Search for the Standard Model Higgs Boson in $e^+ e^-$ Interactions at $161 \leq \sqrt{s} \leq 172$ GeV

L3 Collaboration

Abstract

A search for the Standard Model Higgs boson has been performed with the L3 detector at LEP. The data sample was collected at three centre-of-mass energies, 161.3, 170.3 and 172.3 GeV with integrated luminosities of 10.8, 1.0 and 9.2 pb$^{-1}$, respectively. No Higgs signal is observed. In combination with previous data taken at the Z resonance, a lower Higgs mass limit, $M_H > 69.5$ GeV, is obtained at 95% confidence level.

Submitted to Phys.Lett. B
1 Introduction

The mechanism of spontaneous symmetry breaking [1] in the Standard Model [2] gives rise to a fundamental neutral scalar particle, the Higgs boson. In the Standard Model the couplings of the Higgs to fermions and gauge bosons are known but the mass, $M_H$, is not. Higgs searches have been performed at the Z resonance by L3 [3] and other experiments [4]. Recent limits from LEP2 have also been reported [5]. In this paper we present the results of a Higgs search using a data sample collected at $161 \leq \sqrt{s} \leq 172 \text{ GeV}$.

At LEP2 the main production mechanism is the Higgs-strahlung process:

$$e^+e^- \rightarrow Z^* \rightarrow HZ.$$  

The dominant final states of this reaction for the mass range $60 < M_H < 80 \text{ GeV}$ are summarised in Table 1. In addition to the process (1), there is a small contribution from the $W^+W^-$ and $ZZ$ fusion reactions to the $H\nu\bar{\nu}$ and $He^+e^-$ final states, respectively. The main background to all these final states comes from fermion pair production and from four-fermion final states.

2 Data and Monte Carlo Samples

The data were collected by the L3 detector [6] at LEP in 1996. The integrated luminosities are 10.8, 1.0 and 9.2 pb$^{-1}$ at the centre-of-mass energies $\sqrt{s} = 161.3$, 170.3 and 172.3 GeV, respectively.

The signal cross section is calculated using the HZHA generator [7]. For the efficiency studies a sample of Higgs events have been generated using PYTHIA [8]. For the background studies the following Monte Carlo programs were used: PYTHIA ($e^+e^- \rightarrow q\bar{q}$), KORALW [9] ($e^+e^- \rightarrow W^+W^-$), KORALZ [10] ($e^+e^- \rightarrow \tau^+\tau^-$), PYTHIA and PHOJET [11] ($e^+e^- \rightarrow e^+e^-q\bar{q}$), and EXCALIBUR [12] ($e^+e^- \rightarrow f\bar{f}f\bar{f}$). The number of simulated background events for the most important background channels is typically 100 times the number of collected data events.

The L3 detector response is simulated using the GEANT 3.15 program [13], which takes into account the effects of energy loss, multiple scattering and showering in the detector. The GHEISHA program [14] is used to simulate hadronic interactions in the detector.

3 Analysis Procedures

The search for the Standard Model Higgs boson at LEP2 involves four distinct event topologies produced in the process $e^+e^- \rightarrow HZ$, namely $q\bar{q}q'\bar{q}'$, $q\bar{q}\nu\bar{\nu}$, $q\bar{q}\ell^+\ell^-$ ($\ell = e, \mu$) and $\tau^+\tau^-q\bar{q}$. Each topology requires its own optimised selection criteria. Since it is expected that a large fraction ($\sim$85%) of Higgs decays contain B-hadrons, the selection criteria for hadronic Higgs decays are optimised for the $H \rightarrow b\bar{b}$ final states.

Two independent analyses were carried out: 1) a weight analysis, and 2) a neural network analysis. This allows a cross-check of the validity of the results. The weight analysis uses an optimisation procedure [3] for the selection criteria and constructs a global event weight variable [15]. The neural network analysis first involves event preselection and then makes use of a neural network technique [16] to separate the signal from the background.
3.1 B-tagging

Jets containing b quarks are primarily identified with lifetime information [17]. The confidence level, $C_N$, that a set of $N$ tracks originated from the primary vertex is constructed using the decay length significance of each track. First the crossing point of each track with the closest jet is determined in both the $r\phi$ and $rz$ projections. Then the signed distances between these crossing points and the reconstructed event primary vertex are projected onto the jet axis to determine the decay length, $L$. If the probability that both $r\phi$ and $rz$ measurements are compatible exceeds 5%, then the two are combined. Otherwise, the $r\phi$ projection is used.

The confidence level is calculated in two ways. The first approach takes into account the fraction of tracks with positive decay length,

$$C_N = \frac{\Pi}{2^N} \sum_{i=0}^{N-1} \sum_{j=i+1}^{N} \left(\frac{N}{j}\right) \left(-\log \Pi\right)^i \frac{1}{i!}, \quad \Pi = \prod_{j=1}^{N_+} P_j(L/\sigma_L),$$

where $N_+$ is the number of tracks with positive decay length. The probability that a track originated from the primary vertex, $P(L/\sigma_L)$, is obtained from the control sample of tracks with negative decay length. The second approach weights each $P(L/\sigma_L)$ by a power $\alpha$ depending on the decay length resolution and momentum of the track,

$$C'_N = 1 - \int_{l_{N_+}}^{1} dx_{N_+} \cdots \int_{l_1}^{1} dx_1, \quad l_j = \left(\frac{\mathcal{P}}{\prod_{i=j+1}^{N_+} x_i^{\alpha_i}}\right)^{1/j}, \quad \mathcal{P} = \prod_{j=1}^{N_+} P_{\alpha_j}^j(L/\sigma_L).$$

By construction, the distributions of both variables are flat for events without lifetime, whereas events containing tracks originating from secondary vertices peak at zero.

To improve the tagging efficiency, the two lifetime variables, $C_N$ and $C'_N$, are combined with other discriminating information using a neural network [16]. The network has fourteen inputs. These include variables computed from reconstructed secondary vertices, such as invariant mass and multiplicity; jet shape variables, e.g. boosted sphericity and $\beta\gamma$; and if an identified electron or muon is present, its momentum information.

The neural network output for a set of jets is combined into an event tag by computing the probability that each jet is compatible with the distribution for light quarks determined from Monte Carlo. The $B_{\text{tag}}$ variable is defined as the negative log-likelihood of these probabilities. As an example, the $B_{\text{tag}}$ spectrum for the sample of $e^+e^- \rightarrow Z \rightarrow q\bar{q}$ events taken during the 1996 calibration run at 91 GeV is presented in Figure 1a. The efficiency and purity for $Z \rightarrow b\bar{b}$ events are shown in Figure 1b as functions of the cut on $B_{\text{tag}}$.

3.2 Weight analysis

The weight analysis [18] combines the most important event variables into an event weight. One of the event variables, $\xi_{\text{kin}}$, is constructed [3] using topological observables, and the other two variables are the reconstructed invariant mass of the Higgs boson, $M_{\text{H}}^{\text{rec}}$, and the $B_{\text{tag}}$ variable. The event weight, $W_H$, is defined as a product of the signal-to-background ratios calculated independently for each of these variables. With such a definition, for given $\xi_{\text{kin}}$, $M_{\text{H}}^{\text{rec}}$ and $B_{\text{tag}}$ values, $W_H$ is related to the signal-to-background ratio which cannot be calculated precisely using the statistics-limited Monte Carlo samples.

The weight analysis results in individual weight distributions for the six final states listed in Table 1. These distributions are then combined to give an overall likelihood function, which
is used to evaluate the presence or absence of a Higgs signal. The weight analysis forms the primary analysis of this paper and is used to derive the final results.

### 3.3 Neural network analysis

The second analysis approach [19] uses similar observables as inputs to a neural network and considers its output, \( NN \), for separation of the signal from the background. Events are pre-selected in order to optimise the performance of the neural network. A feed-forward neural network with one layer of input nodes, one layer of hidden nodes and one output node is used to analyse all the final states listed in Table 1. The number of hidden nodes used in the analysis is typically two times the number of input variables. As with the weight analysis, the individual neural network output distributions are combined to give an overall likelihood function in order to evaluate the presence or absence of a Higgs signal. The neural network analysis provides an independent cross-check of the results of the weight analysis.

### 4 Event Selection

An automated procedure is used to optimise the selection criteria. The optimisation is done independently at 161 GeV and 172 GeV. The optimised values of the cuts are rounded to 3 significant digits and the quoted values correspond to the Higgs search at 172 GeV.

#### 4.1 The HZ \( \rightarrow q\bar{q}q\bar{q} \) channel

The signature of these events is four jets. Two of these jets usually contain b quarks and the other two have an invariant mass consistent with the Z mass.

The event selection proceeds in three steps. First, high multiplicity hadronic events with at least 16 tracks and at least 39 calorimetric clusters (see Figure 2a) are selected, with the visible energy \( 0.4\sqrt{s} < E_{\text{vis}} < 1.6\sqrt{s} \) and the visible mass \( M_{\text{vis}} > 0.4\sqrt{s} \). The thrust direction must be at least 11.5° from the beam axis. A cut on the effective centre-of-mass energy \( \sqrt{s'} > 0.752\sqrt{s} \) rejects radiative return events, \( e^+e^- \rightarrow Z\gamma \). If the photon is observed (\( \sim 13\% \) of all selected events), the effective centre-of-mass energy of the hadronic system is reconstructed using the energy of the photon. If the photon is not observed, then \( \sqrt{s'} \) is reconstructed by rescaling jet energies assuming that the photon escapes along the beam direction. The \( \sqrt{s'} \) distribution is presented in Figure 2b. The main background sources at this stage are \( q\bar{q} \) events with hard gluon radiation and hadronic decays of \( W^+W^- \).

Next, events containing at least four jets are selected. Jets are reconstructed using the DURHAM clustering scheme [20] with the parameter \( Y_{\text{cut}} = 0.0056 \). Events containing more than four jets are reconstructed by changing the \( Y_{\text{cut}} \) parameter to the minimum value, \( Y_{34} \), that gives exactly four jets. The \( Y_{34} \) spectrum is shown in Figure 2c. All jets are required to be at least 8.1° from the beam axis and to have an energy above 0.0694\( \sqrt{s} \). Other kinematic variables used in the selection are the maximum difference among the jet energies, \( \Delta E_{ij} < 0.336\sqrt{s} \) (Figure 2d), and the smallest and the largest di-jet masses, \( M_{ij}^{\text{min}} > 0.0967\sqrt{s} \) and \( M_{ij}^{\text{max}} < 0.763\sqrt{s} \), respectively.

A significant fraction of four-jet events from hadronic \( W^+W^- \) decays is then rejected by requiring \( B_{\text{tag}} > 1.2 \) (see Figure 3a). This selection criterion maintains a high efficiency for \( H \rightarrow b\bar{b} \) decays and in addition retains a significant fraction of \( H \rightarrow c\bar{c} \) decays. The numbers
of selected events agree with the Standard Model expectations for both centre-of-mass energies (see Table 2).

The kinematic variable, $\xi_{\text{kin}}$, combines all the variables listed in the selection except the $B_{\text{tag}}$ variable. The $\xi_{\text{kin}}$ distribution after the final selection is shown in Figure 3b. Four-momentum conservation constraints are applied in the kinematic fit to improve the invariant mass resolution of di-jets. To account for the mass configuration, we define $\chi^2 = \frac{(\Sigma_M - M_Z - M_H)^2}{\sigma^2} + \frac{(\delta_M - |M_Z - M_H|)^2}{\sigma^2}$. The variables $\Sigma_M$ and $\delta_M$ are the sum and the difference of the di-jet masses for each of the three possible combinations. Their resolutions are $\sigma_{\Sigma} = 3$ GeV and $\sigma_{\delta} = 5$ GeV, respectively. A negative sign is ascribed to $\chi^2$ if one mass is underestimated and the other is overestimated. If both masses are under- or over-estimated, the sign is positive. The distribution of signed $\chi^2$ is shown in Figure 3c for $M_H = 70$ GeV. The event weight is then constructed using the $\xi_{\text{kin}}$, $\chi^2$ and $B_{\text{tag}}$ variables. The final weight distribution is presented in Figure 3d. No evidence for a Higgs signal is observed.

The neural network analysis uses a preselection with no B-tagging requirement. The number of observed events is 23 at 161 GeV and 38 at 172 GeV, in agreement with the background expectations of 23.1 and 36.6 events, respectively. The signal selection efficiency is estimated to be 58.6% at 161 GeV and 55.5% at 172 GeV. Fourteen variables are used as inputs to the network. The distribution of the network output for the 172 GeV data sample is presented in Figure 4a. The control data sample of four-jet events from $W^+W^-$ decays is used to check the neural network performance. The same inputs are used to train the neural network to identify $W$ pairs. The cross section for $W^+W^-$ production into four jets is measured to be $5.53^{+1.23}_{-1.06}$ pb at 172 GeV, using a binned maximum-likelihood fit to the data presented in Figure 4b. This measurement is consistent with the Standard Model expectations and agrees with the recent L3 measurement of $W^+W^-$ production rates [21].

### 4.2 The HZ $\rightarrow q\bar{q} \nu\bar{\nu}$ channel

The signature of this process is two acoplanar hadronic jets, no isolated leptons, large missing transverse momentum and jets usually containing b quarks.

High multiplicity hadronic events with at least 5 charged tracks and at least 20 calorimetric clusters are selected. The energy in the forward calorimeters is required to be smaller than 10 GeV. All clusters in the event are combined to form two hadronic jets using the DURHAM algorithm. The invariant mass of these jets, $M_{\text{vis}}$, is required to exceed 34 GeV and each jet must be at least 7.4$^\circ$ from the beam axis. These cuts reduce contributions from pure leptonic final states and from two-photon interactions, $e^+e^- \rightarrow e^+e^-q\bar{q}$, while keeping a significant fraction of hadronic events from $Z$ and $W^+W^-$ decays. These background sources are reduced by requiring $E_{\text{vis}} < 86.2$ GeV.

To further reject events from the two fermion production process $e^+e^- \rightarrow q\bar{q}(\gamma)$, the transverse momentum is required to exceed both 8.11 GeV and 12.1% of $E_{\text{vis}}$. The missing momentum vector must be at least 12.9$^\circ$ from the beam axis and the longitudinal momentum must be smaller than 46.9% of $E_{\text{vis}}$. The energy in the 40$^\circ$ sector around the missing momentum direction must be below 19.9 GeV. The acoplanarity angle between the two jets is required to be greater than 0.7$^\circ$. Events containing identified isolated leptons with energies greater than 6 GeV are rejected in order to suppress the remaining background from $e^+e^- \rightarrow W^+W^-$, where one of the $W$ bosons decays into leptons. In addition three jets are reconstructed for every event using the DURHAM algorithm and the angle measuring their aplanarity, $\alpha_{123} = 360^\circ - \alpha_{12} - \alpha_{13} - \alpha_{23}$, is required to be greater than 0.2$^\circ$. 
The numbers of selected events are presented in Table 2 together with the background expectations and signal efficiencies. The spectra of the $B_{\text{tag}}$, $\xi_{\text{kin}}$ and $M_{\text{vis}}$ variables used for the weight calculation are shown in Figures 5a, 5b and 5c, respectively. The final weight distribution is presented in Figure 5d and shows no evidence for a Higgs signal.

The neural network analysis results are as follows: 9 candidate events are selected in the 161 GeV data sample and 15 in the 172 GeV data sample, consistent with the background expectations of 7.0 and 10.6 events, respectively. The signal efficiency for a 70 GeV Higgs boson is 65% at 161 GeV and 77% at 172 GeV. Eight variables are used as inputs to the network: the transverse momentum, the event invariant mass, the maximum of the two jet masses, the energies in the 25° sector and 40° cone around the missing momentum direction, the missing mass, the $B_{\text{tag}}$ variable and the aplanarity angle $\alpha_{123}$. No excess in the signal region is observed.

### 4.3 The $\text{HZ} \rightarrow q\bar{q} \ell^+\ell^-$ ($\ell = e, \mu$) channels

The signature of $\text{He}^+e^-$ and $\text{H}\mu^+\mu^-$ is a pair of high energy electrons or muons, with an invariant mass close to $M_Z$, accompanied by two hadronic jets.

High multiplicity hadronic events are selected with at least 5 tracks, more than 15 calorimetric clusters and a visible energy of at least $0.3 \cdot \sqrt{s}$. A pair of isolated electrons or muons must be present. The energy of each lepton is required to exceed 3 GeV. After a kinematic fit imposing four-momentum conservation, the invariant mass of the lepton pair, $M_{\ell\ell}$, is required to be $58 \text{ GeV} < M_{\ell\ell} < 107 \text{ GeV}$ for electrons and $22 \text{ GeV} < M_{\ell\ell} < 132 \text{ GeV}$ for muons. If there are more than two lepton candidates, the kinematic fit is repeated for every lepton pair with an additional constraint $M_{\ell\ell} = M_Z$, and the pair giving the smallest $\chi^2$ is chosen.

Electron candidates are identified as a track with an associated cluster in the electromagnetic calorimeter. The numbers of selected events in the $\text{He}^+e^-$ channel are presented in Table 2 together with the background expectations and signal efficiencies. As no B-tagging is necessary, the recoil hadronic mass, $M_{\text{had}}$, is used in place of the weights for the combined results. The $M_{\text{had}}$ spectrum is presented in Figure 6a.

Muon candidates are identified as tracks in the muon spectrometer. The numbers of selected $\text{H}\mu^+\mu^-$ candidates are in agreement with the Standard Model background expectations for both centre-of-mass energies (see Table 2). The reconstructed hadronic mass is obtained from a kinematic fit that imposes both four-momentum conservation and the $M_{\ell\ell} = M_Z$ constraint.

The neural network analysis achieves similar results. For instance, in the $\text{He}^+e^-$ channel, after a preselection, 3 candidate events are selected in the 161 GeV data sample and 3 in the 172 GeV data sample, consistent with the background expectation of 2.3 and 3.6 events, respectively. The signal efficiency for a 70 GeV Higgs boson is 68% at 161 GeV and 76% at 172 GeV. Five variables are used as inputs to the neural network: the energies of the two leptons, their invariant mass, the opening angle between the leptons and the invariant mass of the hadronic system. The neural network output for the $\text{He}^+e^-$ channel at 172 GeV is shown in Figure 6b.

### 4.4 The $\text{HZ} \rightarrow \tau^+\tau^- q\bar{q}$ and $\text{HZ} \rightarrow q\bar{q} \tau^+\tau^-$ channels

The signatures of $\text{H}q\bar{q} \rightarrow \tau^+\tau^- q\bar{q}$ and $\text{H}^+\tau^- \rightarrow q\bar{q} \tau^+\tau^-$ events are similar to those of the $\text{H}^+\ell^- (\ell = e, \mu)$ channels. Tau leptons are identified as low multiplicity jets comprising 1, 2 or 3 tracks and at least 2 GeV of calorimetric energy in a cone of 10° half-angle around its
direction.

High multiplicity hadronic events are selected with more than 5 tracks, more than 15 calorimetric clusters and a visible energy greater than $0.3 \cdot \sqrt{s}$. Two tau candidates with an energy $E_\tau > 4 \text{ GeV}$ must be present. In order to separate the hadronic tau candidates from other hadronic jets, the following restrictions are made in a 30° cone around the tau direction: the total additional energy must be below $0.45 \cdot E_\tau$ and a maximum of 3 additional calorimetric clusters is allowed. After energy and momentum conservation is imposed in the kinematic fit, the masses of the tau pair and of the recoiling hadronic system are computed. The invariant mass closest to $M_Z$ is chosen and the event is classified as either $Hq\bar{q}$ or $H\tau^+\tau^-$ depending on whether this mass is made by the jets or the taus. The reconstructed $Z$ boson mass is required to be in the range $78 \text{ GeV} < M_Z < 109 \text{ GeV}$. To reduce the background from $e^+e^-\rightarrow q\bar{q}\gamma$, it is required that $\sqrt{\hat{s}} > 0.6 \cdot \sqrt{s}$. To suppress $W^+W^-$ background, the sum of the energy of the most energetic tau and the missing energy must be smaller than $60 \text{ GeV}$. The numbers of selected events for both $HZ\rightarrow \tau^+\tau^-q\bar{q}$ and $HZ\rightarrow q\bar{q}\tau^+\tau^-$ channels are presented in Table 2 together with the background expectations and signal efficiencies.

The neural network analysis results are as follows. For the $Hq\bar{q}$ channel at $161 \text{ GeV}$ 5 events are selected with 5.7 expected and a signal efficiency of 30% for a 70 GeV Higgs boson. At $172 \text{ GeV}$ 6 events are selected with 5.9 expected and an efficiency of 29%. For the $H\tau^+\tau^-$ channel these numbers are 3 and 4.2 (efficiency of 32%) at $161 \text{ GeV}$, and 5 and 5.0 (efficiency of 30%) at $172 \text{ GeV}$. No excess in the signal region is seen for either the $Hq\bar{q}$ or $H\tau^+\tau^-$ channel.

5 Systematic errors

Detector efficiencies and backgrounds for all the channels are estimated from the Monte Carlo simulation. Uncertainties on these estimates are treated as Gaussian and, for a single channel, assumed to be independent. The correlations between the channels and different centre-of-mass energies are accounted for in the final result.

Experimental uncertainties in the LEP centre-of-mass energy of $\pm 0.03 \text{ GeV}$ [22] and in the luminosity measurements account for 1% systematic error on the number of expected signal events. Theoretical errors on the Higgs boson production cross section due to the uncertainties in $M_{top}$ and $\alpha_s$ [23] ($\sim 0.1\%$), interference effects [24] ($\sim 1\%$) and quark masses [25] ($\sim 1\%$) introduce an additional uncertainty on the predicted number of signal events.

Studies of the possible systematic effects due to the topological cuts are described in detail in [26] for the four-jet channel. Similar studies are performed for the other channels. The net effect on the number of expected signal events is found to be 1% with almost equal contributions from tracking efficiency and uncertainty in the absolute energy scale. The systematic uncertainty due to the B-tagging is estimated using the calibration sample of hadronic events at 91 GeV and assuming that the difference between data and Monte Carlo distributions is entirely due to the systematic effects. The effect on the signal efficiency is evaluated to be $3 - 5\%$, depending on the channel, by reweighting the signal distributions accordingly. The Monte Carlo statistics adds 2% to the error on the signal efficiency. Assuming these uncertainties are independent, the overall systematic error on the number of signal events is estimated to be 4%.

The uncertainty on the background mainly comes from two different sources: Monte Carlo statistics and the normalisation error due to the uncertainty on both the cross section and the selection efficiency for the background processes. The error from the Monte Carlo statistics is relatively large but completely uncorrelated between the different bins of the individual channels; this fact leads to a negligible effect on the confidence level evaluation. The overall
normalisation error is the most important and is estimated to be 10%. It is assumed to be fully correlated between the different channels.

6 Results

In Table 3 are shown the fractions of Higgs events for each selection, \( \sum_{k=1}^{n_{\text{ch}}} B_k \epsilon_k \), where the sum runs over all channels, \( k = (1, n_{\text{ch}}) \), and \( B_k \) and \( \epsilon_k \) are the corresponding branching fractions and efficiencies, respectively. To reduce the effect of the uncertainties on the estimation of the expected background, only the regions in the \( W_H \) variable with a signal-to-background ratio larger than 0.15 are used in the derivation of the confidence level. This implies that the number of data events and the expected background depend on the Higgs boson mass hypothesis. After applying this cut, approximately 10 background events are expected for a \( M_H \) hypothesis in the range from 60 to 70 GeV.

The likelihood as a function of the number of expected signal events is determined from the distributions of the weight variables for all the channels (except He\(^+\)e\(^-\) and H\(\mu^+\mu^-\) channels for which the reconstructed Higgs mass is used) and all centre-of-mass energies. Examples of the corresponding distributions for the Hq\(\bar{q}\), H\(\nu\bar{\nu}\) and He\(^+\)e\(^-\) channels are presented in Figures 3d, 5d and 6a, respectively. Poisson statistics for the number of observed data events, \( N_{i,k} \), is used to define the likelihood function:

\[
\mathcal{L}(\mu^H) = \prod_{k=1}^{n_{\text{ch}}} \prod_{i=1}^{n_{\text{bin}}} \frac{e^{-(\mu^H_{i,k} + \mu^B_{i,k})} (\mu^H_{i,k} + \mu^B_{i,k})^{N_{i,k}}}{N_{i,k}!};
\]

where the product is taken over all search channels, all centre-of-mass energies, \( k = (1, n_{\text{ch}}) \); and all bins, \( i = (1, n_{\text{bin}}) \). The background shapes and normalisations are fixed to the Monte Carlo predictions, \( \mu^B_{i,k} \). The individual signal expectations, \( \mu^H_{i,k} \), are proportional to the total number of signal events, \( \mu^H \), the corresponding branching fractions, \( B_k \), and efficiencies, \( \epsilon_i \).

As no excess in the signal regions is observed in the data, the likelihood has its maximum at \( \mu^H = 0 \). It is concluded that no evidence is observed for Higgs boson production.

The following method [15] is used to derive a confidence level, CL, for the exclusion of the Higgs boson with a mass \( M_H \). First an estimator based on Bayesian statistics is constructed:

\[
P_\mathcal{L}(\mu^H) = \frac{\int_{\mu^H}^{\infty} \mathcal{L}(\mu) d\mu}{\int_{0}^{\infty} \mathcal{L}(\mu) d\mu}.
\]

Then a large number of Monte Carlo experiments is performed such that each experiment generates, based on Poisson statistics, an “observation” using the background and the Higgs signal expectations. The \( (1 - \text{CL}) \) value is obtained as the ratio of the fractions of outcomes with the estimator value less than that of the data for two hypotheses: 1) both the Standard Model Higgs signal and the background and 2) the background only. Defined in this way, the \( (1 - \text{CL}) \) value corresponds to the probability to exclude an existing signal in the framework of classical statistics.

The systematic errors on the signal and background expectations are taken into account during the generation of these Monte Carlo experiments. In each trial experiment candidates are generated according to the signal and background distributions which are smeared to account for the systematic errors. The nominal expected signal and background are then used to calculate the confidence level.

8
The measured and average confidence levels as functions of the Higgs boson mass are shown in Figure 7a. The 95% CL limit on the Higgs mass using only 161 GeV and 172 GeV data is 69.2 GeV. The probability to obtain a better limit is estimated to be 23% using a large number of Monte Carlo experiments. In combination with the data taken at the Z resonance the final result for the Higgs mass limit (Figure 7b) is

\[ M_H > 69.5 \text{ GeV at 95\% CL}. \]

The neural network analysis confirms this result within 0.1 GeV. The new mass limit improves and supersedes our previously published analysis [3].

Acknowledgements

We wish to express our gratitude to the CERN accelerator divisions for the good performance of the LEP machine. We acknowledge the efforts of all engineers and technicians who have participated in the construction and maintenance of this experiment.
References


1) These L3 Notes are freely available on request from: The L3 secretariat, CERN, CH-1211 Geneva 23, Switzerland. Internet: http://hpl3sn02.cern.ch/l3pubanddoc.html
Table 1: Decay channels and branching fractions in $e^+e^- \rightarrow HZ$, for $M_H = 70$ GeV. The $H \rightarrow q\bar{q}$ final states include both $q\bar{q}$ and $gg$.

<table>
<thead>
<tr>
<th>H decay channel</th>
<th>Z decay channel</th>
<th>Branching fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>$q\bar{q}$</td>
<td>$q\bar{q}$</td>
<td>64.4%</td>
</tr>
<tr>
<td>$q\bar{q}$</td>
<td>$\nu\bar{\nu}$</td>
<td>18.4%</td>
</tr>
<tr>
<td>$q\bar{q}$</td>
<td>$e^+e^-$</td>
<td>3.1%</td>
</tr>
<tr>
<td>$q\bar{q}$</td>
<td>$\mu^+\mu^-$</td>
<td>3.1%</td>
</tr>
<tr>
<td>$q\bar{q}$</td>
<td>$\tau^+\tau^-$</td>
<td>3.1%</td>
</tr>
<tr>
<td>$\tau^+\tau^-$</td>
<td>$q\bar{q}$</td>
<td>5.5%</td>
</tr>
</tbody>
</table>

Table 2: The signal efficiencies, expected background and the number of data events for the weight analysis. Each efficiency corresponds to a final state for which the respective selection is optimised.

<table>
<thead>
<tr>
<th>Final state</th>
<th>Efficiency</th>
<th>BG</th>
<th>DATA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\sqrt{s} = 161$ GeV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>Z</td>
<td>$M_H = 65$ GeV</td>
<td>$M_H = 70$ GeV</td>
</tr>
<tr>
<td>bb</td>
<td>$q\bar{q}$</td>
<td>0.622</td>
<td>0.624</td>
</tr>
<tr>
<td>bb</td>
<td>$\nu\bar{\nu}$</td>
<td>0.738</td>
<td>0.639</td>
</tr>
<tr>
<td>qq</td>
<td>$e^+e^-$</td>
<td>0.634</td>
<td>0.579</td>
</tr>
<tr>
<td>qq</td>
<td>$\mu^+\mu^-$</td>
<td>0.508</td>
<td>0.473</td>
</tr>
<tr>
<td>bb</td>
<td>$\tau^+\tau^-$</td>
<td>0.190</td>
<td>0.183</td>
</tr>
<tr>
<td>$\tau^+\tau^-$</td>
<td>$q\bar{q}$</td>
<td>0.278</td>
<td>0.198</td>
</tr>
<tr>
<td></td>
<td>$\sqrt{s} = 172$ GeV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bb</td>
<td>$q\bar{q}$</td>
<td>0.511</td>
<td>0.529</td>
</tr>
<tr>
<td>bb</td>
<td>$\nu\bar{\nu}$</td>
<td>0.786</td>
<td>0.745</td>
</tr>
<tr>
<td>qq</td>
<td>$e^+e^-$</td>
<td>0.667</td>
<td>0.663</td>
</tr>
<tr>
<td>qq</td>
<td>$\mu^+\mu^-$</td>
<td>0.479</td>
<td>0.510</td>
</tr>
<tr>
<td>bb</td>
<td>$\tau^+\tau^-$</td>
<td>0.279</td>
<td>0.268</td>
</tr>
<tr>
<td>$\tau^+\tau^-$</td>
<td>$q\bar{q}$</td>
<td>0.238</td>
<td>0.248</td>
</tr>
</tbody>
</table>
Table 3: Fractions of the total number of signal events, $\sum_{k=1}^{n_{\text{ch}}} B_k \epsilon_k$, that satisfy a given selection. Note that the numbers include feed-ins from other signal channels and fusion diagrams, and events are uniquely assigned to a single selection channel.

<table>
<thead>
<tr>
<th>Selection channel</th>
<th>$\sum_{k=1}^{n_{\text{ch}}} B_k \epsilon_k$</th>
<th>Expected signal events</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M_H = 65$ GeV</td>
<td>$M_H = 70$ GeV</td>
</tr>
<tr>
<td></td>
<td>$\sqrt{s} = 161$ GeV</td>
<td>$\sqrt{s} = 172$ GeV</td>
</tr>
<tr>
<td>q\bar{q}</td>
<td>q\bar{q}</td>
<td>0.3847</td>
</tr>
<tr>
<td>q\bar{q} $\nu\bar{\nu}$</td>
<td>0.1415</td>
<td>0.1293</td>
</tr>
<tr>
<td>q\bar{q} e^+e^-</td>
<td>0.0199</td>
<td>0.0185</td>
</tr>
<tr>
<td>q\bar{q} $\mu^+\mu^-$</td>
<td>0.0164</td>
<td>0.0152</td>
</tr>
<tr>
<td>q\bar{q} $\tau^+\tau^-$</td>
<td>0.0086</td>
<td>0.0098</td>
</tr>
<tr>
<td>$\tau^+\tau^-$ q\bar{q}</td>
<td>0.0159</td>
<td>0.0122</td>
</tr>
<tr>
<td>q\bar{q}</td>
<td>q\bar{q}</td>
<td>0.3121</td>
</tr>
<tr>
<td>q\bar{q} $\nu\bar{\nu}$</td>
<td>0.1504</td>
<td>0.1426</td>
</tr>
<tr>
<td>q\bar{q} e^+e^-</td>
<td>0.0209</td>
<td>0.0208</td>
</tr>
<tr>
<td>q\bar{q} $\mu^+\mu^-$</td>
<td>0.0155</td>
<td>0.0166</td>
</tr>
<tr>
<td>q\bar{q} $\tau^+\tau^-$</td>
<td>0.0149</td>
<td>0.0152</td>
</tr>
<tr>
<td>$\tau^+\tau^-$ q\bar{q}</td>
<td>0.0140</td>
<td>0.0152</td>
</tr>
</tbody>
</table>
Figure 1: (a) The spectrum of the $B_{\text{tag}}$ variable for the sample of $Z \rightarrow q\bar{q}$ events at 91 GeV. (b) The purity of the sample and the efficiency for $Z \rightarrow b\bar{b}$ events are presented as functions of the cut on the $B_{\text{tag}}$ variable.
Figure 2: The most important kinematic variables used in the $q\bar{q}q\bar{q}$ analysis: (a) the number of calorimetric clusters; (b) the scaled effective centre-of-mass energy; (c) the DURHAM parameter $Y_{34}$; (d) the scaled maximal difference of jet energies, $\Delta E_{ij}/\sqrt{s}$. The superimposed hatched histograms correspond to a 70 GeV Higgs signal normalised to the Standard Model cross section. The distribution are shown for a sample of hadronic events at 172 GeV. The corresponding selection cuts for the weight analysis are indicated by arrows.
Figure 3: The variables used to calculate the event weights for the q qqq analysis: (a) the $B_{\text{tag}}$ variable after applying all other selection cuts, (b) the kinematic variable $\xi_{\text{kin}}$, (c) the signed $\chi^2$ representing the consistency of an event with $M_H = 70$ GeV; and (d) the final event weight spectrum for the 172 GeV data sample. The distributions (b), (c) and (d) are shown after imposing all selection criteria. The superimposed hatched histograms correspond to the 70 GeV Higgs boson signal normalised to the Standard Model cross section.
Figure 4: (a) The final $q\bar{q}q\bar{q}$ neural network output spectrum for the 172 GeV data sample. The superimposed hatched histogram corresponds to the 70 GeV Higgs boson signal normalised to the Standard Model cross section. (b) The neural network output spectra for the $W^+W^- \rightarrow q\bar{q}q\bar{q}$ cross section fit at 172 GeV. The hatched area represents the fitted $W^+W^- \rightarrow q\bar{q}q\bar{q}$ contribution and the open area shows the contribution of the QCD background.
Figure 5: The variables used to calculate the event weights for the q$q \bar{q}\nu\bar{\nu}$ analysis: (a) the $B_{tag}$ variable, (b) the kinematic variable $\xi_{kin}$, (c) the invariant mass of the two jets, $M_{vis}$, and (d) the final event weight spectrum for the 172 GeV data sample. The superimposed hatched histograms correspond to the 70 GeV Higgs boson signal normalised to the Standard Model cross section.
Figure 6: (a) The recoiling mass of the hadronic system, $M_{\text{had}}$, calculated from $e^+e^-$ pairs for the $q\bar{q}e^+e^-$ sample as used in the weight analysis. (b) The final $q\bar{q}e^+e^-$ neural network output spectrum for the 172 GeV data sample. The hatched histograms correspond to the 70 GeV Higgs boson signal normalised to the Standard Model cross section.
Figure 7: (a) The (1-CL) line for the actual observation that combines all the data (solid line) and for 161 – 172 GeV data only (dotted line). The combined average expectation is indicated by the dashed line. (b) The number of expected Higgs signal events (solid line) and the number of signal events excluded at 95% confidence level (dashed line) as functions of the Higgs boson mass.