

Low task variance ($\mathcal{A}/g^{(2)}$ applied)

Accumulated reward

- BO-MRL (ours)
- Random initialization
- Optimal task-specific policies

0

1

2

3

4

Number of policy adaptation steps

1.5

1.0

0.5

0.0

