Efficiency

Adam Bzdak

AGH University of Science and Technology, Krakow

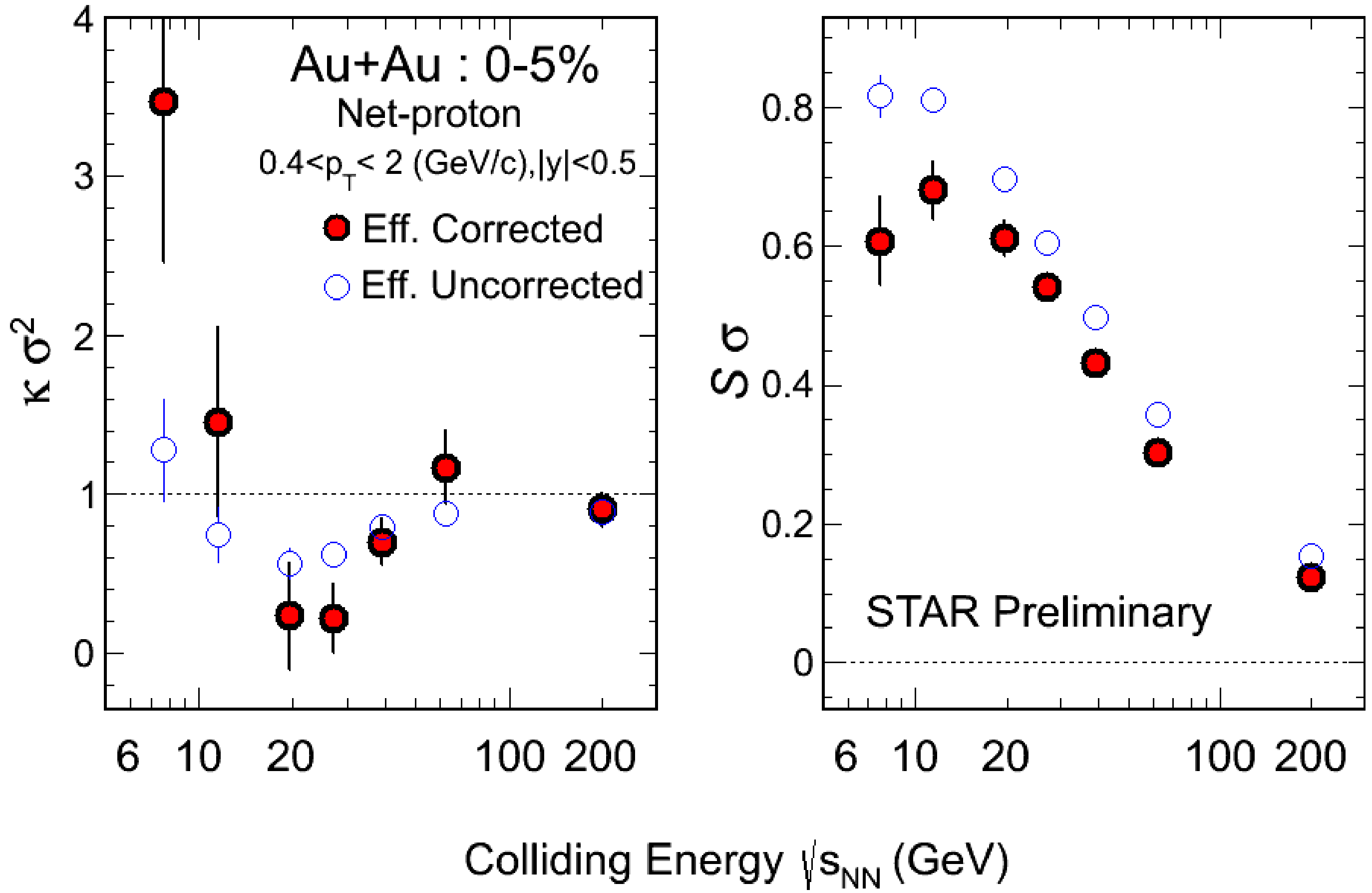
AB, V. Koch,
PRC 86 (2012) 044904
PRC 91 (2015) 027901
in progress

Outline

- Does efficiency matter? **Yes**
- Is correcting with multiplicity independent efficiency good enough? **No**
- Is there any hope? Yes
- So what to do?
- Backup

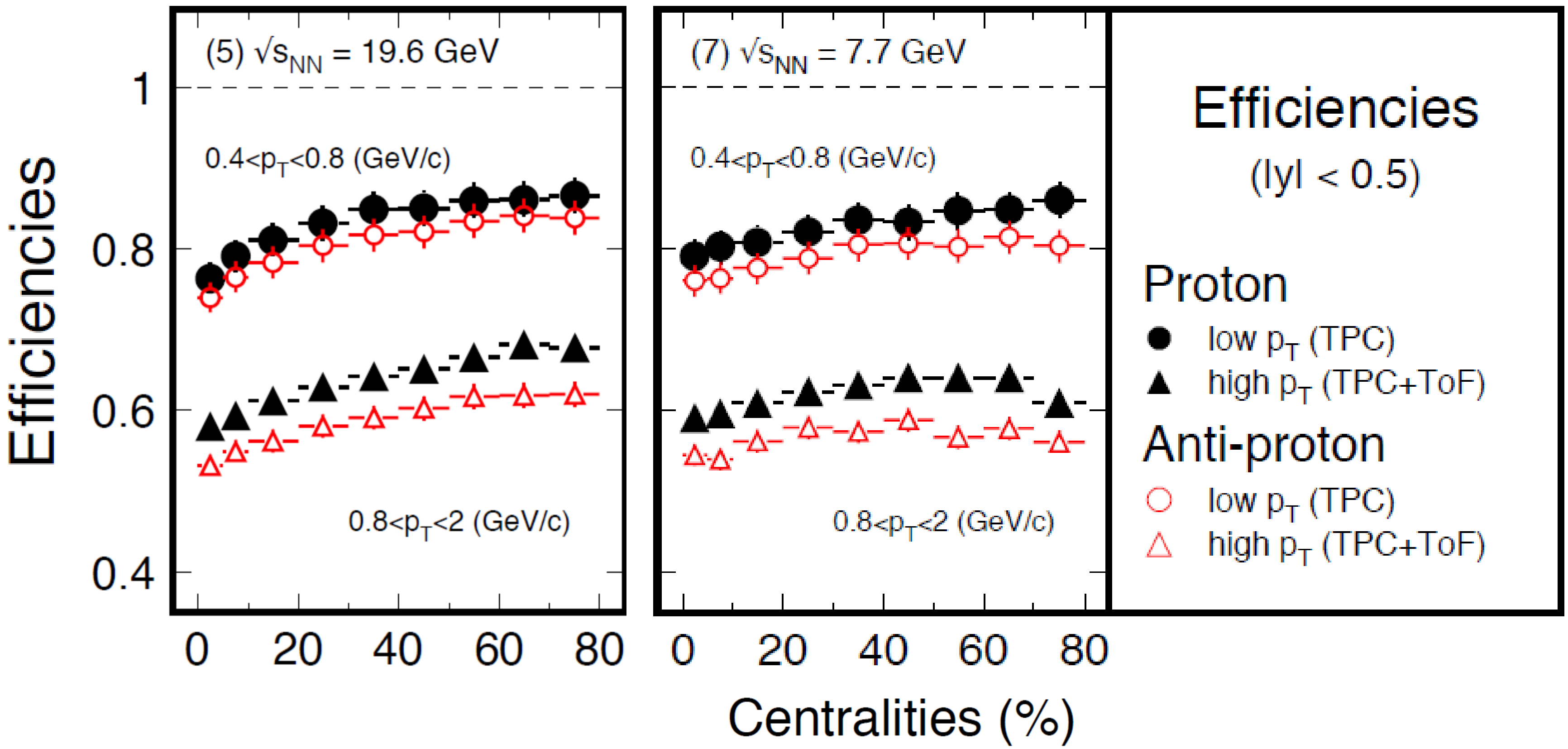
Efficiency correction is important

STAR (thanks to X. Luo)



Efficiencies at 19.6 GeV and 7.7 GeV

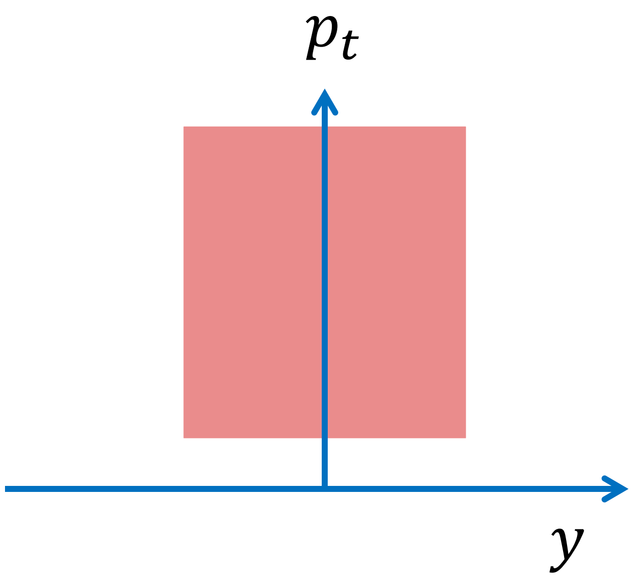
X. Luo [STAR Collaboration] arXiv:1503.02558 [nucl-ex]].



So how can we correct for efficiency?

Efficiencies depend on the number of charged particles.

Definitions



So we decide to measure in some p_t and y range. We produce a certain number of protons but measure only a fraction of them because of imperfect detector.

Moreover we do not see neutrons.

K_n – true cumulants (measured if detector is perfect)

c_n – cumulants that we measure

Calculation

$$p(n_1, n_2) = \sum P(N_1, N_2)\, B(n_1, N_1; p_1) B(n_2, N_2; p_2)$$

$$\downarrow \qquad\qquad\qquad \downarrow$$

c_n $\qquad$ $K_n, F_{i,k}$

what we measure $\qquad$ what we would like to measure

$$\boxed{p_1 = p_2 = 1\colon\ c_n = K_n}$$

factorial moments $F_{i,k} = \left\langle \frac{N_1!N_2!}{(N_1-i)!(N_2-k)!} \right\rangle$

$B(\dots) -$ binomial dist.

It turns out one cannot relate cumulants K_n solely through cumulants c_m

We use factorial moments

$$\left\langle \frac{N_1!\,N_2!}{(N_1-i)!\,(N_2-k)!} \right\rangle = \frac{1}{p_1^i p_2^k} \left\langle \frac{n_1!\,n_2!}{(n_1-i)!\,(n_2-k)!} \right\rangle$$

$$F_{i,k} = \frac{1}{p_1^i p_2^k} f_{i,k}$$

So we express true cumulants through factorial moments $F_{i,k}$, which are known from the above equality ($f_{i,k}$ is measured).

Cumulants vs. factorial moments

$$
\begin{aligned}
K_1 &= \langle N_1 \rangle - \langle N_2 \rangle\,,\\
K_2 &= N - K_1^2 + F_{02} - 2F_{11} + F_{20},\\
K_3 &= K_1 + 2K_1^3 - F_{03} - 3F_{02} + 3F_{12} + 3F_{20} - 3F_{21} + F_{30} - 3K_1(N + F_{02} - 2F_{11} + F_{20}),\\
K_4 &= N - 6K_1^4 + F_{04} + 6F_{03} + 7F_{02} - 2F_{11} - 6F_{12} - 4F_{13} + 7F_{20} - 6F_{21} + 6F_{22} + 6F_{30} - 4F_{31} + F_{40}\\
&\quad + 12K_1^2(N + F_{02} - 2F_{11} + F_{20}) - 3(N + F_{02} - 2F_{11} + F_{20})^2 - 4K_1(K_1 - F_{03} - 3F_{02} + 3F_{12} + 3F_{20} - 3F_{21} + F_{30})
\end{aligned}
$$

$$
\begin{aligned}
K_5 &= K_1 + 24K_1^5 - F_{05} - 10F_{04} - 25F_{03} - 15F_{02} + 15F_{12} + 20F_{13} + 5F_{14} + 15F_{20} - 15F_{21} - 10F_{23} + 25F_{30}\\
&\quad - 20F_{31} + 10F_{32} + 10F_{40} - 5F_{41} + F_{50} - 60K_1^3(N + F_{02} - 2F_{11} + F_{20}) + 30K_1(N + F_{02} - 2F_{11} + F_{20})^2\\
&\quad + 20K_1^2(K_1 - F_{03} - 3F_{02} + 3F_{12} + 3F_{20} - 3F_{21} + F_{30}) - 10(N + F_{02} - 2F_{11} + F_{20})(K_1 - F_{03}\\
&\quad - 3F_{02} + 3F_{12} + 3F_{20} - 3F_{21} + F_{30}) - 5K_1(N + F_{04} + 6F_{03} + 7F_{02} - 2F_{11} - 6F_{12} - 4F_{13} + 7F_{20} - 6F_{21}\\
&\quad + 6F_{22} + 6F_{30} - 4F_{31} + F_{40})
\end{aligned}
$$

$$
\begin{aligned}
K_6 &= N - 120K_1^6 + F_{06} + 15F_{05} + 65F_{04} + 90F_{03} + 31F_{02} - 2F_{11} - 30F_{12} - 80F_{13} - 45F_{14} - 6F_{15} + 31F_{20} - 30F_{21}\\
&\quad + 30F_{22} + 30F_{23} + 15F_{24} + 90F_{30} - 80F_{31} + 30F_{32} - 20F_{33} + 65F_{40} - 45F_{41} + 15F_{42} + 15F_{50} - 6F_{51} + F_{60}\\
&\quad + 360K_1^4(N + F_{02} - 2F_{11} + F_{20}) - 270K_1^2(N + F_{02} - 2F_{11} + F_{20})^2 + 30(N + F_{02} - 2F_{11} + F_{20})^3 - 120K_1^3(K_1\\
&\quad - F_{03} - 3F_{02} + 3F_{12} + 3F_{20} - 3F_{21} + F_{30}) + 120K_1(N + F_{02} - 2F_{11} + F_{20})(K_1 - F_{03} - 3F_{02} + 3F_{12}\\
&\quad + 3F_{20} - 3F_{21} + F_{30}) - 10(K_1 - F_{03} - 3F_{02} + 3F_{12} + 3F_{20} - 3F_{21} + F_{30})^2 + 30K_1^2(N + F_{04} + 6F_{03} + 7F_{02} - 2F_{11}\\
&\quad - 6F_{12} - 4F_{13} + 7F_{20} - 6F_{21} + 6F_{22} + 6F_{30} - 4F_{31} + F_{40}) - 15(N + F_{02} - 2F_{11} + F_{20})(N + F_{04} + 6F_{03}\\
&\quad + 7F_{02} - 2F_{11} - 6F_{12} - 4F_{13} + 7F_{20} - 6F_{21} + 6F_{22} + 6F_{30} - 4F_{31} + F_{40}) - 6K_1(K_1 - F_{05} - 10F_{04} - 25F_{03} - 15F_{02}\\
&\quad + 15F_{12} + 20F_{13} + 5F_{14} + 15F_{20} - 15F_{21} - 10F_{23} + 25F_{30} - 20F_{31} + 10F_{32} + 10F_{40} - 5F_{41} + F_{50}).
\end{aligned}
$$

$$N = \langle N_1 \rangle + \langle N_2 \rangle \qquad K_1 = \langle N_1 \rangle - \langle N_2 \rangle \qquad F_{ik} = \frac{1}{p_1^i p_2^k} f_{ik}$$

Problem is not solved:

We assume that p does not depend on N. **Not good**

We assume binomial distribution. Is it OK? We should test it.
HADES and STAR promised to check it.

By binomial I mean binomial form

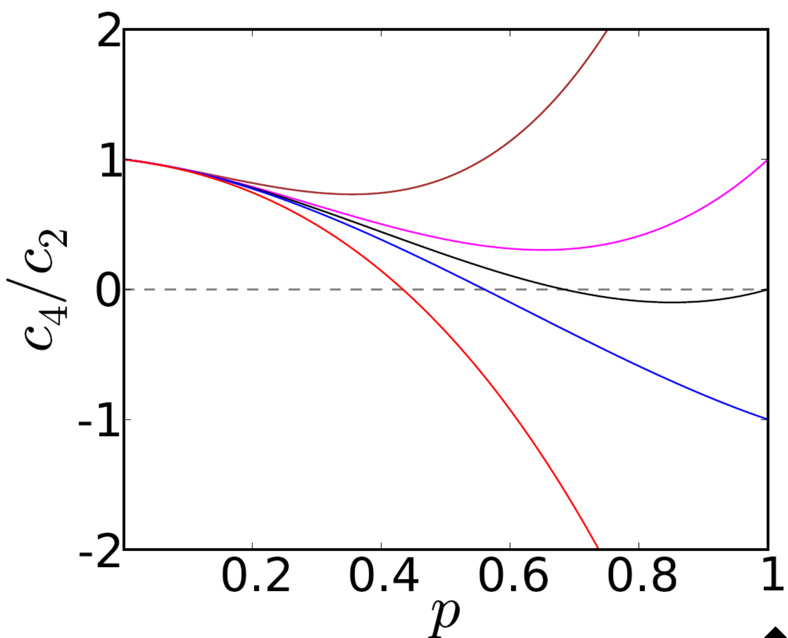
$$B(n, N) = \frac{N!}{n!\,(N-n)!}[p(N, \ldots)]^n[1 - p(N, \ldots)]^{N-n}$$

Net-proton vs net-baryon problem is mathematically equivalent to efficiency problem

M.Kitazawa, M.Asakawa,
PRC 86, 024904 (2012); 069902 (2012)

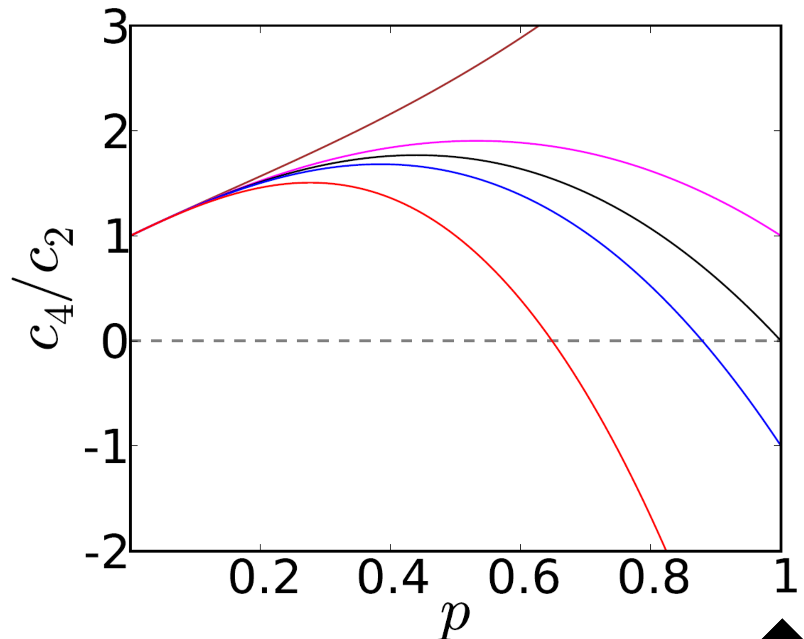
Illustration

multiplicity distr. narrower than Poisson



K_4/K_2

multiplicity distr. broader than Poisson

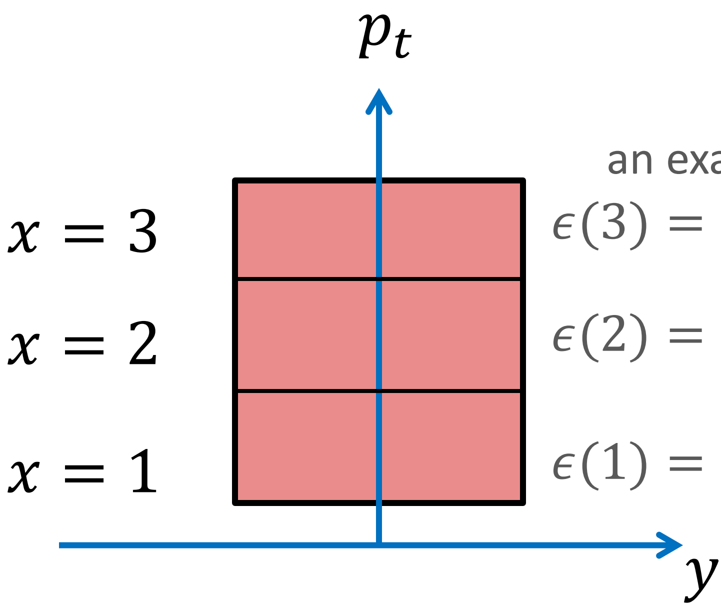


K_4/K_2

$$\frac{K_4}{K_2} = 5, 1, 0, -1, -5$$

Net-proton could be a good approximation of net-baryon if $K_4/K_2 \sim 1$

Suppose our machine detects particles with probabilities that depend on p_t (in general p_t, y, ϕ).



an example:

$\epsilon(3) = 0.5, \bar{\epsilon}(3) = 0.4$

$\epsilon(2) = 0.8, \bar{\epsilon}(2) = 0.8$

$\epsilon(1) = 0.9, \bar{\epsilon}(1) = 0.8$

ϵ ($\bar{\epsilon}$) -- probabilities to detect baryons (antibaryons) or positive (negative) charges

We measure $\underbrace{f_{i,k} \text{ and } c_n}_{\text{observed}}$ but we want to know true $\underbrace{F_{i,k} \text{ and } K_n}_{\text{produced}}$

Calculation

$$\langle N \rangle = \sum_{x=1,2,3} \langle N(x) \rangle \qquad \langle N(x) \rangle = \frac{1}{\epsilon(x)} \langle n(x) \rangle$$

$$F_{1,1} = \langle N\bar{N} \rangle = \sum_{x=1,2,3} \sum_{\bar{x}=1,2,3} \langle N(x)\bar{N}(\bar{x}) \rangle$$

$$\langle N(x)\bar{N}(\bar{x}) \rangle = \frac{1}{\epsilon(x)\bar{\epsilon}(\bar{x})} \langle n(x)\bar{n}(\bar{x}) \rangle$$

Once we know $F_{i,k}$ we can construct cumulnats K_n

See PRC 91 (2015) 027901 for general equations

Let's go back to factorial moments and neglect anti-protons

$$\left\langle \frac{N!}{(N-i)!} \right\rangle = \frac{1}{\epsilon^i} \left\langle \frac{n!}{(n-i)!} \right\rangle \qquad F_i = \frac{1}{\epsilon^i} f_i$$

We assume that ϵ does not depend on N. **Not good**

$$f_i = \sum_N \underline{\epsilon^i(N)} P(N) \frac{N!}{(N-i)!}$$

R. Holzmann, talk at HIC for FAIR Workshop on Fluctuation and Correlation Measures in Nuclear Collisions (2015)

If ϵ depends on N the whole method brakes down.

R. Holzmann,
talk at HIC for FAIR

$$f_i = \sum\nolimits_N \underline{\epsilon^i(N)} P(N) \frac{N!}{(N-i)!}$$

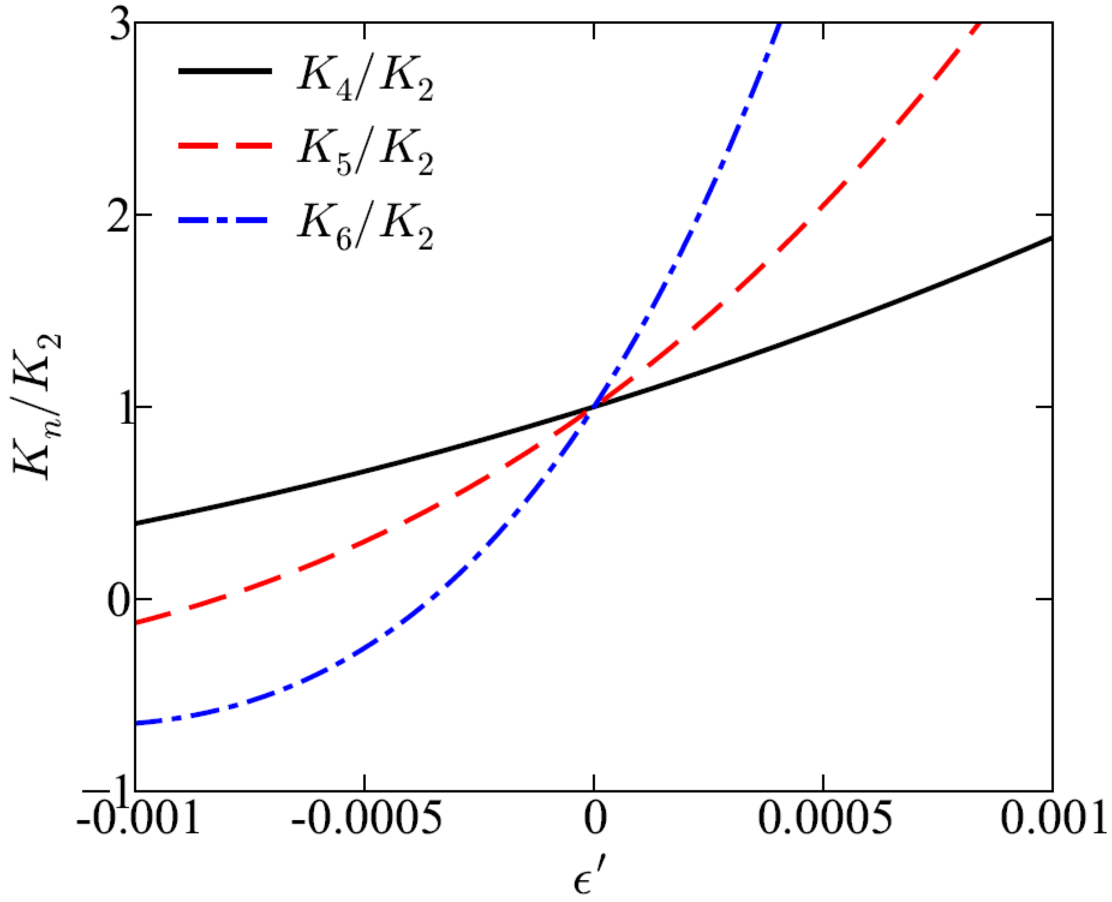
Let's test it. Suppose that

$$P(N) = \frac{\langle N \rangle^N}{N!} e^{-\langle N \rangle},$$

$$\epsilon(N) = \epsilon_0 + \epsilon'(N - \langle N \rangle)$$

with $\langle N \rangle = 40, \epsilon_0 = 0.65$ and plot K_n/K_2 as a function of ϵ'.
We calculate exact f_i and correct using constant efficiency
$F_i = f_i/\epsilon_0^i$.

We obtain



Large corrections for small ϵ'

We need something else.
Why not just solve equations directly, for example:

$$p(n) = \sum_{N=n}^{\infty} P(N) \frac{N!}{n!(N-n)!} \epsilon^n (1-\epsilon)^{N-n}$$

$$\begin{pmatrix} p(0) \\ p(1) \\ p(2) \\ p(3) \\ p(4) \end{pmatrix} = \begin{pmatrix} 1 & 1-\epsilon & (1-\epsilon)^2 & (1-\epsilon)^3 & (1-\epsilon)^4 \\ 0 & \epsilon & 2\epsilon(1-\epsilon) & 3\epsilon(1-\epsilon)^2 & 4\epsilon(1-\epsilon)^3 \\ 0 & 0 & \epsilon^2 & 3\epsilon^2(1-\epsilon) & 6\epsilon^2(1-\epsilon)^2 \\ 0 & 0 & 0 & \epsilon^3 & 4\epsilon^3(1-\epsilon) \\ 0 & 0 & 0 & 0 & \epsilon^4 \end{pmatrix} \begin{pmatrix} P(0) \\ P(1) \\ P(2) \\ P(3) \\ P(4) \end{pmatrix}$$

Triangular equations are trivial to solve.
We can easily use $\epsilon(N)$, matrix is much more complicated but it is not a big deal.

Problem.

For binomial with constant ϵ

$$\det(B) = \prod_{i=0}^{i=M} B(i,i) = \epsilon^{M(M+1)/2}$$

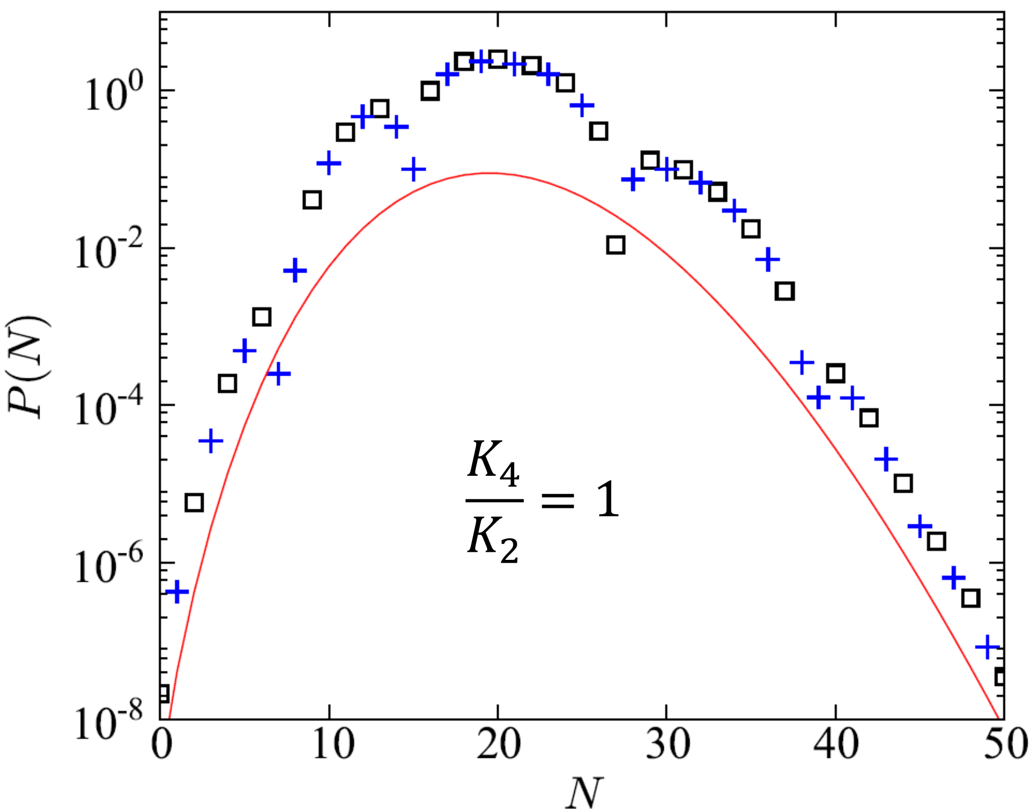
$$M = N_{\max} = n_{\max}$$

For $\epsilon = 0.7$, $M = 100$, $\det(\mathrm{B}) = 10^{-720}$

Matrix is pseudo-singular.

For binomial:

$$P(N) = \sum_{n=0}^{\infty} p(n) \frac{n!}{N!(n-N)!} \frac{1}{\epsilon^n} (-1+\epsilon)^{n-N}$$



We take exact $p(n)$ from slide 16 and replace it by $p(n)[1 + O(10^{-5})]$. Then we calculate above P(N).

We get nonsense but …

it has **correct** cumulants!

red line – input Poisson

black – positive $P(N)$

blue – negative $P(N)$

OK, let's test this method with $\epsilon(N)$. $\epsilon(N) = \epsilon_0 + \epsilon'(N - \langle N \rangle)$

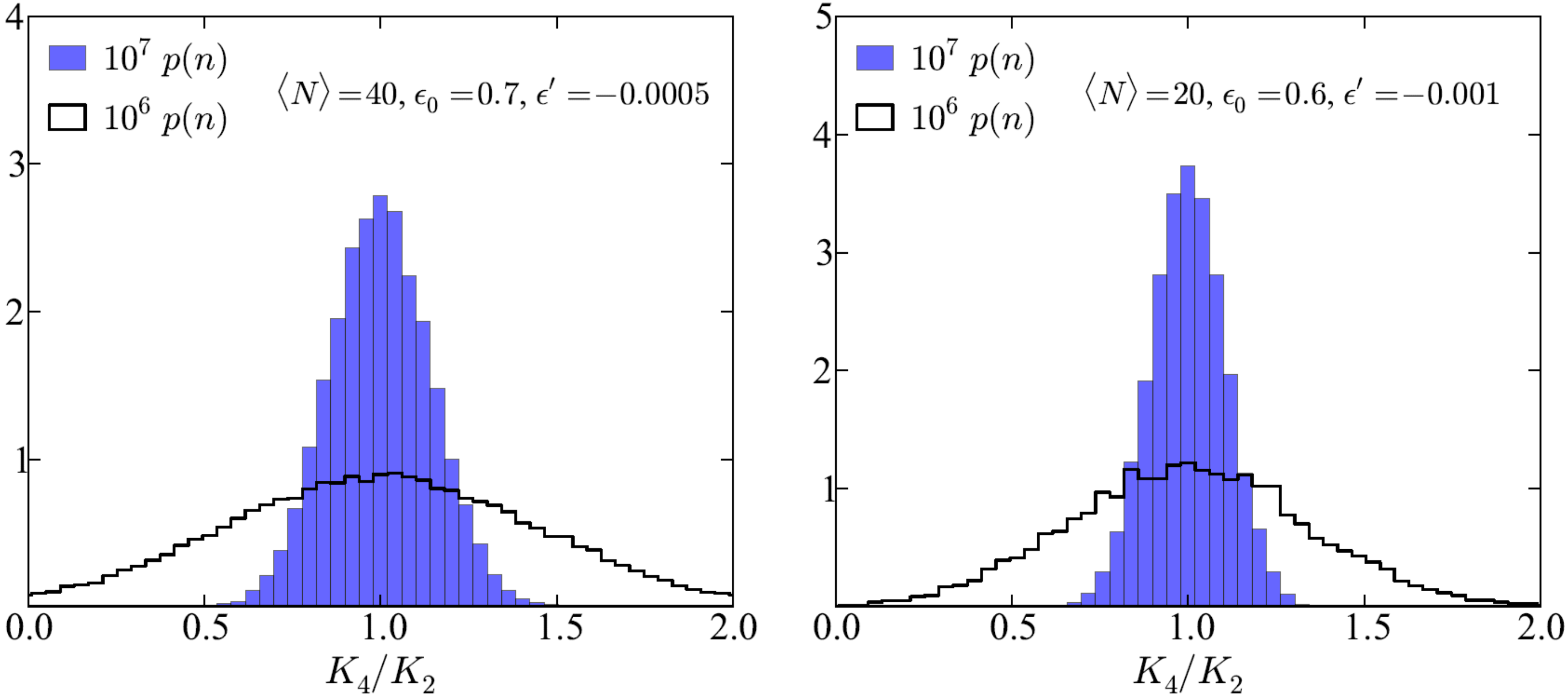
1) we sample N particles from Poisson
2) each particle can be detected with probability $\epsilon(N)$
3) we run 10^7 events and get measured $p(n)$
4) our matrix is given by

$$B(n, N) = \frac{N!}{n!(N-n)!} \epsilon(N)^n \left[1 - \epsilon(N)\right]^{N-n}$$

5) solve triangular equations and obtain $P(N)$. We get nonsense but that's fine
6) we calculate K_4/K_2 and get a number

We repeat the whole exercise many times and plot histogram of obtained K_4/K_2 …

... and this is what we obtain



It works very well, statistical errors are under control

What if we do not know analytical form of $B(n, N)$?

Solution A

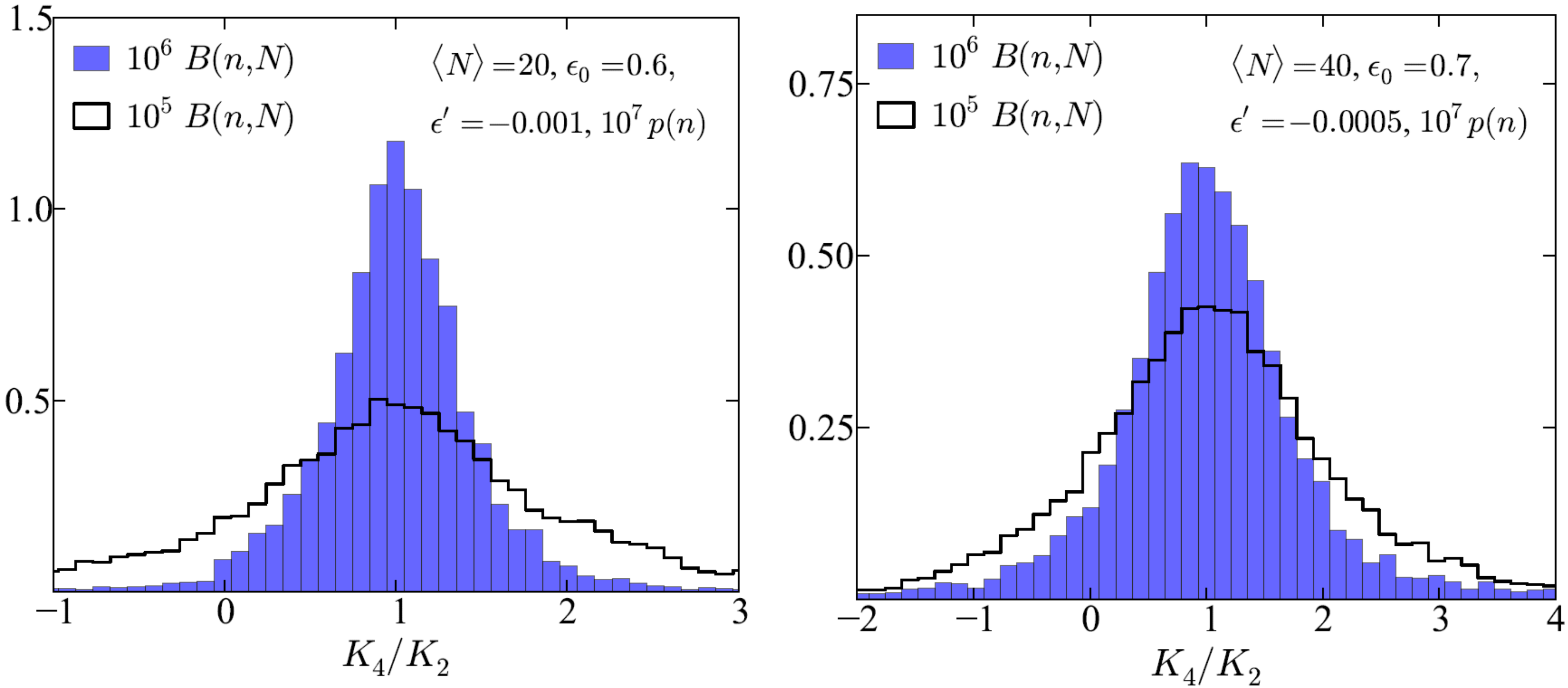
1) generate particles from some reasonable generator
2) run it through a detector simulator
3) we get the matrix $B(n, N)$ but it is not complete
4) each column we fit with some function
5) now we get complete matrix
6) and we go back to the previous case. Done.

What if we do not know analytical form of $B(n, N)$?

Solution B

1) generate particles from some reasonable generator
2) run it through a detector simulator
3) we get the matrix $B(n, N)$ but it is not complete
4) $\det(B) = 0$ exactly
5) calculate the Moore-Penrose pseudoinverse
6) calculate $P(N)$ [which is wrong] and cumulants

The Moore-Penrose pseudoinverse



I am not sure if this method is stable. I would use only if we fail to find something better

Conclusions

Efficiency is very serious problem that makes any interpretation of net-proton cumulants challenging

Solving triangular equation could be helpful

Simulating response matrix and fitting some function seems to be a good idea

We could use the Moore-Penrose pseudoinverse but I am not sure this method is stable (under study)

Backup

Cumulants

$$g(t) = \ln\left(\sum_n P_B(n)e^{nt}\right)$$

cumulant generating function

$$g(t) = \sum_{k=1}^{\infty} c_k \frac{t^k}{k!}$$

n-th derivative with respect to t (at $t = 0$) gives c_n

$P_B(n)$ – net baryon/proton/charge distribution

Relations between K_n and c_n.
Here $p_1 = p_2 = p$.

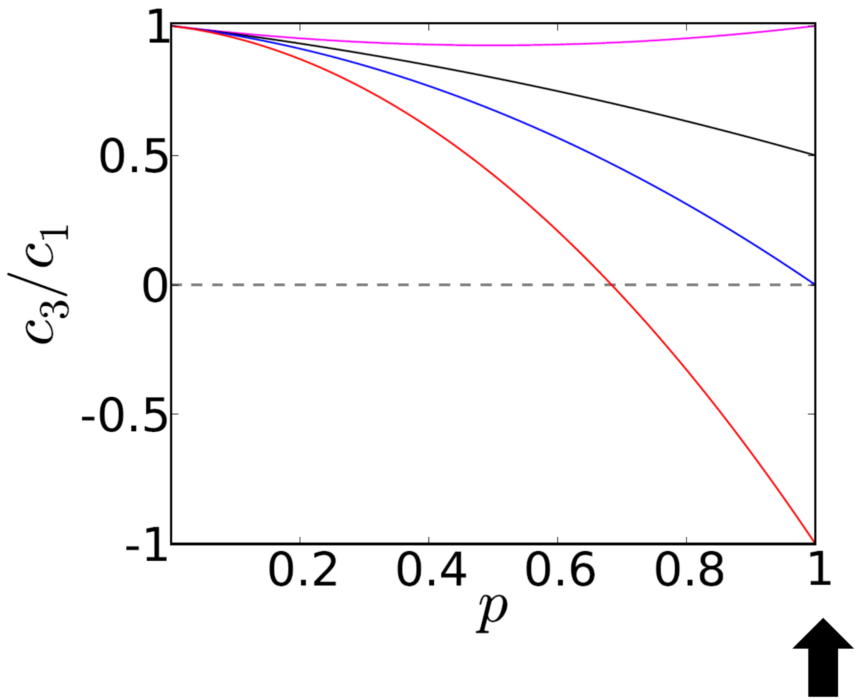
$$
\begin{aligned}
pK_1 &= c_1, \\
p^2 K_2 &= c_2 - n(1-p), \\
p^3 K_3 &= c_3 - c_1(1-p^2) - 3(1-p)(f_{20} - f_{02} - nc_1) \\
p^4 K_4 &= c_4 - np^2(1-p) - 3n^2(1-p)^2 - 6p(1-p) \\
&\quad \times (f_{20} + f_{02}) + 12c_1(1-p)(f_{20} - f_{02}) - (1-p^2) \\
&\quad \times \left(c_2 - 3c_1^2\right) - 6n(1-p)\left(c_1^2 - c_2\right) - 6(1-p) \\
&\quad \times (f_{03} - f_{12} + f_{02} + f_{20} - f_{21} + f_{30}).
\end{aligned}
$$

$f_{i,k}$ – measured factorial moments $\qquad n \equiv \langle n_1 \rangle + \langle n_2 \rangle$

General case $p_1 \neq p_2$, see PRC 86 (2012) 044904

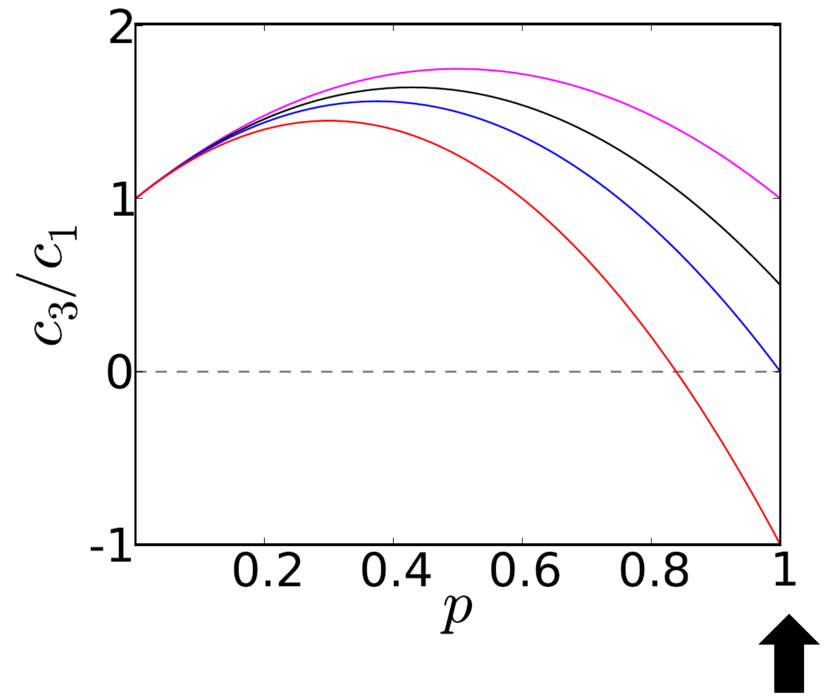
c_3/c_1 as a function of binomial parameter p

multiplicity distr. narrower than Poisson



K_3/K_1

multiplicity distr. broader than Poisson



K_3/K_1

$$\frac{K_3}{K_1} = -1, 0, 1/2, 1$$

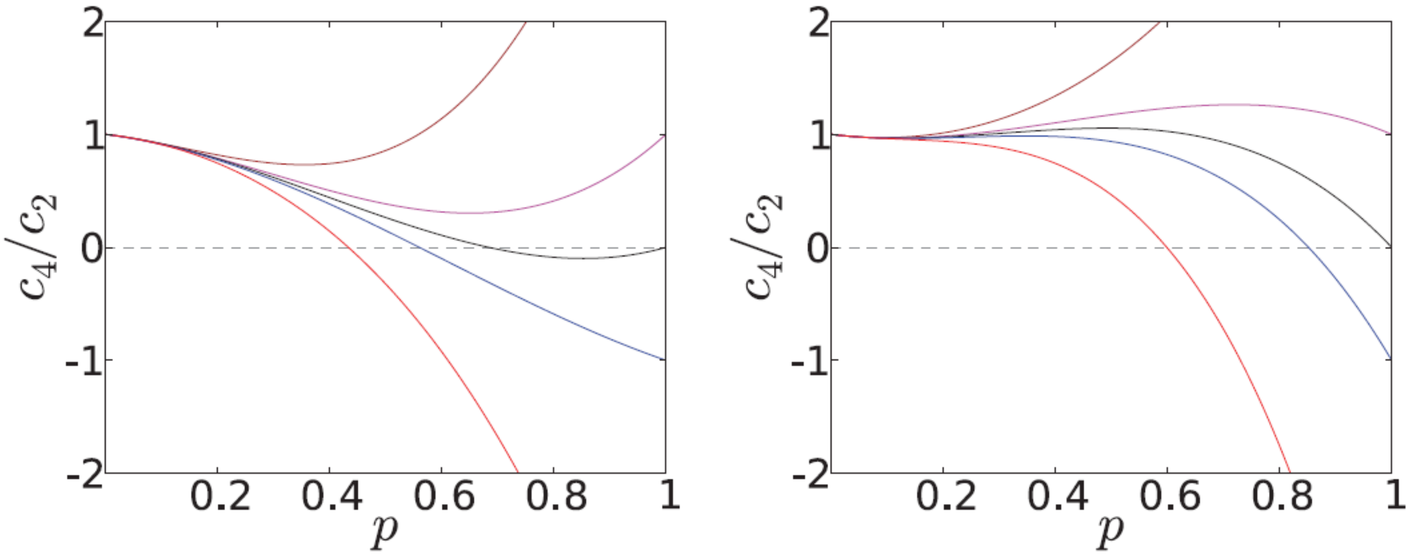


FIG. 3. (Color online) The measured cumulant ratio c_4/c_2 as a function of the acceptance parameter p for the five values $K_4/K_2 = -5, -1, 0, 1$, and 5. K_4/K_2 equals c_4/c_2 at $p = 1$. In both plots $\alpha = -0.1$. In the left plot $\beta = 0.01$, and in the right one $\beta = -0.01$; see Eqs. (23)–(25). For the measurement of the net-proton cumulants at STAR the realistic value of p is smaller than $1/2$ and arguably close to $1/5$.

Local factorial moments

$$F_{2,0} = \langle N(N-1) \rangle = \sum_{x_1=1,2,3} \sum_{x_2=1,2,3} \left\langle N(x_1)[N(x_2) - \delta_{x_1,x_2}] \right\rangle$$

$$\left\langle N(x_1)[N(x_2) - \delta_{x_1,x_2}] \right\rangle = \frac{1}{\epsilon(x_1)\epsilon(x_2)} \left\langle n(x_1)[n(x_2) - \delta_{x_1,x_2}] \right\rangle$$

$\delta_{x_1,x_2} = 1$ if $x_1 = x_2$ and zero otherwise

See backup or PRC 91 (2015) 027901 for general equations

Local efficiency – general expressions

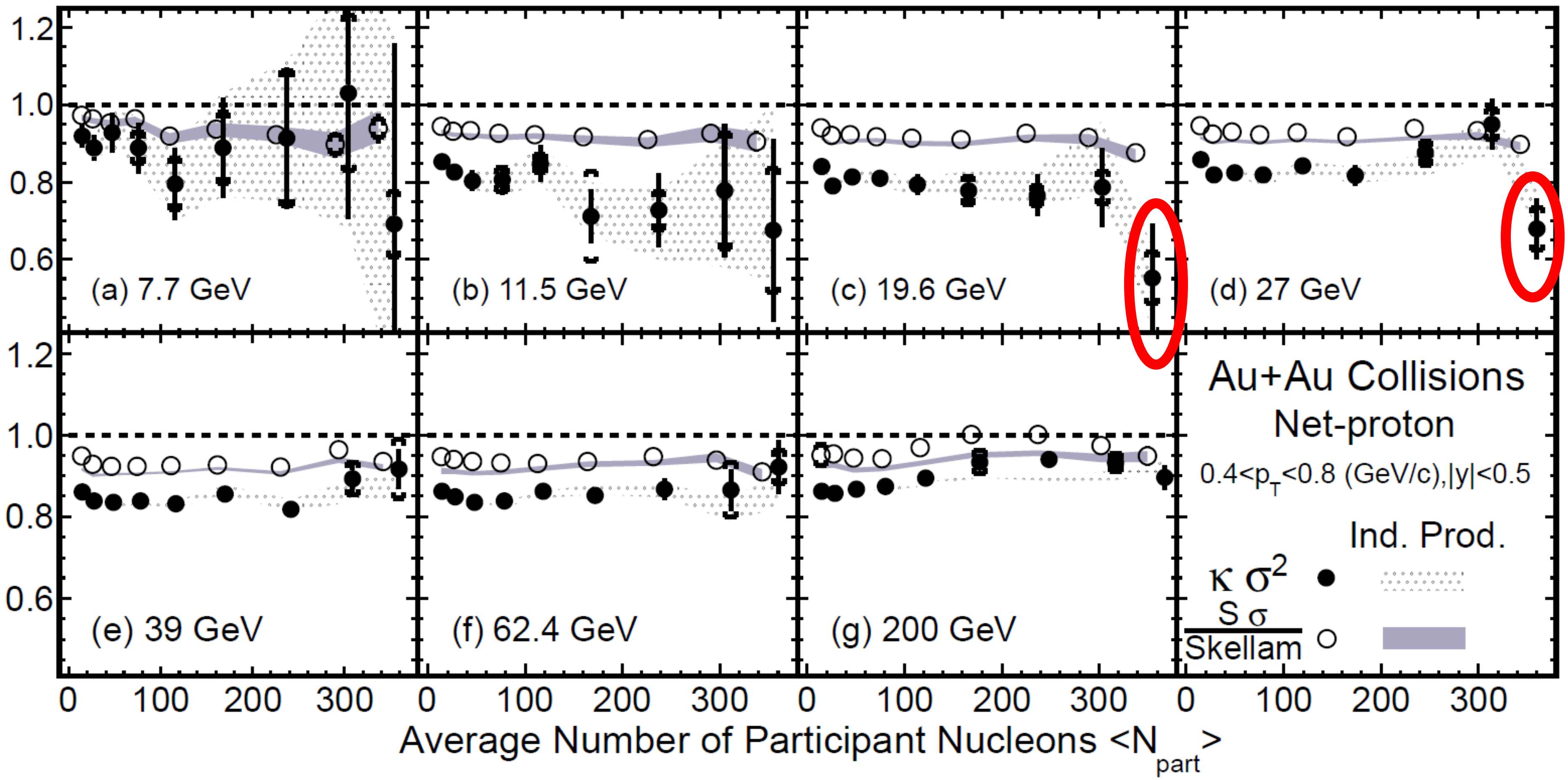
$$A_{i,k}\,(x_1,\ldots,x_i;\bar{x}_1,\ldots,\bar{x}_k) = \langle N(x_1)[N(x_2)-\delta_{x_1,x_2}]\cdots[N(x_i)-\delta_{x_1,x_i}-\cdots-\delta_{x_{i-1},x_i}]$$
$$\bar{N}(\bar{x}_1)[\bar{N}(\bar{x}_2)-\delta_{\bar{x}_1,\bar{x}_2}]\cdots[\bar{N}(\bar{x}_k)-\delta_{\bar{x}_1,\bar{x}_k}-\cdots-\delta_{\bar{x}_{k-1},\bar{x}_k}]\rangle$$

$$a_{i,k}\,(x_1,\ldots,x_i;\bar{x}_1,\ldots,\bar{x}_k) = \langle n(x_1)[n(x_2)-\delta_{x_1,x_2}]\cdots[n(x_i)-\delta_{x_1,x_i}-\cdots-\delta_{x_{i-1},x_i}]$$
$$\bar{n}(\bar{x}_1)[\bar{n}(\bar{x}_2)-\delta_{\bar{x}_1,\bar{x}_2}]\cdots[\bar{n}(\bar{x}_k)-\delta_{\bar{x}_1,\bar{x}_k}-\cdots-\delta_{\bar{x}_{k-1},\bar{x}_k}]\rangle$$

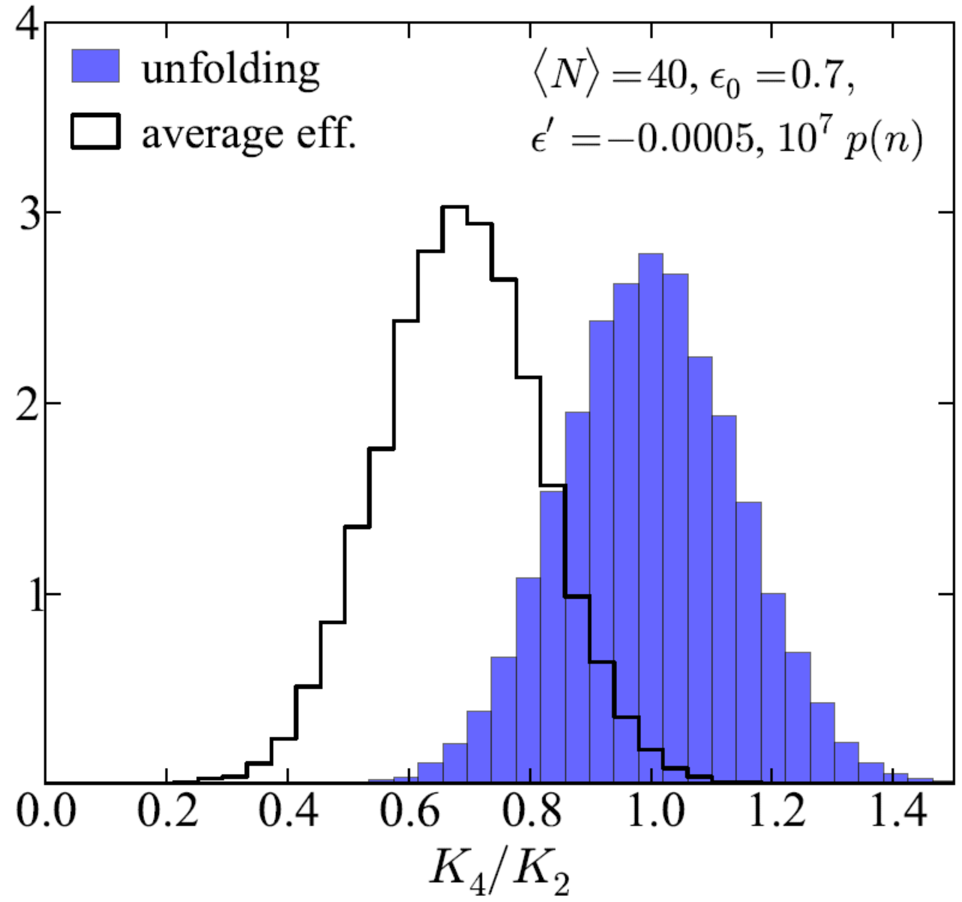
$$a_{i,k} = \epsilon(x_1)\cdots\epsilon(x_i)\bar{\epsilon}(\bar{x}_1)\cdots\bar{\epsilon}(\bar{x}_k)A_{i,k}$$

$$F_{i,k} = \sum_{x_1,\ldots,x_i}\sum_{\bar{x}_1,\ldots,\bar{x}_k}\frac{a_{i,k}\,(x_1,\ldots,x_i;\bar{x}_1,\ldots,\bar{x}_k)}{\epsilon(x_1)\ldots\epsilon(x_i)\bar{\epsilon}(\bar{x}_1)\ldots\bar{\epsilon}(\bar{x}_k)}$$

Peculiar centrality dependence only for 19.6 and 27 GeV

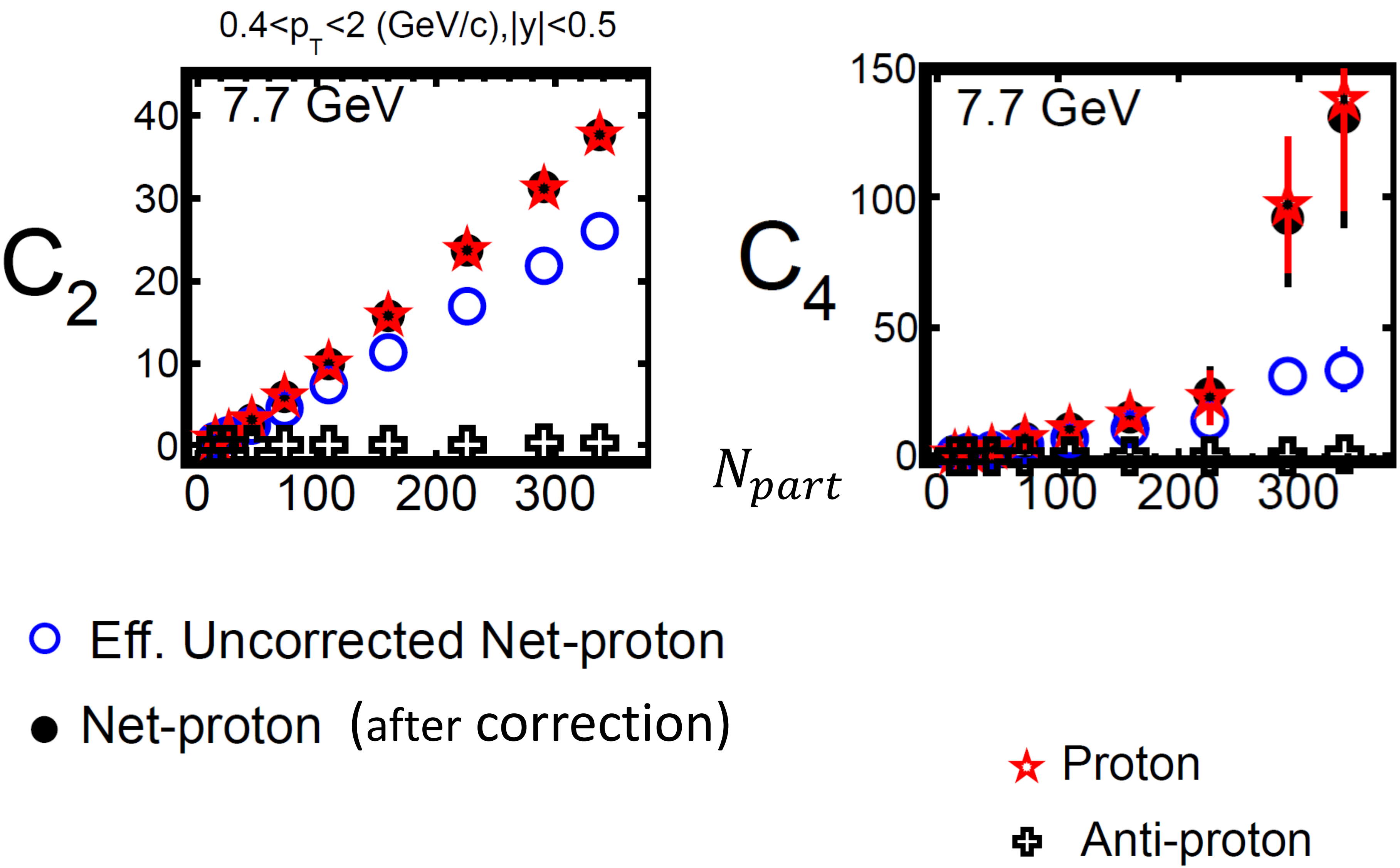


Unfolding vs. average efficiency

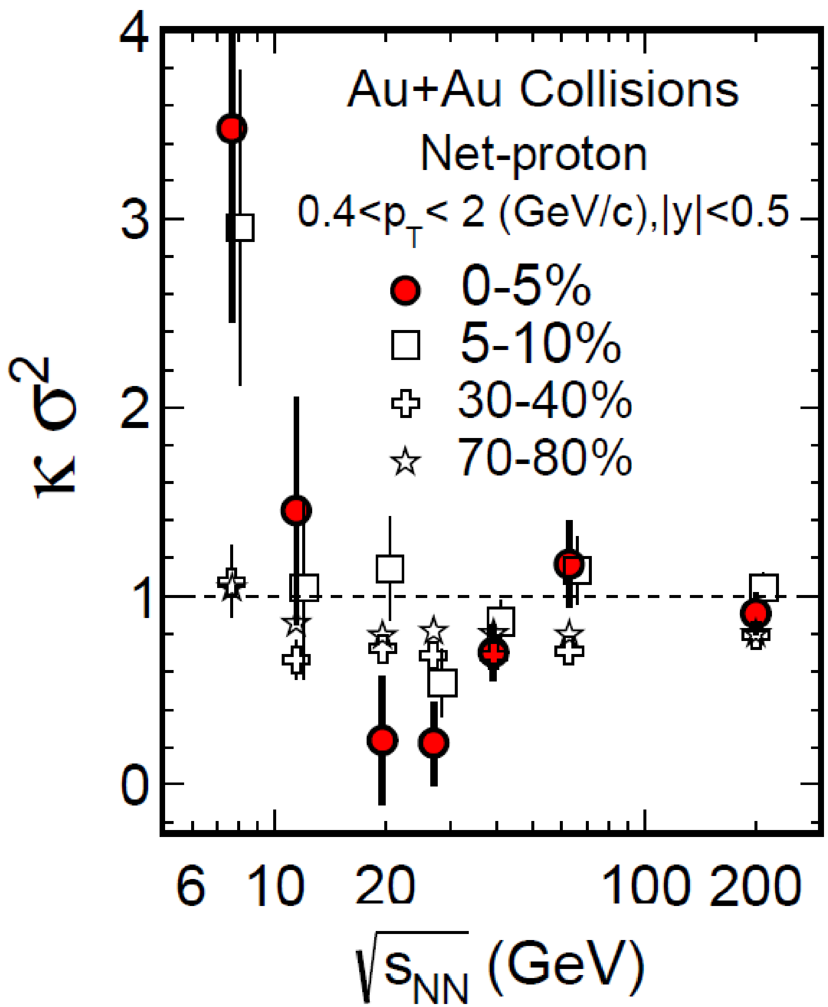


Very similar statistical error bars.

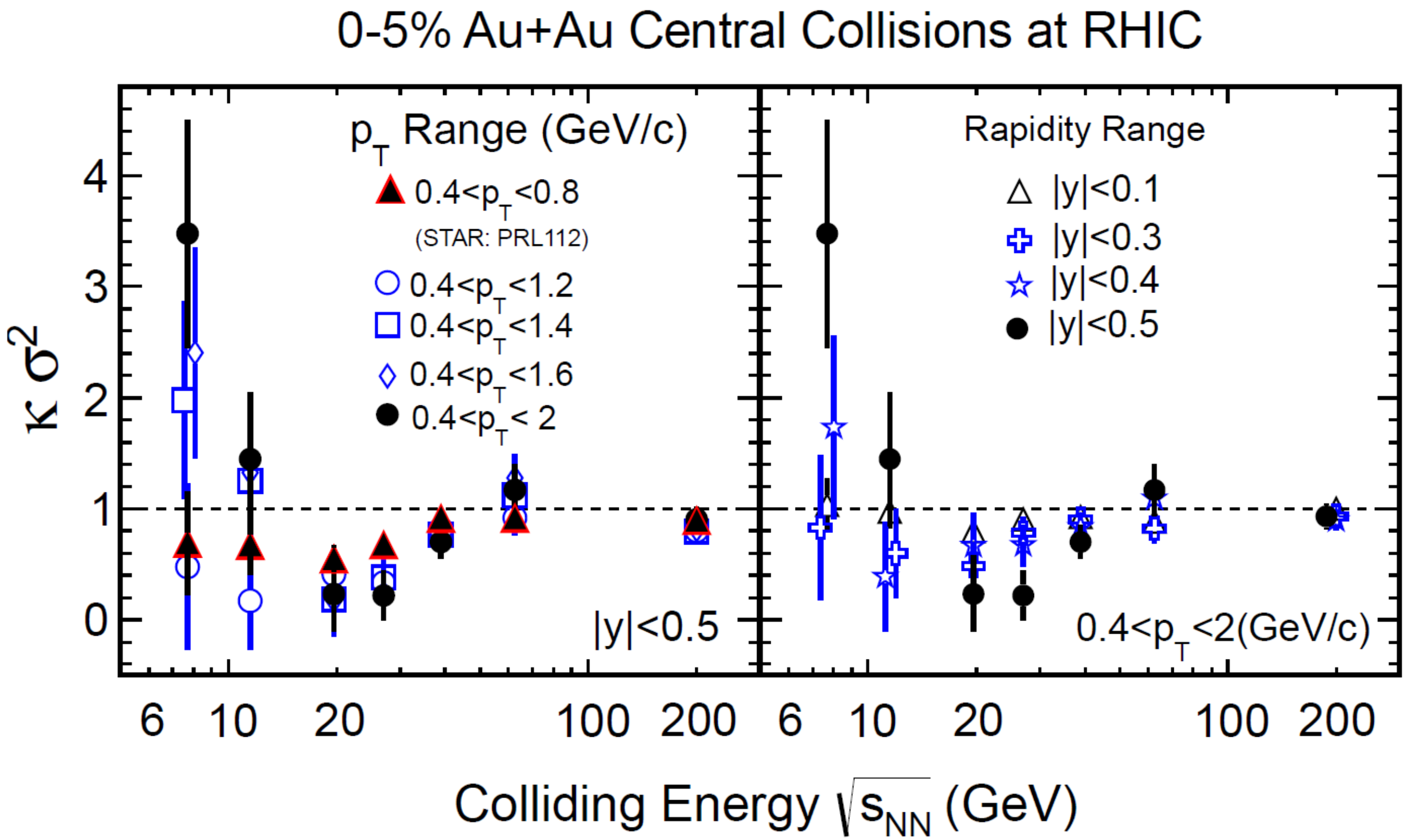
X.Luo, STAR, 1503.02558



New STAR data at 7.7 GeV



Acceptance



Signal is at high p_t, $p_t > 1.2$ GeV. Is it expected from theory?

Signal is pretty sensitive to range of y.

Are stopped protons as good as produced ones?