

PREPARED FOR SUBMISSION TO JHEP

# Matching in $pp \rightarrow t\bar{t}W/Z/h + \text{jet}$ SMEFT studies

Reza Goldouzian,<sup>a</sup> Jeong Han Kim,<sup>a,b</sup> Kevin Lannon,<sup>a</sup> Adam Martin,<sup>a</sup> Kelci Mohrman,<sup>a</sup> and Andrew Wightman<sup>a</sup>

<sup>a</sup>*Department of Physics, 225 Nieuwland Science Hall, University of Notre Dame, Notre Dame, IN 46556, USA*

<sup>b</sup>*Department of Physics, Chungbuk National University, Cheongju, Chungbuk 28644, Republic of Korea*

*E-mail:* [reza.goldouzian@cern.ch](mailto:reza.goldouzian@cern.ch), [jeonghan.kim@cbu.ac.kr](mailto:jeonghan.kim@cbu.ac.kr), [klannon@nd.edu](mailto:klannon@nd.edu), [amarti41@nd.edu](mailto:amarti41@nd.edu), [kmohrman@nd.edu](mailto:kmohrman@nd.edu), [andrew.wightman.2@nd.edu](mailto:andrew.wightman.2@nd.edu)

**ABSTRACT:** In this paper, we explore the impact of extra radiation on predictions of  $pp \rightarrow t\bar{t}X$ ,  $X = h/W^\pm/Z$  processes within the dimension-6 SMEFT framework. While full next-to-leading order calculations are of course preferred, they are not always practical, and so it is useful to be able to capture the impacts of extra radiation using leading-order matrix elements matched to the parton shower. While a matched leading-order calculation for  $t\bar{t}X$  is not expected to reproduce the next-to-leading order inclusive cross section precisely, we show that it does capture the relative impact of the EFT effects by considering the ratio of matched SMEFT inclusive cross sections to Standard Model values,  $\sigma_{\text{SMEFT}}(t\bar{t}X+j)/\sigma_{\text{SM}}(t\bar{t}X+j) \equiv \mu$ . Furthermore, we compare leading order calculations with and without extra radiation and find several cases, such as the effect of the operator  $(\varphi^\dagger i \overleftrightarrow{D}_\mu \varphi)(\bar{t}\gamma^\mu t)$  on  $t\bar{t}h$  and  $t\bar{t}W$ , for which the relative cross section prediction increases by more than 10%—significantly larger than the uncertainty derived by varying the input scales in the calculation, including the additional scales required for matching. Being leading order at heart, matching has the benefit that it can be applied to all operators and processes relevant to  $pp \rightarrow t\bar{t}X$ ,  $X = h/W^\pm/Z + \text{jet}$ , is computationally fast and not susceptible to negative weights. Therefore, it is a useful approach in  $t\bar{t}X+\text{jet}$  studies where complete next-to-leading order results are currently unavailable or unwieldy.

---

## Contents

<b>1</b>	<b>Motivation</b>	<b>2</b>
<b>2</b>	<b>The SMEFT framework</b>	<b>5</b>
<b>3</b>	<b>Jet matching/merging for EFT samples</b>	<b>7</b>
3.1	Theoretical justification	7
3.1.1	$g$ - $g$ - $t$ - $t$ - $h$ and $g$ - $g$ - $t$ - $t$ vertices: a case study on $t\bar{t}h+j$ final state	8
3.1.2	An anomalous $t$ - $t$ - $g$ coupling	9
3.2	Monte Carlo validation of matching	14
<b>4</b>	<b>Results</b>	<b>16</b>
4.1	Comparison methodology	17
4.2	The EFT effects of including an extra parton in the LO calculation	19
4.3	Discussion	22
4.3.1	Discussion of $t\bar{t}h$	23
4.3.2	Discussion of $t\bar{t}Z$	24
4.3.3	Discussion of $t\bar{t}W$	25
<b>5</b>	<b>Matching systematics</b>	<b>25</b>
<b>6</b>	<b>Conclusion</b>	<b>27</b>
<b>A</b>	<b>Comparison of CPU and negative weights for NLO vs LO calculations</b>	<b>29</b>
A.1	CPU comparison	29
A.2	Negative weights	30
<b>B</b>	<b>Feynman rules for <math>t</math>-<math>t</math>-<math>g</math> couplings</b>	<b>30</b>
<b>C</b>	<b>Comparison of fit parameters for <math>t\bar{t}X</math> and <math>t\bar{t}X+j</math> processes</b>	<b>30</b>

---

## 1 Motivation

The Standard Model effective field theory [1–4] (SMEFT) provides a useful, bottom-up framework for new physics searches at the LHC. The SMEFT assumes that all new particles are too heavy to produce on-shell, and encompasses their effects into the coefficients of higher dimensional operators formed from SM fields and their derivatives. The higher dimensional

operators generate new interactions, modify existing interactions, and alter the relationship between Lagrangian parameters and observables.

Recently, the SMEFT approach has been applied to search for new physics in the production of top quarks produced in association with bosons, such as  $t\bar{t}h$ ,  $t\bar{t}W$ , and  $t\bar{t}Z$  [5–8]; see Refs. [9–11] for experimental analyses. Such analyses rely on predictions for the size and nature of SMEFT effects usually obtained from matrix element calculations interfaced to the parton shower, such as **MADGRAPH** [12] interfaced with **PYTHIA** [13, 14]. In such predictions, the minimal production of  $t\bar{t}X$ —where “X” refers to one W, Z, or h—is seen as the leading contribution, while production with one additional jet, e.g.  $t\bar{t}X+j$  is viewed as a higher-order correction.

However, there are many examples of processes for which production with one additional jet turns out to be not just a small correction. An obvious case would be single top production, for which t-channel production dominates over s-channel production. There are also cases for which the EFT contribution arising from the process with one extra jet is significant, or even dominant over the process with no additional jets. One would like to incorporate these radiation effects into SMEFT analyses.

To study the impact of extra jets, the obvious solution is to go to next-to-leading order (NLO). In fact, an NLO implementation containing all bosonic (dimension-6) operators, all operators with two fermions, and four fermion operators with at least one top quark was recently released [15]. However, SMEFT predictions at NLO require additional counterterms and are a significant undertaking; see Refs. [16–20] for  $t\bar{t}X$  specific NLO details. The practical consequences of the counterterms are that the NLO Monte Carlo takes significantly more CPU-time to generate events and tends to produce a substantial fraction of events with negative weights<sup>1</sup>. These considerations can make it challenging to make use of NLO Monte Carlo.

An additional complication arises when generating NLO samples involving dimension-6 operators and processes with electroweak vertices. Calculations in **MADGRAPH** can be done to a fixed order in QCD coupling, QED coupling, new physics coupling, etc. At NLO, fixing the coupling order is a must, as **MADGRAPH** is currently only capable of accounting for NLO QCD effects. Since electroweak loops cannot be included in the NLO calculation, tree-level diagrams with QED order greater than or equal to two plus the lowest QED order tree-level diagrams are not permitted by the code. The addition of new couplings, i.e. the SMEFT Wilson coefficients, introduces a potential ambiguity, as the effective QED order assigned to the new couplings can impact whether or not diagrams involving these couplings can be included in the NLO calculation. Furthermore, the QED orders assigned to SMEFT couplings is somewhat arbitrary and varies among UFO models in the literature. For example, let us consider the dimension-6 SMEFT operator  $\mathcal{O}_{tB}$ . In some models, the coupling associated with  $\mathcal{O}_{tB}$  has a QED order of 0 or 1, while in others it can be 1 or 2 [15, 21]; the way in which  $\mathcal{O}_{tB}$  contributes to a process at a fixed order therefor depends on this subjective

---

<sup>1</sup>For a quantitative example of these differences, see Appendix A.

assignment. These coupling technicalities could potentially be overcome, but this would likely require carefully and manually adjusting the order of each operator. Thus, while a full NLO SMEFT treatment is ideal and available for numerous dedicated processes and operators, its automated use may not always be practical, especially for processes with electroweak gauge bosons.

However, studies of SMEFT in processes with additional radiation can also be carried out at tree-level using matching [22–25]. Tree level calculations, unlike NLO, may be run without any **MADGRAPH** restriction on the coupling order, regardless how new physics couplings are classified (i.e. their QED or QCD order). In the matching procedure, the phase space of the additional radiation is divided up into two regions. Harder emissions are handled by matrix element (**MADGRAPH**, in our case) calculators with the additional parton explicitly listed in the final state, i.e.  $pp \rightarrow t\bar{t}X+j$ , while soft emissions are handled by the parton shower. Events in the ‘wrong’ region, e.g. a partonic event with  $p_T$  below the division, are removed to avoid double counting. Since matching uses tree-level matrix elements only, it is less expensive than the full NLO treatment—both in terms of CPU-time of the computation and also in terms of the fraction of events generated with negative weights—and it avoids complications associated with the lack of electroweak loop corrections. Certainly, matched tree-level calculations will lack NLO precision, but for many cases, the loss of precision is a reasonable trade-off in the face of the practical difficulties of NLO Monte Carlo generation.

Matching does involve some complications. In current Monte Carlo generators, SMEFT operators can be included in the matrix element portion of the calculation, but they cannot be added (easily) into the parton shower. Thus, it would appear that matching with SMEFT always causes a mismatch, as the events deleted from the matrix element are not replaced by anything. While a reasonable worry, we will show that the impact of this mismatch is small and, at least when focusing on  $t\bar{t}X+j$  processes, confined to a single operator.

The goal of this paper is to show that matched calculations capture the relevant physics of extra partons in  $t\bar{t}X$  final states without introducing uncertainties beyond the usual ones associated with LO compared to NLO QCD, such as the uncertainty from the choice of scales for the matching, factorization/renormalization, and ISR/FSR. Because we know that even in the SM case, the predicted  $t\bar{t}X$  inclusive cross section receives non-negligible corrections at NLO compared to LO, we will examine the ratio of the SMEFT cross section to the SM value  $\sigma_{\text{SMEFT}}(t\bar{t}X+j)/\sigma_{\text{SM}}(t\bar{t}X+j)$  – the ‘relative K-factor’, which we will denote as  $\mu$ . We demonstrate the validity of this approach by showing agreement between the cross section ratios calculated for NLO and matched LO. Looking at this ratio has the additional benefit that to determine the overall rate in  $t\bar{t}X$  for some set of SMEFT operator coefficients, starting from a LO calculation, one simply needs to multiply by the (NLO) SM cross section. Once we have established the validity of using matched calculations, we compare the sensitivity of  $t\bar{t}h$ ,  $t\bar{t}W$ ,  $t\bar{t}Z$  matched inclusive cross sections to SMEFT coefficients with and without extra radiation. The extra radiation can come from a SM vertex, or from an additional quark/gluon

emitted from a higher dimensional operator.<sup>2</sup>

The layout of the rest of this paper is as follows. In section 2 we describe the SMEFT framework and the subset of operators that can have a potential impact on  $t\bar{t}X$  and  $t\bar{t}X+j$  processes. Section 3 is devoted to justifying the use of matching with SMEFT for theory calculations and Monte Carlo studies. The following section, section 4 contains our results—plots and tables displaying the impact of SMEFT operators on inclusive  $t\bar{t}X$  and  $t\bar{t}X+j$  processes. We focus on operators whose impact changes dramatically from +0 jets to +1 jet. In several cases, the NLO calculation for the operators we examine is challenging to obtain because too many electroweak vertices are needed. In section 5, we discuss uncertainties and compare their size to the size of the impact from adding additional jets into the calculation. Finally, in section 6, we conclude.

## 2 The SMEFT framework

The most common basis for characterizing SMEFT effects is the so-called Warsaw basis [3]. At dimension-6, there are 59 operators if one assumes complete fermion flavor universality, or 2499 for flavor anarchy [26, 27]<sup>3</sup>. A flavor symmetry assumption that interpolates between these two extremes is to assume the SMEFT operators are invariant under flavor  $U(2)_Q \times U(2)_u \times U(2)_d$ . Under this assumption, first and second generation quarks have universal SMEFT deviations, while operators involving third generation quarks may be different. This pattern allows the top-quark (or, potentially, the whole third generation) to have sizable SMEFT effects without running afoul of strong flavor constraints from e.g. kaon physics. Motivated by the possibility of larger third generation effects, the set of operators containing two or more third generation fermions was compiled in Ref. [21]. In this paper, we continue the mindset of Ref. [21], adopting the reduced set of operators and using the UFO model dim6TopEFT for all **MADGRAPH** purposes. In the dim6TopEFT setup, dimension-6 operators are suppressed by a scale  $\Lambda$  that is conventionally fixed to 1 TeV. This choice renders the Wilson coefficients dimensionless. Additionally, the CKM matrix is assumed to be a unit matrix, the masses of  $u, d, s, c, e, \mu$  fermions are set to zero, and the unitary gauge is used by default.

While the dim6TopEFT model contains 33 operators<sup>4</sup> we will further narrow our focus to operators that can have potential tree-level impact on  $t\bar{t}X$  and  $t\bar{t}X+j$  processes with  $h/W^\pm/Z$  on-shell. Our goal, after all, is to explore how well matching captures the effects of additional radiation in  $t\bar{t}X$  process, rather than perform a complete SMEFT analysis, and a smaller set of

---

<sup>2</sup>There are two important aspects of the UFO model implementation in **MADGRAPH** that need to be accounted for correctly, as outlined in section 3.

<sup>3</sup>At dimension-5, only one operator structure is allowed and it violates lepton number. Therefore, most SMEFT analyses assume lepton number is a good symmetry, making dimension-6 the lowest mass dimension beyond the SM.

<sup>4</sup>This counting refers to the ‘baseline’ model in Ref. [21]. Several stricter flavor symmetry assumptions are possible, such as a ‘top-philic’ scenario. See Ref. [21] for more details.

operators makes this more tractable. This choice limits the number of operators<sup>5</sup> to 9, which are listed below in Table 1. As some of the operators in Table 1 are not self-Hermitian, they may have complex Wilson coefficients. However as imaginary Wilson coefficients inevitably lead to CP violation and are therefore strongly constrained, we consider real coefficients only. Importantly, we include operators that impact  $t\bar{t}X$ ,  $t\bar{t}X+j$  via both QCD and electroweak interactions.

**Table 1.** The operators and corresponding Wilson coefficients considered in this paper. Left-handed fermion doublets are denoted by  $q$ , and right-handed fermion singlets by  $t$  and  $b$ ; the Higgs doublet is denoted by  $\varphi$ . The antisymmetric  $SU(2)$  tensor is denoted by  $\varepsilon \equiv i\tau^2$ . Furthermore,  $\tilde{\varphi} = \varepsilon\varphi^*$ ,  $(\varphi^\dagger \overleftrightarrow{D}_\mu \varphi) \equiv \varphi^\dagger (iD_\mu \varphi) - (iD_\mu \varphi^\dagger)\varphi$ , and  $(\varphi^\dagger \overleftrightarrow{D}_\mu^I \varphi) \equiv \varphi^\dagger \tau^I (iD_\mu \varphi) - (iD_\mu \varphi^\dagger)\tau^I \varphi$ , where  $\tau^I$  are the Pauli matrices. Finally,  $T^A \equiv \lambda^A/2$ , where  $\lambda^A$  are Gell-Mann matrices. The covariant derivative is  $D_\mu = \partial_\mu - ig_s \frac{1}{2} \lambda^A G_\mu^A - ig \frac{1}{2} \tau^I W_\mu^I - ig' Y B_\mu$ , and  $G_{\mu\nu}^A = \partial_\mu G_\nu^A - \partial_\nu G_\mu^A + g_s f^{ABC} G_\mu^B G_\nu^C$ ,  $W_{\mu\nu}^I = \partial_\mu W_\nu^I - \partial_\nu W_\mu^I + g \epsilon_{IJK} W_\mu^J W_\nu^K$ , and  $B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$  are the gauge field strength tensors. The abbreviations  $s_W$  and  $c_W$  denote the sine and cosine of the weak mixing angle (in the unitary gauge). More details about the operators can be found in Ref. [21].

Operator	Definition	Wilson Coefficient
$\mathcal{O}_{t\varphi}$	$\bar{q}t\tilde{\varphi}(\varphi^\dagger\varphi)$	$c_{t\varphi}$
$\mathcal{O}_{\varphi q}^1$	$(\varphi^\dagger \overleftrightarrow{D}_\mu \varphi)(\bar{q}\gamma^\mu q)$	$c_{\varphi Q}^- + c_{\varphi Q}^3$
$\mathcal{O}_{\varphi q}^3$	$(\varphi^\dagger \overleftrightarrow{D}_\mu^I \varphi)(\bar{q}\gamma^\mu \tau^I q)$	$c_{\varphi Q}^3$
$\mathcal{O}_{\varphi t}$	$(\varphi^\dagger \overleftrightarrow{D}_\mu \varphi)(\bar{t}\gamma^\mu t)$	$c_{\varphi t}$
$\mathcal{O}_{\varphi tb}$	$(\tilde{\varphi}^\dagger iD_\mu \varphi)(\bar{t}\gamma^\mu b)$	$c_{\varphi tb}$
$\mathcal{O}_{tW}$	$(\bar{q}\sigma^{\mu\nu}\tau^I t)\tilde{\varphi}W_{\mu\nu}^I$	$c_{tW}$
$\mathcal{O}_{bW}$	$(\bar{q}\sigma^{\mu\nu}\tau^I b)\varphi W_{\mu\nu}^I$	$c_{bW}$
$\mathcal{O}_{tB}$	$(\bar{q}\sigma^{\mu\nu}t)\tilde{\varphi}B_{\mu\nu}$	$(c_W c_{tW} - c_{tZ})/s_W$
$\mathcal{O}_{tG}$	$(\bar{q}\sigma^{\mu\nu}T^A t)\tilde{\varphi}G_{\mu\nu}^A$	$c_{tG}$

We would like to draw attention to several notational choices regarding the operators and Wilson coefficients outlined in Table 1. First, we note that following the convention in Ref. [21], the  $c_{\varphi Q}^3$  and  $c_{tW}$  parameters appear in multiple Wilson coefficients. This choice simplifies the form of  $t$ - $t$ - $Z$  (and  $t$ - $t$ - $Z$ - $h$ ) interactions, so that contributions to  $\bar{t}_L\gamma_\mu t_L Z^\mu$  are set by  $c_{\varphi Q}^-$  alone, contributions to  $\bar{t}_L\gamma_\mu t_R Z^{\mu\nu}$  are set by  $c_{tZ}$  alone, etc. For the rest of this paper, we will refer to each individual parameter as a Wilson coefficient; for example, rather than denoting  $c_{\varphi Q}^3 + c_{\varphi Q}^-$  as the Wilson coefficient of  $\mathcal{O}_{\varphi q}^1$ , we will refer to  $c_{\varphi Q}^3$  and  $c_{\varphi Q}^-$  as two distinct Wilson coefficients, where the Wilson coefficient  $c_{\varphi Q}^3$  is associated with both  $\mathcal{O}_{\varphi q}^1$  and  $\mathcal{O}_{\varphi q}^3$ . Because multiple operators can involve the same Wilson coefficient, we will describe EFT vertices in terms of the Wilson coefficient associated with the vertex (instead of the operator associated with the Wilson coefficient) to avoid potential ambiguities. Finally, to maintain consistency, we will also use the language of Wilson coefficients when describing the

<sup>5</sup>In particular, the requirement of an on-shell  $h/W^\pm/Z$  removes all four fermion operators.

effects of EFT operators. For example, to discuss how the Lagrangian term  $\frac{c_{\varphi t}}{\Lambda^2}\mathcal{O}_{\varphi t}$  affects a particular process, we will refer to the effect of  $c_{\varphi t}$  on the given process.

### 3 Jet matching/merging for EFT samples

When considering processes containing an extra parton in the final state, a matching procedure is required to avoid double counting between matrix element (ME) and parton shower (PS) descriptions of phase space. When the ME is matched to a PS, only widely-separated radiation above a certain matching scale  $Q$  will be modeled by the ME and therefore include the EFT effects, while radiation below the scale  $Q$  is cut off, and not included in the ME description. Radiation below the scale  $Q$ , namely a soft and colinear regime, are generated by the PS which does not include the EFT vertices. Thus, one may worry that a contribution, namely the soft and colinear contribution from EFT vertices, may be missed. If the soft and colinear contributions from EFT vertices are small, they can safely be neglected. However, if they are large, then the matching procedure cannot be used without adding EFT effects to the parton shower.

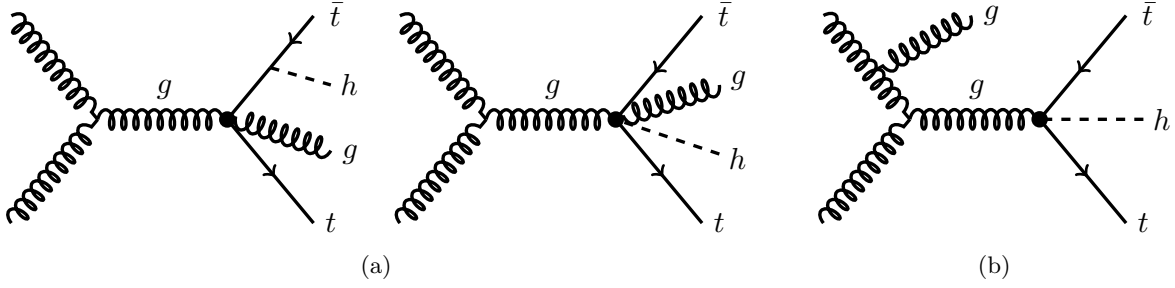
Although there are multiple different ways to implement matching, for these studies we use the  $k_T$ -jet version of the MLM matching scheme, an event-rejection based approach that matches partons generated by **MADGRAPH** to jets clustered by **PYTHIA** [23]. First, the final-state partons in the event produced by **MADGRAPH** are clustered according to the  $k_T$  algorithm, where the  $k_T$  value is required to be above a specified cutoff scale referred to as the **xqcut**. The event is then passed from **MADGRAPH** to **PYTHIA** for parton showering. After showering (but before hadronization), **PYTHIA** clusters the final-state objects using the  $k_T$  algorithm with a cutoff scale called **qCut**; using **qCut** as the maximal  $k_T$  distance between jets and partons, the clustered jets are matched to the ME partons. The event is saved if all jets are successfully matched to partons. The event is otherwise discarded, except in the highest jet multiplicity sample, where extra jets (with  $k_T$  less than the softest ME parton) are permitted since there is no danger of double counting in these events. To avoid missing a region of phase space, **xqcut** is chosen to be less than **qCut**.

In what follows, we explore the validity of matching applied to generation of EFT samples from two perspectives. First, we explore the theoretical considerations that come into play as applied to the processes and operators of interest in this paper. Then, we check the results of applying matching to specific samples generated via **MADGRAPH** to see if any signs of missing contributions are visible.

#### 3.1 Theoretical justification

As mentioned above, a major concern with matching as applied to EFT samples is that SMEFT operators are included in the matrix element, but they are not implemented in the parton shower. In this section, we will demonstrate that the impact of this mismatch is small, as EFT contributions for the subset of Warsaw basis operators we are interested in (see section 2) to soft and colinear regime are generically suppressed.





**Figure 1.** The  $t\bar{t}h+j$  final state when the gluon is emitted from (a) the  $g-g-t-t-h$  and  $g-g-t-t$  vertices generated by the  $\mathcal{O}_{tG}$  operator, and (b) the ordinary QCD vertex.

Extra QCD radiation requires gluons, so can only arise in operators with gluon field strengths or derivatives of quark fields. Working within the Warsaw basis, where derivatives have been systematically removed in favor of operators with more fields, and focusing on CP-even operators<sup>6</sup>, our options are limited to  $\mathcal{O}_G = f_{ABC} G_\mu^{A,\nu} G_\nu^{B,\rho} G_\rho^{C,\mu}$ ,  $\mathcal{O}_{\phi G} = \phi^\dagger \phi G^{A,\mu\nu} G_{\mu\nu}^A$ ,  $\mathcal{O}_{dG}$  and  $\mathcal{O}_{uG}$ . If we limit ourselves to the operators that are particularly important for  $t\bar{t}X$  observables, as laid out by Ref. [8, 21], the only operator that remains is  $\mathcal{O}_{tG}$ <sup>7</sup>.

$$\frac{c_{tG}}{\Lambda^2} \mathcal{O}_{tG} = \frac{g_s c_{tG}}{\Lambda^2} (\bar{q} \sigma^{\mu\nu} T^a t) \tilde{\phi} G_{\mu\nu}^a, \quad (3.1)$$

Notice the explicit factor of  $g_s$  in  $\mathcal{O}_{tG}$ . This factor is necessary in order for **MADGRAPH** to identify  $\mathcal{O}_{tG}$  as a ‘QCD-type’ operator and include it correctly.

To address the concern of omitting  $\mathcal{O}_{tG}$  from the parton shower, we need to understand the circumstances under which  $\mathcal{O}_{tG}$  leads to a soft or collinear gluon. There are two scenarios where we can imagine  $\mathcal{O}_{tG}$  may be problematic, depending on whether we extract one or two gluons from the field strength. If we extract two gluons, we have new  $g-g-t-t-h$  and  $g-g-t-t$  QCD vertices. Connecting one of the gluons to the initial partons, the second gluon may be soft or collinear with a final state  $t$  or  $\bar{t}$ . On the other hand, if we set the Higgs to its vacuum expectation value ( $v$ ) and pull one gluon out of the field strength, we are left with a new (non-SM) three-point interaction between two top quarks and a gluon. Including this vertex on any top quark line, we have another potential source of soft/collinear gluon radiation. In the next two sections we explore each of these scenarios in detail.

### 3.1.1 $g-g-t-t-h$ and $g-g-t-t$ vertices: a case study on $t\bar{t}h+j$ final state

In this subsection, we study the  $g-g-t-t-h$  and  $g-g-t-t$  vertices within  $\mathcal{O}_{tG}$ . While  $g-g-t-t$  can impact any of the  $t\bar{t}X+j$  processes we are interested in,  $g-g-t-t-h$  is specific<sup>8</sup> to  $t\bar{t}h+j$ ,

<sup>6</sup>CP-odd operators will not interfere with the SM and therefore have no effect at  $1/\Lambda^2$ , so we ignore them.

<sup>7</sup>The impact of  $\mathcal{O}_G$  in matching has been addressed in the literature in Ref. [25]. The other operators,  $\mathcal{O}_{\phi G}$  and  $\mathcal{O}_{dG}$  play little role in  $t\bar{t}X$  and are better constrained by Higgs and  $t\bar{t}b\bar{b}$  observables respectively [8, 28]. Additionally, due to the chirality structure of  $\mathcal{O}_{dG}$ , it is reasonable to assume  $c_{dG} \propto y_b \ll 1$ , further suppressing any effects.

<sup>8</sup>Provided we ignore the Yukawa couplings of first and second generation quarks.



therefore we will use on  $t\bar{t}h+j$  as a test case; some representative Feynman diagrams for  $t\bar{t}h+j$  illustrating how  $\mathcal{O}_{tG}$  vertices can enter are shown in Fig. 1(a). As mentioned above, in the matching procedure radiation softer than the scale  $Q$  will be cut out of the matrix element calculation and left to the parton shower. However, as  $\mathcal{O}_{tG}$  effects are absent in the parton shower, the matched calculation is mis-modeling some part of the low  $Q$  phase space. To quantify how severe an issue this mis-modeling is, we need to understand the momentum distribution of gluons from  $g-g-t-t-h$  and  $g-g-t-t$  vertices.

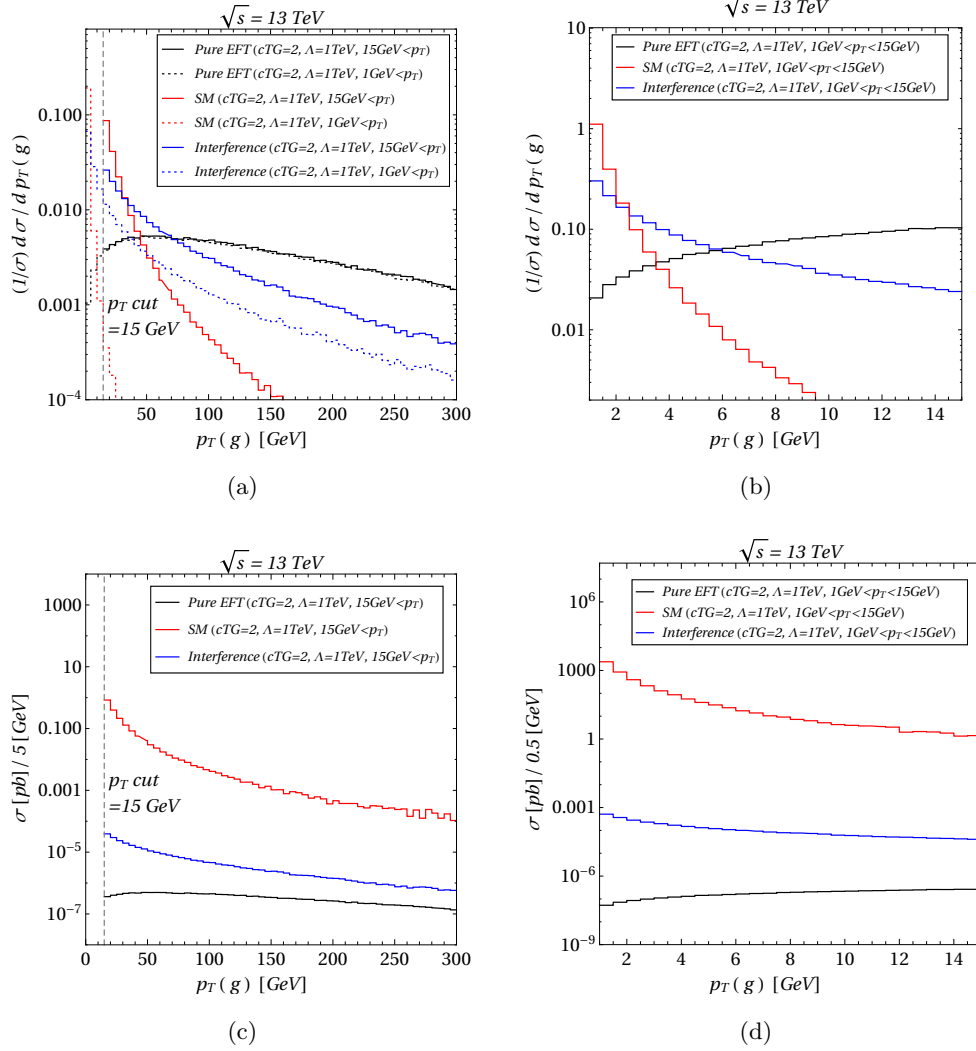
Generating  $t\bar{t}h+j$  at parton level within the dim6TopEFT model [21] using MADGRAPH [12], the normalized  $p_T$  spectrum of the additional gluon is shown below in Fig. 2(a) and 2(b). The only non-zero Wilson coefficient is  $c_{tG}$ , the coefficient of  $\mathcal{O}_{tG}$ . To focus on the effects of the  $g-g-t-t-h$  and  $g-g-t-t$  vertices of  $\mathcal{O}_{tG}$ , we split up the calculation into two pieces: i,) the sum of the SM contribution and all  $\mathcal{O}_{tG}$  effects *except* the  $g-g-t-t-h$  and  $g-g-t-t$  vertices (e.g from the diagram illustrated in Fig. 1(b)), and ii,) contributions from the  $g-g-t-t-h$  and  $g-g-t-t$  vertices; technically, we achieve this splitting by adjusting the coupling order of the vertices within  $\mathcal{O}_{tG}$ . The differential cross section from the SM and non- $g-g-t-t-h$  and non- $g-g-t-t$  parts of  $\mathcal{O}_{tG}$  is shown in red, the interference of this piece with diagrams including the  $g-g-t-t-h$  and  $g-g-t-t$  vertices is shown in blue, and the quadratic contribution from  $g-g-t-t-h$  and  $g-g-t-t$  vertices is shown in black. The events were generated assuming 13 TeV center of mass energy and using default parton distribution functions NNPDF2.3QED [29]. As benchmark values, we choose  $c_{tG} = 2$  and  $\Lambda = 1$  TeV.

In the left panel, Fig. 2(a), we compare the three contributions to  $\frac{1}{\sigma}d\sigma/dp_{T,g}$  using minimum cut of  $p_T > 15$  GeV (solid lines) and  $p_T > 1$  GeV (dashed lines) for the radiated gluon. For the purely EFT contribution (black line), the hard gluon is emitted with  $p_T$  ranging from tens of GeV to even hundreds of GeV. In contrast, the gluon from the SM QCD vertex (solid red) is substantially softer with  $p_T$  of a few tens of GeV. The gluon from the interference (solid blue) between two diagrams, Figure 1(a) and 1(b), stands in the middle. In order to quantify the relative size of each contribution, in Figure 2(c) and 2(d),  $p_T$  distributions are normalized to each individual cross section.

Comparing the red, blue, and black lines, if we lower the selection cut down to  $p_T > 1$  GeV, relatively negligible portions of the interference and purely EFT gluons show up, while the SM piece blows up. For instance, a relative size of the interference piece in  $p_T$  range of [1, 15] GeV is only an  $\mathcal{O}(10^{-4}\%)$  correction to the SM contribution. This numerical result provides a support for the claim that the mismatch caused by the absence of EFT  $g-g-t-t-h$  and  $g-g-t-t$  vertices in the soft and collinear regime is negligibly small.

### 3.1.2 An anomalous $t-t-g$ coupling

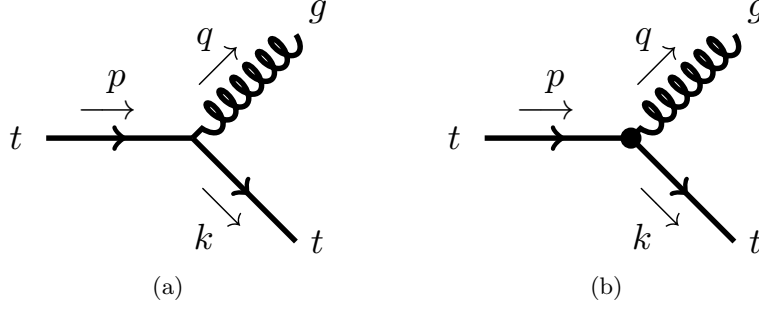
Another avenue to study the EFT radiations is the anomalous  $t-t-g$  coupling induced by the  $\mathcal{O}_{tG}$  operator as shown in Figure 3(b). Unlike the SM coupling (Figure 3(a)), this vertex flips the chirality of the top quark, and can contribute to any  $pp \rightarrow t\bar{t}X$ ,  $X = h/W^\pm/Z$  processes involving an extra parton. Since the parton shower does not include the anomalous  $t-t-g$



**Figure 2.** (a) Unit normalized  $p_T$  distributions of a gluon with  $p_T > 15$  GeV from the EFT vertex (solid black), the SM QCD vertex (solid red), and the interference (solid blue). The  $p_T$  distributions of a gluon with a lower minimum cut  $p_T > 1$  GeV are shown in dotted lines. (b) Normalized  $p_T$  distributions of a gluon in a lower  $p_T$  range,  $1 < p_T < 15$  GeV, from the EFT vertex (solid black), the SM QCD vertex (solid red), and the interference (solid blue). In (c) and (d),  $p_T$  distributions are normalized to each individual cross section.

vertex, the information about the EFT below the matching scale  $Q$  is missing, resulting in an mismatch.

To scrutinize the properties of gluons from the EFT (Figure 3(b)) vertex, we can analytically compute the  $q \rightarrow q + g$  splitting function [30–33]; Feynman rules for the  $t$ - $t$ - $g$  couplings are summarized in Appendix B. We consider the case where two final particles are almost



**Figure 3.** The splitting of gluon from the (a) SM  $t$ - $t$ - $g$  vertex, and (b) anomalous  $t$ - $t$ - $g$  vertex generated by the  $\mathcal{O}_{tG}$  operator.

collinear, with a small relative transverse momentum  $p_T$ . We choose the coordinate such that the initial top quark is aligned to the  $z$ -axis, and all particles are on-shell. The four-momenta of three particles can be written as:

$$\begin{aligned} p &= (E, 0, 0, p), \\ q &= (zE, p_T, 0, \sqrt{z^2 E^2 - p_T^2}), \\ k &= ((1-z)E, -p_T, 0, \sqrt{(1-z)^2 E^2 - m_t^2 - p_T^2}), \end{aligned} \quad (3.2)$$

where  $z = E_q/E_p$  is a dimensionless energy-sharing variable, and  $m_t = 173$  GeV is the top quark mass. The splitting function is defined as

$$\frac{d\mathcal{P}}{dz dp_T^2} = \frac{z\bar{z}}{16\pi^2 \tilde{p}_T^4} \left( \frac{1}{2} \frac{1}{3} \sum |\mathcal{M}|^2 \right), \quad (3.3)$$

where  $\tilde{p}_T^2 \equiv p_T^2 + z^2 m^2$  and  $\bar{z} \equiv 1 - z$ , and  $\mathcal{M}$  denotes a splitting amplitude shown in Figure 3. The sum inside the bracket runs over all polarizations, spins, and colors. The squared amplitude is averaged over initial helicities and colors<sup>9</sup>. Summing the effect of the SM and  $\mathcal{O}_{tG}$  vertices and taking the limit  $E \gg p_T, m_t$ , the splitting functions are:

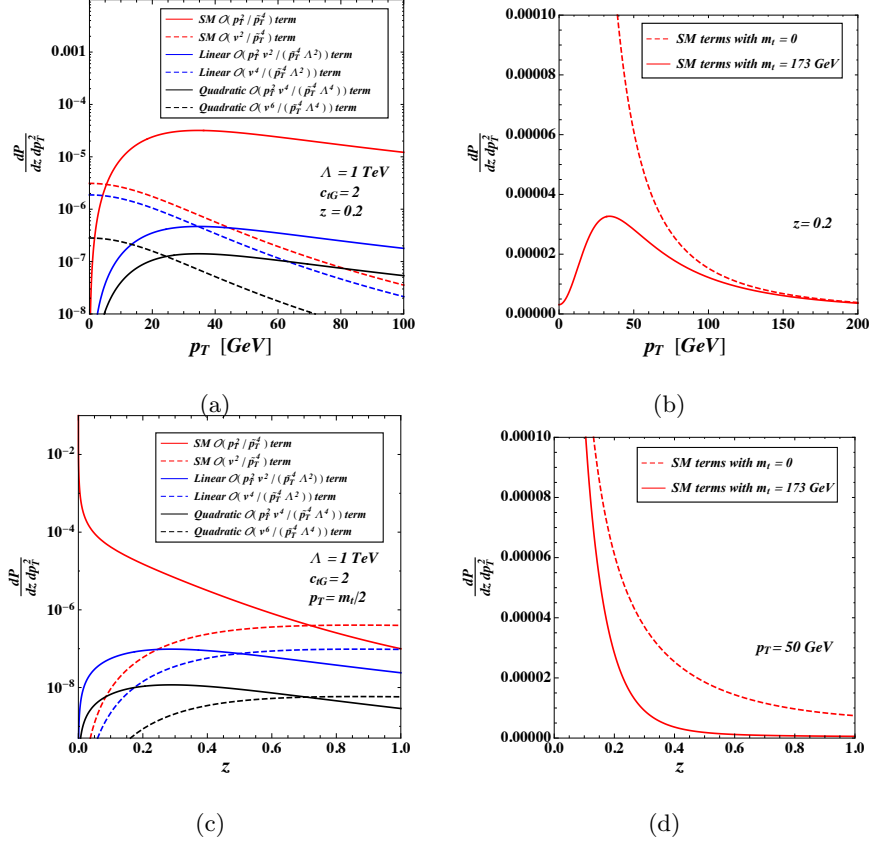
$$\frac{d\mathcal{P}_{SM}}{dz dp_T^2} = \frac{2\alpha_s}{3\pi} \left( \frac{p_T^2}{\tilde{p}_T^4} \right) \frac{(1 + \bar{z}^2)}{z} + \frac{2\alpha_s}{3\pi} \left( \frac{m_t^2}{\tilde{p}_T^4} \right) z^3, \quad (3.4)$$

$$\frac{d\mathcal{P}_{linear}}{dz dp_T^2} = \frac{4\sqrt{2}\alpha_s c_{tG}}{3\pi} \left( \frac{p_T^2 v m_t}{\tilde{p}_T^4 \Lambda^2} \right) z + \frac{4\sqrt{2}\alpha_s c_{tG}}{3\pi} \left( \frac{v m_t^3}{\tilde{p}_T^4 \Lambda^2} \right) z^3, \quad (3.5)$$

$$\frac{d\mathcal{P}_{quadratic}}{dz dp_T^2} = \frac{8\alpha_s c_{tG}^2}{3\pi} \left( \frac{p_T^2 v^2 m_t^2}{\tilde{p}_T^4 \Lambda^4} \right) z + \frac{4\alpha_s c_{tG}^2}{3\pi} \left( \frac{v^2 m_t^4}{\tilde{p}_T^4 \Lambda^4} \right) z^3, \quad (3.6)$$

where we have separated the result into the SM piece, the interference term, and the term quadratic in  $c_{tG}$ .

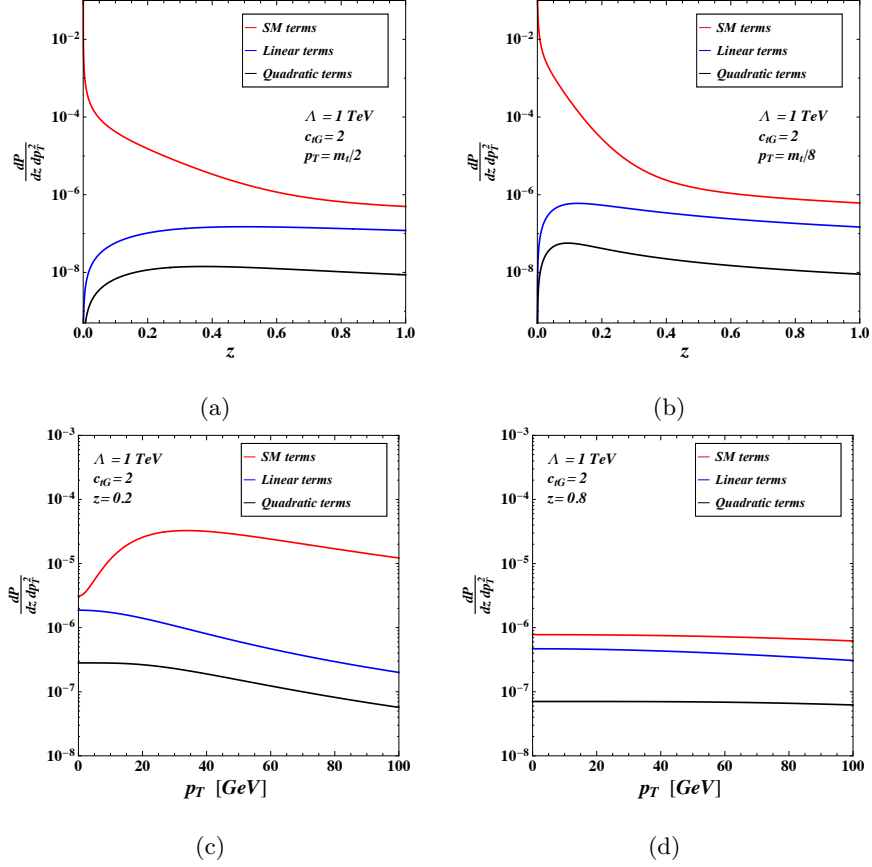
<sup>9</sup>For the detailed expressions of spinors and polarization vectors, see the Refs.[30, 34].



**Figure 4.** Splitting functions for the SM, the interference, and the quadratic EFT pieces in Eq.(3.4,3.5,3.6) respectively. We show the components of each splitting function in terms of (a)  $p_T$  with fixed  $z = 0.2$  and (c)  $z$  with fixed  $p_T = m_t/2$ . To see the mass effect, the SM splitting function is shown with and without including the mass, in terms of (b)  $p_T$  with fixed  $z = 0.2$  and (d)  $z$  with fixed  $p_T = 50 \text{ GeV}$ .

The first term of the SM splitting function in Eq.(3.4) scales like  $\mathcal{O}(p_T^2/\tilde{p}_T^4)$  and is relatively enhanced at high  $p_T$  with respect to the second term which scales like  $\mathcal{O}(v^2/\tilde{p}_T^4)$ . The second term, however, is more enhanced at small  $p_T$ , and hence named as the ultra-collinear splitting [30, 35]. This can be seen in Figure 4(a) that at small  $p_T$  the ultra-collinear term (dashed red) is enhanced, whereas the first term (solid red) is relatively suppressed. At high  $p_T$ , on the other hand, the first term dominates over a wide range of  $p_T$ . Note that the mass-dependent factor  $\tilde{p}_T^4$  in the denominator regulates a collinear divergence at small  $p_T$ . As shown in Figure 4(b), in the massless limit ( $m_t = 0$ ), the SM splitting function (dashed red) diverges as  $p_T \rightarrow 0$  with fixed value of  $z = 0.2$ . On the other hand, if we include the mass term, the splitting function (solid red) is regulated at the infrared region.

The splitting function for the interference piece in Eq.(3.5) receives an overall suppression



**Figure 5.** Splitting functions for the SM, the interference, and the quadratic EFT pieces in Figure 3. The splitting functions in terms of  $z$  variable for fixed (a)  $p_T = m_t/2$  and (b)  $p_T = m_t/8$ . The splitting functions in terms of  $p_T$  for fixed (c)  $z = 0.2$  and (d)  $z = 0.8$ . As benchmark values, we choose  $c_{tG} = 2$  and  $\Lambda = 1 \text{ TeV}$ .

factor of  $\mathcal{O}(v^2/\Lambda^2)$ . The first term (solid blue) scales as  $\mathcal{O}(p_T^2 v^2/(\tilde{p}_T^4 \Lambda^2))$ , and dominates over the second term at high  $p_T$  as shown in Figure 4(a). The second term (dashed blue) is of an order of  $\mathcal{O}(v^4/(\tilde{p}_T^4 \Lambda^2))$  that shows the similar ultra-collinear behavior as it is enhanced at small  $p_T$ . Although the SM and interference splitting functions behave similarly in terms of  $p_T$  up to the overall suppression factor, there is a clear disparity between them in terms of  $z$ . Note that the first terms of Eq.(3.4) and Eq.(3.5) have a different  $z$ -dependence. Their distinctive  $z$ -dependences are plotted in Figure 4(c) where the interference piece (solid blue) is highly suppressed as  $z \rightarrow 0$  while the SM term (solid red) is dramatically enhanced. The splitting function for the pure EFT piece in Eq.(3.6) is qualitatively similar to the interference one, up to the overall suppression factor of  $\mathcal{O}(v^4/\Lambda^4)$ , and hence highly suppressed.

Figure 5 shows total splitting functions for the SM, the interference, and the quadratic

EFT terms. As can be seen in the upper panels, the amount of the interference and the pure EFT contribution is highly suppressed near the  $z = 0$  region, while the SM contribution is much more enhanced. For  $z = 0.05$ , the size of the SM splitting function is  $\sim 1200$  ( $4200$ ) times larger than the size of the interference (pure EFT) splitting function. This means that it is less likely that the soft gluon is emitted due to the interference and the pure EFT portions, while a majority of soft gluon is generated from the SM vertex. As we move away from the  $z = 0$  region, the both interference and the pure EFT splitting functions rapidly grow, closing a wide gap with the SM. For  $z = 0.8$ , the SM splitting function is only  $\sim 2$  ( $10$ ) times larger than the interference (pure EFT) one.

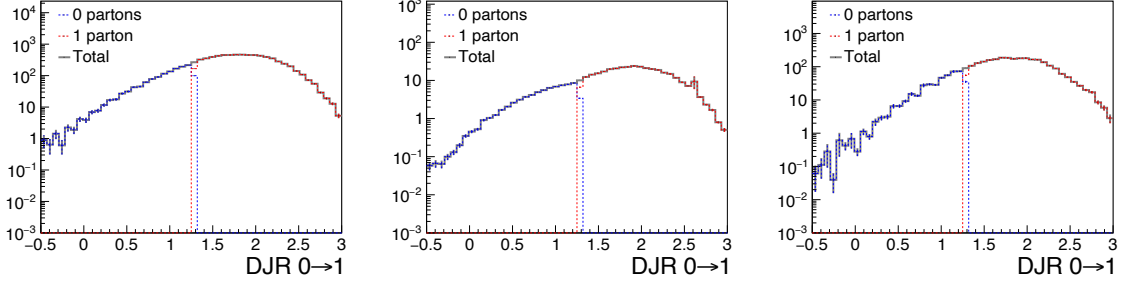
### 3.2 Monte Carlo validation of matching

The prior section demonstrated that the impact of neglecting EFT contributions in the parton shower, at least for the processes of interest in this study, is small. To quantify the impact of applying matching, we look at the concrete case of specific physics processes generated with non-zero values for the Wilson coefficients. We will use the differential jet rate (DJR) to characterize whether the matching has introduced any discontinuities [24, 36]. For the  $k_T$  algorithm, the DJR histogram represents the distribution of  $k_T$  values for which an  $n$  jet event transitions to an  $n + 1$  jet event. A smooth transition between the  $n$  and  $n + 1$  curves indicates that the chosen matching scales have allowed **MADGRAPH** and **PYTHIA** to work together to smoothly populate the overlapping region of phase space.

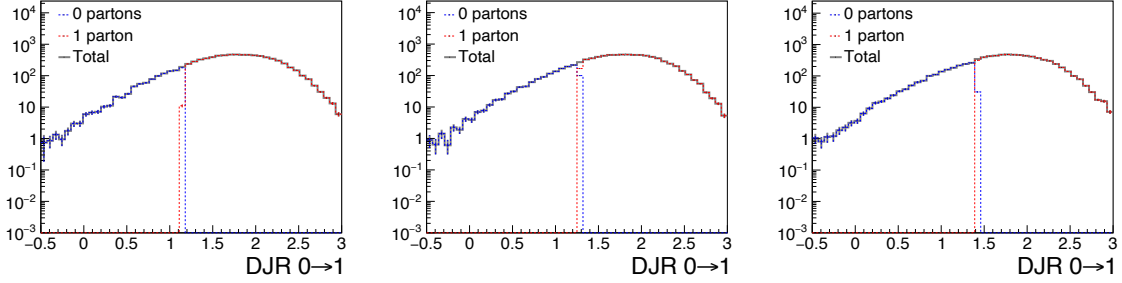
The matched samples used in this section are generated with a model based on the one described in Ref. [21], with an extra factor of  $g_s$  applied to the  $c_{tG}$  vertices, as described in section 3.1. The five-flavor scheme is used, so the mass of the  $b$  quark is set to zero. We use the PDF set NNPDF3.1 [37] and the default dynamical scale choice for the renormalization and factorization scales; we do not make any parton level cuts. The samples are generated with an **xqcut** of 10 and a **qCut** of 19.

Figure 6 shows DJR plots for the three processes considered in this study. For each process, the transition from the 0 parton line to the 1 parton line is smooth, indicating that the matching is being handled appropriately. If the **xqcut** and **qCut** are varied around their nominal values of **xqcut** = 10, **qCut** = 19, the DJR plots should also remain smooth. Figure 7 and Figure 8 show that the  $t\bar{t}h$  DJR plots remain consistently smooth as we vary the **qCut** and **xqcut**, respectively.

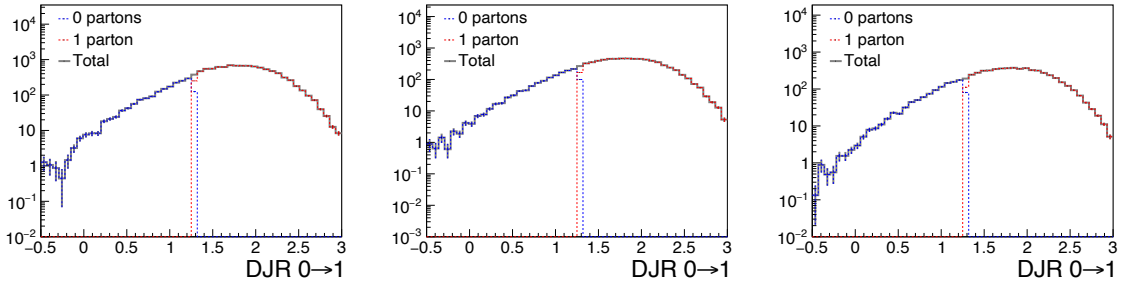
In addition to being useful for identifying problems associated with matching and validating matching parameters, DJR plots can also be useful for discovering problems with simulated samples that are unrelated to matching. For example, when we began this study, we were using an older version of the model described in Ref. [21] that did not include five particle vertices. This is important for  $t\bar{t}h$ , since the  $\mathcal{O}_{tG}$  operator gives rise to a five particle vertex involving two top quarks, two gluons, and a Higgs boson. Before the inclusion of this missing vertex, the  $t\bar{t}h$  DJR plot showed a significant discontinuity when  $c_{tG}$  was set to non-zero values. In fact, this discontinuity is what initially alerted us to the fact the five-particle  $c_{tG}$  vertex was missing from the model file. After the missing  $c_{tG}$  five-particle



**Figure 6.** DJR plots for  $x_{\text{qcut}} = 10$ ,  $q_{\text{Cut}} = 19$  for  $t\bar{t}h$  (left),  $t\bar{t}W$  (middle) and  $t\bar{t}Z$  (right). Here, all 9 Wilson coefficients considered in this study are set to non-SM values. The x axis shows the log base 10 of the scale at which an  $n$  jet event transitions into an  $n+1$  jet event. The line labeled “0 partons” refers to the contribution from the parton shower, while the line labeled “1 parton” refers to the contribution from the matrix element. The line labeled “Total” is the sum of the two contributions.



**Figure 7.**  $t\bar{t}h$  DJR plots for  $x_{\text{qcut}} = 10$ ,  $q_{\text{Cut}} = 15$  (left), 19 (middle) and 25 (right). Here all Wilson coefficients are set to non-SM values. The x axis and the lines in the plots are the same as described in Figure 6.

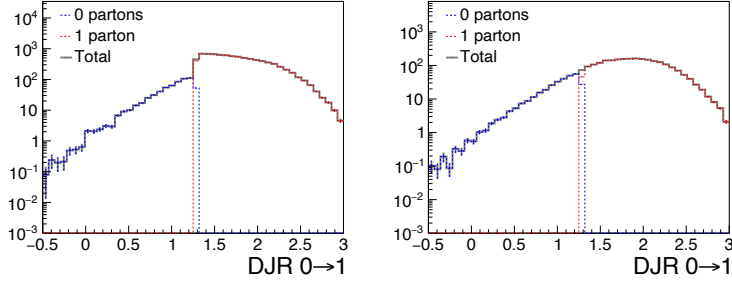


**Figure 8.**  $t\bar{t}h$  DJR plots for  $q_{\text{Cut}} = 19$ ,  $x_{\text{qcut}} = 5$  (left), 10 (middle) and 15 (right). Here all Wilson coefficients are set to non-SM values. The x axis and the lines in the plots are the same as described in Figure 6.

vertex was added to the model, the  $t\bar{t}h$  DJR plots were found to be smooth even when  $c_{tG}$  was set to non-zero values. The DJR plots produced with and without the five-particle  $c_{tG}$



vertex included in the model are shown in Figure 9.



**Figure 9.** DJR plots for  $t\bar{t}h$  with the  $g\text{-}g\text{-}t\text{-}t\text{-}h$  vertex missing (left) and with the vertex included (right). In both plots, the value of  $c_{tG}$  is set to 2, and all other Wilson coefficients are set to 0. The x axis and the lines in the plots are the same as described in Figure 6.

## 4 Results

Having established that including extra partons with matching in an LO calculation provides a valid approach to approximating the impact of higher-order QCD corrections for the operators considered here, we will now look at the results obtained with a matched LO calculation. In particular, we are interested in evaluating whether there are any cases where a matched LO calculation produces different results from an LO calculation that does not include an additional parton. After all, if the matched calculation is always essentially equivalent to the simple LO calculation, then it might not be advantageous to introduce the complexities of the matching procedure. On the other hand, if including an additional parton results in significant differences in some of the EFT predictions, then it will be important to ensure that these contributions are always included.

The LO matched samples used in this section are the same as those described in section 3.2, and the samples produced without an additional parton are generated similarly, but with the matching turned off. We will also look at a few selected processes and operators at NLO. The NLO samples are generated with the SMEFT NLO model [15]; consistent with the LO samples, we use the PDF set NNPDF3.1 [37], the default dynamical scale choice for the renormalization and factorization scales, and we do not make any parton level cuts.

The range over which each Wilson coefficient is varied is taken from the marginalized limits presented in Ref. [8]. We make this choice to ensure that we focus on differences in the predictions in an experimentally relevant range for the Wilson coefficients. We choose to consider the marginalized limits rather than the individual ones—which tend to be significantly more narrow—because extracting marginalized limits requires that the dependence of the cross section on the Wilson coefficients over that entire range is well modeled.

In our calculations, we consider diagrams involving a single EFT vertex only. As such, the amplitudes will depend linearly (at most) on the Wilson coefficients, and the cross section

will depend quadratically:

$$\sigma(c_i) = s_0 + s_1 c_i + s_2 c_i^2, \quad (4.1)$$

where  $c_i$  is a Wilson coefficient,  $s_0$  is the standard model contribution,  $s_1$  represents the interference between the EFT and the standard model, and  $s_2$  is the pure EFT contribution. In principle,  $s_2$  can contain the effects of interference between two different EFT operators. However, in this study we will turn on only one Wilson coefficient at a time, so the EFT cross section is a quadratic function of a single  $c_i$ .<sup>10</sup>

#### 4.1 Comparison methodology

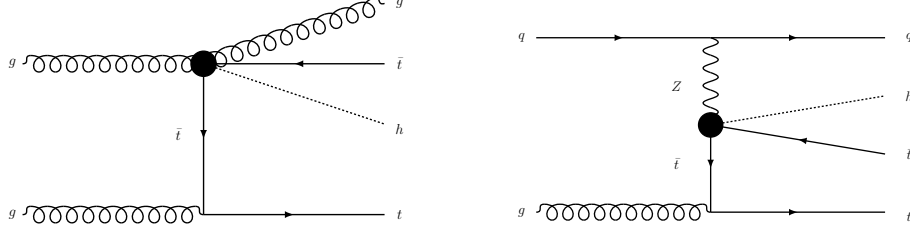
As we examine these results, it is important to define carefully what quantities are interesting to consider. For example, it is well known that the overall normalization obtained from LO calculations, even including the contributions from extra partons with matching, falls short of the NLO value, resulting in the standard practice of applying “K-factors” to correct the normalization of LO samples, for example, when comparing to experimental data. As we look at the impacts of including extra partons with matching compared to LO without matching or NLO calculations, we do not want these known inclusive normalization differences to obscure more relevant differences in how the various samples model the dependence of the cross section on the Wilson coefficients. Therefore, in this paper, we will always look at the value of the inclusive cross section at a given Wilson coefficient relative to the SM cross section value, calculated with the same method; for example, the LO matched samples will be normalized to the SM value predicted by the LO matched calculation. This ratio can be thought of as a “relative K-factor” in that it can be converted into a physical cross section by multiplying by the SM cross section. We will denote this “relative K-factor” with the symbol  $\mu$ . An added benefit to this approach is that sources of uncertainties that just impact the overall normalization, even in the SM case, cancel out, so we can directly visualize the impact of uncertainties on the EFT dependence.

As a starting point, we use the methodology described above to compare a set of matched LO calculations to full NLO calculations. The primary motivation for this is to highlight how our method of comparison is insensitive to the overall normalization, while making clear the comparison in inclusive cross section dependence on the Wilson coefficients between different calculations. Due to the challenges described in section 1 associated with including higher-order QED effects in **MADGRAPH** NLO calculations, we will focus on Wilson coefficients that can enter the processes we consider at the lowest QED order. For example, the effect of  $c_{tG}$  on  $t\bar{t}h$  falls into this category, as  $c_{tG}$  vertices can enter into the  $t\bar{t}h$  process when the QED order is constrained to be less than or equal to 1; an example of a  $t\bar{t}h$  diagram where QED=1

---

<sup>10</sup>Strictly adhering to EFT power counting, keeping the  $c_i^2$  terms in the cross section requires we keep the  $c_i^2$  terms in the amplitude and add the effects from dimension-8 operators interfering with the SM. As the number of dimension-8 operators is large [27, 38], a consistent  $\mathcal{O}(c_i^2)$  result is unwieldy, especially as our goal in this work is to show the effects of matching. We will therefore use Eq. (4.1) with the understanding that the EFT effects it models do not fully describe the complete SMEFT calculation.

that involves a  $c_{tG}$  vertex is shown in Figure 10. Combinations of operators and processes where the operator cannot contribute to the process at the lowest QED order are not the focus of this section due to the complications associated with modeling these effects at NLO.

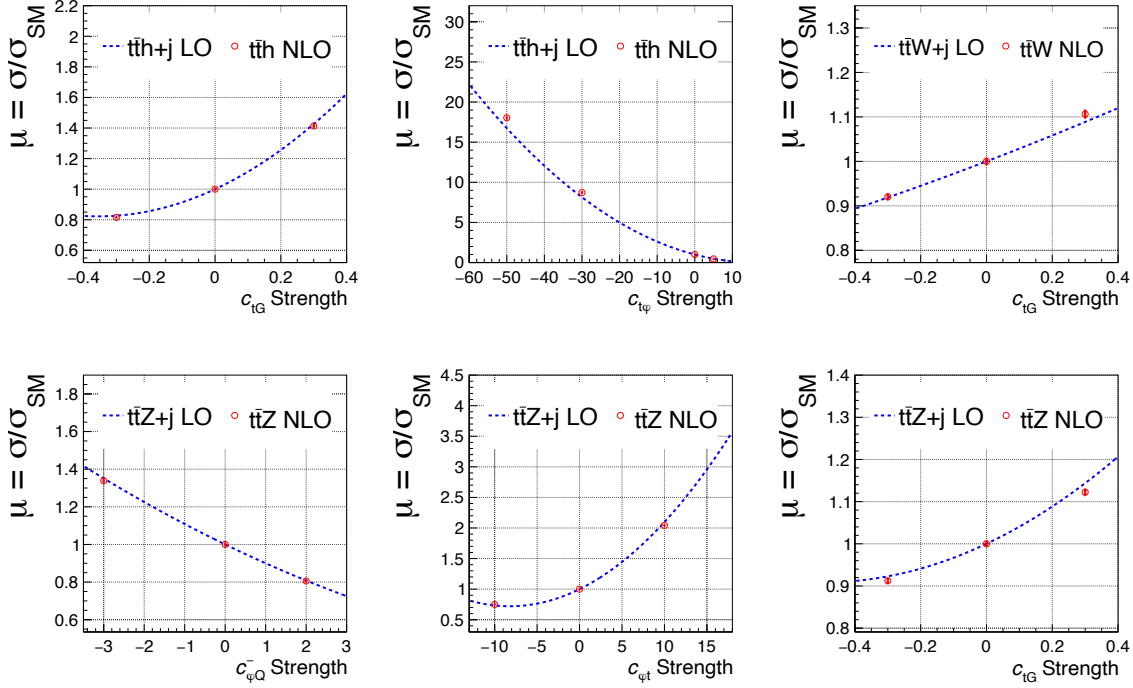


**Figure 10.** Example diagrams that contribute to  $t\bar{t}h$ . The EFT vertex in the diagram on the left is a  $c_{tG}$  vertex, and the QED order of the diagram is 1; we can therefore take this contribution into account at NLO. For comparison, the diagram on the right contains a  $c_{\varphi t}$  vertex, and the QED order of the diagram is 3; we therefore cannot easily take this contribution into account at NLO.

The results of matched LO predictions compared with NLO predictions are shown in Figure 11. As mentioned above, we plot the ratio of the inclusive cross section at a given Wilson coefficient value ( $\sigma$ ) to the cross section predicted by the same calculation for the SM point ( $\sigma_{\text{SM}}$ ), and refer to this ratio as  $\mu$ . Therefore, the y axis indicates the relative scaling of the cross section with respect to the SM as we vary the given Wilson coefficient. For the matched LO calculation, the results are shown as curves generated by fitting a quadratic function to inclusive cross sections calculated at a number of different Wilson coefficient values. The MC statistics are sufficiently large such that the statistical error on the fitted curves is negligible. Since generating NLO MC samples is significantly more computationally expensive (as discussed in Appendix A), we calculate the NLO cross section at just three distinct points (always including the SM as one of the points) and plot those points (scaled to the NLO SM point) with the statistical uncertainty for each point indicated with an error bar.

In general, Figure 11 shows that the dependence of the inclusive cross section on the Wilson coefficients, relative to the SM prediction, are in good agreement between the matched LO and NLO calculations for the processes and operators shown here. Of course, one should remember that the overall normalization of the two calculations is different, but by comparing ratios, that difference drops out. Therefore, any differences visible in the plot of  $\mu$  represent a difference in the quadratic dependence of the cross section on the Wilson coefficient. The largest discrepancies between matched LO and NLO are associated with the dependence of  $t\bar{t}h$  on  $c_{t\varphi}$ , and are 7% or less for the range of NLO points calculated. This amount of discrepancy is reasonable given that the matched LO calculation includes only additional radiation contribution to the corrections, but not the loop contributions.

The fact that the NLO and matched LO calculations agree reasonably well in terms of the inclusive cross section as a function of the Wilson coefficient relative to the SM expectation is perhaps not surprising; it is also known [19, 20] that the relative dependence of the cross



**Figure 11.** Example quadratic cross section plots showing the LO samples (produced with an additional parton) compared with the NLO samples. The y axis shows the ratio of the cross section to the SM cross section; the LO fit is scaled by the LO SM prediction, and the NLO points are scaled by the value of the SM NLO point. As described in the text, the range for each Wilson coefficient is taken from the marginalized limits presented in Ref. [8].

section for  $t\bar{t}h$  and  $t\bar{t}Z$  on these operators calculated at LO without matching is in reasonable agreement with the NLO calculation as well. As mentioned above, this is essentially equivalent to saying that the NLO K-factor is roughly the same for the SM, EFT, and interference contributions.

#### 4.2 The EFT effects of including an extra parton in the LO calculation

In this section, we will compare the LO samples produced with and without matching in order to determine if there are any combinations of operators and processes for which the inclusion of an additional jet with matching changes the observed dependence of the inclusive cross section on the Wilson coefficient. This can be viewed as an interesting way to probe when an NLO K-factor applied to a LO calculation might not be a suitable way to model the EFT dependence. Furthermore, because matched LO calculations are computationally easier to obtain, especially when higher-order QED corrections are involved, we can use these calculations to explore cases that are not currently feasible at NLO.

To avoid preconceptions, we check each of the operators of interest for the  $t\bar{t}h$ ,  $t\bar{t}W$ , and  $t\bar{t}Z$  processes over the ranges from Ref. [8], as described above. For each combination of

operator and process considered, we compare the dependence on the inclusive cross section normalized to the SM expectation as a function of the operator’s Wilson coefficient. The cross section for the matched sample ( $\sigma_{t\bar{t}X+j}$ ) scaled by that sample’s SM expectation is referred to as  $\mu_{t\bar{t}X+j}$ . The cross section for the sample calculated without an extra parton ( $\sigma_{t\bar{t}X}$ ) normalized to that sample’s SM prediction is referred to as  $\mu_{t\bar{t}X}$ . Since these samples are all generated at LO, we place no limits on the QED or QCD order of the diagrams included by **MADGRAPH**. Using each processes’ quadratic parameterizations, we find the normalized cross section at the positive and negative limits of the ranges from Ref. [8], and if  $\mu_{t\bar{t}X+j}$  is more than 10 percent different than  $\mu_{t\bar{t}X}$ , we classify the combination of process and coefficient as being significant impacted by the inclusion on an additional parton.

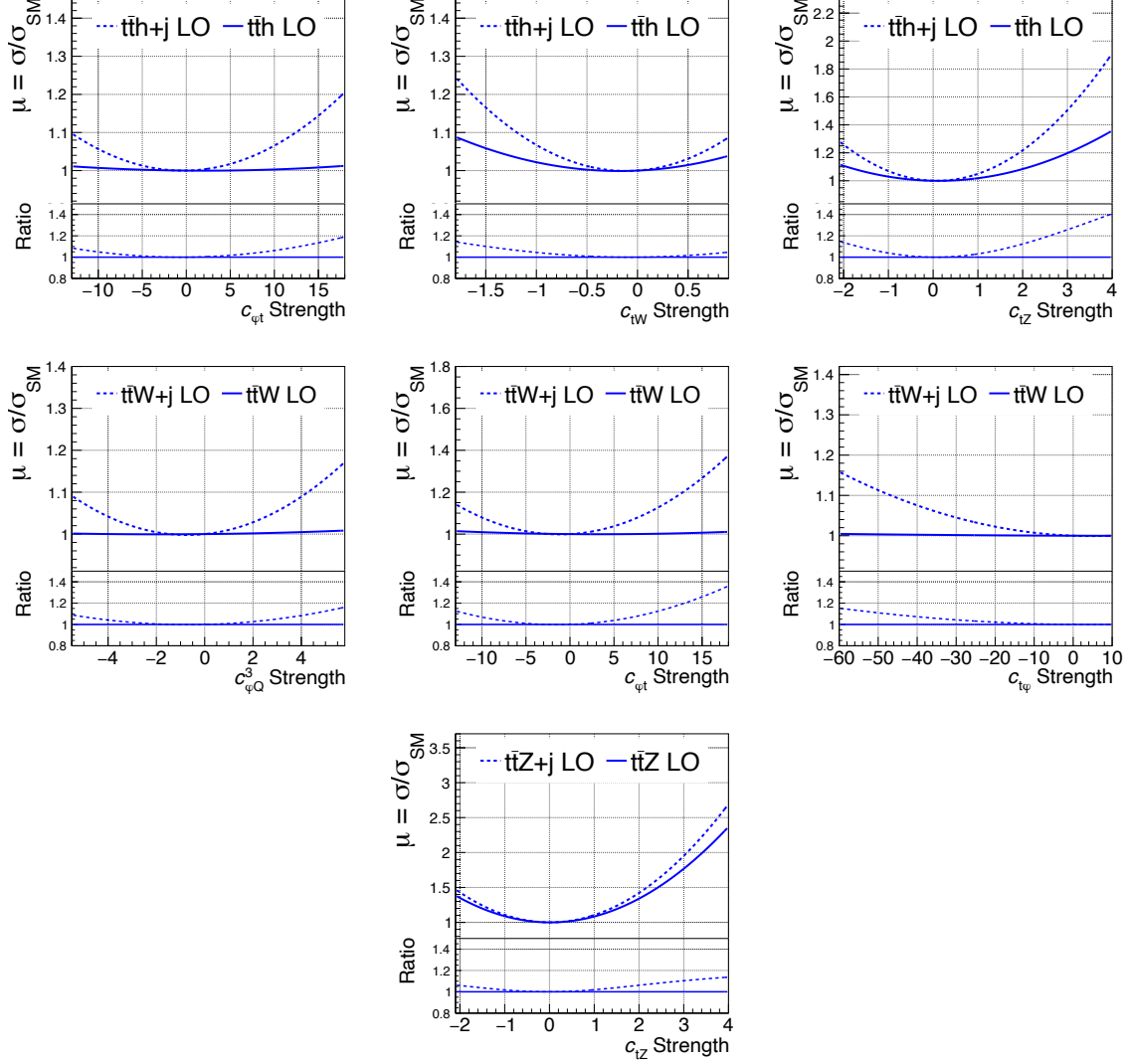
The processes showing large differences for each operator are summarized in Table 2. Note that in the case of  $\mathcal{O}_{tB}$ , while for  $t\bar{t}h$  both  $c_{tW}$  and  $c_{tZ}$  show a significant impact, for  $t\bar{t}Z$ , only  $c_{tZ}$  shows a large difference according to the criteria defined above; additionally, we note that for  $\mathcal{O}_{\varphi q}^1$ , only  $c_{\varphi Q}^3$  has a large impact on  $t\bar{t}W$  according to the criteria above. Section 4.3 provides a discussion of these differences.

**Table 2.** List of operators and corresponding Wilson coefficients considered in this paper indicating for which of the processes we observe a large  $\mu_{t\bar{t}X+j}/\mu_{t\bar{t}X}$  ratio as described in the text. For operators that are associated with two coefficients, if the dependence on either of the coefficients is significantly impacted by the inclusion of an additional parton, the process is listed.

Operator	Definition	Wilson Coefficient	Processes with large $\mu_{t\bar{t}X+j}/\mu_{t\bar{t}X}$
$\mathcal{O}_{t\varphi}$	$\bar{q}t\tilde{\varphi}(\varphi^\dagger\varphi)$	$c_{t\varphi}$	$t\bar{t}W$
$\mathcal{O}_{\varphi q}^1$	$(\varphi^\dagger i\overleftrightarrow{D}_\mu\varphi)(\bar{q}\gamma^\mu q)$	$c_{\varphi Q}^- + c_{\varphi Q}^3$	$t\bar{t}W$
$\mathcal{O}_{\varphi q}^3$	$(\varphi^\dagger i\overleftrightarrow{D}_\mu^I\varphi)(\bar{q}\gamma^\mu\tau^I q)$	$c_{\varphi Q}^3$	$t\bar{t}W$
$\mathcal{O}_{\varphi t}$	$(\varphi^\dagger i\overleftrightarrow{D}_\mu\varphi)(\bar{t}\gamma^\mu t)$	$c_{\varphi t}$	$t\bar{t}h, t\bar{t}W$
$\mathcal{O}_{\varphi tb}$	$(\tilde{\varphi}^\dagger iD_\mu\varphi)(\bar{t}\gamma^\mu b)$	$c_{\varphi tb}$	-
$\mathcal{O}_{tW}$	$(\bar{q}\sigma^{\mu\nu}\tau^I t)\tilde{\varphi}W_{\mu\nu}^I$	$c_{tW}$	$t\bar{t}h$
$\mathcal{O}_{bW}$	$(\bar{q}\sigma^{\mu\nu}\tau^I b)\varphi W_{\mu\nu}^I$	$c_{bW}$	-
$\mathcal{O}_{tB}$	$(\bar{q}\sigma^{\mu\nu}t)\tilde{\varphi}B_{\mu\nu}$	$(c_W c_{tW} - c_{tZ})/s_W$	$t\bar{t}h, t\bar{t}Z$
$\mathcal{O}_{tG}$	$(\bar{q}\sigma^{\mu\nu}T^A t)\tilde{\varphi}G_{\mu\nu}^A$	$c_{tG}$	-

Plots of the inclusive cross section relative to the SM as a function of the relevant Wilson coefficient for these seven combinations of operator and process are shown in Figure 12. Within the ranges of Wilson coefficient values considered, the largest  $\mu_{t\bar{t}X+j}/\mu_{t\bar{t}X}$  ratios are seen in the effect of  $c_{tZ}$  on  $t\bar{t}h$  and the effect of  $c_{\varphi t}$  on  $t\bar{t}W$ . Tables 3 lists the  $\mu_{t\bar{t}X+j}/\mu_{t\bar{t}X}$  ratios (evaluated at the upper and lower limits identified in Ref. [8]) for each of the seven combinations of processes and coefficients for which the inclusion of an additional parton significantly changes the dependence of the process on the given coefficient. The coefficients for the  $t\bar{t}X$  and  $t\bar{t}X+j$  quadratic fits for all of the Wilson coefficients considered in this paper

are shown in Appendix C.



**Figure 12.** Inclusive cross section for  $t\bar{t}h$ ,  $t\bar{t}W$ , and  $t\bar{t}Z$  (with and without an additional parton) as a function of the Wilson coefficient. The y axis shows the ratio of the EFT prediction to the SM. The quadratic parameterizations are generated by fitting a quadratic function to inclusive cross sections calculated at a number of different Wilson coefficient values; the MC statistics are sufficiently large such that the statistical error on the fitted curves is negligible. The ratio plots show the  $\mu_{t\bar{t}X+j}/\mu_{t\bar{t}X}$  ratios. As described in the text, the range for each Wilson coefficient is taken from the marginalized limits presented in Ref. [8].

**Table 3.** Comparison of  $\mu_{t\bar{t}X+j}$  and  $\mu_{t\bar{t}X}$  at the limits identified in [8] for  $t\bar{t}h$ ,  $t\bar{t}W$ , and  $t\bar{t}Z$ . Wilson coefficients whose  $\mu_{t\bar{t}X+j}/\mu_{t\bar{t}X}$  ratio is greater than 1.10 at either the lower or upper limit are included in the table.

Process	WC	Limits from [8]	$\mu_{t\bar{t}X+j}/\mu_{t\bar{t}X}$ at limits
$t\bar{t}h$	$c_{\varphi t}$	[-13, 18]	1.09, 1.19
$t\bar{t}h$	$c_{tW}$	[-1.8, 0.9]	1.14, 1.05
$t\bar{t}h$	$c_{tZ}$	[-2.1, 4.0]	1.15, 1.41
$t\bar{t}W$	$c_{\varphi Q}^3$	[-5.5, 5.8]	1.09, 1.16
$t\bar{t}W$	$c_{\varphi t}$	[-13, 18]	1.13, 1.36
$t\bar{t}W$	$c_{t\varphi}$	[-60, 10]	1.15, 1.00
$t\bar{t}Z$	$c_{tZ}$	[-2.1, 4.0]	1.06, 1.14

### 4.3 Discussion

In Section 4.2, we performed a comparison between the LO samples produced with and without matching, finding that the ratio  $\mu_{t\bar{t}X+j}/\mu_{t\bar{t}X}$  is large for several of the Wilson coefficients and processes considered in this paper. In this section, we will discuss some of the factors that may contribute to how large of an impact the inclusion of an additional parton has on the dependence of a particular process' cross section for a given Wilson coefficient.

The first factor we consider is how the parton distribution functions (PDFs) in the processes involving an additional parton compare to the PDFs in processes that do not involve an additional parton. Without an additional parton, only gluon-gluon and quark-anti-quark initial states are possible for the  $t\bar{t}X$  processes we consider. At the relevant LHC energy ranges, contributions from gluon-gluon initiated processes are significantly larger than quark-anti-quark initiated processes. The inclusion of an additional parton allows for new quark-gluon initiated EFT diagrams. This can lead to a large  $\mu_{t\bar{t}X+j}/\mu_{t\bar{t}X}$  ratio, especially in cases where there are no gluon-gluon initiated EFT diagrams available before the additional parton is included.

Another potential contribution to the differences between  $\mu_{t\bar{t}X+j}$  and  $\mu_{t\bar{t}X}$  is related to the energy scaling of the EFT operators. Since all dimension-six coefficients are suppressed by  $1/\Lambda^2$ , a quantity with units of energy squared is required in the numerator to keep the coupling dimensionless; these powers of energy can either come from the Higgs vacuum expectation value ( $v$ ), or from the energy flowing through the vertex ( $\sim\sqrt{\hat{s}}$ ). Numerators containing  $\sqrt{\hat{s}}$  will scale with the energy of the process, and numerators proportional to  $\hat{s}$  will depend even more strongly on the energy of the process, so at high energies, the contributions from these types of vertices will be enhanced with respect to contributions proportional to  $v^2$ . The inclusion of an additional parton allows us to put more energy into the system, i.e. a  $t\bar{t}X$  system can now recoil against an ISR jet, so we may expect to observe larger  $\mu_{t\bar{t}X+j}/\mu_{t\bar{t}X}$  ratios for contributions involving a stronger dependence on the energy scale of the process. To determine the energy scale dependence of a particular vertex, we should examine the structure



of the operator that gives rise to the vertex of interest. For example, we note that both  $c_{tZ}$  and  $c_{\varphi t}$  have a  $t$ - $t$ - $Z$ - $h$  vertex, but the operator associated with  $c_{tZ}$  contains one Higgs field, while the operator associated with  $c_{\varphi t}$  contains two. Because of this, the  $c_{tZ}$   $t$ - $t$ - $Z$ - $h$  vertex does not contain a  $v$ , and thus scales as  $\hat{s}/\Lambda^2$ . On the other hand, for the  $c_{\varphi t}$   $t$ - $t$ - $Z$ - $h$  vertex, one of the Higgs fields is set to its vev, and the vertex thus scales as  $v\sqrt{\hat{s}}/\Lambda^2$ . Since the  $c_{tZ}$  contribution scales like  $\hat{s}$ , it will grow faster with increasing energy than the  $c_{\varphi t}$  contribution, and we may potentially see a larger impact from the inclusion of an additional parton.

An additional consideration is the color structure of the final state fermions. Except for  $c_{tG}$ , the operators considered in this paper contain EW bosons, so the  $t\bar{t}$  pairs produced by these operators are color singlets. These final states cannot interfere with color octet pairs, such as those produced by gluon splitting. However, when we account for additional radiation in the final state, we have the potential for more complex color structure, which may open up interference with some SM contributions and lead to a larger  $\mu_{t\bar{t}X+j}$  compared to  $\mu_{t\bar{t}X}$ .

The final consideration we will discuss is the chirality of the final state fermions. In SM  $t\bar{t}h$  production, the top quarks are of opposite chirality (i.e. one top quark is a  $t_R$  while the other is a  $t_L$ ). For SM  $t\bar{t}W$  and  $t\bar{t}Z$  processes, the top quarks are of the same chirality. However, the processes involving EFT vertices contain both same-chirality and opposite-chirality contributions, so in the cases where the chirality does not agree with the SM chirality, interference with the SM will be suppressed. Since these contributions enter proportional to the mass of the given particle, the suppression will be less significant for interactions involving top quarks of the wrong chirality, and more significant for interactions involving bottom quarks of the wrong chirality.

In this section, we will step through each of the three processes and discuss how the factors outlined above may affect the cross section's dependence on the relevant Wilson coefficients when an additional parton is included in the process. We note that this discussion does not represent a definitive explanation for the  $\mu_{t\bar{t}X+j}/\mu_{t\bar{t}X}$  ratios; due to the competing influences of the multifarious effects described above, it is challenging to determine which combinations of processes and coefficients will be strongly impacted by the inclusion of an additional parton. However, since we have shown that the extra parton can affect the predicted EFT dependence in many cases, we argue that one should consider accounting for the additional parton whenever possible to avoid inadvertently neglecting significant EFT contributions.

#### 4.3.1 Discussion of $t\bar{t}h$

In section 4.2 we found that the three Wilson coefficients whose effects on  $t\bar{t}h$  significantly change when an additional parton is included are  $c_{tW}$ ,  $c_{tZ}$ , and  $c_{\varphi t}$ . For all three of these cases, the EFT diagrams involving an additional parton can have quark-gluon initial states, which is not possible without an additional parton; furthermore, there are no gluon-gluon initiated diagrams available for the sample produced without an extra parton for these coefficients. We therefore posit that these new quark-gluon initial states are largely responsible for the

significant difference in EFT dependence observed. However, there are several subtleties to this argument that should be addressed.

Since the operators associated with  $c_{\varphi Q}^3$  and  $c_{\varphi Q}^-$  are structurally very similar to the operator associated with  $c_{\varphi t}$ , we might initially expect to see a corresponding increase in the cross section's dependence on these coefficients when considering an additional parton, but no such increase is reported in section 4.2. For  $c_{\varphi Q}^3$ , the apparent discrepancy can be explained by the fact that the quark in the new quark-gluon initiated diagrams must be a b quark, so the effect is suppressed. However, this is not the case for  $c_{\varphi Q}^-$ , as the quarks in the quark-gluon initial states are not required to be b quarks; the apparent discrepancy is instead resolved by noting that the limits on  $c_{\varphi Q}^-$  are much tighter than the limits on  $c_{\varphi t}$ . When we consider a comparable range of values for the coefficients, we see a comparable difference between the  $t\bar{t}h$  and  $t\bar{t}h+j$  curves, and there is thus no discrepancy between the effects we observe for  $c_{\varphi Q}^-$  and  $c_{\varphi t}$ .

### 4.3.2 Discussion of $t\bar{t}Z$

In section 4.2 we found that the only Wilson coefficient whose effect on  $t\bar{t}Z$  significantly changed with the inclusion of an additional parton is  $c_{tZ}$ . To understand why this is the case, we begin by considering the initial states that become available when an additional parton is included. For  $c_{tZ}$ , the additional parton allows new quark-gluon initiated diagrams (where the quark does not have to be a bottom quark). However, there are gluon-gluon initiated  $t\bar{t}Z$  diagrams involving  $c_{tZ}$  already available without an additional parton, so we may expect the impact of including an additional parton to be smaller than those reported for  $t\bar{t}h$ . It should be noted that  $c_{tZ}$  is not the only coefficient for which the inclusion of an additional parton opens up new quark-gluon initiated diagrams where the quark does not have to be a bottom quark;  $c_{tG}$ ,  $c_{tW}$ ,  $c_{\varphi t}$ ,  $c_{\varphi Q}^-$ , and  $c_{t\varphi}$  all share this feature. Furthermore, for  $c_{tW}$  and  $c_{t\varphi}$ , there are no gluon-gluon contributions involving these coefficients when an extra parton is not included, so we may expect the effect of the additional parton to be even more significant than for  $c_{tZ}$ .

To understand the cause of these apparent discrepancies, we will first consider the limits on the relevant coefficients. The limits on  $c_{tG}$  and  $c_{tW}$  are indeed tighter than the  $c_{tZ}$  limits, and when comparable ranges are considered, the  $\mu_{t\bar{t}X+j}/\mu_{t\bar{t}X}$  ratio are in fact comparable to (or even larger than) the  $c_{tZ}$  ratio as we would expect. However, the limits on  $c_{\varphi t}$ ,  $c_{\varphi Q}^-$ , and  $c_{t\varphi}$  are comparable to or looser than the  $c_{tZ}$  limits, so a different explanation is required. For  $c_{\varphi t}$  and  $c_{\varphi Q}^-$ , let us consider the energy dependence. As discussed previously, the operators associated with  $c_{\varphi t}$  and  $c_{\varphi Q}^-$  involve two Higgs fields, while the operator associated with  $c_{tZ}$  only contains one; though all three coefficients are associated with a  $t$ - $t$ - $Z$  vertex, the  $c_{tZ}$  vertex will scale more strongly with energy, so we may expect to see larger impact for  $c_{tZ}$  when including an additional parton. The final coefficient to discuss is  $c_{t\varphi}$ , which has essentially no impact on  $t\bar{t}Z$  before the inclusion of an additional parton and very little effect after. One factor that may explain this observation is that  $t$ - $t$ - $h$  is the only  $c_{t\varphi}$  vertex that

contributes to  $t\bar{t}Z$ , which we would not expect to have a large effect on  $t\bar{t}Z$  due to the Yukawa couplings required.

### 4.3.3 Discussion of $t\bar{t}W$

For  $t\bar{t}W$ , we note that the inclusion of an additional parton leads to new quark-gluon initiated diagrams (where the quark does not have to be a bottom quark) for all nine of the coefficients considered in this paper; furthermore, since  $t\bar{t}W$  does not involve any gluon-gluon initiated contributions, we might expect that the inclusion of an additional parton should have a large impact for all of the Wilson coefficients we consider. In section 4.2, we noted that within the limits identified in Ref. [8], the additional parton indeed leads to a significant change in the dependence of  $t\bar{t}W$  on  $c_{\varphi Q}^3$ ,  $c_{\varphi t}$ , and  $c_{t\varphi}$ . It should be noted that  $c_{tW}$  is also strongly affected, but the comparatively tight limits on the coefficient yield a  $\mu_{t\bar{t}X+j}/\mu_{t\bar{t}X}$  ratio that does not meet the 10% cutoff used in section 4.2 when we evaluate the ratio at these tight upper and lower limits.

The dependence of  $t\bar{t}W$  on the other coefficients is also impacted to varying degrees. Factors that may contribute to the relative sizes of these effects include the energy scaling of the vertices and the chirality and color structure of the final state fermions. In general, the fact that many coefficients are affected is consistent with what we expect based on the quark-gluon initial states that become available at when one additional parton is included. However, we should note that there are two coefficients that are notably not impacted by the additional parton; namely, both  $c_{bW}$  and  $c_{\varphi tb}$  do not have an affect on  $t\bar{t}W$  either before or after the inclusion of the additional parton. This is likely due to the fact that the  $t$ - $b$ - $W$  vertices associated with these coefficients are unique in that they involve right-handed bottom quarks, so the interference with SM  $t\bar{t}W$  production is suppressed.

## 5 Matching systematics

In the previous section, we examined the effect of including an extra parton in the  $t\bar{t}h$ ,  $t\bar{t}W$ , and  $t\bar{t}Z$  processes, finding relatively large differences from the LO without matching in several cases. However, it is fair to ask how the size of these differences compares with the uncertainties on the LO prediction. If these effects are not significantly larger than the uncertainties, then it would still be safe to ignore them. Conversely, if the differences observed when adding an extra parton with matching are large compared to the uncertainties, then these contributions should not be ignored.

In this section, we will explore the size of systematic uncertainties associated with the matched samples. As we did in prior sections, we will focus on  $\mu = \sigma/\sigma_{\text{SM}}$ . This will allow us to focus on uncertainties that change the quadratic shape of the inclusive cross section's dependence on the Wilson coefficients without becoming distracted by effects impacting only the overall normalization, even in the SM. The following uncertainties are included:

- We varied the nominal matching scale (i.e.  $q\text{Cut} = 19 \text{ GeV}$ ) between 15 GeV and 25 GeV.

- We varied the renormalization ( $\mu_R$ ) and factorization scale ( $\mu_F$ ) scales of the hard process. Each scale was varied independently up or down by a factor of two and an envelope was constructed from the various combinations of up/down variations of the two scales.
- We varied the initial and final state radiation (ISR and FSR, respectively) scales in the parton shower up and down by a factor of  $\sqrt{2}$ .
- We use the PDF set NNPDF3.1, so the uncertainties are computed with eigenvector PDF members using the LHAPDF tools as described in Ref. [39].

We will consider how these uncertainties affect several example combinations of processes and Wilson coefficients, as a function of the Wilson coefficient.

Three combinations of Wilson coefficients and operators are chosen for this study. For  $t\bar{t}h$ , we select  $c_{\phi t}$  as an example of a Wilson coefficient whose effect on  $t\bar{t}h$  is significantly different when an extra parton is included. As an example of a Wilson coefficient whose effect on the process does not significantly change when an additional parton is included in the process, we also study the effect of  $c_{tG}$  on  $t\bar{t}h$ . It is also interesting to consider  $c_{tG}$  because it is connected with the only operator among the subset under consideration whose vertices involve gluons. To broaden the study by including a process besides  $t\bar{t}h$ , we also examine the effect of  $c_{\varphi Q}^3$  on  $t\bar{t}W$  as an example of an operator that affects  $t\bar{t}W$  significantly differently when an extra parton is included in the calculation.

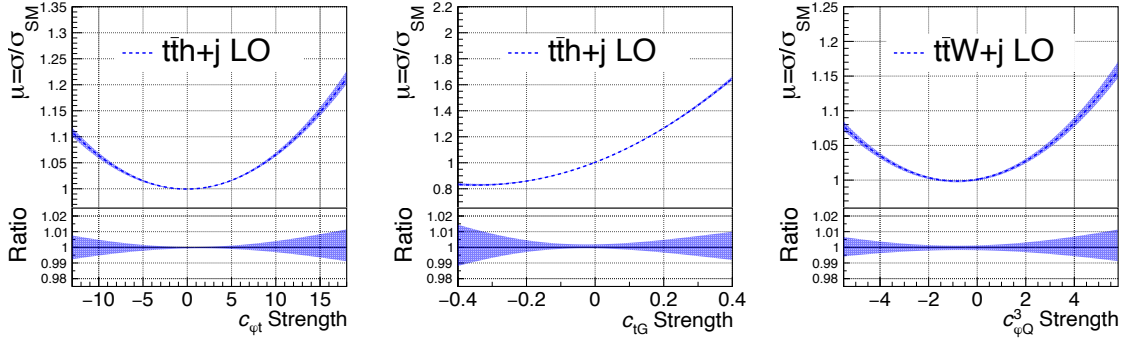
For each of these three representative combinations of processes and Wilson coefficients, we vary the systematic uncertainties listed above for a range of Wilson coefficient values. Fitting a quadratic to the points, we can determine how the uncertainties scale with the Wilson coefficient; the effects of the systematics are combined in quadrature and shown as an error bands on the quadratic curves in Figure 13. Normalized to the SM, these plots show how the effects of the systematic uncertainties evolve as we move away from the SM; sources of uncertainties that just impact the overall normalization cancel out.

Within the range of values considered for each Wilson coefficient (taken from the marginalized limits presented in Ref. [8]) the combined effects of the systematic uncertainties are small, on the order of one percent. The relative size of each systematic is shown in Table 4.

Recall that in section 4.2, we found that including an additional parton can have a sizeable effect, with a  $\mu_{t\bar{t}X+j}/\mu_{t\bar{t}X}$  ratio of 10% or larger for several combinations of processes and Wilson coefficients. As seen in Table 4, the size of the systematic uncertainties for the  $t\bar{t}X+j$  processes (relative to the SM) are less than 2%. This is much smaller than the effects described section 4.2. In Figure 14, we plot both the  $t\bar{t}X$  and  $t\bar{t}X+j$  curves together with the uncertainty bands on both predictions. It is easy to see that the size of effect of including an extra parton is much larger than the size of the associated systematic uncertainties.

**Table 4.** Examples of the relative sizes of the systematic uncertainties considered in this section. For each Wilson coefficient shown, the symmetrized fractional uncertainty for each systematic at the limits from [8] is listed.

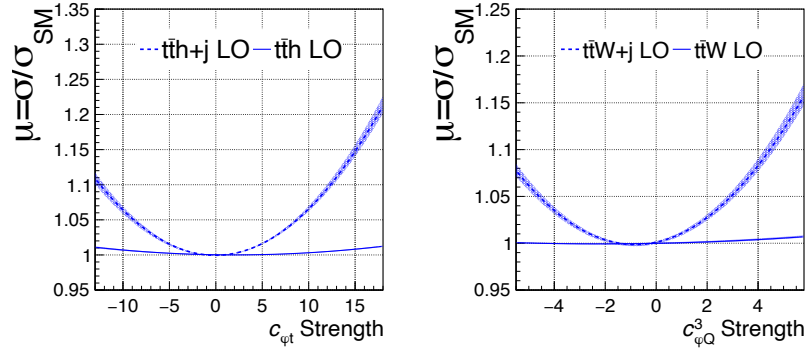
Process	WC value	$\mu_R$	$\mu_F$	PDF	ISR	FSR	qCut	Total
$t\bar{t}h$	$c_{\varphi t} = 18$	0.5%	0.6%		0.1%	0.2%	0.5%	1.0%
$t\bar{t}h$	$c_{\varphi t} = -13$	0.4%	0.4%		0.1%	0.1%	0.2%	0.8%
$t\bar{t}h$	$c_{tG} = 0.4$	0.8%	0.1%		<0.1%	<0.1%	0.3%	0.9%
$t\bar{t}h$	$c_{tG} = -0.4$	1.1%	0.2%		0.1%	0.3%	0.5%	1.3%
$t\bar{t}W$	$c_{\varphi Q}^3 = 5.8$	1.0%	0.2%		0.1%	0.2%	0.1%	1.0%
$t\bar{t}W$	$c_{\varphi Q}^3 = -5.5$	0.7%	<0.1%		0.1%	0.2%	<0.1%	0.6%



**Figure 13.** Systematic uncertainties on the inclusive  $t\bar{t}h$  cross sections for  $c_{\varphi t}$  and  $c_{tG}$ , and on the inclusive  $t\bar{t}W$  cross section for  $c_{\varphi Q}^3$ . The cross sections are normalized to the SM, as discussed in the text. Additionally, the range for each Wilson coefficient is taken from the marginalized limits presented in Ref. [8].

## 6 Conclusion

In this paper we have motivated the use of leading order calculations plus matching for SMEFT  $t\bar{t}X+j$ ,  $X = h/W/Z$  studies. Being leading order, the numerical calculations are fast, have no issue with negative event weights, and can be done without restriction on the coupling order, thereby avoiding any model-dependence or subtlety of the QED or QCD order associated to SMEFT couplings. As LO plus matching is only an approximation of higher order effects, we do not expect it to correctly determine the overall normalization. However, it does capture the relative SMEFT vs. SM cross section. Matching introduces new uncertainties in the form of the phase space boundary where the calculation is split between the matrix element and parton shower; however, we find that the uncertainty introduced by varying this boundary is less than 2% in the cases we have considered. The impact of including an additional parton can be much larger than this. For example, the systematic uncertainties related to matching for  $t\bar{t}h$   $c_{\varphi t}$  are approximately 1%, while the inclusion of an



**Figure 14.** Systematic uncertainties on the inclusive  $t\bar{t}h$  and  $t\bar{t}h+j$  cross sections for  $c_{\phi t}$  and the inclusive  $t\bar{t}W$  and  $t\bar{t}W+j$  cross section for  $c_{\psi Q}^3$ . Both  $t\bar{t}X$  and  $t\bar{t}X+j$  curves are normalized to the SM as discussed in the text, and both curves include the uncertainty band, but the uncertainties on the  $t\bar{t}X$  curve are too small to be seen on this plot. We note that the size of the uncertainties are much smaller than the size of the effect of including an extra parton. As described in the text, the range for each Wilson coefficient is taken from the marginalized limits presented in Ref. [8].

additional parton has a 19% effect at the upper limit for  $c_{\phi t}$  identified in Ref. [8].

As SMEFT vertices cannot be added easily to the parton shower, one may worry that certain regions of phase space are being inappropriately treated. While a valid concern, we find that, if we focus on operators that are most strongly constrained by  $t\bar{t}X$  processes, the only operator missing from the parton shower is  $\mathcal{O}_{tG}$ . Studying the impact of this operator in soft and collinear regions of phase space, we find that its impact is minimal due to the chirality and momentum structure of the operator. This conclusion is similar to what Ref. [25] found regarding  $\mathcal{O}_G$ . Monte Carlo studies, in the form of DJR plots, back up our analytical results, provided we make sure to include all relevant five-point vertices in the SMEFT UFO.

To characterize the benefits of adding one additional parton with matching, we compared LO with no extra partons ( $t\bar{t}X$ ) to the matched LO with an extra parton ( $t\bar{t}X+j$ ). For many of the operators considered for these  $t\bar{t}X$  processes, NLO calculations are difficult to obtain due to complications involving the QED order assigned to the EFT vertices, so the matched LO result provides a particularly interesting estimate of the impact of higher-order corrections on the cross section dependence (relative to the SM) on the Wilson coefficients. We find several cases in which the dependence on Wilson coefficient is significantly modified by the addition of an extra parton. The dominant factor accounting for this change is that when including diagrams containing an additional parton, diagrams with quark-gluon initial states are able to contribute, while the diagrams without extra partons are limited only to quark-quark initiated diagrams. Factors such as the range of Wilson coefficient considered, as well as the derivative, chiral and color structure of the operator under consideration also play an important role.

In closing, while there is no doubt that when practical, NLO calculations provide more

accurate predictions, matched LO calculations do provide a computationally affordable alternative. Matched LO calculations are preferable to strictly LO calculations without extra partons, because contributions from diagrams with an extra parton can provide important modifications to the dependence of the cross section on the Wilson coefficient. Given that matched LO calculations evade complications connected to the QED order assigned to various EFT vertices, they can even be seen to provide a useful cross check on the full NLO predictions.

## Acknowledgments

We thank Steve Mrenna for numerous helpful discussions. The work of AM was supported in part by the National Science Foundation under Grant Number PHY-1820860. The work of RG, KL, KM, and AW was supported in part by the National Science Foundation under Grant Number PHY-1914059. The work of JHK was supported by Chungbuk National University Korea National University Development Project (2020). JHK is grateful to Haider Alhazmi for valuable help and discussions.

## A Comparison of CPU and negative weights for NLO vs LO calculations

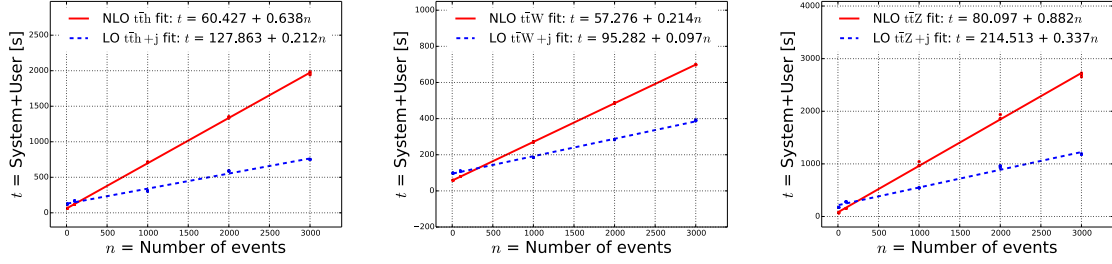
While it is clear that NLO MC are superior in terms of precision and accuracy of modeling QCD effects, there are reasons why matched tree-level calculations might be more practical. For example, the matched tree-level calculations typically require far less CPU-time for event generation. Another consideration is the fact that the NLO calculation produces negatively weighted events, which can affect the statistical power of the samples generated at NLO. Below we quantify some of these differences for a representative example calculation, **MADGRAPH** which can produce both tree-level matched and NLO results.

### A.1 CPU comparison

In this section, we will investigate the differences in CPU required to generate events with both the NLO and LO models. For this comparison, we used **MADGRAPH** to produce  $t\bar{t}X$  gridpacks with both the SMEFT NLO model and our LO model. The LO gridpacks were produced with QED=1, QCD=3 order constraints for consistency with the NLO gridpacks. For each process ( $t\bar{t}h$ ,  $t\bar{t}W$ , and  $t\bar{t}Z$ ), we generated a set number of events from both the NLO and LO gridpacks and recorded the CPU time required. Repeating this procedure for several set numbers of events, we can plot the CPU time required versus the number of events generated to understand how the CPU time scales with the number of events for the NLO and LO models. The same random seed is used for each running of each gridpack. These tests were performed using a single core from an AMD Opteron 6276 2.3 GHz processor.

Figure 15 shows the NLO and LO CPU comparisons  $t\bar{t}h$ ,  $t\bar{t}W$ , and  $t\bar{t}Z$ . Though the overhead time (indicated by the y-intercept of the fit) for the LO samples is somewhat larger than the NLO overhead, the slopes of the NLO fits are larger than the LO slopes by a factor of about two to three, depending on the process.





**Figure 15.** Plots showing how the CPU time scales with the number of events generated for the NLO and LO models. For  $t\bar{t}h$  (left), the slope of the fit to the NLO points is about 3 times larger than the slope of the line fit to the LO points. For  $t\bar{t}W$  (center), the slope of the fit the NLO points is about 2 times larger than the slope of the line fit to the LO points. For  $t\bar{t}Z$  (right), the slope of the line fit to the NLO points is about 3 times larger than the slope of the line fit to the LO points. For consistency with the NLO samples, the LO samples were generated with a QED=1 constraint.

## A.2 Negative weights

It is known that NLO cross sections are not always positive locally in the phase space which would imply events with negative weights in simulations. The negative weighted events reduce the statistical accuracy of the simulated samples at the level of physical observable. For example, the statistical power of a sample with a fraction of 20-30% negative weighted events is 3-5 times less than a sample with the same number of events but all weighted positively [40].

Currently the vast majority of studies about the SMEFT are based on the predictions by the **MADGRAPH**. The NLO simulated samples, generated by the **MADGRAPH**, have negative-weight events which their fractions vary for different processes. In Simulated  $t\bar{t}W$ ,  $t\bar{t}Z$  and  $t\bar{t}h$  samples by the **MADGRAPH** for an arbitrary reference point in WC phase space, far from the SM point, we find 17%, 27%, and 28% negative-weight events, respectively. This means we need to simulate 3-5 times larger sample at NLO for finding similar statistical precision as LO sample which will cause more time and CPU cost for a global EFT search.

## B Feynman rules for $t$ - $t$ - $g$ couplings

Feynman rules for the  $t$ - $t$ - $g$  couplings are shown in Figure 16, where the index  $a$  runs over the eight color degrees of freedom of the gluon field.

## C Comparison of fit parameters for $t\bar{t}X$ and $t\bar{t}X+j$ processes

In tables 5, 6, and 7 we report the coefficients of the quadratic  $\mu_{t\bar{t}X}$  and  $\mu_{t\bar{t}X+j}$  fits. Since the fits have been normalized to the SM, the constant term in each parameterization is 1, and we only report the coefficients for the linear terms (corresponding to the interference between the given Wilson coefficient and the SM) and quadratic terms (corresponding to the purely



**Table 6.** Summary of quadratic fit coefficients for the  $t\bar{t}W$  parameterization for  $\mu_{t\bar{t}W}$  and  $\mu_{t\bar{t}W+j}$ . As discussed in the text, the cross sections are normalized to the SM.

WC	$\mu_{t\bar{t}W}$		$\mu_{t\bar{t}W+j}$	
	Linear coefficient	Quad coefficient	Linear coefficient	Quad coefficient
$c_{tW}$	0.01384	0.02293	0.02332	0.05624
$c_{t\varphi}$	-0.00003	0.00000	-0.00036	0.00004
$c_{\varphi Q}^-$	-0.00085	0.00006	-0.00553	0.00106
$c_{tZ}$	0.00041	0.00139	-0.00043	0.00196
$c_{tG}$	0.28568	0.02875	0.28313	0.04146
$c_{bW}$	0.00000	0.00000	0.00000	0.00000
$c_{\varphi Q}^3$	0.00058	0.00015	0.00590	0.00409
$c_{\varphi tb}$	0.00000	0.00000	0.00000	0.00000
$c_{\varphi t}$	-0.00038	0.00006	0.00237	0.00103

**Table 7.** Summary of quadratic fit coefficients for the  $t\bar{t}Z$  parameterization for  $\mu_{t\bar{t}Z}$  and  $\mu_{t\bar{t}Z+j}$ . As discussed in the text, the cross sections are normalized to the SM.

WC	$\mu_{t\bar{t}Z}$		$\mu_{t\bar{t}Z+j}$	
	Linear coefficient	Quad coefficient	Linear coefficient	Quad coefficient
$c_{tW}$	0.00521	0.00946	0.00780	0.02654
$c_{t\varphi}$	-0.00001	0.00000	-0.00014	0.00001
$c_{\varphi Q}^-$	-0.10464	0.00415	-0.10434	0.00424
$c_{tZ}$	-0.00262	0.08617	-0.00188	0.10632
$c_{tG}$	0.35923	0.26202	0.36826	0.37035
$c_{bW}$	0.00000	0.00154	0.00000	0.00167
$c_{\varphi Q}^3$	0.00368	0.00109	0.00292	0.00106
$c_{\varphi tb}$	0.00000	0.00017	0.00000	0.00016
$c_{\varphi t}$	0.06876	0.00410	0.06820	0.00415

## References

- [1] S. Weinberg, *Baryon and Lepton Nonconserving Processes*, *Phys. Rev. Lett.* **43** (1979) 1566–1570.
- [2] W. Buchmuller and D. Wyler, *Effective Lagrangian Analysis of New Interactions and Flavor Conservation*, *Nucl. Phys. B* **268** (1986) 621–653.
- [3] B. Grzadkowski, M. Iskrzynski, M. Misiak and J. Rosiek, *Dimension-Six Terms in the Standard Model Lagrangian*, *JHEP* **10** (2010) 085, [[1008.4884](#)].
- [4] I. Brivio and M. Trott, *The Standard Model as an Effective Field Theory*, *Phys. Rept.* **793** (2019) 1–98, [[1706.08945](#)].
- [5] A. Buckley, C. Englert, J. Ferrando, D. J. Miller, L. Moore, M. Russell et al., *Constraining top quark effective theory in the LHC Run II era*, *JHEP* **04** (2016) 015, [[1512.03360](#)].
- [6] A. Buckley, C. Englert, J. Ferrando, D. J. Miller, L. Moore, M. Russell et al., *Global fit of top quark effective theory to data*, *Phys. Rev. D* **92** (2015) 091501, [[1506.08845](#)].
- [7] A. Buckley, C. Englert, J. Ferrando, D. J. Miller, L. Moore, K. Nördstrom et al., *Results from TopFitter*, *PoS CKM2016* (2016) 127, [[1612.02294](#)].
- [8] N. P. Hartland, F. Maltoni, E. R. Nocera, J. Rojo, E. Slade, E. Vryonidou et al., *A Monte Carlo global analysis of the Standard Model Effective Field Theory: the top quark sector*, *JHEP* **04** (2019) 100, [[1901.05965](#)].
- [9] CMS collaboration, A. M. Sirunyan et al., *Measurement of the cross section for top quark pair production in association with a W or Z boson in proton-proton collisions at  $\sqrt{s} = 13$  TeV*, *JHEP* **08** (2018) 011, [[1711.02547](#)].
- [10] CMS collaboration, A. M. Sirunyan et al., *Measurement of top quark pair production in association with a Z boson in proton-proton collisions at  $\sqrt{s} = 13$  TeV*, *JHEP* **03** (2020) 056, [[1907.11270](#)].
- [11] ATLAS collaboration, M. Aaboud et al., *Measurement of the  $t\bar{t}Z$  and  $t\bar{t}W$  cross sections in proton-proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, *Phys. Rev. D* **99** (2019) 072009, [[1901.03584](#)].
- [12] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079, [[1405.0301](#)].
- [13] T. Sjöstrand, S. Mrenna and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, *JHEP* **05** (2006) 026, [[hep-ph/0603175](#)].
- [14] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten et al., *An Introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159–177, [[1410.3012](#)].
- [15] C. Degrande, G. Durieux, F. Maltoni, K. Mimasu, E. Vryonidou and C. Zhang, *Automated one-loop computations in the SMEFT*, [2008.11743](#).
- [16] C. Degrande, B. Fuks, K. Mawatari, K. Mimasu and V. Sanz, *Electroweak Higgs boson production in the standard model effective field theory beyond leading order in QCD*, *Eur. Phys. J. C* **77** (2017) 262, [[1609.04833](#)].

- [17] K. Mimasu, V. Sanz and C. Williams, *Higher Order QCD predictions for Associated Higgs production with anomalous couplings to gauge bosons*, *JHEP* **08** (2016) 039, [[1512.02572](#)].
- [18] D. Buarque Franzosi and C. Zhang, *Probing the top-quark chromomagnetic dipole moment at next-to-leading order in QCD*, *Phys. Rev. D* **91** (2015) 114010, [[1503.08841](#)].
- [19] O. Bessidskaia Bylund, F. Maltoni, I. Tsinikos, E. Vryonidou and C. Zhang, *Probing top quark neutral couplings in the Standard Model Effective Field Theory at NLO in QCD*, *JHEP* **05** (2016) 052, [[1601.08193](#)].
- [20] F. Maltoni, E. Vryonidou and C. Zhang, *Higgs production in association with a top-antitop pair in the Standard Model Effective Field Theory at NLO in QCD*, *JHEP* **10** (2016) 123, [[1607.05330](#)].
- [21] D. Barducci et al., *Interpreting top-quark LHC measurements in the standard-model effective field theory*, [1802.07237](#).
- [22] J. Alwall, P. Demin, S. de Visscher, R. Frederix, M. Herquet, F. Maltoni et al., *MadGraph/MadEvent v4: The New Web Generation*, *JHEP* **09** (2007) 028, [[0706.2334](#)].
- [23] J. Alwall et al., *Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions*, *Eur. Phys. J. C* **53** (2008) 473–500, [[0706.2569](#)].
- [24] J. Alwall, S. de Visscher and F. Maltoni, *QCD radiation in the production of heavy colored particles at the LHC*, *JHEP* **02** (2009) 017, [[0810.5350](#)].
- [25] C. Englert, M. Russell and C. D. White, *Effective Field Theory in the top sector: do multijets help?*, *Phys. Rev. D* **99** (2019) 035019, [[1809.09744](#)].
- [26] R. Alonso, E. E. Jenkins, A. V. Manohar and M. Trott, *Renormalization Group Evolution of the Standard Model Dimension Six Operators III: Gauge Coupling Dependence and Phenomenology*, *JHEP* **04** (2014) 159, [[1312.2014](#)].
- [27] B. Henning, X. Lu, T. Melia and H. Murayama, *2, 84, 30, 993, 560, 15456, 11962, 261485, ...: Higher dimension operators in the SM EFT*, *JHEP* **08** (2017) 016, [[1512.03433](#)].
- [28] J. Bramante, A. Delgado, L. Lehman and A. Martin, *Boosted Higgses from chromomagnetic  $b$ 's:  $b\bar{b}h$  at high luminosity*, *Phys. Rev. D* **93** (2016) 053001, [[1410.3484](#)].
- [29] NNPDF collaboration, R. D. Ball, V. Bertone, S. Carrazza, L. Del Debbio, S. Forte, A. Guffanti et al., *Parton distributions with QED corrections*, *Nucl. Phys. B* **877** (2013) 290–320, [[1308.0598](#)].
- [30] J. Chen, T. Han and B. Tweedie, *Electroweak Splitting Functions and High Energy Showering*, *JHEP* **11** (2017) 093, [[1611.00788](#)].
- [31] A. J. Larkoski and M. E. Peskin, *Antenna Splitting Functions for Massive Particles*, *Phys. Rev. D* **84** (2011) 034034, [[1106.2182](#)].
- [32] T. Plehn, *Lectures on LHC Physics*, *Lect. Notes Phys.* **844** (2012) 1–193, [[0910.4182](#)].
- [33] S. Catani, S. Dittmaier and Z. Trocsanyi, *One loop singular behavior of QCD and SUSY QCD amplitudes with massive partons*, *Phys. Lett. B* **500** (2001) 149–160, [[hep-ph/0011222](#)].
- [34] H. Murayama, I. Watanabe and K. Hagiwara, *HELAS: HELicity amplitude subroutines for Feynman diagram evaluations*, .

- [35] J. Chen, P. Ko, H.-N. Li, J. Li and H. Yokoya, *Light dark matter showering under broken dark  $U(1)$  — revisited*, *JHEP* **01** (2019) 141, [[1807.00530](#)].
- [36] P. Lenzi and J. Butterworth, *A Study on Matrix Element corrections in inclusive  $Z/\gamma^*$  production at LHC as implemented in PYTHIA, HERWIG, ALPGEN and SHERPA*, [0903.3918](#).
- [37] NNPDF collaboration, R. D. Ball et al., *Parton distributions for the LHC Run II*, *JHEP* **04** (2015) 040, [[1410.8849](#)].
- [38] L. Lehman and A. Martin, *Low-derivative operators of the Standard Model effective field theory via Hilbert series methods*, *JHEP* **02** (2016) 081, [[1510.00372](#)].
- [39] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht et al., *LHAPDF6: parton density access in the LHC precision era*, *Eur. Phys. J. C* **75** (2015) 132, [[1412.7420](#)].
- [40] R. Frederix, S. Frixione, S. Prestel and P. Torrielli, *On the reduction of negative weights in MC@NLO-type matching procedures*, [2002.12716](#).