

Principal Component Analysis of Correlation Data without Nonflow Effects

Rajeev S. Bhalerao

Department of Theoretical Physics
Tata Institute of Fundamental Research
Mumbai, India

3rd International Conference on the
Initial Stages in High-Energy Nuclear Collisions (IS2016)
Lisbon, Portugal, May 23-27, 2016

(work in progress)

with Jean-Yves Ollitrault, Subrata Pal
(Saclay) (TIFR)

Outline

- Motivation
- Principal Component Analysis (PCA)
- PCA in Pseudorapidity (η) without Nonflow Effects
 - Method
 - Results
- Summary and Outlook

Longitudinal dynamics of the fireball formed in HE collisions:

- Forward-backward rapidity correlations → Mechanism of particle production in HE AA and pA collisions
- Event-plane decorrelation and factorization breakdown for particles of different η were demonstrated recently
- η dependence of v_n → Mechanism underlying collectivity

Origin of Longitudinal Correlations

- Fluctuations in multiplicity $dN/d\eta$ e-by-e. Not just statistical but also due to $N_{part}(\text{proj}) \neq N_{part}(\text{tgt})$.
- Fluctuations in anisotropic flow $v_n(\eta)$ e-by-e. Not just statistical but also due to $\Psi_n(\text{proj}) \neq \Psi_n(\text{tgt})$.
- Symmetry, if any, arises after event averaging.
- $dN/d\eta$ and $v_n(\eta)$ need to be treated **on equal footing**
- **PCA** with its eigenmode expansion (rather than expansion in an arbitrary basis) provides **a natural framework** to study longitudinal correlations.

Principal Component Analysis (PCA), in general

- Statistical procedure to elucidate the underlying covariance structure in the multi-dimensional data
- To identify the directions (PC) where there is the most variance, and possibly reduce the dimension of data
- Diagonalize the covariance matrix: Eigenvector with the largest eigenvalue is the direction of greatest variance; that with the 2nd largest eigenvalue ...
- PCA can be thought of as fitting a hyper-ellipsoid to the cloud of data points — each of its axes representing a PC

PCA — Fluctuations and Correlations in HIC

- PRL **114**, 152301 (2015)
- Divide the detector acceptance into several bins in p_T and/or η . Let p : bin index.
- Flow vector in an event: $Q_n(p) \equiv \sum_{j=1}^{M(p)} \exp(in\phi_j)$
- Covariance matrix $V_{n\Delta}(p_1, p_2)$
 $\equiv \langle Q_n(p_1) Q_n^*(p_2) \rangle - \langle M(p_1) \rangle \delta_{p_1 p_2} - \langle Q_n(p_1) \rangle \langle Q_n^*(p_2) \rangle$
RHS₂: subtracts self-correl. RHS₃: singles out fluct.
- PC are obtained by diagonalizing $V_{n\Delta}(p_1, p_2)$, which is essentially the symmetric real matrix $\langle \cos n\Delta\phi \rangle$.

Flow in a given event is a sum over eigenmodes:

$$V_n(p) = \sum_{\alpha=1}^k \xi^{(\alpha)} V_n^{(\alpha)}(p), \quad (k=2 \text{ or } 3 \text{ suffice})$$

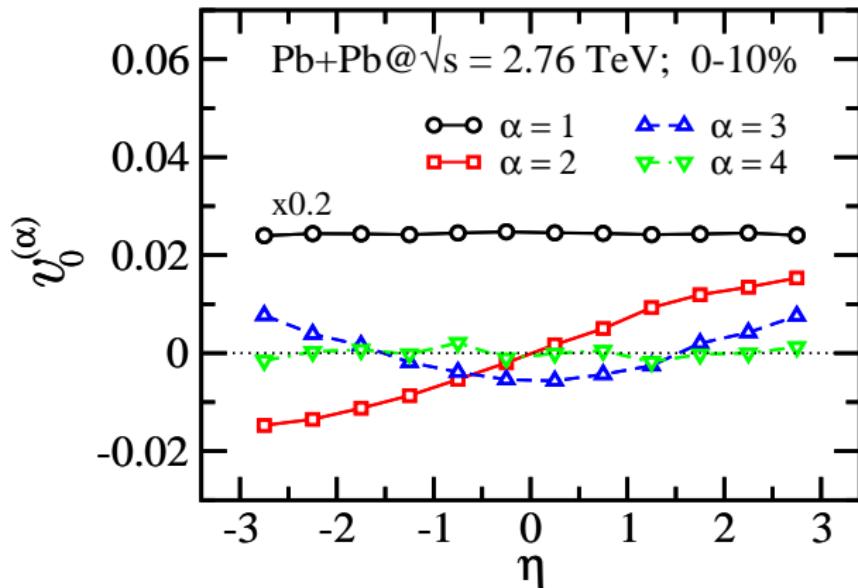
Principal components $V_n^{(\alpha)}(p)$ are **the same for all events**. $\xi^{(\alpha)}$ which **fluctuate from event to event** are complex, random variables with zero mean and unit variance. In particular,

$$\frac{dN}{d\eta} - \left\langle \frac{dN}{d\eta} \right\rangle = \xi^{(1)} V_0^{(1)}(\eta) + \xi^{(2)} V_0^{(2)}(\eta) + \dots$$

PCA — Important Points to Note

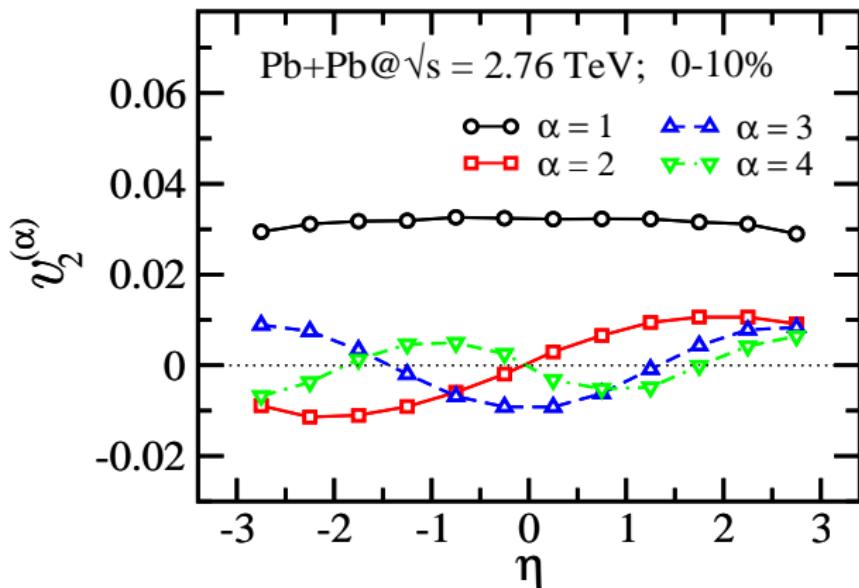
- Particle multiplicity ($n = 0$) and anisotropic flow ($n > 0$) are treated **on the same footing**.
- PCA uses **all** the information in the correl matrix and extracts flow fluctuations directly from data
- PCA yields $V_{n\Delta}(p_1, p_2) \simeq \sum_{\alpha=1}^k V_n^{(\alpha)}(p_1) V_n^{(\alpha)*}(p_2)$
= sum over modes of flow fluct. Here ($k < N_{dim}$).
- Flow fluctuations \Rightarrow Subleading modes \Rightarrow No factorization $V_{n\Delta}(p_1, p_2) \neq V_n^{(1)}(p_1) V_n^{(1)*}(p_2)$
- Leading eigenmodes are essentially equivalent to the familiar measurements of v_2 , v_3 . **Subleading modes of v_2 , v_3 were revealed for the first time**

Principal Components of Relative Multiplicity



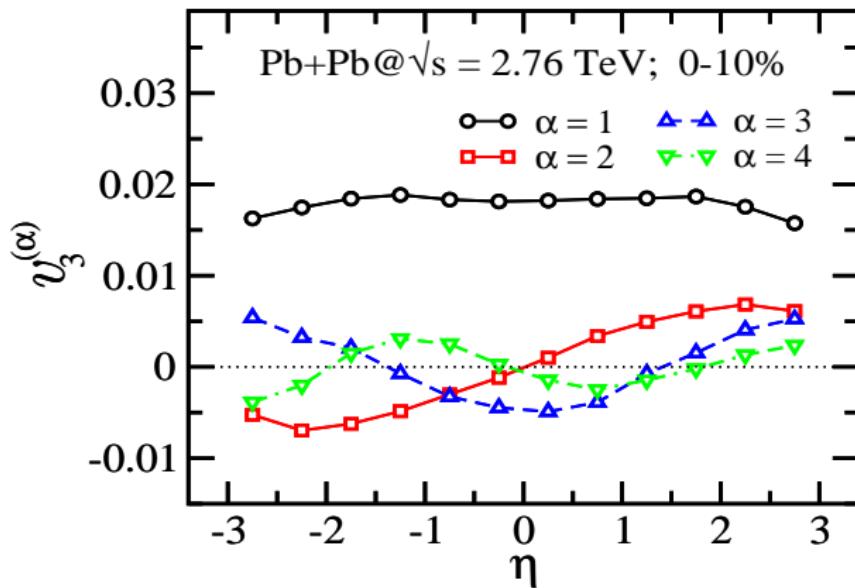
- PCA results for the leading and 3 subleading modes
- PC: Alternating parities; Mutually orthogonal
- Subleading modes \ll Leading modes; Eigenvalues $\lambda^{(3)} \ll \lambda^{(2)} \ll \lambda^{(1)}$

Principal Components of Elliptic Flow



- PCA results for the leading and 3 subleading modes
- PC: Alternating parities; Mutually orthogonal
- Subleading modes \ll Leading modes; Eigenvalues $\lambda^{(3)} \ll \lambda^{(2)} \ll \lambda^{(1)}$

Principal Components of Triangular Flow



- PCA results for the leading and 3 subleading modes
- PC: Alternating parities; Mutually orthogonal
- Subleading modes \ll Leading modes; Eigenvalues $\lambda^{(3)} \ll \lambda^{(2)} \ll \lambda^{(1)}$

Different Eigenmodes \longrightarrow Different Physics

As a function of η

- $v_0^{(1)}$: Global (η -indep) relative multiplicity fluct
- $v_0^{(2)}$: Due to $N_{part}(\text{target}) \neq N_{part}(\text{projectile})$
- $v_{2,3}^{(1)}$: Correspond to the familiar measurements of v_2 and v_3
- $v_{2,3}^{(2)}$: Due to $\Psi_n(\text{target}) \neq \Psi_n(\text{projectile})$
(Event-plane decorrelation and Torqued flow)

Present Work

PCA for Pseudorapidity Correlations

Suppose $\eta = -3$ to 3 and bin size = 0.5. Thus 12×12 correlation matrix $V_{n\Delta}(\eta_1, \eta_2)$. Can diagonalize. But ...

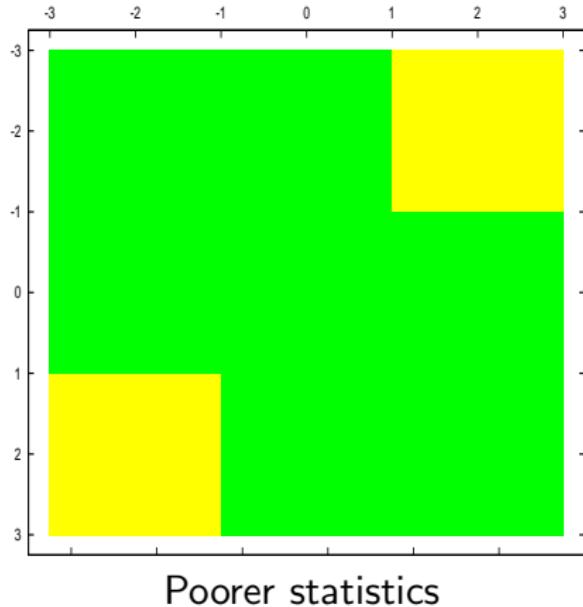
Two-particle correlation data includes both flow (long-range) and nonflow. Our earlier analysis (in η) ignored nonflow.

To reduce short-range nonflow contributions, apply a pseudorapidity gap, i.e., drop diagonal, ... elements.

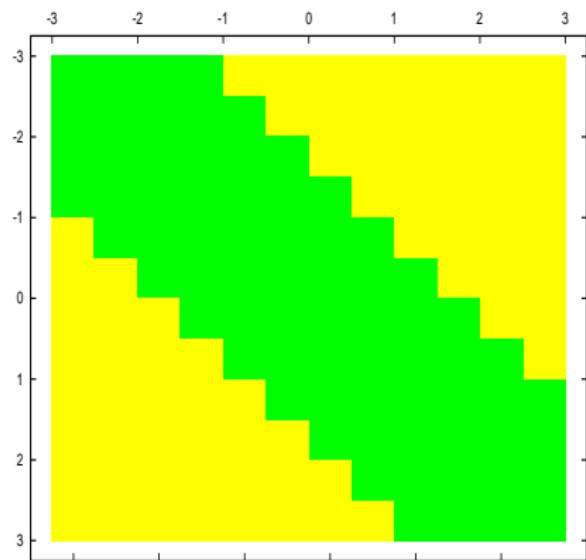
Fit $V_{n\Delta}(\eta_1, \eta_2) \simeq \sum_{\alpha=1}^k V_n^{(\alpha)}(\eta_1) V_n^{(\alpha)}(\eta_2)$,
where $k = 1, 2$ or 3 , treating $V_n^{(\alpha)}$ as parameters.

Two Ways: Fixed Gap versus Mobile Gap

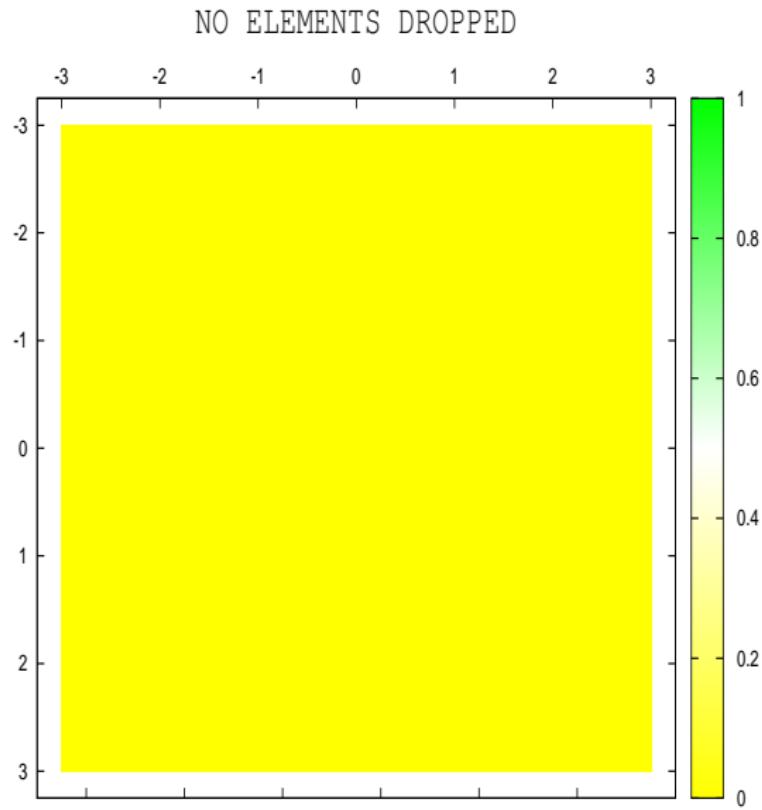
FIXED GAP $|\Delta\eta| > 2$



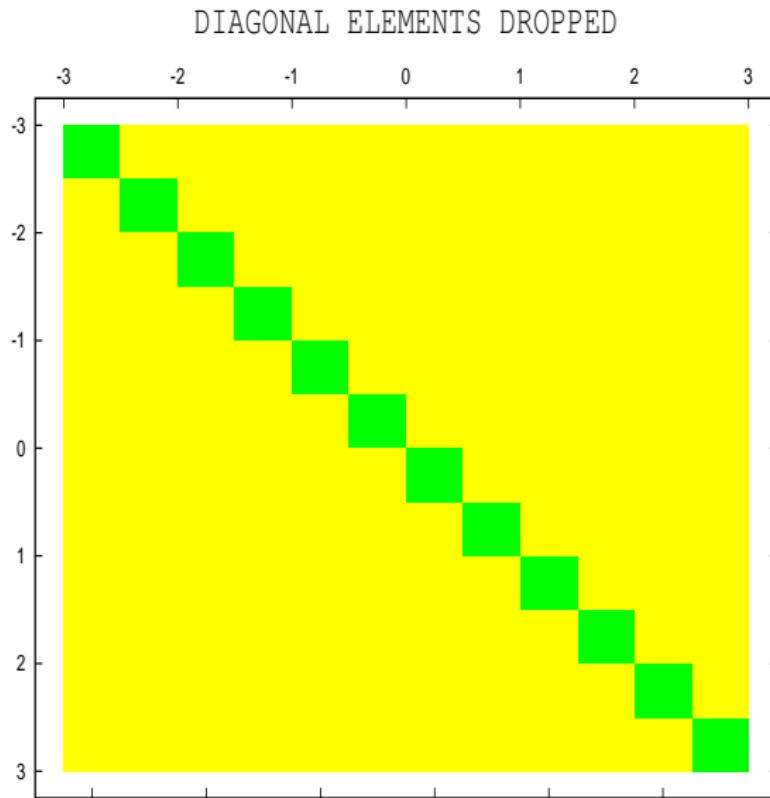
MOBILE GAP $|\Delta\eta| > 2$



Two-Particle Correlation Matrix with $|\Delta\eta| = 0$

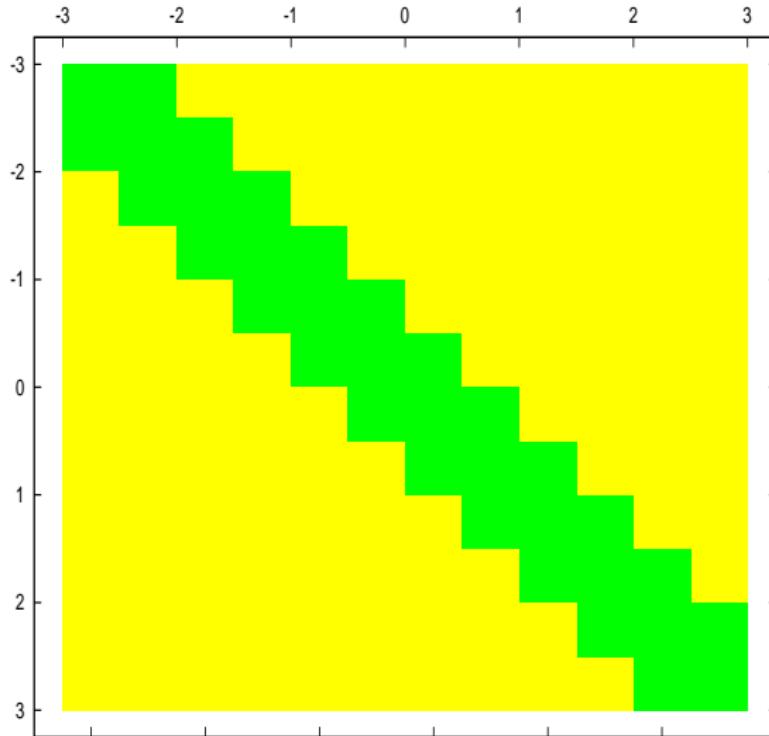


Two-Particle Correlation Matrix with $|\Delta\eta| > 0.5$



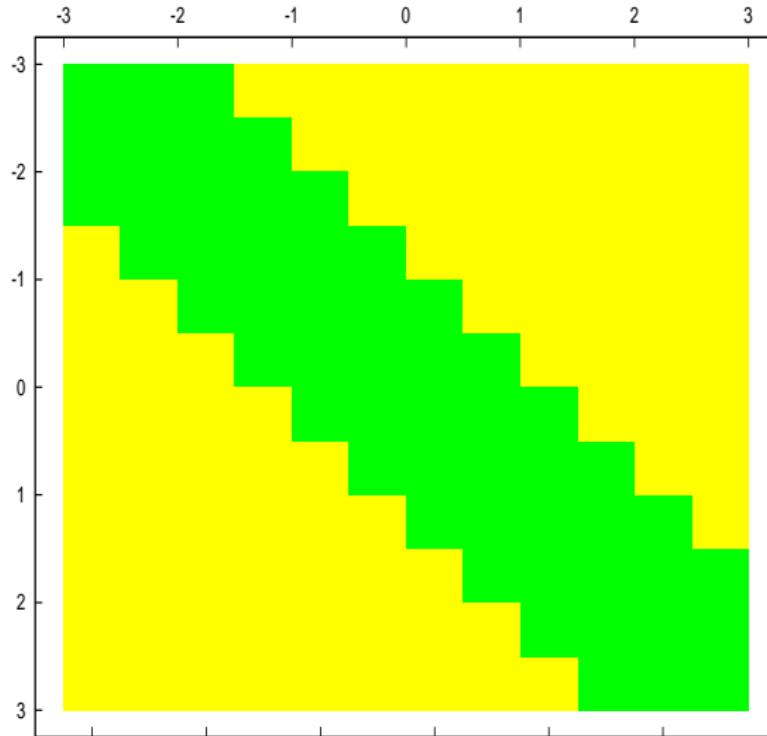
Two-Particle Correlation Matrix with $|\Delta\eta| > 1.0$

DIAGONAL, DIAGONAL ± 1 ELEMENTS DROPPED



Two-Particle Correlation Matrix with $|\Delta\eta| > 1.5$

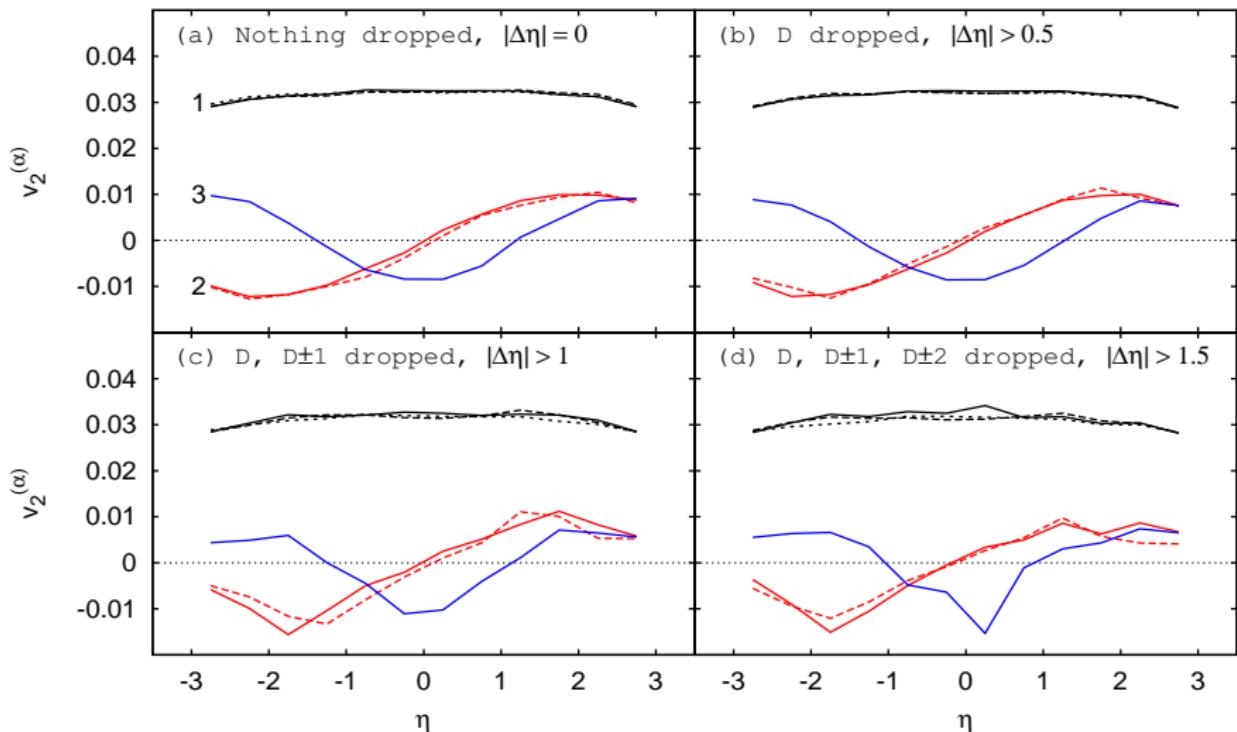
DIAG, DIAG ± 1 , DIAG ± 2 ELEMENTS DROPPED



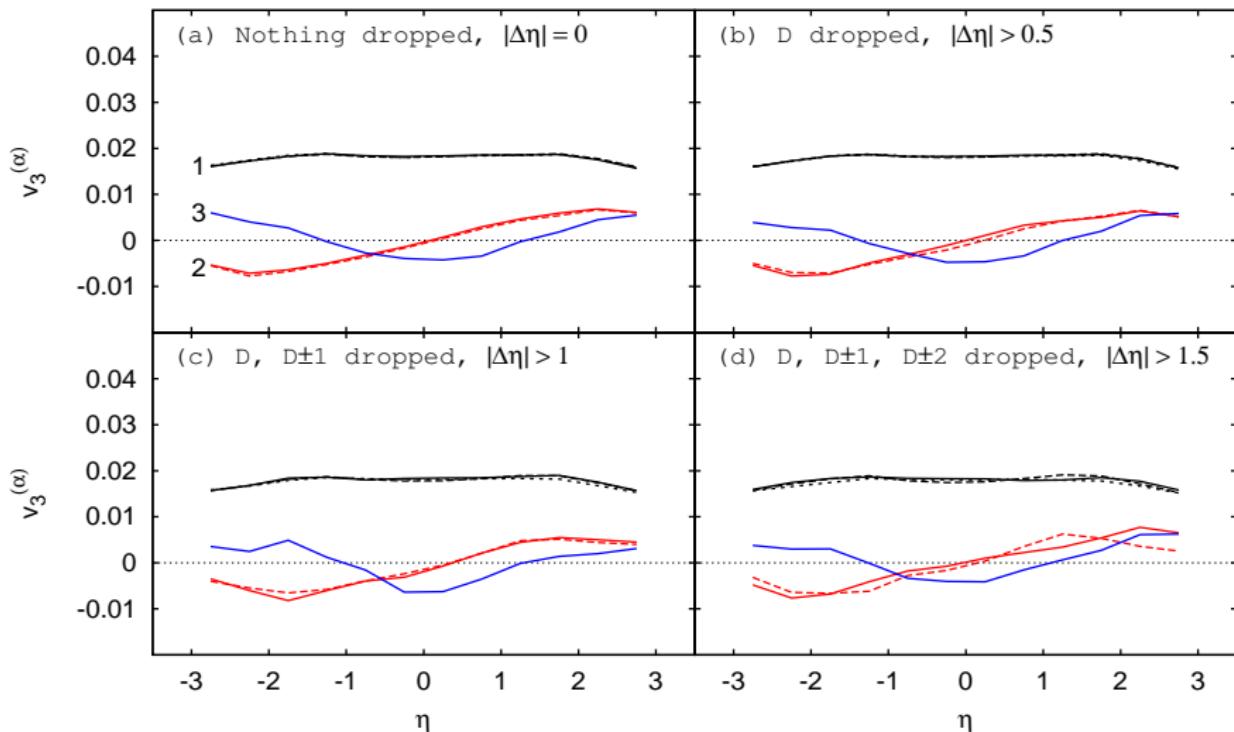
Mobile Gap

- Drop diagonal (D) elements \cong Mobile gap $|\Delta\eta| > 0.5$
- Drop $D, D \pm 1$ elements \cong Mobile gap $|\Delta\eta| > 1.0$
- Drop $D, D \pm 1, D \pm 2$ elements \cong Mobile gap $|\Delta\eta| > 1.5$
- Drop $D, D \pm 1, D \pm 2, D \pm 3$ elements \cong Mobile gap $|\Delta\eta| > 2.0$

Elliptic Flow Modes with a Pseudorapidity Gap



Triangular Flow Modes with a Pseudorapidity Gap



Solid lines: 3-mode fit, dashed lines: 2-mode fit, dotted lines: 1-mode fit

Summary

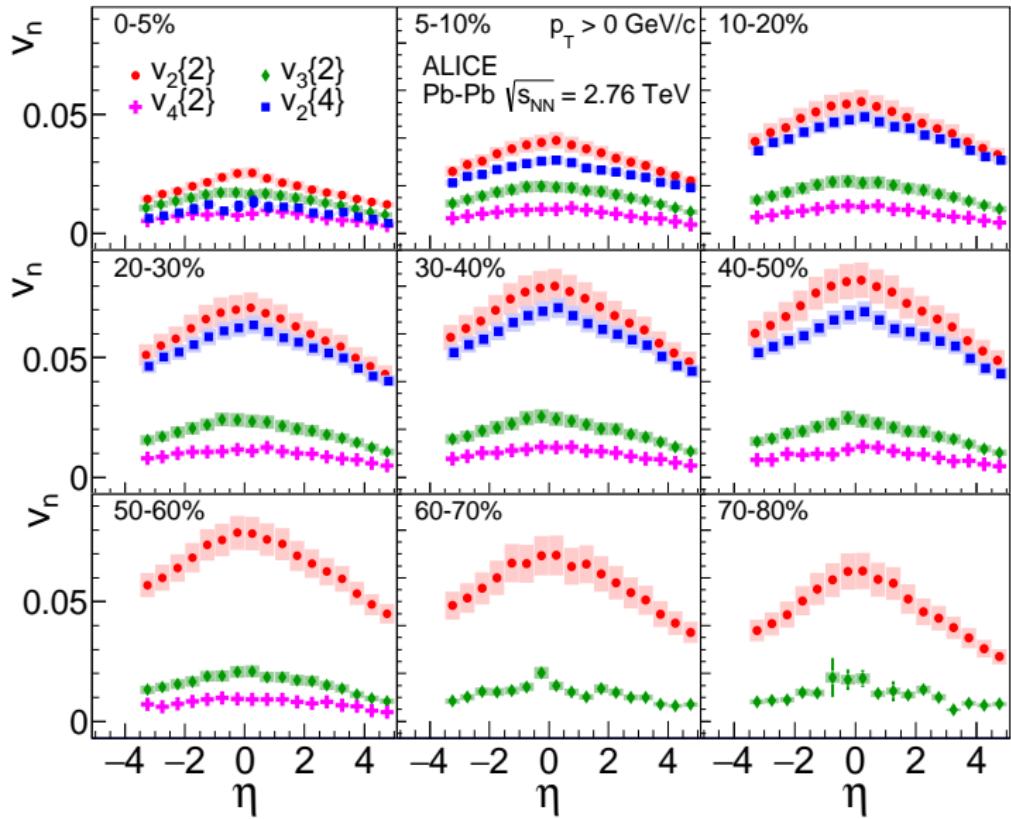
Subleading modes in pseudorapidity are clearly evident and measurable **even after imposing a rapidity gap of $|\Delta\eta| > 1.5$.** (Conclusion based on AMPT simulations.)

ALICE (1605.02035, 6 May 2016):

$v_n(\eta)$ over the widest η range at LHC $-3.5 < \eta < 5.0$.

Agreement with Hydro and AMPT not quite good.

Outlook



Outlook

- PCA in p_T ... J. Milošević (CMS), QM 2015, Kobe
- PCA in η is waiting to be done!

THANK YOU

BACKUP SLIDES

PCA with a Pseudorapidity Gap

Covariance matrix is a 12×12 symmetric matrix. Of the 144 matrix elements: (a) nothing dropped, (b) diagonal elements dropped, (c) diagonal and next-to-diagonal elements dropped, and (d) diagonal, next-to-diagonal and next-to-next-to-diagonal elements dropped. Table shows the number of free elements = $N_{indep} - N_{para}$.

	No. of indep. mat. els. N_{indep}	1-mode fit $N_{para} = 12$	2-mode fit $N_{para} = 24$	3-mode fit $N_{para} = 36$
(a)	78	66	54	42
(b)	66	54	42	30
(c)	55	43	31	19
(d)	45	33	21	9

Fits for $V_{2\Delta}(\eta_1, \eta_2)$. Three leading eigenvalues: 19211.7, 1436.52, 874.472.

1-eigenmode or 12-parameter fit

	χ^2_{red}	λ_1	λ_2	λ_3
(a)	44.79	19229.0		
(b)	38.11	18917.0		
(c)	19.17	18430.3		
(d)	6.19	17972.2		

2-eigenmode or 24-parameter fit

(a)	17.38	19192.8	1484.4	
(b)	15.75	19067.0	1331.4	
(c)	6.50	18798.2	1254.9	
(d)	2.10	18430.9	898.1	

3-eigenmode or 36-parameter fit

(a)	4.131	19201.0	1447.61	898.644
(b)	4.035	19149.1	1387.79	815.306
(c)	1.238	19066.5	1363.74	754.550
(d)	0.559	18933.3	1178.53	886.501

Fits for $V_{3\Delta}(\eta_1, \eta_2)$. Three leading eigenvalues: 6220.91, 483.971, 252.213.

1-eigenmode or 12-parameter fit

	χ^2_{red}	λ_1	λ_2	λ_3
(a)	14.09	6218.36		
(b)	12.87	6139.24		
(c)	6.99	5992.28		
(d)	2.88	5851.58		

2-eigenmode or 24-parameter fit

(a)	4.38	6211.77	494.88	
(b)	4.30	6175.39	459.04	
(c)	2.35	6100.69	374.53	
(d)	0.84	6063.88	395.84	

3-eigenmode or 36-parameter fit

(a)	0.9576	6214.18	484.277	256.007
(b)	0.8649	6204.65	476.458	247.703
(c)	0.4840	6179.52	447.219	250.495
(d)	0.6292	6136.61	464.664	253.376

Flow Picture

- Single-particle distribution in an event:

$$\frac{dN}{dp} = \sum_{n=-\infty}^{\infty} V_n(p) \exp(in\phi)$$

- Pair distribution averaged over events:

$$\begin{aligned} \left\langle \frac{dN_{pairs}}{d\mathbf{p}_1 d\mathbf{p}_2} \right\rangle &= \left\langle \frac{dN}{d\mathbf{p}_1} \frac{dN}{d\mathbf{p}_2} \right\rangle + \mathcal{O}(N) \leftarrow \text{nonflow correl.} \\ &= \sum_{n=-\infty}^{\infty} V_{n\Delta}(p_1, p_2) \exp(in(\phi_1 - \phi_2)) \end{aligned}$$

- Two-particle correlation matrix:

$$V_{n\Delta}(p_1, p_2) = \langle V_n(p_1) V_n^*(p_2) \rangle,$$

neglecting nonflow correlations

Covariance Matrix

- $V_{n\Delta}(p_1, p_2) = \langle V_n(p_1) V_n^*(p_2) \rangle$ is a covariance matrix.
- Covariance matrix is positive semidefinite, and its eigenvalues are non-negative.

Method (contd.)

We define

$$v_n^{(\alpha)}(p) \equiv \frac{V_n^{(\alpha)}(p)}{\langle V_0(p) \rangle}$$

$n = 0$: rel. multiplicity; $n \neq 0$: anisotropic flow

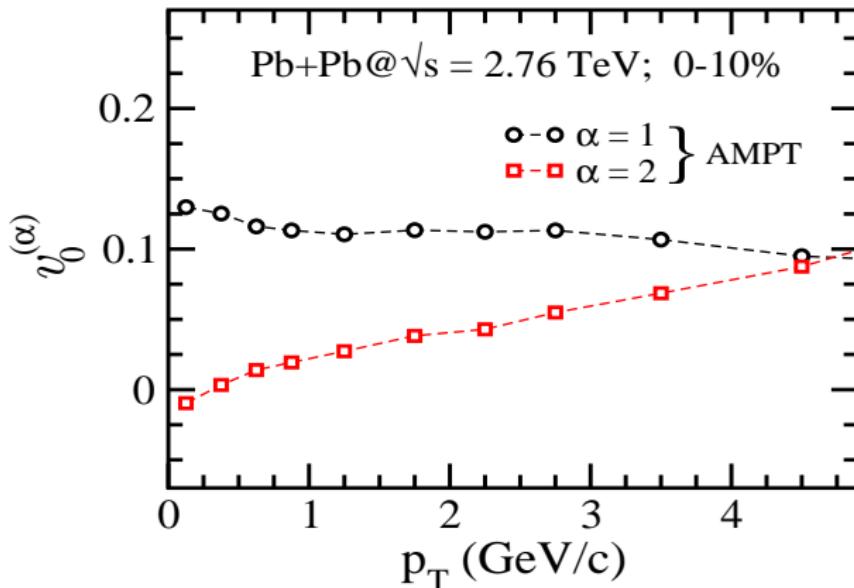
Thus $v_0^{(\alpha)}(p)$: relative multiplicity fluctuations,

$v_n^{(\alpha)}(p)$ for $n \neq 0$: anisotropic flow fluctuations

Orthogonality of Principal Components

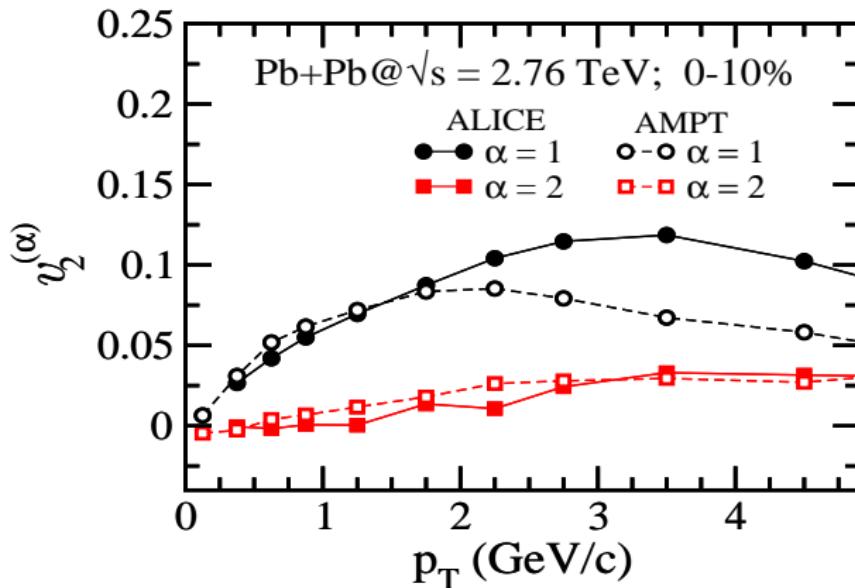
$$\sum_p V_n^{(\alpha)}(p) V_n^{(\beta)*}(p) = 0, \quad \text{if } \alpha \neq \beta$$

Relative Multiplicity Fluctuations vs Transverse Momentum



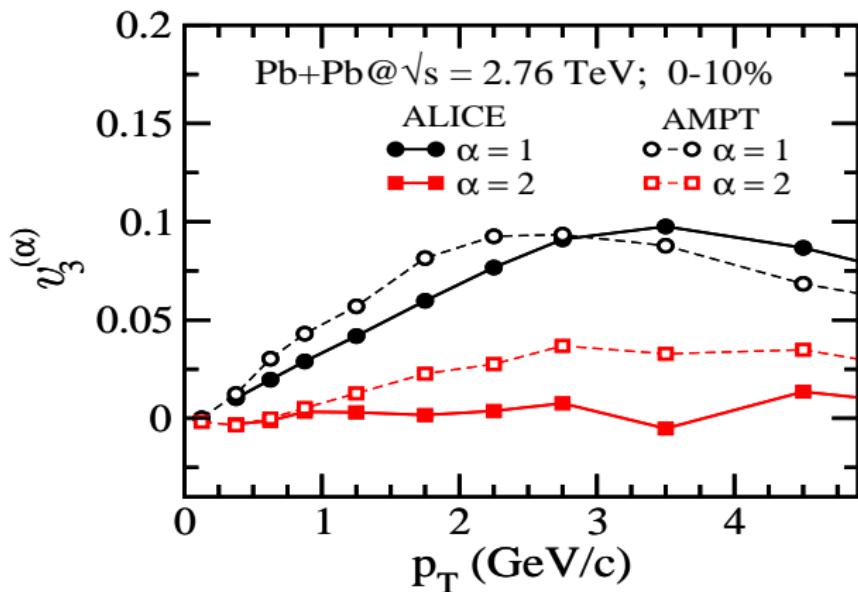
- PCA results for the leading and subleading modes
- Subleading modes \ll Leading modes; Eigenvalue $\lambda^{(2)} \ll \lambda^{(1)}$
- No experimental data available for any result shown so far

Elliptic Flow Fluctuations vs Transverse Momentum



- PCA results for the leading and subleading modes
- Subleading modes \ll Leading modes; Eigenvalue $\lambda^{(2)} \ll \lambda^{(1)}$

Triangular flow fluctuations vs Transverse Momentum



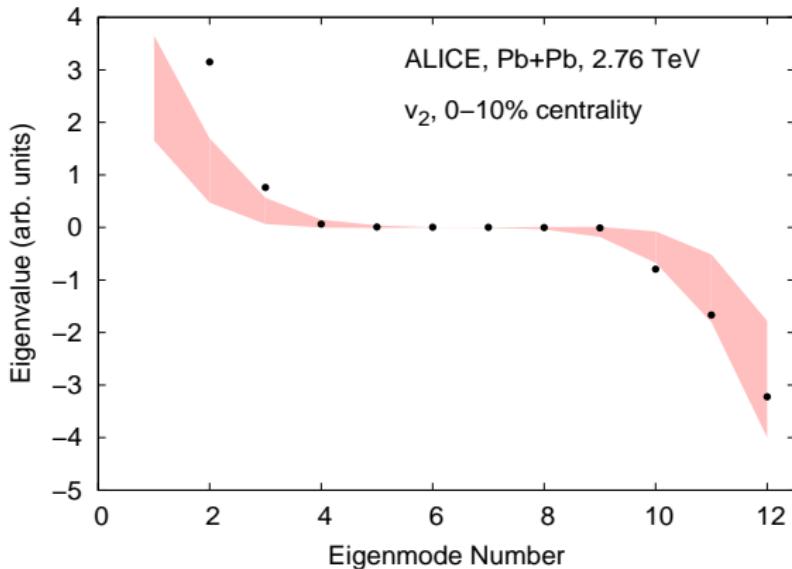
- PCA results for the leading and subleading modes
- Subleading modes \ll Leading modes; Eigenvalue $\lambda^{(2)} \ll \lambda^{(1)}$

Different Eigenmodes → Different Physics

As a function of p_T

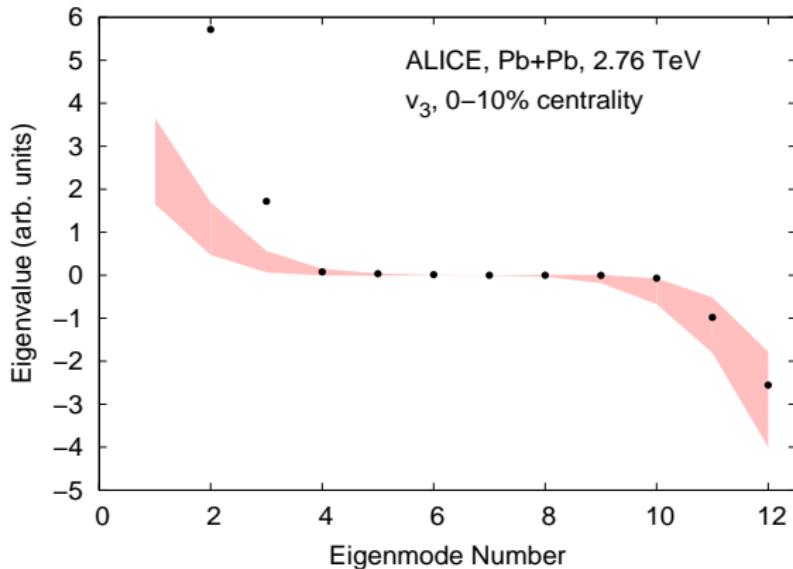
- $v_0^{(1)}$: Global (p_T -indep) relative multiplicity fluct
- $v_0^{(2)}$: Radial flow fluct (expected in hydro model)
- $v_{2,3}^{(1)}$: Correspond to the usual v_2 and v_3
- $v_{2,3}^{(2)}$: Due to p_T -dep Ψ_n (expected in hydro model)

PCA – Eigenvalues for $v_2(p_T)$



- **Band:** PCA applied to purely statistical fluctuations
- Negative eigenvalues of $V_{n\Delta}(p_1, p_2)$ are compatible with those of large random matrices. Can be attributed to stat. fluct.
- Note the few leading eigenmodes which clearly stand out

PCA – Eigenvalues for $v_3(p_T)$



- **Band:** PCA applied to purely statistical fluctuations
- Negative eigenvalues of $V_{n\Delta}(p_1, p_2)$ are compatible with those of large random matrices. Can be attributed to stat. fluct.
- Note the few leading eigenmodes which clearly stand out

- Initial conditions from HIJING 2.0
(Deng, Wang, Xu 2011)
- These contain nontrivial e-by-e fluctuations
- Flow generated mainly as a result of partonic cascade
- Resonance formations & decays → nonflow effects
- In agreement with v_2 to v_6 versus p_T for all centralities at 2.76 TeV. (Except perhaps, ultra-central)
- 0-10% centrality, $\sim 10^4$ events