New Developments in WHIZARD Version 2.6

W. Kilian[1a], S. Brass[2a], T. Ohl[3b], J. Reuter[4c], V. Rothe[5c], P. Stienemeier[6c],
M. Utsch[7a]

[a]*University of Siegen, Department of Physics, D–57068 Siegen, Germany*
[b]*Würzburg University, Faculty of Physics and Astronomy, D–97074 Würzburg, Germany*
[c]*DESY Theory Group, D–22603 Hamburg, Germany*

*Talk presented by W. Kilian at the International Workshop on Future Linear Colliders
(LCWS2017), Strasbourg, France, 23-27 October 2017. C17-10-23.2.*

ABSTRACT

We describe recent additions to the WHIZARD 2 Monte-Carlo event generator
which improve the physics description of lepton-collider event samples and
speed up the calculation time required for cross sections and event generation.

[1]kilian@physik.uni-siegen.de

[2]brass@physik.uni-siegen.de

[3]Thorsten.Ohl@physik.uni-wuerzburg.de

[4]juergen.reuter@desy.de

[5]vincent.rothe@desy.de

[6]pascal.stienemeier@desy.de

[7]manuel.utsch@t-online.de

1 The WHIZARD Monte-Carlo Generator

The WHIZARD Monte-Carlo event generator [1,2] is a stand-alone program that calculates cross sections, distributions, and fully exclusive event samples of perturbative high-energy processes at colliders. The program handles particle-decay processes as well as scattering processes at hadron colliders and for the planned lepton colliders such as ILC and CLIC.

The program has been used to generate simulated event samples for a wide variety of lepton-collider studies. New developments and improvements of the program will allow for producing event samples with a more complete physics description, fueling refined and extended physics and detector studies for the ILC and CLIC collider projects.

Within WHIZARD, functional expressions for the elementary processes are constructed as needed in form of source code when running the program, calling the algebraic matrix-element generator O'Mega [3,4] for tree-level amplitudes. QCD is handled in the color-flow calculational scheme [5]. This code is compiled and linked to the main program on the fly. Beyond leading order in perturbation theory, WHIZARD provides an interface to next-to-leading order (NLO) virtual-amplitude providers OpenLoops [6], GoSam [7], and Recola [8,9]. NLO calculations with WHIZARD [10,11,12,13] have been recently automatized [14], implementing infrared-collinear subtraction via the FKS [15] subtraction scheme.

Precision calculations for e^+e^- colliders require a detailed treatment of interactions of the incoming beams. WHIZARD describes beam structure via the CIRCE1 [16] and CIRCE2 modules. CIRCE2 is a dedicated beam-event generator that is based on fitting detailed simulation results for the beam-beam interactions. The program supports any mode and degree of beam polarization. Furthermore, the universal part of soft-collinear photon radiation from the initial state is accounted for by an inclusive structure-function approach for electrons, positrons, and photons.

The effects of radiation from the final state are handled by an internal call to PYTHIA 6 [17]. The internal PYTHIA 6 implementation also provides the transformation to the fully hadronic final state. Alternatively, events can be showered using WHIZARD's own analytic shower [18], or externally via event files. For communicating with further external processing of events, WHIZARD supports a comprehensive set of event-file formats.

The WHIZARD program is optimized for handling multi-particle final states resulting from hard processes, such as the four- to eight-fermion final states that result from the production of the heavy resonances W, Z, H, and t at future lepton colliders. To this end, WHIZARD follows a multi-channel approach with concurrent phase-space parameterizations that reflect the singularity structure of the matrix element. Integration over phase space is realized by the VAMP module [19], which is a multi-channel extension of the well-known adaptive Monte-Carlo integration algorithm VEGAS.

Regarding the definition of physics models, WHIZARD contains a comprehensive predefined library of models including the SM (Standard Model), supersymmetry [20], Little-Higgs models, or extra dimensions. Beyond these hard-coded models, it contains a model implementation interface to the SARAH [21] and FeynRules [22,23] programs, and it supports the UFO [24] model-definition file format.

For providing input data to the program, steering the workflow, describing cuts and weight factors, and for internal event analysis, WHIZARD implements a specific scripting language,

SINDARIN. The language supports calculations and manipulations of events and subevents, provides an interface to the `FastJet` [25] jet-algorithm package, and enables all I/O of the generator. It allows for conditionals, evaluation loops for parameter scans, concurrent alternative parameter sets, and matrix-element reweighting of event samples. The internal analysis covers free-form selection criteria and enables histograms, tables, and plots within a uniform workflow.

2 Issues in Standard-Model Event Generation

Physics and detector studies for future lepton colliders rely on the availability of validated event samples for all SM processes that are expected to be observable. These include resonant production processes that incorporate two or more massive SM particles, W, Z, or H, or top-quark pairs which subsequently decay into light SM particles. For precise predictions, the calculation has to involve complete matrix elements that include both resonant and non-resonant contributions at a given order in the loop expansion. Current simulation studies generally involve matrix elements at leading order, which in the future are to be replaced by (virtual) higher-order matrix elements.

A large database of SM event samples has been obtained in the past using version 1 of the `WHIZARD` generator. Current and future studies employ version 2 of `WHIZARD`. The second version offers much greater flexibility and convencience as an application, and is set up for improved precision and to cover a wider range of physics models and effects.

In this note, we describe issues that have emerged in the course of this transition to an improved and more versatile framework, their physics impact, and technical solutions. `WHIZARD 2` event samples have been validated by the ILC-CLIC generator group against `WHIZARD 1` event samples in those contexts where this comparison is meaningful. We have identified three areas where improvements were required, such that the new generation of event samples can cover a larger set of processes with reasonable approximations and good efficiency.

1. exclusive generation of semi-hard photons, which are included in the initial-state radiation (ISR) approximation but warrant a nontrivial modification of collinear kinematics;

2. kinematic distortion of resonance shapes due to the approximations involved in the `PYTHIA 6` parton-shower algorithm;

3. a re-implementation of `VAMP` which now supports highly parallel integration and event generation within the message-passing interface (MPI) communication model.

4. A dedicated treatment of NLO contributions and resummed threshold logarithms for off-shell $t\bar{t}$ and $t\bar{t}H$ processes [26,27] is covered in a separate contribution to these Proceedings.

3 Semi-Hard ISR Photons

A fully inclusive ISR description is a convenient method to account for leading-logarithmic effects of multiple photon radiation from the initial state. The description implemented in

`WHIZARD` includes resummed soft photons to all orders, as well as higher-order universal collinear contributions. This is sufficient, given the leading-order approximation in the SM perturbation expansion of the hard matrix elements, to describe the dominant effect of ISR photon emission, namely energy loss and the resulting distortion in the shape of s-channel resonances and thresholds. It also yields the dominant QED corrections to calculated cross sections.

However, in the context of fully exclusive event generation, there is a non-negligible fraction of events where emitted photons are not strictly soft or collinear but carry away a measurable amount of transversal momentum. While this effect can safely be neglected for inclusive quantities, the resulting p_T kick on hard-process kinematics can significantly distort event shapes and distributions of more exclusive observables.

In practice, this p_T mismatch has to be taken into account for both incoming beams simultaneously. The approach in `WHIZARD` 1 to this problem was to sample transverse momentum for both radiated photons independently of each other, according to the logarithmic distribution which prevails over most of the collinear phase space. Regarding the hard process, this was combined to an ad-hoc kinematics modification which did violate, to some extent, either energy or momentum conservation. The algorithm was applied both to cross-section integration and to exclusive event generation. It turns out that the actual inaccuracies were of minor importance but there was some uncontrolled impact on the validity of the prediction near phase-space boundaries.

The implementation of transverse-momentum generation in early versions of `WHIZARD` 2 was intended to describe a more generic chain of radiation and on-shell projections for each beam individually. Unfortunately, such an approach results in more drastic energy-momentum mismatches with unphysical results for various observables. The discrepancy is visible in the comparison between `WHIZARD` 1 and `WHIZARD` 2 samples, such that it should not be tolerated for practical applications.

Therefore, version 2.6 supports a new algorithm for approximating the effect of photon transverse momentum. The kinematics calculation takes into account both beams simultaneously. The new version conserves energy and momentum exactly. The only (inevitable) on-shell projection is applied to the initial partons of the hard process, which themselves are unobservable and not part of the physical event. Furthermore, the integration is now carried out in the strict collinear limit where the ISR approximation is defined. Only in the simulation pass, the generated hard-process events are modified individually according to the logarithmic transverse-momentum distribution. The hard event, and any radiation originating from it, is merely boosted by that effect, while the universal behavior of the radiated photons is correctly described. The approximation loses its validity for p_T of the order of the hard-process scale, where a NLO SM calculation would be required to compute the process-dependent contributions.

4   Resonances and Parton Shower

In the current `WHIZARD` framework, events corresponding to leading-order matrix elements are combined with the `PYTHIA 6` parton shower module. The `WHIZARD/O'Mega` matrix elements

are complete in the sense that they contain all Feynman graphs that connect the initial state to the selected final state. Typically they contain both resonant and non-resonant contributions.

When applying QCD radiation to a colored final state, there are two effects that must be handled in the presence of resonances. For instance, a resonant W boson is a colorless particle, but its $q\bar{q}$ decay state contains color sources and thus initiates a parton shower. (i) If the p_T of a radiated parton off the decay products is larger than the resonance width, the radiation effect shifts the kinematics such that the actual matrix-element value can deviate by a large factor. However, the matrix element is treated in the factorization limit and thus kept unchanged by the shower module. It is obvious that in this situation, the factorization assumption for the parton-shower approximation becomes invalid. (ii) At energy scales above the mass of the resonance, the decaying particle may be considered as a stable, colorless entity. Therefore, radiation in this region is suppressed. A full calculation would show a destructive interference of radiation from both color sources.

The `PYTHIA 6` shower module allows the programmer to mimic both effects by a variant of the shower algorithm that starts evolution at the resonance-mass scale, as opposed to the hard-process scale. Momentum is distributed such that the effective resonance mass is kept unchanged. This addresses both issues described above. There is a continuum (non-resonant) contribution to the process to which this modification would not apply, but regarding the cross section the continuum is a higher-order effect. In fact, the modified algorithm has been successful for describing real LEP data. Therefore, in the `WHIZARD` 1 simulation for e^+e^- studies, each event was assigned a resonance history, and the shower was reorganized accordingly.

However, the accuracy of both the simulation and the expected data for ILC call for a shower algorithm that goes beyond the leading-resonance approximation and also holds in the kinematical regions where neither resonance history applies, or more than one (such as W^+W^-/ZZ with identical final state). Moreover, any explicit NLO-QCD corrections are not matched correctly in the simple scheme described above. To avoid this complication, `WHIZARD` 2 initially did not account for resonance histories at all. Unfortunately, this also results in unphysical resonance-shape distortions which invalidate observables that depend on resonance shapes.

In `WHIZARD` 2.6, we have implemented an algorithm which refines the `WHIZARD` 1 variant as adapted by the experimental lepton-collider collaborations and is well suited to describe showering both on top of a resonance and in the continuum, as well as in the transition region. The effect has been validated by explicit calculation and comparison to `WHIZARD` 1 simulations.

For each given event, the generator does not just compute the complete matrix element, but also any possible factorized matrix element which results from removing all graphs that do not contain a particular set of resonances. The resonance histories are selected according to a suitable kinematics criterion.

Matrix elements that are calculated from a restricted subset of graphs, evaluated off resonance, depend on the chosen electroweak gauge. Fortunately, the resulting ambiguity is controlled. The distinction between gauge bosons of a broken symmetry and massive vector bosons can matter only for scales above the gauge boson masses. Therefore, as long as the involved off-shell distances are small compared to the corresponding resonance masses, the ambiguity is parametrically of higher order. In practice, the resonance criterion measures the distance $p^2 - m^2$ to each given resonance and compares this with $m\Gamma$, where Γ denotes the resonance
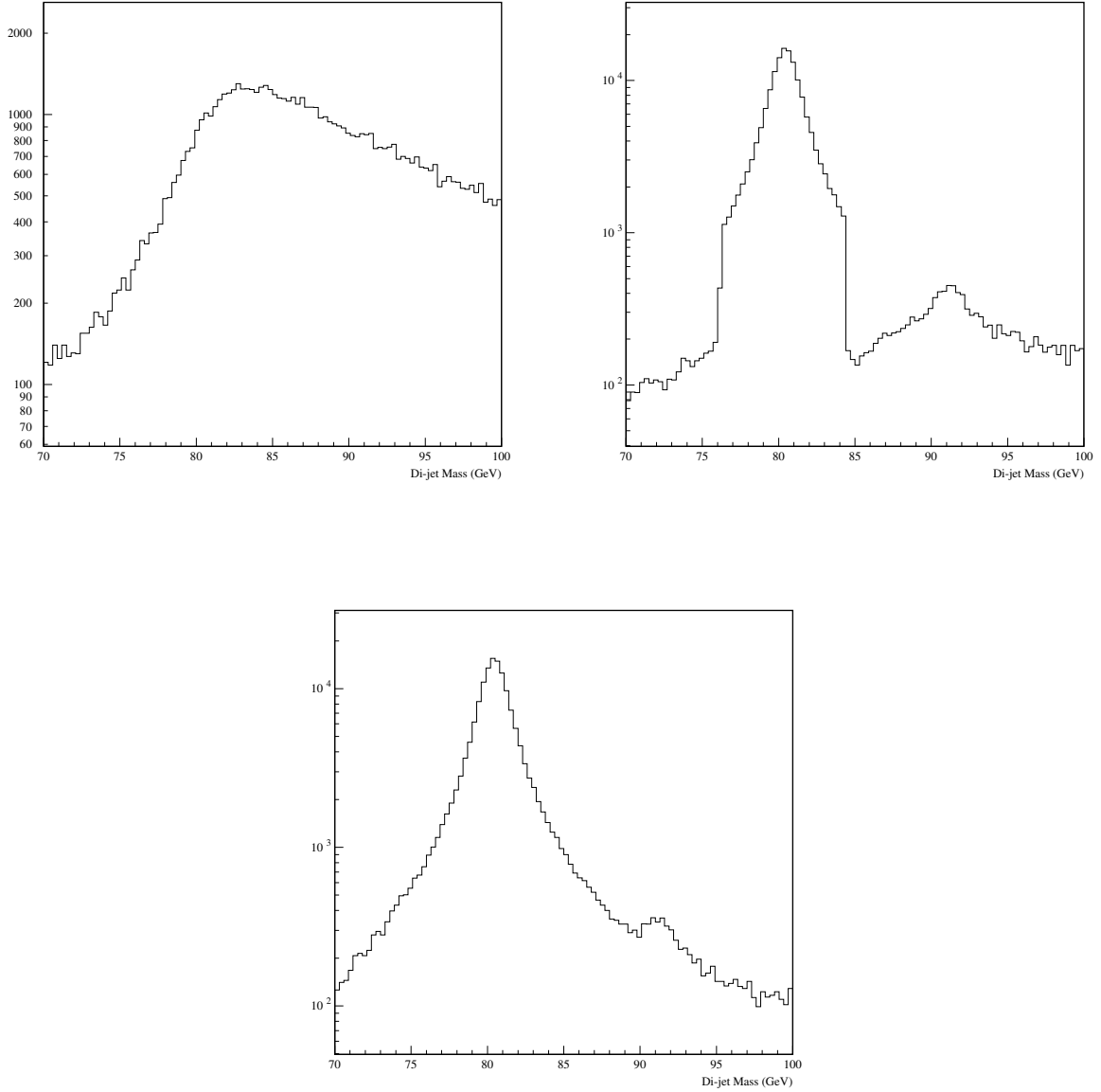
4

Figure 1: *Process $e^+e^- \rightarrow q\bar{q}q\bar{q}$ for $\sqrt{s} = 500$ GeV, with ISR and beamstrahlung. The plots display the dijet invariant mass distribution after* **PYTHIA 6** *shower for three different settings of the resonance-handler parameters. Left top: no resonance history assumed; right top: sharp resonance-history cutoff at $(p^2 - m^2)/m\Gamma = 4$; bottom: resonance-history cutoff at 16, with smooth Gaussian transition factor. Simulation and plots by M. Berggren.*

width. If this ratio exceeds a (tunable) factor, the resonance history is discarded. For the histories that are kinematically allowed, the factorized squared matrix elements are related to the complete squared matrix elements. These ratios are interpreted as probabilities pertaining to the considered histories. For each event, one of the histories, or the continuum case, is selected according to those probabilities. Finally, the selected history is reconstructed in the event record and thus transferred to the shower generator, which will restrict radiation accordingly.

As a refinement, we implement a smooth transition between resonance and continuum off shell by reweighting the factorized matrix elements with a Gaussian. This eliminates artefacts of the off-shell cutoff that otherwise would show up in exclusive distributions.

In Fig. 1, the effect is illustrated quantitatively for the process $e^+e^- \to q\bar{q}q\bar{q}$ which contains contributions from both W^+W^- and ZZ resonance intermediate states. If no resonance history is imposed on the `PYTHIA 6` shower, the generated gluon radiation completely washes out the resonance shapes of the W and Z resonance peaks at 80 and 91 GeV, respectively. If we insert the corresponding resonance histories for events where the partonic dijet invariant mass is close to a resonance peak, say for $(p^2 - m^2)/(m\Gamma) < 4$, shower evolution starts only at the weak-boson mass scale, and radiation kinematics leave the W and Z peaks intact. However, due to the smallness of the non-resonant background for this process, abruptly switching from a resonance assumption to a background assumption at a fixed distance from the peak introduces a step in the distribution as an unphysical artefact.

This problem is eliminated if we implement a smooth transition from the resonance to the background hypothesis, as shown in Fig. 1, bottom. We emphasize that all variants are formally consistent with the given order of the matrix-element calculation and of the parton shower. Nevertheless, only a smooth transition of resonance to continuum hypothesis as an input to the parton shower is expected to actually emulate the true result, which otherwise would require an explicit NLO/NNLO matching calculation.

5   Parallel Evaluation of Adaptive Phase Space

The running time of Monte-Carlo integration and event-generation programs can be reduced by a significant factor if the evaluation exploits the inherent parallelizability of evaluating a large sample of phase-space points. The new `VAMP` implementation in `WHIZARD` 2.6 aims at realizing this potential.

We have embedded the `WHIZARD` program, in particular the multi-channel integration module, in a multi-processing model according to the MPI 3 standard. The program runs on a set of computation nodes with separate associated memory and a protocol for communicating data between nodes.

Besides actually computing matrix elements on distinct phase-space points in parallel, the main issue is reducing the impact of communication, and limiting the parts of the program that have to be evaluated serially. A major communication part, within the `VAMP` adaptive algorithm, is caused by exchanging the grids, i.e., the sets of binning data for individual phase-space parameterization, between nodes. For realistic processes such as $2 \to 6$ or $2 \to 8$ configurations with nontrivial helicity and color structure, the amount of data to be exchanged can become

substantial. We are using the MPI 3 feature of asynchronous communication to minimize the amount of mutual blocking.

Further parallel speedup can be achieved if the multi-core architecture of modern CPUs as computing nodes can be exploited. To this end, `WHIZARD` enables parallel evaluation of distinct helicity configurations on a single node with multiple cores, following the OpenMP shared-memory protocol.

The MPI and OpenMP parallel features are available in the current release version of `WHIZARD` and are being used for computing-intensive studies [28]. It turns out that OpenMP parallelization scales well for standard multi-core processors, as expected given typical values for helicity combinations. The possible speedup due to multi-processor MPI parallelization is currently exhausted up to $O(10 \ldots 100)$ processors, depending on the involved process.

One of the limitations is caused by the algorithm that constructs the phase-space parameterizations, which is inherently serial. A re-implementation of this algorithm, which makes use of the process structure information which is known to the `O'Mega` matrix element generator, eliminates this problem. A systematic study of benchmark processes and the benefit of parallel evaluation is under way, and we expect further improvements in upcoming `WHIZARD` versions.

References

[1] W. Kilian, T. Ohl and J. Reuter, Eur. Phys. J. C **71**, 1742 (2011) [arXiv:0708.4233 [hep-ph]].

[2] URL: `http://whizard.hepforge.org`

[3] M. Moretti, T. Ohl and J. Reuter, hep-ph/0102195.

[4] B. Chokoufe Nejad, T. Ohl and J. Reuter, Comput. Phys. Commun. **196**, 58 (2015) [arXiv:1411.3834 [physics.comp-ph]].

[5] W. Kilian, T. Ohl, J. Reuter and C. Speckner, JHEP **1210**, 022 (2012) [arXiv:1206.3700 [hep-ph]].

[6] F. Cascioli, P. Maierhofer and S. Pozzorini, Phys. Rev. Lett. **108**, 111601 (2012) [arXiv:1111.5206 [hep-ph]].

[7] G. Cullen *et al.*, Eur. Phys. J. C **74**, no. 8, 3001 (2014) [arXiv:1404.7096 [hep-ph]].

[8] S. Actis, A. Denner, L. Hofer, J. N. Lang, A. Scharf and S. Uccirati, Comput. Phys. Commun. **214**, 140 (2017) [arXiv:1605.01090 [hep-ph]].

[9] A. Denner, J. N. Lang and S. Uccirati, arXiv:1711.07388 [hep-ph].

[10] W. Kilian, J. Reuter and T. Robens, Eur. Phys. J. C **48**, 389 (2006) [hep-ph/0607127].

[11] T. Robens, J. Kalinowski, K. Rolbiecki, W. Kilian and J. Reuter, Acta Phys. Polon. B **39**, 1705 (2008) [arXiv:0803.4161 [hep-ph]].

[12] T. Binoth, N. Greiner, A. Guffanti, J. Reuter, J.-P. Guillet and T. Reiter, Phys. Lett. B **685**, 293 (2010) [arXiv:0910.4379 [hep-ph]].

[13] N. Greiner, A. Guffanti, T. Reiter and J. Reuter, Phys. Rev. Lett. **107**, 102002 (2011) [arXiv:1105.3624 [hep-ph]].

[14] J. Reuter, B. Chokoufe, A. Hoang, W. Kilian, M. Stahlhofen, T. Teubner and C. Weiss, J. Phys. Conf. Ser. **762**, no. 1, 012059 (2016) [arXiv:1602.06270 [hep-ph]].

[15] S. Frixione, Z. Kunszt and A. Signer, Nucl. Phys. B **467**, 399 (1996) [hep-ph/9512328].

[16] T. Ohl, Comput. Phys. Commun. **101**, 269 (1997) [hep-ph/9607454].

[17] T. Sjostrand, S. Mrenna and P. Z. Skands, JHEP **0605**, 026 (2006) doi:10.1088/1126-6708/2006/05/026 [hep-ph/0603175].

[18] W. Kilian, J. Reuter, S. Schmidt and D. Wiesler, JHEP **1204**, 013 (2012) [arXiv:1112.1039 [hep-ph]].

[19] T. Ohl, Comput. Phys. Commun. **120**, 13 (1999) [hep-ph/9806432].

[20] T. Ohl and J. Reuter, Eur. Phys. J. C **30**, 525 (2003) [hep-th/0212224].

[21] F. Staub, Comput. Phys. Commun. **185**, 1773 (2014) [arXiv:1309.7223 [hep-ph]].

[22] N. D. Christensen, C. Duhr, B. Fuks, J. Reuter and C. Speckner, Eur. Phys. J. C **72**, 1990 (2012) [arXiv:1010.3251 [hep-ph]].

[23] A. Alloul, N. D. Christensen, C. Degrande, C. Duhr and B. Fuks, Comput. Phys. Commun. **185**, 2250 (2014) [arXiv:1310.1921 [hep-ph]].

[24] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer and T. Reiter, Comput. Phys. Commun. **183**, 1201 (2012) [arXiv:1108.2040 [hep-ph]].

[25] M. Cacciari, G. P. Salam and G. Soyez, Eur. Phys. J. C **72** (2012) 1896 [arXiv:1111.6097 [hep-ph]].

[26] B. Chokoufé Nejad, W. Kilian, J. M. Lindert, S. Pozzorini, J. Reuter and C. Weiss, JHEP **1612**, 075 (2016) [arXiv:1609.03390 [hep-ph]].

[27] F. Bach, B. Chokoufé Nejad, A. Hoang, W. Kilian, J. Reuter, M. Stahlhofen, T. Teubner and C. Weiss, arXiv:1712.02220 [hep-ph].

[28] C. F. Anders *et al.*, arXiv:1801.04203 [hep-ph].