



Gamma/Hadron-Separation bei FACT

Maximilian Sackel

1. Oktober 2017

Experimental Physiks 5b
Astroteilchenphysik

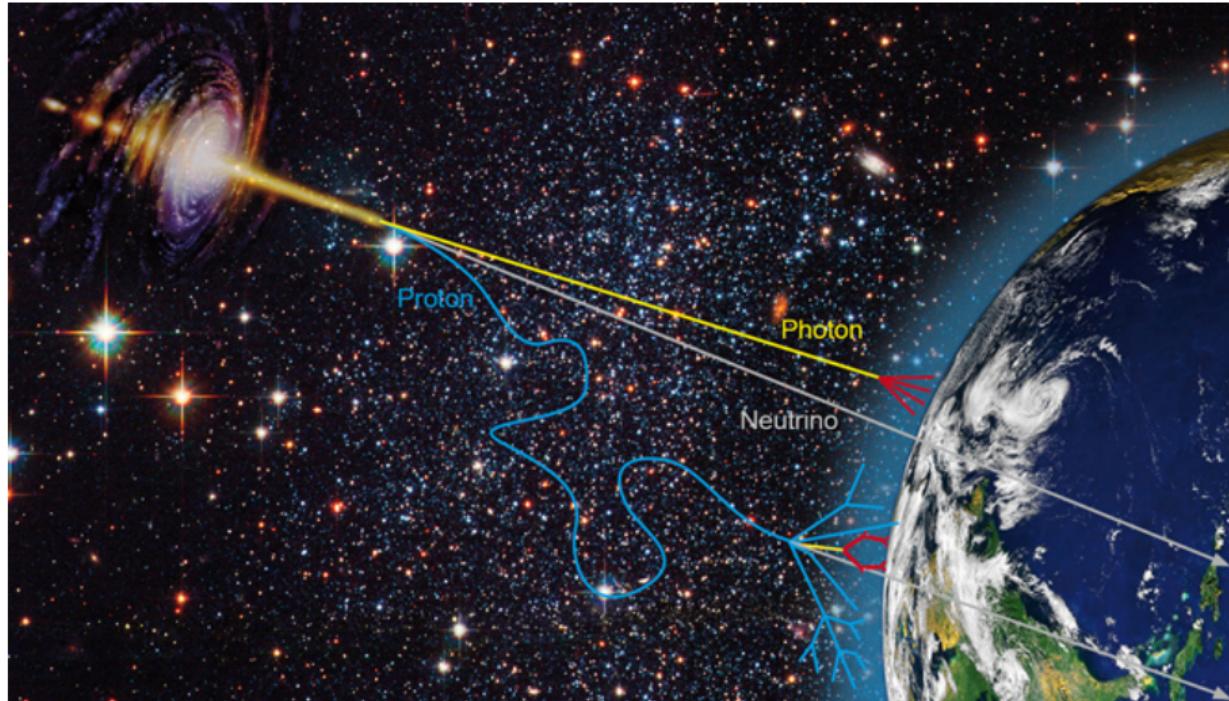
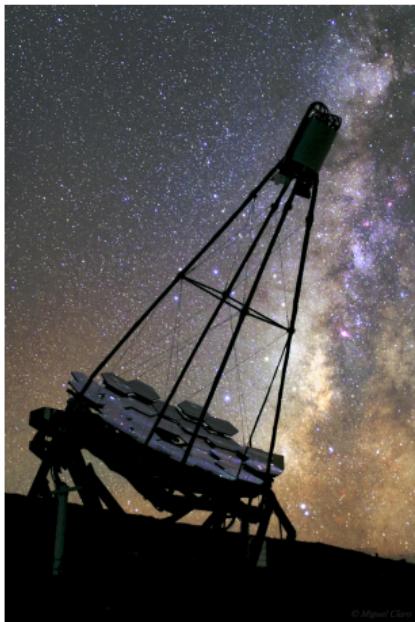


Abbildung: Astroparticle Physik. 2017. URL: <https://astro.desy.de/>



First G-APD Cherenkov Telescope

Abbildung: Carlo, *FACT Cherenkov Telescope in a Milky Way Backlight*



Localisation

Gran Canaria, Los
Rochos, 2200 m

Reflektor

Davis-Cotton-Design,
 4 m^2 Spiegelfläche

Kamera

1440 SiMPs, ro-
bust und sensitiv

Abbildung: Carlo, *FACT Cherenkov
Telescope in a Milky Way Backlight*

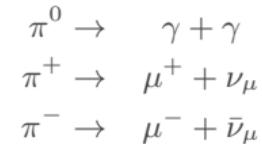
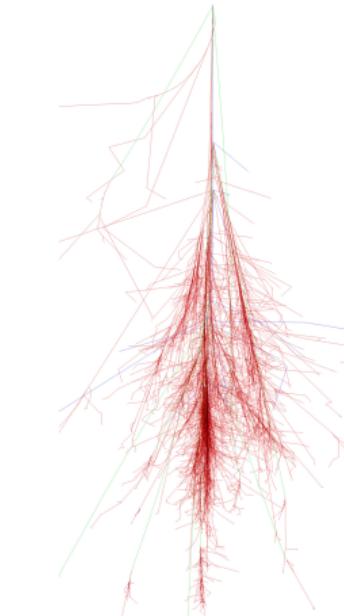
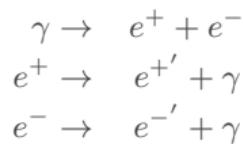
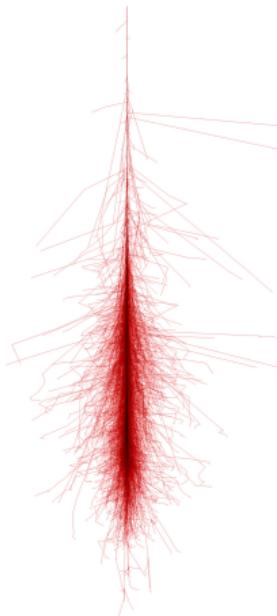


Abbildung: Carlo, FACT
Cherenkov Telescope in a
Milky Way Backlight

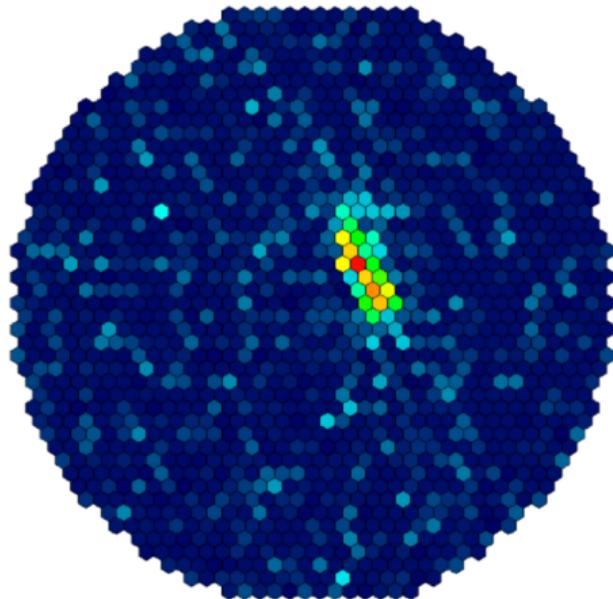


Abbildung: Carlo, FACT Cherenkov Telescope in a Milky Way Backlight

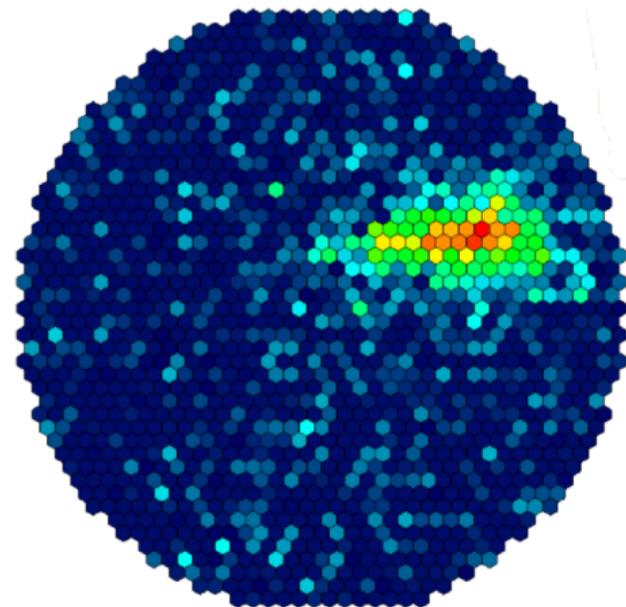
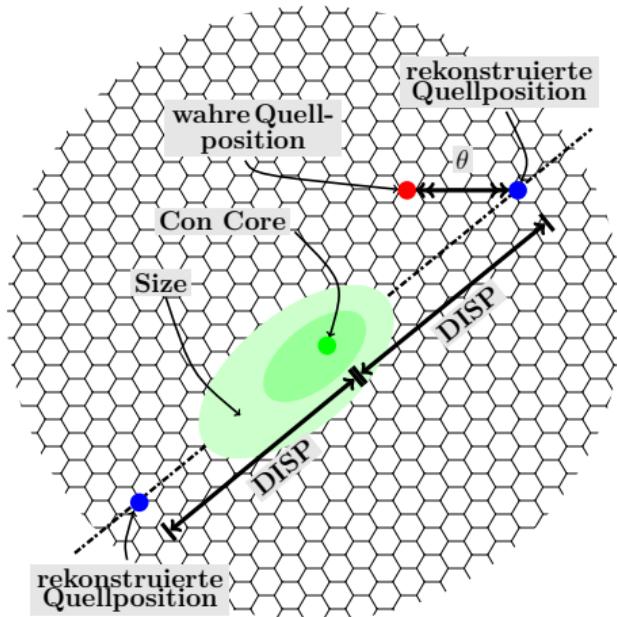
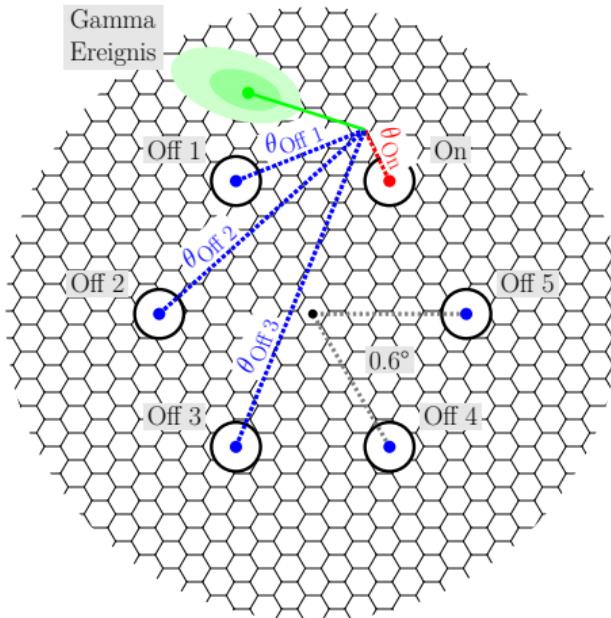


Abbildung: Carlo, FACT Cherenkov Telescope in a Milky Way Backlight

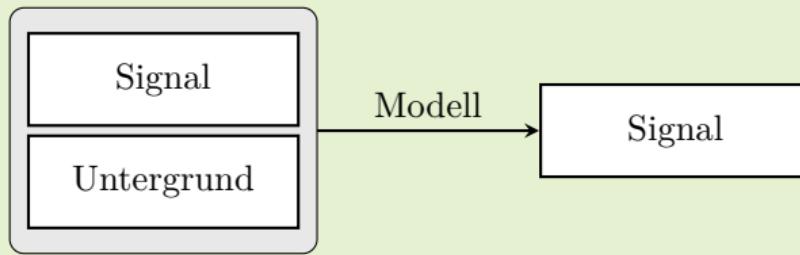


- berechne Feature (Hillas Parameter) des Kamerabildes
- Feature werden zum Klassifizieren benoetigt
- Vorzeichen des Schauers nicht eindeutig

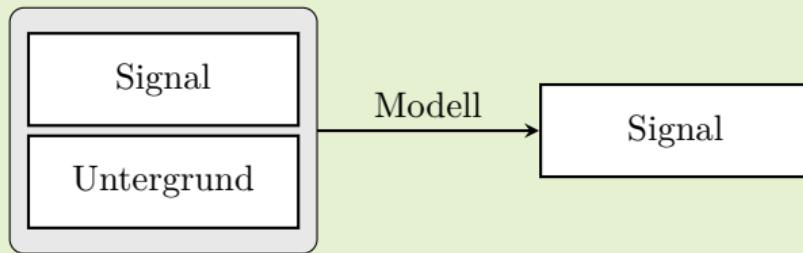


- FACT nimmt keine OFF-Daten
- Daten werden im Wobble Modus genommen

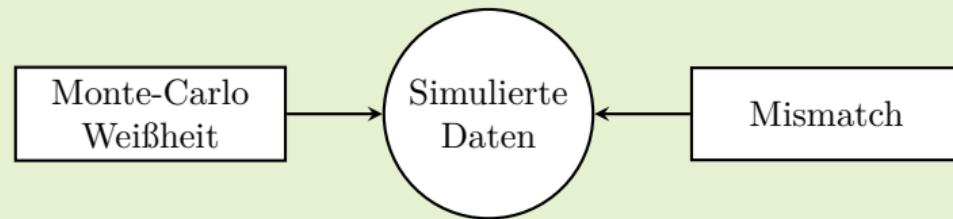
Separation



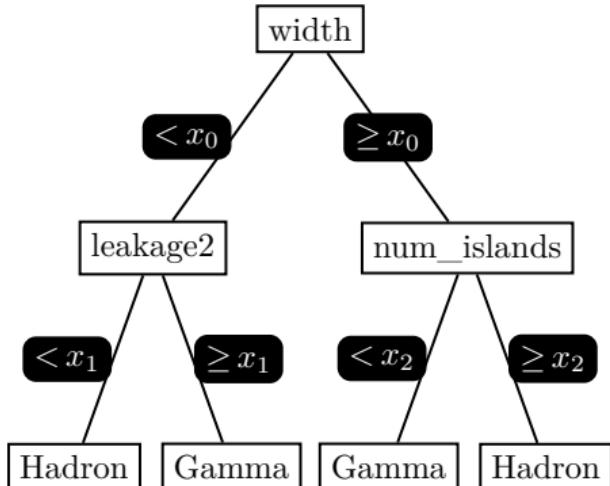
Separation



Simulierte Daten



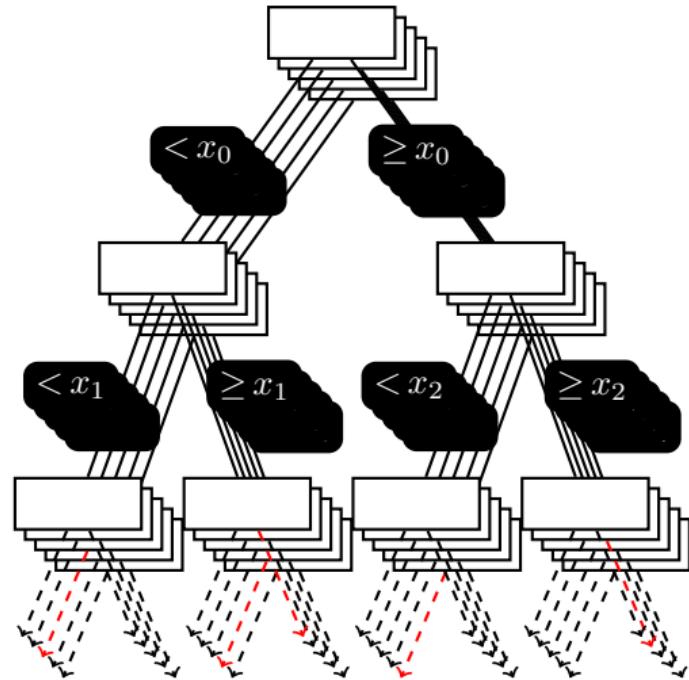
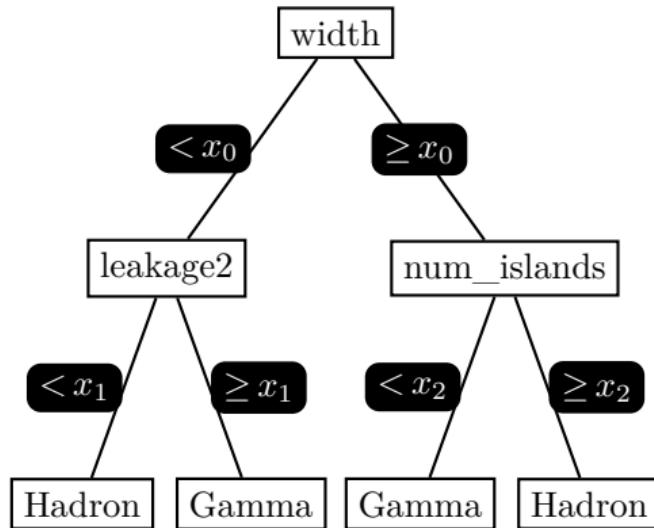
Entscheidungsbaum



- Verknüpfte Abfragen
- Loss-function
- Beschränkung der Komplexität

| Ereignis | width | leakage2 | num_islands | ... | Konfi. |
|----------|-------|----------|-------------|-----|--------|
| 1 | 4.2 | 0.4 | 3 | ... | 0.12 |
| 2 | 3.8 | 0.0 | 2 | ... | 0.56 |
| 3 | 15.3 | 0.8 | 1 | ... | 0.08 |
| 4 | 7.7 | 0.1 | 1 | ... | 0.43 |
| 5 | 6.2 | 0.0 | 1 | ... | 0.85 |

Random Forest



Random Forest

| Ereignis | Konfi. |
|----------|--------|
| 1 | 0.12 |
| 2 | 0.56 |
| 3 | 0.08 |
| 4 | 0.43 |
| 5 | 0.85 |

| Ereignis | Konf ₁ | Konf ₂ | Konf ₃ | ... | Σ_i Konf _i |
|----------|-------------------|-------------------|-------------------|-----|------------------------------|
| 1 | 0.12 | 0.01 | 0.08 | ... | 0.06 |
| 2 | 0.40 | 0.66 | 0.53 | ... | 0.56 |
| 3 | 0.02 | 0.17 | 0.10 | ... | 0.08 |
| 4 | 0.41 | 0.42 | 0.42 | ... | 0.43 |
| 5 | 0.96 | 0.81 | 0.85 | ... | 0.85 |

Boosted Trees



- additives Training
- höhere Gewichtung von Fehlklassifizierungen
- ausgeglichener Vorhersage
- lässt sich nicht Parallelisieren
- Modelle mit geringere Komplexität

Boosted Trees



- additives Training
- höhere Gewichtung von Fehlklassifizierungen
- ausgeglichener Vorhersage
- lässt sich nicht Parallelisieren
- Modelle mit geringere Komplexität

Boosted Trees

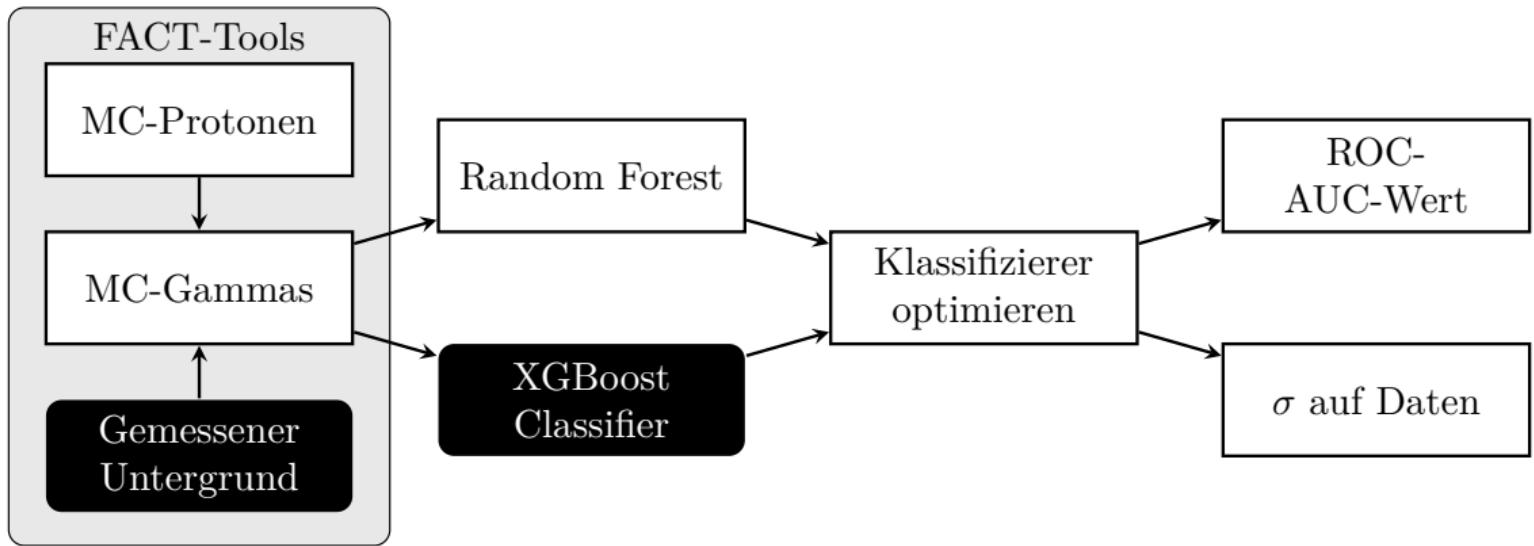


- additives Training
- höhere Gewichtung von Fehlklassifizierungen
- ausgeglichener Vorhersage
- lässt sich nicht Parallelisieren
- Modelle mit geringere Komplexität

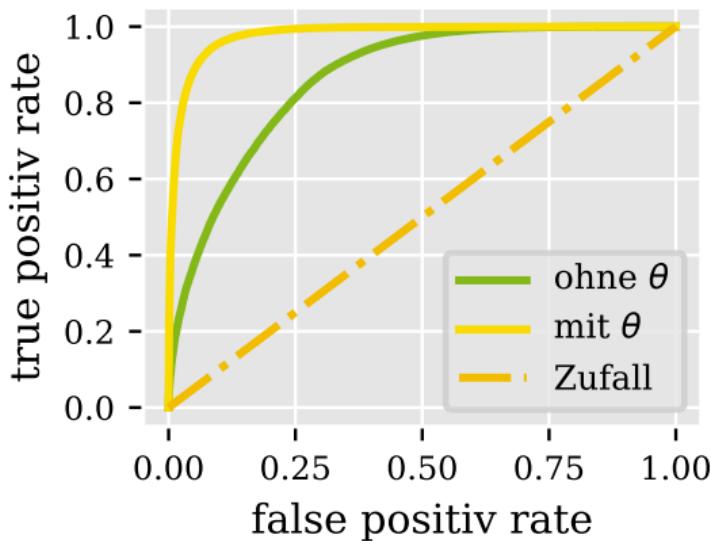
Boosted Trees



- additives Training
- höhere Gewichtung von Fehlklassifizierungen
- ausgeglichener Vorhersage
- lässt sich nicht Parallelisieren
- Modelle mit geringere Komplexität



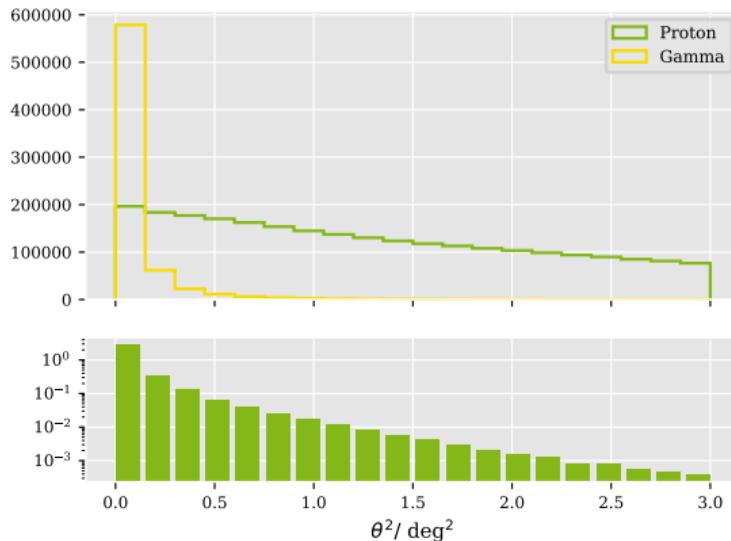
Erstellen Trainingsdatensatzes



Bereinigen der gemessenen Daten von
Gamma-Ereignissen

- ohne Eingang der Monte-Carlo Mismatche
- möglichst Reinen Datensatz
- paar diffuse Gamma lassen sich physikalisch motivieren

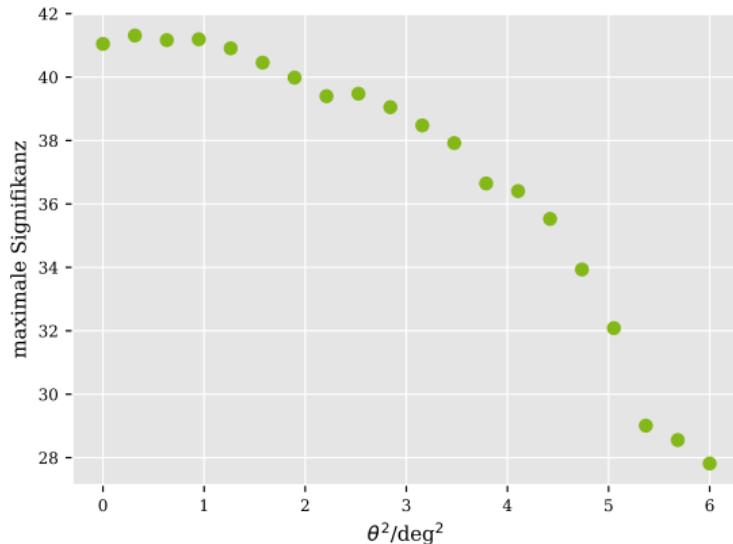
Erstellen Trainingsdatensatzes



Bereinigen der gemessenen Daten von
Gamma-Ereignissen

- ohne Eingang der Monte-Carlo Mismatche
- möglichst Reinen Datensatz
- paar diffuse Gamma lassen sich physikalisch motivieren

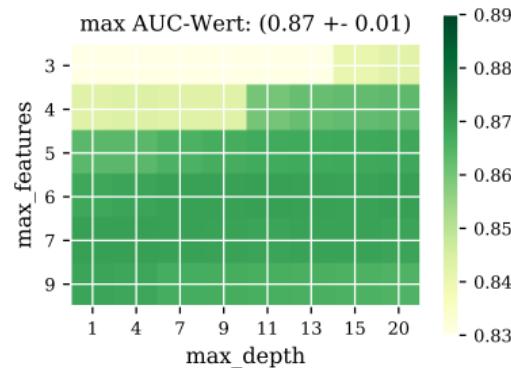
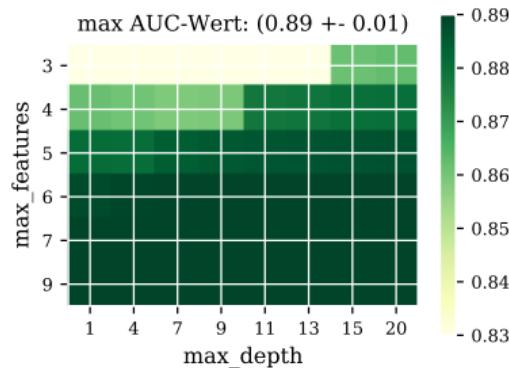
Ueberpruefen Trainingsdatensatzes



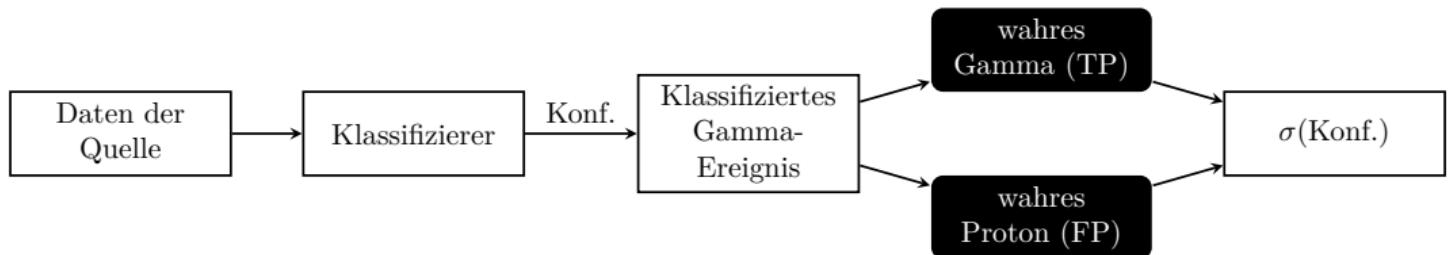
Trainingsdatensatz aus simulierte Gamma- und Proton-Ereignissen

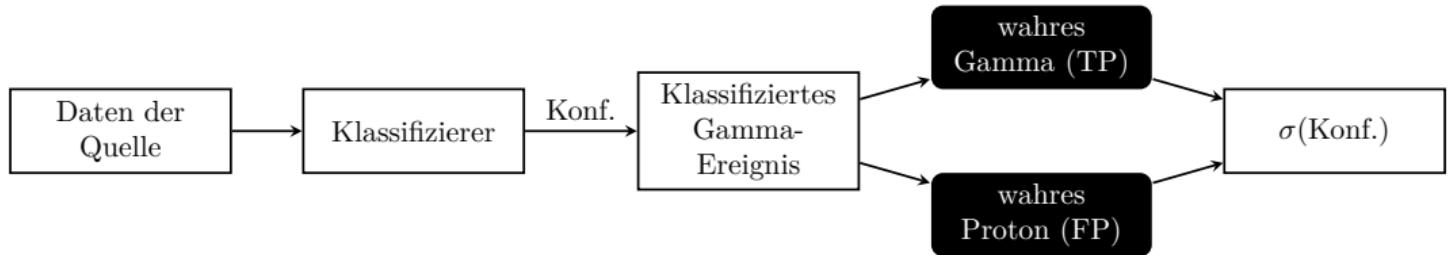
- Detektoreigenschaften fuer groÙe θ signifikant
- Signifikanz abhaengig vom θ -Schnitt
- θ -Schnitt in Abhaengigkeit der Reinheit und korrelation der Signifikanz

Optimieren der Modelle



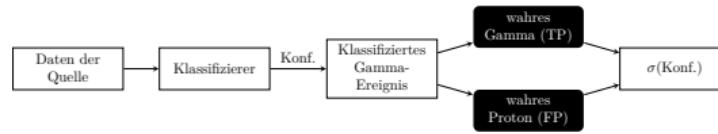
| | MC – Daten | Messdaten |
|------------------------------|------------|-----------|
| XGBoost Classifier (Tiefe 1) | 0.86(2) | 0.869(5) |
| Random Forrest | 0.87(1) | 0.89(1) |

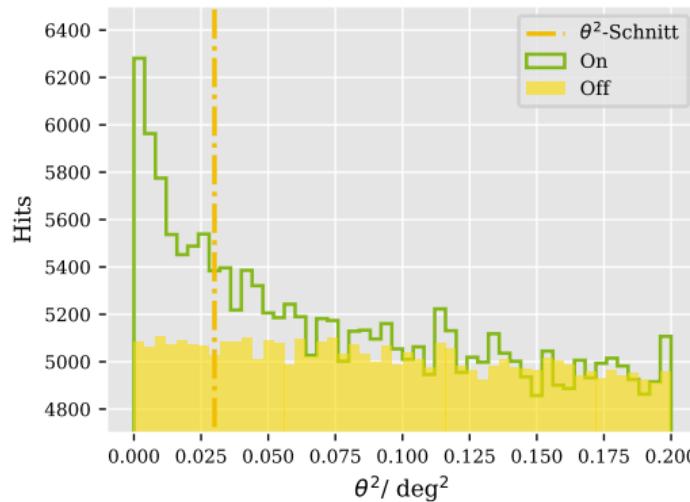
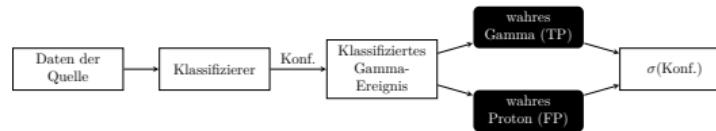


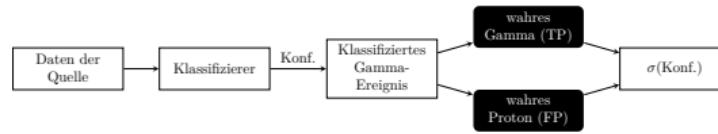


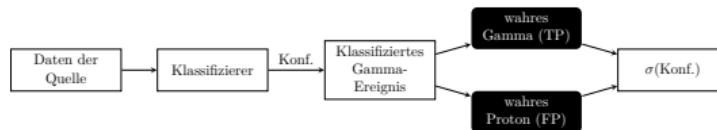
Signifikanz

$$S(N_{\text{on}}, N_{\text{off}}, S) = \sqrt{2} \left(N_{\text{on}} \ln \left[\frac{1+\alpha}{\alpha} \left(\frac{N_{\text{on}}}{N_{\text{on}} + N_{\text{off}}} \right) \right] + N_{\text{off}} \ln \left[(1+\alpha) \left(\frac{N_{\text{off}}}{N_{\text{on}} + N_{\text{off}}} \right) \right] \right)^{1/2}$$







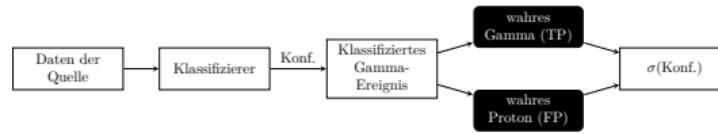


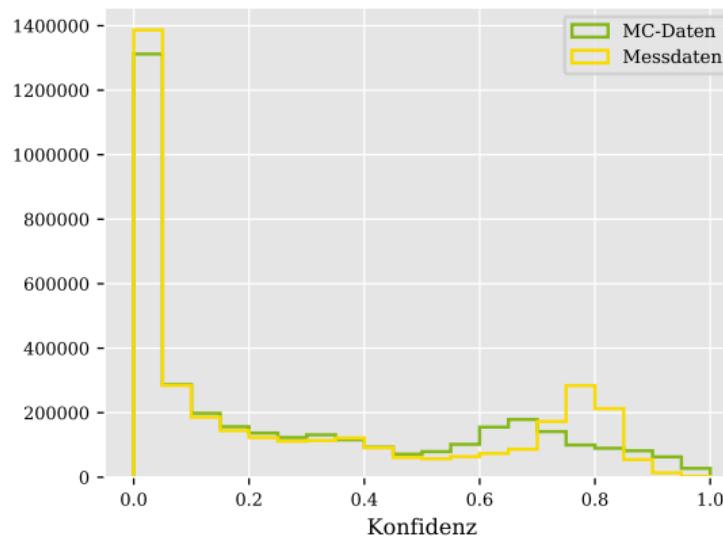
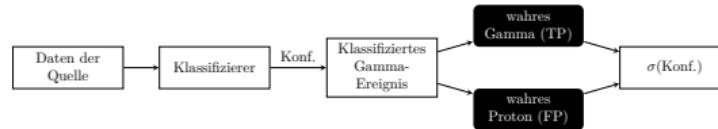
Random Forest

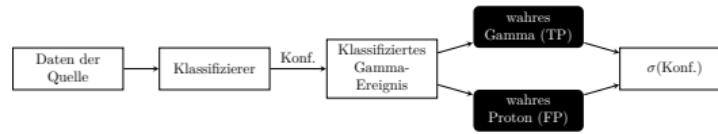
- komplexität
- viele Spezialisten auf ihrem Gebiet
- anfällig auf Training von Mismatche

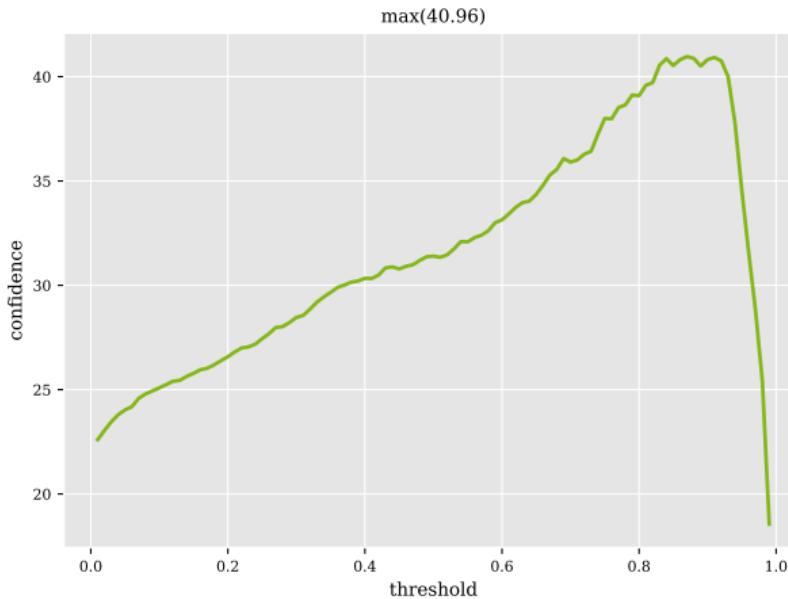
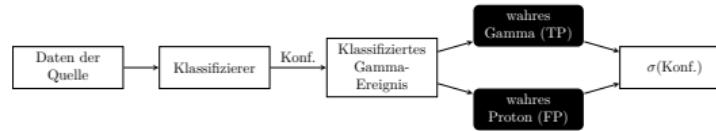
XGBoost Classifier

- geringe Komplexität
- resistent gegen Mismatche

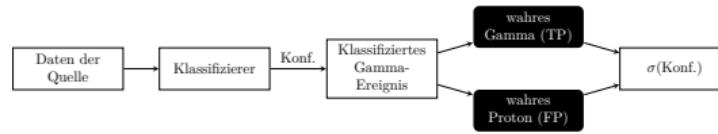


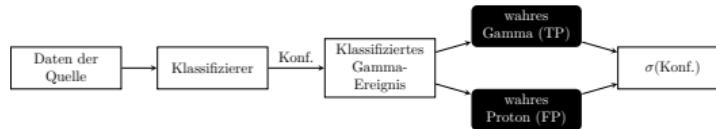






Validieren auf echten Daten





| | Krebsnebel | | Markarian 501 | |
|------------------------|---------------|---------------------|---------------|---------------------|
| | Random Forest | XGBoost (Tiefe = 1) | Random Forest | XGBoost (Tiefe = 1) |
| unklassifizierte Daten | | 21.4σ | | 17.1σ |
| MC-Proton | 41.9σ | 41.3σ | 35.5σ | 35.6σ |
| gemessene Proton | 32.9σ | 37.8σ | 23.6σ | 35.2σ |

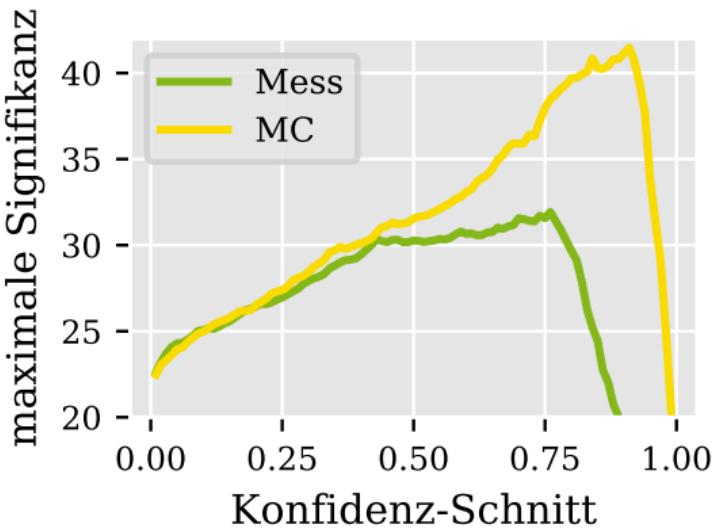
Thesen

- Training auf Montecarlo Mismatches zur Separation
 - komplexere Modelle weisen niedrigere Signifikanzen als einfache Modelle auf.

Thesen

- Training auf Montecarlo Mismatches zur Separation
 - komplexere Modelle weisen niedrigere Signifikanzen als einfache Modelle auf.
- Reduzierung schlecht simulierte Attribute
 - erhöhung der Signifikanz durch Reduzierung von Mismatches

Random Forest



XGBoost Classifier

