Article

Higgs Combination Plot Update & Some Technical Points about Combining Sigmas

Philip E. Gibbs*

Abstract

ATLAS and CMS have updated their Higgs publications based on last years data. These results were released rapidly at the CERN council meeting in December but since then they have had time to polish the reports and as a bonus CMS have added some new data into the diphoton channels. Here I update my previous unofficial Higgs combinations and discuss some important technical points about combining sigmas.

Key Words: Higgs Combination, update, CERN, LHC, ATLAS, CMS.

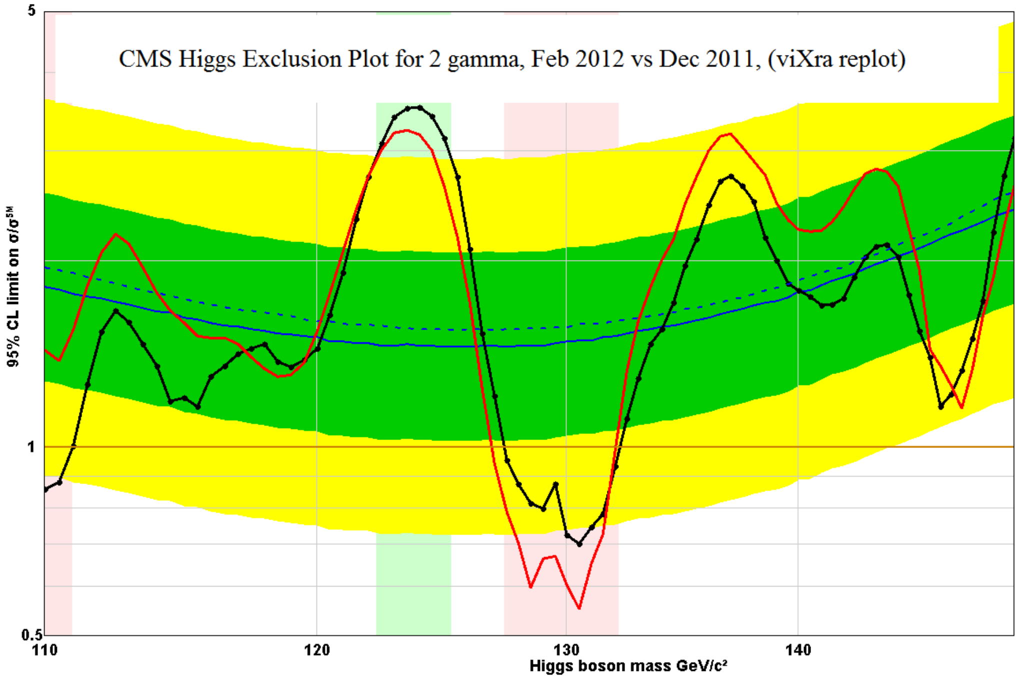
February 9, 2012: Higgs Combination Plot Update

ATLAS and CMS have updated their Higgs publications based on last years data. These results were released rapidly at the CERN council meeting in December but since then they have had time to polish the reports and as a bonus CMS have added some new data into the diphoton channels. This has already been covered nicely on the other blogs including QDS, NEW, TRF, OPS and the best report from Resonaances which gives a nice account of where the extra events come from.

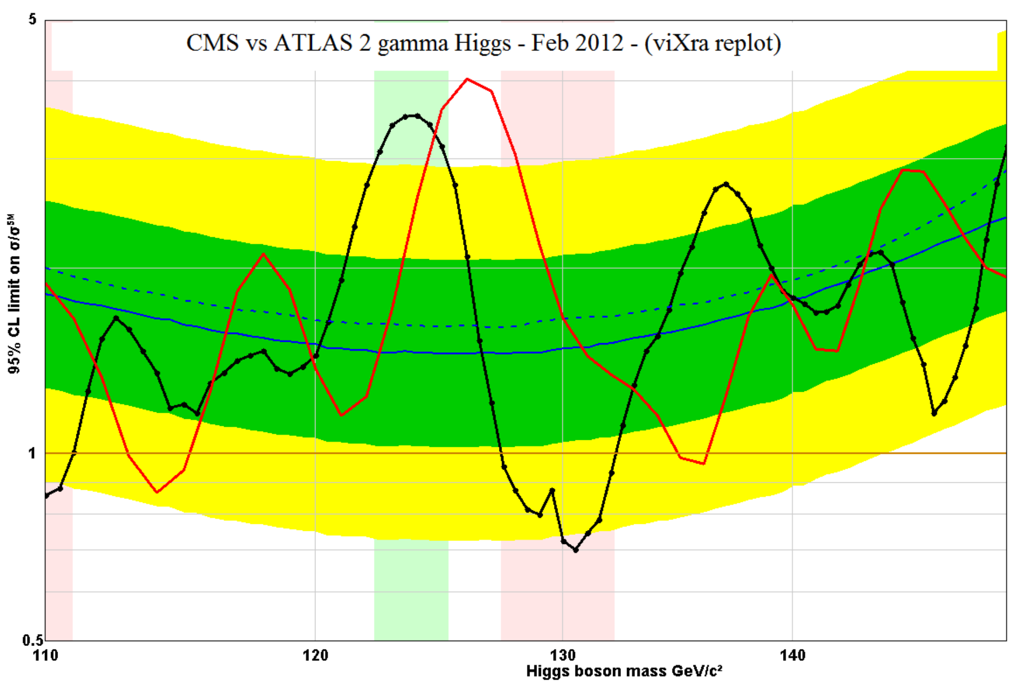
> "*CMS added a new category of events which, apart from 2 photons, contain 2 energetic jets in the forward (closer to the beam pipe) region of the detector. Such events could arise in the so-called vector boson fusion (VBF) process, where each of the 2 colliding quarks emits a W or Z boson which coalesce to create a Higgs boson.*" Jester

The outcome of this is a stronger signal at 124 GeV making it look very similar in strength to the one in ATLAS ar 126 GeV. Here is a plot showing the new CLs for diphoton from CMS with the December version in red for comparison

* Correspondence: Philip E. Gibbs, Ph.D., Independent Researcher, UK. E-Mail: phil@royalgenes.com
Note: This report is adopted from http://blog.vixra.org/2012/02/09/february-higgs-update/ and http://blog.vixra.org/2012/02/10/some-technical-points-about-combining-sigmas/



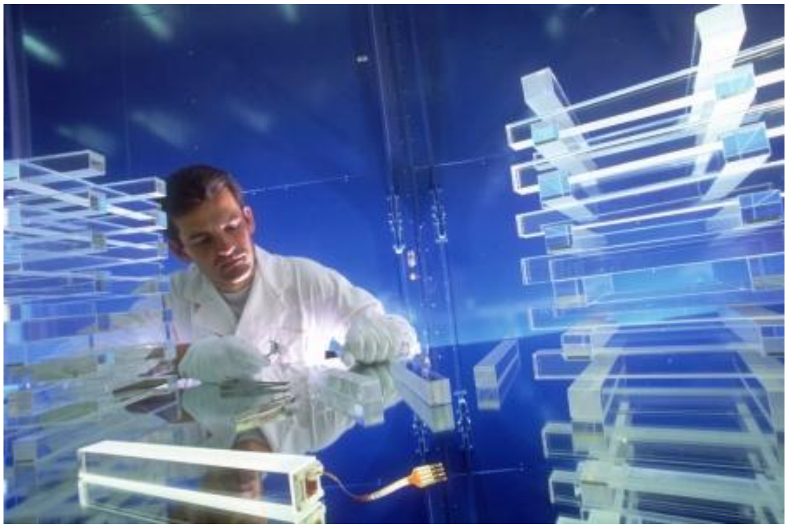
As you can see, the peak at 124 GeV is stronger while the other bumps have gone down. I have plotted this on a log CLs scale rather than the linear scale that CMS use because you need a log scale to read the significance properly. You can see that it has now gone up to 2.5 sigma for this channel alone. For comparison here is the same plot but with the results from ATLAS for comparison (which has not changed since December).



If ATLAS looks better don't forget that its expected value (the dotted blue line) is also higher so in fact the statistical significance is now about the same for both.

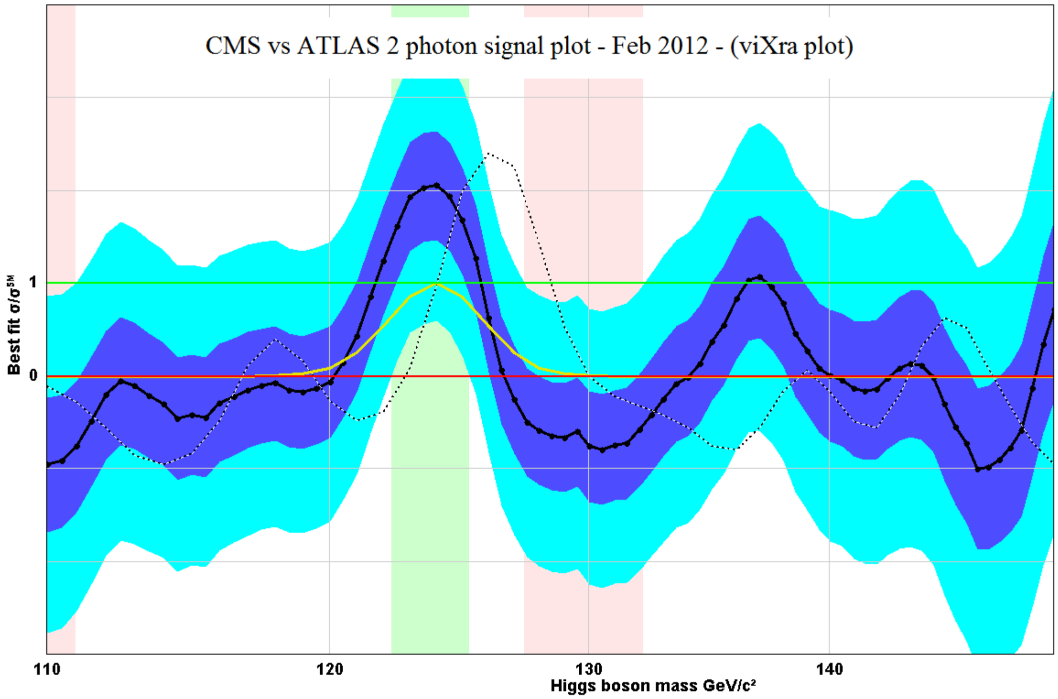
One mystery is why the peaks are about 2 GeV apart. This could simply be a statistical deviation or it may be a sign that they still have some work to do on calibrating the calorimeters which measure the energies of the photons. The two experiments have different systems for detecting these photons. ATLAS use a tank of liquid argon, a clear liquid that looks like this.

CMS on the other had uses some Lead Tungstate crystals which look like glass but are as heavy as lead. See this old article from Symmetry Breaking for the remarkable story behind them.



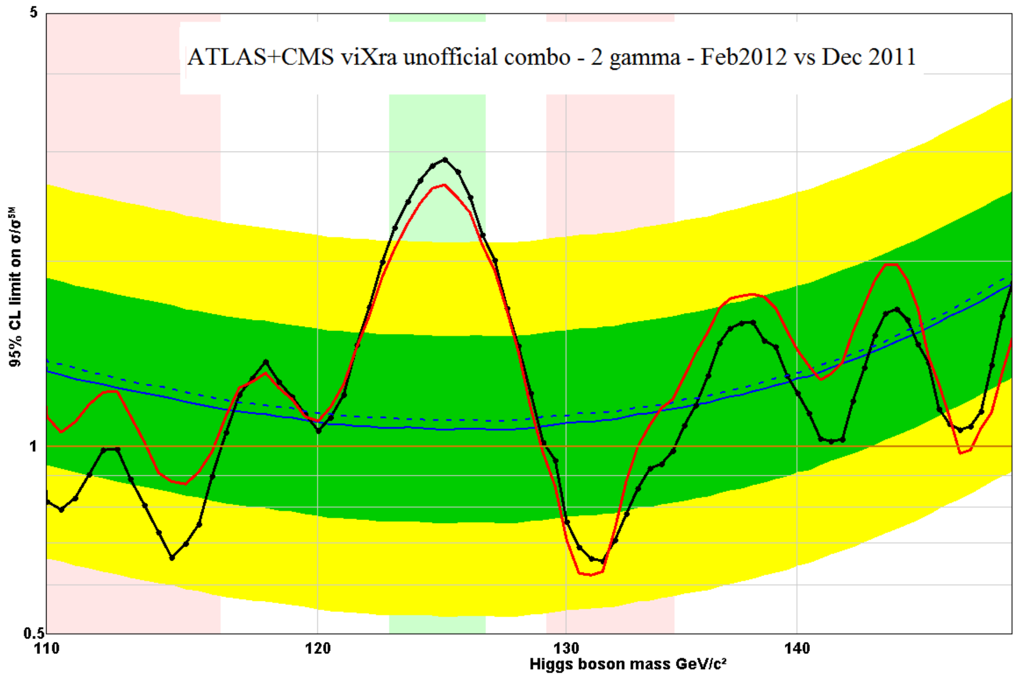
When the CMS detector is finally laid to rest in about 40 years time they will be able to take the crystals out and make a great chandelier out of them as a memorial.

Another mystery about the diphoton results is that the strength of the signal is about twice as strong as expected. This can be seen clearly in this signal plot.



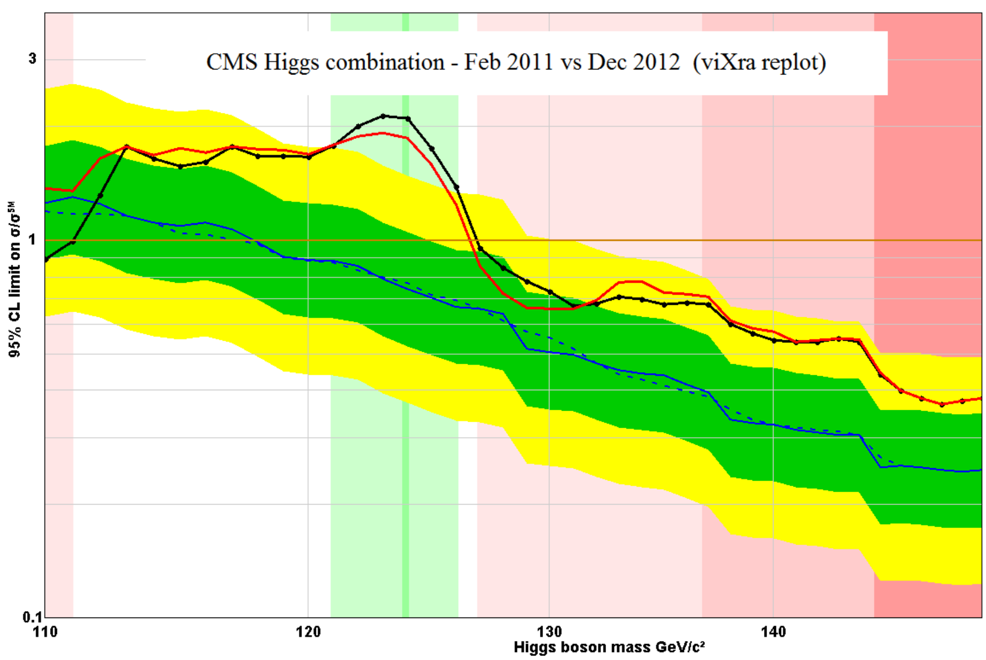
The excess over the standard model Higgs signal is about 1.5 Sigma for each experiment. It could be a sign that the particle is not as standard as expected, but more probably this is just a fluctuation that will go away with more data. It does mean that we should be cautious about how much we should expect the signal to improve if we double the data. It wont be as much as you might think and in fact the signal could get worse. This is why you should not expect conclusive results until all this years data is in.

The unofficial Higgs combinations now need updating and here is the new ATLAS+CMS combo for the diphoton channel with the December version in red for reference.

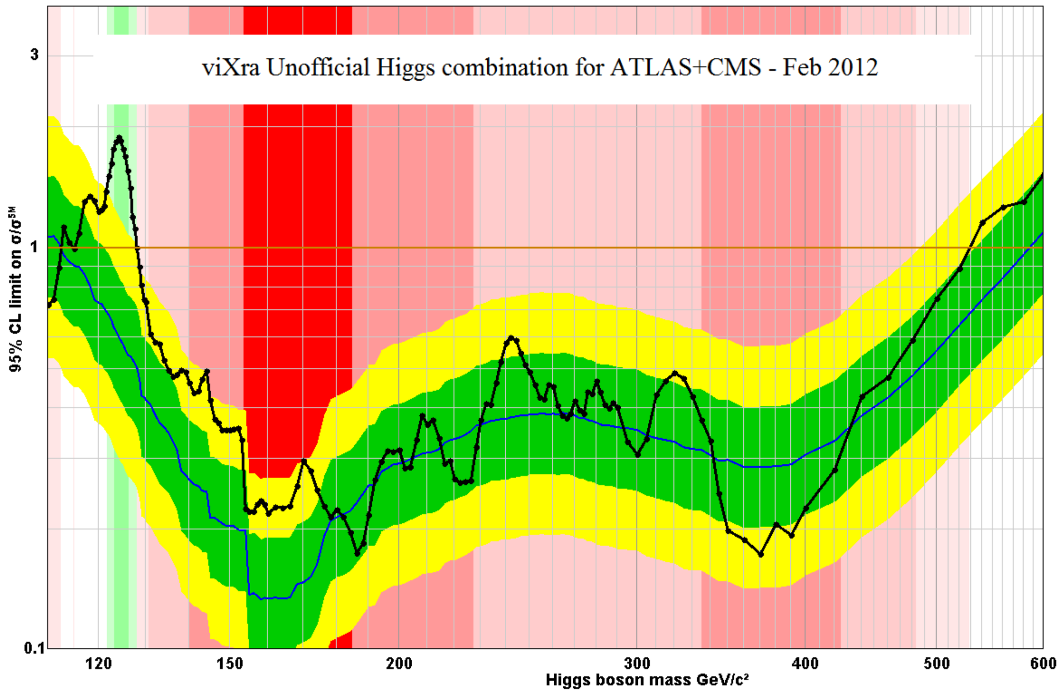


The significance is not as good as it would be if the two excesses coincided better in mass but it is now almost 3 sigma.

The CMS all channel combination has also been updated of course. Here is a replot to show how it has changed vs the December levels in red. there is no change at higher masses.



The significance on this plot has now just passed 3-sigma at the peak of the excess. Which means that the unofficial Higgs combination needs to be updated so here it is



The peak significance here has now reached about 3.5 sigma.

I should remind you that this is an unofficial approximate combination that ignores correlations and un-normal probability distributions. The official version from the Higgs combination group should be due out soon. It will be similar but the differences are important.

February 10, 2012: Some Technical Points about Combining Sigmas

In the latest reports ATLAS is claiming 3.5σ local significance for their combined plot and 2.2σ global significance after 'Look Elsewhere Effect'. For the diphoton channel alone they have 2.8σ local significance and 1.5σ global significance. Meanwhile over at CMS the figures are 3.1σ local significance and 2.1σ global significance for the combined plot, and for the diphoton channel they have also 3.1σ global significance and 2.1σ after LEE.

Now everyone wants to combine these numbers. How can that be done and what is the answer? Concentrating on the combined plots for the moment, a common method is just to add them in quadrature

$$s = \sqrt{s_1^2 + s_2^2}$$

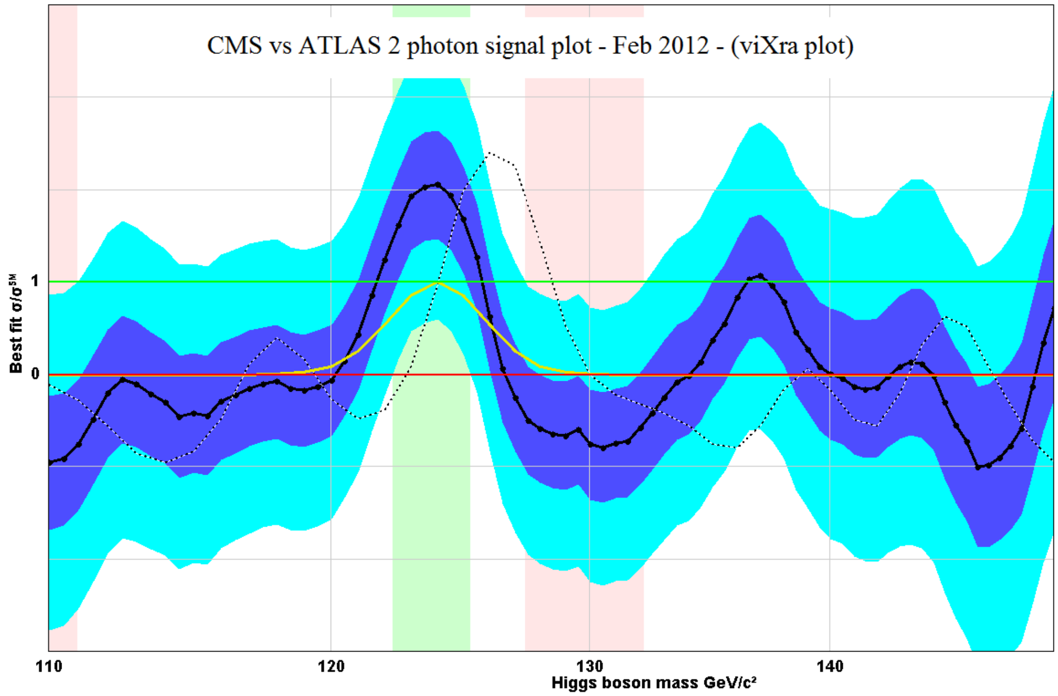
giving $\sqrt{3.5^2 + 2.8^2} = 4.5$ for the combined local significance and $\sqrt{2.2^2 + 1.5^2} = 2.7$ for the global significance. Is this correct?

No that is wrong. Look elsewhere effect must be applied after combining.
The global significance is wrong because we have combined two results with LEE already applied. We should combine the local sigmas first and then apply LEE again. Well LEE is a subjective observer dependent quantity that nobody agrees about how to apply so lets just look at the local significance and let everyone estimate their own LEE afterwards. So are we correct for the global significance?

That is wrong too. The observed excesses were not in the same place.
It's a good point. We can only combine the excesses at the same mass and the peaks of the excess differ by 1 or 2 GeV. If we do this we will get a smaller answer, but is that fair? The difference could be due to a systematic calibration error in one or other of the experiments. In fact this is looking increasingly likely as more data is added and the peaks do not get closer. We will have a much better idea about that when the data is doubled by the summer. So let's be optimistic and just assume that the peaks will nearly coincide after recalibration. In that case we still have 4.5σ. Have we got it right now?

It is still wrong. Combining the numbers in quadrature is not the right formula.
If you think combining sigmas in quadrature is right, or even just approximately right, consider this scenario. Suppose in the first run of data I get a 2 sigma excess at some mass, but it is really just a statistical fluctuation. When the data is double we expect no excess at that mass so combine in quadrature to get $\sqrt{2^2 + 0^2} = 2$ so the excess is still two sigma. Double the data again and we might get a 1 sigma excess so the total significance is $\sqrt{2^2 + 1^2} = 2.2$, even if we double the data again and get a deficit of one sigma below expected we add in quadrature to get $\sqrt{2.2^2 + (-1)^2} = 2.4$ So if you believe that sigmas are added in quadrature you must believe that no excess can ever get smaller as more data is added. In fact they will probably grow like a random walk everywhere. This is obviously rubbish.

The reason for this is that these numbers are not error estimates that are normally added in quadrature. Have a look at this signal plot for the CMS and ATLAS signals in the diphoton channel



The excess for CMS is about 2.1 times the standard model signal plus or minus 0.6. For ATLAS it is 2.4 ± 0.7. These are not sigmas. Those would be given by the size divided by the error so 2.1/0.6 = 3.5 sigma for CMS and 2.4/0.7 = 3.4 sigma for ATLAS (not quite right but I'll come back to that). If we assume flat normal distributions then figures like these have to be averaged weighting by the errors. It is those errors which are combined in quadrature. For equal size data sets the errors should normally be the same which means that the correct formula for combining the sigmas is actually

$$s = \frac{s_1 + s_2}{\sqrt{2}}$$

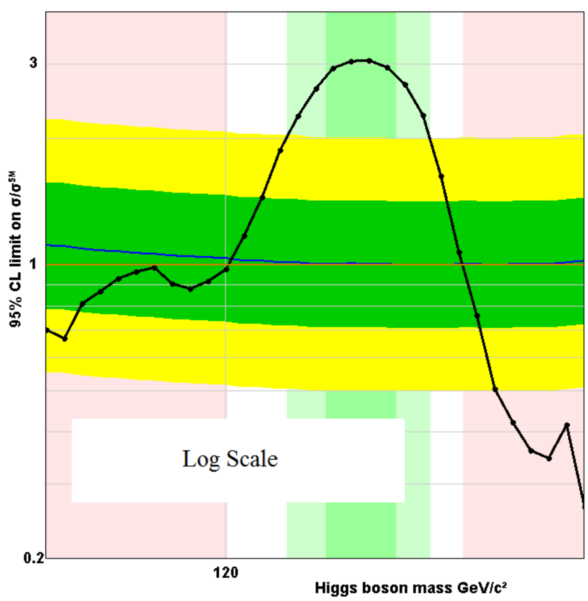
So redoing the calculation we get $(3.5 + 2.8) \times 0.707 = 4.5$, the same answer. In fact if the two sigma levels are similar this formula gives an answer very close to what you would get by adding in quadrature, so why should we care?

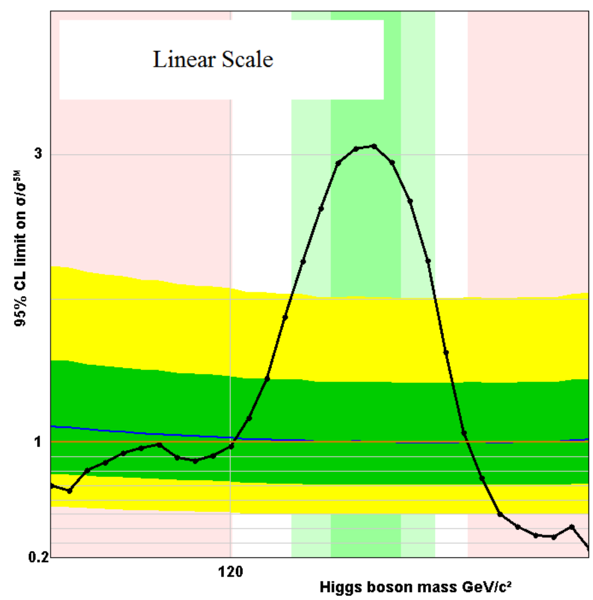
The present excesses in the diphoton channel are larger than predicted by the standard model with the Higgs boson at that mass. This excess is not as big as the excess over the standard model with no Higgs boson. It could be a sign that something non-standard is at work, but let's assume that it is just a statistical fluctuation. In that case when we double the data we expect to get just the standard model signal for the second half of the data. In that case the signal next time will be given by $(2.1 + 1.0) \times 0.707 = 2.2$ for CMS and $(2.4 + 1.0) \times 0.707 = 2.4$ for ATLAS. In other words if the excess is due to a standard model Higgs boson then we should not expect much increase when the data is doubled. Don't get your hopes up for a discovery by the summer. In fact the size of the signal in diphoton could easily go down. Even with quadruple the data it may not grow much bigger. Hopefully the combination with ZZ and WW will fare better because we have not seen the same over-excess in those channels and they will have a discovery by the end of the year, but don't bank on it unless you think the over-excess is a real non-standard effect.

So do we have the right number of sigmas yet?

It is still wrong. You forgot the systematic correlations and have produced NONSENSE.
OK, but get serious. The previous unofficial combinations have shown that the correlations have a negligible effect on the combination when compared to official results. So the combined result of 4.5 sigma still stands.

It is still wrong. The distribution is not flat normal. It is log normal.
Again this has been found to be a good approximation for doing the combination but there is a good point to be made here. Should be we read the number of sigmas off the plot when the CLs scale is linear or logarithmic? Have a look at these two plots which are the same thing on log and linear scale.





Remember that the green and yellow bands show one and two sigma deviations so the excess looks like three sigma on the log scale and four sigma on the linear scale. Which is right? If we assume the flat normal distribution is correct we should be using the linear scale but the bands are more equally spaced on the log scale, so presumably that is more correct. The flat normal approximation is good for generating the plot but we should be careful to read the size of the excess from the log scale. if we do that will the answer be correct.

It's still wrong. For best results use the combined p-value plot.
Have a look at what ATLAS and CMS are quoting for their local significance for the diphoton channel. CMS say 3.1 and ATLAS say it is 2.8. This does not match what you would get from

reading either the linear or logarithmic plots. The numbers come from the p-value plots which are converted to sigma-equivalents. It looks like trying to get these numbers from the exclusion plots or the signal plots will never be that accurate. The bottom line is that we have to wait for the full official combination if we want to know the real answer. Until then adding in quadrature is probably as good as anything else.

References

1. http://blog.vixra.org/2012/02/09/february-higgs-update/

2. http://blog.vixra.org/2012/02/10/some-technical-points-about-combining-sigmas/