PDF4LHC2021 Benchmarking

Thomas Cridge

University College London

12th July 2021

On behalf of PDF4LHC21 Combination Group

ISMD21 Meeting

*Longer recorded talk available for those interested!

Introduction - PDF Landscape

- PDF4LHC15 was a 1 year benchmarking exercise of the CT14, MMHT14, NNPDF3.0 PDFs which resulted in a combination set.
- It has now been more than 5 years since the PDF4LHC15 benchmarking exercise.
- Increasing amounts of data coming out of the LHC, greater precision, more channels, more differential ⇒ changes in PDFs.
- Many theoretical improvements ⇒ full NNLO predictions, methodological improvements (parameterisations, algorithms, etc).
- PDFs now known more accurately and precisely than ever before, but some differences emerging ⇒ benchmarking needed.
- We consider 3 global PDF fits most recent sets, which include much of the recent datasets: MSHT20, CT18, NNPDF3.1.

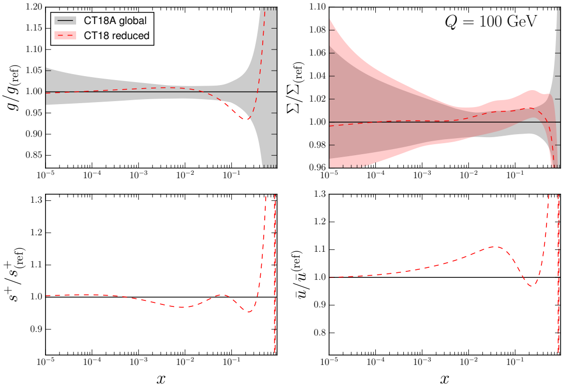
Work undertaken through many useful discussions, many thanks to all members involved.

PDF Benchmarking: Aim and Approach

- Desire to understand origin of differences:
  - Are they due to variations of experimental input, different theory settings, methodologies? Are these equally valid choices?
- Seek to remove as many differences in input/approach as possible:
  - Common input data - Small subset of datasets ⇒ reduced fits.
  - Common theory settings wherever possible.
  - Examine methodological differences in parallel as much as possible.
- Reduced fits offer *ease of comparison at expense of robustness*.
- To benchmark the reduced fits:
  - Compare PDFs directly to look for areas of difference.
  - Compare χ^2 to determine particular datasets showing differences.
  - Compare cross-sections and point-by-point theory predictions.
- Once differences in reduced fits understood, slowly add datasets moving towards global fits, focusing on key areas of differences.
- End result: PDF4LHC21 set of PDFs, central PDFs and Hessian error set (30-50 sufficient) representing the 3 published PDFs.

Reduced Fits: CT18 reduced fit vs CT18A global fit

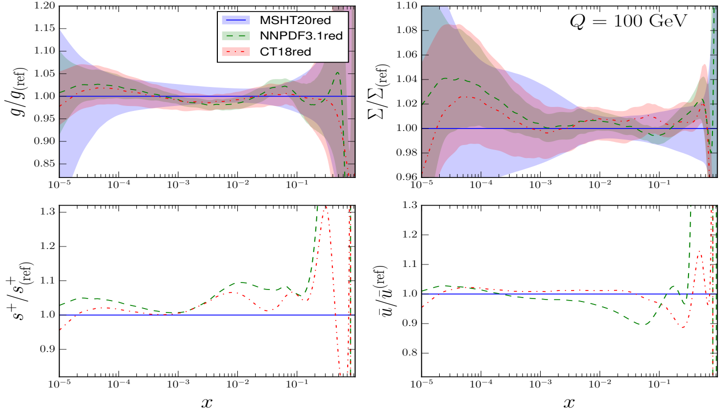
- Current Status:



- Good compatibility with change in high x gluon shape and some increase in $\bar{u}$. Some changes in flavour decomposition.
- Some increase in *nominal* PDF uncertainties, particularly at low x.

Reduced Fits PDF Comparison - central values

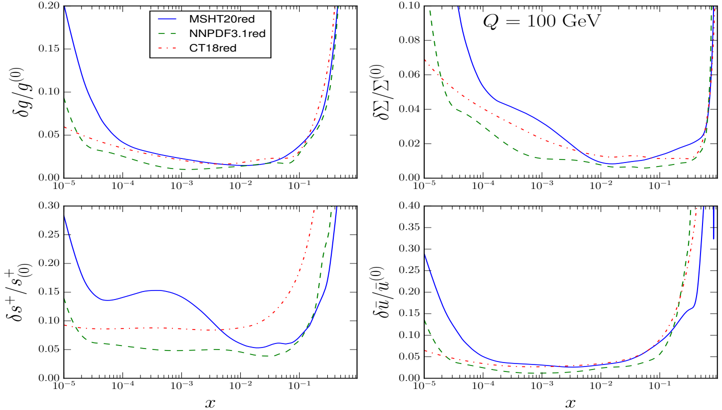
- Current Status:



- Good general agreement within uncertainties, perhaps with the exception of high x flavour decomposition of NNPDF.
- Nonetheless, strangeness and flavour decomposition improved through benchmarking (NuTeV - later). *Note this is without the $t\bar{t}$ added.

Reduced Fits PDF Comparison - uncertainties

- Current Status:



- Similar size uncertainties in data regions, MSHT generally larger errors where constraints lacking in reduced fit.
- Parallel study into differences in uncertainty bands ongoing.

*Note this is without the $t\bar{t}$ added.

Reduced Fits Datasets χ^2 Comparison

- Current status:

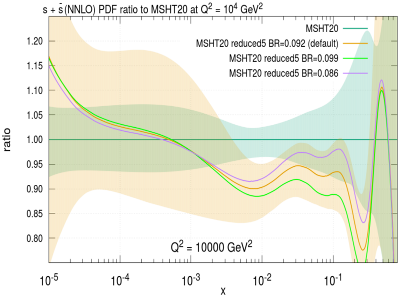
ID	Expt.	N_{pt}	χ^2/N_{pt} (CT)	χ^2/N_{pt} (MSHT)	χ^2/N_{pt} (NNPDF)
101	BCDMS F_2^p	$329/163^{\dagger\dagger}/325^{\dagger}$	1.06	1.00	1.21
102	BCDMS F_2^d	$246/151^{\dagger\dagger}/244^{\dagger}$	1.06	0.88	1.10
104	NMC F_2^d/F_2^p	$118/117^{\dagger}$	0.93	0.93	0.90
124+125	NuTeV $\nu\mu\mu + \bar{\nu}\mu\mu$	38+33	0.79	0.83	1.22
160	HERAI+II	1120	1.23	1.20	1.22
203	E866 $\sigma_{pd}/(2\sigma_{pp})$	15	1.24	0.80	0.43
245+250	LHCb 7TeV& 8TeV W,Z	29+30	1.15	1.17	1.44
246	LHCb 8TeV $Z \to ee$	17	1.35	1.43	1.57
248	ATLAS 7TeV W,Z(2016)	34	1.96	1.79	2.33
260	D0 Z rapidity	28	0.56	0.58	0.62
267	CMS 7TeV eletron A_{ch}	11	1.47	1.52	0.76
269	ATLAS 7TeV W,Z(2011)	30	1.03	0.93	1.01
545	CMS 8TeV incl. jet	$185/174^{\dagger\dagger}$	1.03	1.39	1.30
Total	N_{pt}	—	2263	1991	2256
Total	χ^2/N_{pt}	—	1.14	1.15	1.20

- Similar overall quality of fit in χ^2/N.
- Differences remaining in some datasets:
  - NuTeV agreement improved but difference remains, seen in $s + \bar{s}$.
  - Some differences in NNPDF fit quality to small datasets, e.g. CMS 7 TeV electron asymmetry.

Table from T. Hobbs

Flavour Decomposition - Strangeness and NuTeV

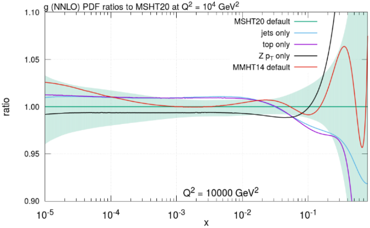
- One of the main differences between the first reduced sets was in the flavour decomposition and strangeness.
- NuTeV dimuon data key driver of this, requires BR(c hadrons $\to \mu$).
- BR($c \to \mu$) anti-correlated with total strangeness, 3 groups have different default values:
  - NNPDF 0.087 ± 0.005
  - MSHT 0.092 ± 0.01 variable.
  - CT 0.099, normalisation uncertainty.



- Choose same BR fixed at 0.092 $\Rightarrow$ better strangeness agreement, largely within uncertainties between all 3 groups.
- Also aids reduction in flavour decomposition differences.

High x gluon - jets and top

- High x gluon of interest to both reduced and global fits.
- Jet, top, Zp_T data, different pulls:
- Not straightforward to fit:
  - Difficulties fitting all bins.
  - Possible tensions.
  - Issue of correlated systematics.
- MSHT, CT, NNPDF - differences in relative importance and fit quality.

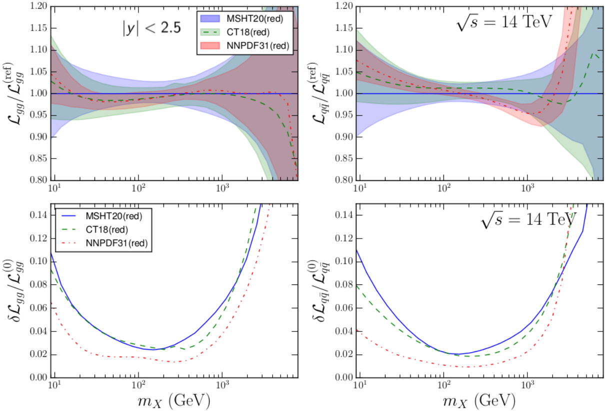


ATLAS 8 TeV multi-differential $t\bar{t}$ lepton+jets - $m_{tt}, y_t, y_{tt}, p_t^T$.

- Weights and tensions with other datasets notably affect fit quality, removing these differences ⇒ similar behaviour can be observed.

Dataset (N)	MSHT reduced (default CMS8j)	MSHT reduced (CMS8j, double weight $t\bar{t}$)	NNPDF reduced (default CMS8j)	MSHT reduced (CMS7j)	MSHT reduced (AT7j)	MSHT reduced (no jets)
χ^2/N	1.15	1.15	1.20	1.11	1.17	1.12
p_t^T (8)	3.8	4.2	7.2	4.0	4.6	4.5
y_t (5)	8.4	5.8	4.3	6.4	5.5	5.2
y_{tt} (5)	12.5	7.4	5.7	7.2	5.2	6.6
m_{tt} (7)	6.4	6.5	2.4	6.4	6.4	7.4
$t\bar{t}$ total	31.2	23.9	19.6	24.0	21.6	23.8

Reduced Fits: Current Status Summary*



- Very good agreement in gluon-gluon, quark-quark and quark-gluon luminosities. (Latter two in backup slides).
- Small difference in quark-antiquark luminosity, still some flavour decomposition differences, although within MSHT uncertainties.

*Note this is without the $t\bar{t}$ added.

Conclusions and Future Work

- New data, theoretical improvements, PDF methodological improvements have meant substantial changes since PDF4LHC15.
- We have been performing a benchmarking exercise of the 3 global fit PDF groups most recent sets: MSHT20, CT18, NNPDF3.1.
- Based on comparing "Reduced Fits" ⇒ very good consistency is now observed between the three groups, particularly in luminosities.
- Sources of differences in reduced fit strangeness largely identified.
- Currently analysing high x gluon region of interest, effects of dataset tensions and weights investigated.
- Overall very good progress towards benchmarking the global fits.
- End result: PDF4LHC21 set of PDFs, central PDFs and Hessian error set (30-50 sufficient) representing the 3 published PDFs.

Many thanks to all those involved in this work/discussions, special thanks to T. Hobbs, T.-J. Hou, L. Harland-Lang, P. Nadolsky, E. Nocera, J. Rojo, R. Thorne for providing tables/plots/fits.

Backup Slides

Introduction - New Datasets (MSHT20)

LHCb W, Z data at high rapidity

CMS W+c

Precision DY data

$\Rightarrow$ Flavour Decomposition

LHC Jet, Zp_T, $t\bar{t}$ data

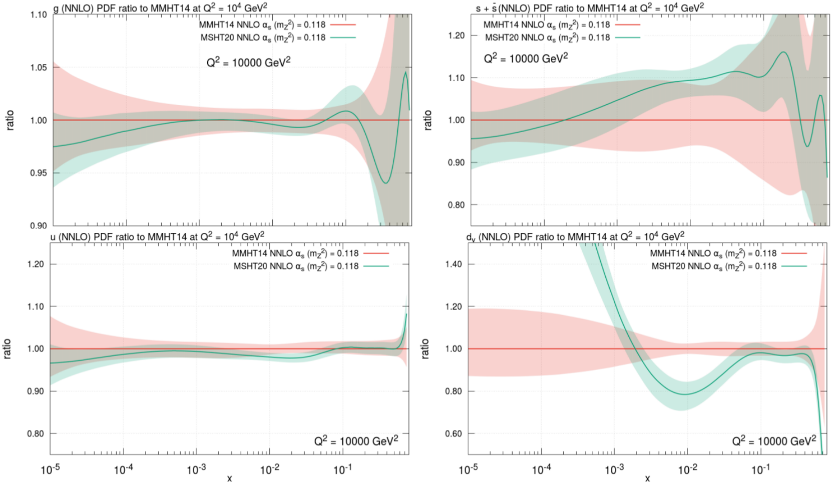
$\Rightarrow$ High x gluon

Data set	Points	NLO χ^2/N_{pts}	NNLO χ^2/N_{pts}
DØ W asymmetry	14	0.94 (2.53)	0.86 (14.7)
$\sigma_{t\bar{t}}$ [93]- [94]	17	1.34 (1.39)	0.85 (0.87)
LHCb 7+8 TeV $W+Z$ [95,96]	67	1.71 (2.35)	1.48 (1.55)
LHCb 8 TeV $Z \to ee$ [97]	17	2.29 (2.89)	1.54 (1.78)
CMS 8 TeV W [98]	22	1.05 (1.79)	0.58 (1.30)
CMS 7 TeV $W+c$ [99]	10	0.82 (0.85)	0.86 (0.84)
ATLAS 7 TeV jets $R = 0.6$ [18]	140	1.62 (1.59)	1.59 (1.68)
ATLAS 7 TeV $W+Z$ [20]	61	5.00 (7.62)	1.91 (5.58)
CMS 7 TeV jets $R = 0.7$ [100]	158	1.27 (1.32)	1.11 (1.17)
ATLAS 8 TeV Z p_T [75]	104	2.26 (2.31)	1.81 (1.59)
CMS 8 TeV jets $R = 0.7$ [101]	174	1.64 (1.73)	1.50 (1.59)
ATLAS 8 TeV $t\bar{t} \to l+j$ sd [102]	25	1.56 (1.50)	1.02 (1.15)
ATLAS 8 TeV $t\bar{t} \to l^+l^-$ sd [103]	5	0.94 (0.82)	0.68 (1.11)
ATLAS 8 TeV high-mass DY [73]	48	1.79 (1.99)	1.18 (1.26)
ATLAS 8 TeV W^+W^-+ jets [104]	30	1.13 (1.13)	0.60 (0.57)
CMS 8 TeV $(d\sigma_{t\bar{t}}/dp_{T,t}dy_t)/\sigma_{t\bar{t}}$ [105]	15	2.19 (2.20)	1.50 (1.48)
ATLAS 8 TeV W^+W^- [106]	22	3.85 (13.9)	2.61 (5.25)
CMS 2.76 TeV jets [107]	81	1.53 (1.59)	1.27 (1.39)
CMS 8 TeV $\sigma_{t\bar{t}}/dy_t$ [108]	9	1.43 (1.02)	1.47 (2.14)
ATLAS 8 TeV double differential Z [74]	59	2.67 (3.26)	1.45 (5.16)
Total, LHC data in MSHT20	1328	1.79 (2.18)	1.33 (1.77)
Total, non-LHC data in MSHT20	3035	1.13 (1.18)	1.10 (1.18)
Total, all data	4363	1.33 (1.48)	1.17 (1.36)

- Lots of new information constraining PDFs.

MSHT20, 2012.04684

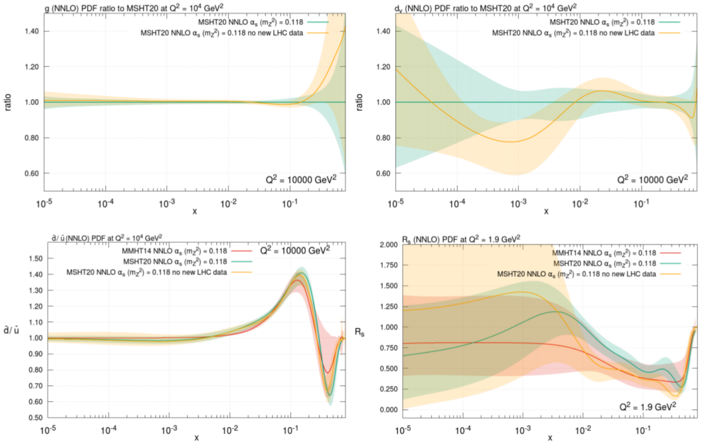
Introduction - Changes in PDFs: MSHT20



- Notable changes in strangeness (ATLAS W, Z data), down valence (new data and parameterisation), gluon (new jets, top, Zp_T data).

More details in R. Thorne's MSHT20 talk.

Effect of new LHC data in MSHT20

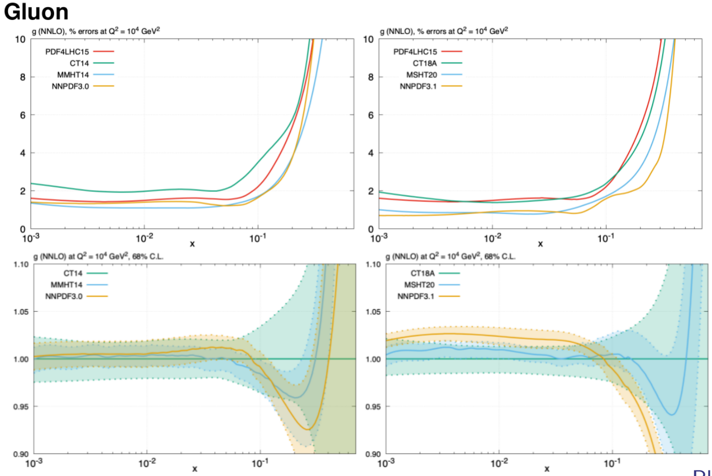


Main effect on details of flavour, i.e. d_V shape, increase in strange quark for $0.001 < x < 0.3$ and $\bar{d}, \bar{u}$ details, though also partially from parameterisation change. Decrease in high-x gluon.

*MSHT20 2012.04684.

Slide from R. Thorne

Introduction - Changes in PDFs



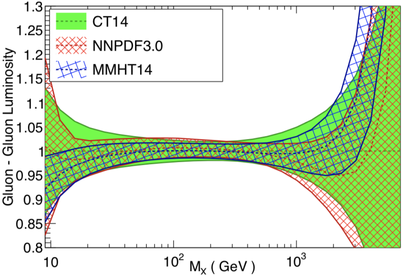
Plots from L. Harland-Lang

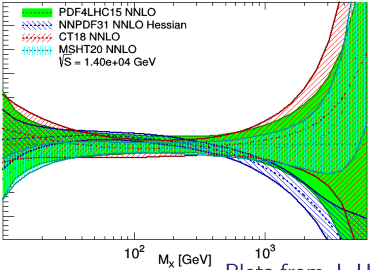
- Reduction in PDF uncertainties seen across all 3 groups.
- Central value agreement not as good, some differences emerging.

Note: CT18A shown for ease of comparison, however CT18 is the default set.

Introduction - Changes in PDFs

N.B. Different baseline for ratio in two plots and different colours.





Plots from J. Huston

- Central value spread effects gluon-gluon luminosity.
- If these were to be combined à la PDF4LHC15, there will be some contribution to uncertainty from spread as well as the uncertainties.
- Motivates understanding these differences and their origin $\Rightarrow$ PDF4LHC21 benchmarking.
- New PDFs CT18, MSHT20, NNPDF3.1 $\Rightarrow$ now is a good time to undertake a benchmarking exercise, ahead of new $\Rightarrow$ PDF4LHC21 combination - feedback on what is ultimately provided is welcome!

PDF Benchmarking: Datasets

- Chosen subset of datasets fit by all 3 groups in (almost) the same way, list is surprisingly small! Small reduced fit set.
- Take most conservative cuts applied by any group for consistency.
- Ensure enough datasets and a sufficient variety of dataset types are fit to have some (but incomplete) constraints on all PDF flavours.

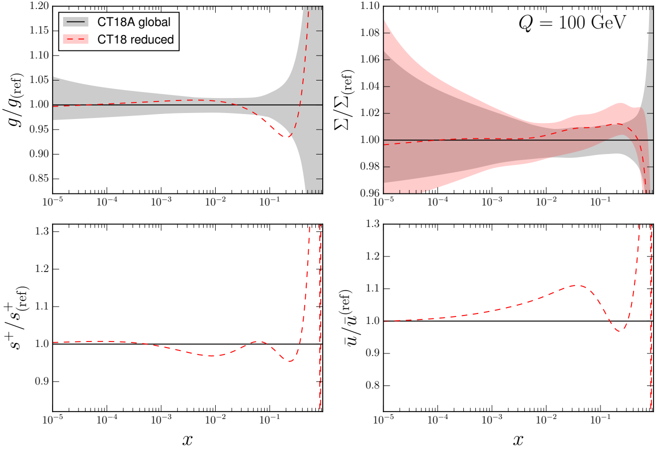
- Overall list:
  - NMC deuteron to proton ratio in DIS.
  - NuTeV dimuon cross-sections.
  - HERA I+II inclusive cross-sections from DIS.
  - E866 fixed target Drell-Yan ratio pd/pp data.
  - D0 Z rapidity distribution.
  - ATLAS W, Z 7 TeV rapidity distribution, only Z peak and central.
  - CMS 7 TeV W asymmetry.
  - CMS 8 TeV inclusive jet data.
  - LHCb 7, 8 TeV W, Z rapidity distributions.
  - BCDMS proton and deuteron DIS data.

PDF Benchmarking: Theory Settings

- Choose common theory settings for simplicity:
  - Same heavy quark masses ($m_c = 1.4$GeV, $m_b = 4.75$GeV) and $\alpha_S(M_Z^2) = 0.118$.
  - No strangeness asymmetry at input scale: $(s - \bar{s})(Q_0) = 0$.
  - Perturbative charm.
  - Positive definite quark distributions (lack of constraint may allow negative fluctuations).
  - No deuteron or nuclear corrections.
  - Fixed branching ratio for charm hadrons to muons.
  - NNLO corrections for dimuon data.
- Note: These are not the chosen settings for any one group, but rather are a compromise to the least common denominator in each case, *we would not recommend them for a full global fit*.

Reduced Fits: CT18 changes - central values

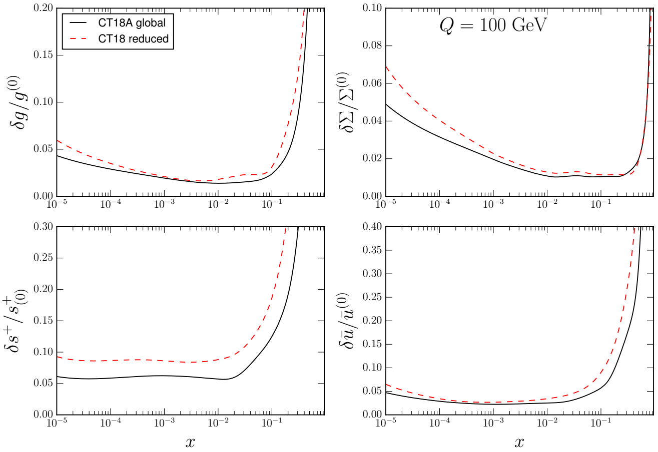
- Current Status:



- Good compatibility with change in high x gluon shape and some increase in $\bar{u}$. Some changes in flavour decomposition.

Reduced Fits: CT18 changes - uncertainties

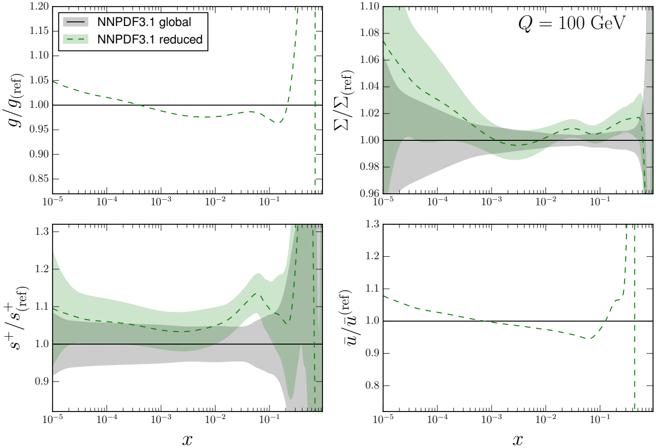
- Current Status:



- Some increase in *nominal* PDF uncertainties, particularly at low x.

Reduced Fits: NNPDF3.1 changes - central values

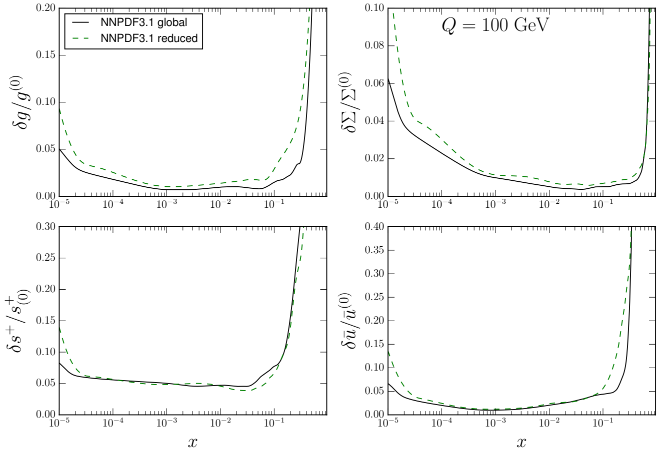
- Current Status:



- Good compatibility, changes in strangeness (see later) and change in large x gluon (removal of top data, addition of CMS 8 TeV jet).

Reduced Fits: NNPDF3.1 changes - uncertainties

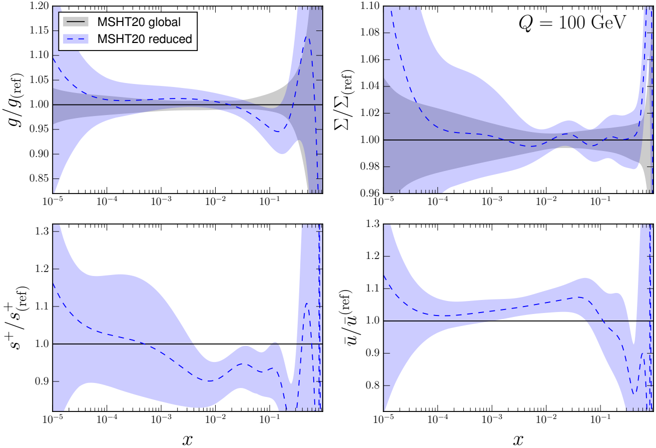
- Current Status:



- Generally slightly increased uncertainties, particularly for the gluon.

Reduced Fits: MSHT20 changes - central values

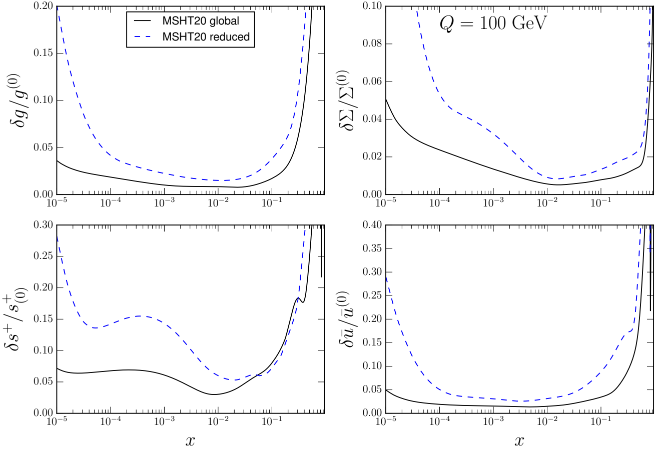
- Current Status:



- Good compatibility, changes in strangeness (removal of 8 TeV ATLAS W, Z data), flavour decomposition and large x gluon.

Reduced Fits: MSHT20 changes - uncertainties

- Current Status:



- General marked increase in uncertainties of reduced fit, particularly outside of regions where there are data.

PDF4LHC15 in Predictions Datasets χ^2 Comparison

- First make predictions with PDF4LHC15 PDFs, identifies any differences in theory/data between groups with fixed PDFs.
- Current status:

Table from T. Hobbs

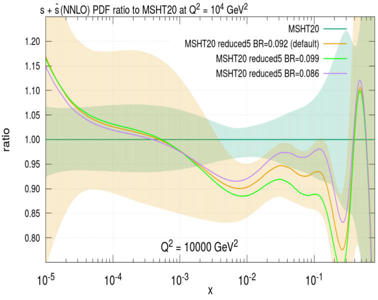
ID	Expt.	N_{pt}	χ^2/N_{pt} (CT)	χ^2/N_{pt} (MSHT)	χ^2/N_{pt} (NNPDF)
101	BCDMS F_2^p	$329/163^{\dagger\dagger}/325^{\dagger}$	1.35	1.2	1.51
102	BCDMS F_2^d	$246/151^{\dagger\dagger}/244^{\dagger}$	0.97	1.27	1.24
104	NMC F_2^d/F_2^p	$118/117^{\dagger}$	0.92	0.93	0.94
124+125	NuTeV $\nu\mu\mu + \bar{\nu}\mu\mu$	38+33	0.75	0.73	0.84
160	HERAI+II	1120	1.27	1.24	1.74
203	E866 $\sigma_{pd}/(2\sigma_{pp})$	15	0.45	0.54	0.59
245+250	LHCb 7TeV& 8TeV W,Z	29+30	1.5	1.34	1.76
246	LHCb 8TeV $Z \to ee$	17	1.35	1.65	1.25
248	ATLAS 7TeV W,Z(2016)	34	6.71	7.46	6.51
260	D0 Z rapidity	28	0.61	0.58	0.61
267	CMS 7TeV eletron A_{ch}	11	0.45	0.5	0.73
269	ATLAS 7TeV W,Z(2011)	30	1.21	1.23	1.31
545	CMS 8TeV incl. jet	$185/174^{\dagger\dagger}$	1.53	1.89	1.78
Total	N_{pt}	—	2263	1991	2256
Total	χ^2/N_{pt}	—	1.31	1.36	1.62

- Similar overall quality of fit for MSHT and CT in χ^2/N, NNPDF significantly larger χ^2/N.
- Differences in some datasets:
  - Difference in NNPDF HERA χ^2 - flavour scheme, disappears in fit.

Flavour Decomposition - Strangeness and NuTeV

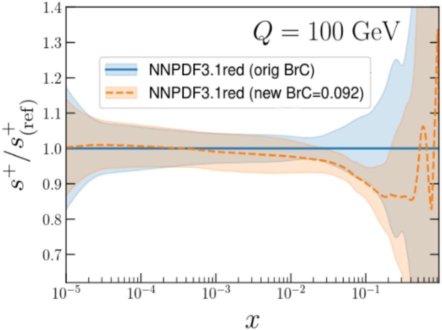
- One of the main differences between the first reduced sets was in the flavour decomposition and strangeness.
- NuTeV dimuon data key driver of this, complicated dataset:
  - Requires knowledge of charm $\rightarrow$ hadrons branching ratio (BR).
  - Non-isoscalar nature of target.
  - Prefers non-zero strangeness asymmetry.
  - Acceptance corrections required.
- BR($c \rightarrow \mu$) anti-correlated with total strangeness, 3 groups have different default values:
  - NNPDF 0.087 ± 0.005
  - MSHT 0.092 ± 0.01 variable.
  - CT 0.099, normalisation uncertainty.
- MSHT20 reduced fit χ^2/N:

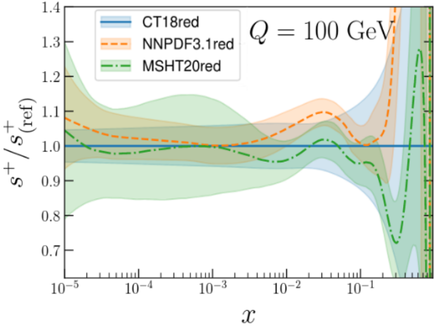
Dataset /BR	0.086	0.092	0.099
NuTeV Dimuon	58.8/71	49.6/71	68.5/71
ATLAS 7 TeV W, Z	60.8/34	65.1/34	57.1/34



Flavour Decomposition - Strangeness and NuTeV

- Setting all variables the same in all 3 fits - same Dimuon BR fixed at 0.092, all treat non-isoscalarity, same acceptance corrections.
- NNPDF strangeness reduced as expected, CT strangeness increases.



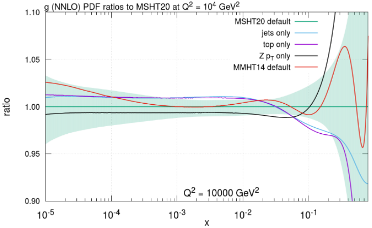


- Better strangeness agreement, certainly in data region, now largely within uncertainties between all 3 groups.
- Also aids reduction in flavour decomposition differences.

Plots from J. Rojo

High x gluon - jets and top

- High x gluon of interest to both reduced and global fits.
- 3 main dataset types - jet data, top data, Zp_T data, different pulls:
- Not straightforward to fit:
  - Difficulties fitting all bins.
  - Possible tensions.
  - Issue of correlated systematics.



- MSHT, CT, NNPDF observe differences in the relative importance of these datasets and the quality of their individual fits

ATLAS 8 TeV multi-differential $t\bar{t}$ lepton+jets - $m_{tt}, y_t, y_{tt}, p_t^T$.

- Several groups have had difficulties fitting this data, either together (MSHT, CT) or y_t or y_{tt} individually (MSHT, CT, ATLAS).
- NNPDF3.0 however able to fit all 4 distributions well individually.

ATLAS 8 TeV multi-differential $t\bar{t}$ lepton+jets

- Start by adding this to the reduced fit, first check theory predictions for PDF4LHC15 read in (no fitting):
  - Data agree and theory agrees to better than 1%.
  - All groups χ^2 in agreement and follow same pattern:

Distribution/N	MSHT	CT	NNPDF
$p_t^T/8$	3.0	3.1	3.4
$y_t/5$	10.6	10.1	9.5
$y_{tt}/5$	17.6	15.3	16.2
$m_{tt}/7$	4.3	4.2	4.1

  - Differences in global fits likely not from $t\bar{t}$ theory implementations.
- Weights and tensions with other datasets notably affect fit quality, removing these differences $\Rightarrow$ similar behaviour can be observed.

Dataset (N)	MSHT reduced (default CMS8j)	MSHT reduced (CMS8j, double weight $t\bar{t}$)	NNPDF reduced (default CMS8j)	MSHT reduced (CMS7j)	MSHT reduced (AT7j)	MSHT reduced (no jets)
χ^2/N	1.15	1.15	1.20	1.11	1.17	1.12
p_t^T (8)	3.8	4.2	7.2	4.0	4.6	4.5
y_t (5)	8.4	5.8	4.3	6.4	5.5	5.2
y_{tt} (5)	12.5	7.4	5.7	7.2	5.2	6.6
m_{tt} (7)	6.4	6.5	2.4	6.4	6.4	7.4
$t\bar{t}$ total	31.2	23.9	19.6	24.0	21.6	23.8

Benchmarking ATLAS 8 TeV $t\bar{t}$ lepton+jets

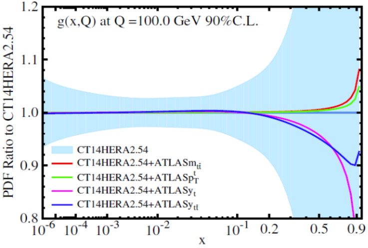
- *What happens when this dataset is added to the reduced fits?*
- Two cases considered - "uncorrelated" (all systematic and statistical correlations between distributions turned off) and "correlated" (including all correlations, produces a very poor fit):

Distribution/N	$p_t^T/8$	$y_t/5$	$y_{tt}/5$	$m_{tt}/7$	Total
MSHT uncorrelated	3.8	8.4	12.5	6.4	31.2
NNPDF uncorrelated	7.2	3.9	5.1	2.5	18.7
CT uncorrelated	3.4	12.9	17.3	6.1	39.7
MSHT correlated	-	-	-	-	130.6
NNPDF correlated	-	-	-	-	122.7
MSHT decorrelated	-	-	-	-	35.3

- MSHT observe usual pattern as in global fits, p_t^T and m_{tt} can be fit but y_t, y_{tt} struggle, although better than in full fit. Awful fit if all correlations included, can fit with parton shower decorrelation.
- CT see usual global fit pattern also, poor fits to rapidities y_t, y_{tt}.
- NNPDF however able to fit rapidity distributions in uncorrelated case, yet correlated case similar to MSHT.

Benchmarking ATLAS 8 TeV $t\bar{t}$ lepton+jets Preliminary!

- Potential explanation division of training and validation in NNPDF.
- Training fraction usually 50%, for small datasets this is unfeasible - all data in training.
- Potentially double-weights small datasets - e.g. ATLAS $t\bar{t}$.
- Affects balance of p_t^T, m_{tt} and y_t, y_{tt}, which have some tension.



Dataset	MSHT uncorrelated	NNPDF uncorrelated	MSHT uncorrelated double weight
Total	2314.1	2731.4	2313.3
χ^2/N	1.15	1.20	1.15
DYratio (15)	9.5	5.2	9.2
CMS W asym. (11)	14.2	8.2	10.2
p_t^T (8)	3.8	7.2	4.2
y_t (5)	8.4	4.3	5.8
y_{tt} (5)	12.5	5.7	7.4
m_{tt} (7)	6.4	2.4	6.5
$t\bar{t}$ total	31.2	19.6	23.9

- May also explain NNPDF better fit of E866 DYratio data and CMS W charge asymmetry data (15 and 11 points respectively):

Benchmarking ATLAS 8 TeV $t\bar{t}$ lepton+jets

Preliminary!

- Additional explanations are other datasets included - tensions?
- NNPDF-3.0 had little jet data. NNPDF-4.0 will have much more, it sees similar issues as MSHT, CT, ATLAS for this dataset.
- Useful to consider different jet datasets as well as CMS 8 TeV jets*:

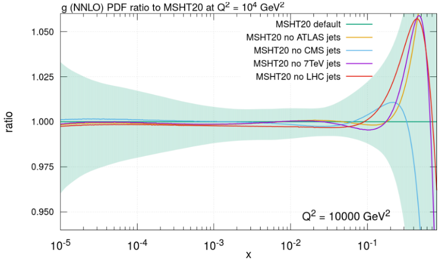
Dataset (N)	MSHT reduced (default CMS8j)	MSHT reduced + CMS7j	MSHT reduced + AT7j	MSHT reduced (CMS7j only)	MSHT reduced (AT7j only)	MSHT reduced (no jets)
χ^2/N	1.15	1.15	1.18	1.11	1.17	1.12
CMS 8 TeV jets (174)	243.6	247.2	249.9	-	-	-
CMS 7 TeV jets (158)	-	163.5	-	156.4	-	-
ATLAS 7 TeV jets (140)	-	-	225.7	-	210.4	-
p_t^T (8)	3.8	4.0	4.3	4.0	4.6	4.5
y_t (5)	8.4	7.3	7.3	6.4	5.5	5.2
y_{tt} (5)	12.5	9.8	10.2	7.2	5.2	6.6
m_{tt} (7)	6.4	6.4	7.0	6.4	6.4	7.4
$t\bar{t}$ total	31.2	27.5	28.8	24.0	21.6	23.8

- Tensions between CMS 8 TeV jets and ATLAS, CMS 7 TeV jets.
- Similar tensions with ATLAS 8 TeV $t\bar{t}$, specifically the rapidity distributions, which favour lower gluon.

*Note "uncorr" case shown, systematic correlations not included, same pattern observed in "corr" case.

High x gluon - Jet tensions

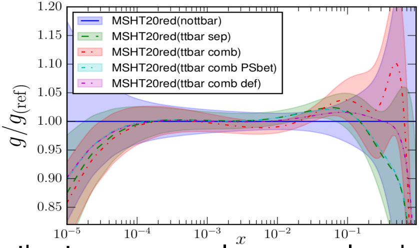
- Not only tensions between different dataset types at high x, also tensions within dataset types, e.g. between different jet measurements.
- ATLAS 7 TeV jets pulls gluon down at high x, whereas CMS jets (mainly 8 TeV) pull gluon up.
- Global fit is a balance between these different pulls and those of Zp_T, $t\bar{t}$ datasets here.



† MSHT20, TC, S. Bailey, L. Harland-Lang, A. Martin, R. Thorne 2012.04684

ATLAS 8 TeV multi-differential $t\bar{t}$ lepton+jets Preliminary!

- What effect does the inclusion of this data in the reduced fit have on the gluon?

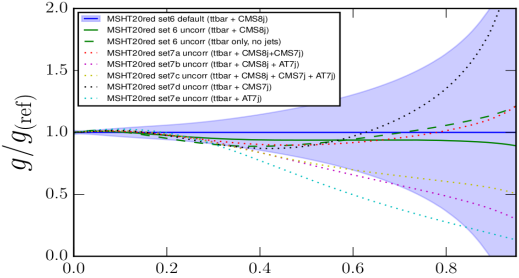


- Fitting all 4 distributions separately, uncorrelated ⇒ gluon moves down at high x, driven by the rapidity data.
- Applying correlations ⇒ gluon raised and shape altered at high x .
- Decorrelating parton shower between distributions ⇒ reverts the gluon to shape obtained when all 4 separately uncorrelated fitted.
- Additionally decorrelating within distributions ⇒ moves gluon closer to fit without $t\bar{t}$ data as its constraining power is reduced.
- Overall, gluon shape moves in direction of global fit gluon.

ATLAS 8 TeV $t\bar{t}$ with various jet datasets: gluon

- What effect does the inclusion of various other jet datasets in addition to the $t\bar{t}$ have on the gluon?

Very Preliminary!

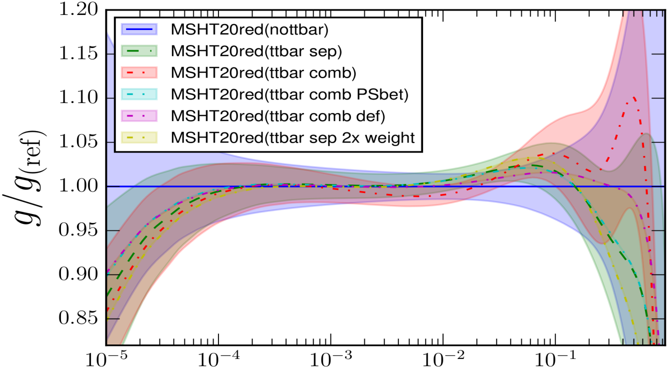


- $t\bar{t}$ improves when gluon pulled down, as does ATLAS 7 TeV jets and CMS 7 TeV jets (although it pulls gluon back up at highest x), CMS 8 TeV jets improves when gluon pulled up.
- Decorrelation reduces impact of data on gluon so it is pulled down less than the "uncorr" case and jet datasets worsen (7 TeV) / improve (8 TeV) accordingly.

ATLAS 8 TeV multi-differential $t\bar{t}$ lepton+jets gluon

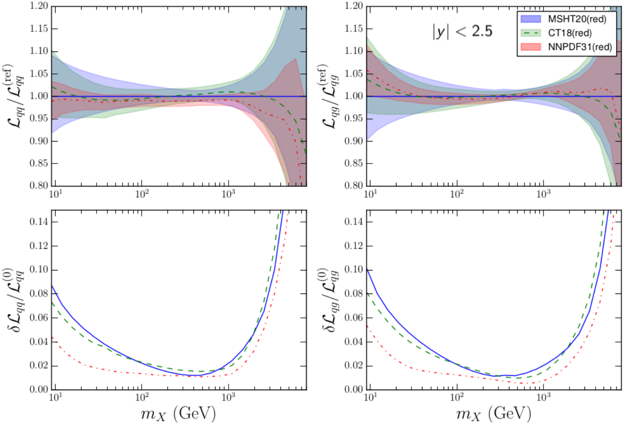
- What effect does the inclusion of this data in the reduced fit have on the gluon?

Preliminary!



- Double weighting (yellow) pulls gluon further in direction of rapidity pull (lower at high x) as expected.

Reduced Fits: Current Status Summary*



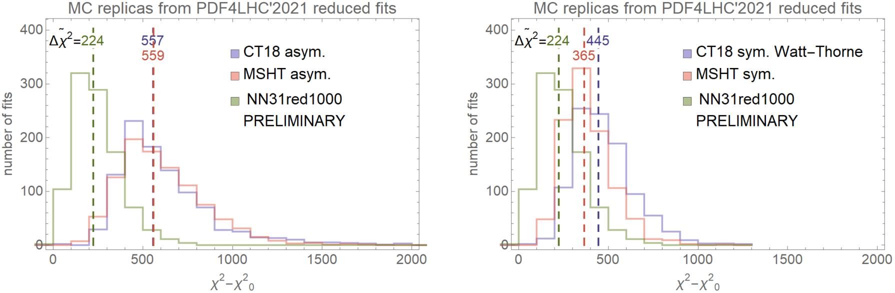
- Very good agreement in the gluon-gluon, quark-quark and quark-gluon luminosities.
- Small difference in quark-antiquark luminosity, still some flavour decomposition differences, although within MSHT uncertainties.

*Note this is without the $t\bar{t}$ added.

Reduced Fits χ^2 replica distributions

Preliminary!

- As well as the overall PDFs, can analyse χ^2.
- Can use CT/MSHT reduced fit eigenvectors to generate replicas.



- Overall distributions of reduced fits replicas similar between CT and MSHT, particularly non-symmetrised versions.
- Symmetrised versions in better agreement with NNPDF but still different.
- Some limited qualitative agreement at least for symmetrised case.

Plots from P. Nadolsky

Deuteron and Nuclear Corrections in MSHT20

- Several older DIS datasets use deuteron or heavy nuclear targets.
- Deuteron data required to fully separate u, d at moderate-large x.
- Heavy nuclear data, via C.C. scattering, required for more constraints on flavour decomposition and strange (dimuon data).
- Deuteron correction is 4-parameter prefactor to usual average of p and n:

$$F^d(x,Q^2) \;=\; c(x)\left[F^p(x,Q^2)+F^n(x,Q^2)\right]/2,$$

$$\begin{aligned} c(x) &= (1+0.01N)\,[1+0.01c_1\ln^2(x_p/x)], && x<x_p,\\ c(x) &= (1+0.01N)\,[1+0.01c_2\ln^2(x/x_p)+0.01c_3\ln^{20}(x/x_p)], && x>x_p, \end{aligned}$$

- Nuclear correction is prefactor*: *de Florian et al arXiv:1112.6324.

$$f^A(x,Q^2) \;=\; R_f(x,Q^2,A)\;f(x,Q^2).$$

- This is multiplied by a 3-parameter modification function to allow penalty-free change in shape and/or normalisation.
- Both deuteron and nuclear corrections prefer modifications of 1%.

More details on all of this in MMHT14 1412.3989, MSHT20 2012.04684.

PDF4LHC21 Benchmarking Summary:

- Great amounts of new data, theoretical improvements, PDF methodological improvements have meant substantial changes since PDF4LHC15.
- We have been performing a benchmarking exercise of the 3 global fit PDF groups most recent sets: MSHT20, CT18, NNPDF3.1.
- Based on comparing "Reduced Fits" with common dataset and common theory settings where possible.
- Goal of exercise is the understanding of differences which have emerged in PDF central values and uncertainties.
  $\Rightarrow$ Good progress.
- End result: PDF4LHC21 set of PDFs, central PDFs and Hessian error set (30-50 sufficient) representing the 3 published PDFs.
- We welcome suggestions, feedback and discussion!

More details on all of this in the slides!

Questions for Experimentalists/Users:

- Are there any lessons from experience with PDF4LHC15 we can take into account?
- Now would be the time to account for these before the benchmarking is finished and combination is performed...
- What form should the output set be provided in? E.g. Hessian and/or MC replicas?
- Are there additional variation sets we could consider providing?
    - Effect of perturbative vs fitted charm: Relevant for NNPDF. Use in providing alternative set with all perturbative charm?
    - Small x resummation effects, effects low x gluon, could be relevant for 100TeV collider? PDF4100TeV?
    - Any need for such sets or others (inclusive jets vs dijets for example)?
- We welcome suggestions, feedback and discussion!