

Monte Carlo methods for TMD analyses

Gunar Schnell^{1,2,a}

¹*Department of Theoretical Physics, University of the Basque Country UPV/EHU, 48080 Bilbao, Spain*

²*IKERBASQUE, Basque Foundation for Science, 48013 Bilbao, Spain*

Abstract. Monte Carlo simulations are an indispensable tool in experimental high-energy physics. Indeed, many discoveries rely on realistic modeling of background processes. In the field of transverse-momentum-dependent parton distribution and fragmentation functions there is a clear lack of a reliable Monte Carlo physics generator that can be used in experimental and phenomenological analyses. The need for such Monte Carlo generators, the status of some solutions and prospects are discussed.

1 Introduction

One key ingredient to modern experimental analyses, but also phenomenology, is the employment of Monte Carlo physics and detector simulations. Many physics generators have emerged over the past decades, some rather specific, others as broad as possible. Little has been available, however, in the field of transverse-momentum-dependent (TMD) parton distribution functions (PDFs) and fragmentation functions (FFs). One of the earlier efforts has focused on the unintegrated gluon distributions at (very) low x [1], not (yet) relevant to most of the work discussed at this workshop. Another example, GMC_TRANS [2], is based on a Gaussian ansatz of the transverse-momentum dependence of those (leading-twist) TMD PDFs and FFs relevant in semi-inclusive deep-inelastic scattering off transversely polarized nucleons. It was mainly used as a tool for systematic studies by the HERMES collaboration, but has evolved to one of the presently more versatile TMD physics generator. A Monte Carlo generator specifically for TMD FFs has been the focus of the work by Matevosyan and collaborators [3], which led to promising results in the transverse-momentum dependence of unpolarized fragmentation functions, but also for the Collins function. Nevertheless, none of the presently available Monte Carlo physics generators is able to take a role in TMD PDF and FF experiments and phenomenology, as for instance the widely used PYTHIA [4] code does in high-energy physics. The latter generator, however, does not include any spin-momentum correlations and thus can not take the role of a TMD PDF and FF generator to be used for the various TMD processes and initial- and final-state polarization possibilities.

In this contribution the merits of Monte Carlo simulations are discussed, with special focus of the necessity to test measurement methods in TMD experiments, but also

the possibility of using Monte Carlo generators for phenomenology.

1.1 General usage and types of Monte Carlo generators

The present sophistication and complexity of the field requires thorough planning when designing new experiments in order to achieve the highest possible precision not only on the side of statistics but much more on systematics. Indeed, many measurements involving multi-particle final states are dominated already now by systematics and not statistical precision. It thus becomes obligatory to present predictions on the experiment's envisaged performance when it comes to the various measurements proposed. For this a realistic simulation of both the physics and the detector response is required. While for the latter a rather sophisticated work package is available in form of GEANT [5], the simulation of the physics signatures requires specific Monte Carlo generators for the various processes. Those can be fed with model predictions and/or fits of PDFs and FFs from previous experiments if the underlying physics process is implemented. Especially in the field of TMD PDFs and FFs the relevant quantities are often unknown and thus difficult to model reliably, making such predictions susceptible to large uncertainties.

Another important field for Monte Carlo simulations is the one of known, e.g., QED, corrections. In deep-inelastic scattering (DIS) those include the typical radiative corrections like emission of real photons from the lepton lines or loop/vertex corrections. For this purpose *external* and *integrated* programs exist, i.e., software packages that are used to correct the measurement *after* correcting for all other experimental effects, or that are integral part of the overall correction procedure, respectively. While mostly giving similar results one should note that radiative corrections and typical experimental effects like acceptance and efficiency corrections do not factorize in general [6].

^ae-mail: gunar.schnell@desy.de

It is thus preferential to include radiative corrections in the simulation from the very beginning, if available.

In general terms Monte Carlo simulations are a powerful tool to develop the methodology to be applied in the measurement. Indeed, radiative corrections might already be one part, but acceptance and efficiency corrections as well as detector smearing in high-precision measurements require sophisticated analysis methods that need to be developed and tested. This can be done best with reliable pseudo-data that comes close to reproducing the actual physics distributions, which again requires employing Monte Carlo physics generators. This equally applies then also to the estimation of systematic uncertainties like acceptance effects. It is close to useless to report measurements in the experimental acceptance if the effect of such limited phase space does not get quantified. Again, such evaluation requires reliable reproduction of the physics process that has to be separated from instrumental effects in the actual measurement.

Last but not least Monte Carlo generators can also be used to extract physics from actual measurements. Parameters of PDF or FF parametrizations can be determined from measurements that involve convoluted quantities by variation of those parameters in the process of reproducing real data in the corresponding Monte Carlo simulation, as done in the tuning process of most Monte Carlo generators (see, e.g., Ref. [7]).

For these tasks there exist basically three different types of Monte Carlo generators. The simplest ones are the *toy* generators, often just sampling the available phase space. Still they can give a pretty good idea of how appropriate a certain setup or methodology is for a specific measurement. However, they are not able to take into account the variety of correlations encountered in most and in particular in TMD PDF and FF related measurements. To avoid overly simplified situations and in return often unrealistic systematics for high-precision experiments *inclusive* and *full-even* generators are required. Only those can disentangle, if at all, the convoluted effects of physics distributions and phase-space limitations imposed by virtually every experimental setup (cf. section 2). While inclusive generators concentrate on the final-state particles of interest, e.g., the scattered lepton in inclusive DIS or in addition one coincident hadron in single-hadron semi-inclusive DIS, full-event generators simulate the entire event, e.g., also the spectator particles in the reaction. While likely not of interest for the physics process under investigation the latter potentially influence the measurement through their correlation with the interesting particles. Among such effects could be trigger biases or particle (mis)identification due to limitations of the apparatus. The advantage of inclusive generators is the applicability of simple parameterizations of the process, e.g. the inclusive DIS cross section can be modeled entirely using its pQCD description together with appropriate PDFs, and that in principle to any order of QCD and QED. For event generators such a pQCD description does not exist. The typical approach is to then generate the corresponding kinematics of the scattered lepton (in case of

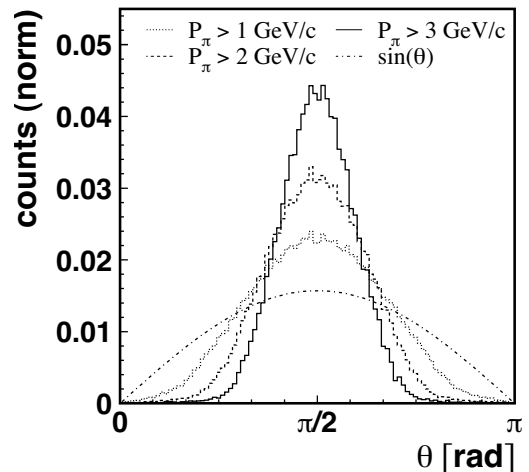


Figure 1. Effect of a momentum cut on the angular distribution of hadrons in di-hadron semi-inclusive DIS. Here, θ represents the polar angle of hadron 1 in the two-hadron center-of-mass system with respect to their combined momentum (boost axis). The distribution is plotted for several choices for the minimum value accepted for P_π , the hadron (pion) momentum in the laboratory system. The *physical distribution* $\sin(\theta)$ gets visually distorted even for the least severe momentum minimum of 1 GeV [11].

lepton-nucleon reactions) from the pQCD description of DIS while leaving the formation of the final hadronic state to some empirical model like JETSET [8]. Naturally the latter require paramount work in tuning the multitude of parameters that control the generation of such involved final states, all subject to verification at each experiment.

2 Detector effects in measurements

In the field of TMD PDFs and FFs one has to state that **no particle-physics experiment has a perfect acceptance**. Even so-called 4π detectors “enjoy” openings for the beam or employ momentum constraints, e.g., the process of particle identification. All those correspond to more or less severe constraints on the phase space sampled and thus lead to potential acceptance effects in the final-particle spectra. As an example, the angular distribution in θ in di-hadron lepto-production (which, for the case of transversely polarized targets, corresponds to a nine-dimensional differential cross section), corresponding to the polar angle in the two-hadron center of mass, is presented in Fig. 1. While following a $\sin\theta$ distribution for the undisturbed case it exhibits a clearly peaking structure when momentum constraints are placed on the final-state hadrons. Other kinds of angular distortions are observed even for the much easier case of single-hadron lepto-production where the angular distribution of the final-state hadron about the virtual photon direction gets affected by either the outer edges or the beam hole(s) of the detector (cf. figure 6 of Ref. [9] or the discussion in Ref. [10]).

How the resulting acceptance effects are handled in the actual measurement is one of the essential questions

in experiments. Distorted angular acceptances can correlate formally orthogonal modulations, e.g., the Fourier decomposition of the single-hadron semi-inclusive DIS cross section, where particular structure functions with distinct Fourier signatures cannot be distinguished any longer when the angular acceptance becomes too limited. But also the integration over other kinematic variables that are affected by the acceptance will result in a wrong statistical weighting of the different phase-space regions, e.g., in the measurement of the asymmetry A integrated over the kinematic phase space Ω restricted by the detection efficiency ϵ :

$$\langle A \rangle_\epsilon \equiv \frac{\int d\Omega A(\Omega) \epsilon(\Omega) \sigma_0(\Omega)}{\int d\Omega \epsilon(\Omega) \sigma_0(\Omega)} \quad (1)$$

$$\neq \frac{\int d\Omega A(\Omega) \sigma_0(\Omega)}{\int d\Omega \sigma_0(\Omega)} \equiv \langle A \rangle_{4\pi}, \quad (2)$$

where $\langle A \rangle_\epsilon$ corresponds to the observed average asymmetry, $\langle A \rangle_{4\pi}$ is the average asymmetry expected for a perfect 4π acceptance, and σ_0 being the unpolarized cross section. This in return means that **acceptance and efficiencies do not cancel in general in asymmetry measurements!** It should be noted that any attempt to extract the corresponding correction factors from Monte Carlo simulations of the detector will almost certainly fail to a certain degree if the Monte Carlo generator employed does not fully describe the underlying unpolarized cross section or if the correction is not done at least fully differential. For asymmetry measurements one is somewhat safe as long as the asymmetries do not exhibit a significant non-linear behavior (or the same for the acceptance function) as then the measured average asymmetry corresponds to the actual asymmetry at the average measured kinematics:

$$\langle A(\Omega) \rangle_\epsilon \equiv \int d\Omega A(\Omega) \epsilon(\Omega) \quad (3)$$

$$= \int d\Omega [A_0 + A_1 \Omega] \epsilon(\Omega) \quad (4)$$

$$= A_0 + A_1 \int d\Omega \Omega \epsilon(\Omega) \quad (5)$$

$$= A_0 + A_1 \langle \Omega \rangle_\epsilon \equiv A(\langle \Omega \rangle_\epsilon), \quad (6)$$

with A_0 (A_1) being the constant (linear) part in Ω of the physical asymmetry. This observation is also the reason behind the recommendation of going as differential as possible in the analysis as for small enough integration regions almost every asymmetry can be approximated by a linear behavior. However, this is in general conflict with the limited available statistics and bin migration due to smearing that ultimately prevail over acceptance effects when going fully differential.

The way out of this dilemma is the usage of reliable Monte Carlo simulations that can be used on one side to evaluate the level of migration in order to subsequently undo the effect by a procedure known as *unfolding* (cf. discussion in Ref. [12]). On the other side the size of disagreement between Eq. (1) and Eq. (2) can in principle be quantified if in the Monte Carlo simulation the shape of

the asymmetry and unpolarized cross section can be reproduced to a sufficient level, either by successive tuning or by employing a data-driven model for the unknown asymmetry, both to be discussed in the following.

3 Monte Carlo generators in TMD analyses

It is unfortunately true that in the field of TMD PDFs and FFs there is a clear lack of sophisticated Monte Carlo generators. Two examples how this situation is dealt with are discussed here. One approach is establishing a new Monte Carlo generator that incorporates the framework of TMD PDFs and FFs. Another possibility, more widely used, is a reweighting (reshuffling) approach in using the PYTHIA event generator to implement in the initially spin-independent Monte Carlo simulation spin effects.

3.1 The TMD generator GMC_TRANS

The HERMES Collaboration developed a Monte Carlo generator that starts from the parton-model expression of the one-hadron semi-inclusive DIS cross section, using several models/parametrization for the various TMD PDFs and FFs. Initially conceived for semi-inclusive pion production on a transversely polarized target with only the Sivers and Collins effect included, several other TMD PDFs have been added as well as charged-kaon production, both for proton and neutron targets (or combinations thereof). In order to be fast an analytic expression for the semi-inclusive DIS cross section was implemented based on the widely used Gaussian ansatz of the transverse-momentum dependences. Direct comparison of the input (model) asymmetry and the measured one allowed a fast computation of the effects of limited phase space due to the detector acceptance. This was facilitated by providing a direct interface to the standard HERMES Monte Carlo package.

More or less obvious short-comings of this approach were the limitation to semi-inclusive one-hadron production, e.g., the absence of event information and thus correlated spectator particles, the entire lack of radiative corrections, the incomplete implementation of the full leading-twist contributions to the cross section, and a missing concept of implementing reliably sub-leading twist contributions (positivity violations being the “biggest nuisance”).

Despite those limitations GMC_TRANS has proven to be very useful in estimates of systematics and in the study of TMD PDFs and FFs. In Fig. 2 an extraction of the Gaussian width of both the unpolarized TMD PDF and FF is illustrated. In the Gaussian approach the resulting transverse-momentum distribution of the produced hadrons (here pions) can be attributed in a factorized way to both the intrinsic transverse momentum of quarks in the nucleon, $\langle p_T^2 \rangle$, and the fragmentation transverse momentum, $\langle K_T^2(z) \rangle$, via

$$\langle P_{h,\perp}^2(z) \rangle = z^2 \langle p_T^2 \rangle + \langle K_T^2(z) \rangle. \quad (7)$$

By fitting the z^2 dependence of the average hadron transverse momentum, where z denotes the target-rest-frame

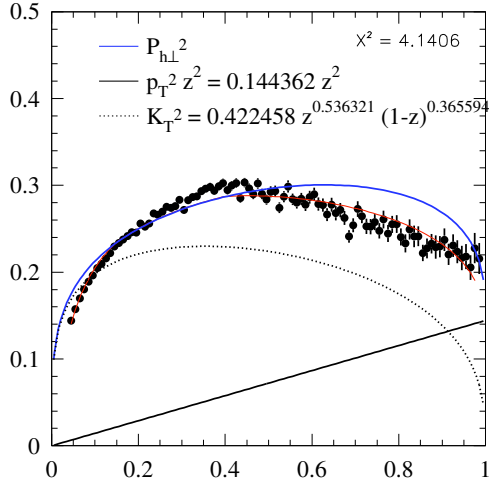


Figure 2. Tentative extraction of intrinsic and fragmentation transverse momentum from the transverse momentum distribution of pions at the HERMES experiment using the Monte Carlo generator GMC_TRANS. Plotted are, as a function of z^2 the average transverse momenta of pions in HERMES acceptance for HERMES data (points) and the corresponding GMC_TRANS simulation (red line) using as input parameters the black full and dotted lines for the intrinsic transverse momentum $\langle p_T^2 \rangle$, and the fragmentation transverse momentum $\langle K_T^2(z) \rangle$, respectively. The blue line is the transverse-momentum width expected for a perfect detector.

energy fraction of the virtual photon carried by the detected hadron, both $\langle p_T^2 \rangle$ and $\langle K_T^2(z) \rangle$ were extracted. The resulting *Hashi* parametrization clearly favored a z dependence of the fragmentation transverse momentum (a linear behavior of $\langle P_{h\perp}^2 \rangle$ with z^2 should have been observed otherwise):

$$\langle p_T^2 \rangle \stackrel{\text{“Hashi set”}}{=} 0.144362 \text{ GeV}^2 \quad (8)$$

$$\langle K_T^2(z) \rangle \stackrel{\text{“Hashi set”}}{=} 0.422458 z^{0.536321} (1-z)^{0.365594} \text{ GeV}^2. \quad (9)$$

The left panel of Fig. 3 shows a comparison of Sivers azimuthal asymmetries from a GMC_TRANS simulation and the ones extracted from HERMES data (both in HERMES acceptance) for charged pions and kaons. The “DSS” fragmentation function set [13] is used together with a slightly rescaled early parametrization of the Sivers function [14]. A rather satisfactory description of the charged pions and positive kaons could be achieved, enough to be used to validate the methodology used at HERMES for extracting both the Collins and Sivers asymmetries at the then available statistical precision [15]. This is demonstrated on the right side where the simulated asymmetries are shown in “ 4π ” acceptance and folded with the HERMES acceptance, with only small differences between the two.

3.2 “Polarizing” PYTHIA

A different approach makes use of an already very good description of the spin-independent semi-inclusive DIS cross section provided by PYTHIA. As PYTHIA events come with event weights equal to unity they are easy to reweigh or to reshuffle. This can be exploited to introduce spin dependence in the otherwise spin-independent event generator. By throwing a random number ρ (between zero and one) a polarization state P can be assigned to each event i according to a model of the spin asymmetry of interest, e.g.,

$$\rho < \frac{1}{2} \left[1 + A_{UT}^{\sin(\phi-\phi_S)}(\Omega^i) \sin(\phi^i - \phi_S^i) \right] \Rightarrow P = +1 \quad (10)$$

$$\rho > \frac{1}{2} \left[1 + A_{UT}^{\sin(\phi-\phi_S)}(\Omega^i) \sin(\phi^i - \phi_S^i) \right] \Rightarrow P = -1 \quad (11)$$

in case of the Sivers azimuthal asymmetry, where $(\Omega^i, \phi^i, \phi_S^i)$ are the fully differential *true* event kinematics for that particular event and $A_{UT}^{\sin(\phi-\phi_S)}$ is a parametrization of choice for the Sivers asymmetry. Equations (10) and (11) can be generalized to include further modulations including double-spin asymmetries.

The power of this approach is that virtually any parametrization of the spin dependence can be implemented (as long as fulfilling positivity constraints) without limiting oneself to, e.g., the Gaussian ansatz. One can also study hadrons for which FFs are not yet available. In addition, the full event will remain available, allowing a more thorough study of systematics due to detector responses. Last but not least, radiative corrections can be used, e.g., the RADGEN package [16] together with PYTHIA. This allows more precise determinations of smearing effects by not only including external radiation but also QED corrections to the process of interest.

Not being limited to existing parametrizations and models opens the door to a data-driven approach in polarizing PYTHIA. An approximate model of reality can be obtained by fitting a Taylor series in all kinematic variables to the various spin asymmetries measured. A maximum likelihood fit can be used to extract the coefficients of the fully differential (truncated) Taylor series for every single azimuthal amplitude appearing in the cross section. At HERMES several such expansions have been examined, e.g., the 22-parameter fit (for both the Collins and Sivers amplitudes) presented in Fig. 4.

Such a parametrization can then be used to assign spin states to a PYTHIA Monte Carlo simulation according to Eqs. (10) and (11) (with the proper inclusion of also the Collins effect). The resulting amplitudes are compared to the actual HERMES data in Fig. 5 for both the Collins and Sivers modulations, and a rather good description could be achieved. Limitation stemming from the truncation of the Taylor series can be spotted, e.g., for the x dependence of the Sivers K^+ asymmetry amplitudes. While not a principle problem including additional terms can become a practical problem, especially when attempting to parameterize

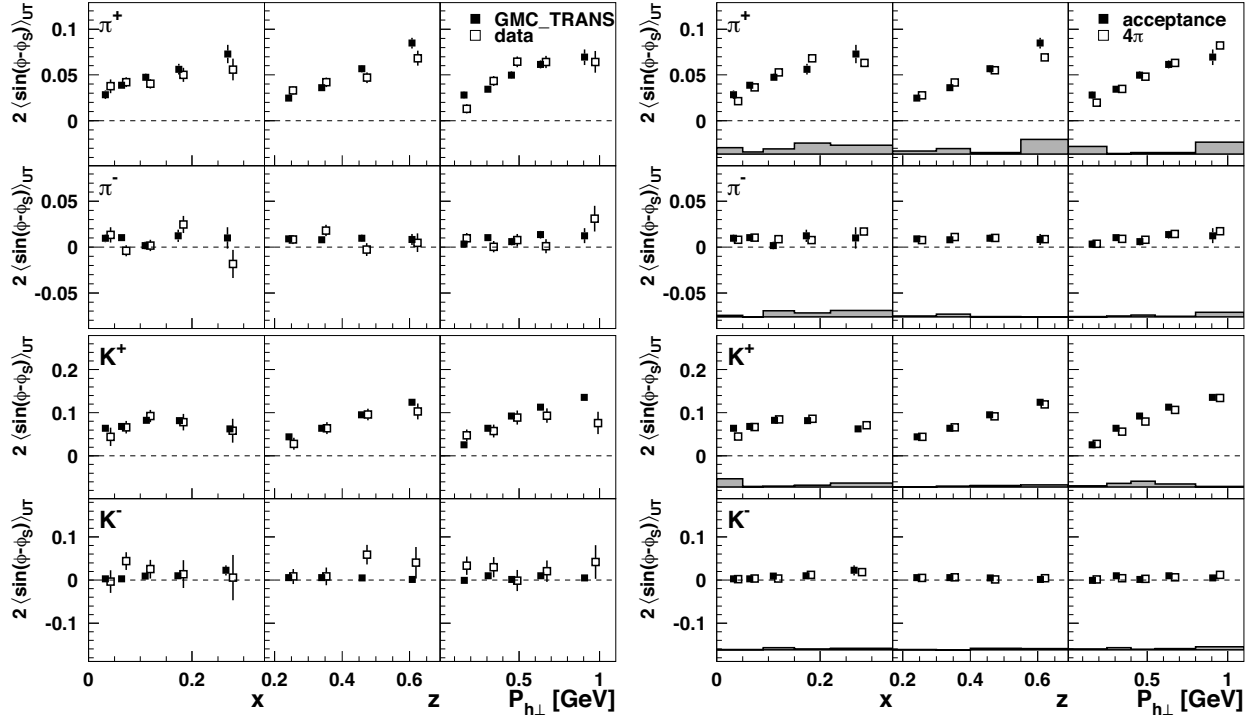


Figure 3. Comparison of a GMC_TRANS simulation of the Sivers effect for charged pions and kaons with the actual measurement, both in HERMES acceptance (left), and of the GMC_TRANS simulation in “4π” and in HERMES acceptance (right).

$$\begin{aligned}
 \overline{[1]}_{22} \sin(\phi \pm \phi_S), h &= \overline{[1]}_{22,1} \sin(\phi \pm \phi_S), h &+ \overline{[1]}_{22,2} \sin(\phi \pm \phi_S), h x' &+ \\
 &\overline{[1]}_{22,3} \sin(\phi \pm \phi_S), h Q^{2'} &+ \overline{[1]}_{22,4} \sin(\phi \pm \phi_S), h z' &+ \\
 &\overline{[1]}_{22,5} \sin(\phi \pm \phi_S), h |\mathbf{P}_{h\perp}|' &+ \overline{[1]}_{22,6} \sin(\phi \pm \phi_S), h x'^2 &+ \\
 &\overline{[1]}_{22,7} \sin(\phi \pm \phi_S), h z'^2 &+ \overline{[1]}_{22,8} \sin(\phi \pm \phi_S), h |\mathbf{P}_{h\perp}|'^2 &+ \\
 &\overline{[1]}_{22,9} \sin(\phi \pm \phi_S), h x' z' &+ \overline{[1]}_{22,10} \sin(\phi \pm \phi_S), h x' |\mathbf{P}_{h\perp}|' &+ \\
 &\overline{[1]}_{22,11} \sin(\phi \pm \phi_S), h z' |\mathbf{P}_{h\perp}|' &+ \overline{[1]}_{22,12} \sin(\phi \pm \phi_S), h x'^3 &+ \\
 &\overline{[1]}_{22,13} \sin(\phi \pm \phi_S), h x' z'^2 &+ \overline{[1]}_{22,14} \sin(\phi \pm \phi_S), h x'^2 z' &+ \\
 &\overline{[1]}_{22,15} \sin(\phi \pm \phi_S), h x'^2 |\mathbf{P}_{h\perp}|' &+ \overline{[1]}_{22,16} \sin(\phi \pm \phi_S), h x' |\mathbf{P}_{h\perp}|'^2 &+ \\
 &\overline{[1]}_{22,17} \sin(\phi \pm \phi_S), h z'^2 |\mathbf{P}_{h\perp}|' &+ \overline{[1]}_{22,18} \sin(\phi \pm \phi_S), h z' |\mathbf{P}_{h\perp}|'^2 &+ \\
 &\overline{[1]}_{22,19} \sin(\phi \pm \phi_S), h x'^2 |\mathbf{P}_{h\perp}|'^2 &+ \overline{[1]}_{22,20} \sin(\phi \pm \phi_S), h z'^2 |\mathbf{P}_{h\perp}|'^2 &+ \\
 &\overline{[1]}_{22,21} \sin(\phi \pm \phi_S), h x' z' |\mathbf{P}_{h\perp}|' &+ \overline{[1]}_{22,22} \sin(\phi \pm \phi_S), h x'^2 z' |\mathbf{P}_{h\perp}|' &.
 \end{aligned}$$

Figure 4. The 22-parameter Taylor expansion used for the description of HERMES data on the Collins and Sivers effects in Fig. 5 [18].

all spin-dependent terms in the semi-inclusive DIS cross section. Already for the case of including the Sivers and Collins modulations only the expansion in Fig. 4 leads to 44 parameters to be fitted, which is close to the usual limit of, e.g., standard MINUIT [17], on how many parameters can be determined simultaneously. Theory guidance might be used to concentrate on the seemingly relevant terms in the expansion.

The approach of assigning polarization states in PYTHIA becomes even more powerful for situations where no rigorous pQCD framework with analytic expressions is available. As long as PYTHIA (or any other suitable Monte Carlo event generator) is able to empirically describe the spin-independent cross section in a satisfactory way it can then be used to also simulate spin dependence. One example is the transverse single-spin asymmetry in inclusive hadron lepto-production, $ep^\uparrow \rightarrow hX$. This

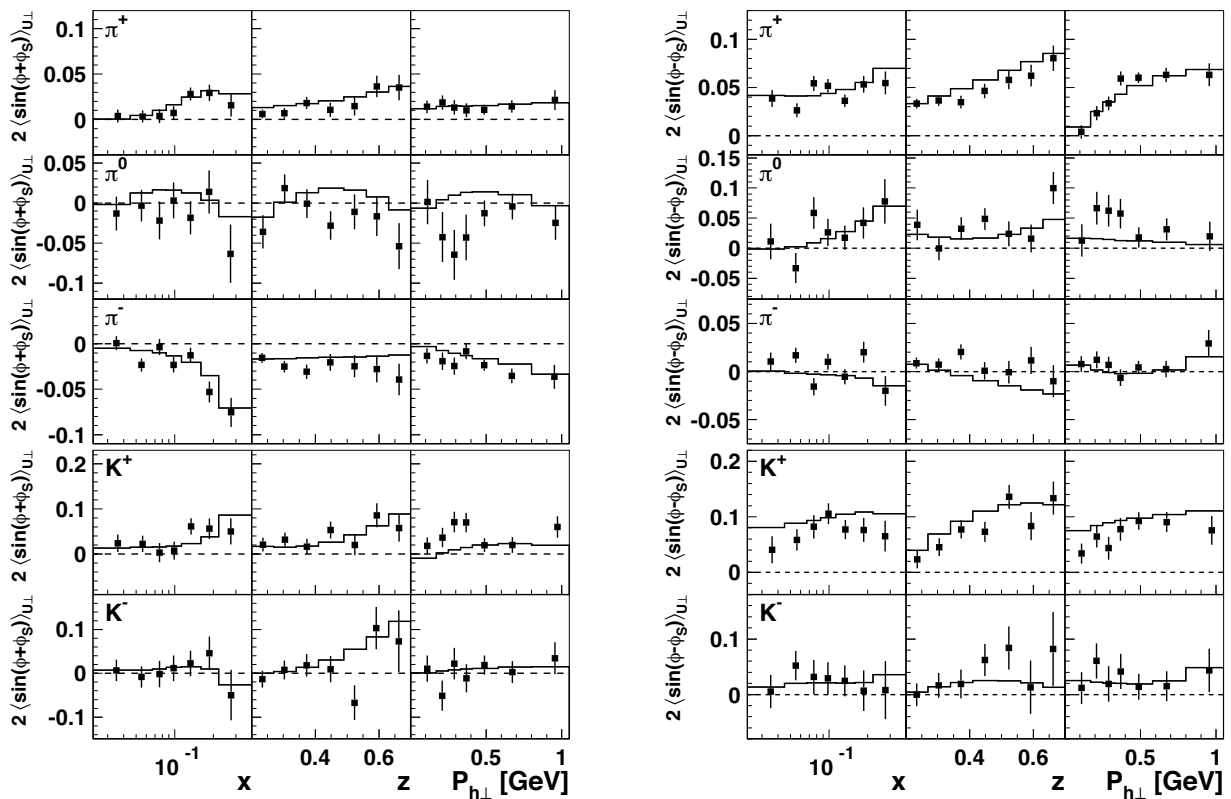


Figure 5. Comparison of the “polarized PYTHIA” simulation (histogram) with HERMES data (data points) in the HERMES acceptance for both Collins (left) and Sivers amplitudes (right) [18].

process to date can not be factorized in terms of PDFs and FFs, thus an empirical model must be employed. On the positive side, this process depends on two kinematic variables only, e.g., Feynman- x and the transverse momentum p_T of the outgoing hadron with respect to the lepton direction, besides the azimuthal angle ψ of the hadron production plane about the lepton-beam axis. Such measurement was done, e.g., by the HERMES [19] and Jefferson Lab Hall A [20] Collaborations, and the resulting parametrizations of the HERMES data for charged pions and kaons are presented in Fig. 6.

Figure 7 illustrates the subsequent extraction of systematics due to detector effects. The “polarized PYTHIA” events can be tracked through a realistic simulation of the respective detector and analyzed in the same way as normal experimental data. The reconstructed asymmetry amplitudes, plotted as points in Fig. 7, can then be compared to the amplitudes expected for a “perfect” detector, i.e., the asymmetry model/parametrization evaluated at the mean reconstructed kinematics in each experimental bin (shown as a line).¹ The difference of the two stem from detector effects like smearing but more importantly on the right side of Fig. 7 from the integration over one of the two kinematic variables. As the asymmetries exhibit a strongly non-linear p_T dependence, the difference between Eq. (1) and Eq. (2) can become sizable in view of limited acceptance

¹The latter is the same as how the data are usually used in fits, e.g., they are interpreted as the true value of the observable for the average kinematics given alongside.

in p_T when integrating over p_T . The two-dimensional, i.e., fully differential, presentation in Ref. [19] does not suffer from such short-coming and is in general the preferred presentation. Also the p_T dependence in the one-dimensional analysis on the left of Fig. 7 can be reconstructed to an excellent level due to the rather weak dependence of the asymmetries on x_F , e.g., due to the correspondences in Eqs. (3)-(6) for up to linear behaviors.

These examples show how powerful the method of assigning spin states to a spin-independent Monte Carlo generator like PYTHIA can be. The main advantages lie in the flexibility and possibility to apply it even to processes where no guidance from theory is available on the shape and magnitude of the effect of interest. Moreover, the whole event topology and track correlations are available when using a full-event generator. However, the method is not without drawbacks. Obviously, one needs a parametrization for the asymmetry. It could be based on model calculations, global analyses, but also to fit to the actual data. When using such data-driven approach, e.g., when no other measurement of that observable has yet become available, the question arises where to stop the Taylor expansion and whether or not one can reliably determine the parameters of all terms in the expansion. The maximum likelihood fit to real data already has folded in effects from smearing, which poses serious limitations on how precise the true physics can be reproduced by the Taylor expansion. It also leaves the question of the goodness of the actual fit. In praxis limitations already occur earlier,

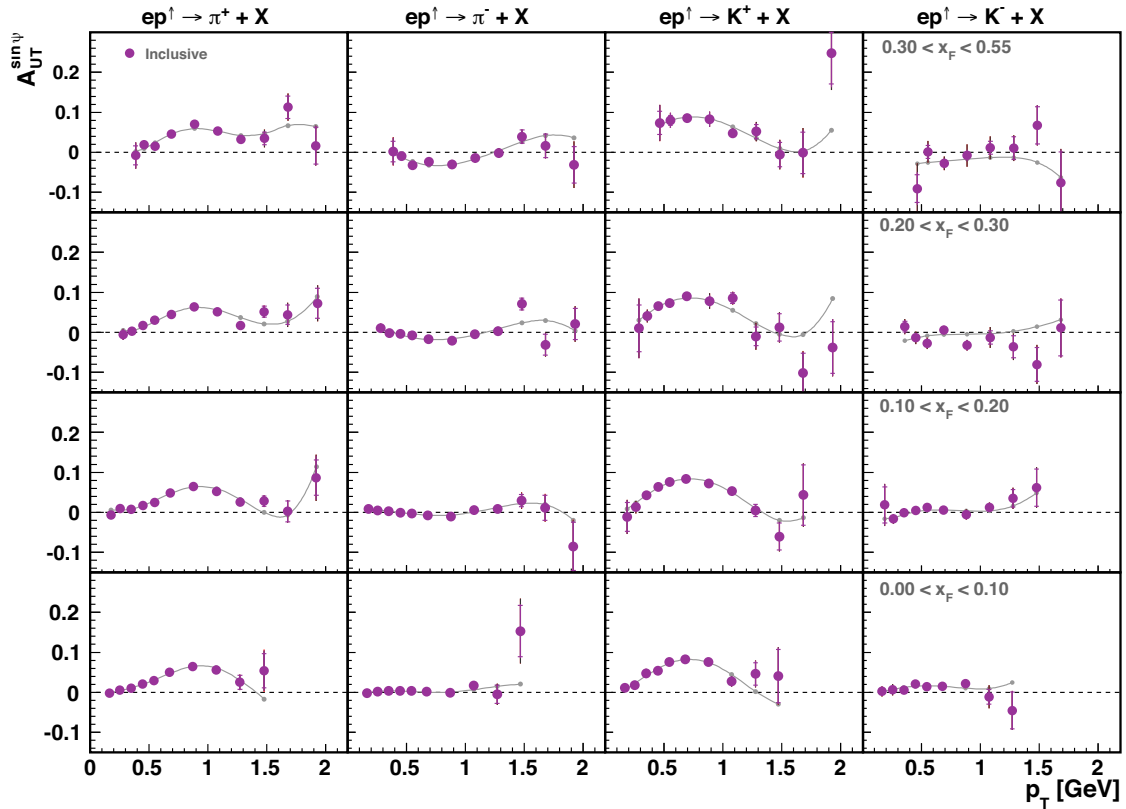


Figure 6. Comparison of a “polarized PYTHIA” simulation (line) with HERMES data (data points) in the HERMES acceptance for inclusive hadron lepto-production from a transversely polarized proton target, as a function of p_T in various slices of x_F .

namely when hitting the limit on the number of parameters to fit. These problems aside, there is one important ingredient that is often difficult to handle. The spin-independent cross section must be well reproduced in the Monte Carlo simulation in order to get a reliable estimate of systematic uncertainties. This is not a priori given for every physics process of interest and must be checked first. And this is the more crucial the further the observable is out of the usual focus of, e.g., the PYTHIA developers.

4 Correcting for event migration

One of the challenges for high-statistics measurements, especially for detectors with almost full acceptance, which suffer less from acceptance effects, is the effect of event migration. The limited resolution of any detection device leads to a mis-reconstruction of the actual kinematics. This can on one side wash out effects. On the other side, when folded with steep kinematic dependences, e.g., the z dependence of FFs or the Q^2 dependence of the DIS cross section, distributions can become heavily distorted as migration becomes prevalently uni-directional. The correction of such effects relies on a precise simulation of the detector response folded with a good description of the physics process in the event generator, and is commonly denoted as *unfolding* [22]. Schematically the migration

problem can be cast in the following form:

$$\mathcal{Y}^{\text{exp}}(\Omega_i) \propto \sum_{j=1}^N S_{ij} \int d\Omega d\sigma(\Omega) + \mathcal{B}(\Omega_i). \quad (12)$$

The reconstructed yield \mathcal{Y}^{exp} in a kinematical bin i is proportional to the *Born* cross section in all experimental bins j that get smeared into bin i (expressed by the smearing matrix S_{ij}) plus additional *background* \mathcal{B} smeared into bin i from outside the experimental kinematic acceptance. Inverting the relation gives, in principle, the wanted Born cross section, once both the background contribution and the smearing matrix is known. The latter can be obtained in an almost model-independent way from a Monte Carlo simulation by basically dividing reconstructed Monte Carlo yields by the corresponding Monte Carlo Born distribution. In the limit of infinitesimally small bins (and fully differential) and/or flat acceptance or cross-section in each bin, the underlying model cross section drops out in the evaluation of the smearing matrix, in principle. In real life a residual dependence on the model survives due to finite bin sizes, but more importantly due to the experimentally inaccessible background contribution. The latter must come entirely from the Monte Carlo simulation. It is thus of importance to have a reliable description of the underlying physics implemented in the Monte Carlo generator. It should be stressed also here that the applicability of the unfolding procedure relies on being fully differential. The more kinematic variables the smearing

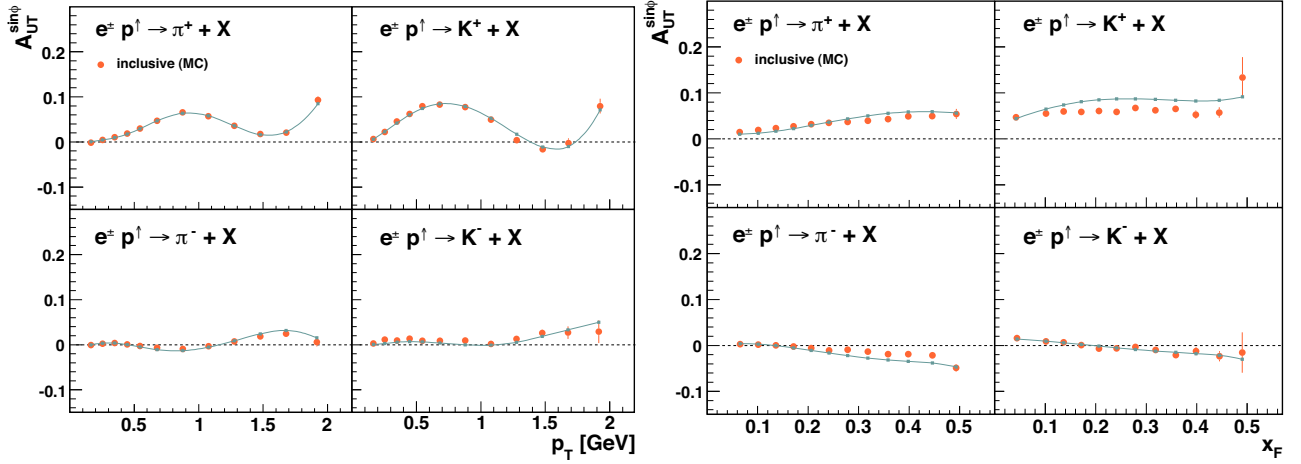


Figure 7. Comparison of the parametrization for inclusive hadron production based on HERMES data, evaluated at the mean reconstructed kinematics in each bin (line), with the reconstructed asymmetries in HERMES acceptance using a PYTHIA simulation “polarized” using the same parametrization evaluated at the true kinematics (points). Shown are the one-dimensional projections, as a function of p_T (left) and of x_F (right) [21].

correction is integrated over the stronger the smearing correction becomes. Even in the case of vanishing smearing, e.g., a reduction of Eq. (12) to a simple diagonal problem, i.e., $S_{ij} \equiv S_i \delta_{ij}$, those diagonal elements can be determined from Monte Carlo model-independently only in case of a fully differential analysis, e.g., not when determining a ϕ correction factor for azimuthal modulations in each ϕ bin independent from the other kinematics (or integrated over at least some of them):

$$\frac{\int d\Omega \epsilon(\Omega, \phi_i) \sigma_{MC}(\Omega, \phi_i)}{\int d\Omega \sigma_{MC}(\Omega, \phi_i)} \neq \frac{\int d\Omega \epsilon(\Omega, \phi_i) \sigma_{true}(\Omega, \phi_i)}{\int d\Omega \sigma_{true}(\Omega, \phi_i)} \neq \int d\Omega \epsilon(\Omega, \phi_i).$$

5 Conclusion

While Monte Carlo generators have been an integral part of nuclear and high-energy physics the status of dedicated generators for TMD physics is poor. No general-purpose generator that incorporates spin-dependence and TMD PDFs and FFs (in a controllable way) is currently available. For the limited precision presently available for most TMD PDF and FF related measurements in lepton-nucleon scattering this might still be acceptable, but in view of the high-statistics measurements planned within the JLab12 program (see corresponding talks at this workshop) or even at a future polarized electron-ion collider this lack might be considered worrisome. Already for the presently available statistics for TMD FF measurements at the $e^+e^- B$ factories, the existing generators may determine the limits on how precise the measurements can be performed. For the meantime introducing spin-dependence into PYTHIA as discussed in Section 3.2 might be a viable option. However, in this respect the dropping of lepton-nucleon scattering by the PYTHIA developers in the current version and missing plans of in-

cluding it in future versions will not make it easier to keep up with the latest developments. It thus falls back to the limited individual efforts of *patching* PYTHIA or alike to include the physics of TMD PDFs and FFs so enthusiastically discussed at this workshop. For the moment such efforts will mainly serve experimentalists in their effort of controlling systematics, proper handling of such issues like evolution etc. will require a much more dedicated effort. Reasons for hope are the dedicated INT workshop hold on the topic of Monte Carlo generators this year [23], but also the work package on Monte Carlo generators for future DIS facilities submitted within the European Commission HadronPhysicsHORIZON proposal [24], which should raise the awareness of the necessity of dedicated Monte Carlo efforts in the field of TMD PDFs and FFs. In that respect, also the ongoing work on the TMD library [25] and individual efforts like the one reported in Ref. [26] are the step in the right direction. Thus at the end, the future of dedicated TMD Monte Carlo generators might still come out as a bright one.

Acknowledgements

The author would like to thank the organizers for their generous support. In addition this work was supported by the Basque Foundation of Science (IKERBASQUE) and the UPV/EHU under program UFI 11/55, as well as the European Community Research Infrastructure Integrating Activity under the FP7 “Study of strongly interacting matter (HadronPhysics3, Grant Agreement number 283286)”.

References

- [1] H. Jung, *Comput. Phys. Commun.* **86**, 147 (1995)
- [2] A. Bacchetta, U. Elschenbroich, Y. Miyachi, G. Schnell, T.A. Shibata, Y. Takubo, H. Tanaka, R. Seidl (2004), *On the way to a decomposition of the*

- flavor dependent Sivers functions: Transverse asymmetries, the decomposition and the transverse MC generator GMC_TRANS*, HERMES Internal Report 04-039
- [3] H.H. Matevosyan, A.W. Thomas, W. Bentz, Phys. Rev. **D83**, 114010 (2011), arXiv:1103.3085 [hep-ph]
- [4] T. Sjöstrand, S. Mrenna, P.Z. Skands, JHEP **0605**, 026 (2006), hep-ph/0603175
- [5] S. Agostinelli et al. (GEANT4), Nucl. Instrum. Meth. **A506**, 250 (2003)
- [6] A. Airapetian et al. (HERMES Collaboration), Phys. Lett. B **567**, 339 (2003)
- [7] A. Bacchetta, H. Jung, A. Knutsson, K. Kutak, F. Samson-Himmelstjerna, Eur. Phys. J. **C70**, 503 (2010), arXiv:1001.4675
- [8] Sjöstrand, T. (1995), *PYTHIA 5.7 and JETSET 7.4: Physics and manual*, hep-ph/9508391
- [9] C. Adolph et al. (COMPASS Collaboration), Nucl. Phys. **B886**, 1046 (2014), arXiv:1401.6284
- [10] A. Bianconi, D. Alberto, M. Bussa, M. de Stefanis, L. Ferrero et al., Eur. Phys. J. **A49**, 42 (2013), arXiv:0911.5493
- [11] P.B. van der Nat, Ph.D. thesis, Vrije Universiteit (2007), HERMES-07-012
- [12] A. Airapetian et al. (HERMES Collaboration), Phys. Rev. **D87**, 012010 (2013), arXiv:1204.4161
- [13] D. de Florian, R. Sassot, M. Stratmann, Phys. Rev. **D75**, 114010 (2007), hep-ph/0703242
- [14] M. Anselmino, M. Boglione, U. D'Alesio, A. Kotzinian, F. Murgia et al., Phys. Rev. **D72**, 094007 (2005), hep-ph/0507181
- [15] A. Airapetian et al. (HERMES Collaboration), Phys. Rev. Lett. **94**, 012002 (2005), hep-ex/0408013
- [16] I. Akushevich, H. Böttcher, D. Ryckbosch (1998), *RADGEN 1.0: Monte Carlo generator for radiative events in DIS on polarized and unpolarized targets*, hep-ph/9906408
- [17] F. James, M. Roos, Comput. Phys. Commun. **10**, 343 (1975)
- [18] M. Diefenthaler, Ph.D. thesis, Friedrich-Alexander University Erlangen Nuremberg (2010), DESY-THESIS-2010-032
- [19] A. Airapetian et al. (HERMES Collaboration), Phys. Lett. **B728**, 183 (2014), arXiv:1310.5070
- [20] K. Allada et al. (Jefferson Lab Hall A Collaboration), Phys. Rev. **C89**, 042201 (2014), arXiv:1311.1866
- [21] A. López Ruiz, Ph.D. thesis, Ghent University (2012), DESY-THESIS-2012-049
- [22] V. Blobel (2002), *An Unfolding method for high-energy physics experiments*, hep-ex/0208022
- [23] <http://www.int.washington.edu/PROGRAMS/14-55w/>
- [24] <http://horizon.hadronphysics.eu>
- [25] <http://tmdlib.hepforge.org>
- [26] M. Aghasyan, H. Avakian, E. De Sanctis, L. Gamberg, M. Mirazita et al. (2014), *Studies of Transverse Momentum Dependent Parton Distributions and Bessel Weighting*, arXiv:1409.0487