# Has CERN Found the God Particle? A Calculation

Philip E. Gibbs[*]

## Abstract

Following the CERN announcement on December 13, 2011, physicists have been giving some very different assessments of the chances that the ATLAS and CMS detectors have seen the Higgs boson. Combining the three things I will consider, I get an overall probability for such a strong signal if there is no Higgs to be about 1 in 30. Perhaps I have failed to account for combinations where more than one of these effects could combine. That requires further coincidences but lets just call the overall result 1 in twenty. In other words, everything considered I take the observed result to be a two sigma effect.

**Key Words:** Higgs Boson, God Particle, CERN, LHC, ATLAS, CMS.

## December 16, 2011:  Has CERN Found the God Particle? A calculation

Yes I know that physicists don't use the term "God particle" but it has entered into popular culture and when the terms "Higgs Boson" and "God Particle" were trending on Twitter and Google earlier this week it was the latter that went the highest. Contrary to what some scientists imagine of the interested public, very few think that there is some religious significance attached to the particle because of this name, it's just a catchy moniker and we need not be afraid to use it.

Following the CERN announcement earlier this week, physicists have been giving some very different assessments of the chances that the ATLAS and CMS detectors have seen the Higgs boson. The CERN DG says merely that they have seen some "interesting fluctuations", while Tommaso Dorigo, (an expert on the statistical aspects of the CMS analysis) calls it "firm evidence". Theorist Lubos Motl is even more positive. He says that it is a "sure thing", but another theorist Matt Strassler has criticised such positive reports. He regards the situation as 50/50 and backed this up with a poll of experimenters that came up 9 to 1 in favour of uncertainty. This contrasts with a similar poll by Bill Murray who is lead Higgs analyst for the ATLAS collaboration. In an interview he reported a 10 to 0 vote that the Higgs had indeed been found.

## What is the question?

So can we make a more objective and quantitative assessment of the current level of uncertainty over the result? You might want to know the probability that the Higgs Boson has been seen for example. Unfortunately this quantity depends on the prior probability that the Higgs Boson exists. Theoretical physicists have a very wide range of opinions on this depending on which theories they favour. Experimenters are supposed to make

[*] Correspondence: Philip E. Gibbs, Ph.D., Independent Researcher, UK. E-Mail: phil@royalgenes.com
   Note: This report is adopted from http://blog.vixra.org/2011/12/16/has-cern-found-the-god-particle-a-calculation/

their assessments independently of such prejudices. So how can we measure the situation objectively?

Luckily there is a different question that is model independent. We can ask for the probability that the experiments would produce results as strong or stronger than those reported if there were no Higgs Boson. This conditional probability removes the theory dependence in the question so the answer should be a number that everyone could in principle agree on. The smaller this probability is, the better the certainty that the Higgs Boson has been found.

Before we can calculate the result we must define precisely what we mean by the "strength of the result". This has to be a single number so it should come from the combined results of both experiments. I will define it to be the maximum value of the CLs likelihood ratio anywhere on the plot. This takes into account both the exclusion side and the signal side of the statistics and is standard use for Higgs searches. Don't worry if you are not familiar with this quantity, it will become clearer in a minute.


## Can we trust the combination?

The Higgs combination groups have tried to spread propaganda that my unofficial combinations cannot be trusted because only people familiar with the inner details of the experimental analysis are capable of doing it correctly. This is not true. I repeatedly acknowledge that my method is an approximation and that only the official combination can be used to claim a discovery, but it is a good approximation and is perfectly acceptable for making a rough assessment of the combined certainty.

They warn that people should not add the event histograms from separate experiments but that is not how my combination is done. They say that only the experts can understand the systematic uncertainties of the detectors well enough to do the combination, but these uncertainties are all built into the individual exclusion plots that they have shown and are therefore taken into account when I combine them. They warned in the past that there are correlations between the background calculations because both experiments use the same algorithms. These correlations are there and must be accounted for to get the most accurate combination possible, but they have been shown to be small. You can ignore these correlations and still get a very good approximation.

In fact the largest source of error comes from the fact that the approximate combination method assumes a flat normal probability distribution at each mass point, when in reality a more complex function based on Poisson distributions would be correct. Happily the central limit theorem says that any error function with a finite variance becomes approximately normal given high enough statistics, so the approximation gets better as more data is added.

When the combination group published their first result in November I was able to compare it with my unofficial combination done in September. This confirmed that the approximation was good. This was no surprise to me because it had already been demonstrated with the Tevatron combinations and some earlier unpublished LHC combinations. I acknowledge that my combinations for some of the individual channels were not so good because the number of events has been low, especially for the ZZ channels. This will have improved for the latest

results because there is now much more data but still these individual channel combinations should be considered less certain than the overall combination.

The assessment I am doing today depends mainly on that, so this is not a big issue, however it is worth showing one further comparison between my combination and the official one for a signal channel. the plot below shows the official combination for the diphoton channel published in November when ATLAS used just 1.1/fb and CMS used just 1.7/fb. The red line is the unofficial result from viXra. It will be interesting to see how much this has improved for 5/fb.
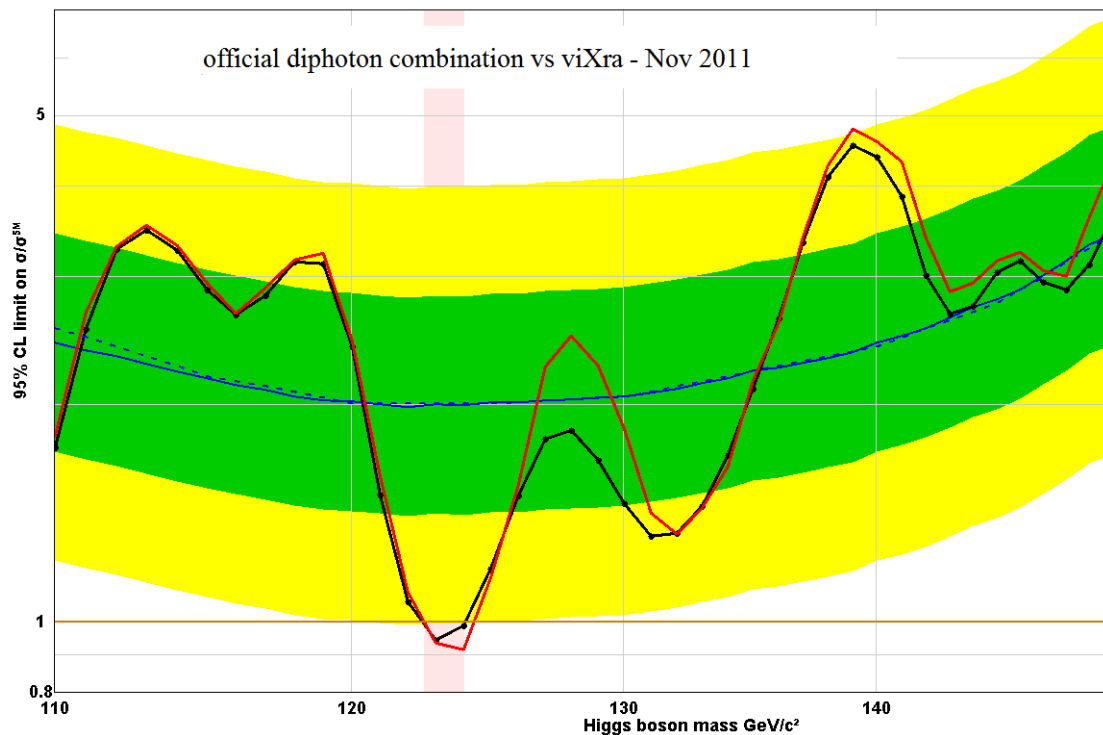


## What must be evaluated?

It is possible to do a systematic evaluation of the probability in question using the combined plot. This takes into account the statistical uncertainties as well as the theoretical uncertainties in the background due to imprecise measurements of the standard model parameters (e.g. W mass) and the approximation methods used in the theory. It also includes the uncertainty in the energy resolution and other similar uncertainties in detector performance. All these things have been considered by the experts from the experimental collaborations and built into the plots, so we don't need to know the details to do the calculation (If anyone tries to claim otherwise they are wrong)

However, there is also the possibility that the experimenters have made some more fundamental kind of error. There may be a subtle fault in the detectors that has not shown up in all the calibration tests which causes an excess on the plot where there should not be one. This should not happen because there are hundreds of people checking for such errors and they are all very competent. Nevertheless bad luck can strike and throw everything out. This

has been the case before and it is probably the case with the OPERA result indicating that neutrinos are faster than light.

A second similar possibility is that the theorists have underestimated the accuracy of some of their calculations so that the background calculation is a little off in one mass range. The analysis involves subtracting a very small signal from a large background, especially in the diphoton channel, so the scope for magnifying any inaccuracy has to be considered. A miscalculation of the signal size is also possible but less likely to lead to a bad result.

As I said, the published plots include all the known experimental and theoretical uncertainties, but these other unknown errors in experiment and theory cannot be accounted for exactly. They can only be estimated based on past experience. Some "expert theorists" say that us more "naive theorists" don't appreciate these facts. Do we really sound so stupid?

## What is the chance of an experimental fault?

How often do experimental faults contribute to a false positive like the excess reported this week? We can only look at past performance but I am not aware of any careful surveys, so a guestimate is required. Someone else may be able to do better. The answer might be one in a hundred but let's be more conservative and say one in ten. If you think it is more common please fell free to reevaluate for yourselves.

However, with the CERN Higgs result we have good evidence that such a fault is not the cause of the excess. That is because there are two independent experiments reporting a very similar result. ATLAS and CMS may seem very similar from the answers they produce, but the detector technologies they use are quite different. The chance of a common fault producing the excess in both detectors must therefore be very small. I am going to assume that this is negligible. If anyone thinks otherwise please explain why.

This means that if the excess is due to such a fault it must be a coincidence that it has a similar effect for both experiments. If there is a one in ten chance of a fault for one experiment, the chance for two independent experiments is one in 100, but even then that is the chance that they would produce the fault at different places. Lets have a look at the two signal plots together.
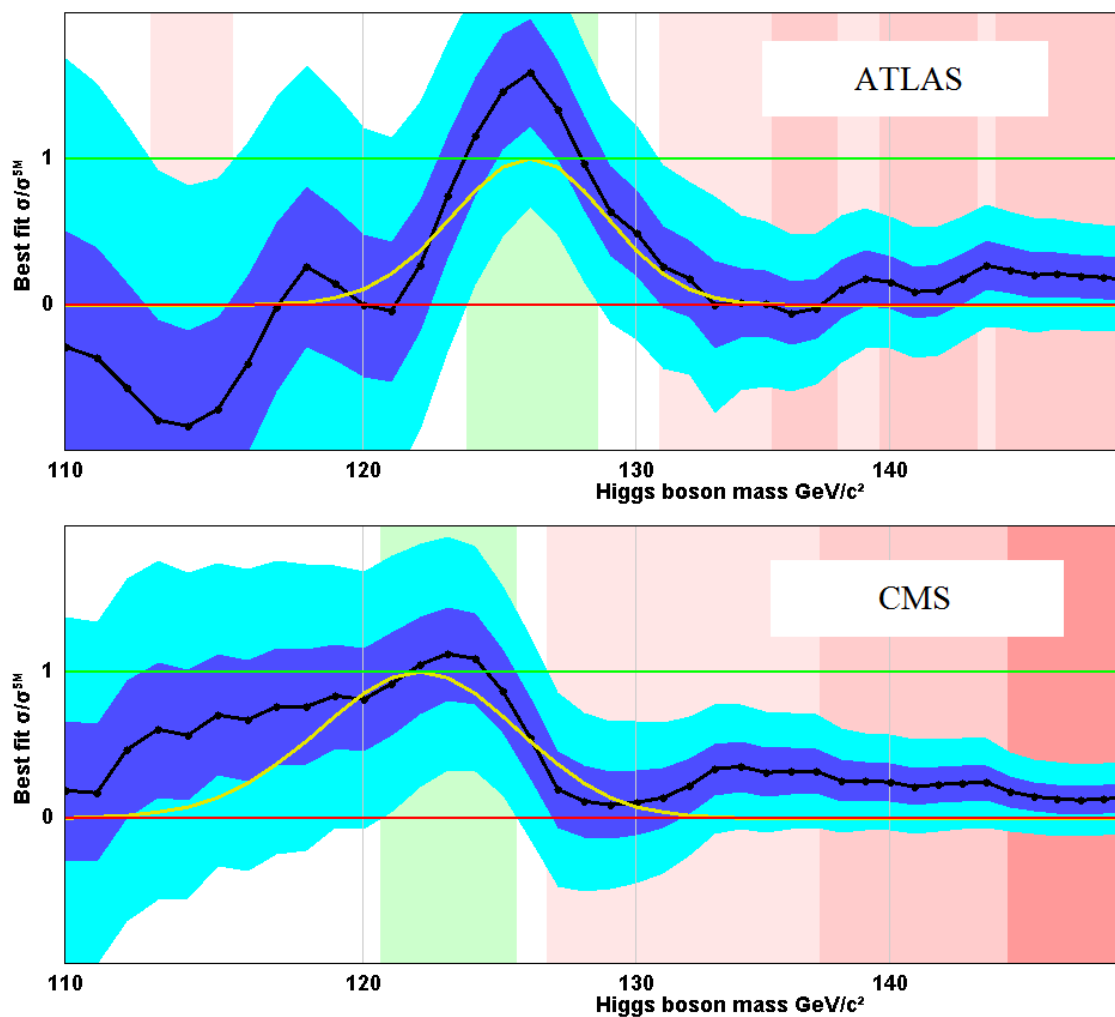
The positions on the maximum excess differ by about 2 GeV but the mass resolution is around 2% so this is not an inconsistency. If these excesses are produced by detector faults then the chance of them lining up so close would be small. How small? That depends on some unknowns. we can't just say the fault could appear anywhere in the mass range, so let's be conservative and just call it a one in three chance.

Overall then we arrive at a one in 300 chance for the observed excess to be explained by a coincidental combination of detector faults. I think this is conservative. Someone else might estimate it to be more probable.

## What is the chance of a theoretical Fault?

The other outside possibility is that the result has been afflicted by a misunderstood background so that the observed excess is really just a subtle effect of the Higgsless standard model that the theorists failed to recognize or estimate correctly. Again this is unlikely but it happens and must be considered. How often does it happen? Once in a hundred perhaps? I will be more cautious and assume one in ten. You may think that is an underestimate in which case you can make your own evaluation.

But again we have more than one place to look. The separate experiments could well be affected by the same theoretical error but the different decay channels are much more independent. There may be some small chance that a single theoretical error could affect all the channels but this would have a small probability, say one in a hundred. If you think it is bigger please justify how that could happen.

So now let's look at the combined signal plots for the three main channels; diphoton, ZZ->4l and WW->lvlv. For the WW plot I can't use the latest CMS results because the plots shown are frankly rubbish quality. I hope they will improve them before publication. However the WW channel has good sensitivity even with less data so I will show the combination from the summer.
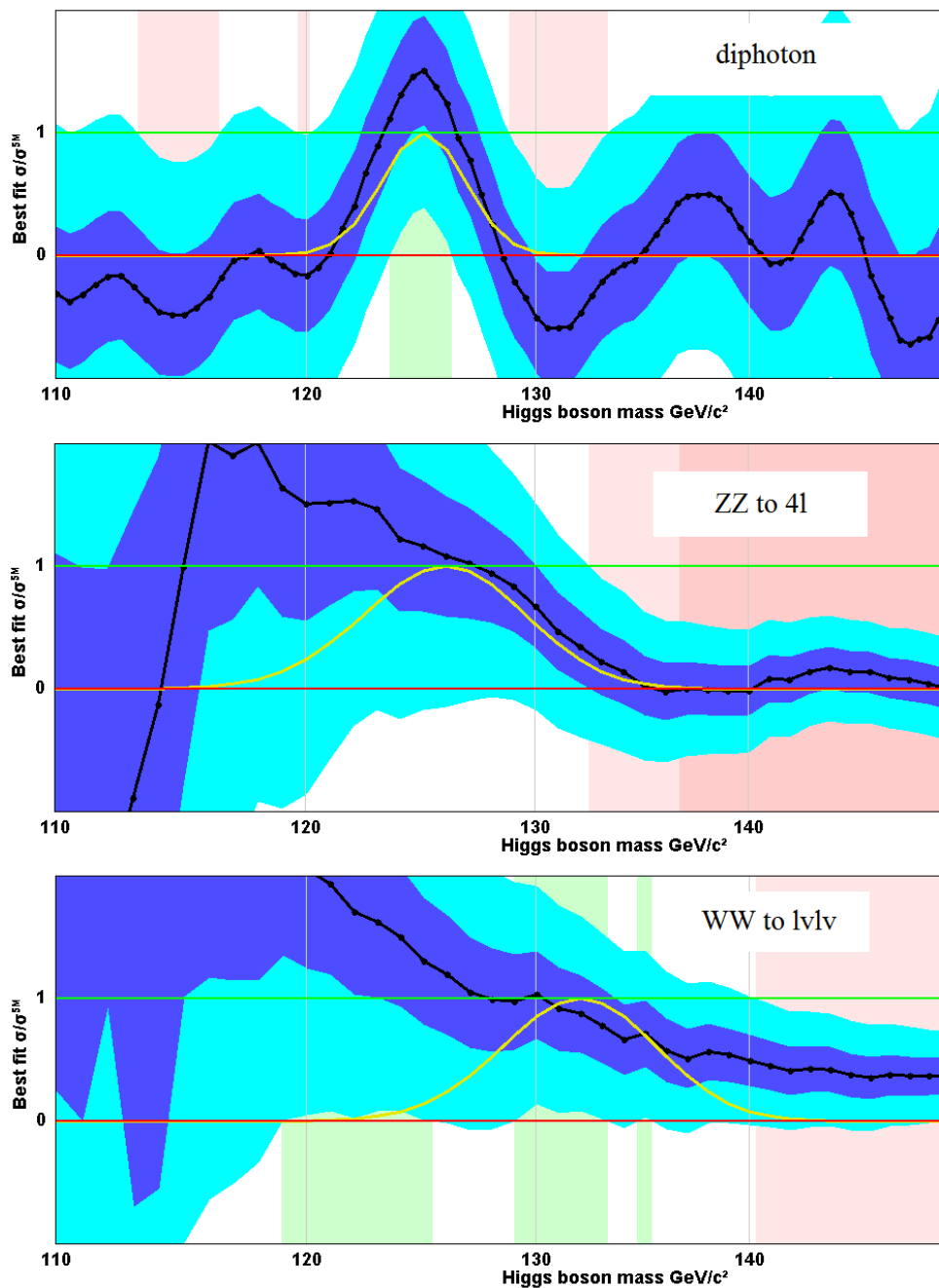
All three channels show an excess in the same low mass region so if this is due to independent faults it would require a coincidence. However, the excess is not as good in ZZ and WW as in the diphoton channel. I am going to put the probability at one in a hundred overall and add to this the probability of one in a hundred for a common fault that affects all three. So the overall chance for a fault from theory is one in 50. Some people will say that this is a low estimate and some people will say that it is low. Others will say that it is nonsense to attempt such an estimate. Never mind, I am just giving it my best honest shot. Let others do the same.

**What is the chance of a statistical fluctuation?**

The last thing to consider is what is the probability of getting s signal as string or stringer than that observed according to the statistical analysis. Actually this also takes into account some theoretical uncertainty and measurement error, but mostly it is statistical. This is a probability that can be worked out more scientifically, but it does include the Look Elsewhere Effect which is partly subjective.

First consider what would be the chance of seeing a signal as strong as the one reported at the fixed mass point of the maximum excess if in fact there was no Higgs Boson. The plot shows a three sigma excess at 124-125 GeV. This would have been much stronger if the peaks from the two experiments had coincided more closely, possibly about 4 sigma. This discrepancy may be due to some detector calibration that could be corrected but it is correct that we do not take that possibility into account. The 3 sigma excess is what we should work with.
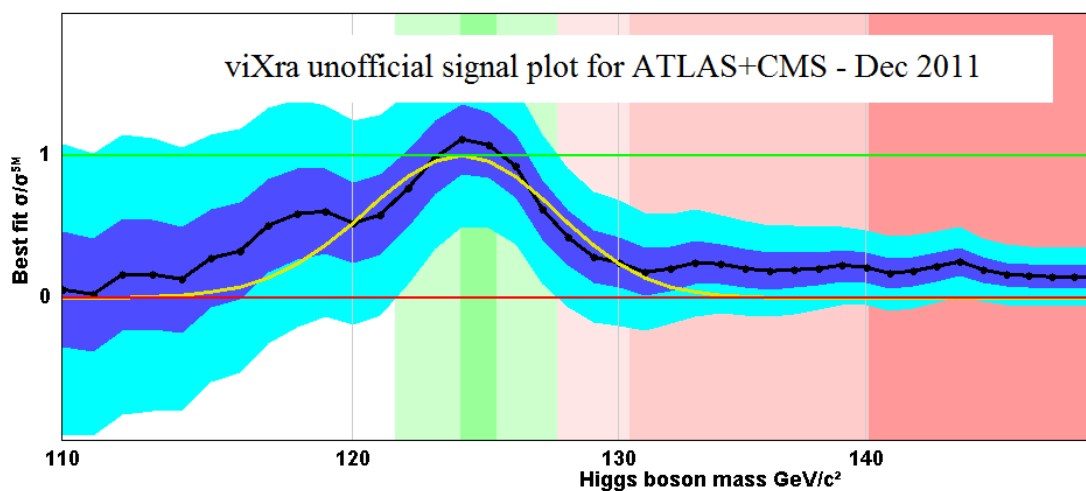


As everyone knows, the probability of a three sigma fluctuation is one in 370, but that allows for fluctuations up or down. So the probability for an excess this big or stronger at this point is one in 740. But we need to know the probability for an excess this strong anywhere on the plot. In other words we need to multiply by the Look Elsewhere Effect factor. Have a look at the plot over the entire range
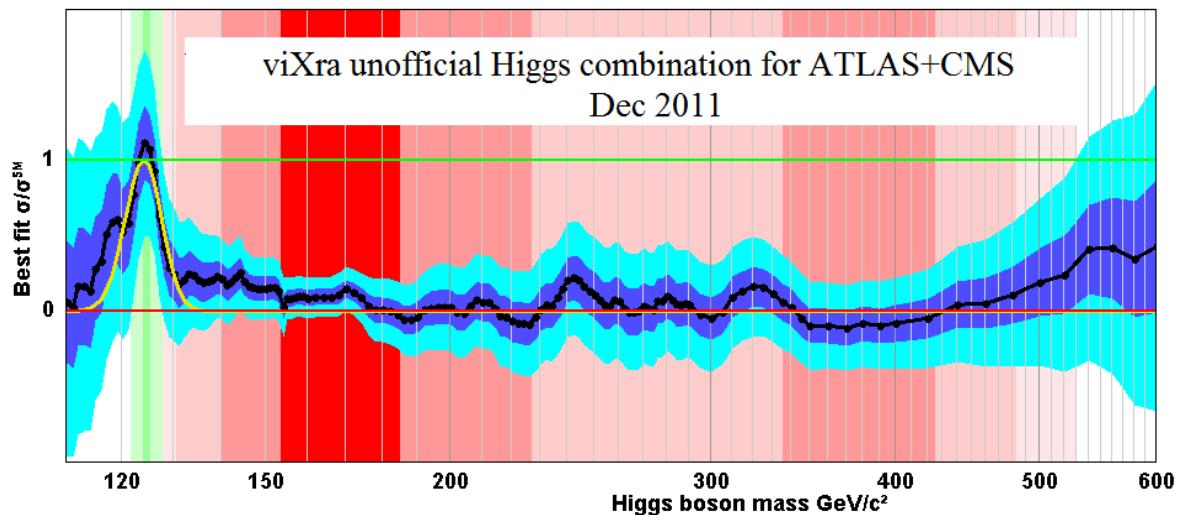
Notice that for the entire range from 130 GeV to 600 GeV the line remains within 2 sigma of the zero line. Big deviations are indeed rare but how rare?

Another point to consider is that if there was a three sigma fluctuation at say 180GeV, the Higgs would still be excluded at that point. This would not count as such a strong signal. This is why I specified that the strength should be measured using the CLs statistic which takes the ration of the probability for the signal hypothesis over the probability from the null hypothesis. This means that the probability of getting a signal as strong in the regions where the Higgs is excluded is much smaller. In fact we can neglect this altogether. So we need only count the regions from 114 GeV (using LEP) to about 135 GeV and perhaps 500 to 600 GeV. Hoe big is the LEE factor for these regions. This depends on the width of the signal which we see to be about 5 GeV in the low mass range due to mass resolution of the detector, and which is much bigger above 500 GeV due to a very large natural width for a high mass Higgs Boson. The LEE factor will therefore be about 6 but let's call it 10 to be extra cautious.

This gives a final answer for the probability of a fluctuation to be about one in 70.

## The final answer?

Combining the three things I have considered I get an overall probability for such a strong signal if there is no Higgs to be about 1 in 30. Perhaps I have failed to account for combinations where more than one of these effects could combine. That requires further coincidences but lets just call the overall result 1 in twenty. In other words, everything considered I take the observed result to be a two sigma effect.

## What about prior probabilities?

there is one more thing you need to take into account when considering how likely a result of any number of sigmas significance is going to stand the test of time. That is your prior estimate for the probability of it being true. The OPERA neutrino observation is a good example of an extreme case. A six sigma effect was observed, but he prior probability of

neutrinos going faster than light would be considered very small by most theoretical physicists. It follows that the probability for this result to go away is quite high despite the statistical significance. An experimental fault is likely to be the biggest contributing factor despite the care of the experimenters.

In fact most 3 sigma excesses for observations beyond the standard model do go away. This is because the prior chance of any one such effect being correct is very small. You can consider this to be part of the Look Elsewhere Effect too. However, the observation of the Higgs Boson is a very different case. Most theoretical physicists would estimate the prior probability for the existence of a Higgs (like) Boson is very high. The standard model provides a very simple explanation of electroweak symmetry breaking but there is no simple way to understand a Higgsless universe. This makes the prior probability high which means that the chance of the 2 sigma result going away is small. There is a bigger chance however that it could move to a different mass.

Not everyone agrees with this. Some people do not think that the Higgs Boson can exist. Stephen hawking is one of them. These people would assign a low value to the prior probability that the signal for the Higgs will be seen and so they will consider it very likely that the present observation will go away. I doubt that there are enough people of this opinion to account for much doubt among the experimenters.

## How long will it take to settle this?

To claim a discovery the combined results must give a 5 sigma excess without considering the Look Elsewhere Effect. How long this takes depends on a certain amount of luck. If the peaks of the excesses comes closer together with more data, then the excess will grow faster than you would otherwise expect. In that case the matter might be settled with just twice as much data and the whole thing will be over by the summer. On the other hand, if they are unlucky it could easily require the full dataset from 2012 to get enough data to finish the job properly. It will then not be until March 2013 when the combination is ready that they will finally be able to declare a discovery.

## References

1. http://blog.vixra.org/2011/12/16/has-cern-found-the-god-particle-a-calculation/