technische universität
dortmund

e5 experimentelle physik 5
astroteilchenphysik

Nested Machine Learning Models for CTA

Master Thesis Half-Time Talk

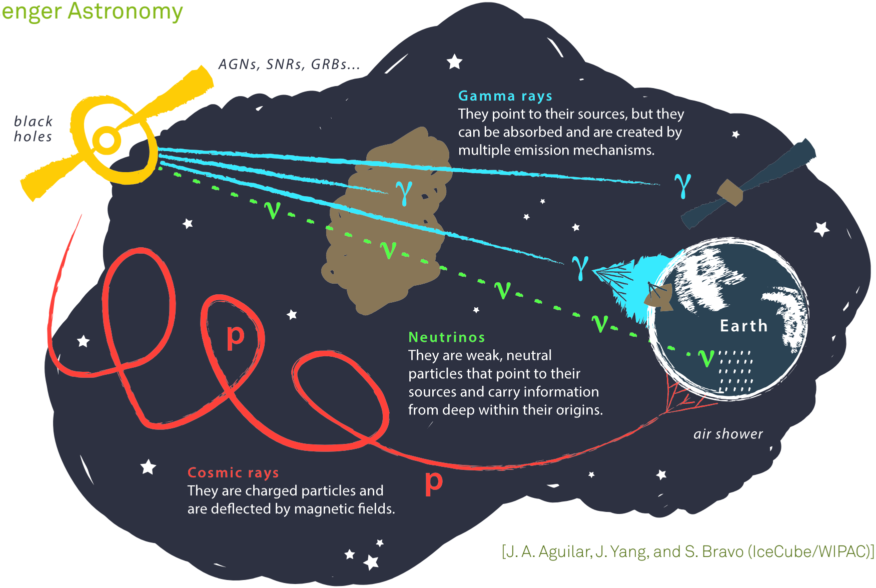
Lukas Beiske

December 16, 2022

E5b Astroparticle Physics
Department of Physics - TU Dortmund

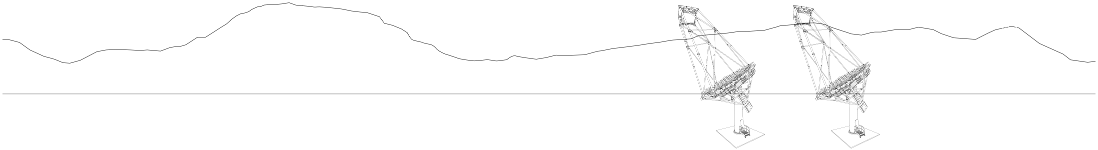
Overview

Multi-Messenger Astronomy



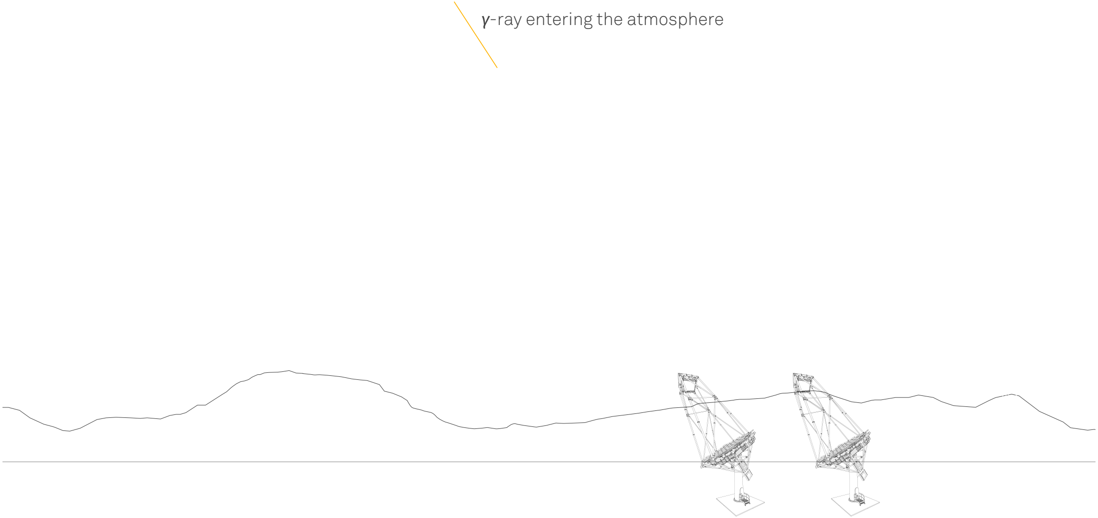
[J. A. Aguilar, J. Yang, and S. Bravo (IceCube/WIPAC)]

Imaging Atmospheric Cherenkov Telescopes (IACTs)



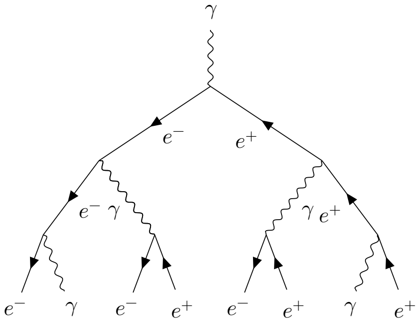
[A. Knierim]

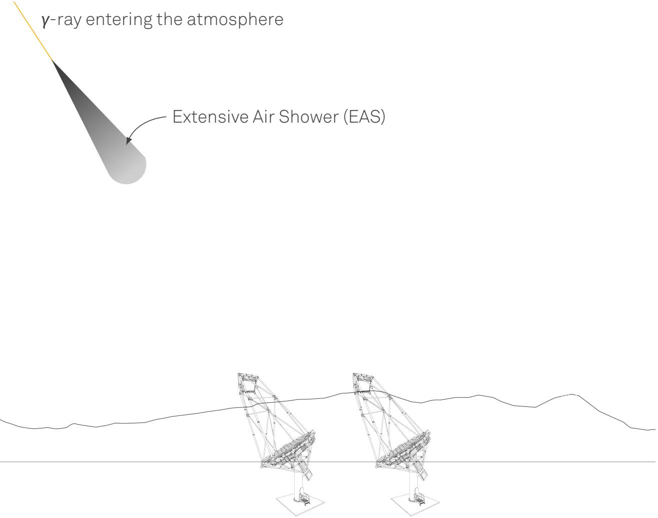
Imaging Atmospheric Cherenkov Telescopes (IACTs)



[A. Knierim]

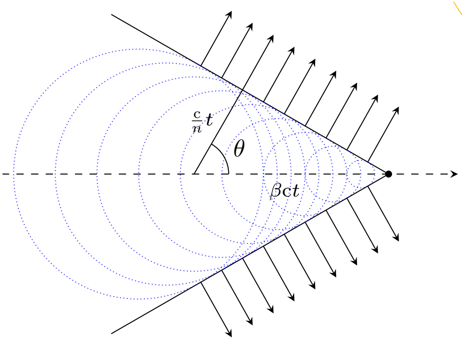
Imaging Atmospheric Cherenkov Telescopes (IACTs)

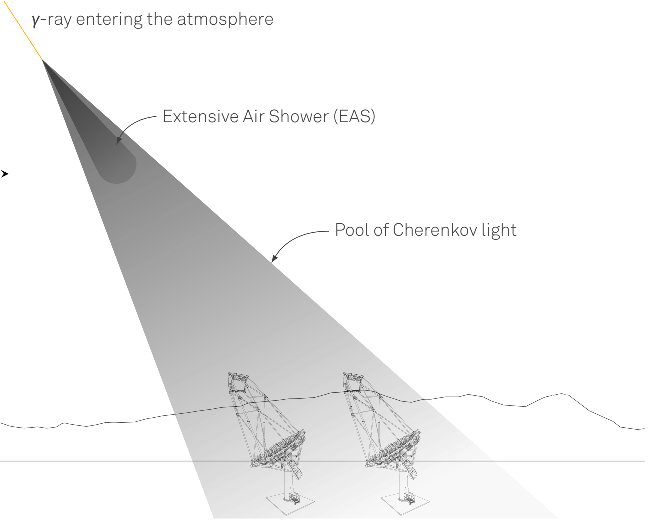




[A. Knierim]

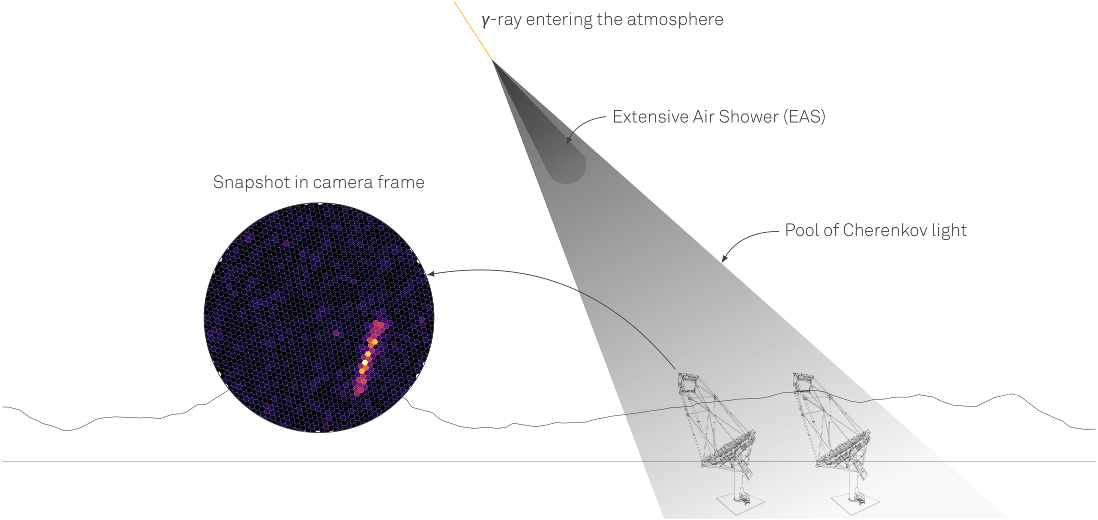
Imaging Atmospheric Cherenkov Telescopes (IACTs)





[A. Knierim]

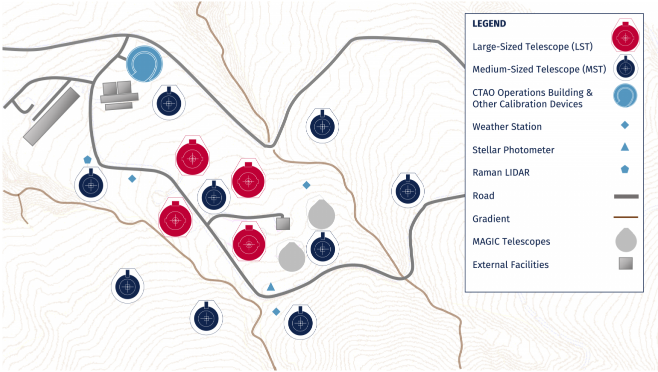
Imaging Atmospheric Cherenkov Telescopes (IACTs)



[A. Knierim]

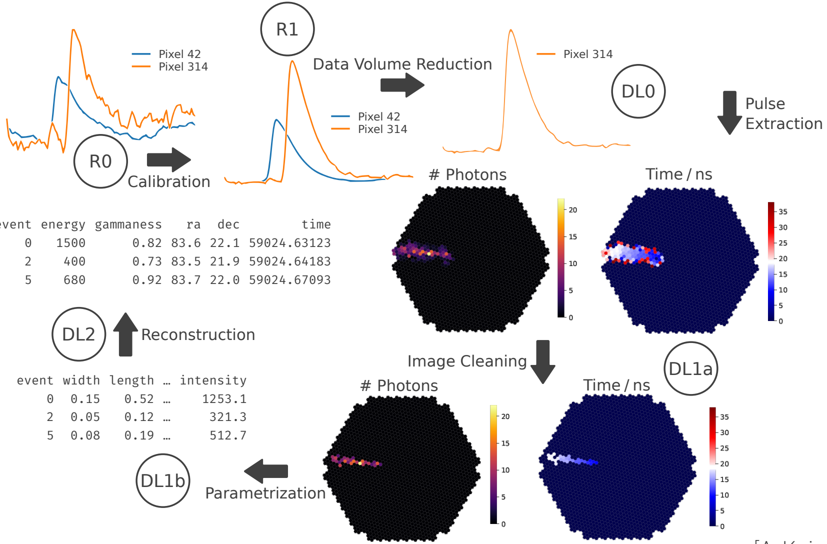
The Cherenkov Telescope Array

- CTA South (Paranal Observatory, Chile)
- CTA North (La Palma)
  - 4 LSTs + 9 MSTs
- Large-Sized Telescope (LST)
  - 23 m mirror diameter
  - 4.3° FoV
- Medium-Sized Telescope (MST)
  - 11.5 m mirror diameter
  - 7.7° FoV



[CTAO]

Dataprocessing using **ctapipe**



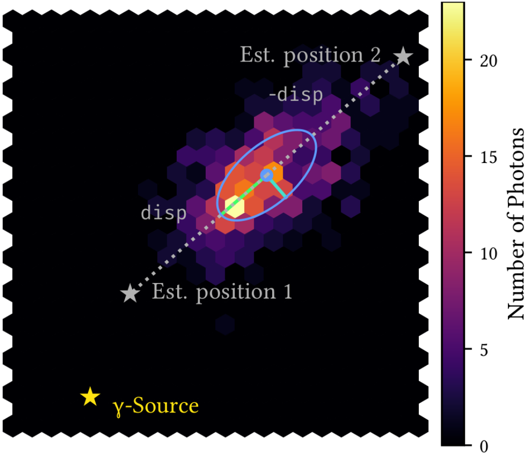
event	energy	gammaness	ra	dec	time
0	1500	0.82	83.6	22.1	59024.63123
2	400	0.73	83.5	21.9	59024.64183
5	680	0.92	83.7	22.0	59024.67093

event	width	length	…	intensity
0	0.15	0.52	…	1253.1
2	0.05	0.12	…	321.3
5	0.08	0.19	…	512.7

[A. Knierim, M. Linhoff]

The disp Method

- Monoscopic origin reconstruction for IACTs
- Assume main shower axis to be correctly reconstructed →source position on axis
- Train regressor to estimate distance from image center of gravity (`|disp|` / `norm`)
- Train classifier to decide between remaining two possibilities (`sign`)



[Max Noethe, PhD thesis]

Scaled Parameters and Feature Selection

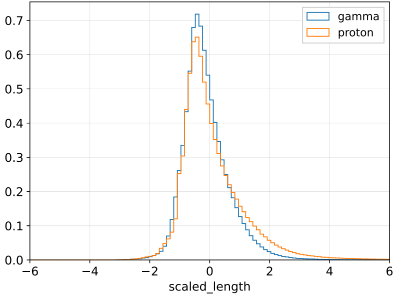
Scaled parameters

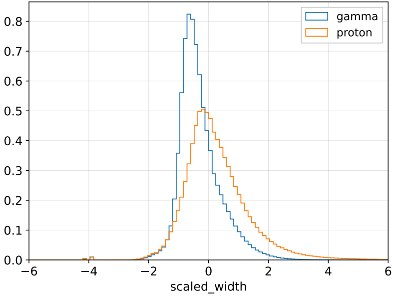
Compare image length (l) and width (w) with expectation value and variance from simulations as function of image charge (q), impact distance (d), and telescope type (t):

$$SL = \frac{l(q,d,t) - \langle l(q,d,t)\rangle}{\sigma_l(q,d,t)} \qquad SW = \frac{w(q,d,t) - \langle w(q,d,t)\rangle}{\sigma_w(q,d,t)}$$

Easy combination for stereo observations:

$$MSL = \frac{\sum_{\text{tels}} SL}{\sqrt{n_{\text{tels}}}} \qquad MSW = \frac{\sum_{\text{tels}} SW}{\sqrt{n_{\text{tels}}}}$$





MRMR feature selection – Energy

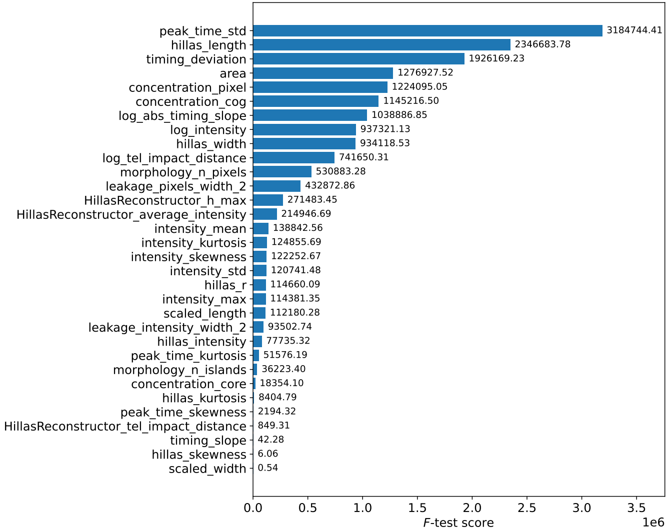
At each iteration ***i*** compute ***score****$_i$* for all ***f*** not yet selected

$$score_i(f) = \frac{relevance(f \mid \text{target})}{redundancy(f \mid f_{\text{already selected}})}$$

and add the feature with the highest score.

Energy features (15)

```
- peak_time_std
- intensity_std
- timing_deviation
- hillas_length
- concentration_pixel
- leakage_intensity_width_2
- log_abs_timing_slope
- area
- concentration_cog
- log_tel_impact_distance
- hillas_width
- log_intensity
- morphology_n_pixels
- HillasReconstructor_h_max
- leakage_pixels_width_2
```



MRMR feature selection – Particle Id

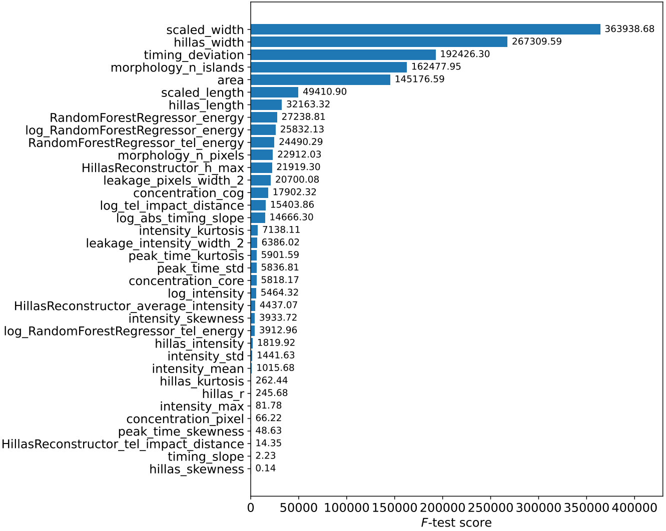
At each iteration ***i*** compute ***score$_i$*** for all ***f*** not yet selected

$$score_i(f) = \frac{relevance(f \mid \text{target})}{redundancy(f \mid f_{\text{already selected}})}$$

and add the feature with the highest score.

Particle Id features (15)

```
- scaled_width
- HillasReconstructor_h_max
- log_tel_impact_distance
- hillas_width
- morphology_n_islands
- timing_deviation
- area
- scaled_length
- RandomForestRegressor_energy
- log_RandomForestRegressor_energy
- leakage_pixels_width_2
- hillas_length
- peak_time_kurtosis
- RandomForestRegressor_tel_energy
- morphology_n_pixels
```



MRMR feature selection – disp

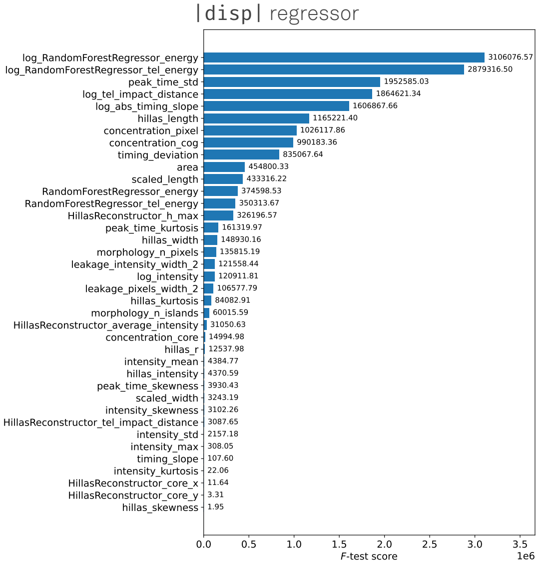
At each iteration i compute $score_i$ for all f not yet selected

$$score_i(f) = \frac{relevance(f \mid \text{target})}{redundancy(f \mid f_{\text{already selected}})}$$

and add the feature with the highest score.

disp features (19)

```
- log_RandomForestRegressor_energy
- log_tel_impact_distance
- log_RandomForestRegressor_tel_energy
- log_abs_timing_slope
- peak_time_std
- concentration_pixel
- hillas_length
- concentration_cog
- timing_deviation
- scaled_length
- HillasReconstructor_h_max
- RandomForestRegressor_energy
- area
- peak_time_kurtosis
- RandomForestRegressor_tel_energy
- timing_slope
- hillas_skewness
- HillasReconstructor_core_x
- HillasReconstructor_core_y
```



MRMR feature selection – disp

At each iteration ***i*** compute ***score**$_i$* for all ***f*** not yet selected

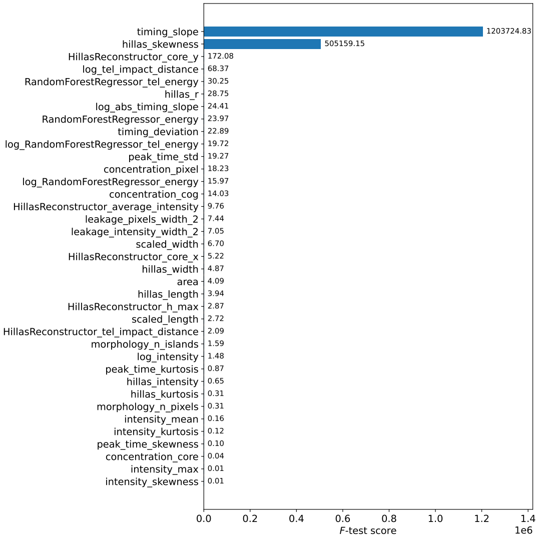
$$score_i(f) = \frac{relevance(f \mid \text{target})}{redundancy(f \mid f_{\text{already selected}})}$$

and add the feature with the highest score.

disp features (19)

- log_RandomForestRegressor_energy
- log_tel_impact_distance
- log_RandomForestRegressor_tel_energy
- log_abs_timing_slope
- peak_time_std
- concentration_pixel
- hillas_length
- concentration_cog
- timing_deviation
- scaled_length
- HillasReconstructor_h_max
- RandomForestRegressor_energy
- area
- peak_time_kurtosis
- RandomForestRegressor_tel_energy
- timing_slope
- hillas_skewness
- HillasReconstructor_core_x
- HillasReconstructor_core_y

sign classifier



Concept

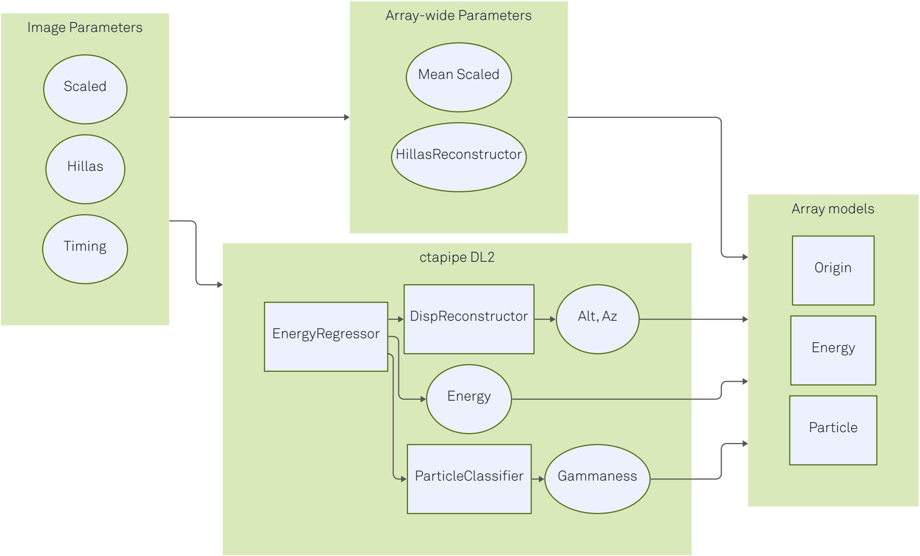


Image Parameters
Scaled
Hillas
Timing
Array-wide Parameters
Mean Scaled
HillasReconstructor
ctapipe DL2
EnergyRegressor
DispReconstructor
Alt, Az
Energy
ParticleClassifier
Gammaness
Array models
Origin
Energy
Particle

Energy Regression

Energy Regression (Telescope)

Randomized hyperparameter optimization yields:

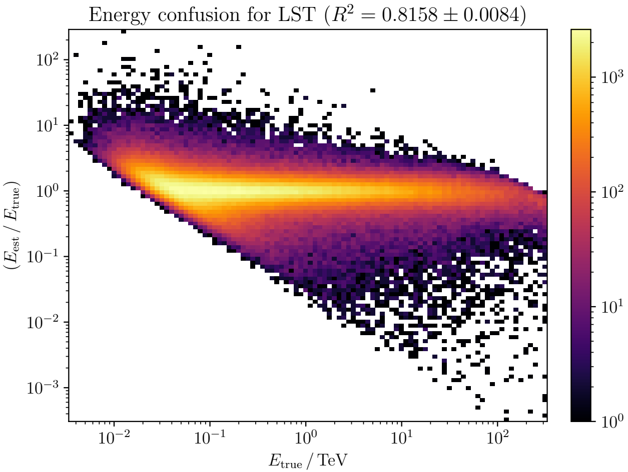
Configuration energy regressor

```
EnergyRegressor:
  model_cls: RandomForestRegressor
  model_config:
    n_estimators: 69
    max_features: 0.5227
    max_samples: 0.7138
    min_samples_leaf: 0.000013
    n_jobs: 40

  log_target: True
```

5-fold cross-validation:

- 551 635 LST events
- 1 199 782 MST events



Energy Regression (Telescope)

Randomized hyperparameter optimization yields:

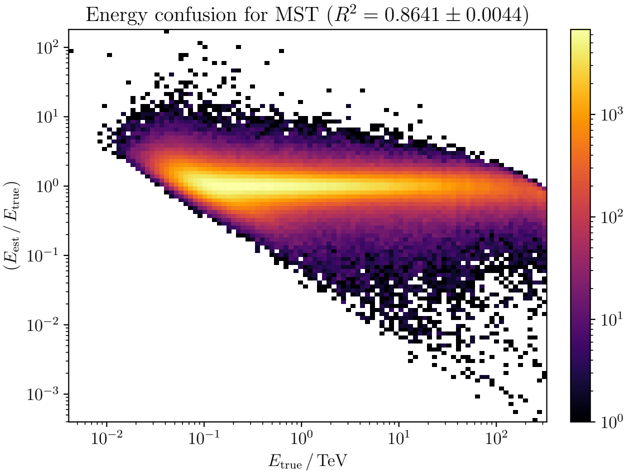
Configuration energy regressor

```
EnergyRegressor:
  model_cls: RandomForestRegressor
  model_config:
    n_estimators: 69
    max_features: 0.5227
    max_samples: 0.7138
    min_samples_leaf: 0.000013
    n_jobs: 40

  log_target: True
```

5-fold cross-validation:

- 551 635 LST events
- 1 199 782 MST events



Energy Regression (Telescope)

Randomized hyperparameter optimization yields:

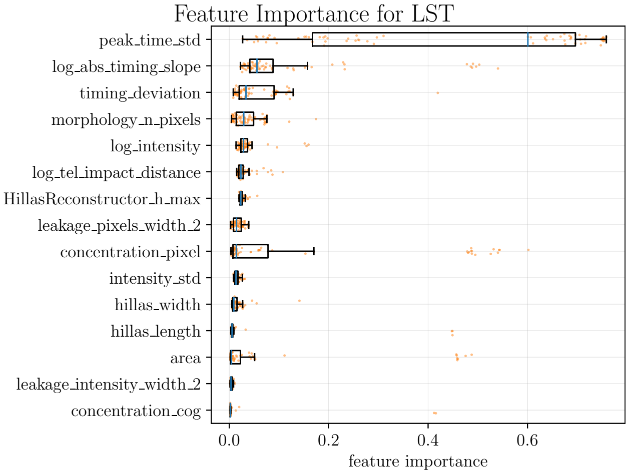
Configuration energy regressor

```
EnergyRegressor:
  model_cls: RandomForestRegressor
  model_config:
    n_estimators: 69
    max_features: 0.5227
    max_samples: 0.7138
    min_samples_leaf: 0.000013
    n_jobs: 40

  log_target: True
```

5-fold cross-validation:

- 551 635 LST events
- 1 199 782 MST events



Energy Regression (Telescope)

Randomized hyperparameter optimization yields:

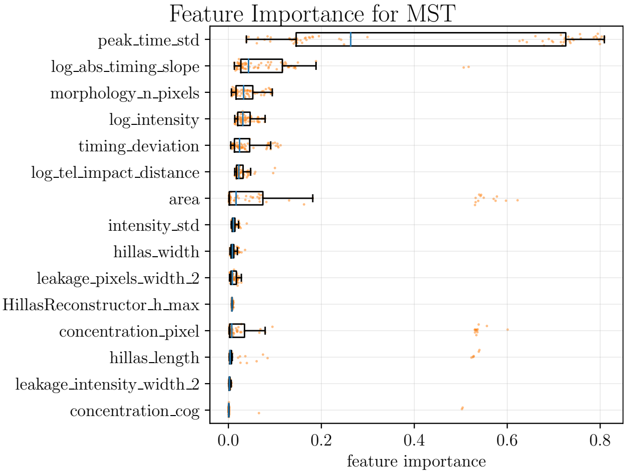
Configuration energy regressor

```
EnergyRegressor:
  model_cls: RandomForestRegressor
  model_config:
    n_estimators: 69
    max_features: 0.5227
    max_samples: 0.7138
    min_samples_leaf: 0.000013
    n_jobs: 40

  log_target: True
```

5-fold cross-validation:

- 551 635 LST events
- 1 199 782 MST events



Energy Regression (Array)

- Use (averaged) telescope predictions and array-wide features
- No hyperparameter optimization (yet)
- 5-fold cross-validation on **461 969** events

Configuration array energy regressor

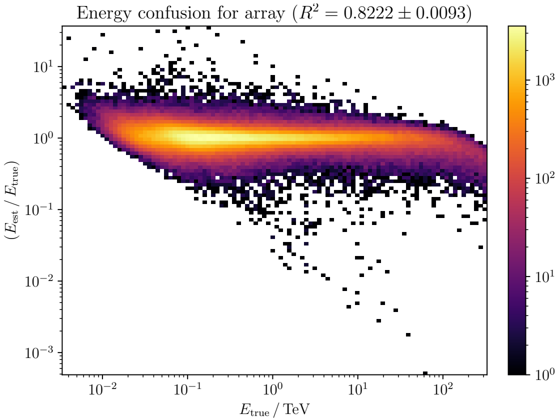
```yaml
model_cls: RandomForestRegressor
model_config:
    n_estimators: 200
    max_features: "sqrt"
    min_samples_leaf: 0.00001
    n_jobs: 40

log_target: True

features:
    - n_telescopes_HillasReconstructor
    - n_LST_HillasReconstructor
    - n_MST_HillasReconstructor
    - mean_scaled_length
    - mean_scaled_width
    - HillasReconstructor_core_x
    - HillasReconstructor_core_y
    - HillasReconstructor_average_intensity
    - HillasReconstructor_h_max
    - HillasReconstructor_alt
    - HillasReconstructor_az
    - RandomForestClassifier_prediction
    - RandomForestRegressor_energy
```

Energy Regression (Array)

- Use (averaged) telescope predictions and array-wide features
- No hyperparameter optimization (yet)
- 5-fold cross-validation on **461 969** events



Configuration array energy regressor

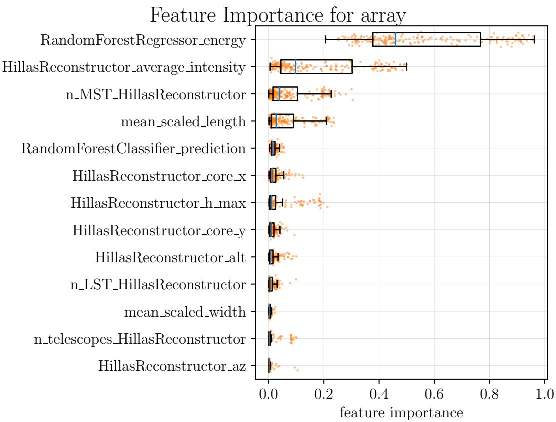
```yaml
model_cls: RandomForestRegressor
model_config:
    n_estimators: 200
    max_features: "sqrt"
    min_samples_leaf: 0.00001
    n_jobs: 40

log_target: True

features:
    - n_telescopes_HillasReconstructor
    - n_LST_HillasReconstructor
    - n_MST_HillasReconstructor
    - mean_scaled_length
    - mean_scaled_width
    - HillasReconstructor_core_x
    - HillasReconstructor_core_y
    - HillasReconstructor_average_intensity
    - HillasReconstructor_h_max
    - HillasReconstructor_alt
    - HillasReconstructor_az
    - RandomForestClassifier_prediction
    - RandomForestRegressor_energy
```

Energy Regression (Array)

- Use (averaged) telescope predictions and array-wide features
- No hyperparameter optimization (yet)
- 5-fold cross-validation on **461 969** events



Configuration array energy regressor

```
model_cls: RandomForestRegressor
model_config:
    n_estimators: 200
    max_features: "sqrt"
    min_samples_leaf: 0.00001
    n_jobs: 40

log_target: True

features:
    - n_telescopes_HillasReconstructor
    - n_LST_HillasReconstructor
    - n_MST_HillasReconstructor
    - mean_scaled_length
    - mean_scaled_width
    - HillasReconstructor_core_x
    - HillasReconstructor_core_y
    - HillasReconstructor_average_intensity
    - HillasReconstructor_h_max
    - HillasReconstructor_alt
    - HillasReconstructor_az
    - RandomForestClassifier_prediction
    - RandomForestRegressor_energy
```

Gamma-Hadron Classification

Gamma-Hadron Classification (Telescope)

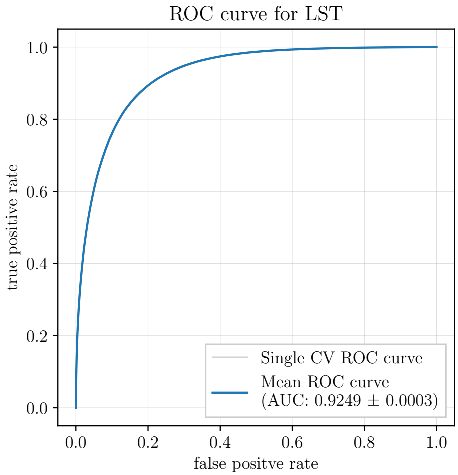
Randomized hyperparameter optimization yields:

Configuration particle classifier

```
ParticleClassifier:
  model_cls: RandomForestClassifier
  model_config:
    n_estimators: 69
    max_features: 0.5227
    max_samples: 0.7138
    min_samples_leaf: 0.000013
    n_jobs: 40
```

5-fold cross-validation:

- LST: **552 754** signal + **561 171** background events
- MST: **1 199 267** signal + **1 122 374** background events



Gamma-Hadron Classification (Telescope)

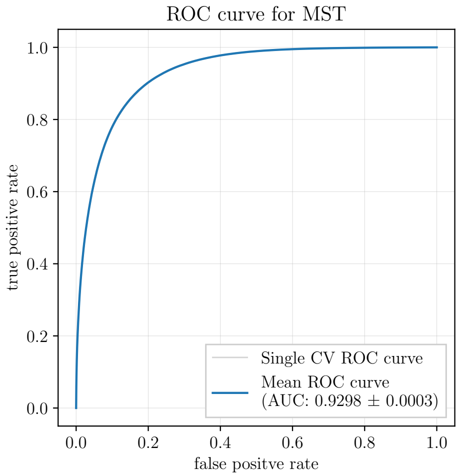
Randomized hyperparameter optimization yields:

Configuration particle classifier

```
ParticleClassifier:
  model_cls: RandomForestClassifier
  model_config:
    n_estimators: 69
    max_features: 0.5227
    max_samples: 0.7138
    min_samples_leaf: 0.000013
    n_jobs: 40
```

5-fold cross-validation:

- LST: **552 754** signal + **561 171** background events
- MST: **1 199 267** signal + **1 122 374** background events



Gamma-Hadron Classification (Telescope)

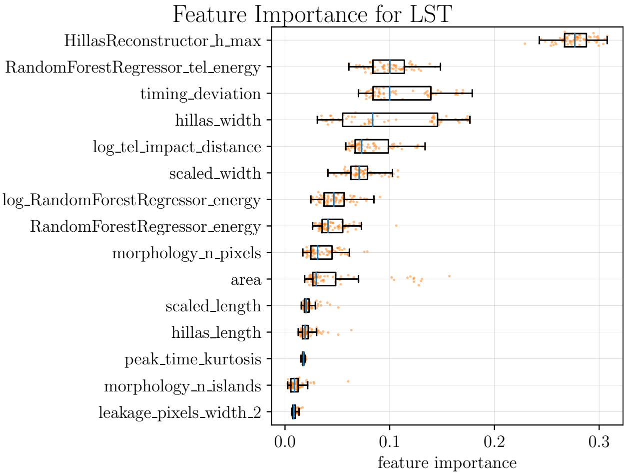
Randomized hyperparameter optimization yields:

Configuration particle classifier

```
ParticleClassifier:
  model_cls: RandomForestClassifier
  model_config:
    n_estimators: 69
    max_features: 0.5227
    max_samples: 0.7138
    min_samples_leaf: 0.000013
    n_jobs: 40
```

5-fold cross-validation:

- LST: **552 754** signal + **561 171** background events
- MST: **1 199 267** signal + **1 122 374** background events



Gamma-Hadron Classification (Telescope)

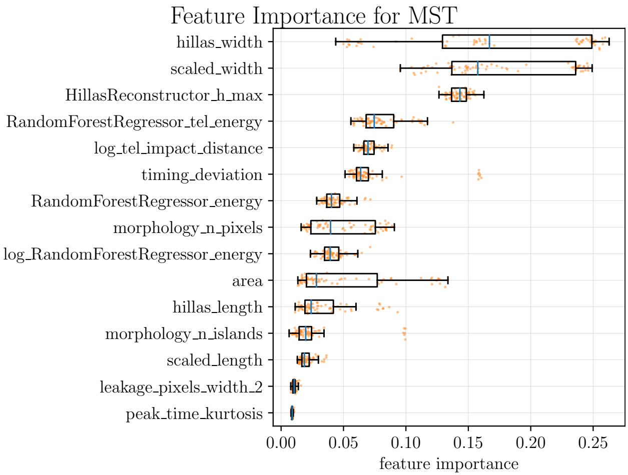
Randomized hyperparameter optimization yields:

Configuration particle classifier

```
ParticleClassifier:
  model_cls: RandomForestClassifier
  model_config:
    n_estimators: 69
    max_features: 0.5227
    max_samples: 0.7138
    min_samples_leaf: 0.000013
    n_jobs: 40
```

5-fold cross-validation:

- LST: **552 754** signal + **561 171** background events
- MST: **1 199 267** signal + **1 122 374** background events



Gamma-Hadron Classification (Array)

- Use (averaged) telescope predictions and array-wide features
- No hyperparameter optimization (yet)
- 5-fold cross-validation on **910 387** events

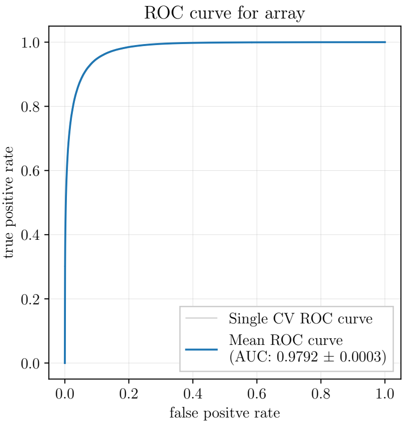
Configuration array particle classifier

```yaml
model_cls: RandomForestClassifier
model_config:
  n_estimators: 200
  max_features: "sqrt"
  min_samples_leaf: 0.00001
  n_jobs: 40

features:
  - n_telescopes_HillasReconstructor
  - n_LST_HillasReconstructor
  - n_MST_HillasReconstructor
  - mean_scaled_length
  - mean_scaled_width
  - HillasReconstructor_core_x
  - HillasReconstructor_core_y
  - HillasReconstructor_average_intensity
  - HillasReconstructor_h_max
  - HillasReconstructor_alt
  - HillasReconstructor_az
  - RandomForestClassifier_prediction
  - RandomForestRegressor_energy
```

Gamma-Hadron Classification (Array)

- Use (averaged) telescope predictions and array-wide features
- No hyperparameter optimization (yet)
- 5-fold cross-validation on **910 387** events



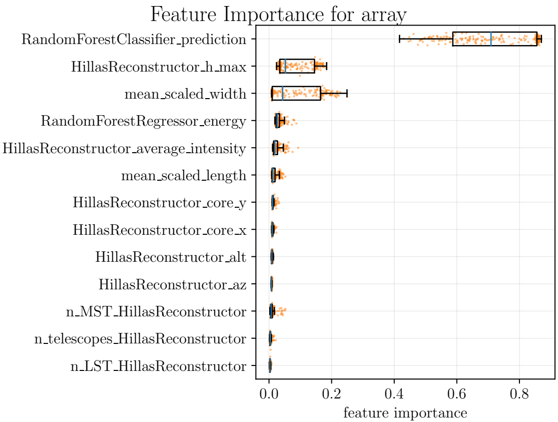
Configuration array particle classifier

```yaml
model_cls: RandomForestClassifier
model_config:
  n_estimators: 200
  max_features: "sqrt"
  min_samples_leaf: 0.00001
  n_jobs: 40

features:
  - n_telescopes_HillasReconstructor
  - n_LST_HillasReconstructor
  - n_MST_HillasReconstructor
  - mean_scaled_length
  - mean_scaled_width
  - HillasReconstructor_core_x
  - HillasReconstructor_core_y
  - HillasReconstructor_average_intensity
  - HillasReconstructor_h_max
  - HillasReconstructor_alt
  - HillasReconstructor_az
  - RandomForestClassifier_prediction
  - RandomForestRegressor_energy
```

Gamma-Hadron Classification (Array)

- Use (averaged) telescope predictions and array-wide features
- No hyperparameter optimization (yet)
- 5-fold cross-validation on **910 387** events



Configuration array particle classifier

```yaml
model_cls: RandomForestClassifier
model_config:
  n_estimators: 200
  max_features: "sqrt"
  min_samples_leaf: 0.00001
  n_jobs: 40

features:
  - n_telescopes_HillasReconstructor
  - n_LST_HillasReconstructor
  - n_MST_HillasReconstructor
  - mean_scaled_length
  - mean_scaled_width
  - HillasReconstructor_core_x
  - HillasReconstructor_core_y
  - HillasReconstructor_average_intensity
  - HillasReconstructor_h_max
  - HillasReconstructor_alt
  - HillasReconstructor_az
  - RandomForestClassifier_prediction
  - RandomForestRegressor_energy
```

Origin Reconstruction

Origin Reconstruction using disp

Randomized hyperparameter optimization yields:

Configuration disp reconstructor

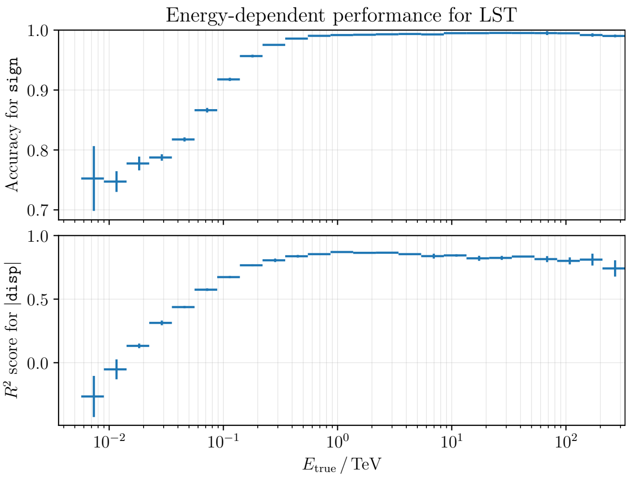
```
norm_cls: RandomForestRegressor
norm_config:
  n_estimators: 69
  max_features: 0.5227
  max_samples: 0.7138
  min_samples_leaf: 0.000013
  n_jobs: 40

log_target: True

sign_cls: RandomForestClassifier
sign_config:
  n_estimators: 343
  max_features: 0.6587
  max_samples: 0.5815
  min_samples_leaf: 0.000035
  n_jobs: 40
```

5-fold cross-validation:

- 552 754 LST events
- 1 199 267 MST events



Origin Reconstruction using disp

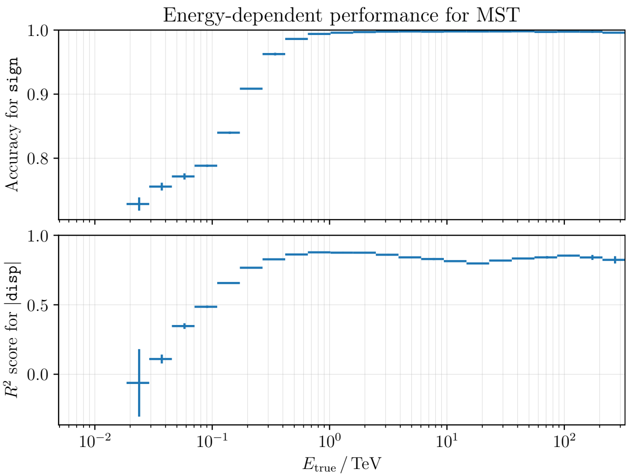
Randomized hyperparameter optimization yields:

Configuration disp reconstructor

```
norm_cls: RandomForestRegressor
norm_config:
  n_estimators: 69
  max_features: 0.5227
  max_samples: 0.7138
  min_samples_leaf: 0.000013
  n_jobs: 40

log_target: True

sign_cls: RandomForestClassifier
sign_config:
  n_estimators: 343
  max_features: 0.6587
  max_samples: 0.5815
  min_samples_leaf: 0.000035
  n_jobs: 40
```



5-fold cross-validation:

- 552 754 LST events
- 1 199 267 MST events

Origin Reconstruction using disp

Randomized hyperparameter optimization yields:

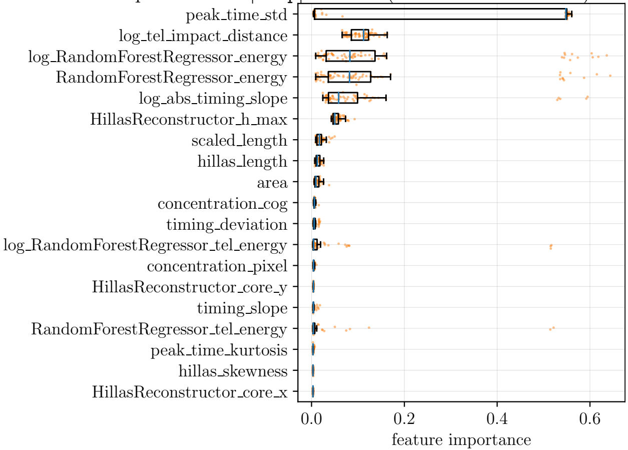
Configuration disp reconstructor

```
norm_cls: RandomForestRegressor
norm_config:
  n_estimators: 69
  max_features: 0.5227
  max_samples: 0.7138
  min_samples_leaf: 0.000013
  n_jobs: 40

log_target: True

sign_cls: RandomForestClassifier
sign_config:
  n_estimators: 343
  max_features: 0.6587
  max_samples: 0.5815
  min_samples_leaf: 0.000035
  n_jobs: 40
```

Feature Importance for |disp| for LST ($R^2 = 0.9525 \pm 0.0013$)



5-fold cross-validation:

- 552 754 LST events
- 1 199 267 MST events

Origin Reconstruction using disp

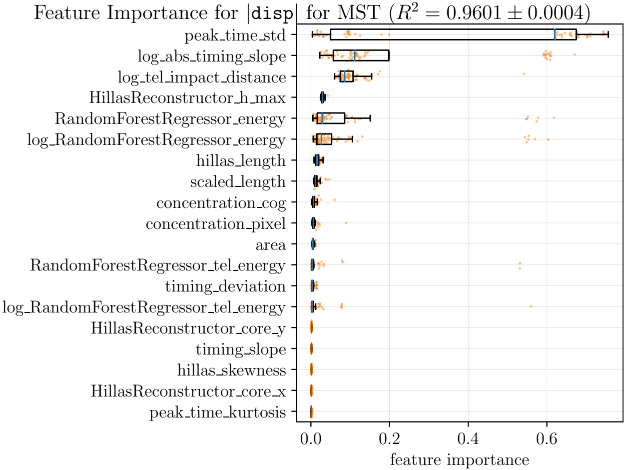
Randomized hyperparameter optimization yields:

Configuration disp reconstructor

```
norm_cls: RandomForestRegressor
norm_config:
  n_estimators: 69
  max_features: 0.5227
  max_samples: 0.7138
  min_samples_leaf: 0.000013
  n_jobs: 40

log_target: True

sign_cls: RandomForestClassifier
sign_config:
  n_estimators: 343
  max_features: 0.6587
  max_samples: 0.5815
  min_samples_leaf: 0.000035
  n_jobs: 40
```



5-fold cross-validation:

- 552 754 LST events
- 1 199 267 MST events

Origin Reconstruction using disp

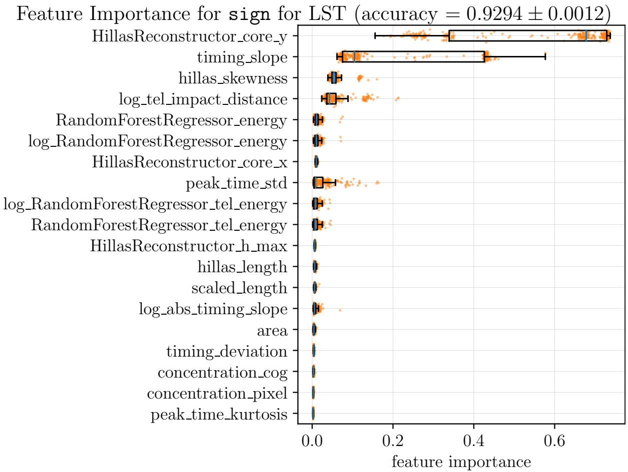
Randomized hyperparameter optimization yields:

Configuration disp reconstructor

```
norm_cls: RandomForestRegressor
norm_config:
  n_estimators: 69
  max_features: 0.5227
  max_samples: 0.7138
  min_samples_leaf: 0.000013
  n_jobs: 40

log_target: True

sign_cls: RandomForestClassifier
sign_config:
  n_estimators: 343
  max_features: 0.6587
  max_samples: 0.5815
  min_samples_leaf: 0.000035
  n_jobs: 40
```



5-fold cross-validation:

- 552 754 LST events
- 1 199 267 MST events

Origin Reconstruction using disp

Randomized hyperparameter optimization yields:

Configuration disp reconstructor

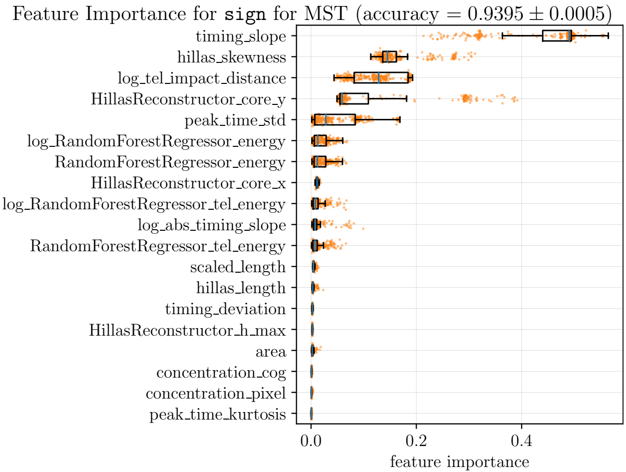
```
norm_cls: RandomForestRegressor
norm_config:
  n_estimators: 69
  max_features: 0.5227
  max_samples: 0.7138
  min_samples_leaf: 0.000013
  n_jobs: 40

log_target: True

sign_cls: RandomForestClassifier
sign_config:
  n_estimators: 343
  max_features: 0.6587
  max_samples: 0.5815
  min_samples_leaf: 0.000035
  n_jobs: 40
```

5-fold cross-validation:

- 552 754 LST events
- 1 199 267 MST events



Origin Reconstruction (Array)

- Predict 3D cartesian position based on unit-sphere $(1, alt, az)$
- Use (averaged) telescope predictions and array-wide features
- No hyperparameter optimization (yet)
- 5-fold cross-validation on **461 212** events

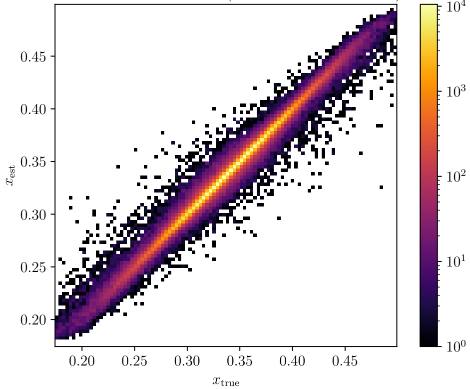
Configuration (every) array origin regressor

```yaml
model_cls: RandomForestRegressor
model_config:
  n_estimators: 200
  max_features: "sqrt"
  min_samples_leaf: 0.00001
  n_jobs: 40

features:
  - n_telescopes_HillasReconstructor
  - n_LST_HillasReconstructor
  - n_MST_HillasReconstructor
  - mean_scaled_length
  - mean_scaled_width
  - HillasReconstructor_core_x
  - HillasReconstructor_core_y
  - HillasReconstructor_average_intensity
  - HillasReconstructor_h_max
  - HillasReconstructor_alt
  - HillasReconstructor_az
  - disp_alt
  - disp_az
  - disp_ang_distance_uncert
  - RandomForestClassifier_prediction
  - RandomForestRegressor_energy
```

Origin Reconstruction (Array)

- Predict 3D cartesian position based on unit-sphere **(1, *alt*, *az*)**
- Use (averaged) telescope predictions and array-wide features
- No hyperparameter optimization (yet)
- 5-fold cross-validation on **461 212** events



Configuration (every) array origin regressor

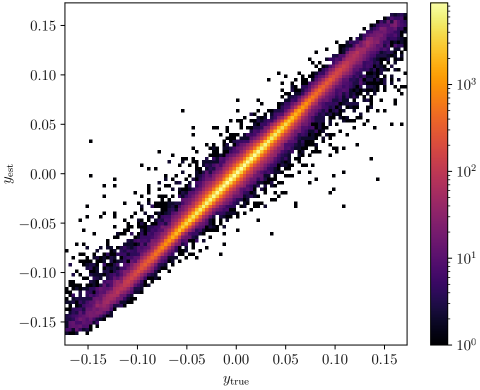
```
model_cls: RandomForestRegressor
model_config:
  n_estimators: 200
  max_features: "sqrt"
  min_samples_leaf: 0.00001
  n_jobs: 40

features:
  - n_telescopes_HillasReconstructor
  - n_LST_HillasReconstructor
  - n_MST_HillasReconstructor
  - mean_scaled_length
  - mean_scaled_width
  - HillasReconstructor_core_x
  - HillasReconstructor_core_y
  - HillasReconstructor_average_intensity
  - HillasReconstructor_h_max
  - HillasReconstructor_alt
  - HillasReconstructor_az
  - disp_alt
  - disp_az
  - disp_ang_distance_uncert
  - RandomForestClassifier_prediction
  - RandomForestRegressor_energy
```

Origin Reconstruction (Array)

- Predict 3D cartesian position based on unit-sphere **(1, *alt*, *az*)**
- Use (averaged) telescope predictions and array-wide features
- No hyperparameter optimization (yet)
- 5-fold cross-validation on **461 212** events

Confusion Matrix of y ($R^2 = 0.9855 \pm 0.0004$)



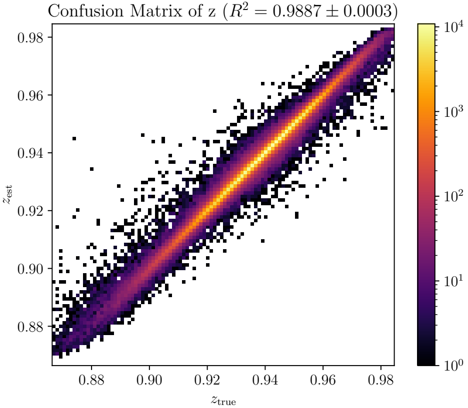
Configuration (every) array origin regressor

```yaml
model_cls: RandomForestRegressor
model_config:
  n_estimators: 200
  max_features: "sqrt"
  min_samples_leaf: 0.00001
  n_jobs: 40

features:
  - n_telescopes_HillasReconstructor
  - n_LST_HillasReconstructor
  - n_MST_HillasReconstructor
  - mean_scaled_length
  - mean_scaled_width
  - HillasReconstructor_core_x
  - HillasReconstructor_core_y
  - HillasReconstructor_average_intensity
  - HillasReconstructor_h_max
  - HillasReconstructor_alt
  - HillasReconstructor_az
  - disp_alt
  - disp_az
  - disp_ang_distance_uncert
  - RandomForestClassifier_prediction
  - RandomForestRegressor_energy
```

Origin Reconstruction (Array)

- Predict 3D cartesian position based on unit-sphere **(1, *alt*, *az*)**
- Use (averaged) telescope predictions and array-wide features
- No hyperparameter optimization (yet)
- 5-fold cross-validation on **461 212** events



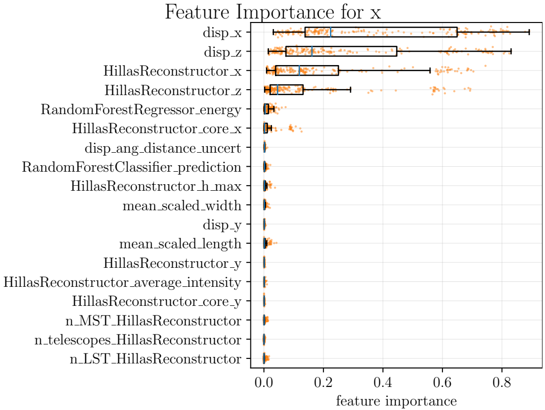
Configuration (every) array origin regressor

```
model_cls: RandomForestRegressor
model_config:
  n_estimators: 200
  max_features: "sqrt"
  min_samples_leaf: 0.00001
  n_jobs: 40

features:
  - n_telescopes_HillasReconstructor
  - n_LST_HillasReconstructor
  - n_MST_HillasReconstructor
  - mean_scaled_length
  - mean_scaled_width
  - HillasReconstructor_core_x
  - HillasReconstructor_core_y
  - HillasReconstructor_average_intensity
  - HillasReconstructor_h_max
  - HillasReconstructor_alt
  - HillasReconstructor_az
  - disp_alt
  - disp_az
  - disp_ang_distance_uncert
  - RandomForestClassifier_prediction
  - RandomForestRegressor_energy
```

Origin Reconstruction (Array)

- Predict 3D cartesian position based on unit-sphere **(1, *alt*, *az*)**
- Use (averaged) telescope predictions and array-wide features
- No hyperparameter optimization (yet)
- **5**-fold cross-validation on **461 212** events



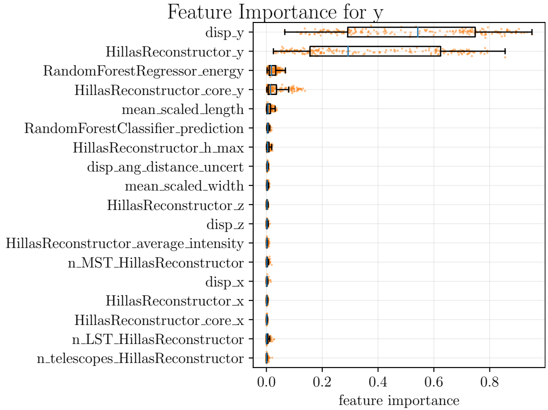
Configuration (every) array origin regressor

```
model_cls: RandomForestRegressor
model_config:
  n_estimators: 200
  max_features: "sqrt"
  min_samples_leaf: 0.00001
  n_jobs: 40

features:
  - n_telescopes_HillasReconstructor
  - n_LST_HillasReconstructor
  - n_MST_HillasReconstructor
  - mean_scaled_length
  - mean_scaled_width
  - HillasReconstructor_core_x
  - HillasReconstructor_core_y
  - HillasReconstructor_average_intensity
  - HillasReconstructor_h_max
  - HillasReconstructor_alt
  - HillasReconstructor_az
  - disp_alt
  - disp_az
  - disp_ang_distance_uncert
  - RandomForestClassifier_prediction
  - RandomForestRegressor_energy
```

Origin Reconstruction (Array)

- Predict 3D cartesian position based on unit-sphere $(1, alt, az)$
- Use (averaged) telescope predictions and array-wide features
- No hyperparameter optimization (yet)
- 5-fold cross-validation on **461 212** events



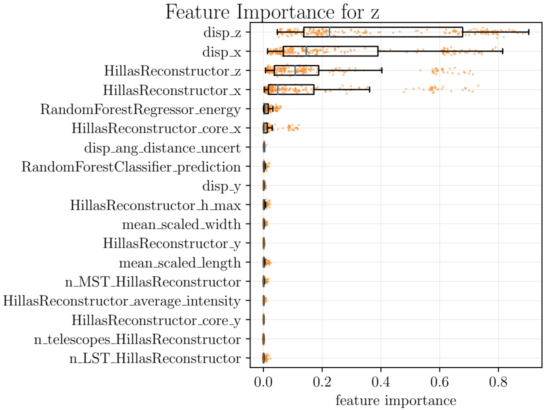
Configuration (every) array origin regressor

```yaml
model_cls: RandomForestRegressor
model_config:
  n_estimators: 200
  max_features: "sqrt"
  min_samples_leaf: 0.00001
  n_jobs: 40

features:
  - n_telescopes_HillasReconstructor
  - n_LST_HillasReconstructor
  - n_MST_HillasReconstructor
  - mean_scaled_length
  - mean_scaled_width
  - HillasReconstructor_core_x
  - HillasReconstructor_core_y
  - HillasReconstructor_average_intensity
  - HillasReconstructor_h_max
  - HillasReconstructor_alt
  - HillasReconstructor_az
  - disp_alt
  - disp_az
  - disp_ang_distance_uncert
  - RandomForestClassifier_prediction
  - RandomForestRegressor_energy
```

Origin Reconstruction (Array)

- Predict 3D cartesian position based on unit-sphere **(1, *alt*, *az*)**
- Use (averaged) telescope predictions and array-wide features
- No hyperparameter optimization (yet)
- **5**-fold cross-validation on **461 212** events



Configuration (every) array origin regressor

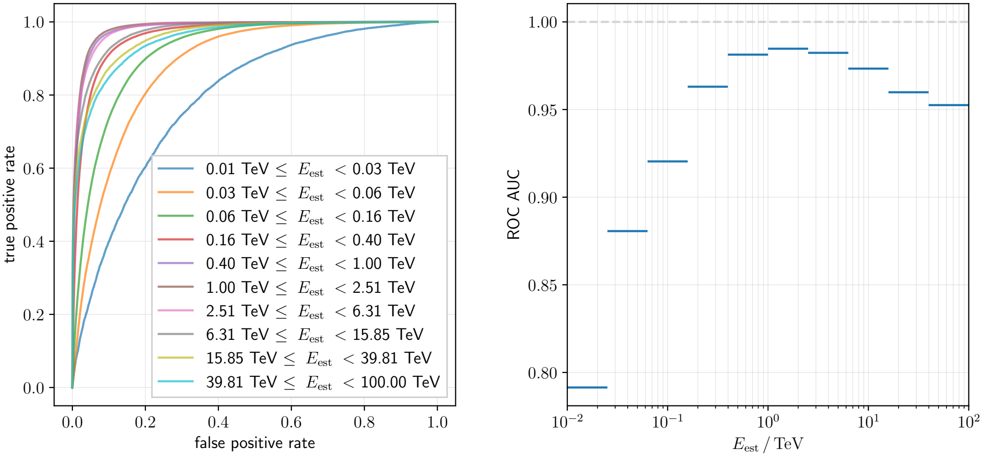
```
model_cls: RandomForestRegressor
model_config:
  n_estimators: 200
  max_features: "sqrt"
  min_samples_leaf: 0.00001
  n_jobs: 40

features:
  - n_telescopes_HillasReconstructor
  - n_LST_HillasReconstructor
  - n_MST_HillasReconstructor
  - mean_scaled_length
  - mean_scaled_width
  - HillasReconstructor_core_x
  - HillasReconstructor_core_y
  - HillasReconstructor_average_intensity
  - HillasReconstructor_h_max
  - HillasReconstructor_alt
  - HillasReconstructor_az
  - disp_alt
  - disp_az
  - disp_ang_distance_uncert
  - RandomForestClassifier_prediction
  - RandomForestRegressor_energy
```

Performance

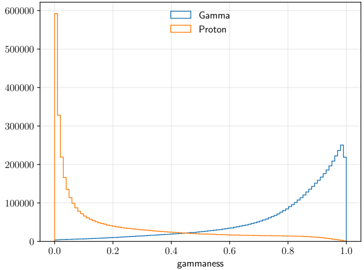
Gamma-Hadron Performance (Energy-Dependent)

Array classifier

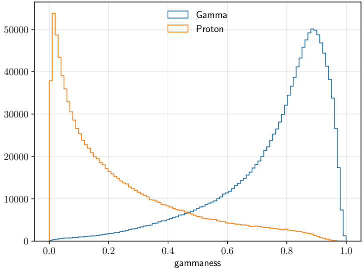


Gamma-Hadron Scores

Single telescope predictions

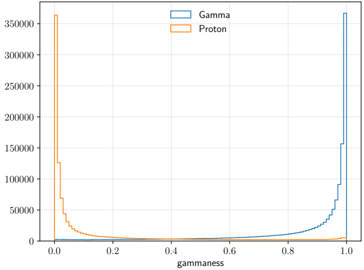


Averaged telescope predictions



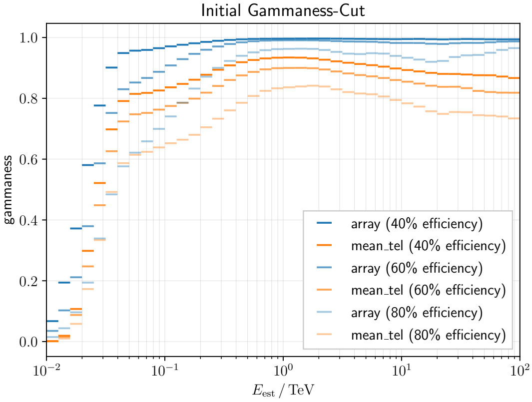
ROC AUC = 0.9442

Array classifier predictions

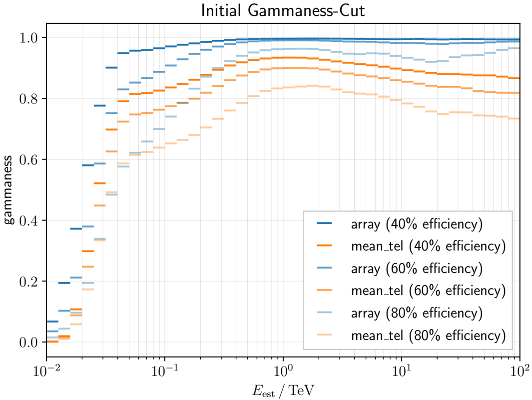


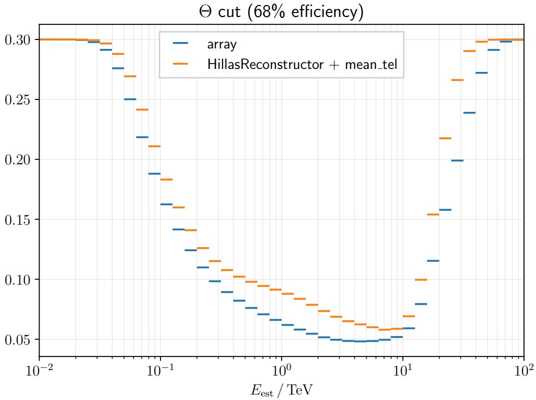
ROC AUC = 0.9538

Gammaness and θ Cuts



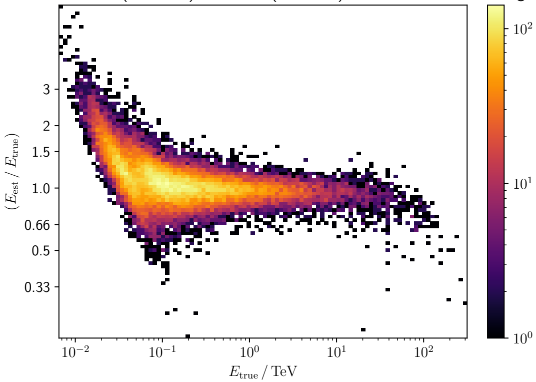
Gammaness and θ Cuts



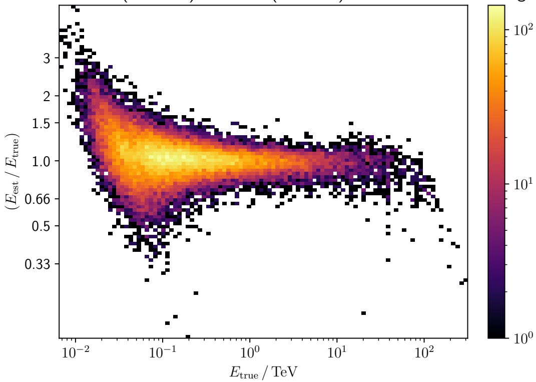


Energy Migration

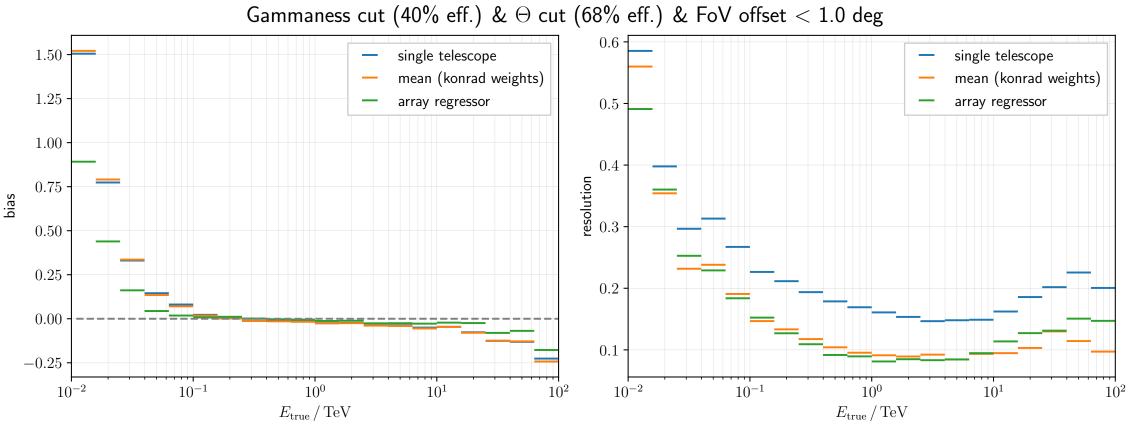
Mean telescope predictions



Array regressor

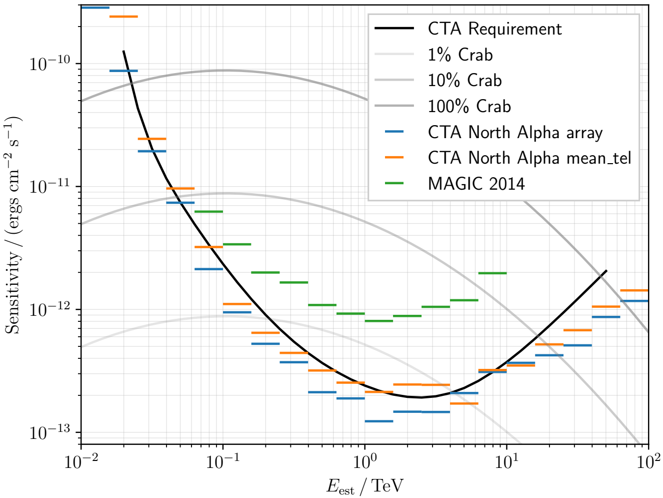


Energy – Bias and Resolution



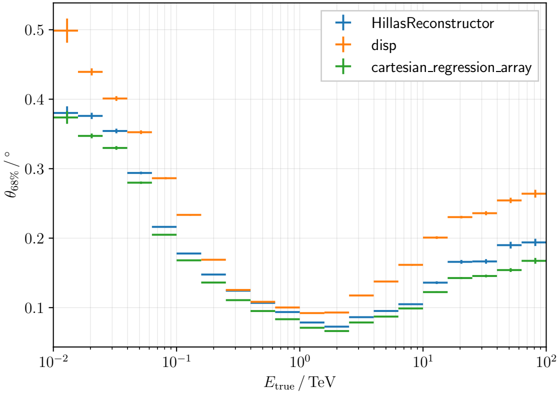
Sensitivity

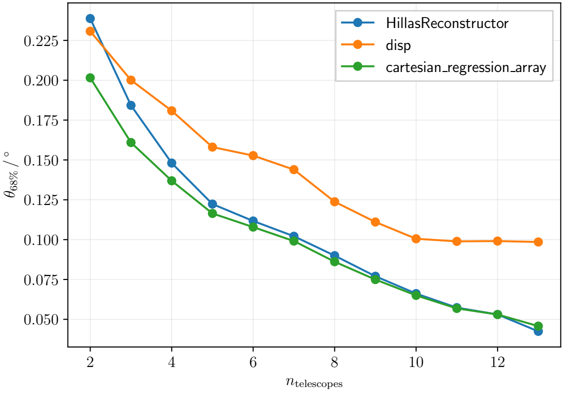
⇒ This and all following plots use gammaness cuts optimized for maximum sensitivity!



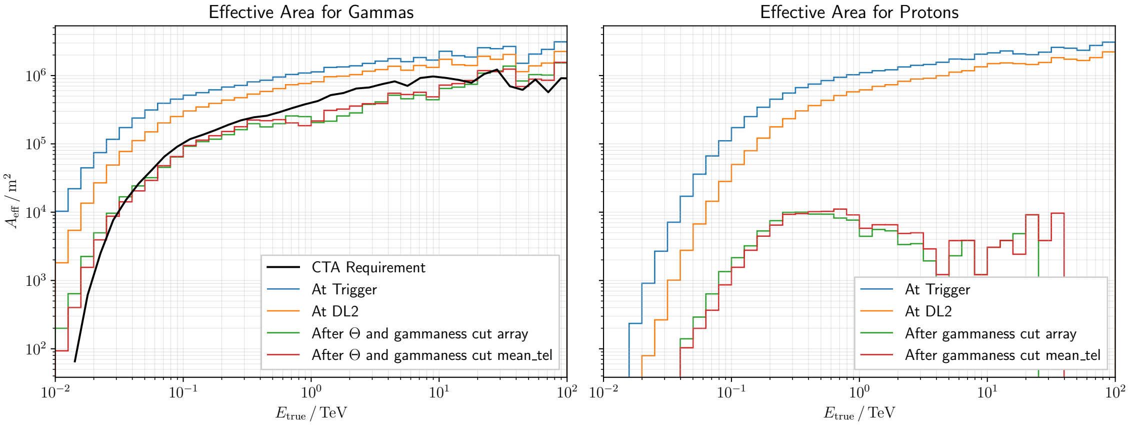
Angular Resolution

Optimized gammaness cut





Effective Area



Outlook

TODO

- Separate dataset for array models → optimize hyperparameters
- Try other methods for averaging disp predictions (→ Lukas' master thesis)
- Telescope models using only mono features → include "mono" events
- Try other ML algorithms (e.g. boosted decision trees) incl. hyperparameter optimization

Summary

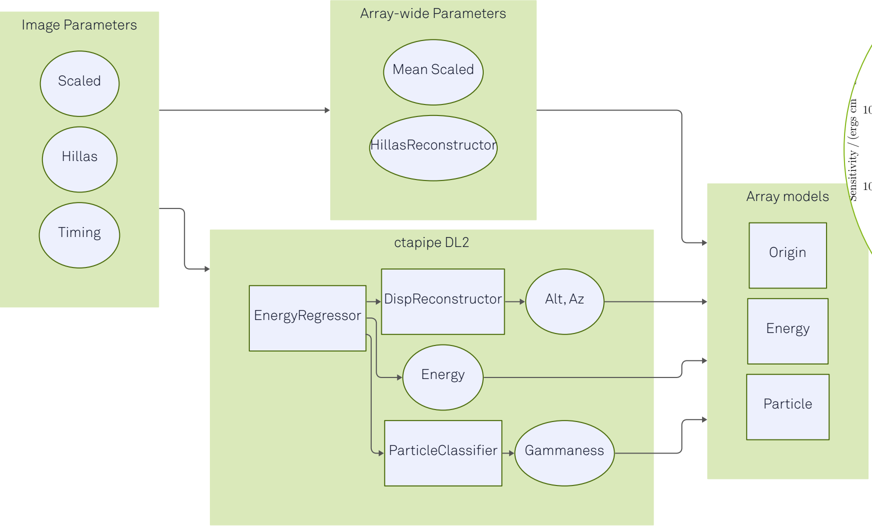
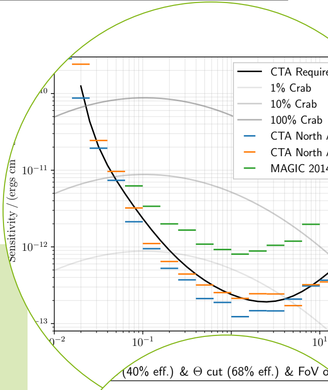
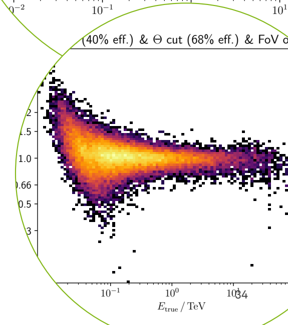


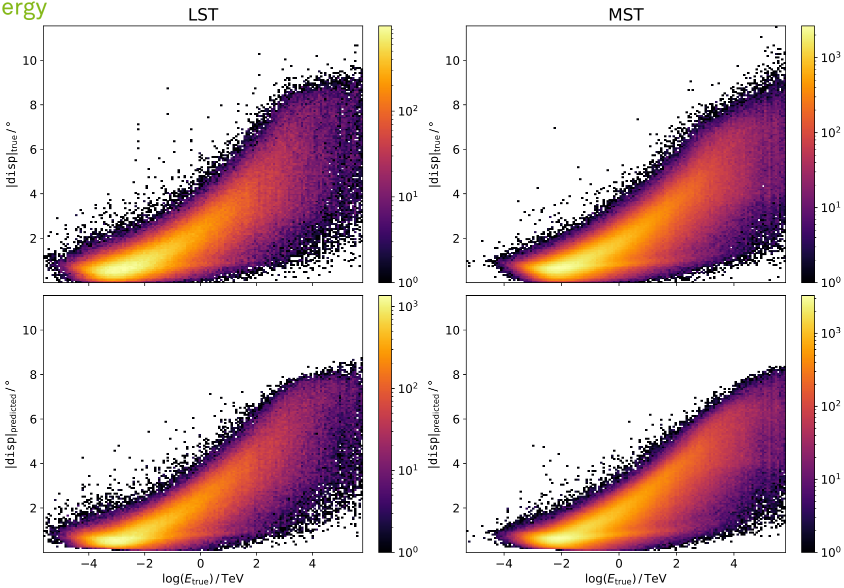
Image Parameters
Scaled
Hillas
Timing
Array-wide Parameters
Mean Scaled
HillasReconstructor
ctapipe DL2
EnergyRegressor
DispReconstructor
Alt, Az
Energy
ParticleClassifier
Gammaness
Array models
Origin
Energy
Particle


1% Crab
10% Crab
100% Crab
CTA North
CTA North
10^{-11}
10^{-12}
10^{-1}

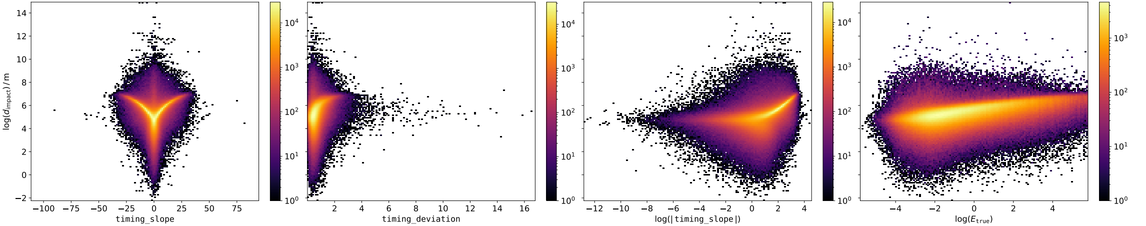

(40% eff.) & Θ cut (68% eff.) & FoV
10^{-1}
10^{0}
E_{true} / TeV

Backup

|disp| vs Energy



Timing Parameters vs Impact Distance

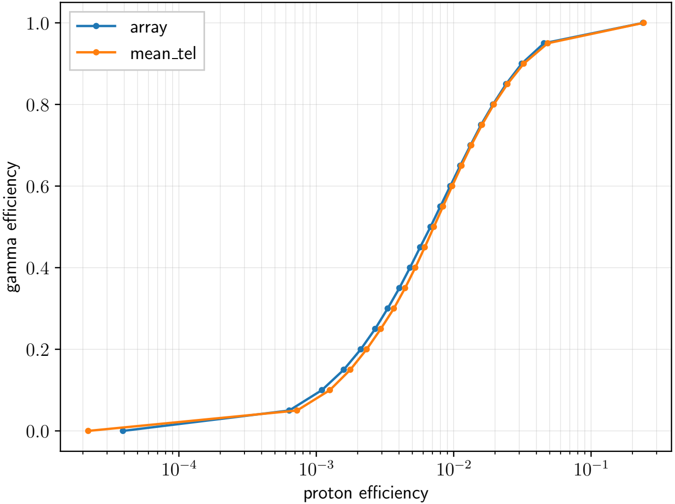


Quality Cuts

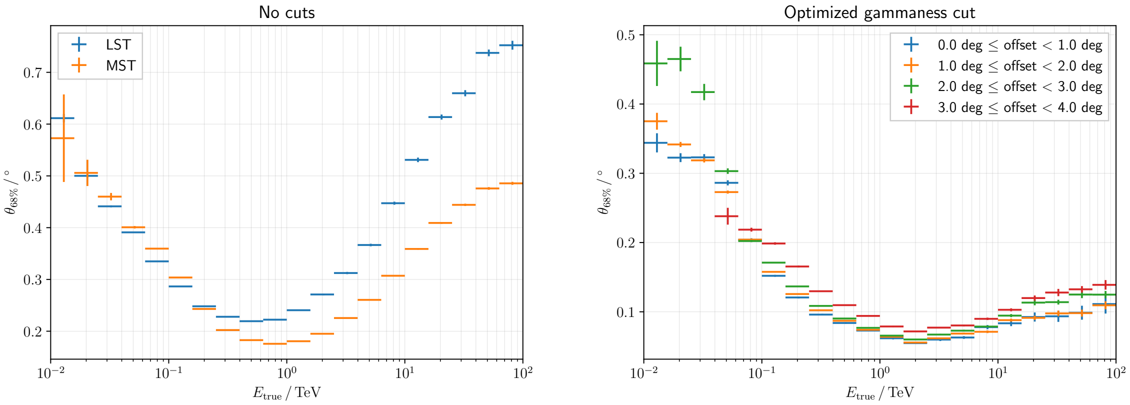
QualityQuery

```
quality_criteria:
  - ["enough intensity", "hillas_intensity > 50"]
  - ["Positive width", "hillas_width > 0"]
  - ["enough pixels", "morphology_n_pixels > 3"]
  - ["not clipped", "leakage_intensity_width_2 < 0.5"]
  - ["HillasValid", "HillasReconstructor_is_valid"]
```

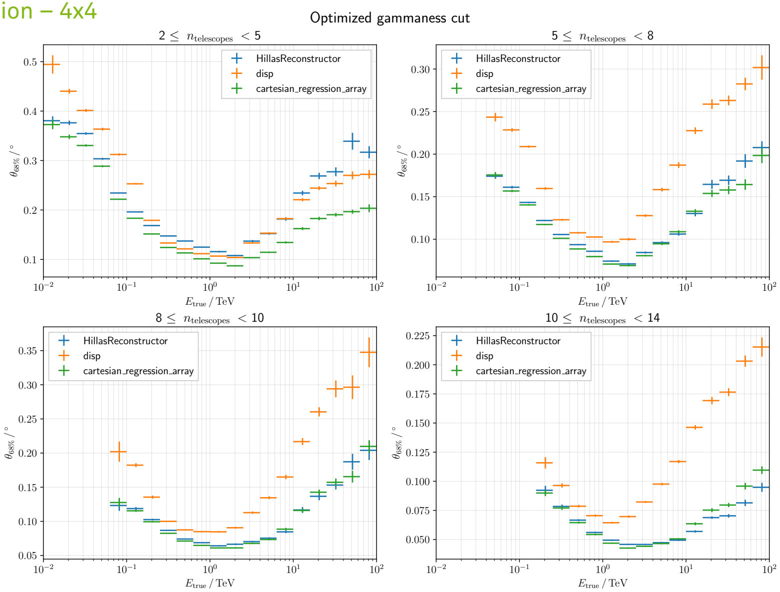
Background Rejection



Angular Resolution – More Plots

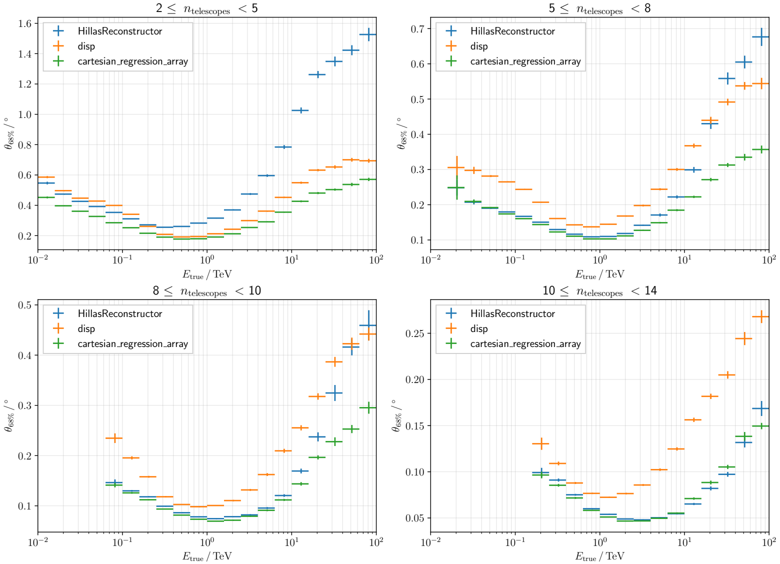


Angular Resolution – 4x4



Angular Resolution – 4x4

No cuts



Error of Mean disp Predicitons