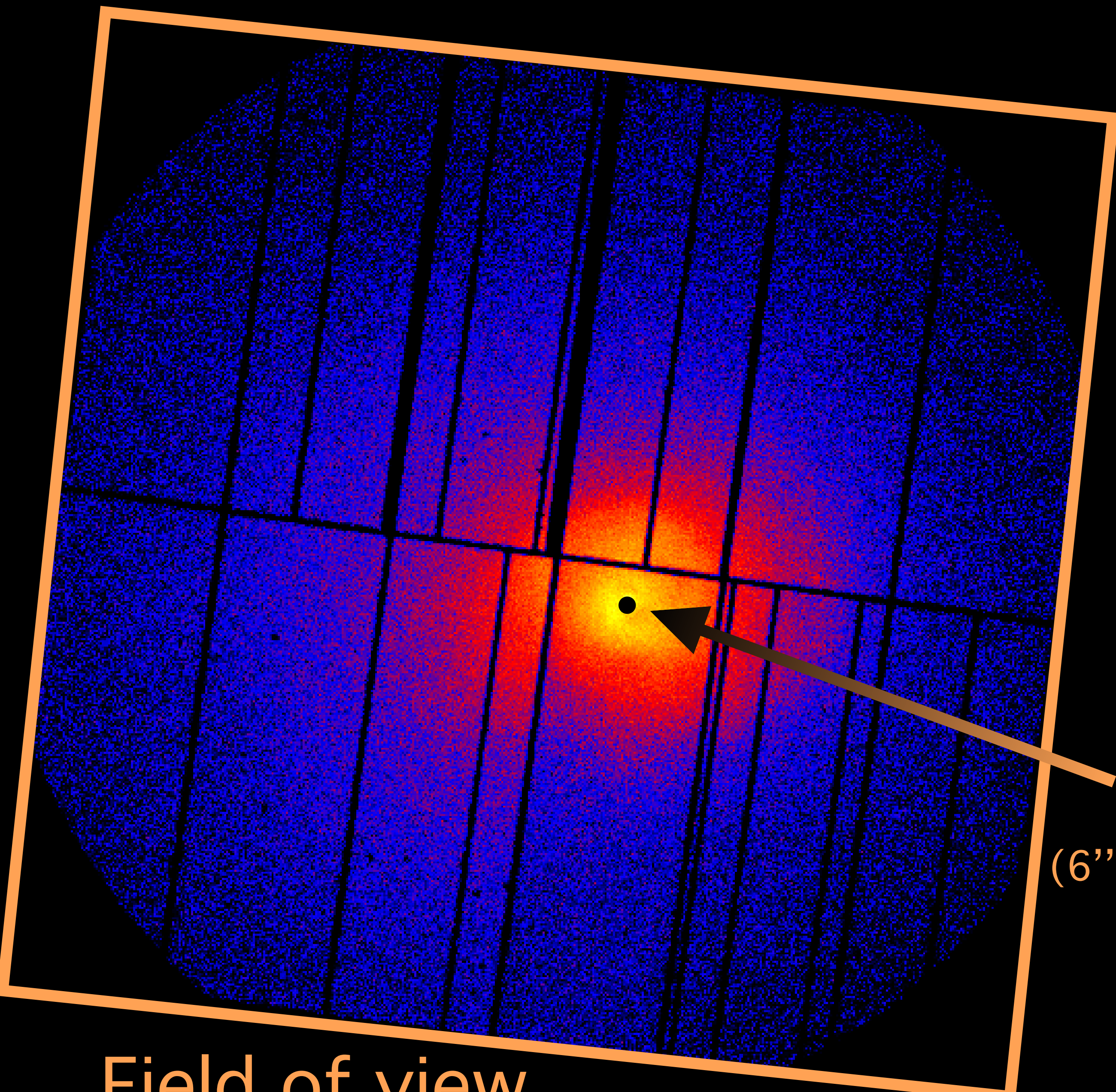
Point spread function (PSF) and field of view



$$\frac{6''}{25'} = 0.4\%$$

XMM-Newton

Point spread function (PSF) and field of view

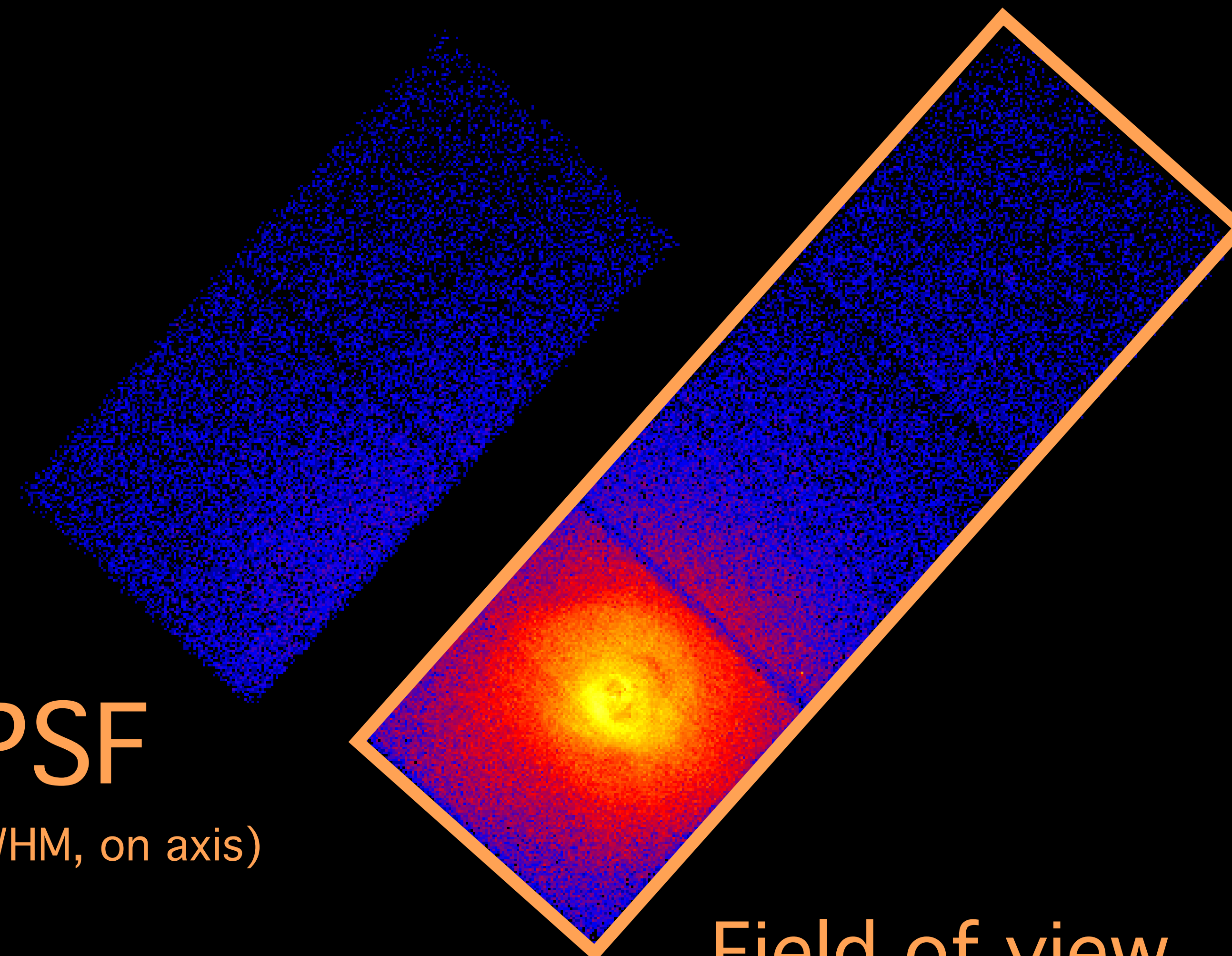


PSF
(6'' FWHM, on axis)

Field of view
(25' x 25' total equivalent)

$$\frac{6''}{25'} = 0.4\%$$

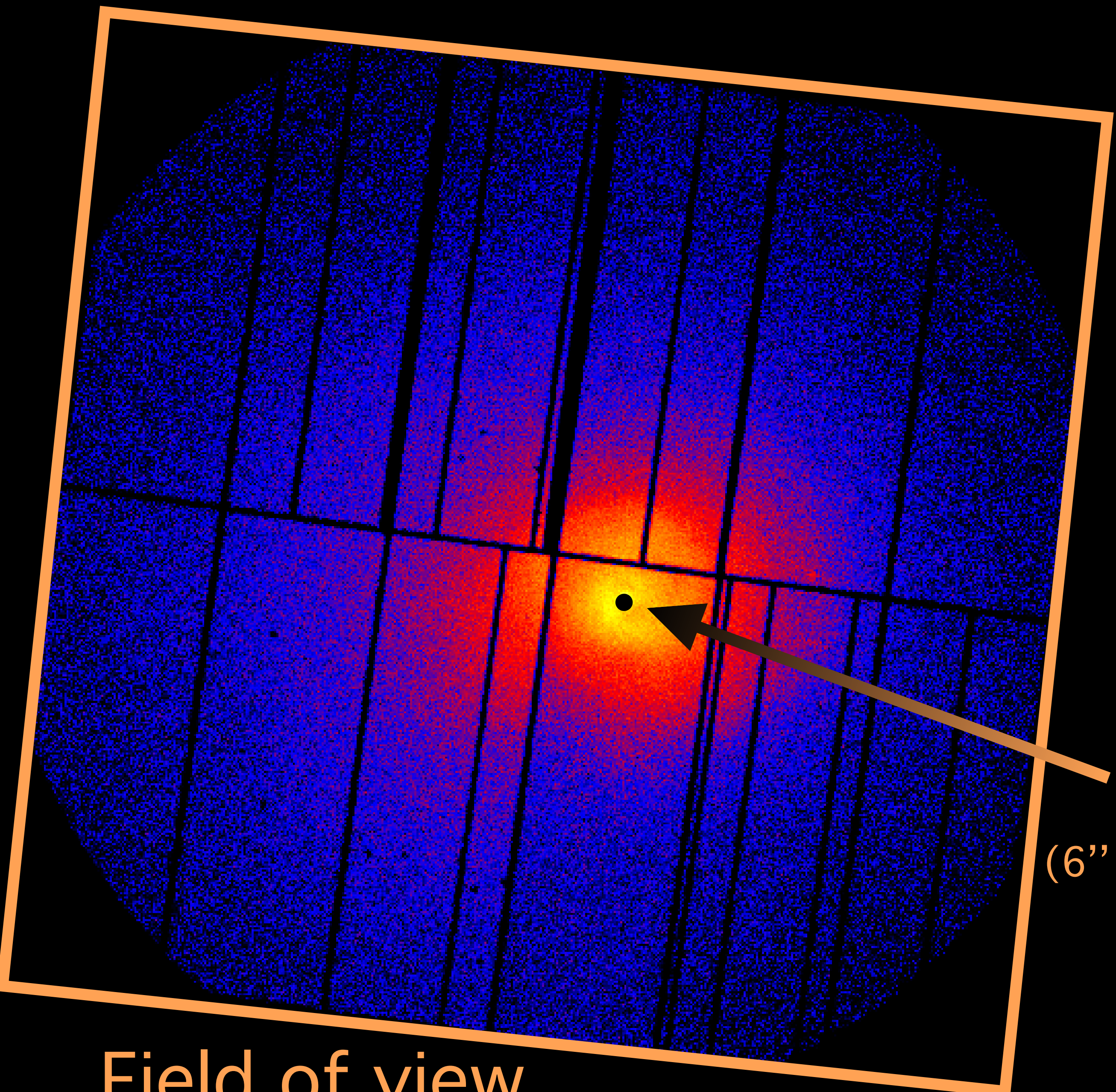
XMM-Newton



Field of view
(8.5' x 8.5' per chip)

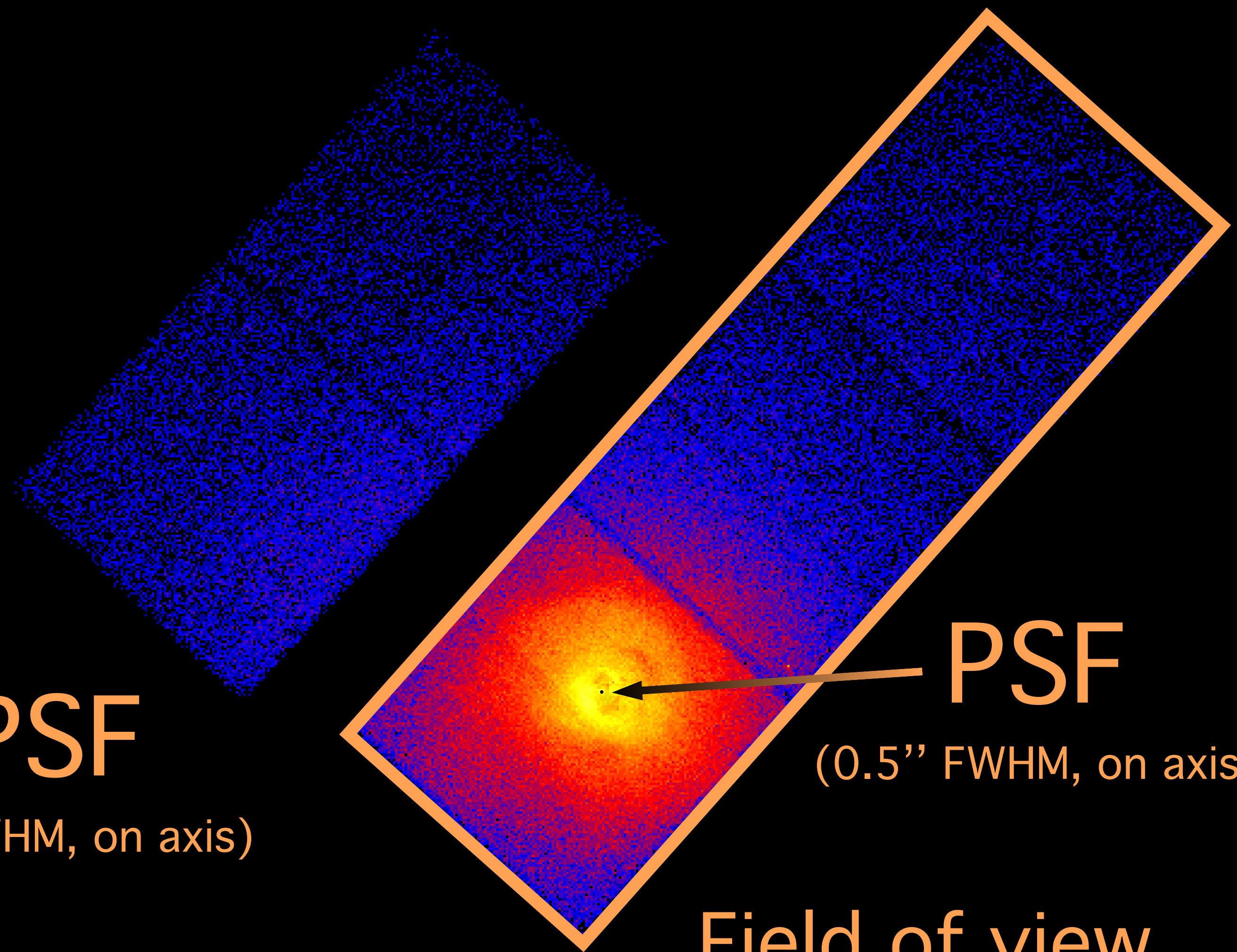
Chandra

Point spread function (PSF) and field of view



$$\frac{6''}{25'} = 0.4\%$$

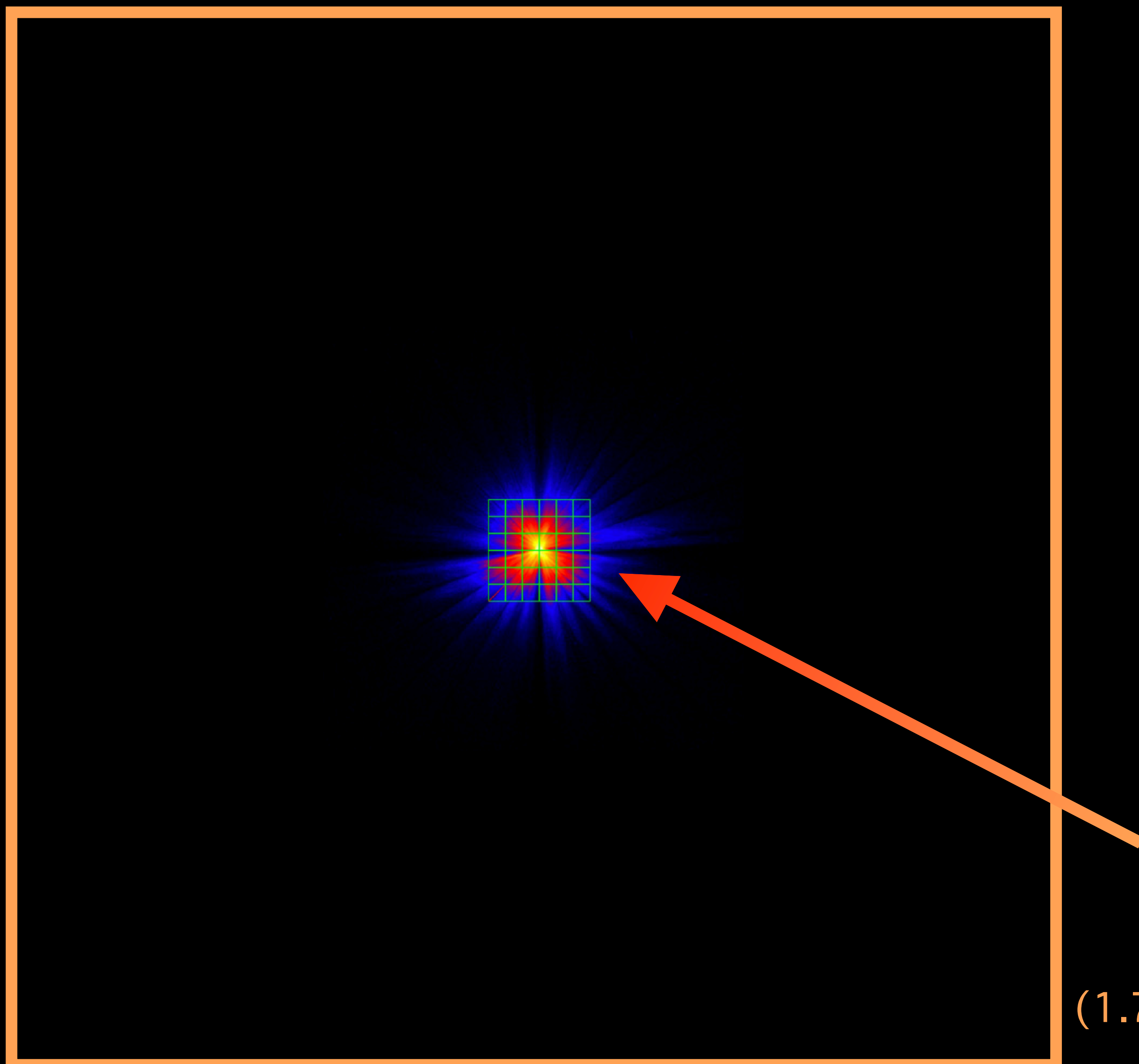
XMM-Newton



$$\frac{0.5''}{8.5'} = 0.1\%$$

Chandra

Point spread function (PSF) and field of view



PSF

(1.7’ HPD, on axis)

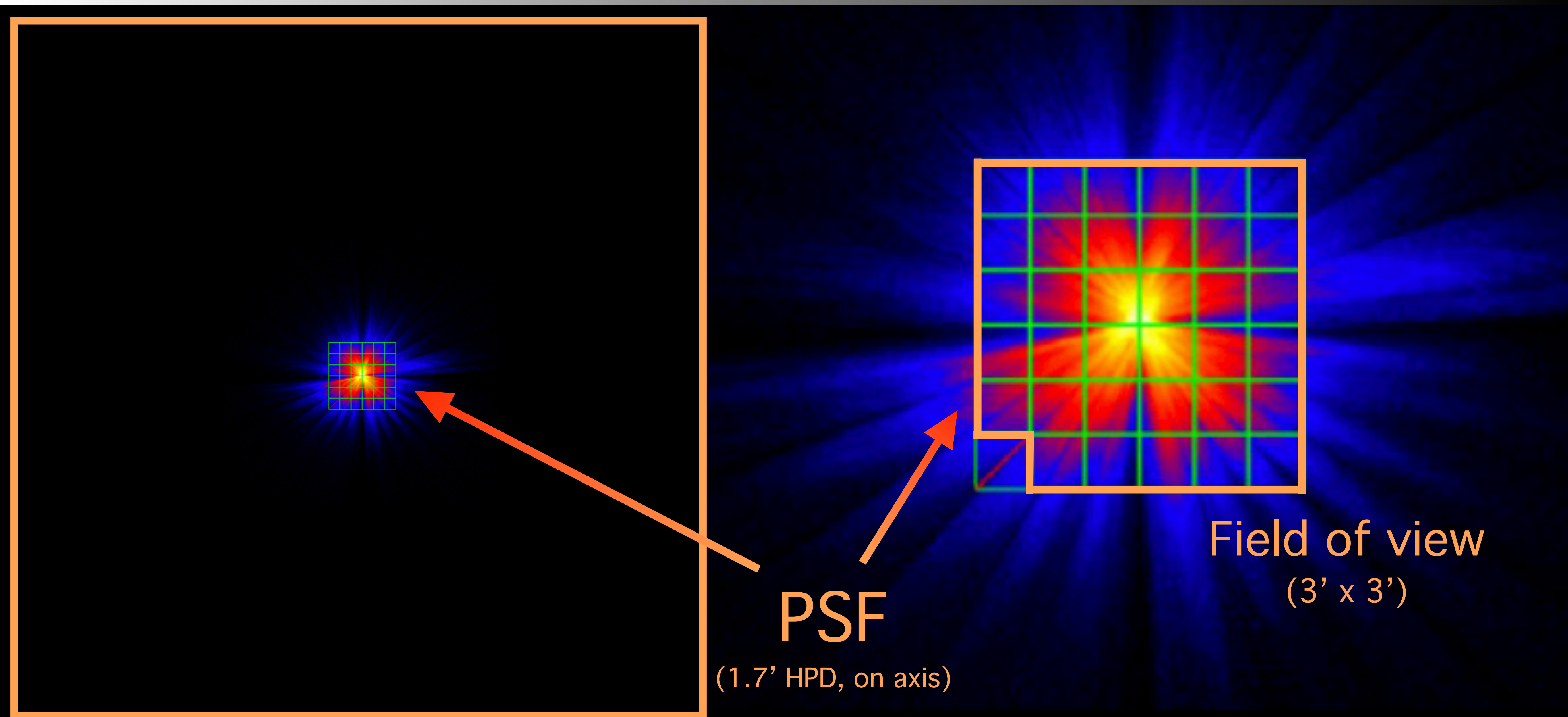
Field of view

(38’ x 38’)

$$\frac{1.7'}{38'} = 4\%$$

XRISM / Xtend

Point spread function (PSF) and field of view



Field of view
(38' x 38')

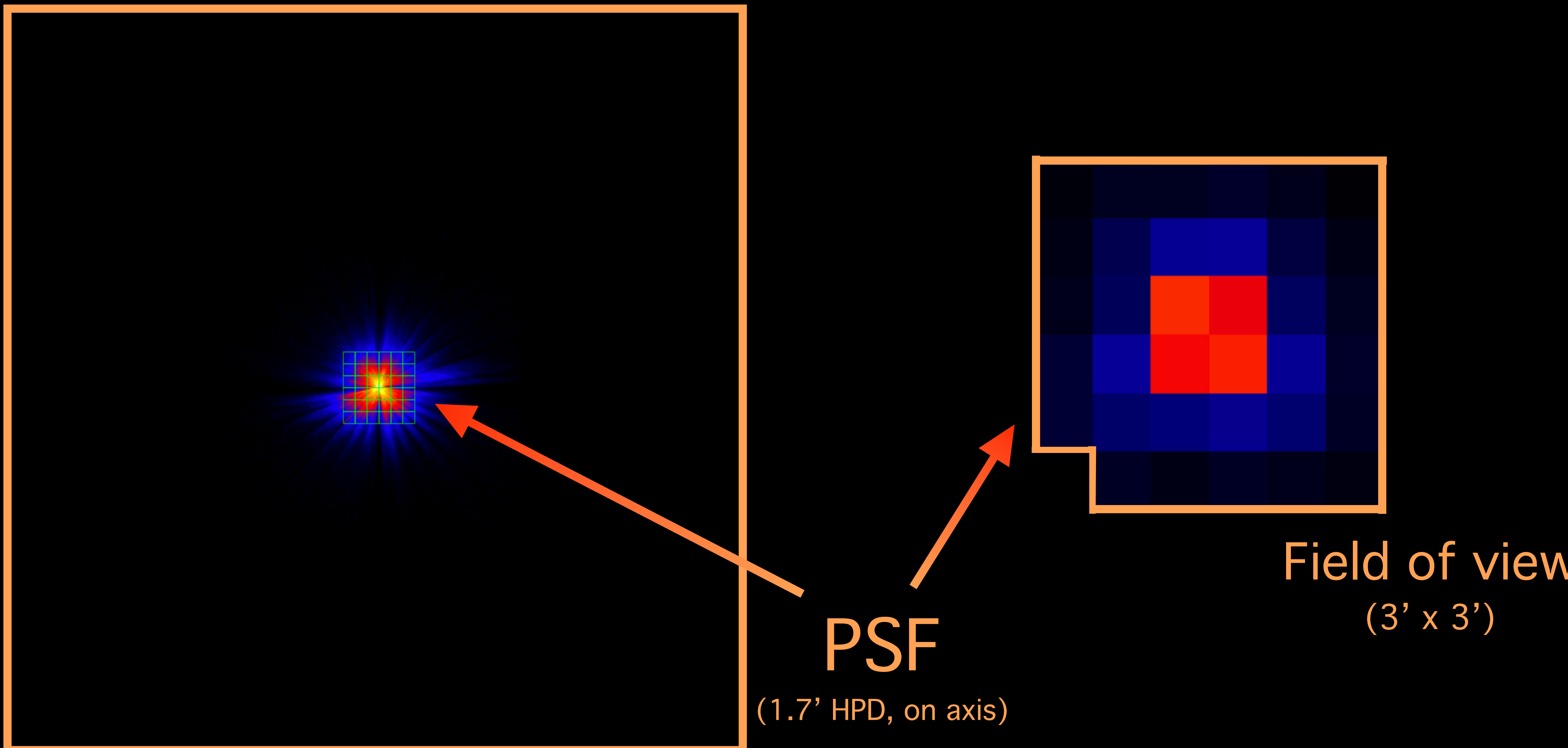
$$\frac{1.7'}{38'} = 4\%$$

$$\frac{1.7'}{3'} = 57\% \ (!)$$

XRISM / Xtend

XRISM / Resolve

Point spread function (PSF) and field of view



Field of view
(30’ x 30’)

$$\frac{1.7'}{38'} = 4\%$$

$$\frac{1.7'}{3'} = 57\%\ (!)$$

XRISM / Xtend

XRISM / Resolve

A bit of nomenclature…

✓ **"Contamination"?**

✓ Need to specify what contaminates what… Moreover, it sounds quite negative (whereas substantial "contamination" is not necessarily bad… see further)

✓ **"Photon leak"?**

✓ Can a large PSF stricto sensu considered as a "leak"?

✓ **"Mixing"?**

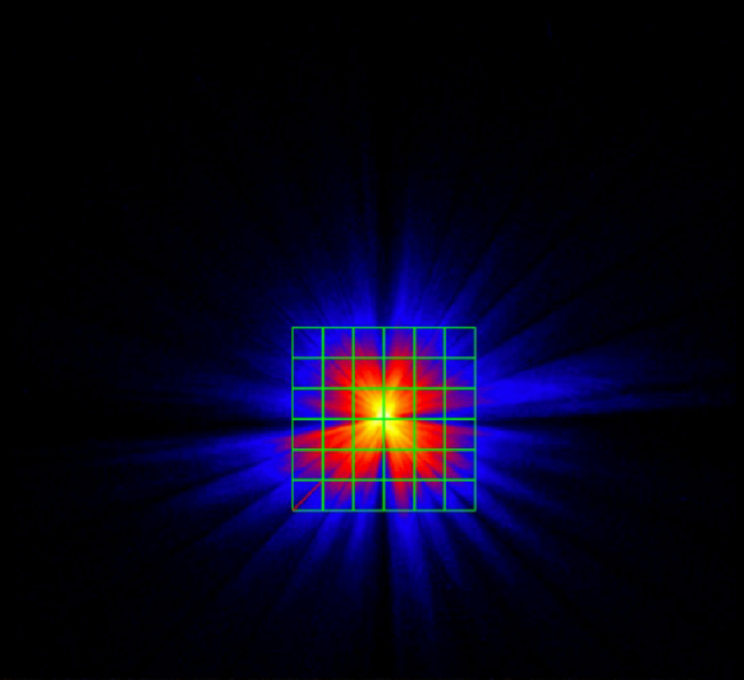
✓ More accurate term (and takes multi-directionality into account) ✅

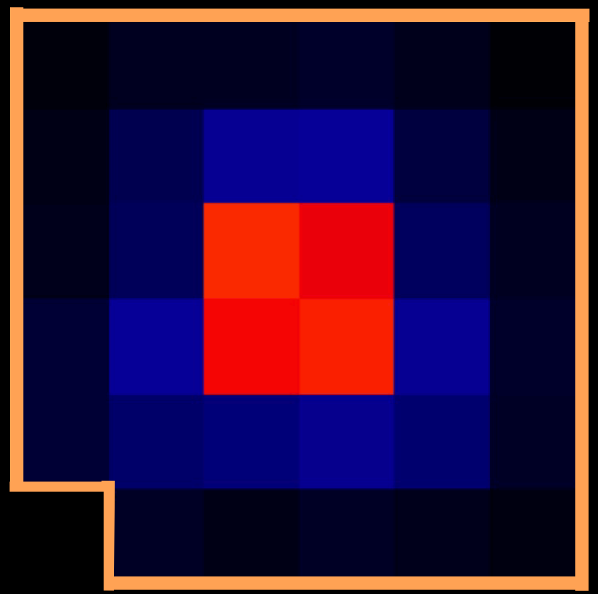
✓ XRISM PSF will mix the data **spatially**… but also **spectrally**!

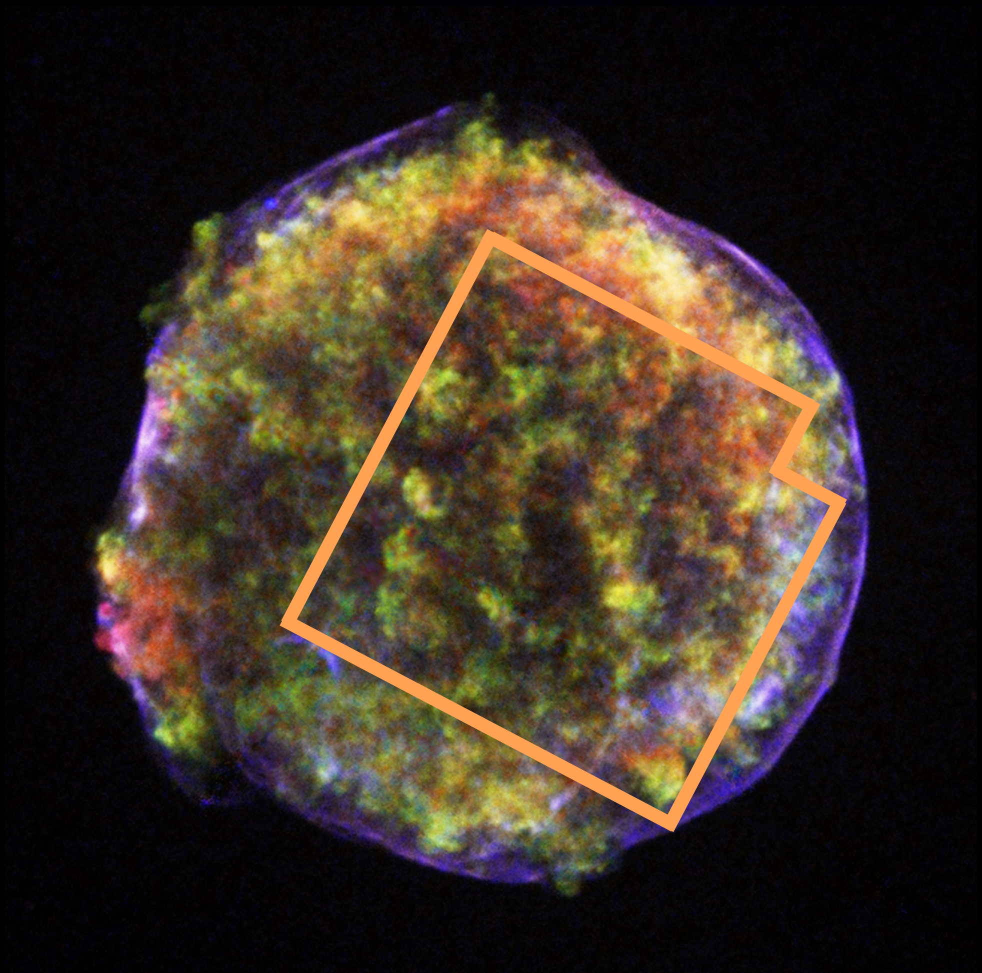
➡ **Spatial-spectral Mixing (SSM)**

✓ The SSM will affect mostly Resolve (…but it may be also substantial in Xtend, depending on the science case / observing strategy)

Spatial-Spectral Mixing: internal vs. external





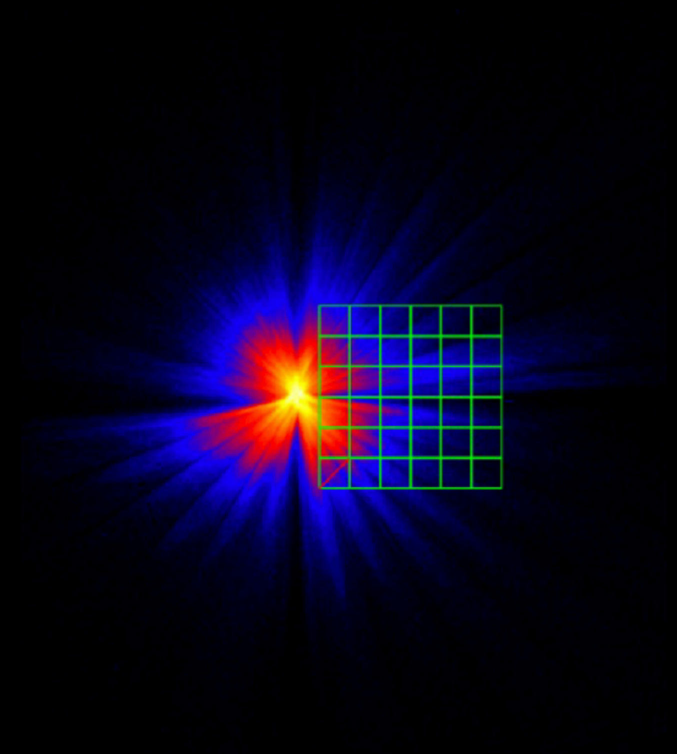


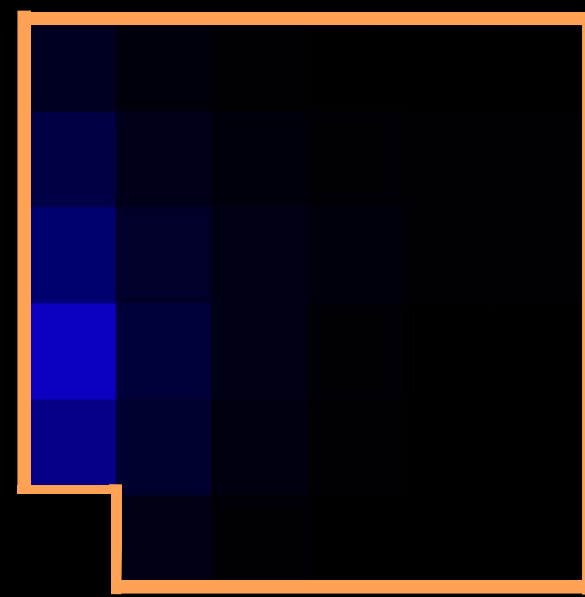
Internal SSM:

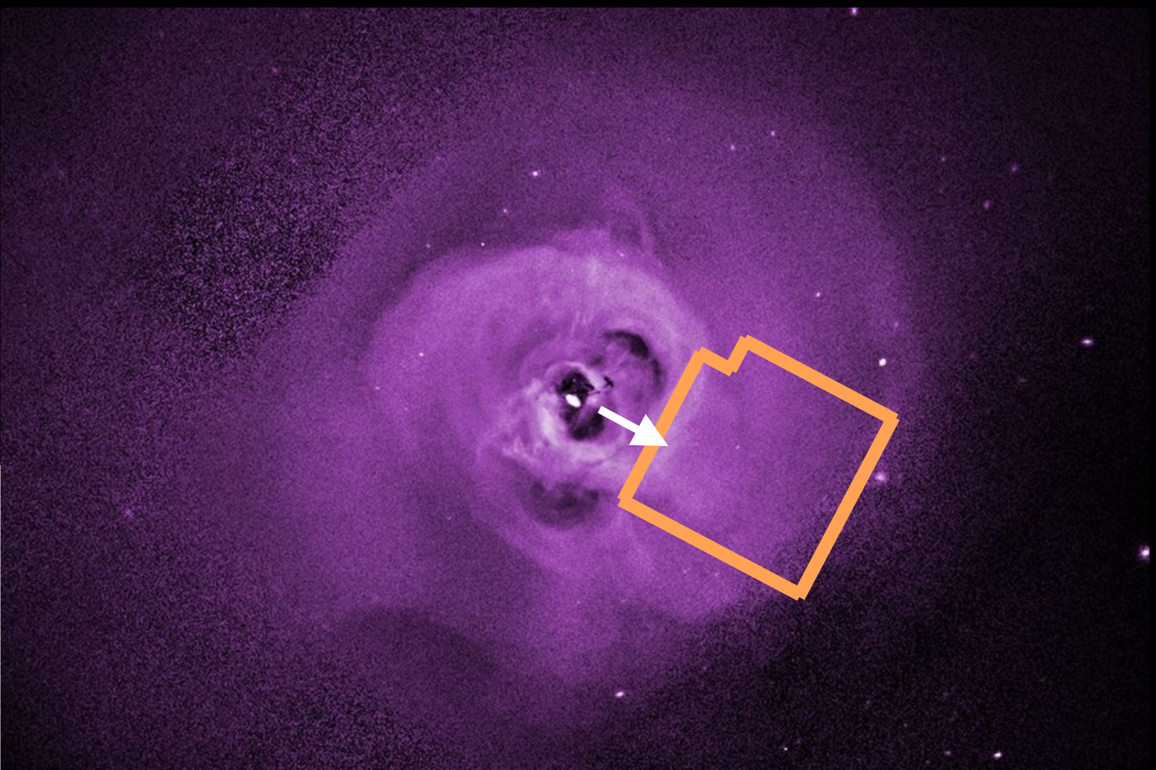
✓ **When the emission within the same field of view mixes across pixels and "contaminate" other regions within the same field of view**

✓ Examples: supernova remnants, complex star-forming regions, clusters with cavities, etc.

Spatial-Spectral Mixing: internal vs. external



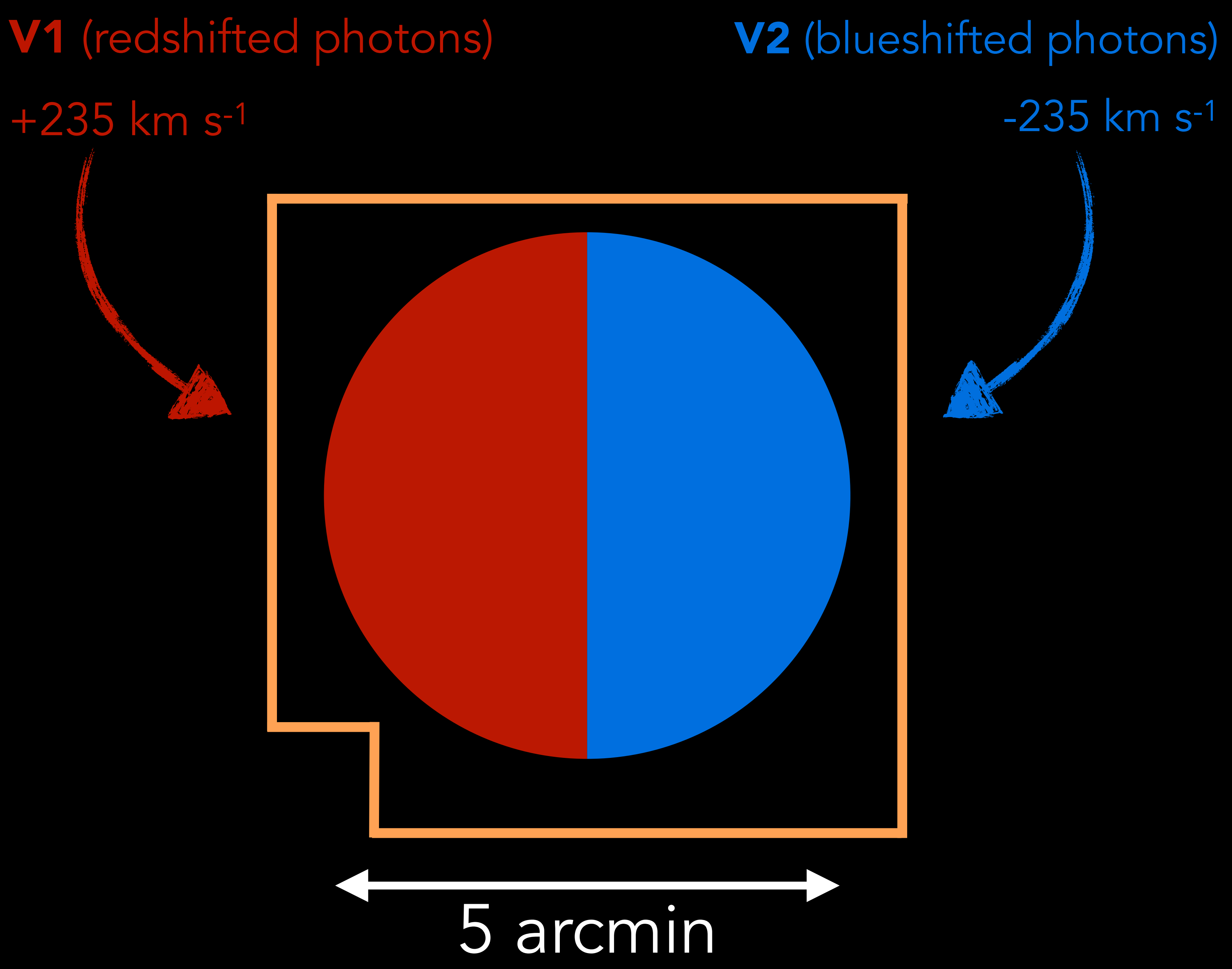




External SSM:

- ✓ **When sources outside the detector contaminate the detector region istelf**
- ✓ Examples: outskirts of a cool-core cluster, bright AGN nearby an extended source, etc.

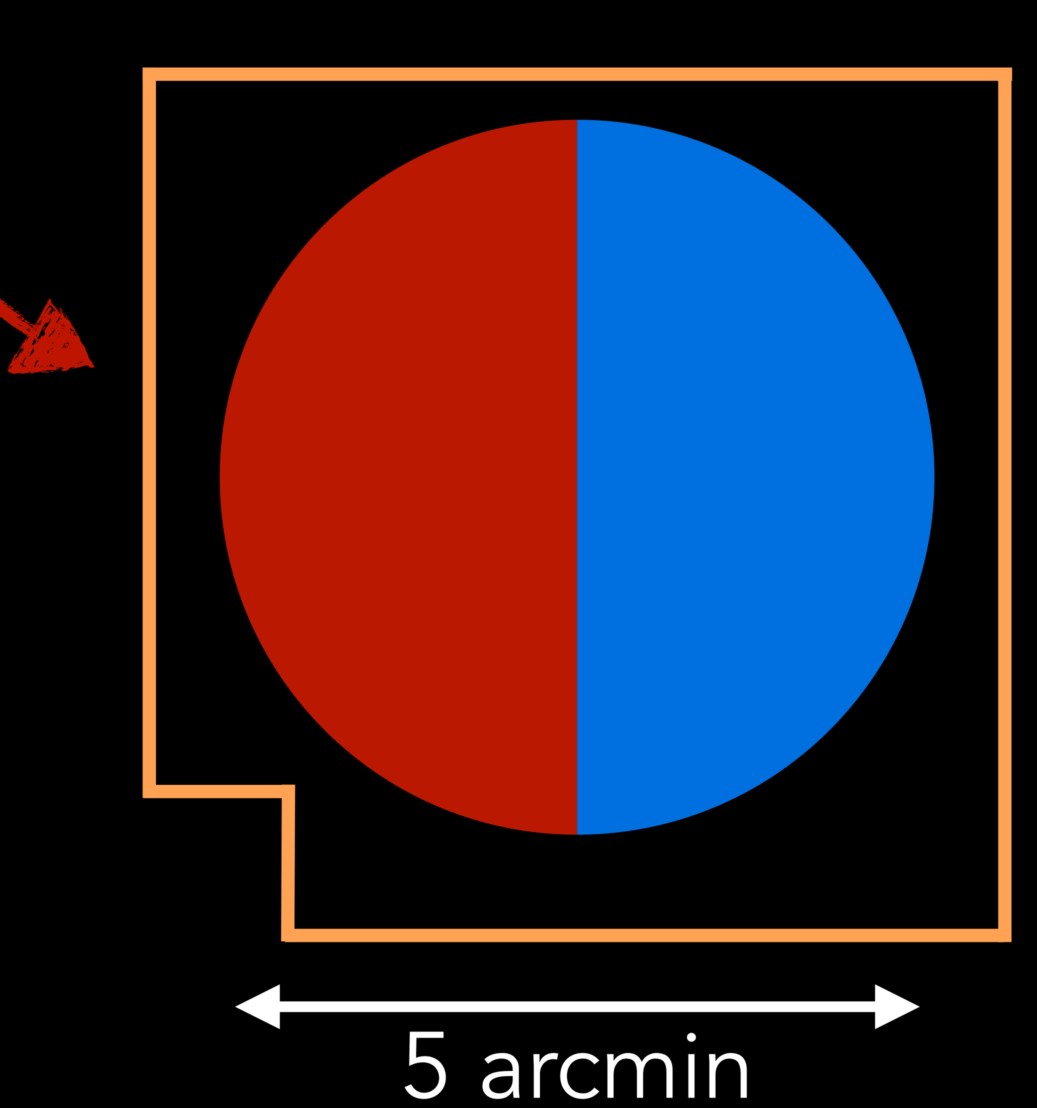
A (very) simple example…

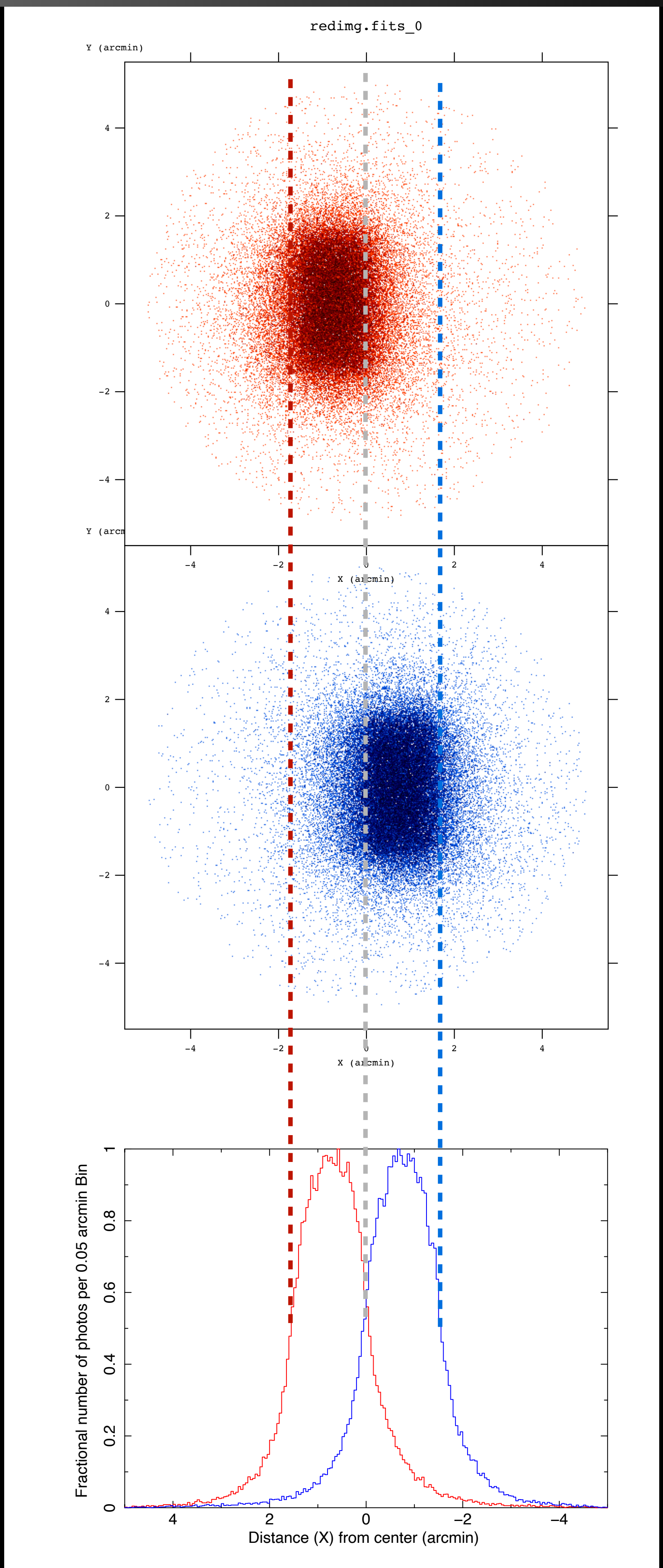


A (very) simple example…

V1 (redshifted photons) | **V2** (blueshifted photons)

+235 km s^{-1} | -235 km s^{-1}

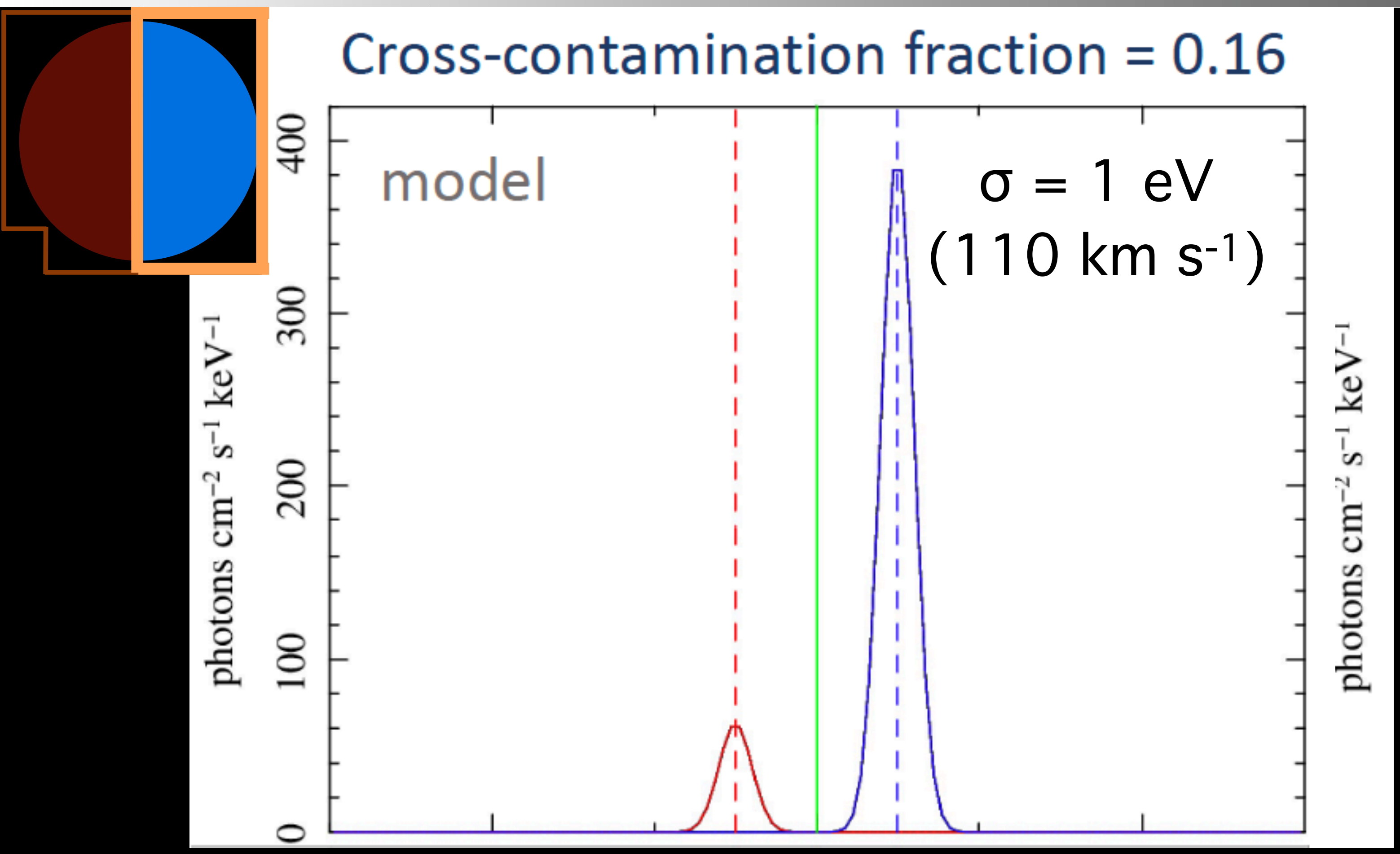




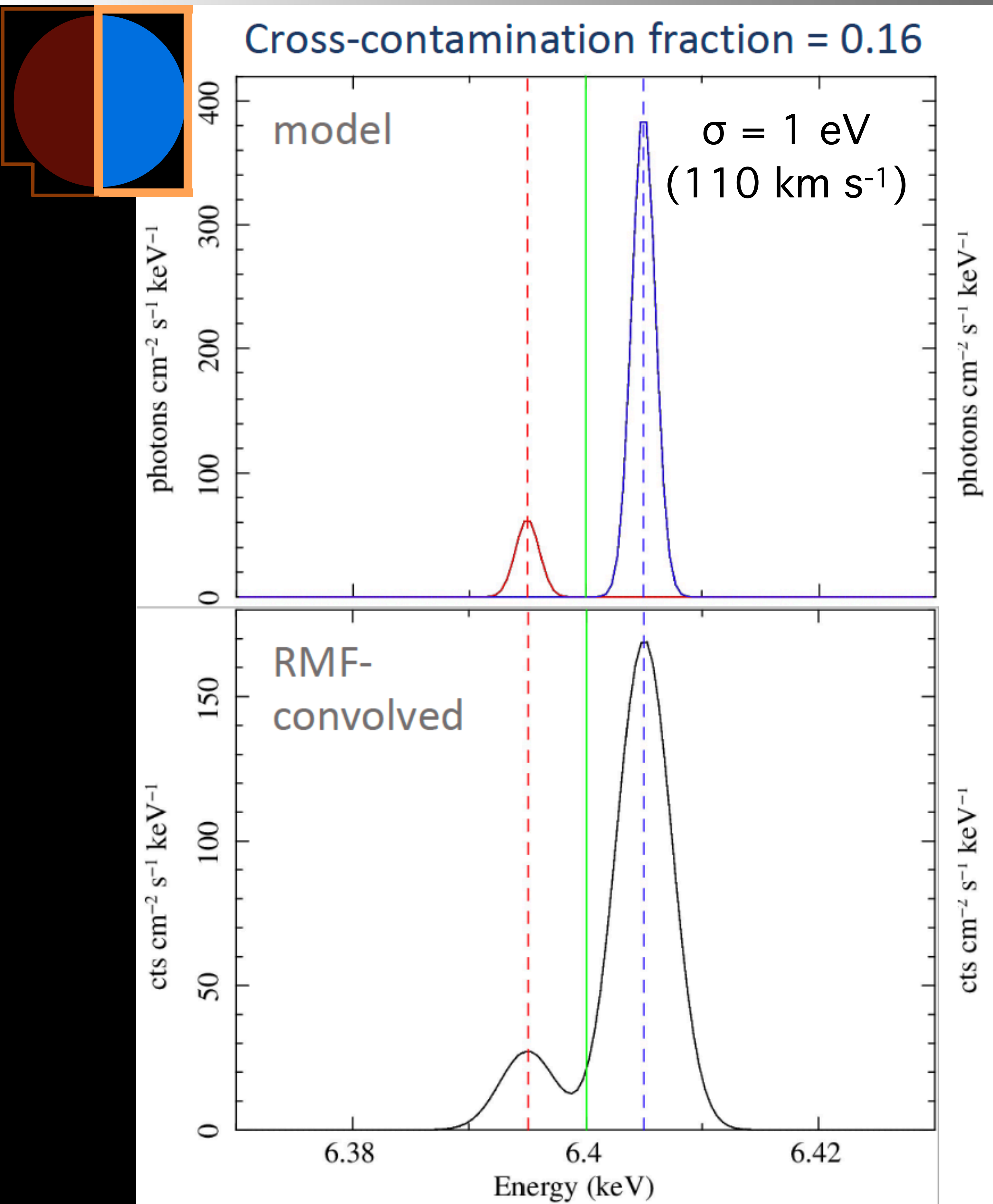
Spatial mixing…

- Fraction of V2 photons mixing into the **V1** region: 15%
- Fraction of V1 photons mixing into the **V2** region: 17%

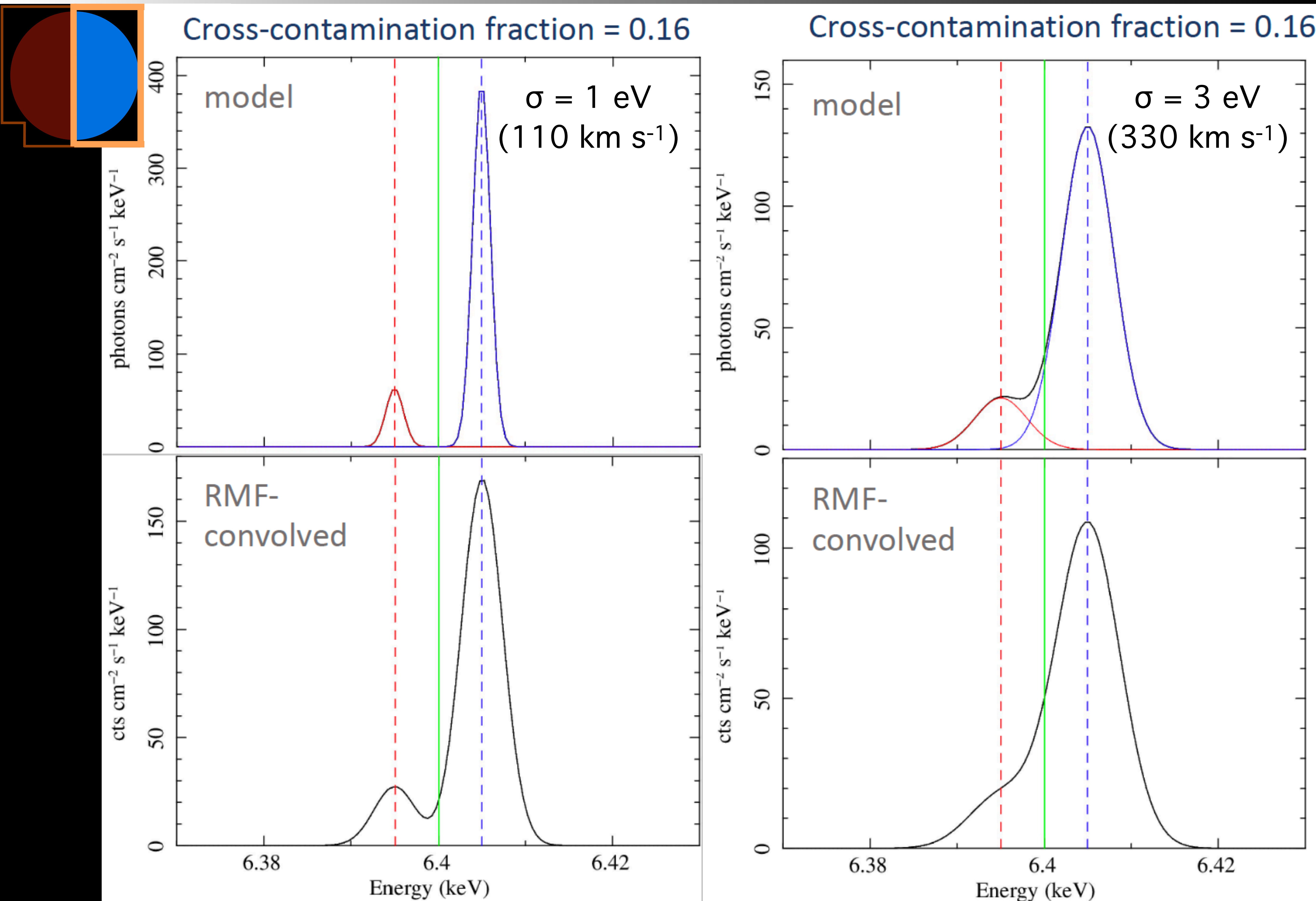
A (very) simple example…



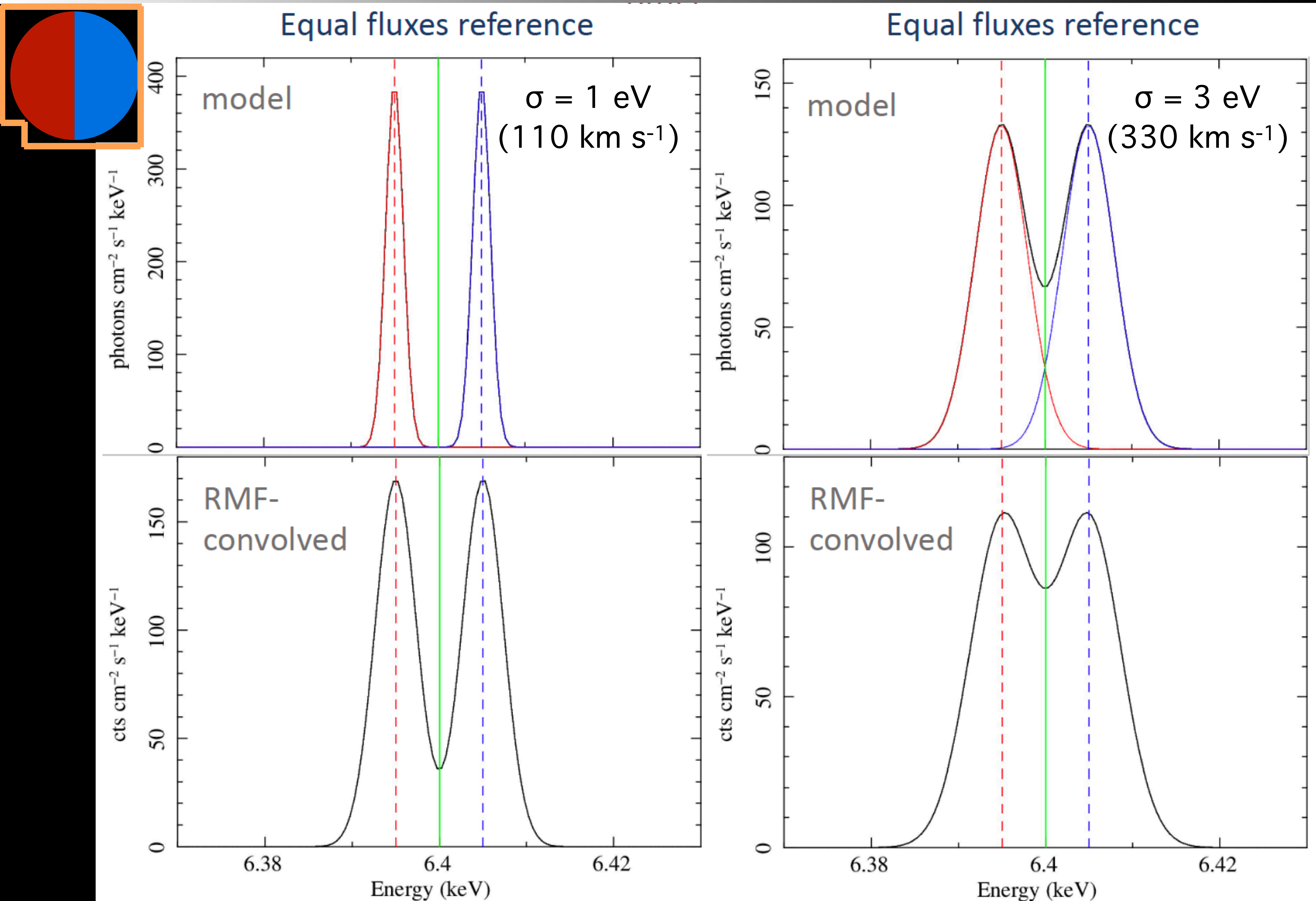
A (very) simple example…



A (very) simple example…



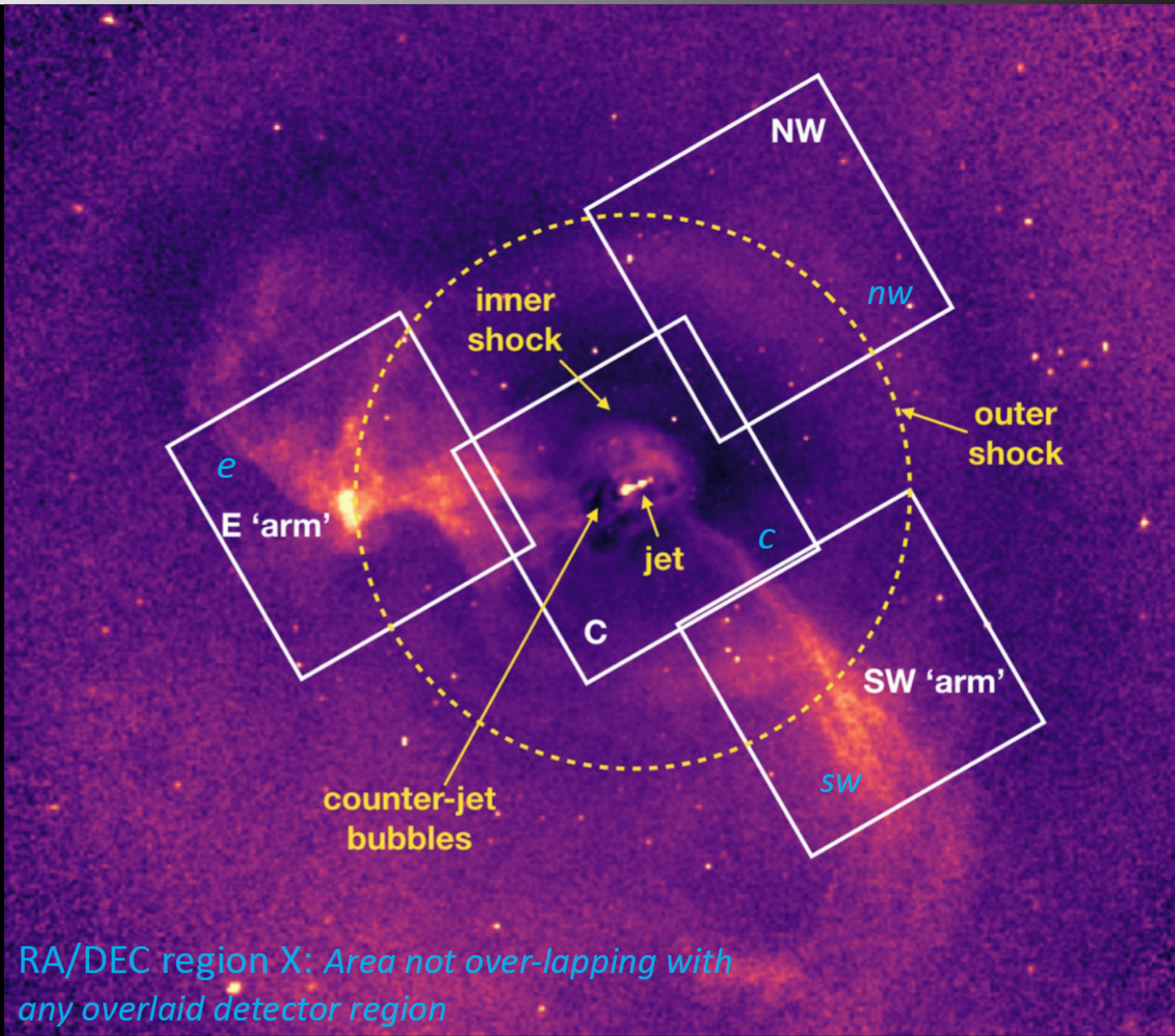
A (very) simple example...



Addressing the problem

Ready to fall into the rabbit hole…?

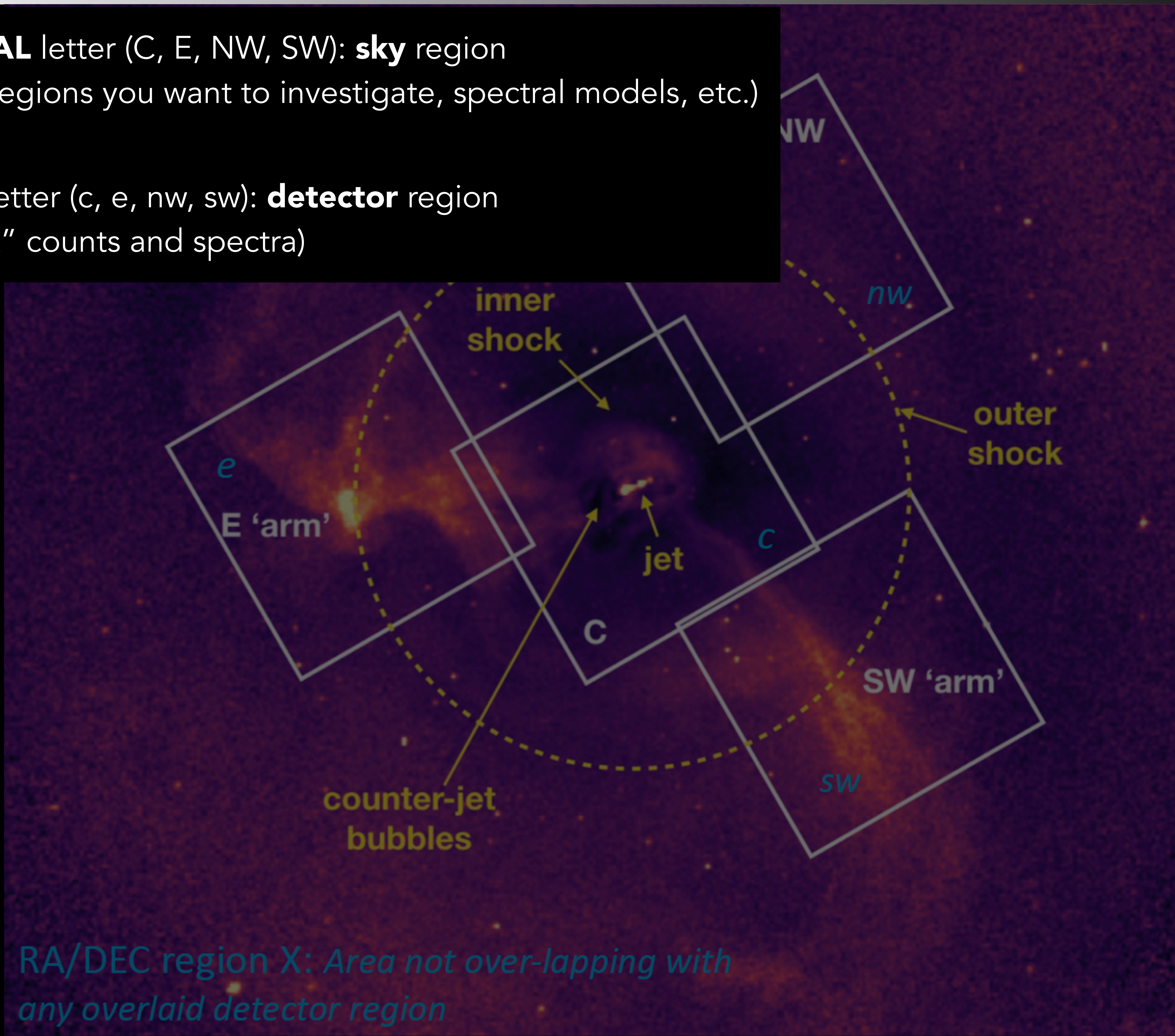
A concrete case (Virgo cluster)



A concrete case (Virgo cluster)

- ✓ **CAPITAL** letter (C, E, NW, SW): **sky** region
  (“true” regions you want to investigate, spectral models, etc.)

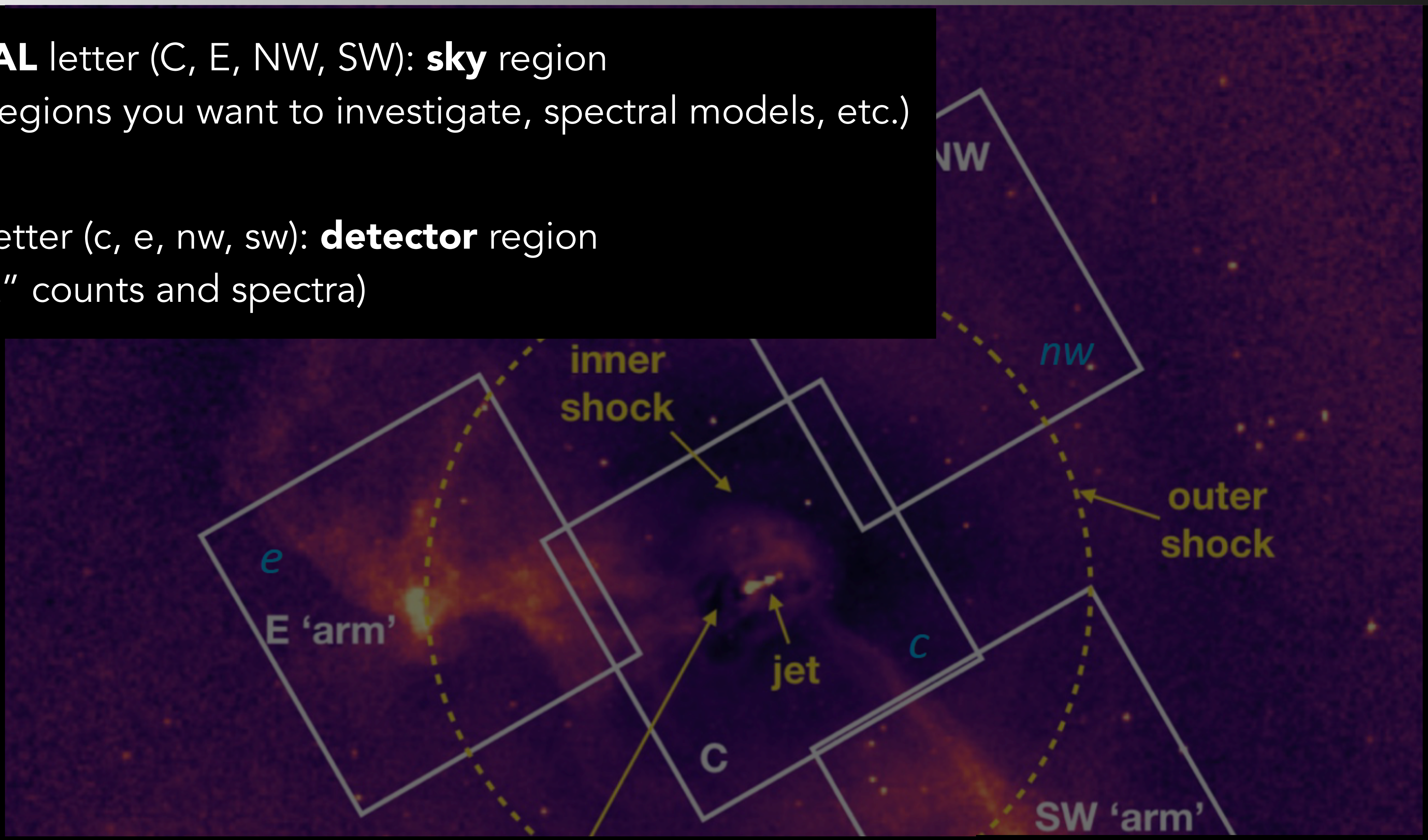
- ✓ **small** letter (c, e, nw, sw): **detector** region
  (“output” counts and spectra)



A concrete case (Virgo cluster)

✓ **CAPITAL** letter (C, E, NW, SW): **sky** region
("true" regions you want to investigate, spectral models, etc.)

✓ **small** letter (c, e, nw, sw): **detector** region
("output" counts and spectra)

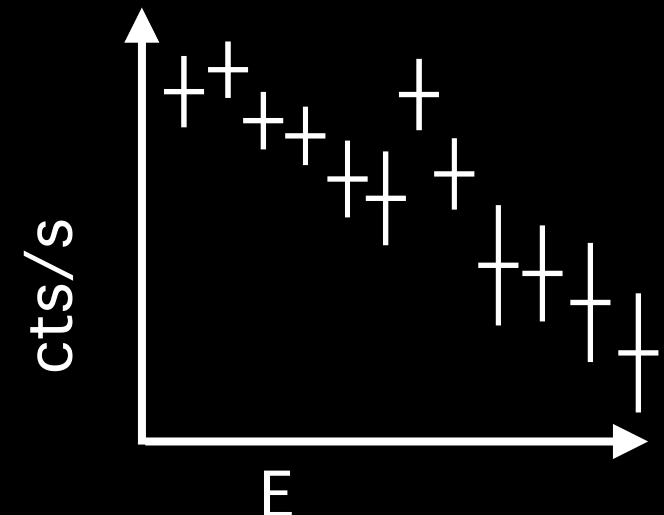


✓ $\mathbf{S_i}$ = observed ***spectrum*** of detector region i

✓ $\mathbf{M_J}$ = ***spectral model*** of sky region J

✓ $\mathbf{R_i}$ = ***response matrix*** (RMF) of detector region i

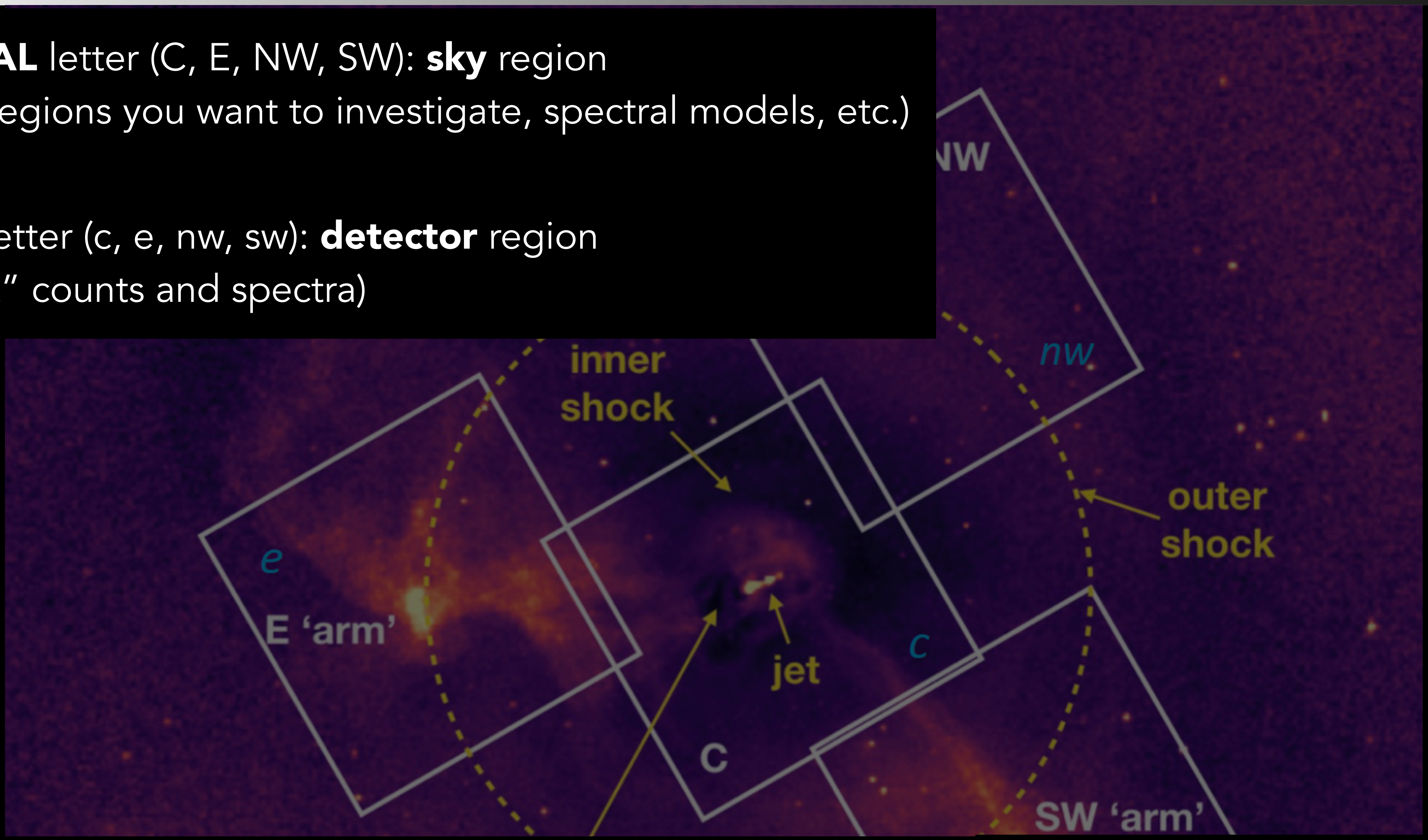
✓ $\mathbf{A_i}$ = ***effective area*** (ARF) of detector region i



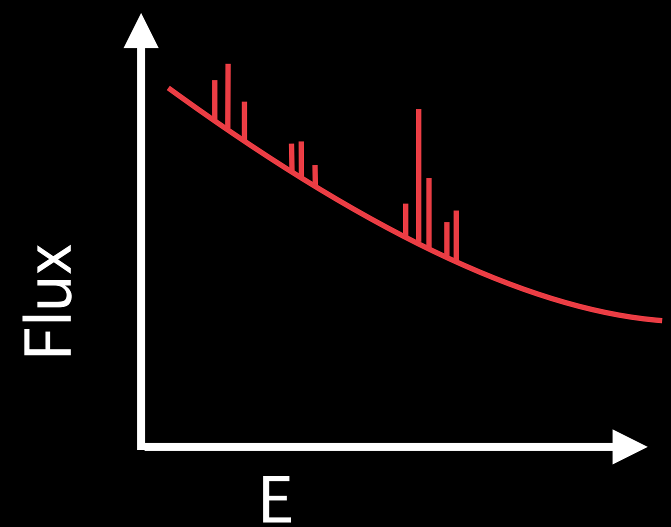
A concrete case (Virgo cluster)

- ✓ **CAPITAL** letter (C, E, NW, SW): **sky** region
  (“true” regions you want to investigate, spectral models, etc.)

- ✓ **small** letter (c, e, nw, sw): **detector** region
  (“output” counts and spectra)



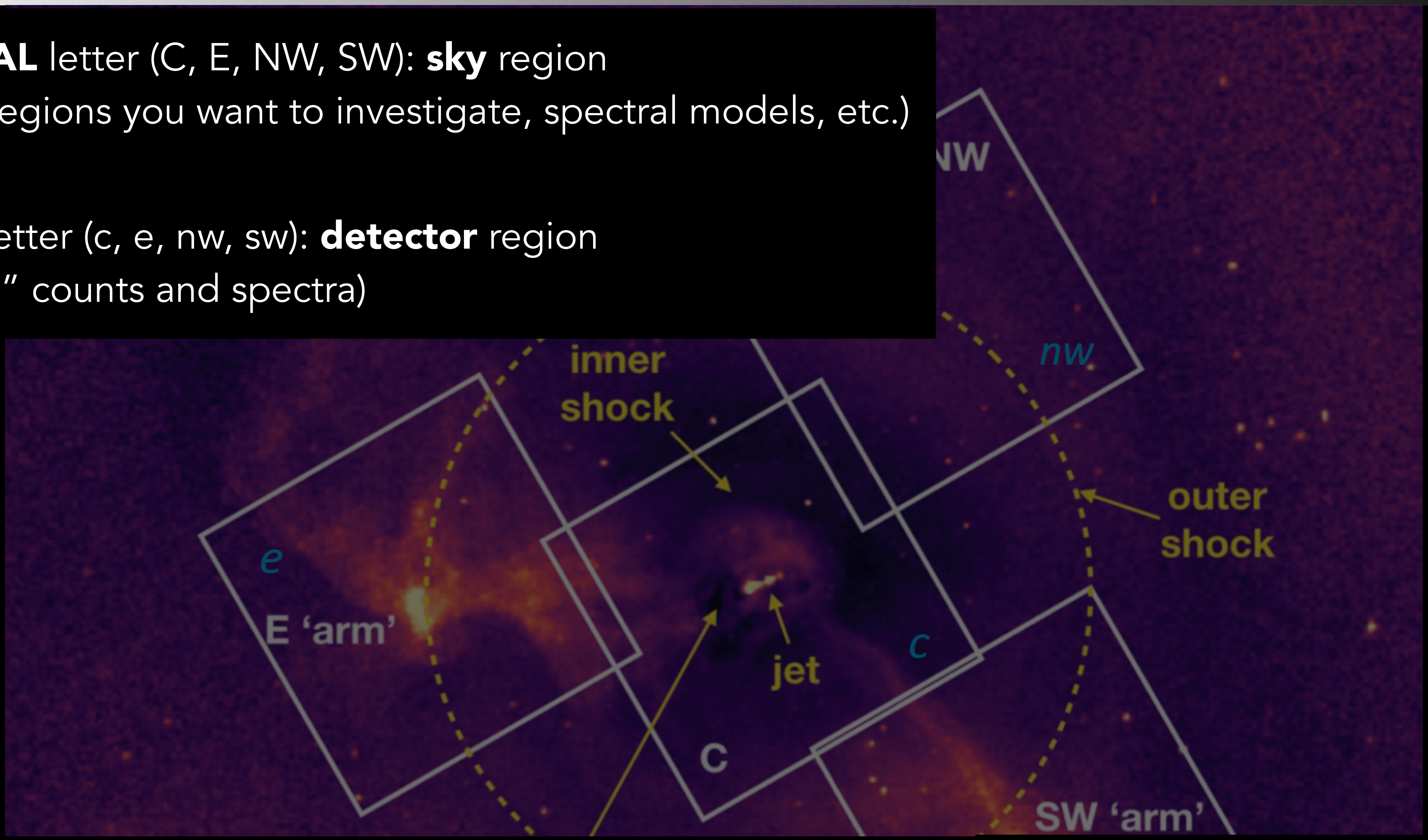
- ✓ S_i = observed ***spectrum*** of detector region i
- ✓ **M_J = *spectral model*** of sky region J
- ✓ R_i = ***response matrix*** (RMF) of detector region i
- ✓ A_i = ***effective area*** (ARF) of detector region i



A concrete case (Virgo cluster)

✓ **CAPITAL** letter (C, E, NW, SW): **sky** region
("true" regions you want to investigate, spectral models, etc.)

✓ **small** letter (c, e, nw, sw): **detector** region
("output" counts and spectra)

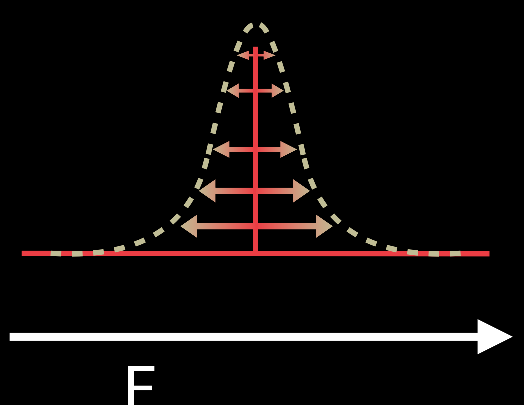


✓ S_i = observed ***spectrum*** of detector region i

✓ M_J = ***spectral model*** of sky region J

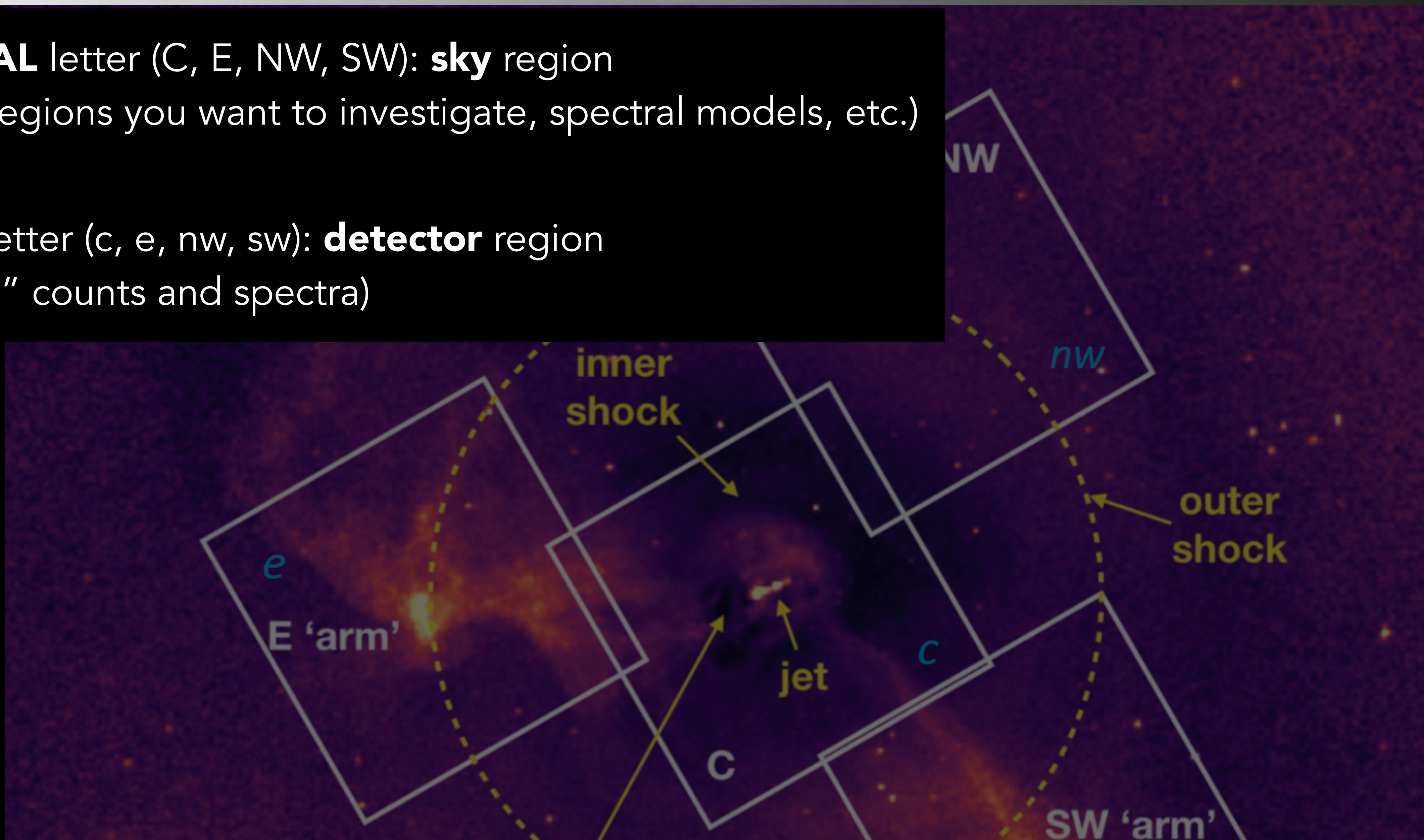
✓ R_i = ***response matrix*** (RMF) of detector region i

✓ A_i = ***effective area*** (ARF) of detector region i

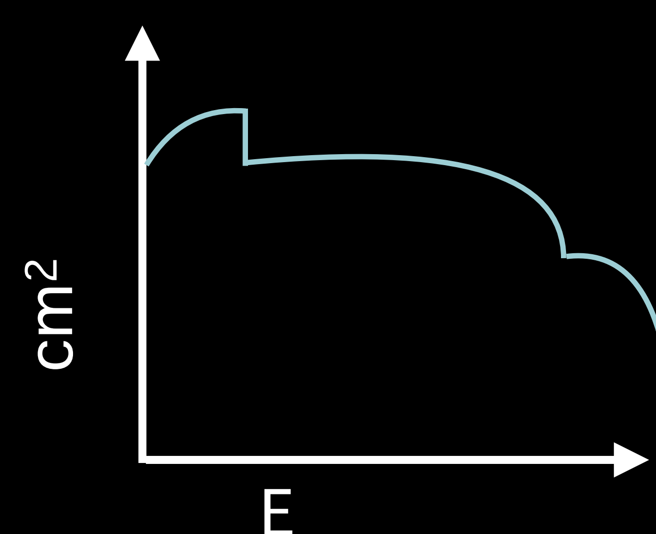


A concrete case (Virgo cluster)

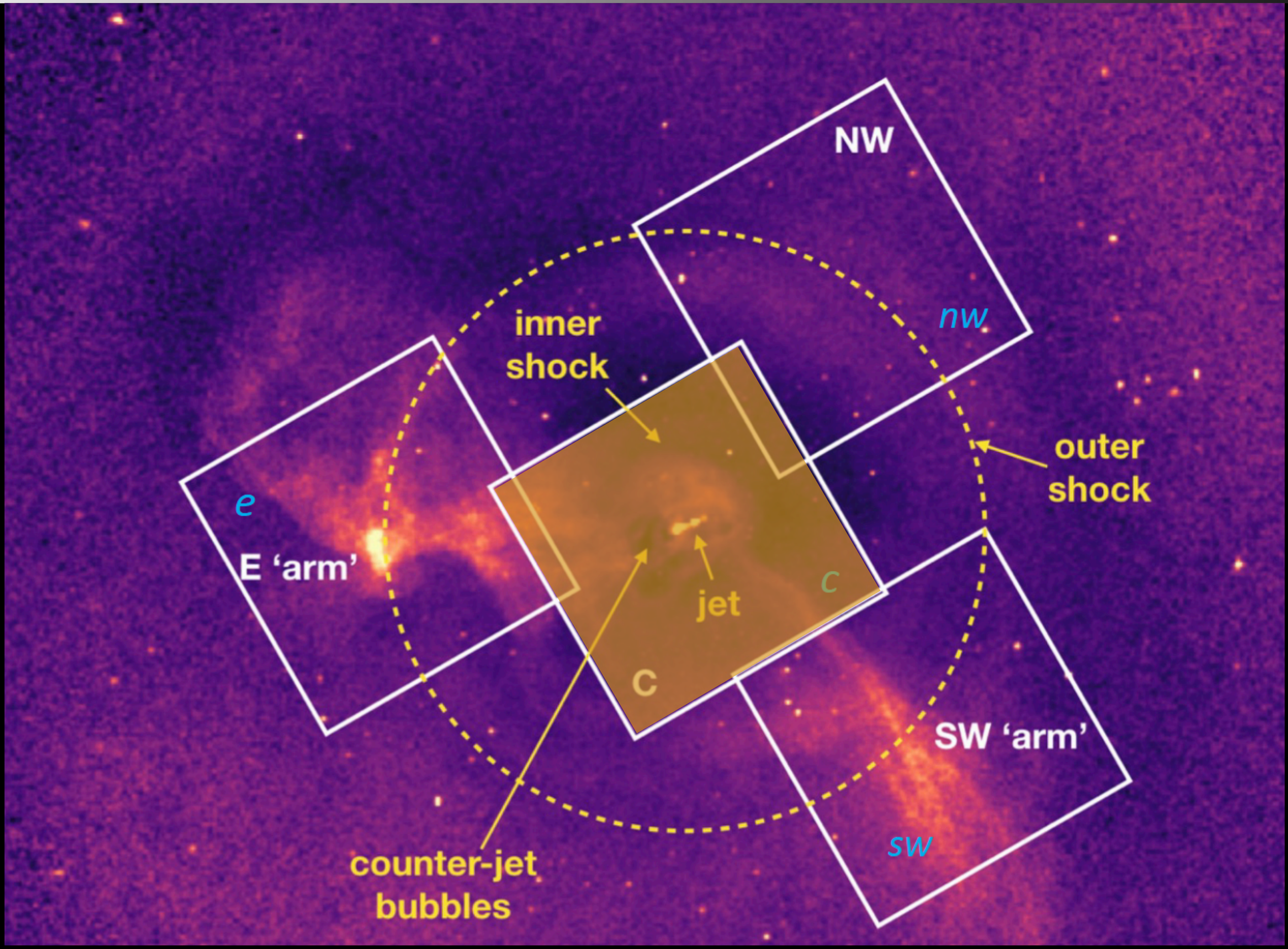
- ✓ **CAPITAL** letter (C, E, NW, SW): **sky** region
  (“true” regions you want to investigate, spectral models, etc.)
- ✓ **small** letter (c, e, nw, sw): **detector** region
  (“output” counts and spectra)



- ✓ **S_i** = observed ***spectrum*** of detector region i
- ✓ **M_J** = ***spectral model*** of sky region J
- ✓ **R_i** = ***response matrix*** (RMF) of detector region i
- ✓ **A_i** = ***effective area*** (ARF) of detector region i

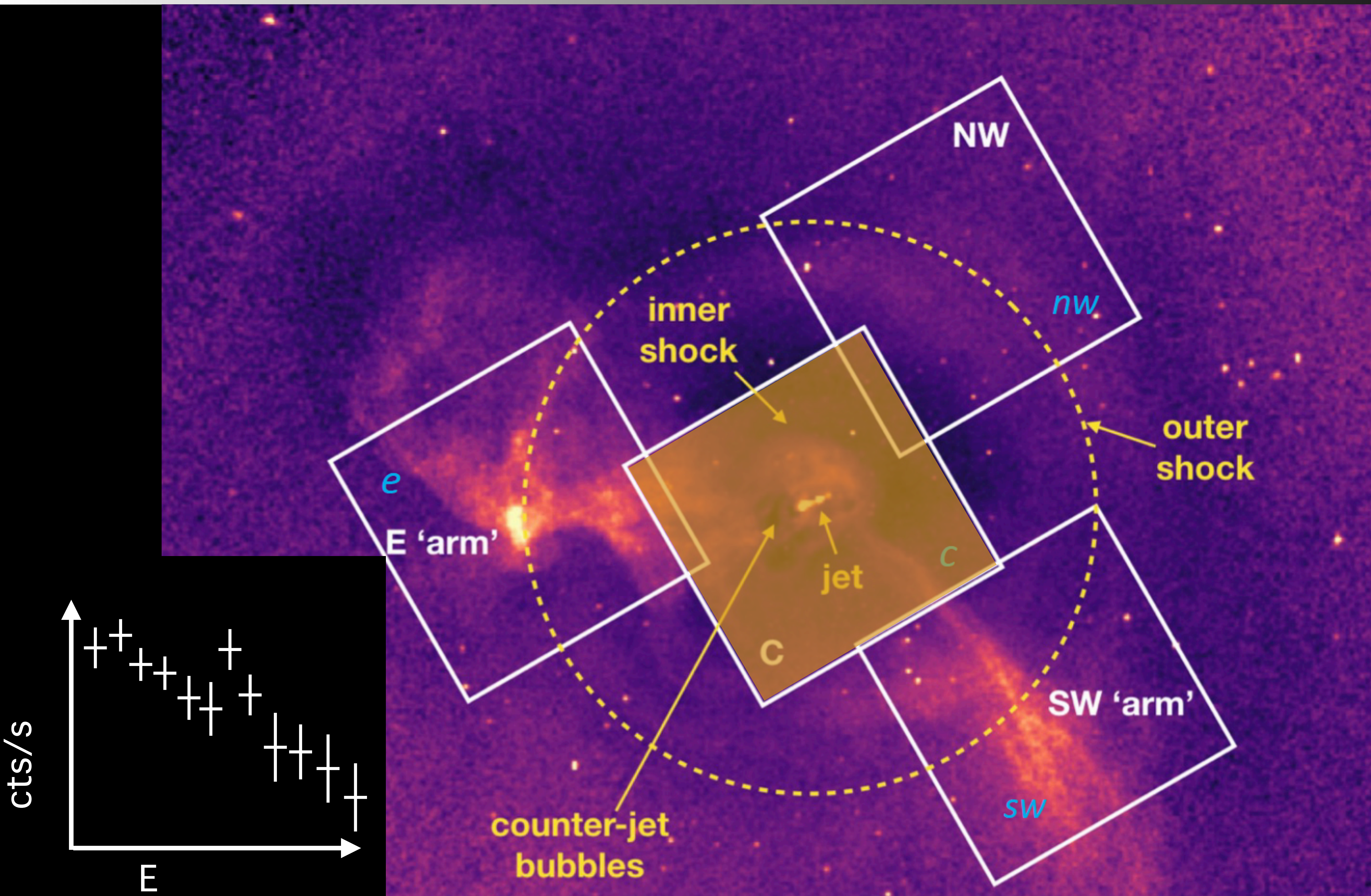


A concrete case (Virgo cluster)



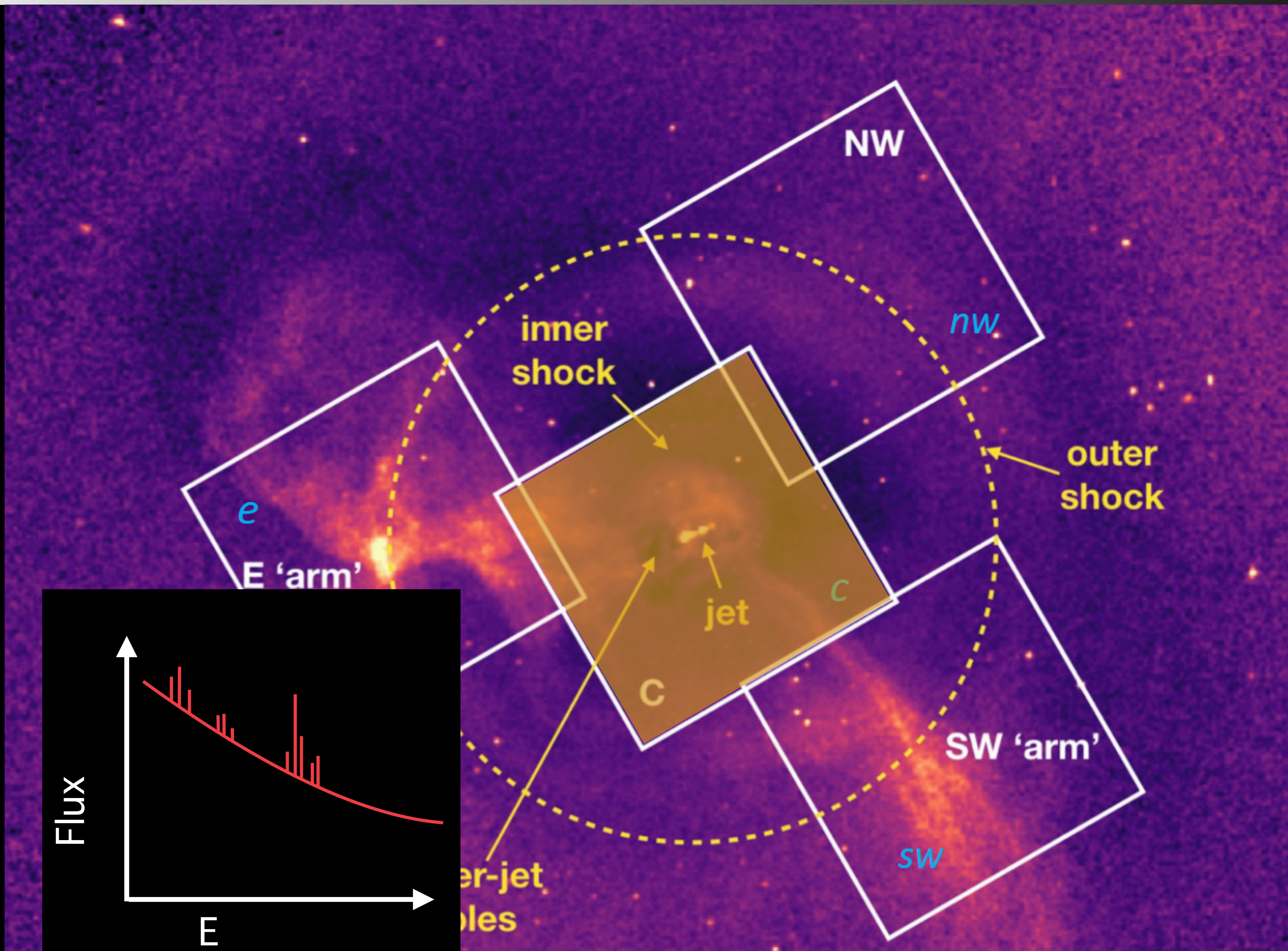
S_c =

A concrete case (Virgo cluster)



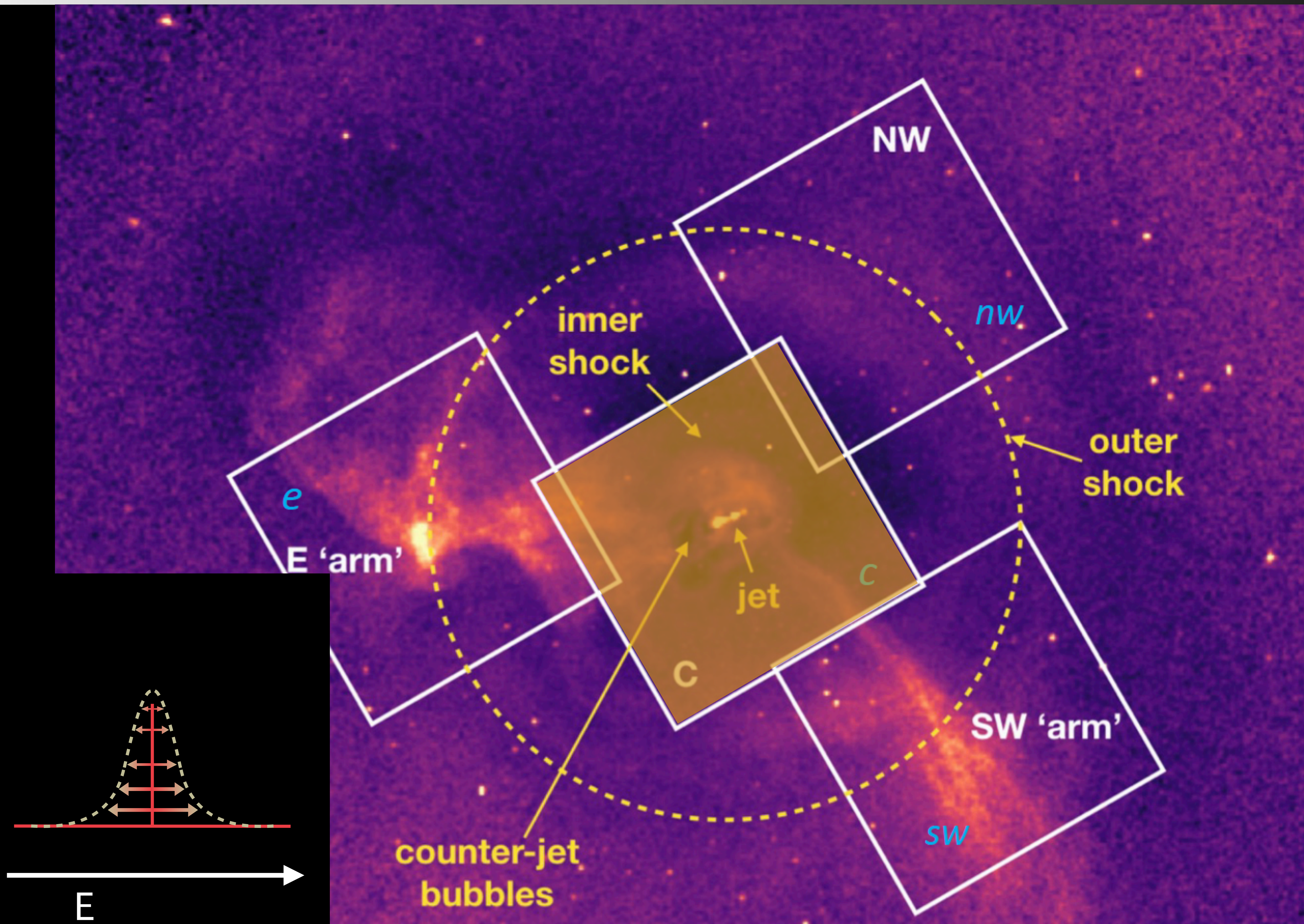
S_c =

A concrete case (Virgo cluster)



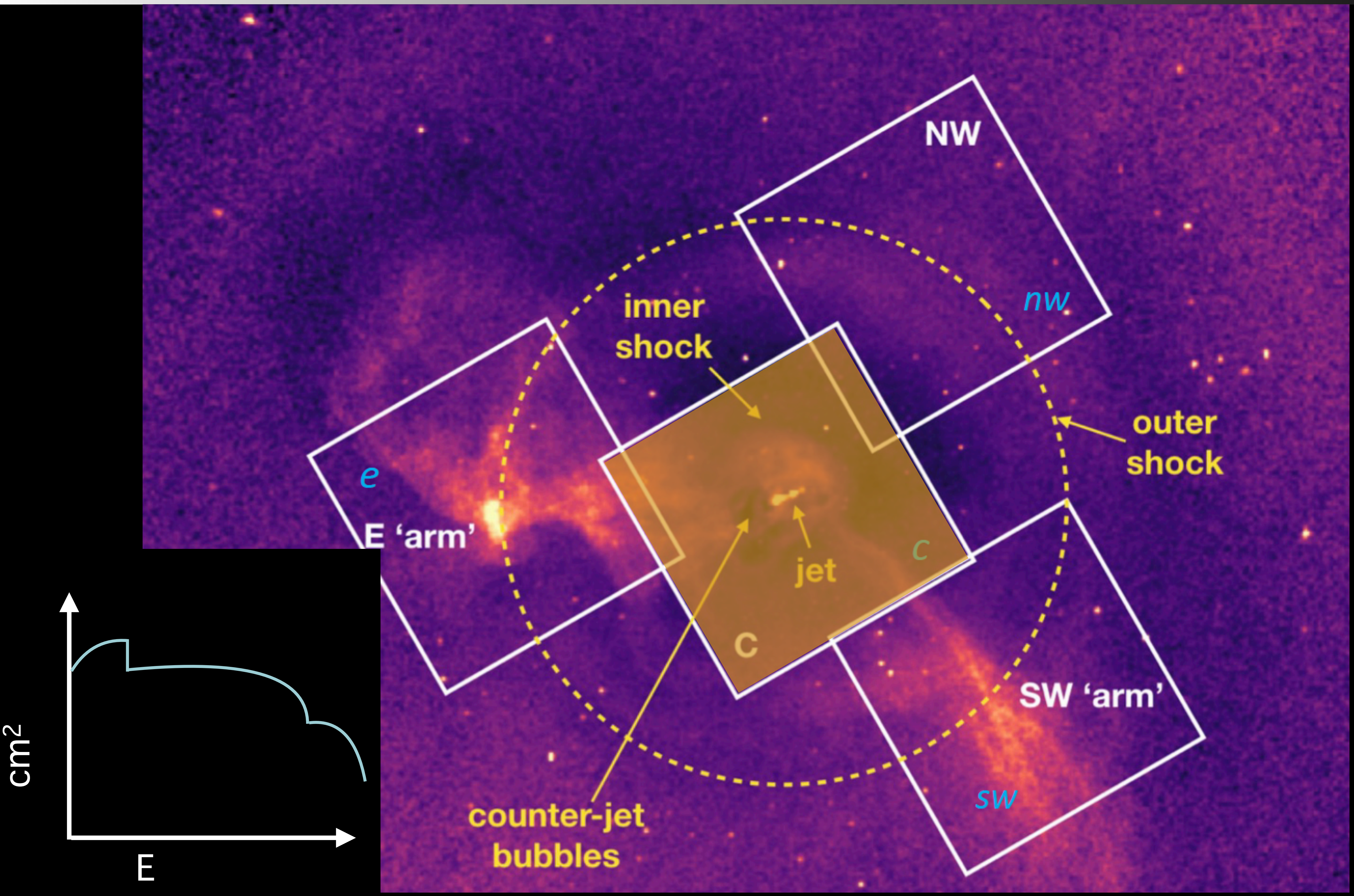
$S_C = A_C R_C$ $\mathbf{M_C}$

A concrete case (Virgo cluster)



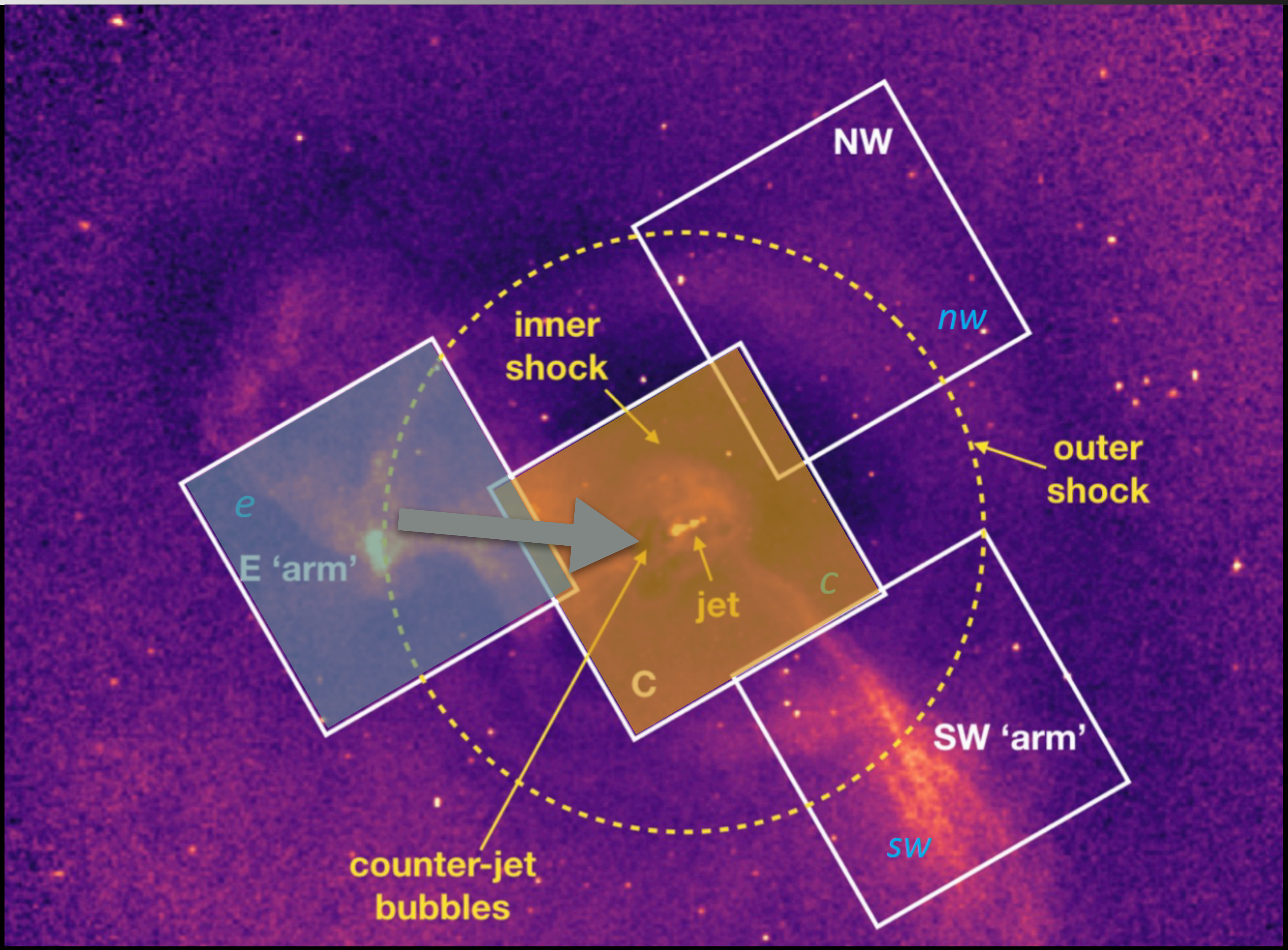
$S_c = A_c\, \mathbf{R_c}$ M_C

A concrete case (Virgo cluster)



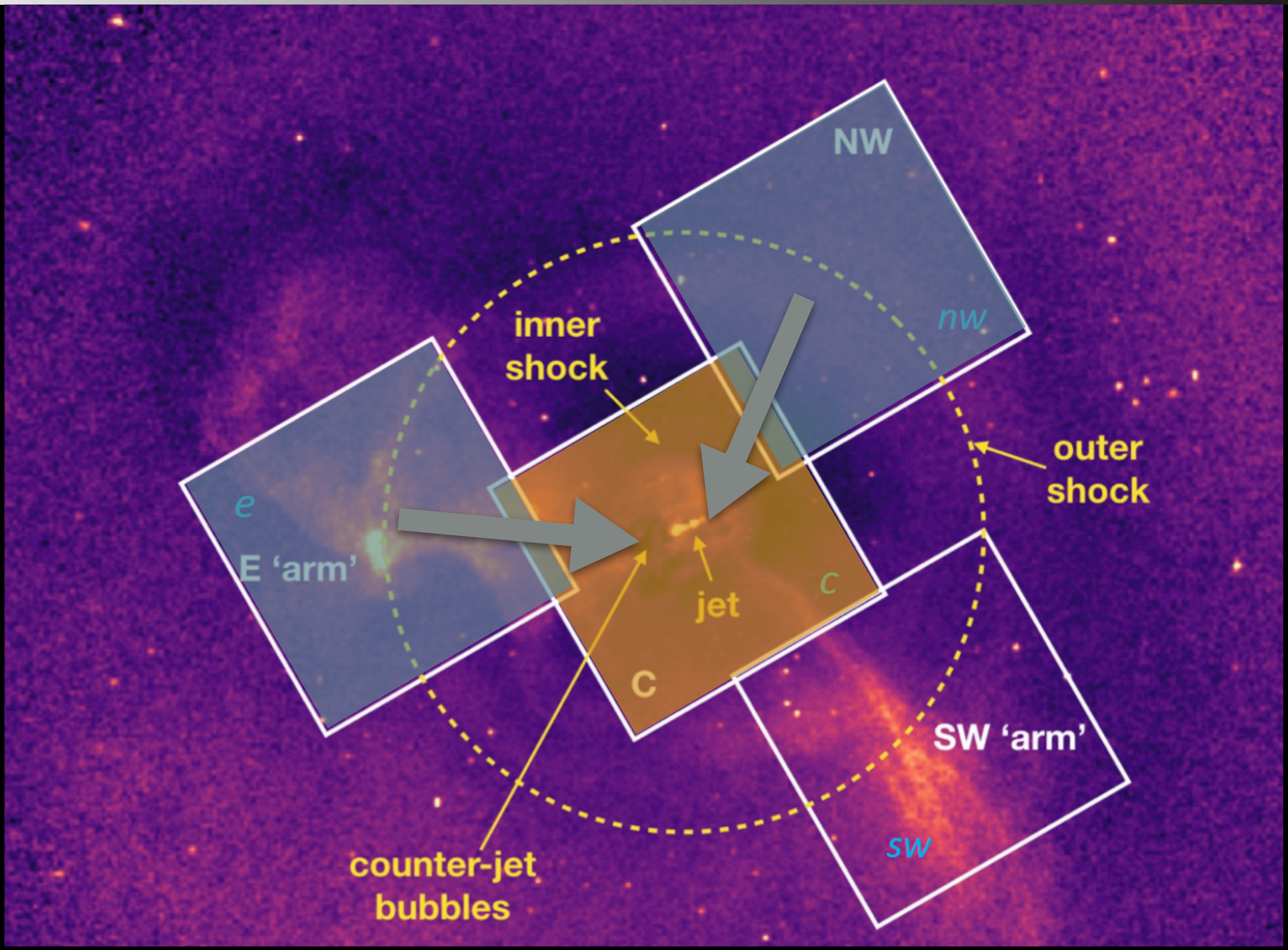
$S_c = \mathbf{A_c} R_c$ M_C

A concrete case (Virgo cluster)



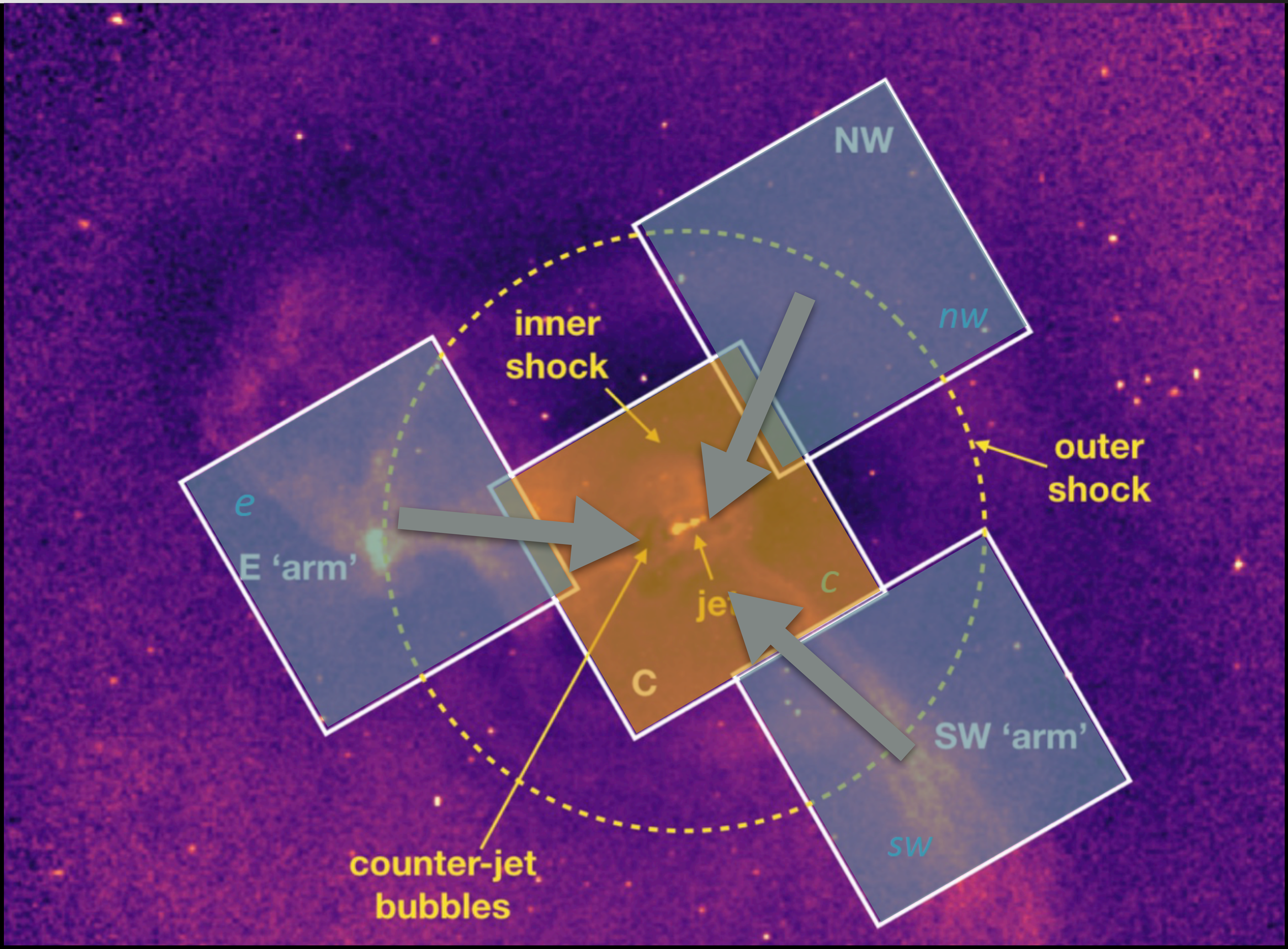
$S_c = A_c R_c [\mathbf{P}_{C\to c} M_C + \mathbf{P}_{E\to c} \mathbf{M}_E$]

A concrete case (Virgo cluster)



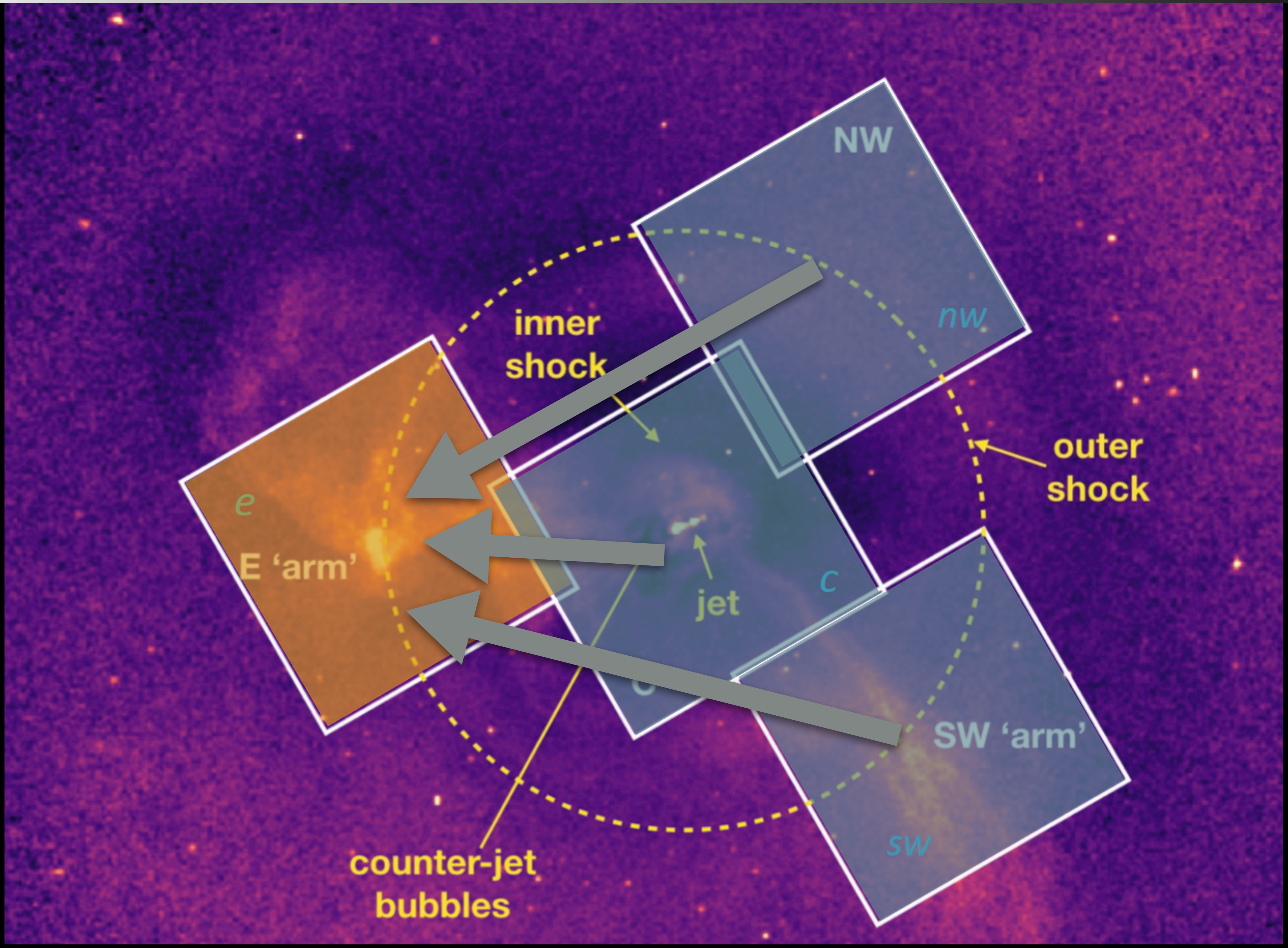
$S_c = A_c R_c [P_{C\to c} M_C + P_{E\to c} M_E + \mathbf{P_{NW\to c} M_{NW}}$ $]$

A concrete case (Virgo cluster)



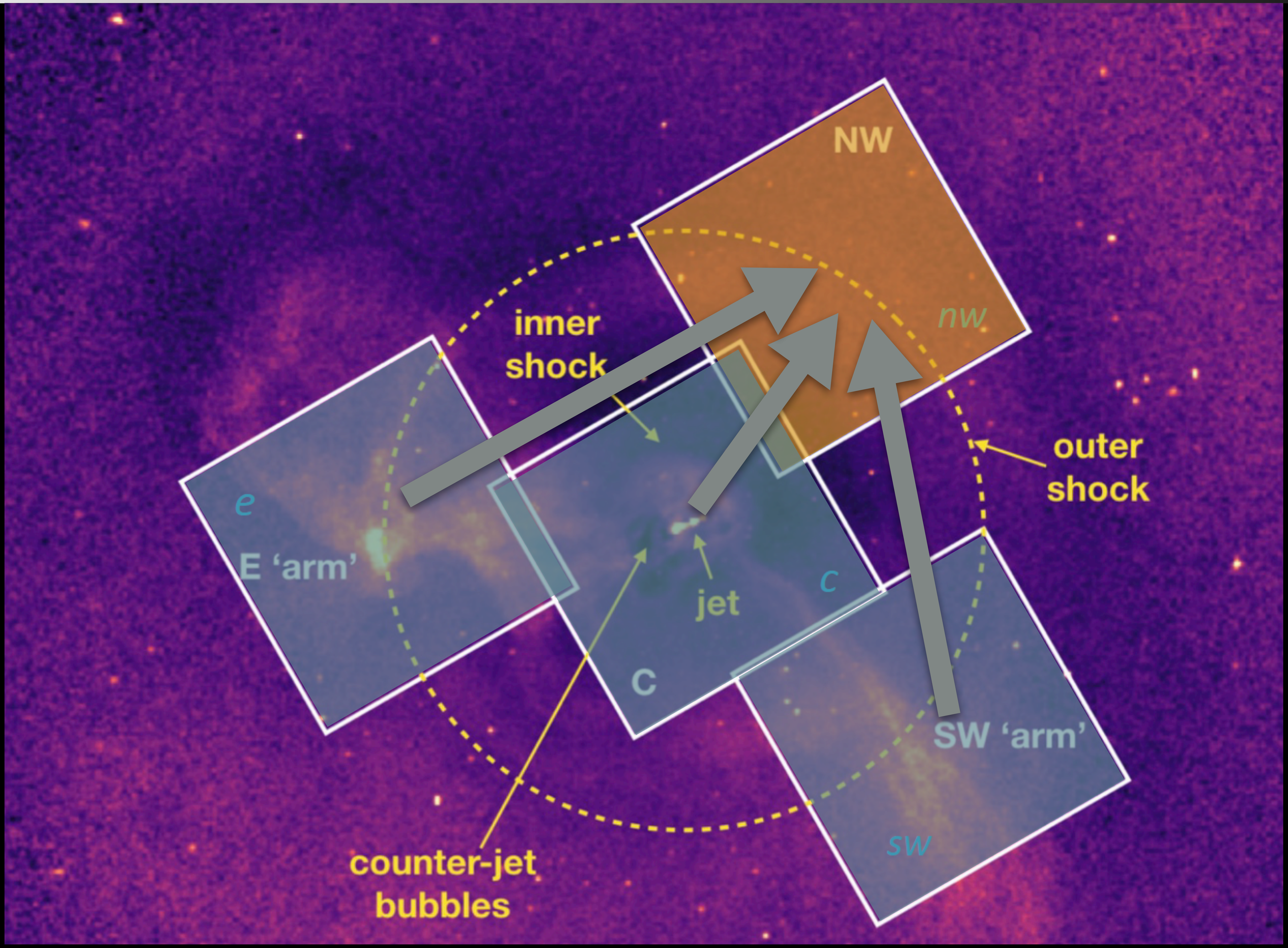
$S_c = A_c R_c [P_{C \rightarrow c} M_C + P_{E \rightarrow c} M_E + P_{NW \rightarrow c} M_{NW} + \mathbf{P_{SW \rightarrow c} M_{SW}}]$

A concrete case (Virgo cluster)



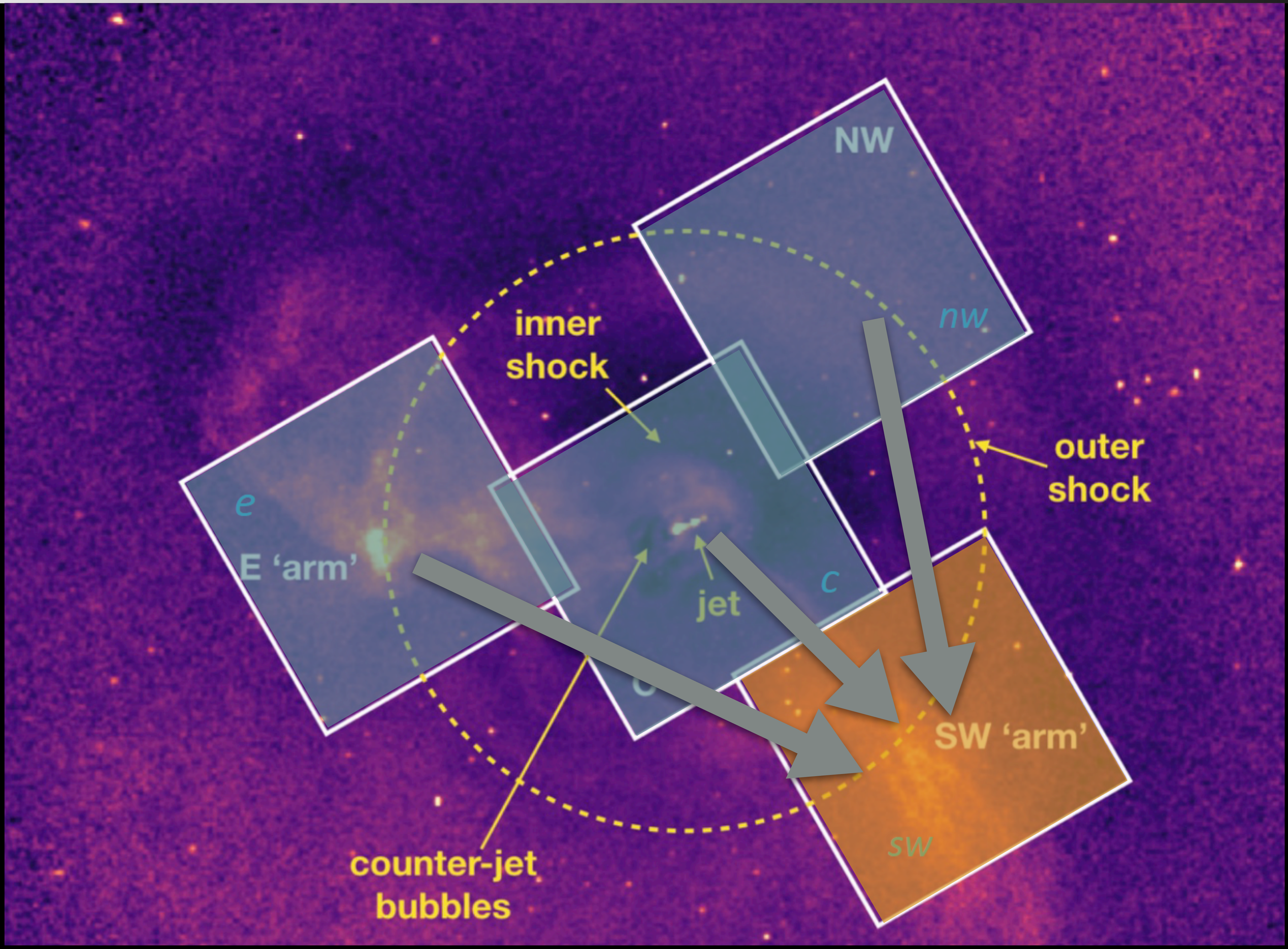
$$S_e = A_e\, R_e\, [P_{C\rightarrow e}\, M_C + P_{E\rightarrow e}\, M_E + P_{NW\rightarrow e}\, M_{NW} + P_{SW\rightarrow e}\, M_{SW}]$$

A concrete case (Virgo cluster)



$$S_{nw} = A_{nw}\, R_{nw}\, [P_{C\rightarrow nw}\, M_C + P_{E\rightarrow nw}\, M_E + P_{NW\rightarrow nw}\, M_{NW} + P_{SW\rightarrow nw}\, M_{SW}]$$

A concrete case (Virgo cluster)



$$S_{sw} = A_{sw}\, R_{sw}\, [P_{C\rightarrow sw}\, M_C + P_{E\rightarrow sw}\, M_E + P_{NW\rightarrow sw}\, M_{NW} + P_{SW\rightarrow sw}\, M_{SW}]$$

A concrete case (Virgo cluster)

$$\begin{bmatrix} S_c = A_c\, R_c\, [P_{C\rightarrow c}\, M_C + P_{E\rightarrow c}\, M_E + P_{NW\rightarrow c}\, M_{NW} + P_{SW\rightarrow c}\, M_{SW}] \\ S_e = A_e\, R_e\, [P_{C\rightarrow e}\, M_C + P_{E\rightarrow e}\, M_E + P_{NW\rightarrow e}\, M_{NW} + P_{SW\rightarrow e}\, M_{SW}] \\ S_{nw} = A_{nw}\, R_{nw}\, [P_{C\rightarrow nw}\, M_C + P_{E\rightarrow nw}\, M_E + P_{NW\rightarrow nw}\, M_{NW} + P_{SW\rightarrow nw}\, M_{SW}] \\ S_{sw} = A_{sw}\, R_{sw}\, [P_{C\rightarrow sw}\, M_C + P_{E\rightarrow sw}\, M_E + P_{NW\rightarrow sw}\, M_{NW} + P_{SW\rightarrow sw}\, M_{SW}] \end{bmatrix}$$

Effective areas of detector region i

Observed spectra of detector region i

Response matrices of detector region i

Spectral models of sky region J

$$S_i = A_i\, R_i \sum_J P_{J\rightarrow i}\, M_J$$

A concrete case (Virgo cluster)

$$\begin{cases} S_c = A_c\, R_c\, [P_{C\to c}\, M_C + P_{E\to c}\, M_E + P_{NW\to c}\, M_{NW} + P_{SW\to c}\, M_{SW}] \\ S_e = A_e\, R_e\, [P_{C\to e}\, M_C + P_{E\to e}\, M_E + P_{NW\to e}\, M_{NW} + P_{SW\to e}\, M_{SW}] \\ S_{nw} = A_{nw}\, R_{nw}\, [P_{C\to nw}\, M_C + P_{E\to nw}\, M_E + P_{NW\to nw}\, M_{NW} + P_{SW\to nw}\, M_{SW}] \\ S_{sw} = A_{sw}\, R_{sw}\, [P_{C\to sw}\, M_C + P_{E\to sw}\, M_E + P_{NW\to sw}\, M_{NW} + P_{SW\to sw}\, M_{SW}] \end{cases}$$

Coefficients of photon mixing from sky region J into detector region i

$$S_i = A_i\, R_i \sum_J P_{J\to i}\, M_J$$

4 x 4 = 16 models fitted simultaneously to 4 observed spectra

A concrete case (Virgo cluster)

How to obtain the $P_{J \to i}$ coefficients?

- ➡ **Method 1**: leave normalizations of the 16 models free
  - ➡ …Bad idea! (Empirical, black box, too many free params, degeneracies, etc.)
- ➡ **Method 2**: estimate coefficients from ray-tracing simulations
  - ➡ …Via ARF generator (part of the future XRISM data reduction software)

A concrete case (Virgo cluster)

➡ **Method 2**: estimate coefficients from ray-tracing simulations

➡ …Via ARF generator (part of the future XRISM data reduction software)

Input "source type" options

point source →

flat circle →

β-model →

Input image (e.g. *Chandra*, sim., etc.)

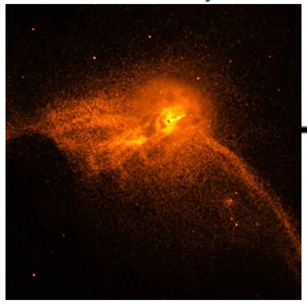


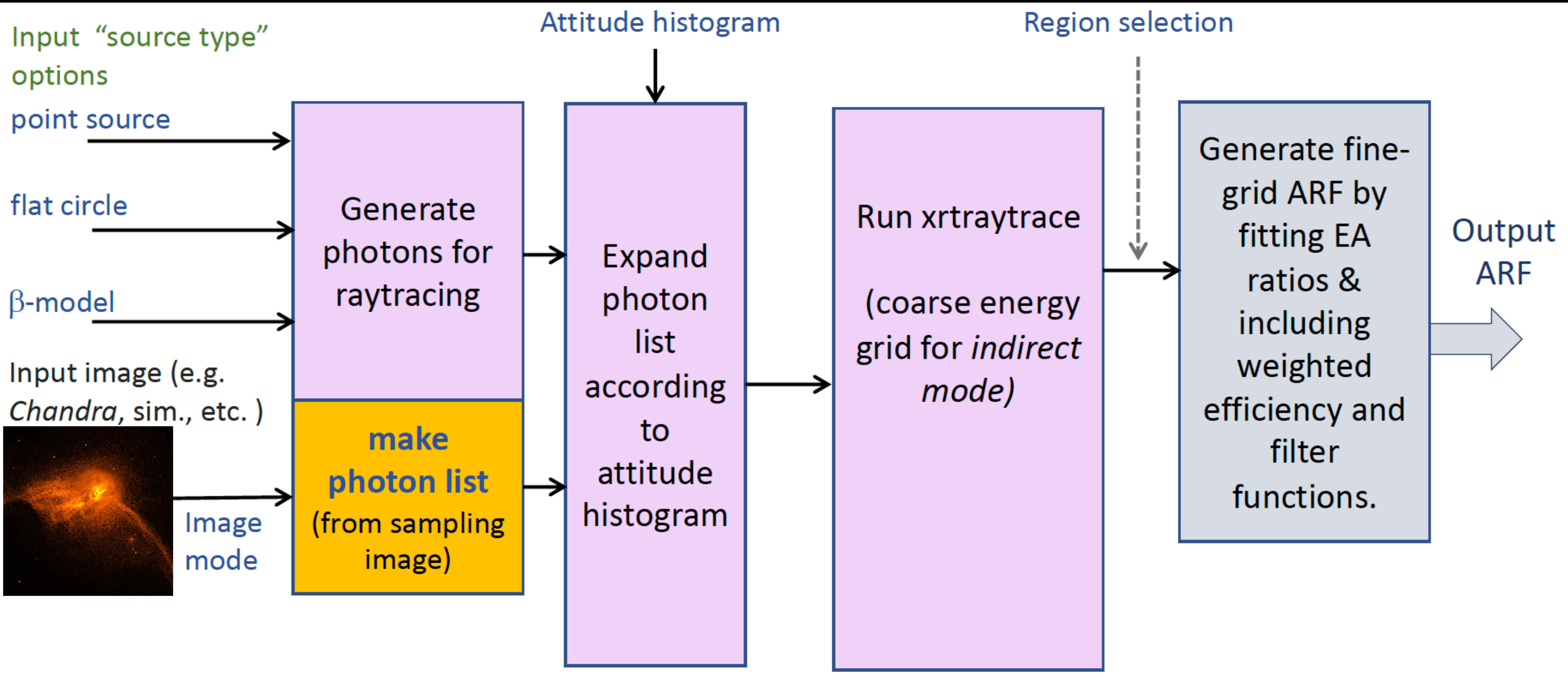
Image mode →

- **aharfgen** (Hitomi, this workshop - see Mike's tutorial)
- **xaarfgen** (XRISM, similar task - name might still change)

Output ARF ⇨

A concrete case (Virgo cluster)

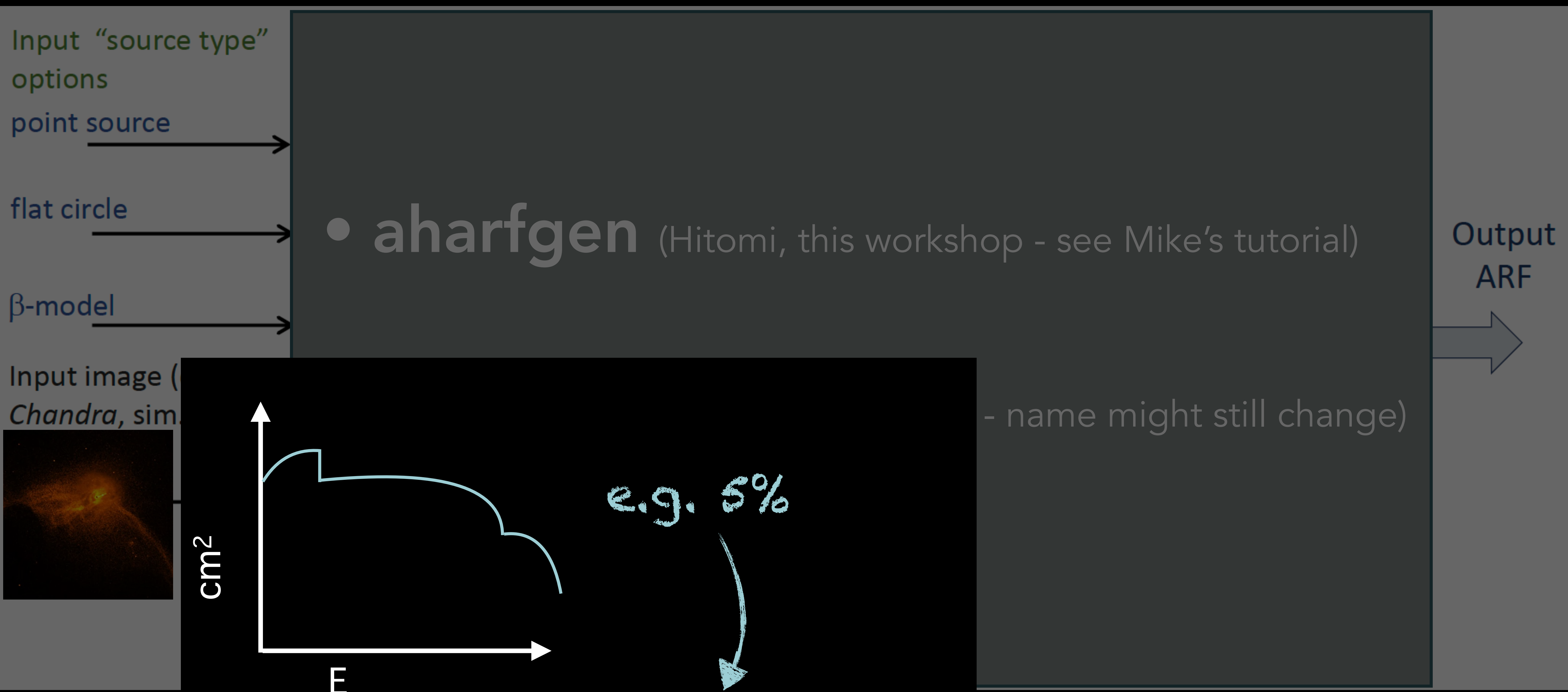
- ➡ **Method 2**: estimate coefficients from ray-tracing simulations
  - ➡ …Via ARF generator (part of the future XRISM data reduction software)



A concrete case (Virgo cluster)

➡ **Method 2**: estimate coefficients from ray-tracing simulations

➡ …Via ARF generator (part of the future XRISM data reduction software)

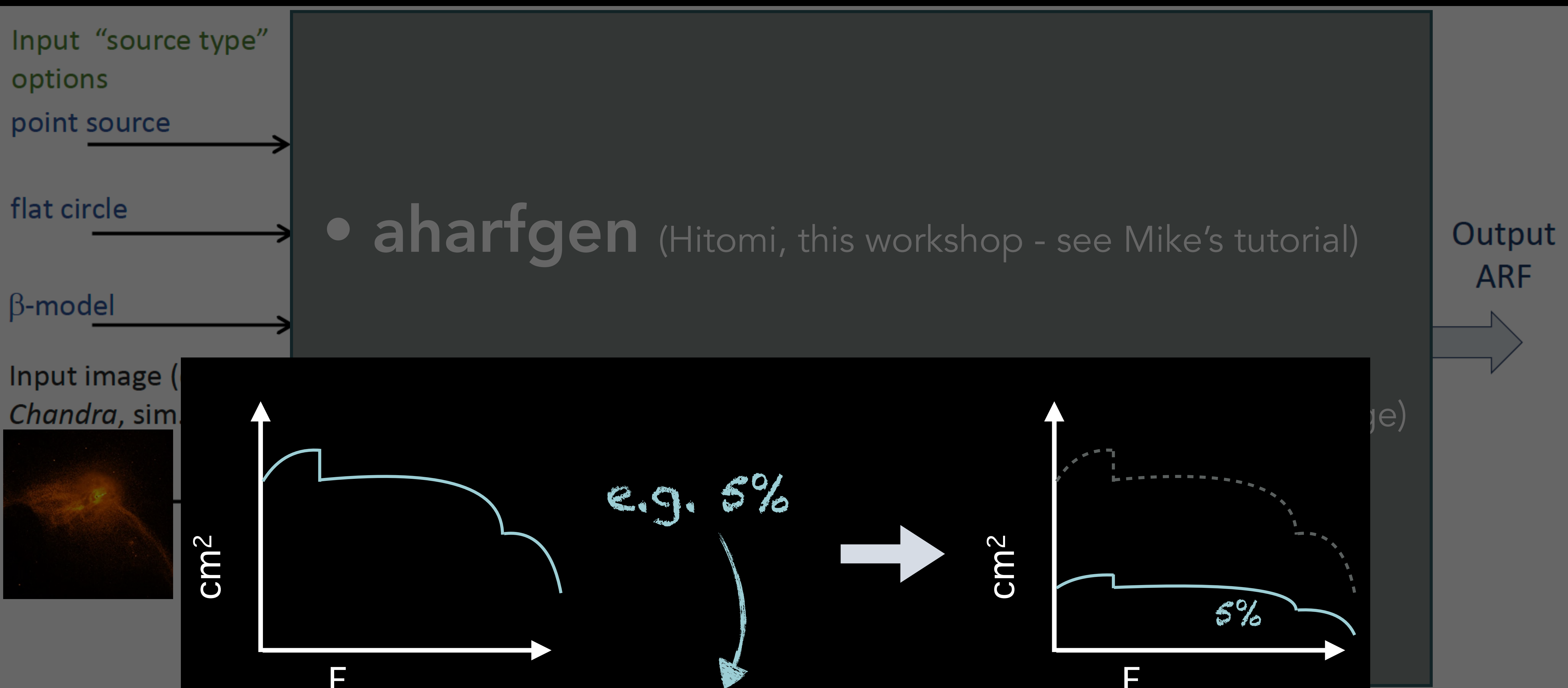


$$(A_i \times P_{J \to i})$$

A concrete case (Virgo cluster)

➡ **Method 2**: estimate coefficients from ray-tracing simulations

➡ …Via ARF generator (part of the future XRISM data reduction software)



$$(A_i \times P_{J\to i}) \Rightarrow A_{J\to i}$$

A concrete case (Virgo cluster)

$$
\begin{bmatrix}
S_c = A_c\, R_c\, [P_{C\to c}\, M_C + P_{E\to c}\, M_E + P_{NW\to c}\, M_{NW} + P_{SW\to c}\, M_{SW}] \\
S_e = A_e\, R_e\, [P_{C\to e}\, M_C + P_{E\to e}\, M_E + P_{NW\to e}\, M_{NW} + P_{SW\to e}\, M_{SW}] \\
S_{nw} = A_{nw}\, R_{nw}\, [P_{C\to nw}\, M_C + P_{E\to nw}\, M_E + P_{NW\to nw}\, M_{NW} + P_{SW\to nw}\, M_{SW}] \\
S_{sw} = A_{sw}\, R_{sw}\, [P_{C\to sw}\, M_C + P_{E\to sw}\, M_E + P_{NW\to sw}\, M_{NW} + P_{SW\to sw}\, M_{SW}]
\end{bmatrix}
$$

Observed spectra of detector region i

Effective areas of detector region i

Response matrices of detector region i

Spectral models of sky region J

$$S_i = A_i\, R_i \sum_J P_{J\to i}\, M_J$$

A concrete case (Virgo cluster)

$$
\begin{bmatrix}
S_c = R_c\,[A_{C\to c}\,M_C + A_{E\to c}\,M_E + A_{NW\to c}\,M_{NW} + A_{SW\to c}\,M_{SW}] \\
S_e = R_e\,[A_{C\to e}\,M_C + A_{E\to e}\,M_E + A_{NW\to e}\,M_{NW} + A_{SW\to e}\,M_{SW}] \\
S_{nw} = R_{nw}\,[A_{C\to nw}\,M_C + A_{E\to nw}\,M_E + A_{NW\to nw}\,M_{NW} + A_{SW\to nw}\,M_{SW}] \\
S_{sw} = R_{sw}\,[A_{C\to sw}\,M_C + A_{E\to sw}\,M_E + A_{NW\to sw}\,M_{NW} + A_{SW\to sw}\,M_{SW}]
\end{bmatrix}
$$

Observed spectra of detector region i

Response matrices of detector region i

Spectral models of sky region J

$$S_i = R_i \sum_J A_{J\to i}\, M_J$$

4 x 4 = 16 models fitted simultaneously to 4 observed spectra

A concrete case (Virgo cluster)

$$\begin{cases} S_c = R_c\,[A_{C\to c} M_C + A_{E\to c} M_E + A_{NW\to c} M_{NW} + A_{SW\to c} M_{SW}] \\ S_e = R_e\,[A_{C\to e} M_C + A_{E\to e} M_E + A_{NW\to e} M_{NW} + A_{SW\to e} M_{SW}] \\ S_{nw} = R_{nw}\,[A_{C\to nw} M_C + A_{E\to nw} M_E + A_{NW\to nw} M_{NW} + A_{SW\to nw} M_{SW}] \\ S_{sw} = R_{sw}\,[A_{C\to sw} M_C + A_{E\to sw} M_E + A_{NW\to sw} M_{NW} + A_{SW\to sw} M_{SW}] \end{cases}$$

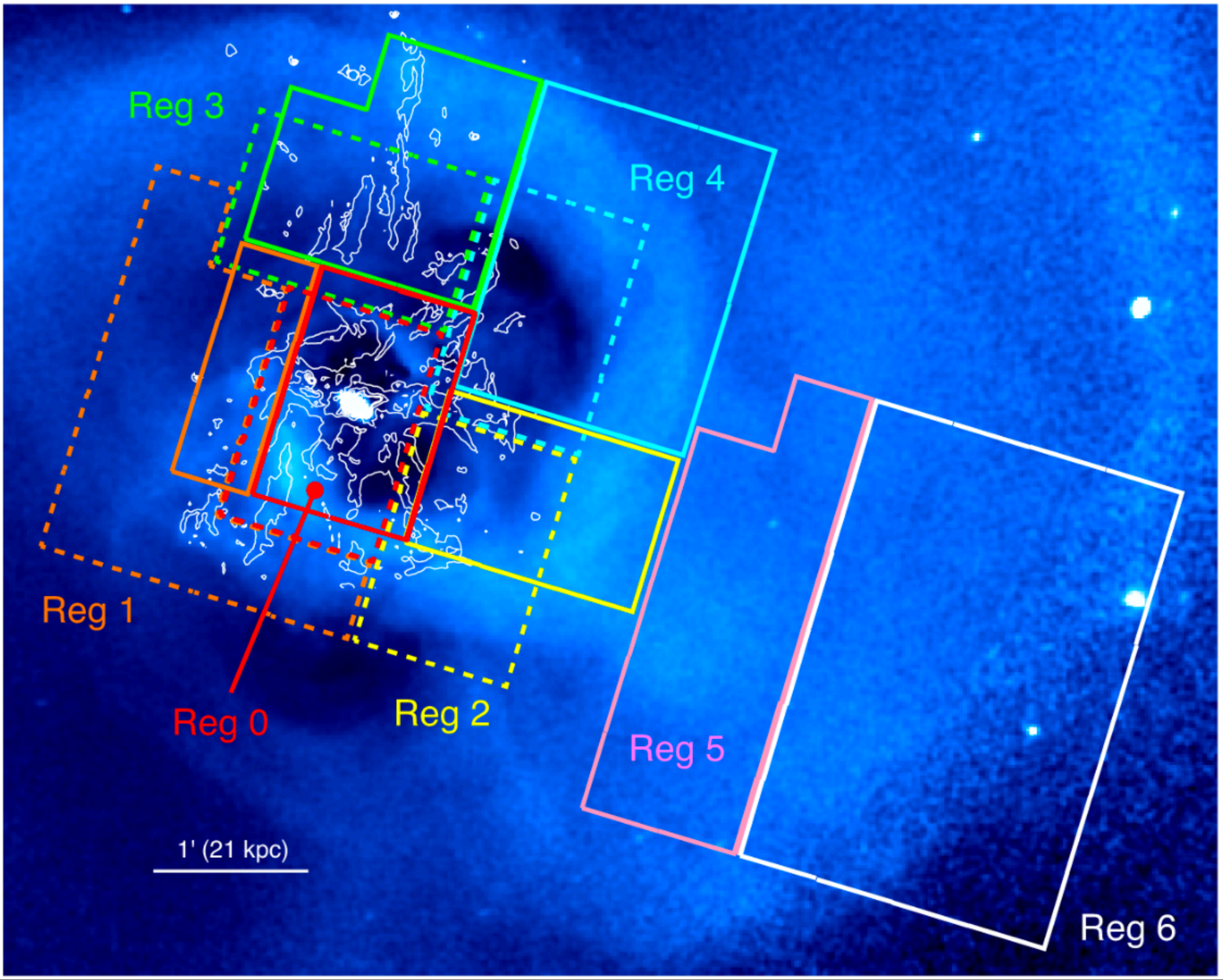
Effective areas of detector region i

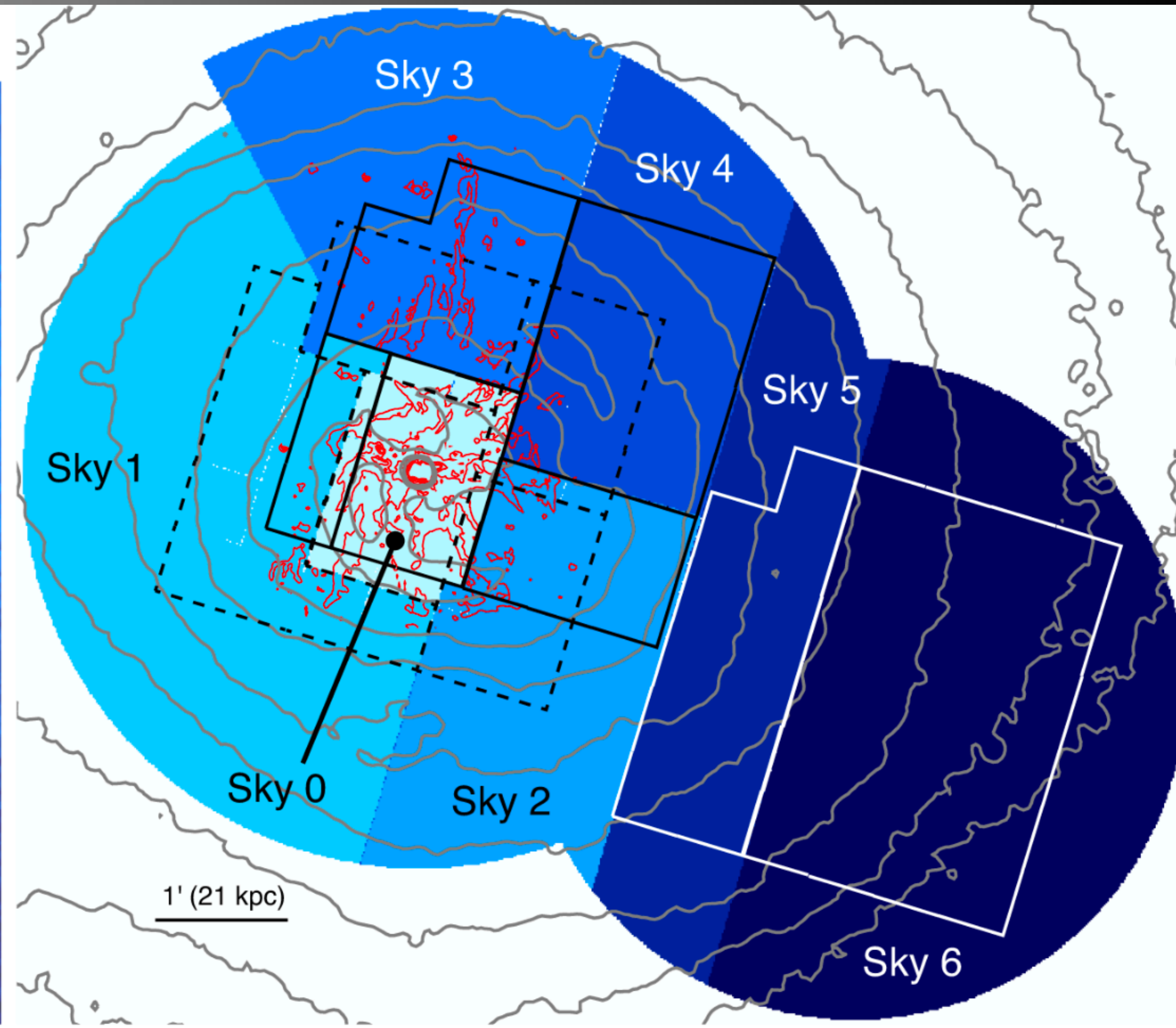
$$S_i = R_i \sum_J A_{J\to i} M_J$$

4 x 4 = 16 models fitted simultaneously to 4 observed spectra

Another concrete case (Perseus cluster)

Hitomi Collaboration et al. (2018)





12 detector regions

(obs1: 2 regions - obs2: 5 regions - obs3: 5 regions)

J = 0,1,...,12

6 sky regions

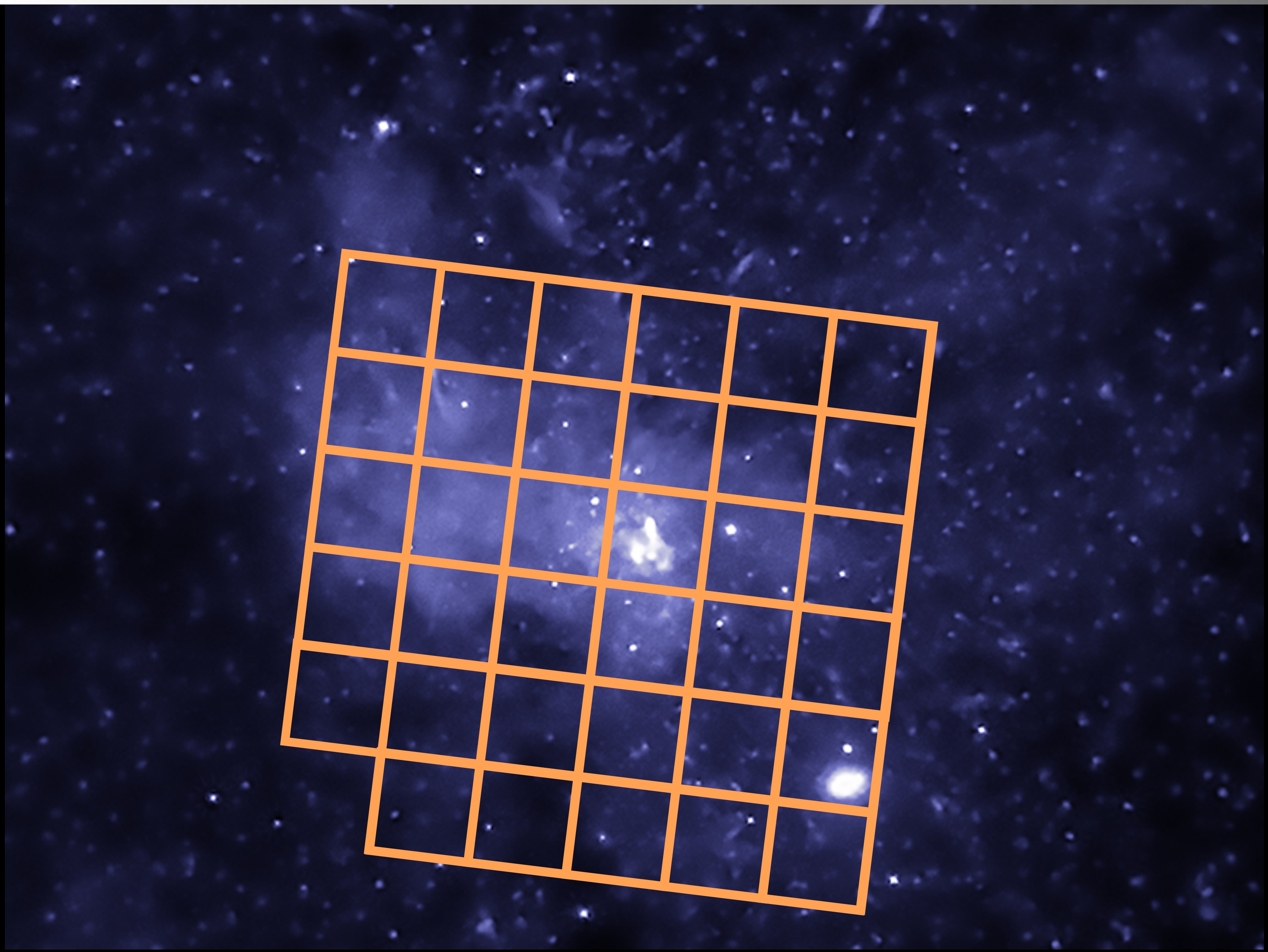
(Larger than detector regions to account for external SSM)

i = 0,1,...,6

$$S_i = R_i \sum_J A_{J\to i} M_J$$

12 x 6 = 72 models fitted simultaneously to 12 observed spectra (!)

Another concrete case (Sagittarius A*)



35 detector regions

J = 0,1,...,35

35 sky regions

i = 0,1,...,35

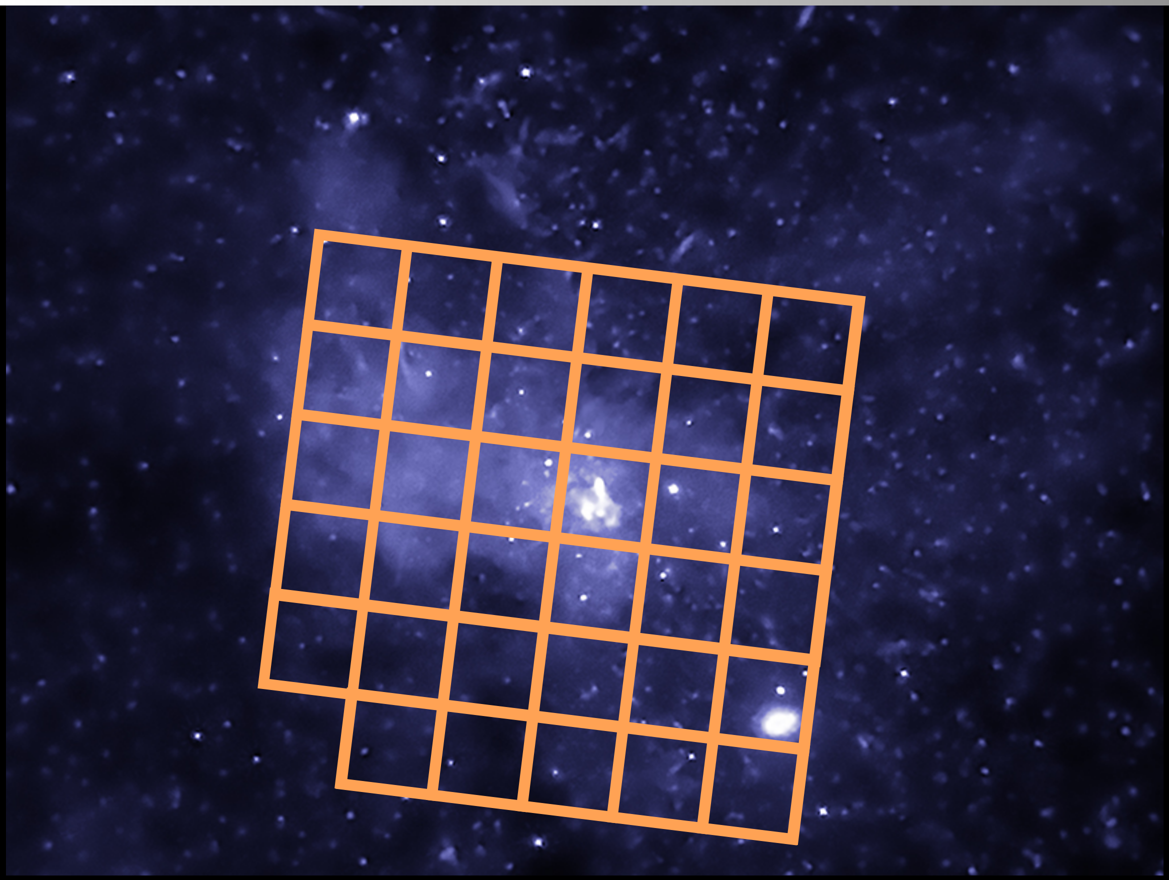
$$S_i = R_i \sum_J A_{J\to i} M_J$$

35 x 35 = 1225 models fitted simultaneously to 35 observed spectra !!

Problem: XRISM ARFs are **very** heavy to generate (>1-1.5 hour per ARF)!

➡ > 51 days on a single-core machine!

Mitigating the long ARF generation times



1. Divide a few regions of interest in a smart way (e.g. regions with similar special features, "optimize" the SSM, etc.)…

2. Treat each pixel separately vs "the rest" (modelled as a block)

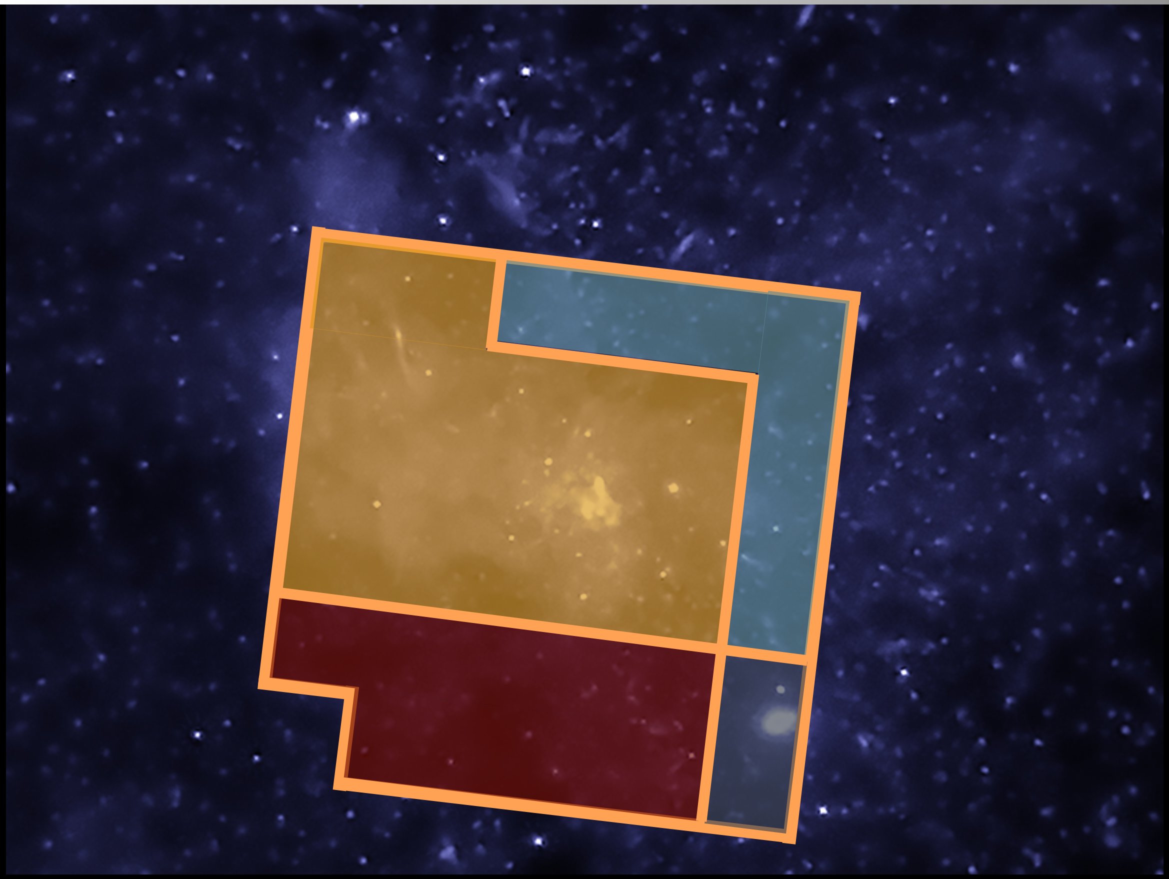
3. Ignore regions with virtually inexistent SSM

4. A library of pre-computed PSFs will be made available some time after the launch (PSFlib)

5. ARFs can be computed over narrow energy ranges too…

6. …other options with your host institute? (Grid computing, computing clusters, etc.)

Mitigating the long ARF generation times



1. Divide a few regions of interest in a smart way (e.g. regions with similar special features, "optimize" the SSM, etc.)…

2. Treat each pixel separately vs "the rest" (modelled as a block)

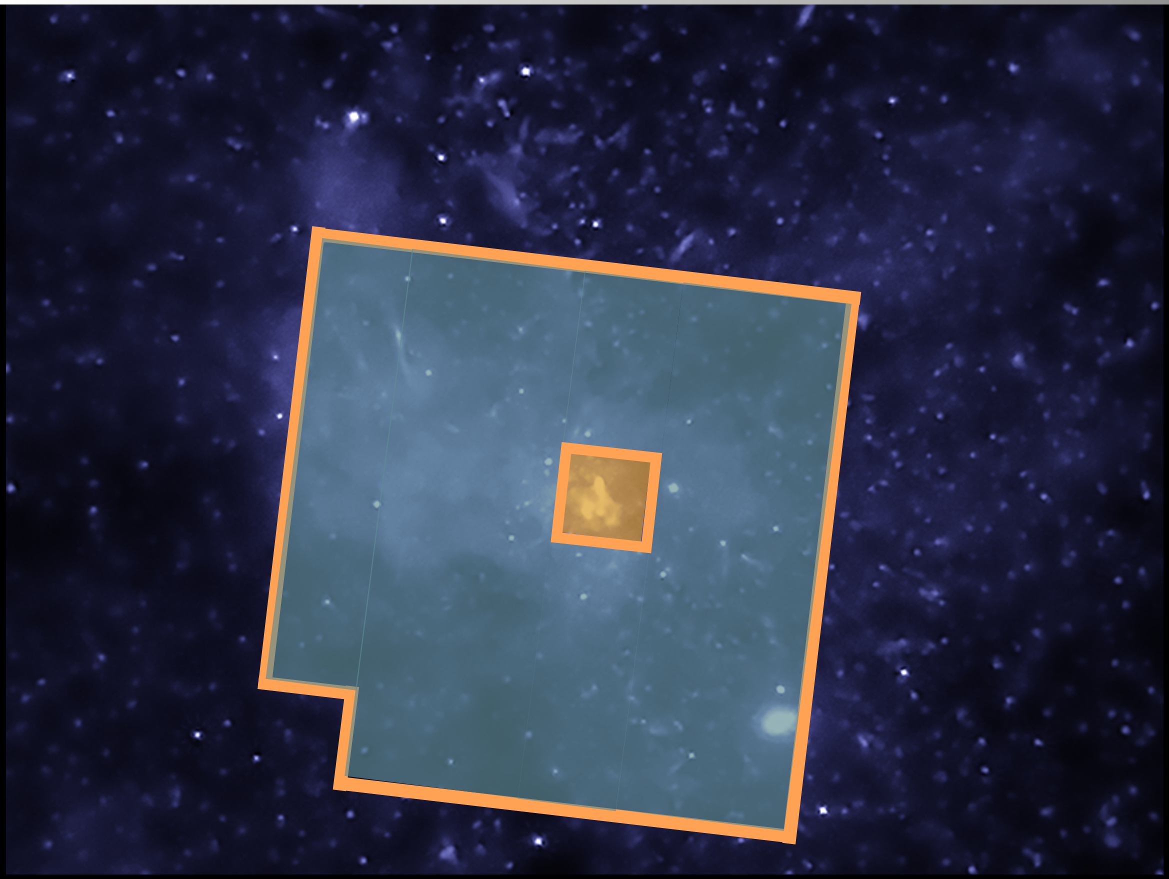
3. Ignore regions with virtually inexistent SSM

4. A library of pre-computed PSFs will be made available some time after the launch (PSFlib)

5. ARFs can be computed over narrow energy ranges too…

6. …other options with your host institute? (Grid computing, computing clusters, etc.)

Mitigating the long ARF generation times



1. Divide a few regions of interest in a smart way (e.g. regions with similar special features, "optimize" the SSM, etc.)…

2. Treat each pixel separately vs "the rest" (modelled as a block)

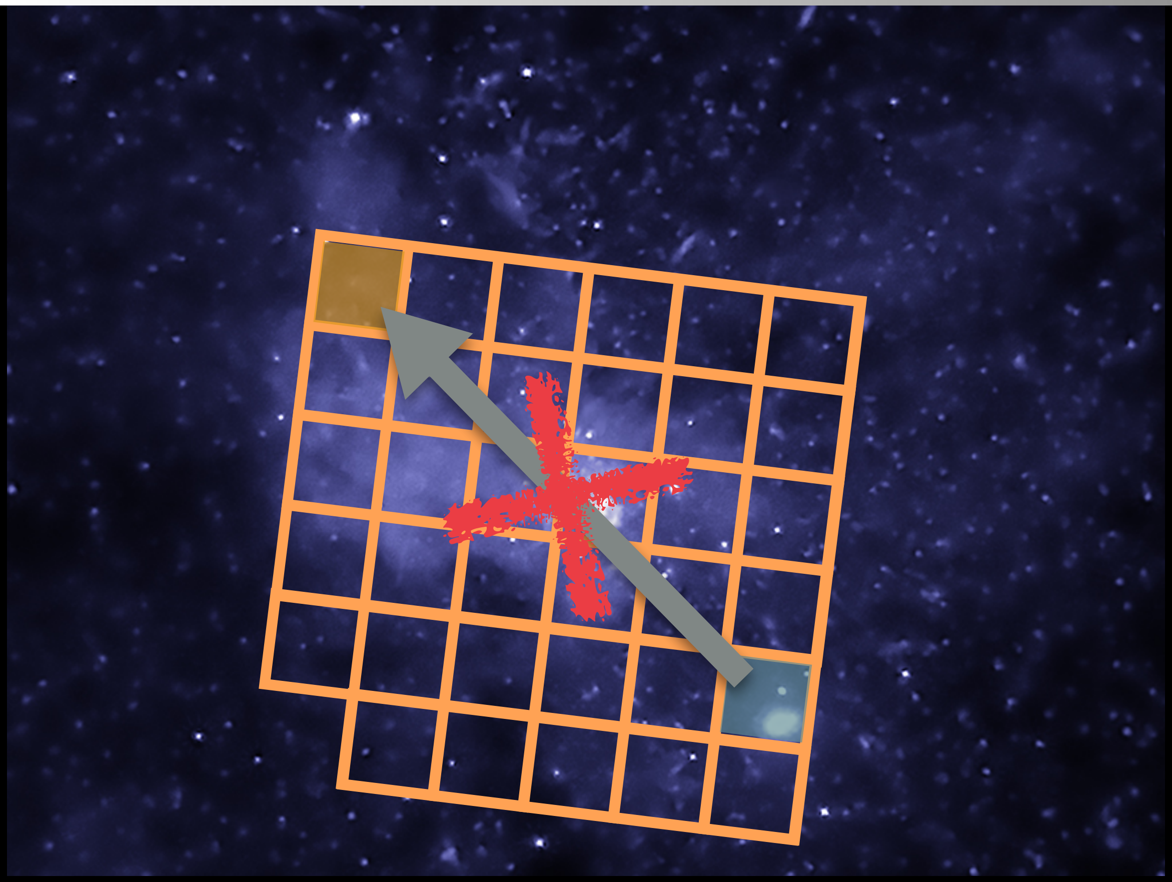
3. Ignore regions with virtually inexistent SSM

4. A library of pre-computed PSFs will be made available some time after the launch (PSFlib)

5. ARFs can be computed over narrow energy ranges too…

6. …other options with your host institute? (Grid computing, computing clusters, etc.)

Mitigating the long ARF generation times



1. Divide a few regions of interest in a smart way (e.g. regions with similar special features, "optimize" the SSM, etc.)…
2. Treat each pixel separately vs "the rest" (modelled as a block)
3. Ignore regions with virtually inexistent SSM
4. A library of pre-computed PSFs will be made available some time after the launch (PSFlib)
5. ARFs can be computed over narrow energy ranges too…
6. …other options with your host institute? (Grid computing, computing clusters, etc.)

Advice and recommendations

If you plan to analyze a source…

- ✓ **Your favorite extended source will (very) likely be affected by SSM!**
- ✓ **Systematic uncertainties** (SSM, background, effective areas, calibration, etc.)
- ✓ Substantial SSM is **not necessarily "bad"** (see above example with ~50% purity)
- ✓ Data analysis will **NOT be a simple XSPEC fit**! (e.g. investigation of spectral signatures of SSM before fitting - see Anna's talk)
- ✓ Recommended to plan a coherent **tiling of regions** to investigate
- ✓ "Brute-force" multi-pixel array spectroscopy is challenging…

If you plan to propose a source…

1. What is the **spatial (and spectral) structure** of the extended source?
   (clues from Chandra / XMM, previous literature, etc.)

2. How is SSM likely to impact the **analysis** of the proposed source?

3. How is SSM expected to impact the **proposed science goals**?

4. Will the source be analyzed over several **spatial regions**? If so, which tiling strategy (and why)?

If you plan to propose a source…

5. What **method(s)** is/are thought to be the most appropriate to tackle internal SSM effects? (e.g. computing precise ARFs directly, using PSFlib, using an alternative or hybrid method, etc.)

6. Is **external SSM** expected to impact the analysis (and output science) of the source?

7. **Other properties** of the source expected to further complicate the analysis? (e.g. pile-up effects, uncertainties in atomic databases impacting spectral features relevant to the science goals, etc.)

Remember…

✓ We are entering a **new era**… lots of challenges and **lessons to learn** (exciting on this aspect too!)

Remember…

✓ We are entering a **new era**… lots of challenges and **lessons to learn** (exciting on this aspect too!)