Neuro-physical dynamic load modeling using differentiable parametric optimization

Shrirang Abhyankar
Electricity Infrastructure and Buildings Division
Pacific Northwest National Laboratory
Richland, United States
slaven.peles@pnnl.gov

Ján Drgoňa
Physics and Computational Sciences Division
Pacific Northwest National Laboratory
Richland, United States
jan.drgona@pnnl.gov

Aaron Tuor
Physics and Computational Sciences Division
Pacific Northwest National Laboratory
Richland, United States
aaron.tuor@pnnl.gov

Andrew August
Physics and Computational Sciences Division
Pacific Northwest National Laboratory
Richland, United States
andrew.august@pnnl.gov

Abstract—In this work, we investigate a data-driven approach for obtaining a reduced equivalent load model of distribution systems for electromechanical transient stability analysis. The proposed reduced equivalent is a neuro-physical model comprising of a traditional ZIP load model augmented with a neural network. This neuro-physical model is trained through differentiable programming. We discuss the formulation, modeling details, and training of the proposed model set up as a differential parametric program. The performance and accuracy of this neuro-physical ZIP load model is presented on a medium-scale 350-bus transmission-distribution network.

Index Terms—Differentiable Optimization, Load Modeling, Neural Networks, Transient stability

I. INTRODUCTION

The electrical power grid is going through a transformation of technologies and processes spurred by the deep proliferation of renewable energy sources, smarter control through increasing use of power electronic sources, and availability of high-fidelity measurement sources such as PMUs. Most of the developments in the electrical grid are emanating from the medium to low-voltage distribution grid. With progressively growing deployment of distributed energy resources (DERs), such as photovoltaics (PV), storage devices, electric vehicles, and microgrids, the characteristics of distribution systems are becoming more complex, both statically and dynamically [1]. Extensive work on load modeling has been carried out by power system researchers [2]–[5] to guide the development of load models through component-based [3], measurement-based approaches [6], [7], dynamic equivalencing [8]–[10], and neural networks [11], [12].

In this paper, we propose a novel neuro-physical load model (reduced equivalent model of distribution systems) for electromechanical stability analysis. The neuro-physical model is a convex combination of the physics-based model and neural network, i.e., it is partly physics-based and partly neural network. The proportions are optimized during training time to realize the benefits of both. Thus, if a given physics-based model has a good fitting to the data then the neuro-physical model will bias the proportion towards it, while it lean towards neural network if the fitting is bad. Moreover, addition of the neural network will capture the missing characteristics of the data that cannot be captured through a physics-based model only because of model or parameter inaccuracies. The limitation of the proposed model, as with data-driven approaches, is the training time and the quality and volume of the training data. However, in our limited experiments, we have observed that the neuro-physical model requires less data for training than using a neural network only as the physical-based model acts a prior.

In this work, the proposed neuro-physical dynamic equivalent comprises of a traditional physics-based ZIP load model augmented with a neural network. This neuro-physical model is set up as a constrained least squares problem and solved through the use of a novel differentiable parametric programming approach that optimizes the physics parameters and the neural network weights together. The neuro-physical model can be used as a surrogate or a reduced-order equivalent of the external area or downstream distribution system to accelerate transient stability simulations of large-scale transmission-distribution grids.

II. NEURO-PHYSICAL DYNAMIC LOAD MODEL

The goal of the work is to develop an a reduced equivalent of the distribution network for stability analysis. Figure 1 illustrates the transmission (T) and distribution (D) networks connected at a single boundary bus. We assume presence of measurement devices (such as PMUs) to measure the voltages $\bar{V}(t) = V(t)\angle\theta_V(t)$ and the currents $\bar{I}(t) = I(t)\angle\theta_I(t)$ at the boundary bus. Using these measurements, an equivalent reduced load model of the external area needs to be discerned given the voltage measurements at the boundary. We assume a following functional form for the power drawn $S = f(\bar{V})$. 

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
where \( S \) is the complex power drawn that consists of two components - real power \( P \) and reactive power \( Q \).

\[
V \angle \theta_V \quad I \angle \theta_I \quad \text{Boundary bus}
\]

**Fig. 1: Measurements at the transmission-distribution interface**

In our work, the proposed neuro-physical load comprises of two components - a physics-based model and a neural network model. We first discuss the ZIP load model which is our choice of physics-based load model used in this paper. Though a ZIP load model is overtly simplistic and may not be an accurate representation of the downstream distribution system, it was solely chosen for implementation ease. Other types of load models, such as motor load models, or composite load models can be also incorporated in the formulation.

**A. Physics-based model**

For the physics based model, we will restrict our discussion to the commonly used ZIP load model. This model expresses the real and reactive power drawn \( P \) and \( Q \) as a 2nd order polynomial function of the voltages.

\[
\hat{P}(t) = \alpha_p \tilde{P}(t_0) + \alpha_i \tilde{P}(t_0) \frac{V(t)}{V(t_0)} + \alpha_z \tilde{P}(t_0) \frac{V(t)^2}{V(t_0)} \quad (1a)
\]

\[
\hat{Q}(t) = \beta_p \tilde{Q}(t_0) + \beta_i \tilde{Q}(t_0) \frac{V(t)}{V(t_0)} + \beta_z \tilde{Q}(t_0) \frac{V(t)^2}{V(t_0)} \quad (1b)
\]

The initial load power \( \tilde{P}(t_0), \tilde{Q}(t_0) \) and the initial voltage \( V(t_0) \) are known parameters. The coefficients of the load model, \( \alpha \) and \( \beta \) need to be determined. Typically, this is done through solving the following constrained least-squares problem for \( \alpha_p, \alpha_i, \alpha_z \) and \( \beta_p, \beta_i, \beta_z \).

\[
\min_{\alpha,\beta} \sum_{k \in K} \sum_{t \in T} (\hat{P}(t, k) - P^*(t, k))^2 + (\hat{Q}(t, k) - Q^*(t, k))^2 \quad (2a)
\]

\[
s.t. \quad \alpha_p + \alpha_i + \alpha_z = 1 \quad (2c)
\]
\[
\beta_p + \beta_i + \beta_z = 1 \quad (2d)
\]
\[
\alpha_p, \alpha_i, \alpha_z \geq 0 \quad (2e)
\]
\[
\beta_p, \beta_i, \beta_z \geq 0 \quad (2f)
\]

Here, \( K \) is the set of time-series trajectories, and \( T \) is the number of time steps for each trajectory \( k \). The measured power references \( P^*(t, k) \) and \( Q^*(t, k) \) at each time-step \( t \) is calculated from the voltage and current measurements at the boundary bus.

\[
P^*(t, k) = V(t, k) I(t, k) \cos(\theta_V(t, k) - \theta_I(t, k)) \quad (3a)
\]
\[
Q^*(t, k) = V(t, k) I(t, k) \sin(\theta_V(t, k) - \theta_I(t, k)) \quad (3b)
\]

**B. Neuro-physical ZIP model**

While the physics-based ZIP load model is simplistic, it lacks the fidelity of capturing the complex dynamics primarily because of its restriction to second-order polynomial form. Employing different load models, such as WECC composite load model, is a possibility but it is an extremely complex model and estimating its parameters is a big challenge. In this work, we instead divert to improving the accuracy of a ZIP model using deep neural networks. Thus, the neuro-physical ZIP model combines the benefits of simplicity and some notion of interpretability from the physics-based model along with the expressiveness of the neural network to discover complex dynamics.

The neuro-physical ZIP model expresses the real \( P(t) \) and reactive \( Q(t) \) powers as a convex combination of the physics-based and neural models as follows:

\[
P_{\text{fit}}(t, k) = a \hat{P}(t, k) + (1 - a) \hat{P}(t, k) \quad (4a)
\]
\[
Q_{\text{fit}}(t, k) = b \hat{Q}(t, k) + (1 - b) \hat{Q}(t, k) \quad (4b)
\]

Here, \( \hat{P} \) and \( \hat{Q} \) is the output from the physics-based model (1a), (1b) and \( \tilde{P} \) and \( \tilde{Q} \) is the output from the neural network model, i.e., \( \{ \tilde{P}(t), \tilde{Q}(t) \} = \pi_\Theta(V(t), \theta_V(t)) \), where \( \pi_\Theta \) is the neural network model with weights \( \Theta \). The overall model is given as a convex combination (4) of the polynomial and neural model where \( a, b \) are scalars that determine the proportion of the physics-based and neural-network outputs. Depending on their values, the proposed neuro-physical ZIP model can range from a pure physics-based model \( (a = 1, b = 1) \) to an only neural network model \( (a = 0, b = 0) \).

To train the model (4) we need to solve the following constrained least squares problem:

\[
\min_{\Theta,\alpha,\beta} \sum_{k \in K} \sum_{t \in T} ((P_{\text{fit}}(t, k) - P^*(t, k))^2 + (Q_{\text{fit}}(t, k) - Q^*(t, k))^2) \quad (5a)
\]

\[
s.t. \quad \alpha_p + \alpha_i + \alpha_z = 1 \quad (5c)
\]
\[
\beta_p + \beta_i + \beta_z = 1 \quad (5d)
\]
\[
\alpha_p, \alpha_i, \alpha_z \geq 0 \quad (5e)
\]
\[
\beta_p, \beta_i, \beta_z \geq 0 \quad (5f)
\]
\[
0 \leq a \leq 1, 0 \leq b \leq 1 \quad (5g)
\]

With optimization variables representing the coefficients of the polynomial model \( \alpha_p, \alpha_i, \alpha_z, \beta_p, \beta_i, \beta_z, a, b \), and the weights \( \Theta \) for the neural network model \( \pi_\Theta \).

**III. DIFFERENTIABLE PARAMETRIC OPTIMIZATION**

In this work, we leverage differentiable parametric optimization (DPO) to learn the solution of the original constrained
optimization problem. A generic formulation of the DPO is given as follows:

$$
\min_{\Theta} \frac{1}{m} \sum_{i=1}^{m} f(x^i, \xi^i) \tag{6a}
$$

subject to:

$$
g(x^i, \xi^i) \leq 0, \tag{6b}
$$

$$
h(x^i, \xi^i) = 0, \tag{6c}
$$

$$
x^i = \pi_{\Theta}(\xi^i), \tag{6d}
$$

$$
\xi^i \in \Xi, \quad \forall i \in \mathbb{N}_1^m \tag{6e}
$$

Where $\Xi$ represents the sampled dataset, and $\xi^i$ represents $i$-th batch of the sampled problem data. The vector $x^i$ represents optimized variables that minimize the loss function while satisfying set of equality (6c) and inequality (6b) constraints. The mapping (6d) parametrized by $\Theta$ represents the solution of the underlying constrained optimization problem. In this work, we use the Neuromancer library [13] for solving the above load modeling given as constrained nonlinear least squares problem (5). Neuromancer is an open-source Python library built on Pytorch [14] infrastructure that allows to formulate and solve generic DPO problems (6) by sampling and gradient optimization using AdamW solver [15].

a) Constraints penalties: A simple approach to penalize the constraints violations is by augmenting the loss function $L_{\text{obj}}$ (6a) with the penalty functions:

$$
L_{\text{con}} = Q_g \|ReLU(g(x^i, \xi^i))\|_l + Q_h \|h(x^i, \xi^i)\|_l \tag{7}
$$

Where $l$ denotes the norm type and $Q_g$, $Q_h$ being the corresponding weight factors. The overall loss then becomes $L = L_{\text{obj}} + L_{\text{con}}$.

b) Neuro-physical ZIP model cast as a DPO problem: To solve the problem (5), we cast it in the form (6). Here, the data samples are represented by the power measurements $\xi^k = \{P^*(t, k), Q^*(t, k)\}$, the least square loss (5b) is captured by (6a), the physics-based model given by (1), (5c), (5d), (5e), (5f) is compactly represented by a set of equality (6c) and inequality (6b) constraints penalized during training via (7). While the parametric solution map to be trained (6d) represents the neural model of the problem (5).

IV. CASE STUDY

A. Power system network setup

The neuro-physical ZIP load model was tested on a synthetic transmission-distribution network made up of 200-bus synthetic network from the ACTIVSg test case repository [16]–[18] connected to a 141-bus distribution network from the MATPOWER [19] library. The two networks are connected at a single bus through an added interface line as shown in Fig. 2. The 200-bus transmission node has dynamic models of synchronous generators, exciters, and turbine-governors as given in the ACTIVSg test case repository.

B. Training and test data

Dynamic simulations of the combined 200-bus transmission+141-bus distribution network were run for three-phase fault on the 141-bus system at different locations. Three scenarios were considered and the following data was collected for training and testing the proposed model.

- Three-phase faults at different locations on the distribution grid. Data for 20 trajectories was collected.
- Three-phase faults on the transmission side at different locations with the distribution loads modeled as ZIP load model. Data for 50 trajectories was collected.
- Three-phase faults on the transmission side at different locations with the distribution loads modeled as a combination of ZIP and motor load. Data for 50 trajectories was collected.

Each fault simulation was for 10 second total with a fault applied for six cycles (or 0.1 seconds). The voltages at the transmission boundary bus and the current flowing to the distribution network were recorded for each trajectory.

C. Solution setup

The data for each scenario mentioned above was split into training and test data sets. 60% of the data, 20% for validation, and 20% for testing. We used the open-source library Neuromancer [13] for solving the problem cast as differential parametric optimization (DPO) problem (6). AdamW [15] with learning rate of 0.01 was used for the optimizer. The relative weights for equality and bounds constraints penalties were set to $Q_g = 1.0$, $Q_h = 1.0$. The neural network model we use has four layers with 20 nodes per layer. Overall, models converge to a solution in about 8 mins on CPU.

D. Results and discussion

1) Scenario 1: Three-phase faults at different locations on the distribution grid: Figures 3 and 4 shows the comparison of real and reactive power from the ZIP and trained Neuro-ZIP model against the reference trajectories. As seen in the figure, the neuro-ZIP model shows a close match with the reference trajectory, particularly for the real power (P).
the trajectories for scenario 1 resulted in reaching a steady-state after 4 seconds and hence we only used data for the first four seconds of the dynamics simulations.

**TABLE I: Prediction accuracy and constraints violations for each scenario**

| Scenario | MSE P(t) | MSE Q(t) | a  | b  | Con. viol. |
|----------|----------|----------|----|----|------------|
| 1        | 0.0004   | 0.0034   | 0.594 | 0.586 | 0.009      |
| 2        | 5.05e-06 | 3.32e-07 | 0.591 | 0.594 | 0.009      |
| 3        | 0.0006   | 0.0009   | 0.579 | 0.569 | 0.002      |

2) **Scenario 2: Three-phase faults on the transmission side at different locations:** In this scenario, the faults were applied on the transmission side. The loads at the distribution nodes were modeled as constant impedances. Figures 5 and 6 compare the performance of the neuro-ZIP model. For this scenario, the proposed model was able to track the reference trajectory very accurately compared to the traditional ZIP model only.

3) **Scenario 3: Three-phase faults on the transmission side at different locations with the distribution loads modeled as combination of ZIP + motor load.** This scenario is the same as scenario 2 with the exception that the loads on the distribution side are a combination of constant impedance and motor. Figure 7 and 8 has a comparison of the output from the ZIP and the trained neuro-ZIP model against the reference trajectory. The trained model is able to track the slow moving dynamics to a good degree, but it has some discrepancies in tracking the transient.

V. CONCLUSIONS AND FUTURE WORK

In this paper, we presented a neuro-physical load modeling approach for determining a dynamic load model of distribution grids for transient stability analysis. We formulate this problem
as a differentiable constrained nonlinear least-squares problem by learning a convex combination of a physics-based model and neural network. Our results show the proposed neuro-augmented modeling approach can greatly improve the prediction accuracy of the simplified physics-based model while satisfying the imposed constraints.

Extensions of the work include finding equivalent of larger networks with multiple boundary buses and use of more complex physics-based load models. The differentiable parametric optimization (DPO) method used in this paper represents a soft-constrained approach for obtaining an approximate parametric solution to the constrained optimization problems. In the future work, employing methods for hard-constraints satisfaction, such as the method presented in [20].

ACKNOWLEDGMENTS

This work was supported through the Data Model Convergence (DMC) initiative at Pacific Northwest National Laboratory (PNNL).

REFERENCES

[1] G. T. Heydt, “The next generation of power distribution systems,” IEEE Transactions on Smart Grid, vol. 1, no. 3, pp. 225–235, 2010.
[2] X. Wu, “New approaches to dynamic equivalent of active distribution network for transient analysis,” 2016.
[3] W. W. Price, K. A. Wriggau, A. Murdoch, J. V. Mitsche, E. Vaahedi, and M. El-Kady, “Load modeling for power flow and transient stability computer studies,” IEEE Trans. Power Syst., vol. 3, no. 1, pp. 180–187, Feb. 1988.
[4] S. Son, S. H. Lee, D.-H. Choi, K.-B. Song, J.-D. Park, Y.-H. Kwon, K. Hur, and J.-W. Park, “Improvement of composite load modeling based on parameter sensitivity and dependency analyses,” IEEE Transactions on Power Systems, vol. 29, no. 1, pp. 242–250, 2014.
[5] K. Zhang, H. Zhu, and S. Guo, “Dependency analysis and improved parameter estimation for dynamic composite load modeling,” IEEE Transactions on Power Systems, vol. 32, no. 4, pp. 3287–3297, 2017.
[6] M. Sadeghi and G. A. sarvi, “Determination of zip parameters with least squares optimization method,” in IEEE Electrical Power Energy Conference (EPEC), Oct. 2009, pp. 1–6.
[7] S. Z. Zhu, J. H. Zheng, L. Li, S. D. Shen, and G. M. Luo, “Effect of load modeling on voltage stability,” in IEEE Power Engineering Society Summer Meeting, vol. 1, Jul. 2000, pp. 395–400 vol. 1.
[8] H. Zhao, X. Lan, and H. Ren, “Nonlinear power system model reduction based on empirical gramians,” Journal of Electric Engineering, vol. 68, no. 6, 2017.
[9] D. Osipov, “Adaptive nonlinear model reduction for fast power system simulation,” Ph.D. dissertation, University of Tennessee, Knoxville, 2018.
[10] Z. Jiang, X. Zhang, Y. Xue, A. G. Tarditi, and Y. Liu, “Measurement-based power system dynamic model reduction using arx equivalents,” in 2019 IEEE Power Energy Society General Meeting (PESGM), 2019, pp. 1–5.
[11] M. A.-E.-A. Abd-El-Rehim, I. Helal, and M. A. Hassan Omar, “Multi-machine power system dynamic equivalents using artificial intelligence (ann),” in 2006 Eleventh International Middle East Power Systems Conference, vol. 1, 2006, pp. 197–207.
[12] F. Ma and V. Vittal, “A hybrid dynamic equivalent using ann-based boundary matching technique,” IEEE Transactions on Power Systems, vol. 27, no. 3, pp. 1494–1502, 2012.
[13] A. Tuor, J. Drgona, and E. Skomski, “NeuroMANCER: Neural Modules with Adaptive Nonlinear Constraints and Efficient Regularizations,” 2021. [Online]. Available: https://github.com/pnnl/neuromancer
[14] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga et al., “Pytorch: An imperative style, high-performance deep learning library,” in Advances in Neural Information Processing Systems, 2019, pp. 8024–8035.
[15] I. Loshchilov and F. Hutter, “Decoupled weight decay regularization,” arXiv preprint arXiv:1711.05101, 2017.
[16] A. B. Birchfield, T. Xu, K. M. Gegner, K. S. Shetye, and T. J. Overbye, “Grid structural characteristics as validation criteria for synthetic networks,” IEEE Transactions on Power Systems, vol. 32, no. 4, pp. 3258–3265, 2017.
[17] Creation of synthetic electric grid models for transient stability studies, 2017.
[18] T. Xu, A. B. Birchfield, and T. J. Overbye, “Modeling, tuning, and validating system dynamics in synthetic electric grids,” IEEE Transactions on Power Systems, vol. 33, no. 6, pp. 6501–6509, 2018.
[19] R. Zimmerman and C. Murillo-Sanchez, “Matpower 4.0 users manual,” Power Systems Engineering Research Center, Tech. Rep., 2011.
[20] P. L. Donit, D. Rolnick, and J. Z. Kolter, “DC3: A learning method for optimization with hard constraints,” CoRR, vol. abs/2104.12225, 2021. [Online]. Available: https://arxiv.org/abs/2104.12225