Exploring the SMEFT at dimension eight with Drell-Yan transverse momentum measurements

Radja Boughezal,[1] Yingsheng Huang[1,2] and Frank Petriello[1,2]

[1]*HEP Division, Argonne National Laboratory, Argonne, Illinois 60439, USA*
[2]*Department of Physics & Astronomy, Northwestern University, Evanston, Illinois 60208, USA*

We demonstrate that measurements of the neutral-current Drell-Yan transverse momentum distribution binned in invariant mass are sensitive to unexplored dimension-eight parameters of the Standard Model effective field theory (SMEFT). These distributions are sensitive to four-fermion operators with additional QCD field strength tensors. The determination of the Wilson coefficients of these operators provides a useful diagnostic tool that distinguishes possible ultraviolet completions of the SMEFT. We study how well these effects can be probed by current LHC data, and explore the sensitivity of the future high-luminosity LHC (HL-LHC) to these operators. We find that the HL-LHC data has the potential to strongly probe this sector of the SMEFT.

I. INTRODUCTION

The Standard Model (SM) of particle physics successfully describes phenomena ranging from low-energy nuclear physics to high-energy collisions. However, since it does not contain neutrino masses nor dark matter, and cannot explain certain observations such as the matter-antimatter asymmetry in the Universe, undiscovered physics beyond the SM that explains these mysteries must exist. Experiments at the Large Hadron Collider (LHC) and elsewhere are probing the SM at the TeV scale, searching for solutions to these outstanding problems. Since no conclusive deviation from SM predictions has yet been found, a major theme of current research is to understand how heavy new physics can be indirectly probed and constrained by available and upcoming data. This effort helps guide searches for new physics by suggesting in what channels measurable deviations from SM predictions may occur given the current bounds. In the event of a discovery it would also indicate what measurements can serve as diagnostic tools to distinguish between different models of new physics.

A convenient theoretical framework for investigating indirect signatures of heavy new physics is the SM effective field theory (SMEFT). The SMEFT is formed by adding higher-dimensional operators to the SM Lagrangian that are consistent with the SM gauge symmetries and formed only from SM fields. The higher-dimensional operators in the SMEFT are suppressed by appropriate powers of a characteristic energy scale Λ below which heavy new fields are integrated out. The SMEFT encapsulates a broad swath of new physics models, making it easier to simultaneously study numerous theories without focusing on details of their ultraviolet completions that do not matter at low energies. The use of the SMEFT framework to analyze LHC data is similar in spirit to the use of the S and T parameters at LEP to bound entire classes of new physics models, and the global fitting of SMEFT parameters at the LHC promises to provide as powerful probe of beyond the SM theories as the global electroweak precision fit did at LEP. Complete, nonredundant bases for the dimension-six [1–3] and dimension-eight operators [4,5] have been constructed. Odd-dimensional operators violate lepton number and are not considered here. It is an ongoing effort to analyze the numerous available data within the SMEFT framework, primarily in partial analyses of individual SMEFT sectors [6–29]. Recent work has been devoted to performing a global, simultaneous fit of all data available [30–41], and to study the interplay between SMEFT fits and the extraction of parton distributions from data [40,42].

Our focus in this manuscript is on semileptonic four-fermion operators in the SMEFT. These coefficients are not constrained by current global fits of top quark, Higgs boson and electroweak data [38,41]. While they can be probed by low-energy data [15], the strongest bounds come from Drell-Yan data at the LHC. Previous results have shown that existing Drell-Yan data is precise enough to probe dimension-eight operators in the SMEFT [43–45]. These works focused on measurements of the invariant mass

distribution of the lepton pair in the Drell-Yan process. A motivation of our paper is to demonstrate that LHC datasets not originally intended as new physics searches can be sensitive to unprobed regions of the SMEFT parameter space, and therefore have unexpected sensitivity to physics beyond the SM. In particular, we focus on the recent CMS measurement of the doubly differential distribution of invariant mass and transverse momentum in the Drell-Yan process [46], intended as a probe of QCD dynamics. The measurement of transverse momentum makes this dataset sensitive to partonic processes containing emission of gluons. These gluons can either be radiated from external legs, or directly from some heavy state that carries QCD color. In this second case they match to semileptonic four-fermion operators with an additional QCD field-strength tensor. Such operators first appear at dimension eight in the SMEFT and are unconstrained by other datasets.

To illustrate our results we focus on a representative example in which only operators containing right-handed fields have nonzero Wilson coefficients. In this scenario our parameter space consists of three categories of operators: a dimension-six four-fermion operator, two momentum-dependent dimension-eight operators that grow with energy and that have been considered in previous work [43,44], and a single *CP*-even semileptonic four-fermion operator with a gluon field-strength tensor that we henceforth label a gluonic operator. We show that a joint measurement of invariant mass and transverse momentum allows the gluonic operator to be probed independently of the other operators, as it has a distinct dependence on transverse momentum. We stress that the determination of Wilson coefficients for all three operator categories provides a useful diagnostic tool that distinguishes possible ultraviolet (UV) completions of the SMEFT. Although our primary interest is in the bottom-up analysis of the possible SMEFT parameter space, we consider the matching of example Z' and vector leptoquark states to this sector of the SMEFT, and show that they lead to very different patterns of Wilson coefficients for these three operator categories. Although current LHC data provides only weak constraints on the gluonic operator, we study the potential of the high-luminosity LHC (HL-LHC) to probe these effects, and find that significant bounds on all three categories of effects can be obtained. We encourage this measurement to be performed again as larger LHC datasets become available.

Our paper is organized as follows. We review in Sec. II details of the SMEFT needed for our analysis. In Sec. III we study the matching of example UV states onto the four-fermion sector of the SMEFT. Our emphasis in this section is to show that very different patterns of Wilson coefficients can be obtained from different UV states, motivating the measurement of all possible operator types. In Sec. IV we show that the doubly differential distribution in invariant mass and transverse momentum can simultaneously probe both the regular and gluonic semileptonic four-fermion operators. We perform fits to the current data in Sec. V, and to simulated HL-LHC data in Sec. VI. We conclude in Sec. VII.

II. REVIEW OF THE SMEFT

We review in this section aspects of the SMEFT relevant for our analysis of the Drell-Yan process. The SMEFT is an effective field theory extension of the SM that includes terms suppressed by an energy scale Λ. Beyond this scale the ultraviolet completion of the EFT becomes important, and new particles beyond the SM appear. In our study we keep terms through dimension eight in the $1/\Lambda$ expansion, and ignore operators of odd dimension which violate lepton number. Our Lagrangian becomes

$$\mathcal{L} = \mathcal{L}_{\text{SM}} + \frac{1}{\Lambda^2}\sum_i C_i^{(6)}\mathcal{O}_i^{(6)} + \frac{1}{\Lambda^4}\sum_i C_i^{(8)}\mathcal{O}_i^{(8)} + \cdots, \quad (1)$$

where the ellipsis denotes operators of higher dimensions. The Wilson coefficients defined above are dimensionless. Cross sections computed through $\mathcal{O}(1/\Lambda^4)$ will have contributions from the square of dimension-six operators, as well as interferences between dimension-eight operators and the SM.

The categories of operators contributing to the Drell-Yan process through dimension eight were extensively cataloged in Ref. [44]. At the dimension-six level three categories of operators contribute: corrections to the three-point vertices of gauge bosons with fermions, four-fermion operators, and dipole operators coupling fermions to gauge bosons. The vertex corrections lead to effects that scale with energy as $\mathcal{O}(v^2/\Lambda^2)$, where v denotes the Higgs vacuum expectation value. These are subleading at high energies compared to the four-fermion operators that scale as $\mathcal{O}(s/\Lambda^2)$, and are strongly constrained by Z-pole observables [47]. We therefore neglect these terms in our analysis. We additionally assume minimal flavor violation for the structure of our Wilson coefficients. This assumption makes all dipole operators, as well as all scalar and tensor four-fermion operators, proportional to SM Yukawa couplings. These couplings are small for the processes considered here, and can be safely neglected. This leaves us with only vectorlike four-fermion operators contributing at dimension six. The contributing terms are summarized below in Table I. q and l denote left-handed quark and lepton doublets, while u, d and e denote right-handed singlets for the up quarks, down quarks and leptons, respectively. τ^I denote the SU(2) Pauli matrices.

Several classes of operators contribute at the dimension-eight level. Considering first the leading-order four-fermion process, we again have corrections to the $f\bar{f}V$ vertices, four-fermion operators with Higgs insertions, and four-fermion operators with derivative insertions. The first two

TABLE I. Dimension-six four-fermion operators contributing to Drell-Yan at leading order in the coupling constants.

$\mathcal{O}_{lq}^{(1)}$	$(\bar{l}\gamma^\mu l)(\bar{q}\gamma_\mu q)$	$\mathcal{O}_{lu}$	$(\bar{l}\gamma^\mu l)(\bar{u}\gamma_\mu u)$
$\mathcal{O}_{lq}^{(3)}$	$(\bar{l}\gamma^\mu \tau^I l)(\bar{q}\gamma_\mu \tau^I l q)$	$\mathcal{O}_{ld}$	$(\bar{l}\gamma^\mu l)(\bar{d}\gamma_\mu d)$
$\mathcal{O}_{eu}$	$(\bar{e}\gamma^\mu e)(\bar{u}\gamma_\mu u)$	$\mathcal{O}_{qe}$	$(\bar{q}\gamma^\mu q)(\bar{e}\gamma_\mu e)$
$\mathcal{O}_{ed}$	$(\bar{e}\gamma^\mu e)(\bar{d}\gamma_\mu d)$		

categories of operators were shown in Ref. [44] to be negligible for reasonable values of the Wilson coefficients, consistent with their energy scaling: $\mathcal{O}(v^4/\Lambda^4)$ for the first category and $\mathcal{O}(sv^2/\Lambda^4)$ for the second. The four-fermion operators with derivative insertions scale as $\mathcal{O}(s/\Lambda^4)$ and are non-negligible. They are shown in Table II. We note that the type-II operators lead to novel angular dependence [43], but vanish upon integration over angles up to small corrections due to acceptance cuts [44]. While we discuss them when matching specific UV examples to the SMEFT, we do not consider them in our numerical analysis since the distributions considered here show little sensitivity to these effects. A proposal for a series of angular measurements to probe these terms was given in [43,48].

In our study we will consider the transverse momentum spectrum in Drell-Yan. In this case we also need to consider dimension-eight operators with a gluon field strength inserted. The possible operators of this form were enumerated in Refs. [4,5]. We organize them according to their *CP* transformation properties in Table III. These transformation rules can be obtained using results from any standard QFT text, or from studying the structure of an explicit amplitude calculation. Since the *CP*-odd operators

TABLE II. Dimension-eight two-derivative four-fermion operators that contribute to the Drell-Yan process. The type-I operators affect inclusive invariant mass and transverse momentum distributions, while the type-II operators affect only angular distributions and vanish when integrated inclusively over angles. The parentheses in the superscripts denote symmetrization over the enclosed indices, while $\overleftrightarrow{D}_\nu = \overrightarrow{D}_\nu - \overleftarrow{D}_\nu$.

	Type I		Type II
$\mathcal{O}_{l^2q^2D^2}^{(1)}$	$D^\nu(\bar{l}\gamma^\mu l)D_\nu(\bar{q}\gamma_\mu q)$	$\mathcal{O}_{l^2q^2D^2}^{(2)}$	$(\bar{l}\gamma^{(\mu}\overleftrightarrow{D}^{\nu)} l)(\bar{q}\gamma_{\mu(}\overleftrightarrow{D}_{\nu)} q)$
$\mathcal{O}_{l^2q^2D^2}^{(3)}$	$D^\nu(\bar{l}\gamma^\mu \tau^I l)D_\nu(\bar{q}\gamma_\mu \tau^I q)$	$\mathcal{O}_{l^2q^2D^2}^{(4)}$	$(\bar{l}\tau^I\gamma^{(\mu}\overleftrightarrow{D}^{\nu)} l)(\bar{q}\tau^I\gamma_{\mu(}\overleftrightarrow{D}_{\nu)} q)$
$\mathcal{O}_{e^2u^2D^2}^{(1)}$	$D^\nu(\bar{e}\gamma^\mu e)D_\nu(\bar{u}\gamma_\mu u)$	$\mathcal{O}_{e^2u^2D^2}^{(2)}$	$(\bar{e}\gamma^{(\mu}\overleftrightarrow{D}^{\nu)} e)(\bar{u}\gamma_{\mu(}\overleftrightarrow{D}_{\nu)} u)$
$\mathcal{O}_{e^2d^2D^2}^{(1)}$	$D^\nu(\bar{e}\gamma^\mu e)D_\nu(\bar{d}\gamma_\mu d)$	$\mathcal{O}_{e^2d^2D^2}^{(2)}$	$(\bar{e}\gamma^{(\mu}\overleftrightarrow{D}^{\nu)} e)(\bar{d}\gamma_{\mu(}\overleftrightarrow{D}_{\nu)} d)$
$\mathcal{O}_{l^2u^2D^2}^{(1)}$	$D^\nu(\bar{l}\gamma^\mu l)D_\nu(\bar{u}\gamma_\mu u)$	$\mathcal{O}_{l^2u^2D^2}^{(2)}$	$(\bar{l}\gamma^{(\mu}\overleftrightarrow{D}^{\nu)} l)(\bar{u}\gamma_{\mu(}\overleftrightarrow{D}_{\nu)} u)$
$\mathcal{O}_{l^2d^2D^2}^{(1)}$	$D^\nu(\bar{l}\gamma^\mu l)D_\nu(\bar{d}\gamma_\mu d)$	$\mathcal{O}_{l^2d^2D^2}^{(2)}$	$(\bar{l}\gamma^{(\mu}\overleftrightarrow{D}^{\nu)} l)(\bar{d}\gamma_{\mu(}\overleftrightarrow{D}_{\nu)} d)$
$\mathcal{O}_{q^2e^2D^2}^{(1)}$	$D^\nu(\bar{q}\gamma^\mu q)D_\nu(\bar{e}\gamma_\mu e)$	$\mathcal{O}_{q^2e^2D^2}^{(2)}$	$(\bar{e}\gamma^{(\mu}\overleftrightarrow{D}^{\nu)} e)(\bar{q}\gamma_{\mu(}\overleftrightarrow{D}_{\nu)} q)$

TABLE III. Dimension-eight four-fermion operators with a gluon field that contribute to the Drell-Yan transverse momentum spectrum, organized according to their *CP* transformation properties. $\tilde{G}$ denotes the dual field-strength tensor.

	CP even			*CP* odd	
$\mathcal{O}_{l^2q^2\tilde{G}}^{(1)}$	$(\bar{l}\gamma^\mu l)(\bar{q}\gamma^\nu T^A q)\tilde{G}_{\mu\nu}^A$		$\mathcal{O}_{l^2q^2G}^{(1)}$	$(\bar{l}\gamma^\mu l)(\bar{q}\gamma^\nu T^A q)G_{\mu\nu}^A$	
$\mathcal{O}_{l^2q^2\tilde{G}}^{(2)}$	$(\bar{l}\tau^I\gamma^\mu l)(\bar{q}\tau^I\gamma^\nu T^A q)\tilde{G}_{\mu\nu}^A$		$\mathcal{O}_{l^2q^2G}^{(2)}$	$(\bar{l}\tau^I\gamma^\mu l)(\bar{q}\tau^I\gamma^\nu T^A q)G_{\mu\nu}^A$	
$\mathcal{O}_{e^2u^2\tilde{G}}$	$(\bar{e}\gamma^\mu e)(\bar{u}\gamma^\nu T^A u)\tilde{G}_{\mu\nu}^A$		$\mathcal{O}_{e^2u^2G}$	$(\bar{e}\gamma^\mu e)(\bar{u}\gamma^\nu T^A u)G_{\mu\nu}^A$	
$\mathcal{O}_{e^2d^2\tilde{G}}$	$(\bar{e}\gamma^\mu e)(\bar{d}\gamma^\nu T^A d)\tilde{G}_{\mu\nu}^A$		$\mathcal{O}_{e^2d^2G}$	$(\bar{e}\gamma^\mu e)(\bar{d}\gamma^\nu T^A d)G_{\mu\nu}^A$	
$\mathcal{O}_{l^2u^2\tilde{G}}$	$(\bar{l}\gamma^\mu l)(\bar{u}\gamma^\nu T^A u)\tilde{G}_{\mu\nu}^A$		$\mathcal{O}_{l^2u^2G}$	$(\bar{l}\gamma^\mu l)(\bar{u}\gamma^\nu T^A u)G_{\mu\nu}^A$	
$\mathcal{O}_{l^2d^2\tilde{G}}$	$(\bar{l}\gamma^\mu l)(\bar{d}\gamma^\nu T^A d)\tilde{G}_{\mu\nu}^A$		$\mathcal{O}_{l^2d^2G}$	$(\bar{l}\gamma^\mu l)(\bar{d}\gamma^\nu T^A d)G_{\mu\nu}^A$	
$\mathcal{O}_{q^2e^2\tilde{G}}$	$(\bar{e}\gamma^\mu e)(\bar{q}\gamma^\nu T^A q)\tilde{G}_{\mu\nu}^A$		$\mathcal{O}_{q^2e^2G}$	$(\bar{e}\gamma^\mu e)(\bar{q}\gamma^\nu T^A q)G_{\mu\nu}^A$	

do not interfere with the tree-level SM amplitudes we do not consider them in this study.

III. EXAMPLE UV MODELS

Our primary interest in this paper is the study of the SMEFT from the bottom up, without reference to explicit UV models. Without further experimental guidance as to the form of new physics it is important to fully explore the possible parameter space without the introduction of theoretical biases. Another motivation of our work is to determine what particular experimental datasets, in this case the measurement of the Drell-Yan transverse momentum spectrum at high invariant mass, can teach us about different sectors of the SMEFT. However, we do wish to demonstrate how the operators enumerated in Sec. II can be obtained from explicit UV examples. In particular this study will show that the dimension-eight effects can serve as a useful diagnostic tool to distinguish between different UV states. We will study only example heavy particles here, and briefly mention how they can be embedded into full UV models. More detailed examples of the matching of full UV models to the dimension-eight level are discussed in the literature [49,50]. In our numerical studies we will focus on the example of all-singlet fermion operators. We consider two UV examples that can lead the operators of interest here: a Z' and a vector leptoquark.

A. Right-handed Z' model

We first study a Z' boson coupled to SM singlets. The Lagrangian for this state is given by

$$\mathcal{L}_{Z'} = -\frac{1}{4}Z'^{\mu\nu}Z'_{\mu\nu} + \frac{M_{Z'}^2}{2} - g_{Z'}\sum_f g_R^f \bar{\psi}_f\gamma^\mu P_R\psi_f Z'_\mu. \quad (2)$$

Here, g_R^f denotes the charge of fermion f under the U(1) gauge group, while $g_{Z'}$ is an overall coupling strength of the Z' that is extracted from the charges. Although it is not our

intent here to discuss full UV models we note that this state can be embedded into an anomaly-free U(1) gauge theory with additional fermionic matter that can be taken heavy, in which cases the charges g_R^f are fixed [51]. To determine the Wilson coefficients this Lagrangian leads to for the operators introduced in Sec. II we compute the process $u_1 \bar{u}_2 \to l_3 \bar{l}_4$, to fix C_{eu} and $C^{(1)}_{e^2 u^2 D^2}$. We also compute $u_1 \bar{u}_2 \to l_3 \bar{l}_4 g_5$ to fix $C_{e^2 u^2 \tilde{G}}$. A straightforward calculation of the amplitude expanded in the limit $s \ll M_{Z'}^2$, where $s = (p_1 + p_2)^2$ is the usual partonic Mandelstam invariant, leads to

$$
\begin{aligned}
&\mathcal{M}_{Z'}(u_1 \bar{u}_2 \to l_3 \bar{l}_4) \\
&= -g_{Z'}^2 g_R^q g_R^e \left\{ \frac{1}{M_{Z'}^2} + \frac{s}{M_{Z'}^4} + \cdots \right\} \bar{u}_3 \gamma^\mu P_R v_4 \bar{v}_2 \gamma_\mu P_R u_1.
\end{aligned}
$$

(3)

From this we can read off the Wilson coefficients:

$$
\begin{aligned}
\frac{C_{eu}}{\Lambda^2} &= -\frac{g_{Z'}^2 g_R^u g_R^e}{M_{Z'}^2}, \\
\frac{C^{(1)}_{e^2 u^2 D^2}}{\Lambda^4} &= -\frac{g_{Z'}^2 g_R^u g_R^e}{M_{Z'}^4}.
\end{aligned}
$$

(4)

An identical matching calculation for the down-quark channel fixes C_{ed} and $C^{(1)}_{e^2 d^2 D^2}$. For simplicity, in the rest of this work we neglect these down-quark Wilson coefficients, and focus on the up-quark sector. We note that upon factoring out the dependence on the dimensionful scale $M_{Z'}$ the Wilson coefficients at dimension six and dimension eight are identical in magnitude; there is no suppression of the dimension-eight coefficient.

We now consider the process $u_1 \bar{u}_2 \to l_3 \bar{l}_4 g_5$. At tree level two diagrams contribute, with the additional gluon radiated from either initial-state quark. In both cases there is no hard scale in the virtual quark propagator, and the only expansion possible is for the Z' propagator. This indicates that this process is completely determined by the emission of a gluon after the insertion of either $\mathcal{O}_{eu}$ or $\mathcal{O}_{e^2 u^2 D^2}$, and consequently

$$
\frac{C_{e^2 u^2 \tilde{G}}}{\Lambda^4} = 0.
$$

(5)

B. Vector leptoquark model

We now consider a vector leptoquark coupled to right-handed leptons and quarks. The general Lagrangian for such a state is given in Refs. [52,53]. We assume a leptoquark coupled to QCD, and coupled to right-handed up quarks and leptons, for which the Lagrangian takes the form

$$
\begin{aligned}
\mathcal{L}_U = &-\frac{1}{2} G_{\mu\nu}^{i\dagger} G_i^{\mu\nu} + M_U^2 U_\mu^{i\dagger} U_i^\mu + h_U(\bar{\psi}_u^i \gamma^\mu P_R l) U_\mu^i \\
&- i g_s (1 - \kappa_U) U_\mu^\dagger T^a U_\nu \mathcal{G}^{a\mu\nu}.
\end{aligned}
$$

(6)

Here, the Roman indices i, j denote color indices in the fundamental representation. The quantity $\mathcal{G}$ denotes the field strength tensor of the SM gluon field. The field strength tensor and covariant derivatives of the leptoquark are given by

$$
\begin{aligned}
D_\mu^{ik} &= \partial_\mu \delta^{ij} - i g_s T_A^{ij} G_\mu^A, \\
G_{\mu\nu}^i &= D_\mu^{ik} U_{\nu k} - D_\nu^{ik} U_{\mu k}.
\end{aligned}
$$

(7)

The coupling κ_U is related to the magnetic moment of the leptoquark. We note that it has been argued that complete leptoquark models generically contain Z' bosons as well [54]. These lead to Wilson coefficient contributions similar to those discussed in Sec. III A, and are not explicitly considered here since we focus instead on the unique aspects of the vector leptoquark. We also note that leptoquarks have received renewed interest recently due to their possible role in resolving outstanding flavor anomalies [54].

We compute the same partonic processes as before to determine the Wilson coefficients. A straightforward calculation of $u_1 \bar{u}_2 \to l_3 \bar{l}_4$ leads, after a Fierz rearrangement, to the amplitude

$$
\mathcal{M}_U = \frac{h_U^2}{M_U^2} \bar{u}_3 \gamma^\mu P_R v_4 \bar{v}_2 \gamma_\mu P_R u_1 \left\{ 1 + \frac{t}{M_U^2} + \cdots \right\}.
$$

(8)

In order to match this amplitude to the operators considered previously we have to decompose the t in the numerator according to the contributions to the operators $\mathcal{O}^{(1)}_{e^2 u^2 D^2}$ and $\mathcal{O}^{(2)}_{e^2 u^2 D^2}$. Doing so we arrive at the Wilson coefficients for the operators considered here:

$$
\begin{aligned}
\frac{C_{eu}}{\Lambda^2} &= \frac{h_U^2}{M_U^2}, \\
\frac{C^{(1)}_{e^2 u^2 D^2}}{\Lambda^4} &= -\frac{h_U^2}{4 M_U^4}.
\end{aligned}
$$

(9)

We note that the leptoquark also contributes to $C^{(2)}_{e^2 u^2 D^2}$, unlike the Z'. This is the first example of how dimension-eight coefficients can help distinguish between models. A nonzero $C^{(2)}_{e^2 u^2 D^2}$, which can be determined from an analysis of the type considered in [43], would disfavor a Z' UV model.

Another difference between the Z' and leptoquark comes when we consider the gluonic process $u_1 \bar{u}_2 \to l_3 \bar{l}_4 g_5$. There is a trilinear coupling UUg which means that the gluon can be emitted from the t-channel leptoquark. Upon expanding around large M_U, this leads to a local

contribution described by the operator $\mathcal{O}_{e^2u^2\tilde{G}}$. We relegate details of the matching calculation to Appendix A, and simply note the result here:

$$\frac{C_{e^2u^2\tilde{G}}}{\Lambda^4} = -\frac{h_U^2 g_s(1-\kappa_U)}{2M_U^4}. \tag{10}$$

This again illustrates our point that measurement of the complete sector of four-fermion operators can discriminate between UV models. $C_{e^2u^2\tilde{G}}$ is not induced in Z' models, but it is in vector leptoquark models. We note as well the following points regarding the Wilson coefficients found in this calculation: upon removal of the dimensionful quantity M_U the dimension-six and dimension-eight Wilson coefficients are similar in size, and the coupling $C_{e^2u^2\tilde{G}}$ can be larger than $C_{e^2u^2D^2}^{(1)}$ for negative κ_U. We will refer to both of these points during discussions in later sections. It has been pointed out that positivity bounds on dimension-eight Wilson coefficients can be derived from the underlying principles of quantum field theory [55,56]. We note that the simplest elastic positivity constraints do not restrict the parameter space of the $C_{e^2u^2D^2}^{(1)}$ and $C_{e^2u^2\tilde{G}}$ coefficients [48].

IV. MOTIVATION FOR THE DOUBLY DIFFERENTIAL DRELL-YAN DISTRIBUTION

A main motivation of our work is to demonstrate that LHC datasets not originally intended as new physics searches can be sensitive, sometimes in novel ways, to the SMEFT parameter space, and therefore have unexpected sensitivity to physics beyond the SM. As an example, the CMS experiment has measured the lepton-pair transverse momentum spectrum in several invariant bins ranging up to 1 TeV [46]. Results are normalized to the Z-peak region in order to reduce the dependence of the measurement on systematic errors. The measurement was performed at 13 TeV with 36.3 fb^{-1} of integrated luminosity. Before performing fits of this data to the SMEFT framework to demonstrate its sensitivity we describe why this dataset is particularly interesting to probe the set of operators described in the previous sections. We show below in Figs. 1–3 the ratio of SMEFT corrections to the SM result as a function of transverse momentum for the two highest invariant mass bins available in the measurement of Ref. [46], turning on the three Wilson coefficients C_{eu}, $C_{e^2u^2D^2}$, and $C_{e^2u^2\tilde{G}}$ separately (we drop the superscript on $C_{e^2u^2D^2}$ henceforth, since we do not consider $C_{e^2u^2D^2}^{(2)}$ further in this work). The SM has been computed at next-to-leading order (NLO) in QCD using the program MCFM [57]. We have set each Wilson coefficient to unity when making these plots. Since the coefficient C_{eu} contributes first at dimension six it shows the largest deviations from the SM result. However, the deviation does not vary significantly with p_T. This is consistent with the structure of the dimension-six EFT correction, which does not depend on the momentum flow into the effective vertex. The deviations for both $C_{e^2u^2D^2}$ and $C_{e^2u^2\tilde{G}}$ increase with p_T, consistent with the fact that the effective vertex depends on the momentum flow. While the $C_{e^2u^2D^2}$ deviation increases moderately with transverse momentum, $C_{e^2u^2\tilde{G}}$ increases rapidly. The EFT vertex for this operator is directly proportional to the gluon momentum, and therefore the p_T of the lepton pair by momentum conservation. This is the motivation for the analysis of this dataset within the SMEFT framework: the p_T distribution offers additional sensitivity to gluonic operators not present with invariant mass distributions alone.

V. CALCULATIONAL FRAMEWORK AND FITS TO THE CURRENT DATA

We begin by performing a fit of the current CMS measurement to the SMEFT framework. Although we will
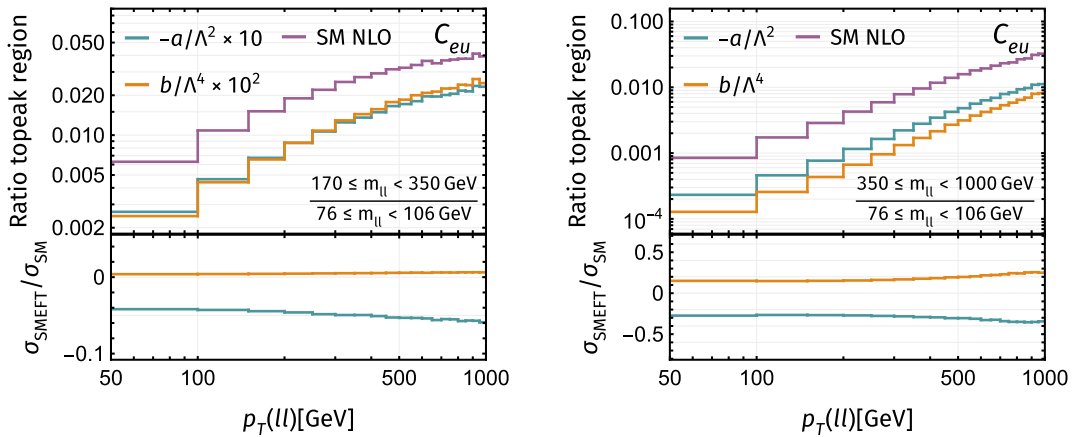

FIG. 1. Ratio of the SMEFT correction assuming nonzero C_{eu} to the SM result as a function of p_T for the upper two invariant mass bins of the CMS measurement [46]. The results have been normalized to the Z-peak region. The a term corresponds to the linear $1/\Lambda^2$ correction while the b term corresponds to the quadratic $1/\Lambda^4$ correction.
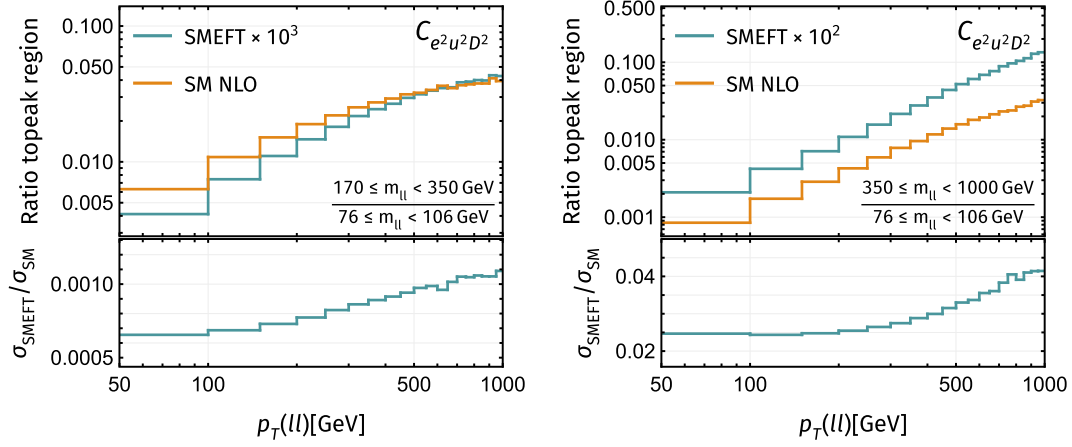
FIG. 2. Ratio of the SMEFT correction assuming nonzero $C_{e^2u^2D^2}$ to the SM result as a function of p_T for the upper two invariant mass bins of the CMS measurement [46]. The results have been normalized to the Z-peak region.

find that there is limited sensitivity to the gluonic operators at this point, we find it useful to quantify the current sensitivity and to also establish our notation for later sections. We will for simplicity focus on operators that contain right-handed fermions only. Possible UV models leading to such operators were discussed in Sec. III, where we introduced the possibility of using $C_{e^2u^2\tilde{G}}$ to distinguish between them. We focus on invariant mass bins above the Z peak, since the SMEFT-induced corrections grow with energy. We also focus on transverse momentum bins above 50 GeV, following the same logic as for invariant mass, and also to avoid phase space regions where p_T resummation may play a role. This leaves us with the eight bins shown below in Table IV. We note that these bins are normalized by the experiment to the Z-peak region $76 \leq m_{ll} \leq 106$ GeV.

We now give the details of our calculational framework. We compute the SM cross section at NLO in QCD using the MCFM program [57]. We use the NNPDF 3.1 NLO parton distribution functions (PDF) [58]. To compute the

PDF errors we follow the standard procedure for Monte Carlo replica sets [58]. To estimate the error arising from higher-order QCD corrections we set the renormalization and factorization scales to the central value

$$\mu_0 = \sqrt{m_{ll}^2 + p_T(ll)^2} \tag{11}$$

and vary them around this value in an uncorrelated way according to

$$\frac{1}{2} \leq \mu_{R,F}/\mu_0 \leq 2, \qquad \frac{1}{2} \leq \mu_R/\mu_F \leq 2. \tag{12}$$

We find the largest variation within this range, and form a symmetric scale uncertainty using this largest variation. We note that this technique leads to slightly more conservative errors than the usual approach, in which the width of the scale variation band without symmetrization is used. We note that the PDF uncertainties are strongly correlated
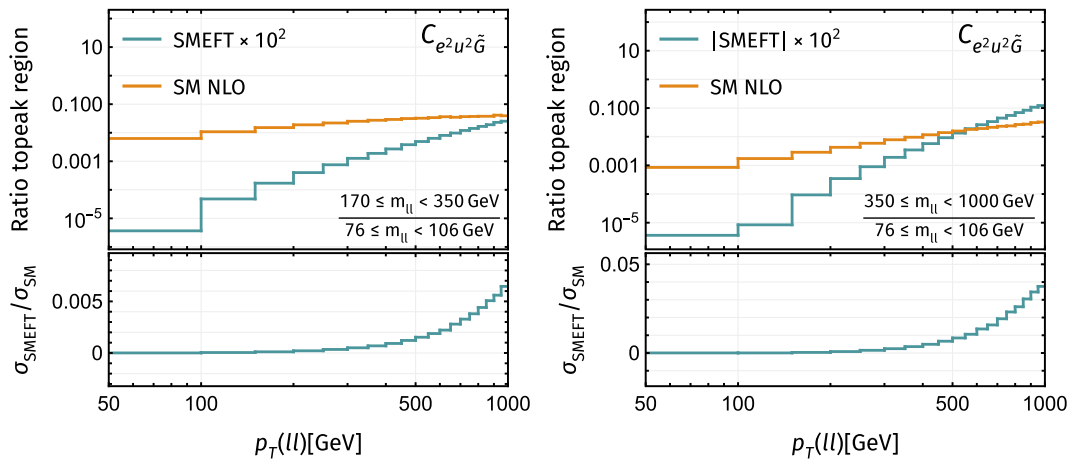

FIG. 3. Ratio of the SMEFT correction assuming nonzero $C_{e^2u^2\tilde{G}}$ to the SM result as a function of p_T for the upper two invariant mass bins of the CMS measurement [46]. The results have been normalized to the Z-peak region.

TABLE IV. Bins in invariant mass and transverse momentum used in Ref. [46]. The p_T values refer to bin boundaries. There is an upper cut of 1 TeV on p_T in all bins.

$m_{ll} = [106, 170]$ GeV	$p_T = [52, 85, 160, 250]$ GeV
$m_{ll} = [170, 350]$ GeV	$p_T = [52, 160]$ GeV
$m_{ll} = [350, 1000]$ GeV	$p_T = [52, 160]$ GeV

between different bins. We assume that the scale uncertainties are uncorrelated between bins.

Since this data involves bins at high energies, it is important to quantify the effect of electroweak Sudakov corrections. We are unaware of a publicly available code that computes the electroweak Sudakov logarithms as a function of p_T off the Z peak. To estimate the impact of these corrections we compute the next-to-leading-logarithmic electroweak Sudakov corrections [59] for each invariant mass bin integrated inclusively over p_T, apply this correction to each of the p_T bins for that invariant mass, and assign half of this correction as an additional theoretical error. The lower bin boundaries in p_T, which provide the largest contributions to each bin, do not go above 160 GeV. For the higher two invariant mass bins in our analysis this p_T value is less than the invariant mass. Since the Sudakov logarithms grow with the Mandelstam variable s that enters the process, and s is dominated by the lepton invariant mass for the reason stated above, we believe that this is a reasonable estimate. We note that their effect ranges from 1% to 4% as we increase the invariant mass bin, and has little effect on the quality of the fit to data. For the SMEFT cross section, we work at leading order in QCD.

We use the experimental uncertainties as provided by the CMS collaboration. These uncertainties range from 1.5% to 8.9%, increasing with both invariant mass and transverse momentum, and contain a mix of both correlated and uncorrelated errors. The systematic uncertainties are dominant in the high invariant mass bins. We define a χ^2 test

to quantify the deviation of the SMEFT cross sections from the SM:

$$\chi^2 = \sum_{i,j}^{\# \text{ of bins}} \frac{(\sigma_i^{\text{SM}} - \sigma_i^{\text{SMEFT}})(\sigma_j^{\text{SM}} - \sigma_j^{\text{SMEFT}})}{\Delta\sigma_{ij}^2}, \quad (13)$$

where $\Delta\sigma_{ij}^2$ signifies the error matrix composed of both theoretical and experimental uncertainties. We then extract the 95% C.L. bounds of the Wilson coefficients based on χ^2 fits. Before studying the SMEFT we note that the SM furnishes an acceptable fit to the data, with a χ^2 per degree of freedom of 1.4.

We consider turning on only a single operator at a time, turning on pairs of operators, and turning on all three. Figure 4 shows the 95% C.L. ellipses of C_{eu} together with either $C_{e^2u^2\tilde{G}}$ or $C_{e^2u^2D^2}$, as well as the bounds with only one Wilson coefficient enabled. Although the data is less sensitive to $C_{e^2u^2\tilde{G}}$ than to $C_{e^2u^2D^2}$, the circular nature of the ellipse in Fig. 4 indicates that $C_{e^2u^2\tilde{G}}$ has little correlation with C_{eu}. The stretched narrow ellipse of C_{eu} and $C_{e^2u^2D^2}$ shows a strong correlation between these two operators, with the effects of the two coefficients indistinguishable with the current data.

A potential issue that must be addressed when studying these constraints is the convergence of the EFT expansion. As can be seen from the examples in Sec. III we have the following rough relations between parameters in UV models and those appearing the SMEFT:

$$\frac{C}{\Lambda^2} \sim \frac{g^2}{M^2} \quad \text{for dimension six,}$$

$$\frac{C}{\Lambda^4} \sim \frac{g^2}{M^4} \quad \text{for dimension eight.} \quad (14)$$

We want any potential resonance to lie above the scales probed experimentally. We take this constraint to be
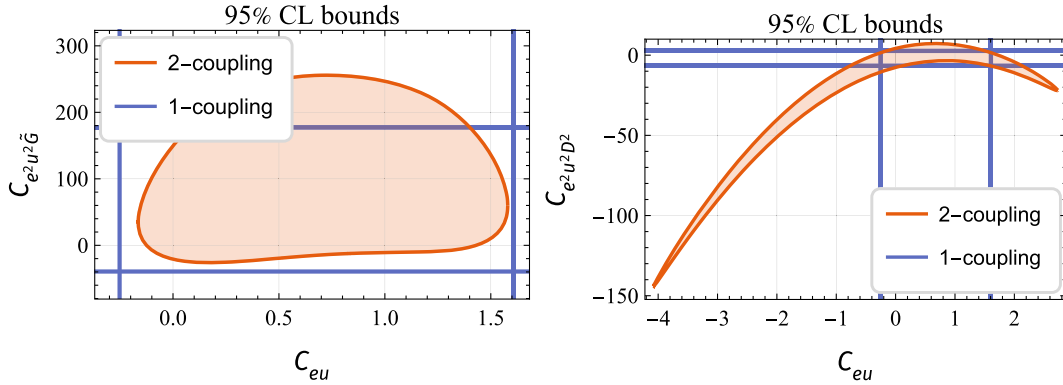

FIG. 4. 95% C.L. ellipses for the Wilson coefficients C_{eu}, $C_{e^2u^2\tilde{G}}$ and $C_{e^2u^2D^2}$ using current CMS data. For each diagram, one dimension-six operator and one dimension-eight operator are enabled. The blue lines denote the bounds with only one of the operators enabled. The energy scale Λ is set to 2 TeV.

TABLE V. 95% C.L. bounds for the Wilson coefficients C_{eu}, $C_{e^2u^2\tilde{G}}$ and $C_{e^2u^2D^2}$ from current CMS data. The first column shows the bounds assuming one operator is enabled at a time. The second column shows the bounds on a given coefficient with the other enabled operators allowed to vary as well. The third column shows these bounds with the dimension-eight coefficients restricted according to the discussion in the main text.

Wilson coefficient	Single coupling	Marginalized	Marginalized[*]
	C_{eu} and $C_{e^2u^2D^2}$		
C_{eu}	[−0.252, 1.61]	[−3.44, 2.37]	[−3.34, 2.33]
$C_{e^2u^2D^2}^{(1)}$	[−6.65, 2.64]	[−90.9, 47.1]	[−83.7, 41.9]
	C_{eu} and $C_{e^2u^2\tilde{G}}$		
C_{eu}	[−0.252, 1.61]	[−0.062, 1.45]	[−0.071, 1.45]
$C_{e^2u^2\tilde{G}}$	[−39.4, 177]	[−16.9, 220]	[−8.95, 197]
	All three operators		
C_{eu}	[−0.252, 1.61]	[−3.75, 2.60]	[−3.49, 2.46]
$C_{e^2u^2D^2}$	[−6.65, 2.64]	[−113, 57.2]	[−92.7, 42.9]
$C_{e^2u^2\tilde{G}}$	[−39.4, 177]	[−6.35, 251]	[3.94, 201]

$M > 1$ TeV, the highest scale probed in this data, which translates to the bounds

$$\frac{\Lambda}{\sqrt{|C|}} > \frac{1\text{ TeV}}{g} \quad \text{for dimension six,}$$

$$\frac{\Lambda}{\sqrt[4]{|C|}} > \frac{1\text{ TeV}}{\sqrt{g}} \quad \text{for dimension eight.} \quad (15)$$

The exact numerical value of this constraint depends on the coupling g, and therefore on the details of the UV model. We take the strong coupling limit $g = \sqrt{4\pi}$, leading to the least stringent constraint, in order to avoid ruling out allowed parameter space. This slightly restricts the allowed parameter space in the joint C_{eu}, $C_{e^2u^2\tilde{G}}$ in Fig. 4. It does not affect the joint C_{eu}, $C_{e^2u^2D^2}$ result.

We list the 95% C.L. bounds with only one operator enabled in Table V. We also calculate the bounds with multiple operators by marginalizing over the couplings. We further impose the effective scale constraint on the marginalized bounds in the final column of this table. The results are also listed in Table V. We observe that $C_{e^2u^2D^2}$ has a strong impact to the bounds on C_{eu}, while $C_{e^2u^2\tilde{G}}$ only mildly changes this limit. Turning on $C_{e^2u^2D^2}$ significantly weakens the bounds on C_{eu}, an effect also observed with LHC invariant mass distributions in [44]. In general there is limited sensitivity of this dataset to the dimension-eight coefficients with Wilson coefficients reaching $\mathcal{O}(100)$ still allowed.

VI. FITS TO SIMULATED HL-LHC DATA

We found in the previous section that the current data shows little sensitivity to $C_{e^2u^2\tilde{G}}$. We consider next the

potential the high-luminosity LHC to probe this SMEFT parameter space through a similar analysis. Since no data is yet available we resort to pseudodata generated with the NLO SM cross section. The HL-LHC pseudodata is generated under similar conditions as the CMS measurement [46]. We assume the center-of-mass energy $\sqrt{s} = 14$ TeV, an integrated luminosity 3 ab^{-1} and a dilepton transverse momentum cut $p_T \geq 100$ GeV. Since the eventual HL-LHC binning is unknown, we consider two possible sets of bins in the dilepton invariant mass and transverse momentum. The binning for the dilepton invariant mass m_{ll} is motivated by the simulation in [60]. For the binning of the dilepton transverse momentum $p_T(ll)$, we enforce that the relative statistical uncertainty of each bin cannot exceed 10%. As such, we discard the highest m_{ll} bin in [60] where $2600 \leq m_{ll} \leq 14000$ GeV. Next, two different binning strategies are applied: a coarse binning where the relative statistical uncertainty of each bin should be smaller than 5% if possible[1]; a fine binning where the relative statistical uncertainty of each bin must be smaller than 10%. We show the explicit bins used in Appendix B, with the coarse binning shown in Table VI and the fine binning in Table VII.

We assume that the pseudodata is affected by three sources of experimental uncertainties: the statistical uncertainty $\Delta\sigma_{\text{stat}}$, the uncorrelated systematic uncertainty $\Delta\sigma_{\text{uncorr}}$ and the fully correlated systematic uncertainty $\Delta\sigma_{\text{corr}}$. We construct the cross section of bin b using

$$\sigma_b^{\text{pseudo}} = \sigma_b^{\text{SM}} + r_b\sqrt{\Delta\sigma_{\text{stat},b}^2 + \Delta\sigma_{\text{uncorr},b}^2} + r'\sqrt{\Delta\sigma_{\text{corr},b}^2},$$

$$(16)$$

where r_b and r' are random numbers generated with a normal distribution of mean 0 and standard deviation 1. For uncorrelated uncertainties, a separate random number is chosen for each bin. For the correlated uncertainty, a single random number r' is used across all bins. We assume the relative uncorrelated systematic uncertainty $\Delta\sigma_{\text{uncorr},b}/\sigma_b^{\text{SM}} = 1\%$ and the relative correlated systematic uncertainty $\Delta\sigma_{\text{corr},b}/\sigma_b^{\text{SM}} = 2\%$. These choices are consistent with the current values found in the CMS measurement [46]. We normalize the cross section to the Z-peak region $76 \leq m_{ll} \leq 106$ GeV. We emphasize that the relative uncorrelated uncertainty of the Z-peak region is still assumed as 1%, but any correlated uncertainties in this region are disregarded. The error matrix contains the experimental uncertainties, the theoretical PDF uncertainties and the scale uncertainties. The experimental part is constructed with $\Delta\sigma_{\text{stat}}$, $\Delta\sigma_{\text{uncorr}}$ and $\Delta\sigma_{\text{corr}}$. The PDF and

---

[1]The only exception is when $2000 \leq m_{ll} \leq 2600$ GeV, where the largest possible p_T bin is $100 \leq p_T \leq 7000$ GeV. The relative statistical uncertainty of this bin is larger than 5%, but still smaller than 10%.

scale uncertainties only contain SM contributions. For each set of random numbers r_b and r' generated, we perform a χ^2 fit as described in Sec. V. Each set of random numbers signifies one pseudoexperiment, and for each pseudoexperiment e, a set of best-fit Wilson coefficients $\{C_{i,e}\}$ is obtained. A total number of 1000 pseudoexperiments are evaluated, and the average best-fit values are obtained by

$$
\begin{pmatrix} \bar{C}_1 \\ \bar{C}_2 \\ \vdots \end{pmatrix} = \left[\sum_{e=1}^{N_{\exp}} (V^{-1})_e \right]^{-1} \left[\sum_{e=1}^{N_{\exp}} (V^{-1})_e \begin{pmatrix} \bar{C}_{1,e} \\ \bar{C}_{2,e} \\ \vdots \end{pmatrix} \right]. \quad (17)
$$

The covariance matrix for each pseudoexperiment is given by

$$
(V^{-1})_{ij} = \frac{1}{2} \frac{\partial \chi^2}{\partial C_i \partial C_j}. \quad (18)
$$

The 95% C.L. bounds on the Wilson coefficients are extracted by

$$
\begin{pmatrix} C_1 - \bar{C}_1 \\ C_2 - \bar{C}_2 \\ \vdots \\ C_N - \bar{C}_N \end{pmatrix}^T V^{-1} \begin{pmatrix} C_1 - \bar{C}_1 \\ C_2 - \bar{C}_2 \\ \vdots \\ C_N - \bar{C}_N \end{pmatrix} = \Delta\chi^2, \quad (19)
$$

where N is the number of Wilson coefficients. For $N = 1$, 2, 3, $\Delta\chi^2 = 3.841$, 5.991, 7.815, respectively. The average inverse covariance matrix is defined as

$$
V^{-1} = \frac{1}{N_{\exp}} \sum_{e=1}^{N_{\exp}} (V^{-1})_e. \quad (20)
$$

We now consider fits to the pseudodata with one, two or three Wilson coefficients enabled. Figure 5 shows the 95% C.L. bounds on C_{eu} and $C_{e^2u^2\tilde{G}}$ with either one, two or all three operators enabled. Figure 6 shows the 95% C.L. bounds on C_{eu} and $C_{e^2u^2D^2}$. The round shape of the ellipse in Fig. 5 and the narrow shape in Fig. 6 confirm what we learned from the existing CMS data: there is little correlation between C_{eu} and $C_{e^2u^2\tilde{G}}$, and a stronger correlation between C_{eu} and $C_{e^2u^2D^2}$. We observe that the inclusion of $C_{e^2u^2D^2}$ loosens the bounds on C_{eu}, while the inclusion of $C_{e^2u^2\tilde{G}}$ has much less impact on C_{eu}. This is consistent with the previous observation that the inclusion of transverse momentum data provides a separate handle on the gluonic operators. We also observe that the fine binning leads to tighter bounds than the coarse binning. The gray area in Fig. 5 shows the region where the effective scale constraint of the dimension-eight operator is violated, and the EFT expansion is no longer valid. We demand the constraint $M > 3$ TeV, consistent with the upper limit of our invariant mass binning, so that the dimension-eight Wilson coefficients must satisfy $\frac{\Lambda}{\sqrt[4]{|C|}} > \frac{3\ \text{TeV}}{\sqrt{g}}$. We use dark gray to indicate $g = \sqrt{4\pi}$, a choice discussed in the previous section.

We observe that the HL-LHC data has the potential to measure these three couplings separately. Although some correlation between C_{eu} and $C_{e^2u^2D^2}$ remains, it is weaker than found with the current data, and only weakens the C_{eu} bounds by a factor of 2. Referring to the fine binning
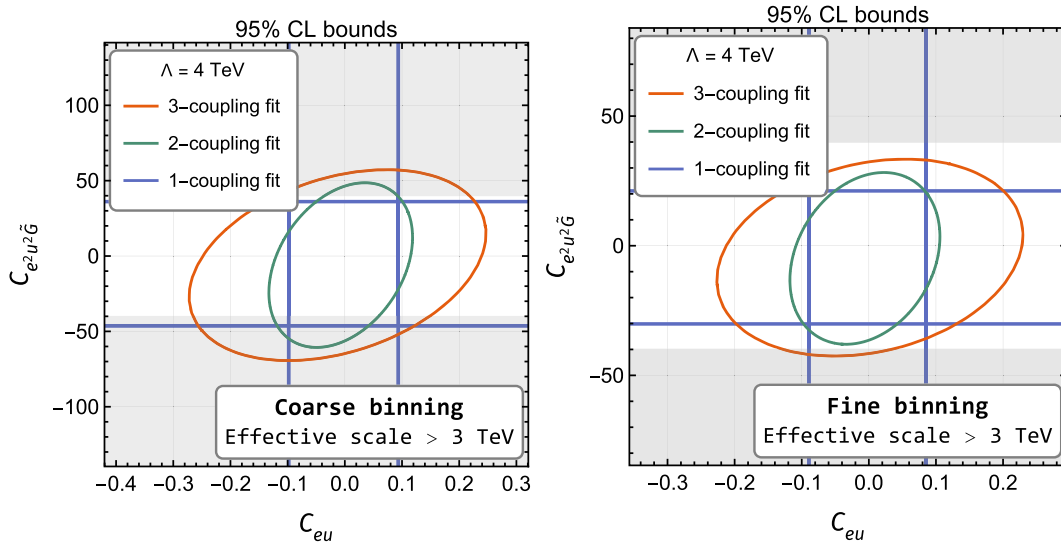

FIG. 5. 95% C.L. ellipses for the Wilson coefficients C_{eu} and $C_{e^2u^2\tilde{G}}$ at the HL-LHC. The left diagram shows the bounds with the coarse binning, while the right one shows the bounds with the fine binning. The blue lines denote the bounds with only one of the operators enabled, the green line denotes the bounds with two operators enabled, and the orange line denotes the bounds with $C_{e^2u^2D^2}$ also enabled. The energy scale Λ is set to 4 TeV, and the effective scale constraint is set to 3 TeV. The dark gray area shows the effective scale constraint with $g = \sqrt{4\pi}$.
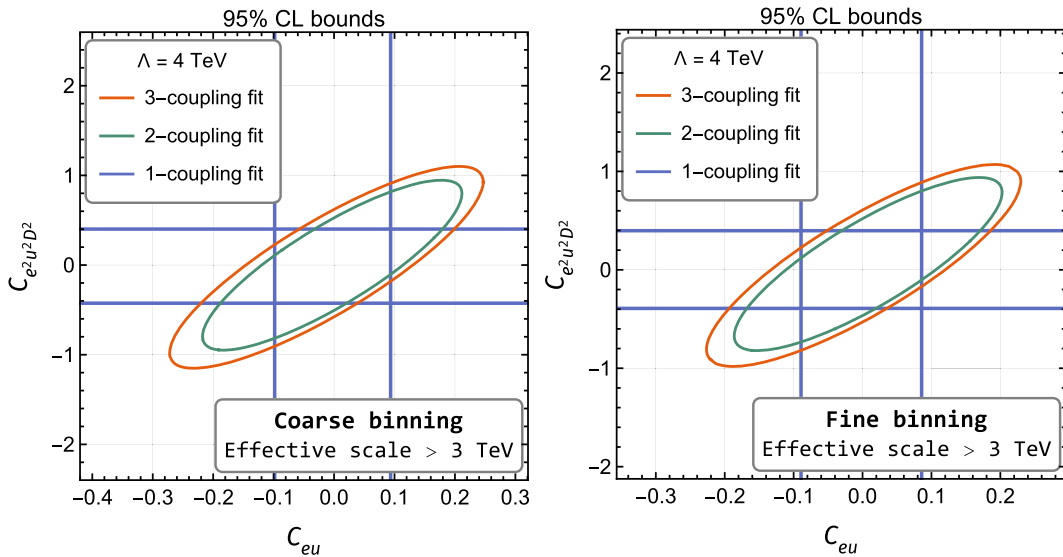
FIG. 6. 95% C.L. ellipses for the Wilson coefficients C_{eu} and $C_{e^2u^2D^2}$ at the HL-LHC. The left diagram shows the bounds with the coarse binning, while the right one shows the bounds with the fine binning. The blue lines denote the bounds with only one of the operators enabled, the green line denotes the bounds with two operators enabled, and the orange line denotes the bounds with $C_{e^2u^2\tilde{G}}$ also enabled. The energy scale Λ is set to 4 TeV, and the effective scale constraint is set to 3 TeV.
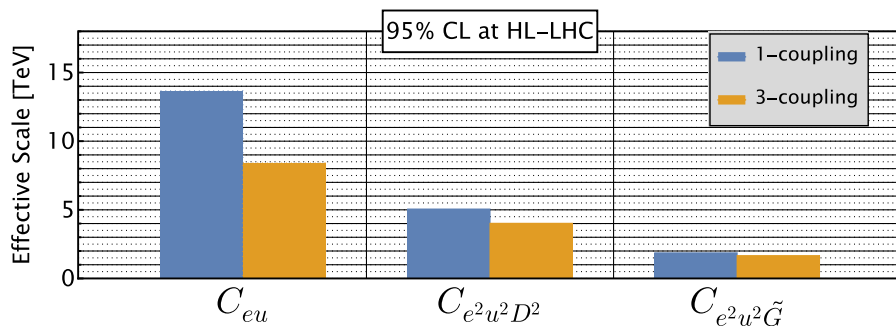

FIG. 7. 95% C.L. bounds on the effective scale of all three operators when only a single one is enabled, and when all three are simultaneously enabled, at the HL-LHC. We recall that the effective scale is defined as $\Lambda/\sqrt{C}$ for dimension six and $\Lambda/\sqrt[4]{C}$ for dimension eight.

results, we observe that there is a hierarchy in the sensitivities to these three coefficients: C_{eu} values of $\mathcal{O}(0.1)$ can be probed, $C_{e^2u^2D^2}$ values of $\mathcal{O}(1)$ can be probed, while the sensitivity to $C_{e^2u^2\tilde{G}}$ drops to $\mathcal{O}(10)$. Considering that the EFT expansion parameter is chosen as $\Lambda = 4$ TeV, these results indicate sensitivity reaching into the multi-TeV region for all three operators. We recall that $C_{e^2u^2\tilde{G}}$ can be enhanced in certain regions of leptoquark parameter space, as discussed in Sec. III B. This indicates that the HL-LHC doubly differential Drell-Yan data can serve as a useful diagnostic tool for realistic UV states. We summarize the potential of the HL-LHC by showing in Fig. 7 the effective UV scales that can be probed for each parameter when only a single coupling is turned on, and when all three are turned on. We recall from Eq. (14) that the effective scale is related to the heavy resonance mass in the UV theory scaled by either g (for dimension six) or $\sqrt{g}$

(for dimension eight) as shown in Sec. V. Effective scales approaching 10 TeV can be probed for C_{eu}, while sensitivities reaching several TeV are possible for the dimension-eight coefficients.

VII. CONCLUSIONS

In this paper we have studied probes of the semileptonic four-fermion sector of the SMEFT that are possible with neutral-current Drell-Yan measurements at the LHC. We have extended previous studies by including dimension-eight operators with additional gluon field-strength tensors. These operators directly modify the high transverse momentum region in Drell-Yan production. A motivation for this work is a recent CMS measurement of the transverse momentum distribution for the Drell-Yan process further binned in invariant mass. Although this work was

intended primarily as a QCD study, it has novel beyond standard model (BSM) sensitivity as well, and provides direct access to this previously unexplored sector of the SMEFT.

To motivate our study we have demonstrated that example UV models can lead to very different patterns of Wilson coefficients for these gluonic operators; some states generate potentially sizable Wilson coefficients for this dimension-eight operator, while others do not generate these operators. This ability to discriminate between different UV completions of the SMEFT would be missed if the SMEFT expansion was truncated at the dimension-six level; the example models considered here would only match to a single dimension-six operator. Measurement of the entire suite of semileptonic four-fermion coefficients through dimension eight can therefore help distinguish between different models of new physics. We have considered fits of the SMEFT framework to both the current CMS measurement and to simulated future HL-LHC data. While the current data shows little sensitivity to the gluonic

operator, there are good prospects for probing this effect with future data. We encourage this measurement to be performed with future data and its BSM potential to be further explored.

APPENDIX A: LEPTOQUARK MATCHING

In this Appendix we study the matching of a vector leptoquark to the SMEFT in the process $u_1 \bar{u}_2 \to l_3 \bar{l}_4 g_5$ in order to determine the Wilson coefficient $C_{e^2 u^2 \tilde{G}}$. There are three contributing diagrams, two where the gluon is emitted from an initial quark, and one where it is emitted from the leptoquark. The amplitudes for each diagram take the form

$$
\begin{aligned}
i\mathcal{M}_1 &= \frac{i h_U^2 g_s T_{ij}^a}{t_{15}(t_{24}-M_U^2)} \left\{ \bar{v}_2^i \gamma_\mu P_R v_4 \bar{u}_3 \gamma^\mu P_R u_1^j \frac{2 p_1 \cdot \epsilon_5^a}{t_{15}} - \bar{v}_2^i \gamma_\mu P_R v_4 \bar{u}_3 \gamma^\mu \not{p}_5 \not{\epsilon}_5^a P_R u_1^j \right\} \\
i\mathcal{M}_2 &= -\frac{i h_U^2 g_s T_{ij}^a}{t_{25}(t_{13}-M_U^2)} \left\{ \bar{v}_2^i \gamma_\mu P_R v_4 \bar{u}_3 \gamma^\mu P_R u_1^j \frac{2 p_2 \cdot \epsilon_5^a}{t_{25}} - \bar{v}_2^i \gamma^\mu \not{\epsilon}_5^a \not{p}_5 P_R v_4 \bar{u}_3 \gamma_\mu P_R u_1^j \right\} \\
i\mathcal{M}_3 &= \frac{i h_U^2 g_s T_{ij}^a}{(t_{13}-M_U^2)(t_{24}-M_U^2)} \{ \bar{v}_2^i \gamma_\mu P_R v_4 \bar{u}_3 \gamma^\mu P_R u_1^j (p_1 - p_2 - p_3 + p_4) \cdot \epsilon_5^a \\
&\quad + (1-\kappa_U) \bar{v}_2^i \gamma^\nu P_R v_4 \bar{u}_3 \gamma^\mu P_R u_1^j [p_{5\mu} \epsilon_{5\nu}^a - p_{5\nu} \epsilon_{5\mu}^a] \}
\end{aligned} \tag{A1}
$$

using the Lagrangian presented in Sec. III. We note that $t_{15} = (p_1 - p_5)^2$, etc. We can expand these expressions in the large M_U. The dimension-six contribution comes from the first two diagrams; the t_{15}, t_{25} in the denominator make it clear that these match to the emission of a gluon off of a dimension-six four-fermion operator. We focus here on the expansion to dimension eight. The contribution from each diagram is

$$
\begin{aligned}
i\mathcal{M}_1^{(8)} &= -\frac{i h_U^2 g_s T_{ij}^a}{M_U^4} \left\{ \frac{2 p_1 \cdot \epsilon_5^a t_{24}}{t_{15}} \bar{v}_2^i \gamma_\mu P_R v_4 \bar{u}_3 \gamma^\mu P_R u_1^j - \frac{t_{24}}{t_{15}} \bar{v}_2^i \gamma_\mu P_R v_4 \bar{u}_3 \gamma^\mu \not{p}_5 \not{\epsilon}_5^a P_R u_1^j \right\} \\
i\mathcal{M}_2^{(8)} &= \frac{i h_U^2 g_s T_{ij}^a}{M_U^4} \left\{ \frac{2 p_2 \cdot \epsilon_5^a t_{13}}{t_{25}} \bar{v}_2^i \gamma_\mu P_R v_4 \bar{u}_3 \gamma^\mu P_R u_1^j - \frac{t_{13}}{t_{25}} \bar{v}_2^i \gamma^\mu \not{\epsilon}_5^a \not{p}_5 P_R v_4 \bar{u}_3 \gamma_\mu P_R u_1^j \right\} \\
i\mathcal{M}_3^{(8)} &= \frac{i h_U^2 g_s T_{ij}^a}{M_U^4} \{ \bar{v}_2^i \gamma_\mu P_R v_4 \bar{u}_3 \gamma^\mu P_R u_1^j (p_1 - p_2 - p_3 + p_4) \cdot \epsilon_5^a + (1-\kappa_U) \bar{v}_2^i \gamma^\nu P_R v_4 \bar{u}_3 \gamma^\mu P_R u_1^j [p_{5\mu} \epsilon_{5\nu}^a - p_{5\nu} \epsilon_{5\mu}^a] \}. \quad \text{(A2)}
\end{aligned}
$$

It is clear that the first two diagrams come from emitting a gluon from dimension-eight four-fermion operators, and do not match to a local operator with a gluon. This can be seen from the t_{15}, t_{25} in the denominator. The same is true for the first term of $\mathcal{M}_3^{(8)}$. This can be determined most simply by demanding gauge invariance: the amplitude must vanish upon replacing $\epsilon_5 \to p_5$. The first two diagrams are not

invariant themselves. Only upon adding the first term of diagram three is gauge invariance satisfied. The last term of $\mathcal{M}_3^{(8)}$ is separately gauge invariant.

This leaves the last term of $\mathcal{M}_3^{(8)}$ to match to a local dimension-eight operator $q\bar{q}l\bar{l}g$. To simplify this we apply the following Fierz identity:

$$\bar{v}_2^i \gamma^\nu P_R v_4 \bar{u}_3 \gamma^\mu P_R u_1^j$$
$$= \frac{1}{2} \{ -g^{\mu\nu} g^{\rho\sigma} + g^{\mu\rho} g^{\nu\sigma} + g^{\mu\sigma} g^{\nu\rho} - i\epsilon^{\mu\nu\rho\sigma} \}$$
$$\times \bar{v}_2^i \gamma^\sigma P_R u_1^j \bar{u}_3 \gamma^\rho P_R v_4. \qquad (A3)$$

Only the antisymmetric term survives when we plug this into the amplitude, leaving us with

$$i\mathcal{M}_{\text{local}}^{(8)} = \frac{h_U^2 (1-\kappa_U) g_s T_{ij}^a}{M_U^4} p_{5\mu} \epsilon_{5\nu}^a \epsilon^{\mu\nu\rho\sigma} \bar{v}_2^i \gamma^\sigma P_R u_1^j \bar{u}_3 \gamma^\rho P_R v_4. \qquad (A4)$$

This matches to the local dimension-eight operator

$$\bar{e} \gamma^\mu e \bar{u} \gamma^\nu T^a u \tilde{G}_{\mu\nu}^a \qquad (A5)$$

with the Wilson coefficient:

$$C_{e^2 u^2 \tilde{G}} = -\frac{h_U^2 (1-\kappa_U) g_s}{2M_U^4}. \qquad (A6)$$

APPENDIX B: HL-LHC BINNING

We present in this Appendix the two choices for HL-LHC binning used in our analysis: a coarse binning where the relative statistical uncertainty of each bin is smaller than 5%, and a fine binning where the relative statistical uncertainty of each bin must be smaller than 10%. The coarse binning is shown in Table VI, while the fine binning is shown in Table VII.

TABLE VI. The coarse binning where the relative statistical uncertainty of each bin should be smaller than 5% if possible. The first column shows the ranges of the m_{ll} bins, and the second column shows the boundaries of the p_T bins.

m_{ll} [GeV]	p_T [GeV]
300–360	[100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 250, 270, 290, 310, 330, 360, 380, 410, 440, 490, 570, 7000]
360–450	[100, 110, 120, 130, 140, 150, 160, 170, 180, 200, 230, 250, 270, 290, 310, 330, 350, 370, 400, 440, 490, 580, 7000]
450–600	[100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 210, 230, 250, 270, 290, 320, 340, 360, 390, 430, 480, 580, 7000]
600–800	[100, 110, 120, 130, 150, 170, 200, 220, 250, 290, 320, 360, 420, 520, 7000]
800–1100	[100, 110, 120, 150, 170, 200, 230, 270, 330, 430, 7000]
1100–1500	[100, 200, 290, 7000]
1500–2000	[100, 7000]
2000–2600	[100, 7000]

TABLE VII. The fine binning where the relative statistical uncertainty of each bin must be smaller than 10%. The first column shows the ranges of the m_{ll} bins, and the second column shows the boundaries of the p_T bins.

m_{ll} [GeV]	p_T [GeV]
300–360	[100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 250, 270, 290, 310, 330, 350, 370, 400, 420, 440, 470, 500, 530, 560, 600, 660, 760, 7000]
360–450	[100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 240, 260, 290, 310, 330, 350, 370, 390, 410, 440, 470, 500, 530, 560, 610, 670, 770, 7000]
450–600	[100, 110, 120, 130, 140, 150, 160, 190, 210, 230, 250, 270, 290, 320, 340, 370, 390, 420, 460, 490, 520, 550, 580, 620, 680, 780, 7000]
600–800	[100, 110, 120, 130, 150, 170, 200, 220, 240, 260, 280, 310, 340, 380, 410, 440, 470, 510, 550, 620, 730, 7000]
800–1100	[100, 110, 120, 140, 160, 180, 200, 220, 250, 270, 300, 330, 360, 410, 460, 540, 660, 7000]
1100–1500	[100, 130, 160, 190, 230, 270, 320, 400, 520, 7000]
1500–2000	[100, 210, 330, 7000]
2000–2600	[100, 7000]

[1] W. Buchmüller and D. Wyler, Nucl. Phys. **B268,** 621 (1986).
[2] C. Arzt, M. B. Einhorn, and J. Wudka, Nucl. Phys. **B433,** 41 (1995).
[3] B. Grzadkowski, M. Iskrzynski, M. Misiak, and J. Rosiek, J. High Energy Phys. 10 (2010) 085.
[4] C. W. Murphy, J. High Energy Phys. 10 (2020) 174.
[5] H.-L. Li, Z. Ren, J. Shu, M.-L. Xiao, J.-H. Yu, and Y.-H. Zheng, Phys. Rev. D **104,** 015026 (2021).
[6] Z. Han and W. Skiba, Phys. Rev. D **71,** 075009 (2005).
[7] V. Cirigliano, M. González-Alonso, and M. L. Graesser, J. High Energy Phys. 02 (2013) 046.
[8] C.-Y. Chen, S. Dawson, and C. Zhang, Phys. Rev. D **89,** 015016 (2014).
[9] J. Ellis, V. Sanz, and T. You, J. High Energy Phys. 07 (2014) 036.
[10] J. D. Wells and Z. Zhang, Phys. Rev. D **90,** 033006 (2014).
[11] A. Falkowski and F. Riva, J. High Energy Phys. 02 (2015) 039.
[12] V. Cirigliano, W. Dekens, J. de Vries, and E. Mereghetti, Phys. Rev. D **94,** 034031 (2016).
[13] J. de Blas, M. Ciuchini, E. Franco, S. Mishima, M. Pierini, L. Reina, and L. Silvestrini, J. High Energy Phys. 12 (2016) 135.
[14] C. Hartmann, W. Shepherd, and M. Trott, J. High Energy Phys. 03 (2017) 060.
[15] A. Falkowski, M. González-Alonso, and K. Mimouni, J. High Energy Phys. 08 (2017) 123.
[16] S. Alioli, V. Cirigliano, W. Dekens, J. de Vries, and E. Mereghetti, J. High Energy Phys. 05 (2017) 086.
[17] S. Alioli, M. Farina, D. Pappadopulo, and J. T. Ruderman, Phys. Rev. Lett. **120,** 101801 (2018).
[18] S. Alioli, W. Dekens, M. Girard, and E. Mereghetti, J. High Energy Phys. 08 (2018) 205.
[19] A. Biekoetter, T. Corbett, and T. Plehn, SciPost Phys. **6,** 064 (2019).
[20] C. Grojean, M. Montull, and M. Riembau, J. High Energy Phys. 03 (2019) 020.
[21] J. Baglio, S. Dawson, S. Homiller, S. D. Lane, and I. M. Lewis, Phys. Rev. D **101,** 115004 (2020).
[22] R. Boughezal, F. Petriello, and D. Wiegand, Phys. Rev. D **101,** 116002 (2020).
[23] R. Boughezal, C.-Y. Chen, F. Petriello, and D. Wiegand, Phys. Rev. D **103,** 055015 (2021).
[24] R. Torre, L. Ricci, and A. Wulzer, J. High Energy Phys. 02 (2021) 144.
[25] A. Horne, J. Pittman, M. Snedeker, W. Shepherd, and J. W. Walker, J. High Energy Phys. 03 (2021) 118.
[26] R. Boughezal, F. Petriello, and D. Wiegand, Phys. Rev. D **104,** 016005 (2021).
[27] J. J. Ethier, R. Gomez-Ambrosio, G. Magni, and J. Rojo, Eur. Phys. J. C **81,** 560 (2021).
[28] R. Boughezal, A. Emmert, T. Kutz, S. Mantry, M. Nycz, F. Petriello, K. Şimşek, D. Wiegand, and X. Zheng, Phys. Rev. D **106,** 016006 (2022).
[29] Anisha *et al.*, arXiv:2111.05876.
[30] A. Pomarol and F. Riva, J. High Energy Phys. 01 (2014) 151.
[31] S. Di Vita, C. Grojean, G. Panico, M. Riembau, and T. Vantalon, J. High Energy Phys. 09 (2017) 069.
[32] E. da Silva Almeida, A. Alves, N. Rosa Agostinho, O. J. P. Éboli, and M. C. Gonzalez-Garcia, Phys. Rev. D **99,** 033001 (2019).
[33] J. Ellis, C. W. Murphy, V. Sanz, and T. You, J. High Energy Phys. 06 (2018) 146.
[34] N. P. Hartland, F. Maltoni, E. R. Nocera, J. Rojo, E. Slade, E. Vryonidou, and C. Zhang, J. High Energy Phys. 04 (2019) 100.
[35] I. Brivio, S. Bruggisser, F. Maltoni, R. Moutafis, T. Plehn, E. Vryonidou, S. Westhoff, and C. Zhang, J. High Energy Phys. 02 (2020) 131.
[36] S. van Beek, E. R. Nocera, J. Rojo, and E. Slade, SciPost Phys. **7,** 070 (2019).
[37] R. Aoude, T. Hurth, S. Renner, and W. Shepherd, J. High Energy Phys. 12 (2020) 113.
[38] J. Ellis, M. Madigan, K. Mimasu, V. Sanz, and T. You, J. High Energy Phys. 04 (2021) 279.
[39] S. Dawson, S. Homiller, and S. D. Lane, Phys. Rev. D **102,** 055012 (2020).
[40] A. Greljo, S. Iranipour, Z. Kassabov, M. Madigan, J. Moore, J. Rojo, M. Ubiali, and C. Voisey, J. High Energy Phys. 07 (2021) 122.
[41] J. J. Ethier *et al.*, J. High Energy Phys. 11 (2021) 089.
[42] S. Carrazza, C. Degrande, S. Iranipour, J. Rojo, and M. Ubiali, Phys. Rev. Lett. **123,** 132001 (2019).
[43] S. Alioli, R. Boughezal, E. Mereghetti, and F. Petriello, Phys. Lett. B **809,** 135703 (2020).
[44] R. Boughezal, E. Mereghetti, and F. Petriello, Phys. Rev. D **104,** 095022 (2021).
[45] T. Kim and A. Martin, arXiv:2203.11976.
[46] CMS Collaboration, arXiv:2205.04897.
[47] S. Dawson, P. P. Giardino, and A. Ismail, Phys. Rev. D **99,** 035044 (2019).
[48] X. Li *et al.*, arXiv:2204.13121.
[49] T. Cohen, X. Lu, and Z. Zhang, SciPost Phys. **10,** 098 (2021).
[50] S. Dawson, D. Fontes, S. Homiller, and M. Sullivan, arXiv:2205.01561.
[51] M. Carena, A. Daleo, B. A. Dobrescu, and T. M. P. Tait, Phys. Rev. D **70,** 093009 (2004).
[52] J. Blumlein and R. Ruckl, Phys. Lett. B **304,** 337 (1993).
[53] J. Blumlein, E. Boos, and A. Kryukov, Z. Phys. C **76,** 137 (1997).
[54] M. J. Baker, J. Fuentes-Martín, G. Isidori, and M. König, Eur. Phys. J. C **79,** 334 (2019).
[55] C. Zhang and S.-Y. Zhou, Phys. Rev. Lett. **125,** 201601 (2020).
[56] C. de Rham, S. Kundu, M. Reece, A. J. Tolley, and S.-Y. Zhou, Snowmass white paper: UV constraints on IR physics, in 2022 Snowmass Summer Study (2022), arXiv:2203.06805.
[57] J. Campbell and T. Neumann, J. High Energy Phys. 12 (2019) 034.
[58] R. D. Ball *et al.* (NNPDF Collaboration), Eur. Phys. J. C **77,** 663 (2017).
[59] A. Denner, B. Jantzen, and S. Pozzorini, Nucl. Phys. **B761,** 1 (2007).
[60] G. Panico, L. Ricci, and A. Wulzer, J. High Energy Phys. 07 (2021) 086.