Search for resonant diphoton production with the D0 detector

We present a search for a narrow resonance in the inclusive diphoton final state using $\sim 2.7 \text{ fb}^{-1}$ of data collected with the D0 detector at the Fermilab Tevatron $pp$ Collider. We observe good agreement between the data and the background prediction, and set the first 95% C.L. upper limits on the production cross section times the branching ratio for decay into a pair of photons for resonance masses between 100 and 150 GeV. This search is also interpreted in the context of several models of electroweak symmetry breaking with a Higgs boson decaying into two photons.

PACS numbers: 14.80.Bn, 13.85.Rm, 13.85.Qk

At a hadron collider, diphoton ($\gamma\gamma$) production allows detailed studies of the Standard Model (SM) [1], as well
as searches for new phenomena, such as new heavy resonances [2], extra spatial dimensions or cascade decays of heavy new particles [3]. Within the SM, continuum $\gamma\gamma+X$ production is characterized by a steeply-falling $\gamma\gamma$ mass ($M_{\gamma\gamma}$) spectrum, on top of which a heavy resonance decaying into $\gamma\gamma$ can potentially be observed. In particular, this is considered one of the most promising discovery channels for a light SM Higgs boson at the LHC [4], despite the small branching ratio of $B(H \rightarrow \gamma\gamma) \sim 0.2\%$ for $110 < M_H < 140$ GeV [5, 6]. At the Tevatron, the dominant SM Higgs boson production mechanism is gluon fusion ($gg \rightarrow H$, or GF), followed by associated production with a $W$ or $Z$ boson ($qq' \rightarrow VH, V=W, Z$), and vector boson fusion ($VV \rightarrow H$, or VBF) [7, 8, 9]. While the SM Higgs production rate at the Tevatron is not sufficient to observe it in the $\gamma\gamma$ mode, the $Hgg$ and $H\gamma\gamma$ couplings, being loop-mediated, are particularly sensitive to new physics effects. Furthermore, in some models beyond the SM [2], $B(H \rightarrow \gamma\gamma)$ can be enhanced significantly relative to the SM prediction.

In this Letter, we present a search for a narrow resonance in the $M_{\gamma\gamma}$ spectrum using a data sample collected by the D0 detector [10] at the Fermilab Tevatron Collider. The selection of an inclusive $\gamma\gamma+X$ sample and the use of the $M_{\gamma\gamma}$ spectrum make the results of this search quasi-model independent. We use the SM Higgs boson ($H_{SM}$) with $H \rightarrow \gamma\gamma$ as a reference model, resulting in the first such search at the Tevatron, and a forerunner to similar planned searches at the LHC. Additionally, we consider other models of electroweak symmetry breaking (EWSB) with enhanced $B(H \rightarrow \gamma\gamma)$ [2], as a consequence of suppressed Higgs couplings to either (i) all fermions (known as bosonic or fermiophobic Higgs boson, $H_I$); (ii) only down-type fermions ($H_u$, as it gives mass to up-type fermions); or (iii) only top and bottom quarks (known as electroweak Higgs boson, $H_{ew}$). All models considered have SM-like production cross sections, with the exception of the $H_I$ and $H_{ew}$ models, where GF is absent or has a negligibly small cross section.

The subdetectors most relevant to this analysis are: the central tracking system, composed of a silicon microstrip tracker (SMT) and a central fiber tracker (CFT) embedded in a 2 T solenoidal magnetic field, the central preshower (CPS), and the liquid-argon and uranium sampling calorimeter. The CPS is located immediately before the inner layer of the calorimeter and is formed of one radiation length of absorber followed by several layers of scintillating strips. The calorimeter consists of a central section with coverage of $|\eta| < 1.1$ [11], and two end calorimeters covering up to $|\eta| \simeq 4.2$. The electromagnetic (EM) section of the calorimeter is segmented into four longitudinal layers (EM$i$, $i = 1, 4$) with transverse segmentation of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ [11], except in EM3, where it is $0.05 \times 0.05$. The calorimeter is well-suited for a precise measurement of electron and photon energies, providing a resolution of $\sim 3.6\%$ at energies of $\sim 50$ GeV. The data used in this analysis were collected using triggers requiring at least two clusters of energy in the EM calorimeter and correspond to an integrated luminosity of $2.7 \pm 0.2$ fb$^{-1}$ [12].

Events are selected by requiring at least two photon candidates with transverse momentum $p_T > 25$ GeV and $|\eta| < 1.1$, for which the trigger requirements are fully efficient. The photon $p_T$ is computed with respect to the reconstructed event primary vertex (PV) with the highest number of associated tracks, which is required to be within 60 cm of the geometrical center of the detector along the beam axis. The PV reconstruction efficiency in $\gamma\gamma+X$ events is $\sim 98\%$, with $\sim 95\%$ probability to match the true vertex. Photons are selected from EM clusters reconstructed within a cone with radius $R = \sqrt{\Delta\eta^2 + (\Delta\phi)^2} = 0.2$ by requiring: (i) $\geq 97\%$ of the cluster energy is deposited in the EM calorimeter; (ii) the calorimeter isolation variable $I = |E_{isol}(0.4) - E_{EM}(0.2)|/E_{EM}(0.2) < 0.1$, where $E_{isol}(R)$ ($E_{EM}(R)$) is the total EM energy in a cone of radius $R$; (iii) the energy-weighted shower width in the $r-\phi$ plane in EM3 is $< \sqrt{14}$ cm; and (iv) the scalar sum of the $p_T$ of all tracks ($p_T^{\text{jet}}$) originating from the primary vertex in an annulus of $0.05 < R < 0.4$ around the cluster is $< 2$ GeV. To suppress electrons misidentified as photons, the EM clusters are required to not be spatially matched to tracker activity, either a reconstructed track, or a density of hits in the SMT and CFT consistent with that of an electron [13]. To suppress jets misidentified as photons, a neural network (NN) is trained using a set of variables sensitive to differences between photons and jets in the tracker activity and in the energy deposits in the calorimeter and CPS: $p_T^{\text{jet}}$, the numbers of cells above a threshold $E_{EM}$ in EM1 within $R < 0.2$ and $0.2 < R < 0.4$ of the EM cluster, the number of EM clusters within $R < 0.1$ of the EM cluster, and the squared-energy-weighted width of the energy deposit in the CPS. The NN is trained using $\gamma\gamma$ and dijet Monte Carlo (MC) samples and its performance is verified using a data sample of $Z \rightarrow \ell^+\ell^-\gamma$ ($\ell = e, \mu$) events. Figure 1a compares the NN output ($O_{NN}$) spectrum for photons and jets. Photon candidates are required to have $O_{NN} > 0.1$, which is $\sim 98\%$ efficient for real photons and rejects $\sim 50\%$ of misidentified jets. Finally, $M_{\gamma\gamma}$, computed from the two highest $p_T$ photons, is required to be $> 60$ GeV. In total, 5608 events are selected in data.

All MC samples used in this analysis are generated using PYTHIA [14] with CTEQ6L [15] parton distribution functions (PDFs), and processed through a GEANT-based [16] simulation of the D0 detector and the same reconstruction software as the data. Signal samples are generated separately for GF, VH and VBF production and normalized using the theoretical cross sections [7, 8, 9] and branching ratio predictions from HDECAY [5].

This analysis is affected by instrumental backgrounds.
such as $\gamma$+jet, dijet and $Z/\gamma^{*} \rightarrow e^{+}e^{-}$ (ZDY) production, with jets or electrons misidentified as photons, as well as an irreducible background from direct diphoton production (DDP). All backgrounds, except for ZDY, are estimated directly from data.

The ZDY background is estimated using the MC simulation, normalized to the next-to-next-to-leading-order cross section \cite{17}. The selection efficiencies determined by the MC simulation are corrected to the corresponding values measured in the data. On average each electron has a 2\% probability to satisfy the photon selection criteria, mainly due to the inefficiency of the track-match veto requirements. The total contribution from ZDY is estimated to be 88 ± 10 events.

Backgrounds due to $\gamma$+jet and dijet events are directly estimated from data by using a 4 x 4 matrix background estimation method \cite{18}. After final event selection, a tightened $O_{NN}$ requirement ($O_{NN} > 0.75$) is used to classify the events into four categories depending on whether the two highest-$p_T$ photons, only the leading photon, only the trailing photon or neither of the two photons, satisfy this requirement. The corresponding numbers of events, after subtraction of the estimated ZDY contributions, are denoted as $N_{pp}$, $N_{pf}$, $N_{fp}$ and $N_{ff}$. The different relative efficiency of the $O_{NN} > 0.75$ requirement between real photons and jets allows the estimation of the sample composition by solving a linear system of equations: $(N_{pp}, N_{pf}, N_{fp}, N_{ff})^T = \mathcal{E} \times (N_{\gamma\gamma}, N_{\gamma j}, N_{jj}, N_{jj})^T$, where $N_{\gamma\gamma}$ ($N_{jj}$) is the number of $\gamma\gamma$ (dijet) events and $N_{\gamma j}$ ($N_{jj}$) is the number of $\gamma$+jet events with the leading (trailing) cluster as the photon. The 4 x 4 matrix $\mathcal{E}$ contains the efficiency terms (parameterized as a function of $|\eta|$), estimated in photon and jet MC samples and validated in data. The estimated sample composition is $N_{\gamma\gamma} = 3155 \pm 125$ (stat), $N_{\gamma j} = 1680 \pm 149$ (stat) and $N_{jj} = 685 \pm 93$ (stat).

The shape of the $M_{\gamma\gamma}$ spectrum for the sum of the $\gamma$+jet and dijet backgrounds is obtained from an independent control data sample by requiring $O_{NN} < 0.1$ for one of the photon candidates, and is parameterized with an exponential function. The resulting shape is found to be in excellent agreement with that derived by directly applying the 4 x 4 matrix method bin-by-bin in the final selected sample, but has smaller statistical fluctuations, especially in the high $M_{\gamma\gamma}$ region.

After subtraction of the ZDY, $\gamma$+jet and dijet background contributions, the $M_{\gamma\gamma}$ spectrum is examined for the presence of a narrow resonance. For each assumed $M_H$ value (between 100 and 150 GeV, in steps of 5 GeV), the search region is defined to be $(M_H - 15$ GeV, $M_H + 15$ GeV), where 15 GeV corresponds to about five times the expected $M_{\gamma\gamma}$ resolution. The DDP background is estimated by performing a sideband fit to the $M_{\gamma\gamma}$ spectrum in the 70 to 200 GeV range (this excludes the search region) using an exponential function (see Fig. 1b). Such a parameterization has been validated using a next-to-leading-order for this process \cite{19}.

Systematic uncertainties affecting the normalization and shape of the $M_{\gamma\gamma}$ spectrum are estimated for both signal and backgrounds. Uncertainties affecting the ZDY background normalization include: integrated luminosity (6.1\%), electron misidentification rate (14.3\%) and ZDY cross section (3.9\%). Such uncertainties are propagated, via the 4 x 4 matrix method, to the estimated normalization of the $\gamma$+jet and dijet background contributions, affected in addition by the uncertainty on the $O_{NN} > 0.75$ selection efficiency for photons (2\%) and jets (10\%). The uncertainty in the shape of the $\gamma$+jet and dijet $M_{\gamma\gamma}$ spectrum is given by the statistics of the control data sample used to parameterize it. The above uncertainties, as well as the statistical uncertainties of the sideband fitting method, result in systematic uncertainties in the normalization and shape of the DDP background contribution. Uncertainties affecting the signal normalization include: integrated luminosity (6.1\%), acceptance due to the photon identification efficiency (6.8\%) and PDFs
<table>
<thead>
<tr>
<th>$M_H$ (GeV)</th>
<th>100</th>
<th>110</th>
<th>120</th>
<th>130</th>
<th>140</th>
<th>150</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z/\gamma^* \rightarrow e^+ e^-$</td>
<td>55±7</td>
<td>17±3</td>
<td>64±2</td>
<td>5±1</td>
<td>4±1</td>
<td>3±1</td>
</tr>
<tr>
<td>$\gamma \gamma$</td>
<td>742±62</td>
<td>481±42</td>
<td>324±34</td>
<td>236±30</td>
<td>161±28</td>
<td>124±22</td>
</tr>
<tr>
<td>$\gamma j + jj$</td>
<td>540±66</td>
<td>319±39</td>
<td>204±25</td>
<td>133±16</td>
<td>89±11</td>
<td>61±8</td>
</tr>
<tr>
<td>total background</td>
<td>1337±29</td>
<td>817±26</td>
<td>534±19</td>
<td>374±12</td>
<td>254±7</td>
<td>188±5</td>
</tr>
<tr>
<td>data</td>
<td>1385</td>
<td>827</td>
<td>544</td>
<td>357</td>
<td>270</td>
<td>202</td>
</tr>
<tr>
<td>$H_{SM}$ signal</td>
<td>1.62±0.11</td>
<td>1.61±0.11</td>
<td>1.51±0.10</td>
<td>1.26±0.08</td>
<td>0.90±0.06</td>
<td>0.54±0.04</td>
</tr>
<tr>
<td>acceptance (%)</td>
<td>19.9,18.8,20.3</td>
<td>20.4,19.9,21.6</td>
<td>21.0,20.6,22.3</td>
<td>21.5,21.2,22.9</td>
<td>21.8,22.0,23.5</td>
<td>22.1,22.2,24.1</td>
</tr>
</tbody>
</table>

TABLE I: Numbers of selected events in data, expected backgrounds, expected $H_{SM}$ signal and signal acceptance (for each production mechanism: GF, VH, VBF), in the search region for different $M_H$ values. The expected signal includes contributions from GF, VH and VBF processes, the latter two representing $\sim 21-24\%$ of the total signal.

<table>
<thead>
<tr>
<th>$M_H$ (GeV)</th>
<th>100</th>
<th>105</th>
<th>110</th>
<th>115</th>
<th>120</th>
<th>125</th>
<th>130</th>
<th>135</th>
<th>140</th>
<th>145</th>
<th>150</th>
</tr>
</thead>
<tbody>
<tr>
<td>exp. $\sigma \times B$</td>
<td>88</td>
<td>78</td>
<td>71</td>
<td>63</td>
<td>58</td>
<td>53</td>
<td>49</td>
<td>45</td>
<td>41</td>
<td>39</td>
<td>36</td>
</tr>
<tr>
<td>obs. $\sigma \times B$</td>
<td>78</td>
<td>110</td>
<td>112</td>
<td>76</td>
<td>50</td>
<td>46</td>
<td>42</td>
<td>53</td>
<td>56</td>
<td>66</td>
<td>52</td>
</tr>
</tbody>
</table>

TABLE II: Observed and expected 95% C.L. upper limits on $\sigma \times B$ (in fb) for different $M_H$ values. The expected limit is defined as the median of the distribution of limits in background-only pseudo-experiments.

(1.7-2.2%) [15]. Finally, the location of the peak in the $M_{\gamma\gamma}$ spectrum for signal is affected by the uncertainty in the relative data to MC photon energy scale (0.6%).

Table I shows the number of events in data, expected background and expected $H_{SM}$ signal in six different search regions. The inset in Fig. 1b illustrates the $M_{\gamma\gamma}$ spectrum in the search region for $M_H = 130$ GeV, found to be in good agreement with the background prediction. The $M_{\gamma\gamma}$ spectrum in the search region is used to derive upper limits on the production cross section times branching ratio for $H \rightarrow \gamma\gamma (\sigma \times B)$ as a function of $M_H$.

The SM prediction for the ratio of the production cross sections for the three signal production mechanisms is assumed. Limits are calculated at the 95% C.L. using the modified frequentist approach with a Poisson log-likelihood ratio test statistic [20, 21]. The impact of systematic uncertainties is incorporated via convolution of the Poisson probability distributions for signal and background with Gaussian distributions corresponding to the different sources of systematic uncertainty. The correlations in systematic uncertainties are maintained between signal and backgrounds.

The resulting limits on $\sigma \times B$ are given in Table II, and displayed in Fig. 1c. Although the SM Higgs boson is used as a reference model, the fact that the signal acceptance is found to be almost independent of the production mechanism (see Table I), makes the estimated limits applicable to other models of new physics with a narrow resonance decaying into $\gamma\gamma$. In the context of models of EWSB with enhanced $B(H \rightarrow \gamma\gamma)$, the current search excludes a $H_1$ boson with $M_H < 101$ GeV at 95% C.L., improving (slightly) upon previous results at the Tevatron [22]. While none of the other EWSB scenarios explored can currently be excluded, the expected sensitivity is within less than a factor of four of the prediction for the $H_a$ model for $M_H < 110$ GeV, and only a factor $\sim 20$ above the SM prediction for $115 \leq M_H \leq 130$ GeV.

As a result, this search contributes to the overall sensitivity of the SM Higgs boson search at the Tevatron from the combination of multiple channels [23]. Assuming the same integrated luminosity in all channels and a single Tevatron experiment, this analysis is expected to improve the combined upper limit on the SM Higgs production cross section by $\sim 5\%$ for $115 \leq M_H \leq 130$ GeV. Finally, this search is used to derive 95% C.L. upper limits on $B(H \rightarrow \gamma\gamma)$ between 14.1% and 33.9% for $M_H$ in the range 100-150 GeV, in the case of models where the Higgs boson does not couple to the top quark. Conversely, for models where the GF production mode is available, this inclusive search allows improvement of the upper limits on $B(H \rightarrow \gamma\gamma)$ to 3.4%-7.2% in the same mass range. These represent the most stringent limits on $B(H \rightarrow \gamma\gamma)$ for $M_H$ in the range 100-150 GeV to date, significantly improving upon previous results from LEP and the Tevatron [22].

In summary, we have performed an inclusive search for a narrow resonance with mass between 100 and 150 GeV decaying into $\gamma\gamma$ at the Tevatron. This channel is used to increase the overall sensitivity of the SM Higgs boson search program at the Tevatron [23] and allows the probe of new physics models predicting an enhanced rate for $H \rightarrow \gamma\gamma$.

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); 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 (Germany).
[a] Visitor from Augustana College, Sioux Falls, SD, USA.
[b] Visitor from Rutgers University, Piscataway, NJ, USA.
[c] Visitor from The University of Liverpool, Liverpool, UK.
[d] Visitor from II. Physikalisches Institut, Georg-August-University, Göttingen, Germany.
[e] Visitor from Centro de Investigacion en Computacion - IPN, Mexico City, Mexico.
[f] Visitor from ECFM, Universidad Autonoma de Sinaloa, Culiacan, Mexico.
[g] Visitor from Helsinki Institute of Physics, Helsinki, Finland.
[h] Visitor from Universität Bern, Bern, Switzerland.
[i] Visitor from Universität Zürich, Zürich, Switzerland.
[j] Deceased.

[6] Throughout this Letter we adopt units where $c = 1$.
[11] Pseudorapidity is defined as $\eta = -\ln[\tan(\theta/2)]$, where $\theta$ is the polar angle relative to the proton beam direction. $\phi$ is defined as the azimuthal angle in the plane transverse to the proton beam direction.