Improving CTAO event reconstruction at the highest energies

CTAO-Australia Summer Meeting 2024

Violet M. Harvey[1]
Sabrina Einecke Gavin Rowell

High Energy Astrophysics Group, The University of Adelaide, Australia

14 November 2024

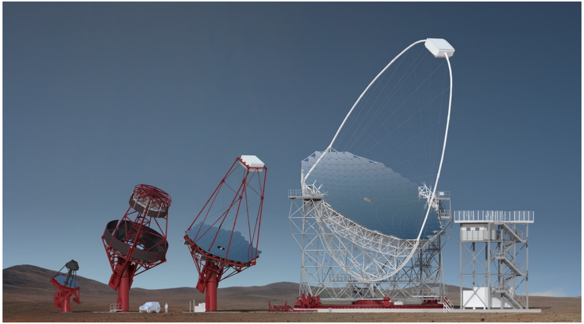
THE UNIVERSITY of ADELAIDE

CTAO Australia

[1] violet.harvey@adelaide.edu.au

Project aims

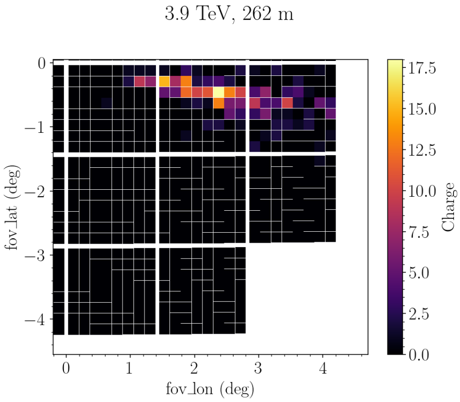
- Optimise the reconstruction/identification of the highest energy γ-rays.
- When within or near the array, existing stereoscopic techniques should handle well.
- However, being very bright, many events will trigger from **long distances** ($>$500 m).
- These might be seen by only one telescope ("**mono**"), or shower images may be "**truncated**" by edge of FOV of telescopes.
- The Small-Sized Telescopes (**SSTs**) are key to this work.
  - Effective mirror area: $\sim$5 m^2.
  - FOV: 8.8°.
  - Energy range: 5 TeV to 300 TeV.

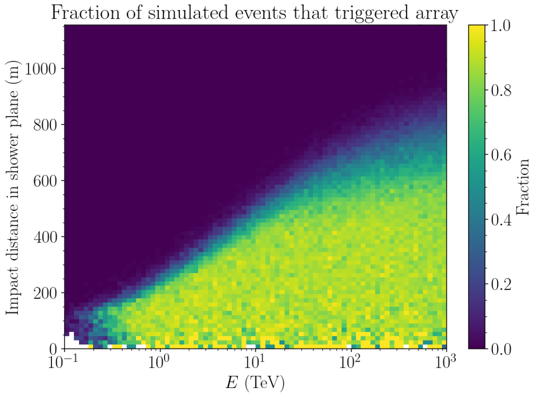


CTAO telescope scales.
Credit: Gabriel Pérez Diaz (IAC).

Air showers in Cherenkov telescopes

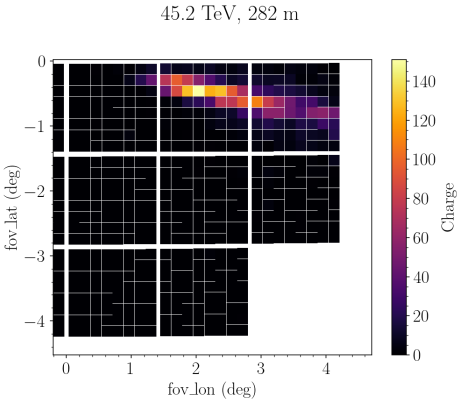
- γ-ray events are generally 10s of nanoseconds long.
- SST camera has 1 ns time resolution.
- Air shower images appear elliptical when summed up in time.
  - Shape depends on energy and impact distance.

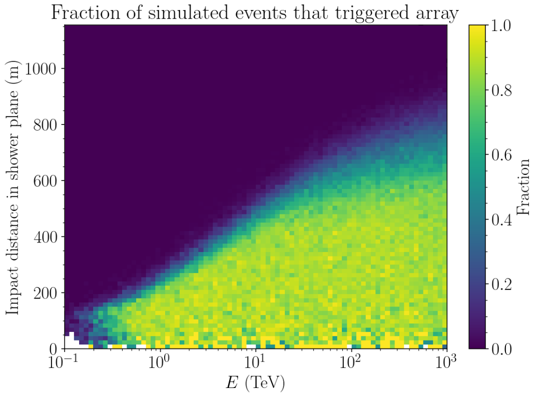




Air showers in Cherenkov telescopes

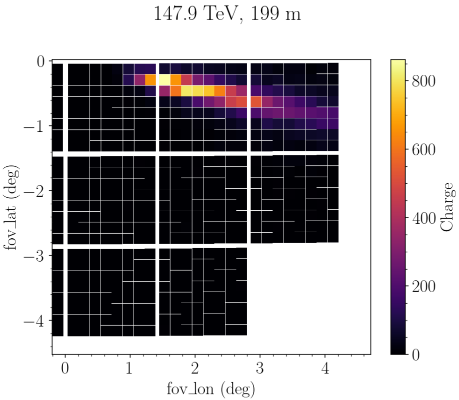
- γ-ray events are generally 10s of nanoseconds long.
- SST camera has 1 ns time resolution.
- Air shower images appear elliptical when summed up in time.
  - Shape depends on energy and impact distance.

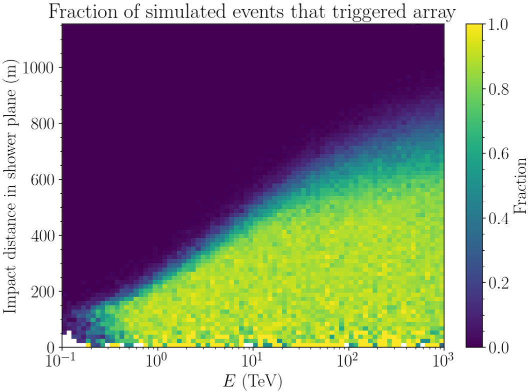




Air showers in Cherenkov telescopes

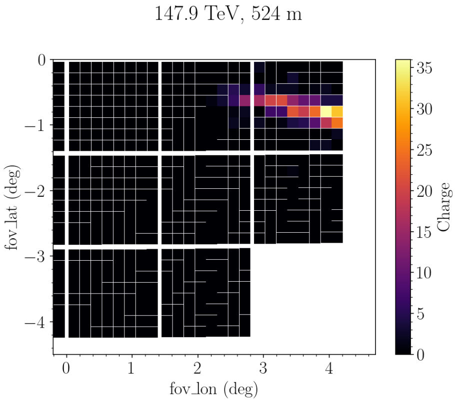
- γ-ray events are generally 10s of nanoseconds long.
- SST camera has 1 ns time resolution.
- Air shower images appear elliptical when summed up in time.
  - Shape depends on energy and impact distance.

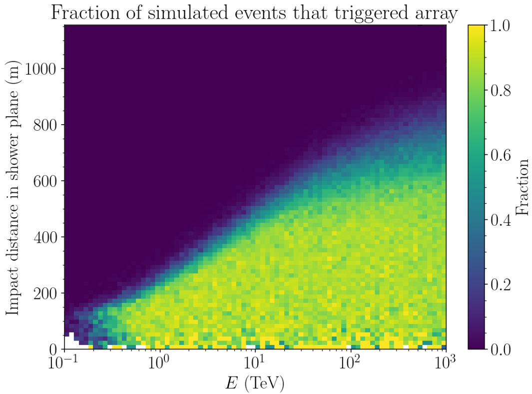




Air showers in Cherenkov telescopes

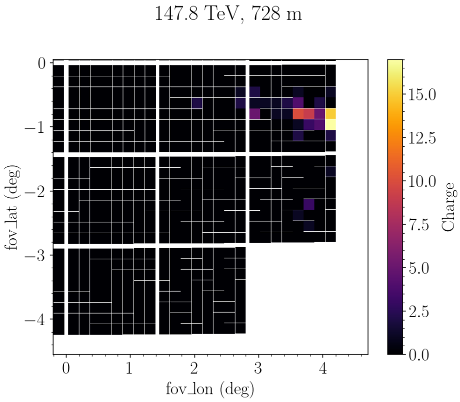
- γ-ray events are generally 10s of nanoseconds long.
- SST camera has 1 ns time resolution.
- Air shower images appear elliptical when summed up in time.
  - Shape depends on energy and impact distance.

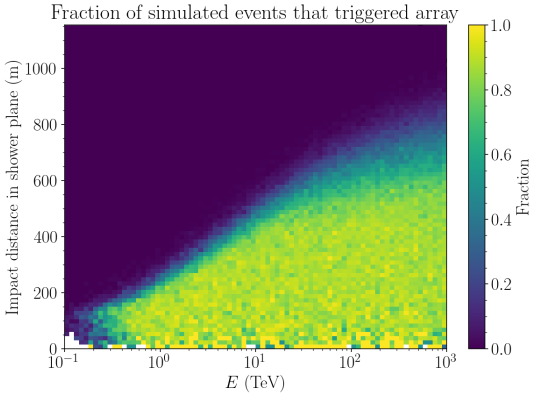




Air showers in Cherenkov telescopes

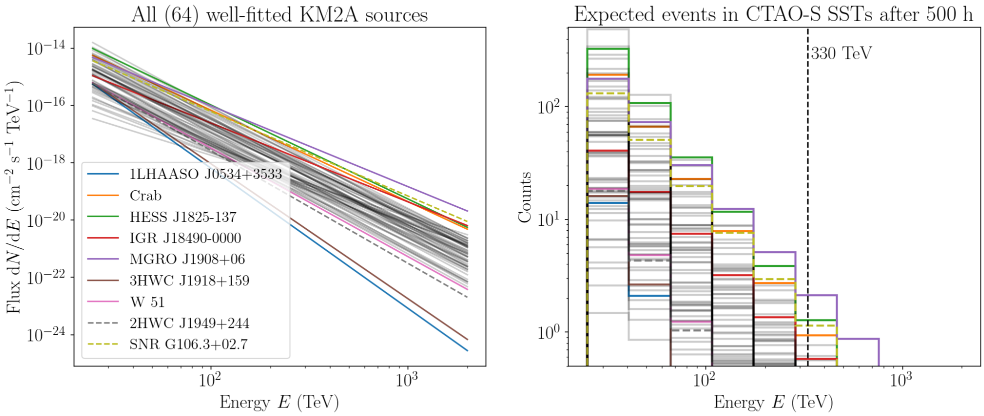
- γ-ray events are generally 10s of nanoseconds long.
- SST camera has 1 ns time resolution.
- Air shower images appear elliptical when summed up in time.
  - Shape depends on energy and impact distance.





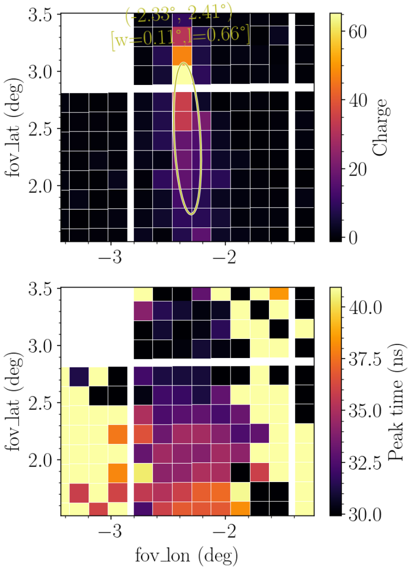
Frequency of high-energy events

- Real γ-ray source spectra tend to follow a **power-law** $\frac{dN}{dE} = N_0 \left(\frac{E}{E_0}\right)^{-\Gamma}$ with Γ around 3.5.
- Using spectral energy distributions from 1LHAASO, I roughly calculated **expected event counts** with CTAO-South SSTs in 500 hours. Not very many above 100 TeV.



Review of IACT analysis

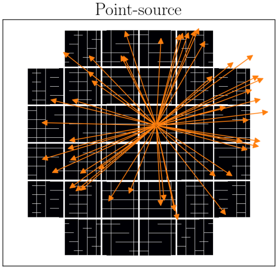
1. Calibration
2. Signal extraction:
   - Integrate total charge from waveform (trace).
3. Image cleaning:
   - Find signal pixels, remove noise pixels (NSB, electronic).
4. **Feature extraction:**
   - **Hillas parametrisation** (PCA):
     - Width, length, asymmetry.
     - Intensity, axis angle.
     - Centre of gravity (COG).
   - **Time gradient**.
   - Leakage.
   - Concentration.
   - Stereo reconstruction for source/impact position.

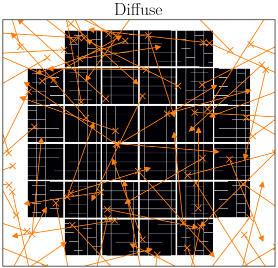


Review of IACT analysis

CTAO

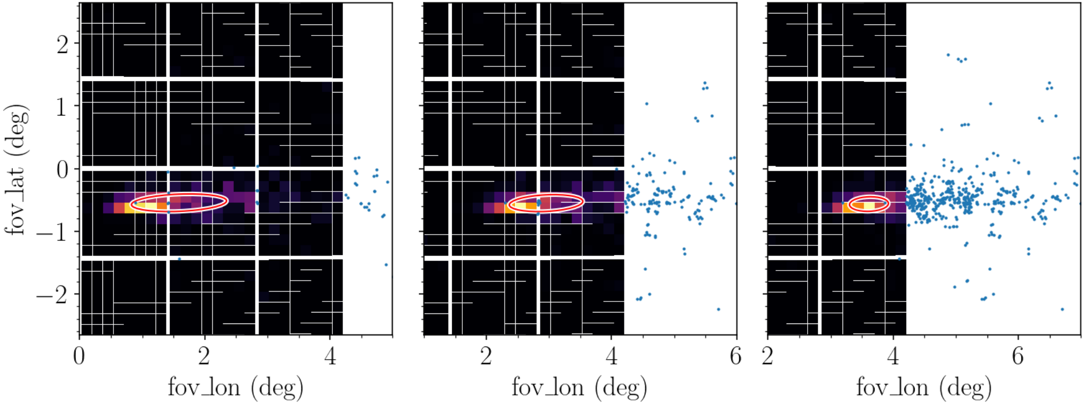
5. **Energy regression and γ–h classification:**
   - Reconstruction is done with machine learning: usually **random forest** (RF).
   - RF is trained on Monte Carlo simulations where the truth is known.
     - Provided the *features*, it learns to predict the *energy* or *gammaness*.
   - RF is trained on **diffuse** γ-rays but performance is tested on **point-source** γ-rays.
     - **Real data** is (usually) point-source.
6. Event lists / IRFs
7. Science!





Review of IACT analysis – Truncated images

- Hillas parametrisation becomes distorted as shower image nears edge of camera.
- Intensity also underestimated due to missing charge.



Blue points mark Cherenkov photons that did not land on a pixel.

Typical handling of truncated images

- To **quantify** truncation, a variable called **intensity leakage** is used: fraction of total collected charge that landed in edge pixels.
- No standardised approach... But similar practices between IACT experiments.
- *Typically*, leakage is used for two things:

Selection cut

- A selection cut is placed on the leakage.
- E.g., typical cut is leakage < 0.2.

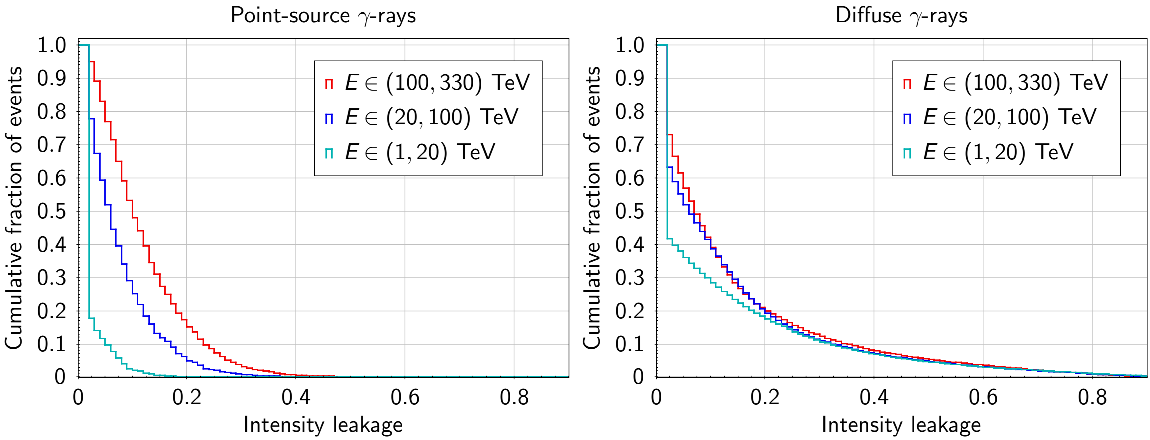
RF feature

- Leakage is also passed to the RF as an image feature.

Let's look closer at the consequences of both of these.

Typical handling of truncated images

What fraction of events are removed by a cut at 0.2?



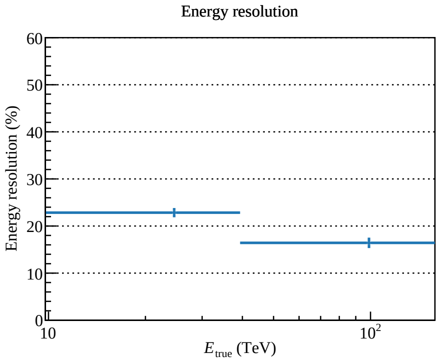
Depends if γ-rays are point-source or diffuse, but up to 20 % of data may be removed.

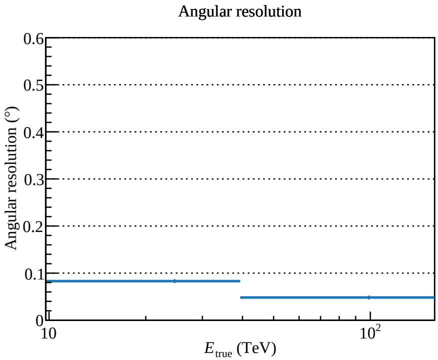
Typical handling of truncated images

How do truncated images perform in the RF analysis?

Disclaimer: Mono reconstruction, very poor statistics! Working on it!

leakage ≈ 0





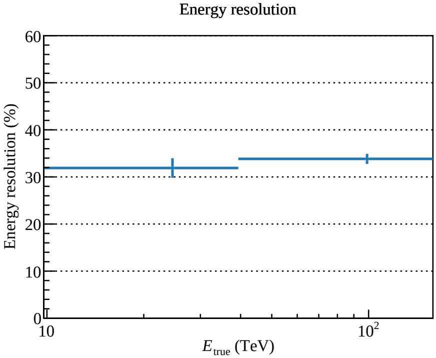
Still, it’s not too bad. The RF compensates fairly well.

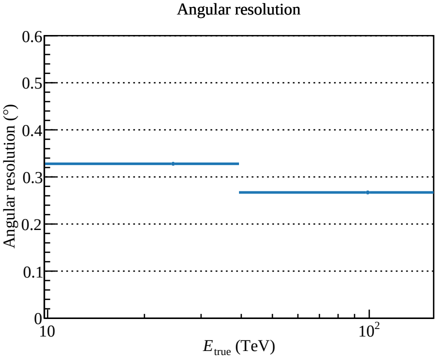
Typical handling of truncated images

How do truncated images perform in the RF analysis?

Disclaimer: Mono reconstruction, very poor statistics! Working on it!

$0.1 < \text{leakage} < 0.2$





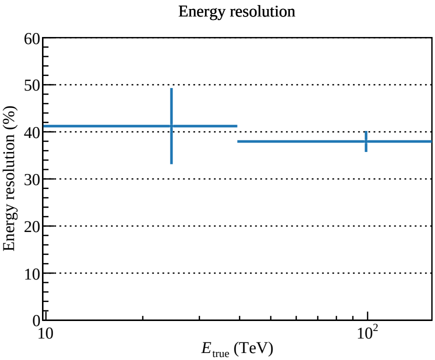
Still, it's not too bad. The RF compensates fairly well.

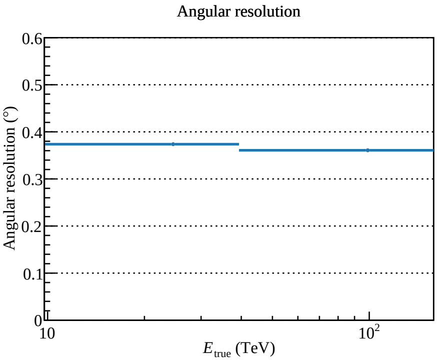
Typical handling of truncated images

How do truncated images perform in the RF analysis?

Disclaimer: Mono reconstruction, very poor statistics! Working on it!

0.2 < leakage < 0.4





Still, it's not too bad. The RF compensates fairly well.

Complications of the leakage definition

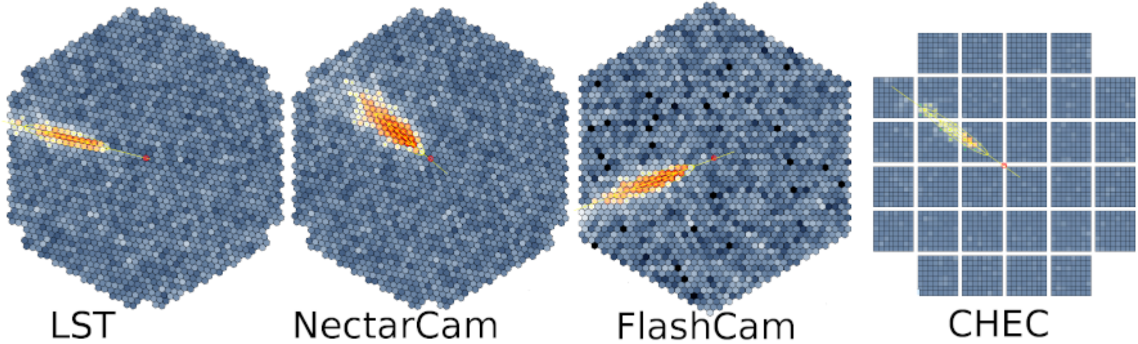
Reminder: intensity leakage $=$ *fraction of total collected charge that landed in edge pixels.*

What we've learnt: This definition of leakage has some problems.

1. SST camera has rather unique "edges".
2. Diffuse γ-rays cross edges at all different angles.

Unique issues for the SST

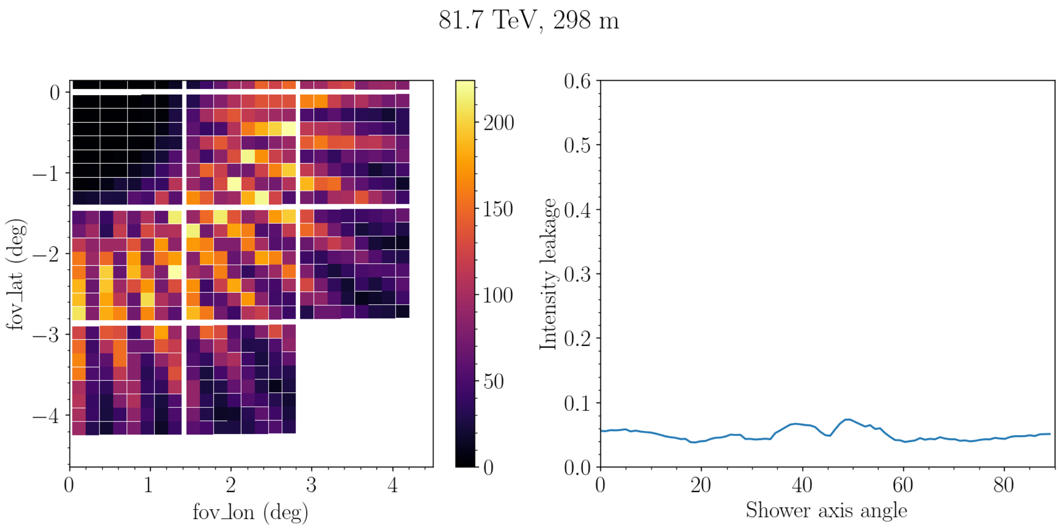
- Cherenkov telescopes typically use hexagonal pixels, arranged in a “circular” fashion.
- The SST camera (CHEC) is something of a first in being so square.



Credit: Konrad Bernlöhr (CTAO).

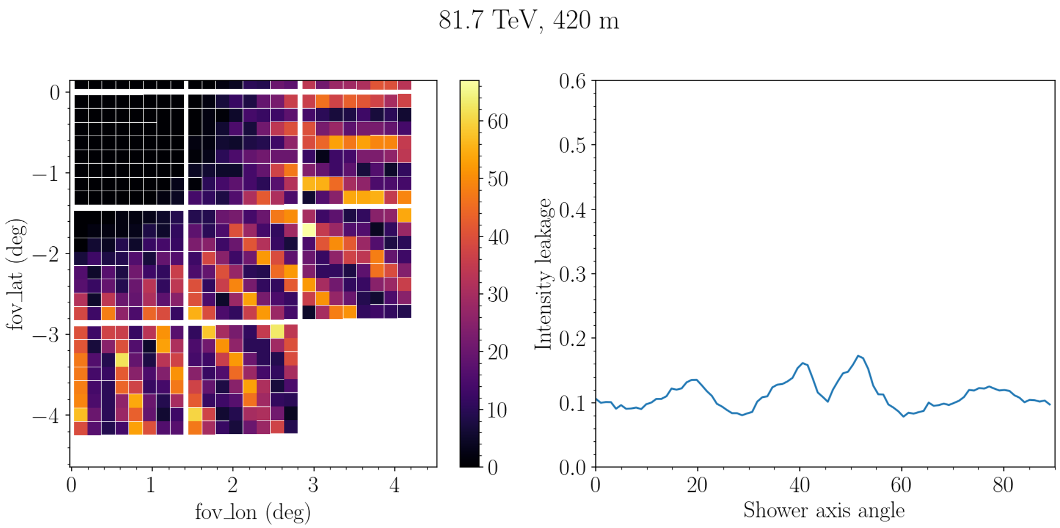
Unique issues for the SST

Same shower at different axis angles.



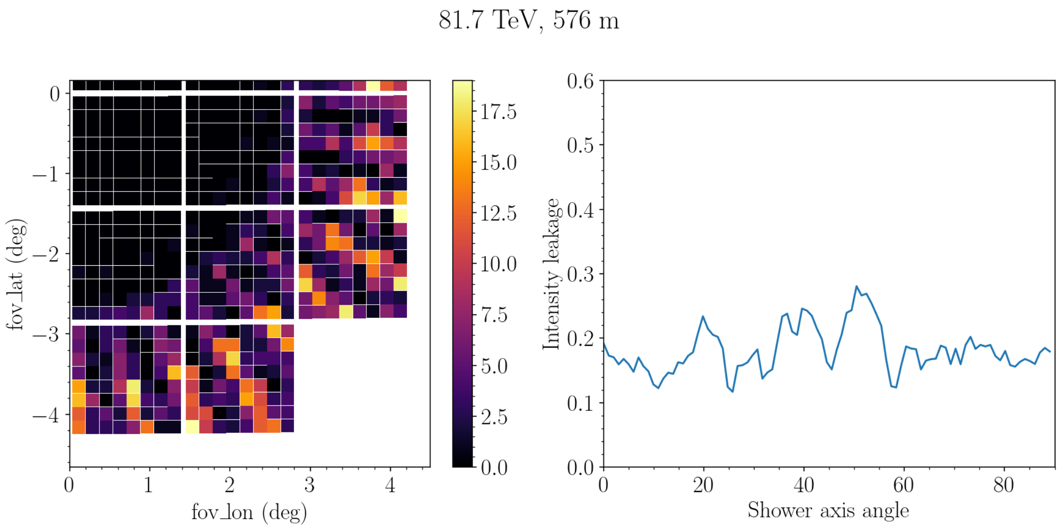
Unique issues for the SST

Same shower at different axis angles. Move it to different impact distances... (2/6)



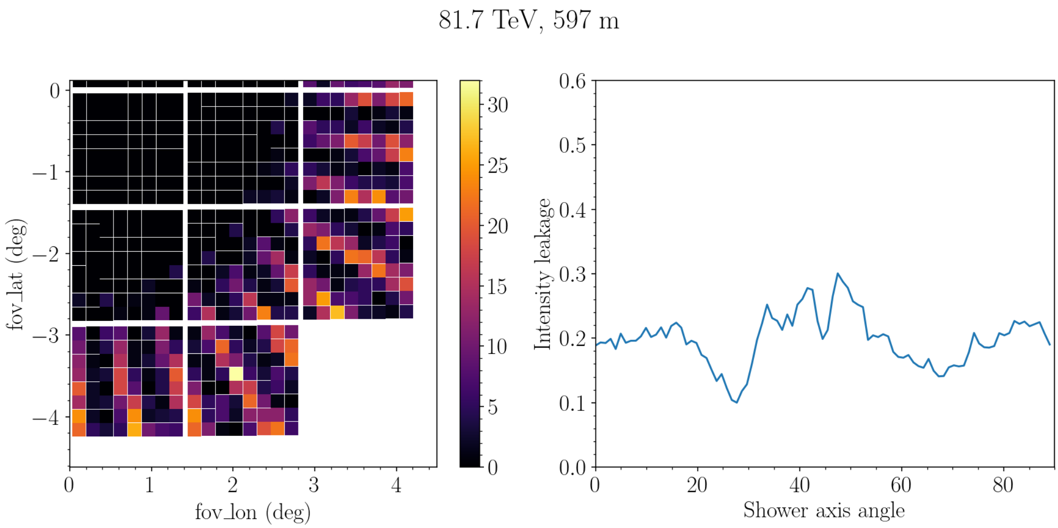
Unique issues for the SST

Same shower at different axis angles. Move it to different impact distances... (3/6)



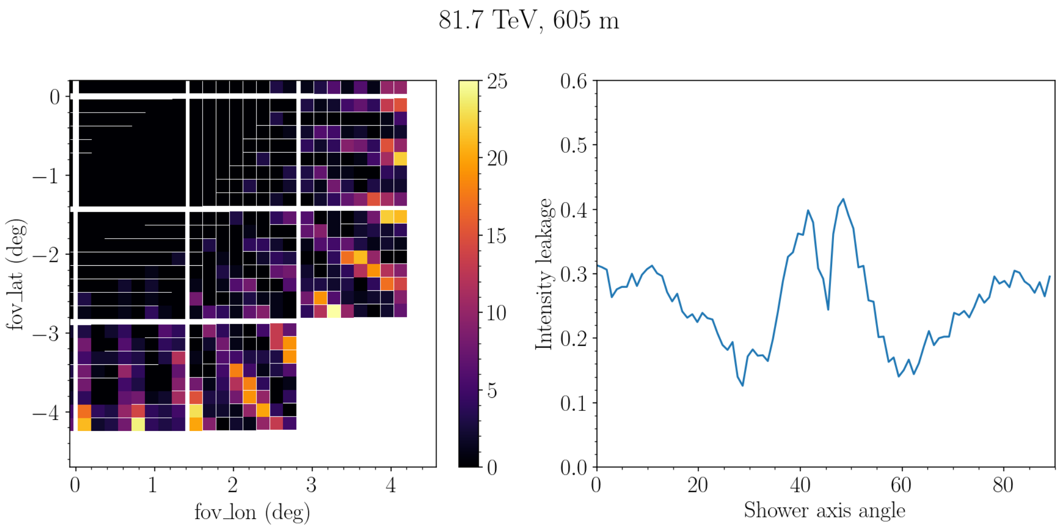
Unique issues for the SST

Same shower at different axis angles. Move it to different impact distances... (**4/6**)



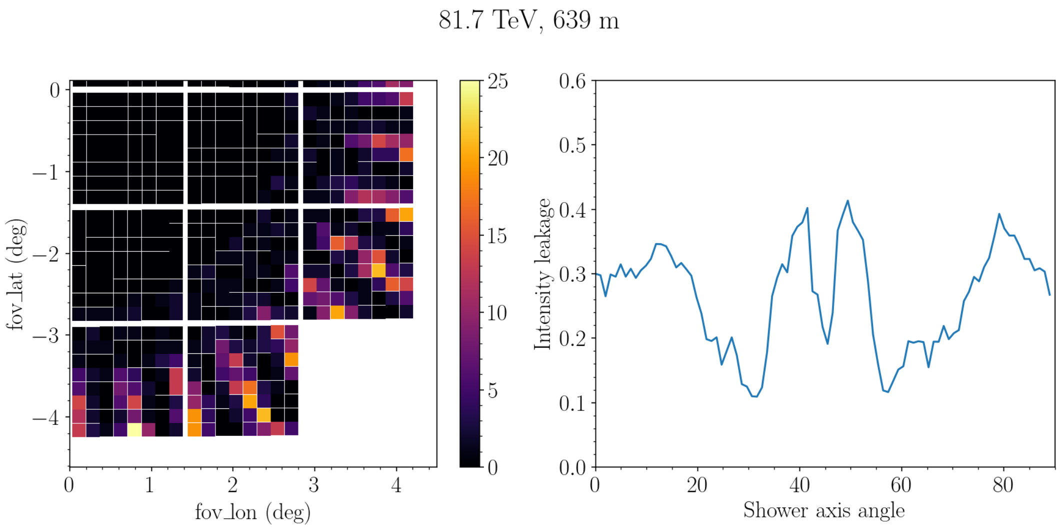
Unique issues for the SST

Same shower at different axis angles. Move it to different impact distances... (5/6)



Unique issues for the SST

Same shower at different axis angles. Move it to different impact distances... (6/6)

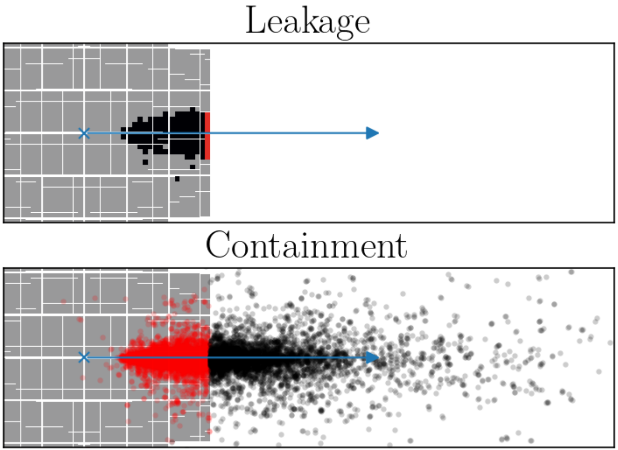


Truthing the leakage parameter

Leakage: Fraction of the total **collected** charge that landed in **edge** pixels.

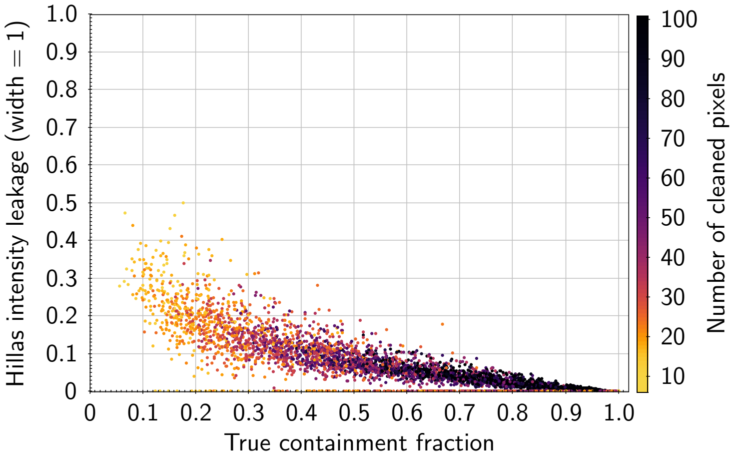
Containment: Fraction of **total** charge that was **collected**.
(Can only be calculated in MC.)

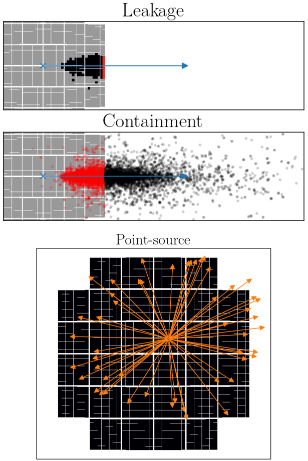
Containment a better proxy for how much data we have $\rightarrow$ how good the reconstruction might be.



Truthing the leakage parameter

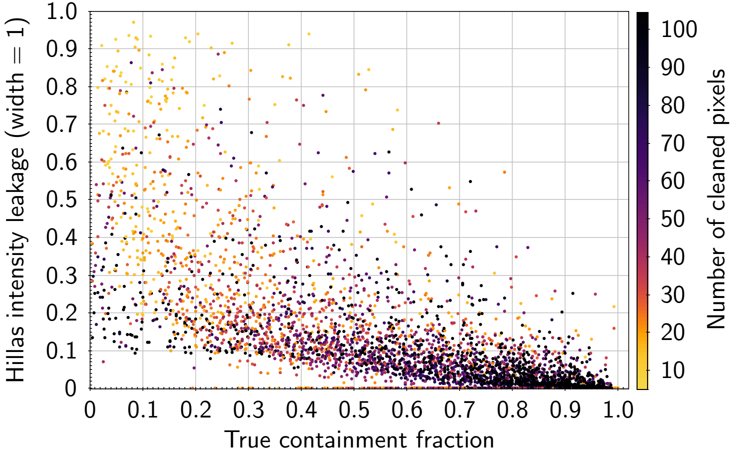
- For **point-source** γ-rays, leakage does somewhat correlate with containment.
- Scatter due to the corner effects...

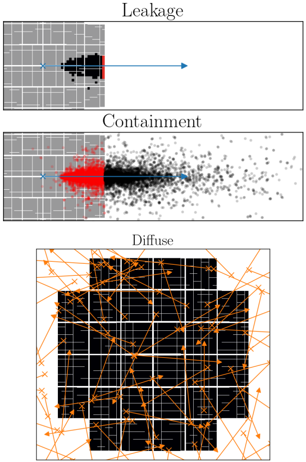




Truthing the leakage parameter

- For **diffuse** γ-rays, leakage is a *mess*!
- Quite little correlation with any concept of "reconstruction quality".





Summary so far...

Leakage:

- Highest energy events are often truncated.
- Typical way to quantify image truncation is with leakage.
- Leakage is generally used for quality cuts and as a random forest feature.

Problems:

- Leakage has unique behaviour for the SST. **This is not well understood.**
- Leakage almost doesn't make any sense at all for diffuse γ-rays, which are critical for training.

Performance:

- Despite this, the RF seems to do well at compensating truncated images using leakage.
- **Can we make it even better?**

Summary so far...

Leakage:

- Highest energy events are often truncated.
- Typical way to quantify image truncation is with leakage.
- Leakage is generally used for quality cuts and as a random forest feature.

Problems:

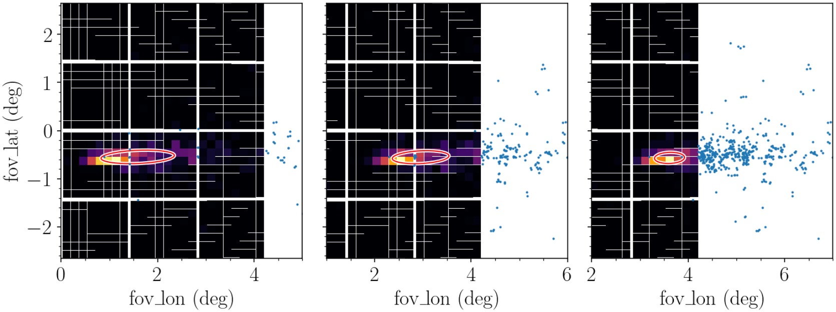
- Leakage has unique behaviour for the SST. **This is not well understood.**
- Leakage almost doesn't make any sense at all for diffuse γ-rays, which are critical for training.

Performance:

- Despite this, the RF seems to do well at compensating truncated images using leakage.
- **Can we make it even better?** Maybe, but not today ;)

An alternative approach

- Is there an alternative to Hillas parametrisation?
- What if our features didn't *need* compensating for the effect of truncation?



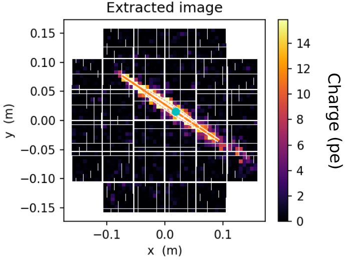
- An ongoing project, developing this code with Sabrina.

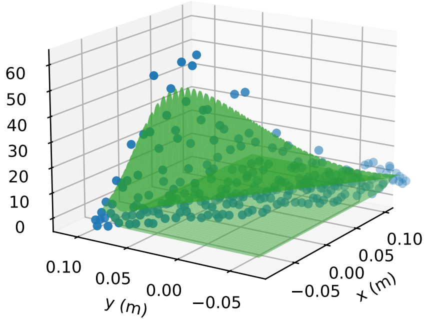
Likelihood analysis

CTAO

Shower model

Replace Hillas parametrisation with a fit of a **2D skew normal distribution**. Assume pixel times are linear with position along shower axis.





Model parameters:

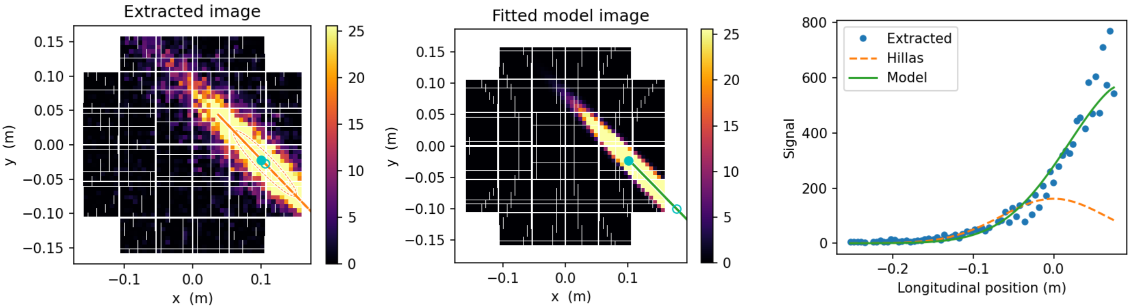
- Amplitude,
- Centre position (x, y),
- Length, width,
- Rotation angle,
- Skewness,
- Time gradient.

Provide these model parameters to the RF as usual.

Likelihood analysis

Hillas parameters (orange) can’t go outside of the camera.

2D model fit (green) can better describe these events.



Still, need smart priors and constraints to keep the fits physical.

Likelihood analysis

CTAO

Innovations

- Prior work by de Naurois and Rolland (2009), Alispach (2020).
- We use 1D fits to initialise the 2D fit, avoid strong dependence on first-guess Hillas parameters.
- Time gradient method avoids need for a (potentially wrong) pulse shape model.

Log-likelihood minimisation

$$s_i = \text{charge in pixel}$$
$$t_i = \text{peak time of pixel trace}$$
$$\vec{\Theta} = \text{model parameters}$$

$$\ln \mathcal{L} = -2 \sum_{i}^{n_{\text{pixels}}} \ln \mathcal{L}_{\text{pixel},i}$$

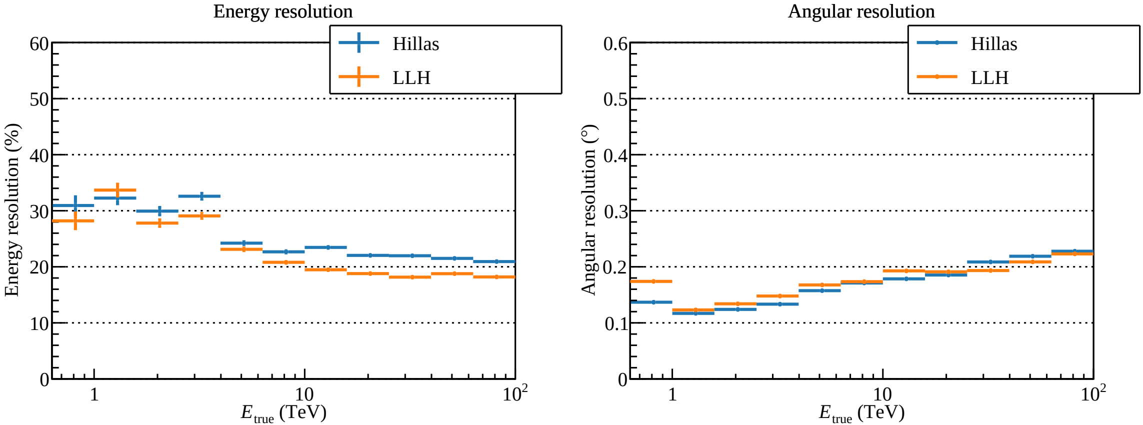
$$\mathcal{L}_{\text{pixel},i} = P_{\text{charge}}\left(s_i \,\middle|\, \vec{\Theta}\right) \cdot P_{\text{time}}\left(t_i \,\middle|\, \vec{\Theta}\right)$$

Search for the set of parameters $\vec{\Theta}$ which minimise $\ln \mathcal{L}$ for a given event.

P_{charge} and P_{time} are the **probabilities** that the observed charge and pulse peak time are explained by the model (mixture of **Poisson** and **Gaussian** prob. density functions).

Likelihood analysis

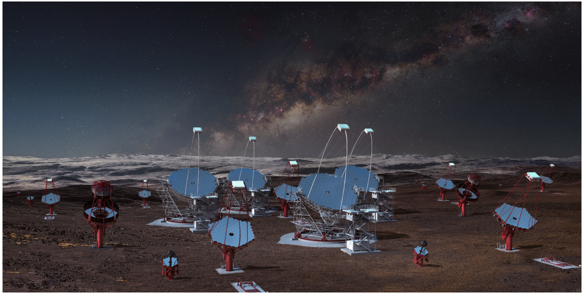
Disclaimer: Mono reconstruction, fully contained images (NO truncation!).



Can’t comment on the performance on truncated images just yet.
But, we seem to have a good starting point.

Conclusion

- Highest energy γ-rays are rare and often truncated.
- Existing measures of image truncation are complicated and sub-optimal.
  - The SST design especially makes for new challenges.
- Plenty of ways this could be improved, with close attention.
- Random forest already seems to be doing better than we might have thought.
- Lots of work going into the likelihood analysis that I haven't shown here.
- Not exactly certain where it will end up yet.
  - It may be able to prove itself on truncated images.
  - Otherwise, there are other useful avenues it unlocks, such as event-by-event posteriors.



Credit: Gabriel Pérez Diaz (IAC) / Marc-André Besel (CTAO) / ESO / N. Risinger (skysurvey.org).