A Learning-boosted Quasi-Newton Method for AC Optimal Power Flow

Dr. Kyri Baker, www.kyrib.com

Assistant Professor

University of Colorado Boulder

What is “AC Optimal Power Flow (OPF)”?

Power flowing throughout the electric power grid flows according to the **AC power flow equations,** a set of nonconvex expressions of power and complex voltage.

Grid operators need to solve a series of complex equations to determine how much electricity various power plants should produce to satisfy the demand throughout the grid.

Ideally, they’d like to solve the following, true, AC OPF problem:

Minimize generation cost across all generators

$$\min_{\mathbf{v},\mathbf{p}_g} \sum_{j \in \mathcal{G}} a_j p_{g,j}^2 + b_j p_{g,j} + c_j$$

s.t :

$$\left.\begin{aligned} |v_i| \sum_{m \in \mathcal{N}} |v_m| (G_{im} \cos(\theta_{im}) + B_{im} \sin(\theta_{im})) &= p_{l,i} - \sum_{k \in \mathcal{G}_i} p_{g,k}, \quad \forall i \in \mathcal{N} \\ |v_i| \sum_{m \in \mathcal{N}} |v_m| (G_{im} \sin(\theta_{im}) - B_{im} \cos(\theta_{im})) &= q_{l,i} - \sum_{k \in \mathcal{G}_i} q_{g,k}, \quad \forall i \in \mathcal{N} \end{aligned}\right\}$$

AC power flow equations (thetas, v’s, p_g and q_g’s are variables)

$$\left.\begin{aligned} \underline{p}_{g,j} \leq p_{g,j} \leq \overline{p}_{g,j}, \quad \forall j \in \mathcal{G} \\ \underline{q}_{g,j} \leq q_{g,j} \leq \overline{q}_{g,j}, \quad \forall j \in \mathcal{G} \end{aligned}\right\}$$

Active/reactive power output limits of generators *G*

$$\underline{|v|} \leq |v_i| \leq \overline{|v|}, \quad \forall i \in \mathcal{N}.$$

This optimization problem needs to be solved faster.

Unfortunately, the AC OPF problem can be huge (hundreds of thousands of variables and constraints), and is nonconvex. Thus, grid operators **linearize or convexify** the problem. Why is this problematic?

1. **Billions** of dollars annually are lost due to these suboptimalities [1].
2. **~500 million metric tons of carbon dioxide** can be cut by improving global grid efficiencies [2]
3. Many linear approximations will **not satisfy** physical grid constraints [3].

We’d like to solve it faster so grid operators can use the actual AC OPF problem to operate grids in real time (5-minutes or less).

Learning optimal solutions.

Newton-Raphson (aka “Newton’s Method”) is the most popular and widely used method to solve the OPF problem. It is computationally expensive due to many iterations involving forming large matrices and inverting them.

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha J^{-1}(\mathbf{x}^k) d(\mathbf{x}^k)$$

Key Idea: Let’s train a neural network to *learn* what the next Newton iteration should be, given the current guess, and avoid large matrix inversions (or factorizations): $\mathbf{x}^{k+1} = F_R(\mathbf{x}^k)$

A RNN.

Pros/Cons of using learning to optimize?

Pros:

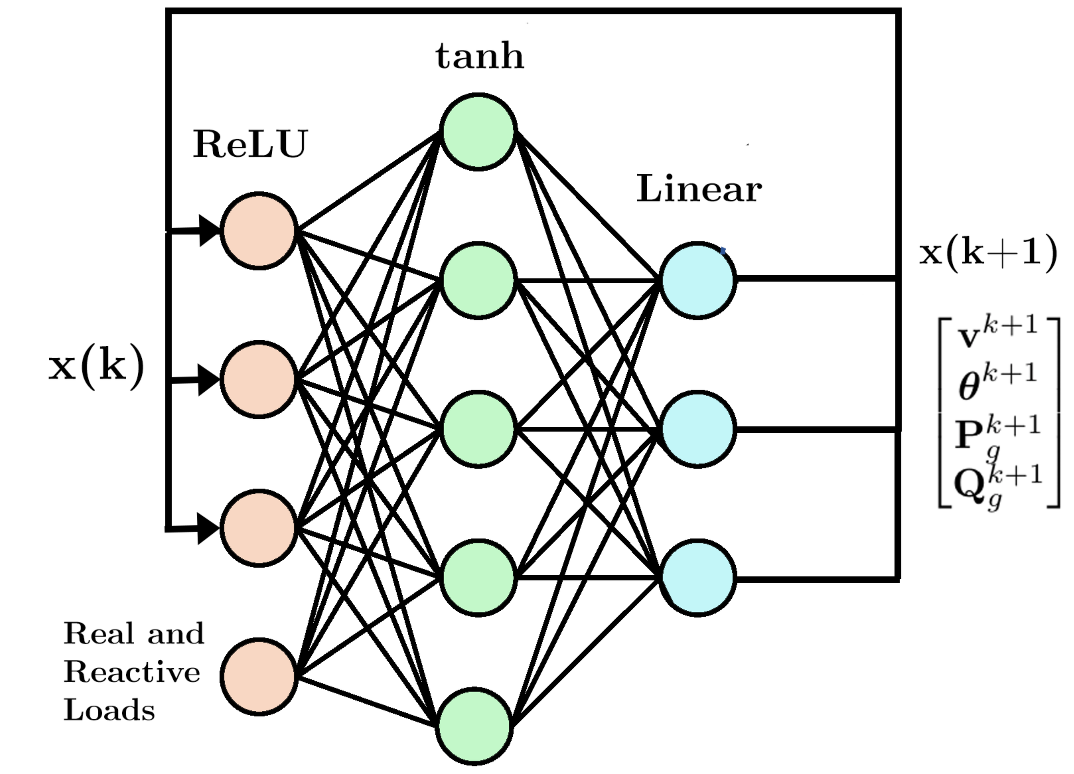
You can design the recurrent neural network to **guarantee convergence** [4]

Training data is easily generated offline or can be used from actual grid data

Inference is extremely fast compared to Newton-Raphson iterations

Cons:

Generally, feasibility of the original AC OPF is **not guaranteed** (but then again, the method grids are using right now isn’t either! [3]).



Recurrent Neural Network emulating Newton-Raphson iterations.

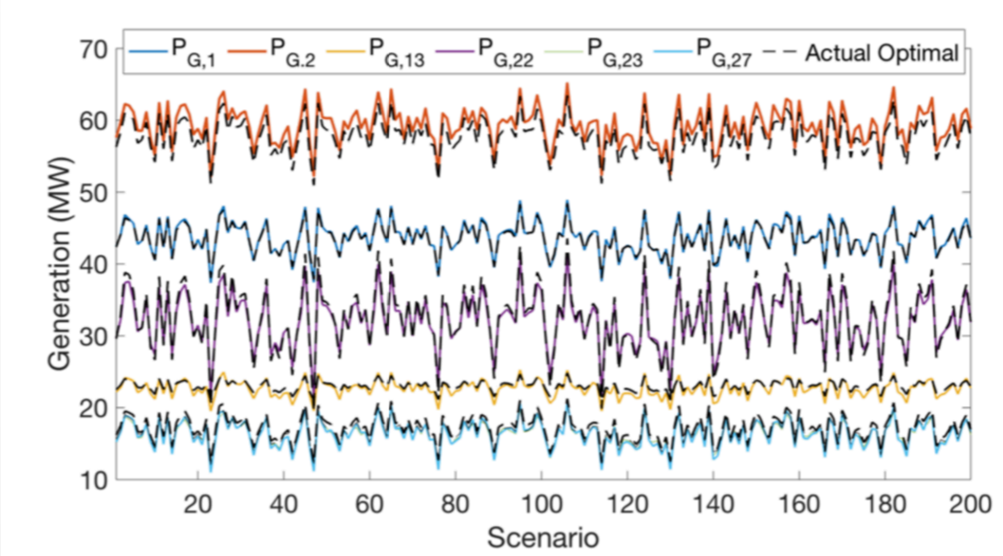
How well did it do?

Case	Mean Time (s)	Max Time (s)	Variance in Time (s)	Mean Speedup
30-bus MIPS	0.04	0.41	4.52e-04	
30-bus NN	0.06	0.42	3.53e-04	**-0.66x**
300-bus MIPS	1.09	10.87	2.09	
300-bus NN	0.03	0.42	0.001	**36.3x**
500-bus MIPS	1.46	2.96	0.49	
500-bus NN	0.08	0.45	3.25e-4	**18.3x**
1,354-bus MIPS	7.64	19.89	15.55	
1,354-bus NN	0.34	0.69	0.0016	**22.5x**

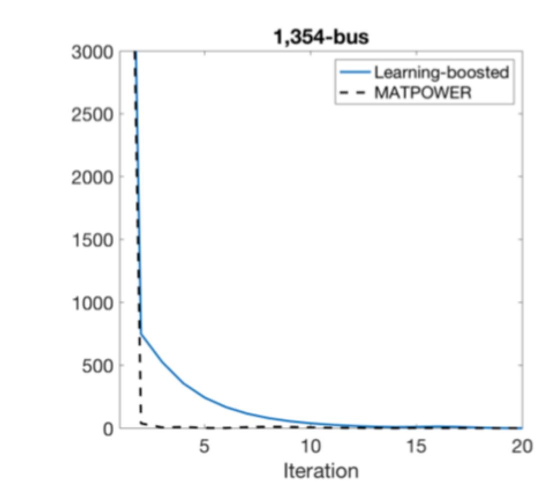
Time to convergence for various AC OPF scenarios (1,000 per network).

Case	MAE: Voltage Magnitude (pu)	MAE: Active Power (MW)	MAPE: Cost (%)
30-bus	0.004 pu	0.64 MW	0.29%
300-bus	0.009 pu	10.47 MW	0.65%
500-bus	0.099 pu	0.62 MW	0.66%
1,354-bus	0.019 pu	7.55 MW	1.16%

Error in predicting AC OPF solutions (1,000 per network).



Predicted generation optimums (colors) and actual optimums for the most varying generator in the 30-bus network



Convergence takes longer with the approximate learning-based method, but each iteration is significantly faster than commercial toolbox (MATPOWER)

References

[1] M. Cain, R. P. O’Neill, and A. Castillo, “History of optimal power flow and formulations,” *FERC Technical Report*, Aug. 2013

[2] S. Jordaan and K. Surana, https://theconversation.com/we-calculated-emissions-due-to-electricity-loss-on-the-power-grid-globally-its-a-lot-128296, 2019

[3] K. Baker, “Solutions to DC OPF are never AC feasible,” https://arxiv.org/pdf/1912.00319.pdf 2020.

[4] J. E. Steck, “Convergence of recurrent networks as contraction map- pings,” in *Intl. Joint Conference on Neural Networks (IJCNN)*, Jun. 1992.