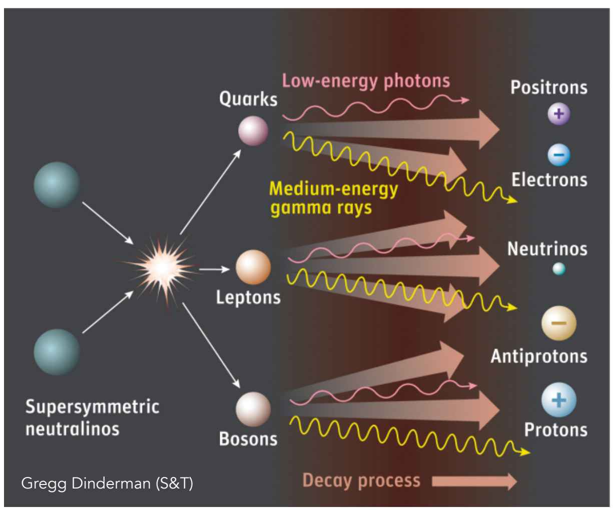
Enhancing the sensitivity to DM signatures in the VHE γ-ray band through machine learning

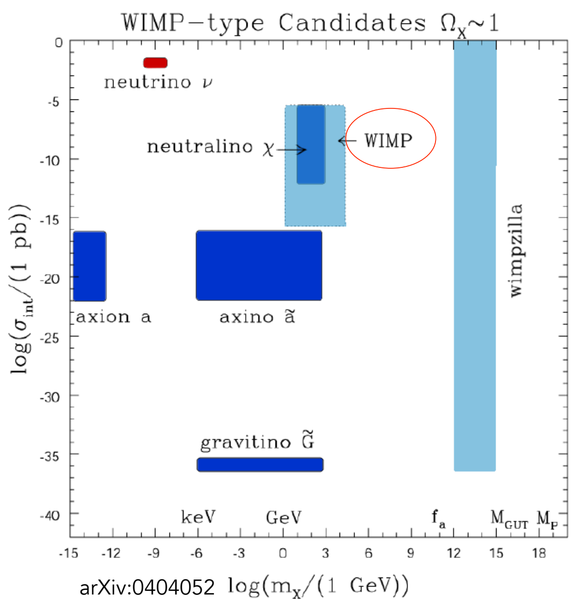
D. Nieto
on behalf of the Cherenkov Telescope Array Consortium

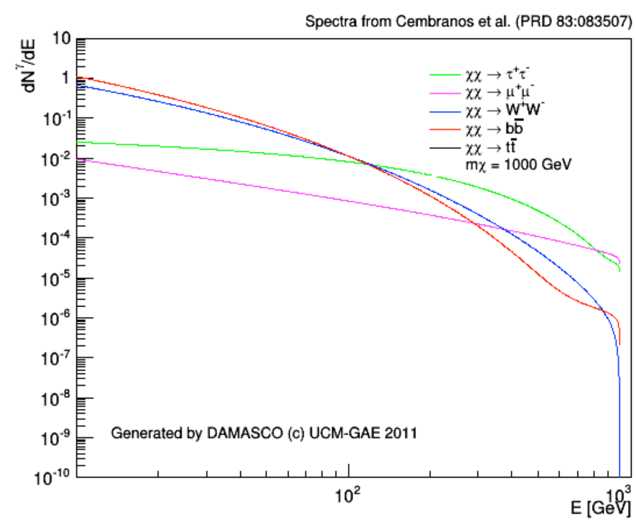
Statistical Challenges in the Search for Dark Matter
February 2018
Banff, Alberta

- Dark matter searches in the (VHE) γ-ray band
- Imaging atmospheric Cherenkov technique
- Machine learning & current-generation IACTs
- Machine learning & next-generation IACTs

- Basis: Detection of DM annihilation or decay products (SM particles)
- In most cases, entangled with CR and subdominant
- WIMPs with masses in the ~100 GeV range are good DM particle candidates
- Photons are privileged messengers
    - No deflection by B-fields, trace back to source
    - Observation of astrophysical targets
    - Characteristic spectral shape: identification







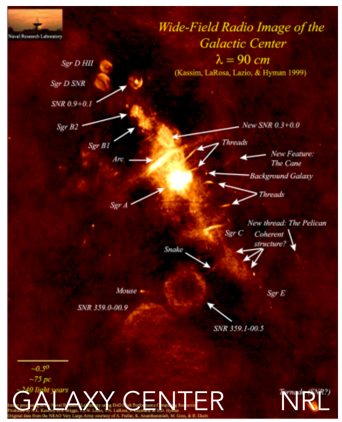
Expected spectrum from annihilating DM

$$\frac{d\Phi}{dE} = \boxed{J(\Delta\Omega)} \times \boxed{\frac{d\Phi^{PP}}{dE}} = \boxed{\int_{l.o.s,V} \rho^2_{DM}(l)\, d\Omega\, dl} \times \boxed{\frac{1}{4\pi} \frac{\langle \sigma_{ann} v \rangle}{2 m^2_{DM}} \sum_i B_i \frac{dN^\gamma_i}{dE}}$$

Key concepts: ρ_{DM}, distance, background

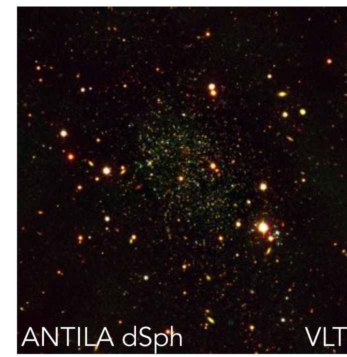
Galactic Center & Halo

- High flux
- Background Issues



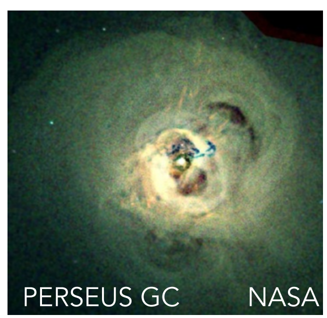
Dwarf Galaxies

- Large M/L
- No background
- Low flux



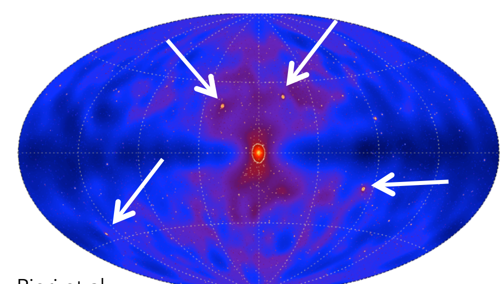
Galaxy Clusters

- Huge DM content
- Large distance
- High background



Unassociated HE Sources:

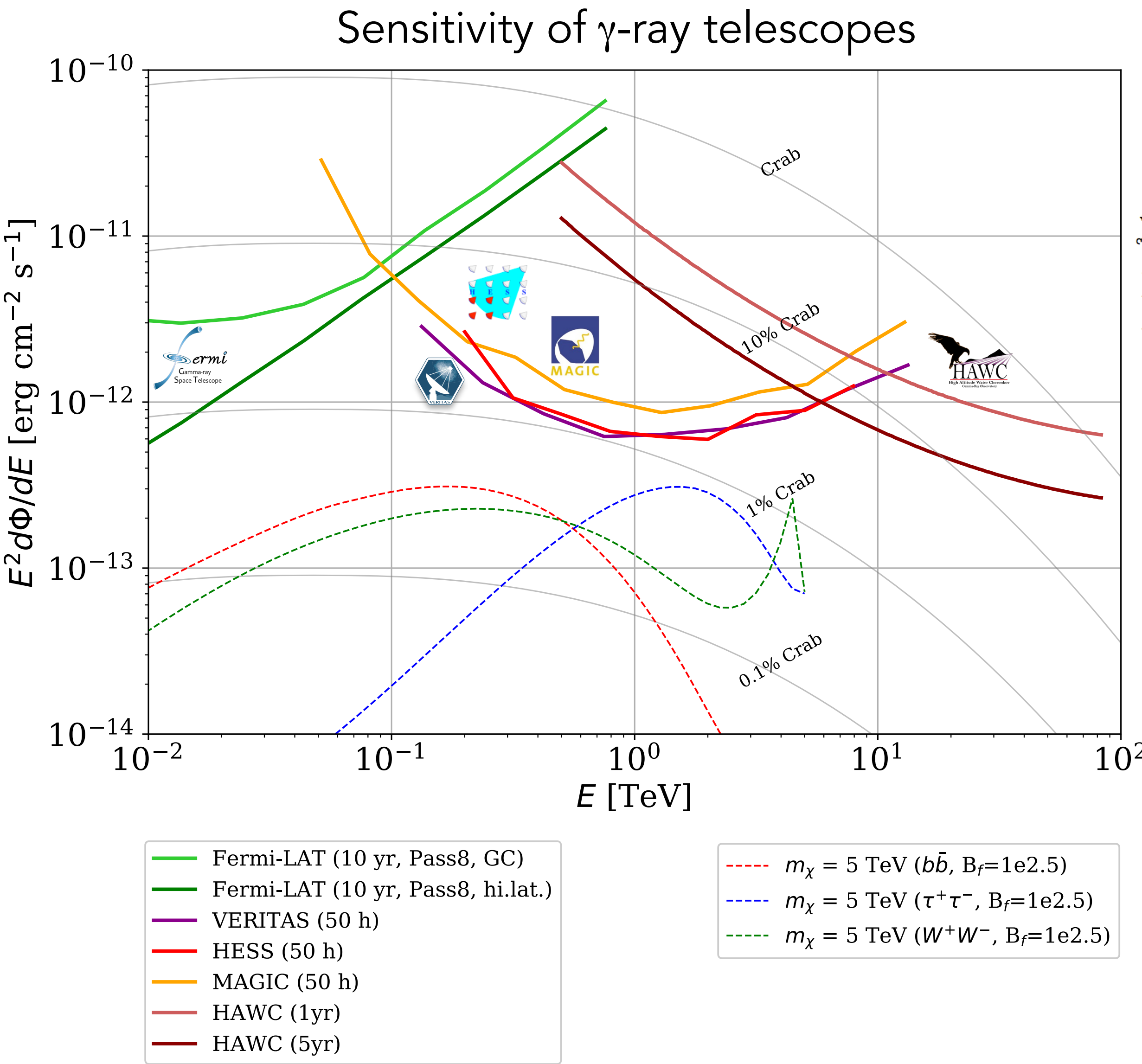
- DM Subhalos?

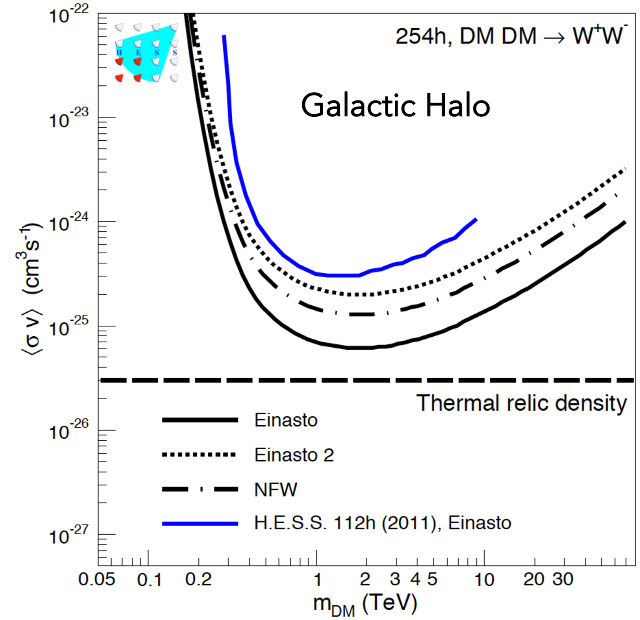


Pieri et al.
PRD 83:0235, 2008

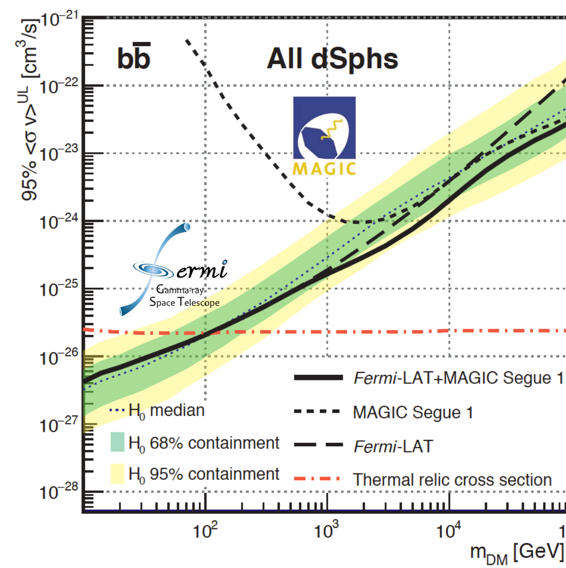
$\chi\chi \rightarrow b\bar{b},\ m_\chi = 40$ GeV

Indirect Dark Matter Searches in the γ-Ray Band



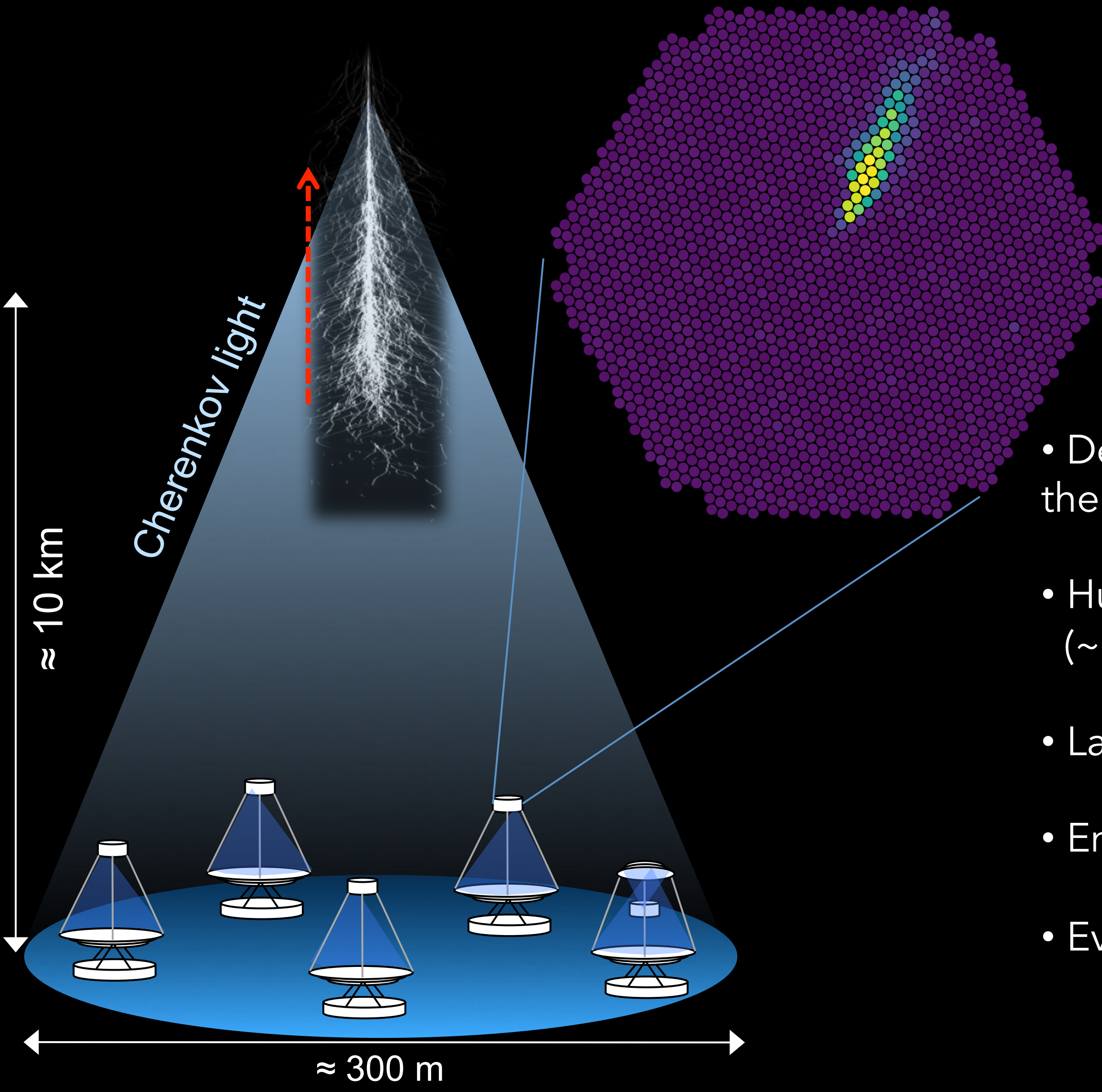


Abdallah et al., PRL 117, 111301 (2016)



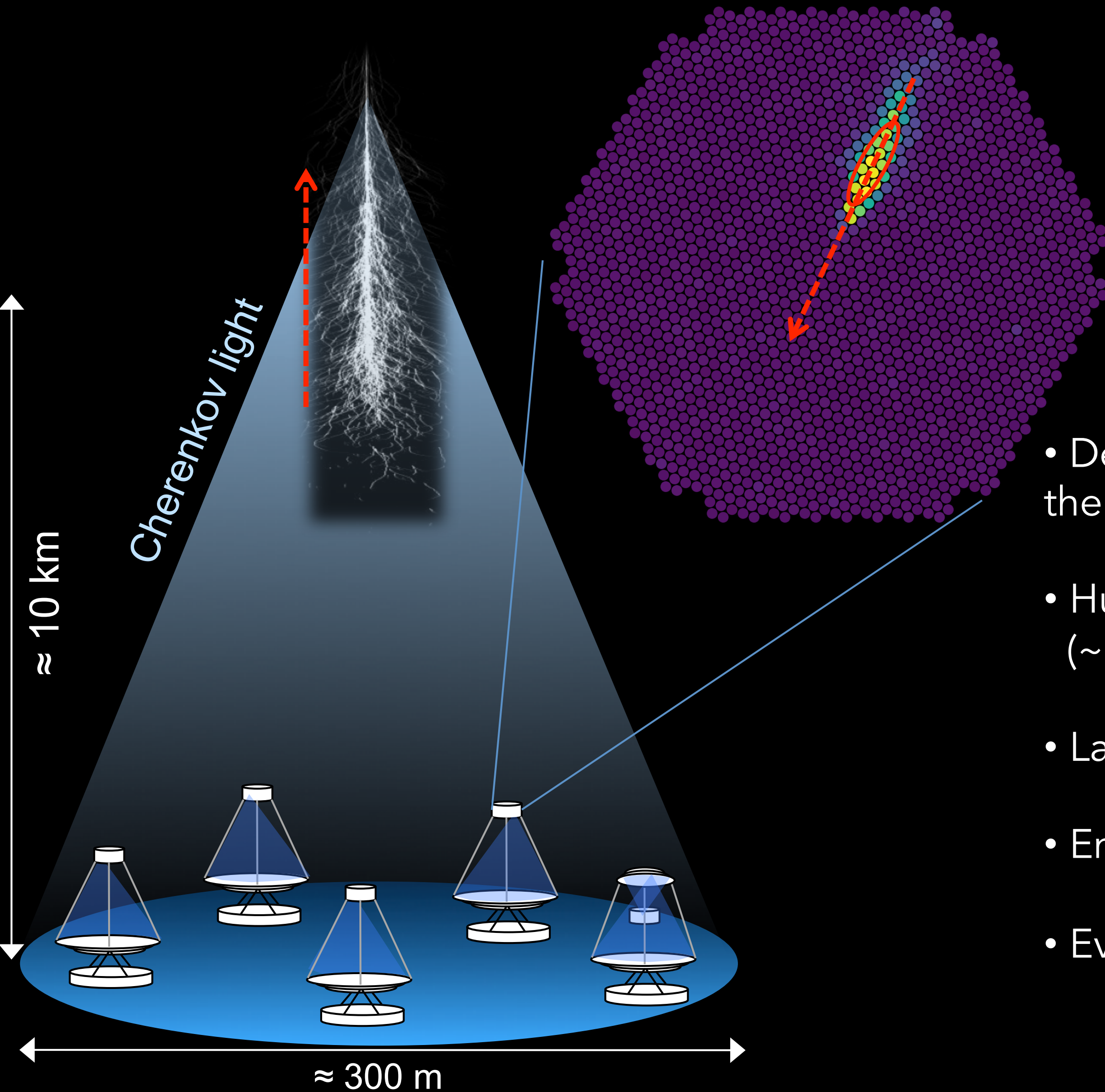
Ahnen et al., JCAP 02 (2016) 039

Imaging Atmospheric Cherenkov Technique



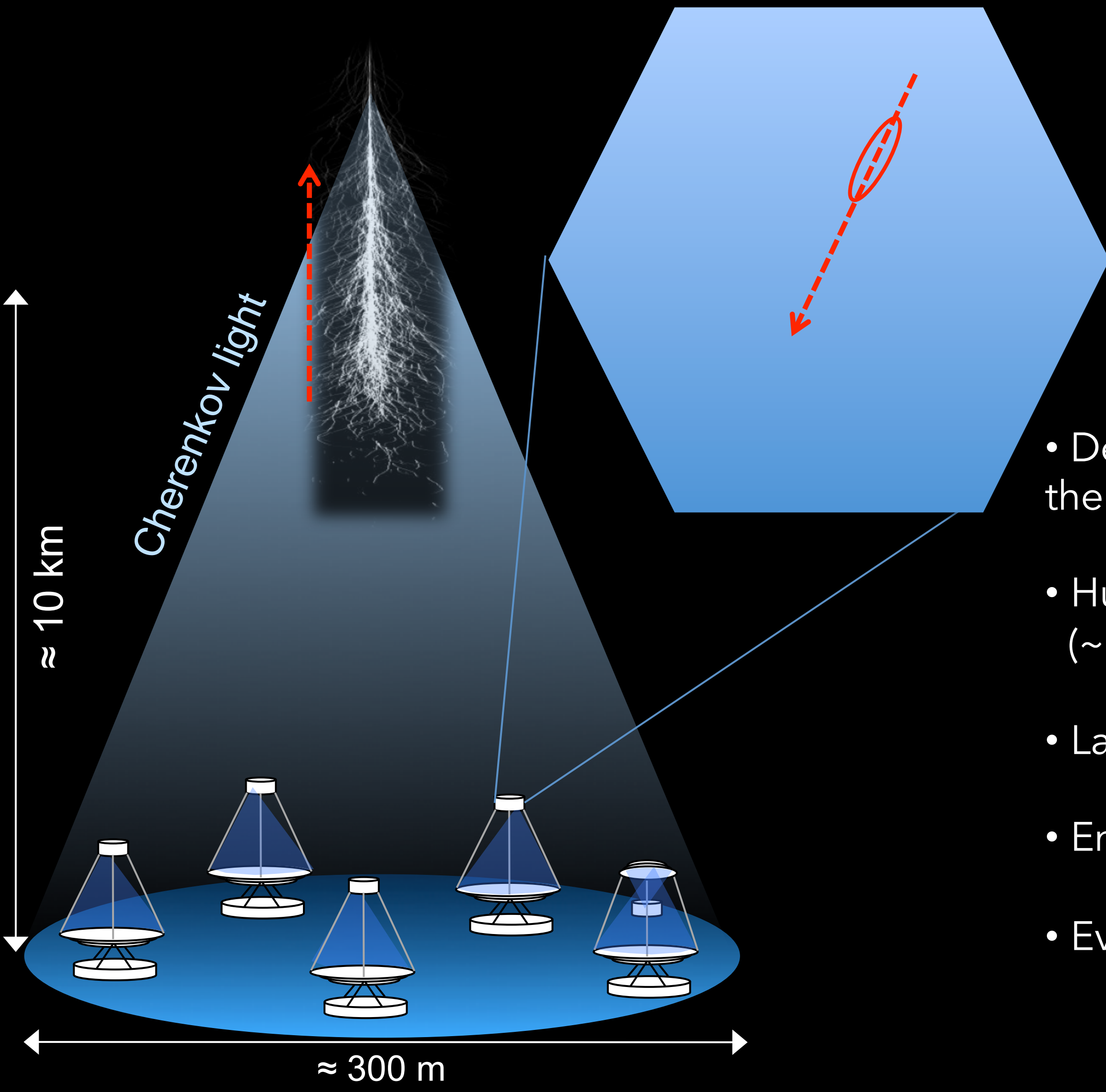
- Detection of extended air showers using the atmosphere as a calorimeter
- Huge γ–ray collection area ($\sim 10^5$ m^2)
- Large background from charged CR
- Energy window: tens GeV - tens TeV
- Event reconstruction from image:
  - Type of primary event
  - Primary energy estimation
  - Primary arrival direction

Imaging Atmospheric Cherenkov Technique

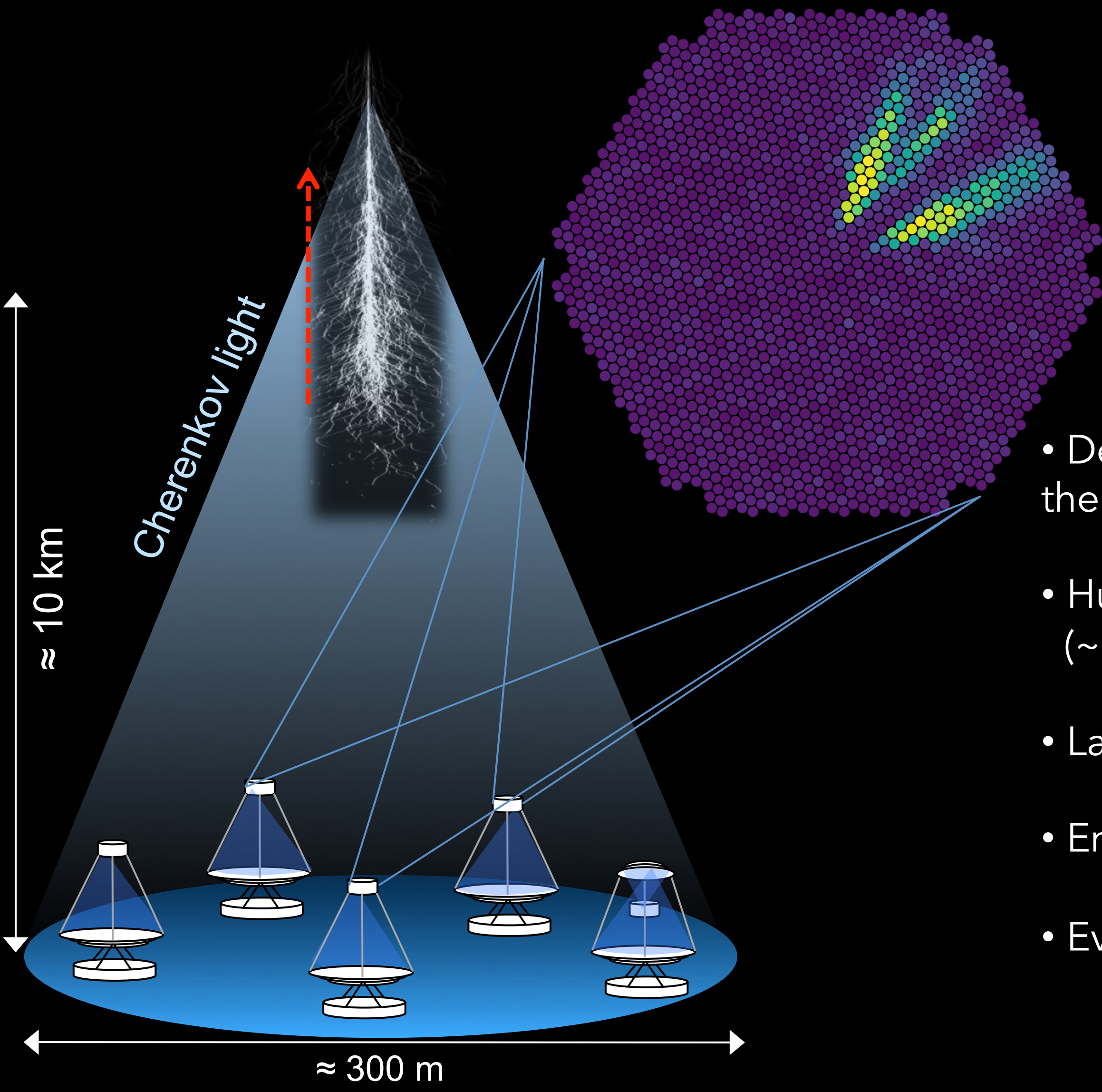


- Detection of extended air showers using the atmosphere as a calorimeter
- Huge γ–ray collection area ($\sim 10^5$ m^2)
- Large background from charged CR
- Energy window: tens GeV - tens TeV
- Event reconstruction from image:
  - Type of primary event
  - Primary energy estimation
  - Primary arrival direction

Imaging Atmospheric Cherenkov Technique

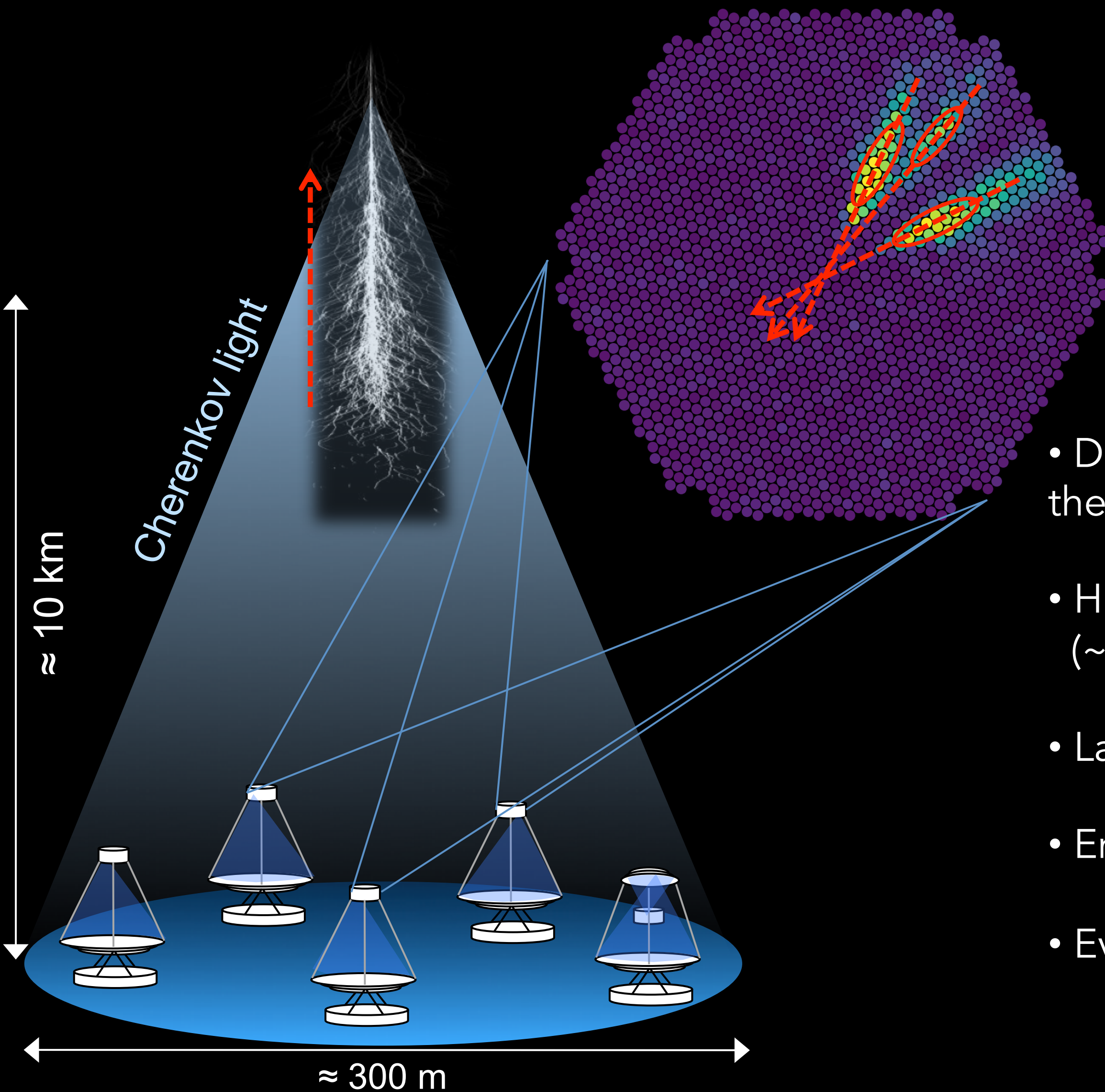


- Detection of extended air showers using the atmosphere as a calorimeter
- Huge γ–ray collection area ($\sim 10^5$ m^2)
- Large background from charged CR
- Energy window: tens GeV - tens TeV
- Event reconstruction from image:
  - Type of primary event
  - Primary energy estimation
  - Primary arrival direction

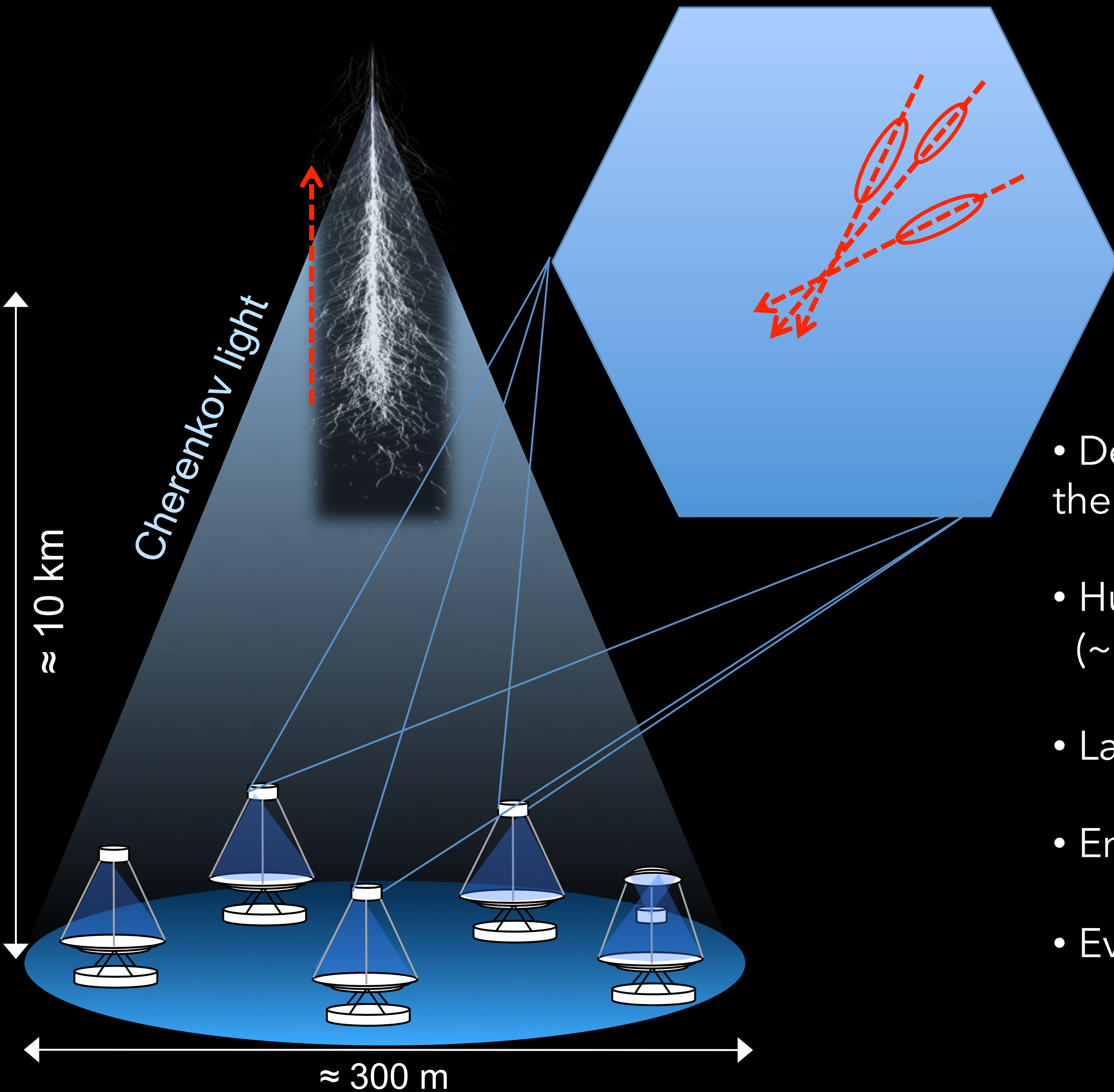


- Detection of extended air showers using the atmosphere as a calorimeter
- Huge γ–ray collection area ($\sim 10^5$ m^2)
- Large background from charged CR
- Energy window: tens GeV - tens TeV
- Event reconstruction from image:
  - Type of primary event
  - Primary energy estimation
  - Primary arrival direction

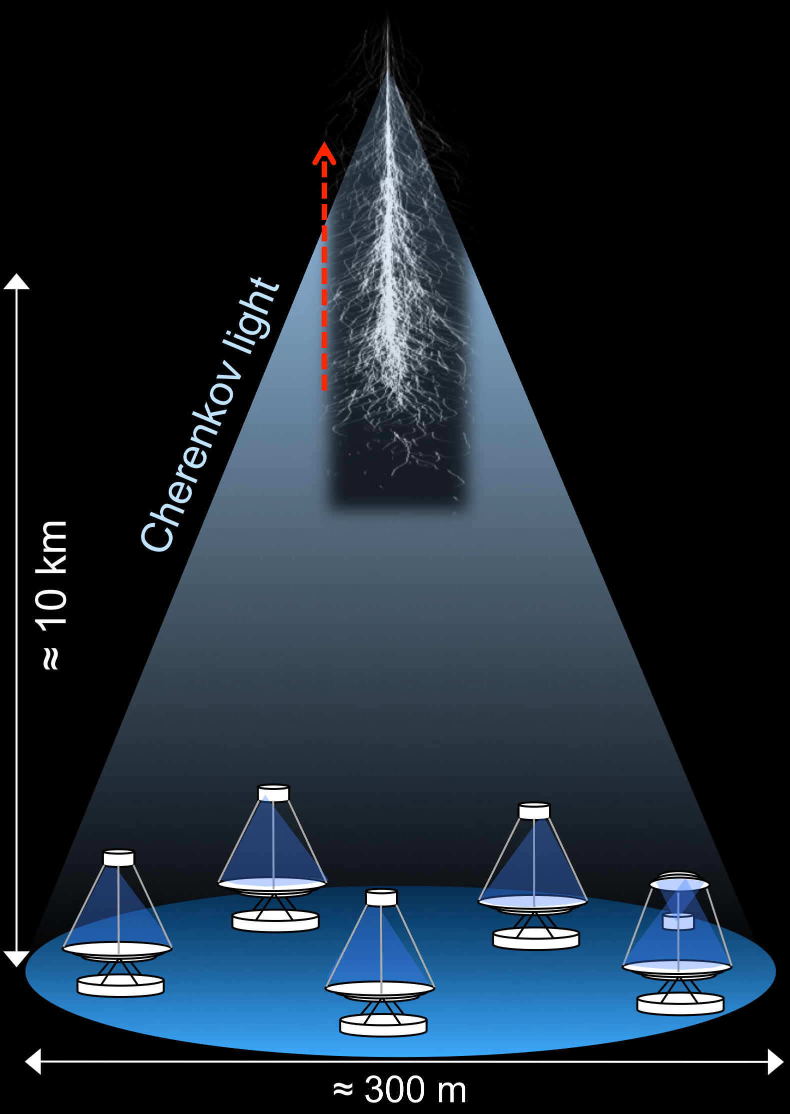
Imaging Atmospheric Cherenkov Technique



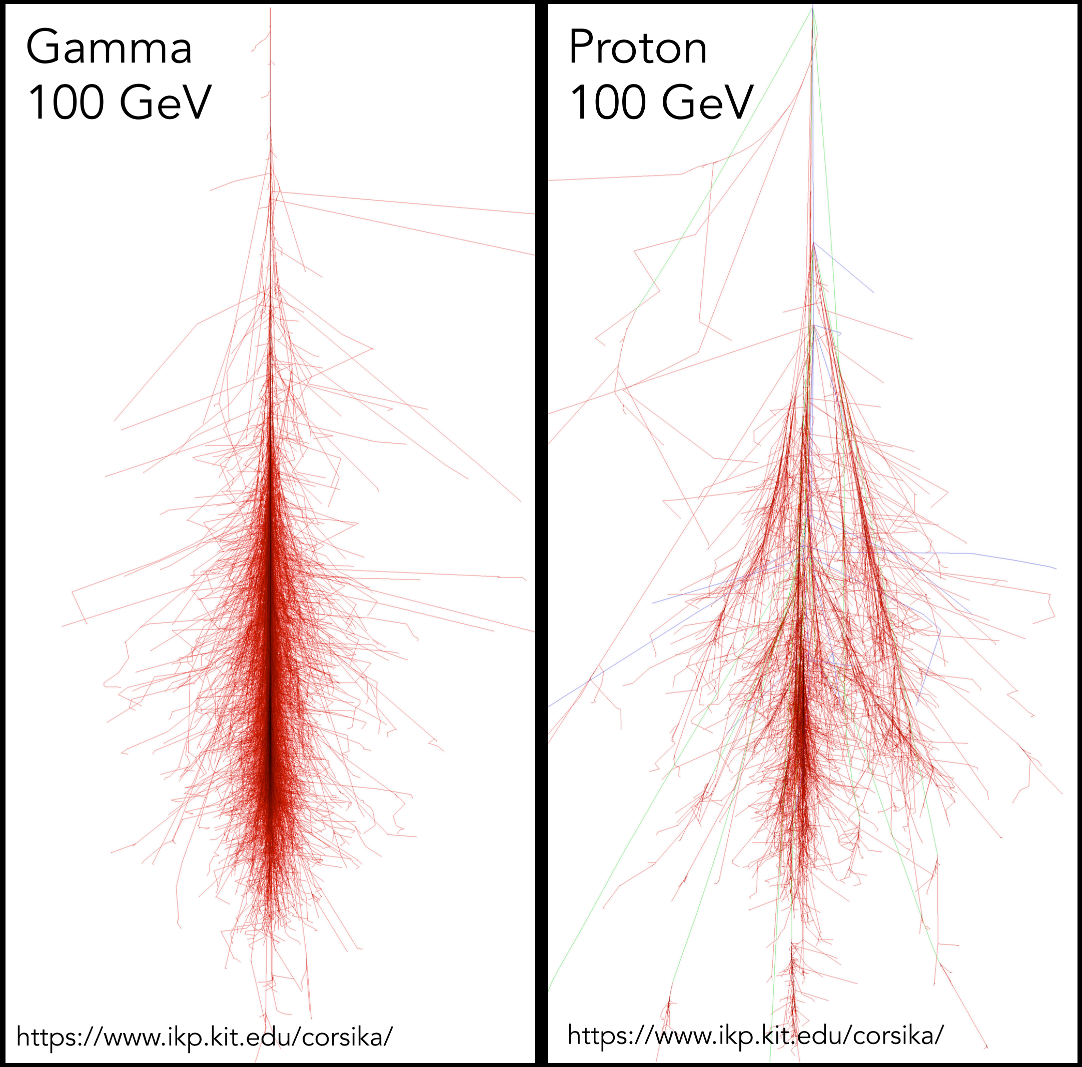
- Detection of extended air showers using the atmosphere as a calorimeter
- Huge γ–ray collection area ($\sim 10^5$ m^2)
- Large background from charged CR
- Energy window: tens GeV - tens TeV
- Event reconstruction from image:
  - Type of primary event
  - Primary energy estimation
  - Primary arrival direction



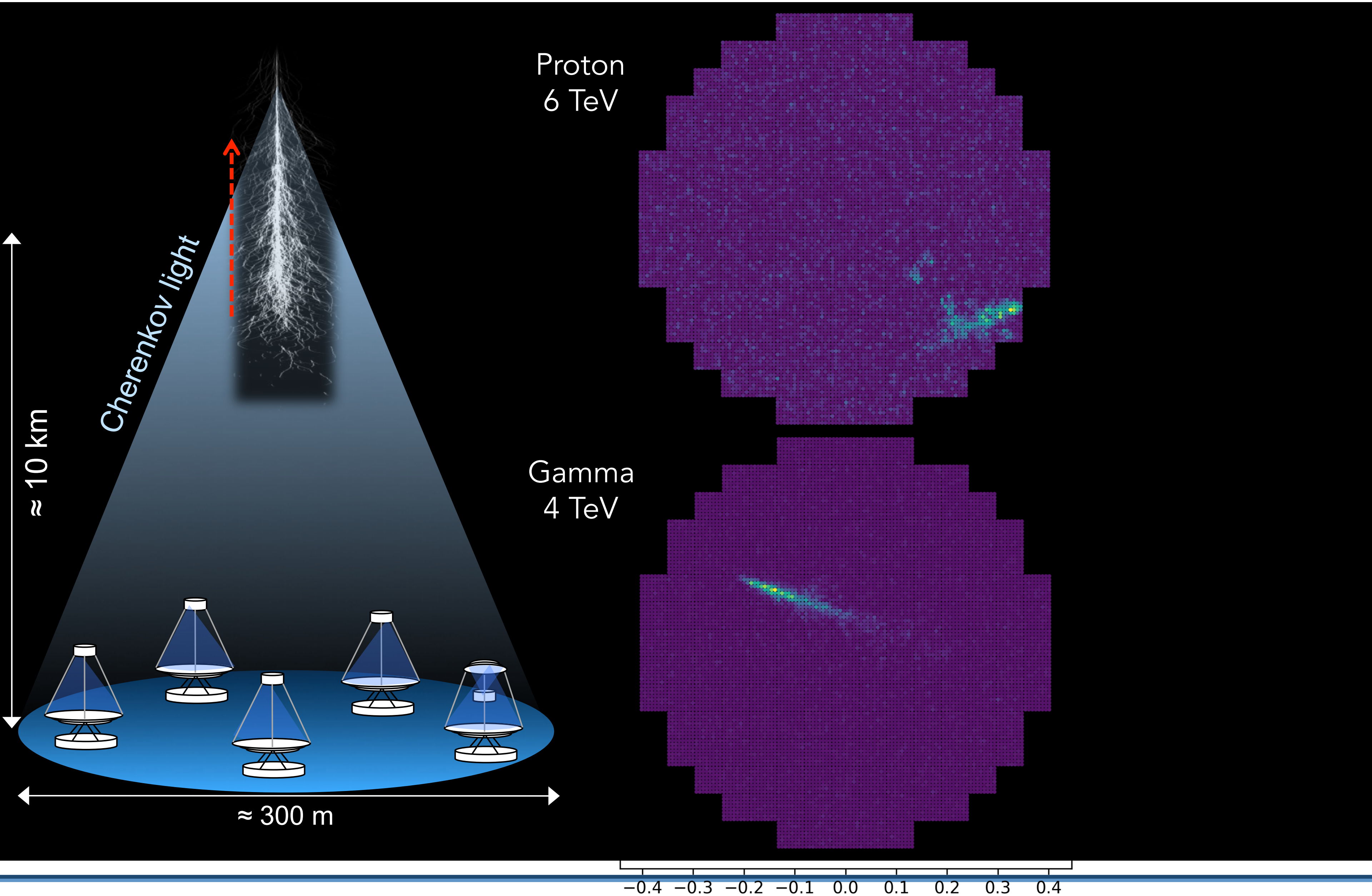
- Detection of extended air showers using the atmosphere as a calorimeter
- Huge γ–ray collection area (~10^5 m^2)
- Large background from charged CR
- Energy window: tens GeV - tens TeV
- Event reconstruction from image:
  - Type of primary event
  - Primary energy estimation
  - Primary arrival direction



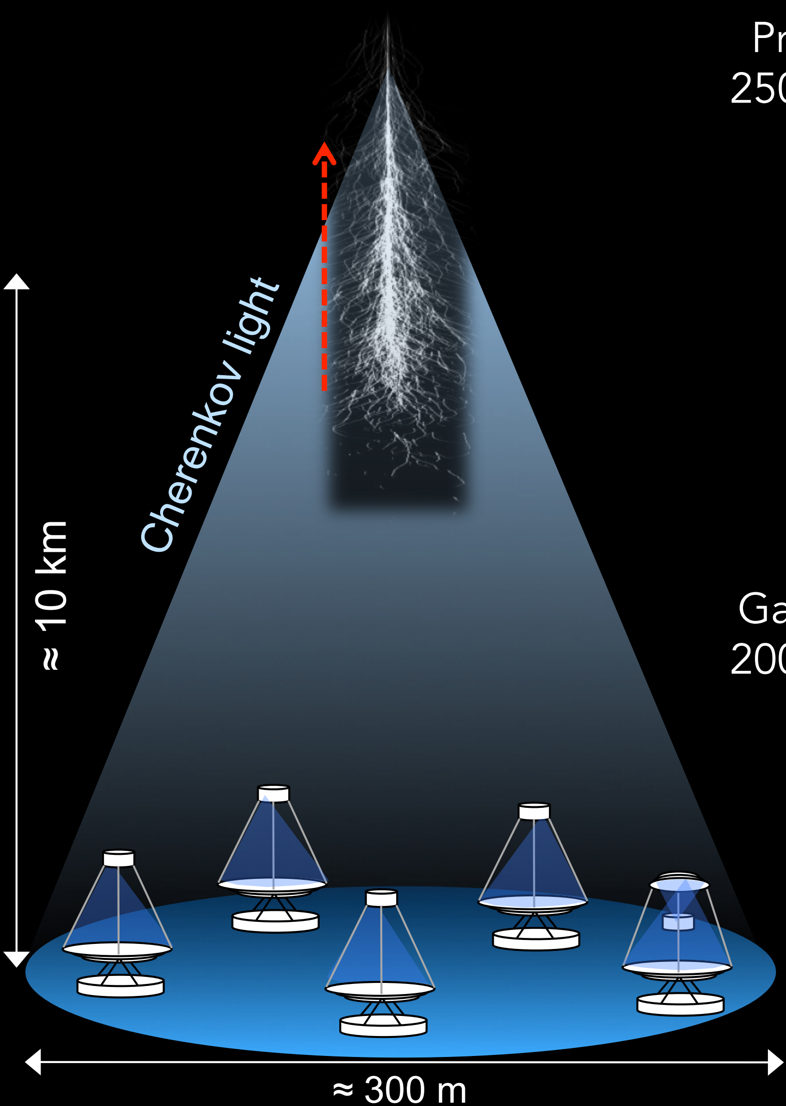
Cherenkov light
≈ 10 km
≈ 300 m

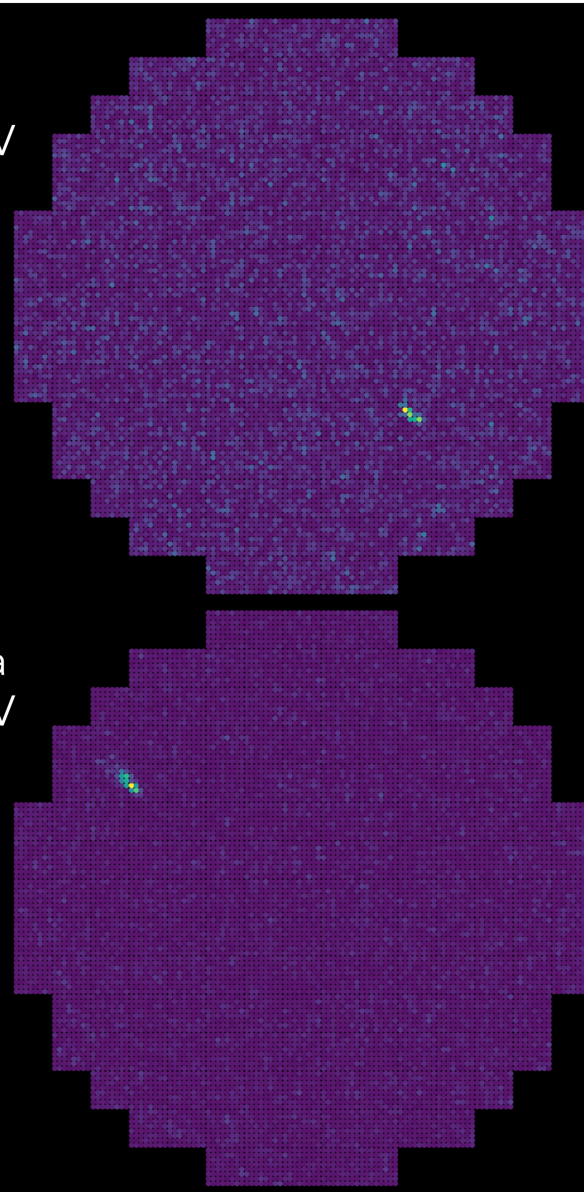


Gamma
100 GeV
https://www.ikp.kit.edu/corsika/
Proton
100 GeV
https://www.ikp.kit.edu/corsika/

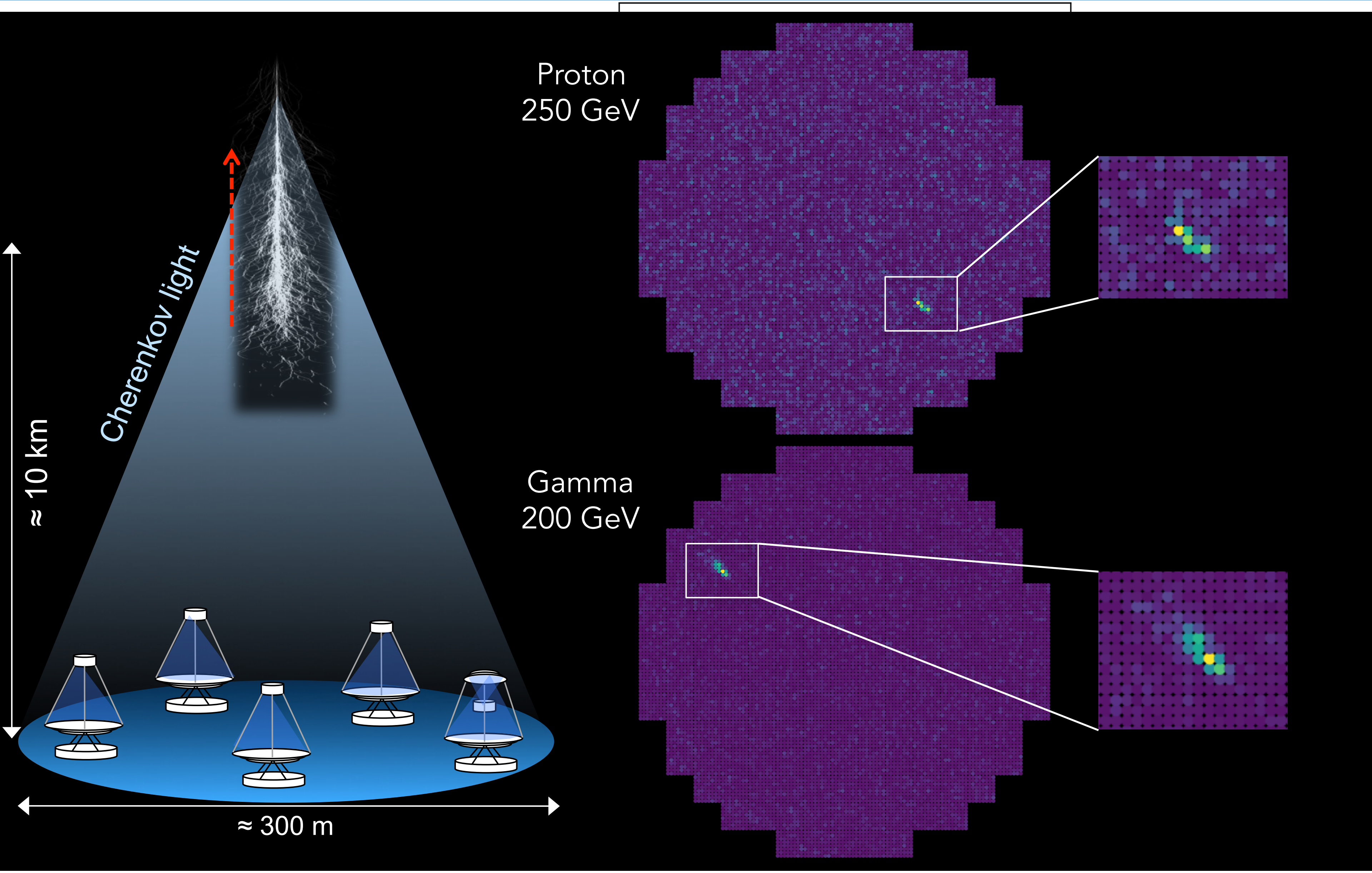

Cherenkov light
≈ 10 km
≈ 300 m
Proton
6 TeV
Gamma
4 TeV
−0.4 −0.3 −0.2 −0.1 0.0 0.1 0.2 0.3 0.4

Imaging Atmospheric Cherenkov Technique

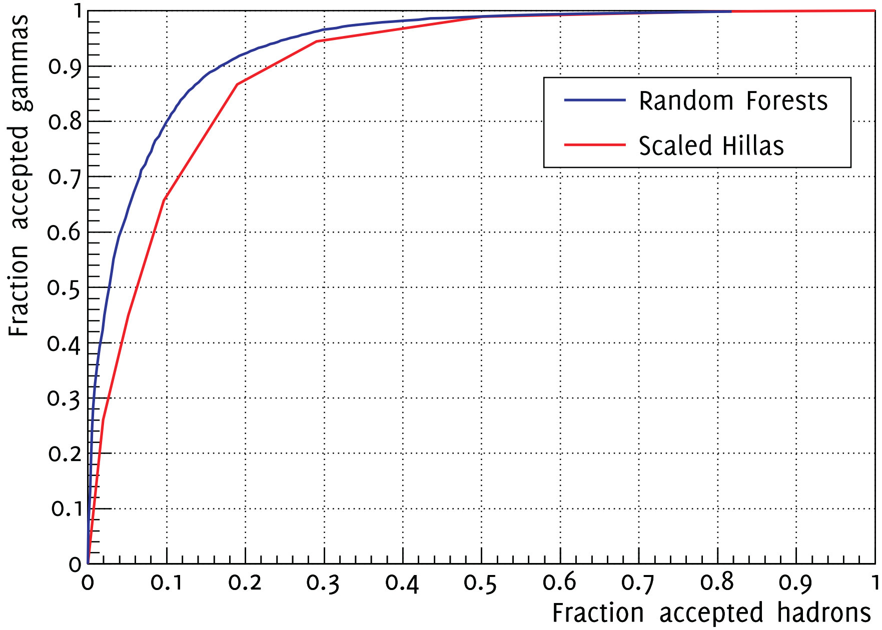




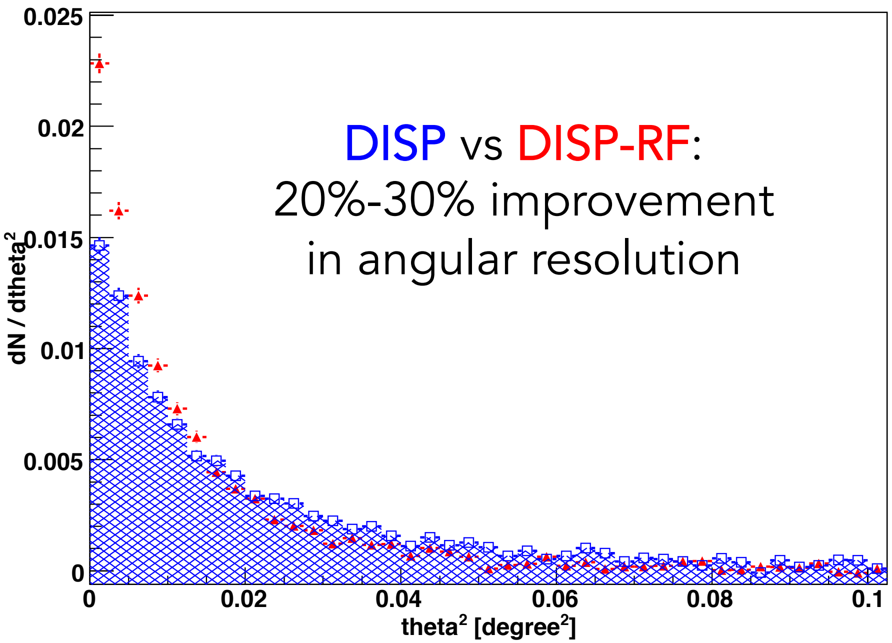
Imaging Atmospheric Cherenkov Technique



- ML method: Random Forest (RF)
- Applied to: background rejection, arrival direction



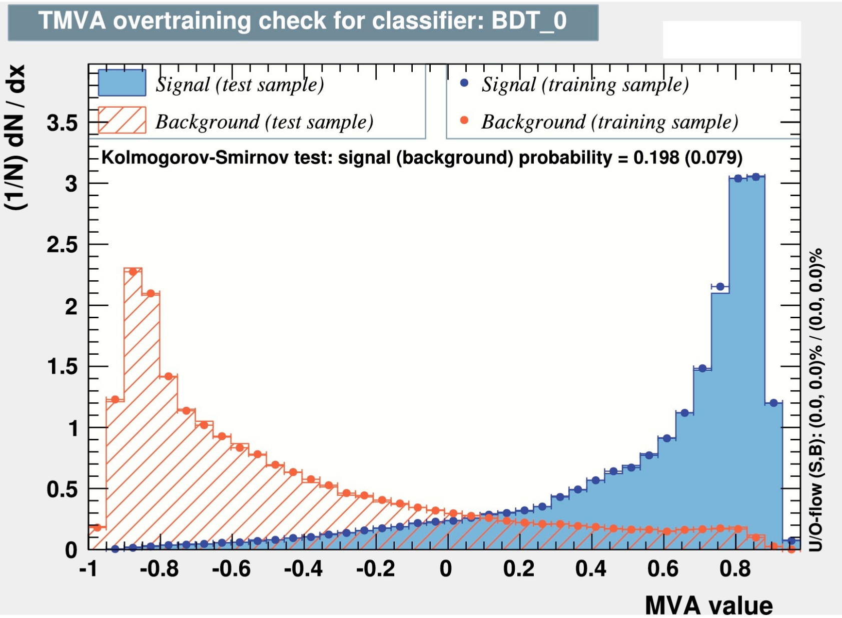
Albert et al., NIM-A 588:424-432 (2008)

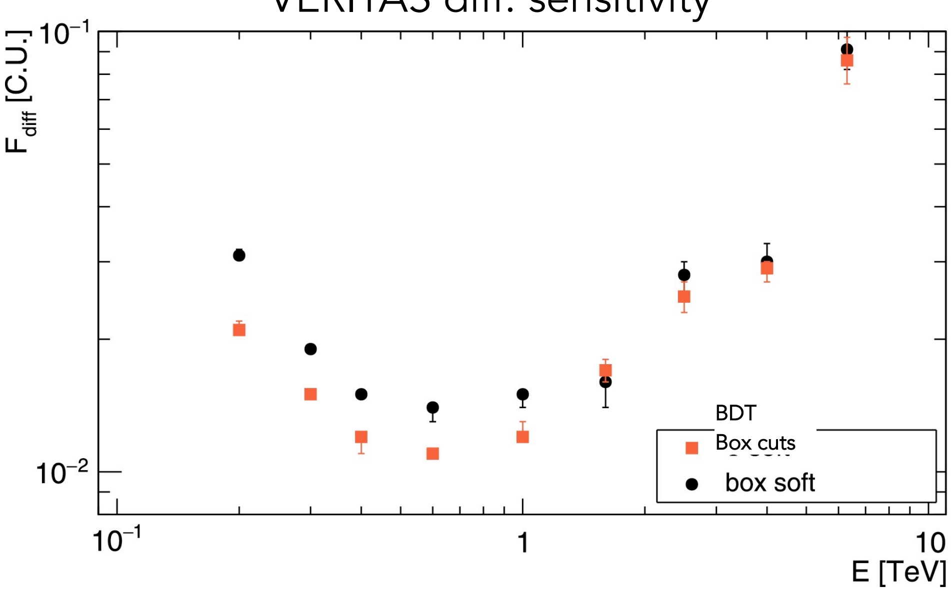


Aleksic et al., A&A 524 A77 (2010)

Machine learning & current generation IACT

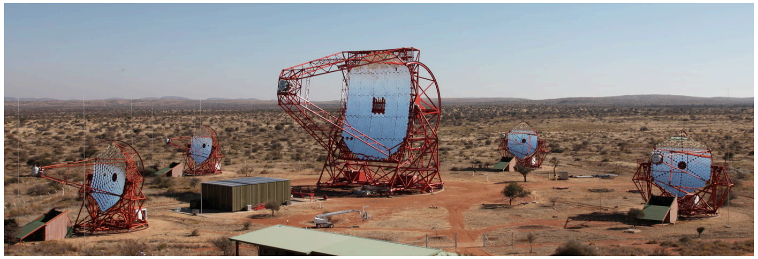
- ML method: Boosted Decision Trees (BDT)
- Applied to: background rejection



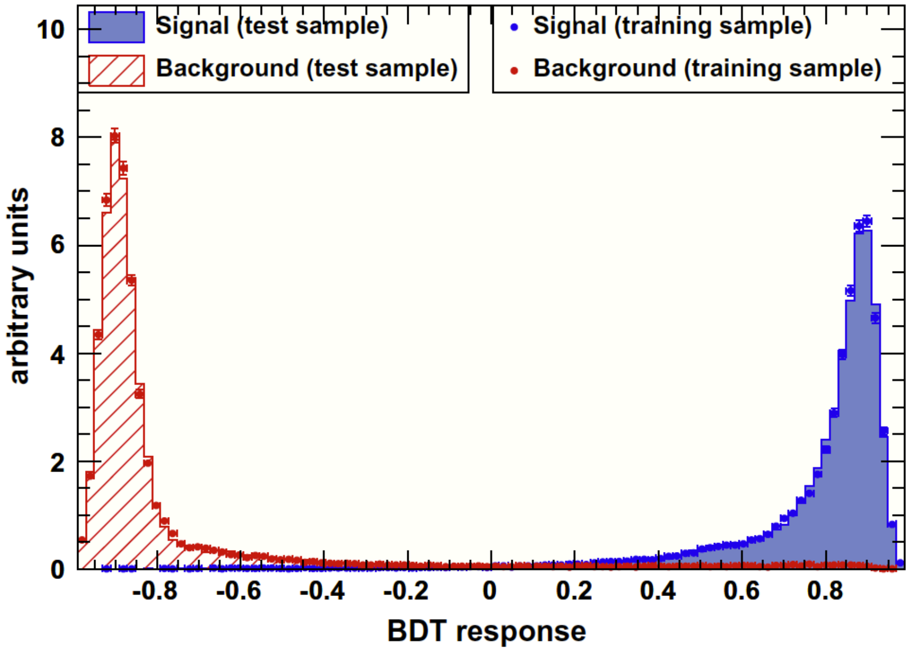


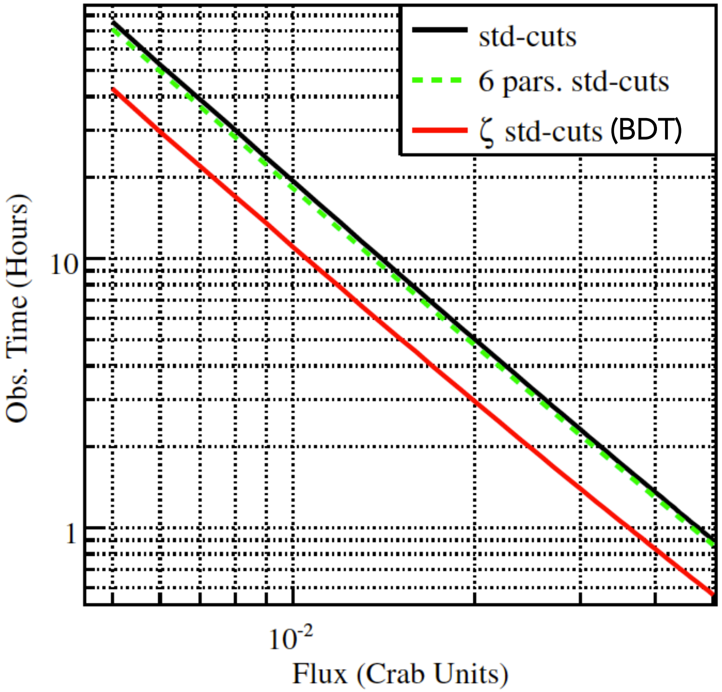
Krause et al., APP V89 P1-9 (2017)

H.E.S.S.



- ML method: Boosted Decision Trees (BDT)
- Applied to: background rejection





Becherini et al., APP V34-12 P858-870 (2011)

Ohm et al., APP V31-5 P383-391 (2009)

(Results for H.E.S.S. I only)

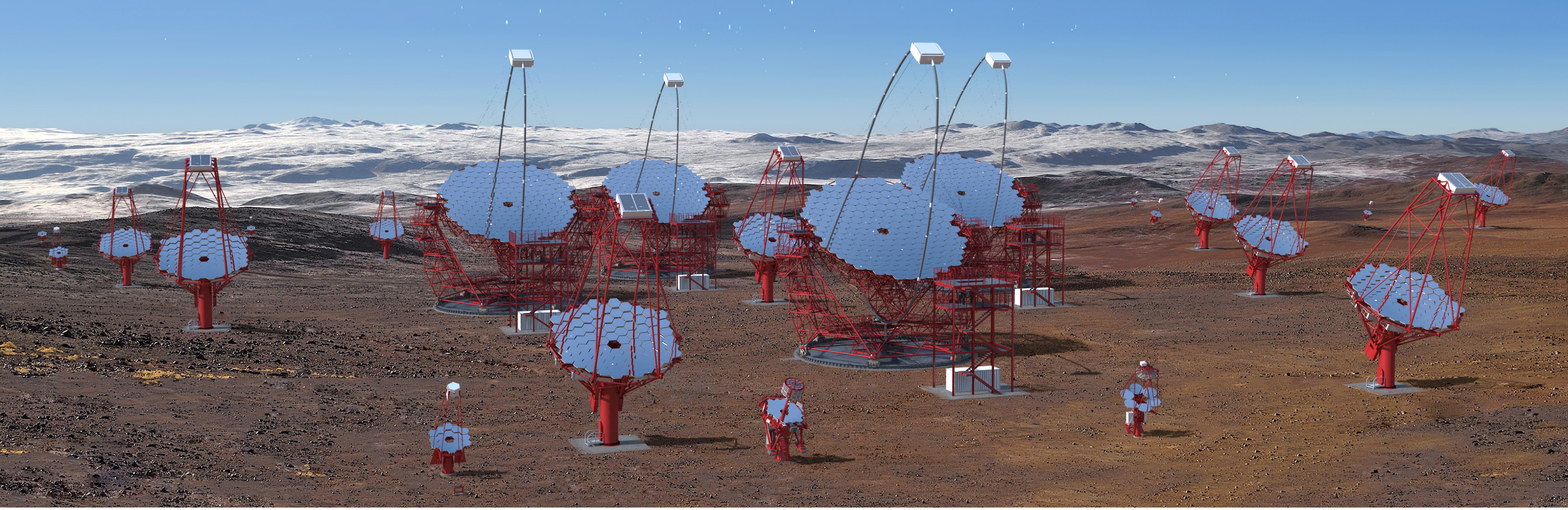
The Cherenkov Telescope Array

- 5-20 fold better sensitivity w.r.t. current IACTs
- 4 decades of energy coverage: tens of GeV to >100 TeV
- Improved angular and energy resolution
- Two arrays (North/South)

Low-energy range:
23 m ø
Parabolic reflector
4° - 5° FoV
Energy threshold ~tens GeV

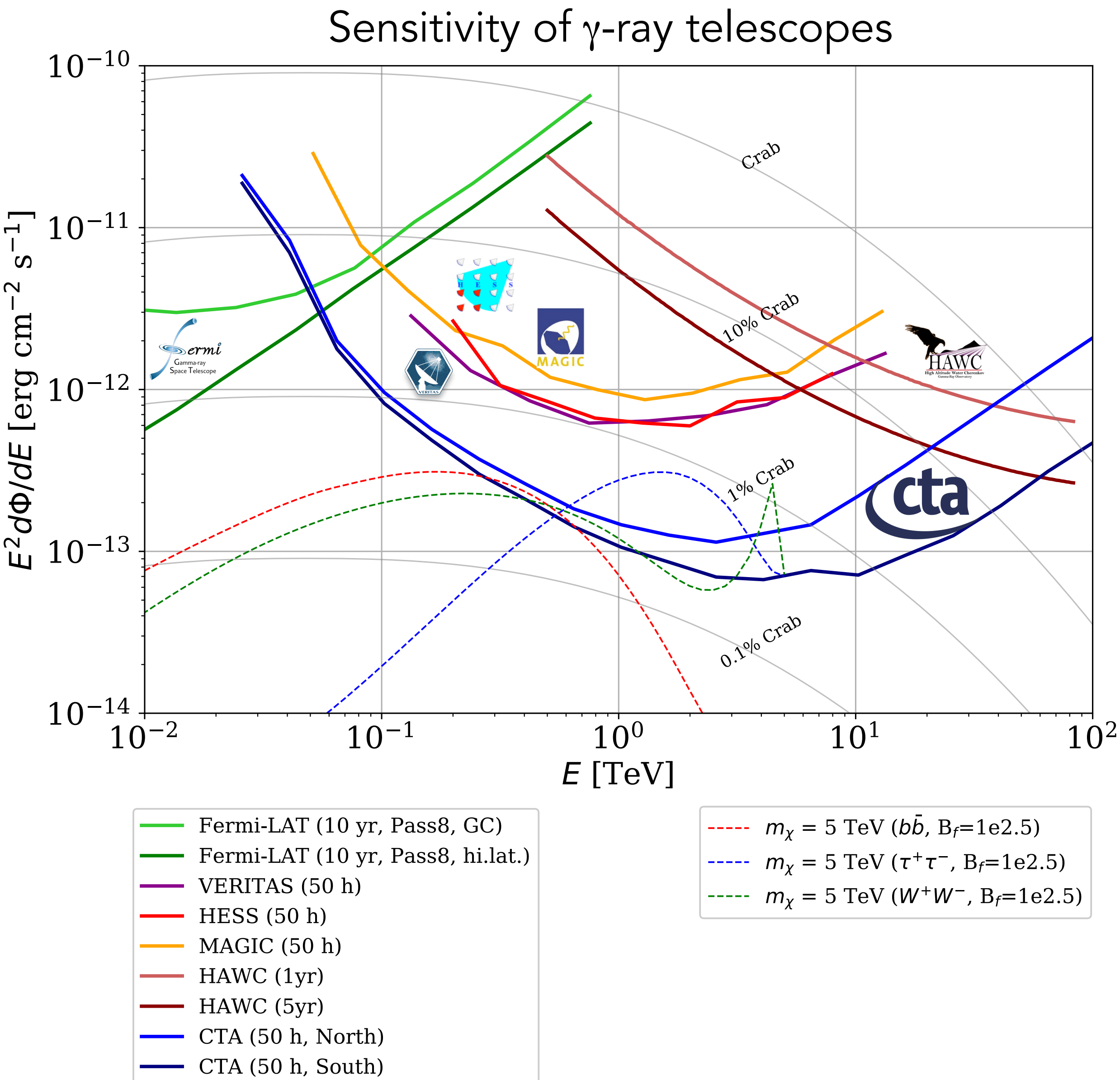
Mid energy-range:
12 m ø Davies-Cotton reflector
9 m ø Schwarzschild-Couder reflector
7° - 8° FoV
mCrab sensitivity in the
100 GeV – 10 TeV range

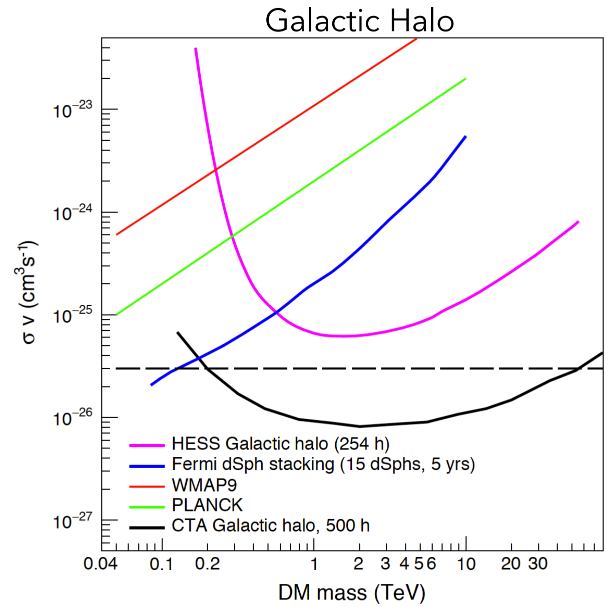
High-energy range:
4 m ø Davies-Cotton reflector
4 m ø Schwarzschild-Couder reflector
9 - 10° FoV
Several km^2 area at
multi-TeV energies

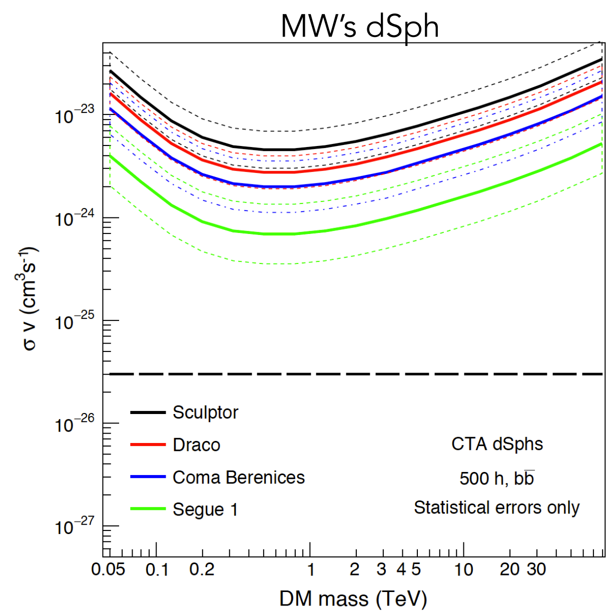


www.cta-observatory.org/ *Science with CTA*, arXiv:1709.07997

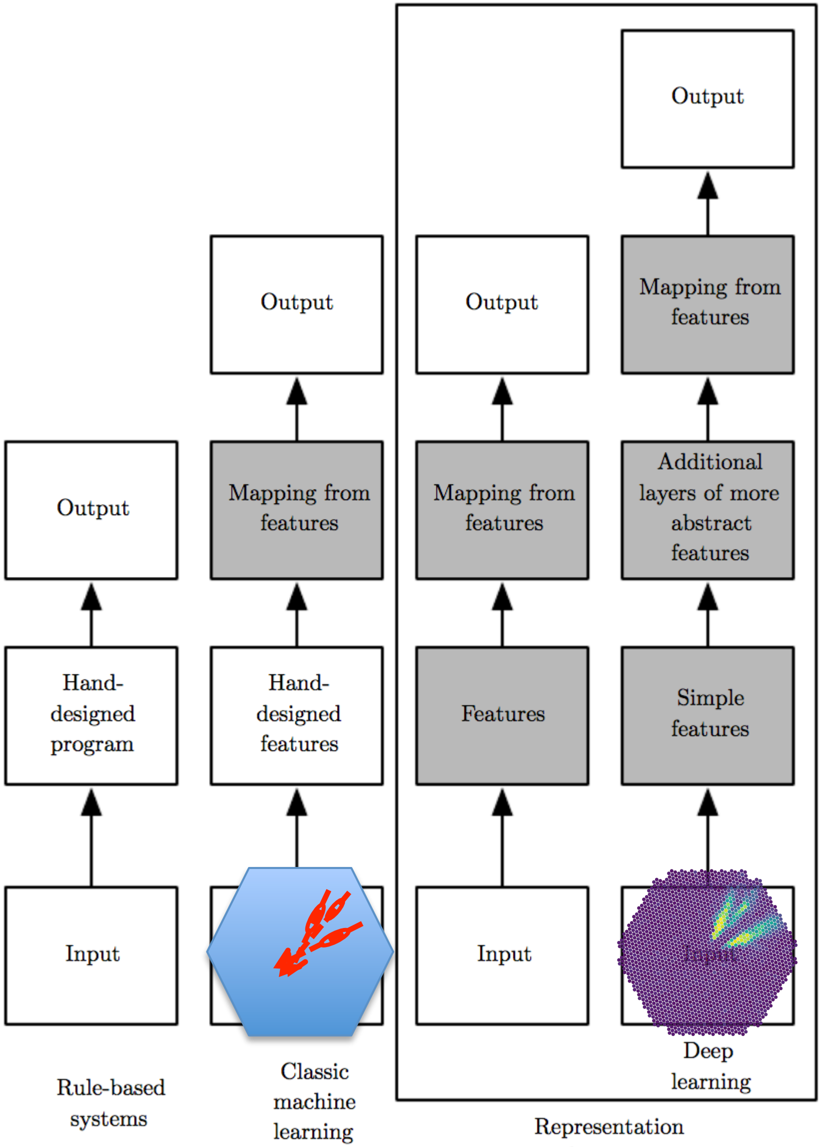
Dark matter searches with CTA



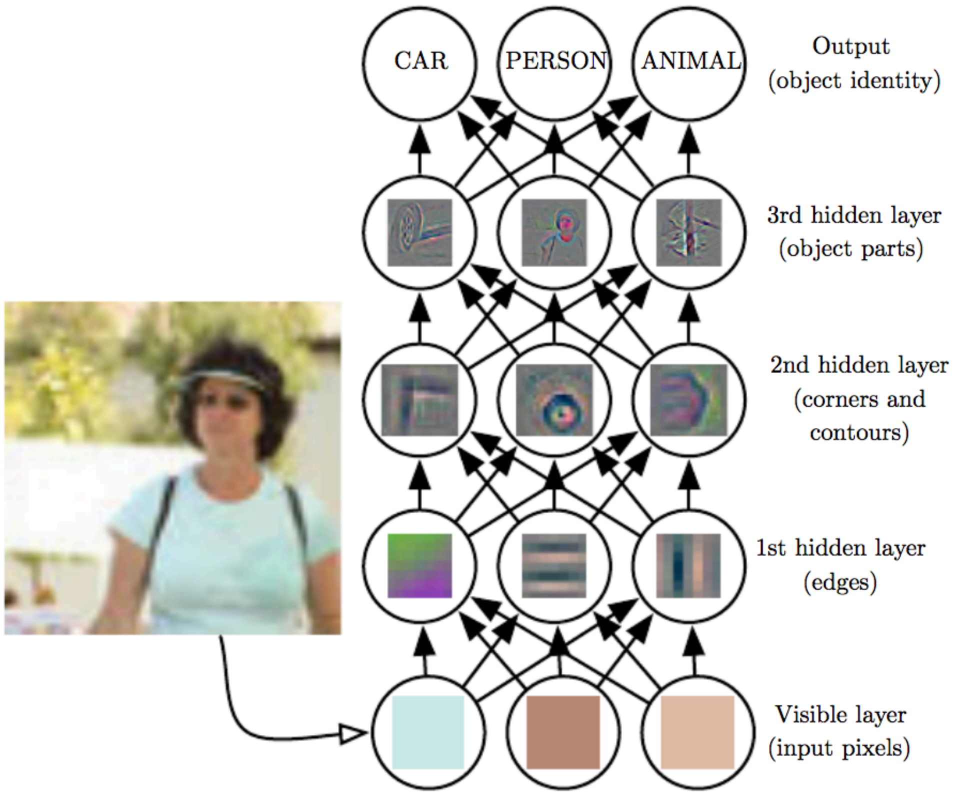




Science with CTA, arXiv:1709.07997

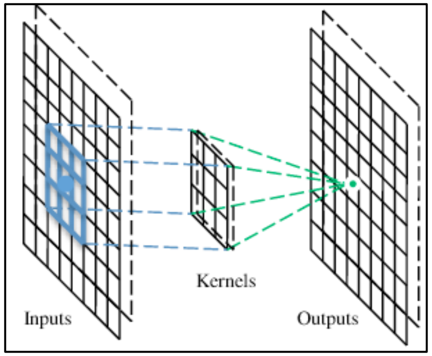


E.g.: RF & BDT



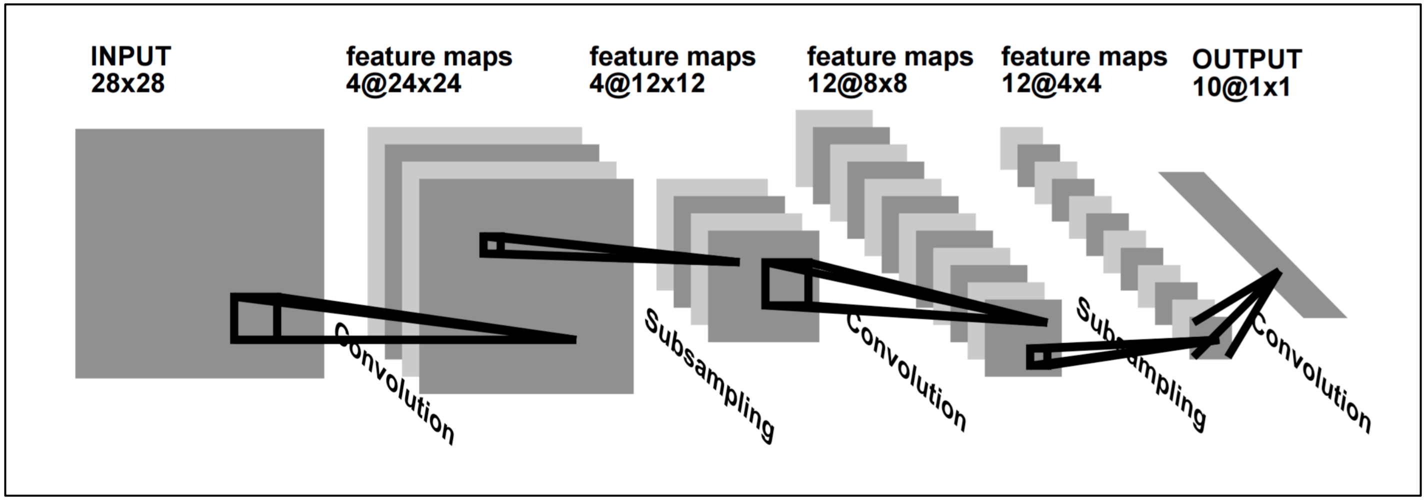
Deep Learning, Goodfellow et al.

Convolution



Guo et al.

Convolutional Neural Network (CNN)



LeCunn et al.

- DL capable of ***extracting*** and mapping image features automatically with unprecedented classification accuracy. Hyper-active CS research field constantly improving
- Many HEP/Astro experiments already exploring/utilizing the technique (LIGO, LHC, MicroBooNe, NOVa, etc…)

Method:

- Use deep learning to reconstruct CTA events from non-parameterized images
  - Performance enhancement -> better sensitivity to DM

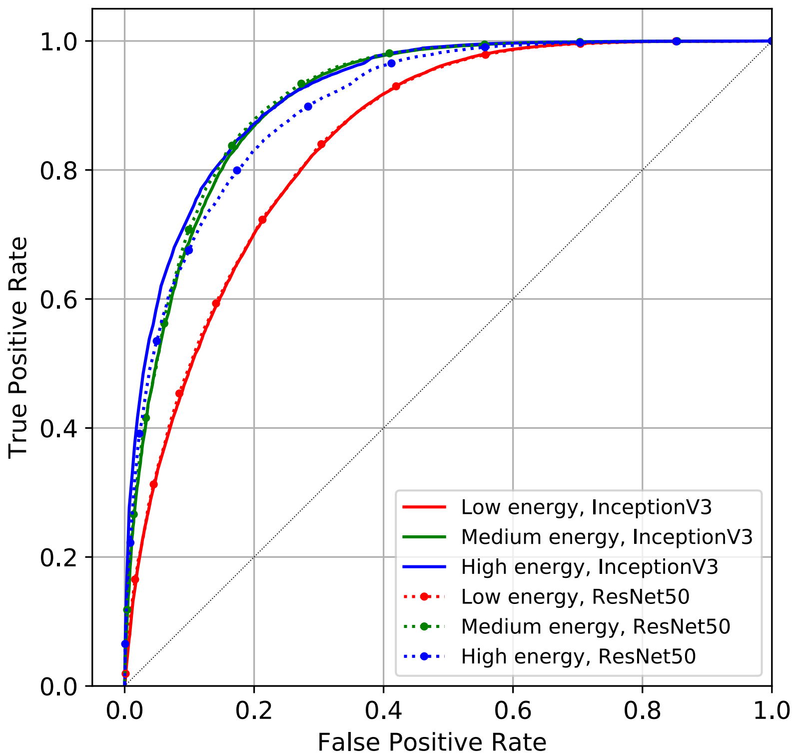
But there are risk…

- MC reliability (e.g. network selecting some features from your MC not present in real data)

Proof of concept

- Classification is happening!

(Note: results for single-telescope images)



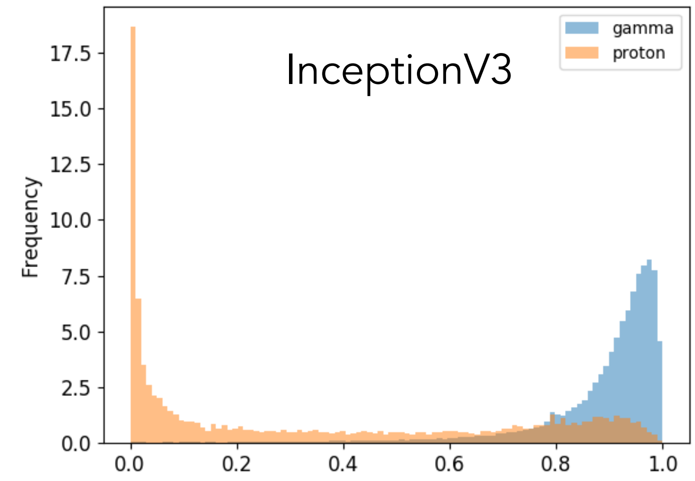
Nieto et al., PoS(ICRC2017)809

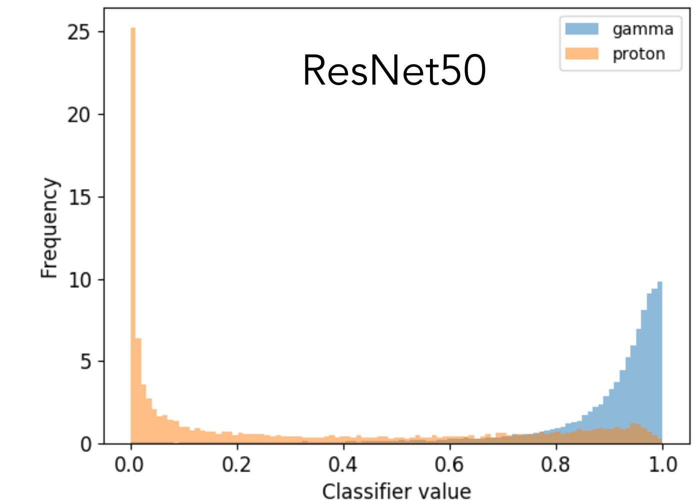
Area Under the Curve

Model/Energy	Low E.	Med. E.	High E.
InceptionV3	84.7%	91.1%	92.0%
ResNet50	84.8%	91.4%	90.2%

100% -> perfect classification
50% -> random classification

Medium energies
(0.3 TeV < E < 1 TeV)





Next step -> find the best performing model for event reconstruction

The curse of dimensionality haunts us here too!

- Hyperparameter space for deep learning architecture design
  - Number of CNN layers
  - Kernel size
  - Activation function
  - Dropout rate
  - Number of FC layers
  - Batch size
  - Learning rate
  - Optimizer
  - …

- Optimization strategies
  - Grid searches
  - Random searches
  - Bayesian optimization
  - Evolutionary algorithms
  - Reinforcement learning
  - …

+ Not that many works on models taking stereoscopic images…

Conclusions & Outlook

- Gamma-ray telescopes and IACTs in particular are competitive DM probes
- Current-generation IACTs have enhanced their performances through ML
- Next-generation IACT may profit from latest developments in ML
  - Any gain in performance can be translated into better sensitivity to DM

- Ongoing efforts to exploit deep learning as an event reconstruction method for CTA
  - Background rejection happens over non-parametrized single images
  - Working on optimizing architectures:
    - That take advantage of stereoscopic information
    - That work for energy and arrival direction estimation (regression)

Backup

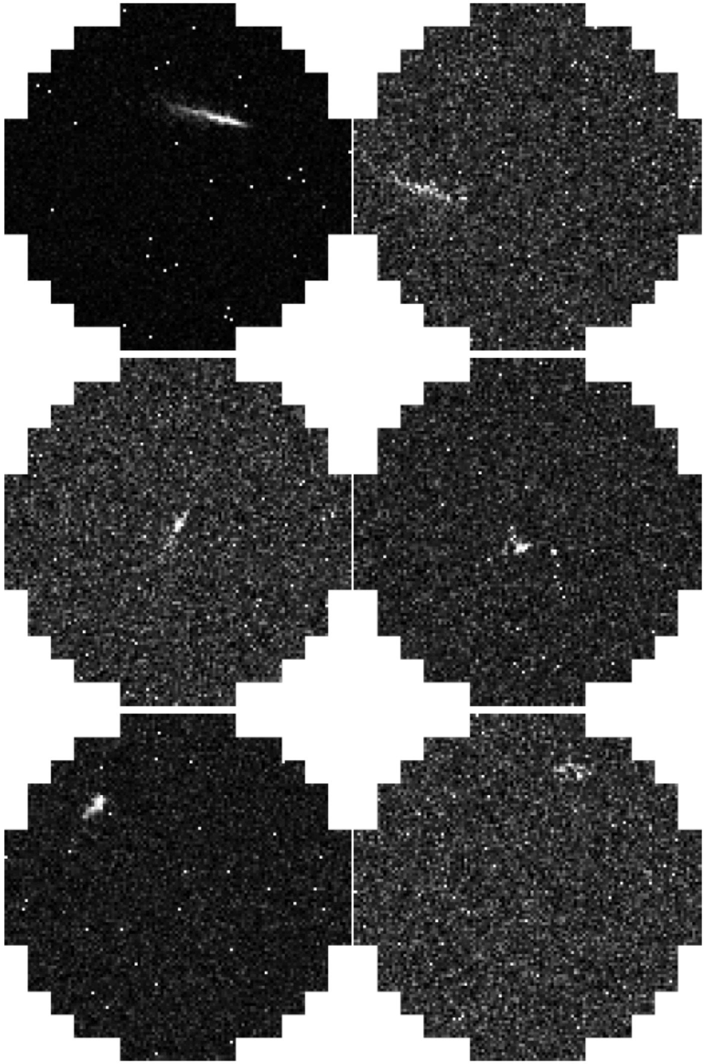
Nieto et al., PoS(ICRC2017)809

- Simulation run:
  - **Diffuse** {gamma, proton}
- Telescope array:
  - 8x SC-MST
- Three energy bins:

Bin	E_{min} [TeV]	E_{max} [TeV]	N_{gamma}	N_{proton}
Total			4160578	6518742
Low Energy	0.1	0.31	727316	499909
Medium Energy	0.31	1	657397	245912
High Energy	1	10	642034	147012

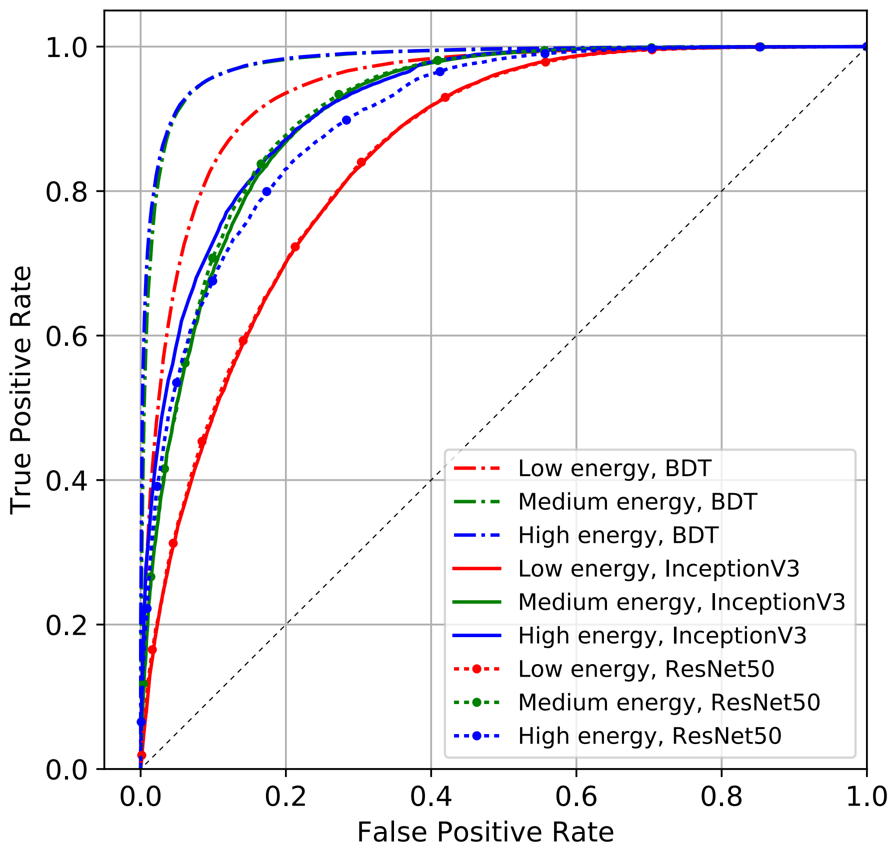
- Default ED sanity cuts prior to BDT training:

Cut
$0 \leq \sqrt{MCxoff^2 + MCyoff^2} \leq 3$
$-2 < MSCW < 2$
$-2 < MSCL < 5$
$EChi2S \geq 0$
$ERecS > 0$
$0 < EmissionHeight < 50$
$dES \geq 0$



Proof of concept: DL target performance

ROC



Accuracy

Model	Low E.	Med. E.	High E.
ResNet50	81.1%	90.1%	91.2%
Inception V3	81.4%	90.1%	91.6%

AUC

Model/Energy	Low E.	Med. E.	High E.
ResNet50	84.8%	91.4%	90.2%
InceptionV3	84.7%	91.1%	92.0%

- Expected trends in performance as a function of energy observed
- Inception V3 similar to ResNet50

- BDT ROCs shown as **reference** and a milestone to overtake
- BDT vs DL ~= 8 SCT array vs single SCT, thus a direct comparison between the two methods is **NOT** on an equal footing