

Unravelling $t\bar{t}h$ via the Matrix Element Method

Pierre Artoisenet,¹ Priscila de Aquino,² Fabio Maltoni,³ and Olivier Mattelaer^{3,4}

¹*Nikhef Theory Group, Science Park 105, 1098 XG Amsterdam, Netherlands*

²*Theoretische Natuurkunde and IIHE/ELEM, Vrije Universiteit Brussel,
and International Solvay Institutes, Pleinlaan 2, B-1050 Brussels, Belgium*

³*Centre for Cosmology, Particle Physics and Phenomenology (CP3), Université catholique de Louvain,
Chemin du Cyclotron 2, B-1348 Louvain-la-Neuve, Belgium*

⁴*Department of Physics, University of Illinois at Urbana-Champaign, 1110 West Green Street, Urbana, Illinois 61801, USA*

(Received 30 April 2013; published 27 August 2013)

Associated production of the Higgs boson with a top-antitop pair is a key channel to gather further information on the nature of the newly discovered boson at the LHC. Experimentally, however, its observation is very challenging due to the combination of small rates, difficult multijet final states, and overwhelming backgrounds. In the standard model, the largest number of events is expected when $h \rightarrow b\bar{b}$, giving rise to a $W^+W^-b\bar{b}b\bar{b}$ signature, deluged in $t\bar{t} + \text{jets}$. A promising strategy to improve the sensitivity is to maximally exploit the theoretical information on the signal and background processes by means of the matrix element method. We show how, despite the complexity of the final state, the method can be efficiently applied to discriminate the signal against combinatorial and $t\bar{t} + \text{jets}$ backgrounds. Remarkably, we find that a moderate integrated luminosity in the next LHC run will be enough to make the signature involving both W 's decaying leptonically as sensitive as the single-lepton one.

DOI: [10.1103/PhysRevLett.111.091802](https://doi.org/10.1103/PhysRevLett.111.091802)

PACS numbers: 14.80.Bn, 12.38.Bx, 13.85.Ni

Introduction.—Evidence for the recently discovered new heavy boson to be the long-sought-for Higgs particle of the standard model (SM) is already quite compelling [1,2]. Rates and distributions are compatible with the predictions of a scalar particle coupling to other SM particles with a strength proportional to their mass. The current sensitivities and accuracies of the golden production-decay modes, however, are not sufficient to draw a final answer on the strength and the structure of the couplings without additional hypotheses. Other channels need to be investigated.

Prominent among the yet-to-be-explored production modes is the $t\bar{t}h$ associated production. The main interest of this channel stems from the fact that the rate is manifestly proportional to the square of the SM Yukawa coupling to the top quarks. In addition, more differential observables could bring information on the coupling structure [3] and on the Higgs parity [4]. This channel, however, is notoriously difficult for several reasons. The first is that production rates at hadron colliders are quite small due to the need of a large c.m.s. collision energy for the initial partons, strongly suppressed by parton distribution functions. Next-to-leading-order (NLO) calculations [4–7] predict an SM rate of 0.137 and 0.632 pb with $O(10\%)$ uncertainty at the LHC for $\sqrt{s} = 8$ and 14 TeV, respectively. Current searches mainly focus on the dominant decay mode $h \rightarrow b\bar{b}$ and therefore on a $W^+W^-b\bar{b}b\bar{b}$ final state, other decays, such as $h \rightarrow W^+W^-$ [8], $h \rightarrow \tau^+\tau^-$ [9], and eventually, $h \rightarrow \gamma\gamma$ [10], being much rarer demand larger integrated luminosities. The second reason is that the $W^+W^-b\bar{b}b\bar{b}$ signature is affected by two different types of challenging backgrounds: on the one hand,

$t\bar{t} + \text{light- or heavy-flavor jets}$ because of the enormous rates [11–14], and on the other hand, the intrinsic combinatorial background that stems from the difficulty of correctly identifying out of four b jets the two from the Higgs decay. Finally, the complexity of the final state makes its kinematic reconstruction not straightforward mainly due to finite jet energy resolution, missing energy, and the ubiquity of extra QCD radiation.

Because of the above intrinsic difficulties, the prospects of first using this channel for discovery or just for observation have been constantly deteriorating as more accurate predictions and simulations were available to the LHC community. More recently, the attention on this channel was revived by Plehn *et al.* [15], who suggested that while drastically lowering the rates, boosted tops and Higgs boson in the final state would make the combinatorial background much less severe, improving the significance S/\sqrt{B} of the SM Higgs observation at large enough luminosities. Nevertheless, strategies based on advanced multivariate analyses are currently employed [16,17] to observe this channel at low p_T , i.e., where the bulk of the cross section resides.

In this work, we argue that the sensitivity to $t\bar{t}h$ can be further enhanced at low p_T by means of the matrix element reweighting method, improving the prospects for observation of this channel at the LHC in the coming years. The matrix element method is able to efficiently reduce the combinatorial problem for the single-lepton final states and even more for the dilepton final state, bringing the two to a comparable level of sensitivity already for moderate integrated luminosities.

The matrix element method.—The matrix element reweighting method (MEM), originally introduced in Ref. [18], assigns matrix-element-based probabilities to competing hypotheses, e.g., signal versus signal + background, given a sample of experimental events. The method, implemented using tree-level matrix elements, has been successfully applied to a number of key results in collider physics: from the most precise top mass determination [19,20], to the single top observation [21,22] at the Tevatron, to the Higgs boson discovery [23] and characterization in the $H \rightarrow ZZ \rightarrow 4\ell$ channel [24,25]. Efforts to include next-to-leading QCD corrections, at least for simple final states, have started [26,27].

The probability density evaluated at an experimental event \mathbf{x} given a set of hypotheses α is called a weight and is defined as

$$P(\mathbf{x}|\alpha) = \frac{1}{\sigma_\alpha} \int d\Phi(\mathbf{y}) |M_\alpha|^2(\mathbf{y}) W(\mathbf{x}, \mathbf{y}), \quad (1)$$

where $|M_\alpha|^2(\mathbf{y})$ is the leading-order matrix element (giving the parton-level probability), $d\Phi(\mathbf{y})$ is the phase-space measure [including the parton distribution functions $f_1(q_1)dq_1$ and $f_2(q_2)dq_2$], and $W(\mathbf{x}, \mathbf{y})$ is the transfer function which describes the evolution of (the final state) parton-level configuration in \mathbf{y} into a reconstructed event \mathbf{x} in the detector. The definition of the weight in Eq. (1) implicitly includes a sum over all possible parton-jet assignments, each of them weighted by the matrix element convoluted with the transfer function. As a result, the MEM suppresses the combinatorial background by using all information encoded in the matrix elements.

As evident from the definition in Eq. (1), the calculation of each weight involves a nontrivial multidimensional integration of complicated functions over the phase space. The problem of computing the weights for arbitrary models and processes was tackled in Ref. [28] by implementing a general algorithm in a specifically designed code named MADWEIGHT. A series of key technical improvements has been recently achieved [29] that allows the method to be applied to challenging final states such as those from $t\bar{t}h$ production. We stress that very fact of automatically, reliably, and quickly calculating weights in these cases is a significant technical result on its own that provides key evidence on the generality and flexibility of the MADWEIGHT approach.

One of the main limitations of the method is that the matrix elements are considered at the leading order only, and therefore extra QCD radiation effects must be handled in some effective way. In our study, we are inclusive on extra transverse radiation and consider only the hardest jets to be matched with the corresponding partons in the matrix element. The transverse momentum of these partons (including isolated leptons) is assumed to be balanced against the transverse momentum of extra radiation in the event [30].

Technical aspects.—Parton-level events for signal and backgrounds are generated with MADGRAPH 5 [31] and passed to PYTHIA 6 [32] for showering and hadronization, employing the k_T merging procedure implemented in MADGRAPH [33]. The underlying event is simulated with PYTHIA, whereas pileup effects are not modeled in our simulation. Detector response simulation is performed using DELPHES [34], with the input parameters tuned to the values associated with the CMS detector.

Only the dominant background $t\bar{t} + \text{jets}$ is taken into account and is modeled by generating parton-level $t\bar{t}$ processes with up to two extra partons in the five-flavor scheme. For the signal, the parton-level processes $t\bar{t}h$ and $t\bar{t}h + 1$ parton are considered. Inclusive samples for the signal and the background have been normalized to the total cross section at NLO from Refs. [4,35], respectively. Spin correlation effects in the decays of the tops, which for signal shapes are more important than NLO QCD corrections [36], have been retained.

The event selection procedure is modeled after that adopted by the CMS Collaboration for the measurement of the $t\bar{t}$ cross section in the dilepton channel [37]. Single-lepton (dilepton) events are required to have at least one (one pair of opposite-charge) lepton(s). Only isolated electrons or muons are lepton candidates in our analysis. They are required to have a transverse momentum $p_T > 20$ GeV and a pseudorapidity $|\eta| < 2.4$.

Jets are reconstructed via the anti- k_T algorithm [38] (with a cone radius $R = 0.5$) as implemented in FASTJET [39] and applied on the calorimeter cells fired by the generated stable or quasistable particles. Jet candidates are required to have $p_T > 30$ GeV and $|\eta| < 2.5$, and not to overlap with any selected leptons. b jets are identified with an efficiency and a mistag rate typical of those obtained with the combined secondary vertex algorithm [40], i.e., a b -tagging efficiency $\epsilon_b = 0.7$, a mistag rate for charm quarks $\epsilon_c = 0.2$, and a mistag rate for light partons $\epsilon_j = 0.015$. Transverse momentum and rapidity dependences in the b -tagging rates are neglected. At least four b jets are required.

The cross sections for signal and backgrounds together with the final efficiencies of the adopted minimal selection procedure are collected in Table I.

As in Ref. [41], only transfer functions for the jet energies are taken with a finite resolution and are parametrized through a double-Gaussian shape. Only parton-jet

TABLE I. Total cross sections at the LHC 14 TeV and corresponding efficiency factors of the applied selection.

Process	Inclusive σ	Efficiency	σ^{rec}
$t\bar{t}h$, single lepton	111 fb	0.0485	5.37 fb
$t\bar{t}h$, dilepton	17.7 fb	0.0359	0.634 fb
$t\bar{t} + \text{jets}$, single lepton	256 pb	0.463×10^{-3}	119 fb
$t\bar{t} + \text{jets}$, dilepton	40.9 pb	0.168×10^{-3}	6.89 fb

assignments consistent with the b -tagging information are considered in the calculation of the weights.

Results.—For a generic event i with kinematics \mathbf{x}_i , the MEM-based observable D_i is defined as follows:

$$D_i = \frac{P(\mathbf{x}_i|S)}{P(\mathbf{x}_i|S) + P(\mathbf{x}_i|B)}. \quad (2)$$

Expected (normalized) distributions of signal and background events with respect to this observable are named D_S and D_B , and are shown in Fig. 1 (left). The plots show that for the same number of signal events, the MEM-based observable delivers a higher discriminating power in the case of the dilepton channel. This is manifest in the right-hand plot of the same figure where the ϵ_s versus ϵ_b efficiencies resulting from a cut on the observable $D > D_{\min}$ are shown. This may seem surprising at first sight, given that the dilepton channel is characterized by two missing particles in the final state, against only one in the single-lepton channel. However, the dilepton channel is much cleaner, with only b jets required in the final state, a lower probability of erroneously including extra QCD radiation, and eventually, a more manageable combinatorial background.

In order to assess the expected significance at the LHC $\sqrt{s} = 14$ TeV for a given luminosity \mathcal{L} , we consider a large number of pseudoexperiments, each with a number of events set to $N = \sigma_{\text{bg}}^{\text{rec}} \mathcal{L}$ (with $\sigma_{\text{bg}}^{\text{rec}}$ the reconstructed cross section, see Table I, last column). In the B -only hypothesis, the numbers of signal and background events are set to $s = 0$ and $b = N$. In the $S + B$ hypothesis, s and b are generated under the constraint $s + b = N$, according to the product of Poisson distributions with mean values $Ns_0/(s_0 + b_0)$ and $Nb_0/(s_0 + b_0)$, respectively. Here, s_0 and b_0 are the expected numbers of reconstructed events after rescaling the signal cross section by a parameter μ , i.e., $b_0 = \sigma_{\text{bg}}^{\text{rec}} \mathcal{L}$ and $s_0 = \mu \sigma_{\text{sig}}^{\text{rec}} \mathcal{L}$. For each event, the corresponding D_i value is generated according to the probability law D_S (in the case of a signal event) or D_B (in the case of a background event) shown in Fig. 1. This procedure is used to generate 10^4 pseudoexperiments under each hypothesis (B only or $S + B$) at a given luminosity \mathcal{L} .

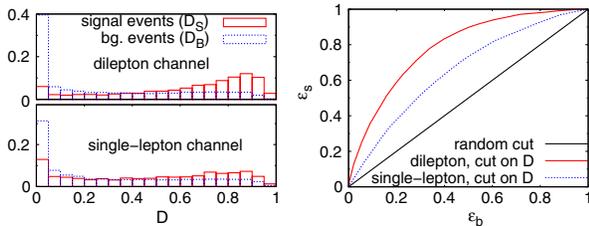


FIG. 1 (color online). Left: Normalized distributions of events with respect to the MEM-based observable D for the dilepton (top) and single-lepton (bottom) channels. Right: Efficiency of selecting signal versus background using a $D > D_{\min}$ cut.

For each pseudoexperiment, the likelihood ratio L^R is calculated as follows:

$$L^R = \prod_i^N \frac{r_0 P(\mathbf{x}_i|S) + (1 - r_0) P(\mathbf{x}_i|B)}{P(\mathbf{x}_i|B)} \\ = \prod_i^N \frac{r_0 D_i + (1 - r_0)(1 - D_i)}{(1 - D_i)}, \quad (3)$$

with $r_0 = s_0/(s_0 + b_0)$. The resulting B -only and $S + B$ distributions of pseudoexperiments with respect to $\ln(L^R)$ are shown in Fig. 2 (left) in the case of the dilepton channel, with $\mathcal{L} = 32 \text{ fb}^{-1}$ and $\mu = 1$. The two distributions are shifted towards positive values of $\ln(L^R)$, which indicates that the MEM weights do not exactly describe the phase-space distributions of background and signal events. This bias originates from the approximations inherent to the calculation of the weights, e.g., the assumed parametrization of the transfer function and the effective treatment of beyond-leading-order QCD radiations.

By smearing the value of b_0 according to a log-normal distribution (mean = b_0 , std = $0.2b_0$) before generating s and b in each pseudoexperiment, we also verified that systematic uncertainties on the background normalization have a negligible impact on the distributions of pseudoexperiments with respect to $\ln(L^R)$. On the other hand, already a 20% uncertainty on b_0 hampers a counting analysis based on the number of events to be available at the LHC, unless $s/b \gg 0.2$.

We repeat this exercise with different values of μ until the median of the B -only distribution cuts 5% of the left-hand tail of the $S + B$ distribution. Such a value of μ provides us with the estimate $\mu \times \sigma(t\bar{t}h)$ of the expected upper bound on the signal cross section at 95% C.L. in the absence of signal. Figure 2 (right) shows our estimate of the parameter μ as a function of the luminosity \mathcal{L} , separately for the dilepton and single-lepton channels. We observe that the sensitivity achieved in the dilepton channel is slightly better than the one in the single-lepton channel at large luminosities.

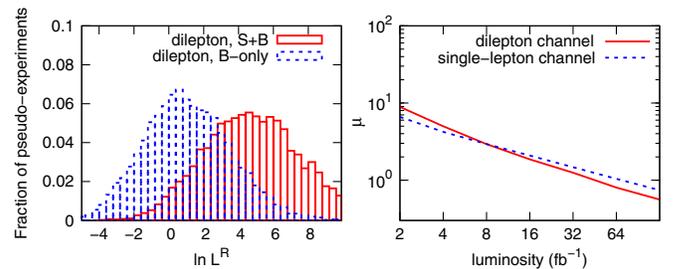


FIG. 2 (color online). Left: Normalized log-likelihood profiles in the case of the dilepton channel, assuming a luminosity of 32 fb^{-1} at 14 TeV and setting $\mu = 1$ (SM cross section). Right: Expected upper bound μ on the $t\bar{t}h$ cross section (in units of SM cross section) at 95% C.L.

Conclusions.—In this work, we have applied the matrix element reweighting method to the observation of Higgs production in association with the $t\bar{t}$ pair, in particular, to final state signatures involving the decay of the Higgs boson to bottom quarks. First, we have verified that the general algorithm implemented in MADWEIGHT provides the possibility of automatically, reliably, and quickly calculating weights for final states as complex as those featured in $t\bar{t}h$. This technical result, by itself, is an important one as it opens the door to applications of the MEM to a much wider set of processes and analyses than what has been done so far. Second, we have applied the method to $t\bar{t}h$ with both one- and two-lepton final states. Our simple analysis, which is based on four b -tagged final states only, indicates that in the absence of signal, rejection at 95% C.L. of SM $t\bar{t}h$ events is possible in each channel separately with a luminosity of around 40 fb^{-1} at 14 TeV. We have also found that the dilepton final state, although penalized by a smaller number of expected events and possibly more difficult to reconstruct due to the presence of two neutrinos in the final state, becomes competitive with the single-lepton channel already after a moderate integrated luminosity. Assuming SM production rates for signal and background processes, our MEM-based analysis leads to an expected 3σ (respectively, 5σ) observation in the dilepton channel with a luminosity of 120 fb^{-1} (respectively, 420 fb^{-1}) at 14 TeV. We reckon that these results, while based on Monte Carlo simulations, are rather robust and encourage more refined investigations. Only fully fledged experimental analyses can assess the final gains and correctly include the systematic uncertainties that have been mainly neglected here. For example, pileup effects could impact more the signature with the largest number of jets in the final state and possibly degrade its expected significance. Conversely, refinements in the matrix element weights (e.g., by increasing the resolution on jet energies) as well as in the event selection procedure may lead to significant improvements. For instance, relaxing the number of requested b tags to three in the dilepton final state would bring a significant increase in the statistics, yet not of the combinatorial background, leading to a further relative gain with respect to the single-lepton final state.

In conclusion, the search for SM $t\bar{t}h$ and in particular the dilepton final state is a perfect illustration of the power of the matrix element method in providing additional leverage in difficult analyses. Further investigations concerning the possibility of using the matrix element method in $t\bar{t}h$ to access more detailed information on the structure of the couplings of the new boson or in other very challenging production channels, such as thj [42], are foreseen.

We are grateful to J. D’Hondt, K. Cranmer, R. Frederix, and to many CP3 members for very useful information exchange and discussions. P.A. is supported by a Marie Curie Intra-European Fellowship (No. PIEF-GA-2011-299999 PROBE4TeVSCALE). P. de A. is supported in

part by the Belgian Federal Science Policy Office through the Interuniversity Attraction Pole P7/37, in part by the “FWO-Vlaanderen” through Project No. G.0114.10N, and in part by the Strategic Research Program “High Energy Physics” and the Research Council of the Vrije Universiteit Brussel. O.M. is supported by the Belgian American Education Foundation.

-
- [1] CMS Collaboration, CMS Reports No. CMS-HIG-13-003, No. CMS-HIG-13-004, No. CMS-HIG-13-006, and No. CMS-HIG-13-009, 2013.
 - [2] ATLAS Collaboration, ATLAS Reports No. ATLAS-CONF-2013-009, Report No. ATLAS-CONF-2013-010, Report No. ATLAS-CONF-2013-012, and Report No. ATLAS-CONF-2013-013, 2013.
 - [3] C. Degrande, J. Gerard, C. Grojean, F. Maltoni, and G. Servant, *J. High Energy Phys.* **07** (2012) 036.
 - [4] R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, R. Pittau, and P. Torrielli, *Phys. Lett. B* **701**, 427 (2011).
 - [5] W. Beenakker, S. Dittmaier, M. Kramer, B. Plumper, M. Spira, and P. Zerwas, *Phys. Rev. Lett.* **87**, 201805 (2001).
 - [6] S. Dawson, L. H. Orr, L. Reina, and D. Wackerroth, *Phys. Rev. D* **67**, 071503 (2003).
 - [7] M. Garzelli, A. Kardos, C. Papadopoulos, and Z. Trocsanyi, *Europhys. Lett.* **96**, 11001 (2011).
 - [8] F. Maltoni, D.L. Rainwater, and S. Willenbrock, *Phys. Rev. D* **66**, 034022 (2002).
 - [9] A. Belyaev and L. Reina, *J. High Energy Phys.* **08** (2002) 041.
 - [10] C. Buttar, S. Dittmaier, V. Drollinger, S. Frixione, A. Nikitenko *et al.*, [arXiv:hep-ph/0604120](https://arxiv.org/abs/hep-ph/0604120).
 - [11] G. Bevilacqua, M. Czakon, C.G. Papadopoulos, and M. Worek, *Phys. Rev. D* **84**, 114017 (2011).
 - [12] A. Bredenstein, A. Denner, S. Dittmaier, and S. Pozzorini, *Phys. Rev. Lett.* **103**, 012002 (2009).
 - [13] A. Bredenstein, A. Denner, S. Dittmaier, and S. Pozzorini, *J. High Energy Phys.* **03** (2010) 021.
 - [14] A. Kardos and Z. Trocsanyi, [arXiv:1303.6291](https://arxiv.org/abs/1303.6291).
 - [15] T. Plehn, G.P. Salam, and M. Spannowsky, *Phys. Rev. Lett.* **104**, 111801 (2010).
 - [16] ATLAS Collaboration, CERN Report No. ATLAS-CONF-2012-135, 2012.
 - [17] CMS Collaboration, *J. High Energy Phys.* **05** (2013) 145.
 - [18] K. Kondo, *J. Phys. Soc. Jpn.* **57**, 4126 (1988).
 - [19] V. Abazov *et al.* (D0 Collaboration), *Nature (London)* **429**, 638 (2004).
 - [20] A. Abulencia *et al.* (CDF Collaboration), *Phys. Rev. D* **75**, 031105 (2007).
 - [21] T. Aaltonen *et al.* (CDF Collaboration), *Phys. Rev. Lett.* **103**, 092002 (2009).
 - [22] V. Abazov *et al.* (D0 Collaboration), *Phys. Rev. Lett.* **103**, 092001 (2009).
 - [23] CMS Collaboration, *Phys. Lett. B* **716**, 30 (2012).
 - [24] CMS Collaboration, *Phys. Rev. Lett.* **110**, 081803 (2013).
 - [25] ATLAS Collaboration, CERN Report No. ATLAS-CONF-2013-013, 2013.
 - [26] J.M. Campbell, W.T. Giele, and C. Williams, *J. High Energy Phys.* **11** (2012) 043.

- [27] J. M. Campbell, R. K. Ellis, W. T. Giele, and C. Williams, *Phys. Rev. D* **87**, 073005 (2013).
- [28] P. Artoisenet, V. Lemaître, F. Maltoni, and O. Mattelaer, *J. High Energy Phys.* **12** (2010) 068.
- [29] P. Artoisenet and O. Mattelaer, “Automatic Matrix Element Method in MADGRAPH 5” (to be published).
- [30] J. Alwall, A. Freitas, and O. Mattelaer, *Phys. Rev. D* **83**, 074010 (2011).
- [31] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, *J. High Energy Phys.* **06** (2011) 128.
- [32] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *J. High Energy Phys.* **05** (2006) 026.
- [33] J. Alwall, S. de Visscher, and F. Maltoni, *J. High Energy Phys.* **02** (2009) 017.
- [34] S. Ovin, X. Rouby, and V. Lemaître, [arXiv:0903.2225](https://arxiv.org/abs/0903.2225).
- [35] M. Cacciari, M. Czakon, M. Mangano, A. Mitov, and P. Nason, *Phys. Lett. B* **710**, 612 (2012).
- [36] P. Artoisenet, R. Frederix, O. Mattelaer, and R. Rietkerk, *J. High Energy Phys.* **03** (2013) 015.
- [37] CMS Collaboration, *Eur. Phys. J. C* **73**, 2339 (2013).
- [38] M. Cacciari, G. P. Salam, and G. Soyez, *J. High Energy Phys.* **04** (2008) 063.
- [39] M. Cacciari, G. P. Salam, and G. Soyez, *Eur. Phys. J. C* **72**, 1896 (2012).
- [40] CMS Collaboration, *JINST* **8**, P04013 (2013).
- [41] A. Abulencia *et al.* (CDF Collaboration), *Phys. Rev. D* **74**, 032009 (2006).
- [42] M. Farina, C. Grojean, F. Maltoni, E. Salvioni, and A. Thamm, [arXiv:1211.3736](https://arxiv.org/abs/1211.3736).