2023

Toward a Generative Modeling Analysis of CLAS Exclusive 2◇ Photoproduction

T. Alghamdi
Old Dominion University

Y. Alanazi
Thomas Jefferson National Accelerator Facility

M. Battaglieri
INFN sezione di Genova

Ł. Bibrzycki
AGH University of Krakow

A. V. Golda
Lomonosov Moscow State University

See next page for additional authors

Authors

T. Alghamdi, Y. Alanazi, M. Battaglieri, Ł. Bibrzycki, A. V. Golda, A. N. Hiller Blin, E. L. Isupov, Y. Li, L. Marsicano, W. Melnitchouk, V. I. Mokeev, G. Montaña, A. Pilloni, N. Sato, A. P. Szczepaniak, and T. Vittorini

Toward a generative modeling analysis of CLAS exclusive 2π photoproduction

T. Alghamdi,[1,2,*] Y. Alanazi,[3] M. Battaglieri,[4] Ł. Bibrzycki,[5] A. V. Golda,[6] A. N. Hiller Blin,[7]
E. L. Isupov,[8] Y. Li,[1] L. Marsicano,[4] W. Melnitchouk,[3,9] V. I. Mokeev,[3] G. Montaña,[3]
A. Pilloni,[10,11] N. Sato,[3] A. P. Szczepaniak,[3,12,13] and T. Vittorini[4,14]

[1]*Department of Computer Science, Old Dominion University, Norfolk, Virginia 23529, USA*
[2]*Department of Computer Science, College of Computer Science and Information Technology,*
Al-Baha University, Al-Baha, Alaqiq 65779, Saudi Arabia
[3]*Thomas Jefferson National Accelerator Facility, Newport News, Virginia 23606, USA*
[4]*INFN sezione di Genova, I-16146 Genova, Italy*
[5]*AGH University of Krakow, Faculty of Physics and Applied Computer Science,*
PL-30-059 Kraków, Poland
[6]*Faculty of Physics, Lomonosov Moscow State University, 119991 Moscow, Russia*
[7]*Institute for Theoretical Physics, Tübingen University, D-72076 Tübingen, Germany*
[8]*Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, 119234 Moscow, Russia*
[9]*CSSM and CDMPP, Department of Physics, University of Adelaide, 5005 Australia*
[10]*Dipartimento di Scienze Matematiche e Informatiche, Scienze Fisiche e Scienze della Terra,*
Università degli Studi di Messina, I-98166 Messina, Italy
[11]*INFN Sezione di Catania, I-95123 Catania, Italy*
[12]*Department of Physics, Indiana University, Bloomington, Indiana 47405, USA*
[13]*Center for Exploration of Energy and Matter, Indiana University, Bloomington, Indiana 47403, USA*
[14]*Università di Genova, I-16146 Genova, Italy*

AI-supported algorithms, particularly generative models, have been successfully used in a variety of different contexts. This work employs a generative modeling approach to unfold detector effects specifically tailored for exclusive reactions that involve multiparticle final states. Our study demonstrates the preservation of correlations between kinematic variables in a multidimensional phase space. We perform a full closure test on two-pion photoproduction pseudodata generated with a realistic model in the kinematics of the Jefferson Lab CLAS `g11` experiment. The overlap of different reaction mechanisms leading to the same final state associated with the CLAS detector's nontrivial effects represents an ideal test case for AI-supported analysis. Uncertainty quantification performed via bootstrap provides an estimate of the systematic uncertainty associated with the procedure. The test demonstrates that GANs can reproduce highly correlated multidifferential cross sections even in the presence of detector-induced distortions in the training datasets, and provides a solid basis for applying the framework to real experimental data.

I. INTRODUCTION

Photoproduction of two pions, with photon energies in the few-GeV range, is an important process in hadron spectroscopy. It has been widely used to address several fundamental quests, such as the "missing baryons"

problem, and to demonstrate that multiparticle final states are necessary to determine the spectrum of unstable excited states. While copious data are available for single-pion photoproduction, and the correspondent phenomenology is well-understood, the addition of a third particle in the final state makes the description of this reaction considerably more complicated. At fixed photon energy, the unpolarized single-pion photoproduction cross section is described by a single independent variable, while for two pions three additional variables are needed. At beam energies of a few GeV, the highest statistics data sample is available from the Jefferson Lab Hall B CLAS experiment `g11` [1]. Even in this case, some bins in the multidimensional space are

unpopulated or subject to large statistical fluctuations. This results in large uncertainties in extracting the underlying reaction mechanisms.

The problem has been addressed by studying one or two variables at a time, while integrating over the others. During integration, correlations between variables, which in turn contain relevant physics information, are partially lost, making the results strongly model dependent. In this context, generative models based on machine learning (ML), which learn the original data distribution and create new so-called synthetic data that mimic the original distribution, can provide new opportunities for extracting the physics information preserving correlations. Furthermore, these models can provide another way to extract the 'true' values from experimental data removing detector effects, with a procedure known as unfolding. The application of machine learning (ML) for unfolding has gained significant attention in recent years within the field of particle physics. A notable method in this area has been proposed in Ref. [2], and has been applied to several analyses [3–6]. In the present study, we distinguish our work from previous investigations, such as those mentioned in [2], by explicitly employing generative modeling techniques. Our research aligns in scope with the work outlined in [7].

Recently, an event-level unfolding analysis using generative adversarial networks (GANs) in inclusive electroproduction was performed [8]. The analysis was able to reconstruct accurately single-variable cross sections. Here, we extend our analysis framework to a multiparticle final state, demonstrating for the first time that the proposed GAN based framework can be used to reproduce scattering reactions in a higher-dimensional phase space. Specifically, we optimize our ML analysis framework to the case of two-pion photoproduction at CLAS g11 kinematics.

This study serves as an excellent testing ground for evaluating the effectiveness of the ML analysis framework in a highly nontrivial case. The presence of baryon and meson resonances with diverse production mechanisms, which overlap within a limited phase space, generate intricate structures and correlations. Moreover, the CLAS detector's highly nonuniform response introduces additional complexities and distortions, adding another layer of complication to the analysis. To test and validate the framework, we generate Monte Carlo (MC) pseudodata with a realistic model of two-pion photoproduction. We produce a synthetic copy with an "unfolding" GAN trained on pseudodata that incorporate detector effects through GEANT simulations [9]. This would be equivalent to train the GAN with experimental data. The detector effects are unfolded using a "detector-simulation" GAN, independently trained on a second MC pseudodata sample generated according to phase space and passed through the GEANT model of the detector. We test the quality of the procedure by a quantitative comparison between the generated MC data and its synthetic copy. This closure test, based on MC

pseudodata, is a necessary step before applying our analysis framework to experimental data.

The paper is organized as follows: in Sec. II we review the importance of two-pion photoproduction in hadron spectroscopy and provide a detailed description of the g11 kinematics. In Sec. III we describe the MC framework used to generate pseudodata and incorporate the CLAS detector response. In Sec. IV we present the ML framework used for reproducing the detector effects and unfold the "true" distributions from the reconstructed pseudodata. The GAN results are reported in Sec. V, where we compare the generated events with the synthetic copy. Finally, in Sec. VI we summarize the procedure and outline work in progress to extend the current framework to the analysis of real CLAS data from Jefferson Lab.

II. TWO-PION PHOTOPRODUCTION

A. The physics case

The $\pi\pi N$ final state is one of the largest contributors to the total photoproduction cross section off protons at center-of-mass (CM) energies $W \lesssim 2.5$ GeV. Studies of this final state have considerably extended the available information on the spectrum of the excited states of the nucleon (N^*) and their photoexcitation amplitudes. The quantum numbers of these resonances can be assessed by studying the correlations between the invariant mass and the angular dependencies of their decay products. Theoretical estimates based on phenomenological approaches [10–14], continuum Schwinger methods [13,14] as well as from first principles within lattice QCD calculations [15], have predicted more states than apparently observed in experiments (for reviews, see Refs. [16–19]), which is referred to as the "missing baryons" problem.

A strategy to improve the sensitivity to the most elusive states is to impose consistency constraints by performing combined analyses of several final states at once, with $\pi\pi N$ playing a pivotal role for the resonances heavier than 1.6 GeV. This allows one to disentangle process-dependent nonresonant contributions, and extract the resonance properties in a nearly model-independent manner [20]. Furthermore, combining photoproduction and electroproduction data has recently proven to be effective in identifying overlapping resonances with the same quantum numbers, as in the case of the $N(1720)$ and $N'(1720)$ states [21–23].

In the same reaction, by looking at the invariant mass distribution of the $\pi\pi$ pair, one can study meson resonances, such as the ρ or the $f_2(1270)$. While the properties of these resonances are well-known, a detailed understanding of their production mechanisms is still missing. At low $W \lesssim 2$ GeV one can study how each N^* state contributes to the meson production process. At higher energies, above the N^* resonance region, the reaction is well-described in terms of Regge theory [24,25]. The two energy regimes are

smoothly connected, making it nontrivial to study the intermediate region rigorously. A formalism to do so has been proposed recently for the production of single π or η mesons [26,27]. The extension to two-pseudoscalar final states requires having the full multidimensional dependence under control [28]. In particular, a complete understanding of meson production mechanisms in the $\pi\pi N$ final state, where resonances are well-known, is necessary before facing the more complicated $\eta\pi N$ and $\eta'\pi N$ channels, where exotic hadrons are expected to appear [29].

B. $\gamma p \to \pi^+\pi^- p$ kinematics

Measurement of the three-body final state in two-pion photoproduction represents a significant challenge to experiment. Recently a large body of data on $\pi^+\pi^- p$ photoproduction observables has become available from measurements by the CLAS Collaboration, with $W \leq 2.9$ GeV [17,22,30–32]. For a given collision energy, the differential cross section for this process depends on five independent variables, which can be chosen to be the invariant masses of the two pions, $M_{\pi^+\pi^-}$, and the proton-π^- pair, $M_{p\pi^-}$, and three angles in the CM frame. Two of the angles are the polar angle θ_{π^+}, with the z-axis along the photon three-momentum, and the angle $\alpha_{[\pi^+ p][\pi^- p']}$ between the plane containing the initial target proton p and π^+ three-momenta and the plane containing the π^- and recoiling proton p' three-momenta. An equivalent choice would replace θ_{π^+} with the invariant momentum transferred t_{π^+}, defined as the difference squared between the photon and π^+ four-momenta. The fifth variable ϕ is the azimuthal angle of π^- with respect to the plane containing the photon three-momentum and the polarization vector, and is relevant only in experiments with polarized beam or target. For unpolarized data, one can still define ϕ by pointing the polarization vector in an arbitrary direction, resulting in a ϕ-independent cross section. Other possible choices for variables are $M_{p\pi^+}$ (invariant mass of the proton-π^+ pair), t_{π^-} (momentum transferred between photon and π^-), t (momentum transferred between target and recoil protons), or $\cos\theta$ (cosine of the angle between target and recoil protons in the CM frame).

Multidimensional analyses are becoming standard, albeit computationally difficult, in modern high statistics experiments [33,34]. However, some specific reactions can suffer from limited statistics. In particular, the direct extraction of $\pi^+\pi^- p$ photoproduction events at a given W value, on a 5D grid (or 4D, if integrated over the angle ϕ) with a bin size acceptable for physics analyses, is quite challenging. Even the highest statistics $\pi^+\pi^- p$ photoproduction sample collected with CLAS [22,30] results in a limited number of counts in the 4D cells (typically < 10 events per cell). In Ref. [22], theoretical curves were fitted to the marginal 1D distributions, determined by integrating the acceptance- and efficiency-corrected 5D distribution over the remaining four variables. This procedure largely washes out the
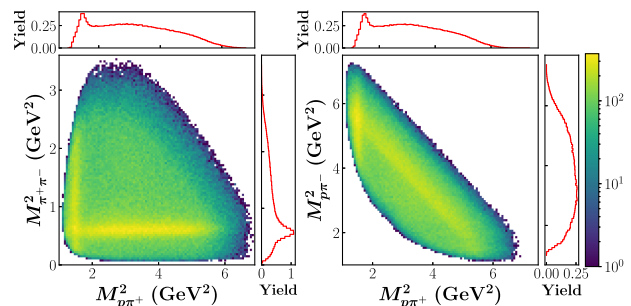

FIG. 1. Examples of 2D normalized yield distributions and their 1D projections measured in CLAS g11 experiment (before unfolding detector effects) [30]. Plots show invariant mass distributions for the yields $M_{\pi^-\pi^+}^2$ versus $M_{p\pi^+}^2$ (left) and $M_{p\pi^-}^2$ versus $M_{p\pi^+}^2$ (right). For each 2D plot, the correspondent 1D projections are shown on top and right. The distributions are dominated by the prominent contributions from Δ^{++} and ρ resonances, which appear as bands in the 2D distributions and peaks in the appropriate 1D projection. In particular, the Δ^{++} appears as a vertical band in the 2D plots, and as a peak in the 1D $M_{p\pi^+}^2$ projection. The ρ is the horizontal band in the left 2D plot, the diagonal band in the right 2D plot, and the peak in the 1D $M_{\pi^+\pi^-}^2$ yields. From the right plots one can see that, choosing $M_{p\pi^+}^2$ and $M_{p\pi^-}^2$ as independent variables, the existence of the ρ is lost if one studies the 1D projections only.

correlations present in the original data, leading to a significant loss of relevant information contained in the joint distribution. In this paper we aim to overcome this problem with ML techniques.

To illustrate this, in Fig. 1 we show two examples of 2D distributions and their 1D projections, as measured in CLAS g11 experiment without efficiency corrections [30]. From these distributions one immediately sees the presence of intermediate resonances that appear as enhancements in the invariant mass of the system in which they decay. For example, the band at $M_{p\pi^+}^2 \simeq 1.5$ GeV^2 corresponds to the $\Delta(1232)$ baryon resonance, which appears as an intermediate unstable state in the reaction $\gamma p \to \Delta^{++}\pi^- \to p\pi^+\pi^-$. The band centered at $M_{\pi^+\pi^-}^2 \simeq 0.6$ GeV^2 corresponds instead to the $\rho(770)$ meson resonance, in the reaction $\gamma p \to p\rho^0 \to p\pi^+\pi^-$. The two resonances are clearly visible as bumps in the respective 1D projections. However, looking at the 1D projections only, one can easily miss the presence of a resonance if the relevant invariant mass distribution is not explicitly considered; for example, no peak is visible in the $M_{p\pi^-}^2$ invariant mass distribution shown in the right panel of Fig. 1. This is an example of loss of information that is contained in correlations. Moreover, because of quantum interference, the production of ρ^0 and Δ^{++} are not independent processes, and it is impossible to associate one event exclusively with either process. This interference appears in the correlations between the invariant masses, and can be partially lost in the 1D projections.

C. Two-pion photoproduction with CLAS

The CLAS spectrometer in Hall B at Jefferson Lab was based on a ~ 1.25 T toroidal magnet which bends charged particles produced in the hadronic interaction along the polar angles θ_{lab} (the z-axis along the photon beam), while preserving azimuthal angles ϕ_{lab}. The polarity of the field determined if positive/negative charges were bent towards/away from the beam line into the acceptance of the detector. A system of three layers of multiwires drift chambers [35] provided momentum information with the resolution, σ_p / p, ranging from 0.5% to 1.0%, depending on the kinematics. Charged hadron identification was obtained by time-of-flight scintillators [36]. Photoproduction experiments were conducted with a bremsstrahlung photon beam produced by the CEBAF continuous electron beam impinging on 8×10^{-5} radiation lengths thickness gold foil. A bremsstrahlung tagging system [37] with a photon energy resolution of 0.1% was used to measure the photon energy in each recorded event. The target cell was a 4 cm in diameter and 40 cm long Mylar cylinder, filled with liquid hydrogen at 20.4 K.

The experimental conditions reported in this paper, and simulated in the framework described in Sec. III, correspond to the `g11` experiment that ran in CLAS in 2004. During the experiment, the torus field was such that positive particles were bent away from the beam line. The detector geometrical acceptance for each positive particle in the relevant kinematic region was about 40% and somewhat less for negative particles (bent towards the beamline and out of the detector acceptance). The primary electron beam energy was 4.02 GeV, providing a tagged photon beam in the energy range from 0.8 to 3.8 GeV. For this analysis we focus on the highest energy region, 3.0–3.8 GeV, that was analyzed in Ref. [30].

The exclusive reaction $\gamma p \to \pi^+ \pi^- p$ was isolated by detecting the proton and the π^+ in the CLAS spectrometer, while the π^- was reconstructed from detected particle four-momenta using the missing-mass technique. In this way, the exclusivity of the reaction was ensured, keeping the contamination from the multipion background to a minimum level. Only events within a fiducial volume were retained in the analysis, in order to avoid the regions at the edge of the detector acceptance. Cuts were defined on the minimum proton momentum and the hadron minimum and maximum polar angle. After all the cuts, approximately 40 M events were identified as produced in exclusive two-pion photoproduction, making the `g11` dataset the largest statistics sample of this reaction in the above photon energy range. Details of the `g11` analysis can be found in Ref. [30].

III. MC SIMULATION FRAMEWORKS

In this section we describe the simulation frameworks used to perform the closure test. Pseudodata corresponding to two-pion photoproduction in the kinematics of the

`g11` experiment were generated using two different MC event generators that produce the four-momenta of the final state particles. A realistic GEANT simulation was used to reproduce the finite resolution and limited acceptance of the CLAS detector.

Detector effects were assessed with a first MC generator based on a pure phase-space distribution. To perform the closure test, we deployed a second MC generator based on a realistic physics model. The use of two different MC generators minimizes the model dependence in the extraction of the original information and mimics a real situation, where the detector effects are estimated with simulations that are similar but not identical to the experimental distributions.

A. Two-pion event generators

The two MC generators simulate the interaction of an incoming unpolarized photon beam with a bremsstrahlung spectrum, in the energy range 3.0–3.8 GeV, with a target proton at rest. With the choice of variables described in Sec. II B, the yields are proportional to the differential cross section, and thus to the squared of the production amplitude A summed over polarizations,

$$
\frac{\mathrm{d}^5 \sigma}{\mathrm{d}M_{p\pi^-}^2 \, \mathrm{d}M_{\pi^+\pi^-}^2 \, \mathrm{d}t_{\pi^+} \, \mathrm{d}\alpha_{[\pi^+ p][\pi^- p']} \, \mathrm{d}\phi}
$$
$$
\propto \left[\left(W^2 - (M_{p\pi^-} + m_\pi)^2 \right) \left(W^2 - (M_{p\pi^-} - m_\pi)^2 \right) \right]^{-1/2}
$$
$$
\times \sum_{\text{pol}} \left| A\left(M_{p\pi^+}^2, M_{\pi^+\pi^-}^2, \cos\theta_{\pi^-}, \alpha_{[\pi^+ p][\pi^- p']} \right) \right|^2. \quad (1)
$$

The first MC generator, referred to as phase space or PS-MC, distributes final-state events according to the $\pi^+ \pi^- p$ phase space. This corresponds to assuming that the production amplitude is a constant. This is clearly unrealistic since, as discussed above, two-pion photoproduction has a much more complicated structure. However, it has the advantage of being well-defined, agnostic to physics models, and distributes events uniformly across the full reaction kinematics. The 1D-projected PS-MC event distributions are shown in Fig. 2, while the 2D distributions are illustrated in Fig. 3.

The second MC event generator, which we refer to as realistic or RE-MC, considers the amplitude squared as an incoherent sum of the three dominant intermediate resonances observed, $\gamma p \to (p\rho^0, \Delta^{++}\pi^-, \Delta^0\pi^+) \to \pi^+ \pi^- p$, added to a $\sim 10\%$ constant that mimics the nonresonant two-pion photoproduction contribution. Each process has been weighted with the corresponding contribution to the total cross section as reported in Ref. [38]. The angular distributions relative to resonance production are parametrized from measured differential cross sections reported in the same database. The decays $\rho \to \pi\pi$ and $\Delta \to p\pi$ are described using the correct spin structure with the decay
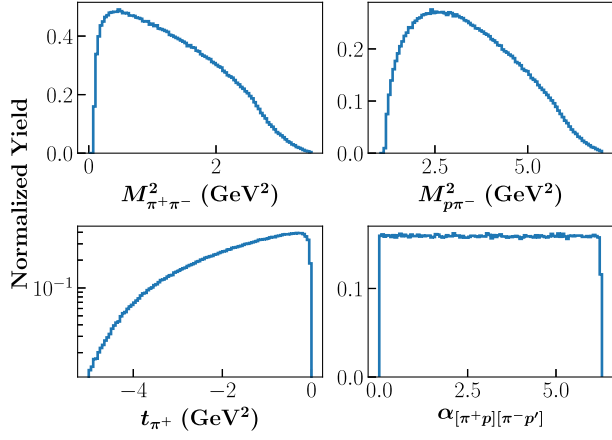
FIG. 2.   Normalized 1D projections of events generated with PS-MC: $\pi^+\pi^-$ invariant mass squared (top left), $p\pi^-$ invariant mass squared (top right), square of the four-momentum transferred from the target to the recoil proton, t_{π^+} (bottom left), the angle $\alpha_{[\pi^+p][\pi^-p']}$ in the CM frame between the plane containing the initial target p and π^- three-momenta and the plane formed by π^+ and the scattered proton three-momenta (bottom right).
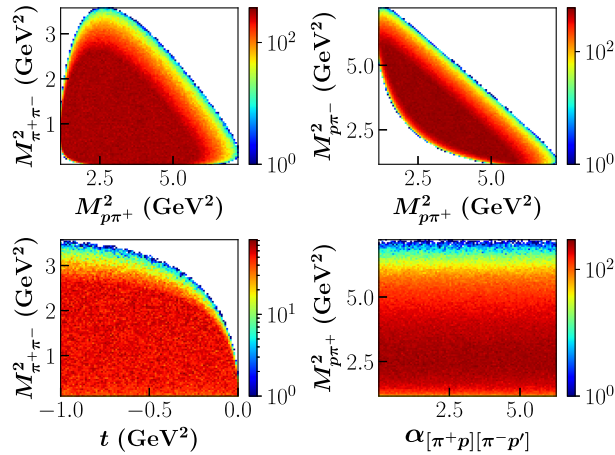

FIG. 3.   Normalized 2D distributions generated with PS-MC: invariant mass distributions $M^2_{\pi^+\pi^-}$ versus $M^2_{p\pi^+}$ (top left), and $M^2_{p\pi^-}$ versus $M^2_{p\pi^+}$ (top right), $M^2_{\pi^+\pi^-}$ invariant mass versus momentum transfer t (bottom left), and CM angle versus $M^2_{p\pi^+}$ (bottom right).

matrix elements detailed in Ref. [39]. The resulting 1D and 2D projections for events generated by RE-MC are shown in Figs. 4 and 5, respectively. We note that this model neglects the interference terms between the intermediate resonances. Despite this, the resulting distribution provides a reasonable description of the experimental data, showing resonance structures in the invariant masses and the correct angular behavior of particles in the final states.

B. CLAS detector simulation

The CLAS detector response has been simulated using the standard GEANT Monte Carlo simulation package,
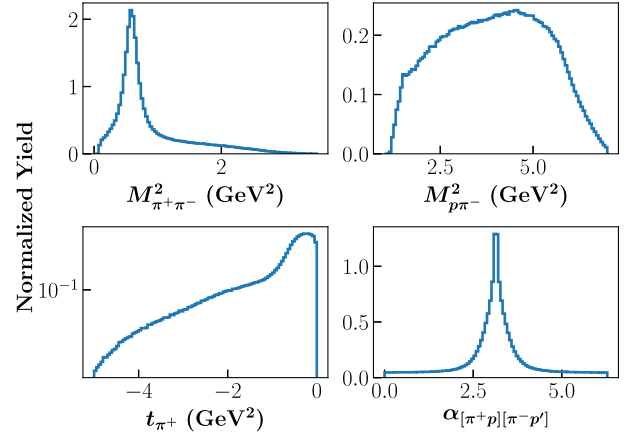

FIG. 4.   Normalized 1D projections of events generated with RE-MC. The panel descriptions are as in Fig. 2.
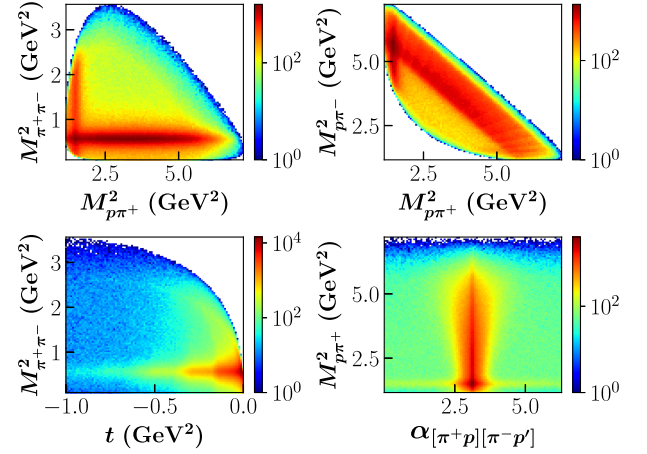

FIG. 5.   2D distributions generated with RE-MC. The panel descriptions are as in Fig. 3.

GSIM, used by the CLAS Collaboration [40]. It consists of a central steering and control package that calls a number of independent detector geometry and response packages. A postprocessing code (GSIM-Post-Processor or GPP) has been used to fine tune the GSIM output to match the tails of the experimental resolution and other effects, such as the detector's dead channels, not described by the idealized GEANT-based simulation. The GSIM output has been fed to the same reconstruction code, RECSIS, used to process experimental data. We will refer to REC or detector-level events to identify the set of pseudodata as processed by the detector simulation, while GEN or vertex-level will identify the "true" events as generated by the MC code.

As reported in Sec. II C, the CLAS detector has a nonuniform acceptance, reduced in the azimuthal angle ϕ_{lab} (around the beam) by the presence of the six coils of the toroidal magnet, and in the polar angle θ_{lab} (with respect to the beam direction) by the limited area covered by the drift chambers, calorimeter and time-of-flight systems.
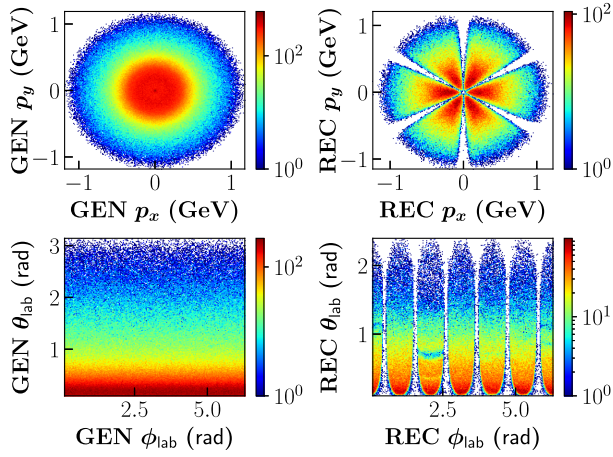
FIG. 6.   The π^+ kinematic variables in the laboratory reference frame as GENerated with RE-MC (left) and REConstructed by CLAS (right): p_x versus p_y (top panels) and θ_{lab} versus ϕ_{lab} (bottom panels).

FIG. 7.   The smearing of π^+ kinematic variables in the laboratory reference frame. Top: relative momentum resolution $(p_{REC} - p_{GEN})/p_{GEN}$ as a function of p_{GEN}. Bottom: relative angular resolution defined as $(\theta_{REC} - \theta_{GEN})/\theta_{GEN}$ as a function of p_{GEN}. The right panels display the 1D projections of the 2D distributions as histograms, and the red lines represent double Gaussian fits. Pseudodata events were generated with RE-MC.

A further limitation concerns the minimum accepted momenta of charged hadrons, due to the energy loss in materials crossed along the track and to the effect of the toroidal magnetic field that bends low-momentum particles out of the detector acceptance. The limited CLAS acceptance results in a reduced yield in REC with respect to GEN events, since not all generated events are reconstructed. The effect of the CLAS acceptance on the π^+ variables in the laboratory frame is shown in Fig. 6.

As any detector, CLAS has finite resolution, which "smears" the measured kinematic variables resulting in a difference between REC and GEN, even when the event is accepted. The smearing affects the reconstructed three-momenta of any detected particle within the CLAS acceptance with a distortion depending on the three-momentum of the particle. Figure 7 shows the resolution on the detected (REC) π^+ momentum and polar angle as a function of the "true" (GEN) momentum, along with the projections in 1D corresponding to the CLAS relative momentum and angular resolution. Fitting the two curves to a double Gaussian line, we obtained $\delta p/p \sim 0.8\%$ and $\delta\theta/\theta \sim 0.5\%$. A similar smearing affects the kinematic variables of the detected proton.

The resolution of the CLAS detector is sufficiently high so as to allow the use of the missing mass technique to identify the exclusive two-pion reaction against the multi-pion background. The technique uses knowledge of the initial state and of the detected particles to calculate the invariant mass of the undetected system to fulfill energy-momentum conservation, within detector resolution. If all particles are detected, the missing mass is zero. If a single particle is undetected, its mass appears as a peak in the missing mass spectrum. If two or more particles are lost, the missing mass of the system is unconstrained and does not pea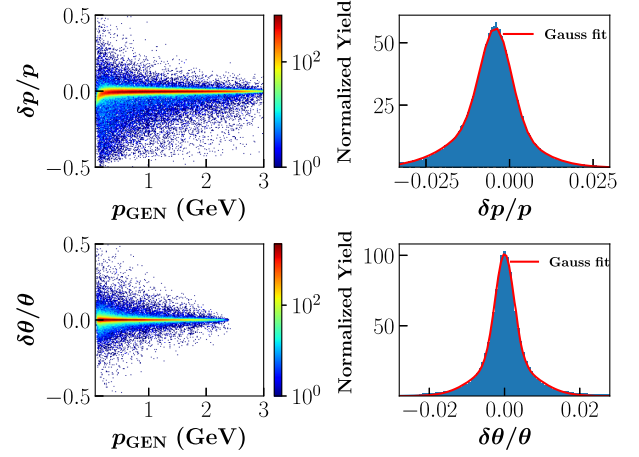k, but rather distributes smoothly. The technique is only applicable if the experimental resolution is sufficient to disentangle the missing mass peak from this multiparticle background. Clearly, the more particles are detected, the lower is the resolution for the missing mass due the error propagation, limiting the validity of the technique to reactions with a small number of particles in the final state. When the missing particle has been identified, its four-momentum is determined by energy and momentum conservation, and the final state can be fully reconstructed.

In two-pion photoproduction, the requirement of at most a single undetected particle corresponds to the following topologies (missing particle in parentheses): $p\pi^+(\pi^-)$, $p\pi^-(\pi^+)$, $\pi^+\pi^-(p)$ and $\pi^+\pi^- p$ (all three detected). Considering the CLAS acceptance, the yield of different topologies is quite different, with a ratio of $(100{:}37{:}30{:}35)$ for the respective topologies. Since the $p\pi^+(\pi^-)$ is by far the dominant contribution to REC data, we focus on this topology, although similar conclusions also hold for the others. Each topology is in one-to-one correspondence with different areas of the allowed phase space, and a combination of different topologies would therefore extend the kinematic coverage of the measurement, mitigating the effect of the limited detector acceptance.

Figure 8 shows the missing mass distribution of the $p\pi^+(\pi^-)$ topology. This exclusive final state is identified by selecting events with missing mass in the peak. Since these simulations only contain the two-pion final state, no multiparticle background populates the plot. The equivalent distribution for g11 data shows a significant multipion background [30] that populates the positive side of the missing mass spectrum, and is rejected during the analysis to assure the reaction exclusivity.
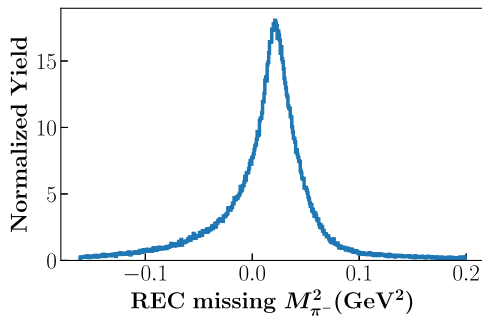
FIG. 8. The π^- missing mass spectrum for the $p\pi^+(\pi^-)$ topology from REC data. Pseudodata events were generated with RE-MC.

IV. GAN-BASED UNFOLDING METHODOLOGY

GANs, a type of neural networks that have gained significant attention in recent years, are powerful generative models highly effective in generating high-quality, realistic data in various fields [41]. The architecture of a typical GAN involves a generator network that learns to produce data and a discriminator network that learns to differentiate between the generated and reference data. The two networks are trained alternately in a competitive setting, where the generator tries to produce more realistic data to fool the discriminator, and the discriminator tries to correctly identify the generated data. This iterative process leads to the generation of data that are progressively more realistic, with the ultimate goal of producing synthetic data that are indistinguishable from the reference data.

GANs have been widely applied in many domains, such as image synthesis [42], text generation [43], music composition [44], and videos [45], and have demonstrated impressive results. In image synthesis, GANs have been used to generate highly realistic images visually indistinguishable from real images, which has numerous practical applications in fields such as gaming, film, and art.

Successfully training GANs can be notoriously challenging, however. Numerous GAN models experience significant issues, such as mode collapse, nonconvergence, model parameter oscillation, destabilization, vanishing gradients, and overfitting, resulting in an unbalanced training of the generator and discriminator [46–49]. In contrast to typical GAN applications, the success of a GAN-based event generator in nuclear and particle physics depends on its ability to accurately reproduce correlations among the momenta of the particles, which becomes increasingly challenging beyond two dimensions. Moreover, the multi-dimensional momentum distributions of events associated with nuclear and high-energy physics reactions, such as the two-pion photoproduction process considered in this work, exhibit highly complex patterns and range over orders of magnitude across the phase space. The task of developing an appropriate GAN architecture that is able to simultaneously reproduce all the correlations among particle momenta, and

accurately reproduce multidimensional histograms, is therefore rather difficult.

Machine learning event generators have gained prominence as efficient fast simulation tools in various scientific fields, including high-energy and nuclear physics [50–56]. Unlike traditional simulation methods that rely on a theoretical framework for the underlying reaction, machine learning event generators learn from large datasets and use this knowledge to produce new events with high fidelity. GANs have emerged as powerful tools in the field of fast simulation, where they learn to generate events that closely resemble reference data, capturing the underlying physics processes and their distributions [54,55,57–59].

Furthermore, GANs have been employed to address the challenge of simulating detector effects in fast simulation [8,60]. This application of GANs helps bridge the gap between simulated and reference data, enabling more realistic and precise simulations for experimental analyses. A comprehensive survey of existing ML-based event generators can be found in Ref. [61]. In this study, we employ the architectural framework of the Least Squares GAN, which involves substituting the cross entropy loss function in the discriminator component of a conventional GAN with a least square term. For further details, see Ref. [62].

In the following, we describe the GAN architecture used to generate the synthetic data that reproduce the $\gamma p \to \pi^+ \pi^- p$ RE-MC pseudodata. As mentioned above, two different GANs were developed and combined. The detector simulation GAN (DS-GAN) was trained on PS-MC pseudodata to learn the detector effects, and was later inserted between the generator and the discriminator of the unfolding GAN (UNF-GAN) to unfold the GEN vertex-level information from REC pseudodata.

A. Detector simulation GAN (DS-GAN)

In order to capture the detector effects, we have developed an ML-based detector simulation using a conditional GAN [63], as illustrated in Fig. 9. Our approach involves training a conditional GAN generator to simulate the detector's smearing effect so that it generates synthetic REC detector-level events from input noise and PS-MC GEN events. The GEN PS-MC accepted events are passed through the GEANT chain to obtain REC pseudodata. As proposed in Ref. [60], both the synthetic REC and REC pseudodata are "concatenated" with original GEN events and fed to the GAN discriminator as input to facilitate convergence. In the training phase, from the event-by-event comparison of GEN and REC pseudodata, the DS-GAN learns how the detector impacts on the original data. After successful training, the DS-GAN generator serves as the ML detector surrogate that will be integrated into the UNF-GAN architecture.

Summarizing the model architecture of the DS-GAN, the generator, conditioned on accepted events (GEN), takes in
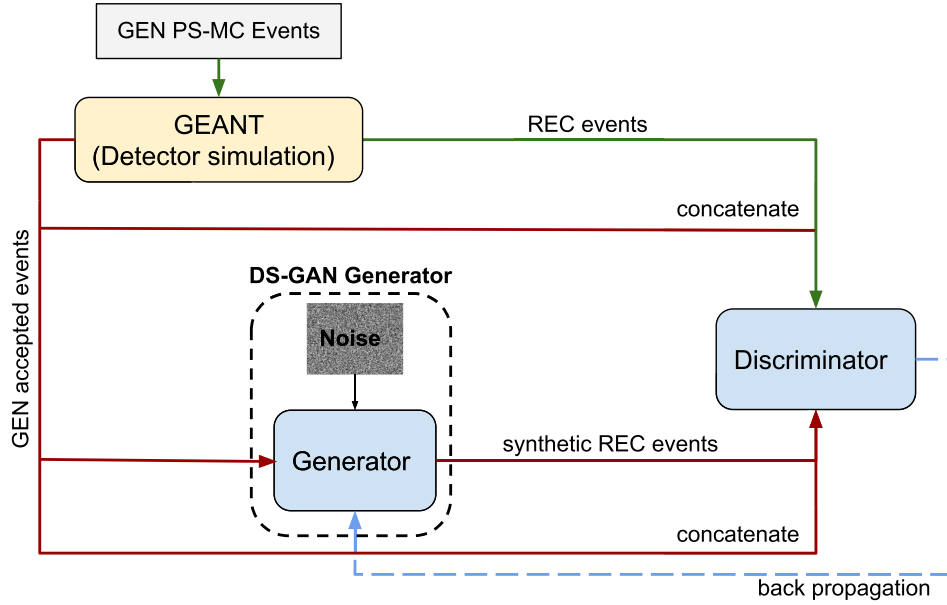
FIG. 9. Schematic view of the ML detector simulation GAN (DS-GAN), where the GAN generator converts input GEN vertex-level events features and noise to REC detector-level events. The training is performed on PS-MC pseudodata passed through the GEANT simulation. Synthetic REC and REC pseudodata are concatenated with GEN PS-MC events and fed to the discriminator.

as input a 100-dimensional array of random values with a mean of 0 and a standard deviation of 1. The generator network consists of five hidden layers, each with 128 neurons, using a leaky rectified linear unit (ReLU) activation function. The final hidden layer is connected to a four-neuron output layer, which uses a linear function to represent the generated features. At the end of the training, the DS-GAN generator learns how to convert the GEN accepted events into REC events, effectively mimicking the smearing due to the detector as described by GEANT.

The discriminator is made of a neural network with five hidden dense layers. The first three layers have 256 neurons each, while the fourth has 128 neurons and the fifth has 32 neurons. A leaky ReLU activation function is used for all the layers. To prevent overfitting during training, a 5% dropout rate is implemented for each hidden layer. The last hidden layer is fully connected to a single-neuron output, activated by a linear function, where "1" indicates a true event and "0" is a fake event. The DS-GAN was trained using about 1M two-pion event samples for 80K adversarial epochs, with an epoch defined as one pass through the training dataset. Both the generator and discriminator were trained using the Adam optimizer [64] with a learning rate of 10^{-5} and exponential decay rates for the moment estimates ($\beta 1 = 0.5$, and $\beta 2 = 0.9$).

B. Unfolding GAN (UNF-GAN)

The training process for the UNF-GAN is illustrated in Fig. 10, which depicts the variation of a typical GAN model structure consisting of a conditional generator and a discriminator. The generator takes as input the photon energy generated by the RE-MC, along with a 100-dimensional white noise vector centered at zero with a unit standard deviation. This combination of inputs allows the generator, implemented as a deep neural network, to transform the noise and photon energy into a minimal set of event features/variables that effectively describe the two-pion photoproduction reaction.

To strike a balance between execution time and convergence, the generator network is designed with seven hidden dense layers. The number of neurons in each layer follows the sequence: 16, 32, 64, 128, 256, 512, and 1024, all of which are activated by the ReLU function. The last hidden layer is fully connected to a 4-neuron output layer, activated by a linear function. This output layer represents the independent variables $M^2_{\pi^+\pi^-}$, $M^2_{p\pi^-}$, t_{π^+}, and $\alpha_{[\pi^+ p][\pi^- p']}$ that are specifically chosen to describe the reaction.

The synthetic GEN event features, generated by the conditional GAN generator, are then fed into the DS-GAN to incorporate the detector effects, and then compared to REC pseudodata obtained by passing the GEN RE-MC pseudodata through Geant. The training process involved utilizing approximately 400k two-pion event samples for a duration of around 200 k adversarial epochs per UNF-GAN model. Consistent configuration parameters for the Adam optimizer were maintained, utilizing the same settings as employed for the DS-GAN. During the training, the generator and the discriminator engage in an adversarial competition, with both updating their parameters throughout the process. Eventually, the generator is able to generate synthetic REC samples that are indistinguishable from the REC pseudodata samples. This means that the
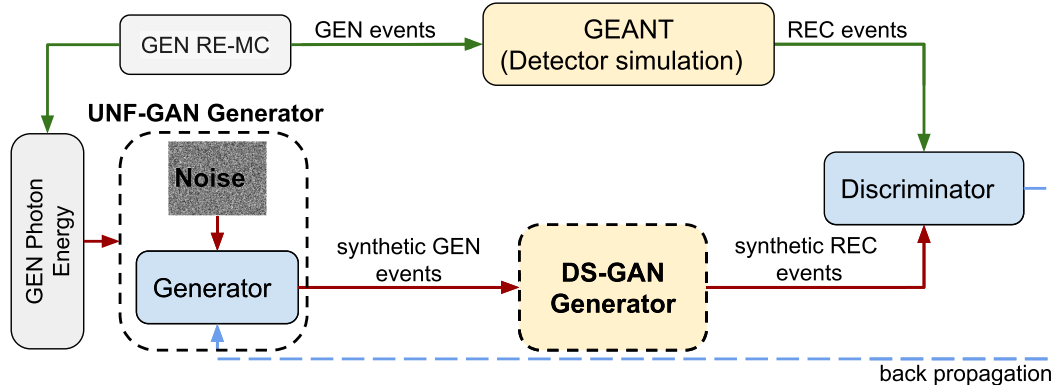
FIG. 10. Schematic view of the UNF-GAN training framework: the UNF-GAN utilizes a generator that converts a GEN photon energy and random noise into synthetic GEN event features, which are passed through the DS-GAN to incorporate detector effect, and are converted into synthetic REC event features. The generator is updated using gradients constructed by a deep neural network discriminator, which compares the features of synthetic and reference REC detector-level events obtained through Geant.

discriminator's ability to correctly classify whether a sample is genuine or synthetic approximates random chance.

C. Uncertainty quantification

As neural networks become increasingly employed in physics analysis, it becomes crucial to accurately assess the reliability of ML predictions. The statistics of the synthetic samples can be made arbitrarily high, so that there is no need to consider a statistical uncertainty. However, it is important to quantify the systematic uncertainty related to the training procedure, and for this a bootstrap resampling technique was employed. For the DS-GAN, the procedure involved training a total of 20 neural networks independently from the beginning. Each one was trained on a different random sample set drawn from the original dataset with replacement, resulting in datasets of the same size but with potentially different observations.

For the UNF-GAN a similar procedure was adopted, with 20 different networks trained independently using the same bootstrap resampling technique. Moreover, each of the 20 UNF-GANs used a different DS-GAN of the 20 discussed above. In this way, the systematic uncertainties associated with the DS- and UNF-GANs are effectively combined. While it is possible that using a higher number of bootstraps could potentially lead to more precise uncertainty estimates, we found that training 20 GANs provided reasonably stable and consistent results.

It is important to note that the specific number of bootstraps can vary depending on the characteristics of the problem, available data, and desired level of uncertainty quantification. In this particular case, 20 bootstraps were deemed sufficient for accurately capturing and quantifying the uncertainties associated with the observables. Furthermore, changing the network architecture was not essential because the convergence we achieved, along with the estimated error and uncertainty quantification, clearly indicate that this architecture is capable of

accurately reproducing the data without introducing further systematic uncertainties.

V. RESULTS

In this section we now discuss the DS-GAN and UNF-GAN performance, comparing synthetic to the REC and GEN pseudodata. We use the nomenclature REC_{SYN} and GEN_{SYN} to indicate synthetic data at the detector and vertex levels, respectively. To visualize the comparison, we build marginal 1D and 2D histograms for some kinematic variables. To show that correlations are correctly accounted for, we also study the distribution of one variable in some slices of the other variables.

Synthetic data are generated with the bootstrap procedure detailed in Sec. IV C, so that the standard deviation σ_{SYN} corresponds to the systematic uncertainty. In all our results, the average μ_{SYN} is shown as a solid line, together with an error band of width $\pm 1\sigma_{SYN}$, while pseudodata are represented by dots with their statistical uncertainty $\sigma_{pseudodata}$. To quantify the level of agreement between the synthetic data and pseudodata, we plot the pull for each bin, defined as

$$\text{pull} = \frac{\mu_{SYN} - \mu_{pseudodata}}{\sqrt{\sigma_{SYN}^2 + \sigma_{pseudodata}^2}}, \qquad (2)$$

where $\mu_{pseudodata}$ denotes the mean of the pseudodata.

A. DS-GAN

The DS-GAN is trained on four independent variables: the invariant masses $M_{p\pi^-}^2$ and $M_{\pi^+\pi^-}^2$, t_{π^+}, and the angle $\alpha_{[\pi^+ p][\pi^- p']}$. The comparison between REC_{SYN} and pseudodata PS-MC REC distributions is shown Fig. 11. In Fig. 12 the comparison is extended to other physics-relevant distributions not used in the training and derived from
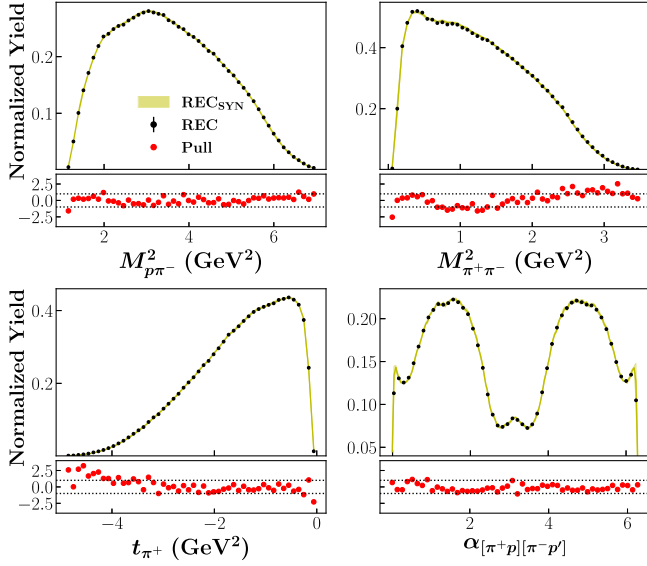
FIG. 11. Training variable PS-MC REC pseudodata distributions (black points) are compared to the synthetic data produced by the DS-GAN (yellow band). The band size reflects the uncertainty estimated using the bootstrap procedure, while the pull distributions (bottom of panels) quantify the agreement between the two, with the horizontal dotted lines corresponding to $\pm 1\sigma$.
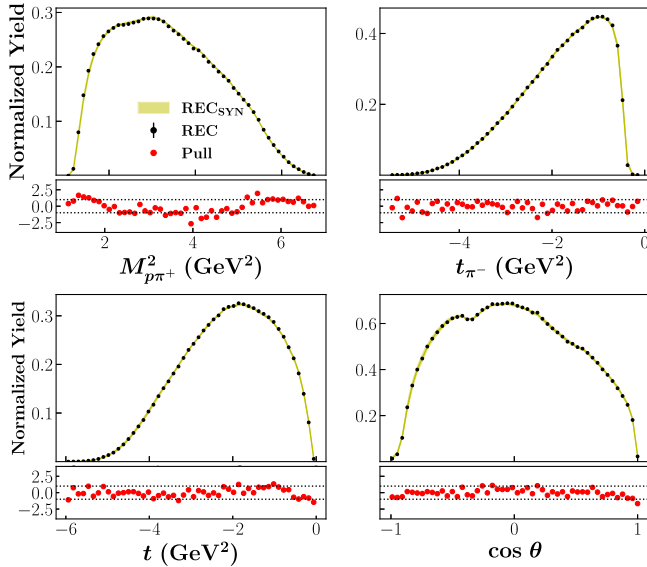

FIG. 12. Same as Fig. 11, but for derived quantities not used in the training process.

the four above-mentioned variables, namely $M^2_{p\pi^+}$, t_{π^-}, t, and $\cos\theta$. The agreement, quantified by the pull distributions shown at the bottom of each plot, is remarkable, in both cases, with most of the points lying within 1σ. This indicates that the DS-GAN is indeed able to learn the CLAS detector effects. Bidimensional distributions from MC and synthetic data are shown in Fig. 13.
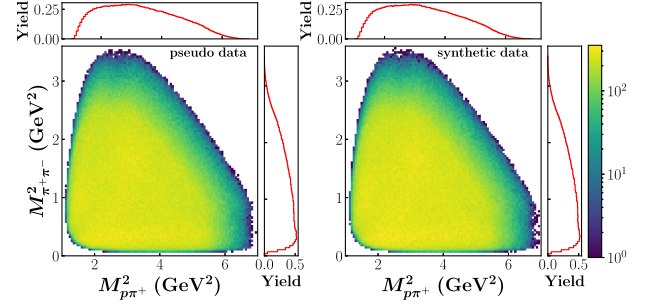
FIG. 13. 2D distributions of $\pi^+\pi^-$ and $p\pi^+$ invariant masses squared in PS-MC REC pseudodata (left) and synthetic data output of one of the 20 DS-GANs (right).
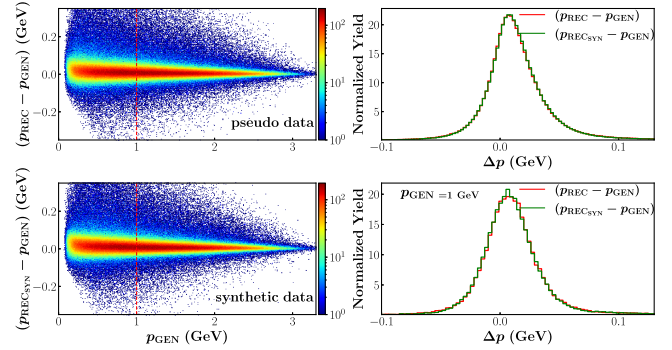

FIG. 14. Left: Comparison of π^+ absolute momentum resolution for PS-MC REC pseudodata (top) and DS-GAN synthetic data (bottom) as a function of p_{GEN}. Right: The same but integrated over p_{GEN} (upper) or for a narrow slice at $p_{\text{GEN}} = 1$ GeV (lower).

The π^+ absolute momentum resolution as obtained from pseudodata (REC − GEN) is shown in Fig. 14, along with synthetic data (REC$_{\text{SYN}}$ − GEN). The two distributions are in very good agreement, indicating that synthetic data incorporate the correct resolution of the detector. Similar results hold for other kinematic variables of all particles.

These comparisons demonstrate the ability of the DS-GAN to learn and reproduce detector effects in a multidimensional space, even in the tails of the distributions. This confirms that generative models can indeed be used as an efficient and fast proxy for more computational expensive GEANT simulations [60].

B. UNF-GAN

As described in Sec IV B, the final step in the closure test is to use REC RE-MC pseudodata to train the UNF-GAN, extract the GEN$_{\text{SYN}}$ distributions, and compare them with GEN pseudodata. Figure 15 shows the comparison between GEN and GEN$_{\text{SYN}}$ for the four training variables. We can see very good agreement between pseudo- and synthetic data at the vertex level, despite the fact that the UNF-GAN
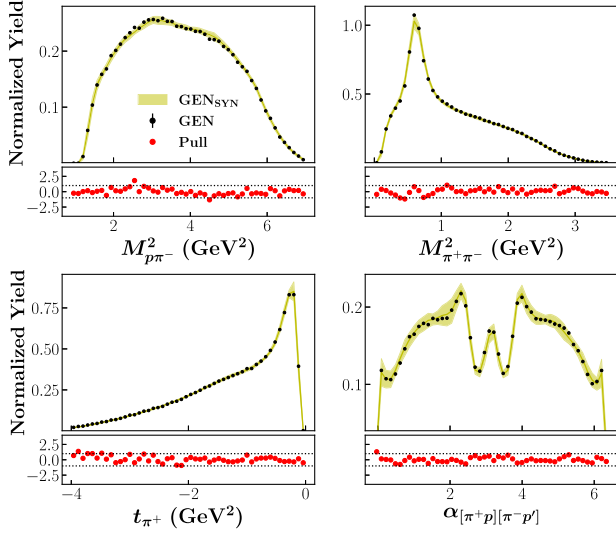
FIG. 15. Comparison of GEN and GEN$_{SYN}$ distributions for the four training variables obtained by the UNF-GAN. Pseudodata were generated using RE-MC.

was trained on detector-level pseudodata. This clearly demonstrates the success of the unfolding procedure. The vast majority of pulls lie within $\pm 1\sigma$, indicating that the uncertainty quantification is appropriate. One observes residual correlations, particularly in the high statistics regions. In other regions, the pulls appear to be randomly scattered. At a practical level, this implies that the quoted 1σ uncertainty is not necessarily purely statistical, but also includes a systematic component.

The key point of this closure test is to demonstrate that synthetic data maintain the correlations of the original pseudodata. This is indeed the case, as seen in Fig. 16 which shows an example of 2D distributions featuring strong correlations. A quantitative determination of the success of the procedure is seen in the calculated pulls in Fig. 17. The highest values of the pulls are randomly scattered in the 2D-distributions (left panel), so no evident systematic trend is present. This is also confirmed by the
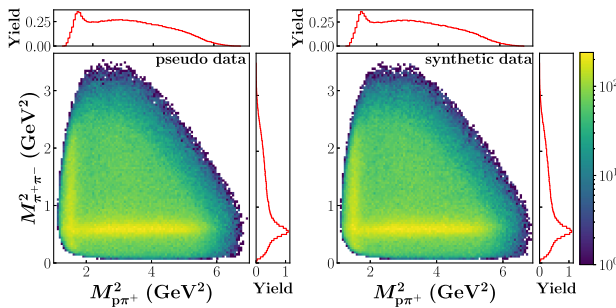

FIG. 16. 2D distributions of $\pi^+\pi^-$ and $p\pi^+$ invariant masses squared in GEN RE-MC pseudodata (left) and synthetic data output of one of the 20 UNF-GANs (right).
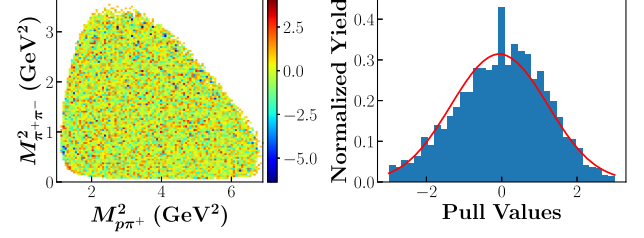

FIG. 17. Left panel: 2D pulls of the distributions shown in Fig. 16. Color palette shows the value of the pull. Right panel: histogram of the 2D pulls with a fitted Gaussian distribution centered at zero with a standard deviation of 1.3.

histogram of the 2D-pulls (right panel), which appears to be normally distributed.

The good agreement and preservation of correlations remains valid for derived kinematic variables that were not used for training. Examples are shown in Fig. 18 for invariant and CM variables, and in Fig. 19 for variables in the lab frame. It is worth noting that in the lab frame the GEN pseudodata exhibits sharp features due to detector acceptance. These features cannot be properly captured by the GANs, which is trained on invariant variables. Even so, this results in a $\lesssim 2\sigma$ local discrepancy in the 1D projections. If better agreement is needed, lab frame variables can be added to the training set.

Finally, in Fig. 20 we compare 1D distributions in a given bin of the other variables. The success of this test shows that correlations underlying the multidifferential cross section are correctly reproduced in the synthetic datasets.
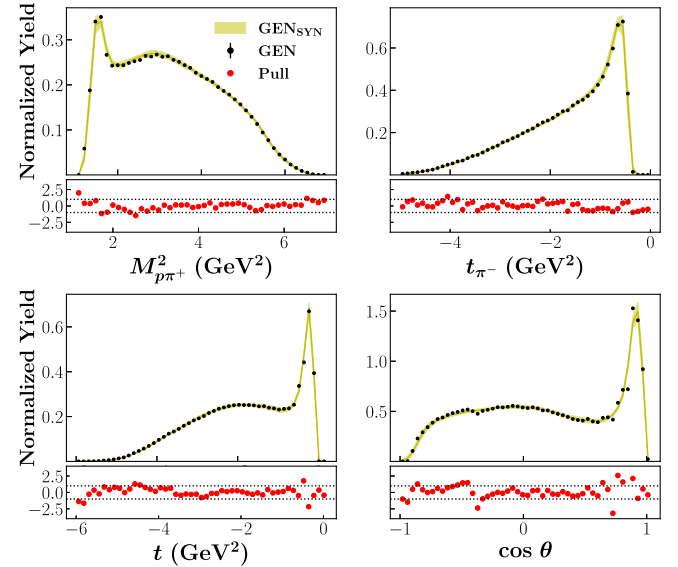

FIG. 18. Comparison of GEN and GEN$_{SYN}$ for four derived variables $M^2_{p\pi^+}$, t_{π^-}, t, and $\cos\theta$. Events were generated using RE-MC.
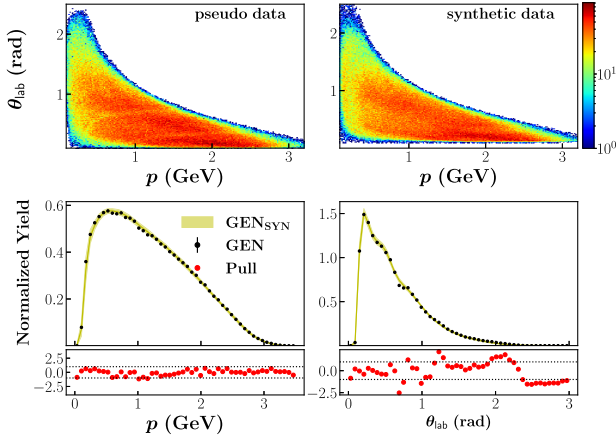
FIG. 19. Comparison of GEN and GEN$_{SYN}$ for p momentum components and θ_{lab} in the laboratory reference frame using RE-MC data.
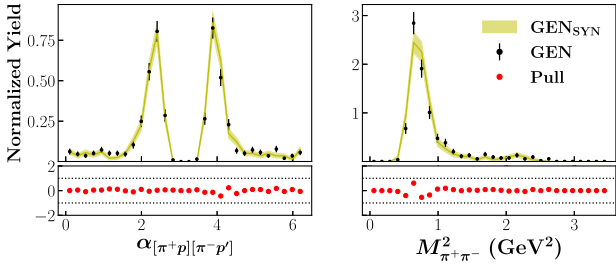

FIG. 20. 1D histograms for fixed slice of the other variables ($2.55 < W < 2.60$ GeV, $-0.7 < t_{\pi^+} < -0.3$ GeV^2, $2.5 < M^2_{p\pi^-} < 3.3$ GeV^2). Left panel: $\alpha_{[\pi^+ p][\pi^- p']}$ distribution for $0.6 < M^2_{\pi^+\pi^-} < 0.9$ GeV^2. Right panel: $M^2_{\pi^+\pi^-}$ distribution for $2.0 < \alpha_{[\pi^+ p][\pi^- p']} < 2.5$.

VI. CONCLUSIONS AND OUTLOOK

One of the central results of this paper is the demonstration that a generative adversarial network can be used to reproduce a realistic multibody physics reaction. As a case study, we have used two-pion photoproduction in the kinematics of the Jefferson Lab CLAS g11 experiment. This process represents an ideal test case, where several baryon and meson production mechanisms overlap, resulting in rich and complex observable distributions. The nonuniformity of the CLAS detector response further adds complication to the challenge.

In order to validate the framework, we have performed a closure test to demonstrate that synthetic data correctly reproduce the multidifferential cross section preserving

correlations between kinematic variables. Detector effects were also correctly unfolded by the procedure. We deployed two MC event generators, one distributed according to pure phase space, and the other incorporating a realistic physics model. Generated pseudodata were fed into a Geant-based detector model to realistically take into account the detector response. Phase-space pseudodata were used to train a GAN-based proxy to learn the detector effects, and realistic pseudodata were then used to train the unfolding GAN and generate synthetic copies of MC events.

The uncertainty quantification of the entire procedure was assessed by combining a bootstrap for the two NNs. Comparison between the true and GAN-generated samples demonstrated that, within the quoted systematic error, the NN is able to reproduce training and derived kinematic variables, as well as to unfold the detector effects in multiple dimensions.

This work represents a first step towards a full AI-supported analysis of CLAS exclusive two-pion photoproduction data. It demonstrates that the same analysis framework, trained on CLAS data, can provide a synthetic copy of the experimental data, preserving correlations between kinematic variables and unfolding the detector effects. Physics interpretation in term of production mechanisms, separating different contributions and extracting resonance parameters from the unfolded data, will follow. An extension of this framework to include the different topologies and extrapolating in a controlled (albeit model-dependent) outside detector acceptance is also in progress.

[1] M. Battaglieri, R. De Vita, V. Kubarovsky, J. Price, P. Stoler, and D. Weygand, Study of pentaquark states in photo-production off protons (2004), http://www.jlab.org/exp_prog/proposals/04/PR04-017.pdf.

[2] A. Andreassen, P. T. Komiske, E. M. Metodiev, B. Nachman, and J. Thaler, Phys. Rev. Lett. **124**, 182001 (2020).

[3] V. Andreev *et al.* (H1 Collaboration), Phys. Rev. Lett. **128**, 132002 (2022).

[4] V. Andreev *et al.* (H1 Collaboration), Phys. Lett. B **844**, 138101 (2023).

[5] R. Aaij *et al.* (LHCb Collaboration), Phys. Rev. D **108**, L031103 (2023).

[6] P. T. Komiske, S. Kryhin, and J. Thaler, Phys. Rev. D **106**, 094021 (2022).

[7] M. Bellagente, A. Butter, G. Kasieczka, T. Plehn, and R. Winterhalder, SciPost Phys. **8**, 070 (2020).

[8] Y. Alanazi, P. Ambrozewicz, M. Battaglieri, A. N. Hiller Blin, M. P. Kuchera, Y. Li, T. Liu, R. E. McClellan, W. Melnitchouk, E. Pritchard, M. Robertson, N. Sato, R. Strauss, and L. Velasco, Phys. Rev. D **106**, 096002 (2022).

[9] R. Brun, F. Bruyant, M. Maire, A. C. McPherson, and P. Zanarini, *Geant: User's Guide Geant 3.10, Geant 3.11; Rev. version* (CERN, Geneva, 1987).

[10] S. Capstick and W. Roberts, Prog. Part. Nucl. Phys. **45**, S241 (2000).

[11] E. Klempt and J.-M. Richard, Rev. Mod. Phys. **82**, 1095 (2010).

[12] M. M. Giannini and E. Santopinto, Chin. J. Phys. (Taipei) **53**, 020301 (2015).

[13] S.-X. Qin, C. D. Roberts, and S. M. Schmidt, Few-Body Syst. **60**, 26 (2019).

[14] C. Chen, G. I. Krein, C. D. Roberts, S. M. Schmidt, and J. Segovia, Phys. Rev. D **100**, 054009 (2019).

[15] R. G. Edwards, J. J. Dudek, D. G. Richards, and S. J. Wallace, Phys. Rev. D **84**, 074508 (2011).

[16] V. Crede and W. Roberts, Rep. Prog. Phys. **76**, 076301 (2013).

[17] D. G. Ireland, E. Pasyuk, and I. Strakovsky, Prog. Part. Nucl. Phys. **111**, 103752 (2020).

[18] R. Beck and U. Thoma, EPJ Web Conf. **134**, 02001 (2017).

[19] V. Burkert, E. Klempt, and U. Thoma, arXiv:2211.12906.

[20] V. I. Mokeev and D. S. Carman (CLAS Collaboration), Few-Body Syst. **63**, 59 (2022

[21] M. Ripani *et al.* (CLAS Collaboration), Phys. Rev. Lett. **91**, 022002 (2003).

[22] E. Golovatch *et al.* (CLAS Collaboration), Phys. Lett. B **788**, 371 (2019).

[23] V. I. Mokeev *et al.*, Phys. Lett. B **805**, 135457 (2020).

[24] J. Nys, V. Mathieu, C. Fernández-Ramírez, A. N. Hiller Blin, A. Jackura, M. Mikhasenko, A. Pilloni, A. P. Szczepaniak, G. Fox, and J. Ryckebusch (JPAC Collaboration), Phys. Rev. D **95**, 034014 (2017).

[25] V. Mathieu, J. Nys, C. Fernández-Ramírez, A. N. Hiller Blin, A. Jackura, A. Pilloni, A. P. Szczepaniak, and G. Fox (JPAC Collaboration), Phys. Rev. D **98**, 014041 (2018).

[26] V. Mathieu, A. Pilloni, M. Albaladejo, L. Bibrzycki, A. Celentano, C. Fernández-Ramírez, and A. P. Szczepaniak (JPAC Collaboration), Phys. Rev. D **102**, 014003 (2020).

[27] V. Mathieu, J. Nys, C. Fernández-Ramírez, A. Jackura, A. Pilloni, N. Sherrill, A. P. Szczepaniak, and G. Fox (JPAC Collaboration), Phys. Rev. D **97**, 094003 (2018).

[28] M. Albaladejo *et al.* (JPAC Collaboration), Prog. Part. Nucl. Phys. **127**, 103981 (2022).

[29] P. Pauli (GlueX Collaboration), Rev. Mex. Fis. Suppl. **3**, 0308002 (2022).

[30] M. Battaglieri *et al.* (CLAS Collaboration), Phys. Rev. D **80**, 072005 (2009).

[31] S. Strauch *et al.* (CLAS Collaboration), Phys. Rev. Lett. **95**, 162003 (2005).

[32] M. Battaglieri *et al.* (CLAS Collaboration), Phys. Rev. Lett. **87**, 172002 (2001).

[33] R. Aaij *et al.* (LHCb Collaboration), Phys. Rev. Lett. **127**, 082001 (2021).

[34] R. Aaij *et al.* (LHCb Collaboration), Phys. Rev. Lett. **131**, 031901 (2023).

[35] M. D. Mestayer *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **449**, 81 (2000).

[36] E. S. Smith *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **432**, 265 (1999).

[37] D. I. Sober *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **440**, 263 (2000).

[38] M. R. Whalley, Comput. Phys. Commun. **57**, 536 (1989).

[39] K. Schilling, P. Seyboth, and G. E. Wolf, Nucl. Phys. **B15**, 397 (1970); **B18**, 332(E) (1970).

[40] Gsim user's guide, https://www.jlab.org/Hall-B/document/gsim/userguide.html.

[41] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, in *Advances in Neural Information Processing Systems*, edited by Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K. Q. Weinberger (Curran Associates, Inc., Red Hook, NY, 2014), Vol. 27, https://proceedings.neurips.cc/paper_files/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf.

[42] T. Karras, S. Laine, and T. Aila, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2019), Vol. 43, pp. 4401–4410, arXiv:1812.04948.

[43] Y. Li, Q. Pan, S. Wang, T. Yang, and E. Cambria, Inf. Sci. **450**, 301 (2018).

[44] G. Brunner, Y. Wang, R. Wattenhofer, and S. Zhao, in *Proceedings of the 2018 IEEE 30th International Conference on Tools with Artificial Intelligence (ICTAI)* (IEEE, New York, 2018), pp. 786–793, arXiv:1809.07575.

[45] A. Clark, J. Donahue, and K. Simonyan, arXiv:1907.06571.

[46] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen, in *Proceedings of the Advances in Neural Information Processing Systems 29 (NIPS 2016)* (2016), Vol. 29, arXiv:1606.03498.

[47] S. Arora and Y. Zhang, arXiv:1706.08224.

[48] M. Arjovsky and L. Bottou, arXiv:1701.04862.

[49] D. Bang and H. Shim, in *Proceedings of the IEEE/CVF International Conference on Computer Vision* (2021), pp. 2347–2356, arXiv:1804.04391.

[50] C. Gao, S. Höche, J. Isaacson, C. Krause, and H. Schulz, Phys. Rev. D **101**, 076002 (2020).

[51] S. Otten, S. Caron, W. de Swart, M. van Beekveld, L. Hendriks, C. van Leeuwen, D. Podareanu, R. Ruiz de Austri, and R. Verheyen, Nat. Commun. **12**, 2985 (2021).

[52] M. Paganini, L. de Oliveira, and B. Nachman, Phys. Rev. D **97**, 014021 (2018).

[53] Y. Alanazi, N. Sato, T. Liu, W. Melnitchouk, P. Ambrozewicz, F. Hauenstein, M. P. Kuchera, E. Pritchard, M. Robertson, R. Strauss, L. Velasco, and Y. Li, in *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21* (International Joint Conferences on Artificial Intelligence Organization, Menlo Park, CA, 2021), p. 2126.

[54] B. Hashemi, N. Amin, K. Datta, D. Olivito, and M. Pierini, arXiv:1901.05282.

[55] A. Butter, T. Plehn, and R. Winterhalder, SciPost Phys. **7**, 075 (2019).

[56] R. Di Sipio, G. M. Faucci, H. S. Ketabchi, and S. Palazzo, J. High Energy Phys. 08 (2020) 110.

[57] M. Paganini, L. de Oliveira, and B. Nachman, Phys. Rev. Lett. **120**, 042003 (2018).

[58] L. de Oliveira, M. Paganini, and B. Nachman, Comput. Software Big Sci. **1**, 4 (2017).

[59] P. Musella and F. Pandolfi, Comput. Software Big Sci. **2**, 8 (2018).

[60] M. Bellagente, A. Butter, G. Kasieczka, T. Plehn, and R. Winterhalder, SciPost Phys. **8**, 070 (2020).

[61] Y. Alanazi, N. Sato, P. Ambrozewicz, A. Hiller-Blin, W. Melnitchouk, M. Battaglieri, T. Liu, and Y. Li, in *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21* (International Joint Conferences on Artificial Intelligence Organization, Menlo Park, CA, 2021), pp. 4286–4293.

[62] X. Mao, Q. Li, H. Xie, R. Y. K. Lau, Z. Wang, and S. P. Smolley, in *Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV)* (2017), pp. 2813–2821, arXiv:1611.04076.

[63] M. Mirza and S. Osindero, arXiv:1411.1784.

[64] D. P. Kingma and J. Ba, in *Proceedings of the 3rd International Conference for Learning Representations (ICLR)* (2014), arXiv:1412.6980.