Active Learning for Physics Informed Neural Networks on AC-OPF Problems

Osanda Wewalwala

33905142

Submitted in partial satisfaction of the requirements for the degree of

C3001 Bachelor of Computer Science Advanced (Honours) Computer Science

Department of Computer Science

Monash University

Primary Advisor: Dr. Terrence Mak

2025

Table of Contents

Acknowledgements

I want to thank Terrence Mak, Allen Zhong and Markus Wagner for their invaluable assistance during this work.

If you have any troubles, requests, or issues please get in touch with

`owew0001@student.monash.edu`.

Abstract

AC-Optimal Power Flow (AC-OPF) is a fundamental problem within power system operation that looks to find the optimum generation setpoint for buses within a system. The complexity of this problem means that power system operators must use approximations which often leads to an increase in the financial and environmental burdens of power systems. Machine Learning has proved to be a promising avenue for AC-OPF solvers, specifically Physics Informed Neural Networks (PINN) have showed promise in their ability to increase the computational efficiency and enforce the constraints of a real power system. However, one of the biggest issues facing this avenue is the lack of quality data for use in training machine learning models. This study looks to employ the use of active learning techniques to investigate their effect on the training of PINN models. This research concludes that the active learning techniques employed on the model did not have a substantial effect on the training of the PINN model, both when varying the learning period and the number of models used to form the training dataset. Overall, this study finds inconclusive evidence towards the beneficial effects of active learning techniques on the training of PINN models for AC-OPF problems.

Chapter 1

Introduction

The AC- Optimal Power Flow (AC-OPF) problem is what lies at the heart of the power industry and is instrumental in the deliverance of energy to households and businesses worldwide. Machine learning techniques have showed much promise in the world of AC-OPF solvers, namely Physics Informed Neural Network models (Raissi et al., 2019) have shown that they are able to compute solutions to AC-OPF problems faster whilst enforcing the numerous constraints of the problem. This work analyses the effect of active learning techniques on PINN models for AC-OPF problems.

1.1 AC-OPF

When operating a power system, it is necessary to distribute the generation of power throughout the various generators of that system. The AC-OPF problem is fundamentally an optimisation problem that looks at how we can distribute this generation such that we minimise the cost of operation. A deceptively complex task due to the various non-linear constraints placed on this optimisation making the problem itself non-linear and non-convex (Molzahn and Hiskens, 2019).

AC-OPF is formulated as follows:

Let:

- N_b,N_l and N_g denote the number of buses, lines and generators in the system respectively
- v_n denote the active and reactive voltage at bus n
- p_n^g and q_n^g denote the active and reactive power generation, and p_n^d and q_n^d denote the active and reactive power demand at bus n respectively
- ℓ_{mn} denote the current in the line from bus m to bus n

The objective function of the AC-OPF problem can be formulated as:

$$\min_{p_g, q_g} \sum_{i=1}^{N_g} \mathbf{c}(p_i^g) + \mathbf{c}(q_i^g) \qquad (\mathbf{1})$$

Where $c(x)$ is the cost function.

Such that:

$$p_n = p_n^g - p_n^d \qquad \forall n \in N_b \quad (\mathbf{2})$$

$$q_n = q_n^g - q_n^d \qquad \forall n \in N_b \quad (\mathbf{3})$$

Where p_n and q_n are the active and reactive power injection at bus n. This constraint looks to preserve Kirchoff's Law

$$\underline{v}_n \leq v_n \leq \bar{v}_n \qquad \forall n \in N_b \quad (\mathbf{4})$$

Where $\underline{v}_n$ and $\bar{v}_n$ are the upper and lower voltage limits at bus n.

$$\ell_{mn} \leq \bar{\ell}_{mn} \qquad \forall mn \in N_l \quad (\mathbf{5})$$

Where $\bar{l}_{mn}$ is the upper line limits on the line from bus m to bus n.

Because of it's computaional complexity, power system operators often utilise various approximations of AC-OPF such as the DC-OPF formulation. Although widely used within industry, approximations of AC-OPF result in numerous complications including a large overhead in the cost and a significant increase in the environmental burden of the power industry. Estimates state that tens of billions of dollars (Cain et al., 2013) and over 1 billion metric tons of greenhouse gas emissions (Surana and Jordaan, 2019) could be mitigated with an efficient implementation of AC-OPF within practical applications.

1.2 Physics Informed Neural Networks

Traditional approaches to OPF problems such as the Gauss Seidel method, Non-linear Programming and the Newton-Raphson method often utilise iterative methods for finding solutions. However, due to the complexity of the problem these methods are time consuming and often unfeasible in a real-time context.

As of recent, there has been a vested interest in the implementation of machine learning techniques in order to solve AC-OPF problems. Work by Pourahmadi et al.(2025) and Xu et al., (2025) implement architecture based machine learning techniques in the form of SVMs and Convolutional Neural Networks (CNNs) respectively. Although these methods display a substantial increase in computational efficiency, they do not present any mitigation of constraint violation and as such may present unfeasible solutions when placed in pragmatic applications. These works do however convey that deep learning is a very promising avenue for AC-OPF solvers.

This work builds on implementations by Nellikkath and Chatzivasileiadis (2022) who deploy Physics Informed Neural Networks (PINNs) in order to formulate solutions to AC-OPF problems. PINNs are able to penalise models for constraint violations. Through a convex relaxation of AC-OPF, we are able to obtain a Lagrangian Dual Function for the AC-OPF problem. Accordingly, we are able to place Karush-Kuhn-Tucker conditions within the objective function of the model thus incentivising models to formulate solutions that do not violate the physical constraints of the power system.

Chapter 2

Methodology

2.1 Dataset

This study employed the IEEE 162 bus test case (Babaeinejadsarookolaee et al., 2021) in order to construct a dataset comprising of 2000 data points. Each data point was generated by scaling the original active and reactive loads p_n^d and q_n^d from the original test case by a random variable $X \sim U(0.8, 1.2)$.

These rescaled load profiles were then processed using the MATPOWER Interor Point Solver (Wang et al., 2007). If these rescaled demands yielded a successful AC-OPF solution, they were incorporated into the dataset for training the PINN.

Each data point consisted of the vector of all active and reactive power demands as the input, and the vector of all active and reactive power generation, together with the real and imaginary components of voltage as the output.

2.2 Implementing PINNs

The key way in which Physics Informed Neural Network models differentiate themselves from other machine learning techniques employed for AC-OPF solvers, is the fact that they mostly focus on the objective function of the network (Raissi et al., 2019) and not the architecture. Often altering the function to take into account the physical realities of the world. The ways in which these physical realities are encoded into the objective function changes from domain to domain.

Nellikkath and Chatzivasileiadis (2022) choose to encode the restrictions of a given power system by utilising a convex relaxtion of the AC-OPF problem, thus allowing them to access the Lagrangian dual function of the optimisation problem. They formulate the Lagrangian function as follows: [1]

Nellikkath and Chatzivasileiadis simplify the objective function $(\mathbf{1})$ to

$$\min_{\mathbf{v},\mathbf{G}} \mathbf{c}^{\mathbf{T}}\mathbf{G}$$

where $\mathbf{c}$ is the linear cost coefficients and $\mathbf{G}$ is the set of all active and reactive power generations.

The equality constraints $(\mathbf{2}),(\mathbf{3})$ are simplified to

$$\mathbf{v}^{\mathbf{T}}\mathbf{F_l}\mathbf{v} = a_l\mathbf{G} + b_l\mathbf{D}$$

where $l = [0 : \text{Number of equality constraints}]$ and $\mathbf{v}$ is the real and imaginary components of the voltage.

The inequality constraints $(\mathbf{4}),(\mathbf{5})$ are also simplified to

$$\mathbf{v}^{\mathbf{T}}\mathbf{M_k}\mathbf{v} = d_k\mathbf{D} + f_k$$

where $k = [0 : \text{Number of inequality constraints}]$

The Lagrangian Function of the AC-OPF can be written as:

$$\mathcal{L}(G, v, \lambda, \mu, D) = \mathbf{c}^{\mathbf{T}}\mathbf{G} + \sum_{l=1}^{L} \lambda_l(\mathbf{v}^{\mathbf{T}}\mathbf{F_l}\mathbf{v} - a_l\mathbf{G} - b_l\mathbf{D}) + \sum_{k=1}^{K} \mu_k(\mathbf{v}^{\mathbf{T}}\mathbf{M_k}\mathbf{v} - d_k^T\mathbf{D} - f_k)$$

Where L, K is the number of equality and inequality constraints respectively

[1] Full Derivation can be observed in (Nellikkath and Chatzivasileiadis, 2022)

Consequently, the KKT conditions can written as:

$$STAT_i = |c - \sum \hat{\lambda}_{l,i} a_l| + |\sum \hat{\lambda}_{l,i} F_l + \sum \hat{\mu}_{m,i} M_m|$$

$$COMP_i = \sum |\hat{\mu}_{m,i}(\hat{\mathbf{v}}^{\mathbf{T}}\mathbf{M_m}\hat{\mathbf{v}} - d_m^T\mathbf{D} - f_m)|$$

$$DUAL_i = \sigma(\hat{\mu}_{m,i})$$

$$PRIM_i = \sum |\hat{\mathbf{v}}^{\mathbf{T}}\mathbf{G}\hat{\mathbf{v}} - a_l^T\mathbf{G} - b_l^T\mathbf{D}|$$

Therefore the final loss function of the PINN model can be written as:

$$\sum_{i=1}^{N_t} \Lambda_G|\hat{\mathbf{G}} - \mathbf{G}| + \Lambda_V|\hat{\mathbf{v}} - \mathbf{v}| + \Lambda_L|\hat{\mathbf{L}} - \mathbf{L}|$$
$$+ \sum_{i=1}^{N_t+N_c} STAT_i + COMP_i + DUAL_i + PRIM_i$$

Where N_t, N_c are the number of data points and the number of collocation points respectively and $\Lambda_G, \Lambda_V, \Lambda_L, \Lambda_{KKT}$ are the weight of the generation loss, voltage loss, Lagrangian variable loss and KKT condition loss respectively.

2.3 Implementing Active Learning

Active Learning is a machine learning technique that aims to reduce the amount of labelled data required to train a model. The basic premise is that allowing models to 'choose' which data instances they are exposed to allows them to build much more robust models whilst minimizing the amount of labelled data required to train them.

This study implements a pool-based active learning cycle such that the model is initially trained on D data instances on the first epoch and selects the c data points incurring the highest loss across the m models. These data points are then incorporated into the pool of training data in the subsequent epoch. This means that at epoch n there will be $D + (n-1)c$ data points in the training pool.

Let:

- $\mathcal{U}$ denote the unlabeled pool of data
- L_0 denote the initial labeled dataset with $|L_0| = D$
- $\mathcal{M} = M_1, M_2, ..., M_m$ denote the **m** models in the active training loop
- n denote the n-th epoch of training

At epoch n the models $\mathcal{M}$ are trained on the labeled dataset L_{n-1}. Then, for every unlabeled instance $x_i \in \mathcal{U}$, each model M_j makes a prediction and incurs a loss of $l(M_j(x_i))$. Let $L(x_i)$ denote the cumulative loss of data point x_i across **m** models

$$L(x_i) = \sum_{j=0}^{\mathbf{m}} l(\mathbf{M}_j(x_i))$$

Let S_n be the c data points that incurred the highest loss

$$S_n = \text{argmax}^{\text{c}}\,(L(x_i)) \qquad x_i \in \mathcal{U}$$

S_n is then incorporated into the labelled dataset

$$L_n = L_{n-1} \cup S_n$$

Where $|L_n| = D + (n-1)c$

The reason that this methodology was chosen was under the assumption that the data points that incurred the highest loss across many models would also be the data points that were most difficult for the PINN model to regress. Thus contributing more towards the learning of the underlying patterns than a data point that was redundant within the data space.

Chapter 3

Results and Discussion

This section shows the analysis of active learning techniques on PINN models for AC-OPF problems. The chosen PINN model backpropogated on 3 Feed Forward Neural Networks predicting generation ($\mathbf{G}$), voltage ($\mathbf{v}$) and the dual variables ($\mathbf{L}$). Each of the networks had 3 hidden layers with 20,30 and 50 hidden nodes in each layer respectively as shown below in table 3.2.

Table 3.1: Neural Network modules and the size of their hidden layers

Neural Network	$\lvert h_1 \rvert$	$\lvert h_2 \rvert$	$\lvert h_3 \rvert$
Generation ($\mathbf{G}$)	20	20	20
Voltage ($\mathbf{v}$)	30	30	30
Dual Variables ($\mathbf{L}$)	50	50	50

These predicted variables were then input into a final PINN layer that calculated the physical loss of the proposed solution utilising the above mentioned loss function. The calculated loss was used to backpropogate across the 3 networks utilsing the Adam optimiser.

The above mentioned active learning technique outlined in section 2.3 was then tested across 2 primary parameters

1. **Learning Period** - that is the number of epochs for which new data points were added to the training pool of data.

2. **Population of Models** - that is the number of models $|\mathcal{M}|$ used to calculate which data points got added to the training pool

3.1 Variation in Learning Period

Presented below is the loss curves for the PINN models varying in the length of their learning period.

Table 3.2: Initial dataset size and number of additional data points added every epoch c

Length of Learning Period	D	c	**Final Dataset Size**
0 epochs	1000	0	1000
10 epochs	100	90	1000
15 epochs	100	60	1000
20 epochs	100	50	1000

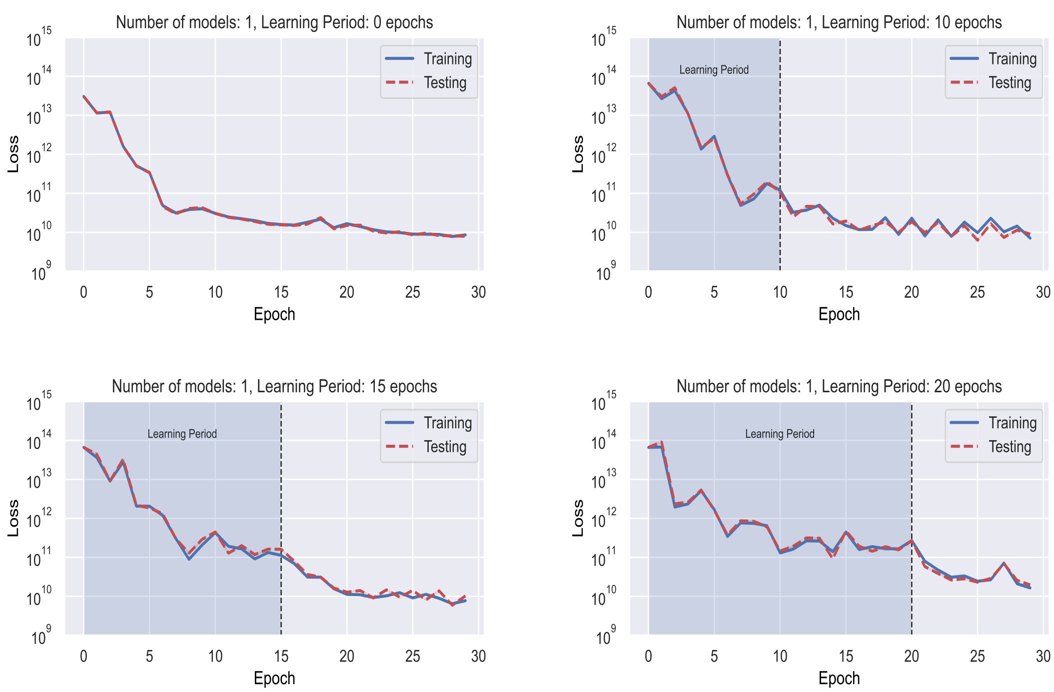

Figure 3.1: PINN Models varying in learning period

It is observed primarily that the introduction of active learning techniques decreases the smoothness of the loss curve, this is likely due to the fact that the sample distribution to which the models are adapting to gets altered every epoch, thus the model's overfitting to the previous epoch's training pool causes an increase in the loss of the current epoch.

Although a general downwards trend is observed in all iterations of the model, employing active learning appears to slow down the decrease of the loss within the learning period, this is especially observed within the model for which the learning period was 20 epochs. This may have been a consequence of the fact that the models with active learning were given significantly less training data initally and data was slowly added across the epochs. Whereas the model not employing active learning was given all data points from the beginning thus being able to develop deeper instances of pattern recognition across it's training dataset

It appears as though the model not employing active learning decreased the fastest and smoothest. From this we can conclude that the length of the learning period does not have a substantial effect on the training of the PINN model.

3.2 Variation in Population of Models

Presented in Fig 3.2 is the loss curves for the PINN models varying in the number of models $|\mathcal{M}|$ utilised. Shown in bold is the average testing and training curves across the $|\mathcal{M}|$ models.

It is observed from Fig 3.2 that as $|\mathcal{M}|$ increases, the epochs after the learning period become smoother. This is likely a consequence of the fact that as the number of models we average across increases, the curve extrapolated becomes more representative of the population model and thus smoother.

Although smoother, an increase in $|\mathcal{M}|$ did not decrease the loss incurred by the model overall. It appears as though all models averaged approximately the same loss value of 1×10^{10} after 30 epochs.

From these points of analysis, we are able to conclude that varying the number of

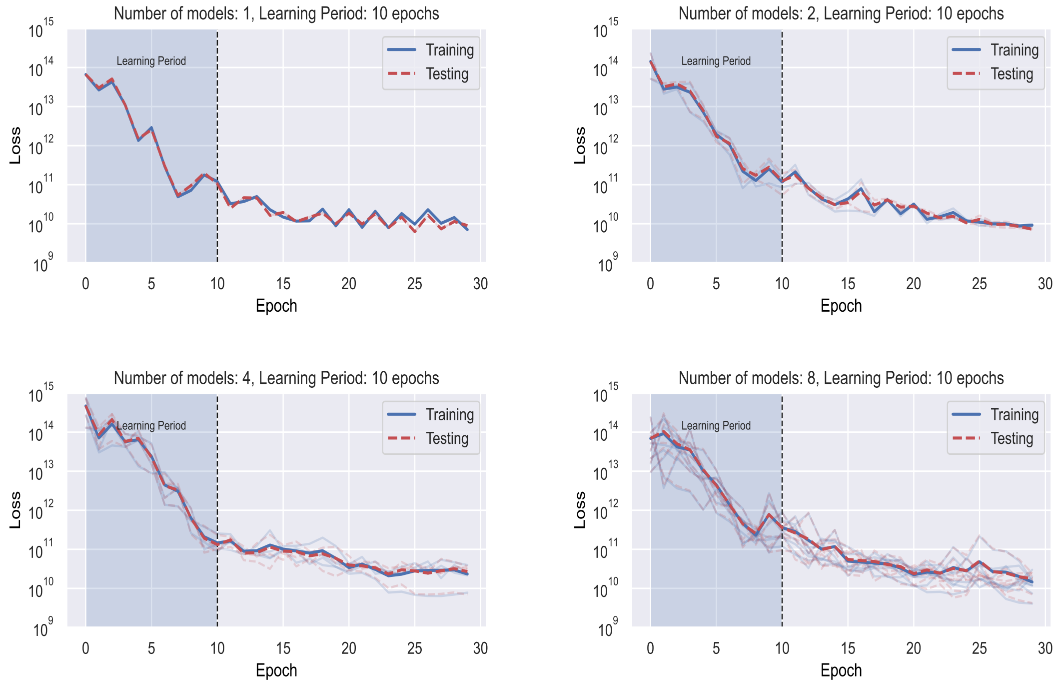

Figure 3.2: PINN Models varying in the number of models for active learning

models from which the training pool of data is built from does not have a substantial effect on the training of the PINN models.

3.3 Overall

The results from section 3.1 and 3.2 both lead to the conclusion that the employed active learning technique did not have a meaningful impact on the training of the PINN models. This brings into question the validity of the active learning technique within the context of PINN models. Additionally, the lack of readily available test cases may have been a limiting factor in the experiment leading to an unreliable training dataset being built.

Chapter 4

Conclusion and Future Work

4.1 Project summary

This study aimed to investigate the impacts of active learning techniques on Physics Informed Neural Networks for AC-OPF problems. By utilising an implementation of PINNs (Nellikkath and Chatzivasileiadis, 2022), analysis was conducted on the effects of varying learning period and the number of models utilised to build the training data pool. The active learning techniques implemented did not appear to have any substantial improvements on the training of the PINN model both when varying the length of the learning period and the number of models.

A significant hurdle within the research process was the lack of reliable test cases which may have lead to the results observed within this work. Overall, this study has discovered inconclusive evidence towards the effectiveness of active learning techniques on PINN Models for AC-OPF

4.2 Future work

There are numerous avenues for continued exploration within this topic. Perhaps the biggest point of further study could be an exploration into how to produce higher quality AC-OPF datasets. Additionally, further exploration into the specific boundaries of demand values where PINNs struggle to produce viable solutions may prove to be quite useful.

Chapter 5

References

- Babaeinejadsarookolaee, S., Birchfield, A., Christie, R.D., Coffrin, C., DeMarco, C., Diao, R., Ferris, M., Fliscounakis, S., Greene, S., Huang, R., Josz, C., Korab, R., Lesieutre, B., Maeght, J., Mak, T.W.K., Molzahn, D.K., Overbye, T.J., Panciatici, P., Park, B., Snodgrass, J., Tbaileh, A., Hentenryck, P.V., Zimmerman, R., 2021. The Power Grid Library for Benchmarking AC Optimal Power Flow Algorithms. https://doi.org/10.48550/arXiv.1908.02788
- Cain, M.B., O'Neill, R.P., Castillo, A., 2012. History of Optimal Power Flow and Formulations.
- Huang, W., Pan, X., Chen, M., Low, S.H., 2022. DeepOPF-V: Solving AC-OPF Problems Efficiently. IEEE Trans. Power Syst. 37, 800–803. https://doi.org/10.1109/TPWRS.2021.3114092 MATPOWER Interior Point Solver (MIPS) User's Manual [WWW Document], n.d. URL https://zenodo.org/records/4073324 (accessed 11.7.25).
- Molzahn, D.K., Hiskens, I.A., 2019. A Survey of Relaxations and Approximations of the Power Flow Equations. Found. Trends® Electr. Energy Syst. 4, 1–221. https://doi.org/10.1561/3100000012
- Nellikkath, R., Chatzivasileiadis, S., 2022. Physics-Informed Neural Networks for AC Optimal Power Flow. Electr. Power Syst. Res. 212, 108412. https://doi.org/10.1016/j.epsr.2022.108412

- Pourahmadi, F., Kazempour, J., 2025. Unit Commitment Predictor With a Performance Guarantee: A Support Vector Machine Classifier. IEEE Trans. Power Syst. 40, 715–727. https://doi.org/10.1109/TPWRS.2024.3400405
- Raissi, M., Perdikaris, P., Karniadakis, G.E., 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J. Comput. Phys. 378, 686–707. https://doi.org/10.1016/j.jcp.2018.10.045
- Surana, K., Jordaan, S.M., 2019. The climate mitigation opportunity behind global power transmission and distribution. Nat. Clim. Change 9, 660–665. https://doi.org/10.1038/s41558-019-0544-3
- Wang, H., Murillo-Sanchez, C.E., Zimmerman, R.D., Thomas, R.J., 2007. On Computational Issues of Market-Based Optimal Power Flow. IEEE Trans. Power Syst. 22, 1185–1193. https://doi.org/10.1109/TPWRS.2007.901301
- Xu, K., Dai, X., Qiu, L., 2025. OPFormer: Real-Time Optimal Power Flow with CNN-Based Transformer, in: ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Presented at the ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1–5.

  https://doi.org/10.1109/ICASSP49660.2025.10888727