CERN/LEPC 97-11
LEPC/M 115
November 3, 1997

ALEPH 97-101 PHISIC 97-100
DELPHI 97-167 PHYS 743
L3 Note 2167
OPAL Technical Note TN-518

Lower bound for the SM Higgs boson mass: combined result from the four LEP experiments

Report of the LEP working group for Higgs boson searches [1]

Abstract

A lower limit for the mass of the Standard Model Higgs boson is reported. It is based on data collected by the four LEP experiments, ALEPH, DELPHI, L3 and OPAL, at centre-of-mass energies of 161 GeV and 170-172 GeV. The experimental inputs and the statistical methods used for combining the results are described. A lower limit of 77.5 GeV/c^2 is obtained at the 95% confidence level.

[1]Contributions from P. Bock, J. Carr, S. De Jong, F. Di Lodovico, E. Gross, P. Igo-Kemenes, P. Janot, W. Murray, M. Pieri, A. Read, V. Ruhlmann-Kleider, A. Sopczak.

1 Introduction

After five years of running LEP at centre-of-mass energies close to the Z resonance (LEP1), the energy of the e^+e^- collider has been raised (LEP2) in successive steps to 172 GeV. The data accumulated by the four LEP experiments at the new centre-of-mass energies allow the mass range of Higgs boson searches to be extended well beyond the bounds obtained at LEP1 [1]. The sensitivity of the searches is particularly high when the results of individual experiments are statistically combined.

Using the data collected at energies between 130 GeV and 172 GeV in conjunction with LEP1 data, the four experiments have reported no evidence for Higgs boson production. The published limits for the Standard Model (SM) Higgs boson mass, m_{H}, from the four experiments are 70.7 GeV/c^2 (ALEPH [2]), 66.2 GeV/c^2 (DELPHI [3]), 69.5 GeV/c^2 (L3 [4]) and 69.4 GeV/c^2 (OPAL [5]).

The LEP working group for Higgs boson searches has compared several strategies for combining the results from individual experiments. The problem is rather complex since each experiment uses data collected at various energies, looks for a signal in various final states and uses different statistical methods for combining channels [2]. Also, the experiments use different methods for data selection.

As a first application, a new, combined, 95% confidence level (c.l.) limit for the SM Higgs boson mass is produced. Combined limits for Higgs bosons predicted by the Minimal Supersymmertic Standard Model (MSSM) will be presented later in a separate note.

This note is structured as follows: Section 2 summarizes the experimental inputs, separately for each experiment. Detailed information can be found in the original publications and notes. Section 3 describes the statistical methods employed by each experiment internally and which are proposed for combining the results. Comparative tests, based on Monte Carlo studies, are also described in that section. The combined results are presented in Section 4 and summarized in Section 5.

2 Experimental input

The data used in this note have been collected at 161 GeV and 170-172 GeV energies. The integrated luminosities at each of the two energy settings are ≈ 10 pb^{-1}, for a total of ≈ 20 pb^{-1} per experiment. LEP1 data, and LEP2 data taken at 130-136 GeV are not used since their contributions to the combined mass limit would be small (< 200 MeV/C^2).

Each experiment has performed searches for Higgsstrahlung, $e^+e^- \to \mathrm{ZH}$, in the following

[2] Throughout this note, the word "channel" may designate different search channels at the same energy, or the same search channels at different energies, or even the same search channels at the same energy but looked at by different experiments.

main final states:

- (Z$\rightarrow$q$\bar{\text{q}}$)(H$\rightarrow$q$\bar{\text{q}}$) (q = any hadron flavour), with branching fraction $B \approx$ 64%;
- (Z$\rightarrow\nu\bar{\nu}$)(H$\rightarrow$q$\bar{\text{q}}$), $B \approx$ 18%;
- (Z$\rightarrow$e$^+$e$^-$, $\mu^+\mu^-$)(H$\rightarrow$q$\bar{\text{q}}$), $B \approx$ 6.2%;
- (Z$\rightarrow\tau^+\tau^-$)(H$\rightarrow$q$\bar{\text{q}}$), $B \approx$ 3.1%;
- (Z$\rightarrow$q$\bar{\text{q}}$)(H$\rightarrow\tau^+\tau^-$), $B \approx$ 5.4%.

As the Higgs boson decays dominantly ($\sim$ 86%) into b$\bar{\text{b}}$, most searches are restricted to H$\rightarrow$b$\bar{\text{b}}$ final states. Some are also sensitive to decays into other quark flavours or even into gluon pairs. The searches in the (Z$\rightarrow\nu\bar{\nu}$)H and (Z$\rightarrow$e$^+$e$^-$)H channels are also sensitive to small (<10%) contributions to Higgs boson production by WW and ZZ fusion; these are taken into account in most analyses.

2.1 The ALEPH input

The ALEPH searches are described in Ref. [2]. The data samples correspond to integrated luminosities of 10.9 pb^{-1} at 161 GeV and 10.6 pb^{-1} at 170-172 GeV energies. Table 1 summarizes, channel by channel, the expected backgrounds, the signal selection efficiencies and the expected numbers of signal events for a Higgs boson of 70 GeV/c^2 mass. The variation of the selection efficiencies and signal expectations as functions of the Higgs boson mass are presented in Figure 1. In the (Z$\rightarrow$q$\bar{\text{q}}$)H channel, events in which the Higgs boson decays to τ leptons are explicitly removed, as this final state is selected by the (H$\rightarrow\tau^+\tau^-$)q$\bar{\text{q}}$ analysis.

Final State	Background		Efficiency(%)		$\langle n_{\text{s}} \rangle$	
Hq$\bar{\text{q}}$	0.17	0.23	21.1	21.9	0.24	1.12
H$\nu\bar{\nu}$	0.06	0.09	26.3	42.9	0.11	0.70
H(e$^+$e$^-$, $\mu^+\mu^-$)	0.06	0.11	64.2	74.8	0.08	0.40
H$\tau^+\tau^-$	0.02	0.02	18.8	20.4	0.01	0.05
(H$\rightarrow\tau^+\tau^-$)q$\bar{\text{q}}$	0.05	0.03	17.4	17.4	0.02	0.07
Total	0.36	0.48			0.46	2.39

Table 1: *ALEPH selection: expected background, signal efficiency, and the expected number of signal events for a Higgs boson mass of 70* GeV/c^2*. In each column, the entries to the left are for 161 GeV and those to the right for 170-172 GeV. The expected numbers of signal events take into account small contributions from* WW *and* ZZ *fusion.*

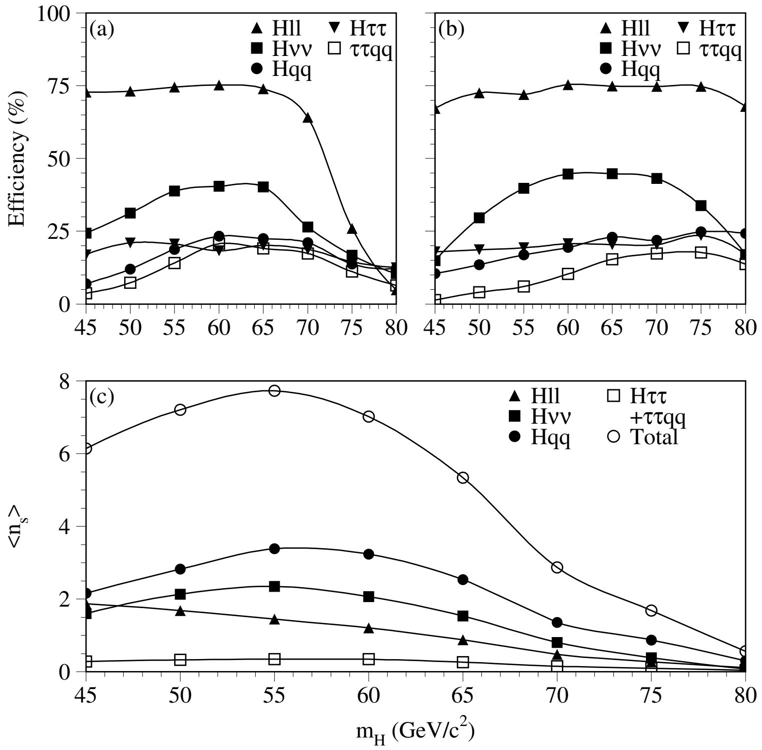

Figure 1: *ALEPH selection: the detection efficiency as a function of the Higgs boson mass at (a) 161 GeV and (b) 170-172 GeV energies; (c) the numbers of expected events, all energies combined, as functions of the Higgs boson mass.*

No candidate events are retained in any of the selections, in agreement with 0.84 events expected from SM background processes. In the absence a signal, the results of the five selections are combined using the statistical procedure outlined in Section 3.1 below. Including only the data collected at energies of 161 GeV and 170-172 GeV, ALEPH obtains a 95% c.l. lower limit of 69.6 GeV/c^2 for the SM Higgs boson mass.

Figure 2 shows distributions of the reconstructed Higgs boson mass in (a) the (H$\rightarrow$q$\bar{\text{q}}$)q$\bar{\text{q}}$ and (b) the(H$\rightarrow$q$\bar{\text{q}}$)$\nu\bar{\nu}$ channels, for the case of 170-172 GeV energy, at the end of the data selection, for the simulated SM background and for a signal with $m_{\rm H}$=70 GeV/c^2. In these two channels, which have the largest contribution to the expected signal, such a Higgs boson would be reconstructed with mass resolutions of 2.8 and 2.9 GeV/c^2, respectively, as estimated by Gaussian fits to the central parts of the mass spectra.

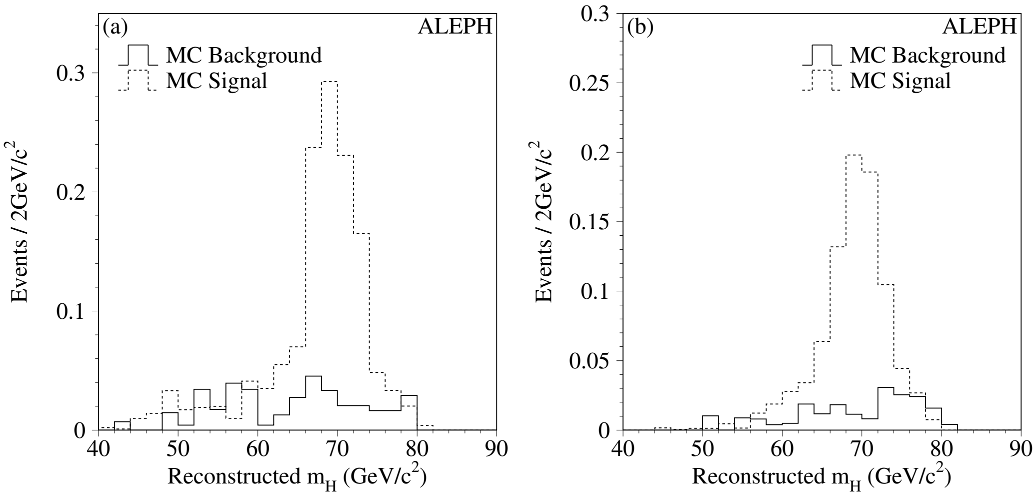

Figure 2: *ALEPH selection: reconstructed Higgs boson mass at the end of the data selection, for the case of 170-172 GeV energy, (a) in the* $(\mathrm{H}\rightarrow \mathrm{q\bar{q}})\mathrm{q\bar{q}}$ *and (b) in the* $(\mathrm{H}\rightarrow \mathrm{q\bar{q}})\nu\bar{\nu}$ *channels. Full-line: distribution for the SM background; dashed: for a signal with* m_{H}=70 GeV/c^2. *The histograms are scaled to the integrated luminosity of the data.*

2.2 The DELPHI input

The experimental selections of DELPHI are described in Ref. [3]. The data samples correspond to integrated luminosities of 10.0 pb^{-1} at 161 GeV and 10.0 pb^{-1} at 170-172 GeV energies. Table 2 summarizes, channel by channel, the expected backgrounds, the signal selection efficiencies and the expected numbers of signal events for a Higgs boson of 70 GeV/c^2 mass. The first four channels listed in the table include all decays of the Higgs boson; this is not the case for the two channels with decays into τ leptons. It has been checked, however, that the $\mathrm{Hq\bar{q}}$ selection is inefficient for the $(\mathrm{H} \rightarrow \tau^+\tau^-)$ decay; therefore, there is no double counting in the expected numbers of signal events. The variation of the selection efficiencies and signal expectations as functions of the Higgs boson mass are presented in Figure 3.

Figure 4 shows distributions of the reconstructed Higgs boson mass in the $\mathrm{Hq\bar{q}}$ and the $\mathrm{H}\nu\bar{\nu}$ channels for the case of 170-172 GeV energy, at the end of the data selection. In the $\mathrm{Hq\bar{q}}$ channel the Higgs boson mass is estimated as the sum of the dijet masses minus the Z mass while in the $\mathrm{H}\nu\bar{\nu}$ channel it is the event mass after rescaling of the energy and momentum of the visible system assuming m_{Z} for the recoil mass. The figure shows distributions for the remaining data, the simulated SM background and for a signal with m_{H}=70 GeV/c^2. In these two channels the Higgs boson mass is reconstructed with resolutions of 2.8 and 3.8 GeV/c^2, respectively, as estimated by Gaussian fits to the central parts of the mass spectra.

Two events are selected in the data: one in the $\mathrm{H}\nu\bar{\nu}$ channel at 161 GeV, with a mass of 64.6 GeV/c^2, and one in the $\mathrm{Hq\bar{q}}$ channel at 172 GeV, with a mass of 58.7 GeV/c^2. Using only

Final State	Background		Efficiency(%)		$\langle n_{\mathrm{s}} \rangle$	
$\mathrm{Hq\bar{q}}$	0.30	0.50	21.6	23.6	0.25	1.21
$\mathrm{H}\nu\bar{\nu}$	0.65	0.61	36.3	42.8	0.12	0.63
$\mathrm{He^+e^-}$	0.13	0.20	41.7	37.2	0.02	0.09
$\mathrm{H}\mu^+\mu^-$	0.04	0.13	69.0	69.8	0.04	0.17
$(\mathrm{H} \to \mathrm{q\bar{q}})\tau^+\tau^-$	0.31	0.22	22.9	24.4	0.01	0.06
$(\mathrm{H} \to \tau^+\tau^-)\mathrm{q\bar{q}}$	0.32	0.91	22.1	24.4	0.02	0.10
Total	1.74	2.50			0.46	2.26

Table 2: *DELPHI selection: expected background, signal efficiency, and the expected number of signal events for a Higgs boson mass of 70* $\mathrm{GeV/c^2}$. *The left-hand entries are for* 161 GeV *and the right-hand entries are for* 170–172 GeV. *The total background is corrected for a small overlap between the two* τ *channels. The expected numbers of signal events do not take into account the contributions from* WW *and* ZZ *fusion.*

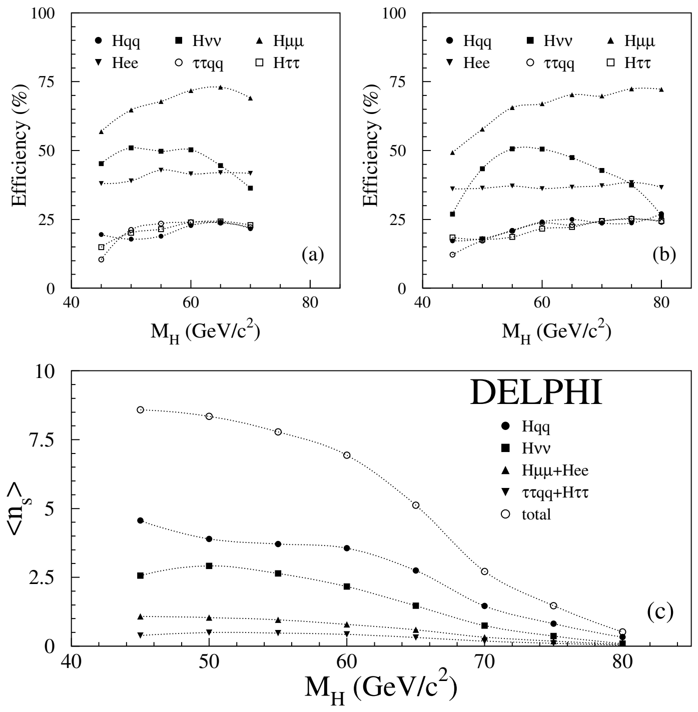

Figure 3: *DELPHI selection: the detection efficiency as a function of the Higgs boson mass at 161 GeV and 170-172 GeV energies; lower part: the number of expected events, all energies combined, as a function of the Higgs boson mass.*

the data collected at energies of 161 GeV and 170-172 GeV, DELPHI obtains a 95% c.l. lower limit of 65.9 GeV/c^2 for the SM Higgs boson mass, using the statistical procedure described in Section 3.2 below, which is different from the one used in Ref. [3].

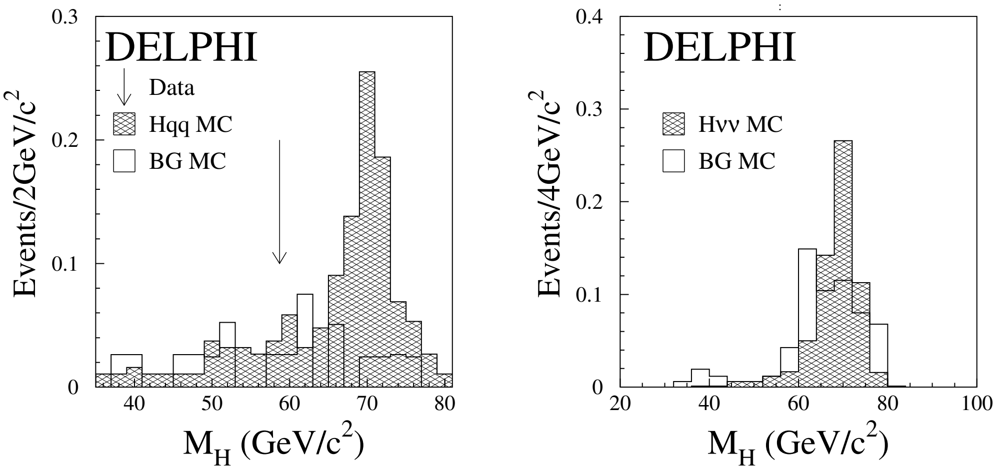

Figure 4: *DELPHI selection: reconstructed Higgs boson mass at the end of the data selection at 170-172 GeV; in the* $\mathrm{Hq\bar{q}}$ *channel and* $\mathrm{H}\nu\bar{\nu}$ *channels. White histograms: distributions for the SM background; hatched histograms: for a signal with* m_{H}*=70 GeV/c*^{2}*. The arrow indicates the mass of one candidate event. The histograms are scaled to the experimental integrated luminosity.*

2.3 The L3 input

The experimental selections of L3 are described in Ref. [4]. Details of the analysis can also be found in Refs. [6, 7]. The data samples correspond to integrated luminosities of 10.8 pb^{-1} at 161 GeV and 10.2 pb^{-1} at 170-172 GeV energies.

For events which satisfy the selection criteria, variables with high discriminating power, such as topological observables, the reconstructed invariant mass of the Higgs boson, and a b-tag variable, are combined to form an event "weight", W_{H}, which is used for the final separation of the signal from the background. A sliding cut, depending on the hypothetical Higgs boson mass, is applied in W_{H} such as to provide a signal-to-background ratio larger than 0.15 in all cases. This cut considerably reduces the uncertainty in the estimation of the expected background; it implies, however, that the number of selected candidates and the expected SM background depend on the Higgs boson mass hypothesis. After applying this cut, approximately 10 background events are expected for any mass hypothesis between 60 and 70 GeV/c^2. Distributions of W_{H} in the $(\mathrm{H}\to\mathrm{q\bar{q}})\mathrm{q\bar{q}}$ and $(\mathrm{H}\to\mathrm{q\bar{q}})\nu\bar{\nu}$ channels at 170-172 GeV energy, after a cut which corresponds to m_{H}=70 GeV, are presented in Figure 5.

Table 3 summarizes the expected backgrounds, the signal selection efficiencies and the ex-

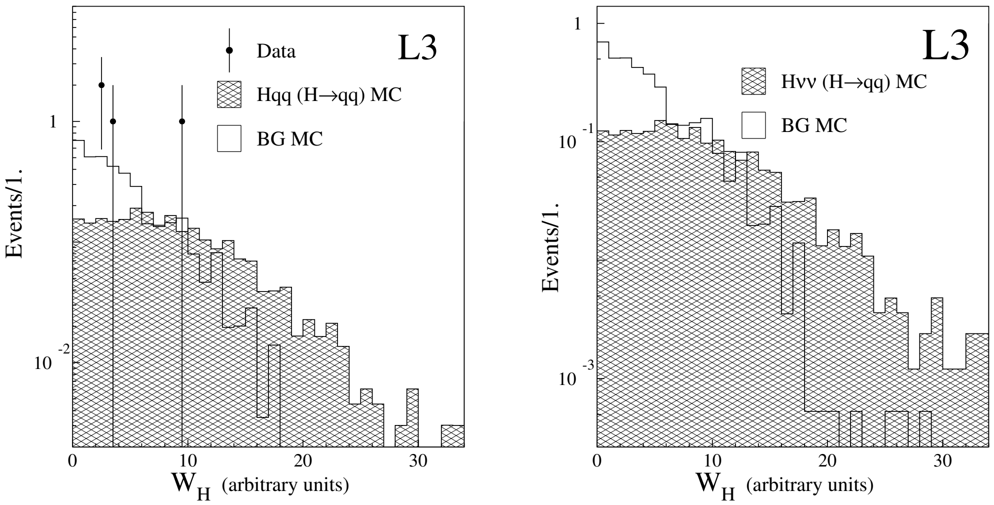

Figure 5: *The L3 event weight spectrum (arbitrary units) for the (H→*$q\bar{q}$*)*$q\bar{q}$ *and (H→*$q\bar{q}$*)*$\nu\bar{\nu}$ *channels, at the end of the event selection applied at 170-172 GeV energies. White histograms: SM background; hatched histograms: Higgs boson signal for 70* GeV/c^2 *mass. All histograms are normalized to the integrated luminosity of the data. The dots with error bars indicate selected candidates.*

Selection	Background		Efficiency(%)		$\langle n_s \rangle$	
$(H\rightarrow q\bar{q})q\bar{q}$	0.77	3.68	28.1	38.5	0.37	1.87
$(H\rightarrow q\bar{q})\nu\bar{\nu}$	0.40	1.46	46.0	69.4	0.17	0.97
$(H\rightarrow q\bar{q})e^+e^-$	0.03	0.18	45.5	65.8	0.03	0.15
$(H\rightarrow q\bar{q})\mu^+\mu^-$	0.04	0.15	34.4	48.3	0.02	0.11
$(H\rightarrow q\bar{q})\tau^+\tau^-$	0.008	0.23	13.5	34.9	0.01	0.08
$(H\rightarrow \tau^+\tau^-)q\bar{q}$	0.0	0.25	0.0	17.7	0.0	0.07
Total	1.25	5.96			0.60	3.26

Table 3: *L3 selection: expected background, selection efficiency, and the expected numbers of signal events for a Higgs boson mass of 70* GeV/c^2 *for the six topological selections. The selection efficiency is defined as the ratio of the total number of signal events selected to the number of signal events produced in the channel for which the selection is devised. In each column, the entries to the left are for 161 GeV and those to the right for 170-172 GeV. The expected numbers of signal events take into account small contributions from* WW *and* ZZ *fusion.*

pected numbers of signal events for the case of a Higgs boson of 70 GeV/c^2 mass, after the corresponding cut in W_{H}, for the six topological selections used. Since each selection is applied to all Higgs boson final states, as explained in [4, 6], the selection efficiencies are defined as the number of signal events retained by a specific selection, divided by the number of Higgs events produced in that channel.

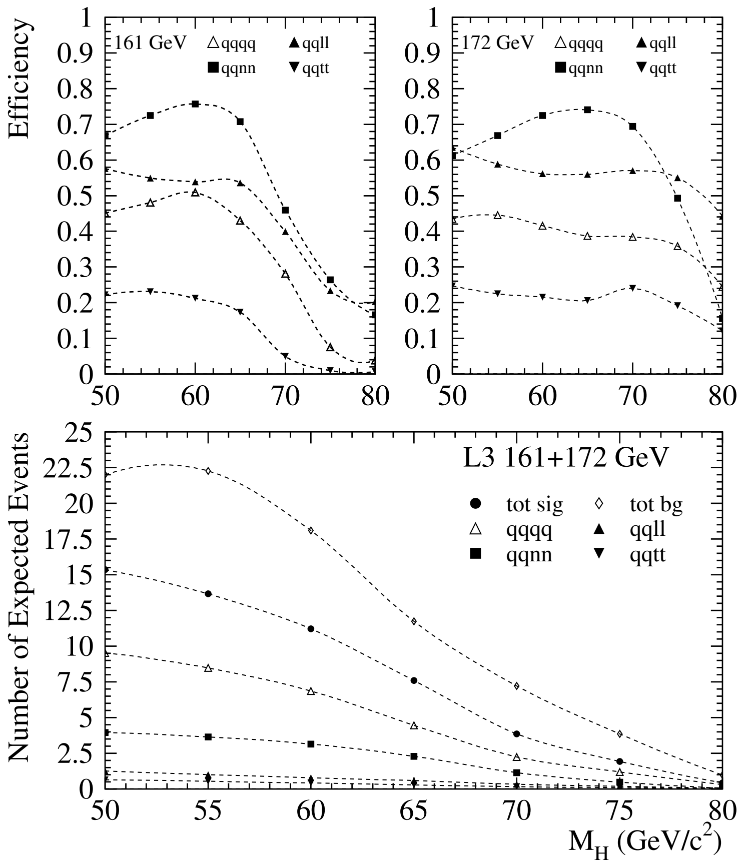

Figure 6: *L3 selection: the detection efficiency as a function of the Higgs boson mass at 161 GeV and 170-172 GeV energies. Lower part: the number of expected events for all energies as a function of the Higgs boson mass. Since, in the L3 selection, the background also depends on* m_{H}*, the total background is also shown. In the legend,* $\ell\ell$ *includes* $\mathrm{e^+e^-}$ *and* $\mu^+\mu^-$*; the channel* $\mathrm{q\bar{q}}\tau^+\tau^-$ *includes both cases, where either the H or the Z bosons decay into* $\tau^+\tau^-$ *pairs.*

The energies, channels of occurence and reconstructed masses of the six events selected, assuming a hypothetical mass of 70 GeV/c^2, are listed in Table 4. The candidates are ordered according to decreasing expected signal-to-noise ratio. The variation of the signal selection efficiencies, of the numbers of expected signal and background events, as functions of the Higgs boson mass, are presented in Figure 6. Using only the data collected at 161 GeV and 170-

Event	Energy	Channel	Higgs mass (GeV/c^2)
1	170-172 GeV	(H→q$\bar{\mathrm{q}}$)q$\bar{\mathrm{q}}$	73.3
2	170-172 GeV	(H→q$\bar{\mathrm{q}}$)q$\bar{\mathrm{q}}$	57.0
3	170-172 GeV	(H→q$\bar{\mathrm{q}}$)q$\bar{\mathrm{q}}$	70.9
4	170-172 GeV	(H→q$\bar{\mathrm{q}}$)q$\bar{\mathrm{q}}$	68.2
5	161 GeV	(H→q$\bar{\mathrm{q}}$)q$\bar{\mathrm{q}}$	73.9
6	161 GeV	(H→q$\bar{\mathrm{q}}$)q$\bar{\mathrm{q}}$	73.7

Table 4: *Data events selected by L3, after the cut in W_{H} for a Higgs boson of 70* GeV/c^2 *mass. The candidates are ordered from high to low signal-to-noise ratio expected. The Higgs boson masses are those recontructed by the 4C kinematic fit for the jet combination which minimizes the χ^2 for a 70* GeV/c^2 *Higgs boson mass hypothesis (see Ref. [4, 6]).*

172 GeV energies, L3 obtains a 95% c.l. lower limit of 69.3 GeV/c^2 for the SM Higgs boson mass, using the statistical procedure described in Section 3.3 below.

2.4 The OPAL input

The OPAL analysis is described in Ref. [5]. Details of the event selections in the various search channels can also be found in Ref. [8]. The data samples correspond to integrated luminosities of 10.0 pb^{-1} at 161 GeV and 10.4 pb^{-1} at 170-172 GeV energies. Table 5 summarizes the expected background, the signal selection efficiencies and the expected numbers of signal events for a Higgs boson of 70 GeV/c^2 mass. Curves of the selection efficiencies and signal expectations as functions of the Higgs boson mass are presented in Figure 7.

Figure 8 shows distributions of the reconstructed Higgs boson mass in the (H→b$\bar{\mathrm{b}}$)q$\bar{\mathrm{q}}$ and (H→q$\bar{\mathrm{q}}$)$\nu\bar{\nu}$ channels, for the case of 170-172 GeV energy, at the end of the data selection. The figure shows distributions for the SM background and for a signal with m_{H}=70 GeV/c^2 mass. The mass of one selected candidate is also indicated. In these two channels, the Higgs boson mass is reconstructed with resolutions of 3.0 and 3.6 GeV/c^2, respectively, as estimated by Gaussian fits to the central parts of the mass spectra.

Two candidates are selected: the one indicated in Figure 8, with m_H=75.6 GeV/c^2, in the (H→b$\bar{\mathrm{b}}$)q$\bar{\mathrm{q}}$ channel at $\sqrt{s}$=170-172 GeV and one, with m_{H}=39.3 GeV/c^2, in the (H→q$\bar{\mathrm{q}}$)$\nu\bar{\nu}$ channel at $\sqrt{s}$=161 GeV. Using only the data collected at 161 GeV and 170-172 GeV energies, OPAL obtains a 95% c.l. lower limit of 68.9 GeV/c^2 for the SM Higgs boson mass. This limit is obtained using the statistical procedure described in Section 3.4 below, which differs from the method employed in Ref. [5] only in the treatment of the residual background.

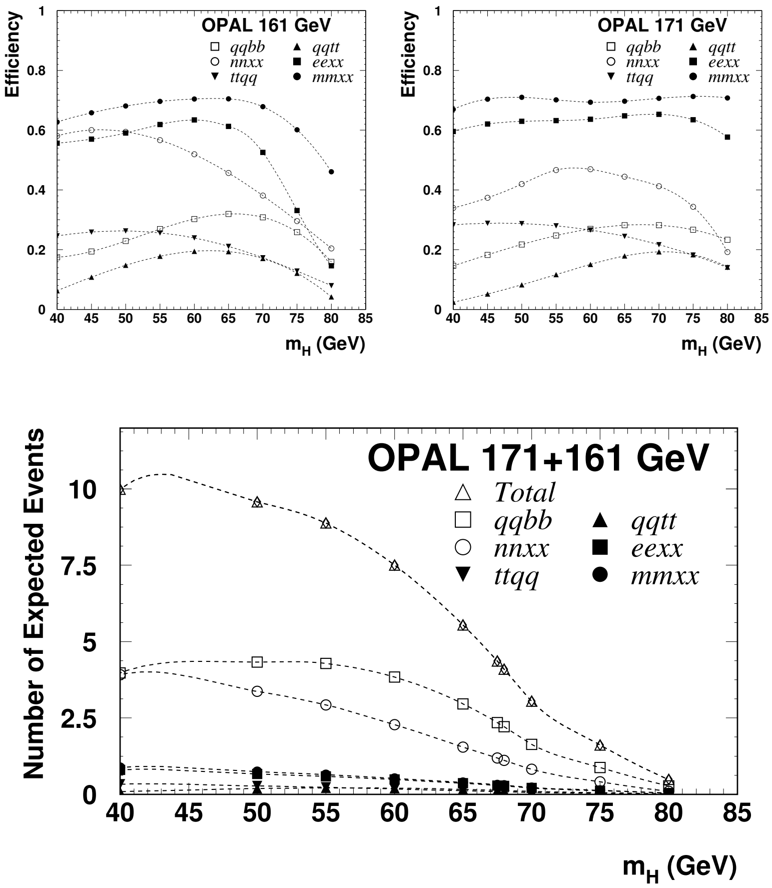

Figure 7: *OPAL selection: the detection efficiency as a function of the Higgs boson mass at 161 GeV and 170-172 GeV energies. Lower part: the numbers of expected events, all energies combined, as a function of the Higgs boson mass.*

Final State	Background		Efficiency(%)		$\langle n_s \rangle$	
$(H\to b\bar{b})q\bar{q}$	0.75	0.88	30.8	28.2	0.35	1.29
$(H\to q\bar{q})\nu\bar{\nu}$	0.90	0.55	38.1	41.3	0.16	0.66
He^+e^-	0.06	0.08	52.6	65.3	0.04	0.17
$H\mu^+\mu^-$	0.04	0.06	67.8	70.6	0.04	0.18
$(H\to q\bar{q})\tau^+\tau^-$	0.10	0.41	17.3	21.7	0.01	0.05
$(H\to \tau^+\tau^-)q\bar{q}$	0.06	0.18	17.0	19.3	0.02	0.08
Total	1.91	2.16			0.62	2.43

Table 5: *OPAL selection: expected background, signal efficiency, and the expected numbers of signal events for a Higgs boson mass of 70* GeV/c^2*. In each column, the entries to the left are for 161 GeV and those to the right for 170-172 GeV. The expected numbers of signal events take into account small contributions from* WW *and* ZZ *fusion.*

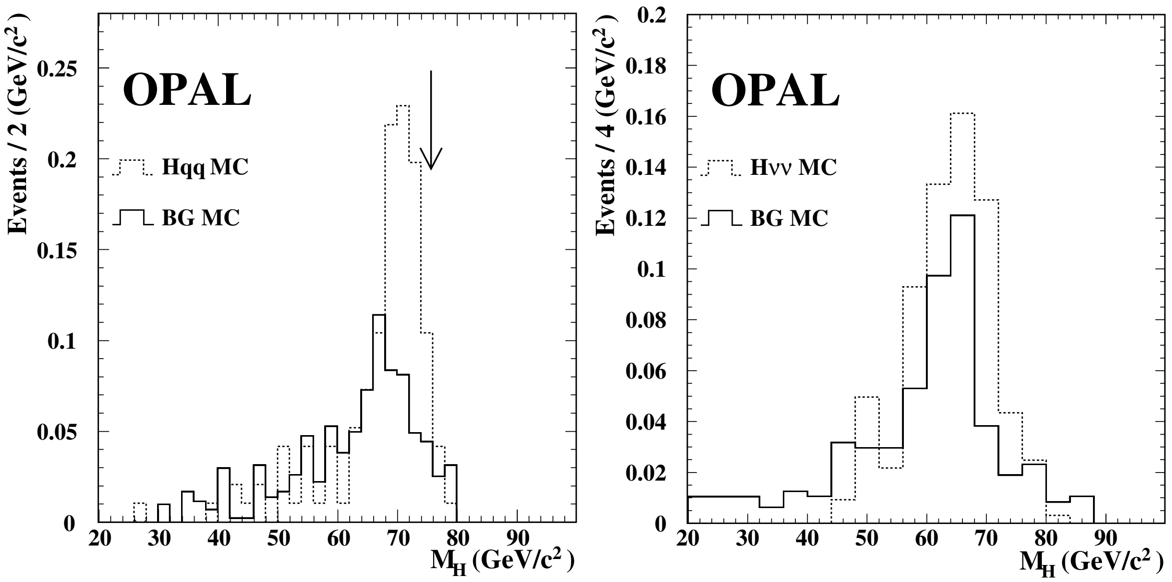

Figure 8: *OPAL selection: reconstructed Higgs boson mass at the end of the data selection, for the case of 170-172 GeV energy, in the* $(H\to q\bar{q})q\bar{q}$ *and* $(H\to q\bar{q})\nu\bar{\nu}$ *channels. Full-line histograms: distributions for the SM background; dashed-line histograms: for a signal with* $m_H=70\ \mathrm{GeV}/c^2$*. The arrow indicates the mass of one selected candidate. The histograms are scaled to the integrated luminosity of the data.*

2.5 Summary, experimental inputs

In Table 6 some results of the four selections are listed. The table shows the total backgrounds,

Experiment	Background		$\langle n_{\mathrm{s}} \rangle$		N_{obs}		m_{95}^{lim} (GeV/c^2)
ALEPH	0.36	0.48	0.46	2.39	0	0	69.6
DELPHI	1.74	2.50	0.46	2.26	1	1	65.9
L3	1.25	5.96	0.60	3.26	4	2	69.3
OPAL	1.91	2.16	0.62	2.43	1	1	68.9

Table 6: *Expected numbers of background events, expected numbers of signal events for a Higgs boson mass of 70* GeV/c^2*, the numbers of observed candidates and the 95% c.l. lower limit for the SM Higgs boson mass obtained by each experiment. In columns with two entries, the ones to the left are for 161 GeV and those to the right for 170-172 GeV.*

the expected numbers of signal events for m_{H}=70 GeV/c^2 and the numbers of observed candidates. In the last column, the 95% c.l. lower limit for m_{H} obtained by each experiment using only the 161 GeV and 170-172 GeV data are listed. The relatively low limit obtained by DELPHI is mainly due to a candidate event with reconstructed Higgs mass close to the limit.

3 Statistical procedures

Internally, the four collaborations use their own, preferred, statistical procedures to derive mass limits. These procedures, which are also proposed as possible schemes for combining the experiment results, have been investigated by the working group. In all cases, the procedures originally proposed have been adapted to the particular signal and background situation characterizing Higgs boson searches at LEP2.

Each statistical procedure starts with the construction of a single test-statistic [3], X, from the number and pattern of events selected. The test-statistic observed [4], X_{obs}, is compared to distributions of the same test-statistic, expected on the basis of "gedanken experiments" for particular Higgs boson mass hypotheses. These are obtained either analytically or using Monte Carlo simulation. A given mass hypothesis is excluded at the confidence level $1-CL$, where CL is the probability for gedanken experiments to obtain $X_{s+b} < X_{obs}$, and X_{s+b} is the test-statistic expected for the *signal and background* hypothesis. This is the classical, frequentist definition:

$$CL_{s+b} = P(X_{s+b} < X_{obs}). \tag{1}$$

[3] A *statistic* is a function of observables (e.g. the reconstructed Higgs boson mass or any other discriminating variable) of a population sample and is independent of population parameters. A *test-statistic* is a function of the observables and of parameters (e.g. the true Higgs boson mass) which is used in hypothesis testing.

[4] Throughout this note, the substcript "*obs*" stands for *observed* and "*exp*" for *expected* on the basis of gedanken experiments.

One should note that the above procedure is not perfect since it does not consider the possibility that a particular observation might be such that even the background, with no signal, is excluded. This shortcoming is corrected in two ways: following the Particle Data Group (PDG) prescription [9], DELPHI, L3 and OPAL use, instead of CL_{s+b}, the normalized signal confidence level

$$CL_s = \frac{P(X_{s+b} < X_{obs})}{P(X_b < X_{obs})} \tag{2}$$

while ALEPH uses

$$CL_s = P(X_s < X_{obs}). \tag{3}$$

Here X_b is the test-statistic expected assuming *background only* and X_s is that expected assuming *signal only*. The latter choice corresponds to the conservative attitude of performing no background subtraction. Both approaches allow the claimed confidence level to be reached, or a better one; however, the former approach tends to give results closer to the intended confidence level, especially if the background is relatively high.

The 95% c.l. lower limit for the Higgs boson mass is the value of the hypothetical mass which yields $CL_s = 0.05$.

There are important practical differences between the statistical combination schemes proposed by DELPHI, L3 and OPAL on one side and ALEPH on the other. While, in the DELPHI, L3 and OPAL schemes, all channels from all experiments at all energies are combined in a single step, in the ALEPH case this is done in two steps: first, each experiment combines its own channels (using its preferred statistical method); second, the *results* of the four experiments are combined. While the DELPHI, L3 and OPAL schemes allow more mutual insight, since detailed experimental information are exchanged, the ALEPH method has the virtue of simplicity since only the outcome of each experiment, in a well defined formulation, is to be exchanged. In the case of equal performances the argument of simplicity may determine the choice of the working group when existing results have to be combined using new data.

3.1 The ALEPH statistical procedure

The test-statistic in the ALEPH procedure makes use of the number of observed candidate events together with a characteristic variable x which discriminates between signal and background events. This variable can be, for instance, the reconstructed Higgs boson mass, but also any other discriminating variable obtained e.g. from a multivariate analysis. More explicitly, the test-statistic for a single channel reads

$$\mathcal{E} = \sum_{k=0}^{n} \left[\exp(-s) \frac{s^k}{k!} \right] \mathcal{P}_k^n, \tag{4}$$

where n is the number of selected candidate events, the factor inside the square braquets is the Poisson probability for k events coming from signal, and $\mathcal{P}_k^n$ is the probability for these k signal events to be less signal-like than observed, according to the density distributions of the variable x for signal and background. The exact formulation of $\mathcal{P}_k^n$ can be found in [10].

A confidence level $c(m_{\mathrm{H}})$ is determined as a function of the hypothetical Higgs boson mass m_{H}: it is the fraction of gedanken experiments, with *signal only*, for which the test-statistic is smaller than that of the real experiment.

The method also uses the expected confidence level for the null hypothesis, $\langle c\rangle_\infty(m_{\mathrm{H}})$, which is the average of $c(m_{\mathrm{H}})$ for the *background only* hypothesis. [5] The mass for which $\langle c\rangle_\infty(m_{\mathrm{H}})$ crosses the 5% level is the mass which, on average, is excluded at the 95% confidence level if the Higgs boson mass is out of reach.

3.1.1 Merging of channels without background subtraction

The procedures described below to merge different channels is proposed also for the merging of experiments. Let $c_i(m_{\mathrm{H}})$ be the confidence levels determined from the data in $i = 1, \ldots, N$ different channels, as functions of m_{H}. If there is no information regarding the intrinsic capabilities of each of the N channels (e.g. mass resolution), a new test-statistic,

$$f(m_{\mathrm{H}}) = \prod_{i=1}^{N} c_i(m_{\mathrm{H}}), \tag{5}$$

can be defined, which attributes equal "weight" to all channels. The combined confidence level is obtained analytically [10] as the fraction of gedanken experiments with *signal only* that lead to a value of f less than observed. This "democratic" approach can, however, lead to a dilution (on average) of the performance of a superior analysis by the addition of an inferior one. This situation is avoided in an "elitist" approach which attributes different weights to each channel in the merging process. The corresponding test-statistic may be written

$$f(m_{\mathrm{H}}) = \prod_{i=1}^{N} c_i^{a_i}(m_{\mathrm{H}}), \tag{6}$$

and the optimal weights ($a_i \neq 1$) are obtained by minimizing the expected combined confidence level for the null hypothesis, calculated from the individual confidence levels $\langle c_i\rangle_\infty(m_{\mathrm{H}})$. These weights guarantee that the confidence level, and hence the mass limit, is never degraded, *on average*, by the addition of a new channel. In what follows, the more realistic, elitist, approach is adopted and the right hand side of Eq. 6 is identified to be the test-statistic X in Eq's. 1-3.

3.1.2 Merging of channels with background subtraction

The use of *signal only* experiments in the ALEPH confidence level determination is conservative since it implies that no background subtraction is attempted. To perform background subtraction (in channels where the background is well known), the confidence levels $c_i(m_{\mathrm{H}})$ and $\langle c_i\rangle_\infty(m_{\mathrm{H}})$ have to be determined for the *signal+background* hypothesis instead of the *signal*

[5]In ALEPH, the optimization of the data selection is done, in fact, by minimizing $\langle c\rangle_\infty(m_{\mathrm{H}})$ with respect to the selection cuts, m_{H} being chosen close to the edge of the expected sensitivity domain.

only hypothesis. The merging of channels is done, otherwise, exactly as above. The combined confidence level is renormalized to the fraction of gedanken experiments with *background only* that lead to a test-statistic smaller than observed. This fraction, $c_i^b(m_{\mathrm{H}})$, has to be determined separately for each of the N channels, and the merged fraction is obtained as the fraction of *background only* gedanken experiments that lead to products of $(c_i^b)^{a_i}(m_{\mathrm{H}})$ smaller than measured.

3.2 The DELPHI statistical procedure

In the framework of the LEP Higgs working group DELPHI proposes a new statistical procedure [11] with respect to that used in their latest publication [3]. The new method uses the ratio of likelihoods (more correctly the ratio of probability densities) corresponding to the *signal+background* and the *background only* hypotheses. The likelihood ratio is a test-statistic commonly used in hypothesis testing [12] to maximize the probability of rejecting a false hypothesis at a given confidence level. Confidence levels are computed by a large number of Monte Carlo experiments according to the SM predictions and corresponding to the *signal+background* and the *background only* hypotheses, and counting the fraction of experiments with values of the likelihood ratio less than observed. Since the combined likelihood ratio is simply a product of the likelihood ratios of individual channels, the inclusion of new channels is straightforward.

In the SM, the Higgs boson production cross-section is a function of the mass only; therefore, the likelihood ratio, $Q(m_H)$, for experiments with N channels and with measurements of a single discriminating variable for each candidate, is written as

$$Q(m_{\mathrm{H}}) = \frac{\prod_{i=1}^{N} \frac{e^{-(s_i(m_H)+b_i)}(s_i(m_H)+b_i)^{n_i}}{n_i!}}{\prod_{i=1}^{N} \frac{e^{-b_i} b_i^{n_i}}{n_i!}} \; \frac{\prod_{j=1}^{n_i} \frac{s_i(m_H)S_i(m_H,m_{ij})+b_i B_i(m_{ij})}{s_i(m_H)+b_i}}{\prod_{j=1}^{n_i} B_i(m_{ij})} \tag{7}$$

which can be simplified to

$$Q(m_{\mathrm{H}}) = e^{-s_{tot}(m_H)} \prod_{i=1}^{N} \prod_{j=1}^{n_i} \left(1 + \frac{s_i(m_H)S_i(m_H, m_{ij})}{b_i B_i(m_{ij})}\right). \tag{8}$$

Here n_i is the number of observed candidates in channel i, m_{ij} is the invariant mass (or other discriminating variable) of candidate j, $s_i(m_H)$ and b_i are the integrated signal and background rates per channel, s_{tot} is the total signal rate for all channels, and $S_i(m_H, m_{ij})$ and $B_i(m_H)$ are the probability distribution functions (p.d.f.-s) for the signal and background, respectively.

The p.d.f.-s S_i and B_i can be extended to include other information about the candidates which help distinguish signal from the background (such as b-tagging probabilities, production angles, etc.) resulting in lower expected confidences in the absence of a Higgs boson signal and more significant discoveries in the presence of an observable signal, at the cost of increased complexity of the analysis to treat the correlations between the different variables. Some of the properties inherent to this method are:

- The procedure simplifies directly to the PDG formula [9] for counting experiments with only one channel.
- For large event samples, $-2\ \ln(Q)$ is well approximated by changes in χ^2, meaning that confidence levels can also be extracted from standard tables or computer functions of χ^2 probabilities.
- Only slight changes to the method are required to carry out a maximum likelihood analysis of the Higgs boson mass and cross-section in the presence of a signal, what is required to compute confidence levels for discovery (i.e. the probability that background processes alone cannot account for the observation).
- Monte Carlo integration techniques provide a simple and direct manner for including uncertainties on the experimental inputs.

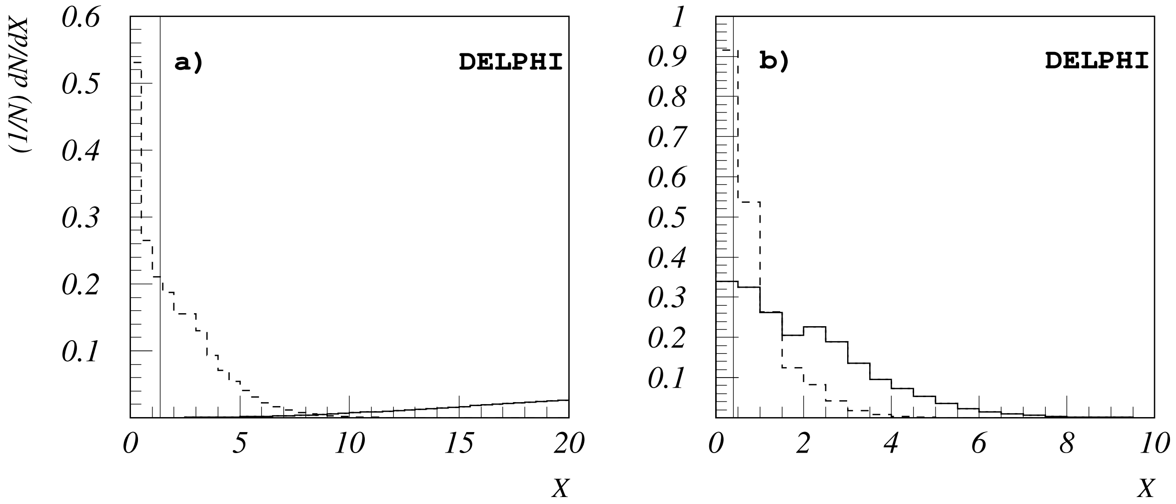

Figure 9: *The probability distributions for* $X = \ln(Q(m_H)) + s$ *for signal+background (solid line) and background only experiments (dashed line) for (a)* m_{H}=*45* GeV/c^2 *and (b)* m_{H}=*77* GeV/c^2*, for the DELPHI Higgs boson search. The vertical lines indicate the experimental result* $X_{obs}(m_{\mathrm{H}})$*. The areas under the curves to the left of the experimental values are the confidence levels* CL_{s+b} *(solid line) and* CL_b *(dashed line) in Eq's. (1-3).*

A simple derivation shows that the likelihood ratio method is effectively based on counting weighted events. Since $P(Q \leq Q_{obs}) = P(\ln(Q) \leq \ln(Q_{obs}))$ we can write

$$\ln(Q) = -s_{tot} + \sum_{k=1}^{n} n_k w_k \tag{9}$$

where n is the total number of events observed in all channels and the weight for each candidate k is given by

$$w_k = \ln\left(1 + \frac{s_k}{b_k}\frac{S_k(M_h, m_k)}{B_k(m_k)}\right). \tag{10}$$

(Note that the weight attributed to candidate k depends on the channel in which the candidate is observed.) Since the constant s_{tot} appears on both sides of the expression $\ln(Q) \leq \ln(Q_{obs})$, the test-statistic X in Eq.'s 1-3 can be identified to be $X = ln(Q) + s_{tot}$ for this method. It is illustrated in Figure 9 where the p.d.f.-s of $\ln(Q) + s(m_H)$, emerging from the DELPHI searches at 161 GeV and 170-172 GeV energies, are shown for two extreme mass hypotheses, $m_{\rm H}$=45 GeV/c^2 and 77 GeV/c^2, along with the experimental observation. The confidence levels CL_{s+b} and CL_b, for the *signal+background* and *background only* hypotheses, are obtained from the figure by integrating the relevant p.d.f.-s over the range below the observed value. From these examples it appears that the DELPHI search easily excludes a Higgs boson with 45 GeV/c^2 mass while for $m_{\rm H}$=77 GeV/c^2 the *signal+background* and *background only* distributions are rather similar, not allowing any conclusion to be drawn from the DELPHI experiment alone.

3.3 The L3 statistical procedure

In the method proposed by L3, the test-statistic used [13] is an extension of the PDG prescription [9] to the case where several channels are combined, with background subtraction [14]. The combined likelihood function for the *signal + background* hypothesis can be written as

$$F(s,b) = e^{-(s+b)} \prod_{j=1}^{k} \frac{(s \cdot f_j + b \cdot g_j)^{n_j}}{n_j!}.$$

The index j runs over all k channels; f_j, g_j are the fractions of the total signal s and total background b, respectively, and n_j is the number of candidate events which fall into channel j.

The test-statistic is the Bayesian confidence level:

$$X = \frac{\int_\sigma^\infty F(s,b)ds}{\int_0^\infty F(s,b)ds} \tag{11}$$

where σ is the number of expected signal events, depending on the Higgs boson mass hypothesis. The signal confidence level CL_s is obtained according to Eq. (2).

It has been shown [15] that the inclusion of channels with low signal-to-background ratio (S/B) may improve considerably the overall sensitivity, provided that background subtraction is performed and the uncertainties which affect the expected signal and background are not excessive. This is demonstrated in the example below where the number of expected signal events is fixed to $\sigma = 5$ and in which the signal and background expectations per channel are listed in Table 7; the channels are ordered with decreasing S/B. Figure 10 shows the average, expected, signal confidence level $< CL_s >$ obtained when channels are added one by one, starting with the channel with the highest S/B ratio. When background is not subtracted (curve (a)), the highest sensitivity (lowest $< CL_s >$) is obtained when only channels with S/B>1 are included. If background subtraction is performed (curves (b) to (e)), the improvement of sensitivity persists when channels with lower S/B ratios are included, provided that the errors attached to the signal efficiency and the residual background are not excessive (a degradation of the sensitivity is observed when both errors are 40%, see curve (c)).

Channel	Expected signal	Expected background	S/B
1	1.5	0.1	15.000
2	1.0	0.2	5.000
3	0.5	0.5	1.000
4	0.5	1.0	0.500
5	0.5	2.0	0.250
6	0.5	5.0	0.100
7	0.25	10.0	0.025
8	0.15	10.0	0.015
9	0.10	10.0	0.010

Table 7: *Example of a search in 9 channels, with a total of 5 signal events expected. The channels are ordered according to the signal to background ratio (last column). The expected numbers of signal and background events per channel are listed.*

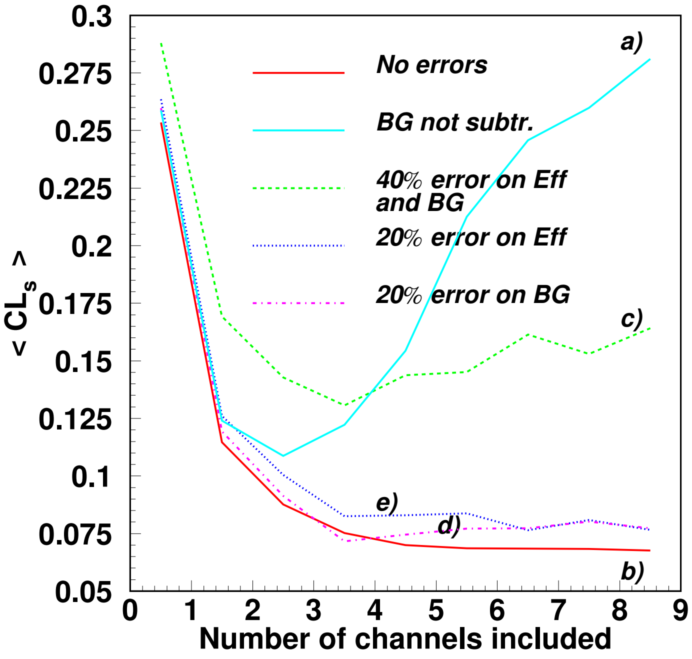

Figure 10: *Average confidence level* $< CL_s >$ *as a function of the number of channels included. The curves show the behaviours in the cases of (a) no background subtraction, (b) background subtraction with no errors assigned to the expected signal efficiencies and residual backgrounds and (c-e) various errors assigned to them, as indicated in the figure.*

In the L3 event selection [4] background subtraction is performed and cuts are applied which limit the signal-to-background ratio to values larger than 0.15 in all channels, for hypothetical Higgs boson masses up to 80 GeV/c^2.

3.4 The OPAL statistical procedure

The OPAL statistical method is based on fractional event counting [16]. A weight is assigned to each selected candidate, i, according to a filter function $w_i = F(m_i)$. The argument m_i is the reconstructed Higgs boson mass in the event; however, any other discriminating observable can be used instead. The weights attributed to candidate events are added in all channels resulting in a total weight $w = \sum_i w_i$, which is the test-statistic X in Eq's. (1-3).

The filter function, $F(m_i)$, depends on the hypothetical Higgs boson mass, m_{H}; therefore, the total weight w, is also a function of m_{H}: it becomes large at values of m_{H} where candidate events cluster. The probability distribution of the total weight, $P(w)_{s+b}$, including signal and background events, forms the basis for the evaluation of confidence levels.

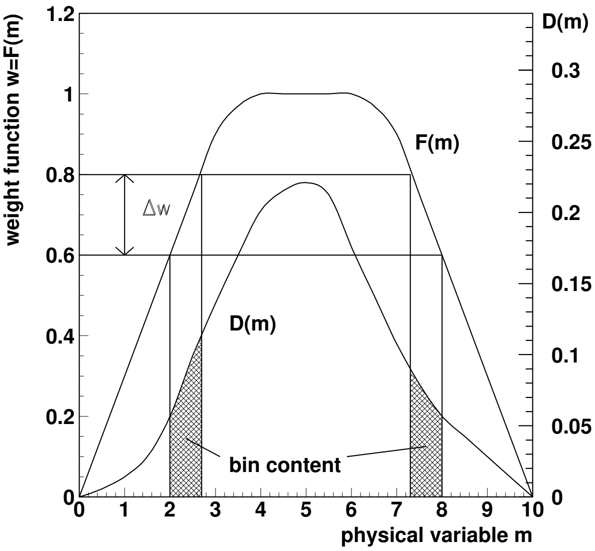

Figure 11: *Relation between the mass distribution $D(m)$, the filter function $F(m)$, and the probability $P_1(w)\Delta w$ (hatched area) for one candidate event to have a weight $w \pm \Delta w/2$. In the example chosen $w \pm \Delta w/2 = 0.7 \pm 0.1$.*

In principle, the filter function $F(m)$ is arbitrary. The construction of the weight distribution function, $P_1(w)$, for the case of one signal event and no background event, is illustrated in Figure 11, assuming a specific form for $F(m)$. In the figure, the function $D(m)$ describes the actual mass distribution and the hatched areas indicate the probability $P_1(w)\Delta w$ for a Higgs particle to yield a weight inside an interval Δw. The $P(w)$ distributions for fixed numbers of signal events and no background event can be obtained from $P_1(w)$ by iterative integration, and the total probability density $P_s(w)$ is obtained by adding the whole set of functions, weighted with the Poisson distribution for the expected number of Higgs bosons. The distribution $P_s(w)$ is thus sensitive to the Higgs boson production rate.

An analogous procedure leads to the w distribution, $P_b(w)$, of background events, which can be combined with the signal distribution, $P_s(w)$, to evaluate the overall distribution, $P_{s+b}(w)$. The function $P_{s+b}(w)$ includes then the background predicted by the SM and depends on the signal rate which is a function of m_{H}.

The confidence level is computed, in accordance with Eq. (2), as

$$CL_s = \frac{\int_0^{w_{obs}} P_{s+b}(w)dw}{\int_0^{w_{obs}} P_b(w)dw}. \tag{12}$$

The denominator is introduced to avoid situations with excessively low limits arising when less candidates are observed than expected, see Section 3.

The filter function is chosen to provide good discrimination between signal and background events. For a single decay channel and a flat background distribution it is adequate to use the mass distribution itself: $F(m) = \frac{D(m)}{D_{max}}$, where D_{max} is the maximum value of $D(m)$. When several channels with different efficiencies and backgrounds are to be combined, a channel-dependent filter function is introduced to account for the differing signal-to-background ratios:

$$F_k(m) = K \cdot \frac{1}{C + \frac{b_k(m)\cdot\sum_k S_k}{D_{k,max}\cdot S_k}} \cdot \frac{D_k(m)}{D_{k,max}}. \tag{13}$$

Here, S_k is the expected signal rate in channel k and b_k the differential background rate. The factor K is chosen such as to fix the largest value of $F_k(m)$ to unity. The constant C is a free parameter which can be optimized by minimizing the signal confidence level in *background only* gedanken experiments, as a function of the Higgs boson mass hypothesis.

The above filter function has the following properties:

- In the limit of very low background, $F_k(m)$ approaches a universal value K/C at the mass peak positions, independently of the channel number k;
- $F_k(m)$ is invariant under arbitrary splitting of any decay channel;
- If the background is large compared to the signal, $F_k(m)$ is proportional to the local signal to background ratio;
- For C equal to the total number of expected events, the maximum of $F_k(m)$ is proportional to the probability for a candidate in channel k to be a signal event.

3.5 Discussion, comparative tests

Since each of the four statistical methods uses different test-statistics, the results are expected to deviate slightly from each other even though the same input is used. Exceptions are cases where no candidates are observed and where the results should be identical since all methods boil down to simple event counting.

Before applying the proposed methods to the data, the following properties have been checked:

1. If a channel is subdivided, or if the order in combining the channels is changed (relevant only for the ALEPH method since the DELPHI, L3 and OPAL methods treat all channels simultaneously), the result is invariant;
2. Candidate events which have a reconstructed Higgs boson mass (or other discriminating variable) highly incompatible with a specific Higgs mass hypothesis do not significantly alter the limit on the signal rate at that hypothetical mass;
3. The quoted 95% c.l. limits obtained are true *at least*. They may be slightly conservative but never optimistic on average (considering a large number of gedanken experiments).

The last two properties are discussed below on the basis of fictive Monte Carlo experiments.

Systematic errors on detection efficiencies and background levels can be considered in various ways. For the size of the effects, ALEPH quotes a shift of -10 MeV/c^2 of the mass limit when incorporating the error on the signal detection efficiency according to the convolution method of Ref. [17]; OPAL observes a shift of -300 MeV/c^2 when decreasing the predicted background by one standard deviation while L3 obtains a shift of -200 MeV/c^2 when incorporating the systematic error on the background in the generation of the gedanken experiments (the shift due to the error on the signal detection efficiency is found by L3 to be much smaller). Finally, DELPHI observes a shift smaller than 50 MeV/c^2 in their mass limit when incorporating uncertainties on the background rates and signal efficiencies using the same method as L3.

3.5.1 A fictive experiment

A fictive experiment has been considered in which three events are observed, with reconstructed Higgs boson masses of 34, 35 and 55 GeV/c^2, while the background, fixed at 4 events in total, is assumed to be flat in mass between zero and 100 GeV/c^2. The mass resolution is assumed to decrease linearly from 10.5 GeV/c^2 at 10 GeV/c^2 mass to 3.3 GeV/c^2 at 80 GeV/c^2 mass.

Figure 12 shows the 95% c.l. "experimental" upper bound for the number of signal events, as a function of the hypothetical Higgs boson mass, obtained by the four methods.

In the cases of the ALEPH and OPAL schemes, the "peaks" at the observed masses are higher than in the cases of the DELPHI and L3 schemes; however, they are also sharper, thus

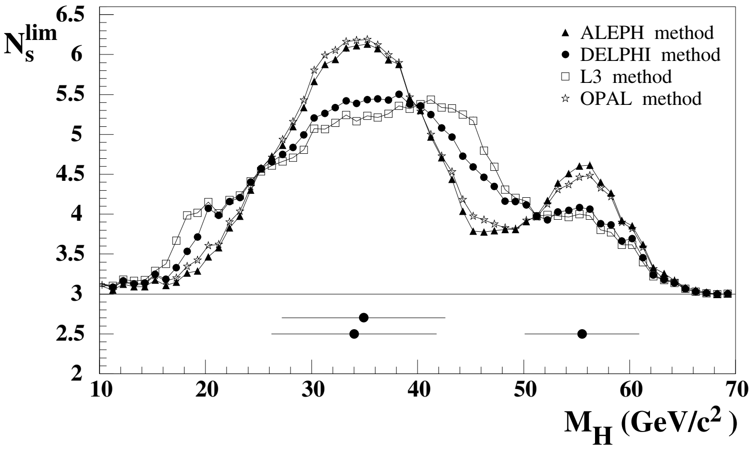

Figure 12: *The 95% c.l. "experimental" upper bound for the number of signal events, as a function of the hypothetical Higgs boson mass, obtained by the four methods.*

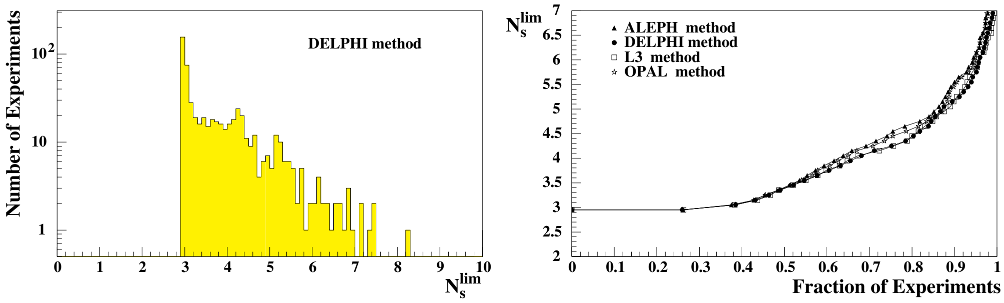

Figure 13: *The distribution of limits obtained for the Higgs boson signal rate, N_s^{lim} in 600 gedanken experiments, for m_{H} fixed arbitrarily to 40 $\mathrm{GeV/c^2}$. Left hand side: distribution obtained using the DELPHI method as an example; Right hand side: the integrated distributions, for all four methods, which is the fraction of cases in which this limit or a better one is set.*

affecting the limit in a more restricted mass range. These local differences, of the order of 15%, increase slightly when the background is decreased.

It is important to compare limit algorithms not by the results which they produce in any particular case but on average. In order to do that, six hundred Monte Carlo experiments were made, each with randomly chosen events from the background distribution of Figure 12. In each case the 95% c.l. limit on the signal event rate, at $m_{\rm H}$=40 GeV/c^2, was calculated using all four algorithms. These limits are plotted (from the DELPHI algorithm only, for clarity) in Figure 13 (a). The minimum rate limit is at 3 events, from cases where the influence of candidates is small, while the peak at around 4.3 events corresponds to being close to one candidate. To compare the four algorithms, the integrated version of this plot is also made, in Figure 13 (b), which shows the probability of getting a rate limit smaller than any particular value. The two likelihood-based algorithms, of DELPHI and L3, set limits on the signal rate which are lower by about 3%, on average, than the ones set by the ALEPH and OPAL methods.

3.5.2 False exclusion of an existing signal

The probability for falsely excluding a signal which is actually existing has also been compared for the four methods. Figure 14 shows this false exclusion probability, for $m_{\rm H}$=77 GeV/c^2, as a function of the number of signal events expected. The curves obtained by the four methods

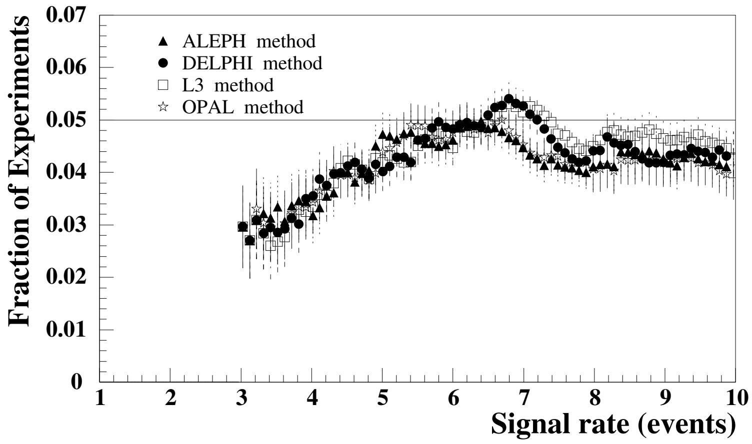

Figure 14: *The fraction of gedanken experiments in which an existing Higgs boson, with 77* GeV/c^2 *mass, is actually excluded by the four statistical methods, as a function of the signal event rate.*

have very similar behaviours. With the strict frequentist definition of Eq. 1 all curves would be flat at 5%.

3.5.3 Summary of the tests

The Monte Carlo investigations did not reveal any unexpected, large, differences between the four statistical schemes, and the working group could not identify any reason to invalidate one method or the other. It is still debated whether there might be small improvements of the methods; but it seems clear that potential gains in exclusion power are rather small. So far the working group has only examined the methods regarding their abilities of producing accurate combined exclusion limits; the performances in combined discovery, that is in claiming the correct confidence level for a signal, have not yet been considered in detail.

4 Combined mass limits

In this section we describe the derivation of combined mass limits for the SM Higgs boson using the four proposed statistical procedures.

As explained earlier, the ALEPH method analytically combines the *results* from individual experiments without recalculating them. The inputs required for the elitist prescription with background subtraction are the following:

- the *measured* confidence level, $c(m_{\mathrm{H}})$, obtained by each experiment with its own, preferred, statistical method to combine channels, with respect to the *signal+background* hypothesis, and the number of expected signal events, $s(m_{\mathrm{H}})$, as a function of the hypothetical Higgs boson mass m_{H}. This information is needed for the combination of the four confidence levels;
- the *expected* confidence level, $\langle c\rangle_{\infty}(m_{\mathrm{H}})$, obtained by each experiment with its own statistical method, with respect to the *signal+background* hypothesis, and the number of expected background events, $b(m_{\mathrm{H}})$, as a function of the hypothetical Higgs boson mass. This information is needed for the weight optimization in the elitist prescription;
- from each experiment, the fraction of *background only* gedanken experiments less signal-like than observed, $c^b(m_{\mathrm{H}})$, as a function of the Higgs boson mass. This is needed for the renormalization of the confidence levels, using background subtraction.

The DELPHI, L3 and OPAL methods consider each channel of each experiment independently and combine them in a single step. These methods require from the experiments detailed, channel-by-channel, input information. Using the notations of Section 3.4, these inputs are:

- the numbers of expected signal events, S_k, in each channel k;
- the mass distributions (or those of the discriminating variable chosen) for the signal, $D_k(m)$, in the form of histograms or, in case of Gaussian mass distributions, the resolutions in each channel;

- the normalized mass distributions (or those of the discriminating variable chosen) for the background, $b_k(m)$;
- the list of selected candidates, with their reconstructed Higgs boson masses (or discriminating variable values) and the channels where they are observed.

Each of the four experiments have provided full information allowing individual and combined limits to be calculated with any of the proposed methods. The results are summarized below, separately for the four methods.

The DELPHI, L3 and OPAL methods generate, as an intermediate step, the confidence levels for the *signal+background* and the *background only* hypotheses which are also used as inputs to the ALEPH method. The comparison of these, experiment by experiment and method by method, provides an important check of consistency. Such a comparison is performed in Section 5 on the *observed* confidence levels (fraction of gedanken experiments with test-statistics smaller than observed) as well as on the *expected* confidence levels (averaged over a large number of gedanken experiments for the *background only* hypothesis).

4.1 Combined limit using the ALEPH method

The inputs delivered by the four experiments to the ALEPH method are shown in Figures 15, 16 and 17. Figure 15 displays the expected numbers of signal and background events. Figures 16 and 17 show the observed and expected confidence levels versus m_{H} for *signal+background* and *background only* hypotheses. In the lowest parts of Figures 16 and 17, the combined confidence levels obtained using the ALEPH method are shown.

Experiment	Exp. lim. (GeV/c^2)	Obs. lim. (GeV/c^2)	$CL_b(m_H^{95})$ (%)
LEP	75.7	77.9	29.8

Table 8: *Expected and observed 95% c.l. lower limits on the SM Higgs boson mass from the four experiments combined using the ALEPH method. Column 3 is the confidence level for the background only hypothesis at the observed mass limits.*

Figure 18 shows for each experiment separately and for the four experiments combined the confidence levels after renormalization by the fraction of *background only* experiments that would yield a test-statistic less signal-like than observed (see Section 3.1.2). The intersections of the curves with the 5% horizontal lines define the observed and expected 95% c.l. combined mass limits obtained using the ALEPH method (see also Table 8). It should be noted that the expected combined mass limit is obtained analytically using an *a priori* shape for the CL distributions of *background only* experiments [10], and is therefore an approximation of the expected mass limit.

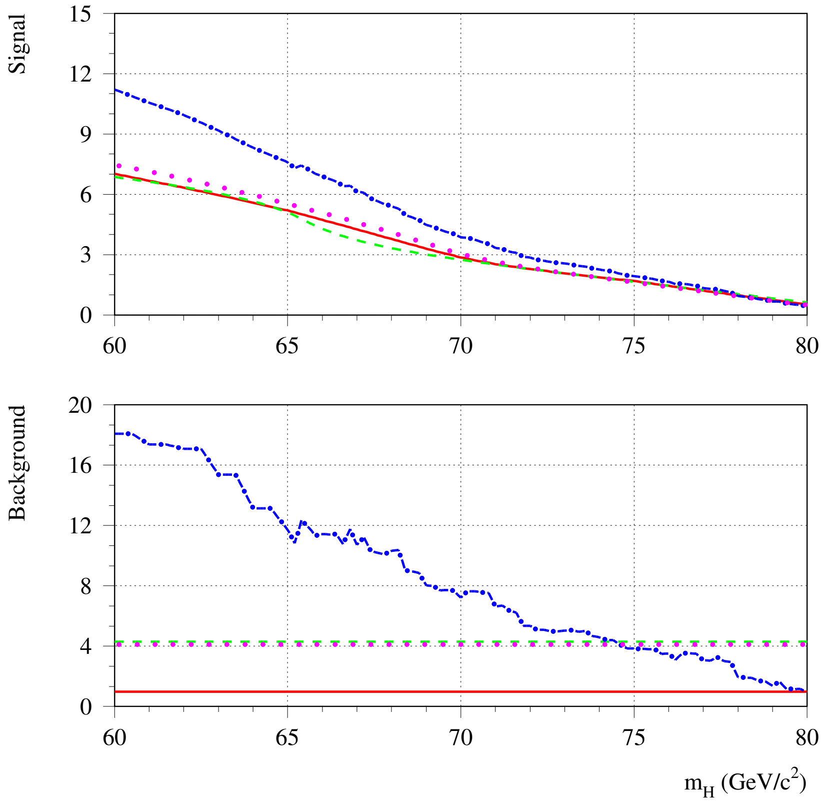

Figure 15: *Expected numbers of signal and background events, inputs to the ALEPH combination scheme, from ALEPH (full lines), DELPHI (dashed), L3 (dash-dotted) and OPAL (dotted lines), as functions of the Higgs boson mass.*

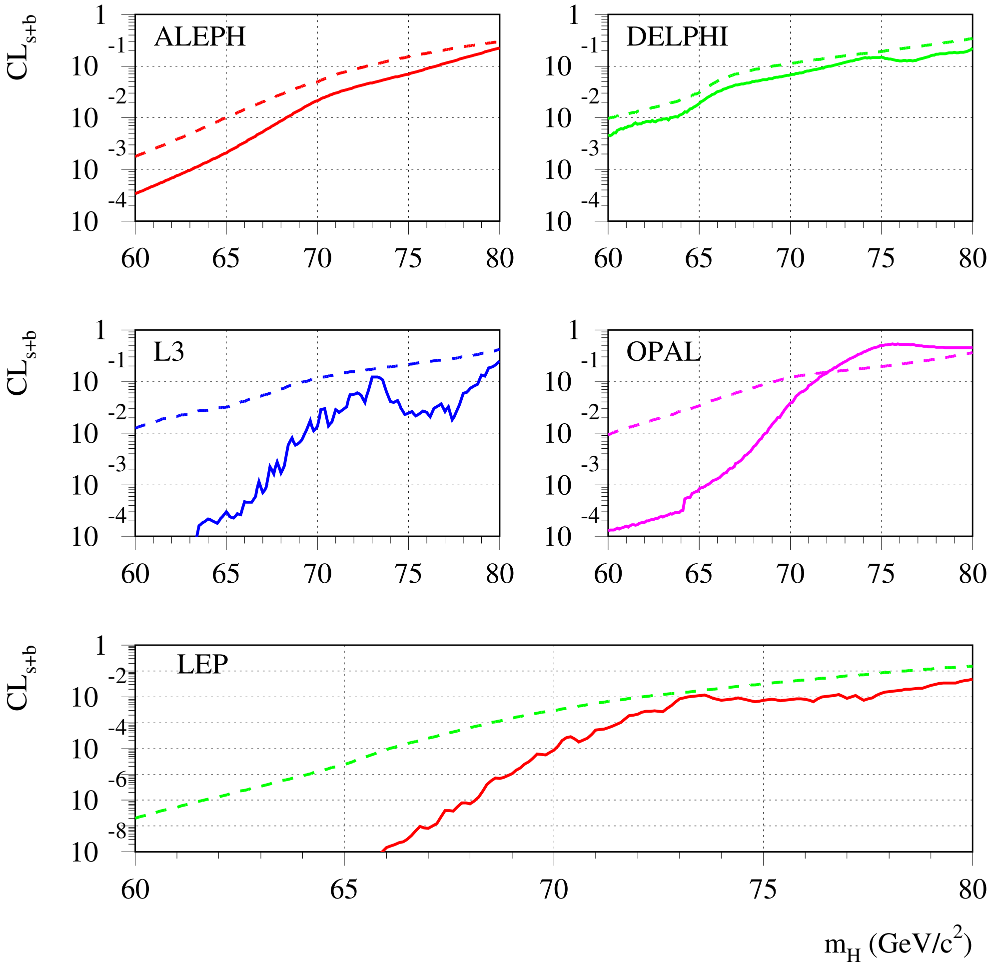

Figure 16: *Observed (full line) and expected (dashed line) confidence levels for the signal+background hypothesis: inputs provided by the four experiments, and the combined confidence levels (lower part) obtained using the ALEPH method.*

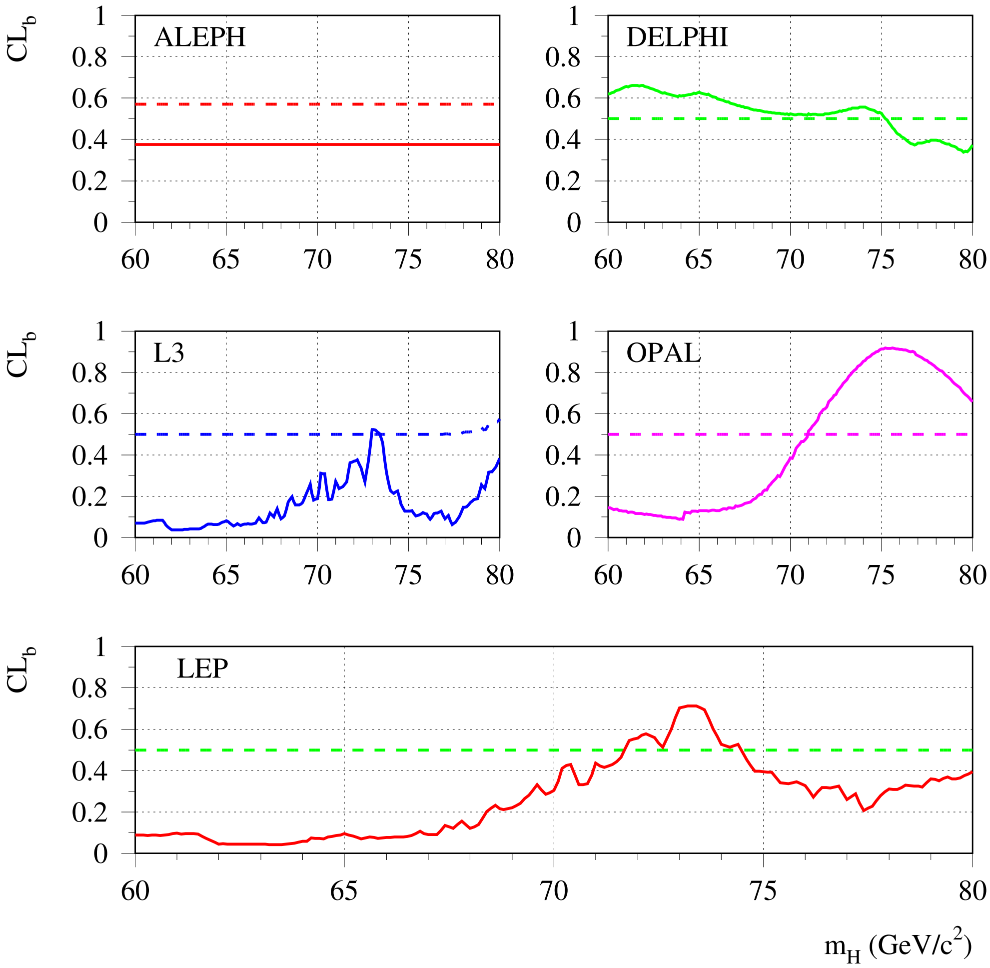

Figure 17: *Observed (full line) and expected (dashed line) confidence levels for the background only hypothesis: inputs provided by the four experiments, and the combined confidence levels (lower part) obtained using the ALEPH method.*

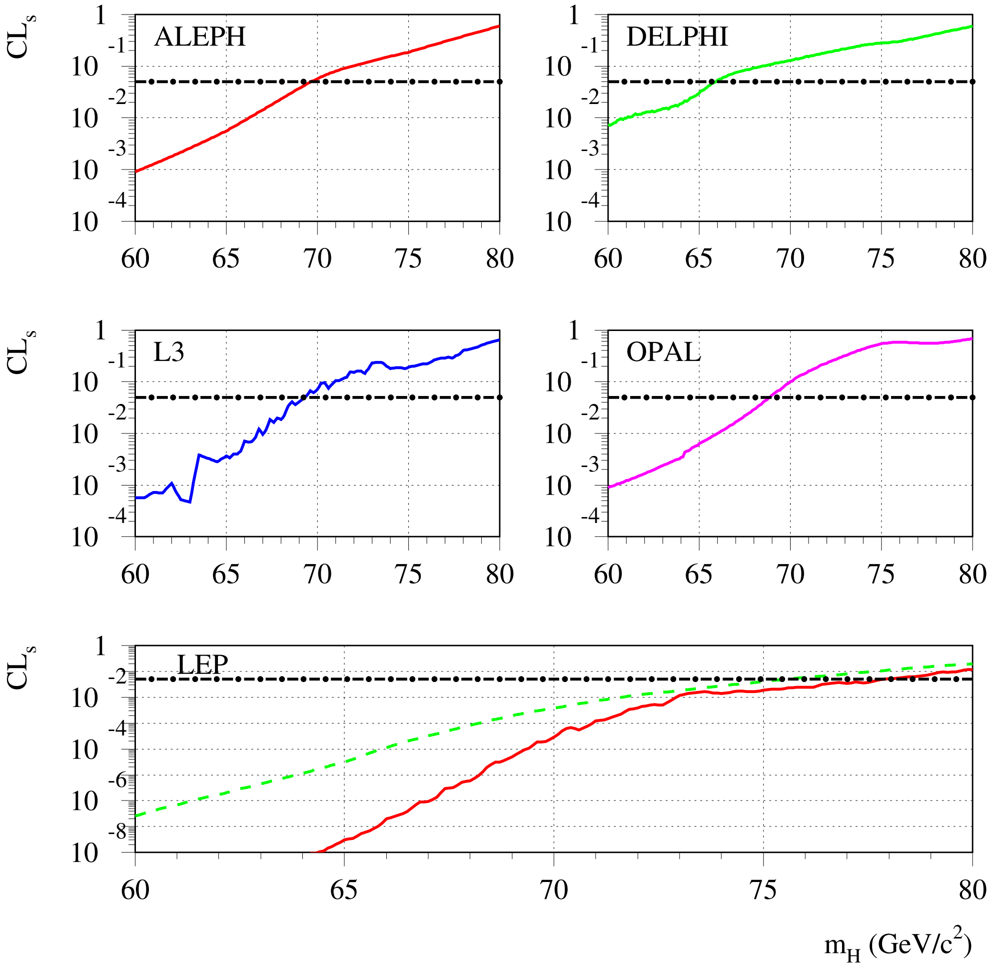

Figure 18: *Confidence levels, as a function of the Higgs boson mass, after renormalization by the fraction of gedanken experiments which yield a test-statistic less signal-like than observed, in individual experiments and in LEP combined, using the ALEPH method. The intersections of the horizontal lines at 5% with the curves define the expected and observed 95% c.l. lower limits on the Higgs boson mass.*

Additional outputs of the ALEPH method are the exponents a_i in Eq. 6 which depend on the hypothetical mass. They are related to the expected signal-to-background ratio and their values, between zero and one, regulate the probability for a candidate selected in channel i to be a signal event. In the vicinity of the combined LEP mass limit these weights are about 50% lower for DELPHI, L3 and OPAL than for ALEPH. The optimal values of a_i have also been shown [10] to be insensitive to the assumed shape of the CL_{s+b} distribution.

4.2 Combined limit using the DELPHI method

The inputs from the four experiments are used in the DELPHI likelihood ratio method to derive observed and expected confidence levels, versus m_H, for the *signal+background* and *background only* hypotheses. These are shown in Figures 19 and 20, for each experiment separately and for the four experiments combined. The information pertaining to the DELPHI experiment is in fact part of the DELPHI input to the ALEPH method.

Fig. 21 shows the ratio $CL_s \equiv CL_{s+b}/CL_b$ versus m_{H} which is the modified frequentist confidence level for *signal only* experiments. The intersections of these ratios with the lines at 5% define the expected and observed 95% c.l. lower limits on the SM Higgs boson mass, for the four experiments separately and for the four experiments combined. These are listed in Table 9, together with the observed confidence level for the *background only* hypothesis, at the mass limits set for the individual experiments and at the combined LEP limit.

Experiment	Exp. lim. (GeV/c^2)	Obs. lim. (GeV/c^2)	$CL_b(m_H^{95})$ (%)
ALEPH	68.5	69.6	48.6
DELPHI	65.4	65.9	60.2
L3	66.1	69.4	21.3
OPAL	65.9	69.0	25.8
LEP	75.8	77.5	35.1

Table 9: *Expected and observed 95% c.l. lower limits on the SM Higgs boson mass for the individual experiments and for the four experiments combined, based on the likelihood ratio method used by DELPHI. Column 3 is the confidence level for the background only hypothesis at the observed mass limits.*

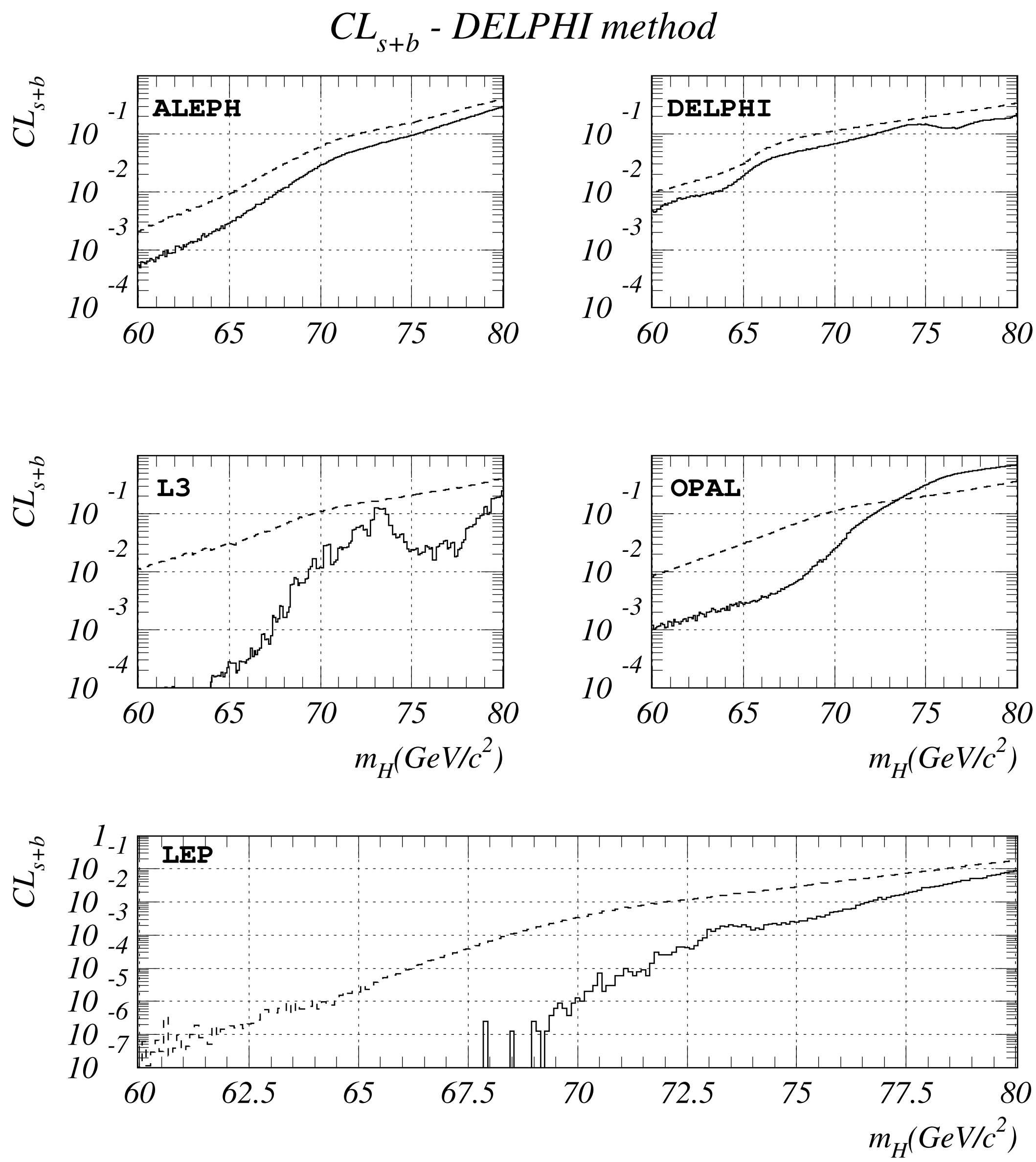

Figure 19: *Observed (full line) and expected (dashed line) confidence levels for the signal+background hypothesis, as functions of the Higgs boson mass, for the individual experiments and for the four experiments combined, based on the DELPHI method.*

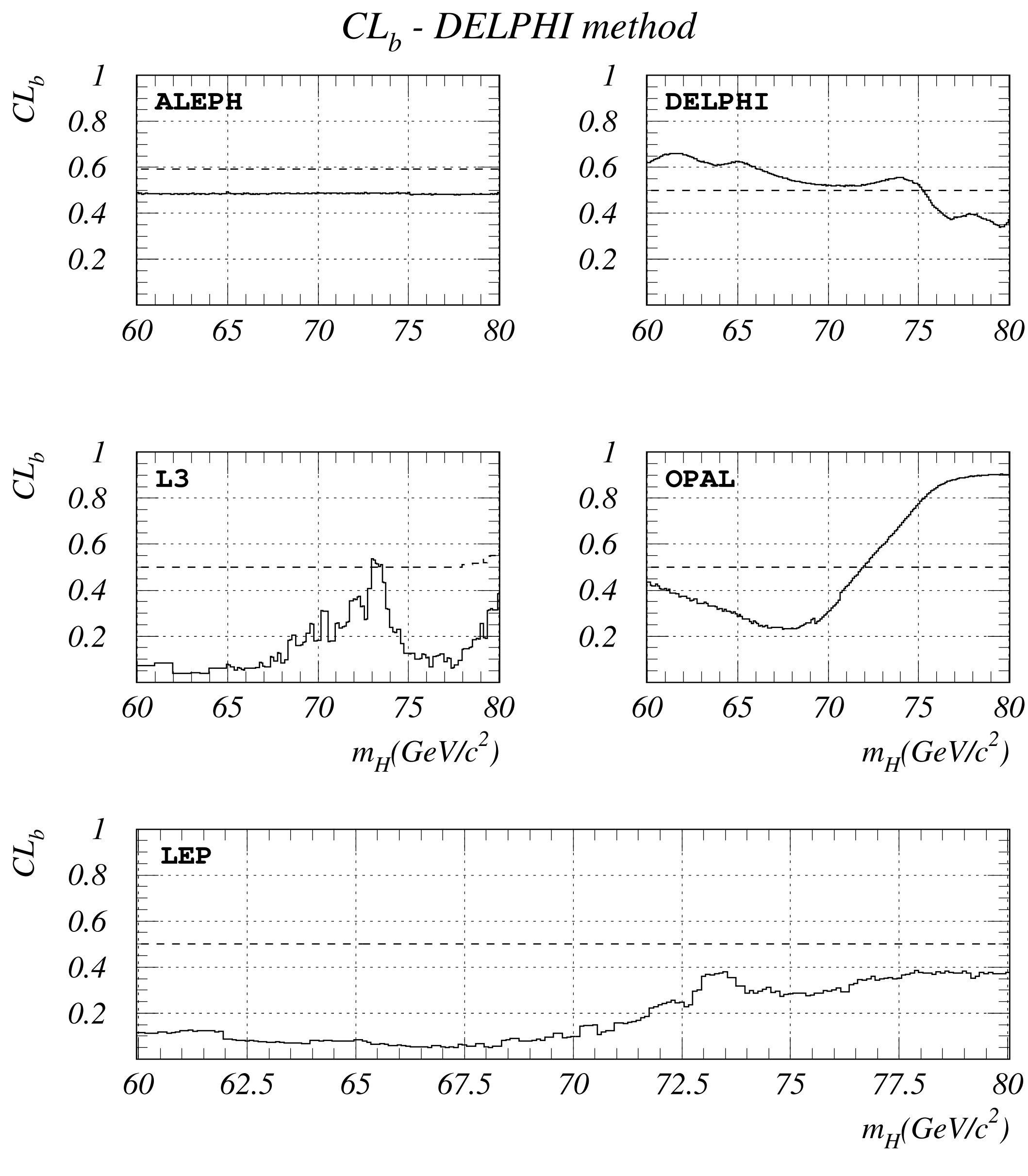

Figure 20: *Observed (full line) and expected (dashed line) confidence levels for the background only hypothesis, as functions of the Higgs boson mass, for the individual experiments and for the four experiments combined, based on the DELPHI method.*

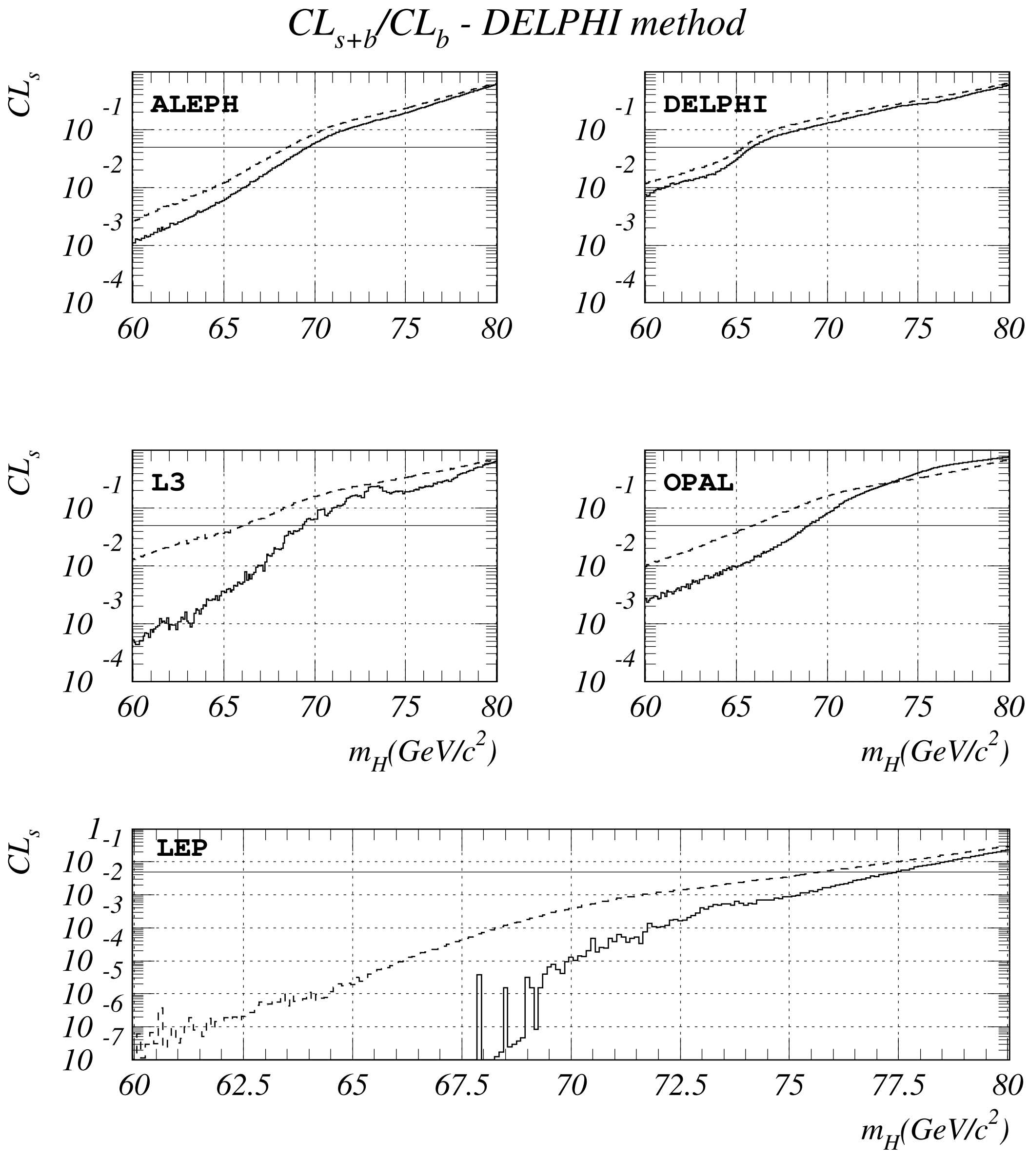

Figure 21: *The observed (full line) and expected (dashed line) confidence level ratios, $CL_s \equiv CL_{s+b}/CL_b$, for the individual experiments and for the four experiments combined, based on the method used by DELPHI. The intersections of the horizontal lines at 5% with the curves define the 95% c.l. lower limits on the Higgs boson mass.*

4.3 Combined limit using the L3 method

The inputs from the four experiments are used in the L3 method based on the Bayesian confidence level to derive observed and expected confidence levels, versus m_H, for the *signal+background* and *background only* hypotheses. These are shown in Figures 22 and 23, for each experiment separately and for the four experiments combined. The information pertaining to the L3 experiment is in fact part of the L3 input to the ALEPH method.

Fig. 24 shows the ratio $CL_s \equiv CL_{s+b}/CL_b$ versus m_{H} which is the confidence level for *signal only* experiments. The intersections of these ratios with the lines at 5% define the expected and observed 95% c.l. lower limits on the SM Higgs boson mass, for the four experiments separately and combined. These are listed in Tab. 10 together with the observed confidence level for the *background only* hypothesis, at the mass limits set for the individual experiments and at the combined LEP limit.

Experiment	Exp. lim. (GeV)	Meas. lim. (GeV)	$CL_b(m_H^{95})$ (%)
ALEPH	68.8	69.6	52.4
DELPHI	65.3	65.9	60.6
L3	65.7	69.3	19.2
OPAL	65.6	68.6	30.8
LEP	76.0	77.8	34.6

Table 10: *Expected and observed 95% CL lower limits on the SM Higgs boson mass for the individual experiments and for the four experiments combined, based on the Bayesian likelihood ratio method used by L3. Column 3 is the confidence level for the background only hypothesis at the observed mass limits.*

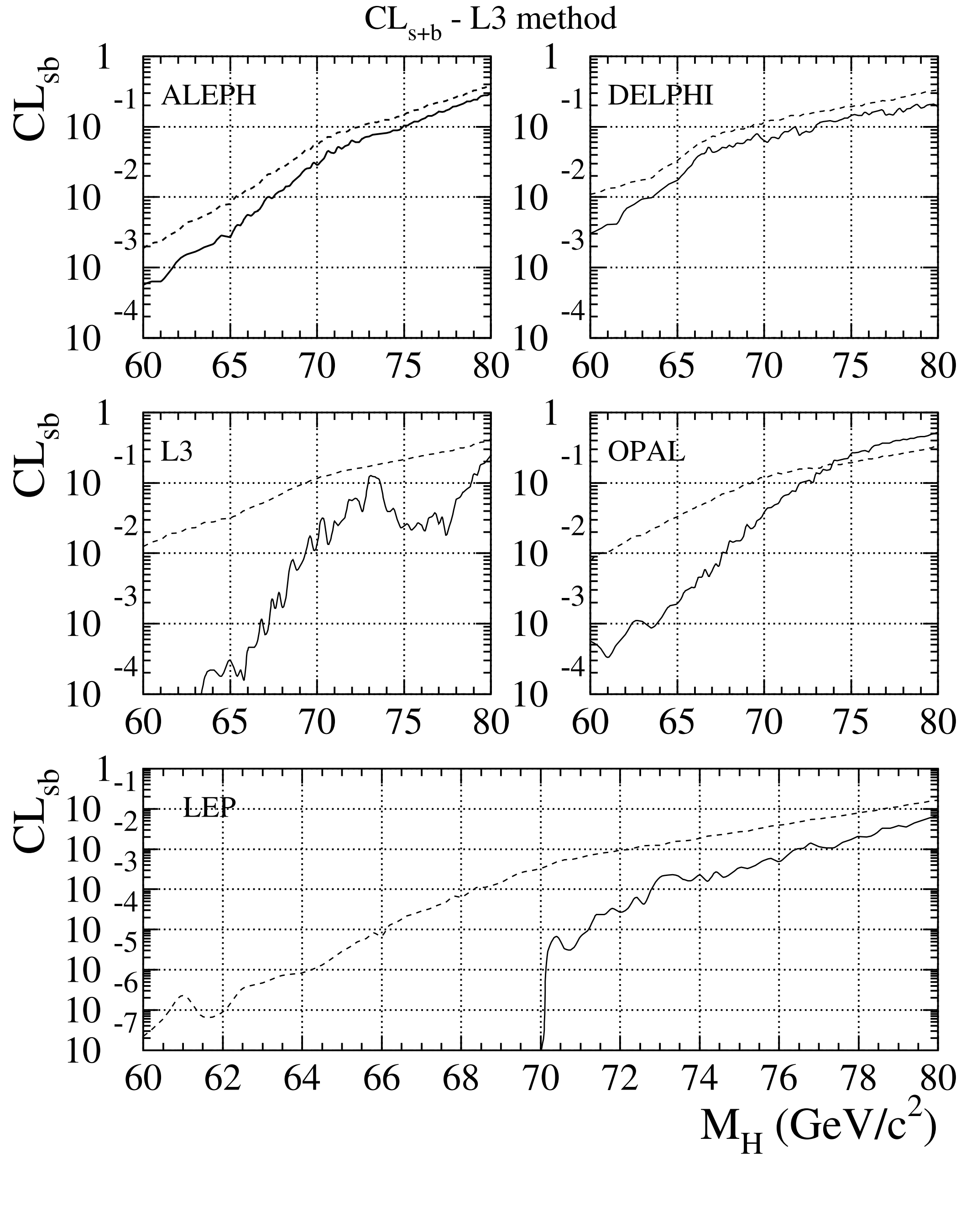

Figure 22: *Observed (full line) and expected (dashed line) confidence levels for the signal+background hypothesis, as functions of the Higgs boson mass, for the individual experiments and for the four experiments combined, based on the L3 method.*

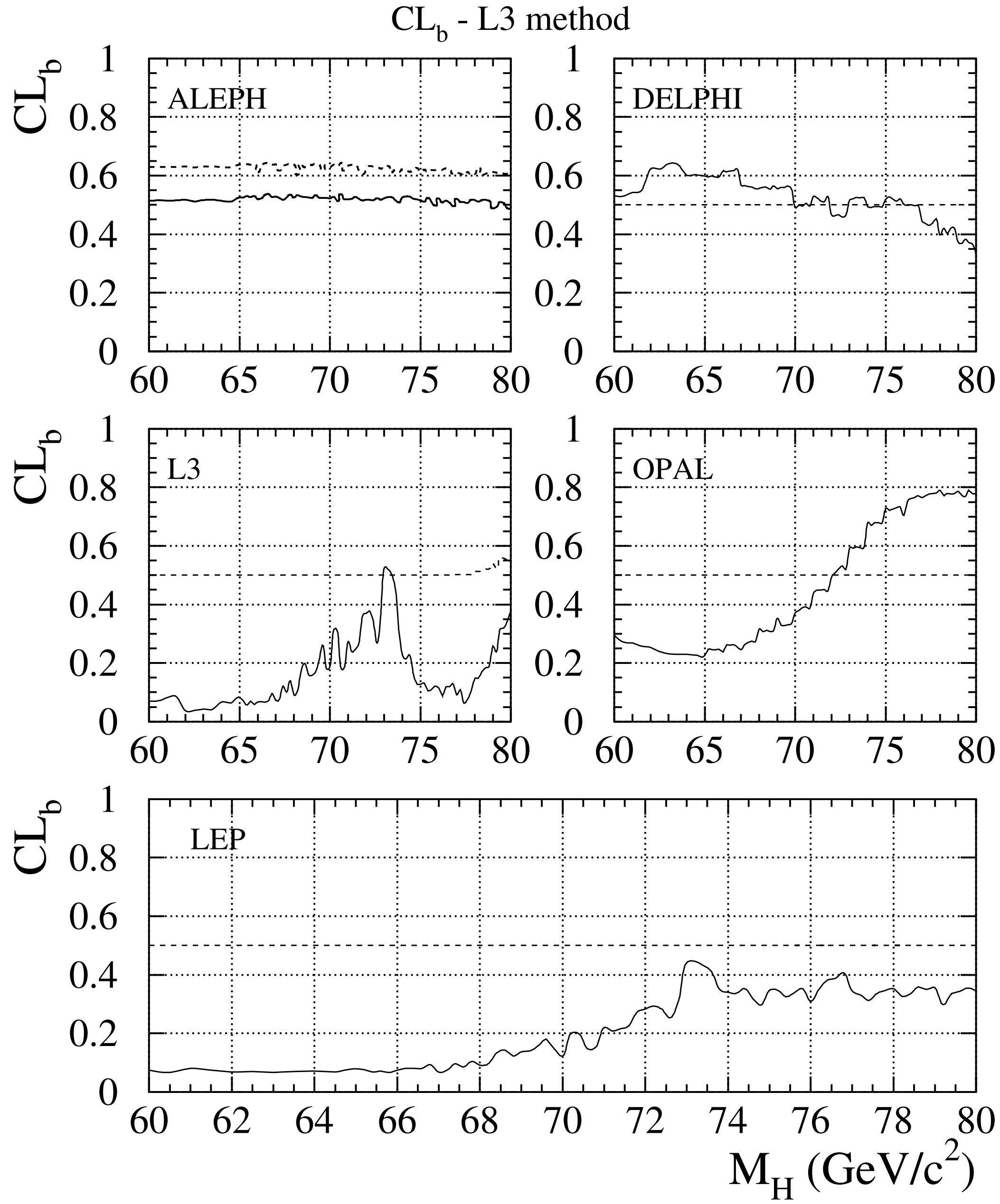

Figure 23: *Measured (full line) and expected (dashed line) confidence levels for the background only hypothesis, as functions of the Higgs boson mass, for the individual experiments and for the four experiments combined, based on the L3 method.*

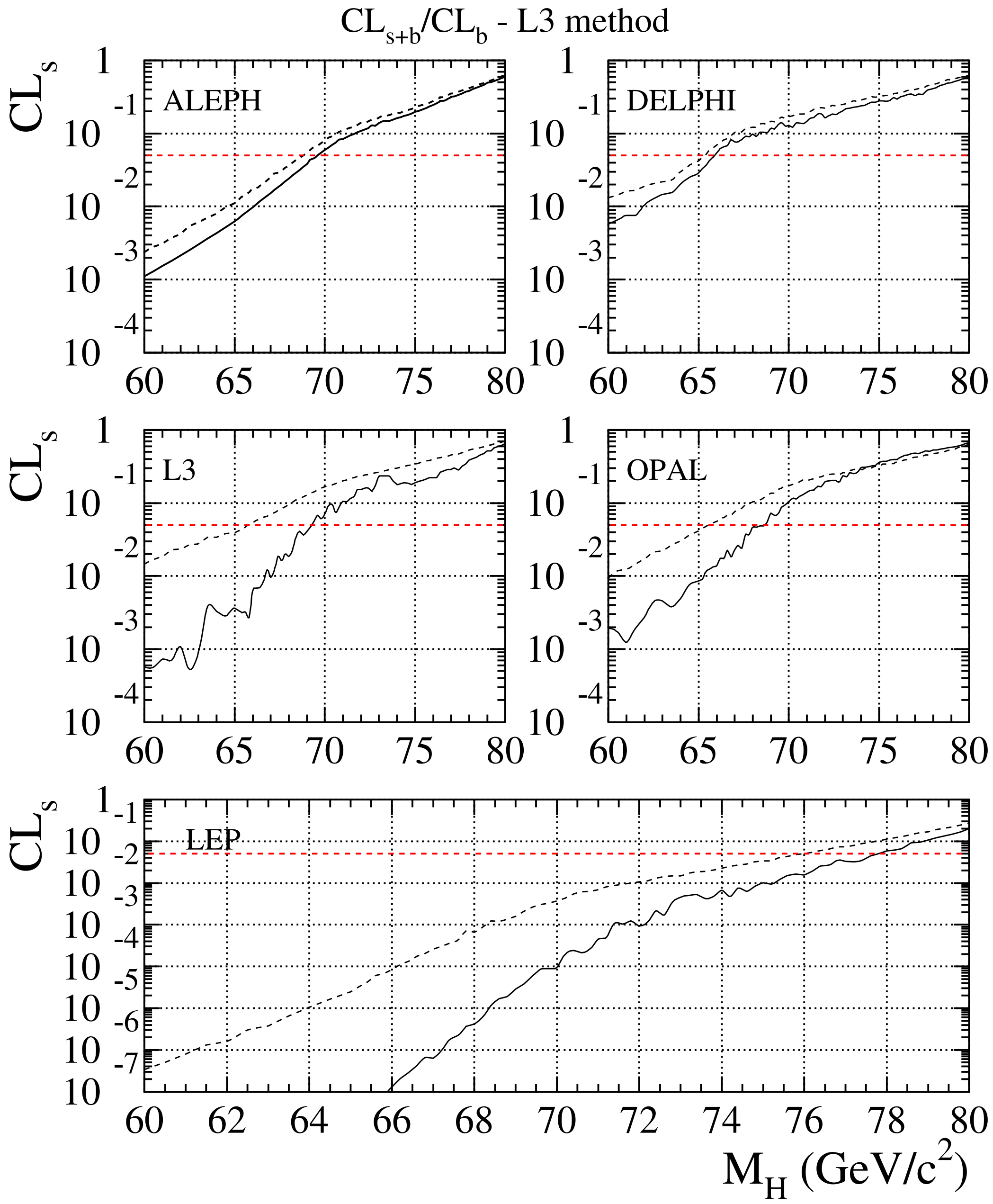

Figure 24: *The observed (full line) and expected (dashed line) confidence level ratios $CL_s \equiv CL_{s+b}/CL_b$, for the individual experiments and for the four experiments combined, based on the method used by L3. The intersections of the horizontal lines at 5% with the curves define the expected and observed 95% c.l. lower limits on the Higgs boson mass.*

4.4 Combined limit using the OPAL method

The inputs from the four experiments are used to derive observed and expected confidence levels, versus m_H, for the *signal+background* and *background only* hypotheses using the OPAL method based on fractional event counting. These are shown in Figures 25 and 26, for each experiment separately and combined. The information pertaining to the OPAL experiment is in fact part of the OPAL input to the ALEPH method.

The free parameter C in Eq. 13 is optimized for each experiment separately and for the four experiments combined, as a function of the assumed Higgs boson mass. The values $C = 0.5$ and $C = 2$, which were found to be optimal in the vicinity of the individual and combined mass limits, respectively, have been used.

Fig. 27 shows the ratio $CL_s \equiv CL_{s+b}/CL_b$ versus m_{H} which is the confidence level for *signal only* experiments. The intersections of these ratios with the lines at 5% define the expected and observed 95% c.l. lower limits on the SM Higgs boson mass for the four experiments separately and combined. These are listed in Table 11 together with the observed confidence level for the *background only* hypothesis, at the mass limits set for the individual experiments and at the combined LEP limit.

Experiment	Exp. lim. (GeV/c^2)	Obs. lim. (GeV/c^2)	$CL_b(m_H^{95})$ (%)
ALEPH	68.6	69.6	40.5
DELPHI	65.1	65.5	64.3
L3	65.0	69.2	18.6
OPAL	65.3	68.9	25.4
LEP	75.6	77.7	26.8

Table 11: *Expected and observed 95% c.l. lower limits on the SM Higgs boson mass for the individual experiments and for the four experiments combined based on the fractional event counting method used by OPAL. Column 3 is the confidence level for the background only hypothesis at the observed mass limits.*

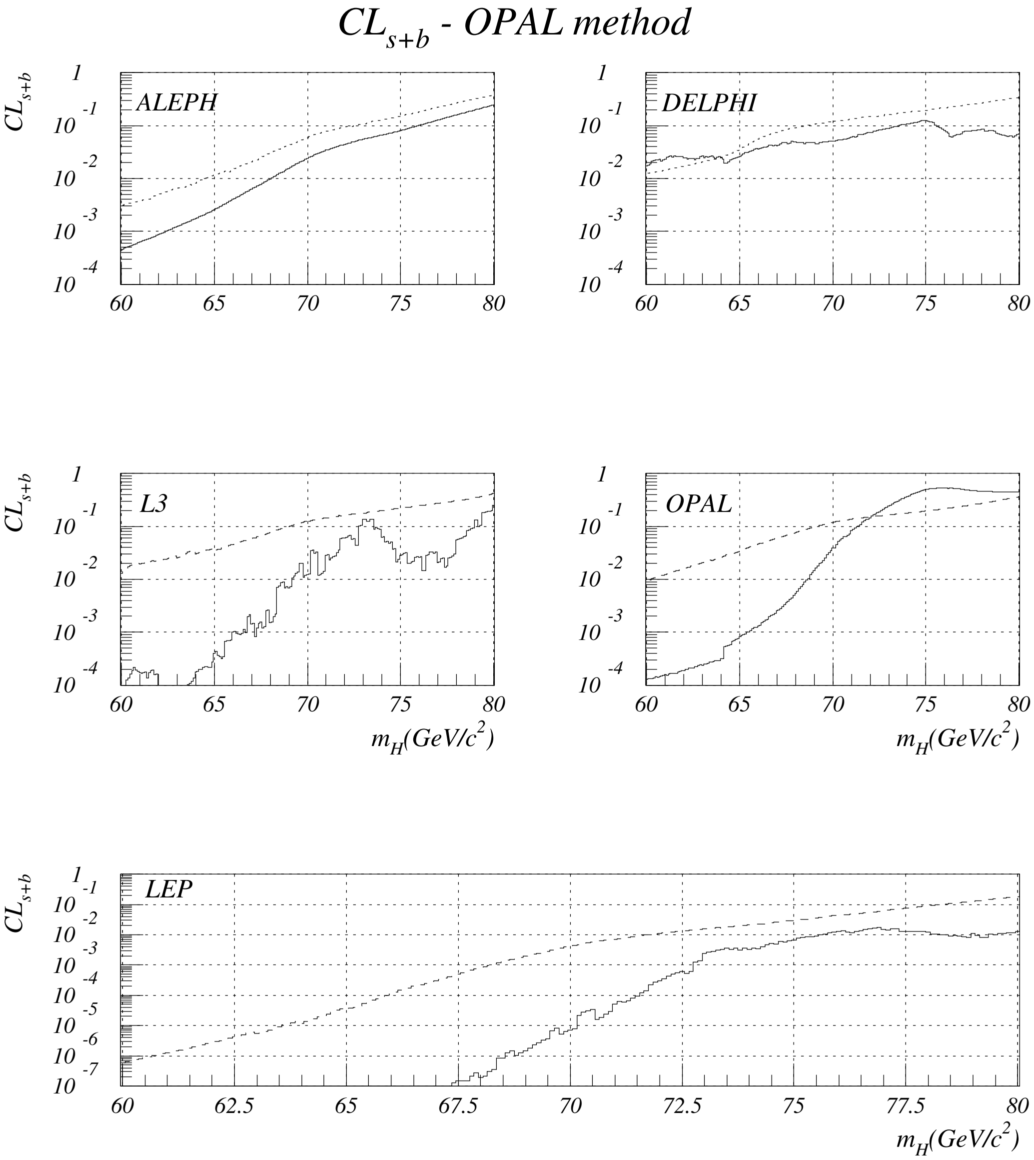

Figure 25: *Observed (full line) and expected (dashed line) confidence levels for the signal+background hypothesis, as functions of the Higgs boson mass, for the individual experiments and for the four experiments combined, based on the OPAL method.*

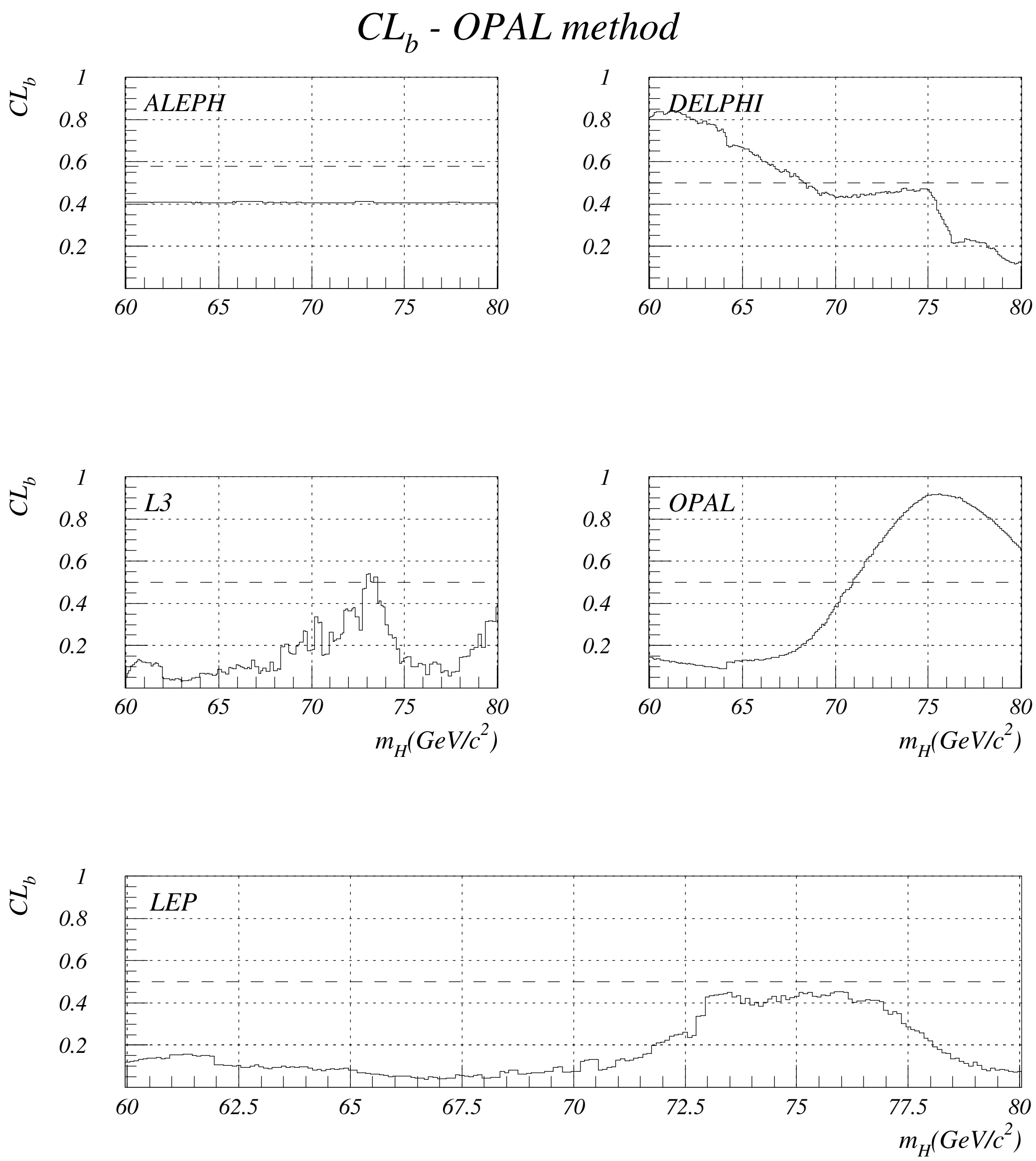

Figure 26: *Observed (full line) and expected (dashed line) confidence levels for the background only hypothesis, as functions of the Higgs boson mass, for the individual experiments and for the four experiments combined, based on the OPAL method.*

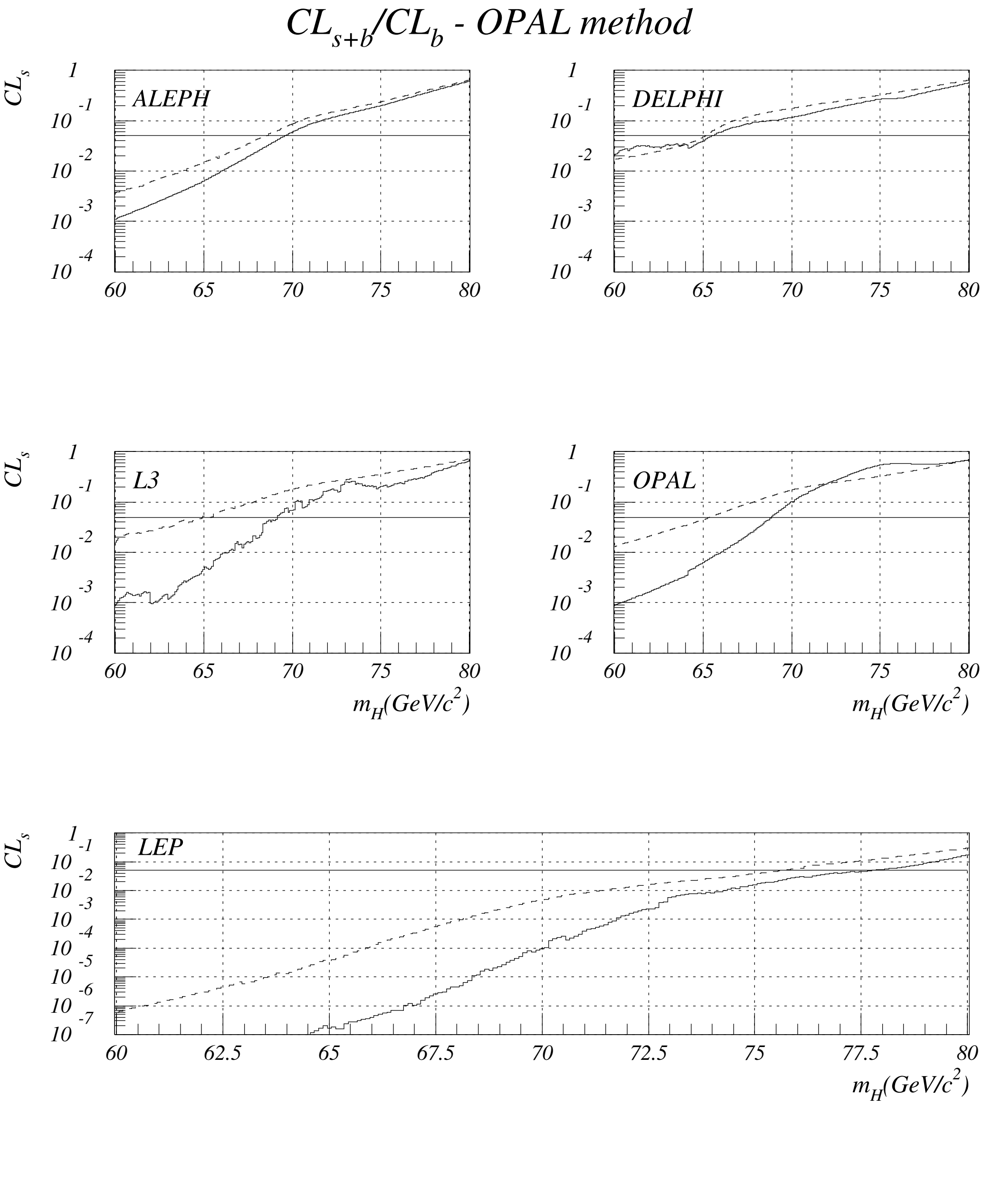

Figure 27: *The observed (full line) and expected (dashed line) confidence level ratios $CL_s \equiv CL_{s+b}/CL_b$, for the individual experiments and for the four experiments combined, based on the method used by OPAL. The intersections of the horizontal lines at 5% with the curves define the expected and observed 95% c.l. lower limits on the Higgs boson mass.*

4.5 Discussion of the results

From comparing the mass limits listed in Tables 8 to 11 one is led to the following observations:

(1) The observed limits are systematically higher than those expected from the *background only* gedanken experiments. This is mainly due to the fact that in each experiment the numbers of observed candidates, although compatible with the SM predictions, are smaller than expected (see Tables 1 to 5). This may be a statistical fluctuation (after all there are only four experiments) but it may also be an indication for a systematic bias.

(2) Comparing, in the first four lines of Table 12, the *observed* 95% c.l. mass limits from single experiments using the four statistical methods, one observes spreads which are listed in the last column. (In the column attributed to the ALEPH method the DELPHI, L3 and OPAL entries are empty since the individual results were not recalculated using the ALEPH method.) The results of ALEPH calculated by the four methods (first line) show no spread, as expected, since ALEPH observes no candidate, in which case all methods have to give identical results. The spread in the results of the other three experiments and in the combined results (last line) [6] is of a few hundred MeV/c^2. It is attributed, in part, to differences in processing the same experimental inputs within the different statistical algorithms and, in part, to intrinsic differences of the test-statistics used, for example in their ways of dealing with candidate events close to the mass limits (see Section 3.5.1). Pertinently, OPAL observes a candidate event in the $(H\rightarrow b\bar{b})q\bar{q}$ channel with a reconstructed Higgs boson mass of 75.6 GeV/c^2 (see Section 2.4).

Experiment	Statistical method				Spread (MeV/c^2)
	ALEPH	DELPHI	L3	OPAL	
ALEPH	69.6	69.6	69.6	69.6	–
DELPHI	-	65.9	65.9	65.5	±200
L3	-	69.4	69.3	69.2	±100
OPAL	-	69.0	68.6	68.9	±200
LEP	77.9	77.5	77.8	77.7	±200

Table 12: *The observed 95% c.l. mass limits (*GeV/c^2*) for the individual experiments and for the four experiments combined using, in each case, the four statistical methods. The last column indicates the spread in MeV/c^2. The entries in the first column pertaining to the DELPHI, L3 and OPAL experiments are empty since the ALEPH method was not used to recalculate the individual limits.*

The lowest of the four combined limits, 77.5 GeV/c^2, can be considered as a conservative estimate of the 95% c.l. lower bound for the SM Higgs boson mass.

[6] This residual spread may be compared to the much larger difference, approximately 4 GeV/c^2, between the 90% and 99% c.l. combined mass limits.

Experiment	Statistical method				Spread (MeV/c^2)
	ALEPH	DELPHI	L3	OPAL	
ALEPH	68.5	68.5	68.8	68.6	±150
DELPHI	-	65.4	65.3	65.1	±150
L3	-	66.1	65.7	65.0	±550
OPAL	-	65.9	65.6	65.3	±300
LEP	75.7	75.8	76.0	75.6	±200

Table 13: *The expected 95% c.l. mass limits (*GeV/c^2*) for the individual experiments and for the four experiments combined using, in each case, the four statistical methods. The last column indicates the spread in MeV/c^2. The entries in the first column pertaining to the DELPHI, L3 and OPAL experiments are empty since the ALEPH method was not used to recalculate the individual limits.*

(3) A similar behaviour of the spreads of individual and combined results is seen in Table 13 where the *expected* 95% c.l. limits are summarized. Again, the spreads can be attributed, in part, to differences in processing the same experimental inputs within the different algorithms and, in part, to differences in the test-statistics used.

The combined signal confidence levels CL_s, expected and observed, as functions of the assumed Higgs boson mass, obtained by the four statistical procedures, are shown again in Figure 28. In these figures, the intersections of the curves with the horizontal lines at 5% define the 95% c.l. lower bounds, expected and observed, on the Higgs boson mass.

4.6 Combining with electroweak data

The high-precision electroweak data from LEP1, in conjunction with measurements of the top quark mass [18] at the Tevatron and of the W boson mass [19] provide a stringent test of the SM [20] which constrains the Higgs boson mass, indirectly, from below and from above. The combined limit on m_{H} derived in this note and based on direct searches can be used as an additional input to testing the SM. It will affect, indirectly, the lower and upper constraints on m_{H} since in Ref. [20] a method based on Bayesian likelihood intervals has been chosen.

However, the curves of confidence level, $CL_s(m_{\mathrm{H}})$, presented in Figure 28 are not directly applicable; they have to be transformed into curves of $\Delta\chi^2(m_{\mathrm{H}})$ representing the variation of χ^2 with m_{H}. Since the confidence levels CL_s, as defined in Eq's. 2 and 3, represent upper limits for the signal probability, they would provide lower bounds for $\Delta\chi^2$. These would be acceptable (since conservative) for deriving a *lower* bound of m_{H} in the context of the SM-test but not for an *upper* bound.

A suitable approach is provided, for example, by the DELPHI statistical method (Section

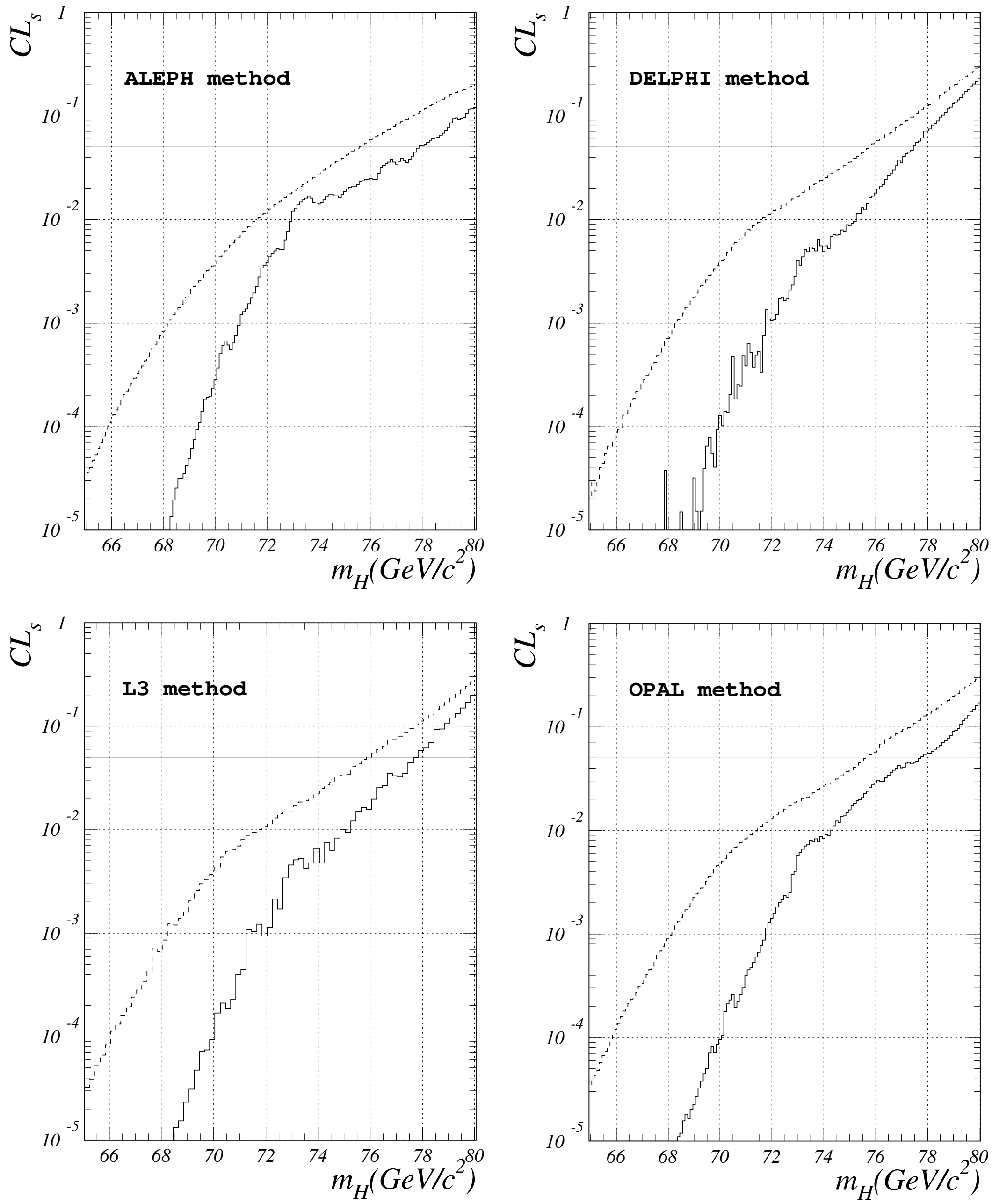

Figure 28: *Expected and observed confidence levels, CL_s, as functions of the hypothetical Higgs boson mass, obtained from combining the results of the LEP experiments using the four statistical methods. The intersections of the curves with the 5% horizontal line define the 95% c.l. lower bounds, expected and observed, for the mass of the SM Higgs boson.*

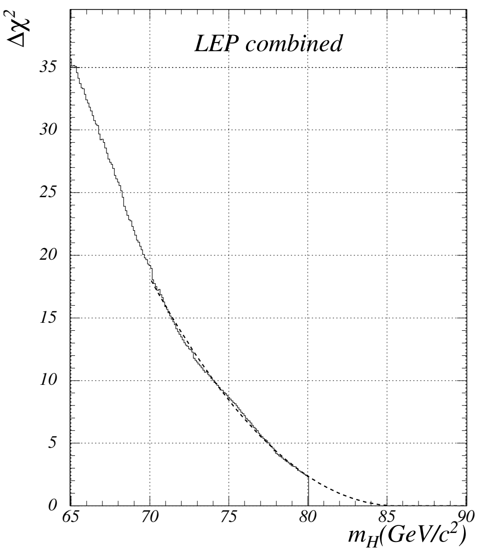

Figure 29: *Full line: $\Delta\chi^2$ as a function of the Higgs boson mass, obtained from direct searches of the four LEP experiments combined, limited to $m_{\mathrm{H}} < 80$ GeV/c^2; dashed line: extrapolation to $m_{\mathrm{H}} > 80$ GeV/c^2, obtained by a parabola fitted in the domain 70 GeV/c^2 $< m_{\mathrm{H}} <$ 80 GeV/c^2.*

3.2), which allows $\Delta\chi^2(m_{\mathrm{H}})$ to be derived directly from the likelihood ratio $Q(m_{\mathrm{H}})$ of Eq. 8:

$$\Delta\chi^2(m_{\mathrm{H}}) = -2\ \ln Q(m_{\mathrm{H}}). \tag{14}$$

In Figure 29, the curve of $\Delta\chi^2(m_{\mathrm{H}})$ obtained in this manner is shown (full line) up to $m_{\mathrm{H}} =$ 80 GeV/c^2 (the direct searches and their combination were not pursued to higher values). For $m_{\mathrm{H}} >$ 80 GeV/c^2, an extrapolation up to the kinematic limit of the $\mathrm{e^+e^-} \to \mathrm{ZH}$ process is provided (dashed line) by a parabolic fit in the domain 70 GeV/$c^2 < m_{\mathrm{H}} <$ 80 GeV/c^2:

$$\Delta\chi^2(m_{\mathrm{H}}) \approx 0.0743\ (m_{\mathrm{H}} - 85.7)^2. \tag{15}$$

It should be noted that the confidence levels computed for this direct search are only approximated by the χ^2 probabilities that may be extracted from Figure 29, due to the small statistics; however, the approximation improves when combined with $-2\ln Q$ curves from other measurements and searches.

5 Summary

In this note, a combined 95% confidence level lower limit for the mass of the Standard Model Higgs boson, from the four LEP experiments, is presented. The limit is based on data collected at centre-of-mass energies of 161 GeV and 170-172 GeV, with integrated luminosities of $\approx$ 20 pb^{-1} per experiment. In their individual searches, none of the four experiments have reported a signal.

The LEP working group has studied four different strategies for combining the results. After testing, each of the four schemes was found suitable for statistically combining the currently available search results of the four experiments. One of the methods, proposed by ALEPH, may however be singled out for future use due to its relative simplicity and the relatively small amount of information that needs to be exchanged when new data are to be added.

Each experiment has provided full sets of experimental information allowing combined mass limits using each of the proposed schemes to be determined. On the basis of the results obtained, a 95% c.l. lower limit of 77.5 GeV/c^2 can be quoted for the mass of the Standard Model Higgs boson. This result can be incorporated into the global test of the Standard Model based on high-precision electroweak data.

Acknowledgements

The LEP working group for Higgs boson searches expresses its gratitude to the LEP Collaborations for making available their experimental results, in some cases prior to their own publications.

References

[1] ALEPH Collaboration, R. Barate et al., Phys. Lett. **B384** (1996) 427;
DELPHI Collaboration, P. Abreu et al., Nucl. Phys. **B421** (1994) 3;
L3 Collaboration, M. Acciarri et al., Phys. Lett. **B385** (1996) 454;
OPAL Collaboration, G. Alexander et al., Z. Phys. **C73** (1997) 189.

[2] ALEPH Collaboration, R. Barate et al., *Search for the Standard Model Higgs Boson in* e^+e^- *Collisions at* $\sqrt{s}$*=161, 170 and 172 GeV*, CERN PPE/97-070, submitted to Physics Letters **B**.

[3] DELPHI Collaboration, P. Abreu et al., *Search for neutral and charged Higgs bosons in* e^+e^- *collisions at* $\sqrt{s}$*=161 and 172 GeV*, CERN-PPE/97-85, submitted to Z. Phys. **C**.

[4] L3 Collaboration, M. Acciarri et al., *Search for the Standard Model Higgs Boson in* e^+e^- *Interactions at* $161 \leq \sqrt{s} \leq 172$ *GeV*, CERN-PPE/97-81, submitted to Physics Letters **B**.

[5] OPAL Collaboration, K. Ackerstaff et al., *Search for the Standard Model Higgs Boson in* e^+e^- *Collisions at* $\sqrt{s}$*=161, 170 and 172 GeV*, CERN-PPE/97-115, submitted to Z. Phys. **C**.

[6] J. Branson *et al.*, L3 Note 2107, June 30, 1997.

[7] M. Acciarri *et al.*, L3 Collaboration, Phys.Lett. **B 385** (1996) 454.

[8] OPAL Collaboration, K. Ackerstaff *et al.*, Phys. Lett. **B393** (1997) 231.

[9] Particle Data Group: R. M. Barnett et al., *Review of Particle Physics*, Phys. Rev. **D54** 1, (1996).

[10] ALEPH method: P. Janot and F. Le Diberder, *"Combining 'Limits' "*, CERN PPE 97-053 and LPNHE 97-01, to be submitted to *Nucl. Instrum. Methods.*

[11] DELPHI method: A. L. Read, *Optimal statistical analysis of search results based on the likelihood ratio and its application to the search for the MSM Higgs boson at 161 and 172 GeV*, DELPHI note 97-158 PHYS 737.

[12] A. G. Frodesen, O. Skjeggestad, and H. Tøfte, *Probability and Statistics in Particle Physics*, Universitetsforlaget, 1979, ISBN 82-00-01906-3.
S. L. Meyer, *Data Analysis for Scientists and Engineers*, John Wiley and Sons, 1975, ISBN 0-471-59995-6.

[13] V.F. Obraztsov, NIM **A316** (1992) 388

[14] O. Helene, NIM **A212** (1983) 319

[15] L3 method: A. Favara and M. Pieri, *Confidence level estimation and analysis optimisation*, Firenze University preprint DFF-278/4/1997, E-preprint hep-ex 9706016.

[16] OPAL method: P. Bock, *Determination of exclusion limits for particle production using different decay channels with different efficiencies, mass resolutions and backgrounds*, Heidelberg University preprint HD-PY-96/05 (1996), to be submitted to Nucl. Instr. Meth.

[17] R. D. Cousins and V. L. Highland, Nucl. Instr. Meth. **320** (1992) 331.

[18] A. Yagil *Top physics*, contribution to the HEP'97 Conference, Jerusalem, August 19-26, 1997.

[19] D. Ward, *Test of the Standard Model W mass and WWZ coupling*, contribution to the HEP'97 Conference, Jerusalem, Aug. 19-26, 1997.

[20] J. Erler and P. Langacker, *The 1997 off-year partial update for the 1998 edition of Particles Data*, available on the PDG Web-pages, (URL: http://pdg.lbl.gov/).