Discriminating Higgs production mechanisms using jet energy profiles

Vikram Rentala,1,∗ Natascia Vignaroli,1† Hsiang-nan Li,2,3,4‡ Zhao Li,5§ and C.-P. Yuan1¶

1Department of Physics & Astronomy, Michigan State University, E. Lansing, MI 48824, USA
2Institute of Physics, Academia Sinica, Taipei, Taiwan 115, Republic of China
3Department of Physics, National Cheng-Kung university, Tainan, Taiwan 701, Republic of China
4Department of Physics, National Tsing-Hua university, Hsin-Chu, Taiwan 300, Republic of China
5Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, China

Abstract

We present a new tool for precision measurements of the Higgs boson production mechanisms at the LHC. We study events with a Higgs boson produced with two forward jets. Even with fairly stringent cuts, one expects a significant contamination of gluon fusion (GF) in addition to vector-boson fusion (VBF) in the event sample. By measuring the jet energy profile of the most central jet, we find that SM production can be distinguished from either pure VBF or pure GF at the 5σ level with 100 fb−1 of luminosity at the 14 TeV LHC. Moreover, this discrimination technique can be used to validate or rule out new physics models that predict similar observable branching fractions as the 125 GeV SM Higgs but have different production mechanisms.

PACS numbers: 14.80.Bn; 14.80.Ec; 12.38.Bx; 12.38.Qk

I. INTRODUCTION

The discovery of a 125 GeV resonance by the ATLAS and CMS collaborations [1, 2] has raised many interesting questions about the nature of electroweak symmetry breaking and the Higgs mechanism. While we are still in the relatively early stages of collecting data about this particle, its properties appear to be consistent with that of a Standard Model (SM) Higgs boson. Given the apparent lack of evidence of weak-scale supersymmetry or compositeness, if this particle’s properties continue to agree with those of a SM Higgs boson after the accumulation of even more data, then the problems of naturalness and hierarchy come to the forefront.

It is therefore crucial to measure every aspect of the purported Higgs boson in the coming years at the LHC and possibly at future colliders. Currently, most techniques have focussed on a global fit of the decay modes of the Higgs boson in order to determine its coupling to SM particles [3–5]. In all studies of the Higgs decay modes, the experimental measurement is strictly speaking a product of the production cross-section times the decay branching ratio to a particular mode. This limits the discriminating power of using only the branching ratios to compare consistency with the SM Higgs. Moreover, decays to final states involving gluons (about 80% of the time) or light quarks cannot directly be tested due to the large QCD background.
At the LHC, the SM Higgs boson, with a mass around 125 GeV, has two dominant production mechanisms. The first is gluon fusion (GF) and the second is vector boson fusion (VBF). Thus, a verification of the Higgs production modes allows us to probe the Higgs coupling to gluons and weak bosons independently of the decay branching ratios.

Currently, attempts to separate the GF contribution from the VBF contribution to Higgs production rely on imposing kinematic cuts on the inclusive process \( pp \rightarrow H + 2 \) jets with \( H \rightarrow \gamma \gamma \). At least two hard jets separated by a large rapidity gap are required. However, in spite of these cuts, there is usually a sizeable contamination (\( \sim 20\% \)) from gluon fusion and also an \( O(1) \) background contribution when looking for the Higgs decaying to two photons. In order to determine the gluon and the weak bosons couplings to the Higgs, one needs to disentangle the GF contribution from this sample.

We make the simple observation that the jets associated with VBF are initiated by quarks at the parton level, whereas the jets associated with GF are dominantly gluon-like. Thus, if we can measure the ratio of gluon to quark jets in the data (by considering the more central jet of the two leading \( p_T \) jets, say), we can effectively measure the VBF production and GF cross-sections independently.

The main goal of this work is to demonstrate the effectiveness of separating the GF from the VBF contribution of SM Higgs production by examining the average energy profile of the more central jet. We construct a discrimination variable \( f_V \), the fraction of VBF produced Higgs bosons in a given sample, and explain how to measure it. One could in principle look at any sample of \( H + 2 \) jets (not just in the above kinematic regime) and measure \( f_V \) to check consistency with the SM prediction. This variable can be used not just to verify the Standard Model but to discriminate models of new physics that predict different Higgs production mechanisms. We remark that although the discriminating variable \( f_V \) is treated independently in this paper, one should include it in a global analysis to get even stronger constraints on Higgs couplings to gluons and vector bosons.

This paper is organized as follows: In section II we describe the current discriminating power of Higgs production mechanisms at the LHC using conventional kinematic variables. In section III we discuss the use of jet energy profiles to discriminate quark from gluon jets. In section IV we describe the procedure to measure these jet energy profiles in Higgs production. In section V we describe how to construct the discriminating variable \( f_V \). We simulate trials of LHC data to estimate the errors on the jet energy profiles and on \( f_V \). The results of this numerical analysis are shown in section VI. The numerical results (with error bars) enable us to estimate the luminosity needed to discriminate SM Higgs production from pure VBF production or pure GF production at the 5\( \sigma \) level or more using 100 fb\(^{-1}\) of luminosity at the 14 TeV LHC. In section VII we comment on the role of background on our analysis. While a full background study is beyond the scope of this work, we make some estimates of the effect of error from background on the discriminating power of our technique. Finally, we conclude in section VIII and make some suggestions for future applications of jet energy profiles to new physics searches.

### II. KINEMATIC SEPARATION

We consider the process \( pp \rightarrow H + 2 \) jets with \( H \rightarrow \gamma \gamma \). Based on the CMS analysis with di-jet tag in [2], we will consider the tight cut selection, in which the two photons satisfy the following requirements on their transverse momenta \( (p_T) \) and pseudorapidity \( (\eta) \):

\[
p_T^{\gamma_1} > m_{\gamma\gamma}/2, \quad p_T^{\gamma_2} > m_{\gamma\gamma}/4, \quad |\eta_{\gamma_i}| < 2.5,
\]  

(1)
where \( m_{\gamma\gamma} \) denotes the invariant mass of the two photons. The two jets satisfy:

\[
p_T^{j_1} > 30 \text{ GeV}, \quad p_T^{j_2} > 30 \text{ GeV}, \quad |\eta_j| < 4.7, \quad \Delta \eta_{jj} > 3.5,
\]

(2)

where \( \Delta \eta_{jj} \) is the rapidity separation between the two jets. The following requirement on the invariant mass of two leading jets is also imposed:

\[
M_{jj} > 500 \text{ GeV}.
\]

(3)

For our analysis, we will also identify a second category of cuts, where we lower the cut on \( M_{jj} \) to:

\[
M_{jj} > 250 \text{ GeV}.
\]

(4)

The following conditions are applied to further reduce the background and enhance selection of VBF-like events: the difference between the average pseudorapidity of the two jets and the average pseudorapidity of the diphoton system is required to be less than 2.5, and the difference in azimuthal angle between the diphoton system and the di-jet system is required to be greater than 2.6 radians [2].

As mentioned in the introduction, in spite of these cuts, there is usually a sizeable contamination \( \sim O(20\%) \) from gluon fusion and also an \( \sim O(100\%) \) background contribution when looking for the Higgs decaying to two photons.

As a first attempt to separate the GF and VBF contributions in this regime, one could try to use kinematic discriminants. Fig. [1] shows the \( p_T \) of the central jet for GF and for VBF. Both graphs have been normalized to have the same area.

We point out two interesting features of the \( p_T \) spectra.

1. Towards the lower \( p_T \) region, VBF jets are expected to typically have \( p_T \sim M_W/2 \), where \( M_W \) is the \( W \)-boson mass. However, for GF, the associated jets have a lower \( p_T \) spectrum.

2. In the high \( p_T \) region the jets from GF have a more slowly decaying tail compared to the jets from VBF. This can be understood from the large invariant mass cut imposed on the two leading jets. The large pseudorapidity separation for jets in VBF ensures a large invariant mass (even for lower \( p_T \) jets), however GF events are required to have harder jets in order to generate a large invariant mass.

Fig. [2] shows a contour plot of the \( p_T \) of the central jet versus the rapidity separation \( \Delta \eta \) of the two jets. We can see the larger pseudorapidity separation for VBF, as alluded to in point 2 above.

Fig. [3] is a contour plot showing the \( p_T \) of the central jet versus the \( p_T \) of the less central jet for GF and VBF. We see that more often than not the leading \( p_T \) jet is the same as the central jet.
We summarize this section by noting that although the kinematic distributions highlight interesting differences between GF and VBF production mechanisms, the distributions look quite similar once tight kinematic cuts are applied. This shows the difficulty of trying to further separate GF and VBF by simply strengthening the kinematic cuts. New discriminating tools, such as the study of jet shapes described in this paper can provide a more effective way to achieve this separation.

III. A NEW DISCRIMINANT: JET ENERGY PROFILES

In this paper we will use jet energy profiles (JEPs) to statistically distinguish quark jets from gluon jets. For a jet of size $R$, the

\[ \psi(r) = \frac{\sum_{r'<R} |p_T(r')|}{\sum_{r'<R} p_T(r')} \]  

These JEPs were measured at the Tevatron CDF experiment [6]. More recently, JEPs have been measured at ATLAS [7] and at CMS [8]. It was found that the experimental data can be described by a carefully tuned event generator, PYTHIA Tune A [9, 10].

Gluon initiated jets are expected to spread more due to more radiation and thus have a slowly rising JEP. Quark initiated jets on the other hand radiate less, and thus accumulate a larger fraction of their energy for fairly small $r$ and have a quickly rising JEP. Both profiles approach unity as $r \to R$ because of the normalization.
One can calculate these profiles in perturbative QCD (pQCD) for quark or gluon initiated jets. It has been found that the next-to-leading order (NLO) profiles overshoot the data \[11\] at small \(r\), where the logarithmic corrections \(\alpha_s \log(R/r)\) are important and need to be resummed. The invalidity of fixed-order analysis of jet observables motivates the recent development of the resummation formalism: an improved next-to-leading-logarithm (NLL) resummation calculation of the JEP was performed in \[11, 12\], which organizes terms of the form \(\alpha_s^n (\log(R/r))^{2n}\) and \(\alpha_s^n (\log(R/r))^{2n-1}\) to all orders in \(\alpha_s\). The JEP also changes with the total \(p_T\) of the jet due to QCD scaling violation. Because the running coupling constant decreases with \(p_T\), the resummation effect is expected to be minor at high \(p_T\). This is the reason why the resummation predictions approach the NLO ones as \(p_T\) increases \[11, 12\].

However at (large) \(p_T > 200\) GeV, some deviations were observed at low \(r \sim 0.1\), where the resummation predictions fall a bit below the data. This deviation may be attributed to non-perturbative effects from hadronization and underlying events, or to higher-power effects in the resummation formalism \[12\]. For example, the phase space of the soft gluons that contribute to the anomalous dimension of the resummation was overestimated, and the overestimate, being proportional to \(r\) \[12\], was regarded as a power correction to the energy profile. The energy profile is normalized to unity at \(r = R\), so the overestimate actually causes suppression of the distribution at small \(r\), explaining the slight drop of the resummation predictions in comparison with the data. The overestimate is more pronounced at larger \(p_T\) due to the narrowness of the jet, explaining why the deviation becomes more obvious at high \(p_T\).

These uncertainties should be kept in mind when comparing our predictions to the data.

As discussed in \[11, 12\], the factorization of the soft gluons that contribute to the anomalous dimension of the resummation for an energetic jet can be regarded equivalently as associating these soft gluons with the clustered jet. This prescription is exactly what was adopted in the anti-\(k_T\) algorithm. In order to match theoretical calculations to the data consistently, we will consider anti-\(k_T\) jets \[13\] with a cone size of \(R = 0.7\).

An important point, about comparing the theoretical prediction to the experimental results, is that the theory calculation involves some (arbitrary) scale parameters which are introduced into pQCD calculations to estimate the effect of the yet-to-be calculated sub-leading logarithmic contributions. Hence, the variation in the theory prediction for different values of the scale parameters can be taken as the theoretical error in our calculation. It is expected that the experimental errors at the LHC will be smaller than this theoretical uncertainty.

Thus, before applying our theory prediction to the energy profile of the jets produced in association with the Higgs boson observed at the LHC, we could test our prediction with precision experimental data from a known process, such as \(Z + 2\) jets events, in which the kinematics of the observed jets could be chosen to be similar to those in the \(H + 2\) jets events. In other words, we could use the comparison to the \(Z + 2\) jets data to calibrate our prediction for the JEP of the jets associated with the Higgs boson produced at the LHC.

### IV. ANALYSIS PROCEDURE

We now explain how to obtain a numerical prediction for the JEPs for the central jet in \(H + 2\) jets events with \(H \rightarrow \gamma\gamma\) in the Standard Model. We also consider two hypothesis models for comparison. The first is pure VBF production and the second is pure GF production. For these test scenarios, we
will rescale the total cross-sections to agree with the SM Higgs boson production rates after imposing the relevant kinematic cuts.

First, we simulate $H + 2$ jets events using MadGraph v5 [14]. The results of our leading-order (LO) simulations for the 8 TeV LHC are compared to the CMS results in [2] for the tight selection cuts with $M_{jj} > 500$ GeV, introduced in Sec. II. In order to incorporate higher-order corrections to the cross-section, we use a "$k$-factor" correction, that rescales the simulated MadGraph cross-section to agree with the CMS results [2].

We will use these same $k$-factors, quoted in Table I, for the simulation at 14 TeV. Table II shows the expected cross-sections at the 14 TeV LHC, using tight cuts with $M_{jj} > 500$ GeV and with $M_{jj} > 250$ GeV.

### Table I: CMS cross-sections at the 8 TeV LHC

| $M_{jj} > 500$ GeV | $M_{jj} > 250$ GeV |
|-------------------|-------------------|
| **8 TeV**         | **14 TeV**        |
| CMS               | MG $\times K_{CMS}^f$ |
| GF 77%            | GF 68%            |
| VBF 23%           | VBF 38%           |
| 0.38 fb           | 1.4 fb            |
| 0.11 fb           | 0.88 fb           |

$K_{CMS}^f$ is the correction factor needed to rescale the MadGraph cross-sections to agree with the CMS data.

### Table II: SM expected cross-sections at the 14 TeV LHC

| $M_{jj} > 500$ GeV | $M_{jj} > 250$ GeV |
|-------------------|-------------------|
| **14 TeV**        | **14 TeV**        |
| CMS               | CMS               |
| GF 32%            | GF 38%            |
| VBF 68%           | VBF 62%           |
| 0.57 fb           | 1.2 fb            |
| 1.2 fb            | 1.4 fb            |

A. Calculating the JEPs

Now, we are ready to calculate the theoretical predictions from pQCD for the JEP of the central jet using the formulae derived in [11, 12]. We take jet four-momenta from MadGraph and convolve with the resummed jet functions to arrive at the numerical JEPs.

The JEP of the most central jet in the $H + 2$ jets events passing the cuts with $M_{jj} > 500$ GeV is shown in Fig. 4 (upper panel) for pure VBF production, pure GF production and for the SM. The lower panel of Fig. 4 shows, for comparison, these same profiles when we consider the leading $p_T$ jet, instead of the most central jet. We see that the discriminating power is somewhat reduced in this case. This retroactively justifies our use of the central jet for discrimination.

We also observe a larger separation between the VBF, GF and SM profiles when the cut on $M_{jj}$ is lowered to 250 GeV (Fig. 5). The reason for this effect is an increase of the fraction ($f_g$) of events where the most central jet is a gluon in pure GF (and in the SM) when the $M_{jj}$ cut is lowered (Fig. 6). However, for a milder $M_{jj}$ cut, we also expect a larger contamination from background; this effect will be discussed in Section VII.

B. Estimating the errors on the JEPs

In order to determine how efficiently these profiles can be distinguished from each other, we need to estimate the statistical uncertainties on the theoretical profiles for a given integrated luminosity at the 14 TeV LHC. These uncertainty sources include the scale uncertainty on the LO cross-section, the jet energy resolution uncertainty and the $k$-factor uncertainty. 

1 Specifically, we use the so-called “Higgs Effective Field Theory” model with a 125 GeV Higgs and with the corresponding LO width.

2 As a sanity check, we also find good agreement with the CMS data using the same $k$-factors but looking in the “loose” cut region identified in [2].
FIG. 4: Energy profile of the most central jet (upper panel) and of the leading $p_T$ jet (lower panel) in the $H + 2$ jets events that have passed the tight cuts with $M_{jj} > 500$ GeV, in the SM and in the hypothetical cases of a Higgs produced via pure GF and via pure VBF.

FIG. 5: Energy profile of the most central jet in the $H + 2$ jets events which satisfy tight cuts with $M_{jj} > 250$ GeV, in the SM and in the hypothetical cases of a Higgs produced via pure GF and via pure VBF.

FIG. 6: Upper Panel: Fraction of $H + 2$ jets events where the most central jet is a gluon, $f_g$, as function of the $M_{jj}$ cut (all the other cuts are the same as in the tight selection), in the SM and in the hypothetical case of a Higgs produced via pure GF. Lower Panel: Cross-sections as function of the $M_{jj}$ cut.

statistical errors are obtained by studying the substructure of the reconstructed more central jet (of the two leading $p_T$ jets) in the full event sample including effects from parton showering and hadronization through Pythia v6.4 [9] with the default tune.

We simulate Higgs $+1,2,3$ jets events with MadGraph, then we pass them to Pythia for showering and hadronization and we apply the MLM prescription [15] for matching. Jets are reconstructed using SpartyJet [17], a wrapper for FastJet [18], using the anti-$k_t$ algorithm with $R = 0.7$. We first apply the selection cuts described in Sec. II to the two

---

3 We have used a cutoff scale $Q_{CUT}=15$ GeV and a $x_{qcut}=10$ GeV scale. We refer the reader to [16] for details on how to use matching inside MadGraph.
leading jets in the final state\textsuperscript{4}.

Next, using our Pythia event sample, we examine the central jets and, for a given sub-cone of size $r$, we calculate the mean of the integrated energy distribution, $\psi(r)$. In Fig. 7 we show the energy profile of the central jet for SM events, obtained from Pythia showering, compared to that from the theoretical calculation. Notice the large difference between the theoretical and the Pythia prediction, which depends on the specific Pythia tune considered.

\textit{Due to this tune-dependence, we will rely on our theoretical calculations to determine the central value of the JEP, however we will use the Pythia results to estimate the errors on the JEPs.}

Using the same Pythia event sample, we can calculate the statistical variation on the sample-mean of the JEP, $\psi(r)$, for a given number of events. We find, as expected, that the variations follow Gaussian distributions and that the errors scale as the square root of the number of events.

\textsuperscript{4} We find that the final cross-sections obtained in this manner are in good agreement (within 5\%) with the parton-level final cross-sections listed in Table II.

Fig. 8 shows our final result for the JEPs in the SM and in the hypothetical cases of a Higgs produced via pure GF and via pure VBF. The mean values of $\psi(r)$ are derived from the theoretical perturbative calculation, the error bars are the $1\sigma$ variations estimated from Pythia simulations by using the method explained above and considering the number of events predicted for the SM with 100 fb$^{-1}$ of integrated luminosity at the 14 TeV LHC.
TABLE III: Fit parameters $a$ and $b$ for the central JEPs of SM, pure VBF and pure GF events using tight cuts with $M_{jj} > 250$ GeV. The parameters are to be used in the analytic approximation, Eq. (6), for the JEPs.

|     | $a$  | $b$  |
|-----|------|------|
| SM  | 5.3  | 0.87 |
| VBF | 6.0  | 0.83 |
| GF  | 4.4  | 1.1  |

V. THE DISCRIMINATING VARIABLE $f_V$

JEPS can be well fitted by a two-parameter function of $(a,b)$,

$$\psi(r) = \frac{1 - be^{-ar}}{1 - be^{-aR}},$$

where $R = 0.7$ is the jet cone size. The values of $a$ and $b$ for SM, VBF and GF events are shown in Table III. We will use these analytical expressions to approximate the numerical JEPs shown in Fig. 8 in what follows.

We define a one parameter series of JEPs that linearly interpolate between the VBF and GF profiles,

$$\psi_{fV}(r) = f_V \psi_{VBF}(r) + (1 - f_V) \psi_{GF}(r).$$

Here, $f_V$ is the parameter and $\psi_{VBF}(r)$ and $\psi_{GF}(r)$ are the pure VBF and pure GF profiles that are determined from the MadGraph simulation convolved with jet functions, as described in Sec. IV.

For any experimentally measured JEP, we can perform a fit to the one-parameter family of curves and find the corresponding value of $f_V$. We see that $f_V$ has a clear physical meaning as the fractional amount of VBF contribution to $H + 2$ jet production. Explicitly, $f_V = 1$ for pure VBF and $f_V = 0$ for pure GF production. The central values of $f_V$ for the SM for different cut choices are shown in Table IV. These values agree with the expected VBF fractions from Table II as they should.

Strictly speaking, the profiles are also functions of the jet $p_T$ as explained in Sec. III. Hence, Eq. (7) should be viewed as being valid only for a narrow $p_T$ band, and the parameter $f_V$ should also be regarded as a $p_T$-dependent parameter. In practice, we can ignore this subtlety for two reasons. Firstly, a factor-of-a-few change in $p_T$ is required to see appreciable changes in the profile. Secondly, by examining the $p_T$ distributions for the events under consideration (Fig. 1), we see that the bulk of the events are picked up in a fairly narrow $p_T$ band and thus we can safely ignore the $p_T$ dependence of Eq. (7) in what follows.

VI. RESULTS

Now using our simulated JEPs with error bars shown for 100 fb$^{-1}$ of integrated luminosity at the 14 TeV LHC (Fig. 8), we can translate the errors on the JEPs into errors on the measurement of the fitted $f_V$. The results are shown in Table IV for a SM sample and pure VBF, pure GF samples. Note that the errors are Gaussian, and scale as the square root of the number of events (or equivalently the integrated luminosity).

From the results of Table IV, we can determine our ability to discriminate SM events from a pure VBF sample or a pure GF sample. For 100 fb$^{-1}$ of integrated luminosity, we calculate the difference between the best-fit $f_V$ for VBF or GF and the best-fit $f_V$ for

|     | $M_{jj} > 500$ GeV | $M_{jj} > 250$ GeV |
|-----|-------------------|-------------------|
| SM  | 0.68 ± 0.05       | 0.62 ± 0.04       |
| VBF | 1.00 ± 0.04       | 1.00 ± 0.03       |
| GF  | 0.00 ± 0.06       | 0.00 ± 0.05       |
TABLE V: Expected $\sigma$–level distinction between SM and pure GF or VBF event samples using 100 fb$^{-1}$ of luminosity at the 14 TeV LHC.

| $M_{jj} > 500$ GeV | $M_{jj} > 250$ GeV |
|-------------------|-------------------|
| GF VBF            | GF VBF            |
| $\sigma$-level    | 8.7   | 9.7   |
|                   | 5.0    | 7.6    |

TABLE VI: Integrated luminosity required to distinguish SM from pure GF or VBF event samples at the 5$\sigma$ level.

| $M_{jj} > 500$ GeV | $M_{jj} > 250$ GeV |
|-------------------|-------------------|
| Lum [fb$^{-1}$]    | 33 100            |
|                   | 27 43             |

VII. EFFECT OF BACKGROUND

So far we have neglected the effect of contamination from background on our results. A precise estimate of this effect needs accurate simulations and/or fits to 14 TeV LHC data and is beyond the scope of this study. Nevertheless, we will estimate the effect of the background on the discriminating power of our technique.

We can safely assume that the background JEP, $\psi_B(r)$, can be reconstructed from the data and that the Higgs signal profile, $\psi_S(r)$, can be obtained from the observed profile, $\psi_{obs}(r)$, by subtracting the background contribution:

$$\psi_S(r) = \psi_{obs}(r) + \frac{B}{S} (\psi_{obs}(r) - \psi_B(r)) , \quad (9)$$

where $B/S$ is the background-to-signal ratio. Note that we are once again neglecting the $p_T$ dependence of the JEPs.

This subtraction procedure introduces a correction to the statistical errors on the signal profiles derived in Sec. IV. We make a conservative assumption that the error on background is the same size as the error on the signal for the same number of events. This introduces a scaling correction to the errors obtained in the previous section, given by a factor:

$$\sqrt{1 + 2 \frac{B}{S}}.$$

In order to get a rough estimate of the $B/S$ ratio at the 14 TeV LHC, we simulate with MadGraph the irreducible $\gamma\gamma jj$ QCD background and we apply the selection cuts of Sec. II. Table VII shows the results of simulations at the 8 TeV LHC, compared to the CMS results \[2\]. Given the good agreement with the 8 TeV LHC data, we use the LO simulation to derive the signal-to-background ratio for the different selection categories at the 14 TeV LHC. Our results are shown in Table VII. We see that the background contamination leads to an increase in the error by about 37% using tight cuts with $M_{jj} > 500$.
8 TeV Background

| M_{jj} > 500 GeV | M_{jj} > 250 GeV |
|------------------|------------------|
| CMS              | 0.25 fb          |
| MG               | 0.23 fb          |

14 TeV Background

| M_{jj} > 500 GeV | M_{jj} > 250 GeV |
|------------------|------------------|
| MG               | 0.78 fb          |
| S/B              | 2.3              |

| M_{jj} > 500 GeV | M_{jj} > 250 GeV |
|------------------|------------------|
| 100 fb^{-1} GF VBF | 6.4          |
| 5σ                 | 6.4          |
| Lum [fb^{-1}] | 61 190 |
| S/B | 1.5 |

TABLE VII: Upper Table: Background cross-sections extracted from the number of background events using 5.3 fb^{-1} of data from CMS (listed in Table 2 of [2]) and the MadGraph (MG) prediction for the irreducible γγjj QCD background after tight selection cuts and after applying an additional cut, 124 < m_{γγ} < 126 GeV. Lower Table: Estimated background cross-sections and signal-background ratios at the 14 TeV LHC.

GeV and an increase of about 52% for the M_{jj} > 250 GeV category.

Table VIII shows the corresponding ability of the 14 TeV LHC to discriminate the GF and VBF hypotheses from the SM, including the effect of background. We see that even with the inclusion of background, it is still better to consider a lower M_{jj} cut of 250 GeV, to increase the discriminating power of our analysis.

VIII. SUMMARY AND CONCLUSIONS

Separation of Higgs production modes is important in order to directly measure the Higgs couplings to gluons and vector bosons. Conventional methods of separating the VBF contribution to Higgs production use cuts on Higgs boson events with two forward jets. However, a sizeable fraction of GF events still remains in this sample. Kinematic discriminators provide only a modest ability to further separate GF from VBF events.

In this paper, we made the observation that the jets associated with VBF are initiated by quarks whereas the jets associated with GF are dominantly initiated by gluons. We then presented a new tool for discriminating Higgs boson production mechanisms based on the analysis of the JEP associated with the central jet in H + 2 jets events. We used the profiles to construct a discriminating variable, f_V, which can be regarded as the fraction of VBF events in a given sample. We constructed two test scenarios, where the Higgs is produced via pure VBF and pure GF, and used these to benchmark our ability to discriminate against the SM.

We found the central value for f_V using a theoretical pQCD calculation of the JEP (Table IV). We estimated the expected errors on the measurement of f_V at the 14 TeV run of the LHC for different luminosities using Pythia simulations, including the effect of background.

The main results of our paper are shown in Table VIII where we show how well the pure GF and pure VBF hypotheses can be separated from the SM by using the measured values of f_V. We find that with a 100 fb^{-1} of luminosity at the 14 TeV LHC, both the pure GF and via pure VBF hypotheses
can be excluded at the 5\(\sigma\) level. Our method should be included in a global analysis to further strengthen its discriminatory power.

The use of JEPs to separate various operators contributing to Higgs production is the novel feature of this work. A similar technique can be applied to probe new physics models. A couple of applications are:

- Separation of \(QQ\chi\chi\) versus \(GG\chi\chi\) contact operator coefficients in dark matter mono-jet searches.
- Distinction between different types of dijet resonances (colorons, Z-primes, etc.).

IX. ACKNOWLEDGMENTS

We would like to thank J. Huston and C. Vermilion for useful discussions about SparxyJet. This work was supported by the U.S. National Science Foundation under Grant Nos. PHY-0855561 and PHY-0854889; the National Science Council of R.O.C. under Grant No. NSC-101-2112-M-001-006-MY3; and the National Natural Science Foundation of China under Grant No. NSFC11245007. V.R. would like to thank the KITP where part of this work was completed with the support of Grant No. NSF PHY11-25915.

[1] G. Aad et al. (ATLAS Collaboration), Phys.Lett. B716, 1 (2012), hep-ex/1207.7214.
[2] S. Chatrchyan et al. (CMS Collaboration), Phys.Lett. B716, 30 (2012), hep-ex/1207.7235.
[3] Tech. Rep. ATLAS-CONF-2013-034, CERN, Geneva (2013).
[4] Tech. Rep. CMS-PAS-HIG-13-005, CERN, Geneva (2013).
[5] See, for example, J. Ellis and T. You (2013), hep-ph/1303.3879, and references therein.
[6] D. Acosta et al. (CDF Collaboration), Phys.Rev. D71, 112002 (2005), hep-ex/0505013.
[7] G. Aad et al. (Atlas Collaboration), Phys.Rev. D83, 052003 (2011), hep-ex/1101.0070.
[8] Tech. Rep. CMS-PAS-QCD-10-013, CERN, 2010. Geneva (2010).
[9] T. Sjostrand, S. Mrenna, and P. Z. Skands, Comput.Phys.Commun. 178, 852 (2008), hep-ph/0710.3820.
[10] PYTHIA-Tune A. Monte Carlo samples are generated using the following tuned parameters in PYTHIA: PARP(67) = 4.0, MSTP(82) = 4, PARP(82) = 2.0, PARP(84) = 0.4, PARP(85) = 0.9, PARP(86) = 0.95, PARP(89) = 1800.0, PARP(90) = 0.25.
[11] H.-n. Li, Z. Li, and C.-P. Yuan, Phys.Rev.Lett. 107, 152001 (2011), hep-ph/1107.4535.
[12] H.-n. Li, Z. Li, and C.-P. Yuan, Phys.Rev. D87, 074025 (2012), hep-ph/1206.1344.
[13] M. Cacciari, G. P. Salam, and G. Soyez, JHEP 0804, 063 (2008), hep-ph/0802.1189.
[14] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, JHEP 1106, 128 (2011), hep-ph/1106.0522.
[15] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, JHEP 0701, 013 (2007), hep-ph/0611129.
[16] J. Alwall, S. Hache, F. Krauss, N. Lavesson, L. Lonnblad, et al., Eur.Phys.J. C53, 473 (2008), hep-ph/0706.2569.
[17] P.-A. Delsart, K. L. Geerlings, J. Huston, B. T. Martin, and C. K. Vermilion (2012), hep-ex/1201.3617.
[18] M. Cacciari, G. P. Salam, and G. Soyez, Eur.Phys.J. C72, 1896 (2012), hep-ph/1111.6097.