Bayesian approach for centrality determination in nucleus-nucleus collisions with forward hadron calorimeter at the BM@N experiment

Idrisov Dim, Fedor Guber, Nikolay Karpushkin, Parfenov Peter

INR RAS, Moscow, Russia

08/10/24

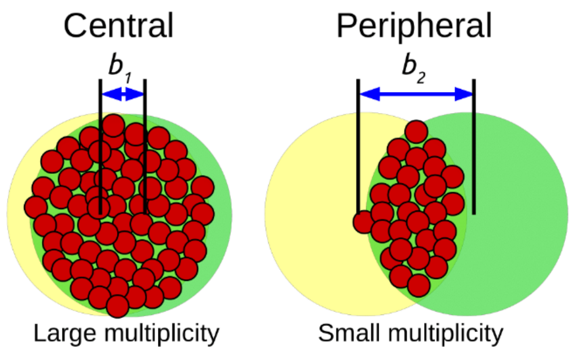
Centrality

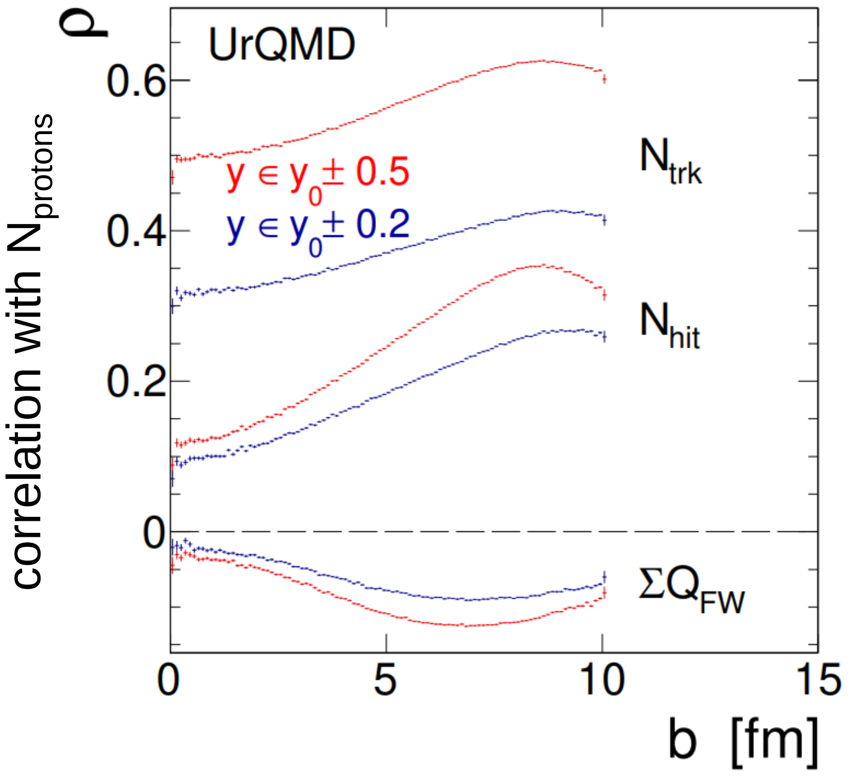
- Evolution of matter produced in heavy-ion collisions depend on its initial geometry

- Centrality procedure maps initial geometry parameters with measurable quantities (multiplicity or energy of the spectators)

- **This allows comparison of the future BM@N results with the data from other experiments (STAR BES, NA49/NA61 scans) and theoretical models**

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

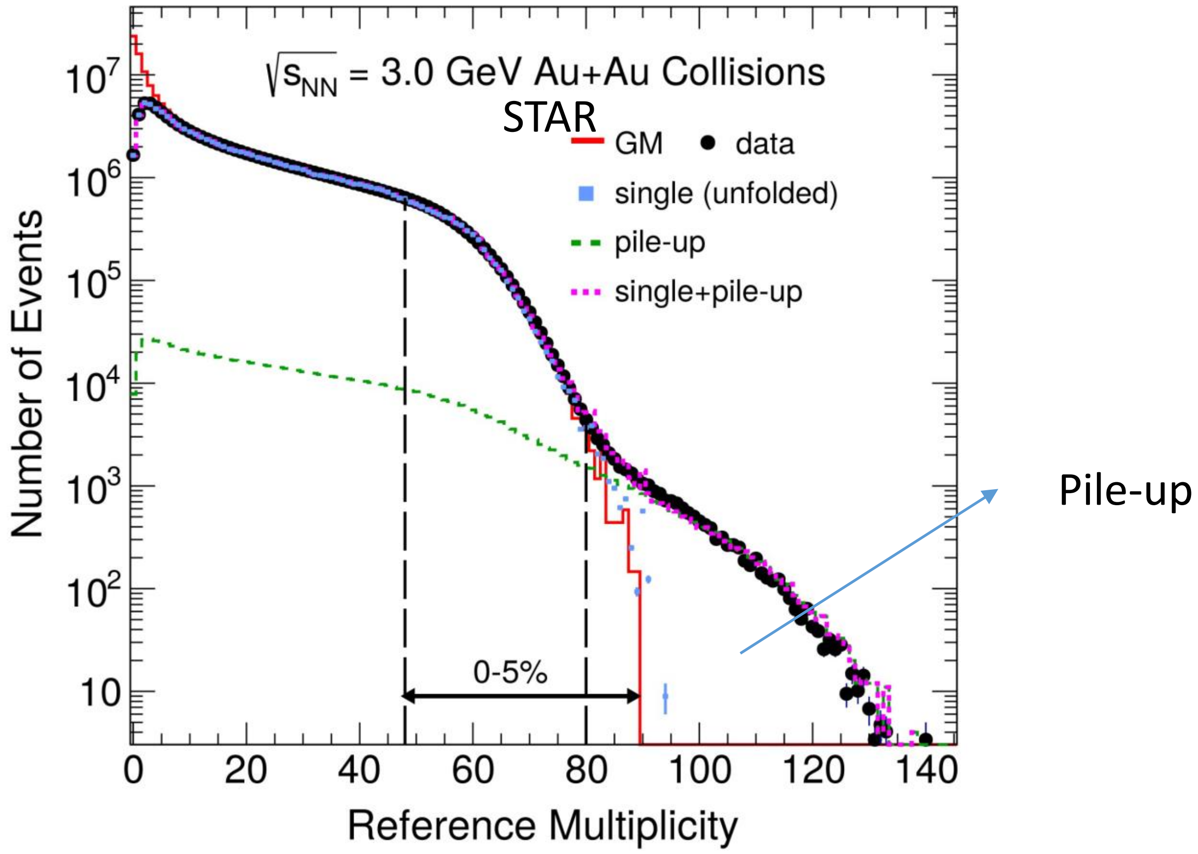




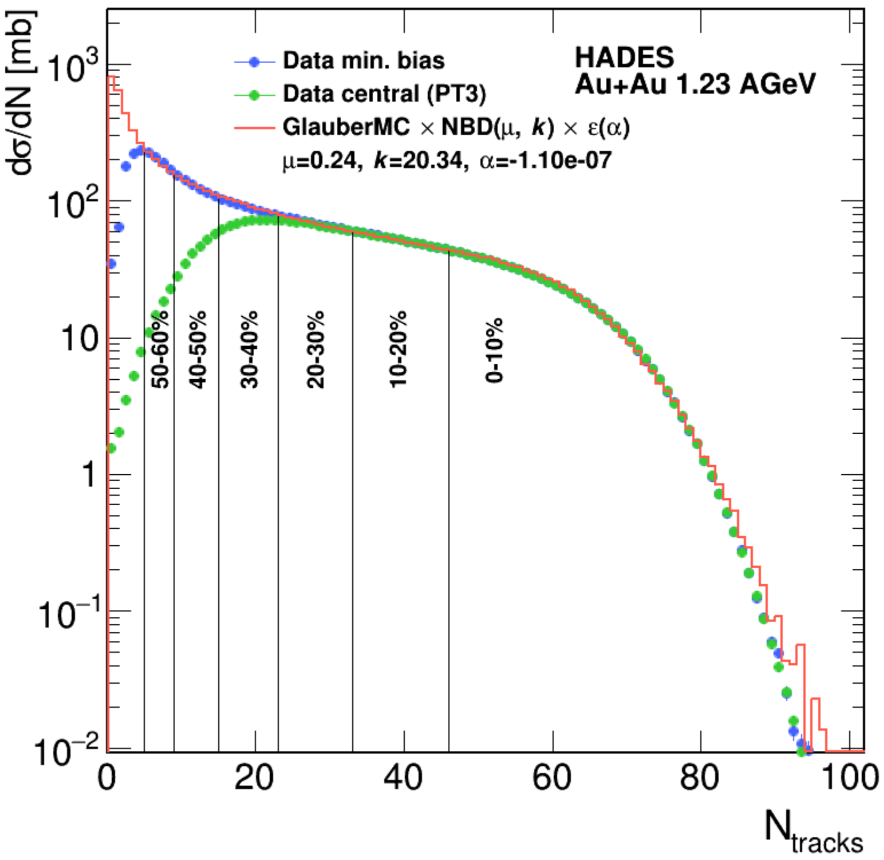
HADES; Phys.Rev.C 102 (2020) 2, 024914

- 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

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.



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/2112.00240

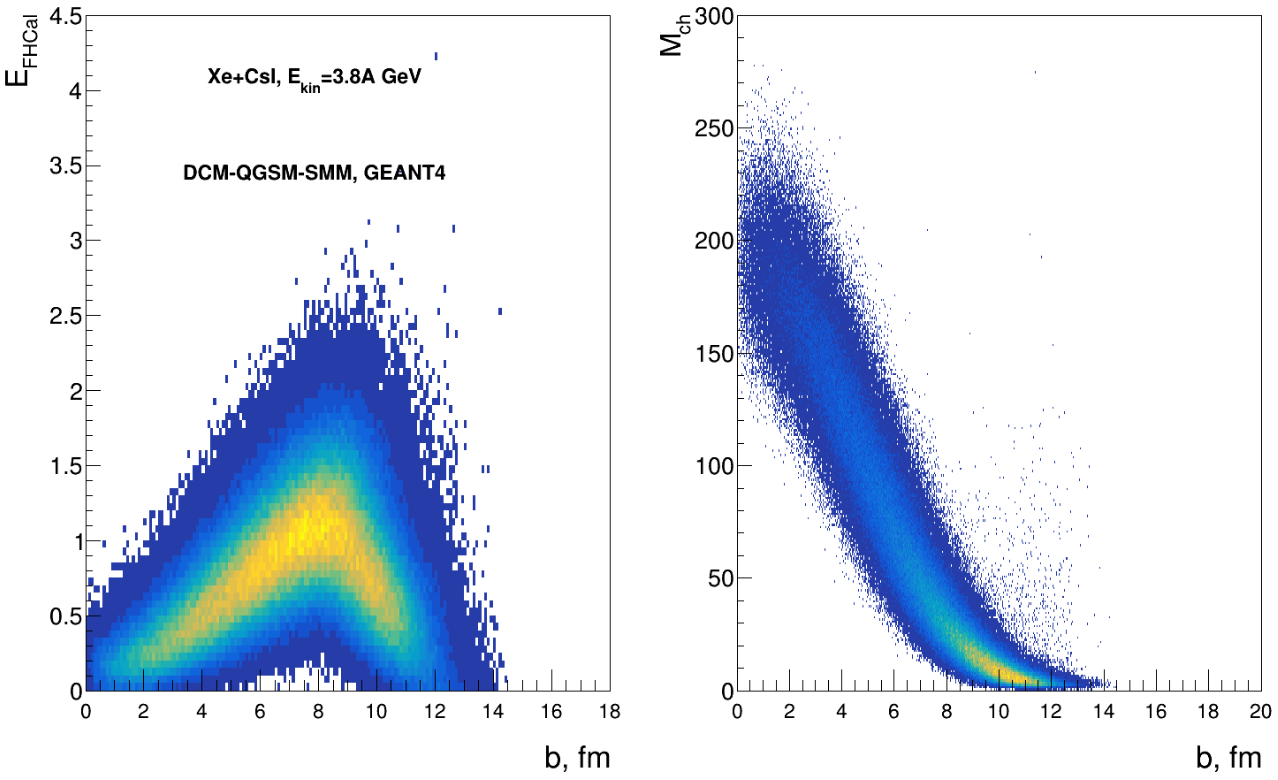
https://arxiv.org/abs/1712.07993

Centrality determination in BM@N

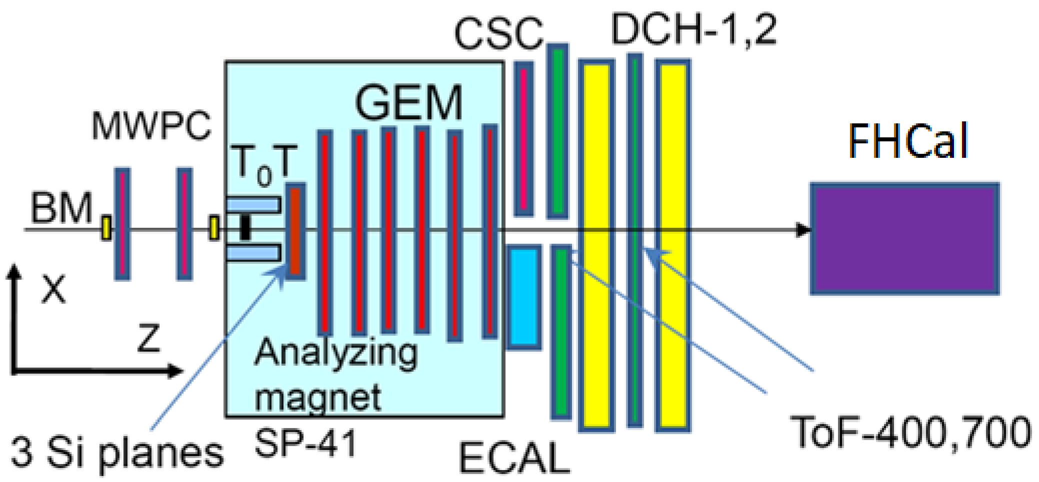
DCM-QGSM-SMM → GEANT4 → Reconstruction → Centrality determination

Centrality determination: Multiplicity of produced charged particles in tracking system

Simulated data sets: Xe+Cs, Nev=500k



Relation between impact parameter and track multiplicity



BM@N setup overview

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

- The fluctuation kernel for multiplicity at fixed impact parameter can be describe by 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

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

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

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

- 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

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

Corrections for efficiency and pileup

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

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

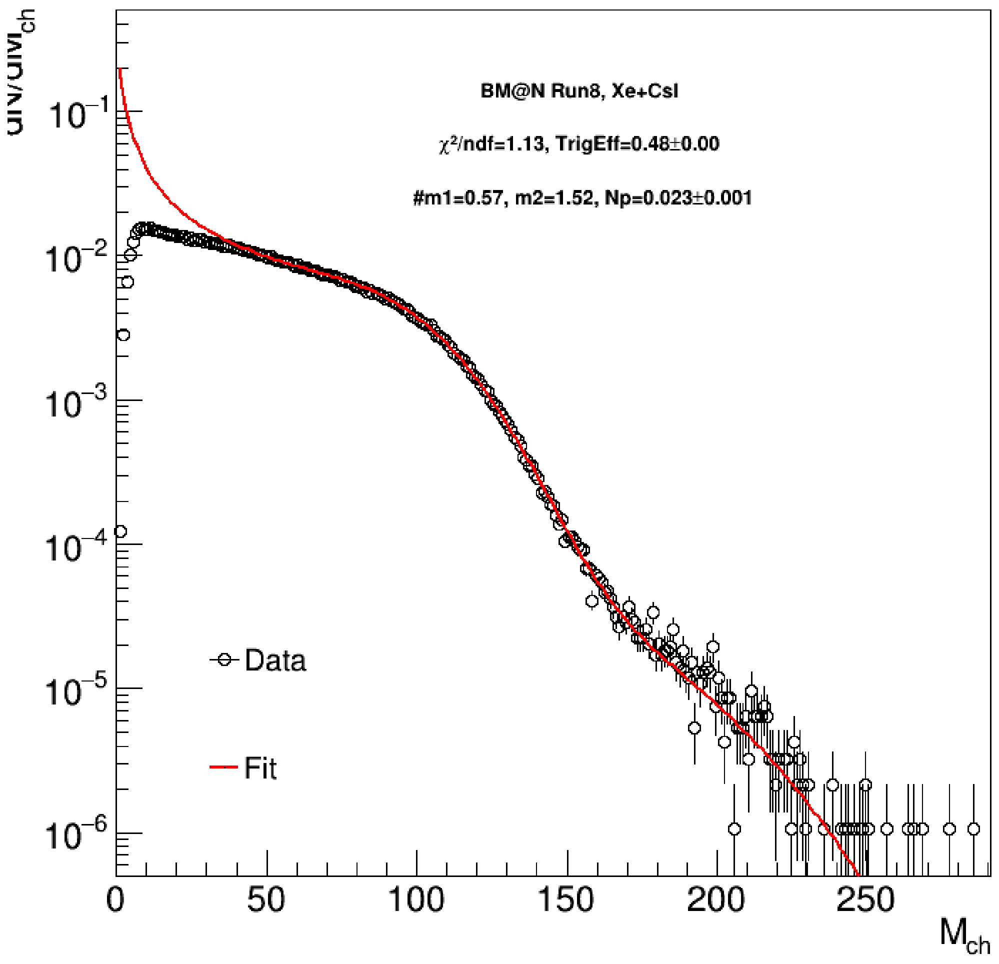
$$Eff=\frac{N_{raw}^{ev}}{N_{ideal}^{ev}}=\frac{1}{K}$$ integral efficiency

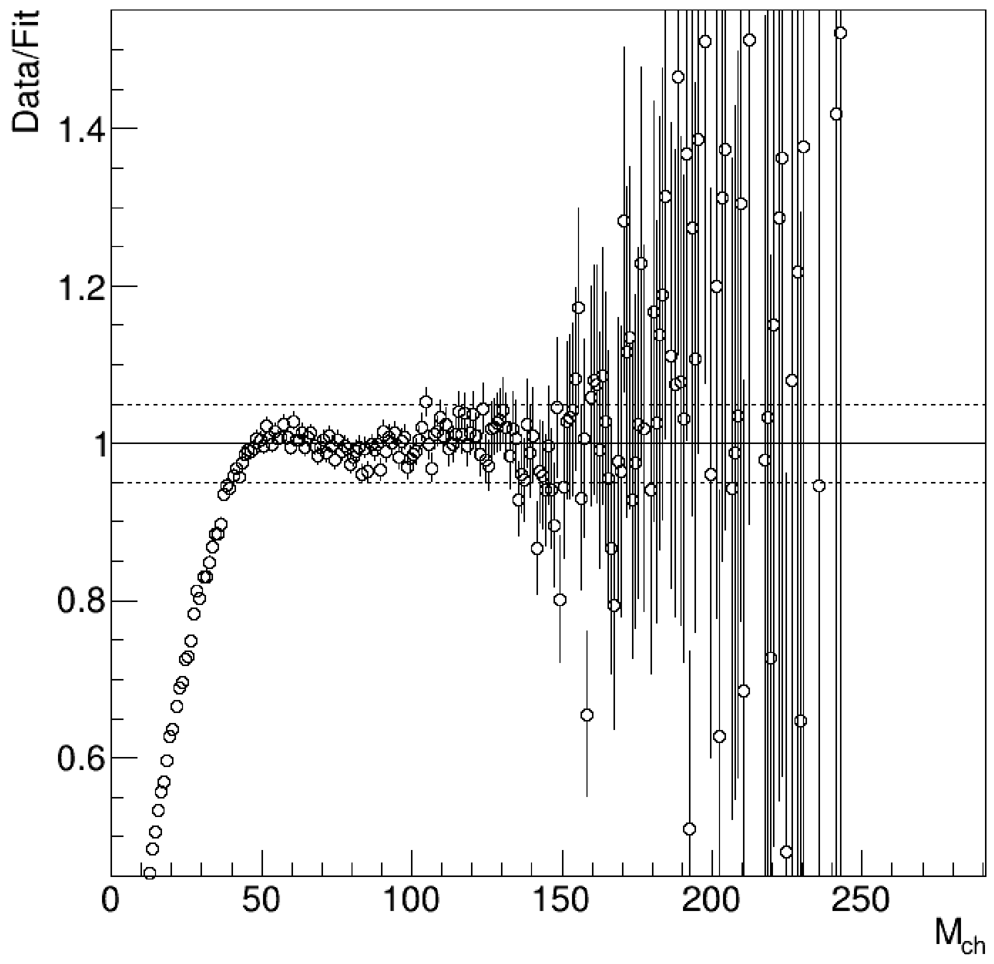
- **Fit function for multiplicity distribution P(M)**

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

m_1,m_2,K,N_p- fit parameters, $F(M)-$ fit function, corrected for efficiency and pileup

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

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.:

$$P(E,M \mid c_b) = G_{2D}(E,M,\langle E\rangle,\langle M\rangle,D(E),D(M),R)$$

$$c_b = \int_0^b P(b')db'$$ – 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

$$R(E,M) = \varepsilon_1 \cdot m_1 \cdot R(E',M')\sqrt{\frac{D(E')D(M')}{D(E)D(M)}}$$

$\varepsilon_0, \varepsilon_1, \varepsilon_2, m_1, m_2$ - fit parameters

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

$$\langle E\rangle = \varepsilon_1\langle E'(c_b)\rangle + \varepsilon_0, \quad D(E) = \varepsilon_2 D(E'(c_b))$$

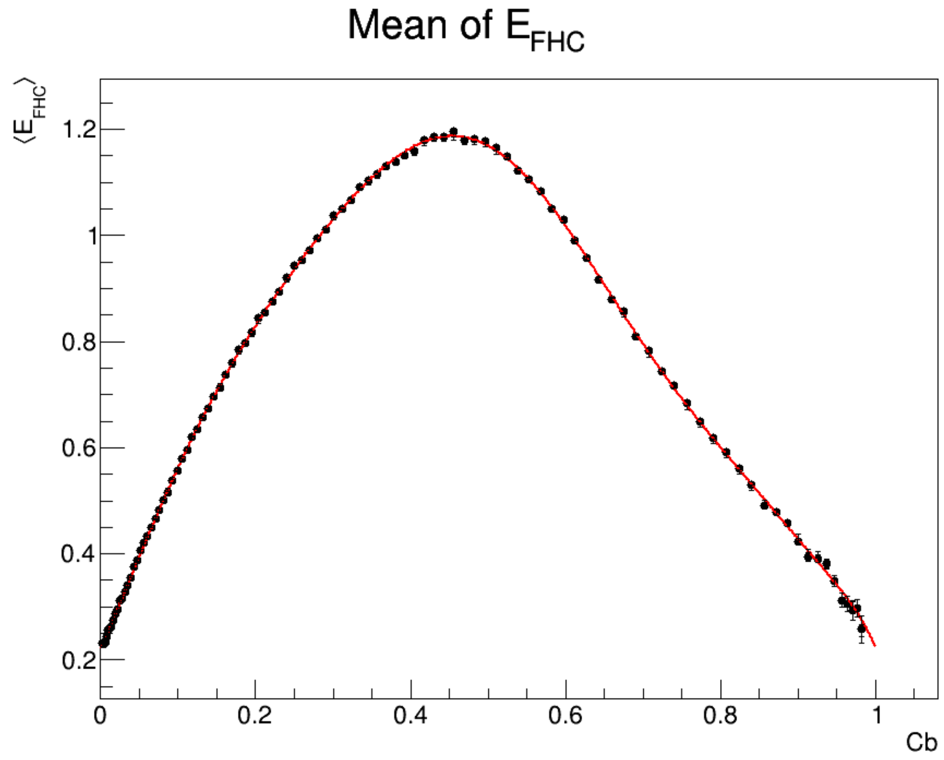
$$\langle M\rangle = m_1\langle M'(c_b)\rangle, \quad D(M) = m_1^{\,2}\cdot D(M') + m_1\cdot m_2\langle M'\rangle$$

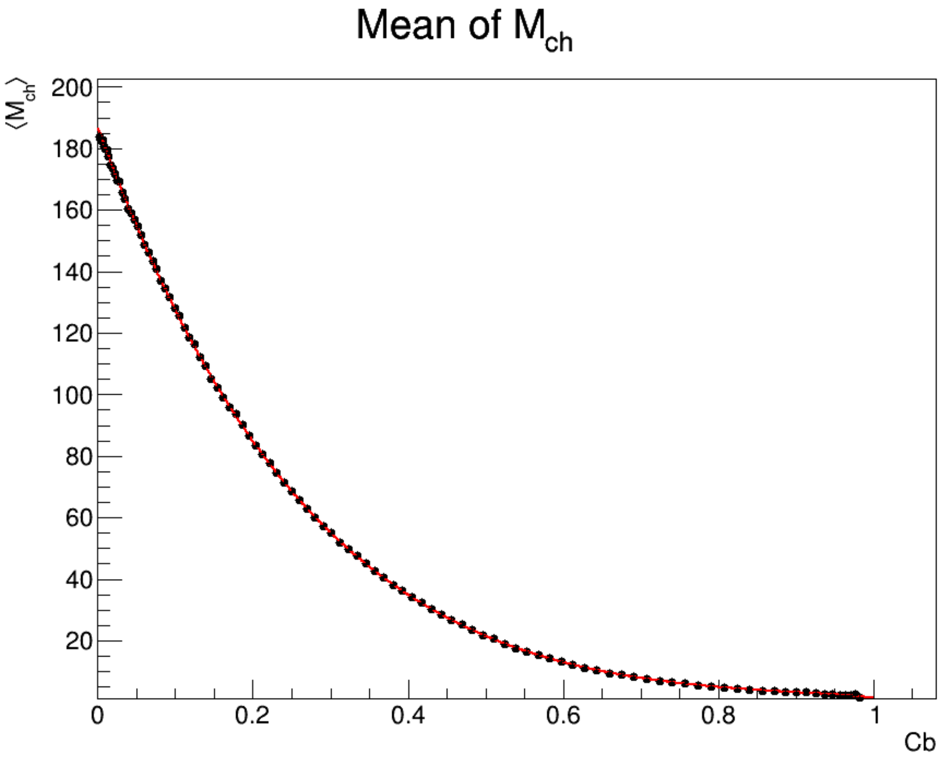
$\langle E'(c_b)\rangle, \quad D(E'(c_b))$ - can be approximated by polynomials

$$\langle E'(c_b)\rangle = \sum_{j=1}^{12} a_j c_b^j, \quad D(E'(c_b)) = \sum_{j=1}^{19} b_j c_b^j$$

$$\langle M'(c_b)\rangle = \sum_{j=1}^{12} a_j c_b^j, \quad D(M'(c_b)) = \sum_{j=1}^{6} b_j c_b^j$$

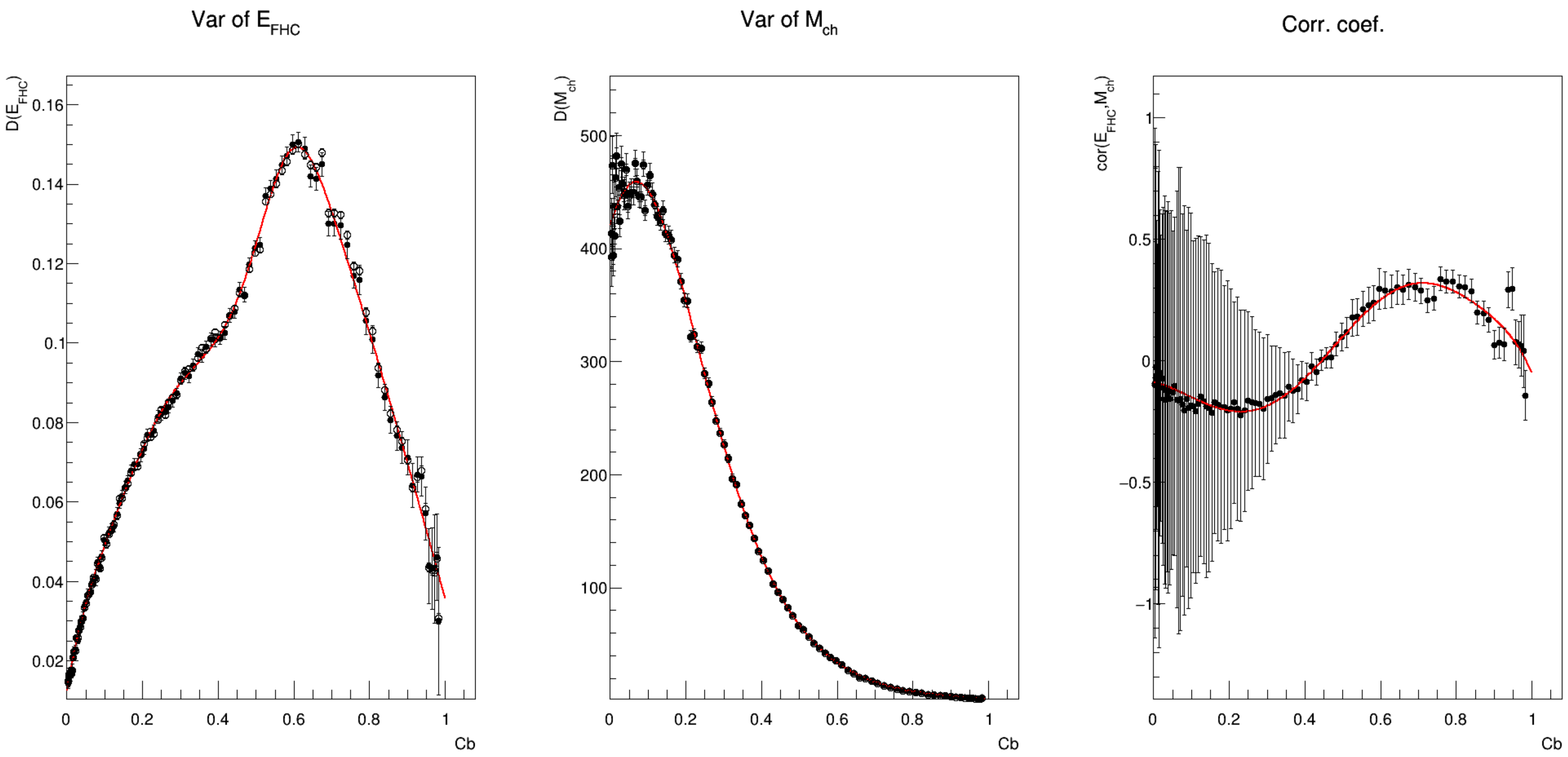
Dependence of the average value of multiplicity and energy on centrality





Good fit quality

Dependence of the variance of multiplicity and energy on centrality



Good fit quality

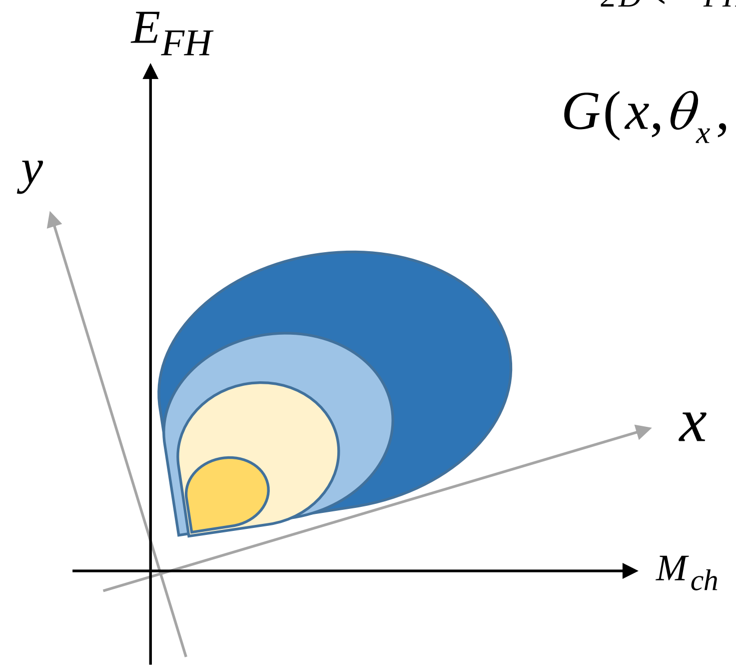
2D Gamma distribution

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

Then the two-dimensional distribution in the new coordinate system will be

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

$$G(x,\theta_x,k_x)\cdot G(y,\theta_y,k_y)=\frac{(x)^{k_x(c_b)-1}e^{-x/\theta_x}}{\Gamma(k_x(c_b))\theta_x^2}\cdot\frac{(y)^{k_y(c_b)-1}e^{-y/\theta_y}}{\Gamma(k_y(c_b))\theta_y^2}$$



$$\theta_x=\frac{D(x)}{\langle x\rangle},\quad k_x=\frac{\langle x\rangle^2}{D(x)},\ \theta_y=\frac{D(y)}{\langle y\rangle},\quad k_y=\frac{\langle y\rangle^2}{D(y)}$$

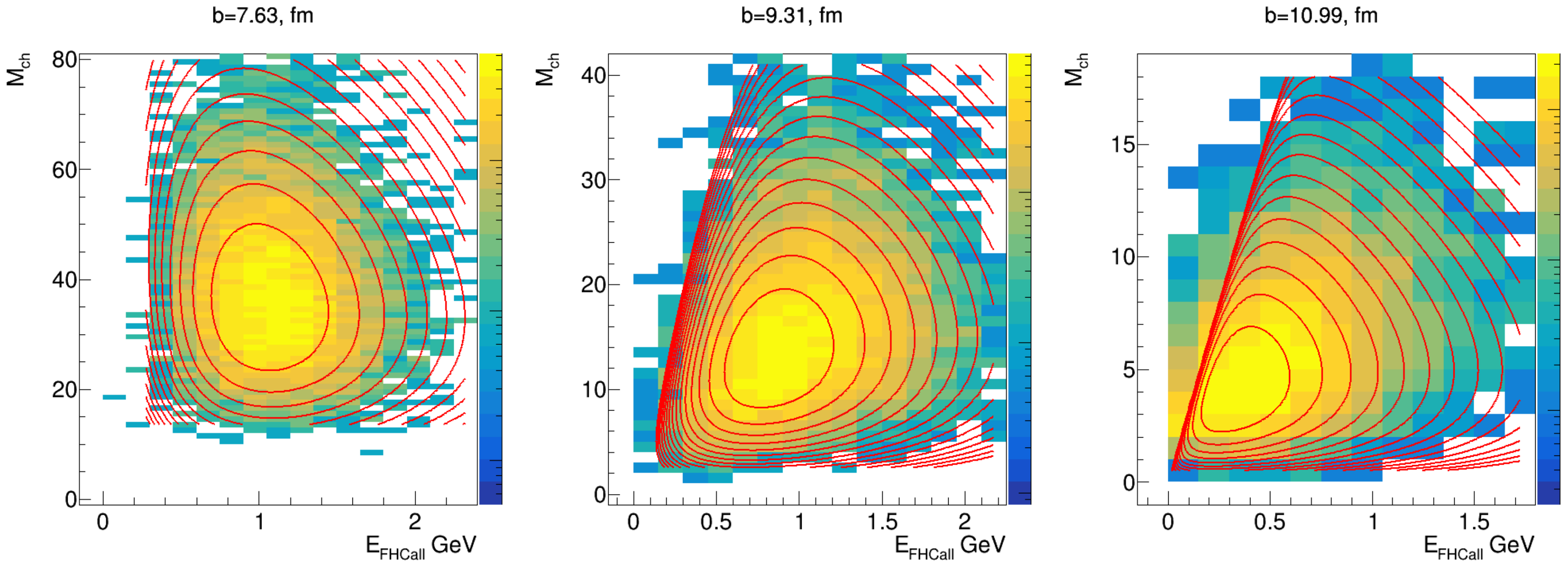
$$\alpha=\arctan\left(\frac{2\sqrt{D(E)D(M)}R(E,M)}{D(E)-D(M)}\right)$$

mean value and variance in the new coordinate system

$$\langle x\rangle=\cos(\alpha)\langle E\rangle+\sin(\alpha)\langle M\rangle \qquad D(x)=D(E)\cos(\alpha)^2+R(E,M)\sqrt{D(E)D(M)}\sin(2\alpha)+D(M)\sin(\alpha)^2$$

$$\langle y\rangle=-\sin(\alpha)\langle E\rangle+\cos(\alpha)\langle M\rangle \qquad D(y)=D(E)\sin(\alpha)^2-R(E,M)\sqrt{D(E)D(M)}\sin(2\alpha)+D(M)\cos(\alpha)^2$$

The fluctuation of energy and multiplicity at fixed impact

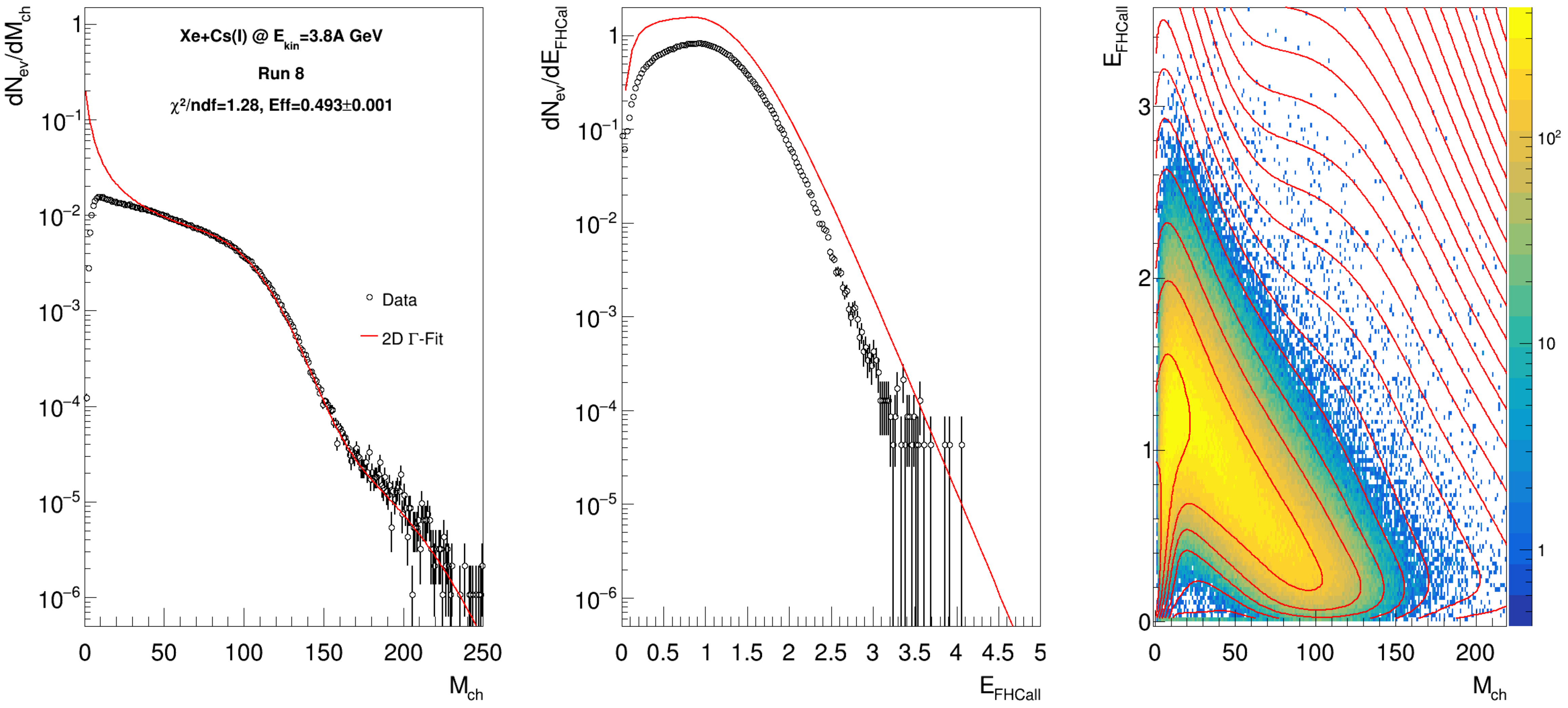


The distribution of energy and multiplicity at a fixed impact parameter is well described by the gamma distribution

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

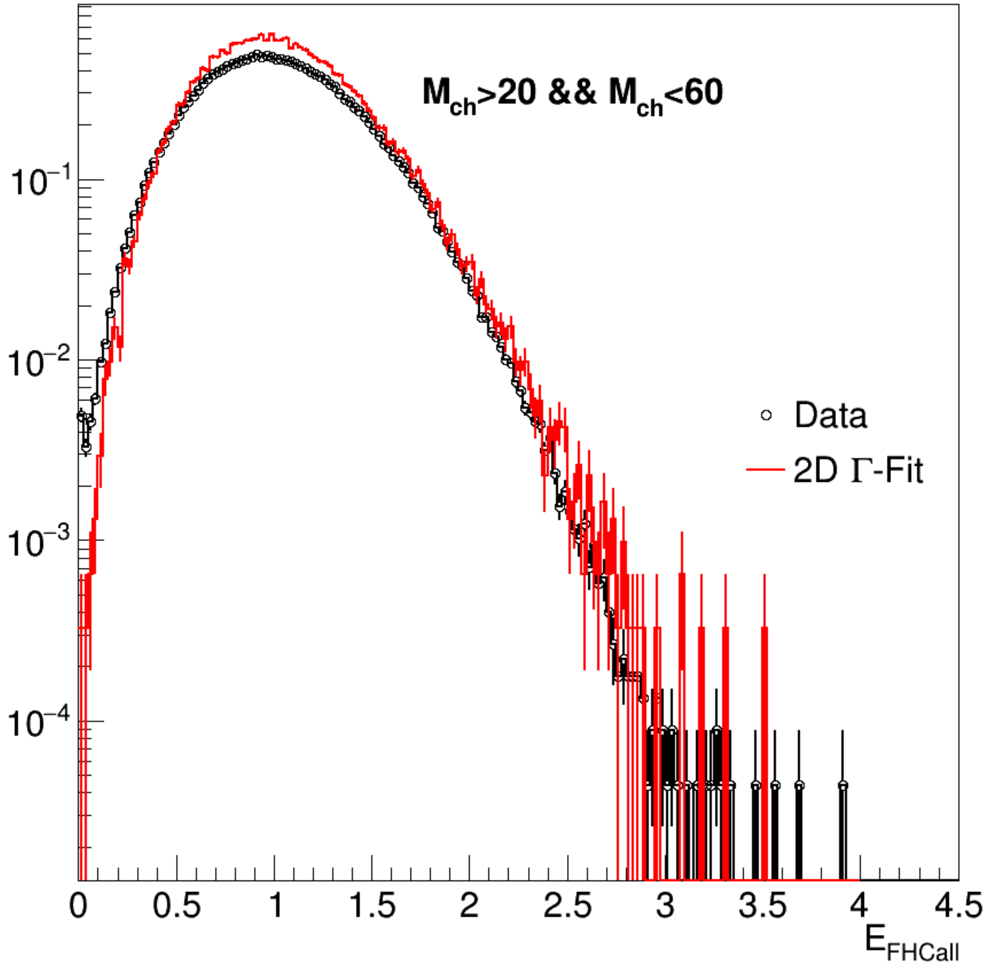
$$P(b \mid E_1 < E < E_2, M_1 < M < M_2) = P(b)\frac{\int_{E_1}^{E_2}\int_{M_1}^{M_2} P(E,M \mid c_b)\,dM\,dE}{\int_{E_1}^{E_2}\int_{M_1}^{M_2}\int_{0}^{1} P(E,M \mid c_b)\,dM\,dE\,dc_b}$$

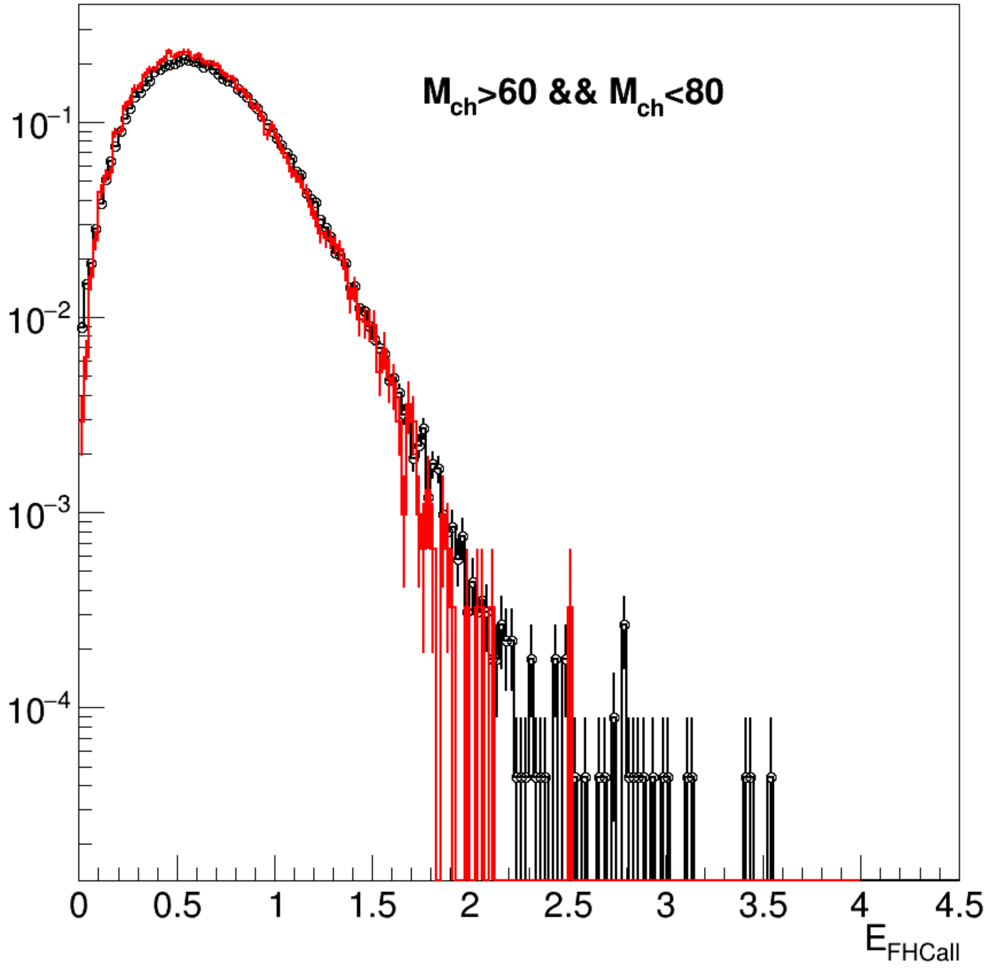
2D fit results



Good agreement between fit and data.

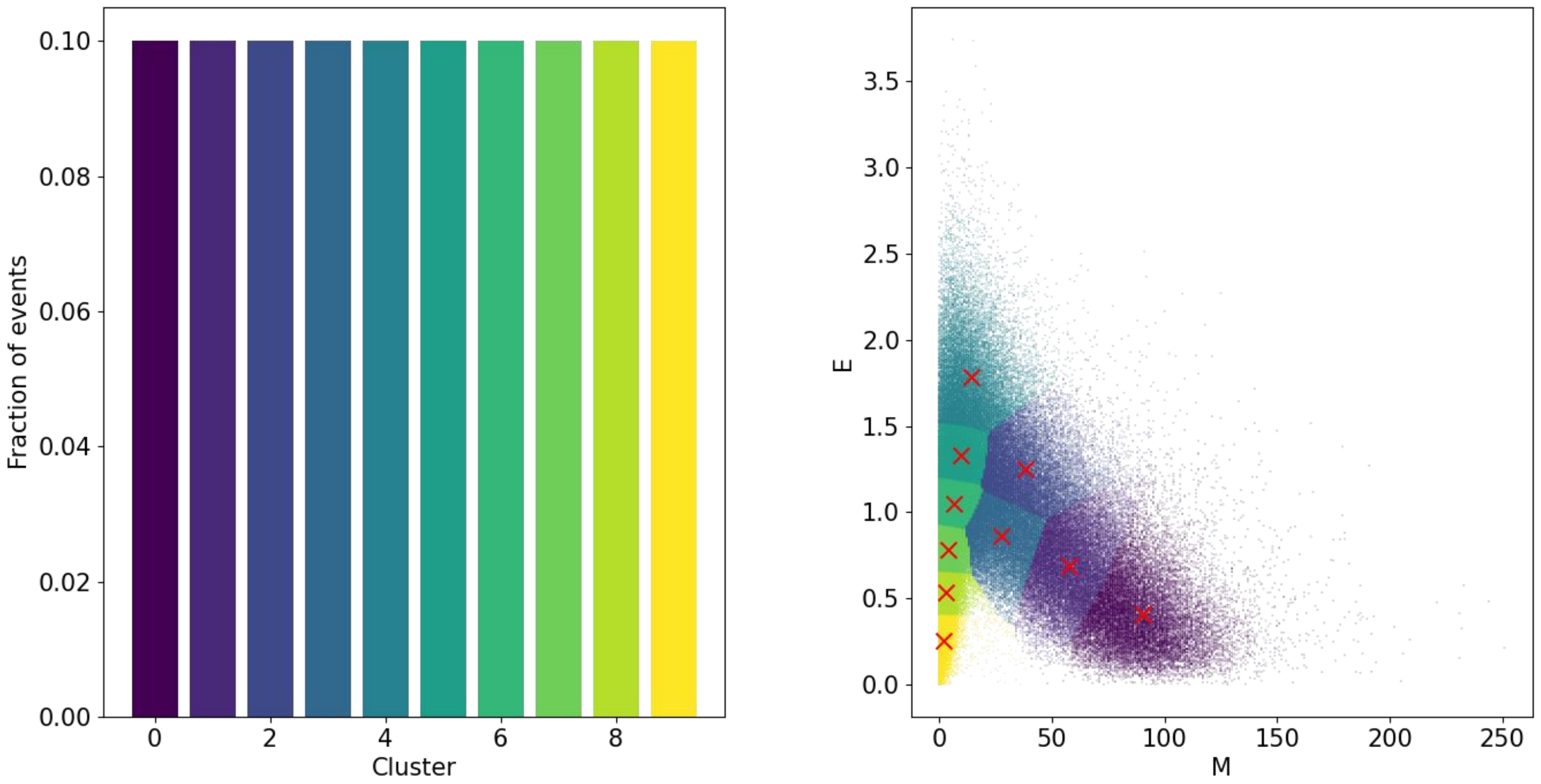
Energy distribution





Good agreement between fit and data.

Clusterization with k means for centrality classes



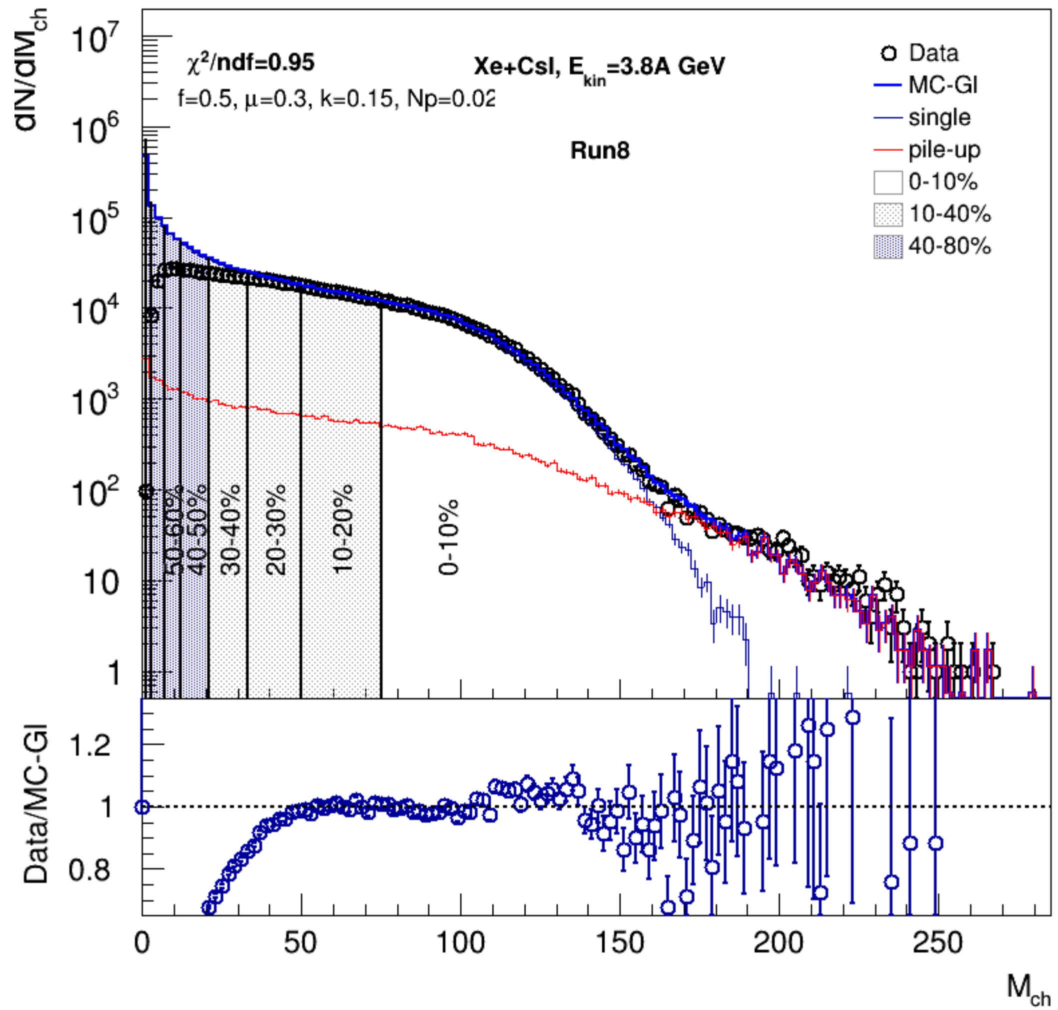
the bivariate fit distribution was divided into 10 centrality classes

MC-Glauber based centrality framework

MC Glauber data → Evaluate N_a: $N_a = fN_{part}+(1-f)N_{coll}$ → Call $NBD(\mu,k) \times N_a \times \varepsilon(\alpha)$ → Build multiplicity fitting function

Input multiplicity distribution

Build multiplicity fitting function + Input multiplicity distribution → Evaluate χ^2 → Minimize χ^2 to find α, μ, k



NBD – negative binomial distribution

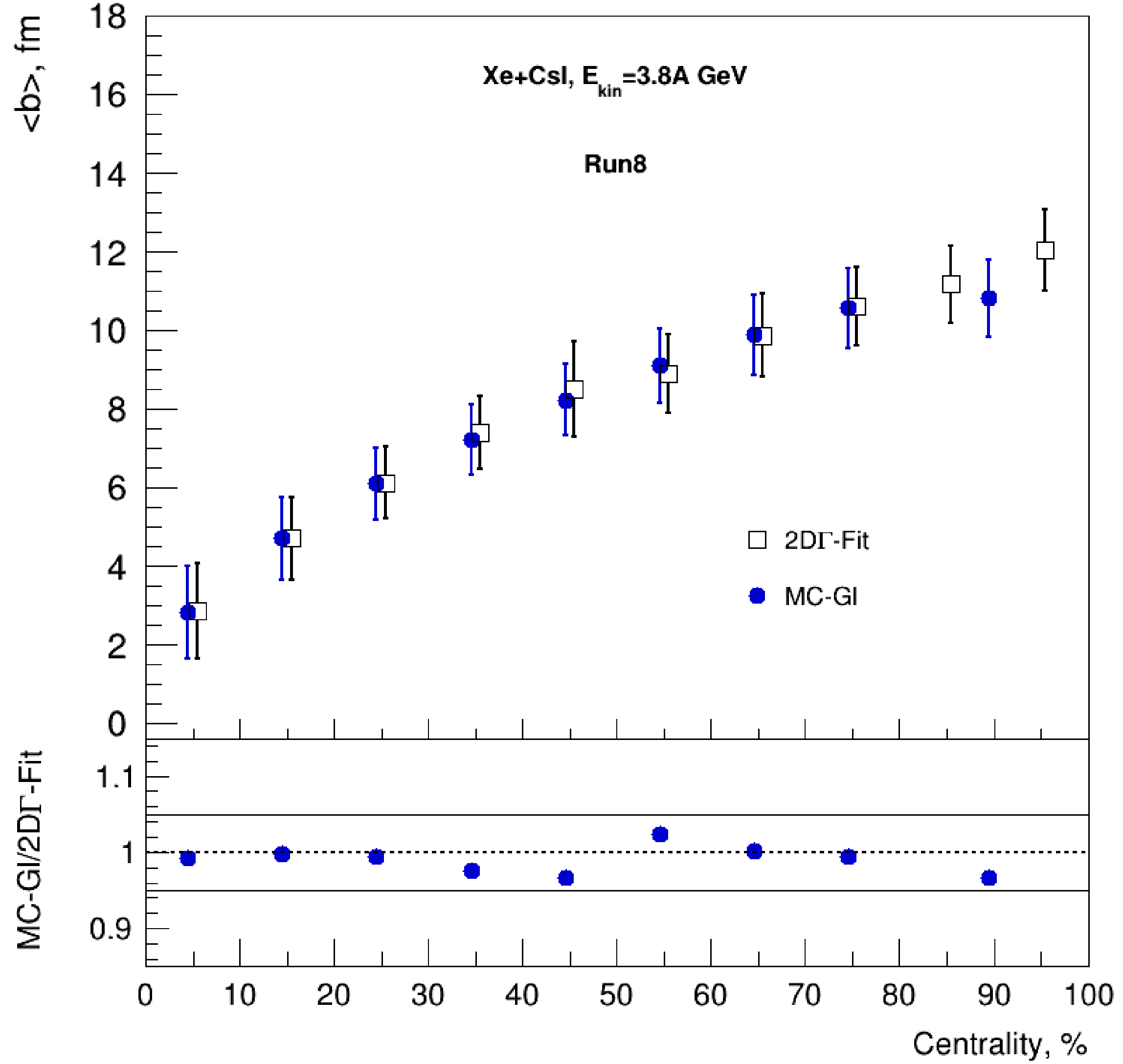
Parameters of the fit:

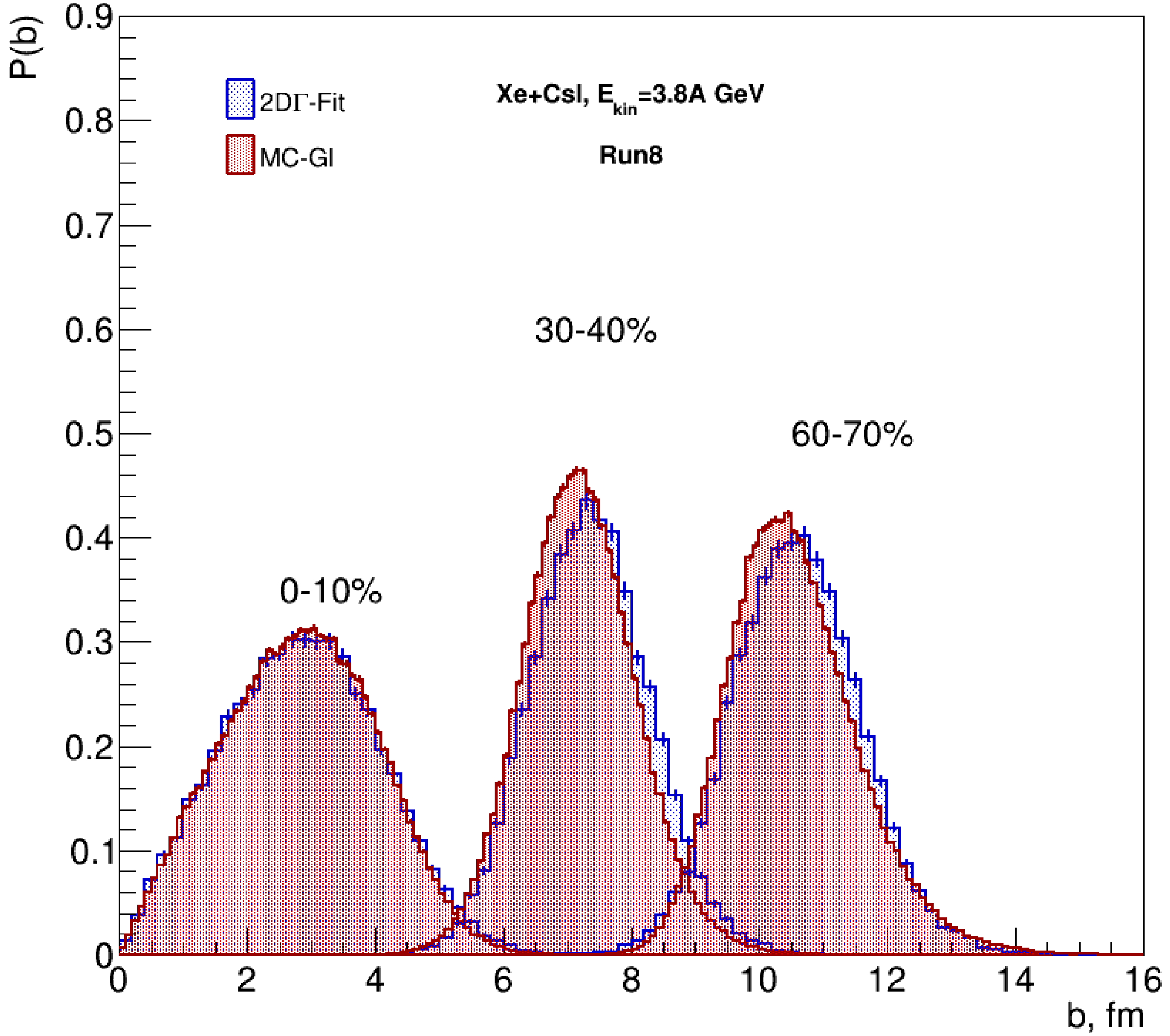
- **α** – coefficient in efficiency function
- **μ** – mean multiplicity value
- **k** – width of the multiplicity distribution, can be connected to the fluctuations

Implamantation for MPD: https://github.com/FlowNICA/CentralityFramework

P. Parfenov, et al., Particles. 2021; 4(2):275-287

Comparison with MC-Glauber fit





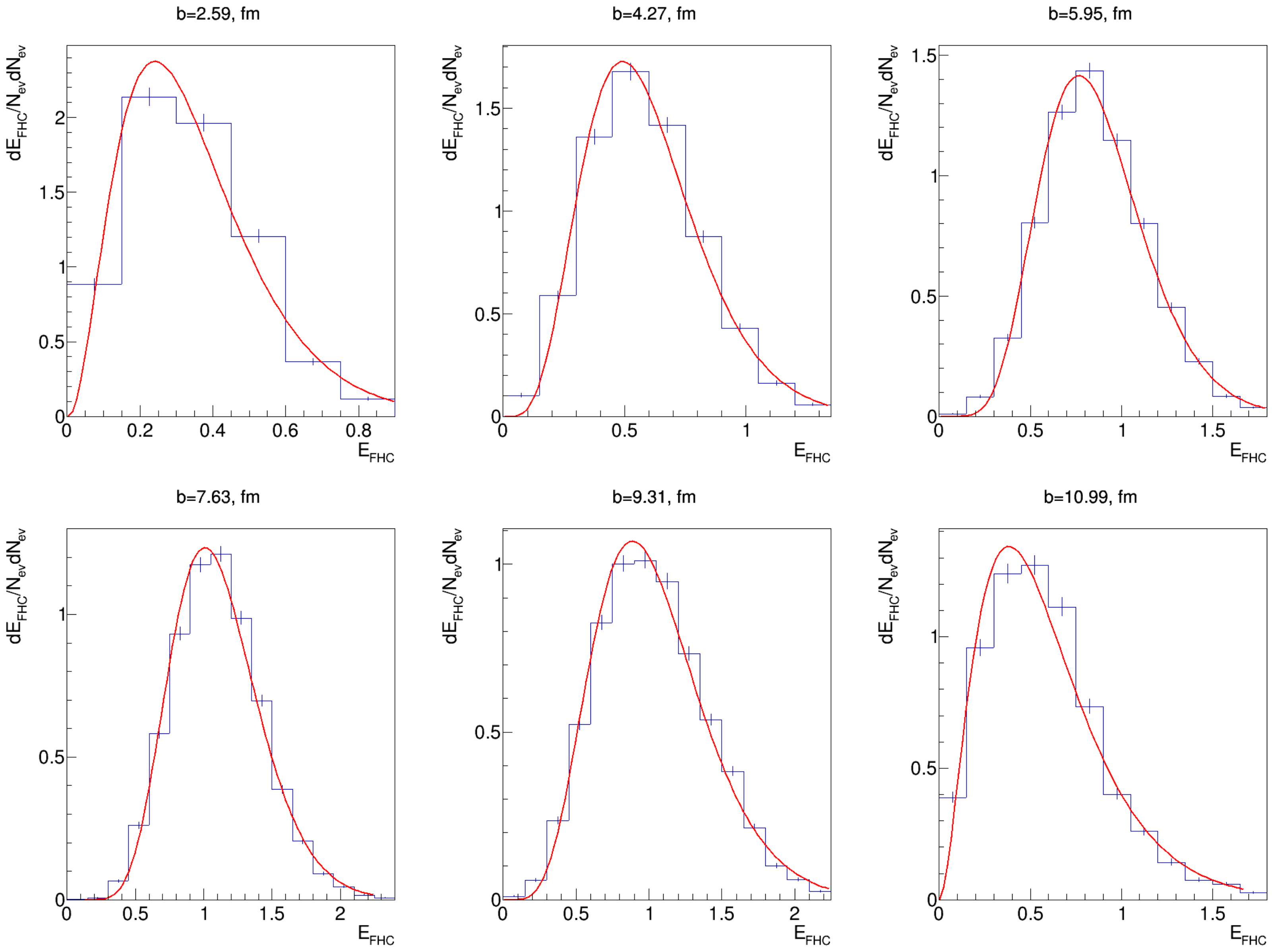
There is agreement within 5%.

Summary and outlook

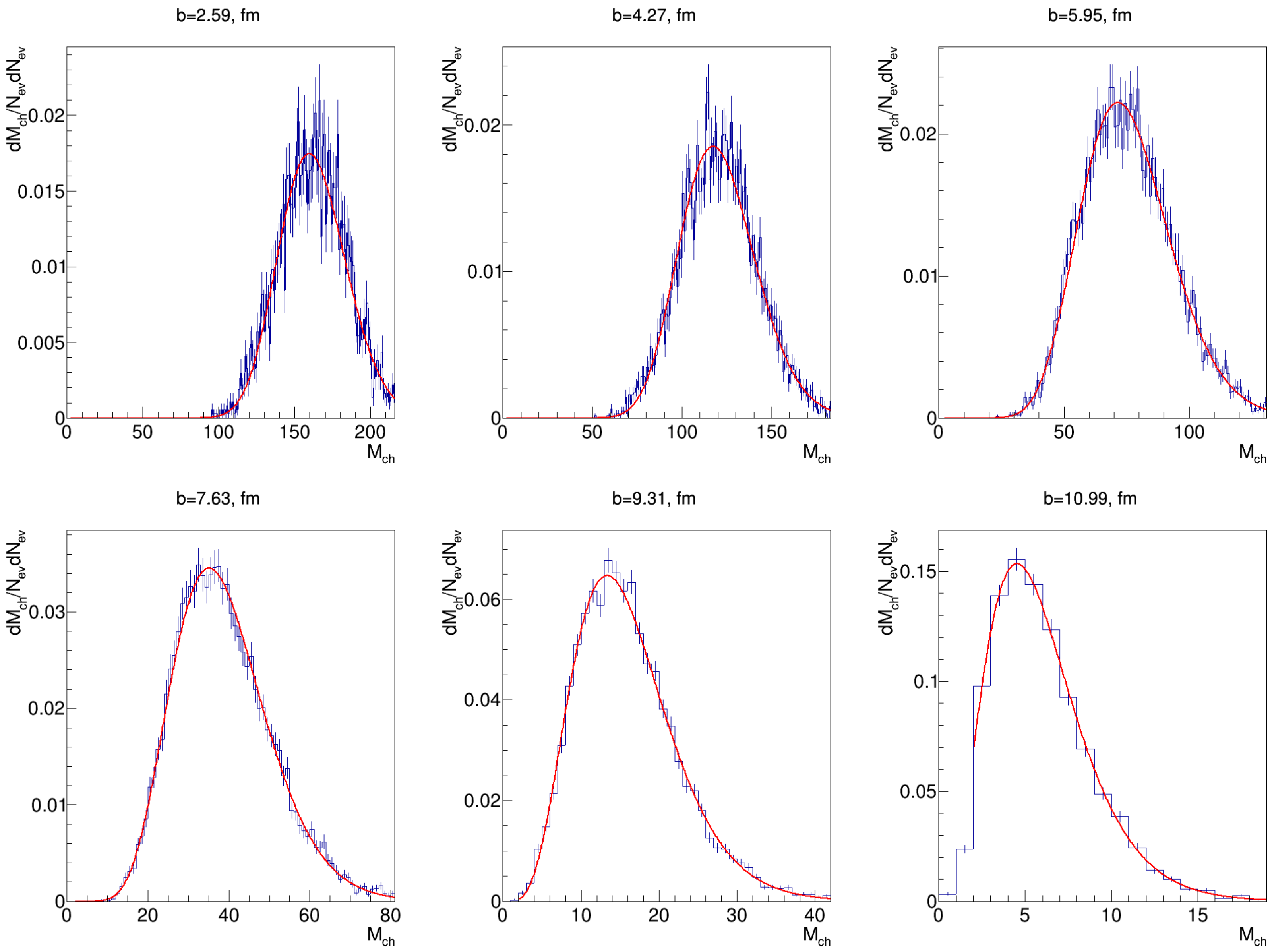
- A new approach to accounting for efficiency and pileup is considered
- The Bayesian inversion method reproduce charged particle multiplicity for fixed-target experiment at BM@N
- A new approach for centrality determination with energy of spectators and multiplicity of charged particles was proposed
- The proposed method was applied to the data from BM@N experiment
- It is planned to create a two-dimensional method based on a signal from a hodoscope and energy from the FHCal

Thank you for your attention!

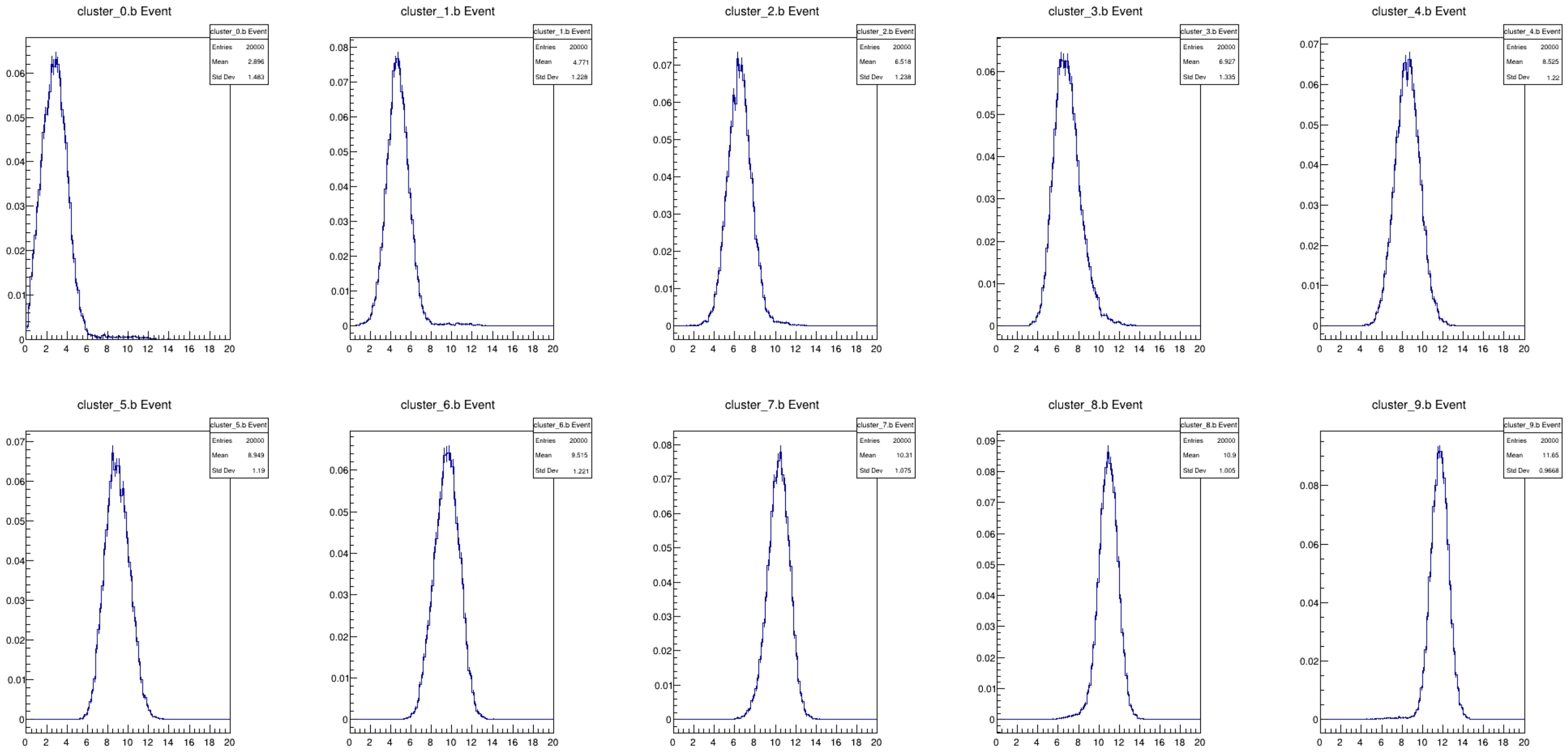
Energy distr. fit



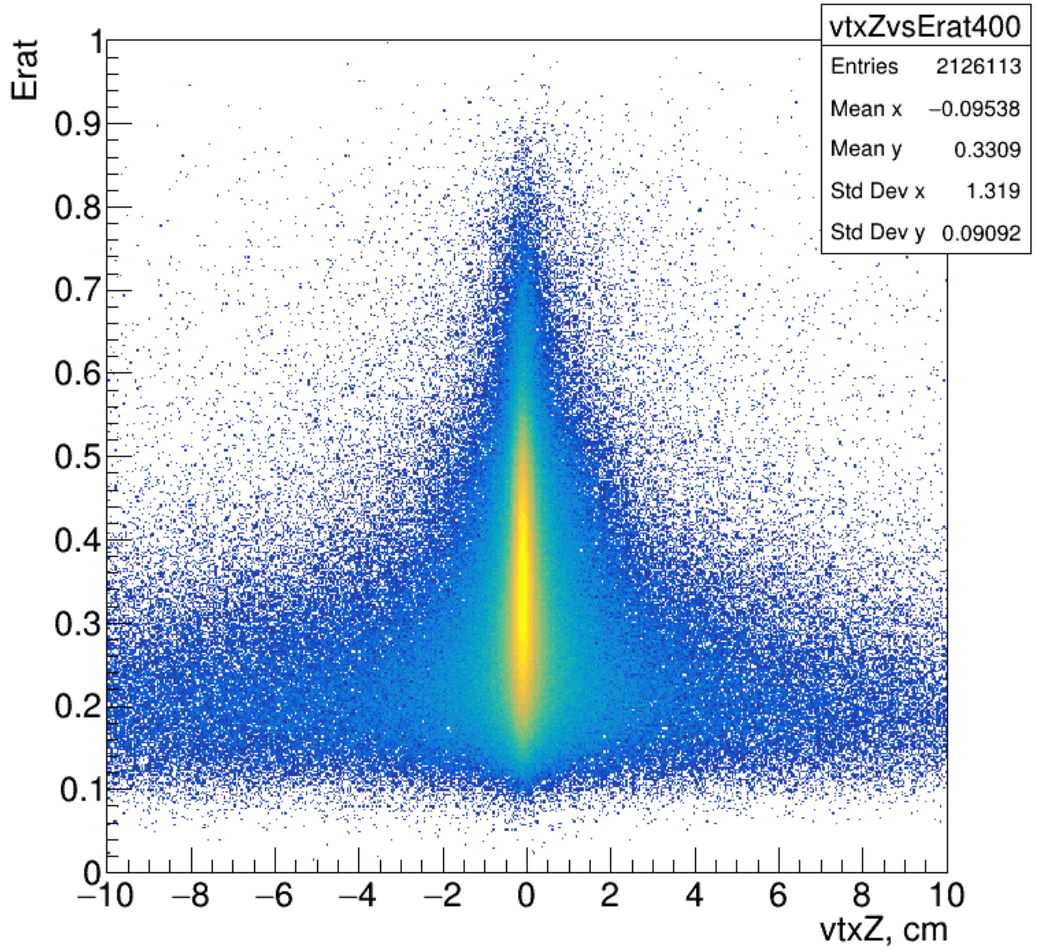
Mult distr. fit

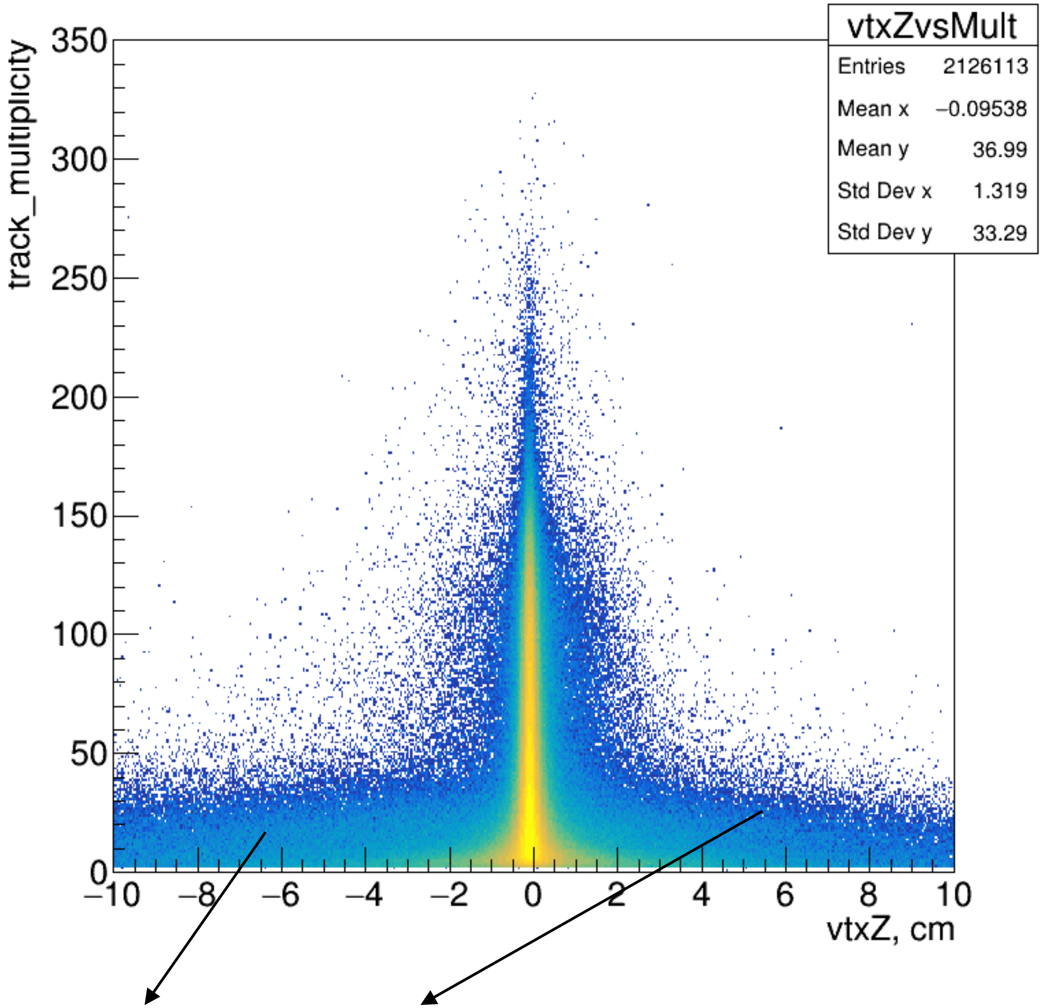


Impact parameter distribution for centrality classes



Event cleaning

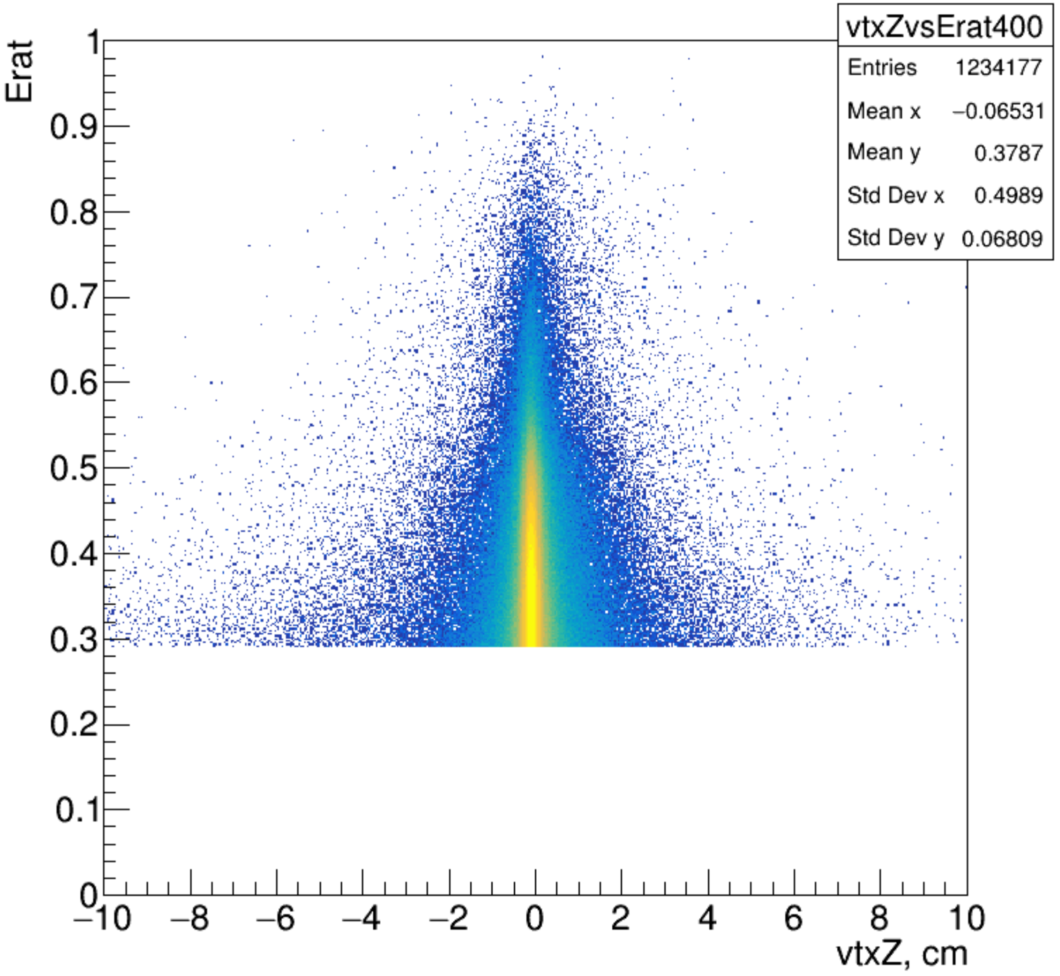


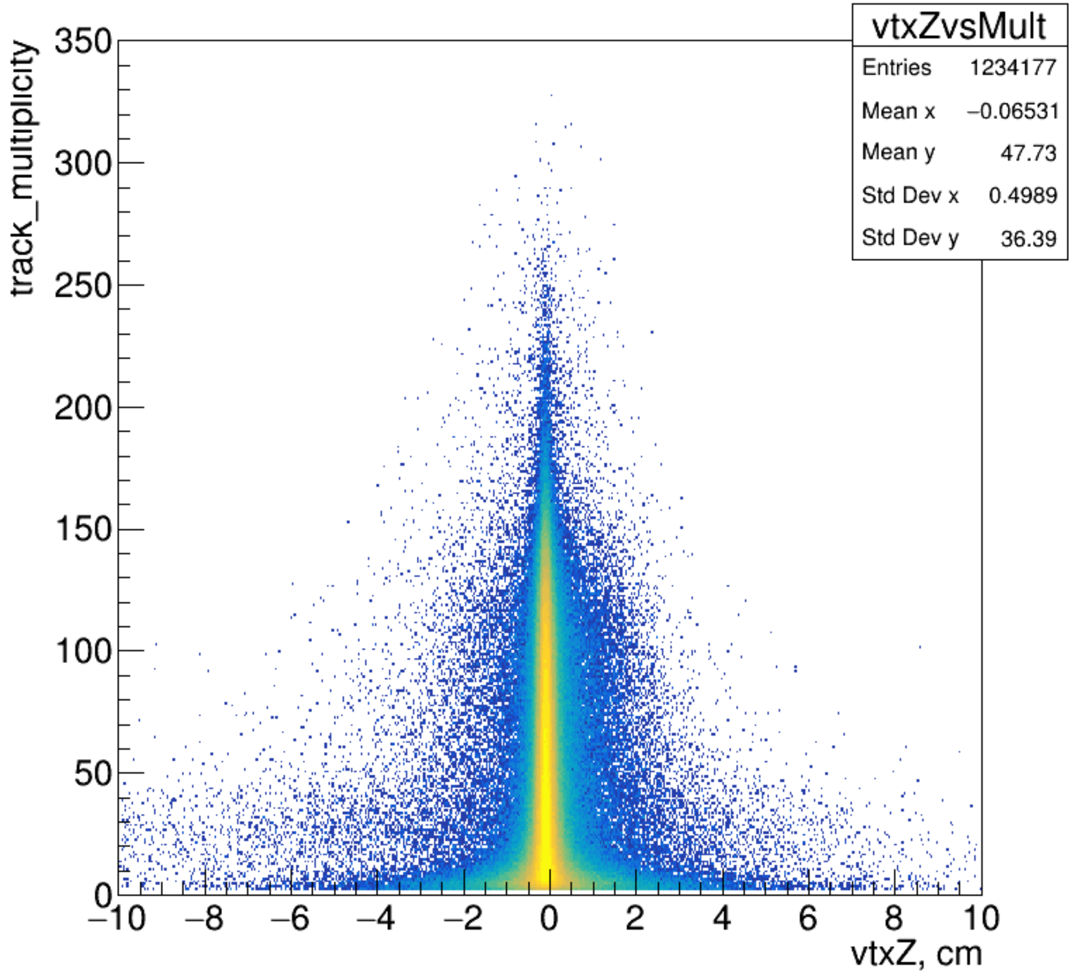


$Erat = \sum E_t / \sum E_l$ – ratio of transverse energy to longitudinal

background due to the interaction with a pipe or kapton

Event cleaning



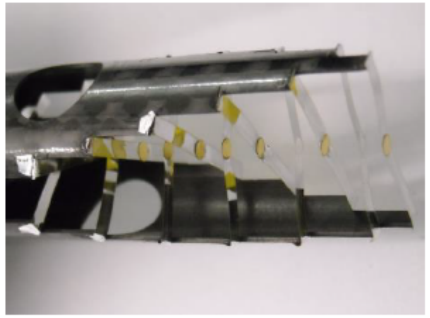


The most of the background has been suppressed after cuts for Erat >0.29 and vertex position $(V_x-0.3)^2+(V_y-0.14)^2<1$ cm

Event cleaning in HADES

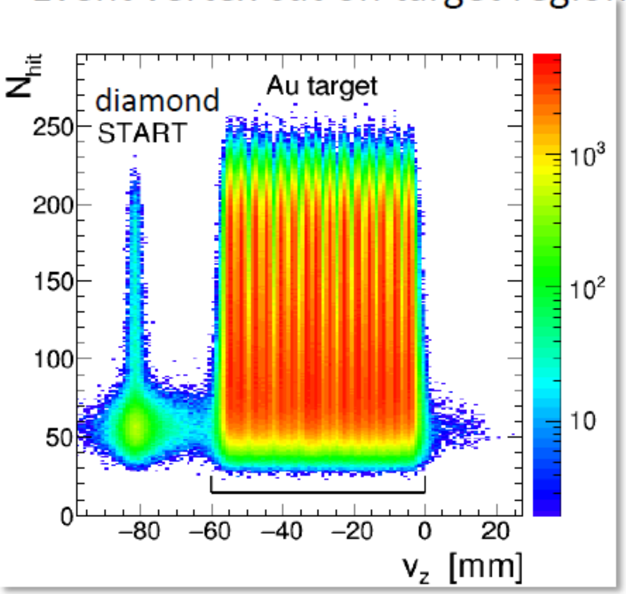
Segmented gold target:

- ^{197}Au material
- 15 discs of Ø = 2.2 mm mounted on kapton strips
- Δz = 3.6 mm
- 2.0% interaction prob.

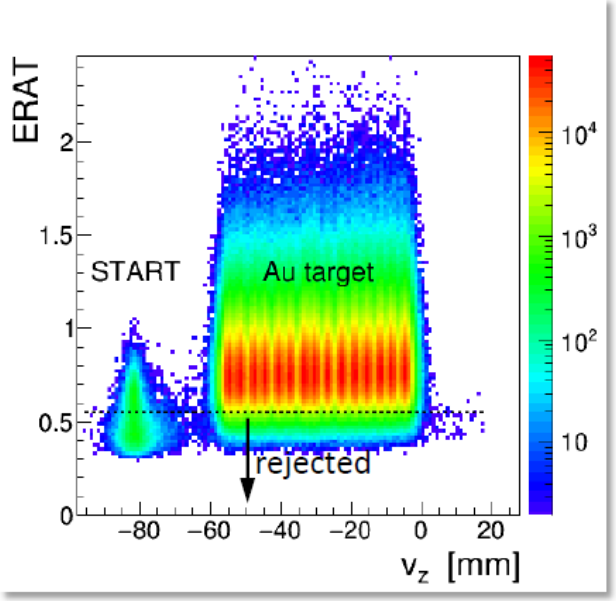


Kindler et al.,
NIM A 655 (2011) 95

Event vertex cut on target region



Remove Au+C bkgd on the kapton with a cut on $ERAT = \sum E_t / \sum E_l$



beam direction

http://indico.oris.mephi.ru/event/221/session/1/contribution/1/material/slides/0.pdf

Reconstruction of *b*

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

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

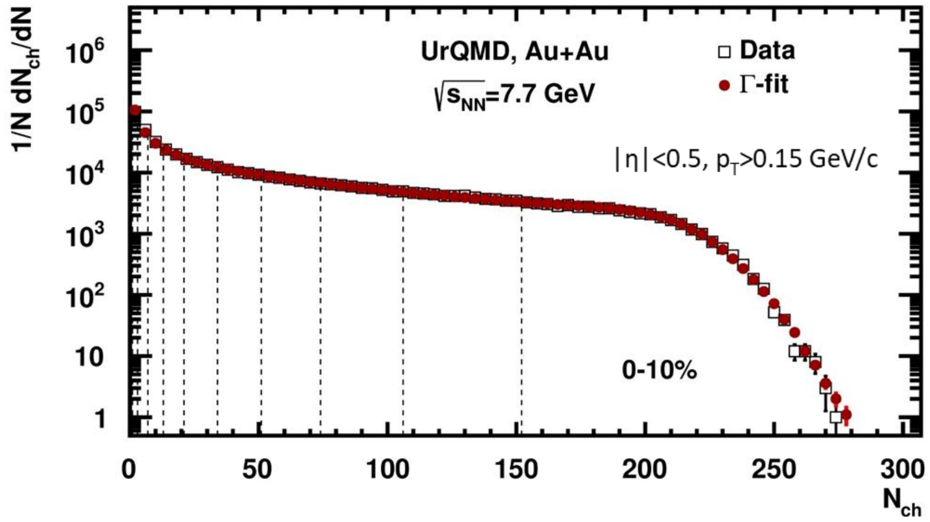
- Find probability of *b* for fixed range of N_{ch} using Bayes' theorem:

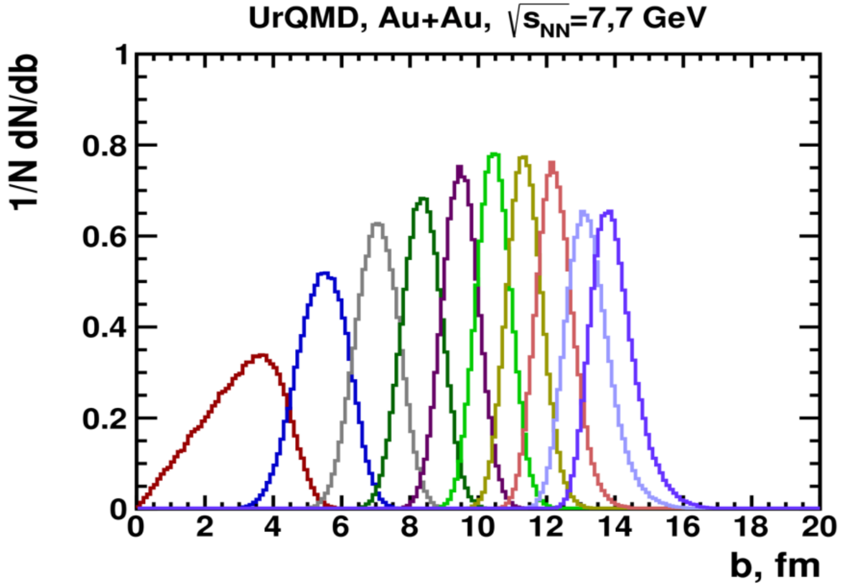
$$P(b|n_1 < N_{ch} < n_2) = P(b)\frac{\int_{n_1}^{n_2} P\,(N_{ch}|b)dN_{ch}}{\int_{n_1}^{n_2} P\,(N_{ch})dN_{ch}}$$

- **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

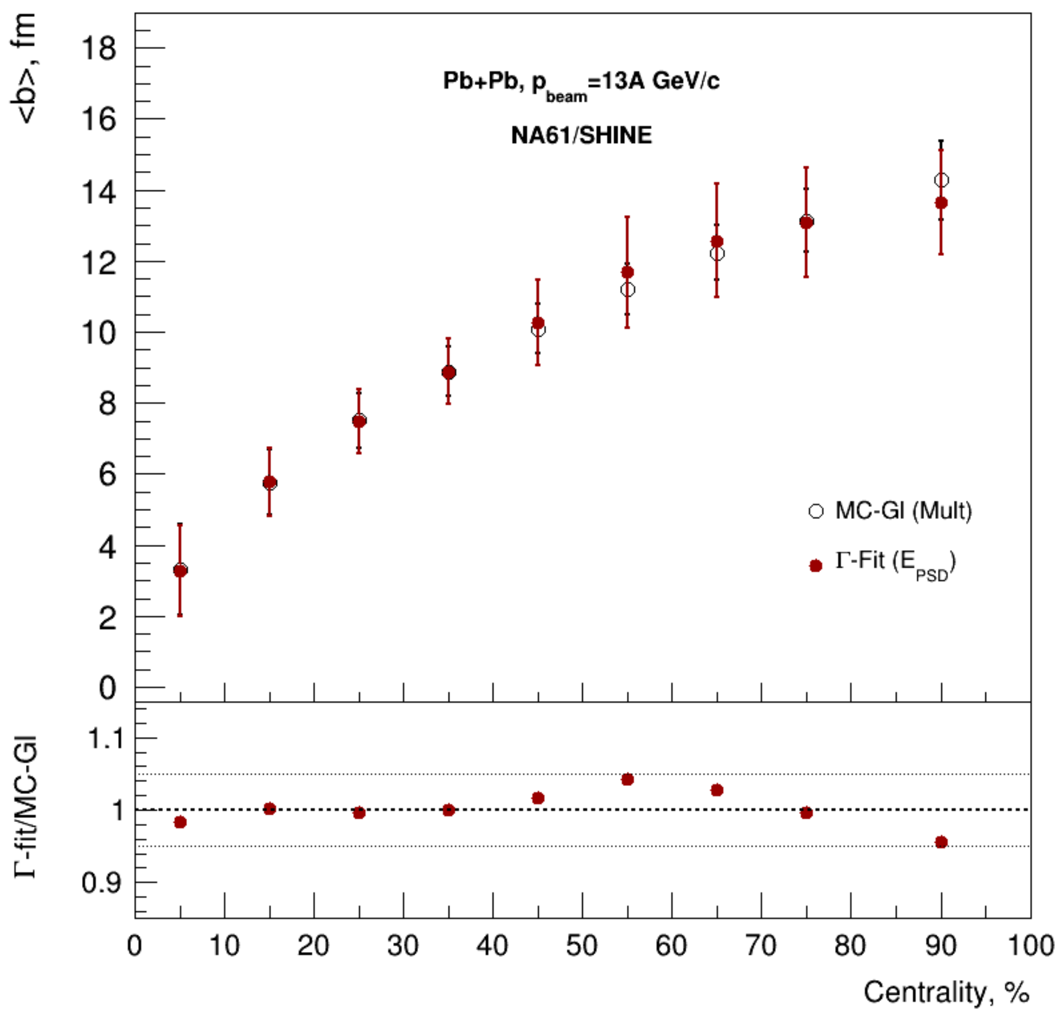




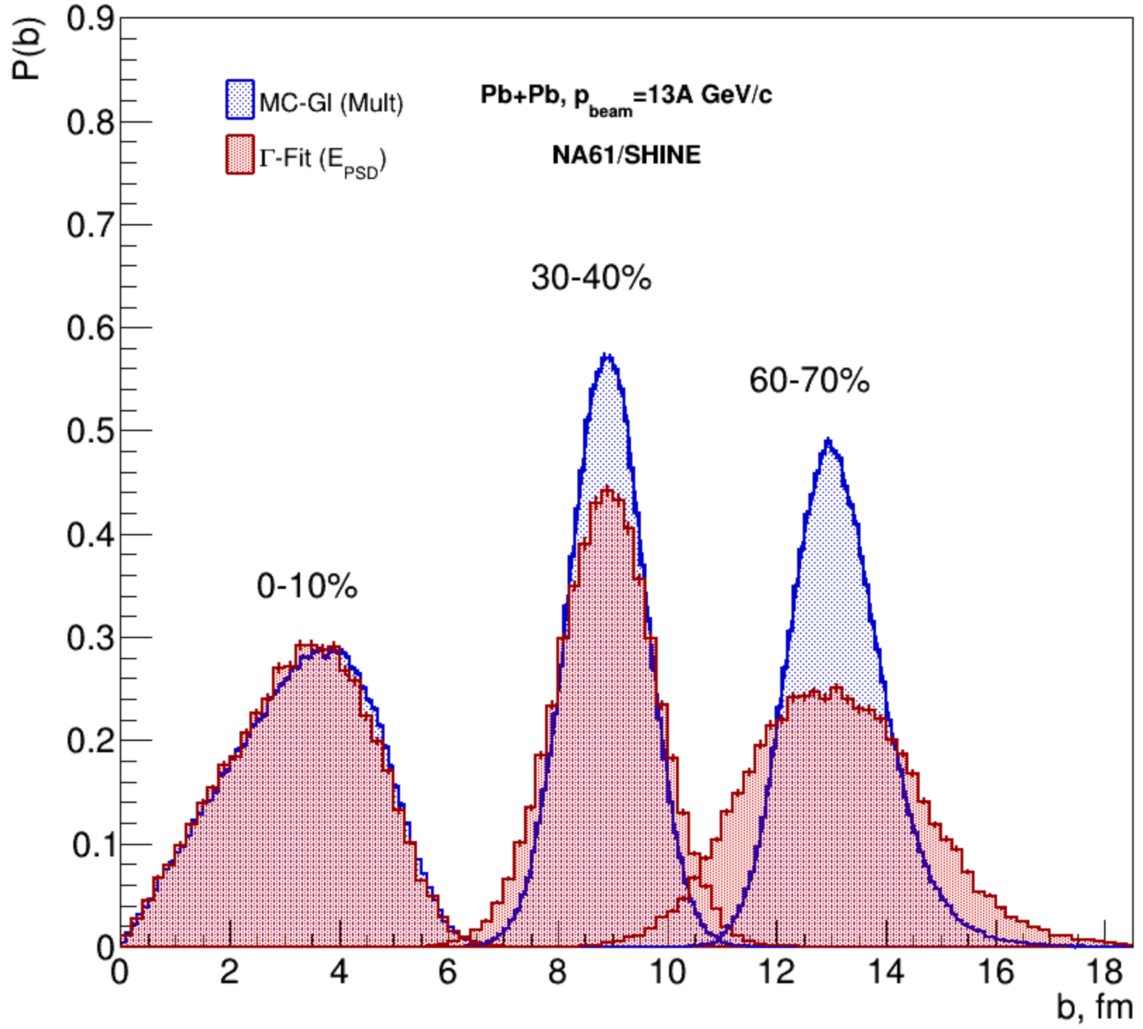
R. Rogly, G. Giacalone and J. Y. Ollitrault, Phys.Rev. C98 (2018) no.2, 024902

Implementation in MPD: https://github.com/Dim23/GammaFit

Comparison with MC-Glauber fit(пример)



Good agreement between fit and data.



There is agreement within 5%.