Influence of uncertainty in hadronic interaction models on the sensitivity estimation of Cherenkov Telescope Array

Michiko Ohishi, L. Arbeletche[a], V. de Souza[a], G. Maier[b], K. Bernlöhr[c], A. Moralejo[d], J. Bregeon[e], L. Arrabito[e], T. Yoshikoshi for the CTA Consortium

ICRR, Univ. of Tokyo, Universidade de São Paulo[a], DESY[b], MPIK[c], IFAE[d], LUPM, Université de Montpellier, CNRS/IN2P3[e]

This work was conducted in the context of CTA Analysis and Simulation Working Group.

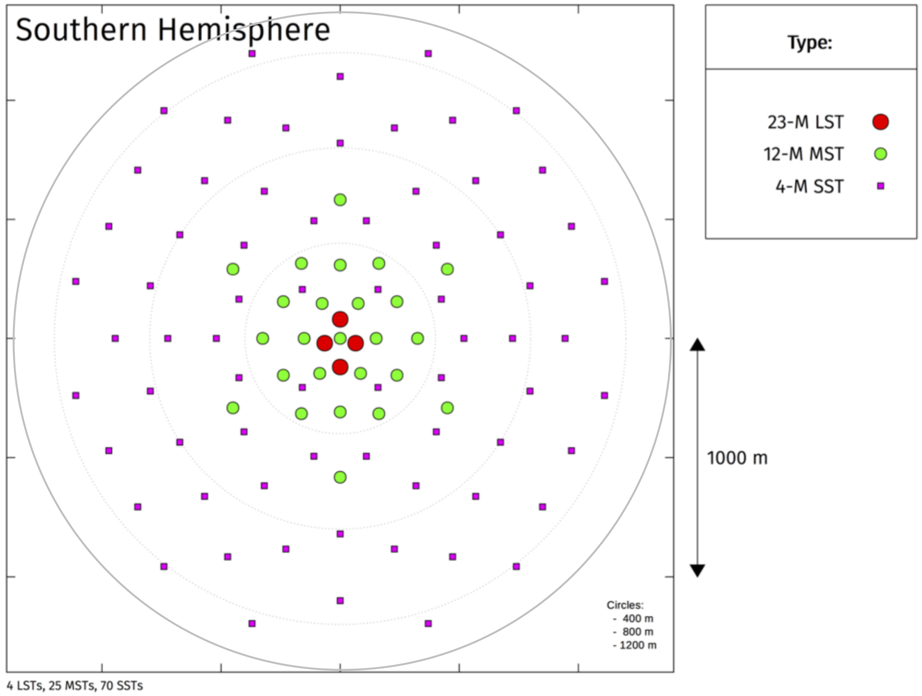
Outline

- **Introduction** γ-ray sensitivity of CTA and cosmic-ray backgrounds
- **Difference of hadronic interaction models in shower particles**
  - π^0 spectrum
  - energy fraction consumed in electromagnetic (EM) components
- **CTA simulation and analysis**
  - Energy scale and shower rate of cosmic-ray proton
  - Basic shower parameters and γ-hadron separation MVA parameters
  - Differential sensitivity
- **In the viewpoint of model verification**
  - Difference in γ–ray-like event rate
  - Contribution from heavy nuclei
- **Summary**

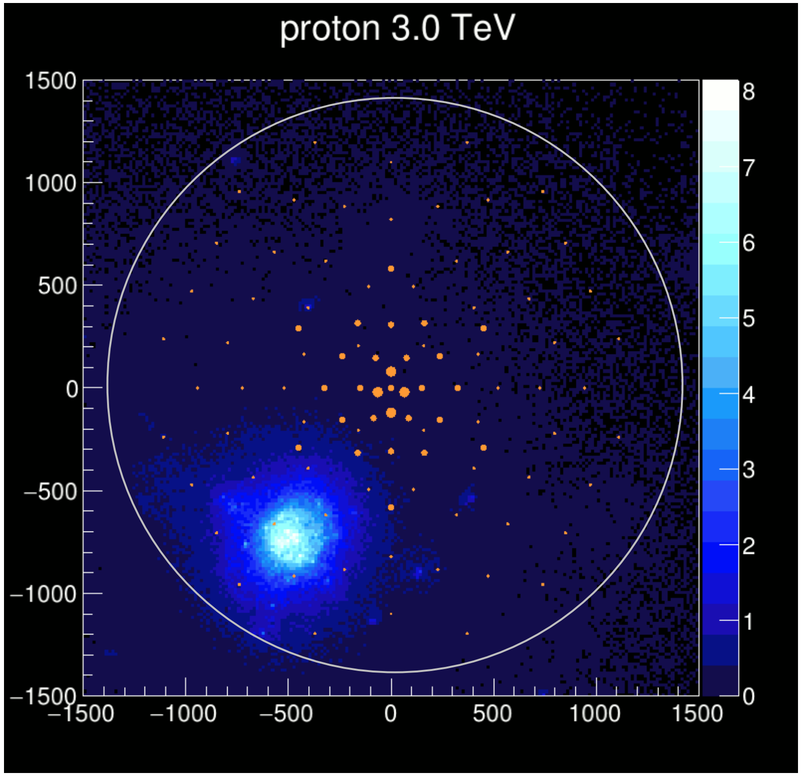
Current IACT systems and CTA (array scale)

- **Current IACT arrays** (H.E.S.S., VERITAS, MAGIC): coverage of ~0.03 km^2
- **CTA** : ~ 4 km^2 for South site (99 telescopes)
  ~ 0.6 km^2 for North site (19 telescopes)

→ Full containment of Cherenkov photons from γ-ray and proton showers



Array configuration (South site), public at https://www.cta-observatory.org/science/cta-performance/

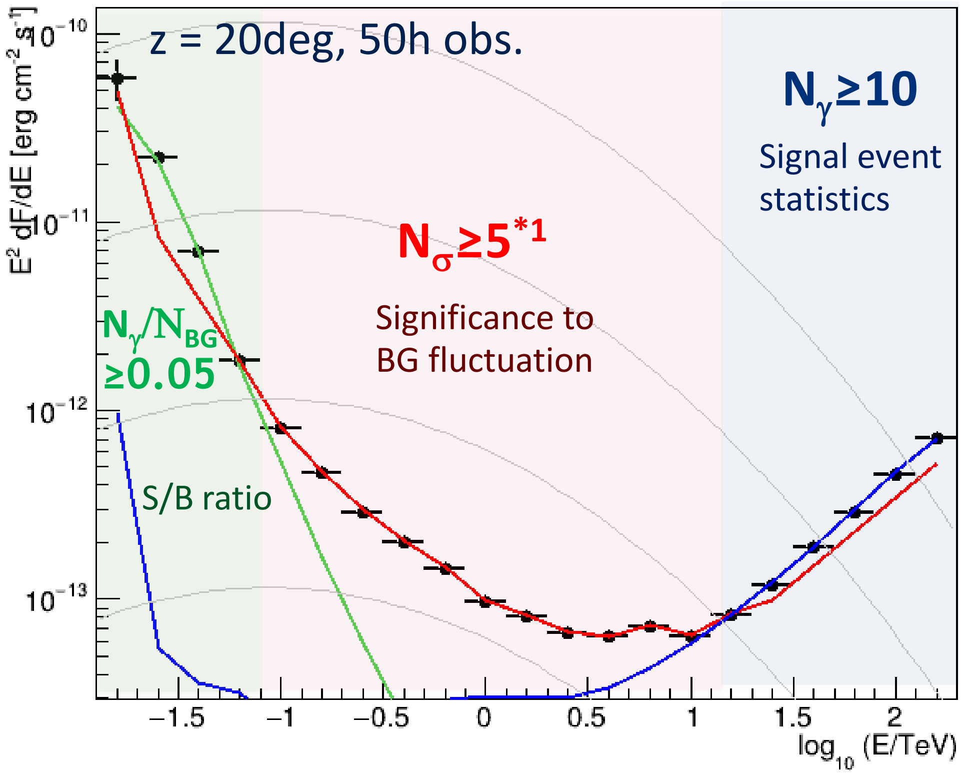


Array scale and Cherenkov light-pool size

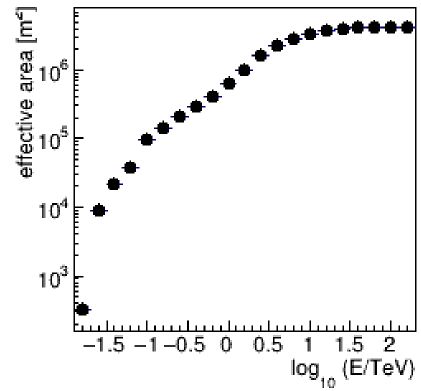
γ-ray sensitivity of CTA

- γ-ray sensitivity of an IACT system is mostly determined by
  - Significance of signal events to the background fluctuation ($\geq 5\sigma$)
  - Signal-to-background ratio ($\geq 5\%$)

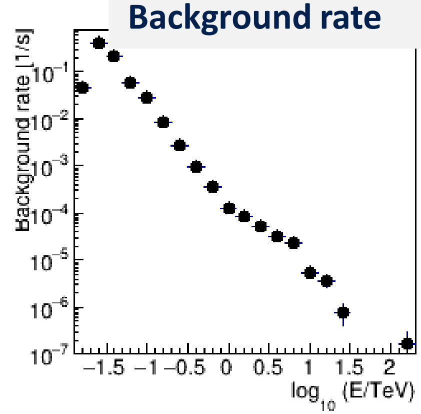
Differential Sensitivity of CTA South array



γ-ray effective area



Background rate



"Background" ≈ CR proton + electron

CTA Instrument Response Functions (IRFs), public at https://www.cta-observatory.org/science/cta-performance/

*1 Significance def. in Li & Ma (1983), Eq. (17)

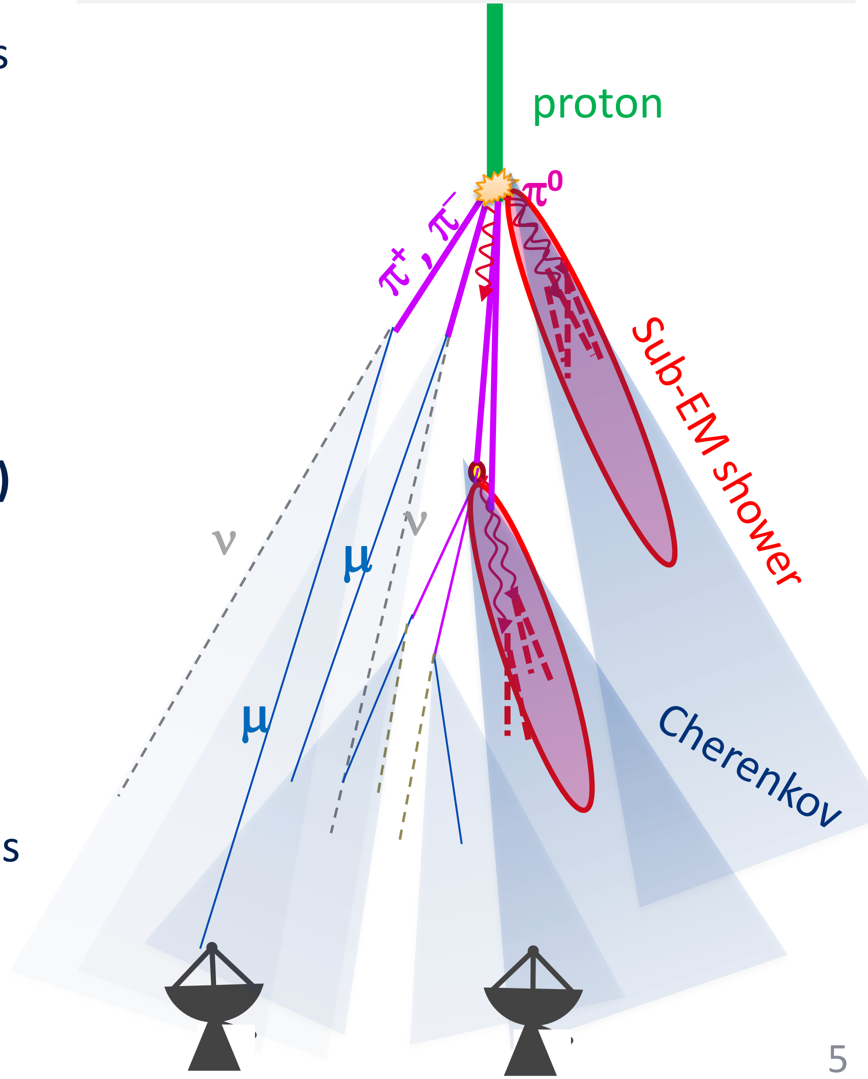
Estimation of background level in IACT systems

- **Current IACT systems**
  - Real cosmic-ray data (“**OFF-source” data**) are used as background samples
  - Real OFF-source data are used in both of training of machine learning for γ-hadron separation and estimation of residual background
- **CTA** (and systems in design/construction phase)
  - **Monte Carlo (MC) simulation** data are used for background estimation
  - Usually cosmic-ray **protons** and **electrons** are simulated as backgrounds
  - As for **proton**: currently interaction between cosmic-ray proton and nuclei in very-high-energy region is not perfectly understood
    - several **hadronic interaction models (QSGJET, EPOS, SIBYLL…)** are in use in VHE/UHE CR field
    - Improvement of models with feedback from collider and CR experiments is ongoing

Hadronic shower and IACT observation

- IACT array detects Cherenkov photons from sub-EM showers (primarily from π^0) and muons contained in a hadronic shower
- Energy spectra and angular distribution of secondary particles are different from model to model
- Related studies in IACT field :
  - Cherenkov photon density **(Parsons+ 2011)**
  - Muon flux on the ground **(Mitchell+ 2019)**
  - Nature of γ-ray-like proton events (sub-EM showers mimic gamma-ray showers) **(Maier+ 2007, Sitarek+ 2017)**
- Discrimination ability of model difference depends on the array performance - this study is focused on CTA, testing QGSJET-II-03 (currently used in CTA) and recent post-LHC models

Schematic diagram of a hadronic shower

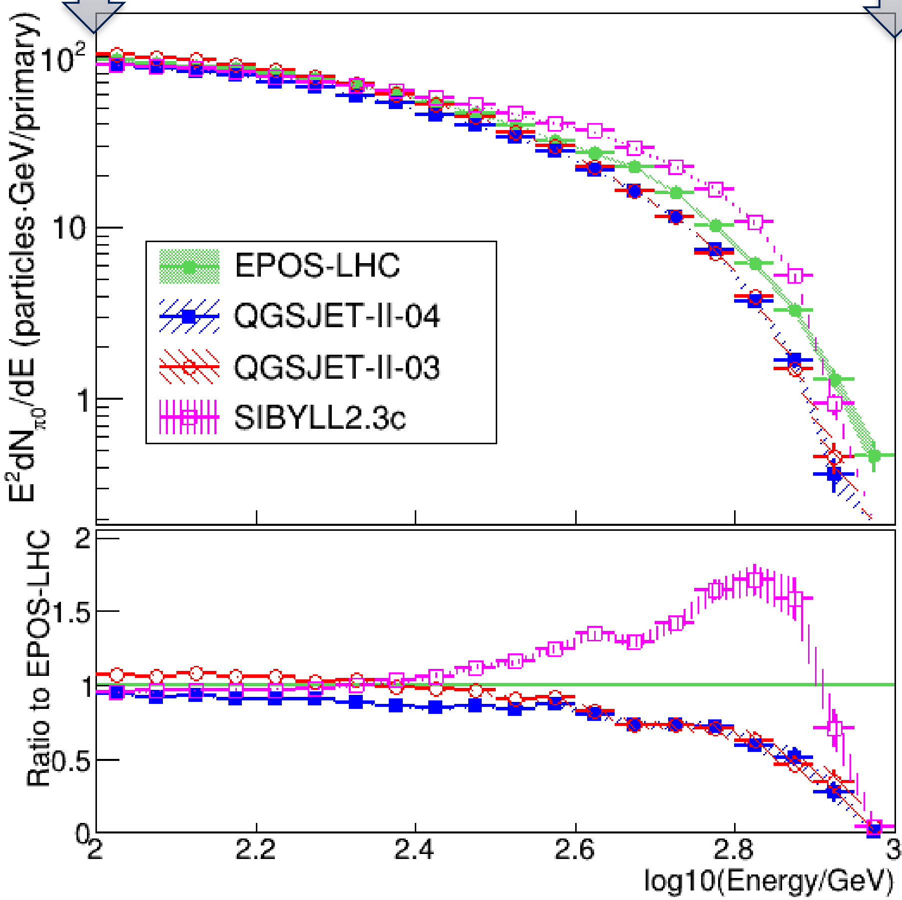


Difference of models in shower particles
- π^0 spectrum -

10% of primary energy

primary proton energy

π^0 spectra, E_p=1 TeV (mono)



- Air shower simulation with CORSIKA to investigate difference of secondary particles between different models

- Used models:
  - **QGSJET-II-03** in CORSIKA6.99 (currently used in CTA)
  - **QGSJET-II-04, EPOS-LHC, SIBYLL2.3c** in CORSIKA7.69
  - E<80 GeV: fixed low energy model UrQMD (for all cases)
- **π^0 spectrum**
  - Spectrum at high energy end can affect the rate of γ-ray-like events
  - Harder spectrum tends to give more γ-ray-like BG events:

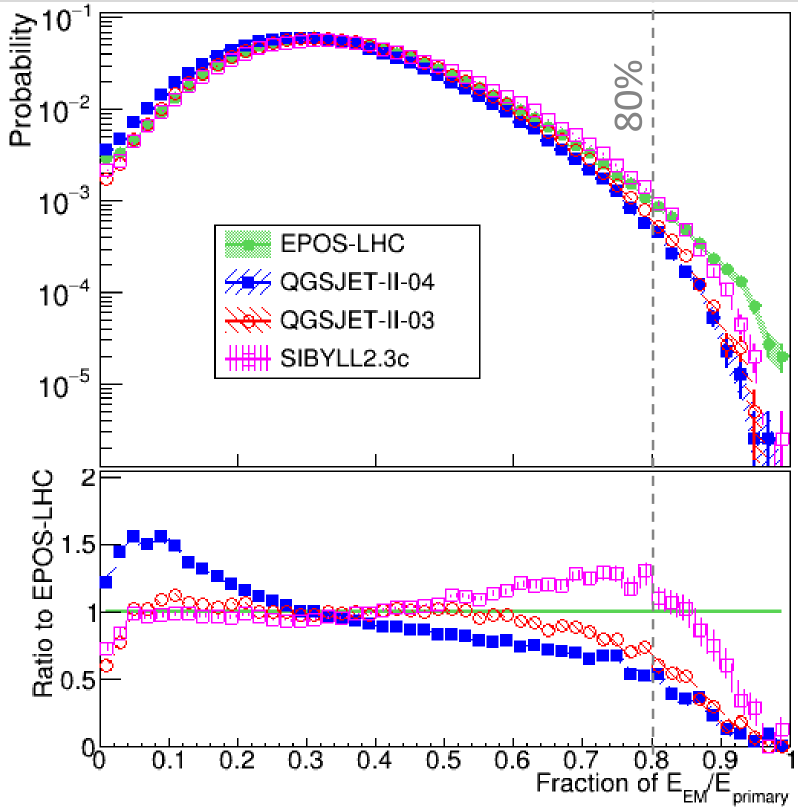
    EPOS $\rightarrow$ SIBYLL

    $\rightarrow$ QGSJET-II-03 $\approx$ QGSJET-II-04

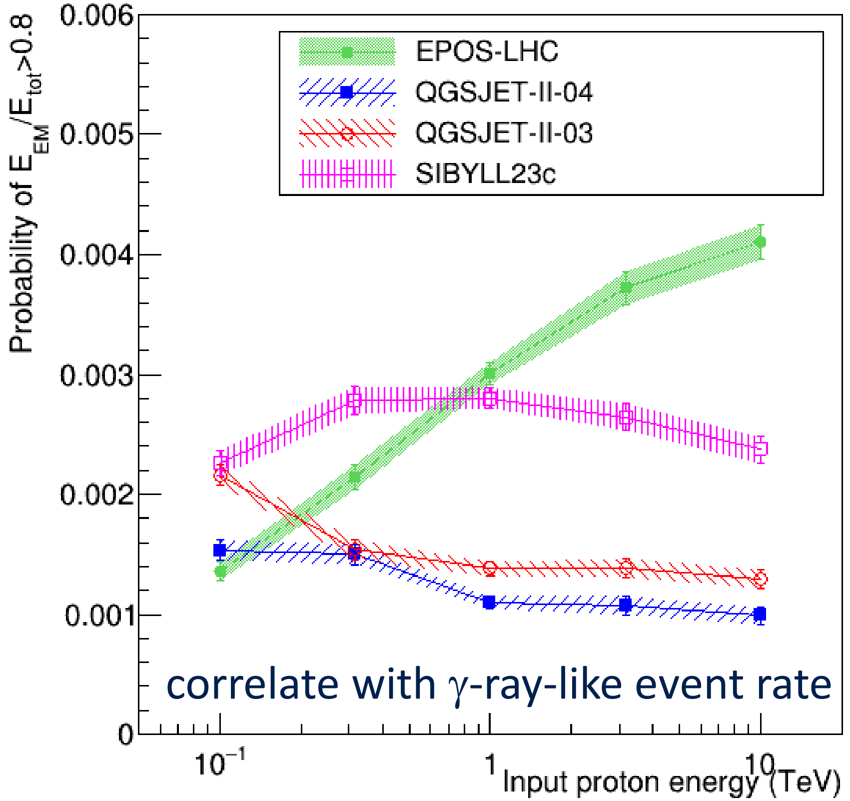
Difference of models in shower particles
- Energy fraction in EM components -

- Energy fraction carried by γ +e⁻+e⁺ (EM components) after the 3rd interaction (as for γ-ray primary case, this fraction is close to 100%)
- Similar pattern as π^0 spectrum is seen; relation between model changes at ~1 TeV
- Energy fraction in EM which will be regarded as "γ-ray-like" event depends on the array performance -- 80% was used in this study for CTA

$E_{EM}/E_{primary}$ distribution, E_p=1 TeV case

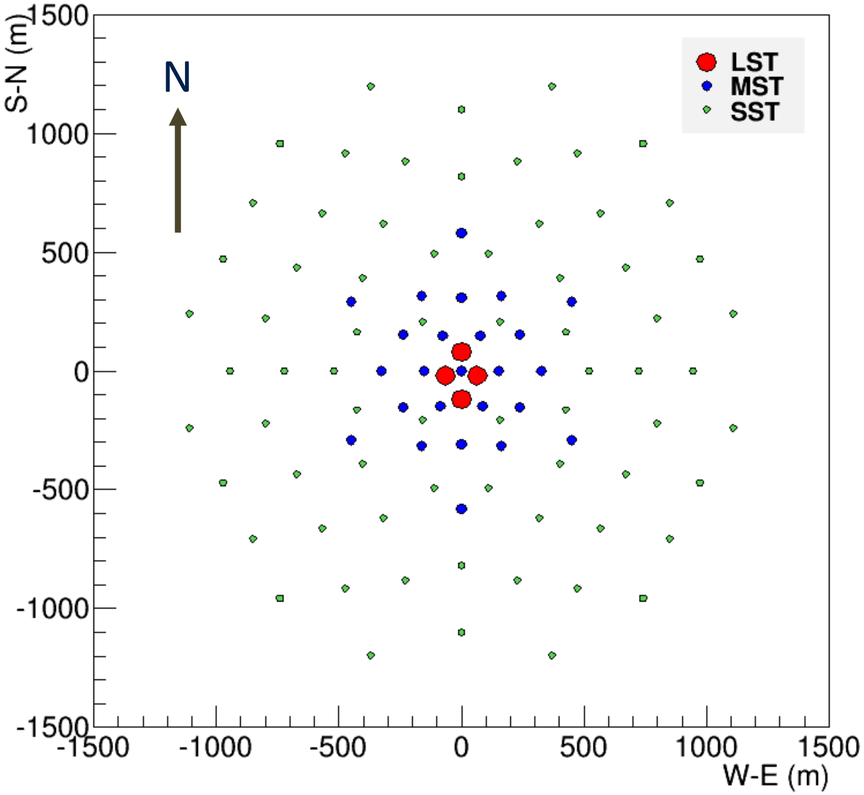


Prob. of high EM fraction events VS true E



CTA simulation

Array configuration, South Site



Analysis tool: EventDisplay v500-rc04

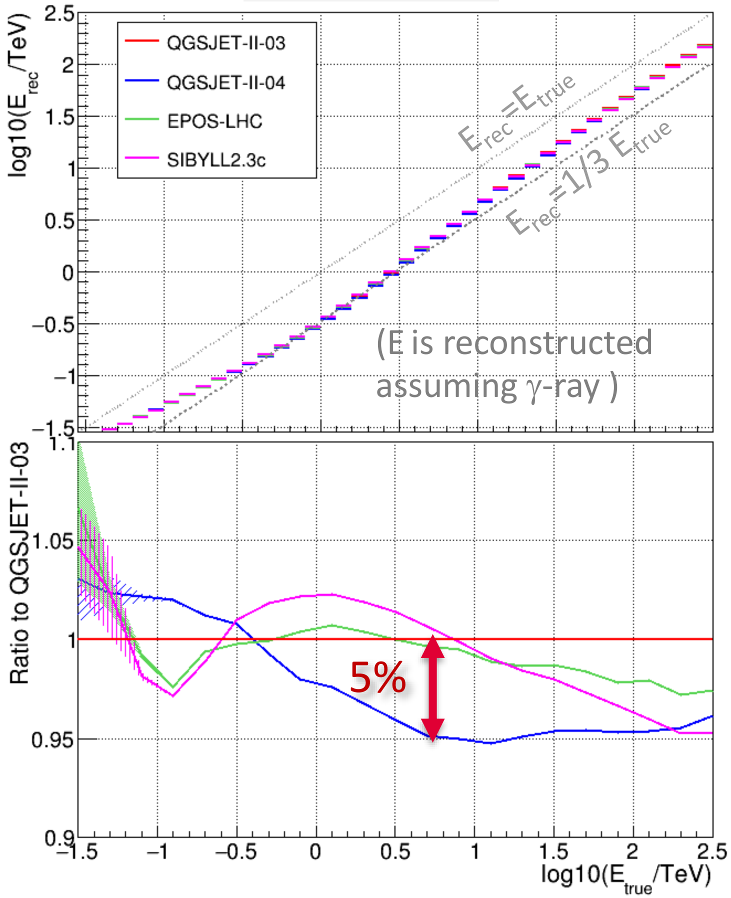
Site	Paranal (Chile)
Array	4 LSTs, 25 MSTs, 70 SSTs (configuration shown left)
Particle	Gamma, e-, proton: QGSJET-II-03 *1 proton :QGSJET-II-04 EPOS-LHC v3.4 /SIBYLL2.3c*2 Low Energy Model (E<80 GeV) : fixed as UrQMD
Core range	2500 m
Viewcone	0 - 10 deg
Energy range	0.003 - 330 TeV (e-, gamma) 0.004 - 600 TeV (proton)
Spec. index	-2.0 *3

*1 in CORSIKA 6.99, produced on GRID system in EU
*2 in CORSIKA 7.69, produced on cluster in Japan
*3 Reweighted in the analysis

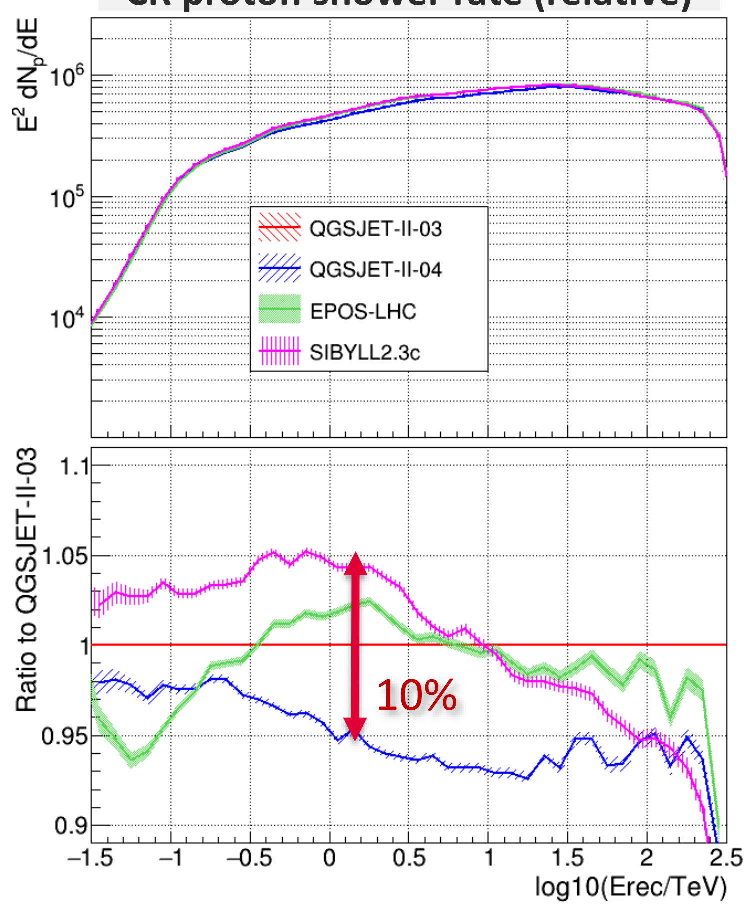
Energy scale and shower rate

- Difference in π^0 production can lead to difference in E scale and CR proton rate
- **~5%** difference in reconstructed energy and **~10%** difference in CR proton rate between models (before gamma-ray selection cuts)

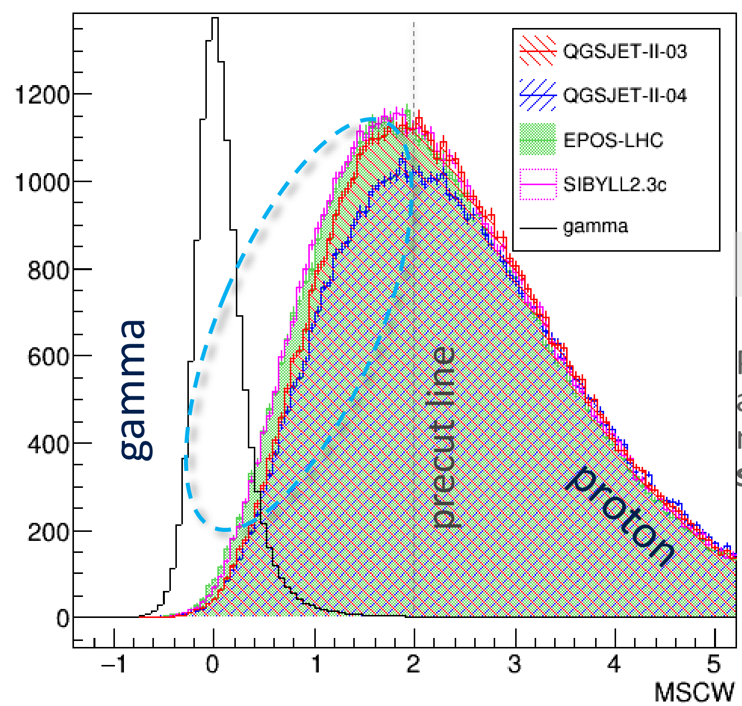
E_{rec} VS E_{true}



CR proton shower rate (relative)

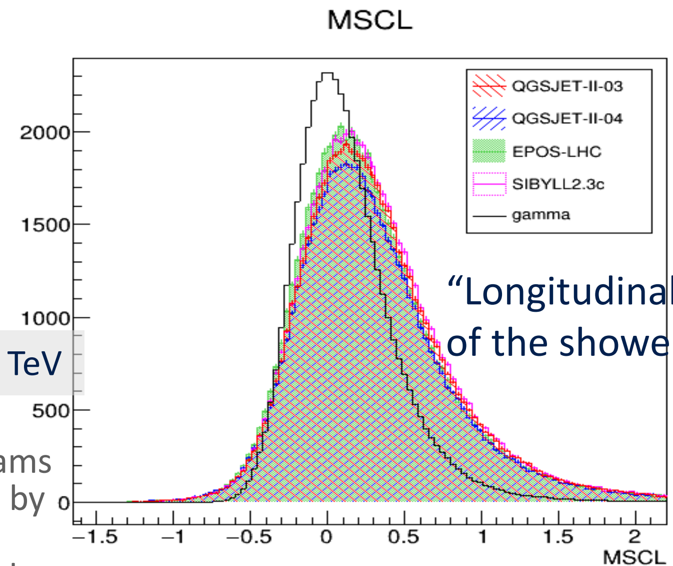


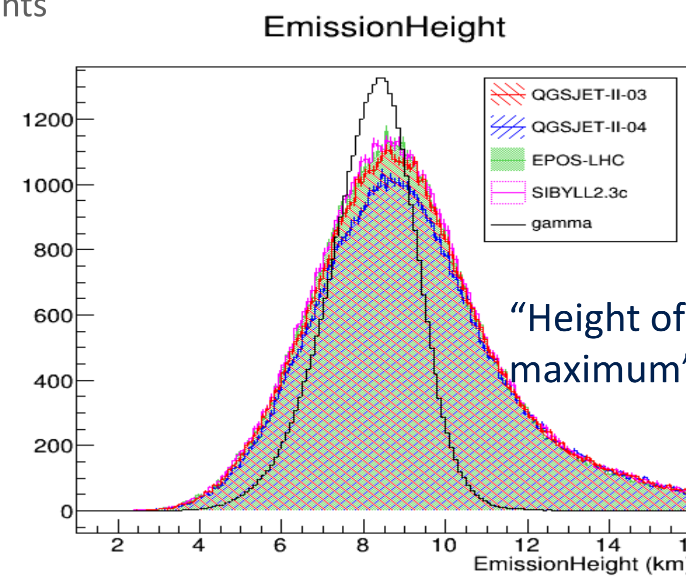
Difference in basic shower parameter distribution



1 TeV<E_{rec}<10 TeV

Proton histograms are normalized by number of **simulated** events



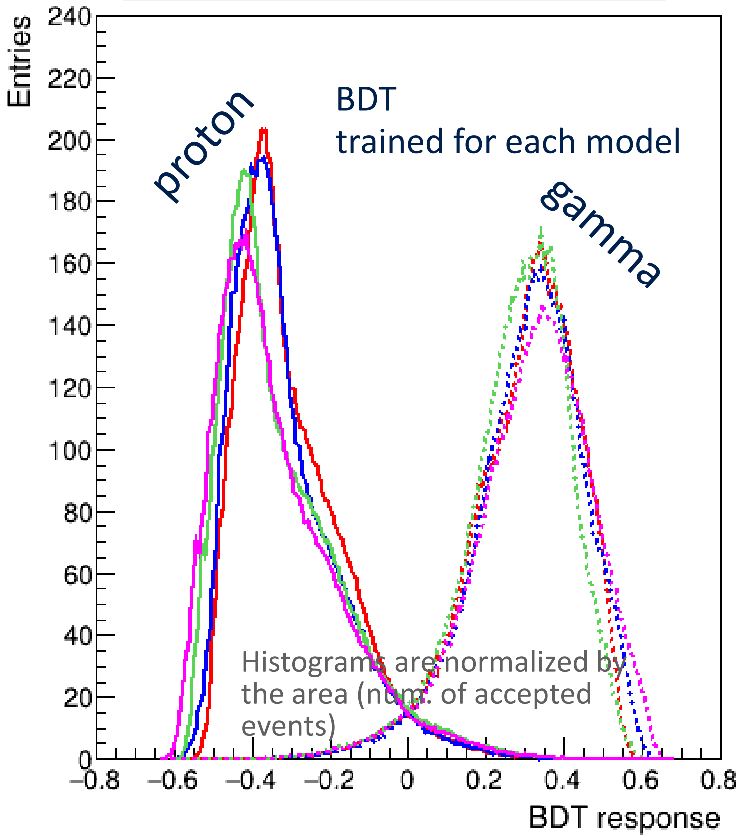


- Most important shower characteristics for γ-hadron separation : **WIDTH** (lateral size of the shower)
- MSCW : corrected and normalized WIDTH
- Difference between models is seen at small MSCW (γ-ray-like region)

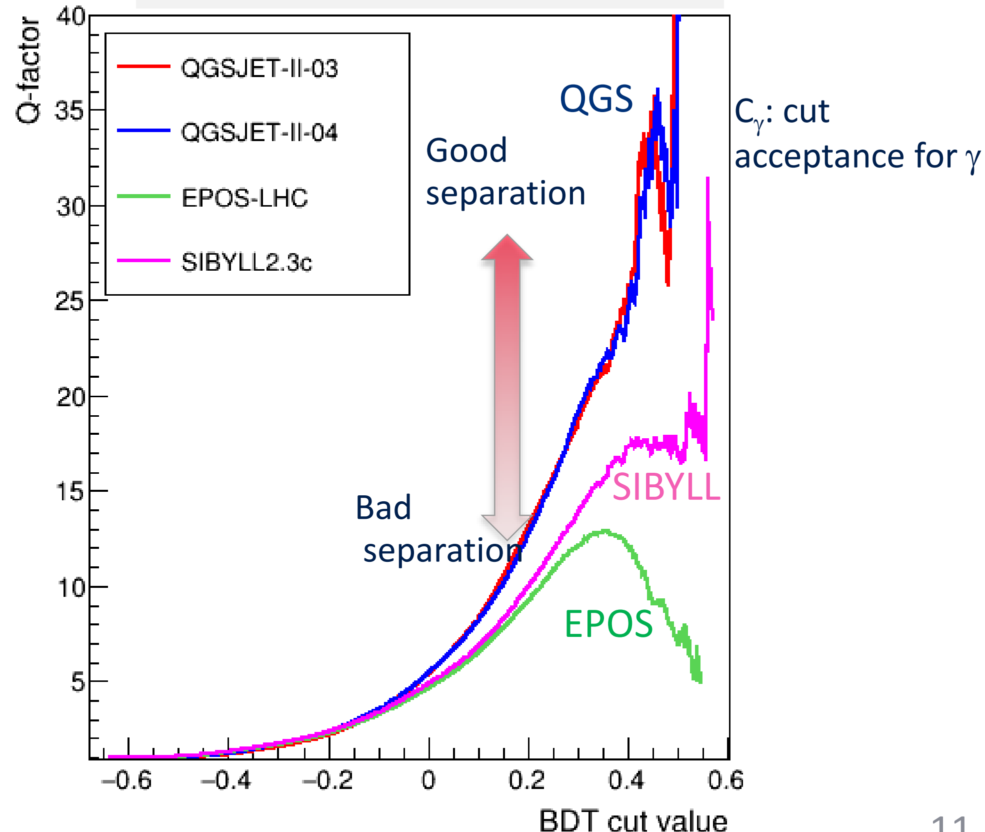
MVA parameters for γ-hadron separation

- Multivariate analysis (MVA) to introduce a single index of "gammaness" (or hadroness)
  - Boosted Decision Tree is used here, with precuts in basic shower parameters
- **EPOS** and **SIBYLL** show worse separation, with more γ-like events than QGS as expected

MVA parameter distribution
1.0 TeV <E_{rec}< 5.6 TeV



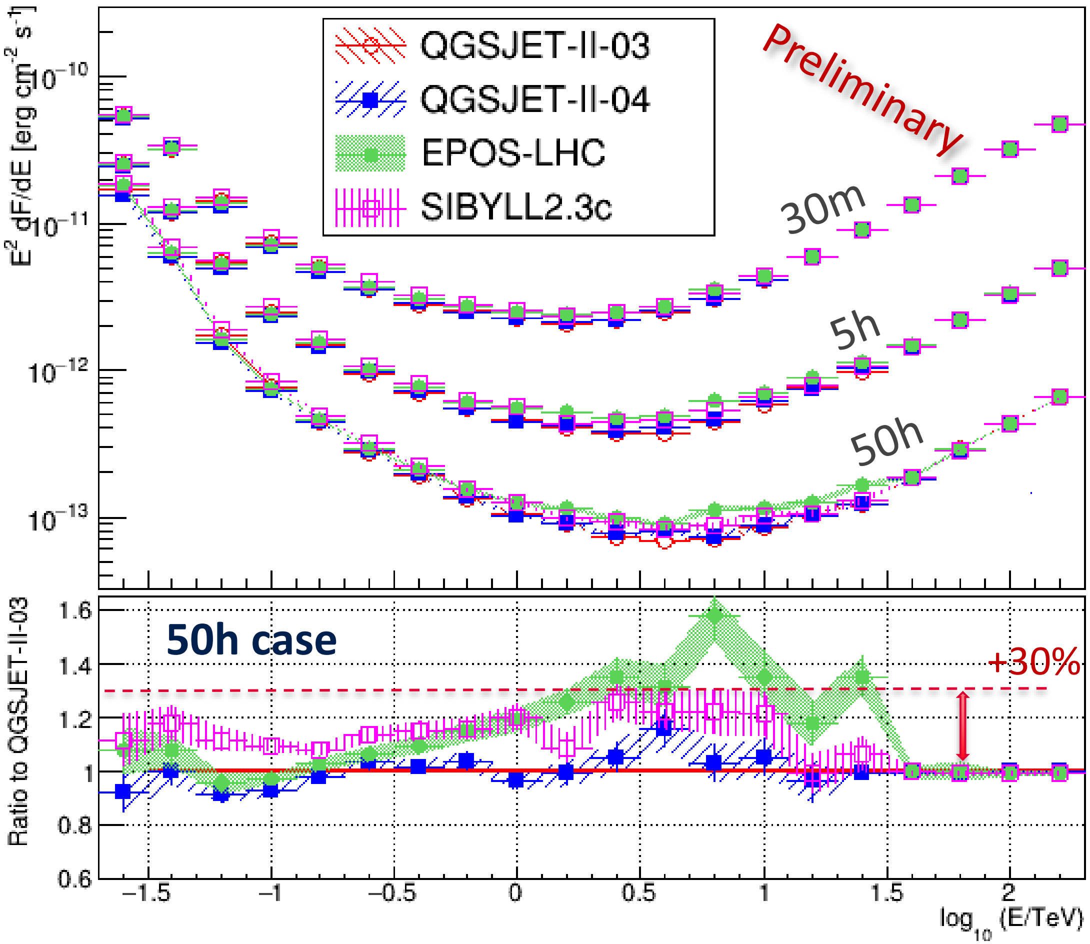
$Q \equiv C_\gamma / \sqrt{C_p}$ **VS BDT cut value**



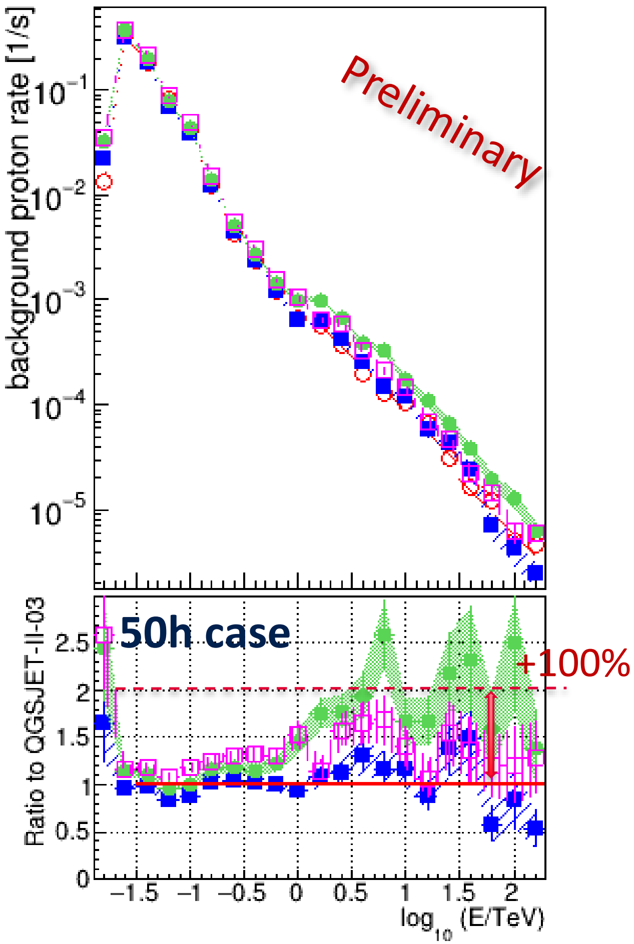
Differential sensitivity

South site, LST+MST+SST array, z=20deg, average of North+South pointing

Differential sensitivity

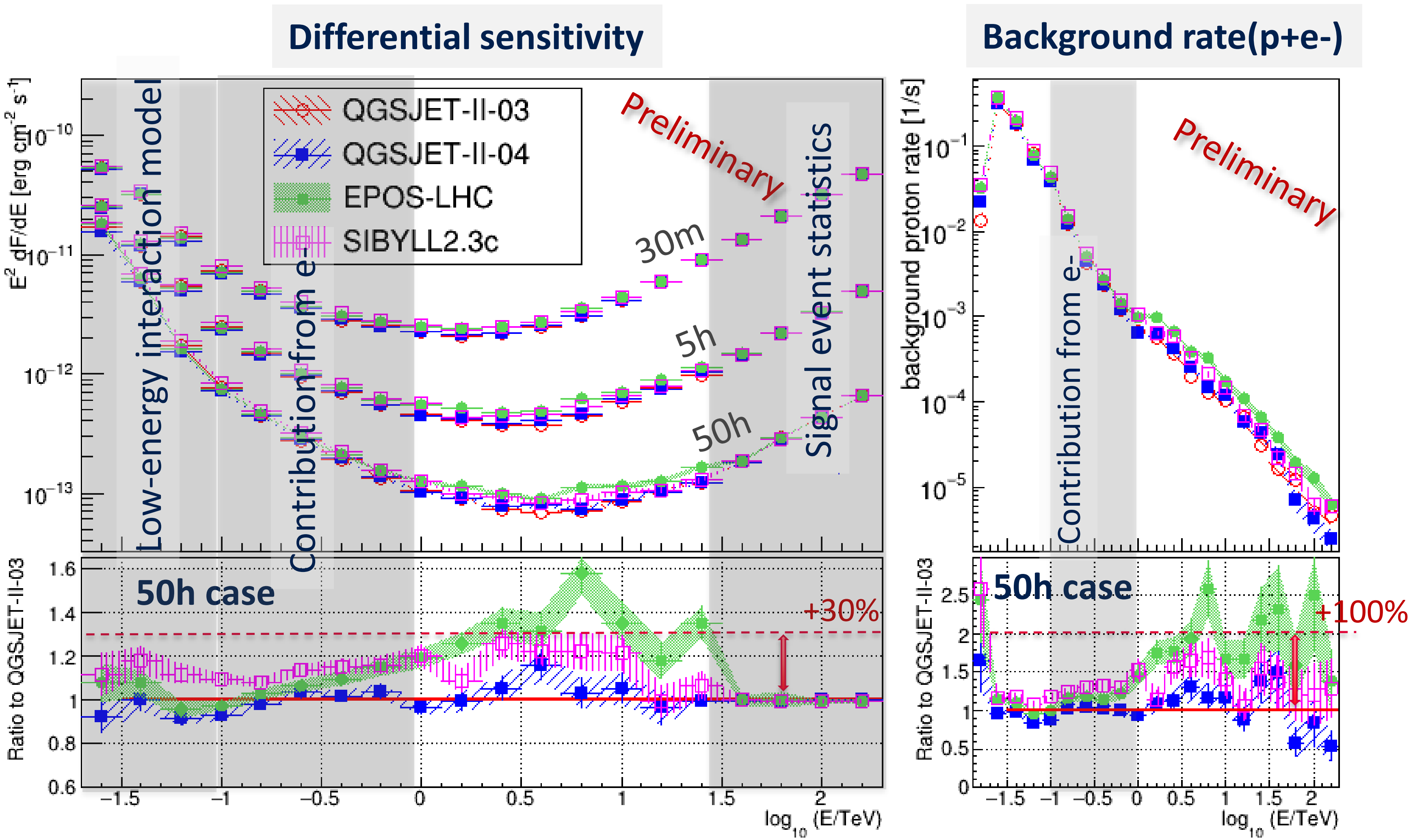


Background rate(p+e-)



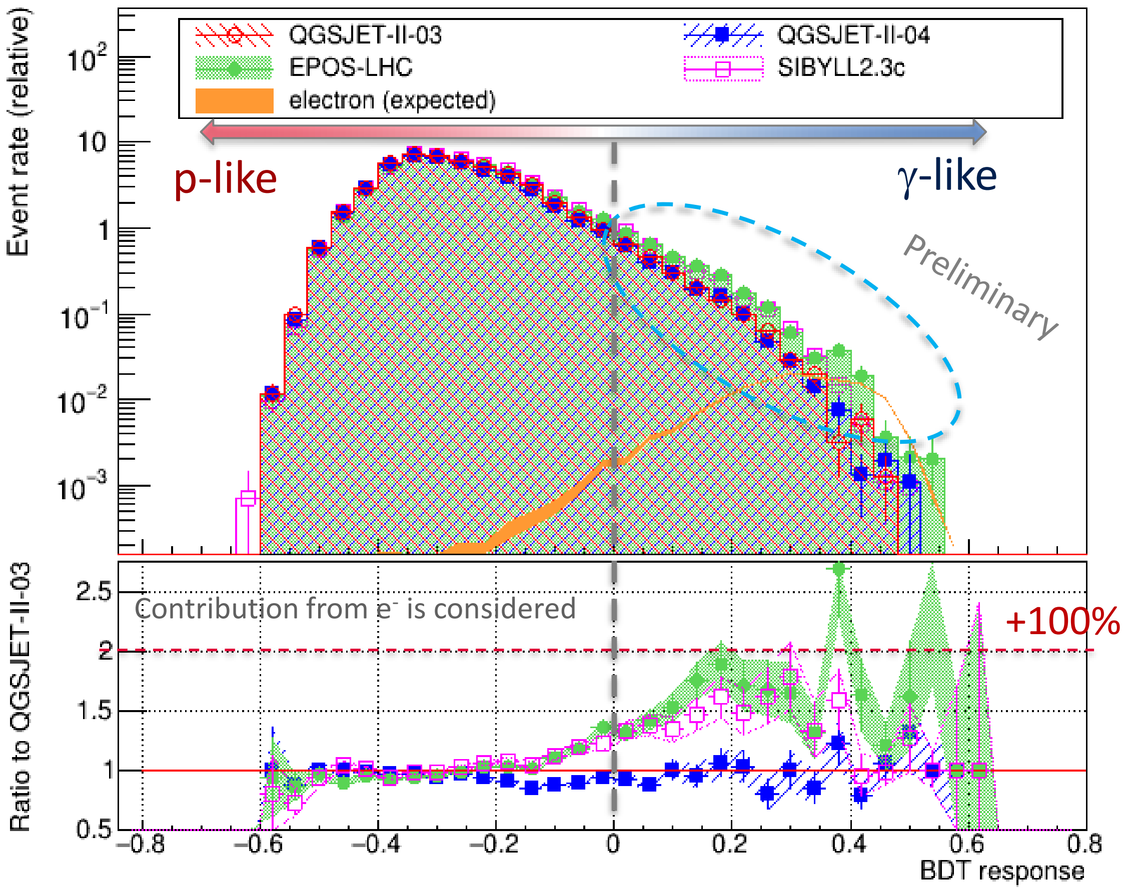
Differential sensitivity

South site, LST+MST+SST array, z=20deg, average of North+South pointing



In the viewpoint of model verification with IACTs

MVA parameter distribution (1 TeV < E_{rec} < 10 TeV)

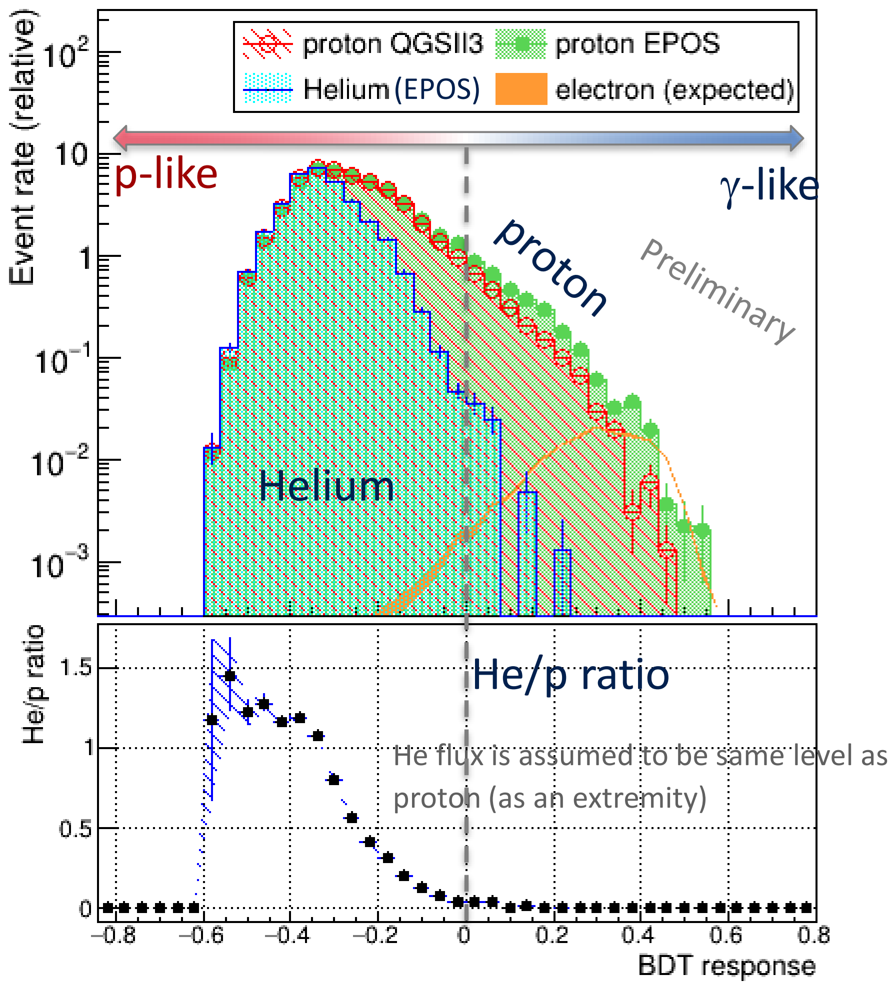


identical trained BDT (QGSJETII-03) is used for all models

- Once we have real CR data, we can test which model is the closest to the reality by comparing MC and real data :
  - Event rate
  - Shower param. dist.
  - γ-hadron separation parameter dist. (relatively large factor ~2 difference)
- **Current IACT systems can also contribute** to model verification, though model discrimination ability depends on the array performance (worse than CTA).

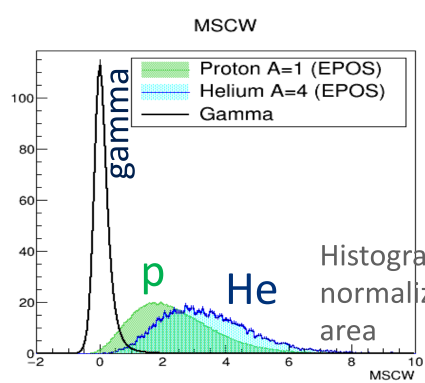
Model verification: contribution from heavy nuclei?

MVA parameter distribution (1 TeV < E_{rec} < 10 TeV)



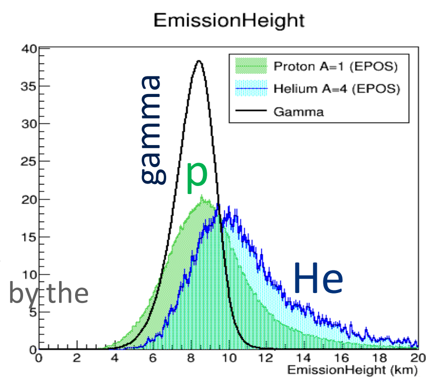
- Uncertainty in **CR composition** can affect the model verification accuracy
- As far as treating γ-ray-like events, contribution from heavy nuclei is negligibly small
  → **good** verification **measure**
- Helium and heavier nuclei do not mimic γ-rays because of their lateral size and shower maximum height

Lateral size of shower



Histograms normalized by the area

Height of shower max.



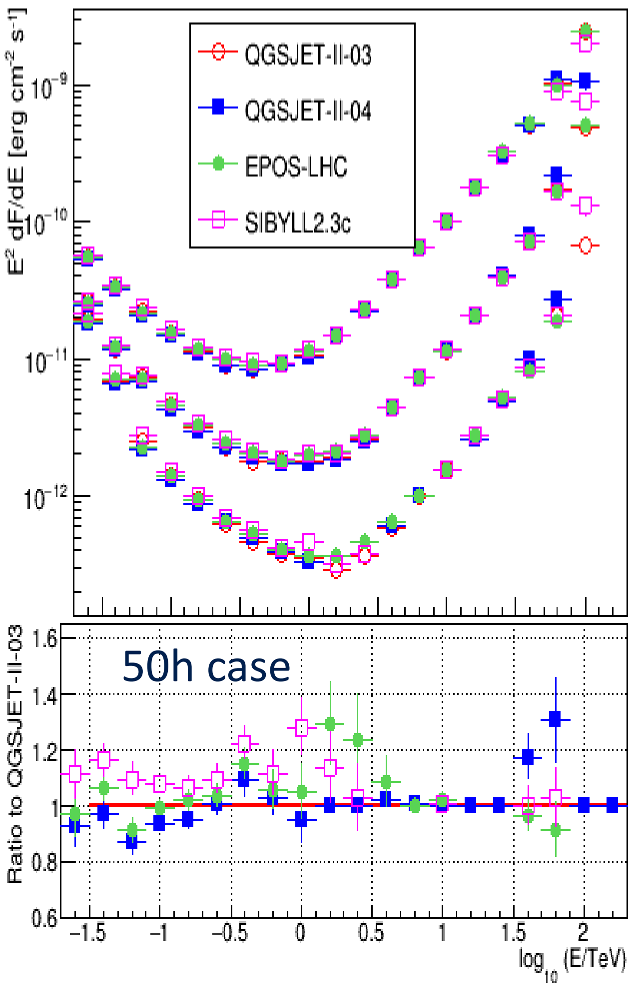
Summary

- Effect of difference in hadronic interaction models on gamma-ray sensitivity of **CTA south array** (99tels, 4-LSTs + 25-MSTs + 70-SSTs) was estimated with MC simulation data
- Tried models:
  - **QGSJET-II-03** in CORSIKA6.99 (currently used for CTA IRF)
  - **QGSJET-II-04**, **EPOS-LHC**, **SIBYLL2.3c** in CORSIKA7.69 (post-LHC models)
- **~5%** level difference in energy scale and **~10%** level difference in proton shower rate were seen.
- As a preliminary result, difference in γ-ray sensitivity between models was estimated to be ~30% level (with ±10% statistical error from MC data); Relation between models is consistent with π^0 spectrum and EM fraction
- In the viewpoint of **model verification**, γ-ray-like event rate is a relatively good measure :
  - almost free from uncertainty of cosmic-ray nuclei **composition**
  - relatively large (factor ~2) difference between models
- **Current IACT systems** can also contribute to model testing (discrimination ability depends on the array performance)

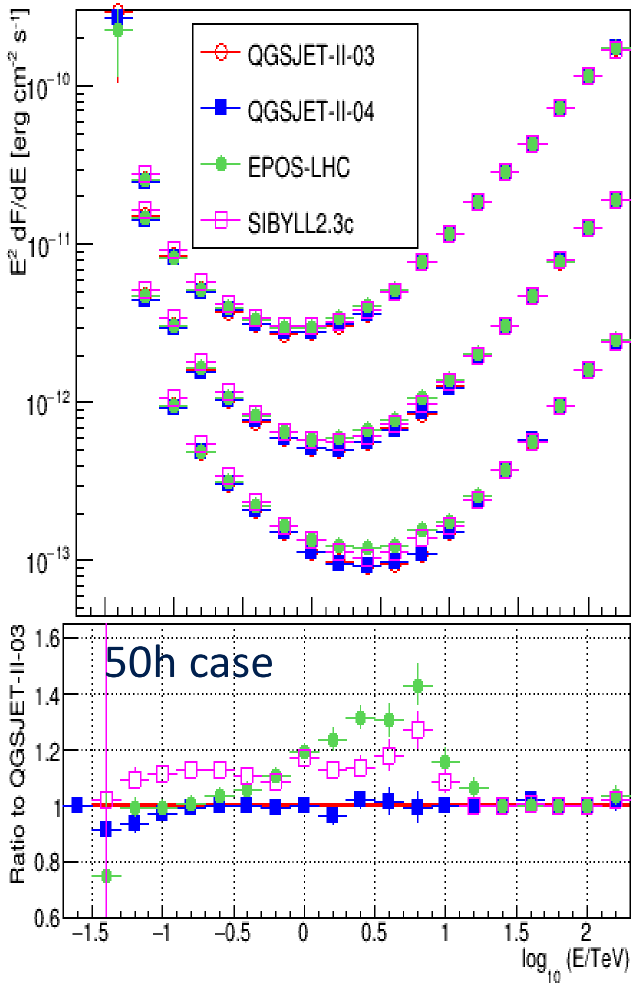
Backup slides

Differential sensitivity – subsystems -

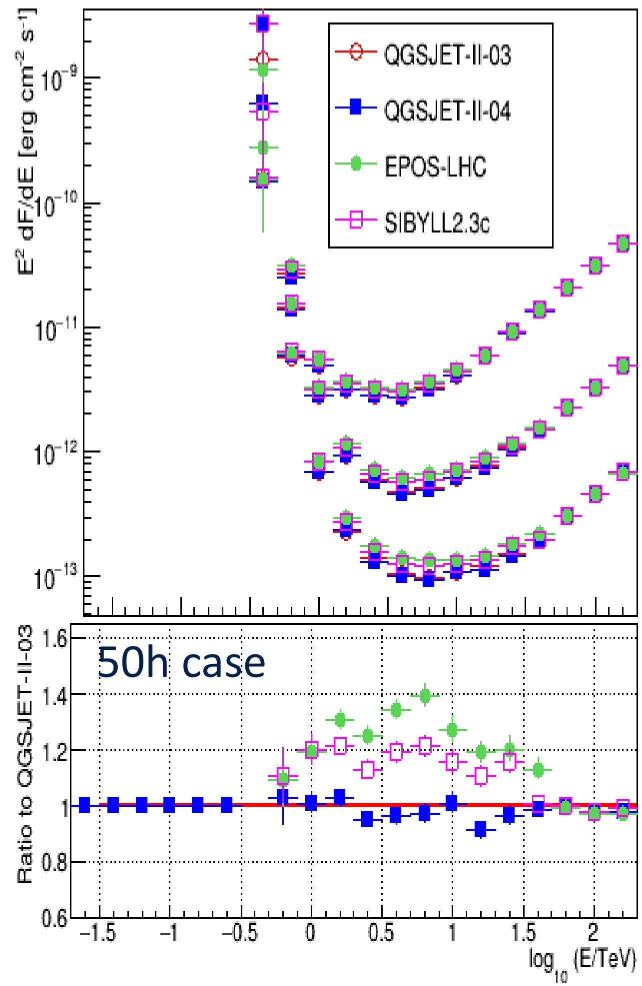
LST only



MST only



SST only



CR spectra used in the background calculation

- **CR proton**

$$\frac{dN}{dE} = I_0 \left(\frac{E}{E_C}\right)^{-\Gamma}$$

$$I_0 = 9.8 \times 10^{-6}\ \mathrm{cm^{-2}\ s^{-1}\ TeV^{-1}\ str^{-1}},\ E_C = 1.0\ \mathrm{TeV}, \Gamma = 2.62$$

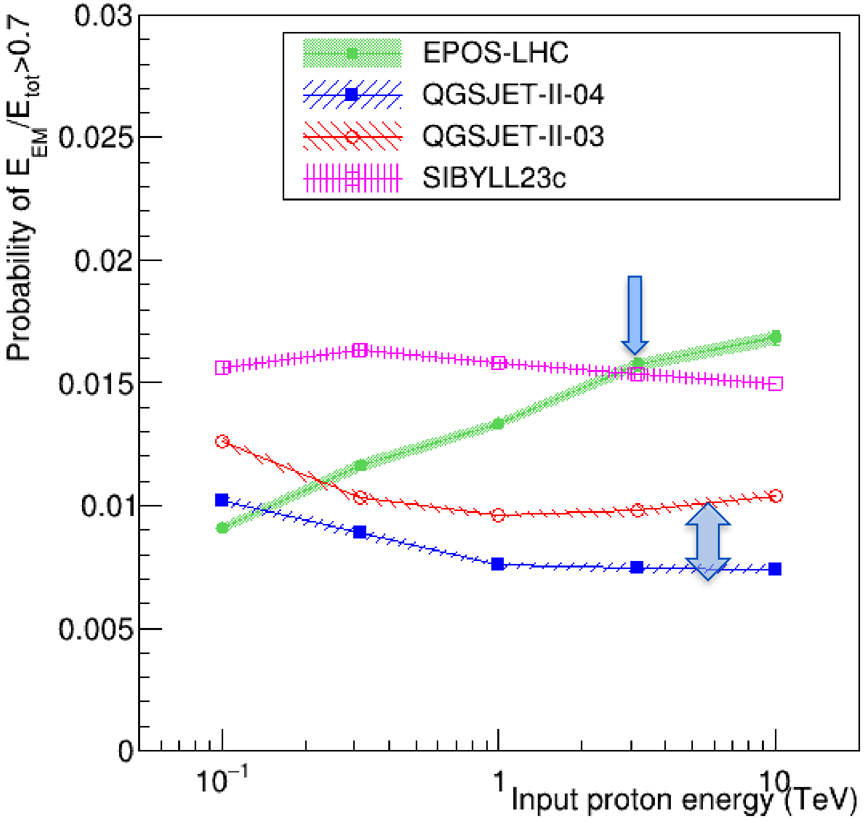
- **CR electron**

$$E^3 \frac{dN}{dE} = I_0 \left(\frac{E}{E_C}\right)^{-\Gamma} \times (1 + f \times (\exp(\exp(-\frac{(\log_{10}(E/E_C) - \mu)^2}{2\sigma^2}))\text{-}1))$$

$$I_0 = 2.385 \times 10^{-9}\ \mathrm{cm^{-2}\ s^{-1}\ TeV^{-1}\ str^{-1}},\ E_C = 1.0\ \mathrm{TeV},$$
$$\Gamma = 3.43, \quad \mu = -0.101, \sigma = 0.741, f = 1.950$$

High EM fraction event prob. VS trueE

EM frac>70% prob.



EM frac>60% prob.