Topology-Aware Reinforcement Learning for Tertiary Voltage Control

– Supplementary Material –

Balthazar Donon, François Cubelier
Efthymios Karangelos, Louis Wehenkel
Department of EE & CS - Montefiore Institute
Université de Liège, Belgium
{balthazar.donon, f.cubelier}@uliege.be
{e.karangelos, l.wehenkel}@uliege.be

Laure Crochepierre, Camille Pache
Lucas Saludjian, Patrick Panciatici
Research & Development Department
RTE (Réseau de Transport d'Électricité), France
{laure.crochepierre, camille.pache}@rte-france.com
{lucas.saludjian, patrick.panciatici}@rte-france.com

The present document is a supplementary material to our paper *Topology-Aware Reinforcement Learning for Tertiary Voltage Control* submitted at the PSCC 2024 conference. It provides information to those who would like to replicate our work, as well as extensive experimental results.

I. H2MGNODE Equations Detail

In the main paper, we state that ν should be bounded. By choosing a hyperbolic tangent mapping, we enforce $\frac{dh_a}{dt}$ to remain between -1 and 1. The differential system being integrated between $\tau = 0$ and 1, this enforces final latent variables $(h_a(1))_{a \in A_x}$ to remain between -1 and 1. As will be evoked later in this supplementary material, neural networks are very sensitive to the order of magnitude of their inputs. Enforcing latent variables to remain in this range of values improves the stability and speed of the training process.

II. Data Generating Process

This Section details the process of generating the datasets used in the experiments.

At first, three distinct variants (namely *Standard*[1], *Condenser*[2] and *Reduced*[3]) were crafted from the *case60nordic* test case [1] as detailed in Figure 1. We believe these variants to be significantly different: the *Condenser* variant has an additional synchronous condenser, while the *Reduced* variant has lost its north-eastern part.

A full dataset is then generated from each of the three operating conditions, based on the following process, also illustrated in Figure 2.

1) *Sample topology:* Either 0, 1, 2, 3 or 4 transmission lines are disconnected, with equal probability.
2) *Sample total load:* Total active consumption P_{tot}^{new} is sampled uniformly between 50% and 120% of the initial total load P_{tot}^{old}.

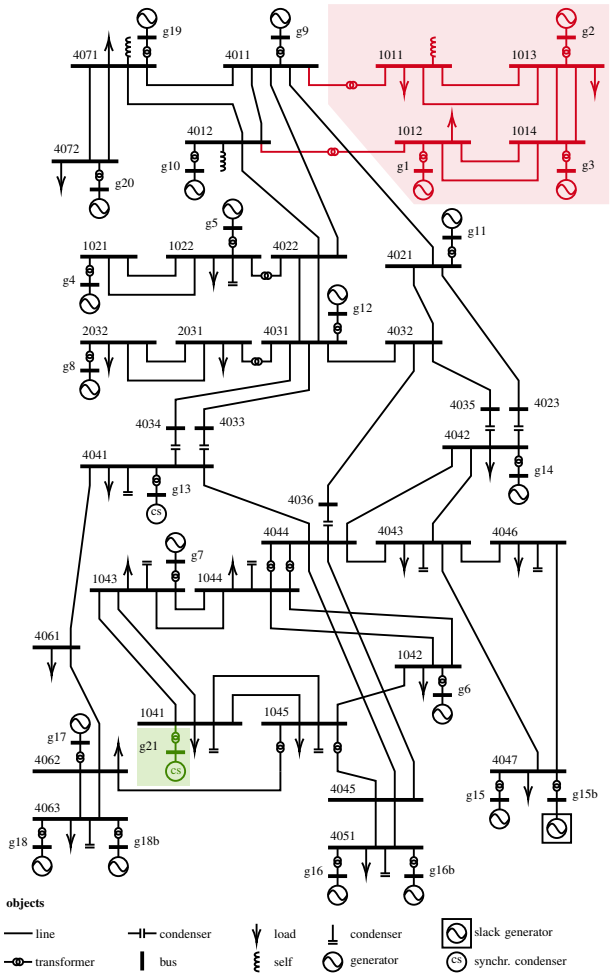


Fig. 1. Case60nordic test case. The *Condenser* dataset includes the whole system, the *Standard* dataset does not have the synchronous condenser g21 (green), and the *Reduced* dataset has neither the synchronous condenser g21 (green) nor the north-east region (red).

3) *Sample individual active loads:* Individual active loads are given by $\forall n \in \mathcal{E}^{load}, P_n^{new} = (\xi_n - \frac{\sum \xi_{n'}}{|\mathcal{E}^{load}|} + \frac{P_n^{old}}{P_{tot}^{old}}) \times$

[1] https://zenodo.org/doi/10.5281/zenodo.8367764
[2] https://zenodo.org/doi/10.5281/zenodo.8367613
[3] https://zenodo.org/doi/10.5281/zenodo.8367756

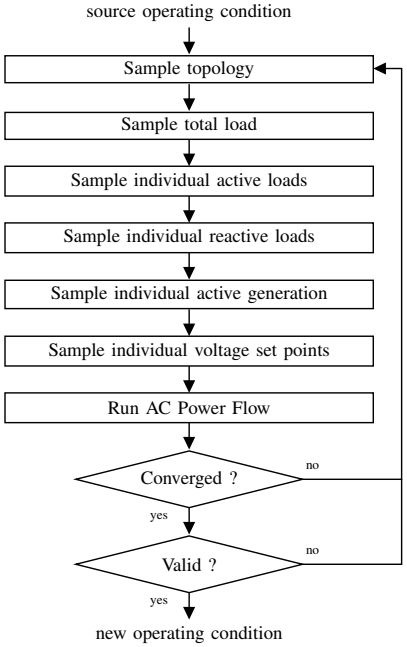


Fig. 2. Data generation pipeline.

P_{tot}^{new}, where $\xi_n \sim \mathcal{N}(0, 0.05)$.

4) *Sample individual reactive loads:* Individual reactive loads follow the heuristics described in [2]. Power factors are sampled uniformly between 0.8 and 1 and there is a 10% chance that the reactive load has negative sign.
5) *Sample individual active generation:* Each generator is associated with a cost uniformly sampled between 1.0 and 2.0. A DC-OPF is solved, where flows are constrained to remain below 80% of thermal limits, and all minimal power generation are set to $0MW$. After this DC-OPF, all generators whose power is below their initial minimal power (usually around $10MW$) are disconnected. Another DC-OPF is then run, enforcing flows to remain below 80% of thermal limits, and enforcing active generation to respect actual minimal power values.
6) *Sample individual generation voltage set points:* Generator voltage set points are all initially set to $1p.u.$. In addition, all capacitors are disconnected, while all inductors are connected.
7) *Run AC Power Flow:* Operating conditions that do not converge are rejected.

After the power flow, a series of checks is run over the resulting operating conditions, and improper ones are rejected.

- *Extreme voltage:* All voltages should remain between $0.85p.u.$ and $1.15p.u.$.
- *Overflows:* All flows should remain below thermal limits.
- *Disconnected components:* Only one connected component is allowed. Branch disconnections should not split the grid in multiple parts.
- *Negative loads:* Only positive loads are allowed.

Each rejection triggers a new snapshot generation, until the required amount of valid snapshots is reached.

For each of the three initial snapshots, exactly $100,000$ samples are generated for the train set, $2,000$ for the validation set and $10,000$ for the test set.

Table I displays the number of loops required to generate each of the three datasets, while underlining the number of rejections caused either by a DCOPF non-convergence, an AC power flow non-convergence or a check failure. For both the *Standard* and *Condenser* datasets, approximately 25% of initiated generations result in a viable operating condition. For the *Reduced* dataset, we observe a much larger check failure rate (mostly due to overflows and overvoltages), which leads to a 16% rate of valid operating conditions. Notice that the different causes of check failure do not add up to the total amount of check failures, because it is common for a single snapshot to display multiple causes for rejection at the same time. Notice that the table mentions $120,000$ operating conditions: $10,000$ operating conditions were generated for the validation sets, but only $2,000$ of them were actually retained so as not to slow the training process too much.

TABLE I
SUMMARY OF THE DATASET GENERATION PROCESS.

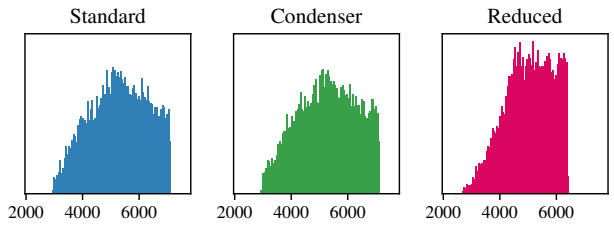
	Standard	Condenser	Reduced
Loop count	473,137 (100%)	458,724 (100%)	725,618 (100%)
DCOPF Div.	11,751 (2,5%)	11,376 (2,5%)	16,216 (2,2%)
PF Div.	100,581 (21%)	95,728 (21%)	104,630 (14%)
Check Failed	240,805 (51%)	231,620 (51%)	484,772 (67%)
$V > 1.15$ p.u.	113,429 (24%)	108,799 (24%)	298,845 (41%)
$V < 0.85$ p.u.	53,232 (11%)	47,860 (10%)	62,042 (8,6%)
Overflow	181,181 (38%)	172,768 (38%)	410,802 (57%)
Disconnected	7,971 (1,7%)	7,765 (1,7%)	14,598 (2,0%)
Loads < 0	1,763 (0,37%)	1,719 (0,37%)	3,798 (0,52%)
Valid samples	120,000 (25%)	120,000 (26%)	120,000 (16%)

A fourth dataset called *All* is made up of the union of the three previous datasets.

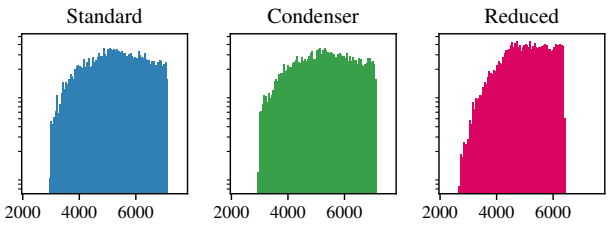
III. Datasets Analysis

In the context of this study, it is essential to consider datasets of operating conditions which are *challenging* enough in terms of tertiary voltage control. In order to assess the relevance of our three datasets, Figure 3 provides a statistics summary for each test set., which we describe as follows.

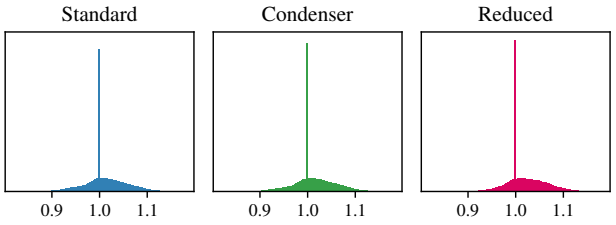
- *Total active load (a – b):* In the three datasets, the sum of individual loads vary between 3 GW and 7 GW. Contrarily to the initial sampling distribution, these are not uniform because low values are more likely to be rejected by our data generation process.
- *Bus voltages (c – d):* These histograms display statistics over all buses of all operating conditions in each test set. The peak that occurs at exactly 1 p.u. is caused by generators voltage setpoints which are all initialized at this value. We observe a tendency towards overvoltages (*i.e.* larger than 1.1 p.u.), although the logarithmic scale unveils the occurrence of undervoltage events (*i.e.* smaller than 0.9 p.u.). There is no value outside the $[0.85, 1.15]$ range because of the previously described filtering.



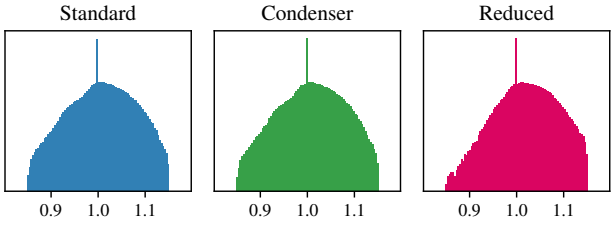
(a) Total active load (MW) per operating condition – Linear scale



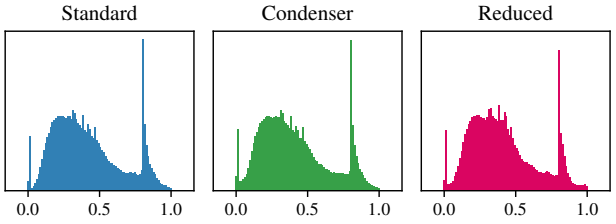
(b) Total active load (MW) per operating condition – Logarithmic scale



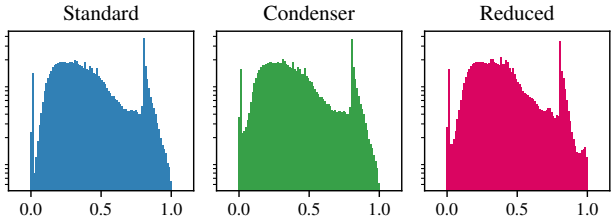
(c) Bus voltages (p.u.) – Linear scale



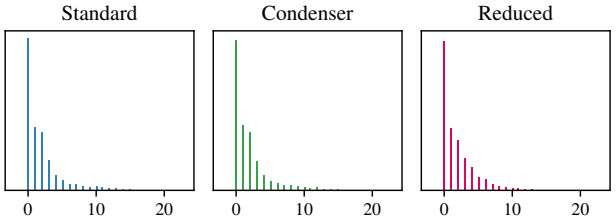
(d) Bus voltages (p.u.) – Logarithmic scale



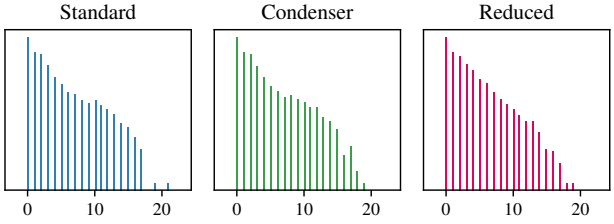
(e) Branch usage ratio – Linear scale



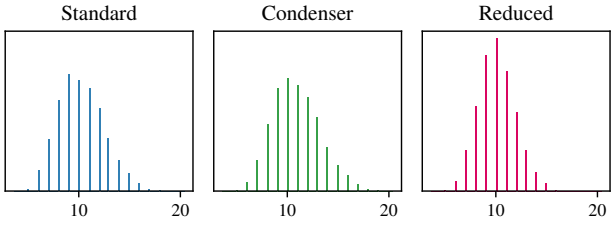
(f) Branch usage ratio – Logarithmic scale



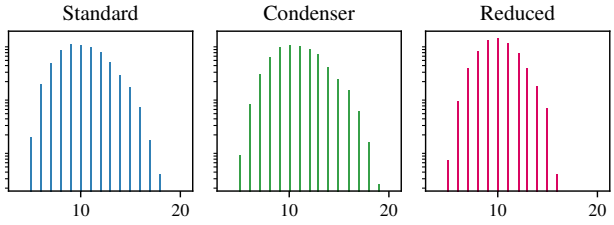
(g) Voltage violations per operating condition – Linear scale



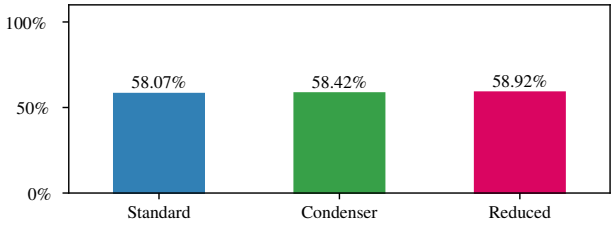
(h) Voltage violations per operating condition – Logarithmic scale



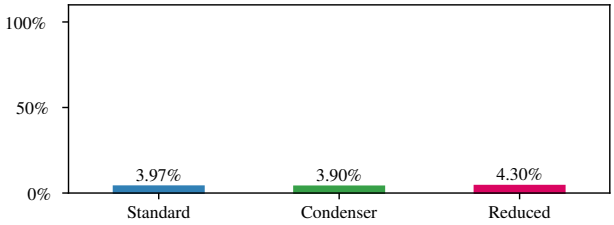
(i) Connected generators per operating condition – Linear scale



(j) Connected generators per operating condition – Logarithmic scale

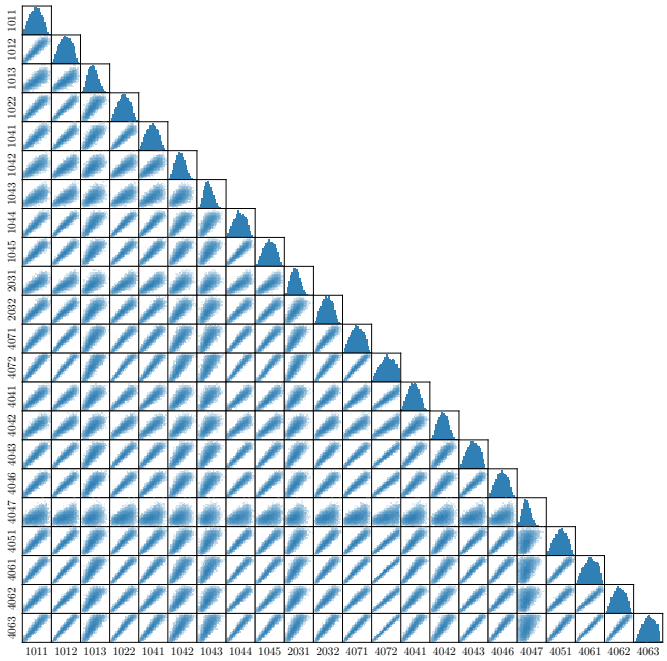


(k) Operating conditions w/ violation

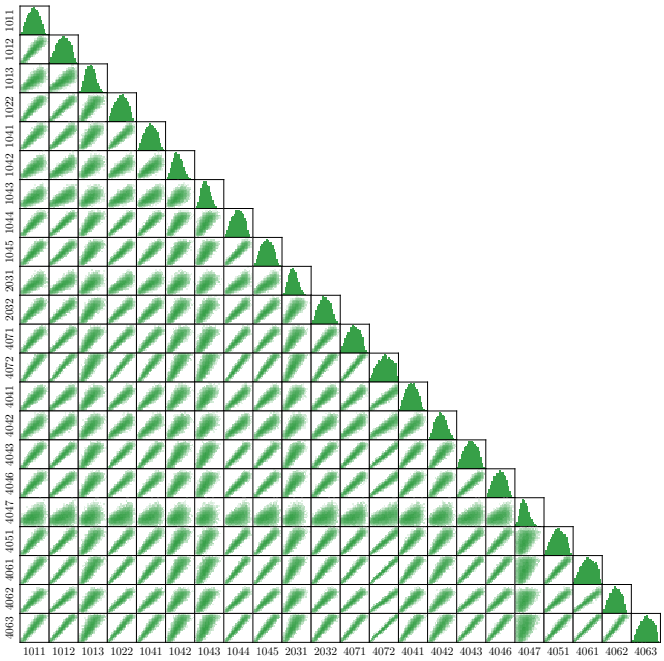


(l) Buses w/ violation

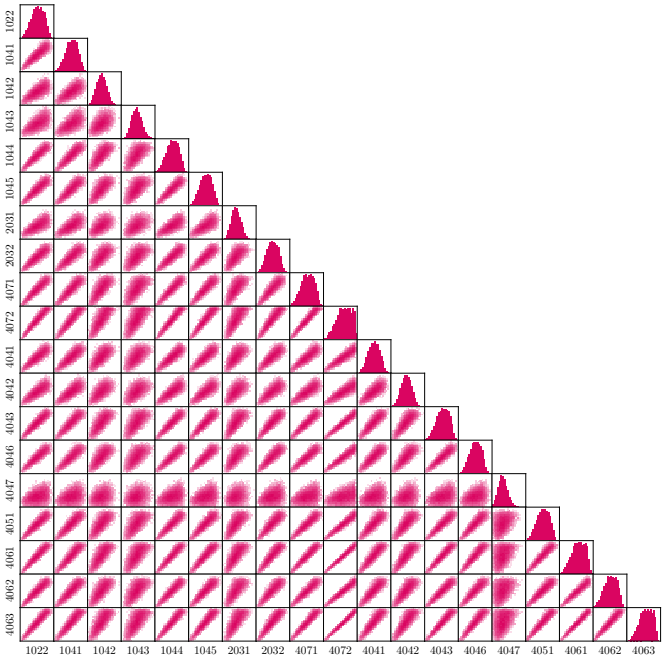
Fig. 3. Statistics summary for the three datasets, with the default 1.0 p.u. voltage setpoint.



(a) Standard test set.



(b) Condenser test set.



(c) Reduced test set.

Fig. 4. Two-dimensional histograms for each pair of loads throughout each test set. Colors intensity scale logarithmically with the frequency of occurrence for the sake of readability.

- *Branch currents (e – f):* These histograms display statistics over all branches of all operating conditions in each test set. The peak that can be observed at 80% of thermal limits (*i.e.* 0.8 on the histogram) is due to the DCOPF step, which tried to keep current below this threshold, as previously described. The actual currents that are then computed using an AC power flow end up going beyond this value, although a filter has been applied to keep all currents below thermal limits (*i.e.* 1.0 on the histogram).
- *Voltage violations per operating condition (g – h):* While most operating conditions have less than 5 voltage violations, some end up with more than 20.
- *Connected generators per operating conditions (i – j):* Generators that are not used after the DCOPF step (with

random costs) are disconnected from the grid. Although around 20 generators are available, it appears that most operating conditions only make use of less than 15 of them. Notice that operating conditions with fewer generators also have fewer levers for tertiary voltage control.
- *Operating conditions with violation (k):* Regardless of the considered dataset, around 58% of operating conditions display a voltage violation when generators voltage setpoints are set to 1.0 p.u.
- *Buses with violation (l):* The number of buses – over all operating conditions – that are outside of their permanent limits is around 4%. Notice that these are responsible for the 58% of operating conditions with a violation, which means that when a voltage violation occurs in an operating conditions, only a few buses have an issue.

Figure 4 displays 2-dimensional histograms for all pairs of loads. Although loads are roughly correlated, they do not scale homothetically with one another.

IV. Optimization Baseline

In the main paper, we compare our proposed approach to an ACOPF solver based on an interior-point method. The ACOPF problem formulation has been omitted for the sake of readability. In the present section, we describe the different sets, parameters and variables, and frame the tertiary voltage control as an ACOPF problem through equations (1 – 14).

Sets

$\mathcal{E}^{bus}$	Buses.
$\mathcal{E}^{gen}$	Generators.
$\mathcal{E}^{branch}$	Transmission Branches (lines & transformers).

Subscripts define relevant subsets (*e.g.* $\mathcal{E}^{gen}_n \subseteq \mathcal{E}^{gen}$ is the subset of generators connected at bus $n \in \mathcal{E}^{bus}$). n^* denotes the slack bus.

Generation Parameters, $\forall i \in \mathcal{E}^{gen}$

P^g_i	Active power generation.
$\overline{Q}_i, \underline{Q}_i$	Generator reactive power bounds.

Bus Parameters, $\forall n \in \mathcal{E}^{bus}$

P^d_n, Q^d_n	Nodal active and reactive power demand.
g^s_n	Nodal shunt conductance.
b^s_n	Nodal shunt susceptance.
$\overline{V}_n, \underline{V}_n$	Nodal voltage bounds.

Branch Parameters, $\forall \ell \in \mathcal{E}^{branch}$

g_ℓ	Branch series conductance.
b_ℓ	Branch series susceptance.
t_ℓ	Branch tap ratio.
ϕ_ℓ	Branch phase shift angle.
$\overline{I}_\ell$	Branch current rating.

Variables

Q_i	Reactive power generation.
dP	Slack bus active power generation adjustment.
V_n	Nodal voltage magnitude.
θ_n	Nodal voltage angle.
$I_{\ell,n\to m}$	Branch current magnitude from bus n to m.
$P_{\ell,n\to m}$	Branch active power flow from bus n to m.
$Q_{\ell,n\to m}$	Branch reactive power flow from bus n to m.

$$\min \Bigg[\frac{\sum_{\ell\in\mathcal{E}^{branch}} \left(P_{\ell,n\to m} + P_{\ell,n\leftarrow m}\right)}{\sum_{n\in\mathcal{E}^{bus}} P^d_n} + \frac{\lambda}{|\mathcal{E}^{bus}|} \sum_{n\in\mathcal{E}^{bus}} \left(v^2_{n,+} + v^2_{n,-}\right) + \frac{\lambda}{|\mathcal{E}^{branch}|} \sum_{\ell\in\mathcal{E}^{branch}} \left(\frac{i^2_{\ell,f} + i^2_{\ell,t}}{2}\right) \Bigg] \tag{1}$$

subject to:

$$\theta_{n^*} = 0 \tag{2}$$

$$\sum_{\ell\in\mathcal{E}^{branch}_{n^*}} P_{\ell,n^*\to m} = dP + \sum_{i\in\mathcal{E}^{gen}_{n^*}} P^g_i - P^d_{n^*} - g^s_{n^*}\cdot V^2_{n^*} \tag{3}$$

$$\forall n \neq n^* \quad \sum_{\ell\in\mathcal{E}^{branch}_{n}} P_{\ell,n\to m} = \sum_{i\in\mathcal{E}^{gen}_{n}} P^g_i - P^d_n - g^s_n\cdot V^2_n \tag{4}$$

$$\forall n, \quad \sum_{\ell\in\mathcal{E}^{branch}_{n}} Q_{\ell,n\to m} = \sum_{i\in\mathcal{E}^{gen}_{n}} Q^g_i - Q^d_n - b^s_n\cdot V^2_n \tag{5}$$

$$\forall i, \quad Q^{\min}_i \leq q_i \leq Q^{\max}_i \tag{6}$$

$$\forall \ell, \quad P_{\ell,n\to m} = (g_\ell + g^s_n)\left(V_n/t_\ell\right)^2 - \frac{V_n V_m}{t_\ell}\left[g_\ell cos\left(\theta_{nm} - \phi_\ell\right) - b_\ell sin\left(\theta_{nm} - \phi_\ell\right)\right] \tag{7}$$

$$\forall \ell, \quad P_{\ell,n\leftarrow m} = (g_\ell + g^s_m)\, V^2_m - \frac{V_n V_m}{t_\ell}\left[g_\ell cos\left(\theta_{mn} + \phi_\ell\right) - b_\ell sin\left(\theta_{mn} + \phi_\ell\right)\right] \tag{8}$$

$$\forall \ell, \quad Q_{\ell,n\to m} = -(b_\ell + b^s_n)\left(V_n/t_\ell\right)^2 + \frac{V_n V_m}{t_\ell}\left[g_\ell sin\left(\theta_{nm} - \phi_\ell\right) + b_\ell cos\left(\theta_{nm} - \phi_\ell\right)\right] \tag{9}$$

$$\forall \ell, \quad Q_{\ell,n\leftarrow m} = -(b_\ell + b^s_m)\, V^2_m + \frac{V_n V_m}{t_\ell}\left[g_\ell sin\left(\theta_{mn} + \phi_\ell\right) + b_\ell cos\left(\theta_{mn} + \phi_\ell\right)\right] \tag{10}$$

$$\forall n, \quad v_{n,+} \geq (V_n - \underline{V}_n)/(\overline{V}_n - \underline{V}_n) - 1 + \epsilon \tag{11}$$

$$v_{n,-} \geq \epsilon - (V_n - \underline{V}_n)/(\overline{V}_n - \underline{V}_n) \tag{12}$$

$$\forall \ell, \quad i_{\ell,f} \geq \frac{\left(P_{\ell,n\to m}\right)^2 + \left(Q_{\ell,n\to m}\right)^2}{V^2_n\cdot\left(\overline{I}_\ell\right)^2} - 1 + \epsilon \tag{13}$$

$$i_{\ell,t} \geq \frac{\left(P_{\ell,n\leftarrow m}\right)^2 + \left(Q_{\ell,n\leftarrow m}\right)^2}{V^2_m\cdot\left(\overline{I}_\ell\right)^2} - 1 + \epsilon \tag{14}$$

Notice that there is a discrepancy between this ACOPF formulation and the one used in the main paper in the way currents are penalized. In the present formulation, we consider current magnitudes at both ends of each branch, penalize them equally, and then take their mean (see equation 1). In the main

paper, we only penalize the end of each branch that is the most loaded (*w.r.t.* its thermal limit). While we acknowledge the discrepancy, we believe that it does not impact experimental results in a meaningful way.

V. Data Normalization

Neural networks being extremely sensitive to the scale of input data [3], it is common to preprocess their inputs using a bijective *normalization* function η as follows:

$$x' = \eta(x), \tag{15}$$

where x is a sample and x' serves as input to the neural network. The purpose of η is to make inputs of neural networks follow a somewhat uniform distribution between -1 and 1.

The normalization mapping η is usually fine-tuned to fit the train set $\mathcal{D}_{train}$ before the training process starts. Once fixed, η should be considered as an integral part of the neural network, even though it is not updated during training.

In our case, we need η to be compatible with operating conditions with various topologies (including the number of assets and their interconnection pattern). We propose to consider a different normalization function per hyper-edge class:

$$\eta = (\eta^c)_{c\in\mathcal{C}}, \tag{16}$$

which can be applied to an operating condition x as follows:

$$x' = (\eta^c(x_e^c))_{c\in\mathcal{C}, e\in\mathcal{E}}. \tag{17}$$

As commonly done in the deep learning literature, each η^c is decomposed into a series of bijective functions from $\mathbb{R}$ to $\mathbb{R}$, one per component of x_e^c. For the sake of readability and without loss of generality, let us assume that all classes of hyper-edge bear one-dimensional features.

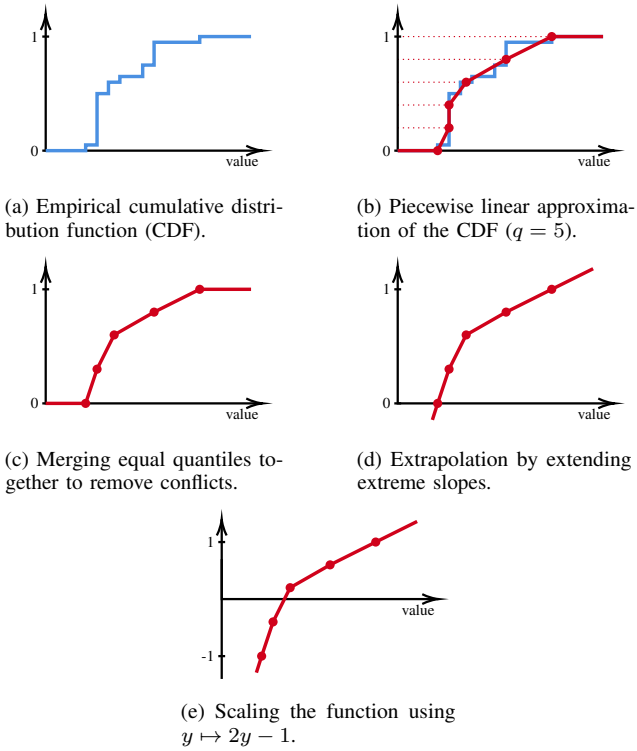
Let us consider a class $c \in \mathcal{C}$ and denote by $\mathcal{D}^c_{train}$ the set of values taken by all hyper-edges of class c in $\mathcal{D}_{train}$,

$$\mathcal{D}^c_{train} = \{x_e^c | x \in \mathcal{D}_{train}, e \in \mathcal{E}_x^c\}. \tag{18}$$

Our goal is to find a bijective function η^c such that $\eta^c(\mathcal{D}^c_{train})$ is roughly uniformly distributed between -1 and 1. A common approach would be to use $x \mapsto (x - \underline{x})/(\overline{x} - \underline{x})$ where $\overline{x}$ and $\underline{x}$ are respectively the maximal and minimal values in $\mathcal{D}^c_{train}$. However, this amounts to assuming that feature values already follow a nearly uniform distribution. Such an assumption does not hold in our power systems application. To give a concrete example, the distribution of line resistances has multiple modes, one per voltage level.

This shortcoming pushed us to develop a slightly different approach, based on a piecewise-linear approximation of the Cumulative Distribution Function (CDF). As illustrated in Figure 5, our approach is made of the following steps:

(a) First, let us gather all values x_e^c of all hyper-edges of class $c \in \mathcal{C}$ encountered in $\mathcal{D}^c_{train}$. The empirical CDF of $\mathcal{D}^c_{train}$ provides a function that maps $\mathcal{D}^c_{train}$ to a uniformly distributed subset of $[0, 1]$. However, it is not bijective, and requires to store a whole dataset.



(a) Empirical cumulative distribution function (CDF).

(b) Piecewise linear approximation of the CDF ($q = 5$).

(c) Merging equal quantiles together to remove conflicts.

(d) Extrapolation by extending extreme slopes.

(e) Scaling the function using $y \mapsto 2y - 1$.

Fig. 5. Creation of a normalization function (in red) from an empirical cumulative distribution function (in blue).

(b) Instead, we build a piecewise linear approximation of the CDF by storing the q-quantiles, where $q \in \mathbb{N}$ is a hyper-parameter of the process.

(c) It is likely that multiple quantiles end up being equal, which prevents the piecewise linear interpolation from being well-defined. We thus propose to merge conflicting quantiles together by taking their mean.

(d) In order to make a bijective mapping from $\mathbb{R}$ to $\mathbb{R}$ out of our piecewise linear function, we propose to extend its extreme slopes to $\pm\infty$.

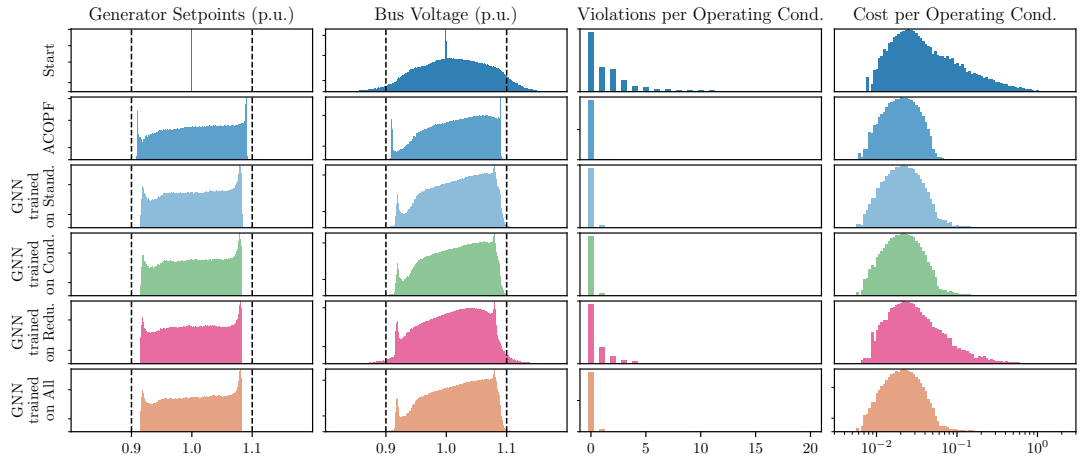
(e) Finally, we apply $y \mapsto 2y - 1$ to its output to make $\eta^c(\mathcal{D}^c_{train})$ evenly spread between -1 and 1.

The resulting function η^c is a bijective piecewise linear function that requires to store at most q quantiles from the train set. In our experiments, we used $q = 200$.

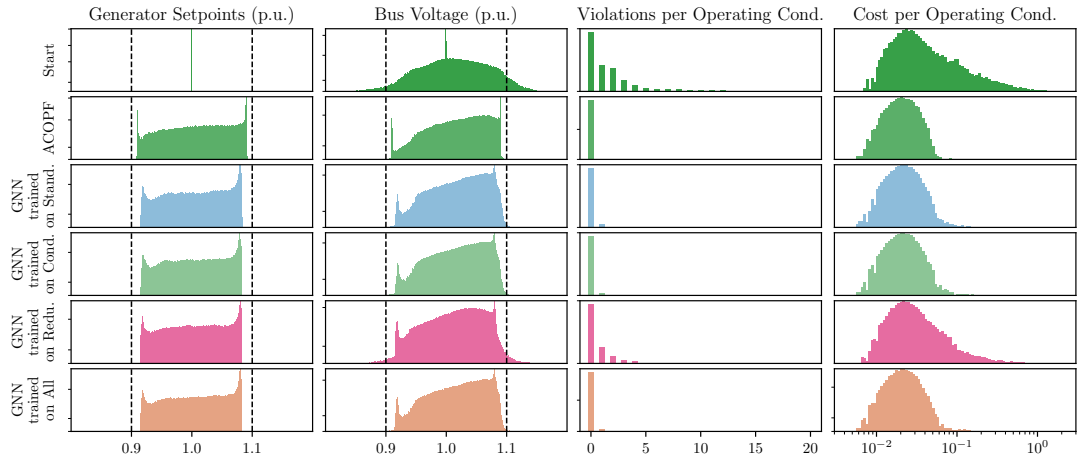
VI. Results Analysis

This Section provides additional experimental results. Figure 6 provides extensive histograms for each of the three test sets, when one uses:

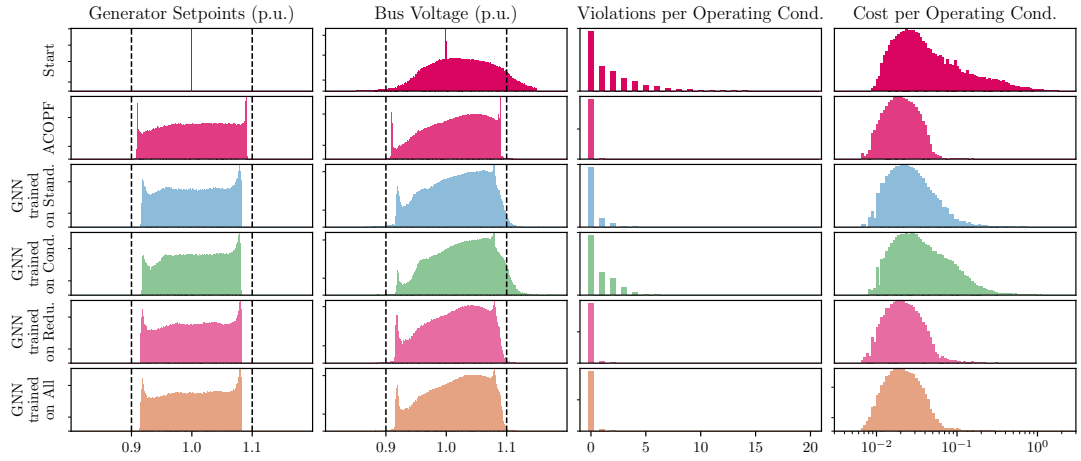
- the initial 1.0 p.u. value to all generators (denoted by *Start*);
- the ACOPF solver baseline (denoted by *ACOPF*);
- the GNN policy trained on the *Standard* train set (denoted by *GNN trained on Stand.*);
- the GNN policy trained on the *Condenser* train set (denoted by *GNN trained on Cond.*);
- the GNN policy trained on the *Reduced* train set (denoted by *GNN trained on Redu.*);



(a) *Standard* test set.



(b) *Condenser* test set.



(c) *Reduced* test set.

Fig. 6. Comparison of the impact of the four different policies and of the ACOPF solver baseline over the three different test sets. The first column displays the histogram of all voltage setpoints (control variables) over all generators and all operating conditions. The second column shows the histogram of resulting bus voltage magnitudes over all buses and all operating conditions. The third column contains histograms of the violation count (current or voltage) per operating conditions. The fourth column presents histograms of costs for all operating conditions. The ordinate axis is represented in *symlog* scale (*i.e.* $y \mapsto \ln(y+1)$).

- the GNN policy trained on the union of all train sets (denoted by *GNN trained on All*).

All policies are compatible with all test sets, even if the number of generators, lines, loads, etc. vary from one dataset to the other.

Histograms of bus voltages show that most policies manage to keep voltages within reasonable bounds, at the exception of the one trained on the *Reduced* train set. This assertion is further supported by violation counts per operating condition, which show that remaining violations are significantly less severe than with default 1.0 p.u. voltage setpoints.

REFERENCES

[1] F. Capitanescu, "Suppressing ineffective control actions in optimal power flow problems," *IET Generation, Transmission & Distribution*, vol. 14, pp. 2520–2527, 2020.

[2] B. L. Thayer and T. J. Overbye, "Deep Reinforcement Learning for Electric Transmission Voltage Control," in *2020 IEEE Electric Power and Energy Conference (EPEC)*, pp. 1–8, 2020.

[3] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.