ALEPH 92- 64
PHYSIC 92- 55
30-Apr-1992
J.-F. Grivaz
F. Le Diberder

**On the determination of a mass lower limit for the Higgs boson
in the presence of candidate events.**

Abstract

A method is presented in which the mass resolution is taken into account to derive a mass lower limit for the neutral Higgs boson when candidate events are retained by the selection cuts.

In the various analyses performed by the ALEPH experiment^[1] until the winter of 1992 in order to search for the standard model Higgs boson H produced in the reaction $e^+e^- \rightarrow HZ^*$, no candidate events survived all the selection cuts in a sample of almost 500 000 hadronic Z decays. Therefore, setting a 95% CL lower limit on m_H , the Higgs boson mass, has been a trivial exercise, once the search efficiencies determined as a function of m_H . However, it has been stated in Ref. 1 that about one background event was already expected to be found by the search in the Hl^+l^- channel. The bulk of this expected background is due to the process $e^+e^- \rightarrow l^+l^-q\bar{q}$ and is largely irreducible. It is therefore highly likely that, in the future, the problem will arise of how to optimally handle such background events in order not to lose too much sensitivity. Fortunately, the resolution on m_H , calculated as the mass recoiling against the l^+l^- system, is very good, typically $0.5 \text{ GeV}/c^2$ if $l = \mu$ and $0.75 \text{ GeV}/c^2$ if $l = e$ when $m_H \sim 50 \text{ GeV}/c^2$. This feature has already been used by other LEP experiments; for instance, L3 have recently set a 95% CL lower limit^[2] at $52 \text{ GeV}/c^2$ by simply not considering as valid candidates events with reconstructed masses of 31 and $70 \text{ GeV}/c^2$. The purpose of this note is to indicate how advantage can be taken of the reconstructed mass resolution even when candidate events show up near the boundary of the sensitivity domain.

Consider a Higgs boson with a mass m_H such that the number n_H of signal events expected to be detected has an average value ν . The distribution of n_H around ν is of the Poisson type:

$$\mathcal{P}_\nu(n_H) = e^{-\nu} \frac{\nu^{n_H}}{n_H!}.$$

If a number n of candidate events is actually observed, the usual statement, which is actually correct, is that ν , and therefore m_H , are excluded at some confidence level CL if:

$$\sum_{i=0}^n \mathcal{P}_\nu(i) < 1 - CL.$$

In particular, if no events are selected, the mass lower limit is set to the m_H value such that the corresponding ν satisfies:

$$\mathcal{P}_\nu(0) = 1 - CL,$$

and, if one event is selected, it is set to the m_H value such that the corresponding ν satisfies:

$$\mathcal{P}_\nu(0) + \mathcal{P}_\nu(1) = 1 - CL.$$

In the latter case, to take the information on the mass m_1 of the candidate event into account, one can simply replace the above expression by:

$$\mathcal{P}_\nu(0) + \mathcal{P}_\nu(1) \times G_1(m_1; m_H) = 1 - CL,$$

where $G_1(m_1; m_H)$ is the probability that, if a Higgs event is produced, its reconstructed mass will be “further away” from m_H than m_1 . For the sake of simplicity, this is illustrated

in the case of a Gaussian probability density distribution for the reconstructed mass, with a uniform resolution σ . Then

$$G_1(m_1; m_H) = P\left(\frac{(m_1 - m_H)^2}{\sigma^2}, 1\right)$$

where $P(\chi^2, ND)$ is the χ^2 probability for ND degrees of freedom. It can be shown easily that the exclusion confidence level obtained this way is actually CL , which means that a Higgs boson with mass m_H will be wrongly excluded in a fraction of the cases not exceeding $1 - CL$. In this procedure, no assumption is made as to the mass distribution of the expected background.

When $n > 1$ events are selected, the usual mass lower limit is set to the m_H value such that the corresponding ν satisfies:

$$\sum_{i=0}^n \mathcal{P}_\nu(i) = 1 - CL.$$

It is less obvious now how to modify this expression in order to take into account the information on the candidate event masses m_j ($1 \leq j \leq n$). There is no reason to touch $\mathcal{P}_\nu(0)$, and changing $\mathcal{P}_\nu(n)$ into

$$\mathcal{P}_\nu(n) \times G_n(m_1, \dots, m_n; m_H)$$

is as justified as in the case $n = 1$. With a Gaussian probability density distribution,

$$G_n(m_1, \dots, m_n; m_H) = P(\chi^2, n)$$

with

$$\chi^2 = \sum_{j=1}^n \frac{(m_j - m_H)^2}{\sigma^2}.$$

But there is no unique fully satisfactory way to modify the terms $\mathcal{P}_\nu(1)$ to $\mathcal{P}_\nu(n-1)$.

One possibility is simply not to modify them at all. The estimator thus obtained

$$\mathcal{F}_n^\nu(m_j, j = 1, n; m_H) = \sum_{i=0}^{n-1} \mathcal{P}_\nu(i) + \mathcal{P}_\nu(n) \times G_n(m_j, j = 1, n; m_H)$$

has the advantage that excluding m_H if $\mathcal{F}_n^\nu(m_j, j = 1, n; m_H) < 1 - CL$ automatically leads to an exclusion confidence level equal to (or better than) CL .

On the other hand, the information on the candidate event masses has been used only when it is assumed that all of them actually come from the Higgs boson production process. If one wishes to test whether only $n - 1$ of them are likely to come from the signal while attributing the remaining one to the background, but still with no *a priori* assumption as

to how the background is distributed in mass, one may consider it preferable to replace $\mathcal{P}_\nu(n-1)$ in $\mathcal{F}_n^\nu(m_j, j=1, n; m_H)$ by

$$\mathcal{P}_\nu(n-1) \times B_{n-1}(m_j, j=1, n; m_H),$$

with

$$B_{n-1}(m_j, j=1, n; m_H) = \text{Sup}_{k=1, n} G_{n-1}(m_j, j=1, n, j \neq k; m_H).$$

This corresponds to calling background the event among the n candidates which is the “furthest away” from m_H , and B_{n-1} is a measure of the goodness of the remaining “best” set of $n-1$ events. Similarly, $\mathcal{P}_\nu(n-k)$ can be replaced by

$$\mathcal{P}_\nu(n-k) \times B_{n-k}(m_j, j=1, n; m_H),$$

where $B_{n-k}(m_j, j=1, n; m_H)$ is obtained by trying all sets of $n-k$ events and keeping the “best” one. This is simply achieved by removing from the sample the k events with the smallest $G_1(m_j; m_H)$ ($j=1, n$); in the case of a Gaussian probability density distribution, these are the k events with the largest χ^2 .

This method has the advantage that the information on the masses of the candidate events is taken into account whatever the number of events assumed to be due to the signal and to the background. On the other hand, it has the (minor) disadvantage that the estimator built in this way

$$\mathcal{E}_n^\nu(m_j, j=1, n; m_H) = \sum_{i=0}^n \mathcal{P}_\nu(i) \times B_i(m_j, j=1, n; m_H)$$

does not automatically lead to an exclusion confidence level equal to (or better than) CL when a cut on $\mathcal{E}_n^\nu(m_j, j=1, n; m_H)$ at $1-CL$ is applied. However, for any given ν and n , it is easy to derive once and for all from the *a priori* distribution of $\{\mathcal{E}_k^\nu; k \leq n\}$ for the signal the cut value which provides the required exclusion confidence level; an example of such a distribution is displayed in Fig. 1, and the cut values for $CL = 95\%$ are given in Fig. 2 as a function of ν for $n = 1$ to 4. It can be shown that, in the presence of background, the true exclusion confidence level can only be higher. The power of this method can be appreciated when considering as an example the case where an average of 3 background events is expected, with a uniform mass distribution within $\pm 10\sigma$ of m_H . Then, if no signal is present, $\nu = 4$ will be excluded at 95% CL in 75% of the experiments instead of 5% from simple event counting, and $\nu = 6$ will be excluded in 95% of the experiments instead of 20%.

With this estimator, the effect of the occurrence of candidate events on the 95% CL lower limit on m_H is demonstrated in Fig. 3. As expected, Hl^+l^- candidates affect the limit only in a very narrow mass range. In the example shown, it can be seen that the 95% CL upper limit on the number of events expected reaches the value 5.2 at the location of the candidate event masses. This value would become 4.7, as could be guessed, if only one candidate event were detected (but it would hardly change if additional candidate

events were detected with masses more than a few standard deviations away). In spite of the poorer mass resolution, the use of this estimator can also be beneficial in the case of candidate events in the $H\nu\bar{\nu}$ channel, as also shown in Fig.3.

Up to now, no assumption has been made as to the background intensity or as to its mass distribution. If these can be accurately predicted, it is certainly more efficient to use a maximum likelihood method. But some discriminating power can be gained as soon as the background mass distribution can be expected to be more uniform than the corresponding signal one. For instance, a cut at $\pm 3.3\sigma$ in a Gaussian distribution would remove only 0.1% of the signal while restricting the source for potential background events to a mass range of less than $5 \text{ GeV}/c^2$, for a typical recoil mass resolution in the Hl^+l^- channel. The effect of such a procedure can be seen in Fig. 3. In the example shown, the maximum 95% CL upper limit on the number of signal events is reduced from 5.2 to 4.7; the mass limit, however, is essentially unchanged.

The method presented here can be used for any signal search in which potential backgrounds can be expected. The confidence level determination relies only on the possibility to define, from a set of variables characterizing a candidate event, the probability for a signal event to show up in a “less likely” configuration. No assumptions need to be made on the backgrounds, neither regarding their intensities nor their distributions in the chosen variables.

References.

1. Results presented by the ALEPH Collaboration at the 1992 winter conferences (La Thuile, Moriond, APS); P. Janot and J.-F. Grivaz, "Improved mass limit for the standard model Higgs boson", note ALEPH 92-37, PHYSIC 92-34, 5 March 1992.
2. B. Adeva et al., L3 Coll., "Search for the neutral Higgs boson at LEP", CERN-PPE/92-40, 18 March 1992.

Figure Captions.

1. For $\nu = 6$ and $n = 4$, *a priori* distribution of $\{\mathcal{E}_k^\nu; k \leq n\}$ from 0. to 1. (top) and from 0. to 0.1 (bottom). The normalization corresponds to 100 000 experiments. The vertical line indicates the value corresponding to a 95% confidence level.
2. As a function of ν , cut value to be applied to $\{\mathcal{E}_k^\nu; k \leq n\}$ when a 95% confidence level is required. The curves are labelled after the corresponding n value ($n = 1$ to 4).
3. The number of Higgs events expected, taken from Ref. 1, as a function of the Higgs boson mass (dashed curve). With no candidates observed, the 95% CL limit is set at 53 GeV/ c^2 , corresponding to 3.0 events expected. Had an Hl^+l^- candidate been observed at 53 GeV/ c^2 and two others at 48 and 56 GeV/ c^2 , with a 0.75 GeV/ c^2 Gaussian mass resolution, the limit would have been set at 52 GeV/ c^2 as indicated by curve (a), instead of 50 GeV/ c^2 (48 GeV/ c^2) considering one (two) candidate event(s) and using simple event counting. With an additional mass cut at $\pm 3.3\sigma$, the limit remains at 52 GeV/ c^2 , as indicated by curve (b). Had a single $H\nu\bar{\nu}$ candidate been observed at 55 GeV/ c^2 , assuming for simplicity a 5 GeV/ c^2 Gaussian mass resolution, the limit would have been set at 51 GeV/ c^2 as indicated by curve (c), instead of 50 GeV/ c^2 from simple event counting.

Figure 1

Figure 2

ν

Figure 3