ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

Physics-Aware Normalizing Flows: Leveraging Electric Circuit Models in Adversarial Learning

Benjamin Schindler[1], and Thomas Schmid[231]

1- Universität Leipzig, Machine Learning Group
Augustusplatz 10, Leipzig - Germany

2- Martin-Luther-Universität Halle-Wittenberg
Magdeburger Str. 8, Halle (Saale) - Germany

3- Lancaster University in Leipzig
Nikolaistrasse 10, Leipzig - Germany

Abstract. We introduce Physics-Aware Normalizing Flows, a novel framework combining data-driven generative modeling with a physical layer based on an Electric Circuit Model, ensuring adherence to electricity laws, sample fidelity, and explainability. Four existing Normalizing Flow architectures, including Real-NVP and NSF, were adapted to our adversarial regime and evaluated with promising results for the ad hoc determination of value ranges of physical quantities and the generation of labeled measurements based on an unlabeled dataset. By extensive data generation according to our self-explainable approach, Random Forest regressions of underlying physical quantities could be improved significantly, compared to the original dataset including omitted ground truth labels.

1 Motivation

Lately, generative artificial intelligence has attracted a high level of public attention. Despite the emergence of high-performing models, however, significant weaknesses in these methods have been revealed, such as hallucination [1] and violation of causal relationships and real-world conditions [2]. Approaches that integrate explicitly represented knowledge with purely statistical generative techniques have therefore gained increasing interest recently [3].

An important approach in this regard is the development of physics-aware generative modeling [2], which strengthens the credibility of the generated data as well as the explainability of the models themselves. To this end, invertible flow-based generative models, i.e. *Normalizing Flows (NF)* [4], deserve special attention due to their built-in self-explainability with support for both sampling and probability density estimation [5], as well as continuous improvements in modeling capacity and quality of the generated samples [6, 7, 5].

One way to strengthen the integrity of generated data is to embed physical models in NFs. This allows ad hoc synthesis of labeled training data using the physical forward model and an unlabeled dataset. Particularly in the field of impedance spectroscopy, where electrical circuit models are used to explain observed measurements, the determination of the corresponding parameter values remains a challenge, but one that can be solved by supervised learning [8]. In the following, a novel approach to combine NFs with physical models is presented, exemplary for electric circuits, but adaptable for a wide range of forward models.

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

2 Normalizing Flows

2.1 Preliminaries

Normalizing Flows (NFs) are a class of generative models that define a transformation of a simple probability distribution (usually a multivariate Gaussian) with density p_Z into a more complex distribution with density p_Y [4]. Given a random variable Z with probability density function $p_Z(\mathbf{z})$ defined over $\mathbb{R}^D$, an NF is a parametric function g_θ such that $g_\theta(\mathbf{z}) = \mathbf{y}$, where $\mathbf{y}$ is the transformed variable, subject to the following conditions:

i) g_θ is invertible and differentiable.

ii) $\mathbf{y}$ follows a probability distribution $p_Y(\mathbf{y})$ over $\mathbb{R}^D$.

Analogous to neural networks, the capacity of NFs can be increased by stacking several layers making g a composition of several sub-functions g_i, each of which must be invertible and differentiable. More intricate interrelations within $p_Y(\mathbf{y})$ can be captured via *coupling*, where in each layer g_i one subset of dimensions is treated as independent while the other is conditioned on the former, with this relationship mapped by a neural network with trainable parameters.

p_Z is tractable, as we can both draw samples using $g_\theta(\mathbf{z}) = \mathbf{y}$ as well as compute the probability density $p_Y(\mathbf{y})$ of a sample $\mathbf{y}$ using the inverse function $f = g^{-1}$ according to the change of variable formula:

$$p_Y(\mathbf{y}) = p_Z(f(\mathbf{y})) \cdot |\det D(f(\mathbf{y}))| \tag{1}$$

where $D(f(\mathbf{y}))$ is the Jacobian Matrix of $f(\mathbf{y})$.

NFs can therefore be utilized bidirectionally, with g representing the 'generating' direction and f the 'normalizing' direction, which is essential for computing likelihoods and the classical training of NFs.

2.2 Training of Normalizing Flows

Classically, NFs are trained using maximum likelihood estimation (MLE), where the expected negative log-likelihood of the observed data under the model is minimized, commonly using stochastic gradient descent:

$$\min_\theta \quad - \mathbb{E}_{\mathbf{y}\sim p_Y} \left[\log f_\theta(\mathbf{y})\right]. \tag{2}$$

NFs can be trained without using f, instead by employing an *adversarial* neural network, the *discriminator* d_ϕ, which learns to differentiate the data distributions of synthetic and real data by approximating their Wasserstein-1 distance [9]. d_ϕ is counter-optimized, leading to the following minimax objective:

$$\min_\theta \max_{\phi \in F} \quad \mathbb{E}_{\mathbf{y}\sim p_Y}[d_\phi(\mathbf{y})] - \mathbb{E}_{\mathbf{z}\sim p_Z}[d_\phi(g_\theta(\mathbf{z}))] \tag{3}$$

where F is defined such that d_ϕ is 1-Lipschitz.

Adversarial training shows superior sample quality but poorer distribution coverage of the training data compared to classic NF [9]. Furthermore, adversarial training is notorious for being difficult to train, which is particularly evident in non-convergence and mode collapse [10].

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

3 Physics-Aware Normalizing Flows

3.1 Integrating Electric Circuits as a Physical Layer

In this work, an *Equivalent Circuit Model* (ECM, Fig. 1) is used as an explanatory model for spectral impedance data. ECMs play an important role in the investigation of charging processes and the state of health of batteries, through to the characterization of biological tissue and inflammatory diseases. In accordance with the given ECM (Fig. 1), impedance spectra can be derived as a function of the circuit elements R^e, R^s and C^e as well as a given set of AC frequencies $\omega = \{\omega_0, \ldots, \omega_K\}$:

$$\mathcal{Z}_\omega(R^e, R^s, C^e) = \left(\frac{R^e}{1 + j\omega_k C^e R^e} + R^s \right)_{k=0,\ldots,K} \quad \text{where} \quad j = \sqrt{-1}. \tag{4}$$

If we (i) constrain the codomain of g_θ to $\mathbb{R}_+^D$, and (ii) set $D = 3$, i.e. adjust the dimension of the NF to the number of model parameters of the ECM, you can now add equation (4) as an additional layer to the NF:

$$\tilde{g}_{\theta,\omega} = \mathcal{Z}_\omega \circ g_\theta \quad \text{where} \quad g_\theta : \mathbb{R}^3 \to \mathbb{R}_+^3 \quad \text{and} \quad \mathcal{Z}_\omega : \mathbb{R}_+^3 \to \mathbb{C}^K. \tag{5}$$

Note that $\tilde{g}_{\theta,\omega}$ is still differentiable but, unlike g, no longer invertible, yet we retain the ability to train $\tilde{g}_{\theta,\omega}$ within an adversarial regime:

$$\min_\theta \max_{\phi \in F} \quad \mathbb{E}_{\mathbf{y} \sim p_Y}[d_\phi(\mathbf{y})] - \mathbb{E}_{\mathbf{z} \sim p_Z}[d_\phi(\tilde{g}_\theta(\mathbf{z}))]. \tag{6}$$

This principled framework applies to all NF architectures and aligns with work towards the generalization of adversarial learning [11]. Also, $\mathcal{Z}_\omega$ can be exchanged with any other non-invertible but differentiable function.

3.2 Implementing the Adversaries

Using the *nflows* python package four NFs were extended according to 3.1:

a) *DiagonalFlows*[5] perform a dimension-wise scaling and shifting.
b) *LinearFlows*[5] model co-dependencies as a linear transformation.
c) *RealNVP*[6] employ multiple affine transformations.
d) *NSF*[7] uses rational quadratic splines as invertible transformations.

Real-NVPs and *NSFs* were implemented with *coupling* to map dimension-interrelations, parameterized by dense *Resnets*. An additional scaling layer, analogous to a *DiagonalFlow*, was appended at the end of the other NFs to improve training stability.

Adversarial Learning is carried out by Wasserstein-Discriminators with gradient penalty [12], implemented as *multilayer perceptrons*. Minibatch-discrimination [13], sample packing [14], gaussian instance noise [10], spectral normalization [15], and multiple discrimination steps have proven to be essential for successful training. Based on previous work, finite differences were used as additional discriminatory features [11]. Additional penalty terms were introduced to enforce positive capacitances and resistances (satisfying equation (5)) and to avoid peaked distributions with low deviation.

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

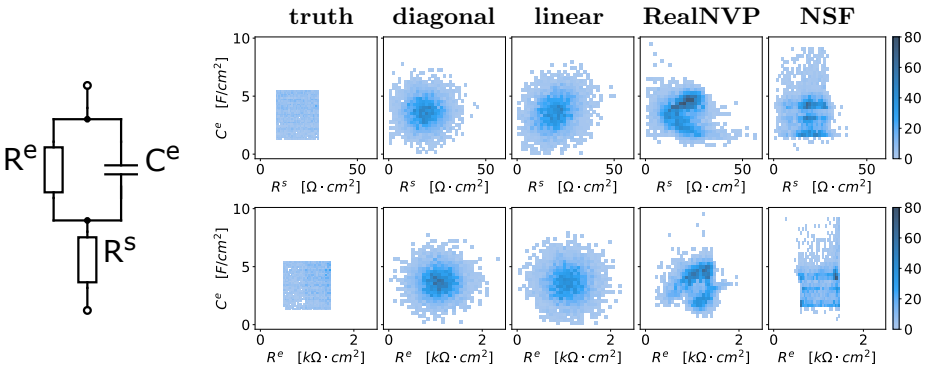

Fig. 1: Left: Equivalent Circuit Model employed within this work, consisting of two resistors and one capacitor. Right: Capacitance C^e plotted against Resistances R^e and R^s for truth and modeled data as 2D histograms (5000 Samples each, 30×30 bins).

4 Proof-of-Concept Study

4.1 Setup and Evaluation

We employ published spectral impedance measurements modeled after the cell line *HT29B6* under control conditions [8], which provides access to the ground truth ECM values during validation. To reflect real data availability, only 5000 random samples were used for training, whereas 50,000 original and generated samples were used for validation by computing the Kolmogorov-Smirnoff statistic (*1D-KS*), a two-dimensional KS approximation [16] (*2D-KS*), and the Correlation Similarity using Pearson Coefficients (*CS*). We assess the Supervised Learning performance (*SLP*) of each generated dataset with *Random Forests*, consisting of 50 estimators, trained on the generated and tested on the original data, by computing the Mean Absolute Percentage Errors (*MAPE*) in predicting all ECM values, with *SLP* being the average over all targets. For each NF architecture, hyperparameter tuning was performed using *optuna* and *manual search* to find a suitable configuration (i.e. flow and neural architectures, activations, regularizations, extracted features, and optimizer settings).

4.2 Results

True and modeled ECM value distributions are shown in Fig. 1. All NF architectures are able to capture the value ranges of the ECM based on the unlabeled dataset. Rigid value ranges and low correlations can be seen for the original data. Distributions derived from *DiagonalFlows* and *LinearFlows* remain Gaussians, whereas *RealNVP* and *NSF* are capable of modeling non-gaussian distributions.

Quantified evaluation results of the tuned NFs can be seen in Table 1. Looking at *1D-KS* and *2D-KS*, *DiagonalFlow* performs best in replicating the original distribution. *RealNVP* lead to the best *CS*, whereas *LinearFlow* generates the

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

Flow	Evaluation Metric			
	1D-KS	2D-KS	CS	SLP
Diagonal	<u>0.928</u>	<u>0.874</u>	0.558	0.946
Linear	0.919	0.857	0.567	<u>0.949</u>
RealNVP	0.925	0.851	<u>0.581</u>	0.934
NSF	0.902	0.821	0.576	0.887

Train Set	N	*MAPE* [%]		
		R^s	R^e	C^e
original	5000	5.74	<u>0.27</u>	14.2
generated	5000	5.76	0.45	17.5
generated	25000	3.52	0.33	14.3
generated	50000	<u>3.47</u>	0.30	<u>12.4</u>

Table 1: Left: Performances of adversarial NF architectures derived from comparing the validation set with the same amount of generated measurements. Evaluation metrics have been scaled to $[0, 1]$ with 1 being the best. Right: Mean Absolute Percentage Error (*MAPE*) of Random Forests in predicting ECM values trained on original and generated datasets with sample size N. Generated data is derived from *LinearFlows* trained on the original data without labels.

best training data for supervised learning, which we have examined in more detail in Table 1. The original data leads to superior predictions compared to the same amount of generated data, whereas any number of training samples can be generated. As the number of generated samples increases, the *MAPE* decreases significantly, so that with a tenfold sample size, the predictions could be even improved compared to the labeled original data for R^s (from 5.7% to 3.5%) and C^e (from 14.2% to 12.4%), and brought to line for R^e (0.27% and 0.30%).

4.3 Discussion

We show that various NFs, including state-of-the-art *NSFs* and *RealNVPs*, can be used for physics-aware modeling using a physical ECM layer and adversarial training. In contrast, classical NFs would not be able to model the underlying quantities of interest. Since the generated samples follow the ECM, we get certainty about sample quality and compliance with the physical laws, as well as real-world coverage through the feedback of the discriminator with access to a dataset, without any prior knowledge about the true value ranges of the physical quantities. Finding suitable learning configurations through tuning remains challenging, as multiple regularization techniques and batch-information are necessary, which aligns to general observations of adversarial learning [9, 13, 10].

For our dataset, however, we recommend rather simple NFs like *DiagonalFlow* and *LinearFlow* as their capacities are sufficient and much easier to train. This is especially true for the goal of training data generation for supervised learning, as generating some samples outside the original value ranges is less bad than the risk of omitting entire regions. It is also worth noting that our method trained on unlabeled data can generate a better-performing training dataset than the original, including the omitted labels.

While ECMs already find diverse applications in engineering and medicine, physics-aware adversarial learning extends to any domain featuring differentiable, non-invertible forward models. Thanks to auto-differentiation in modern deep learning frameworks, no explicit derivations need to be formulated.

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

5 Conclusions and Outlook

With the present study, we have introduced a novel approach for physics-aware NFs that incorporates the physical characteristics of a given electric circuit. We have demonstrated the ability of this approach to effectively determine underlying physical quantities of measurements and generate labeled training data. In future work, the application of the presented method to more complex ECMs and datasets, the embedding of physical forward models from other domains, as well as its application to experimental measurements, will be of special interest.

References

[1] Ziwei Ji, Nayeon Lee, Rita Frieske, Tiezheng Yu, Dan Su, Yan Xu, Etsuko Ishii, Ye Jin Bang, Andrea Madotto, and Pascale Fung. Survey of hallucination in natural language generation. *ACM Comput. Surv.*, 55(12), mar 2023.

[2] L. Manduchi et al. On the challenges and opportunities in generative ai. *arXiv preprint arXiv:2403.00025*, 2024.

[3] M. De Boer, Q. T. S. Smit, M. A. van Bekkum, A. Meyer-Vitali, and T. Schmid. Modular design patterns for generative neuro-symbolic systems. In *ESWC 2024 / Generative Neuro-Symbolic AI Workshop (GeNeSy 2024)*, Crete, Greece (in press), 2024.

[4] D. Rezende and S. Mohamed. Variational inference with normalizing flows. In *International conference on machine learning*, pages 1530–1538. PMLR, 2015.

[5] I. Kobyzev, S. D. Prince, and M. A. Brubaker. Normalizing flows: An introduction and review of current methods. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 43(11):3964–3979, nov 2021.

[6] L. Dinh, J. Sohl-Dickstein, and S. Bengio. Density estimation using real NVP. In *International Conference on Learning Representations*, 2017.

[7] C. Durkan, A. Bekasov, I. Murray, and G. Papamakarios. Neural spline flows. *Advances in neural information processing systems*, 32, 2019.

[8] B. Schindler, D. Günzel, and T. Schmid. Transcending two-path impedance spectroscopy with machine learning: A computational study on modeling and quantifying electric bipolarity of epithelia. *Int J Adv Lif Sci*, 13(3-4):134–148, 2021.

[9] A. Grover, M. Dhar, and S. Ermon. Flow-gan: Combining maximum likelihood and adversarial learning in generative models. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32, 2018.

[10] L. Mescheder, A. Geiger, and S. Nowozin. Which training methods for gans do actually converge? In *International conference on machine learning*, pages 3481–3490, 2018.

[11] B. Schindler, D. Günzel, and T. Schmid. Neural noise module: Automated error modeling using adversarial neural networks. In *International Conference on Bioelectromagnetism, Electrical Bioimpedance, and Electrical Impedance Tomography 2022*, 2022.

[12] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A.C. Courville. Improved training of wasserstein gans. *Advances in neural information processing systems*, 30, 2017.

[13] Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, and Xi Chen. Improved techniques for training gans. *Advances in neural information processing systems*, 29, 2016.

[14] Z. Lin, A. Khetan, G. Fanti, and S. Oh. Pacgan: The power of two samples in generative adversarial networks. *Advances in neural information processing systems*, 31, 2018.

[15] T. Miyato, T. Kataoka, M. Koyama, and Y. Yoshida. Spectral normalization for generative adversarial networks. In *International Conference on Learning Representations*, 2018.

[16] G. Fasano and A. Franceschini. A multidimensional version of the kolmogorov–smirnov test. *Monthly Notices of the Royal Astronomical Society*, 225(1):155–170, 1987.