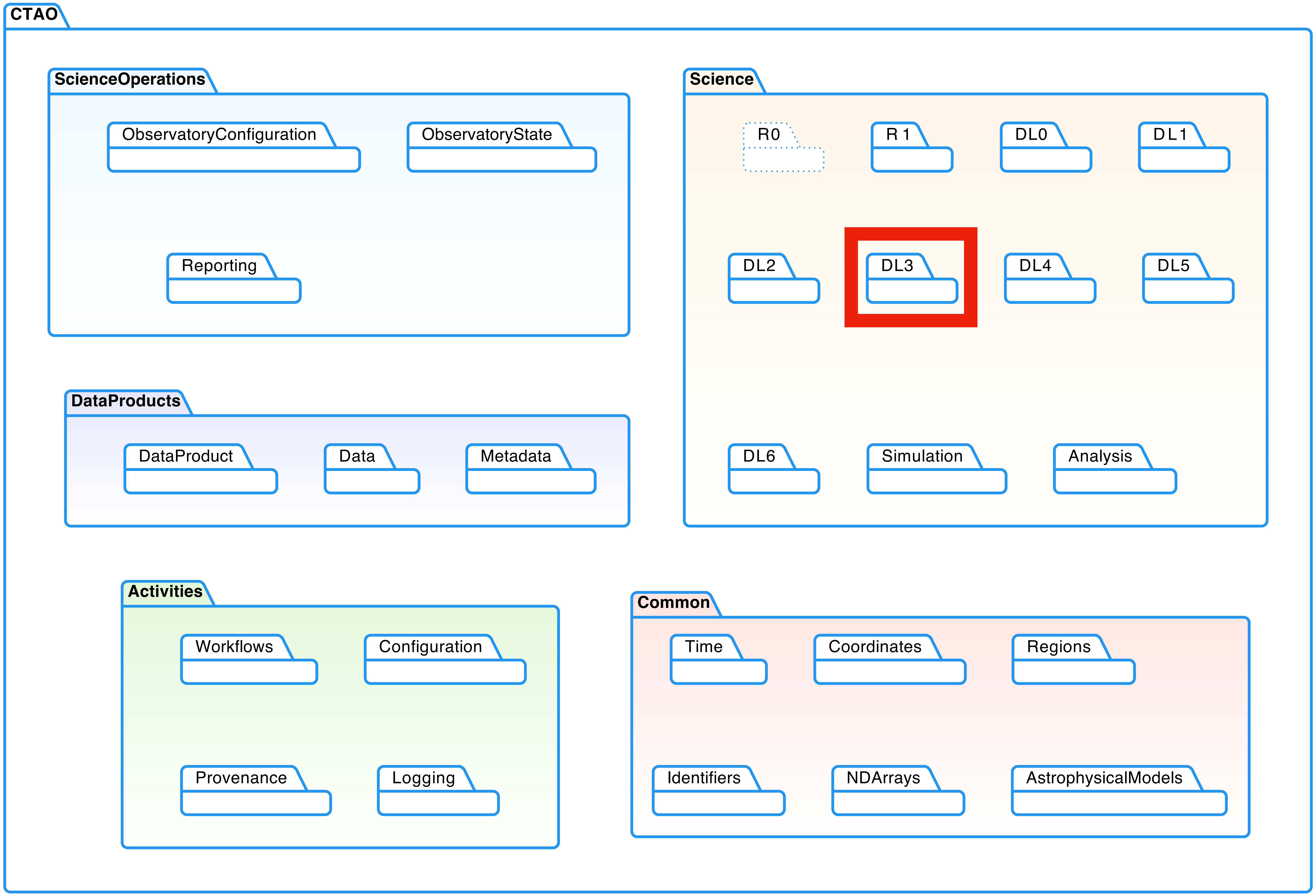
IRFs for High-energy Astrophysics

In the context of the CTAO Data Model

Karl Kosack
for the CTAO Data Model Group

IVOA HEIG meeting, online 2023-10-18



CTAO top-level data model draft_2a_dr1-43-gd389bbb

Event-counting instruments

"Events" here are detections of a single photon or background particle

≈ 10,000 / s for CTAO! (gammas + background) in the raw data

Measure many *reconstructed* properties per event:

- Physical:
  - Position on the sky (RA/Dec, Alt/Az)
  - Energy
  - Arrival Time
  - + uncertainty estimates on these
- *Optional* Instrumental / Shower physics
  - telescope multiplicity
  - Impact parameter w.r.t. telescopes
  - Height of Shower-maximum
  - Width of shower (Molière radius)
  - "Gammaness" (probability to be signal)
  - Quality of reconstruction

CTAO DL3 Data Model Specification, v1b

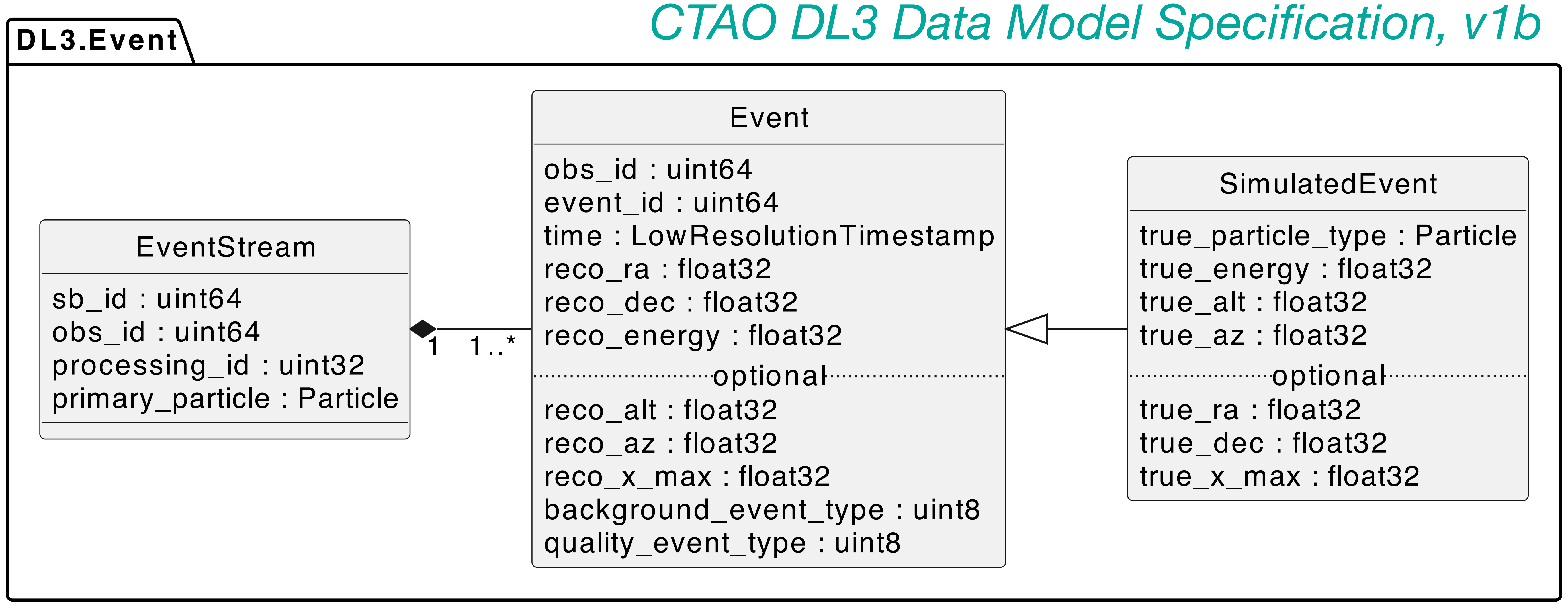


Figure 2.2 – UML Diagram of the DL3/Event Data Model.

Event-lists *not* Photon-lists!

Event-lists *not* Photon-lists!

For Imaging Atmospheric Cherenkov Telescopes
(and also Water Cherenkov Telescopes):

- Strong **irreducible background** from **mis-reconstructed cosmic rays**
- Mostly isotropic (except for instrumental effects)

Event-lists *not* Photon-lists!

For Imaging Atmospheric Cherenkov Telescopes
(and also Water Cherenkov Telescopes):

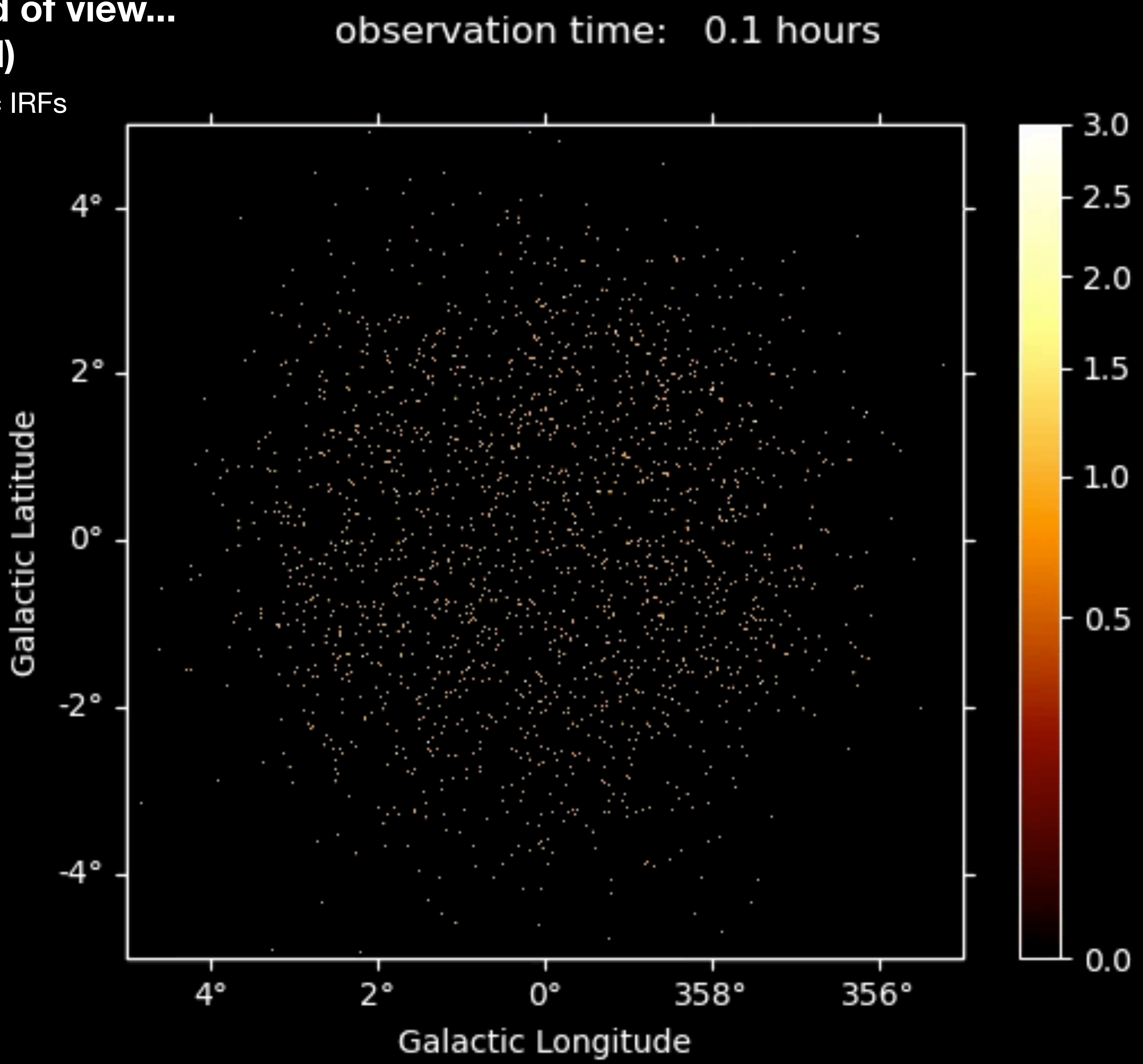
- Strong **irreducible background** from **mis-reconstructed cosmic rays**
- Mostly isotropic (except for instrumental effects)

Photons are only statistical!

- Can compute "Excess photon-like events" within a region of space/time/energy.
- Background model is needed (or off-source measurement in some cases)
- Can **never** say if an individual event is a photon or not.

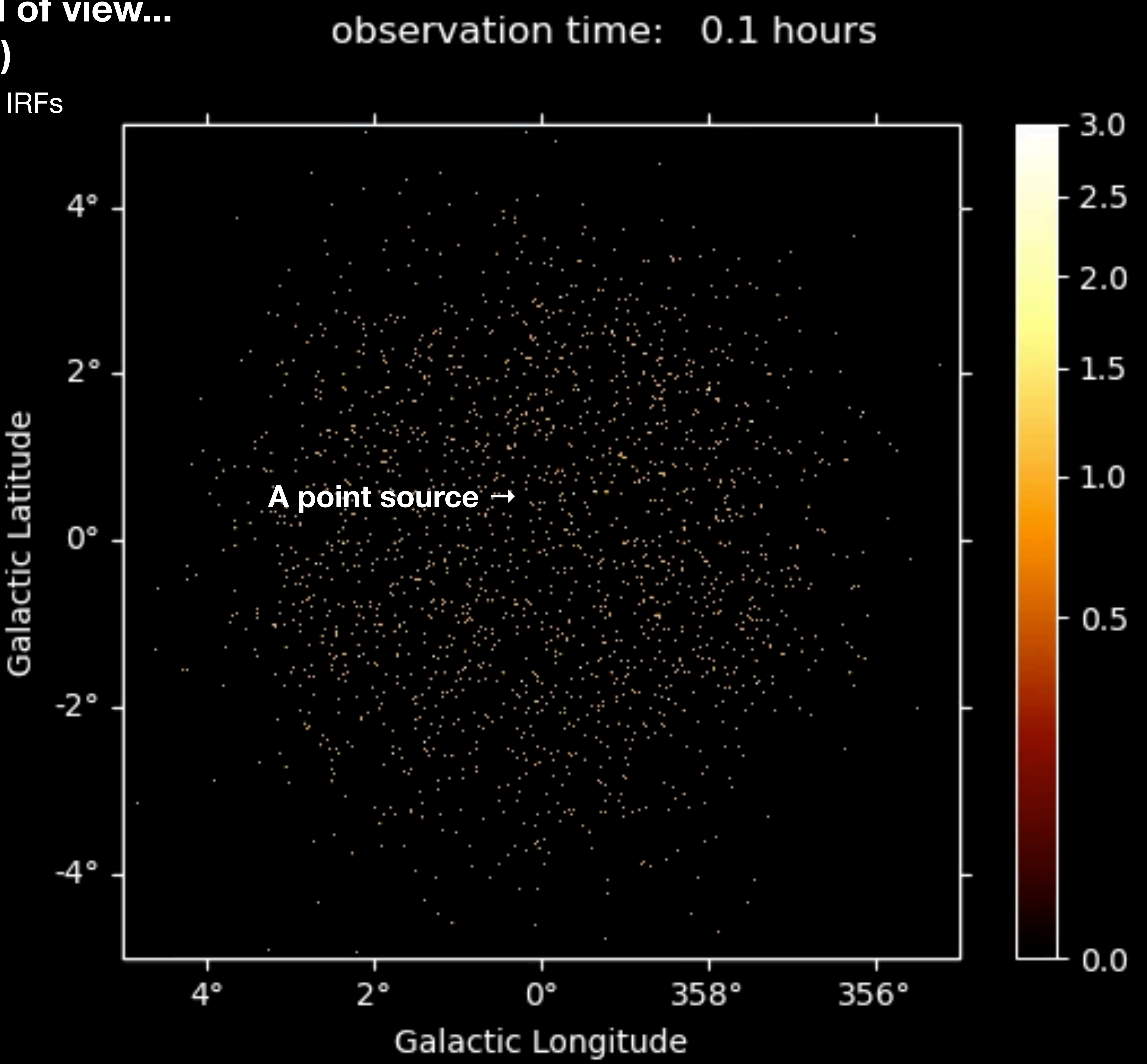
A mostly blank field of view... (simulated)

CTA-South public IRFs



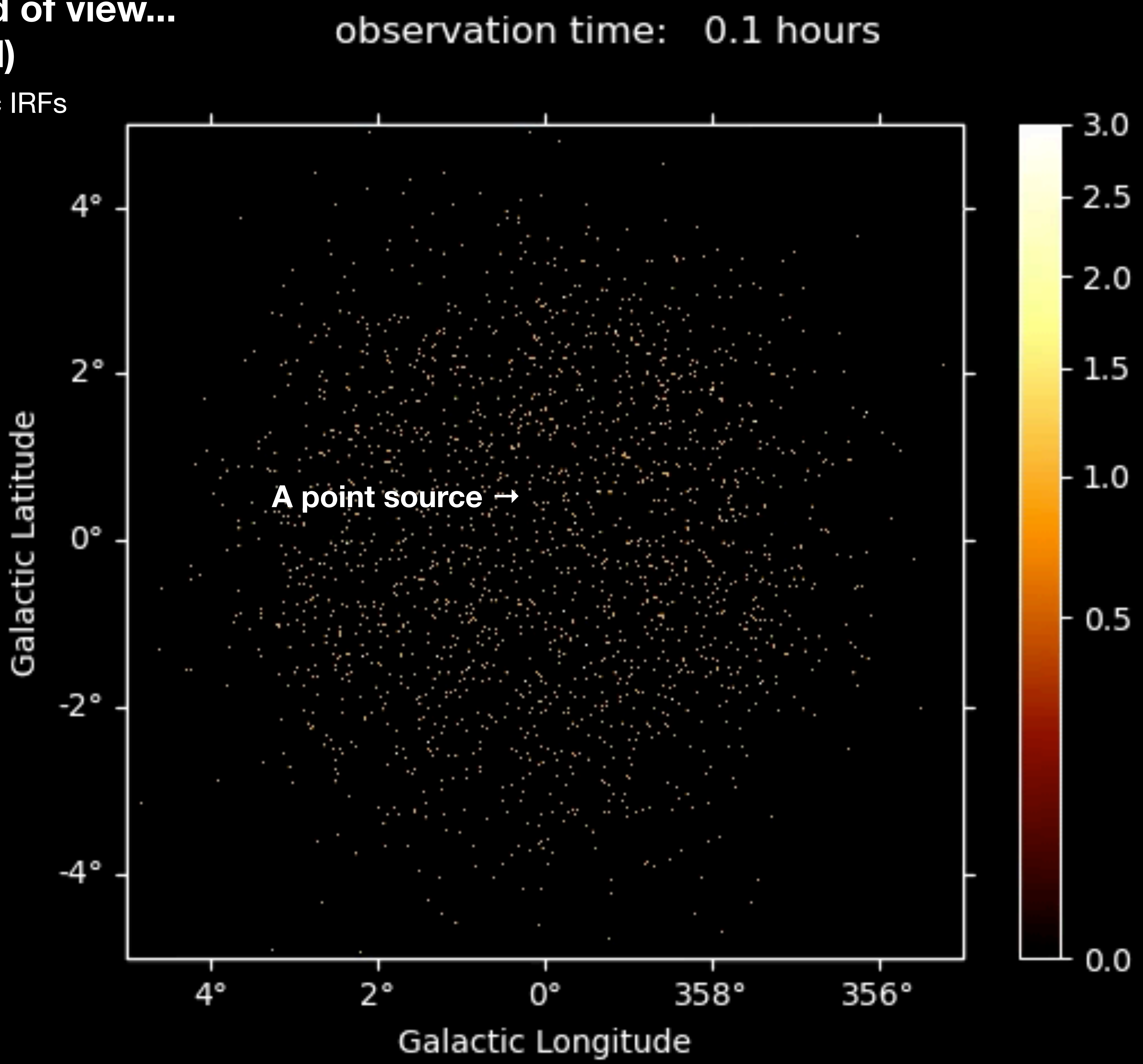
A mostly blank field of view...
(simulated)

CTA-South public IRFs



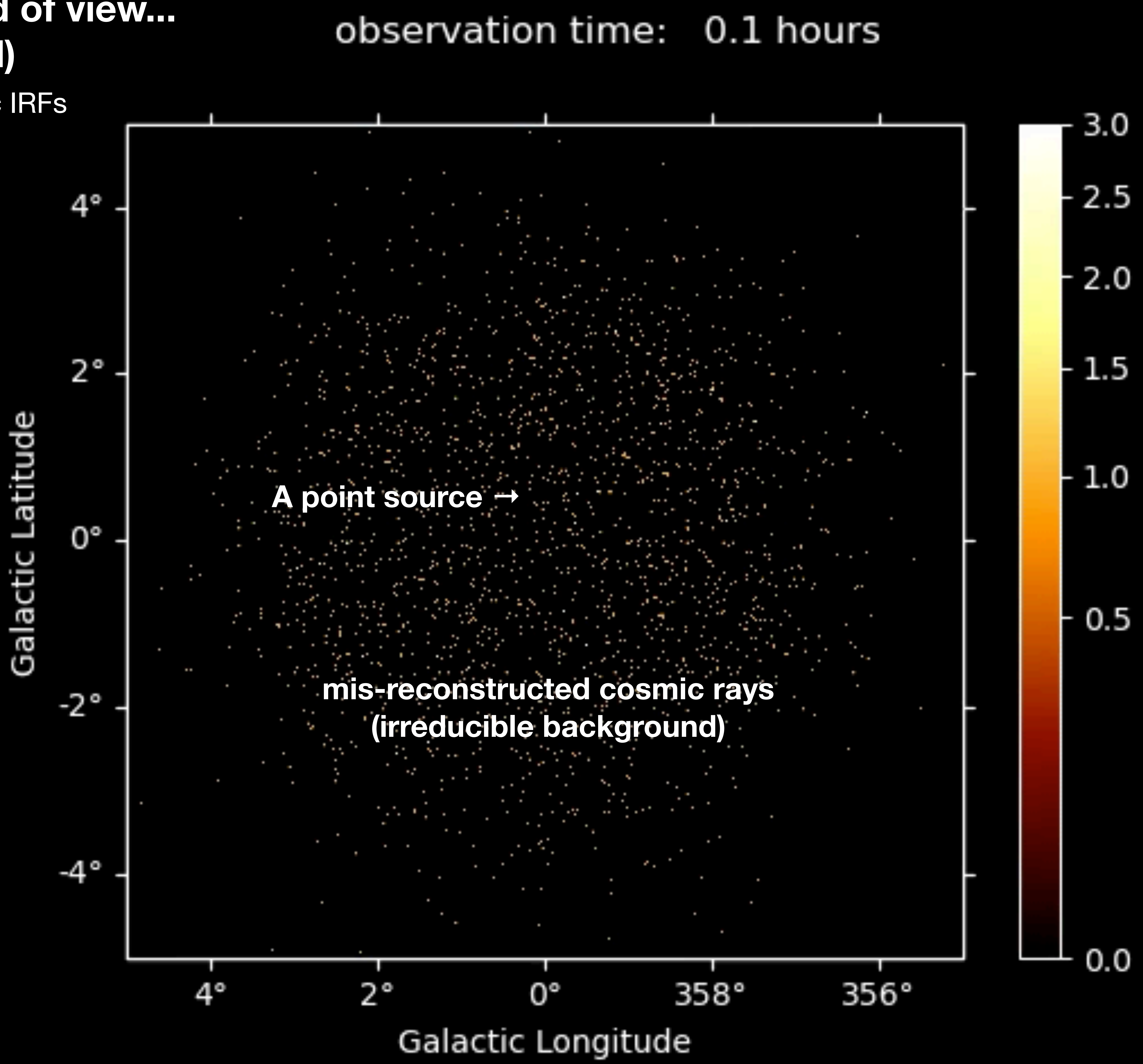
A mostly blank field of view…
(simulated)

CTA-South public IRFs



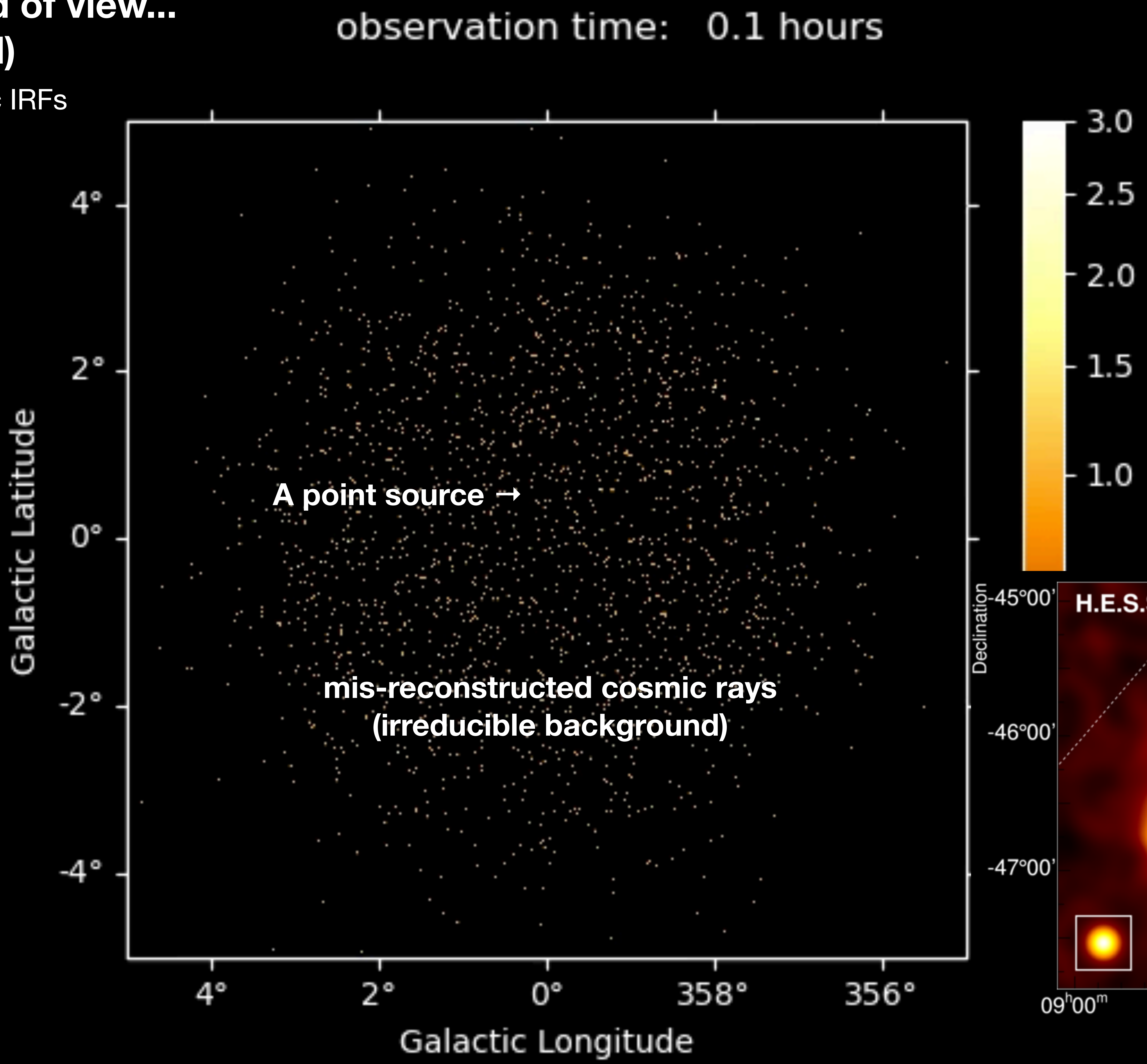
A mostly blank field of view...
(simulated)

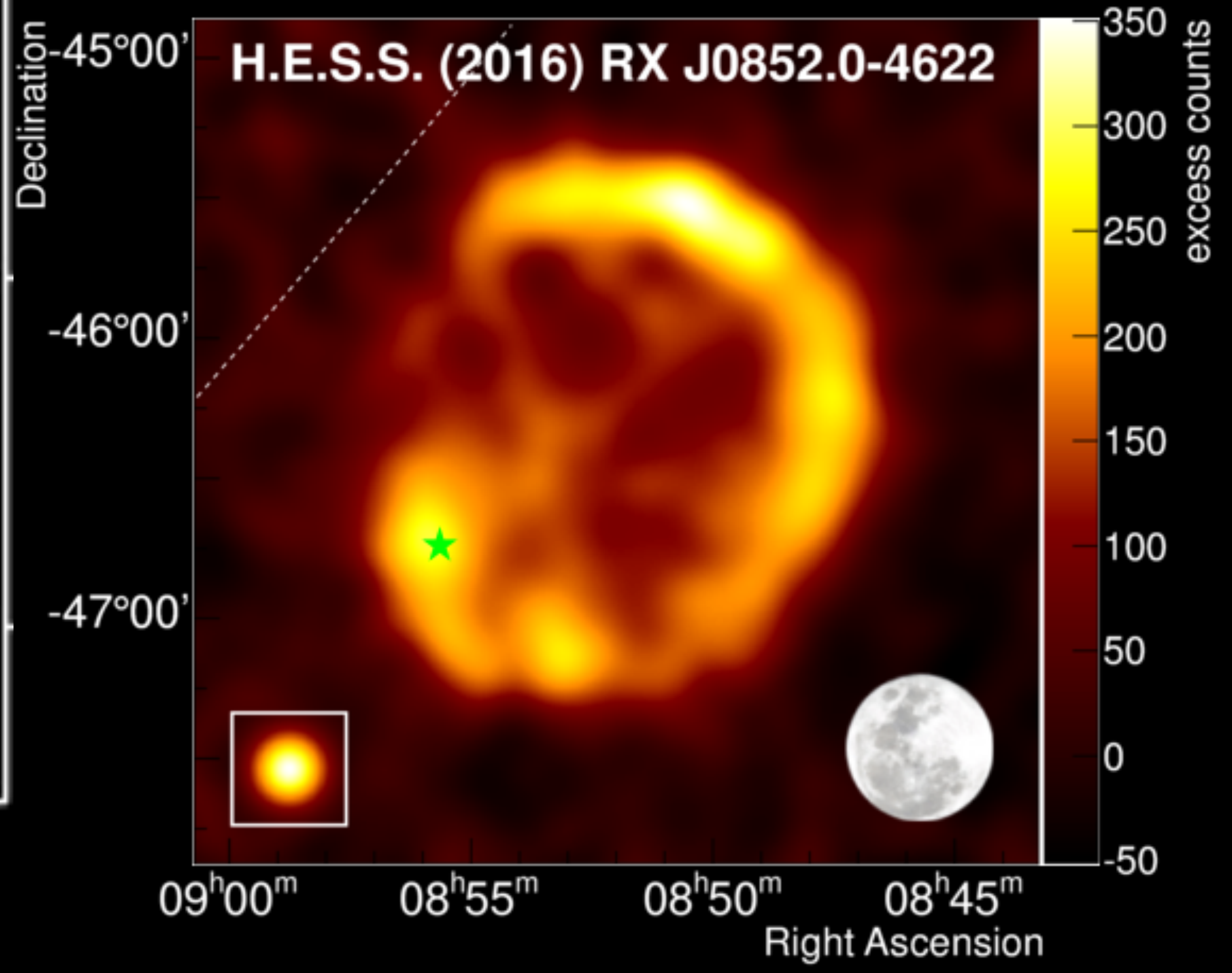
CTA-South public IRFs



A mostly blank field of view… (simulated)

CTA-South public IRFs





What we *really* want:

To make a **hypothesis** about the gamma-ray emission in a region of the sky and to **test that hypothesis**.

Define Source Model

Compare Model to Data

Physical Quantities

$F(E_{\mathrm{true}}, \vec{p}_{\mathrm{true}}, t_{\mathrm{true}})$

Flux

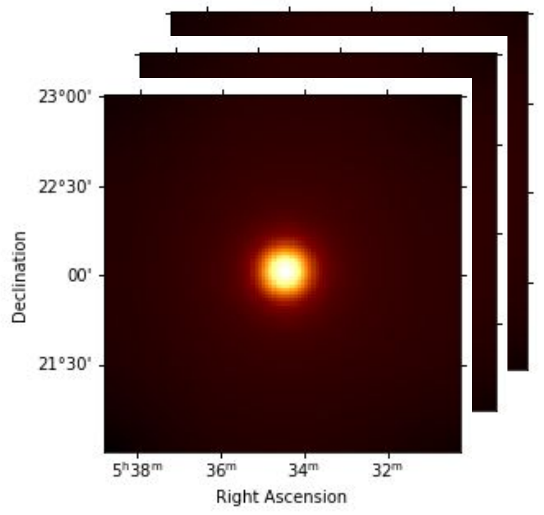
Reconstructed Quantities

$N_{\mathrm{events}}(E_{\mathrm{reco}}, \vec{p}_{\mathrm{reco}}, t_{\mathrm{reco}})$

Event Counts
within some bin or region

We are missing one piece of information:
how to go between true and measured (reconstructed) quantities?

Forward-Folded Analysis



1D: spectum or lightcurve
2D: image or spectral-evolution
3D: image-energy cube
4D: image-energy-time cube

Source model

×

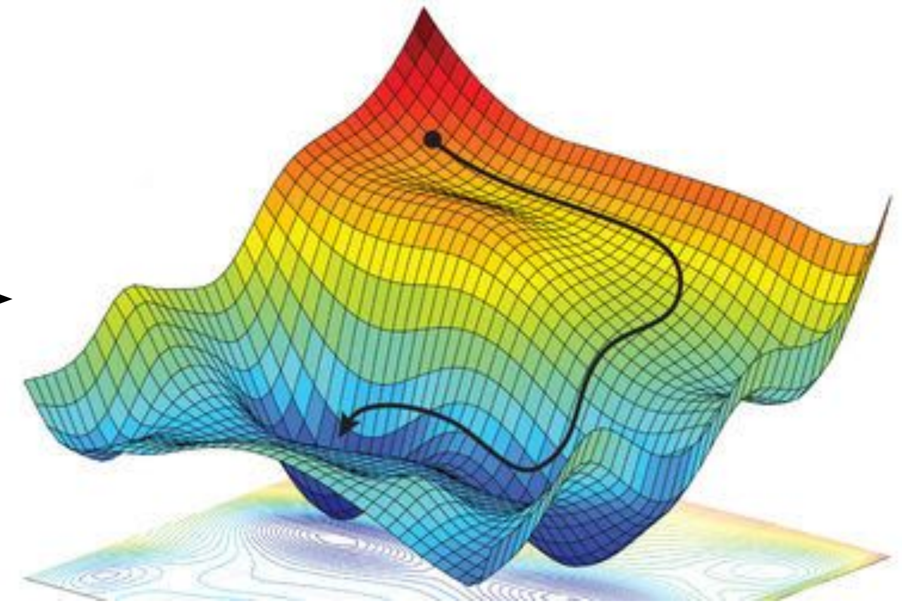
IRFs

+

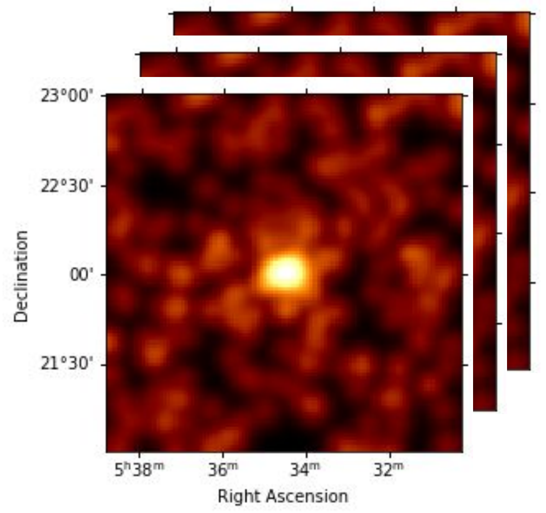
Background model

Predicted counts cube

$\mathscr{L}$ →



Data cube



Hypothesis testing

$$\mathrm{TS} = -2\ln\left(\frac{\mathscr{L}^{\max}_{\text{null}}}{\mathscr{L}^{\max}_{\text{alt.}}}\right)$$

Luca Giunti, thesis presentation 2021

Forward-Folded Analysis

Source model

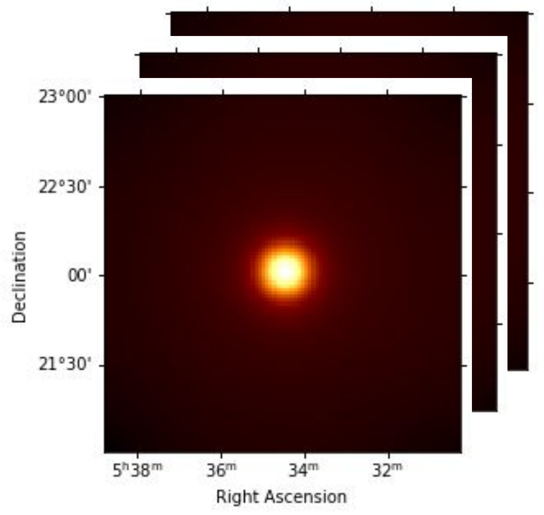
×

IRFs

+

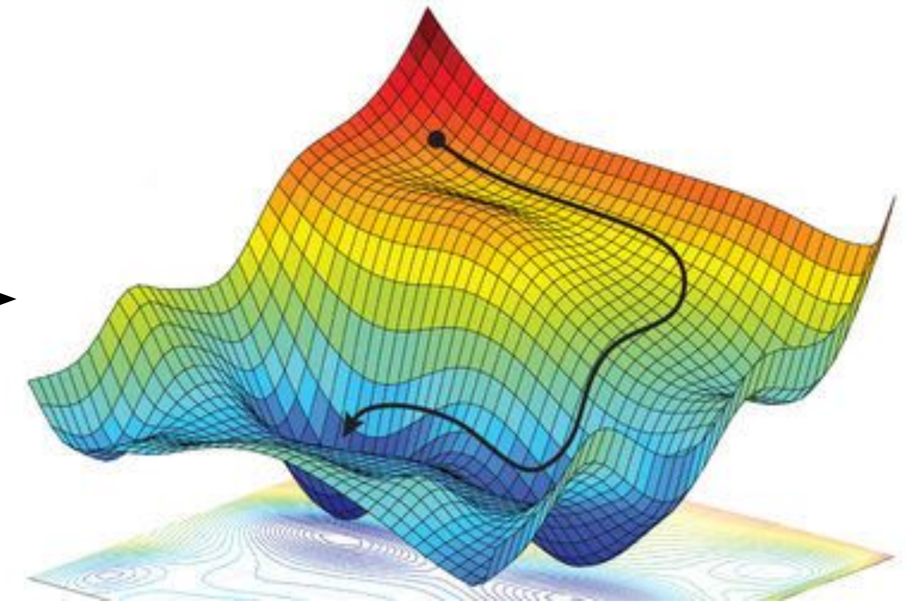
Background model

Predicted counts cube

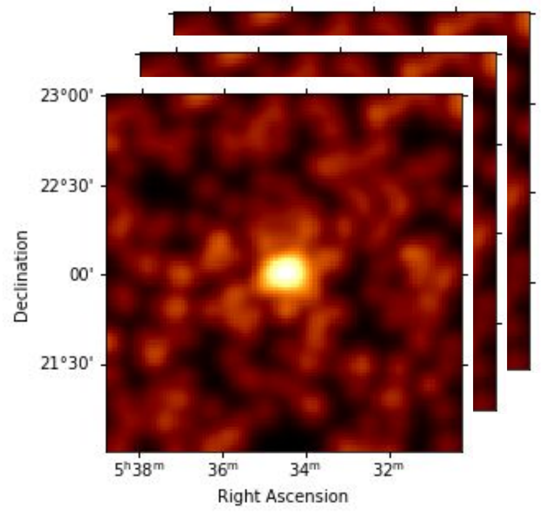


1D: spectum or lightcurve
2D: image or spectral-evolution
3D: image-energy cube
4D: image-energy-time cube

$\mathscr{L}$ →



Data cube



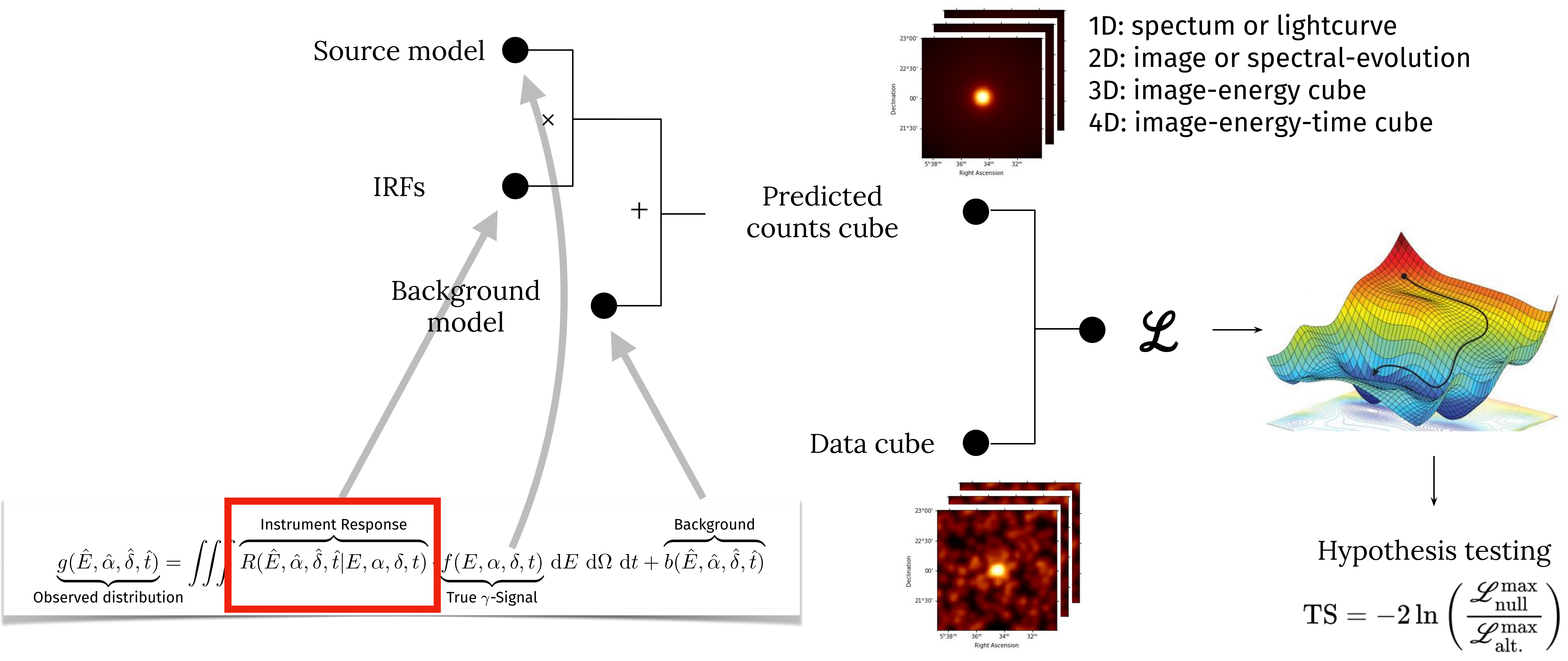
$$\underbrace{g(\hat{E},\hat{\alpha},\hat{\delta},\hat{t})}_{\text{Observed distribution}} = \iiint \overbrace{R(\hat{E},\hat{\alpha},\hat{\delta},\hat{t}|E,\alpha,\delta,t)}^{\text{Instrument Response}} \cdot \underbrace{f(E,\alpha,\delta,t)}_{\text{True }\gamma\text{-Signal}}\,\mathrm{d}E\,\mathrm{d}\Omega\,\mathrm{d}t + \overbrace{b(\hat{E},\hat{\alpha},\hat{\delta},\hat{t})}^{\text{Background}}$$

Hypothesis testing

$$\mathrm{TS} = -2\ln\left(\frac{\mathscr{L}^{\max}_{\text{null}}}{\mathscr{L}^{\max}_{\text{alt.}}}\right)$$

Luca Giunti, thesis presentation 2021

Forward-Folded Analysis



Luca Giunti, thesis presentation 2021

Instrumental Response Decomposition

$$R(\hat{E},\hat{\alpha},\hat{\delta}|E,\alpha,\delta,t) = \underbrace{A_{\text{eff}}(E,\alpha,\delta,t)}_{\text{Effective Area}} \cdot \overbrace{M(\hat{E}|E,\alpha,\delta,t)}^{\text{Energy Migration}} \cdot \underbrace{\text{PSF}(\hat{\alpha},\hat{\delta}|E,\alpha,\delta,t)}_{\text{Point Spread Function}} .$$

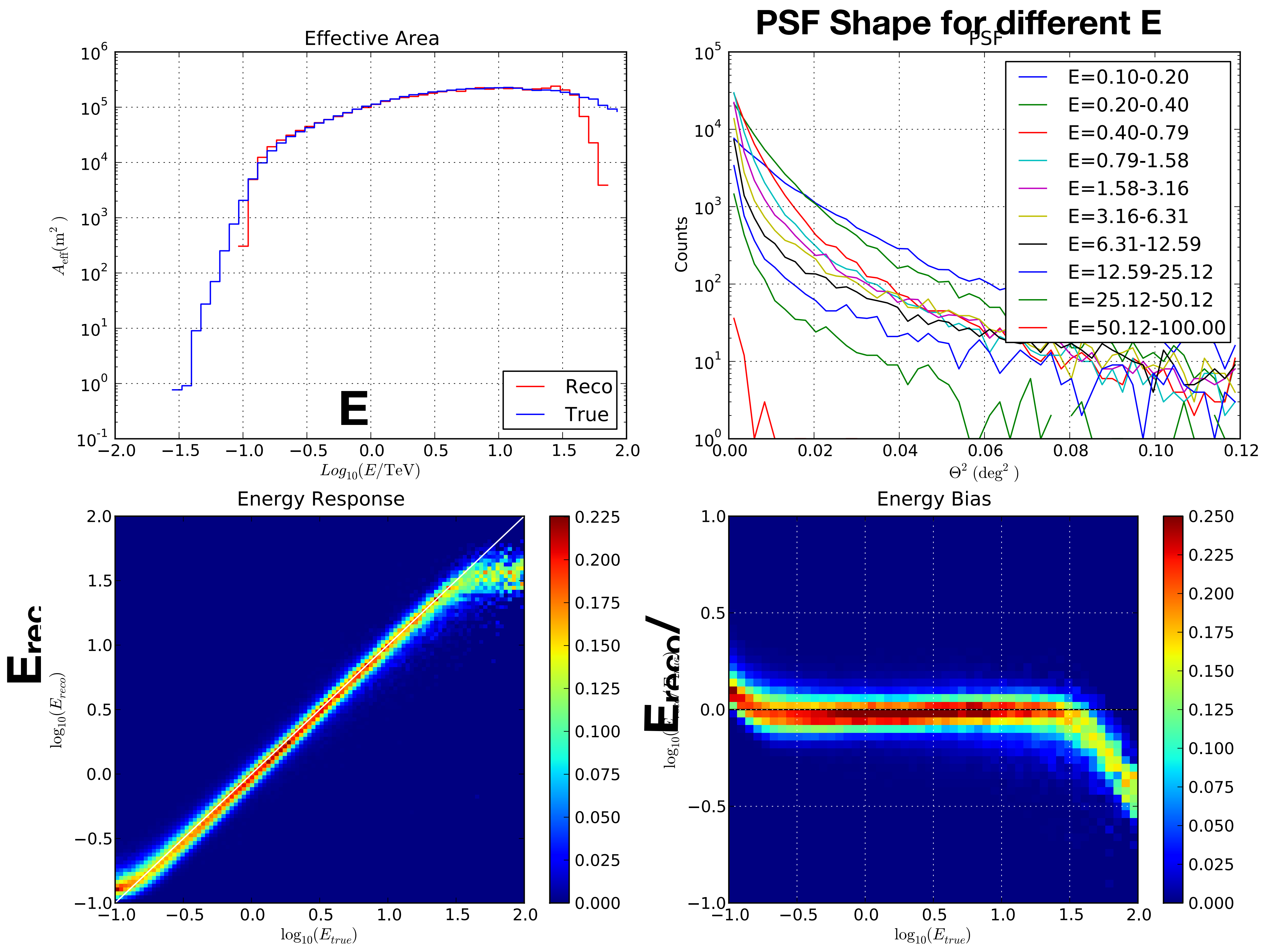
+ cross terms!
Currently not treated

Treat components independently:

- **Effective Collection Area** (A_{eff}): combines optical effective area and efficiency of detection **of photons**

$$A_{eff} = \frac{N_{det}}{N_{sim}} A_{sim}$$

- **Point-Spread function** (PSF): probability of reconstructing event at position **p'** given true position **p** and true energy E .
- **Energy Dispersion**/Migration (Edisp): Probability of reconstructing energy E' given true energy E
- **Background Rate**: rate of expected **background** at reconstructed position P' and reconstructed energy E')



Examples

Outputs from gammapy

Event List

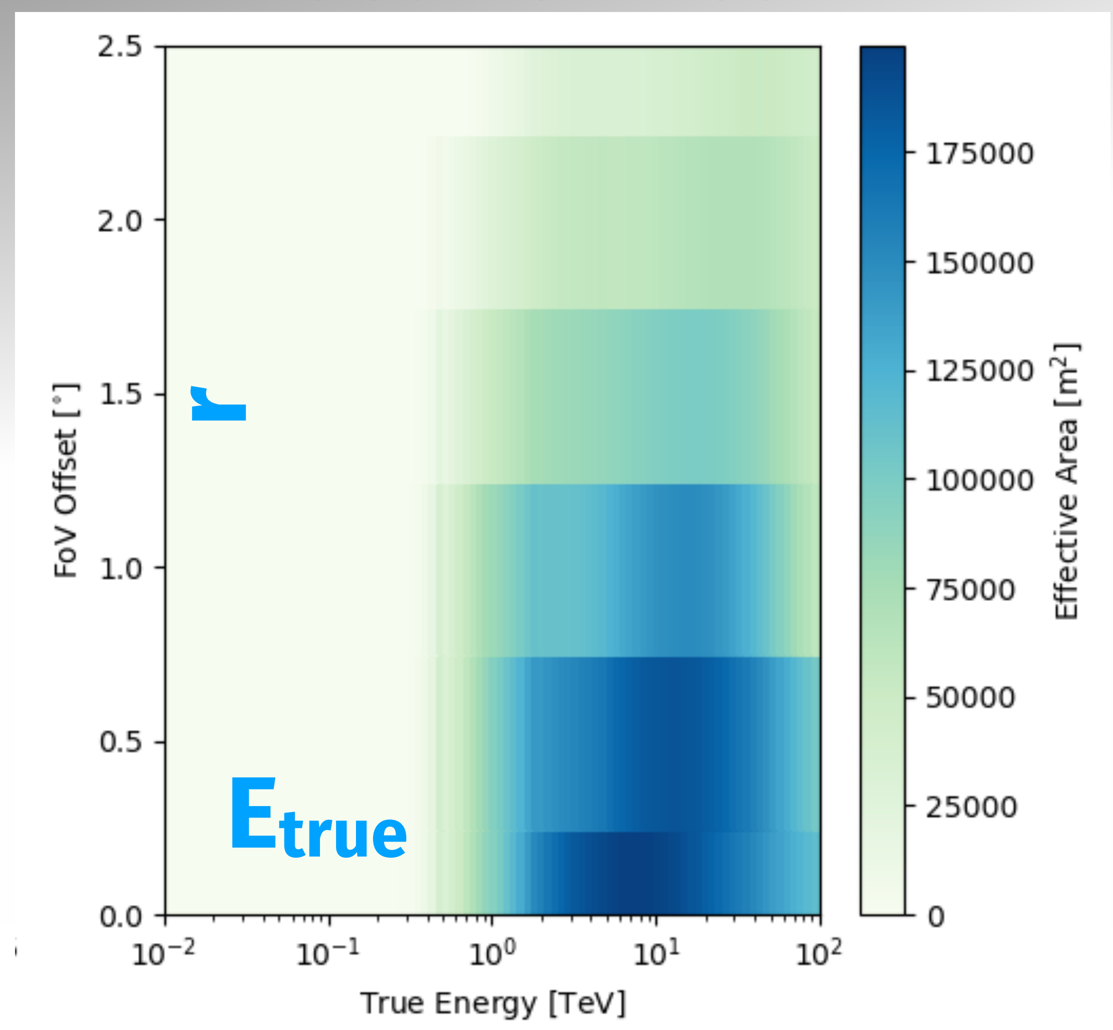
```
[5]: tab = Table.read('hess_events_085356.fits.gz')
```

```
[6]: tab
```

[6]: *Table length=499*

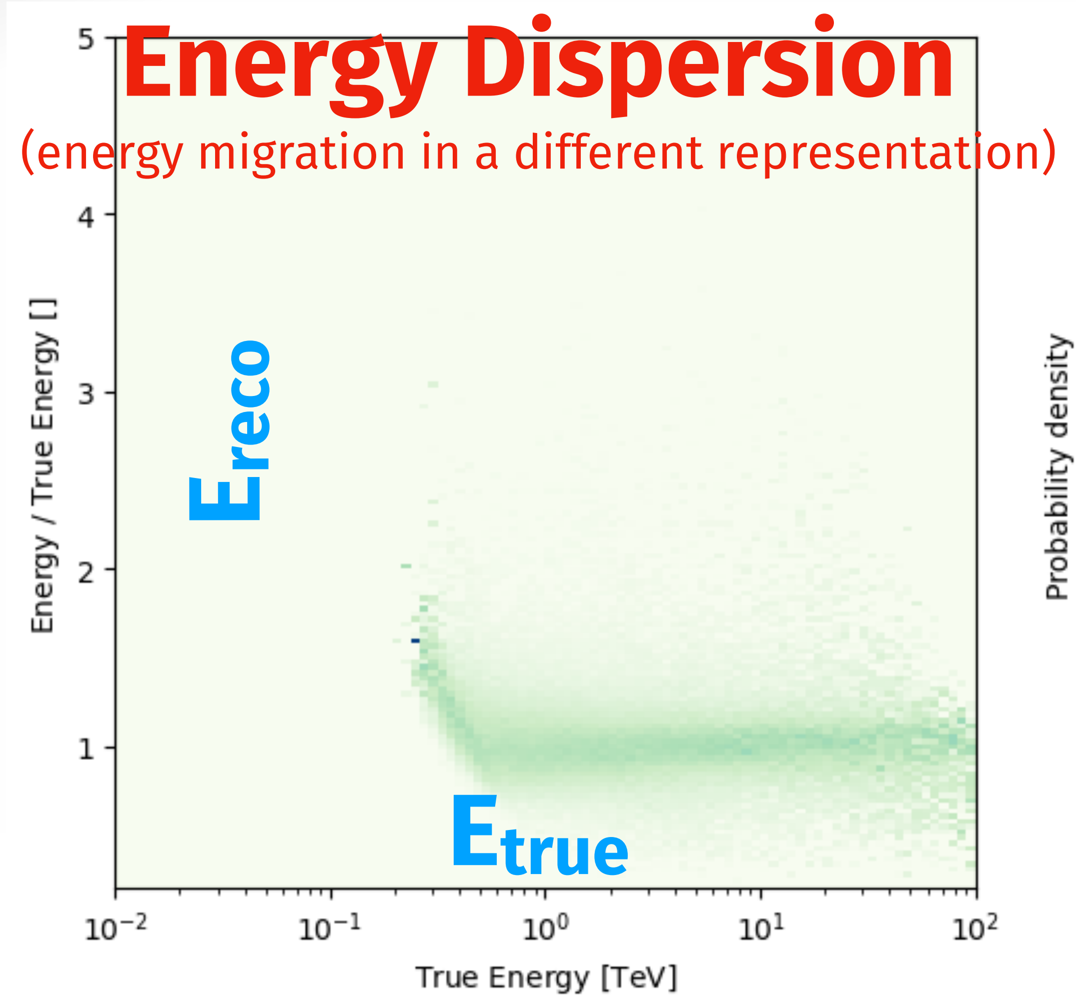
EVENT_ID	OBS_ID	TIME	MULTIP	RA	DEC	DIR_ERR	SKYX_RADEC	SKYY_RADEC
		s		deg	deg	deg	deg	deg
int64	uint32	float64	int16	float32	float32	float32	float32	float32
24189255811677	85356	392695107.16595197	3	329.76013	-28.378927	0.0	-0.8467294	0.03804477
24193550778756	85356	392695109.406718	2	329.7613	-28.11467	0.0	-1.1110648	0.03915714
24197845746261	85356	392695114.1640556	2	329.71042	-28.54954	0.0	-0.6760932	-0.0056832614
24202140713433	85356	392695117.68513155	2	329.14612	-26.46658	0.0	-2.7600195	-0.5115584
24206435680287	85356	392695119.8251052	2	330.07318	-30.546444	0.0	1.3215607	0.30692154
24210730647732	85356	392695124.4254937	2	331.64767	-29.174904	0.0	-0.03684649	1.6862341
24215025615341	85356	392695129.84283924	2	328.74268	-28.294678	0.0	-0.927548	-0.85798883
24215025615494	85356	392695130.539845	2	330.48044	-27.899296	0.0	-1.3244383	0.6749976
24215025615622	85356	392695131.2239411	3	327.54605	-29.050547	0.0	-0.15758944	-1.8983405

Effective Area

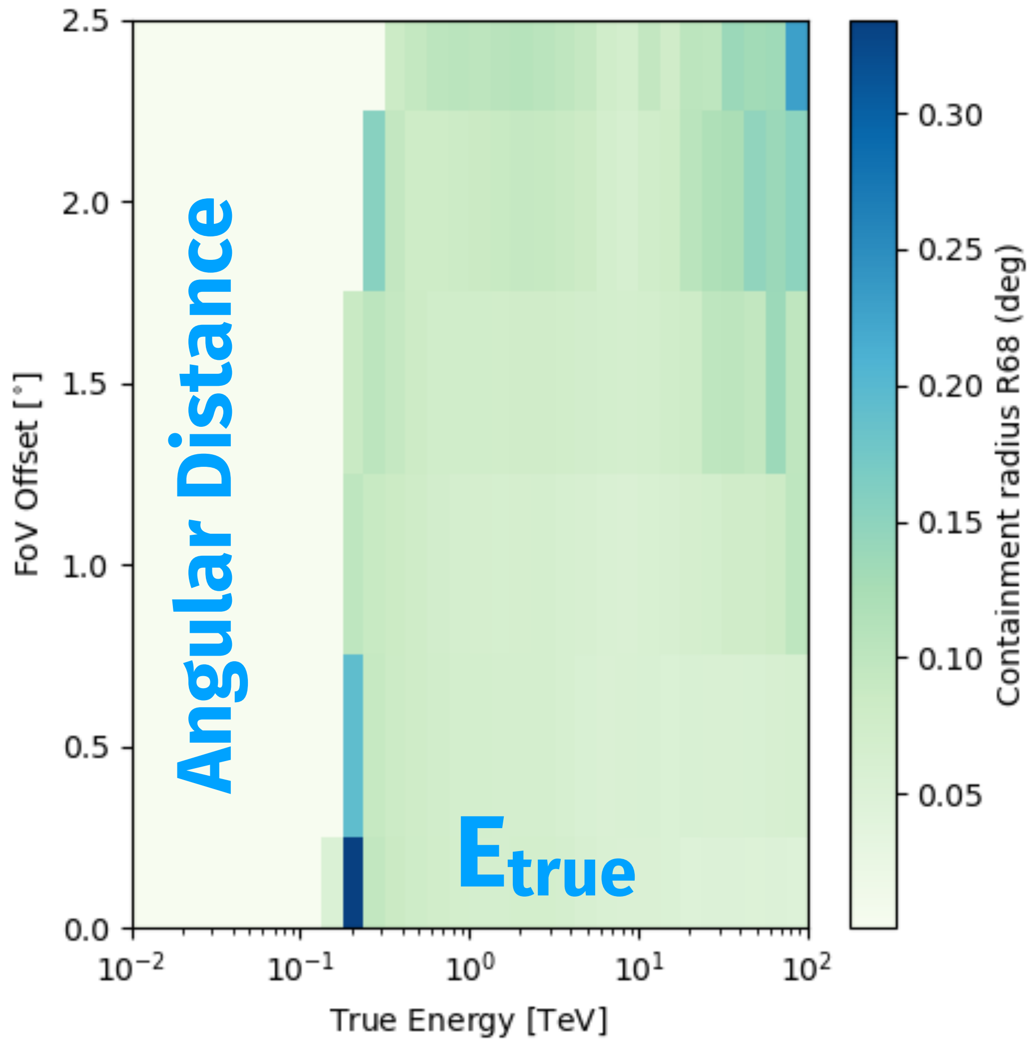


Energy Dispersion

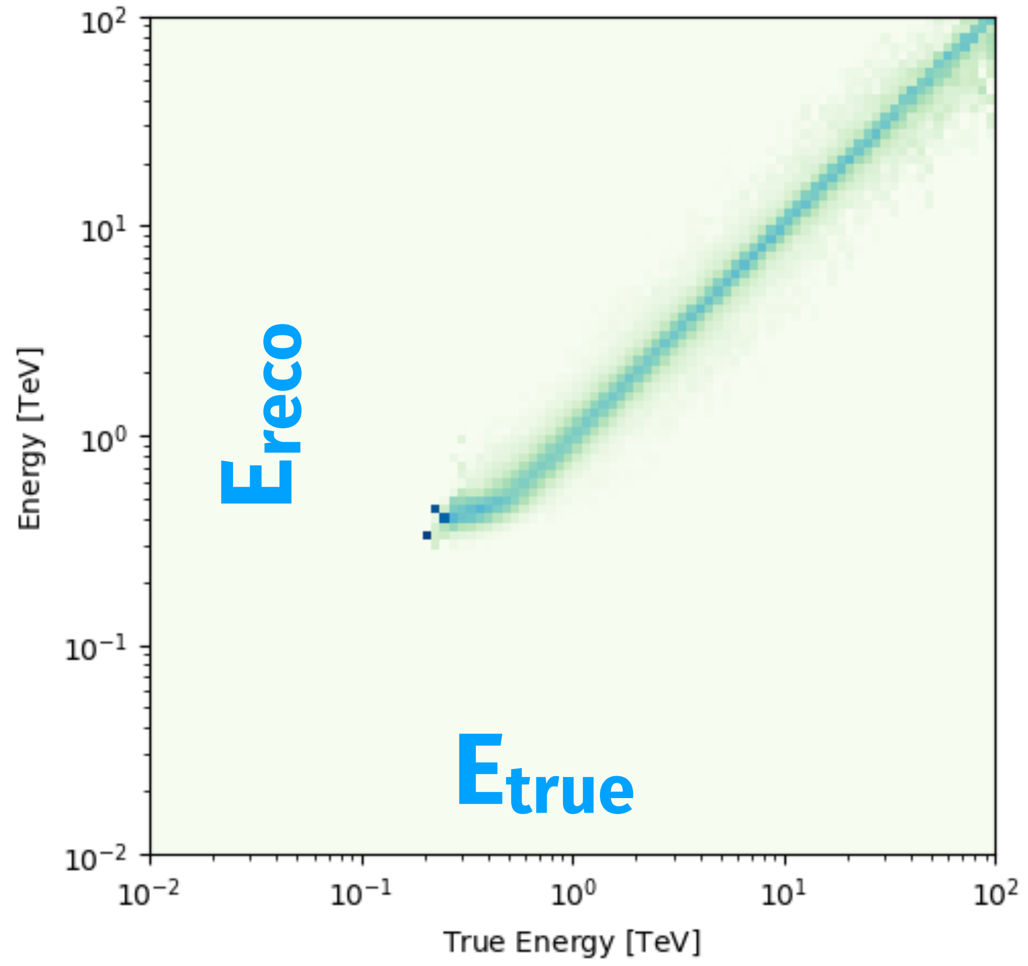
(energy migration in a different representation)



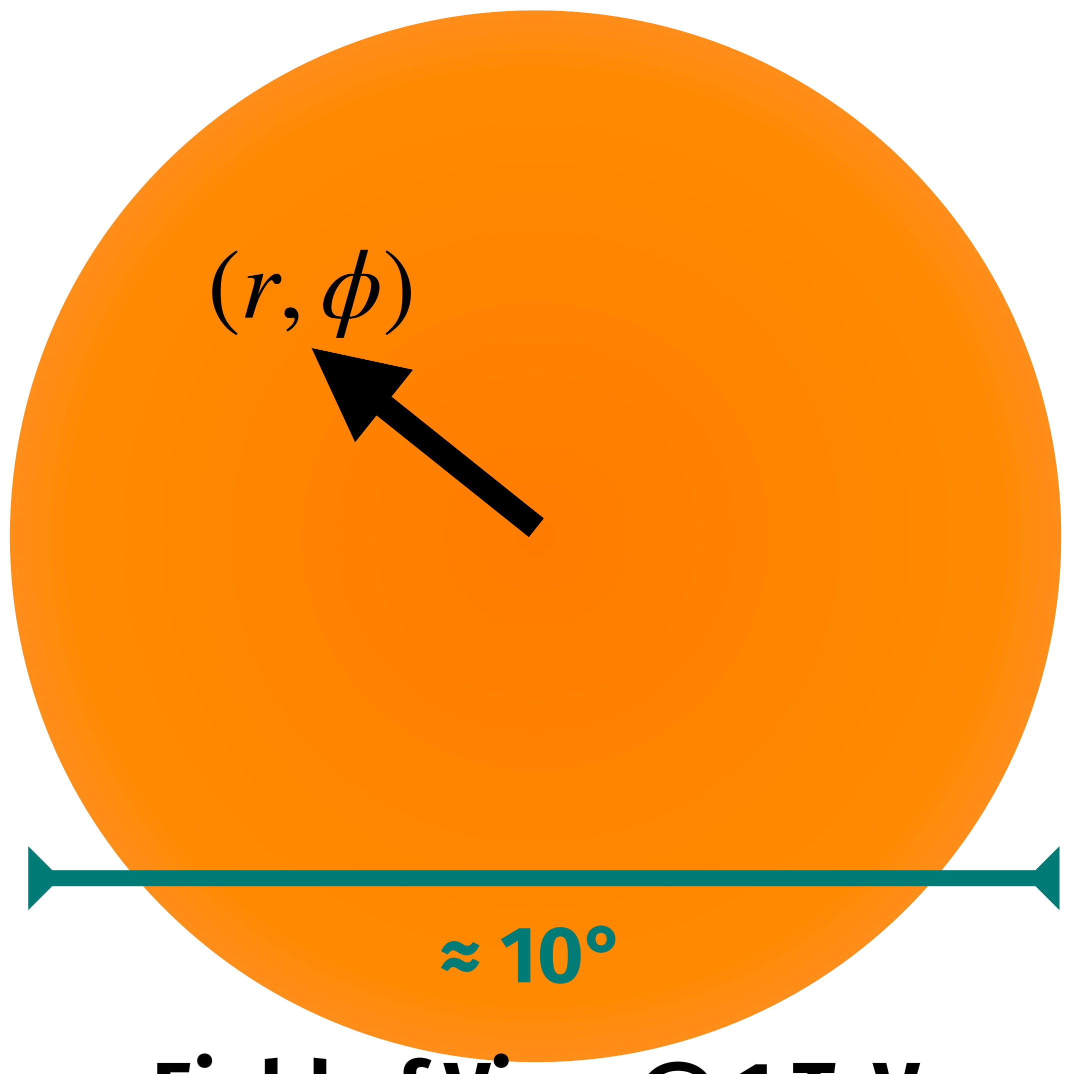
Point-Spread



Energy Migration



IRF Dimensionality

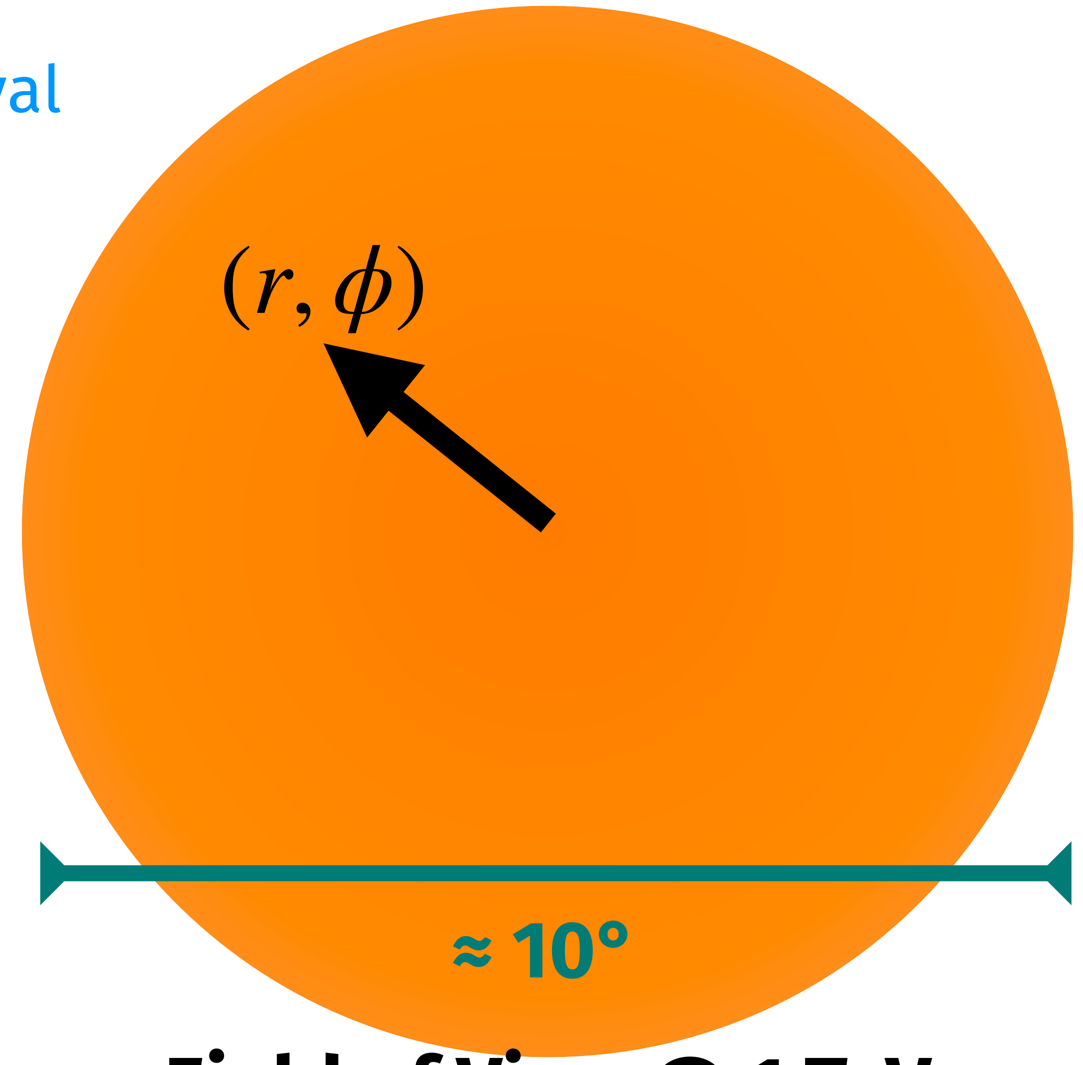


Field of View @ 1 TeV
(Energy dependent!)

IRF Dimensionality

Internal Dimensionality (3D):

- **Position** in Field-of-View (r, ϕ)
- **Energy** → large differences! even for FOV!
- ~~**Time**~~ → Usually taken into account by using 1 IRF per short time interval



Field of View @ 1 TeV
(Energy dependent!)

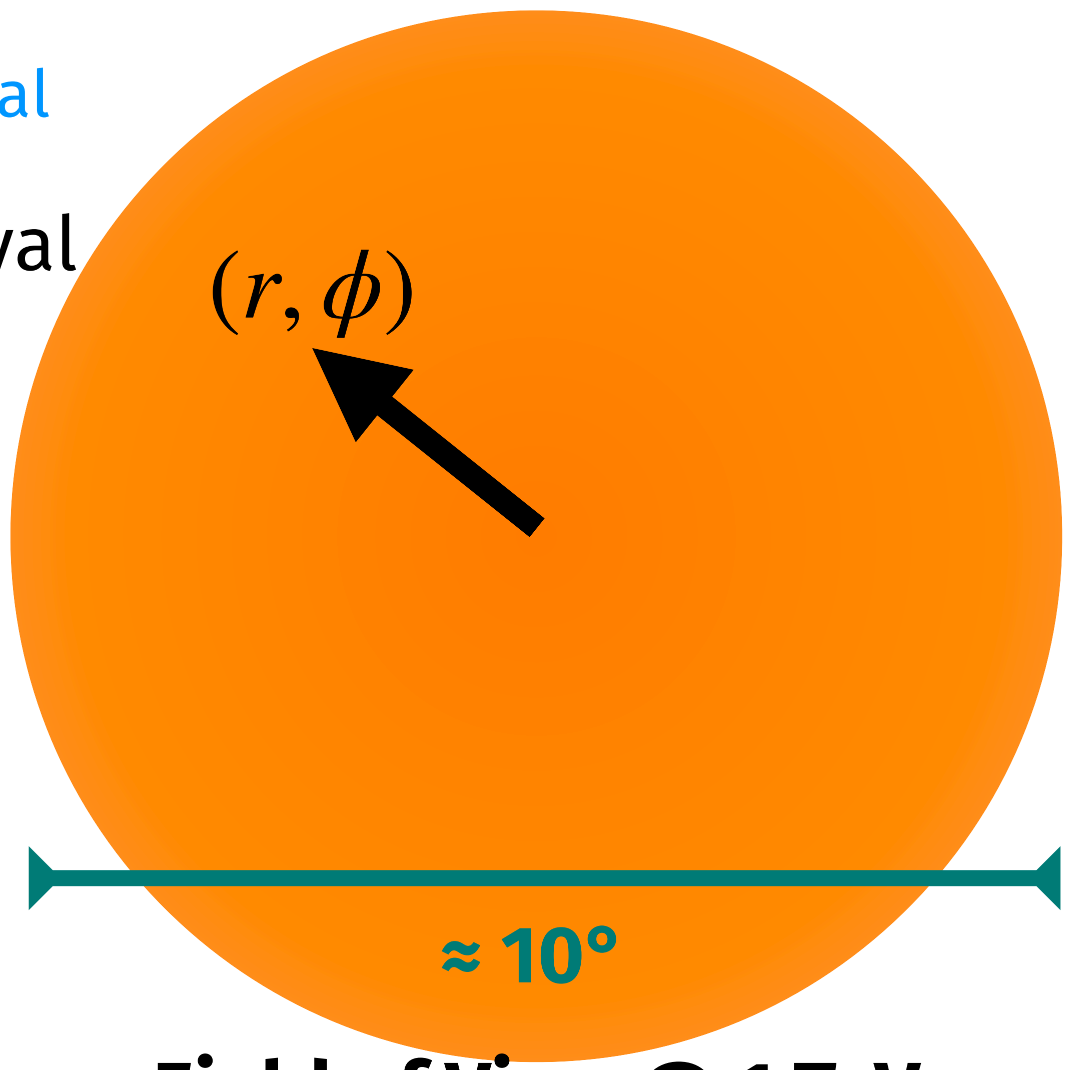
IRF Dimensionality

Internal Dimensionality (3D):

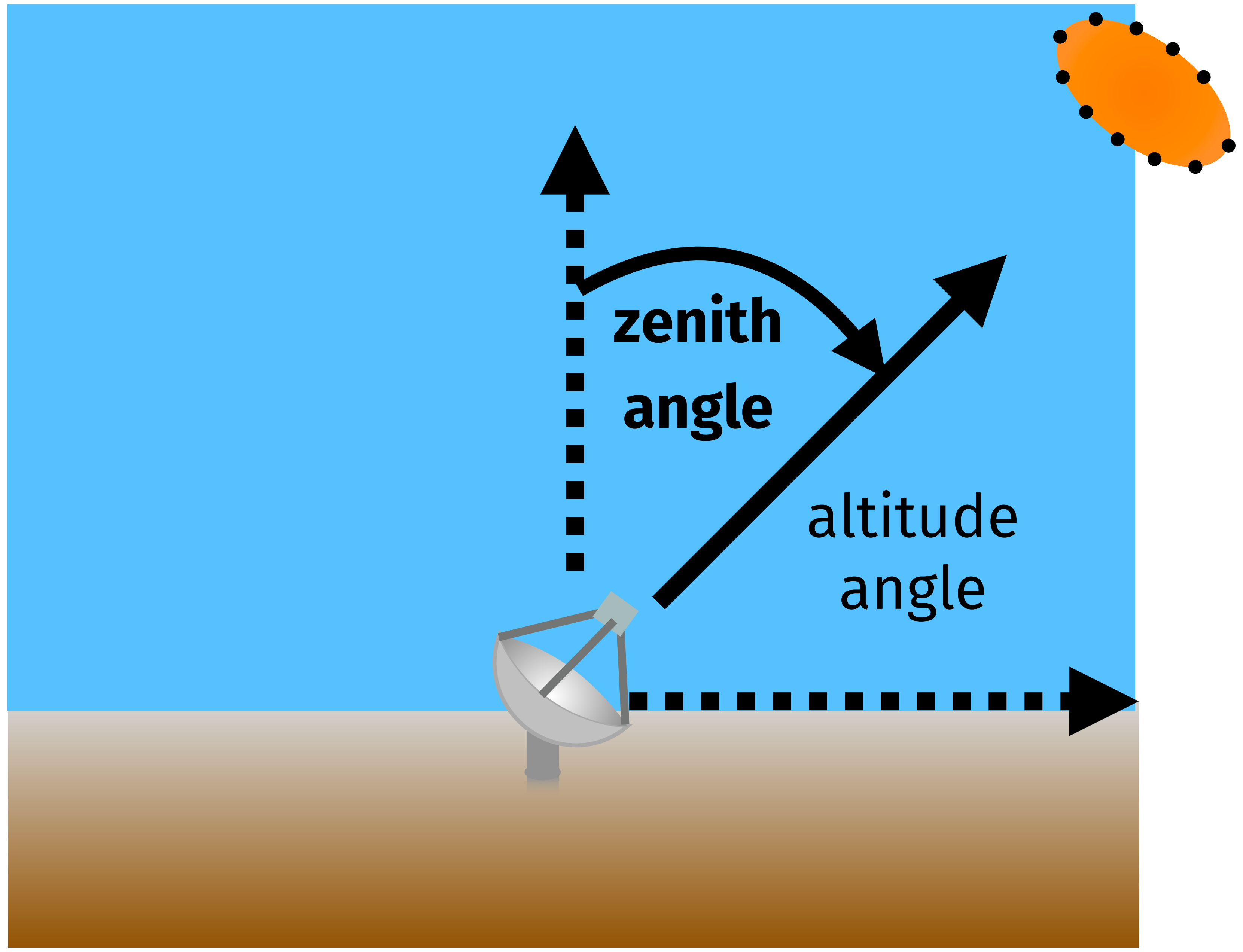
- **Position** in Field-of-View (r, ϕ)
- **Energy** → large differences! even for FOV!
- ~~**Time**~~ → Usually taken into account by using 1 IRF per short time interval

External Dimensionality assumed constant within a stable interval (≈30min):

- **Pointing direction** on AltAz mount (α, ζ) *:
  - altitude angle (atmosphere density)
  - azimuthal angle (magnetic field orientation)
- **Night-Sky-Background** light intensity/distribution
  - Galactic Plane variations, stars (field-rotation!)
  - Zodiacal light
  - Human sources (cities, etc)
- Seasonal **atmosphere** changes, instrumental **aging**



Field of View @ 1 TeV
(Energy dependent!)



* this works fine for HESS/MAGIC/VERITAS with smallish FOV, but for CTAO (>10°) and HAWC/SWGO (>>10°), **pointing direction must be an internal dimension**

IRF Dimensionality

Internal Dimensionality (3D):

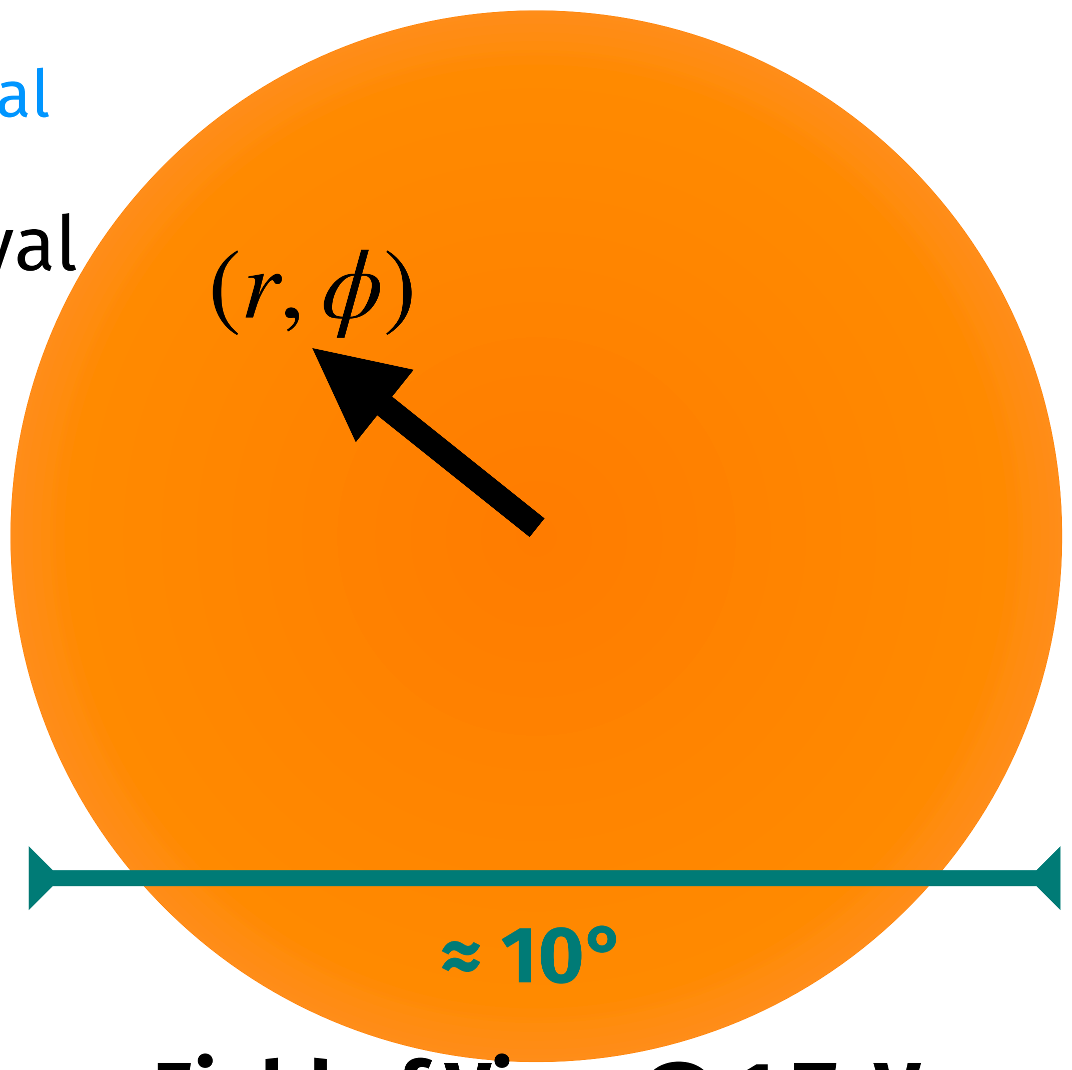
- **Position** in Field-of-View (r, ϕ)
- **Energy →** large differences! even for FOV!
- ~~**Time**~~ **→** Usually taken into account by using 1 IRF per short time interval

External Dimensionality assumed constant within a stable interval (≈30min):

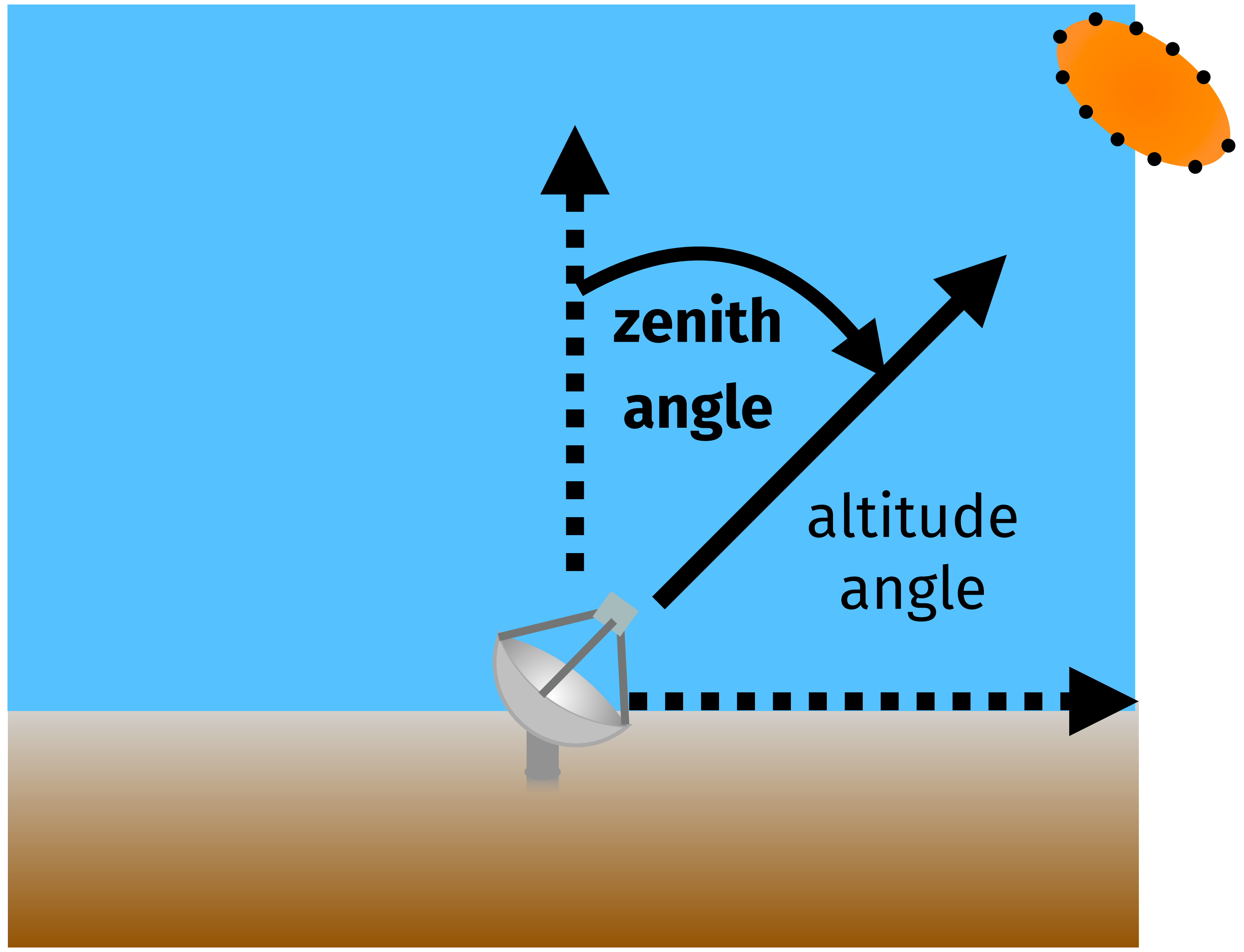
- **Pointing direction** on AltAz mount (α, ζ) *:
  - altitude angle (atmosphere density)
  - azimuthal angle (magnetic field orientation)
- **Night-Sky-Background** light intensity/distribution
  - Galactic Plane variations, stars (field-rotation!)
  - Zodiacal light
  - Human sources (cities, etc)
- Seasonal **atmosphere** changes, instrumental **aging**

Non-continuous Dimensions

- **Event Type** (reconstruction quality, background quality)
- **Subarray Choice**
- **pointing mode** (tracking, drifting, moon)



Field of View @ 1 TeV
(Energy dependent!)



* this works fine for HESS/ MAGIC/VERITAS with smallish FOV, but for CTAO (>10°) and HAWC/SWGO (>>10°), **pointing direction must be an internal dimension**

Important note:

X-ray an analysis also uses forward-folding

- PI channel ≈ E_{reco}
- **ARF** ≈ 1D effective area IRF (with assumed source region), ours more like *instmap*?
- **RMF** ≈ energy dispersion IRF, but we often store as log(Etrue/Ereco) vs Etrue to reduce sparsity
- No instrumental background model needed (?)

For Gamma rays we often simultaneously model a combination of :

- Energy (energy-only == 1D spectra)
- Morphology (position-only point-source == 2D image)
- Time (time-only == 1D light-curve)

"Data Cubes" are important:

- **Binned** data in Space/Time/Energy (and also instrumental dimensions like event-type!)
- We *do not use the term "Cube"* to talk about event lists (re: proposed Cube data model) as event-lists are not constrained in bins or boundaries.

Data Model

Observation ≈ 30 min exposure
Stable Observation Interval (SOI) ≈ 1-few per observation
One IRF per SOI, decomposed into multiple components.

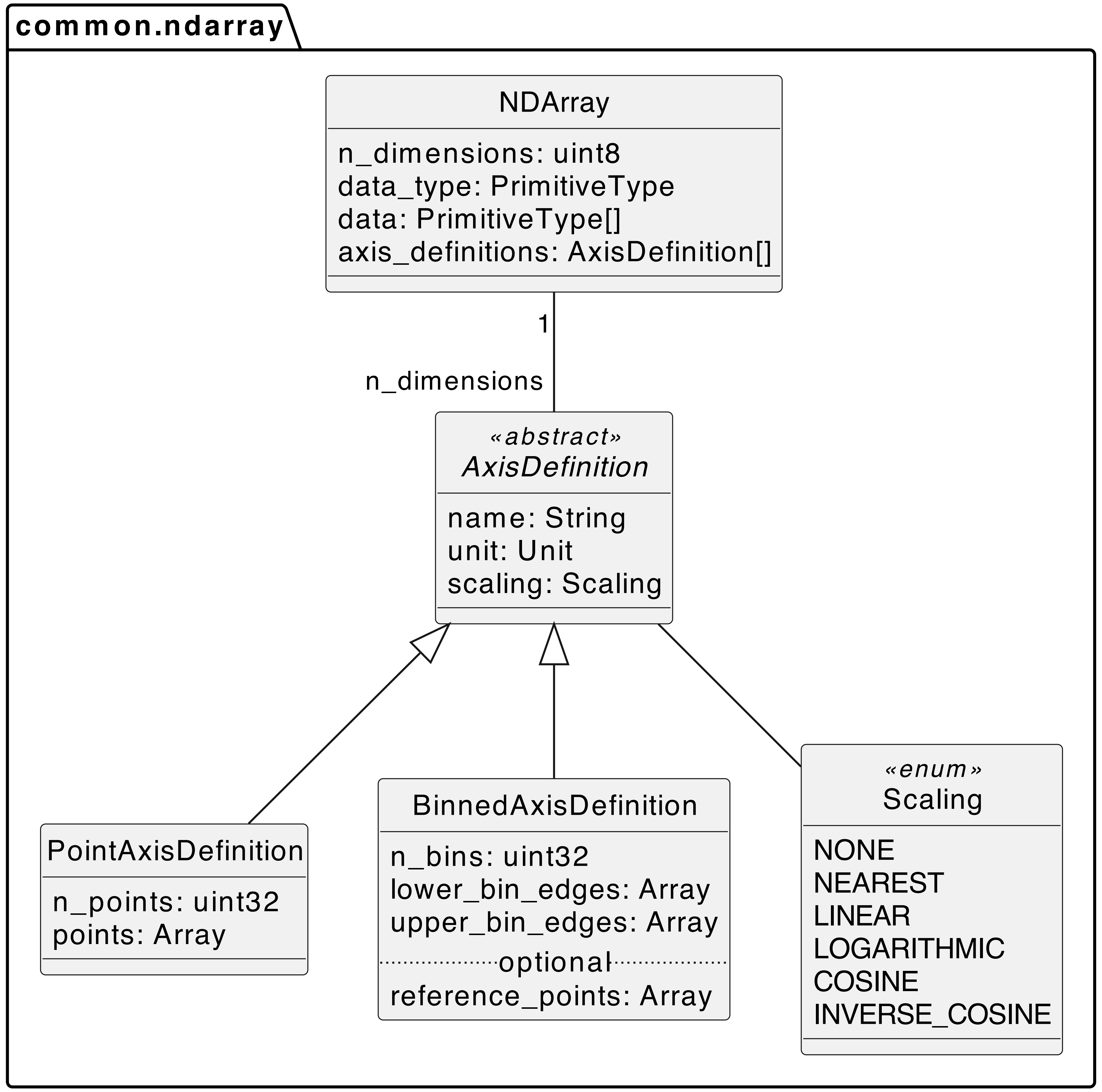


Figure 3.1 – UML Diagram of the NDArray data model.
CTAO Common Data Model Specification, v1a

Data Model

Observation ≈ 30 min exposure
Stable Observation Interval (SOI) ≈ 1-few per observation
One IRF per SOI, decomposed into multiple components.

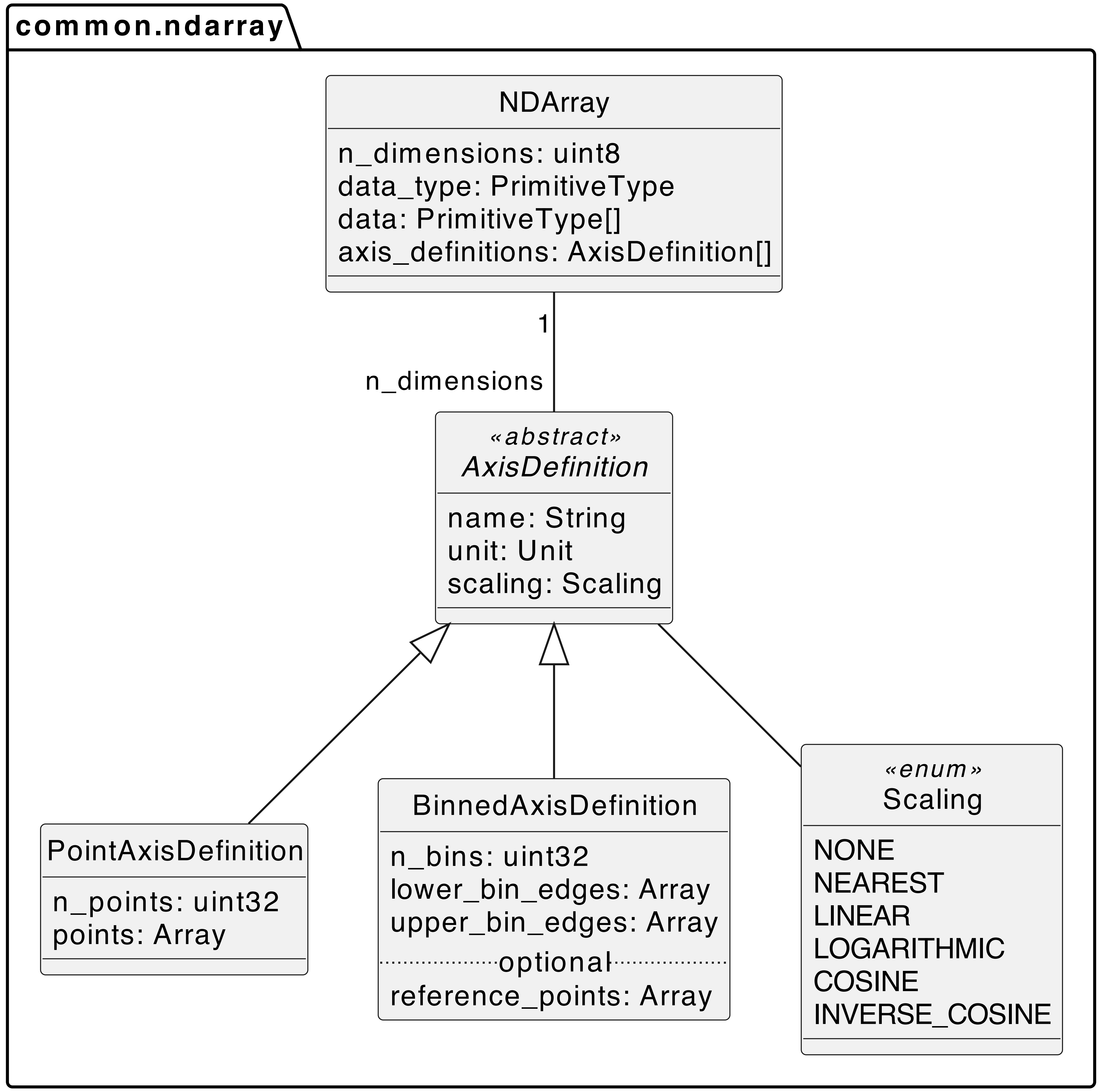


Figure 3.1 – UML Diagram of the `NDArray` data model.
CTAO Common Data Model Specification, v1a

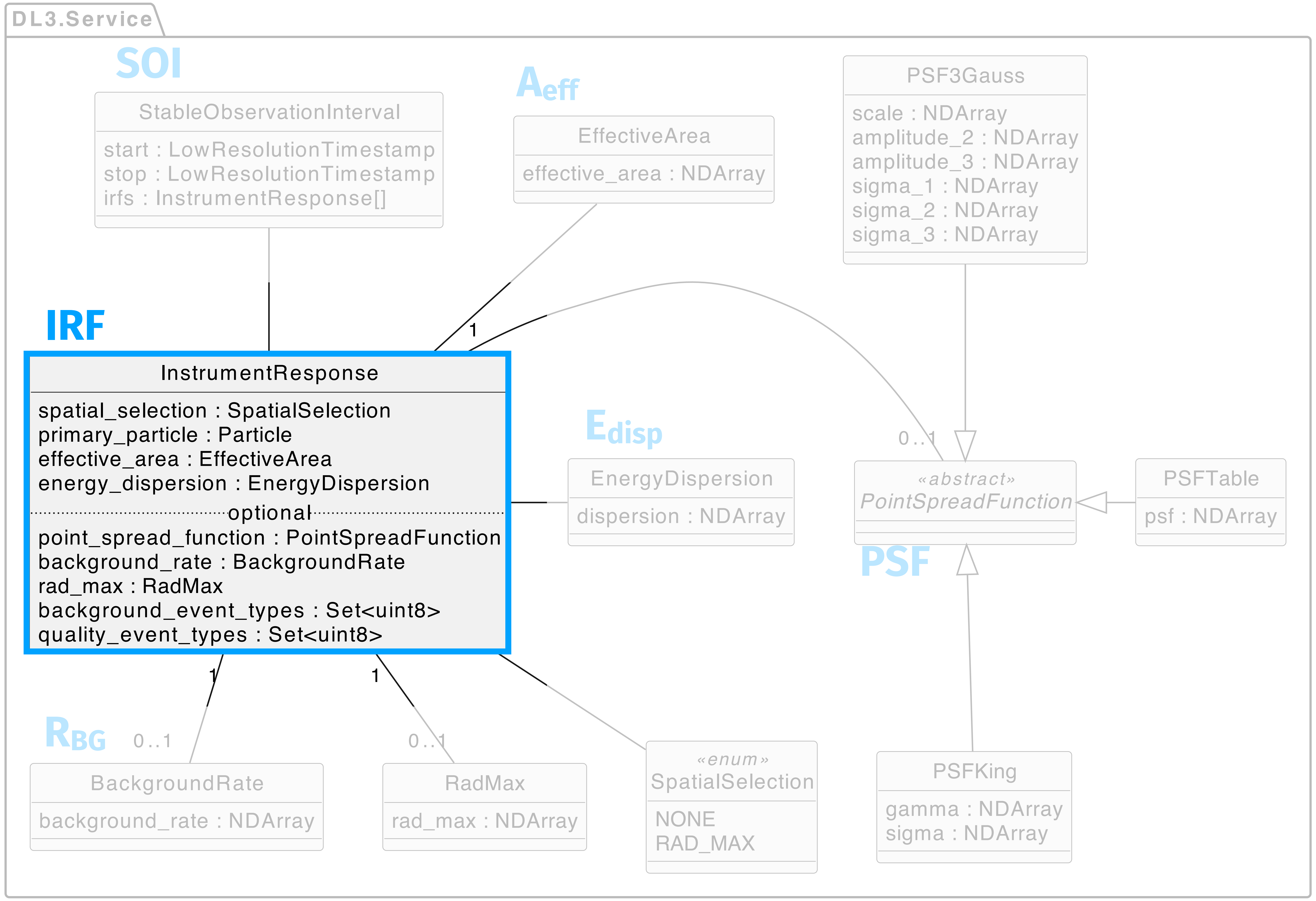


Figure 2.3 – UML Diagram of the DL3/Service/IRF Data Model.
CTAO DL3 Data Model Specification, v1b

Data Model

Observation ≈ 30 min exposure
Stable Observation Interval (SOI) ≈ 1-few per observation
One IRF per SOI, decomposed into multiple components.

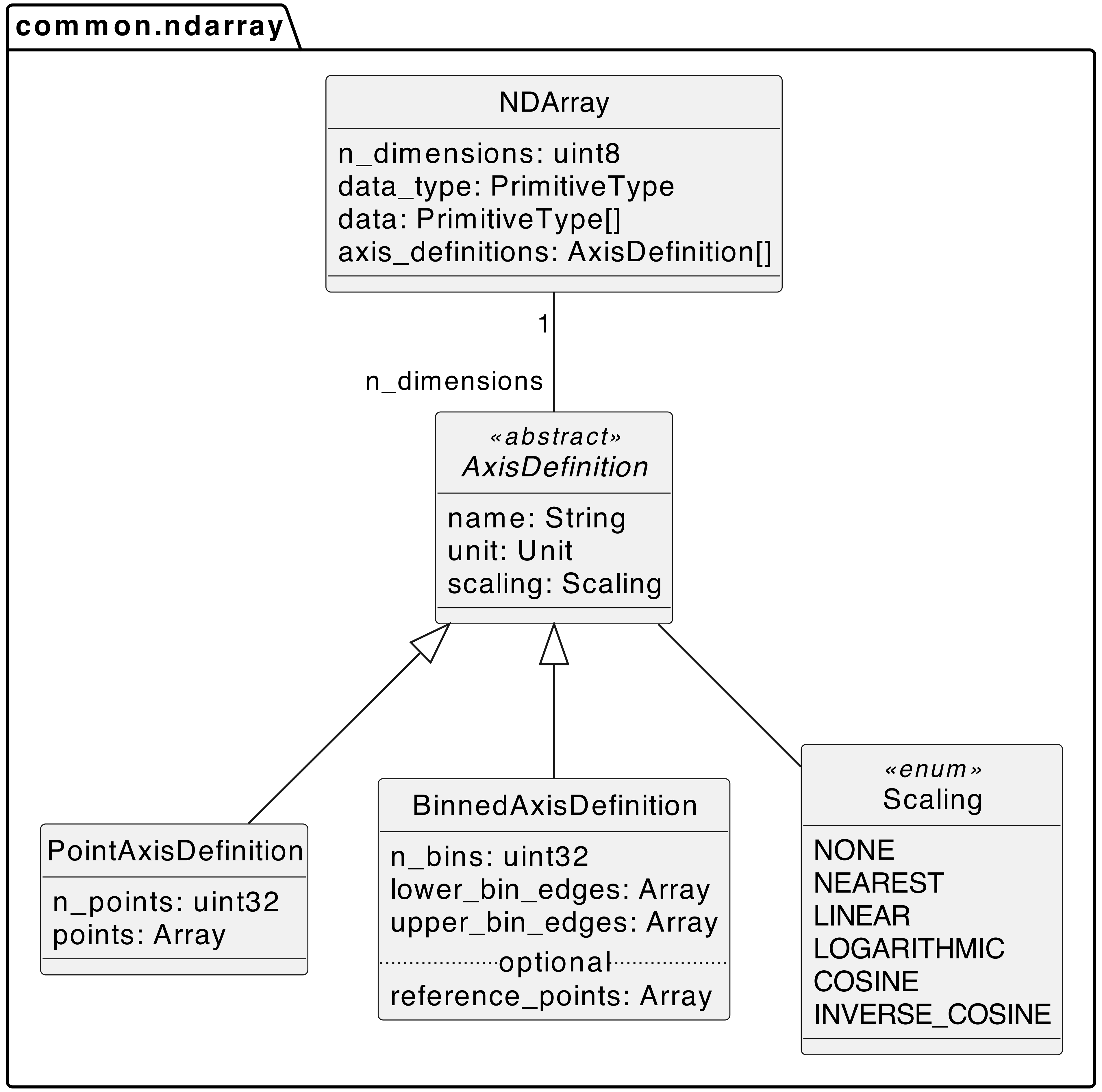


Figure 3.1 – UML Diagram of the `NDArray` data model.
CTAO Common Data Model Specification, v1a

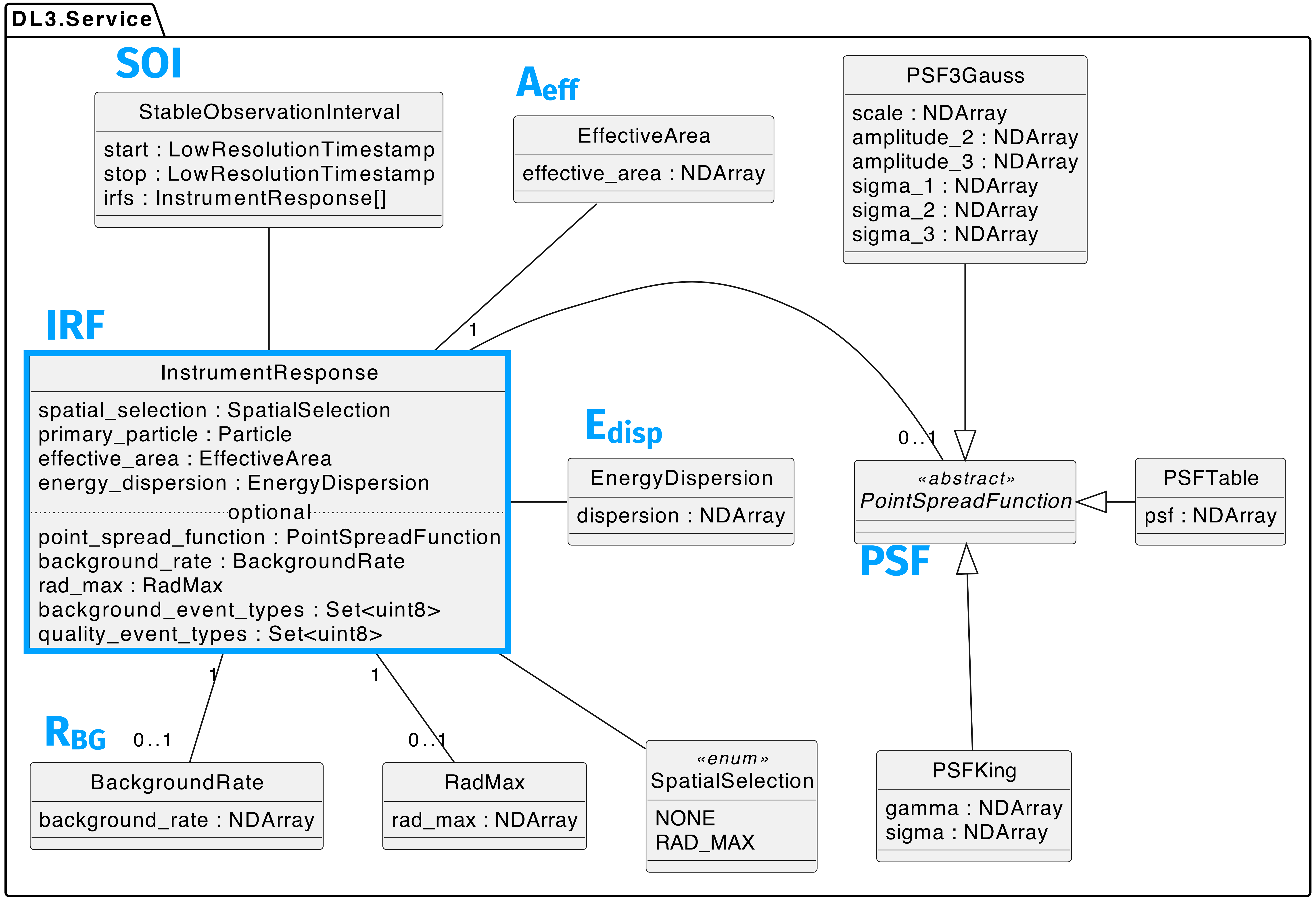


Figure 2.3 – UML Diagram of the DL3/Service/IRF Data Model.
CTAO DL3 Data Model Specification, v1b

A Caveat: Point-like vs General IRFs

For some science cases, cut away background by selecting events near to source, optimized by the PSF: "Point-Like IRFs"

- Assumes a **point-like source model** with known test position
- PSF effects (spillover of flux outside of region) taken into account in other IRFs
- Simpler analysis: no need to model PSF or source position/morphology
- BUT: is not general (extended emission, simultaneous fit of position and spectrum not supported)

We support this "type" of IRF as well:

- include a table of "spatial cut" vs energy
- other IRFs are generated with this cut already applied.
- User must select which type to use, based on their analysis case.