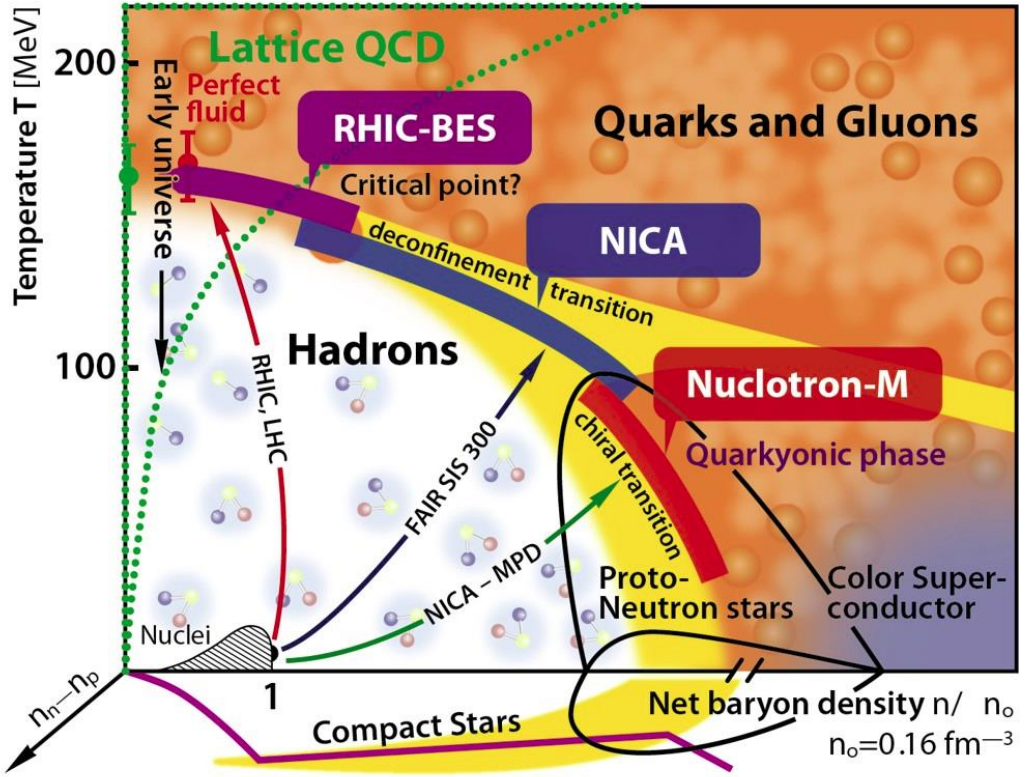
Байесовский подход в определении центральности в ядро-ядерных в столкновениях на ускорительном комплексе NICA

Идрисов Дим, Губер Федор, Парфенов Петр, Карпушкин Николай

ИЯИ РАН, Москва

10/04/25

Ускорительный комплекс NICA



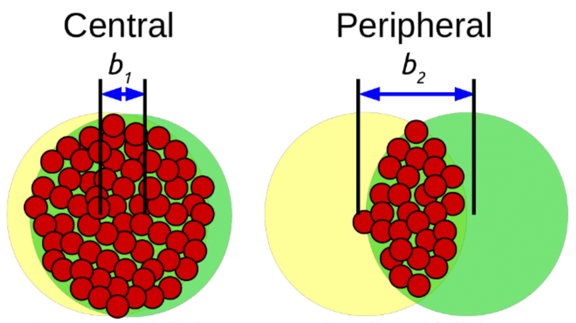
- **Высокие энергии ($\sqrt{s_{NN}}$>100 GeV):**
  - Высокая T, $\mu_B \approx 0$, Эволюция ранней вселенной
- **Низкие энергии (2.4<$\sqrt{s_{NN}}$<11 GeV):**
  - Средние T, высокая μ_B, внутренняя структура компактных звезд, слияния нейтронных звезд

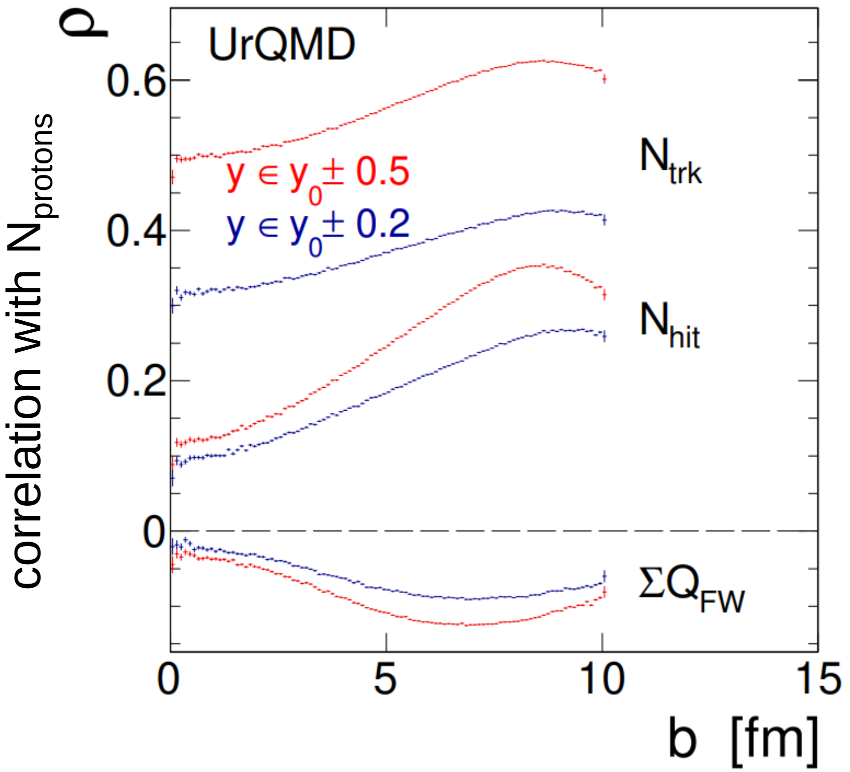
- **Baryonic Matter at Nuclotron (BM@N) –** эксперимент с фиксированной мишенью, первый физический сеанс Xe+CsI 2022-2023
- **Multi-Purpose Detector (MPD) –** запуск 2025-2026
- **Spin Physics Detector (SPD) –** возможность работать с поляризованными пучками дейтронов

Centrality

- Evolution of matter produced in heavy-ion collisions depend on its initial geometry
- Centrality procedure maps initial geometry parameters with measurable quantities (multiplicity or energy of the spectators)
- **This allows comparison of the future BMAN results with the data from other experiments (STAR BES, NA49/NA61 scans) and theoretical models**

$$c(b) = \frac{\int_0^b \frac{d\sigma}{db'} db'}{\int_0^\infty \frac{d\sigma}{db'} db'} = \frac{1}{\sigma_{A-A}} \int_0^b \frac{d\sigma}{db'} db'$$

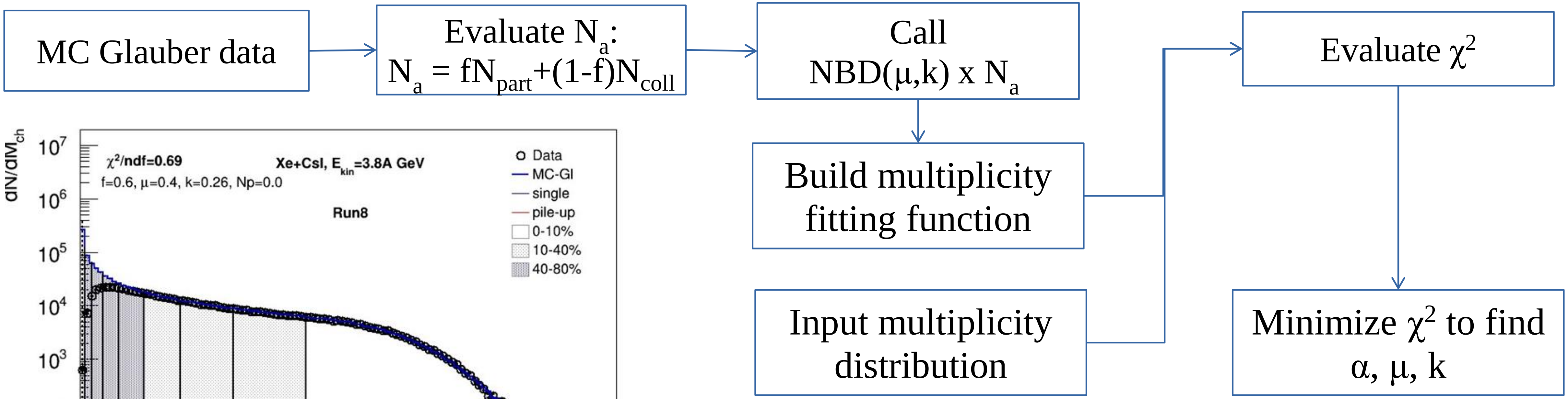


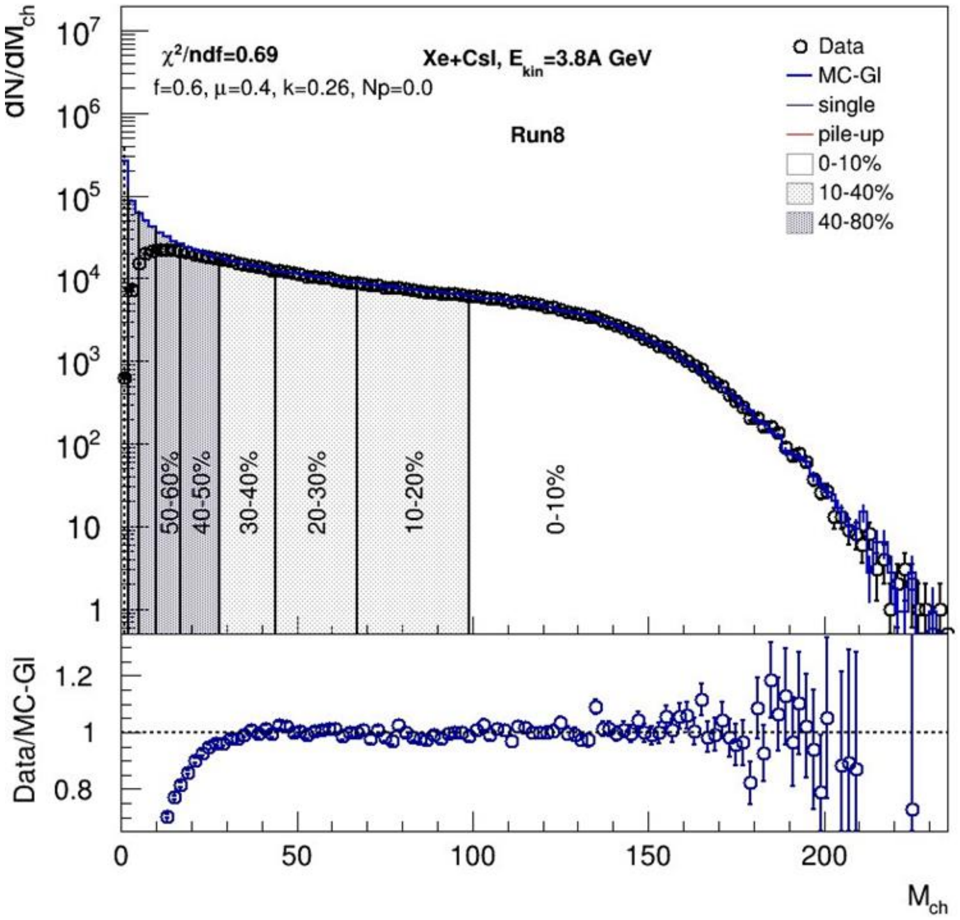


HADES; Phys.Rev.C 102 (2020) 2, 024914

- A number of produced protons is stronger correlated with the number of produced particles (track & RPC+TOF hits) than with the total charge of spectator fragments (FW)
- to suppress self-correlation biases, it is necessary to use spectators fragments for centrality estimation

MC-Glauber based centrality framework



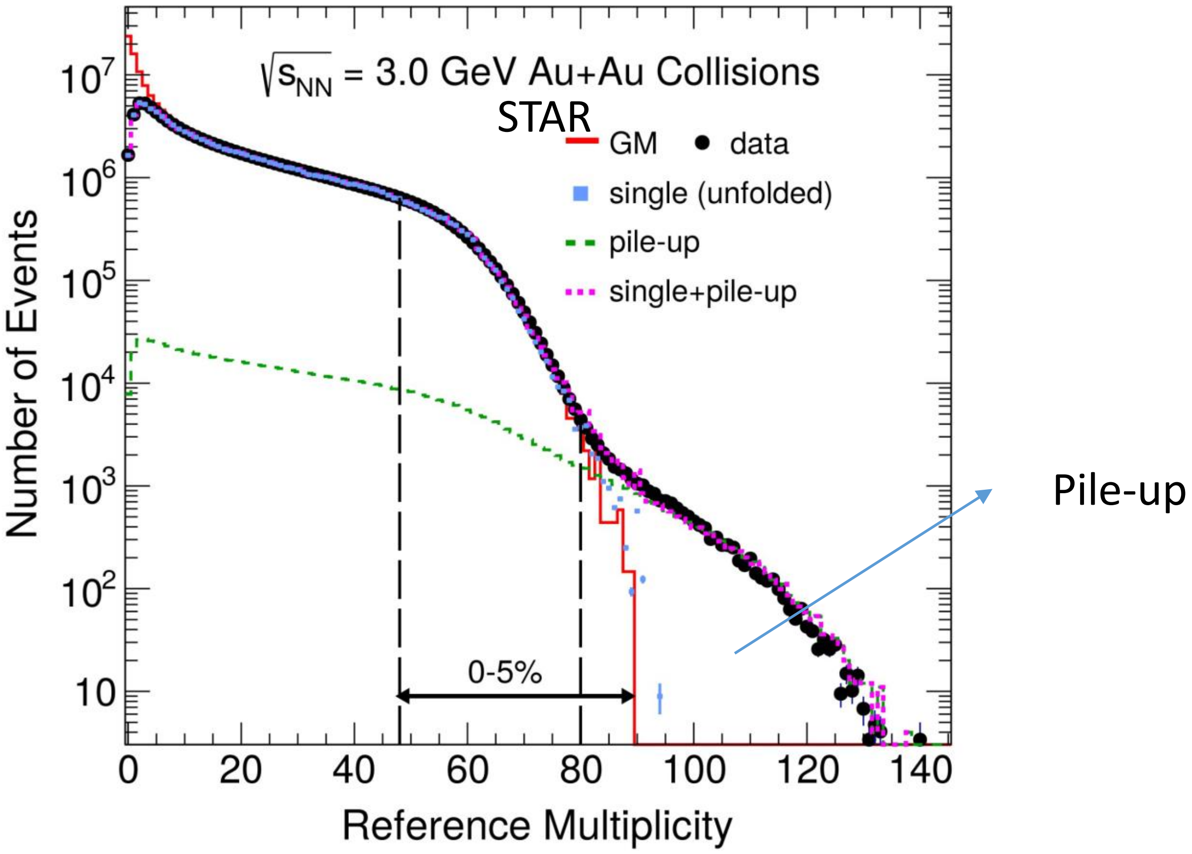


NBD – negative binomial distribution
Parameters of the fit:

- **f** – the fraction of multiplicity from the soft component
- **μ** – mean multiplicity value
- **k** – width of the multiplicity distribution, can be connected to the fluctuations

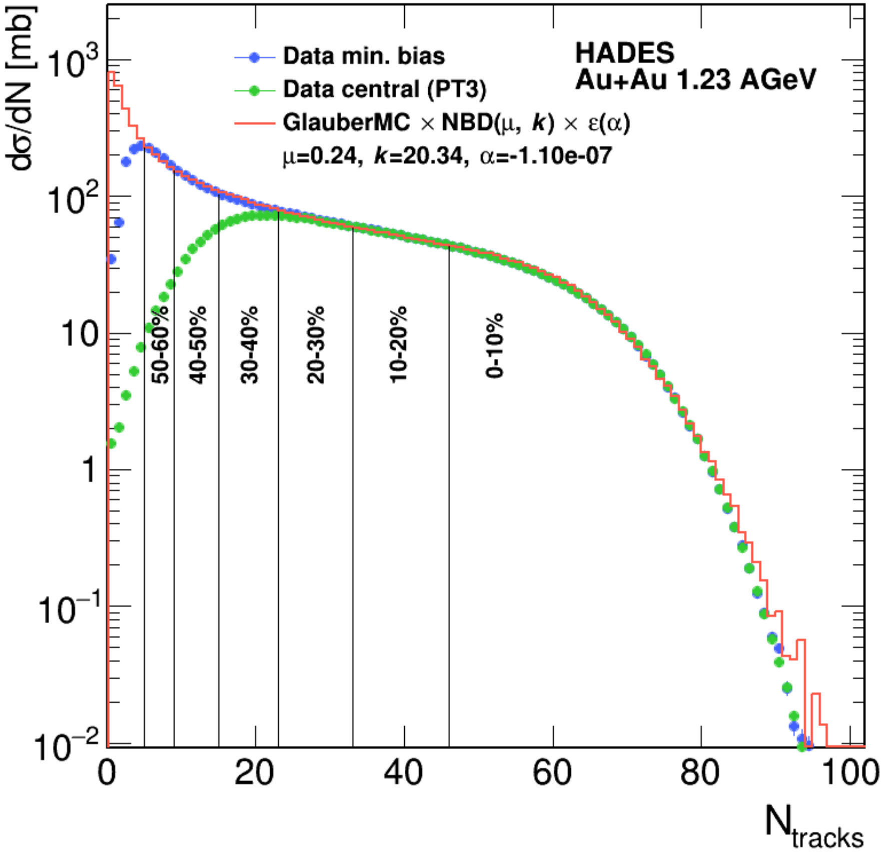
Implemantation for MPD: https://github.com/FlowNICA/CentralityFramework
P. Parfenov, et al., Particles. 2021; 4(2):275-287

Centrality determination in the FIX-target experiments



Reference multiplicity distributions (black markers) in the kinematic acceptance within −0.5 < y < 0 and 0.4 < pT < 2.0 GeV/c ,
GM (red histogram), and single and pile-up contributions from unfolding.

https://arxiv.org/abs/2112.00240



The cross section as a function of Ntracks for minimum bias (blue symbols) and central (PT3 trigger, green symbols) data in comparison with a fit using the Glauber MC model (red histogram).

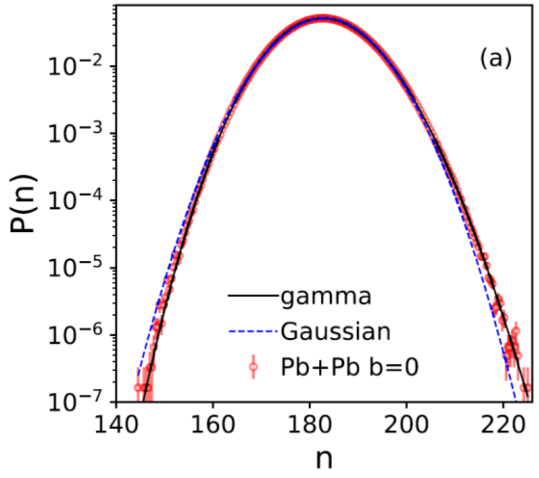
https://arxiv.org/abs/1712.07993

The Bayesian inversion method (Γ-fit): main assumptions

- Relation between multiplicity N_{ch} and impact parameter b is defined by the fluctuation kernel:

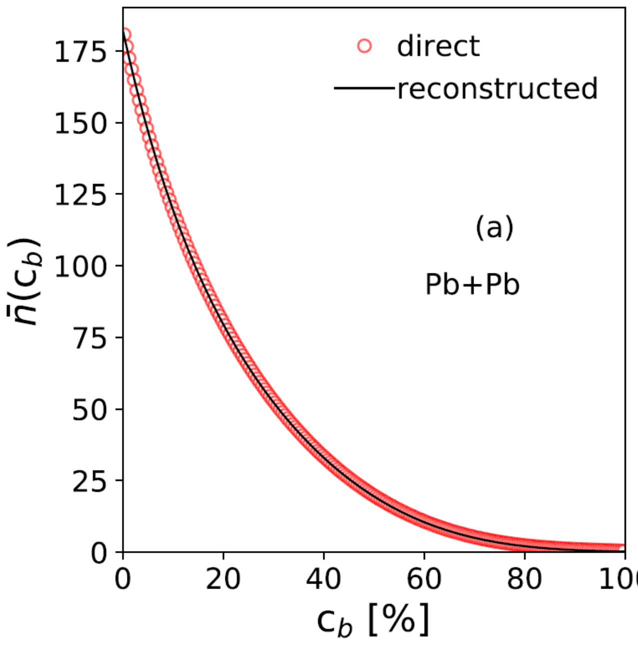
$$P(N_{ch}|c_b) = \frac{1}{\Gamma(k(c_b))\theta^k} N_{ch}^{k(c_b)-1} e^{-n/\theta}$$

$$c_b = \int_0^b P(b')db' \simeq \frac{\pi b^2}{\sigma_{inel}}$$ – centrality based on impact parameter



The results of fitting the multiplicity distribution for a fixed impact parameter TRENTo model, $\sqrt{S_{nn}} = 5.02\ TeV$

R. Rogly, G. Giacalone and J. Y. Ollitrault, Phys.Rev. C98 (2018) no.2, 024902



The dependence of the average value of multiplicity on centrality and the results of its fit

$$\frac{\sigma^2}{\langle N_{ch} \rangle} = \theta \simeq const$$

$$\langle N_{ch} \rangle = N_{knee} \exp\left(\sum_{j=1}^{3} a_j c_b^j \right)$$

$$, k = \frac{\langle N_{ch} \rangle}{\theta}$$

Five fit parameters N_{knee}, θ, a_j

Reconstruction of *b*

- Normalized multiplicity distribution $P(N_{ch})$

$$P(N_{ch}) = \int_0^1 P(N_{ch}|c_b)dc_b$$

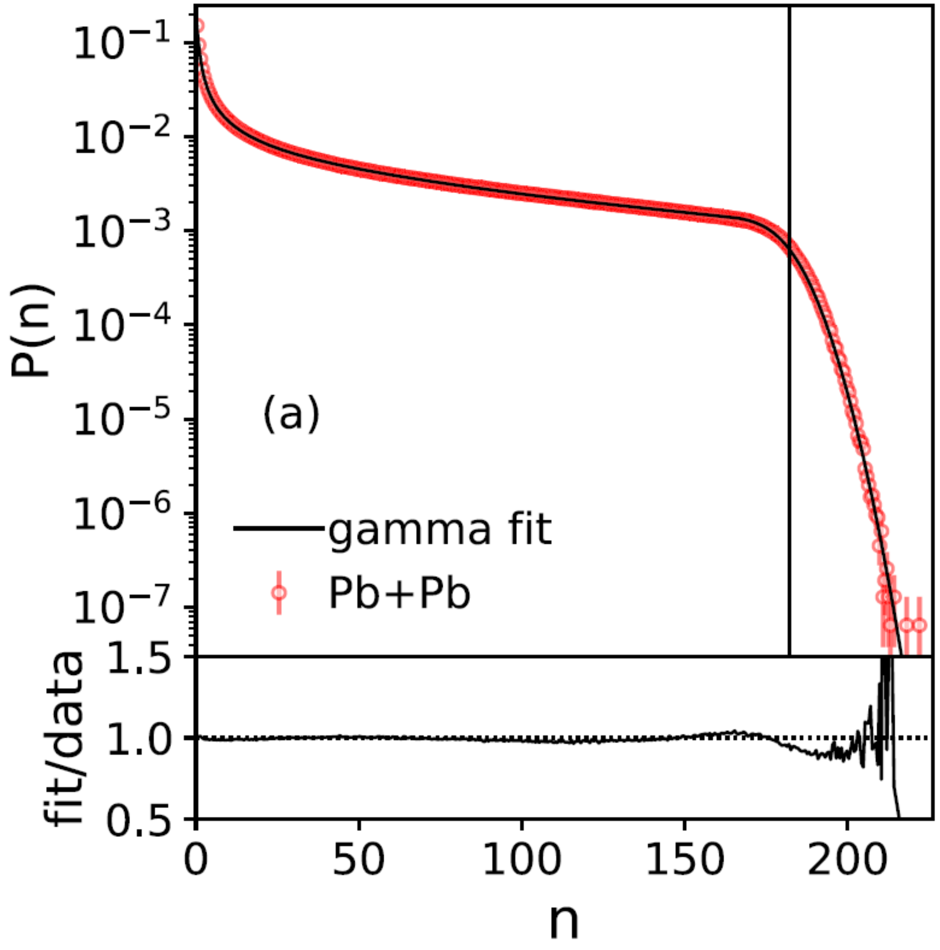
- Find probability of *b* for fixed range of N_{ch} using Bayes' theorem:

$$P(b|n_1 < N_{ch} < n_2) = P(b)\frac{\int_{n_1}^{n_2} P\,(b|N_{ch})dN_{ch}}{\int_{n_1}^{n_2} P\,(N_{ch})dN_{ch}}$$

- **The Bayesian inversion method consists of 2 steps:**

–Fit normalized multiplicity distribution with $P(N_{ch})$

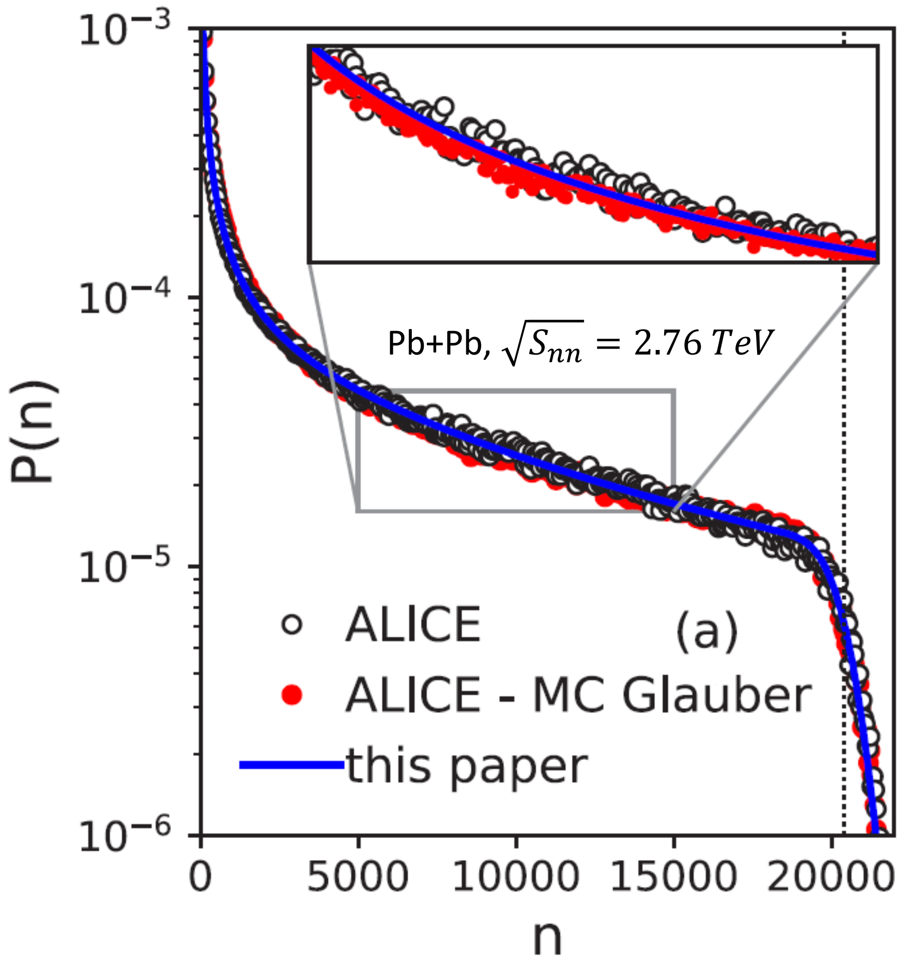
–Construct $P(b|N_{ch})$ using Bayes' theorem with parameters from the fit



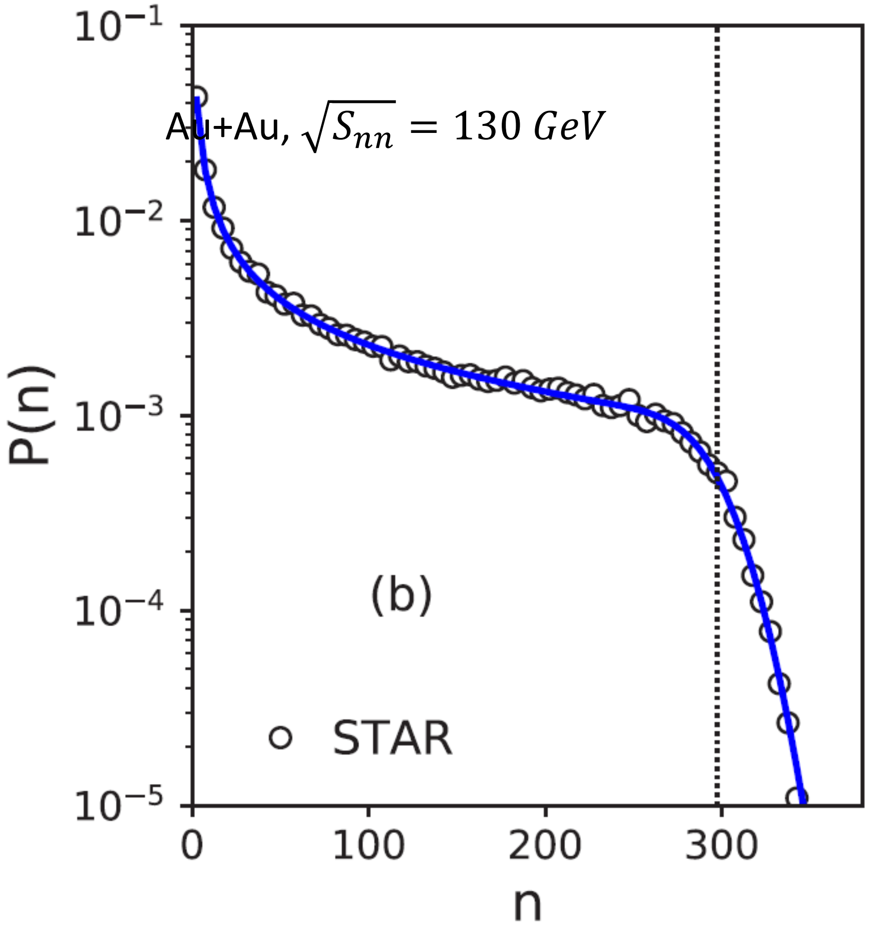
TRENTo model, $\sqrt{S_{nn}} = 5.02\ TeV$

R. Rogly, G. Giacalone and J. Y. Ollitrault, Phys.Rev. C98 (2018) no.2, 024902

Application to experimental data



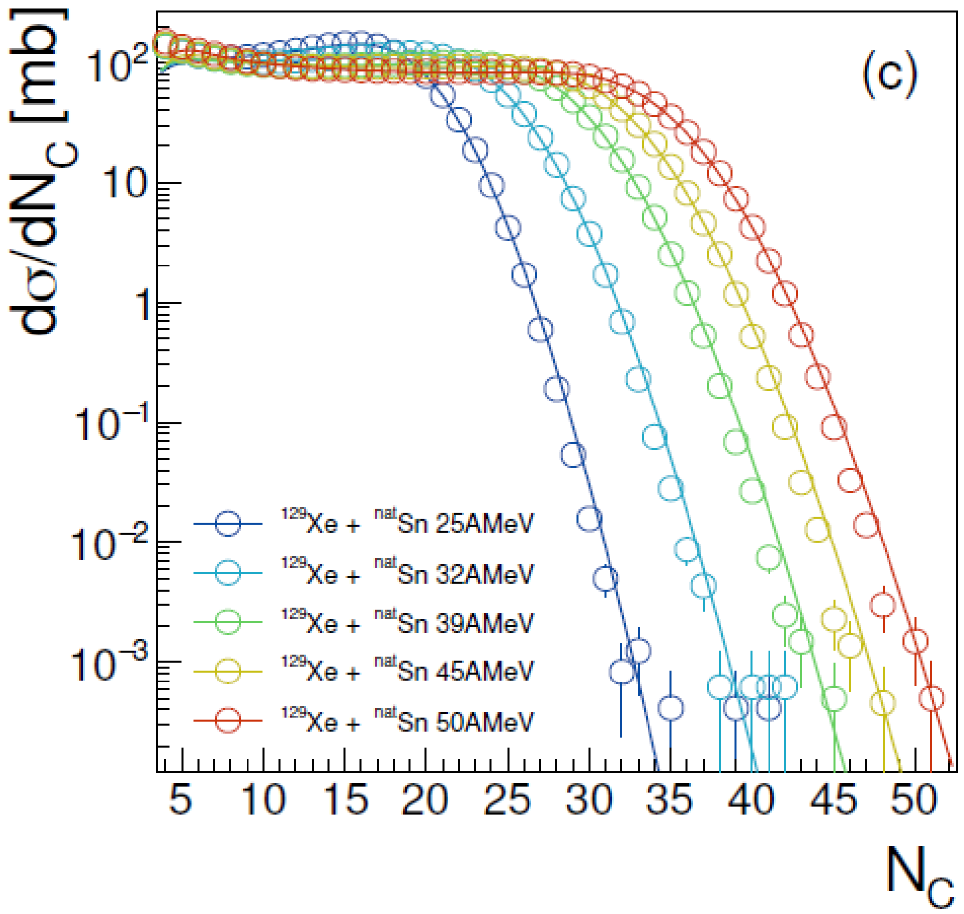
Empty symbols show distribution of the VZERO amplitude

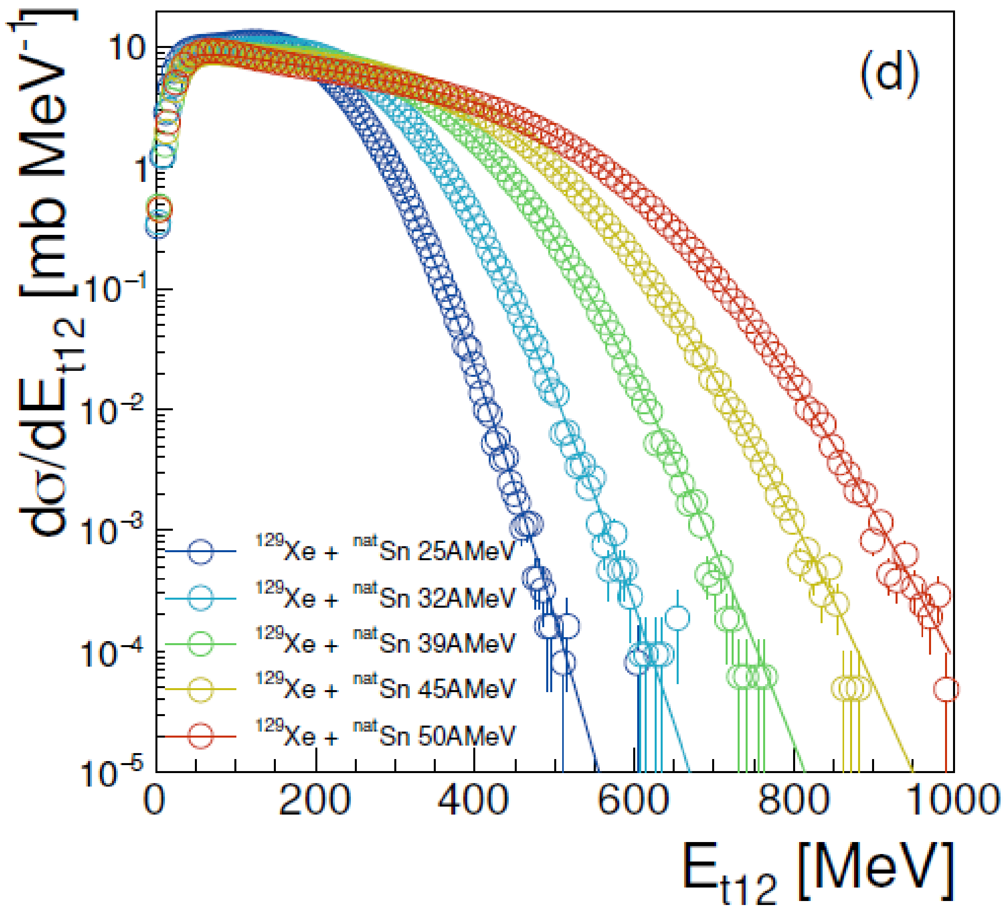


Open symbols denote the multiplicity of negatively charged particles
The vertical line in both panels indicates the position of the knee.
The method was also applied to data from LHC, CMS, and ATLAS experiments.

Rogly, R., Giacalone, G. & Ollitrault, J. Y. *Phys. Rev. C* 98, 1–9 (2018).

Application to experimental data: INDRA





Results of fits to the multiplicity of charged particles and transvers energy Et12 for the 129Xe+natSn data.
The Bayesian approach was applied to a very wide range of data for different collisions measured with INDRA between 25 MeV/nucleon and 100 MeV/nucleon

NA61/SHINE experimental setup

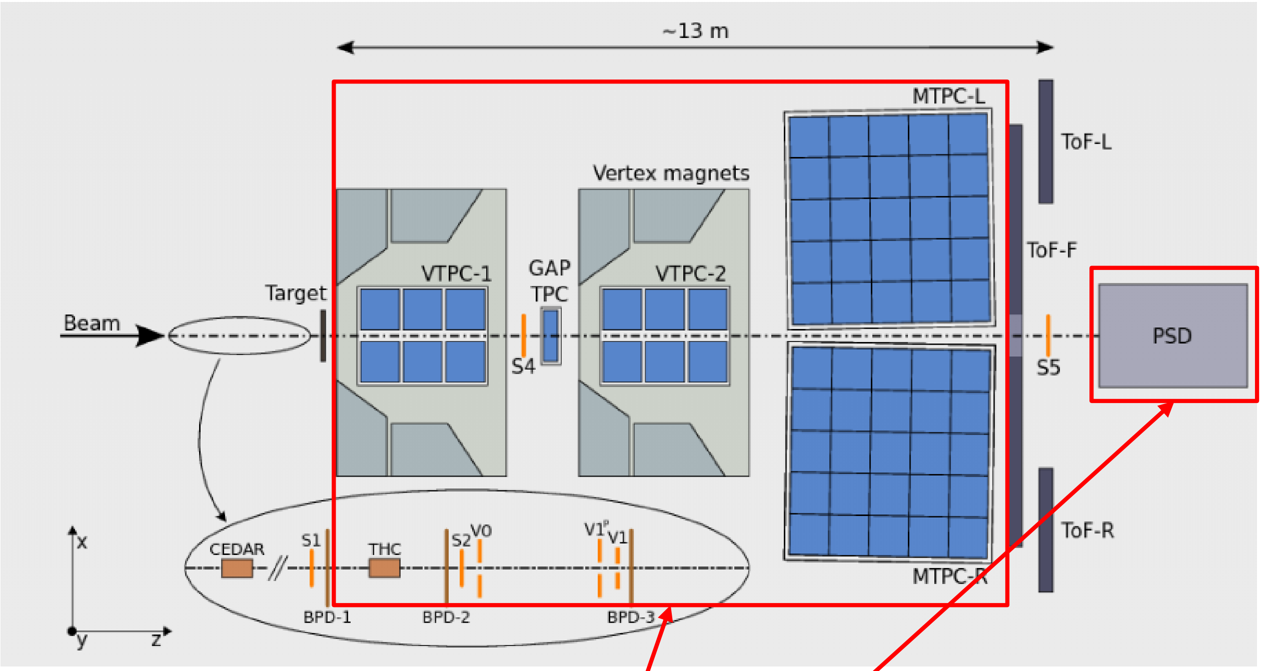
Data samples:

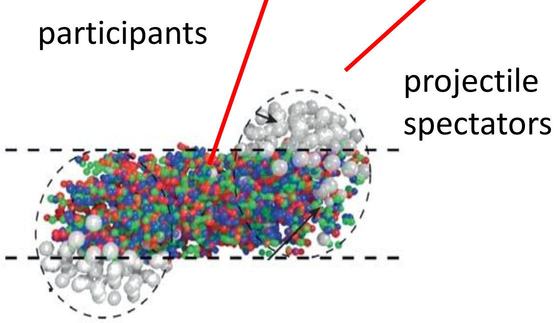
- Pb-Pb @ p_{beam} = 13A GeV/c
- data from 2016 physics run
- DCM-QGSM-SMM x Geant4

  M.Baznat et al. PPNL 17 (2020) 3, 303

Subsystems

- Multiplicity: TPCs
- Spectators energy: PSD





The Bayesian inversion method (Γ-fit): main assumptions

• Relation between energy E and impact parameter b is defined by the fluctuation kernel:

$$P(E \mid c_b) = \frac{1}{\Gamma(k(c_b))\theta^2} E^{k(c_b)-1} e^{-E/\theta}$$

$$c_b = \int_0^b P(b')db'$$ – centrality based on impact parameter

$$\theta = \frac{D(E)}{\langle E \rangle}, \quad k = \frac{\langle E \rangle}{\theta}$$

$\langle E \rangle, \quad D(E)$ – average value and variance of energy

$$\langle E \rangle = \mu_1 \langle E'(c_b) \rangle + \lambda_1, \quad D(E) = \mu_2 D(E'(c_b))$$

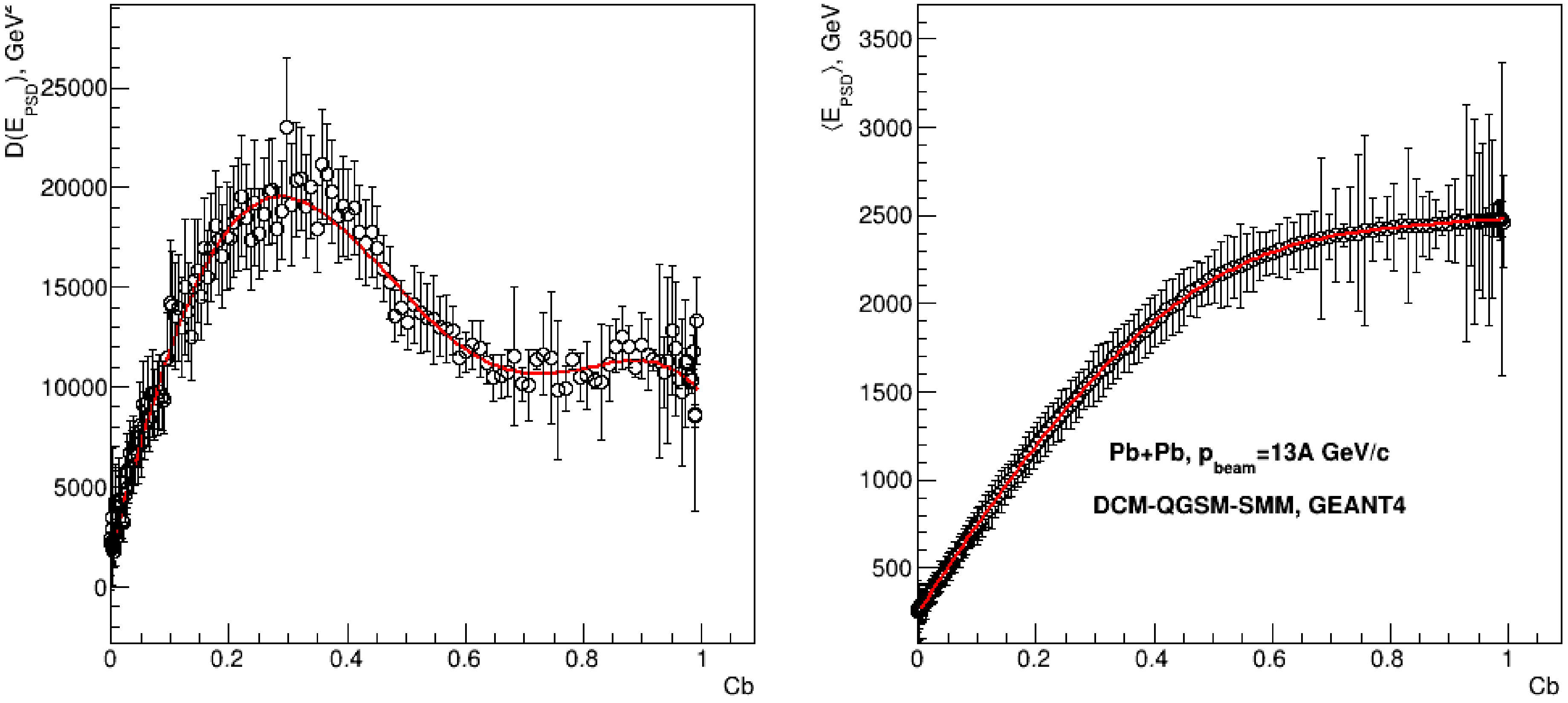
$\langle E'(c_b) \rangle, \quad D(E'(c_b))$ – average value and variance of energy from the rec. model data

Three fit parameters μ_1, μ_2, λ_1

$\langle E'(c_b) \rangle, \quad D(E'(c_b))$ - can be approximated by polynomials

$$\langle E'(c_b) \rangle = \sum_{j=1}^{8} a_j c_b^j, \quad D(E'(c_b)) = \sum_{j=1}^{6} b_j c_b^j$$

Dependence of the average value and variance of energy on centrality



The average value and dispersion of energy from the DCM-QGSM-SMM model are well described by polynomials

Reconstruction of *b*

- Normalized energy distribution P(E)

$$P(E) = \int_0^1 P(E \mid c_b) dc_b$$

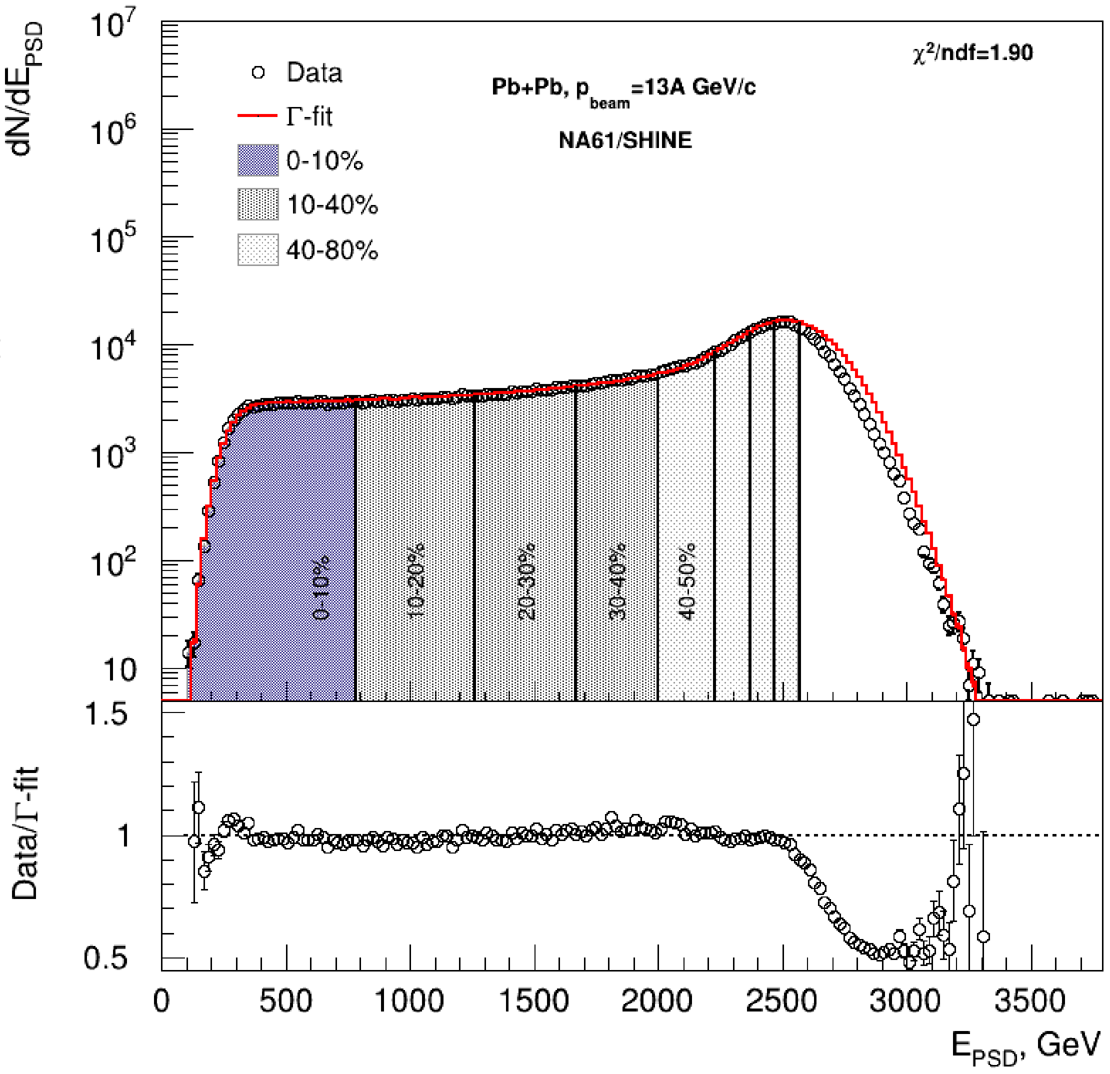
- Find probability of *b* for fixed range of E using Bayes' theorem:

$$P(b \mid E_1 < E < E_2) = P(b) \int_{E_1}^{E_2} P(b \mid E) dE \Big/ \int_{E_1}^{E_2} P(E) dE$$

- **The Bayesian inversion method consists of 2 steps:**

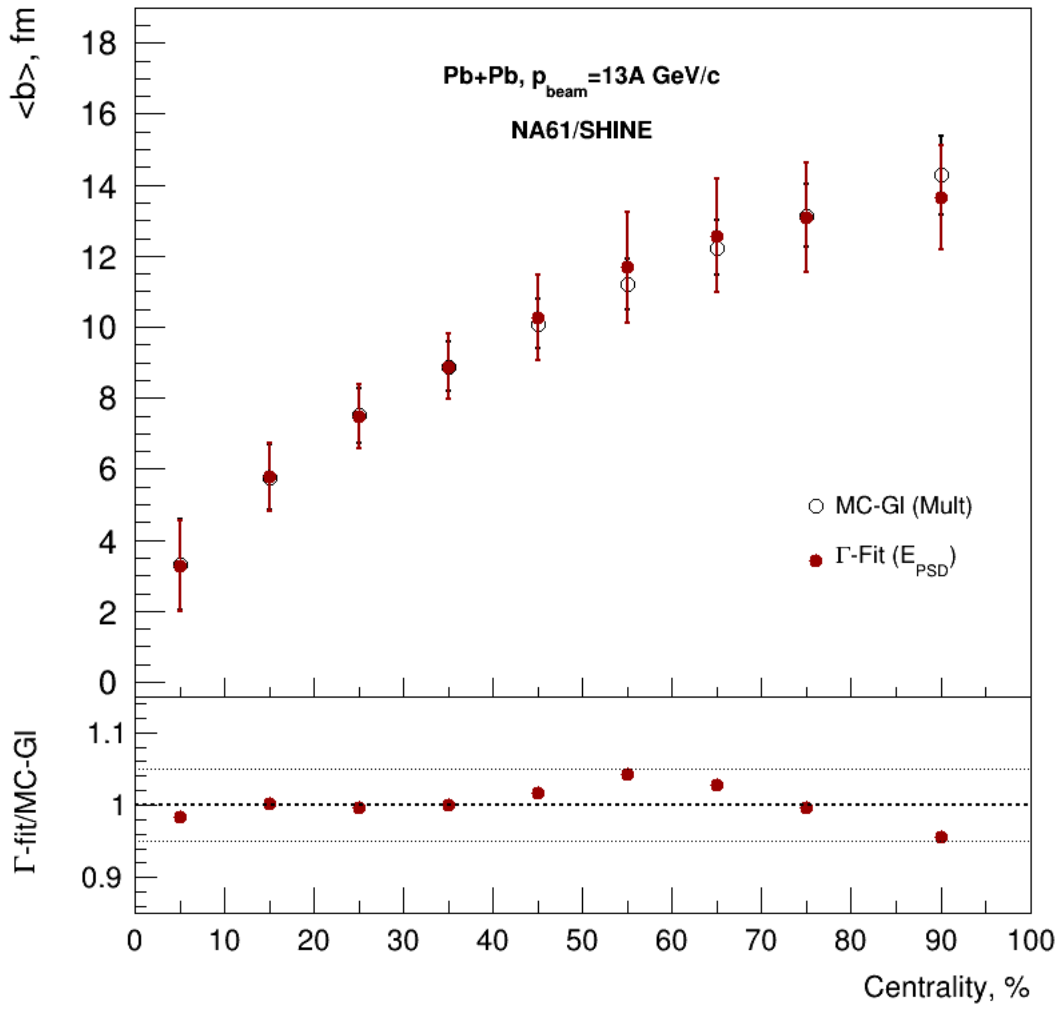
–Fit normalized energy distribution with P(E)

–Construct P(*b*|E) using Bayes' theorem with parameters from the fit

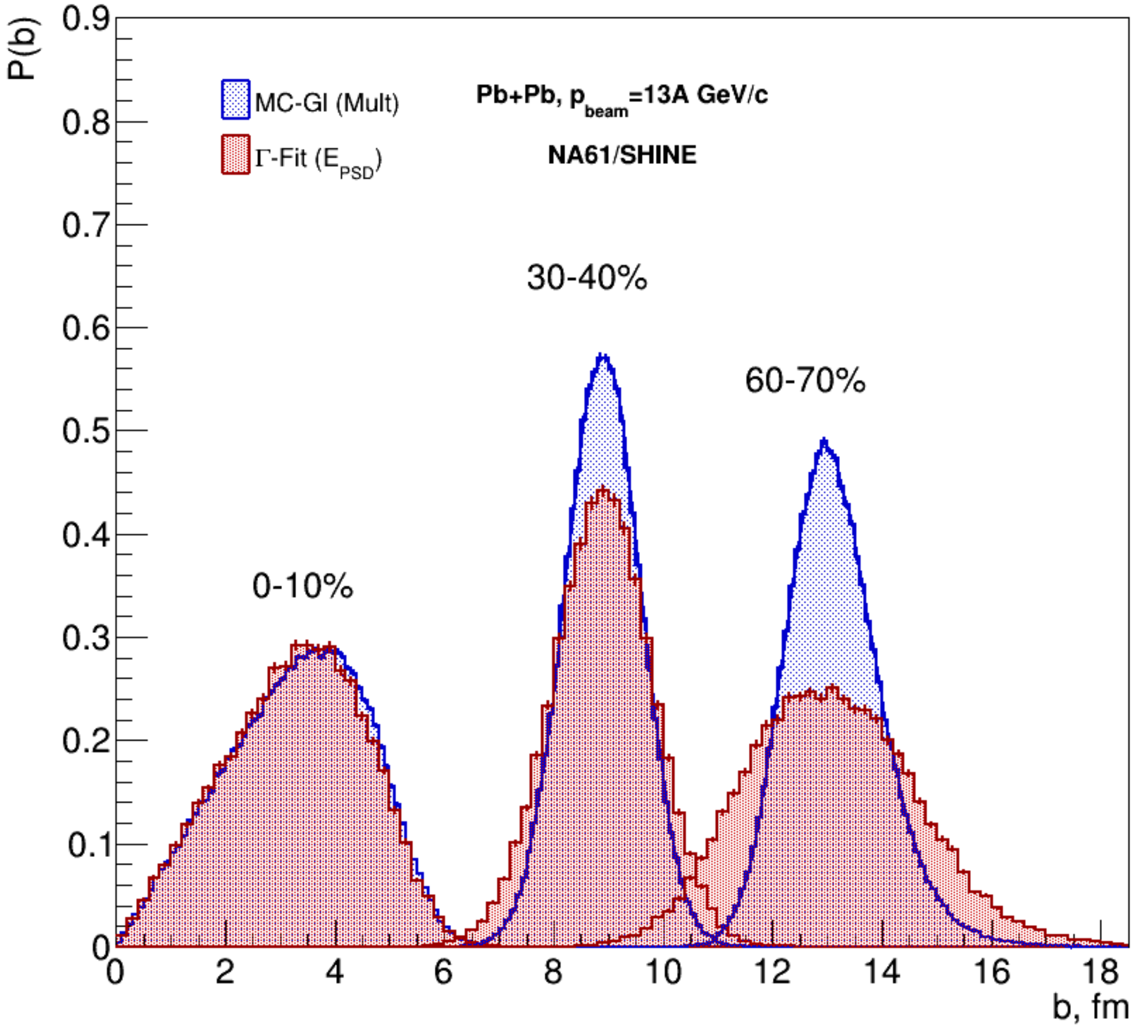


Good agreement between fit and data in wide energy range

Comparison with MC-Glauber fit



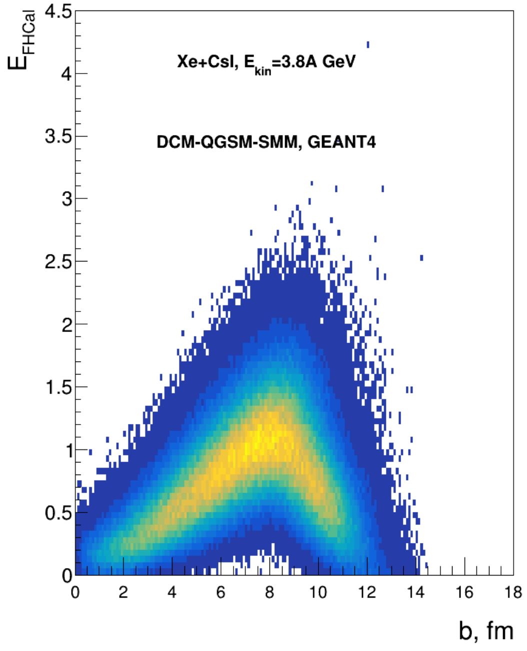
Good agreement between fit and data.

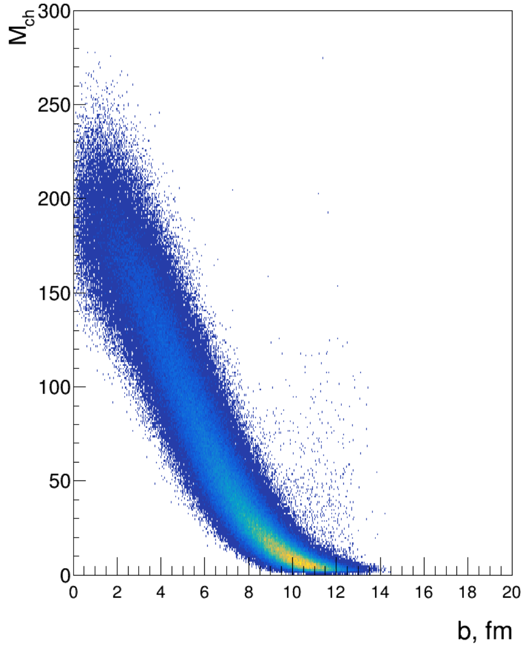


There is agreement within 5%.

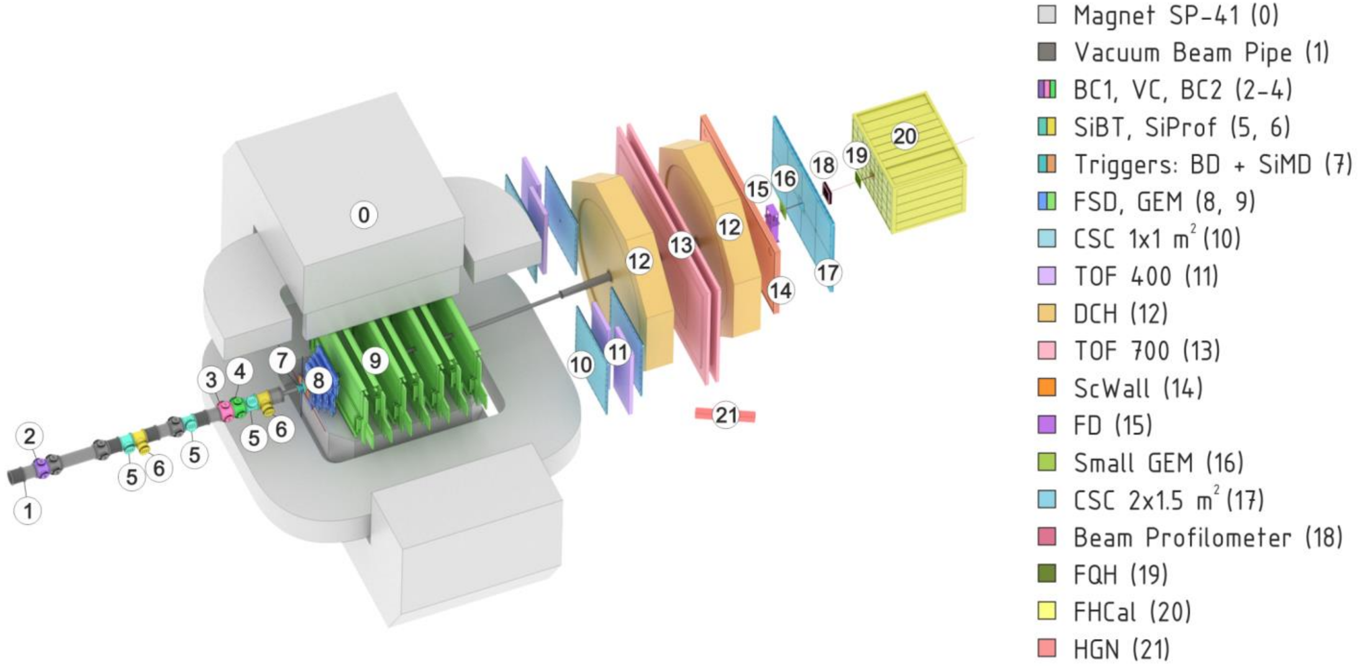
Centrality determination in BM@N

DCM-QGSM-SMM → GEANT4 → Reconstruction → Centrality determination





Dependence of energy in FHCal and track multiplicity on the impact parameter



BM@N setup overview

The Bayesian inversion method (Γ-fit): DCM-QSM-SMM based

- The fluctuation kernel Gamma distr.:

$$P(M \mid c_b) = \frac{1}{\Gamma(k(c_b))\theta^2} M^{k(c_b)-1} e^{-M/\theta}$$

$$c_b = \int_0^b P(b')db'$$ – centrality based on impact parameter

$$\theta = \frac{D(M)}{\langle M \rangle}, \quad k = \frac{\langle M \rangle}{\theta}$$

$\langle M \rangle, D(M)$ – average and variance of Multiplicty

$$P(M) = \int_0^1 P(M \mid c_b)dc_b$$

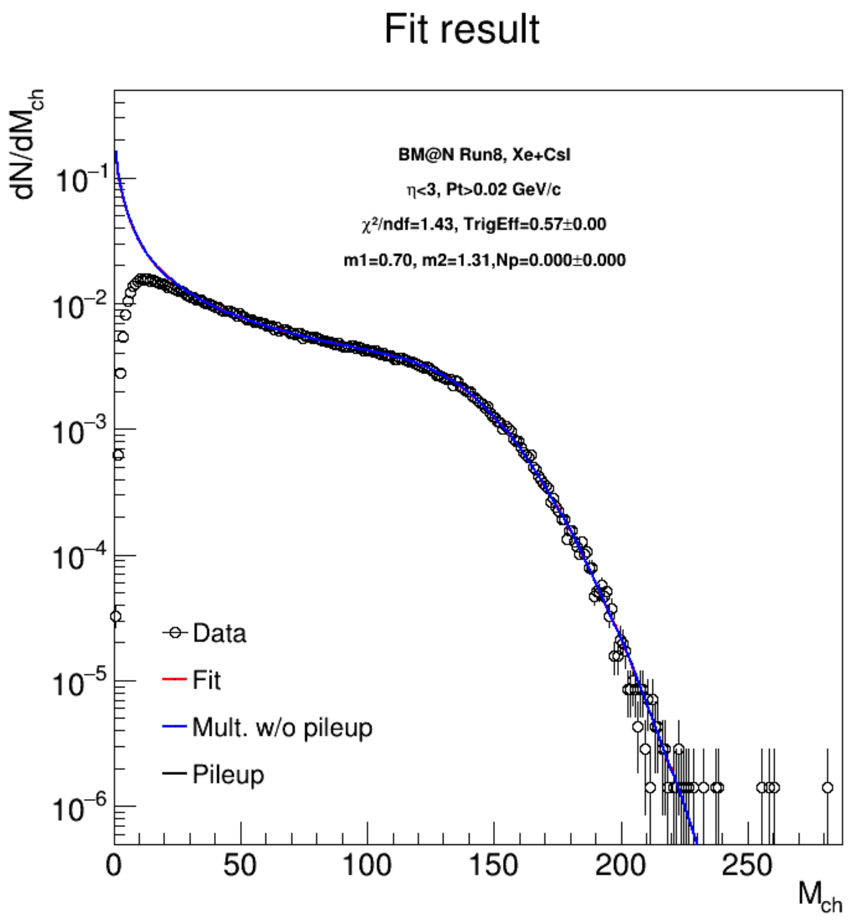
$$\langle M \rangle = m_1 \cdot \langle M' \rangle$$

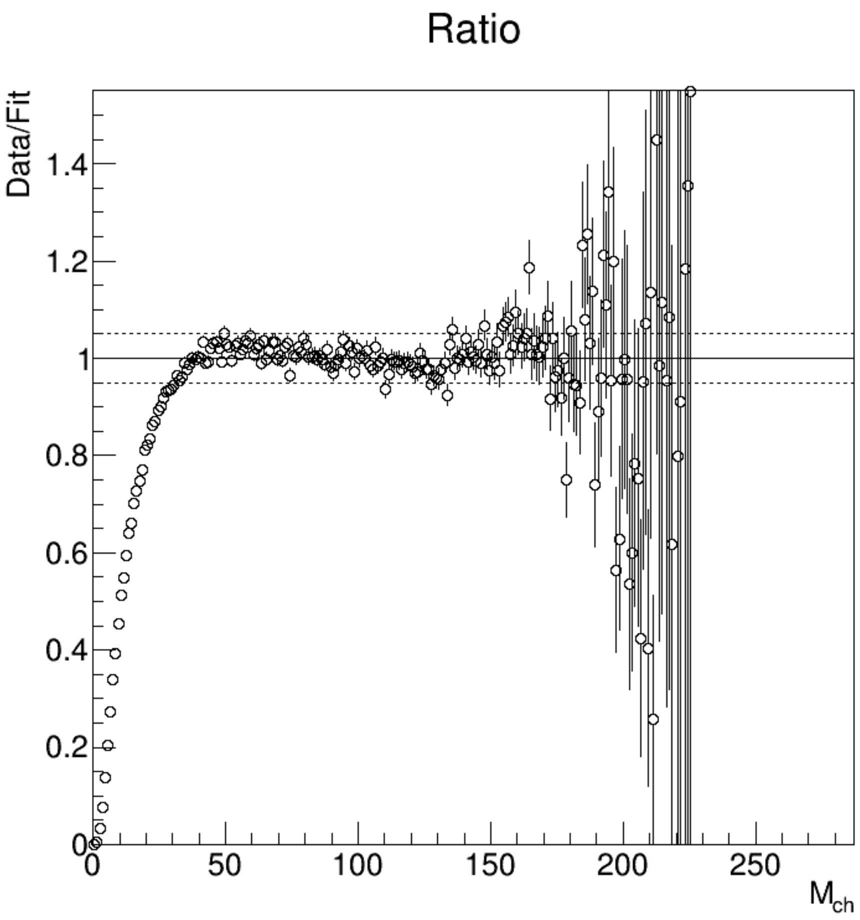
$$D(M) = m_1^2 \cdot D(M') + m_1 \cdot m_2 \langle M' \rangle$$

$\langle M'(c_b) \rangle$, $D(M'(c_b))$ – average value and var. of energy/mult. from the rec. model data

- can be approximated by polynomials and exponential polynomial

Multiplicty-based Γ-fit: BM@N Xe+CsI





Vertex Cuts: CCT2, N_{vtxTr}>1,|$V_{x,y}$ –(0.3,0.14)|<1 cm, |V_z -0.07|< 0.2 cm

Good agreement with data

Track selection: Nhit>4, η<3, Pt>0.05 GeV/c

The Bayesian inversion method (Γ-fit): 2D fit

- The fluctuation kernel for energy and multiplicity at fixed impact parameter can be describe by 2D Gamma distr.:

$$P(E,M\,|\,c_b)=G_{2D}(E,M)$$

$$c_b=\int_0^b P(b')db'$$ – centrality based on impact parameter

$\langle E\rangle, D(E)$ – average value and variance of energy

$\langle M\rangle, D(M)$ – average value and variance of mult.

$R(E,M)$ – Pirson correlation coefficient

$$R(E,M)=\frac{\varepsilon_1^2 m_1^2}{\varepsilon_2 m_2}R(E',M')$$

$\varepsilon_1,\varepsilon_2,m_1,m_2$ - fit parameters

$\langle E'(c_b)\rangle$, $D(E'(c_b))$ – average value and var. of energy/mult. from the rec. model data

$$\langle E\rangle=\varepsilon_1\langle E'(c_b)\rangle,\quad D(E)=\varepsilon_2 D(E'(c_b))$$

$$\langle M\rangle=m_1\langle M'(c_b)\rangle,\quad D(M)=m_2 D(M'(c_b))$$

$\langle E'(c_b)\rangle,\quad D(E'(c_b))$ - can be approximated by polynomials

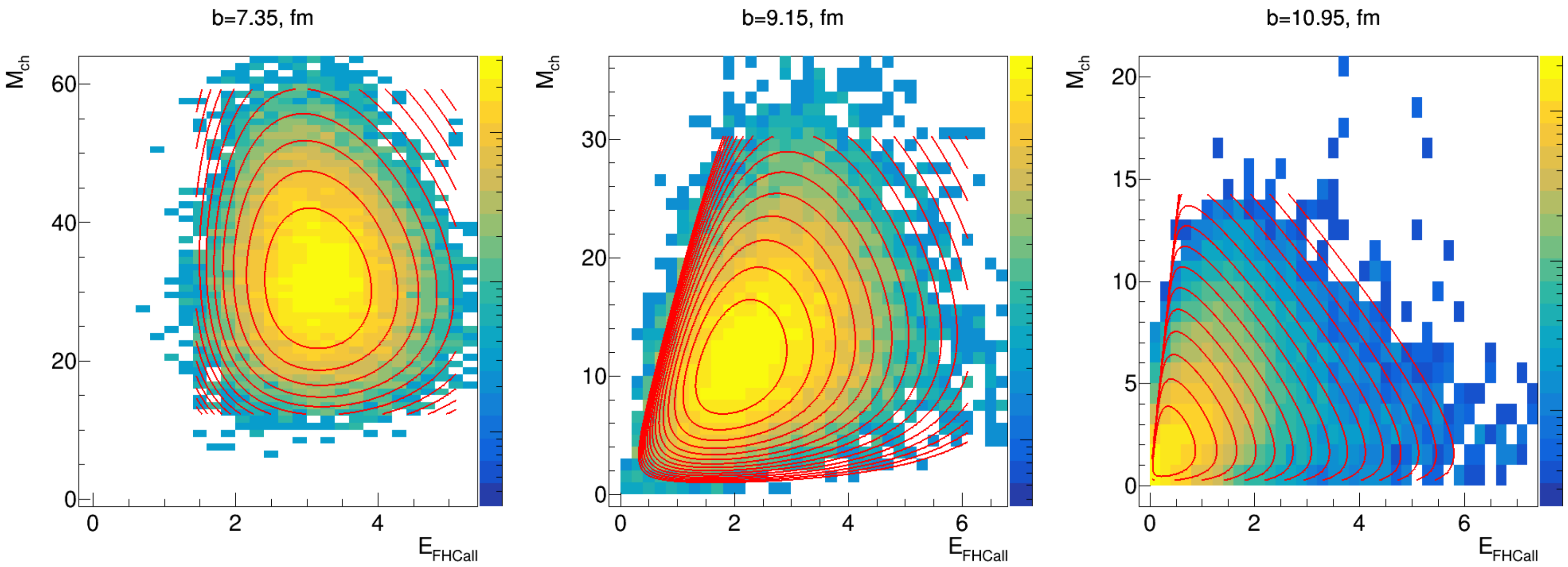
The fluctuation of energy and multiplicity at fixed impact parameter

It is possible to find such a rotation angle of the system that $\operatorname{cov}(x, y) = 0$

$$\langle x \rangle = \cos(\alpha)\langle E \rangle + \sin(\alpha)\langle M \rangle$$
$$\langle y \rangle = -\sin(\alpha)\langle E \rangle + \cos(\alpha)\langle M \rangle$$

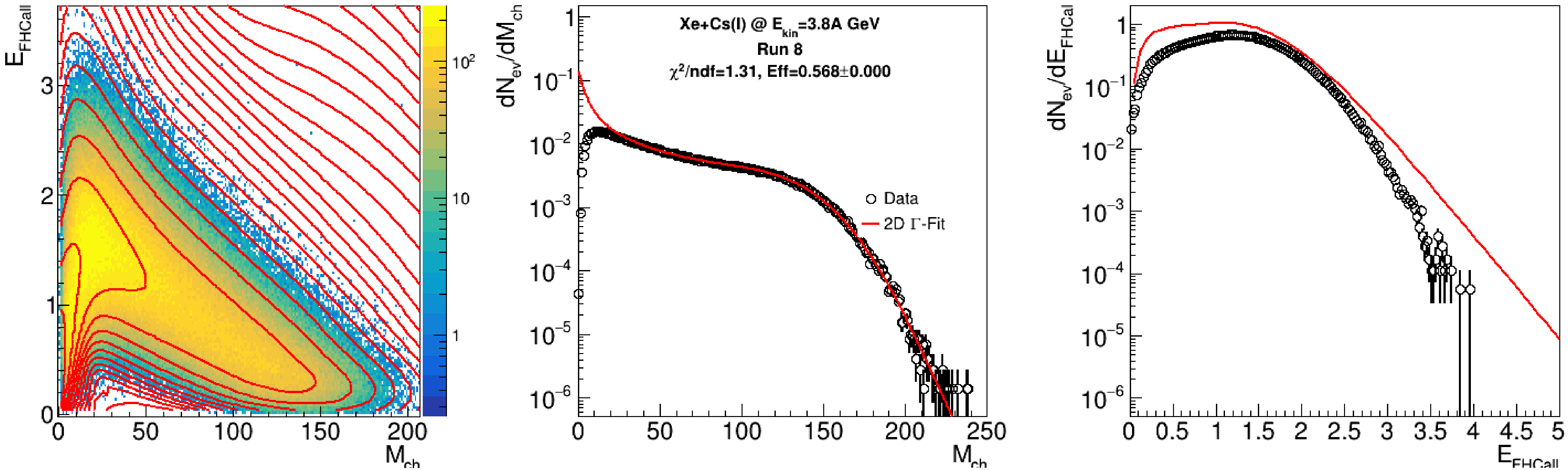
$$\alpha = \arctan\left(\frac{2\sqrt{D(E)D(M)}R(E,M)}{D(E) - D(M)}\right)$$

$$G_{2D}(E_{FH}, M_{ch}) = G(x, \theta_x, k_x) \cdot G(y, \theta_y, k_y)$$



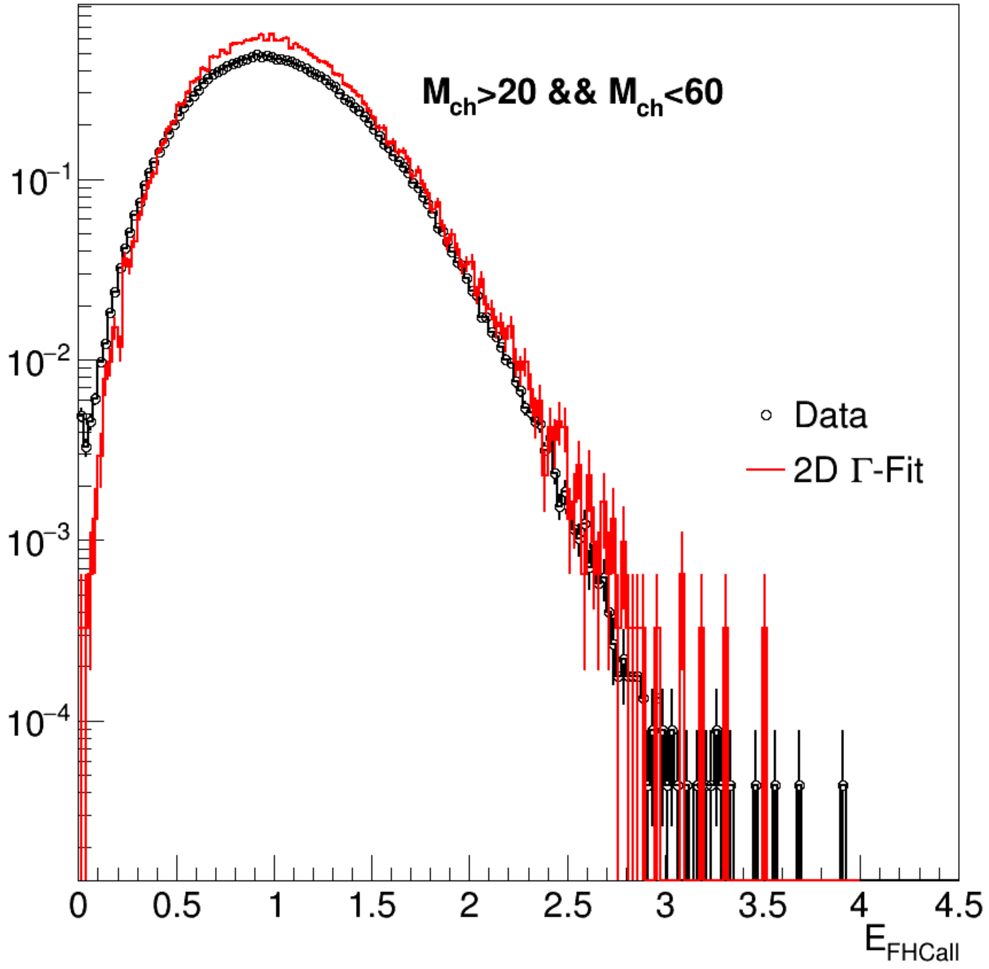
The distribution of energy and multiplicity at a fixed impact parameter is well described by the gamma distribution

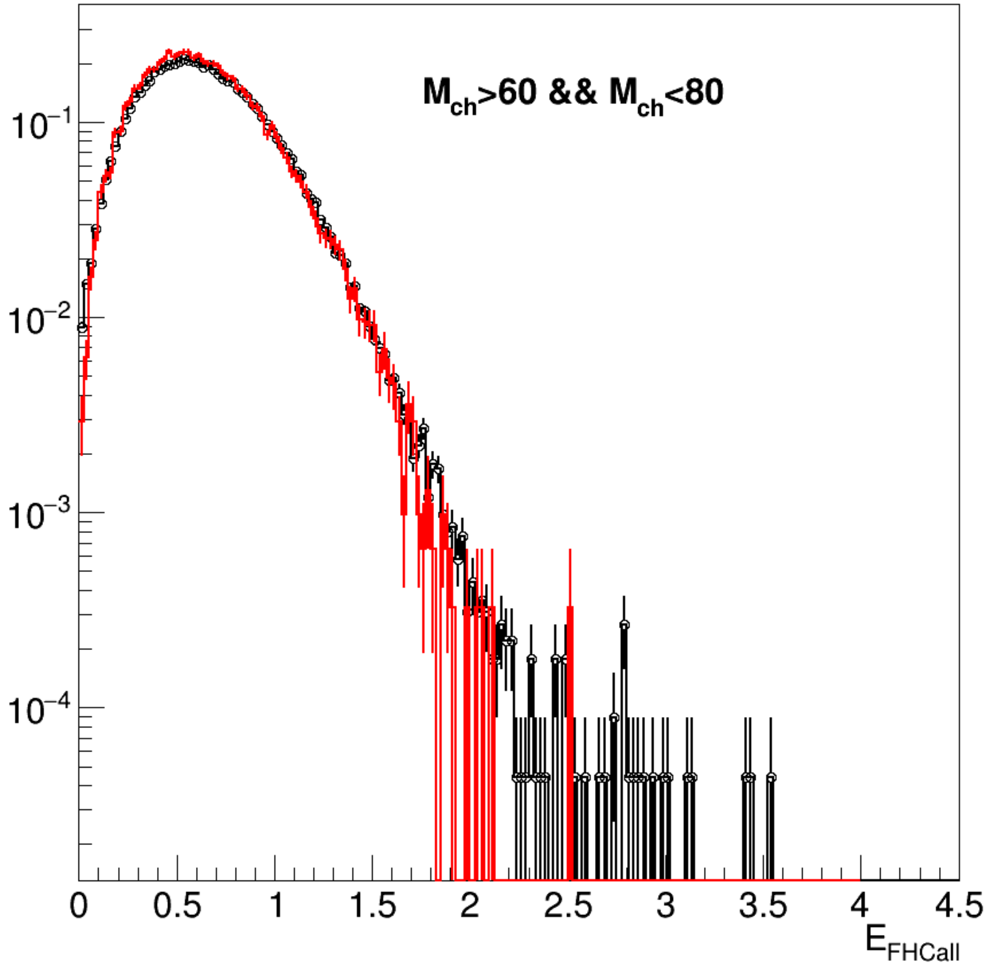
2D fit results



The fit function qualitatively reproduces the multiplicity-energy correlation from FHCal

Energy distributions from FHCal for different multiplicity cuts

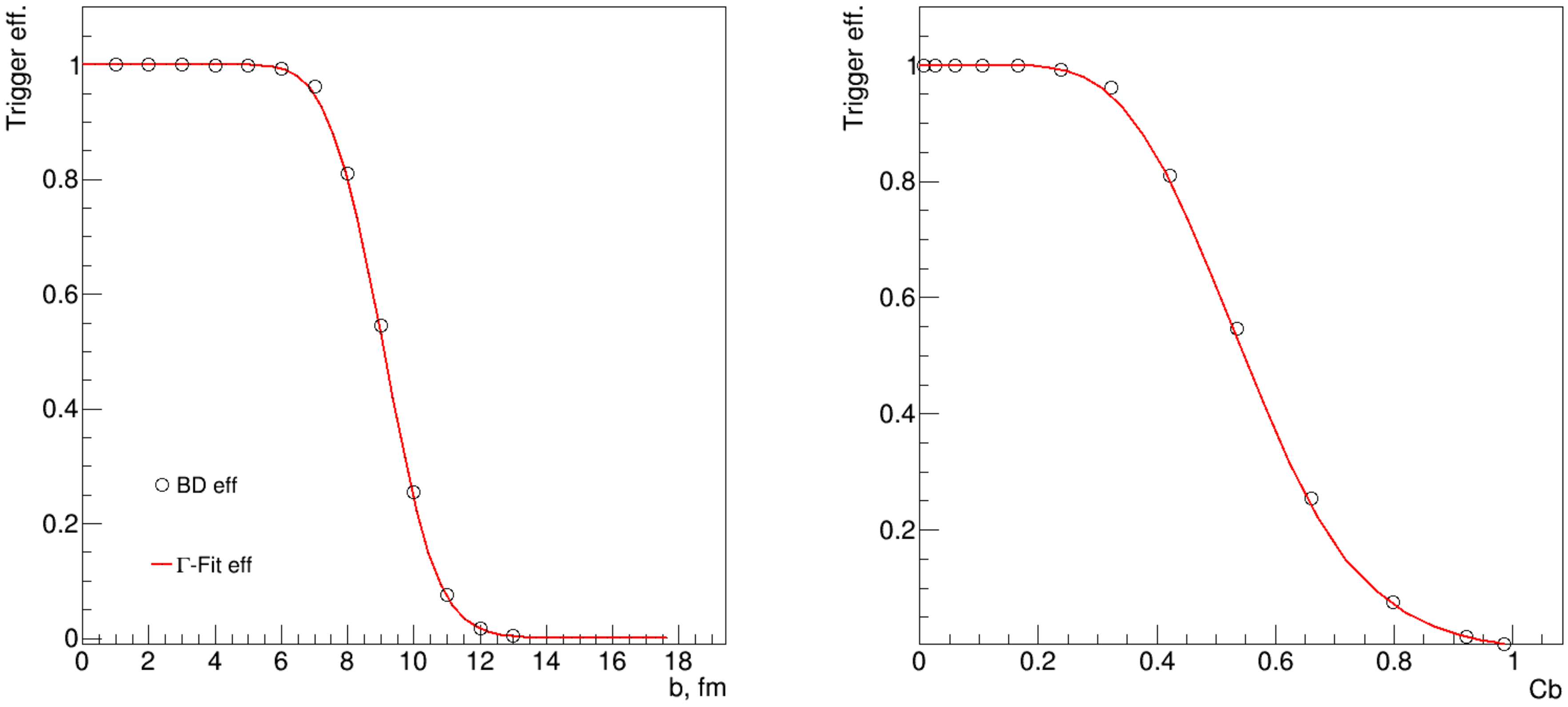




Good agreement between fit and data for the area below the anchorpoint

The total efficiency of event registration

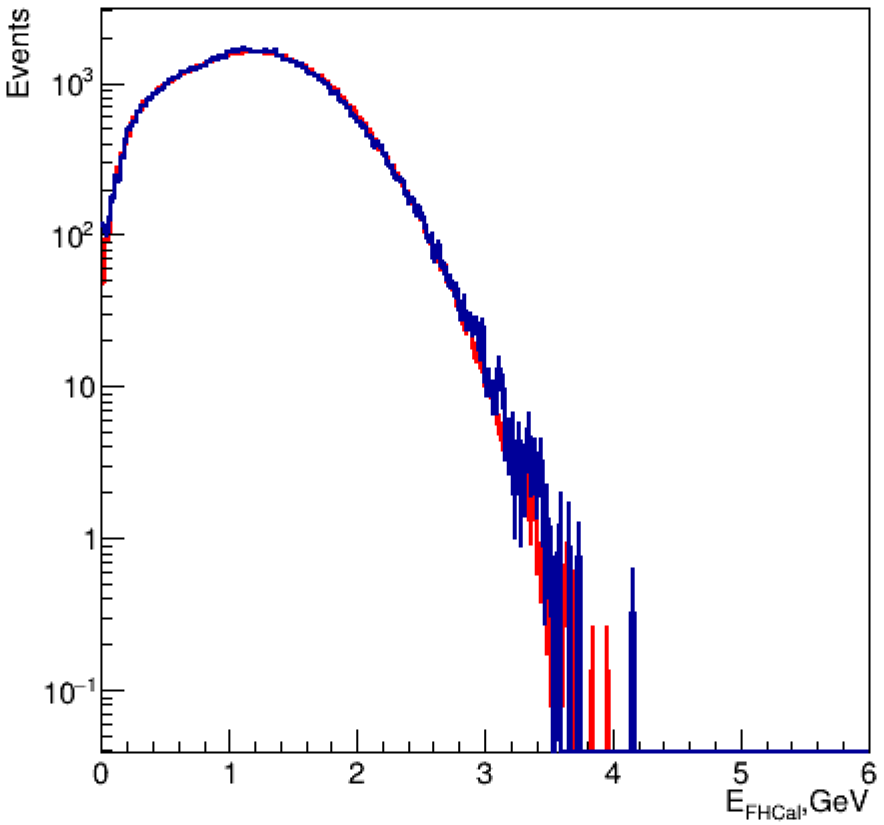
$$P_{eff}(b) = \int_0^{M_{\max}} P_{eff}(\mathrm{M}) P(\mathrm{M}\,|\,\mathrm{b})\, dM$$

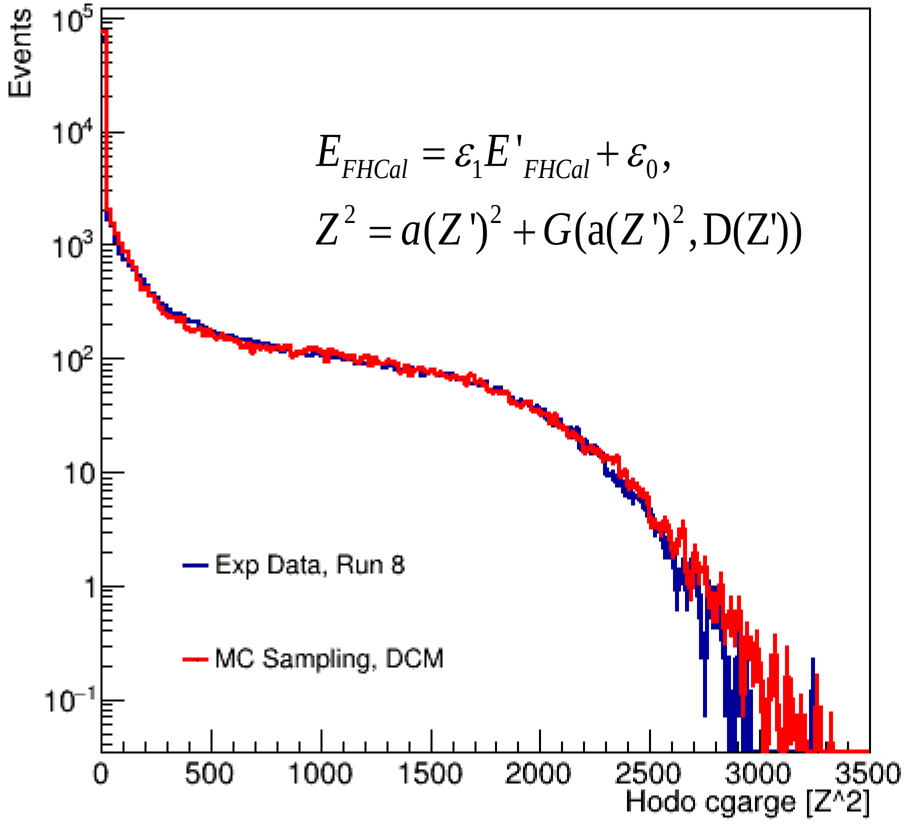


The trigger efficiency obtained from the Bayesian approach is consistent with the results, obtained on the basis of simulations

https://indico.jinr.ru/event/4762/contributions/28478/attachments/20298/35273/lashmanov_report_BM@N-meeting_Oct-2024.pdf

The results of the fit signals from the calorimeter and hodoscope

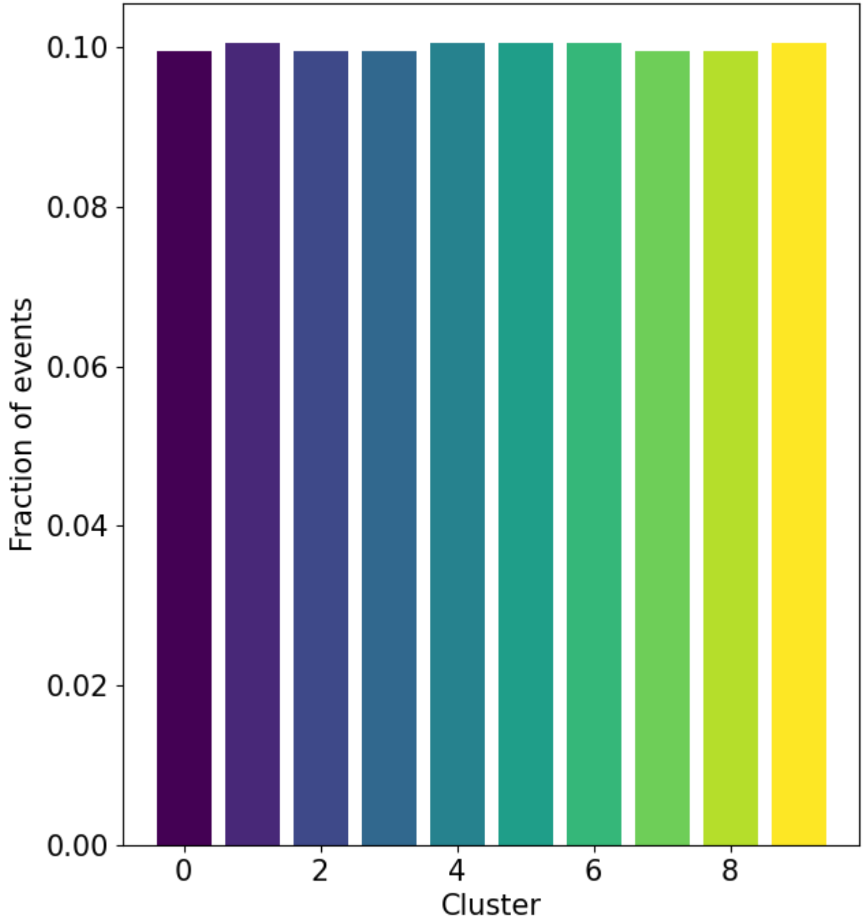


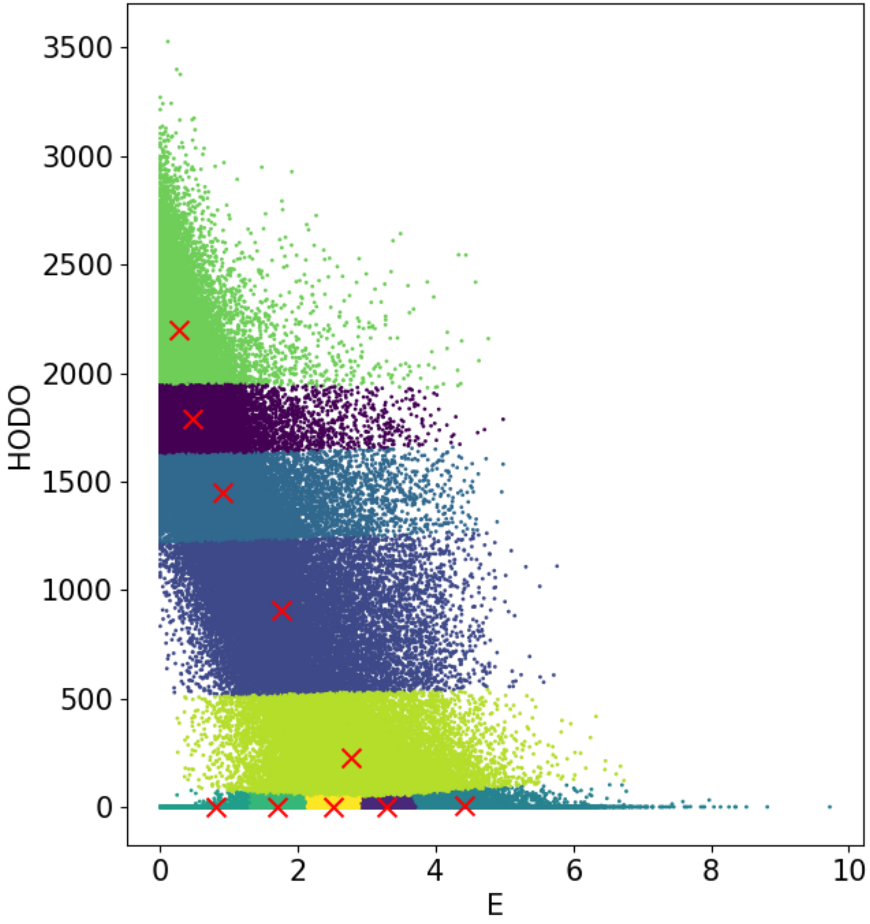


Good agreement of fit results for the calorimeter

The fit procedure for the hodoscope is in the process of developing

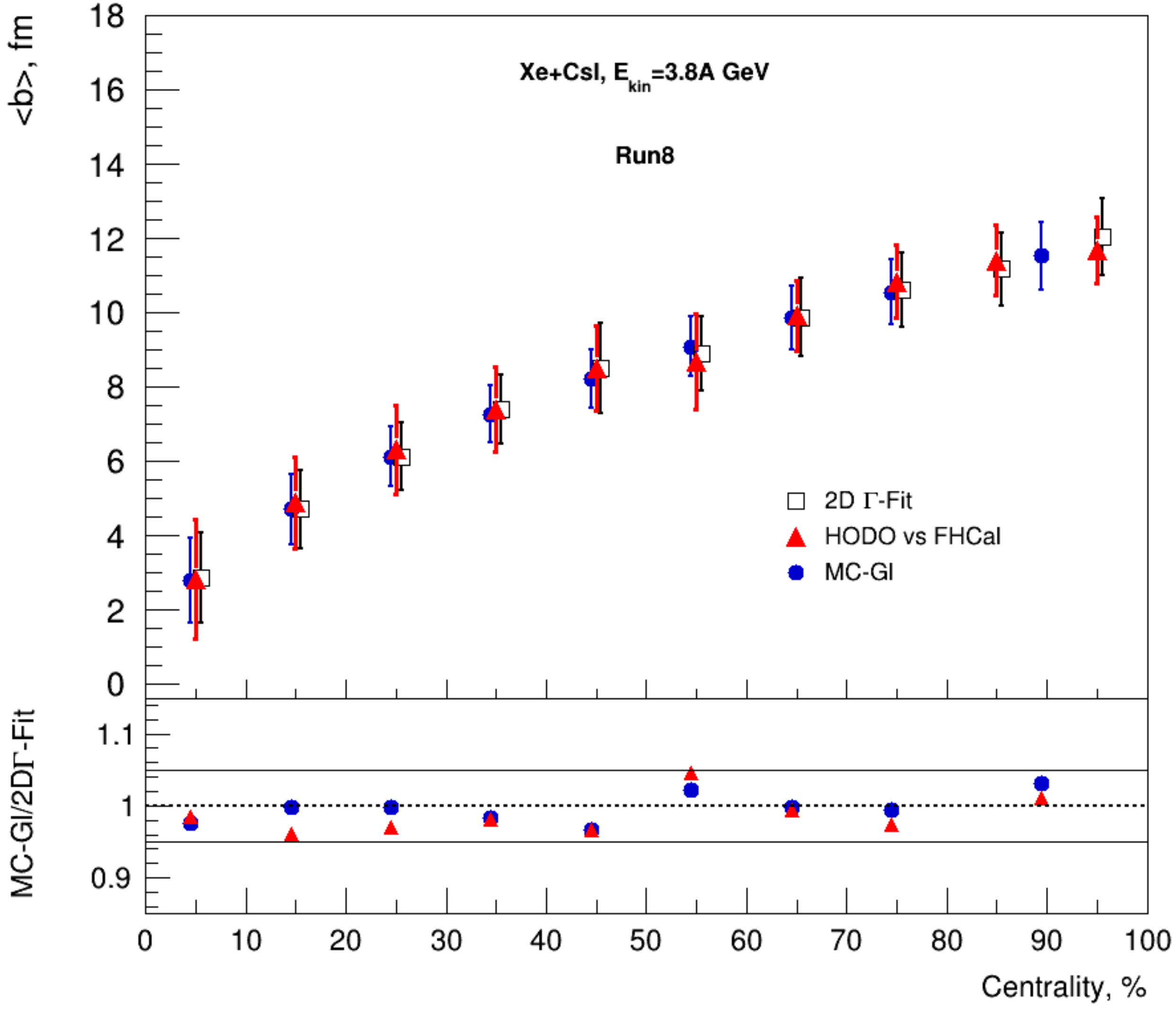
Centrality determination using an forward calorimeter and hodoscope

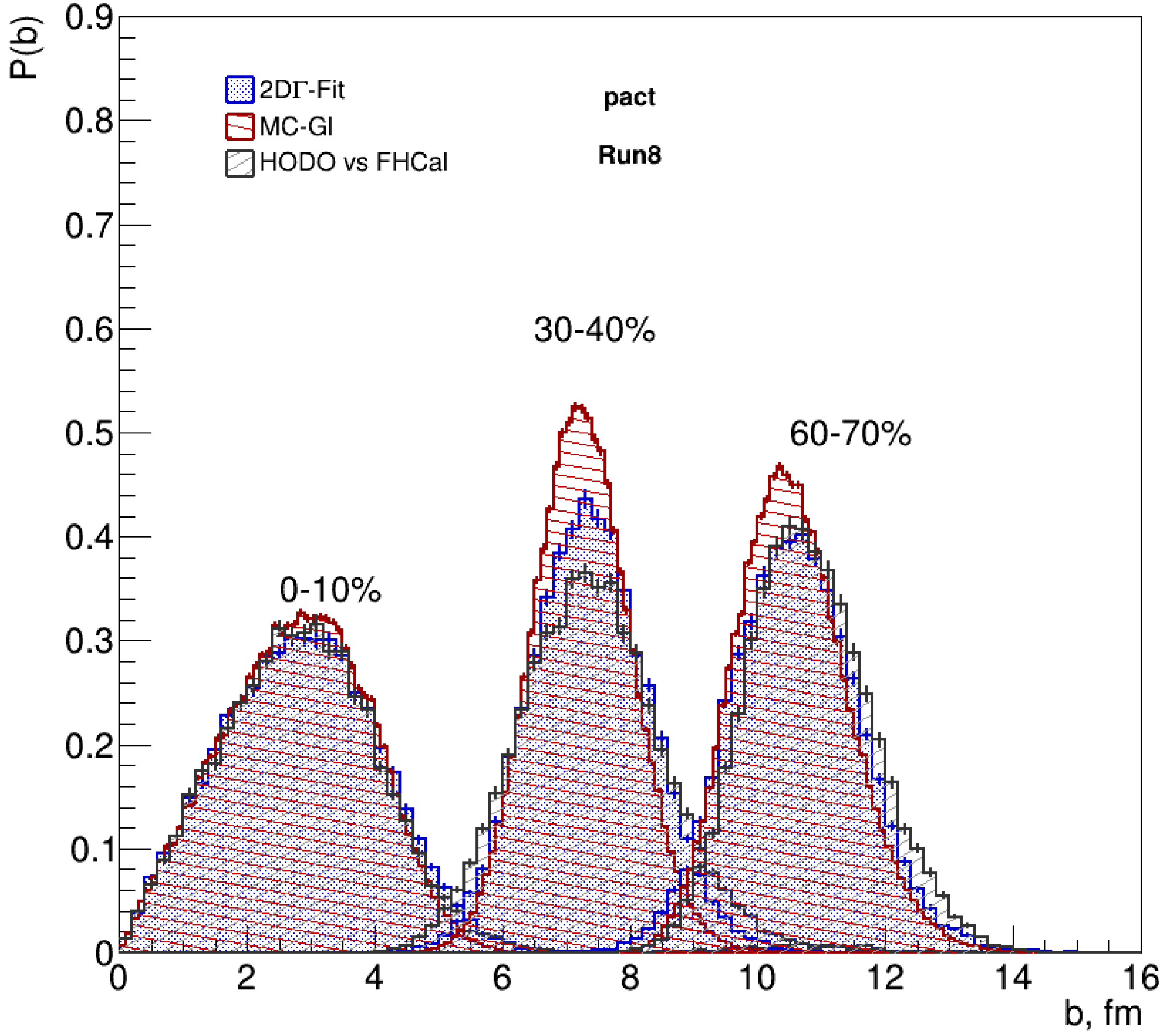




The K-means method allows to divide a two-dimensional distribution into centrality classes.
In order to correctly apply the class boundaries, it is necessary to match the simulation results with the experiment

Comparison with MC-Glauber fit

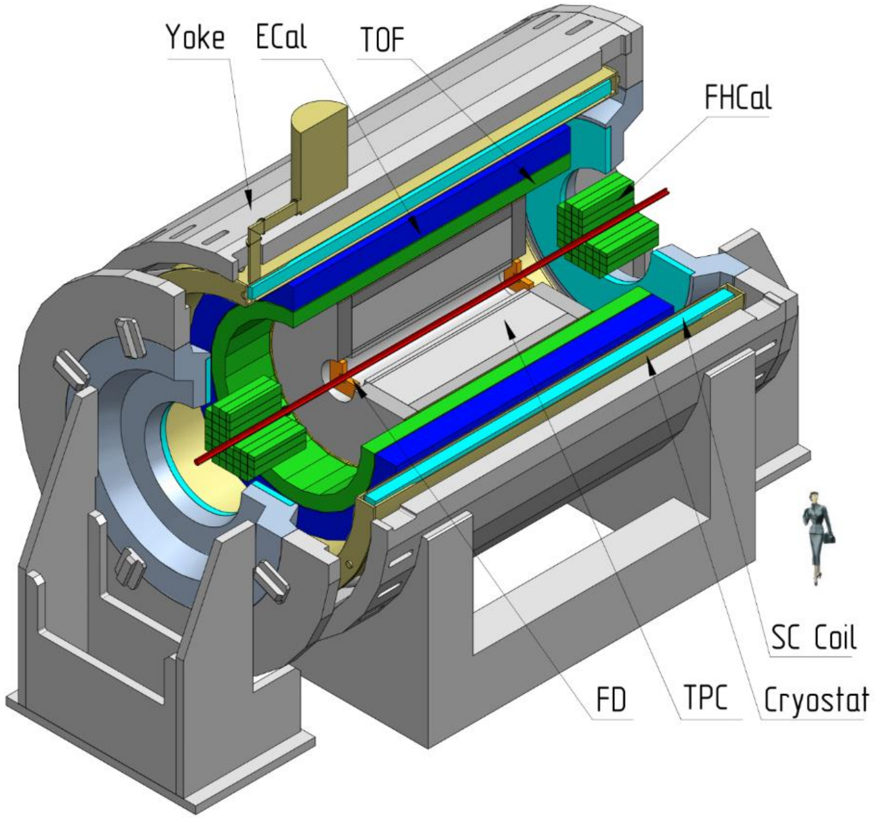




There is agreement within 5%.

MPD Experiment at NICA

UrQMD → GEANT4 → Reconstruction → Centrality determination

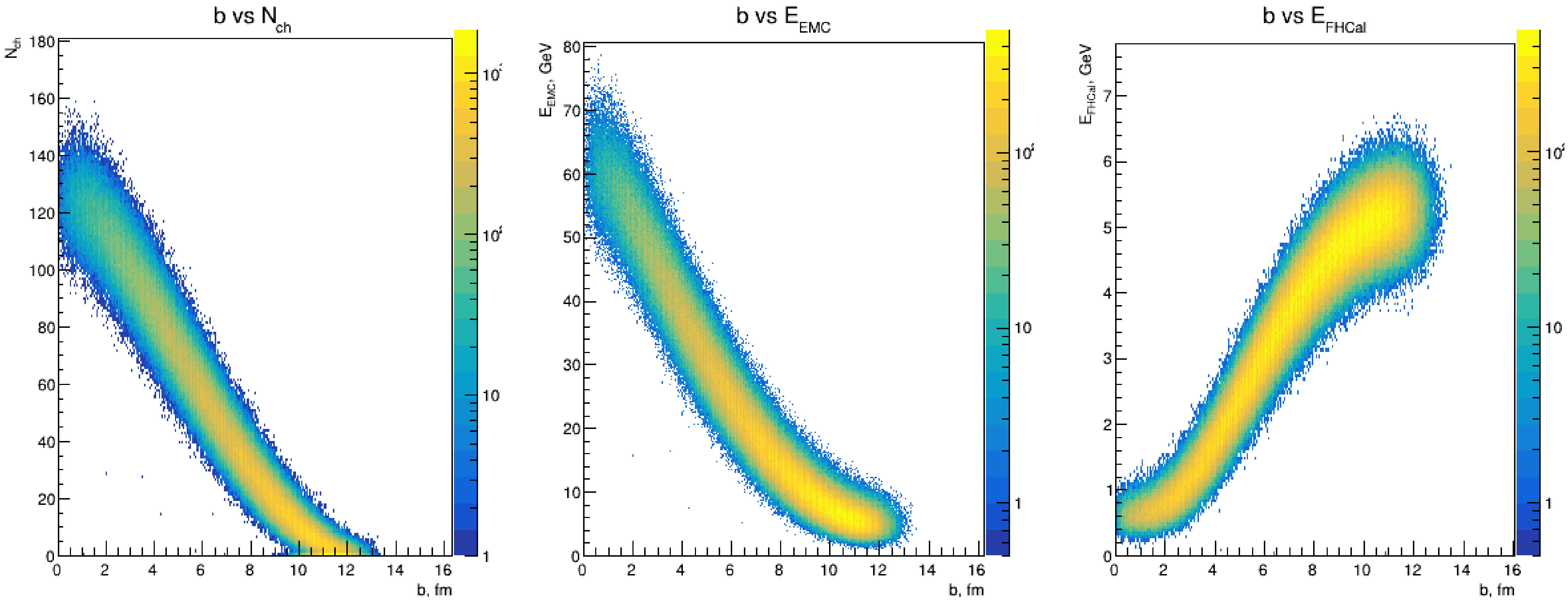


Multi-Purpose Detector (MPD) setup overview

- **Centrality determination:** Multiplicity of produced charged particles in TPC, energy from EMC and FHCal
- **Event plane determination:** TPC
- **Track selection:**
  - Primary tracks
  - $N_{TPC\ hits} \geq 10$
  - $0.2 < p_T < 2.0$ GeV/c
  - $0.5<\eta< 2$

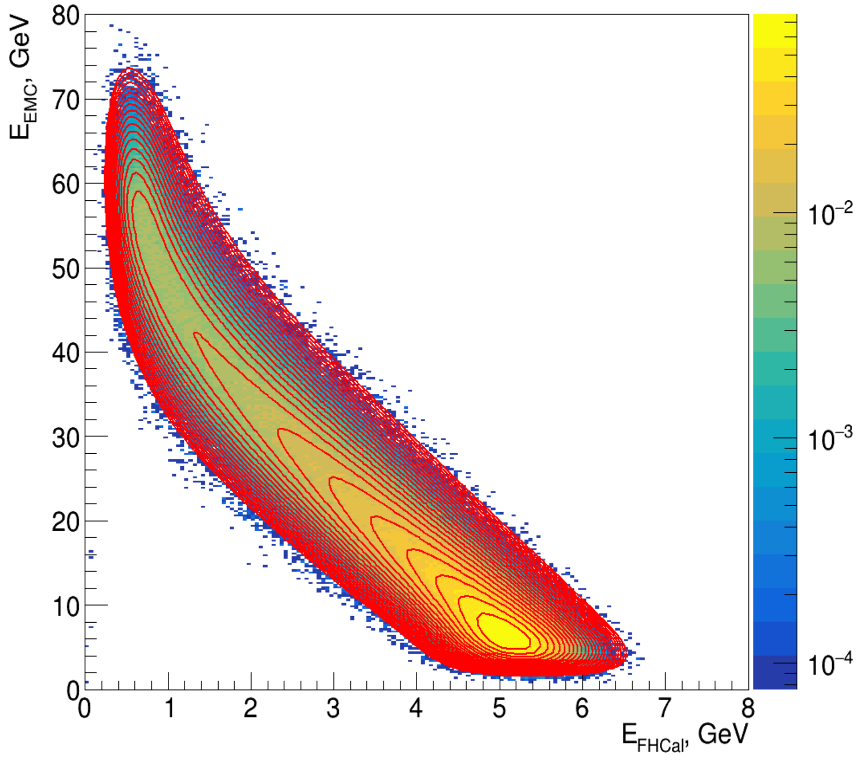
Centrality determination in MPD

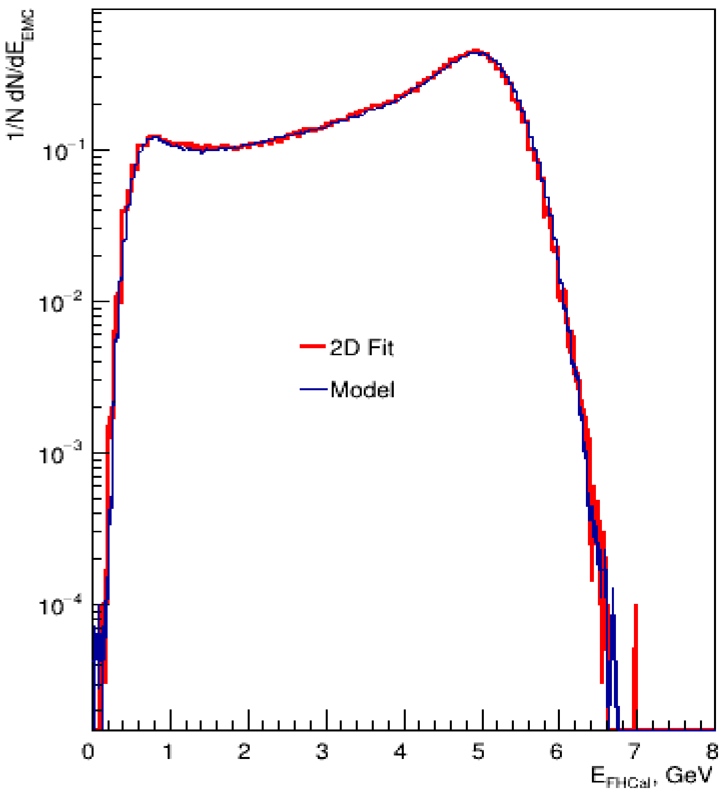
UrQMD → GEANT4 → Reconstruction → Centrality determination

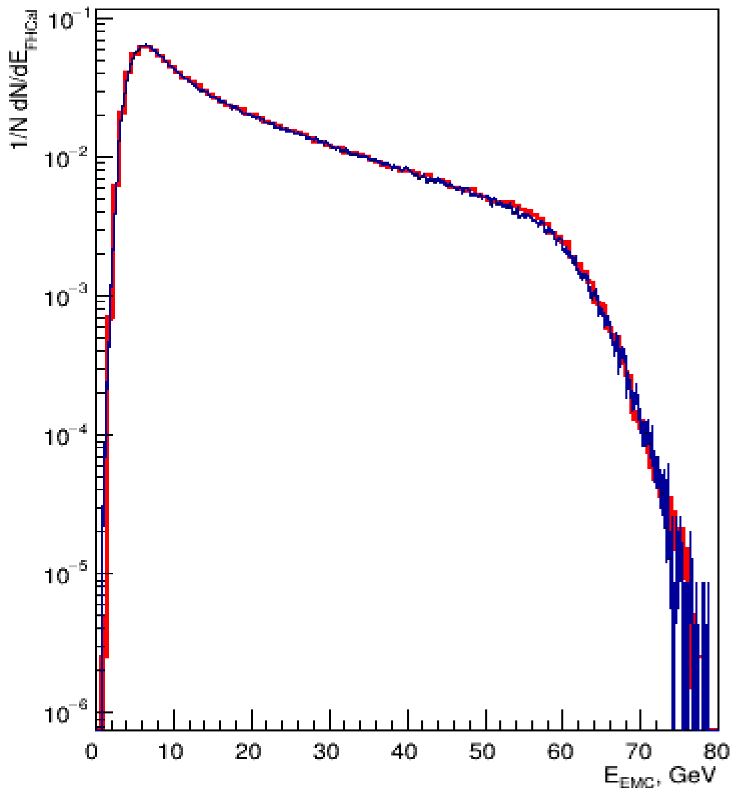


Relation between impact parameter and observables

2D Bayesian approach: results

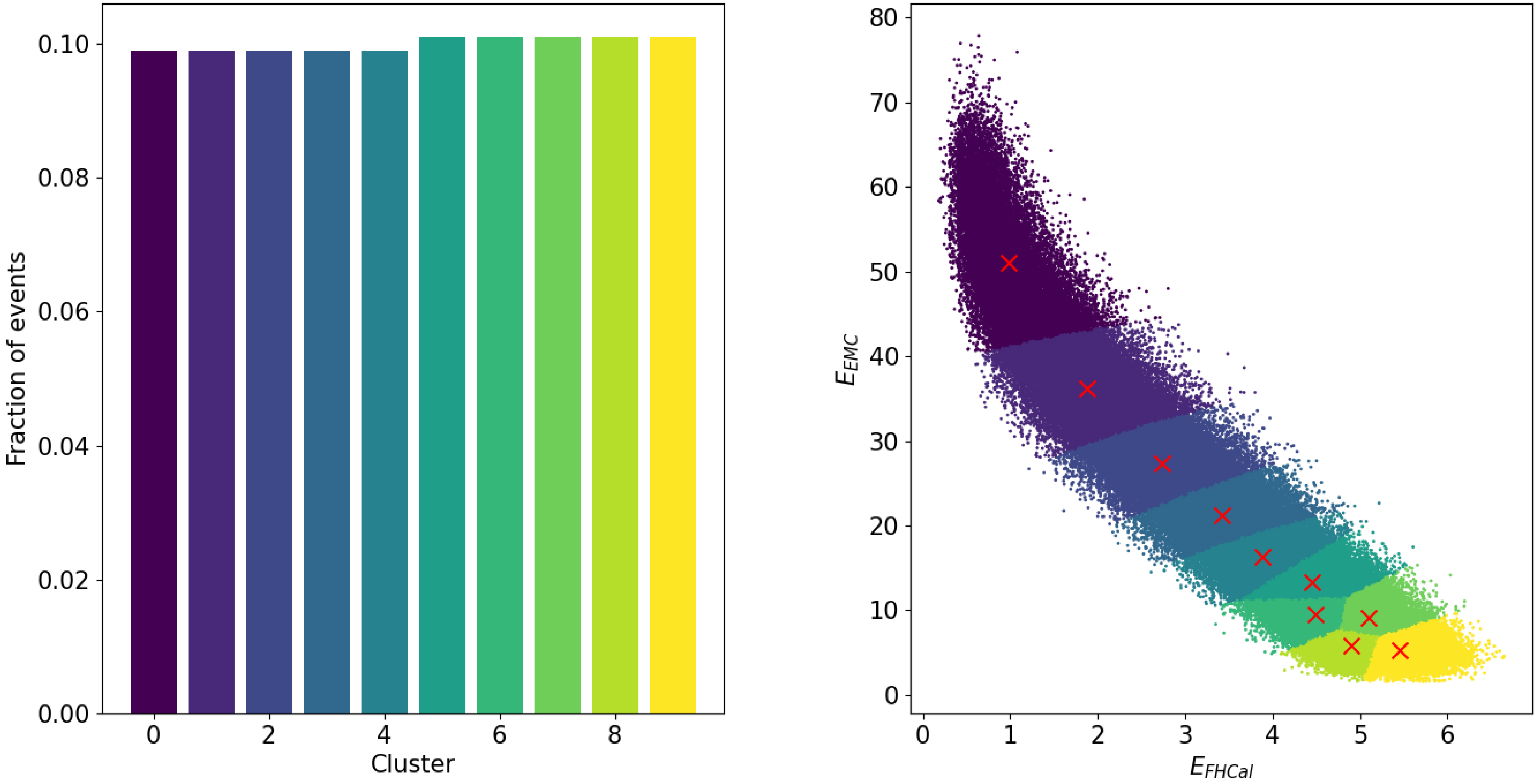






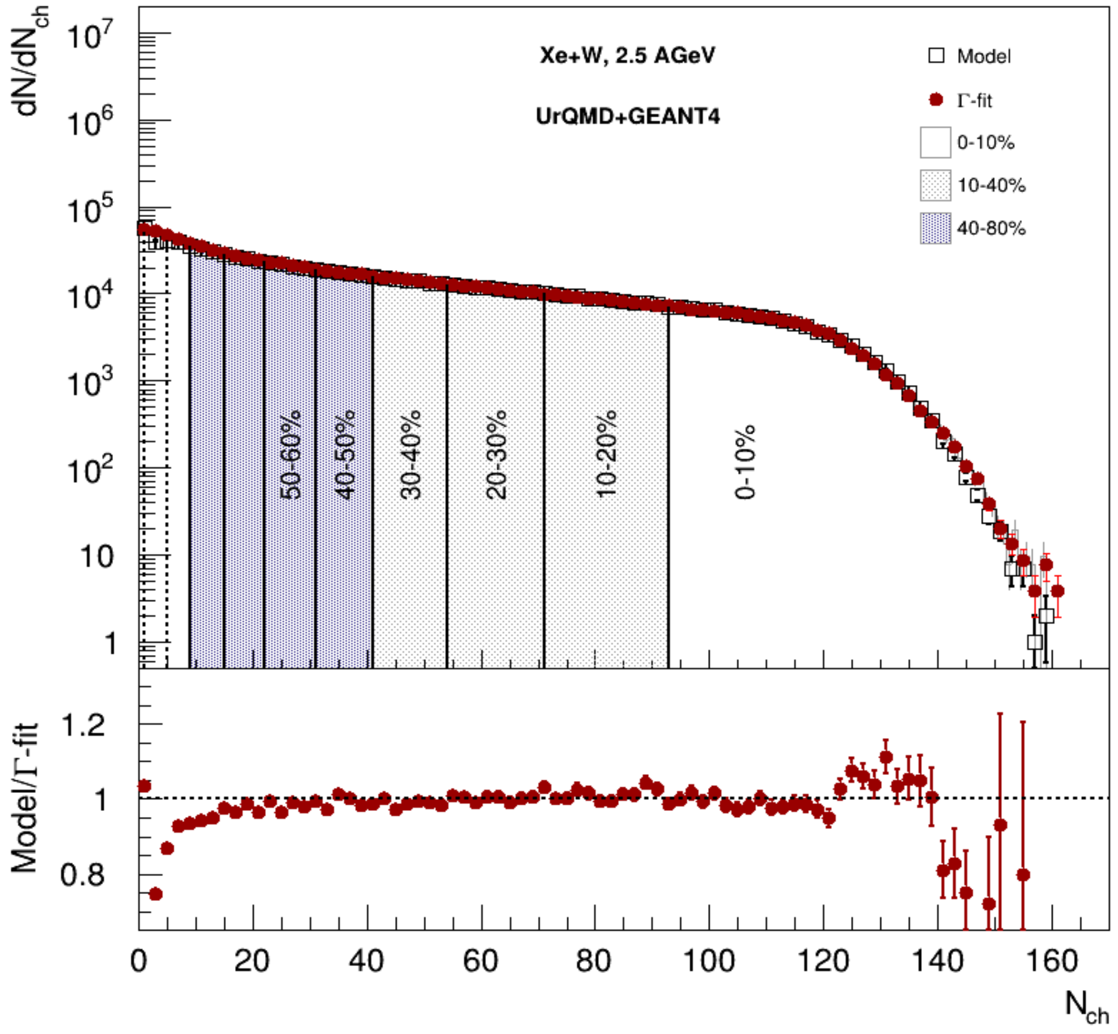
Good agreement between fit and data.

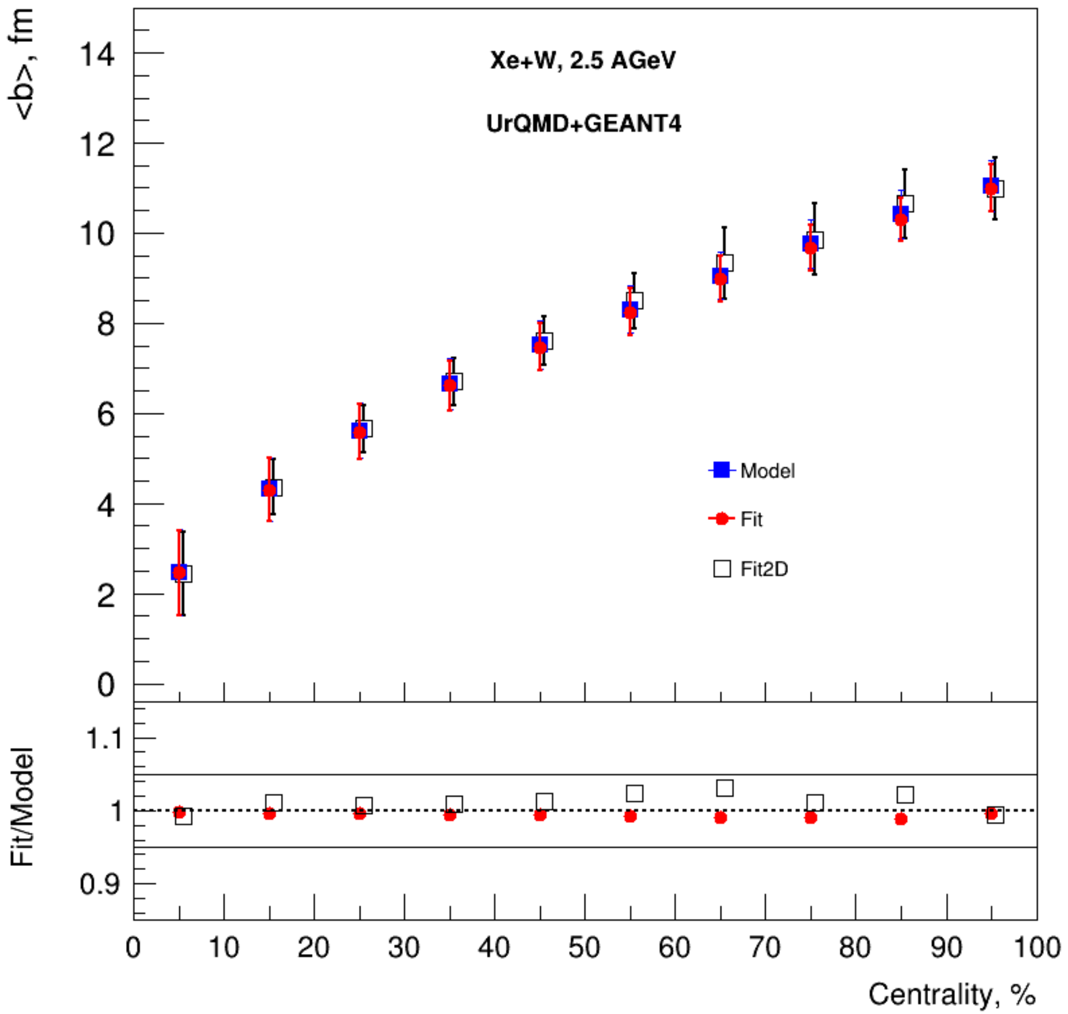
Centrality classes



Centrality classes were obtained using the k-means algorithm

Comparison centrality determination methods





Good agreement with fit

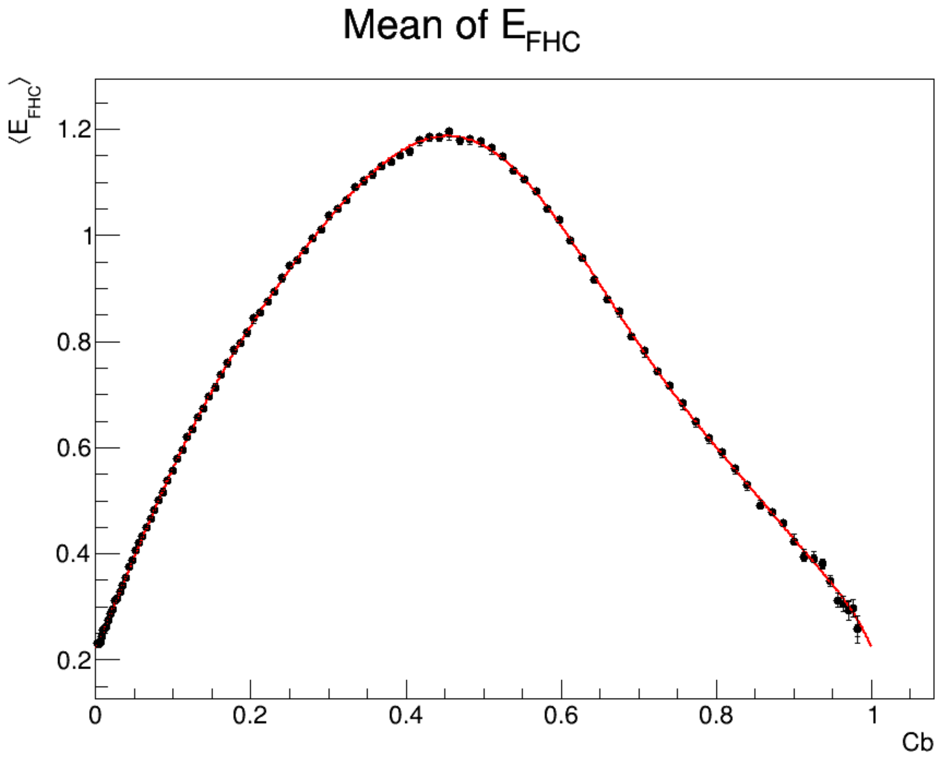
There is agreement within 5%.

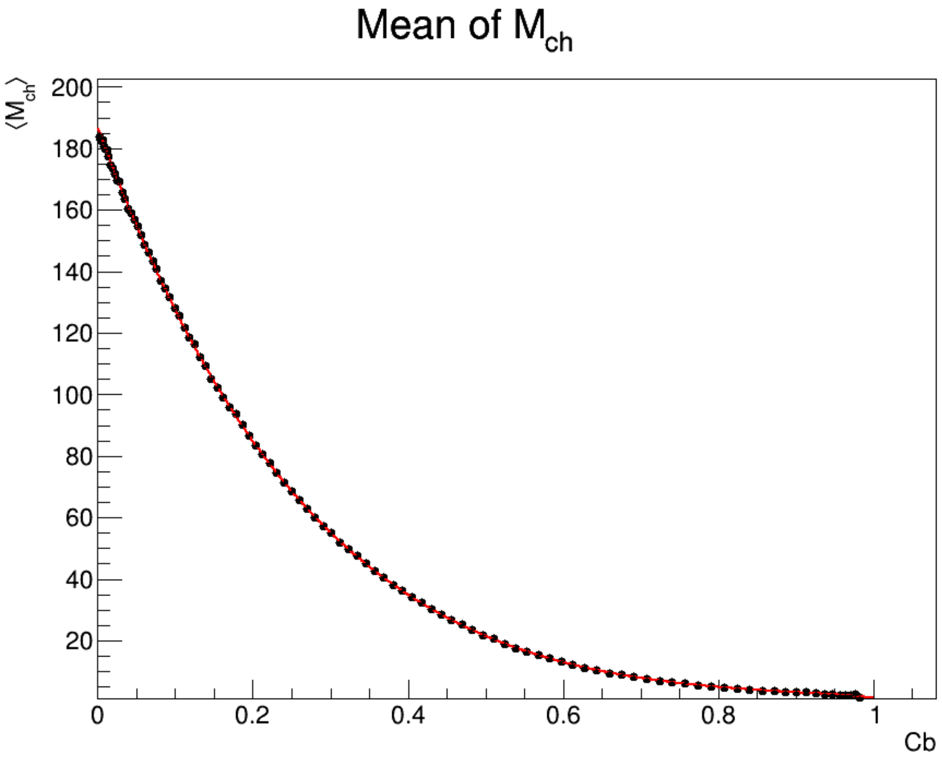
Summary and outlook

- A new approach for centrality determination with energy of spectators and multiplicity of charged particles was proposed.
- The proposed method was applied to the data from BM@N experiment:
  - Multiplicity-based and 2D approaches using Q^2_{Hodo} and E_{FHCal} describe experimental data reasonably well
- The Bayesian inversion method reproduce observables for fixed-target mode at MPD:
  - Multiplicity-based and 2D approaches using E_{EMC} and E_{FHCal} show consistent results with model data
- To do:
  - Systematic study of different models with and without realistic fragmentation
  - Study of the spectator contribution in the centrality determination based on multiplicity

Thank you for your attention!

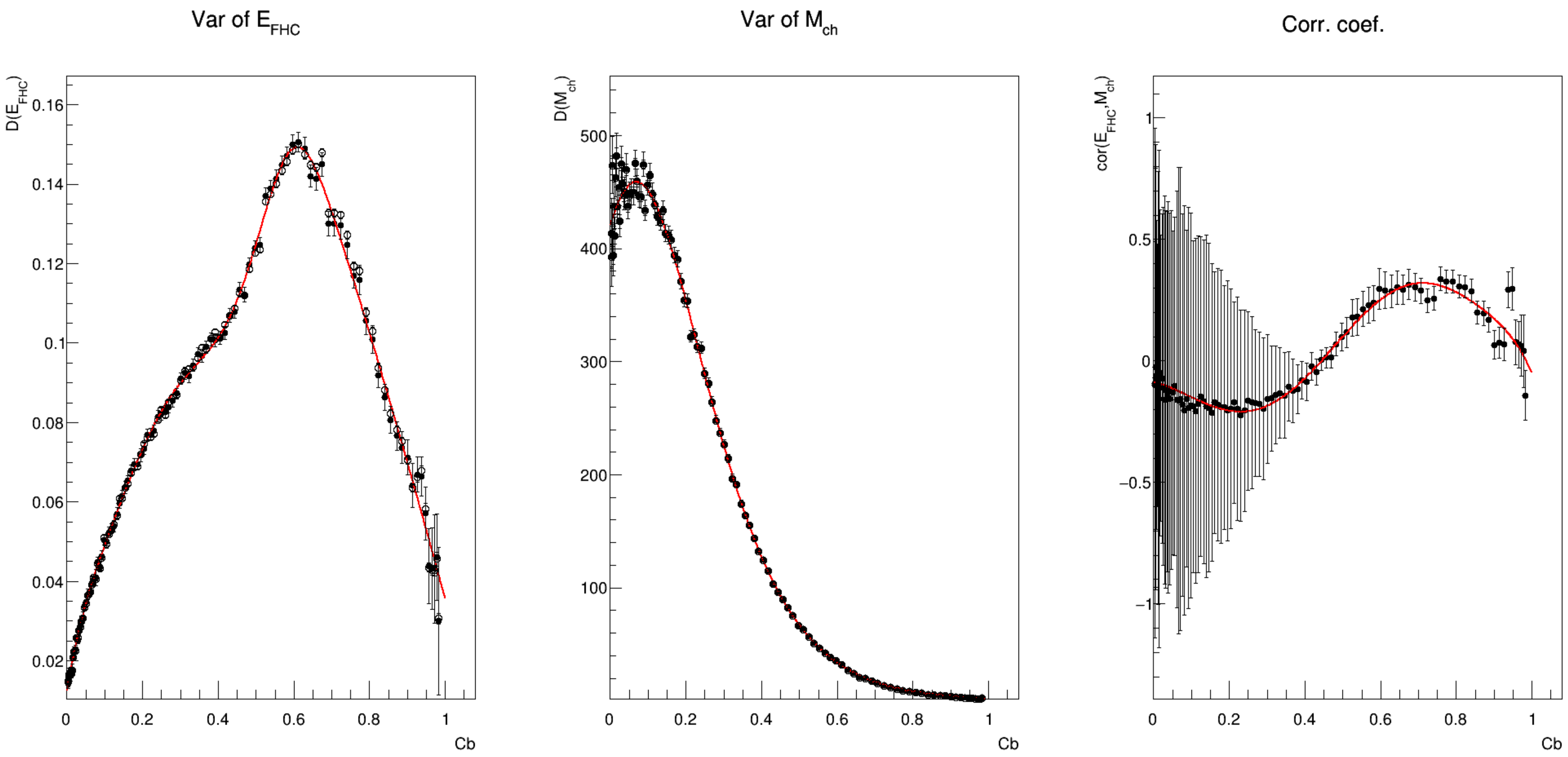
Dependence of the average value of multiplicity and energy on centrality





Good fit quality

Dependence of the variance of multiplicity and energy on centrality



Good fit quality

Probabilistic model of pileup

$M_{pu}(b_1,b_2) = M_1(b_1) + M_2(b_2)$ - pileup as two independent events, with impact parameters b_1, b_2

$$\langle M_{pu}(b_1,b_2)\rangle = \langle M_1(b_1)\rangle + \langle M_2(b_2)\rangle, \quad D(M_{pu}(b_1,b_2)) = D(M_1(b_1)) + D(M_2(b_2))$$

$$P_{pu}(M_{pu} \mid b_1,b_2) = \frac{1}{\Gamma(k_p)\theta_p^{\;2}} M_{pu}^{\;k_p-1} e^{-M_{pu}/\theta_p}$$

- The fluctuation of multiplicity can be describe by Gamma distribution

$$\theta_p = \frac{D(M(b_1,b_2))}{\langle M(b_1,b_2)\rangle}, \quad k_p = \frac{\langle M(b_1,b_2)\rangle}{\theta_p}$$

- The parameters of Gamma distribution

$P_{pu}(M_{pu})$ – the probability distribution of pileup can be calculated as

$$P_{pu}(M_{pu}) = \int_0^{b_{\max}}\int_0^{b_{\max}} P(M_{pu} \mid b_1,b_2)P(b_1)P(b_2)db_1db_2 = \int_0^{c_{b1}}\int_0^{c_{b2}} P_{pu}(M_{pu} \mid c_{b1},c_{b2})dc_{b1}dc_{b2}$$

Corrections for efficiency and pileup

- **Correction for efficiency of normalized multiplicity distribution P(M)**

$$P(M) = \frac{dN}{dM} / N_{ideal}^{ev} = \frac{N_{raw}^{ev}}{N_{ideal}^{ev}} \cdot \frac{1}{N_{raw}^{ev}} \frac{dN_r}{dM} = \frac{1}{K} \cdot Norm.Histogr$$

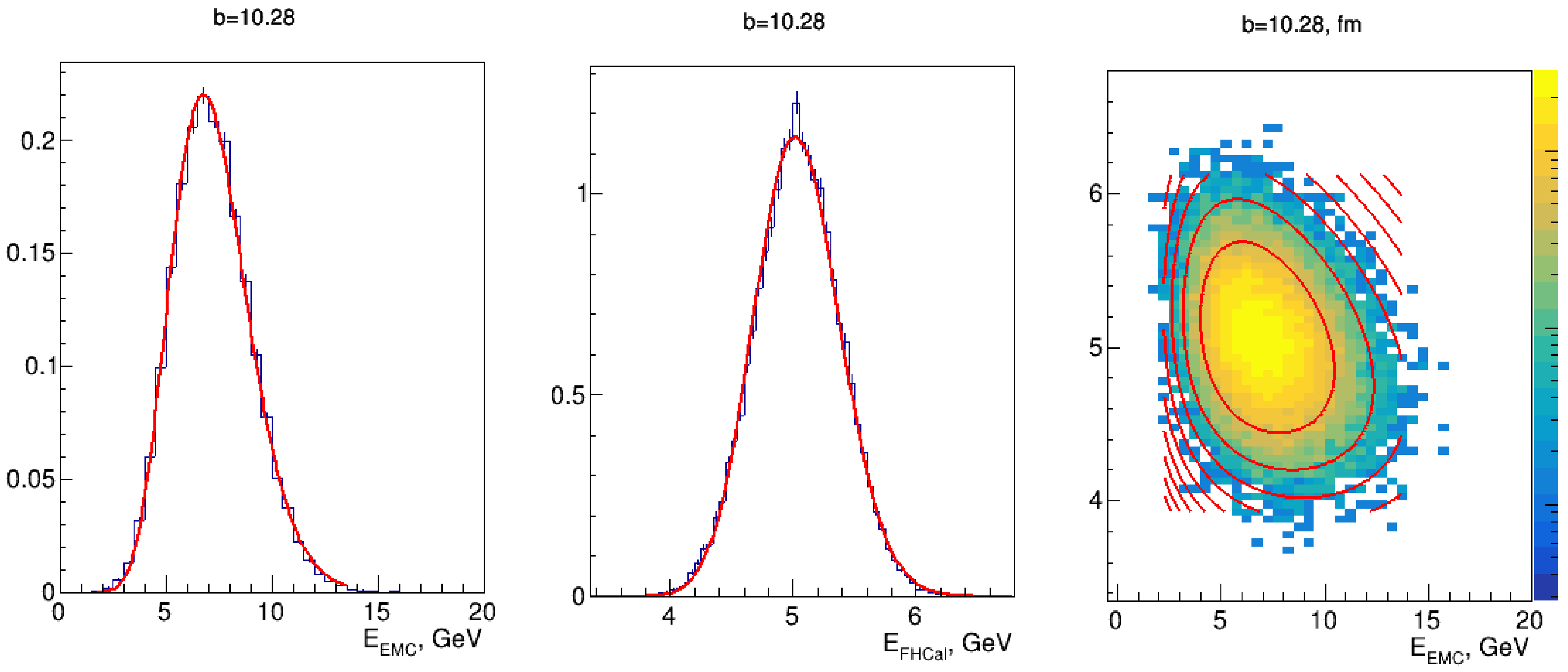
$$Eff = \frac{N_{raw}^{ev}}{N_{ideal}^{ev}} = \frac{1}{K}$$ integral efficiency

- **Fit function for multiplicity distribution P(M)**

$$F(M) = K \cdot P_{total}(M), \;\; P_{total}(M) = N_p \cdot P_{pu}(M) + (1 - N_p) \cdot P(M)$$

μ, f, k, K, N_p - fit parameters, $F(M) -$ fit function, corrected for efficiency and pileup

The fluctuation of observables at fixed impact parameter

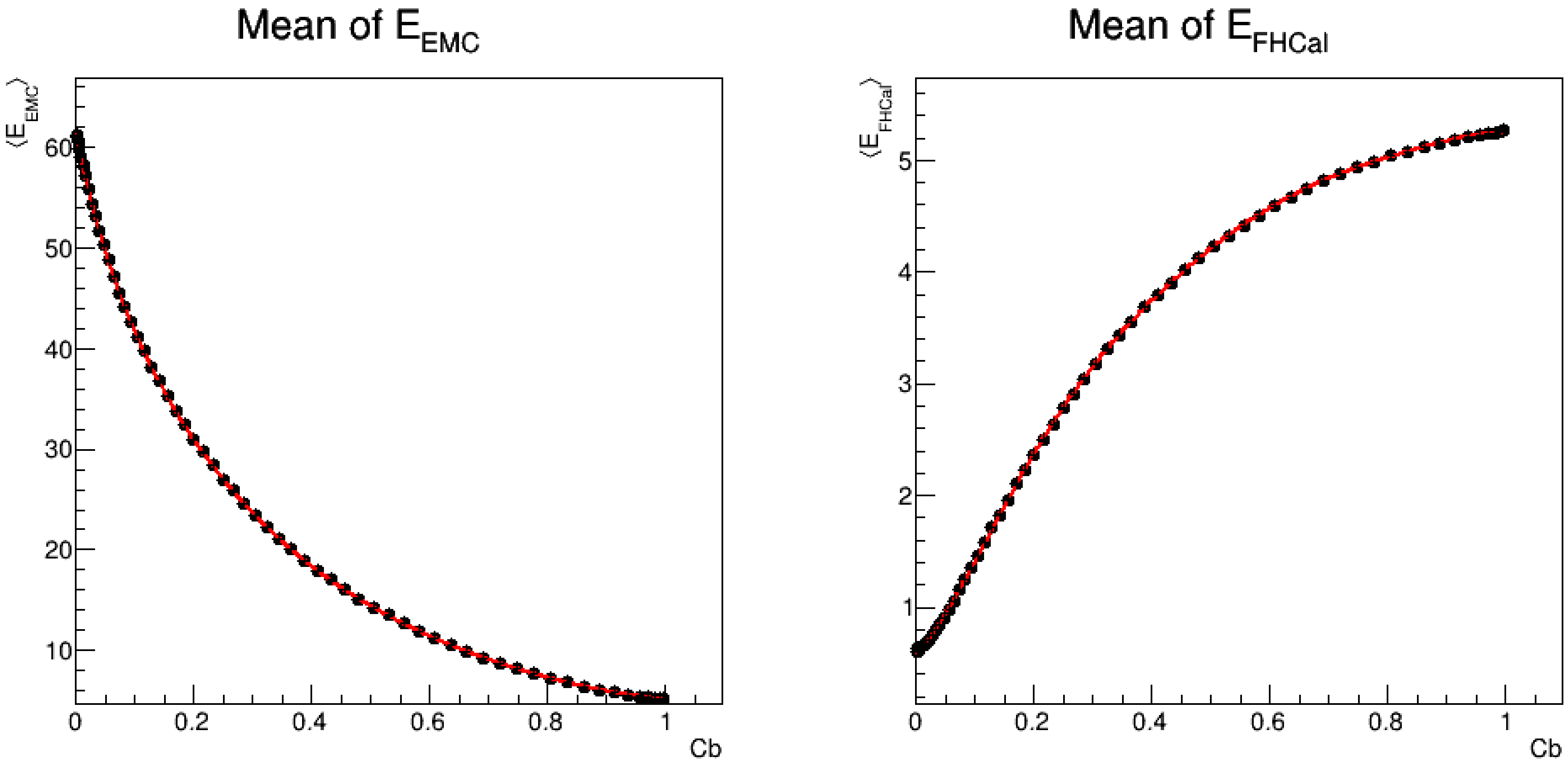


The distribution of observables at a fixed impact parameter is well described by the 2D gamma distribution

- Find probability of b for fixed range of observables using Bayes' theorem:

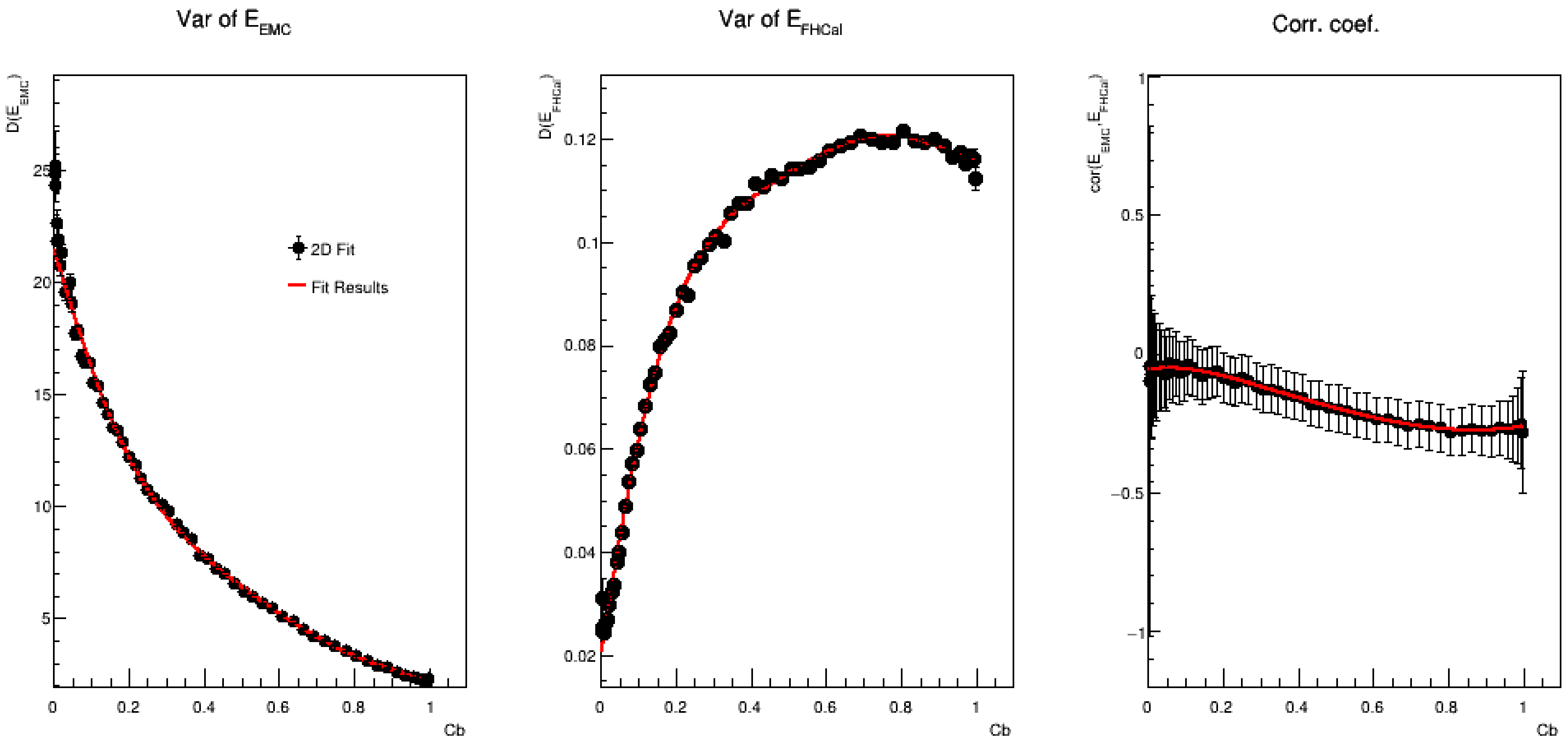
$$P(b \mid E_1 < E < E_2, \mathrm{E}_F^1 < E_F < E_F^2) = P(b) \frac{\int_{E_1}^{E_2} \int_{M_1}^{M_2} P(E, E_F \mid c_b) dE_F dE}{\int_{E_1}^{E_2} \int_{E_F^1}^{E_F^2} \int_0^1 P(E, E_F \mid c_b) dE_F dE dc_b}$$

Dependence of the average value of energy in EMC and FHCal on centrality



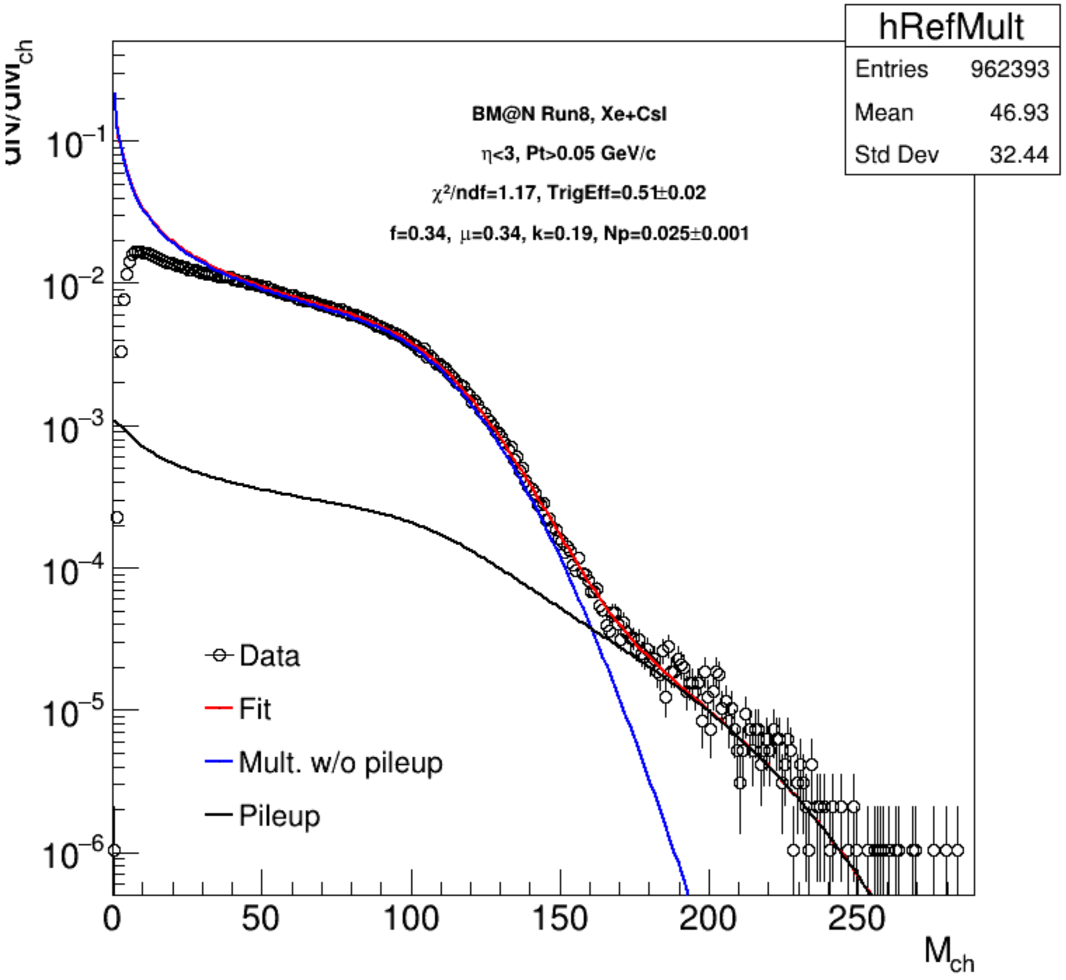
Mean values for E_{EMC} and E_{FHCal} from UrQMD and parametrization are in a good agreement over all c_b

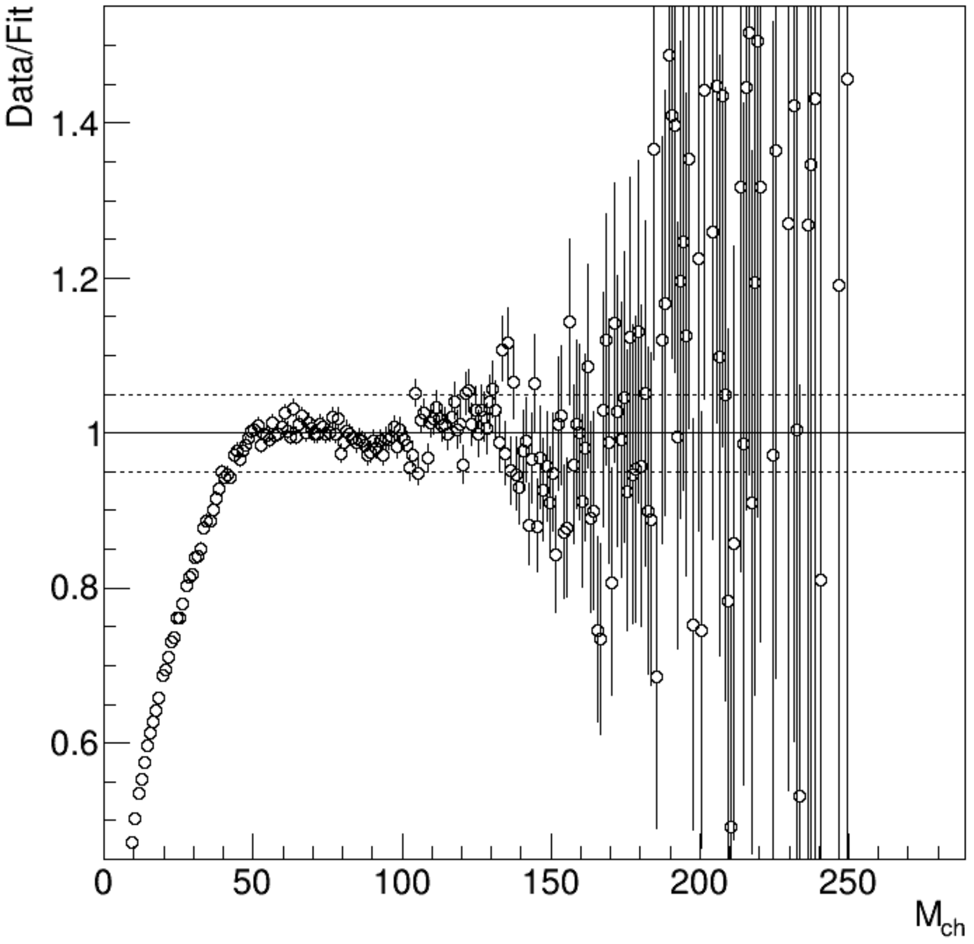
Dependence of the variance of observables on centrality



Variance for E_{EMC} and E_{FHCal} from UrQMD and parametrization are in a good agreement over all c_b

Fit results



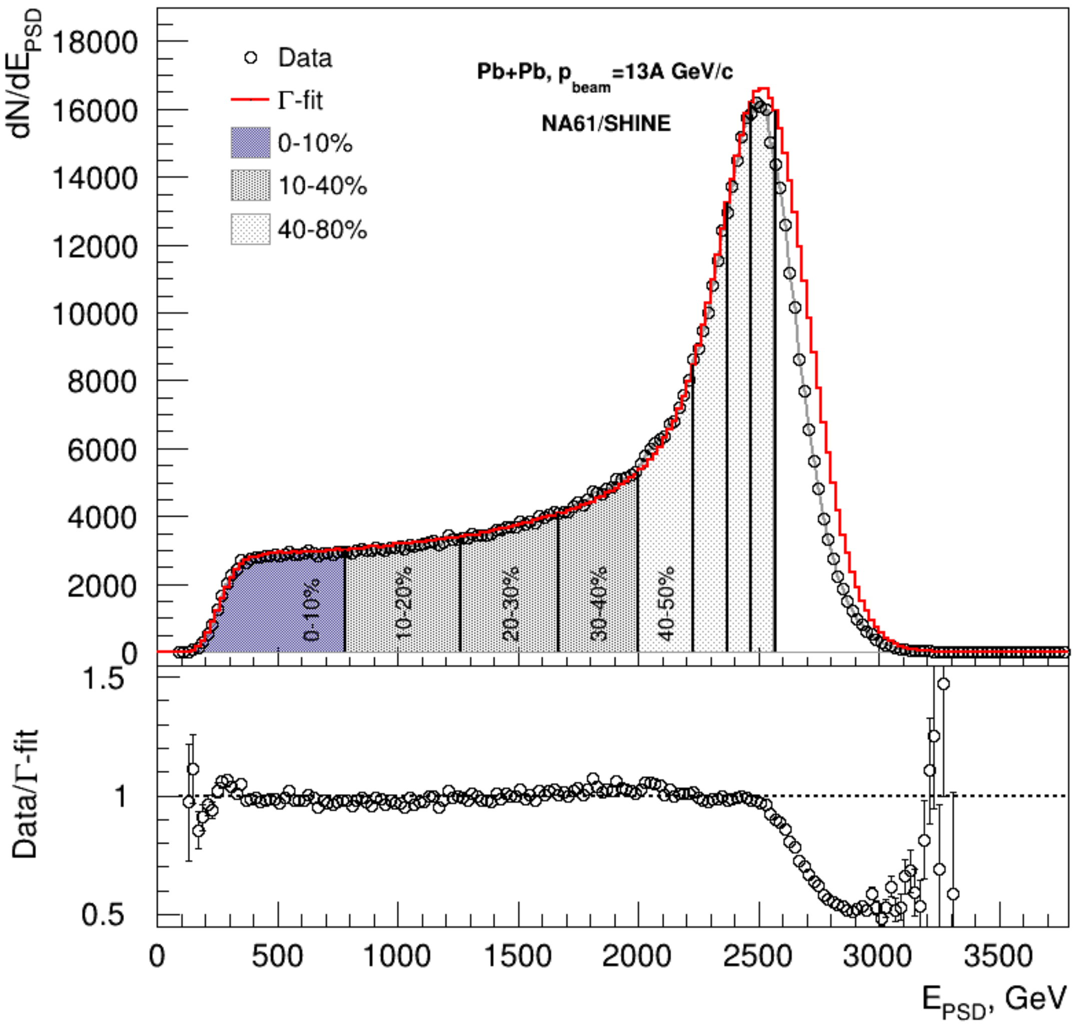


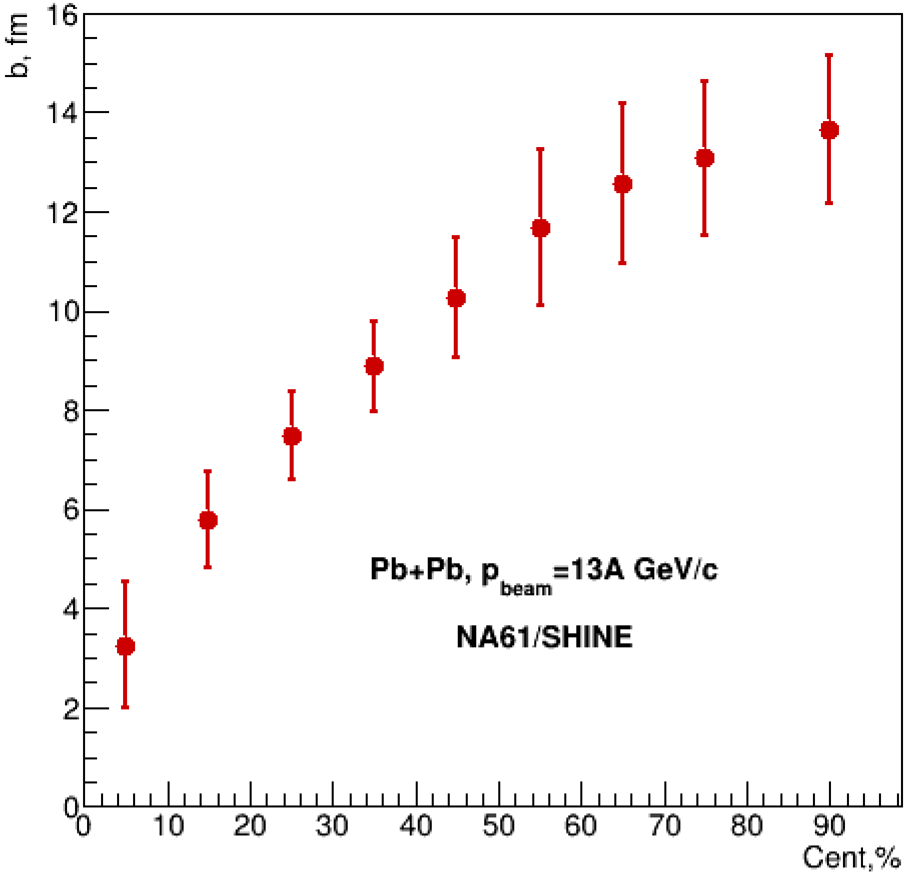
Vertex Cuts: CCT2, N_{vtxTr}>1,|$V_{x,y}$ –(0.3,0.14)|<1 cm, |V_z -0.07|< 0.2 cm

Good agreement with fit

Track selection: Nhit>4, η<3, Pt>0.05 GeV/c

Fit results for NA61





The method reproduces the energy distribution well.
The difference in the peripheral region is due to the trigger efficiency

Find probability of *b* using Bayes' theorem:

$$P(b\,|\,E_1 < E < E_2) = P(b)\int_{E_1}^{E_2} P(b\,|\,E)dE \Bigg/ \int_{E_1}^{E_2} P(E)dE$$