A general procedure for detector response correction of higher order cumulants

Toshihiro Nonaka[1], Masakiyo Kitazawa[2], ShinIchi Esumi[3]
Central China Normal University[1], Osaka University[2],
University of Tsukuba[3]

NA61-theory virtual meeting

July 19, 2018

Based on arXiv:1805.00279, submitted to Nucl.Instrum.Meth. A

Outline

- Introduction
- Efficiency correction with binomial assumption
- More general procedures
- Summary

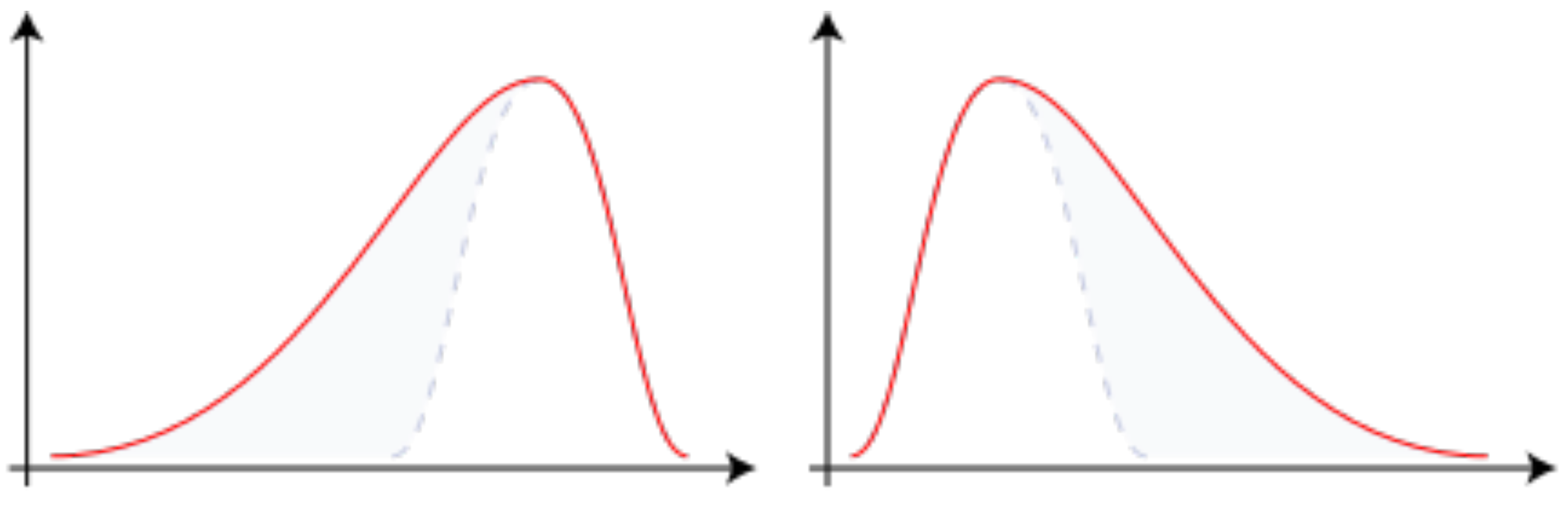
Introduction

Higher order fluctuation

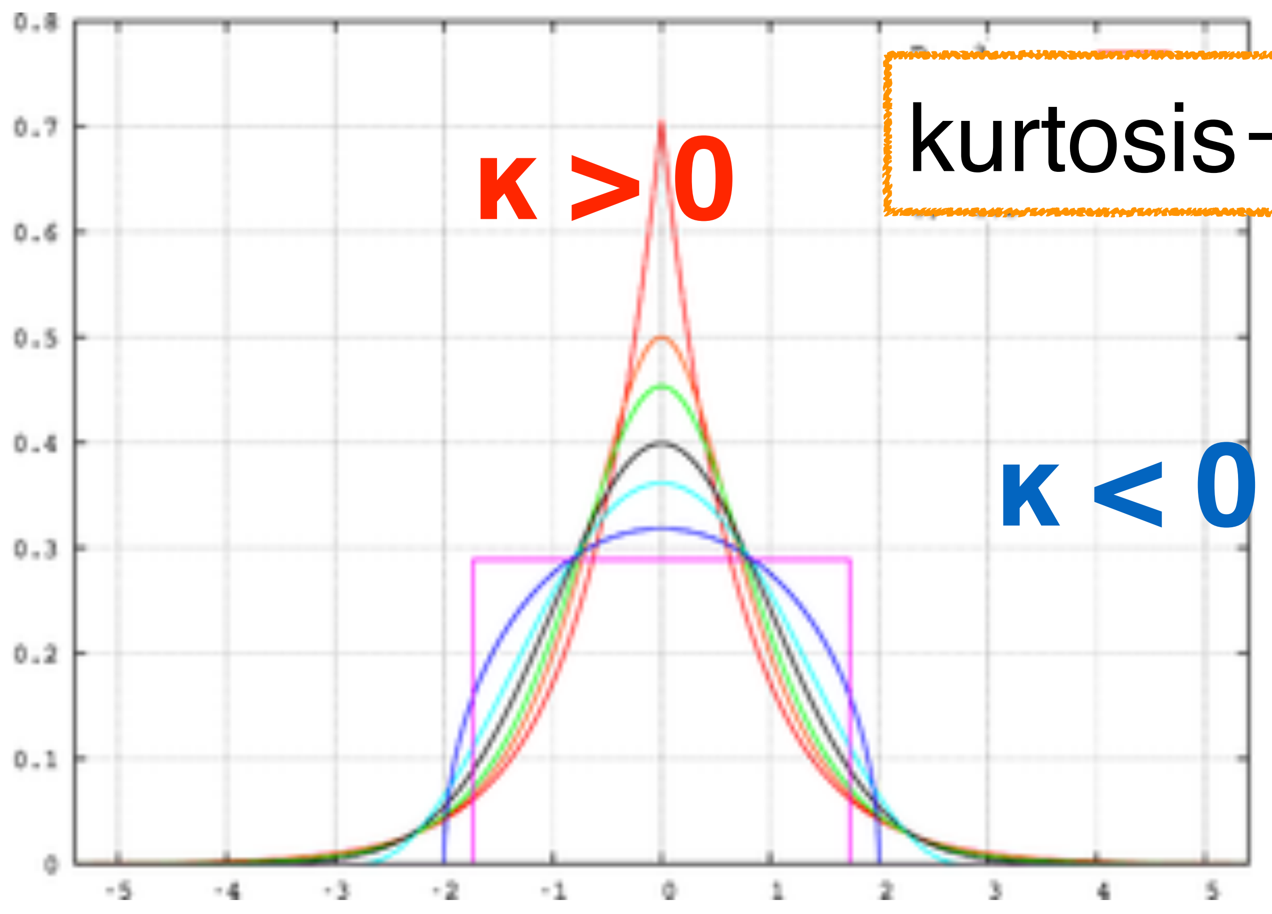
✦ **Moments and cumulants are mathematical measures of "shape" of a distribution which probe the fluctuation of observables.**

- ✓ **Moments: mean (*M*), standard deviation (*σ*), skewness (*S*) and kurtosis (κ).**
- ✓ ***S and κ are non-gaussian* fluctuations.**

skewness→asymmetry



kurtosis→sharpness



from wikipedia

✓ **Cumulant ⇄ Moment**

$$< \delta N >= N - < N >$$
$$C_1 = M =< N >$$
$$C_2 = \sigma^2 =< (\delta N)^2 >$$
$$C_3 = S\sigma^3 =< (\delta N)^3 >$$
$$C_4 = \kappa\sigma^4 =< (\delta N)^4 > -3 < (\delta N)^2 >^2$$

✓ **Cumulant : additivity**

$$C_n(X+Y) = C_n(X) + C_n(Y)$$

→ **proportional to volume**

Cumulants of conserved quantities

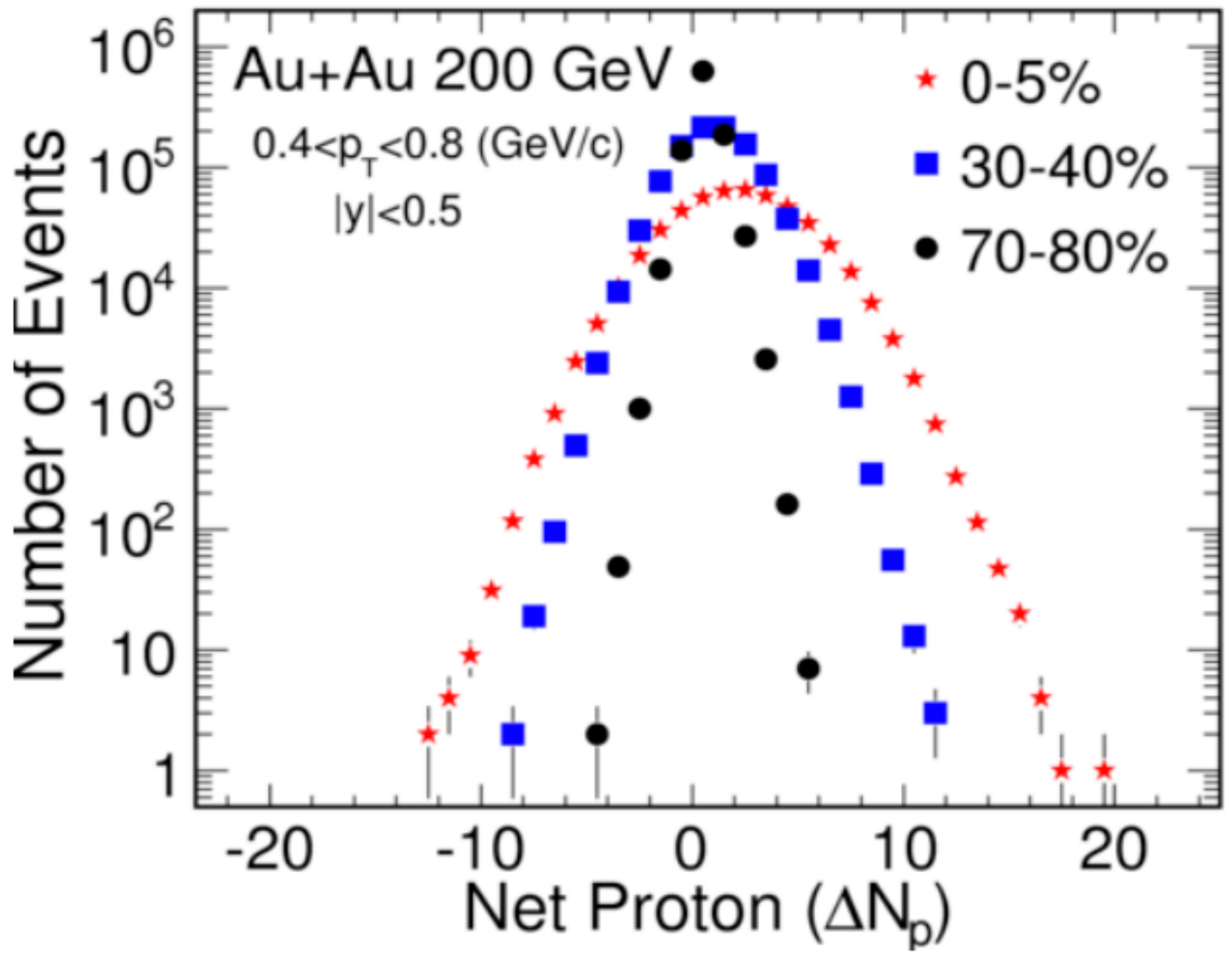
✦ **Net-baryon, net-charge and net-strangeness**

"Net" : positive - negative

$$\Delta N_q = N_q - N_{\bar{q}}, \quad q = B, Q, S$$

No. of positively charged particles in one collision	No. of negatively charged particles in one collision

Fill in histograms over many collisions →



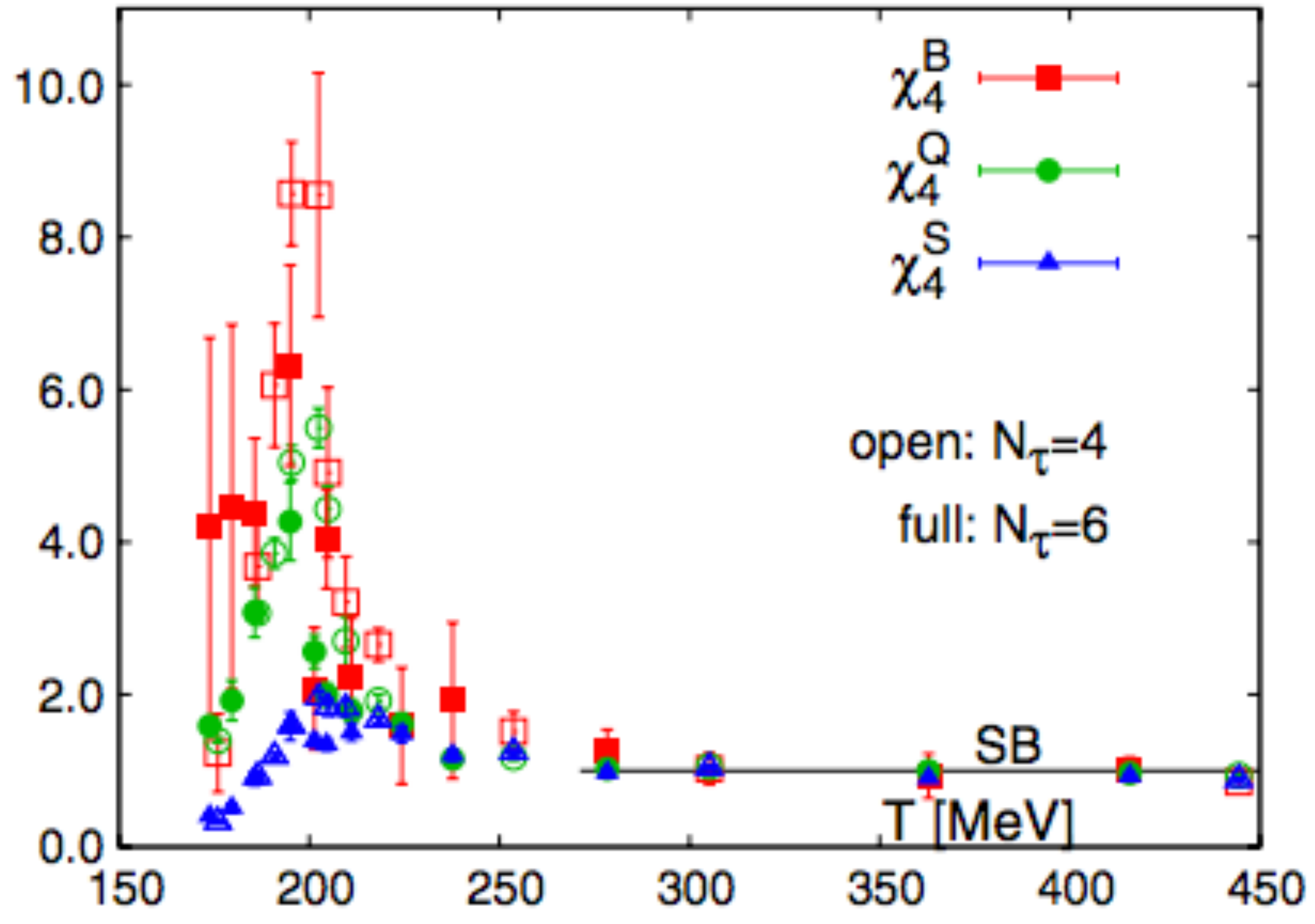
→neutrons cannot be measured

(1) Sensitive to correlation length

$$C_2 =< (\delta N)^2 >_c \approx \xi^2 \qquad C_5 =< (\delta N)^5 >_c \approx \xi^{9.5}$$
$$C_3 =< (\delta N)^3 >_c \approx \xi^{4.5} \qquad C_6 =< (\delta N)^6 >_c \approx \xi^{12}$$
$$C_4 =< (\delta N)^4 >_c \approx \xi^7$$

(2) Direct comparison with susceptibilities.

M. Cheng et al, PRD 79, 074505 (2009)



$$S\sigma = \frac{C_3}{C_2} = \frac{\chi_3}{\chi_2} \qquad \kappa\sigma^2 = \frac{C_4}{C_2} = \frac{\chi_4}{\chi_2}$$

$$\chi_n^q = \frac{1}{VT^3} \times C_n^q = \frac{\partial^n p/T^4}{\partial \mu_q^n}, \quad q = B, Q, S$$

***Volume* dependence can be canceled by taking ratio.**

Two difficulties

✓ Two issues which have not been fully understood yet.

- Initial volume fluctuation

- It is know that event-by-event participant (impact parameter) fluctuation affects the value of cumulants.
- Several model dependent correction methods.

- Non-binomial detector effects

- Detectors miss some particles with finite probability called “efficiency”, which distorts the value of cumulants.
- Conventional correction methods relies on the assumption that efficiencies are binomial.
- **Experimentally, however, it may not be binomial.**

Efficiency correction with binomial assumption

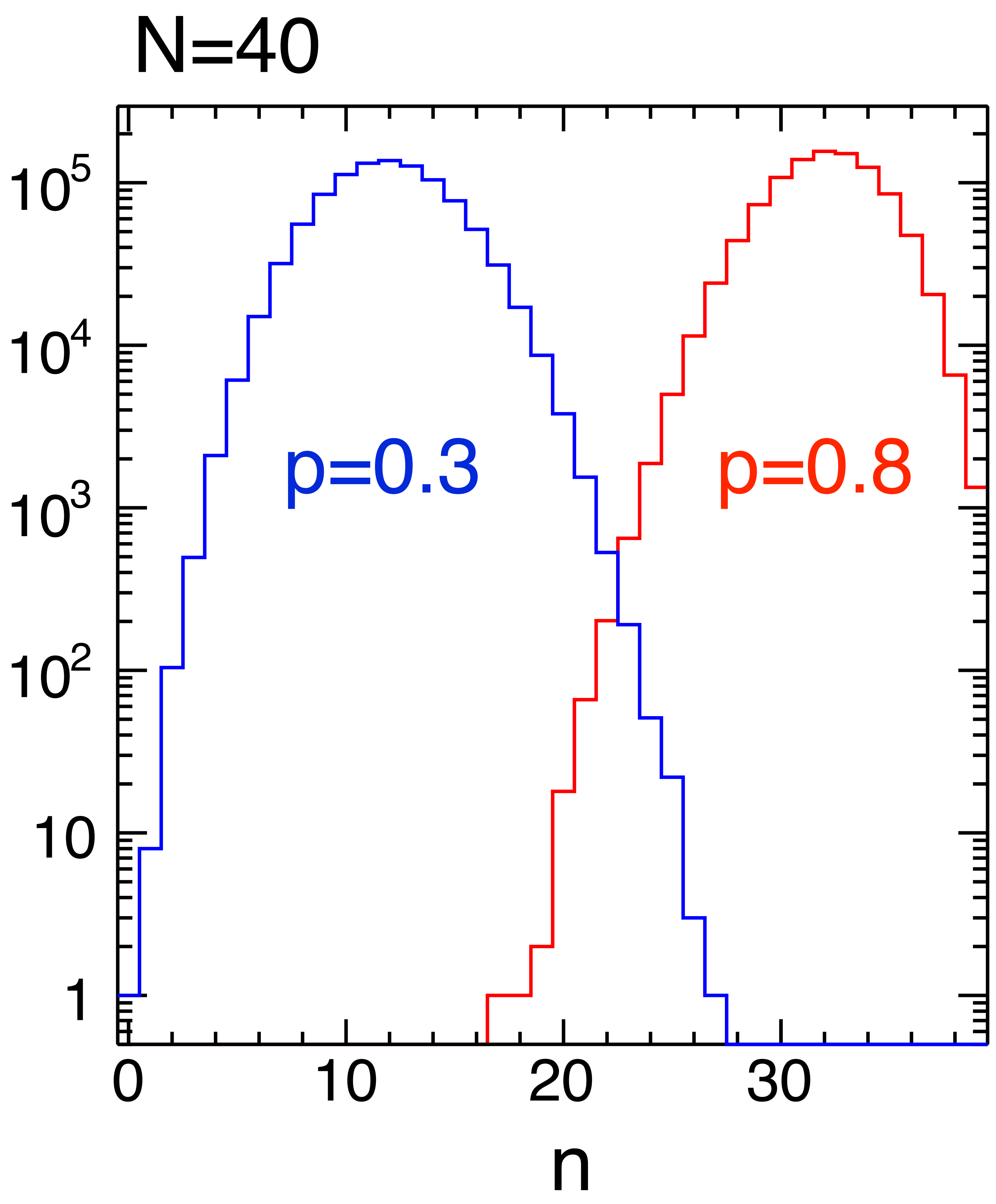
$$B(n;p,N) = p^n(1-p)^{N-n}\frac{N!}{n!(N-n)!}$$

M. Kitazawa and M. Asakawa : PRC.86.(2012)024904

$$\langle N\rangle = \xi_1^{-1}\langle n\rangle,$$
$$\langle N^2\rangle_c = \xi_1^{-2}\langle n^2\rangle_c - \xi_2\xi_1^{-3}\langle n\rangle,$$
$$\langle N^3\rangle_c = \xi_1^{-3}\langle n^3\rangle_c - 3\xi_2\xi_1^{-4}\langle n^2\rangle_c + (3\xi_2^2\xi_1^{-5} - \xi_3\xi_1^{-4})\langle n\rangle,$$
$$\langle N^4\rangle_c = \xi_1^{-4}\langle n^4\rangle_c - 6\xi_2\xi_1^{-5}\langle n^3\rangle_c + (15\xi_2^2\xi_1^{-6} - 4\xi_3\xi_1^{-5})\langle n^2\rangle_c - (15\xi_2^3\xi_1^{-7} - 10\xi_2\xi_3\xi_1^{-6} + \xi_4\xi_1^{-5})\langle n\rangle$$

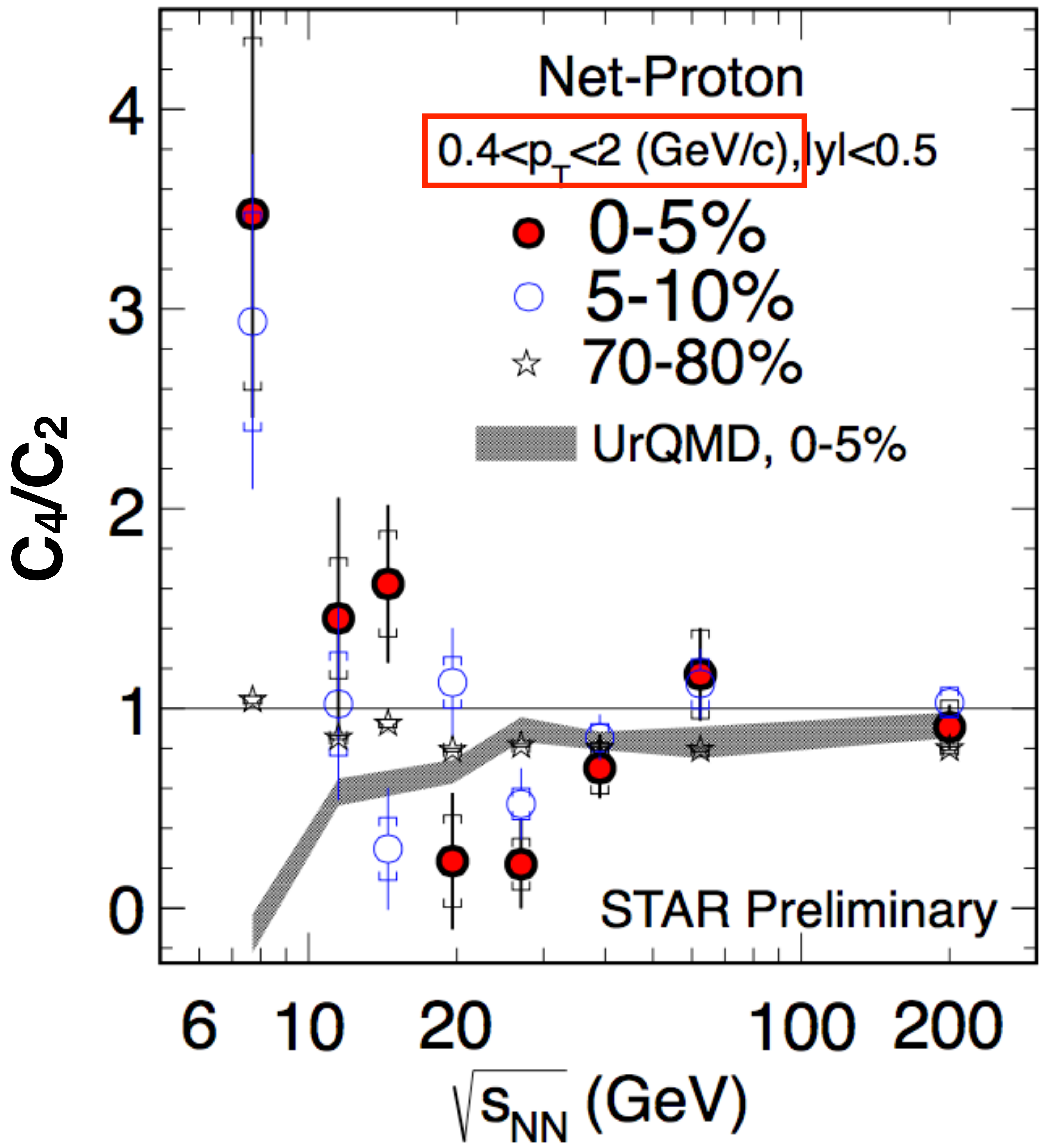
A. Bzdak and V. Koch : PRC.(2012)86.044904

$$pK_1 = c_1,$$
$$p^2K_2 = c_2 - n(1-p),$$
$$p^3K_3 = c_3 - c_1(1-p^2) - 3(1-p)(f_{20} - f_{02} - nc_1),$$
$$p^4K_4 = c_4 - np^2(1-p) - 3n^2(1-p)^2 - 6p(1-p)(f_{20}+f_{02}) + 12c_1(1-p)(f_{20}-f_{02}) - (1-p^2)(c_2 - 3c_1^2) - 6n(1-p)(c_1^2 - c_2) - 6(1-p)(f_{03} - f_{12} + f_{02} + f_{20} - f_{21} + f_{30}).$$



Efficiency correction with binomial assumption

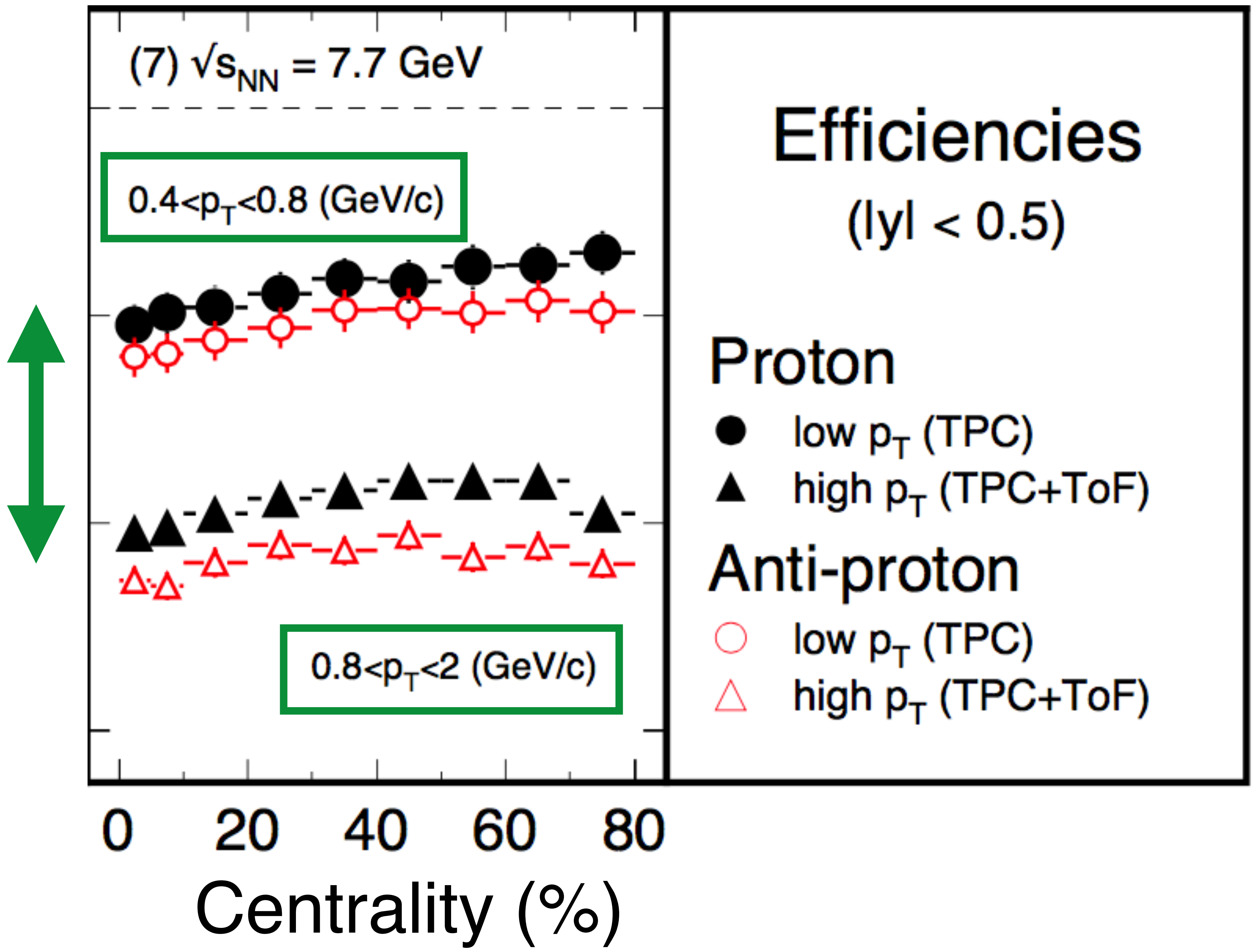
X. Luo (STAR collaboration) arXiv:1503.02558v2



✓ **Efficiency correction was extended to multi-variable case (many efficiency bins).**

A. Bzdak and V. Koch : PRC.91.(2015)027901

Different efficiency between two p_T regions due to different PID



Efficient formulas

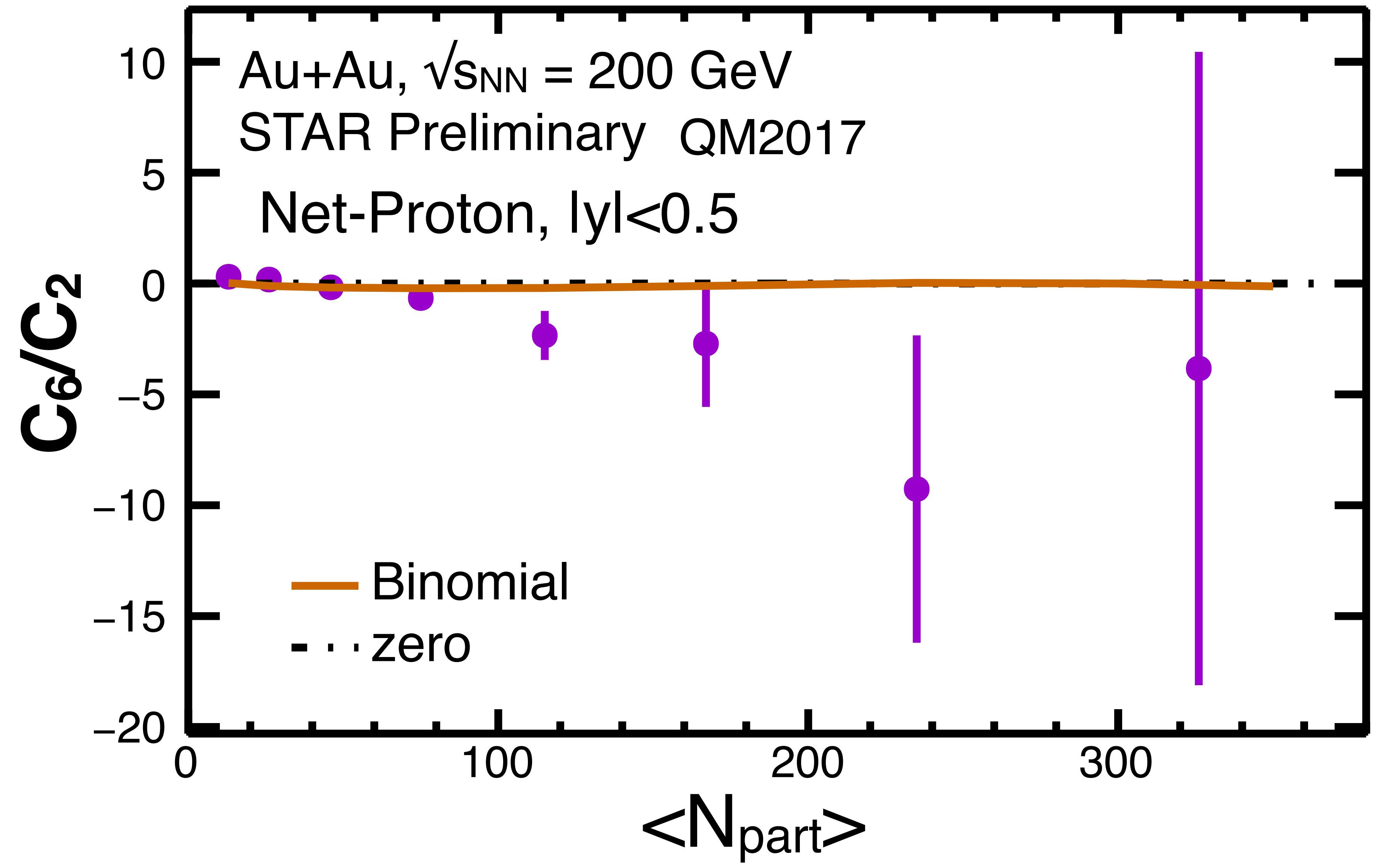
✓ **Huge calculation cost for higher order with many efficiency bins.**

$\sim M^m$ for large M

m : order of cumulant
M : # of efficiency bins

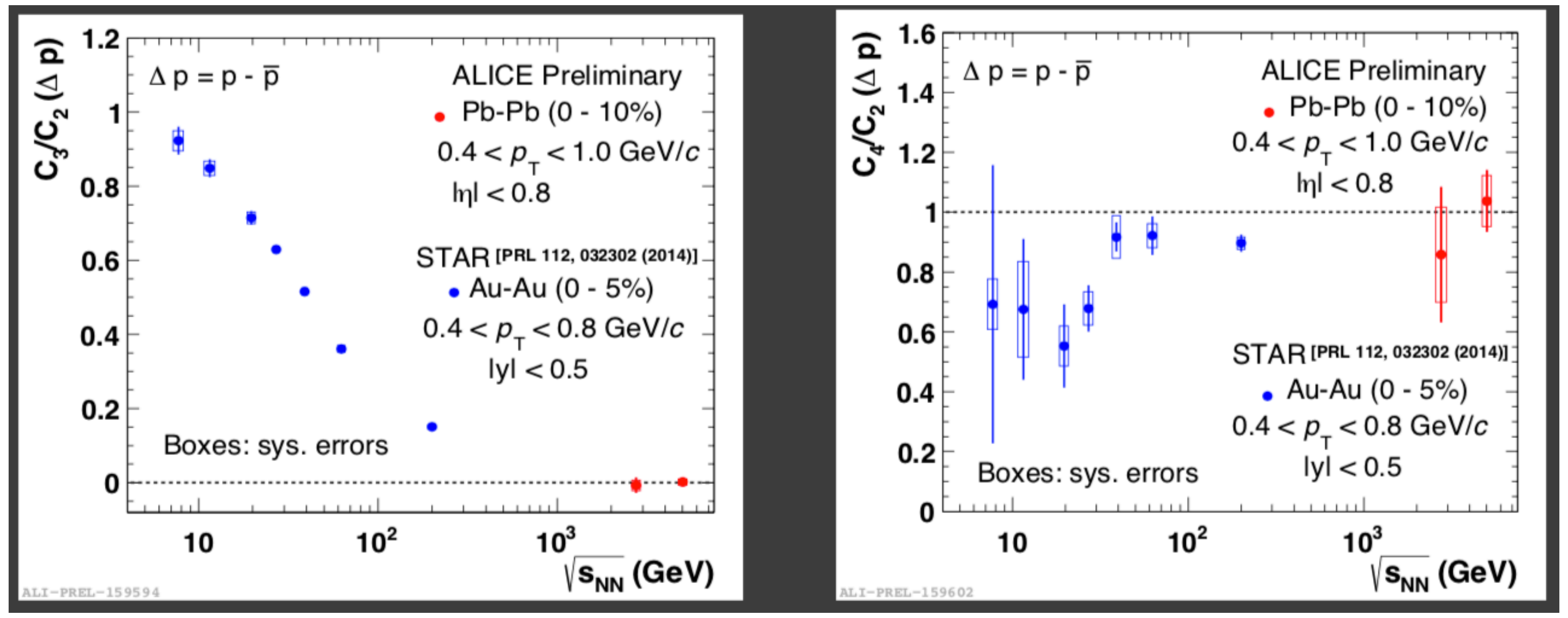
✓ **More efficient formulas have been developed.**

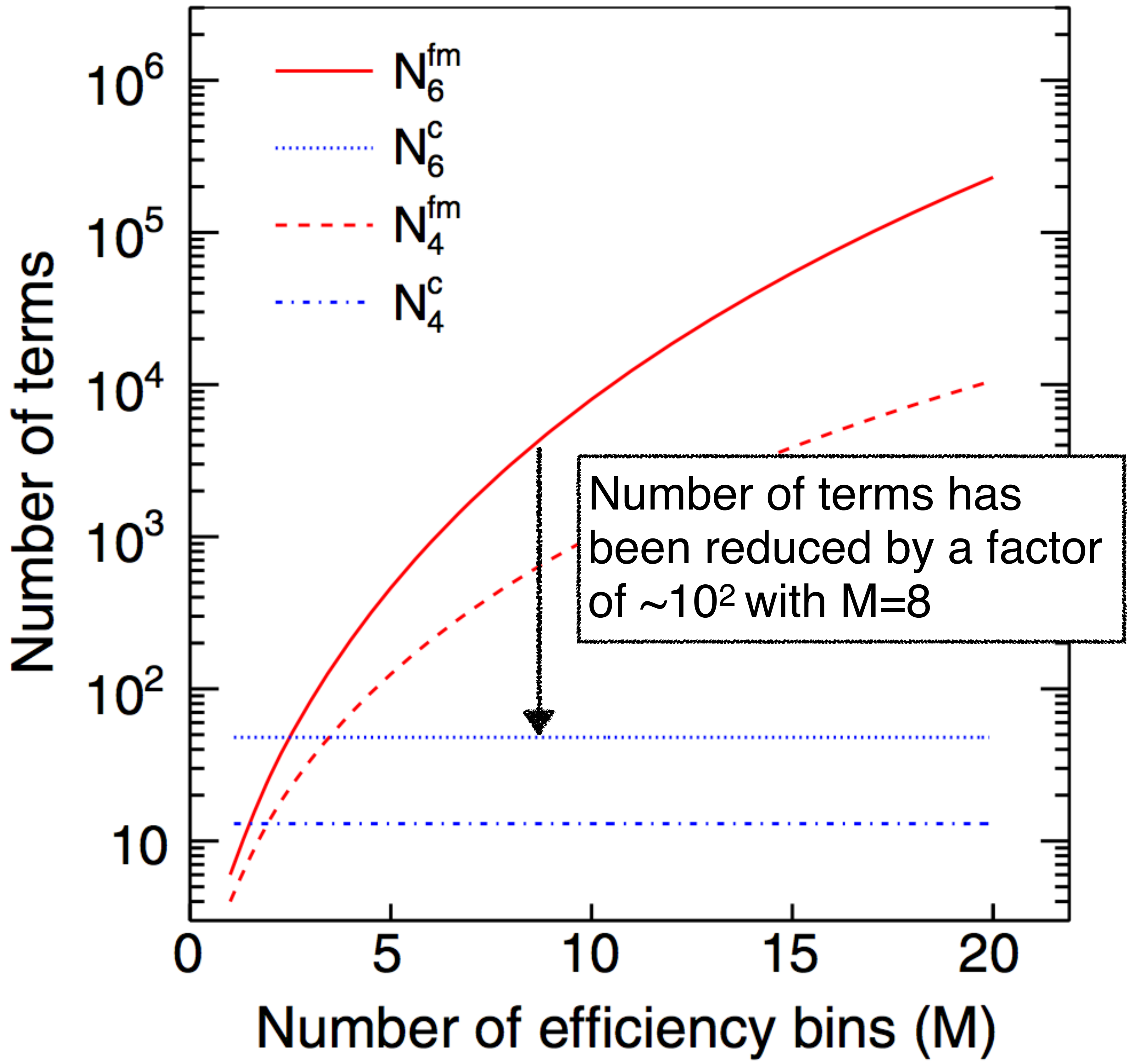
- **M. Kitazawa : PRC.93.(2016)044911**
- **T. Nonaka, M. Kitazawa, S. Esumi, PRC.94.(2017)034909**
- **Used in STAR C_6 and ALICE C_4**



T. Nonaka, QM2017, poster

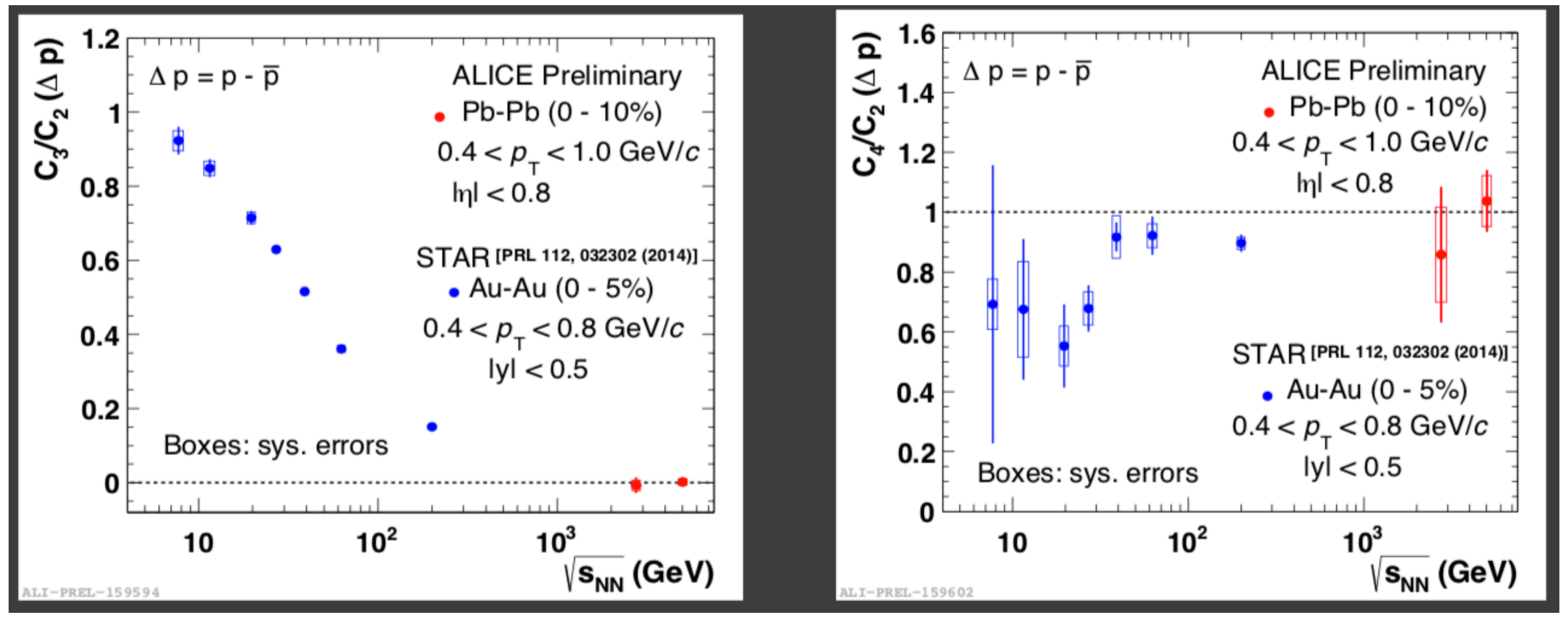
N. K. Behera (ALICE), QM2018 → ~20 efficiency bins





Efficient formulas

N. K. Behera (ALICE), QM2018 → ~20 efficiency bins



✓ **Huge calculation cost for higher order with many efficiency bins.**

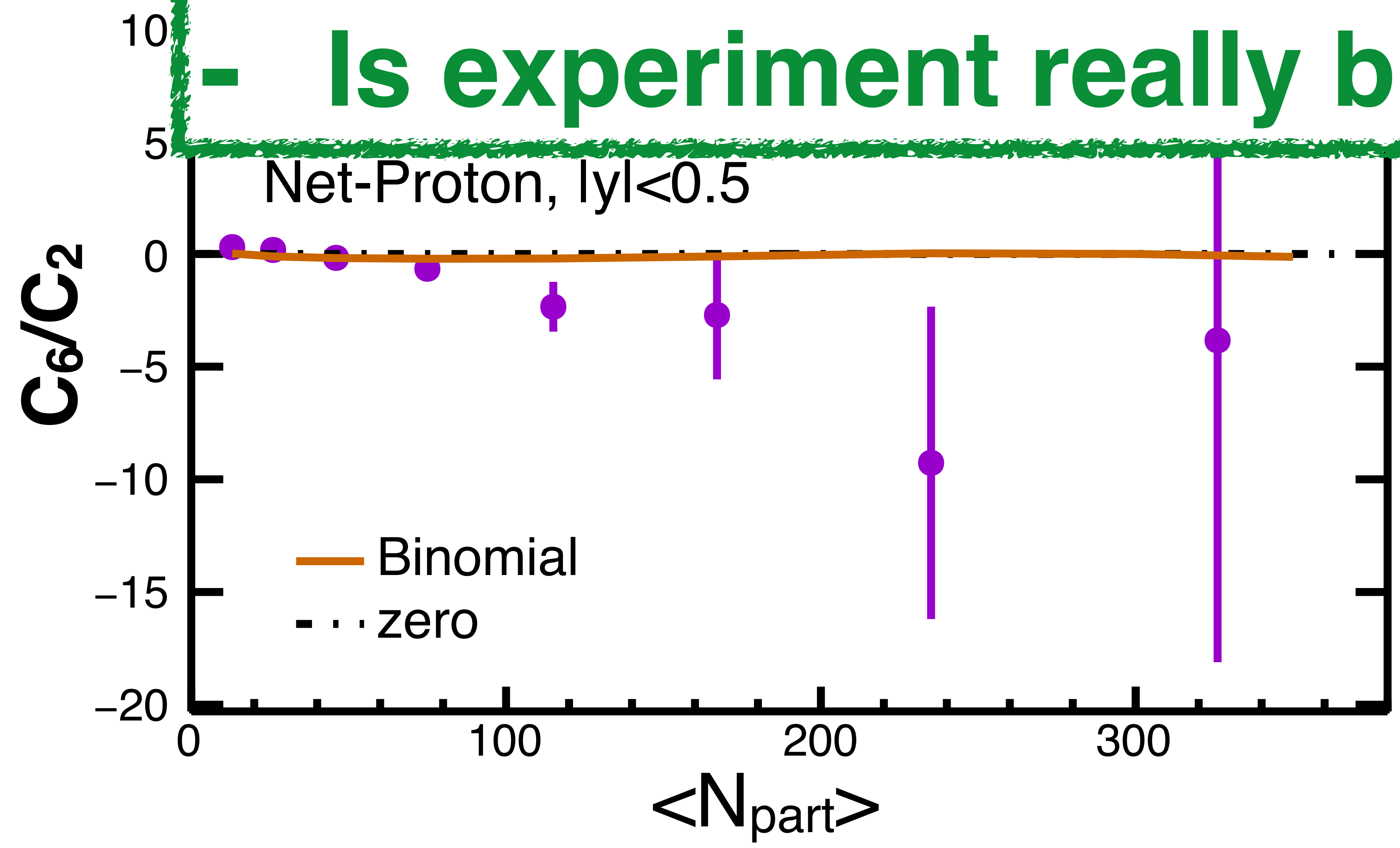
$\sim M^m$ for large M

m : order of cumulant
M : # of efficiency bins

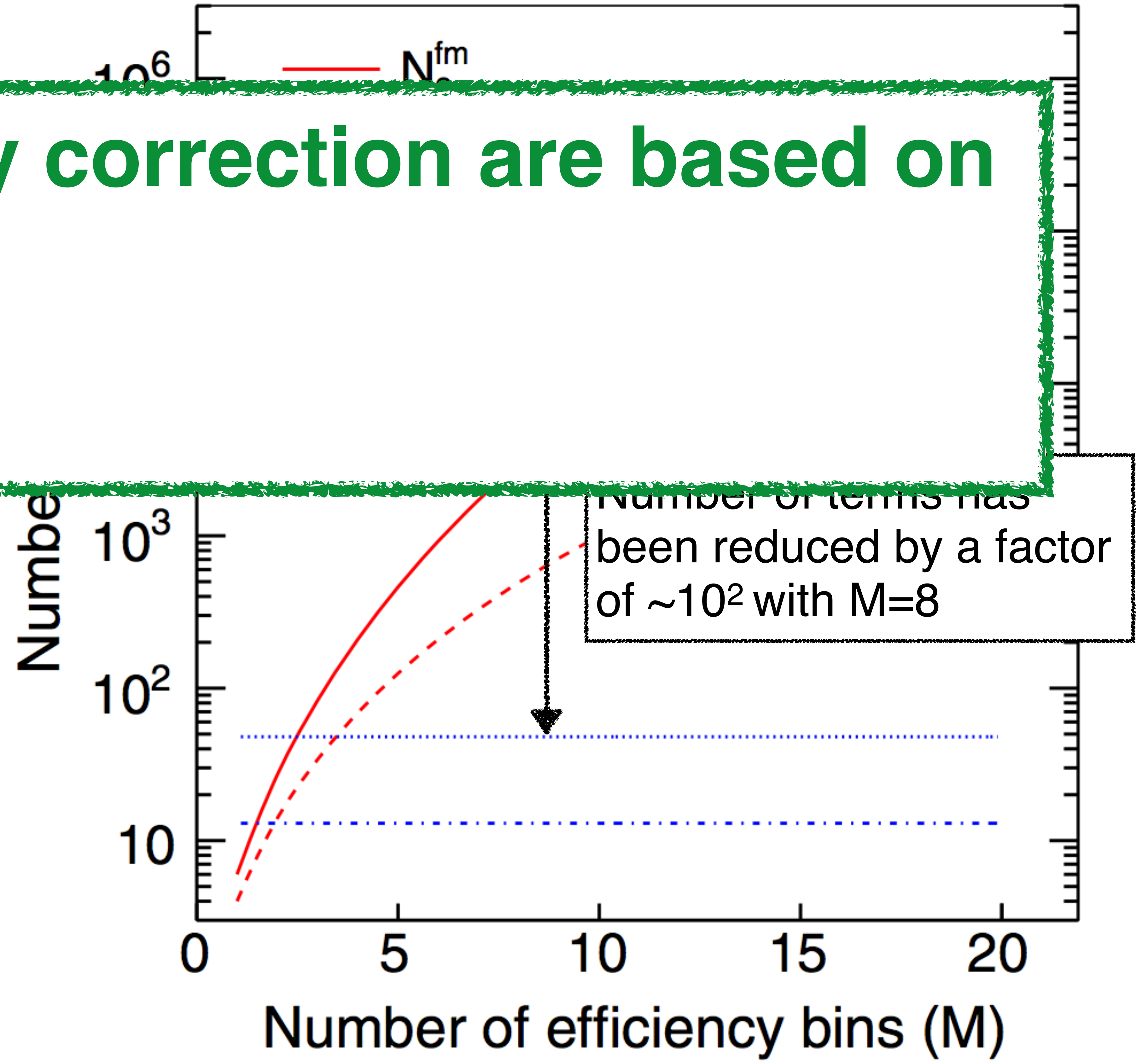
✓ **More efficient formulas have been developed.**

- **M. Ki**
- **T. Nor**
- **Used**

- **So far, all methods for efficiency correction are based on the binomial assumption.**
- **Is experiment really binomial?**



T. Nonaka, QM2017, poster

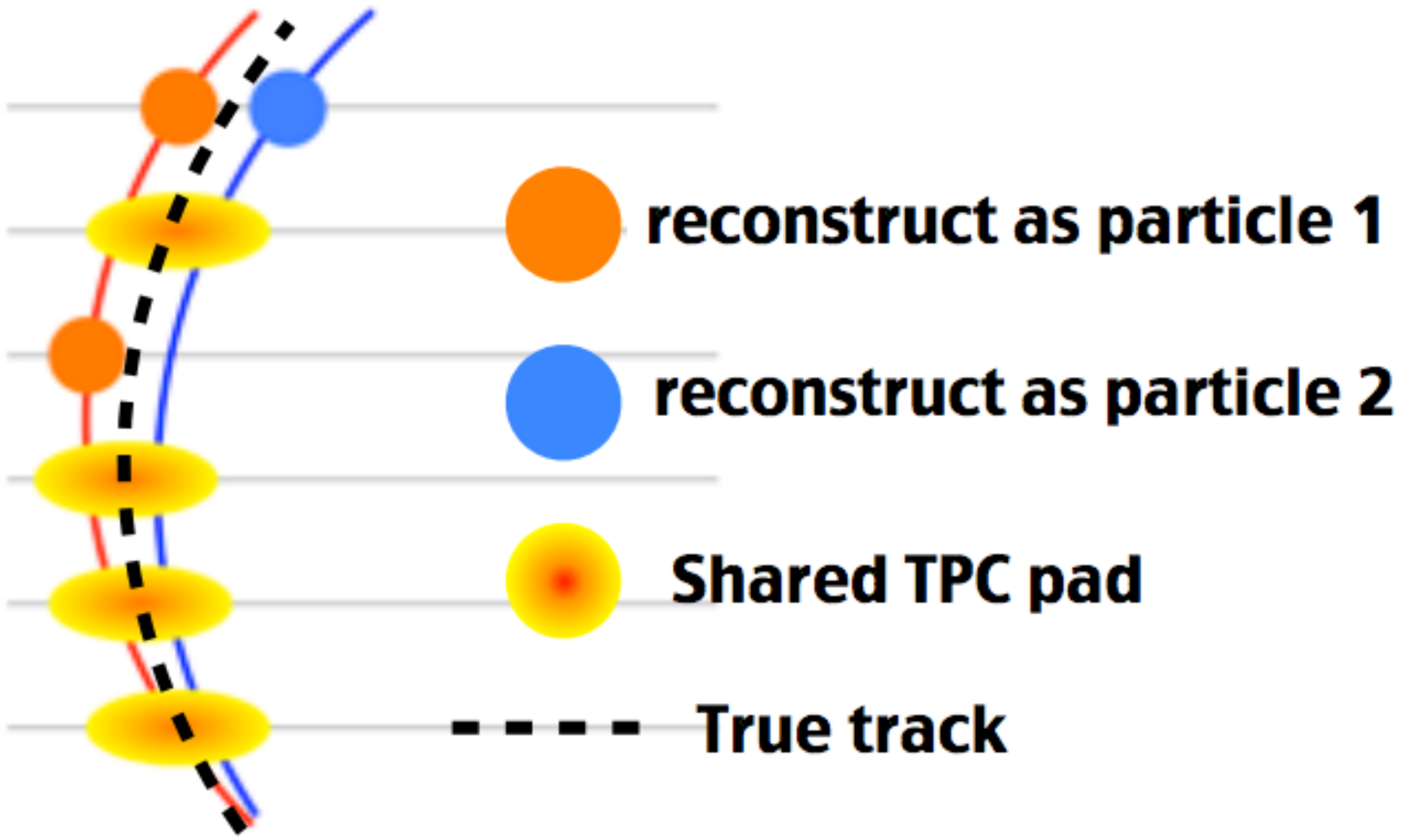


Where does non-binomial come from?

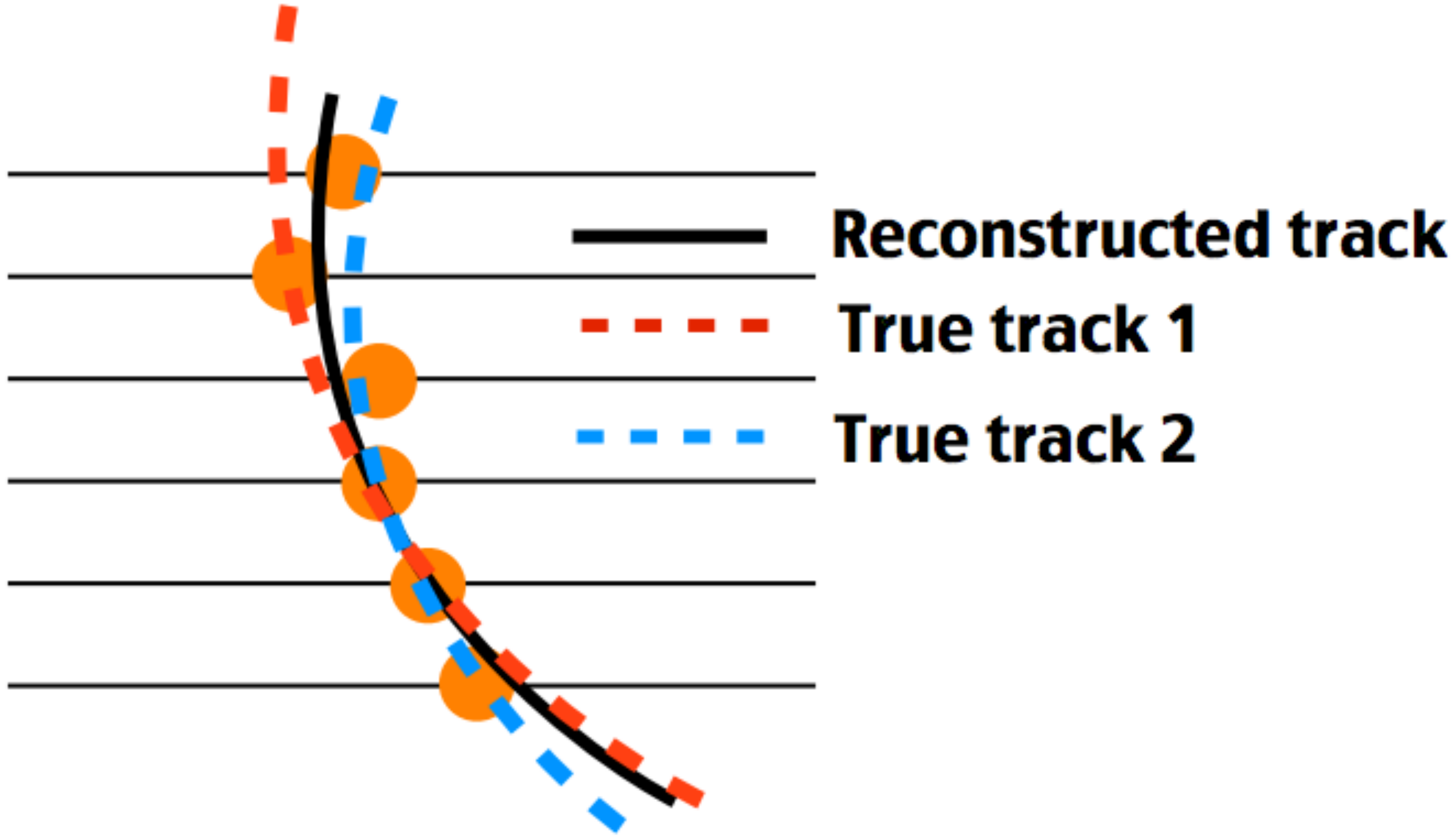
✓ **There might be some reasons to make the response matrix non-binomial, which can be understood only by MC.**

- Track splitting/merging
- Time dependence of efficiency
- Particle misidentification

Track splitting



Track merging



If experiment is non-binomial…

A.Bzdak, R. Holzmann, V. Koch : PRC.94.(2016)064907

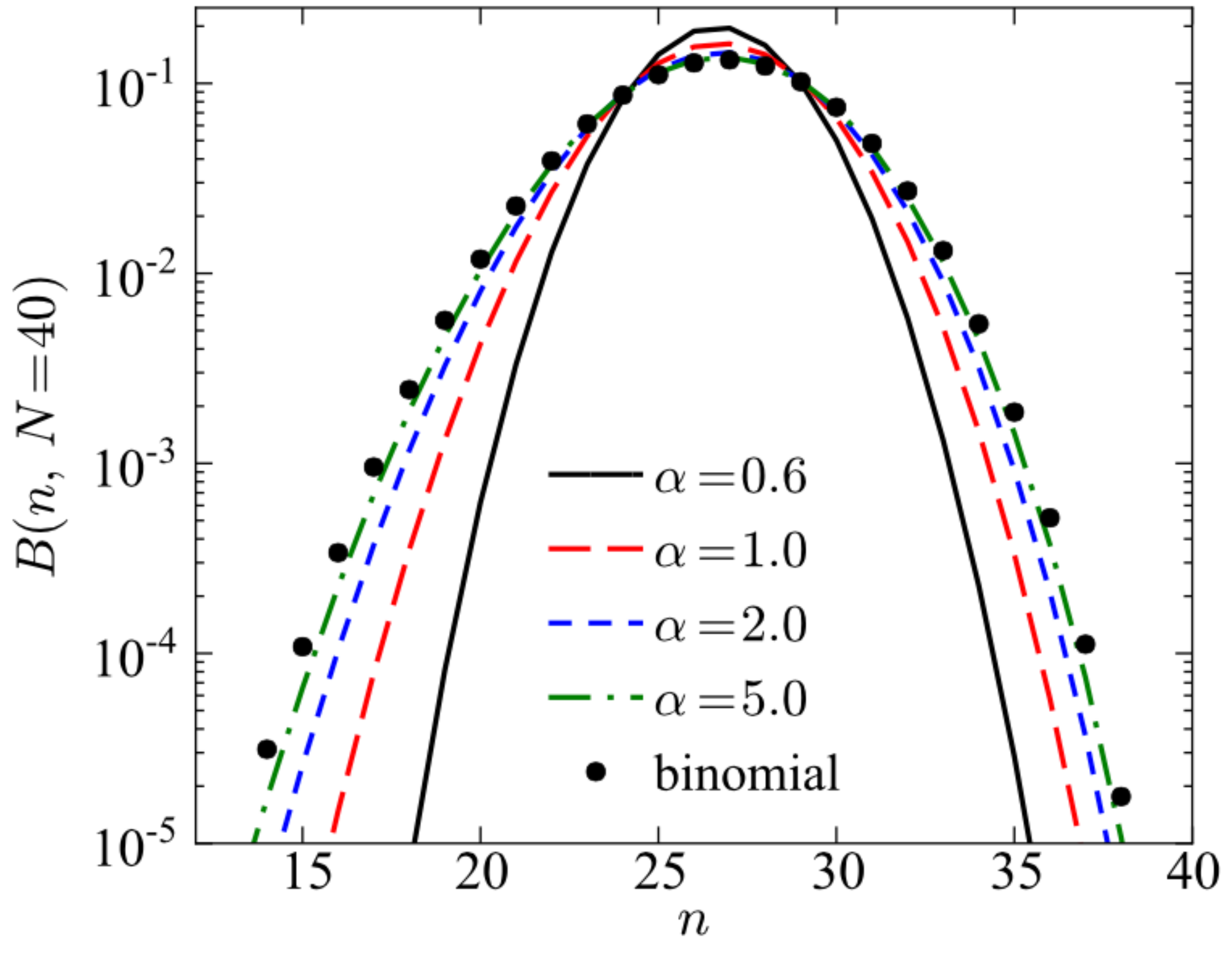


FIG. 2. The hypergeometric distribution for different values of α compared with the binomial distribution (black points). Here $N = 40$ and $\epsilon = 2/3$.

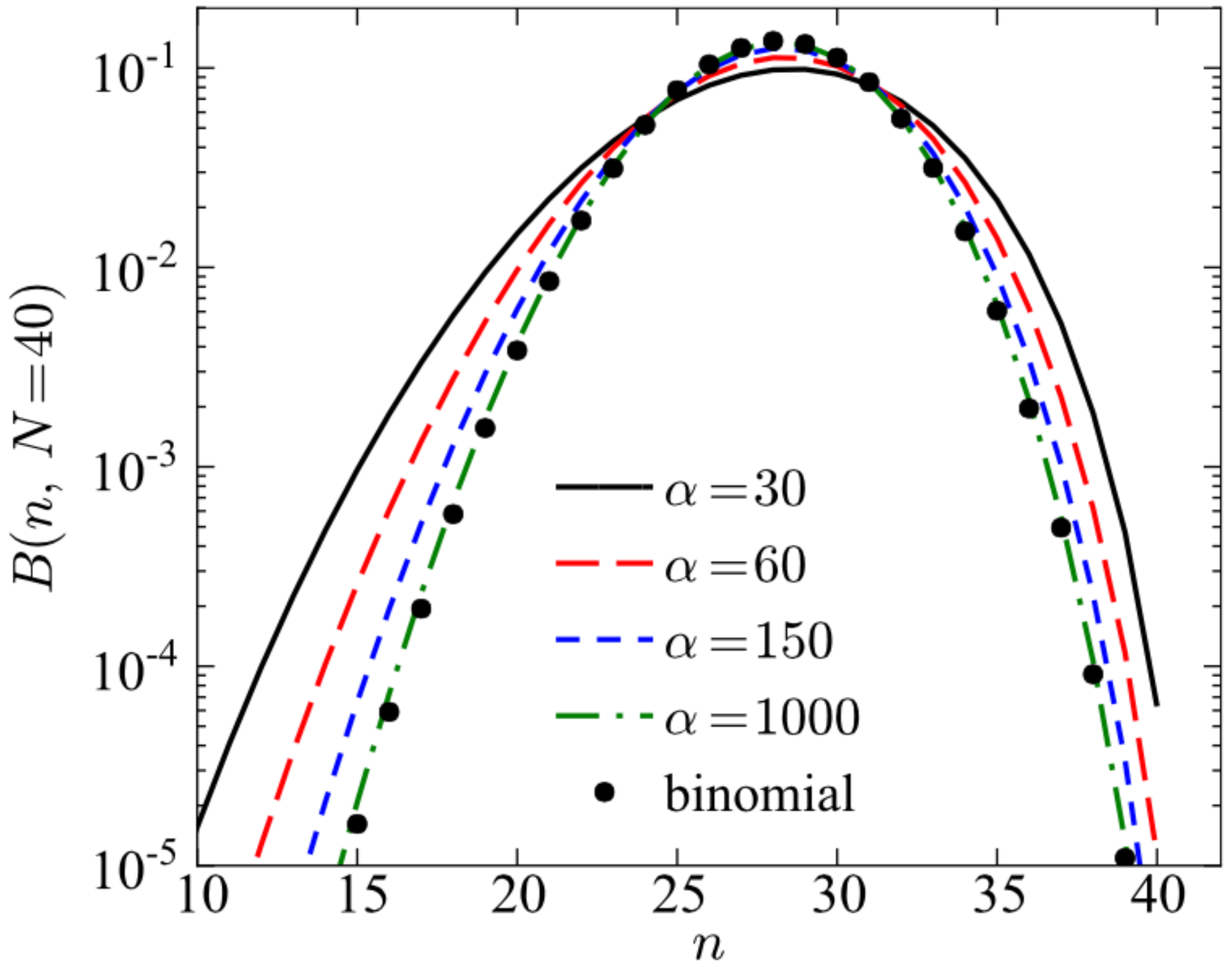


FIG. 3. The beta-binomial distribution for different values of α compared with the binomial distribution (black points). Here $N = 40$ and $\epsilon = 0.7$.

Hypergeometric

TABLE I. The obtained values of K_n/K_2 for the hypergeometric distribution, using $F_i = f_i/\epsilon^i$ with $\epsilon = 2/3$, for different values of α as presented in Fig. 2.

Hypergeometric	$\alpha = 0.6$	$\alpha = 1.0$	$\alpha = 2.0$	$\alpha = 5.0$
K_3/K_2	1.16	1.12	1.07	1.03
K_4/K_2	0.66	0.88	0.98	1.00
K_5/K_2	2.19	1.68	1.23	1.05
K_6/K_2	−3.99	−1.38	0.31	0.89

Beta-binomial

TABLE II. The obtained values of K_n/K_2 for the beta-binomial distribution, using $F_i = f_i/\epsilon^i$ with $\epsilon = 0.7$, for different values of α as presented in Fig. 3.

beta binomial	$\alpha = 30$	$\alpha = 60$	$\alpha = 150$	$\alpha = 1000$
K_3/K_2	1.28	1.24	1.13	1.02
K_4/K_2	0.82	1.45	1.35	1.07
K_5/K_2	−1.11	1.15	1.63	1.16
K_6/K_2	5.71	−0.44	1.80	1.32

True cumulants are unity, but all results deviate.

✓ **Unfolding is necessary. Reconstruct the distribution itself numerically.**

- **Estimation of systematic uncertainties is difficult.**
- **Huge calculation cost due to iterations.**
- **Distributions are not so necessary.**

More general method to correct for detector effects by cumulants directly?

Notations

$$\tilde{P}(n) = \sum_N \mathcal{R}(n;N) P(N). \quad (1)$$

Response matrix

$$\langle\!\langle n^m \rangle\!\rangle = \sum_n n^m \tilde{P}(n),$$

Measured moments

$$\langle N^m \rangle = \sum_N N^m P(N).$$

True moments

$$\langle\!\langle n^m \rangle\!\rangle = \sum_N P(N) \sum_n n^m \mathcal{R}(n;N) = \sum_N P(N) R_m(N), \quad (4)$$

$$R_m(N) = \sum_n n^m \mathcal{R}(n;N).$$

Moments of response matrix with fixed N

$$R_m(N) = \sum_{j=0}^{\infty} r_{mj} N^j. \quad (6)$$

- ✓ **If Eq. (6) closes at finite order, Eq. (9) can be exactly solved.**
- ✓ **If not, some truncations are necessary.**

$$\langle\!\langle n^m \rangle\!\rangle = r_{m0} + \sum_{j=1}^{\infty} r_{mj} \langle N^j \rangle. \quad (7)$$

Property of the detector

$$\begin{bmatrix} \langle N \rangle \\ \langle N^2 \rangle \\ \vdots \end{bmatrix} = \boldsymbol{R}^{-1} \begin{bmatrix} \langle\!\langle n \rangle\!\rangle \\ \langle\!\langle n^2 \rangle\!\rangle \\ \vdots \end{bmatrix} - \boldsymbol{R}^{-1} \begin{bmatrix} r_{10} \\ r_{20} \\ \vdots \end{bmatrix}. \quad (9)$$

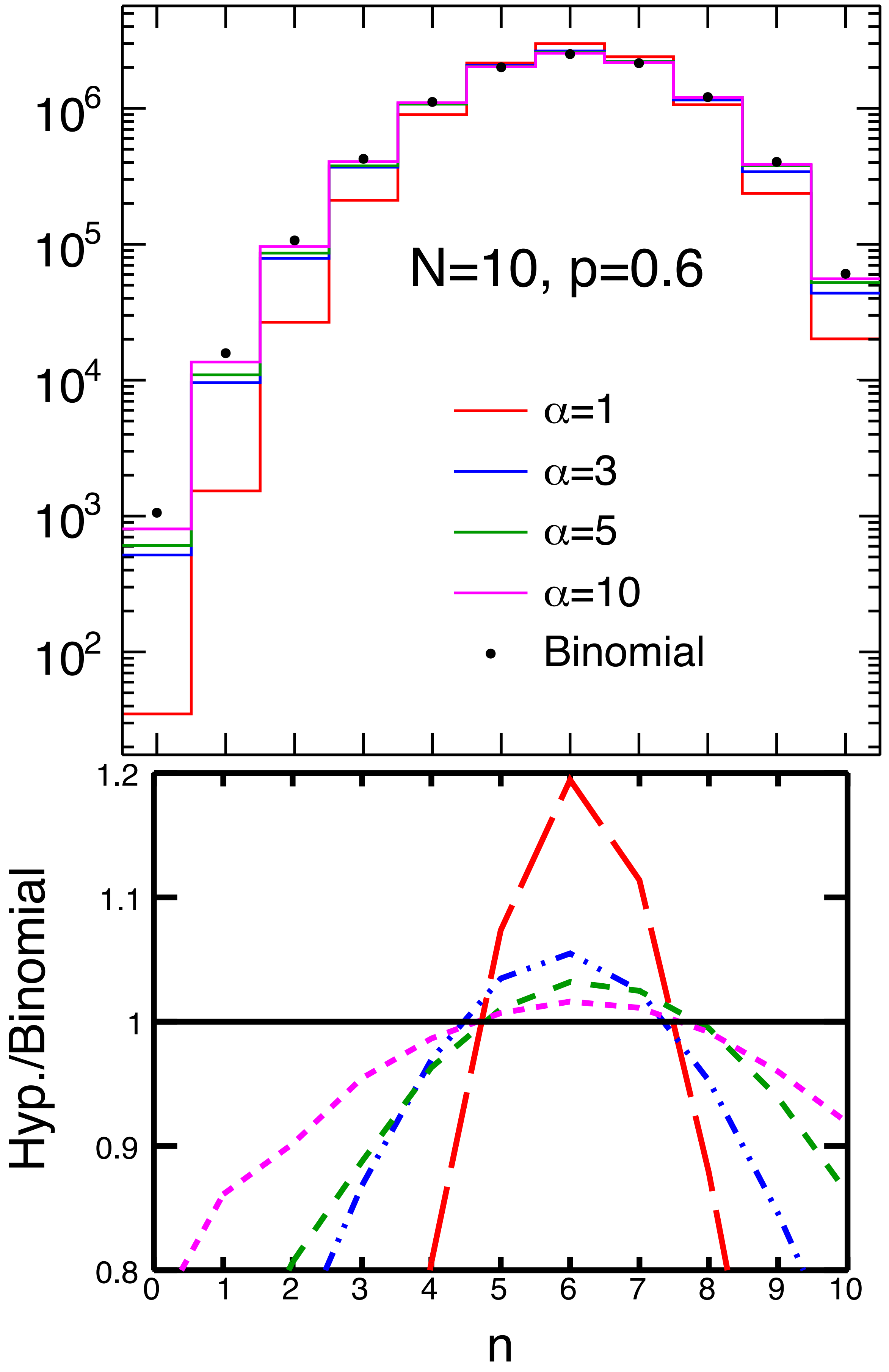
General procedures

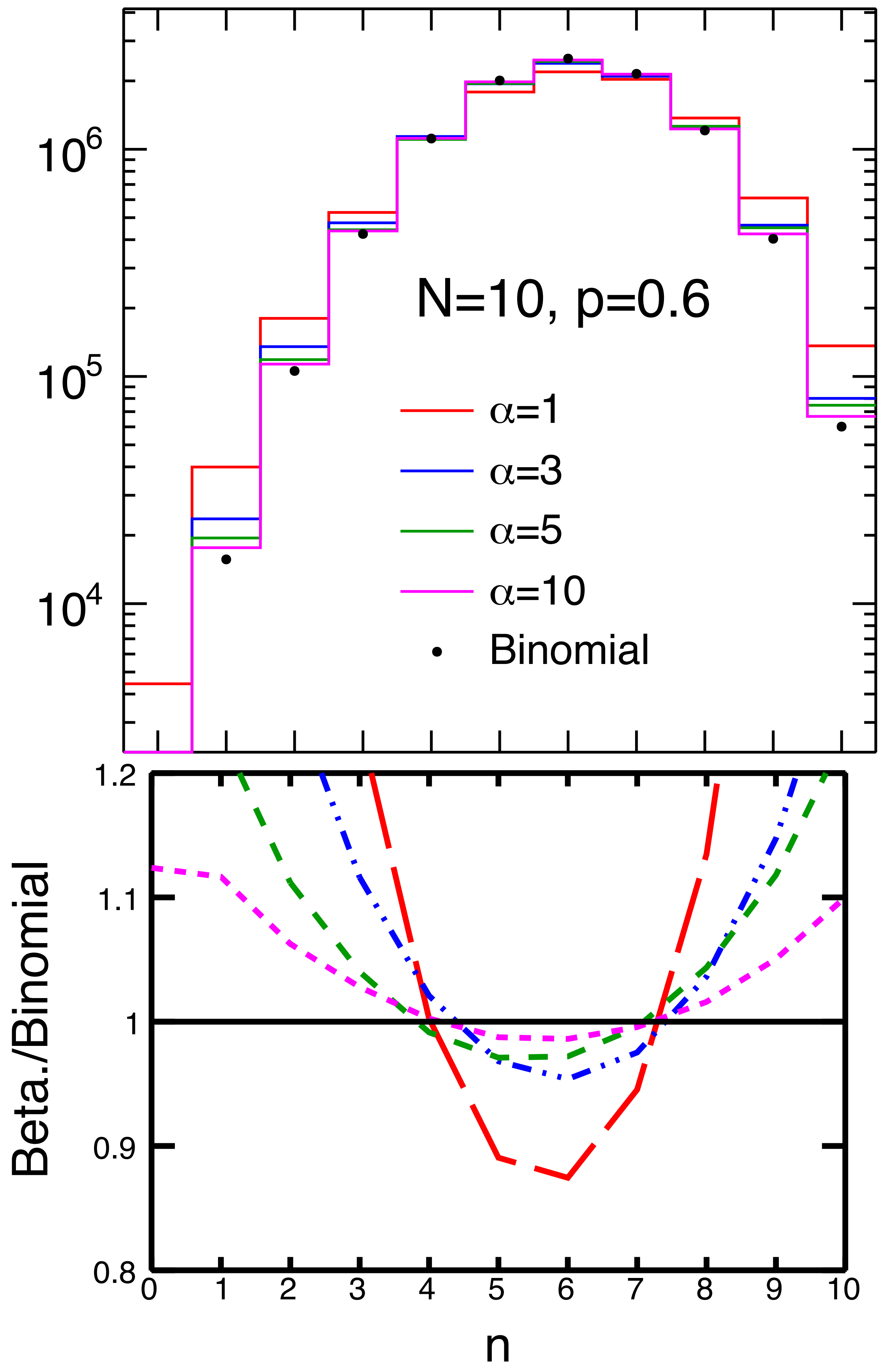
- Exactly solvable case
- Truncation

Non-binomial distribution

Hypergeometric distribution
Draw a ball from urn, if it is white, count particle. This is repeated without replacement.

Beta-binomial distribution
Draw a ball from urn, if it is white, count particle. And return two white balls to urn (similar for black balls).





- $X = N_w$, $Y = N_{tot} = N_w + N_b$
- $p = N_w/N_{tot} = X/Y$

$$\mathcal{R}_{\rm HG}(n;N) = H(n;N,X,Y), \quad (18)$$

$$\mathcal{R}_{\beta}(n;N) = \beta(n;N,X,Y-X), \quad (19)$$

Close to binomial with Y→∞

✓ **Use these two response distributions to check our new method.**

✓ **Also, these examples can be solved analytically, which is discussed in Appendix D.**

Non-binomial RM

- ✓ Assume non-binomial response matrix, hypergeometric and beta-binomial distribution, with efficiency p=0.7.
- ✓ Generate P(N) with Poisson distribution, then sample $\tilde{P}(n)$ by hypergeometric and beta-binomial response function.
- ✓ 10 times larger statistics are generated for MC.

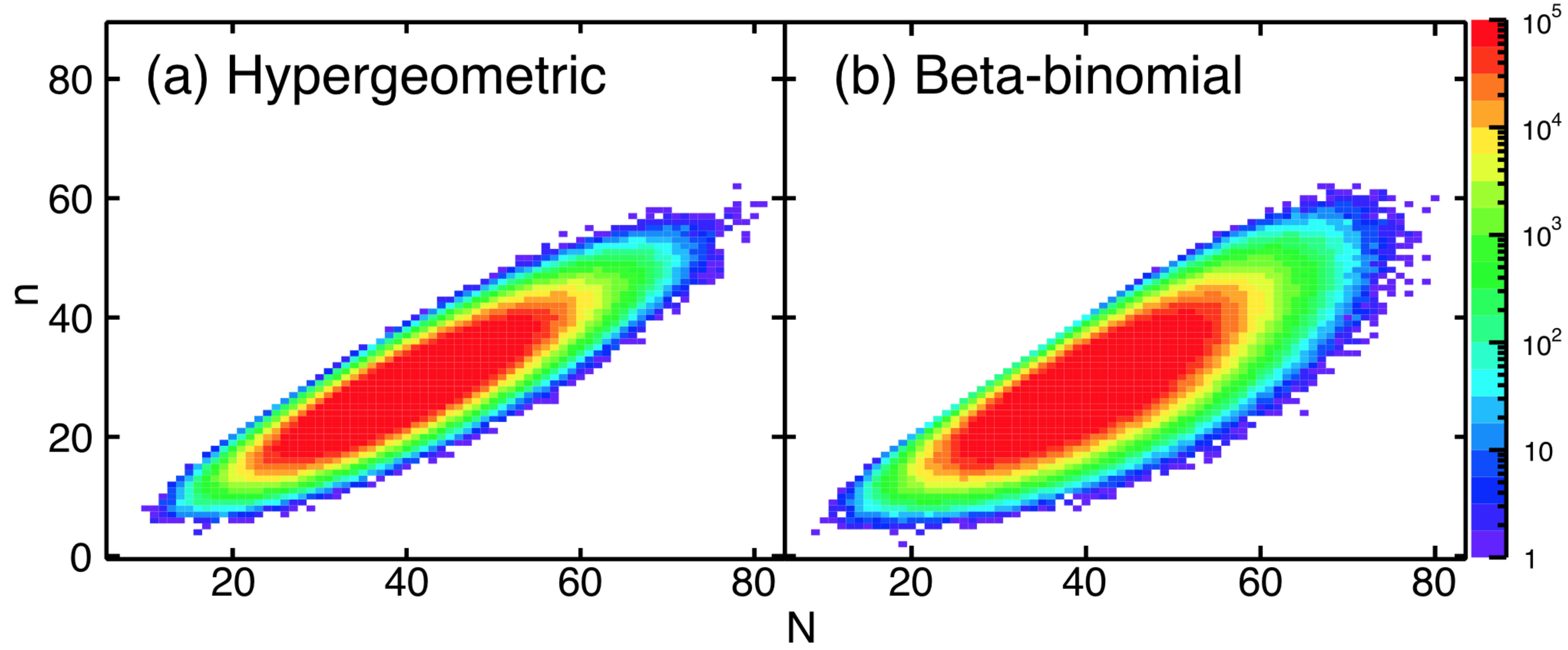
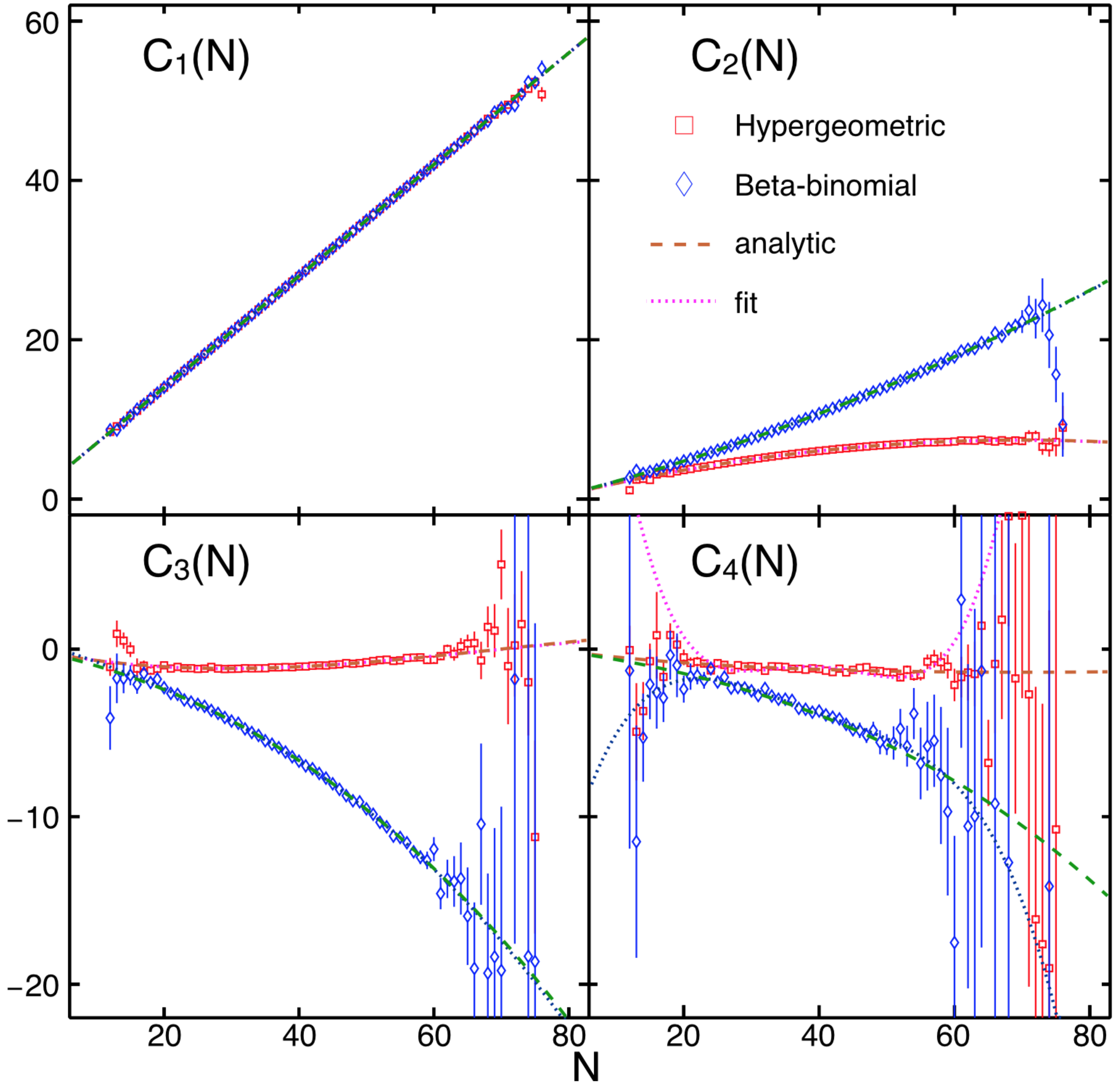


FIG. 1. Correlation between n and N on the sample events, i.e. the magnitude of $\mathcal{R}(n;N)P(N)$, for the response matrices $\mathcal{R}_{\rm HG}(n;N)$ (hypergeometric) and $\mathcal{R}_{\beta}(n;N)$ (beta-binomial) with $p = X/Y = 0.7$ and $Y = 140$.

Extract detector properties



- ✓ **Calculate m^th order moments of n at each N.**
  - Cumulants are shown here for better presentation.
- ✓ **Fit by m^th polynomial, to extract r^mj**

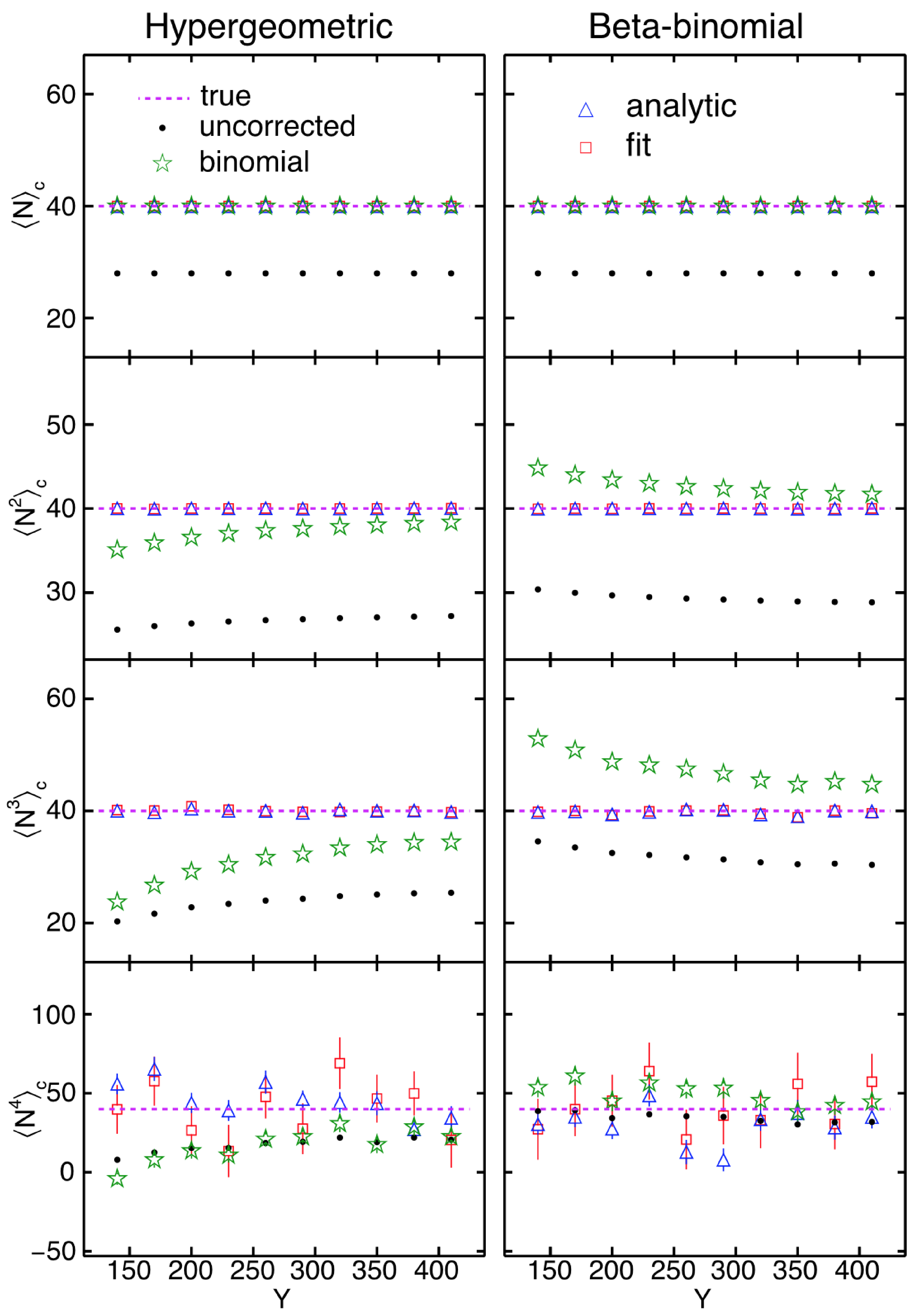
$$R_m(N) = \sum_{j=0}^{m} r_{mj} N^j. \qquad (10)$$

- ✓ **Solve the equations.**

$$\langle\!\langle n^m \rangle\!\rangle = r_{m0} + \sum_{j=1}^{\infty} r_{mj} \langle N^j \rangle. \qquad (7)$$

$$\begin{aligned}
\langle\!\langle n \rangle\!\rangle &= r_{10} + r_{11}\langle N \rangle \\
\langle\!\langle n^2 \rangle\!\rangle &= r_{20} + r_{21}\langle N \rangle + r_{22}\langle N^2 \rangle \\
\langle\!\langle n^3 \rangle\!\rangle &= r_{30} + r_{31}\langle N \rangle + r_{32}\langle N^2 \rangle + r_{33}\langle N^3 \rangle \\
\langle\!\langle n^4 \rangle\!\rangle &= r_{40} + r_{41}\langle N \rangle + r_{42}\langle N^2 \rangle + r_{43}\langle N^3 \rangle + r_{44}\langle N^4 \rangle
\end{aligned}$$

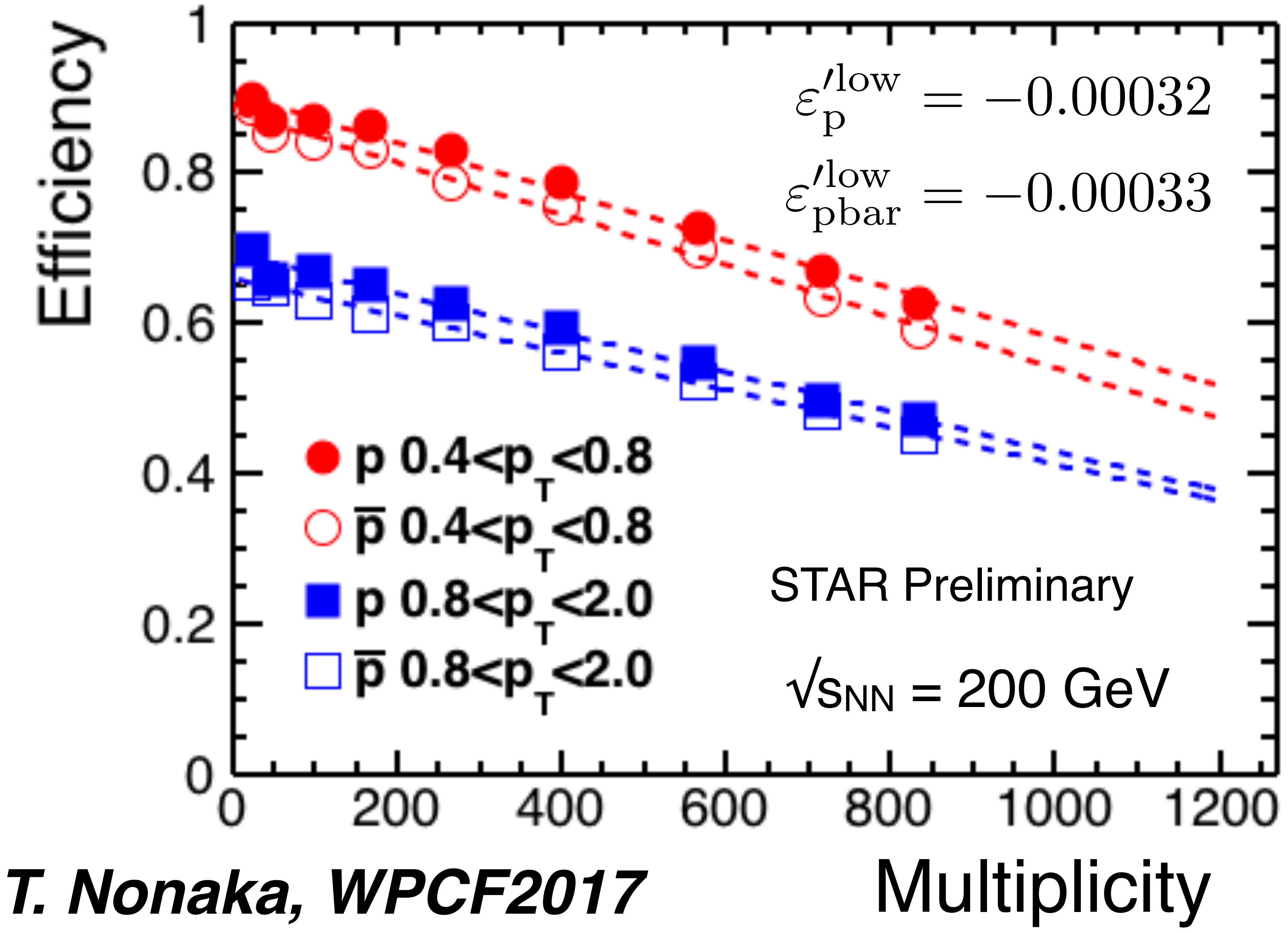
Results



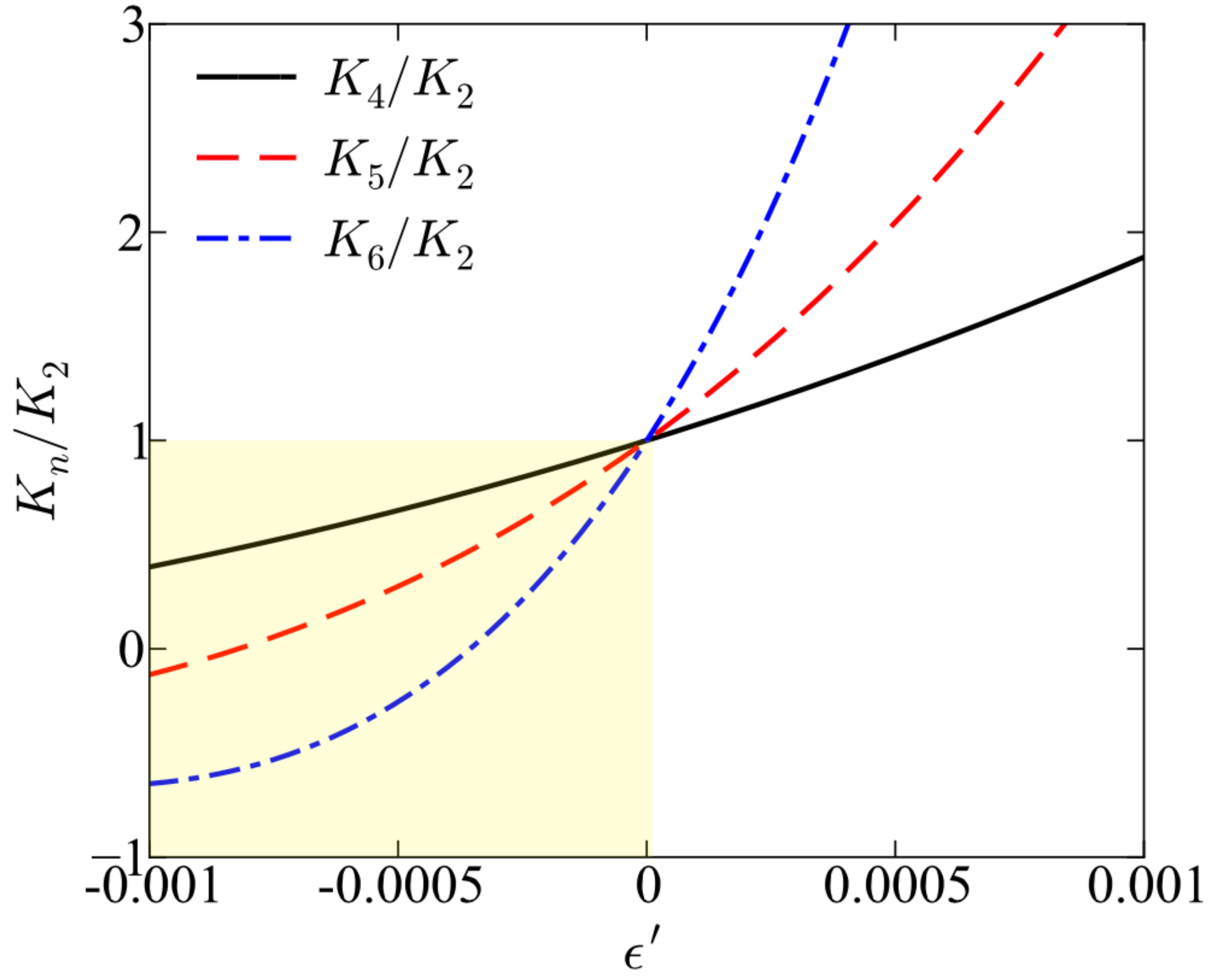
- ✓ **Check the correction as a function of different non-binomial parameter Y.**
- ✓ **Conventional correction assuming binomial efficiencies does not work.**
- ✓ **Our new method works well.**
- ✓ **Hypergeometric and beta-binomial examples can be also solved analytically.**

Multiplicity dependent efficiency

- ✓ **Efficiency decreases with increasing multiplicity, which artificially decreases the value of cumulants.**
- ✓ **Residual dependence of efficiency inside one multiplicity bin (for centrality) needs to be taken into account.**



T. Nonaka, WPCF2017



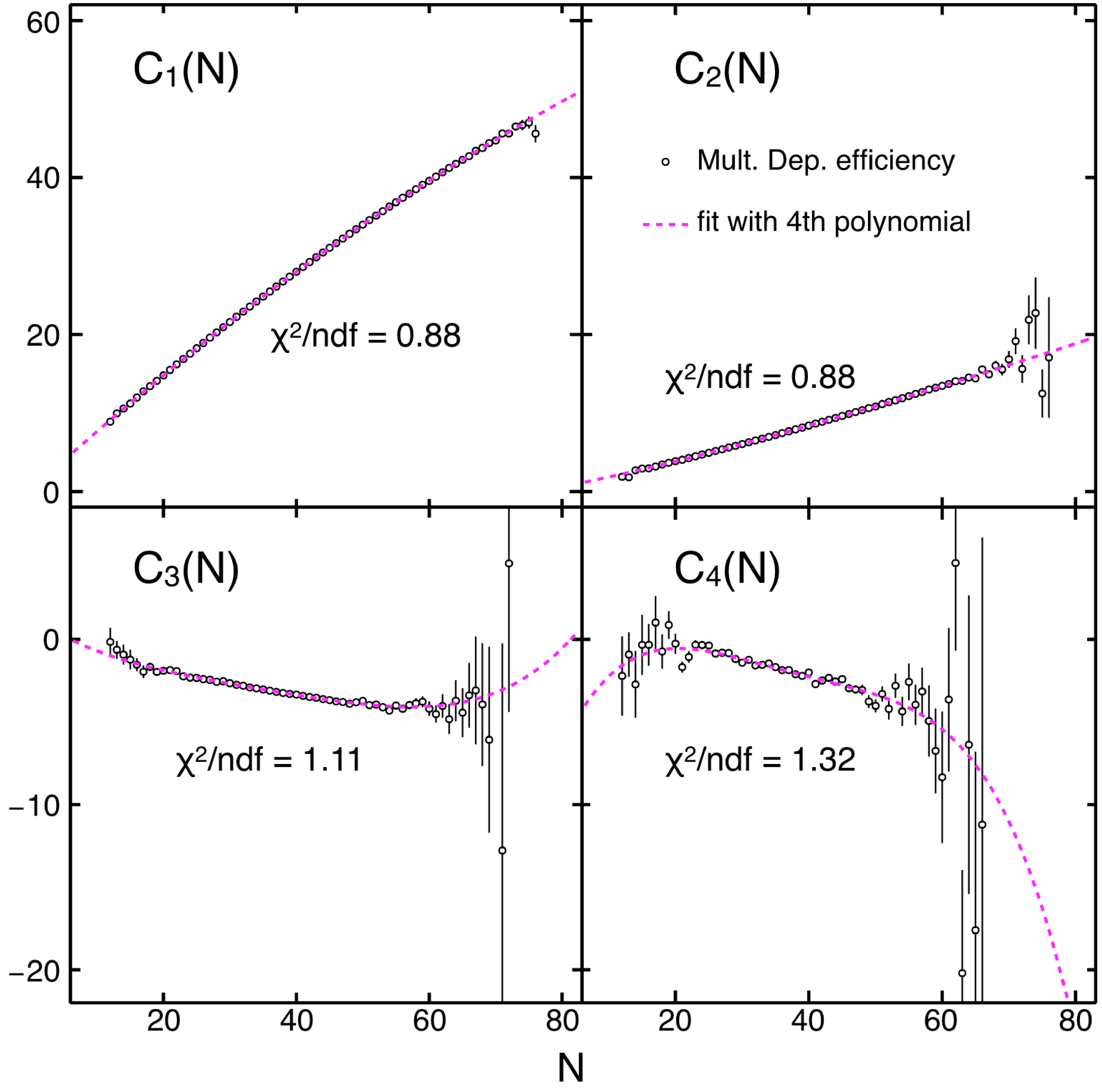
A. Bzdak, R. Holzmann, V. Koch : PRC.94.064907

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

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

Multiplicity dependent efficiency

$p(N) = 0.7 - \varepsilon(N - \langle N \rangle) \quad \varepsilon = 0.002$



- ✓ **$C_1(N) \propto N^2$, $C_2(N) \propto N^4$, $C_3(N) \propto N^6$……**
  - Cannot be exactly solved.
  - Truncation is necessary.
- ✓ **Fit by 4th polynomial, to extract r_{mj}**

L=4

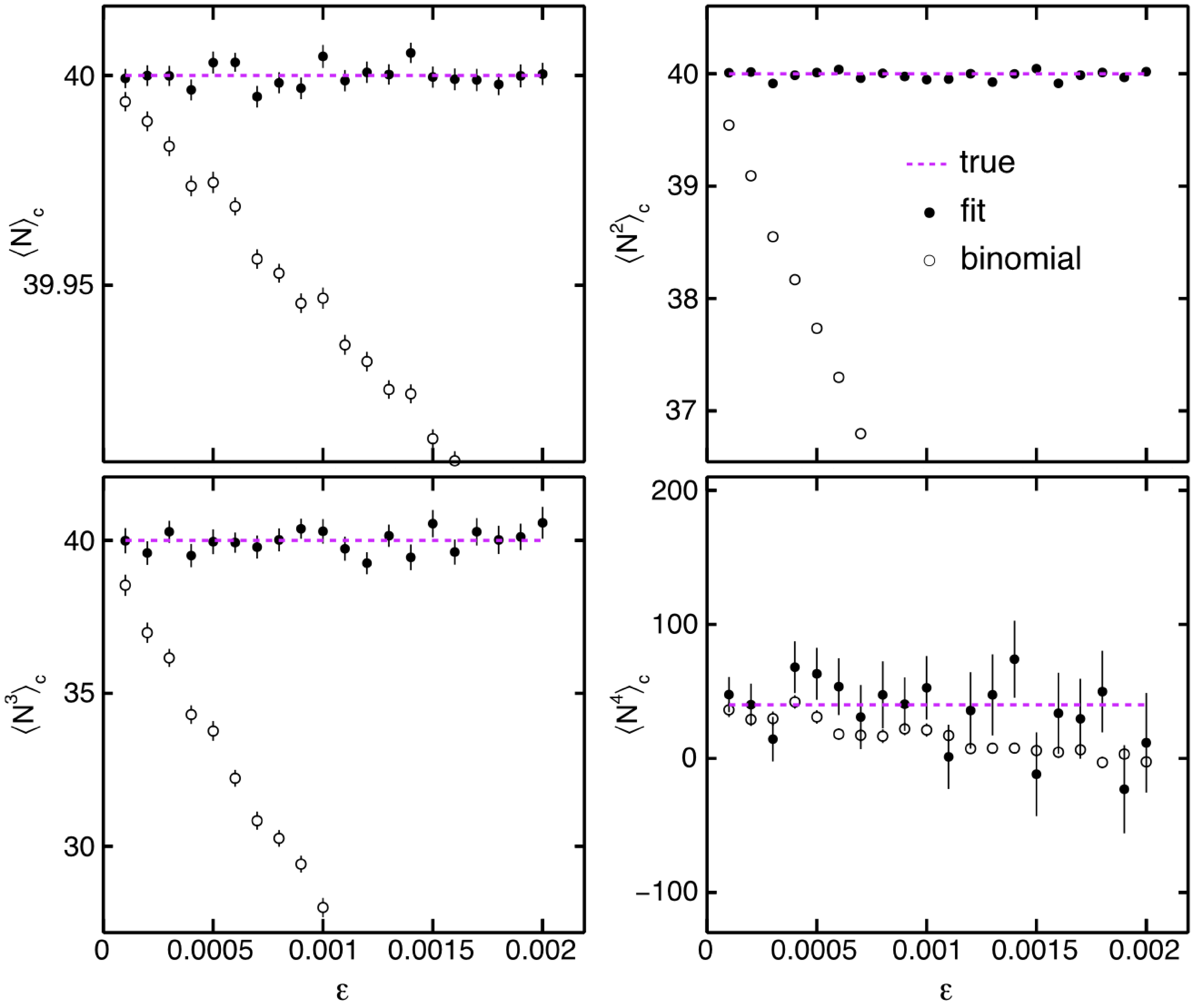
$$R_m(N) = \sum_{j=0}^{m} r_{mj} N^j. \qquad (10)$$

- ✓ **R will be L x L matrix**

$$\begin{bmatrix} \langle\langle n \rangle\rangle \\ \langle\langle n^2 \rangle\rangle \\ \vdots \\ \langle\langle n^L \rangle\rangle \end{bmatrix} = \begin{bmatrix} r_{10} \\ r_{20} \\ \vdots \\ r_{L0} \end{bmatrix} + \boldsymbol{R} \begin{bmatrix} \langle N \rangle \\ \langle N^2 \rangle \\ \vdots \\ \langle N^L \rangle \end{bmatrix}, \qquad (17)$$

Multiplicity dependent efficiency

$p(N) = 0.7 - \varepsilon(N - \langle N \rangle) \quad \varepsilon = 0.002$



- ✓ **Check the correction as a function of efficiency slope ε.**
- ✓ **Conventional method dose not work.**
- ✓ **Truncation at 4th order works well.**
- ✓ **With high statistics in MC, the trunction at 4th order would lose it's validity**
- ✓ **Validity of truncation can be checked by changing the truncation order like Lth, L+1th, L+2th order.**

Additional comments

- **In the case of net-particle distribution, 2-D fitting is necessary.**

- **We assumed $P_{MC}(N) = P(N)$, which is not the case in the real experiment, since we don't know true P(N).**

  - Correct results were obtained by using a Gauss distribution C_1=40 and C_2=√40 for $P_{MC}(N)$ with respect to the Poisson distribution for P(N) with C_1=40.

- **We should try several models to determine *R*(n;N), which should be included as systematic uncertainties**

  - HIJING+GEANT?? Or embedding simulation??

- **It would be better if data driven method will be developed in the future.**

Summary

- General procedures to correct non-binomial detector effects were presented.

- Truncation will be necessary if efficiency depends on the multiplicity.

- Systematic studies using different models would be needed.

Thank you!

Back up

Cumulants of binomial distribution

✓ **Cumulants of binomial distribution is proportional to N.**

M. Asakawa and M. Kitazawa, Prog. Part. Nucl. Phys. 90, 299 (2016), arXiv:1512.05038

$$B_{p,N}(m) = {}_N C_m p^m (1-p)^{N-m}, \qquad {}_N C_m = \frac{N!}{m!(N-m)!},$$

$$\langle m^n \rangle_c = \xi_n N,$$

$$\xi_1 = p, \quad \xi_2 = p(1-p), \quad \xi_3 = p(1-p)(1-2p), \quad \xi_4 = p(1-p)(1-6p+6p^2).$$

Cumulants of non-binomial distributions

$$\langle n \rangle_{\rm c}^{\rm HG} = Np, \tag{D4}$$

$$\langle n^2 \rangle_{\rm c}^{\rm HG} = \frac{N(N-Y)(-1+p)p}{-1+Y}, \tag{D5}$$

$$\langle n^3 \rangle_{\rm c}^{\rm HG} = \frac{N(2N^2-3NY+Y^2)(-1+p)p(-1+2p)}{(-2+Y)(-1+Y)}, \tag{D6}$$

$$\begin{aligned}\langle n^4 \rangle_{\rm c}^{\rm HG} = &[N(N-Y)(-1+p)p(6N^2(-1-6(-1+p)p+Y(1+5(-1+p)p)) \\ &-6NY(-1-6(-1+p)p+Y(1+5(-1+p)p)) \\ &+(-1+Y)Y(1+Y(1+6(-1+p)p)))]/\left[(-3+Y)(-2+Y)(-1+Y)^2\right],\end{aligned} \tag{D7}$$

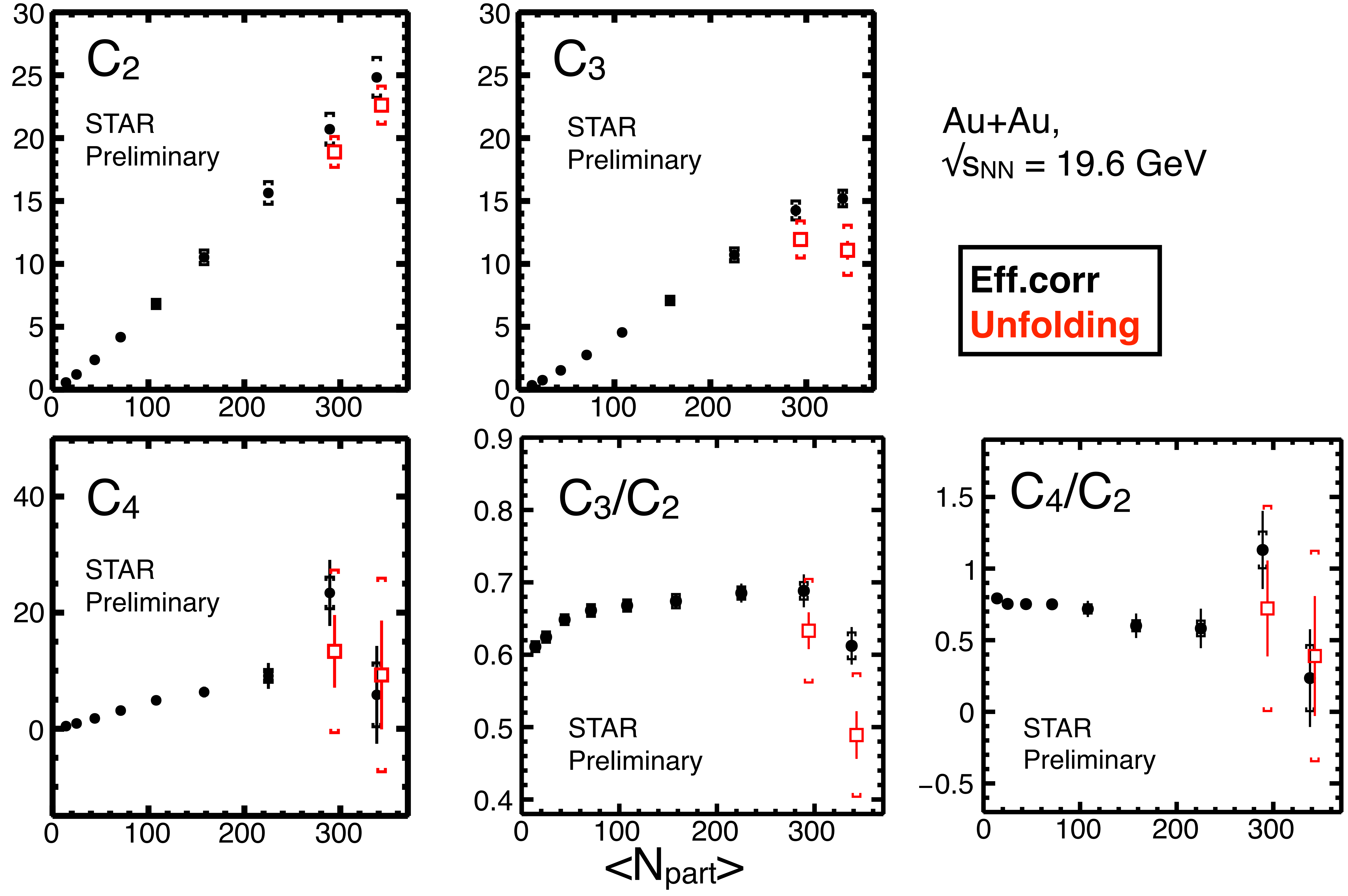
$$\langle n \rangle_{\rm c}^{\beta} = Np, \tag{D8}$$

$$\langle n \rangle_{\rm c}^{\beta} = -\frac{N(N+a+b)(-1+p)p}{1+a+b}, \tag{D9}$$

$$\langle n \rangle_{\rm c}^{\beta} = \frac{N(N+a+b)(2N+a+b)(-1+p)p(-1+2p)}{(1+a+b)(2+a+b)}, \tag{D10}$$

$$\begin{aligned}\langle n \rangle_{\rm c}^{\beta} = &-((N(N+a+b)(-1+p)p(6N^2(1+6(-1+p)p+(a+b)(1+5(-1+p)p)) \\ &+6N(a+b)(1+6(-1+p)p+(a+b)(1+5(-1+p)p))+(a+b)(1+a+b) \\ &(-1+(a+b)(1+6(-1+p)p))))/((1+a+b)^2(2+a+b)(3+a+b))).\end{aligned} \tag{D11}$$

Non-binomial efficiency



QM2018, T. Nonaka (STAR)