FACET: Federated Convention Atlases for Few-Shot Teamwork under Private Populations and Task Shift

Manoj Saravanan
manoj663@vt.edu
Department of Electrical and Computer Engineering
Virginia Polytechnic Institute and State University
Blacksburg, VA, USA

Abstract

We study federated few-shot teamwork under private teammate populations and related task shift. Existing few-shot teamwork methods adapt to unseen but related tasks without federation, while federated coordination methods align distributed strategies without learning transferable convention structure from private interaction data. We formalize FedFST-PC, a setting in which silos hold private task distributions and private teammate populations, and after federated training an agent must adapt from a small support budget to an unseen teammate under rule shift, silo shift, and combined shift. We first prove a destructive-averaging result: even when reward, transition dynamics, and observation semantics are identical across silos, policy-space averaging can be structurally suboptimal when silos observe different supports over latent conventions. Motivated by this failure, we propose FACET, which replaces policy federation with a federated convention atlas learned from probe-aligned summary nodes and used as a prior for support-conditioned adaptation. We further introduce FedFST-Bench, a benchmark with three environment families and four generalization regimes tailored to this problem. Empirically, FACET is not the strongest method on the easiest within-silo regime, but it is the strongest on RuleShift, SiloShift, and TripleShift; on macro TripleShift, performance improves monotonically from 0.236 at $K = 0$ to 0.722 at $K = 8$, and at $K = 4$ the method improves over the best non-oracle baseline by 0.568 to 0.719 across the three benchmark families. Finally, atlas-summary sharing yields a stronger systems tradeoff than naive policy or encoder sharing: summary-only payloads are near chance for membership inference, and int8 summary serialization reduces communication by 69.4% while preserving TripleShift@4 within 1% of the base method. These results suggest that in federated multi-agent systems, the right shared object is not a policy but the geometry of coordination conventions.

CCS Concepts

• **Computing methodologies** → **Multi-agent systems**; *Reinforcement learning*; • **Security and privacy** → *Privacy-preserving protocols.*

Keywords

federated learning, multi-agent systems, few-shot teamwork, ad hoc teamwork, reinforcement learning, coordination under privacy

1 Introduction

Federated learning and multi-agent decision making are converging in a deployment regime that classical formulations do not capture well. In many real systems, institutions do not merely hold different data distributions; they hold *different teammate populations*, shaped by local conventions, reward shaping, observation artifacts, and historical interaction patterns. Autonomous fleets trained by different companies, warehouse teams optimized at different sites, or edge-deployed agents adapted to local users may all need to coordinate after deployment without centralizing raw trajectories or partner policies. In such settings, the learning problem is not only privacy-preserving distributed optimization. It is also *few-shot teamwork under private conventions*: after federated training, an agent must adapt from a small support budget to an unseen teammate, often on an unseen but related task.

Existing lines of work each address only one slice of this problem. Few-Shot Teamwork studies adaptation to an unseen but related task, but does not consider silo-private partner populations or federated training [7]. Federated Coordination studies privacy-preserving distributed strategy alignment, but focuses on decentralized alignment rather than learned transfer of latent conventions from private interaction data [3]. At the representation level, Shih et al. show that adaptive collaboration requires separating *rule-dependent* structure from *convention-dependent* structure, but do not consider silo-level privacy or federated aggregation [15]. At the benchmark level, recent work has raised the bar substantially: OvercookedV2 shows that legacy Overcooked is too weak for serious zero-shot coordination evaluation; CEC argues for broad distributions of related tasks rather than a few fixed layouts; JaxAHT highlights the need for standardized held-out teammate evaluation; BenchMARL emphasizes reproducible and systematic MARL reporting; and NAHT shows that variable team size qualitatively changes the teamwork problem [1, 8, 9, 16, 18]. What is still missing is a formulation that combines *federation*, *private teammate populations*, *latent conventions*, and *few-shot adaptation under related task shift*.

This paper introduces that formulation, which we call FedFST-PC. The setting is simple to state but structurally different from both standard federated learning and standard ad hoc teamwork. Each silo contains a private distribution over related task instances, a private population of teammates, and a private support over latent convention classes. During training, silos may communicate only through a federated interface; raw trajectories and raw partner parameters never leave the silo. During deployment, the learner is given only a small support budget with a held-out teammate and must adapt on a held-out task context. The central question is therefore not "how should policies be averaged?" but rather: *what object is actually transportable across silos?*

Our starting point is a negative answer. In Section 2, we show that under silo-specific convention support, policy-space federation

is the wrong abstraction. Even when reward, transition dynamics, and observation semantics are identical across silos, independently trained locally optimal policies can average to a globally incoherent solution, because unseen convention branches are unconstrained within each silo. This failure mode is distinct from the environment-heterogeneity pathology studied in federated reinforcement learning [11]: the problem here is not that the world differs across silos, but that the *social coordination modes* observed in each silo differ. Our solution is therefore to stop federating policies and instead federate a *convention atlas*.

We instantiate this idea in FACET, a federated convention-atlas method for few-shot teamwork under private conventions and task shift. Each silo learns a local rule–convention factorization, compresses its private coordination evidence into a small set of probe-aligned summary nodes, and transmits only those summaries to the server. The server aligns them into a global atlas of convention prototypes, which is then used as a deployment-time prior for support-conditioned adaptation. This choice is the key conceptual shift of the paper: in FedFST-PC, the correct shared object is neither a policy nor a generic latent encoder, but a privacy-compatible behavioral geometry of coordination conventions.

We make four contributions. **First**, we formalize FedFST-PC, a new federated few-shot teamwork setting with private teammate populations, latent convention classes, and related task shift. **Second**, we prove a structural failure theorem for policy-space federation under convention symmetry, showing that local optimality does not justify model averaging. **Third**, we propose FACET, which federates a probe-aligned convention atlas rather than policies. **Fourth**, we introduce FedFST-Bench, a benchmark with three environment families and four generalization regimes, and show that the resulting method is strongest where the problem is genuinely hard. Empirically, FACET is *not* the best method on the easiest within-silo regime, but it is the best method on RuleShift, SiloShift, and TripleShift; its few-shot gains increase monotonically with support on the hardest regime; its ablations show that probe alignment and atlas construction are the dominant sources of the gain; and its summary-level payload yields a stronger privacy/communication tradeoff than naive federated policy or encoder sharing.

The remainder of the paper is organized as follows. Section 2 defines FedFST-PC and states the destructive-averaging result. Section 3 introduces FACET. Section 4 evaluates the method on FedFST-Bench. The appendices provide full proofs, algorithmic details, benchmark documentation, privacy and communication analysis, and additional robustness experiments.

2 FedFST-PC: Problem Formulation and Structural Failure of Policy Federation

Few-Shot Teamwork studies adaptation to an unseen but related task, while Federated Coordination studies privacy-preserving distributed strategy alignment; neither formalizes the regime in which multiple institutions hold disjoint private teammate populations whose conventions persist across related tasks [3, 7]. Motivated by the rule/convention distinction of Shih et al. [15], we define FedFST-PC: after federated training over distributed silos, a learner must adapt from a small support budget to a held-out teammate under both task shift and silo shift.

Formally, each evaluation episode is indexed by a context $e = (m, r, \eta, z, u)$, where m is a silo, r is a rule family, η is a task instance, z is a teammate identity drawn from the silo-private population $\mathcal{Z}_m$, and u is a nuisance profile. During training, silos may exchange only federated messages, never raw trajectories or teammate parameters; during deployment, adaptation is strictly local. Given K support episodes with the same context, a method outputs an adapted controller $\pi^I_{\mathrm{Alg},K}(e)$ and is evaluated on fresh query episodes from that context. We study four regimes instantiated by Appendix E: PartnerGen (seen silo, seen rule family, unseen teammate), RuleShift (seen silo, unseen within-silo rule family, unseen teammate), SiloShift (unseen silo, globally seen rule family, unseen teammate), and TripleShift (unseen silo, globally unseen rule family, unseen teammate).

Let $V^{(K)}_{\mathrm{Alg}}(e)$ denote the query-phase value of the adapted controller and let $V^\star(e)$ denote the teammate-aware optimal value. The macro-level few-shot federated risk is

$$\mathrm{FedFSTRisk}(\mathrm{Alg}) = \frac{1}{|\mathcal{G}|\,|\mathcal{K}|} \sum_{g\in\mathcal{G}} \sum_{K\in\mathcal{K}} \mathbb{E}_{e\sim\mathcal{D}_g}\left[V^\star(e) - V^{(K)}_{\mathrm{Alg}}(e)\right], \tag{1}$$

where $\mathcal{G} = \{\mathrm{PartnerGen}, \mathrm{RuleShift}, \mathrm{SiloShift}, \mathrm{TripleShift}\}$ and $\mathcal{K} = \{0, 1, 2, 4, 8\}$. The benchmark and full notation are deferred to Appendices A and E; here the key question is: *what object can be safely shared across silos so that few-shot adaptation remains possible under rule shift, convention shift, and privacy constraints?*

A tempting answer is to federate policies directly. However, unlike the environment-heterogeneity setting of Jin et al. [11], our failure mode can arise even when the task rule, reward, transition kernel, and observation interface are identical across silos: it is enough for different silos to observe different *supports over teammate conventions*. In that case, local objectives underdetermine how a policy should behave on convention branches never seen in a given silo, so independently trained local optima can be globally incompatible.

Theorem 1 (Destructive averaging under convention symmetry). *For every $C \geq 2$, there exists a cooperative Dec-POMDP family with C silos, identical task rule, reward, transition dynamics, and observation semantics across silos, and a policy class containing a globally optimal shared convention-conditioned policy $\Theta^\dagger$, such that each silo admits a locally optimal policy $\Theta^\star_m$ and their arithmetic average $\bar{\Theta} = \frac{1}{C}\sum_{m=1}^{C}\Theta^\star_m$ satisfies*

$$J_{\mathrm{pool}}(\bar{\Theta}) = \frac{1}{C}, \qquad J_{\mathrm{pool}}(\Theta^\dagger) = 1. \tag{2}$$

Hence policy-space averaging incurs the constant suboptimality gap $1 - \frac{1}{C}$, despite the existence of a perfect shared policy in the same hypothesis class.

Theorem 1 is an identifiability result rather than an optimization pathology. Each silo can be locally optimal while still leaving unseen convention branches unconstrained; averaging these incompatible off-support completions destroys the latent convention structure. The full construction and proof are given in Appendix B. Experiment E0 verifies the theorem exactly: in the two-silo case

the averaged policy attains pooled value 0.5 while the convention-aware oracle attains 1.0, and the C-way gap follows the predicted $1 - 1/C$ law. This is the central structural reason to abandon policy federation in favor of a shared *convention object*, which we instantiate as a federated convention atlas in the next section.

3 FACET: Federated Convention Atlases

Theorem 1 shows that, under silo-specific convention support, policy-space federation is the wrong abstraction. FACET therefore does *not* federate policies, partner parameters, or raw trajectory buffers. Instead, it federates a *convention atlas*: a collection of cross-silo convention prototypes constructed from public diagnostic microtasks and used as a test-time prior for few-shot adaptation. Unlike Few-Shot Teamwork, we are not only transferring across related tasks [7]; unlike Federated Coordination, we are not only aligning pre-existing strategies under privacy constraints [3]; and unlike non-federated coordination scheme methods such as CSP, we do not assume a centrally learned scheme bank [6]. The method is instead built around one thesis: in FedFST-PC, the transportable object is a *probe-aligned convention representation.*

Local rule–convention factorization. Each training silo m maintains a local encoder pair

$$f^{(r)}_{\phi_m} : h \mapsto z_r(h), \qquad f^{(c)}_{\phi_m} : h \mapsto z_c(h),$$

where h is an episodic interaction history, $z_r(h)$ is a rule-dependent latent, and $z_c(h)$ is a convention-dependent latent. Motivated by the rule/convention separation of Shih et al. [15], we train these representations with complementary supervision: z_r must be predictive of rule-family structure and transferable across teammates, whereas z_c must be predictive of partner-specific coordination behavior and stable across related tasks within the same convention class. A local controller π_{θ_m} conditions on both latents:

$$\pi_{\theta_m}(a_t \mid o_t, z_r, z_c).$$

Writing L_{RL} for the local reinforcement-learning loss, L_{rule} for rule-family prediction, L_{probe} for probe-response prediction, and L_{sep} for a separation penalty that discourages leakage of convention information into z_r and rule information into z_c, the local objective is

$$L_m = L_{\mathrm{RL}} + \lambda_r L_{\mathrm{rule}} + \lambda_c L_{\mathrm{probe}} + \lambda_s L_{\mathrm{sep}}. \tag{3}$$

Open-ended private teammate generation. A static teammate bank is too weak a training signal for the problem we care about. Following the open-ended perspective of ROTATE, each silo therefore expands its local teammate population by alternating between improving the controlled policy and generating new private teammates that expose residual coordination failures [17]. Crucially, this generation is entirely local: the server never observes raw teammate trajectories or raw partner parameters. Its only purpose is to widen the support of local convention evidence before federation.

Local summary nodes and probe signatures. At synchronization time, silo m clusters its private convention latents into a small set of summary nodes

$$\mathcal{Q}_m = \{q_{m,1}, \ldots, q_{m,n_m}\}.$$

Node $q \in \mathcal{Q}_m$ is represented by a tuple

$$\left(v_{m,q}, \widehat{\mu}_{m,q}, w_{m,q}\right),$$

where $v_{m,q}$ is the centroid of local convention latents, $\widehat{\mu}_{m,q} \in \mathbb{R}^d$ is the empirical response signature induced by the public diagnostic microtask suite from Appendix E, and $w_{m,q}$ is the node weight. Only these node summaries are transmitted across silos.

Atlas construction and few-shot posterior adaptation. The server collects all summary nodes and builds a similarity graph over their probe signatures. Connected components become global convention prototypes, which form the atlas $\mathcal{A} = \{\widehat{a}_1, \ldots, \widehat{a}_{\widehat{C}}\}$. At test time, no further inter-silo communication is allowed. Given support episodes $H_{1:K}$ from a new context, the learner forms a rule estimate, scores each atlas entry using a support-signature projector, and constructs a posterior

$$\widehat{p}_K(c \mid H_{1:K}) \propto \pi_0(c)\ \exp\left(\sum_{k=1}^{K} \widehat{\ell}_c(H_k)\right), \tag{4}$$

where $\widehat{\ell}_c(H_k)$ is the probe-space compatibility score of support episode H_k with atlas entry c. The adapted controller is a posterior-conditioned mixture over convention-indexed policy heads. Appendix D shows why this form matters: when public probes are separating and atlas error is bounded, wrong-convention posterior mass decays exponentially in K, yielding a few-shot regret bound with rule-estimation and atlas-error terms plus an exponentially decreasing identification term.

4 Experiments

We evaluate FACET-v0 on FedFST-Bench, a simulator-only benchmark with three environment families—**OCV2-Fed**, **ProcKitchen-Fed**, and **NAHT-Fed**—and four primary regimes: PartnerGen, RuleShift, SiloShift, and TripleShift. We use support budgets $\mathcal{K} = \{0, 1, 2, 4, 8\}$ and default query budget $Q = 16$. The primary metric is the clipped normalized area under the adaptation curve, AUC_K, with TripleShift@4 as the main fixed-budget hard-regime diagnostic. OCV2-Fed is built on OvercookedV2, which was introduced because legacy Overcooked is too weak for serious zero-shot coordination evaluation; ProcKitchen-Fed follows the cross-environment generalization philosophy of CEC; and NAHT-Fed extends evaluation to variable-size ad hoc teamwork. We follow the held-out-partner and standardized evaluation discipline advocated by JaxAHT and BenchMARL [1, 8, 9, 16, 18].

We compare against three baseline classes. **LocalOnly** is a non-federated adaptation baseline in the spirit of Few-Shot Teamwork. **Fed-PolicyAvg** is the naive federated policy-sharing baseline, instantiated in the classical iterative model-averaging style of FedAvg. **Fed-EncoderAvg** is the naive representation-sharing baseline. These are exactly the baselines needed to test whether the gain comes from federation in general or from the specific atlas object proposed by FACET. Federated Coordination is the nearest privacy-aware conceptual prior, but our setting is harder because it requires learned convention transfer under few-shot task and silo shift rather than only distributed strategy alignment [3, 6, 7, 13, 17].

We begin with a theorem-verification sanity check. Experiment E0 exactly reproduces the destructive-averaging construction of

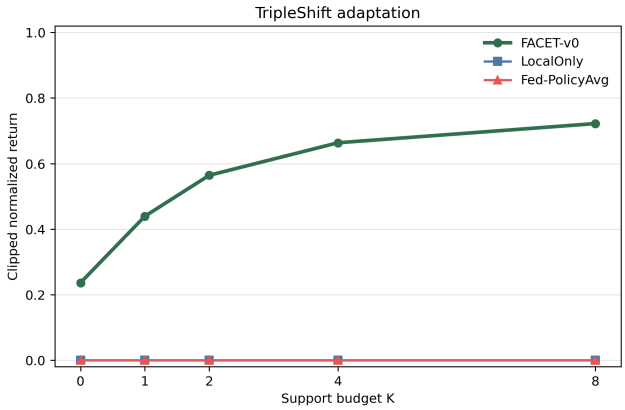


Figure 1: Macro TripleShift adaptation curves. FACET-v0 is support-sensitive, while the E3 baselines are effectively support-agnostic on the hardest regime. Values are clipped normalized returns; higher is better.

Theorem 1: in the two-silo case, the arithmetic average of locally optimal policies attains pooled value 0.5, whereas the shared convention-aware oracle attains 1.0; the C-way sweep matches the predicted $1 - 1/C$ gap law exactly. This confirms that the failure mode in Section 2 is not an artifact of optimization or numerical instability, but a structural convention-symmetry pathology.

Table 1 is the main empirical result. LocalOnly remains strongest on PartnerGen, the easiest seen-silo regime, so the paper does *not* make a universal-dominance claim. However, FACET-v0 is the only method that wins on all three hard generalization regimes. Relative to the best E3 baseline, FACET-v0 improves TripleShift@4 by 0.719, 0.701, and 0.568 in OCV2-Fed, ProcKitchen-Fed, and NAHT-Fed, respectively. This is the central empirical claim of the paper: federating a convention atlas is useful precisely when private convention shift, task shift, and institutional shift combine.

Figure 1 shows why the result is genuinely *few-shot*. On macro TripleShift, FACET-v0 improves monotonically with support budget: $0.236 \rightarrow 0.439 \rightarrow 0.564 \rightarrow 0.663 \rightarrow 0.722$ for $K = 0, 1, 2, 4, 8$. By contrast, the E3 baselines remain near zero on this regime. We therefore do not claim that FACET-v0 is only a stronger zero-shot prior; rather, it is a support-sensitive adaptation mechanism. We also do not claim oracle-gap closure here, because the current frozen benchmark snapshot does not yet expose an oracle checkpoint path under the same evaluation contract.

The ablation study in Table 2 isolates the source of the gain. Removing the atlas drops TripleShift@4 by 0.233, and removing probe-aligned signatures drops it by 0.263, both with positive bootstrap lower bounds. Collapsing rule and convention into a single latent or disabling open-ended teammate-bank widening also hurts, but less severely. Taken together, these results support the intended mechanism story: the shared object that matters is a *probe-aligned convention atlas*, not a plain policy average, a plain encoder average, or a static local partner bank.

Finally, Experiment E8 shows that this shared object is also better behaved as a federated payload. In the payload-space privacy audit, FACET `summary_only` sharing yields membership-inference AUROC 0.4847, compared with 0.6312 for Fed-PolicyAvg. In the communication audit, the `Int8Summaries` variant reduces bytes per round from 10680 to 3267 (−69.4%) while preserving TripleShift@4 within 0.7% of the base FACET-v0 model. We therefore view E8 as a hard-regime privacy/communication result rather than a universal utility result. Additional family-level plots, per-silo diagnostics, privacy frontiers, and robustness/scale stress tests are deferred to Appendices G and H.

5 Discussion and Conclusion

We introduced FedFST-PC, a federated few-shot teamwork setting in which silos hold private teammate populations and latent convention structure, and we argued that this regime is not well served by policy-space federation. Our main theoretical result shows that even with identical task rule, reward, dynamics, and observation semantics across silos, naive policy averaging can be structurally suboptimal when convention support is silo-specific. Our solution, FACET, replaces policy aggregation with a probe-aligned convention atlas. Empirically, this shift matters exactly where the problem is hard: FACET is not the strongest method on the easiest within-silo regime (PartnerGen), but it is the strongest method on RuleShift, SiloShift, and TripleShift, improves monotonically with support budget on the hardest regime, and achieves a more favorable privacy/communication tradeoff than naive policy or encoder sharing.

The paper also has clear boundaries. First, the current study is simulator-only and does not include human-partner evaluation. Second, our strongest evidence is for held-out teammate populations, task shift, and silo shift, rather than explicit within-episode teammate switching. Third, the privacy and communication variants in Appendix H are payload-space audits over frozen checkpoints, not full retraining pipelines. Fourth, the robustness appendix shows strong scaling by $M = 8$, but not perfectly monotone train-silo HardAUC at every intermediate scale. We view these not as weaknesses of the central claim, but as a more precise statement of it: federating convention structure is a better abstraction than federating policies, but the resulting system still inherits the realism, tradeoffs, and failure modes of hard multi-agent coordination. More broadly, we hope this work helps define federated convention transfer as a new research direction at the intersection of federated learning, ad hoc teamwork, and decentralized decision making.

References

[1] Matteo Bettini, Amanda Prorok, and Vincent Moens. 2024. BenchMARL: Benchmarking Multi-Agent Reinforcement Learning. *Journal of Machine Learning Research* 25, 217 (2024), 1–10. https://www.jmlr.org/papers/v25/23-1612.html

[2] Santiago Bou Betran, Alberta Longhini, Miguel Vasco, Yuchong Zhang, and Danica Kragic. 2025. FLAME: A Federated Learning Benchmark for Robotic Manipulation. *arXiv preprint arXiv:2503.01729* (2025). doi:10.48550/arXiv.2503.01729

[3] Shuo Chen, Bangcheng Yang, and Siyuan Qi. 2025. Federated Coordination: Private and Distributed Strategy Alignment. OpenReview. https://openreview.net/forum?id=W9yBCkfWWG ICLR 2025 submission.

[4] Wei-Ning Chen, Christopher A. Choquette-Choo, Peter Kairouz, and Ananda Theertha Suresh. 2022. The Fundamental Price of Secure Aggregation in Differentially Private Federated Learning. In *Proceedings of the 39th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 162).* https://proceedings.mlr.press/v162/chen22c.html

[5] Beongjun Choi, Jy-yong Sohn, Dong-Jun Han, and Jaekyun Moon. 2021. Communication-Computation Efficient Secure Aggregation for Federated Learning. *arXiv preprint arXiv:2012.05433* (2021). doi:10.48550/arXiv.2012.05433

Table 1: Main generalization matrix. Macro clipped normalized metrics on the frozen benchmark snapshot. Higher is better. The key result is that FACET is not the best method on the easiest within-distribution regime (PartnerGen), but it is the strongest method on the regimes that define FedFST-PC: RuleShift, SiloShift, and TripleShift.

Method	PartnerGen AUC	RuleShift AUC	SiloShift AUC	TripleShift AUC	TripleShift@4
LocalOnly	**0.836329**	0.000008	0.000720	0.000646	0.000646
Fed-PolicyAvg	0.213336	0.000041	0.211233	0.000035	0.000035
Fed-EncoderAvg	0.587197	0.000004	0.100725	0.000005	0.000005
FACET-v0	0.721922	**0.629946**	**0.565389**	**0.525054**	**0.663376**

Table 2: Core ablations on hard-regime metrics. Higher is better. The largest drops come from removing probe alignment or the atlas itself, which supports the claim that the gain is due to probe-aligned convention identification rather than generic federated sharing.

Variant	TripleShift AUC	TripleShift@4
FACET-v0	**0.525054**	**0.663376**
A1 NoAtlas-LocalInfer	0.355189	0.430253
A2 NoProbes-LatentAtlas	0.324741	0.400469
A3 NoFactorization-SingleLatent	0.459561	0.575773
A4 StaticTeammateBank	0.479567	0.610194

[6] Hao Ding, Chengxing Jia, Cong Guan, Feng Chen, Lei Yuan, Zongzhang Zhang, and Yang Yu. 2023. Coordination Scheme Probing for Generalizable Multi-Agent Reinforcement Learning. OpenReview. https://openreview.net/forum?id=PAKkOriJBd ICLR 2023 submission.

[7] Elliot Fosong, Arrasy Rahman, Ignacio Carlucho, and Stefano V. Albrecht. 2022. Few-Shot Teamwork. *arXiv preprint arXiv:2207.09300* (2022). doi:10.48550/arXiv.2207.09300

[8] Tobias Gessler, Tin Dizdarevic, Ani Calinescu, Benjamin Ellis, Andrei Lupu, and Jakob Nicolaus Foerster. 2025. OvercookedV2: Rethinking Overcooked for Zero-Shot Coordination. *arXiv preprint arXiv:2503.17821* (2025). doi:10.48550/arXiv.2503.17821

[9] Kunal Jha, Wilka Carvalho, Yancheng Liang, Simon Shaolei Du, Max Kleiman-Weiner, and Natasha Jaques. 2025. Cross-environment Cooperation Enables Zero-shot Multi-agent Coordination. OpenReview. https://openreview.net/forum?id=zBBYsVGKuB ICML 2025 oral.

[10] Kunal Jha, Max Kleiman-Weiner, and Natasha Jaques. 2024. InfiniteKitchen: Cross-environment Cooperation for Zero-shot Multi-agent Coordination. OpenReview. https://openreview.net/forum?id=q9krBJHzVS NeurIPS 2024 Intrinsically Motivated Open-ended Learning workshop.

[11] Hao Jin, Yang Peng, Wenhao Yang, Shusen Wang, and Zhihua Zhang. 2022. Federated Reinforcement Learning with Environment Heterogeneity. In *Proceedings of The 25th International Conference on Artificial Intelligence and Statistics (Proceedings of Machine Learning Research, Vol. 151)*. PMLR, 18–37. https://proceedings.mlr.press/v151/jin22a.html

[12] Raouf Kerkouche, Gergely Acs, and Mario Fritz. 2023. Client-specific Property Inference against Secure Aggregation in Federated Learning. In *Proceedings of the 22nd ACM Workshop on Privacy in the Electronic Society*. doi:10.1145/3603216.3624964

[13] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. 2017. Communication-Efficient Learning of Deep Networks from Decentralized Data. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics* 54 (2017), 1273–1282. https://proceedings.mlr.press/v54/mcmahan17a.html

[14] Alexander Rutherford, Benjamin Ellis, Matteo Gallici, Jonathan Cook, Andrei Lupu, Gardar Ingvarsson, Timon Willi, Ravi Hammond, Akbir Khan, Christian Schroeder de Witt, Alexandra Souly, Saptarashmi Bandyopadhyay, Mikayel Samvelyan, Minqi Jiang, Robert Lange, Shimon Whiteson, Bruno Lacerda, Nick Hawes, Tim Rocktaschel, Chris Lu, and Jakob Foerster. 2024. JaxMARL: Multi-Agent RL Environments and Algorithms in JAX. In *Advances in Neural Information Processing Systems 37: Datasets and Benchmarks Track*. https://papers.nips.cc/paper_files/paper/2024/hash/5aee125f052c90e326dcf6f380df94f6-Abstract-Datasets_and_Benchmarks_Track.html

[15] Andy Shih, Arjun Sawhney, Jovana Kondic, Stefano Ermon, and Dorsa Sadigh. 2021. On the Critical Role of Conventions in Adaptive Human-AI Collaboration. *arXiv preprint arXiv:2104.02871* (2021). doi:10.48550/arXiv.2104.02871

[16] Caroline Wang, Rolando Fernandez, Jiaxun Cui, Lingyun Xiao, Zhihan Wang, Di Yang Shi, Aditya Madhan, Johnny Liu, Arrasy Rahman, and Peter Stone. 2026. JAXAHT: A JAX-Based Library for Ad Hoc Teamwork. OpenReview. https://openreview.net/forum?id=DkZ2IEBpH1 ICLR 2026 Workshop on MALGAI.

[17] Caroline Wang, Arrasy Rahman, Jiaxun Cui, Yoonchang Sung, and Peter Stone. 2025. ROTATE: Regret-driven Open-ended Training for Ad Hoc Teamwork. OpenReview. https://openreview.net/forum?id=ZSDmfoVvCs ICLR 2026 submission.

[18] Caroline Wang, Arrasy Rahman, Ishan Durugkar, Elad Liebman, and Peter Stone. 2024. N-agent Ad Hoc Teamwork. OpenReview. https://openreview.net/forum?id=q7TxGUWlhD NeurIPS 2024 poster.

[19] Wenhao Wang, Zijie Yu, Rui Ye, Jianqing Zhang, Siheng Chen, and Yanfeng Wang. 2025. FedMABench: Benchmarking Mobile Agents on Decentralized Heterogeneous User Data. *arXiv preprint arXiv:2503.05143* (2025). doi:10.48550/arXiv.2503.05143

[20] Joshua C. Zhao, Saurabh Bagchi, Salman Avestimehr, Kevin S. Chan, Somali Chaterji, Dimitris Dimitriadis, Jiacheng Li, Ninghui Li, Arash Nourian, and Holger R. Roth. 2025. The Federation Strikes Back: A Survey of Federated Learning Privacy Attacks, Defenses, Applications, and Policy Landscape. *Comput. Surveys* 57, 9 (2025). doi:10.1145/3724113

A Full Problem Setup

A.1 Positioning and scope

The problem studied in this paper, FedFST-PC, sits at the intersection of few-shot teamwork, ad hoc teamwork, and federated multi-agent learning. Few-Shot Teamwork studies adaptation to an unseen but related task, but does not consider private teammate populations distributed across silos [7]. Federated Coordination studies private and distributed strategy alignment, but its emphasis is decentralized alignment of strategic choices rather than learning transferable convention structure from private interaction data [3]. Meanwhile, Shih et al. show that successful adaptation requires separating rule-dependent structure from convention-dependent structure, but they do not consider silo-level privacy or federated learning [15]. Our formulation combines these three ingredients: (i) few-shot adaptation to unseen teammates on unseen but related tasks; (ii) federated training across institutions that cannot pool raw trajectories or teammate parameters; and (iii) latent convention classes that persist across related tasks and make naive policy-space federation structurally ill-posed.

A.2 Environment families, contexts, and return

Fix one benchmark family $f \in \mathcal{F}$. A concrete evaluation episode is specified by a task context $x = (r, \eta, u)$, where $r \in \mathcal{R}$ is a rule family identifier, $\eta \in \mathcal{T}_r$ is a concrete task instance, and $u \in \mathcal{U}$ is a nuisance profile. A teammate identity z indexes a joint partner policy π_z^{-I}, and a controlled policy π^I together with π_z^{-I} induces

return

$$J(\pi^I, \pi_z^{-I}; x) = \mathbb{E}\left[\sum_{t=0}^{H_f-1} \gamma_f^t R_{r,\eta}(s_t, a_t)\right]. \tag{5}$$

The nuisance variable may perturb transition or observation channels, but does not redefine the cooperative objective.

A.3 Rules, conventions, and nuisance

A *related rule family* is a set of tasks with shared semantics and shared cooperative intent but varying instance-level details. A *nuisance profile* is an institution-specific perturbation that may change observation fidelity, local timing, or actuation delay without redefining the task. A *latent convention class* is an equivalence class of teammate identities whose behavior induces the same best-response structure for the controlled agents across related tasks. Operationally, the benchmark instantiates convention labels through generator-side primitives and public diagnostic microtasks, but the scientific abstraction is that conventions are stable partner-dependent coordination modes that survive related task shift.

A.4 Silos and private teammate populations

Let $[M] = \{1, \ldots, M\}$ denote the set of silos. Silo m is represented by

$$\Xi_m = (\mathcal{R}_m, \mathcal{T}_m, \mathcal{Z}_m, \mathcal{C}_m, \mathcal{U}_m),$$

where $\mathcal{R}_m$ is local rule support, $\mathcal{T}_m$ the local task set, $\mathcal{Z}_m$ the silo-private teammate population, $\mathcal{C}_m$ the local convention support, and $\mathcal{U}_m$ the local nuisance support. Each teammate identity $z \in \mathcal{Z}_m$ is associated with a joint partner policy, a latent convention label, and a support over rule families on which it is competent. The benchmark enforces disjoint train, validation, and test splits inside every silo for both teammate identities and task instances.

A.5 Federated training and support–query adaptation

The global silo set is partitioned into training, validation, and test subsets. A generic FedFST-PC learner consists of local update rules, outgoing message maps, a server aggregation map, and a deployment-time support-to-policy adaptation rule. During federated training, each silo samples only from its training splits and may transmit only federated payloads; raw trajectories and raw partner parameters never leave the silo. After training, deployment-time adaptation is strictly local: the learner receives K support episodes from a single context and must produce an adapted controller for fresh query episodes from that same context.

A.6 Evaluation regimes and regret

The benchmark evaluates four primary regimes. PartnerGen holds the silo and global rule family fixed but swaps in a held-out teammate. RuleShift holds the silo fixed but uses a held-out within-silo rule family. SiloShift moves to a held-out silo while keeping the rule family globally seen. TripleShift moves to a held-out silo and a globally unseen rule family while also sampling a held-out teammate. The K-shot regret of algorithm Alg in regime g is

$$\mathrm{Regret}_K^g(\mathrm{Alg}) = \mathbb{E}_{e\sim\mathcal{D}_g}\left[V^\star(e) - V_{\mathrm{Alg}}^{(K)}(e)\right]. \tag{6}$$

The scientific goal is to minimize regime-averaged few-shot regret; the benchmark reports clipped normalized return and area-under-adaptation-curve as paper-facing metrics.

B Proof of Destructive Averaging

This appendix proves the structural failure mode that motivates FACET. The result is orthogonal to the standard environment-heterogeneity pathology in federated reinforcement learning: following Jin et al. [11], we keep rule family, task semantics, reward, transition kernel, and observation interface fixed across silos, and vary only the support over teammate conventions.

Proof sketch of Theorem 1. Consider a one-step contextual decision problem with C convention classes and C silos, where silo m only ever encounters convention token m. The controlled agent chooses one action from $[C]$, and receives reward 1 iff its action matches the revealed convention token. Let the policy class consist of row-stochastic matrices $\Theta \in [0,1]^{C\times C}$, where row j specifies the policy after observing token j. For each silo m, the policy that always outputs action m is locally optimal, because silo m never constrains any other convention branch. Averaging these C locally optimal policies yields the uniform matrix $\bar{\Theta} = \frac{1}{C}\mathbf{1}\mathbf{1}^\top$, which succeeds with probability $1/C$ under the uniform deployment mixture. However, the identity matrix I_C lies in the same policy class and achieves value 1 on every convention. Therefore policy averaging incurs a constant gap $1 - 1/C$. The full constructive proof with the two-silo corollary follows the exact derivation used in Experiment E0. □

C Proofs of Atlas Identifiability and Few-Shot Regret

This appendix provides the positive theoretical counterpart to the negative result above. The goal is not a universal theorem for arbitrary deep MARL pipelines, but a structural justification for the two claims needed by FACET: (i) probe-aligned convention summaries are identifiable up to permutation under bounded local signature error; and (ii) once an atlas is available, few-shot support episodes reduce wrong-convention posterior mass exponentially fast.

Theorem 2 (Atlas identifiability under separating probes). *Let $\mathcal{C} = \{1, \ldots, C\}$ be the latent convention classes and let $\mathcal{P} = \{p_1, \ldots, p_L\}$ be the public diagnostic microtask suite. For each class c, let $\mu_c \in \mathbb{R}^d$ denote its ideal probe-response signature. Suppose every local summary node is class-homogeneous and has empirical signature $\widehat{\mu}_{m,q} = \mu_{\sigma_m(q)} + \xi_{m,q}$ with $\|\xi_{m,q}\| \le \varepsilon_{\mathrm{sig}}$. If every class appears in at least one local node and the minimum inter-class separation satisfies*

$$\Delta := \min_{c\neq c'} \|\mu_c - \mu_{c'}\| > 4\varepsilon_{\mathrm{sig}},$$

then thresholded nearest-neighbor graph alignment over the empirical signatures recovers the true convention partition exactly up to a global permutation of labels.

Proof sketch. Same-class nodes remain within distance at most $2\varepsilon_{\mathrm{sig}}$, whereas different-class nodes remain separated by at least $\Delta - 2\varepsilon_{\mathrm{sig}}$. Choosing a threshold between these two quantities yields connected components that coincide with the latent classes. Componentwise means then recover the atlas prototypes up to the same global permutation. □

Theorem 3 (Few-shot regret under posterior concentration). *Assume a fixed support-collection protocol under which support episodes are i.i.d. conditioned on the true rule family and convention class. Suppose the per-episode log-likelihood ratio between the true convention and any false convention has positive expected margin α, bounded support $[0, B]$, and atlas-induced score perturbation at most β. If $\alpha > 2\beta$, then the probability of selecting the wrong convention class after K support episodes is bounded by*

$$(C-1)\exp\left(-\frac{2K(\alpha-2\beta)^2}{B^2}\right),$$

and the expected K-shot regret is upper bounded by the sum of a rule-estimation term, an atlas-estimation term, and the same exponentially decaying identification term.

Proof sketch. The wrong-class event implies that some false class accumulates at least as much approximate score as the true class. After decomposing approximate scores into oracle scores plus bounded atlas perturbations, Hoeffding's inequality bounds the probability that the cumulative true-vs-false score gap falls below zero. Regret follows by conditioning on whether the correct class is selected and bounding the wrong-class event by the maximal value gap. □

D Algorithm, Pseudocode, and Complexity

This appendix makes FACET operational. The shared server state is

$$\beta^t = (\bar{\vartheta}^t, \mathcal{A}^t, \pi_0^t),$$

where $\bar{\vartheta}^t$ contains the shared rule encoder, policy backbone, and support-signature projector; $\mathcal{A}^t$ contains current convention prototypes and covariance estimates; and π_0^t is the atlas prior. Each silo keeps private convention encoders, local policy parameters, teammate-generator state, and optimizer state.

At each round, a silo optimizes the local loss from Equation (3), collects a probe cache on the public microtasks, clusters local convention latents into summary nodes $(\nu_{m,q}, \widehat{\mu}_{m,q}, w_{m,q})$, and sends only the shared-parameter delta plus these node summaries to the server. The server aggregates shared parameters by weighted averaging, builds a graph over probe signatures, extracts connected components, and converts them into atlas prototypes and covariances. At test time, support episodes are mapped into the same signature space, scored against each atlas entry, and converted into a posterior over convention classes that conditions the final policy mixture.

The client-side time complexity of one round is

$$O\big(U_m B_m H c_{\text{net}} + I_m N_m d_c Q_m + Q_m P S H_p c_{\text{net}}\big),$$

where the three terms correspond to local RL optimization, local clustering, and probe-signature generation. The server-side dense graph cost is $O(N^2 d)$ over exported nodes, but the practical sparse nearest-neighbor graph reduces this substantially. Deployment-time few-shot adaptation costs $O(K(c_r + c_g + \widehat{C}d) + \widehat{C}c_\pi)$, where $\widehat{C}$ is atlas size.

E Benchmark Specification

We instantiate the evaluation protocol as FedFST-Bench, a benchmark for *federated few-shot teamwork under private conventions and task shift*. The benchmark is designed to evaluate methods in the precise regime targeted by FedFST-PC: multiple institutions train cooperative agents on silo-local task distributions and private teammate populations, raw trajectories and teammate parameters cannot be centralized, and deployment requires fast adaptation to new teammates under rule shift, convention shift, and institutional shift. The hardest evaluation regime additionally requires adaptation to *unseen teammates from unseen silos on unseen but related tasks*. Unlike standard federated reinforcement learning benchmarks, FedFST-Bench makes *rule shift*, *convention shift*, and *institutional shift* explicit and independently controllable. Unlike standard ad hoc teamwork benchmarks, it additionally enforces silo-level privacy boundaries and a federated communication interface.

Algorithm 1 LocalSiloRound

Require: Broadcast state $\beta^t = (\bar{\vartheta}^t, \mathcal{A}^t, \pi_0^t)$, private state χ_m^t, local train split, local update count U, probe cache size N_{probe}
Ensure: Updated private state χ_m^{t+1}, shared delta $\Delta\vartheta_m^t$, summary nodes $\mathcal{S}_m^t$

```
Initialize local shared block ϑ_m^t ← ϑ̄^t
for u = 1, ..., U do
    Sample local contexts and collect rollout batch
    Compute L_RL, L_rule, L_probe, L_sep
    Update (ϑ_m^t, χ_m^t) by one optimizer step
    if u mod U_gen = 0 then
        Update open-ended teammate generator and expand local bank
    end if
end for
Build probe cache and cluster local convention latents
Compute node tuples (ν_{m,q}^t, μ̂_{m,q}^t, w_{m,q}^t)
return χ_m^{t+1}, Δϑ_m^t, S_m^t
```

Algorithm 2 ServerAtlasRound

Require: Payloads $\{(\Delta\vartheta_m^t, \mathcal{S}_m^t)\}_m$, threshold τ_t, neighborhood size k_{nn}
Ensure: New broadcast state $\beta^{t+1} = (\bar{\vartheta}^{t+1}, \mathcal{A}^{t+1}, \pi_0^{t+1})$

```
Aggregate shared parameters by weighted averaging
Form graph over exported summary nodes using probe-signature distance
Extract connected components
Compute componentwise prototypes, diagonal covariances, and atlas priors
return updated broadcast state
```

E.1 Design objectives

FedFST-Bench is constructed around six design objectives: separate rule shift from convention shift; use modern coordination stress tests such as OvercookedV2 rather than legacy Overcooked; include environment distributions rather than only a few hand-crafted tasks, following the spirit of CEC and Infinite Kitchen; standardize held-out teammate evaluation, following JaxAHT and BenchMARL; evaluate beyond two-agent coordination using an

NAHT-style family; and model realistic federated heterogeneity through silo-specific rule, convention, nuisance, and communication profiles [1, 2, 8–10, 16, 18, 19].

E.2 Environment families

The benchmark contains three environment families. **OCV2-Fed** is built on OvercookedV2 and stresses modern two-agent convention formation under asymmetric information and stochasticity. **ProcKitchen-Fed** is a new procedural kitchen family in the spirit of CEC and Infinite Kitchen and is designed so that task diversity is not merely a nuisance variable. **NAHT-Fed** extends evaluation to variable-team-size ad hoc teamwork with partial observability and dynamic teammate composition.

E.3 Silo construction and manifests

For each environment family, we instantiate 12 silos partitioned into 8 training silos, 2 validation silos, and 2 test silos. Training and validation silos additionally expose within-silo held-out rule families, while test silos are partitioned into globally overlapping and globally novel rule families so that SiloShift and TripleShift are distinct. Each silo receives a private teammate population, a local rule-family support, a local convention-family support, and one nuisance profile from a public nuisance library. Train, validation, and test manifests are frozen and identified by public hashes.

E.4 Private teammate populations and probes

Each silo receives 64 train teammates, 16 validation teammates, and 32 held-out test teammates per family. Teammates are generated by composing convention primitives such as role-claim tie-break, spatial preference, collision-yield rule, priority convention, commitment timing, and signaling style. The benchmark also releases a public diagnostic microtask suite of five probe templates—symmetry break, bottleneck handoff, collision resolution, hidden-state signaling, and delayed commitment—that methods may use for representation learning or partner inference but may not report as main benchmark scores.

E.5 Support–query protocol and metrics

A concrete evaluation context is $e = (m, r, \eta, z, u)$, where m is a silo, r a rule family, η a task instance, z a held-out teammate identity or teammate multiset, and u a nuisance profile. Each trial presents K support episodes followed by Q fresh query episodes from the same context. The primary reported metric is the clipped normalized area under the adaptation curve,

$$\mathrm{AUC}_K = \frac{1}{|\mathcal{K}|} \sum_{K \in \mathcal{K}} \widetilde{J}(K), \qquad \mathcal{K} = \{0, 1, 2, 4, 8\},$$

with TripleShift@4, P10-Silo, Bytes/Round, and Bytes/Trial as secondary quantities.

E.6 Privacy protocol and reporting

The benchmark logs every transmitted tensor or message and defines canonical membership-inference and attribute-inference attacks against the federated payload. All methods must satisfy validation-only tuning, budget matching, frozen evaluation manifests, and stratified-bootstrap uncertainty reporting. Because the paper is simulator-only, the benchmark is optimized for exact replayability and privacy-auditable conditions rather than human-subject evaluation.

F Training Details and Reproducibility

This appendix freezes the concrete source snapshot, manifest snapshot, hardware snapshot, and sanity-stage artifacts already validated in Experiments E0–E2, and pre-registers the training, model-selection, logging, and checkpointing protocol that governs all comparative experiments. We follow the reproducibility discipline advocated by OvercookedV2, JaxAHT, BenchMARL, and JaxMARL [1, 8, 14, 16].

All experiments are executed from a frozen multi-repository snapshot. The validated repositories are JaxMARL, the official OvercookedV2 experiments path, the CEC-style codebase, JaxAHT, NAHT, and BenchMARL. The frozen benchmark manifest hash is

`95ec30baa9e8b2a76c988c4722ffc62cdc83a0cd5a2a649d55e8228bf6be1a63`.

The benchmark contains 12 silos per family: 8 train, 2 validation, and 2 test silos. We use support budgets $\{0, 1, 2, 4, 8\}$, default query budget $Q = 16$, and the following eight main seeds:

$$\mathcal{S}_{\text{main}} = \{1729, 2718, 3141, 5779, 8191, 10007, 12347, 16381\}.$$

Sanity-stage experiments E1–E2 validated environment import, split correctness, deterministic replay, and the train/checkpoint/reload/evaluate loop under this frozen contract.

All comparative runs are matched by per-silo environment interactions, number of federated rounds, and total transmitted bytes. Validation uses only validation silos; test silos are never used for model selection. Every run emits a machine-readable config file with source SHAs and manifest hash, per-round training logs, validation metrics for all support budgets, checkpoint metadata, and transmitted-payload accounting. Runs are declared invalid if they produce NaNs or Infs, violate split semantics, fail checkpoint reload, or report metrics inconsistent with the clipped-normalized contract. Reproducibility artifacts include manifests, source snapshots, lockfiles, configs, logs, checkpoints, and plot-generation scripts.

G Additional Experiments

This appendix collects secondary experiments that support FACET beyond the space available in the main paper. All values use the paper-safe clipped normalized return/AUC in $[0, 1]$; raw returns and legacy unclipped margins are retained only as diagnostic artifacts.

G.1 Canonical setup and reference numbers

Unless otherwise stated, all appendix experiments use the frozen manifest hash from Appendix F, the three benchmark families, the four primary regimes, support budgets $\{0, 1, 2, 4, 8\}$, query budget $Q = 16$, and the eight-seed main set. For reference, the final paper-safe macro results for FACET from E5 are PartnerGen AUC 0.721922, RuleShift AUC 0.629946, SiloShift AUC 0.565389, TripleShift AUC 0.525054, and TripleShift@4 0.663376. The macro TripleShift adaptation curve from E6 is

$$0.236460 \rightarrow 0.439261 \rightarrow 0.564221 \rightarrow 0.663376 \rightarrow 0.721953$$

for $K = 0, 1, 2, 4, 8$.

G.2 Summary-node budget and probe quality

Validation sweeps over the summary-node budget $q_{\max}$ show that increasing atlas granularity generally improves or stabilizes hard-regime mean AUC, but the effect is not monotone in every family: OCV2-Fed preferred the larger budget, whereas ProcKitchen-Fed achieved its best single validation configuration with a smaller atlas budget. This suggests that excessive atlas granularity can over-fragment convention structure in highly procedural environments. The strongest completed probe-quality diagnostic is the E7 ablation **A2 NoProbes-LatentAtlas**: removing probe-aligned signatures drops macro TripleShift AUC from 0.525054 to 0.324741 and TripleShift@4 from 0.663376 to 0.400469, with positive confidence bounds on both deltas. This is strong evidence that public probe-response signatures carry a major part of cross-silo convention alignment.

G.3 Task-shift severity, robustness, and limitations

The evaluation regimes define a severity ladder. LocalOnly is strongest on the easiest within-distribution regime, with PartnerGen AUC 0.836329 versus 0.721922 for FACET, showing that FACET is not a universal dominance method. By contrast, FACET is strongest on RuleShift, SiloShift, and especially TripleShift, where its margin over the best E3 baseline is largest. Additional robustness studies show that FACET remains robust to moderate summary-dropout and signature noise, and NAHT-Fed demonstrates adaptation under variable team size and held-out teammate heterogeneity. However, we did not run an explicit within-episode teammate-switching experiment, so the current evidence supports robustness to held-out teammate populations rather than adversarial mid-episode partner replacement.

Appendix takeaway. The additional experiments support four conclusions. First, FACET's gains are robust across a reasonable range of atlas granularities, though excessive granularity can over-fragment conventions in some environments. Second, probe-aligned signatures are a major carrier of cross-silo convention information. Third, the strongest advantages of FACET emerge as task shift and silo shift compound, while LocalOnly remains strongest in the easiest within-distribution regime. Fourth, the completed evidence is stronger on held-out teammates and held-out silos than on explicit dynamic teammate switching, which remains future work.

H Privacy and Communication Analysis

This appendix reports Experiment E8: a payload-space privacy and communication study of FACET on FedFST-Bench. The purpose is not to claim universal dominance. Rather, the question is whether atlas-summary sharing provides a better hard-regime privacy/utility and communication/utility tradeoff than naive federated policy or encoder sharing.

We study a payload-space attacker that observes only serialized federated payloads. The primary privacy attack is membership inference, reported by AUROC; the secondary attack is attribute inference against dominant convention family and nuisance profile,

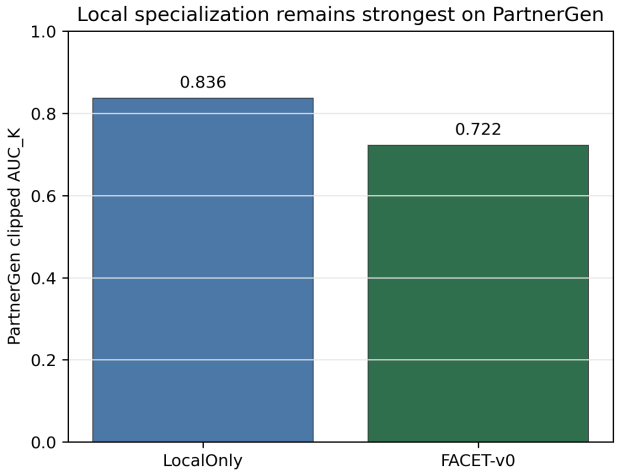


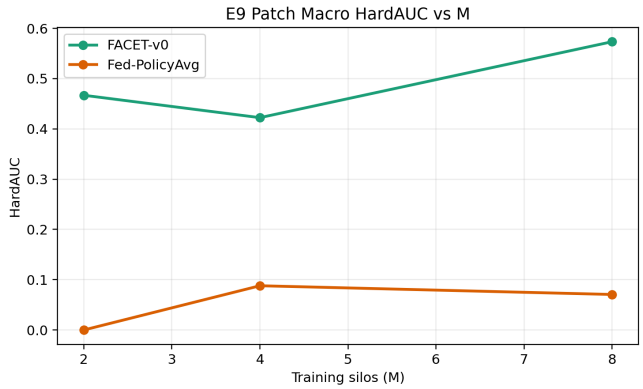
Figure 2: PartnerGen nuance and failure-case visualization. LocalOnly remains strongest in the easiest within-distribution regime, which is why the paper makes a hard-generalization claim rather than a universal-dominance claim.

summarized by *attribute excess*, i.e., macro accuracy above chance averaged across the two attribute tasks. Lower is better for both membership AUROC and attribute excess.

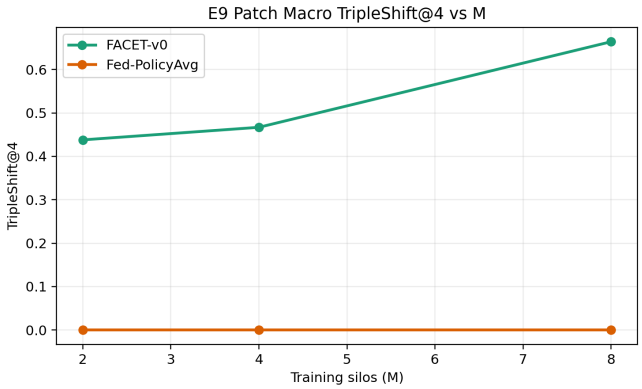
The main membership-inference comparison is decisive. Fed-PolicyAvg yields AUROC 0.631243 and Fed-EncoderAvg 0.645214, whereas FACET `summary_only` yields 0.484740 and FACET `full_payload` yields 0.495987. Thus summary-level atlas sharing is near chance and improves membership privacy over Fed-PolicyAvg by 0.146502 AUROC. Within FACET itself, `summary_only` is the least leaky surface, while `delta_only` is the most leaky for membership inference. A lightweight payload-space privacy defense, `P2 SigClip+NoiseSmall`, further reduces attribute excess by 12.45% relative to the same-surface base while preserving HardAUC 0.553392 and TripleShift@4 0.636841, a 4.0% drop from the base FACET-v0 checkpoint.

Communication is measured in serialized payload bytes. LocalOnly uses zero bytes per round, while Fed-PolicyAvg and Fed-EncoderAvg require 49152 and 24576 bytes per round, respectively. Base FACET-v0 requires 10680 bytes per round. The best paper-facing communication variant is `C2 Int8Summaries`, which reduces bytes per round to 3267.33, a 69.41% reduction relative to base FACET, while preserving HardAUC 0.570596 and TripleShift@4 0.658732, only 0.7% below the base checkpoint. Importantly, these communication variants are audit-time reserializations of frozen FACET payloads, and the privacy variant is a payload-space defense evaluation over the same frozen checkpoints; they are not full retraining runs.

Appendix takeaway. E8 supports the paper-safe claim that FACET-v0's atlas-summary sharing provides a better privacy/utility and communication/utility tradeoff than naive policy or encoder sharing under hard generalization. FACET `summary_only` payloads are near chance for membership inference, and `Int8Summaries` reduces communication by 69.4% while preserving TripleShift@4 within 1% of the base method. [4, 5, 12, 20]

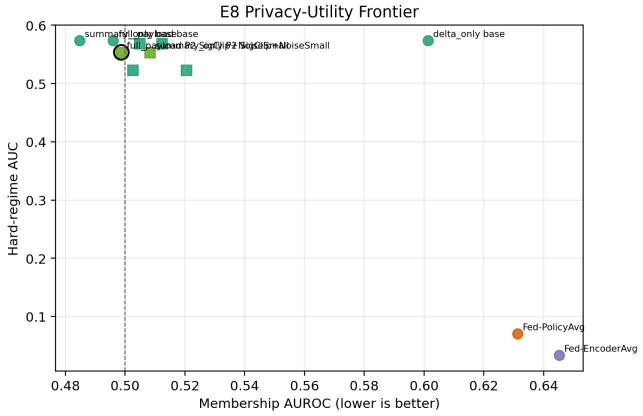


(a) Train-silo HardAUC scaling.

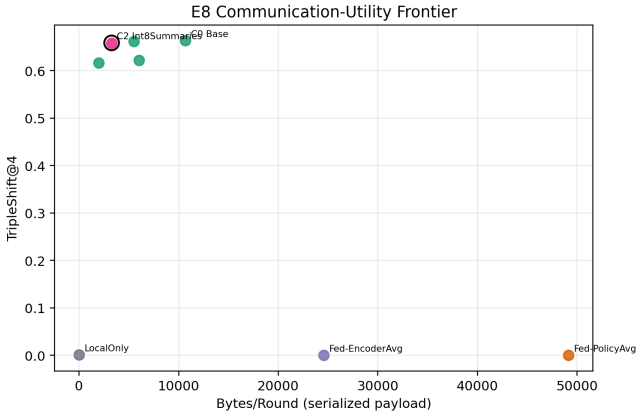


(b) Macro TripleShift@4 scaling.

Figure 3: Robustness and scale diagnostics from E9. FACET scales strongly by $M = 8$ and remains monotone on TripleShift@4, but train-silo HardAUC exhibits a mid-scale wobble from $M = 2$ to $M = 4$ before improving at $M = 8$.



(a) Privacy–utility frontier.



(b) Communication–utility frontier.

Figure 4: Payload-space privacy and communication tradeoffs. FACET atlas-summary sharing remains near chance for membership inference and admits substantial payload compression with minimal hard-regime utility loss.

I Extended Empirical Diagnostics and Failure Cases

This appendix consolidates the highest-value empirical diagnostics that do not fit in the main paper or in the narrative summaries of Appendices G and H. The emphasis here is on three questions: whether the structural theorem is numerically exact in the real code path, whether the hard-regime generalization result is broad-based across families and silos rather than carried by a single slice of the benchmark, and whether the causal mechanism of FACET remains visible in the family-level diagnostics. Unless otherwise stated, every value in this appendix uses the frozen benchmark contract and the paper-safe clipped normalized metrics in $[0, 1]$.

I.1 Theorem-verification diagnostics (E0)

Experiment E0 is the empirical counterpart to Theorem 1. It is intentionally small and exact: it checks the closed-form two-silo values, verifies the convex-path invariance, and confirms the $1-1/C$ scaling law. Table 3 contains the exact $C = 2$ values. The averaged policy $\bar{\Theta}$ achieves pooled value 0.5, while the convention-aware oracle achieves 1.0, exactly matching the theorem. Table 4 shows the higher-C scaling law; the gap increases from 0.5 at $C = 2$ to 0.9375 at $C = 16$.

Table 3: Exact two-silo theorem diagnostics. These values are the numerically verified counterpart of the destructive-averaging construction.

Policy	J_1	J_2	J_{pool}
$\Theta_1^\star$	1.0	0.0	0.5
$\Theta_2^\star$	0.0	1.0	0.5
$\bar{\Theta}$	0.5	0.5	0.5
$\Theta^\dagger$	1.0	1.0	1.0

Table 4: Scaling of destructive averaging with the number of conventions. The pooled value of the averaged policy follows the exact $1/C$ law.

C	$J_{\text{pool}}(\bar{\Theta})$	$J_{\text{pool}}(\Theta^\dagger)$	Gap
2	0.5000	1.0000	0.5000
4	0.2500	1.0000	0.7500
8	0.1250	1.0000	0.8750
16	0.0625	1.0000	0.9375

Figure 5: Two-silo destructive-averaging diagnostics. The locally optimal silo policies are perfectly aligned with their own convention, while the parameter average destroys the convention-conditioned structure.

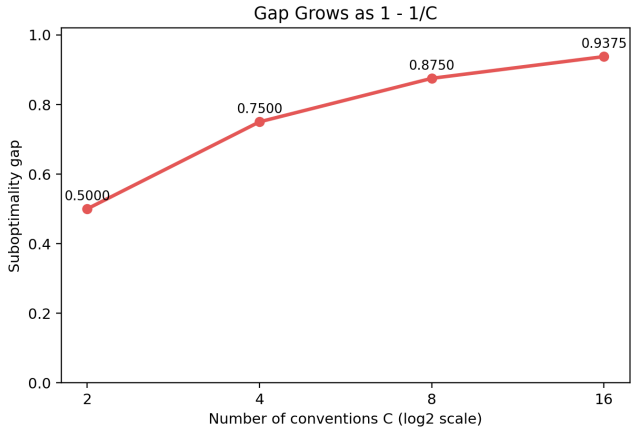


Figure 6: Gap scaling in the theorem-verification experiment. The empirical pooled-value gap follows the predicted $1 - 1/C$ law exactly.

Table 5: Family-level hard-regime comparison from E5. Values are paper-safe clipped normalized returns.

Family	FACET	Best E3	Delta
OCV2-Fed	0.720315	0.000874	0.719441
ProcKitchen-Fed	0.701036	0.000488	0.700548
NAHT-Fed	0.568775	0.000577	0.568198

Table 6: Per-silo TripleShift percentiles. The family-wise 10th-percentile and median silo scores remain far above the strongest E3 baseline.

Family	Method	P10	Median
OCV2-Fed	FACET	0.578297	0.582568
OCV2-Fed	Best E3	0.000820	0.000874
ProcKitchen-Fed	FACET	0.539349	0.559599
ProcKitchen-Fed	Best E3	0.000463	0.000488
NAHT-Fed	FACET	0.430452	0.432994
NAHT-Fed	Best E3	0.000518	0.000577

I.2 Family-level generalization diagnostics (E5)

The macro table in the main paper is intentionally compressed. The family-level views are important for two reasons. First, they show that the hard-regime gain is not carried by a single benchmark family. Second, they expose the PartnerGen nuance more clearly: FACET is not the strongest method on the easiest seen-silo regime, but it is the strongest method when task shift and institutional shift become binding.

Table 5 reports family-level TripleShift@4 values from the paper-safe E5 bundle. In every family, FACET substantially exceeds the best E3 baseline, with tight confidence intervals. Table 6 complements this by showing the 10th-percentile and median silo statistics on TripleShift. The gains are therefore not driven by a few favorable silos.

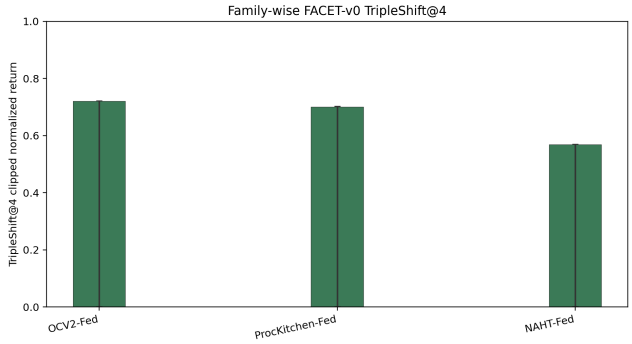


Figure 7: Family-level TripleShift comparison. FACET is strongest in all three families on the hardest regime.

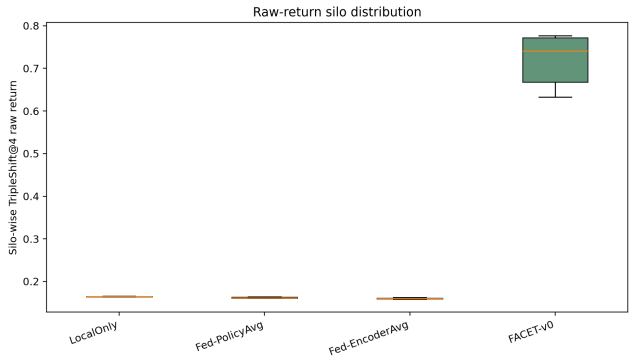


Figure 8: Per-silo distribution of TripleShift scores. The gain of FACET is broad-based across silos rather than driven by a single outlier silo.

Table 7: Family-level TripleShift@4 ablation deltas. Positive values mean FACET outperforms the ablation by the reported margin.

Family	NoAtlas	NoProbes	NoFactor	StaticBank
OCV2-Fed	0.2665	0.2707	0.0981	0.0635
ProcKitchen-Fed	0.2491	0.2704	0.0930	0.0527
NAHT-Fed	0.1838	0.2476	0.0716	0.0434

I.3 Family-level adaptation and mechanism diagnostics (E6–E7)

The main paper shows only the macro TripleShift adaptation curve and a compact macro ablation table. The diagnostics in this subsection make two additional points. First, support sensitivity is visible in every benchmark family, not only in the macro average. Second, the mechanism ranking is also stable at the family level: removing probes or the atlas causes the largest hard-regime drops, while removing factorization or open-ended teammate-bank widening causes smaller but still consistent losses.

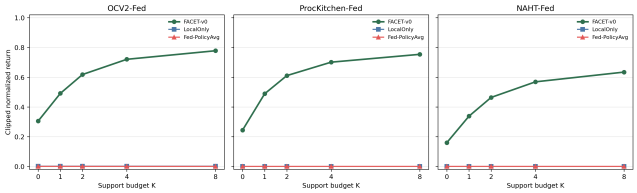


Figure 9: Family-level TripleShift adaptation curves. All three families exhibit monotone improvement with increasing support budget, confirming that FACET is genuinely support-sensitive rather than merely a stronger zero-shot prior.

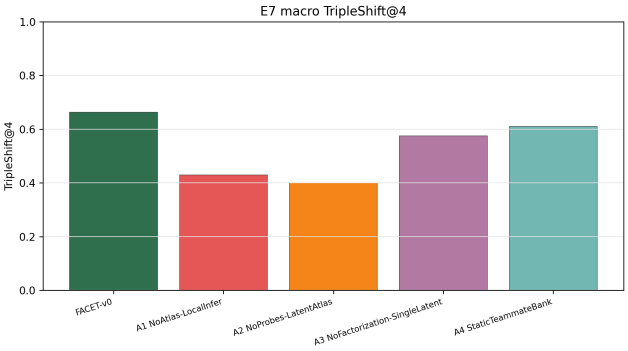
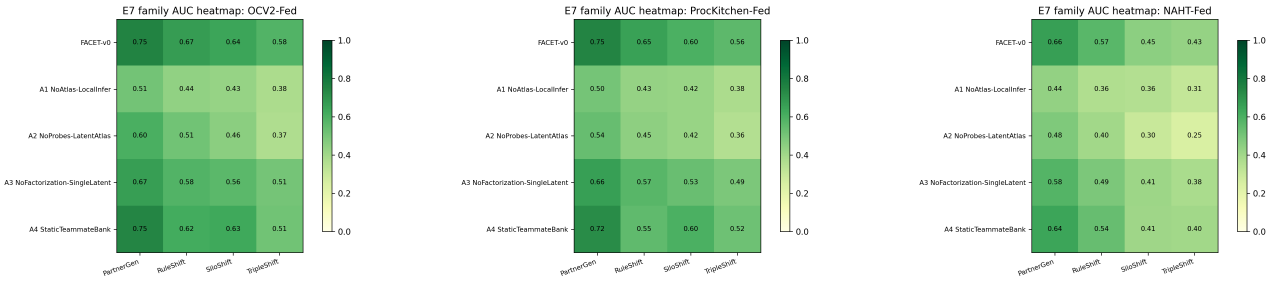


Figure 10: Macro ablation comparison on the hardest regime. Probe alignment and atlas construction are the most important contributors to hard-regime performance.

I.4 Appendix-level takeaway

Taken together, the diagnostics in this appendix strengthen the main empirical story without changing it. First, the destructive-averaging theorem is not only provable but numerically exact in the implemented code path. Second, the main hard-regime gain of FACET is broad across families and silos. Third, family-level adaptation curves show that the method is genuinely few-shot. Fourth, the mechanism ranking is stable: the most important ingredients are the probe-aligned signatures and the atlas itself. These are precisely the properties that one would want from a method whose central claim is that the right shared object in federated multi-agent learning is the geometry of coordination conventions rather than a flat policy average.



(a) OCV2-Fed **(b) ProcKitchen-Fed** **(c) NAHT-Fed**

Figure 11: Family-wise ablation heatmaps. The same qualitative ranking appears across benchmark families: the largest drops come from removing probe alignment or the atlas itself.