

# Synthetic Topology Generation for Generalizable Machine Learning in Low Voltage Distribution Grids

Markus de Koster  
Institute of Electrical Power Engineering  
Cologne Institute for Renewable Energy  
Technische Hochschule Köln, Cologne, Germany  
email: markus.de\_koster@th-koeln.de

Patrick Lehnen, Patrick Mack, Eberhard Waffenschmidt, Ingo Stadler  
Institute of Electrical Power Engineering  
Cologne Institute for Renewable Energy  
Technische Hochschule Köln, Cologne, Germany

## ABSTRACT

Generating synthetic power grids for training of machine learning models that can generalize to unseen topologies requires more than replication of existing real-world grids. Instead, representative structures that capture real-world properties while at the same time exhibiting high variance are crucial for effective model training. In this paper, a probabilistic method for generating synthetic low voltage distribution grids is proposed. The method employs Markov chains to mimic realistic structures and assigns attributes to nodes and edges based on configurable probability distributions. The resulting grids are validated using load flow calculations and through comparison of key graph-theoretic metrics.

## KEYWORDS

Distribution Grid, Synthetic Grid Generation, Machine Learning, Markov Chain, Quasi-Monte-Carlo

## INTRODUCTION

The recent shift towards distributed energy resources (DER) and increasing loads in low voltage (LV) distribution grids necessitate additional monitoring and control strategies to identify and mitigate potential issues.

Increasingly, algorithms based on machine learning (ML) are being developed to aid in this transformation process. ML models for power flow and power quality state estimation are used to estimate the system state, compensating the lack of real-time measurements in LV grids through their training on synthetic data [1–3]. Fault localization algorithms leverage ML techniques to find the most probable error sources based on a limited set of measurements [4]. However, those models are typically trained specifically for a single grid topology. Changes to the topology are likely to require training of a new model and unnoticed changes such as line outages may result in erroneous outputs of the models. Moreover, a model trained for a specific topology can't easily be reused for a different topology since the inputs – typically measured voltages or currents at specific nodes or lines – are entangled with their location. Likely, the model embeds the relationship between voltages at neighbouring nodes and hereby the line impedance in the weights of its neurons. Similarly, several works actively embed the power grid topology in the model architecture in physics-aware neural networks ([5, 6]) and graph neural networks ([7, 8]). For all those cases, the fixed representations would

require training of a new model if subjected to a different underlying topology – a process that is both costly and unsustainable.

In order to create a model that works for multiple, possibly changing topologies and to enable generalization capabilities where the model generalizes from topologies seen during training to unknown topologies, training data of diverse sets of power grid are required.

However, power grid data is not readily available due to privacy and security concerns. Additionally, LV grid topologies are often not well documented and rarely digitalized. Hence, numerous works have been conducted on generating synthetic power grids. Most of those works aim at replicating real world power grids as closely as possible. The authors of [9] build a reference network model for all voltage levels based on geographical location, contracted customer demand, substation and DER capacity and location. Other works use publicly available data such as street maps to estimate power line locations and cluster loads together, often choosing cluster centroids as substation locations [10–16].

Aside from methods based on geographical data, only few works exist. A random graph model is used in [17] to generate highly interconnected synthetic transmission grids. The authors of [18] aim at creating statistically correct random topologies through a small-world model ([19]), adding lattice connections and then assigning line impedances to synthesize a transmission grid. They highlight the difficulties of generating realistic power grid topologies through a small-world model since power grid node degree does not significantly change with grid size.

For different domains such as chemistry, medicine or computer vision various works on graph generation have been conducted that use ML models including graph recurrent neural networks, generative adversarial networks, variational autoencoders (VAE) or graph diffusion models [20–22]. However, those models have been designed for different domains that exhibit different properties than LV grids as will be explained in the upcoming section. Additionally, training ML models using the output of other ML models requires special considerations as the output of the first model is already less likely to reflect real world properties than measured or observed data.

In conclusion, none of the existing literature aims at synthesizing power grid topologies for the purpose of creating machine learning models with generalization capabilities to unseen topologies.

### **Generalization capabilities**

In general, machine learning models perform optimally when trained on datasets that encapsulate the full range of variability present in the target domain. In the context of power grid modeling, reliance on a limited set of grid configurations can lead to models that overfit to specific topologies, thereby reducing their generalization capabilities. This overfitting arises when models capture noise or specific patterns unique to the training data, resulting in poor performance on unseen configurations.

To mitigate overfitting and improve generalization, several regularization techniques are commonly employed. These include dropout, early stopping, and structural simplification of the model by reducing the number of neurons or layers or through pruning [23]. While these methods act directly on the model architecture, the focus of this work lies in generating diverse power grid topologies that facilitate robust training, rather than model optimization itself. For this purpose, domain randomization is utilized, wherein the diversity of the data distribution is enhanced by systematically varying input parameters [24].

Moreover, generating synthetic grid topologies with high variability is a complex task. The vast number of possible configurations and parameter combinations leads to a high-dimensional input space. As the dimensionality increases, the volume of the space expands exponentially, making it challenging to sample the space adequately and train models that

generalize well. For example, the amount of possible binary tree topologies with  $n$  internal nodes is given in equation (1) and for a grid with  $n=20$  amounts to 6564120420 possibilities.

$$T(n) = (2n)! / ((n+1)! * n!) \quad (1)$$

Under various constraints such as only considering non-isomorphic trees, this number can be reduced but remains too big to simulate all possible combinations.

Thus, the challenge is to find representative topologies that capture the essence of the graph structure of LV power grids.

### Graph metrics

Several metrics can be used to evaluate structural properties of graphs. First, the total number of edges or links  $L$  is described in equation (2) as the sum of degree  $k$  of node  $i$  divided by 2 across all nodes  $N$ . Here the node degree is the number of edges connected to that node.

$$L = \sum_{i=1}^N (k_i) / 2 \quad (2)$$

Second, the clustering coefficient for a single node  $C_i$  is defined in equation (3).

$$C_i = (2e_i) / (k_i(k_i - 1)) \quad (3)$$

The global clustering is the average of cluster coefficient across all nodes and defined in equation (4).

$$C = (1/N) * \sum_{i=1}^N (C_i) \quad (4)$$

Further, the graph diameter  $D$  is defined as the maximum shortest path length between any pair of nodes in the graph. In equation (5),  $d(u, v)$  is the shortest path distance between two nodes (or vertices  $V$ )  $u, v$ .

$$D = \max(d(u, v) \text{ for } u, v \in V) \quad (5)$$

Table 1 shows a comparison of the abovementioned graph metrics for exemplary power grids found in literature for different voltage levels compared to graphs of different domains such as protein interactions or graphs describing references between papers sourced from the OpenGraphBenchmark (OGB) [25]. LV grids are characterized by a relatively small number of nodes and edges, an average node degree of slightly below 2 and an average clustering coefficient of 0. This is reflective of radial feeder structures with no interconnecting (or lattice) lines. In this evaluation open switches are not included as edges. In comparison, high voltage (HV) and MV grids show a higher average node degree and hereby an increased clustering coefficient. Graph diameters are smaller relative to the number of nodes since more interconnections exist.

Table 1. Comparison of different graph domains and metrics.

Category	Name	Average #Nodes	Average #Edges	Average Node Deg.	Average Clust. Coeff.	Graph Diameter
LV power grids [26–28]	CIGRE LV	44	43	1.95	0.00	23
	Kerber Vorstadt	386	385	1.99	0.00	95
	Dickert LV	47	46	1.96	0.00	25
MV /HV power grids [29]	IEEE 24	24	34	2.83	0.03	7
	IEEE 33	33	32	1.94	0.00	20
	IEEE 57	57	78	2.74	0.12	12
	IEEE 89	89	206	4.63	0.19	9
	IEEE 118	118	179	3.03	0.17	14
	Illinois 200	200	245	2.45	0.04	18
	Trans. France	6470	8066	2.49	0.05	34
	Trans. Europe	9241	14207	3.07	0.10	91
OGB Node	products	2,449,029	61,859,140	50.5	411	27
	proteins	132,53	39,561,252	597.0	280	9
	papers100M	111,059,956	1,615,685,872	29.1	85	25
OGB Link	wikikg2	2,500,604	17,137,181	12.2	168	26
OGB Graph	molhiv	25.5	27.5	2.2	2	12.0

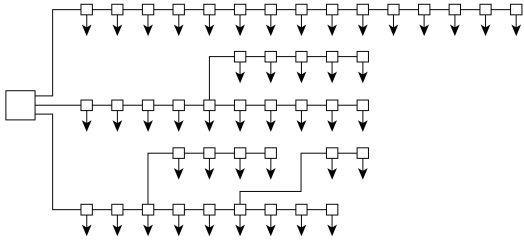


Figure 1: Dickert LV Grid showing the tree-like structure of distribution grids with no interconnections and a high graph diameter

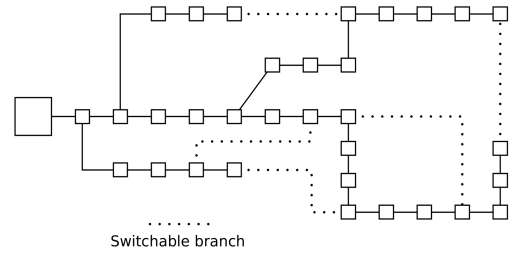


Figure 2: IEEE33 MV grid with open switches as lattice connections between feeders.

Figures 1 and 2 show exemplary one-line diagrams of an LV and an MV grid. While the LV grid has long feeders with no lattice connections, the MV grid has 5 switchable lines that interconnect feeders. In the LV grid loads are shown as small arrows with a load at every node representing residential households.

This paper aims at providing a proof of concept and guidance for filling the gap in synthetic grid generation capturing essential graph structural attributes for downstream machine learning applications.

## METHOD

We propose a method for generating synthetic power grid topologies utilizing a statistical top-down approach that incorporates Markov chains and quasi-Monte Carlo (QMC) sampling for parameter variation. Notably, all probabilities and parameters of the algorithm that will be described in the following sections can be adjusted and were chosen based on observed patterns and abovementioned statistical analysis.

### Topology generation

The topology generation begins at the medium voltage (MV) section, creating one or multiple feeders that are connected through transformers to the low voltage grid.

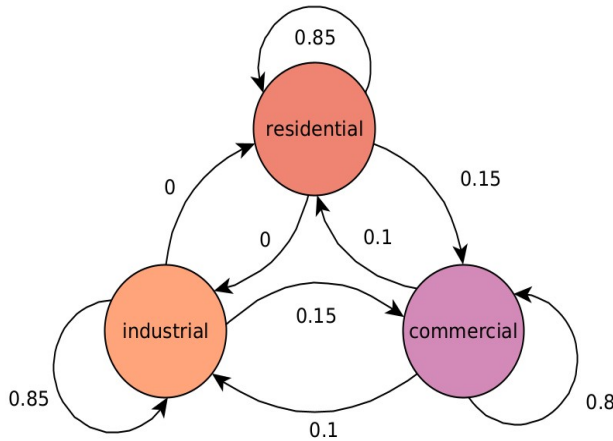


Figure 3: Markov chain for transitions between residential, commercial and industrial consumer structure within urban grids. Self-transition are favoured to reflect real-world distributions such as residential or commercial areas.

Each feeder is assigned an initial consumer structure from residential, commercial, or industrial categories, sampled according to user-defined feeder structure probabilities (e.g., in urban settings: 60% residential, 30% commercial, 10% industrial). This initial structure serves as the starting point for a Markov chain process, which governs transitions between consumer types at each subsequent service drop node.

The transition matrix for urban grids, for instance, strongly favors self-transitions such as residential-to-residential (0.85), commercial-to-commercial (0.8), and industrial-to-industrial

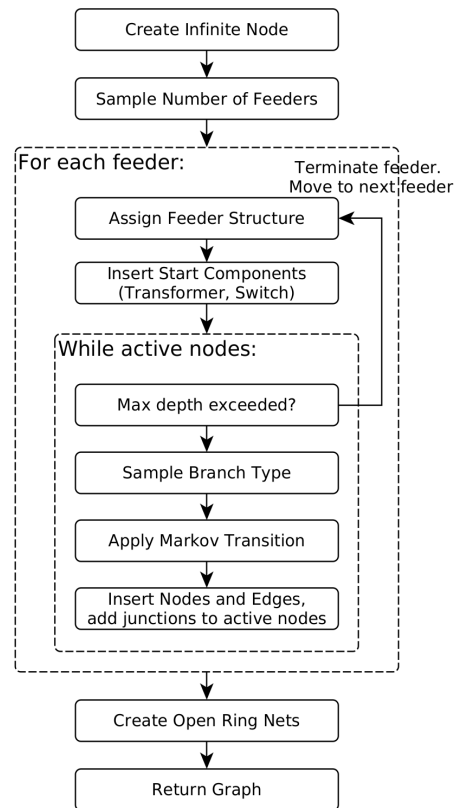


Figure 4: Flow diagram of the proposed probabilistic algorithm for synthetic grid generation.

(0.85), while allowing for lower-probability transitions to other structures. This Markov Chain is shown in figure 3.

Following the initialization of the feeder's first node, each subsequent node is selected based on a branching distribution. At each distribution node, the algorithm probabilistically decides to either

1. create a service drop and continue the feeder (70%),
2. introduce a junction that spawns multiple branches (20%), or
3. terminate the branch with a final service drop (10%).

These choices control the branching behavior and overall depth of the feeder, which is constrained by a user-defined maximum. Loads are placed at service drop nodes and are sampled from context-specific probability distributions defined by their associated consumer structure. For example, residential loads in rural settings may range from 2–20 kW (mean: 5 kW,  $\sigma$ : 5 kW), while commercial and industrial loads follow different statistical profiles. Power factors are also sampled from bounded distributions per consumer type. Furthermore, generators are optionally placed at service drops, with a structure-dependent attachment probability. Generator types such as photovoltaic systems or fuel cells and capacities are drawn from structured probability distributions reflecting realistic deployment scenarios.

Line lengths are sampled from probabilistic parameter sets that differentiate between service drops and distribution lines. Transformers are selected from a set of standard configurations depending on the associated load. Line types are selected based on an estimated maximal line current calculated from the rated power of all loads in downstream grid sections.

### QMC sampling

Finally, additional variation is introduced via quasi-Monte Carlo (QMC) sampling on a selected subset of parameters. Due to the exponential growth of the state space with the number of varied dimensions, this subset must be carefully limited to maintain tractability. Priority is given to parameters with the highest impact on grid behaviour. In particular, line lengths are varied, as they strongly influence voltage drops and power flows, thereby affecting the relationship between measurements and latent states. Load parameters are also included due to their central role in determining the overall operating condition of the network.

By altering these parameters, additional variation is introduced into the dataset. As previously discussed this domain randomization aids in improving generalization capabilities of models trained on this data.

### Computational complexity and scaling

The generation of synthetic power grid topologies requires a careful balance between realism, scalability, and computational feasibility. The random tree algorithm employed in this paper incrementally constructs a tree by attaching nodes based on a stochastic rule set. Its computational complexity scales linearly with the number of nodes, i.e.,  $O(N)$ , making it highly suitable for generating large-scale grid datasets.

In contrast, topology generation methods based on geospatial data such as street map data are computationally intensive due to external data processing and querying. Deep generative models including graph diffusion, variational autoencoders and generative adversarial networks require large existing datasets and training. Thus, among the considered methods, the random tree algorithm exhibits the best computational scalability and minimal dependency on external data.

### Verification

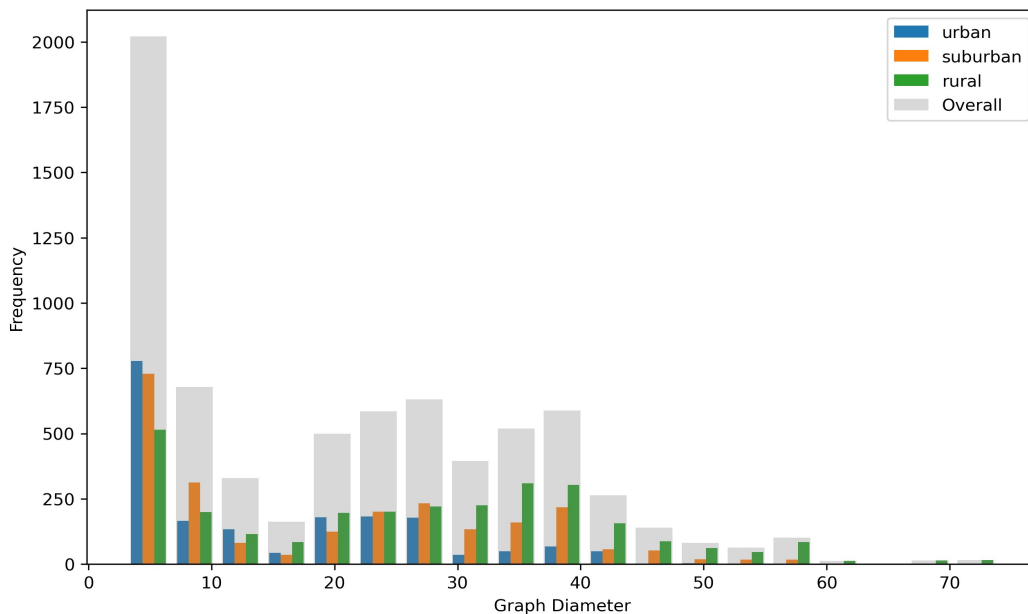
To verify the feasibility of the generated grid topology, the graph is converted to an OpenDSS grid description, and a power flow is calculated. During the power flow simulation, all loads

are set to their rated maximum and generators are shut off. If the power flow solution converges, the apparent power under maximal load flow at each transformer is compared to its rated apparent power. The ratio of measured to rated apparent power is then saved. Similarly measured line currents are compared to the rated maximum thermal line current at each line and the ratio is again saved. For each graph in the dataset a label is added whether it converged with the given parameters along with the respective ratios. Feasible grid topologies are those that converged, did not exceed thermal currents and whose transformer usage is high but not above its rated power.

## RESULTS

In total, 520 different graph structures were generated using the proposed algorithm and associated with realistic yet diverse component parameters. Additionally, selected attributes (load apparent power, line length) were varied via QMC sampling and added to the set of topologies resulting in a total of 9976 power grid topologies. For 9681 of these topologies the load flow simulation using OpenDSS converged and 7101 topologies of these were selected based on realistic transformer usage and line current ratios. The 7101 topologies are itself again divided into 2848 rural, 2390 suburban and 1863 urban topologies.

On average a topology contained 111.37 nodes with the smallest generated grid at a size of 4 nodes and the largest with 690 nodes. The average node degree is 1.989 with an average clustering of 0. Graph diameters range from 3 to 74 with the average at 20.67.



*Figure 5: Histogram of graph diameters per structure combined into bins. Especially in urban structures graph diameters are small while rural structures also show larger graph diameters.*

Figure 5 shows the frequency of occurrence of different graph diameters in the dataset. The graph diameter of rural grids is larger compared to that of urban and suburban grids, reflecting the properties of rural grids with long feeders that are barely interconnected. In general, many of the generated graphs have very small diameters. This is likely a side-effect of a relatively high value for the probability of terminating the current branch during the branch type sampling. However, this effect could as well be mitigated by introducing a depth-dependent branch type probability where at low depths within the feeder the branch is less likely to get terminated.

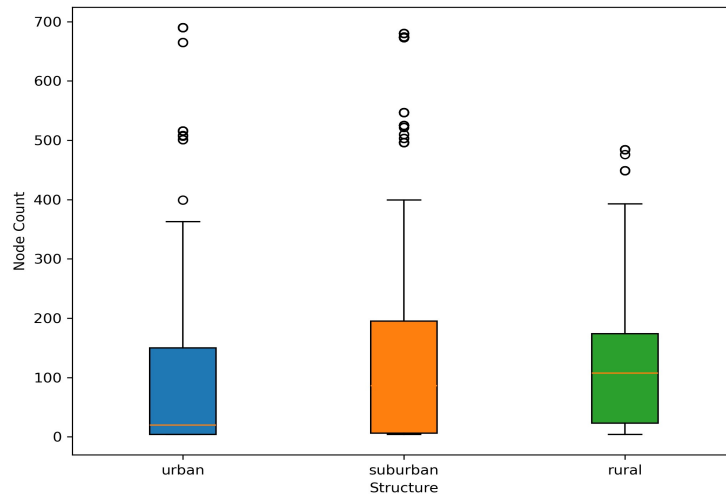


Figure 6: Boxplot of the amount of nodes per structure. For all structures, grids in varying sizes are generated with the interquartile range extending from just a few nodes to almost 200 for suburban structures.

Figure 6 shows the locality and spread of the number of nodes per grid relative to the overall grid structure. For all structures, grids in various sizes were generated with a small skew towards smaller grids especially within urban structures. This is mostly caused by larger structures being filtered out during the line current and transformer rated power check since significantly higher loads were configured for urban settings. In urban settings apartment buildings and multi-unit buildings dominate the grid structure who both exhibit higher power demand than single-family homes in rural settings.

As mentioned earlier, the average node degree is an important measure for realism of power grid topologies as it directly reflects the radial open ring net structure of LV grids.

Figure 7 shows as a scatter plot of the average node degree over the number of nodes in the graph. Radial distribution grids without meshed (lattice) interconnections exhibit an average

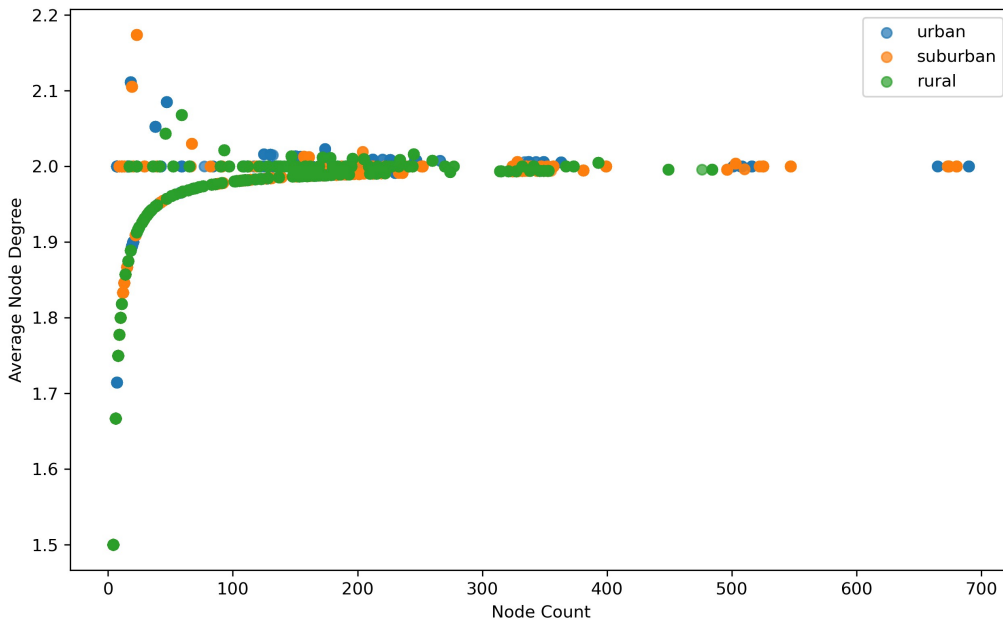


Figure 7: Scatter plot of grids with amount of nodes on x-axis and average node degree on y-axis. The average node degree converges towards 2 for larger grids.

node degree slightly below 2, as expressed by Equation (2). Terminal nodes, representing endpoints of service drops, are connected to a single edge. Internal nodes typically connect to three edges: one upstream edge, one downstream edge, and either a service drop or, in the



case of junctions, an additional downstream edge. Given that for a feeder without interconnections the number of terminal service drop nodes equals the number of internal nodes minus one (corresponding to the feeder’s terminal node), the average node degree is given by  $2(N-1)/N$  resulting in the curve observed in the lower part of figure 7. Those points that do not follow this expected distribution are grids with interconnections between feeders resulting in fewer terminal service drops. For larger grids the curve converges towards an average node degree of 2, a desired property for LV grids that using specific different generation techniques was not achieved [18].

Finally, two representative examples of generated graphs are shown in figures 8 and 9 that represent rural and urban structures. Internal nodes that are neither at service drops nor the starting node are depicted in grey. Those nodes in the current algorithm do not have loads or generators attached to them directly only through service drops.

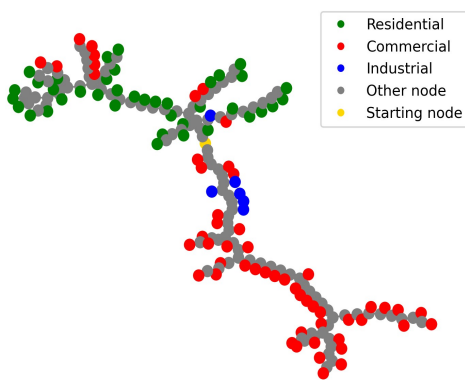


Figure 8: Generated grid in rural grid structure with long feeders and no interconnections.

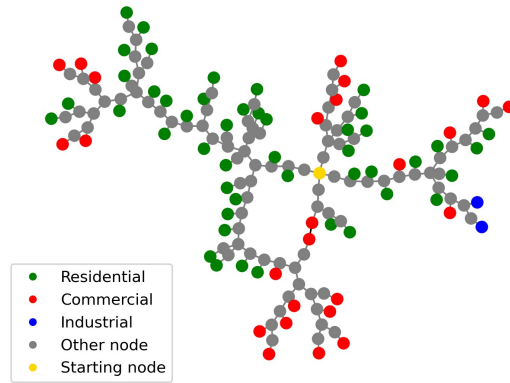


Figure 9: Generated grid in urban structure with more junctions and higher degree of interconnection.

## DISCUSSION

The presented algorithm enables the generation of realistic and diverse synthetic power grid topologies. Realism is confirmed through load flow convergence, by verifying line currents and transformer load and by comparing graph structural properties to existing grid descriptions. Additionally, the generated grids are enriched with diverse parameters, drawn from pre-defined probability distributions and some varied through QMC sampling.

Despite its advantages, several limitations and future directions remain. No comprehensive parameter optimization has yet been performed to determine probability distributions that best reflect empirical LV grid characteristics. Such optimization is computationally intensive due to the high dimensionality of the parameter space.

Moreover, literature supports the assertion that domain randomization and diverse training data facilitate the generalization capability of machine learning models. However, due to the current absence of publicly available baseline or benchmark models specifically designed for tasks on unknown or unseen power grid topologies, direct empirical comparisons remain challenging.

Further graph generation techniques could be explored and compared to the presented algorithm. Transferring knowledge gained from graph generation in different domains to the power grid domain – albeit difficult due to the different structural properties – may be a further avenue for research. Especially, VAEs and graph diffusion models could be of interest for future research into synthetic power grid generation as they can provide graphs along with properties. Moreover,

it is possible to encode meaning in the latent space of the models in a way that allows generating grids based on user input – much like current image generation models generate images based on text provided by the users. However, the applicability of such models is currently constrained by the scarcity of suitable training datasets in the power systems domain.

This paper has presented an algorithm that enables fast and large-scale generation of synthetic grids with controlled structural variability and diverse parameters, making it particularly well-suited for data-driven applications where a high volume of diverse training samples is needed.

## ACKNOWLEDGMENT

This research was funded by the German Federal Ministry for Economic Affairs and Climate Action grant number 03EI4048R.

## NOMENCLATURE

### Acronyms

DER – Distributed Energy Resources  
 LV – Low Voltage  
 MV – Medium Voltage  
 HV – High Voltage  
 ML – Machine Learning  
 OGB – OpenGraphBenchmark  
 QMC – Quasi-Monte-Carlo  
 VAE – Variational Autoencoders

### Symbols

$\sigma$  – Standard deviation  
 $k_i$  – Node degree of node  $i$   
 $C_i$  – Clustering coefficient of node  $i$   
 $D$  – Graph Diameter

## REFERENCES

- [1] A. S. Zamzam, X. Fu, and N. D. Sidiropoulos, “Data-Driven Learning-Based Optimization for Distribution System State Estimation,” *IEEE Transactions on Power Systems*, vol. 34, no. 6, pp. 4796–4805, Nov. 2019
- [2] Y. Wang, H. Ma, X. Xiao, Y. Wang, Y. Zhang, and H. Wang, “Harmonic State Estimation for Distribution Networks Based on Multi-Measurement Data,” *IEEE Transactions on Power Delivery*, vol. 38, no. 4, pp. 2311–2325, Aug. 2023
- [3] M. Fotopoulou, S. Petridis, I. Karachalios, and D. Rakopoulos, “A Review on Distribution System State Estimation Algorithms,” *Applied Sciences*, vol. 12, no. 21, p. 11073, Nov. 2022
- [4] A. Mabrouk and R. Rajagopal, “Distribution Grid Topology Estimation: A New Approach-based on Bayesian Network Models,” in *2022 IEEE Eighth International Conference on Big Data Computing Service and Applications (BigDataService)*, Newark, CA, USA: IEEE, Aug. 2022, pp. 124–131.
- [5] P. Mack, M. de Koster, P. Lehnen, E. Waffenschmidt, and I. Stadler, “Power Quality State Estimation for Distribution Grids Based on Physics-Aware Neural Networks—Harmonic State Estimation,” *Energies*, vol. 17, no. 21, Art. no. 21, Jan. 2024
- [6] X. Hu, H. Hu, S. Verma, and Z.-L. Zhang, “Physics-Guided Deep Neural Networks for Power Flow Analysis,” *IEEE Transactions on Power Systems*, vol. 36, no. 3, pp. 2082–2092, May 2021

- [7] S. Park, F. Gama, J. Lavaei, and S. Sojoudi, “Distributed Power System State Estimation Using Graph Convolutional Neural Networks”
- [8] O. Kundacina, M. Cosovic, D. Miskovic, and D. Vukobratovic, “Distributed Nonlinear State Estimation in Electric Power Systems using Graph Neural Networks,” Sep. 08, 2022, *arXiv*: arXiv:2207.11465.
- [9] C. Mateo Domingo, T. Gomez San Roman, A. Sanchez-Miralles, J. P. Peco Gonzalez, and A. Candela Martinez, “A Reference Network Model for Large-Scale Distribution Planning With Automatic Street Map Generation,” *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 190–197, Feb. 2011
- [10] M. Grzanic, M. G. Flammini, and G. Pretico, “Distribution Network Model Platform: A First Case Study,” *Energies*, vol. 12, no. 21, Art. no. 21, Jan. 2019
- [11] R. Meyur *et al.*, “Ensembles of realistic power distribution networks,” *Proc Natl Acad Sci U S A*, vol. 119, no. 42, p. e2205772119, Oct. 2022
- [12] B. Abhilash, C. Syranidou, J. Linssen, and D. Stolten, “Geo-referenced synthetic low-voltage distribution networks: A data-driven approach,” in *2021 IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe)*, Oct. 2021, pp. 1–6.
- [13] W. Medjroubi, U. P. Müller, M. Scharf, C. Matke, and D. Kleinhans, “Open Data in Power Grid Modelling: New Approaches Towards Transparent Grid Models,” *Energy Reports*, vol. 3, pp. 14–21, Nov. 2017
- [14] J. Amme, G. Pleßmann, J. Bühler, L. Hülk, E. Kötter, and P. Schwaegerl, “The eGo grid model: An open-source and open-data based synthetic medium-voltage grid model for distribution power supply systems,” *J. Phys.: Conf. Ser.*, vol. 977, no. 1, p. 012007, Feb. 2018
- [15] J. Kays, A. Seack, T. Smirek, F. Westkamp, and C. Rehtanz, “The Generation of Distribution Grid Models on the Basis of Public Available Data,” *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 2346–2353, May 2017
- [16] B. Reveron Baecker, S. Candas, D. Tepe, and A. Mohapatra, “Generation of low-voltage synthetic grid data for energy system modeling with the pylovo tool,” *Sustainable Energy, Grids and Networks*, vol. 41, p. 101617, Mar. 2025
- [17] F. Giacomarra, G. Bet, and A. Zocca, “Generating Synthetic Power Grids Using Exponential Random Graph Models,” *PRX Energy*, vol. 3, no. 2, p. 023005, Jun. 2024
- [18] Z. Wang, A. Scaglione, and R. J. Thomas, “Generating Statistically Correct Random Topologies for Testing Smart Grid Communication and Control Networks,” *IEEE Transactions on Smart Grid*, vol. 1, no. 1, pp. 28–39, Jun. 2010
- [19] D. J. Watts and S. H. Strogatz, “Collective dynamics of ‘small-world’ networks,” vol. 393, 1998
- [20] C. Vignac, I. Krawczuk, A. Siraudin, B. Wang, V. Cevher, and P. Frossard, “DiGress: Discrete Denoising diffusion for graph generation,” May 23, 2023, *arXiv*: arXiv:2209.14734.
- [21] M. Simonovsky and N. Komodakis, “GraphVAE: Towards Generation of Small Graphs Using Variational Autoencoders,” Feb. 09, 2018, *arXiv*: arXiv:1802.03480.
- [22] M. Zhang *et al.*, “A Survey on Graph Diffusion Models: Generative AI in Science for Molecule, Protein and Material,” 2023
- [23] J. Kukačka, V. Golkov, and D. Cremers, “Regularization for Deep Learning: A Taxonomy,” Oct. 29, 2017, *arXiv*: arXiv:1710.10686.
- [24] J. Tobin, R. Fong, A. Ray, J. Schneider, W. Zaremba, and P. Abbeel, “Domain Randomization for Transferring Deep Neural Networks from Simulation to the Real World,” Mar. 20, 2017, *arXiv*: arXiv:1703.06907.
- [25] W. Hu *et al.*, “Open Graph Benchmark: Datasets for Machine Learning on Graphs,” Feb. 24, 2021, *arXiv*: arXiv:2005.00687.
- [26] Conseil international des grands réseaux électriques, Ed., *Benchmark systems for network integration of renewable and distributed energy resources*. Paris: CIGRÉ, 2014.

- [27] G. Kerber, “Aufnahmefähigkeit von Niederspannungsverteilnetzen für die Einspeisung aus Photovoltaikkleinanlagen,” 2011.
- [28] J. Dickert, M. Domagk, and P. Schegner, “Benchmark low voltage distribution networks based on cluster analysis of actual grid properties,” in *2013 IEEE Grenoble Conference*, Grenoble, France: IEEE, Jun. 2013, pp. 1–6.
- [29] Fraunhofer IEE and Universität Kassel, “Power System Test Cases — pandapower 2.14.9 documentation.”