Constraint-Driven Deep Learning for N-k Security Constrained Optimal Power Flow

Bastien N. Giraud, Ali Rajaei, Jochen L. Cremer
Department of Electrical Sustainable Energy, Delft University of Technology
Delft, the Netherlands
bastien.g@live.nl, a.rajaei@tudelft.nl, j.l.cremer@tudelft.nl

***Abstract*—The transition to green energy is reshaping the energy landscape, marked by increased integration of renewables, distributed resources, and the electrification of other energy sectors. These changes challenge grid security, particularly regarding the N-1 security criterion, a crucial factor in preventing blackouts. This necessitates studying the security constrained optimal power flow (SCOPF) problem with multiple line outages (N-k). Conventional methods exhibit poor scalability as k increases. This paper proposes a constraint-driven machine learning (ML) approach using line outage distribution factors (LODF). The method shows promise in its ability to scale effectively to N-k contingencies. Key contributions include a deterministic approach for N-k security and a probabilistic security assessment. Case studies on the IEEE 39-bus and the IEEE-118 bus systems show the approach's effectiveness in identifying violating post-contingency cases, with up to 173× speedups and close to optimal dispatch costs.**

***Index Terms*—Neural Network, N-k SCOPF, LODF, constraint-driven**

I. Introduction

The energy sector is swiftly evolving towards sustainability, driven by the increasing integration of renewable energy sources, expanding distributed energy assets, and electrification across diverse domains [1]. These developments pose a challenge to grid security, or operational reliability, making it harder to meet the N-1 security criterion, the failure of which has led to numerous blackouts. Furthermore, unforeseen weather events, now more frequent due to climate change, account for the second major cause of these blackouts [2, 3]. The growing grid complexity highlights the importance of N-k outages, where k lines fail simultaneously, necessitating enhanced N-k security and increased resilience against high impact low probability (HILP) weather triggered events [4]. While the N-1 criterion historically balanced security and computational efficiency, recent developments advocate for a probabilistic risk-based assessment, encompassing N-k contingencies with k >1, to enhance power system operational reliability [5]. Conventional methods encounter difficulties in effectively tackling the computationally demanding N-k security due to its inherent combinatorial complexity. This paper focuses on the optimization problem faced by transmission system operators (TSOs), known as N-k security constrained optimal power flow (SCOPF), that seeks to find a preventive dispatch which minimizes the thermal flow limit violation in the case of k line outages.

In the context of solving N-k SCOPF problems, different approaches have been introduced to reduce problem size. In [6, 7], Benders decomposition (BD) [8] is used, to decompose the problem into a master problem and subproblems representing contingency cases. Furthermore, [9] employs a column-and-constraint generation algorithm (C&CGA) to iteratively add the most violating contingencies. However, these approaches face challenges like slow convergence when dealing with numerous contingencies or the risk of generating infeasible predictions when too many constraints are introduced. Alternatively, [10] employs an iterative contingency screening method to identify the most critical contingencies, allowing the SCOPF to be solved with a selected subset of contingencies. In [11, 12], this screening process leverages line outage distribution factors (LODFs) due to their computational efficiency in managing multiple contingencies. The proposed approaches in [6, 7, 9–12] do not scale with the number of outages k, as demonstrated in [13]. A comparison between BD, a combination of C&CGA with robust optimization (RO), and LODFs reveals issues such as timeouts and slow convergence for N-3. This is particularly limiting for addressing the N-k SCOPF problem that requires multiple daily resolutions.

The recurrent nature of the SCOPF problem has prompted exploration of machine learning (ML) techniques to approximate the solution. Reference [14] predicts generator setpoints from power system loads for N-1 SCOPF using a neural network (NN). Similarly, reference [15] uses a NN to classify binding constraints. Both approaches in [14, 15] do not scale with increasing k. This is because they rely on training data, which becomes exceedingly difficult to generate for k >1. To address these limitations, [16] proposes a C&CGA-ML algorithm in combination with RO, iteratively adding constraints, succesfully solving N-1 SCOPFs, but unable to solve N-k SCOPFs. Moreover, in [17], an adversarially robust ML approach is introduced, using a min-max strategy where the worst-case contingency is computed and incorporated in the SCOPF. However, the success of [17] is highly reliant on the ability to identify the worst-case contingency. To address the issue of prediction feasibility, recent attention has focused on the integration of constrained optimization and ML techniques [18, 19]. However, these methods also struggle in finding secure and feasible predictions while maintaining scalability with increasing k. Therefore, there is a pressing need to develop novel, practical methodologies that can effectively

approximate the N-k SCOPF problem.

This paper proposes a constraint-driven ML approach to approximate the N-k SCOPF problem, considering all possible line contingencies. The approach uses the computationally efficient linearized DC OPF, which is widely adopted in transmission network operation to investigate the thermal flow limits [20]. A NN learns the mapping from loads to generator setpoints. With the use of LODFs, all post-contingency flows are efficiently computed through a simple matrix multiplication. The LODF matrices are expressed in sparse format to significantly reduce memory cost and accelerate the computation. An algorithm is developed to reduce the computational graph that guides the backpropagation process. The proposed constraint-driven approach obviates the need for labeled training data, thereby eliminating the computational burden associated with solving the N-k SCOPF. Furthermore, to formulate a risk-based security criterion, contingency probabilities are included during training. These probabilities are computed considering a spatial correlation between individual line outages, a characteristic feature of extreme weather events [21, 22]. The performance of the security assessment is evaluated using loss of load expectation (LOLE) and expected energy not supplied (EENS). The main contributions of this paper are:

- The deterministic constraint-driven approach to approximate N-k SCOPFs, considering all line contingencies using LODFs. The proposed approach is weakly supervised without the need for labeled training data.
- The computational graph memory reduction for fast and efficient implementation.
- The probabilistic security assessment to formulate a N-k risk-based security criterion, providing an alternative to the current deterministic N-1 security criterion.

The rest of the paper is organized as follows. Section II presents the constraint-driven approach. Section III explains the security assessment approach. Section IV presents the case studies on the IEEE 39-bus and IEEE 118-bus systems. Section V concludes the paper.

II. Proposed Constraint-Driven Learning Approach

A. Method Overview

The proposed approach uses ML to approximate the solution of the N-k SCOPF optimization. Fig. 1 shows the workflow of the proposed approach, which consists of 3 steps: prediction, feasibility restoration, and post-contingency.

1) Prediction: A NN learns the mapping between the loads to generator setpoints as:

$$\alpha_G = \Phi_W\big(P_D\big) \tag{1a}$$

$$\widehat{P}_G = \alpha_G \cdot (P_G^{max} - P_G^{min}) + P_G^{min} \tag{1b}$$

where P_D is the load vector, $\widehat{P}_G$ is the vector of generator setpoints, P_G^{min} and P_G^{max} are the minimum and maximum generation limit, and W represents the weight matrices of the NN. To ensure the NN adheres to the generator limits, the NN outputs a scaling vector $\alpha_G \in [0, 1]$ in (1a), and $\widehat{P}_{G_n}$ is computed based on (1b).

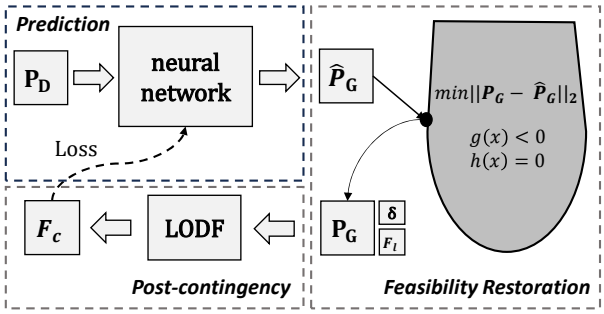

Fig. 1: The workflow of the proposed constraint-driven approach. A NN maps the loads to generator setpoints. A feasibility restoration layer restores base case infeasible setpoints. Using LODFs, all post-contingency flows are obtained.

2) Feasibility restoration: The NN prediction of generator setpoints must satisfy the physical constraints of the system during normal operation. Thus, a feasibility restoration layer is developed. In case of any violation of the DC power flow (PF) equations, the feasibility restoration layer projects the NN prediction into the feasible region by solving the following optimization:

$$\min_{P_{G_n}} \sum_{n \in \Omega^G} ||P_{G_n} - \widehat{P}_{G_n}||_2 \tag{2a}$$

s.t.

$$\boldsymbol{B} \cdot \boldsymbol{\delta} = \boldsymbol{P_G} - \boldsymbol{P_D} \tag{2b}$$

$$-F_l^{max} \leq \frac{1}{x_l}(\delta_i - \delta_j) \leq F_l^{max} \quad \forall i, j \in \Omega^B, \forall l \in \Omega^L \tag{2c}$$

$$P_{G_n}^{min} \leq P_{G_n} \leq P_{G_n}^{max} \quad \forall n \in \Omega^G \tag{2d}$$

where the objective function (2a) minimizes the L2 norm between the predicted setpoints ($\widehat{P}_{G_n}$) and the setpoints adhering to the feasible region (P_{G_n}). The set of buses is $i \in \Omega^B$, set of lines is $l \in \Omega^L$, set of generators is $n \in \Omega^G$ and set of loads $d \in \Omega^D$. x_l is the line reactance and F_l^{max} is the line thermal limit. The restoration optimization (2) outputs other operational variables including phase angles (δ_i) and line flows (F_l). Restoration optimization (2) is implemented as a NN layer using CVXPYlayer [23], a convex solver that computes the gradients through the solution of the optimization.

3) Post-contingency: The set of all post-contingency flows for all of the N-k outages is computed using LODFs:

$$F_l^c = F_l^0 + LODF_{N-k} \times F_l^0 \quad \forall l \in \Omega^L \tag{3}$$

where F_l^0 is flow at normal state, and F_l^c is the flow at contingency case c. The predicted base case flows before the restoration layer are computed using power transfer distribution factors (PTDFs) as:

$$\widehat{F_l^0} = PTDF(\widehat{P}_G - P_D) \tag{4}$$

Finally, the NN is trained using backpropagation on the following loss function:

$$\begin{aligned} Loss &= \lambda_c c_G^T P_G + \lambda_0 \|ReLU(|\widehat{F_l^0}| - F_l^{max})\|_1 \\ &+\lambda_1 \|ReLU(|F_l^c| - F_l^{max})\|_1 + \lambda_2 \|\Sigma \widehat{P}_G - \Sigma P_D\|_1 \end{aligned} \tag{5}$$

where c_G is the vector of generator cost and λ_x are hyper-parameters of the model. The loss function (5) considers the dispatch cost, the predicted base case flow violations, the predicted generation-demand imbalance and the post-contingency flow violations. Therefore, the loss function of the proposed approach is constraint-driven and eliminates the need for labeled training data.

B. Line Outage Distribution Factors

The LODFs are computed using the method of generalized line outage distribution factors [24].

$$LODF_{M,O} = PTDF^0_{M,O}(E - PTDF^0_{O,O})^{-1} \tag{6}$$

The letter M denotes the set of monitored lines, the letter O denotes the set of outaged lines and the superscript 0 indicates the pre-contingency state. Here, $E \in \mathbb{R}^{|\Omega^L| \times |\Omega^L|}$ is the identity matrix. The set of outaged lines is denoted with $\Omega^{\hat{L}}$. We define a bus-to-monitored line incidence matrix $A \in \{-1, 1\}^{|\Omega^L| \times |\Omega^B|}$, where 1 indicates the 'from bus' and -1 indicates the 'to bus'. $\hat{A}$ is the bus-to-tripped line incidence matrix where $\hat{A} \in \{-1, 1\}^{|\Omega^L| \times |\Omega^B|}$. The matrix $B \in \mathbb{R}^{|\Omega^B| \times |\Omega^B|}$ is the susceptance matrix which is computed using the line reactances x_{ij} as follows:

$$B_{ii} = \sum_{j=1} \frac{1}{x_{ij}} \quad \forall i, j \in \Omega^B \tag{7a}$$

$$B_{ij} = -\frac{1}{x_{ij}} \quad \forall i, j \in \Omega^B \tag{7b}$$

To avoid singularity, the row and the column corresponding to the slack bus are removed from the susceptance matrix B. The matrix X is the inverse of the susceptance matrix $X = B^{-1}$. Finally, the matrix $B_{br} \in \mathbb{R}^{|\Omega^B| \times |\Omega^B|}$ is a diagonal matrix with the line reactances on the diagonal entries, and zeros elsewhere. Subsequently, $PTDF^0_{M,O}$, $PTDF^0_{O,O}$ and $PTDF$ are computed as follows.

$$PTDF^0_{M,O} = B_{br} \times A \times X \times \hat{A}^T \tag{8}$$

$$PTDF^0_{O,O} = B_{br} \times \hat{A} \times X \times \hat{A}^T \tag{9}$$

$$PTDF = B_{br} \times A \times X \tag{10}$$

For each contingency case, a square single LODF ($SLODF$) matrix of size $\mathbb{R}^{|\Omega^L| \times |\Omega^L|}$ is constructed. As Fig. 2 shows, these $SLODF$ matrices are then vertically stacked to form a larger rectangular matrix, which is called the full LODF ($FLODF$) matrix.

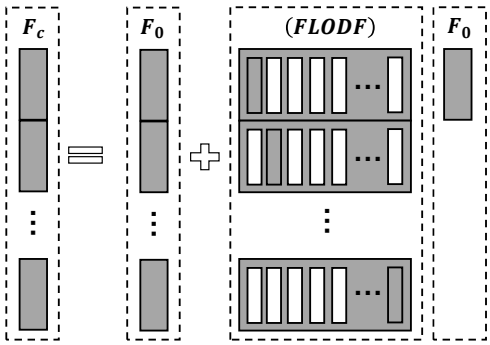

Fig. 2: FLODF created by stacked SLODFs for a N-1 case. The white columns are zero columns, the grey columns contain the LODFs.

C. Memory Reduction with Sparsity FLODF

As visualized in Fig. 2, the construction of the $FLODF$ matrix is designed to enable a single matrix multiplication to obtain all post-contingency flows. The $FLODF$ matrix is sparse as the number of columns containing non-zero elements equals the number of outages k. The number of contingency cases $|\Omega^C|$ holds combinatorial complexity and can be computed as:

$$\Omega^C(l, k) = \frac{|\Omega^L|!}{k!(|\Omega^L| - k)!} \tag{11}$$

Each $SLODF$ contains k non-zero columns. Each column has $|\Omega^L|$ entries, with ($|\Omega^L|$ - k) possible non-zero entries. The maximum possible number of non-zero elements in a $SLODF$ is $k(|\Omega^L| - k)$. The maximum possible amount of nonzero elements for the $FLODF$ is $k|\Omega^C|(|\Omega^L| - k)$. The number of elements in the $FLODF$ is $|\Omega^C||\Omega^L|^2$. A lower bound of sparsity can be computed for the $FLODF$ matrix:

$$sparsity = 1 - \frac{k(|\Omega^L| - k)}{|\Omega^L|^2} \tag{12}$$

This is a theoretical lower bound since for certain cases the LODF is zero due to the network topology, and in some contingency cases a part of the network gets disconnected. For these islanding cases, the LODFs are intentionally set to zero. Islanding is recognized by the singularity of the matrix $(E - PTDF^0_{O,O})^{-1}$. As k increases, more islanding cases occur, especially for smaller systems. To reduce the memory of the $FLODF$ matrix, the matrix is expressed in sparse coordinate (COO) format, storing only the non-zero elements along with their row and column indices. Consequently, a sparse matrix multiplication is performed between the $FLODF$ and the base case flows.

D. Computational Graph Memory Reduction

A computational graph is a directed graph that visually represents a mathematical or computational model, illustrating the data flow and interconnection of various operations. Computational graphs find significant utility in deep learning, where they guide the gradient flow for backpropagation in tensors. Fig. 3 shows the computational graph of the proposed

Algorithm 1 Reducing the computational graph

1: Input: P_D
2: **for** e in Epochs **do**
3: **NN:** $\widehat{P}_G \leftarrow \Phi_W(P_D)$ Eq. (1)
4: **function** RESTORATION(P_D, $\widehat{P}_G$) Eq. (2)
5: **Input:** P_D, $\widehat{P}_G$
6: **return** P_G, F_l^0, δ_i
7: **end function**
8: requires grad = False
9: $F_l^c \leftarrow F_l^0 + FLODF \times F_l^0$ Eq. (3)
10: **if** $F_l^c < F_l^{max}$ **then**
11: Obtain row indices non-violating F_l^c
12: Remove non-violating rows $FLODF$
13: $FLODF_{reduced} \leftarrow FLODF$
14: **end if**
15: requires grad = True
16: $F_l^c \leftarrow F_l^0 + FLODF_{reduced} \times F_l^0$
17: **Loss:** Eq. (5)
18: Perform backward pass
19: **end for**

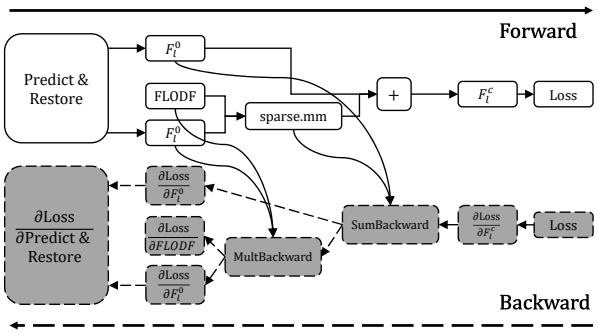

Fig. 3: The computational graph of the proposed approach.

approach. When dealing with extensive tensors, such as those involving large $FLODF$ matrices and multiple samples, substantial memory is required. To reduce memory requirements, Algorithm 1 is presented. Line 3 is the forward pass of the NN in (1). Lines 4-7 show the feasibility restoration step in (2). Lines 8-9 compute the post-contingency flows based on (3) without the construction of the computational graph (requires grad = False). Lines 10-14 remove rows corresponding to non-violating flows from the $FLODF$ matrix. Lines 15-16 construct the computational graph (requires grad = True) only for post-contingency flows with a violation, which is often a small portion of the contingency cases, post-contingency, therefore reducing memory requirements. Finally, the NN weights are updated by a backward pass on the loss function in (5).

III. Probabilistic Security Assessment

The integration of probability considerations within the proposed approach facilitates decision-making in accordance with an N-k risk-based security criterion. The goal is to achieve N-k security and heightened resilience to changing outage probabilities caused by unforeseen HILP events, thus enhancing power system operational resilience [3].

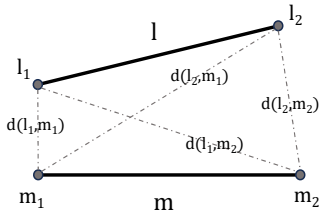

Fig. 4: The distances between any pair of buses of two lines. The shortest distance is used, in the Figure d(l_1,m_1).

A. Joint Probabilities with Copulas

To compute the joint line outage probabilities, the Gaussian copula function introduced in [21] is used. It is assumed that the inverse stress on transmission line l follows a Gaussian distribution and is given by a random variable $\mathbf{Y_l} = N(\mu_l, \sigma_l)$. μ_l and σ_l are the mean and standard deviation respectively and the cumulative distribution function is given by:

$$\mathcal{F}_{Y_l}(y_l) = \frac{1}{2}\left[1 + erf\left(\frac{y_l - \mu_l}{\sigma_l\sqrt{2}}\right)\right] \tag{13}$$

For each line l, μ_l and σ_l are chosen so that $F_{Y_l}(0) = p_l$ where p_l is the independent probability of line l going out. Assuming $\mu_l = 1\ \forall l$, σ_l is determined as:

$$\sigma_l = \frac{-1}{erf^{-1}(2p_l - 1)\sqrt{2}} \tag{14}$$

To compute the covariance between two probabilities, we assume exponential decay of correlation with distance between lines l and m $\rho_{l,m} = \rho_0 e^{-d_{l,m}/L}$ for $l, m \in \Omega^L$. Here, ρ_0 represents maximal correlation at zero distance, L is the characteristic length, and $d_{l,m}$ is simplified as the shortest distance between any pair of buses of the lines ($d_{l,m} = min\{d(l_1, m_1), d(l_1, m_2), d(l_2, m_1), d(l_2, m_1)\}$) as shown in Fig. 4. This exponential decay models spatial behavior seen in events like earthquakes, tornadoes, and hurricanes, which often exhibit geographic correlations. The covariance between lines is calculated as $cov(l, m) = \rho_{l,m}\sigma_l\sigma_m$. This equation allows for the construction of the covariance matrix $\mathbf{C}$. The copula is formed using the probability density function of the multivariate normal distribution given in Eq. (15).

$$f(y) = \frac{1}{\sqrt{(2\pi)^k|\mathbf{C}|}} exp\left(-\frac{1}{2}(\mathbf{y} - \mu)^T\mathbf{C}^{-1}(\mathbf{y} - \mu)\right) \tag{15}$$

By integrating Eq. (15) over the region in the joint probability distribution representing k = $\{2, 3\}$ outages, the joint probability of those outages is obtained. For this work, the joint probabilities for all possible combinations of k = $\{2, 3\}$ are computed to form vector π_{N-k}, containing the probabilities for all N-k contingencies.

B. Expected Violations

By multiplying the contingency probabilities with the post-contingency flow violations, the expected post-contingency flow violations are obtained. Using these expected violations the probability of the system being secure or not is assessed.

$$\mathbb{E}[F_l^{viol}] = \|\pi_{N-k} \cdot ReLU(|F_l^c| - F_l^{max})\|_1 \quad (16)$$

The loss function employed to minimize the expected post-contingency flow violations is similar as (5). The only modification is the expected violations are used instead of the actual violations.

$$\begin{aligned} PLoss = \lambda_c c_G^T P_G + \lambda_0 \|ReLU(|\widehat{F_l^0}| - F_l^{max})\|_1 \\ +\lambda_1 \|\pi_{N-k} \cdot ReLU(|F_l^c| - F_l^{max})\|_1 \\ +\lambda_2 \|\Sigma \widehat{P}_G - \Sigma P_D\|_1 \end{aligned} \quad (17)$$

The NN is trained using the probabilistic loss function in (17), which accounts for all contingencies across multiple security levels including their probabilities.

Probabilistic reliability indicators, such as loss of load expectation (LOLE) and expected energy not supplied (EENS), are metrics used to assess power system's reliability. The LOLE represents the expected amount of time per period that the demand cannot be fully supplied. Moreover, EENS quantifies the amount of energy that is expected not to be supplied during a given time period, serving as a quantitative addition to LOLE [25]. We evaluate these indices during normal operation, and we assess the effect on these reliability indices when our proposed approach undergoes an HILP event. This assessment enhances our understanding of short-term resilience to extreme weather events [2]. Furthermore, these indices reveal the impact of the HILP event in the absence of any recovery actions [26].

IV. Case Studies

A. Settings and Test Systems

The deterministic approach is tested on the IEEE 39-bus and the IEEE 118-bus systems, while the probabilistic approach is tested on the IEEE 39-bus system [27]. Demand points are sampled around $\pm$ 0.25 of the nominal load. The samples are drawn from a Kumaraswamy(1.6,2.8) distribution with a Pearson correlation coefficient of 0.75 to model the correlation pattern between loads. Both systems are trained using a multi layer perceptron (MLP) with three hidden layers, dropout (0.2), ReLU activation and a scaled tanh activation in the output. For the 39-bus system, 1000 samples are used with a training-validation split of 4/1. 8 neurons per hidden layer result in 410 parameters. For the 118-bus system, 4000 samples are used with the same training-validation split. 16 neurons per hidden layer result in a total of 2467 parameters. Testing is done on a not previously seen data set of 1000 samples.

For the probabilistic study, testing is also done on a not previously seen data set of 1000 samples. Equal outage probabilities are assumed for all lines, where $p_l = 0.005$ for $l \in \Omega^L$.

TABLE I: Sparsity, element reduction and islanding cases

System	k	Sparsity [%]	Reduction COO [%]	Islanding [%]
39-bus	1	98.5	-95.6	6.06
	2	97.6	-92.7	17.80
	3	97.5	-92.4	34.86
	4	98.1	-94.2	55.29
	5	98.9	-96.7	75.24
118-bus	1	99.6	-98.8	4.84
	2	99.3	-97.8	9.90
	3	99.0	-97.0	15.12

These probabilities are then scaled according to individual line lengths $p_l = 0.005 \frac{L_{len} - \mu_{len}}{\sigma_{len}}$ for $l \in \Omega^L$, assuming that longer lines have a higher outage probability than shorter lines. Joint probabilities for $k > 1$ are computed using the previously described copula analysis. We set the characteristic length L to 50 km and the maximum correlation ρ_0 to 0.15. An outage frequency of 0.0051 /cctkmyr (circuit-kilometer-year) is considered for all lines with an average repair time of 44h per individual line in outage [25]. It is assumed that for 10% line overloads no load shedding is necessary, but when a line is more than 10% overloaded, load shedding has to take place to avoid an outage.

All tests are performed on the DelftBlue supercomputer with 16GB RAM and NVIDIA Volta V100S GPUs [28]. Both the DCOPF and SCOPF are implemented in the CVXPY 1.3.0 library with ECOS 2.0.11 solver. The feasibility restoration layer is implemented using CVXPYlayers 0.1.5. The NN is implemented in PyTorch 1.13.1. The code to reproduce the case studies of this paper can be accessed in GitHub [29].

The following baselines are compared against the proposed deterministic approach:

- Contingency screening (CS) baseline: the N-k SCOPF optimization with iterative CS. At each iteration, the 20 most critical contingency cases are added to the SCOPF. We repeat the CS for 3 iterations.
- Heuristic (H) baseline: a set of critical contingency set is created offline on a separate dataset. The N-k SCOPF optimization is solved considering only this set of critical contingencies.

The proposed probabilistic approach is compared against:

- N-1 SCOPF considering all N-1 contingencies.
- N-2 SCOPF considering all N-1 contingencies and 20 most critical N-2 contingencies selected through CS process.

B. Memory Reduction

This case study assesses the achieved memory reduction. Table I presents the percentage of islanding cases, the sparsity of the $FLODF$ matrix, and the achieved reduction in elements through its expression in COO format. A noteworthy decrease in the number of elements is evident.

Fig. 5 compares the proposed "reduced graph" algorithm with the "full graph" baseline with unmodified tensors using

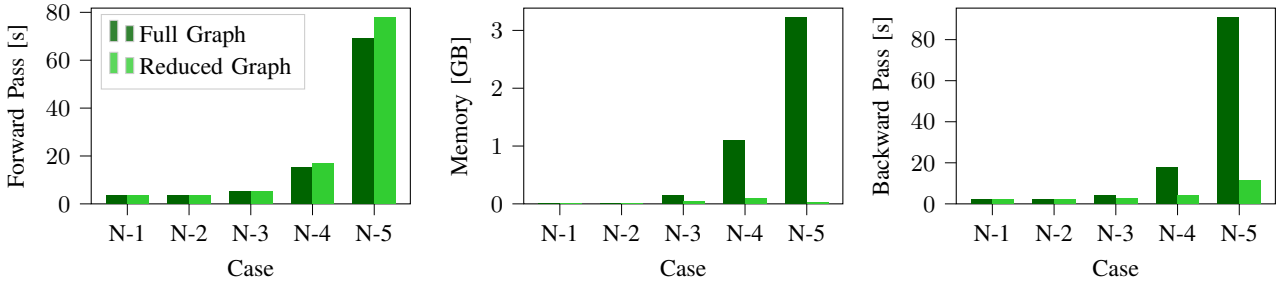

Fig. 5: Comparing speed of the forward pass, backward pass and the memory consumption of the full computational graph against the reduced computational graph.

TABLE II: Proposed deterministic approach against CS baseline

System	Model	#violations [%]		
		N-1	N-2	N-3
39-bus	Proposed approach	0.31	1.33	2.16
	CS	0.06	0.74	3.30
118-bus	Proposed approach	1.08	1.70	2.64
	CS	0.91	1.03	5.88

TABLE III: Proposed deterministic approach against H baseline

System	Model	#violations [%]		
		N-1	N-2	N-3
39-bus	Proposed approach	0.25	1.19	2.00
	H	0.01	0.60	2.87
118-bus	Proposed approach	0.89	1.53	2.42
	H	0.86	1.01	4.52

a batch size of 100 samples for the 39-bus system. For the N-5 case, a batch size of 50 is used due to memory issues of the 'full graph' baseline. Fig. 5 shows that the 'reduced graph' algorithm is slightly slower during the forward pass as a result of the added computations to identify violating contingencies. The increase in time is small, as less than 1% of post-contingency flows violate constraints resulting in quick matrix multiplications with small matrices. In contrast, the N-5 case achieves a 100-fold memory reduction and 8x speedup in the backward pass by reducing the computational graph and the size of tensors involved in backpropagation.

TABLE IV: Proposed deterministic approach cost and speedup compared to baselines

		39-bus		118-bus	
		Cost [%]	Speedup	Cost [%]	Speedup
CS	N-1	+1.44	12×	+2.51	158×
	N-2	+1.17	49×	-6.13	173×
	N-3	+0.51	41×	-5.29	27×
H	N-1	+1.45	15×	+2.88	76×
	N-2	+1.22	21×	-6.41	165×
	N-3	+0.10	21×	-4.92	15×

C. Deterministic Post-Contingency Violations

This case study assesses how many post-contingency cases violate the network's flow limits. Table II and Table III show the number of post-contingency violations of the proposed deterministic approach compared to the CS and H baseline respectively. "#violations" for each method refers to the number of testing samples where the dispatch of that method cause any post-contingency violations. Table IV shows the relative changes in cost and speed. Regarding the baselines, the H baseline leverages the pre-existing contingency list and significantly reduces computation time compared to the CS baseline for the 118-bus systems. Fig. 6 displays the percentage of identified violating cases for the model trained at individual security levels. Fig. 6 investigates the effect of training the NN of the proposed approach on different N-k security levels and testing against k outages.

In the case of the IEEE 39-bus system, for the N-1 case our proposed deterministic approach is 15× faster than the H baseline, with a slight increase in detected post-contingency violations and a 1.45% cost increase. For the N-2 case, our approach is 21× faster than the H baseline, with a slight increase in detected post-contingency violations and a 1.22% cost increase. For the N-3 case, our approach is 21× faster than the H baseline, notably reducing violating post-contingency cases by 0.87% in absolute terms and a 0.10% cost increase. Fig. 6 (a) reveals that the model trained on N-3 contingencies is comparable to N-4 and N-5 models, primarily due to a notable number of islanding cases in N-4 and N-5 shown in Table I.

In the case of the IEEE 118-bus system, for the N-1 case our proposed deterministic approach is 76× faster than the H baseline, with a slight increase in detected post-contingency violations and a 2.88% cost increase. For the N-2 case, our approach is 165× faster than the H baseline, achieving a 6.41% cost improvement with a small increase in violating post-contingency cases. For the N-3 case, our approach outperforms all baselines, notably reducing violating post-contingency cases by 2.10% in absolute terms. Fig. 6 (b) reveals that the N-3 trained model exhibits enhanced security against k $=\{1, 2\}$ outages compared to models specifically trained for those levels. Furthermore, the N-3 trained model achieves a dispatch cost similar to the N-2 trained model and only incurs 2.5% cost increase compared to the N-1 trained model. This demonstrates that for a marginal cost increase, the N-3 model significantly enhances overall system security.

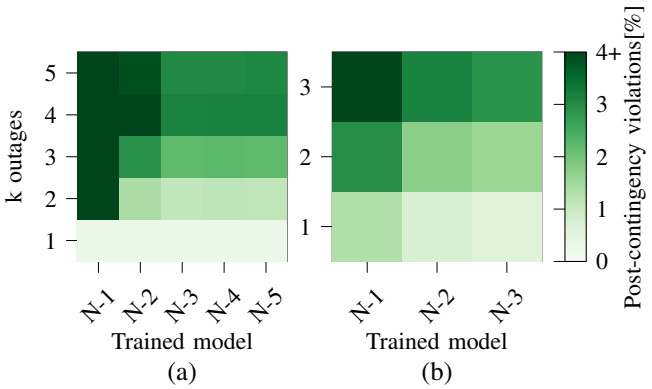

Fig. 6: Evaluating the percentage of violating post-contingency cases for k outages, for models trained on specific (N-k) levels of security for (a) the 39-bus system, (b) the 118-bus system.

TABLE V: LOLE and EENS comparison probabilistic approach

	LOLE [h/y]	EENS [MWh/y]
N-1 SCOPF	9.94	13.20
N-2 SCOPF	1.08	1.85
Probabilistic approach	0.99	1.40

D. Probabilistic N-k Security Assessment

This case study assesses N-k SCOPFs with probability considerations. The NN of the proposed approach is trained based on the probabilistic loss function in (17), which considers the contingency probabilities for $k = \{1, 2, 3\}$. Table V and and Table VI compare the probabilistic approach with the baseline N-1 and N-2 SCOPF baselines. The N-1 SCOPF finds the cheapest and least secure dispatch, lacking security against N-2 and N-3 contingencies. The proposed probabilistic approach demonstrates strong performance, showing slightly lower LOLE and EENS values compared to the N-2 baseline, with only a 1.86% increase in dispatch cost and a $17\times$ speedup.

E. Operational Resilience to HILP Events

To test the resilience of the proposed probabilistic approach against HILP events, a case study is performed where the contingency probabilities suddenly change due to an earthquake.

The earthquake simulation is represented using three concentric circles. Within the innermost circle, the individual probabilities of the lines increase significantly, while in the outermost circle, the individual probabilities show the least increase. The probabilities for $k = \{2, 3\}$ are recomputed after the earthquake. The probabilistic approach that is trained in the previous section is used. Table VII compares the proposed approach against the baseline in terms of reliability indicators after the earthquake. The proposed approach and the baseline exhibit similar LOLE values, but the proposed approach achieves a 21.5% reduction in EENS, highlighting its enhanced operational resilience against HILP events.

F. Discussion

We consider this approach as a first step toward using deep learning to enhance resiliency and reliability. The memory reduction method efficiently reduced the computational graph by 100-fold and achieved $8\times$ speedup in the backward pass. Our deterministic approach is $21\times$ and $165\times$ faster than the baselines in the 39-bus and 118-bus systems, respectively, while remaining competitive in terms of post-contingency violations. However, considering all N-k line contingencies in real-practice could lead to over conservative solutions and operational costs. In this respect, the proposed probabilistic approach offers a balanced solution by considering all line contingencies with their respective probability. This approach effectively reduces the LOLE and EENS, particularly in the case of HILP event. This approach requires models of the likelihood of simultaneous equipment failures. While some weather triggered events may be modelled through joint probability distributions, others require more deterministic models. For example, cascading failures may not follow joint distributions, but deterministic sequences defined through protection systems.

TABLE VI: Computational time and dispatch cost between probabilistic approach and baselines

	N-1 SCOPF	N-2 SCOPF	Probabilistic approach
Time [s]	61	139	8
Cost [$]	1968	2038	2076

TABLE VII: LOLE and EENS after HILP probabilistic approach

	After HILP	
	LOLE [h/y]	EENS [MWh/y]
N-2 SCOPF	13.37	28.28
Probabilistic approach	13.75	22.20

Nevertheless, some limitations of the approach should be mentioned. While our approach penalizes post-contingency flow violations through the loss function and satisfies PF constraints during normal operations via the feasibility restoration layer, enforcing flow limits in post-contingency cases remains a challenge. Relying solely on ML poses security risks, as in our previous work [30], hence restoration or considering these risks remains important. The proposed approach uses NNs to predict the dispatches for a fixed network topology. However, the network topology is frequently changing, which also changes the LODFs, necessitating NN retraining. Currently, the proposed approach only considers one objective: dispatch cost. In reality, system operators have to consider multiple objectives. While multiple objectives can easily be combined in this workflow, once trained, the objective itself cannot be changed without retraining. The proposed approach finds a preventive dispatch that minimizes flow limit violations of N-k line outages. The DC PF approximation enables fast evaluation of line thermal limits using LODFs. However, relying on DC approximation can result in voltage limit violations. AC formulations may be investigated in the feasibility restoration layer to ensure physical feasibility.

V. Conclusion

This paper proposes a constraint-driven learning approach to approximate N-k SCOPF optimization considering all line contingencies. The approach uses LODFs and eliminates the need for labeled training samples. The proposed memory reduction method significantly reduces the computational graph and achieves 8× speedup. The proposed deterministic approach is 21× and 165× faster than the baselines in the 39-bus and 118-bus systems, respectively. The approach effectively identifies violating post-contingency cases, particularly for k>2 with minimum optimally loss. When k increases, the baseline's approaches prioritize the identification of feasible solutions over achieving optimality in terms of dispatch cost, particularly as the feasible region becomes smaller. Within this set of case studies, the proposed probabilistic approach found a reliable and optimal dispatch for real-time operation and showed increased resilience against HILP events. The proposed approach enables, for the first time, a ML approximation of the N-k SCOPF that efficiently considers all line contingencies, offering both computational efficiency and security enhancements. This work will be strengthened by future work on developing topology adaptive ML approximations with feasibility guarantees, AC PF modeling, and further memory reduction methods for larger k values. Future research can also investigate generator outages using PTDFs, and corrective actions to improve system security against these important outages.

References

[1] Florin Capitanescu. Are we prepared against blackouts during the energy transition?: Probabilistic risk-based decision making encompassing jointly security and resilience. *IEEE Power and Energy Magazine*, 21(3):77–86, 2023.

[2] Mathaios Panteli and Pierluigi Mancarella. Modeling and evaluating the resilience of critical electrical power infrastructure to extreme weather events. *IEEE Systems Journal*, 11(3):1733–1742, 2015.

[3] Mingguo Hong, Xiaochuan Luo, Slava Maslennikov, and Late Eugene Litvinov. Enhancement of operational resilience with the online weather look-ahead study. In *2021 IEEE Power & Energy Society General Meeting (PESGM)*, pages 01–05. IEEE, 2021.

[4] Narayan Bhusal, Michael Abdelmalak, Md Kamruzzaman, and Mohammed Benidris. Power system resilience: Current practices, challenges, and future directions. *IEEE Access*, 8:18064–18086, 2020.

[5] Emanuele Ciapessoni, Diego Cirio, Gerd Kjølle, Stefano Massucco, Andrea Pitto, and Marino Sforna. Probabilistic risk-based security assessment of power systems considering incumbent threats and uncertainties. *IEEE Transactions on Smart Grid*, 7(6):2890–2903, 2016.

[6] Sanja Cvijic and Jinjun Xiong. Security constrained unit commitment and economic dispatch through benders decomposition: A comparative study. In *2011 IEEE Power and Energy Society General Meeting*, pages 1–8. IEEE, 2011.

[7] Qin Wang, James D McCalley, Tongxin Zheng, and Eugene Litvinov. Solving corrective risk-based security-constrained optimal power flow with lagrangian relaxation and benders decomposition. *International Journal of Electrical Power & Energy Systems*, 75:255–264, 2016.

[8] J.F. Benders. Partitioning procedures for solving mixed-variables programming problems '. *Numerische mathematik*, 4(1):238–252, 1962.

[9] Alexandre Velloso, Pascal Van Hentenryck, and Emma S Johnson. An exact and scalable problem decomposition for security-constrained optimal power flow. *Electric Power Systems Research*, 195:106677, 2021.

[10] Alinson Santos Xavier, Feng Qiu, Fengyu Wang, and Prakash R Thimmapuram. Transmission constraint filtering in large-scale security-constrained unit commitment. *IEEE Transactions on Power Systems*, 34(3):2457–2460, 2019.

[11] Diego A Tejada-Arango, Pedro Sánchez-Martın, and Andres Ramos. Security constrained unit commitment using line outage distribution factors. *IEEE Transactions on power systems*, 33(1):329–337, 2017.

[12] P Kaplunovich and K Turitsyn. Fast and reliable screening of n-2 contingencies. *IEEE Transactions on Power Systems*, 31(6):4243–4252, 2016.

[13] Liping Huang, Chun Sing Lai, Zhuoli Zhao, Guangya Yang, Bang Zhong, and Loi Lei Lai. Robust nk security-constrained optimal power flow incorporating preventive and corrective generation dispatch to improve power system reliability. *CSEE Journal of Power and Energy Systems*, 2022.

[14] Xiang Pan, Tianyu Zhao, Minghua Chen, and Shengyu Zhang. Deepopf: A deep neural network approach for security-constrained dc optimal power flow. *IEEE Transactions on Power Systems*, 36(3):1725–1735, 2020.

[15] Sidhant Misra, Line Roald, and Yeesian Ng. Learning for constrained optimization: Identifying optimal active constraint sets. *INFORMS Journal on Computing*, 34(1):463–480, 2022.

[16] Alexandre Velloso and Pascal Van Hentenryck. Combining deep learning and optimization for preventive security-constrained dc optimal power flow. *IEEE Transactions on Power Systems*, 36(4):3618–3628, 2021.

[17] Priya Donti, Aayushya Agarwal, Neeraj Vijay Bedmutha, Larry Pileggi, and J Zico Kolter. Adversarially robust learning for security-constrained optimal power flow. *Advances in Neural Information Processing Systems*, 34:28677–28689, 2021.

[18] James Kotary, Ferdinando Fioretto, Pascal Van Hentenryck, and Bryan Wilder. End-to-end constrained optimization learning: A survey. *CoRR abs/2103.16378*, 2021.

[19] Priya L. Donti, David Rolnick, and J Zico Kolter. DC3: A learning method for optimization with hard constraints. In *International Conference on Learning Representations*, 2021.

[20] Stephen Frank, Ingrida Steponavice, and Steffen Rebennack. Optimal power flow: A bibliographic survey i: Formulations and deterministic methods. *Energy systems*, 3:221–258, 2012.

[21] Laurence A Clarfeld, Paul DH Hines, Eric M Hernandez, and Margaret J Eppstein. Risk of cascading blackouts given correlated component outages. *IEEE Transactions on Network Science and Engineering*, 7(3):1133–1144, 2019.

[22] Laurence A Clarfeld, Margaret J Eppstein, Paul DH Hines, and Eric M Hernandez. Assessing risk from cascading blackouts given correlated component failures. In *2018 Power Systems Computation Conference (PSCC)*, pages 1–7. IEEE, 2018.

[23] A. Agrawal, B. Amos, S. Barratt, S. Boyd, S. Diamond, and Z. Kolter. Differentiable convex optimization layers. In *Advances in Neural Information Processing Systems*, 2019.

[24] Teoman Guler, George Gross, and Minghai Liu. Generalized line outage distribution factors. *IEEE Transactions on Power systems*, 22(2):879–881, 2007.

[25] Bart W Tuinema, JL Rueda Torres, Alexandru I Stefanov, Francisco M Gonzalez-Longatt, and MA van der Meijden. *Probabilistic Reliability Analysis of Power Systems*. Springer, 2020.

[26] Maedeh Mahzarnia, Mohsen Parsa Moghaddam, Payam Teimourzadeh Baboli, and Pierluigi Siano. A review of the measures to enhance power systems resilience. *IEEE Systems Journal*, 14(3):4059–4070, 2020.

[27] Sogol Babaeinejadsarookolaee, Adam Birchfield, Richard D Christie, Carleton Coffrin, Christopher DeMarco, Ruisheng Diao, Michael Ferris, Stephane Fliscounakis, Scott Greene, Renke Huang, et al. The power grid library for benchmarking ac optimal power flow algorithms. *arXiv preprint arXiv:1908.02788*, 2019.

[28] Delft High Performance Computing Centre (DHPC). DelftBlue Supercomputer (Phase 1). https://www.tudelft.nl/dhpc/ark:/44463/DelftBluePhase1, 2022.

[29] Bastien Giraud, Ali Rajaei, and Jochen Cremer. Implementation of constraint-driven deep learning for N-k security constrained optimal power flow. [Online]. Available: https://github.com/TU-Delft-AI-Energy-Lab/Constraint-Driven-SCOPF, 2024.

[30] Jochen L Cremer and Goran Strbac. A machine-learning based probabilistic perspective on dynamic security assessment. *International Journal of Electrical Power & Energy Systems*, 128:106571, 2021.