DØ Note 6181-CONF

Search for a Fermiophobic Higgs Boson in the di-photon final state using 8.2 fb^{-1} of DØ data

The DØ Collaboration
URL http://www-d0.fnal.gov
(Dated: March 10, 2011)

This note describes a search for a Fermiophobic Higgs boson in the di-photon final state using DØ data based on an integrated luminosity of 8.2 fb^{-1}, collected at the Fermilab Tevatron collider from April 2002 to December 2010. We set 95% C.L. upper limits on the cross section times the branching ratio ($\sigma \times BR(h_f \to \gamma\gamma)$) and the branching ratio ($BR(h_f \to \gamma\gamma)$) for different assumed Higgs masses from 100 GeV to 150 GeV, well beyond the sensitivity of the combined LEP experiments. We exclude Fermiophobic Higgs particles with a mass $M_{h_f} < 112$ GeV, which is currently the best limit.

Preliminary Results for Winter 2011 Conferences

I. INTRODUCTION

A search for standard model (SM) Higgs bosons in $H \to \gamma\gamma$ channel is difficult for its small branching ratio. For instance, it is $\sim 0.22\%$ for a Higgs boson with a mass of 130 GeV. However, the branching ratio can be enhanced significantly in some models beyond the SM [1]. In the model of Fermiophobic Higgs bosons (h_f), which assumes zero coupling of the Higgs boson to fermions, the branching ratio can be enhanced by about an order of magnitude (see Table. I). This hypothesis has been searched for at LEP [2] and Tevatron [3]. At Tevatron, the main SM Higgs boson production mechanism, gluon-gluon fusion that involves a top quark loop, is suppressed in the model of Fermiophobic Higgs bosons, which leaves associate vector boson ($h_f + V \to \gamma\gamma + V$, $V = W, Z$) and vector boson fusion (VBF $h_f \to \gamma\gamma$) processes as the main production mechanisms in this model. The coupling strength of a Fermiophobic Higgs boson to V is assumed to be the same as that of a SM Higgs boson.

In this analysis, we use 8.2 fb^{-1} of DØ Run II data and adopt the same analysis technique as Ref [4]. Multivariate Analysis Technique (MVA) [5] is used to combine five kinematic variables to build a final discriminant between signal and background to bring more sensitivity.

m_{h_f} (GeV)	100	110	120	130	140	150
$BR(H \to \gamma\gamma)$	0.0015	0.0019	0.0022	0.0022	0.0019	0.0014
$BR(h_f \to \gamma\gamma)$	0.18	0.062	0.028	0.019	0.0061	0.0020
$BR(h_f \to \gamma\gamma)/BR(H \to \gamma\gamma)$	120	33	13	9	3	1.4

TABLE I: Branching Ratio comparison for a SM Higgs boson and a Fermiophobic Higgs boson into two photons. $BR(H \to \gamma\gamma)$ and $BR(h_f \to \gamma\gamma)$ stand for the branching ratios of a SM Higgs boson and a Fermiophobic boson into two photons respectively.

II. DØ DETECTOR AND DATA SAMPLE

The DØ detector comprises a central tracking system in a 2 T superconducting solenoidal magnet, a liquid-argon/uranium calorimeter, a preshower detector and a muon spectrometer [6]. The major parts of the DØ detector used in the 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). The calorimeter has a central section (CC) covering up to $|\eta_{det}| \approx 1.1$ [7], and two end components (EC) extending coverage to $|\eta_{det}| \approx 4.2$. 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 $\eta_{det} - \phi_{det}$ [7] space except in the third layer, where it is 0.05×0.05. Immediately before the inner layer of the central EM calorimeter, there is 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. 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 energy (E_T) above 25 GeV, the trigger efficiency is close to 100%. The data sample used in this analysis was collected between April 2002 and December 2010 and corresponds to an integrated luminosity of 8.2 ± 0.5 fb^{-1} after applying standard data quality requirements.

The SM Higgs Monte Carlo (MC) samples, are generated using PYTHIA [8] with CTEQ6L [9] parton distribution functions (PDFs), and processed through a GEANT-3 based [10] simulation of the DØ detector. In order to accurately model the effects of multiple $p\bar{p}$ interactions and detector noise, the MC samples are overlaid with events from random $p\bar{p}$ crossings with a similar instantaneous luminosity profile as data, and then reconstructed using the same reconstruction software as data. Samples corresponding to the two dominant SM Higgs boson production mechanisms discussed above are generated, and normalized to the next-to-next-to-leading order (NNLO) theoretical cross sections for VH, and NLO for VBF [11]. The branching ratio predictions are from HDECAY [12].

III. EVENT SELECTION

Events are selected by requiring at least two photon candidates with $E_T > 25$ GeV and $|\eta_{det}| < 1.1$. Photons are selected from EM clusters reconstructed within a cone of radius $\mathcal{R} = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.2$ and satisfying the following requirements: (i) At least 95% of the cluster energy is deposited in the EM calorimeter; (ii) The calorimeter isolation variable $I = [E_{\rm tot}(0.4) - E_{\rm EM}(0.2)]/E_{\rm EM}(0.2)$ is less than 0.1, where $E_{\rm tot}(0.4)$ is the total energy in a cone of radius $\mathcal{R} = 0.4$ and $E_{\rm EM}(0.2)$ is the EM energy in a cone of radius $\mathcal{R} = 0.2$; (iii) The energy-weighted cluster width

in EM3 is consistent with an EM shower; (iv) The scalar sum of the transverse momentum p_T of all tracks originating from the primary vertex in an annulus of $0.05 < \mathcal{R} < 0.4$ around the cluster ($\sum p_T^{\rm trk}$) is less than 2 GeV; (v) In order to suppress electrons misidentified as photons, the EM clusters are required not to be spatially matched to tracker activity (track veto), i.e. either to a reconstructed track, or to a set of hits in the SMT and CFT consistent with that of an electron trajectory; (vi) The contribution of jets misidentified as photons is reduced by combining the information from a set of variables that are sensitive to differences between photons and jets on the tracker activity and the energy distributions in the calorimeter and CPS, using an artificial neural network (ANN) [13]. Fig. 1 compares the ANN output (O_{NN}) distribution for photons and jets. Photon candidates are required to have an ANN output O_{NN} larger than 0.1. Such a requirement is almost 100% efficient for photons while rejecting $\sim$ 40% of misidentified jets. (vii) Additionally, the di-photon invariant mass is required to be larger than 60 GeV, and the azimuthal angle between the two photon candidates is required to be larger than 0.5.

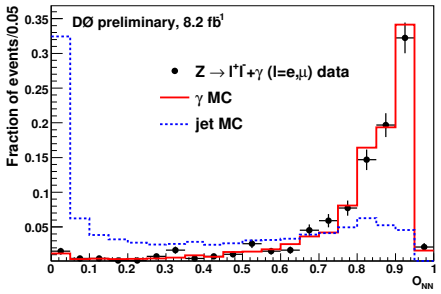


FIG. 1: Normalized distributions of O_{NN} value from real and fake photons.

IV. BACKGROUNDS

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

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 [14] is used for the normalization. From the $Z/\gamma* \to ee$ MC samples, we find that 2.0% of the electrons can satisfy the photon selection requirements described in Section III due to tracker inefficiencies. The contribution from the Drell-Yan process on the total background is found to be 699 $\pm$ 15 (stat.) events.

B. γ+jet and di-jet background

We estimate the γ + jet and di-jet contributions from the data after the event selection(see section III) by using a 4×4 matrix method described in detail in Ref. [15]. In this analysis, we use $O_{NN} = 0.75$ as a boundary to classify events into four categories:

- N_{pp} have both photon candidates with $O_{NN} > 0.75$;
- N_{pf} have the leading photon candidate with $O_{NN} > 0.75$, but the sub-leading $O_{NN} < 0.75$;
- N_{fp} have the leading photon candidate with $O_{NN} < 0.75$, but the sub-leading $O_{NN} > 0.75$;
- N_{ff} 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 simulation and are removed. The pass-fail vector (N_{pp}, N_{pf}, N_{fp}, N_{ff}) thus obtained is related to 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 γ+γ 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:

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

where $\epsilon_{\gamma 1}$ and $\epsilon_{\gamma 2}$ are the fractions of the leading and sub-leading photons that have passed the event selection and have $O_{NN} > 0.75$; 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 data.

Data	20734
N_{DY}	699 ± 15
$N_{\gamma\gamma}$	10905 ± 234
$N_{\gamma j} + N_{j\gamma}$	6207 ± 343
N_{jj}	2923 ± 183

TABLE II: The number of $\gamma\gamma$, γ+jet and di-jet events in the data sample from the 4x4 matrix method. The number of DY events is estimated from MC. The quoted uncertainties are statistical only.

Due to the limited statistics of non-$\gamma\gamma$ (sum of γ+jet and di-jet) events from the 4×4 matrix method, especially in the high mass region, we reverse the event selection cut $O_{NN} > 0.1$ on one or both of the two photon candidates to obtain an enriched γ+jet or di-jet orthogonal sample from data. The kinematic distributions from such "reversed-O_{NN}" samples are in good agreement with the non-$\gamma\gamma$ events from the 4×4 matrix method. Therefore we use the "reversed-O_{NN}" samples normalized by the corresponding numbers of non-$\gamma\gamma$ events from the 4x4 matrix method to model the non-$\gamma\gamma$ background.

C. Direct QCD di-photon production

We use SHERPA [16] Monte Carlo to model the direct QCD $\gamma\gamma$ background. The inclusive cross section of SHERPA has leading order accuracy while the photon fragmentation function is modeled by an interleaved QCD+QED parton shower including higher-order real-emission matrix elements.

V. SYSTEMATIC UNCERTAINTIES

Systematic uncertainties affecting the normalization and shape of the MVA output are estimated for both signal and backgrounds.

- 6.1% uncertainty on the integrated luminosity [17] affecting signal and Drell-Yan background normalization;
- 3% uncertainty on the photon identification efficiency affecting signal and Drell-Yan background normalization;
- 0.7% − 1.0% uncertainty on the signal acceptance from the PDFs uncertainty, estimated using the CTEQ6M [18] error functions;
- 9% uncertainty on the track veto efficiency and 4% on the cross section affecting the Drell-Yan background normalization;

- 3% uncertainty on the photon $O_{NN} > 0.75$ efficiency in the 4×4 matrix method affecting non-$\gamma\gamma$ and $\gamma\gamma$ backgrounds normalization and shape;
- 10% uncertainty on the jet $O_{NN} > 0.75$ efficiency in the 4×4 matrix method affecting non-$\gamma\gamma$ and $\gamma\gamma$ backgrounds normalization and shape.
- 10-20% uncertainty from the factorization and renomalization scale of SHERPA affecting $\gamma\gamma$ background shape.

VI. FINAL DISCRIMINANT DISTRIBUTIONS AND LIMITS

A. Final discriminant distributions

To improve the overall sensitivity, we use the Gradient Boosted Decision Tree method (BDTG) from the Toolkit for Multivariate Analysis [5] that combines five kinematic variables to build a final discriminant between the signal and background. The five kinematic variables we used are,

- leading photon transverse energy, E_T^1;
- sub-leading photon transverse energy, E_T^2
- di-photon invariant mass, $M_{\gamma\gamma}$;
- di-photon transverse momentum, $p_T^{\gamma\gamma}$;
- azimuthal angle between the two photon candidates, $\Delta\phi_{\gamma\gamma}$.

Fig. 2 shows these five kinematic distributions from data, backgrounds and the $M_{h_f} = 110$ GeV signal. The signal and total background samples are trained for every mass point displayed in Table III using events within the [M_{h_f} - 30GeV, M_{h_f} + 30GeV] mass window. At the 2.5 GeV mass points we interpolate the MVA input from the neighbouring 5 GeV points using the fact that the selection efficiency is almost independent of the di-photon mass and the mass resolution is approximately constant ($\sim$ 3 GeV). As an illustration, we show the MVA output distributions for six of the hypothetic Fermiophobic Higgs masses in Fig. 3.

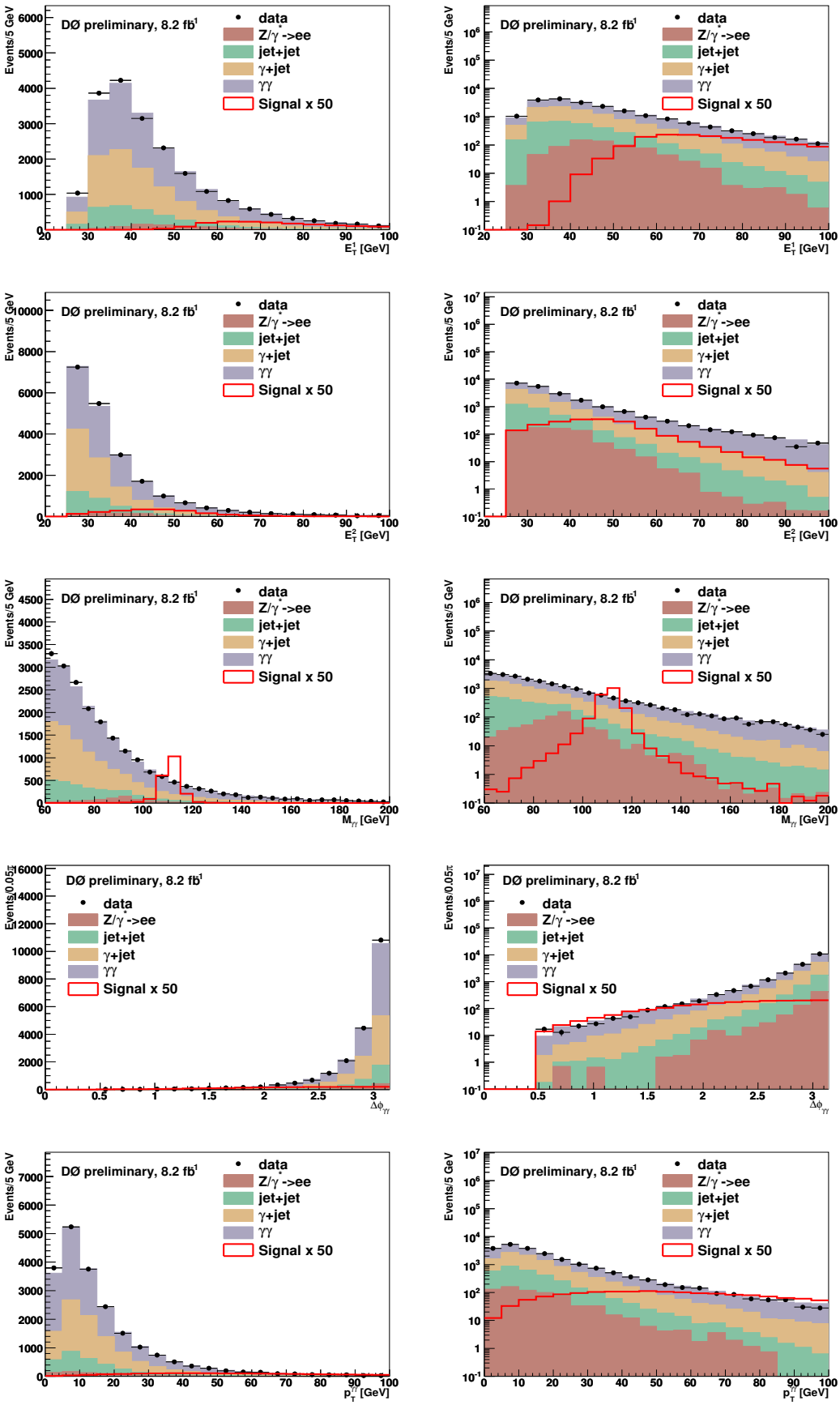
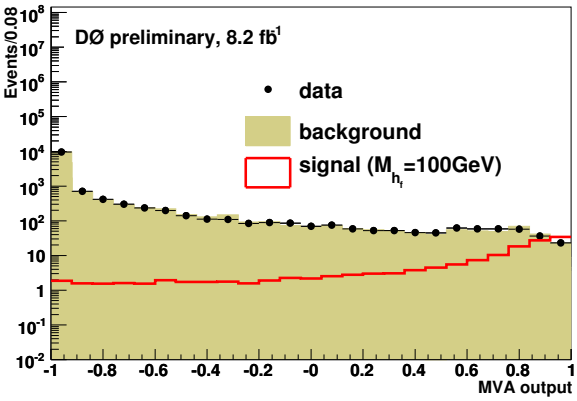
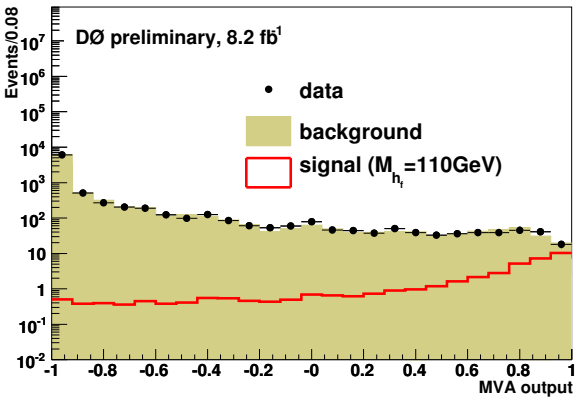


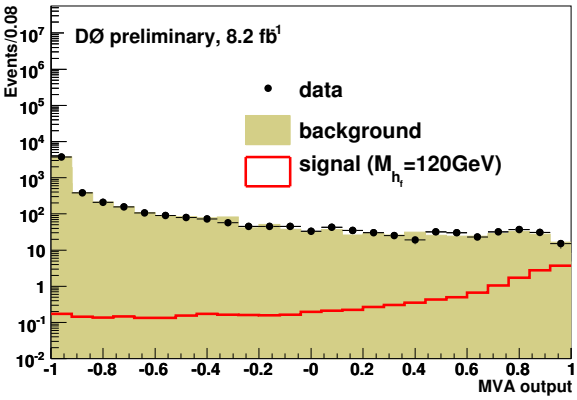
FIG. 2: Data and background modeling comparisons in terms of E_T^1, E_T^2, $M_{\gamma\gamma}$, $\Delta\phi_{\gamma\gamma}$ and $p_T^{\gamma\gamma}$ for the mass region [60, 200] GeV. A signal for $M_H = 110$ GeV is also shown. The plots in the left column are in linear scale and the plots in the right column are in log scale.



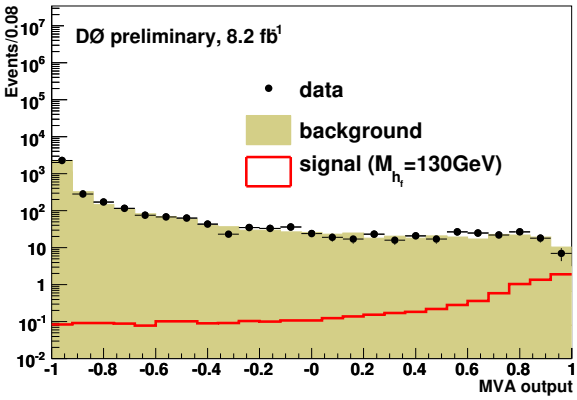
(a) $M_{h_f} = 100$ GeV



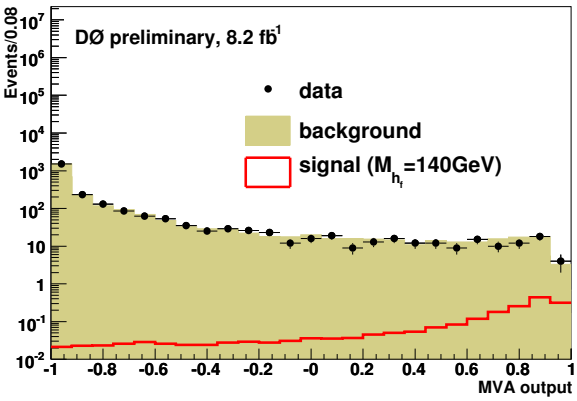
(b) $M_{h_f} = 110$ GeV



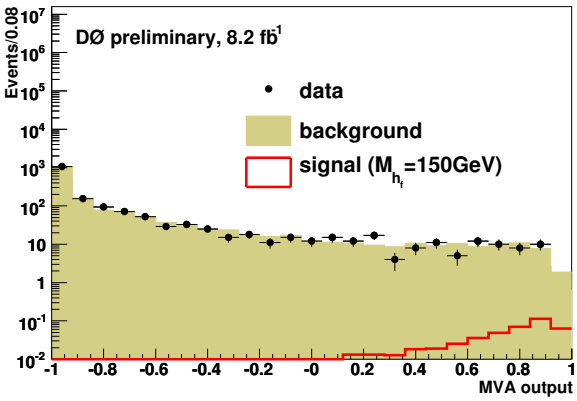
(c) $M_{h_f} = 120$ GeV



(d) $M_{h_f} = 130$ GeV



(e) $M_{h_f} = 140$ GeV



(f) $M_{h_f} = 150$ GeV

FIG. 3: MVA output distributions for $M_{h_f} = 100 - 150$ GeV in 10 GeV intervals. Each mass point has a mass window selection of ± 30 GeV.

Higgs mass(GeV)	100	102.5	105	107.5	110	112.5	115	117.5	120	122.5
observed limit(fb)	26.9	27.6	32.7	37.6	23.1	18.7	24.0	22.9	24.4	18.9
expected limit(fb)	24.7	23.1	20.6	20.5	19.2	19.8	18.8	17.9	17.3	15.2

Higgs mass(GeV)	125	127.5	130	132.5	135	137.5	140	142.5	145	147.5	150
observed limit(fb)	20.9	20.0	14.7	10.5	14.6	14.4	10.2	9.2	8.7	8.9	6.3
expected limit(fb)	14.1	13.9	14.4	13.6	12.8	12.5	11.7	12.0	11.1	10.2	10.1

TABLE III: 95% C.L. limits on $\sigma \times BR(h_f \to \gamma\gamma)$ for different Fermiophobic Higgs masses.

Higgs mass(GeV)	100	102.5	105	107.5	110	112.5	115	117.5	120	122.5
observed limit(%)	4.8	5.3	6.7	8.3	5.5	4.8	6.7	6.8	7.8	6.4
expected limit(%)	4.4	4.4	4.2	4.5	4.5	5.1	5.3	5.3	5.5	5.2

Higgs mass(GeV)	125	127.5	130	132.5	135	137.5	140	142.5	145	147.5	150
observed limit(%)	7.6	7.8	6.0	4.6	6.8	7.1	5.3	5.1	5.1	5.6	4.2
expected limit(%)	5.1	5.4	5.9	6.0	6.0	6.2	6.1	6.6	6.6	6.4	6.7

TABLE IV: 95% C.L. limits on $BR(h_f \to \gamma\gamma)$ for different Fermiophobic Higgs masses.

B. Limit setting

We set upper limits on the cross section times branching ratio and the branching ratio for a Fermiophobic Higgs boson decaying into a pair of photons, using the MVA output distributions for each mass point in the interval of [M_{h_f} - 30 GeV, M_{h_f} + 30 GeV]. The limits are calculated at the 95% confidence level using the modified frequentist CL$_S$ approach with a Poisson log-likelihood ratio test statistic [19]. Systematic uncertainties are treated as nuisance parameters constrained by their priors, and the best fits of these parameters to data are determined at each value of M_{h_f} by maximizing the likelihood ratio [20]. The correlations of the systematic uncertainties are maintained.

As an illustration, the background subtracted data distribution is shown in Fig. 4.

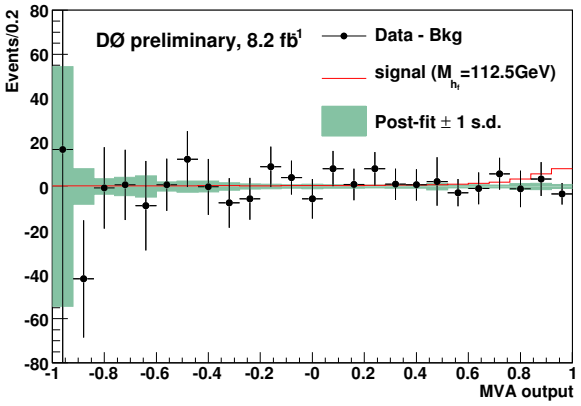


FIG. 4: Post-fit background subtracted data distribution for $M_{h_f} = 112.5$ GeV.

Table III and Fig. 5(left) show the upper limits on $\sigma \times BR(h_f \to \gamma\gamma)$ for the different hypothetic Fermiophobic Higgs masses. By assuming the SM cross section for the associated and vector boson fusion Higgs production mechanism, we also derive the upper limits on the $BR(h_f \to \gamma\gamma)$ as a function of the Fermiophobic Higgs mass (see Table IV and Fig. 5(right)).

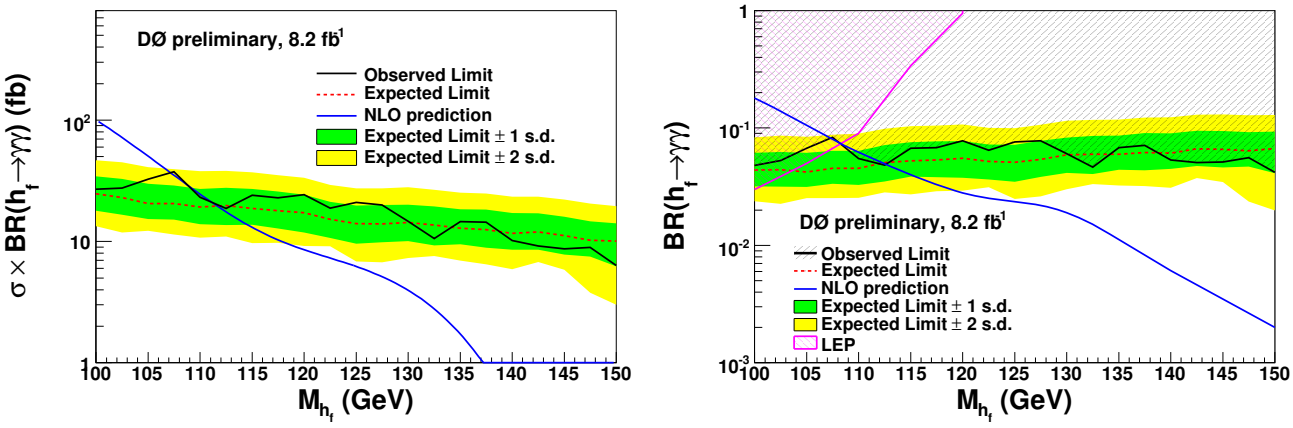


FIG. 5: 95% C.L. upper limits on $\sigma \times BR$ (left) and BR (right) as a function of Fermiophobic Higgs masses. The observed limit is shown as a solid black line while the expected limit under the background-only hypothesis is shown as a dashed red line. The green and yellow areas correspond to the 1 and 2 standard deviations (s.d.) around the expected limit.

VII. SUMMARY

We have presented a search for Fermiophobic Higgs bosons in the di-photon channel using 8.2 fb^{-1} DØ Run II data. We set 95% C.L. upper limits on the $\sigma \times BR(h_f \to \gamma\gamma)$ and $BR(h_f \to \gamma\gamma)$ for hypothetic Fermiophobic Higgs masses. By comparision with the 4.2 fb^{-1} results [21], the expected limits are improved typically by a factor of ~ 1.7, which is better than what is expected from the luminosity increase ($\sim 40\%$) due to the MVA technique. Moreover, the sensitivity is beyond that of the combined LEP experiments. We exclude Fermiophobic Higgs particles with a mass of $M_{h_f} < 112$ GeV, which is currently the best limit.

Acknowledgments

We thank the staffs at Fermilab and collaborating institutions, and acknowledge support from the DOE and NSF (USA); CEA and CNRS/IN2P3 (France); FASI, Rosatom and RFBR (Russia); CAPES, CNPq, FAPERJ, FAPESP and FUNDUNESP (Brazil); DAE and DST (India); Colciencias (Colombia); CONACyT (Mexico); KRF and KOSEF (Korea); CONICET and UBACyT (Argentina); FOM (The Netherlands); STFC (United Kingdom); MSMT and GACR (Czech Republic); CRC Program, CFI, NSERC and WestGrid Project (Canada); BMBF and DFG (Germany); SFI (Ireland); The Swedish Research Council (Sweden); CAS and CNSF (China); and the Alexander von Humboldt Foundation.

[1] S. Mrenna and J. Wells, Phys. Rev. D **63**, 015006 (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 **6177-CONF** (2011). http://www-d0.fnal.gov/Run2Physics/WWW/documents/Run2Results.htm
[5] A. Hoecker *et al.*, TMVA 4 Toolkit for Multivariate Data Analysis with ROOT Users Guide, arXiv:physics/0703039 (2007).
[6] V. M. Abazov *et al.*, Nucl. Instrum. Methods A **565**, 463 (2006).
[7] The polar angle θ and the azimuthal angle ϕ are defined with respect to the positive z axis, which is along the proton beam direction. Pseudorapidity is defined as $\eta = ln[tan(\theta/2)]$. Also, η_{det} and ϕ_{det} are the pseudorapidity and the azimuthal angle measured with respect to the center of the detector.
[8] T. Sjöstrand *et al.*, Comput. Phys. Commun. **135**, 238 (2001). We use PYTHIA version v6.323 with Tune A.

[9] H.L. Lai *et al.*, Phys. Rev. D **55**, 1280 (1997).
[10] R. Brun and F. Carminati, CERN Program Library Long Writeup W5013, 1993 (unpublished).
[11] The Tevatron New Phenomena and Higgs Working Group, "Cross Section and Branching Ratio Recommendations for Tevatron Higgs Searches", (2011).
[12] A. Djouadi, J. Kalinowski and M. Spira, Comput. Phys. Commun. **108**, 56 (1998).
[13] DØ Collaboration, DØ Note **5858-CONF** (2007). http://www-d0.fnal.gov/Run2Physics/WWW/documents/Run2Results.htm
[14] R. Hamberg, W. L. van Neerven and T. Matsuura, Nucl. Phys. B **359**, 343 (1991) [Erratum-ibid. B **644**, 403 (2002)].
[15] Y. Liu, Ph.D. thesis, FERMILAB-THESIS-2004-37 (2004).
[16] T. Gleisberg and S. Hoche and F. Krauss and M. Schonherr and S. Schumann and F Siegert and J. Winter, Event generation with Sherpa 1.1, JHEP 02 (2009) 007 [arXiv:0811.4622].
[17] T. Andeen *et al.*, FERMILAB-TM-2365 (2007).
[18] J. Pumplin *et al.*, JHEP **0207**, 012 (2002).
[19] T. Junk, Nucl. Intrum. Methods A **434**, 435 (1999); A. Read, CERN 2000-005 (2000).
[20] W. Fisher, FERMILAB-TM-2386-E (2006).
[21] DØ Collaboration, DØ Note **5880-CONF** (2009). http://www-d0.fnal.gov/Run2Physics/WWW/documents/Run2Results.htm