3 Parton distribution functions[6]

3.1 PDF set updates

Several of the PDF sets which were discussed in the previous Yellow Report [7] have been updated since. NNPDF2.1 [95] is an updated version of NNPDF2.0 [96] which uses the FONLL general mass VFN scheme [97] instead of a zero-mass scheme. There are also now NNPDF NNLO and LO sets [98], which are based on the same data and methodology as NNPDF2.1. The HERA PDF group now use HERAPDF1.5 [99], which contains more data than HERAPDF1.0 [100], a wider examination of parameter dependence, and an NNLO set with uncertainty. This set is available in LHAPDF, however it is partially based on preliminary HERA-II structure-function data.

The current PDF4LHC prescription [101, 102] for calculating a central value and uncertainty for a given process should undergo the simple modification of using the most up-to-date relevant set from the relevant group, i.e., NNPDF2.0 should be replaced with NNPDF2.1 [95] and CTEQ6.6 [103] with CT10 [104]. At NNLO the existing prescription should be used, but with the uncertainty envelope calculated using the up-to-date sets noted above.

3.2 Correlations

The main aim of this section is to examine the correlations between different Higgs production processes and/or backgrounds to these processes. The PDF uncertainty analysis may be extended to define a *correlation* between the uncertainties of two variables, say $X(\vec{a})$ and $Y(\vec{a})$. As for the case of PDFs, the physical concept of PDF correlations can be determined both from PDF determinations based on the Hessian approach and on the Monte Carlo approach. For convenience and commonality, all physical processes were calculated using the MCFM NLO program (versions 5.8 and 6.0) [105, 106] with a common set of input files for all groups.

We present the results for the PDF correlations at NLO for Higgs production via gluon–gluon fusion, vector-boson fusion, in association with W or with a $\mathrm{t\overline{t}}$ pair at masses $M_{\mathrm{H}} = 120$ GeV, $M_{\mathrm{H}} = 160$ GeV, $M_{\mathrm{H}} = 200$ GeV, $M_{\mathrm{H}} = 300$ GeV, $M_{\mathrm{H}} = 500$ GeV. We also include a wide variety of background processes and other standard production mechanisms, i.e., W, WW, WZ, Wγ, $\mathrm{Wb\overline{b}}$, $\mathrm{t\overline{t}}$, $\mathrm{t\overline{b}}$ and the t-channel $\mathrm{t}(\rightarrow \overline{\mathrm{b}}) + \mathrm{q}$, where W denotes the average of W^+ and W^-.

For MSTW2008 [107], CT10 [104], GJR08 [108, 109], and ABKM09 [110] PDFs the correlations of any two quantities X and Y are calculated using the standard formula [111]

$$\rho\left(X,Y\right) \equiv \cos\varphi = \frac{\sum_i (X_i - X_0)(Y_i - Y_0)}{\sqrt{\sum_i (X_i - X_0)^2 \sum_i (Y_i - Y_0)^2}}. \tag{7}$$

The index in the sum runs over the number of free parameters and X_0, Y_0 correspond to the values obtained by the central PDF value. When positive and negative direction PDF eigenvectors are used this is equivalent to (see e.g., Ref. [103])

$$\begin{aligned} \cos\varphi &= \frac{\vec{\Delta}X \cdot \vec{\Delta}Y}{\Delta X \Delta Y} = \frac{1}{4\Delta X \Delta Y} \sum_{i=1}^{N} \left(X_i^{(+)} - X_i^{(-)}\right)\left(Y_i^{(+)} - Y_i^{(-)}\right), \\ \Delta X &= \left|\vec{\Delta}X\right| = \frac{1}{2}\sqrt{\sum_{i=1}^{N}\left(X_i^{(+)} - X_i^{(-)}\right)^2}, \end{aligned} \tag{8}$$

where the sum is over the N pairs of positive and negative PDF eigenvectors. For MSTW2008 and CT10 the eigenvectors by default contain only PDF parameters, and α_{s} variation may be considered separately. For GJR08 and ABKM09, α_{s} is one of the parameters used directly in calculating the correlation coefficient, with the central value and variation determined by the fit.

[6] S. Forte, J. Huston and R. S. Thorne (eds); S. Alekhin, J. Blümlein, A.M. Cooper-Sarkar, S. Glazov, P. Jimenez-Delgado, S. Moch, P. Nadolsky, V. Radescu, J. Rojo, A. Sapronov and W.J. Stirling.

Due to the specific error calculation prescription for HERAPDF1.5 which includes parametrisation and model errors, the correlations can not be calculated in exactly the same way. Rather, a formula for uncertainty propagation can be used in which correlations can be expressed via relative errors of compounds and their combination:

$$\left[\sigma\left(\frac{X}{\sigma(X)}+\frac{Y}{\sigma(Y)}\right)\right]^2 = 2+2\cos\varphi, \tag{9}$$

where $\sigma(O)$ is the PDF error of observable O calculated using the HERAPDF prescription.

The correlations for the NNPDF prescription are calculated as discussed in Ref. [112], namely

$$\rho\left(X,Y\right) = \frac{\langle XY\rangle_{\rm rep} - \langle X\rangle_{\rm rep}\langle Y\rangle_{\rm rep}}{\sigma_X\sigma_Y} \tag{10}$$

where the averages are performed over the $N_{\rm rep} = 100$ replicas of the NNPDF2.1 set.

For all sets the correlations are for the appropriate value of $\alpha_{\rm s}$ for the relevant PDF set.

3.2.1 Results for the correlation study

Our main result is the computation of the correlation between physical processes relevant to Higgs production, either as signal or as background. It is summarised in Tables 10 and 11, where we show the PDF4LHC average for each of the correlations between signal and background processes considered. These tables classify each correlation in classes with $\Delta\rho = 0.2$, that is, if the correlation is $1 > \rho > 0.9$ the processes is assigned correlation 1, if $0.9 > \rho > 0.7$ the processes is assigned correlation 0.8 and so on. The class width is typical of the spread of the results from the PDF sets, which are in generally very good, but not perfect agreement. The average is obtained using the most up-to-date PDF sets (CT10, MSTW2008, NNPDF2.1) in the PDF4LHC recommendation, and it is appropriate for use in conjunction with the cross-section results obtained in Ref. [7], and with the background processes listed; the change of correlations due to updating the prescription is insignificant in comparison to the class width.

We have also compared this PDF4LHC average to the average using all six PDF sets. In general there is rather little change. There are quite a few cases where the average moves into a neighbouring class but in many cases due to a very small change taking the average just over a boundary between two classes. There is only a move into the next-to-neighbouring class, i.e., a change of more than 0.2, in a very small number of cases. For the VBF–Wγ correlation at $M_{\rm H} = 120, 160$ GeV it reduces from 0.6 to 0.2, for the VBF–W correlation at $M_{\rm H} = 500$ GeV it reduces from 0.4 to 0, for the Wγ–$\rm t\overline{b}$ correlation it reduces from 0.8 to 0.4, for WZ–$\rm t\overline{t}H$ at $M_{\rm H} = 120$ GeV it increases from -0.4 to 0.0, and for $\rm Wb\overline{b}$–$\rm t\overline{t}H$ at $M_{\rm H} = 200$ GeV it increases from -0.2 to 0.2.

A more complete list of processes, with results for each individual PDF set, may be found in the tables on the webpage at the LHC Higgs Cross Section Working Group TWiki [113]. Note, however, that there is a high degree of redundancy in the approximate correlations of many processes. For example W production is very similar to Z production, both depending on partons (quarks in this case) at very similar hard scales and x values. Similarly for WW and ZZ, and the subprocesses $\rm gg \to WW(ZZ)$ and Higgs production via gluon–gluon fusion for $M_{\rm H} = 200$ GeV.

More detailed results are presented for the MSTW2008, NNPDF2.1, CT10, HERAPDF1.5, GJR08, and ABKM09 PDFs in Figures 6–21. The result using each individual PDF sets is compared to the (updated) PDF4LHC average. There is usually a fairly narrow clustering of the individual results about the average, with a small number of cases where there is one, or perhaps two outliers. The sets with the largest parametrisations for the PDFs generally tend to give smaller magnitude correlations or anticorrelations, but this is not always the case, e.g., NNPDF2.1 gives the largest anti-correlation for VBF–$\rm t\overline{t}H$. There are some unusual features, e.g., for HERAPDF1.5 and high values of $M_{\rm H}$, the $\rm t\overline{t}H$ correlations with quantities depending on the high-x gluon, e.g., $\rm ggH$ and $\rm t\overline{t}$ is opposite to the other sets

Table 10: The up-to-date PDF4LHC average for the correlations between all signal processes with other signal and background processes for Higgs production considered here. The processes have been classified in correlation classes, as discussed in the text.

$M_{\mathrm{H}} = 120$ GeV	ggH	VBF	WH	$\mathrm{t\bar{t}H}$
ggH	1	−0.6	−0.2	−0.2
VBF	−0.6	1	0.6	−0.4
WH	−0.2	0.6	1	−0.2
$\mathrm{t\bar{t}}H$	−0.2	−0.4	−0.2	1
W	−0.2	0.6	0.8	−0.6
WW	−0.4	0.8	1	−0.2
WZ	−0.2	0.4	0.8	−0.4
Wγ	0	0.6	0.8	−0.6
$\mathrm{Wb\bar{b}}$	−0.2	0.6	1	−0.2
$\mathrm{t\bar{t}}$	0.2	−0.4	−0.4	1
$\mathrm{t\bar{b}}$	−0.4	0.6	1	−0.2
$\mathrm{t}(\rightarrow \mathrm{\bar{b}})\mathrm{q}$	0.4	0	0	0

$M_{\mathrm{H}} = 160$ GeV	ggH	VBF	WH	$\mathrm{t\bar{t}H}$
ggH	1	−0.6	−0.4	0.2
VBF	−0.6	1	0.6	−0.2
WH	−0.4	0.6	1	0
$\mathrm{t\bar{t}H}$	0.2	−0.2	0	1
W	−0.4	0.4	0.6	−0.4
WW	−0.4	0.6	0.8	−0.2
WZ	−0.4	0.4	0.8	−0.2
Wγ	−0.4	0.6	0.6	−0.6
$\mathrm{Wb\bar{b}}$	−0.2	0.6	0.8	−0.2
$\mathrm{t\bar{t}}$	0.4	−0.4	−0.2	0.8
$\mathrm{t\bar{b}}$	−0.4	0.6	1	0
$\mathrm{t}(\rightarrow \mathrm{\bar{b}})\mathrm{q}$	0.6	0	0	0

$M_{\mathrm{H}} = 200$ GeV	ggH	VBF	WH	$\mathrm{t\bar{t}H}$
ggH	1	−0.6	−0.4	0.4
VBF	−0.6	1	0.6	−0.2
WH	−0.4	0.6	1	0
$\mathrm{t\bar{t}H}$	0.4	−0.2	0	1
W	−0.6	0.4	0.6	−0.4
WW	−0.4	0.6	0.8	−0.2
WZ	−0.4	0.4	0.8	−0.2
Wγ	−0.4	0.4	0.6	−0.6
$\mathrm{Wb\bar{b}}$	−0.2	0.6	0.8	−0.2
$\mathrm{t\bar{t}}$	0.6	−0.4	−0.2	0.8
$\mathrm{t\bar{b}}$	−0.4	0.6	0.8	0
$\mathrm{t}(\rightarrow \mathrm{\bar{b}})\mathrm{q}$	0.6	−0.2	0	0

$M_{\mathrm{H}} = 300$ GeV	ggH	VBF	WH	$\mathrm{t\bar{t}H}$
ggH	1	−0.4	−0.2	0.6
VBF	−0.4	1	0.4	−0.2
WH	−0.2	0.4	1	0.2
$\mathrm{t\bar{t}H}$	0.6	−0.2	0.2	1
W	−0.6	0.4	0.4	−0.6
WW	−0.4	0.6	0.8	−0.2
WZ	−0.6	0.4	0.6	−0.4
Wγ	−0.6	0.4	0.4	−0.6
$\mathrm{Wb\bar{b}}$	−0.2	0.4	0.8	−0.2
$\mathrm{t\bar{t}}$	1	−0.4	0	0.8
$\mathrm{t\bar{b}}$	−0.4	0.4	0.8	−0.2
$\mathrm{t}(\rightarrow \mathrm{\bar{b}})\mathrm{q}$	0.4	−0.2	0	−0.2

$M_{\mathrm{H}} = 500$ GeV	ggH	VBF	WH	$\mathrm{t\bar{t}H}$
ggH	1	−0.4	0	0.8
VBF	−0.4	1	0.4	−0.2
WH	0	0.4	1	0
$\mathrm{t\bar{t}H}$	0.8	−0.2	0	1
W	−0.6	0.4	0.2	−0.6
WW	−0.4	0.6	0.6	−0.4
WZ	−0.6	0.4	0.6	−0.4
Wγ	−0.6	0.4	0.2	−0.6
$\mathrm{Wb\bar{b}}$	−0.4	0.4	0.6	−0.4
$\mathrm{t\bar{t}}$	1	−0.4	0	0.8
$\mathrm{t\bar{b}}$	−0.4	0.4	0.8	−0.2
$\mathrm{t}(\rightarrow \mathrm{\bar{b}})\mathrm{q}$	0.2	−0.2	0	−0.2

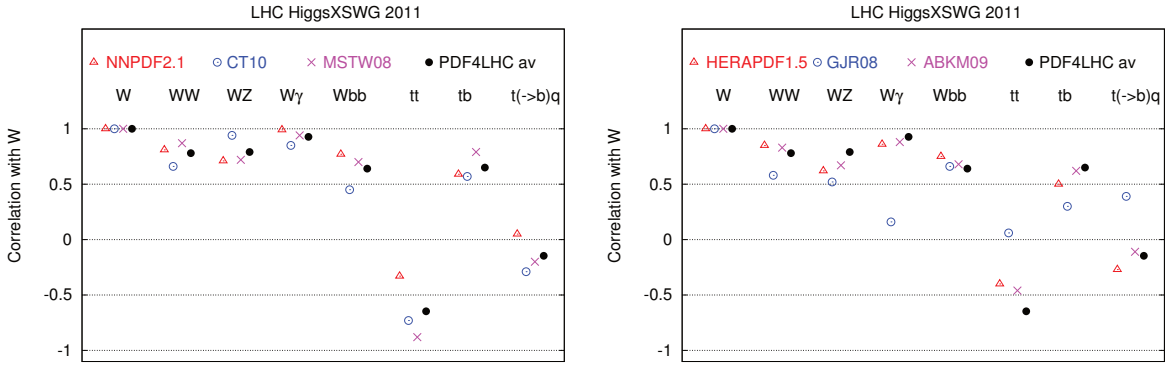

Fig. 6: The correlations between W production and the other background processes considered. We show the results for NNPDF2.1, CT10 and MSTW2008 in the left plot, and HERAPDF, JR and ABKM in the right plot. In both cases we show the up-to-date PDF4LHC average result.

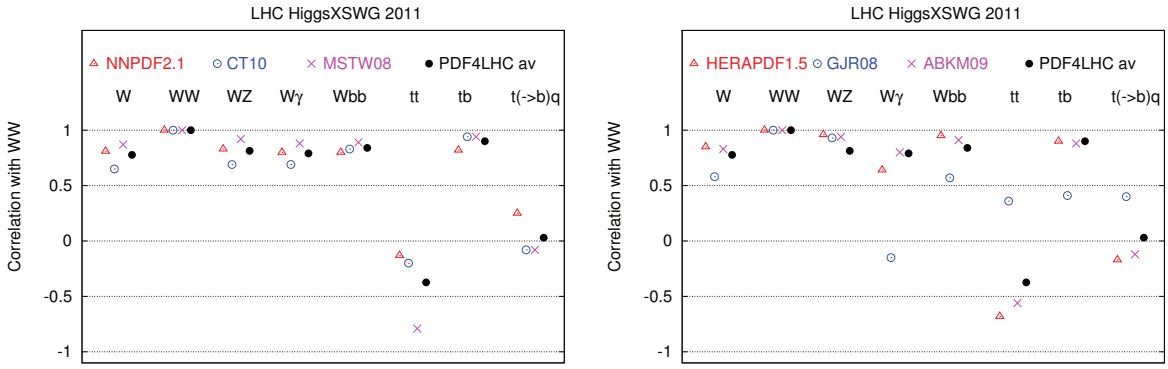

Fig. 7: The same as Figure 6 for WW production.

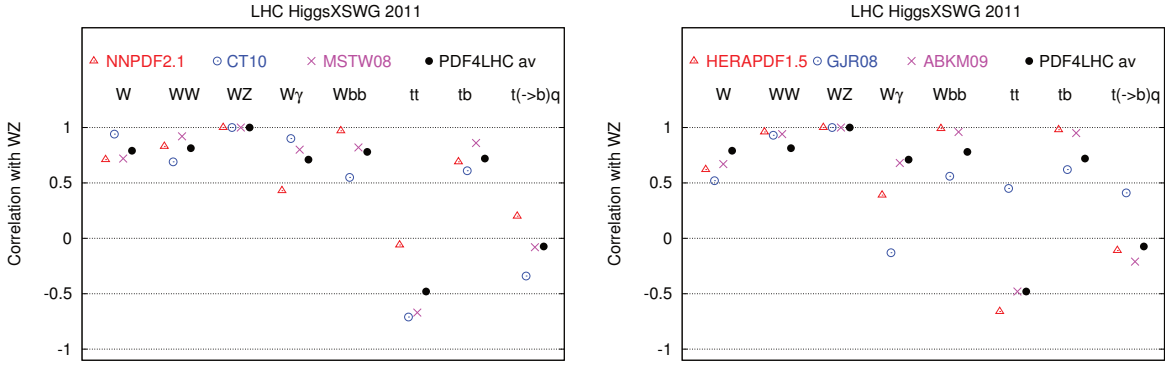

Fig. 8: The same as Figure 6 for WZ production.

Table 11: The same as Table 10 for the correlations between background processes.

	W	WW	WZ	Wγ	$Wb\bar{b}$	$t\bar{t}$	$t\bar{b}$	$t(\to \bar{b})q$
W	1	0.8	0.8	1	0.6	−0.6	0.6	−0.2
WW	0.8	1	0.8	0.8	0.8	−0.4	0.8	0
WZ	0.8	0.8	1	0.8	0.8	−0.4	0.8	0
Wγ	1	0.8	0.8	1	0.6	−0.6	0.8	0
$Wb\bar{b}$	0.6	0.8	0.8	0.6	1	−0.2	0.6	0
$t\bar{t}$	−0.6	−0.4	−0.4	−0.6	−0.2	1	−0.4	0.2
$t\bar{b}$	0.6	0.8	0.8	0.8	0.6	−0.4	1	0.2
$t(\to \bar{b})q$	−0.2	0	0	0	0	0.2	0.2	1

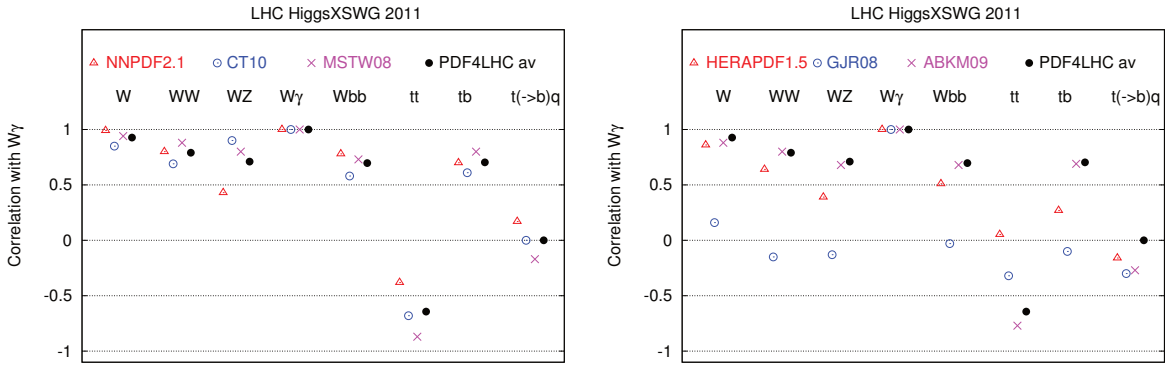

Fig. 9: The same as Figure 6 for Wγ production.

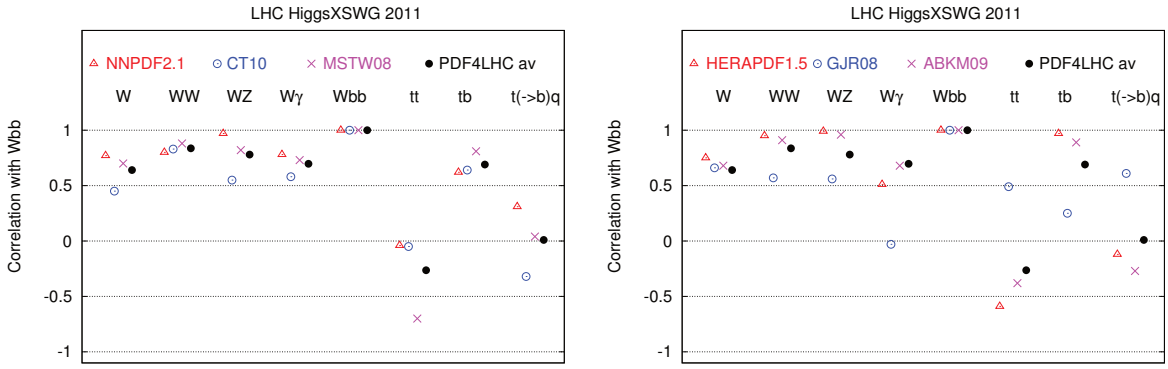

Fig. 10: The same as Figure 6 for $Wb\bar{b}$ production.

and the correlations with quantities depending on high-x quarks and antiquarks, e.g., VBF and WW, is stronger. This is possibly related to the large high-x antiquark distribution in HERAPDF which contributes to $t\bar{t}H$ but not ggH or very much to $t\bar{t}$. GJR08 has a tendency to obtain more correlation between some gluon dominated processes, e.g., ggH and $t\bar{t}$ and quark dominated processes, e.g., W and WZ, perhaps because the dynamical generation of PDFs couples the gluon and quark more strongly.

We can now also see the origin of the cases where the averages move two classes. For VBF–Wγ at lower masses GJR08, ABKM09, and HERAPDF1.5 all lie lower than the (updated) PDF4LHC average, but not too different to CT10. For VBF–W at $M_H = 500$ GeV, ABKM09, and GJR08 give a small anticorrelation, while others give a correlation, though it is only large in the case of MSTW2008. For Wγ–$t\bar{b}$ the change is due to slightly lower correlations for GJR08. However, although the change is two

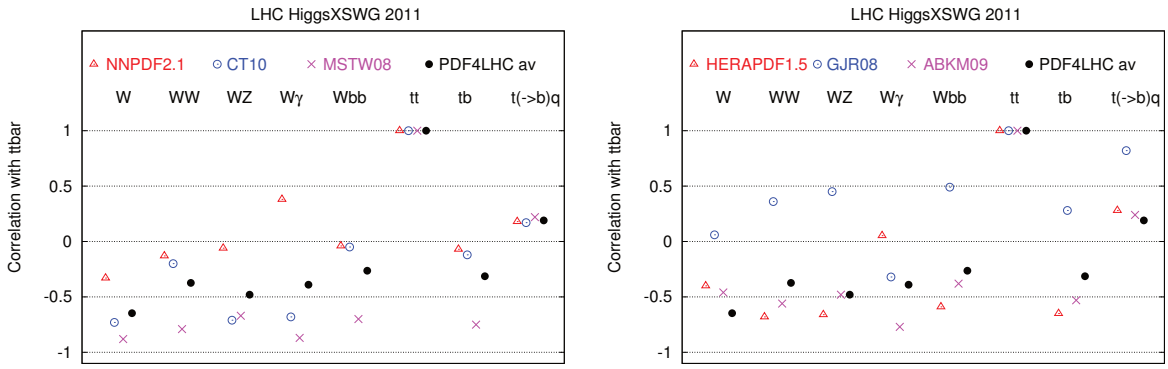

Fig. 11: The same as Figure 6 for $\mathrm{t\bar{t}}$ production.

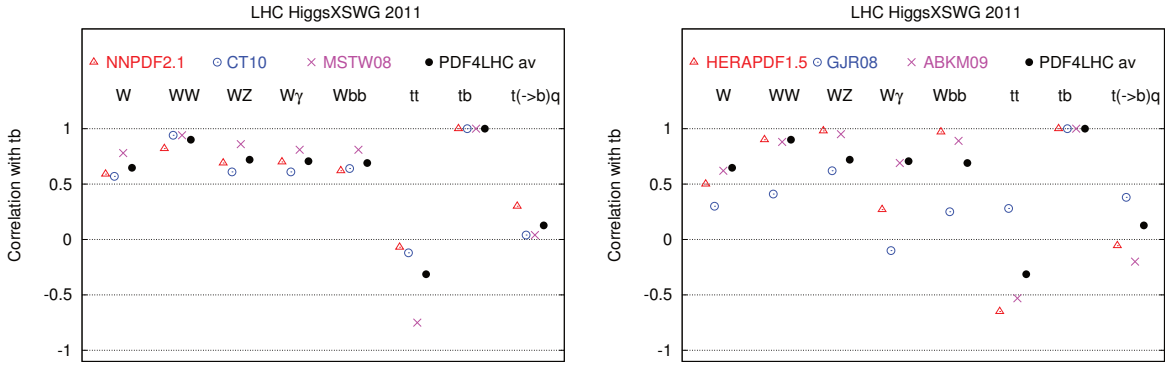

Fig. 12: The same as Figure 6 for $\mathrm{t\bar{b}}$ production.

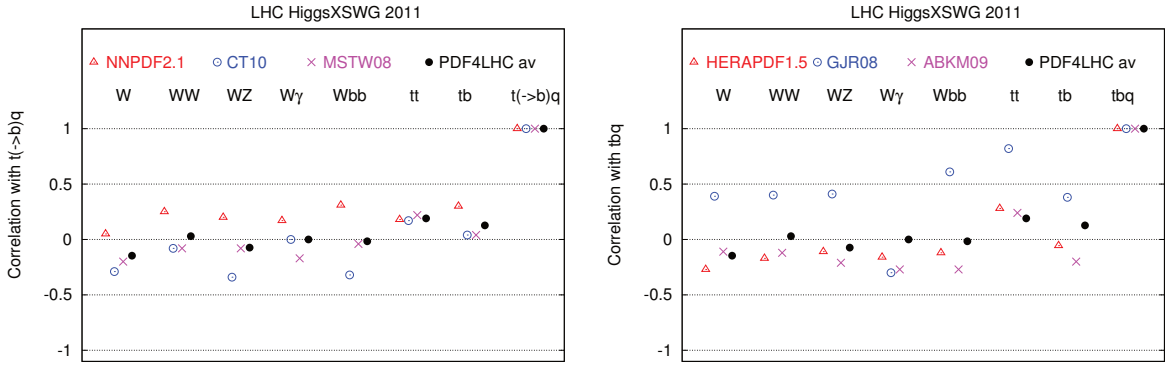

Fig. 13: The same as Figure 6 for t-channel $\mathrm{t}(\to \mathrm{\bar{b}})\mathrm{q}$ production.

classes, in practice it is barely more than 0.2. For WZ–$\mathrm{t\bar{t}H}$ at $M_{\mathrm{H}} = 120$ GeV both HERAPDF1.5 and GJR08 have a larger correlation. This increases with M_{H} for HERAPDF1.5 as noted, whereas GJR08 heads closer to the average, but the move at $M_{\mathrm{H}} = 120$ GeV is only marginally two classes, and is only one class for other masses. For $\mathrm{Wb\bar{b}}$–$\mathrm{t\bar{t}H}$ at $M_{\mathrm{H}} = 200$ GeV the situation is similar, and MSTW2008 gives easily the biggest pull in the direction of anticorrelation.

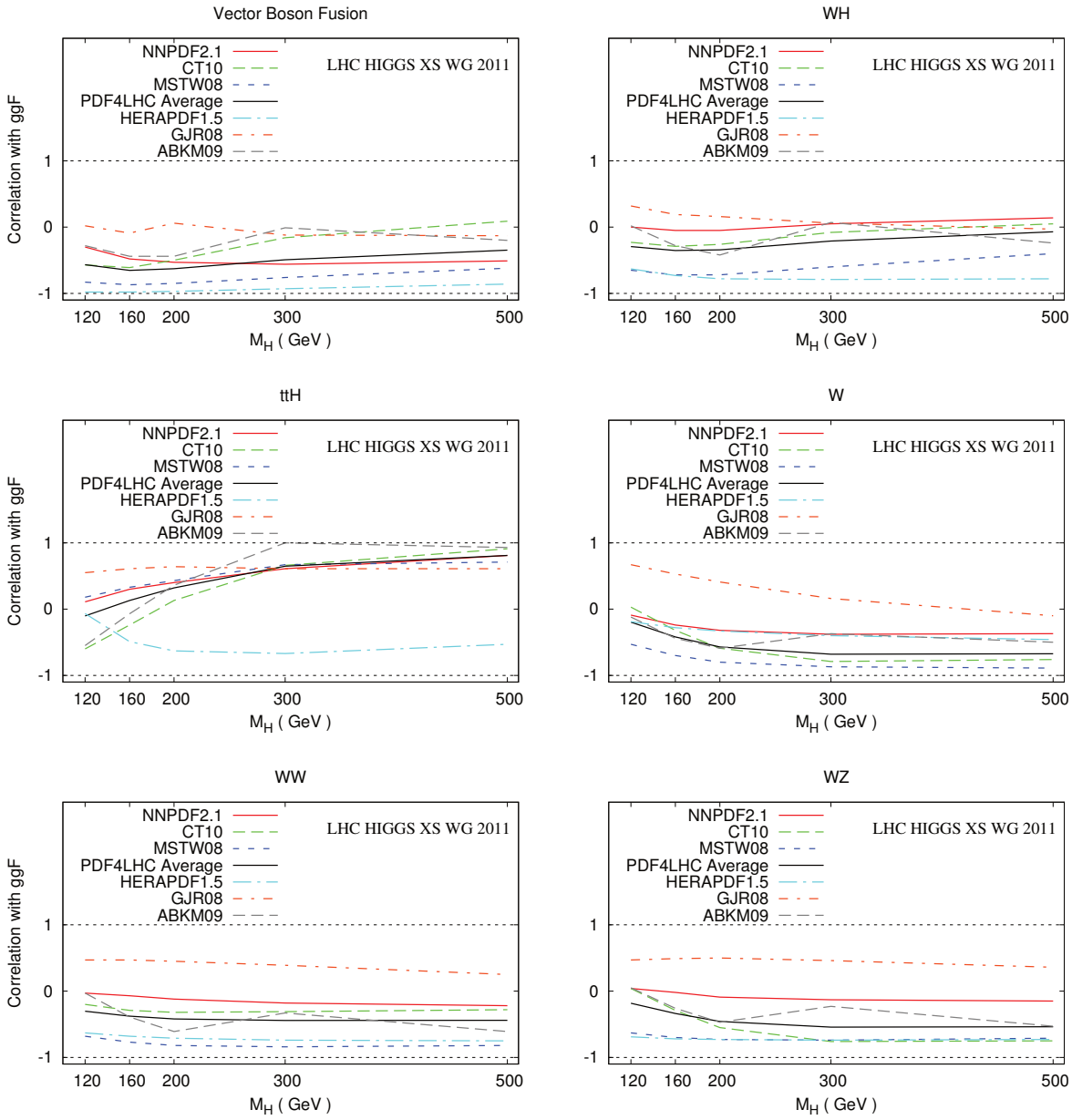

Fig. 14: Correlation between the gluon-fusion $\mathrm{gg} \to \mathrm{H}$ process and other signal and background processes as a function of M_{H}. We show the results for the individual PDF sets as well as the up-to-date PDF4LHC average.

3.2.2 Additional correlation studies

The inclusion of the α_{s} uncertainty on the correlations, compared to PDF only variation, was also studied for some PDF sets, e.g., MSTW2008 using the approximation that the PDF sets for the upper and lower α_{s} values [114] simply form another pair of orthogonal eigenvectors (show to be true in the quadratic approximation [115]). This increases the correlation between some processes, i.e., W production and Higgs via gluon fusion, because the former increases with α_{s} due to increased evolution of quarks while the latter increases due to direct dependence on α_{s}. Similarly for e.g., Higgs production via gluon fusion and $\mathrm{t\bar{t}}$ production since each depends directly on α_{s}. This may also contribute to some of the stronger correlations seen using GJR08 in some similar cases. In a handful of processes it reduces correlation, e.g., W and $\mathrm{Wb\bar{b}}$ since the latter has a much stronger α_{s} dependence. In most cases the change compared

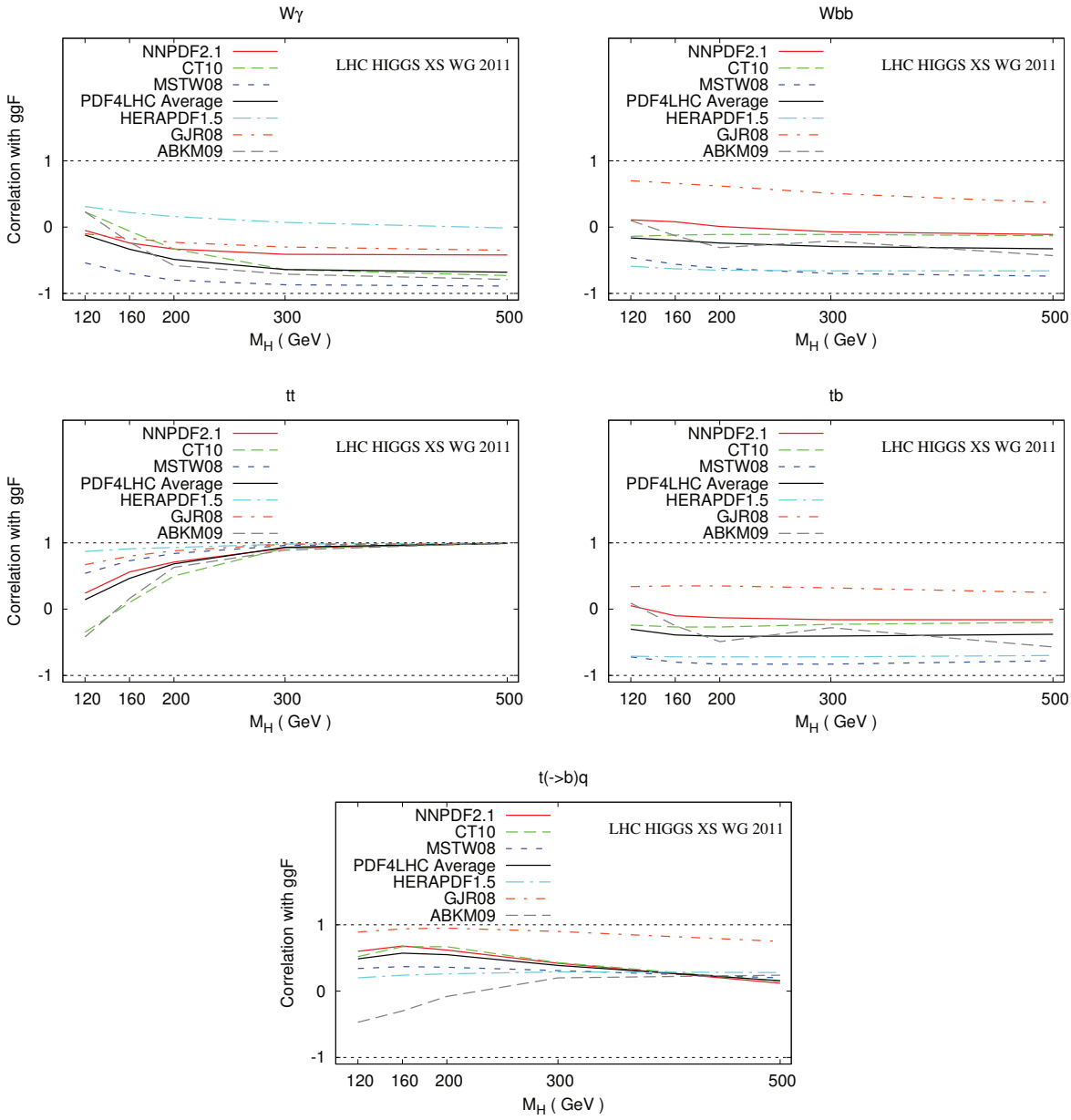

Fig. 15: Correlation between the gluon fusion $\mathrm{gg} \to \mathrm{H}$ process and other signal and background processes as a function of M_{H}. We show the results for the individual PDF sets as well as the up-to-date PDF4LHC average.

to PDF-only correlation for a given PDF sets small, and it is not an obvious contributing factor to the cases where the (updated) PDF4LHC average is noticeably different to the average using six sets, except possibly to $\mathrm{Wb\overline{b}}$–$\mathrm{t\overline{t}H}$ at $M_{\mathrm{H}} = 200$ GeV, where α_{s} does increase correlation.

A small number of correlations were also calculated at NNLO for MSTW2008 PDFs, i.e., W, Z and $\mathrm{gg} \to \mathrm{H}$ for the same range of M_{H}. The correlations when taking into account PDF uncertainty alone were almost identical to those at NLO, variations being less than 0.05. When α_{s} uncertainty was included the correlations changed a little more due to $\mathrm{gg} \to \mathrm{H}$ having more direct dependence on α_{s}, but this is a relatively minor effect. Certainly the results in Tables 10 and 11, though calculated at NLO, can be used with confidence at NNLO.

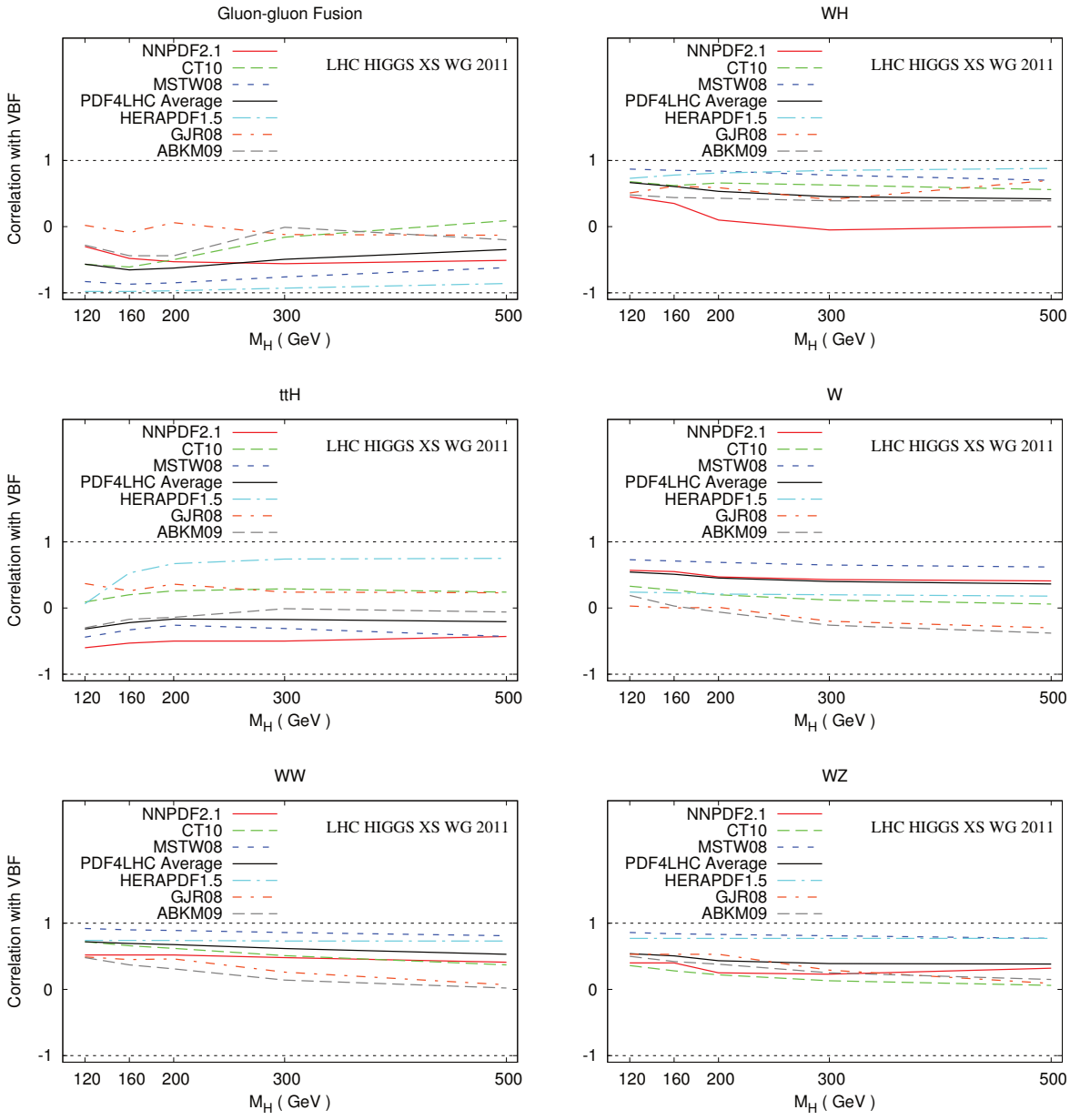

Fig. 16: The same as Figure 14 for the vector boson fusion process.

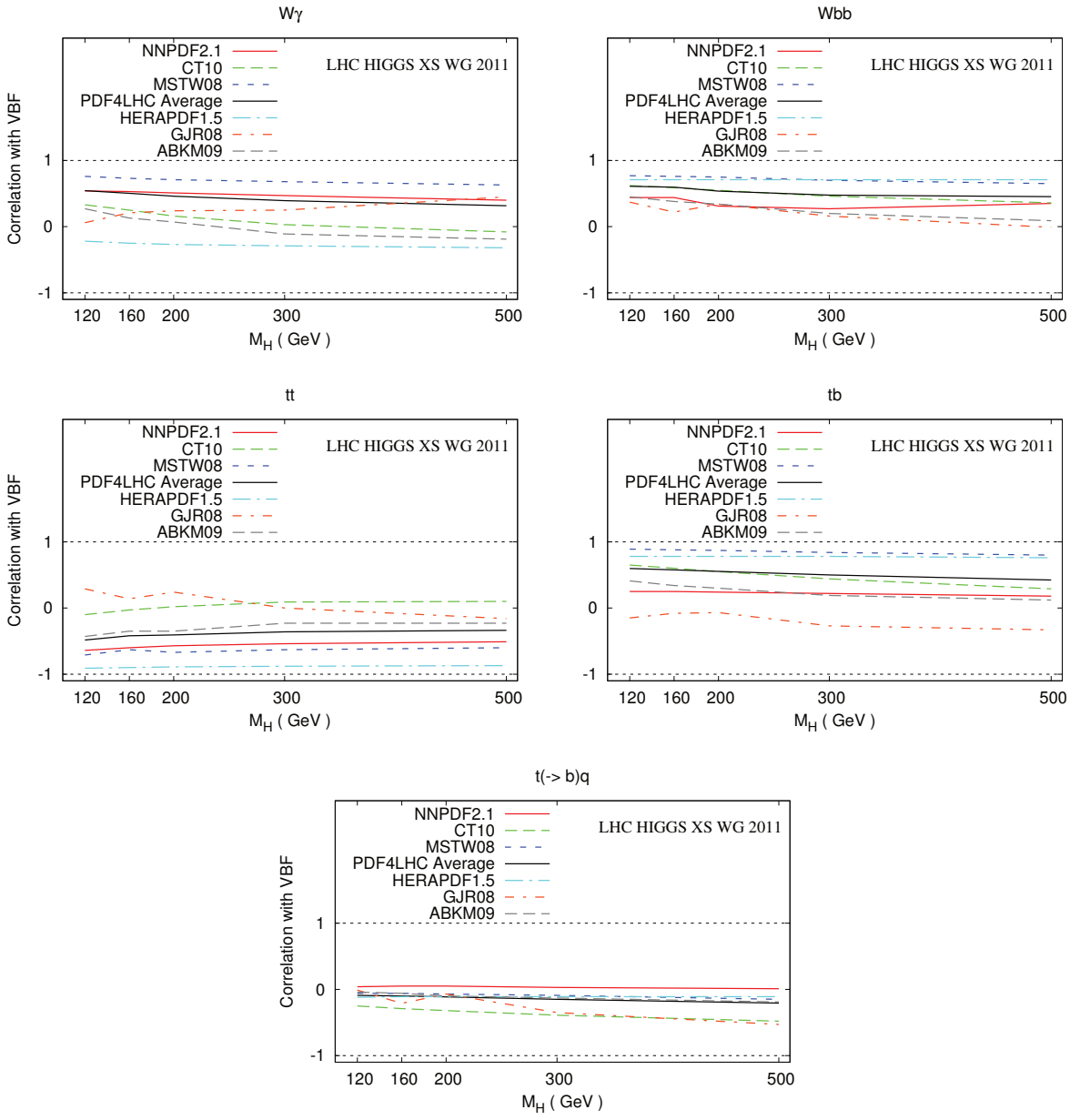

Fig. 17: The same as Figure 15 for the vector boson fusion process.

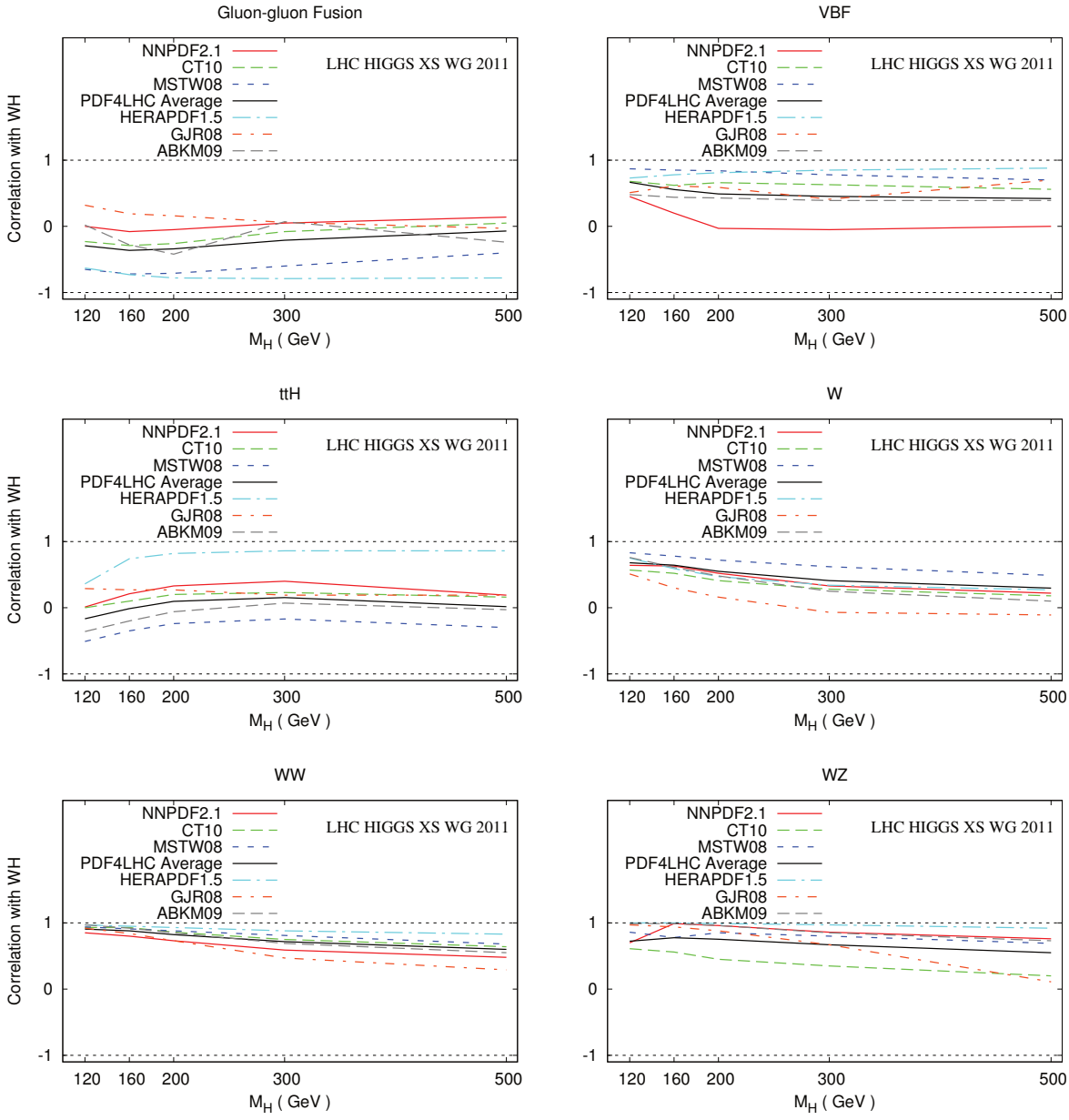

Fig. 18: The same as Figure 14 for the WH production process.

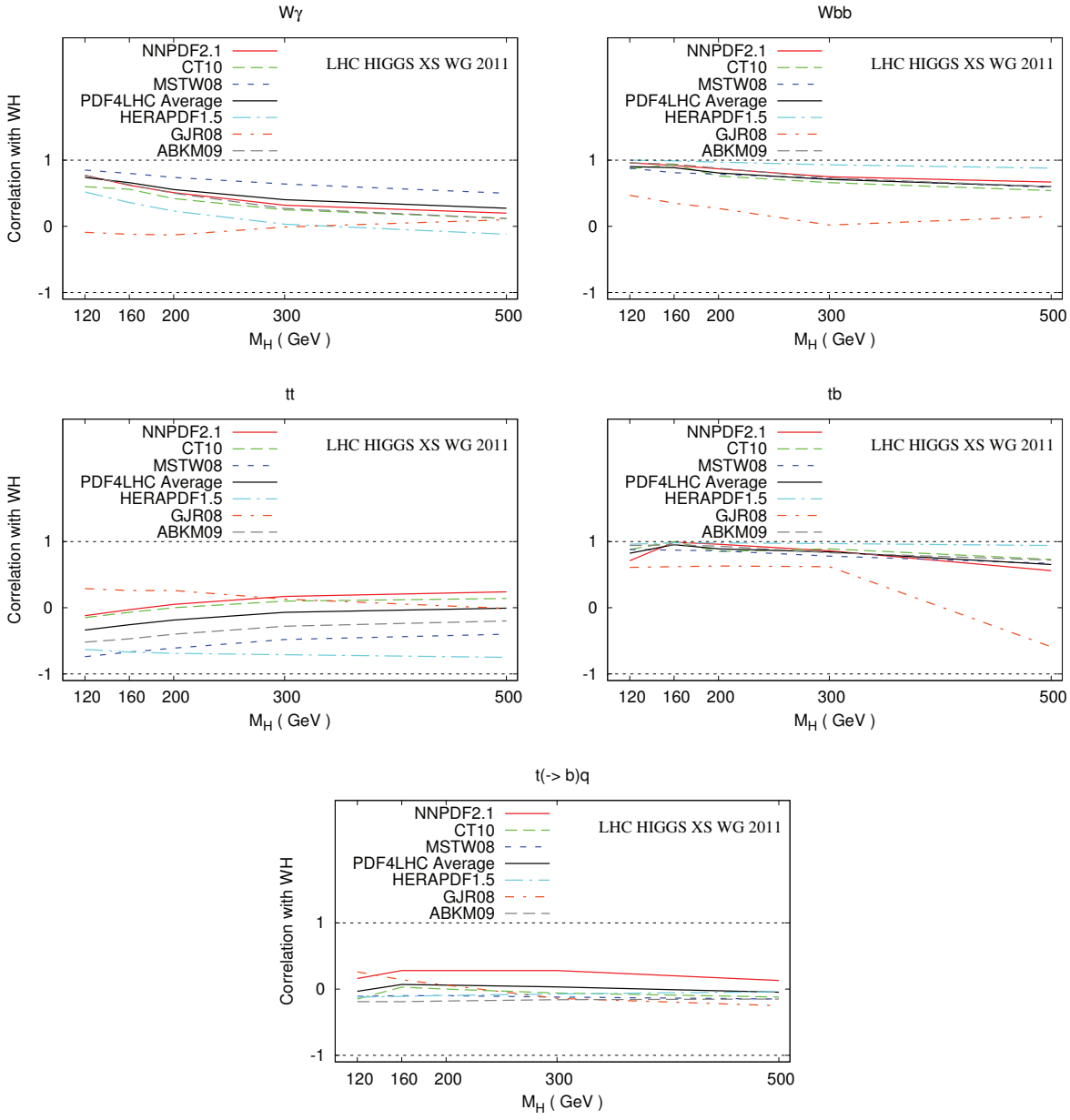

Fig. 19: The same as Figure 15 for the WH production process.

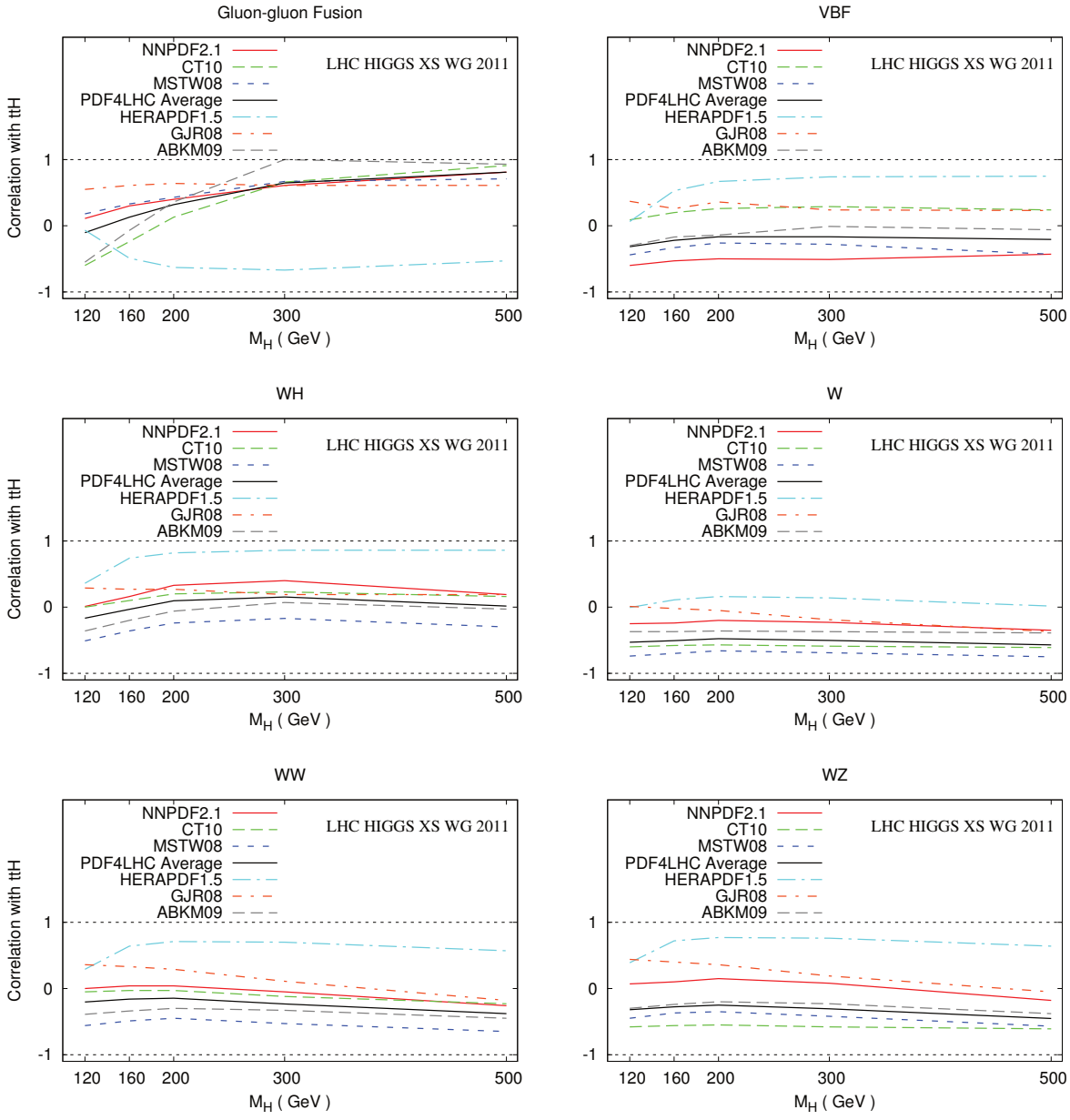

Fig. 20: The same as Figure 14 for the $\mathrm{t\bar{t}H}$ production process.

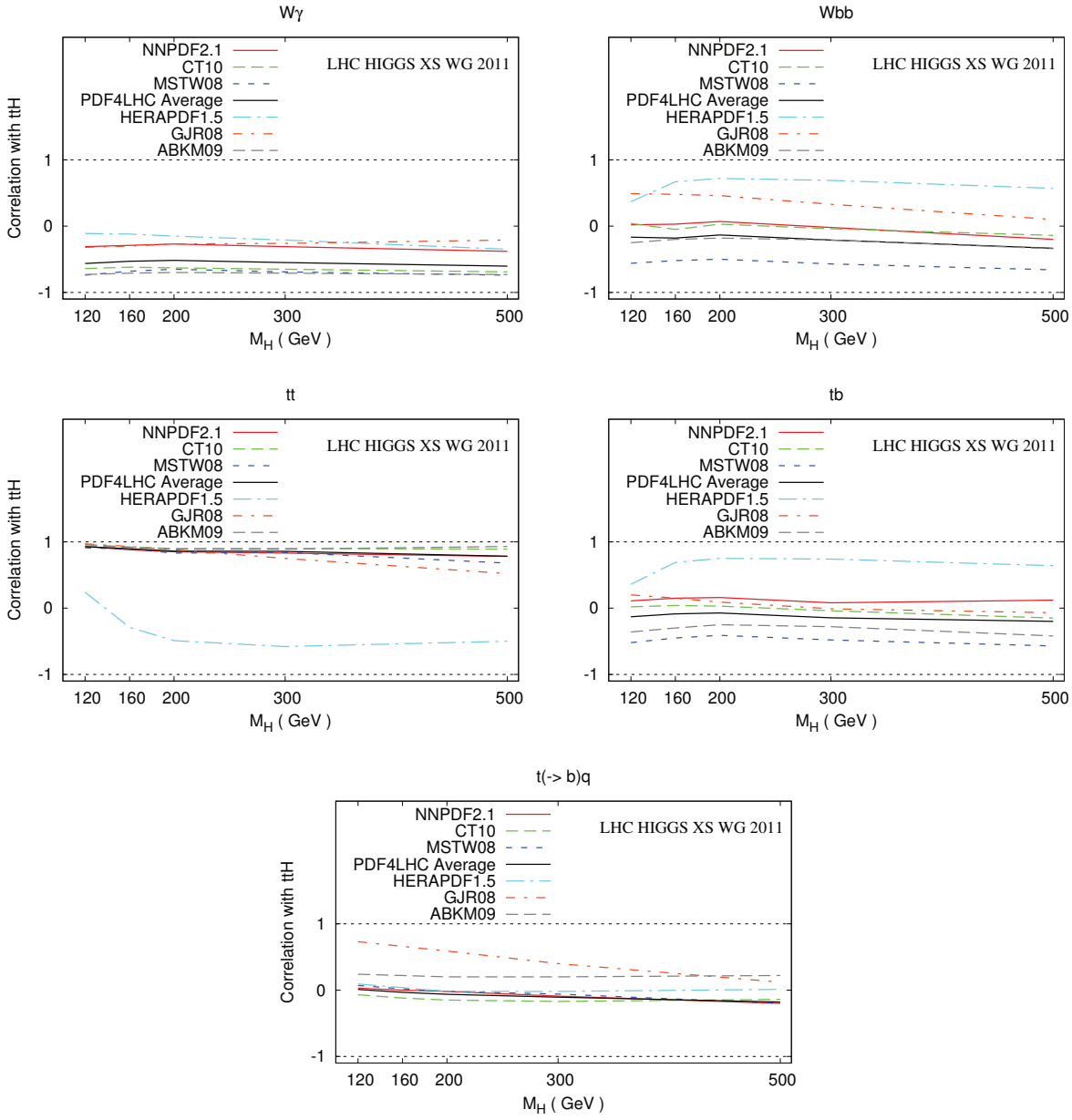

Fig. 21: The same as Figure 15 for the $\mathrm{t\bar{t}H}$ production process.