Vector boson scattering: Recent experimental and theory developments


a University of Antwerp, Antwerpen, Belgium
b Physics Department, Brookhaven National Laboratory, Upton, NY, USA
c CEA/Saclay-IRFU, France
d Deutsches Elektronen-Synchrotron, Hamburg, Germany
e University College Dublin, Ireland
f Institute for Particle Physics Phenomenology, Department of Physics, University of Durham, UK

⁎ Corresponding author.
E-mail address: Pietro.Govoni@mib.infn.it (G. Pietro).

https://doi.org/10.1016/j.revip.2018.11.001

Available online 27 November 2018
2405-4283/ © 2018 Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/BY-NC-ND/4.0/).
ARTICLEINFO

Keywords:
LHC  
VBS  
HEP  
EWSB (Electroweak symmetry breaking)

ABSTRACT

This document summarises the talks and discussions happened during the VBSCan Split17 workshop, the first general meeting of the VBSCan COST Action network. This collaboration is aiming at a consistent and coordinated study of vector-boson scattering from the phenomenological and experimental point of view, for the best exploitation of the data that will be delivered by existing and future particle colliders.

1. Introduction

In the past years, the study of the vector-boson scattering (VBS) attracted lots of interest in the theory and experimental communities (see e.g. References [1,2] for recent reviews). If the Standard Model (SM) is a partial description of Nature and its completion happens at energies which are experimentally unreachable, the cross section of VBS processes could increase substantially between the Higgs boson mass and the scale at which new physics mechanisms intervene to restore the unitarity of the process. The study of VBS offers then the unique opportunity to seize this scenario of “delayed unitarity cancellations” even if the energy scale at which the processes beyond the SM (BSM) enter into play goes beyond our experimental reach, either today or in the next future. This measurement, though, presents several challenges, both on the theory side and on the experimental one. For example difficulties arise, at hadron colliders, because of the complexity of the final state already at tree level, where six fermions arise from the interaction vertex of the initial state partons. Four of them are expected to be the product of vector-boson disintegrations, while two to be remnants of the colliding protons. In this complex environment, the exact calculation of the process needs to be kept as reference to validate any approximations performed to understand the main features of the process. Effective field theory parameterisations of BSM theories have to cope with the multiplex final state, and higher order corrections are not easily calculated either. Experimentally, several
processes give rise to signatures similar to the VBS ones, and the events typically span a large angle in rapidity, involving the entire particle detectors in the measurement. Therefore, algorithms need to be optimised to reject backgrounds at best, involving all the sub-elements of each measuring apparatus, and featuring the most advanced data analysis techniques.

Only a coherent action in the experimental and theoretical community will grant that all these challenges will be met and that the data delivered by current and future particle colliders will be exploited at best. The VBSCan COST Action is a four-year project, funded by the Horizon 2020 Framework Programme of the European Union, aiming at a consistent and coordinated study of VBS from the phenomenological and experimental points of view, gathering all the interested parties in the high-energy physics community, together with experts of data mining techniques.

The community is organised in five working groups, three of which are focussing on the scientific aspects of the collaboration. One is dedicated to the theoretical understanding of the VBS process (WG1), which targets a detailed description of the signal and relative backgrounds in the SM, as well as effective field theory (EFT) modelling of BSM effects. A second one focuses on analysis techniques (WG2), studying the definition of data analysis protocols and performances to maximise the significance of the VBS analyses at hadron colliders, promoting the communication between theory and experiments. A third one fosters the optimal deployment of the studies in the hadron collider experiments data analyses (WG3). The remaining two working groups address the knowledge exchange and cross-activities (WG4), and the implementation of the COST inclusiveness policies (WG5), respectively.

The Network is composed of theoretical and experimental physicists from both the ATLAS and CMS experiments, as well as data analysis experts and industrial partners. The first general meeting marked the start of the activities. It happened at the end of June 2017 in Split [3] and was dedicated to reviews of the data analysis status of the art, as well as of the theoretical and experimental instruments relevant for VBS studies. This report contains a summary of the talks presented1 divided into sections corresponding to the Network working groups. A review of the theoretical understanding and the parameterisation of the impact of new physics in the VBS domain through effective field theory expansion is given in Section 1; an outline of the existing experimental results at the time of the Split workshop, and the overview of future perspectives are presented in Section 2; we conclude in Section 3 with an overview of the existing techniques for the identification of jets in the ATLAS and CMS experiments, which are the most complicated physics objects to be dealt with in VBS and, therefore, the main focus of WG3.

2. WG1: Theoretical understanding

2.1. Complete NLO corrections to $W^+W^-$ scattering2

The first NLO process that has been observed during the run I of the LHC is the same-sign WW production [4–6]. This observation has already been confirmed by a measurement of the CMS collaboration at the 13 TeV run II [7]. In view of the growing mole of data which will be collected by the experiments, and of the consequent reduction of the uncertainties affecting these measurements, precise theoretical predictions become necessary.

In that respect next-to-leading order (NLO) QCD and electroweak (EW) corrections to such signatures should be computed. So far, NLO computations have focused on NLO QCD corrections to the VBS process [1,8–10] and its QCD-induced irreducible background process [1,11–14]. No NLO EW corrections had been computed and the NLO QCD computations relied on the so-called VBS approximation. In Reference [15], for the first time, all leading order (LO) and NLO contributions to the full pp → $\mu^+\mu^-e^+\nu$ process have been reported.3 As the full amplitudes are used, this amounts to computing three LO contributions and four NLO contributions. At LO, the three contributions are the EW process [order $O(\alpha)$], its QCD-induced counterpart [order $O(\alpha_s^2\alpha)$] as well as the interference [order $O(\alpha\alpha_s)$]. Due to the VBS event selection applied to the final state, the full process is dominated by the purely EW contribution (see Table 1). This EW contribution features the proper VBS diagrams but also background diagrams where, for example, the W bosons are simply radiated off the quark lines.

At NLO, the four contributions arise at the orders $O(\alpha^2)$, $O(\alpha_s\alpha)$, $O(\alpha_s^2\alpha^2)$, and $O(\alpha_s^3\alpha^3)$. An interesting feature is that the orders $O(\alpha\alpha_s)$ and $O(\alpha_s^2\alpha^3)$ receive both EW and QCD corrections. Thus, at NLO (as opposed to LO) it is not possible to strictly distinguish the EW process from the QCD-induced one. As it can be seen from Table 2, at the level of the fiducial cross section, the largest contributions are the ones of order $O(\alpha^2)$. These are the NLO EW corrections to the EW processes.

This is also reflected in two differential distributions in the transverse momentum for the hardest jet and invariant mass for the two leading jets in Fig. 1. At LO, the QCD-induced process as well as the interference are rather suppressed due to the typical VBS event selection (as for the fiducial cross section). This is exemplified in the invariant mass of the two leading jets where, at high invariant mass, the QCD-induced background becomes negligible. At NLO, the bulk of the corrections originates from the large EW corrections to the EW process at order $O(\alpha^2)$. In particular, they display the typical behaviour of Sudakov logarithms that grow large in the high-energy limit. Note that the contribution from initial-state photon is also shown but is not taken into account in the definition of the NLO predictions. This contribution is rather small and relatively constant in shape over the whole range studied here. Finally, as at NLO it is not possible to isolate the EW production from its irreducible backgrounds, a global measurement of the full process $pp \rightarrow \mu^+\mu^-e^+\nu$ with all components is desirable.

We conclude by stressing that usually EW corrections are particularly sizeable only in phase space regions which are dominated

---

1 All the presentations can be found at https://indic0.cern.ch/event/629638/.
2 Speaker: M. Pellen.
3 The $O(\alpha^2)$ corrections were computed previously in Reference [16].
by large scales and hence are suppressed. Therefore the impact at the level of the total fiducial cross section is usually rather limited. This is not the case here where the corrections are already large at the level of the cross section and reach $\sim 17\%$ [16]. The origin of these large EW corrections are virtual corrections and in particular the ones corresponding to the insertion of massive vector particles in the scattering process [16]. Hence, large NLO EW corrections are an intrinsic feature of VBS at the LHC. As the EW corrections are particularly large, it might be possible to measure them at a high luminosity LHC, hence probing the EW sector of the SM to very high precision. This is illustrated on the right-hand side of Fig. 2 where the band represents the estimated statistical error for a high-luminosity LHC collecting $3000$ fb$^{-1}$.

2.2. Monte Carlo comparisons for $W^+W^-$ scattering

This talk presents some preliminary results of a study which has appeared during the publication of these proceedings [17]. In the

---

**Table 1**

Fiducial cross section from Reference [15] at $\sqrt{s} = 13$ TeV at LO for the process $pp \to \mu^+\mu^-e^+e^-$, at orders $O(\alpha^2)$, $O(\alpha,\alpha^2)$, and $O(\alpha^2,\alpha^4)$. The sum of all the LO contributions is in the last column and all contributions are expressed in femtobarn. The statistical uncertainty from the Monte Carlo integration on the last digit is given in parenthesis.

<table>
<thead>
<tr>
<th>Order</th>
<th>$O(\alpha^2)$</th>
<th>$O(\alpha,\alpha^2)$</th>
<th>$O(\alpha^2,\alpha^4)$</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma$ [fb]</td>
<td>1.4178(2)</td>
<td>0.04815(2)</td>
<td>0.17229(5)</td>
<td>1.6383(2)</td>
</tr>
</tbody>
</table>

**Table 2**

NLO corrections from Reference [15] for the process $pp \to \mu^+\mu^-e^+e^-$ at the orders $O(\alpha^2)$, $O(\alpha,\alpha^2)$, $O(\alpha^2,\alpha^4)$, and $O(\alpha^4)$. The sum of all the NLO contributions is in the last column. The contribution $\delta\sigma_{NLO}$ corresponds to the absolute correction while $\delta\sigma_{NLO}/\sigma_{LO}$ gives the relative correction normalised to the sum of all LO contributions. The absolute contributions are expressed in femtobarn while the relative ones are expressed in per cent. The statistical uncertainty from the Monte Carlo integration on the last digit is given in parenthesis.

<table>
<thead>
<tr>
<th>Order</th>
<th>$O(\alpha^2)$</th>
<th>$O(\alpha,\alpha^2)$</th>
<th>$O(\alpha^2,\alpha^4)$</th>
<th>$O(\alpha^4)$</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta\sigma_{NLO}$ [fb]</td>
<td>$-21.69(3)$</td>
<td>$-0.0568(5)$</td>
<td>$0.00032(13)$</td>
<td>$-0.0063(4)$</td>
<td>$-0.2804(7)$</td>
</tr>
<tr>
<td>$\delta\sigma_{NLO}/\sigma_{LO}$ [%]</td>
<td>$-13.2$</td>
<td>$3.5$</td>
<td>$0$</td>
<td>$-0.4$</td>
<td>$17.1$</td>
</tr>
</tbody>
</table>

---

4 Speaker: M. Zaro.
last decade many codes capable of performing VBS simulations have appeared. Within a network such as VBSCan it is therefore natural to perform a quantitative comparison of these programs, both to cross-validate results and to assess the impact of the different approximations which are used. In fact, already at LO, when considering the process \( pp \to \mu^+ \mu^- e^+ e^- jj \) at order \( O(\alpha^3) \), the various implementations of VBS simulations are different. They differ, for example, by the (non-)inclusion of diagrams with vector bosons in the \( s \)-channel or by the treatment of interferences between diagrams. The reason of these differences is that, when typical signal cuts for VBS are imposed, these effects turn out to be small on rates and distributions.

In the comparison, the following codes are used:

- The program BONSAY [18] (contact person: C. Schwan) consists of a general-purpose Monte Carlo integrator and matrix elements taken from several sources: Born matrix elements are adapted from the program LUSIFER [19] for the partonic processes, real matrix elements are written by Marina Billoni, and virtual matrix elements by Stefan Dittmaier. One loop integrals are evaluated using the COLLIER library [20,21].

- MADGRAPH5_AMC@NLO [22] (contact person: M. Zaro) is a meta-code, automatically generating the source code to simulate any scattering process including NLO QCD corrections both at fixed order and including matching to parton showers. It makes use of the FKS subtraction method [23,24] (automated in the module MADFKS [25,26]) for regulating IR singularities. The computations of one-loop amplitudes are carried out by switching dynamically between two integral-reduction techniques, OPP [27] or Laurent-series expansion [28], and TIR [29–31]. These have been automated in the module MADLOOP [32], which in turn exploits CUTTOOLS [33], NINJA [34,35], or IREGI [36], together with an in-house implementation of the OPENLOOPS optimisation [37]. The simulation of VBS at NLO-QCD accuracy can be performed by issuing the following commands in the program interface:
  > set complex_mass_scheme #1
  > import model loop_qcd_qed_sm_Gmu #2
  > generate p p > e+ ve mu+ vm jj QCD=0 [QCD] #3
  > output #4

With these commands the complex-mass scheme is turned on #1, then the NLO-capable model is loaded #2
3, finally the process code is generated #3 (note the QCD=0 syntax to select the purely-electroweak process) and written to disk #4. Because of some internal limitations, which will be lifted in the future version capable of computing both QCD and EW corrections, only loops with QCD-interacting particles are generated.

- VBFNLO [14,38,39] (contact person: M. Rauch) is a flexible parton-level Monte Carlo for processes with electroweak bosons. It allows the calculation of VBS processes at NLO QCD in the VBF approximation also including the \( s \)-channel tri-boson contribution, neglecting interferences between the two. Besides the SM, also anomalous couplings of the Higgs and gauge bosons can be simulated.

- The POWHEG-BOX [40,41] (contact person: A. Karlberg) is a framework for matching NLO-QCD calculations with parton showers. It

\(^5\) Despite the loop_qcd_qed_sm_Gmu model also includes NLO counterterms for computing electroweak corrections, it is not yet possible to compute such corrections with the current version of the code.
relies on the user providing the matrix elements and Born phase space, but will automatically construct FKS subtraction terms and the phase space for the real emission. For the VBS processes all matrix elements are being provided by a previous version of VBFNLO \([14,38,39]\) and hence the approximations used in the POWHEG-BOX are similar to those used in VBFNLO.

- The program \texttt{RECOLA+MOCANO} (contact person: M. Pellen) is composed of a flexible Monte Carlo program dubbed \texttt{MOCANO} \([42]\) and the general matrix element generator \texttt{RECOLA} \([43,44]\). To numerically evaluate the one-loop scalar and tensor integrals, \texttt{RECOLA} relies on the \texttt{COLLIER} library \([20,21]\). These tools have been successfully used for the computation of the full NLO corrections for VBS \([15,16]\).

- \texttt{WHIZARD} \([45,46]\) (contact person: V. Rothe) is a multi-purpose event generator with the LO matrix element generator \texttt{OMEGA}. It provides FKS subtraction terms for any NLO process, while virtual matrix elements are provided externally by \texttt{OPENLOOPS} \([37]\) (alternatively, \texttt{RECOLA} \([43,44]\) can be used as well). \texttt{WHIZARD} allows to simulate a huge number of BSM models as well, in particular for new physics in the VBS channel in terms of both higher-dimensional operators as well as explicit resonances.

The complete comparison of the codes will be published in a separate work. Here, we present some preliminary results obtained at LO \(O(\alpha^2)\) and including NLO QCD corrections at fixed-order \(O(\alpha^2 \alpha_s)\), for the process \(pp \rightarrow \mu^+ \nu e^+ \nu jj\). In Table 3 the details of the various codes are reported. In particular, it is specified whether:

- all \(s\)- and \(t/u\)-channel diagrams that lead to the considered final state are included;
- interferences between diagrams are included at LO;
- diagrams which do not feature two resonant vector bosons are included;
- the so-called non-factorisable (NF) QCD corrections, that is the corrections where (real or virtual) gluons are exchanged between different quark lines, are included;
- EW corrections to the \(O(\alpha^2 \alpha_s)\) interference are included. These corrections are of the same order as the NLO QCD corrections to the \(O(\alpha^2)\) term.

We simulate VBS production at the LHC, with a centre-of-mass energy \(\sqrt{s} = 13\) TeV. We assume five massless flavours in the proton, and employ the NNPDF 3.0 parton density \([47]\) with NLO QCD evolution (the \texttt{hhad} in LHAPDF6 \([48]\) for this set is 260,000) and strong coupling constant \(\alpha_s(M_Z) = 0.118\). Since the employed PDF set has no photonic density, photon-induced processes are not considered. Initial-state collinear singularities are factorised with the \(\overline{\text{MS}}\) scheme, consistently with what is done in NNPDF.

We use the following values for the mass and width of the massive particles:

\[
\begin{align*}
    m_t &= 173.21\ \text{GeV}, & \Gamma_t &= 0\ \text{GeV}, \\
    M_Z^{\text{OS}} &= 91.1876\ \text{GeV}, & \Gamma_Z^{\text{OS}} &= 2.4952\ \text{GeV}, \\
    M_W^{\text{OS}} &= 80.385\ \text{GeV}, & \Gamma_W^{\text{OS}} &= 2.085\ \text{GeV}, \\
    M_H &= 125.0\ \text{GeV}, & \Gamma_H &= 4.07 \times 10^{-3}\ \text{GeV}.
\end{align*}
\]

The pole masses and widths of the W and Z bosons are obtained from the measured on-shell (OS) values \([49]\) according to

\[
\begin{align*}
    M_V &= \frac{M_V^{\text{OS}}}{\sqrt{1 + (\Gamma_V^{\text{OS}}/M_V)^2}}, & \Gamma_V &= \frac{\Gamma_V^{\text{OS}}}{\sqrt{1 + (\Gamma_V^{\text{OS}}/M_V)^2}}.
\end{align*}
\]

The EW coupling is renormalised in the \(G_\mu\) scheme \([50]\) where

\[
G_\mu = 1.16637 \times 10^{-5}\ \text{GeV}^{-2}.
\]

The derived value of the EW coupling \(\alpha\), corresponding to our choice of input parameters, is

\[
\alpha = 7.555310522369 \times 10^{-3}.
\]

We employ the complex-mass scheme \([51,52]\) to treat unstable intermediate particles in a gauge-invariant manner.

Cross sections and distributions are computed within the following VBS cuts inspired by experimental measurements \([4–7]\):
The two same-sign charged leptons are required to have
\[ p_{T,\ell} > 20 \text{GeV}, \quad |y_\ell| < 2.5, \quad \Delta R_{\ell\ell} > 0.3. \]  
(5)

The total missing transverse energy, computed from the vectorial sum of the transverse momenta of the two neutrinos in the event, is required to be
\[ E_{T,\text{miss}} = p_{T,\text{miss}} > 40 \text{ GeV}. \]  
(6)

QCD partons (quarks and gluons) are clustered together using the anti-\(k_T\) algorithm \cite{53} with distance parameter \( R = 0.4 \). Jets are required to have
\[ p_{T,j} > 30 \text{ GeV}, \quad |y_j| < 4.5, \quad \Delta R_{j\ell} > 0.3. \]  
(7)

On the two jets with largest transverse-momentum the following invariant-mass and rapidity-separation cuts are imposed
\[ m_{jj} > 500 \text{ GeV}, \quad |\Delta y_{jj}| > 2.5. \]  
(8)

When EW corrections are computed, real photons and charged fermions are clustered together using the anti-\(k_T\) algorithm with radius parameter \( R = 0.1 \). In this case, leptons and quarks mentioned above must be understood as dressed fermions.

Table 4

<table>
<thead>
<tr>
<th>Code</th>
<th>σ [fb]</th>
</tr>
</thead>
<tbody>
<tr>
<td>BONSAY</td>
<td>1.5524 ± 0.0002</td>
</tr>
<tr>
<td>MG5_AMC</td>
<td>1.547 ± 0.001</td>
</tr>
<tr>
<td>FORTHEG</td>
<td>1.5573 ± 0.0003</td>
</tr>
<tr>
<td>RECODA + MOCANLO</td>
<td>1.5503 ± 0.0003</td>
</tr>
<tr>
<td>VBFNLO</td>
<td>1.5538 ± 0.0002</td>
</tr>
<tr>
<td>WHIZARD</td>
<td>1.5539 ± 0.0004</td>
</tr>
</tbody>
</table>

In Table 4 we report the total rates at LO accuracy obtained with the set-up described above, and in Fig. 3 we show the results for the tagging jets (left) and lepton-pair (right) invariant-mass distributions. In both cases we show the absolute distributions in the main frame of the figures, while in the inset the ratio over VBFNLO is displayed. For both observables we find a relatively good agreement among the various tools, which confirms the fact that contributions from s-channel diagrams as well as from non-resonant configurations are strongly suppressed in the fiducial region. The same level of agreement is found for all other differential observables. At NLO, rates show slightly larger discrepancies, as it can be observed in Table 5. This is most likely due to low dijet invariant-mass configurations, where s-channel diagrams and interferences are less suppressed than at LO, because of the presence of extra QCD radiation.

We conclude this section by recalling that the results presented must be regarded as preliminary. In the coming months, this work will be enlarged to include comparison of predictions at NLO QCD matched to parton shower or with EW corrections, as well as to study the effect of changing the VBS cuts. The QCD-induced background will also be studied.

2.3. Polarisation of vector bosons\(^6\)

Processes related to new physics could disturb the delicate balance which preserves unitarity in VBS between longitudinally polarised vector bosons, and lead to potentially large enhancements of the VBS rate, making it the ideal process for searches of deviations from the SM and hints of new physics. Developing methods which allow the separation of the different vector polarisations is, therefore, of primary relevance. A new technique has been proposed and applied to the scattering of two W bosons of opposite charge \cite{54}: the investigated process is \( pp \rightarrow jje^+\mu^-\nu
\nu \) at the LHC@13TeV, after VBS-like kinematic cuts.

The underlying formalism was established long time ago (see e.g. References \cite{55,56}). The polarisation tensor in the W propagator can be expressed in terms of polarisation vectors:

\[ -g^{\mu\nu} + \frac{k^\mu k^\nu}{M^2} = \sum_{\lambda=1}^4 v_\lambda^\mu(k) v_\lambda^\nu(k). \]  
(9)

\(^6\) Speaker: E. Maina.
The decay amplitudes of the W depend on its polarisation. In the rest frame of the ℓν pair, they are:

\[ M^{0} = \pm i g E \sin \theta, \quad M^{R/L} = \pm i g E (1 \pm \cos \theta) e^{i \phi}, \]

where 0, R, L refer to the longitudinal, right, and left polarisations and (θ, φ) are the charged lepton angles relative to the boson direction. Hence, each physical polarisation is uniquely associated with a specific angular distribution of the charged lepton.

Two issues, however, remain unresolved:

- With few exceptions, electroweak boson production processes are described by amplitudes including nonresonant diagrams, which cannot be interpreted as production times decay of any vector boson. These diagrams are essential for gauge invariance and cannot be ignored. For them, separating polarisations is simply unfeasible.
- Since the W's are unstable particles, the decays of the individual polarisations interfere among themselves. In order to define amplitudes with definite W polarisation, it is necessary to devise an accurate approximation to the full result that only involves double resonant diagrams. Reference [54] employed an on-shell projection (OSP) method, similar to the procedure employed for the calculation of EW radiative corrections to W⁺W⁻ production in Reference [50].

It consists in substituting the momentum of the ℓν pair with a momentum on the W mass shell, while the denominator in the W propagator is left untouched. However, this projection is not uniquely defined. In order to have an unambiguous prescription one can choose to conserve: the total four–momentum of the WW system (thus, also M_{WW} is conserved); the direction of the two W bosons in the WW centre of mass frame; the angles of each charged lepton, in the corresponding W centre of mass frame, relative to the boson direction in the laboratory. This procedure is gauge invariant.

If, for instance, one considers a polarised W⁺ boson and a non-polarised W⁻ one, once all non double resonant diagrams have been dropped and the resonant ones have been projected, the squared amplitude becomes:

\[ |M|^2 = \sum_{j,k} |M_{jj}\|^2 + \sum_{k\neq j} M_{jk}^* M_{kj}, \]

where the sum runs over all resonant diagrams with coherent and incoherent contributions, and \( |M_{jj}\|^2 \) represent the squared amplitudes of resonant diagrams with identical final states. The interference terms \( M_{jk}^* M_{kj} \) represent the contributions of resonant diagrams with different final states.
where $\lambda$ is the $W^-$ polarisation. In the absence of cuts on the final state leptons, the interference terms in Eq. (11) cancel upon integration and the projected cross section is simply the sum of singly polarised cross sections. In the $W$ centre of mass frame the charged lepton angular distribution is:

$$
\frac{1}{\sigma(X)} \frac{d\sigma(\ell, X)}{d\cos \theta} = \frac{3}{8} (1 + \cos \theta)^2 f_L + \frac{3}{8} (1 - \cos \theta)^2 f_R + \frac{3}{4} \sin^2 \theta f_0.
$$

(12)

The coherent sum of polarised amplitudes, obtained via direct computation (OSP method), differs by about 1% from the exact cross section. The differential distributions which do not depend on decay products of the $W$ bosons are equally well described. Other variables, like the transverse momentum or the angle $\phi$ of the electron, show sizeable differences between the exact distributions and the projected ones. The polarisation fractions obtained expanding the full result on the first three Legendre polynomials and through direct computation of singly polarised processes agree.

The approach proposed in Reference [54] can be applied also when standard acceptance cuts ($p_T^{\ell} > 20$ GeV and $|\eta^{\ell}| < 2.5$) are imposed on both charged leptons. In this case, the coherent sum of OSP polarised amplitudes differs from the full result by about 2%. The interference among polarisations is generally small. The distributions obtained from the incoherent sum of three OSP distributions agree well with the full result for variables which do not depend on the $W$ decay products. The individual polarisations are not affected equally by the cuts. Typically, the cross section for right–handed $W$'s is reduced the most, followed by the cross section for longitudinally polarised $W$'s. Left–handed $W$ bosons seem to be the least sensitive to acceptance cuts.

The full angular distribution, Fig. 4 (top), is approximated within a few percent, over the full range, by the sum of the unpolarised distributions. The exact result, shown by the black histogram, however, is not of the form of Eq. (12). This becomes clear expanding the full result on the first three Legendre polynomials which yields the blue, green, and orange smooth curves in Fig. 4: their sum (smooth grey curve) fails to describe the correct distribution, thus the Legendre expansion cannot be applied when lepton cuts are imposed.

Computing the polarised cross section for different bins of $WW$ invariant mass, we obtain the polarisation fractions as functions of $M_{WW} > 300$ GeV.

**Fig. 4.** Distribution of $\cos \theta$ in the $W^-$ reference frame (top), polarisation fractions as functions of $M_{WW}$ (bottom). Lepton cuts: $p_T^{\ell} > 20$ GeV, $|\eta^{\ell}| < 2.5.$
M_{\text{NEW}}, shown in the bottom panel of Fig. 4: even in the presence of cuts, the fraction of longitudinally polarised $W^-$'s is well above 10%.

The fact that the exact distribution is well described by the incoherent sum of the polarised differential distributions allows for a measurement of the polarisation fractions, within a single model, even in the presence of cuts on the charged leptons, using Monte Carlo templates for the fit.

This analysis demonstrates that it is possible to study the polarisation of massive gauge bosons in a well defined set-up for VBS processes. The method has been implemented in the code Phantom [57].

2.4. Effective field theory for vector-boson scattering

VBS processes represent a particularly interesting probe of new physics, as they give a unique access to the couplings of gauge bosons. Without committing to a specific model, a convenient instrument for testing experimental data against the presence of BSM effects is that of effective field theories (EFTs).

In a EFT approach, the SM is assumed to be the low energy limit of an unknown UV completion, whose typical scale $\Lambda$ is well separated from the electroweak one. In this scenario, the new physics sector is decoupled and its impact onto observables measured at $E \ll \Lambda$ can be parametrised without specifying any property of the UV completion, by means of a Lagrangian that contains only the SM fields and respects the SM symmetries. New physics effects are organised in a Taylor expansion in $E/\Lambda$, i.e. they are encoded in an infinite series of gauge-invariant operators ordered by their canonical dimension. This is often called SMEFT (SM EFT) Lagrangian and, neglecting lepton number violating terms, it reads

$$L_{\text{SMEFT}} = L_{\text{SM}} + \frac{1}{\Lambda^4} L_{\text{dim-6}} + \frac{1}{\Lambda^6} L_{\text{dim-8}} + \cdots,$$

(13)

with the dots standing for higher orders. The SMEFT Lagrangian constitutes a convenient theoretical tool for probing the presence of new physics, as it provides the only systematic parameterisation of its effects that is both well-defined as a field theory and model-independent, in that it can be matched onto any UV completion compatible with the SM symmetries and field content.

One can restrict to leading deviations from the SM cutting the series at dimension 6 which reads

$$L_{\text{dim-6}} = \sum_i C_i O_i,$$

(14)

Here $\{O_i\}$ is a set of gauge-invariant dimension-6 operators that form a complete basis and $\{C_i\}$ are the corresponding Wilson coefficients. Any evidence for a non-zero Wilson coefficient would represent a smoking gun of new physics. Further, knowing which terms are non-vanishing can allow to characterise the new physics states and help designing more effective direct search strategies.

A complete basis for dimension-6 terms contains 59 independent structures (+ their Hermitian conjugates) that in complete generality are associated to 2499 independent parameters [58]. This number can be significantly reduced by assuming CP conservation and/or an approximate $U(3)^5$ flavour symmetry. Choosing convenient kinematic cuts in the experimental measurements can also help to restrict the set of relevant operators. Different basis choices for $L_{\text{dim-6}}$ have been proposed in the literature, that are related by equation-of-motion and integration-by-parts transformations. Despite containing different sets of operators (often distributing the effects differently among fermion and boson couplings), all the bases give equivalent parametrisations for physical $S$-matrix elements, i.e. once a complete process with stable external states is computed. The so-called Warsaw basis [59] is sometimes preferred, due to the fact that this was the first complete basis in the literature and that its renormalisation group evolution (RGE) and one-loop renormalisation are completely known [58,60–64].

Assuming CP conservation and a $U(3)^5$ flavour symmetry, VBS processes receive corrections from 14 dimension-6 operators. To keep the analysis as general as possible and to have a well-defined IR limit of a given underlying UV sector, these should be all considered simultaneously in the fit. Setting a subset of the Wilson coefficients to zero cannot be done arbitrarily. For example, this may spoil strong correlations hidden in the parametrisation and artificially remove blind directions. In particular, including anomalous fermion couplings may have a significant impact on the analysis, despite the strong constraints imposed by LEP measurements (see e.g. Reference [67,68] for a recent study in the context of $W^+ W^-$ production at the LHC). A reduction of the number of parameters may be necessary, nonetheless, for the technical feasibility of the analysis. In this case the removal of some (combination of) operators may be very carefully considered in the future.

The possibility of extending the EFT analysis with dimension-8 operators has also been discussed, as these terms can introduce important decorrelation effects between triple and quartic gauge couplings. Although this is an interesting avenue, exploring it in a consistent way is a challenging task due to the extremely large number of parameters involved (considering one fermion generation, there are 895 B-conserving independent operators at $d = 8$, among which up to 86 can contribute to quartic gauge couplings (QGCs).
and TGCs [69]) and to the fact that a complete basis of dimension-8 operators is not available to date. Therefore it is advisable to defer this study to a later stage. A more compelling alternative is rather performing an analysis in the basis of the Higgs EFT (HEFT), for which complete bases have been presented in References [70,71] (see references therein for further theoretical details and previous phenomenological studies). The HEFT differs from the SMEFT in that it is not constructed with the Higgs doublet, but rather embedding the Goldstone fields into a dimensionless matrix \( U = \exp(i\theta^a u^a / v) \) (analogously to the pion fields in chiral perturbation theory) and treating the physical Higgs as a gauge singlet. The HEFT is more general than the SMEFT and it matches the low energy limit, for instance, of some theories with a strongly interacting electroweak symmetry breaking sector in the UV, such as composite Higgs models. Such an analysis would be highly motivated as the scattering of longitudinal gauge bosons constitutes one of the best probes for UV scenarios matching the HEFT (see e.g. References [72,73] for recent studies), and they are among the observables that may allow to disentangle it from the SMEFT. The number of relevant Wilson coefficients for VBS in the HEFT (in the CP conserving, \( U (3^c) \) symmetric limit) is about 30, which is larger than for the SMEFT but much lower than for including a complete dimension-8 set of operators, which makes this analysis an ideal follow-up to the SMEFT one.

One of the main points to be addressed in the EFT analysis is that of its validity: as mentioned above, adopting a dimension-6 parametrisation is theoretically justified only for \( \Lambda \) sufficiently larger than the Higgs vev \( v \). Namely the impact of dimension-8 terms \( \sim (E/\Lambda)^8 \) should be roughly smaller than the experimental uncertainty. When analysing experimental data, however, the cutoff scale \( \Lambda \) is unknown and the actual energy \( E \) exchanged in the process is often inaccessible too. Extracting \( E \) is particularly complex for VBS at the LHC, with various scales entering at different stages in the (sub-)process(es). Thus the validity of the EFT cannot be established a priori: at best it can be verified a posteriori, checking that the energy range of the distributions used for the fit does not exceed the lower limit obtained for the cutoff. Some methods of this kind have been discussed in the literature (see e.g. References [68,74–81]) and could also be applied to VBS studies. If a constraint is found to be incompatible with the validity of the EFT itself, it should be rejected. Attention should be paid to the application of unitarisation methods, that are often employed to correct the divergences obtained in the kinematic distributions of Monte Carlo generated signals. Introducing a damping of the distribution tails, these techniques may alter the behaviour of the Taylor series in a way that does not reflect the correct behaviour of the EFT at high energies (which is indeed divergent where the expansion breaks down) and lead to an incorrect estimation of the constraints.

The first step of the EFT-VBS program is an accurate theoretical study of VBS in the SMEFT at dimension-6, which includes agreeing on a given parametrisation, evaluating the necessity of reducing the number of operators considered and testing the capabilities of available theoretical tools (Monte Carlo generators etc.) This will be conducted in parallel with a preliminary study of the experimental constraints that could be obtained. One of the primary goals of these studies, in which both theorists and experimentalists will participate, is to define an optimal way to report data (cross sections and differential distributions) that maximises the transparency and versatility of the results. Finally, further avenues are worth exploring in subsequent stages, among which the analysis of the HEFT basis (and later on, if possible, of dimension-8 operators) and a comparison of the impact of VBS processes with that of other data sets, with the possibility of considering a combination of different measurements in the fit.

3. WG2: Analysis techniques

3.1. Experimental overview

At the time of the workshop, a number of experimental results in VBS have been made available, all of them from the LHC experiments CMS and ATLAS (see Table 6). The highlight among these results is a measurement from the CMS experiment in the \( W^\pm W^\mp \) channel, which for the first time observes the VBS contribution in EW processes with a significance above 5 \( \sigma \) (5.5 \( \sigma \) observed) [7]. It is interesting to notice that apart from this observation only two other evidences have been found, one for the VBS production in the \( Z\gamma \) channel and one for the exclusive production \( \gamma\gamma \rightarrow WW \). The lack of evidences and observations of the VBS process is a consequence of its very small cross-section (order of 1 fb), but also of the large size of systematic uncertainties coming from background evaluation and the jet reconstruction in the forward region of detectors.

By now, a large fraction of the different possible final state boson combinations have been studied by at least one experiment, with the notable absence of the \( \gamma\gamma \) and \( W^\pm W^{\mp} \) channels in the non-exclusive channel, which are much more complex to reach given the amount of experimental background associated to them. This wide coverage of channels has been shown to be very helpful when constraining aQGCs, as the different channels show varying sensitivity to different operators, as shown in Figs. 5 and 6. With the larger datasets collected during the second phase of exploitation of LHC (Run2), experiments will be able to reach more easily the VBS phase space and be able to study channels that are not accessible yet.

The presentation and interpretation of results in the VBS studies reviewed during the workshop show some notable differences among the various analyses and experiments.

The first difference is on the treatment of interference between electroweak and QCD amplitudes in the predictions for SM cross sections. Most of the time, the interference is derived from the \texttt{MADGRAPH} [92] or \texttt{PHANTOM} [57] generators, and is treated as systematic uncertainty on the signal yield, leading to 5–10% uncertainties. In some cases, on the other hand, the interference is treated as a signal (\( W^\pm W^{\mp} \) for the ATLAS experiment) or neglected when found to be too small (generally below 4%). Such differences could complicate the combination of results. Even if, with respect to the current level of experimental accuracy, the differences in treatment of interference effects are still small, with the continued data-taking at the LHC a common approach is desirable.

\textsuperscript{9} Speaker: N. Lorenzo Martinez.
Another difference is the framework in which the aQGCs results are interpreted. There are mainly two approaches: one based on the CP-conserving dimension-8 EFT operators that maintains SU(2)\textsubscript{L} × U(1)\textsubscript{Y} gauge symmetry of the type \(\phi^4/λ^4\) (for the ATLAS experiment the Zγ channel, for the CMS on the Zγ, Wγ, W±W±, and ZZ zones), the other one based on the \(α^4\) and \(α^5\) coefficients of the two linearly independent dimension-4 operators contributing to aQGCs (for the ATLAS experiment the WZ, W±W±, WV semileptonic channels). Even if the conversion between the two frameworks can be done, it can be better for quick comparison and combination to adopt an unique framework.

Finally, another difference lies in the treatment of unitarity issues that can arise when the analysis sensitivity to potential aQGCs is not high enough to exclude aQGC values within the energy range where the EFT can be considered unitary. In the \(α^4\), \(α^5\) framework, the unitarisation condition is imposed with the so-called K-matrix method, within the WHIZARD\[46\] generator. In the dimension-8 EFT operator framework, different approaches are used by ATLAS and CMS, respectively. ATLAS scales the spectra with analytical form factors of the type \(f_j/(1 + s/Λ_{\text{cut}})\)\textsuperscript{n}, where \(n = 2\) and \(Λ_{\text{cut}}\) is the cut-off scale. CMS, on its side, provides a validity bound, i.e. the scattering energy at which the observed limit would violate the unitarity, derived with the VBFNLO\[38\] generator. Some of the results published by the CMS collaboration show that at the LHC scales, already many limits are set in the unitarity unsafe region. Choosing different approaches could severely complicate the combination of different aQGCs limits measurements. Performing such combinations could substantially increase the statistical power of the total dataset and could help to break degeneracies between the effects of different operators which may affect a single channel in similar ways. Providing recommendations to unify the treatments mentioned above is an important goal for WG2.

Another lesson learned from the experimental review concerns the modelling of the main background producing the same final

### Table 6
Summary of all published experimental results on VBS processes by final state with the details on luminosity and energy at the centre of mass \(\sqrt{s}\) used for the measurements. When available both expected and observed significances are provided. Channels for which “Only aQGC limits” were studied are indicated in the significance column.

<table>
<thead>
<tr>
<th>Channel</th>
<th>(\sqrt{s})</th>
<th>Luminosity [fb\textsuperscript{-1}]</th>
<th>Observed (expected) significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Z(\ell\ell)\gamma)</td>
<td>8 TeV</td>
<td>ATLAS 20.2, CMS 19.7</td>
<td>2.0σ (1.8σ) [82], 3.0σ (2.1σ) [83]</td>
</tr>
<tr>
<td>(Z(\ell\ell)\gamma)</td>
<td>8 TeV</td>
<td>ATLAS 20.2, CMS 19.7</td>
<td>Only aQGC limit [82]</td>
</tr>
<tr>
<td>(W^±W^±)</td>
<td>8 TeV</td>
<td>ATLAS 20.3, CMS 19.4</td>
<td>3.6σ (2.3σ) [45], 2.0σ (3.1σ) [6]</td>
</tr>
<tr>
<td>(W^±W^±)</td>
<td>13 TeV</td>
<td>ATLAS 35.9, CMS 35.9</td>
<td>5.5σ (5.7σ) [7]</td>
</tr>
<tr>
<td>(W(\ell\nu)\gamma)</td>
<td>8 TeV</td>
<td>ATLAS 20.2, CMS 19.7</td>
<td>Only aQGC limit [86], N/A [6]</td>
</tr>
<tr>
<td>(W(\ell\nu)(Z\gamma))</td>
<td>8 TeV</td>
<td>ATLAS 20.2, CMS 19.4</td>
<td>Only aQGC limit [87], N/A [6]</td>
</tr>
<tr>
<td>(γγ → WW)</td>
<td>7 TeV</td>
<td>ATLAS 20.2, CMS 24.8</td>
<td>3.0σ [89], 3.4σ (2.8σ) [90]</td>
</tr>
</tbody>
</table>

### Fig. 5.
Limits on dimension-8 mixed transverse and longitudinal parameters \(f_M, i\) [91].
state via QCD interactions, which is most of the time very important. To control its impact on the analyses, experiments use control regions (generally low dijet invariant mass) in which they constrain and verify the QCD background normalization and shape. While until now the precision is not enough to constrain these quantities, with more statistics it will become more relevant and QCD modelling issues could become one limiting factor. At the same time, care should be taken in the signal definition as well, because of the presence of the interference between the electroweak and QCD production of the VBS final state.

The measurements are currently dominated by statistical uncertainty, due to the very low VBS cross-sections. Then generally follow the uncertainty on the jet energy scale and resolution, the uncertainty on background estimation and theory uncertainties (scales, parton distribution functions). The experimental uncertainties together with the previous point will be important to mitigate in the future: this is one of the goals of WG2 and WG3.

Finally, it is interesting to notice the importance of the WV channel, where V is a W or Z boson decaying hadronically. Thanks to advanced jet substructure techniques, this channel brings the tightest constraint on EFT charged parameters, since the boosted topology allows to reach higher energy regimes. The WG3 activities will focus on such techniques as well.

### 3.2. Common selection criteria

In order to facilitate feasibility studies and similar forward looking analyses in a way that allows for a fair comparison between such studies, it is useful to define a common baseline selection for a fiducial phase space. This baseline criteria aim at selecting the two bosons and the two jets that are found in the final state of VBS processes, without entering yet in a VBS-enriched region (this will be a subject for later discussions in WG1, WG2 and WG3). A single definition cannot serve as the base for every study for the following reasons:

- Different VBS channels and different boson decay modes may require notably different selection criteria.
- The experiments don’t always have the same geometrical acceptances and efficiencies in the final state objects identification and reconstruction (depends on the calorimeter and muon systems),
- The experiments will evolve due to hardware upgrades planned for the high luminosity phase of the LHC, necessitating different assumptions on the detector acceptance depending on the integrated luminosity used for the forward looking study.

A simple example can be taken from studies of VBS channels with signals large enough to be amenable to a simple “cut & count” style analysis, notably the $W^\pm W^\pm$ channel. Based on published results, a phase-space region close to CMS as well as ATLAS studies has been identified (see Table 7). This phase space definition can be used for forward-looking comparison studies within the experiments reaching up to, but excluding the high luminosity phase of the LHC, when experiments will undergo upgrades that will improve their acceptance.

---

**Fig. 6.** Limits on dimension-8 transverse parameters $f_{T, i}$ \([91]\).
Table 7
Proposed phase space for studies in the $W^+W^−$ channel.

<table>
<thead>
<tr>
<th></th>
<th>Electrons</th>
<th>Muons</th>
<th>Jets</th>
<th>Photons</th>
</tr>
</thead>
<tbody>
<tr>
<td>$</td>
<td>\eta</td>
<td>\leq$</td>
<td>$&lt; 2.5$</td>
<td>$&lt; 2.4$</td>
</tr>
<tr>
<td>$p_{T\text{lead}}$</td>
<td>$&gt; 25$ GeV</td>
<td>$&gt; 25$ GeV</td>
<td>$&gt; 30$ GeV</td>
<td>$&gt; 25$ GeV</td>
</tr>
<tr>
<td>$p_{T\text{sublead}}$</td>
<td>$&gt; 15$ GeV</td>
<td>$&gt; 15$ GeV</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

These numbers differ from the ones quoted in Section 2.2, that were used only for a Monte Carlo comparison purpose of the same VBS process, and then did not need to follow exactly the experimental constraints (for example the current lepton trigger energy thresholds).

Further extrapolation into the future suffers from the problem that the design work for the envisioned detector upgrades is not entirely completed yet, so that work in this area is necessarily somewhat speculative. Nevertheless a few general conclusions can be drawn from existing detector design documents [93–98]. Both experiments plan to improve their trigger systems to keep the thresholds for lepton triggers at a similar level to current running conditions, so that $p_T$ thresholds should remain close to current ones. Both experiments also plan to extend coverage of their respective tracking detectors up to $|\eta| \sim 4$.

For studies of channels with very low cross section, branching fraction or efficiency, for example the $ZZ \rightarrow 4\ell$ channel [85], the phase-space defined above is not suitable, and the analysis is performed in a more inclusive phase-space (by lowering the threshold on the lepton transverse momentum).

3.3. Prospects

The presentation touched on several major topics interesting for future studies. First among these was the study of the final state bosons polarisation fractions. The scattering of longitudinal vector bosons violates unitarity in the absence of a standard model Higgs boson. By looking at the scattering of electroweak gauge bosons we will be probing the Higgs boson properties. At the LHC by the time of the workshop the $W^\pm W^\mp \rightarrow t\bar{t}4\ell$ channel was the only one with observation of weak boson scattering. The polarisation is accessible through the angular distributions of the boson decay products in the boson rest frame, but in this particular channel the decay products include two neutrinos, which prevent a straight-forward reconstruction of the boson decay angular distribution. Other channels, which allow for the reconstruction of these angular distributions, on the other hand, suffer from much larger backgrounds. Several potential approaches to address the issue of the missing information were presented. One of them is to use a number of mass-like variables [99,100] that have shown to be able to distinguish between the $W$ polarisations. Fig. 7 shows the contributions from the different polarisation states as a function of two of this potential mass-like variables $M_{1\gamma}$ and the $M_{3\gamma}$. The use of this variables may allow to extract polarisation information even from final states with two neutrinos and might also be sensitive to new physics effects. Another possibility to access the polarisation will be the use of a regression technique for pulling out the missing information [101]. In this sense it might be interesting to use a deep learning technique to output the true polarization values [102] using as input measurable quantities like leptons, jets and missing energy ($p_T, \eta, \phi$ etc.).

As a topic for future studies, the presentation discussed potential information to be gained by measuring the relative cross sections of different VBS channels, in particular channels related by charge symmetry. Beyond the naive expectations related to the valence quark content of the proton, the charge cross section ratio can provide useful constraints on the effects of the underlying parton density functions [103]. It was shown as well that these ratios can provide sensitivity to BSM processes [104]. As an example Fig. 8 shows for the $W^\pm Zjj$ channel the charge cross section ratio using different unitarization prescriptions and different aQGC parameters. For increasing aQGC parameters the charge ratio increases up to a plateau separating from the SM prediction.

Additionally, the interpretation of experimental results in terms of BSM effects was discussed. While the EFT approach discussed above has the advantage of being generic and provides a complete description of a very wide range of BSM models, there is also a number of shortcomings. However, anomalous couplings of a size that will cause observable effects at low scales, will commonly violate unitarity at high scales, so that ad-hoc approaches to unitarization are applied which introduce a new model-dependence: the dependence on the unitarization scheme. Beyond this, there is the issue that the EFT is only valid in the approximation that the scale of the observed scattering is much smaller than the scale $\Lambda$. Empirical studies, comparing explicit resonance models to EFTs [105], show that the amount by which $\Lambda$ has to exceed the described region of data can be substantial, to the point where visible effects that are accurately described by an EFT would correspond to theories so strongly coupled, that their treatment in perturbation theory would be questionable. A way to avoid these problems would be an interpretation using a simplified resonance parameterization [106,107], which does not run into unitarity issues, but is by construction less general than the EFT approach.

---

11 Speaker: M. Kobel.
12 $M_{1\gamma} = \frac{1}{2}(p_T(\ell_1)+p_T(\ell_2)) - (p_T(\ell_1)+p_T(\ell_2)) + |p_{T\text{miss}}|^2$.
13 $M_{3\gamma} = \frac{1}{2}(\sqrt{M_{1\gamma}^2 + p(\ell_1) + p(\ell_2) + |p_{T\text{miss}}|^2} - (p(\ell_1) + p(\ell_2) + |p_{T\text{miss}}|^2$.
At the high centre of mass energies of the Large Hadron Collider even the heaviest known SM particles can be observed with large transverse momenta, in the so called boosted topology. Boosted W, Z, and Higgs bosons, and top quarks that decay to quarks will have highly collimated decay products and can therefore be reconstructed in a single jet with large radius parameter $R$. In ATLAS jets are reconstructed with the anti-$k_{T}$ algorithm, usually requiring a transverse momentum of $p_T > 200$ GeV and a radius parameter of $R = 1.0$, and are trimmed \cite{108} with the parameters $R_{\text{sub}} = 0.2$ and $f_{\text{cut}} = 5\%$.

To distinguish signal, e.g. real W bosons, from QCD induced jet backgrounds one of the main observables is the jet mass, calculated from the jet constituents. In Fig. 9 the distribution of the jet mass in data and simulation is shown after a lepton + jet selection that aims at selecting $t\bar{t}$ events. It shows a clear peak at the expected SM W boson mass. Using a further discriminating variable W boson taggers are built that have a fixed signal efficiency of 50% and reduce the background by a factor of 50.

Recently, the possibilities of adding more information and exploiting multi-dimensional correlations have been explored by using Boosted Decision Trees (“BDT”) and Deep Neural Networks (“DNN”) to tag boosted hadronically decaying W bosons. Compared to simple 2-variable tagging approaches the multivariate ones perform better, as can be seen in Fig. 10.

\cite{104} Speaker: C. F. Anders.
Jet substructure techniques for VBS in CMS

Jet identification techniques that make use of jet substructure information are important tools for the measurement of VBS. A brief summary of the existing tools used by the CMS experiment and future prospects for the HL-LHC is given in the following.

To probe high $WW$, $ZZ$ or $WZ$ invariant masses, special jet identification techniques for $W$ and $Z$ bosons decaying to quarks are needed, since for high momentum $W$ and $Z$ bosons, the shower of hadrons originating from the quark anti-quark pair merges into a single large radius jet of particles \cite{111–113}. The maximum angular separation between the quark and anti-quark is given by $R_{qq\bar{q}} = \sqrt{\frac{m^2}{p_T^2}}$, where $m$ and $p_T$ are the mass and transverse momentum, respectively, of the $W$ or $Z$ boson. For a $W$ boson with $p_T = 1$ TeV, an angular separation of $R_{qq\bar{q}} = 0.2$ is expected, which is well below the typical jet size parameter of 0.4 used by CMS. At even higher $p_T = 3.5$ TeV, the angular distance $\Delta R_{qq\bar{q}} = 0.05$ between the decay products of a $W$ boson is even smaller than the granularity of the hadron calorimeter of CMS with cell sizes of $\Delta \eta \times \Delta \phi = 0.087 \times 0.087$ in the barrel region of the detector. CMS thus employs a particle-flow event algorithm \cite{114} to measure jet substructure, that reconstructs and identifies each individual particle with an optimized combination of information from the various elements of the CMS detector, benefiting from spatial and energy resolution of all sub-detectors.

Fig. 11(left) shows the main observable used by CMS to distinguish $W$ and $Z$ boson jets from quark and gluon initiated jets, the softdrop jet mass, which is the mass of the jet after iteratively removing soft radiation with the modified mass-drop...
This “softdrop” algorithm [117] reduces the mass of quark and gluon jets and improves the mass resolution of W and Z boson jets. In addition, the substructure of the jet is explored with an N-subjettiness [118] ratio that distinguishes the W and Z boson jets composed of two hard subjets from single quark and gluon jets. With this combination of observables, a mistag rate of 1% at an efficiency of 50% is achieved for a broad range of jet transverse momenta from $p_T > 200$ GeV up to at least 3.5 TeV. The jet mass and substructure observables are calibrated using a data sample of top pair production in the single lepton final state containing high-$p_T$ W bosons, achieving uncertainties of the order of 1% on jet mass scale and 10% on jet mass resolution and jet substructure tagging efficiency, which increase at higher jet $p_T$ where simulation is used for extrapolation. Particles from additional interactions happening in the same pp bunch crossing, called pileup interactions, can significantly distort these observables. CMS thus employs dedicated particle-based pileup removal techniques that correct not only jet momenta, but also jet shape and substructure observables. Charged particles that are identified by the tracking detector to originate from pileup interaction vertices are removed before jet clustering (this procedure is called charged hadron subtraction, CHS). For neutral particles a probability weight based on the distribution of surrounding particles following the pileup per particle identification (PUPPI) algorithm [119] is applied to the particle four momenta [120]. With these pileup suppression techniques, the performance of W and Z boson identification is constant up to at least 40 pileup interactions, and with the higher granularity tracking detector planned for the HL-LHC, performance is maintained up to 200 pileup interactions. Notable for VBS studies, the longitudinal and transverse polarization of W boson jets can be separated using subjet information, yielding a mistag rate of 30% at an efficiency of 50% [113].

Also the two forward jets from VBS require an analysis of their substructure, in order to suppress the background from (possibly overlapping) jets originating from pileup interactions [111,121]. In a scenario of 25 pileup interactions (about half of the pileup expected in Run II of the LHC), without pileup mitigation ~50% of VBS selected forward jets come from pileup. Since for $|\eta| > 3.0$ no tracking information is available to suppress pileup particles and also jet shape differences are difficult to resolve due to coarse calorimeter granularity ($\Delta\eta \times \Delta\phi = 0.175 \times 0.175$), a multivariate analysis (MVA) is needed to distinguish quark jets from pileup and gluon background. Fig. 11 (right) shows the pileup jet MVA discriminator that allows to achieve a pileup mistag rate of ~30% at a quark efficiency of 50%. At the HL-LHC, for which CMS tracking and vertex identification will be extended from $|\eta| < 2.5$ to $|\eta| < 4$ and a high granularity endcap calorimeter will be installed in the region $1.5 < |\eta| < 3$, VBS jet identification can rely on the much more powerful CHS and PUPPI pileup rejection techniques.

5. Summary

The VBSCan COST Action aims at a consistent and coordinated study of VBS from the phenomenological and experimental points of view, gathering all the interested parties in the high-energy physics community, together with experts of data mining techniques. In fact, the complex structure of the process calls for a joint effort to exploit at best the data sets that will be delivered by the LHC in the following years. This will require a close interaction between the experimental community and the theory one, as well as the deployment of the most advanced data analysis techniques to maximise the reach of the ATLAS and CMS collaborations. This document contains the summary of the kickoff meeting of the VBSCan Action, happened in June 2017, and paves the way for the activities of the Network.

Acknowledgements

The authors would like to acknowledge the contribution of the COST Action CA16108. We would like to thank the Split-FESB team for the great hospitality in Split. Kristin Lohwasser, Hannes M¨ildner, and Philip Sommer are supported by the European Union Horizon 2020 research and innovation programme under ERC grant agreement no. 715871. Nigel Glover, Raquel Gomez-Ambrosio,
Giulia Gonella, Markus Schumacher acknowledge the support of the Research Executive Agency (REA) of the European Union under the grant agreement PITN-GA-2012-316704 ("HiggsTools"). Stefan Dittmaier and Christopher Schwan acknowledge support by the state of Baden-Württemberg through bwHPC and the DFG through grant nos. INST 39/963-1 FUGG and DI 784/3. Benedikt Biedermann, Ansgar Denner, and Mathieu Pellen acknowledge financial support by the German Federal Ministry for Education and Research (BMBF) under contract no. 05H15WCCA1 and the German Research Foundation (DFG) under reference number DE 623/6-1.

Andreas Hinzmann gratefully acknowledges funding in the Emmy-Noether program (HI 1952/1-1) of the German Research Foundation DFG. Ilaria Brivio and Michael Trott acknowledge support from the Villum Foundation, NBLA, the Discovery Center at Copenhagen University and the Danish National Science Foundation (DNRF91). Qi Jiang Li is supported in part by the Swiss National Science Foundation (SNF) under contract 200020-175595. The work of Marc-André Pleier was supported by the DOE contract no. DE-SC0012704.

References

4fermionsindouble-poleapproximation:the
8TeVwiththeATLAS


[91] CMS Collaboration, Kolion Limits on anomalous triple and quartic gauge couplings, https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSMPatGCC.


