Studies of correlations between measurements of jet observables

Andrii Verbytskyi[a]

[a]Max-Planck-Institut für Physik
For the VCES2016, 1-2 December 2016, Vienna, Austria

See also arXiv:1609.06898

Introduction and goals

In high energy particle collisions the partons and hadrons are produced in collimated bunches called *jets*. The studies of jet production in e^+e^-, $e^{\pm}p$ and pp collisions are important for validation of the Quantum Chromodynamics (QCD) theory. Comparisons of corresponding measurements to fixed order or resummed perturbative QCD predictions are used as an ultimate test of the theory. For this reason a good understanding of the experimental uncertainties of the measurements and correlations between them is needed. The aim of this work is to describe a method for robust estimation of these from data only and demonstrate the method using toy Monte Carlo (MC) simulated samples [1].

Jet algorithms

Jet finding algorithm is a procedure to reconstruct kinematics of hard interaction by combining momenta and energy of charged and neutral hadrons. It can be briefly demonstrated with the e^+e^- k_T (Durham) algorithm [2]. A distance measure is defined for pairs of particles i and j via their energies E_i, E_j and angle between them θ_{ij} and the total visible energy in the event E_{vis}.

$$y_{ij} = \frac{2min(E_i^2, E_j^2)(1-\cos\theta_{ij})}{E_{\text{vis}}^2}. \quad (1)$$

Starting from the smallest y_{ij}, the particles i and j are combined until $y_{ij} < y$. The final combinations of particles are desired jets.

Different quantities of interest can be measured with jet algorithms, e.g. e^+e^- **$\mathcal{N}$-jet rates** $R_{\mathcal{N}}$ – the multijet cross sections with $e^+e^- - k_T$ algorithm [3, 4, 5, 6] as functions of y normalised to the total hadronic cross section; pp **multijet cross sections** in bins of transverse momentum of leading jets(LJ) [7] reconstructed with the pp-anti-k_T algorithm [8].

In the measurements every event can contribute to several p_T bins or y points, leading to correlation between them. The difficulties with the calculation of the covariance matrices in this case are solved with introduction of **'event classes'**.

Event classes

Let us fix as set of n cut parameters $y_1 < y_2 \cdots < y_n$ or $p_{T1} < \cdots < p_{Tn}$. Every event can be assigned to a **class**, a set of natural numbers $r_1 \dots r_n$, where r_i is the number of jets with cut y_i, e.g. event with 4 jets at y_1, 3 jets at y_2, 3 jets at y_3 and 2 jets at y_4 belongs to class $\{4332\}$.

- **Classes can be constructed for any algorithm in e^+e^-, $e^{\pm}p$ and pp collisions.**
- **Each event belongs to one class only.**
- **Number of events in classes are linearly related to $\mathcal{N}$-jet rates, jet multiplicities, multijet cross sections etc., e.g. for e^+e^-, $n=4$ with k_T algorithm:**

$$R_3(y_3) = 1/N_{\text{events}} \sum_{r_3=3} \{r_1 r_2 r_3 r_4\} =$$
$$= 1/N_{\text{events}}(\{3332\} + \{3333\} + \{4333\} + \{4433\}).$$

Hereby, the construction of **exact** covariance matrix $\boldsymbol{W}$ for derived quantities from the covariance matrices of event classes is trivial.

- **Events in the same class have similar topology and are expected to have similar detector corrections.**
- **All corrections (e.g. detector or hadronisation) applied to classes are simultaneously and consistently propagated to derived quantities.**

Application to MC simulated samples

100000 e^+e^-, $\sqrt{s} = 91$GeV events by SHERPA2.2 [9] with default parameters and $\alpha_s(M_Z) = 0.12$. Durham jet algorithm is applied with y points $10^{-3.2}, 10^{-2.4}, 10^{-1.6}, 10^{-0.8}$. The analysis is similar to one from ALEPH [3]. R_4 includes also higher multiplicities.

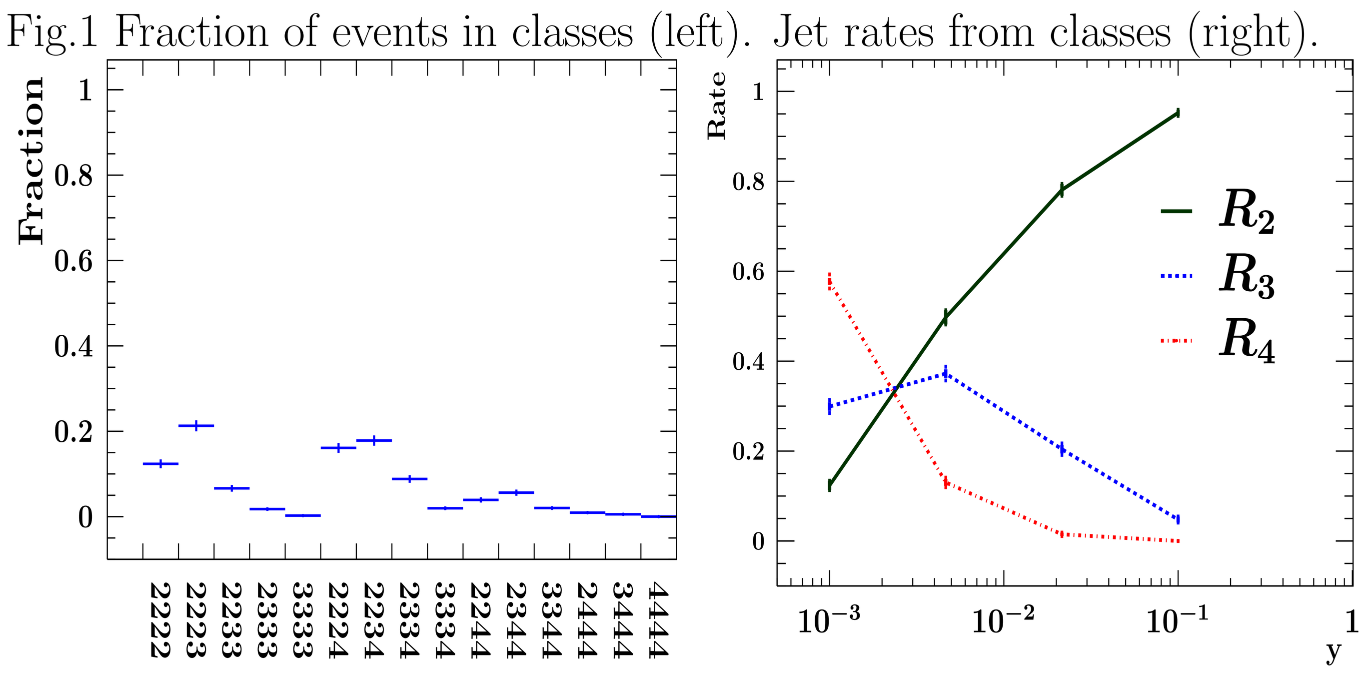


Fig.1 Fraction of events in classes (left). Jet rates from classes (right).

Tab.1 Corr. matrix W^R for jet rates.

+1.00	+0.38	+0.20	+0.08	-0.25	-0.29	-0.19	-0.08	-0.44	-0.15	-0.05	-0.00	$R_2(y_1)$
	+1.00	+0.53	+0.22	+0.28	-0.77	-0.50	-0.22	-0.51	-0.38	-0.12	-0.00	$R_2(y_2)$
		+1.00	+0.42	+0.24	-0.23	-0.96	-0.42	-0.36	-0.45	-0.23	-0.01	$R_2(y_3)$
			+1.00	+0.12	-0.04	-0.38	-1.00	-0.17	-0.27	-0.19	-0.02	$R_2(y_4)$
				+1.00	-0.11	-0.22	-0.12	-0.76	-0.25	-0.08	-0.00	$R_3(y_1)$
					+1.00	+0.27	+0.04	+0.30	-0.30	-0.09	-0.00	$R_3(y_2)$
						+1.00	+0.38	+0.33	+0.37	-0.06	-0.00	$R_3(y_3)$
							+1.00	+0.17	+0.27	+0.19	-0.00	$R_3(y_4)$
								+1.00	+0.33	+0.10	+0.00	$R_4(y_1)$
									+1.00	+0.32	+0.01	$R_4(y_2)$
										+1.00	+0.04	$R_4(y_3)$
											+1.00	$R_4(y_4)$
$R_2(y_1)$	$R_2(y_2)$	$R_2(y_3)$	$R_2(y_4)$	$R_3(y_1)$	$R_3(y_2)$	$R_3(y_3)$	$R_3(y_4)$	$R_4(y_1)$	$R_4(y_2)$	$R_4(y_3)$	$R_4(y_4)$	

250000 pp, $\sqrt{s} = 7$TeV events by SHERPA2.2 [9] with default parameters, $\alpha_s(M_Z) = 0.12$ and HERAPDF2.0 NNLO PDFs [10]. Anti-k_T jet algorithm with $R = 0.4$ is applied with p_T bins $60, 80, 110, 160, 210$GeV. The analysis is similar to one from ATLAS [7].

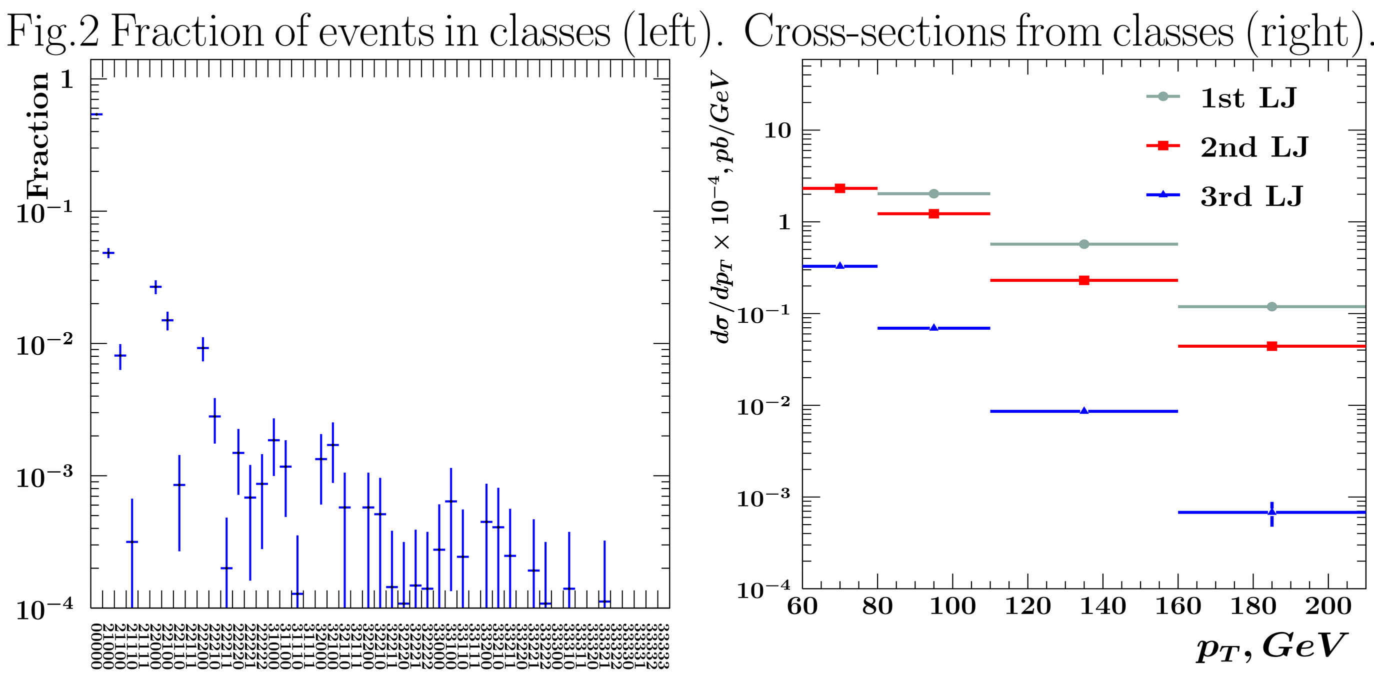


Fig.2 Fraction of events in classes (left). Cross-sections from classes (right).

Tab.2 Corr. matrix W^{CS} for multijet cross sections.

+1.00	-0.06	-0.03	+0.71	+0.43	-0.04	-0.02	+0.10	+0.00	-0.01	-0.00	1LJ, b.2
	+1.00	-0.02	+0.16	+0.39	+0.43	-0.01	+0.18	+0.10	+0.02	-0.00	1LJ, b.3
		+1.00	-0.00	+0.07	+0.35	+0.36	+0.16	+0.15	+0.09	-0.00	1LJ, b.4
			+1.00	-0.06	-0.03	-0.01	+0.12	-0.01	-0.01	-0.00	2LJ, b.1
				+1.00	-0.03	-0.01	+0.17	+0.10	-0.01	-0.00	2LJ, b.2
					+1.00	-0.01	+0.10	+0.17	+0.11	-0.00	2LJ, b.3
						+1.00	+0.05	+0.09	+0.12	+0.03	2LJ, b.4
							+1.00	-0.00	-0.00	-0.00	3LJ, b.1
								+1.00	-0.00	-0.00	3LJ, b.2
									+1.00	-0.00	3LJ, b.3
										+1.00	3LJ, b.4
1LJ, b.2	1LJ, b.3	1LJ, b.4	2LJ, b.1	2LJ, b.2	2LJ, b.3	2LJ, b.4	3LJ, b.1	3LJ, b.2	3LJ, b.3	3LJ, b.4	

Comparison to sampling method

The sampling method for covariance matrix **estimation**:

- Build N_{sub} subsamples from measured or MC simulated events (in case of limited data statistics).
- Estimate W as

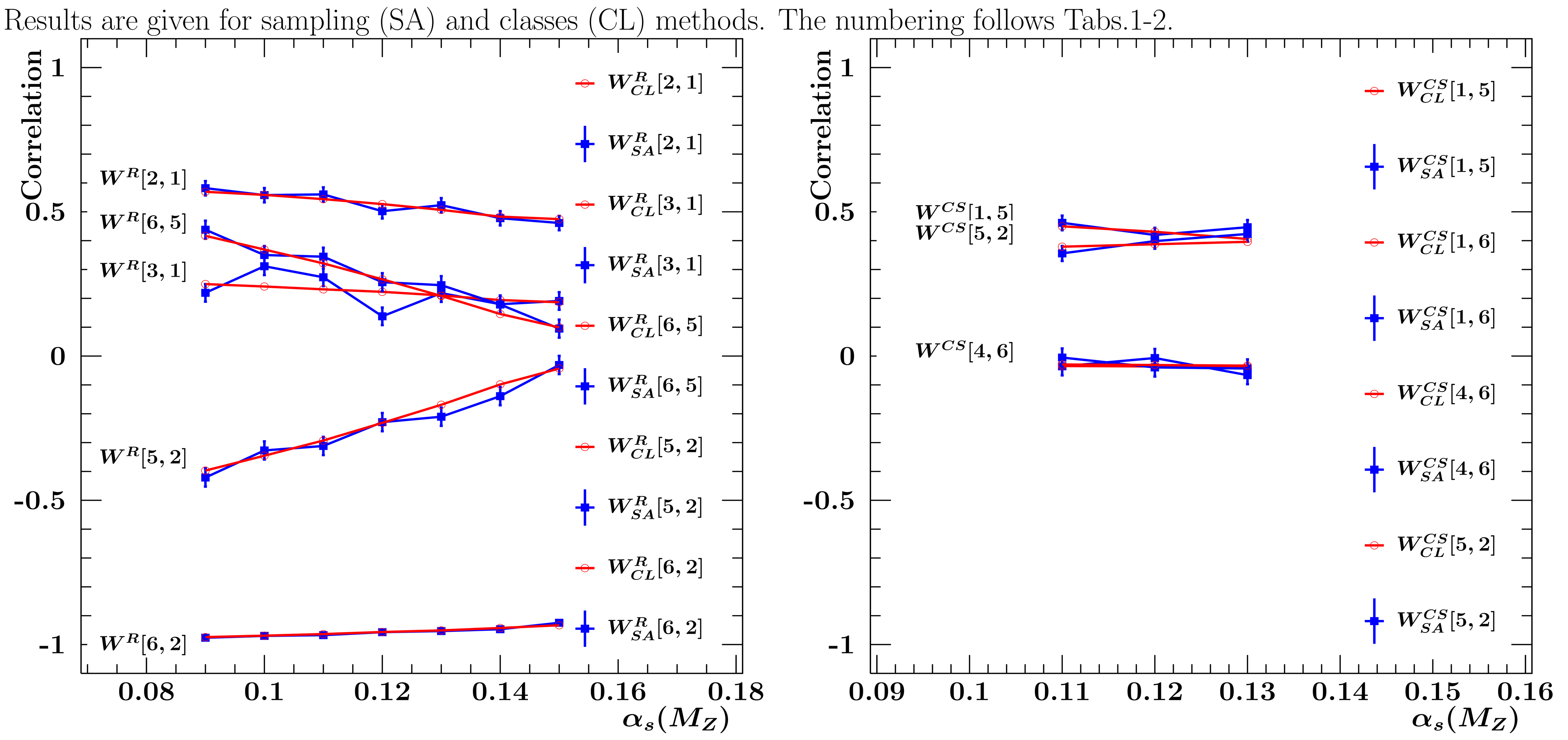
$$W^R[i,j] = 1/(N_{\text{sub}}-1) \sum_{k=1\dots N_{\text{sub}}} (R_{i,k} - \bar{R}_i)(R_{j,k} - \bar{R}_j).$$

- Delivers **only estimation** with uncertainty [11]

$$\delta(arctanh(W^R[i,j])) \propto 1/\sqrt{N_{\text{sub}} - 3}$$

- Used in earlier studies, e.g. in Refs. [5, 6].
- Computing (CPU) demanding.
- If the MC events are used for sampling, the result depends on the MC model.

Fig.3 Dependence of selected elements of correlation matrices on the $\alpha_s(M_Z)$ in simulation of e^+e^- (left) and pp (right) samples. Results are given for sampling (SA) and classes (CL) methods. The numbering follows Tabs.1-2.



For every $\alpha_s(M_Z)$ value 100000 e^+e^-/250000 pp events are simulated and used to build 1000 subsamples. **Results from SA and CL methods are numerically close. CL results are much more stable. Dependence of results on MC model emphasises a need of model independent estimation of the correlations, which requires sampling of the data with huge number of subsamples or application of CL method.**

Conclusions

A new type of jet observables, classes, was introduced. A method to calculate correlations between measurements of jet observables with classes was demonstrated. The method is applicable to any jet algorithm, has high precision, provides robust results, is not computing demanding, does not rely on the MC simulations and is simple to implement.

References

[1] A. Verbytskyi, Studies of correlations between measurements of jet observables. (2016). arXiv:1609.06898.

[2] S. Catani et al., New clustering algorithm for multi-jet cross-sections in e^+e^- annihilation. Phys. Lett. **B269**, 432 (1991).

[3] ALEPH Coll., A. Heister et al., Studies of QCD at e^+e^- centre-of-mass energies between 91GeV and 209GeV. Eur. Phys. J. **C35**, 457 (2004).

[4] OPAL Coll., G. Abbiendi et al., Determination of α_s using OPAL hadronic event shapes at $\sqrt{s} = 91 - 209$GeV and resummed NNLO calculations. Eur. Phys. J. **C71**, 1733 (2011). arXiv:1101.1470.

[5] JADE Coll., J. Schieck et al., Measurement of the strong coupling α_s from the three-jet rate in e^+e^- - annihilation using JADE data. Eur. Phys. J. **C73**, 2332 (2013). arXiv:1205.3714.

[6] M.E. Zomorrodian, M. Hasheminia and A. Mirjalili, Measurement of strong coupling α_s in e^+ e^- annihilation using jet rate and event shape. Indian J. Phys. **90**, 469 (2016).

[7] ATLAS Coll., G. Aad et al., Measurement of multi-jet cross sections in proton-proton collisions at a 7TeV center-of-mass energy. Eur. Phys. J. **C71**, 1763 (2011). arXiv:1107.2092.

[8] M. Cacciari, P.G. Salam and G. Soyez, The Anti-k_T jet clustering algorithm. JHEP **04**, 063 (2008). arXiv:0802.1189.

[9] T. Gleisberg et al., Event generation with SHERPA 1.1. JHEP **02**, 007 (2009). arXiv:0811.4622.

[10] ZEUS and H1 Colls., H. Abramowicz et al., Combination of measurements of inclusive deep inelastic $e^{\pm}p$ scattering cross sections and QCD analysis of HERA data. Eur. Phys. J. **C75**, 580 (2015). arXiv:1506.06042.

[11] R.A. Fisher, Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. Biometrika **10**, 507 (1915).

Thanks to S. Kluth, O. Zenaiev and O. Behnke