PHYSICAL REVIEW D **87,** 073005 (2013)

Finding the Higgs boson in decays to $Z\gamma$ using the matrix element method at next-to-leading order

John M. Campbell,[*] R. Keith Ellis,[†] Walter T. Giele,[‡] and Ciaran Williams[§]

Theoretical Physics Department, Fermilab, P. O. Box 500, Batavia, Illinois 60510, USA

(Received 4 March 2013; published 10 April 2013)

We illustrate how the matrix element method at next-to-leading order can be used to discriminate between events arising from the production of a Higgs boson, which subsequently decays to a final state consisting of $\ell^+\ell^-\gamma$, and the background production of the same final state. We illustrate how the method could be used in an experimental analysis by devising cuts on the signal (P_S) and background (P_B) weights that are computed event by event in this approach. We find that we can increase the $S/\sqrt{B}$ ratio by around 50% compared to an invariant mass fit on its own. Considering only statistical uncertainty, this is equivalent to recording a factor of around two times more integrated luminosity.

DOI: 10.1103/PhysRevD.87.073005 PACS numbers: 14.80.Bn, 13.85.Qk

I. INTRODUCTION

The recent discovery of a new boson with properties consistent with that of the Standard Model (SM) Higgs boson [1,2] has indicated that the discovery of the electroweak (EW) symmetry breaking mechanism may be at hand. In order to confirm whether the new boson is indeed the SM Higgs, it is crucial to measure both its properties and branching ratios for the largest number of experimentally viable decay channels. These analyses could result in tension with the SM Higgs prediction; for instance, the boson may differ in parity from the SM Higgs or even be a mixture of CP-odd and -even states. An additional possibility is that the rate for one or more measured decay channels is different from the SM prediction. The most obvious mechanism for such a scenario is an enhancement/suppression in loop-induced decays that are naturally sensitive to couplings to new virtual particles, for instance, the decay to two photons ($\gamma\gamma$). This would thus be evidence for beyond the Standard Model physics.

Another loop-induced Higgs decay is the decay to a final state containing a Z boson and a photon ($H \to Z\gamma$) [3,4]. Since it is a loop-induced process, the branching ratio is small, and the decay of the Z boson to well-measured final-state particles ($\mu^+\mu^-$ or e^+e^-) means that the decay $H \to \ell^+\ell^-\gamma$ is a very rare SM process. However, this is not necessarily the case in extensions of the SM. In addition, the ratio of $\gamma\gamma$ to $Z\gamma$ branching ratios can be used to discriminate between certain models of new physics [5–7].

Observing the Higgs in the $Z\gamma$ final state is a difficult feat. First, the background production of $Z\gamma$ is several orders of magnitude larger than the Higgs induced rate. The exact value of this ratio depends on the cuts defining the background cross section. In the region of invariant mass near the Higgs mass ($m_{\ell\ell\gamma} \sim 125$ GeV), with typical LHC cuts, one would usually expect around 500–1000 background events for each signal event. Second, the kinematics of the decay limit the final-state photon to a challenging region of phase space. At leading order (LO), the maximum transverse momentum (p_T) of the final-state photon is restricted, since the final-state invariant mass is close to m_H and includes a pair of charged leptons of mass close to the Z mass (m_Z). Since m_Z is not too far from m_H, the remaining energy to be imparted to the photon lies in a limited range. Therefore, the p_T spectrum of the photon peaks around 30 GeV.

Typically, in such a soft region of phase space, QCD can provide large backgrounds to searches. This increases the difficulty in separating signal from background when compared with $H \to \gamma\gamma$, for which the photon p_T from the signal is significantly harder ($p_T^\gamma \sim 60$ GeV). Indeed, once the full detector simulation has been included, there are only small differences between the signal and background shapes in the transverse variables [8]. In terms of final-state kinematics, the main discriminating variable is the angle between the direction of the photon and the beam [9,10]. A spin-0 scalar is isotropic in this variable, while the background matrix element prefers emission in the forward region. A recent CMS study with around 10 fb^{-1} of 7 and 8 TeV data [8] set a limit around 10 times the SM cross section. This result already disfavors scenarios in which the new boson is a pure pseudoscalar since, in some of these models, the branching ratio to $Z\gamma$ can be enhanced by up to 170 times the SM prediction [11].

An experimental search for $H \to Z\gamma$ should thus utilize as much theoretical information as possible in order to effectively reduce the unwanted background $Z\gamma$ events. One such method is the matrix element method (MEM) [12–16]. This method uses the matrix element associated with a given theoretical hypothesis to assign a probabilistic weight to an experimental event. Comparing weights obtained by varying the theoretical hypothesis allows one to identify the most favorable one. Originally, the method

[*]johnmc@fnal.gov
[†]ellis@fnal.gov
[‡]giele@fnal.gov
[§]ciaran@fnal.gov

was used in order to perform a measurement of a known theoretical parameter, say, for instance, the top mass [17–20]. More recently, the method has gained favor as an event-by-event discriminant [9,21–25]. By using the matrix element, one naturally includes all of the kinematic correlations present in the observed final state and thus gains a large amount of theoretical information. Until recently, a major drawback of the MEM was its restriction to LO matrix elements. However, in Ref. [26], a new version of the MEM was proposed that can be extended to higher orders in perturbation theory.[1] Using the next-to-leading order (NLO) method provides a much greater degree of theoretical reliability and control over the theoretical systematics.

Recently, the MEM has been used in searches for, and studies of, the Higgs boson. The MEM at LO has been applied to Higgs searches in the ZZ [21] and $Z\gamma$ [9] channels. The MEM has also recently been applied to study the properties of the Higgs (decaying to two photons) via vector boson fusion [22] and to investigate its role in unitarizing WW scattering [23]. The $Z\gamma$ search [9] used an implementation of the MEM that is restricted to LO and considered the $Z\gamma$ decay in an effective field theory approach. The authors found only a marginal improvement between the MEM and a simpler approach that only used $m_{\ell\ell\gamma}$ as the discriminant. With the recent CMS study to guide us [8], the aim of this paper is to reinvestigate the channel using the matrix element method at next-to-leading order (MEM@NLO) algorithm and the full loop matrix element for the decay.

This paper proceeds as follows. In Sec. II, we discuss the $H \to Z\gamma$ calculation and the form of the matrix elements used in our analysis. Section III provides a brief overview of the MEM@NLO technique, and we present our results in Sec. IV. Finally, in Sec. V, we draw our conclusions.

II. THE HIGGS DECAY TO $Z\gamma$

In this section, we briefly discuss the calculation of $H \to Z\gamma$ as it is implemented in the code MCFM [27,28]. The $H \to Z\gamma$ decay was first considered over twenty years ago [3,4]. We consider the process,

$$H(p_0) \to \ell^+(p_3) + \ell^-(p_4) + \gamma(p_5), \tag{1}$$

where the momenta are shown in parentheses. The squared matrix element for this loop-induced process has the following form:

$$|M_{H\to\ell^+\ell^-\gamma}|^2 = \frac{e^8 s_{34}(s_{35}^2 + s_{45}^2)}{2\sin^2\theta_W (16\pi^2 m_W)^2}(|\mathcal{F}_L|^2 + |\mathcal{F}_R|^2). \tag{2}$$

[1]With the caveat that the final state of interest should consist of only EW particles.

In this expression, we have introduced the electroweak coupling e, and the weak mixing angle θ_W and kinematic invariants are defined through $s_{ij} = (p_i + p_j)^2$. The left-handed and right-handed amplitudes are defined by

$$\mathcal{F}_{L,R} = \frac{4Q_t N_c m_t^2}{s_{34}}\left(Q_t Q_\ell + \frac{1}{2}(v_L^t + v_R^t)v_{L,R}\mathcal{P}_Z\right)F_t + (Q_\ell + v_{L,R}^\ell \cot\theta_W \mathcal{P}_Z)F_W, \tag{3}$$

in terms of the charge of the leptons and top quarks (Q_ℓ, Q_t), the number of colors (N_c), and the top mass (m_t). The vector couplings are defined as

$$v_L^\ell = \frac{-1 - 2Q_\ell \sin^2\theta_W}{2\sin 2\theta_W}, \quad v_R^\ell = -\frac{2Q_\ell \sin^2\theta_W}{\sin 2\theta_W}, \tag{4}$$

$$v_L^t = \frac{1 - 2Q_t \sin^2\theta_W}{2\sin 2\theta_W}, \quad v_R^t = -\frac{2Q_t \sin^2\theta_W}{\sin 2\theta_W}. \tag{5}$$

Finally, the function $\mathcal{P}_Z$ describes the Z propagator (with width Γ_Z),

$$\mathcal{P}_Z = \frac{s_{34}}{s_{34} - m_Z^2 + i\Gamma_Z m_Z}. \tag{6}$$

The loop integral functions are contained in F_W and F_t, which are defined by

$$F_W = 2\left[\frac{s_{345}}{m_W^2}\left(1 - 2\frac{m_W^2}{s_{34}}\right) + 2\left(1 - 6\frac{m_W^2}{s_{34}}\right)\right]C_2(p_5, p_{34}, m_W, m_W, m_W) + 4\left(1 - 4\frac{m_W^2}{s_{34}}\right)C_0(p_5, p_{34}, m_W, m_W, m_W) \tag{7}$$

and

$$F_t = C_0(p_5, p_{34}, m_t, m_t, m_t) + 4C_2(p_5, p_{34}, m_t, m_t, m_t). \tag{8}$$

In these expressions, C_0 and C_2 are standard Passarino-Veltman tensor integrals. Further details can be found in Refs. [29–31].

In addition to the decay described above, MCFM contains NLO Higgs and QCD $Z\gamma$ production, including the $gg \to Z\gamma$ loop induced process [28]. We will use these matrix elements to calculate our weights.

III. THE MEM@NLO TECHNIQUE

This section provides a brief overview of the MEM technique developed in Ref. [26], to which we refer the interested reader for a more complete discussion. The crux of the MEM method is to provide an event—by—event

weight using the matrix element. At LO, one defines each event to be weighted by the following quantity:

$$\tilde{P}_{\mathrm{LO}}(\tilde{\phi}) = \frac{1}{\sigma_{\mathrm{LO}}} \int dx_1 dx_2 d\phi\, \delta(x_1 x_2 s - Q^2) \times f^j(x_1) f^i(x_2) \mathcal{B}_{ij}(x_1, x_2, \phi) W(\phi, \tilde{\phi}). \quad (9)$$

Here, $f(x)$ represent the parton distribution function (PDF) with momentum fraction x, Q^2 is the invariant mass of the EW final state, $\mathcal{B}_{ij}$ represents the LO matrix element, which depends on the final-state phase space point ϕ, which is derived from the input event from data ($\tilde{\phi}$), via the transfer function $W(\phi, \tilde{\phi})$. The weights are normalized by the LO cross section σ_{LO}. Often in this paper, we will use the following weight, which is defined in the limit of a perfect detector setup, i.e., $W(\phi, \tilde{\phi}) = \delta(\phi - \tilde{\phi})$,

$$P_{\mathrm{LO}}(\phi) = \frac{1}{\sigma_{\mathrm{LO}}} \int dx_1 dx_2 \delta(x_1 x_2 s - Q^2) \times f^j(x_1) f^i(x_2) \mathcal{B}_{ij}(x_1, x_2, \phi). \quad (10)$$

This weight has the advantage of requiring fewer Monte Carlo integrations, and hence it is less computationally expensive. However, one must be confident that the analysis is not sensitive to such a simplifying assumption. For instance, in this study, the narrow width of the Higgs would spoil this assumption, since any event with $m_{\ell\ell\gamma} = m_H$ would result in a large weight compared to the remaining events in the sample, yielding unrealistic results. Therefore, in order to use the above definition, care must be taken with variables (in this case, the invariant mass) that are extremely sensitive to detector resolution. We shall discuss this further in the next section.

The observed EW final state typically recoils against hadronic activity that is not modeled in the leading-order calculation. In order for the weights to be well-defined and unique, one must, therefore, perform a boost to ensure that the final-state ϕ is balanced in p_T. Then one can apply the PDF weighting assuming two beams colliding in the z direction. Since there are multiple Lorentz transformations satisfying these requirements which are connected by longitudinal boosts to each other, we integrate over the allowed range. This results in the corresponding integration over x_1 (or x_2) in Eq. (9). We refer to the set of p_T balanced final-state frames collectively as the MEM frame. We note that failure to perform this boost and subsequent integration results in either an ill-defined (no-boost) or nonunique (no integration over boosts) weights and hence a theoretically unreliable weight.

The MEM frame allows the calculation of weights accurate to NLO defined as [26]

$$\tilde{P}_{\mathrm{NLO}}(\tilde{\phi}) = \frac{1}{\sigma_{\mathrm{NLO}}} \int d\phi (V(\phi) + R(\phi)) W(\phi, \tilde{\phi}). \quad (11)$$

The virtual corrections are expressed as

$$V(\phi) = \int dx_1 dx_2 \delta(x_1 x_2 s - Q^2) \times f^j(x_1) f^i(x_2) \hat{V}_{ij}(x_1, x_2, \phi), \quad (12)$$

where $\hat{V}$ represents the contributions from the Born matrix element, the interference between the Born and one-loop amplitudes and the integrated form of a relevant subtraction term (in this work, we use a slightly modified Catani-Seymour [32] dipole approach). The real radiation pieces involve integration over an unresolved emission for which we use the forward-branching phase space (FBPS) generator described in Refs. [33,34],

$$R(\phi) = \int dx_1 dx_2 d\phi_{\mathrm{FBPS}} \delta(x_1 x_2 s - Q^2_{\mathrm{FBPS}}) \times f^j(x_1) f^i(x_2) \hat{\mathcal{R}}_{ij}(x_1, x_2, \phi, \phi_{\mathrm{FBPS}}). \quad (13)$$

In the above, $\hat{\mathcal{R}}_{ij}$ represents the matrix element for the Born amplitude plus one additional parton, rendered finite by the corresponding subtraction terms. Note that the constraining delta function for the PDFs has changed definition with respect to LO; Q^2_{FBPS} is the invariant mass of the Born final state plus the NLO emission. For full details of the FBPS and subtraction setup, we refer the reader to Ref. [26].

The main difference between the LO and NLO MEM is the integration over the real phase space. Some events in the lab frame, when mapped to the MEM frame, no longer lie in the fiducial region defined by the lab frame cuts (which the weights are normalized to) and therefore are assigned zero weight [26]. At NLO, these events can have nonzero weights since the real emission contributions can boost these events back into the fiducial region. A simple way of interpreting this phenomenon is that NLO covers a larger kinematic phase space than LO. This larger phase space manifests itself as an ability to accept events that do not possess the correct kinematics to have arisen from our LO discriminant. This is one of the primary advantages of the MEM@NLO method compared to the MEM@LO (in addition to the usual increase in confidence in the understanding of the theoretical systematic error from using an NLO prediction). Typically, one finds that $\mathcal{O}(\alpha_S)$ (i.e., of order 10%) of the events fail the LO cuts.

Each event in the data set can now be assigned a unique LO or NLO weight associated with a theoretical hypothesis controlled by the underlying matrix element. In our case, we will assign it a weight based on the signal matrix element, or the irreducible background production of $Z\gamma$. Throughout the paper, we will refer to these as signal and background weights as follows:

$$P_S = P^{H\to Z\gamma}_{\mathrm{NLO}}, \qquad P_B = P^{Z\gamma}_{\mathrm{NLO}}. \quad (14)$$

The superscript refers to the matrix element that is implicitly used in Eq. (11), and we take $W(\phi, \tilde{\phi}) = \delta(\tilde{\phi} - \phi)$. One can use these individual quantities to build

discriminants. Unless stated otherwise, our default is to use NLO matrix elements in our weight calculations.

IV. RESULTS

In order to study the MEM@NLO for the $H \to Z\gamma$ decay mode, we generate samples of signal and background events. We do this using the SHERPA event generator [35]. For the background, we generate a CKKW [36] matched sample of $Z\gamma$ events. For the signal events, we generate NLO matched Higgs events for $m_H = 125$ GeV. These Higgs events are then subsequently decayed to the $Z\gamma$ final state using the MCFM implementation that is described in Sec. II. Since the Higgs is a scalar particle, production and decay are uncorrelated. This allows us to simply calculate the decay using MCFM in the rest frame of the Higgs and then boost it so that it has the four-momentum of the SHERPA event. Throughout our studies, we will use the CT10 PDF set [37].

The above procedure produces events at the particle level. However, in order to study the light Higgs in a meaningful way, one must include some kind of detector simulation. This is because the light Higgs has such a narrow width that the $m_{\ell\ell\gamma}$ spectrum is dominated by the detector resolution. For example, the CMS technical design report [38] estimates a resolution of photon energy using the following:

$$\left(\frac{\Delta E_\gamma}{E_\gamma}\right) = \frac{3.6\%}{\sqrt{E_\gamma/\mathrm{GeV}}} \oplus \frac{18.5\%}{E_\gamma/\mathrm{GeV}} \oplus 0.66\%, \qquad (15)$$

where ΔE_γ represents the width of the Gaussian smearing and $\oplus$ indicates that the quantities are to be added in quadrature. At $E = 30$ GeV, this provides a width of around 0.3 GeV. Using this smearing (and the equivalent leptonic quantity), we find a Higgs boson line shape that is too narrow compared to that recently reported by CMS [8], where the effective Gaussian width quoted is around 3–4 GeV. Therefore, in order to match the results of this paper, we inflate our Gaussian smearing to

$$\Delta E_\gamma = 2 \text{ GeV}, \qquad \Delta E_\ell = 0.5 \text{ GeV}. \qquad (16)$$

Our resulting line shape for the Higgs is now in good agreement with Ref. [8]. Note that, since our enhanced width is around a factor of 6 greater than that arising from the energy-dependent piece, Eq. (15), we drop the energy dependence for simplicity.

After smearing our events as described above, we apply the following lab frame cuts:

$$\begin{aligned} &p_T^\gamma > 15 \text{ GeV}, \qquad |\eta^\gamma| < 2, \qquad p_T^\ell > 20 \text{ GeV}, \\ &|\eta^\ell| < 2, \qquad 60 < m_{\ell\ell} < 120 \text{ GeV}, \\ &115 < m_{\ell\ell\gamma} < 135 \text{ GeV}, \qquad R_{\ell\gamma} > 0.7. \end{aligned} \qquad (17)$$

Note that we have kept the cuts on the lepton and photon p_T loose. Part of the attraction of the MEM discriminant is that it will naturally select events that have the correct kinematics to be signal events; therefore, one does not have to spend time attempting to optimize the kinematic selection criteria. Of course, if some observable clearly discriminates the signal from the background, cuts on this quantity should be applied in order to reduce the overall computational load. For this reason, we impose a tight cut on the invariant mass of the $Z\gamma$ system, centered on 125 GeV.

A. Generation of reducible background events

In order to effectively simulate LHC conditions, we must also consider events that do not arise from the irreducible $Z\gamma$ background but instead are misidentified in the detector. Since they are naturally very dependent on the exact detector setup and modeling, such events are difficult to accurately simulate in our study. However, since they are a large fraction of the resulting event sample [8], it is necessary to attempt to provide an estimate of our discriminant on a ''fake'' sample. To this end, we generate a sample of fake events in the following way. We assume that the dominant component of the fake events results from Z + jet events in which the leptons from the Z decay are clean, but the jet is mismeasured as a photon. As a crude model, we use SHERPA to generate Z + jet events and then smear the p_T and η of the (leading) jet by Gaussian functions with a width of 10 GeV and 0.5, respectively. Our event sample is then generated by applying the cuts described in the previous section to the smeared events, treating the smeared jet as a photon.

B. The MEM@NLO as a kinematic discriminant

We first discuss our definition of the Higgs signal hypothesis, which is particularly important because of the very narrow SM width. One approach is to define a weight for each event by integrating over transfer functions that model the detector resolution, as in Eq. (9). This approach allows one to test a single Higgs mass hypothesis for a given set of events but requires additional integrations per event. An alternative approach is to change the Higgs mass hypothesis on an event-by-event basis by choosing $m_H = m_{\ell\ell\gamma}$. In this scenario, one effectively changes the propagator in the matrix element to

$$\frac{1}{(s - m_H)^2 + i\Gamma_H m_H} \to \frac{1}{i\Gamma_H m_{\ell\ell\gamma}}. \qquad (18)$$

Since Γ_H is very small, this approach makes each event have a large P_S, and the collection of events can no longer define a probability density function. In addition, $P_S \gtrsim P_B$ even for background event samples. However, one still expects discriminating power between the signal and

background since P_S arising from the events that match the signal hypothesis will be larger than that for P_S from the background. Finally, we note that the (signal) normalization is defined uniquely for each event by $\sigma(m_H = m_{\ell\ell\gamma})$.

This technique has been used extensively in studies involving kinematic discriminants in $H \to ZZ \to 4\ell$ [24,25]. The advantage of this technique is that there are less integrations per event, and thus the weights are computationally less expensive. One can then restore the

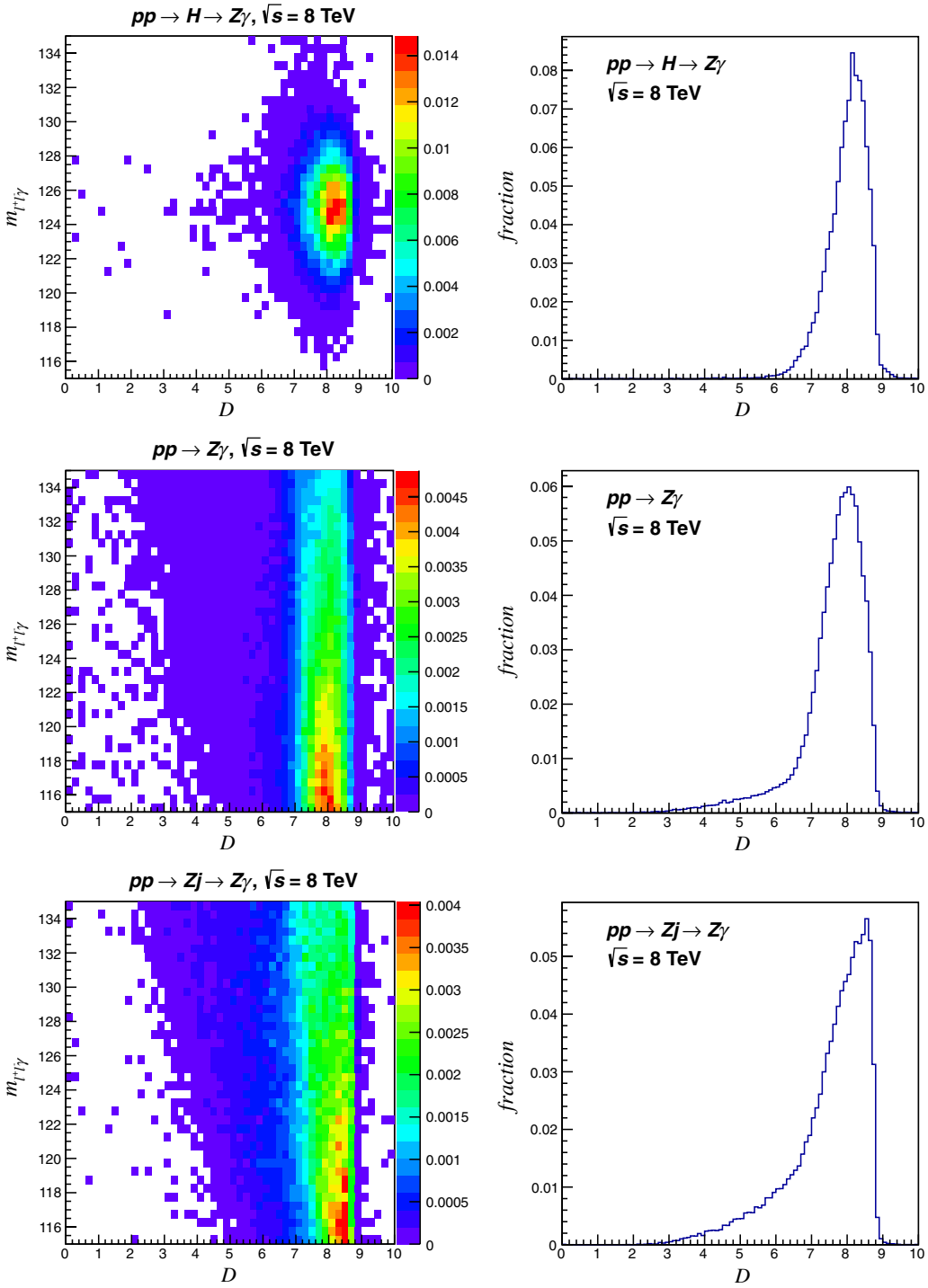

FIG. 1 (color online). Distribution of events in terms of the invariant mass of the final state $m_{\ell\ell\gamma}$ and our discriminant $\mathcal{D}$ defined in Eq. (19). The two-dimensional histograms (left) present the density of events in the plane of $\mathcal{D}$ and $m_{\ell\ell\gamma}$. The right-hand panels represent the distribution of $\mathcal{D}$ for our signal (top row), background (middle row), and fake (bottom row) samples. Each sample is normalized to the total number of events in the sample.

invariant mass $m_{\ell\ell\gamma}$ as an additional discriminant since Higgs events will cluster in invariant mass while the background will be more diffuse. We will adopt this approach for the remainder of this paper.

With the event samples generated as described in the previous sections, we can now introduce our discriminant $\mathcal{D}$. There is a range of possibilities, but, in this paper, we will choose

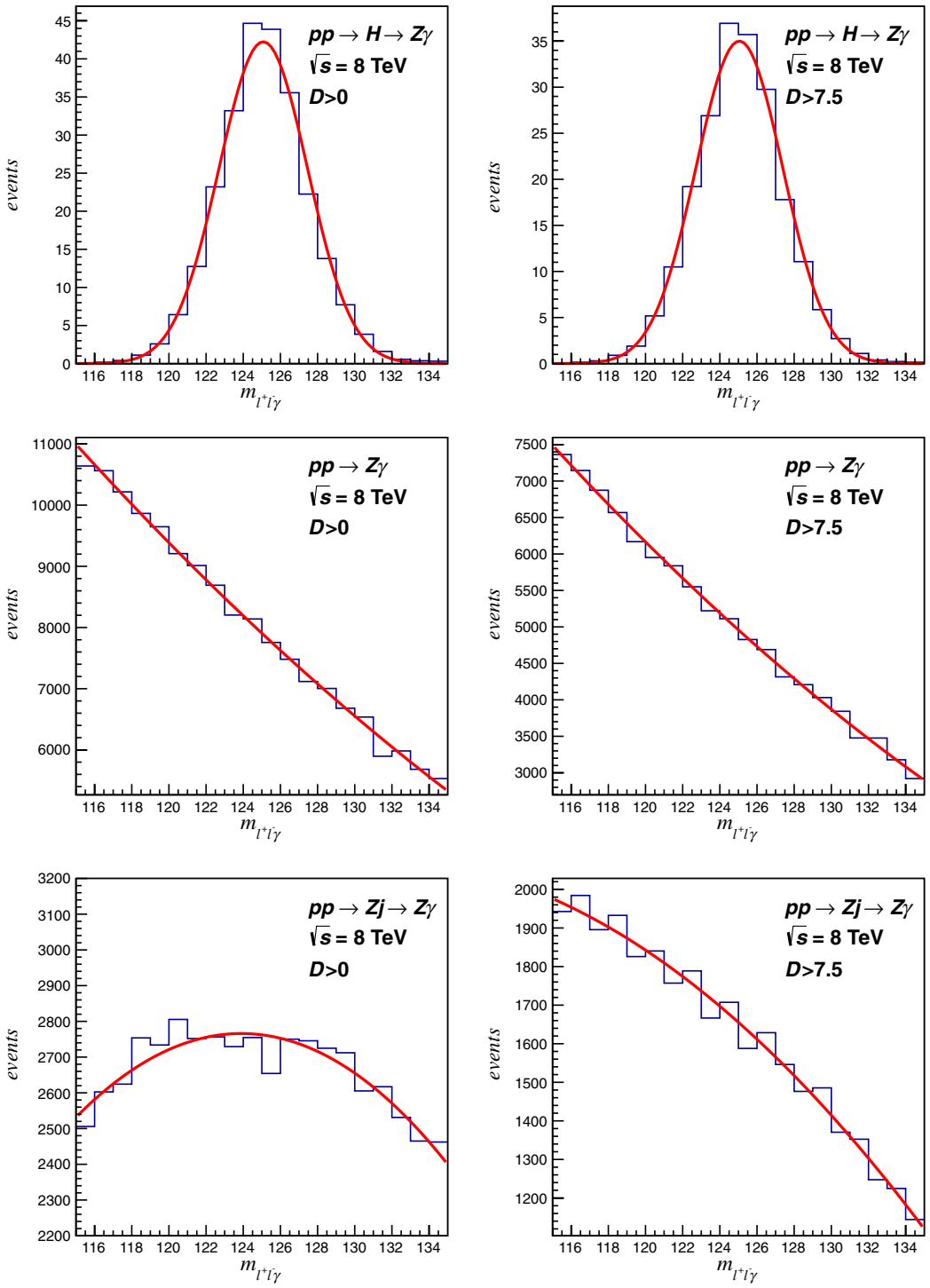

FIG. 2 (color online). Invariant mass distributions (for $m_{\ell\ell\gamma}$) before (left panels) and after (right panels) our analysis cut $\mathcal{D} > 7.5$. Distributions are shown for the signal events (top row), the irreducible background (middle row), and fakes (final row). The number of events in each distribution is normalized to the irreducible background sample without application of any cuts on $\mathcal{D}$, as described in the text. The red curve indicates a Gaussian (polynomial) fit to the signal (background) data.

$$\mathcal{D} = -\log\left(\frac{P_B}{P_S + P_B}\right). \quad (19)$$

Events that arise from the background should have larger P_B than P_S so the ratio in the logarithm is near one. As a result, events with $\mathcal{D}$ nearer zero should be more backgroundlike than the signal.

We present results for $\mathcal{D}$ for our three different event sample classes in Fig. 1. The results are shown as two-dimensional histograms, binned both by the discriminant $\mathcal{D}$ and the invariant mass $m_{\ell\ell\gamma}$. In addition, we also show one-dimensional projections of these histograms, as a function of $\mathcal{D}$ only. As expected, the signal events peak at larger $\mathcal{D}$ than the corresponding background distribution. The background and fake samples have roughly similar shapes (indicating some of the similarities between Z + jet and $Z\gamma$). Although the signal shapes are similar to the background, there are still significant regions that are only populated by background events (but which may still have an invariant mass in the Higgs window). In particular, both the background and fake samples have significant tails in the lower $\mathcal{D}$ region, whereas the signal sample does not. For example, there are barely any ($\sim 0.5\%$) signal events with $\mathcal{D} < 6$. On the other hand, around 10% of the background events lie in this region. Cutting at $\mathcal{D} > 6$ would thus be an almost zero-cost reduction in signal at the expense of a nontrivial background number of events.

The scales on the two-dimensional histograms illustrate the stark differences between the signal and background events in the $(\mathcal{D}, m_{\ell\ell\gamma})$ plane. The area of highest density for the signal events (around the truth value, $m_{\ell\ell\gamma} = 125$ GeV) is around three times greater than the corresponding highest density region for the background (which is at a much lower invariant mass). Retaining only the events that satisfy $\mathcal{D} > 7$, one rejects 21% of the irreducible background events and keeps 93% of the signal. A higher cut, $\mathcal{D} > 8$, rejects 64% of the background and retains 55% of the signal. In an experimental analysis, one would thus choose the optimal value of $\mathcal{D}$ at which to cut in order to optimize the signal-to-background ratio. Since our model of the fakes is less developed than our signal and background models, we optimize our cut on the discriminant on the combination of signal and irreducible background samples. We find a value of the cut at $\mathcal{D} > 7.5$ corresponds to a signal efficiency of 81%, with a background rejection of 37%. We note that, here, we have chosen a fairly simple cut on $\mathcal{D}$ that optimizes $S/\sqrt{B}$. One could instead perform counting experiments using contours in the $(\mathcal{D}, m_{\ell\ell\gamma})$ plane, although such a study is beyond the scope of this work.

We plot the invariant mass $m_{\ell\ell\gamma}$ before and after our cut ($\mathcal{D} > 7.5$) for our three samples in Fig. 2. Each sample is weighted to reflect the number of events expected, given the total number of irreducible background $Z\gamma$ events. We weight our signal sample by the ratio of cross sections (including a NLO to next-to-next-to-leading-order [NNLO] K factor of 1.2 [39]). We normalize our fake sample to be compatible with that reported by CMS [8], namely, by fixing the number of fake events to be one-third of the irreducible background. Our cuts have altered the shape of the background and fake samples, while maintaining the overall shape of the signal.

Ultimately, we would like to investigate the efficiency of this method in the vicinity of the Higgs signal at $m_H = 125$ GeV. We therefore define a window

$$122 < m_{\ell\ell\gamma} < 128 \text{ GeV}, \quad (20)$$

where the width has been optimized for the analysis below. We attempt to quantify the improvement the cut on $\mathcal{D}$ has made in the following way. We define the quantity

$$\alpha = \frac{\sqrt{N_{Z\gamma} + N_{\text{fakes}}}}{N_H}, \quad (21)$$

where N_X represents the expected number of events for process X. Our measure includes no treatment of systematic errors and instead only assumes the $S/\sqrt{B}$ scaling of the statistical uncertainty. In spite of its shortcomings compared to the full experimental analysis, α can provide us with an estimate of the improvement one might envisage after applying our cut. We find

$$\frac{\alpha_{D>0}}{\alpha_{D>7.5}} = 1.52. \quad (22)$$

Since α scales as $\mathcal{L}^{-1/2}$, using a cut of $D > 7.5$ is (statistically) equivalent to taking 2.31 times more data.

Before concluding, we will briefly consider the impact of using the leading order method, MEM@LO, rather than the NLO one. We find that the fraction of events that fail the fiducial cuts at LO is larger for the Higgs signal than for the irreducible background. As a result, the MEM@NLO produces slightly better signal over background ratios than the MEM@LO. For example, we find

$$\frac{\alpha^{\text{LO}}_{D>0}}{\alpha^{\text{LO}}_{D>7.5}} = 1.41, \quad (23)$$

which is 7% smaller than the corresponding NLO value. This small difference, between the LO and NLO analyses, indicates that the method is perturbatively stable, and the theoretical systematic uncertainty is under good control.

V. CONCLUSIONS

In this paper, we have presented an application of the MEM@NLO to searches for the Higgs boson in the decay channel $Z\gamma$. This channel is extremely challenging experimentally as can be seen from the preliminary results from CMS [8]. The reasons for these difficulties are twofold. First, the $H \to Z\gamma$ branching ratio is already very small, even before the requirement that the Z-boson

decays to muons and electrons only. Since the background production of Z in association with a photon is large, one naturally has low signal-to-background ratios. Second, the kinematics of the decay for a Higgs boson with mass $m_H \sim 125$ GeV force the final-state photon to have a relatively soft p_T. The matrix element has a soft singularity as $p_T^\gamma \to 0$, and therefore the background is very large in the region in which the Higgs signal peaks. Once detector effects are included, there is very little difference between the signal and background events in the transverse variables.

Given these difficulties, it is essential to utilize all the remaining differences between the signal and background processes. One approach, the matrix element method, uses a theoretically defined matrix element to assign a weight to each experimental event. When there is a good match between the theoretical hypothesis and the input events, the weights become larger. Therefore, one can use the MEM to produce samples of events that increase the signal-to-background ratio for a certain theoretical hypothesis. Recently, the MEM has been extended to NLO for electroweak final states [26]. We used this MEM@NLO to calculate signal and background discriminants for a sample of events generated using SHERPA. Our event sample included showered and hadronized Higgs signal and SM background events as well as a crude model of $Z\gamma$ fake events. Higgs decays to $Z\gamma$ were generated using the MCFM implementation. Our estimates of resolution effects and fake rates were guided by the recent results from CMS presented in Ref. [8].

We used the MEM@NLO to construct a discriminant ($\mathcal{D}$) from the event-by-event weights P_S (using the signal matrix element) and P_B (the background matrix element). In defining these weights, we removed the invariant mass as a discriminating variable. As a result, we were subsequently able to create a two-dimensional discriminant in $\mathcal{D}$ and $m_{\ell\ell\gamma}$. In this plane, the signal events cluster around m_H and at higher $\mathcal{D}$ compared to those arising from the background and fakes. Therefore, by cutting on $\mathcal{D}$ and $m_{\ell\ell\gamma}$, we were able to improve our measure of the signal significance, $S/\sqrt{B}$. We found that $S/\sqrt{B}$ increased by around a factor of 1.5 compared to the value obtained without any cut on $\mathcal{D}$, suggesting that roughly half as much data would be needed to obtain the same limit on $H \to Z\gamma$. We found that the MEM@LO algorithm is also able to provide $S/\sqrt{B}$ improvements by a factor of around 1.4, approximately 10% less efficient than the MEM@NLO.

This search has provided an example of the power of the matrix element method in a worst-case scenario for a traditional analysis. We hope that the ideas presented in this paper are useful to our experimental colleagues in the hunt for the Higgs boson in this difficult channel. The code which calculates the weights described in this paper is available upon request.

ACKNOWLEDGMENTS

Fermilab is operated by Fermi Research Alliance, LLC, under Contract No. DE-AC02-07CH11359 with the United States Department of Energy.

[1] G. Aad *et al.* (ATLAS Collaboration), Phys. Lett. B **716**, 1 (2012).
[2] S. Chatrchyan *et al.* (CMS Collaboration), Phys. Lett. B **716**, 30 (2012).
[3] R. Cahn, M. S. Chanowitz, and N. Fleishon, Phys. Lett. **82B**, 113 (1979).
[4] L. Bergstrom and G. Hulth, Nucl. Phys. **B259**, 137 (1985).
[5] C.-W. Chiang and K. Yagyu, Phys. Rev. D **87**, 033003 (2013).
[6] C.-S. Chen, C.-Q. Geng, D. Huang, and L.-H. Tsai, arXiv:1301.4694.
[7] J. Cao, L. Wu, P. Wu, and J. M. Yang, arXiv:1301.4641.
[8] S. Chatrchyan *et al.* (CMS Collaboration), Report No. bCMS-PAS-HIG-12-049, 2012.
[9] J. S. Gainer, W.-Y. Keung, I. Low, and P. Schwaller, Phys. Rev. D **86**, 033010 (2012).
[10] S. Choi, M. Muhlleitner, and P. Zerwas, Phys. Lett. B **718**, 1031 (2013).
[11] B. Coleppa, K. Kumar, and H. E. Logan, Phys. Rev. D **86**, 075022 (2012).
[12] K. Kondo, J. Phys. Soc. Jpn. **60**, 836 (1991).
[13] R. Dalitz and G. R. Goldstein, Phys. Rev. D **45**, 1531 (1992).
[14] J. Alwall, A. Freitas, and O. Mattelaer, Phys. Rev. D **83**, 074010 (2011).
[15] P. Artoisenet, V. Lemaitre, F. Maltoni, and O. Mattelaer, J. High Energy Phys. 12 (2010) 068.
[16] I. Volobouev, arXiv:1101.2259.
[17] V. Abazov *et al.* (D0 Collaboration), Nature (London) **429**, 638 (2004).
[18] V. Abazov *et al.* (D0 Collaboration), Phys. Rev. D **74**, 092005 (2006).
[19] A. Abulencia *et al.* (CDF-Run II Collaboration), Phys. Rev. D **75**, 031105 (2007).
[20] A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. Lett. **99**, 182002 (2007).
[21] J. S. Gainer, K. Kumar, I. Low, and R. Vega-Morales, J. High Energy Phys. 11 (2011) 027.
[22] J. R. Andersen, C. Englert, and M. Spannowsky, Phys. Rev. D **87**, 015019 (2013).
[23] A. Freitas and J. Gainer, arXiv:1212.3598.
[24] S. Chatrchyan *et al.* (CMS Collaboration), Phys. Rev. Lett. **110**, 081803 (2013).

[25] P. Avery, D. Bourilkov, M. Chen, T. Cheng, A. Drozdetskiy *et al.*, arXiv:1210.0896.
[26] J. M. Campbell, W. T. Giele, and C. Williams, J. High Energy Phys. 11 (2012) 043.
[27] J. M. Campbell and R. K. Ellis, Phys. Rev. D **60**, 113006 (1999).
[28] J. M. Campbell, R. K. Ellis, and C. Williams, J. High Energy Phys. 07 (2011) 018.
[29] J. F. Gunion, H. E. Haber, G. L. Kane, and S. Dawson, Front. Phys. **80**, 1 (2000).
[30] J. F. Gunion, H. E. Haber, G. L. Kane, and S. Dawson, arXiv:hep-ph/9302272.
[31] A. Djouadi, V. Driesen, W. Hollik, and A. Kraft, Eur. Phys. J. C **1**, 163 (1998).
[32] S. Catani and M. Seymour, Nucl. Phys. **B485**, 291 (1997).
[33] W. Giele, E. Glover, and D. A. Kosower, Nucl. Phys. **B403**, 633 (1993).
[34] W. T. Giele, G. C. Stavenga, and J.-C. Winter, arXiv:1106.5045.
[35] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter, J. High Energy Phys. 02 (2009) 007.
[36] S. Catani, F. Krauss, R. Kuhn, and B. Webber, J. High Energy Phys. 11 (2001) 063.
[37] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, and C.-P. Yuan, Phys. Rev. D **82**, 074024 (2010).
[38] S. Chatrchyan *et al.* (CMS Collaboration), Report No. CMS-TDR-008-1 (2006).
[39] C. Anastasiou, S. Buehler, C. Duhr, and F. Herzog, J. High Energy Phys. 11 (2012) 062.