Adaptive Privacy-Preserving Hierarchical Coordination for Multi-Robot Systems

Improving Privacy–Utility Trade-offs and Byzantine Robustness

WiML Symposium @ ICML 2026

Maryam Fatima *Independent Researcher*

ICML International Conference On Machine Learning 2026

WiML Women in Machine Learning

1 MOTIVATION

Multi-robot systems must coordinate using **sensitive state information**.

Differential Privacy (DP) protects data but often hurts coordination.

Fixed DP budgets are either **too strict** (low utility) or **too loose** (weak privacy).

We propose a **rule-based, adaptive DP** scheme that allocates privacy budget **per robot, per step, and per context layer**.

Result: higher task success, better privacy–utility trade-offs, and improved robustness to Byzantine attacks.

2 HIERARCHICAL CONTEXTUAL ALIGNMENT FRAMEWORK

Each robot maintains three context embeddings at different scales.

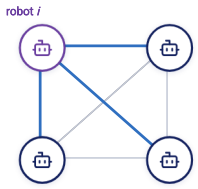
Global layer (c_g)
Fleet centroid, task centroid, unfinished-task fraction.

Zone layer (c_s)
Positions & active-task stats in robot i's quadrant.

Individual layer (c_i)
Robot i's position & distance to nearest unfinished task.

Pairwise alignment between robots i and j

$$\alpha_{ij} = \lambda_g \cos(c_g^i, c_g^j) + \lambda_s \cos(c_s^i, c_s^j) + \lambda_i \cos(c_i^i, c_i^j)$$



Coordination group for robot i:
$\mathcal{G}_i = \{ j \neq i : \alpha_{ij} > \tau \}$

Group feature g_i: mean of individual contexts in $\mathcal{G}_i$ (defaults to c_i if $\mathcal{G}_i = \varnothing$).

coordination group
other pairs

Policy input per robot: $[c_g^i, c_s^i, c_i^i, g_i]$
Critic input: concatenated joint features

Key hyperparameters:
$d = 32$, $\tau = 0.95$, $\lambda_g = \lambda_s = \lambda_i = \frac{1}{3}$

3 ADAPTIVE DIFFERENTIAL PRIVACY

Layer-wise Gaussian DP

Add Gaussian noise to each layer input:

$$\tilde{x}_\ell = x_\ell + \mathcal{N}(0, \sigma_\ell^2 I), \quad \ell \in \{g, s, i\}$$

Sensitivity $\Delta \leq 1$ (inputs in $[0, 1]$).

Per-step DP guarantee (ϵ_{step}, $\delta = 10^{-5}$):

$$\epsilon_{\text{step}} = \frac{\Delta\sqrt{2\ln(1.25/\delta)}}{\sigma_\ell}$$

At $\sigma = 0.1$: $\epsilon_{\text{step}} \approx 48.4$ At $\sigma = 1.0$: $\epsilon_{\text{step}} \approx 4.85$

(We report per-step ϵ; composition deferred.)

Per-robot, per-step adaptive rules

Deadline pressure ($\Delta t \geq 75$ steps since last completion, $\sim$4.2–5.0% of robot-steps)
$(\sigma_i, \lambda_g, \lambda_s, \lambda_i) = (0.25\,\sigma_{\text{base}},\ 0.25,\ 0.25,\ 0.50)$
→ Emphasize individual layer

Local crowding ($p_j \geq 3$ robots in quadrant, $\sim$3.1–5.2% of robot-steps)
$(\sigma_s, \lambda_g, \lambda_s, \lambda_i) = (0.5\,\sigma_{\text{base}},\ 0.20,\ 0.50,\ 0.30)$
→ Emphasize zone layer

Otherwise: all $\sigma_\ell = \sigma_{\text{base}}$, $\lambda_\ell = \frac{1}{3}$

4 MAIN RESULTS — ALL-TASKS-DONE RATE

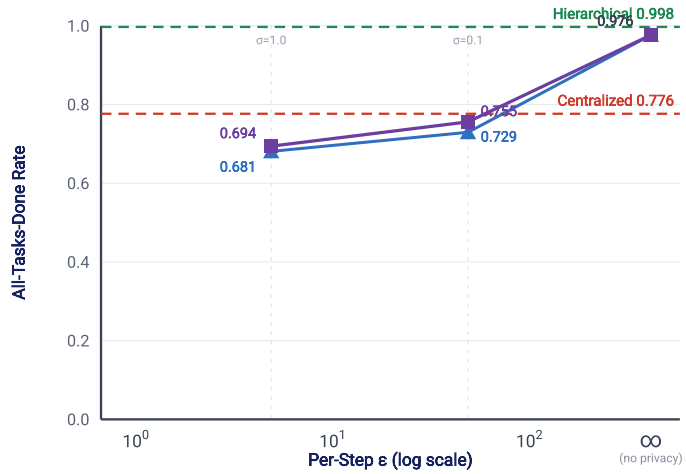
Higher is better. ε is per-step Gaussian DP guarantee at δ = 10^{-5}, Δ = 1. Critical cells use n = 8 seeds; others n = 3 seeds. Mean ± std.

Method	σ	ε/step	All-Tasks-Done (↑)
Centralized (no hierarchy)	–	∞	0.776 ± 0.004
Hierarchical (no privacy)	–	∞	0.998 ± 0.001
Fixed-DP (no noise)	0.00	∞	0.976 ± 0.007
Fixed-DP	0.10	48.4	0.729 ± 0.014
Adaptive-DP	0.10	48.4	0.755 ± 0.013
+2.63 pp improvement at matched privacy budget (Welch's p ≈ 0.003, n = 8 seeds)			
Fixed-DP	1.00	4.85	0.681 ± 0.015
Adaptive-DP	1.00	4.85	0.694 ± 0.010
+1.34 pp improvement (Welch's p ≈ 0.07, not significant)			

Hierarchy retains a clear advantage over the independent baseline at **all privacy levels** (full results in supplementary).

5 PRIVACY–UTILITY PARETO FRONTIER

X-axis: per-step ε (log scale; larger ε = weaker privacy). Right edge denotes no privacy (ε = ∞).



Adaptive-DP
Fixed-DP
Hierarchical (no privacy)
Centralized (no hierarchy)

+2.63 pp at ε ≈ 48.4 (σ = 0.1) p ≈ 0.003

+1.34 pp at ε ≈ 4.85 (σ = 1.0) p ≈ 0.07

6 ABLATION STUDIES ($\sigma_{\text{base}} = 0.5$)

All-Tasks-Done Rate (n = 3)

Variant	Rate (mean ± std)
Deadline rule only	0.702 ± 0.003
Crowding rule only	0.702 ± 0.013
Both rules (Adaptive)	0.687 ± 0.018

Findings

- Each rule contributes similarly (~+1.5 pp over both).
- Combining rules does not add additively — bounded headroom (combined firing ≤ 10% of robot-steps).
- Realized ε within 3% of fixed baseline.

Rule fire rates (empirical):
Deadline: 4.2–5.0% • Crowding: 3.1–5.2% of robot-steps

7 BYZANTINE ROBUSTNESS (σ = 0.1, 40% compromise)

Alignment-spoofing attack (white-box) + random actions. Defence: trim-mean (drop highest & lowest α peers), n = 10 seeds.

Method	Condition	All-Tasks-Done (mean ± std)
Fixed-DP	Clean	0.795 ± 0.033
	Byzantine (no defence)	0.717 ± 0.054
	Byzantine + trim-mean	0.720 ± 0.040
Adaptive-DP	Clean	0.797 ± 0.058
	Byzantine (no defence)	0.759 ± 0.063
	Byzantine + trim-mean	0.764 ± 0.054

Attack drop: Fixed-DP −7.8 pp (p ≈ 0.0014), Adaptive-DP −3.8 pp (p ≈ 0.18).

Adaptive-DP > Fixed-DP under attack (+4.2 pp, p ≈ 0.13, in expected direction).

Trim-mean does not recover significantly at 40% (single-trim limit with up to 2 adversaries).

8 KEY TAKEAWAYS

Three-layer hierarchical alignment enables strong coordination **(0.998 all-tasks-done)** without explicit messaging.

Rule-based adaptive DP improves privacy–utility at low noise: **+2.63 pp** at matched budget (σ = 0.1, p ≈ 0.003).

Hierarchy remains better than independent partial-view agents at **all privacy levels**.

Adaptive DP shows **better resilience** to alignment-spoofing attacks at 40% compromise.

Future work: larger fleets, stronger attacks & defences, learned adaptation, and full DP composition accounting.

EXPERIMENTAL SETUP
5 robots on 10×10 grid; 10 tasks, deadline = 100 steps without new completion, horizon = 200 steps.

TRAINING
PPO: 2×10^6 steps, 32 envs; 128 steps/rollout, 4 epochs, clip ε = 0.2; shared hyperparameters across methods.

METRICS
Primary: all-tasks-done rate; report mean ± std over final 10% of training.