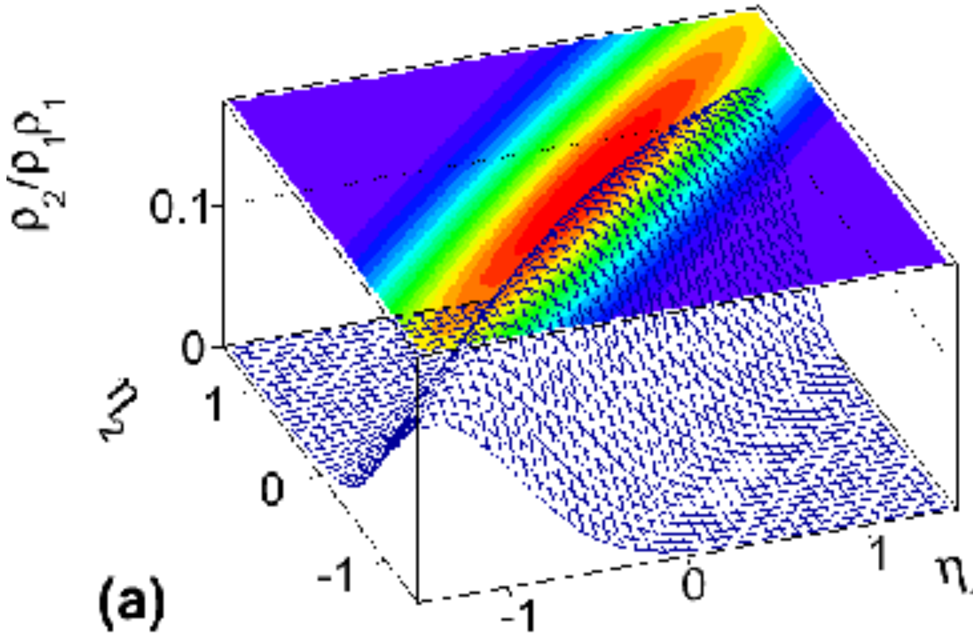
Wayne State University
College of Liberal Arts & Sciences
Department of Physics and Astronomy

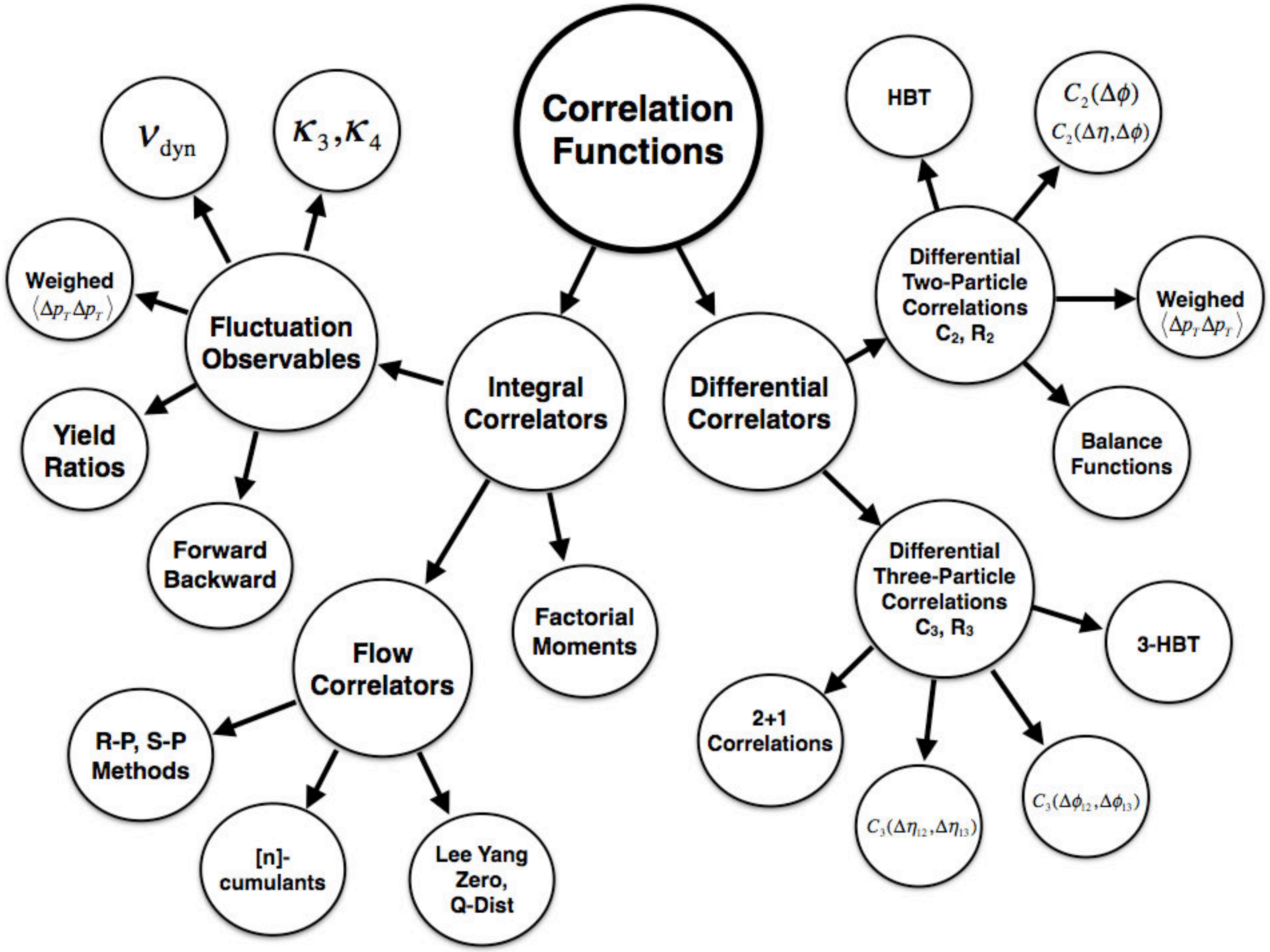


Correlation Functions
An experimental and technical perspective

Lecture at University of Peking
Oct 2017

Claude A. Pruneau
Wayne State University

Goal of this (technical) talk:



Provide you with a basis to understand …

- the notion of correlation function
- the link between integral and differential correlation functions.
- how to measure them i.e., how to correct for instrumental effects.

And much more …

Outline

- Part I: What is a correlation function?
  - Correlation functions as covariance.
- Part II: Correlation Function Formal Definition
  - Integral and Differential Correlation Functions
  - The Multiple facets of Correlation Functions
  - Moments, Cumulants, Factorial Moments, Factorial Cumulants
- Part III: Why Measure Differential Correlation Functions?
  - Emphasis on Cumulants
- Part IV: Multi-Facets of Correlation Functions
- Part V: Experimental Considerations
  - Acceptance
  - Efficiency
  - Other instrumental effects

Part 1: What's a correlation function?

- **Definition of Correlation Functions** as an **extension of the notion of covariance**.
  - Introduction based of two-particle cross-sections.
  - Could be formulated in more general terms as generic functions describing fields, yields, intensity, in multi-dimensional spaces.
- We will formally introduce correlation functions based on cumulants of cross-sections during the next segment.

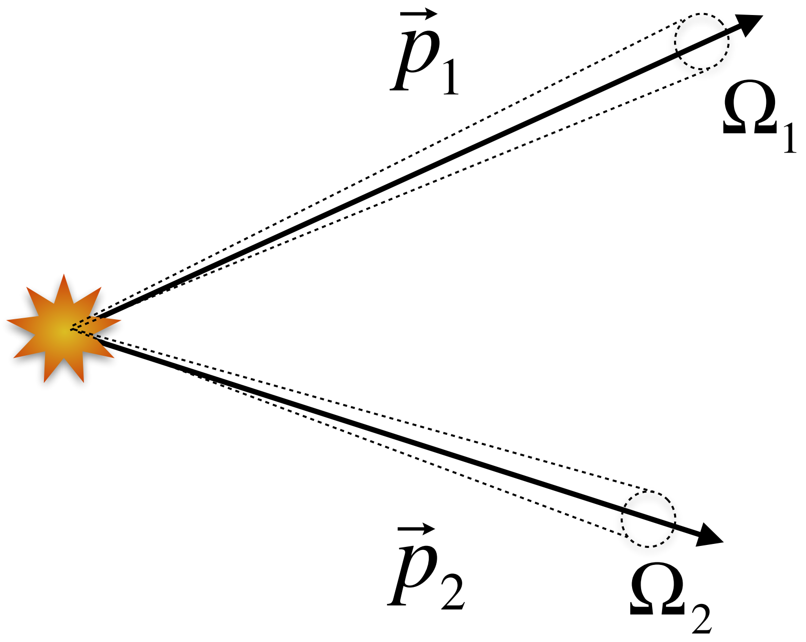
Setup

- Consider a measurement of the number of particles produced at two distinct momenta $\vec{p}_1$ and $\vec{p}_2$
- Let N_i represent the number of particles produced in volumes $\mathbf{\Omega}_i$, i=1, 2, in ranges "centered" on $\vec{p}_1$ and $\vec{p}_2$

$$p_{\mathrm{T},i}^{\min} \le p_{\mathrm{T},i} \le p_{\mathrm{T},i}^{\max}$$

$$\eta_i^{\min} \le \eta_i \le \eta_i^{\max}$$

$$\phi_i^{\min} \le \phi_i \le \phi_i^{\max}$$



Average Yields

- Given the stochastic nature of particle production, the yields N_i are expected to fluctuate event-by-event — *even for identical collision parameters*.
- For a given type of particle, collision, etc, one can consider the averages $\langle N_i \rangle$
- These averages are determined by the particle production cross-section of the specific process considered:

$$\langle N_i \rangle = \int_{\Omega_i} \frac{d^3 N_i}{dp_T d\phi d\eta} dp_T \, d\phi \, d\eta$$

- Bracket notation <O> used to denote ensemble/population average.

RMS Yield and Covariance

- Fluctuations characterized by the variance of N_i:

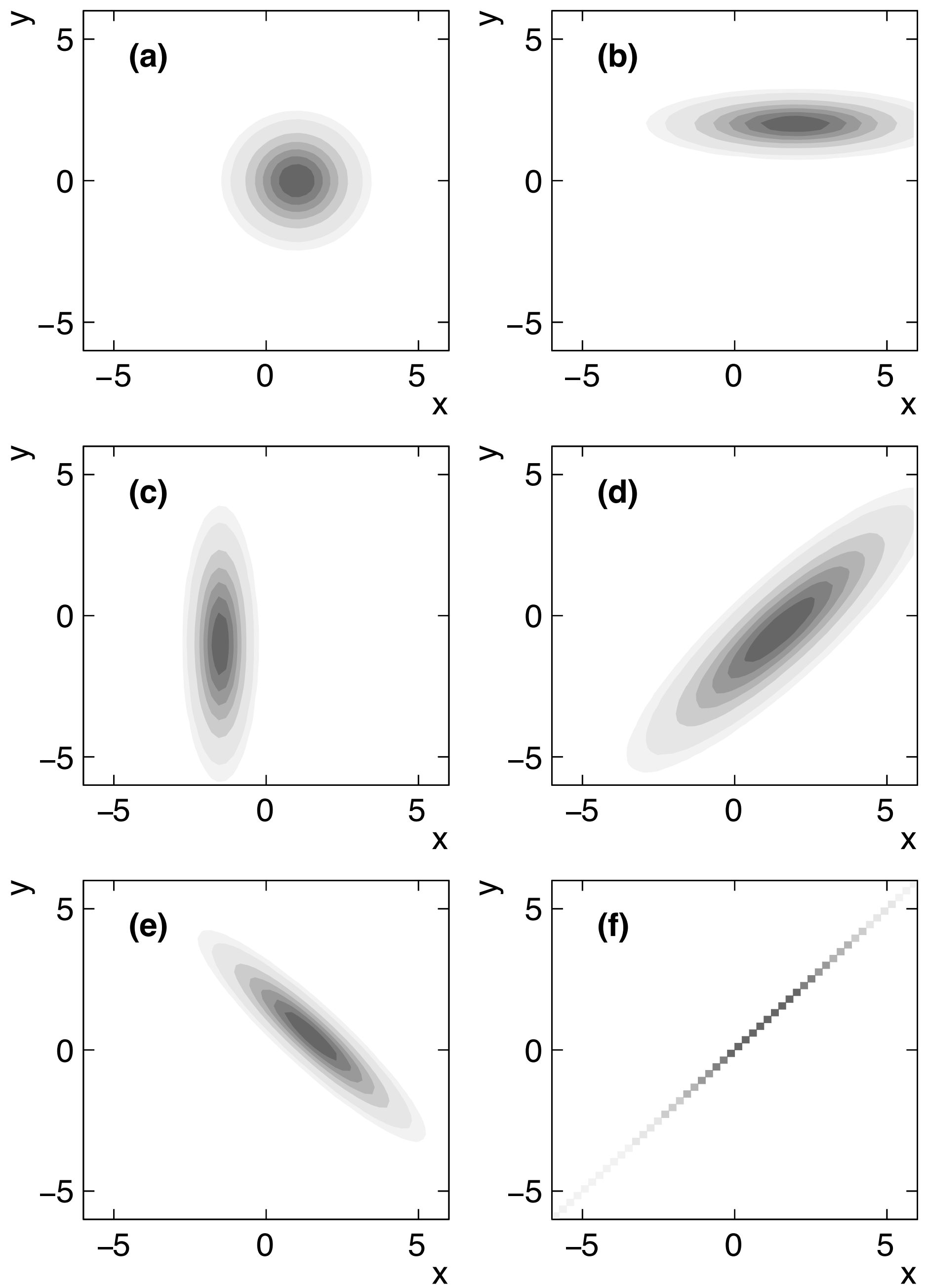
$$Var[N_i] = \langle N_i^2 \rangle - \langle N_i \rangle^2$$

- More informative to study the covariance of these two yields

$$Cov[N_1, N_2] = \langle N_1 N_2 \rangle - \langle N_1 \rangle \langle N_2 \rangle$$

- $Cov[N_1,N_2]$ depends on the size of the bins Ω_1 and Ω_2 used to measure the yields N_1 and N_2, respectively,
- Also a function of the coordinates $\vec{p}_1$ and $\vec{p}_2$ at which the particle emission is considered.

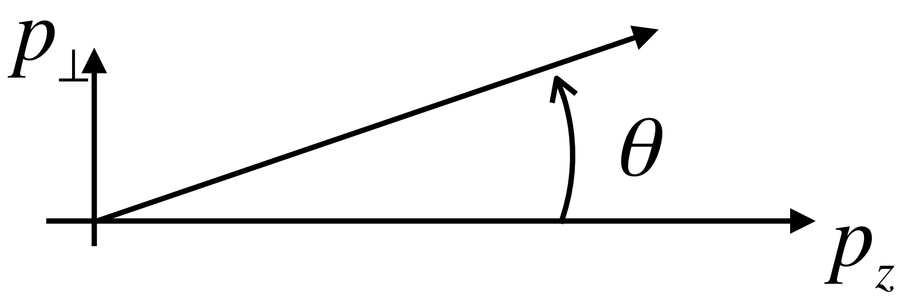
$$Cov[x,y] = \langle xy \rangle - \langle x \rangle \langle y \rangle$$

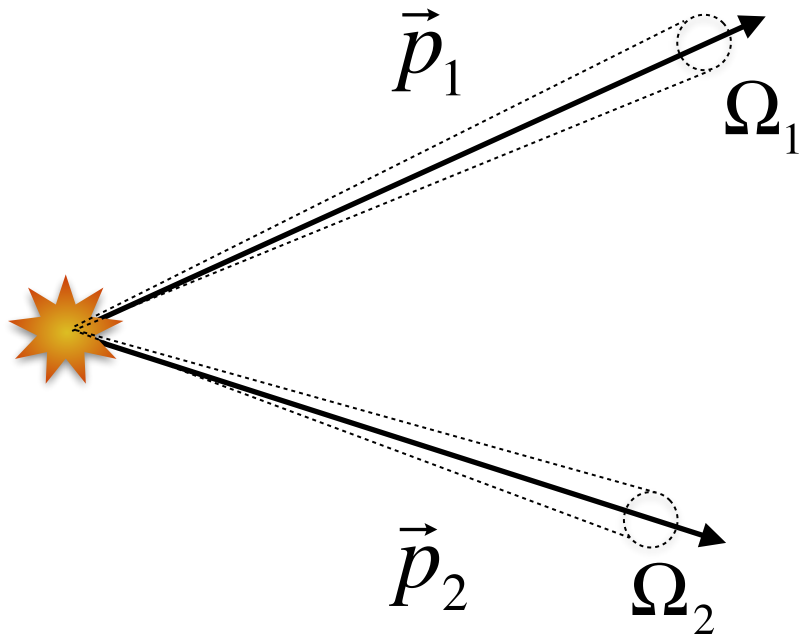


Pair Yield Covariance

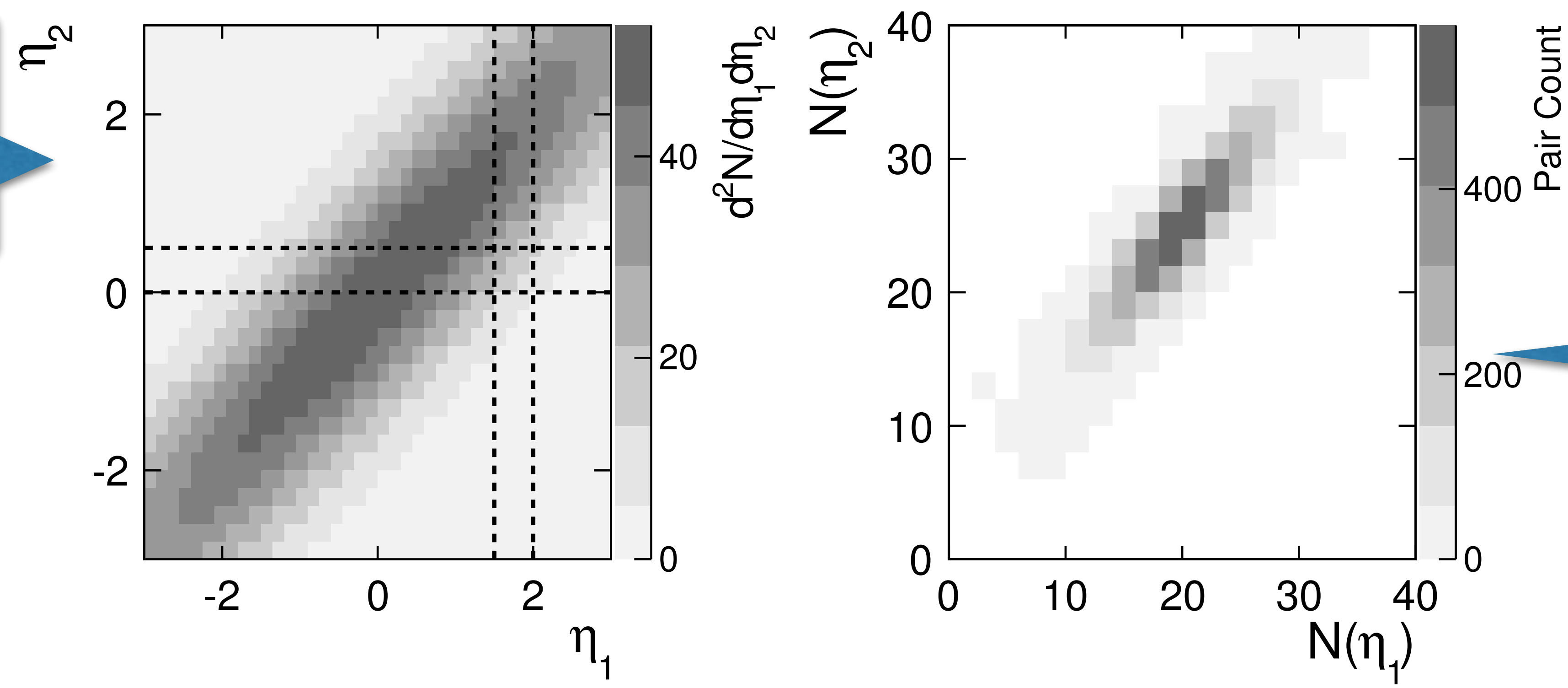
Pseudorapidity

η = –ln (tan(θ/2)) "eta"





Correlation Function



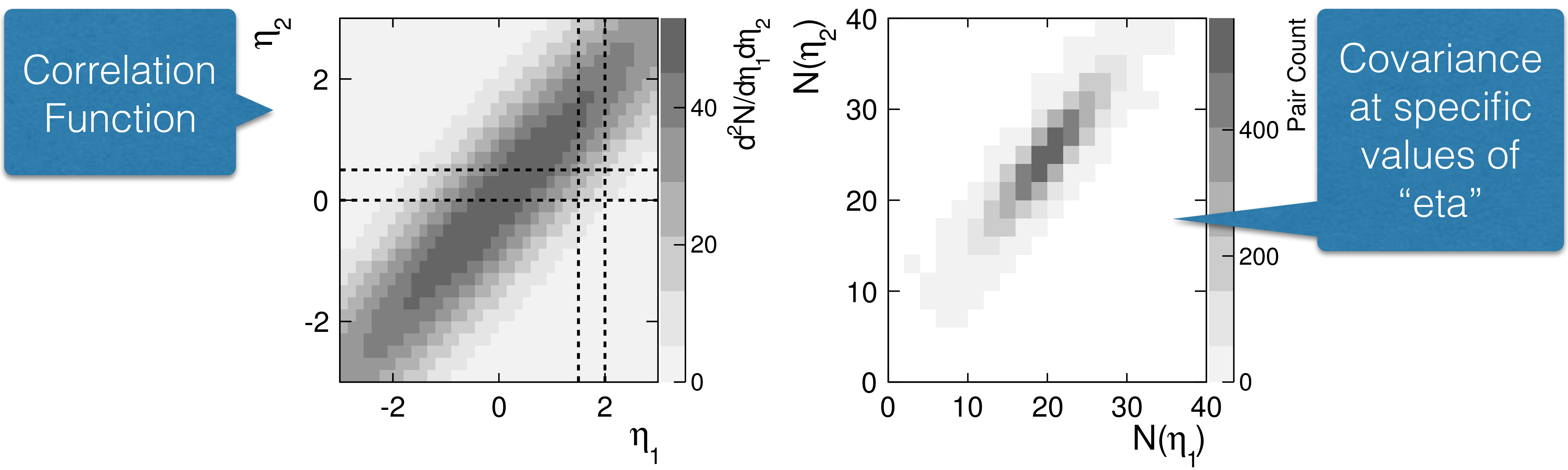
Covariance at specific values of "eta"

Correlation function

- Natural to introduce the notion of correlation function at $\vec{p}_1$ and $\vec{p}_2$ based on

$$C(\vec{p}_1,\vec{p}_2)=\frac{1}{\Omega_1\Omega_2}\left[\langle N(\vec{p}_1)N(\vec{p}_2)\rangle-\langle N(\vec{p}_1)\rangle\langle N(\vec{p}_2)\rangle\right]$$

- Defined in the limit in which the bin sizes Ω_1 and Ω_2 vanish.



Wayne State University
College of Liberal Arts & Sciences
Department of Physics and Astronomy

- Note the sight change in notation.

Single Particle Density Estimator

- The average yield $\langle N(\vec{p}_i)\rangle$ normalized by the bin size Ω_i constitutes an **estimator** of the (single) **particle density** at $\vec{p}_i$.

$$\hat{\rho}_1(\vec{p}_i) = \frac{\langle N(\vec{p}_i)\rangle}{\Omega_i}$$

- In the limit $\Omega_i \rightarrow 0$ and infinite statistics, one gets the single particle cross-section:

$$\lim_{\Omega_i \to 0} \hat{\rho}_1(\vec{p}_i) = \rho_1(\vec{p}_i) = \frac{d^3 N_i}{dp_T d\phi d\eta}(\vec{p}_i)$$

Two-Particle Density

- The average yield $\langle N(\vec{p}_1)N(\vec{p}_2)\rangle$ normalized by the product of bin sizes, $\Omega_1 \times \Omega_2$, constitutes an estimator of the joint- or two-particle density at $\vec{p}_1$ and $\vec{p}_2$.

$$\hat{\rho}_2(\vec{p}_1,\vec{p}_2) \equiv \frac{\langle N(\vec{p}_1)N(\vec{p}_2)\rangle}{\Omega_1\Omega_2}$$

joint pair density (estimator)

- In the limit $\Omega_i \longrightarrow 0$ and infinite statistics

$$\lim_{\Omega_i \to 0} \hat{\rho}_2(\vec{p}_1,\vec{p}_2) = \rho_2(\vec{p}_1,\vec{p}_2) = \frac{d^6 N_{pairs}}{dp_{T,1}d\phi_1 d\eta_1 dp_{T,2}d\phi_2 d\eta_2}(\vec{p}_1,\vec{p}_2)$$

Correlation Function Definition

- In the limit $\Omega_1,\Omega_2 \longrightarrow 0$, one has a **correlation function**

$$C(\vec{p}_1,\vec{p}_2) = \rho_2(\vec{p}_1,\vec{p}_2) - \rho_1(\vec{p}_1)\rho_1(\vec{p}_2)$$

- which is the "**most general**" form a two-particle correlation function (i.e., **6 momentum components**)
- choice of coordinate representation is somewhat arbitrary
  - cartesians: p_{1x}, p_{1y}, p_{1z}, p_{2x}, p_{2y}, p_{2z}
  - rapidity: y_1, ϕ_1, p_{1T}, y_2, ϕ_2, p_{2T}
  - pseudorapidity: η_1, ϕ_1, p_{1T}, η_2, ϕ_2, p_{2T}
  - etc.

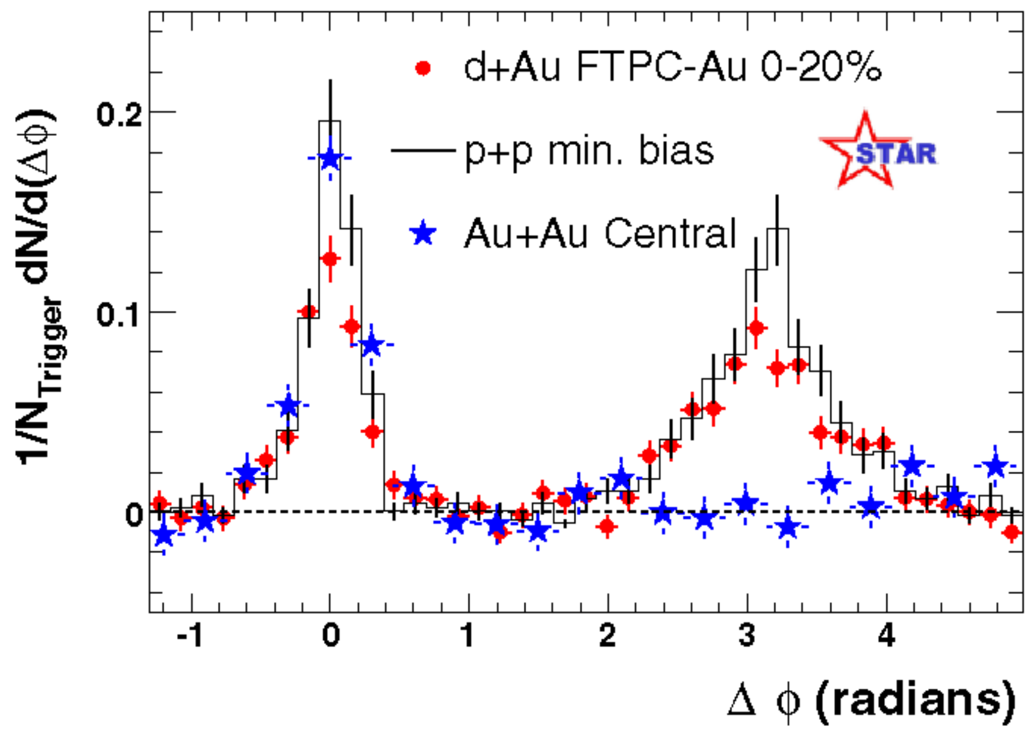
Parameter Marginalization

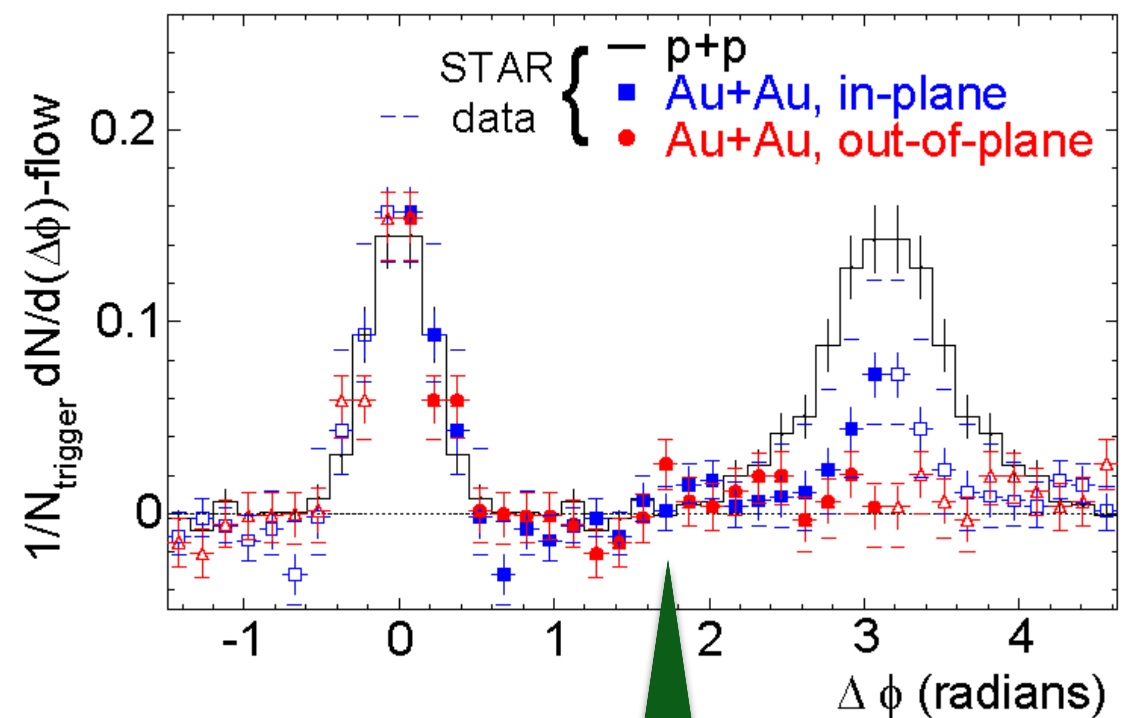
- A measurement of correlation function can be reduced to a smaller number of coordinates of interest by **integrating, or averaging,** called **marginalization by statisticians,** over variables that are not of interest.
- Common to study correlation functions of produced particles as a function of
  - the relative angle $\Delta\phi = \phi_1 - \phi_2$, or
  - the difference in pseudorapidity $\Delta\eta = \eta_1 - \eta_2$,
  - or both,
  - for specific types of particles (e.g., all charge hadrons, positive particles only, or only pions, etc.), and within a specific range of transverse momentum, and for events (i.e., collisions) satisfying specific conditions.

Example:

STAR, White Paper, Nuclear Physics A 757 (2005) 102–183

Jet Quenching Discovery





"background" subtracted (AuAu)

Important Remarks

- **Particle yields** are by definition **non-negative** (i.e., positive or null),
- But the function $C(\vec{p}1,\vec{p}2)$ may be **positive**, **null**, or **even negative**.
- As for covariances, a positive value indicates that a **rise** of the particle yield at $\vec{p}_1$ is, on average, *accompanied by* a **rise** of the yield at $\vec{p}_2$. The yields are said to be **correlated**.
- A negative value corresponds to an **anti-correlation**, so that the **rise** of the yield at one momentum is accompanied by a **decline** at the other momentum.
- A null value, of course, implies that the two yields, at the given momenta $\vec{p}_1$ and $\vec{p}_2$, are seemingly independent.

$$C(\vec{p}_1,\vec{p}_2)=0 \quad \Rightarrow \quad \rho_2(\vec{p}_1,\vec{p}_2)=\rho_1(\vec{p}_1)\rho_1(\vec{p}_2)$$

Is this condition sufficient to conclude the production at the two momenta is statistically independent?

Part 1: Summary

- Used the yields N_1 and N_2 of particle production at two momenta $\vec{p}_1$ and $\vec{p}_2$ in solid angles Ω_1 and Ω_2.
- Considered the covariance of N_1 and N_2.
- Showed that in the limit Ω_1 —> 0, the covariance defines a function of $\vec{p}_1$ and $\vec{p}_2$ which expresses the covariance of the pair density at these momenta.
- This function is called **correlation function** of the pair yield vs. $\vec{p}_1$ and $\vec{p}_2$
- The correlation function can be **marginalized against several of its variables**.

Part II: Formal Definition of Corr Fct

- Goal:
  - Obtain tools to determine whether detected particles are correlated.
- Define
  - probability density of particle emission.
  - number densities.
  - factorial moments.
  - cumulants.
- Derive formula usable towards the extraction of cumulants from measured densities.

Where are correlations from?

- Conservation Laws
  - Energy, Momentum
  - Quantum Numbers
    - Charge, Strangeness, Baryon Number
- Geometry (System Shape)
  - Opacity
  - Thermal Motion (Decays)
  - Pressure Gradients (e.g. radial flow, anisotropic flow)

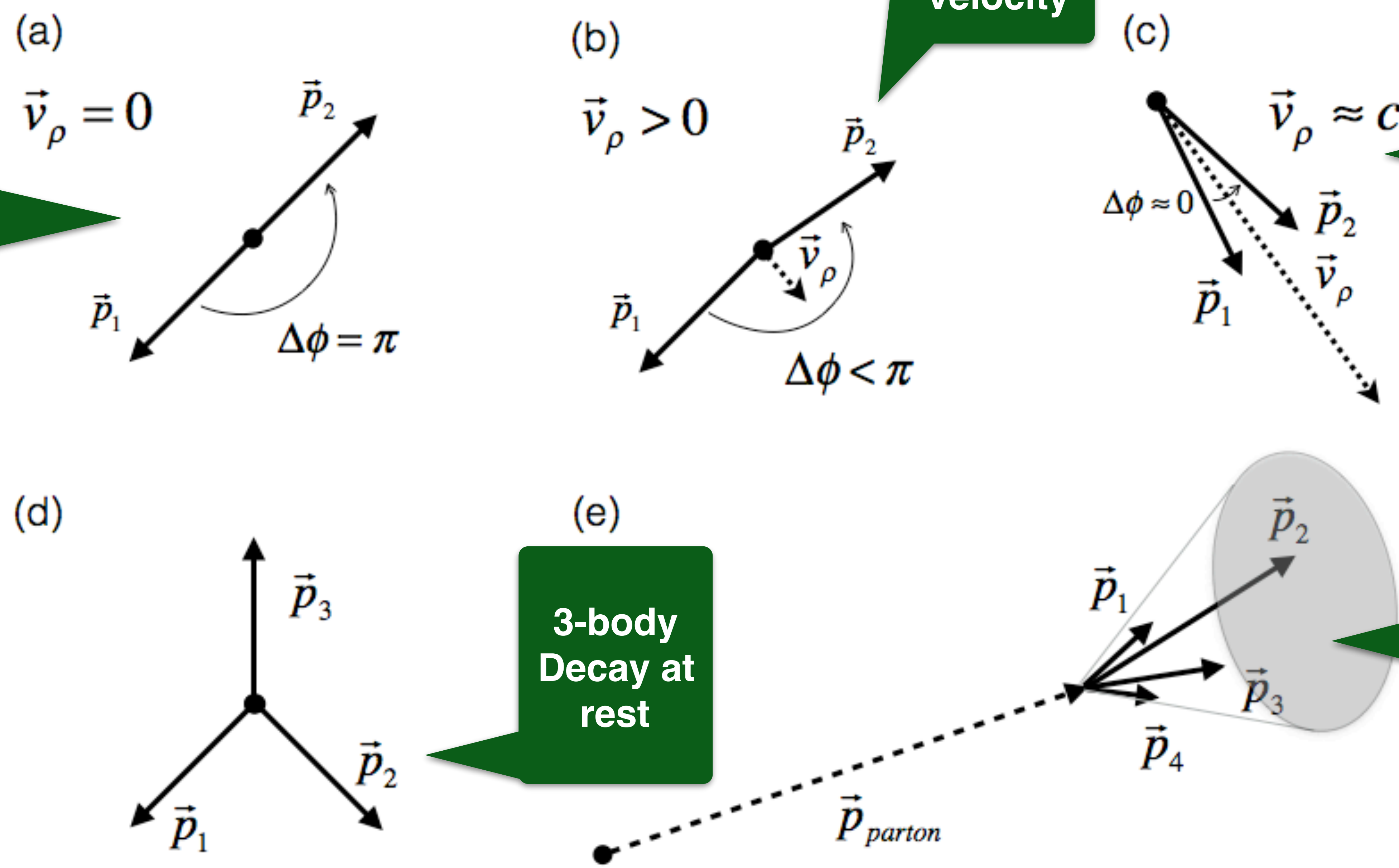
What's the cause of correlations?

Energy Momentum Conservation
e.g., Resonance Decays, Jets

2-body Decay at rest in the lab!

2-body Decay at small velocity

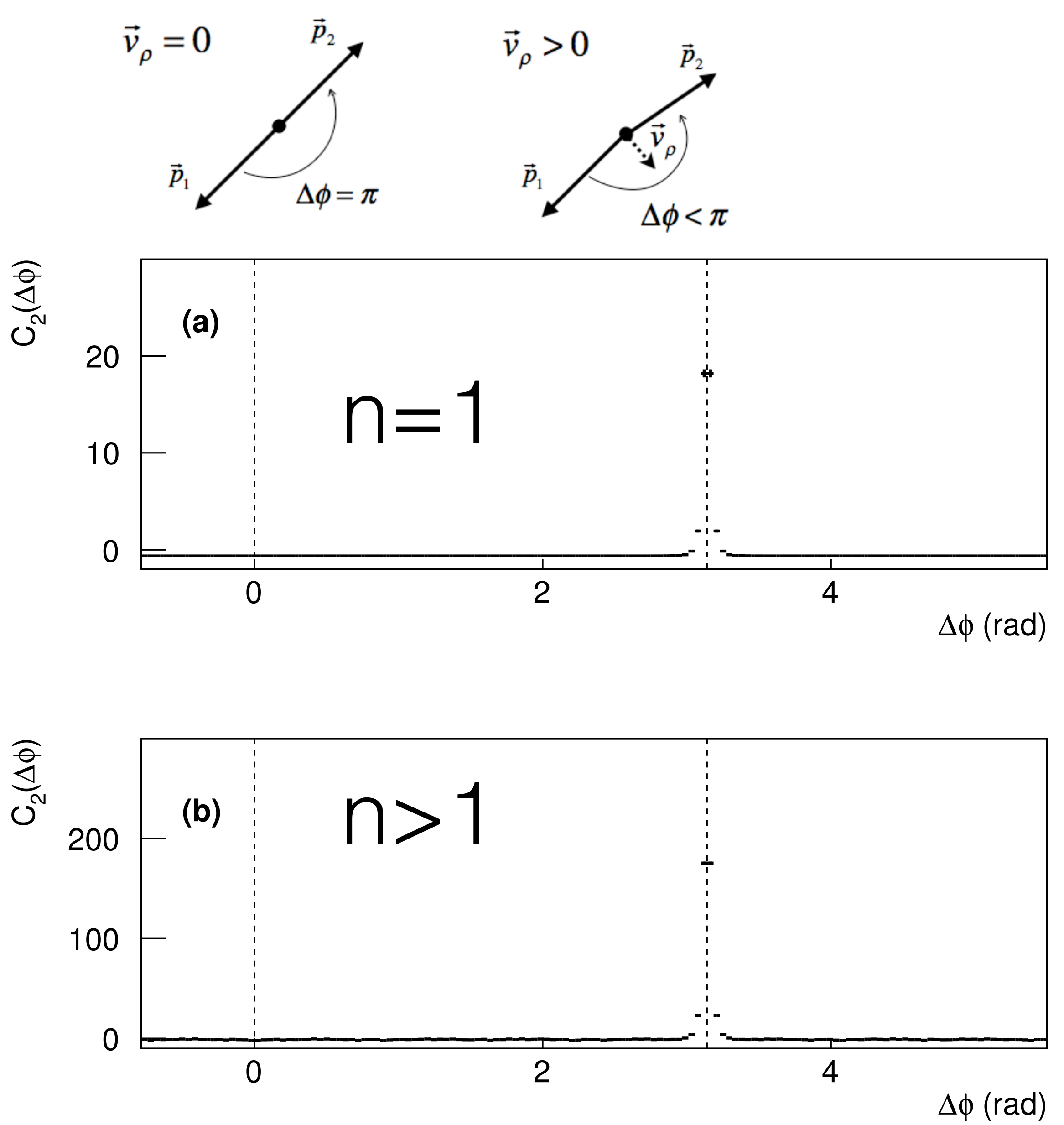
2-body Decay at high velocity

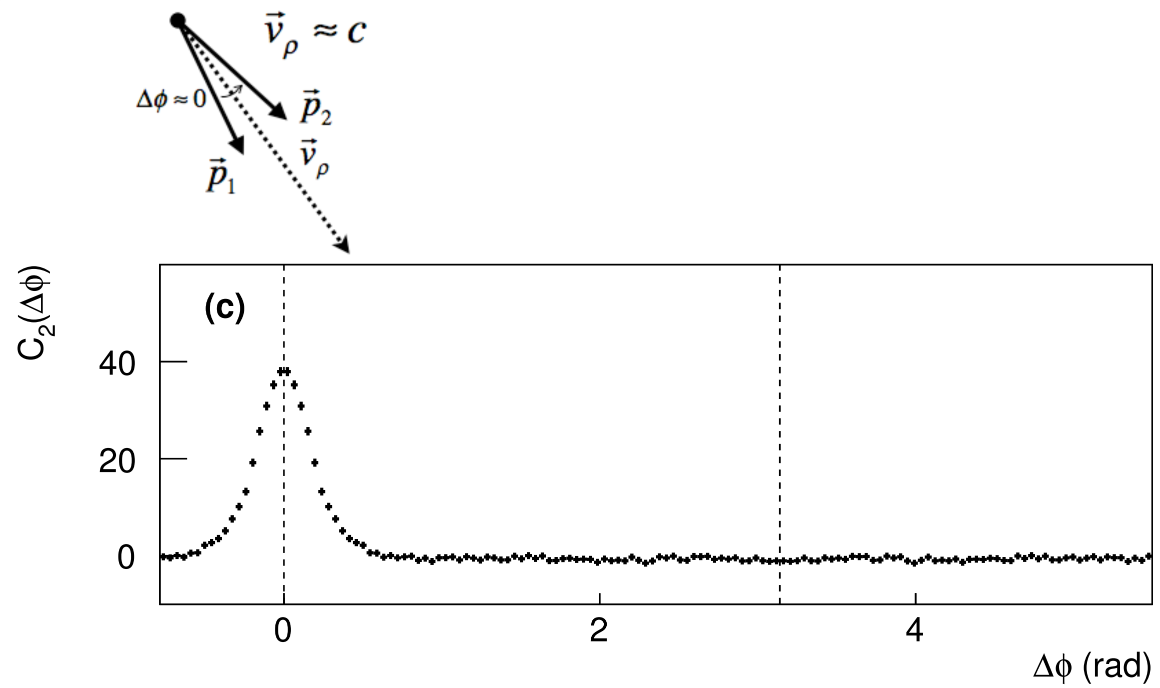


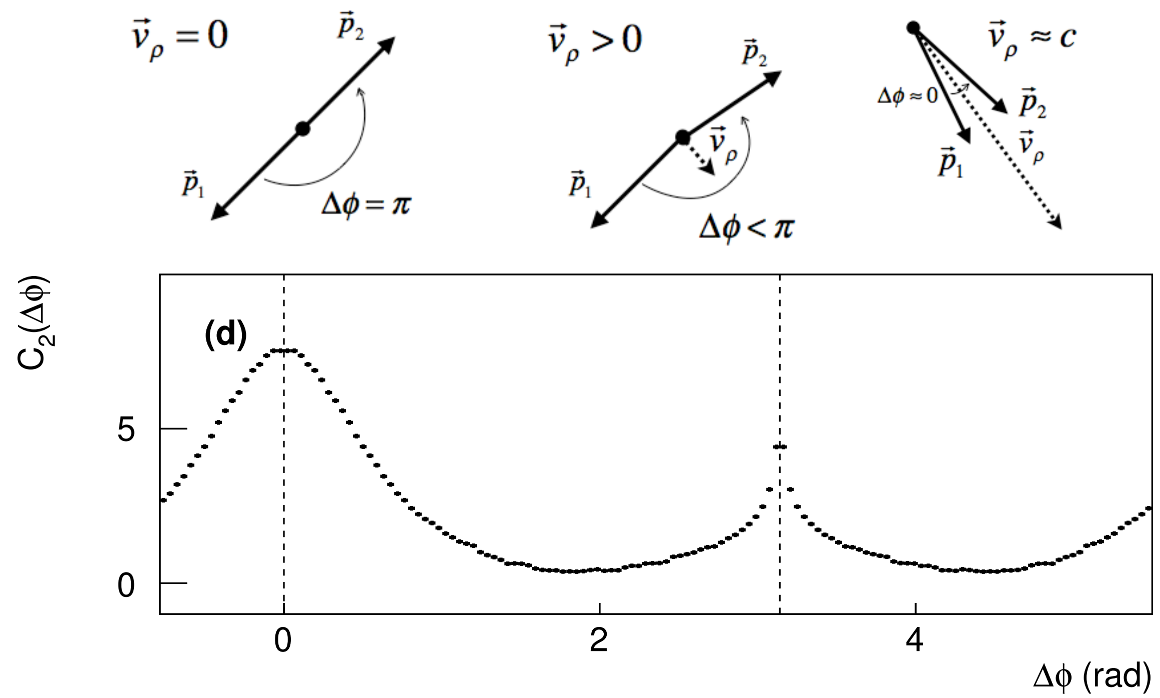
3-body Decay at rest

Jet

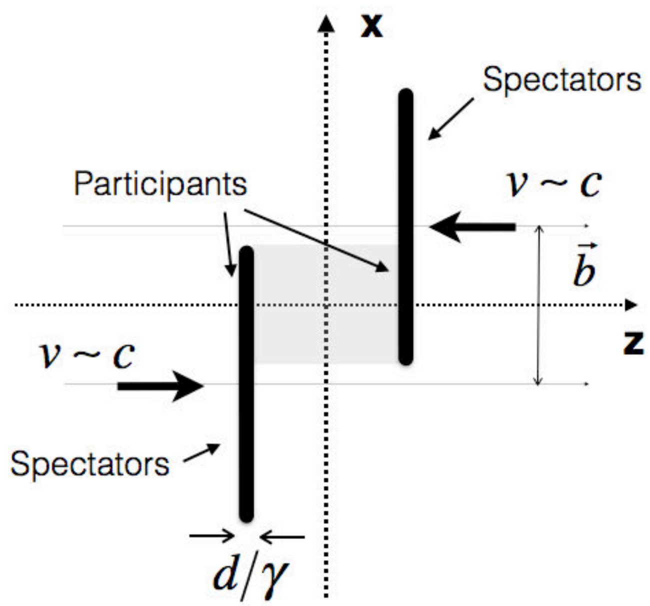
Resonance Decay: An Example

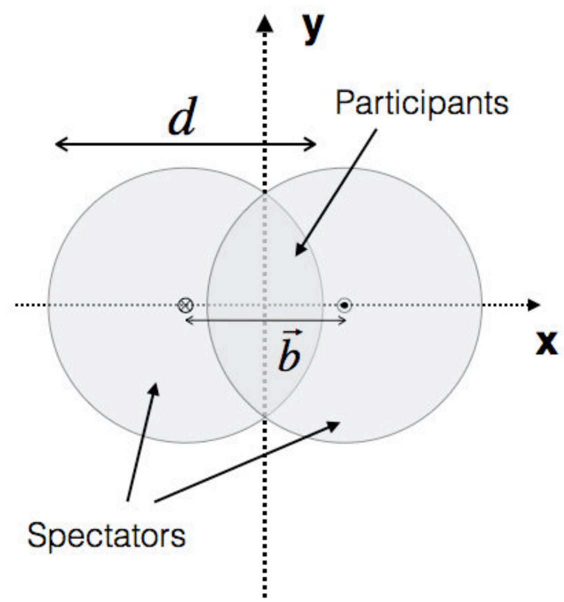


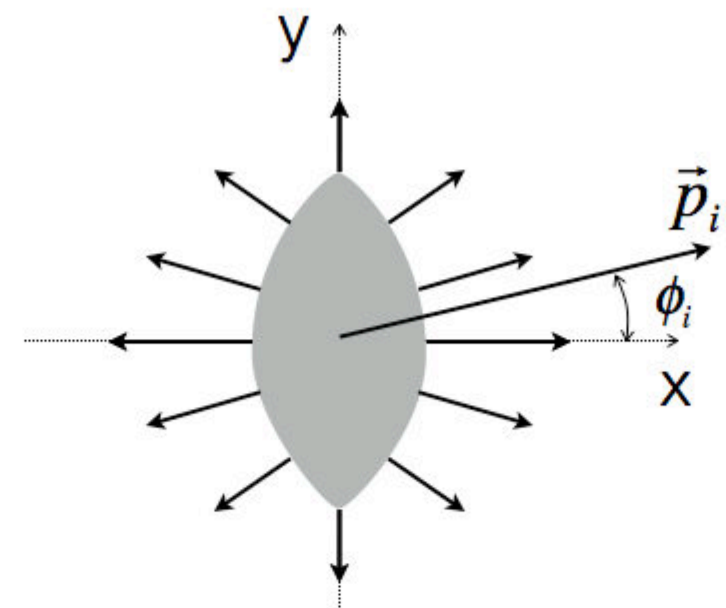


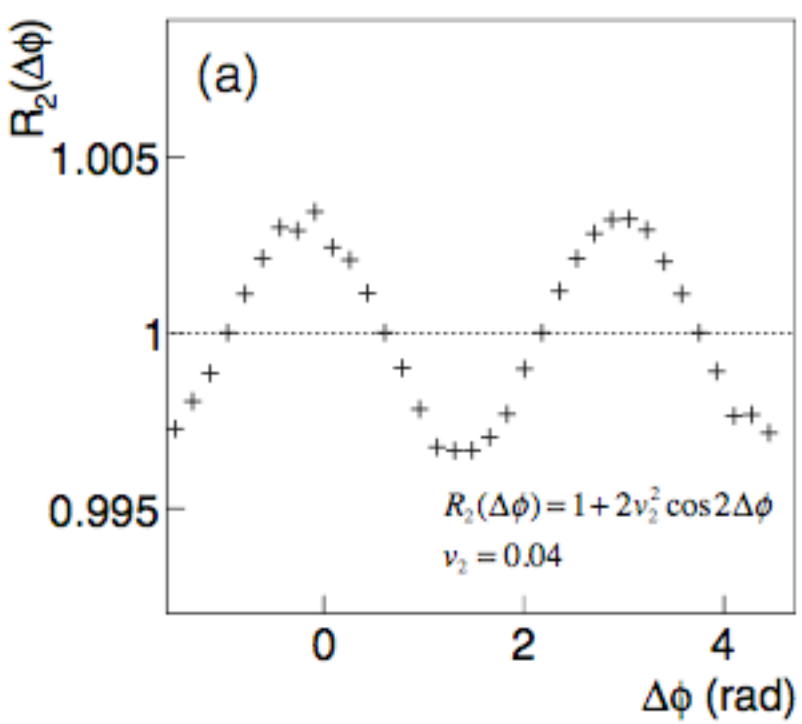


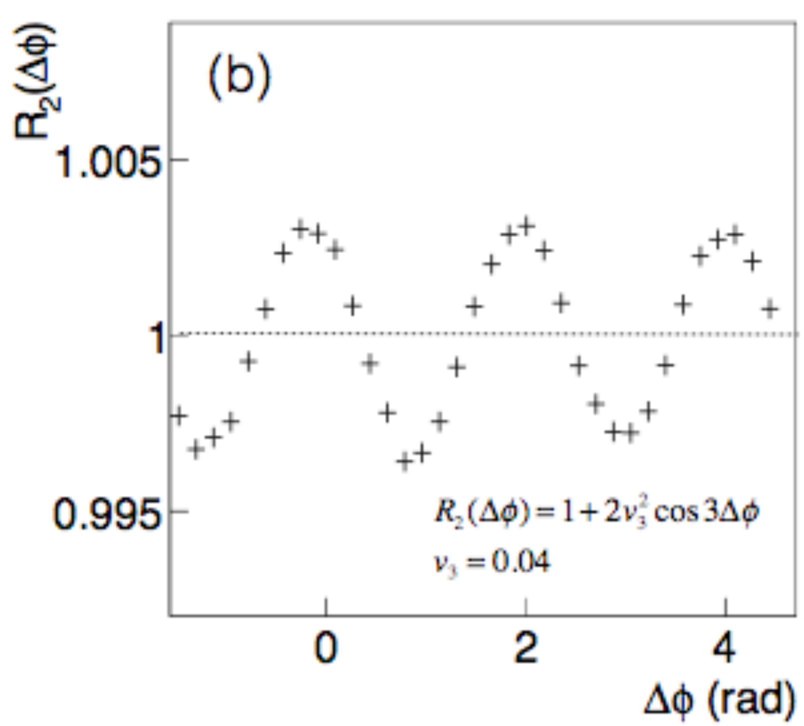
Flow

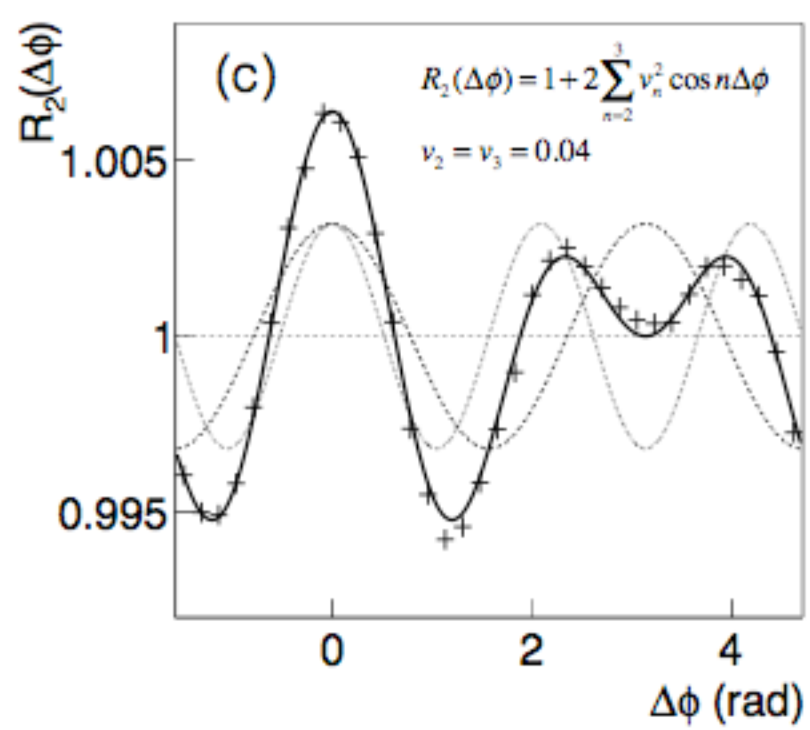

x
Spectators
Participants
$v \sim c$
$\vec{b}$
z
$v \sim c$
Spectators
d/γ


y
d
Participants
$\vec{b}$
x
Spectators


y
$\vec{p}_i$
ϕ_i
x


(a)
$R_2(\Delta\phi)$
1.005
1
0.995
$R_2(\Delta\phi) = 1 + 2v_2^2 \cos 2\Delta\phi$
$v_2 = 0.04$
0
2
4
Δφ (rad)


(b)
$R_2(\Delta\phi)$
1.005
1
0.995
$R_2(\Delta\phi) = 1 + 2v_3^2 \cos 3\Delta\phi$
$v_3 = 0.04$
0
2
4
Δφ (rad)


(c)
$R_2(\Delta\phi)$
$R_2(\Delta\phi) = 1 + 2\sum_{n=2}^{3} v_n^2 \cos n\Delta\phi$
$v_2 = v_3 = 0.04$
1.005
1
0.995
0
2
4
Δφ (rad)

Introducing number densities

- **Inclusive number densities ρ_n** are proportional to the n-probabilities (probability to find particle at some momentum coordinates).
- They yield a sequence of **inclusive differential functions**:

$$\frac{1}{\sigma_{\text{inel}}} d\sigma = \rho_1(y)dy,$$

$$\frac{1}{\sigma_{\text{inel}}} d^2\sigma = \rho_2(y_1, y_2)dy_1 dy_2,$$

$$\frac{1}{\sigma_{\text{inel}}} d^3\sigma = \rho_3(y_1, y_2, y_3)dy_1 dy_2 dy_3,$$

$$\ldots$$

n-particle densities or inclusive cross-sections.

Factorial Moments

- Integration over the momentum volume Ω yields

$$\int_{\Omega} \rho_1(y)dy = \int_{\Omega} \frac{d^3 N_i}{p_T dp_T d\phi d\eta} p_T\, dp_T\, d\phi\, d\eta = \langle N \rangle$$

$$\int_{\Omega} \rho_2(y_1, y_2) dy_1\, dy_2 = \langle N(N-1) \rangle$$

$$\int_{\Omega} \rho_3(y_1, y_2, y_3) dy_1\, dy_2\, dy_3 = \langle N(N-1)(N-2) \rangle$$

$$\cdots$$

$$\int_{\Omega} \cdots \int_{\Omega} \rho_n(y_1, \ldots, y_n) dy_1 \cdots dy_n = \langle N(N-1) \cdots (N-n+1) \rangle$$

Note:
Factorial Moments are integral of differential quantities.

- where $\langle N(N-1) \cdots (N-n+1) \rangle$ coefficients are called **factorial moments of order n**.

Number vs. Probability Densities

- Normalization of probability densities:

$$\int P(y_1, y_2, \ldots, y_n)\, dy_1\, dy_2 \ldots dy_n = 1$$

- But: $$\int_\Omega \cdots \int_\Omega \rho_n(y_1, \ldots, y_n)\, dy_1 \cdots dy_n = \langle N(N-1)\cdots(N-n+1)\rangle$$

- One can then write:

$$\rho_n(y_1, \ldots, y_n) = \langle N(N-1)\cdots(N-n+1)\rangle P_n(y_1, \ldots, y_n)$$

Factorial Moments

Differential Density

Probability Density

Independent Particle Emission

Absence of Correlations

- Two variables are said to be statistically independent iff their joint-probability density factorizes.
- Implies: Two particles are said to be statistically independent iff their joint-number density (which is proportional to a probability density) also factorizes.
- Example: For two particles, **Statistical Independence** is verified ONLY iff:

$$\rho_2(y_1, y_2) = \rho_1(y_1)\rho_1(y_2)$$

Wayne State University
College of Liberal Arts & Sciences
Department of Physics and Astronomy

Independent Particle Emission (2)

- With more than two particles: Statistical Independence is similarly verified ONLY iff:

$$\rho_n(y_1,\ldots,y_2) = \rho_1(y_1)\cdots\rho_1(y_n)$$

- But the emission of n particles may involve a superposition (sum) of processes leading to some correlated and uncorrelated particles.
- How do we extract the components corresponding to correlated particles??

Correlated and Uncorrelated Particle Production

- In general, inclusive n-particle densities $\rho_m(y_1, y_2, ..., y_m)$ are the result of a superposition of several subprocesses.
- Although the n particles might be produced by a single and specific subprocess, it is also quite possible that they originate from two or more distinct subprocesses.
- The n particles might in fact originate from n distinct and uncorrelated subprocesses.
- An n-tuplets of particles may then feature a broad variety of correlation sources associated with a plurality of dynamic processes.
- It is a common goal of multi-particle production measurements to identify and study these correlated emission as distinct (sub)processes.
- Accomplished by invoking correlation functions known as (factorial) cumulant functions, expressed either in terms of integral correlators or as differential functions of one or more particle coordinates.

Introducing cumulants, C_m

- **Cumulants of order m**, noted C_m, are defined as m-particle densities representing the emission (production) of **m correlated particles originating from a common production process**.
- Various notations used in the literature. We will use:

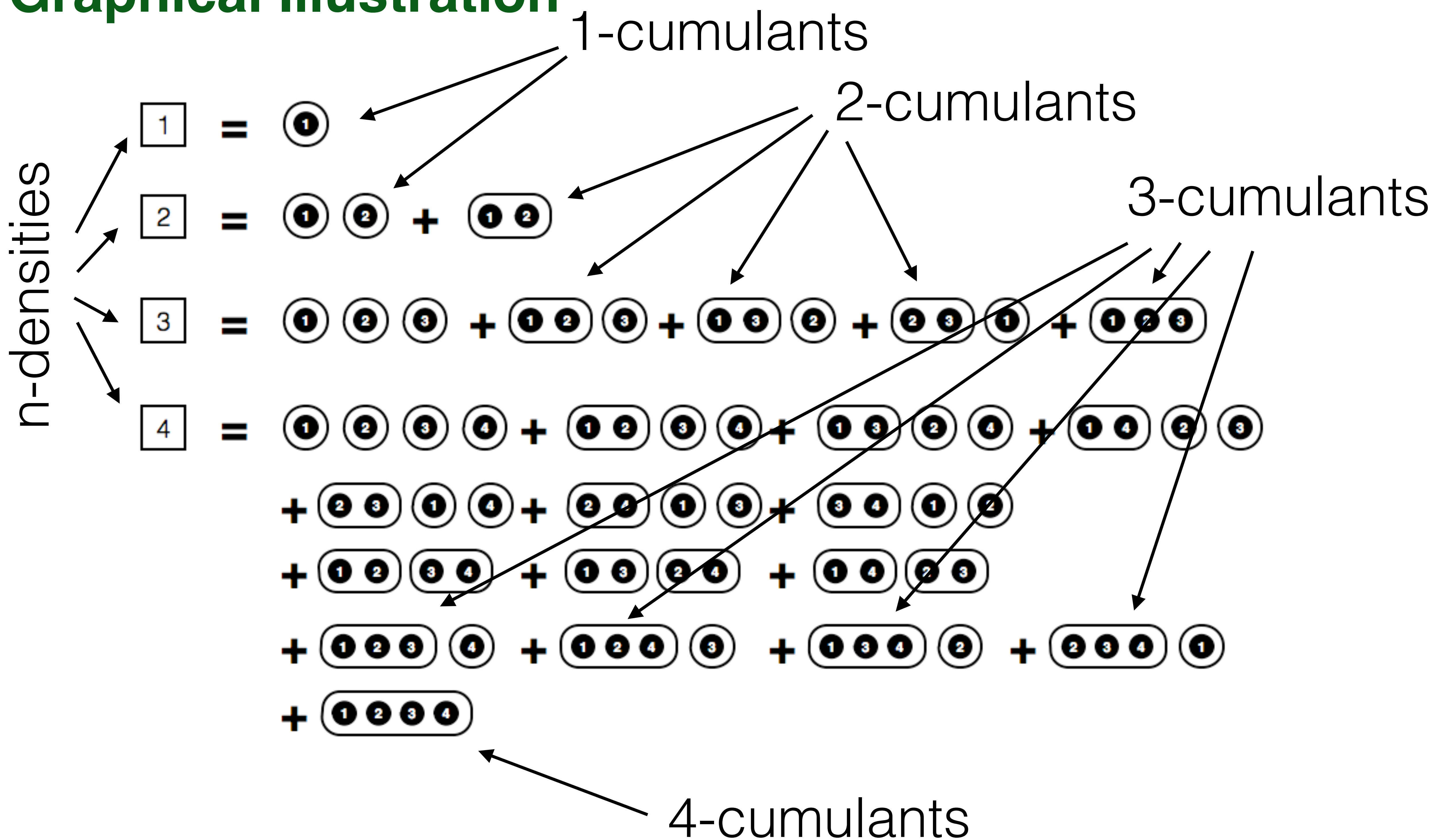
$$C_m \equiv \hat{\rho}_m$$

Multi-particle Densities

- Emission of n particles with n > m can be regarded as a superposition (sum) of several processes that together concur to produce a total of n particles.
- Let the term m-cluster refer to a group of m correlated particles produced a single process.
- There are, in principle, several ways to cluster n particles.
- An n-particle density can then be expressed as a sum of several terms yielding n particles, but each with its own “cluster” decomposition into products of cumulants.

Decomposition of Multi-particle Densities

Graphical Illustration



Wayne State University
College of Liberal Arts & Sciences
Department of Physics and Astronomy

Decomposition of Multi-particle Densities

- Mathematically…
- Using shorthand notation y_i —> i
- 1-Density: $\rho_1(1) = C_1(1)$
- 2-Density: $\rho_2(1,2) = C_1(1)C_1(2) + C_2(1,2)$
- 3-Density: $\rho_3(1,2,3) = C_1(1)C_1(2)C_1(3)$

$$+C_2(1,2)C_1(3) + C_2(1,3)C_1(2) + C_2(2,3)C_1(1)$$

$$+C_3(1,2,3)$$

Single correlated processes

Combinatorial processes

Single process

Decomposition of Multi-particle Densities

- *4*-Density:

$$
\begin{aligned}
\rho_4(1,2,3,4) = {} & C_1(1)C_1(2)C_1(3)C_1(4) \\
& +C_2(1,2)C_1(3)C_1(4)+C_2(1,3)C_1(2)C_1(4)+C_2(1,4)C_1(2)C_1(3) \\
& +C_2(2,3)C_1(1)C_1(4)+C_2(2,4)C_1(1)C_1(3)+C_2(3,4)C_1(2)C_1(3) \\
& +C_2(1,2)C_2(3,4)+C_2(1,3)C_2(2,4)+C_2(1,4)C_2(2,3) \\
& +C_3(1,2,3)C_1(4)+C_3(1,2,4)C_1(3)+C_3(1,3,4)C_1(2)+C_3(2,3,4)C_1(1) \\
& +C_4(1,2,3,4)
\end{aligned}
$$

Decomposition of Multi-particle Densities

- *Higher-densities*

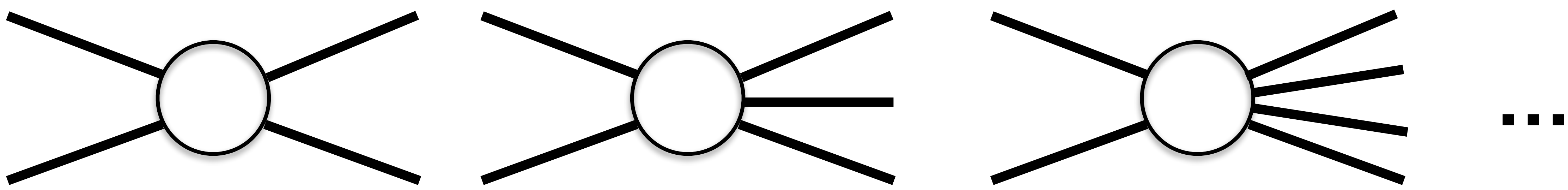
$$\rho_m(1,...,m) = C_m(1,...,m) + \sum_{perm.} C_1(1)C_{m-1}(2,...,m)$$
$$+ \sum_{perm} C_1(1)C_1(2)C_{m-2}(3,...,m)$$
$$+ \sum_{perm} C_2(1,2)C_{m-2}(3,...,m)$$
$$+ \cdots$$
$$+ \prod_{i=1}^{m} C_1(i)$$

"perm" indicates permutations of all particle indexes yielding distinct terms.

Formula such as this one can be obtained from cumulant generating functions…

Cumulants: Theory vs. Experiments

- m-cumulants represent fractions of the particle production cross-section **associated with processes yielding m (correlated) particles**.
- Theoretically: Calculated "directly" based on specific production models.



- **Experimentally:** Measured quantities are **densities**, not cumulants.

Measurements of Cumulants

- **Cumulants are not measured directly**.
- **Densities are first obtained** from measured particles.
- **Cumulants must be "extracted"** from measured densities.
- For instance:
  - 1-cumulant: $\rho_1(1) = C_1(1)$ (Trivial)
  - 2-cumulant: $\rho_2(1,2) = C_1(1)C_1(2) + C_2(1,2)$

by inversion: $C_2(1,2) = \rho_2(1,2) - \rho_1(1)\rho_1(2)$

Wayne State University
College of Liberal Arts & Sciences
Department of Physics and Astronomy

Measurements of Cumulants (2)

- Higher cumulants obtained recursively.
- For instance:

$$\rho_3(1,2,3) = C_1(1)C_1(2)C_1(3) + C_2(1,2)C_1(3) + C_2(1,3)C_1(2) + C_2(2,3)C_1(1) + C_3(1,2,3)$$

$$\rho_3(1,2,3) = \rho_1(1)\rho_1(2)\rho_1(3) + \left[\rho_2(1,2) - \rho_1(1)\rho_1(2)\right]\rho_1(3) + \left[\rho_2(1,3) - \rho_1(1)\rho_1(3)\right]\rho_1(2) + \left[\rho_2(2,3) - \rho_1(2)\rho_1(3)\right]\rho_1(1) + C_3(1,2,3)$$

by inversion:

$$C_3(1,2,3) = \rho_3(1,2,3) - \rho_2(1,2)\rho_1(3) - \rho_2(1,3)\rho_1(2) - \rho_2(2,3)\rho_1(1) + 2\rho_1(1)\rho_1(2)\rho_1(3)$$

Measurements of Cumulants (3)

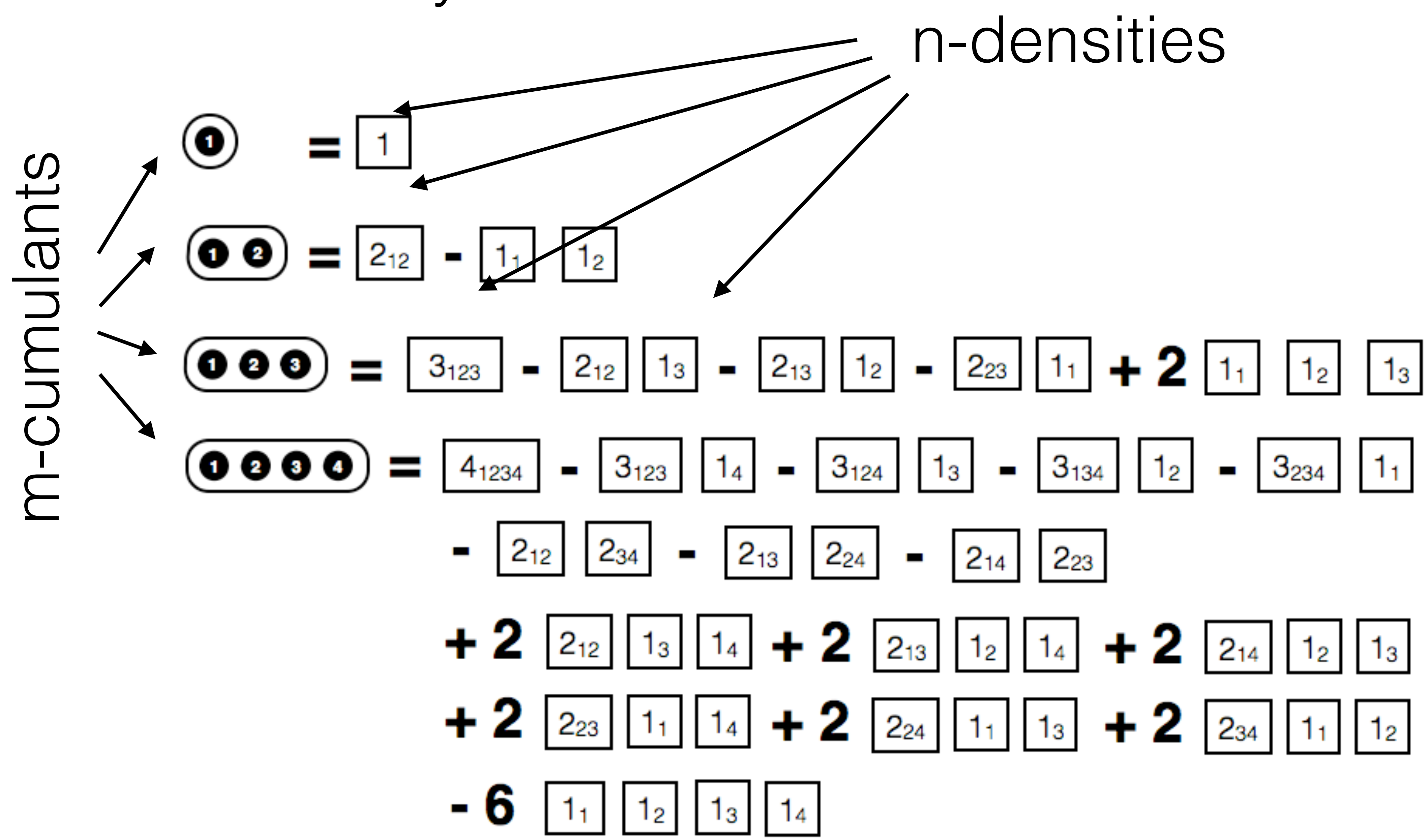
- And for 4-cumulants, one gets

$$
\begin{aligned}
C_4(1,2,3,4) = \rho_4(1,2,3,4) &- \sum_{(4)} \rho_1(1)\rho_3(2,3,4) \\
&- \sum_{(3)} \rho_2(1,2)\rho_2(3,4) + 2\sum_{(6)} \rho_1(1)\rho_1(2)\rho_2(3,4) \\
&- 6\rho_1(1)\rho_1(2)\rho_1(3)\rho_1(4)
\end{aligned}
$$

(n) indicates permutations of all particle indexes yielding distinct terms.

Measurements of Cumulants (4)

- Schematically…



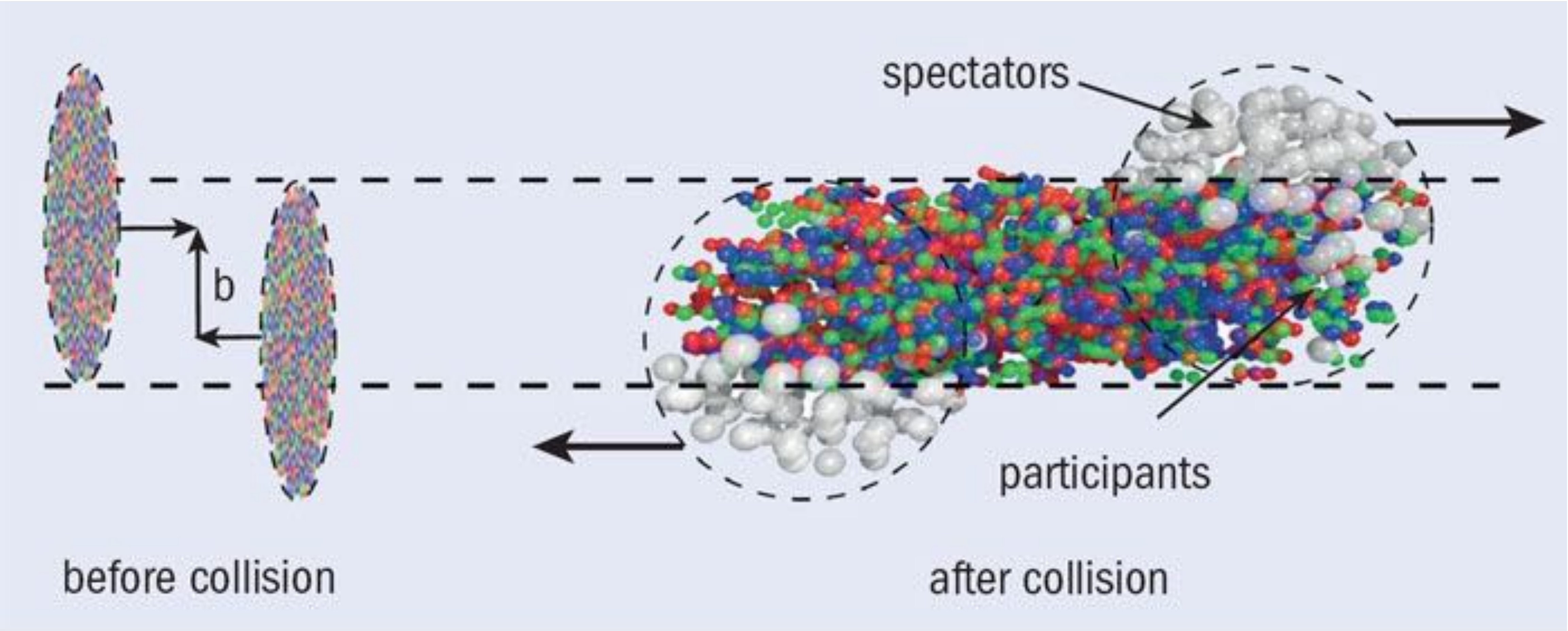
Important Remarks (1)

- **Densities** …
  - are **non-negative quantities**.
  - vary in amplitude according to the number of particles produced (n), the number of processes that yield particles, and the relative probability of these processes.
- **Cumulants** …
  - are **extracted** by adding/subtracting densities.
  - are **NOT positive definite**.
  - can be ***arbitrarily small compared to densities***.

Important Remarks (2)

- Measurements of Cumulants …
  - required (much) more statistics than densities of same order.
  - statistical errors of cumulant may be challenging to extract.
  - systematic errors can be a nightmare…

Part III: Cumulant Scaling Properties



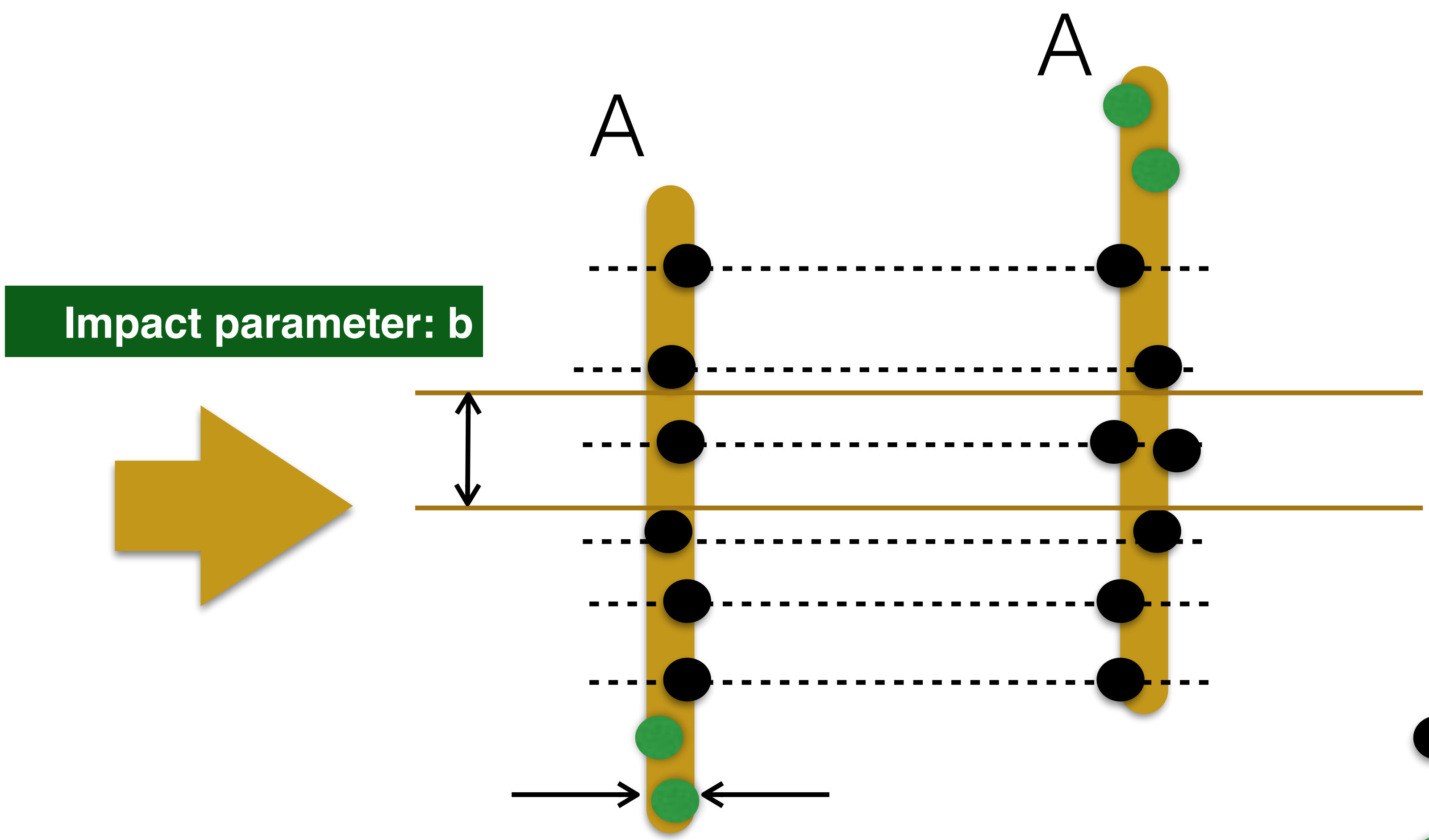
Cumulants $C_n(y_1,\ldots, y_n)$ feature a simple scaling property for collision systems consisting of a superposition of m_s independent (but otherwise identical) subsystems.

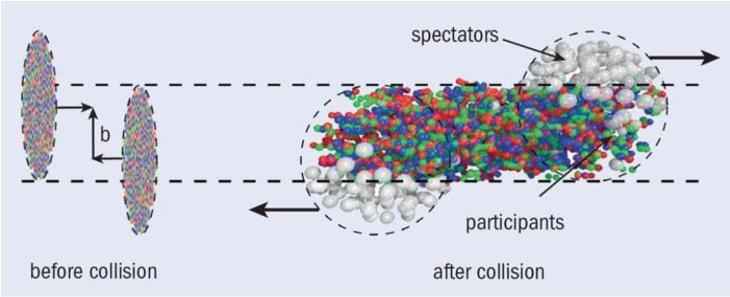
Context

- To **first approximation**, heavy ion collisions (HIC) can be regarded as a **superposition of**
  - independent nucleon-nucleon (n-n) collisions, or
  - independent constituent quark-quark (q-q) collisions, or
  - identical subsystems (whatever they might be)
  - with no re-scattering of produced particles.
- This approximation provides a **baseline** for the study of HIC: **how do actual HIC differ from a simple superposition of independent n-n scatterings?**

Independent Collision Approximation in HIC

- Observed cross-sections (densities) and cumulants are determined by the number of "**binary collisions**", which are considered, on average, to all be identical.





● **participants**

● **spectators**

$\gamma \sim 100$ at RHIC

$\gamma \sim 1000$ at LHC

Setup & Reasoning

- Consider a collision of two large nuclei (A - A collisions) at a specific energy.
- Assume that it can be reduced, to first-order approximation, to a superposition of m_s proton-proton (p - p) interactions, which each produce clusters consisting of n correlated particles.
- Assume the production of such clusters in p - p may be described by cumulants C_m^{pp}.
- At given impact parameter b, collisions should involve an average of $\langle m_s \rangle$ n - n interactions.
- Let us calculate the cumulants C_m^{AA} in A - A collisions.

Source Multiplicity Scaling - Cumulants

- m_S fluctuates collision-by-collision, but for a given value of m_S, one expects that the number of clusters of correlated particles of size n should be, on average, m_S times larger than in n - n collisions.
- The n-cumulant for A - A collisions, **at fixed** m_S, may thus be written

$$C_n^{AA}(y_1, y_3, y_2, \ldots, y_n) = m_s C_n^{pp}(y_1, y_2, y_3, \ldots, y_n)$$

- Given that m_S fluctuates event-by-event, averaging over all A - A collisions consequently yields

$$C_n^{AA}(y_1, y_3, y_2, \ldots, y_n) = \langle m_s \rangle C_n^{pp}(y_1, y_2, y_3, \ldots, y_n)$$

- for A - A collisions consisting of a superposition of independent and unmodified p - p collisions.

Scaling of total multiplicity

- Total multiplicity of particles produced in A - A collisions consisting of m_s independent and unmodified n - n collisions features the same scaling with m_s.
- Average multiplicity obtained in A - A for a given (fixed) value of m_s should simply be the product of m_s by the average particle multiplicity produced in n – n:

$$\rho_1^{AA}(y) = m_s \rho_1^{pp}(y)$$

$$\langle n \rangle_{AA} = m_s \langle n \rangle_{pp}$$

- since $\rho_1(y) = C_1(y)$
- and $C_1^{AA}(y) = m_s C_1^{pp}(y)$

Scaling of n>1 densities

- First consider pairs of particles.
  - In an A - A collision consisting of m_s independent n - n interactions, one can form m_s times the pairs from individual n-n collisions.
  - But one can also mix particles from different n-n interactions. Since there are $m_s(m_s-1)$ ways of doing that, one can write

$$\rho_2^{AA}(y_1,y_2) = m_s \rho_2^{pp}(y_1,y_2) + m_s(m_s-1)\rho_1^{pp}(y_1)\rho_1^{pp}(y_2)$$

- One obtains the same result using a cumulant decomposition:

$$\begin{aligned}
\rho_2^{AA}(y_1,y_2) &= C_1^{AA}(y_1)C_1^{AA}(y_2) + C_2^{AA}(y_1,y_2) \\
&= m_s^2 C_1^{pp}(y_1)C_1^{pp}(y_2) + m_s C_2^{pp}(y_1,y_2) \\
&= m_s^2 \rho_1^{pp}(y_1)\rho_1^{pp}(y_2) + m_s\left[\rho_2^{pp}(y_1,y_2) - \rho_1^{pp}(y_1)\rho_1^{pp}(y_2)\right] \\
&= m_s(m_s-1)\rho_1^{pp}(y_1)\rho_1^{pp}(y_2) + m_s \rho_2^{pp}(y_1,y_2)
\end{aligned}$$

Scaling of n>1 densities (2)

- At fixed value of m_s, integration over y_1 and y_2 yields:

$$\langle n(n-1)\rangle_{AA} = m_s\langle n(n-1)\rangle_{pp} + m_s(m_s-1)\langle n\rangle_{pp}^2$$

- For large m_s, the scaling of the number of pairs produced in A - A is dominated by the term in $m_s(m_s-1)$, which involves uncorrelated, combinatorial pairs from particle produced by different n-n collisions.

$$\begin{aligned}\rho_2^{AA}(y_1,y_2) &= C_1^{AA}(y_1)C_1^{AA}(y_2) + C_2^{AA}(y_1,y_2)\\ &= m_s^2 C_1^{pp}(y_1)C_1^{pp}(y_2) + m_s C_2^{pp}(y_1,y_2)\\ &= m_s^2\rho_1^{pp}(y_1)\rho_1^{pp}(y_2) + m_s\left[\rho_2^{pp}(y_1,y_2) - \rho_1^{pp}(y_1)\rho_1^{pp}(y_2)\right]\\ &= \boxed{m_s(m_s-1)\rho_1^{pp}(y_1)\rho_1^{pp}(y_2)} + m_s\rho_2^{pp}(y_1,y_2)\end{aligned}$$

Combinatorial Term

True Correlation Term

That's why one needs cumulants

Scaling of n>1 densities (3)

- Previous reasoning easily extended to n-densities with n > 2.
- Higher-order density measured in A - A amount to a combination of several n-n terms.
- The dominant terms is also "the most combinatoric"

$$\rho_n^{AA}(y_1, y_2, \ldots, y_n) = m_s(m_s - 1)\cdots(m_s - n + 1)\rho_1^{pp}(y_1)\cdots\rho_1^{pp}(y_n) + \cdots$$

- and n-cumulants would be the weakest term.

Wayne State University
College of Liberal Arts & Sciences
Department of Physics and Astronomy

3-Densities

$$
\begin{aligned}
\rho_3^{AA}(1,2,3) &= m_s^3 C_1^{pp}(1) C_1^{pp}(2) C_1^{pp}(3) \\
&\quad + m_s^2 \sum_{perms.} C_1^{pp}(1) C_2^{pp}(2,3) \\
&\quad + m_s C_3^{pp}(1,2,3) \\
&= \left(m_s^3 - m_s^2 + 2m_s\right) \rho_1^{pp}(1)\rho_1^{pp}(2)\rho_1^{pp}(3) \\
&\quad + \left(m_s^2 - m_s\right) \sum_{perms.} \rho_1^{pp}(1)\rho_2^{pp}(2,3) \\
&\quad + m_s \rho_3^{pp},
\end{aligned}
$$

Combinatorial Terms

True Correlation Term

$$
\begin{aligned}
\langle n(n-1)(n-2)\rangle_{AA} &= \left(m_s^3 - m_s^2 + 2m_s\right) \langle n\rangle_{pp}^3 \\
&\quad + 3\left(m_s^2 - m_s\right) \langle n(n-1)\rangle_{pp} \langle n\rangle_{pp} \\
&\quad + m_s \langle n(n-1)(n-2)\rangle_{pp}
\end{aligned}
$$

Again, we see that the combinatorial terms dominate over the most correlated terms for large values of m_s.

Wayne State University
College of Liberal Arts & Sciences
Department of Physics and Astronomy

Normalized Densities & Cumulants

- Convenient to divide densities and cumulants by products of one-particle densities.
- Leads to the definition of normalized inclusive densities and normalized cumulants:
- Normalized Densities:

$$r_n(y_1,\ldots,y_n) = \frac{\rho_n(y_1,\ldots,y_n)}{\rho_1(y_1)\cdots\rho_1(y_n)}$$

Often also called reduced densities.

- Normalized Cumulants:

$$R_n(y_1,\ldots,y_n) = \frac{C_n(y_1,\ldots,y_n)}{\rho_1(y_1)\cdots\rho_1(y_n)}$$

Often also called reduced cumulants.

- $R_2(y_1,y_2)$ correlation functions are quite commonly studied in HIC at RHIC and LHC.

Sorry: No standard/universal notations for these quantities.

Normalized Factorial Moments

- Also quite convenient/common to consider **normalized factorial moments**.

$$f_n = \frac{\langle N(N-1)\cdots(N-n+1)\rangle}{\langle N\rangle^n}$$

Often also called reduced factorial moments.

Sorry: No standard/universal notations for these quantities.

Scaling Behavior of Normalized Cumulants

- Interesting/convenient to consider the scaling behavior of normalized cumulants for systems consisting of a superposition of m_s identical sub-processes or sources.
- Based on the scaling of cumulants, one gets

$$
\begin{aligned}
R_n^{(m)}(y_1,\ldots,y_n) &= \frac{C_n^{(m)}(y_1,\ldots,y_n)}{\rho_1^{(m)}(y_1)\cdots\rho_1^{(m)}(y_n)} \\
&= \frac{mC_n^{(1)}(y_1,\ldots,y_n)}{m^n\rho_1^{(1)}(y_1)\cdots\rho_1^{(1)}(y_n)} \\
&= \frac{1}{m^{n-1}}\frac{C_n^{(1)}(y_1,\ldots,y_n)}{\rho_1^{(1)}(y_1)\cdots\rho_1^{(1)}(y_n)} \\
&= \frac{1}{m^{n-1}}R_n^{(1)}(y_1,\ldots,y_n)
\end{aligned}
$$

Normalized cumulant for m sources

Cumulant for m sources

Cumulant for 1 source

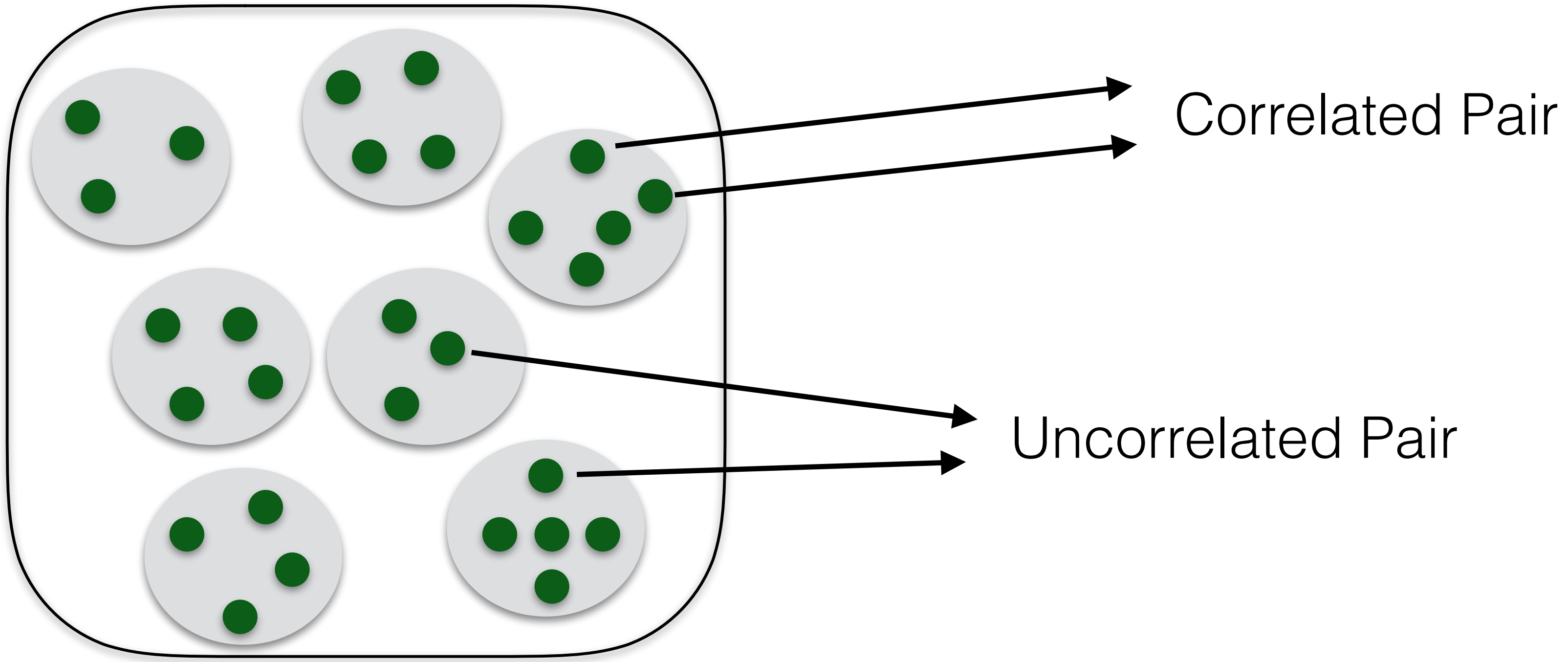
Normalized cumulant for 1 source

Scaling Factor or Dilution Factor

Remark

$$R_n^{(m)}(y_1,\ldots,y_n) = \frac{1}{m^{n-1}} R_n^{(1)}(y_1,\ldots,y_n)$$

- The inverse (n-1)th power of m implies the strength of m-cumulants shall (in general) monotonically decrease with the system size i.e., for systems consisting of “sum” of m identical subsystems.
- The normalized cumulants are said to be **diluted by a power m^{n-1}** relative to the elementary systems composing the large system.
- Dilution is due to combinatorial effects: with m sources, there are far many ways to make uncorrelated pairs than correlated ones.
- An important effect (or consideration) in heavy ion collisions because spatial correlation lengths are relatively small, and the collision systems very short lived.

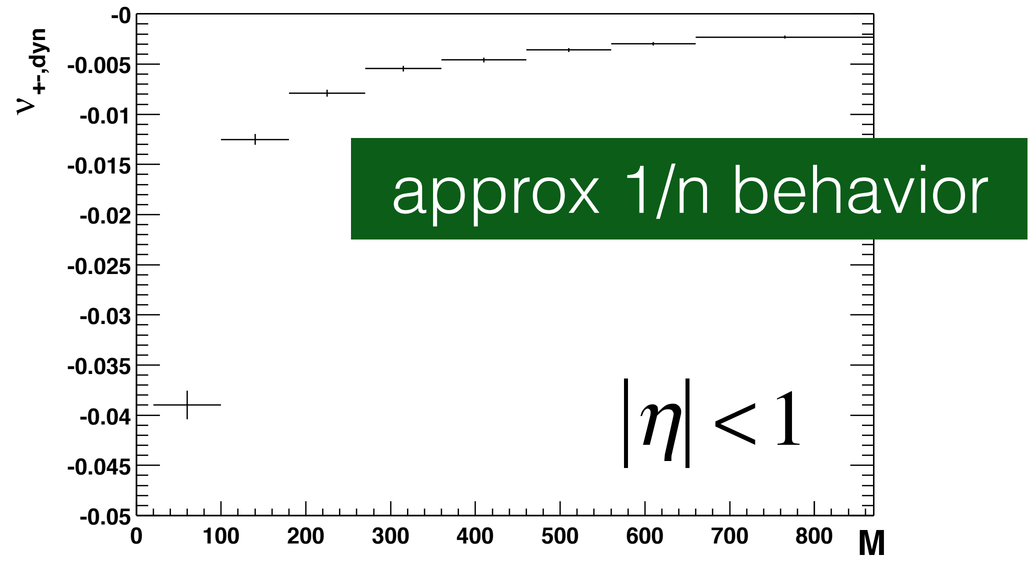


Example: Nu-Dyn

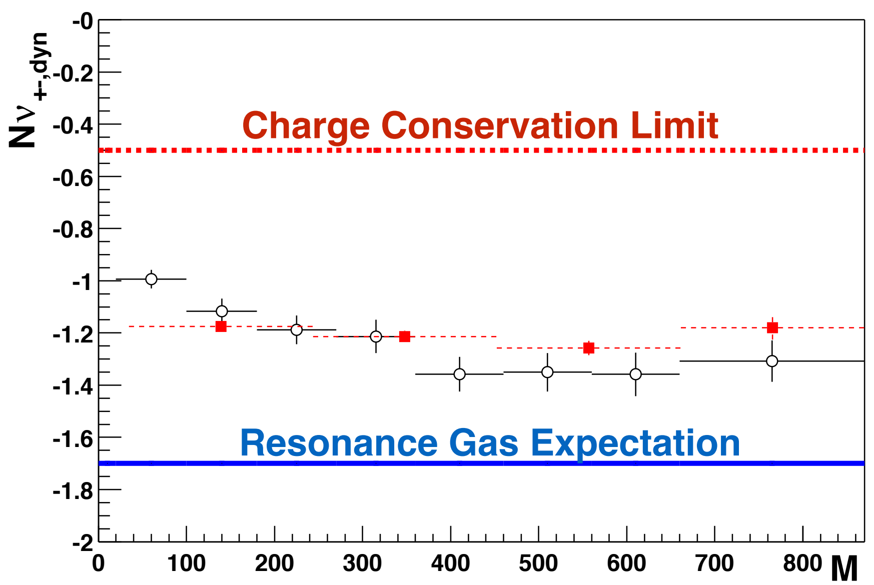
Conservation laws and particle production processes underlie correlations.

Multiplicity fluctuations in Au+Au collisions at $\sqrt{s_{NN}} = 130$ GeV

STAR, Phys.Rev. C68 (**2003**) 044905



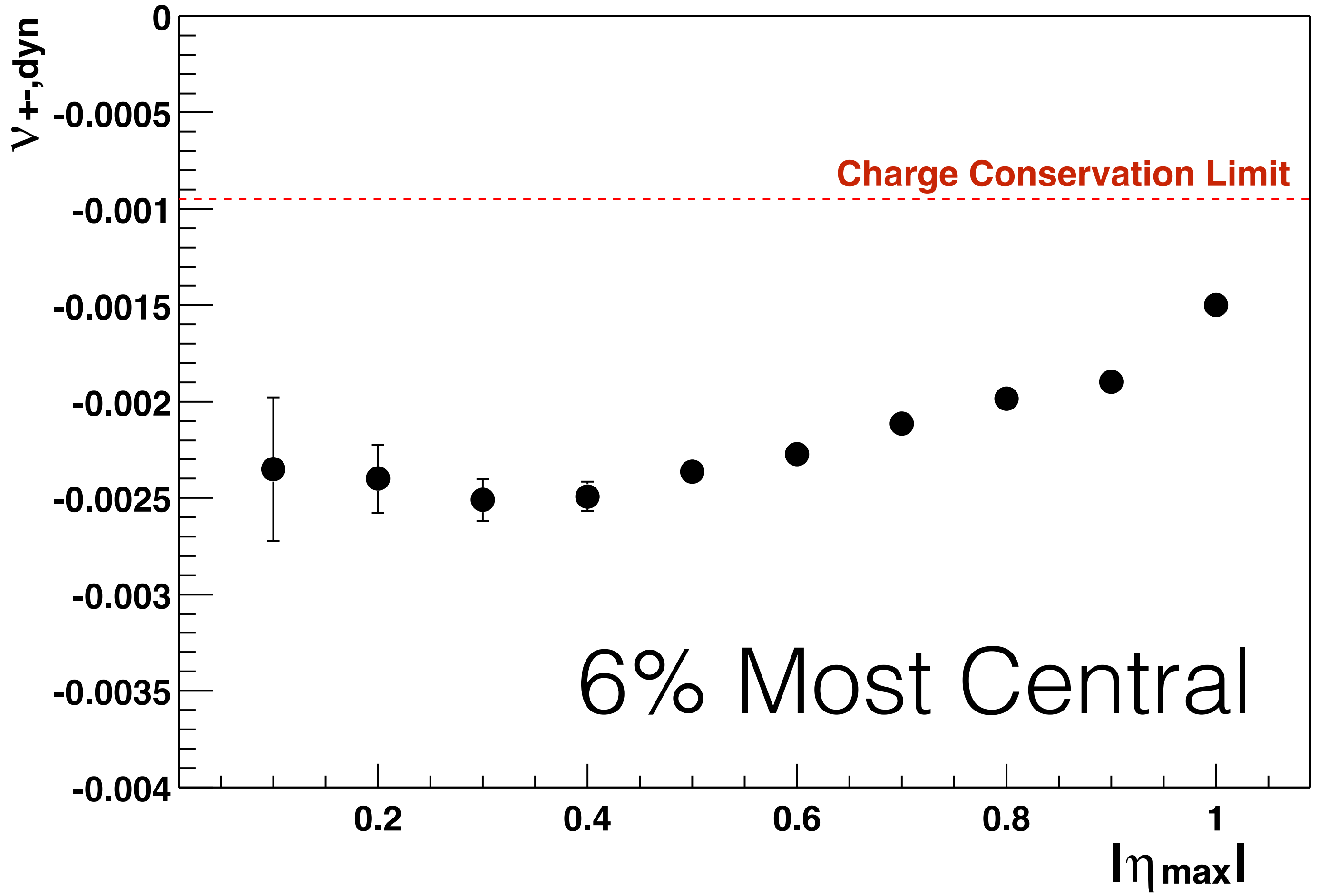
(a)



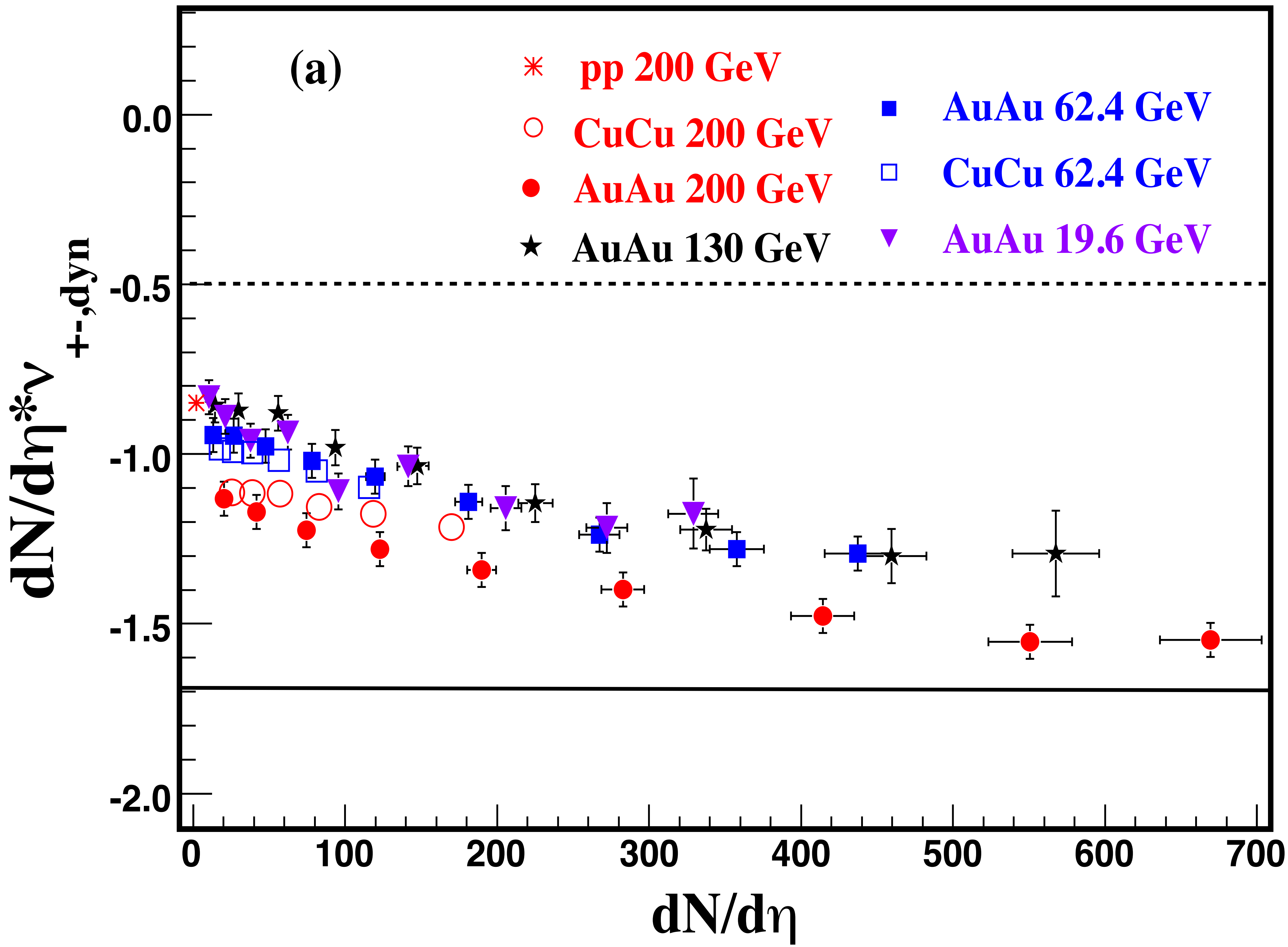
(b)

$$\nu_{+-,dyn} = \bar{R}_{++} + \bar{R}_{--} - 2\bar{R}_{+-}$$

$$\bar{R}_{ab} = \frac{\int_{\Delta\eta} R_{2,ab}(\eta_a,\eta_b)\rho_{1,a}(\eta_a)\rho_{1,b}(\eta_b)d\eta_a d\eta_b}{\int_{\Delta\eta}\rho_{1,a}(\eta_a)d\eta_a \int_{\Delta\eta}\rho_{1,b}(\eta_b)d\eta_b}$$



Width of the Acceptance Matters



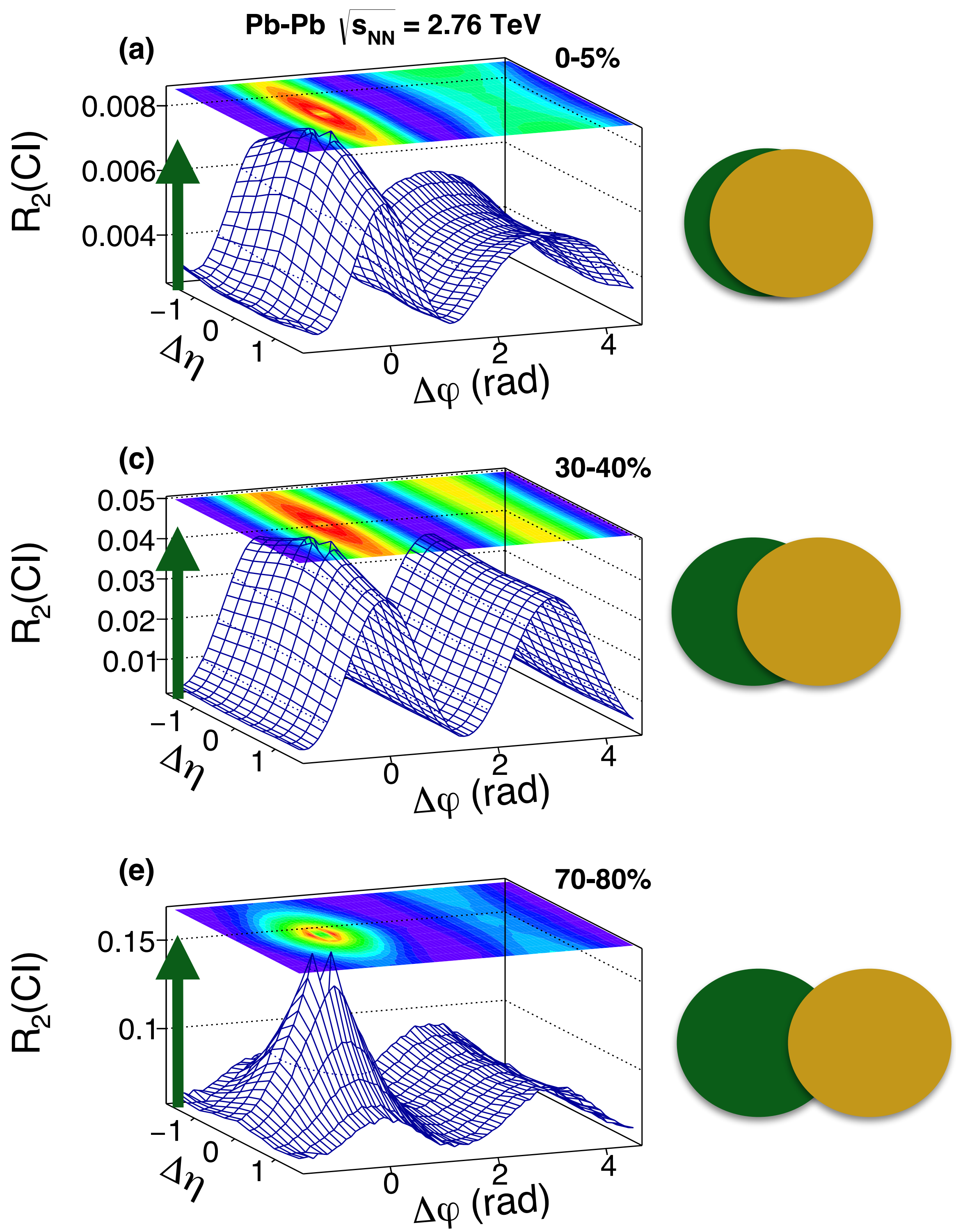
(a)
pp 200 GeV
CuCu 200 GeV
AuAu 200 GeV
AuAu 130 GeV
AuAu 62.4 GeV
CuCu 62.4 GeV
AuAu 19.6 GeV
0.0
-0.5
-1.0
-1.5
-2.0
0
100
200
300
400
500
600
700
dN/dη
dN/dη*ν+-,dyn

Remark (3)

$$R_2^{(m)}(y_1,y_2)=\frac{1}{m}R_2^{(1)}(y_1,y_2)$$

- R_2 expected to scale approximately as 1/m.
- Decrease of correlation actually observed in RHIC and LHC
- Also **observed a change in the shape of the correlation function**
  - **Indicative of a modification of the correlation dynamics, i.e., the processes that produce the particles.**
- Measurements of higher order cumulant require lots of statistics — because the actual strength of the cumulant is much weaker than that of the density.

$$\rho_2^{AA}(y_1,y_2)=m_s(m_s-1)\rho_1^{pp}(y_1)\rho_1^{pp}(y_2)+m_s\rho_2^{pp}(y_1,y_2)$$



Pujahari et al., ALICE

Cumulants are Statistics Hungry

- Measurements of cumulant require lots of statistics because the actual strength of the cumulant is much weaker than that of the density.

$$R_2(y_1,y_2) = \frac{\rho_2(y_1,y_2) - \rho_1(y_1)\rho_1(y_2)}{\rho_1(y_1)\rho_1(y_2)} = \frac{\rho_2(y_1,y_2)}{\rho_1(y_1)\rho_1(y_2)} - 1$$

$$R_2^{(m)}(y_1,y_2) = \frac{1}{m} R_2^{(1)}(y_1,y_2)$$

$$\frac{\rho_2(y_1,y_2)}{\rho_1(y_1)\rho_1(y_2)} \approx 1$$

- ρ_2 and ρ_1 ρ_1 have approximate same magnitude
- Their difference is nearly zero
- Their ratio is of order unity
- R_2 thus has larger relative statistical errors than ρ_2.

Probability Densities & Statistical Independence

- Integration of particle densities $\rho_n(y_1, ..., y_n)$ over the momentum volume Ω provides a natural and convenient normalization to define particle probability densities:

$$P_n(y_1,\ldots,y_n) = \frac{\rho_n(y_1,\ldots,y_n)}{\langle N(N-1)\cdots(N-n+1)\rangle}$$

- Expresses the probability of finding n particles jointly at y_1, y_2, …, y_n.
- Reduction of these probabilities by products of single particle probability densities yields

$$q_n(y_1,\ldots,y_n) = \frac{P_n(y_1,\ldots,y_n)}{P_1(y_1)P_1(y_2)\cdots P_n(y_n)}$$

- which must equal unity if the particles are emitted/produced independently.

Strength of Correlations

- Normalized densities written in terms of normalized factorial moments and the function q.

$$r_n(y_1,\ldots,y_n)=\frac{\langle N(N-1)\cdots(N-n+1)\rangle}{\langle N\rangle^n}q_n(y_1,\ldots,y_n)$$

- which tells us that the strength of correlation depends both on multiplicity fluctuations through

$$\frac{\langle N(N-1)\cdots(N-n+1)\rangle}{\langle N\rangle^n}\neq 1$$

- and the shape and magnitude of $q_n(y_1, \ldots, y_n)$.

Correlation function Normalization

2-Cumulant: $C_2(y_1, y_2) = \rho_2(y_1, y_2) - \rho_1(y_1)\rho_1(y_1)$

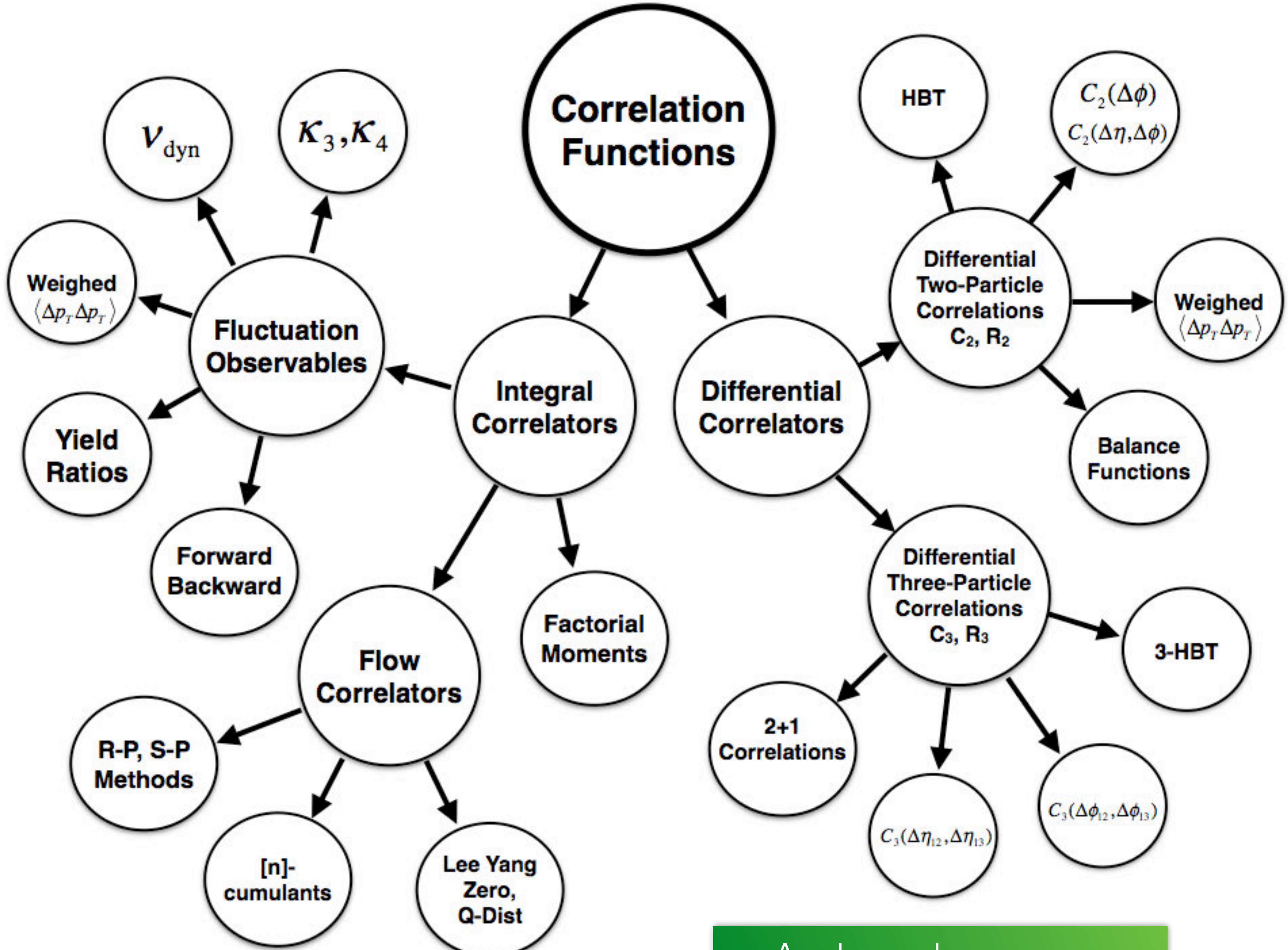
Normalized 2-Cumulant: $R_2(y_1, y_2) = \dfrac{C_2(y_1, y_2)}{\rho_1(y_1)\rho_1(y_2)}$ $\begin{cases} >0 \text{ correlation} \\ =0 \text{ no correlation} \\ <0 \text{ anti-correlation} \end{cases}$

Normalized 2-Density: $r_2(y_1, y_2) = 1 + \dfrac{C_2(y_1, y_2)}{\rho_1(y_1)\rho_1(y_2)}$ $\begin{cases} >1 \text{ correlation} \\ =1 \text{ no correlation} \\ <1 \text{ anti-correlation} \end{cases}$

But not a per trigger ratio: $K_2(y_1, y_2) = \dfrac{\rho_2(y_1, y_2)}{\rho_1(y_1)}$

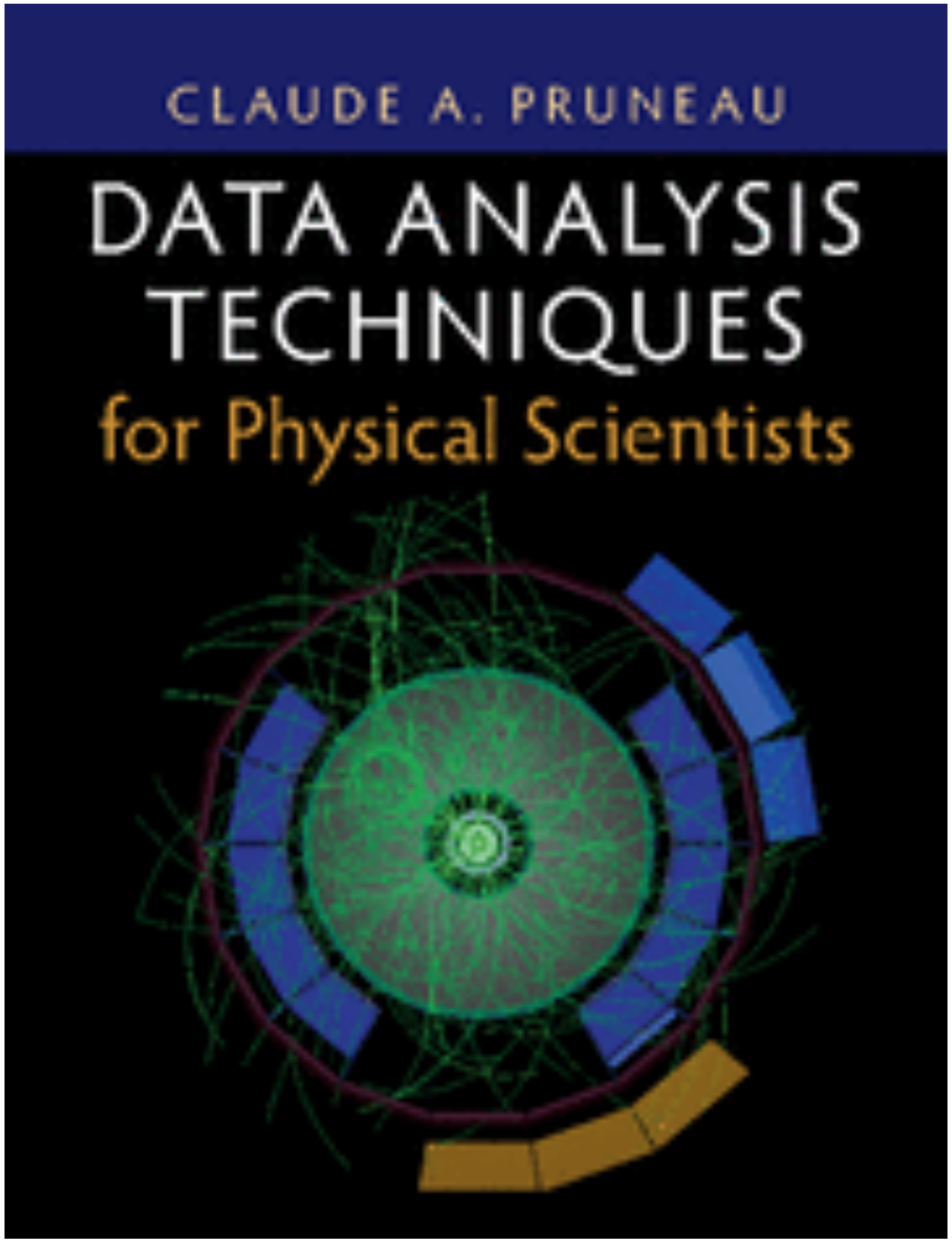
>0 always
No proper reference level!
Problematic!

Part IV: The Multiple Facets of Correlation Functions



And much more …

Source of Information



CAMBRIDGE

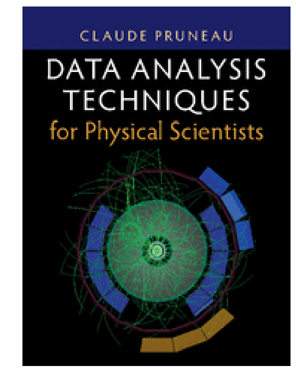
20% Discount *on this title*

Expires 8 August 2018

Data Analysis Techniques for Physical Scientists

Claude A. Pruneau

Wayne State University, Michigan



A comprehensive guide to data analysis techniques for physical scientists, providing a valuable resource for advanced undergraduate and graduate students, as well as seasoned researchers. The book begins with an extensive discussion of the foundational concepts and methods of probability and statistics under both the frequentist and Bayesian interpretations of probability. It next presents basic concepts and techniques used for measurements of particle production cross-sections, correlation functions, and particle identification. Much attention is devoted to notions of statistical and systematic errors, beginning with intuitive discussions and progressively introducing the more formal concepts of confidence intervals, credible range, and hypothesis testing. The book also includes an in-depth discussion of the methods used to unfold or correct data for instrumental effects associated with measurement and process noise as well as particle and event losses, before ending with a presentation of elementary Monte Carlo techniques.

September 2017

246 x 189 mm c.730pp 195 b/w illus. 20 tables

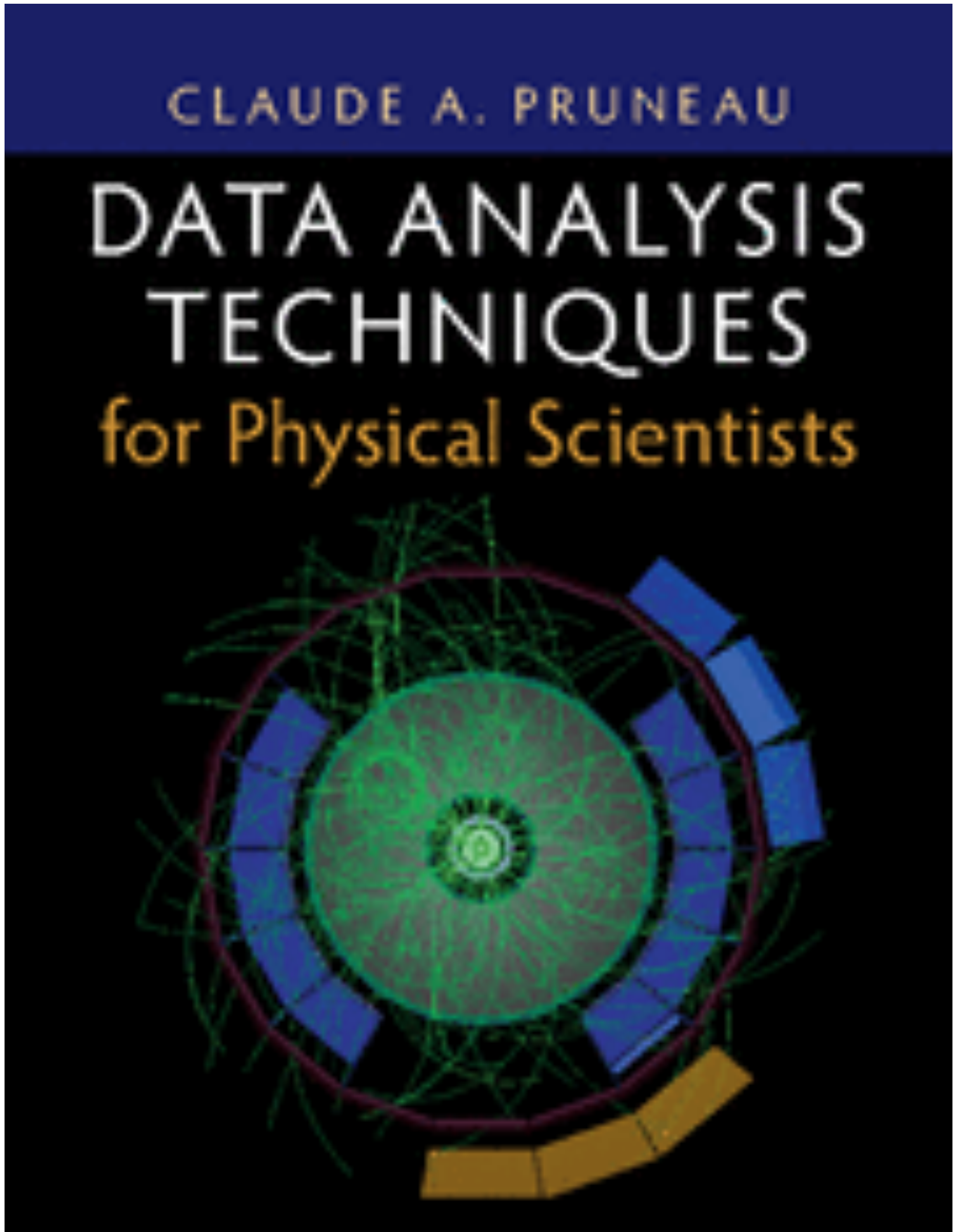
Hardback 978-1-108-41678-8

Original price	*Discount price*
£69.99	£55.99
$89.99	$71.99

Advance praise: 'This ambitious book provides a comprehensive, rigorous, and accessible introduction to data analysis for nuclear and particle physicists working on collider experiments, and outlines the concepts and techniques needed to carry out forefront research with modern collider data in a clear and pedagogical way. The topic of particle correlation functions, a seemingly straightforward topic with conceptual pitfalls awaiting the unaware, receives two full chapters. Professor Pruneau presents these concepts carefully and systematically, with precise definitions and extensive discussion of interpretation. These chapters should be required reading for all practitioners working in this area.'

Peter Jacobs,
Lawrence Berkeley National Laboratory

Source of Information



CAMBRIDGE

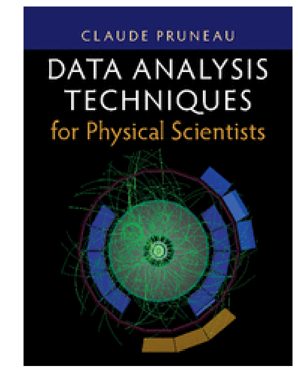
20% Discount *on this title*

Expires 8 August 2018

Data Analysis Techniques for Physical Scientists

Claude A. Pruneau

Wayne State University, Michigan



A comprehensive guide to data analysis techniques for physical scientists, providing a valuable resource for advanced undergraduate and graduate students, as well as seasoned researchers. The book begins with an extensive discussion of the foundational concepts and methods of probability and statistics under both the frequentist and Bayesian interpretations of probability. It next presents basic concepts and techniques used for measurements of particle production cross-sections, correlation functions, and particle identification. Much attention is devoted to notions of statistical and systematic errors, beginning with intuitive discussions and progressively introducing the more formal concepts of confidence intervals, credible range, and hypothesis testing. The book also includes an in-depth discussion of the methods used to unfold or correct data for instrumental effects associated with measurement and process noise as well as particle and event losses, before ending with a presentation of elementary Monte Carlo techniques.

September 2017

246 x 189 mm c.730pp 195 b/w illus. 20 tables

Hardback 978-1-108-41678-8

Original price	*Discount price*
£69.99	£55.99
$89.99	$71.99

Advance praise: 'This ambitious book provides a comprehensive, rigorous, and accessible introduction to data analysis for nuclear and particle physicists working on collider experiments, and outlines the concepts and techniques needed to carry out forefront research with modern collider data in a clear and pedagogical way. The topic of particle correlation functions, a seemingly straightforward topic with conceptual pitfalls awaiting the unaware, receives two full chapters. Professor Pruneau presents these concepts carefully and systematically, with precise definitions and extensive discussion of interpretation. These chapters should be required reading for all practitioners working in this area.'

Peter Jacobs,
Lawrence Berkeley National Laboratory

For more information, and to order, visit:
www.cambridge.org/9781108416788
and enter the code PRUNEAU2017 at the checkout

CAMBRIDGE
UNIVERSITY PRESS

Part V Experimental Considerations

- Detector Acceptance
- Detection Efficiency
- Momentum Smearing
- Signal Contamination
  - Physical Backgrounds
  - Instrumental Backgrounds

Efficiency Losses

- Introduce the notion of efficiency in the context of multiplicity measurements.
- Method for the correction of Mean Number of Particles in a given acceptance (multiplicity).
- Issues with the variance.

Mean Particle Production

- Theoretically: average integrated yield, ⟨N⟩, over a specific kinematic domain, Ω, determined by the particle production cross-section

$$\langle N \rangle = \int_{\Omega} \frac{d^3N}{dp^3} dp^3$$

- Experimentally: number of particles fluctuates collision by collision owing to the stochastic nature of the particle production process.
- Fluctuations described by a probability function, $P_{prod}(N)$, determined by the dynamics and correlations involved in the particle production process

$$\langle N \rangle = \int_{\Omega} P_{\text{prod}}(N) N \, dN$$

Particle Losses

- Measurements of particle production are usually subject to losses.
- For large detectors, one can usually assume that the probability of detecting one particle is independent of the probability of detecting others. One can then model the detection of a single-particle with a Bernoulli distribution

$$\begin{aligned} P_{\text{single}}(n|\varepsilon) &= 1-\varepsilon \quad \text{probability of not observing, } n=0 \\ &= \varepsilon \quad \text{probability of observing, } n=1. \end{aligned}$$

- We can then express the probability of simultaneously detecting n particles in the domain Ω as a binomial distribution with success probability ε:

$$P_{\text{det}}(n|N,\varepsilon) = \frac{N!}{n!(N-n)!}\varepsilon^{N}(1-\varepsilon)^{N-n}$$

Wayne State University
College of Liberal Arts & Sciences
Department of Physics and Astronomy

Average & Variance at Fixed N

- For a fixed produced multiplicity N, the measured average is then

$$\langle n \rangle_{\mathrm{N}} = \mathrm{E}[n] = \int P_{\mathrm{det}}(n|N,\varepsilon) n dn = \varepsilon N,$$

- The measured variance at fixed N

$$\begin{aligned}\langle (n - \langle n \rangle)^2 \rangle_{\mathrm{N}} &= \int P_{\mathrm{det}}(n|N,\varepsilon)\,(n - \langle n \rangle)^2\, dn \\ &= N\varepsilon(1-\varepsilon).\end{aligned}$$

Average w/ varying N

- The probability of observing n particles when N are produced

$$P_{\text{meas}}(n|\varepsilon) = \int dN P_{\text{det}}(n|N,\varepsilon) P_{prod}(N).$$

- The mean measured multiplicity <n> is then

$$\begin{aligned}\langle n\rangle &= \int dn n P_{\text{meas}}(n|\varepsilon),\\ &= \int dn n \int dN P_{\text{det}}(n|N,\varepsilon) P_{prod}(N).\end{aligned}$$

- Interchanging the order of integrations

$$\begin{aligned}\langle n\rangle &= \int dN P_{prod}(N) \int dn n P_{\text{det}}(n|N,\varepsilon),\\ &= \varepsilon \int dN P_{prod}(N) N,\\ &= \varepsilon \langle N\rangle.\end{aligned}$$

The observed mean is proportional to the produced mean. The proportionality factor is the efficiency.

Wayne State University
College of Liberal Arts & Sciences
Department of Physics and Astronomy

Efficiency Correction

- If smearing can be neglected, correction for particle losses is simply accomplished according to:

$$\langle N \rangle = \frac{\langle n \rangle}{\varepsilon}.$$

Unfriendly Variance

- The second moment of the measured multiplicity is

$$\begin{aligned}\langle n^2\rangle &= \int dN P_{\rm Prod}(N) \int dn n^2 P_{\rm det}(n|N,\varepsilon),\\ &= \int dN P_{\rm Prod}(N) N\varepsilon(1-\varepsilon+N\varepsilon),\\ &= \varepsilon(1-\varepsilon)\langle N\rangle + \varepsilon^2\langle N^2\rangle,\end{aligned}$$

- The variance of the measured distribution is thus

$$\mathrm{Var}[n] = \varepsilon^2 \mathrm{Var}[N] + \varepsilon(1-\varepsilon)\langle N\rangle.$$

- The variance CANNOT be corrected by a simple factor!

Friendly Factorial Moments

- Moments:

$$\langle n \rangle = \varepsilon \langle N \rangle$$

$$\langle n^2 \rangle = \varepsilon (1-\varepsilon) \langle N \rangle + \varepsilon^2 \langle N^2 \rangle$$

- Factorial Moments:

$$\langle n(n-1) \rangle = \langle n^2 \rangle - \langle n \rangle = \varepsilon (1-\varepsilon) \langle N \rangle + \varepsilon^2 \langle N^2 \rangle - \varepsilon \langle N \rangle$$

$$= -\varepsilon^2 \langle N \rangle + \varepsilon^2 \langle N^2 \rangle$$

$$= \varepsilon^2 \langle N(N-1) \rangle$$

- R_2:

$$R_2^M = \frac{\langle n(n-1) \rangle}{\langle n \rangle^2} = \frac{\varepsilon^2 \langle N(N-1) \rangle}{\varepsilon^2 \langle N \rangle^2} = \frac{\langle N(N-1) \rangle}{\langle N \rangle^2} = R_2^T$$

Measured R_2

True R_2

- Same properties for higher factorial moments
- Measurements of factorial moments ratios intrinsically more robust!!

What if the efficiency changes w/ time?

- Assume that an experiment can be divided into two time periods featuring particle detection efficiencies ε_1 and ε_2.
- Let us also assume that the probability of observing the events during the two time periods is unmodified by this change,
- Let us denote the number of events detected in the two periods as N_1^{ev} and N_2^{ev}.
- The average efficiency is calculated as a weighted average of the efficiencies of the two periods

$$\varepsilon_{\text{avg}} = \frac{N_1^{ev}\varepsilon_1 + N_2^{ev}\varepsilon_2}{N_1^{ev} + N_2^{ev}}$$

What if the efficiency changes? (II)

- The multiplicity measured across the two periods is:

$$\langle n \rangle = \varepsilon_{\text{avg}} \langle N \rangle .$$

- Extraction of the true mean multiplicity ⟨N⟩ can thus be obtained for either time periods

$$\langle N \rangle = \frac{\langle n \rangle_1}{\varepsilon_1} = \frac{\langle n \rangle_2}{\varepsilon_2}$$

- or globally from the average of the two periods

$$\langle N \rangle = \frac{\langle n \rangle}{\varepsilon_{\text{avg}}}$$

What if the efficiency changes? (III)

$$
\begin{aligned}
\langle n \rangle_{\text{avg}} &= \frac{N_1^{ev}\langle n \rangle_1 + N_2^{ev}\langle n \rangle_2}{N_1^{ev} + N_2^{ev}} \\
&= \frac{N_1^{ev}\varepsilon_1\langle N \rangle + N_2^{ev}\varepsilon_2\langle N \rangle}{N_1^{ev} + N_2^{ev}} \\
&= \frac{N_1^{ev}\varepsilon_1 + N_2^{ev}\varepsilon_2}{N_1^{ev} + N_2^{ev}}\langle N \rangle \\
&= \varepsilon_{\text{avg}}\langle N \rangle.
\end{aligned}
$$

This conclusion can be generalized to multiple time periods when the detection efficiency might have taken different values. For measurements of <N>, it does not matter that the experimental response changes over time as long as one can track these changes and estimate the detection efficiency during each period independently or globally for the entire data-taking run.

What if the efficiency changes within the acceptance?

- Split the measurement acceptance into two parts of size Ω_1 and Ω_2 with respective efficiencies ε_1 and ε_2.

$$\langle n_i \rangle = \varepsilon_i \langle N_i \rangle,$$

$$\langle N_i \rangle = \int_{\Omega_i} \frac{d^3N}{dp^3} dp^3$$

$$\Omega = \sum_{i=1}^{2} \Omega_i.$$

- The average number of produced particles can be properly determined by summing corrected yields in part 1 and 2 individually

$$\langle N \rangle = \sum_{i=1}^{2} \langle N_i \rangle = \sum_{i=1}^{2} \frac{\langle n_i \rangle}{\varepsilon_i}$$

What if the efficiency changes within the acceptance?

- If the fractions $f_i = \langle N_i \rangle / \langle N \rangle$ of the total yield produced in the two parts of the acceptance are known a priori, one can write an average efficiency (as in the case of the time-varying efficiency discussed above):

$$\varepsilon_{\text{avg}} = \frac{f_1\varepsilon_1 + f_2\varepsilon_2}{f_1 + f_2} = f_1\varepsilon_1 + f_2\varepsilon_2 \qquad f_1 + f_2 = 1$$

- Unfortunately, the fractions f_i are in general not known a priori, and it is thus not possible to formally define a model independent average efficiency across the full acceptance Ω.
- However, in cases where the production cross-section is nearly constant within the experimental acceptance, one can write

$$f_i = \frac{\int_{\Omega_i} \frac{d^3N}{dp^3} dp^3}{\int_{\Omega} \frac{d^3N}{dp^3} dp^3} \approx \frac{\int_{\Omega_i} dp^3}{\int_{\Omega} dp^3} = \frac{\Omega_i}{\Omega} \qquad f_1 + f_2 = 1$$

Two-Particle Case

- Single and Pair Yields

$$\hat{n}_1(x) = [f_1\varepsilon_1(x) + f_2\varepsilon_2(x)]\,\rho_1(x),$$

$$\hat{n}_2(x_1, x_2) = [f_1\varepsilon_1(x_1)\varepsilon_1(x_2) + f_2\varepsilon_2(x_1)\varepsilon_2(x_2)]\,\rho_2(x_1, x_2).$$

- Normalized Pair Yields

$$\hat{r}_2^{\text{meas}}(x_1, x_2) = \xi(x_1, x_2)\hat{r}_2(x_1, x_2), \qquad \hat{r}_2(x_1, x_2) \equiv \frac{\rho_2(x_1, x_2)}{\rho_1(x_1)\rho_1(x_2)},$$

Robustness Function

$$\xi(x_1, x_2) = \frac{f_1\varepsilon_1(x_1)\varepsilon_1(x_2) + f_2\varepsilon_2(x_1)\varepsilon_2(x_2)}{[f_1\varepsilon_1(x_1) + f_2\varepsilon_2(x_1)]\,[f_1\varepsilon_1(x_2) + f_2\varepsilon_2(x_2)]}.$$

Integrating over x_1 and x_2 does not make this equal to unity.

Relation Between Integral and Differential Correlations

- Definition: Single & Pair Densities

Histogram — number of singles per event normalized per bin width

Single Density: $\rho_1(\phi_i,\eta_i) = \langle N(\phi_i,\eta_i)\rangle / \Delta\phi\Delta\eta$

Histogram — number of pairs per event normalized per bin width

Pair Density: $\rho_2(\phi_1,\eta_1,\phi_2,\eta_2) = \langle N(\phi_1,\eta_1)N(\phi_2,\eta_2)\rangle / \Delta\phi^2\Delta\eta^2$

- Factorize average yield and kinematic dependence

Single Probability Distribution

$$\rho_1(\phi_i,\eta_i) = \langle N\rangle P_1(\phi_i,\eta_i)$$

Avg Multiplicity

$$\langle N\rangle = \int_{accept} \rho_1(\phi_i,\eta_i)\, d\phi_i\, d\eta_i$$

$$1 = \int_{accept} P_1(\phi_i,\eta_i)\, d\phi_i\, d\eta_i$$

$$\rho_2(\phi_1,\eta_1,\phi_2,\eta_2) = \langle N(N-1)\rangle P_2(\phi_1,\eta_1,\phi_2,\eta_2)$$

Avg Number of Pairs

Pair Probability Distribution

$$\langle N(N-1)\rangle = \int_{accept} \rho_2(\phi_1,\eta_1,\phi_2,\eta_2) d\phi_1 d\eta_1 d\phi_2 d\eta_2$$

$$1 = \int_{accept} P_2(\phi_1,\eta_1,\phi_1,\eta_1) d\phi_1 d\eta_1 d\phi_2 d\eta_2$$

Differential Correlation Functions

- Two-Particle Cumulant

$$C(\phi_1,\eta_1,\phi_2,\eta_2) = \rho_2(\phi_1,\eta_1,\phi_2,\eta_2) - \rho_1(\phi_1,\eta_1)\rho_1(\phi_2,\eta_2)$$

- Normalized Cumulant

$$R_2(\phi_1,\eta_1,\phi_2,\eta_2) = \frac{\rho_2(\phi_1,\eta_1,\phi_2,\eta_2) - \rho_1(\phi_1,\eta_1)\rho_1(\phi_2,\eta_2)}{\rho_1(\phi_1,\eta_1)\rho_1(\phi_2,\eta_2)} = \frac{\rho_2(\phi_1,\eta_1,\phi_2,\eta_2)}{\rho_1(\phi_1,\eta_1)\rho_1(\phi_2,\eta_2)} - 1$$

- Factorization of probability and integral

Factorizes in the absence of correlations.

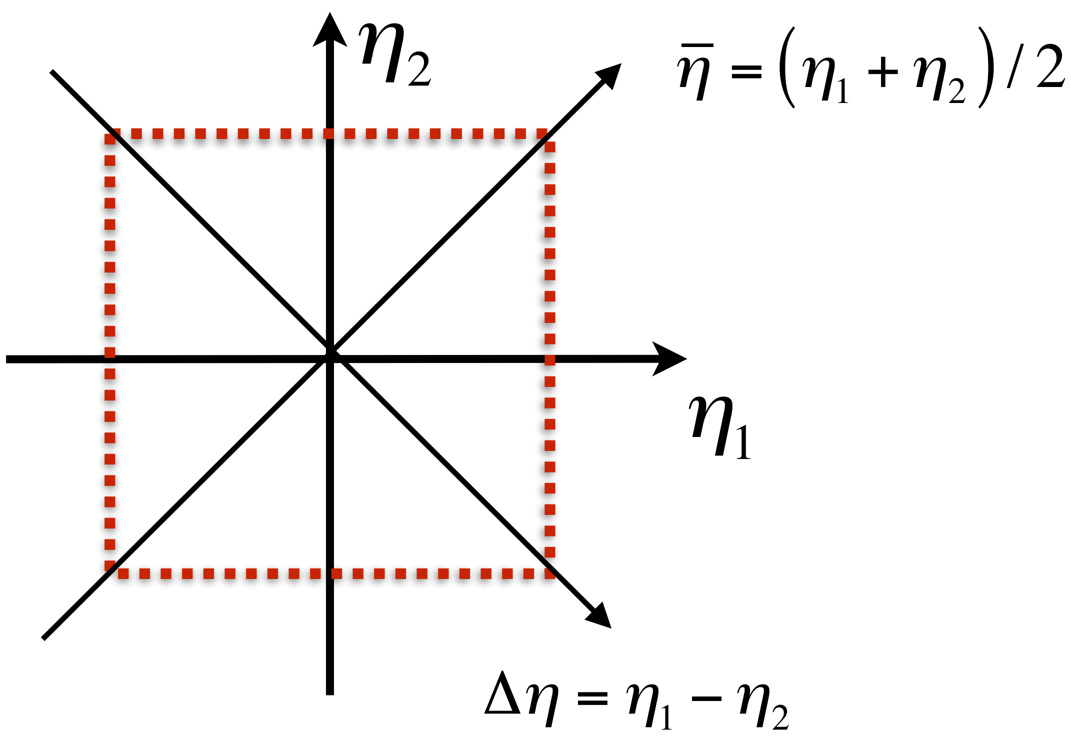
$$R_2(\phi_1,\eta_1,\phi_2,\eta_2) = \frac{\rho_2(\phi_1,\eta_1,\phi_2,\eta_2)}{\rho_1(\phi_1,\eta_1)\rho_1(\phi_2,\eta_2)} - 1 = \frac{\langle N(N-1)\rangle}{\langle N\rangle^2}\frac{P_2(\phi_1,\eta_1,\phi_2,\eta_2)}{P_1(\phi_1,\eta_1)P_1(\phi_2,\eta_2)} - 1$$

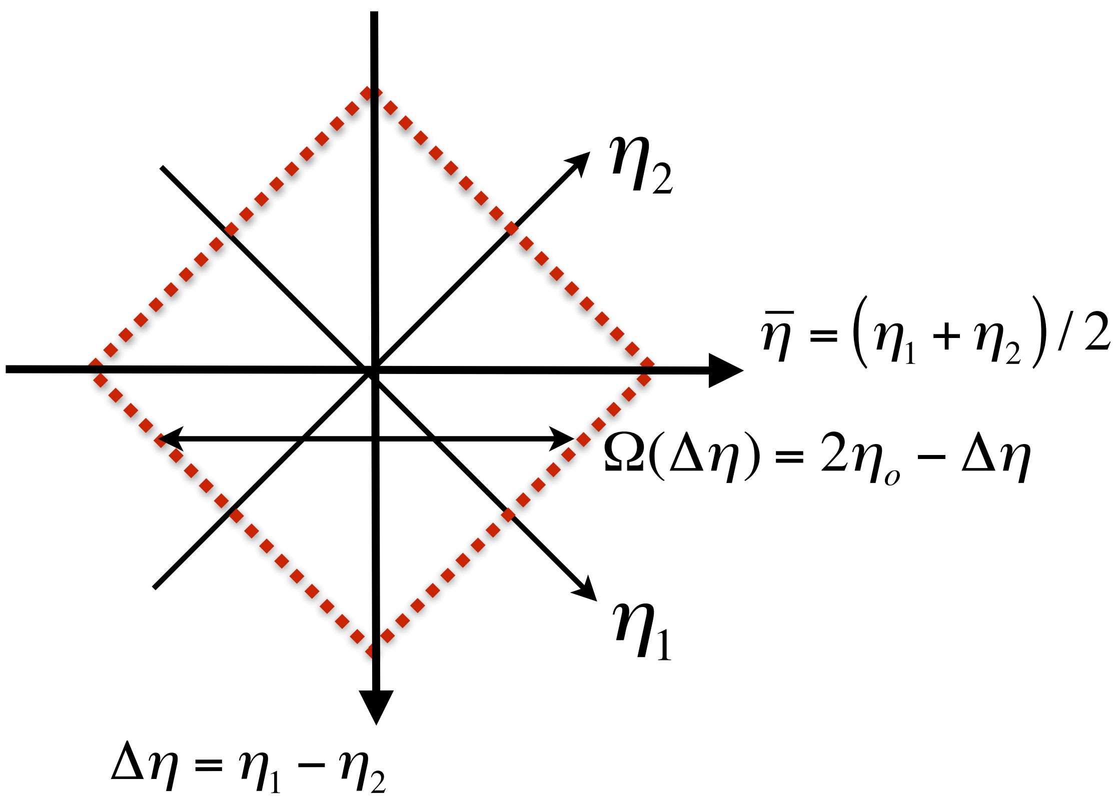
- Readily extended to 3 or more particles

Acceptance Averaging

- Two-Particle Correlation Fcts
  - Most general case: 6 coordinates
  - Most common analyses: vs. Δφ or vs. Δη or vs. Δφ, Δη or vs η1, η2

$$C(\Delta\eta) = \frac{1}{\Omega(\Delta\eta)} \int_{-(\eta_o - \Delta\eta/2)}^{\eta_o - \Delta\eta/2} C(\Delta\eta, \bar{\eta}) d\bar{\eta}$$





Note: This is an **acceptance average NOT a correction**

Efficiency & Robustness (I)

- Model the Probability of observing n particles given N (in a given "bin") were produced with binomial distribution.

$$P_{\det}(n \mid N;\varepsilon) = \frac{\varepsilon^{N}(1-\varepsilon)^{N-n}}{n!(N-n)!}$$

- Model the Probability of observing particle fluctuations…
  - Singles

$$P_M(n(\eta_1) \mid N(\eta_1);\varepsilon_1,) = \sum_{N_1=1}^{\infty} P_T(N(\eta_1)) \frac{\varepsilon_1^{N(\eta_1)}(1-\varepsilon_1)^{N(\eta_1)-n(\eta_1)}}{n(\eta_1)!(N(\eta_1)-n(\eta_1))!}$$

Measured Probability distribution

True Probability distribution

  - Pairs

$$P_M(n(\eta_1),n(\eta_2) \mid N(\eta_1),N(\eta_2);\varepsilon_1,\varepsilon_2) = \sum_{N_1,N_2=1}^{\infty} P_T(N(\eta_1),N(\eta_2)) \frac{\varepsilon_1^{N(\eta_1)}(1-\varepsilon_1)^{N(\eta_1)-n(\eta_1)}}{n(\eta_1)!(N(\eta_1)-n(\eta_1))!} \frac{\varepsilon_2^{N(\eta_2)}(1-\varepsilon_2)^{N(\eta_2)-n(\eta_2)}}{n(\eta_2)!(N(\eta_2)-n(\eta_2))!}$$

Measured Probability distribution

True Probability distribution

Efficiency & Robustness (1)

- Singles Average
  - True $$\langle N \rangle = \int P_T(N) N dN$$
  - Measured $$\langle n \rangle = \int P_M(n) n dn$$

$$\langle n \rangle = \int P_T(N) dN \int n P_{\det}(n \mid N; \varepsilon) dn = \varepsilon \int P_T(N) N \, dN$$

$$\langle n \rangle = \varepsilon \langle N \rangle$$

- Pair Averages
  - True $$\langle N_1 N_2 \rangle = \int P_p(N_1, N_2) N_1 N_2 dN_1 dN_2$$
  - Measured $$\langle n_1 n_2 \rangle = \int P_m(n_1, n_2) n_1 n_2 dn_1 dn_2$$

$$\langle n_1 n_2 \rangle = \varepsilon_1 \varepsilon_2 \langle N_1 N_2 \rangle$$

Correct for any P_T PDF
Only requires binomial sampling.

Efficiency & Robustness (III)

- Correlation function measurement
- Goal:

$$C_2^{(True)}(\eta_1,\eta_2) = \rho_2(\eta_1,\eta_2) - \rho_1(\eta_1)\rho_1(\eta_2)$$

Produced

- "Raw" Measurement

$$C_2^{(measured)}(\eta_1,\eta_2) = \frac{1}{\Delta\eta^2}\langle n(\eta_1)n(\eta_2)\rangle - \langle n(\eta_1)\rangle\langle n(\eta_2)\rangle$$

$$= \frac{1}{\Delta\eta^2}\varepsilon_1(\eta_1)\varepsilon_2(\eta_2)\left\{\langle N(\eta_1)N(\eta_2)\rangle - \langle N_1(\eta_1)\rangle\langle N_2(\eta_2)\rangle\right\}$$

Measured

- Ratio Fct

$$R_2^{(Measured)}\eta_1,\eta_2) = \frac{\langle n(\eta_1)n(\eta_2)\rangle}{\langle n(\eta_1)\rangle\langle n(\eta_2)\rangle} - 1$$

$$= \frac{\varepsilon_1(\eta_1)\varepsilon_1(\eta_2)\langle N(\eta_1)N(\eta_2)\rangle}{\varepsilon_1(\eta_1)\varepsilon_1(\eta_2)\langle N(\eta_1)\rangle\langle N(\eta_2)\rangle} - 1$$

$$= \frac{\langle N(\eta_1)N(\eta_2)\rangle}{\langle N(\eta_1)\rangle\langle N(\eta_2)\rangle} - 1$$

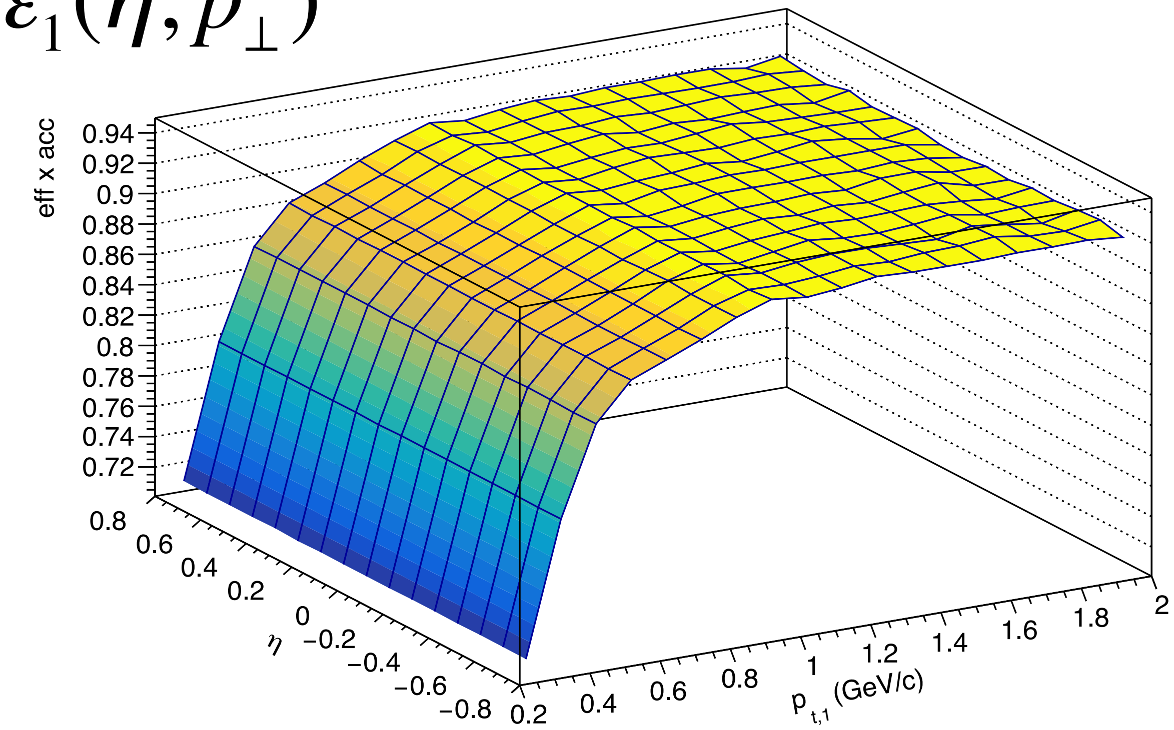
$$= R_2^{(True)}(\eta_1,\eta_2)$$

Efficiencies cancel >>> Robust Observable

Efficiency vs. Momentum Coordinates

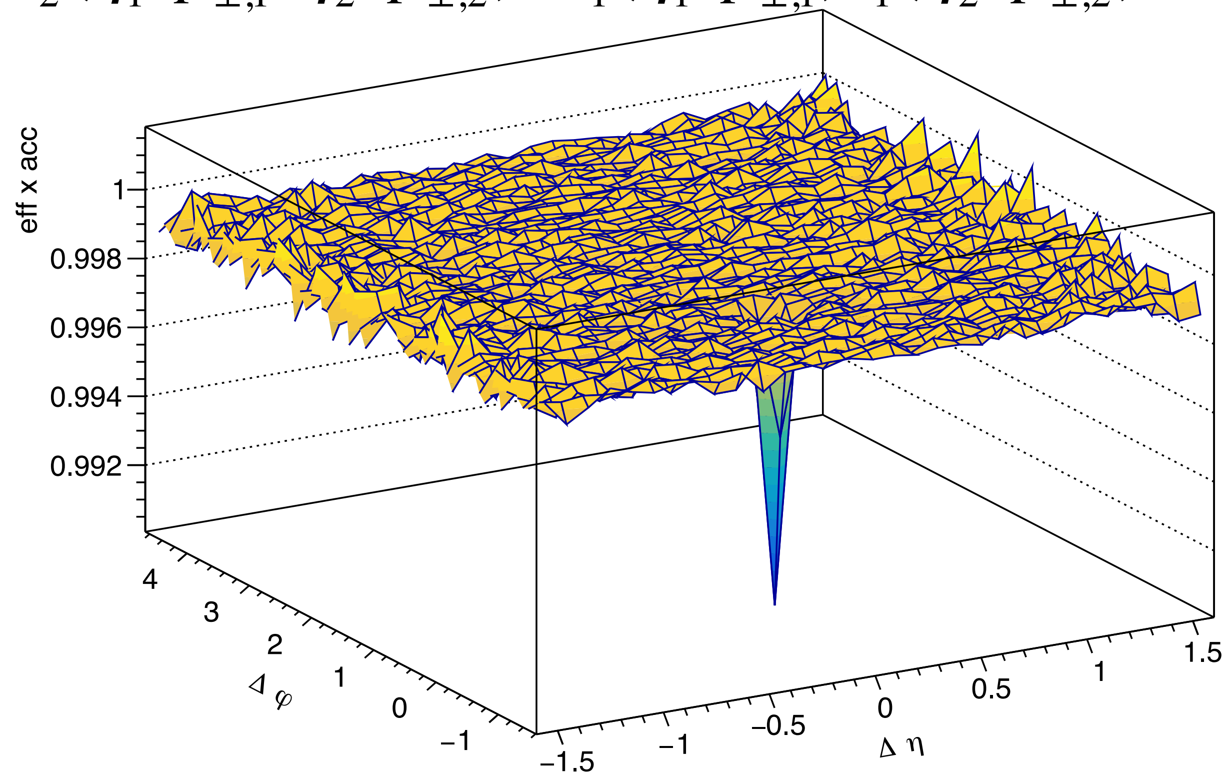
- Example for ALICE detector
- Determined from HIJING events propagated through detector simulation with GEANT and detector response simulator

$\varepsilon_1(\eta, p_\perp)$



Single Particle Efficiency

$\varepsilon_2(\eta_1, p_{\perp,1}, \eta_2, p_{\perp,2}) / \varepsilon_1(\eta_1, p_{\perp,1})\varepsilon_1(\eta_2, p_{\perp,2})$



Pair Efficiency

Folding of Singles vs Event Mixing

- Ratio R requires product of single yields
  - Can be obtained from actual singles

$$R_M(\eta_1,\eta_2) = \frac{\langle n_1(\eta_1) n_2(\eta_2)\rangle}{\langle n_1(\eta_1)\rangle\langle n_2(\eta_2)\rangle} - 1$$

No event mixing required

  - Can be obtained from mixed events

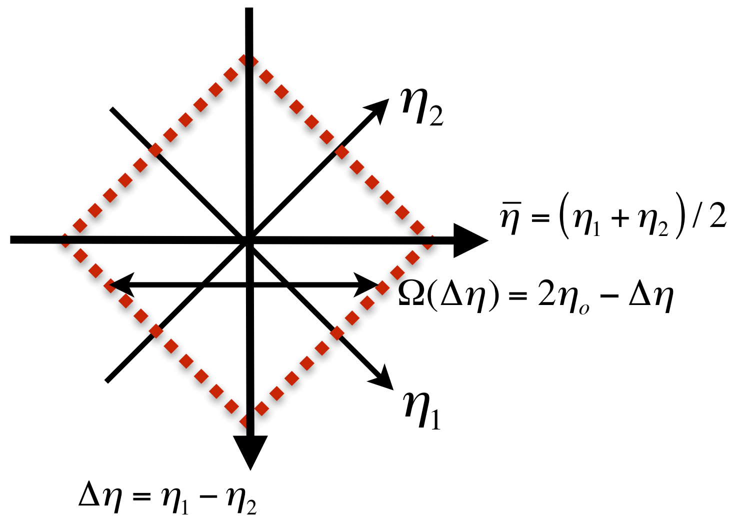
$$R_m(\eta_1,\eta_2) = \frac{\langle n_1 n_2(\eta_1,\eta_2)\rangle}{\langle n_1(\eta_1)\rangle\langle n_2(\eta_2)\rangle} = \frac{\langle n_1 n_2(\eta_1,\eta_2)\rangle_{\text{same}}}{\langle n_1 n_2(\eta_1,\eta_2)\rangle_{\text{mixed}}} - 1$$

Greater flexibility w/ cuts

Two Methods

- Method 1: Ratio of averages (Common Approach)
  - Measure pair yields (same and mixed) directly vs Δη.
  - Calculate R(Δη) by taking the ratio of same to mixed.

$$R_M(\Delta\eta) = \frac{\frac{1}{\Omega(\Delta\eta)}\int_{accept} \rho_2(\Delta\eta,\overline{\eta})\,d\overline{\eta}}{\frac{1}{\Omega(\Delta\eta)}\int_{accept} \rho_1 \otimes \rho_1(\Delta\eta,\overline{\eta})\,d\overline{\eta}} - 1$$



- Method 2: Average of Ratio
  - Measure R(η_1,η_2) by taking the ratio of same to mixed.
  - Average out $\overline{\eta}$ dependence, i.e. project onto Δη to get R(Δη)

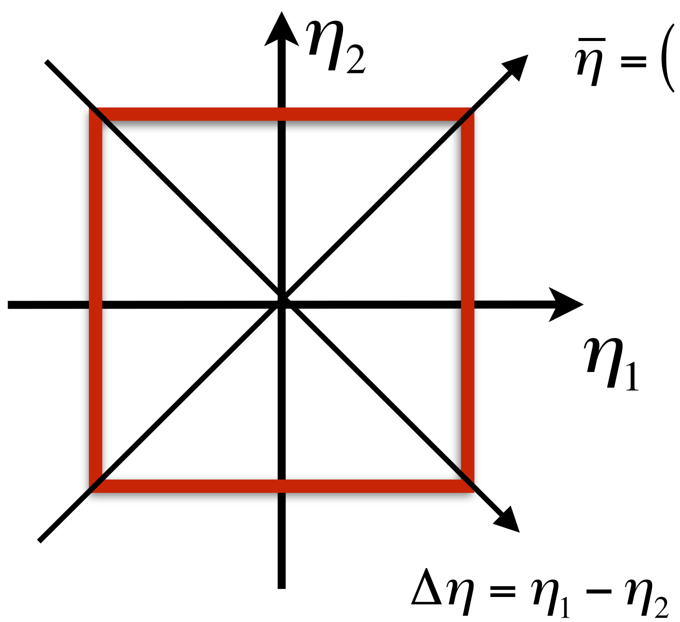
$$R_M(\Delta\eta) = \frac{1}{\Omega(\Delta\eta)}\int_{accept} R_2(\Delta\eta,\overline{\eta})\,d\overline{\eta} = \frac{1}{\Omega(\Delta\eta)}\int_{accept}\left(\frac{\rho_2(\Delta\eta,\overline{\eta})}{\rho_1 \otimes \rho_1(\Delta\eta,\overline{\eta})} - 1\right)d\overline{\eta}$$

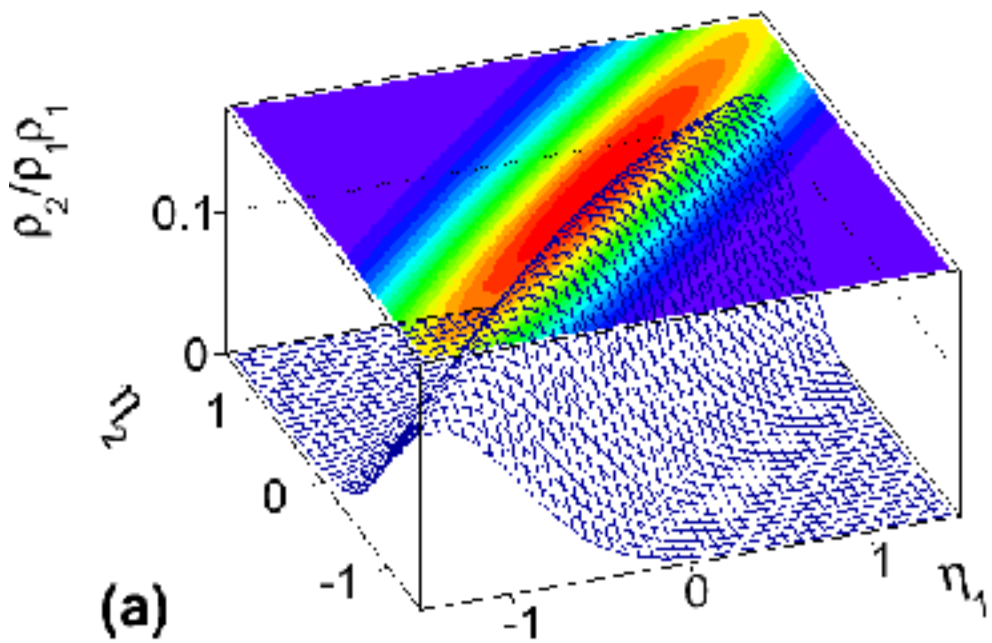
Method 1 vs. Method 2: Correlation Model

- Correlation Model:
  - Longitudinal Model w/ Two-particle emission correlated vs.

$$C(\Delta\eta,\bar{\eta}) \propto \exp\left(-\frac{\Delta\eta^2}{2\sigma_{\Delta\eta}^2}\right)\exp\left(-\frac{\bar{\eta}^2}{2\sigma_{\bar{\eta}}^2}\right)$$

- Assumed factorization of the dependence on the relative and average pseudorapidity.
- **Factorization may not be realized in practice**





Method 1 vs. Method 2: Efficiency Model

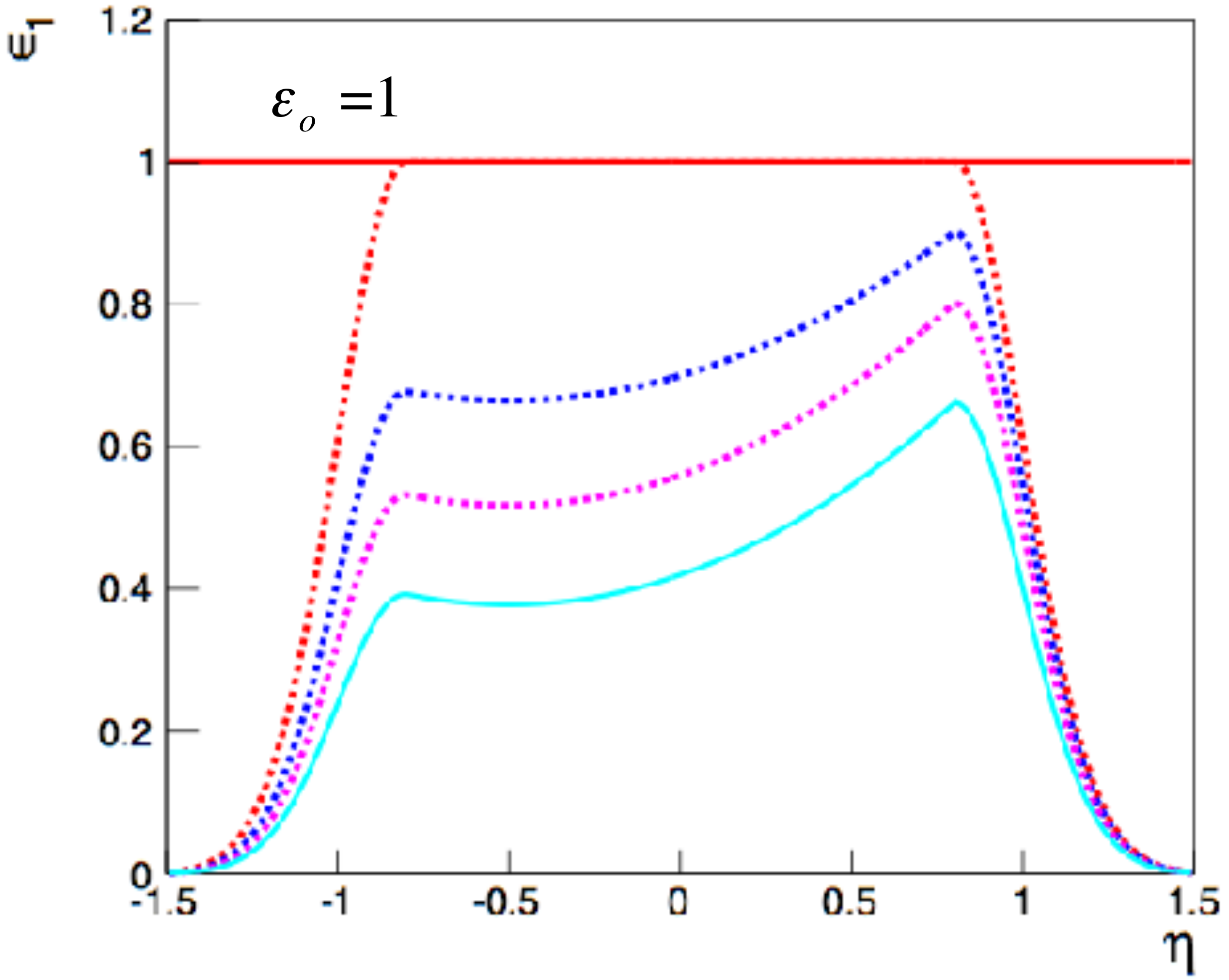
- Use a simple but non trivial correlation model
- Use a simple model of the detection efficiency and edge effects.

$$\varepsilon(\eta) = \varepsilon_q(\eta)\exp\left(-\frac{(\eta-\eta_<)^2}{2\sigma_\varepsilon^2}\right) \quad \text{for} \quad \eta < \eta_<$$

$$= \varepsilon_q(\eta) \quad \text{for} \quad \eta_< < \eta < \eta_>$$

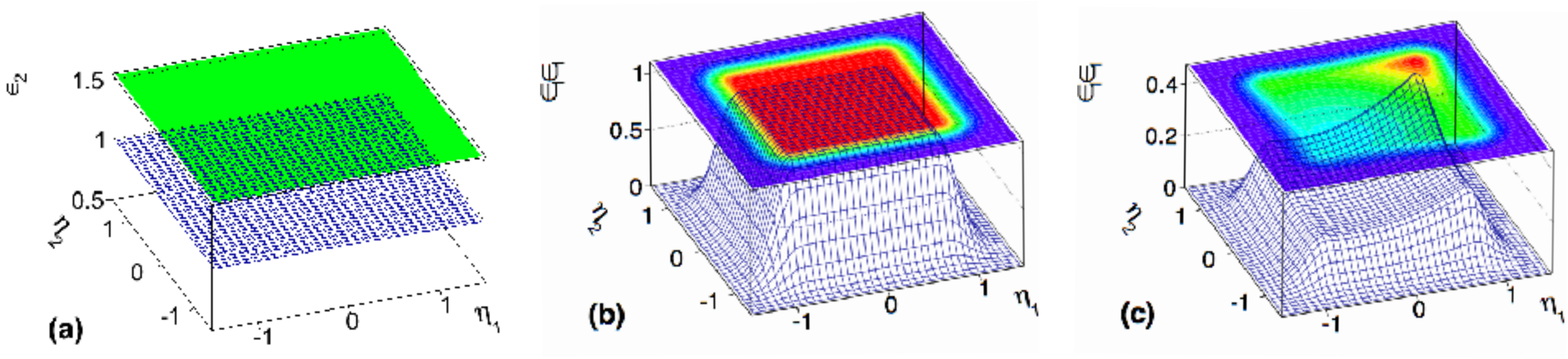
$$= \varepsilon_q(\eta)\exp\left(-\frac{(\eta-\eta_>)^2}{2\sigma_\varepsilon^2}\right) \quad \text{for} \quad \eta > \eta_>$$

$$\varepsilon_q(\eta) = 1 + \alpha(\eta-\eta_o) + \beta(\eta-\eta_o)^2$$

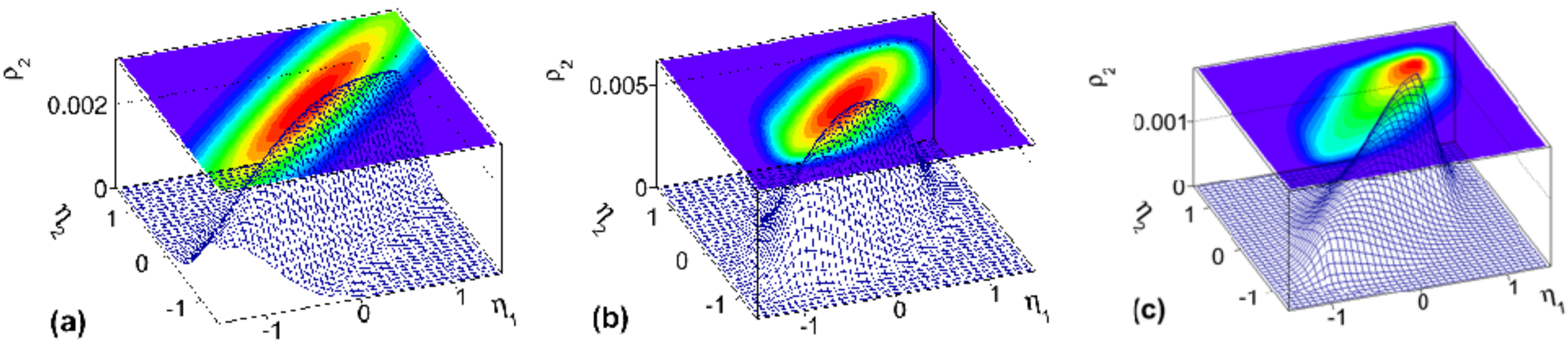


Efficiency, Pair Yield

- Efficiency

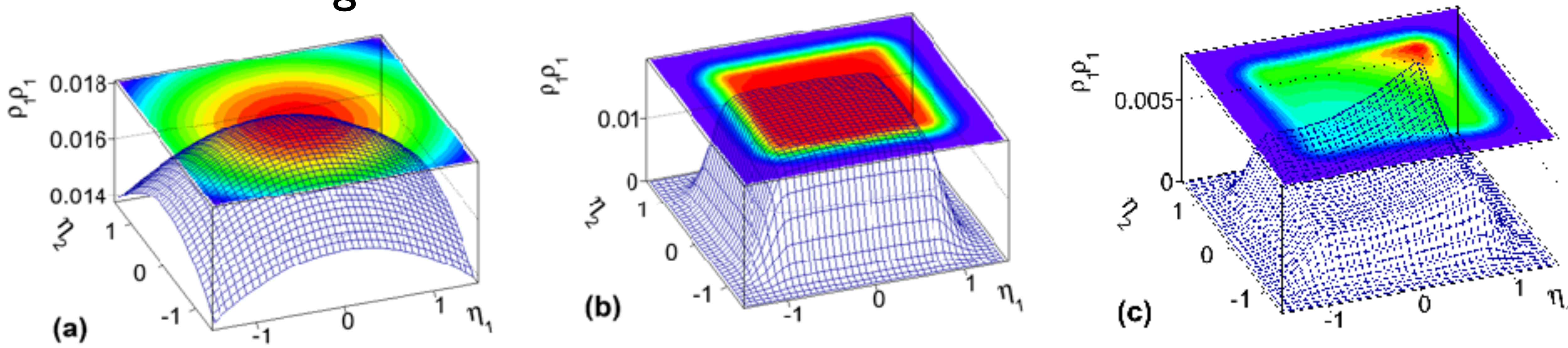


- Pair Yield



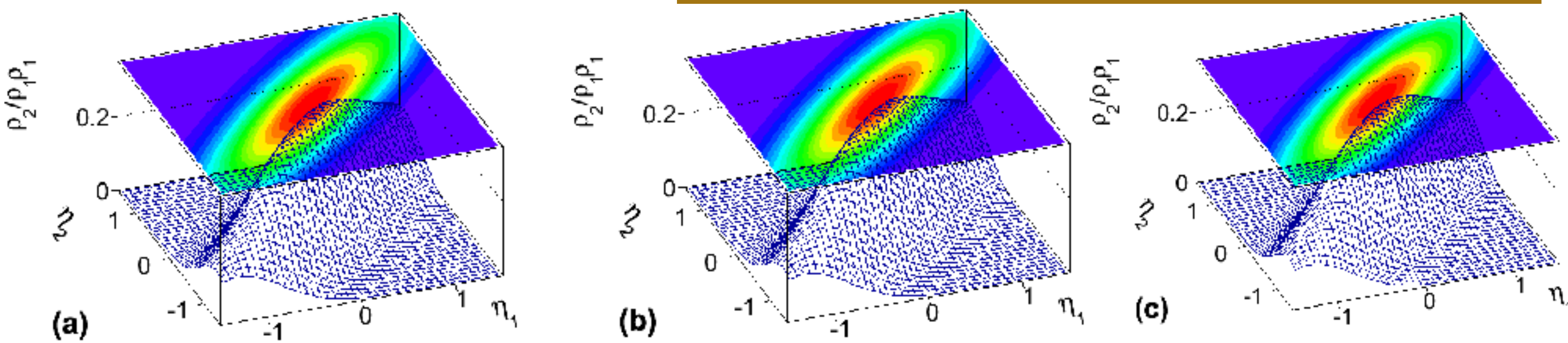
Method 2: Results

Product of singles

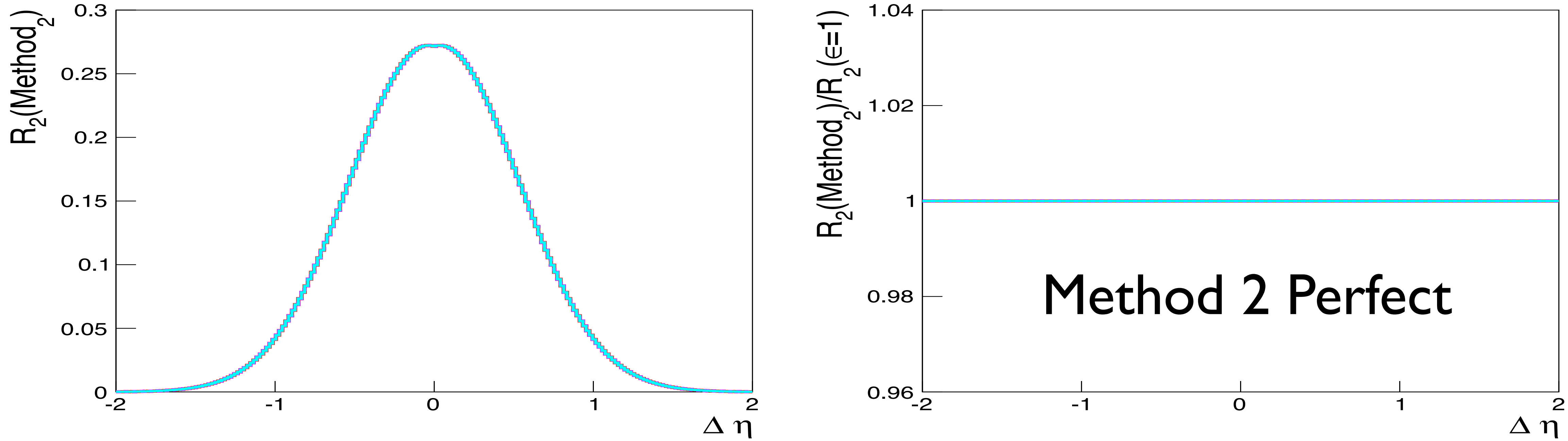


R2 (Method 2)

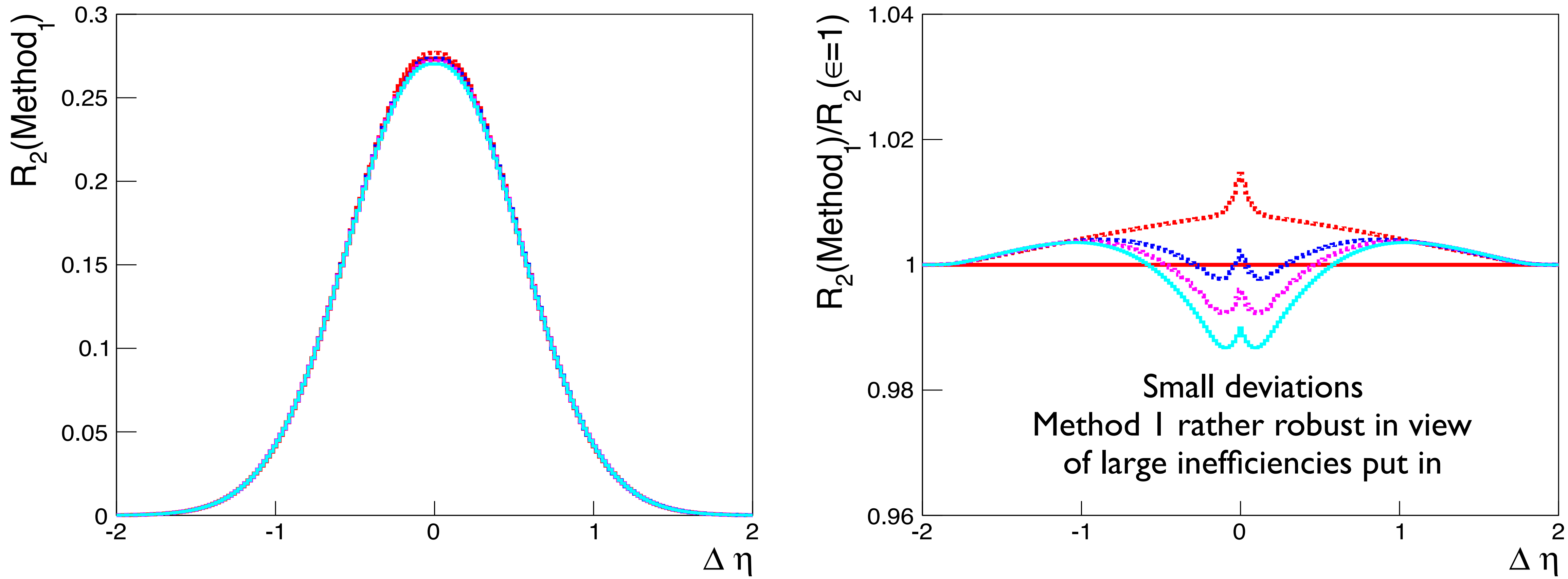
Perfect Reconstruction for any factorized efficient model w/ sufficient statistics



R2(Δη)   Method 2



R2(Δη)   Method 1



Why?

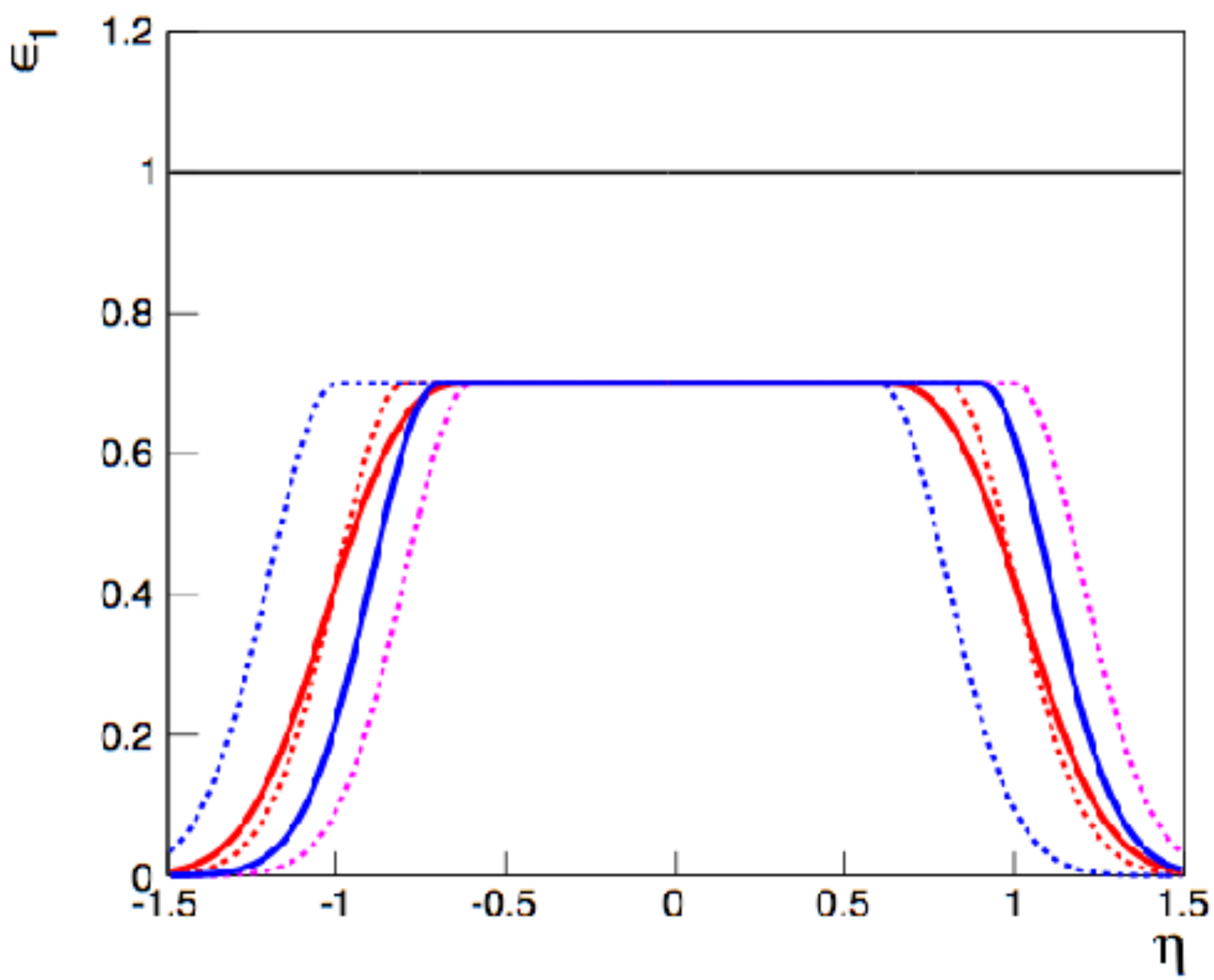
$$R_2(\Delta\eta)^{Method1} = \frac{\int g(\Delta\eta,\overline{\eta}) R_2^{true}(\Delta\eta,\overline{\eta}) d\overline{\eta}}{\int g(\Delta\eta,\overline{\eta}) d\overline{\eta}}$$

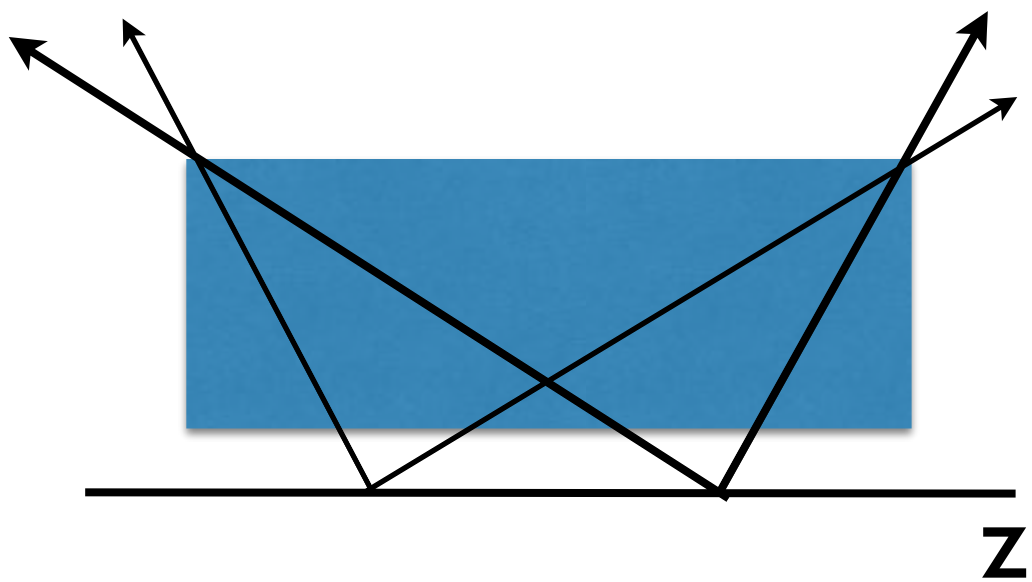
$$g(\Delta\eta,\overline{\eta}) = \epsilon_1 \times \epsilon_1 \times \rho_1 \times \rho_1(\Delta\eta,\overline{\eta})$$

- If efficiency, yield, or correlation varies with avg-rapidity, then g or R2 cannot be factorized out of the integrals.
  - The numerator and denominator are in general NOT equal.
- Method 1 is only approximately robust - for slow varying functions
- Note: not a problem in azimuthal correlation because of periodic boundary conditions.

Dependence on z-vertex

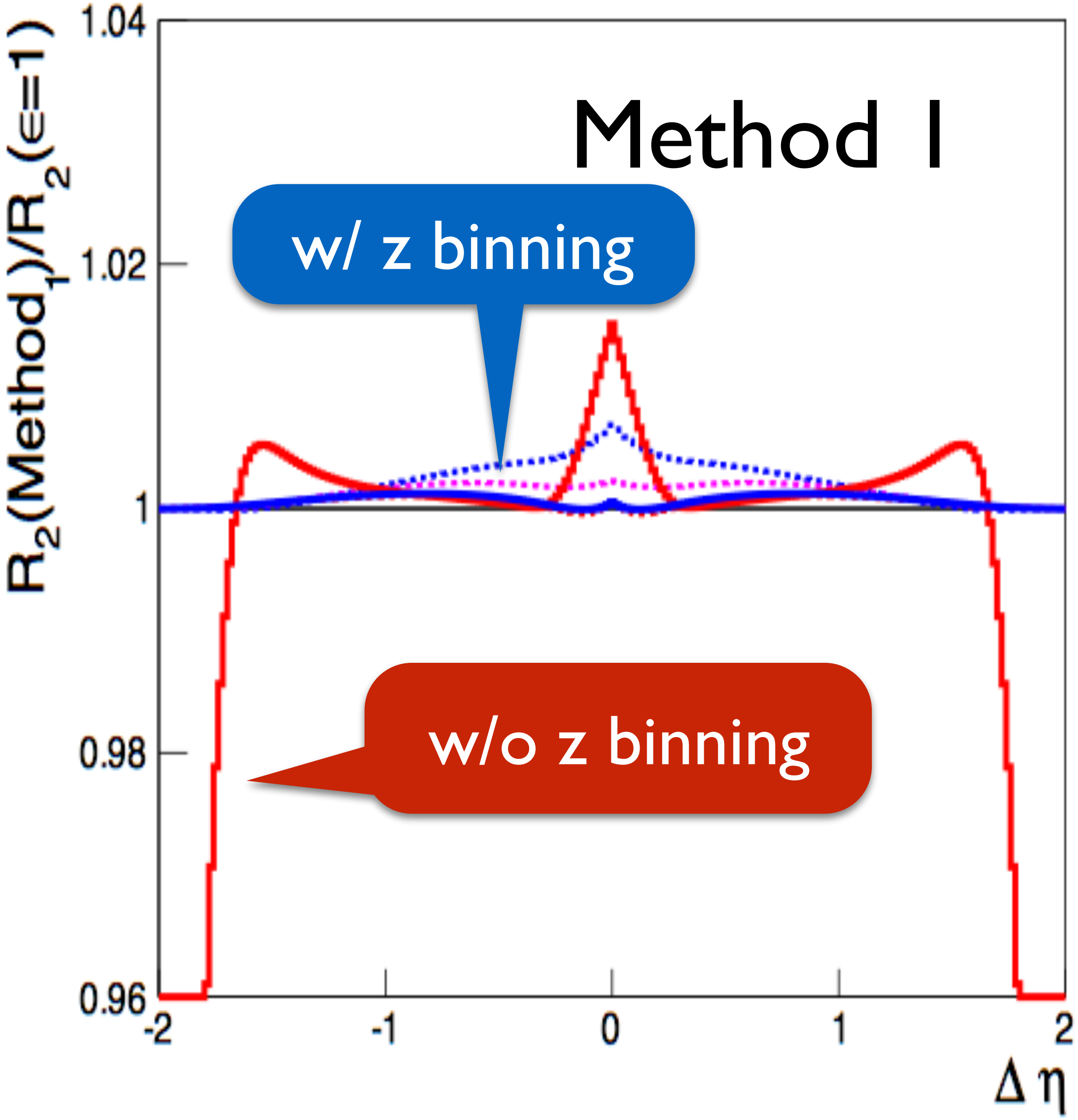
- ALICE, STAR Acceptances are functions of the vertex position.
- Use a simple model as before...

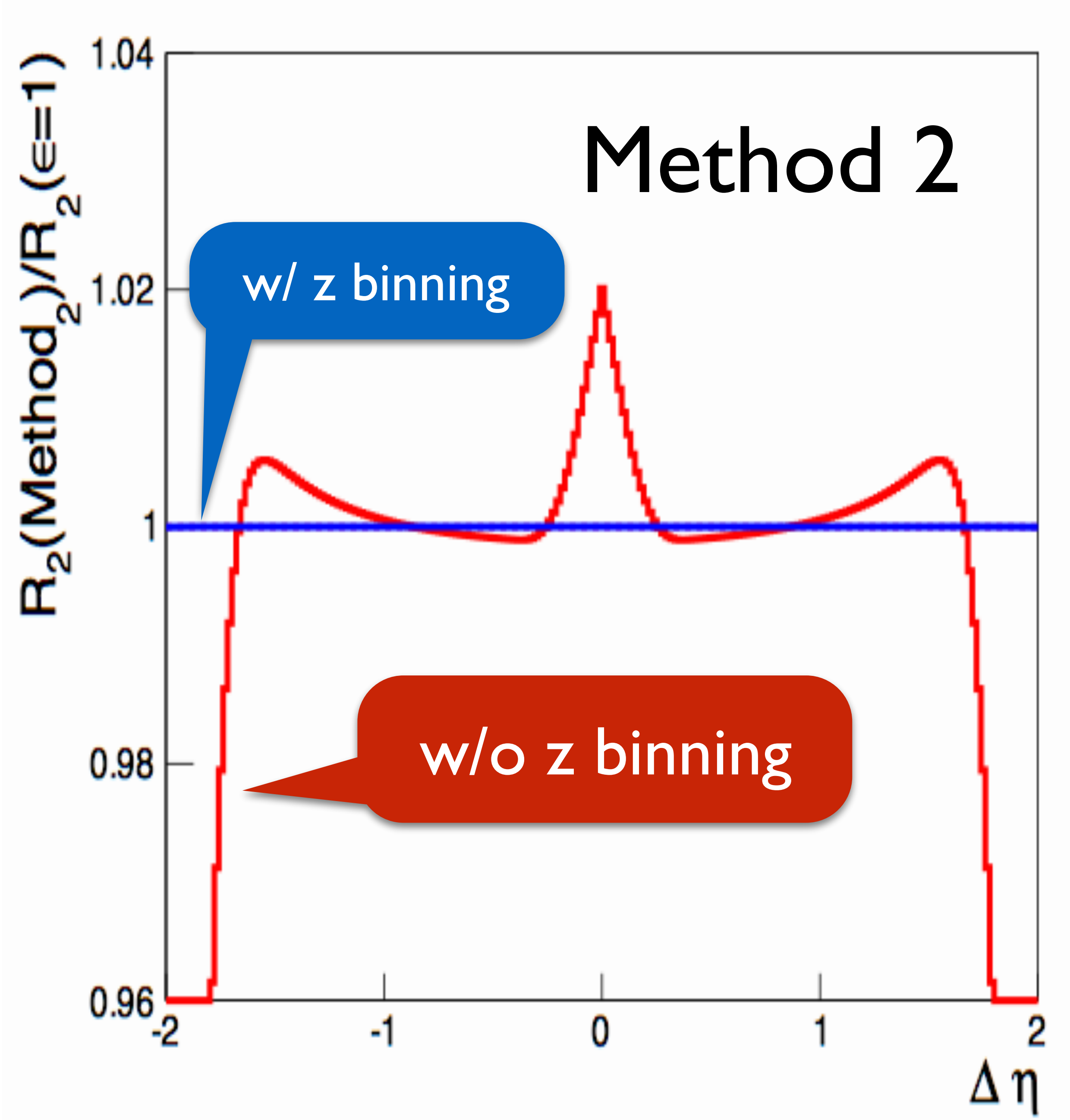




Method 1 and 2

Efficiency dependence on "z-vertex", with gaussian edges, but **quadratic** dependence on eta in the fiducial volume.





Both methods fail if efficiency is dependent on "z".
Approximate recovery with fine z-bins using Method 1
Complete recovery with fine z-bins using Method 2

Part IV: Summary

- Method 2 Robust
  - unless efficiency has dependence on z-vertex
  - but recovery possible for analysis in narrow z-bins
- Method 1 Only Approximately Robust
  - Robustness lost if singles, correlation, or efficiency are function of avg-eta
  - Approximate Robustness lost if dependence on z-vertex
  - "Partial" recovery possible for analysis in narrow z-bins

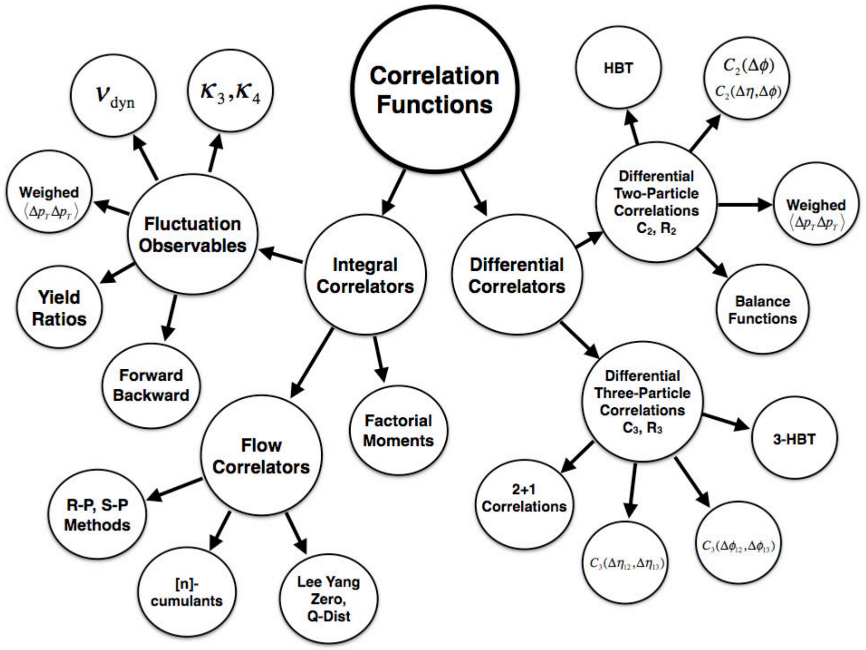
- **Bigger Point:**
  - With differential correlations, it is possible to identify detector features more readily than with integral correlations.
  - Integral correlations average over detector issues, they DO NOT eliminate them.

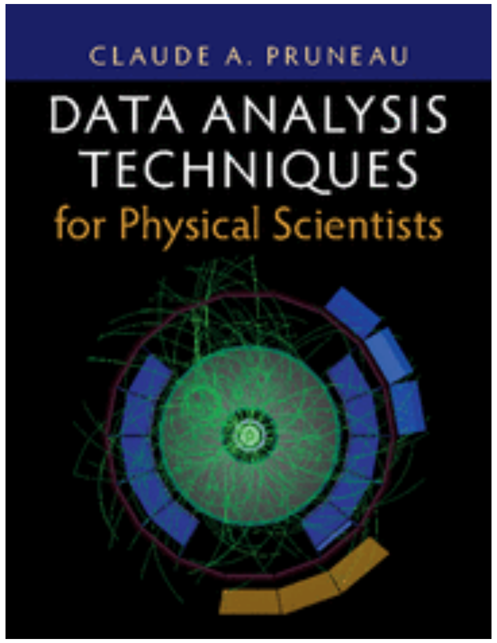
Summary/Conclusions

Measurements in Heavy Ion Collisions predominantly based on two and multi-Particle Correlation Function

Sensitive to a wide range of phenomena or aspects of collisions

Here focused on the structure and basic properties of correlation function and their measurement.





For more information, and to order, visit:

www.cambridge.org/9781108416788

and enter the code PRUNEAU2017 at the checkout