A method for inferring signal strength modifiers by conditional invertible neural networks

Máté Zoltán Farkas, Svenja Diekmann, Niclas Eich and Martin Erdmann

RWTH Aachen University, III. Physikalisches Institut A, Otto-Blumenthal-Str., 52074 Aachen, Germany

E-mail: mate.zoltan.farkas@cern.ch, svenja.diekmann@cern.ch, niclas.steve.eich@cern.ch, martin.erdmann@cern.ch

Abstract. The continuous growth in model complexity in high-energy physics (HEP) collider experiments demands increasingly time-consuming model fits. We show first results on the application of conditional invertible networks (cINNs) to this challenge. Specifically, we construct and train a cINN to learn the mapping from signal strength modifiers to observables and its inverse. The resulting network infers the posterior distribution of the signal strength modifiers rapidly and for low computational cost. We present performance indicators of such a setup including the treatment of systematic uncertainties and highlight the features of cINNs estimating a signal strength for HEP-data on simulations.

1. Introduction

Performing many-parameter maximum likelihood fits in high-energy physics (HEP) can be time-consuming. Parameter inference with conditional invertible neural networks (cINNs) is ultrafast and thanks to the underlying normalizing flow, the trained network model is continuous and continuously differentiable in both directions providing additional generative properties. These versatile networks have already been successfully applied to a broad range of challenges outside of HEP ranging from scientific model inversion to image colorization and guided image generation [1, 2]. In high-energy physics, cINNs have been applied to stellar evolution parameter estimation, cosmic-ray source property determination and detector effect simulations [3, 4, 5]. In this paper, the applicability of cINNs to infer signal strength modifier parameters (defined as the ratio of cross sections $\mu = \sigma/\sigma_{SM}$) is demonstrated and the performance measures of the resulting model are described.

2. Vector Boson associated Higgs Production: Analysis Strategy and Setup

The network has been trained to infer the signal strength of the gluon-fusion ($gg \to ZH$) and quark-initiated VH-final state ($qq \to ZH$ and $qq \to WH$) processes using simulated Monte Carlo (MC) samples at a collider experiment. In the analysis workflow one-lepton and two-lepton final states with $H \to b\bar{b}$ are considered and the events are assigned to categories according to the number of reconstructed leptons and b-jets. These categorized events are then further classified with a feed-forward deep neural network (DNN) into 13 categories for the 3 relevant signal and 10 other background processes. The assigned DNN scores for each event are then histogrammed

and used as fit observable. In cINN terminology, these histograms represent the conditions c and used as an input for the network.

3. Neural Network Setup

3.1. Introduction to cINNs

Conditional invertible neural networks have an input node, an output node and a condition node. The network maps the inputs to the output space through alternating layers of affine coupling blocks (ACBs) and permutation layers. The ACBs encode the normalizing flow, such that the resulting network model is continuous and continuously differentiable in both directions. The condition node is connected to each of the ACBs, serving as an additional input. The output space is called latent space.

The aim of the training is to approximate the unknown true posteriors $p(x|c)$ of the physics parameters x under the observation of c with the network parameters ϕ, i.e. to obtain

$$p_\phi(x|c) \approx p(x|c). \tag{1}$$

This can be done by minimizing the Kullback-Leibler divergence between the approximated posterior and the true posterior

$$\mathbb{KL}\left[p \mid\mid p_\phi\right] = \mathbb{E}_{x\sim p(x|c)}\left[\log\left(\frac{p(x|c)}{p_\phi(x|c)}\right)\right] = -\mathbb{E}_{x\sim p(x|c)}\left[\log p_\phi(x|c)\right] + \text{const.} \tag{2}$$

The Kullback-Leibler divergence in eq. 2 can be rewritten using the network output variables $f(x) = z$. Setting the latent space variable distribution to a normal distribution, the result of eq. 2 becomes

$$L = \mathbb{E}_{x\sim p(x|c)}\left[\frac{z^2}{2} - \log\left|\det\frac{\partial z}{\partial x}\right|\right] \tag{3}$$

and can be used as a loss function for the training, if the ACB's determinants are feasible to evaluate. The GLOW coupling block [6] provides such a determinant since it consists of two consecutive transformations with a triangular Jacobian matrix each. Its determinant is thus just the product of the diagonal entries of the matrix. A sketch of the GLOW coupling block is shown in fig. 1 for both the forward and the backward pass with input- and output variables u and v, respectively. The intermediate transformations s_i and t_i do not have to be invertible themselves and can be parameterised with feed-forward networks.

3.2. Synthetic Dataset Generation

Reliable parameter inference requires correct modelling of the expectations in the training dataset. The effects of the physics parameters, systematic and statistical fluctuations are all modelled in the conditions. For shape-changing systematic effects where only templates of the 1σ up- and down variations are accessible a histogram template morphing technique [7] has been applied to produce samples with different nuisance parameters for each uncertainty. The resulting histogram bin contents for each process are varied to represent the limited MC sample size.

For each histogram sample the signal processes have been scaled with random numbers drawn from a Γ-distributed prior. This particular prior enables a more refined sampling in the region $\mu \lesssim 10$ where the signal strength modifiers are mostly expected in the data. The long decay of the Γ-distribution also enables to maintain some sensitivity for a higher-than-expected signal scenario. The nuisance parameters of each background process normalizing uncertainty have been drawn from a lognormal prior with mean 1; for the luminosity nuisance parameter a

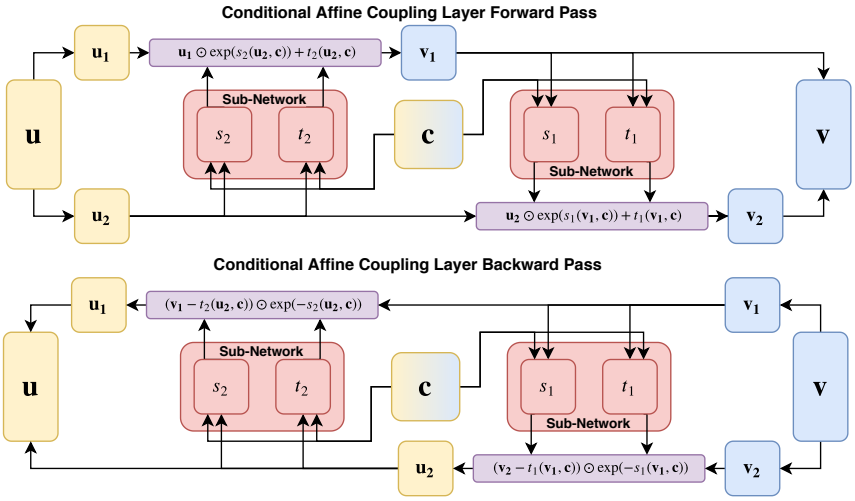

Figure 1. The GLOW coupling block [3]. The two consecutive transformations are written in the purple boxes.

narrower lognormal prior has been chosen and affects each process equally. Since the expected data is also Poisson-distributed, the bin content of each bin has been varied accordingly as well.

With the 16 physics signal and background modifier parameters and the luminosity nuisance parameter the total dimension of network input parameters is then 17. A sketch of the network setup with the inputs and nuisance parameters is shown in fig. 2.

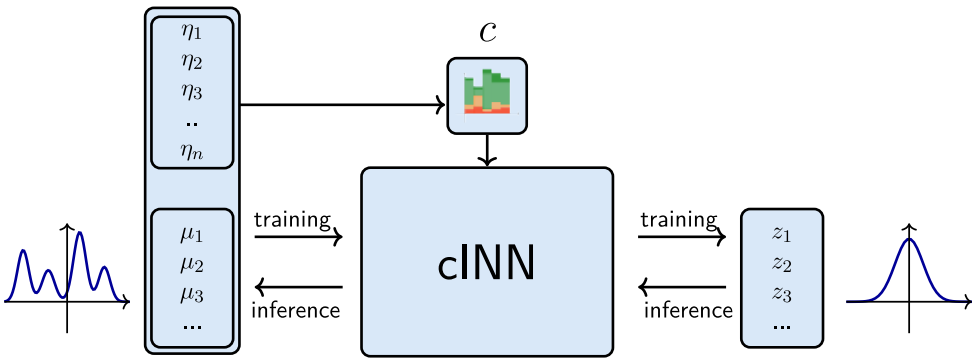

Figure 2. The network setup, training and inference process. The nuisance parameters for the shape-changing systematic uncertainties η_i effects are exclusively represented in the conditions and are not used as a network input.

3.3. Network Setup

The training and validation datasets consist of 1.5 million and 150,000 samples, respectively. The cINN model has been implemented in PyTorch [8] using the Framework for Easily Invertible Architectures (FrEIA) [1] library. The cINN has 12 GLOW ACBs and 12 permutation layers. The sub-networks in the GLOW blocks have 3 layers with 128 nodes with ReLU activation each. The network has been trained for 11,000 epochs with the Adam optimizer [9] and the model with the lowest validation loss has been evaluated further. For the learning rate, an initial value of 10^{-3} has been gradually decreased to 10^{-5} through a cosine learning rate scheduler.

4. Signal Strength Modifier Parameter Inference

The trained network's performance can be characterized by comparing the validation set's distribution in the latent space to the expected normal distribution; by computing the calibration errors and evaluating the median calibration error; by comparing the predicted μ to the MC truth and, by comparing the posterior samples to the prior distributions.

4.1. Latent Space Distribution

Since the posterior samples are produced through the backward evaluation of the normal-distributed latent space samples, it is crucial to check whether the network maps the input samples to the latent space correctly. The latent space distribution for three selected output nodes are shown in fig. 3. For all output nodes the output distribution closely follows the required normal distribution. For this reason, sampling from the normal distribution for posterior inference is justified.

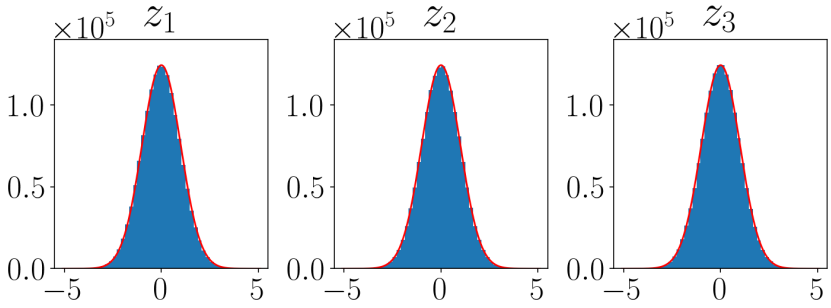

Figure 3. The latent space distribution histogram (blue) and the expected normal distribution $\mathcal{N}(0, 1)$ (red).

4.2. Calibration Curves

The posteriors for a fixed condition are obtained by sampling from the latent space and inverting the network. For this reason, it is important to check whether the trained model is free of biases in the predicted posterior distributions. In case of a perfectly calibrated network, the number of predicted posterior samples containing the true MC value in the q quantile (N_{in}) divided by the total number of posteriors N equals q. The calibration error e_{cal} is defined as

$$e_{cal}(q) = \frac{N_{in}}{N} - q \tag{4}$$

and its absolute median is a measure of miscalibration of the model. The calibration curves (the fraction of histograms N_{in}/N as a function of q) for all parameters show similar features where those of four selected parameters are shown in fig. 4. As the calibration errors for all parameters vary around zero, the calculated blue curves fluctuate around the gray identity line. The maximum median calibration error among all parameters is $e_{cal} \approx 4\%$. Thus, the trained model is not showing any strong biases towards a parameter region.

4.3. Predictions and Posterior Distributions

Compared to maximum likelihood fits, parameter inference is ultrafast. While the maximum likelihood fit required around 12 minutes for parameter inference, the cINN needed only roughly 0.4 seconds. The parameter's predictions and uncertainties are determined from the mean of the posteriors and the corresponding 68% quantiles, respectively. The inferred parameters can be classified into three categories:

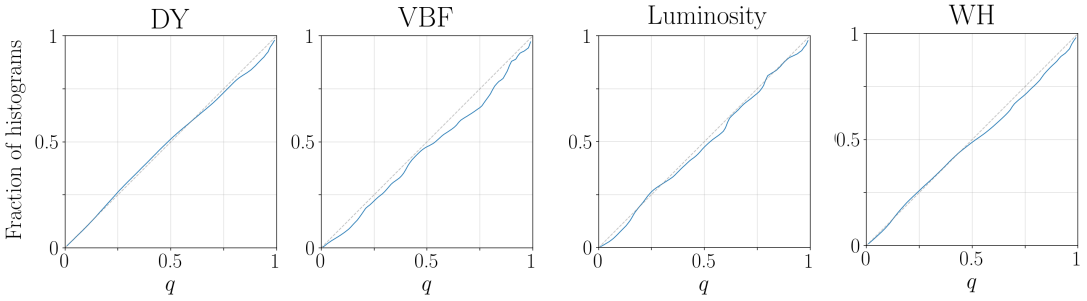

Figure 4. The fraction of histograms N_{in}/N as a function of q.

- *Well-reconstructed parameters*: these parameters' posteriors are narrow distributions around the true MC value and have low uncertainty. These parameters have a large impact on the input histogram, which the cINN can fully identify and reconstruct.
- *Weakly-reconstructed parameters*: the network has become sensitive to the parameter but yields predictions of low confidence. These parameters induce moderate changes in the histogram or the cINN cannot fully resolve their effects.
- *Unrecognized parameters*: for these parameters, the network has no sensitivity. As these parameters have a marginal impact on the conditions c, it fits the parameter value to the average of the prior distribution and returns the prior as the posterior distribution. The latter can be seen from the weak statistical dependence of the condition on the parameter, which directly yields

$$p(x|c) \approx p(x). \tag{5}$$

The 2D histograms of the true value plotted against the selected predicted parameters is shown in fig. 5. The corresponding posterior distributions for the Standard Model expectation for these parameters are shown in fig. 6 qualitatively.

For the well-reconstructed parameters (such as the nuisance parameter of the dominant Drell-Yan background process) all predictions scatter closely around the true MC value. The posteriors are narrow, hence the uncertainties are small. For these parameters, the trained network shows the highest sensitivity. Conversely, the unrecognized parameters' (as the low VBF background's) predictions scatter around the average of the prior distribution while the predicted posterior resembles the prior. Weakly-reconstructed parameters (as the luminosity nuisance) show an in-between state with some sensitivity developing from the mean of priors. Depending on the region, the posterior distributions still show some similarities to the priors; nevertheless some deviations in the overall posterior shape from the prior can already be observed.

For the WH signal process, the signal strength modifier parameters cannot be inferred with the same sensitivity over the whole prior region. For this reason, the predictions for smaller μ differ from the MC truth and the posteriors are broader. This behaviour results in a tail-like structure in the 2D histogram. The overall posterior histogram shape has some similarities with that of the priors in this low-sensitivity region. At the same time, the posteriors become more peaked and symmetric in the more sensitive region.

5. Conclusion

In this work, a method for signal strength modifier parameter inference with conditional invertible neural networks has been presented. The trained model reconstructs these parameters and does so several orders of magnitude faster than many-parameter maximum likelihood fits, with an observed 1800-fold speed-up. The trained model's latent space distribution matches the

expectation and the calibration curves do not show any strong biases. The network's resolution can be characterized with the widths of the posteriors, from which the parameter uncertainties can be obtained. As the network model is continuous and continuously differentiable cINNs posses generative properties and have the potential to be used in differentiable analysis workflows as well.

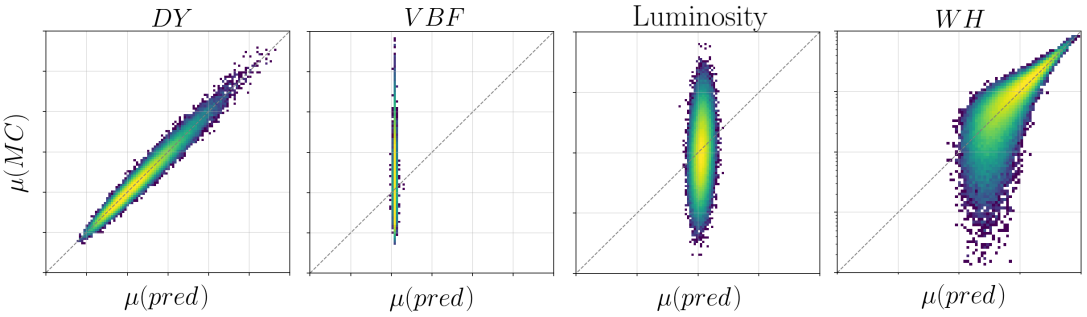

Figure 5. Comparison of the predicted μ to the true MC values for three nuisance parameters and a signal strength modifier.

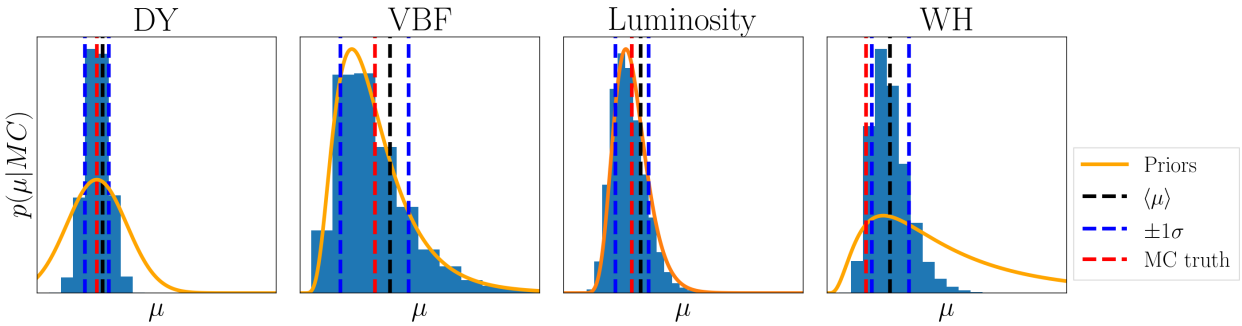

Figure 6. The posterior (blue) and prior distributions (orange) with the location of the MC truth, predictions and 68% quantile edges, qualitatively.

References

[1] Ardizzone L, Lüth C, Kruse J, Rother C and Köthe U 2019 Guided Image Generation with Conditional Invertible Neural Networks *Preprint* arXiv:1907.02392

[2] Radev S, Mertens U, Voss A, Ardizzone L and Köthe U 2022 BayesFlow: Learning complex stochastic models with invertible neural networks *IEEE Trans. Neural Netw. Learn. Syst.* **33** 1452-66

[3] Ksoll V, Ardizzone L, Klessen R, Koethe U, Sabbi E, Robberto M, Gouliermis D, Rother C, Zeidler P and Gennaro M 2020 Stellar parameter determination from photometry using invertible neural networks *Monthly Notices of the Royal Astronomical Society* **499** 5447-85

[4] Bister T, Erdmann M, Köthe U, Schulte J 2022 Inference of cosmic-ray source properties by conditional invertible neural networks *Eur. Phys. J.* C **82** 171

[5] Bellagente M, Butter A, Kasieczka G, Plehn T, Rousselot A, Winterhalder R, Ardizzone L and Köthe U 2020 Invertible networks or partons to detector and back again *SciPost Phys.* **9** 74

[6] Kingma D and Dhariwal P 2018 *Advances in Neural Information Processing Systems* vol 31, ed Bengio S *et al* (Montréal, CA: Curran Associates, Inc.)

[7] Baak M, Gadatsch S, Harrington R and Verkerke W 2015 Interpolation between multi-dimensional histograms using a new non-linear moment morphing method *Nucl. Instrum. Methods Phys. Res. Section* A **771** 39-48

[8] Paszke et al. 2019 *Advances in Neural Information Processing Systems* vol 32, ed Wallach H *et al* (Vancouver, CA: Curran Associates, Inc.)

[9] Kingma D and Ba J 2017 Adam: A Method for Stochastic Optimization *Preprint* arXiv:1412.6980