

Robot
State



$$V \in \mathbb{R}^{N_2 \times H}$$

Embedding

$$K \in \mathbb{R}^{N_2 \times H}$$

$$F_r \in \mathbb{R}^{N_1 \times H}$$

Latent Space
Generation

Latent Space
Generation

$$\bar{V} \in \mathbb{R}^{(N_1 \times N_2) \times H}$$

$$\bar{K} \in \mathbb{R}^{(N_1 \times N_2) \times H}$$

learnable mask

learnable mask

$$V^u \in \mathbb{R}^{N' \times H}$$

$$K^u \in \mathbb{R}^{N' \times H}$$

Latent Space
Compression

