① **Warming-up common FE**

Client i

$\mathcal{D}_i$

$\{\theta_i^f, \theta_{i,m=1}^h\}$

$\{\theta_j^f, \theta_{j,m=2}^h\}$

Client j

$\mathcal{D}_j$

$\theta_g^f \leftarrow \text{AGGREGATE}(\theta_k^f)$

② **Generative knowledge aggregation**

Client i θ_g^f

$\mathcal{D}_i$

$G(z_i|y_i, \mathbf{w}_i)$

Feature generator

Client j θ_g^f

$\mathcal{D}_j$

$G(z_i|y_i, \mathbf{w}_j)$

Feature generator

$\mathbf{w}_g \leftarrow \text{AGGREGATE}(\mathbf{w}_k)$

$G(z_i|y_i, \mathbf{w}_g)$

Feature generator

③ **Target network training**

Client i

$\mathbf{w_g}$

Generated features

Real samples

Client j

$\mathbf{w_g}$

Generated features

Real samples

Model-agnostic FL with feature-generative models