

Toponium hunter's guide

J. A. Aguilar-Saavedra 

Instituto de Física Teórica, IFT-UAM/CSIC, c/ Nicolás Cabrera 13-15, 28049 Madrid, Spain



(Received 10 August 2024; accepted 5 September 2024; published 23 September 2024)

We address the discovery and characterization of toponium production at the Large Hadron Collider. In the dilepton decay mode, multivariate analyses of spin and color observables could provide evidence that an excess of events present near the $t\bar{t}$ threshold corresponds to a spin-zero color singlet. The semileptonic decay mode may also exhibit an excess near threshold, but is not expected to play any role in the toponium characterization.

DOI: [10.1103/PhysRevD.110.054032](https://doi.org/10.1103/PhysRevD.110.054032)

I. INTRODUCTION

Measurements of quantum entanglement between the spins of top quark pairs produced at the Large Hadron Collider (LHC) have recently been performed by the ATLAS [1] and CMS [2] collaborations, in a kinematical region near the $t\bar{t}$ production threshold. The ATLAS measurement exhibits a quite sizeable discrepancy with respect to next-to-leading order (NLO) predictions of the Standard Model (SM). On the other hand, the CMS collaboration has found that, remarkably, data is very well explained provided the SM NLO prediction is supplemented by the production of toponium, a yet unobserved $t\bar{t}$ bound state of extremely short lifetime. Bound states of $t\bar{t}$ are a prediction of the SM, and are essentially a non-perturbative phenomenon. In hadron collisions, the formation of a $J^P = 0^-$ color-singlet resonance is predicted [3–6], while in e^+e^- collisions a $J^P = 1^-$ resonance is expected near threshold [7–9].

The renewed interest in toponium production motivates a thorough investigation of strategies towards its potential discovery at the LHC. At hadron colliders the presence of a toponium “signal” may be spotted by deviations with respect to the predictions of perturbative QCD near the $t\bar{t}$ threshold. These deviations comprise the following:

- (i) an event excess;
- (ii) differences in observables characterizing the spin of the $t\bar{t}$ pair, which arise because $t\bar{t}$ from toponium decay are produced in a spin-singlet state;
- (iii) differences in observables characterizing the $t\bar{t}$ color connection: for toponium decay the $t\bar{t}$ pair is a color

singlet, while $t\bar{t}$ in the continuum is dominated by the color-octet state.¹

Previous work [10] has addressed the prospects to discover toponium in the dilepton decay channel $t\bar{t} \rightarrow \ell\nu b\ell\nu\bar{b}$, with $\ell = e, \mu$, as an excess of events near threshold. This excess can be enhanced via suitable kinematical cuts on the dilepton invariant mass $m_{\ell\ell}$ and the laboratory-frame azimuthal angle difference $\Delta\phi_{\ell\ell}$. These two variables are related to the $t\bar{t}$ spin correlation: the $t\bar{t}$ spin-singlet state tends to produce closer leptons, which also have smaller invariant mass. However, that simple strategy does not seem sufficient to provide strong evidence for toponium production. It also has the disadvantage that $\Delta\phi$ is quite sensitive to boosts of the $t\bar{t}$ pair in the transverse plane.² Lepton angles defined in the rest frame of the parent top quark are experimentally more challenging but also more robust from the theoretical point of view.

The discovery and characterization of toponium is, admittedly, a formidable task that requires a very accurate theoretical modeling. As a first step, it is useful to investigate the observables that would reveal its properties, thereby allowing us to characterize an excess near threshold as “toponium.” This is the purpose of this work. Our goal is not to give precise sensitivity estimations: these would not be realistic in the absence of systematic uncertainties, which can only be evaluated by the experiments. Instead, our focus is on the strategy that experiments could follow, including a comparative study of the observables that might be used for that characterization.

¹For brevity we will often refer to $t\bar{t}$ production in perturbative QCD as production in the continuum, as opposed to resonant toponium production.

²The $\Delta\phi_{\ell\ell}$ distribution is known to exhibit a mismodeling [11,12] that cannot be attributed to the toponium contribution, since it is present far above threshold. Next-to-next-to-leading order corrections [13] improve the agreement, but still data and predictions exhibit some discrepancies [14].

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

The strategy proposed is, in any case, well motivated by the estimations we provide of the statistical sensitivity to observe deviations.

The main focus of this work is the dilepton decay mode. Provided the systematic uncertainties are under control, significant deviations with respect to the predictions of perturbative QCD could be measured near threshold, in the total number of events, and spin/color observables. For the analysis of the latter two, a multivariate analysis could be very useful, as we will show. The semileptonic decay mode of the top quark pair, which has a larger branching ratio than the dilepton one, may also be useful to spot an event excess near threshold, but not for the characterization of the toponium properties. These points are further elaborated in an Appendix.

II. EVENT GENERATION, SELECTION, AND RECONSTRUCTION

The production of $t\bar{t}$ in the continuum is modeled with the SM process $pp \rightarrow b\bar{b}WW$, which includes $t\bar{t}$ as well as nonresonant diagrams. It is generated with MadGraph [15] at the leading order, using NNPDF 3.1 [16] parton density functions and setting as factorization and renormalization scale half the transverse mass, $Q = 1/2 \sum_i (m_i^2 + p_{Ti}^2)^{1/2}$, with p_T the transverse momentum in the usual notation and i labeling the different particles. We set the top mass to $m_t = 172.5$ GeV. Toponium production is modeled as a pseudoscalar resonance η_t with mass $m_{\eta_t} \simeq 2m_t - 2 = 343$ GeV, width $\Gamma_{\eta_t} \simeq 2\Gamma_t = 3$ GeV [17], and interactions

$$\mathcal{L} = -g_{g\eta_t} G_{\mu\nu}^a \tilde{G}^{\mu\nu a} \eta_t - ig_{t\eta_t} \bar{t} \gamma_5 t \eta_t. \quad (1)$$

The effective $g\eta_t$ interaction stands for the triangle loop diagram with a top quark. For our study this is a good approximation because being a spin-zero particle, the only feature required from the production is the cross section, which is fitted from calculations of $t\bar{t}$ production near threshold with nonrelativistic effects [4–6]. The spatial size of toponium also gives rise to differences in top momentum distributions, which can be implemented via a Green function reweighting [10]. However, these differences are rather small, cf. Figs. 13, 14 from Ref. [6], and not relevant for our analysis.

We generate high-statistics samples with 3×10^7 events for $pp \rightarrow b\bar{b}W^+W^- \rightarrow b\bar{b}\ell^+\nu\ell^-\nu$, with $\ell = e, \mu$, and 3×10^6 events for $pp \rightarrow \eta_t \rightarrow b\bar{b}W^+W^- \rightarrow b\bar{b}\ell^+\nu\ell^-\nu$, with $\ell = e, \mu$. Hadronization and parton showering is performed with Pythia [18] and detector simulation with Delphes [19] using the default card for the CMS detector. Jets are reconstructed with FastJet [20] using the anti- k_T algorithm [21] with radius $R = 0.4$. A probabilistic b -tagging is applied corresponding to the 70% efficiency working point [22]. We apply the kinematical selection

criteria of the CMS entanglement measurement [2], which are inherited from the spin correlation measurement [12]:

- (1) Two opposite-sign charged leptons with pseudorapidity $|\eta| \leq 2.4$, the leading one with transverse momentum $p_T \geq 25$ GeV, and the trailing one with $p_T \geq 20$ GeV. Their invariant mass $m_{\ell\ell}$ has to be larger than 20 GeV.
- (2) Two jets with $|\eta| \leq 2.4$ and $p_T \geq 30$ GeV, one of them b tagged.
- (3) When the two charged leptons have the same flavor, the invariant mass window $76 \leq m_{\ell\ell} \leq 106$ GeV is excluded, and a lower cut $E_T \geq 40$ GeV is placed on the missing transverse energy (MET).

The overall efficiency of this event selection is 0.14.

The final state is reconstructed assuming the kinematics of nearly on-shell production of a top quark pair that decays $t\bar{t} \rightarrow bW^+\bar{b}W^- \rightarrow b\ell^+\nu\bar{b}\ell^-\nu$. Furthermore, it is assumed that the MET originates from the two escaping neutrinos. A minimization is performed to find the values of the neutrino momenta that are most compatible with the assumed $t\bar{t} \rightarrow bW\bar{b}W$ kinematics. For given neutrino momenta p_{ν_1}, p_{ν_2} (three unknowns for each momentum), the W and top quark momenta are reconstructed as

$$\begin{aligned} p_{W_1} &= p_{\ell_1} + p_{\nu_1}, & p_{t_1} &= p_{\ell_1} + p_{\nu_1} + p_{b_1}, \\ p_{W_2} &= p_{\ell_2} + p_{\nu_2}, & p_{t_2} &= p_{\ell_2} + p_{\nu_2} + p_{b_2}, \end{aligned} \quad (2)$$

and their reconstructed invariant masses are labeled as $m_{W_{1,2}}, m_{t_{1,2}}$. In events with two b -tagged jets we select them for the reconstruction, and in events with only one b -tagged jet we attempt the reconstruction selecting also one of the two untagged jets with largest p_T . The labeling of the two neutrinos ν_1, ν_2 is defined by the accompanying charged lepton, while there are two possible pairings of charged leptons and b quarks to reconstruct the top quarks. For each pairing, the neutrino momenta are chosen as the ones that minimize the quantity

$$\begin{aligned} \chi^2 &= \frac{(m_{W_1} - M_W)^2}{\sigma_W^2} + \frac{(m_{W_2} - M_W)^2}{\sigma_W^2} + \frac{(m_{t_1} - m_t)^2}{\sigma_t^2} \\ &+ \frac{(m_{t_2} - m_t)^2}{\sigma_t^2} + \frac{[(p_{\nu_1})_x + (p_{\nu_2})_x - (E_T)_x]^2}{\sigma_p^2} \\ &+ \frac{[(p_{\nu_1})_y + (p_{\nu_2})_y - (E_T)_y]^2}{\sigma_p^2} + \frac{[(p_{t_1})_T - (p_{t_2})_T]^2}{\sigma_p^2}, \end{aligned} \quad (3)$$

and among the two possible pairings between charged leptons and b quarks, we select the pairing with smallest χ^2 . The first four terms favor solutions where the reconstructed masses are close to the true ones, taken as $M_W = 80.4$ GeV, $m_t = 172.5$ GeV, but without explicitly requiring that any of the particles is on its mass shell.

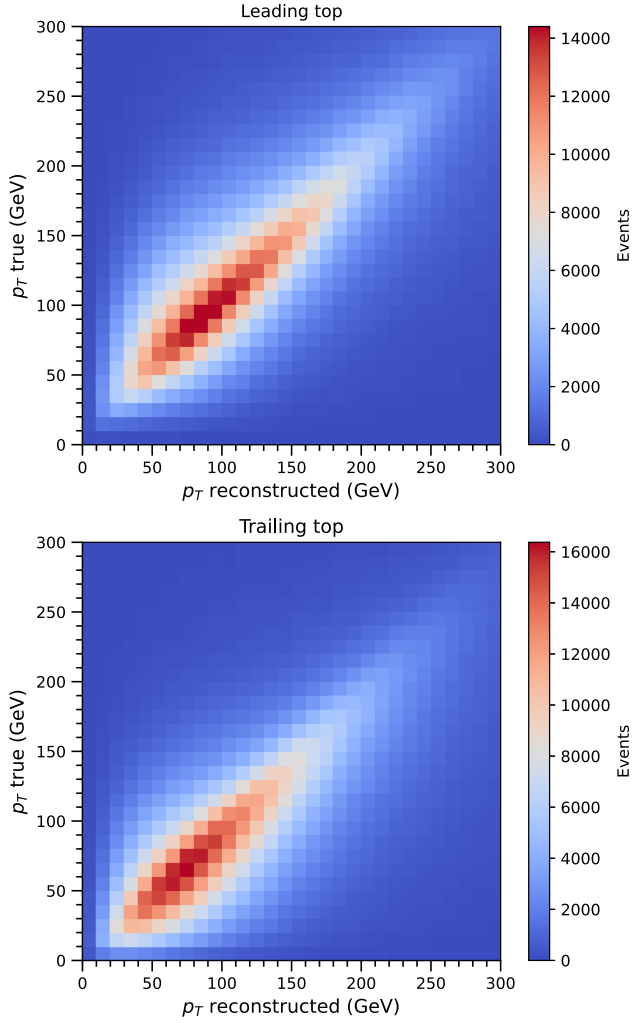


FIG. 1. Reconstructed versus true transverse momenta of the top quarks, for the $t\bar{t}$ process.

The denominators $\sigma_W = 7.5$ GeV, $\sigma_t = 11.5$ GeV [23] represent typical experimental resolutions for the reconstructed width of the top quarks and W bosons. The fifth and sixth terms account for our assumption that MET results from the two neutrinos, while also considering potential mismeasurements. The last term avoids solutions with large transverse momentum imbalance for the two top quarks, with $\sigma_p = 20$ GeV [23] a typical value of the experimental resolution for top transverse momenta. Variations of these values basically give the same results.

Radiation may result in energy loss for b quark jets. To address this, our minimization process includes the possibility of scaling the b quark four-momenta to higher values, up to one standard deviation of the expected jet energy resolution of 15% [24]. This results in a total number of eight variables for the minimization. The stability of the solution for the minimization problem is verified by starting with three different initial values: (i) all

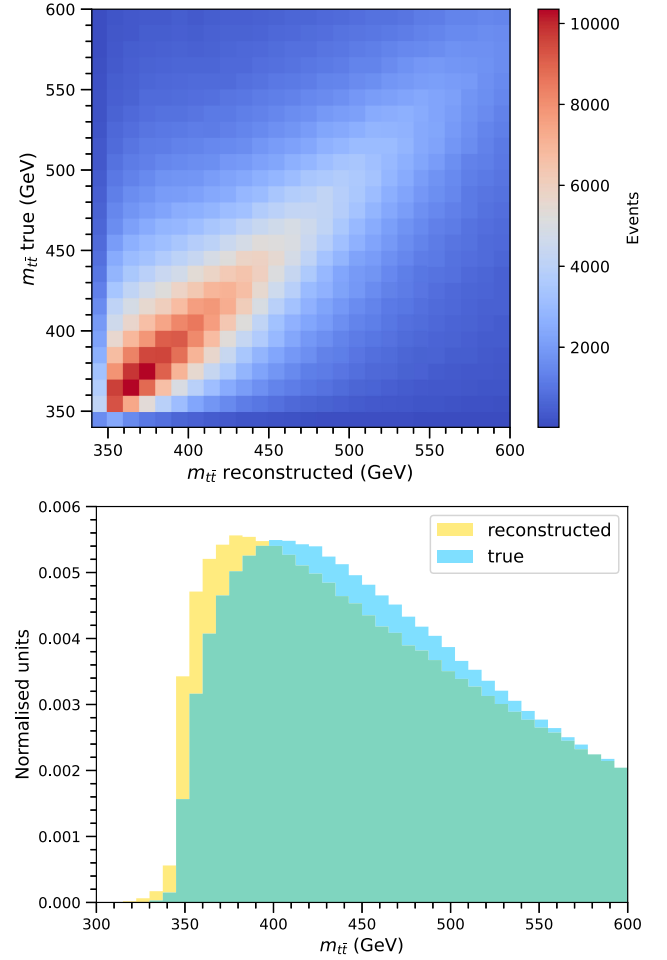


FIG. 2. Top: reconstructed versus true $t\bar{t}$ invariant mass for the $t\bar{t}$ process. Bottom: one-dimensional projection.

the missing energy corresponds to ν_1 , and $(p_{\nu_2})_z = 0$; (ii) the same with $\nu_1 \leftrightarrow \nu_2$; and (iii) the missing energy is equally shared by the two neutrinos. Although there are small numerical differences in the minima found, the resulting kinematical distributions are nearly identical.

Figures 1–3 show the performance of the reconstruction for $pp \rightarrow t\bar{t} \rightarrow bW^+\bar{b}W^- \rightarrow b\ell^+\nu\bar{b}\ell^-\nu$ events, with two nearly on-shell top quarks. In Fig. 1 we present two-dimensional plots for the reconstructed p_T versus true p_T for the leading and trailing top quark. In Fig. 2 we present the true versus reconstructed invariant mass $m_{t\bar{t}}$. We observe a small shift of the reconstructed versus the true value; however, we do not attempt to correct for it because we address the observability of toponium at the detector level. Figure 3 shows the reconstructed W boson and top quark masses.

This $t\bar{t}$ reconstruction is applied to $b\bar{b}WW$ and toponium events. The W boson and top quark reconstructed masses for $b\bar{b}WW$ are presented in Fig. 4. Although this process has diagrams that correspond to single top production, we observe that events are quite compatible with a $t\bar{t}$

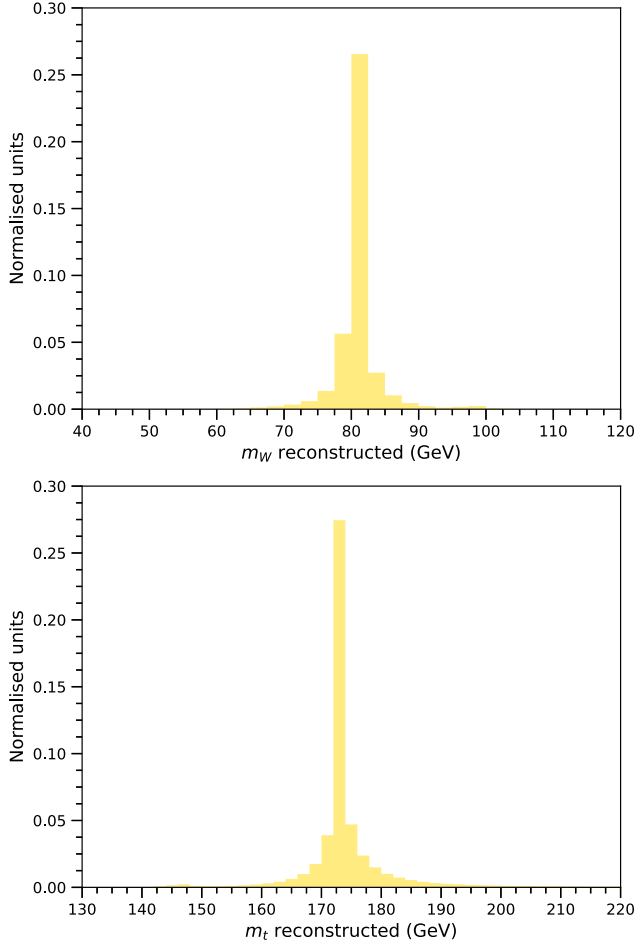


FIG. 3. Reconstructed W boson and top quark masses, for the $t\bar{t}$ process.

kinematics (compare with Fig. 3). Therefore, we do not apply any “quality” cut on the reconstruction, but keep the whole samples fulfilling the selection criteria.

III. MULTIVARIATE ANALYSIS FOR TOPONIUM CHARACTERISATION

As previously discussed, the presence of a toponium “signal” may be verified by an event excess, and deviations in observables characterizing the spin and color of the $t\bar{t}$ pair. While the event excess can be spotted by a simple counting, deviations in spin and color observables require a more sophisticated analysis. To this end, we use an unbinned multivariate test as proposed in Ref. [25] that proceeds in two steps. First, a neural network (NN) is trained to discriminate between “signal” (toponium) and “background” ($b\bar{b}WW$) events, using variables that we describe in the following subsections. Subsequently, the NN score evaluated on signal plus background, and background-only samples, provides two 1D distributions to which a Kolmogorov-Smirnov test [26,27] can be applied to determine the statistical compatibility of both. Using this

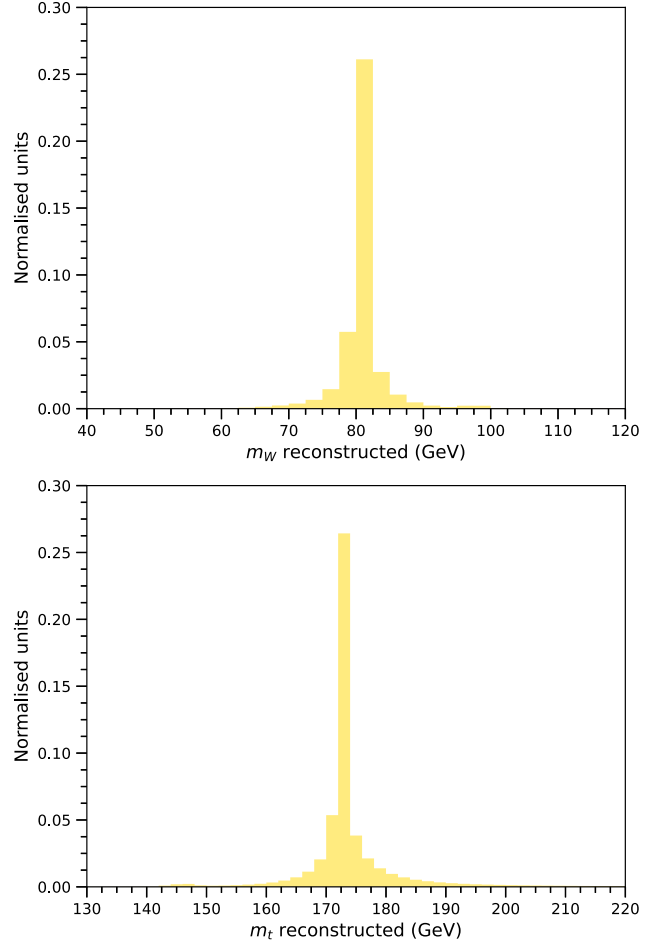


FIG. 4. Reconstructed W boson and top quark masses, for the $b\bar{b}WW$ process.

approach in an experimental analysis, the NNs would be trained with a Monte Carlo simulation, and the background-only sample test would be generated with a Monte Carlo simulation as well. Other statistical methods [28] could be used as well, which skip the need of NN training on signal and background. Preliminary tests show that the expected statistical sensitivity is similar.³

For the training and testing of each NN a standardization of the inputs, based on the background distributions, is performed. Each NN is trained with two samples of 1.5×10^4 signal and 1.5×10^4 background events. The architecture of the NNs is not crucial for the discrimination. We use NNs with two hidden layers of 1024 and 128 nodes, with rectified linear unit activation for the hidden layers and a sigmoid function for the output one. For simplicity we keep the same NN architecture even if the number of inputs (5 for spin, 38 for color) is very different. We have tested that other NN architectures provide similar results. The NN

³I thank Gaia Grosso for performing some tests using the NPLM method.

optimization relies on the binary cross-entropy loss function, using the Adam [29] optimizer. Overtraining is avoided by monitoring the NN performance on validation samples, of the same size as the training samples, and stopping the training when the performance ceases to improve. The NNs are implemented using Keras [30] with a TensorFlow backend [31].

A. Spin variables

The spin state of the $t\bar{t}$ pair is fully characterized by the four-dimensional distribution of polar and azimuthal angles of the two charged leptons in the rest frame of their parent top quark. As reference system we use the so-called helicity basis [32] $(\hat{r}, \hat{n}, \hat{k})$, with the axes defined as

- (1) K axis (helicity): \hat{k} is a normalized vector in the direction of the top quark three-momentum in the $t\bar{t}$ rest frame.
- (2) R axis: \hat{r} is in the production plane and defined as $\hat{r} = (\hat{p}_p - \cos\theta_t \hat{k}) / \sin\theta_t$, with $\hat{p}_p = (0, 0, 1)$ the momentum of one of the initial protons and θ_t the production angle in the c.m. frame.
- (3) N axis: $\hat{n} = \hat{k} \times \hat{r}$ is orthogonal to the production plane.

The same basis is used for the top quark and antiquark, and with respect to this basis, the rest-frame angles (θ_1, ϕ_1) for ℓ^+ and (θ_2, ϕ_2) for ℓ^- are defined. From the kinematical distribution of these four quantities, the spin density operator of the $t\bar{t}$ pair can be determined. However, close to the threshold where the top quarks are nearly at rest in the $t\bar{t}$ rest frame, an accurate determination of the helicity basis is difficult. Therefore, in addition to $\theta_{1,2}$ and $\phi_{1,2}$ we include in our set the angle between the two lepton momenta θ_{12} , which is basis independent. We present this distribution in Fig. 5 for events with reconstructed invariant mass $m_{t\bar{t}} \leq 400$ GeV.

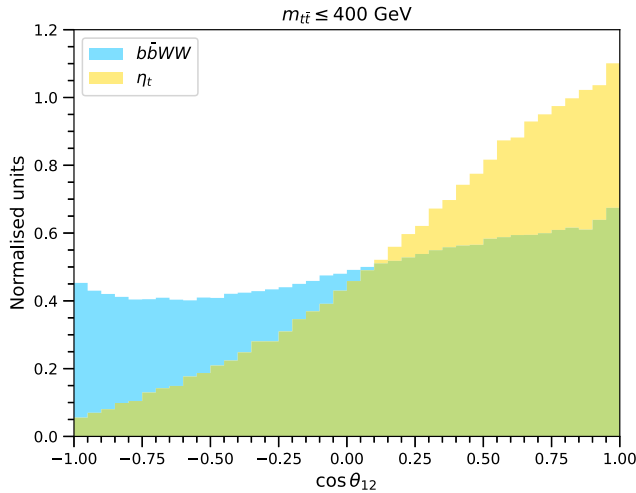


FIG. 5. Kinematical distribution of $\cos\theta_{12}$ for $b\bar{b}WW$ and the toponium signal.

B. Color variables

The discrimination between $b\bar{b}WW$ and η_t is mainly based on the global shape of the event, which is parametrized by a set of variables adapted from the ones introduced to characterize jet substructure [33–35]. We define them as

$$\tau_n^{(\beta)} = \frac{1}{\sum_i E_i} \sum_i p_{Ti} \min \{ \Delta R_{1i}^\beta, \Delta R_{2i}^\beta, \dots, \Delta R_{ni}^\beta \}, \quad (4)$$

with i labeling the detector-level particles in the event, and p_{Ti}, E_i as their transverse momentum and energy, respectively; $\Delta R_{ki} = [\Delta\eta_{ki}^2 + \Delta\phi_{ki}^2]^{1/2}$ is the lego-plot distance between the momentum of the particle i and the axis $k = 1, \dots, n$. In contrast to their widespread use for jet substructure, here we use $\tau_n^{(\beta)}$ to characterize the global structure of the event, with $n = 1, \dots, 6$ and $\beta = 1, 2, 0.5$. Note that the variables in (4) are not equivalent to the N jettiness introduced in Ref. [36]. Additionally, we use the subjettiness of the two jets that are identified by the kinematical reconstruction (cf. Sec. II) as corresponding to the hadronization of the two b quarks. In this case, we use $n = 1, \dots, 3$ and $\beta = 1, 2, 0.5$. The kinematical distributions of $\tau_n^{(\beta)}$ are not very different for $b\bar{b}WW$ and toponium, but altogether these sets of variables provide a good discrimination between the color-singlet and color-octet configurations. For illustration, Fig. 6 shows $\tau_1^{(1)}$ of the event and the b -tagged jet, for $b\bar{b}WW$ and the toponium signal. We have verified that including higher-order $\tau_n^{(\beta)}$ does not improve the results.

Additional variables providing some discrimination power are the jet multiplicity (for which we consider $p_T \geq 20$ GeV), and the total number of detector-level particles in the event. The jet pull [37] is an observable designed to probe color flow between jets, which has been used by the D0 [38] and ATLAS [39] collaborations to measure color flow in the hadronic decays of W bosons produced in the semileptonic decay of $t\bar{t}$ pairs. However, we find that these variables are strongly dependent on $m_{t\bar{t}}$, and bin migrations wash out the differences between $b\bar{b}WW$ and toponium, rendering their kinematical distributions nearly identical at the reconstructed level.

IV. STATISTICAL SENSITIVITY TO TOPONIUM

The toponium signal may be spotted by its effect near the $t\bar{t}$ threshold, which motivates an analysis in different bins of $m_{t\bar{t}}$. The bin size is a compromise between sample size and signal to background ratio. The Kolmogorov-Smirnov test used to detect deviations is more powerful for larger samples but, on the other hand, the signal (toponium) to background ($b\bar{b}WW$) ratio is larger near threshold. Additionally, the energy resolution broadens the toponium

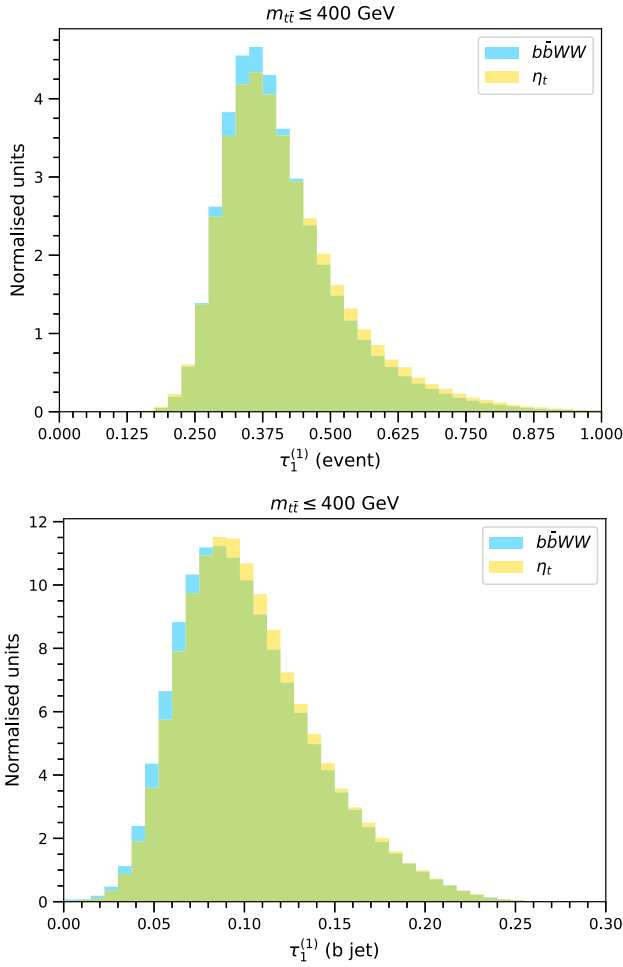


FIG. 6. $\tau_1^{(1)}$ of the event (top) and of the b -tagged jet (bottom), for $b\bar{b}WW$ and the toponium signal.

signal to higher values of $m_{\bar{t}t}$. We therefore use bins of 20 GeV with the first bin $m_{\bar{t}t} \leq 360$ GeV.

For the toponium signal, Ref. [10] used calculations in [6] to extract a toponium cross section $\sigma(\eta_t) = 6.43$ pb at 13 TeV. On the other hand, Ref. [17] used calculations in [5] to extract a toponium cross section $\sigma(\eta_t) = 3.6$ pb at 14 TeV (at 13 TeV the cross section is 15% smaller). We use the former as our baseline benchmark, since the entanglement measurement performed by the CMS collaboration [2] obtained very good agreement with data using that value, and provide alternative results for the lower cross section. For the $b\bar{b}WW$ background there are no calculations of the cross section beyond the leading order (LO); at LO its cross section is 1.1 times larger than for $\bar{t}t$. At NLO the $\bar{t}t$ cross section using NNPDF 3.1 parton density functions is 671 pb. Then, we assume a next-to-leading order cross section of $k \times 671$ pb for $b\bar{b}WW$, with $k = 1.1$ our baseline choice. We also explore $k = 1.5$ as conservative background estimate. We note that the $\bar{t}t$ cross section is known to next-to-next-to-leading order [40]. However, since the predictions in Refs. [5,6] from which

TABLE I. Expected number of toponium and $b\bar{b}WW$ events in selected bins of $m_{\bar{t}t}$, for the baseline benchmark.

| $m_{\bar{t}t}$ (GeV) | Run 2 | | Run 2 + 3 | | S/B |
|----------------------|----------|--------------|-----------|--------------|-------|
| | η_t | $b\bar{b}WW$ | η_t | $b\bar{b}WW$ | |
| ≤ 360 | 1640 | 38160 | 4370 | 101680 | 0.041 |
| 360–380 | 1180 | 57000 | 3140 | 151890 | 0.020 |
| 380–400 | 740 | 58000 | 1970 | 154540 | 0.013 |
| 400–420 | 500 | 55160 | 1340 | 146950 | 0.009 |

the toponium cross sections are extracted are at NLO in perturbative QCD, for consistency we also use background predictions at NLO. For the baseline benchmark the expected number of events in each $m_{\bar{t}t}$ bin is collected in Table I, for luminosities corresponding to LHC run 2 (140 fb^{-1}) and run 2 + 3 (350 fb^{-1}). We also include the signal (S) to background (B) ratio in the last column.

For the excess events the expected statistical significance in each bin is computed as S/\sqrt{B} . The expected statistical significance of the deviations in the multidimensional probability density functions of spin and color variables is calculated from pseudoexperiments. In each pseudoexperiment, two samples are compared: (i) a sample consisting of S signal and B background events, which is taken as pseudodata; (ii) a sample of $S + B$ background events. Both signal and random events are randomly chosen from large pools of Monte Carlo generated events. Repeating the pseudoexperiments 10^4 times, we obtain the probability density function (PDF) for the expected significance, from which we take the mean value as expected significance. Likewise, one can perform pseudoexperiments comparing two independent samples of S signal and B background events, to verify the level of agreement of pseudodata with toponium production. As an example, we show the results

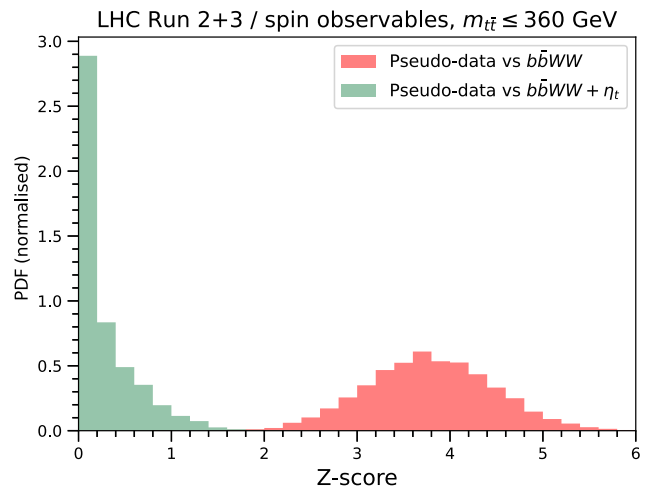


FIG. 7. Example of distributions for the statistical significance of the deviations between pseudodata (with injected toponium) and the $b\bar{b}WW$ and $b\bar{b}WW + \eta_t$ hypotheses, obtained from pseudoexperiments.

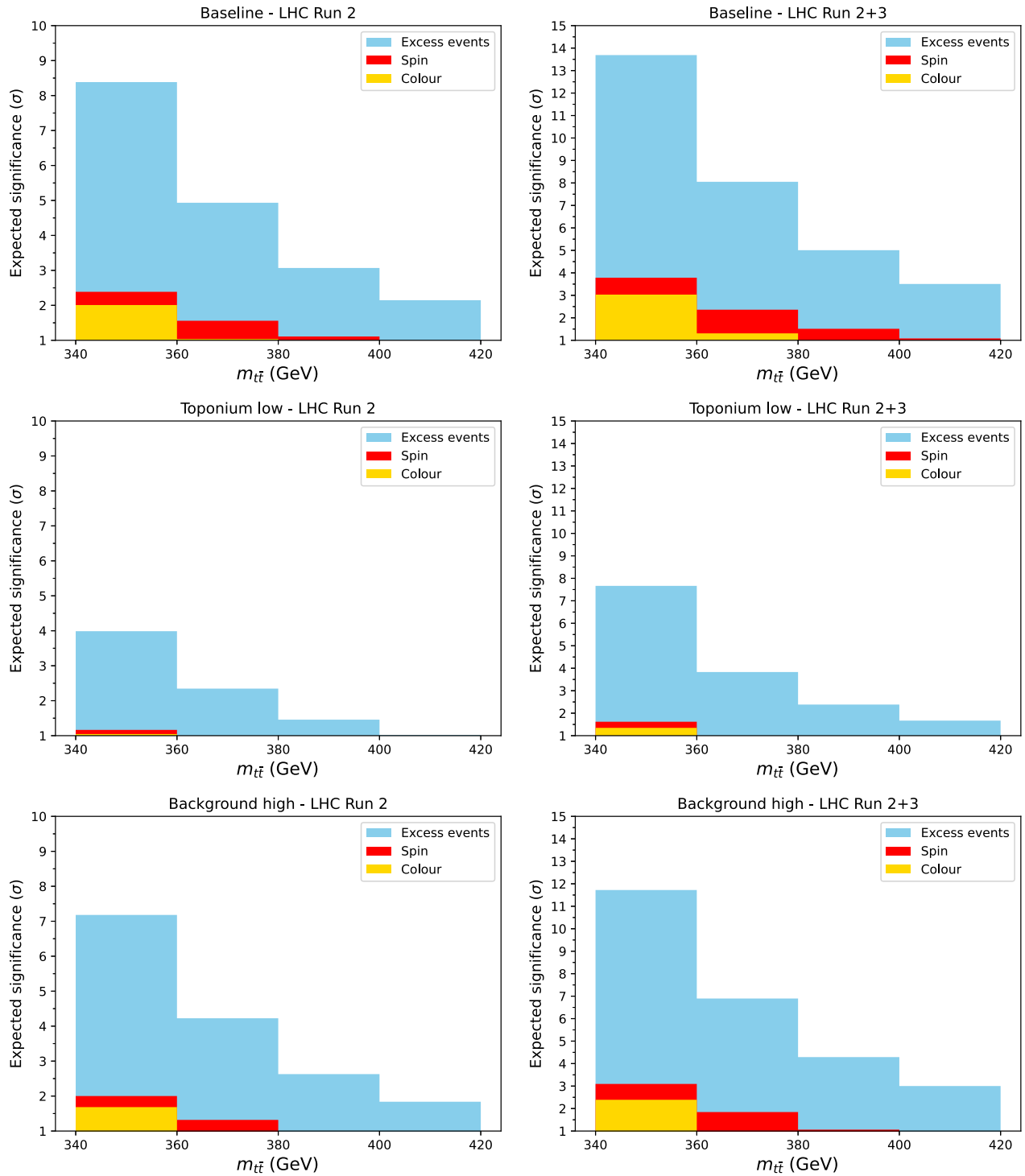


FIG. 8. Expected statistical significance for the toponium signal with run 2 and run 2 + 3 data for three benchmarks described in the text.

for spin variables in the bin $m_{t\bar{t}} \leq 360$ GeV in Fig. 7. The size of the samples corresponds to run 2 + 3 luminosity in the baseline benchmark. The red distribution shows the disagreement of pseudodata with the $b\bar{b}WW$ hypothesis, while the green distribution exhibits the agreement with the $b\bar{b}WW + \eta_t$ hypothesis. Of course, the latter depends on

the toponium cross section assumed, but this could be fixed from data by the excess events observed.

The statistical significance of the deviations with respect to the $b\bar{b}WW$ hypothesis is presented in Fig. 8 for luminosities corresponding to LHC run 2 and run 2 + 3, and the three benchmarks mentioned:

(i) Baseline: $\sigma(\eta_t) = 6.43$ pb, $k = 1.1$;
(ii) “toponium low,” with $\sigma(\eta_t) = 3.06$ pb, $k = 1.1$;
(iii) “background high,” with $\sigma(\eta_t) = 6.43$ pb, $k = 1.5$.
For each bin of $m_{\bar{t}t}$ the height of the bars indicate the expected statistical significance of the toponium signal. We discard significances lower than 1σ . As it can be readily observed, the toponium signal produces an excess of events near threshold whose spin and color properties are not consistent with continuum $\bar{t}t$ production but would agree with the production of a color-singlet scalar. This excess could be visible and identified as toponium, provided the experimental and modeling uncertainties are under control, as discussed in the next section.

V. DISCOVERY STRATEGIES

The discovery and characterization of toponium is a formidable task that requires a very accurate theoretical modeling. The dilepton decay mode is particularly clean with quite small backgrounds. Therefore, one expects that the main source of systematic uncertainties will be the theoretical modeling, in addition to experimental uncertainties.

The toponium signal leads to localized deviations close to the $\bar{t}t$ threshold, as it can be seen in Fig. 8. These deviations decrease quickly, and for $m_{\bar{t}t} \geq 400$ GeV the toponium effect on spin and color observables is below 1σ (statistical) even for run 2 + 3 data. Then, it is conceivable that experiments could use a “control” region $m_{\bar{t}t} \geq 400$ GeV to tune the theoretical predictions for the relevant observables, so that their uncertainty in the “measurement” region $m_{\bar{t}t} \leq 360$ GeV can be reduced and the toponium signal is visible. We note that these are detector-level values, and the difference with the parton-level $m_{\bar{t}t}$ is in the 10–20 GeV ballpark, see Fig. 2. Therefore, the region $m_{\bar{t}t} \leq 360$ GeV at the detector level is not much stricter than the ones used by the ATLAS and CMS collaborations in the entanglement measurements [11,12].

The identification of the event excess as toponium requires to verify that the features of the excess are (i) not compatible with $b\bar{b}WW$ production; (ii) compatible with $b\bar{b}WW + \eta_t$ production. In order to verify the latter, which requires as input the toponium cross section, one can directly use data, i.e., the size of the event excess. Theoretical predictions of the toponium cross section have a large uncertainty, as evidenced by the different cross sections $\sigma(\eta_t) = 6.43$ pb, $\sigma(\eta_t) = 3.06$ pb obtained from Refs. [5,6], respectively. And, in any case, the toponium cross section extracted from the dilepton and semileptonic decay modes would have to be in agreement with theoretical expectations.

The theoretical modeling of the observables used to characterize color seems especially difficult. However, we point out that a multidimensional NN discriminant based on $\tau_n^{(\beta)}$ variables may be more resilient against mismodeling than the input variables themselves. An example of this

nice behavior was found in Ref. [41] for jet taggers. While the subjettness variables used in the NN exhibit differences depending on the hadronization scheme, the NN discriminant turns out to be rather insensitive.

In summary, the experimental discovery of toponium and characterization of its properties seems a quite difficult endeavor, yet not impossible, as we have argued. Improved modeling and data-driven calibration of theoretical predictions might render visible the expected deviations in the number of events, spin, and color observables that we have pointed out in this work.

ACKNOWLEDGMENTS

I gratefully acknowledge many useful discussions with M. L. Mangano and help with toponium modeling, and the CERN Theory Department for hospitality during the realization of this work. I also thank C. Severi for correspondence regarding Ref. [17]. This work has been supported by the Spanish Research Agency (Agencia Estatal de Investigación) through Projects No. PID2022-142545NB-C21, and No. CEX2020-001007-S funded by MCIN/AEI/10.13039/501100011033, and by Fundação para a Ciência e a Tecnologia (FCT, Portugal) through Project No. CERN/FIS-PAR/0019/2021.

APPENDIX: ANALYSIS OF THE SEMILEPTONIC DECAY CHANNEL

Despite the larger branching ratio, the semileptonic final state $b\bar{b}\ell\nu q\bar{q}'$ is not competitive with the dilepton one for toponium characterization. We focus here on the lowest invariant mass bin $m_{\bar{t}t} \leq 360$ GeV. We generate toponium (10^6 events) and $b\bar{b}WW$ (6×10^6 events) in the semileptonic decay mode, with a generator-level upper cut on the total invariant mass $M \leq 400$ GeV. Events are processed with Pythia and Delphes as described for the dilepton final state. For event selection we require (see, e.g., Ref. [42]):

- (1) One charged lepton ℓ with $|\eta| \leq 2.4$ and $p_T \geq 30$ GeV.
- (2) Two b -tagged jets with $|\eta| \leq 2.4$ and $p_T \geq 30$ GeV.
- (3) At least two untagged jets with $|\eta| \leq 2.4$ and $p_T \geq 30$ GeV.

The reconstruction of the W and top quark momenta is done as follows. The momentum of the W boson decaying hadronically is reconstructed as $p_{W_h} = p_{j_1} + p_{j_2}$, with j_1, j_2 the two nontagged jets, among the three with largest p_T which have invariant mass closest to M_W . The neutrino momentum is reconstructed defining $(p_\nu)_{x,y} = (\vec{E}_T)_{x,y}$ and solving the quadratic equation $(p_\ell + p_\nu)^2 = M_W^2$ for $(p_\nu)_z$. (If the equation does not have real solutions then we set the discriminant to zero.) The momentum of the W boson decaying leptonically is then $p_{W_l} = p_\ell + p_\nu$, and the momenta of the two top quarks

$$\begin{aligned} p_{t_l} &= p_{W_l} + p_{b_1}, \\ p_{t_h} &= p_{W_h} + p_{b_2}. \end{aligned} \quad (\text{A1})$$

In general there are two solutions for $(p_\nu)_z$ and two possible pairings in (A1). We chose the ones that minimize the quantity

$$\chi^2 = (m_{t_l} - m_t)^2 + (m_{t_h} - m_t)^2, \quad (\text{A2})$$

with $m_{t_l} = \sqrt{p_{t_l}^2}$, $m_{t_h} = \sqrt{p_{t_h}^2}$ the invariant masses of the leptonically and hadronically decaying top quark, respectively. Spin variables are defined as outlined in Sec. III A, but using for the hadronic top decay the optimal polarimeter introduced in Ref. [43]. Color observables are defined as in Sec. III B.

The expected number of events with $m_{t\bar{t}} \leq 360$ GeV for LHC run 2 is 95830 for $b\bar{b}WW$ and 4150 for toponium, i.e., around 2.5 times larger than in the dilepton decay mode, cf. Table I. Note that the ratio $S/B = 0.041$ is quite close to

the one in the dilepton channel—leaving aside the fact that other non- $t\bar{t}$ backgrounds are larger in this channel than in the dilepton mode. Still, the sensitivity to toponium using spin and color observables is much smaller, 0.8σ and 1.5σ , respectively. This is quite as expected. For spin observables the use of the hadronic polarimeter, with smaller spin analyzing power than the charged lepton, leads to smaller differences in the distributions. In addition, near threshold where the jets are not very energetic, it is found that often one of the two untagged jets used to reconstruct the hadronically decaying W boson actually does not correspond to the W decay—which spoils the rest-frame angular distributions. (A quality cut on the reconstructed hadronic W mass does not improve the sensitivity.) This smaller sensitivity to toponium is also in agreement with the smaller entanglement significance found by the CMS collaboration in the semileptonic mode near threshold [44]. For color observables, the smaller sensitivity can be attributed to the additional hadronic activity that washes out the differences due to the color connection between the top quarks.

-
- [1] G. Aad *et al.* (ATLAS Collaboration), Observation of quantum entanglement in top-quark pairs using the ATLAS detector, [arXiv:2311.07288](https://arxiv.org/abs/2311.07288).
- [2] CMS Collaboration, Observation of quantum entanglement in top quark pair production in proton-proton collisions at $\sqrt{s} = 13$ TeV, [arXiv:2406.03976](https://arxiv.org/abs/2406.03976).
- [3] V. S. Fadin, V. A. Khoze, and T. Sjostrand, On the threshold behavior of heavy top production, *Z. Phys. C* **48**, 613 (1990).
- [4] K. Hagiwara, Y. Sumino, and H. Yokoya, Bound-state effects on top quark production at hadron colliders, *Phys. Lett. B* **666**, 71 (2008).
- [5] Y. Kiyo, J. H. Kuhn, S. Moch, M. Steinhauser, and P. Uwer, Top-quark pair production near threshold at LHC, *Eur. Phys. J. C* **60**, 375 (2009).
- [6] Y. Sumino and H. Yokoya, Bound-state effects on kinematical distributions of top quarks at hadron colliders, *J. High Energy Phys.* **09** (2010) 034; **06** (2016) 037.
- [7] V. S. Fadin and V. A. Khoze, Threshold behavior of heavy top production in e^+e^- collisions, *JETP Lett.* **46**, 525 (1987).
- [8] M. J. Strassler and M. E. Peskin, The heavy top quark threshold: QCD and the Higgs, *Phys. Rev. D* **43**, 1500 (1991).
- [9] Y. Sumino, K. Fujii, K. Hagiwara, H. Murayama, and C. K. Ng, Top quark pair production near threshold, *Phys. Rev. D* **47**, 56 (1993).
- [10] B. Fuks, K. Hagiwara, K. Ma, and Y. J. Zheng, Signatures of toponium formation in LHC run 2 data, *Phys. Rev. D* **104**, 034023 (2021).
- [11] M. Aaboud *et al.* (ATLAS Collaboration), Measurements of top-quark pair spin correlations in the $e\mu$ channel at $\sqrt{s} = 13$ TeV using pp collisions in the ATLAS detector, *Eur. Phys. J. C* **80**, 754 (2020).
- [12] A. M. Sirunyan *et al.* (CMS Collaboration), Measurement of the top quark polarization and $t\bar{t}$ spin correlations using dilepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV, *Phys. Rev. D* **100**, 072002 (2019).
- [13] A. Behring, M. Czakon, A. Mitov, A. S. Papanastasiou, and R. Poncelet, Higher order corrections to spin correlations in top quark pair production at the LHC, *Phys. Rev. Lett.* **123**, 082001 (2019).
- [14] See <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCtopWGSummaryPlots>
- [15] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations, *J. High Energy Phys.* **07** (2014) 079.
- [16] R. D. Ball *et al.* (NNPDF Collaboration), Parton distributions from high-precision collider data, *Eur. Phys. J. C* **77**, 663 (2017).
- [17] F. Maltoni, C. Severi, S. Tentori, and E. Vryonidou, Quantum tops at circular lepton colliders, *J. High Energy Phys.* **09** (2024) 001.
- [18] T. Sjostrand, S. Mrenna, and P. Z. Skands, A brief Introduction to Pythia 8.1, *Comput. Phys. Commun.* **178**, 852 (2008).
- [19] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaitre, A. Mertens, and M. Selvaggi, Delphes 3,

- A modular framework for fast simulation of a generic collider experiment, *J. High Energy Phys.* **02** (2014) 057.
- [20] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet* user manual, *Eur. Phys. J. C* **72**, 1896 (2012).
- [21] M. Cacciari, G. P. Salam, and G. Soyez, The anti- k_r jet clustering algorithm, *J. High Energy Phys.* **04** (2008) 063.
- [22] CMS Collaboration, Identification of b-quark jets with the CMS experiment, *J. Instrum.* **8**, P04013 (2013).
- [23] J. A. Aguilar-Saavedra, M. C. N. Fiolhais, P. Martín-Ramiro, J. M. Moreno, and A. Onofre, A template method to measure the $t\bar{t}$ polarisation, *Eur. Phys. J. C* **82**, 134 (2022).
- [24] V. Khachatryan *et al.* (CMS Collaboration), Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV, *J. Instrum.* **12**, P02014 (2017).
- [25] M. Williams, How good are your fits? Unbinned multivariate goodness-of-fit tests in high energy physics, *J. Instrum.* **5**, P09004 (2010).
- [26] A. N. Kolmogorov, Sulla determinazione empirica di una legge di distribuzione, *G. Ist. Ital. Attuari* **4**, 83 (1933).
- [27] N. V. Smirnov, Table for estimating the goodness of fit of empirical distributions, *Ann. Math. Stat.* **19**, 279 (1948).
- [28] G. Grosso, M. Letizia, M. Pierini, and A. Wulzer, Goodness of fit by Neyman-Pearson testing, *SciPost Phys.* **16**, 123 (2024).
- [29] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, [arXiv:1412.6980](https://arxiv.org/abs/1412.6980).
- [30] F. Chollet, Keras: Deep learning for Python (2015), <https://github.com/fchollet/keras>.
- [31] M. Abadi *et al.*, TensorFlow: Large-scale machine learning on heterogeneous systems (2015), <http://tensorflow.org/>.
- [32] W. Bernreuther, D. Heisler, and Z. G. Si, A set of top quark spin correlation and polarization observables for the LHC: Standard Model predictions and new physics contributions, *J. High Energy Phys.* **12** (2015) 1.
- [33] J. Thaler and K. Van Tilburg, Identifying boosted objects with N-subjettiness, *J. High Energy Phys.* **03** (2011) 015.
- [34] J. Thaler and K. Van Tilburg, Maximizing boosted top identification by minimizing N-subjettiness, *J. High Energy Phys.* **02** (2012) 093.
- [35] K. Datta and A. Larkoski, How much information is in a jet?, *J. High Energy Phys.* **06** (2017) 073.
- [36] A. Banfi, M. Dasgupta, K. Khelifa-Kerfa, and S. Marzani, Non-global logarithms and jet algorithms in high- p_T jet shapes, *J. High Energy Phys.* **08** (2010) 064.
- [37] J. Gallicchio and M. D. Schwartz, Seeing in color: Jet superstructure, *Phys. Rev. Lett.* **105**, 022001 (2010).
- [38] V. M. Abazov *et al.* (D0 Collaboration), Measurement of color flow in $t\bar{t}$ events from $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV, *Phys. Rev. D* **83**, 092002 (2011).
- [39] G. Aad *et al.* (ATLAS Collaboration), Measurement of colour flow with the jet pull angle in $t\bar{t}$ events using the ATLAS detector at $\sqrt{s} = 8$ TeV, *Phys. Lett. B* **750**, 475 (2015).
- [40] M. Czakon, P. Fiedler, and A. Mitov, Total top-quark pair-production cross section at hadron colliders through $O(\alpha_s^4)$, *Phys. Rev. Lett.* **110**, 252004 (2013).
- [41] J. A. Aguilar-Saavedra, Taming modeling uncertainties with mass unspecific supervised tagging, *Eur. Phys. J. C* **82**, 270 (2022).
- [42] A. Tumasyan *et al.* (CMS Collaboration), Search for CP violation using $t\bar{t}$ events in the lepton + jets channel in pp collisions at $\sqrt{s} = 13$ TeV, *J. High Energy Phys.* **06** (2023) 081.
- [43] B. Tweedie, Better hadronic top quark polarimetry, *Phys. Rev. D* **90**, 094010 (2014).
- [44] CMS Collaboration, Measurements of polarization, spin correlations, and entanglement in top quark pairs using lepton + jets events from pp collisions at $\sqrt{s} = 13$ TeV, Report No. CMS-PAS-TOP-23-007.