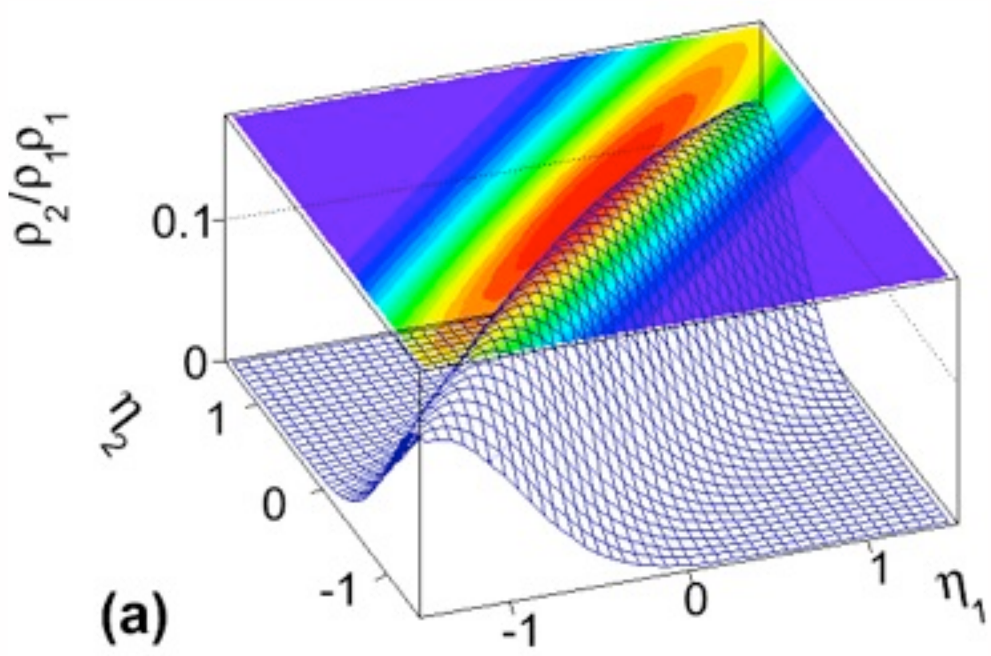
Wayne State University
College of Liberal Arts & Sciences
Department of Physics and Astronomy

Correcting Correlation Function Measurements

Claude A. Pruneau
with Shantam Ravan, Prabhat Pujahari, Sidharth Prasad

Available: arXiv:1311.3915
Accepted in PRC

- Acceptance
- Efficiency
- And some other instrumental effects



Correlation - Definition

- Two "events" A and B can be considered independent (statistical independence) if their joint probability factorizes.

$$P(A,B) = P(A) \times P(B)$$

- In the context of particle production measurements, this translates in having the probability (and yield) of particle production at two points of phase space being factorizable.

$$P(N(\phi_1,\eta_1),N(\phi_2,\eta_2)) = P(N(\phi_1,\eta_1)) \times P(N(\phi_2,\eta_2))$$

- Practical evaluation: Covariance of yields produced at 2 points of phase space

$$Cov[N(\phi_1,\eta_1),N(\phi_2,\eta_2)] = \langle N(\phi_1,\eta_1)N(\phi_2,\eta_2)\rangle - \langle N(\phi_1,\eta_1)\rangle\langle N(\phi_2,\eta_2)\rangle$$

Correlation Function - Definition (1)

- Covariance of yields vs. 2 points of phase space

$$C(\phi_1,\eta_1,\phi_2,\eta_2)=\langle N(\phi_1,\eta_1)N(\phi_2,\eta_2)\rangle-\langle N(\phi_1,\eta_1)\rangle\langle N(\phi_2,\eta_2)\rangle$$

  - <> represent average over ensemble of events
  - Null value at one given point does not guarantee statistical independence.

- Correlation Fct as a ratio

$$R(\phi_1,\eta_1,\phi_2,\eta_2)=\frac{C(\phi_1,\eta_1,\phi_2,\eta_2)}{\langle N(\phi_1,\eta_1)\rangle\langle N(\phi_2,\eta_2)\rangle}=\frac{\langle N(\phi_1,\eta_1)N(\phi_2,\eta_2)\rangle}{\langle N(\phi_1,\eta_1)\rangle\langle N(\phi_2,\eta_2)\rangle}-1$$

  - Same "information"

- Convenient Notation

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

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

Correlation Function - Definition (2)

- Correlation Function

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

- As a ratio:

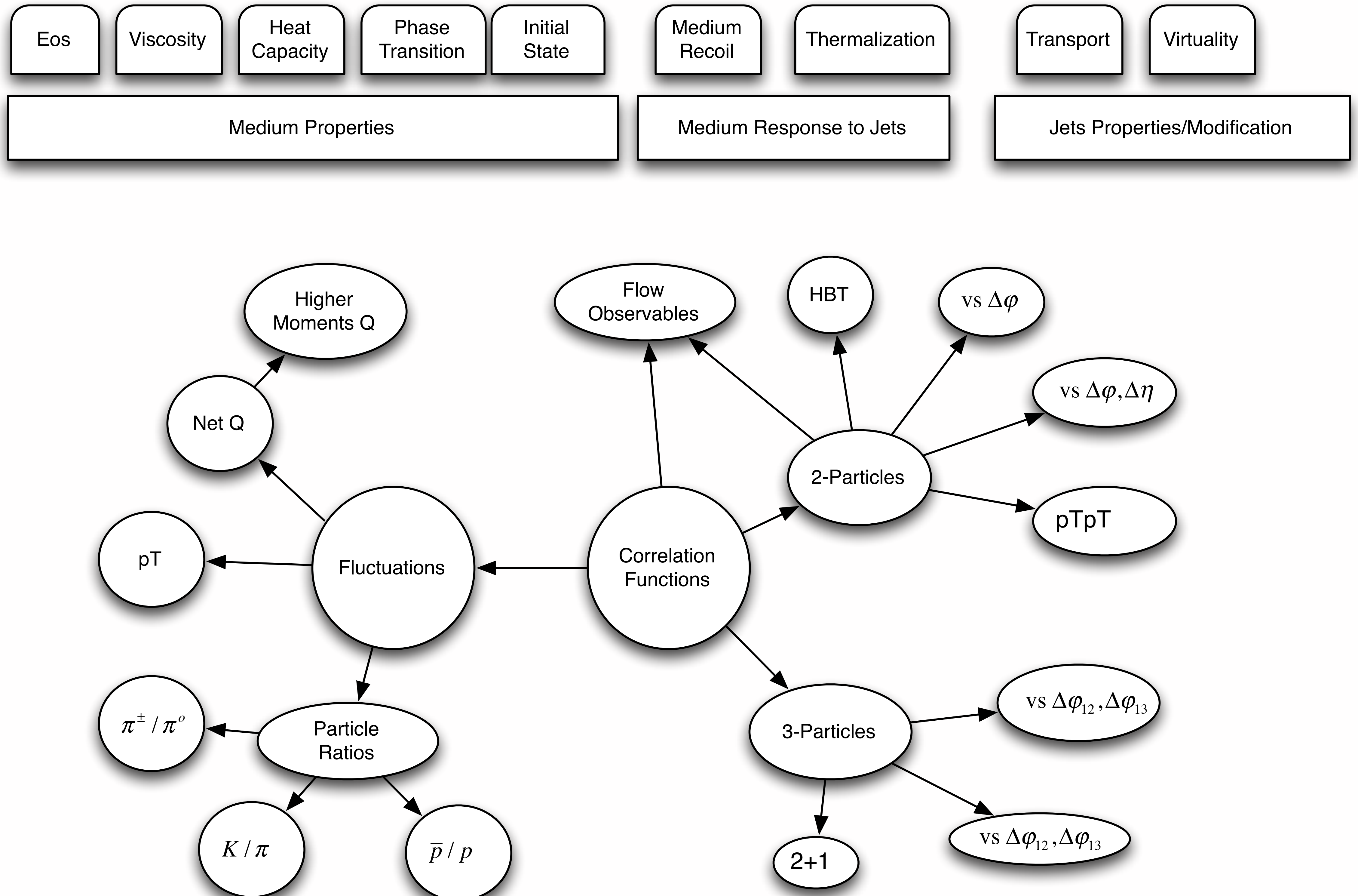
$$R(\phi_1,\eta_1,\phi_2,\eta_2) = \frac{C(\phi_1,\eta_1,\phi_2,\eta_2)}{\rho_1(\phi_1,\eta_1)\rho_1(\phi_2,\eta_2)} = \frac{\rho_2(\phi_1,\eta_1,\phi_2,\eta_2)}{\rho_1(\phi_1,\eta_1)\rho_1(\phi_2,\eta_2)} - 1$$

- Alternatively

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

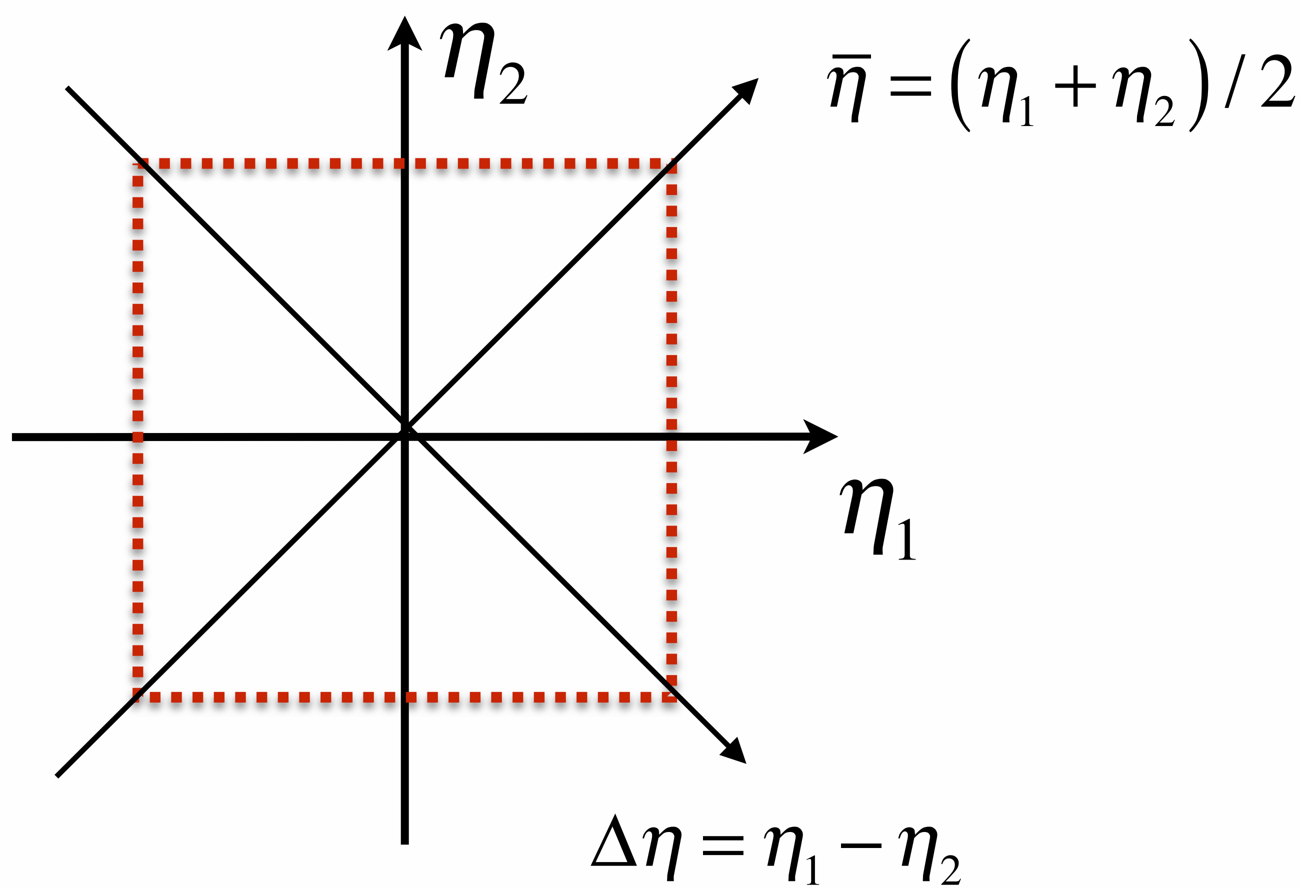
Same information in C or R

The many facets of correlation functions



“Projections”

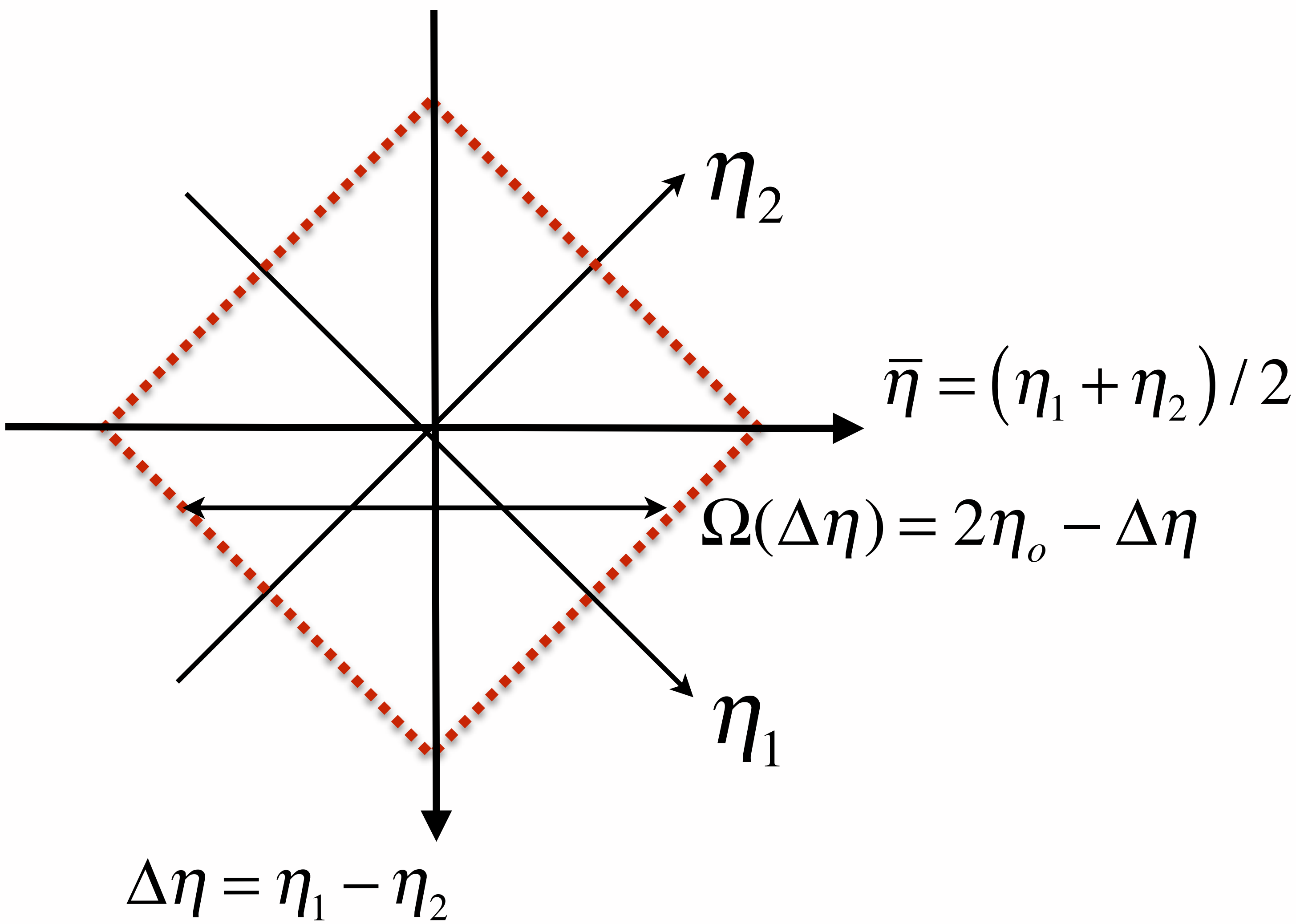
- Two-Particle Correlation Fcts
  - Most general case: 6 coordinates
  - Most common analyses:
    - vs. Δφ
    - vs. Δη
    - vs. Δφ, Δη
  - Less common
    - vs η1, η2



Longitudinal Correlation Function

- Average of the rapidity-average

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



Per Trigger Correlation

- Correlation yield per trigger

$$C_{PerTrig}(\Delta\eta) = \frac{1}{N_{Trig}} C_{PerTrig}(\Delta\eta) = \frac{1}{\int \rho_{Trig}\, d\eta} C_{PerTrig}(\Delta\eta)$$

Efficiency & Robustness (1)

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

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

- Produced Multiplicity Distribution (in given bin): $P_p(N)$
- Measured Multiplicity Distribution (in given bin): $P_m(n)$
- Average Yields

Produced

$$\langle N \rangle = \int P_p(N) N dN$$

Measured

$$\langle n \rangle = \int P_m(n) n dn$$

$$P_m(n) = \int P_p(N) P_{\det}(n \mid N;\varepsilon) dN$$

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

$$\boxed{\langle n \rangle = \varepsilon \langle N \rangle}$$

Efficiency & Robustness (II)

- Model the Probability of observing n_1 and n_2 particles given N_1 and N_2 (in two given "bins") similarly...

$$P_{\mathrm{det}}(n_1, n_2 \mid N_1, N_2; \varepsilon_1, \varepsilon_2) = P_{\mathrm{det}}(n_1 \mid N_1; \varepsilon_1) P_{\mathrm{det}}(n_2 \mid N_2; \varepsilon_2)$$
$$= \frac{\varepsilon_1^{N_1}(1-\varepsilon_1)^{N_1-n_1}}{n_1!(N_1-n_1)!} \frac{\varepsilon_2^{N_2}(1-\varepsilon_2)^{N_2-n_2}}{n_2!(N_2-n_2)!}$$

- Average Pair Yields

Produced

$$\langle N_1 N_2 \rangle = \int P_p(N_1, N_2) N_1 N_2 dN_1 dN_2$$

Measured

$$\langle n_1 n_2 \rangle = \int P_m(n_1, n_2) n_1 n_2 dn_1 dn_2$$

$$\boxed{\langle n_1 n_2 \rangle = \varepsilon_1 \varepsilon_2 \langle N_1 N_2 \rangle}$$

Efficiency & Robustness (III)

- Correlation function measurement
- Goal:

$$C_p(\eta_1,\eta_2) = \rho_2(\eta_1,\eta_2) - \rho_1(\eta_1)\rho_1(\eta_2)$$

Produced

- "Raw" Measurement

$$\begin{aligned} C_m(\eta_1,\eta_2) &= \langle n_1 n_2(\eta_1,\eta_2)\rangle - \langle n_1(\eta_1)\rangle\langle n_2(\eta_2)\rangle \\ &= \varepsilon_1(\eta_1)\varepsilon_2(\eta_2)\left\{\langle N_1 N_2(\eta_1,\eta_2)\rangle - \langle N_1(\eta_1)\rangle\langle N_2(\eta_2)\rangle\right\} \end{aligned}$$

Measured

- Ratio Fct

$$\begin{aligned} R_m(\eta_1,\eta_2) &= \frac{\langle n_1 n_2(\eta_1,\eta_2)\rangle}{\langle n_1(\eta_1)\rangle\langle n_2(\eta_2)\rangle} - 1 \\ &= \frac{\varepsilon_1(\eta_1)\varepsilon_2(\eta_2)\langle N_1 N_2(\eta_1,\eta_2)\rangle}{\varepsilon_1(\eta_1)\varepsilon_2(\eta_2)\langle N_1(\eta_1)\rangle\langle N_2(\eta_2)\rangle} - 1 \\ &= \frac{\langle N_1 N_2(\eta_1,\eta_2)\rangle}{\langle N_1(\eta_1)\rangle\langle N_2(\eta_2)\rangle} - 1 \\ &= R_p(\eta_1,\eta_2) \end{aligned}$$

Efficiencies cancel >>> Robust Observable

Note

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

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

  - Can be obtained from mixed events

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

No event mixing required

Greater flexibility w/ cuts

Two Methods

- Method 2:
  - Measure single and pair yields vs η_1 and η_2,
  - Compute $R(\eta_1,\eta_2)$
  - Average out $\overline{\eta}$ dependence, i.e. project onto $\Delta\eta$ to get $R(\Delta\eta)$
- Method 1: (Standard Method)
  - Measure pair yields (same and mixed) directly vs $\Delta\eta$.
  - Calculate $R(\Delta\eta)$ by taking the ratio of same to mixed.

Comparison of the two Methods

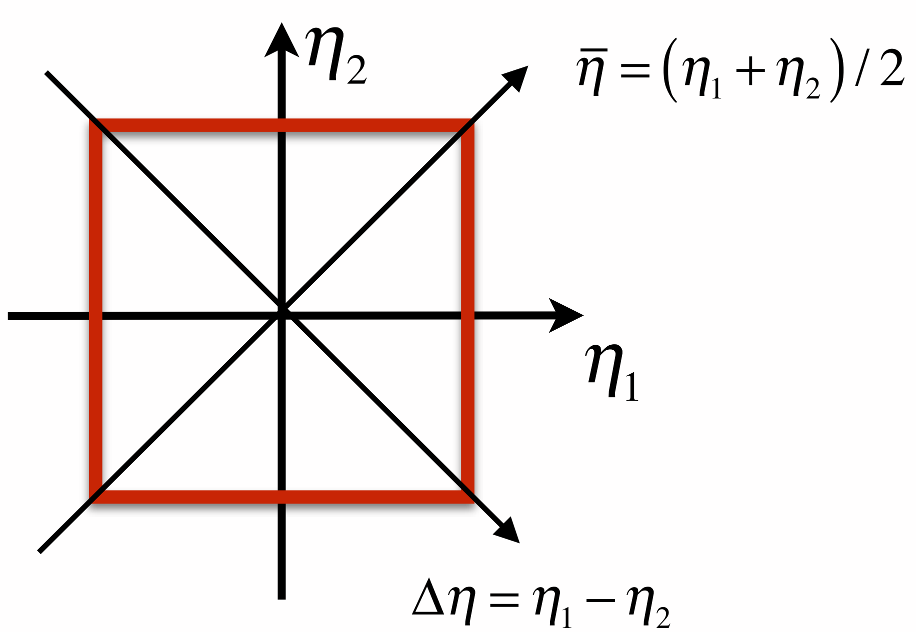
- Use a simple but non trivial correlation model
- Use a simple model of the detection efficiency and edge effects.

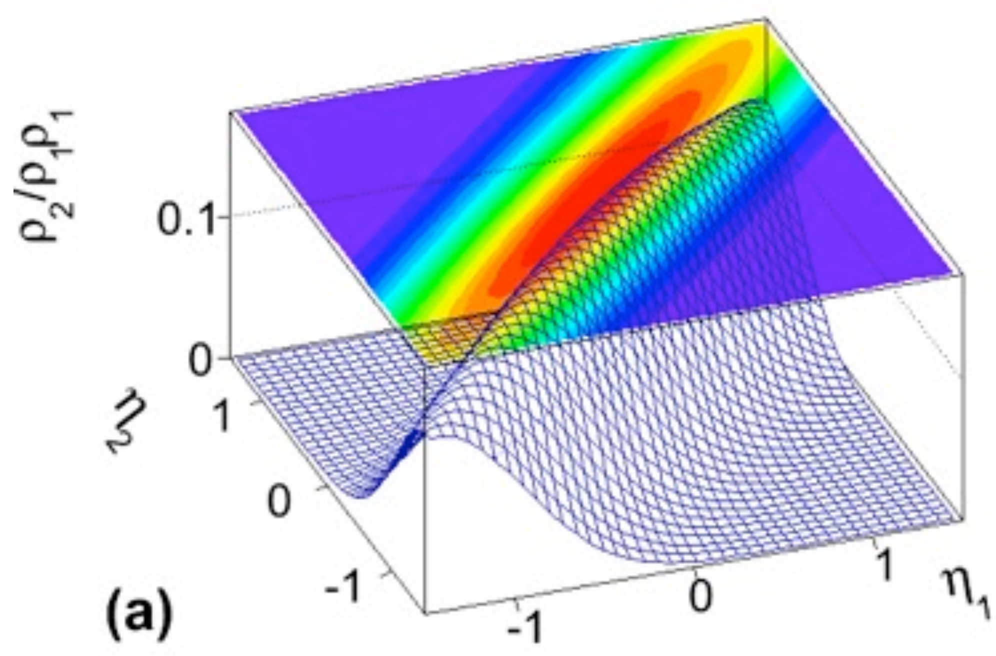
Illustrative Correlation Model

- Longitudinal Model w/ Two-particle emission correlated vs. relative and average rapidity.

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

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



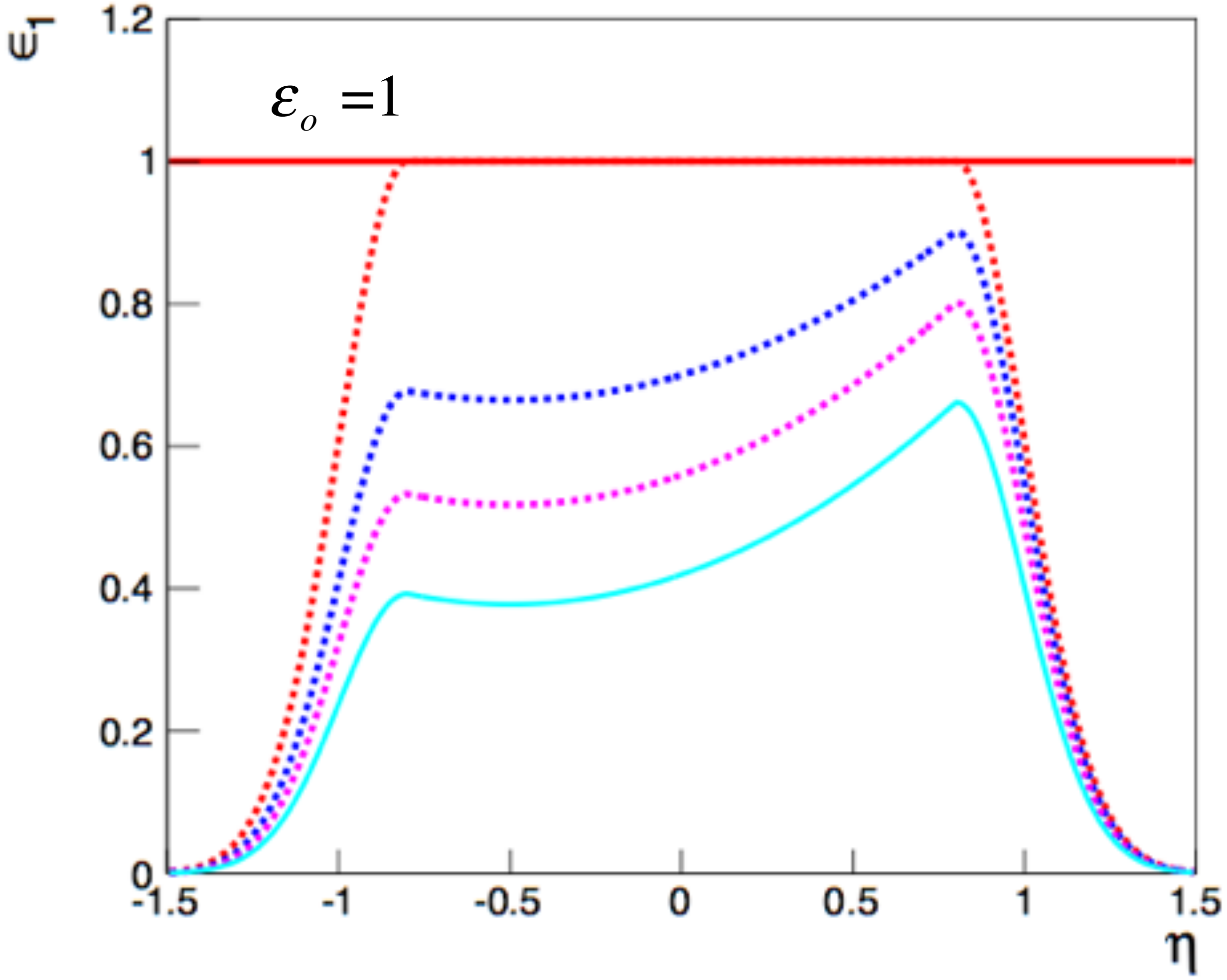


Efficiency Model

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

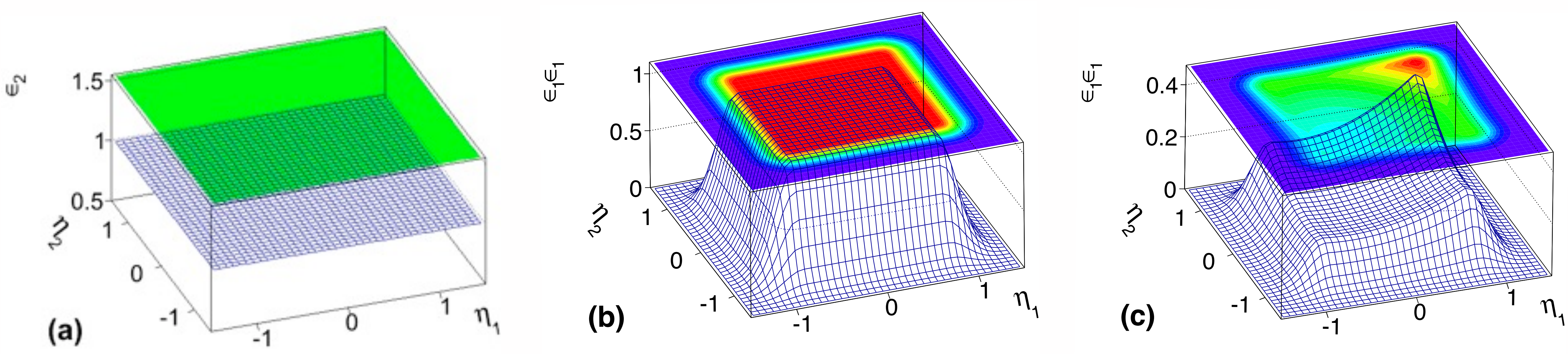
$$
\begin{aligned}
\varepsilon(\eta) &= \varepsilon_q(\eta)\exp\left(-\frac{(\eta-\eta_<)^2}{2\sigma_\varepsilon^2}\right) && \text{for} \quad \eta < \eta_< \\
&= \varepsilon_q(\eta) && \text{for} \quad \eta_< < \eta < \eta_> \\
&= \varepsilon_q(\eta)\exp\left(-\frac{(\eta-\eta_>)^2}{2\sigma_\varepsilon^2}\right)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

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

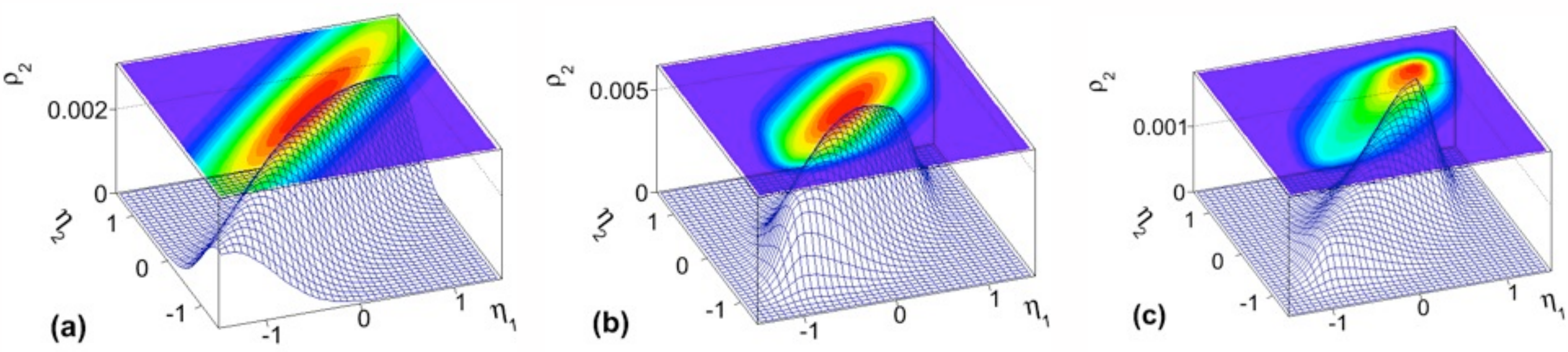


Efficiency

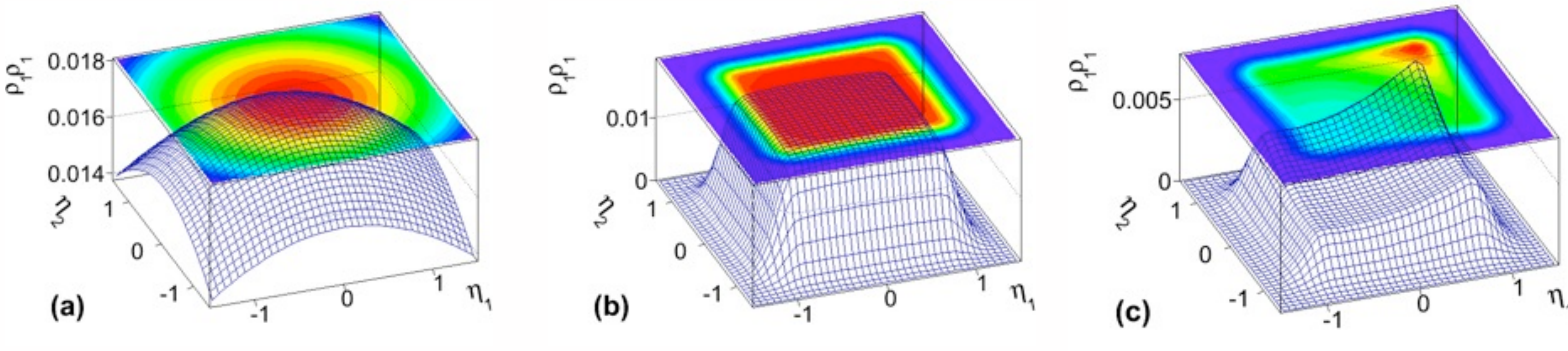
- Efficiency



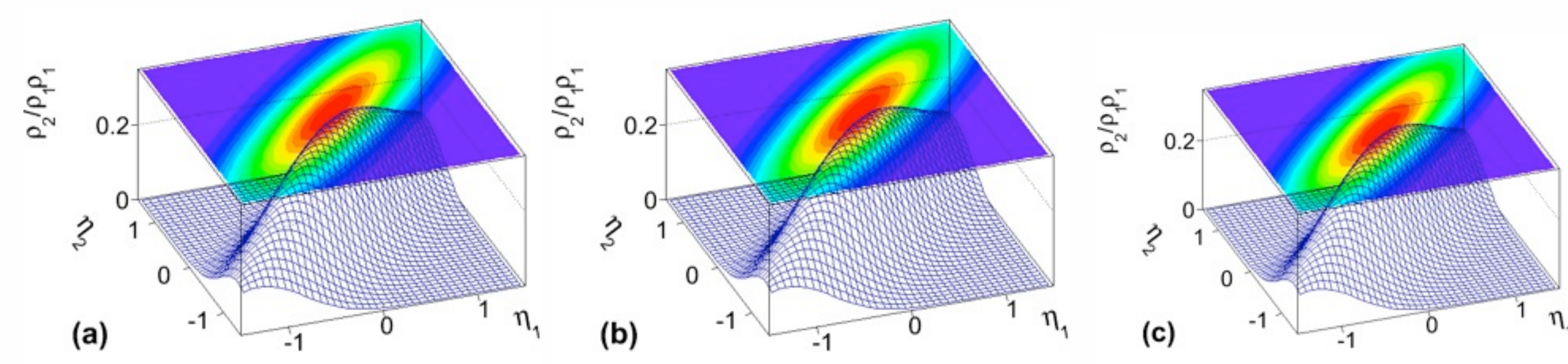
- Pair Yield



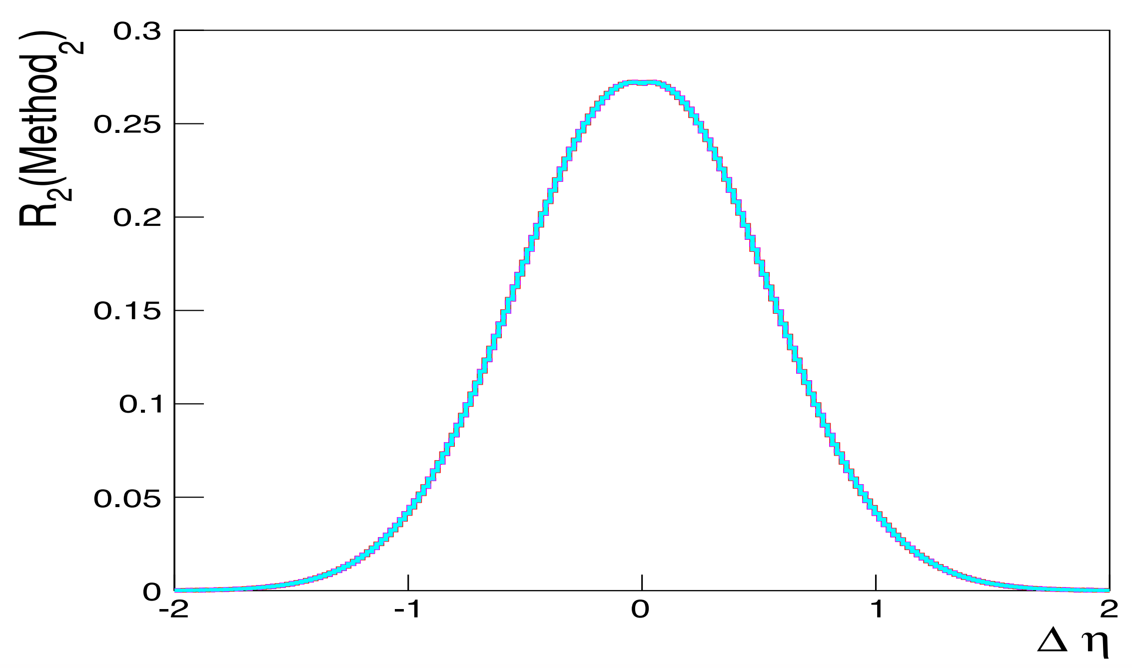
Product of singles

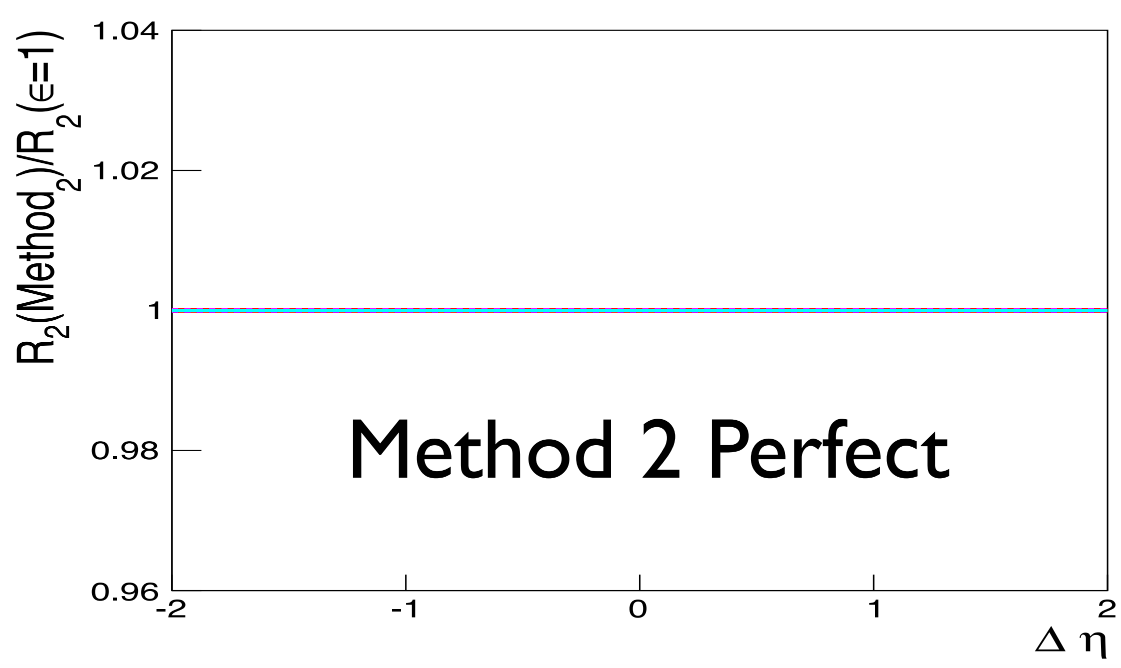


R2 (Method 2)

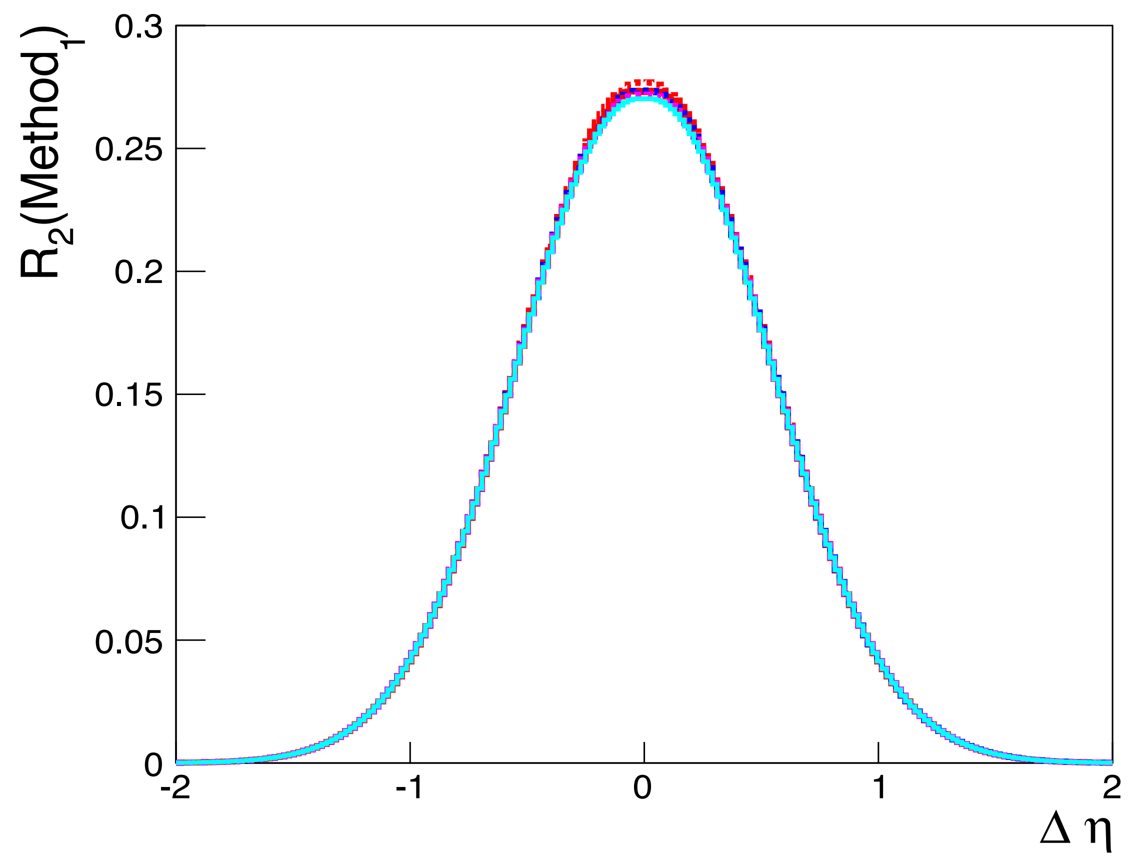


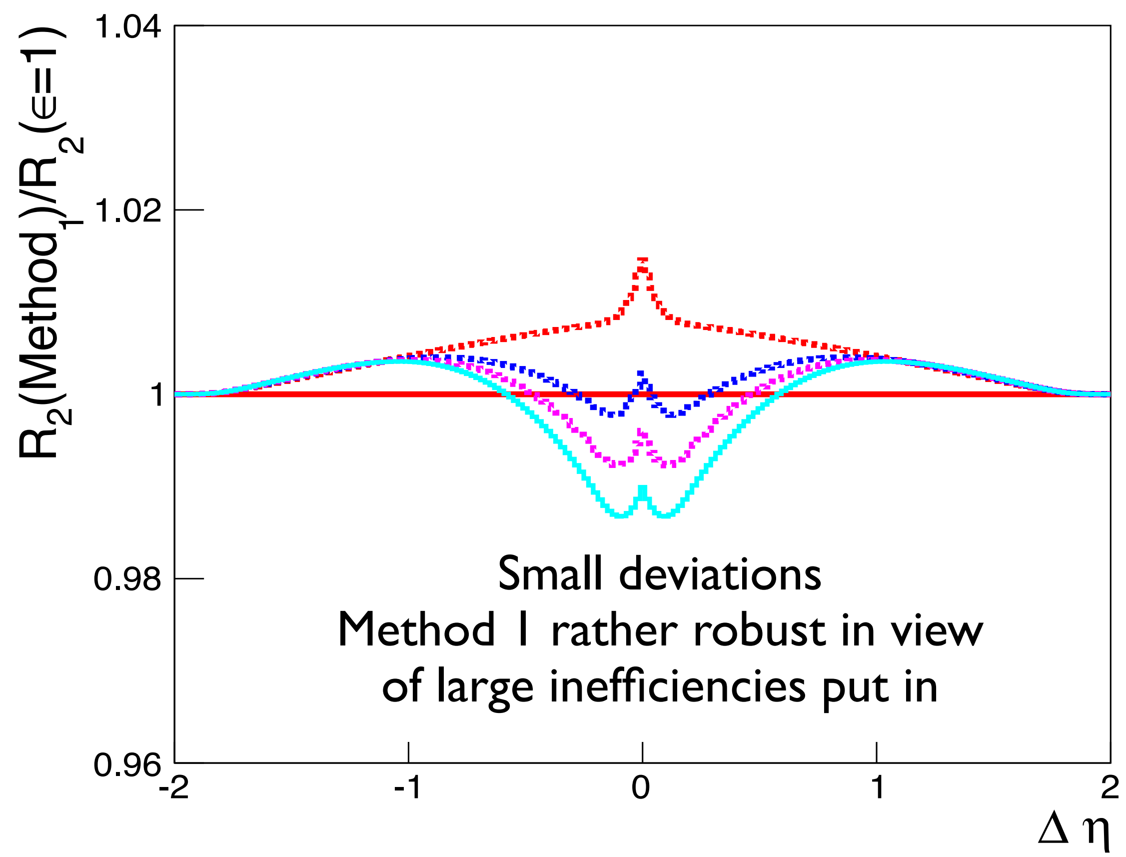
R2(Δη) Method 2





R2(Δη) Method 1





Why?

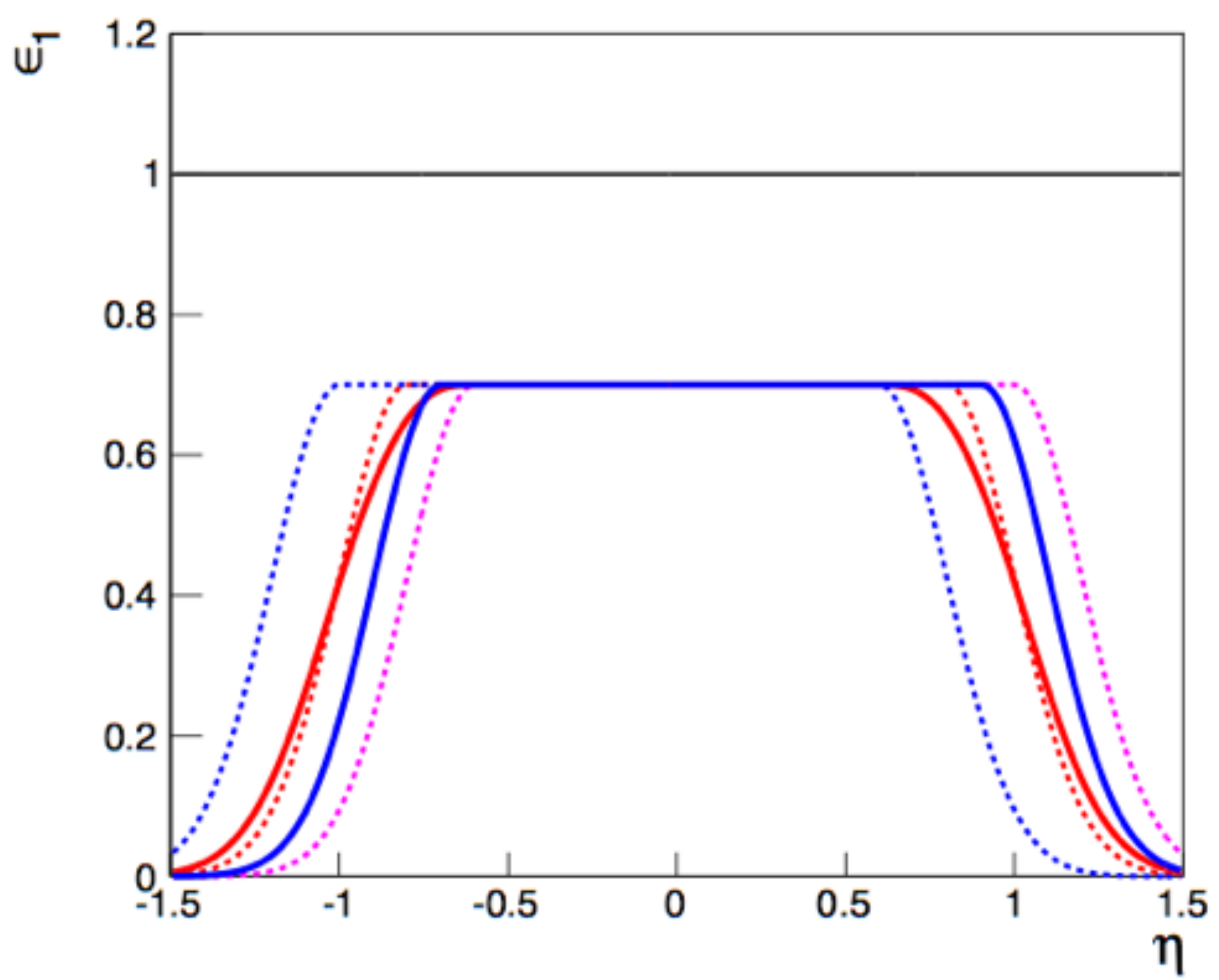
$$R_2(\Delta\eta)^{Method1} = \frac{\int g(\Delta\eta,\overline{\eta})R_2^{true}(\Delta\eta,\overline{\eta})d\overline{\eta}}{\int g(\Delta\eta,\overline{\eta})d\overline{\eta}}$$

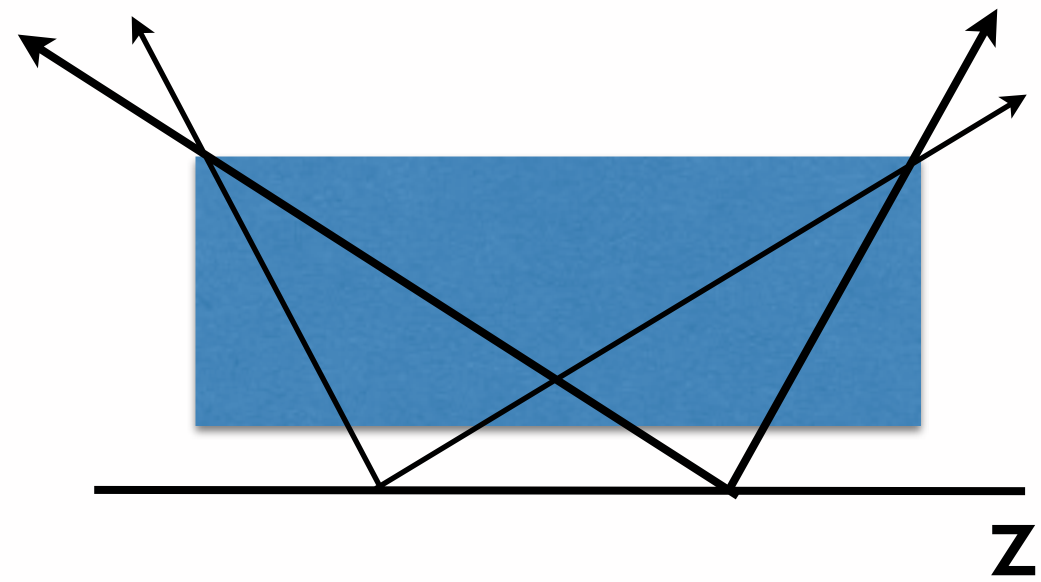
$$g(\Delta\eta,\overline{\eta}) = \epsilon_1 \times \epsilon_1 \times \rho_1 \times \rho_1(\Delta\eta,\overline{\eta})$$

- 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 Acceptance is a function of the vertex position.
- Use a simple model as before...





Efficiency Factorization???

- Local Factorization:

$$\epsilon_{pair}(\eta_1, \eta_2|z) = \epsilon_1(\eta_1|z) \times \epsilon_1(\eta_2|z)$$

- Loss of "Global" Factorization:

$$\begin{aligned}
\langle n_1(\eta_1)\rangle &= K\int_{z_{min}}^{z_{max}} P_c(z)\epsilon(\eta_1|z)\langle N_1(\eta_1)\rangle dz = \langle N_1(\eta_1)\rangle f_1(\eta_1) \\
\langle n_2(\eta_1,\eta_2)\rangle &= K\int_{z_{min}}^{z_{max}} P_c(z)\epsilon(\eta_1|z)\times\epsilon(\eta_2|z)\langle N_2(\eta_1,\eta_2)\rangle dz = \langle N_2(\eta_1,\eta_2)\rangle f_2(\eta_1,\eta_2)
\end{aligned}$$

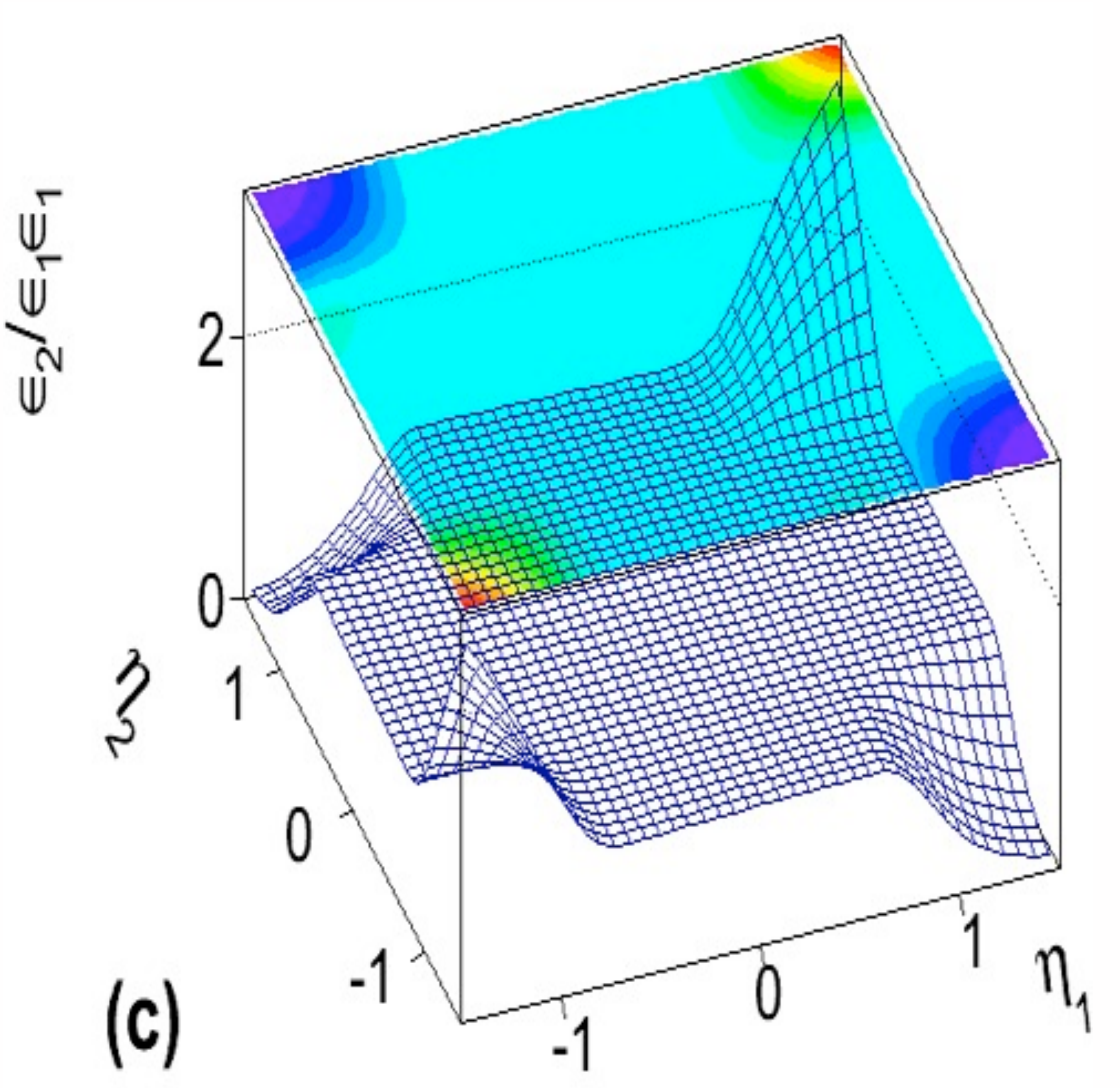
$$\begin{aligned}
f_1(\eta_1) &= K\int_{z_{min}}^{z_{max}} P_c(z)\epsilon(\eta_1|z)dz \\
f_2(\eta_1,\eta_2) &= K\int_{z_{min}}^{z_{max}} P_c(z)\epsilon(\eta_1|z)\times\epsilon(\eta_2|z)dz
\end{aligned}$$

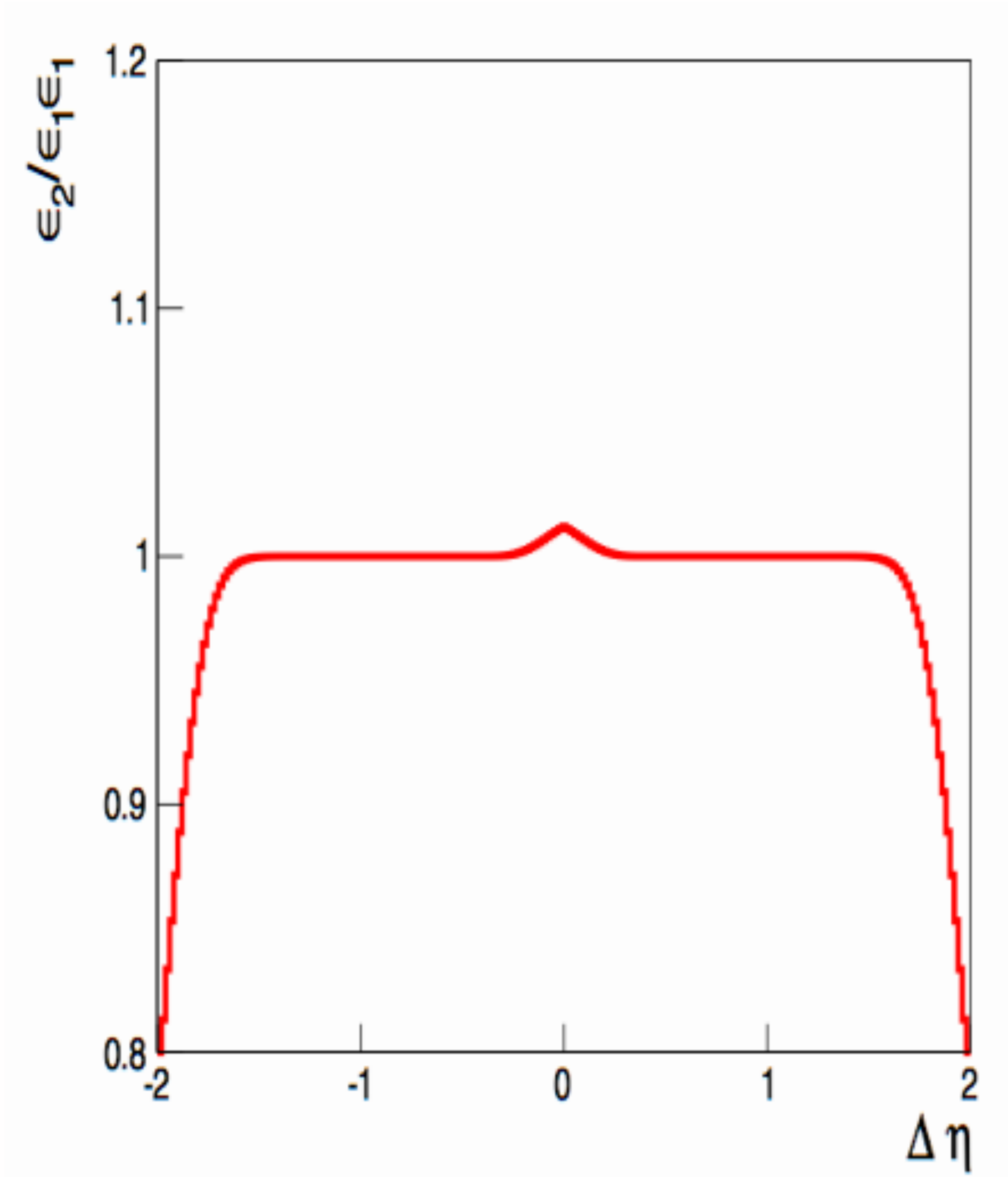
$$K^{-1} = \int_{z_{min}}^{z_{max}} P_c(z)dz$$

$$R_2(\eta_1,\eta_2) = \frac{f_2(\eta_1,\eta_2)}{f_1(\eta_1)f_1(\eta_2)}\frac{\langle N_2(\eta_1,\eta_2)\rangle}{\langle N_1(\eta_1)\rangle\langle N_1(\eta_2)\rangle}$$

Neither Methods Robust

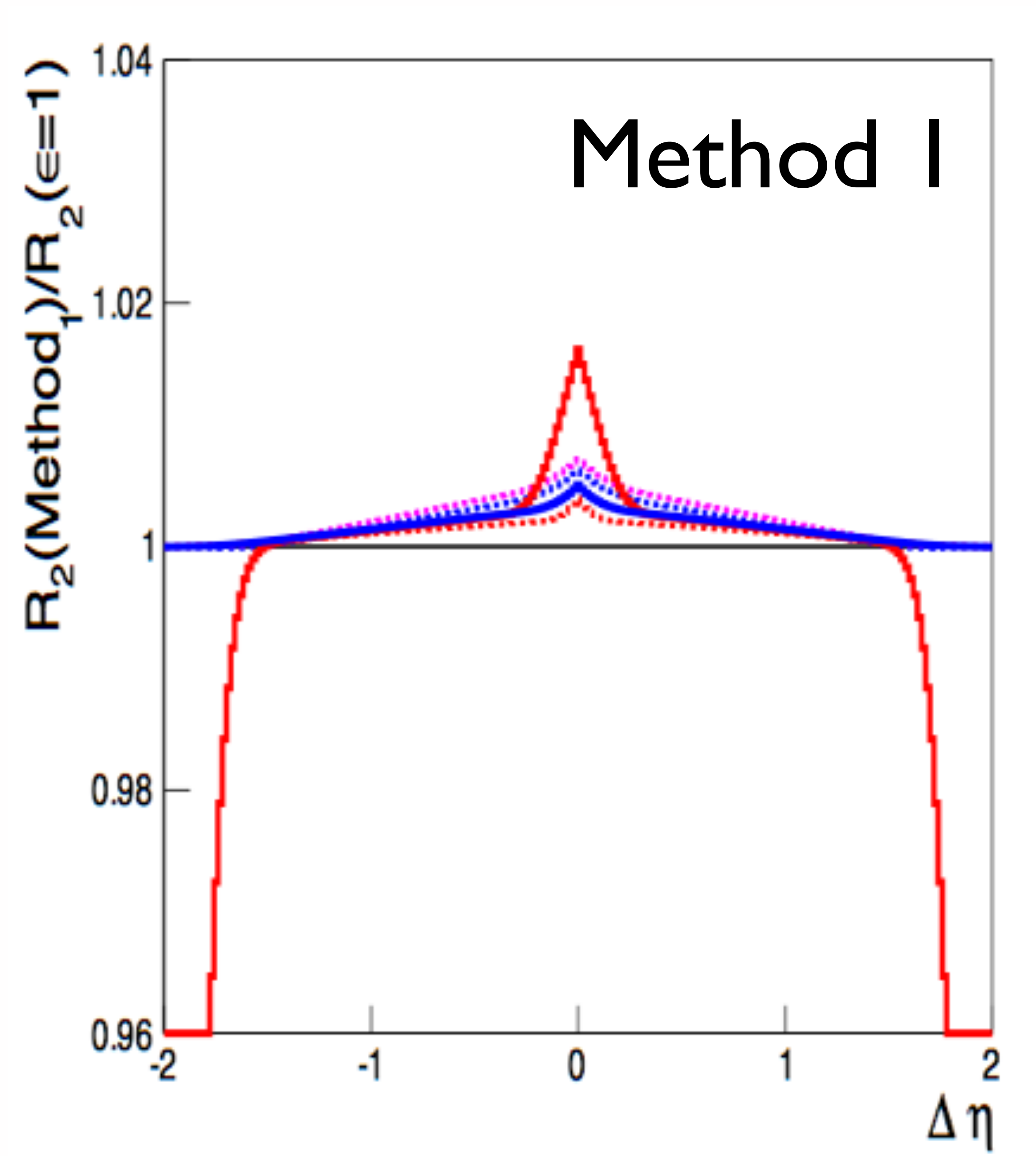
Pair and Single Efficiencies

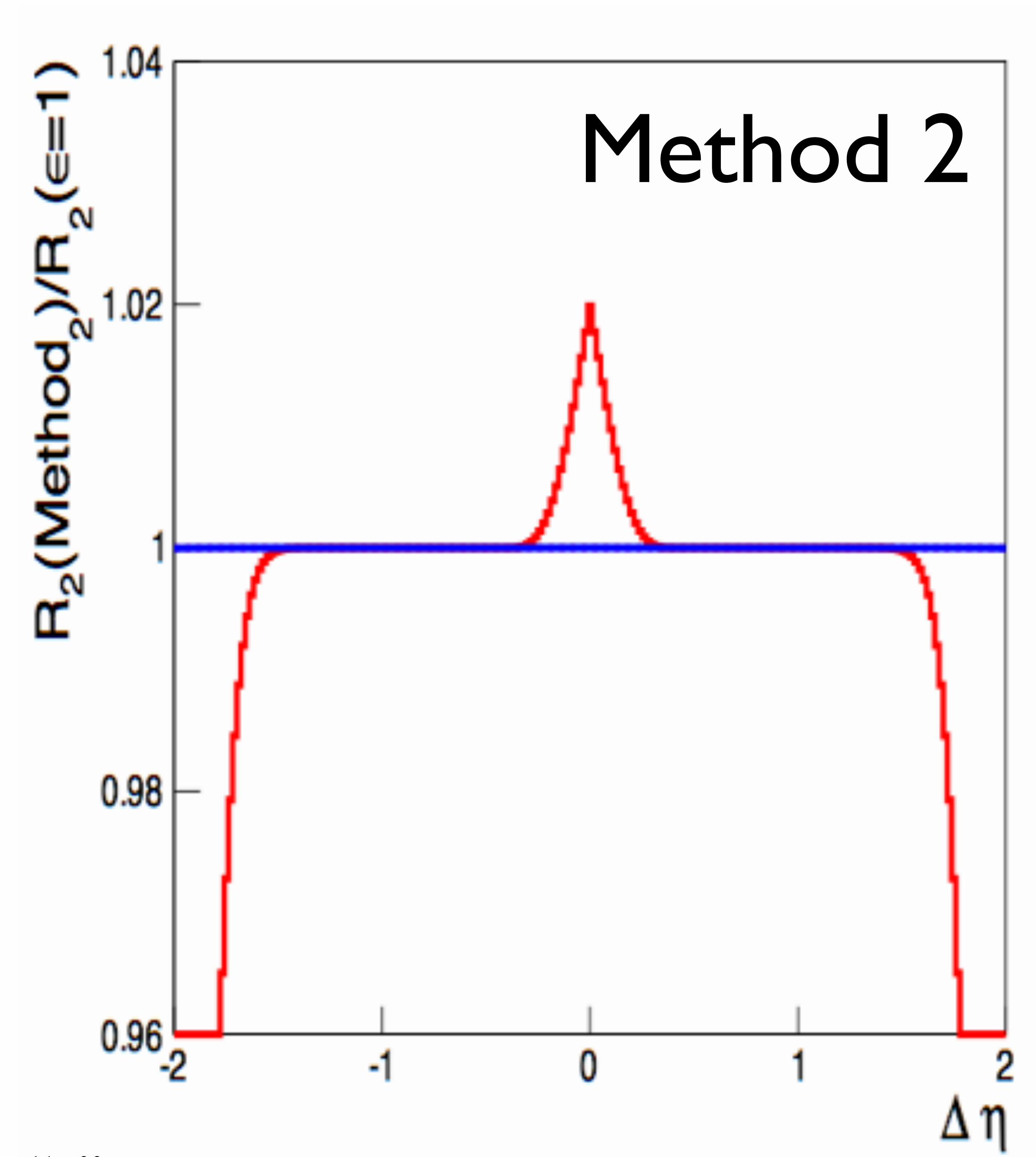




Comparison of Methods 1 and 2

Efficiency dependence on "z-vertex", with gaussian edges, but flat 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

Recovery...

- Method 2:
  - Carry analysis in fine (narrow) z bins.
  - Apply local efficiency factorization.

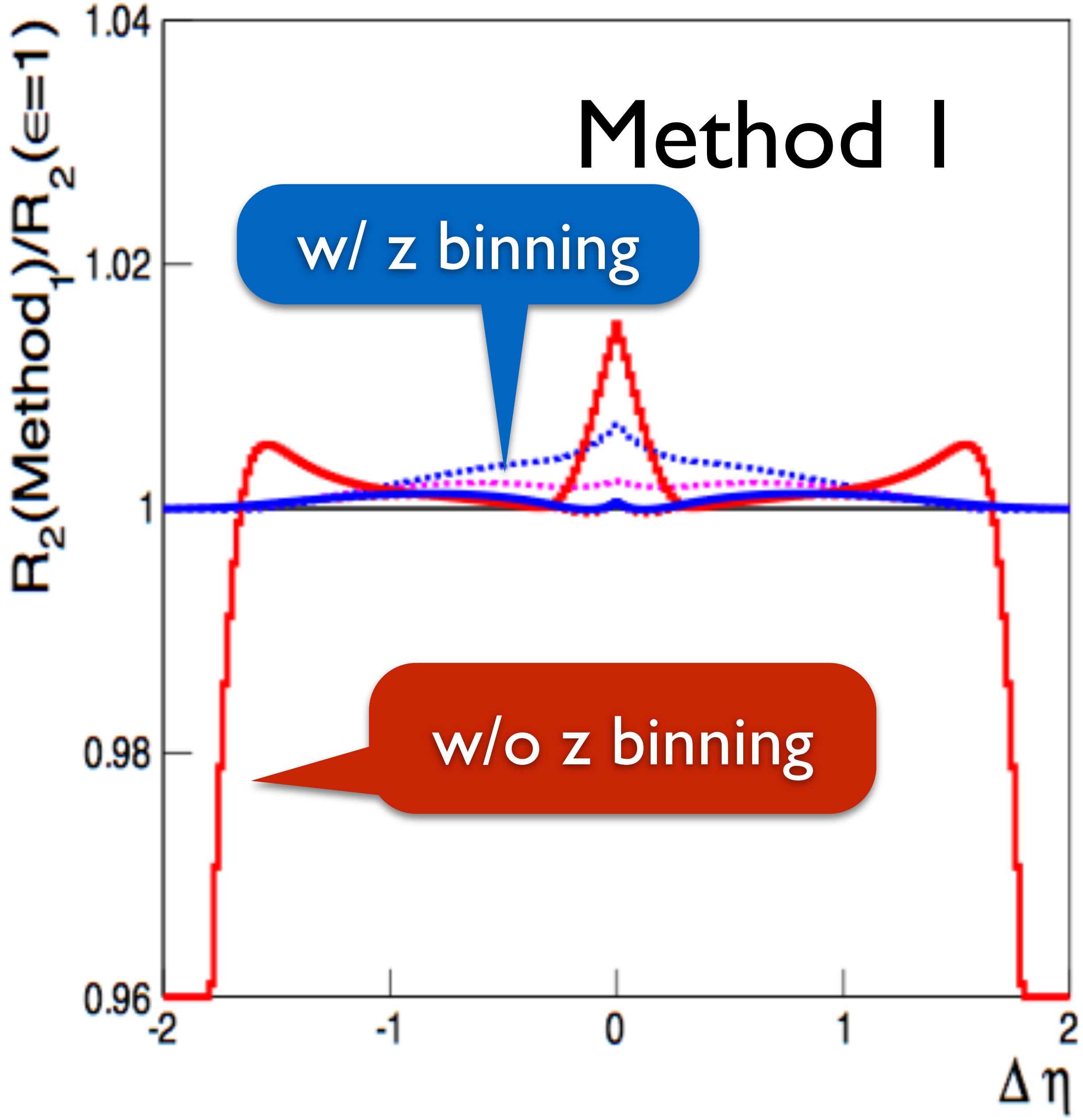
$$R_2(\eta_1, \eta_2) = K \int_{z_{min}}^{z_{max}} P_c(z) \times \frac{\epsilon(\eta_1|z) \times \epsilon(\eta_2|z)\langle N_2(\eta_1, \eta_2)\rangle}{\epsilon(\eta_1|z) \times \epsilon(\eta_2|z)\langle N_1(\eta_1)\rangle\langle N_1(\eta_2)\rangle} dz \quad (49)$$

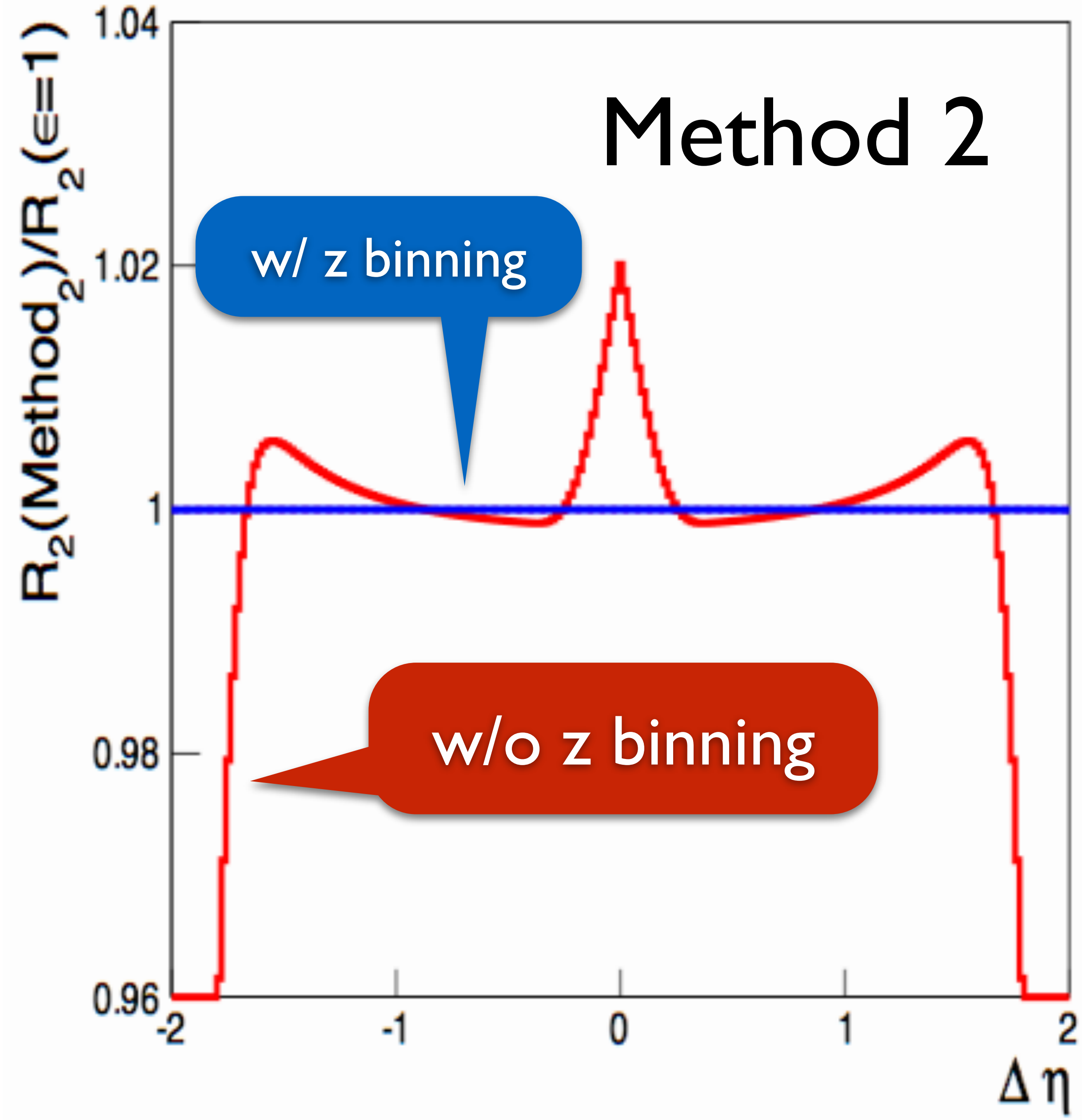
$$\begin{aligned} R_2(\eta_1, \eta_2) &= K \frac{\langle N_2(\eta_1, \eta_2)\rangle}{\langle N_1(\eta_1)\rangle\langle N_1(\eta_2)\rangle} \int P_c(z) dz \\ &= \frac{\langle N_2(\eta_1, \eta_2)\rangle}{\langle N_1(\eta_1)\rangle\langle N_1(\eta_2)\rangle} \end{aligned}$$

- Method 1:
  - This “recipe” not strictly valid for method 1

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

Conclusions

- 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