

Developing AI tools for hadron spectroscopy

GIORGIO FOTI (UNIME & INFN
CATANIA)

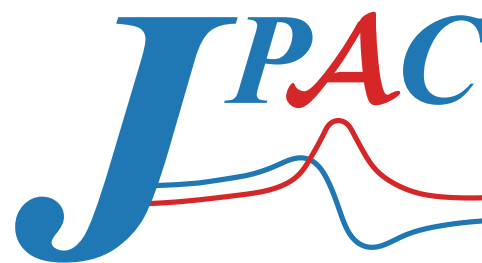
WORKSHOP JPAC & JLAB12

FEB 18 - 20, 2026

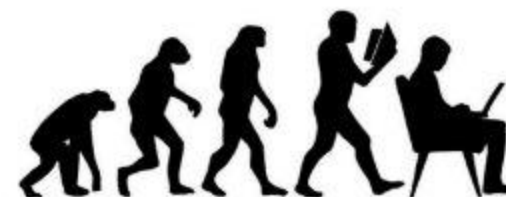
DIPARTIMENTO DI FISICA,
UNIVERSITÀ DI GENOVA



Università
degli Studi di
Messina



Istituto Nazionale di Fisica Nucleare



A(i)DAPT

AI for Data Analysis and PreservaTion

Summary

- Unfolding the detector effect through generative AI
- How a diffusion model work
- Status of closure test for CLAS detector
- Conclusions & Next directions



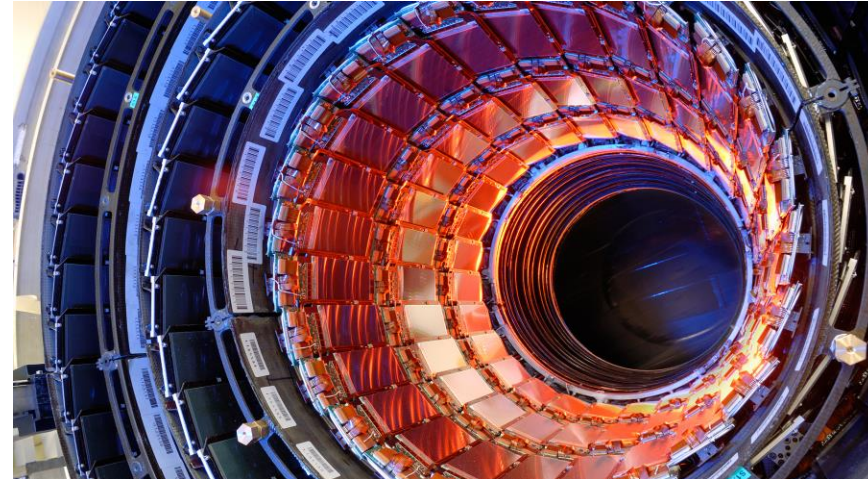
The response of a detector

Detector introduce resolution effects, acceptance gaps and inefficiencies that distort the true particle distributions.

$$\text{det}(y) = \int_x P(y|x) \text{true}(x)$$

Ground truth
("vertex" level):
E = 10.0 GeV

Detector level:
E=10.1 GeV



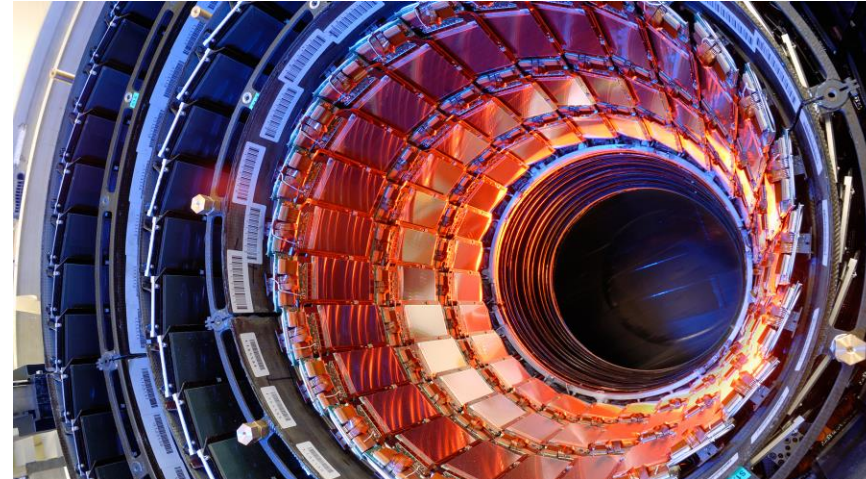
Unfolding the detector effect

We want to infer the posterior probability $P(x|y)$ given the measured $\text{det}(y)$ ("unfolding"). Prior $\text{true}(x)$ is unknown: this is an ill-posed inverse problem

$$P(x|y) = \frac{P(y|x) \text{true}(x)}{\text{det}(y)}$$

Ground truth
("vertex" level):
E = 10.0 GeV

Detector level:
E=10.1 GeV



Traditional Unfolding = matrix inversion

$$\det(y) = \int_x P(y|x) \text{true}(x)$$



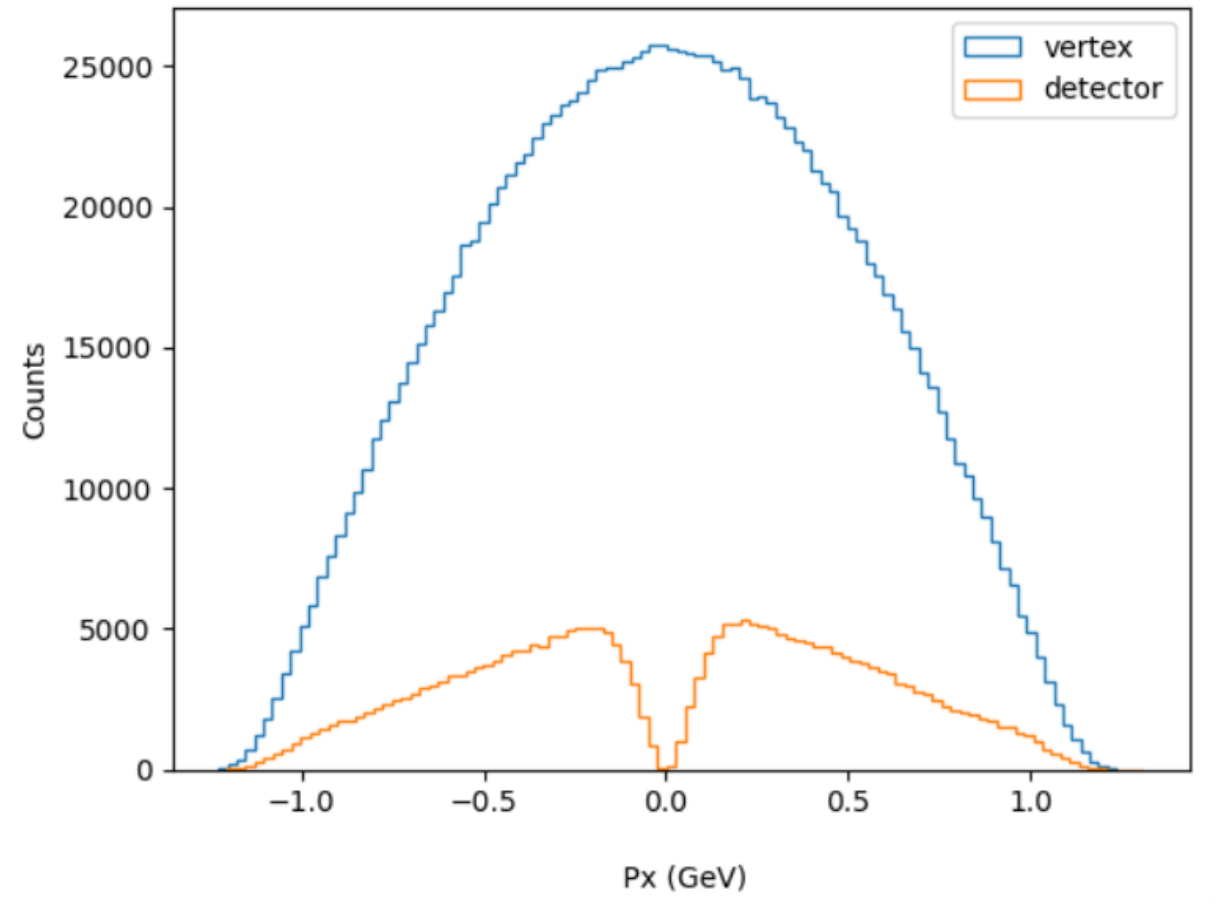
Binning

$$\det_i = \sum_j M_{ij} \text{true}_j$$



Response matrix inversion

$$\text{true}_i = \sum_j M_{ij}^{-1} \det_j$$



Limitations of Traditional Unfolding

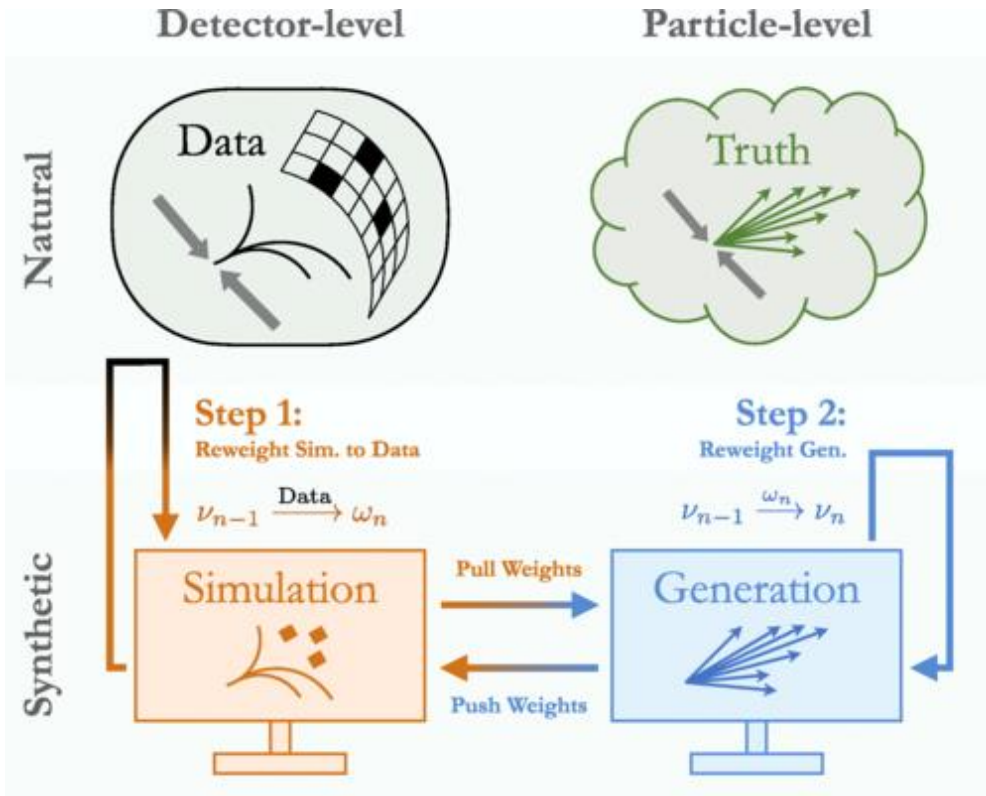
[Unfolding Methods
in Particle Physics for Quantum
Information Observables
Michele Pinamonti
(https://agenda.infn.it/event/34555/contributions/212423/attachments/112947/161553/Unfolding_Pinamonti10nov2023.pdf)]

- Standard techniques typically requires **binning** data and assume a **specific response matrix** (e.g. linear)
- **Background** need to be subtracted before unfolding (misidentification & other reducible sources – irreducible ones are eliminated with the successive physical analysis)
- Struggle with **systematics uncertainties**: they enter in the response matrix choice and in the background estimation
- Bad times with unfolding **many correlated observables**
- For many algorithms, different hypotheses for **underlying physics** could in general change the response function
- Large "**oscillations**" in the inversion: amplification of statistical fluctuations due to off-diagonal entries of M (many different ways to mitigate them)



These challenges motivate alternative approaches, like **AI-based unfolding**, which can handle high-dimensional input (e.g. full event kinematics or images) and learn complex detector effects without manual binning.

ML-based alternative on the market



E.g. OmniFold: A Method to Simultaneously Unfold All Observables

Pros:

- Based on iterative reweighting of a simulated dataset, using machine learning to capitalize on all available information
- Unbinned approach
- Works for arbitrarily high-dimensional data, and naturally incorporates information from the full phase space

Cons:

- Tuned on LHC experiments, typically much more inclusive than ours of hadronic physics
- We need something more suitable for exclusive experiments like CLAS (GlueX, Compass...)

Generative AI-Based Unfolding Approaches

Toward a generative modeling analysis of CLAS exclusive 2π photoproduction

T. Alghamdi^{1,2,*}, Y. Alanazi³, M. Battaglieri⁴, L. Bibrzycki⁵, A. V. Golda⁶,
A. N. Hiller Blin⁷, E. L. Isupov⁸, Y. Li¹, L. Marsicano⁴, W. Melnitchouk^{3,9}, V. I. Mokeev³,
G. Montaña³, A. Pilloni^{10,11}, N. Sato³, A. P. Szczepaniak^{3,12,13} and T. Vittorini^{4,14}

- Adversarial and likelihood-based models have been applied: e.g. in A(i)DAPT, generative adversarial networks (GANs), (conditional) normalizing flows
- Diffusion models: recent work show that they are **outperforming GANs**. These can, in principle, model complex posteriors and provide uncertainty quantification.

Each AI method has trade-offs; diffusion models are attractive for their strong theoretical foundations, **flexibility and bijectivity**

Summary

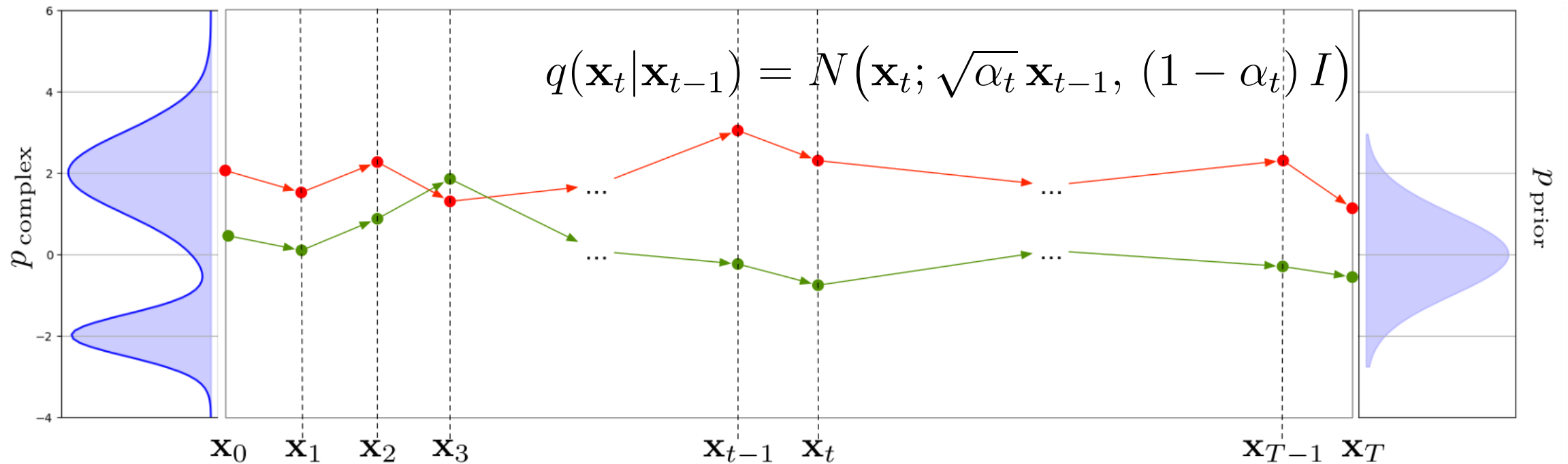
- Unfolding the detector effect through generative AI
- How a diffusion model work
- Status of closure test for CLAS detector
- Conclusions & Next directions



Denoising Diffusion Probabilistic Models (DMs)

A DM is a generative model is trained in two phases: a **deterministic forward diffusion** (noise addition) process and a **learned reverse** (denoising) process:

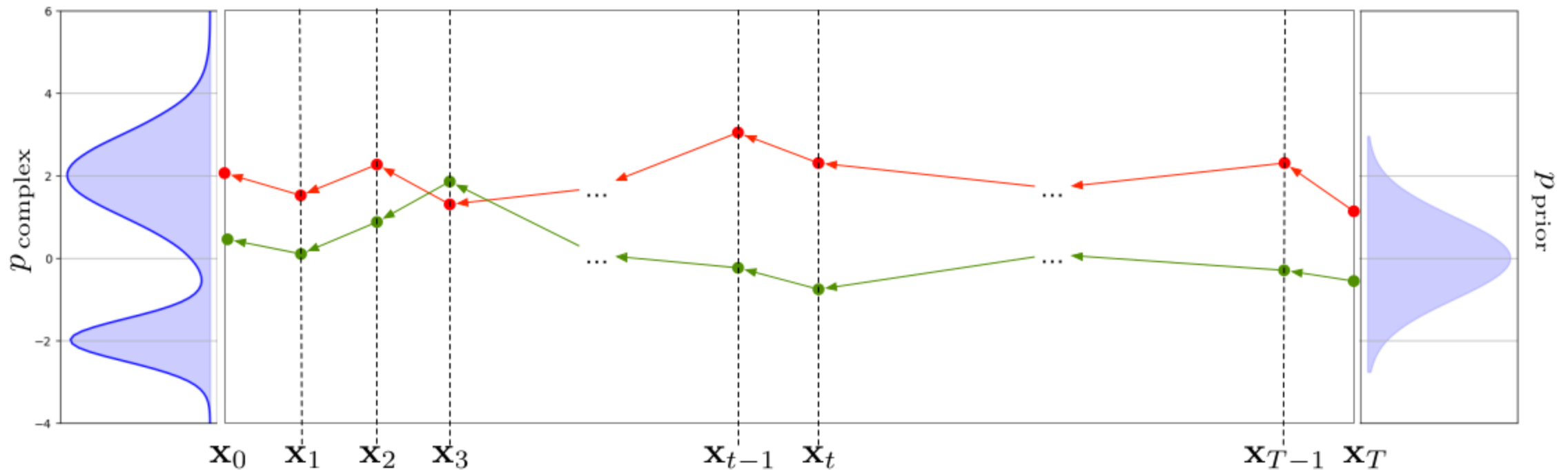
- In the forward pass, one gradually corrupts a clean data sample x_0 by injecting Gaussian noise over T timesteps, producing $x_T \approx N(0, I)$ (α_t is a variance schedule):



Denoising Diffusion Probabilistic Models (DMs)

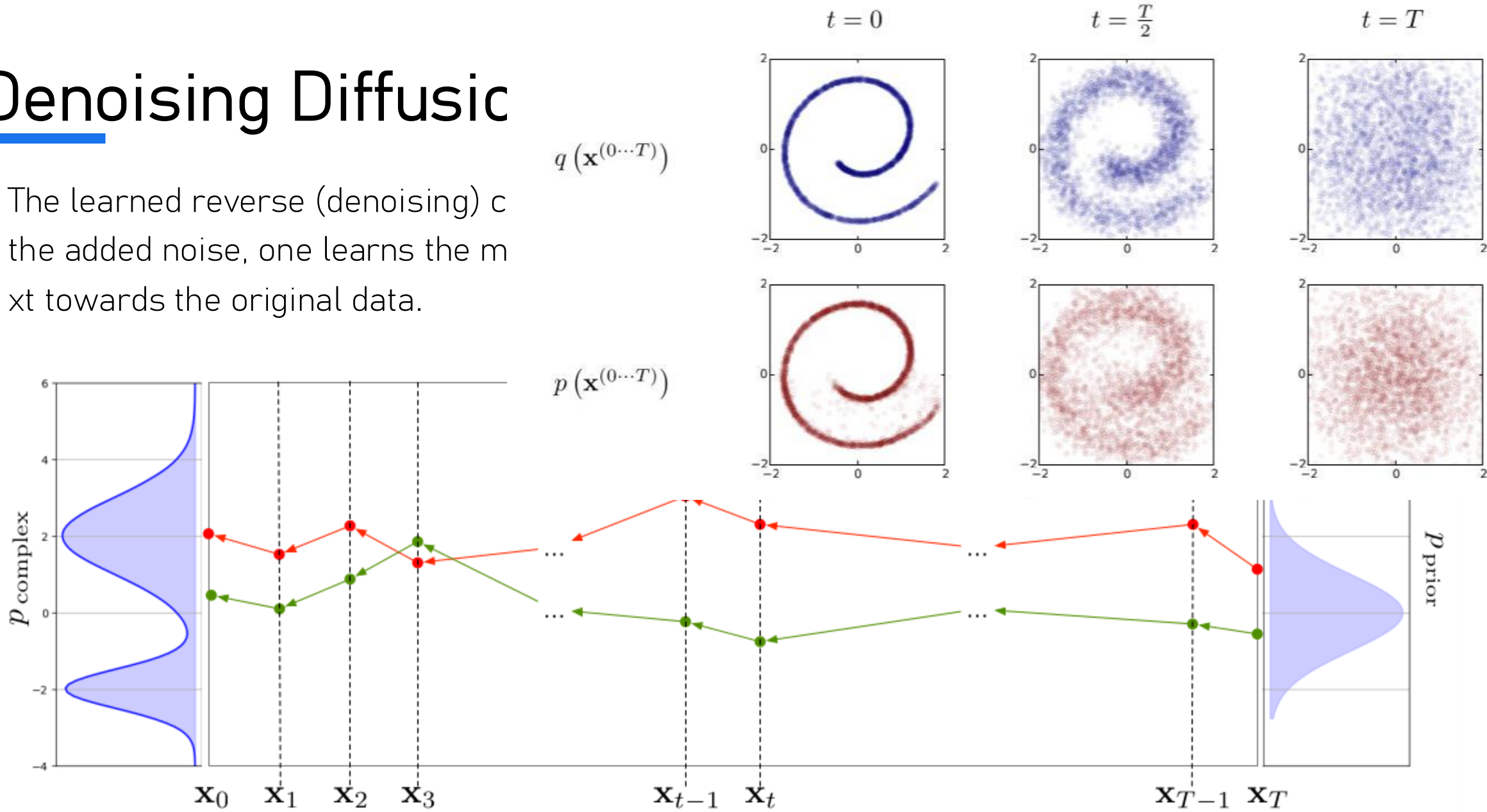
- The learned reverse (denoising) chain is parameterized by θ : by training the model to predict the added noise, one learns the mean of the Gaussian transition, $\mu_\theta(\mathbf{x}_t, t)$, that maps any noisy \mathbf{x}_t towards the original data.

$$p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t) = N(\mathbf{x}_{t-1}; \mu_\theta(\mathbf{x}_t, t), \sigma_t^2 I)$$



Denoising Diffusion

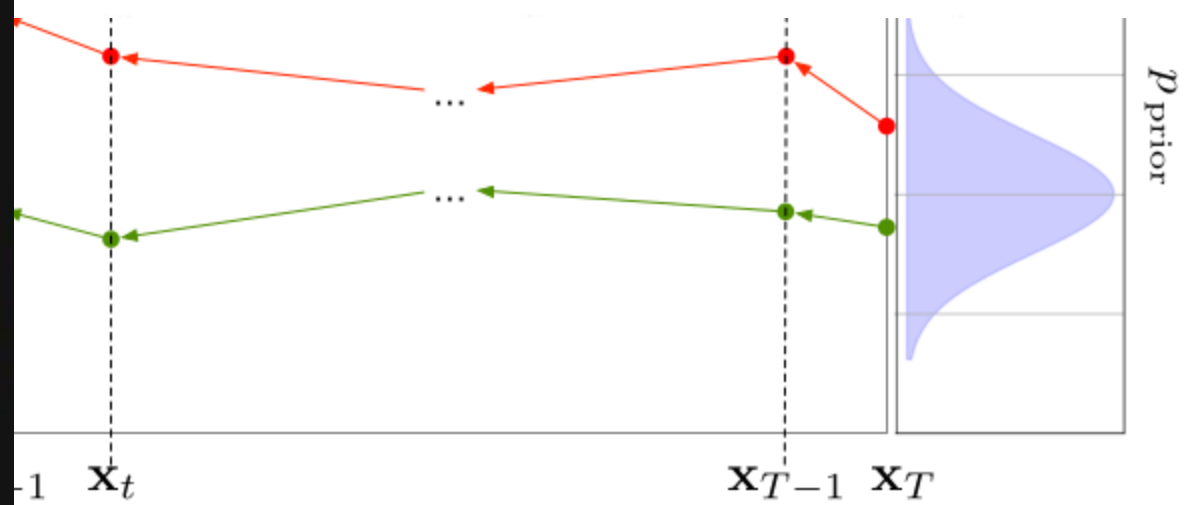
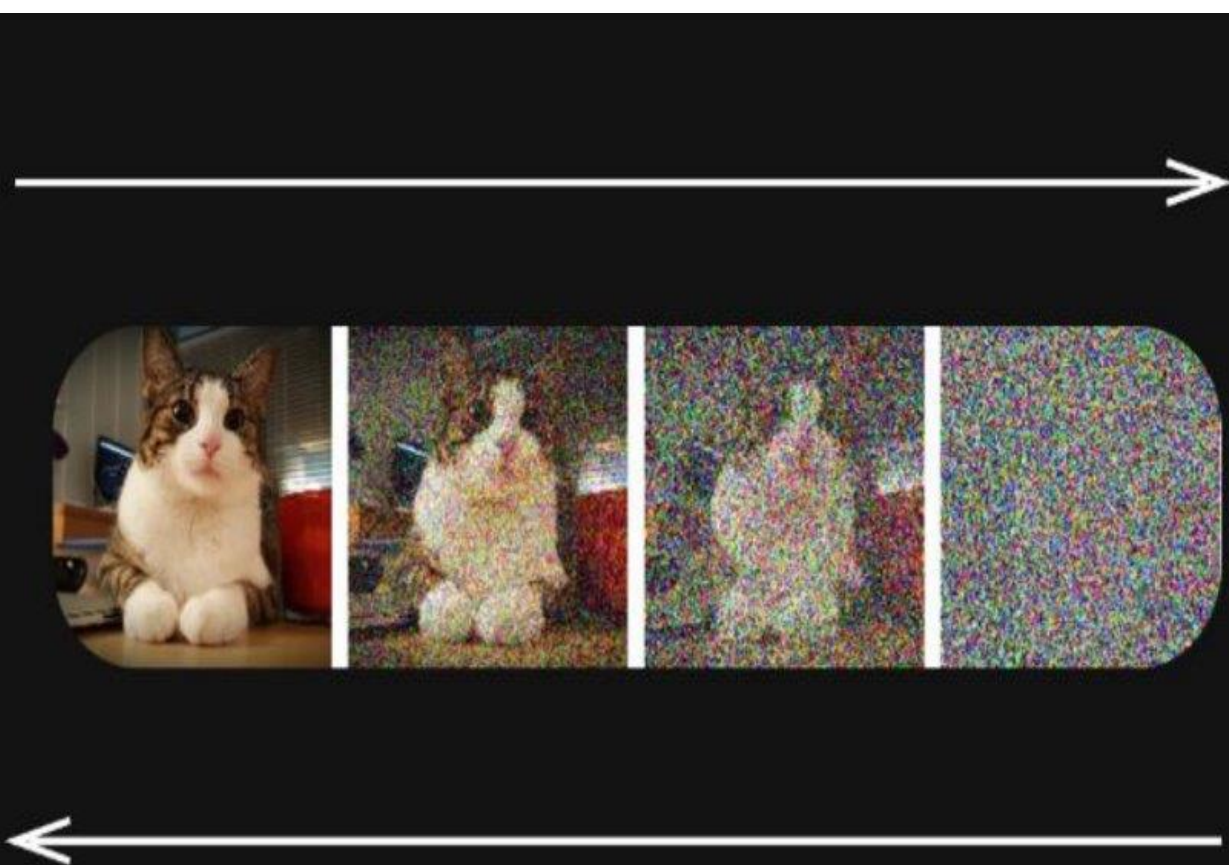
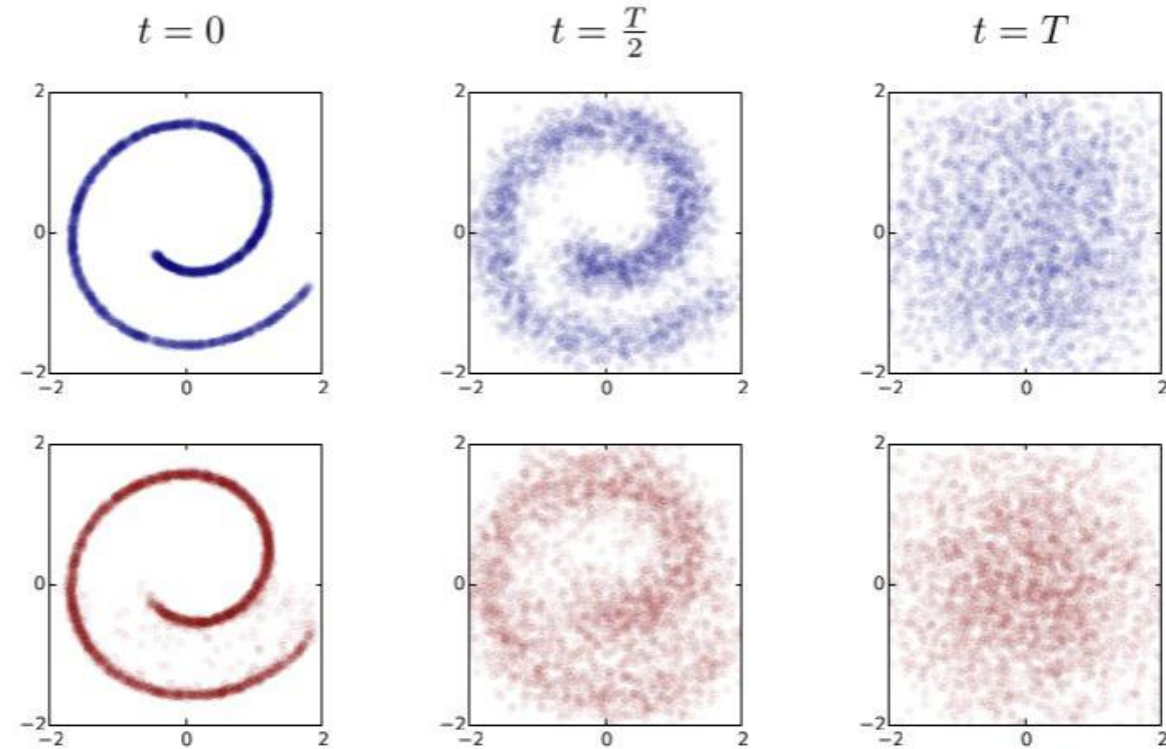
- The learned reverse (denoising) process, by adding noise, one learns the model towards the original data.



Denoising Diffusion

- The learned reverse (denoising) process removes the added noise, one learns the model towards the original data.

$$q(\mathbf{x}^{(0:T)})$$



Diffusion for Unfolding: conditioning

To apply diffusion to unfolding, we use a **conditional diffusion model**. Here the reverse process is conditioned on the detector level y :

$$p_{\theta}(x_{t-1}|x_t, y) = N(x_{t-1}; \mu_{\theta}(x_t, y, t), \sigma_t^2 I)$$

- The forward noising process remains unchanged, but **the denoiser network now takes y as an input or context** (this context can be given by labels, features etc.)
- In practice, the output x_0 is a sample from the posterior $P(x|y)$. Conditioning directly on y makes the **generative process data-dependent** and enables it to **approximate the Bayesian inverse** $P(x|y)$ without explicitly computing the prior

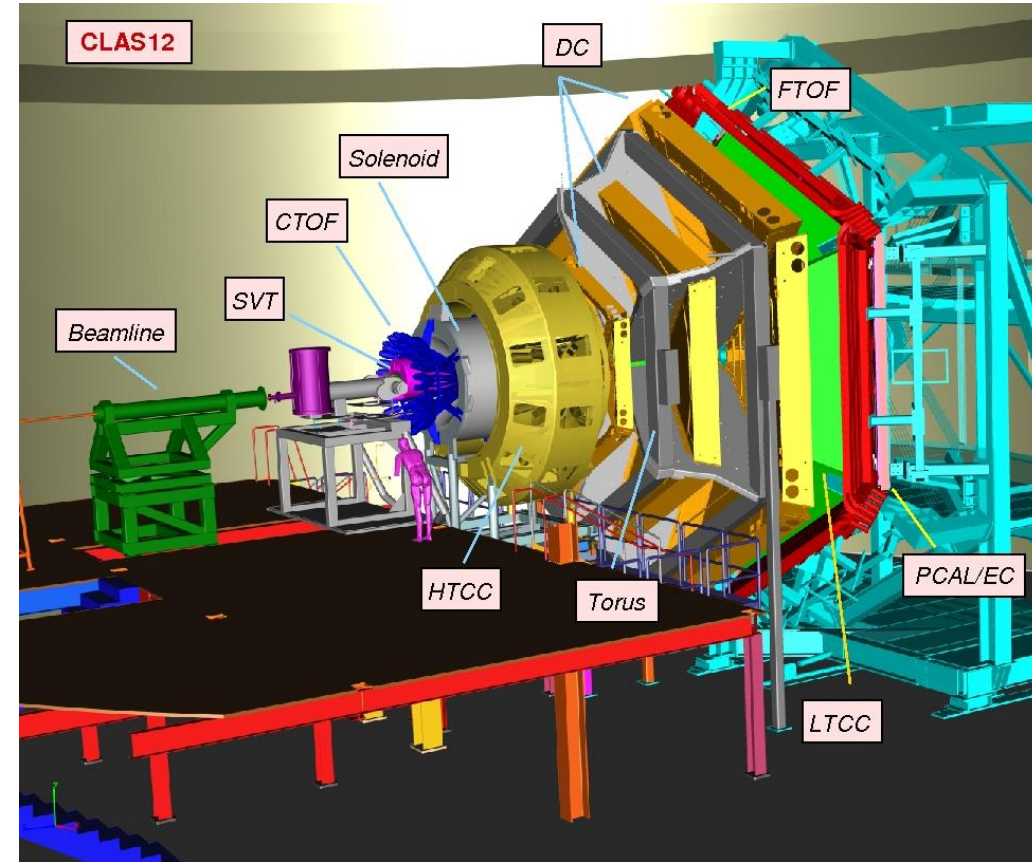
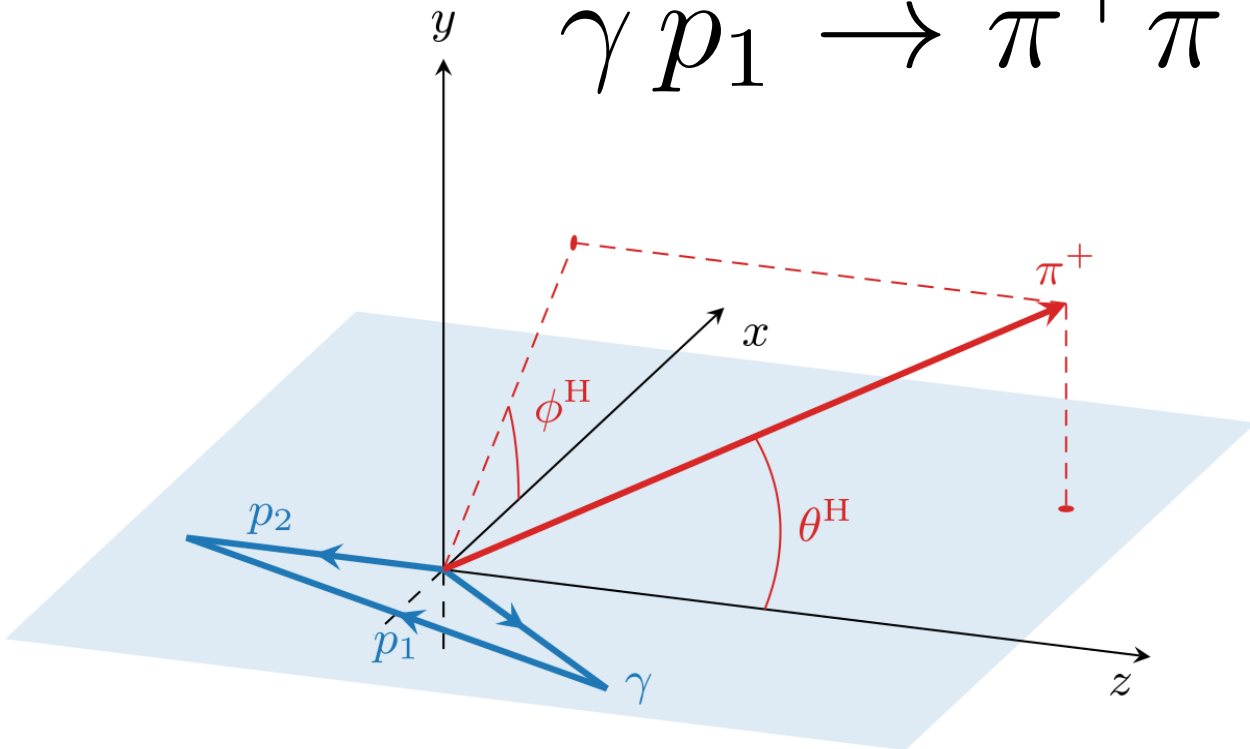
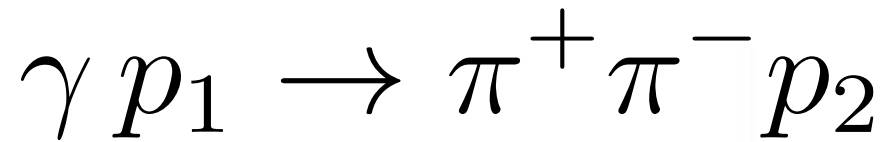
Summary

- Unfolding the detector effect through generative AI
- How a diffusion model work
- Status of closure test for CLAS detector
- Conclusions & Next directions



A closure test for unfolding CLAS detector

Photoproduction of two pions on the proton, measured by CLAS



Let's define the moments

$$\frac{d\sigma}{dt d\sqrt{s_{12}} d\Omega^H} = \kappa \sum_{\lambda_1 \lambda_\gamma \lambda_2} |\mathcal{M}_{\lambda_\gamma \lambda_1 \lambda_2}(s, t, s_{12}, \Omega^H)|^2,$$

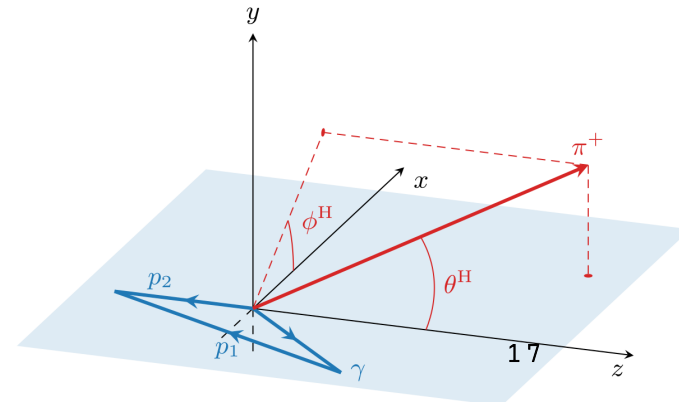
$$\langle Y_M^L \rangle = \sqrt{4\pi} \int d\Omega^H \frac{d\sigma}{dt d\sqrt{s_{12}} d\Omega^H} \text{Re}\{Y_M^L(\Omega^H)\}$$

$$\mathcal{M}_{\lambda_\gamma \lambda_1 \lambda_2}(s, t, s_{12}, \Omega^H) = \sum_{lm} \mathcal{M}_{\lambda_\gamma \lambda_1 \lambda_2 m}^l(s, t, s_{12}) Y_m^l(\Omega^H),$$

$$\begin{aligned} \langle Y_M^L \rangle &= \sqrt{4\pi} \kappa \sum_{lm'l'm'} A_{Mmm'}^{Ll'l'} \\ &\times \sum_{\lambda_\gamma \lambda_1 \lambda_2} \mathcal{M}_{\lambda_\gamma \lambda_1 m' \lambda_2}^{l'*}(s, t, s_{12}) \mathcal{M}_{\lambda_\gamma \lambda_1 m \lambda_2}^l(s, t, s_{12}), \end{aligned}$$

where

$$A_{Mmm'}^{Ll'l'} = \int d\Omega^H Y_m^l(\Omega^H) Y_{m'}^{l'*}(\Omega^H) \text{Re}\{Y_M^L(\Omega^H)\}.$$



Let's define the moments

Photoproduction of $\pi^+\pi^-$ meson pairs on the proton

Moment extraction
from real data by
laboration

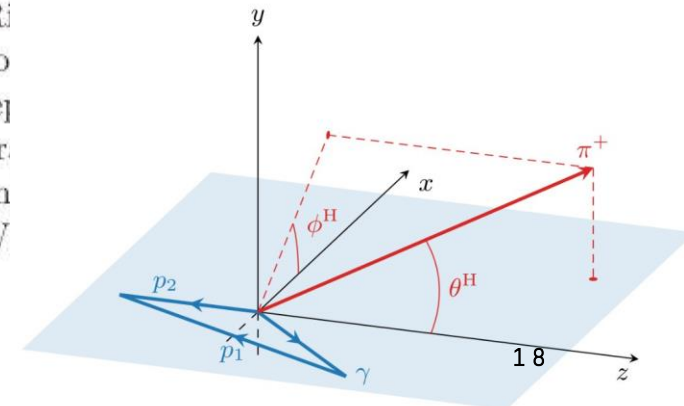
$\frac{d\sigma}{d\Omega^H}$

M. Battaglieri,¹ R. De Vita,¹ A. P. Szczepaniak,² K. P. Adhikari,³⁵ M.J. Amarian,³⁵ M. Anghinolfi,¹
 H. Baghdasaryan,⁴⁵ I. Bedlinskiy,²² M. Bellis,⁷ L. Bibrzycki,²⁹ A.S. Biselli,^{13,36} C. Bookwalter,¹⁵
 D. Branford,¹² W.J. Briscoe,¹⁶ V.D. Burkert,⁴² S.L. Careccia,³⁵ D.S. Carman,⁴² E. Clinton,²⁸ P.L. Cole,¹⁸
 P. Collins,⁴ V. Crede,¹⁵ D. Dale,¹⁸ A. D'Angelo,^{20,38} A. Daniel,³⁴ N. Dashyan,⁴⁷ E. De Sanctis,¹⁹ A. Deur,⁴²
 S. Dhamija,¹⁴ C. Djalali,⁴¹ G.E. Dodge,³⁵ D. Doughty,^{10,42} V. Drozdov,¹ H. Egiyan,^{30,42} P. Eugenio,¹⁵
 G. Fedotov,⁴⁰ S. Fegan,¹⁷ A. Fradi,²¹ M.Y. Gabrielyan,¹⁴ L. Gan,³² M. Garçon,⁹ A. Gasparian,³³ G.P. Gilfoyle,³⁷
 K.L. Giovanetti,²⁴ F.X. Girod,^{9,*} O. Glamazdin,²⁶ J. Goett,³⁶ J.T. Goetz,⁵ W. Gohn,¹¹ E. Golovatch,^{40,1}
 R.W. Gothe,⁴¹ K.A. Griffioen,⁴⁶ M. Guidal,²¹ L. Guo,^{42,†} K. Hafidi,³ H. Hakobyan,^{44,47} C. Hanretty,¹⁵
 N. Hassall,¹⁷ K. Hicks,³⁴ M. Holtrop,³⁰ C.E. Hyde,³⁵ Y. Ilieva,^{41,16} D.G. Ireland,¹⁷ E.L. Isupov,⁴⁰ J.R. Johnstone,¹⁷
 K. Joo,¹¹ D. Keller,³⁴ M. Khandaker,³¹ P. Khetarpal,³⁶ W. Kim,²⁷ A. Klein,³⁵ F.J. Klein,⁸ M. Kossov,²²
 A. Kubarovsky,³⁵ V. Kubarovsky,⁴² S.V. Kuleshov,^{44,22} V. Kuznetsov,²⁷ J.M. Laget,^{42,9} L. Lesniak,²⁹
 K. Livingston,¹⁷ H.Y. Lu,⁴¹ M. Mayer,³⁵ M.E. McCracken,⁷ B. McKinnon,¹⁷ C.A. Meyer,⁷ K. Mikhailov,²²
 T. Mineeva,¹¹ M. Mirazita,¹⁹ V. Mochalov,²³ V. Moiseev,^{40,42} K. Moriya,⁷ E. Munevar,¹⁶ P. Nadel-Turonski,⁸
 I. Nakagawa,³⁹ C.S. Nepali,³⁵ S. Niccolai,²¹ I. Niculescu,²⁴ M.R. Niroula,³⁵ M. Osipenko,^{1,40} A.I. Ostrovidov,¹⁵
 K. Park,^{41,27,*} S. Park,¹⁵ M. Paris,^{16,42} E. Pasyuk,⁴ S. Anefalos Pereira,¹⁹ S. Pisano,²¹ N. Pivnyuk,²²
 O. Pogorelko,²² S. Pozdniakov,²² J.W. Price,⁶ Y. Prok,^{45,‡} D. Protopopescu,¹⁷ B.A. Raue,^{14,42} G. Ri
 M. Ripani,¹ B.G. Ritchie,⁴ G. Rosner,¹⁷ P. Rossi,¹⁹ F. Sabatié,⁹ M.S. Saini,¹⁵ C. Salgado,³¹ D. Scho
 R.A. Schumacher,⁷ H. Seraydaryan,³⁵ Y.G. Sharabian,⁴² D.I. Sober,⁸ D. Sokhan,¹² A. Stavinsky,²² S. Stej
 S. S. Stepanyan,²⁷ P. Stoler,³⁶ I.I. Strakovsky,¹⁶ S. Strauch,^{41,16} M. Taiuti,¹ D.J. Tedeschi,⁴¹ A. Teymur
 S. Tkachenko,³⁵ M. Ungaro,^{11,36} M.F. Vineyard,⁴³ A.V. Vlassov,²² D.P. Watts,^{17,§} L.B. Weinstein
 D.P. Weygand,⁴² M. Williams,⁷ E. Wolin,⁴² M.H. Wood,⁴¹ L. Zana,³⁰ J. Zhang,³⁵ B. Zhao,^{11,¶} and Z.W

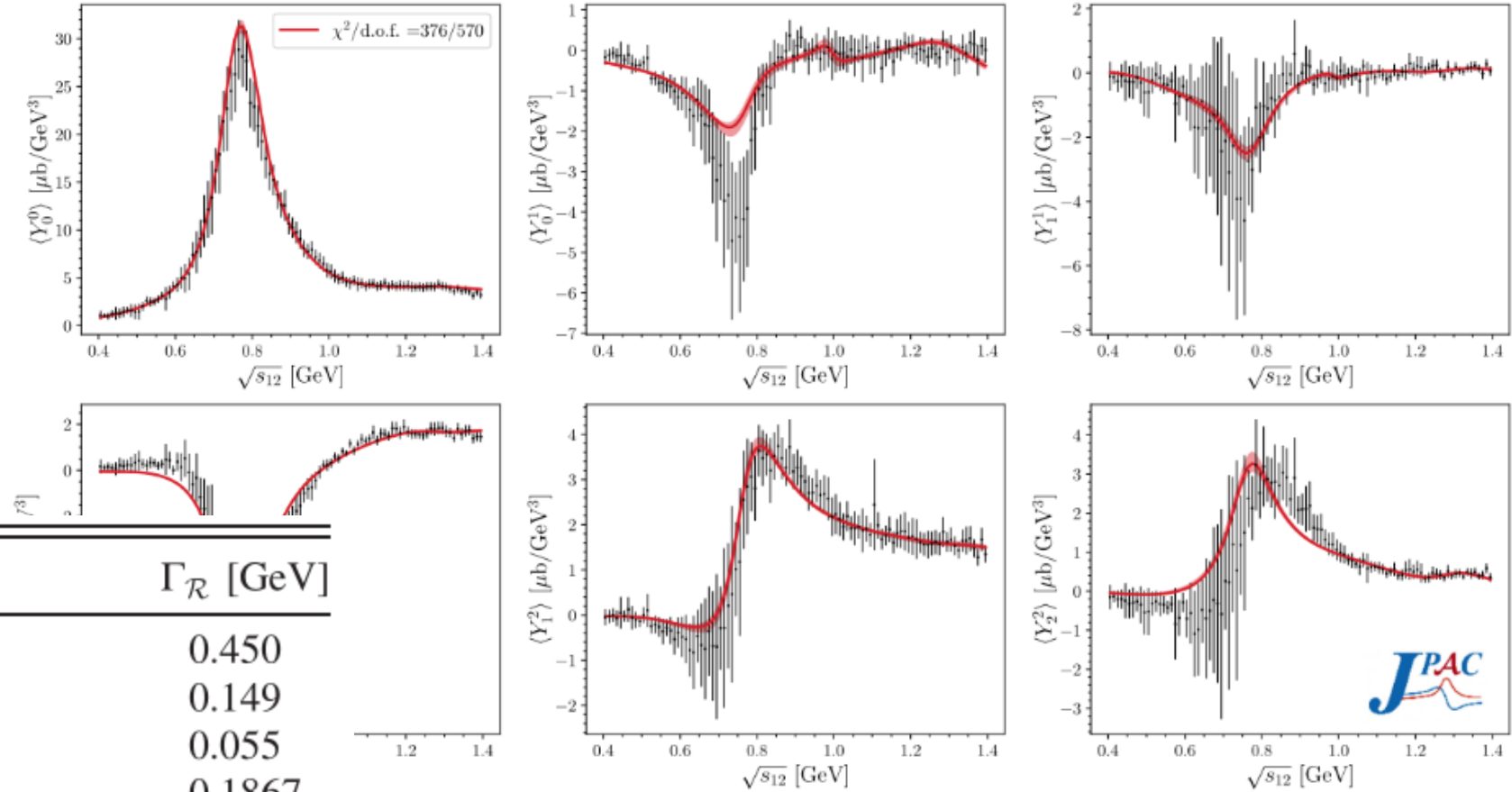
(The CLAS Collaboration)

$$A_{Mmm'}^{Ll'l'} = \int d\Omega^H Y_m^l(\Omega^H) Y_{m'}^{l'*}(\Omega^H) \text{Re}\{Y_M^L(\Omega^H)\}.$$

$$\frac{d\sigma}{2d\Omega^H} \text{Re}\{Y_M^L(\Omega^H)\}$$



Fit results by JPAC collaboration



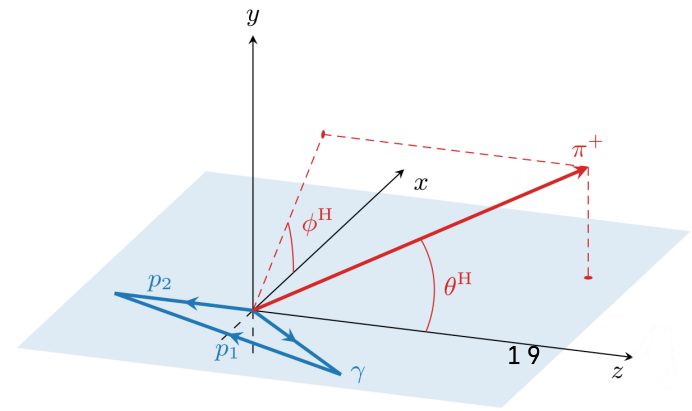
\mathcal{R}	J	$m_{\mathcal{R}}$ [GeV]	$\Gamma_{\mathcal{R}}$ [GeV]
$f_0(500)$	0	0.500	0.450
$\rho(770)$	1	0.775	0.149
$f_0(980)$	0	0.990	0.055
$f_2(1270)$	2	1.2755	0.1867
$f_0(1370)$	0	1.370	0.350

te model fitted to experimental measurements from Ref. [21] of two-pion angular moments $\langle Y_M^L \rangle$ for $E_\gamma = 3.7$ GeV and $t = -0.45$ GeV². Since all data shown here are fit simultaneously, this corresponds to 600 – 30 = 570 degrees of freedom (d.o.f.).

Studying $\pi^+ \pi^-$ photoproduction beyond Pomeron exchange

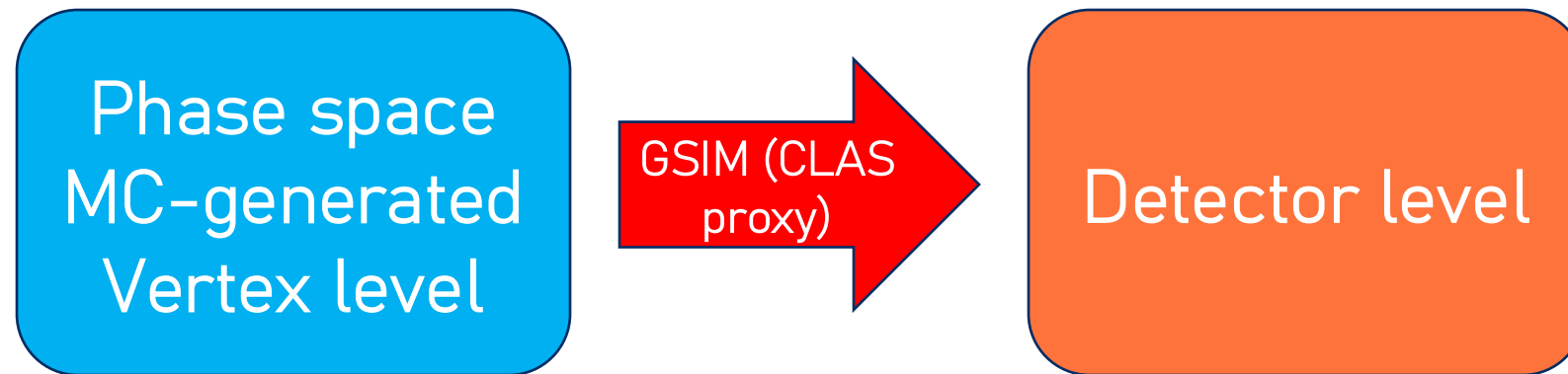
Łukasz Bibrzycki^{1,*}, Nadine Hammoud^{2,3,†}, Vincent Mathieu³, Robert J. Perry^{3,‡}, Alex Akridge^{4,5},
 César Fernández-Ramírez⁶, Gloria Montaña⁷, Alessandro Pilloni^{8,9}, Arkaitz Rodas^{7,10}, Vanamali Shastry^{4,5},
 Wyatt A. Smith^{11,12,13}, Daniel Winney¹⁴ and Adam P. Szczepaniak^{7,5,4}

(Joint Physics Analysis Center)

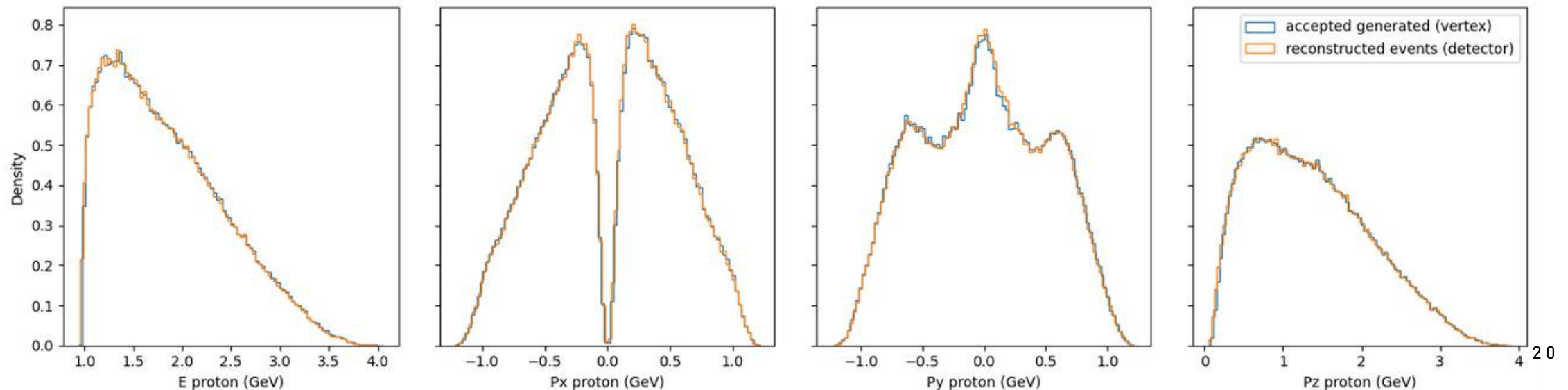


Our strategy: generating data

Using MC generated 4-vectors sampled from different models and correspondent GSIM (the *father* of Geant4) reconstructed and train the DM to learn how to do the inverse process:

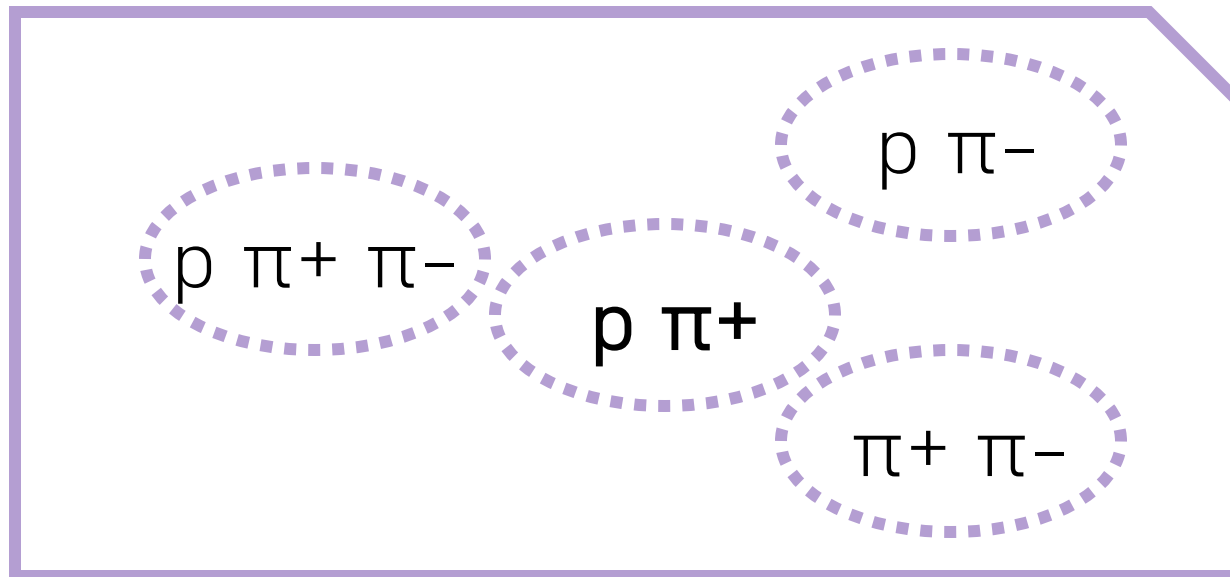


Alert: we are learning just smearing, not acceptance



Our strategy: generating data

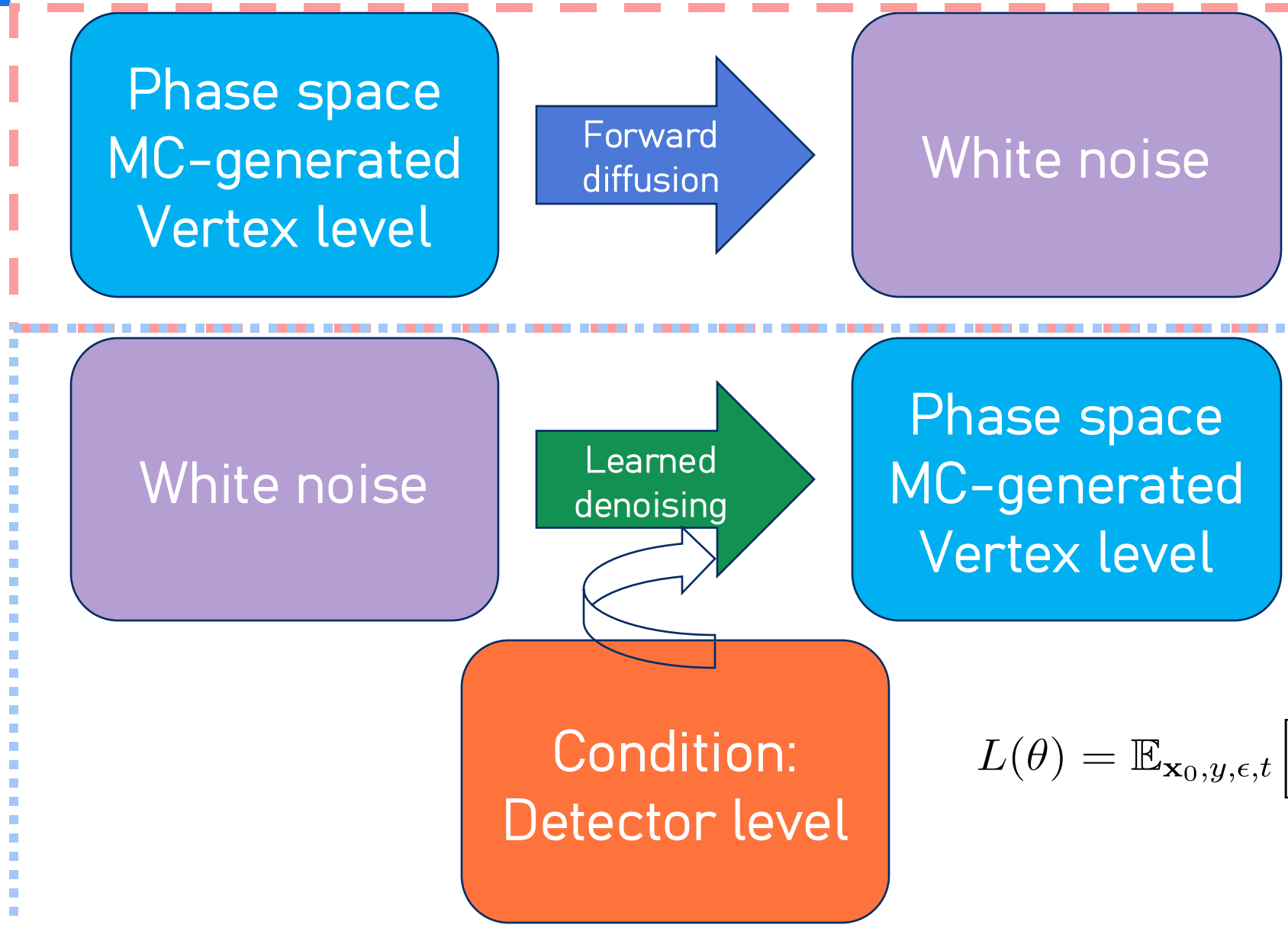
- Let's start from the most populated topology ($\rho \pi^+$)



Phase space region covered by CLAS

- Then we will exploit DMs flexibility for moving from one topology to the other

Our strategy: training the Conditional DM



Pseudocode for training a DM

- Start with a batch of detector-vertex pairs (x_0, y)
- Sample a random t and noise $\epsilon \sim N(0, I)$
- Compute

$$\mathbf{x}_t = \sqrt{\alpha_t} \mathbf{x}_0 + \sqrt{1 - \alpha_t} \epsilon$$

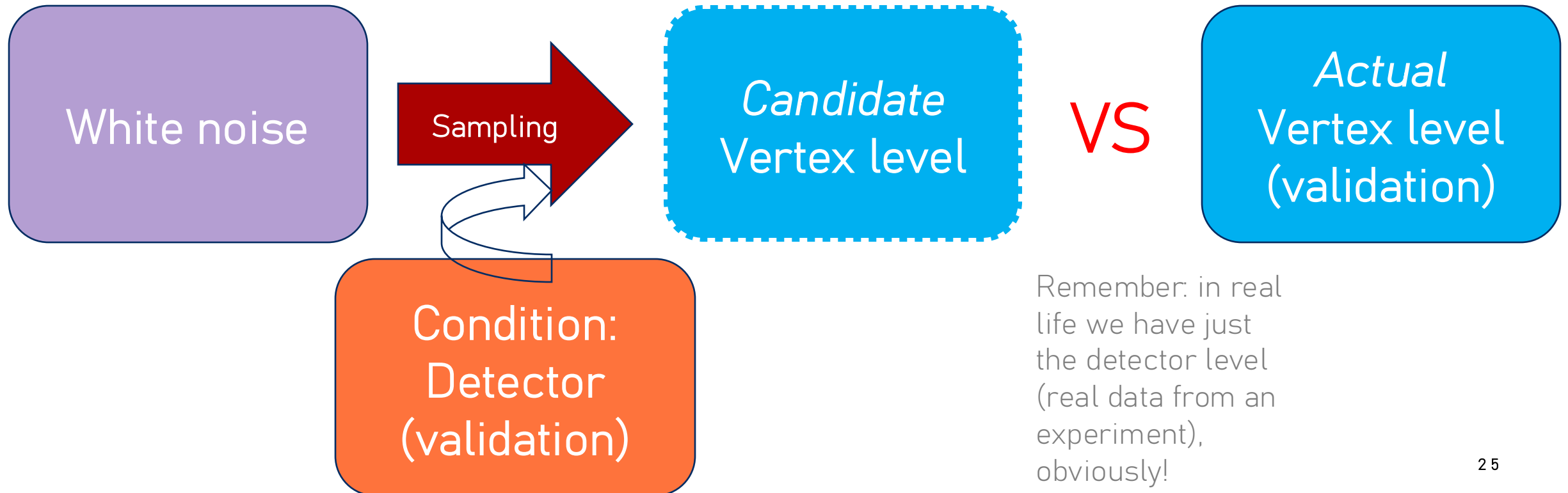
- The network predicts $\epsilon_{\theta}(\mathbf{x}_t, y, t)$
- Compute L and backpropagate to update θ
- Iterate until convergence

$$L(\theta) = \mathbb{E}_{\mathbf{x}_0, y, \epsilon, t} \left[\|\epsilon - \epsilon_{\theta}(\mathbf{x}_t, y, t)\|^2 \right]$$

Our strategy: results validation

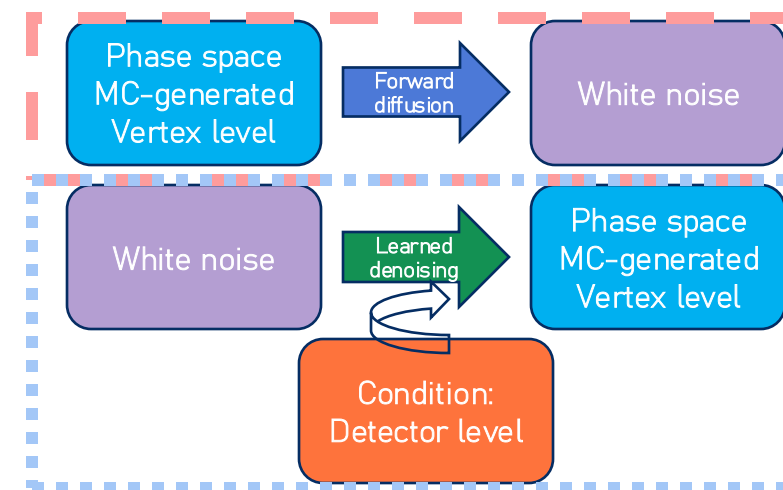
Once you have a trained model, you can use it to generate as much vertex samples as you want, giving it a validation detector set – independent from the training set – and comparing it with its actual vertex level, to complete the closure test:

- Compute pulls distributions (pulls = residues normalized to standard deviation)
- Perform statistical tests like Kolmogorov-Smirnov (unbinned! and analogue alternatives)



Initial explorative test

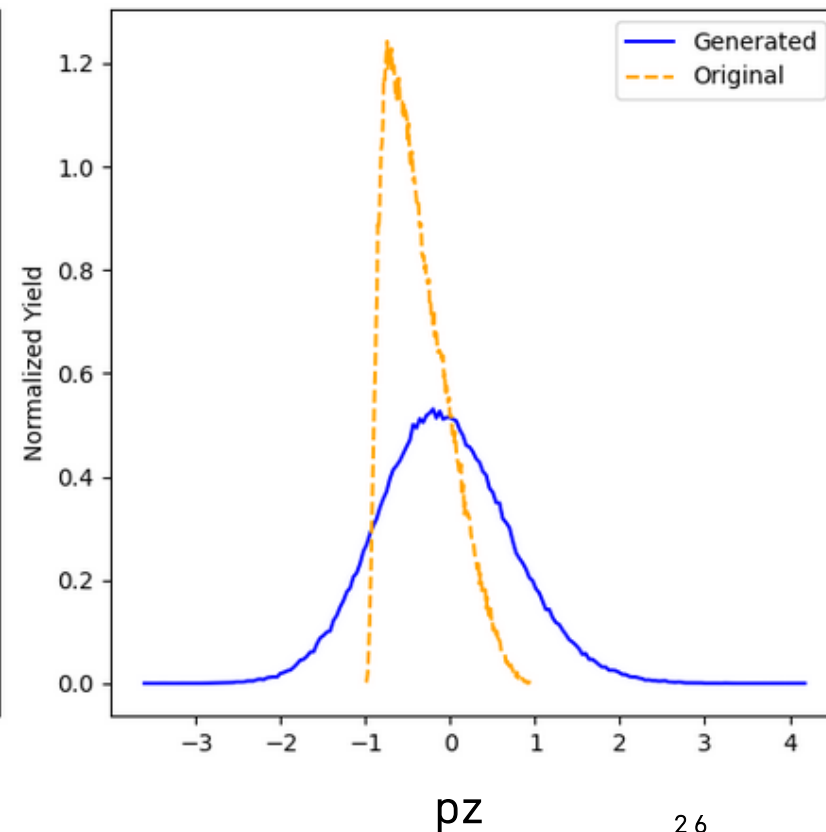
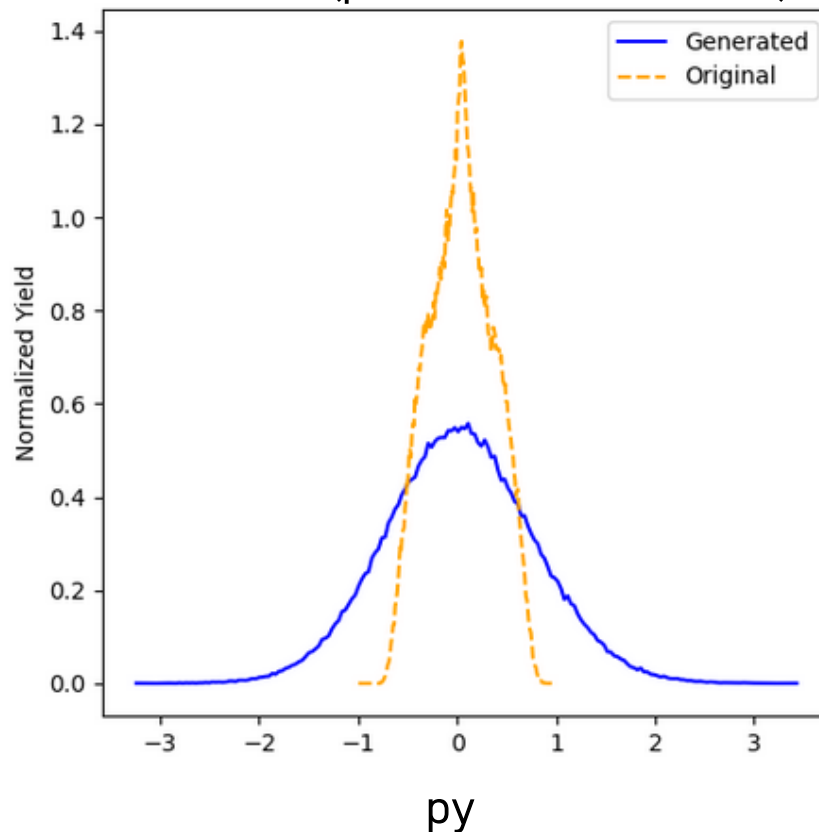
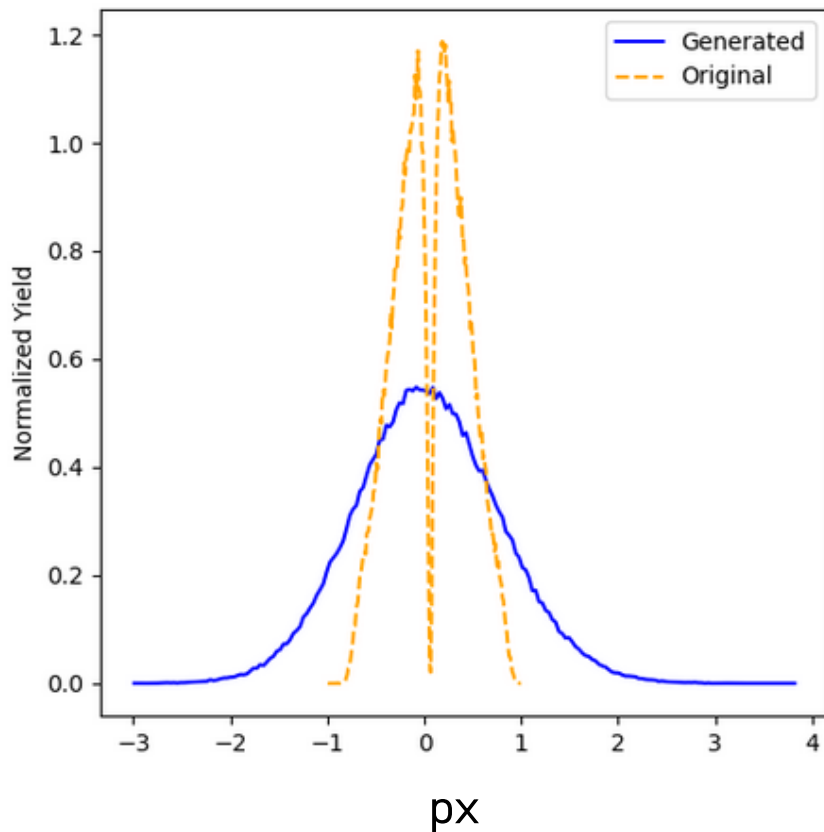
- We started training the diffusion model on simulations based on phase space only: generated in the acceptance + reconstructed (gsim detector proxy) (re-mapped into $[-1,1]$ interval)



Epoch 5/10000 [Loss: 0.209704]

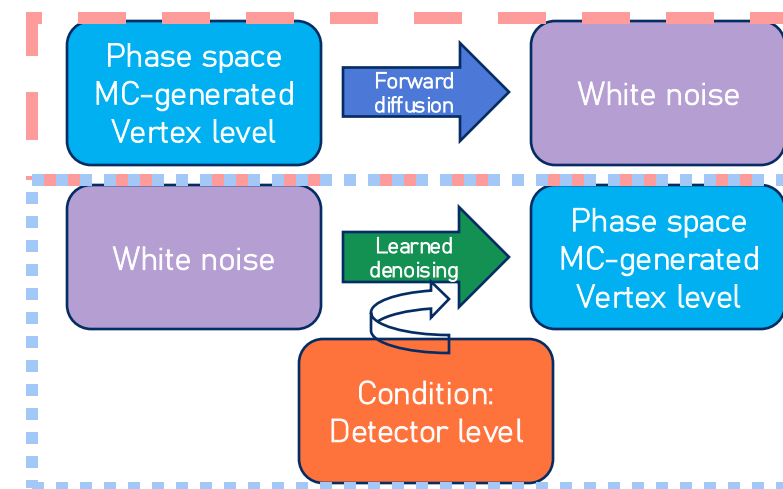
Sampling: 100% | 100/100 [00:34<00:00, 2.92it/s]

(proton 3-momentum)



Initial explorative test

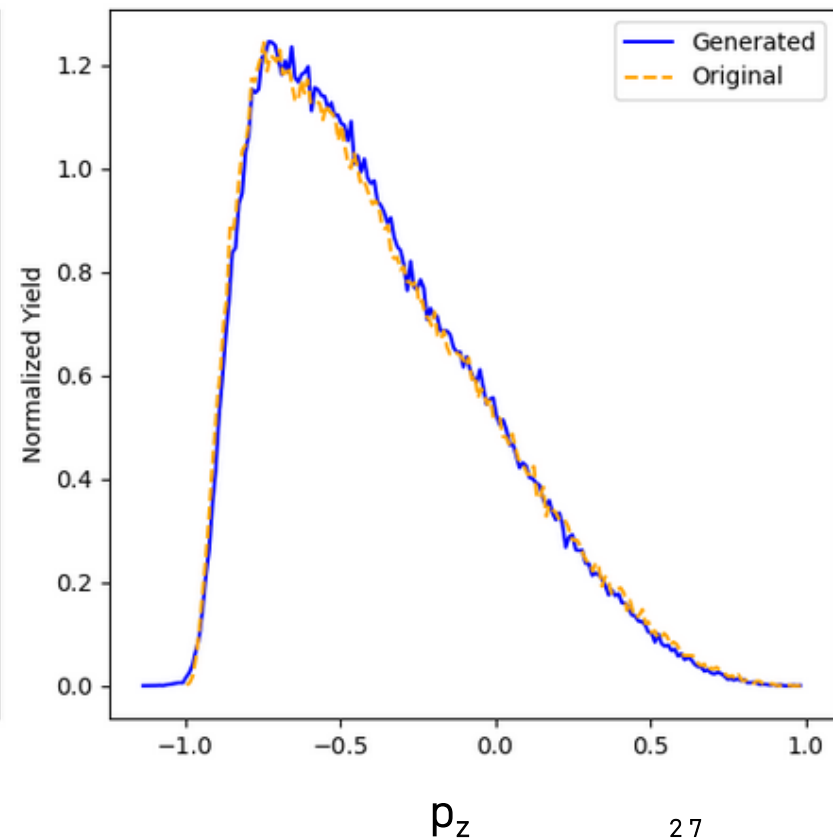
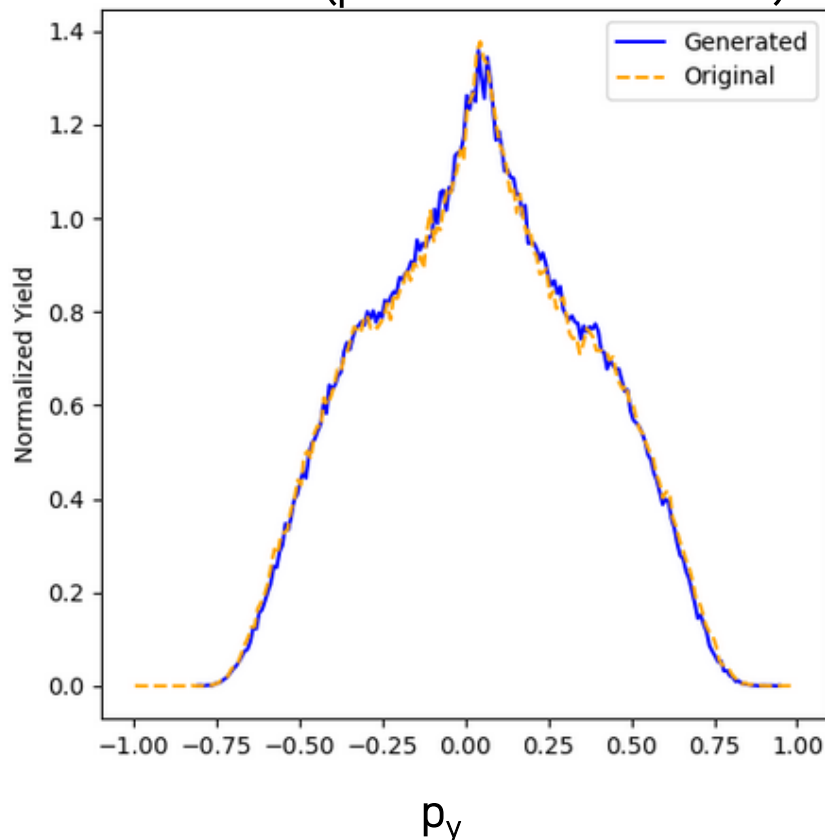
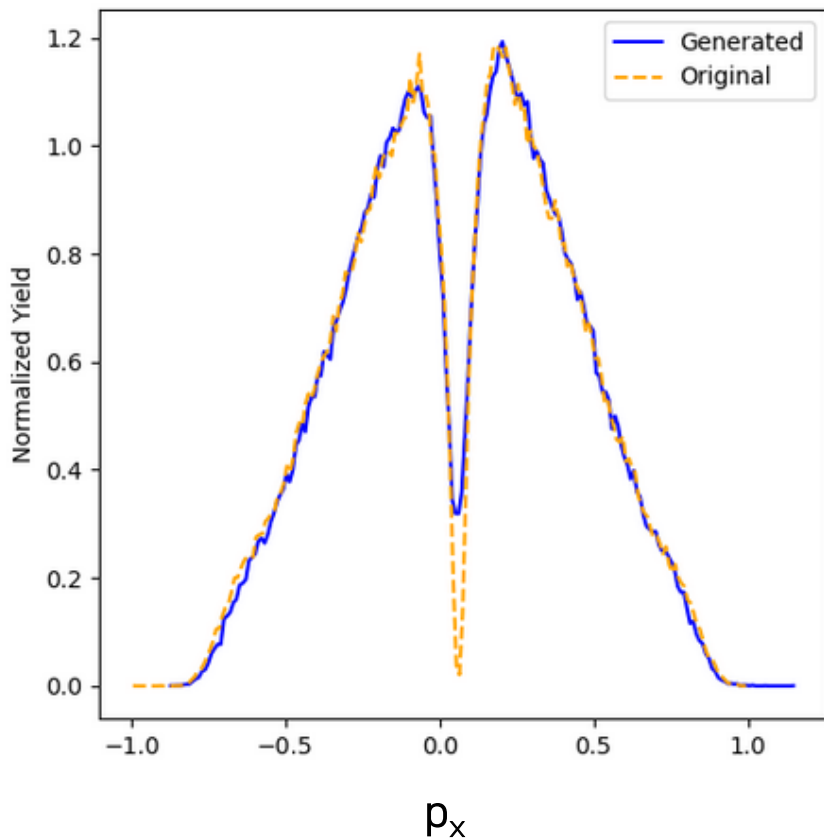
- We observed rapid convergence and huge flexibility: training other topologies we maintained the same architecture



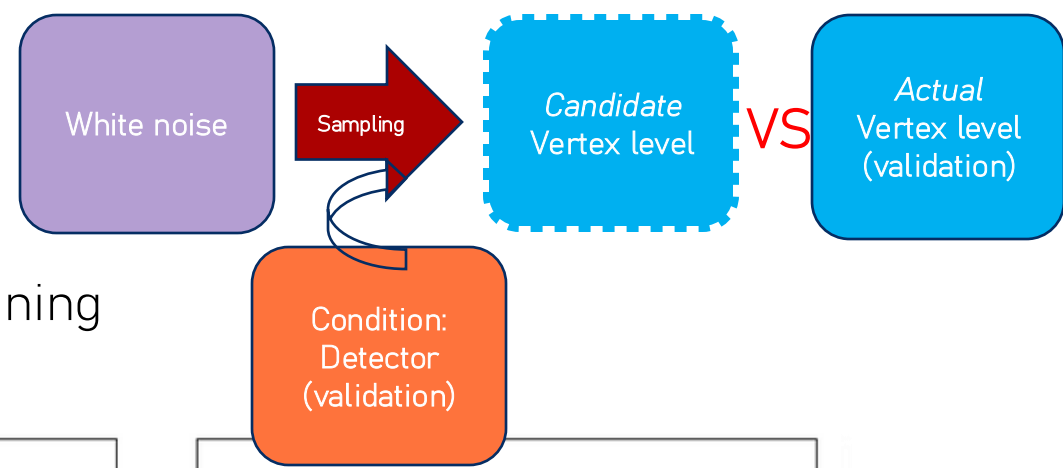
Epoch 95/10000 [Loss: 0.0181241]

Sampling: 100% | ██████████ | 100/100 [00:33<00:00, 2.95it/s]

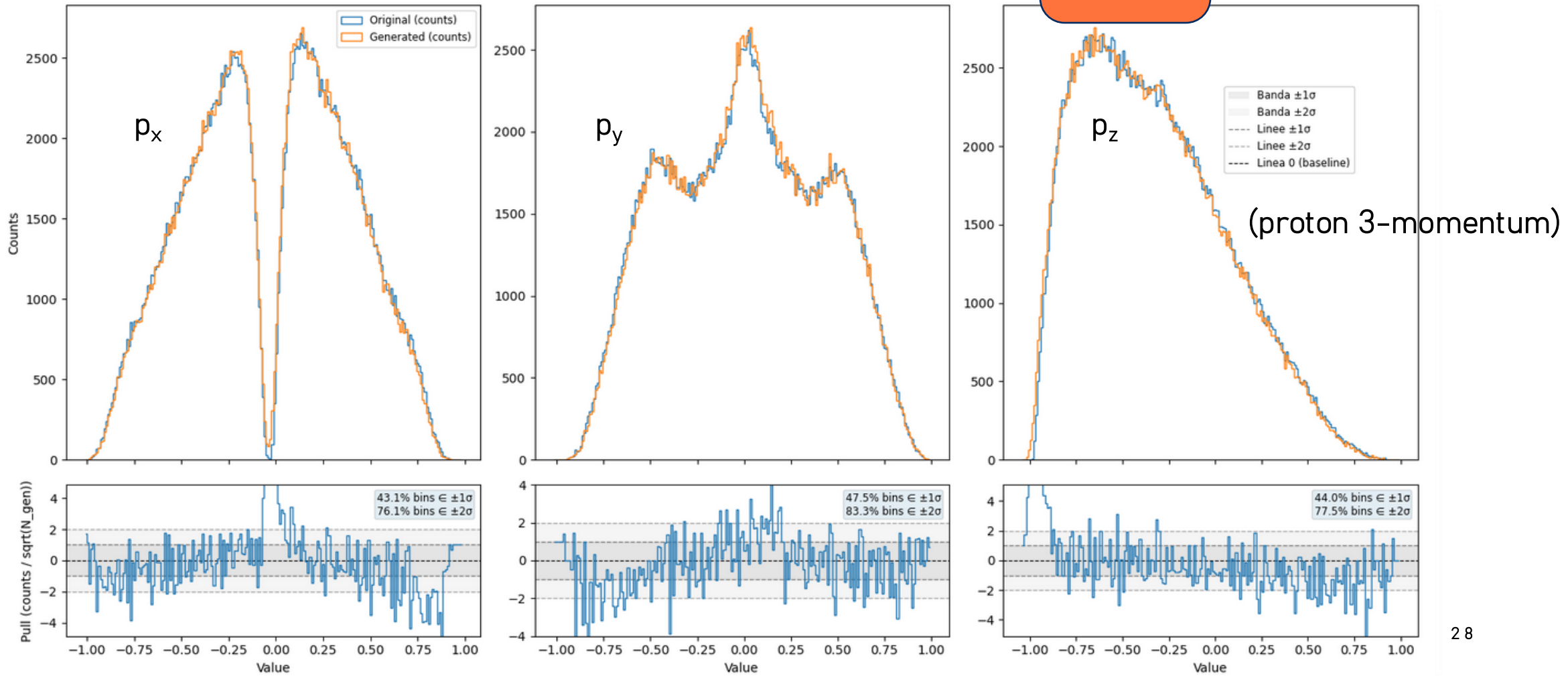
(proton 3-momentum)



Initial explorative test



- We observed rapid convergence and huge flexibility: training other topologies we maintained the same architecture



Current status of the project

- We want to have under control the mapping between the vertex and the detector in the **entire measured phase space**, so we are focusing on a complete set of Mandelstam invariants (they are 5, for a 2->3 process)

- Training and validation sets have the same size.

- Training observables:

$$s_{23} = (p_{\pi^-} + p_f)^2$$

$$u_2 = (p_{\pi^-} - p_i)^2$$

$$s_{12} = (p_{\pi^+} + p_{\pi^-})^2$$

$$s_{13} = (p_{\pi^+} + p_f)^2$$

$$u = (p_{\pi^+} - p_i)^2.$$

$$\gamma p_1 \rightarrow \pi^+ \pi^- p_2$$

- To eliminate outliers, we cut in s in $[5.9, 8.4] \text{ GeV}^2$. (s_{\max} is fixed by the energy of the beam $\sim 3.7 \text{ GeV}$).
- Every variable is transformed according to the map $x \rightarrow \text{sign}(x) \cdot (1.0 - (1.00000001 - |x|)^\alpha)$, $\alpha = 0.4$

to reduce the slope at the borders of the domain.

Current status of the project

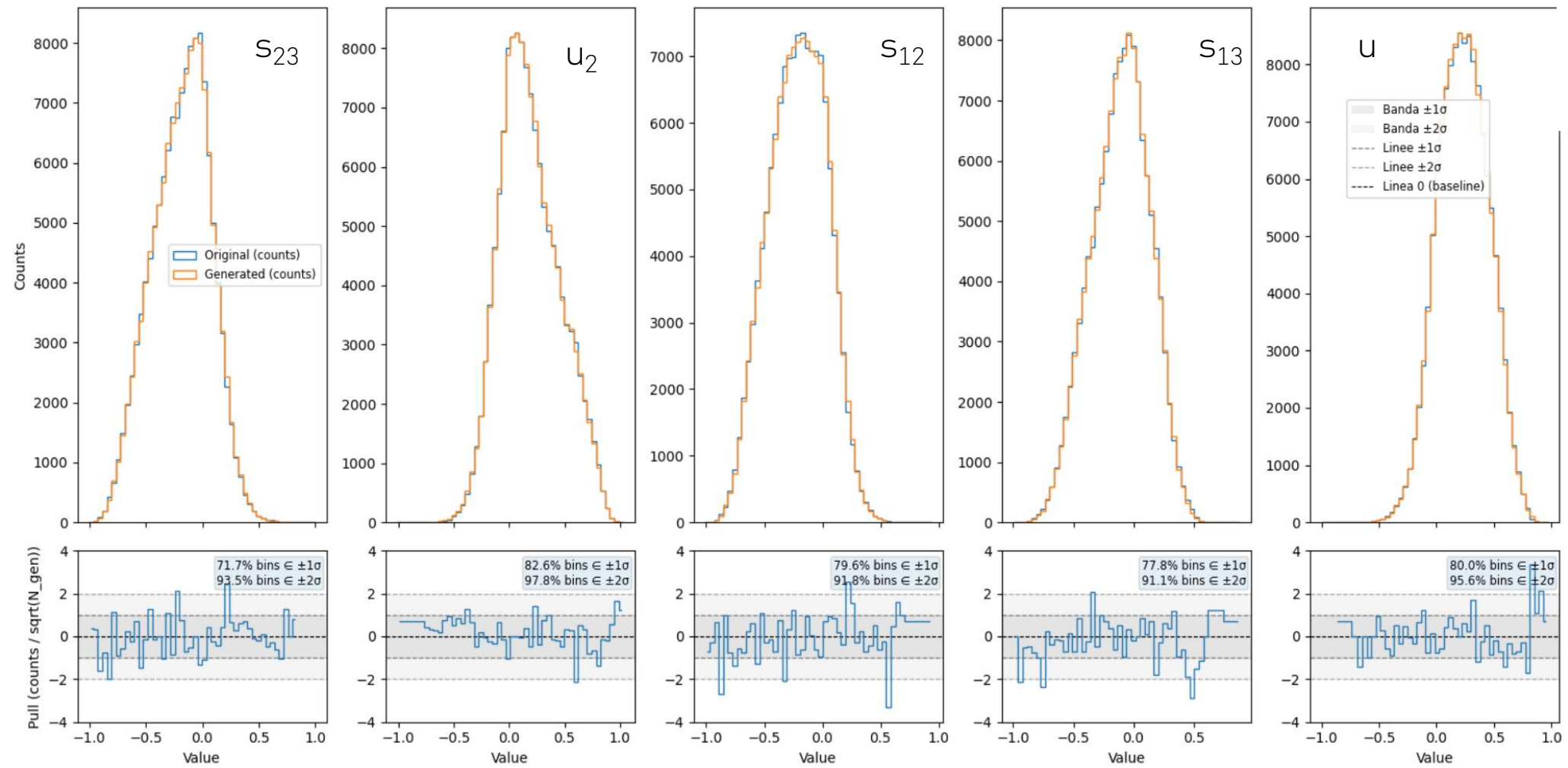
$$s_{23} = (p_{\pi^-} + p_f)^2$$

$$u_2 = (p_{\pi^-} - p_i)^2$$

$$s_{12} = (p_{\pi^+} + p_{\pi^-})^2$$

$$s_{13} = (p_{\pi^+} + p_f)^2$$

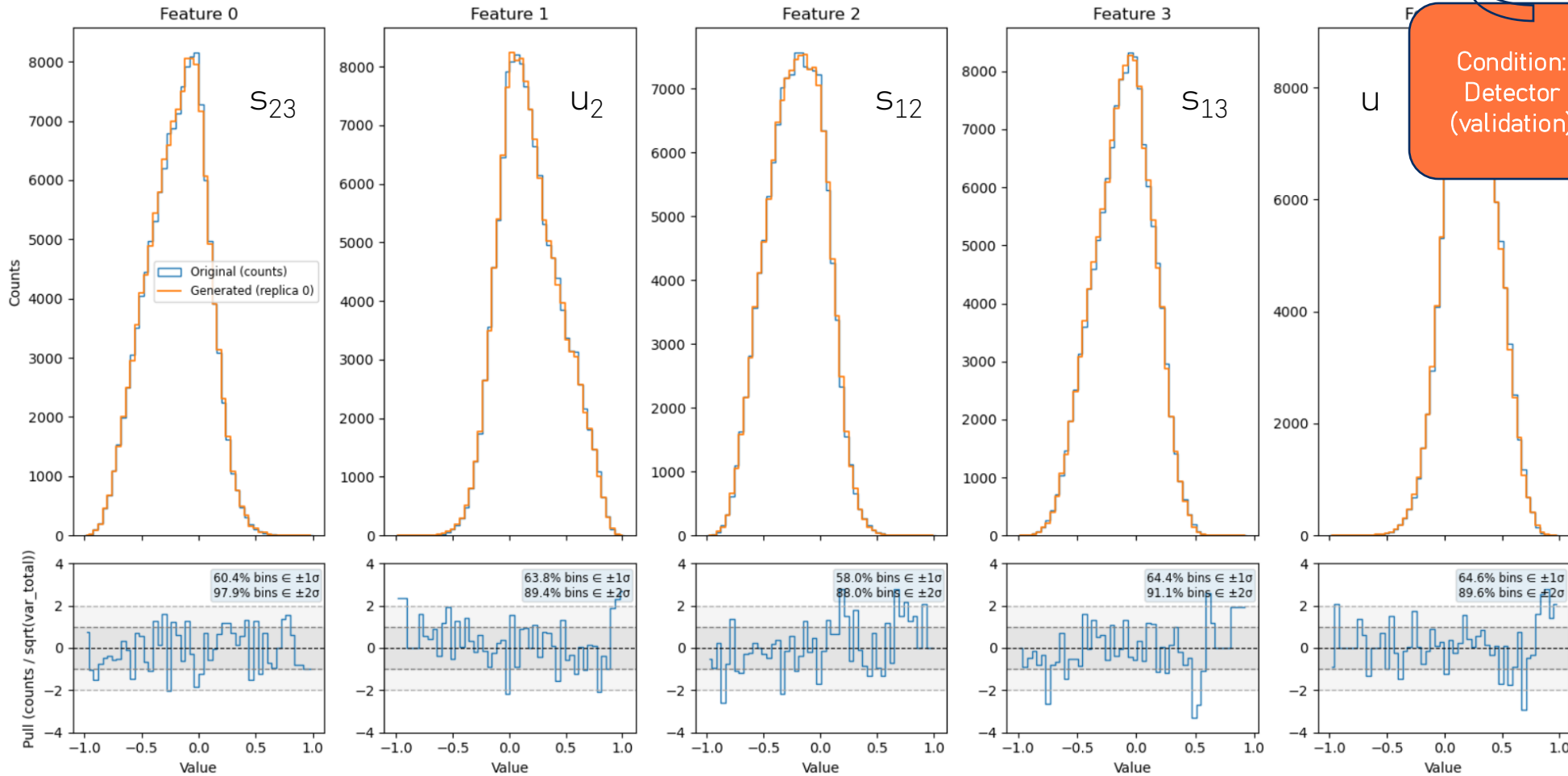
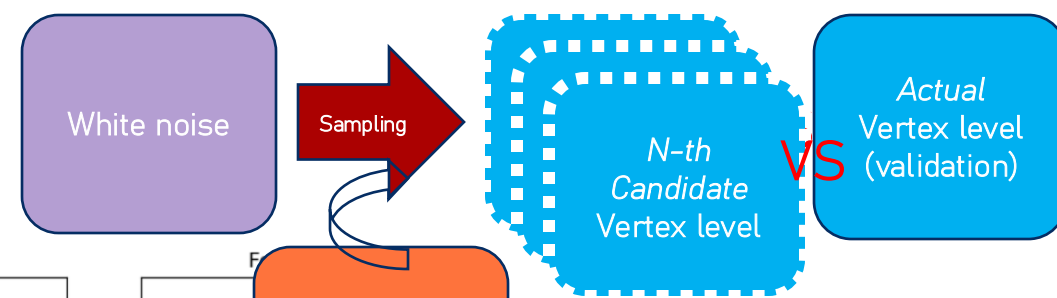
$$u = (p_{\pi^+} - p_i)^2$$



KS-test: p-value = { 0.70, 0.82, 0.23, 0.92, 0.26 } : all greater than the threshold for rejecting the null hypothesis, 0.05!

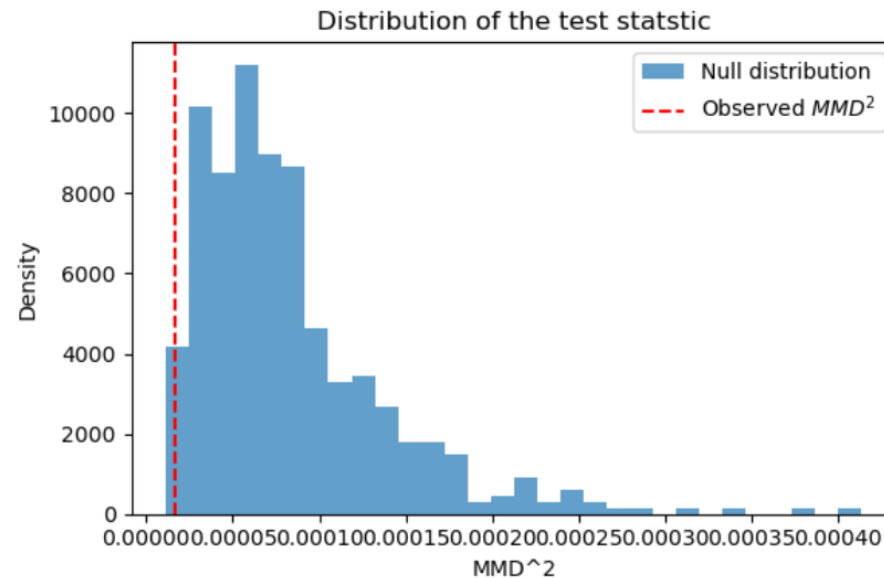
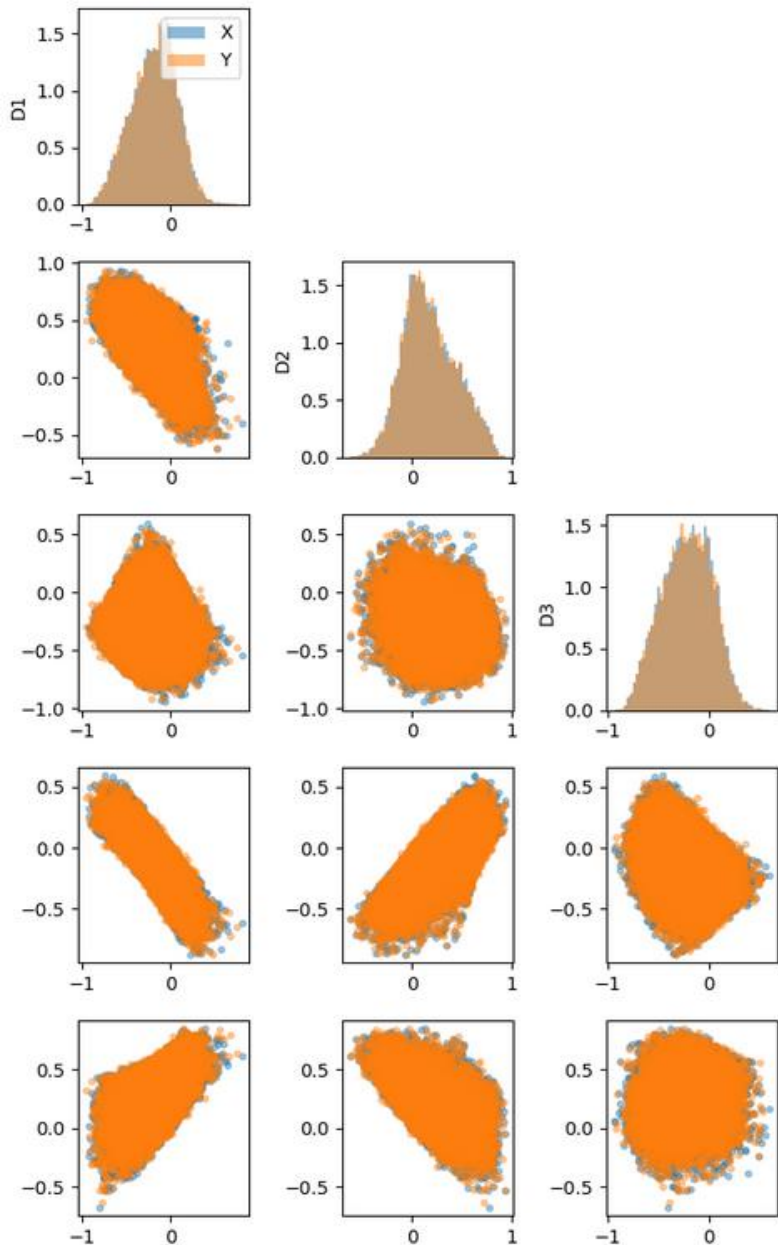
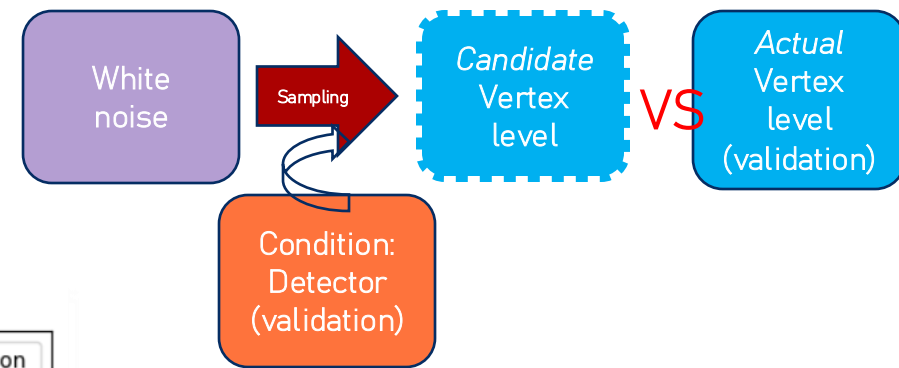
More than 90% of sampled data **within 2 sigma**, and we haven't included the systematics errors (for sure dominant over the statistical one) yet!

Current status of the project



- This is a set of vertex replicas generated from the same detector level
- The statistical error associated to the sampling procedure is \ll statistical error
- The KS p-value is stably > 0.05 within the whole bunch of replicas

Current status of the project

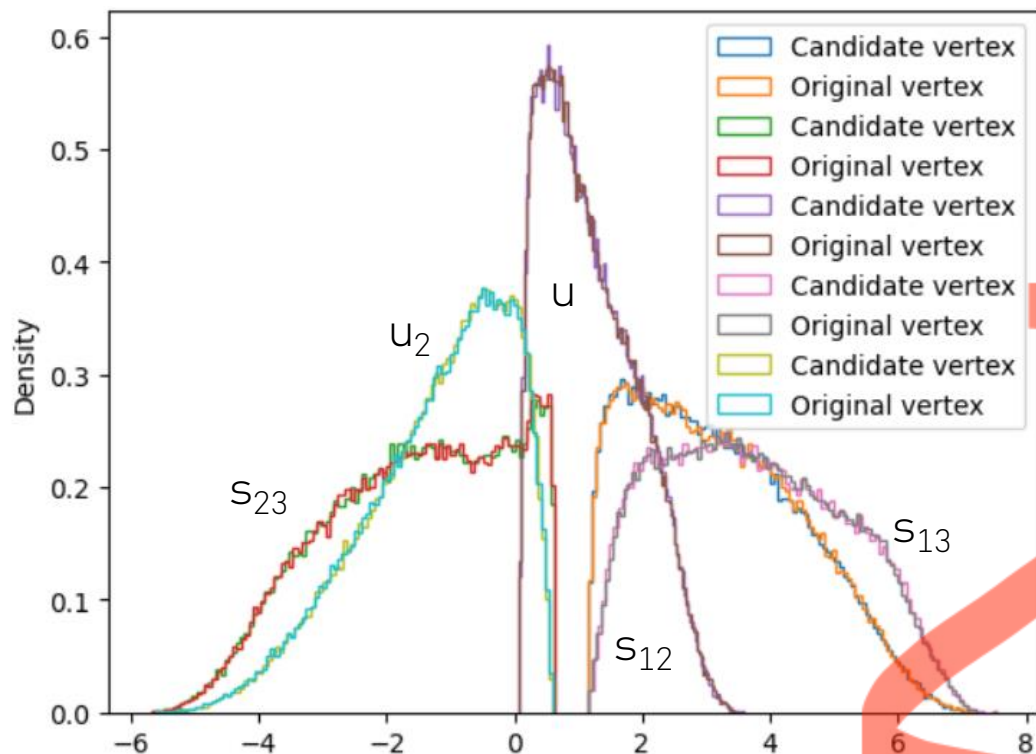
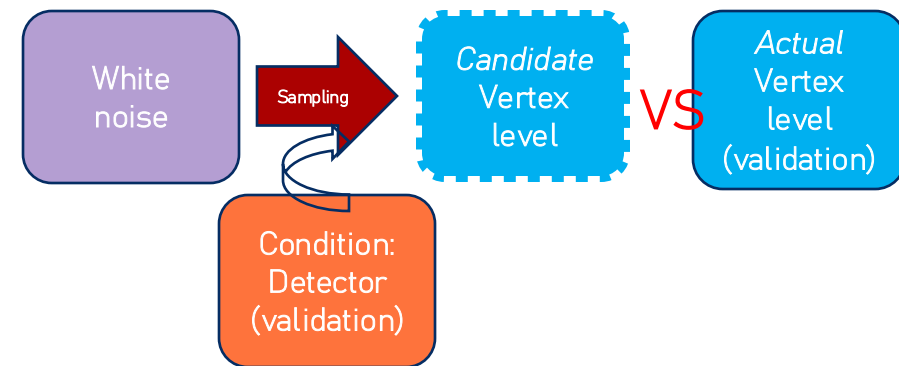


We can do better than simple KS-test to evaluate how good our model is:

- KS is a univariate test, so doesn't allow us to take care of the correlations between sampled observables
- But Nyström MMD Permutation Test developed by Marco Letizia (MaLGa centre) does!

Number of nystrom centers used: 282
 Output of the test: 0
P-value: 0.99
 Observed value of $rMMD^2$: 1.6421587121545935e-05

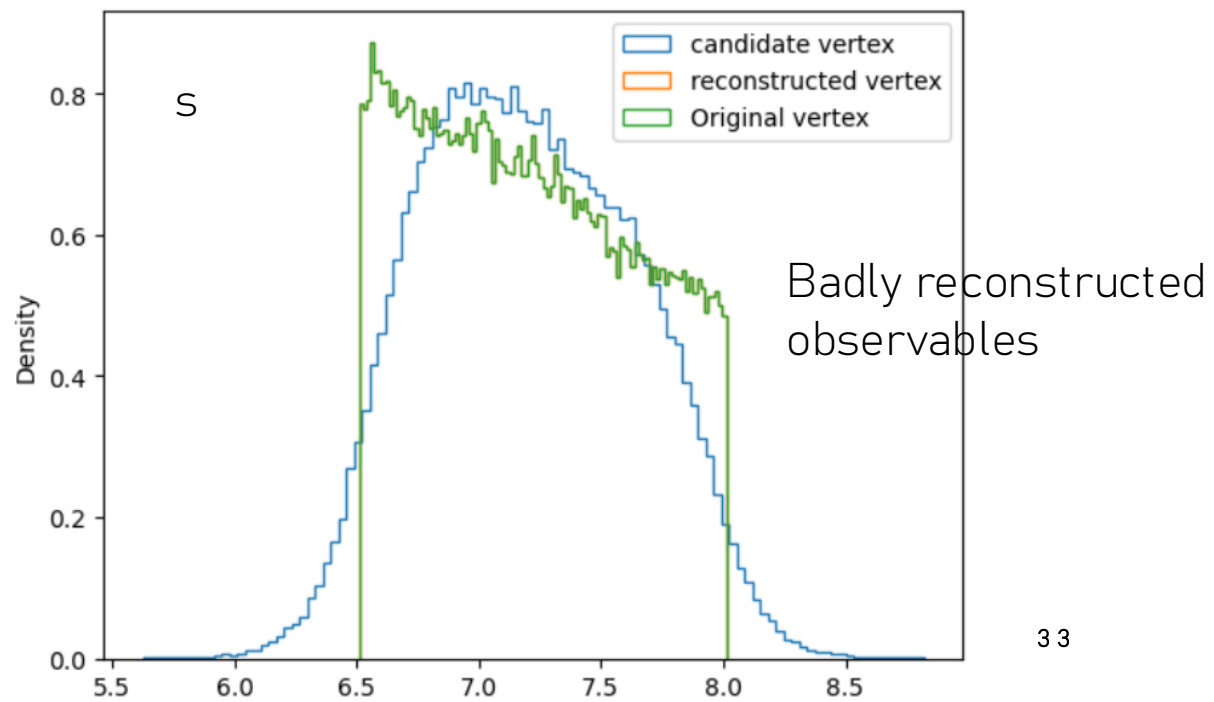
Current status of the project



De-normalized invariants well sampled, with correlations respected

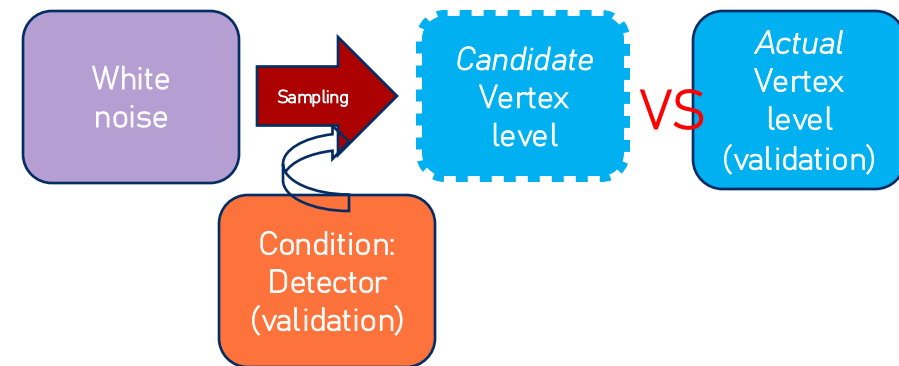


Currently trying to understand why we are not good in reconstructing through linear combinations and transformations the other invariants and the other observables in the lab frame, although the tests is telling us that our model caught correctly the correlations



$$s = s_{12} + s_{13} + s_{23} - \sum_{k=1,2,3} m_k^2$$

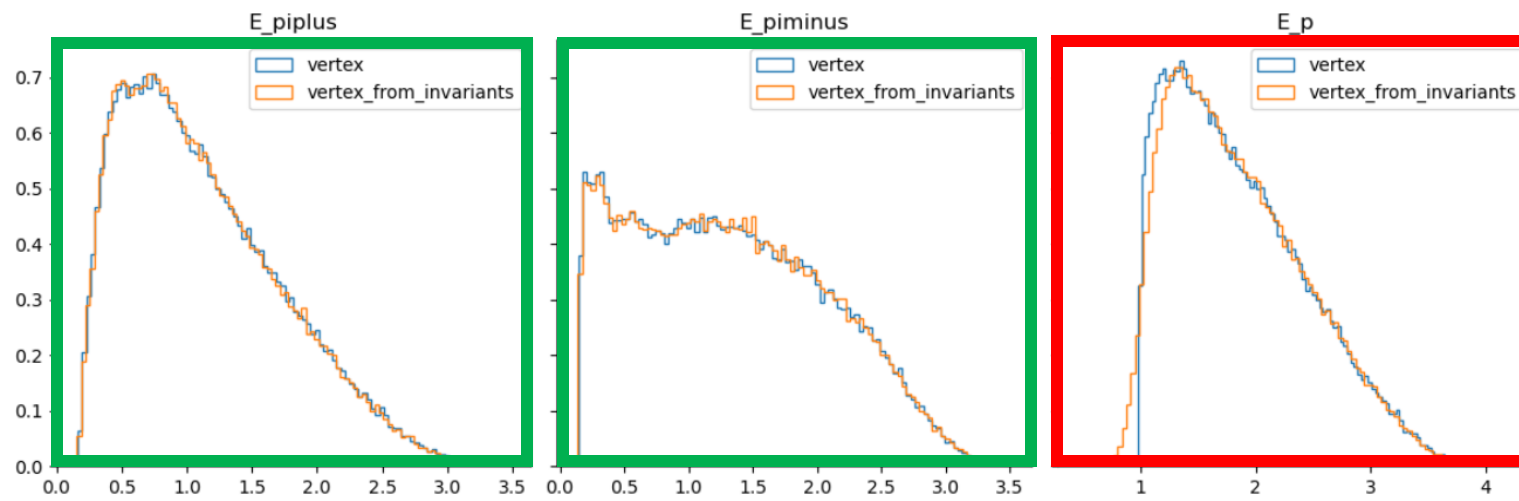
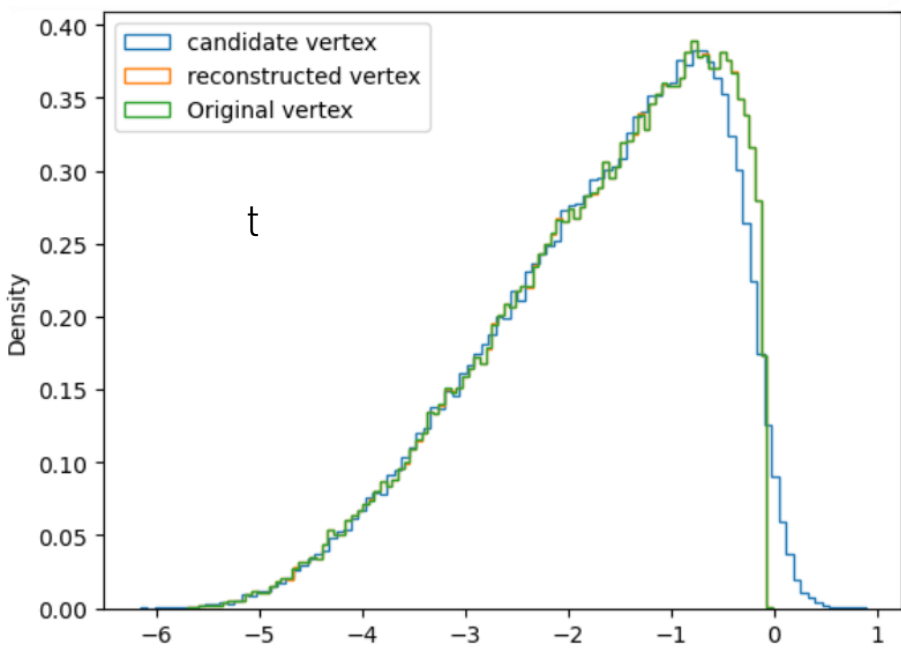
Current status of the project



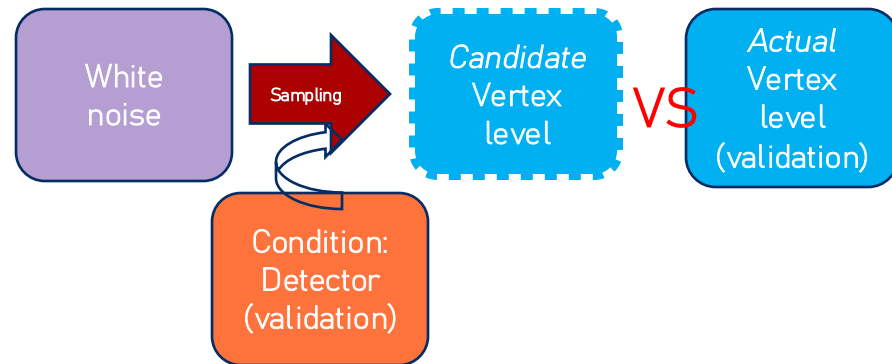
And so on for t and the observables in the lab frame (energies and 3-momenta components):



$$\sum_{k=1,2,3} t_k = 2m_a^2 + m_b^2 + \sum_{k=1,2,3} m_k^2 - s$$



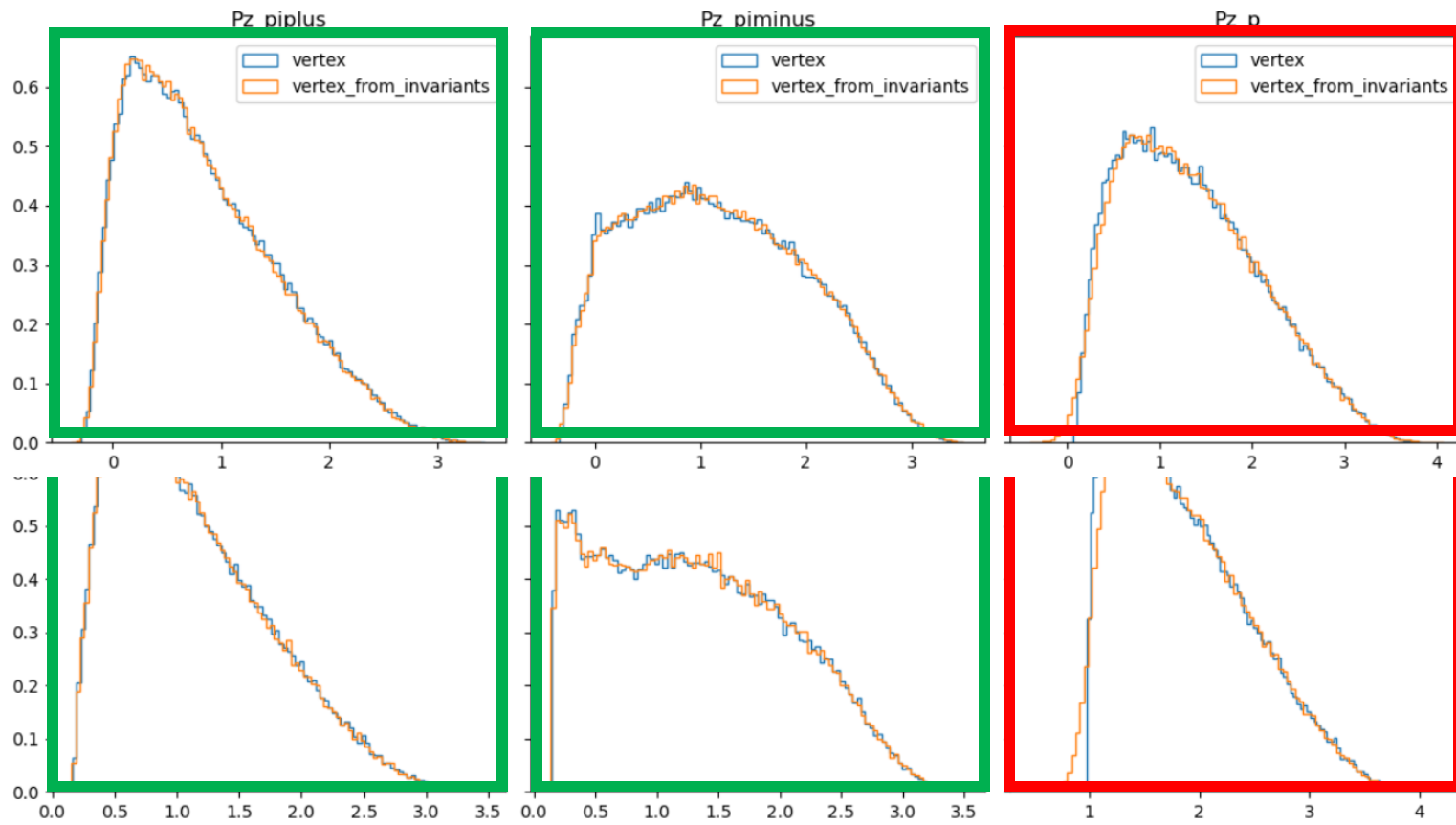
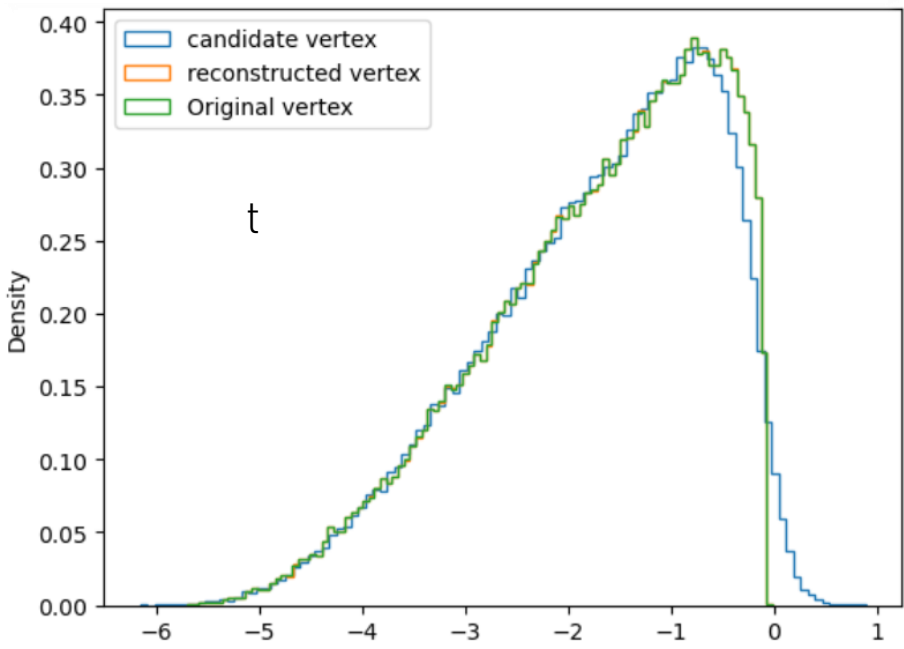
Current status of the project



And so on for t and the observables in the lab frame (energies and 3-momenta components):



$$\sum_{k=1,2,3} t_k = 2m_a^2 + m_b^2 + \sum_{k=1,2,3} m_k^2 - s$$

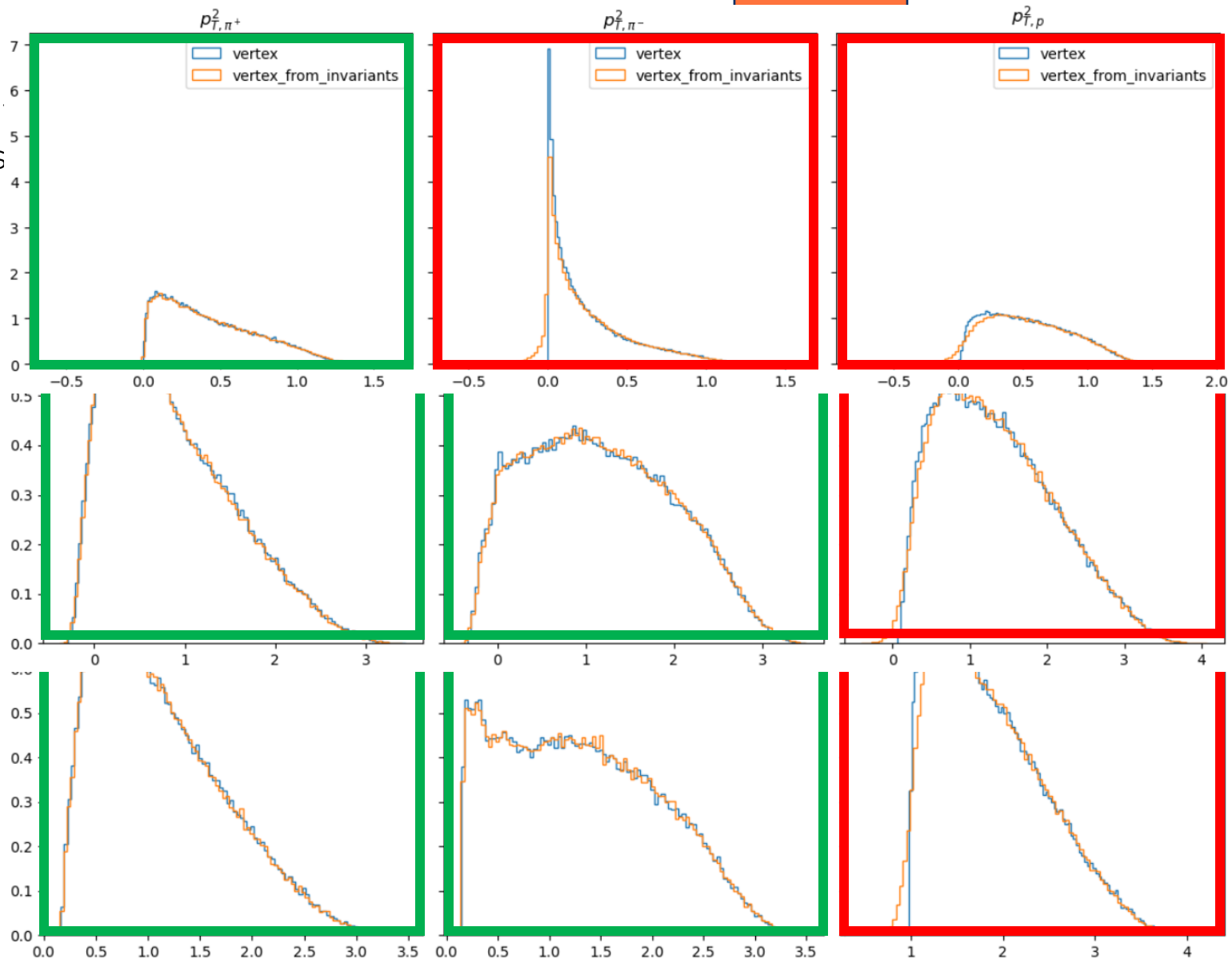
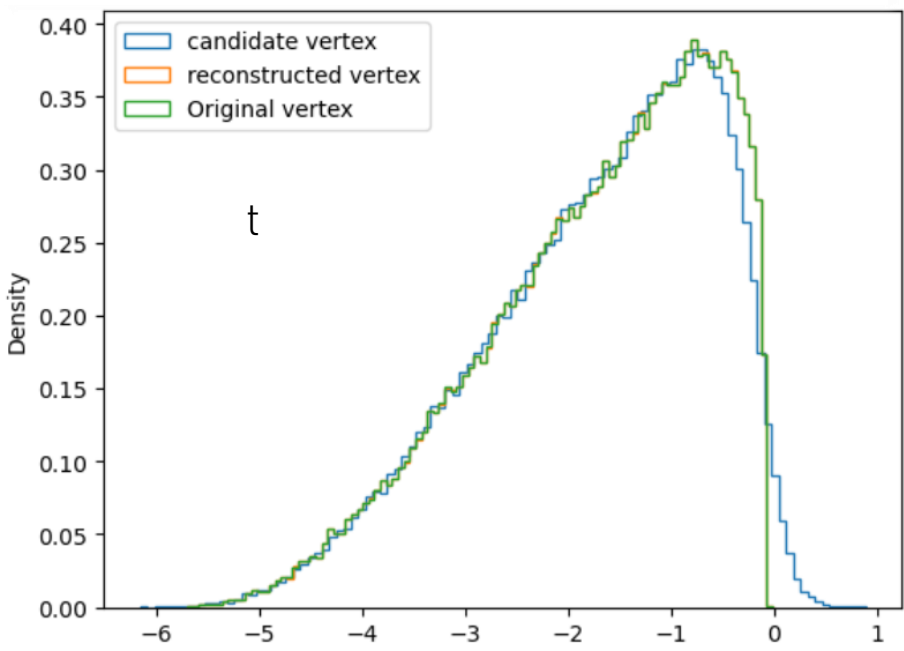


Current status of the project



And so on for t and the observables in (energies and 3-momenta components)

$$\sum_{k=1,2,3} t_k = 2m_a^2 + m_b^2 + \sum_{k=1,2,3} m_k^2 - s$$



Summary

- Unfolding the detector effect through generative AI
- How a diffusion model work
- Status of closure test for CLAS detector
- Conclusions & Next directions



Conclusions & Next directions

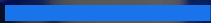
We trained a DM to map vertex and detector level related to the entire phase space simulated for the reaction of photoproduction of a charged pion pair on proton, smeared by the GSIM proxy of CLAS detector

- We can substitute the real vertex distribution with the analogue sampled by our DM within 2 sigma, including just statistical uncertainties (road to a digital twin) and systematic related to sampling variance (completely negligible), according both to **simple pulls** and to the sophisticated **Nyström MMD Permutation Test**, developed by Marco Letizia et al.
- Understand how to **improve the reconstruction of the other invariants and lab observables**
- Refine the estimation of the **statistical error bars** and the treatment of **challenging events**: special attention must be given to particles outside detector acceptance or rare phase-space regions (e.g. the regions at the borders of two different topologies). The model should ideally be able to properly treat also the outliers.
- Re-do the same procedure on the **other topologies**, to construct a wrapper of models that can work in the entire measured phase-space
- **Closure test with JPAC model**: now we are using data generated from a simple model containing just phase-space (no dynamics included), and we will do the closure test not only on s , t etc. as we are trying to do now, but also on other significant invariants, like the **moments of the angular distribution**
- **Application of our DM to real CLAS data & extract moments from them**
- **Collaboration with GlueX** people to enlarge it way more including the phase space measured by them: we will be able to compare moments extracted at the vertex level by us and them in the region overlapped by the two detectors. This – as far as we know – never tried approach can **reduce a lot the model dependence** that we introduce when we extrapolate the cross section in the unmeasured phase space region



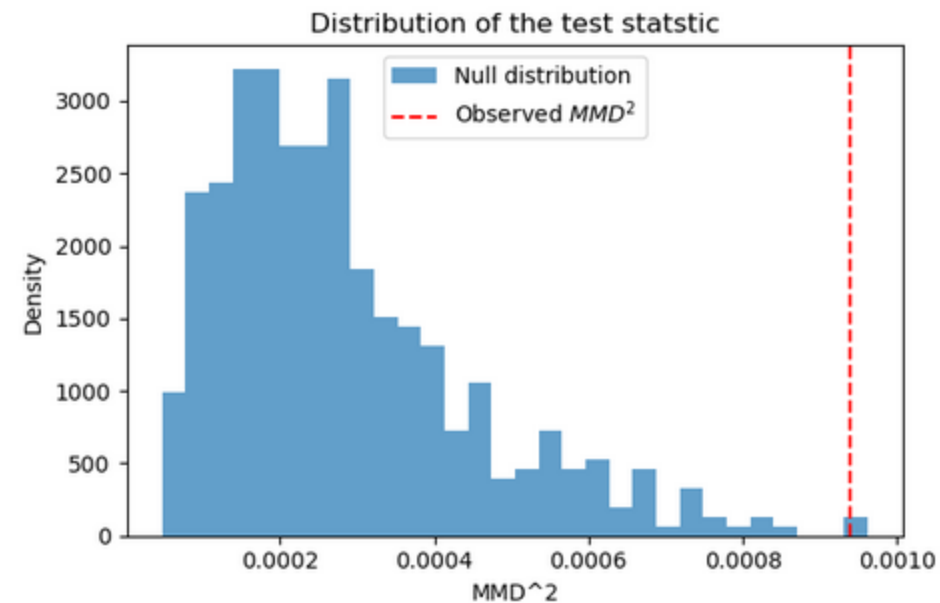
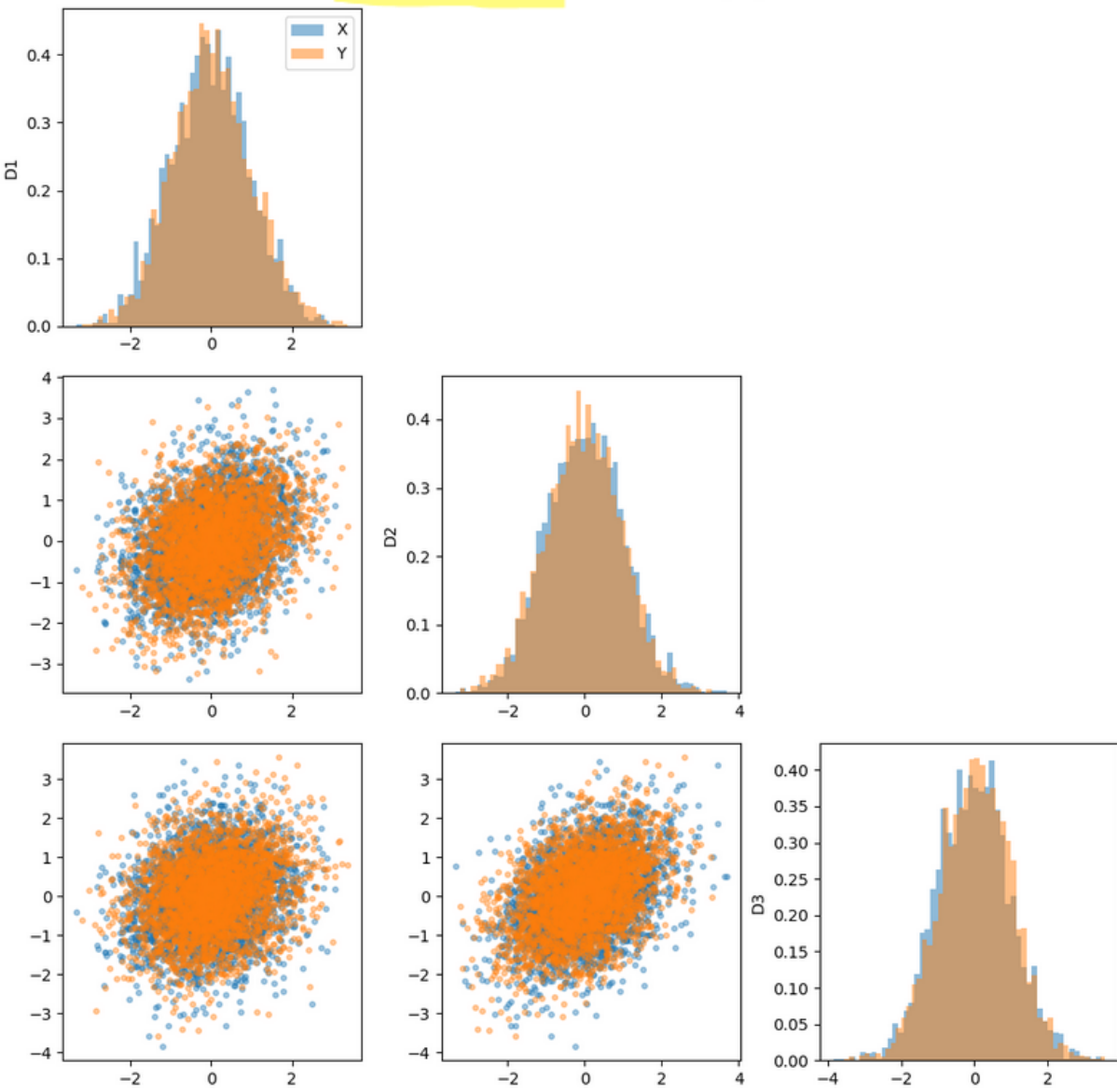
Thank you for your
attention!

Backup



More on the Nyström MMD Permutation Test

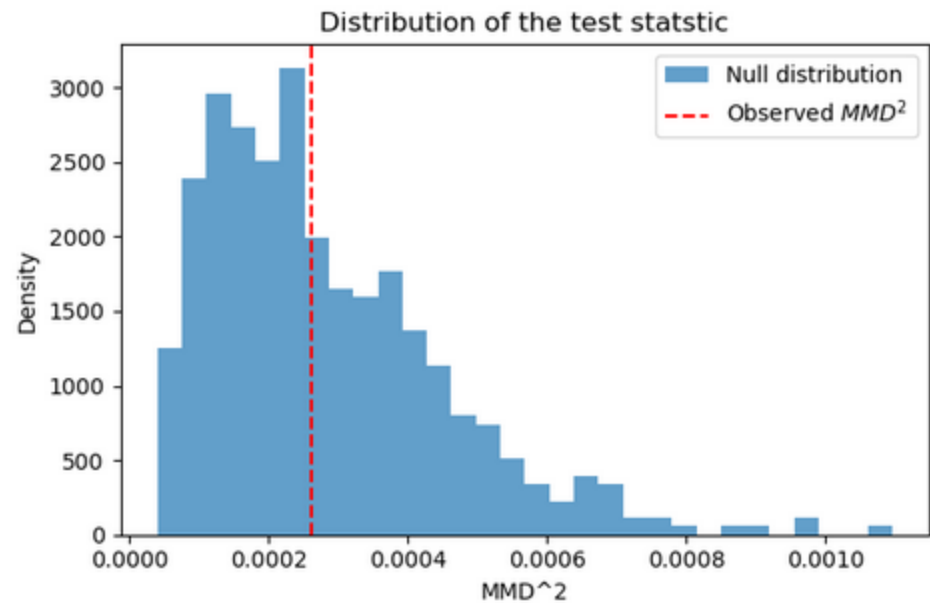
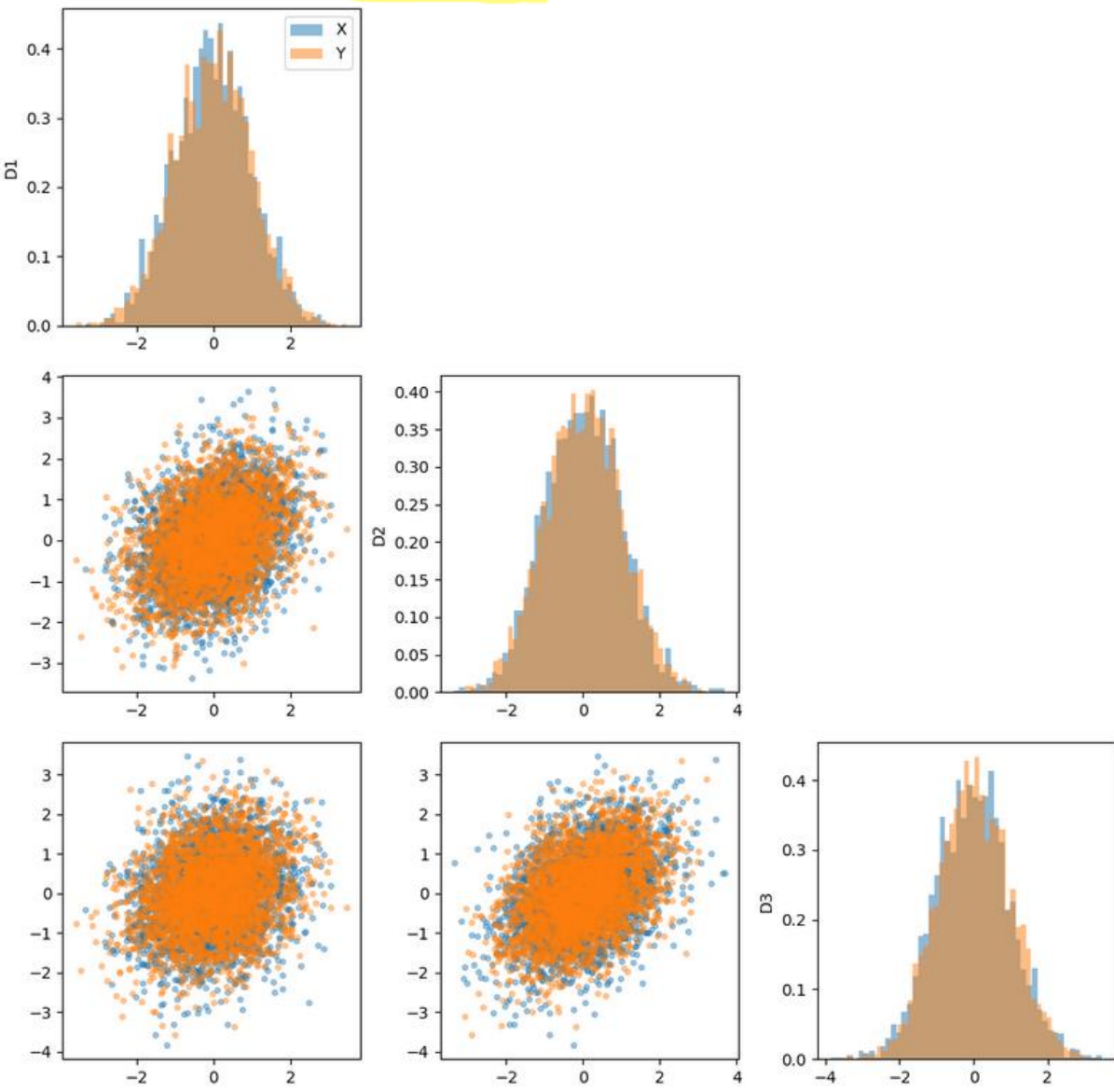
Different distribution - Scatter Matrix (3D)



Number of nystrom centers used: 154
Output of the test: 1
P-value: 0.004
Observed value of rMMD²: 0.0009392642708340842

More on the Nyström MMD Permutation Test

Same distribution - Scatter Matrix (3D)



Number of nystrom centers used: 154
Output of the test: 0
P-value: 0.454
Observed value of $rMMD^2$: 0.00026143954739286007

Toy case

- 2 variables (x,y), uniformly distributed in $[-1,1]$; we want to reproduce a 3rd variable (z) is a simple linear combination of them ($z = -(x+y)$)
- Difficulties in reproducing the sharpest features, without a smoothing transformation (more expressive architectures where tested)

Comparison visualization saved for epoch 240
Model checkpoint saved (epoch 240)
Epoch 245/10000 [Loss: 0.062275]

Sampling: 100% | ██████████ | 500/500 [00:15<00:00, 33.30it/s]

