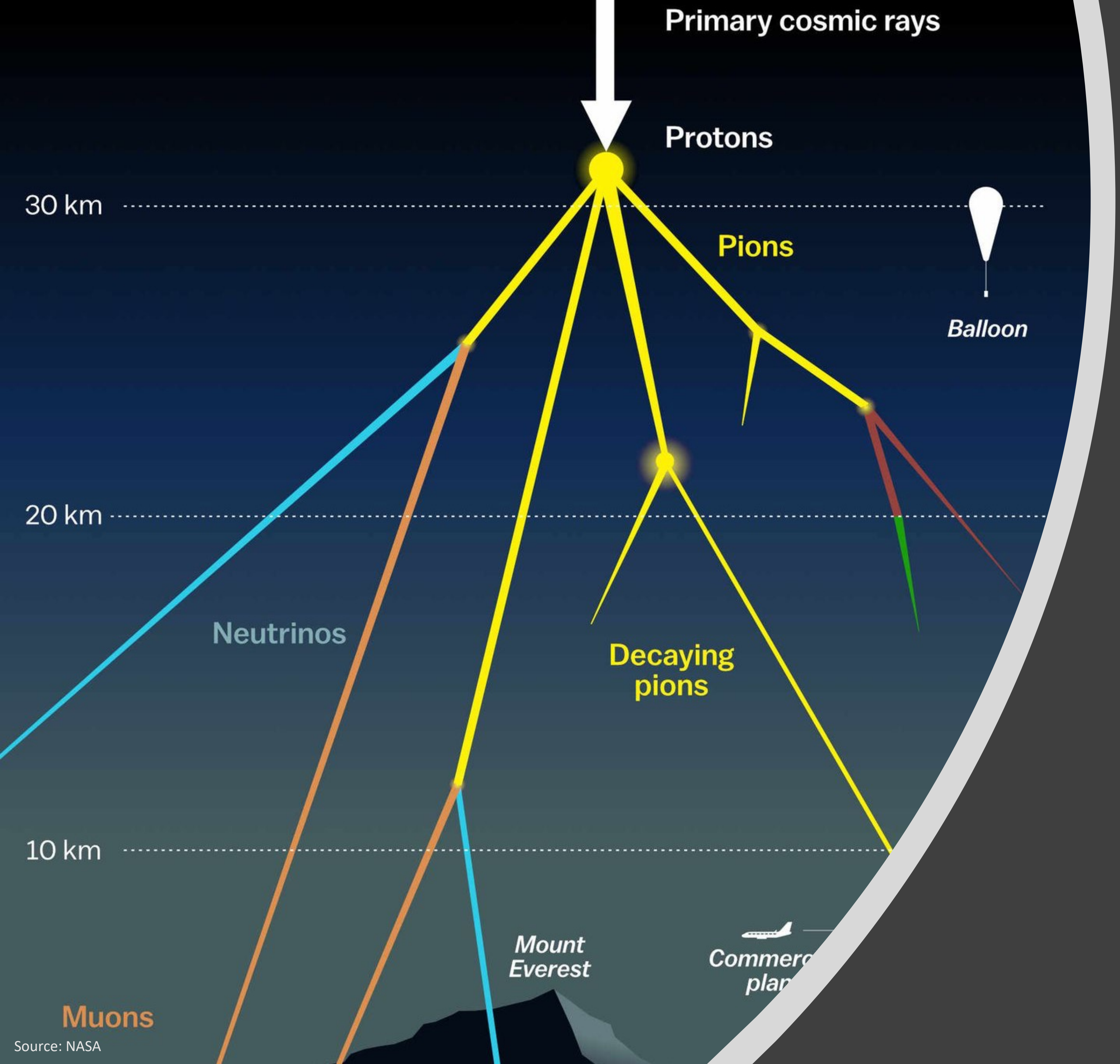


Source: NASA

Update on prompt muon analysis – reconstructions and selection

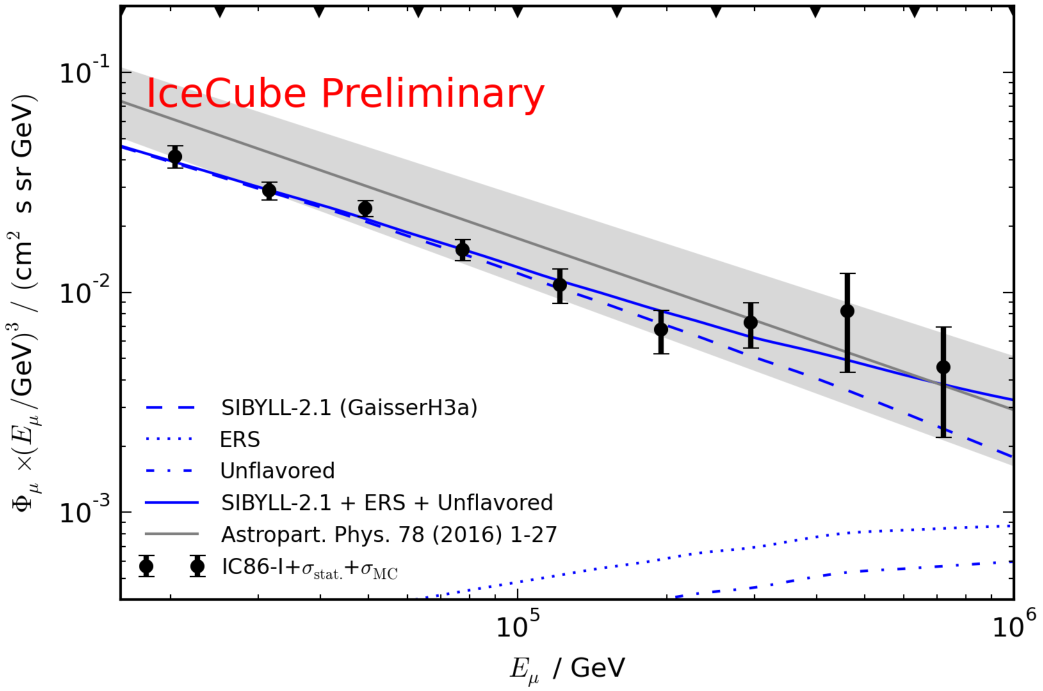
Pascal Gutjahr and Mirco Hünnefeld

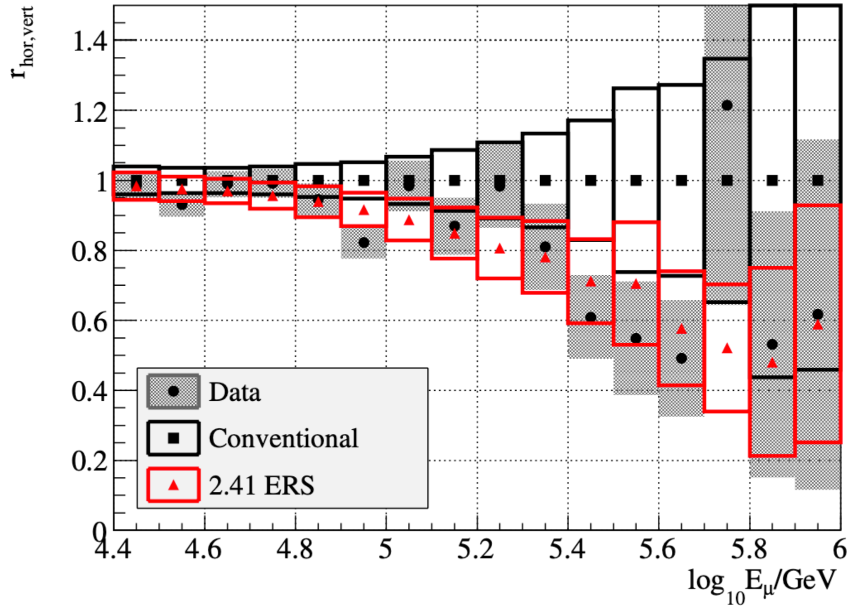
CR-WG March 1, 2024

Motivation

- Detect and measure normalization of prompt component of atmospheric muon flux
- Constrain uncertainties on hadronic interaction models at very high energies

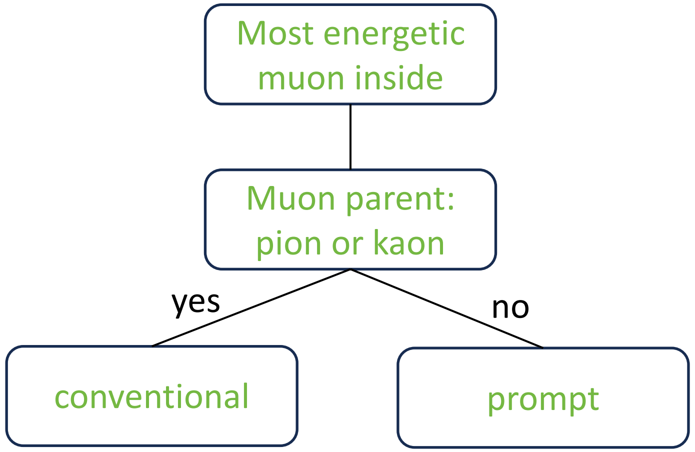
- Former analyses:
  - Leading muon analysis: limited MC statistics (by Tomasz Fuchs, https://wiki.icecube.wisc.edu/index.php/Analysis_of_Leading_Muons)
  - Characterization of the muon flux: zenith problem (by Patrick Berghaus, https://arxiv.org/abs/1506.07981)





Sample	Best Fit (ERS)	1σ Interval (90% CL)	$\sigma(\Phi_{\mathrm{prompt}} > 0)$
Uncorrected	4.93	4.05-5.87 (3.55-6.56)	9.43
Marginalized Ang. Corr.	3.19	1.64-5.48 (0.98-7.26)	3.46

- Good analyses with some challenges
- Prompt MC truth not available -> new CORSIKA simulations



Overview

Final goals

- ❑ Measure normalization of the atmospheric prompt muon flux
- ❑ Unfold muon energy spectrum

Steps for normalization measurement

- ☑ Verify CORSIKA extended history simulations
- ☑ Tag prompt muons
- ☑ Comparisons to MCEq
- ☑ Set up preliminary analysis chain
- ★ Reconstruct muon energy and direction
- ★ Data/MC comparisons
- ❑ Include systematics
- ❑ Run full statistics simulation

Additional steps for unfolding

- ❑ Effective area

Terminology

- Muon bundle: all muons in a bundle
- Leading muon: most energetic muon in a muon bundle
- Single muon: no single muons at high energies
- Prompt muon: parent is not pion or kaon

New (preliminary) CORSIKA Ehist simulation

- CORSIKA 77420
- SIBYLL 2.3d
- 600 GeV – 50 EeV
- /data/sim/IceCube/2023/generated/CORSIKA_EHISTORY/

Intention

For today

1. General update

For the upcoming collaboration meeting

1. Ask for WG reviewer
2. Storage request for large-scale CORSIKA extended history simulation

What we have done

- Reconstruction
  - Trained networks to reconstruct several properties
- Selection
  - Level 3 (L2MuonFilter + 200 TeV bundle energy cut at surface)
  - Level 4 (add reconstructions)
- Data-MC
  - Check several properties
  - Largest mismatch in z-vertex
- Forward fit
  - Test NNMFit for analysis
- Unfolding
  - Unfold event rate
  - Calculate effective area
- New simulations
  - Preparation for large scale IceProd simulation (latest software, switch some options)
- Wiki
  - Created and uploaded

Today: ★

Münster: ⏭

DNN reconstructions

Reconstructed properties

Energy

- ❑ entry_energy: Leading muon energy at the detector entry
- ❑ bundle_energy_at_entry: Muon bundle energy at the detector entry
- ❑ muon_energy_first_mctree: Leading muon energy at surface
- ❑ bundle_energy_in_mctree: Muon bundle energy at surface

Network 1

Track geometry

- ❑ Length: Propagation length of muon in the ice
- ❑ LengthInDetector: Propagation length of muon in the detector
- ❑ center_pos_x: Closest x position of muon to center of the detector
- ❑ center_pos_y: Closest y position of muon to center of the detector
- ❑ center_pos_z: Closest z position of muon to center of the detector
- ❑ center_pos_t: Time of closest approach to the center of the detector
- ❑ entry_pos_x: x position of muon at the detector entry
- ❑ entry_pos_y: y position of muon at the detector entry
- ❑ entry_pos_z: z position of muon at the detector entry
- ❑ entry_pos_t: Time of muon at the detector entry

Network 2

Direction

- ❑ zenith: Zenith angle of muon
- ❑ azimuth: Azimuth angle of muon

Network 3

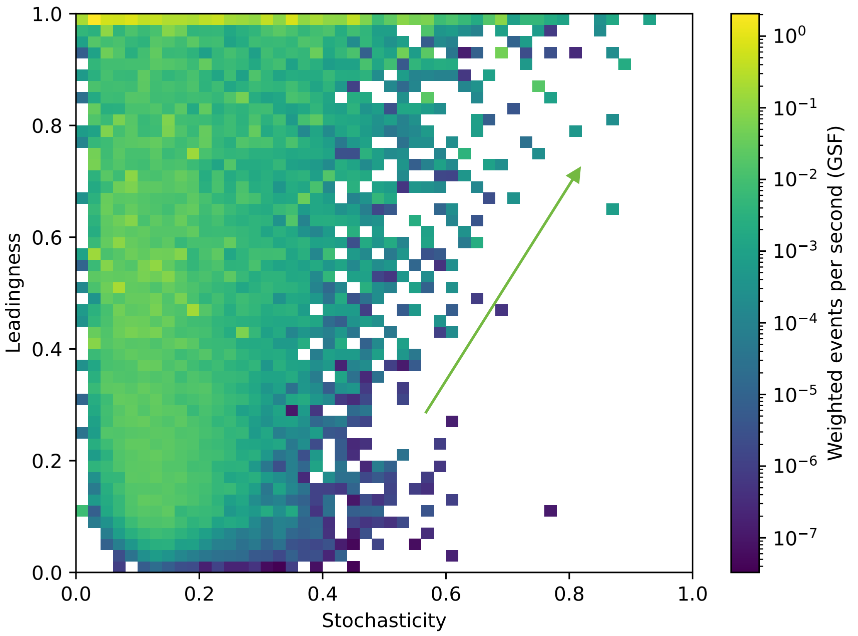
DNN Reconstructions

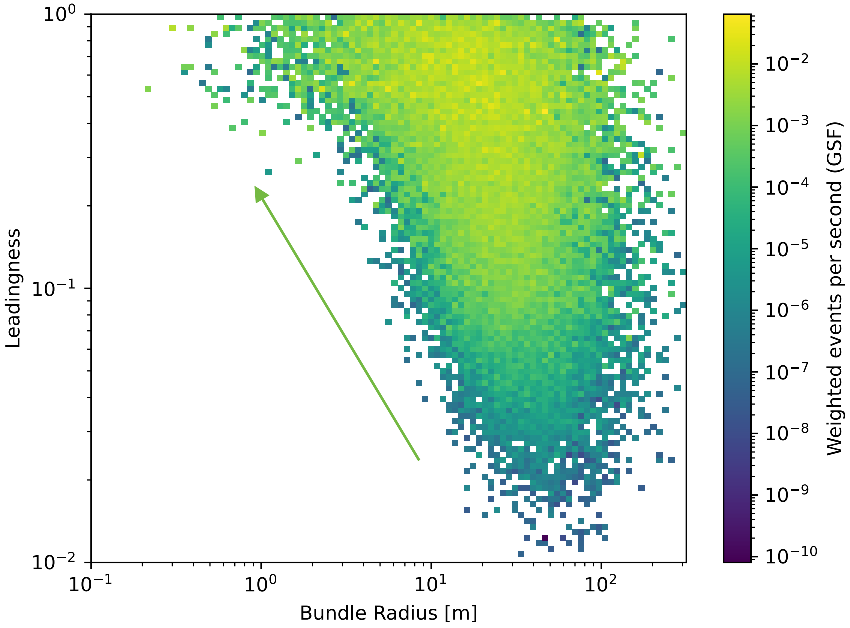
Reconstruct

- ❑ Energy
- ❑ Track geometry
- ❑ Direction

Physics motivation

- Muons lose energy stochastically
  - High leadingness: energy depositions are dominated by large stochastic losses
  - Low leadingness: stochastic losses sum up and appear continuously
- High energies: forward production
  - High leadingness: small bundle radius
  - Low leadingness: larger bundle radius

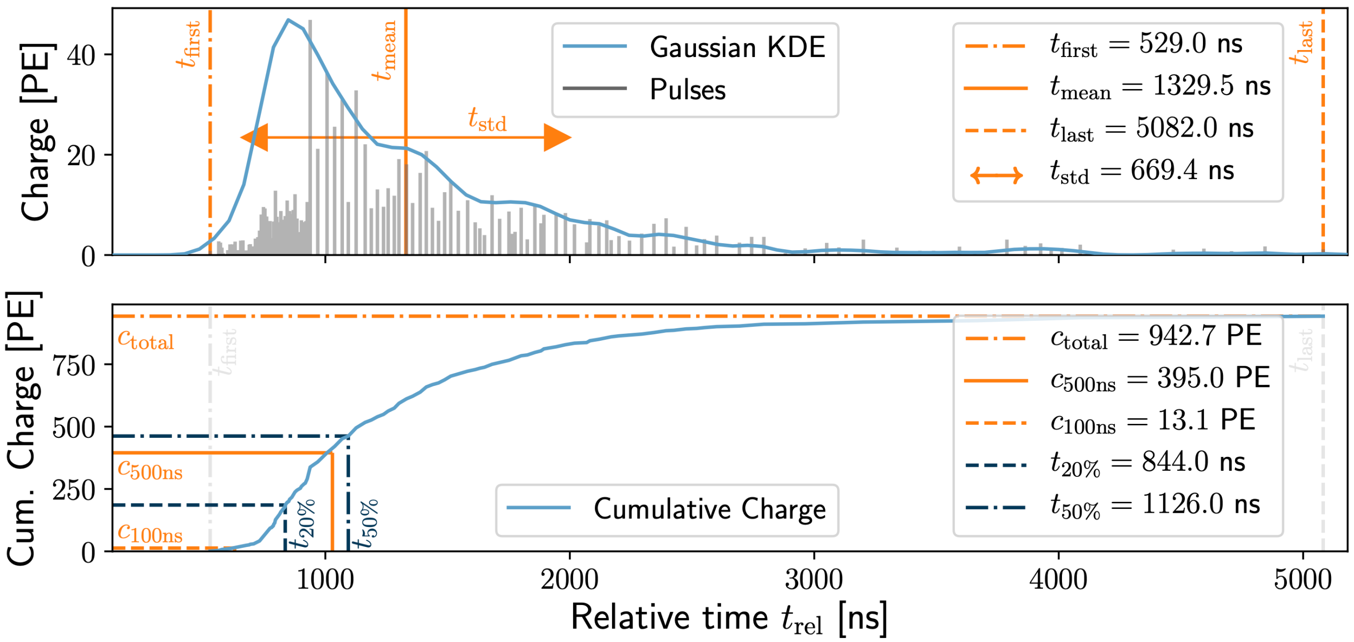




pascal.gutjahr@tu-dortmund.de

Input data per DOM

10.1088/1748-0221/16/07/P07041



3 inputs

- c_{total}: Total charge
  - Sum of charge
- t_{first}: Relative time of first pulse
  - Relative to total time offset, calculated as the charge weighted mean time of all pulses
- t_{std}: Standard deviation of first pulse
  - Charge weighted standard deviation of pulse times relative to total time offset

9 inputs

- t_{last}: Relative time of last pulse
  - Relative to total time offset, calculated as the charge weighted mean time of all pulses
- $t_{20\%}$: Relative time of 20% charge
  - Relative to total time offset, calculated as the charge weighted mean time of all pulses
- $t_{50\%}$: Relative time of 50% charge
  - Relative to total time offset, calculated as the charge weighted mean time of all pulses
- t_{mean}: Mean time
  - Charge weighted mean time of all pulses relative to total time offset
- c_{500ns}: Charge at 500ns
  - Sum of charge after 500ns
- c_{100ns}: Charge at 100ns
  - Sum of charge after 100ns

Input pulses

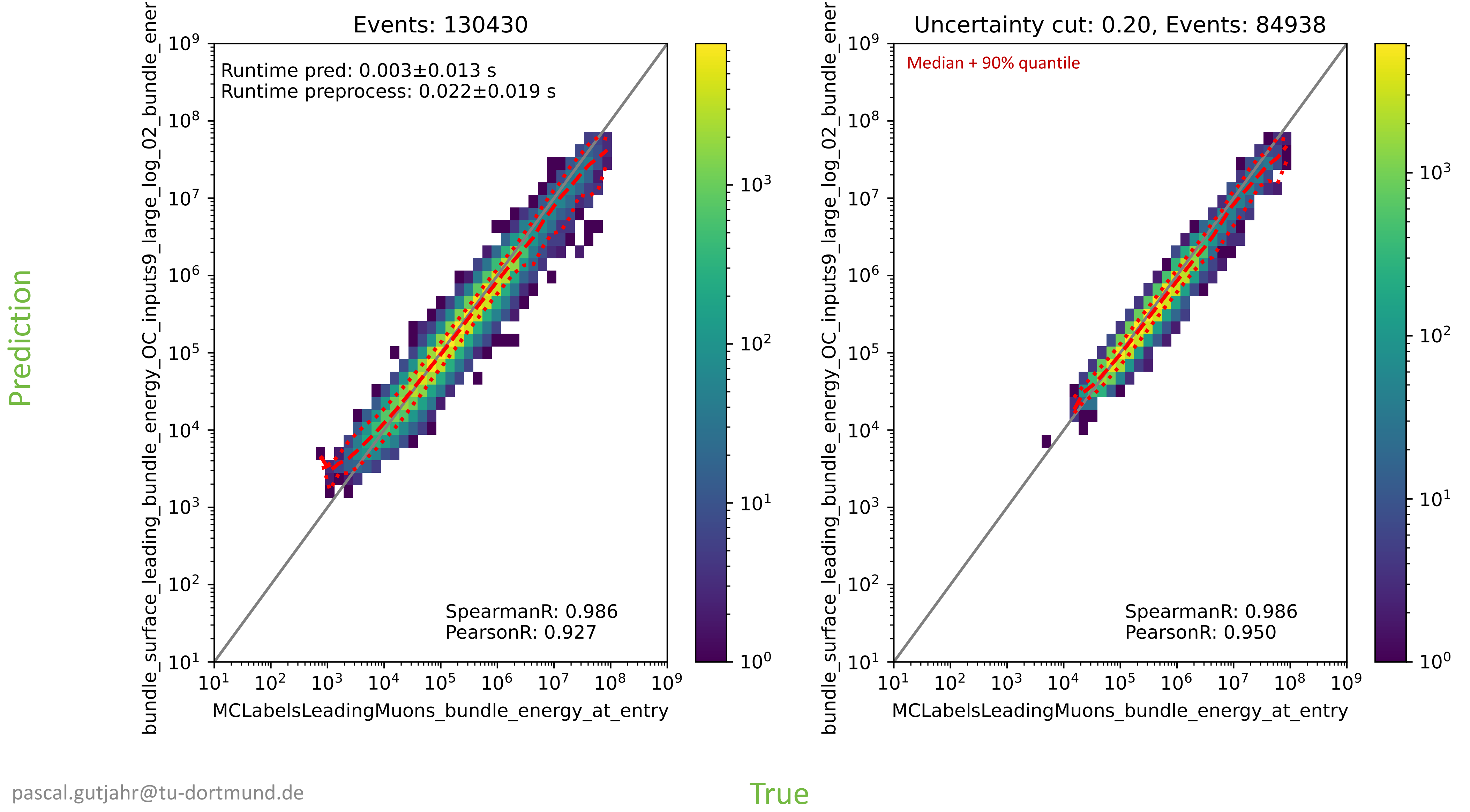
- SplitInIceDSTPulses
- SplitInIceDSTPulsesTWCleaning6000ns
- (DNN framework performs an internal cleaning)

Training datasets

- 20904
- 21962
- 22020
- 22187

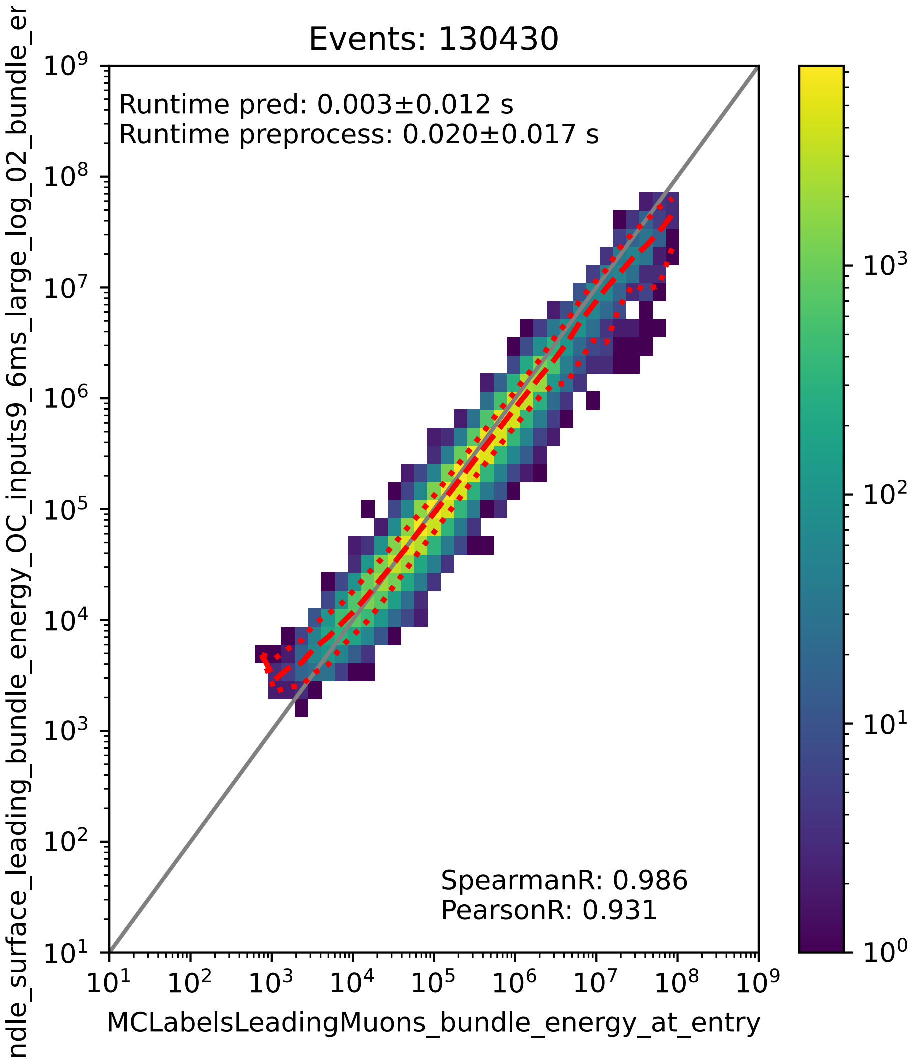
Network evaluation - Energy

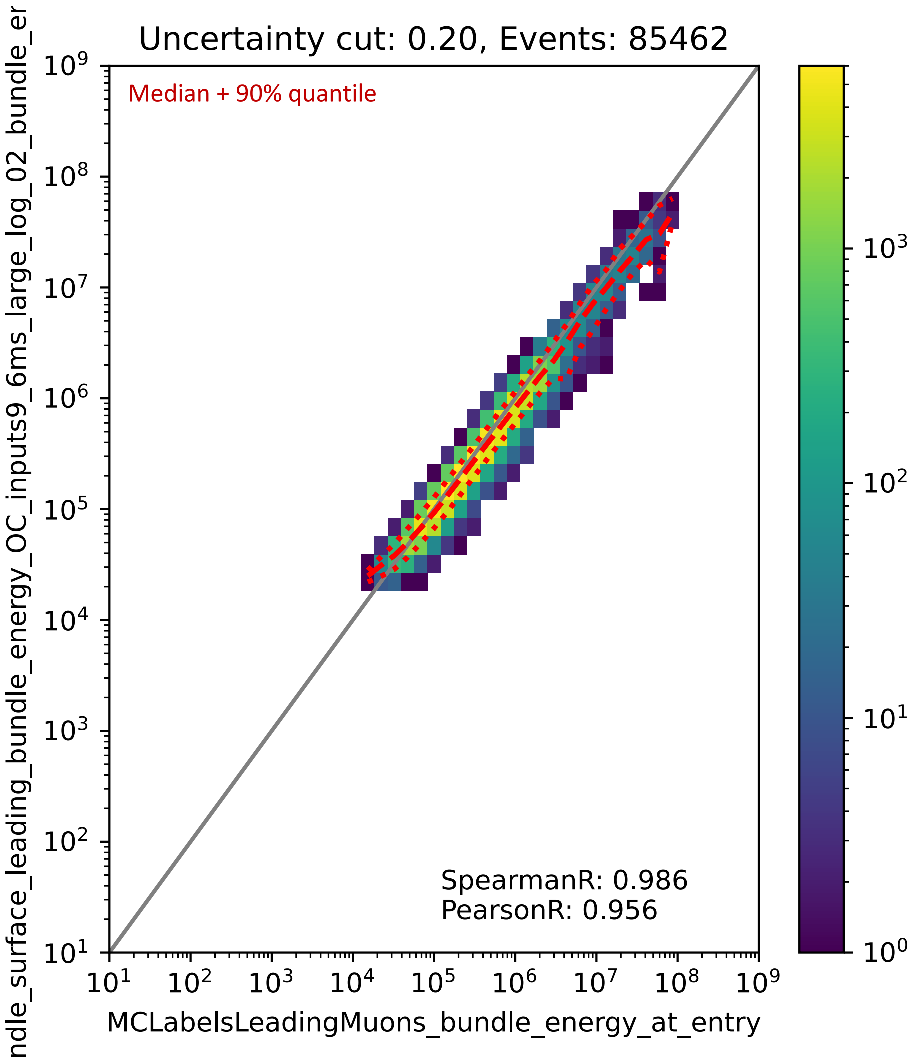
Bundle energy at entry – internal cleaned pulses



Bundle energy at entry – 6 ms cleaned pulses

Prediction

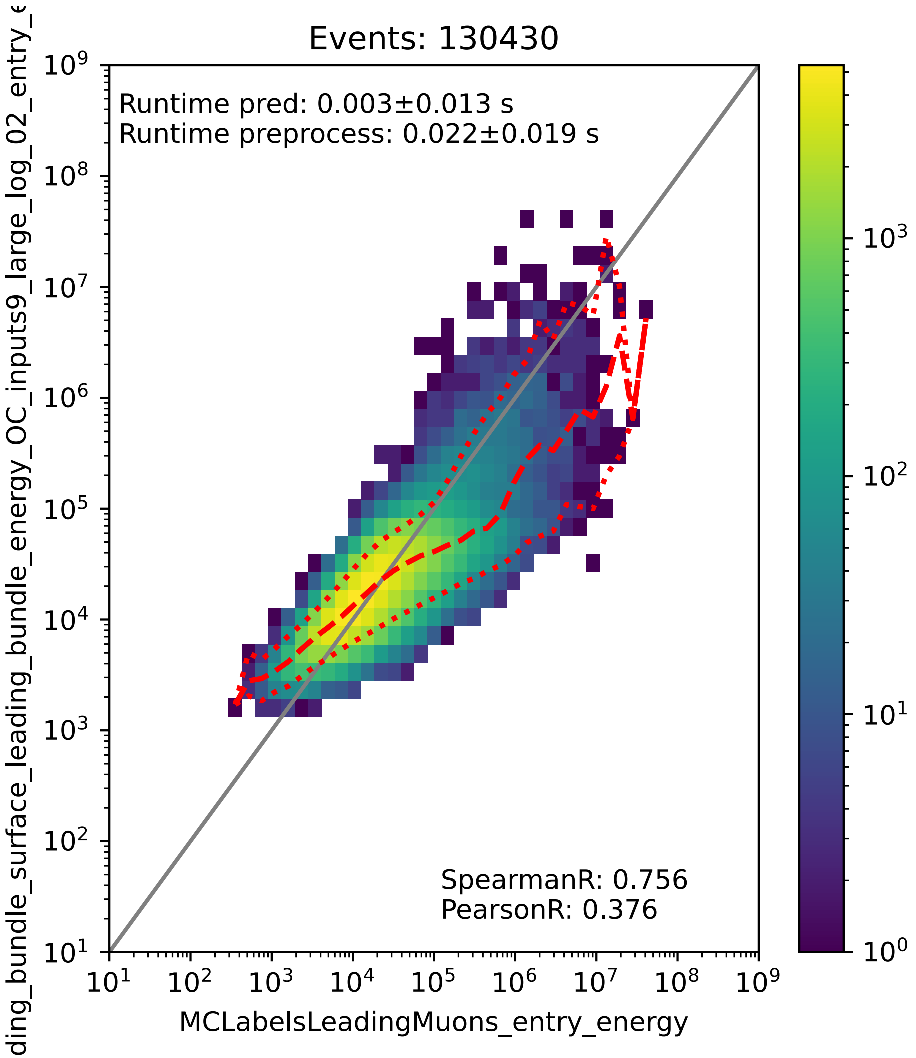


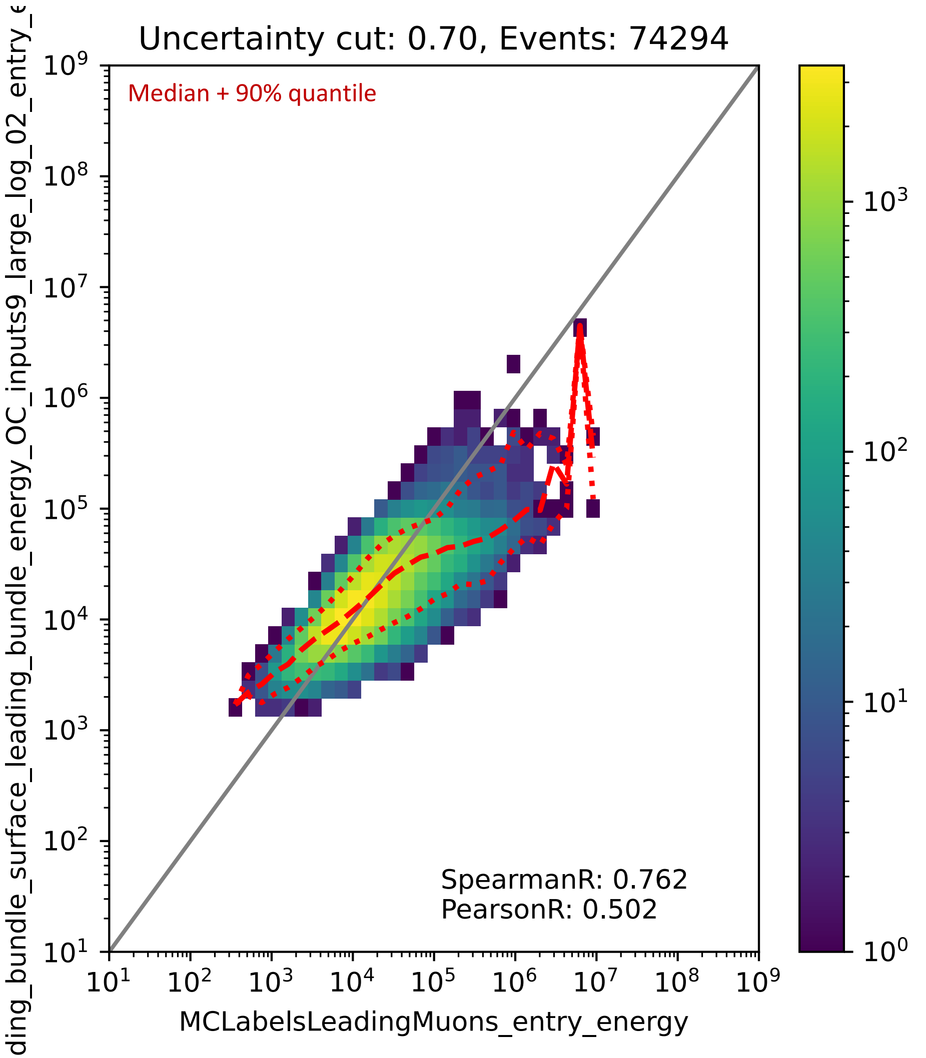


True

Leading muon energy at entry – internal cleaned pulses

Prediction

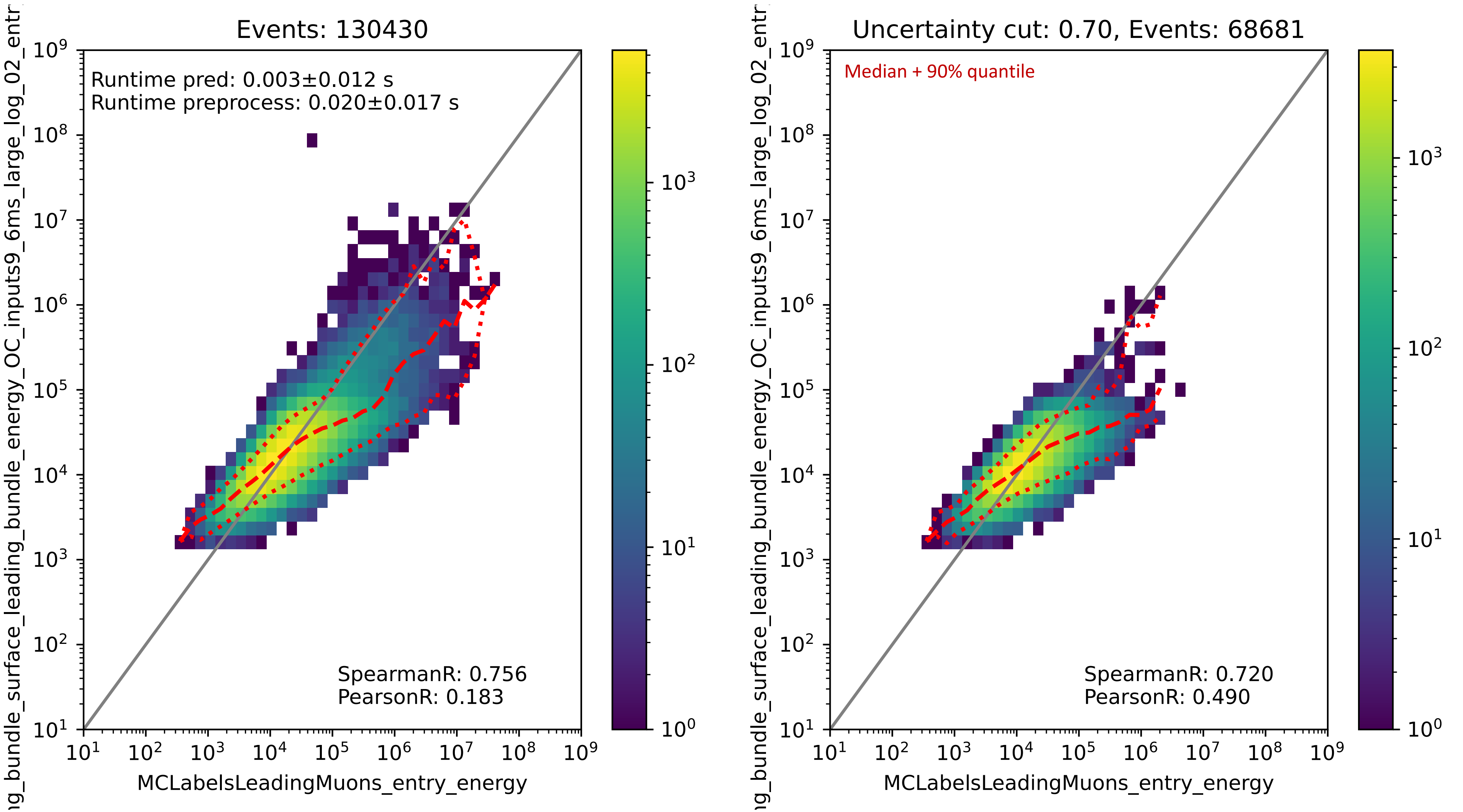




True

Leading muon energy at entry – 6 ms cleaned pulses

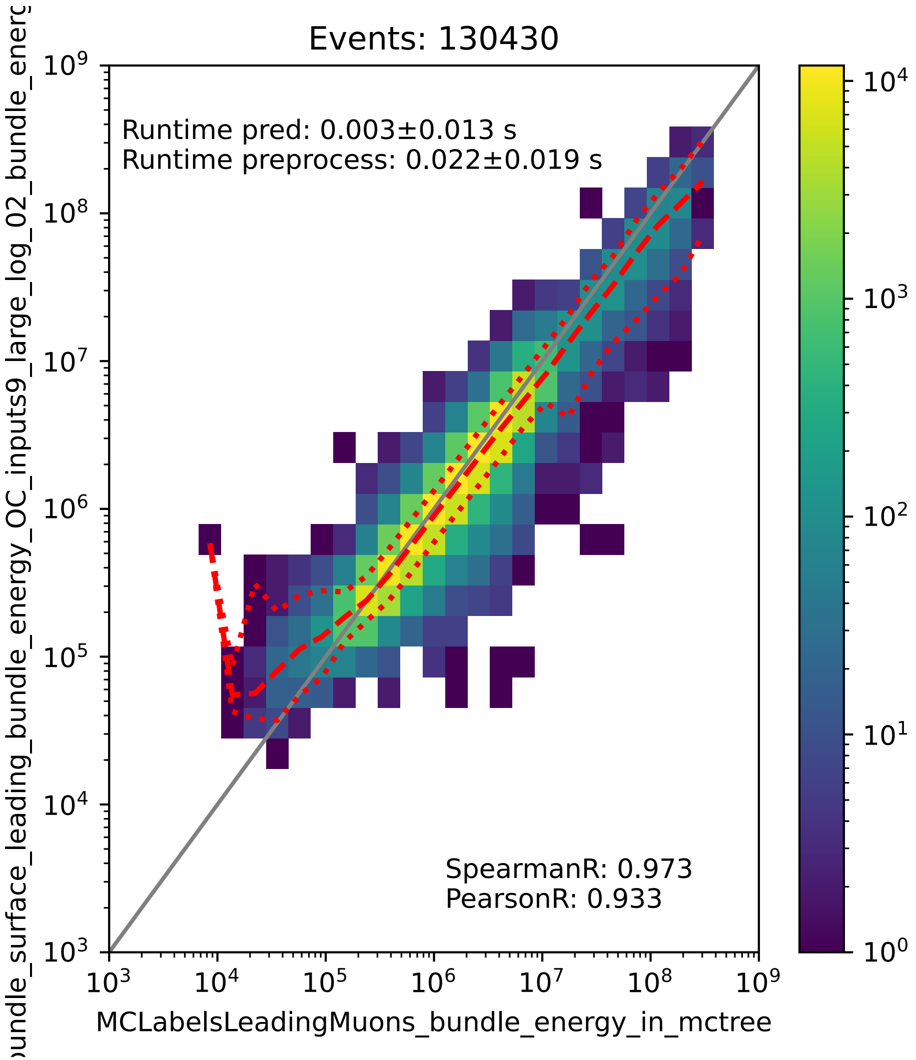
Prediction

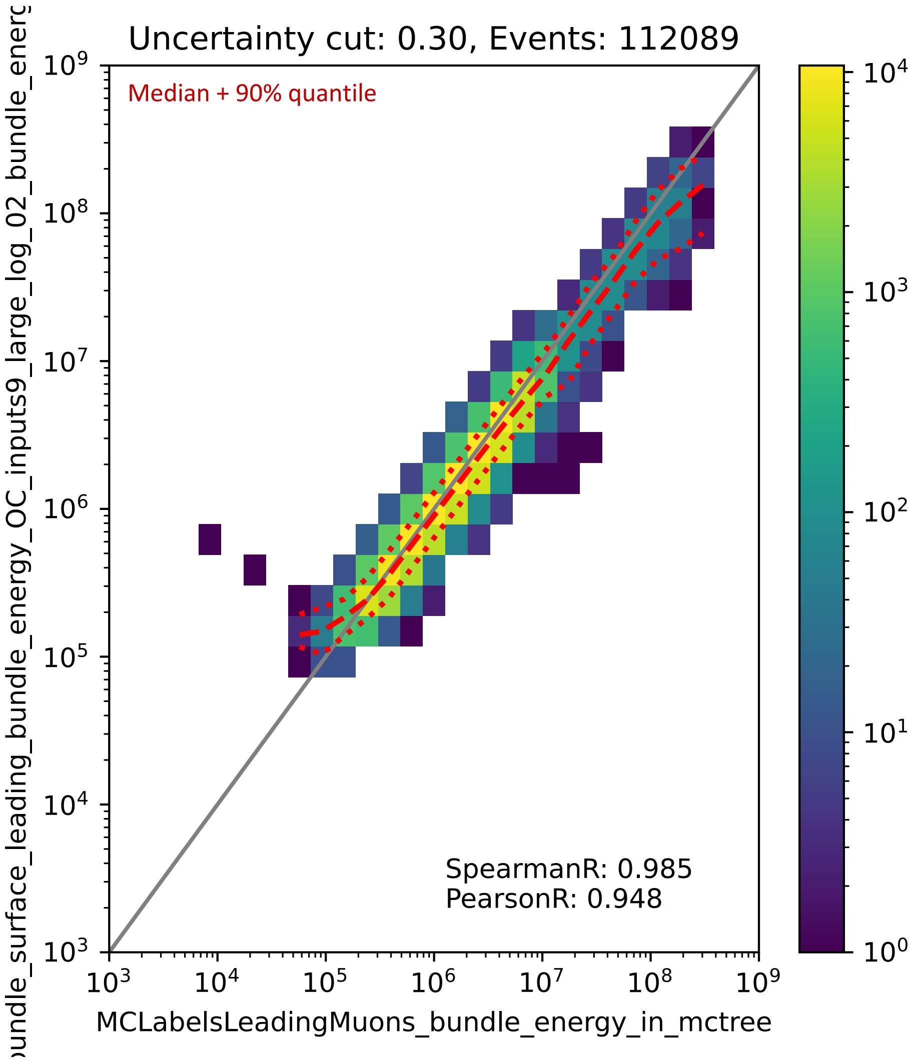


True

Bundle energy at surface – internal cleaned pulses

Prediction

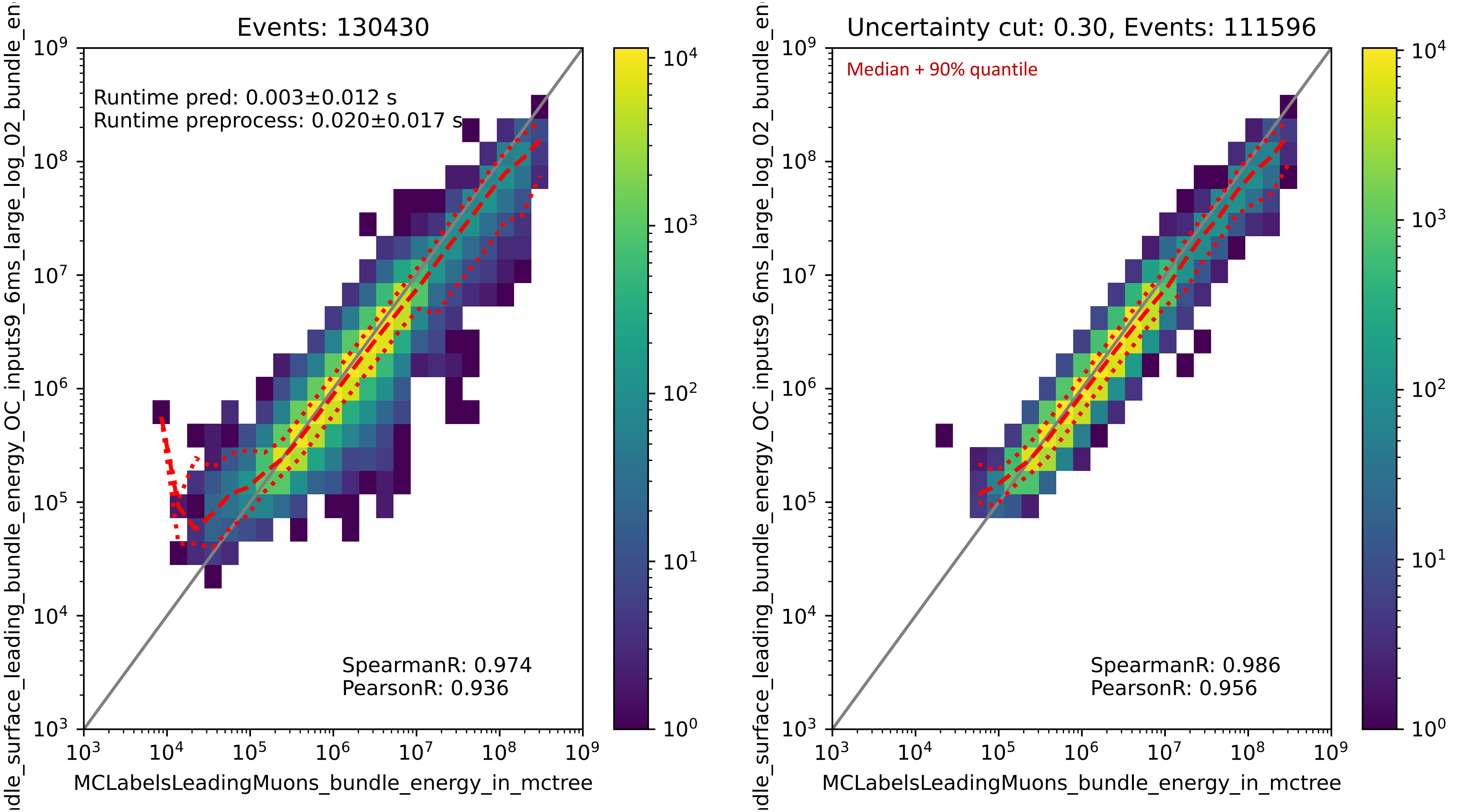




True

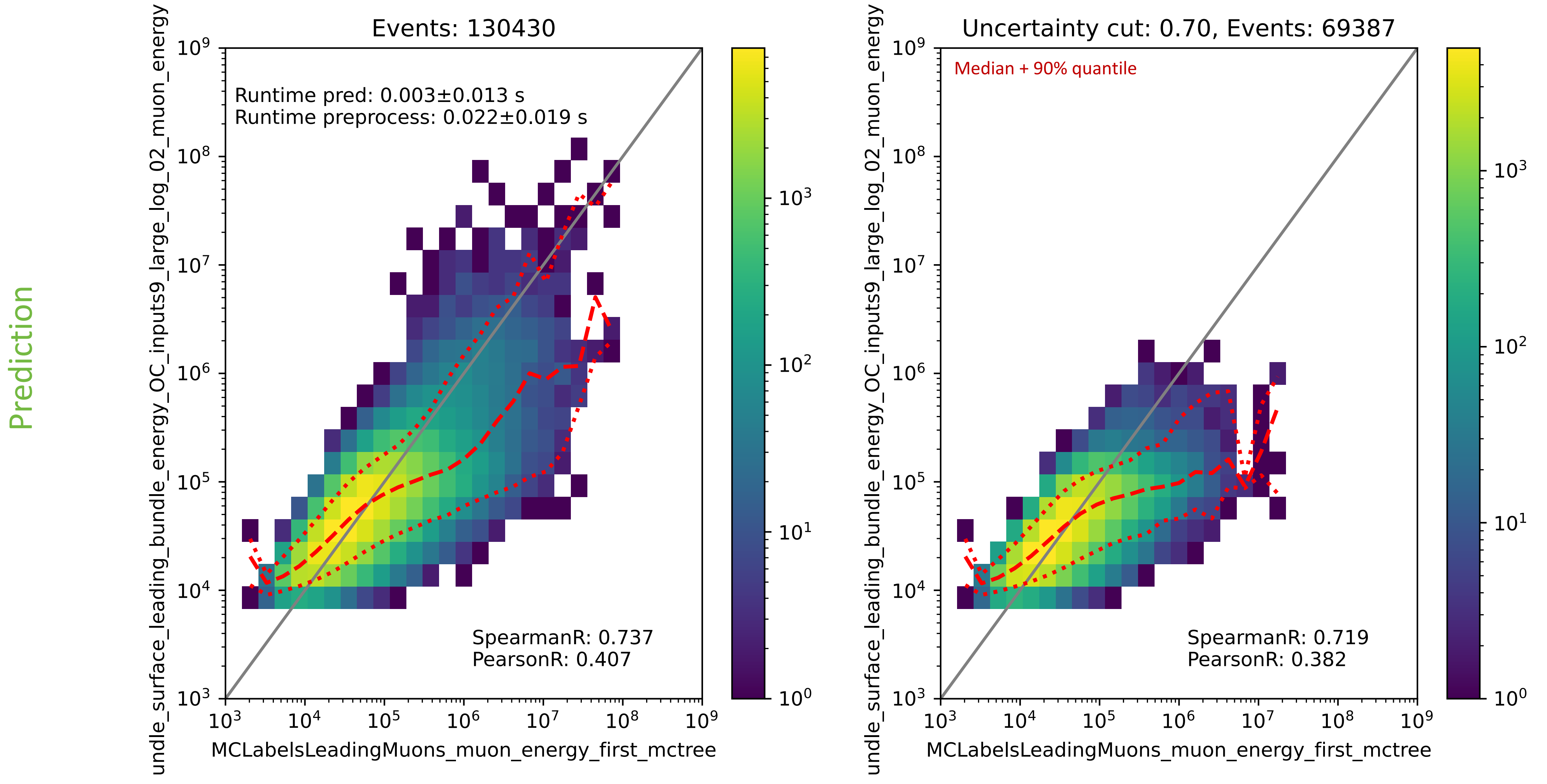
Bundle energy at surface – 6 ms cleaned pulses

Prediction



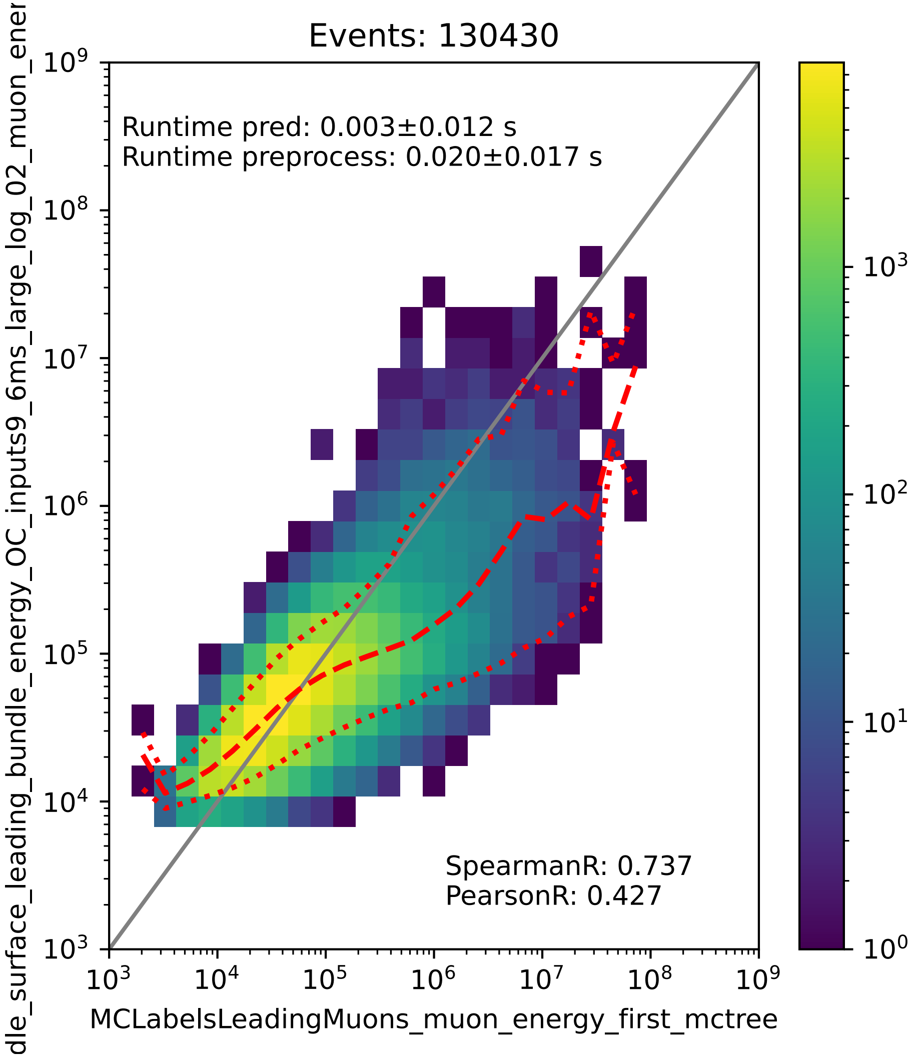
True

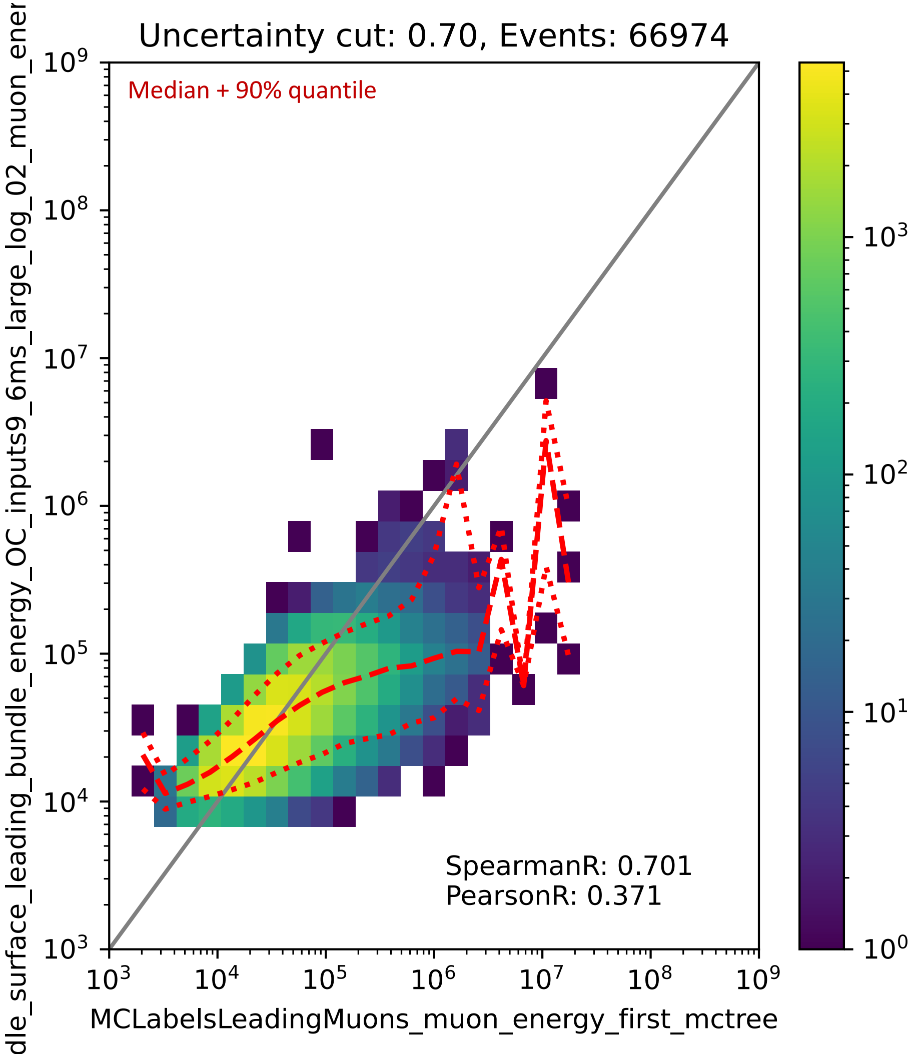
Leading muon energy at surface – internal cleaned pulses



Leading muon energy at surface – 6 ms cleaned pulses

Prediction

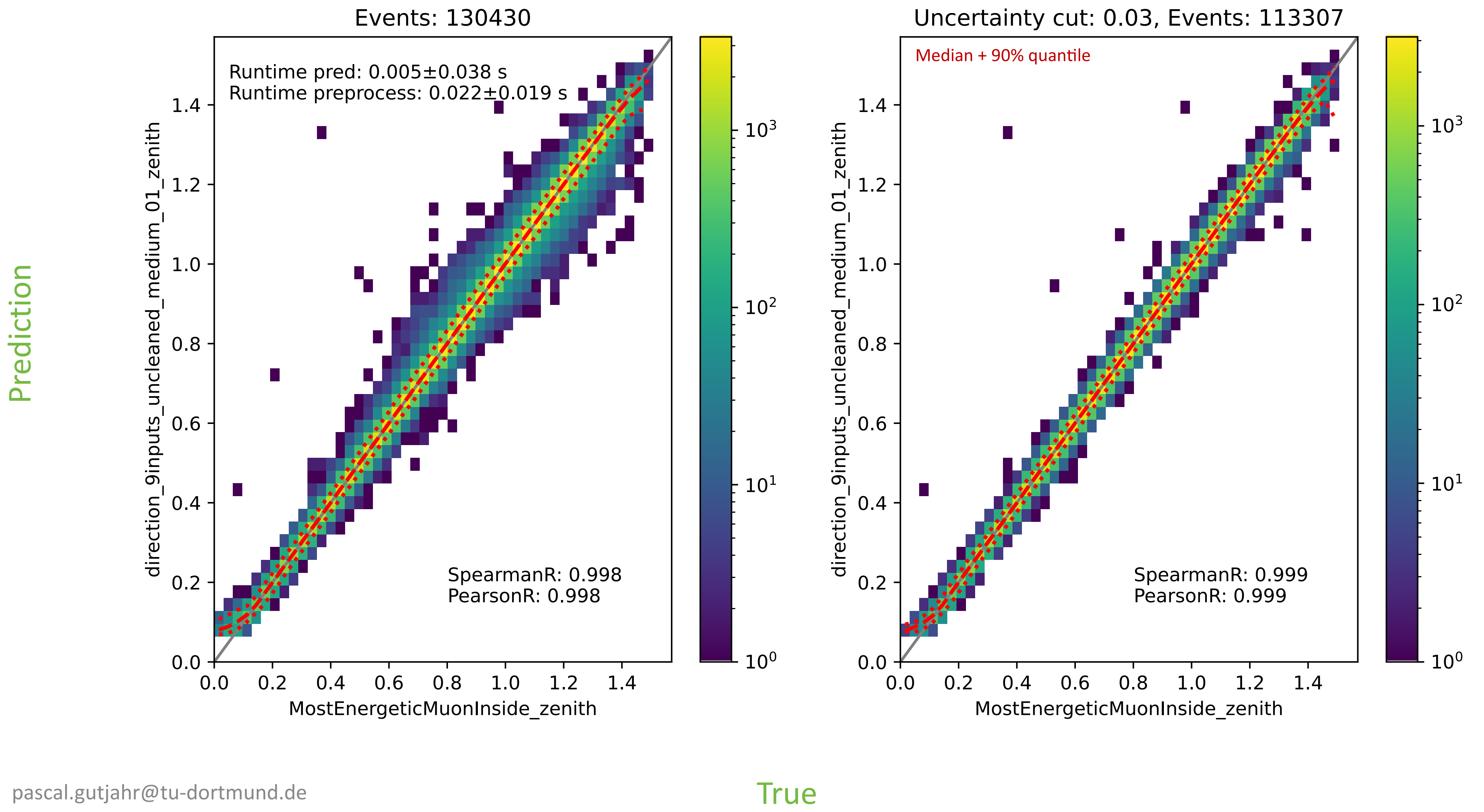




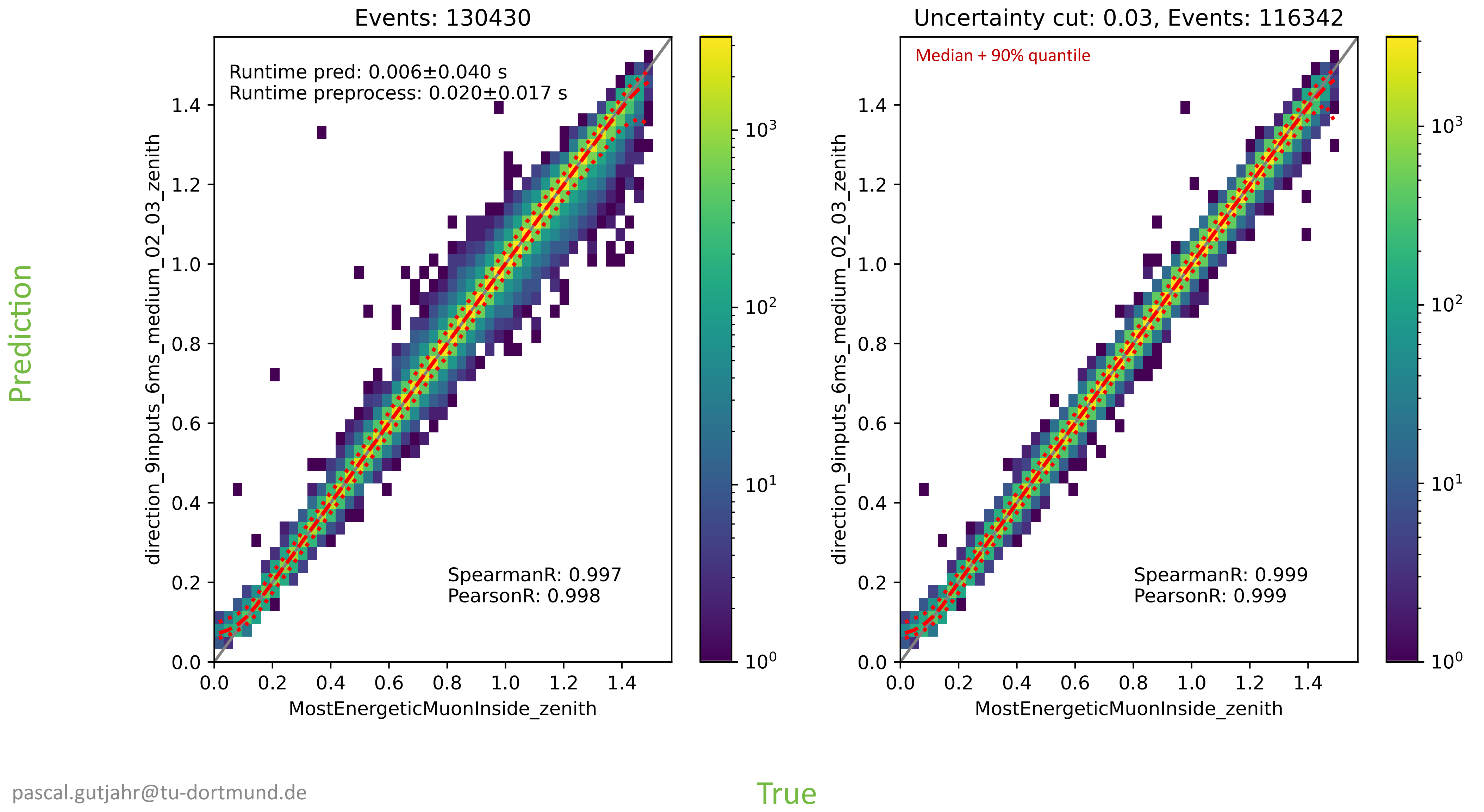
True

Network evaluation - Direction

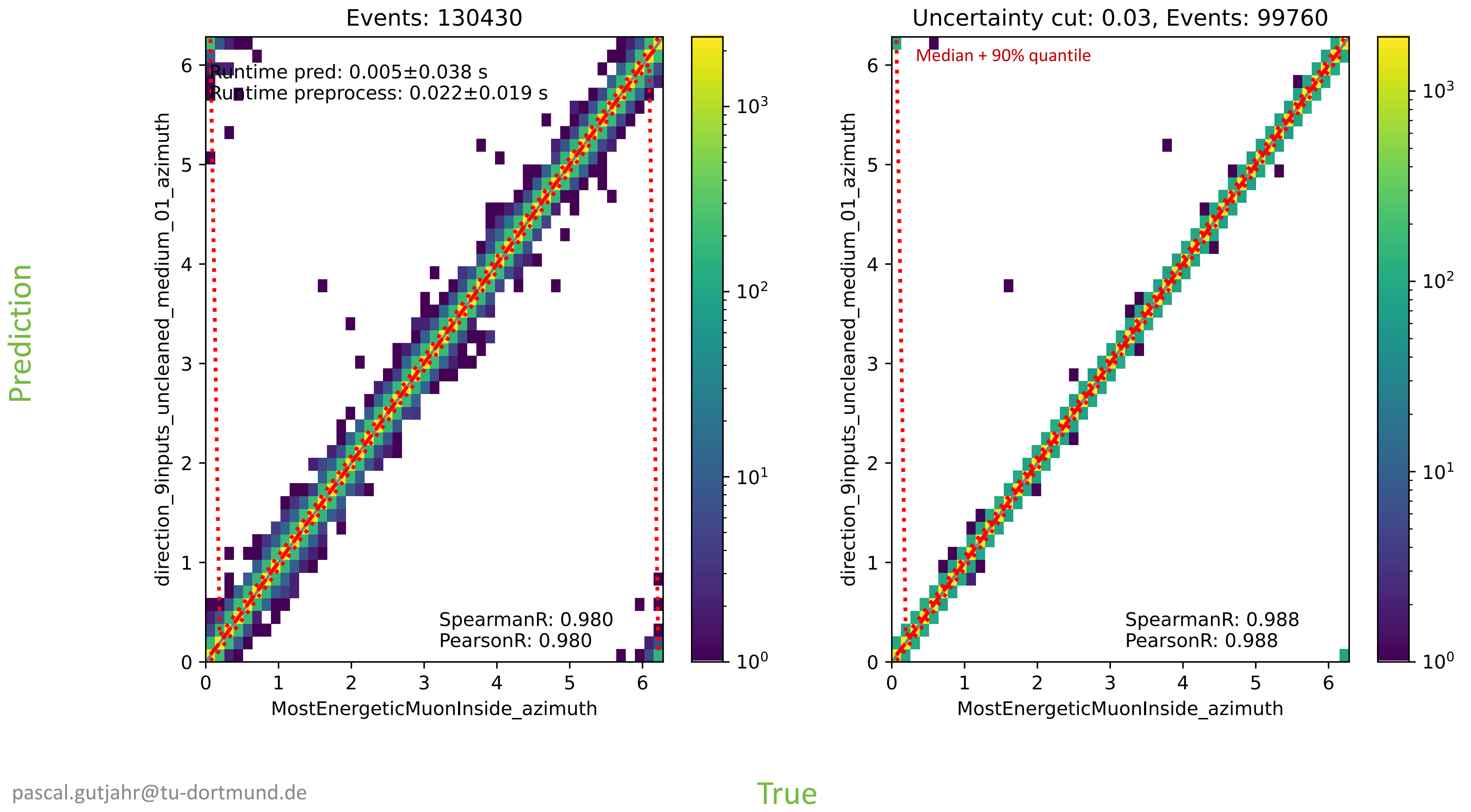
Zenith – internal cleaned pulses



Zenith – 6 ms cleaned pulses



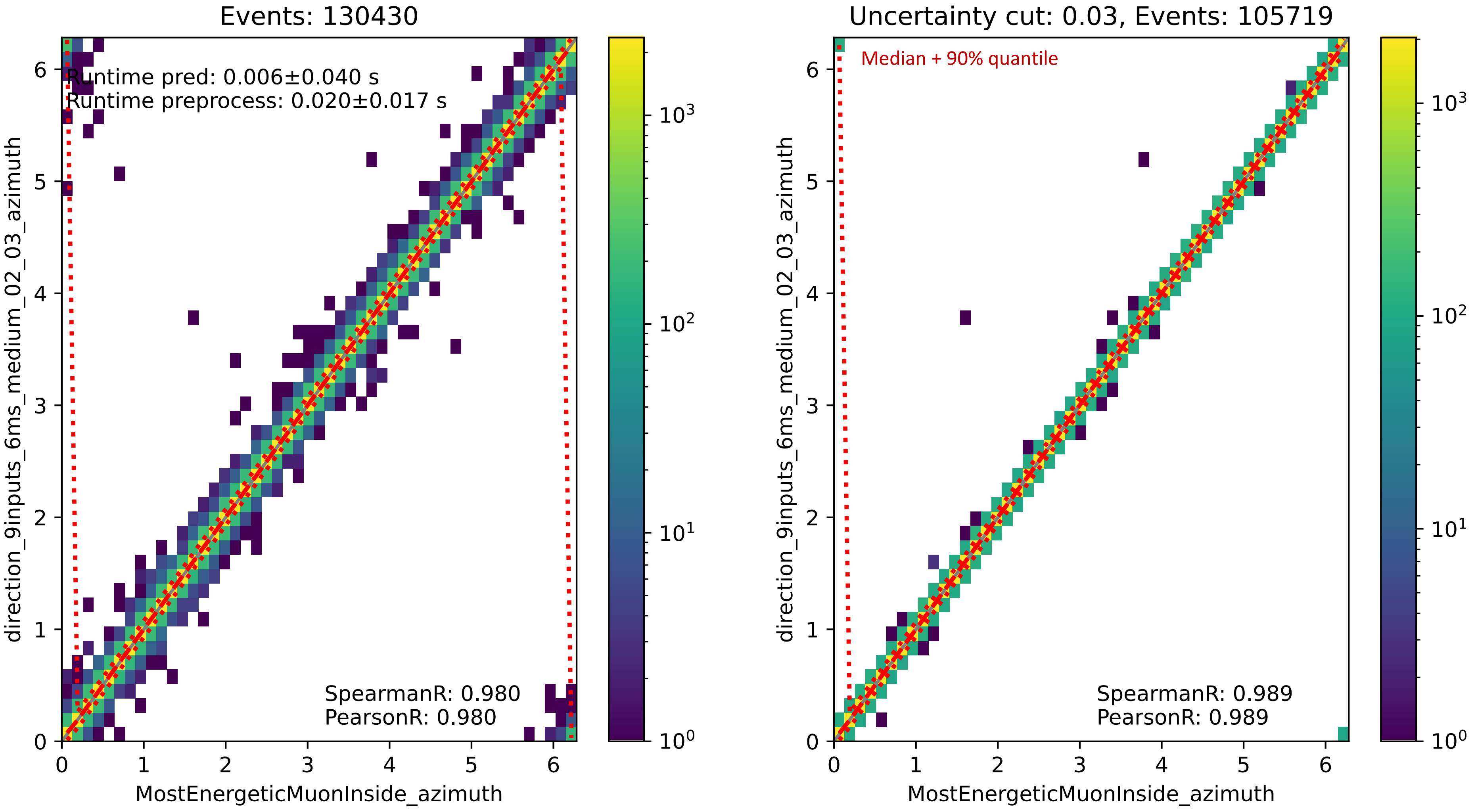
Azimuth – internal cleaned pulses



tu dortmund university

IceCube Neutrino Observatory

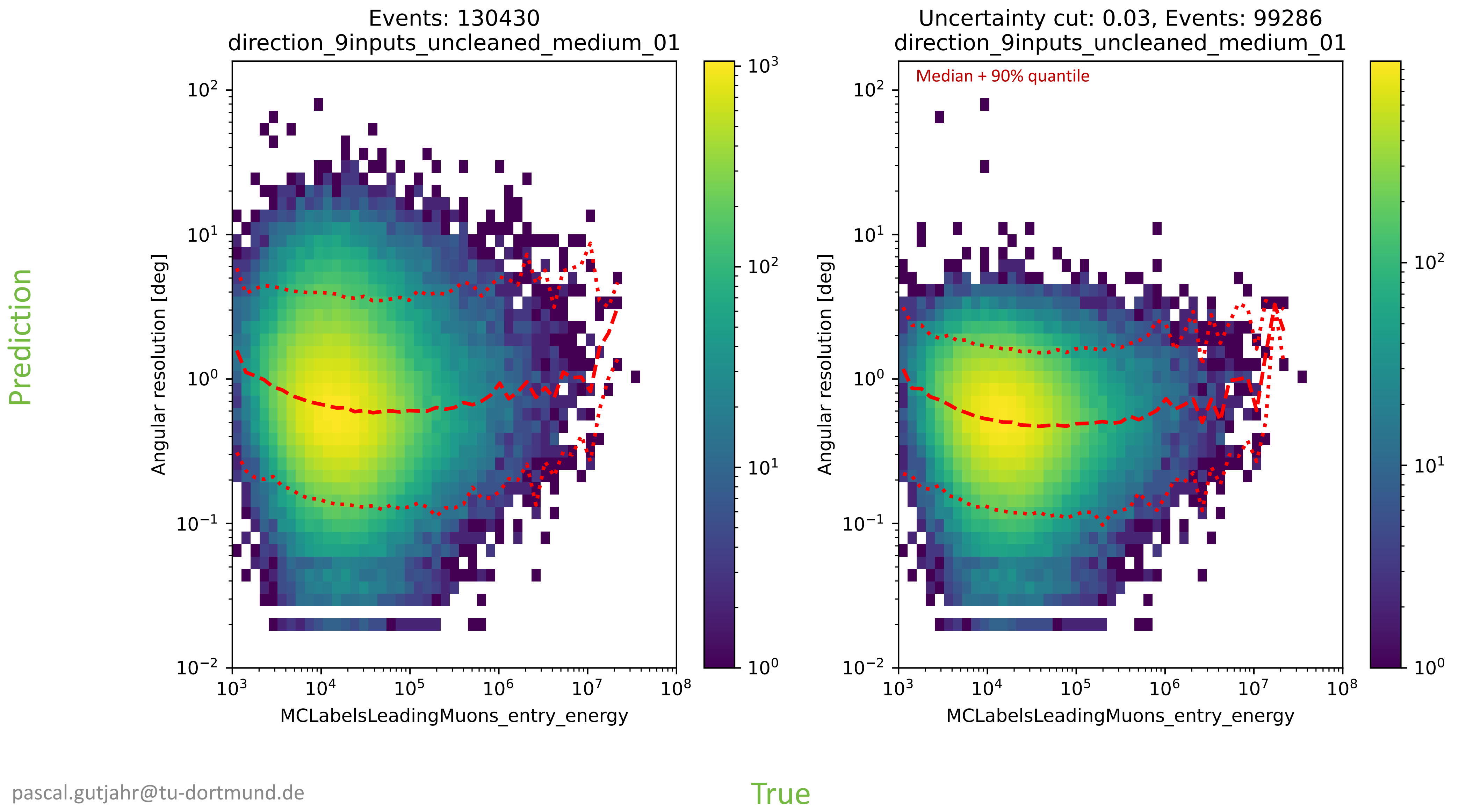
Azimuth – 6 ms cleaned pulses



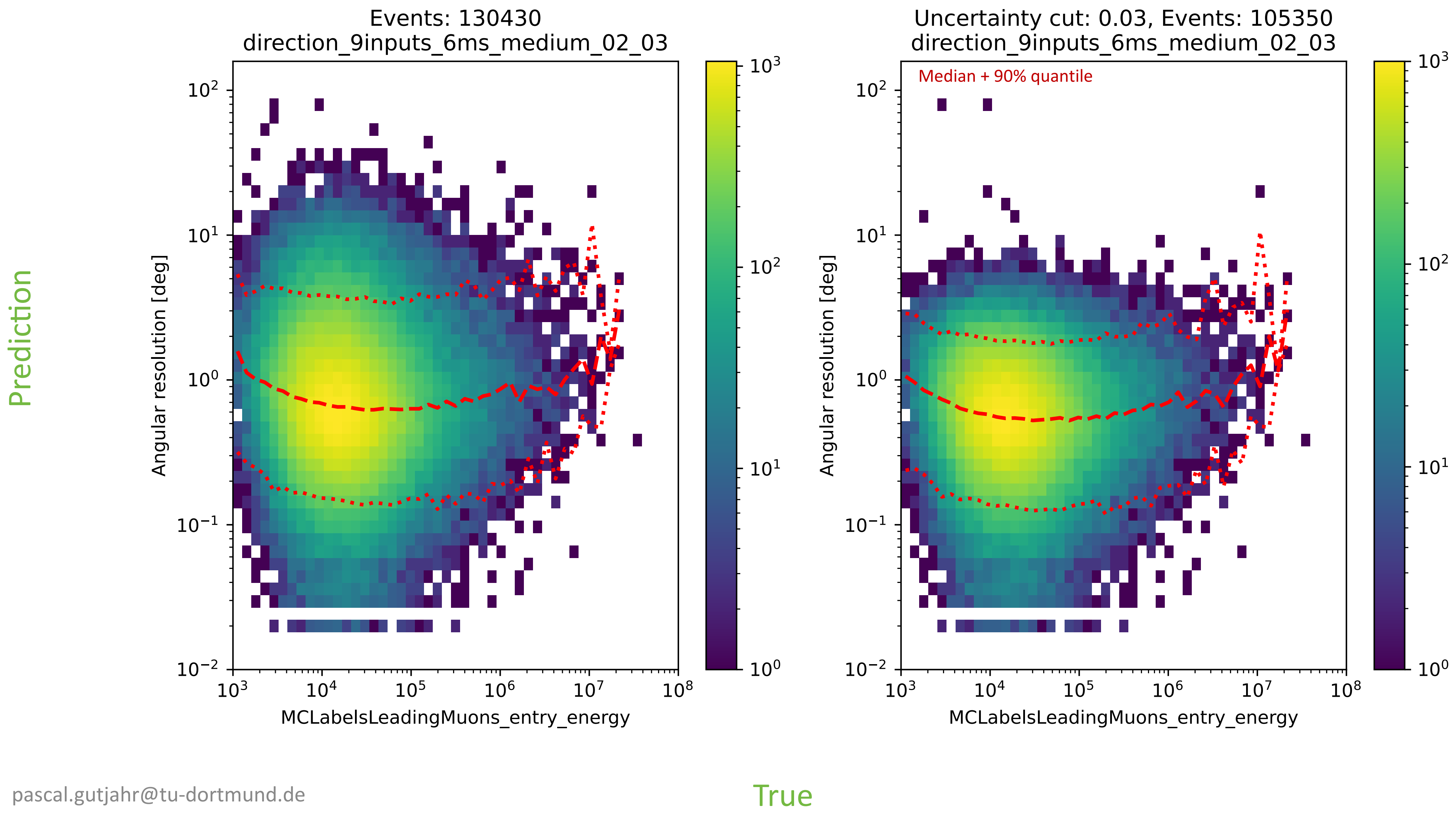
Prediction

True

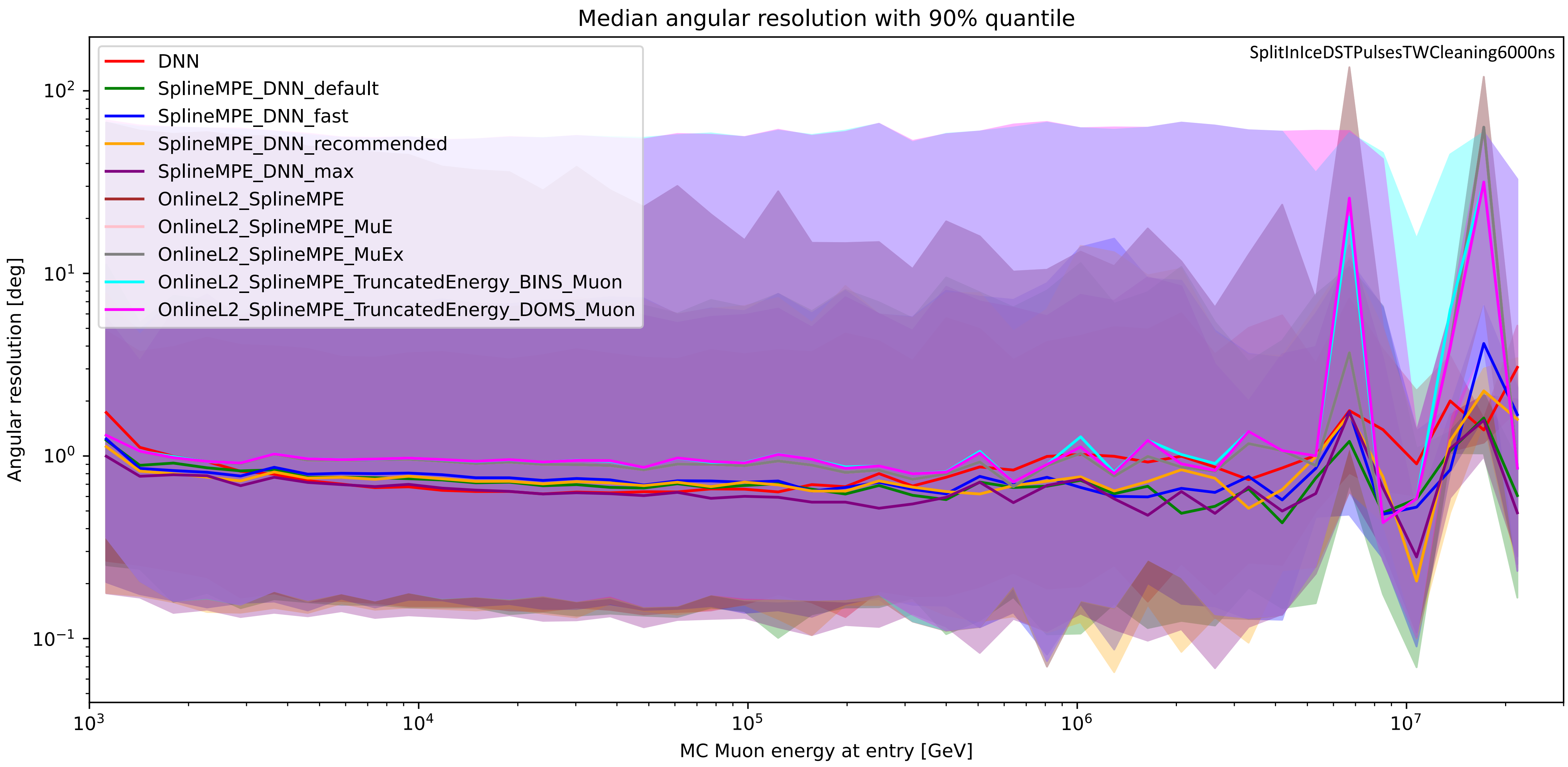
Angular resolution – internal cleaned pulses



Angular resolution – 6 ms cleaned pulses



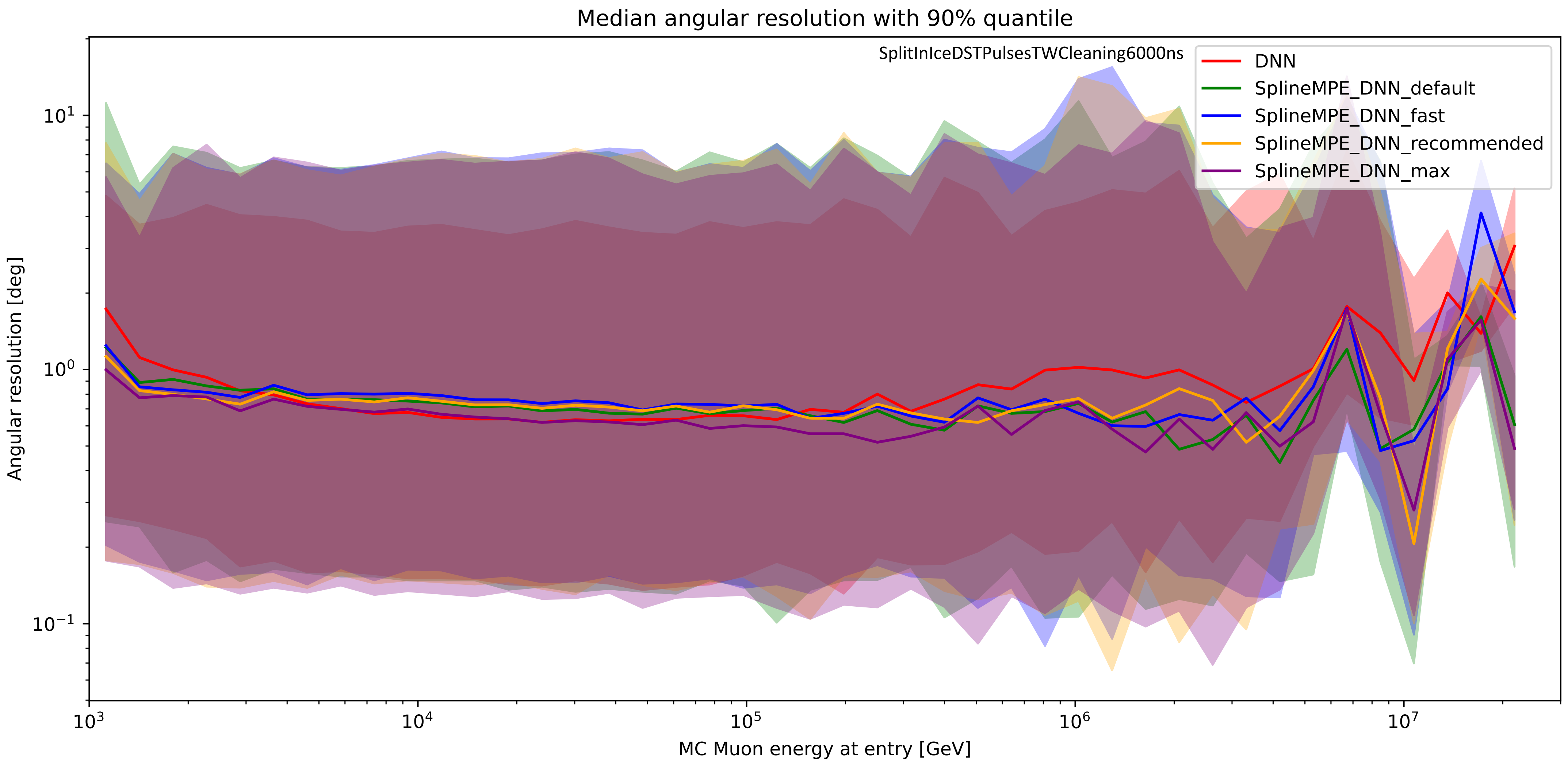
SplineMPE – DNN and conventional seeds



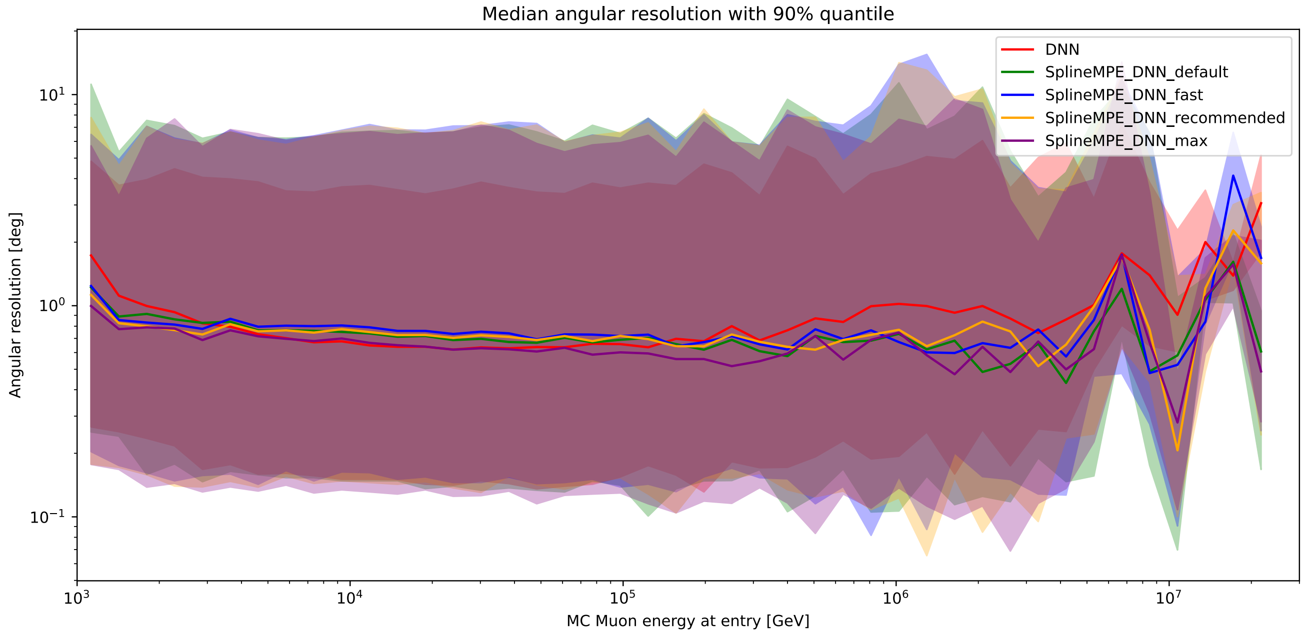
tu dortmund university

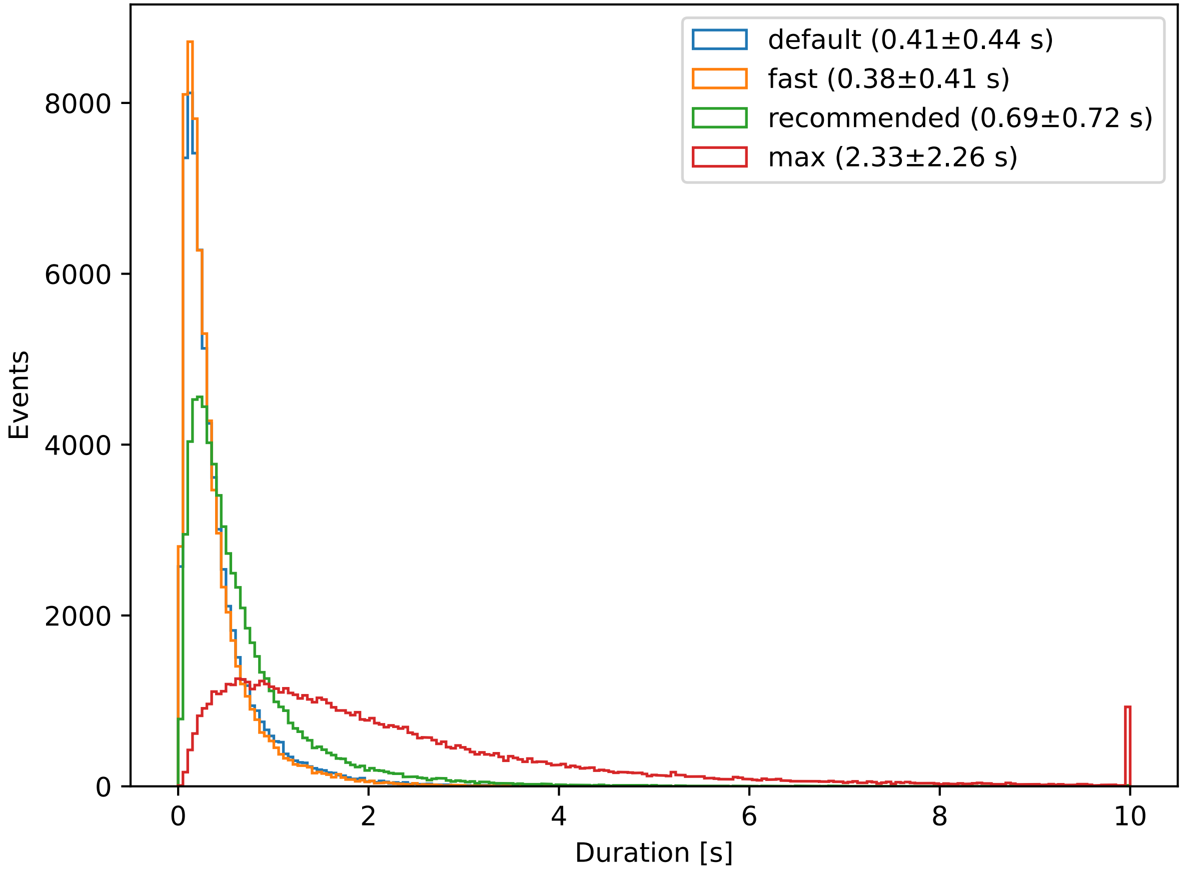
ICECUBE NEUTRINO OBSERVATORY

SplineMPE – DNN seeds



SplineMPE – duration per event





- Only small improvement at energies around 1 PeV
- Countours are larger
- Additional runtime

-> Use only DNN reconstruction, since we do not need the best angular resolution

Network evaluation – Track geometry

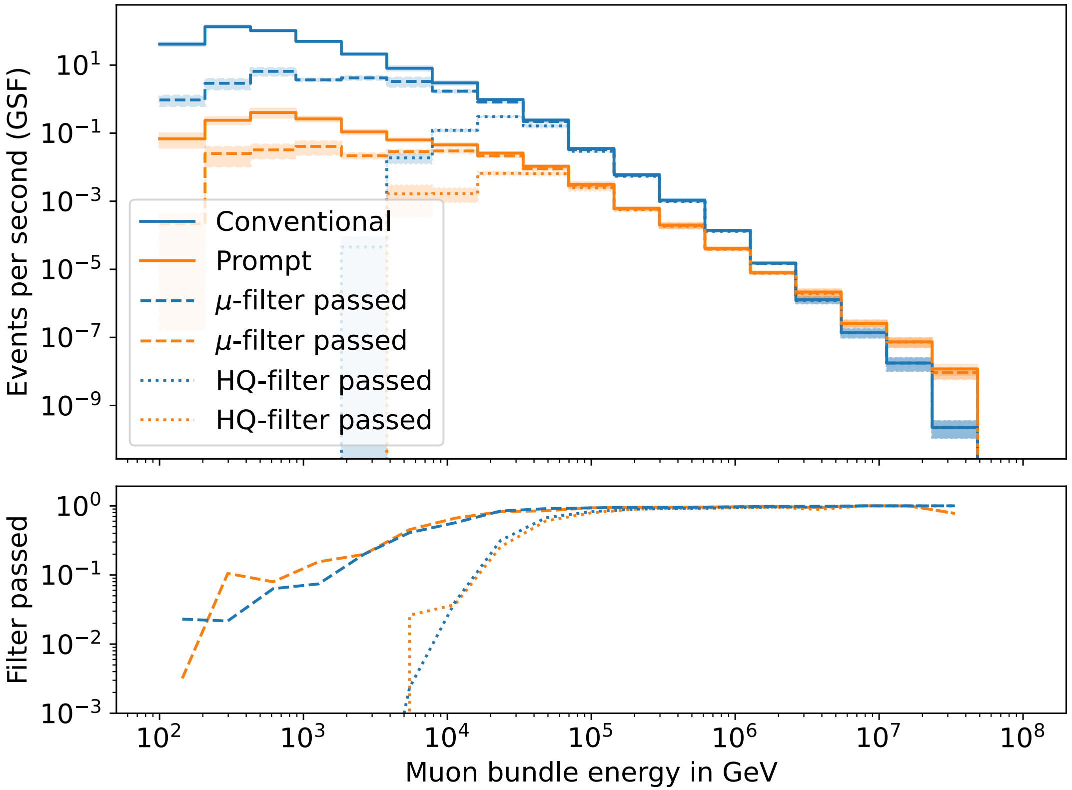
The evaluation of the track geometry works fine and can be found in the wiki

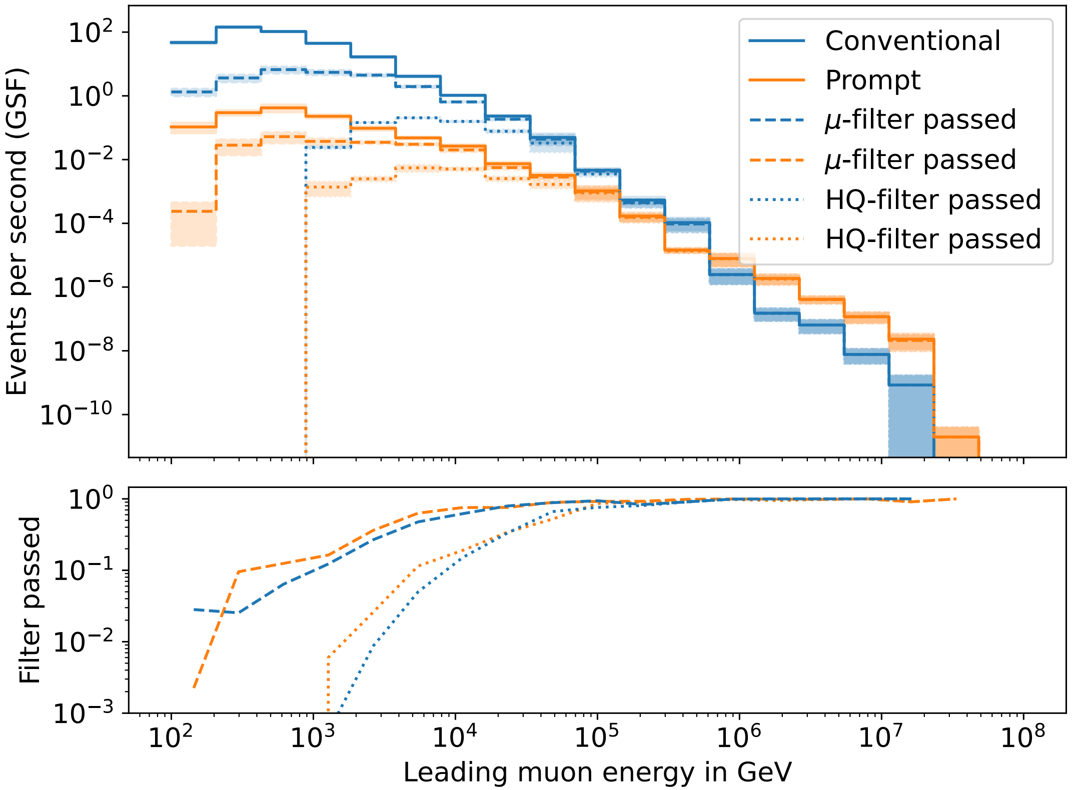
Selection – Level 3

1. L2 muon filter
2. 200 TeV bundle energy cut at surface
3. Add labels

L2 filter

Fraction rejected events	All energies	Leading energy > 10 TeV	Leading energy > 100 TeV
MuonFilter	0.93	0.28	0.06
HQFilter	0.99	0.74	0.18

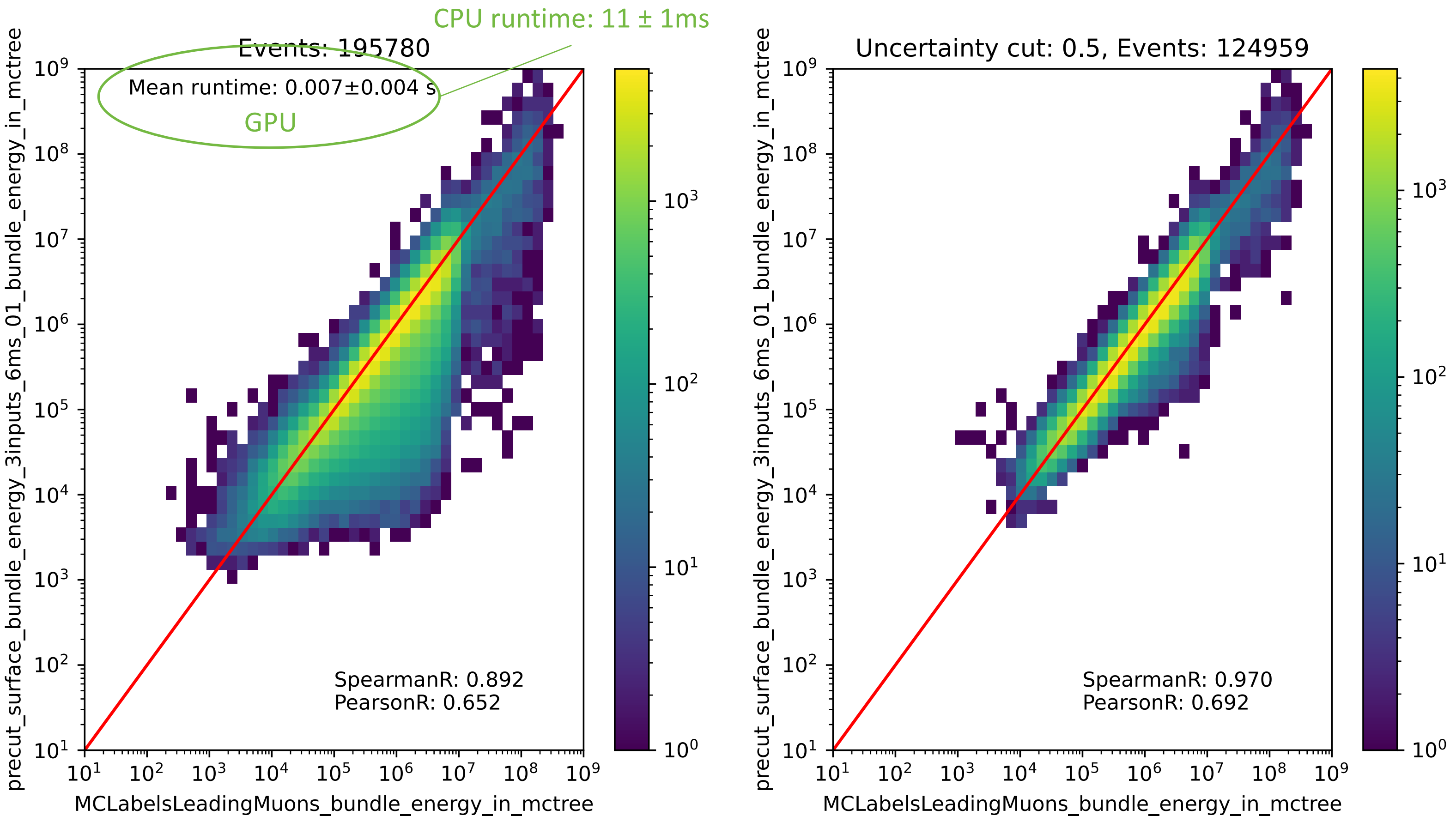




- L2 rate: 369.43 Hz
- Rate after muon filter: 24.62 Hz
- ➢ Choose muon filter to remove large amount of statistics at low energies

Bundle energy cut

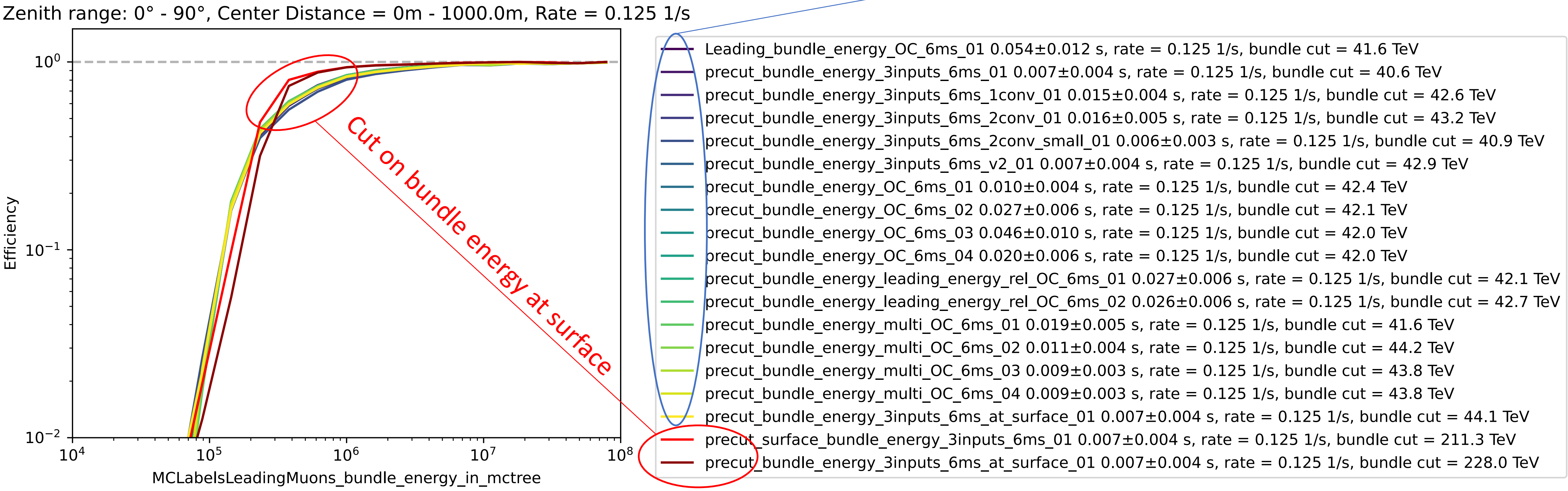
- Rate after muon filter: 24.62 Hz
- If the process of 1 event needs 1 second, 8h run takes 200h -> needs to be reduced!
- Use small, fast network to remove low energy events -> target rate 125 mHz



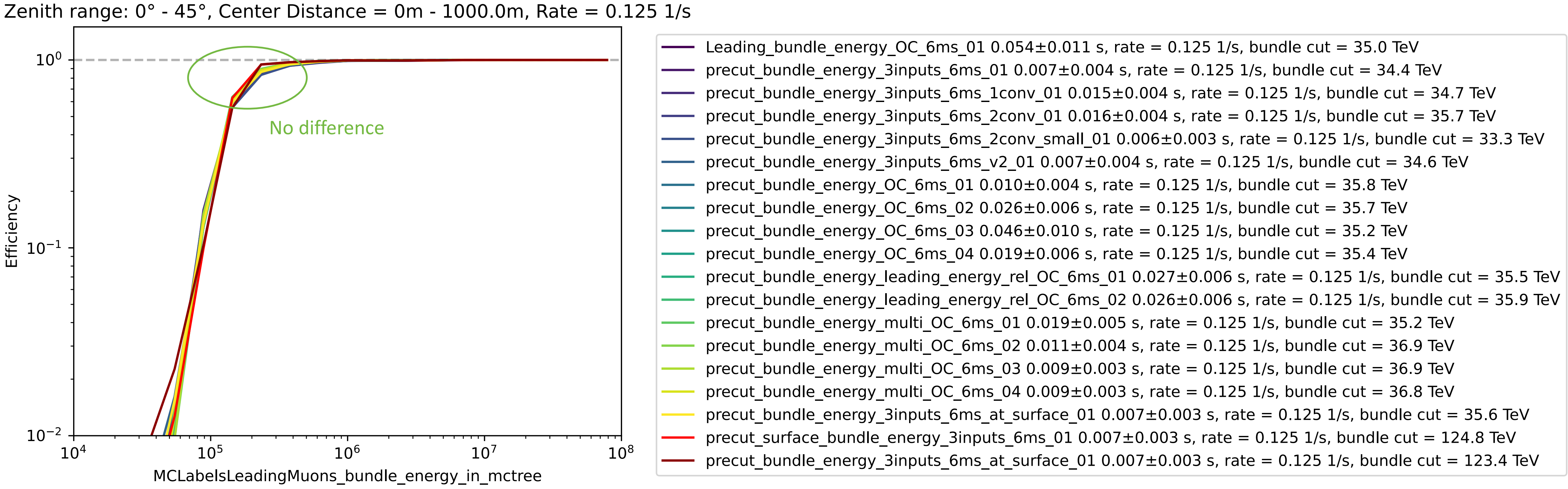
Network: precut_surface_bundle_energy_3inputs_6ms_01

Rejection efficiency - all events

Cut on bundle energy at entry

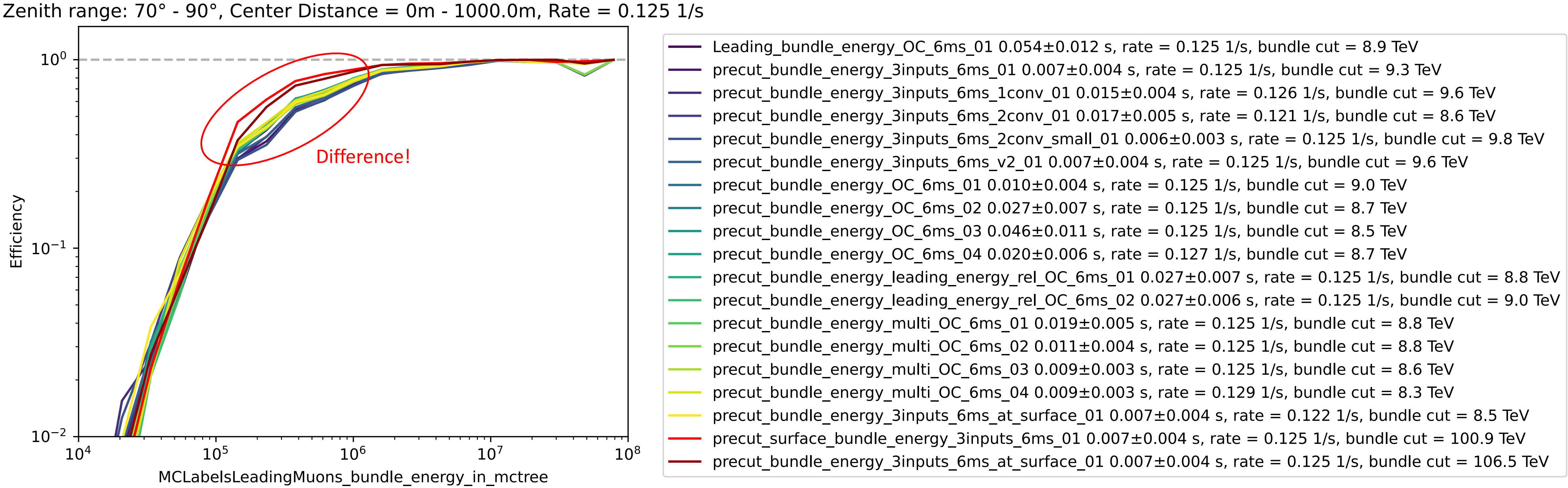


Rejection efficiency – small zenith



Rejection efficiency – high zenith

- At high zenith angles, events with high energies are removed with a cut on the bundle energy at entry



- Rate after muon filter: 24.62 Hz
- Rate after bundle cut: 144 mHz
- Perform 200 TeV cut on bundle energy at surface

Selection – Level 4

1. Add DNN network reconstructions

Add DNN networks

- Internal DNN pulse cleaning leads to slightly better results

Network	Preprocess / ms	CPU / ms	GPU / ms
direction_9inputs_uncleaned_medium_01	22 ± 20	106 ± 42	5 ± 38
leading_bundle_surface_leading_bundle_energy_OC_inputs9_large_log_02	22 ± 20	144 ± 56	3 ± 13
track_geometry_9inputs_uncleaned_01	22 ± 20	106 ± 42	3 ± 10
precut_surface_bundle_energy_3inputs_6ms_01 (added in level 3)	22 ± 20	11 ± 1	7 ± 4

What we have discussed today

- [x] Reconstruction
    - [x] Trained networks to reconstruct several properties
- [x] Selection
    - [x] Level 3 (L2MuonFilter + 200 TeV bundle energy cut at surface)
    - [x] Level 4 (add reconstructions)
- [ ] Data-MC
    - [ ] Check several properties
    - [ ] Largest mismatch in z-vertex
- [ ] Forward fit
    - [ ] Test NNMFit for analysis
- [ ] Unfolding
    - [ ] Unfold event rate
    - [ ] Calculate effective area
- [ ] New simulations
    - [ ] Preparation for large scale IceProd simulation (latest software, switch some options)
- [x] Wiki
    - [x] Created and uploaded

Collaboration meeting

- Ask for WG Reviewer
- Storage request large-scale simulation

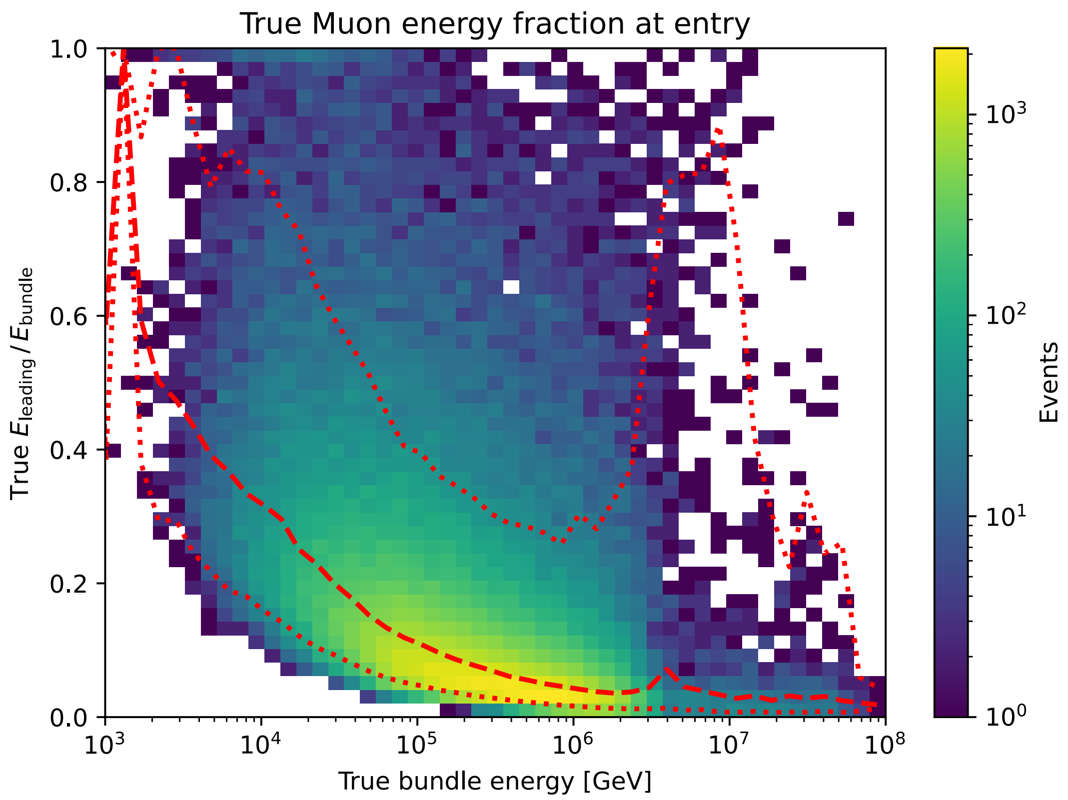
Thank you for your attention ☺️

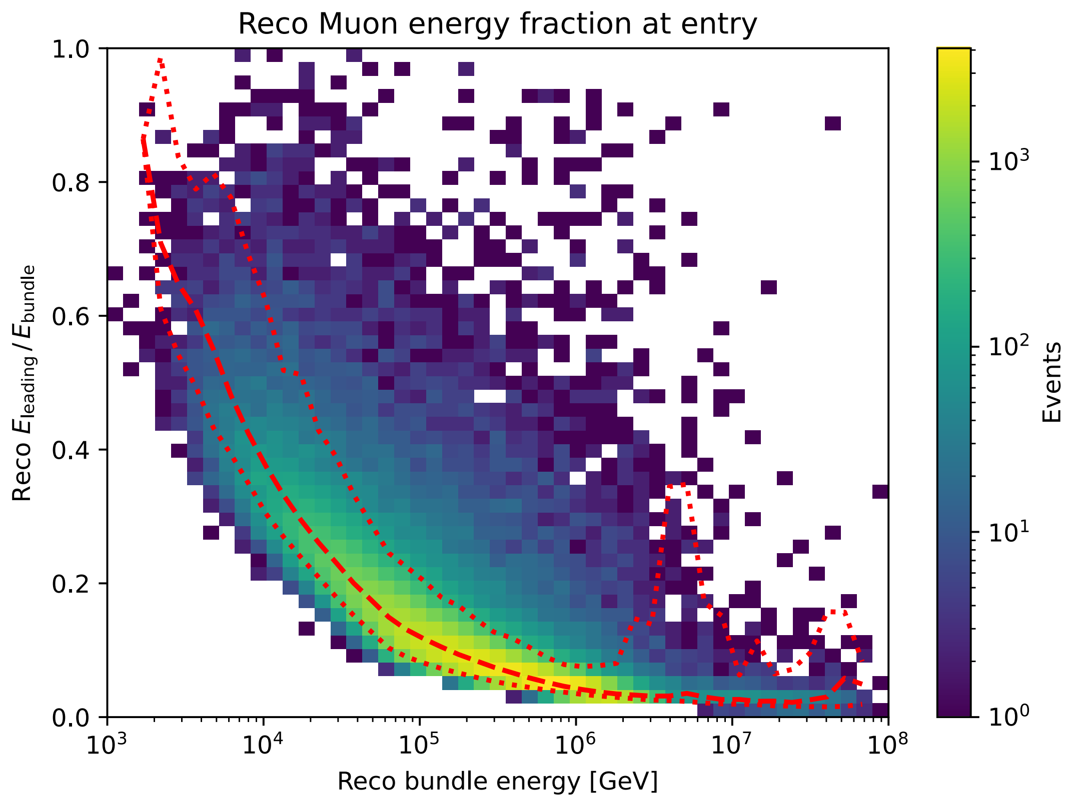
Backup

tu dortmund university

Leading muon energy fraction - leadingness

- True muon energy fraction is smeared
- Network tries to predict the median of the distribution

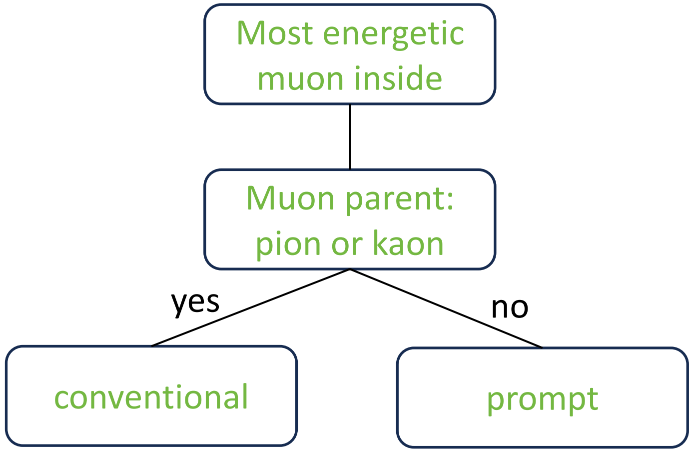




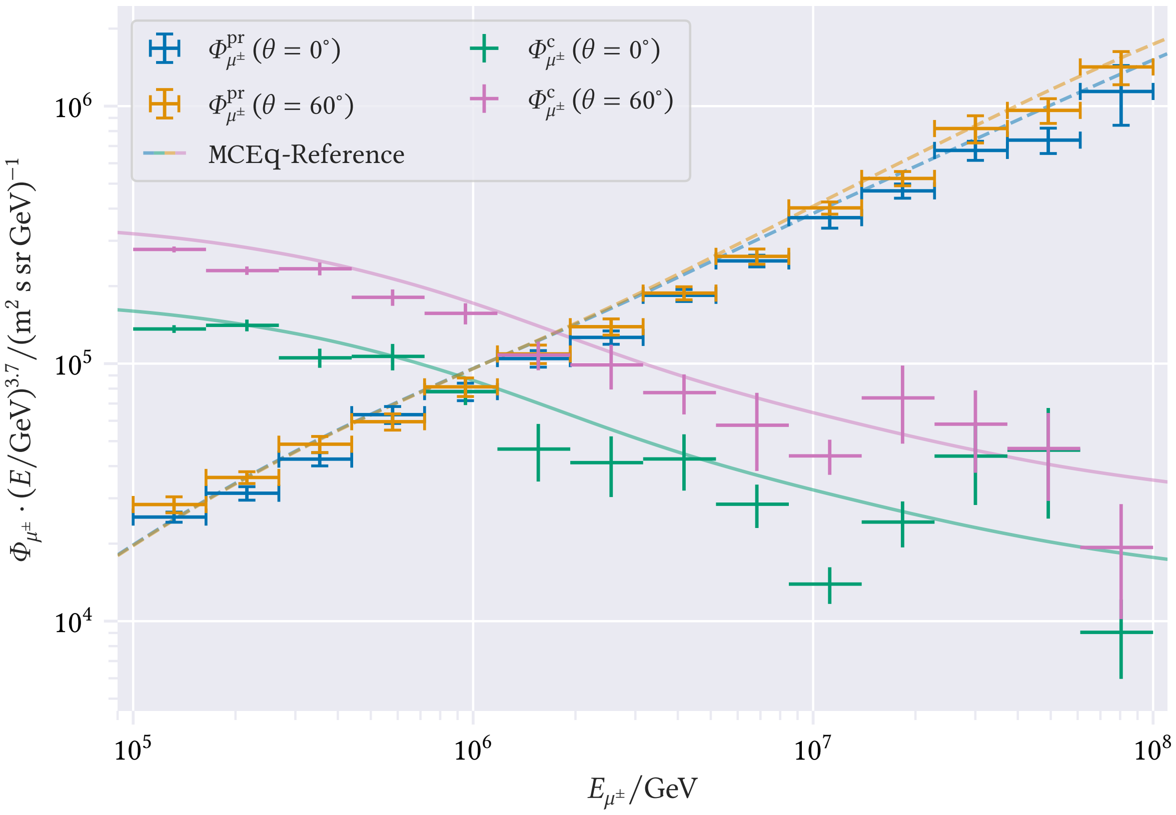
CORSIKA vs. MCEq

➢ Presented by Ludwig Neste

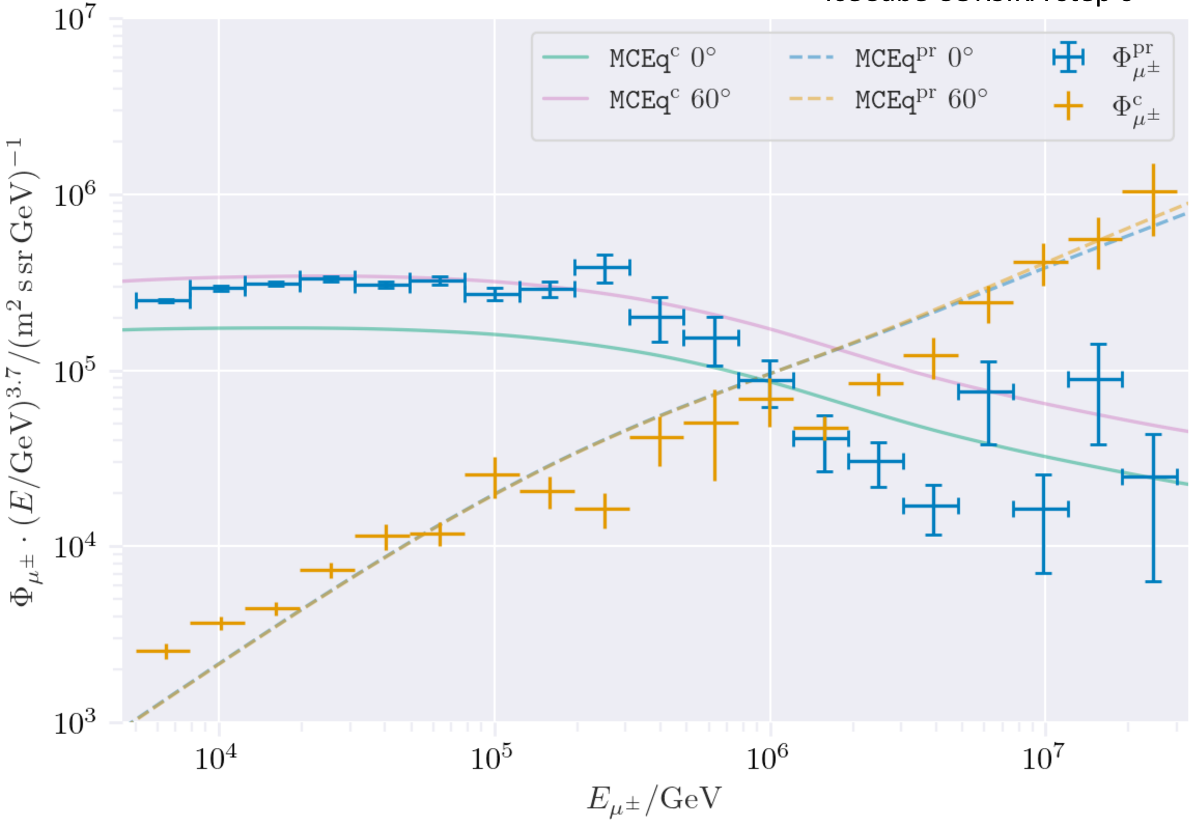
➢ Good agreement



CORSIKA standalone

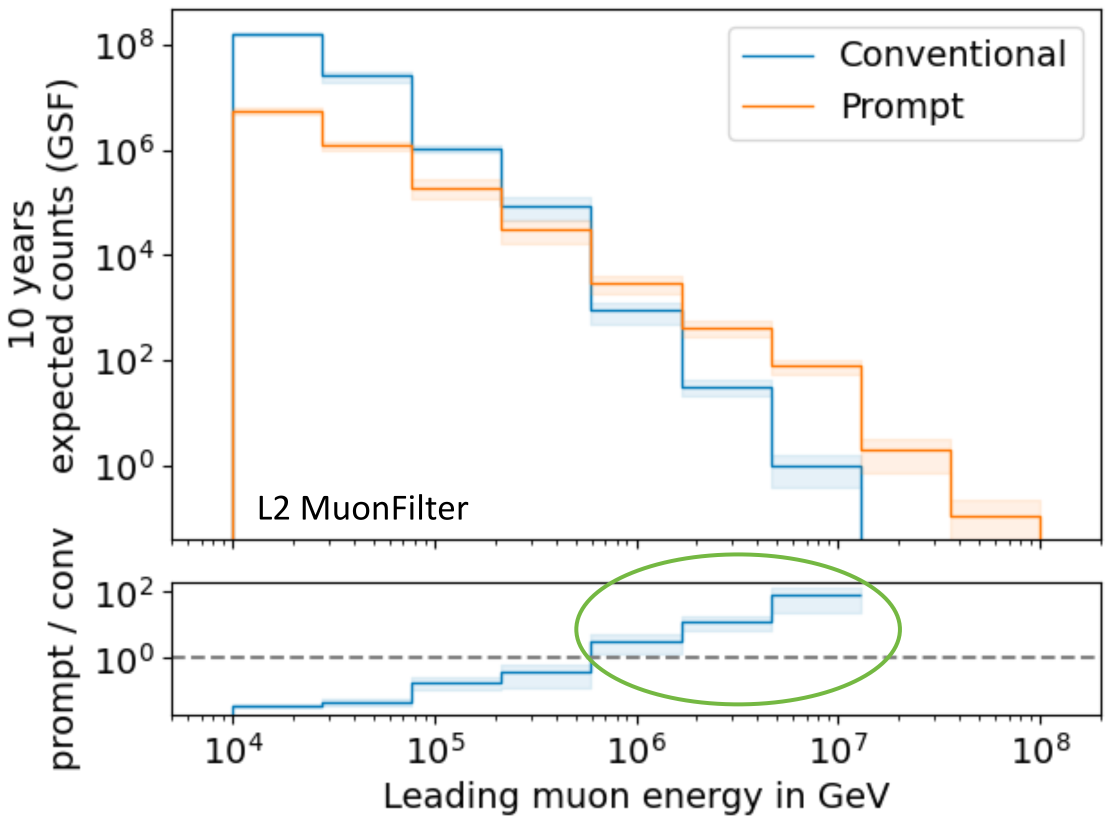


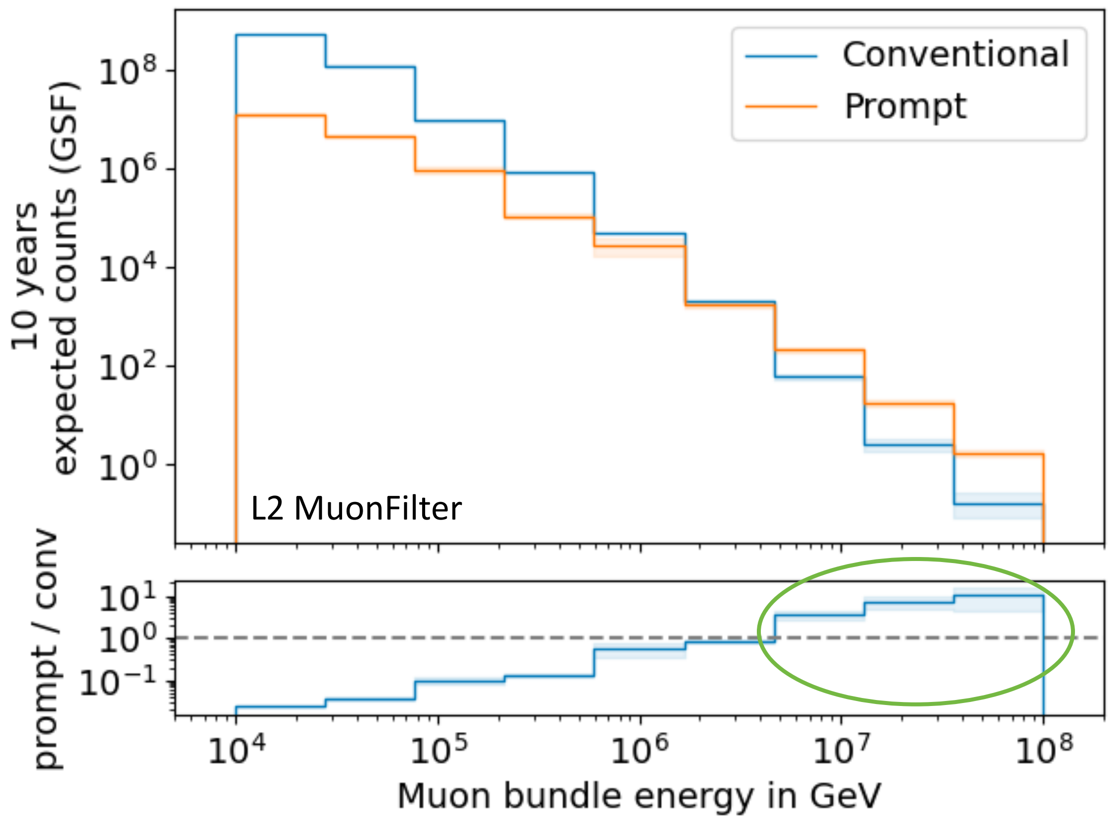
IceCube CORSIKA step 0



MC data exploration

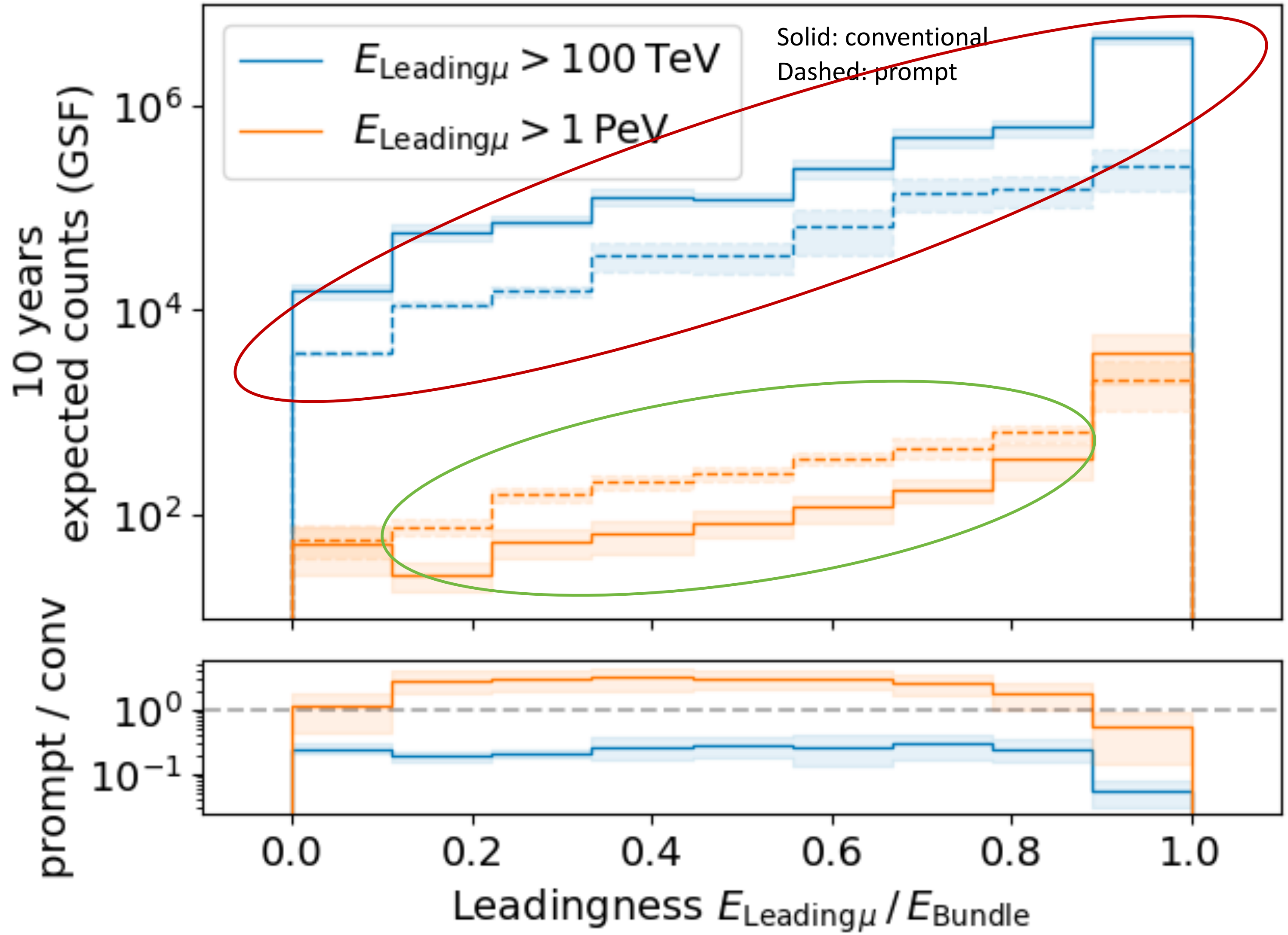
Expected muons for 10 years: leading vs. bundle energy (GSF)





- Both leading and bundle energy are sensitive to detect prompt
- Leading muon energy is more sensitive

Leading muon energy fraction



- ➢ Prompt dominates for energies > 1 PeV
- ➢ Leading energy sweet spot: 0.1 – 0.9

Leading muon contribution

Patrick Berghaus

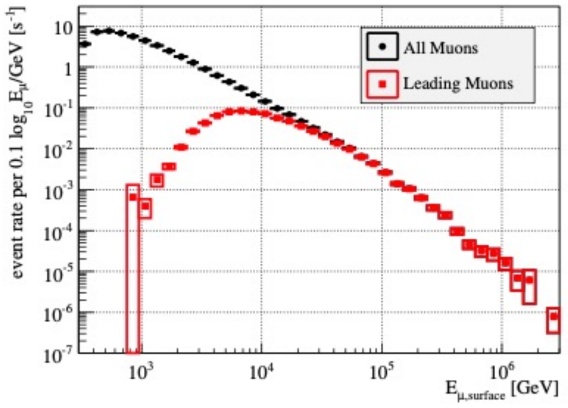
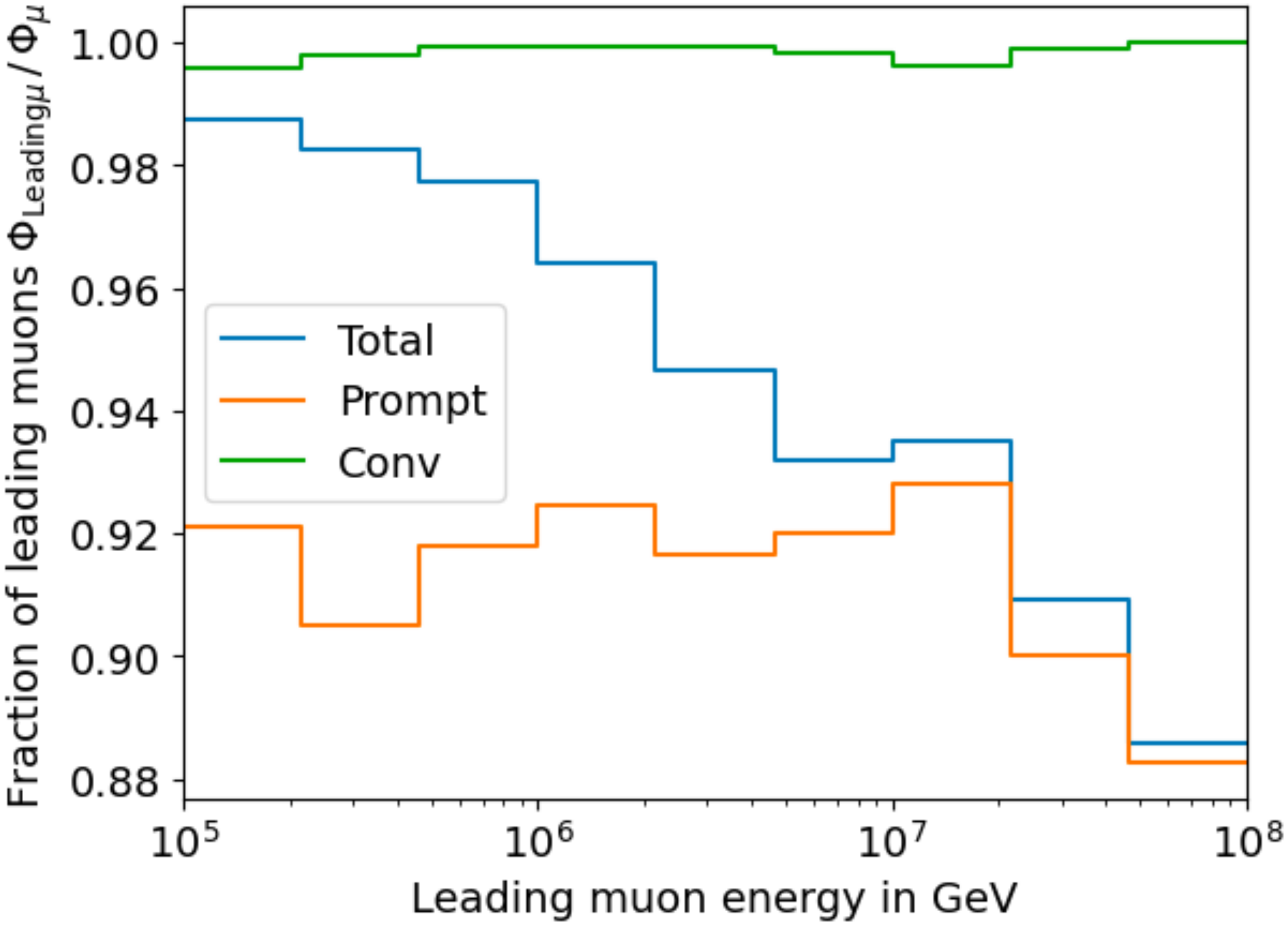


Figure 10: Surface energy distribution for all and most energetic ("leading") muons in simulated events with a total of more than 1,000 registered photo-electrons in IceCube.

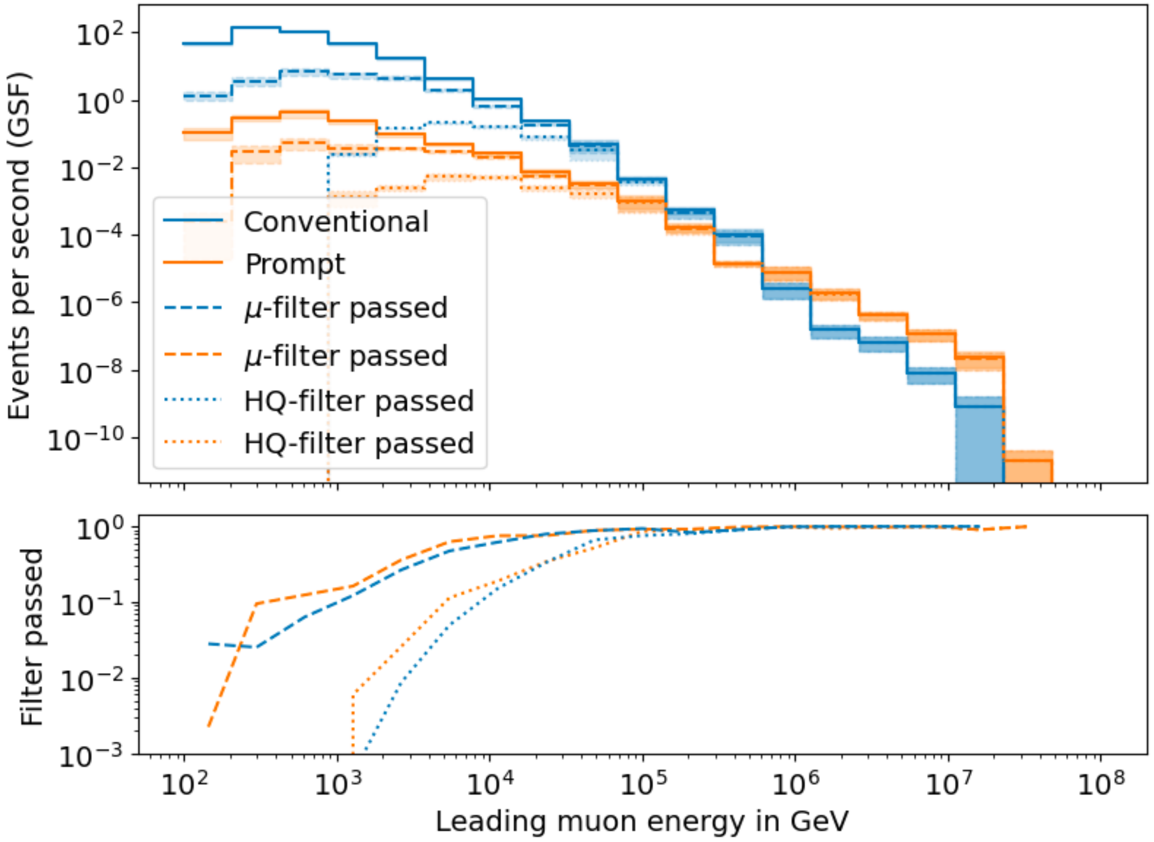


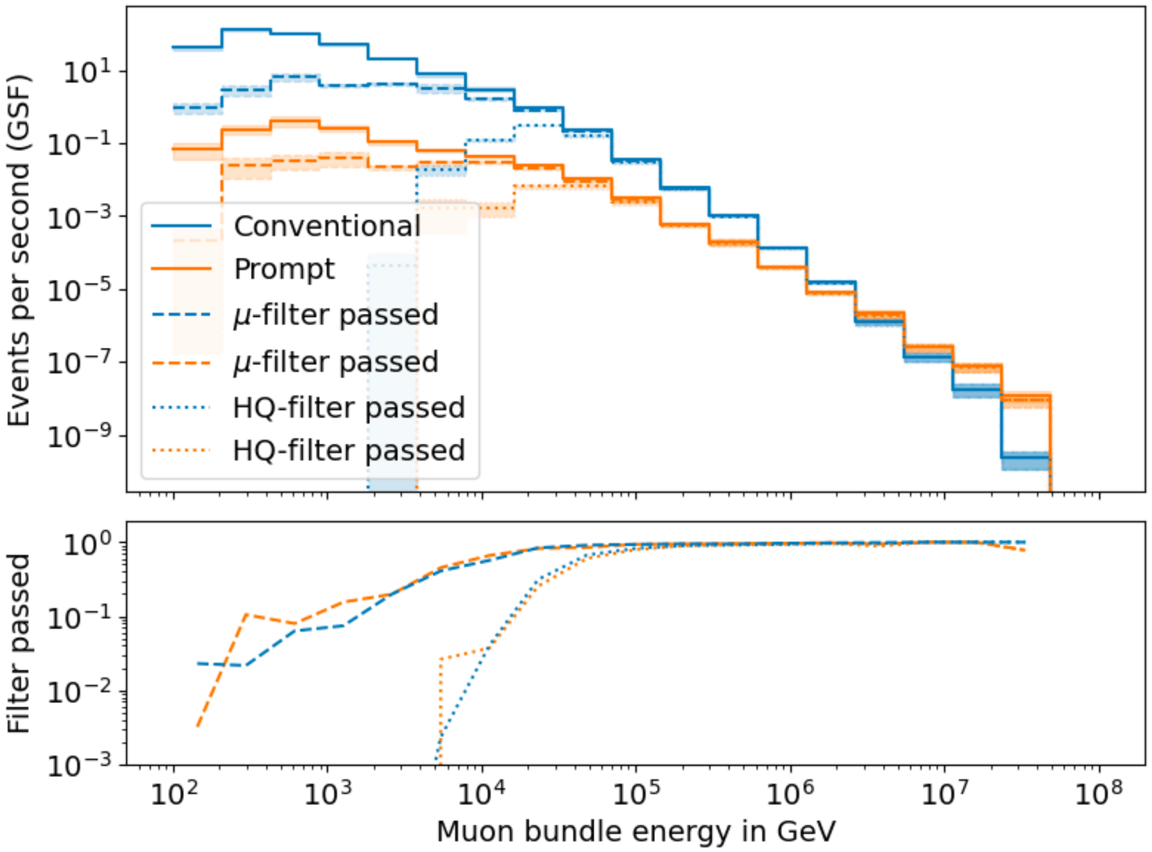
- Muons with energies between 100 TeV and 50 PeV dominate the bundle by more than 90%
- In average conventional muons are more dominant than prompt
- But: at high energies, there are more prompt than conventional events

➢ High leading energy fraction does not lead to more sensitivity to detect prompt

L2 Filters

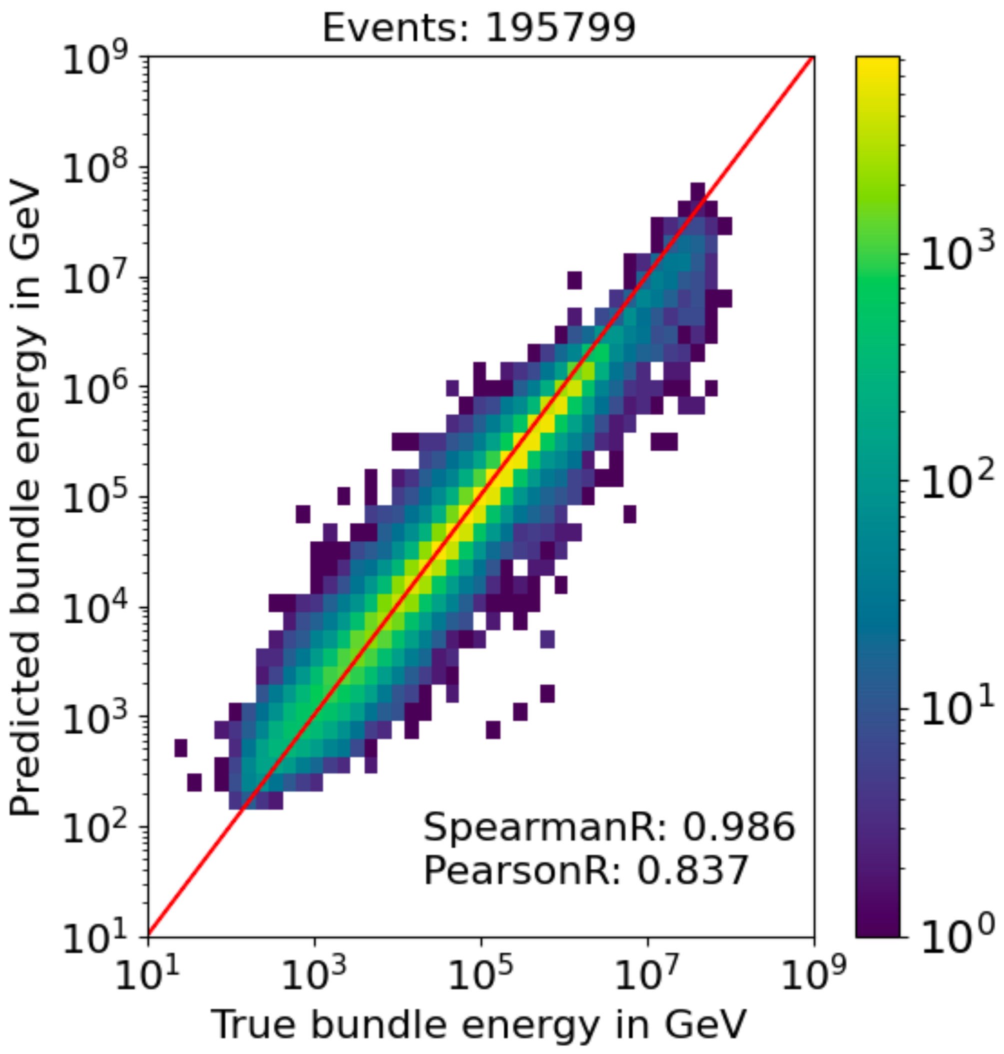
Fraction events rejected	All energies	Leading energy > 10 TeV	Leading energy > 100 TeV
MuonFilter	0.93	0.28	0.06
HQFilter	0.99	0.74	0.18

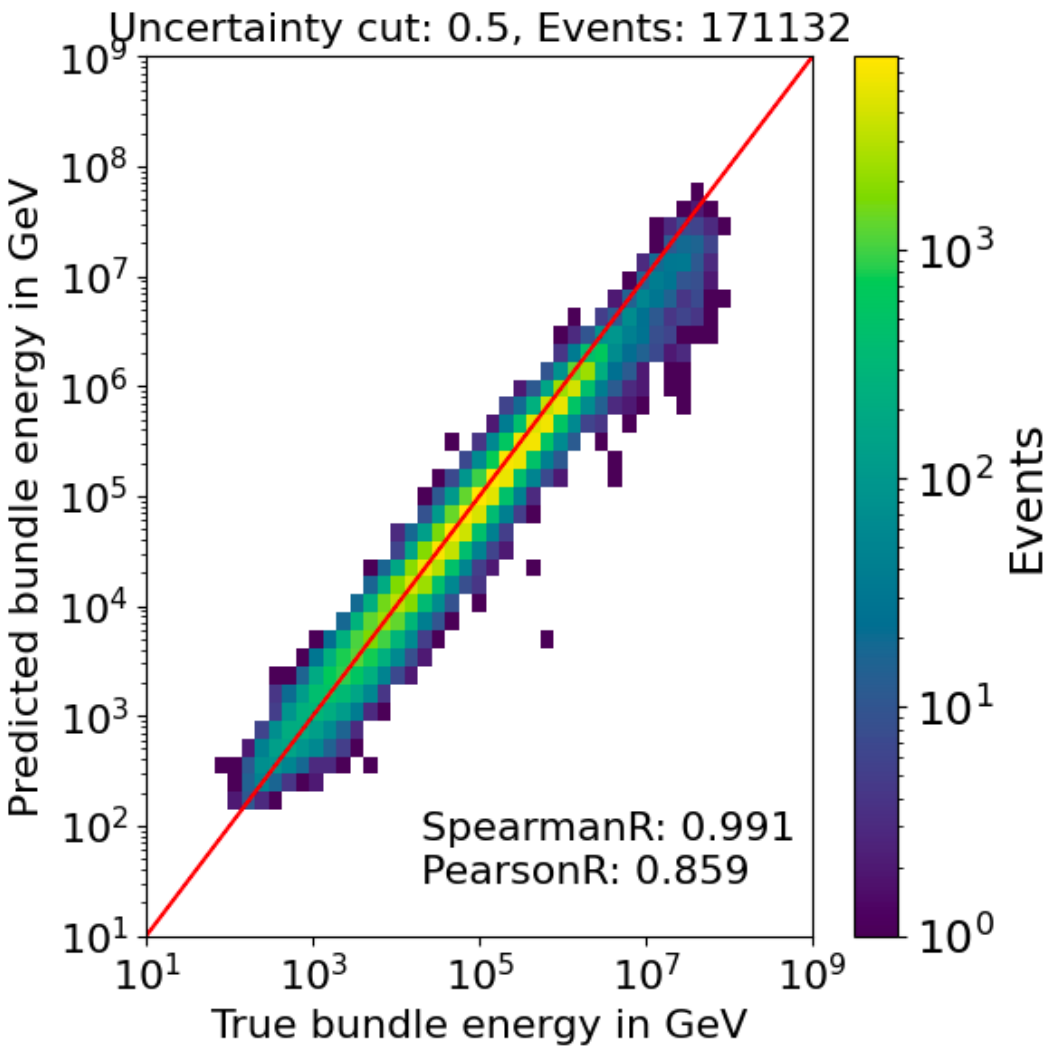




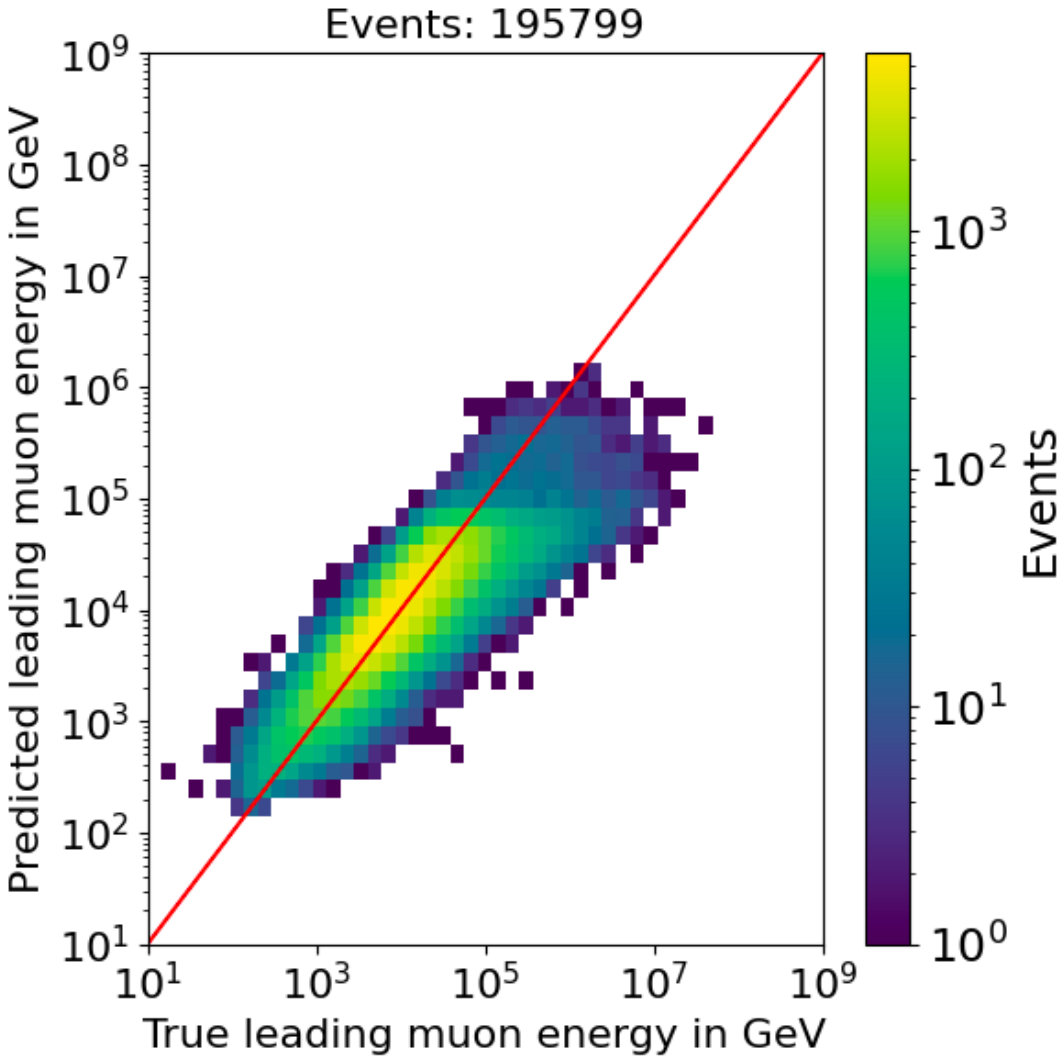
➢ Choose muon filter to remove large amount of statistics at low energies

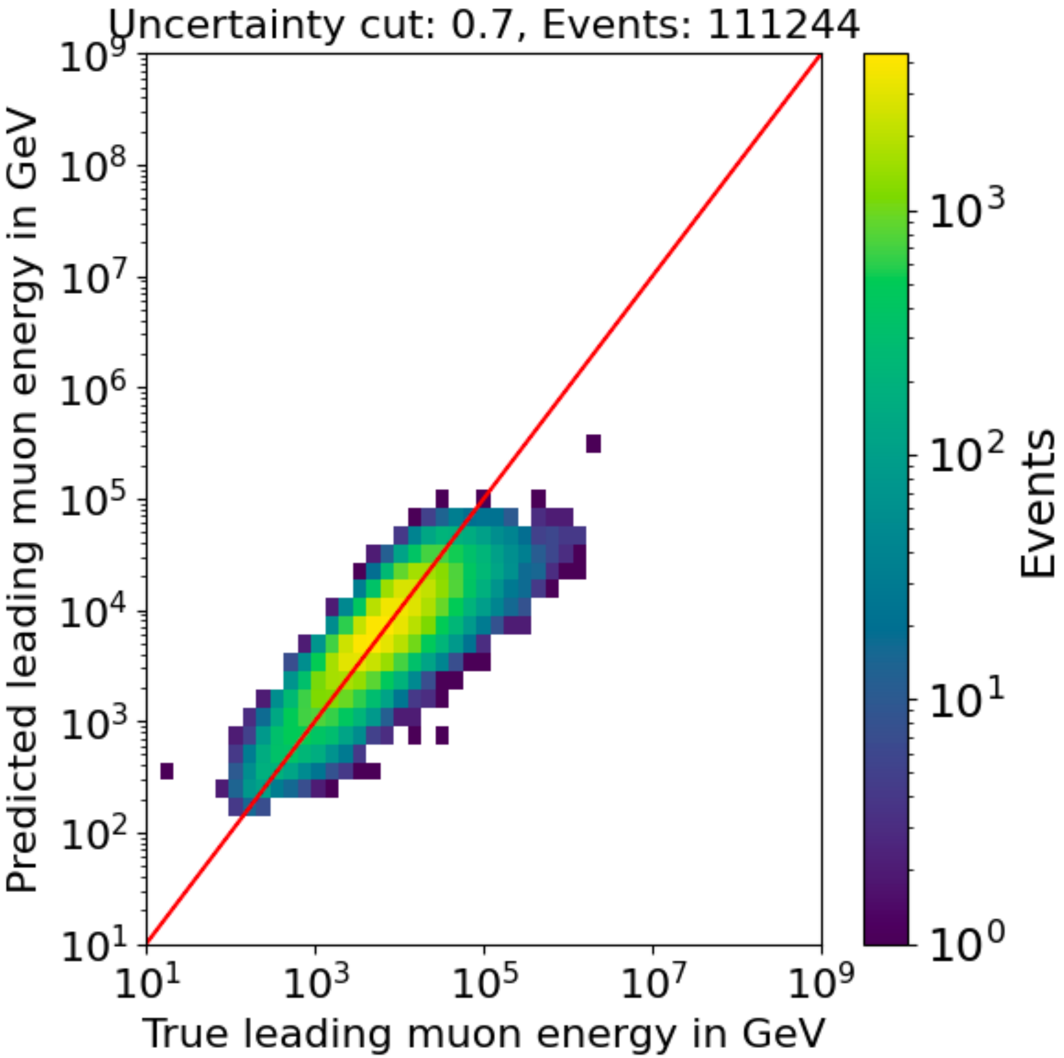
Bundle energy reconstruction





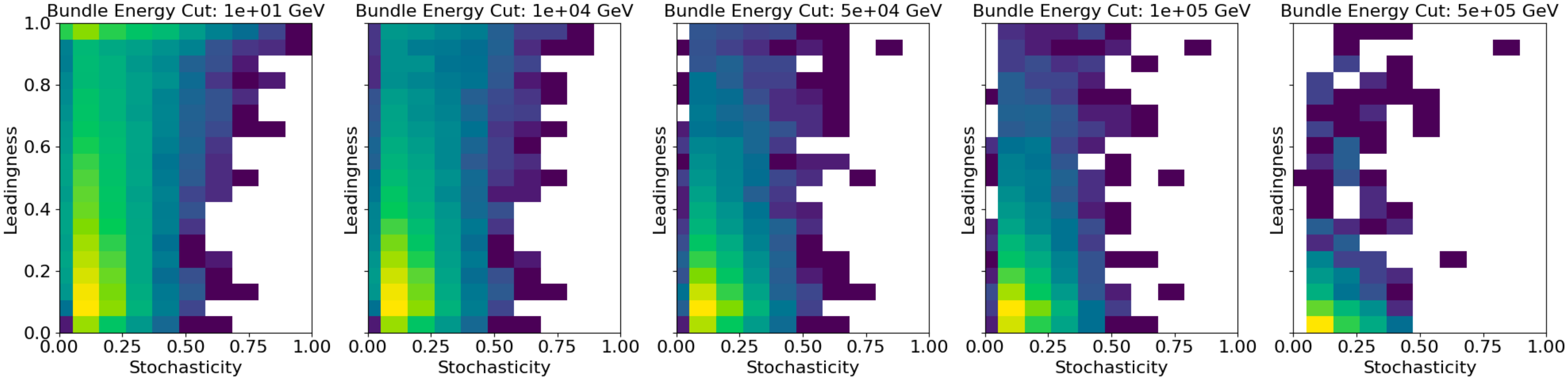
Leading energy reconstruction

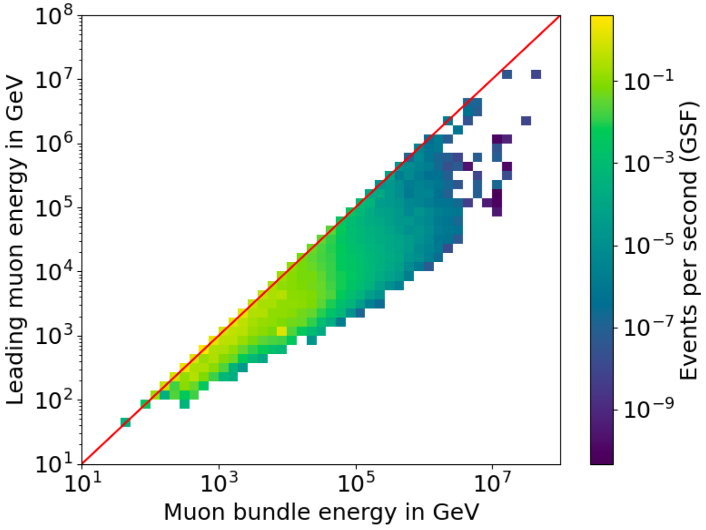


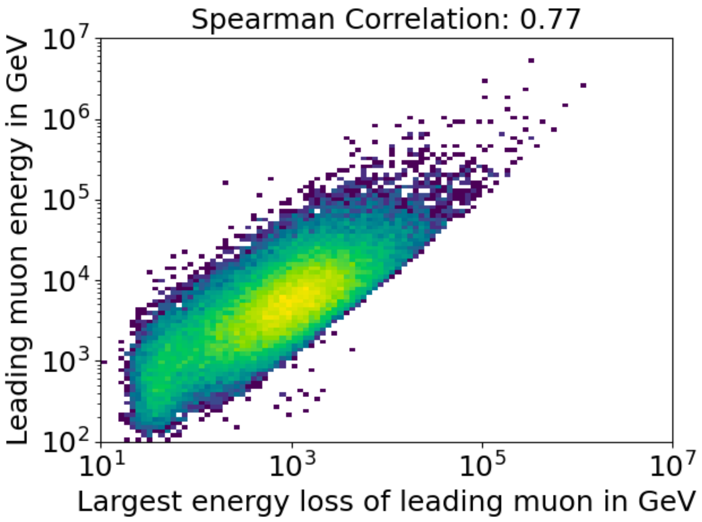


Bundle energy cut

Stochasticity



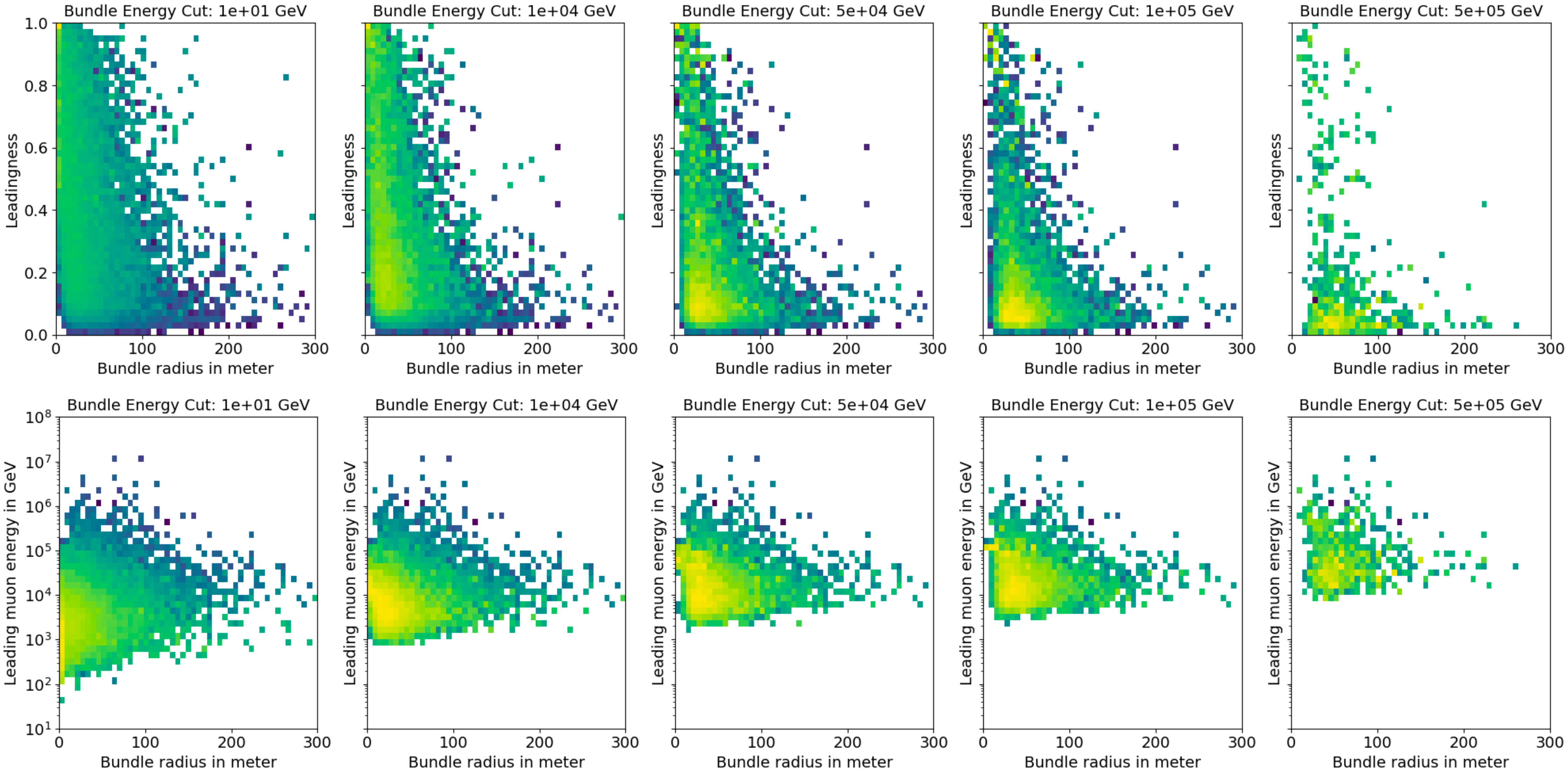




- High stochasticity leads to high leadingness, but only for a small number of events
- Leading muon energy smears out at large bundle energies
- Largest energy loss of the leading muon correlates with the leading muon energy

Bundle energy cut

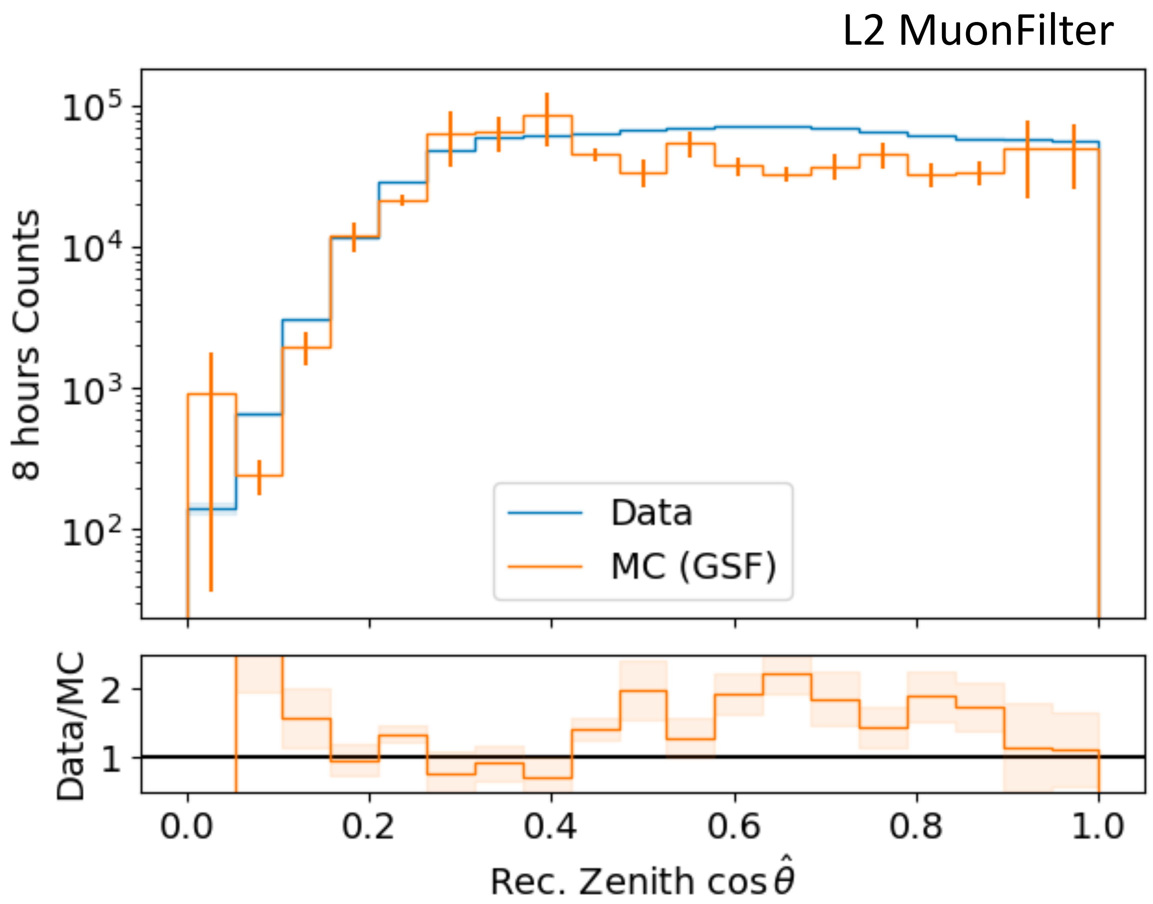
Bundle radius

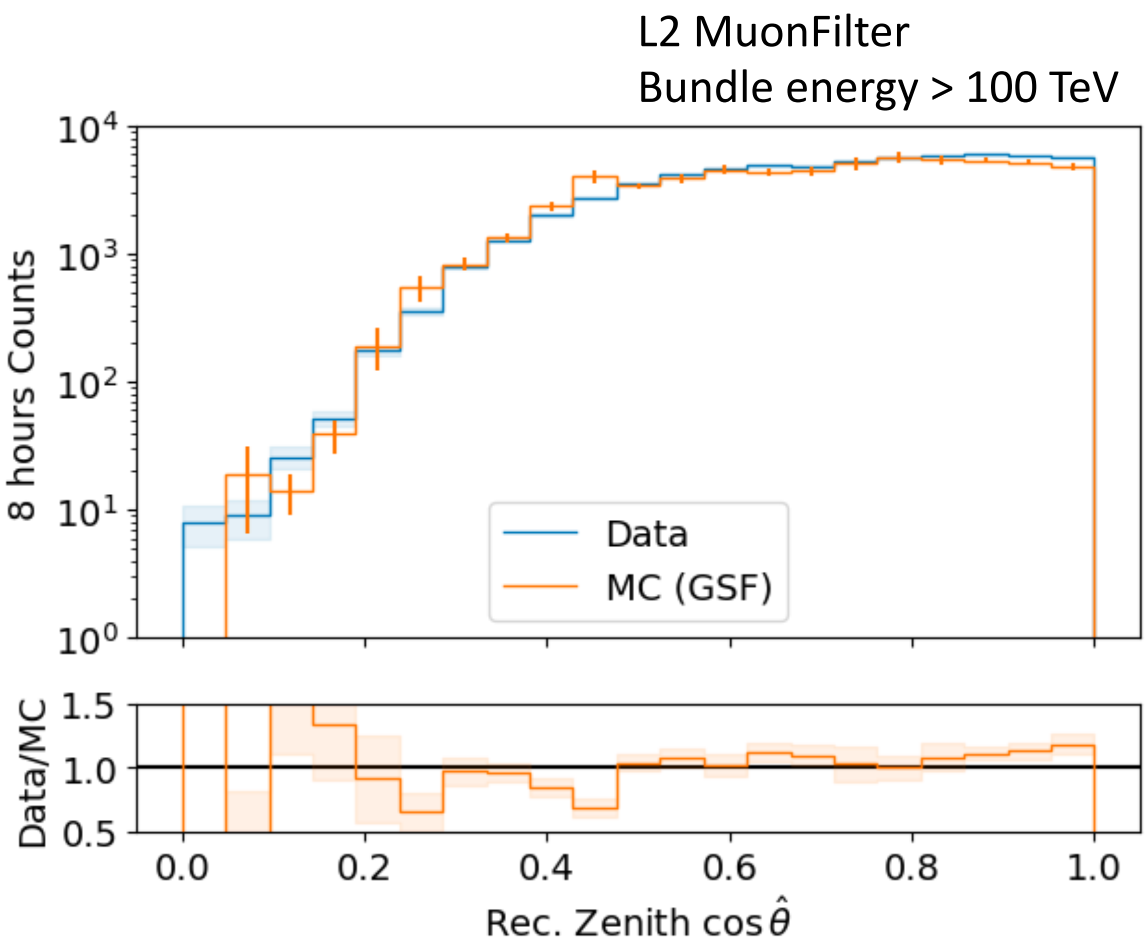


➢ Large bundle radius leads to low leadingness

Data/MC

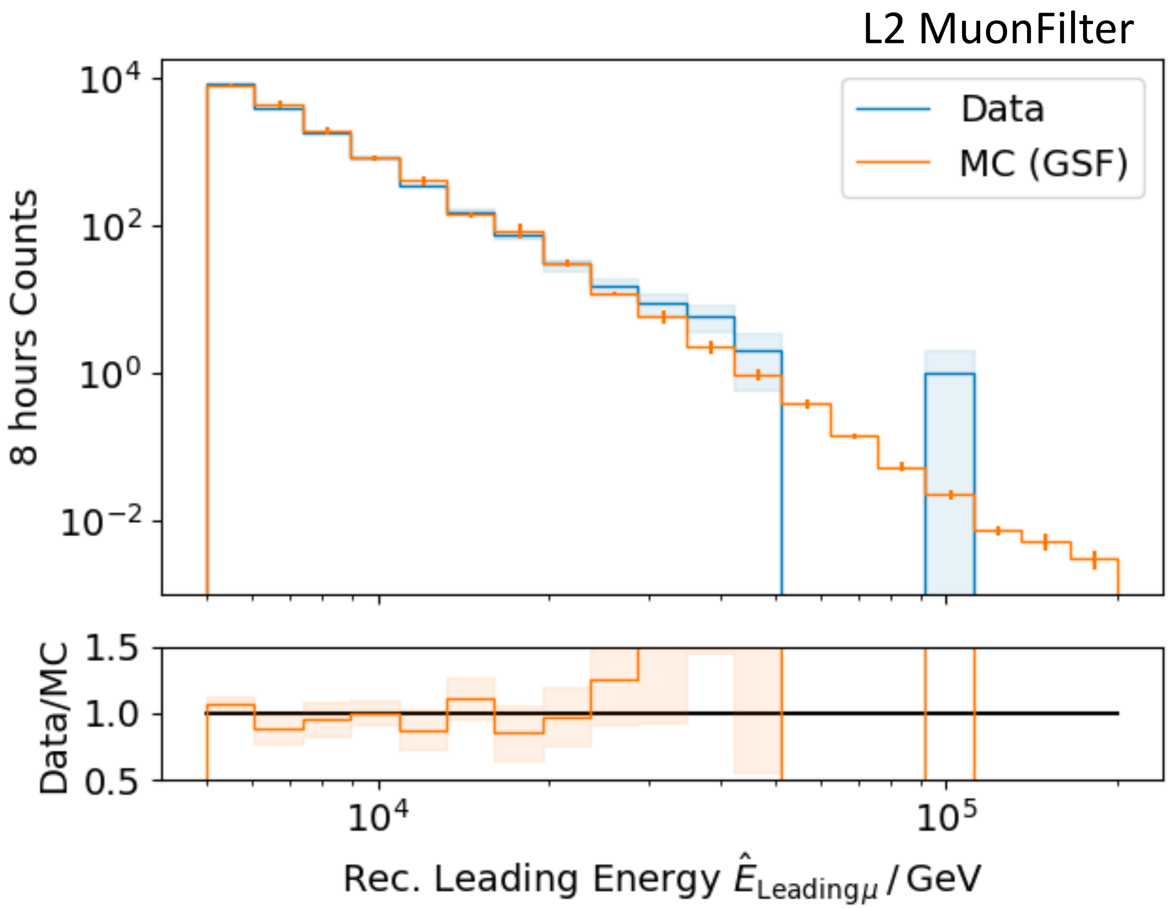
Data-MC: cos(zenith)

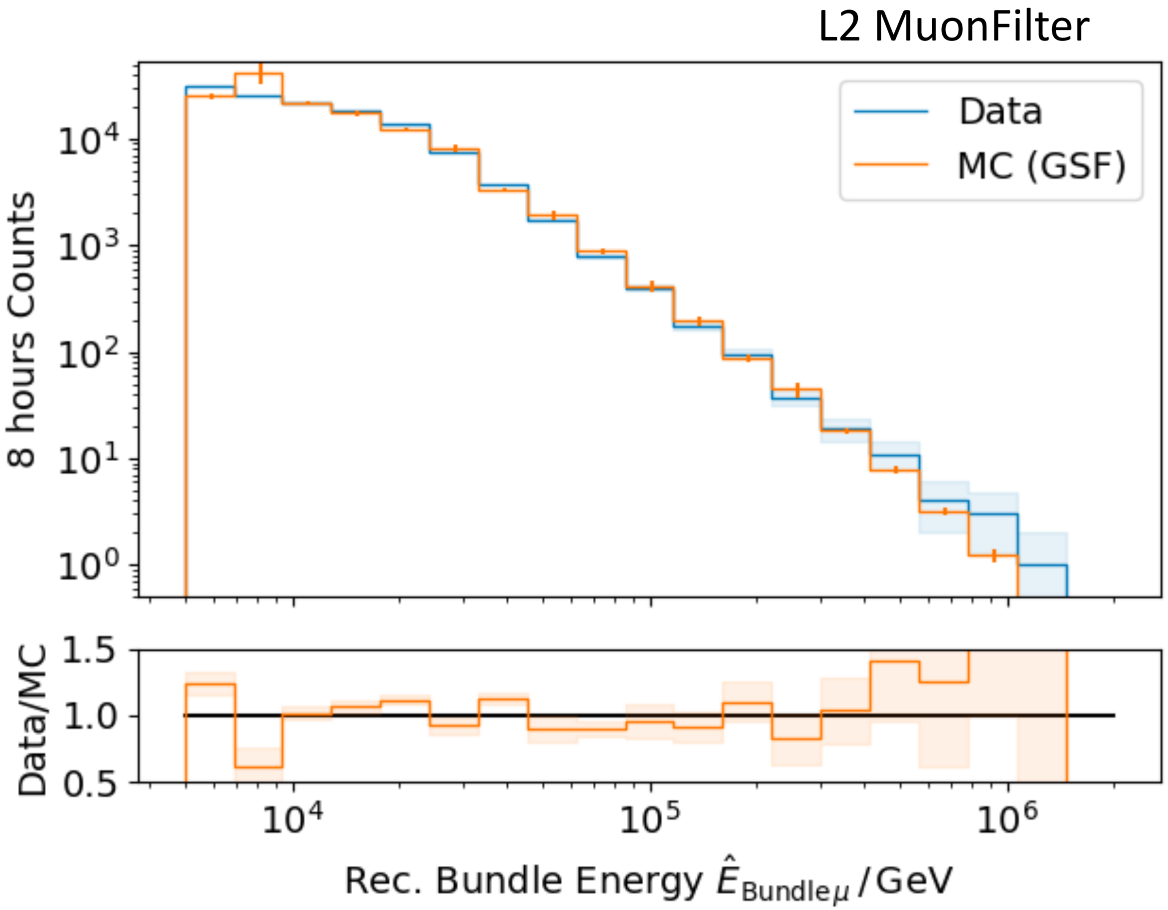




- Deviations at low cos(zenith), but very small statistics
- More data at cos(zenith) > 0.5
- Less data at cos(zenith) ~ 0.3

Data-MC: energy spectrum



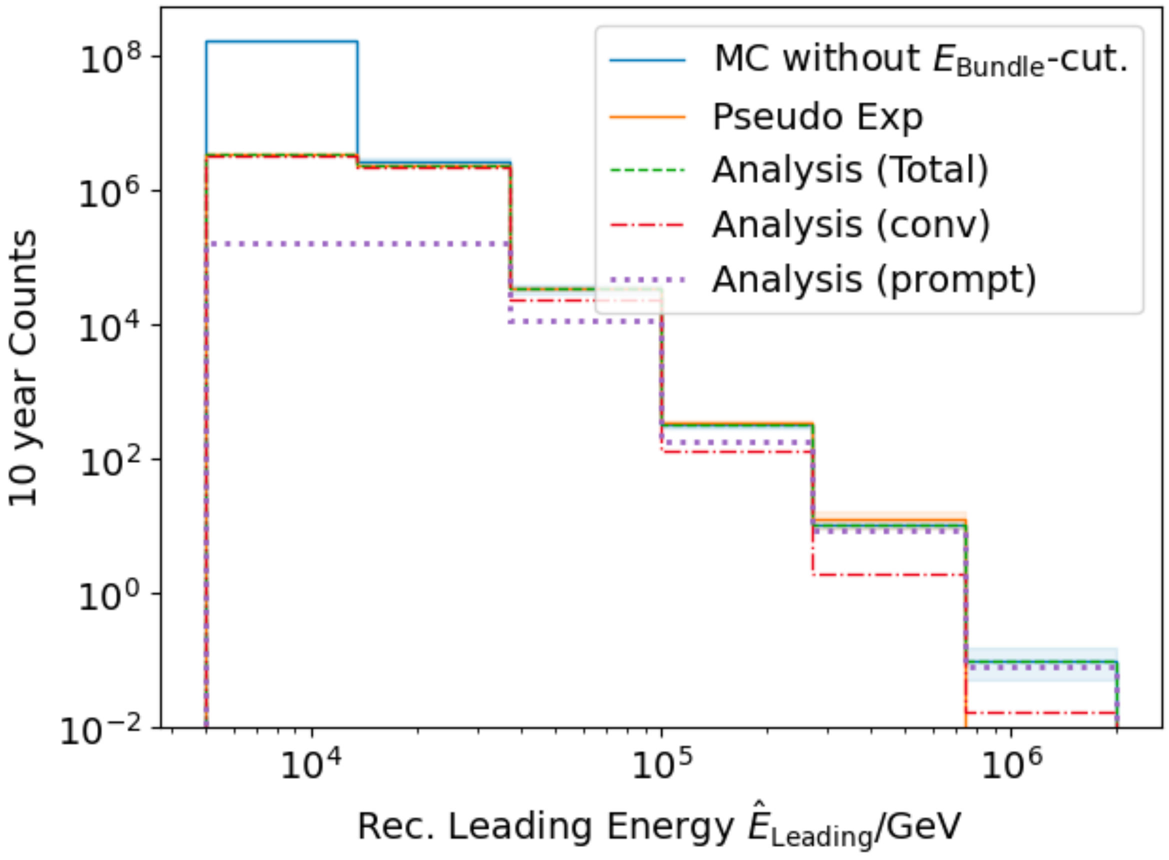


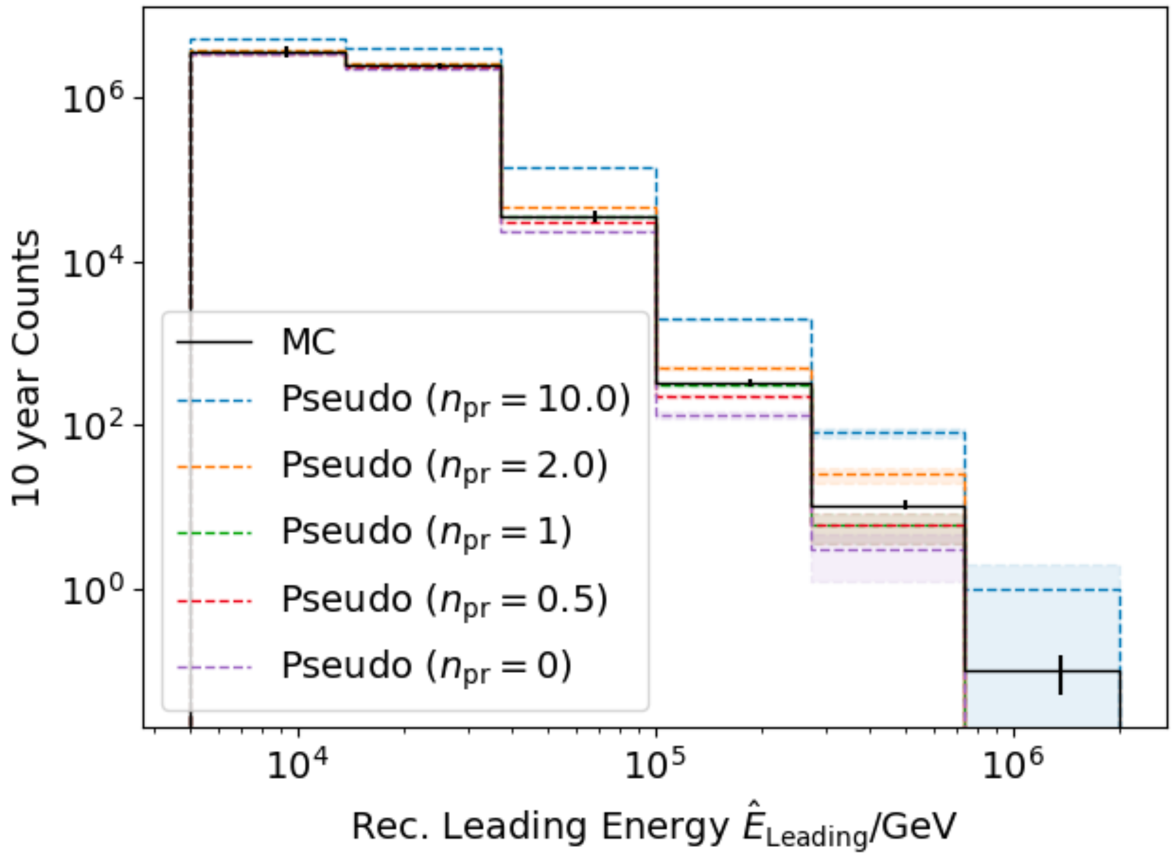
- Bundle energy: good agreement with GSF up to 300 TeV
- But insufficient statistics

Pseudo analysis

Cuts:
L2 MuonFilter
Bundle energy > 100 TeV

Pseudo data sampling





➢ Tagging allows scaling of prompt by factor n_{pr}

Poisson likelihood fit performed in leading muon energy

Prompt scaling/normalization

Conv norm = 1

MC counts per bin i

$$C_1^{\mathrm{MC}} = n_{\mathrm{pr}} C_1^{\mathrm{MC,pr}} + n_{\mathrm{conv}} C_1^{\mathrm{MC,conv}}, \dots, C_M^{\mathrm{MC}} = n_{\mathrm{pr}} C_M^{\mathrm{MC,pr}} + n_{\mathrm{conv}} C_M^{\mathrm{MC,conv}}$$

Experimental counts

$$p\left(C_i\right) = p_{\mathrm{poisson}}\left(C_i; \lambda\left(n_{\mathrm{pr}}\right) = C_i^{\mathrm{MC}}\left(n_{\mathrm{pr}}\right)\right) = \frac{\lambda\left(n_{\mathrm{pr}}\right)^{C_i} e^{-\lambda\left(n_{\mathrm{pr}}\right)}}{C_i!}$$

Maximize likelihood

$$\mathscr{L}\left(n_{\mathrm{pr}}\right) = \prod_{i=1}^{M} p\left(C_i; n_{\mathrm{pr}}\right)$$

With a constant conv norm:
bin counts depend only on prompt norm
= expectation value per bin

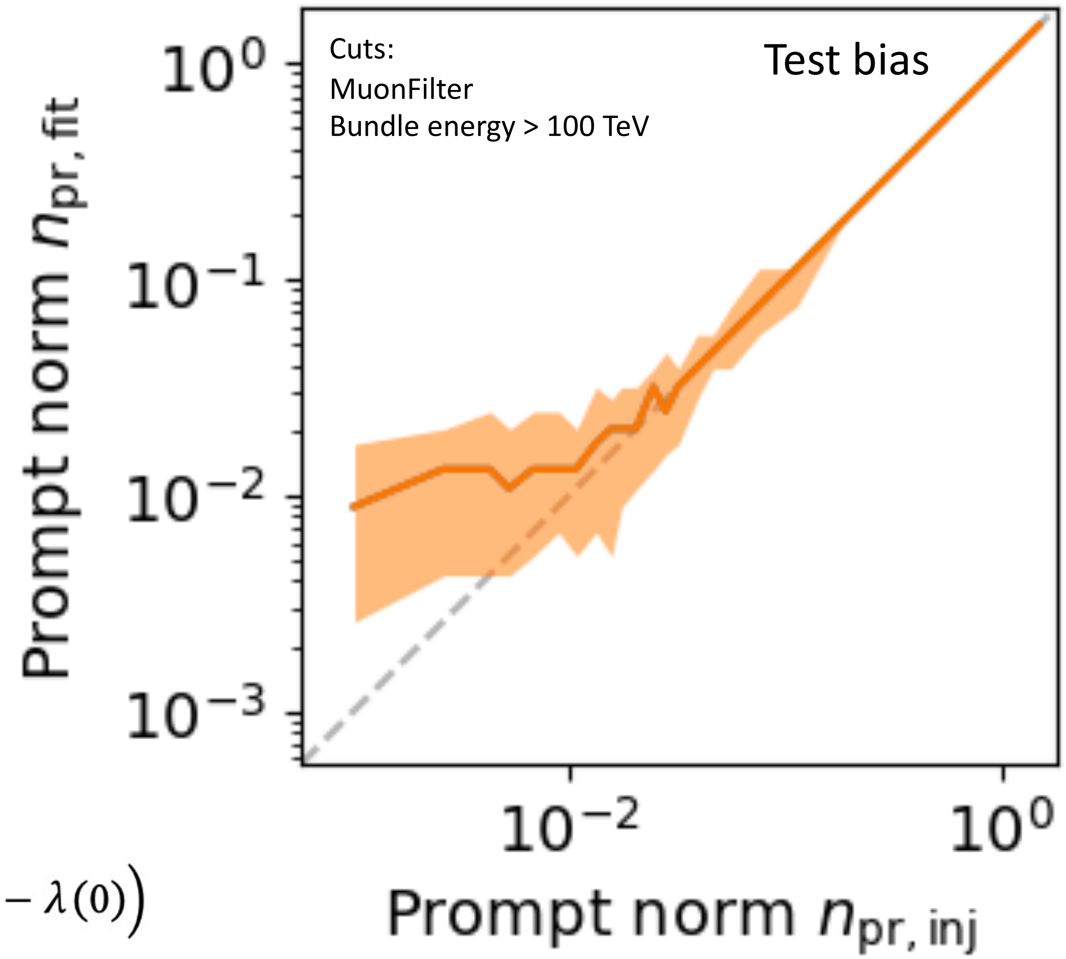
Easier:
minimize negative
log-likelihood

$$-\ln \mathscr{L} = -\sum_{i=1}^{M} C_i \ln \lambda\left(n_{\mathrm{pr}}\right) - \lambda\left(n_{\mathrm{pr}}\right) - \ln C_i!$$

$$\Lambda = -2 \ln \frac{\mathscr{L}\left(n_{\mathrm{pr}} = \hat{n}_{\mathrm{pr}}\right)}{\mathscr{L}\left(n_{\mathrm{pr}=0}\right)} = -2 \sum_{i=1}^{M} C_i \left(\ln \lambda\left(\hat{n}_{\mathrm{pr}}\right) - \ln \lambda(0)\right) - \left(\lambda\left(n_{\mathrm{pr}}\right) - \lambda(0)\right)$$

Test statistic for Wilks' theorem

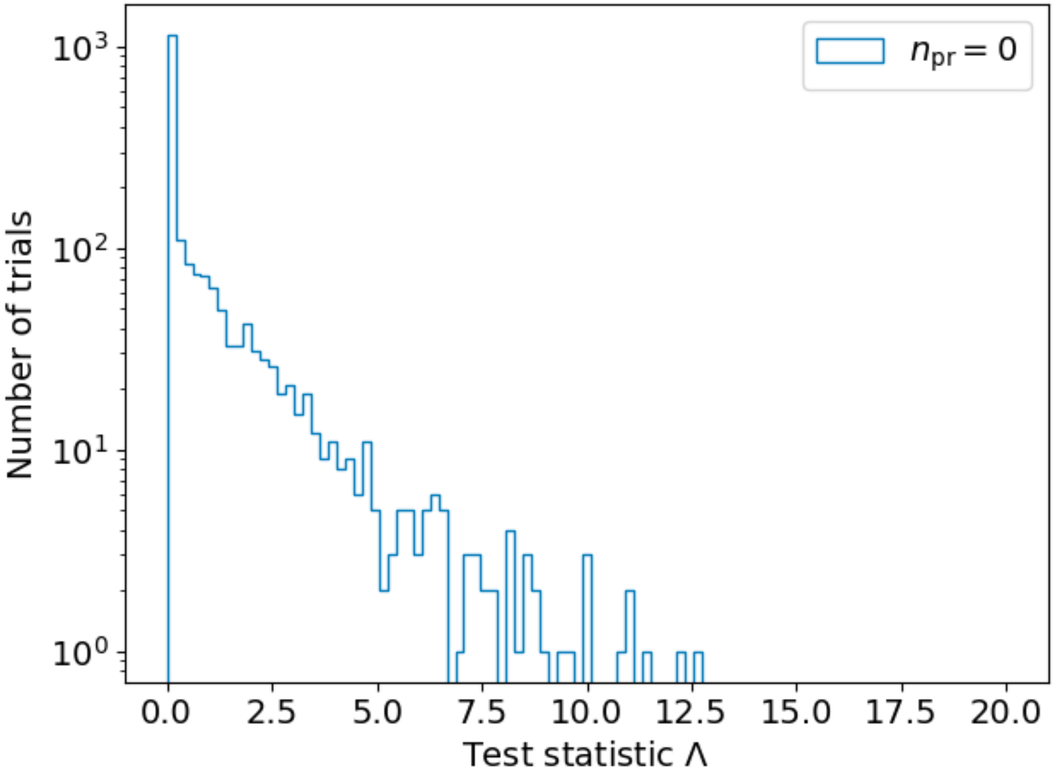
Null hypothesis: no prompt

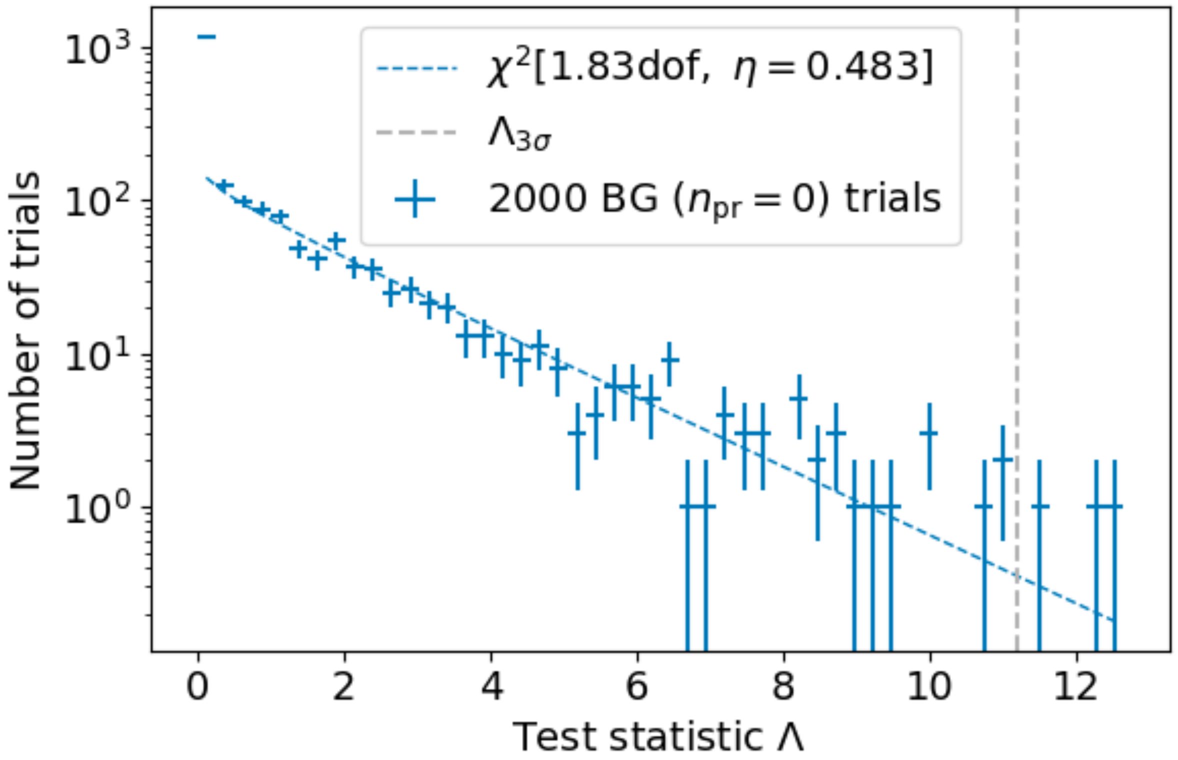


- Bias starts at a prompt normalization of 0.1

Test background statistics

Cuts:
L2 MuonFilter
Bundle energy > 100 TeV





- ➢ Background statistic is χ^2 – distributed
- ➢ Assume Wilks' theorem for test statistics

tu dortmund university

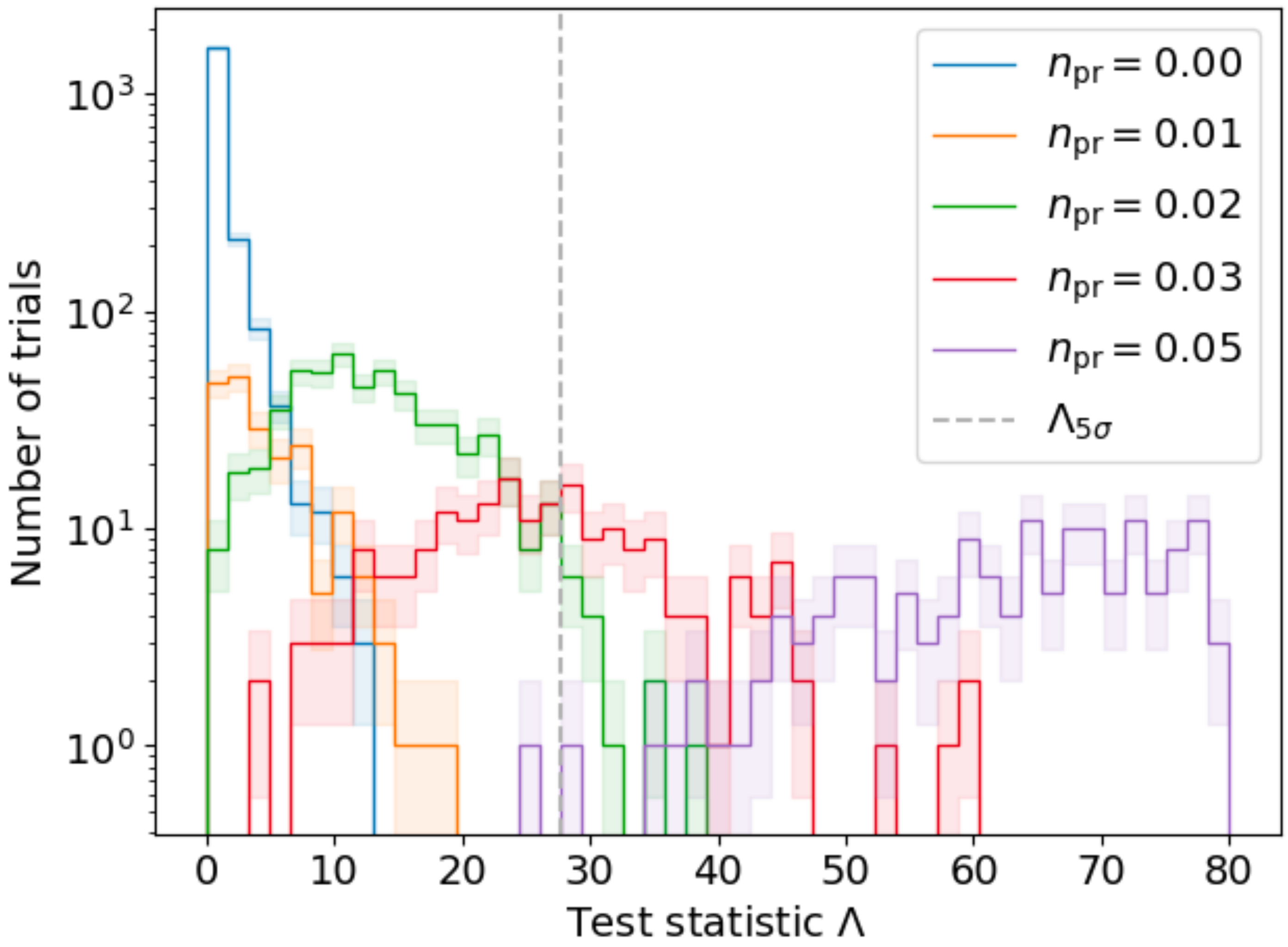
Cuts:
L2 MuonFilter
Bundle energy > 100 TeV

Discovery potential and sensitivity

Expectation for 1 year:
- 5 sigma discovery potential: 0.102 ± 0.005
- Sensitivity: 0.024 ± 0.001

Expectation for 10 years:
- 5 sigma discovery potential: 0.032 ± 0.001
- Sensitivity: 0.007 ± 0.000

Caution:
Limited MC statistics -> events are oversampled in pseudo dataset



Conclusion and outlook

Next goal

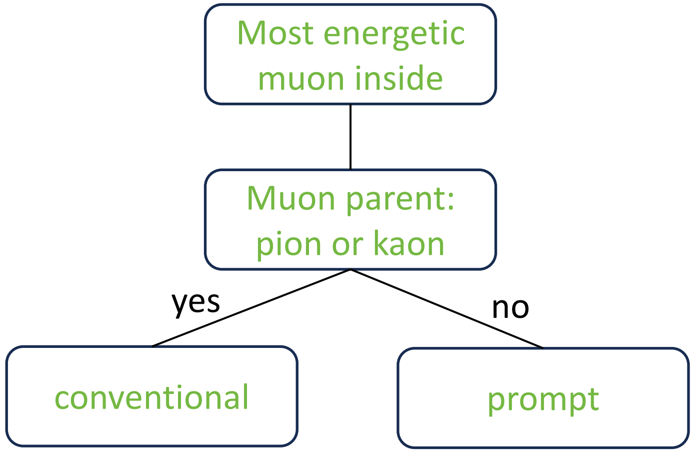
- ❑ Measure normalization of the atmospheric prompt muon flux

Steps for normalization

- ☑ Verify CORSIKA extended history simulations
- ☑ Tag prompt muons (see talk by Ludwig Neste)
- ☑ Comparisons to MCEq (see talk by Ludwig Neste)
- ☑ Set up preliminary analysis chain
- ★ Reconstruct muon energy and direction
- ★ Data/MC comparisons
- ❑ Include systematics
- ❑ Run full statistics simulation

Some definitions and wording…

- Leading muon
  - The most energetic muon inside a bundle (no minimum fraction required)
- Single muon
  - Except for stopping and very low energetic muons, there are never any single muons (almost every event is muon bundle)
- Prompt muon
  - Muon parent is not pion or kaon



Suggestion:

To avoid confusion regarding different leading muon definitions we can introduce a "leadingness" (For example: Tomasz used a leadingness of 50%, …)

Intention

1) Detect prompt component of the atmospheric muon flux significantly
   - Measure the normalization
   - Get handle on hadronic interaction models
2) Unfold an energy spectrum

Idea:

- New CORSIKA simulations with extended history
- Tag muons by parent → prompt or conventional
- Scale amount of prompt particles
  - Scaling saves time and resources instead of doing multiple simulations with different interaction models
  - Perform forward fit of the prompt normalization

Definition of the muon flux

$$\Phi_{\text{tot}} = \Phi_{\text{conventional}} + \Phi_{\text{prompt}}$$

$\pi, K \propto$ E$^{-3.7}$

"not" $\pi, K \propto$ E$^{-2.7}$

Conventional component:

$$\underset{\propto E^{-2.7}}{p} \rightarrow \overbrace{\pi^{+}}^{\tau = 2.6\cdot 10^{-8}\,\mathrm{s}} \rightarrow \underset{\propto E^{-3.7}}{\mu^{+}} + \nu_{\mu}$$

prompt component:

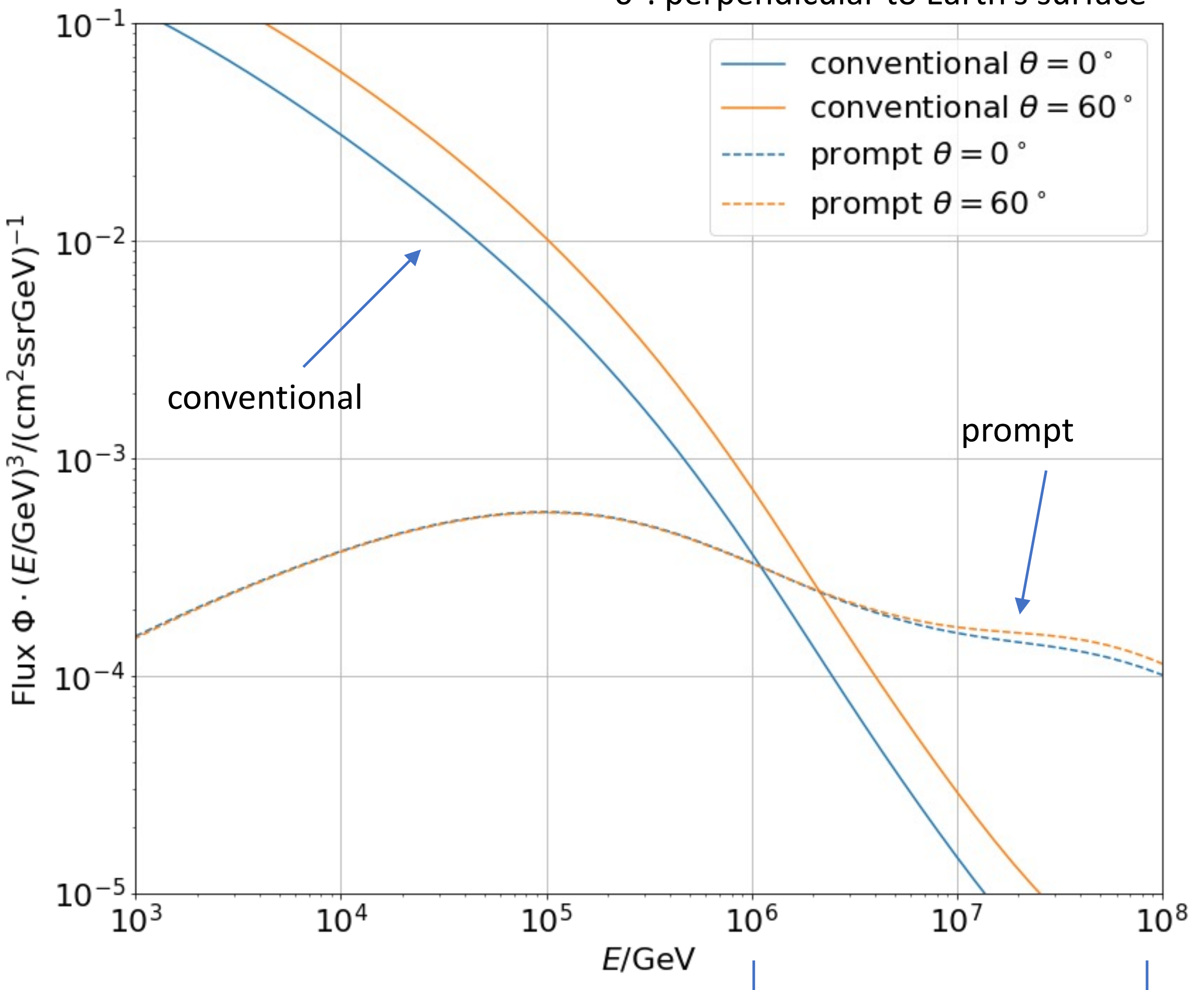
$$\underset{\propto E^{-2.7}}{p} \rightarrow \overbrace{D^{+}}^{\tau = 1.04\cdot 10^{-12}\,\mathrm{s}} \rightarrow \bar{K}^{0} + \underset{\propto E^{-2.7}}{\mu^{+}} + \nu_{\mu}$$

Muon flux

$$\Phi_{\text{tot}} = \Phi_{\text{conv}} + \Phi_{\text{prompt}}$$

- Prompt dominates at energies larger than PeV
- Conventional particle flux depends on zenith angle

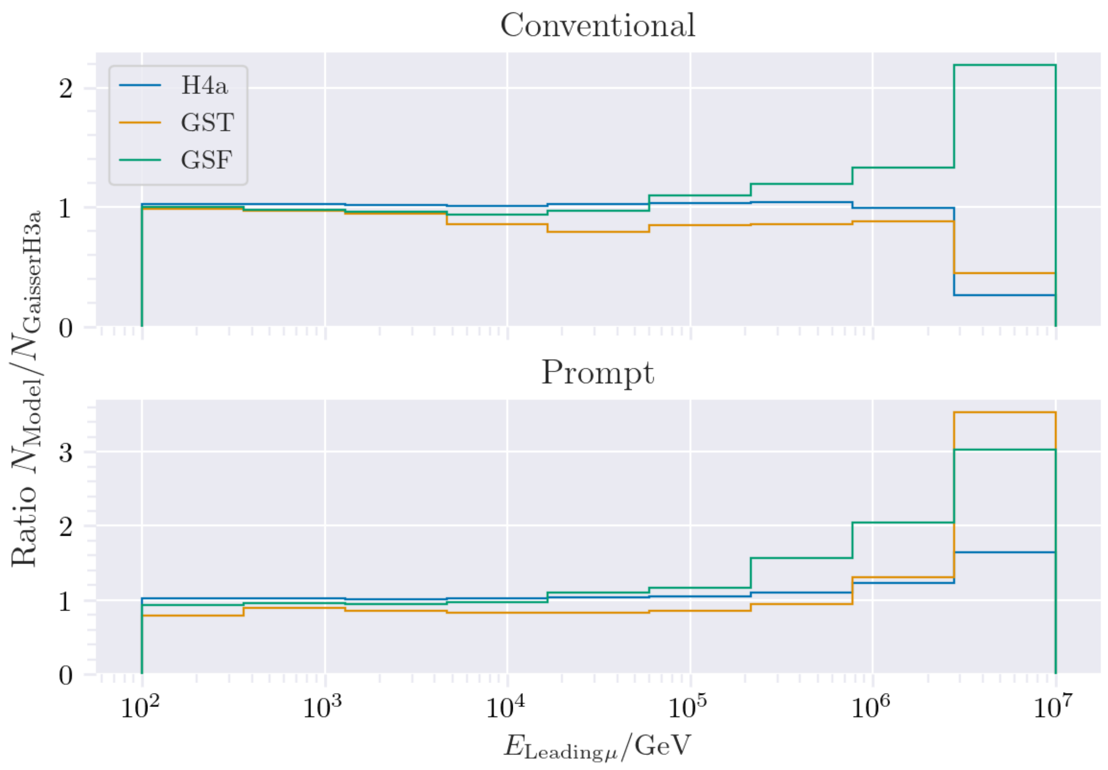
0°: perpendicular to Earth's surface



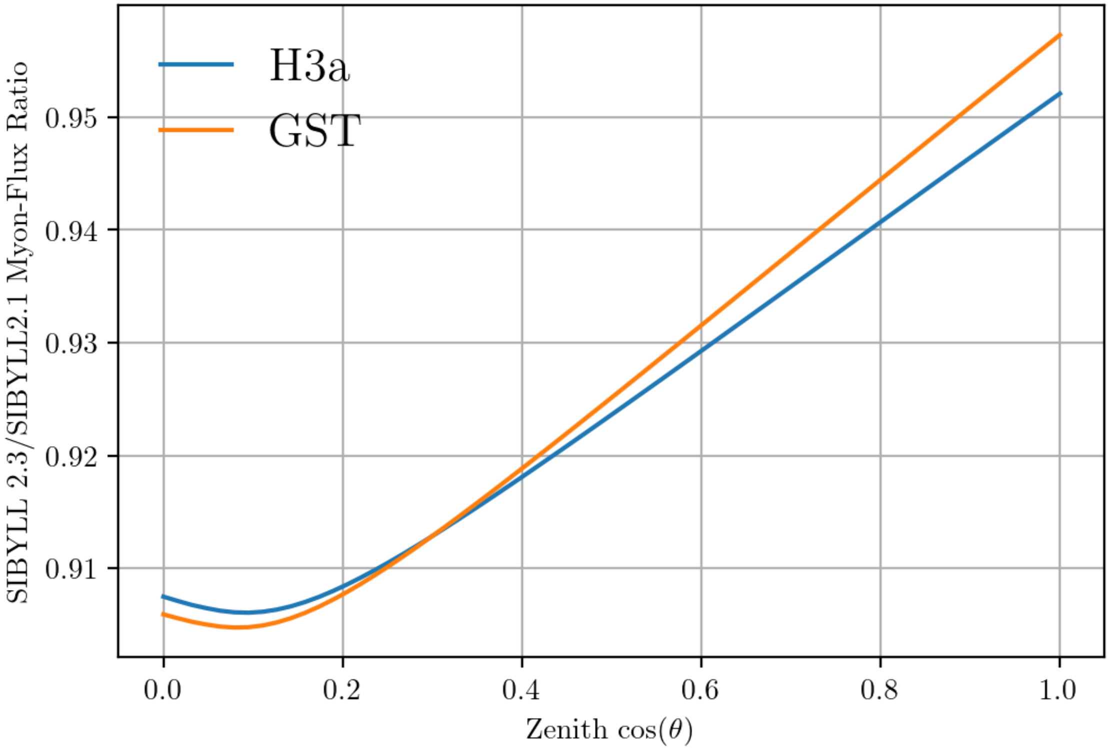
We can measure prompt muon energies from ~1 PeV to ~100 PeV

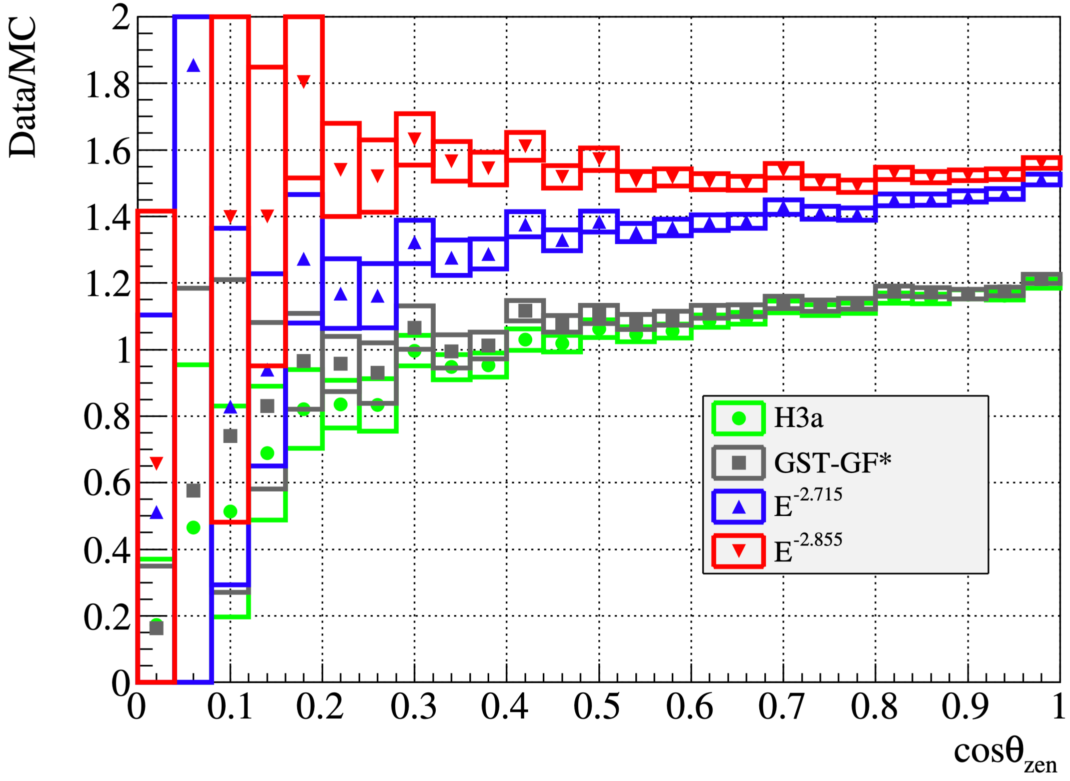
Muon production – different weightings

GST predicts most prompt



Solution to zenith problem?

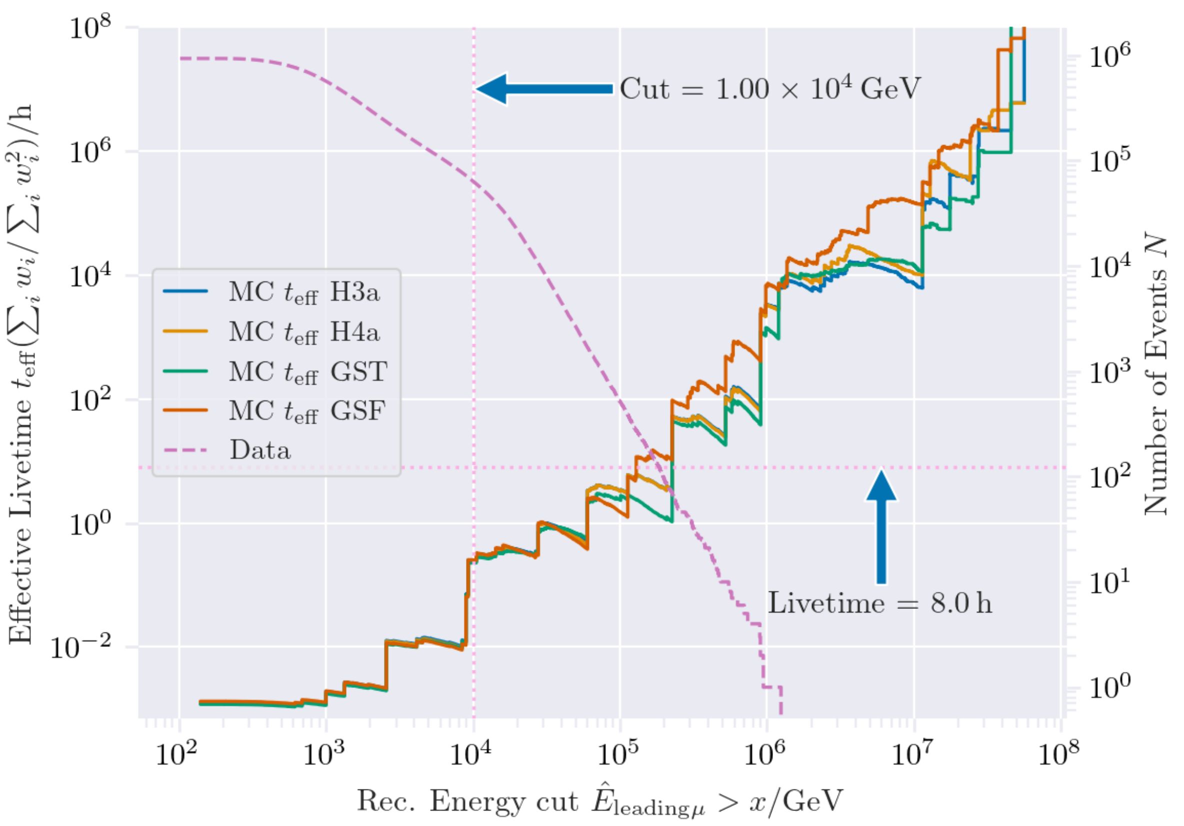




- ➢ No complete solution, but a step in the right direction

New CORSIKA extended history simulations

- CORSIKA 77420
- SIBYLL 2.3d
- Icetray 1.5.1
- 5 components (p, He, N, Al, Fe)
- Polyplopia: True
- Trimshower: True
- Ecuts1: 273 GeV (hadron min energy)
- Ecuts2: 273 GeV (muon min energy)
- Ecuts3: 10^{20} GeV (electron min energy)
- Ecuts4: 10^{20} GeV (photon min energy)
- 4 datasets:
    - 30010: 600 GeV – 1 PeV
    - 30011: 1 PeV – 100 PeV
    - 30012: 100 PeV – 1 EeV
    - 30013: 1 EeV – 50 EeV
- /data/sim/IceCube/2023/generated/CORSIKA_EHISTORY/

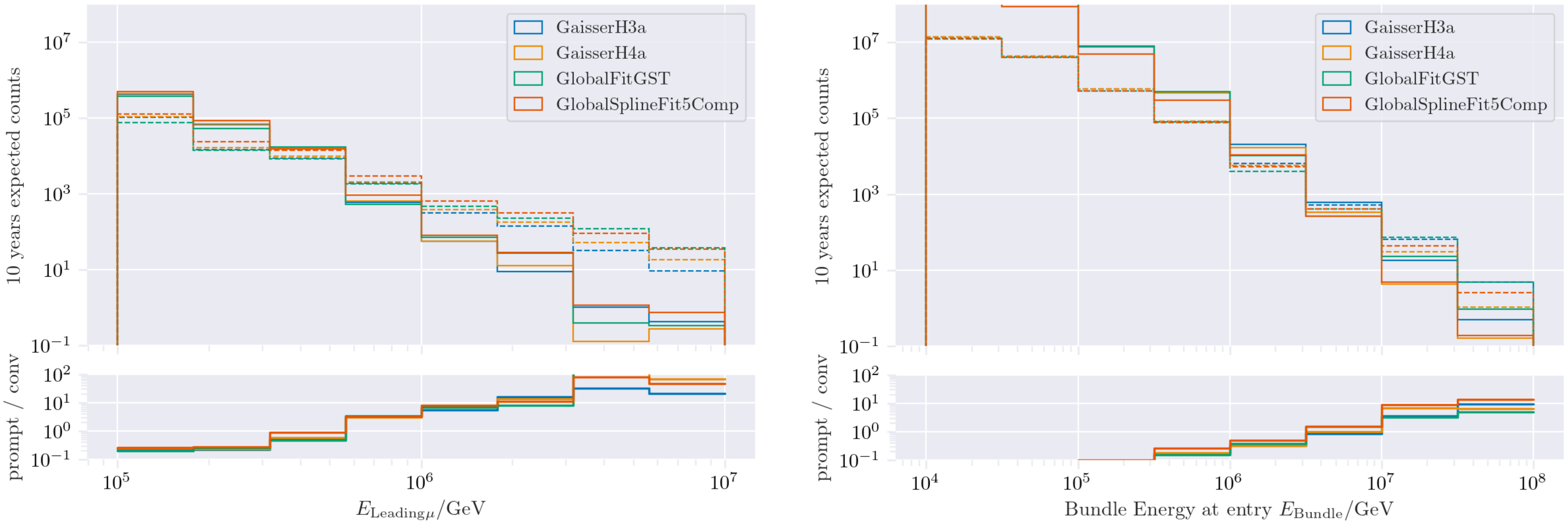


Please go ahead and test the datasets

- ➢ Sufficient statistics above 1 PeV
- ➢ Too few statistics at lower energies

Expected muons for 10 years: leading vs. bundle energy

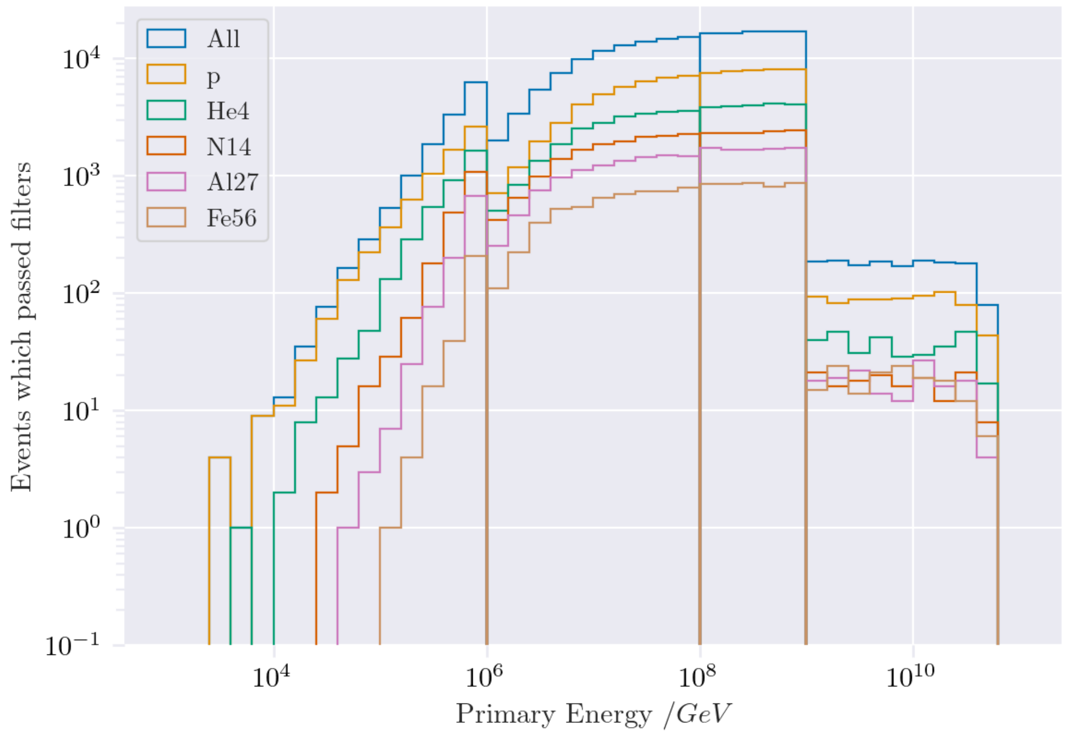
Solid: conventional
Dashed: prompt



- Different primary fluxes lead to different prompt fluxes
- Bundle energy extends to higher energies

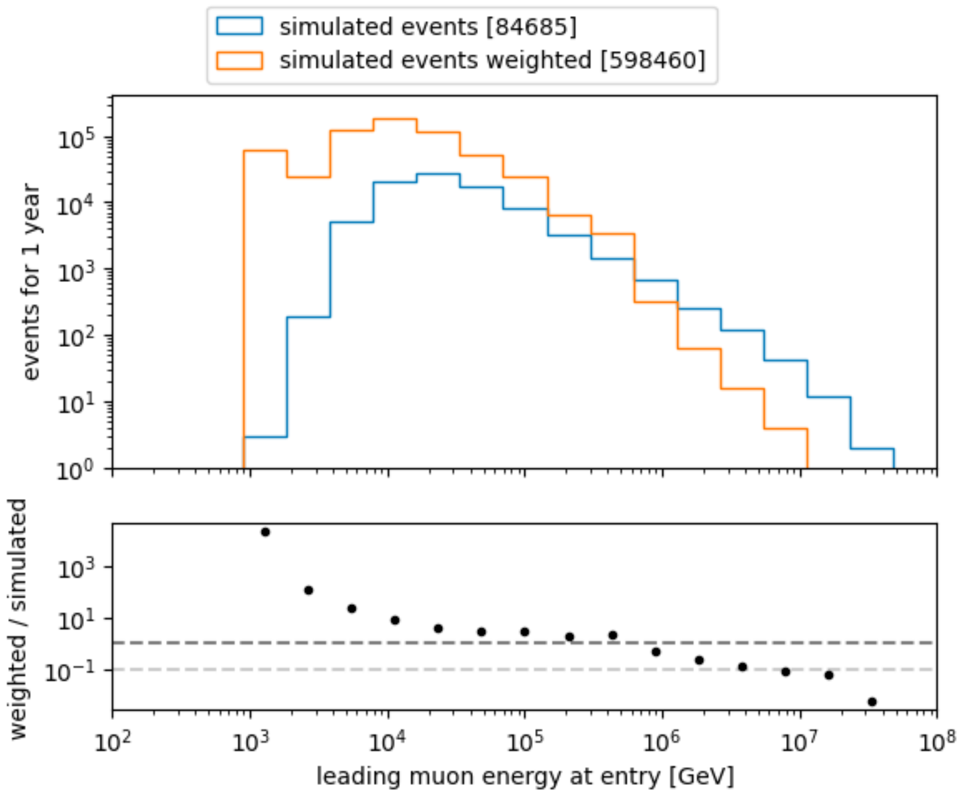
Simulated events

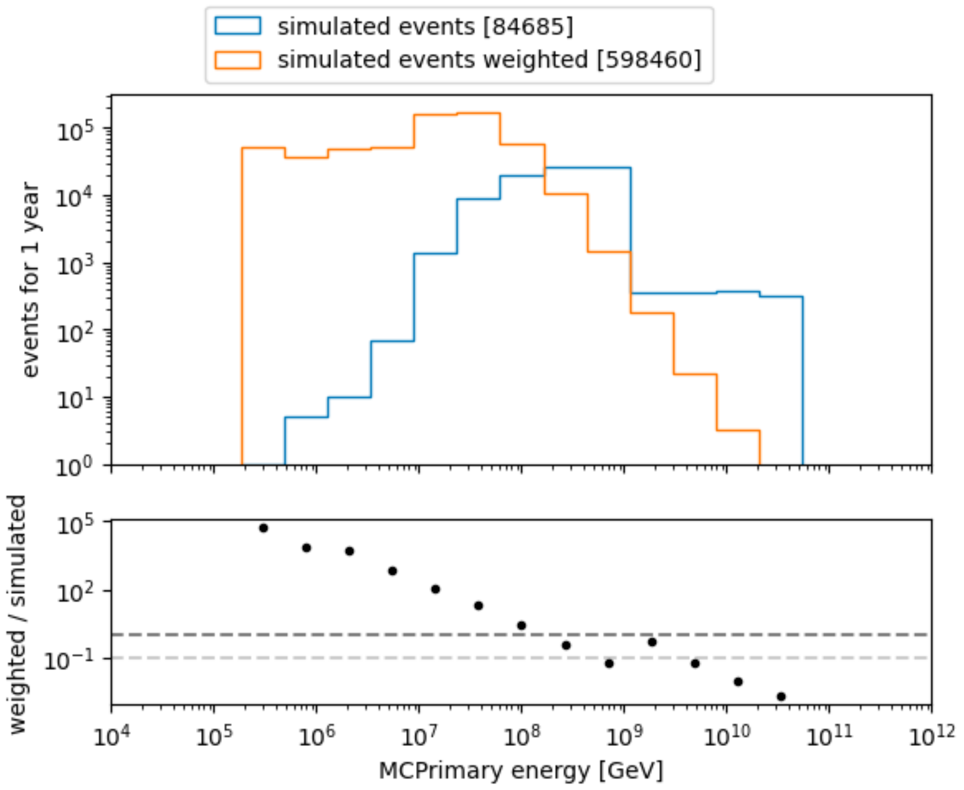
Passed muon filter



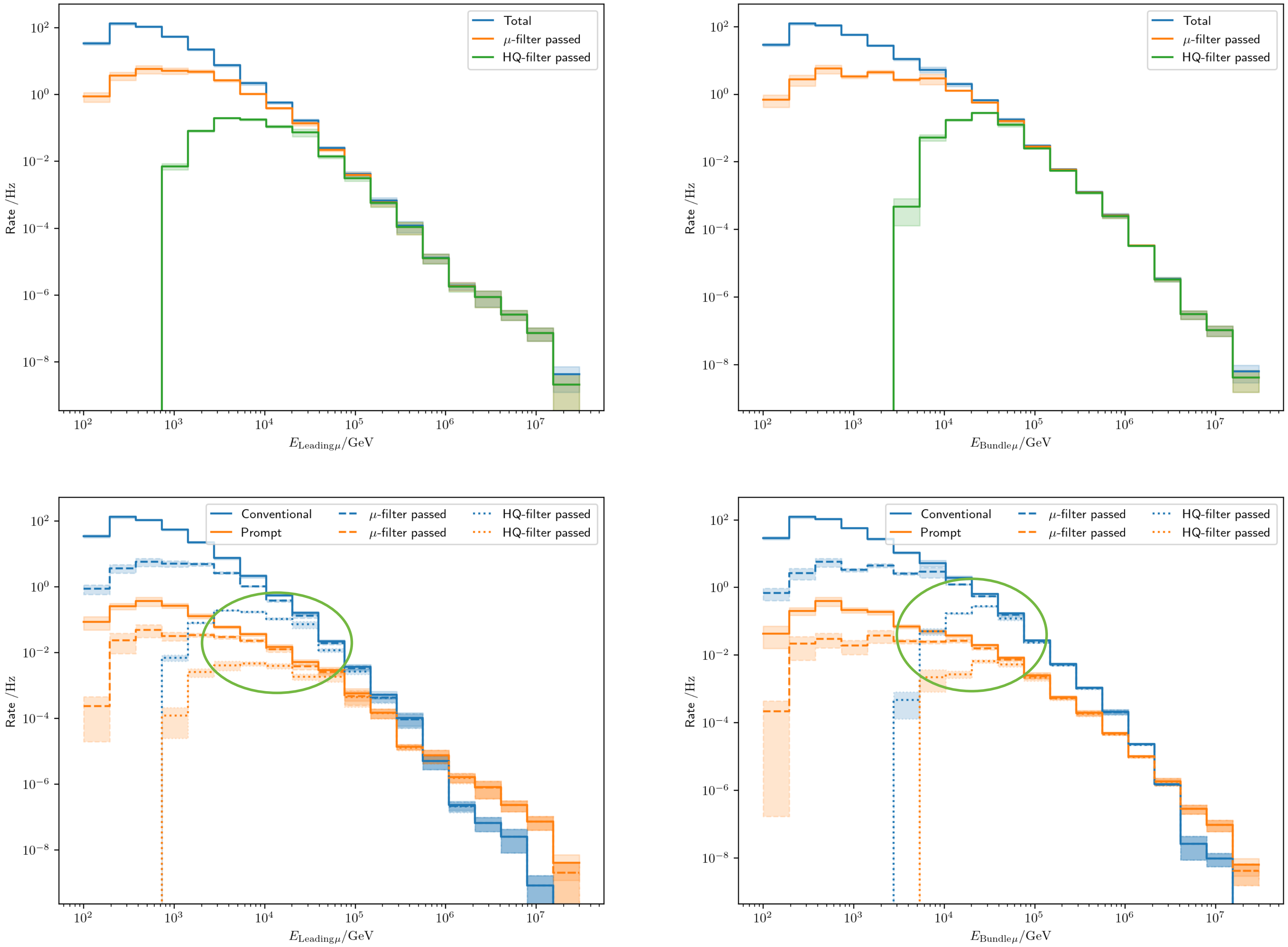
Which energies do we need to simulate?

L2 MuonFilter
Bundle energy > 100 TeV



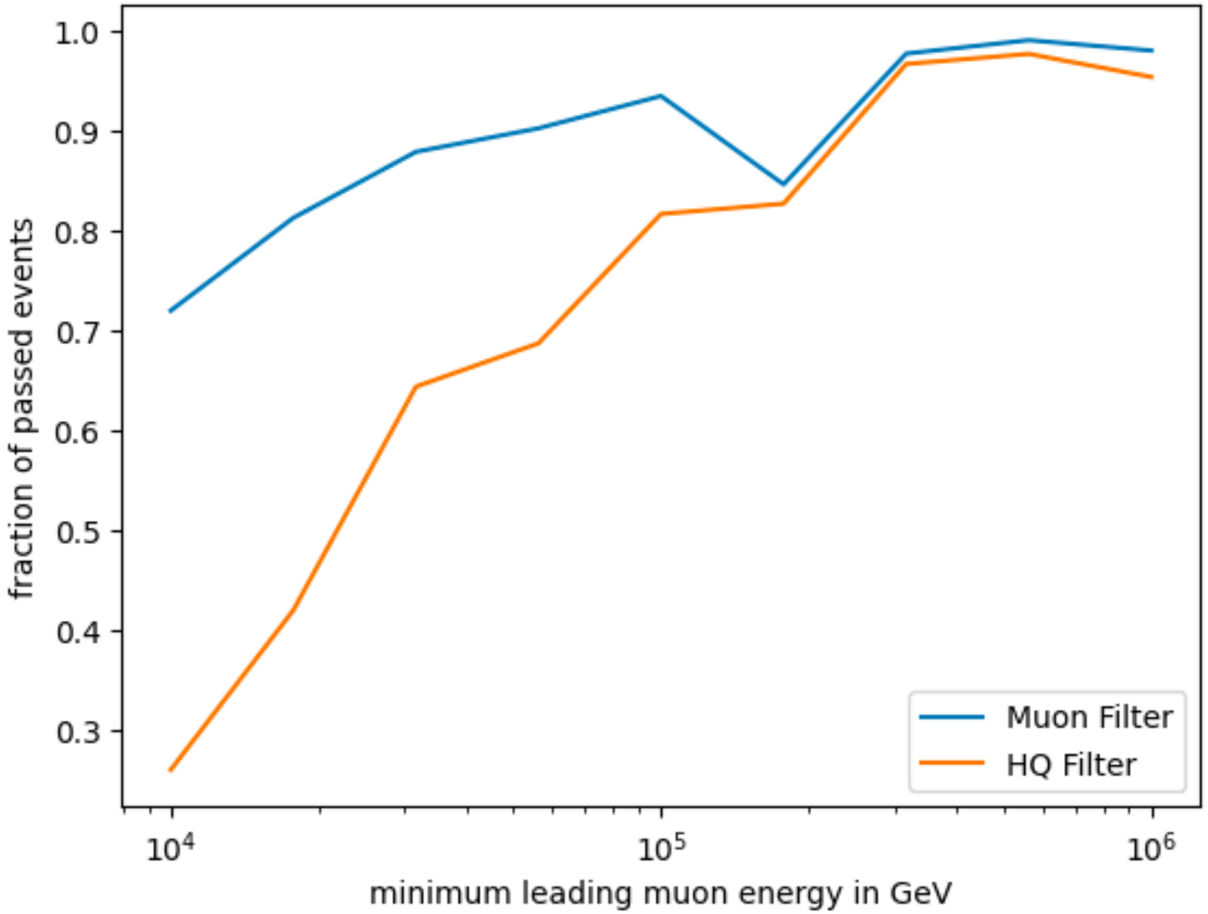


Filters

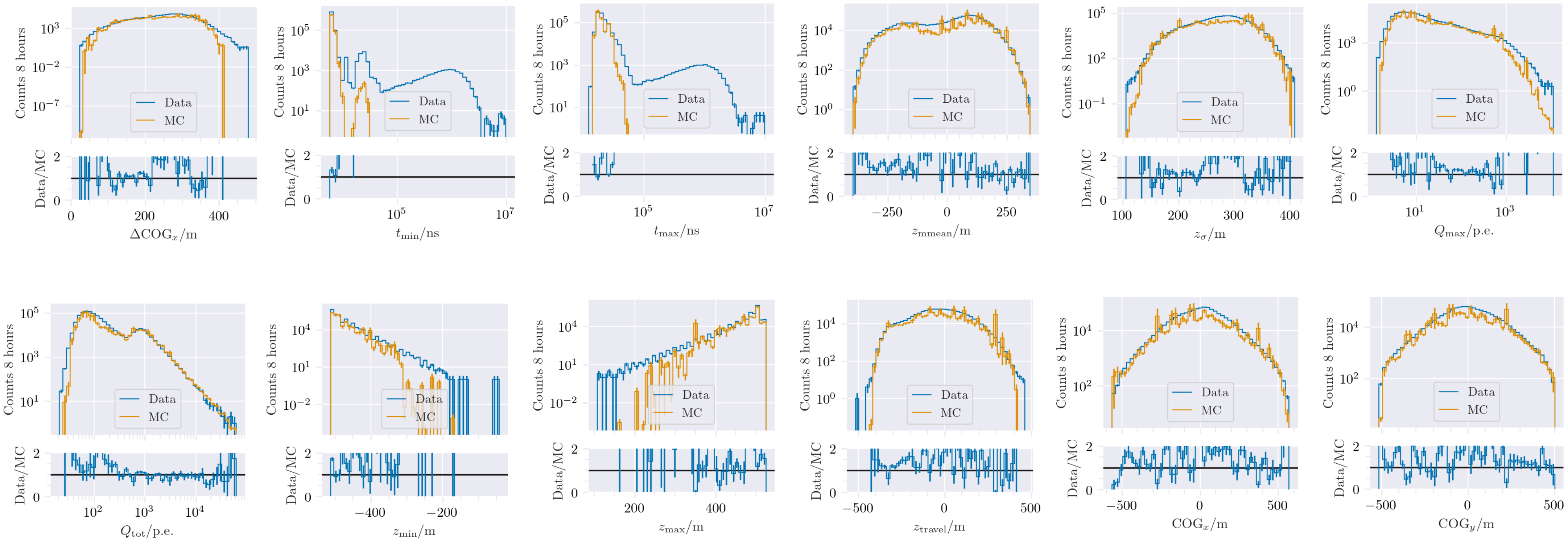


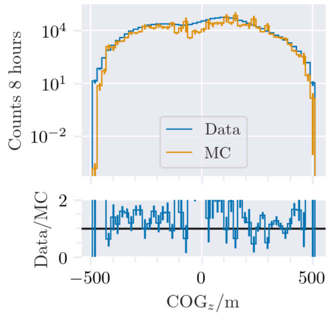
Choose muon filter, larger statistics at 10 TeV

Filters – passed events per leading energy



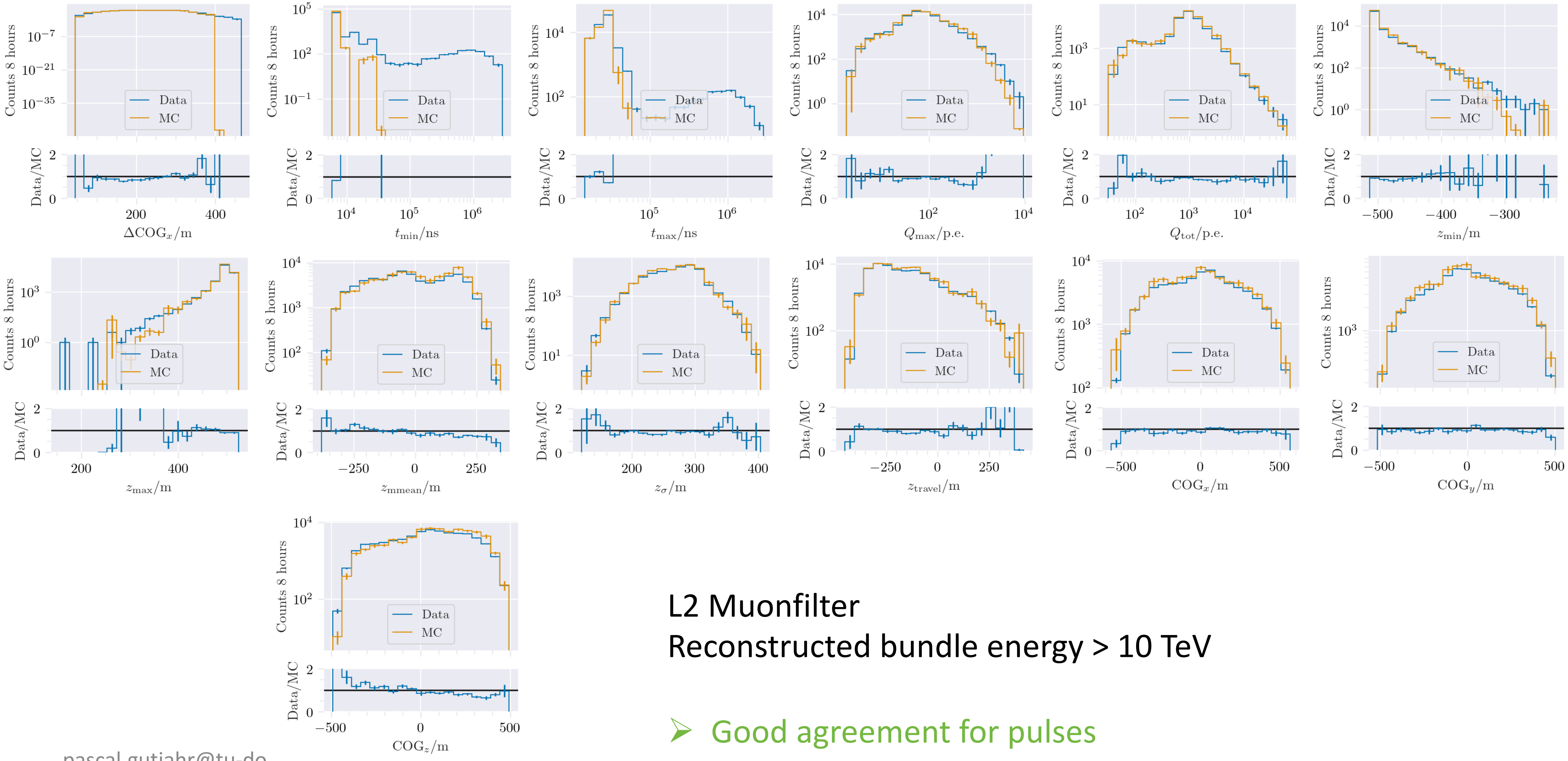
Data-MC: HitStatistics - SplitInIceDSTPulses





L2 MuonFilter
No further exclusions

Data-MC: HitStatistics- SplitInIceDSTPulses



L2 Muonfilter

Reconstructed bundle energy > 10 TeV

➢ Good agreement for pulses

Evaluation of models used in pseudo analysis

- Note: the DNN reconstruction models are still investigated! The models used here are in an early stage. So far, the two bachelor students have trained models with better performance.
- Updates on the models and reconstructions are provided in the near future

Data-MC agreements

- Two bachelor students worked on reconstructions using the dnn_reco framework (thesis available in english):
    - Leander Flottau
    - Benjamin Brandt

Reconstructions:

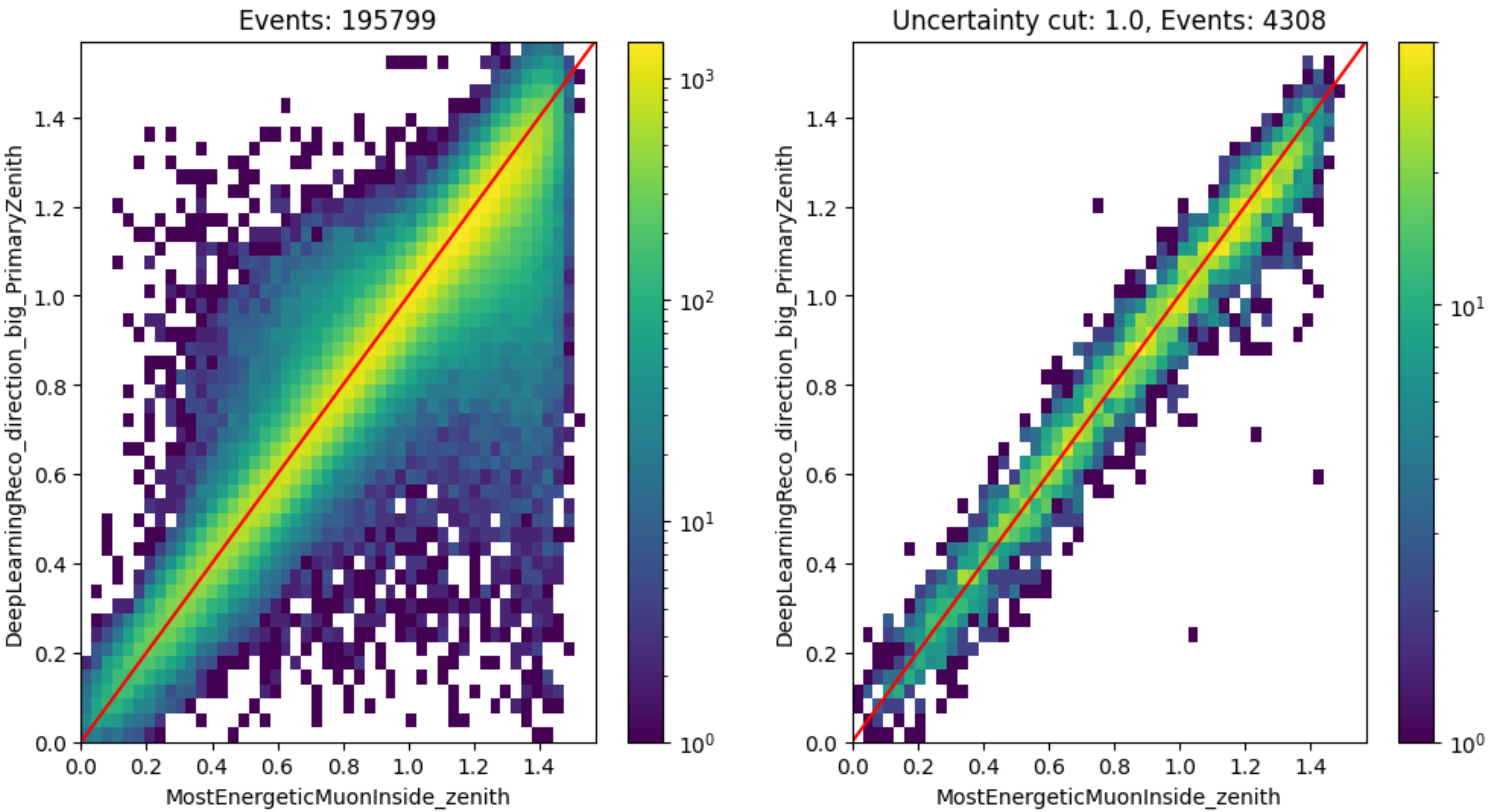
- Leading energy
- Leading fraction
- Bundle energy
- Multiplicity
- Azimuth
- Zenith

Work in progress

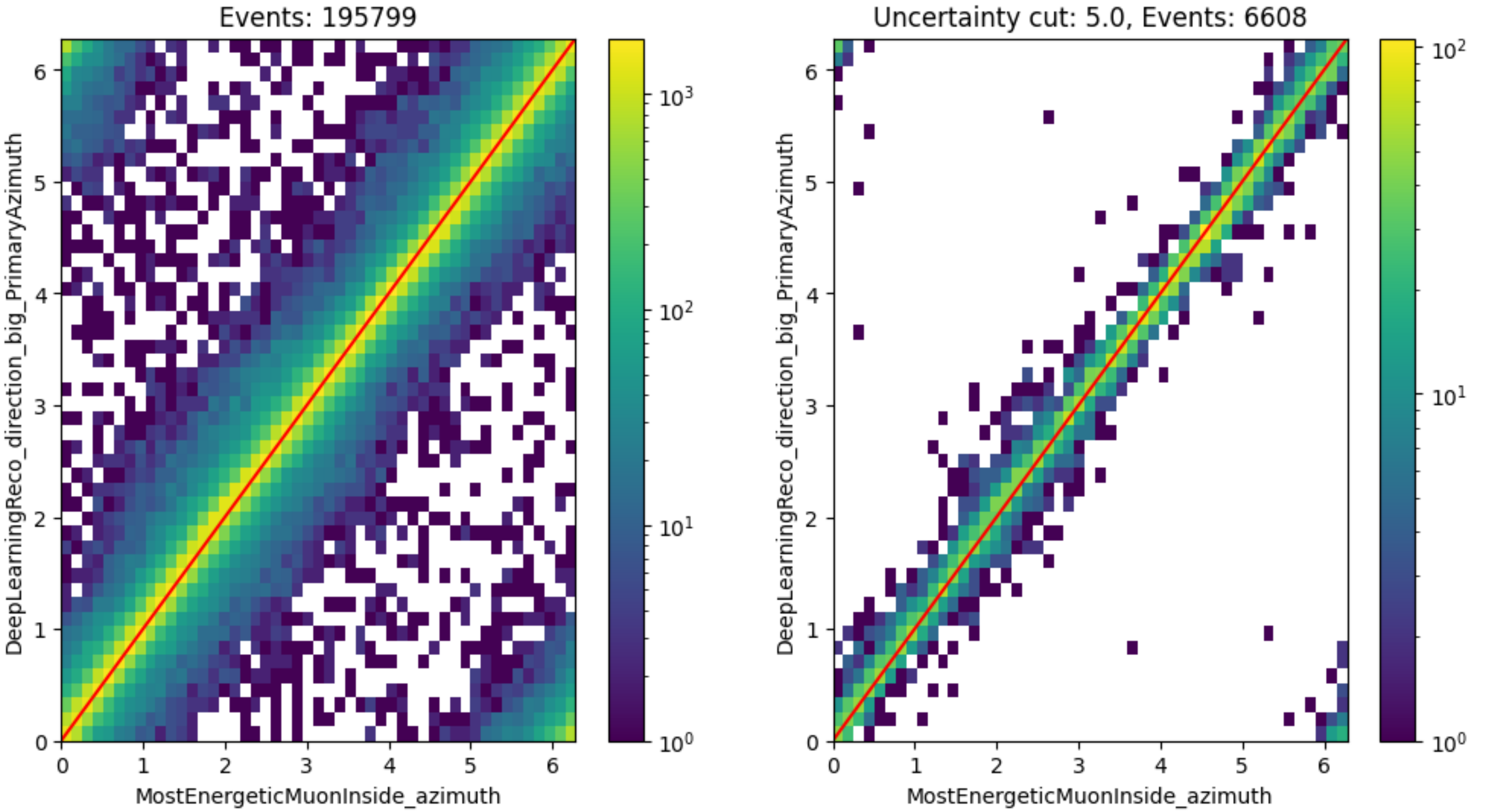
General:

- Trained on uncleaned and 6µs cleaned muon pulses
- Processed 8h of experimental data (June 4th, 2020)

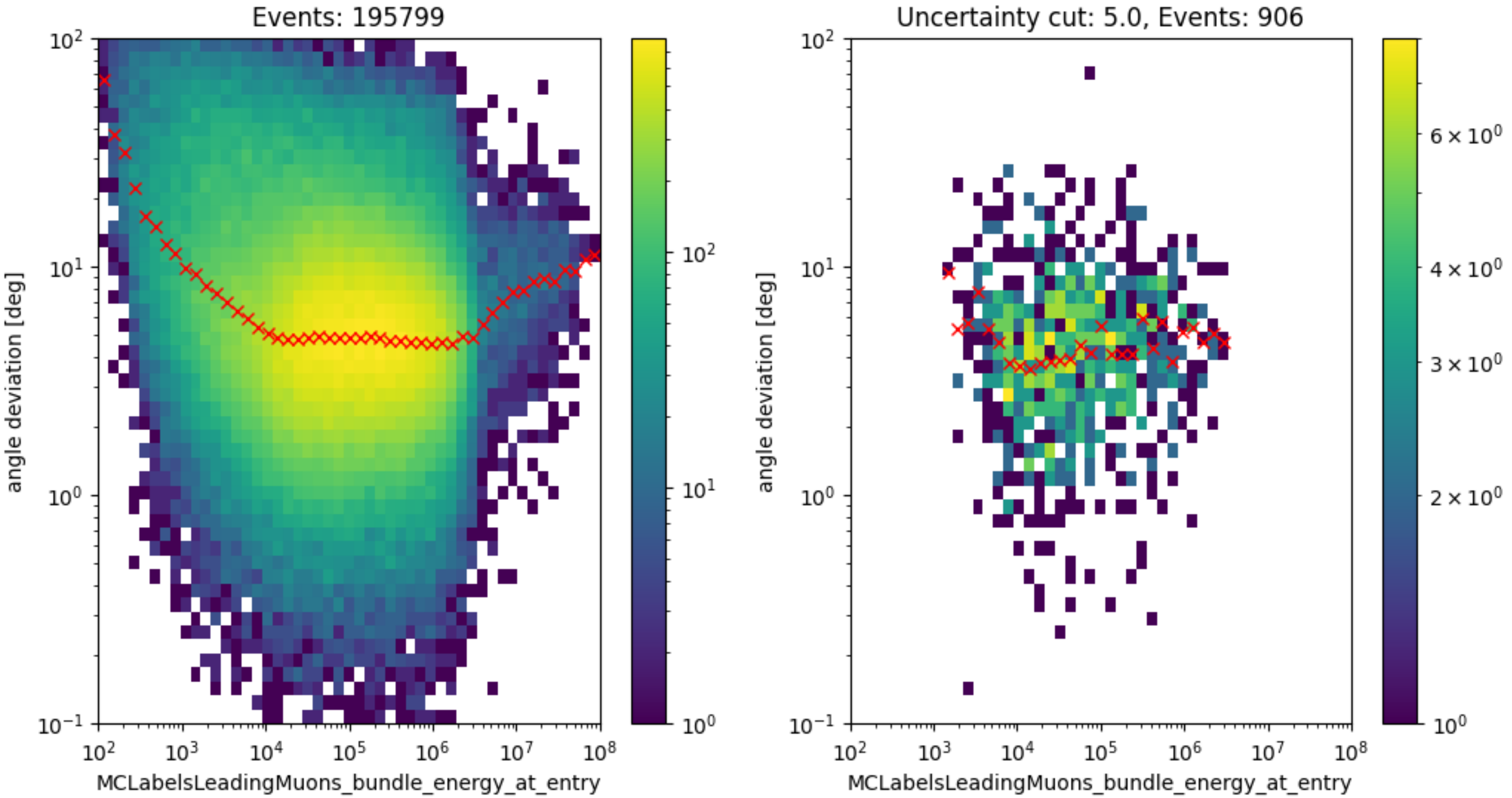
Zenith reconstructions



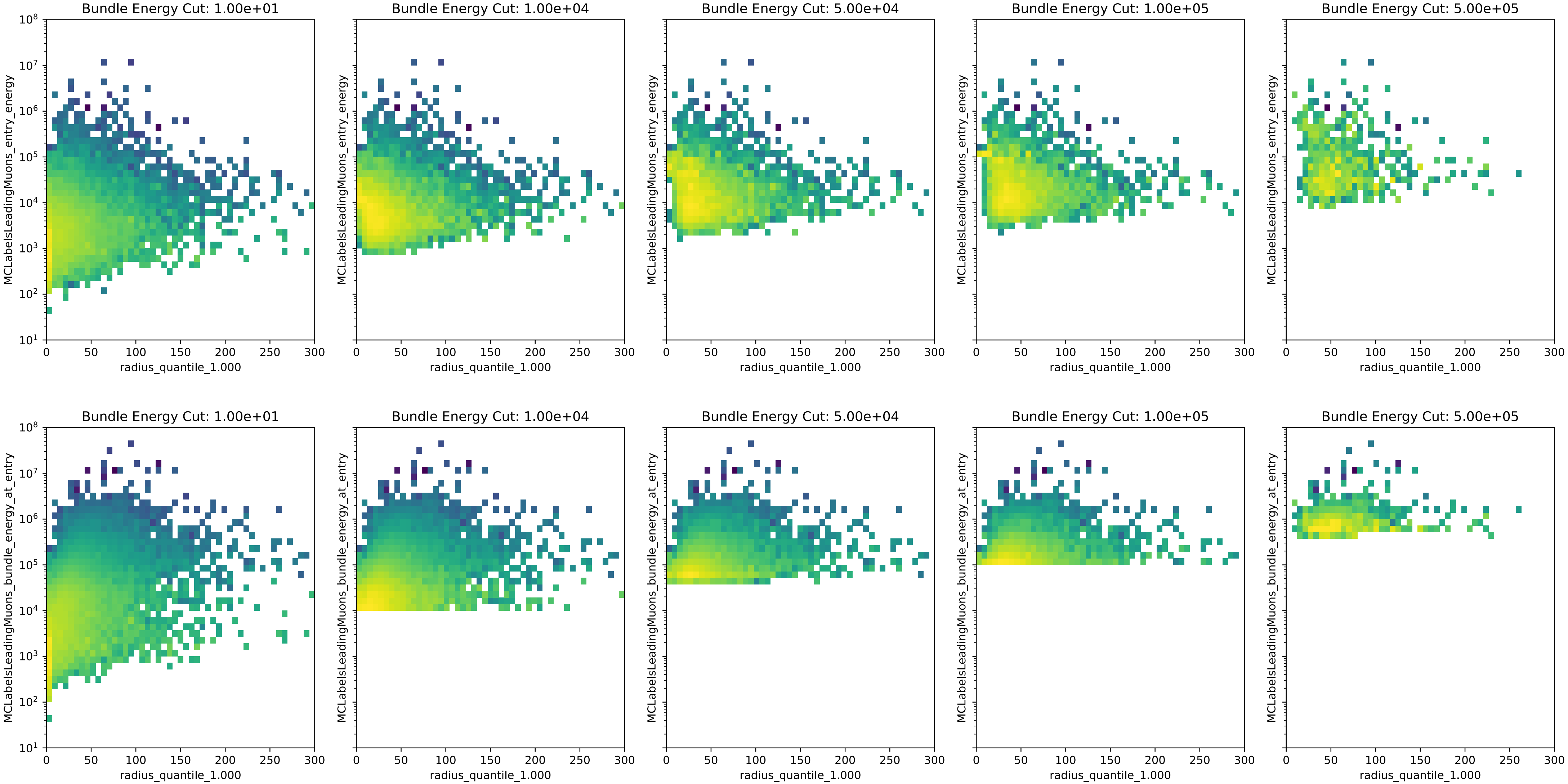
Azimuth reconstructions



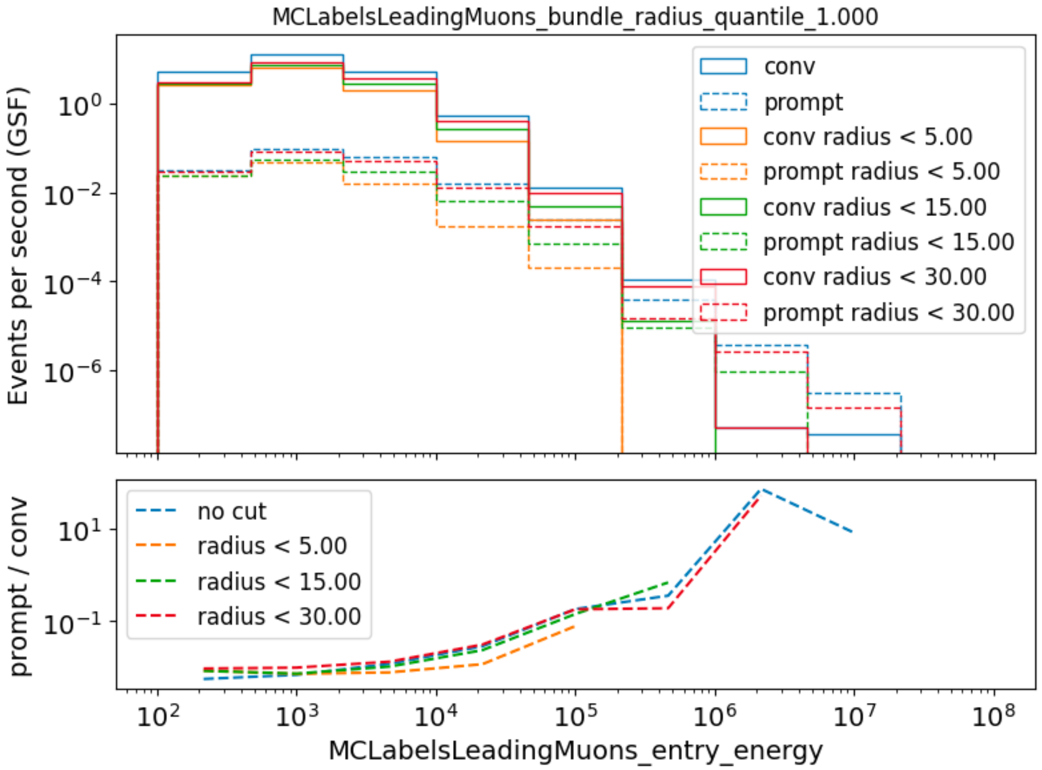
Angular resolution



Bundle radius - energy



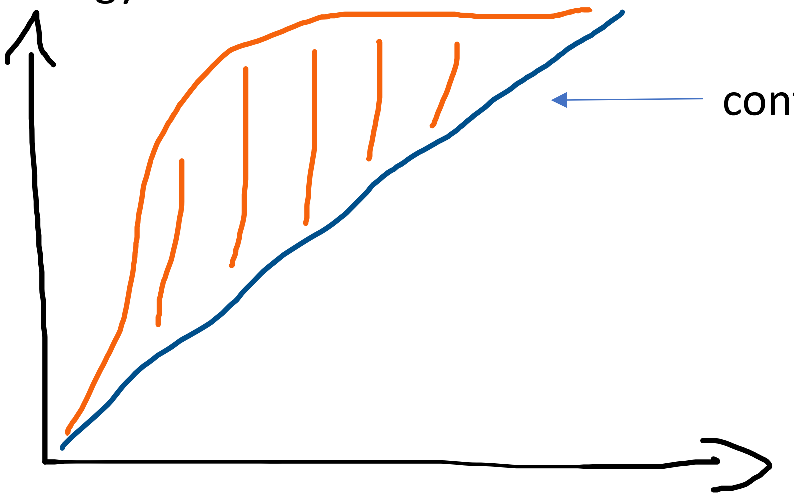
Muon flux – bundle radius



Stochasticity

muon that loses much energy at the beginning of the propagation

cumulated energy loss



continuous energy loss

track length

- calculate the area between the "continuous muon energy loss" and the accumulated muon energy losses and normalize it
- high stochasticity: area = 1
- no stochasticity: area = 0

Bundle radius

- calculate perpendicular distance between leading muon and closest point to detector center
  - reference point
- calculate the distance between the leading muon and all the other muons at the reference point
- analyze the distances weighted by their energy (100%, 99%, 95%,... energy containment)