Parton Distributions - Correlations for Higgs Cross Sections and Backgrounds

Robert Thorne

November 28th, 2011

University College London

Thanks to ABM, CT10, GJR, HERAPDF, MSTW, NNPDF for correlations results

PDF Correlations

The PDF uncertainty analysis may be extended to define a *correlation* between the uncertainties of two variables, say $X(\vec{a})$ and $Y(\vec{a})$.

The correlations were calculated using the MCFM NLO program (versions 5.8 and 6.0) with a common set of input files for all groups. Each group did their own calculations.

For the groups using a Hessian approach the correlations were calculated using

$$\cos\varphi = \frac{\vec{\Delta}X \cdot \vec{\Delta}Y}{\Delta X \Delta Y} = \frac{1}{4\Delta X \Delta Y} \sum_{i=1}^{N} \left(X_i^{(+)} - X_i^{(-)}\right) \left(Y_i^{(+)} - Y_i^{(-)}\right)$$

$$\Delta X = \left|\vec{\Delta}X\right| = \frac{1}{2}\sqrt{\sum_{i=1}^{N} \left(X_i^{(+)} - X_i^{(-)}\right)^2}$$

or some similar variation. This included the most up-to-date published sets for each group, i.e. , ABKM09, CT10, GJR 08, MSTW08. The basic results for CT10 and MSTW08 are PDF only, whereas ABKM09 and GJR08 include α_S as a parameter in the error matrix.

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

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

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

The correlations for the NNPDF prescription are calculated using

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

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

The averaging was done and diagrams made by J. Rojo.

Backgrounds

	Z	W	ZZ	WW	WZ	Wγ	WQQ	ZQQ	ggWW	ggZZ	ttbar	tW	tb	tbq
Z	1	0.95	0.67	0.70	0.95	0.9	0.43/0.53	0.08	-0.67	-0.75	-0.74	-0.81	0.59	-0.29
W	0.95	1	0.52/0.69	0.60/0.71	0.88/1.0	0.90/0.80	0.39/0.50	0.08	-0.67	-0.74	-0.73	-0.8	0.57	-0.29
ZZ	0.67	0.52/0.69	1	0.97	0.54/0.73	0.62	0.78/0.87	-0.09	-0.36	-0.34	-0.17	-0.81	0.9	-0.23
WW	0.70	0.60/0.71	0.97	1	0.63/0.75	0.69	0.80/0.86	-0.02	-0.34	-0.33	-0.20	-0.33	0.94	-0.08
WZ	0.95	0.88/1.0	0.54/0.73	0.63/0.75	1	0.9	0.55	0.1	-0.64	-0.71	-0.71	-0.73	0.61	-0.34
Wγ	0.9	0.90/0.80	0.62	0.69	0.9	1	0.63/0.53	0.32	-0.44	-0.54	-0.68	0.61	0.61	0
WQQ	0.43/0.53	0.39/0.50	0.78/0.87	0.80/0.86	0.55	0.63/0.53	1	0.08	-0.12	-0.12	-0.05	-0.15	0.64	-0.32
ZQQ	0.08	0.08	-0.09	-0.02	0.1	0.32	0.08	1	0.54	0.36	-0.26	-0.05	-0.03	0.59
ggWW	-0.67	-0.67	-0.36	-0.34	-0.64	-0.44	-0.12	0.54	1	0.98	0.65	0.81	-0.28	0.63
ggZZ	-0.75	-0.74	-0.34	-0.33	-0.71	-0.54	-0.12	0.36	0.98	1	0.79	0.91	-0.27	0.55
ttbar	-0.74	-0.73	-0.17	-0.20	-0.71	-0.68	-0.05	-0.26	0.65	0.79	1	0.97	-0.12	0.17
tW	-0.81	-0.8	-0.81	-0.33	-0.73	0.61	-0.15	-0.05	0.65	0.91	0.97	1	-0.25	0.31
tb	0.59	0.57	0.9	0.94	0.61	0.61	0.64	-0.03	-0.28	-0.27	-0.12	-0.25	1	0.04
tbq	-0.29	-0.29	-0.23	-0.08	-0.34	0	-0.32	0.59	0.63	0.55	0.17	0.31	0.04	1

Full study involves range of backgrounds shown by Huston at PDF4LHC- July 2011. Will be found at

https://twiki.cern.ch/twiki/bin/view/LHCPhysics/PDFCorrelations

m_H=120

	ggH	VBF	WH	ZH	ttH	Z	W+/W-	ZZ	WW	WZ	Wγ	WQQ	ZQQ	ggWW	ggZZ	ttbar	tW	tb	tbq
ggH	1	-0.57	-0.23	-0.14	-0.6	0.01	0.03	0.02	-0.20	0.04	0.23	-0.14	0.95	0.47	0.28	-0.35	-0.12	-0.24	0.52
VBF	-0.57	1	0.63/0.73	0.76	0.09	0.43	0.26/0.41	0.79	0.72	0.28/0.43	0.28/0.37	0.52/0.71	-0.41	-0.47	-0.4	-0.10	-0.28	0.65	-0.25
WH	-0.23	0.63/0.73	1	0.93	0	0.62	0.52/0.64	0.92	0.93	0.65/058	0.65/0.56	0.79/0.95	-0.02	-0.29	-0.28	-0.15	-0.28	0.99/0.77	0.05/-0.30
ZH	-0.14	0.76	0.93	1	0.03	0.64	0.53/0.66	0.99	0.99	0.55/0.71	0.63	0.83	-0.07	-0.31	-0.3	-0.14	-0.28	0.93	-0.14
ttH	-0.6	0.09	0	0.03	1	-0.61	-0.6	0	-0.05	-0.58	-0.64	0.04	-0.5	0.03	0.56	0.94	0.84	0.02	-0.07

m_H=160

	ggH	VBF	WH	ZH	ttH	Z	W+/W-	ZZ	WW	WZ	Wγ	WQQ	ZQQ	ggWW	ggZZ	ttbar	tW	tb	tbq
ggH	1	-0.61	-0.29	-0.35	-0.24	-0.32	-0.32	-0.35	-0.29	-0.29	-0.06	-0.12	0.9	0.82	0.68	0.1	0.33	-0.27	0.67
VBF	-0.61	1	0.62	0.74	0.2	0.35	0.19/0.34	0.75	0.66	0.20/0.36	0.19/0.28	0.46/0.70	-0.47	-0.46	-0.37	-0.03	-0.22	0.6	-0.29
WH	-0.29	0.62	1	0.93	0.1	0.55	0.52	0.9	0.93	0.56	0.56	0.93	-0.07	-0.26	-0.23	-0.07	-0.21	1	0.03
ZH	-0.35	0.74	0.93	1	0.16	0.54	0.43/0.58	0.98	0.97	0.45/0.63	0.52	0.93	-0.14	-0.29	-0.25	-0.04	-0.2	0.91	-0.16
ttH	-0.24	0.2	0.1	0.16	1	-0.59	-0.58	0.03	-0.03	-0.56	-0.62	-0.05	-0.54	0.33	0.51	0.92	0.8	0.04	-0.12

m_H=200

	ggH	VBF	WH	ZH	ttH	Z	W+/W-	ZZ	WW	WZ	Wγ	WQQ	ZQQ	ggWW	ggZZ	ttbar	tW	tb	tbq
ggH	1	-0.5	-0.26	-0.3	0.13	-0.59	-0.59	-0.36	-0.32	-0.55	-0.33	-0.11	0.68	0.98	0.93	0.5	0.69	-0.27	0.67
VBF	-0.5	1	0.60/0.73	0.72	0.26	0.28	0.13/0.28	0.7	0.62	0.15/0.30	0.12/0.20	0.40/0.69	-0.52	-0.44	-0.34	0.02	-0.17	0.55	-0.32
WH	-0.26	0.60/0.73	1	0.92	0.2	0.44	0.44/0.38	0.89	0.86	0.48/0.41	0.47/0.36	0.78/0.74	-0.15	-0.24	-0.2	0	-0.15	0.98/0.69	0
ZH	-0.3	0.72	0.92	1	0.24	0.46	0.34/0.51	0.95	0.93	0.37/0.56	0.43	0.74/0.85	-0.19	-0.3	-0.22	0.02	-0.14	0.88	-0.2
ttH	0.13	0.26	0.2	0.24	1	-0.57	-0.57	0.03	-0.03	-0.55	-0.63	0.03	-0.56	0.29	0.48	0.9	0.78	0.03	-0.15

And similar for signals. However, too detailed for concise presentation when averaging/comparing, and precision much higher than spread between groups.

Full list also not vital since 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 $\mathrm{gg} \to \mathrm{WW(ZZ)}$ and $\mathrm{gg} \to \mathrm{H}$ for $M_H = 200$ GeV.

The up-to-date PDF4LHC average (CT10, MSTW08, NNPDF2.1) 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 of width 0.2.

120 GeV				
	ggH	VBF	WH	$t\bar{t}H$
ggH	1	-0.6	-0.4	-0.2
VBF	-0.6	1	0.6	-0.4
WH	-0.2	0.6	1	-0.2
$t\bar{t}H$	-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
$W\gamma$	0	0.6	0.8	-0.6
$Wb\bar{b}$	-0.2	0.6	1	-0.2
$t\bar{t}$	0.2	-0.4	-0.4	1
$t\bar{b}$	-0.4	0.6	1	-0.2
$t(\to \bar{b})q$	0.4	0	0	0

160 GeV				
	ggH	VBF	WH	$t\bar{t}H$
ggH	1	-0.6	-0.4	0.2
VBF	-0.6	1	0.6	-0.2
WH	-0.4	0.6	1	0
$t\bar{t}H$	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
$W\gamma$	-0.4	0.6	0.6	-0.6
$Wb\bar{b}$	-0.2	0.6	0.8	-0.2
$t\bar{t}$	0.4	-0.4	-0.2	0.8
$t\bar{b}$	-0.4	0.6	1	0
$t(\to \bar{b})q$	0.6	0	0	0

200 GeV				
	ggH	VBF	WH	$t\bar{t}H$
ggH	1	-0.6	-0.4	0.4
VBF	-0.6	1	0.4	-0.2
WH	-0.4	0.6	1	0
$t\bar{t}H$	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
$W\gamma$	-0.4	0.4	0.6	-0.6
$Wb\bar{b}$	-0.2	0.6	0.8	-0.2
$t\bar{t}$	0.6	-0.4	-0.2	0.8
$t\bar{b}$	-0.4	0.6	0.8	0
$t(\rightarrow \bar{b})q$	0.6	-0.2	0	0

300 GeV				
	ggH	VBF	WH	$t\bar{t}H$
ggH	1	-0.4	-0.2	0.6
VBF	-0.4	1	0.4	-0.2
WH	-0.2	0.4	1	0.2
$t\bar{t}H$	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
$W\gamma$	-0.6	0.4	0.4	-0.6
$Wb\bar{b}$	-0.2	0.4	0.8	-0.2
$t\bar{t}$	1	-0.4	0	0.8
$t\bar{b}$	-0.4	0.4	0.8	-0.2
$t(\rightarrow \bar{b})q$	0.4	-0.2	0	-0.2

500 GeV	ggH	VBF	WH	$t\bar{t}H$
ggH	1	-0.4	0	0.8
VBF	-0.4	1	0.4	-0.2
WH	0	0.4	1	0
$t\bar{t}H$	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
$W\gamma$	-0.6	0.4	0.2	-0.6
$Wb\bar{b}$	-0.4	0.4	0.6	-0.4
$t\bar{t}$	1	-0.4	0	0.8
$t\bar{b}$	-0.4	0.4	0.8	-0.2
$t(\rightarrow \bar{b})q$	0.2	-0.2	0	-0.2

Generally the results expected, i.e. gluon dominated processes correlated with each other as are quark dominated processes. Little correlation between the two.

However, see that breakdown of correlation between gluon probed at different x values, e.g $gg \rightarrow H$ for $M_H = 120$ GeV and $t\bar{t}$ since from momentum conservation gluon changes in one place (high x) are compensated by those in another (low x), and the crossing point is between 0.01 and 0.1 but varies slightly between groups.

The same for the correlations between background processes.

	W	WW	WZ	$W\gamma$	$Wb\bar{b}$	$t\bar{t}$	$t\bar{b}$	$t(\rightarrow \bar{b})q$
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
$W\gamma$	1	0.8	0.8	1	0.6	-0.4	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
$t(\rightarrow \bar{b})q$	-0.2	0	0	0	0	0.2	0.2	1

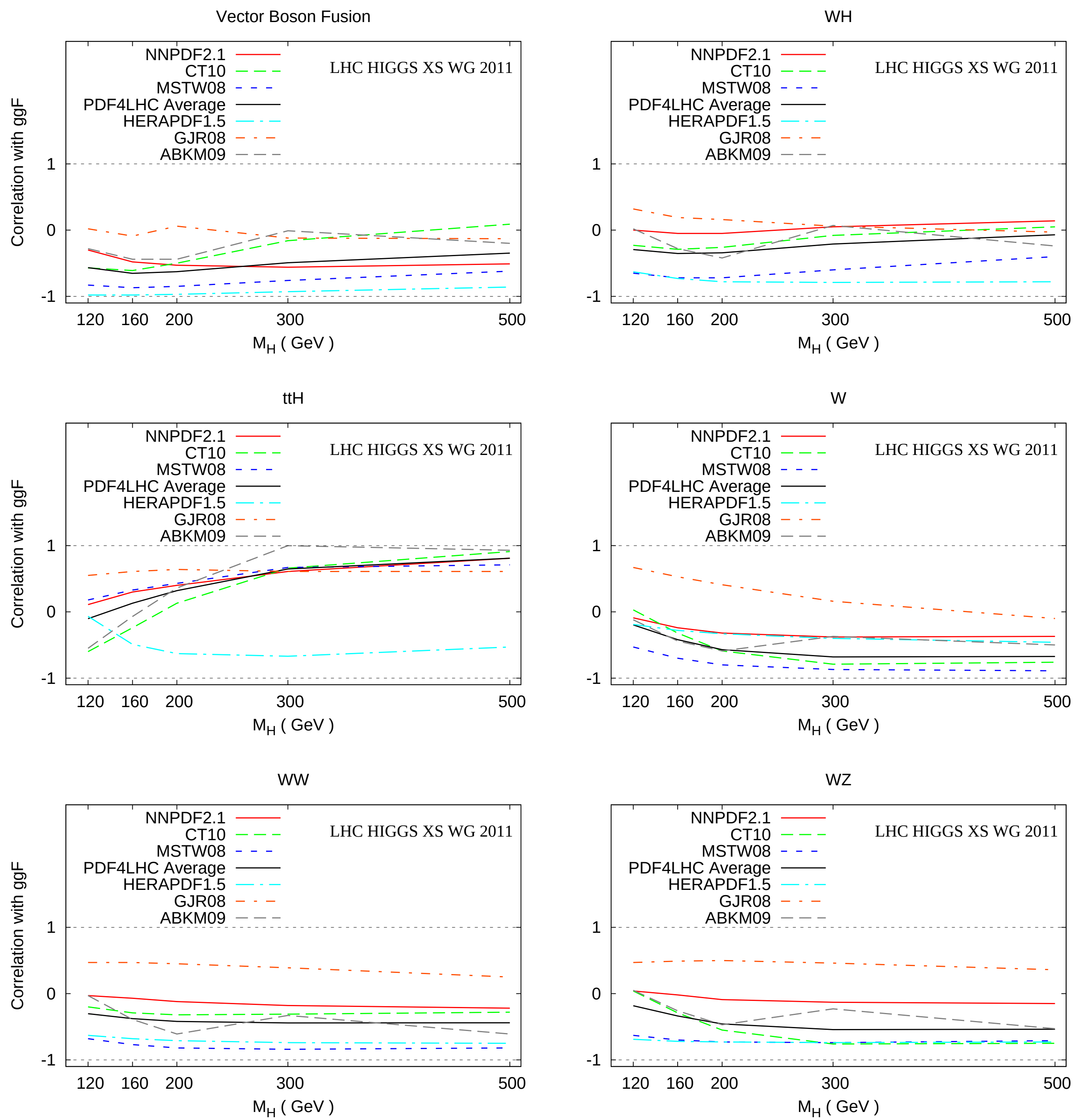
Similar conclusions as for signals.

What about variations between groups?

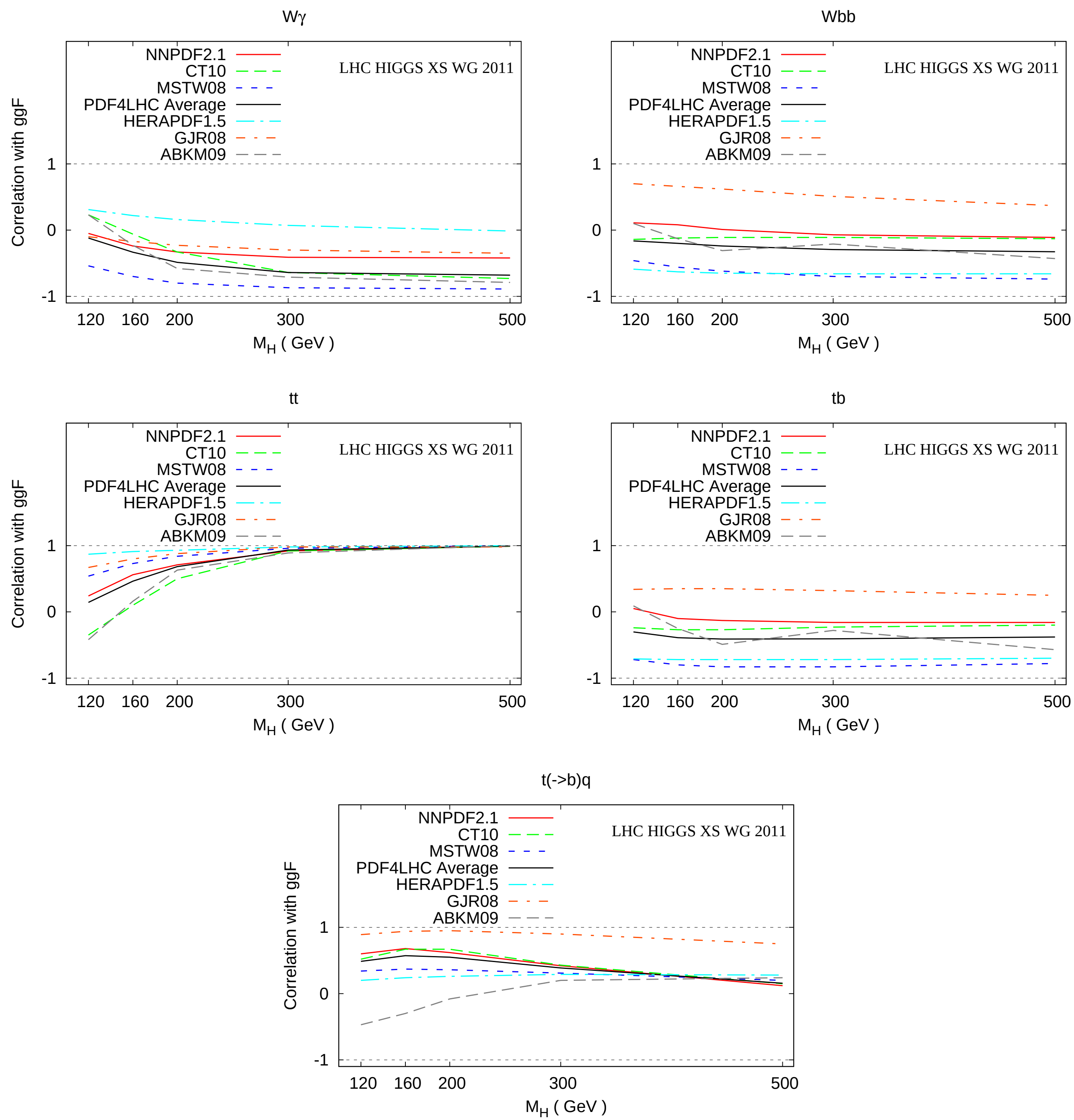
Not a very large amount.

Correlation between the gluon fusion $gg \to \mathrm{H}$ process and the other signal and background processes as a function of M_H.

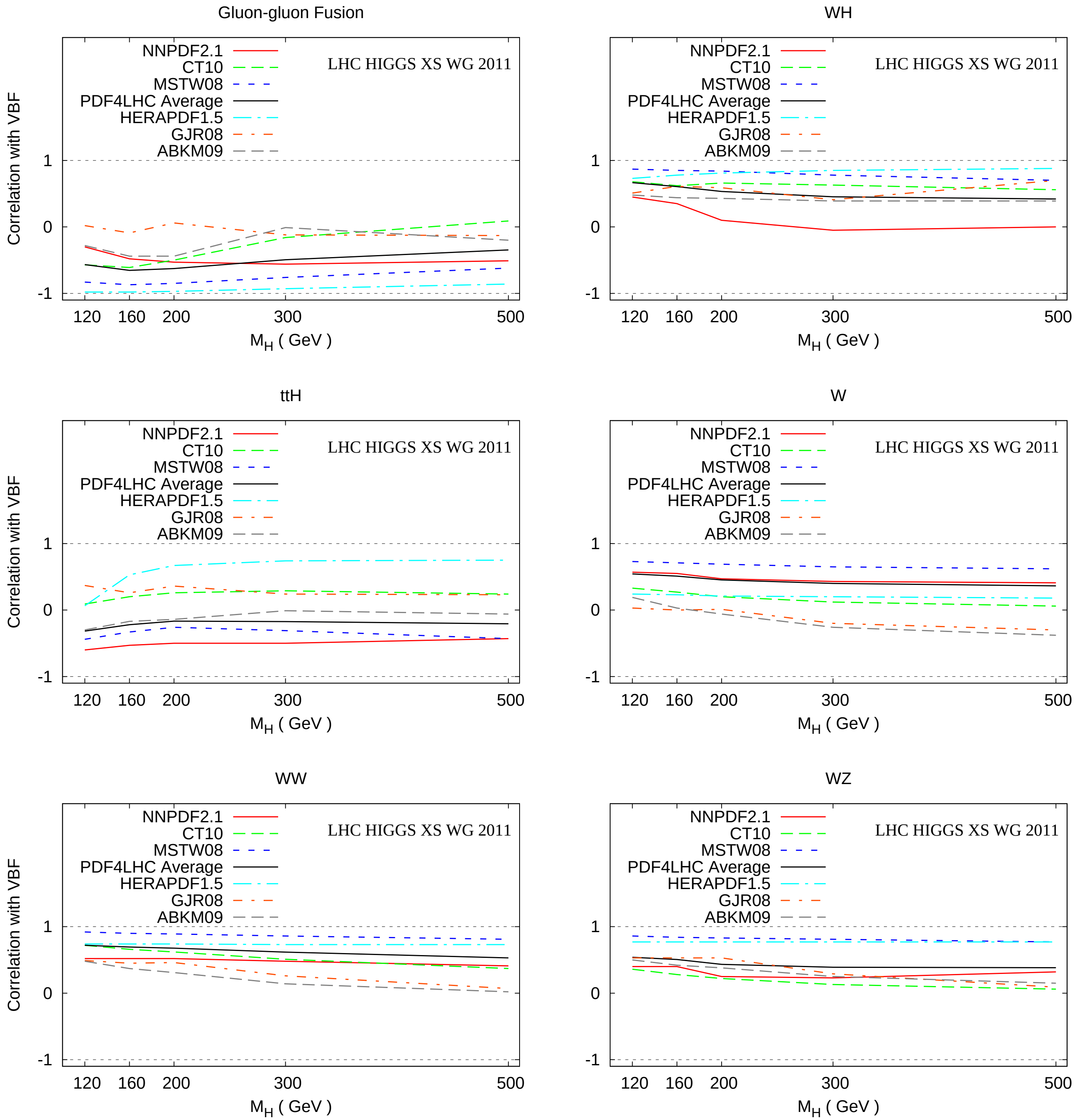
The class width of 0.2 is typical of the scatter of most deviations between groups.



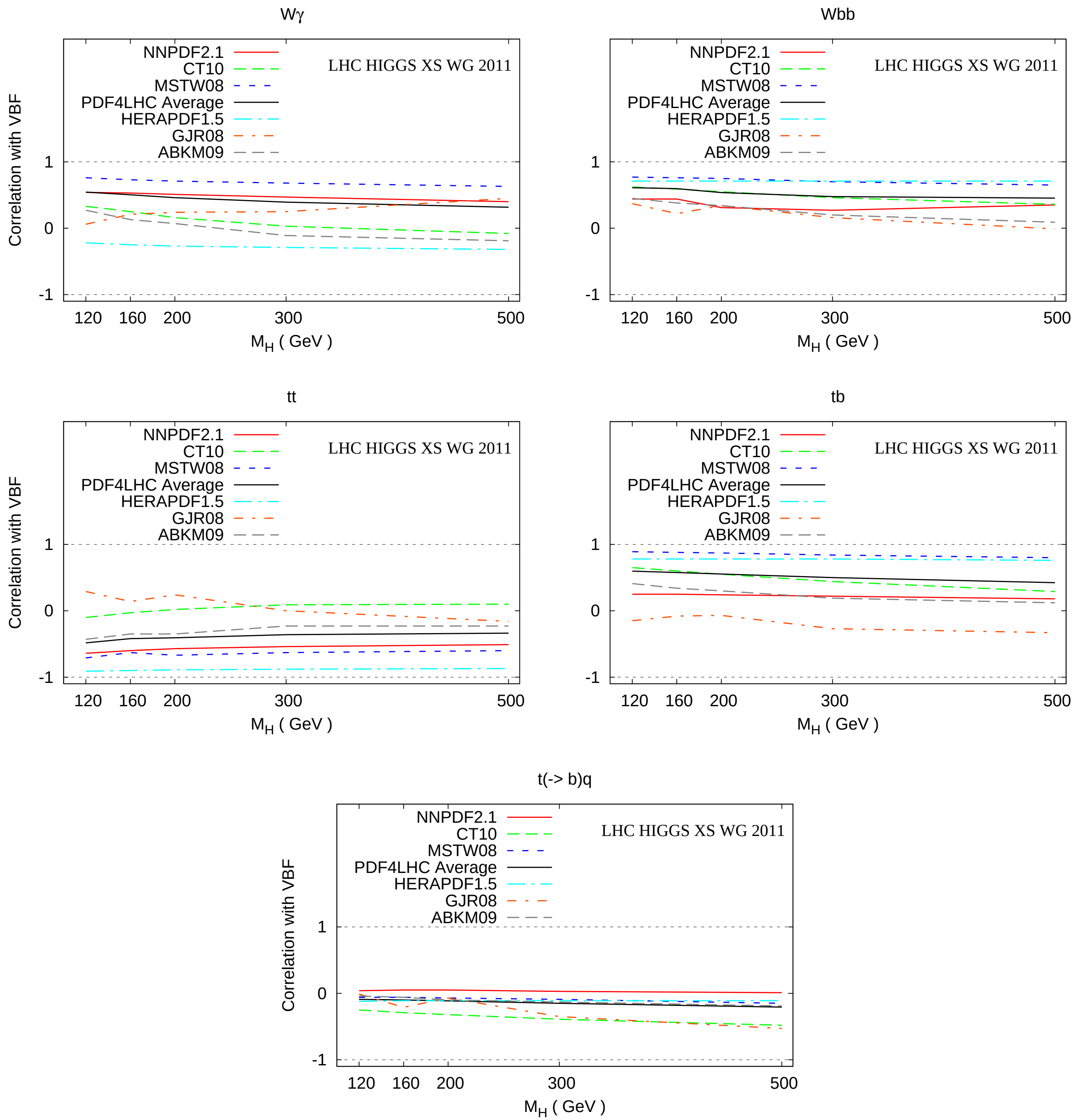
Correlation between the gluon fusion $gg \to H$ process and the other signal and background processes as a function of M_H.



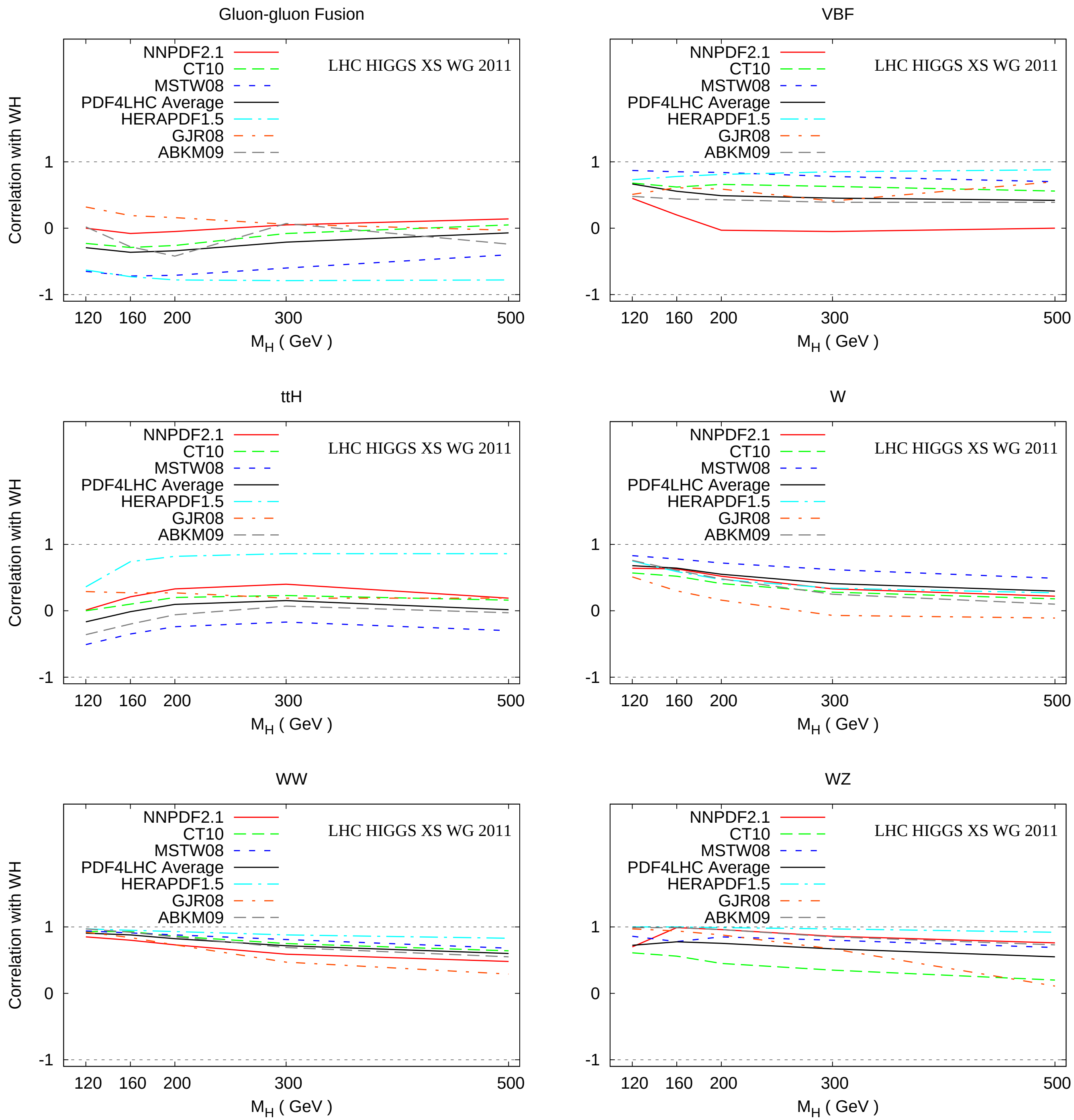
Correlation between the vector boson fusion process and the other signal and background processes as a function of M_H.



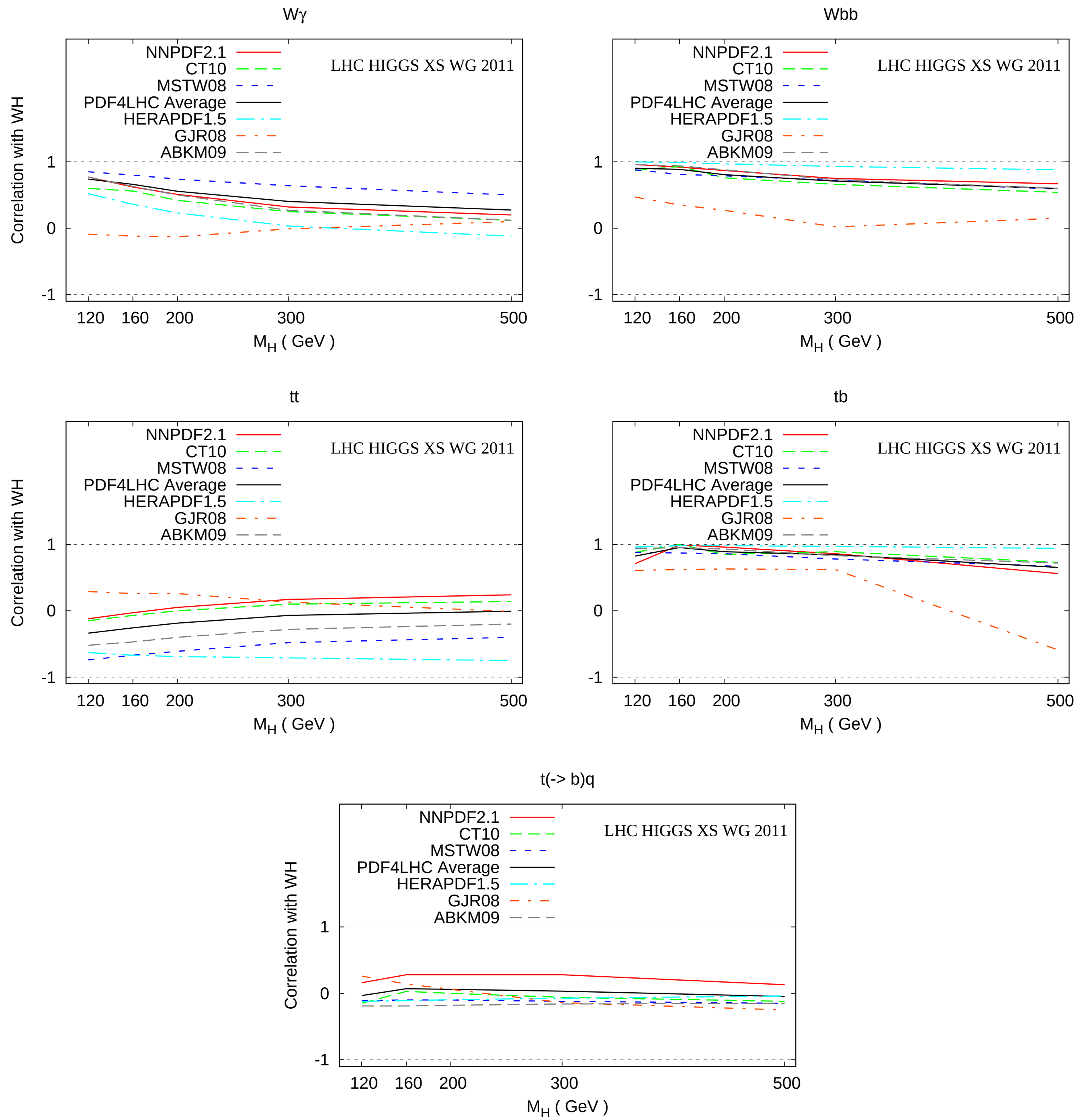
Correlation between the vector boson fusion process and the other signal and background processes as a function of M_H.



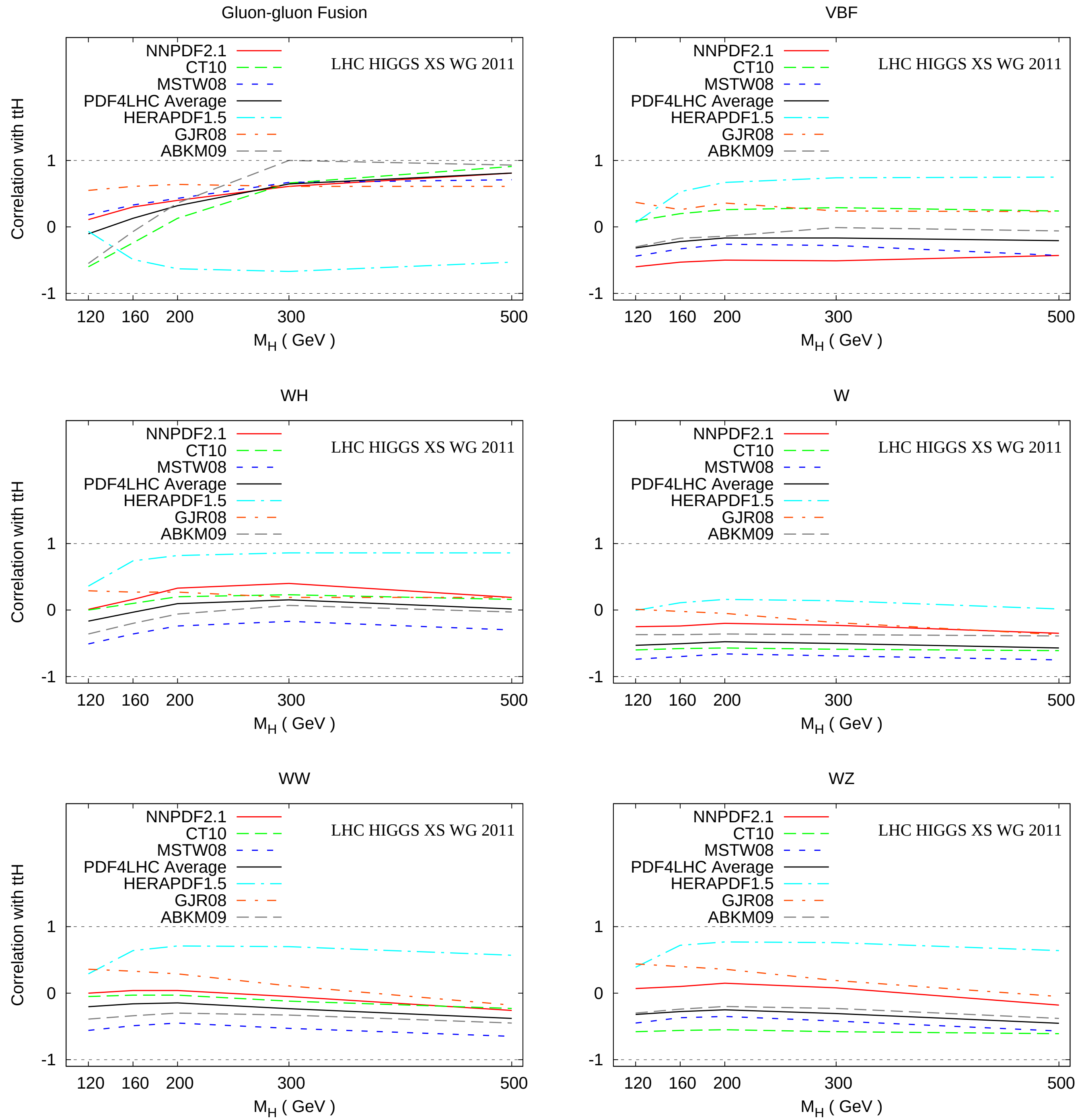
Correlation between the WH production process and the other signal and background processes as a function of M_H.



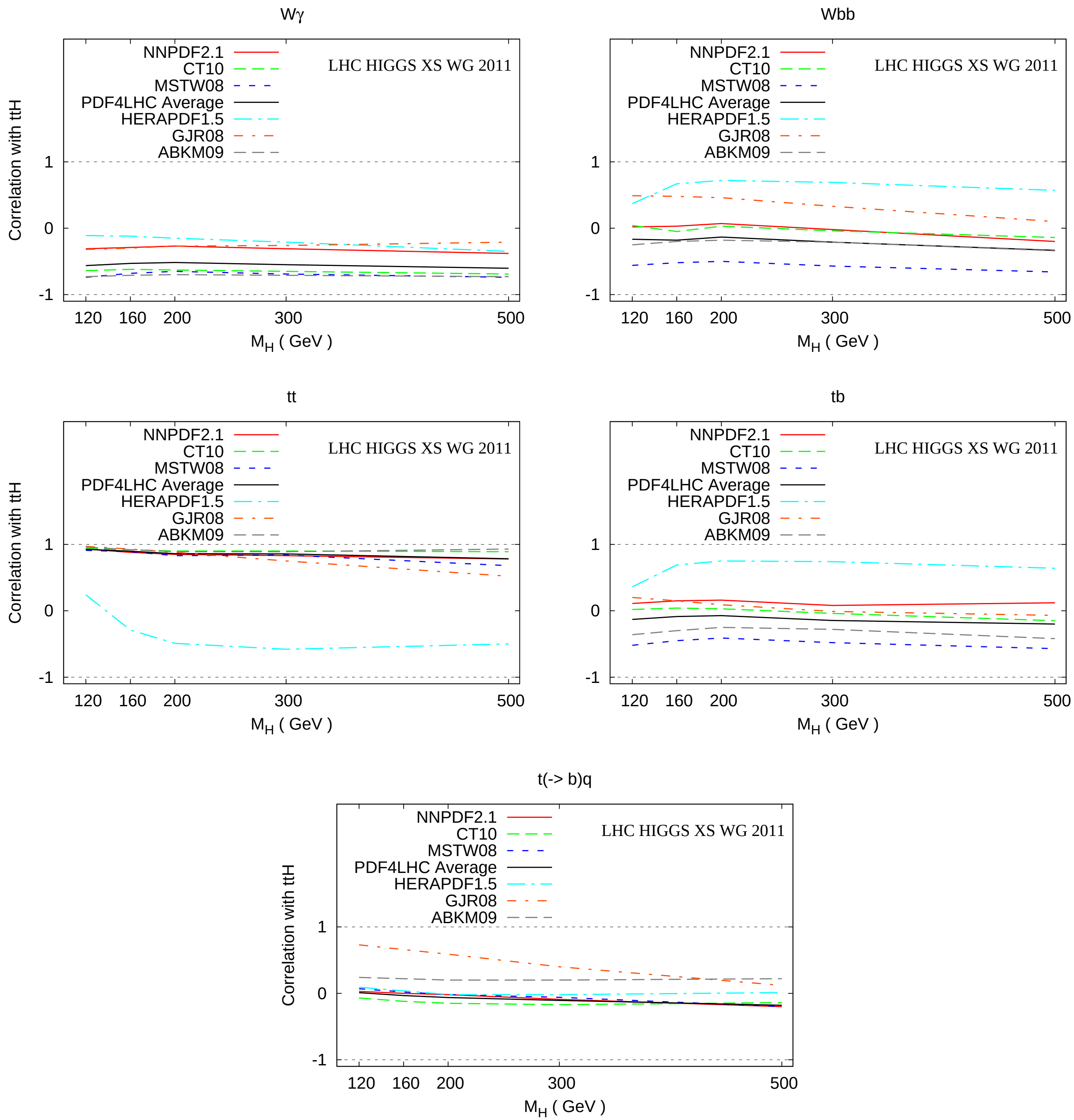
Correlation between the WH production process and the other signal and background processes as a function of M_H.

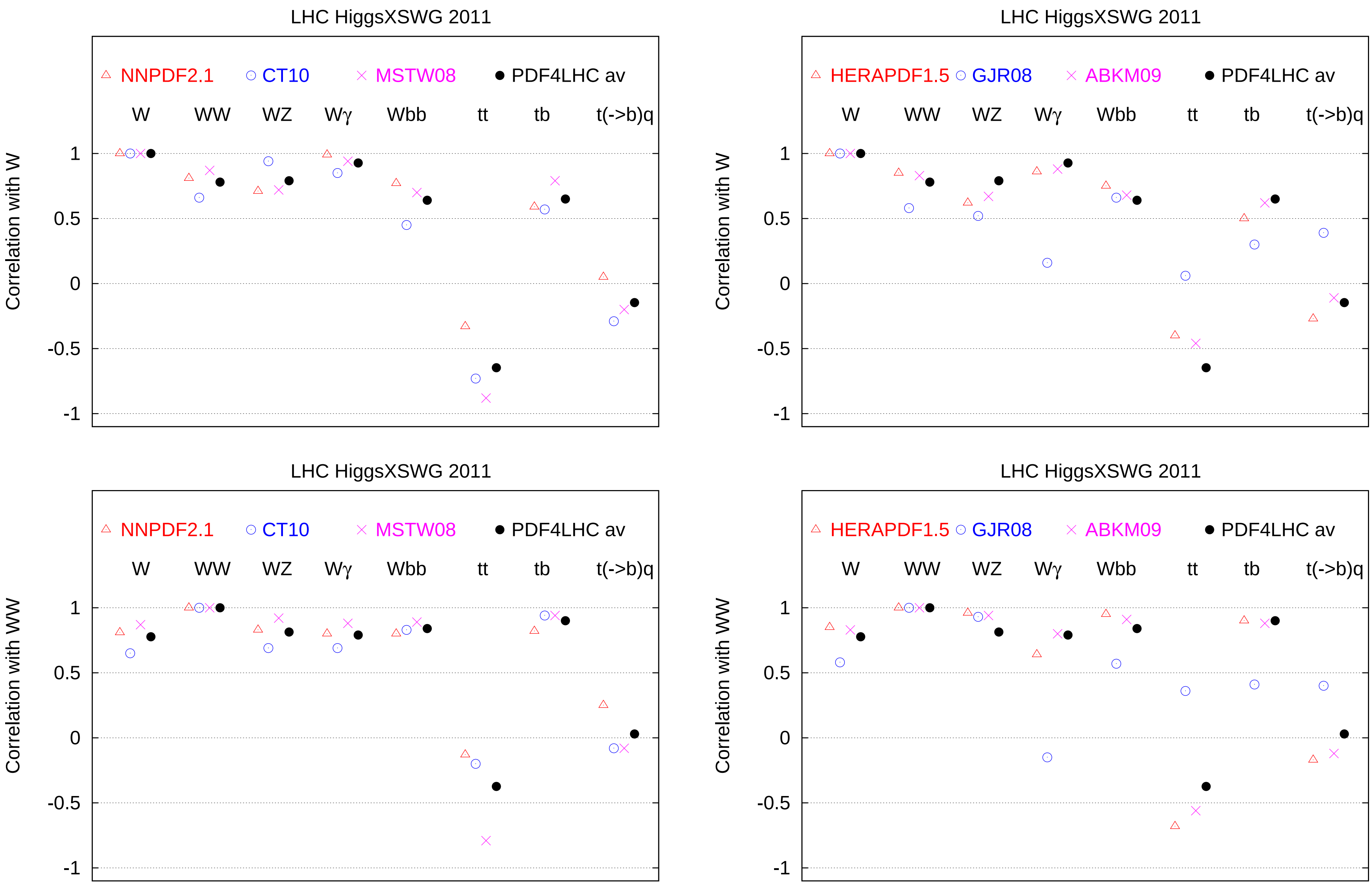


Correlation between the $t\bar{t}H$ production process and the other signal and background processes as a function of M_H.

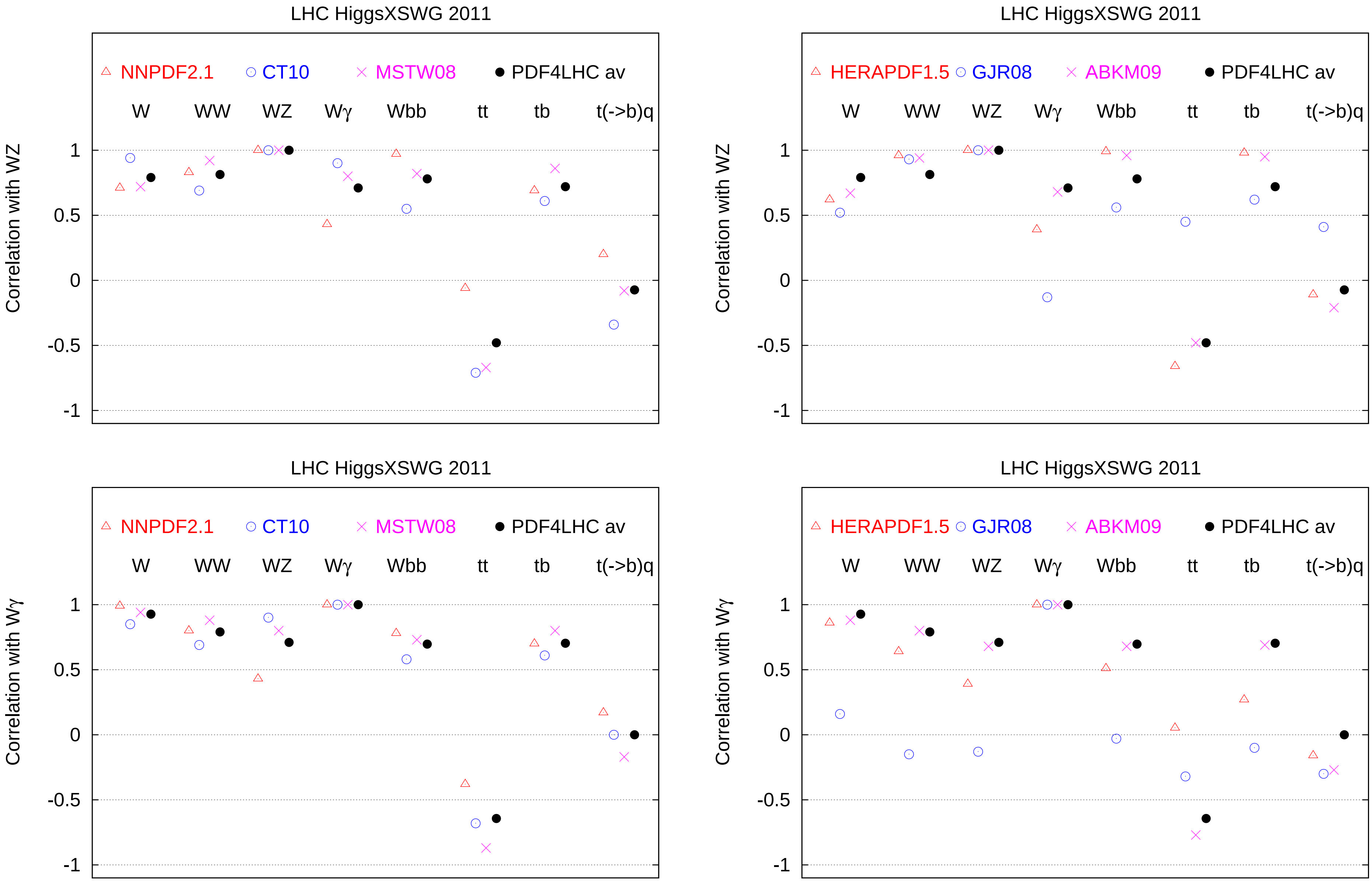


Correlation between the $t\bar{t}H$ production process and the other signal and background processes as a function of M_H.

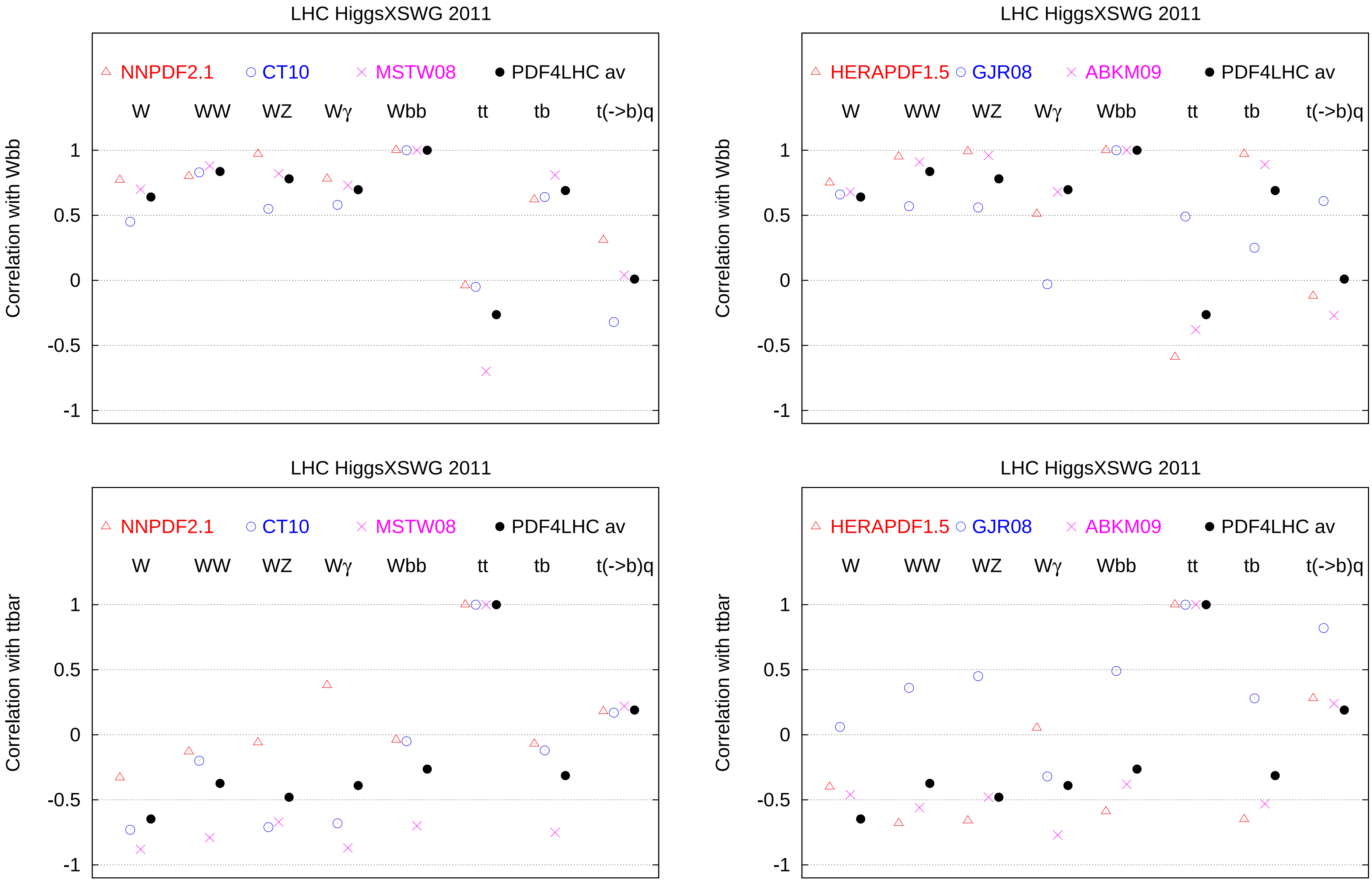




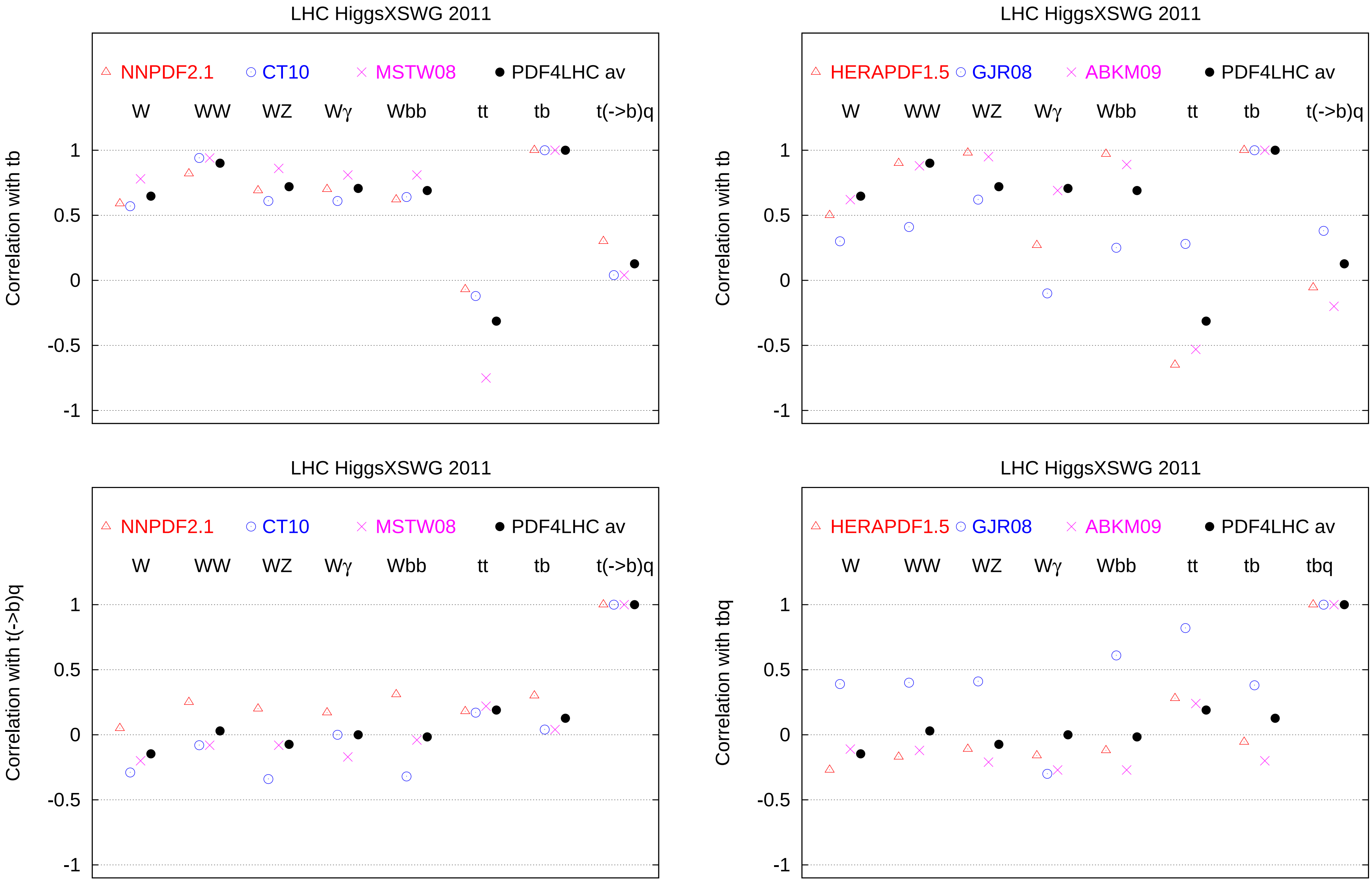
The correlations between W production and WW production and the other background processes considered.



The correlations between WZ production and $W\gamma$ production and the other background processes considered.



The correlations between $\mathrm{Wb\bar{b}}$ production and $\mathrm{t\bar{t}}$ production and the other background processes considered.



The correlations between $\mathrm{t\bar{b}}$ production and $\mathrm{t(\rightarrow \bar{b}) + q}$ production and the other background processes considered.

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 averages using all 6 sets are usually the same as, and nearly always within one class of the PDF4LHC average.

The sets with the largest parameterisations 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-$\mathrm{t\bar{t}H}$.

There are some unusual features, e.g. for HERAPDF1.5 and high values of M_H the $\mathrm{t\bar{t}H}$ correlations with quantities depending on the high-x gluon, e.g. $\mathrm{gg \rightarrow H}$ and $\mathrm{t\bar{t}}$ is opposite to the other sets 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 HERAPDF1.5 (at NLO) which contributes to $\mathrm{t\bar{t}H}$ but not $\mathrm{gg \rightarrow H}$ or very much to $\mathrm{t\bar{t}}$.

GJR08 has a tendency to obtain more correlation between some gluon dominated processes, e.g. $\mathrm{gg \rightarrow H}$ and $\mathrm{t\bar{t}}$ and quark dominated processes, e.g. W and WZ, perhaps because the dynamical generation of PDFs couples the gluon and quark more strongly.

Some additional results

The inclusion of the α_S uncertainty on the correlations, compared to PDF only variation, was also studied for MSTW2008, using the approximation that the PDF sets for the upper and lower α_S values simply form another pair of orthogonal eigenvectors.

This can increase the correlation between some processes, i.e. W production and $\mathrm{gg} \to \mathrm{H}$, 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. $\mathrm{gg} \to \mathrm{H}$ and $\mathrm{t\bar{t}}$ production since each depends directly on α_S.

In a handful of processes it can reduce correlation, e.g. W and $\mathrm{Wb\bar{b}}$ since the latter has a much stronger α_S dependence.

α_S dependence is not a very obvious contributing factor to the differences between groups.

Correlations at NNLO

A small number of correlations were also calculated at NNLO for MSTW2008 PDFs, i.e. W, Z and $\mathrm{gg} \rightarrow \mathrm{H}$ 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 $\mathrm{gg} \rightarrow \mathrm{H}$ having more direct dependence on α_S, but this is a relatively minor effect.

The NLO tables can be used at NNLO with good accuracy.

Summary

No major surprises about which processes are correlated, anti-correlated or uncorrelated in the vast majority of cases.

Generally very good agreement between groups. We think we understand the small number of systematic differences.

Some variation in precisely what values of x are correlated or anti-correlated for the gluon, i.e. effective "fixed-point" varies a little between groups.

Would like the more detailed results, along with any accompanying explanation, to put on the LHC Higgs Cross Section Working Group Twiki before the Yellow Report is published, i.e. before Christmas.