Supplementary material

Shuo Yang, *Student Member, IEEE,* Zhengshuo Li, *Senior Member, IEEE*

I. Select Appropriate Measurements for (3)

Below we explain how to select appropriate measurements for (3).

●*Determine the required number of samples for (3)*

We introduce the concept of Rademacher Complexity to characterize the generalization capability of the model. First, we consider a predefined set of linear mapping set $\mathcal{F}$, in paper, it refers to the constraints (5). For the training dataset $\mathcal{Z} = \{z_1, z_2, \dots, z_n\}$, where z_i refers to the measurement in paper, all mappings $f_j \in \mathcal{F}$ satisfy the constraint $f_j\ z_i\ = a_j^T z_i + b_j \le 0$ for all training samples:

The training error $\widehat{R}_S(\mathcal{F})$ can be expressed as:

$$\widehat{R}_S\ \mathcal{F}\ = \frac{1}{n}\sum_{i=1}^{n}\sum\nolimits_j f_j\ z_i\ \le 0. \tag{R1}$$

Assume that the generalization error on the test set is $R\ \mathcal{F}\ $. Here, it can also be interpreted as the probability of constraint violation on the test set. Through the Rademacher complexity, we can obtain:

$$R\ \mathcal{F}\ \le \widehat{R}_S\ \mathcal{F}\ + 2\mathfrak{R}_n\ \mathcal{F}\ + \sqrt{\frac{\log\ (1/\delta)}{2n}}, \tag{R2}$$

where $\mathfrak{R}_n\ \mathcal{F}\ $ is the Rademacher complexity of $\mathcal{F}$, and δ represents the confidence parameter (e.g., $\delta = 0.05$ represents the confident level is 95%). Combining (R1) and (R2), we can get (R3):

$$R\ \mathcal{F}\ \le \frac{1}{n}\sum_{i=1}^{n}\sum\nolimits_j f_j\ z_i\ + 2\mathfrak{R}_n\ \mathcal{F}\ + \sqrt{\frac{\log\ 1/\delta}{2n}}. \tag{R3}$$

For $\mathcal{F}$ contains linear functions, its Rademacher complexity is related to $\|a_j\|_2$ and $\|b_j\|_2$, and we can get:

$$\mathfrak{R}_n\ \mathcal{F}\ \le \frac{\max\{\|a_j\|_2\}\max\{\|b_j\|_2\}}{\sqrt{n}}. \tag{R4}$$

To ensure that the probability of constraint violation $R\ \mathcal{F}\ < \epsilon$ on the test set, we need to ensure that:

$$\frac{1}{n}\sum_{i=1}^{n}\sum\nolimits_j f_j\ z_i\ + 2\frac{\max\{\|a_j\|_2\}\max\{\|b_j\|_2\}}{\sqrt{n}} + \sqrt{\frac{\log\ 1/\delta}{2n}} \le \epsilon. \tag{R5}$$

In summary, the number of the required samples for LV DN model recovery can be determined by (R5).

●*The process of selecting proper input data*

First, there are multiple channels for collecting measurement data in LV DNs. Smart meters provide data on customer-side voltage, current, and power factor; distribution transformer monitoring terminals and feeder terminals collect transformer voltage and power, as well as line voltage, current, and power data; distributed PV inverters monitor PV generation power and point-of-connection voltage; additionally, smart sockets, electric vehicle charging piles, and similar devices can also serve as data sources, providing node-level voltage and power information.

Second, data cleaning is necessary to address issues such as outliers and missing values, thereby ensuring data quality. For example, measurement data may be missing due to sensor failures or communication interruptions, in which case interpolation is typically applied to fill in the missing values; measurements identified as obvious bad data, such as sudden abnormal spikes, should be excluded.

Third, data selection and construction are required. Taking the data needed for the LV DN model recovery in this paper as an example, the method requires nodal voltage magnitudes, nodal net active power, and nodal net reactive power. If the aforementioned data cannot be directly obtained, one possible approach to address this issue is to reconstruct them from the available measurements.

Fourth, data normalization and standardization are essential. Data with different physical dimensions should be normalized based on selected base values to unify the scales of variables with differing units or magnitudes. Additionally, the data should be categorized according to their characteristics.

The above process can be represented by Fig. R-1. Following the above processing steps, we obtain the voltage magnitudes, net active power, and net reactive power at each node that can be directly used in the LV DN model recovery.

Fig. R-1. The measurement data processing steps.

II. Supplementary Descriptions of the Model.

In this part the supplementary descriptions about regulating devices and three systems are given.

A. Supplementary Description about the Equation (8)

For capacitor shunt at node i, according to the symbol definition in the manuscript, the output $\boldsymbol{Q}_{C,i,t}^{3\phi}$ for a capacitor shunt can be modeled as follows:

$$\begin{cases} \boldsymbol{Q}_{C,i}^{3\phi} = \sum\nolimits_K q_{Ck}\boldsymbol{z}_{c,i,k} \circ \boldsymbol{U}_i^{3\phi}, \\ \sum\nolimits_K \boldsymbol{z}_{c,i,k} \le \mathbf{1}, \boldsymbol{z}_{c,i,k} \in \{0,1\}^{\,\boldsymbol{z}_{c,i,k}}. \end{cases} \tag{R6}$$

Here, the McCormick relaxation is leveraged to linearize the bilinear term $\boldsymbol{z}_{c,i,k} \circ \boldsymbol{U}_i^{3\phi}$. Each bilinear term in the objective function is replaced by an auxiliary variable $\boldsymbol{\lambda}_{c,i,k}$, and the variable $\boldsymbol{\lambda}_{c,i,k}$ is subjected to McCormick's constraints, and thus (R6) is equivalently replaced by (R7) (notice that the element in $\boldsymbol{z}_{c,i,k}$ equal to 1 implies the corresponding element in $\boldsymbol{\lambda}_{c,i,k}$ equal to U_i^{ϕ}, and the element in $\boldsymbol{z}_{c,i,k}$ equal to 0 implies the corresponding element in $\boldsymbol{\lambda}_{c,i,k}$ equal to 0):

$$\begin{cases} \boldsymbol{Q}_{C,i}^{3\phi} = \sum\nolimits_K q_{Ck}\boldsymbol{\lambda}_{c,i,k}, \\ \begin{cases} \boldsymbol{U}_i^{3\phi} - U_i^{max}(\mathbf{1} - \boldsymbol{z}_{c,i,k}) \le \boldsymbol{\lambda}_{c,i,k}, k = 1,2,\dots,K, \\ \boldsymbol{U}_i^{3\phi} - U_i^{min}(\mathbf{1} - \boldsymbol{z}_{c,i,k}) \ge \boldsymbol{\lambda}_{c,i,k}, k = 1,2,\dots,K, \\ U_i^{min}\boldsymbol{z}_{c,i,t,k} \le \boldsymbol{\lambda}_{c,i,k} \le U_i^{max}\boldsymbol{z}_{c,i,k}, \end{cases} \\ \sum\nolimits_K \boldsymbol{z}_{c,i,k} \le \mathbf{1}, \boldsymbol{z}_{c,i,k} \in \{0,1\}^{\,\boldsymbol{z}_{c,i,k}}. \end{cases} \tag{R7}$$

B. Supplementary Description about the Equation (15d)-(15f)

The literature [1] provides a linearized three-phase power flow equation. Building upon this, we derive equations (15d)–(15f) in the manuscript, i.e., without accounting for the variations introduced by the OLTC, obtain the more general forms (R8a) and (R8b):

$$\boldsymbol{U}_i^{3\phi} - \boldsymbol{U}_j^{3\phi} = 2(\overline{\boldsymbol{R}}_{ij}\boldsymbol{P}_{ij}^{3\phi} + \overline{\boldsymbol{X}}_{ij}\boldsymbol{Q}_{ij}^{3\phi}) \tag{R8a}$$

$$\begin{cases} \bar{\boldsymbol{R}}_{ij} = \mathcal{Re}(\boldsymbol{\alpha}\boldsymbol{\alpha}^H) \circ \boldsymbol{R}_{ij} + \mathcal{Im}(\boldsymbol{\alpha}\boldsymbol{\alpha}^H) \circ \boldsymbol{X}_{ij}, \\ \bar{\boldsymbol{X}}_{ij} = \mathcal{Re}(\boldsymbol{\alpha}\boldsymbol{\alpha}^H) \circ \boldsymbol{X}_{ij} + \mathcal{Im}(\boldsymbol{\alpha}\boldsymbol{\alpha}^H) \circ \boldsymbol{R}_{ij}, \\ \boldsymbol{\alpha} = \left[1, e^{-j2\pi/3}, e^{j2\pi/3}\right]^T, \end{cases} \tag{R8b}$$

Based on [1], we can obtain the following equations:

$$\begin{aligned} & diag\left(\begin{bmatrix} v_i^a{}^2 \\ v_i^b{}^2 \\ v_i^c{}^2 \end{bmatrix} - \begin{bmatrix} (v_j^a)^2 \\ (v_j^b)^2 \\ (v_j^c)^2 \end{bmatrix}\right) \\ = \boldsymbol{\alpha}\boldsymbol{\alpha}^H & \left(diag\begin{pmatrix} P_{ij}^a + jQ_{ij}^a \\ P_{ij}^b + jQ_{ij}^b \\ P_{ij}^c + jQ_{ij}^c \end{pmatrix}\right) \begin{bmatrix} z_{ij}^{aa} & z_{ij}^{ab} & z_{ij}^{ac} \\ z_{ij}^{ba} & z_{ij}^{bb} & z_{ij}^{bc} \\ z_{ij}^{ca} & z_{ij}^{cb} & z_{ij}^{cc} \end{bmatrix}^H \\ + \begin{bmatrix} z_{ij}^{aa} & z_{ij}^{ab} & z_{ij}^{ac} \\ z_{ij}^{ba} & z_{ij}^{bb} & z_{ij}^{bc} \\ z_{ij}^{ca} & z_{ij}^{cb} & z_{ij}^{cc} \end{bmatrix} & \left(\boldsymbol{\alpha}\boldsymbol{\alpha}^H \left(diag\begin{pmatrix} P_{ij}^a + jQ_{ij}^a \\ P_{ij}^b + jQ_{ij}^b \\ P_{ij}^c + jQ_{ij}^c \end{pmatrix}\right)\right)^H, \end{aligned} \tag{R9}$$

where $z_{ij} = r_{ij} + jx_{ij}$. Notice that $\boldsymbol{\alpha}\boldsymbol{\alpha}^H{}^H = \boldsymbol{\alpha}\boldsymbol{\alpha}^H$. Substituting these into (R9) yields (R10):

$$\begin{aligned} & diag(\boldsymbol{U}_i^{3\phi} - \boldsymbol{U}_j^{3\phi}) \\ = 2(\mathcal{R}\boldsymbol{e} & \left(\boldsymbol{\alpha}\boldsymbol{\alpha}^H diag\begin{pmatrix} P_{ij}^a + jQ_{ij}^a \\ P_{ij}^b + jQ_{ij}^b \\ P_{ij}^c + jQ_{ij}^c \end{pmatrix}\right) \begin{bmatrix} r_{ij}^{aa} & r_{ij}^{ab} & r_{ij}^{ac} \\ r_{ij}^{ba} & r_{ij}^{bb} & r_{ij}^{bc} \\ r_{ij}^{ca} & r_{ij}^{cb} & r_{ij}^{cc} \end{bmatrix} \\ + \mathcal{I}\boldsymbol{m} & \left(\boldsymbol{\alpha}\boldsymbol{\alpha}^H diag\begin{pmatrix} P_{ij}^a + jQ_{ij}^a \\ P_{ij}^b + jQ_{ij}^b \\ P_{ij}^c + jQ_{ij}^c \end{pmatrix}\right) \begin{bmatrix} x_{ij}^{aa} & x_{ij}^{ab} & x_{ij}^{ac} \\ x_{ij}^{ba} & x_{ij}^{bb} & x_{ij}^{bc} \\ x_{ij}^{ca} & x_{ij}^{cb} & x_{ij}^{cc} \end{bmatrix}) \\ = 2(\mathcal{R}\boldsymbol{e}\ & \boldsymbol{\alpha}\boldsymbol{\alpha}^H\ diag(\boldsymbol{P}_{ij}^{3\phi})\boldsymbol{R}_{ij} + \mathcal{I}\boldsymbol{m}\ \boldsymbol{\alpha}\boldsymbol{\alpha}^H\ diag(\boldsymbol{P}_{ij}^{3\phi})\boldsymbol{X}_{ij} \\ + \mathcal{R}\boldsymbol{e}\ & \boldsymbol{\alpha}\boldsymbol{\alpha}^H\ diag(\boldsymbol{Q}_{ij}^{3\phi})\boldsymbol{X}_{ij} - \mathcal{I}\boldsymbol{m}\ \boldsymbol{\alpha}\boldsymbol{\alpha}^H\ diag(\boldsymbol{Q}_{ij}^{3\phi})\boldsymbol{R}_{ij}), \end{aligned} \tag{R10}$$

where $\boldsymbol{U}_i^{3\phi} = \begin{bmatrix} v_i^a{}^2 \\ v_i^b{}^2 \\ v_i^c{}^2 \end{bmatrix}$, $\boldsymbol{P}_{ij}^{3\phi} + j\boldsymbol{Q}_{ij}^{3\phi} = \begin{pmatrix} P_{ij}^a + jQ_{ij}^a \\ P_{ij}^b + jQ_{ij}^b \\ P_{ij}^c + jQ_{ij}^c \end{pmatrix}$ and $\boldsymbol{R}_{ij} + j\boldsymbol{X}_{ij} = \begin{bmatrix} r_{ij}^{aa} & r_{ij}^{ab} & r_{ij}^{ac} \\ r_{ij}^{ba} & r_{ij}^{bb} & r_{ij}^{bc} \\ r_{ij}^{ca} & r_{ij}^{cb} & r_{ij}^{cc} \end{bmatrix} + j\begin{bmatrix} x_{ij}^{aa} & x_{ij}^{ab} & x_{ij}^{ac} \\ x_{ij}^{ba} & x_{ij}^{bb} & x_{ij}^{bc} \\ x_{ij}^{ca} & x_{ij}^{cb} & x_{ij}^{cc} \end{bmatrix}$. After eliminating the diagonal elements, (R10) can be expressed as (R8a) and (R8b), and further be turned into (15d)-(15f) in the manuscript.

When the network has a missing phase, the impedance of two-phase or single-phase lines can still be represented by a 3×3 matrix according to [2], with the rows and columns corresponding to the missing phase set to zero. Therefore, the model of the DN in this paper is also applicable to networks with missing phases.

C. Supplementary Description about the M4L13 IMLV DN

The diagram of the M4L13 LV DN is shown in Fig. R-2.

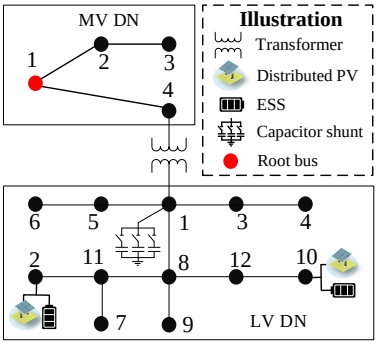

Fig. R-2. The diagram of the M4L13 LV DN.

The line impedance of the MV DN in M4L13 IMLV DN is shown in Table R-I (It is assumed that the lines of the MV DN in IMLV DN are three-phase balanced):

TABLE R-I
THE LINE IMPEDANCE OF THE MV DN IN IMLV DN (p.u.)

From	To	Resistance	Reactance
1	2	0.0074	0.0172
2	3	0.0101	0.0194
1	4	0.0030	0.0060

The line impedance of the LV DN in M4L13 IMLV DN can be found in [3]. Values of relevant regulating devices in M4L13 IMLV DN can be found in Table R-II.

TABLE R-II
VALUES OF RELEVANT PARAMETERS (p.u.)

Name	Values	Name	Values	Name	Values
$v_{LV}^{min}/v_{LV}^{max}$	0.9/1.1	$P_{PV,i}^{max}$	0.1	$S_{PV,i}^{max}$	0.2
$\alpha_{SOP,i}$	0.01	$\underline{\phi}/\overline{\phi}$	-0.9/0.9	$S_{SOP,i}^{max}$	1
$q_{C1/2/3/4}$	0.1/0.2/0.3/0.4	P_D/P_C	0.1/0.1	η_D/η_C	1/1
$a_{OLTC}^{min}/a_{OLTC}^{max}$	0.98/1.02	τ^{min}/τ^{max}	0.95/1.05	$\boldsymbol{Y}_1/\boldsymbol{Y}_2/\boldsymbol{Y}_3/\boldsymbol{Y}_4$	**0**
$\boldsymbol{EA}_1/\boldsymbol{A}_2$	$diag(\boldsymbol{1})$	$v_{MV}^{min}/v_{MV}^{max}$	0.95/1.05		

D. Supplementary Description about the M162L360 IMLV DN

The diagram of the M162L360 IMLV DN is shown in Fig. R-3.

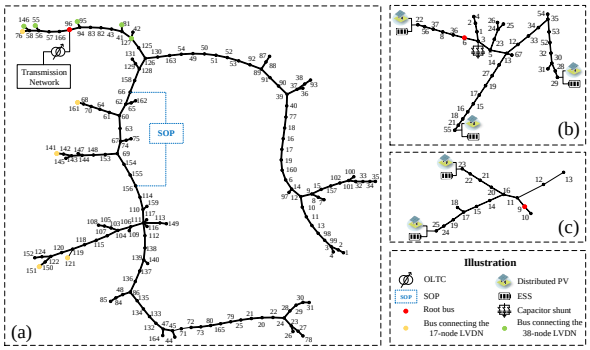

Fig. R-3. (a) The diagram of the modified 162-bus MV DN. (b) The diagram of the modified 17-node LV DN. (c) The diagram of the modified 38-node LV DN.

The line impedance of the M162L360 IMLV DN can be found in 0. Values of relevant regulating devices in M4L13 IMLV DN can be found in Table R-III.

TABLE R-III
VALUES OF RELEVANT PARAMETERS (p.u.)

Name	Values	Name	Values	Name	Values
$v_{LV}^{min}/v_{LV}^{max}$	0.9/1.1	$P_{PV,i}^{max}$	0.1	$S_{PV,i}^{max}$	2
$\alpha_{SOP,i}$	0.01	$\underline{\phi}/\overline{\phi}$	-0.9/0.9	$S_{SOP,i}^{max}$	2
$q_{C1/2/3/4}$	0.1/0.2/0.3/0.4	P_D/P_C	0.1/0.1	η_D/η_C	1/1
$a_{OLTC}^{min}/a_{OLTC}^{max}$	0.98/1.02	τ^{min}/τ^{max}	0.95/1.05	$\boldsymbol{Y}_1/\boldsymbol{Y}_2/\boldsymbol{Y}_3/\boldsymbol{Y}_4$	**0**
$\boldsymbol{EA}_1/\boldsymbol{A}_2$	$diag(\boldsymbol{1})$	$v_{MV}^{min}/v_{MV}^{max}$	0.95/1.05		

III. SUPPLEMENTARY DESCRIPTIONS OF THE UNCERTAINTY MODELING.

When considering only the impact of solar irradiance on the output of distributed PV, many studies have employed the Beta distribution to model solar irradiance [4][5]. The shape of the Beta distribution is determined by two shape parameters, α and β. Depending on the values of these parameters, the distribution can take various forms, such as a uniform distribution, a unimodal symmetric or asymmetric distribution, a U-shaped distribution, or a strictly increasing or decreasing distribution. When modeling solar irradiance S using the Beta distribution, its probability density function is given by equation (R11):

$$f\ S\ =\frac{\Gamma\ \alpha+\beta}{\Gamma\ \alpha\ \Gamma\ \alpha}\left(\frac{S}{S_{max}}\right)^{\alpha-1}\left(1-\frac{S}{S_{max}}\right)^{\beta-1}, \tag{R11}$$

where $\Gamma\ \cdot$ is the gamma function, and S_{max} is the maximum irradiance intensity of the distributed PV generation unit. Based on the solar irradiance, the output power of the distributed PV can be expressed as:

$$P_{PV,i}^{\phi}=\begin{cases}P_{PV,i}^{max}\dfrac{S}{S_{max}^{r}}, & S\le S_{max}^{r}\\ P_{PV,i}^{max}, & S>S_{max}^{r}\end{cases},\phi=\{a,b,c\}. \tag{R12}$$

S_{max}^{r} denotes the irradiance level at which the distributed PV reaches its maximum output power. When the irradiance exceeds this threshold, the output power of the distributed PV no longer increases.

A potential issue with modeling distributed PV output using the Beta distribution is that the constraints are difficult to directly transform into a form that is easy to solve, unlike the case when using the GMM. At this point, the inequality constraints affected by distributed PV output—such as output constraints and capacity limits for distributed PV, and voltage security constraints in low-voltage DNs—can be formulated as chance constraints and subsequently transformed into mixed-integer linear programming (MILP) model using the Sample Average Approximation method. As long as the number of sampling scenarios is sufficiently large, the SAA method can achieve a high accuracy. However, this also results in an excessively large equivalent MILP model, imposing a heavy computational burden.

IV. Supplementary Descriptions of the Improved AD Algorithm.

Although the commonly used AD method occasionally succeeds in solving the problem, it often tends to converge to local optima. Therefore, we introduced the following improvements to the traditional AD method:

1) Introduce random noise to perturb the current solution randomly at each iteration, increasing the probability of escaping from the current optimum and exploring potentially superior regions [6].

2) Accelerate the iterative process using the gradient descent method [7]. Additional explanations were provided regarding the optimization variables in the iterative process. When solving the subproblem (31) using improved AD method, $\boldsymbol{l}_{0,k}^{*}=\boldsymbol{l}_{0,k-1}^{*}\circ\ \mathbf{1}+\boldsymbol{\epsilon}\ +\epsilon_{GD}\times\left(\boldsymbol{\theta}_{2,k-1}^{T}\boldsymbol{I}_{l}\left(\widetilde{\underline{\boldsymbol{L}}}+\widetilde{\underline{\boldsymbol{L}}}\right)\right)$, and when solving (32), $\boldsymbol{\theta}_{1,k}^{*}=\boldsymbol{\theta}_{1,k-1}^{*}\circ\ \mathbf{1}+\boldsymbol{\epsilon}\ ,\boldsymbol{\theta}_{2,k}^{*}=\boldsymbol{\theta}_{2,k-1}^{*}\circ\ \mathbf{1}+\boldsymbol{\epsilon}\ +\ +\epsilon_{GD}\times\left(\mathbf{I}_{l}\left(\widetilde{\underline{\boldsymbol{L}}}+\widetilde{\underline{\boldsymbol{L}}}\right)\circ\boldsymbol{l}_{0,k-1}^{*}\right)^{T},\boldsymbol{\pi}_{k}^{*}=\boldsymbol{\pi}_{k-1}{}^{*}\circ(\mathbf{1}+\boldsymbol{\epsilon})$.

3) To avoid entrapment in local minima and ensure convergence to a global optimum, the AD algorithm typically necessitates iterative executions from multiple initial points. Leveraging the inherent properties of bilinear optimization problems, we introduce an approach for selecting the initial points. i) **Selecting the vertices of the feasible region as initial points**: In the AD method, each iteration involves solving a linear programming subproblem. According to the fundamental properties of linear programming, the optimal solution is guaranteed to reside at a vertex of the feasible region. Therefore, by choosing these vertices as initial points, the solving process can be significantly accelerated, thereby improving the overall convergence rate of the AD algorithm. ii) **Selecting singular vectors obtained from the singular value decomposition of the coefficient matrix of the bilinear terms as initial values**. The reasons are as follows:

Considering bilinear optimization problems with only unit-norm constraints:

$$\max_{\|x\|_2\le1,\|y\|_2\le1}x^{T}Ay. \tag{R14}$$

By leveraging the properties of matrix norms, we can obtain:

$$\max_{\|x\|_2\le1,\|y\|_2\le1}x^{T}Ay\le\|A\|_2\cdot\|x\|_2\cdot\|y\|_2\le\sigma_1\cdot\|u_1\|_2\cdot\|v_1\|_2, \tag{R15}$$

where σ_1 is the largest singular value of A, u_1 and v_1 are the left and right singular vectors of σ_1. When the optimization variable constraints are transformed from norm constraints to box constraints, the principal component decomposition may no longer yield globally optimal solutions under the new feasible set. Nevertheless, it can still offer high-quality initial points that effectively facilitate the acceleration of the AD algorithm's convergence.

We also test the effectiveness of the improved AD method on small-scale studies. We denote the dimensions of the bilinear terms in the test cases by (m,n), and compare the average computation time and the average number of iterations over 100 independent trials. The results are shown in Table R-IV.

TABLE R-IV Performance of the FI-C&CG Algorithm

Test System	Methods	Obtain the optimal solution?	The average computation time	The average number of iterations
(3,2)	AD	√	0.24s	2.21
	Improved AD	√	0.19s	2
(12,162)	AD	√	0.49s	3.17
	Improved AD	√	0.21s	2

It can be observed that the improved AD algorithm requires fewer iterations and achieves faster solution times.

IV. Supplementary Descriptions when Testing the Effectiveness of the Outlier-immune Two-sided CL Model.

The line parameters of the 13-node LV DN and the 38-node Swiss LV DN used for testing can be found in 0 and [8]. It is important to note that we consider the impact of line parameters and do not take into account the regulating devices connected to the system during the test. The load level of the 1000 training sets and 1500 test sets varies within the range of 0.5–1.5 times the original system loads. The original system loads for the two testing systems can be found in Table R-V and Table R-VI. In the 38-node Swiss LV DN system, the three-phase original loads are the same, and only one phase is provided in Table V. The power injection at the root node is set to 0.

TABLE R-V
The Original System Loads for the 13-node LV DN (p.u.)

Node	Active power	Reactive power	Node	Active power	Reactive power	Node	Active power	Reactive power
1a	0.0012	0.0006	5a	0.0024	0.0016	9a	0.0024	0.0014
1b	0.0012	0.0004	5b	0.0012	0.0002	9b	0.0040	0.0120
1c	0.0040	0.0020	5c	0.0012	0.0004	9c	0.0030	0.0014
2a	-0.0040	-0.0024	6a	0.0012	0.0004	10a	-0.0036	-0.0020
2b	-0.0036	-0.0016	6b	0.0018	0.0008	10b	-0.0168	-0.0080
2c	-0.0048	-0.0032	6c	0.0018	0.0008	10c	-0.0168	-0.0080

3a	0.0040	0.0020	7a	0.0018	0.0008	11a	0.0042	0.0020
3b	0.0012	0.0004	7b	0.0018	0.0008	11b	0.0012	0.0008
3c	0.0012	0.0004	7c	0.0018	0.0008	11c	0.0020	0.0012
4a	9e-4	6e-4	8a	0.0012	0.0005	12a	0.0024	0.0016
4b	0.0012	0.0007	8b	0.0012	0.0005	12b	0.0012	0.0004
4c	0.0012	0.0007	8c	0.0012	0.0004	12c	0.0040	0.0020

TABLE R-VI
THE ORIGINAL SYSTEM LOADS FOR THE 38-NODE SWISS LV DN (p.u.)

Node	Active power	Reactive power	Node	Active power	Reactive power	Node	Active power	Reactive power
1	4e-4	1e-4	14	5e-4	1e-4	27	1e-4	3e-5
2	8e-4	2e-4	15	2e-4	6e-5	28	1e-4	3e-5
3	8e-4	2e-4	16	5e-4	2e-4	29	7e-4	2e-4
4	2e-4	9e-4	17	-4e-3	-5e-4	30	5e-3	1e-3
5	4e-4	4e-4	18	7e-4	2e-4	31	0.11	0.03
6	0	0	19	5e-4	1e-4	32	4e-3	1e-3
7	4e-4	9e-4	20	5e-4	1e-4	33	2e-4	4e-5
8	1e-3	3e-4	21	1e-3	2e-4	34	3e-4	9e-5
9	1e-3	2e-4	22	1e-3	4e-4	35	2e-4	4e-5
10	1e-3	3e-4	23	7e-4	2e-4	36	-0.02	-2e-3
11	1e-3	3e-4	24	-5e-3	-5e-4	37	3e-4	8e-5
12	1e-3	2e-4	25	7e-4	2e-4	38	9e-4	2e-4
13	6e-4	1e-4	26	1e-4	3e-5			

V. VOLTAGE MAGNITUDE ERRORS UNDER DIFFERENT M_0.

TABLE R-VII
VOLTAGE MAGNITUDE ERRORS UNDER DIFFERENT M_0

	Errors/sample (%)					
	0		5%		10%	
	Mean	Max	Mean	Max	Mean	Max
Over-estimate	0.0023	0.0124	0.0027	0.0135	0.0027	0.0179
Under-estimate	0.0032	0.0122	0.0065	0.0140	0.0066	0.0143

VI. SUPPLEMENTARY DESCRIPTIONS OF THE LINE PARAMETERS OF THE THREE-PHASE 4-NODE TEST SYSTEM.

The system diagram of the 4-node test system is shown in Fig. R-4.

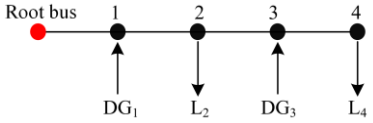

Fig. R-4. The system diagram of the 4-node test system.

The line impedance of the three-phase 4-node test system modified from can be found in Table R-VIII.

TABLE R-VIII
THE LINE IMPEDANCE OF THE 4-NODE TEST SYSTEM (p.u.)

Branch number	From	To	Phase	Resistance	Reactance
1	Root bus	1	a	0.07	0.064
			b	0.08	0.073
			c	0.07	0.064
2	1	2	a	0.07	0.064
			b	0.07	0.064
			c	0.07	0.064
3	2	3	a	0.10	0.066
			b	0.11	0.070
			c	0.12	0.073
4	3	4	a	0.10	0.066
			b	0.10	0.066
			c	0.12	0.073

VII. SUPPLEMENTARY DESCRIPTIONS OF THE TESTED THREE SCENARIOS.

The tested scenarios include:

Scenario 1: System power injection is set as 0 in M4L13 system and M162L360 system.

Scenario 2: Distributed PV units are connected to different single phases of the LV DN, with phase a connected at node 2 and phase c connected at node 10 in M4L13 system and random single phase in M162L360 system.

Scenario 3: The presence or absence of controllable devices in M4L13 system, including capacitor banks, SOPs, and ESS.

REFERENCES

[1] L. Gan and S.H. Low, "Convex relaxations and linear approximation for optimal power flow in multiphase radial network," in *18th Power Systems Computation Conference (PSCC)*, 2014.

[2] Kersting W H. "Distribution System Line Models" in Distribution system modeling and analysis. *Boca Raton: CRC Press LLC*, pp. 125, 2002.

[3] "IEEE13NodeSystem.dss." Accessed: Jun. 28, 2024. [Online]. Available: https://github.com/YYY-maker130/Assessment-of-the-Security-Region.git

[4] S. Zhang, H. Cheng, K. Li, N. Tai, D. Wang, and F. Li, "Multi-objective distributed generation planning in distribution network considering correlations among uncertainties," *Appl. Energy*, vol. 226, pp. 743–755, 2018.

[5] L. Guan, X. Chen, Y. Lv, Z. Tang, and Q. Zhao, "Probability model of PV generation for power system planning and its application," *Electric Power Automation Equipment*, vol. 37, no. 11, pp. 1-7, 2017.

[6] J. Liu, Y. Yuan, "Almost sure convergence rates analysis and saddle avoidance of stochastic gradient methods," *Journal of Machine Learning Research*, vol. 25, no. 271, pp. 1-40, 2024.

[7] W. Hao, "A gradient descent method for solving a system of nonlinear equations," *Appl. Math. Lett.*, vol. 112, 2021, Art. no. 106739.

[8] A. E. Oneto, F. Tettamanti, B. Gjorgiev and G. Sansavini, "Synthetic low- and medium-voltage grids for Switzerland," *Zenodo*, April 2025. doi: 10.5281/zenodo.15167589.