Policy

Concrete(α, τ)

Predictor

| 0 |
| x_2 |
| 0 |
| x_4 |

$\pi(\,\cdot\,; \phi)$

$f(\,\cdot\,; \theta)$

$\hat{y}$

Update masked input

| x_1 |
| x_2 |
| ≈ 0 |
| x_4 |

Repeat for k selection steps