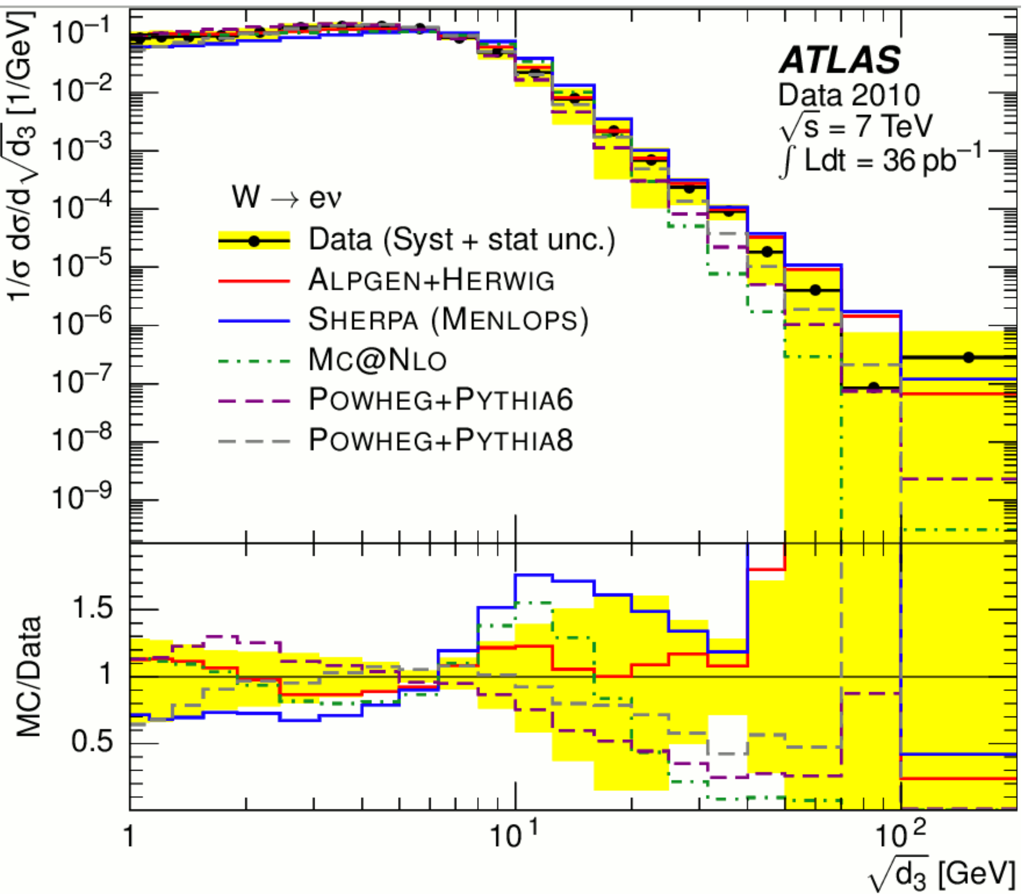
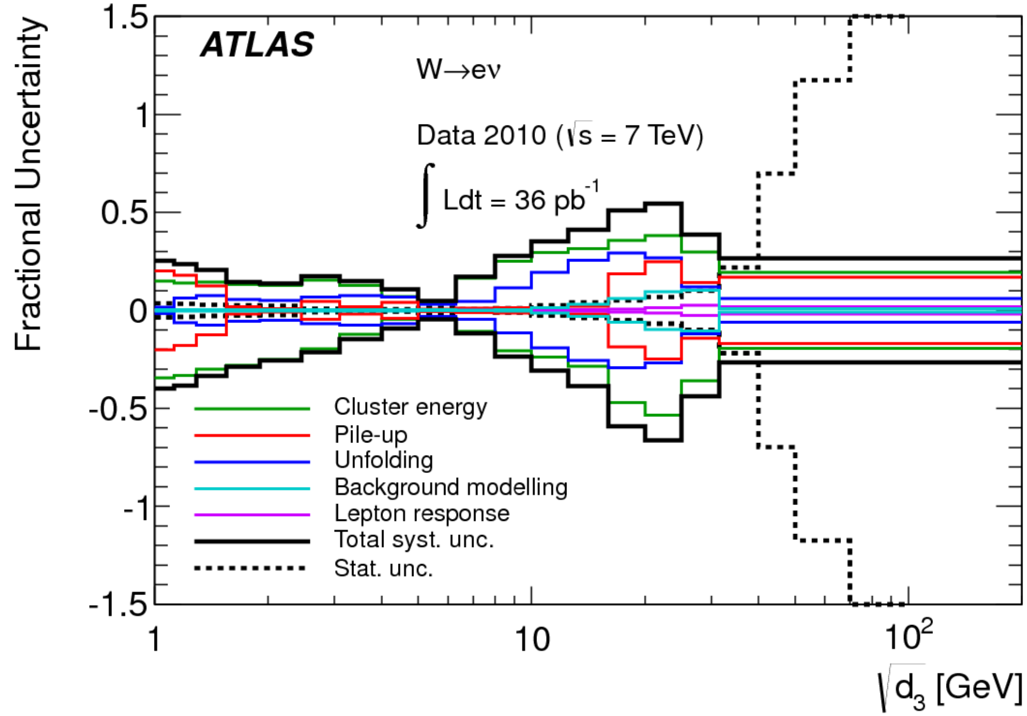
Correlations

many analyses produce systematics-dominated results.

eg: Measurement of kt splitting scales in W->lnu events at sqrt(s) = 7 TeV
with the ATLAS detector

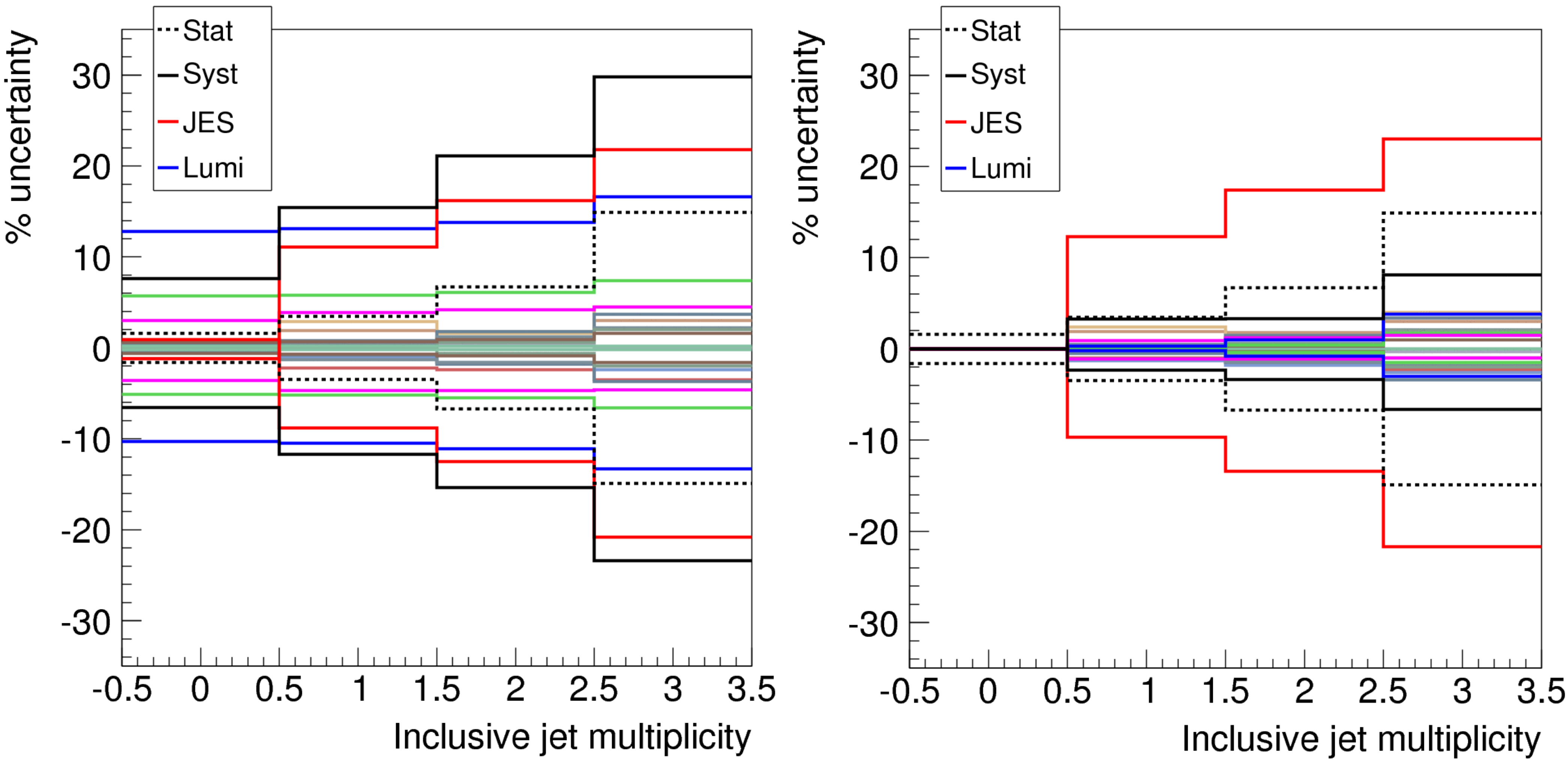




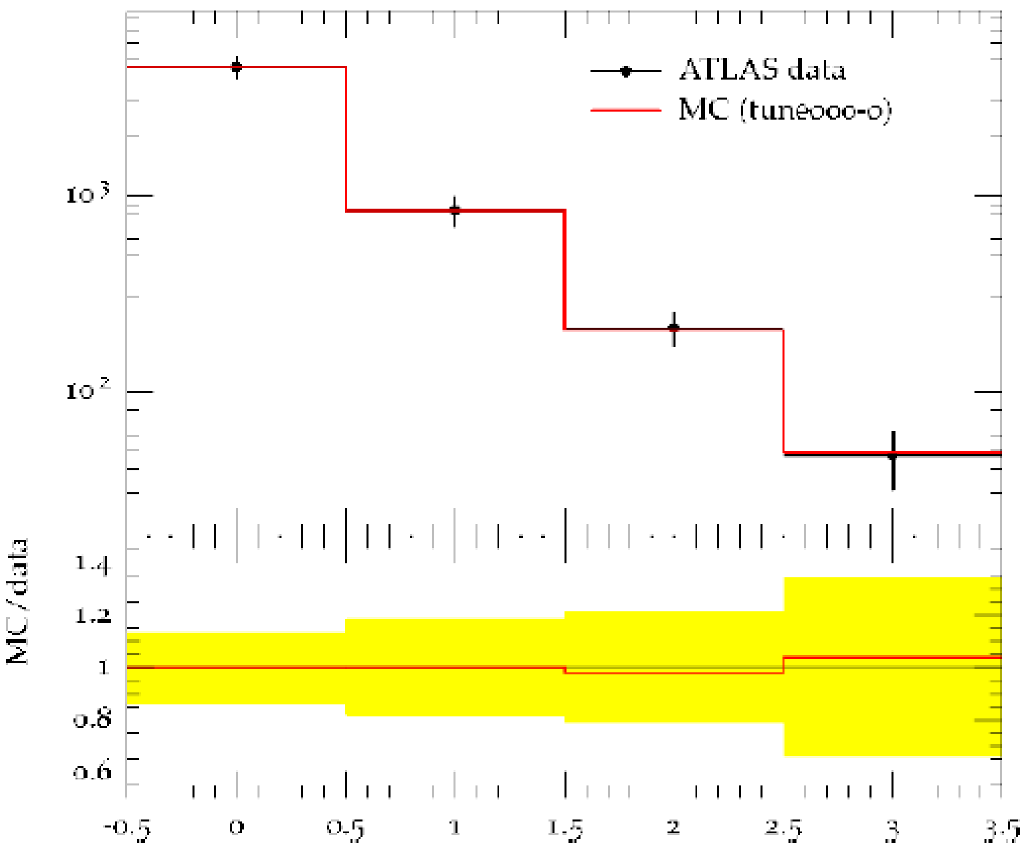
it is incorrect to take the per bin uncertainties as uncorrelated!

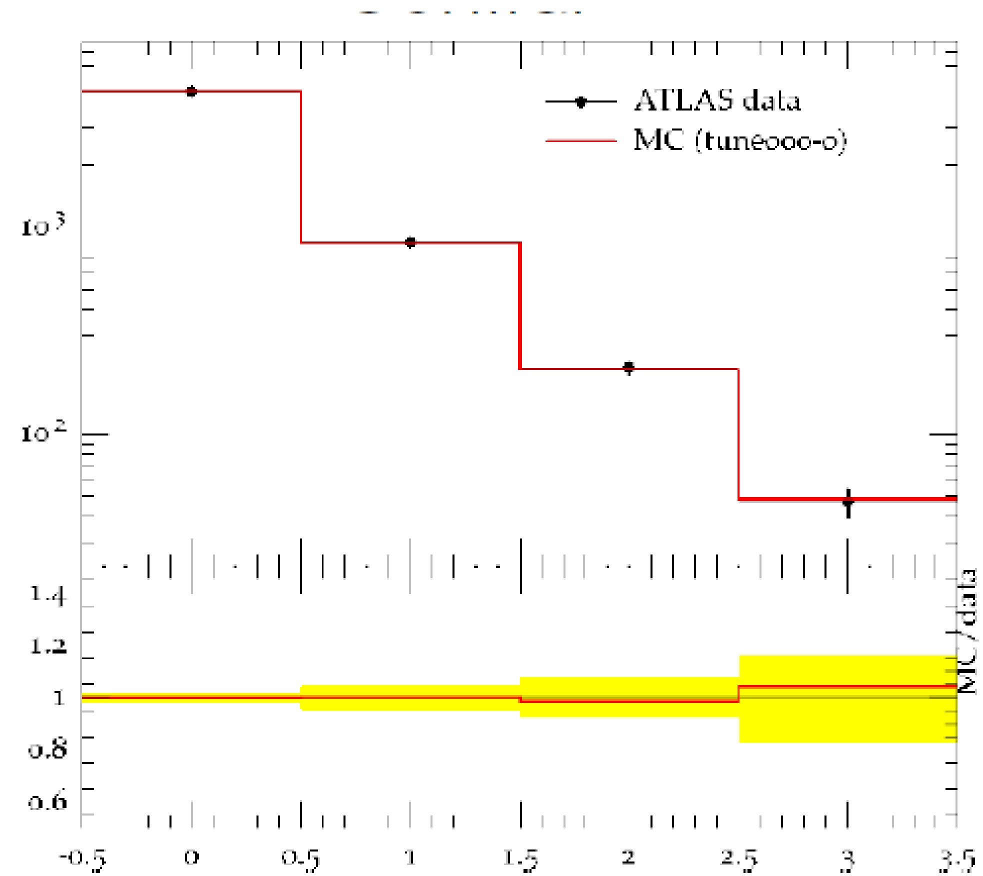
Correlations

An old example (24/3/11), looking at W+jet events



Correlations





Correlations

In the case of a single distribution, a covariance matrix is fine
But for many distributions, probably not:
- eg Z+jets: jet1 pT, jet 2pT, jet 1 y, jet 2 y, dR(j,j),.....
- many distributions with the same correlated uncertainty (JES)
- could build a giant covariance matrix (100's x 100's)
- or publish a list of uncertainty sources

TABLE VII: The measured cross section in bins of $\Delta\phi(Z, \text{jet})$ for Z/γ^*+jet+X events with $p_T^Z > 25$ GeV, normalized to the measured Z/γ^* cross section.

$\Delta\phi$ (rad)	$\langle\Delta\phi\rangle$ (rad)	result (1/rad)	stat. unc. (%)	uncorr. unc. (%)	source 1 (%)	source 2 (%)	source 3 (%)	source 4 (%)	source 5 (%)	source 6 (%)	source 7 (%)	source 8 (%)	source 9 (%)	source 10 (%)
0.0-1.5	1.09	0.000282	±12.	±4.0	16. -20.	0.7 -0.7	-1.9 1.9	-1.0 1.0	8.0 -8.0	-3.2 1.3	-3.8 5.3	2.4 -2.4	-0.3 0.1	-4.4 6.3
1.5-2.2	1.95	0.00422	±6.9	±0.8	3.2 -4.2	0.7 -0.7	-1.0 1.0	-0.5 0.5	2.1 -2.1	-3.5 4.2	-4.4 3.4	1.3 -1.3	-0.8 -0.7	-3.7 2.9
2.2-2.5	2.38	0.0193	±5.5	±0.9	1.4 -1.6	0.5 -0.5	-0.6 0.6	-0.3 0.3	0.5 -0.5	-2.6 1.6	-3.7 3.3	0.1 -0.1	-0.3 -0.2	-2.8 2.4
2.5-2.7	2.61	0.0527	±4.2	±0.8	1.1 -1.4	0.5 -0.5	-0.3 0.3	-0.2 0.2	0.6 -0.6	-1.7 1.9	-2.4 2.3	-1.2 1.2	-0.1 -0.3	-2.9 2.8
2.7-2.9	2.81	0.113	±2.8	±0.6	0.8 -2.1	0.5 -0.5	-0.3 0.3	-0.1 0.1	0.8 -0.8	-1.2 1.1	-1.7 1.8	-0.9 0.8	0.1 0.2	-1.7 2.1
2.9-3.2	3.04	0.332	±1.7	±0.4	-0.3 0.3	0.2 -0.2	-0.5 0.5	0.0 -0.0	-0.4 0.4	2.2 -1.8	-1.0 1.2	-1.1 1.1	0.2 0.0	-1.3 1.3

TABLE VIII: The measured cross section in bins of $|\Delta y(Z, \text{jet})|$ for Z/γ^*+jet+X events with $p_T^Z > 25$ GeV, normalized to the measured Z/γ^* cross section.

$\|\Delta y\|$	$\langle\|\Delta y\|\rangle$	result	stat. unc. (%)	uncorr. unc. (%)	source 1 (%)	source 2 (%)	source 3 (%)	source 4 (%)	source 5 (%)	source 6 (%)	source 7 (%)	source 8 (%)	source 9 (%)	source 10 (%)
0.00-0.40	0.21	0.0791	±2.6	±0.6	0.5 -0.5	0.3 -0.3	-0.4 0.4	-0.1 0.1	0.1 -0.1	0.8 -0.3	-1.6 1.8	-1.1 1.1	0.3 0.2	-1.4 1.6
0.40-0.80	0.61	0.0679	±2.8	±0.6	0.1 -0.1	0.3 -0.3	-0.4 0.4	-0.1 0.1	-0.6 0.6	0.4 -0.4	-1.5 1.6	-0.9 0.9	-0.3 -0.2	-1.6 1.5
0.80-1.20	1.02	0.0568	±3.0	±0.7	-0.1 0.1	0.4 -0.4	-0.4 0.4	-0.1 0.1	-0.0 0.0	0.5 -0.5	-1.4 1.6	-1.1 1.1	0.1 0.1	-1.5 1.6
1.20-1.55	1.37	0.0452	±3.6	±0.9	-0.2 0.2	0.2 -0.2	-0.4 0.4	-0.1 0.1	0.1 -0.1	0.3 -0.8	-1.6 1.4	-0.3 0.3	-0.1 -0.6	-1.8 1.6
1.55-2.05	1.78	0.0274	±3.8	±0.9	-0.4 0.4	-0.2 0.2	-0.4 0.4	-0.1 0.1	1.1 -1.1	1.1 -0.1	-1.3 1.8	-1.2 1.2	0.4 0.4	-1.8 2.5
2.05-4.50	2.89	0.00480	±4.0	±1.1	-0.6 0.6	-1.1 1.1	-0.5 0.5	-0.1 0.1	1.3 -1.3	-0.1 -0.8	-2.0 1.5	-0.5 0.5	-0.0 -0.1	-3.0 2.6

So, can a YODA histogram come with a list of systematic errors per bin?
- do we even need this, or will PROFESSOR use a different format?

alpgen

the same probably holds for madgraph

Breaks down a process into slices based on number of final state partons
eg Z+0lp, Z+1lp, Z+2lp, Z+ >=3lp
- these must be combined to obtain a full sample
- combine by normalising each slice to the same lumi, then adding.

Generating each slice involves running three steps:
1) generate weighed events
2) from this, produce some unweighted events (parton level 4-vectors)
 - produces N events, with a cross section X
 - corresponding to a luminosity L

3) run these 4-vectrors through pythia/herwig/...
 - this is the AGILE step
 - the MLM matching is applied here
   - changes the cross section and the number of events
   - the luminosity is constant
 - the final cross section & number of events per slice Is only known **after** this step.

So, how to implement this in RIVET?

My hack job:

1) Any normalisation must be turned off:
- eg 1/sigma d sigma / d X

this is only known after combining all slices

2) read in the luminosity from step 2, normalise all plots to 1/L

3) run over **all** the events prodiced in step 2
- note you can't specify the number in advance, due to MLM
- specift a large number, or N, even though this many will not be processed.

4) use a custom script to add all the resulting plots from the various slices
- and then calculate eg 1/sigma d sigma / d X