Input: $\mathbf{W^h}$, $\boldsymbol{\theta}^{\text{BN}}$, $\mathbf{U_W}$, $\mathbf{U_\theta}$, $(\mathbf{x}, \mathbf{y})$, m, η.
Output: $\mathbf{W^h}$, $\boldsymbol{\theta}^{\text{BN}}$, $\mathbf{U_W}$, $\mathbf{U_\theta}$.
1: $\mathbf{W^b} \leftarrow \text{Sign}(\mathbf{W^h})$ ▷Computing binary weights
2: $\hat{\mathbf{y}}$, cache $\leftarrow$ Forward$(\mathbf{x}, \mathbf{W^b}, \boldsymbol{\theta}^{\text{BN}})$ ▷Perform inference
3: $C \leftarrow \text{Cost}(\hat{\mathbf{y}}, \mathbf{y})$ ▷Compute mean loss over the batch
4: $(\partial_{\mathbf{W}} C, \partial_{\boldsymbol{\theta}} C) \leftarrow$ Backward$(C, \hat{\mathbf{y}}, \mathbf{W^b}, \boldsymbol{\theta}^{\text{BN}}, \text{cache})$
   ▷Cost gradients
5: $(\mathbf{U_W}, \mathbf{U_\theta}) \leftarrow \text{Adam}(\partial_{\mathbf{W}} C, \partial_{\boldsymbol{\theta}} C, \mathbf{U_W}, \mathbf{U_\theta})$
6: **for** W^h in $\mathbf{W^h}$ **do**
7:   **if** $U_W \cdot W^b > 0$ **then** ▷If U_W prescribes to decrease$|W^b|$
8:     $W^h \leftarrow W^h - \eta U_W \cdot f_{\text{meta}}(m, W^h)$ ▷Metaplastic update
9:   **else**
10:     $W^h \leftarrow W^h - \eta U_W$
11:   **end if**
12: **end for**
13: $\boldsymbol{\theta}^{\text{BN}} \leftarrow \boldsymbol{\theta}^{\text{BN}} - \eta \mathbf{U_\theta}$
14: **return** $\mathbf{W^h}$, $\boldsymbol{\theta}^{\text{BN}}$, $\mathbf{U_W}$, $\mathbf{U_\theta}$