$$\boldsymbol{L} = \boldsymbol{MAPE}(\boldsymbol{C}_1(\boldsymbol{X}\cdot\boldsymbol{X}) + \boldsymbol{C}_2(\boldsymbol{X}\cdot\boldsymbol{X}),\ (\boldsymbol{C}_3\boldsymbol{X})\cdot(\boldsymbol{C}_4\boldsymbol{X})) \tag{16}$$

As expressed in (16), $\boldsymbol{X}$ represents the actual value of $[\hat{\boldsymbol{V}}_{\text{sin,pu}} \quad \hat{\boldsymbol{V}}_{\text{squ,pu}} \quad \hat{\boldsymbol{V}}_{\text{cos,pu}}]$ obtained in Fig. 2, and the physics loss is actually related to $(V_iV_j\cos\theta_{ij})^2 + (V_iV_j\sin\theta_{ij})^2 = V_i^2\cdot V_{\boldsymbol{j}}^2$ represented by (9) in the paper.

$$\boldsymbol{C}_1\left(\begin{bmatrix}\hat{\boldsymbol{V}}_{\text{sin}}\\ \hat{\boldsymbol{V}}_{\text{squ}}\\ \hat{\boldsymbol{V}}_{\text{cos}}\end{bmatrix}\bullet\begin{bmatrix}\hat{\boldsymbol{V}}_{\text{sin}}\\ \hat{\boldsymbol{V}}_{\text{squ}}\\ \hat{\boldsymbol{V}}_{\text{cos}}\end{bmatrix}\right) + \boldsymbol{C}_2\left(\begin{bmatrix}\hat{\boldsymbol{V}}_{\text{sin}}\\ \hat{\boldsymbol{V}}_{\text{squ}}\\ \hat{\boldsymbol{V}}_{\text{cos}}\end{bmatrix}\bullet\begin{bmatrix}\hat{\boldsymbol{V}}_{\text{sin}}\\ \hat{\boldsymbol{V}}_{\text{squ}}\\ \hat{\boldsymbol{V}}_{\text{cos}}\end{bmatrix}\right) = \left(\boldsymbol{C}_3\begin{bmatrix}\hat{\boldsymbol{V}}_{\text{sin}}\\ \hat{\boldsymbol{V}}_{\text{squ}}\\ \hat{\boldsymbol{V}}_{\text{cos}}\end{bmatrix}\right)\bullet\left(\boldsymbol{C}_4\begin{bmatrix}\hat{\boldsymbol{V}}_{\text{sin}}\\ \hat{\boldsymbol{V}}_{\text{squ}}\\ \hat{\boldsymbol{V}}_{\text{cos}}\end{bmatrix}\right) \tag{9}$$

Specifically, the variable correspondence is described as (1*)-(3*):

$$\boldsymbol{C}_1\left(\begin{bmatrix}\hat{\boldsymbol{V}}_{\text{sin}}\\ \hat{\boldsymbol{V}}_{\text{squ}}\\ \hat{\boldsymbol{V}}_{\text{cos}}\end{bmatrix}\bullet\begin{bmatrix}\hat{\boldsymbol{V}}_{\text{sin}}\\ \hat{\boldsymbol{V}}_{\text{squ}}\\ \hat{\boldsymbol{V}}_{\text{cos}}\end{bmatrix}\right) = [(V_iV_j\cos\theta_{ij})^2]_{m\times 1} \tag{1*}$$

$$\boldsymbol{C}_2\left(\begin{bmatrix}\hat{\boldsymbol{V}}_{\text{sin}}\\ \hat{\boldsymbol{V}}_{\text{squ}}\\ \hat{\boldsymbol{V}}_{\text{cos}}\end{bmatrix}\bullet\begin{bmatrix}\hat{\boldsymbol{V}}_{\text{sin}}\\ \hat{\boldsymbol{V}}_{\text{squ}}\\ \hat{\boldsymbol{V}}_{\text{cos}}\end{bmatrix}\right) = [(V_iV_j\sin\theta_{ij})^2]_{m\times 1} \tag{2*}$$

$$\left(\boldsymbol{C}_3\begin{bmatrix}\hat{\boldsymbol{V}}_{\text{sin}}\\ \hat{\boldsymbol{V}}_{\text{squ}}\\ \hat{\boldsymbol{V}}_{\text{cos}}\end{bmatrix}\right)\bullet\left(\boldsymbol{C}_4\begin{bmatrix}\hat{\boldsymbol{V}}_{\text{sin}}\\ \hat{\boldsymbol{V}}_{\text{squ}}\\ \hat{\boldsymbol{V}}_{\text{cos}}\end{bmatrix}\right) = [V_i^2\cdot V_{\boldsymbol{j}}^{\ 2}]_{m\times 1} \tag{3*}$$

To improve the training effect, the loss shown in (16) can also be further transformed. Let $\tilde{\boldsymbol{V}}_{\cos} = \sqrt{\left(\boldsymbol{C}_3\begin{bmatrix}\hat{\boldsymbol{V}}_{\text{sin}}\\ \hat{\boldsymbol{V}}_{\text{squ}}\\ \hat{\boldsymbol{V}}_{\text{cos}}\end{bmatrix}\right)\bullet\left(\boldsymbol{C}_4\begin{bmatrix}\hat{\boldsymbol{V}}_{\text{sin}}\\ \hat{\boldsymbol{V}}_{\text{squ}}\\ \hat{\boldsymbol{V}}_{\text{cos}}\end{bmatrix}\right) - \boldsymbol{C}_2\left(\begin{bmatrix}\hat{\boldsymbol{V}}_{\text{sin}}\\ \hat{\boldsymbol{V}}_{\text{squ}}\\ \hat{\boldsymbol{V}}_{\text{cos}}\end{bmatrix}\bullet\begin{bmatrix}\hat{\boldsymbol{V}}_{\text{sin}}\\ \hat{\boldsymbol{V}}_{\text{squ}}\\ \hat{\boldsymbol{V}}_{\text{cos}}\end{bmatrix}\right)}$, and $\hat{\boldsymbol{P}}$, $\hat{\boldsymbol{Q}}$ can be obtained based on (4*) :

$$\begin{bmatrix}\hat{\boldsymbol{P}}\\ \hat{\boldsymbol{Q}}\end{bmatrix} = \begin{bmatrix}\boldsymbol{B}^{\boldsymbol{p}} & \boldsymbol{G}^{\Lambda} & \boldsymbol{G}^{\boldsymbol{p}}\\ \boldsymbol{G}^{\boldsymbol{Q}} & \boldsymbol{B}^{\Lambda} & \boldsymbol{B}^{\boldsymbol{Q}}\end{bmatrix}\begin{bmatrix}\hat{\boldsymbol{V}}_{\text{sin}}\\ \hat{\boldsymbol{V}}_{\text{squ}}\\ \tilde{\boldsymbol{V}}_{\text{cos}}\end{bmatrix} \tag{4*}$$

Assigning $\boldsymbol{S}_{real} = [\boldsymbol{P} \quad \boldsymbol{Q}]$, $\boldsymbol{S}_{pred} = [\hat{\boldsymbol{P}} \quad \hat{\boldsymbol{Q}}]$, the improved loss function for network training can be given as (5*):

$$\boldsymbol{L} = \boldsymbol{MAPE}(\boldsymbol{S}_{real}, \boldsymbol{S}_{pred}) \tag{5*}$$