Measuring the anomalous quartic gauge couplings in the $W^+W^- \to W^+W^-$ process at muon collider using artificial neural networks

Ji-Chong Yang,[a,c] **Xue-Ying Han,**[a,c] **Zhi-Bin Qin,**[a] **Tong Li**[b] **and Yu-Chen Guo**[a,c,1]

[a]*Department of Physics, Liaoning Normal University,*
No. 850 Huanghe Road, Dalian 116029, China

[b]*School of Physics, Nankai University,*
94 Weijin Road, Tianjin 300071, China

[c]*Center for Theoretical and Experimental High Energy Physics, Liaoning Normal University,*
No. 850 Huanghe Road, Dalian 116029, China

E-mail: yangjichong@lnnu.edu.cn, 944789960@qq.com, lnnuqzb163@163.com,
litong@nankai.edu.cn, ycguo@lnnu.edu.cn

ABSTRACT: The muon collider provides a unique opportunity to study the vector boson scattering processes and dimension-8 operators contributing to anomalous quartic gauge couplings (aQGCs). Because of the cleaner final state, it is easier to decode subprocess and certain operator couplings at a muon collider. We attempt to identify the anomalous $WWWW$ coupling in the exclusive $WW \to WW$ scattering in this paper. Since one aQGC can be induced by multiple dimension-8 operators, the study of one coupling can help to confine different operators. Meanwhile, singling out the $WW \to WW$ process can help to study the unitarity bounds. The vector boson scattering process corresponding to the anomalous $WWWW$ coupling is $\mu^+\mu^- \to \nu\nu\bar{\nu}\bar{\nu}\ell^+\ell^-$, with four (anti-)neutrinos in the final state, which brings troubles in phenomenological studies. In this paper, the machine learning method is used to tackle this problem. We find that, the artificial neural network is helpful to extract the $W^+W^- \to W^+W^-$ contribution, and reconstruct the center of mass energy of the subprocess which is important in the study of the Standard Model effective field theory. The sensitivities and the expected constraints on the dimension-8 operators at the muon collider with $\sqrt{s} = 30\,\mathrm{TeV}$ are presented. We demonstrate that the artificial neural networks exhibit great potential in the phenomenological study of processes with multiple neutrinos in the final state.

---

[1]Corresponding author.

Contents

1 Introduction

The self-couplings of electroweak (EW) gauge bosons are most closely related to the nature of the electroweak symmetry breaking (EWSB) [1–5]. Any hints for the anomalous gauge couplings would indicate the existence of new physics (NP) beyond the Standard Model (SM). In the framework of the SM effective field theory (SMEFT) [6–9], the dimension-8 operators contribute to the anomalous quartic gauge couplings (aQGCs) [10, 11]. On the other hand, the vector boson scattering (VBS) is one of the most common channels for performing precision measurements of the SM or searching NP beyond the SM at high-energy colliders. The probe of aQGCs through VBS is thus one of the most important topics at the Large Hadron Collider (LHC) and has received great attention [12–26]. Nevertheless, the VBS measurements suffer from the large QCD background at the LHC and it is difficult to decode the initial states of subprocess as the final jets in the forward region are not distinguishable.

Recently, the development of the muon collider has gradually entered the limelight [27–34]. On the high-energy muon collider, the dominant production mode for the SM and

NP particles is VBS or vector boson fusion process [33]. Therefore, the muon collider is also known as a gauge boson collider [34]. Compared to the LHC, there are no composite particles in the initial states at muon collider, and thus the QCD background is not severe. Taking WW initiated scattering as an example, the high-energy muon beams radiate W bosons and turn into neutrinos. Meanwhile, the neutral gauge bosons Z, γ are also radiated under an approximately unbroken SM gauge symmetry and muons are produced in final states. The outgoing muons are extremely forward with a small polar angle of the order $\theta_\mu \sim M_Z/E_\mu \approx 1.2°$ for a Z-initiated process at $10\,\text{TeV}$ [35] and most likely escape the detector. If we require the outgoing muons to be observable in the detector coverage $10° < \theta_\mu < 170°$, the cross sections of neutral gauge bosons initiated scattering would be substantially suppressed by two orders of magnitude [36, 37]. Thus, it is feasible to study the aQGCs induced exclusive $W^+W^- \to W^+W^-$ scattering at a muon collider. In addition, the muon collider can reach both high energy and high luminosity, which will be of great help to precisely measure aQGCs, since the cross-section induced by dimension-8 operators increases significantly with energy. Meanwhile, high luminosity is considered as one of the keys to solve the "EFT triangle" problem [38–41].

In this work, we investigate the sensitivity of $W^+W^- \to W^+W^-$ scattering to the dimension-8 scalar/longitudinal operators contributing to aQGCs at muon colliders.[1] The W bosons in final states are then followed by purely leptonic decay $W^\pm \to \ell^\pm \nu$. One key problem of the process $\mu^+\mu^- \to \nu\nu\bar{\nu}\bar{\nu}\ell^+\ell^-$ is the presence of the (anti-)neutrinos which lead to difficulties in the phenomenological studies. For example, it is difficult to reconstruct the center of mass (c.m.) energy of the subprocess $W^+W^- \to W^+W^-$ (denoted as $\sqrt{\hat{s}}$). In the content of EFT, the Wilson coefficients of effective operators should rely on energy scales [43]. At a high-energy muon collider, the c.m. energy is especially important because the SMEFT is only valid under a certain energy scale and its validity should be taken into account. The unitarity bound [44–48] is often needed to investigate the validity of the SMEFT, and the c.m. energy is necessary information to apply unitarity bounds. To solve the problem of reconstructing $\hat{s}$ in the processes with multiple (anti-)neutrinos, a machine learning approach has been introduced into high energy physics (HEP) community [41]. The machine learning methods have been widely used, and are being rapidly developed in HEP [49–63]. In this paper, we adopt the artificial neural network (ANN) to extract the $W^+W^- \to W^+W^-$ contribution and to reconstruct $\hat{s}$. The complexity caused by the neutrinos just provides a venue to explore the boundaries of ANN capabilities. Based on the ANNs, the sensitivities of the process $\mu^+\mu^- \to \nu\nu\bar{\nu}\bar{\nu}\ell^+\ell^-$ to the dimension-8 operators contributing to aQGCs are investigated, with the focus on the $W^+W^- \to W^+W^-$ contribution.

The rest of this paper is organized as follows. In section 2, the dimension-8 operators contributing to aQGCs are briefly reviewed. The ANN approach to extract the $W^+W^- \to W^+W^-$ contribution is discussed in section 3. In section 4, we discuss the ANN approach to reconstruct $\hat{s}$. The expected constraints on the coefficients of the aQGC operators at the muon collider are estimated in section 5. Section 6 summarizes our main conclusions.

---

[1]A recent Snowmass paper investigated the searches of aQGCs through the production of WW boson pairs at a muon collider with $\sqrt{s} = 6\,\text{TeV}$ and an integrated luminosity of $4\,\text{ab}^{-1}$ [42]. They studied the $WW\nu\nu$ and $WW\mu\mu$ final states with the W bosons decaying hadronically.

2 A brief introduction of the anomalous quartic gauge couplings

The Lagrangian of the dimension-8 operators contributing to aQGCs can be written as [10, 11]

$$\mathcal{L}_{\text{aQGC}} = \sum_{i=0}^{2} \frac{f_{S_i}}{\Lambda^4} O_{S,i} + \sum_{j=0}^{7} \frac{f_{M_j}}{\Lambda^4} O_{M,j} + \sum_{k=0}^{9} \frac{f_{T_k}}{\Lambda^4} O_{T,k} \tag{2.1}$$

with

$$O_{S,0} = \left[(D_\mu \Phi)^\dagger D_\nu \Phi \right] \times \left[(D^\mu \Phi)^\dagger D^\nu \Phi \right], \quad O_{S,1} = \left[(D_\mu \Phi)^\dagger D^\mu \Phi \right] \times \left[(D_\nu \Phi)^\dagger D^\nu \Phi \right],$$
$$O_{S,2} = \left[(D_\mu \Phi)^\dagger D_\nu \Phi \right] \times \left[(D^\nu \Phi)^\dagger D^\mu \Phi \right], \tag{2.2}$$

$$O_{M,0} = \text{Tr} \left[\widehat{W}_{\mu\nu} \widehat{W}^{\mu\nu} \right] \times \left[(D_\beta \Phi)^\dagger D^\beta \Phi \right], \quad O_{M,1} = \text{Tr} \left[\widehat{W}_{\mu\nu} \widehat{W}^{\nu\beta} \right] \times \left[(D_\beta \Phi)^\dagger D^\mu \Phi \right],$$
$$O_{M,2} = \left[B_{\mu\nu} B^{\mu\nu} \right] \times \left[(D_\beta \Phi)^\dagger D^\beta \Phi \right], \quad O_{M,3} = \left[B_{\mu\nu} B^{\nu\beta} \right] \times \left[(D_\beta \Phi)^\dagger D^\mu \Phi \right],$$
$$O_{M,4} = \left[(D_\mu \Phi)^\dagger \widehat{W}_{\beta\nu} D^\mu \Phi \right] \times B^{\beta\nu}, \quad O_{M,5} = \left[(D_\mu \Phi)^\dagger \widehat{W}_{\beta\nu} D^\nu \Phi \right] \times B^{\beta\mu} + h.c., \tag{2.3}$$
$$O_{M,7} = (D_\mu \Phi)^\dagger \widehat{W}_{\beta\nu} \widehat{W}^{\beta\mu} D^\nu \Phi,$$

$$O_{T,0} = \text{Tr} \left[\widehat{W}_{\mu\nu} \widehat{W}^{\mu\nu} \right] \times \text{Tr} \left[\widehat{W}_{\alpha\beta} \widehat{W}^{\alpha\beta} \right], \quad O_{T,1} = \text{Tr} \left[\widehat{W}_{\alpha\nu} \widehat{W}^{\mu\beta} \right] \times \text{Tr} \left[\widehat{W}_{\mu\beta} \widehat{W}^{\alpha\nu} \right],$$
$$O_{T,2} = \text{Tr} \left[\widehat{W}_{\alpha\mu} \widehat{W}^{\mu\beta} \right] \times \text{Tr} \left[\widehat{W}_{\beta\nu} \widehat{W}^{\nu\alpha} \right], \quad O_{T,5} = \text{Tr} \left[\widehat{W}_{\mu\nu} \widehat{W}^{\mu\nu} \right] \times B_{\alpha\beta} B^{\alpha\beta},$$
$$O_{T,6} = \text{Tr} \left[\widehat{W}_{\alpha\nu} \widehat{W}^{\mu\beta} \right] \times B_{\mu\beta} B^{\alpha\nu}, \quad O_{T,7} = \text{Tr} \left[\widehat{W}_{\alpha\mu} \widehat{W}^{\mu\beta} \right] \times B_{\beta\nu} B^{\nu\alpha}, \tag{2.4}$$
$$O_{T,8} = B_{\mu\nu} B^{\mu\nu} \times B_{\alpha\beta} B^{\alpha\beta}, \quad O_{T,9} = B_{\alpha\mu} B^{\mu\beta} \times B_{\beta\nu} B^{\nu\alpha},$$

where Φ denotes the SM Higgs doublet, D_μ is covariant derivative, $\widehat{W} \equiv \vec{\sigma} \cdot \vec{W}/2$ with σ being the Pauli matrix and $\vec{W} = \{W^1, W^2, W^3\}$, B_μ and W_μ^i are $U(1)_Y$ and $SU(2)_I$ gauge fields, and $B_{\mu\nu}$ and $W_{\mu\nu}$ correspond to the gauge invariant field strength tensor. Many NP models can generate effective dimension-8 operators contributing to aQGCs [64–73]. Although the dimension-6 operators have received most studies, the importance of dimension-8 operators was recently emphasized by many groups [1, 74–91]. As VBS processes receive great attention at the LHC, the above operators in the SMEFT have been investigated intensively. The LHC constraints on the coefficients of the operators assuming one operator at a time are listed in table 1. Note that a UV completion model usually does not contribute to only one operator. The assumption that only one operator exists at a time can be used to study the sensitivity of a process and place stringent constraints when NP beyond the SM has not yet been found. The scattering process $W^+ W^- \to W^+ W^-$ can be contributed by $O_{S_{0,1,2}}$, $O_{M_{0,1,7}}$ and $O_{T_{0,1,2}}$ operators, therefore, in the following we concentrate on these operators.

3 Identification of the $W^+ W^- \to W^+ W^-$ scattering

A naive dimensional analysis of the cross-section ignoring inferred divergences and logarithms yields $\sigma_{\text{SM}} \sim 1/s$, $\sigma_{\text{int}} \sim sf/\Lambda^4$ and $\sigma_{\text{NP}} \sim s^3 (f/\Lambda^4)^2$, where σ_{SM}, σ_{int} and σ_{NP} denote the contributions from the SM, interference term and NP squared term, respectively. For

coefficient	constraint	coefficient	constraint
f_{S_0}/Λ^4	$[-2.7, 2.7]$ [24]	f_{T_0}/Λ^4	$[-0.12, 0.11]$ [24]
f_{S_1}/Λ^4	$[-3.4, 3.4]$ [24]	f_{T_1}/Λ^4	$[-0.12, 0.13]$ [24]
f_{S_2}/Λ^4	—	f_{T_2}/Λ^4	$[-0.28, 0.28]$ [24]
f_{M_0}/Λ^4	$[-0.69, 0.70]$ [24]	f_{T_5}/Λ^4	$[-0.5, 0.5]$ [25]
f_{M_1}/Λ^4	$[-2.0, 2.1]$ [24]	f_{T_6}/Λ^4	$[-0.4, 0.4]$ [25]
f_{M_2}/Λ^4	$[-2.8, 2.8]$ [25]	f_{T_7}/Λ^4	$[-0.9, 0.9]$ [25]
f_{M_3}/Λ^4	$[-4.4, 4.4]$ [25]	f_{T_8}/Λ^4	$[-0.43, 0.43]$ [26]
f_{M_4}/Λ^4	$[-5, 5]$ [25]	f_{T_9}/Λ^4	$[-0.92, 0.92]$ [26]
f_{M_5}/Λ^4	$[-8.3, 8.3]$ [25]		
f_{M_7}/Λ^4	$[-3.4, 3.4]$ [25]		

Table 1. The LHC constraints on the coefficients (in unit of TeV^{-4}) of dimension-8 aQGC operators obtained at 95% CL.

$\sqrt{s} = 30$ TeV, one has $\sigma_{\text{int}} \sim \sigma_{\text{NP}}$ when $s^2 f/\Lambda^4 \sim 1$, that is $f/\Lambda^4 \sim 10^{-6}$ TeV^{-4}. For f/Λ^4 above this value, the interference contribution is smaller than σ_{NP}. Besides, as the helicity amplitude grow fast with energy, there is not necessarily a corresponding large helicity amplitude in the SM that interferes with NP. In summary, whether the interference can be neglected needs to be verified for the range of operator coefficients of interest. A numerical justification of this is postponed to section 5.2.

The tree-level Feynman diagrams of the aQGCs contribution to the process $\mu^+\mu^- \rightarrow \ell^+\ell^-\nu\nu\bar{\nu}\bar{\nu}$ are shown in figure 1. They can be categorized into three different types, including tri-boson processes (figure 1 (a)), VBS (figure 1 (b)) and the Yukawa suppressed diagrams involving Higgs boson (figure 1 (c)). Although the VBS processes dominate the contribution of the aQGCs at high energies, there is still significant contribution from the tri-boson process for the above dimension-8 operators [92]. Among the VBS processes, other than $W^+W^- \rightarrow W^+W^-$, there are contributions from neutral gauge boson induced processes together with forward muons. To identify the subprocess $W^+W^- \rightarrow W^+W^-$ contribution, it is necessary to select the charged leptons in central region. In the following, the contribution from diagrams including a $WWWW$ SMEFT interaction is denoted as σ_{4W}, and $\sigma_{\text{no-4W}}$ denotes other SMEFT contributions.

To study the features of the aQGCs contribution, a Monte Carlo (M.C.) simulation is applied with the help of the `MadGraph5_aMC@NLO` toolkit [93, 94]. The events are generated with one operator at a time. Through out the paper, the standard cuts are set as the default ones in `MadGraph5_aMC@NLO`, as

$$p_T^\ell > 10 \text{ GeV}, \ \ |y_\ell| < 2.5, \ \ \Delta R_{\ell\ell} > 0.4, \tag{3.1}$$

where p_T^ℓ is the transverse momentum of charged lepton, y_ℓ is the rapidity of charged lepton, and $\Delta R_{\ell\ell} = \sqrt{\Delta\phi^2 + \Delta y^2}$ with $\Delta\phi$ and Δy being the differences of azimuthal angles and rapidities of two charged leptons.

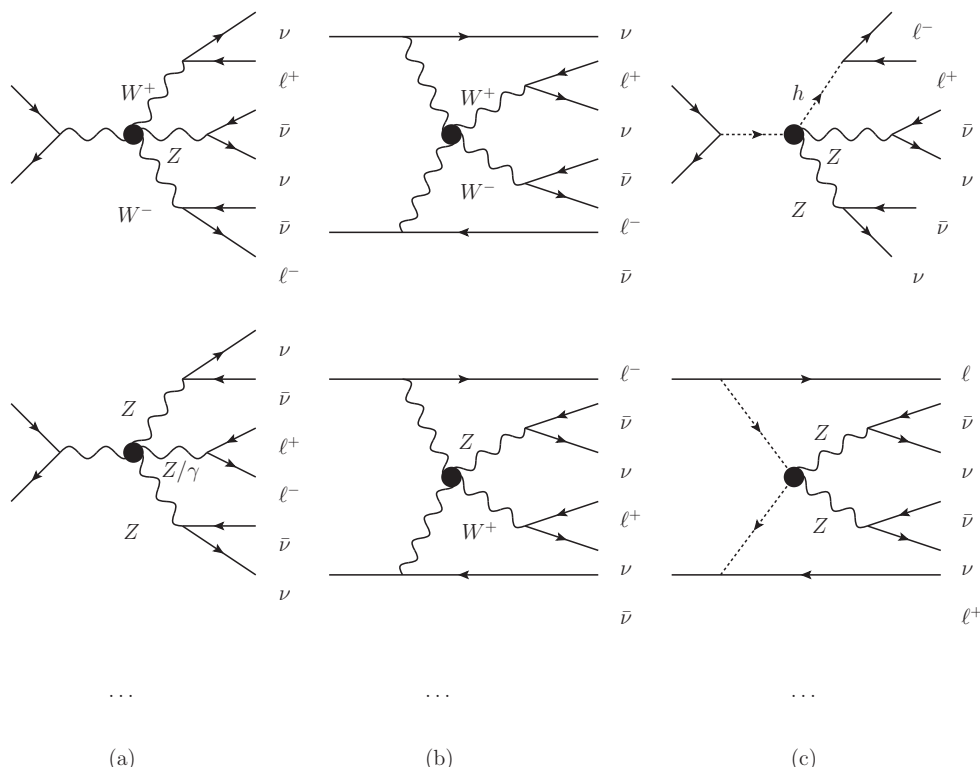**Figure 1**. The tree-level Feynman diagrams of the aQGCs contribution to the process $\mu^+\mu^- \to \ell^+\ell^-\nu\nu\bar{\nu}\bar{\nu}$.

Since the neutrinos are invisible, in principle, one cannot use the information of the neutrinos. However, before we use the machine learning method to identify the contribution of σ_{4W}, it is useful to illustrate the size of σ_{4W} first. This can provide a criterion for the later algorithms which only utilize detectable observables. For this purpose, here we temporarily use the information of neutrinos obtained in the M.C. simulations.

Note that there are always two (anti-)neutrinos from a Z boson decay in no-4W processes. The neutrino flavors from a Z boson decay must be the same. For $W^+W^- \to W^+W^-$, besides the two (anti-)neutrinos ν_μ and $\bar{\nu}_\mu$ along the beam direction, the flavors of the other two neutrinos must correspond to the charged leptons from two W bosons' decay. We denote $m_{\nu\bar{\nu}}$ as the invariant mass of a pair of (anti-)neutrinos with the same flavor whose invariant mass is closest to m_Z among all possible combinations of neutrinos. The events are then separated into two groups, according to the neutrino flavors and the size of mass window $\Delta_m = |m_{\nu\bar{\nu}} - m_Z|$. σ_{4W} is calculated as the cross-section with $\Delta_m > 15$ GeV and the (anti-)neutrino flavors from $W^+W^- \to W^+W^-$. Meanwhile, $\sigma_{\text{no}-4W}$ corresponds to the cross-section with $\Delta_m \le 15$ GeV or with wrong flavors. Their results at $\sqrt{s} = 3, 5, 10,$ 14 and 30 TeV [34] are listed in table 2.

In table 2, using the above selection strategy, we show the M.C. results of σ_{4W} and $\sigma_{\text{no}-4W}$ given by O_{S_0}, O_{M_0} and O_{T_0} for illustration. The Wilson coefficients are taken to be the maximal values allowed by the LHC in table 1. The charged leptons in final states

	$\ell\ell$	3 TeV	10 TeV	14 TeV	30 TeV
$\sigma_{4\mathrm{W}}^{\mathrm{EVA}}$	ee	0.0048	6.62	49.8	4825.7
	ee	0.00409 : 0.00016	5.832 : 0.005	44.14 : 0.03	4286.7 : 0.6
O_{S_0}	$e\mu$	0.0078 : 0.0012	11.7 : 0.2	88.3 : 0.8	8571.6 : 17.4
	$\mu\mu$	0.0037 : 0.0015	5.8 : 0.2	44.0 : 0.8	4280.5 : 17.0
$\sigma_{4\mathrm{W}}^{\mathrm{EVA}}$	ee	0.0058	18.9	172.3	24338.5
	ee	0.0061 : 0.0062	19.2 : 6.3	173.6 : 47.1	24044 : 4515
O_{M_0}	$e\mu$	0.012 : 0.010	38.4 : 13.4	346.6 : 101.2	48302 : 9812
	$\mu\mu$	0.0060 : 0.0092	19.1 : 10.3	173.6 : 76.4	24120 : 7388
$\sigma_{4\mathrm{W}}^{\mathrm{EVA}}$	ee	0.0061	23.7	221.6	32523.9
	ee	0.0069 : 0.0052	26.8 : 6.1	250.9 : 45.6	37083 : 4618
O_{T_0}	$e\mu$	0.014 : 0.011	53.7 : 19.7	502.1 : 157.8	74450 : 17396
	$\mu\mu$	0.0068 : 0.0117	26.8 : 19.9	251.6 : 160.1	37359 : 17673

Table 2. $\sigma_{4\mathrm{W}} : \sigma_{\mathrm{no-4W}}$ (fb) for different operators and different charged lepton flavors. The predictions of effective vector boson approximation (EVA) are denoted as $\sigma_{4\mathrm{W}}^{\mathrm{EVA}}$ for ee final states.

are ee, $\mu\mu$ or $e\mu$. We also evaluate $\sigma_{4\mathrm{W}}$ with ee charged leptons for comparison, using effective vector boson approximation (EVA) [95–97] (denoted as $\sigma_{4\mathrm{W}}^{\mathrm{EVA}}$). The detailed EVA calculation is given in appendix A. Although there are kinematic cuts and unavoidable interference in M.C. simulations, the small discrepancy between $\sigma_{4\mathrm{W}}^{\mathrm{EVA}}$ and $\sigma_{4\mathrm{W}}$ indicates that our selection strategy of $\sigma_{4\mathrm{W}}$ is reliable. Table 2 shows that if the aQGC signal is induced by O_{S_i} operators, when $\sqrt{s} \geq 10$ TeV, one can concentrate on $\sigma_{4\mathrm{W}}$ and neglect $\sigma_{\mathrm{no-4W}}$. This is because the $ZV \to VV$ VBS processes are suppressed by requiring the charged leptons in central region, i.e. the p_T^ℓ and $|y_\ell|$ cuts in eq. (3.1). Besides, the diagrams with the Higgs bosons are suppressed by the Yukawa coupling. As a consequence, $\sigma_{\mathrm{no-4W}}$ is in fact dominated by the tri-boson contribution. Meanwhile, the Feynmann rules for O_{M_i, T_i} operators are typically momentum dependent. This is however not the case for O_{S_i} operators, and therefore the tri-boson contribution is significantly suppressed at high energies. As a result, $\sigma_{\mathrm{no-4W}}$ is negligible for O_{S_i} operators.

For the cases of O_{M_i} and O_{T_i} operators, although the cross-section of the $W^+W^- \to W^+W^-$ contribution grows with $\sqrt{s}$, $\sigma_{\mathrm{no-4W}}$ is not negligible even at $\sqrt{s} = 30$ TeV. We will use an ANN to select the events from the process involving subprocess $W^+W^- \to W^+W^-$.

Based on the results of table 2, we find that, to study the VBS subprocess $W^+W^- \to W^+W^-$, a larger $\sqrt{s}$ is needed. Therefore, we only consider $\sqrt{s} = 30$ TeV below. In the following, we use $d_{4\mathrm{W}}$ to represent the event data-set from $W^+W^- \to W^+W^-$ contribution, and $d_{\mathrm{no-4W}}$ to represent the event data-set from $\sigma_{\mathrm{no-4W}}$ contribution.

3.1 The traditional approach

The mechanism behind the ANN is a "black box". Thus, before using the ANN, we must verify that the data set contains information that allows us to extract the $W^+W^- \to W^+W^-$ contribution.

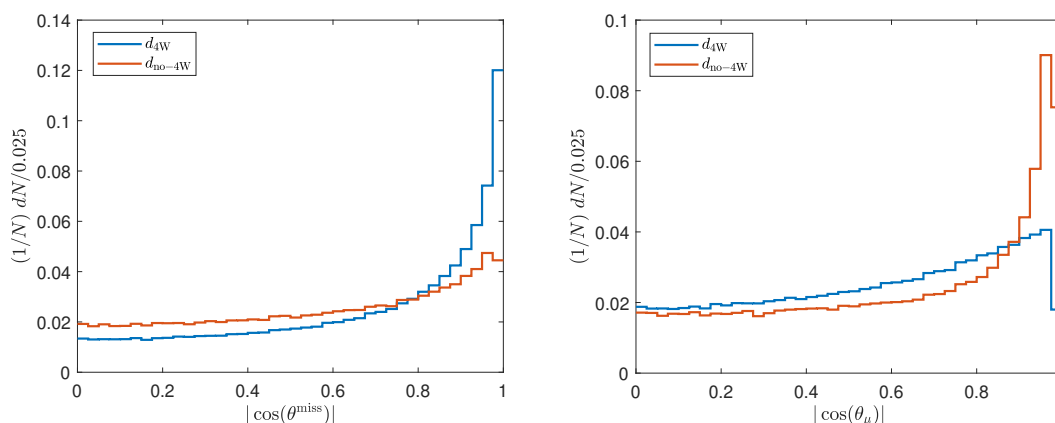**Figure 2**. The normalized distributions of $|\cos(\theta^{\mathrm{miss}})|$ (left panel) and $|\cos(\theta_\mu)|$ (right panel).

For the process $\mu^+\mu^- \to \nu_\mu\bar{\nu}_\mu W^+W^-$, the (anti-)neutrinos tend to be along the muon beam direction and are back-to-back. Meanwhile, when $\hat{s}$ is large, $W^\pm$ are energetic and also back-to-back. Consequently, the (anti-)neutrinos from $W \to \ell\nu$ tend to be along the directions of $W^\pm$ and therefore also back-to-back. One can conclude that, the transverse missing momentum in the $W^+W^- \to W^+W^-$ contribution should be relatively small. At a lepton collider, all the components of the missing momentum can be obtained by using momentum conservation with satisfactory accuracy, and the zenith angle of the missing momentum (denoted as θ^{miss}) can be regarded as an observable. We find that θ^{miss} provides a better discrimination than the transverse missing momentum. At $\sqrt{s} = 30$ TeV the normalized distributions of $|\cos(\theta^{\mathrm{miss}})|$ for O_{M_0} are shown in left panel of figure 2. One can see that $|\cos(\theta^{\mathrm{miss}})|$ is indeed closer to 1 for events from $d_{4\mathrm{W}}$.

The cross-section can also be contributed by other VBS processes in addition to $W^+W^- \to W^+W^-$. Taking the process $\mu^+\mu^- \to \nu\nu\bar{\nu}\bar{\nu}e^\pm\mu^\mp$ as an example, there is also $ZW^\pm \to ZW^\pm$ process as shown in the second Feynman diagram in figure 1 (b). For the $ZW^\pm \to ZW^\pm$ process, the direction of the muon in the final state tends to be along the direction of **z**-axis. Denoting the zenith angle of the muon as θ_μ, the normalized distributions of $|\cos(\theta_\mu)|$ for O_{T_0} are shown in the right panel of figure 2. One can see that $|\cos(\theta_\mu)|$ is closer to 1 for events from $d_{\mathrm{no-4W}}$.

Although we can distinguish events from $d_{4\mathrm{W}}$ and $d_{\mathrm{no-4W}}$ by some observables in this way, the distinction is not very efficient. One has to analyze different processes accordingly. To achieve a desirable efficiency, a complicated analysis is required. Finding patterns from complicated relationship is what an ANN is good at. It can be seen that even without the ANN, there are still clues to distinguish events from $d_{4\mathrm{W}}$ and $d_{\mathrm{no-4W}}$. The ANN simply automates and improves the search for the clues.

3.2 The neural network approach

The ANN is a mathematical model to simulate a human brain [98], which is good at finding the complicated mathematical mapping relationship between input and output. It could
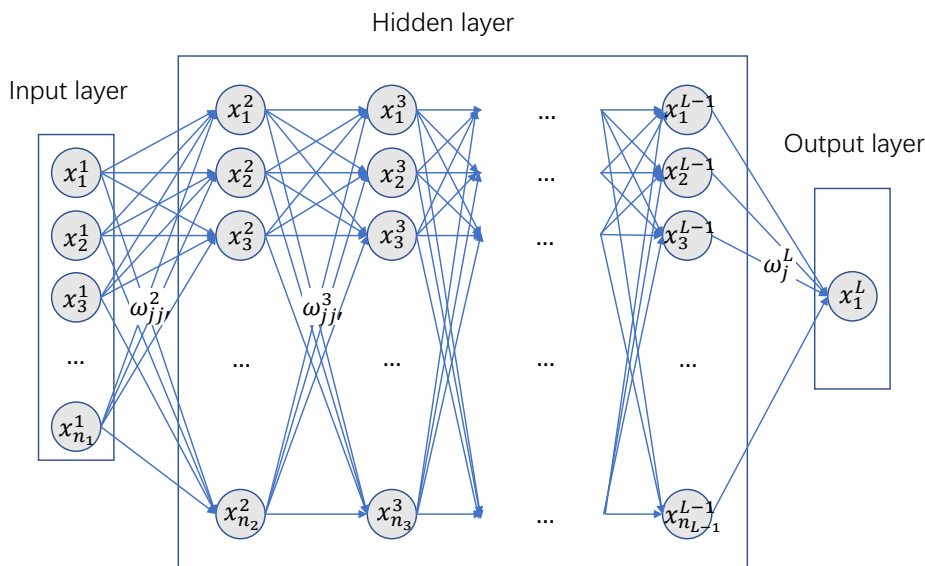
Figure 3. The graphical representation of the ANN.

be utilized for clustering the events. However, a classification is not tunable. For example, to archive a cleaner event set of d_{4W}, some events from the d_{4W} could be allowed to be misidentified as from d_{no-4W}. Therefore, instead of clustering, we treat the identification of the events from d_{4W} as a regression problem.

For an ANN, the relationship between input and output is determined by interconnected nodes and their connection modes. We use a dense connected ANN. An ANN is composed with an input layer, hidden layers and an output layer. Denoting x_j^i as neurons in the i-th layer, where $x_{1 \leq j \leq n_1}^1$ are input neurons, $x_{1 \leq j \leq n_i}^{2 \leq i \leq L-1}$ are in hidden layers and x_1^L is the output neuron, where L is the number of layers and n_i are the number of neurons in the i-th layer, the ANN can be depicted in figure 3.

A value is assigned for each neuron which is also denoted as x_j^i. Then $x_{j'}^{i+1}$ can be related with x_j^i as

$$x_{j'}^{i+1} = f_{j'}^{i+1} \left(\sum_j \omega_{jj'}^{i+1} x_j^i + b_{j'}^{i+1} \right), \tag{3.2}$$

where $\omega_{jj'}^{i+1}$ and $b_{j'}^{i+1}$ are trainable parameters. $\omega_{jj'}^{i+1}$ are called the elements of the weight matrix W^{i+1}, $b_{j'}^{i+1}$ are components of the bias vector, and $f_{j'}^{i+1}$ are activation functions. Except for the output layer, the activation functions are chosen as the parametric rectified linear unit (PReLU) function [99] defined as

$$f(x) = \begin{cases} x, & x \geq 0; \\ \alpha x, & x < 0, \end{cases} \tag{3.3}$$

where α's are also trainable parameters. For the output layer, no activation function (i.e., linear activation function) is used. We use $L = 15$, $n_{10>i>1} = 50$, $n_1 = 14$ the same as the dimension of input data and $n_L = 1$ for the output layer. The architecture is built using `Keras` with a `TensorFlow` [100] backend.
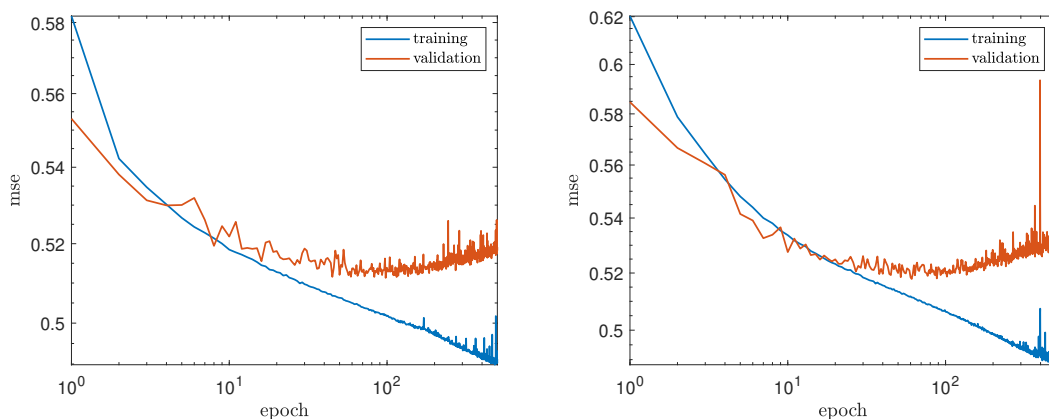
Figure 4. The learning curves of the ANNs trained for W_{score}, corresponding to O_{M_i} (left panel) and O_{T_i} (right panel).

The training and validation data-sets are prepared by using M.C. simulation. The data-sets consist of elements with 15 variables. The first 14 variables make up a 14-dimension vector fed to the input layer (denoted as v_i), and the last variable corresponds to the output layer. 12 components of the 14-dimension vector are the components of 4-momenta of charged leptons, and those of the missing momentum. The other 2 components correspond to the flavors of the charged leptons. The variable for output is set to 1 if the event is from σ_{4W} otherwise is set to 0. The ground truth of the output is determined with the help of non-observables, and the goal of the ANN is to reproduce the ground truth with only observables. In the following, the output predicted by the ANN is denoted as W_{score}. Therefore, if the ANN is well-trained, we expect that W_{score} is close to 1 for the events from σ_{4W} otherwise is close to 0. Two ANNs are trained, one for O_{M_i} operators and the other for O_{T_i} operators. The O_{S_i} operators dominantly contribute to σ_{4W}. Therefore, in this subsection, no ANN is trained to identify the $WW \to WW$ contribution and W_{score} cuts are not applied for O_{S_i} operators. For each operator of O_{M_i} and O_{T_i}, one million events are generated. One half of them form the training data-sets, and the other half form the validation data-sets. The data-sets are normalized using the z-score standardization, i.e., v_i' instead of v_i as input to the ANNs. It is defined as $v_i' = (v_i - \bar{v}_i)/\sigma_{v_i}$, where $\bar{v}_i$ and σ_{v_i} are the mean value and the standard deviation of all i-th variables of the elements in the training data-sets, respectively.

The learning curves are shown in figure 4. One can see that the mean squared errors (mses) stop to decrease for the validation data-sets at about epoches $= 100 \sim 200$. To avoid overfitting, we stop at epoches $= 100$ where the mses of validation data-sets stop to decrease. After training, the W_{score} results for the validation data-sets are shown in figure 5. Compared with figure 2, the W_{score} has stronger discrimination power. In the following, for each operator, we choose one minimal cut of W_{score} as long as the mistag rate of the events from d_{4W} reaches about 5%. The cuts and the effects of the cuts are shown in table 3.
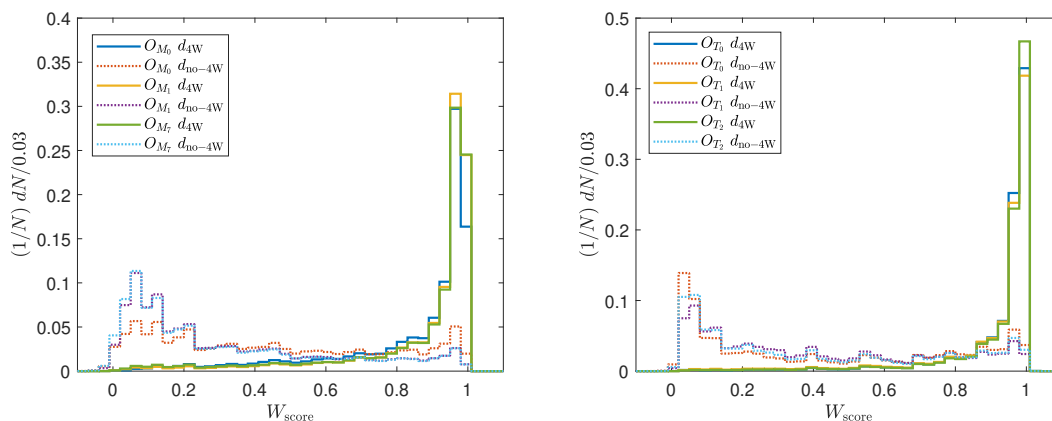
Figure 5. The normalized distributions of W_{score}.

	before cut	$W_{\text{score}} > 0.85$	$W_{\text{score}} > 0.9$	$W_{\text{score}} > 0.95$
O_{M_0}	$96.4 : 21.8$	$64.6 : 3.3$		
O_{M_1}	$132.6 : 64.5$	$99.5 : 5.2$		
O_{M_7}	$98.5 : 74.2$			$53.6 : 2.5$
O_{T_0}	$149.0 : 39.5$		$117.4 : 5.8$	
O_{T_1}	$141.5 : 36.4$	$116.7 : 5.4$		
O_{T_2}	$223.0 : 42.0$	$190.8 : 7.0$		

Table 3. $\sigma_{4\text{W}} : \sigma_{\text{no}-4\text{W}}$ (pb) for different operators after W_{score} cuts.

4 The reconstruction of center of mass energy of the $W^+W^- \to W^+W^-$ subprocess

The $\hat{s}$ of the process $W^+W^- \to W^+W^-$ is important in the study of the SMEFT, because as an EFT the Wilson coefficients should be dependent in energies. On the other hand, the SMEFT is only valid below certain energy scale. The violation of unitarity is often used as a signal that the SMEFT is no long valid, and unitarity bounds depend on the energy scale. In any case, energy scales are important information in the study of the SMEFT. However, the process $\mu^+\mu^- \to \nu\nu\bar{\nu}\bar{\nu}\ell^+\ell^-$ have four (anti-)neutrinos in the final state, which causes problems to reconstruct $\hat{s}$.

4.1 The traditional approach

In traditional approach, one has to analyze the kinematics of the process. With four (anti-)neutrinos, the kinematic feature is difficult to analyze. For the process $\mu^+\mu^- \to \nu\bar{\nu}W^+W^-$, the (anti-)neutrinos tend to be along the beam direction of $\mu^+\mu^-$ and thus the transverse momenta of the (anti-)neutrinos are small. If one neglects the transverse momenta of (anti-)neutrinos along the beam, the transverse missing momentum is the sum of the transverse momenta of the (anti-)neutrinos in the processes $W \to \ell\nu$. Then, further assuming the mass of the W boson is negligible compared with $\hat{s}$, the situation is approximately the

same as that investigated in refs. [41, 79] which deal with the reconstruction of $\hat{s}$ with two (anti-)neutrinos, and $\hat{s}$ can be approximately given by

$$\sqrt{\hat{s}_{\text{lep}}} \approx c_1 + c_2(E_{\ell^+} + E_{\ell^-}) + (c_3 + c_4(E_{\ell^+} + E_{\ell^-}) + c_5 E_{\ell^+} E_{\ell^-}) \cos(\theta_{\ell\ell}), \qquad (4.1)$$

where $E_{\ell^\pm}$ are the energies of charged leptons, $\theta_{\ell\ell}$ is the angle between the charged leptons, and c_i are the parameters to be fitted.

Eq. (4.1) is in fact not obtained by kinematic analysis but by using machine learning approach. With the help of some kinematic analysis and approximations, we translate our problem with four (anti-)neutrinos into the one with two (anti-)neutrinos solved by eq. (4.1). Using kinematic analysis, there is another approximation which is $\hat{s}_{\text{ap}}$ in refs. [41, 79]. $\hat{s}_{\text{ap}}$ is less accurate than $\hat{s}_{\text{lep}}$, therefore is not discussed in this paper. However, the procedure of deriving the $\hat{s}_{\text{ap}}$ shows that, although it is not possible to give the exact $\hat{s}$ by the information in the final state, it is possible to give the most likely one. Using the ANN is merely an improvement and an automation of this complicated procedure.

4.2   The neural network approach

We trained three ANNs to reconstruct $\hat{s}$, which correspond to O_{S_i}, O_{M_i} and O_{T_i} operators. The architecture of the ANNs as well as the 14-dimensional vectors to feed the input are the same as those used in section 3.2. For the output layer, the ground truth of $\hat{s}$ is estimated as $\hat{s}_{\text{tr}} = (p_{\ell^+} + p_{\ell^-} + p_\nu + p_{\bar{\nu}})^2$, where p_ν ($p_{\bar{\nu}}$) is the 4-momentum of the neutrino (anti-neutrino) with the same flavor as ℓ^+ (ℓ^-) and with the direction closest to ℓ^+ (ℓ^-). Since the Les-House event files created by `MadGraph5_aMC@NLO` contains the information of intermediate $W^\pm$ for events from NP squared terms, it is possible to verify the correctness of $\hat{s}_{\text{tr}}$. For the events from O_{S_0} and for all events containing intermediate $W^\pm$, we find the rate of mismatch is smaller than 0.0001%.

To construct the data-sets, one million events are generated for each operator. We only include the events in $\sigma_{4\text{W}}$. They are separated into training data-sets and validation data-sets as done in the previous section. For O_{S_i}, O_{M_i} and O_{T_i} operators, about 1.50, 1.03 and 1.21 million events are included in each data-set, respectively. The learning curves are shown in figure 6. To avoid overfitting, we stop at the epoch when the accuracy of validation data-set starts to fall. Based on the learning curves in figure 6, we choose to stop at 100, 200 and 500 epoches for O_{S_i}, O_{M_i} and O_{T_i} operators. Denoting $\hat{s}_{\text{ann}}$ as the prediction of $\hat{s}$ by the ANNs, for the validation data-sets, the normalized distributions of relative differences defined as $\Delta\sqrt{\hat{s}}/\sqrt{\hat{s}_{\text{tr}}}$ are shown in figure 7. For comparison with $\hat{s}_{\text{lep}}$, c_i in eq. (4.1) are fitted with the training data-sets and the results are listed in table 4. The normalized distributions of relative differences for $\hat{s}_{\text{lep}}$ are also shown in figure 7. For both $\hat{s}_{\text{ann}}$ and $\hat{s}_{\text{lep}}$, the predictions for most events can be smaller than 40%. One can find that the ANNs are able to predict $\hat{s}$ more accurately than $\hat{s}_{\text{lep}}$. If one requires the relative difference to be smaller than 20%, using $\hat{s}_{\text{ann}}$, 76.8%, 81.9% and 81.8% of the O_{S_i}, O_{M_i} and O_{T_i} events satisfy the requirement, compared with 72.6%, 78.6% and 66.7% for using $\hat{s}_{\text{lep}}$.
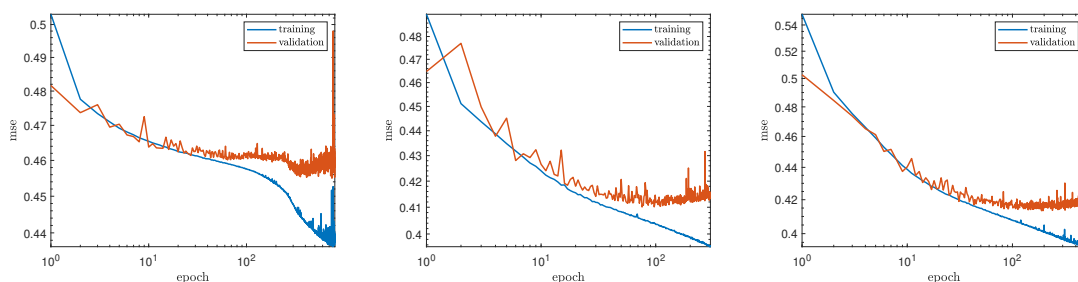
Figure 6. The learning curves of the ANNs trained for $\sqrt{\hat{s}}$. The left (middle) [right] panel corresponds to O_{S_i} (O_{M_i}) $[O_{T_i}]$.
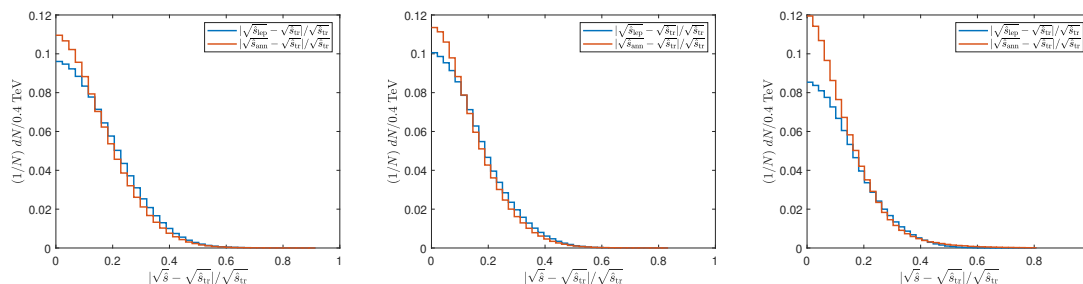

Figure 7. The normalized distributions of $|\Delta\sqrt{\hat{s}}|$. The left (middle) [right] panel corresponds to O_{S_i} (O_{M_i}) $[O_{T_i}]$.

	c_1 (TeV)	c_2	c_3 (TeV)	c_4	c_5 (TeV^{-1})
O_{S_i}	7.39	-0.131	-3.64	-1.11	0.0667
O_{M_i}	9.11	-0.0203	-5.09	-0.680	0.0236
O_{T_i}	10.15	-0.0448	-8.94	-0.141	-0.0351

Table 4. Results of c_i in eq. (4.1) fitted with the training data-sets.

5 Signal significance

5.1 The partial wave unitarity bound

As an EFT, the SMEFT is only valid under the NP energy scale Λ. The large $\hat{s}$ at the muon collider provides a great chance to detect the NP. Meanwhile, the verification of the validity of the SMEFT becomes inevitable. The partial wave unitarity has been widely used in previous studies as an indicator of the SMEFT validation [73, 77, 90, 101–106]. For a VBS process $W^+_{\lambda_1}W^-_{\lambda_2} \to W^-_{\lambda_3}W^+_{\lambda_4}$ with $\lambda_{1,2,3,4} = \pm 1, 0$ corresponding to the helicities of the vector bosons, in the c.m. frame with **z**-axis along the flight direction of W^- in the initial state, the amplitudes can be expanded as [107]

$$\mathcal{M}(W^-_{\lambda_1}W^+_{\lambda_2} \to W^-_{\lambda_3}W^+_{\lambda_4}) = 8\pi \sum_J (2J+1)\, e^{i(\lambda-\lambda')\phi} d^J_{\lambda\lambda'}(\theta) T^J, \tag{5.1}$$

where θ and ϕ are zenith and azimuth angles of the W^- boson in the final state, $\lambda = \lambda_1 - \lambda_2$, $\lambda' = \lambda_3 - \lambda_4$ and $d^J_{\lambda\lambda'}(\theta)$ are the Wigner D-functions. The partial wave unitarity bound is

$|T^J| \leq 2$ [47]. With the helicity amplitudes calculated in appendix B, and assuming one operator at a time, the tightest bounds are

$$\hat{s}^2 < \frac{48\pi\Lambda^4}{|f_{S_0}|}, \quad \hat{s}^2 < \frac{24\pi\Lambda^4}{|f_{S_1}|}, \quad \hat{s}^2 < \frac{24\pi\Lambda^4}{|f_{S_2}|},$$
$$\hat{s}^2 < \frac{32\pi\Lambda^4}{|f_{M_0}|}, \quad \hat{s}^2 < \frac{128\pi\Lambda^4}{|f_{M_1}|}, \quad \hat{s}^2 < \frac{256\pi\Lambda^4}{|f_{M_7}|}, \quad (5.2)$$
$$\hat{s}^2 < \frac{6\pi\Lambda^4}{|f_{T_0}|}, \quad \hat{s}^2 < \frac{8\pi\Lambda^4}{|f_{T_1}|}, \quad \hat{s}^2 < \frac{16\pi\Lambda^4}{|f_{T_2}|}.$$

In this paper, the unitarity bounds are applied using a matching procedure [108, 109] which has been used in previous studies of aQGCs at the LHC [41, 78, 80]. We compare the cross-sections with and without aQGCs under a certain energy scale. It should be emphasized that, such unitrization procedure introduces no extra assumptions. This is important because it has been pointed out that different unitrization methods lead to different results [110], and therefore the unitrization methods introducing extra assumptions actually break the model-independence principle of the SMEFT [1].

In fact, in our approach, no bounds or constraints are applied, despite a misleading "bounds" in the name of this procedure. Studying the Wilson coefficient within an energy range is standard for an EFT because the Wilson coefficients typically depend on energy scales. Namely, even without unitarity bounds, it is a matter of interest to compare the SMEFT and the SM within a certain energy scale. We simply choose the energy scale according to the coefficients such that unitarity is guaranteed.

The suppression of cross-section when unitarity bounds are considered has been noticed in refs. [41, 78, 80], and demonstrates the necessity of the unitarity bounds. The effect of unitarity bounds can be estimated in terms of EVA (see appendix. A). To illustrate the necessity of unitarity bounds, we compared the cases with and without unitarity bounds at $\sqrt{s} = 30$ TeV in figure 8. As a verification, the results from M.C. simulation are also shown with $\sqrt{\hat{s}_{\mathrm{tr}}}$ in section 4.2 used for unitarity bound. It can be seen in figure 8 that, the cross-section is suppressed significantly and is no longer a bilinear function of f_{S_0} after the unitarity bound is applied, which is also seen in refs. [41, 78, 80].

After event selection strategy, the suppressed efficiencies of unitarity bounds will change. Thus, we cannot use eq. (A.6) to estimate the effect of unitarity bounds after the event selection strategy. Also, we cannot use $\sqrt{\hat{s}_{\mathrm{tr}}}$ to apply the unitarity bounds in M.C. simulation because $\sqrt{\hat{s}_{\mathrm{tr}}}$ is not observable. Instead, with ANNs trained to reconstruct $\hat{s}$ at hand, we use the $\hat{s}_{\mathrm{ann}}$ for unitarity bounds. Figure 8 is only for an illustration that the unitarity bounds are necessary.

5.2 Signal and backgrounds

The major SM background is $\mu^+\mu^- \to \ell^+\ell^- + \not{E}$. We consider the processes with up to four (anti-)neutrinos and the Feynman diagrams are shown in figure 9. Using M.C. simulation, at $\sqrt{s} = 30$ TeV, we find the total SM cross-section as $\sigma_{\mathrm{SM}} = 176.8$ fb. The dominant one is $\mu^+\mu^- \to \ell^+\ell^-\nu\bar{\nu}$ whose cross-section is about 155.2 fb.

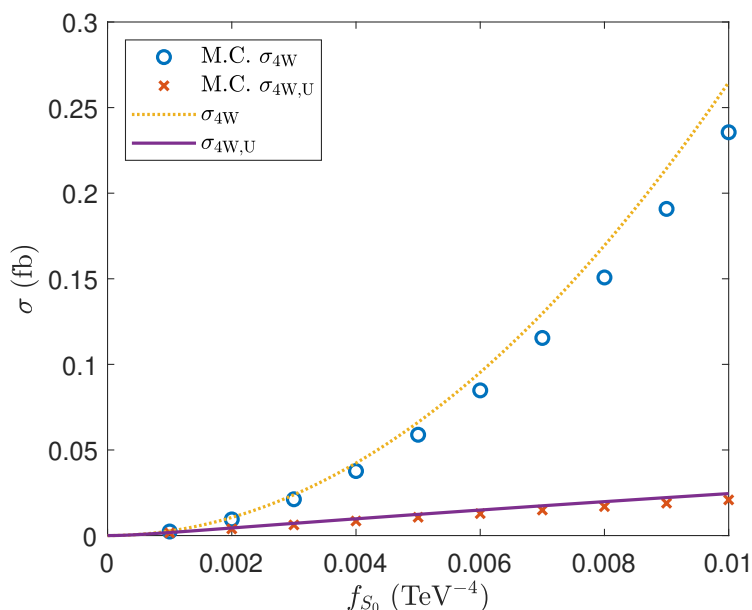**Figure 8**. $\sigma_{4W}(\mu^+\mu^- \to \bar{\nu}\nu W^+W^- \to \bar{\nu}\nu\bar{\nu}\nu\ell^+\ell^-)$ compared with $\sigma_{4W,U}(\mu^+\mu^- \to \bar{\nu}\nu W^+W^- \to \bar{\nu}\nu\bar{\nu}\nu\ell^+\ell^-)$ in eq. (A.6). This is only for an illustration that the unitarity bounds are necessary.

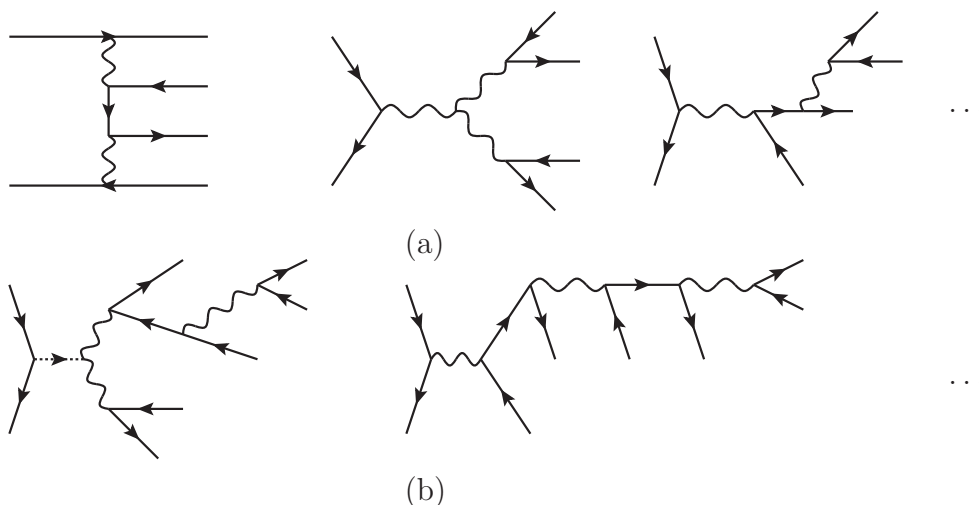**Figure 9**. The tree level Feynman diagrams of the SM contribution to the process $\mu^+\mu^- \to \ell^+\ell^- + \not{E}$.

The signal significance is estimated using the definition $\mathcal{S}_{\mathrm{stat}} = N_s/\sqrt{N_s + N_{\mathrm{bg}}}$ where N_s (N_{bg}) is the number of signal (backgrounds) events. We define the cross-section with unitarity bounds as $\sigma_{4W,U}$ as shown in eq. (A.5). Before performing the M.C. simulation, $\sigma_{4W,U}$ is used to initially predict the constraints on coefficients, as well as to determine the range of parameter space. The expected luminosity of the $\mu^+\mu^-$ collider at $\sqrt{s} = 30$ TeV is about $L = 10$ ab$^{-1} \sim 90$ ab^{-1} [34]. By requiring $\mathcal{S}_{\mathrm{stat}} \approx 2 \sim 5$ before event selection strategies and after the unitarity bounds are applied, we choose the coefficient spaces
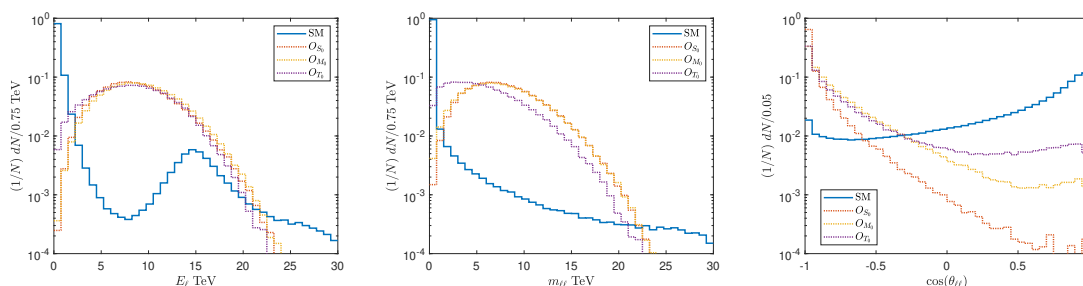
Figure 10. The normalized distributions of E_ℓ (left panel), $m_{\ell\ell}$ (middle panel) and $\cos(\theta_{\ell\ell})$ (right panel).

satisfying $\sigma_{4W,U} \approx 0.1\% \times \sigma_{SM}$. The largest coefficients are calculated according to eqs. (A.6)–(A.13) and listed in table 5. The coefficients for O_{S_i} are much larger than O_{M_i} and O_{T_i}, because the suppression of the unitarity bounds is more important for O_{S_i} operators. This is because the dominant helicity amplitude for each O_{S_i} operator is longitudinal scattering $W_0 W_0 \to W_0 W_0$ as shown in eq. (B.1). However, the luminosity of transverse polarized W bosons from the beam is logarithmically enhanced, as shown in eq. (A.1). As a result, to produce cross-sections of the same order of magnitude, f_{S_i} should be much larger than f_{M_i} and f_{T_i}. Meanwhile, the unitarity bounds are set by the amplitudes of $WW \to WW$, so that the suppression is at the same order of magnitude for all operators. In return, a larger f_{S_i} suffers from a more significant suppression, and an even larger f_{S_i} is required as a consequence.

To study the kinematic features of the signals and the background, in the following, a fast detector simulation is performed by `Delphes` [111] with the muon collider card. We cut off the events that do not contain two opposite-sign leptons or with at least two charged leptons but the hardest two have a same-sign (denoted as N_ℓ cut). The kinematic features are shown after the N_ℓ cut.

Compared with the VBS contribution from the aQGCs, the $\hat{s}$ of the VBS in the SM is smaller. As a result, a smaller $E_\ell = E_{\ell^+} + E_{\ell^-}$ is expected for the SM. For the same reason, the invariant masses of charged leptons $m_{\ell\ell} = \sqrt{(p_{\ell^+} + p_{\ell^-})^2}$ of the signal events are larger. For the signal events, with an energetic W boson, the charged lepton should be approximately collinear to the W boson. The values of $\hat{s}$ of the signal events are large, so that the produced $W^\pm$ bosons tend to be approximately back-to-back for a signal event. As a result, the angle between charged leptons (denoted as $\theta_{\ell\ell}$) should be close to π. Taking O_{S_0}, O_{M_0} and O_{T_0} as examples, the normalized distributions of E_ℓ, $m_{\ell\ell}$ and $\cos(\theta_{\ell\ell})$ are shown in figure 10. It can be seen that E_ℓ, $m_{\ell\ell}$ and $\cos(\theta_{\ell\ell})$ for the signal events are very different from those of the background events.

To constrain the aQGC operators, different event selection strategies can be designed for different operators. The event selection strategies and effects of the cuts are summarized in table 5. Here the W_{score} cuts are applied to filter out the $WW \to WW$ contribution induced by aQGCs from other aQGCs contributions, as shown in table 3. An additional bonus is that the W_{score} cuts can also significantly suppress the SM contribution. The

f_{S_i}/Λ^4	SM	O_{S_0} 0.1	O_{S_1} 0.1	O_{S_2} 0.1
σ	176805	23559	56280	56326
N_ℓ cut	140132	18054	43441	43570
2.2 TeV $< E_\ell <$ 22 TeV	7786.7	17835	42912	43046
2 TeV $< m_{\ell\ell} <$ 20 TeV	4138.8	17594	42331	42460
$\cos(\theta_{\ell\ell}) < -0.3$	2632.3	17378		41758
$\cos(\theta_{\ell\ell}) < -0.8$	1674.8		36049	

f_{M_i}/Λ^4	SM	O_{M_0} 0.002	O_{M_1} 0.004	O_{M_7} 0.008
σ	176805	964.0	714.2	955.5
N_ℓ cut	140132	743.8	548.8	738.0
2.2 TeV $< E_\ell <$ 22 TeV	7786.7	735.9	542.2	728.5
2.2 TeV $< m_{\ell\ell} <$ 20 TeV	3935.0	709.7	516.2	684.7
$W_{\text{score}} > 0.85$	964.9	402.0	273.2	
$W_{\text{score}} > 0.95$	403.8			220.8

f_{T_i}/Λ^4	SM	O_{T_0} 0.0002	O_{T_1} 0.0003	O_{T_2} 0.0005
σ	176805	621.7	944.8	842.8
N_ℓ cut	140132	485.2	726.0	649.4
2.2 TeV $< E_\ell <$ 22 TeV	7786.7	462.2	688.2	603.0
2.8 TeV $< m_{\ell\ell} <$ 18 TeV	3288.0	366.8	553.3	477.1
$W_{\text{score}} > 0.85$	482.0		362.5	346.4
$W_{\text{score}} > 0.9$	299.2	246.2		

Table 5. Cross-sections (ab) after cuts at $\sqrt{s} = 30$ TeV.

O_{S_i} operators dominantly contribute to the $WW \to WW$ process. Therefore, no ANN is trained to identify the $WW \to WW$ contribution and W_{score} cuts are not applied for O_{S_i} operators. We find that the event selection strategies can suppress the backgrounds significantly, while keeping most of the signal events.

The interference has been neglected until now. For the coefficients listed in table 5, the contributions of interference terms compared with those of NP squared terms are shown in table 6. One can see that the interference of O_{T_i} operators is about 10% of the squared term and should be taken into account. However, the interference does not only come from the interference of $WW \to WW$ but also from other diagrams such as tri-boson diagrams, which are meant to be cut off. Apart from that, the event selection strategy designed to cut off background should also suppress the interference contribution. Therefore, taking O_{T_2} as an example (whose interference is the greatest among all operators in study), we investigate the effect of event selection strategy on the interference. After the event selection strategy, σ_{int} becomes 26.4 ab, which is less than 8% of σ_{NP} after event selection strategy is applied.

	f_X/Λ^4 $(10^{-3}\text{ TeV}^{-4})$	σ_{NP}	σ_{int}
O_{S_0}	100	23559	37.2
O_{S_1}	100	56280	-6.3
O_{S_2}	100	56326	-0.81
O_{M_0}	2	964.0	0.28
O_{M_1}	4	714.2	18.5
O_{M_7}	8	955.5	-23.3
O_{T_0}	0.2	621.7	47.7
O_{T_1}	0.3	944.8	76.5
O_{T_2}	0.5	842.8	93.5

Table 6. Cross-sections (ab) of interference terms and squared terms.

5.3 Signal significance and the expected constraints

As introduced, we compare the cross-sections under a certain energy scale corresponding to the coefficients of dimension-8 operators. As a result, due to the energy cuts caused by unitarity bounds depending on the coefficients, the cross-section of the SM appears to be functions of coefficients. From another point of view, the cross-section of the SM does not actually depend on the coefficients of the operators, but on a certain energy scale that we have selected.

The cross-sections after applying unitarity bounds are shown in figure 11. The upper bounds of $\hat{s}$ are calculated using eq. (5.2) and denoted as $\hat{s}_U^{O_X}$ for each operator O_X. It can be found that, in the ranges of coefficients considered in this paper, the cross-sections of the SM are typically $\sim \mathcal{O}(1)$ fb, and the cross-sections of NP are typically less than 0.1 fb.

For luminosities $L = 10$ ab^{-1} and $L = 90$ ab^{-1} [34], the expected constraints on the absolute values of coefficients ($|f_X|$) are obtained with the help of signal significance and assuming one operator at a time. Since there are still errors between $\hat{s}_{\text{ann}}$ and $\hat{s}_{\text{tr}}$, moreover, the EFT probably stops being valid before the unitarity limit. Therefore, the robustness of the results are studied by varying the cut-off on $\hat{s}_{\text{ann}}^2$ by factors $1/2$ and 2, analogous to what has been done with the QCD scales for the study of aQGCs at the LHC [80]. The results are listed in tables 7 and 8. The expected constraints with $2\hat{s}_U^2$ and $\hat{s}_U^2/2$ used as unitarity bounds are presented as systematic errors in tables 7 and 8.

As introduced, the O_{S_i} operators are significantly affected by the unitarity bounds, which can also be seen from tables 7 and 8. Take the expected constraints at $\mathcal{S}_{\text{stat}} = 2$, for O_{S_i} operators in general and O_{M_i} operators at $L = 10$ ab^{-1}, our results only show the orders of magnitude. For O_{T_i} operators in general and O_{M_i} operators at $L = 90$ ab^{-1}, however, our results do not rely on the unitarity bounds and are thus more meaningful for experiments. This is a representation of the "EFT triangle" problem [38–41], which can be solved by high luminosity. As a consequence, one can see that the results for O_{M_i} and O_{T_i} operators at $L = 90$ ab^{-1} are more reliable. Generally, the O_{T_i} operators are seldom affected by the unitarity bounds. This is because the luminosity of the transverse W bosons
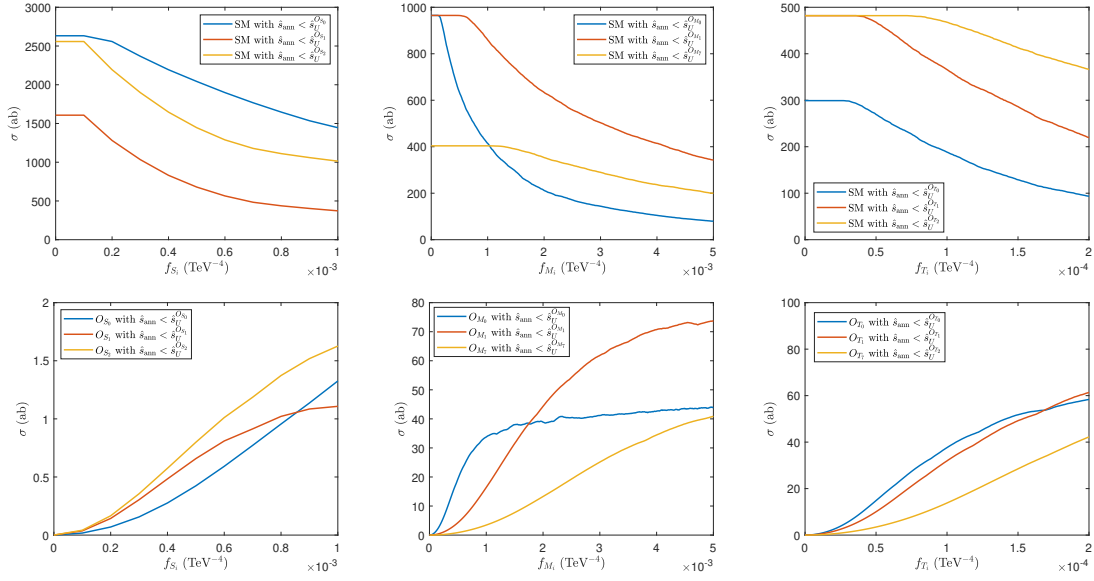
Figure 11. The cross-sections after unitarity bounds are applied.

$\mathcal{S}_{\text{stat}}$	f_{S_0}/Λ^4	f_{S_1}/Λ^4	f_{S_2}/Λ^4
2	$35.3^{+39.5}_{-27.5}$	$61.4^{+21.8}_{-15.6}$	$42.4^{+10.8}_{-0.3}$
3	$72.0^{+20.7}_{-59.5}$	$78.0^{+30.0}_{-5.2}$	$60.0^{+14.6}_{-1.0}$
5	$111.7^{+48.0}_{-87.7}$	$166.3^{+48.6}_{-52.9}$	$95.8^{+12.5}_{-4.3}$
$\mathcal{S}_{\text{stat}}$	f_{M_0}/Λ^4	f_{M_1}/Λ^4	f_{M_7}/Λ^4
2	$0.45^{+0.59}_{-0.01}$	$1.08^{+0.31}_{-0}$	$1.91^{+0.34}_{-0}$
3	$0.59^{+1.28}_{-0.05}$	$1.35^{+1.07}_{-0.03}$	$2.37^{+1.31}_{-0}$
5	$0.99^{+4.15}_{-0.29}$	$1.90^{+3.88}_{-0.19}$	$3.23^{+8.92}_{-0.15}$
$\mathcal{S}_{\text{stat}}$	f_{T_0}/Λ^4	f_{T_1}/Λ^4	f_{T_2}/Λ^4
2	$0.043^{+0.006}_{-0}$	$0.059^{+0.010}_{-0.001}$	$0.101^{+0.012}_{-0}$
3	$0.053^{+0.014}_{-0}$	$0.073^{+0.023}_{-0.001}$	$0.124^{+0.035}_{-0}$
5	$0.072^{+0.044}_{-0.004}$	$0.100^{+0.069}_{-0.005}$	$0.166^{+0.109}_{-0.004}$

Table 7. The expected constraints on the absolute values of coefficients (10^{-3} TeV^{-4}) at $\sqrt{s} = 30$ TeV and $L = 10$ ab^{-1}. The results by varying the unitarity bounds by factors $1/2$ and 2 are presented as systematic errors.

from the beam is logarithmically enhanced, and all the dominant helicity amplitudes of O_{T_i} contain transverse W bosons in the initial state of $WW \to WW$. Compared with table 1 which does not consider unitarity bounds, the coefficients can be narrowed down significantly even with unitarity bounds considered. The expected constraints can be 3 to 4 orders of magnitude stronger than those at the 13 TeV LHC for O_{M_i} and O_{T_i} operators.

$\mathcal{S}_{\text{stat}}$	f_{S_0}/Λ^4	f_{S_1}/Λ^4	f_{S_2}/Λ^4
2	$5.3^{+21.1}_{-2.6}$	$31.3^{+0.8}_{-29.5}$	$5.9^{+19.4}_{-3.8}$
3	$8.7^{+34.6}_{-5.0}$	$40.8^{+1.4}_{-36.8}$	$22.9^{+9.3}_{-19.6}$
5	$18.2^{+41.7}_{-11.9}$	$56.9^{+23.0}_{-17.9}$	$35.0^{+13.8}_{-29.1}$
$\mathcal{S}_{\text{stat}}$	f_{M_0}/Λ^4	f_{M_1}/Λ^4	f_{M_7}/Λ^4
2	$0.26^{+0.01}_{-0}$	$0.62^{+0.01}_{-0}$	$1.12^{+0.01}_{-0}$
3	$0.31^{+0.07}_{-0}$	$0.76^{+0.04}_{-0}$	$1.36^{+0.02}_{-0}$
5	$0.40^{+0.32}_{-0}$	$0.98^{+0.19}_{-0}$	$1.75^{+0.19}_{-0}$
$\mathcal{S}_{\text{stat}}$	f_{T_0}/Λ^4	f_{T_1}/Λ^4	f_{T_2}/Λ^4
2	0.025^{+0}_{-0}	0.034^{+0}_{-0}	0.058^{+0}_{-0}
3	$0.030^{+0.001}_{-0}$	$0.042^{+0.001}_{-0}$	0.072^{+0}_{-0}
5	$0.039^{+0.003}_{-0}$	$0.054^{+0.006}_{-0}$	0.092^{+0}_{-0}

Table 8. The expected constraints on the absolute values of coefficients (10^{-3} TeV^{-4}) at $\sqrt{s} = 30$ TeV and $L = 90$ ab^{-1}. The results by varying the unitarity bounds by factors 1/2 and 2 are presented as systematic errors.

6 Summary

As a gauge boson collider, the muon collider is suitable to study the VBS processes for the aQGCs. Unlike the LHC, with cleaner final states, it is easier to extract aQGCs out of all relevant dimension-8 operators. For example, exclusive $W^+W^- \rightarrow W^+W^-$ can be separated out from $VV \rightarrow W^+W^-$ processes. In this paper, we study the process $\mu^+\mu^- \rightarrow \ell^+\ell^-\nu\nu\bar{\nu}\bar{\nu}$ containing a $W^+W^- \rightarrow W^+W^-$ subprocess induced by dimension-8 operators contributing to aQGCs.

The presence of four (anti-)neutrinos in the final states poses difficulties to the phenomenological study. In this paper, this problem is tackled with a machine learning approach. The ANNs can be used to pick out the contribution including a $W^+W^- \rightarrow W^+W^-$ subprocess. Moreover, the ANNs can also be used to reconstruct $\hat{s}$ of the $W^+W^- \rightarrow W^+W^-$ subprocess.

With the help of the ANNs, the sensitivities of the process $\mu^+\mu^- \rightarrow \ell^+\ell^-\nu\nu\bar{\nu}\bar{\nu}$ to the dimension-8 operators contributing to aQGCs are studied at the muon collider with $\sqrt{s} = 30$ TeV. The kinematic features are investigated, and the event selection strategies are proposed. The unitarity bounds are also considered, which turn out to be necessary in the study of the $W^+W^- \rightarrow W^+W^-$ subprocess at a muon collider. The expected constraints are studied with the help of signal significance, and are found to be as small as 4 orders of magnitude stronger than those at the 13 TeV LHC even with unitarity bounds applied. This shows the great advantage of the muon collider in studying the aQGCs.

Acknowledgments

A Effective vector boson approximation

As a comparison, the contribution of NP involving a $W^+W^- \to W^+W^-$ subprocess is calculated by using EVA. The $W^+W^- \to W^+W^-$ contribution to the process $\mu^+\mu^- \to \nu\nu\bar{\nu}\bar{\nu}\ell^+\ell^-$ can be given by EVA as [95–97]

$$
\sigma_{\mathrm{VBS}}(\mu^+\mu^- \to \bar{\nu}\nu W^+W^-) = \sum_{\lambda_1\lambda_2\lambda_3\lambda_4} \int d\xi_1 \int d\xi_2 f_{W^-_{\lambda_1}/\mu^-}(\xi_1) f_{W^+_{\lambda_2}/\mu^+}(\xi_2) \sigma_{W^+_{\lambda_1}W^-_{\lambda_2}\to W^+_{\lambda_3}W^-_{\lambda_4}}(\hat{s}),
$$

$$
f_{W^-_{+1}/\mu^-_L}(\xi) = f_{W^+_{-1}/\mu^+_L}(\xi) = \frac{e^2}{8\pi^2 s_W^2}\frac{(1-\xi)^2}{2\xi}\log\frac{\mu_f^2}{M_W},
$$

$$
f_{W^-_{-1}/\mu^-_L}(\xi) = f_{W^+_{+1}/\mu^+_L}(\xi) = \frac{e^2}{8\pi^2 s_W^2}\frac{1}{2\xi}\log\frac{\mu_f^2}{M_W},
$$

$$
f_{W^-_0/\mu^-_L}(\xi) = f_{W^+_0/\mu^+_L}(\xi) = \frac{e^2}{8\pi^2 s_W^2}\frac{1-\xi}{\xi},
$$

$$
f_{W^\pm_\lambda/\mu^\pm_R} = 0, \quad f_{W^\pm_\lambda/\mu^\pm_R} = \frac{f_{W^\pm_\lambda/\mu^\pm_L} + f_{W^\pm_\lambda/\mu^\pm_R}}{2},
$$

$$(A.1)$$

where $\sqrt{\hat{s}} = \sqrt{\xi_1\xi_2 s}$ and μ_f is the factorization scale which can be set to be $\sqrt{\hat{s}}/4$ [97], s is the collision energy. At the leading order of s, one has

$$
\sigma_{\mathrm{VBS}}(\mu^+\mu^- \to \bar{\nu}\nu W^+W^-) = \frac{e^4 s^3}{16986931200000\pi^5 s_W^4 \Lambda^8}
$$

$$
\times \Bigg\{ 40\log\left(\frac{s}{16M_W^2}\right)\Big[17385600 f_{M_0}^2 + 2400 f_{M_0}(2759 f_{M_7} - 3622 f_{M_1}) - 1154484 f_{M_1}^2
$$

$$
+ 585684 f_{M_1} f_{M_7} - 134871 f_{M_7}^2 + 128\Big[7500\left(6 f_{S_0}^2 + 11 f_{S_0}(f_{S_1} + f_{S_2}) + 14(f_{S_1} + f_{S_2})^2\right)
$$

$$
+ 3744668 f_{T_0}^2 + 4574508 f_{T_0} f_{T_1} + 2688508 f_{T_0} f_{T_2} + 2078155 f_{T_1}^2 + 2052462 f_{T_1} f_{T_2} + 665294 f_{T_2}^2\Big]
$$

$$
+ 15\log\left(\frac{s}{16M_W^2}\right)\left(9600 f_{M_0}^2 + 2400 f_{M_0}(f_{M_7} - 2 f_{M_1}) + 1004 f_{M_1}^2 - 1004 f_{M_1} f_{M_7} + 301 f_{M_7}^2\right.
$$

$$
+ 128\left(4652 f_{T_0}^2 + 5532 f_{T_0} f_{T_1} + 3532 f_{T_0} f_{T_2} + 3055 f_{T_1}^2 + 3198 f_{T_1} f_{T_2} + 1046 f_{T_2}^2\right)\big)
$$

$$
- 163200 f_{M_0}^2 + 600 f_{M_0}(136 f_{M_1} - 167 f_{M_7}) + 47612 f_{M_1}^2 - 32762 f_{M_1} f_{M_7} + 9403 f_{M_7}^2
$$

$$
- 128\left(139804 f_{T_0}^2 + 169284 f_{T_0} f_{T_1} + 102284 f_{T_0} f_{T_2} + 82295 f_{T_1}^2 + 83076 f_{T_1} f_{T_2} + 27022 f_{T_2}^2\right)\Big]\Bigg\}.
$$

$$(A.2)$$

With $\mathrm{Br}(W \to e\nu_e) \approx \mathrm{Br}(W \to \mu\nu_\mu) \approx 10.8\%$ [112], the predictions of σ_{4W} by eq. (A.2) (denoted as $\sigma_{4W}^{\mathrm{EVA}}$) can be calculated.

The differential cross-section is also calculated to estimate the effect of unitarity bounds, which is

$$
\begin{aligned}
\frac{d\sigma_{4W}(\mu^+\mu^- \to \bar{\nu}\nu W^+W^-)}{d\hat{s}} \\
= -\frac{e^4\hat{s}^2}{31457280\pi^5\Lambda^8 s^2 s_W^4} & \left\{ 8s(A_0+A_2)\left(2(s-\hat{s})+(s+\hat{s})\log\left(\frac{\hat{s}}{s}\right)\right) \right. \\
& + \log^2\left(\frac{\hat{s}}{16M_W^2}\right)\left[(s-\hat{s})(\hat{s}(-(-4A_1+A_2+A_3)))+\left(A_1\left(4s\hat{s}+\hat{s}^2\right)\right)\log\left(\frac{\hat{s}}{s}\right)\right. \quad \text{(A.3)} \\
& \left. + (A_2+A_3)\left(2s^2\log\left(\frac{\hat{s}}{s}\right)+3s(s-\hat{s})\right)\right] \\
& \left. + A_4\log\left(\frac{\hat{s}}{16M_W^2}\right)\left((s-\hat{s})(7s+\hat{s})+4s(s+\hat{s})\log\left(\frac{\hat{s}}{s}\right)\right)\right\},
\end{aligned}
$$

where

$$
\begin{aligned}
A_0 &= 32\left(6f_{S_0}^2 + 11f_{S_0}(f_{S_1}+f_{S_2}) + 14(f_{S_1}+f_{S_2})^2\right), \\
A_1 &= (f_{M_7}-2f_{M_1})^2 + 384\left(4f_{T_0}^2 + 4f_{T_0}(f_{T_1}+f_{T_2}) + 5f_{T_1}^2 + 6f_{T_1}f_{T_2} + 2f_{T_2}^2\right), \\
A_2 &= 480f_{M_0}^2 + 120f_{M_0}(f_{M_7}-2f_{M_1}) + 34f_{M_1}^2 - 34f_{M_1}f_{M_7} + 11f_{M_7}^2, \quad \text{(A.4)} \\
A_3 &= 128\left(184f_{T_0}^2 + 228f_{T_0}f_{T_1} + 128f_{T_0}f_{T_2} + 92f_{T_1}^2 + 87f_{T_1}f_{T_2} + 28f_{T_2}^2\right), \\
A_4 &= 2\left(96f_{M_0}^2 - 12f_{M_0}(4f_{M_1}+f_{M_7}) + 46f_{M_1}^2 - 37f_{M_1}f_{M_7} + 11f_{M_7}^2\right).
\end{aligned}
$$

The cross-section with unitarity bounds is then estimated as

$$
\sigma_{4W,U} = \int_0^{\min\{\hat{s}_{\max},s\}} d\hat{s}\frac{d\sigma_{4W}}{d\hat{s}}, \quad \text{(A.5)}
$$

where $\hat{s}_{\max}$ is the maximally allowed $\hat{s}$ in the sense of unitarity. Assuming one operator at a time, $\hat{s}_{\max}$ is given in eq. (5.2). As an example to illustrate the effect of unitarity bounds, consider only the existence of O_{S_0},

$$
\begin{aligned}
\sigma_{4W}(\mu^+\mu^- \to \bar{\nu}\nu W^+W^-) &= \frac{e^4 f_{S_0}^2 s^3}{2949120\pi^5\Lambda^8 s_W^4}, \\
\sigma_{4W,U}(\mu^+\mu^- \to \bar{\nu}\nu W^+W^-) &= \frac{e^4 f_{S_0}^2 s^3(1-B_{S_0})^3}{2949120\pi^5 s_W^4} \\
& \quad \times \left[1+9B_{S_0}\left(4\log\left(1-B_{S_0}\right)-9\right)-84\log\left(1-B_{S_0}\right)\right], \quad \text{(A.6)} \\
B_{S_0} &= \left(1-\frac{4\sqrt{3}\pi}{s\sqrt{|f_{S_0}|}}\right)\theta\left(s-\frac{4\sqrt{3}\pi}{\sqrt{|f_{S_0}|}}\right),
\end{aligned}
$$

where $\sigma_{4W,U}(\mu^+\mu^- \to \ell^+\ell^-\bar{\nu}\bar{\nu}\nu\nu) = \sigma_{4W,U}(\mu^+\mu^- \to \bar{\nu}\nu W^+W^-) \times (\mathrm{Br}(W \to \ell\nu))^2$, $\sigma_{4W,U}(\mu^+\mu^- \to \ell^+\ell^-\bar{\nu}\bar{\nu}\nu\nu)$ denotes the cross-section of $W^+W^- \to W^+W^-$ induced

$\mu^+\mu^- \to \ell^+\ell^-\nu\nu\bar{\nu}\bar{\nu}$ with unitarity bound applied, $\theta(x)$ is the Heaviside unit step function. Similarly,

$$\sigma_{4W,U}^{O_{S_{1,2}}}(\mu^+\mu^- \to \bar{\nu}\nu W^+W^-) = \frac{7e^4 f_{S_{1,2}}^2 s^3 (1-B_{S_{1,2}})^3}{8847360\pi^5 s_W^4}$$
$$\times \left[1 + 9B_{S_{1,2}}\left(4\log\left(1-B_{S_{1,2}}\right) - 9\right) - 84\log\left(1-B_{S_{1,2}}\right)\right], \tag{A.7}$$

$$\sigma_{4W,U}^{O_{M_0}}(\mu^+\mu^- \to \bar{\nu}\nu W^+W^-)$$
$$= \frac{(B_{M_0}-1)^3 e^4 s^3 f_{M_0}^2}{1769472000\pi^5 s_W^4}\left\{ 122500\log(1-B_{M_0}) \right.$$
$$- \frac{864\sqrt{2\pi}B_{M_0}}{s\sqrt{|f_{M_0}|}}\left[2 - 5\log\left(\frac{(1-B_{M_0})s}{16M_W^2}\right)\right]^2 - 51300 B_{M_0}\log(1-B_{M_0}) \tag{A.8}$$
$$+ 600[36B_{M_0} + 30\log(1-B_{M_0}) - 1]\log^2\left(\frac{(1-B_{M_0})s}{16M_W^2}\right) + 118611 B_{M_0} - 1811$$
$$\left. + 40[243B_{M_0} + 30(11 - 9B_{M_0})\log(1-B_{M_0}) + 17]\log\left(\frac{(1-B_{M_0})s}{16M_W^2}\right)\right\},$$

$$\sigma_{4W,U}^{O_{M_1}}(\mu^+\mu^- \to \bar{\nu}\nu W^+W^-)$$
$$= \frac{(B_{M_1}-1)^3 e^4 s^3 f_{M_1}^2}{4246732800000\pi^5 s_W^4 \left(\frac{8\sqrt{2\pi}}{s\sqrt{|f_{M_1}|}} - 1\right)}\left\{ -9(1607931B_{M_1} + 32069) \right.$$
$$+ 600\log^2\left(\frac{(1-B_{M_1})s}{16M_W^2}\right)\left((251 - 6201B_{M_1}) + \frac{8\sqrt{2\pi}(6975B_{M_1} - 251)}{s\sqrt{|f_{M_1}|}} - \frac{99072\pi B_{M_1}}{s^2|f_{M_1}|}\right.$$
$$\left. - 60\log(1-B_{M_1})\left((103 - 18B_{M_1}) + \frac{8\sqrt{2\pi}(15B_{M_1} - 103)}{s\sqrt{|B_{M_1}|}} + \frac{384\pi B_{M_1}}{s^2|f_{M_1}|}\right)\right)$$
$$+ 40\log\left(\frac{(1-B_{M_1})s}{16M_W^2}\right)\left((11903 - 555903B_{M_1}) + \frac{8\sqrt{2\pi}(489375b - 11903)}{s\sqrt{|B_{M_1}|}} + \frac{8515584\pi B_{M_1}}{s^2|f_{M_1}|}\right.$$
$$\left. + 30\log(1-B_{M_1})\left((10089B_{M_1} - 22189) + \frac{8\sqrt{2\pi}(22189 - 10125B_{M_1})}{s\sqrt{|B_{M_1}|}} + \frac{4608\pi B_{M_1} b}{s^2|f_{M_1}|}\right)\right)$$
$$- 20\log(1-B_{M_1})\left((677557 - 307557B_{M_1}) + \frac{8\sqrt{2\pi}(307125B_{M_1} - 677557)}{s\sqrt{|B_{M_1}|}} + \frac{55296\pi B_{M_1}}{s^2|f_{M_1}|}\right)$$
$$\left. + \frac{72\sqrt{2\pi}(1666875B_{M_1} + 32069)}{s\sqrt{|B_{M_1}|}} - \frac{67903488\pi B_{M_1}}{s^2|f_{M_1}|}\right\}, \tag{A.9}$$

$$\sigma_{4\mathrm{W,U}}^{O_{M_7}}(\mu^+\mu^- \to \bar\nu\nu W^+W^-)$$

$$= \frac{(B_{M_7}-1)^3 e^4 s^3 f_{M_7}^2}{16986931200000\pi^5 s_W^4} \left\{ 33(635913 B_{M_7} + 4087) \right.$$

$$- 600\log^2\left(\frac{(1-B_{M_7})s}{16M_W^2}\right)\left[9B_{M_7}\left(\frac{2176\sqrt{\pi}}{s\sqrt{|f_{M_7}|}} - 889\right)\right.$$

$$\left. + 120\log(1-B_{M_7})\left((9B_{M_7}-64) + \frac{24\sqrt{\pi}B_{M_7}}{s\sqrt{|f_{M_7}|}}\right) + 301\right]$$

$$+ 40\log\left(\frac{(1-B_{M_7})s}{16M_W^2}\right)\left[189B_{M_7}\left(2727 + \frac{5632\sqrt{\pi}}{s\sqrt{|f_{M_7}|}}\right)\right.$$

$$\left. + 30\log(1-B_{M_7})\left(-9639B_{M_7} + \frac{576\sqrt{\pi}B_{M_7}}{s\sqrt{|f_{M_7}|}} + 20639\right) - 9403\right]$$

$$\left. - 20\log(1-B_{M_7})\left((449307b - 1021307) + \frac{6912\sqrt{\pi}B_{M_7}}{s\sqrt{|f_{M_7}|}}\right) - \frac{8487936\sqrt{\pi}B_{M_7}}{s\sqrt{|f_{M_7}|}} \right\},$$

$$\text{(A.10)}$$

$$\sigma_{4\mathrm{W,U}}^{O_{T_0}}(\mu^+\mu^- \to \bar\nu\nu W^+W^-)$$

$$= \frac{(B_{T_0}-1)^3 e^4 s^3 f_{T_0}^2}{33177600000\pi^5 s_W^4\left(\frac{\sqrt{6\pi}}{s\sqrt{|f_{T_0}|}} - 1\right)} \left\{ -\frac{1669248\pi B_{T_0}}{s^2|f_{T_0}|} \right.$$

$$+ 600\log^2\left(\frac{(1-B_{T_0})s}{16M_W^2}\right)\left[(1163 - 33363B_{T_0}) + \frac{\sqrt{6\pi}(39375B_{T_0} - 1163)}{s\sqrt{|f_{T_0}|}}\right.$$

$$\left. - 60\log(1-B_{T_0})\left(2(257-27B_{T_0}) + \frac{\sqrt{6\pi}(45B_{T_0}-514)}{s\sqrt{|f_{T_0}|}} + \frac{54\pi B_{T_0}}{s^2|f_{T_0}|}\right) - \frac{36072\pi B_{T_0}}{s^2|f_{T_0}|}\right]$$

$$+ 40\log\left(\frac{(1-B_{T_0})s}{16M_W^2}\right)\left[(264951B_{T_0} - 34951) + \frac{\sqrt{6\pi}(34951 - 300375B_{T_0})}{s\sqrt{|f_{T_0}|}}\right.$$

$$\left. + 30\log(1-B_{T_0})\left((9983 - 783B_{T_0}) + \frac{\sqrt{6\pi}(675B_{T_0}-9983)}{s\sqrt{|f_{T_0}|}} + \frac{648\pi B_{T_0}}{s^2|f_{T_0}|}\right) + \frac{212544\pi B_{T_0}}{s^2|f_{T_0}|}\right]$$

$$+ (936167 - 2776167B_{T_0}) + \frac{\sqrt{6\pi}(3054375B_{T_0} - 936167)}{s\sqrt{|f_{T_0}|}}$$

$$\left. - 20\log(1-B_{T_0})\left(81B_{T_0}\left(\frac{125\sqrt{6\pi}}{s\sqrt{|f_{T_0}|}} - 141 + \frac{96\pi}{s^2|f_{T_0}|}\right) + 195421\left(1 - \frac{\sqrt{6\pi}}{s\sqrt{|f_{T_0}|}}\right)\right)\right\},$$

$$\text{(A.11)}$$

$$\sigma_{4W,U}^{O_{T_1}}(\mu^+\mu^- \to \bar\nu\nu W^+W^-)$$

$$= \frac{(B_{T_1}-1)^3 e^4 s^3 f_{T_1}^3}{26542080000\pi^5 s_W^4 \left(\frac{2\sqrt{2}\pi}{s\sqrt{|f_{T_1}|}}-1\right)} \left\{ 415631\left(1-\frac{2\sqrt{2}\pi}{s\sqrt{|f_{T_1}|}}\right) \right.$$

$$+600\log^2\left(\frac{(1-B_{T_1})s}{16M_W^2}\right)\left[(611-13491B_{T_1})+\frac{2\sqrt{2}\pi(14535B_{T_1}-611)}{s\sqrt{|f_{T_1}|}}\right.$$

$$\left.-120\log(1-B_{T_1})\left((119-27B_{T_1})+\frac{\sqrt{2}\pi(45B_{T_1}-238)}{s\sqrt{|f_{T_1}|}}+\frac{36\pi B_{T_1}}{s^2|f_{T_1}|}\right)-\frac{8352\pi B_{T_1}}{s^2|f_{T_1}|}\right]$$

$$+40\log\left(\frac{(1-B_{T_1})s}{16M_W^2}\right)\left[30\log(1-B_{T_1})\left((4463-783B_{T_1})+\frac{2\sqrt{2}\pi(675B_{T_1}-4463)}{s\sqrt{|f_{T_1}|}}+\frac{864\pi B_{T_1}}{s^2|f_{T_1}|}\right)\right.$$

$$\left.+351B_{T_1}\left(309-\frac{650\sqrt{2}}{s\sqrt{|f_{T_1}|}}+\frac{128\pi}{s^2|f_{T_1}|}\right)-16459\left(1-\frac{2\sqrt{2}\pi}{s\sqrt{|f_{T_1}|}}\right)\right]$$

$$-27B_{T_1}\left(42653-\frac{88250\sqrt{2}}{s\sqrt{|f_{T_1}|}}+\frac{11776\pi}{s^2|f_{T_1}|}\right)$$

$$\left.-20\log(1-B_{T_1})\left[81B_{T_1}\left(\frac{250\sqrt{2}}{s\sqrt{|f_{T_1}|}}-141+\frac{128\pi}{s^2|f_{T_1}|}\right)+85021\left(1-\frac{2\sqrt{2}\pi}{s\sqrt{|f_{T_1}|}}\right)\right]\right\},$$
$$(A.12)$$

$$\sigma_{4W,U}^{O_{T_2}}(\mu^+\mu^- \to \bar\nu\nu W^+W^-)$$

$$= \frac{(B_{T_2}-1)^3 e^4 s^3 f_{T_2}^2}{66355200000\pi^5 s_W^4}\left\{40\log\left(\frac{(1-B_{T_2})s}{16M_W^2}\right)[-83511B_{T_2}\right.$$

$$+30\log(1-B_{T_2})\left((783B_{T_2}-3583)+\frac{432\sqrt{\pi}B_{T_2}}{s\sqrt{|f_{T_2}|}}\right)+\frac{3456\sqrt{\pi}B_{T_2}}{s\sqrt{|f_{T_2}|}}+13511\Big]$$

$$-600\log^2\left(\frac{(1-B_{T_2})s}{16M_W^2}\right)\left[9B_{T_2}\left(\frac{112\sqrt{\pi}}{s\sqrt{|f_{T_2}|}}-1147\right)\right.$$

$$\left.+120\log(1-B_{T_2})\left((27B_{T_2}-97)+\frac{18\sqrt{\pi}B_{T_2}}{s\sqrt{|f_{T_2}|}}\right)+523\right]$$

$$\left.+(27B_{T_2}(64B_{T_2}+32997)-332647)-20\log(1-B_{T_2})\left(81B_{T_2}\left(141+\frac{64\sqrt{\pi}}{s\sqrt{|f_{T_2}|}}\right)-67421\right)\right\},$$
$$(A.13)$$

where

$$B_{S_{1,2}}=\left(1-\frac{2\sqrt{6}\pi}{s\sqrt{|f_{S_{1,2}}|}}\right)\theta\left(s-\frac{2\sqrt{6}\pi}{\sqrt{|f_{S_{1,2}}|}}\right), \quad B_{M_0}=\left(1-\frac{4\sqrt{2}\pi}{s\sqrt{|f_{M_0}|}}\right)\theta\left(s-\frac{4\sqrt{2}\pi}{\sqrt{|f_{M_0}|}}\right),$$

$$B_{M_1}=\left(1-\frac{8\sqrt{2}\pi}{\sqrt{|f_{M_1}|}s}\right)\theta\left(s-\frac{8\sqrt{2}\pi}{\sqrt{|f_{M_1}|}}\right), \quad B_{M_7}=\left(1-\frac{16\sqrt{2}\pi}{\sqrt{|f_{M_7}|}s}\right)\theta\left(s-\frac{16\sqrt{2}\pi}{\sqrt{|f_{M_7}|}}\right),$$

$$B_{T_0}=\left(1-\frac{\sqrt{6}\pi}{s\sqrt{|f_{T_0}|}}\right)\theta\left(s-\frac{\sqrt{6}\pi}{\sqrt{|f_{T_0}|}}\right), \quad B_{T_1}=\left(1-\frac{2\sqrt{2}\pi}{s\sqrt{|f_{T_1}|}}\right)\theta\left(s-\frac{2\sqrt{2}\pi}{\sqrt{|f_{T_1}|}}\right),$$

$$B_{T_2}=\left(1-\frac{4\sqrt{\pi}}{s\sqrt{|f_{T_2}|}}\right)\theta\left(s-\frac{4\sqrt{\pi}}{\sqrt{|f_{T_2}|}}\right).$$
$$(A.14)$$

B Helicity amplitudes relevant with the unitairty bounds

With a large $\hat{s}$, we only need to focus on helicity amplitudes growing fastest with $\hat{s}$. Denoting $\mathcal{M}(W^+_{\lambda_1} W^-_{\lambda_2} \to W^+_{\lambda_3} W^-_{\lambda_4}) = \mathcal{M}_{\lambda_1 \lambda_2 \lambda_3 \lambda_4} + \mathcal{O}(\hat{s})$, the relevant amplitudes are

$$
\begin{aligned}
\mathcal{M}_{++++} &= \frac{2f_{T_0} + f_{T_1} + f_{T_2}}{8\pi\Lambda^4} D^0_{0,0}, \\
\mathcal{M}_{++--} &= \frac{4f_{T_0} + 3f_{T_1} + f_{T_2}}{12\pi\Lambda^4} D^0_{0,0} - \frac{4f_{T_0} - 2f_{T_1} + f_{T_2}}{96\pi\Lambda^4} D^1_{0,0} + \frac{4f_{T_0} + 6f_{T_1} + f_{T_2}}{480\pi\Lambda^4} D^2_{0,0}, \\
\mathcal{M}_{++00} &= -\frac{-8f_{M_0} - 2f_{M_1} + f_{M_7}}{128\pi\Lambda^4} D^0_{0,0} - \frac{f_{M_7}}{384\pi\Lambda^4} D^1_{0,0}, \\
\mathcal{M}_{+-+-} &= \frac{2f_{T_1} + f_{T_2}}{40\pi\Lambda^4} D^2_{2,2}, \quad \mathcal{M}_{+--+} = \frac{2f_{T_0} + f_{T_1} + f_{T_2}}{40\pi\Lambda^4} D^2_{2,-2}, \\
\mathcal{M}_{+-00} &= -\frac{2f_{M_1} - f_{M_7}}{320\sqrt{6}\pi\Lambda^4} D^2_{2,0}, \quad \mathcal{M}_{+0+0} = \frac{2f_{M_1} - f_{M_7}}{192\pi\Lambda^4} D^1_{1,1}, \\
\mathcal{M}_{+0-0} &= \frac{-4f_{M_0} + f_{M_1} + \frac{f_{M_7}}{3}}{256\pi\Lambda^4} D^1_{1,-1} + \frac{4f_{M_0} - f_{M_1} + f_{M_7}}{1280\pi\Lambda^4} D^2_{1,-1}, \\
\mathcal{M}_{0000} &= \frac{f_{S_0} + 2(f_{S_1} + f_{S_2})}{24\pi\Lambda^4} D^0_{0,0} + \frac{2f_{S_0} - f_{S_1} - f_{S_2}}{96\pi\Lambda^4} D^1_{0,0} + \frac{2f_{S_0} + f_{S_1} + f_{S_2}}{480\pi\Lambda^4} D^2_{0,0},
\end{aligned}
\tag{B.1}
$$

where

$$
D^J_{m_1,m_2} = \hat{s}^2 16\pi \left(J + \frac{1}{2} \right) e^{i(m_1 - m_2)\phi} d^J_{m_1,m_2}(\theta, \phi).
\tag{B.2}
$$

The helicity amplitudes producing duplicated T_J are not shown.

References

[1] D.R. Green, P. Meade and M.-A. Pleier, *Multiboson interactions at the LHC*, *Rev. Mod. Phys.* **89** (2017) 035008 [arXiv:1610.07572] [INSPIRE].

[2] J. Chang, K. Cheung, C.-T. Lu and T.-C. Yuan, *WW scattering in the era of post-Higgs-boson discovery*, *Phys. Rev. D* **87** (2013) 093005 [arXiv:1303.6335] [INSPIRE].

[3] C.F. Anders et al., *Vector boson scattering: Recent experimental and theory developments*, *Rev. Phys.* **3** (2018) 44 [arXiv:1801.04203] [INSPIRE].

[4] C. Zhang and S.-Y. Zhou, *Positivity bounds on vector boson scattering at the LHC*, *Phys. Rev. D* **100** (2019) 095003 [arXiv:1808.00010] [INSPIRE].

[5] Q. Bi, C. Zhang and S.-Y. Zhou, *Positivity constraints on aQGC: carving out the physical parameter space*, *JHEP* **06** (2019) 137 [arXiv:1902.08977] [INSPIRE].

[6] S. Weinberg, *Baryon and Lepton Nonconserving Processes*, *Phys. Rev. Lett.* **43** (1979) 1566 [INSPIRE].

[7] B. Grzadkowski, M. Iskrzynski, M. Misiak and J. Rosiek, *Dimension-Six Terms in the Standard Model Lagrangian*, *JHEP* **10** (2010) 085 [arXiv:1008.4884] [INSPIRE].

[8] S. Willenbrock and C. Zhang, *Effective Field Theory Beyond the Standard Model*, *Ann. Rev. Nucl. Part. Sci.* **64** (2014) 83 [arXiv:1401.0470] [INSPIRE].

[9] E. Masso, *An Effective Guide to Beyond the Standard Model Physics*, *JHEP* **10** (2014) 128 [arXiv:1406.6376] [INSPIRE].

[10] O.J.P. Éboli, M.C. Gonzalez-Garcia and J.K. Mizukoshi, $pp \to jje^{\pm}\mu^{\pm}\nu\nu$ and $jje^{\pm}\mu^{\mp}\nu\nu$ at $\mathcal{O}(\alpha_{\pm}^6)$ and $\mathcal{O}(\alpha_{\pm}^4\alpha_s^2)$ for the study of the quartic electroweak gauge boson vertex at CERN LHC, *Phys. Rev. D* **74** (2006) 073005 [hep-ph/0606118] [INSPIRE].

[11] O.J.P. Éboli and M.C. Gonzalez-Garcia, *Classifying the bosonic quartic couplings*, *Phys. Rev. D* **93** (2016) 093013 [arXiv:1604.03555] [INSPIRE].

[12] ATLAS collaboration, *Evidence for Electroweak Production of $W^{\pm}W^{\pm}jj$ in pp Collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector*, *Phys. Rev. Lett.* **113** (2014) 141803 [arXiv:1405.6241] [INSPIRE].

[13] CMS collaboration, *Measurements of production cross sections of WZ and same-sign WW boson pairs in association with two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Phys. Lett. B* **809** (2020) 135710 [arXiv:2005.01173] [INSPIRE].

[14] ATLAS collaboration, *Studies of $Z\gamma$ production in association with a high-mass dijet system in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, *JHEP* **07** (2017) 107 [arXiv:1705.01966] [INSPIRE].

[15] CMS collaboration, *Measurement of the cross section for electroweak production of $Z\gamma$ in association with two jets and constraints on anomalous quartic gauge couplings in proton-proton collisions at $\sqrt{s} = 8$ TeV*, *Phys. Lett. B* **770** (2017) 380 [arXiv:1702.03025] [INSPIRE].

[16] CMS collaboration, *Measurement of the cross section for electroweak production of a Z boson, a photon and two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV and constraints on anomalous quartic couplings*, *JHEP* **06** (2020) 076 [arXiv:2002.09902] [INSPIRE].

[17] CMS collaboration, *Measurement of electroweak-induced production of $W\gamma$ with two jets in pp collisions at $\sqrt{s} = 8$ TeV and constraints on anomalous quartic gauge couplings*, *JHEP* **06** (2017) 106 [arXiv:1612.09256] [INSPIRE].

[18] CMS collaboration, *Measurement of vector boson scattering and constraints on anomalous quartic couplings from events with four leptons and two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Phys. Lett. B* **774** (2017) 682 [arXiv:1708.02812] [INSPIRE].

[19] CMS collaboration, *Measurement of differential cross sections for Z boson pair production in association with jets at $\sqrt{s} = 8$ and 13 TeV*, *Phys. Lett. B* **789** (2019) 19 [arXiv:1806.11073] [INSPIRE].

[20] ATLAS collaboration, *Observation of electroweak $W^{\pm}Z$ boson pair production in association with two jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Lett. B* **793** (2019) 469 [arXiv:1812.09740] [INSPIRE].

[21] CMS collaboration, *Measurement of electroweak WZ boson production and search for new physics in WZ + two jets events in pp collisions at $\sqrt{s} = 13$ TeV*, *Phys. Lett. B* **795** (2019) 281 [arXiv:1901.04060] [INSPIRE].

[22] CMS collaboration, *Evidence for exclusive $\gamma\gamma \to W^+W^-$ production and constraints on anomalous quartic gauge couplings in pp collisions at $\sqrt{s} = 7$ and 8 TeV, JHEP* **08** (2016) 119 [arXiv:1604.04464] [INSPIRE].

[23] CMS collaboration, *Observation of electroweak production of same-sign W boson pairs in the two jet and two same-sign lepton final state in proton-proton collisions at $\sqrt{s} = 13$ TeV, Phys. Rev. Lett.* **120** (2018) 081801 [arXiv:1709.05822] [INSPIRE].

[24] CMS collaboration, *Search for anomalous electroweak production of vector boson pairs in association with two jets in proton-proton collisions at 13 TeV, Phys. Lett. B* **798** (2019) 134985 [arXiv:1905.07445] [INSPIRE].

[25] CMS collaboration, *Observation of electroweak production of $W\gamma$ with two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV, Phys. Lett. B* **811** (2020) 135988 [arXiv:2008.10521] [INSPIRE].

[26] CMS collaboration, *Evidence for electroweak production of four charged leptons and two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV, Phys. Lett. B* **812** (2021) 135992 [arXiv:2008.07013] [INSPIRE].

[27] D. Buttazzo, D. Redigolo, F. Sala and A. Tesi, *Fusing Vectors into Scalars at High Energy Lepton Colliders, JHEP* **11** (2018) 144 [arXiv:1807.04743] [INSPIRE].

[28] J.P. Delahaye et al., *Muon Colliders*, arXiv:1901.06150 [INSPIRE].

[29] M. Lu et al., *The physics case for an electron-muon collider, Adv. High Energy Phys.* **2021** (2021) 6693618 [arXiv:2010.15144] [INSPIRE].

[30] R. Franceschini and M. Greco, *Higgs and BSM Physics at the Future Muon Collider, Symmetry* **13** (2021) 851 [arXiv:2104.05770] [INSPIRE].

[31] R. Palmer et al., *Muon collider design, Nucl. Phys. B Proc. Suppl.* **51** (1996) 61 [acc-phys/9604001].

[32] S.D. Holmes and V.D. Shiltsev, *Muon Collider*, in *Outlook for the Future*, C. Joshi, A. Caldwell, P. Muggli, S.D. Holmes and V.D. Shiltsev, eds., (Germany), pp. 816–822, Springer-Verlag Berlin Heidelberg (2013), DOI [arXiv:1202.3803] [INSPIRE].

[33] A. Costantini et al., *Vector boson fusion at multi-TeV muon colliders, JHEP* **09** (2020) 080 [arXiv:2005.10289] [INSPIRE].

[34] H. Al Ali et al., *The muon Smasher's guide, Rept. Prog. Phys.* **85** (2022) 084201 [arXiv:2103.14043] [INSPIRE].

[35] T. Han, D. Liu, I. Low and X. Wang, *Electroweak couplings of the Higgs boson at a multi-TeV muon collider, Phys. Rev. D* **103** (2021) 013002 [arXiv:2008.12204] [INSPIRE].

[36] T. Han, Z. Liu, L.-T. Wang and X. Wang, *WIMPs at High Energy Muon Colliders, Phys. Rev. D* **103** (2021) 075004 [arXiv:2009.11287] [INSPIRE].

[37] T. Han, T. Li and X. Wang, *Axion-Like Particles at High Energy Muon Colliders — A White paper for Snowmass 2021*, in *2022 Snowmass Summer Study*, (2022) [arXiv:2203.05484] [INSPIRE].

[38] J. Kalinowski, P. Kozów, S. Pokorski, J. Rosiek, M. Szleper and S. Tkaczyk, *Same-sign WW scattering at the LHC: can we discover BSM effects before discovering new states?, Eur. Phys. J. C* **78** (2018) 403 [arXiv:1802.02366] [INSPIRE].

[39] P. Kozów, L. Merlo, S. Pokorski and M. Szleper, *Same-sign WW Scattering in the HEFT: Discoverability vs. EFT Validity, JHEP* **07** (2019) 021 [arXiv:1905.03354] [INSPIRE].

[40] G. Chaudhary et al., *EFT triangles in the same-sign WW scattering process at the HL-LHC and HE-LHC*, *Eur. Phys. J. C* **80** (2020) 181 [arXiv:1906.10769] [INSPIRE].

[41] J.-C. Yang, J.-H. Chen and Y.-C. Guo, *Extract the energy scale of anomalous $\gamma\gamma \to W^+W^-$ scattering in the vector boson scattering process using artificial neural networks*, *JHEP* **09** (2021) 085 [arXiv:2107.13624] [INSPIRE].

[42] B. Abbott et al., *Anomalous quartic gauge couplings at a muon collider*, in *2022 Snowmass Summer Study*, (2022) [arXiv:2203.08135] [INSPIRE].

[43] R. Contino, A. Falkowski, F. Goertz, C. Grojean and F. Riva, *On the Validity of the Effective Field Theory Approach to SM Precision Tests*, *JHEP* **07** (2016) 144 [arXiv:1604.06444] [INSPIRE].

[44] T.D. Lee and C.-N. Yang, *Theoretical discussions on possible high-energy neutrino experiments*, *Phys. Rev. Lett.* **4** (1960) 307 [INSPIRE].

[45] M. Froissart, *Asymptotic behavior and subtractions in the Mandelstam representation*, *Phys. Rev.* **123** (1961) 1053 [INSPIRE].

[46] G. Passarino, *WW scattering and perturbative unitarity*, *Nucl. Phys. B* **343** (1990) 31 [INSPIRE].

[47] T. Corbett, O.J.P. Éboli and M.C. Gonzalez-Garcia, *Unitarity Constraints on Dimension-Six Operators*, *Phys. Rev. D* **91** (2015) 035014 [arXiv:1411.5026] [INSPIRE].

[48] A. Alboteanu, W. Kilian and J. Reuter, *Resonances and Unitarity in Weak Boson Scattering at the LHC*, *JHEP* **11** (2008) 010 [arXiv:0806.4145] [INSPIRE].

[49] J. Searcy, L. Huang, M.-A. Pleier and J. Zhu, *Determination of the WW polarization fractions in $pp \to W^{\pm}W^{\pm}jj$ using a deep machine learning technique*, *Phys. Rev. D* **93** (2016) 094033 [arXiv:1510.01691] [INSPIRE].

[50] J. Lee et al., *Polarization fraction measurement in same-sign WW scattering using deep learning*, *Phys. Rev. D* **99** (2019) 033004 [arXiv:1812.07591] [INSPIRE].

[51] J. Lee et al., *Polarization fraction measurement in ZZ scattering using deep learning*, *Phys. Rev. D* **100** (2019) 116010 [arXiv:1908.05196] [INSPIRE].

[52] K. Lasocha, E. Richter-Was, D. Tracz, Z. Was and P. Winkowska, *Machine learning classification: Case of Higgs boson CP state in $H \to \tau\tau$ decay at the LHC*, *Phys. Rev. D* **100** (2019) 113001 [arXiv:1812.08140] [INSPIRE].

[53] L. Lönnblad, C. Peterson and T. Rognvaldsson, *Using neural networks to identify jets*, *Nucl. Phys. B* **349** (1991) 675 [INSPIRE].

[54] V. Innocente, Y.F. Wang and Z.P. Zhang, *Identification of tau decays using a neural network*, *Nucl. Instrum. Meth. A* **323** (1992) 647 [INSPIRE].

[55] B. Holdom and Q.-S. Yan, *Searches for the t′ of a fourth family*, *Phys. Rev. D* **83** (2011) 114031 [arXiv:1101.3844] [INSPIRE].

[56] A. Radovic et al., *Machine learning at the energy and intensity frontiers of particle physics*, *Nature* **560** (2018) 41 [INSPIRE].

[57] P. Baldi, P. Sadowski and D. Whiteson, *Searching for Exotic Particles in High-Energy Physics with Deep Learning*, *Nature Commun.* **5** (2014) 4308 [arXiv:1402.4735] [INSPIRE].

[58] J. Ren, L. Wu, J.M. Yang and J. Zhao, *Exploring supersymmetry with machine learning*, *Nucl. Phys. B* **943** (2019) 114613 [arXiv:1708.06615] [INSPIRE].

[59] M. Abdughani, J. Ren, L. Wu and J.M. Yang, *Probing stop pair production at the LHC with graph neural networks*, *JHEP* **08** (2019) 055 [arXiv:1807.09088] [INSPIRE].

[60] R. Iten, T. Metger, H. Wilming, L. del Rio and R. Renner, *Discovering physical concepts with neural networks*, *Phys. Rev. Lett.* **124** (2020) 010508.

[61] J. Ren, L. Wu and J.M. Yang, *Unveiling CP property of top-Higgs coupling with graph neural networks at the LHC*, *Phys. Lett. B* **802** (2020) 135198 [arXiv:1901.05627] [INSPIRE].

[62] Y.-C. Guo, L. Jiang and J.-C. Yang, *Detecting anomalous quartic gauge couplings using the isolation forest machine learning algorithm*, *Phys. Rev. D* **104** (2021) 035021 [arXiv:2103.03151] [INSPIRE].

[63] J.-C. Yang, Y.-C. Guo and L.-H. Cai, *Using a nested anomaly detection machine learning algorithm to study the neutral triple gauge couplings at an e^+e^- collider*, *Nucl. Phys. B* **977** (2022) 115735 [arXiv:2111.10543] [INSPIRE].

[64] D. Espriu and F. Mescia, *Unitarity and causality constraints in composite Higgs models*, *Phys. Rev. D* **90** (2014) 015035 [arXiv:1403.7386] [INSPIRE].

[65] R.L. Delgado, A. Dobado, M.J. Herrero and J.J. Sanz-Cillero, *One-loop $\gamma\gamma \to W_L^+ W_L^-$ and $\gamma\gamma \to Z_L Z_L$ from the Electroweak Chiral Lagrangian with a light Higgs-like scalar*, *JHEP* **07** (2014) 149 [arXiv:1404.2866] [INSPIRE].

[66] S. Fichet and G. von Gersdorff, *Anomalous gauge couplings from composite Higgs and warped extra dimensions*, *JHEP* **03** (2014) 102 [arXiv:1311.6815] [INSPIRE].

[67] T.D. Lee, *A Theory of Spontaneous T Violation*, *Phys. Rev. D* **8** (1973) 1226 [INSPIRE].

[68] J.-C. Yang and M.-Z. Yang, *Effect of the Charged Higgs Bosons in the Radiative Leptonic Decays of B^- and D^- Mesons*, *Mod. Phys. Lett. A* **31** (2015) 1650012 [arXiv:1508.00314] [INSPIRE].

[69] X.-G. He, G.C. Joshi, H. Lew and R.R. Volkas, *Simplest Z-prime model*, *Phys. Rev. D* **44** (1991) 2118 [INSPIRE].

[70] J.-X. Hou and C.-X. Yue, *The signatures of the new particles h_2 and $Z_{\mu\tau}$ at e-p colliders in the $U(1)_{L_\mu - L_\tau}$ model*, *Eur. Phys. J. C* **79** (2019) 983 [arXiv:1905.00627] [INSPIRE].

[71] K. Mimasu and V. Sanz, *ALPs at Colliders*, *JHEP* **06** (2015) 173 [arXiv:1409.4792] [INSPIRE].

[72] C.-X. Yue, M.-Z. Liu and Y.-C. Guo, *Searching for axionlike particles at future ep colliders*, *Phys. Rev. D* **100** (2019) 015020 [arXiv:1904.10657] [INSPIRE].

[73] C.-X. Yue, X.-J. Cheng and J.-C. Yang, *The charged-current non-standard neutrino interactions at the LHC and HL-LHC*, arXiv:2110.01204 [INSPIRE].

[74] 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 [*Erratum ibid.* **09** (2019) 019] [arXiv:1512.03433] [INSPIRE].

[75] T. Han, H.-J. He and C.P. Yuan, *Quartic gauge boson couplings at linear colliders: Interplay of WWZ/ZZZ production and WW fusion*, *Phys. Lett. B* **422** (1998) 294 [hep-ph/9711429] [INSPIRE].

[76] E. Boos, H.J. He, W. Kilian, A. Pukhov, C.P. Yuan and P.M. Zerwas, *Strongly interacting vector bosons at TeV e+- e- linear colliders: Addendum*, *Phys. Rev. D* **61** (2000) 077901 [hep-ph/9908409] [INSPIRE].

[77] G. Perez, M. Sekulla and D. Zeppenfeld, *Anomalous quartic gauge couplings and unitarization for the vector boson scattering process $pp \to W^+W^+jjX \to \ell^+\nu_\ell\ell^+\nu_\ell jjX$, Eur. Phys. J. C* **78** (2018) 759 [arXiv:1807.02707] [INSPIRE].

[78] Y.-C. Guo, Y.-Y. Wang, J.-C. Yang and C.-X. Yue, *Constraints on anomalous quartic gauge couplings via $W\gamma jj$ production at the LHC, Chin. Phys. C* **44** (2020) 123105 [arXiv:2002.03326] [INSPIRE].

[79] Y.-C. Guo, Y.-Y. Wang and J.-C. Yang, *Constraints on anomalous quartic gauge couplings by $\gamma\gamma \to W^+W^-$ scattering, Nucl. Phys. B* **961** (2020) 115222 [arXiv:1912.10686] [INSPIRE].

[80] J.-C. Yang, Y.-C. Guo, C.-X. Yue and Q. Fu, *Constraints on anomalous quartic gauge couplings via $Z\gamma jj$ production at the LHC, Phys. Rev. D* **104** (2021) 035015 [arXiv:2107.01123] [INSPIRE].

[81] M. Born and L. Infeld, *Foundations of the new field theory, Proc. Roy. Soc. Lond. A* **144** (1934) 425.

[82] J. Ellis and S.-F. Ge, *Constraining Gluonic Quartic Gauge Coupling Operators with $gg \to \gamma\gamma$, Phys. Rev. Lett.* **121** (2018) 041801 [arXiv:1802.02416] [INSPIRE].

[83] J. Ellis, N.E. Mavromatos and T. You, *Light-by-Light Scattering Constraint on Born-Infeld Theory, Phys. Rev. Lett.* **118** (2017) 261802 [arXiv:1703.08450] [INSPIRE].

[84] C. Degrande, *A basis of dimension-eight operators for anomalous neutral triple gauge boson interactions, JHEP* **02** (2014) 101 [arXiv:1308.6323] [INSPIRE].

[85] J. Ellis, S.-F. Ge, H.-J. He and R.-Q. Xiao, *Probing the scale of new physics in the $ZZ\gamma$ coupling at e^+e^- colliders, Chin. Phys. C* **44** (2020) 063106 [arXiv:1902.06631] [INSPIRE].

[86] J. Ellis, H.-J. He and R.-Q. Xiao, *Probing new physics in dimension-8 neutral gauge couplings at e^+e^- colliders, Sci. China Phys. Mech. Astron.* **64** (2021) 221062 [arXiv:2008.04298] [INSPIRE].

[87] G.J. Gounaris, J. Layssac and F.M. Renard, *Off-shell structure of the anomalous Z and γ selfcouplings, Phys. Rev. D* **62** (2000) 073012 [hep-ph/0005269] [INSPIRE].

[88] G.J. Gounaris, J. Layssac and F.M. Renard, *Signatures of the anomalous Z_γ and ZZ production at the lepton and hadron colliders, Phys. Rev. D* **61** (2000) 073013 [hep-ph/9910395] [INSPIRE].

[89] A. Senol et al., *Probing the Effects of Dimension-eight Operators Describing Anomalous Neutral Triple Gauge Boson Interactions at FCC-hh, Nucl. Phys. B* **935** (2018) 365 [arXiv:1805.03475] [INSPIRE].

[90] Q. Fu, J.-C. Yang, C.-X. Yue and Y.-C. Guo, *The study of neutral triple gauge couplings in the process $e^+e^- \to Z\gamma$ including unitarity bounds, Nucl. Phys. B* **972** (2021) 115543 [arXiv:2102.03623] [INSPIRE].

[91] J. Ellis, H.-J. He and R.-Q. Xiao, *Probing Neutral Triple Gauge Couplings at the LHC and Future Hadron Colliders*, arXiv:2206.11676 [INSPIRE].

[92] J.-C. Yang, Z.-B. Qing, X.-Y. Han, Y.-C. Guo and T. Li, *Tri-photon at muon collider: a new process to probe the anomalous quartic gauge couplings, JHEP* **22** (2020) 053 [arXiv:2204.08195] [INSPIRE].

[93] J. Alwall 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 [arXiv:1405.0301] [INSPIRE].

[94] N.D. Christensen and C. Duhr, *FeynRules — Feynman rules made easy*, *Comput. Phys. Commun.* **180** (2009) 1614 [`arXiv:0806.4194`] [InSPIRE].

[95] G.L. Kane, W.W. Repko and W.B. Rolnick, *The Effective W+-, Z0 Approximation for High-Energy Collisions*, *Phys. Lett. B* **148** (1984) 367 [InSPIRE].

[96] E. Boos, H.J. He, W. Kilian, A. Pukhov, C.P. Yuan and P.M. Zerwas, *Strongly interacting vector bosons at TeV e^+e^- linear colliders*, *Phys. Rev. D* **57** (1998) 1553 [`hep-ph/9708310`] [InSPIRE].

[97] R. Ruiz, A. Costantini, F. Maltoni and O. Mattelaer, *The Effective Vector Boson Approximation in high-energy muon collisions*, *JHEP* **06** (2022) 114 [`arXiv:2111.02442`] [InSPIRE].

[98] Y. LeCun, Y. Bengio and G. Hinton, *Deep learning*, *Nature* **521** (2015) 436.

[99] K. He, X. Zhang, S. Ren and J. Sun, *Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification*, `arXiv:1502.01852` [InSPIRE].

[100] M. Abadi et al., *TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems*, `arXiv:1603.04467` [InSPIRE].

[101] J. Layssac, F.M. Renard and G.J. Gounaris, *Unitarity constraints for transverse gauge bosons at LEP and supercolliders*, *Phys. Lett. B* **332** (1994) 146 [`hep-ph/9311370`] [InSPIRE].

[102] T. Corbett, O.J.P. Éboli and M.C. Gonzalez-Garcia, *Unitarity Constraints on Dimension-six Operators II: Including Fermionic Operators*, *Phys. Rev. D* **96** (2017) 035006 [`arXiv:1705.09294`] [InSPIRE].

[103] R. Gomez-Ambrosio, *Vector Boson Scattering Studies in CMS: The $pp \to ZZjj$ Channel*, *Acta Phys. Polon. Supp.* **11** (2018) 239 [`arXiv:1807.09634`] [InSPIRE].

[104] E.d.S. Almeida, O.J.P. Éboli and M.C. Gonzalez-Garcia, *Unitarity constraints on anomalous quartic couplings*, *Phys. Rev. D* **101** (2020) 113003 [`arXiv:2004.05174`] [InSPIRE].

[105] W. Kilian, S. Sun, Q.-S. Yan, X. Zhao and Z. Zhao, *Multi-Higgs boson production and unitarity in vector-boson fusion at future hadron colliders*, *Phys. Rev. D* **101** (2020) 076012 [`arXiv:1808.05534`] [InSPIRE].

[106] W. Kilian, S. Sun, Q.-S. Yan, X. Zhao and Z. Zhao, *Highly Boosted Higgs Bosons and Unitarity in Vector-Boson Fusion at Future Hadron Colliders*, *JHEP* **05** (2021) 198 [`arXiv:2101.12537`] [InSPIRE].

[107] M. Jacob and G.C. Wick, *On the General Theory of Collisions for Particles with Spin*, *Annals Phys.* **7** (1959) 404 [InSPIRE].

[108] D. Barducci et al., *Interpreting top-quark LHC measurements in the standard-model effective field theory*, `arXiv:1802.07237` [InSPIRE].

[109] D. Racco, A. Wulzer and F. Zwirner, *Robust collider limits on heavy-mediator Dark Matter*, *JHEP* **05** (2015) 009 [`arXiv:1502.04701`] [InSPIRE].

[110] C. Garcia-Garcia, M. Herrero and R.A. Morales, *Unitarization effects in EFT predictions of WZ scattering at the LHC*, *Phys. Rev. D* **100** (2019) 096003 [`arXiv:1907.06668`] [InSPIRE].

[111] DELPHES 3 collaboration, *DELPHES 3, A modular framework for fast simulation of a generic collider experiment*, *JHEP* **02** (2014) 057 [`arXiv:1307.6346`] [InSPIRE].

[112] PARTICLE DATA GROUP collaboration, *Review of Particle Physics*, *PTEP* **2020** (2020) 083C01 [InSPIRE].