Search for a Fermiophobic Higgs Boson in the diphoton final state using 4.2 fb^{-1} of DØ data

The DØ Collaboration
URL http://www-d0.fnal.gov
(Dated: March 11, 2009)

This note describes a search for a fermiophobic Higgs boson in the di-photon final state using $4.2 \pm 0.3 \ fb^{-1}$ of the DØ Run II data, collected at the Fermilab Tevatron collider from April 2002 to December 2008. Good agreement between the data and the Standard Model (SM) background prediction is observed. We set 95% C.L. limits on the production cross section times the branching ratio $(\sigma \times BR(h_f \to \gamma\gamma))$ for different assumed Higgs masses from 80 GeV to 150 GeV. The results have reached the same sensitivity as a single LEP experiment, setting a lower limit on the Fermiophobic Higgs of $M_{h_f} > 102.5$ GeV ($M_{h_f} > 107.5$ GeV expected) and approaching the sensitivity of the combined LEP limit ($M_{h_f} > 109.7$ GeV). We also provide access to the $M_{h_f} > 125$ GeV region, which was inaccessible to the LEP experiments.

I. INTRODUCTION

In the standard model (SM), the $h \to \gamma\gamma$ branching ratio is very small. The branching ratio is about 0.22% for a Higgs boson with a mass of 130 GeV. However, in some models beyond the SM, the $h \to \gamma\gamma$ branching ratio can be enhanced significantly [1]. For example in the case of a fermiophobic Higgs boson (h_f), which assumes zero couplings of the Higgs boson to fermions, this branching ratio is almost an order of magnitude larger (see Table. I). This model has been tested at LEP [2] and the Tevatron [3]. In the SM the main production mechanism for the Higgs boson in hadronic collisions is the gluon fusion diagram which involves a top quark loop. If the Higgs does not couple to fermions this production mechanism is absent and the Higgs boson is produced mainly in associations with vector bosons ($h_f + V \to \gamma\gamma + V$, $V = W, Z$) or through the vector boson fusion process (VBF $h_f \to \gamma\gamma$). It is assumed that the coupling of the fermiophobic Higgs to V is the same as that of the SM Higgs boson. A distinctive feature of these production mechanisms is that the Higgs boson is produced with large transverse momentum $p_T^{\gamma\gamma}$ (see Fig. 1). Therefore, a cut on the p_T of the diphoton system ($p_T^{\gamma\gamma}$) can be very effective at suppressing the background while keeping most of the signal. Besides this $p_T^{\gamma\gamma}$ cut, we use the same technique as Ref. [4].

The Higgs Monte Carlo (MC) samples, as well as any other MC samples used in this analysis, are generated using PYTHIA [5] with CTEQ6L [6] parton distribution functions (PDFs), and processed through a GEANT-3 based [7] simulation of the DØ detector and the same reconstruction software as the data. Samples corresponding to each of the three dominant SM Higgs boson production mechanisms are generated, and normalized using the next-to-next-to-leading order (NNLO) theoretical cross sections [8, 9] and the branching ratio predictions from HDECAY [10].
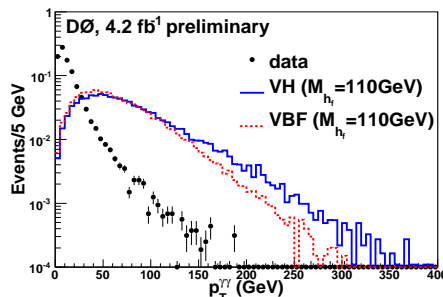

FIG. 1: p_T of di-photon system ($p_T^{\gamma\gamma}$) for data, associated production and a vector boson fusion signal with $M_{h_f} = 110$ GeV. All histograms are normalized to unity.

m_{h_f} (GeV)	Wh_f(NNLO)	Zh_f(NNLO)	VBF(NLO)	$BR(h_f \to \gamma\gamma)$
80	0.581	0.313	0.141	0.70
90	0.406	0.223	0.119	0.41
100	0.2861	0.1668	0.0995	0.18
110	0.2077	0.1233	0.0842	0.062
120	0.1532	0.0926	0.0693	0.028
130	0.1145	0.0704	0.0604	0.019
140	0.0866	0.0541	0.0508	0.0061
150	0.0661	0.0420	0.0441	0.0020

TABLE I: Cross section (pb) and branching ratio (BR) for a fermiophobic Higgs boson. Signal with mass greater than 150 GeV is not considered due to its small BR.

II. DØ DETECTOR AND DATA SAMPLE

The DØ detector is comprised of a central tracking system in a 2 T superconducting solenoidal magnet, a liquid-argon/uranium calorimeter, a central preshower detector and a muon spectrometer [11]. The major parts of the DØ detector used in event selection are the tracking system, the electromagnetic (EM) calorimeter and the central preshower detector (CPS). The tracking system consists of a silicon microstrip tracker (SMT) and an eight-layer scintillating fiber tracker (CFT) mounted on thin coaxial barrels. It provides coverage for charged particles in the pseudorapidity range $|\eta| < 3$ (where the pseudorapidity is defined as $\eta \equiv -\ln[\tan(\frac{\theta}{2})]$, with θ denoting the polar angle

with respect to the proton beam direction.) The calorimeter has a central section (CC) covering up to $|\eta| \approx 1.1$, and two end components (EC) extending coverage to $|\eta| \approx 4.2$. Each section is housed in a separate cryostat, and divided into EM layers on the inside and hadronic layers on the outside. The EM calorimeter has four longitudinal layers and transverse segmentation of 0.1×0.1 in $\eta - \phi$ space (where ϕ is the azimuthal angle), except in the third layer, where it is 0.05×0.05. Immediately before the inner layer of the central EM calorimeter, there is a central preshower detector (CPS) formed of $2X_0$ of absorber followed by several layers of scintillating strips with embedded wavelength-shifting fibers. Luminosity is measured using plastic scintillator arrays located in front of the EC cryostats, covering $2.7 < |\eta| < 4.4$. The data acquisition system consists of a three-level trigger, designed to accommodate the high instantaneous luminosity. For final states containing two photon candidates with transverse momentum (p_T) above 20 GeV, the trigger efficiency is close to 100%. The data samples used in this analysis was collected between April 2002 and December 2008 and corresponds to an integrated luminosity of 4.2 ± 0.3 fb^{-1} after applying the data quality requirements.

III. EVENT SELECTION

Events are selected requiring at least two photon candidates with $p_T > 20$ GeV and $|\eta| < 1.1$, for which the trigger requirements are fully efficient. Photons are selected from EM clusters reconstructed within a cone with radius $\mathcal{R} = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.2$ and requiring that: (i) at least 97% of the cluster energy is deposited in the EM calorimeter, (ii) the calorimeter isolation variable $I = [E_{\text{tot}}(0.4) - E_{\text{EM}}(0.2)]/E_{\text{EM}}(0.2)$ is less than 0.1, where $E_{\text{tot}}(0.4)$ is the total energy in a cone of radius $\mathcal{R} = 0.4$ and $E_{\text{EM}}(0.2)$ is the EM energy in a cone of radius $\mathcal{R} = 0.2$; (iii) the energy-weighted shower width squared in the $r - \phi$ plane in EM3 is less than 14 cm^2; and (iv) the scalar sum of the p_T of all tracks originating from the primary vertex in an annulus of $0.05 < \mathcal{R} < 0.4$ around the cluster $(p_{T\text{trk}}^{\text{sum}})$ is less than 2 GeV. To suppress electrons misidentified as photons, the EM clusters are required to not be spatially matched to tracker activity, i.e. either to a reconstructed track, or to a density of hits in the SMT and CFT consistent with that of an electron. The contribution of jets misidentified as photons is reduced by combining the information from a set of variables sensitive to differences between photons and jets in the tracker activity and in the energy distributions in the calorimeter and CPS, using an artificial neural network (ANN) [12]. The variables used are: $p_{T\text{trk}}^{\text{sum}}$, the number of cells in the first EM calorimeter layer within $\mathcal{R} < 0.2$ and $0.2 < \mathcal{R} < 0.4$ of the EM cluster, the number of associated CPS clusters within $\mathcal{R} < 0.1$ of the EM cluster, and the squared-energy-weighted width of the energy deposition in the CPS. The ANN is trained using diphoton and dijet MC samples and its performance is verified using a data sample of $Z \to \ell^+\ell^-\gamma$ ($\ell = e, \mu$) events. Fig. 2 compares the ANN output (O_{NN}) distribution for photons and jets. Photon candidates are required to have O_{NN} larger than 0.1. Such a requirement is almost 100% efficient for photons while rejecting $\sim 50\%$ of misidentified jets. Finally, the $p_T^{\gamma\gamma}$ (see Fig. 1) is conservatively required to be larger than 35 GeV to further suppress the backgrounds for all signal mass points.
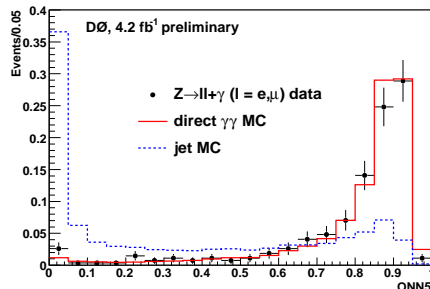

FIG. 2: Normalized distributions of O_{NN} value from real and fake photons.

IV. BACKGROUNDS

There are three major sources of background to the $h \to \gamma\gamma$ signature: (i) Drell-Yan events, where both electrons are misidentified as photons; (ii) γ+jet and di-jet events where the jet(s) are mis-identified as photon(s); (iii) direct di-photon production.

A. Drell-Yan $Z/\gamma* \to ee$ contribution

We use $Z/\gamma* \to ee$ PYTHIA MC samples to estimate the Drell-Yan contribution. The next-to-next-to-leading-order (NNLO) $p\bar{p} \to Z/\gamma* \to ee$ cross section [13] is used for the absolute normalization. The total background contribution from the Drell-Yan process is found to be 25.7 ± 2.3 (stat.) events.

B. γ+jet and di-jet background

We estimate the γ + jet and di-jet contributions from the data with the final event selection applied (see section III) by using a 4×4 matrix background subtraction method. The method is described in Ref. [14]. In this analysis, we use $O_{NN} = 0.75$ as a boundary to classify the candidates into four categories:

- N_{pp} of them have both photon candidates with $O_{NN} > 0.75$;

- N_{pf} of them have the leading photon candidate with $O_{NN} > 0.75$, but the sub-leading $O_{NN} < 0.75$;

- N_{fp} vice versa;

- N_{ff} of them have both photon candidates with $O_{NN} < 0.75$.

The Drell-Yan $Z/\gamma* \to ee$ contributions to $(N_{pp}, N_{pf}, N_{fp}, N_{ff})$ are determined from MC simulations and are removed. The pass-fail vector $(N_{pp}, N_{pf}, N_{fp}, N_{ff})$ thus obtained is related the $(N_{\gamma\gamma}, N_{\gamma j}, N_{j\gamma}, N_{jj})$ vector as follows:

$$\begin{pmatrix} N_{ff} \\ N_{fp} \\ N_{pf} \\ N_{pp} \end{pmatrix} = E \times \begin{pmatrix} N_{jj} \\ N_{j\gamma} \\ N_{\gamma j} \\ N_{\gamma\gamma} \end{pmatrix} \tag{1}$$

where the $N_{\gamma\gamma}$ is the number of $\gamma+\gamma$ events, $N_{\gamma j}$ and $N_{j\gamma}$ are the number of γ+jet events and N_{jj} is the number of di-jet events. The 4×4 matrix E is defined as:

$$E = \begin{pmatrix} (1-\epsilon_{j1})(1-\epsilon_{j2}) & (1-\epsilon_{j1})(1-\epsilon_{\gamma2}) & (1-\epsilon_{\gamma1})(1-\epsilon_{j2}) & (1-\epsilon_{\gamma1})(1-\epsilon_{\gamma2}) \\ (1-\epsilon_{j1})\epsilon_{j2} & (1-\epsilon_{j1})\epsilon_{\gamma2} & (1-\epsilon_{\gamma1})\epsilon_{j2} & (1-\epsilon_{\gamma1})\epsilon_{\gamma2} \\ \epsilon_{j1}(1-\epsilon_{j2}) & \epsilon_{j1}(1-\epsilon_{\gamma2}) & \epsilon_{\gamma1}(1-\epsilon_{j2}) & \epsilon_{\gamma1}(1-\epsilon_{\gamma2}) \\ \epsilon_{j1}\epsilon_{j2} & \epsilon_{j1}\epsilon_{\gamma2} & \epsilon_{\gamma1}\epsilon_{j2} & \epsilon_{\gamma1}\epsilon_{\gamma2} \end{pmatrix} \tag{2}$$

where $\epsilon_{\gamma1}$ and $\epsilon_{\gamma2}$ are the fractions of the leading and sub-leading photons that have passed the event selection and have $O_{NN} > 0.75$, and ϵ_{j1} and ϵ_{j2} are the fractions of jets that have passed the event selection and have $O_{NN} > 0.75$. The photon efficiency (ϵ_γ) is estimated using direct di-photon MC and corrected for small differences between data and the simulation measured in pure samples of photon events from radiative Z decays $Z \to \ell^+\ell^-\gamma$ ($\ell = e, \mu$). The jet efficiency (ϵ_j) is estimated using di-jet MC enriched in jets misidentification as photons, and cross-checked in jet samples in data. Both efficiencies are parameterized as a function of photon pseudorapidity. $(N_{\gamma\gamma}, N_{\gamma j}, N_{j\gamma}, N_{jj})$ can be obtained by solving the linear equation. Table II shows the results after applying the method on the real data.

Total	1084
Total - N_{DY}	1058
$N_{\gamma\gamma}$	745 ± 57
non-$\gamma\gamma$	313 ± 53

TABLE II: The number of $\gamma\gamma$ and non-$\gamma\gamma$ (sum of γ+jet and di-jet) events in the data samples from the 4x4 matrix method. The quoted uncertainties are statistical only.

We reverse the event selection O_{NN} cut (0.1) on one of the two photon candidates to get an enriched non-$\gamma\gamma(\gamma$+jet,di-jet) sample from data. Fig. 3 shows that the shape of the di-photon mass distribution from such "reversed-ONN" sample is in good agreement with the results from the 4×4 matrix method. Given the good agreement between the distribution from the two orthogonal samples and the low statistics of the results from the matrix method, we use the "reversed-ONN" sample to determine the shape of the non-$\gamma\gamma$ background. In order to smooth out statistical fluctuations, we fit the mass distribution with an exponential function, $f(M_{non}) = \exp(p_0 \cdot M_{non}^2 + p_1 \cdot M_{non} + p_2)$, with M_{non} denoting the mass of the two photon candidates. The estimation of the total number of non-$\gamma\gamma$ events from the 4x4 matrix method is used to fix the normalization. The systematic uncertainty from the shape function is taken into account when calculating the limits.
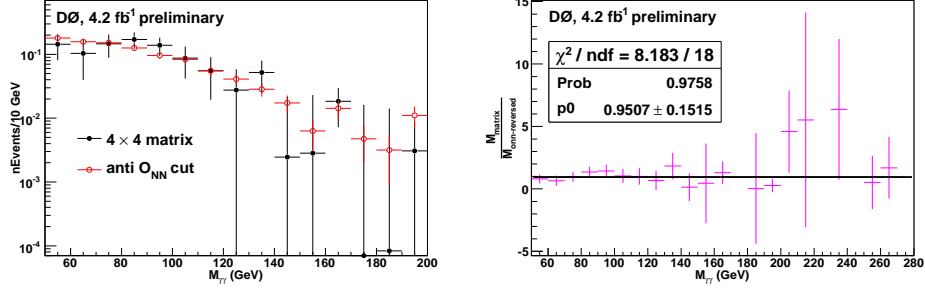
FIG. 3: non-$\gamma\gamma$ component invariant mass distribution (left) from 4×4 matrix background subtraction and from reversing the O_{NN} cut. The area of both histograms are normalized to unity. Right plot shows the corresponding ratio for the two distributions shown in the left plot.

C.   Direct di-photon production

We obtain a di-photon invariant mass distribution (with Drell-Yan, γ+jet, di-jet subtracted) and use the side-band fitting method to determine the direct di-photon background in the signal mass region (M_{h_f}-10 GeV, M_{h_f}+10 GeV). For each assumed Higgs mass (M_{h_f}), we use a simple exponential function, $f(M_{diem}) = \exp(p_0 \cdot M_{diem}^2 + p_1 \cdot M_{diem} + p_2)$, to fit the di-photon mass (M_{diem}) distribution in the $[60, 180]$ GeV range outside of the signal mass region (M_{h_f}-10 GeV, M_{h_f}+10 GeV). We then interpolate the function in the signal region to determine the direct di-photon contribution.

V.   SYSTEMATIC UNCERTAINTIES

The uncertainty of the $O_{NN} > 0.75$ efficiencies for the photon and photon-like jets is the main source of the uncertainty of the background subtraction. We adopt the difference in the number of background events from the mean efficiencies and the upper and lower uncertainty bands as the systematic uncertainty. The influence of the parton distribution functions (PDF) uncertainty on the acceptance is 0.6% - 1.0% depending on the Higgs mass, estimated from CTEQ6M error functions. Table III lists all the systematic uncertainties of this analysis:

source	uncertainty
luminosity	6.1% [15]
trigger	0.1%
PDF for $h_f \rightarrow \gamma\gamma$ acceptance	0.6% - 1.0%
electron misidentification efficiency	19.0%
Z/γ^*(ee) cross section	3.9%
photon identification efficiency	6.8%
background subtraction	shape (10%-15%)
photon energy scale	shape (0.6%)

TABLE III: Systematic uncertainties. The photon energy scale and background subtraction systematic uncertainties change the signal and background shape, and are taken into account appropriately in the final limits setting.

VI.   FINAL EVENT DISTRIBUTIONS AND LIMITS

A.   Final event distributions

For illustration, we show the invariant mass distribution of the events in data together with those of the background estimation for the mass values region [60 GeV, 180 GeV] in Fig. 4, where the Drell-Yan contribution is estimated from MC, the non-$\gamma\gamma$ component is estimated from data by using 4×4 matrix method and the direct di-photon production is estimated from data by using side-band fitting for 130 ± 10 GeV mass window. Figure 5 shows the invariant mass of the two photon candidates in the interval of the (M_{h_f}-10 GeV, M_{h_f}+10 GeV) for some of the assumed Higgs masses.
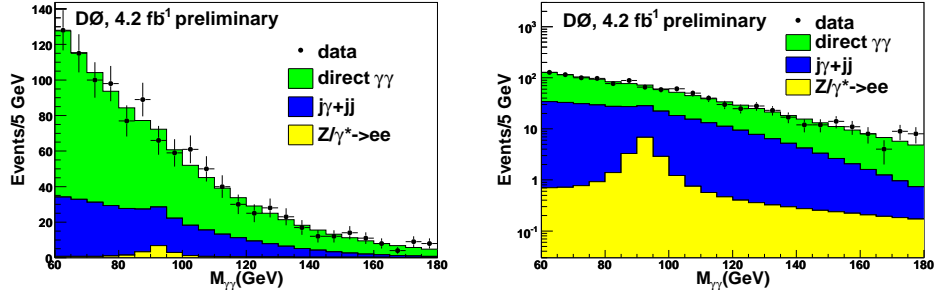
FIG. 4: Invariant mass distribution of two photon candidates in mass region [60 GeV, 180 GeV] with linear and log scales.

The shaded region corresponds to the expected background systematic uncertainty band. Table IV shows the number of events in data, expected background and signal in each mass interval for different SM Higgs mass values.

	80 GeV	90 GeV	100 GeV	110 GeV	120 GeV	130 GeV	140 GeV	150 GeV
ϵ_{sel}(VH)	0.117±0.001	0.132±0.001	0.148±0.001	0.156±0.001	0.161±0.001	0.168±0.001	0.176±0.001	0.176±0.001
ϵ_{sel}(VBF)	0.133±0.001	0.144±0.001	0.152±0.001	0.160±0.001	0.164±0.001	0.169±0.001	0.170±0.001	0.173±0.001
$Z/\gamma^* \to ee$	6±1	14±3	12±2	3±1	2±0	1±0	1±0	1±0
γj+jj	109±23	91±19	73±16	55±12	40±9	27±6	18±4	11±2
direct $\gamma\gamma$	242±16	191±13	140±11	106±8	87±6	65±4	51±3	37±2
total background	357±9	296±9	225±7	164±6	129±4	93±3	70±2	49±2
data	364	291	236	181	123	93	64	49

TABLE IV: Event selection efficiencies (ϵ_{sel}) with their statistical errors for signal, number of events in data, and the background estimation in the mass interval (M_{h_f}-10 GeV, M_{h_f} + 10 GeV) from 80 GeV to 150 GeV in 10 GeV steps, where the systematic uncertainties have been included for the background, the correlation between the different backgrounds has been considered when calculating the uncertainty for the total background.
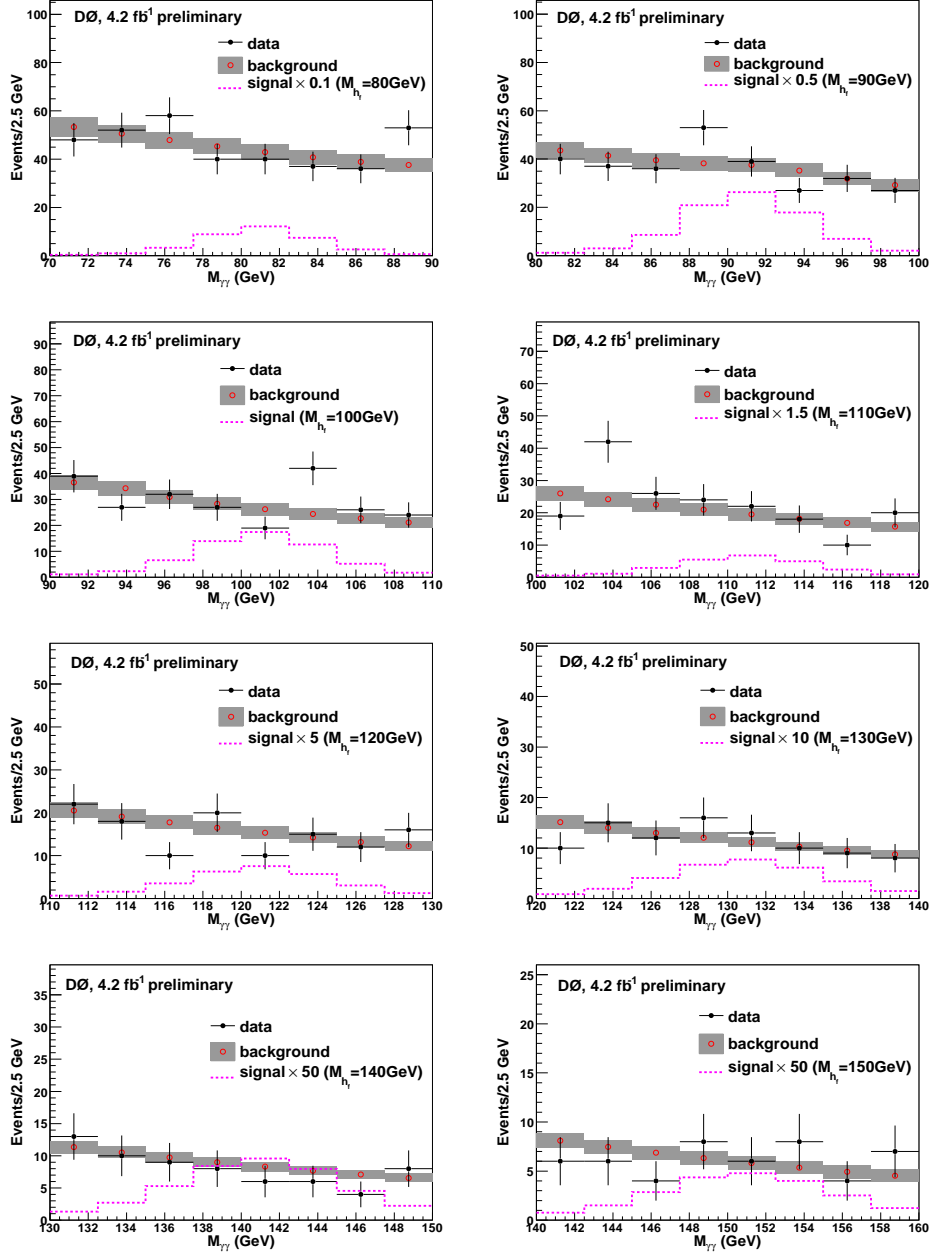
FIG. 5: Invariant mass distribution of the two photon candidates in the mass interval $(M_{h_f}-10 \text{ GeV}, M_{h_f}+10 \text{ GeV})$ for different Higgs mass assumptions.

B. Limit setting

We proceed to set upper limits on the Higgs production cross section times branching ratio for Higgs decaying into a pair of photons. The distributions of invariant mass of the two photon candidates in the interval of (M_{h_f}-10 GeV, M_{h_f} + 10 GeV) (shown in Fig. 5) are used for this purpose. Limits are calculated at the 95% confidence level using the modified frequentist CLs approach with a Poisson log-likelihood ratio test statistic [16, 17]. The impact of systematic uncertainties is incorporated via convolution of the Poisson probability distributions corresponding to the different sources of systematic uncertainty. The correlation in systematic uncertainties are maintained between signal and backgrounds. In this analysis, we use the diphoton invariant mass as the discriminant for limits setting. We find that the selection efficiency varies smoothly and the di-photon mass resolution is almost constant (3 GeV). Therefore, we are able to determine the limits in every 2.5 GeV mass values by interpolation of the efficiencies and mass distributions of the signal using MC samples generated at every 5 GeV mass points. Table V and Fig. 6(left) show the limits on $\sigma \times BR(h_f \rightarrow \gamma\gamma)$ for the different Higgs masses. By assuming the SM cross section for the associated and vector boson fusion Higgs production mechanims, we derive upper limits on the $BR(h_f \rightarrow \gamma\gamma)$ as a function of Higgs mass (see Fig. 6(right)). As it can be appreciated, this search considerably extends the range excluded by LEP and a previous DØ result.

M_{h_f}	80	90	100	102.5	105	107.5	110	112.5	115	117.5	120	122.5	125	127.5	130	132.5	135	137.5	140	142.5	145	147.5	150
$\sigma \times BR_{Obs}$ (fb)	46.5	61.9	54.3	71.7	80.8	63.7	45.4	33.1	24.3	22.7	23.3	26.3	30.6	33.4	31.6	27.5	21.8	17.5	14.7	14.3	16.4	18.1	18.7
$\sigma \times BR_{exp}$ (fb)	58.9	51.5	40.2	38.9	37.8	35.5	33.7	33.6	33.0	31.5	29.7	29.4	28.1	27.0	25.4	24.8	23.9	23.0	21.9	21.3	20.7	20.0	22.2

TABLE V: 95% C.L. limits on $\sigma \times BR$ for different fermiophobic Higgs masses.
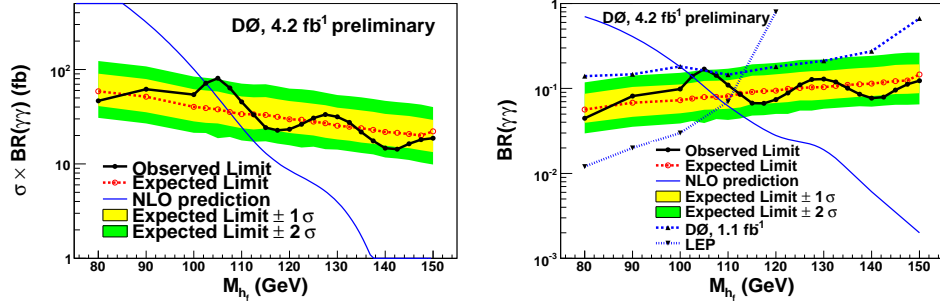

FIG. 6: 95% C.L. limits on $\sigma \times BR$ (left) and BR (right) as a function of fermiophobic Higgs mass.

VII.   SUMMARY

This note describes a search for the fermiophobic Higgs boson in the di-photon channel in 4.2 fb^{-1} DØ Run II data. The data and SM background estimation are consistent, so we set 95% C.L. limits on the $\sigma \times BR$ for different fermiophobic Higgs masses. The results is better than the previous Tevatron results [3] after removing the luminosity effect, and have reached the same sensitivity as a single LEP experiment, setting a lower limit on the fermiophobic Higgs of $M_{h_f} > 102.5$ GeV ($M_{h_f} > 107.5$ GeV expected). We are slightly below the combined LEP limit ($M_{h_f} > 109.7$ GeV). We also provide access to the $M_{h_f} > 125$ GeV regions which is inaccessible by LEP.

Acknowledgments

[1] S. Mrenna and J. Wells, arXiv:hep-ph/0001226 (2000).
[2] A. Heister *et al.* (ALEPH Collaboration), Phys. Lett. B **544**, 16 (2002);
    P. Abreu *et al.* (DELPHI Collaboration), Phys. Lett. B **507**, 89 (2001); Eur. Phys. J. C **35**, 313, (2004);
    G. Abbiendi *et al.* (OPAL Collaboration), Phys. Lett. B **544**, 44 (2002);
    P. Achard *et al.* (L3 Collaboration), Phys. Lett. B **534**, 28 (2002); Phys. Lett. B **568**, 191 (2003);
    LEP (LEP Higgs Working Group) arXiv:hep-ex/0107035 (2001).
[3] B. Abbott *et al.* (DØ Collaboration), Phys. Rev. Lett. **82**, 2244 (1999);
    T. Affolder *et al.* (CDF Collaboration), Phys. Rev. D **64**, 092002 (2001);
    V.M. Abazov *et al.* (DØ Collaboration), Phys. Rev. Lett. **101**, 051801 (2008).
[4] DØ Collaboration, DØ Note 5858-CONF, (2009).
[5] T. Sjöstrand *et al.*, Comput. Phys. Commun. **135**, 238 (2001).
[6] H.L. Lai *et al.*, Phys. Rev. D **55**, 1280 (1997).
[7] R. Brun and F. Carminati, CERN Program Library Long Writeup W5013, 1993 (unpublished).
[8] S. Catani *et al.*, JHEP **0307**, 028 (2003).
[9] K.A. Assamagan *et al.*, arXiv:hep-ph/0406152 (2004).
[10] A. Djouadi, J. Kalinowski and M. Spira, Comput. Phys. Commun. **108**, 56 (1998).
[11] V. M. Abazov *et al.*, Nucl. Instrum. Methods A **565**, 463 (2006).
[12] C. Peterson, T. Rognvaldsson and L. Lonnblad, "JETNET 3.0 A versatile Artificial Neural Network Package", Lund University Preprint LU-TP 93-29. Version 3.5 is used here.
[13] R. Hamberg, W. L. van Neerven and T. Matsuura, Nucl. Phys. B **359**, 343 (1991) [Erratum-ibid. B **644**, 403 (2002)].
[14] Y. Liu, Ph.D. thesis, FERMILAB-THESIS-2004-37 (2004).
[15] T. Andeen *et al.*, FERMILAB-TM-2365 (2007).
[16] W. Fisher, FERMILAB-TM-2386-E (2006).
[17] T. Junk, Nucl. Intrum. Methods A **434**, 435 (1999); A. Read, CERN 2000-005 (2000).