### **The Durham HepData Project**

REACTION DATABASE . DATA REVIEWS . PDF PLOTTER



# HepData status+data

Contact us at: hepdata@projects.hepforge.org

HepData is funded by the UK STFC and hosted at the Durham IPPP. HepData also maintains the UK mirror of the PDG.



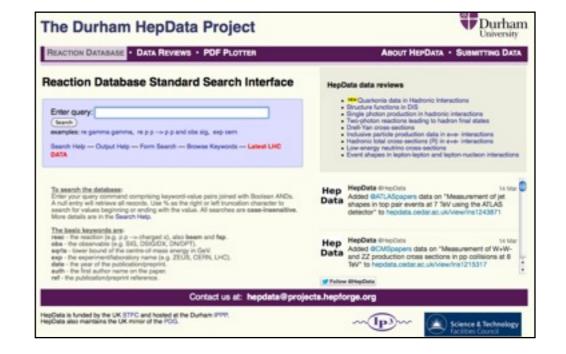
Science & Technology Facilities Council

Durham

University

### Mike Whalley IPPP Durham U., UK

AAHEP7 Information Provider Summit Stony Brook - 2nd/3rd April 2014



# Outline of talk

- Brief introduction to HepData with examples
- How does data get into HepData ?
- How should data get into HepData ?
- The HepData-Inspire connection
- Datasets and DOIs
- User/Data statistics
- Summary and Conclusions

### Brief Introduction to HepData

- Purpose to compile (mainly) published HEP 'cross section' data and make them publicly available.
- Small group, based at the IPPP, Durham U. (UK) -
  - DB Manager/Physicist (MW retiring)
  - New DB Manager/Physicist (GraemeWatt)
  - Non-Physicist Assistant
- STFC(UK) funded (>30 years) presently to October 2016.
- 2009 moved to more modern and long-term maintainable computing system.

# What exactly is HepData ?

- DataBase\* essentially of Standard Model cross sections and measurements dating back over 30 years.
- Files\*\* collections of data of varying format systematic error breakdowns, correlation matrices, slha files, acceptances, efficiencies, etc....
- Linked together via a Web page
- Different output formats eg. html, plain text, xml, plots
- Yoda files for **Rivet** input

\* mysql + Java model using Hibernate for database persistency and Tapestry for web pages
 \*\* resource area on our main server

# The HepData 'home' page

### The Durham HepData Project



REACTION DATABASE . DATA REVIEWS . PDF PLOTTER

#### ABOUT HEPDATA · SUBMITTING DATA

### Reaction Database Standard Search Interface



#### HepData data reviews

- NEW Quarkonia data in Hadronic Interactions
- Structure functions in DIS
- Single photon production in hadronic interactions
- Two-photon reactions leading to hadron final states
- Drell-Yan cross-sections
- Inclusive particle production data in e+e- interactions
- Hadronic total cross-sections (R) in e+e- interactions
- Low-energy neutrino cross-sections
- Event shapes in lepton-lepton and lepton-nucleon interactions

### HepData @HepData

Added @ATLASpapers data on "Measurement of jet shapes in top pair events at 7 TeV using the ATLAS detector" to hepdata.cedar.ac.uk/view/ins1243871

# HepData @HepData

14 Mar Added @CMSpapers data on "Measurement of W+Wand ZZ production cross sections in pp collisions at 8 TeV" to hepdata.cedar.ac.uk/view/ins1215317

. Ŧ

Follow @HepData

communication HepData also maintains the UK mirror of the PDG.





Wednesday, 2 April 2014

### Initial search results

### **The Durham HepData Project**

REACTION DATABASE · DATA REVIEWS · PDF PLOTTER

### Reaction Database Search Result Search: exp Ihc Result: 261 documents found (displaying 1 to 20) First | Previous | Next | Last | All Enter query: exp lbc Search again ... need help with searching? AAD 2014 – Experiment: CERN-LHC-ATLAS (ATLAS) Preprint: CERN-PH-EP-2013-184 Archive: ARXIV:1401.2831 Measurement of the production cross section of prompt J/psi mesons in association with a W boson in pp collisions at sqrt(s) = 7 TeV with the ATLAS detector Full data record | Short data record | INSPIRE | CDS ABELEV 2014 – Experiment: CERN-LHC-ALICE (ALICE) Preprint: CERN-PH-EP-2013-230 Archive: ARXIV:1401.1250 Production of charged pions, kaons and protons at large transverse momenta in pp and Pb-Pb collisions at sqrt(sNN) = 2.76 TeV Full data record | Short data record | INSPIRE | CDS

3. AAD 2013 - Experiment: CERN-LHC-ATLAS (ATLAS) Preprint: CERN-PH-EP-2013-192 Archive: ARXIV:1312.3524

Measurement of dijet cross sections in pp collisions at 7 TeV centre-of-mass energy using the ATLAS detector Full data record | Short data record | INSPIRE | CDS

4. CHATRCHYAN 2014 – Experiment: CERN-LHC-CMS (CMS) Published: PL B729,149 (2014) Preprint: CERN-PH-EP-2013-215 Archive: ARXIV:1311.7667 Inclusive search for a vector-like T quark with charge 2/3 in pp collisions at sqrt(s) = 8 TeV Full data record | Short data record | INSPIRE | CDS

5. AAD 2013 - Experiment: CERN-LHC-ATLAS (ATLAS)



ABOUT HEPDATA · SUBMITTING DATA

# Example of a HepData record

### **The Durham HepData Project**

REACTION DATABASE · DATA REVIEWS · PDF PLOTTER

#### **Reaction Database Full Record Display**

View short record or as: input, plain text, AIDA, PyROOT, YODA, ROOT, mpl, ScaVis or MarcXML

### AAD 2014 — Measurement of the production cross section of prompt J/psi mesons in association with a W boson in pp collisions at sqrt(s) = 7 TeV with the ATLAS detector

Experiment: CERN-LHC-ATLAS (ATLAS) Preprinted as CERN-PH-EP-2013-184 Archived as: ARXIV:1401.2831 Auxiliary Material: http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/BPHY-2012-06/ Record in: INSPIRE Record in: CERN Document Server

CERN-LHC. The process P P --> W J/PSI X provides a powerful probe of the production mechanism of charmonium in hadronic collisions, and is also sensitive to multiple parton interactions in the colliding protons. Using the 2011 ATLAS dataset of 4.5 fb^-1 of sqrt(s) = 7 TeV pp collisions at the LHC, the first observation is made of the production of W + prompt J/psi events in hadronic collisions, using W --> MU NU and J/PSI --> MU MU. A yield of 27.4 +7.5,-6.5 W + prompt J/psi events is observed, with a statistical significance of 5.1 sigma. The production rate as a ratio to the inclusive W boson production rate is measured, and the double parton scattering contribution to the cross section is estimated.

Table 1 (Figure 5.) cor as: input, plain text, AIDA, PyROOT, YODA, ROOT, mpl, ScaVis or MarcXML

The W + prompt J/psi to inclusive W production cross-section ratio (times 10^6) in the J/psi fiducial region (Fiducial), after correction for J/psi acceptance (Inclusive), and after subtraction of the double parton scattering component (DPS-subtracted). The first uncertainty is statistical, the second is systematic, and the third/fourth (where applicable) is the uncertainty up/down due to spin-alignment.

ABS(YRAP(J/PSI))	0.0-2.1	Si 3 <sup>×10<sup>6</sup></sup>
PT(J/PSI)	8.5-30.0 GeV	PD = Deta
RE	P P> W+ < MU+ NU > J/PSI < MU+ MU- > X	
RE	P P> W- < MU- NUBAR > J/PSI < MU+ MU- > X	2 2 Spin alignment uncertainty I C CS voluting 2 feedbarn N 0 CO prediction
SQRT(S)	7000.0 GeV	1.5 ·
Ratio	10**6 * BR(J/PSI -> MU+ MU-) * 1/SIG(W) * D(SIG(W JPSI))/DYRAP(J/PSI)	
Fiducial	0.51 ± 0.13 (stat) ± 0.04 (sys)	a 0.5
Inclusive	1.26 ± 0.32 (stat) ± 0.09 (sys) +0.41,-0.25 (sys,spin-alignment)	0 <sup>2</sup> Fiducial Inclusive DPS-subtracted
DPS-subtracted	0.78 ± 0.32 (stat) ± 0.22 (sys) +0.41,-0.25 (sys,spin-alignment)	



ABOUT HEPDATA · SUBMITTING DATA

### another example....

### **The Durham HepData Project**

REACTION DATABASE · DATA REVIEWS · PDF PLOTTER

#### Reaction Database Full Record Display

View short record or as: input, plain text, AIDA, PyROOT, YODA, ROOT, mpl, ScaVis or MarcXML

#### ABELEV 2014 — Production of charged pions, kaons and protons at large transverse momenta in pp and Pb-Pb collisions at sqrt(sNN) = 2.76 TeV

Experiment: CERN-LHC-ALICE (ALICE) Preprinted as CERN-PH-EP-2013-230 Archived as: ARXIV:1401.1250 Record in: INSPIRE Record in: CERN Document Server

CERN-LHC. Measurement of transverse momentum (pT) spectra of charged pions, kaons and protons up to pT = 20 GeV at mid-rapidity in Pb-Pb and pp collisions at sqrt(sNN) = 2.76 TeV, recorded in 2010 and 2011, respectively, using the ALICE detector at the LHC. At intermediate pT (2-8 GeV) an enhancement of the proton-to-pion ratio with respect to pp collisions is observed and the ratio reaches ~0.80 in central Pb-Pb collisions. The measurement of the nuclear modification factors for charged pions, kaons and protons indicates that within the systematic and statistical uncertainties they are the same at high pT (> 10 GeV).

Table 1 (Figure 2 (left).) cor as: input, plain text, AIDA, PyROOT, YODA, ROOT, mpl, ScaVis or MarcXML Invariant yields of identified pions in central and peripheral Pb-Pb collisions, together with the unscaled pp reference yields.

CENTRALITY	0 TO 5 pct	60 TO 80 pct	÷
ETARAP	-0.8-0.8		
RE	PB PB> (PI+ + PI-) X	PB PB> (PI+ + PI-) X	P P> (PI+ + PI-) X
SQRT(S)/NUCLEON	2760.0 GeV		
PT IN GEV		1/(Nev*2*PI*PT)*D2(N)/DPT/DYRAP IN GEV**	-2
2.0 - 2.2	2.302 ± 0.004 (stat) ± 0.210 (sys)	0.06735 ± 0.00025 (stat) ± 0.00714 (sys)	0.003910 ± 0.000011 (stat) ± 0.000319 (sys)
2.2 - 2.4	1.345 ± 0.002 (stat) ± 0.122 (sys)	0.04154 ± 0.00016 (stat) ± 0.00439 (sys)	0.002432 ± 0.000008 (stat) ± 0.000199 (sys)
2.4 - 2.6	0.8011 ± 0.0015 (stat) ± 0.0728 (sys)	0.02617 ± 0.00011 (stat) ± 0.00276 (sys)	0.001539 ± 0.000006 (stat) ± 0.000126 (sys)
2.6 - 2.8	0.4828 ± 0.0009 (stat) ± 0.0438 (sys)	0.01694 ± 0.00008 (stat) ± 0.00179 (sys)	0.001011 ± 0.000004 (stat) ± 0.000083 (sys)
2.8 - 3.0	0.2960 ± 0.0006 (stat) ± 0.0268 (sys)	0.01124 ± 0.00006 (stat) ± 0.00118 (sys)	0.0006816 ± 0.0000043 (stat) ± 0.0000557 (sys)
3.0 - 3.2	0.1834 ± 0.0006 (stat) ± 0.0166 (sys)	0.007594 ± 0.000058 (stat) ± 0.000796 (sys)	0.0004670 ± 0.0000025 (stat) ± 0.0000382 (sys)
3.2 - 3.4	0.1168 ± 0.0004 (stat) ± 0.0106 (sys)	0.005238 ± 0.000043 (stat) ± 0.000548 (sys)	0.0003254 ± 0.0000019 (stat) ± 0.0000266 (sys)



ABOUT HEPDATA · SUBMITTING DATA

# example showing link to 'resource area' file

REACTION DATABASE · DATA REVIEWS · PDF PLOTTER

ABOUT HEPDATA · SUBMITTING DATA

#### Reaction Database Full Record Display

View short record or as: input, plain text, AIDA, PyROOT, YODA, ROOT, mpl, ScaVis or MarcXML

#### AAD 2013 — Measurement of the high-mass Drell-Yan differential cross-section in pp collisions at sqrt(s)=7 TeV with the ATLAS detector

Experiment: CERN-LHC-ATLAS (ATLAS) Preprinted as CERN-PH-EP-2013-064 Archived as: ARXIV:1305.4192 Record in: INSPIRE

CERN-LHC. Based on an integrated luminosity of 4.9 //b, the differential cross-section in the Z/gamma measured with the ATLAS detector as a function of the invariant mass, Mee, in the range 116 < Mee fiducial region in which both the electron and the positron have transverse momentum pT > 25 GeV a pseudorapidity letal < 2.5..

> Nbkg CDY

Energy scale & res Unfolding

250

300

400

700

1000

1500

250

300

400

500

700

1000

5.2

4.3

5.1

9.4

11

24

0.9

0.7

0.9

2.0

2.0

4.0

7.6

2.4

0.9

1.0

0.8

0.5

3.8

4.1

4.4

4.0

3.1

3.1

1.6

1.6

1.6

1.6

1.7

1.7

2.4

2.4

2.5

2.6

2.6

2.6

2.5

1.2

1.7

1.7

2.4

2.8

3.3

1.5

1.5

1.5

1.5

1.5

1.5

0.9

0.8

0.8

0.8

0.8

0.8

0.8

0.2

0.2

0.3

0.5

0.2

0.2

0.3

0.3

0.2

0.3

0.4

0.3

0.4

0.4

1.8

1.8

1.8

1.8

1.8

1.8

Born level

Link to the full breakdown of the sources of errors

Table 1 (T 2, F 2, 3.) end on as: input, plain text, AIDA, PyROOT, YODA, ROOT, mpl, St modelling Theoretical Measured differential cross sections as a function of the di-electron mass for DY production at the Becami

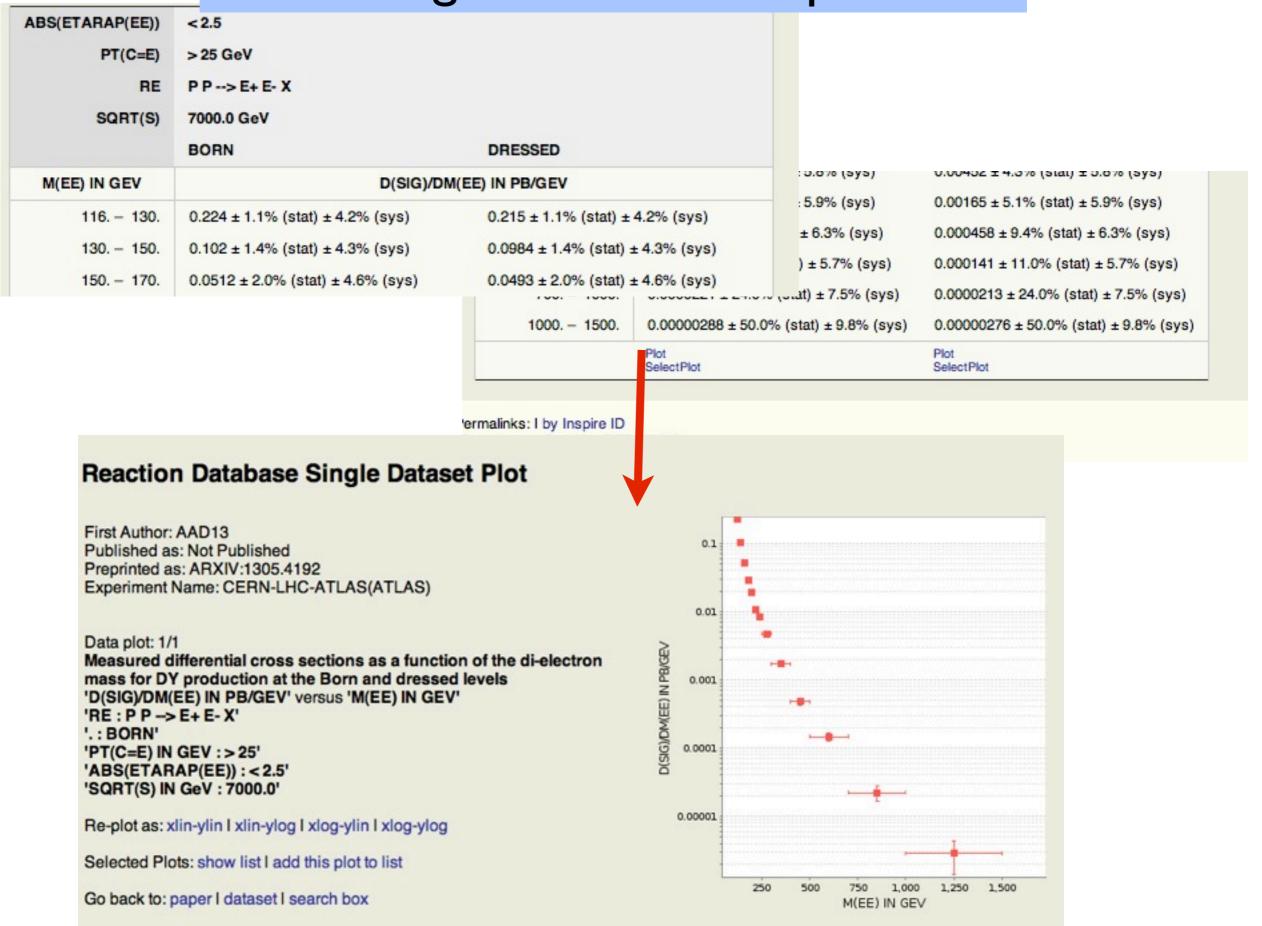
	background events uncertainty MC statistic uncertainty for bin-by-bin unfolding electron reconstruction uncertainty electron identification uncertainty	resource area file
•	uncertainty of the energy scale and energy resolution uncertainty of the unfolding method	
	trigger uncertainty includes all uncertainties of the MC modeling (vertex ;	position, I pT, pile-up, K-factors)

includes all uncertainties of the MC modeling (vertex position, I pT, pile-up, K-factors) uncertainty of theoretical extrapolation to the fiducial region luminosity uncertainty

ABS(ETARAP(EE)) PT(C=E) RE	< 2.5 > 25 GeV P P> E+ E- X	fror	n the database
SQRT(S)	7000.0 GeV BORN		DRESSED
M(EE) IN GEV		D(SIG)/DN	(EE) IN PB/GEV
116 130.	0.224 ± 1.1% (stat) ± 4	.2% (sys)	0.215 ± 1.1% (stat) ± 4.2% (sys)
130 150.	0.102 ± 1.4% (stat) ± 4	.3% (sys)	0.0984 ± 1.4% (stat) ± 4.3% (sys)
150 170.	0.0512 ± 2.0% (stat) ±	4.6% (sys)	0.0493 ± 2.0% (stat) ± 4.6% (sys)

noe min	nee xnax	Stat.	Nbkg stat.	CDY stat.	Nbkg	Reco.	Id.	Energy			MC dell	Theoretical	Lomi
Gel	7	<-unco	rrelated	(8)->	<			-bis-to-b	in correla	ated (1)-			>
	1.1.0												
116	130	1.1	0.1	0.7	1.3	1.6	2.3	2.1	1.5	0.8	0.2	0.3	1.8
130	150	1.4	0.2	0.7	1.8	1.6	2.3	1.7	1.5	0.8	0.5	0.2	1.8
150	170	2.0	0.3	1.0	2.5	1.6	2.3	1.6	1.5	0.8	0.2	0.2	1.8
170	190	2.7	0.4	1.3	2.8	1.6	2.3	1.0	1.5	0.8	0.2	0.2	1.8
190	210	3.0	0.5	1.7	3.4	1.6	2.4	1.5	1.5	0.8	0.3		1.8
210	230	4.4	0.9	2.0	4.1	1.6	2.4	2.0	1.5	0.8	0.8		1.8
230	250	5.2	0.9	2.4	3.8	1.6	2.4	1.2	1.5	0.8	0.2	0.3	1.8
250	300	4.3	0.7	0.9	4.1	1.6	2.4	1.7	1.5	0.8	0.2	0.2	1.8
300	400	5.1	0.9	1.0	4.4	1.6	2.5	1.7	1.5	0.8	0.3	0.3	1.8
400	500	9.4	2.0	0.9	4.0	1.6	2.6	2.3	1.5	0.8	0.5	0.4	1.8
500	700	11	2.0	0.8	3.1	1.6	2.6	2.4	1.5	0.8	0.2	0.3	1.8
700	1000	24	4.0	0.6	4.3	1.6	2.6	2.8	1.5	0.8	0.2	0.4	1.8
1000	1500	50	7.6	0.4	3.1	1.7	2.5	3.3	1.5	0.8	0.3	0.4	1.8
iressed	i level												
100	nee	Stat.	Nbkg	CDY	Nhing	Reco.	Id.	Emergy	Unfolding	Trigger	20C	Theoretical	Loni
min	хлах		stat.	stat.				scaleáre	16	ne	dell:	ing	
Ce1	7	<-unco	rrelated	(8)->	£			-biz-to-b	in correla	ated (1)-			>
116	130	1.1	0.1	0.7	1.3	1.6	2.3	2.1	1.5	0.8	0.2	0.3	1.8
130	150	1.4	0.2	0.7	1.8	1.6	2.3	1.7	1.5	0.8	0.5	0.2	1.8
150	170	2.0	0.3	1.0	2.5	1.6	2.3	1.6	1.5	0.8	0.2	0.2	1.8
170	190	2.7	0.4	1.3	2.8	1.6	2.3	1.0	1.5	0.8	0.2	0.2	1.8
190	210	3.0	0.5	1.7	3.4	1.6	2.4	1.5	1.5	0.9	0.3	0.4	1.8
210	230	4.4	0.9	2.1	4.1	1.6	2.4	2.0	1.5	0.8	0.8	0.5	1.8

### Plotting data within HepData



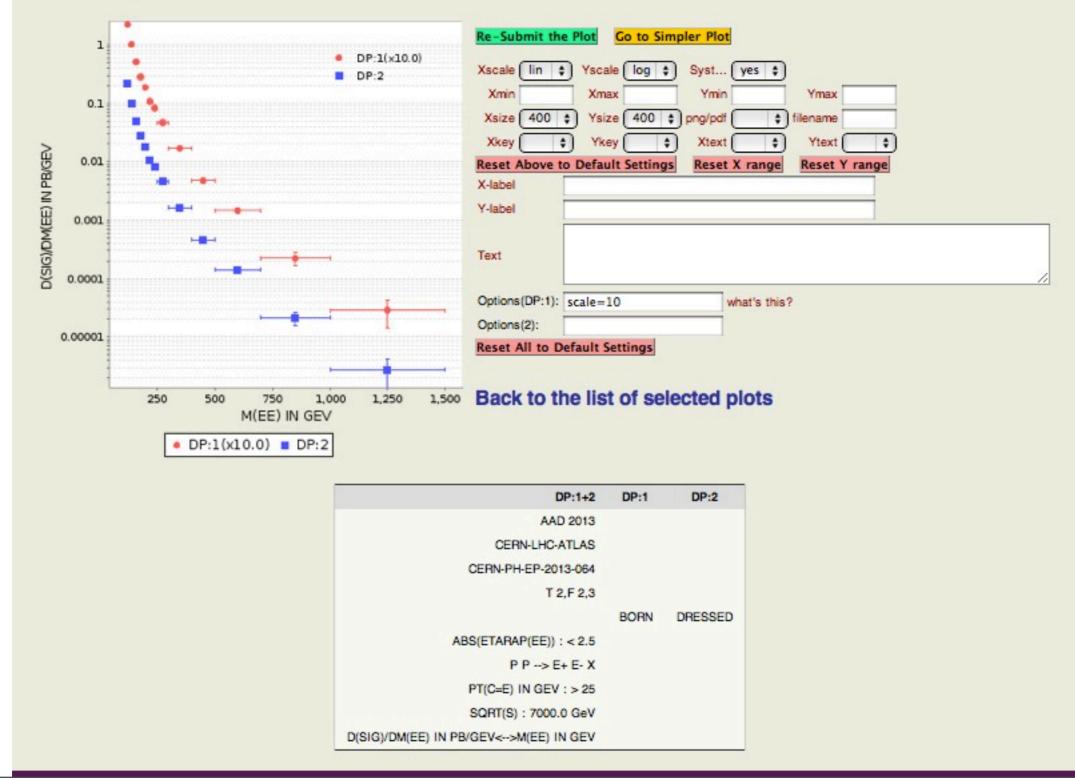
# Plotting multple datasets within HepData

(ETARAP(EE))	< 2.5								
PT(C=E)	> 25 GeV								
RE	P P> E+ E- X								
SQRT(S)	7000.0 GeV								
	BORN		DRESS	ED					
(EE) IN GEV			O(SIG)/DM(EE) IN PB/G	EV		0.0% (Sy	5)	0.00402 I 4.0% (8	siai) ± 0.0% (595)
116 130.	0.224 ± 1.1% (st				± 4.2% (sys)	5.9% (sy	s)	0.00165 ± 5.1% (s	stat) ± 5.9% (sys)
						± 6.3% (s	ys)	0.000458 ± 9.4%	(stat) ± 6.3% (sys)
130 150.	0.102 ± 1.4% (st	at) ± 4.3% (s	sys) 0.0984 ±	: 1.4% (stat	) ± 4.3% (sys)	) ± 5.7% (	sys)	0.000141 ± 11.0%	(stat) ± 5.7% (sys)
150 170.	0.0512 ± 2.0% (s	stat) ± 4.6% (	(sys) 0.0493 ±	2.0% (stat	) ± 4.6% (sys)				% (stat) ± 7.5% (sys)
					0.00000000				
			10	00 1500.		0.0% (stat) ± 9.8%	% (sys)	0.00000276 ± 50.0	0% (stat) ± 9.8% (sys
					Plot SelectPlot			Plot SelectPlot	
Multiple datas	ets from the sar	me, or diffe	ermalinks: I Reaction db I ast updated ted Datasets rent, HepData record	D=8470, RE 20131218 for C	ED=6090 145753) Omposite I combired on a si	ngle plot. This	page pres	sents and	
Multiple datas allows manipu a) from the 'pl b) from the 'pa Options are p pages to add	ets from the sar lation of these o ot' pages using ' aper' pages using rovided here to r more data sets.	me, or diffe datasets wh add this pl g the 'Selec manipulate t More detail	Reaction db I ast updated ted Datasets rent, HepData record hich have been selec lot to list' link	for Constant of the form of th	ED=6090 (45753) Omposite I combined on a si previous pages ei clearing the com	ngle plot. This ther:	eturning to	o the 'paper'	
Multiple datas allows manipu a) from the 'pl b) from the 'pa Options are p pages to add	ets from the sar lation of these of ot' pages using ' aper' pages using rovided here to r more data sets. ok of the plot are	me, or diffe datasets wh <b>add this pl</b> g the <b>'Selec</b> manipulate t More detail provided o	Reaction db l ast updated ted Datasets rent, HepData record nich have been selec lot to list' link tPlot' link the list by deleting an ls of individual data so on the next page.	for Constant of the form of th	ED=6090 (45753) Omposite I combined on a si previous pages ei clearing the com	ngle plot. This ther:	eturning to	o the 'paper'	
Multiple datas allows manipu a) from the 'pl b) from the 'pa Options are p pages to add	ets from the sar lation of these of ot' pages using aper' pages using rovided here to r more data sets. ok of the plot are	me, or diffe datasets wh add this pl g the 'Selec manipulate to More detail provided o	Reaction db I ast updated ted Datasets rent, HepData record nich have been select lot to list' link the list by deleting an ls of individual data so on the next page.	b=8470, RE 20131218 for Colls ted from p n item (x), sets can b	ED=6090 (45753) Omposite I combined on a si previous pages ei clearing the com be displayed by u	plete list, or re sing the 'summ	eturning to	o the 'paper'	
Multiple datas allows manipu a) from the 'pl b) from the 'pa Options are p pages to add modify the loo	ets from the sam lation of these of ot' pages using aper' pages using rovided here to r more data sets. ok of the plot are	me, or diffe datasets wh add this pl g the 'Selec manipulate t More detail provided o	Peaction db l ast updated ted Datasets rent, HepData record nich have been selec lot to list' link the list by deleting an ls of individual data s on the next page.	b=8470, RE 20131218 for Co s can be ted from p n item (x), sets can b	ED=6090 (45753) Omposite I combined on a si previous pages ei clearing the com be displayed by u remov	plete list, or re sing the 'summ	eturning to	o the 'paper'	
Multiple datas allows manipu a) from the 'pl b) from the 'pa Options are p pages to add modify the loo	ets from the sam lation of these of ot' pages using aper' pages using rovided here to r more data sets. ok of the plot are AAD 13 1 2 AAD 13 1	me, or diffe datasets wh add this pl g the 'Selec manipulate t More detail provided o	Peaction db l ast updated ted Datasets rent, HepData record nich have been selec lot to list' link the list by deleting an ls of individual data s on the next page.	b=8470, RE 20131218 for Co s can be ted from p n item (x), sets can b	ED=6090 (45753) Omposite I combined on a si previous pages ei clearing the com be displayed by u remov	plete list, or re sing the 'summ	eturning to	o the 'paper'	

## Plotting multiple datasets within HepData

### **Reaction Database Composite Plot**

This page displays the combined data plot with options to 'replot' the data changing various features of the plot. Mouseover keywords for more information.



# BICEP2 record in HepData

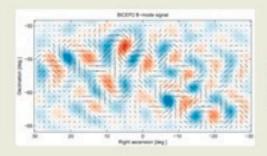
### **Reaction Database Full Record Display**

View short record or as: input, plain text, AIDA, PyROOT, YODA, ROOT, mpl, ScaVis or MarcXML

### ADE 2014 — BICEP2 I: Detection Of B-mode Polarization at Degree Angular Scales

Experiment: BICEP2 (BICEP2) Archived as: ARXIV:1403.3985 Auxiliary Material: http://bicepkeck.org/bicep2\_2014\_release/ Record in: INSPIRE Record in: CERN Document Server

BICEP2. Results are reported from the BICEP2 experiment, a Cosmic Microwave Background (CMB) polarimeter specifically designed to search for the signal of inflationary gravitational waves in the B-mode power spectrum around angular multipole ~ 80. BICEP2 observed from the South Pole for three seasons from 2010 to 2012. An excess of B-mode power over the base lensed-LambdaCDM expectation is found in the range 30 < multipole < 150, inconsistent with the null hypothesis at a significance of > 5 sigma. The observed B-mode power spectrum is well-fit by a lensed-LambdaCDM + tensor theoretical model with tensor/scalar ratio r = 0.20 + 0.07, -0.05, with r = 0 disfavoured at 7.0 sigma. Subtracting the best available estimate for foreground dust modifies the likelihood slightly so that r = 0 is disfavoured at 5.9 sigma.



The links immediately below give access to the original files provided by the BICEP2 experiment (see Auxiliary Material). The information in the first two links is also given further below in Table 1 (corresponding to Figure 2 of the paper) and Table 2 (corresponding to the middle panel of Figure 10 of the paper).

Tabulated likelihood for the tensor-to-scalar ratio, r, computed using the "direct likelihood" calculation	rlikelihood.txt
Bandpowers and statistical uncertainties	bandpowers.txt
Bandpower window functions for each multipole bin	allbins.tgz   bin1.txt   bin2.txt   bin3.txt   bin4.txt   bin5.txt   bin6.txt   bin7.txt   bin8.txt   bin9.txt

## How does data get into HepData ?

- Regular literature scans (arXiv) for new suitable data.
- Experimenters send us text files of numbers either at our request or unsolicited.
- We can also extract the numbers from the latex files on arXiv if the data is in tabular form in the paper.
- We process them into a HepData record.
- Finally we invite the authors to approve the record.

### **The Durham HepData Project**

REACTION DATABASE • DATA REVIEWS • PDF PLOTTER

#### Submitting Data to the Reaction Database

The compilers of the Reaction Database regularly scan the archives and literature to locate suitable new data sets for inclusion in the database, often contacting the experiments directly to obtain the exact numerical data.

If you have, or know of, data which you think should be in the database, then please email us about this.

The criteria for data to be included is that it be data of a publishable form and not preliminary data. Data from conferences often fall into the latter category. If it is possible that data will change following peer review, it may be better to delay HepData submission until after journal acceptance.

The format we accept data in is very wide and generally we require only a flat file containing the numerical values. Postscript and pdf figures are not suitable.

Ideally the format should be:

xlow xhigh y +stat -stat +sys1 -sys1 +sys2 -sys2 ....

where: xlow and xhigh are the xbin edges y is the measured quantity +stat and -stat are the positive and negative statistical errors (could also be +-stat if equal) +sysn and -sysn are any number of positive and negative systematic errors (again could be +-sysn if equal)

You can also include other thinge like xtocus (mean of the distribution within the bin); just specify the column definitions when submitting the data.

Any text files you send us will be processed into a standard format, which can be seen from the "input" link on any of the existing records. Data submissions provided in text formats closer to the "input" format can therefore be entered more quickly into the database. You can start by looking at the "input" file for an existing record that is similar to your submitted data. An annotated sample input file is available, as is the historical 1970s encoding manual (14 MB) defining some nomenclature that is still used in practice. It is useful to provide key metadata for each table such as a caption, the process being considered (initial-state and final-state particles), important kinematic cuts, and headers for the x and y columns specifying the quantity with units. Multiple x and y values can be given in the same row separated by semicolons and non-existent entries can be indicated by dashes "-". Numerical values should be given with an appropriate and not excessive number of significant figures (in general, a maximum of 4), avoiding giving more decimal places for the errors than the central values. An online form is available to test data input. Click on "Browse..." to select your input file, followed by "Upload" and "Process", then click "Display" for the first 10 tables, or "All" for the whole record. Please note that this online form is only for testing purposes, and the final submittal to HepData must still be done by email.

Contact us at: hepdata@projects.hepforge.org

HepData is funded by the UK STFC and hosted at the Durham IPPP. HepData also maintains the UK mirror of the PDG.



Science & Technology Facilities Council

University

ABOUT HEPDATA · SUBMITTING DATA

Wednesday, 2 April 2014

### How should data get into HepData ?

Given the volume of data already from the LHC, and the expected increase from Run 2, we feel there needs to be an increase of input into HepData 'directly' from the experiments.

We have an input web form, recently linked from the main page, which can already be used by anyone to test data input.

Main problem for outsiders is to know the 'meta-data' formats and the 'input number' formats.

Very recently this has been used by PHENIX to upload data to HepData - and we appreciate their comments on how this can be improved.

We intend to develop a more 'bare-bones' and 'flexible' approach in which the user directly uploads, to the DB, tables of data, or even Root files, needing only to put a minimum of extra information, mainly bibliographic and textual comments.

#### REACTION DATABASE . DATA REVIEWS . PDF PLOTTER

#### ABOUT HEPDATA · SUBMITTING DATA

DATA INPUT FORM	000	File Upload		
		0	٩	
ease specifiy a file to upload	* DEVICES	Name highmassDY_HepData.txt_orig	Date Modified A 13/06/2013	
Browse) No the selected. Upload	El iDisk	highmassDY_HepData.txt	13/06/2013	
(proad)	Skype 🚔	5090.dat read.pl_orig	13/06/2013 14/06/2013	
ploading the file 6090.input	* SHARED	5090.input	16/11/2013	
Process the Input Data - then -[Display][All] the Final Record (getstarted)		ead.pl	16/11/2013	
4	Benjamin Thom	sys.html	16/11/2013	
eference: CERN-PH-EP-2013-064 : 2013	Courtenay's Ma			
eference: ARXIV:1305.4192 : 2013	CSAIMac			
tatus: Encoded 31 MAY 2013 by MRW	David Carty's C			a la compañía de la c
title: Measurement of the high-mass Drell-Yan differential cross-sec				ctor 1 is measured with th
comment: CERN-LHC. Based on an integrated luminosity of 4.9 /fb, the TLAS detector as a function of the invariant mass, Mee, in the range				the electron and the
sitron have transverse momentum pT > 25 GeV and absolute pseudorapi	V PLACES			AS Luminosity
easurement (Eur.Phys.J. C73 (2013) 2518) the measured cross-sections accrtainty stays at the previous value of 1.8% durhamId: 6090	Hide extension	Car	open	The luminosity
nspireId: 1234228			10	3
letector: ATLAS				
experiment: CERN-LHC-ATLAS	Policition Databa	E · DATA REVIEWS · PDF PLOT	<b>ED</b>	ABOUT HEPDATA . SUBMITTING D
dataset:	HER TON DATABA			ABOUT NEP DAIX CODMITTING D
location: T 2,F 2,3 dscomment: Measured differential cross sections as a function of the	di-electro Reaction Date	hase Full Record Displ	av	
DY production at the Born and dressed levels			-	
ceackey: P P> E+ E- X	View short record or as:	nput, pain text, AIDA, PyROOT, Y	ODA, ROOT, mpl, ScaVis o	r MarcXML
bbskey: DSIG/DM	AAD 2013 -	Moscurement of the hi	ah-mass Droll-Van	differential cross-sectio
		ins at sqrt(s)=7 TeV with		
	Experiment: CERN-	LHC-ATLAS (ATLAS)		
	Preprinted as CERI Archived as: ARXIV	-PH-EP-2013-064		
		1303.4132		
I - browse for the prepared input file	Record in: INSPIRE			
	Hecoro In: INSPIRE			
	CERN-LHC. Based	on an integrated luminosity of 4.9 /b,		
	CERN-LHC. Based measured with the / fiducial region in wh	TLAS detector as a function of the in ch both the electron and the positron	variant mass, Mee, in the ran	ge 116 < Mee < 1500 GeV, for a
	CERN-LHC. Based measured with the /	TLAS detector as a function of the in ch both the electron and the positron	variant mass, Mee, in the ran	ge 116 < Mee < 1500 GeV, for a
2 - upload the file	CERN-LHC. Based measured with the / fiducial region in wh pseudorapidity letal	TLAS detector as a function of the in ch both the electron and the positron	variant mass, Mee, in the ran	ge 116 < Mee < 1500 GeV, for a
	CERN-LHC. Based measured with the / fiducial region in wh pseudorapidity letal	TLAS detector as a function of the in ch both the electron and the positron < 2.5	variant mass, Mee, in the ran	ge 116 < Mee < 1500 GeV, for a
	CERN-LHC. Based measured with the J fiducial region in wh pseudorapidity letal Link to the full bre Table 1 ( T 2,F 2,3	TLAS detector as a function of the in ch both the electron and the positron < 2.5 akdown of the sources of errors	variant mass, Mee, in the ran have transverse momentum t, AIDA, PyROOT, YODA,	ge 116 < Mee < 1500 GeV, for a pT > 25 GeV and absolute ROOT, mpl, ScaVis or MarcXML
2 - upload the file	CERN-LHC. Based measured with the J fiducial region in wh pseudorapidity letal Link to the full bre Table 1 ( T 2,F 2,3	TLAS detector as a function of the in ch both the electron and the positron < 2.5 akdown of the sources of errors	variant mass, Mee, in the ran have transverse momentum t, AIDA, PyROOT, YODA,	ge 116 < Mee < 1500 GeV, for a pT > 25 GeV and absolute
	CERN-LHC. Based measured with the <i>J</i> fiducial region in wh pseudorapidity letal Link to the full bre Table 1 ( T 2,F 2,3 Measured different	TLAS detector as a function of the in ch both the electron and the positron < 2.5 akdown of the sources of errors ) EXECUTE: or as: input, plain tex ial cross sections as a function of the	variant mass, Mee, in the ran have transverse momentum t, AIDA, PyROOT, YODA,	ge 116 < Mee < 1500 GeV, for a pT > 25 GeV and absolute ROOT, mpl, ScaVis or MarcXML
2 - upload the file	CERN-LHC. Based measured with the / fiducial region in wh pseudorapidity letal Link to the full bre Table 1 ( T 2,F 2,3 Measured different ABS(ETARAP(EE))	TLAS detector as a function of the im ch both the electron and the positron < 2.5 akdown of the sources of errors ) EXECUTION or as: input, plain tex al cross sections as a function of the < 2.5	variant mass, Mee, in the ran have transverse momentum t, AIDA, PyROOT, YODA,	ge 116 < Mee < 1500 GeV, for a pT > 25 GeV and absolute ROOT, mpl, ScaVis or MarcXML
2 - upload the file	CERN-LHC. Based measured with the <i>J</i> fiducial region in wh pseudorapidity letal Link to the full bre Table 1 ( T 2,F 2,3 Measured different	TLAS detector as a function of the im ch both the electron and the positron < 2.5 akdown of the sources of errors ) EXECUTION or as: input, plain tex al cross sections as a function of the < 2.5 > 25 GeV	variant mass, Mee, in the ran have transverse momentum t, AIDA, PyROOT, YODA,	ge 116 < Mee < 1500 GeV, for a pT > 25 GeV and absolute ROOT, mpl, ScaVis or MarcXML
<ul> <li>2 - upload the file</li> <li>3 - process the file</li> </ul>	CERN-LHC. Based measured with the <i>J</i> fiducial region in wh pseudorapidity letal Link to the full bre Table 1 ( T 2,F 2,3 Measured different ABS(ETARAP(EE)) PT(C=E)	TLAS detector as a function of the im ch both the electron and the positron < 2.5 akdown of the sources of errors ) EDECEDENT or as: input, plain tex al cross sections as a function of the < 2.5 > 25 GeV	variant mass, Mee, in the ran have transverse momentum t, AIDA, PyROOT, YODA,	ge 116 < Mee < 1500 GeV, for a pT > 25 GeV and absolute ROOT, mpl, ScaVis or MarcXML
2 - upload the file	CERN-LHC. Based measured with the <i>i</i> fiducial region in wh pseudorapidity letal Link to the full bre Table 1 ( T 2,F 2,3 Measured different ABS(ETARAP(EE)) PT(C=E) RE	TLAS detector as a function of the in ch both the electron and the positron < 2.5 akdown of the sources of errors ) ECCENTRIE or as: input, plain tex al cross sections as a function of the <2.5 > 25 GeV P P -→ E+ E- X	variant mass, Mee, in the ran have transverse momentum t, AIDA, PyROOT, YODA,	ge 116 < Mee < 1500 GeV, for a pT > 25 GeV and absolute ROOT, mpl, ScaVis or MarcXML
<ul> <li>2 - upload the file</li> <li>3 - process the file</li> </ul>	CERN-LHC. Based measured with the <i>i</i> fiducial region in wh pseudorapidity letal Link to the full bre Table 1 ( T 2,F 2,3 Measured different ABS(ETARAP(EE)) PT(C=E) RE	TLAS detector as a function of the in ch both the electron and the positron < 2.5 akdown of the sources of errors ) ECCCONDUCTION or as: input, plain tex ial cross sections as a function of the < 2.5 > 25 GeV P P> E+ E- X 7000.0 GeV	variant mass, Mee, in the ran have transverse momentum t, AIDA, PyROOT, YODA, di-electron mass for DY prod	ge 116 < Mee < 1500 GeV, for a pT > 25 GeV and absolute ROOT, mpl, ScaVis or MarcXML
<ul> <li>2 - upload the file</li> <li>3 - process the file</li> </ul>	CERN-LHC. Based measured with the <i>J</i> fiducial region in wh pseudorapidity letal Link to the full bree Table 1 ( T 2,F 2,3 Measured different ABS(ETARAP(EE)) PT(C=E) RE SQRT(S)	TLAS detector as a function of the in ch both the electron and the positron < 2.5 akdown of the sources of errors ) EDECEMBER or as: input, plain tex al cross sections as a function of the <2.5 > 25 GeV P P> E+ E- X 7000.0 GeV BORN D(SIGYDM(EE)	variant mass, Mee, in the ran have transverse momentum t, AIDA, PyROOT, YODA, di-electron mass for DY prod	ge 116 < Mee < 1500 GeV, for a pT > 25 GeV and absolute ROOT, mpl, ScaVis or MarcXML
<ul> <li>2 - upload the file</li> <li>3 - process the file</li> </ul>	CERN-LHC. Based measured with the <i>J</i> fiducial region in wh pseudorapidity letal Link to the full bree Table 1 ( T 2,F 2,3 Measured different ABS(ETARAP(EE)) PT(C=E) RE SORT(S) M(EE) IN GEV	TLAS detector as a function of the in ch both the electron and the positron < 2.5 akdown of the sources of errors ) EDECEMBER or as: input, plain tex al cross sections as a function of the <2.5 > 25 GeV P P> E+ E- X 7000.0 GeV BORN D(SIGYDM(EE)	variant mass, Mee, in the ran have transverse momentum t, AIDA, PyROOT, YODA, di-electron mass for DY prod	ge 116 < Mee < 1500 GeV, for a pT > 25 GeV and absolute ROOT, mpl, ScaVis or MarcXML

Wednesday, 2 April 2014

### HepData-Inspire connection

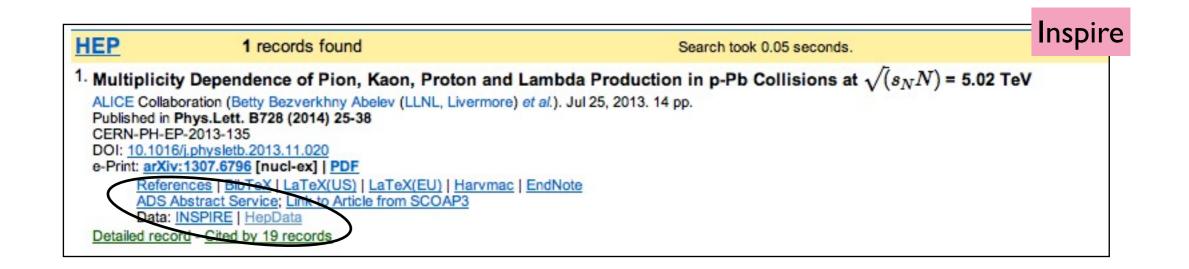
- HepData is linked to Inspire at various levels
- I. Reciprocal links on their respective web pages for common papers.
- 2. Inspire holds a copy of the HepData data records and displays via the HepData tag on their "detailed record" page.
- 3. HepData has the facility to use Inspire searching, for example a word from the title with "title:Higgs".

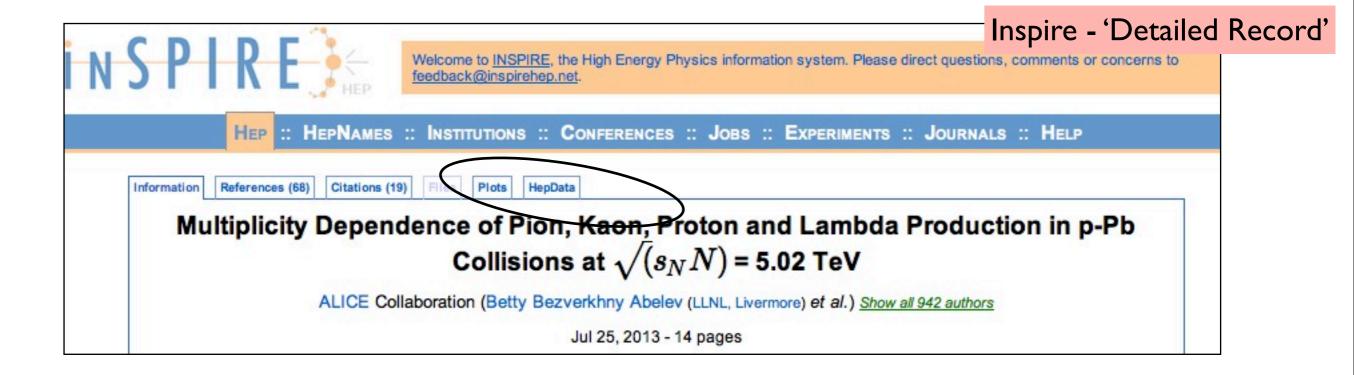
### HepData

#### 1. ABELEV 2013 – Experiment: CERN-LHC-ALICE (ALICE) Preprint: CERN-PH-EP-2013-135 Archive: ARXIV:1307.6796

Multiplicity Dependence of Pion, Kaon, Proton and Lambda Production in p-Pb Collisions at sqrt(s\_NN) = 5.02 TeV

Full data record I Short data record I INSPIRE





Information References Citations Files Plots HepData

Multiplicity Dependence of Pion, Kaon, Proton and Lambda Production in p-Pb Collisions at  $\sqrt{(s_N N)} = 5.02 \text{ TeV}$ - ALICE Collaboration (Abelev, Betty Bezverkhny *et al.*) Phys.Lett. B728 (2014) 25-38 arXiv:1307.6796 [nucl-ex] CERN-PH-EP-2013-135

THIS DATA COMES FROM DURHAM HEPDATA PROJECT

SUMMARY:

CERN-LHC. Measurements of the transverse momentum multiplicity dependences of I Proton-Lead collisions at a centre-of-mass energy/nucleon of 5.02 TeV. The dependen momentum on the charged particle multiplicity density dN/dETARAP is also presented,

#### DATASETS:



Description: pT-differential invariant yield of charged pions in pPb collisions with centre

Go to the reco	ord		
Plain			
		$p Pb \rightarrow (P$	$PI^+ + PI^-)X$
1		y <sub>CM</sub> =	0.0 - 0.5
1	class = 0.0 - 5.0 %	= 5.0 - 10.0 %	= 10.0 - 1
		$\sqrt{s}/nucleon$	= 5020.0 ~GeV
$p_T (GeV)$		$(1/Nev) \cdot (1/(2 \cdot PI^* p_T))$	$) \cdot d^2(N)/dp_T/d$
		†††Collapse	1114
0.1-	93.88 ±0.23(stat)	78.00 ±0.21(stat)	67.39 ±0.
0.12	±6.75(sys)	±5.61(sys)	±4.85(
	82.47 ±0.17(stat)	68.38 ±0.15(stat)	59.00 ±0.
0.12-	±3.73(sys)	±3.09(sys)	±2.67(
0.14			

Inspire uploads data on a regular basis from HepData based on a HepData 'timestamp' indicating the date of last change.

As yet does not upload any of the 'resource area' extra files.

Inspire attempts reformatting items of the text (eg reactions).

### Reaction Database Full Record Display

View short record or as: input, plain text, AIDA, PyROOT, YODA, ROOT, mpl, ScaVis or MarcXML

### ABELEV 2013 — Multiplicity Dependence of Pion, Kaon, Proton in p--Pb Collisions at sqrt(s\_NN) = 5.02 TeV

Experiment: CERN-LHC-ALICE (ALICE) Preprinted as CERN-PH-EP-2013-135 Archived as: ARXIV:1307.6796 Record in: INSPIRE



CERN-LHC. Measurements of the transverse momentum multiplicity dependences of Pion, Kaon, Proto production in Proton-Lead collisions at a centre-of-mass energy/nucleon of 5.02 TeV. The dependences mean transverse momentum on the charged particle multiplicity density dN/dETARAP is also presented

Total number of tables: 22. Displaying: 1 to 10. First | Previous | Next | Last | All

Table 1 (F 1.) Table

RE	P PB> (Pl+ + Pl-) X		
SQRT(S)/NUCLEON	5020.0 GeV		
YRAP(C=CM)	0.0-0.5		
class	0.0-5.0 pct	5.0-10.0 pct	10.0-20.0 pct
PT IN GEV		(1/Nev)*(1/(2*PI*PT))*D2	(N)/DPT/DYRAP IN GEV
0.1 - 0.12	93.88 ± 0.23 (stat) ± 6.75 (sys)	78.00 ± 0.21 (stat) ± 5.61 (sys)	67.39 ± 0.14 (stat) ± 4
0.12 - 0.14	82.47 ± 0.17 (stat) ± 3.73 (sys)	68.38 ± 0.15 (stat) ± 3.09 (sys)	59.00 ± 0.10 (stat) ± 2
0.14 - 0.16	72.69 ± 0.15 (stat) ± 3.29 (sys)	60.08 ± 0.13 (stat) ± 2.72 (sys)	51.80 ± 0.09 (stat) ± 2

# HepData-Elsevier connection

S	cienceDire	CT Journals   Books	Shopping cart   Sign in 🎛   Help	Durham		
	Download PDF	Export  More options  Search ScienceDirect	Advanced search	University		
		Physics Letters B	Recommended article	85 		
	ELSEVIER	Volume 728, 20 January 2014, Pages 562–578	► Citing articles (0)			
	Oceret for		▶ Related reference wo	rk articles		
		new phenomena in photon + jet events colle oton collisions at $\sqrt{s} = 8$ TeV with the ATLAS de		e		
	ATLAS Collabo	Show more	HepDa Reaction	ta data		
	http://dx.doi.org/1	10.1016/j.physletb.2013.12.029 0	Open Access			
	Abstract			The Durham HepData Pr	roject	Dur
	(γ+jet) events of-mass energy o	ribes a model-independent search for the production of new resonance using 20 fb <sup>-1</sup> of proton-proton LHC data recorded with the ATLAS de of $\sqrt{s} = 8$ TeV. The $\gamma$ +jet mass distribution is compared to a backgr	tector at a centre- ound model fit from	REACTION DATABASE · DATA REVIEWS · PDF PL		Univer Авоит НерДата • Submitting
	level on generic	nt deviation from the background-only hypothesis is found. Limits are se Gaussian-shaped signals and two benchmark phenomena beyond th tum black holes and excited quarks. Non-thermal quantum black holes	e Standard Model:	Reaction Database Full Record		
		✓ and excited quarks are excluded below masses of 3.5 TeV.		AAD 2013 — Search for new p proton collisions at sqrt(s) = 8	henomena in photon+jet event	
	1. Introduction			Experiment: CERN-LHC-ATLAS (ATLAS) Published in PL B728,562 (2014) (DOI:10.1016/j.j		
	states. They inclu quirks [7], [8] and past searches [1	oduction mechanisms have been proposed that produce massive photon ide non-thermal quantum black holes (QBHs) [1], [2] and [3], excited qua- [9], Regge excitations of string theory [10], [11] and [12], and topologic [4], [15], [16], [17] and [18], the only LHC search for this signature was ision data obtained at a centre-of-mass energy of $\sqrt{s} = 7$ TeV with the	urks [4], [5] and [6], al pions [13]. Of the done using proton-	Preprinted as CERN-PH-EP-2013-126 Archived as: ARXIV:1309.3230 Auxiliary Material: http://atlas.web.cern.ch/Atlas/G Record in: INSPIRE Record in: CERN Document Server		
	proton (DD 1 Coll	sion data obtained at a centre-of-mass energy of $\sqrt{s} = 1$ iev with the	All AS detector it	of proton-proton LHC data recorded with the ATL photon + jet mass distribution is compared to a ba only hypothesis is found. Limits are set at 95% cm phenomena beyond the Standard Model: non-ther	ackground model fit from data; no significant deviati	f sqrt(s) = 8 TeV. The tion from the background- and two benchmark n-thermal quantum black
					Nain text, AIDA, PyROOT, YODA, ROOT, mpl, passing the final selections. The number of observ imates are rounded to the nearest integer.	

GAMMA JET) IN GEV	Events			
	Observed Data	Fit Estimate		
SORT(S)	8000.0 GeV			
RE	P P -> GAMMA JET X			

14

# Data - what is a dataset? the view from HepData

- Generally data is published in a paper as a result of a specific analysis based on an input dataset.
- Each paper could contain one or several datasets.
- Datasets in papers could be single tables or several tables (or plots).
- Datasets could be spread over several papers.
- Datasets could extend beyond papers (ie extra data)
- HepData codes a single paper as one record.
- HepData's table structure is not necessarily the same as that in the paper.

# Datasets - how to assign DOIs A few thoughts from HepData

- Given the description of datasets on the previous slide it is not obvious at what level DOI's should be assigned in HepData records
- A single DOI per record?
- A single DOI per table?
- Possibly multiple DOIs in a record but not simply following the table structure, more aligned to the real 'data' structure.
- Could DOIs extend over several HepData records?
- Apart from assigning DOIs, other things to consider are: versioning, removal of tables/records

### Some statistics

• Total Records/Papers: 7893

# • LHC 261/662

- \* ATLAS 122 / 263
- \* CMS 59 / 270
- \* ALICE 61 / 76
- \* LHCb 15/44

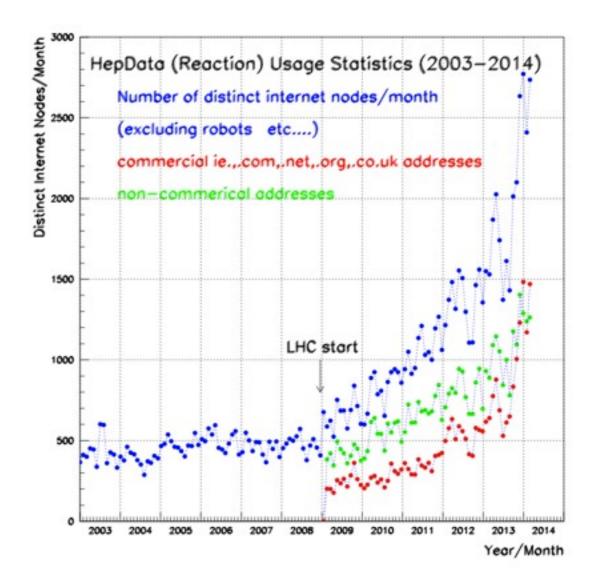
Although apparently only ~30% of LHC data has been entered into HepData this represents a very large part of the 'classic' HepData type data (ie SM type). Searches and more exotic things are less well represented.

# HepData Usage

As a metric to measure the use of HepData we have continued to count the number of distinct internet nodes, excluding robots etc., accessing per month.

This remained steady from ~500 up to 2009, the start of the LHC, then has increased five-fold to the present day.

We show in RED the accesses from 'commercial' sites (ie those ending .com, .net, .org and .co.uk) and in GREEN the rest.



# Summary & Conclusions

### Status

- HepData has continued to add new records, mainly LHC data.
- New DBM, Graeme Watt, in place to replace MW on retirement
- Improvements to web pages, plotting, user documentation are on-going
- We see an apparent continued growth in usage

### Future (for discussion)

- Need to engage experiments more in adding data directly themselves.
- Need to incorporate more data 'types' (eg 3D histograms, corr. matrices) into the Database model.
- Need to resolve the problem of assigning DOIs to data.
- Need to upgrade the Inspire-HepData connection.