PAIRCAST: DIRECT COUNTERFACTUAL PACKET VALUE FOR MULTIRATE SWARM COMMUNICATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Freeze a swarm one control cycle before an occluded neighbor becomes collision-relevant: decoding a short update can redirect several receivers, while an equally novel packet can leave every action unchanged. Scheduling with an integrated controller and packet format therefore requires the control consequence of each available rate, not packet salience alone. Counterfactual packet value (CPV) is the delivery-adjusted, sender-local return gain per attempted bit between successful decoding of a candidate prefix and continued silence. PAIRCAST estimates CPV for nested 1–8-bit prefixes by branching a permutation-equivariant neighborhood world model into paired multi-step futures, then admits the highest-density prefix through a shared reservation price and a one-step-lagged occupancy signal. On 20-agent formation under occlusion, 30-agent dynamic coverage, and 10-versus-3 pursuit-evasion, it reaches 95.9–97.4% normalized return with roughly 48% of Dense Broadcast traffic; at the matched 48.2% operating point, it leads Adaptive self-triggering by 1.9–3.4 taskwise points and a SchedNet-style multirate scheduler by 2.1 pooled points [1.0, 3.1]. Consequence estimation is predictive on 45000 held-out interventions (0.66 Spearman correlation, 78.4% sign accuracy, and 0.18 normalized decision regret), while replacing it with uncertainty scoring costs 4.8 return points and reducing its paired horizon from five steps to one costs 2.1. Shared-prefix execution retains a 1.6-point advantage within an 8.9 ms A100 p99 envelope, and matched-load gains remain 1.8–2.2 points throughout the communication-scarce 40–60% regime. Directly valuing how a decode changes the team's future provides a practical scheduling principle for multirate coordination under shared-medium contention.

1 INTRODUCTION

Consider two futures from the same pre-transmission swarm state. In the first, a two-bit prefix about an occluded agent is decoded by its current neighbors; their next actions open the formation and avoid a collision. In the second, the sender remains silent, receiver beliefs age, and the same reactive controller closes the gap too late. The scheduling decision is the difference between those futures. Neither the packet's magnitude nor the sender's uncertainty specifies it.

This distinction matters after a communication system is integrated. Packet semantics, controller code, decoder, and radio framing may already be fixed, leaving each sender to select silence or one of several payload rates from local information. Partial observability makes occasional updates decisive, while shared-medium contention makes indiscriminate transmission costly (Brambilla et al., 2013; Dorigo et al., 2021). A useful release score must therefore compare the downstream coordination induced by each decoded prefix and charge for both delivery risk and attempted bits.

Existing rules expose complementary but different objects. Event and innovation triggers measure state change, uncertainty rules measure ambiguity, and slot schedulers learn relative access priority (Tabuada, 2007; Heemels et al., 2012; Dimarogonas et al., 2012; Kim et al., 2019; Hu et al., 2023). Gated-ACML uses a communication-conditioned action-value gap for a binary gate (Mao et al., 2020), and CMVC values requested teammate messages at a receiver before aggregation (Gao et al., 2025). The fixed sender-side multirate decision instead requires the multi-step return effect of

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

successfully decoding each available prefix rather than receiving nothing, expressed per attempted bit under current delivery conditions.

Counterfactual packet value (CPV) makes this quantity explicit. For every 1–8-bit prefix, PAIRCAST forks the same sender-local latent and stale neighbor beliefs into a successful-decode branch and a silent branch. A permutation-equivariant world model propagates receiver reactions, physical actions, topology, and common reward for five steps. Candidate delivery probability enters once, after the conditional return difference, and a shared reservation price chooses the highest-value prefix or silence. Execution uses only local history, decoded prefixes and their ages, the previous rate, and one lagged occupancy scalar.

The estimator is directly testable. On 45000 held-out snapshot–payload interventions, predicted successful-decode value attains 0.66 rank correlation, 78.4% sign accuracy, and 0.18 normalized regret. That predictive signal matters downstream: replacing consequence estimation with uncertainty scoring costs 4.8 mean return points, and shortening the paired horizon from five steps to one costs 2.1. At 48.2% of Dense Broadcast traffic, PAIRCAST reaches 97.4, 96.3, and 95.9 normalized return on formation, coverage, and pursuit, leading Adaptive self-triggering by 1.9–3.4 points and SchedNet-style scheduling by 2.1 pooled points [1.0, 3.1]. Shared-prefix execution preserves a 1.6-point gain at 8.9 ms A100 p99, and the ordering is reflected in collision, discovery, and capture outcomes.

The technical advance is to make direct packet consequence both estimable and operational under a fixed sender-side rate interface. The same quantity drives nested payload selection, admits validation against replayed interventions, and supports comparisons from critic accuracy through runtime-bounded control. Gains persist across the 40–60% traffic range and across MPE, heterogeneous-agent, shifted-channel, packet-trace, and ROS 2 studies, establishing successful-decode-versus-silence value as a useful principle for bandwidth-limited multi-agent control.

2 RELATED WORK

Event and self-triggered control. Classical event-triggered methods release updates when a stability or innovation condition crosses a threshold; self-triggered variants predict the next release time (Tabuada, 2007; Heemels et al., 2012; Dimarogonas et al., 2012). ETCNet learns event gates under a bandwidth constraint and retains stale messages through zero-order holding (Hu et al., 2023). Their release statistics capture deviation or urgency. CPV targets a different quantity: the downstream team-return effect of decoding a particular rate candidate under current channel load.

Learned bandwidth allocation. CommNet, DIAL, IC3Net, TarMAC, graph communication, and information-bottleneck methods jointly learn communication and control (Sukhbaatar et al., 2016; Foerster et al., 2016; Singh et al., 2019; Das et al., 2019; Niu et al., 2021; Wang et al., 2020). SchedNet ranks agents for top-k access (Kim et al., 2019), while ETCNet couples an event gate to receiver-side holding. Our experiments preserve these decision structures in multirate adaptations but hold the action policy and codec fixed, separating scheduling from learning a new communication representation.

Value-based message selection. Gated-ACML trains a binary gate from $\Delta Q = Q(\text{communication-conditioned actions}) - Q(\text{independent action})$ and adjusts a threshold to prune a target fraction of messages (Mao et al., 2020). CMVC predicts each teammate message's value from a receiver's local observation, requests messages with positive value, and uses those values in aggregation (Gao et al., 2025). Their native decision contracts differ from ours in ways that affect the estimand: Gated-ACML is binary and conditions on communication-dependent physical actions, while CMVC moves selection to receivers and couples value to message aggregation. The contract studied here requires each sender, without current receiver observations, to choose among nine rates before channel realization. CPV therefore evaluates that complete rate set through successful broadcast decode versus silence, spans H reactive control steps, and includes candidate-dependent delivery and attempted-bit normalization. The closest contract-preserving empirical controls separate generic learned rate scoring (Learned Trigger+Rate), uncertainty-based salience, and learned access priority (SchedNet-style); none receives a changed decoder, action policy, or receiver-request path.

Counterfactual prediction. World models use latent imagination for action planning and policy learning (Ha & Schmidhuber, 2018; Hafner et al., 2020; 2025); graph models provide permutation-aware dynamics for robot teams (Scarselli et al., 2009; Kipf & Welling, 2017; Veličković et al., 2018; Jiang et al., 2020; Li et al., 2020; Tolstaya et al., 2020). PAIRCAST uses imagination to estimate a communication intervention while a reactive controller selects physical actions. Shared-randomness replay then calibrates the predicted successful-decode/silence difference directly, making the novelty-bearing quantity empirically testable.

3 PROBLEM FORMULATION

Our setting is a decentralized partially observable Markov decision process (Dec-POMDP) in which N agents receive the same reward. At timestep t, the local observation available to agent i is $o_t^i \in \mathcal{O}$; the agent then chooses an action $a_t^i \in \mathcal{A}$. The team maximizes the following discounted return:

$$J = \mathbb{E}\left[\sum_{t=0}^{T} \gamma^t r(s_t, \mathbf{a}_t)\right], \tag{1}$$

where s_t is the global state, $\mathbf{a}_t = (a_t^1, \ldots, a_t^N)$ is the joint action, and $\gamma \in (0, 1)$ is the discount factor.

Each agent i maintains a recurrent latent state $z_t^i \in \mathbb{R}^d$ encoding its observation history:

$$z_t^i = f_\theta(z_{t-1}^i, o_t^i, \tilde{m}_{t-1}^{\to i}), \tag{2}$$

where f_θ implements a GRU-based encoder, while $\tilde{m}_{t-1}^{\to i}$ is the set, possibly empty, of messages received at the previous step.

At each step, agent i chooses silence or a payload of $b_t^i \in \{1, \ldots, B_{\max}\}$ bits. A transmission is one local broadcast to the current recipient set $\mathcal{R}_t^i = \mathcal{N}_t(i)$: one channel event delivers the same decoded prefix to every current recipient, or delivers it to none. This coupled-broadcast convention is shared by all schedulers and paired interventions. The delivery model is

$$P(\text{deliver} \mid b_t^i, L_t) = \sigma\big(-\alpha \cdot b_t^i - \beta \cdot L_t + \kappa\big), \tag{3}$$

where $L_t = \sum_j b_t^j$ is attempted payload load and (α, β, κ) are fitted by maximum-likelihood logistic regression on held-out payload/load sweeps for each task. Runtime decisions receive only the previous interval's attempted-bit total, $\hat{L}_t = \sum_j b_{t-1}^j$. A candidate replaces the sender's previous contribution:

$$\hat{L}_t^i(b) = [\hat{L}_t - b_{t-1}^i]_+ + b, \qquad p_t^i(b) = \sigma(-\alpha b - \beta \hat{L}_t^i(b) + \kappa). \tag{4}$$

Simultaneous peer changes leave residual error $\xi_t^i = \sum_{j \neq i}(b_t^j - b_{t-1}^j)$; thus the rule is decentralized in action selection and uses one lagged network scalar rather than current joint decisions.

Agent i maintains at the sender a prediction $\mu_t^{i \to j}$ of each current receiver j and its message age a_t^{ij}; the arrow denotes the candidate delivery direction, not access to j's runtime state. The belief updater U_ψ consumes i's local observation, the last prefix successfully decoded on the corresponding link, a delivery mask, and age. Without a new packet it predicts forward and increments age up to 32. Let $\mathcal{B}_t^i = \{(\mu_t^{i \to j}, a_t^{ij})\}_{j \in \mathcal{N}_t(i)}$. Runtime scheduling depends on $x_t^i = (z_t^i, \mathcal{B}_t^i, \hat{L}_t, b_{t-1}^i)$ and never on current peer latents, global state, future channel outcomes, or simultaneous peer choices.

For a candidate $b > 0$, paired rollouts update every $j \in \mathcal{R}_t^i$ and estimate the common-return gain conditional on successful delivery:

$$\Delta J_{t,\text{succ}}^i(b) = \hat{V}(x_t^i; \text{deliver}(Q_b)) - \hat{V}(x_t^i; \text{silent}). \tag{5}$$

We define CPV as the sender-local direct-benefit density: expected common-return change per attempted payload bit attributable to successful decoding of i's candidate. The paired branches share contemporaneous peer actions, attempted rates, and channel outcomes, then permit all policies to react from $t{+}1$ onward. CPV therefore includes downstream receiver and topology reactions while isolating the current candidate from simultaneous traffic. The reservation price converts this estimand into the operational value-per-bit score:

$$\begin{aligned} &\text{CPV}_t^i(b) = \frac{p_t^i(b) \Delta J_{t,\text{succ}}^i(b)}{b + \epsilon}, \\ &\text{VPB}_t^i(0) = 0, \qquad \text{VPB}_t^i(b) = \text{CPV}_t^i(b) - \lambda_{\text{bit}}, \quad b = 1, \ldots, 8. \end{aligned} \tag{6}$$

Delivery probability appears exactly once in CPV. With $\epsilon = 10^{-6}$, the rule first finds the most efficient positive-rate candidate, $b_{\text{dens}} = \arg\max_{b>0} \text{CPV}_t^i(b)$, and sends it only if its density exceeds λ_{bit}; otherwise it remains silent. Thus the price controls admission while CPV orders positive rates. The shared price is selected once on the pooled calibration split to meet the workload traffic target and then frozen. This density-threshold policy targets the directly attributable decode benefit. Contemporaneous displacement of other packets is outside per-sender CPV; the common lagged-load input adapts subsequent admission to aggregate contention.

4 PAIRCAST: PAIRED COUNTERFACTUAL PACKET VALUATION

PAIRCAST operationalizes CPV with a sender-local neighborhood model, a nested payload, and a paired communication critic. The action policy π_{act} remains reactive while imagination evaluates the communication intervention. Figure 1 summarizes the data flow, and Algorithm 1 gives the execution loop.

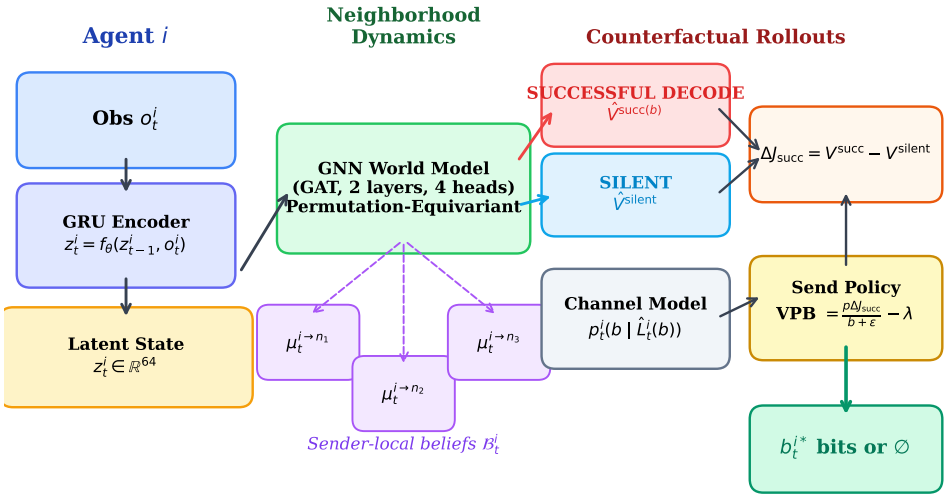

Figure 1: PAIRCAST data flow. Local observations and successfully received prefixes update sender-local neighbor beliefs. Paired H-step rollouts compare a successful candidate decode with silence, producing ΔJ_{succ}. Candidate-conditioned delivery converts this difference into CPV; subtracting the reservation price selects a nested payload or silence while the control policy remains reactive.

4.1 PERMUTATION-EQUIVARIANT NEIGHBORHOOD WORLD MODEL

The local graph is $\mathcal{G}_t^i = (\{z_t^i\} \cup \{\mu_t^{i\to j}, a_t^{ij}\}_{j\in\mathcal{N}_t(i)}, E_t^i)$. Before imagination, U_ψ advances each belief from the sender's observation, last decoded prefix, delivery mask, and age. Its one-step training loss is

$$\mathcal{L}_{\text{bel}} = \mathbb{E}\left[\frac{\sum_{j\in\mathcal{N}_t(i)} \left\| U_\psi(\mu_t^{i\to j}, o_t^i, \tilde{m}_t^{j\to i}, d_t^{ji}, a_t^{ij}) - \text{sg}(z_{t+1}^{j,\text{tgt}}) \right\|_2^2}{|\mathcal{N}_t(i)| + 1}\right]. \tag{7}$$

The rollout then uses the shared graph-attention transition

$$\hat{\mu}_{t+k}^{i\to j} = g_\phi\left(\hat{\mu}_{t+k-1}^{i\to j}, \sum_{l\in\mathcal{N}(j)} \alpha_{jl} W_v \hat{\mu}_{t+k-1}^{i\to l}\right), \tag{8}$$

where α_{jl} are attention weights and the shared GRU transition preserves permutation equivariance. At each imagined step, the frozen reactive policy maps predicted latents to physical actions; g_ϕ

predicts the next local graph and observations, recomputes distance/obstacle-defined edges, and a permutation-invariant reward head predicts the environment's shared team reward. Centralized training supplies stop-gradient peer targets from an exponential-moving-average encoder; these targets are unavailable to and removed from execution. The rollout loss is

$$\mathcal{L}_{\mathrm{wm}} = \frac{1}{H}\sum_{k=1}^{H}\left[\frac{\sum_j \|\hat{\mu}_{t+k}^{i\to j} - \mathrm{sg}(z_{t+k}^{j,\mathrm{tgt}})\|_2^2}{|\mathcal{N}(i)|+1} + \lambda_r(\hat{r}_{t+k} - r_{t+k})^2 + \lambda_\pi D_{\mathrm{KL}}\big(\mathrm{sg}\big[\pi_{\mathrm{act}}(\cdot \mid z_{t+k}^{\mathrm{tgt}})\big] \,\|\, \pi_{\mathrm{act}}(\cdot \mid \hat{\mu}_{t+k})\big)\right]. \tag{9}$$

4.2 Nested Maximum-8-Bit Payload

The shared interface uses a depth-8 binary tree codebook. For projection $y = W_m z_t^i$, each prefix $u \in \{0,1\}^b$ has reconstruction e_u; the selected leaf minimizes average distortion over all depths. Thus $Q_b(y) = c_{1:b}^\star$ and $D_b(Q_b(y)) = e_{c_{1:b}^\star}$, making every 1–8-bit prefix independently decodable. Packet framing carries b, while Bits% counts payload bits only. The nested codebooks use VQ commitment weight 0.25, EMA decay 0.99, recursive 2-means initialization, and dead-node reseeding after 1,000 unused assignments. Strict multirate comparisons expose the identical action set $\{0,\dots,8\}$ and decoder to PairCast, SchedNet-style, ETCNet-style, and Learned Trigger+Rate schedulers. The quantizer objective is

$$\mathcal{L}_Q = \mathbb{E}_{b\sim\mathrm{Unif}\{1,\dots,8\}}\left[\|\mathrm{sg}(y) - e_{c_{1:b}^\star}\|_2^2 + 0.25\|y - \mathrm{sg}(e_{c_{1:b}^\star})\|_2^2\right], \tag{10}$$

with EMA updates for reconstruction centroids. This loss trains every decodable depth rather than only the 8-bit leaf.

4.3 Estimating Counterfactual Packet Value

For each $b > 0$, the send intervention forces $D_b(Q_b(W_m z_t^i))$ into every predicted receiver update:

$$\hat{\mu}_{t+1}^{i\to j,\mathrm{succ}(b)} = f_\theta(\hat{\mu}_t^{i\to j}, \hat{o}_{t+1}^{i\to j}, D_b(Q_b(W_m z_t^i))). \tag{11}$$

The shared silent branch receives no candidate:

$$\hat{\mu}_{t+1}^{i\to j,\mathrm{silent}} = f_\theta(\hat{\mu}_t^{i\to j}, \hat{o}_{t+1}^{i\to j}, \varnothing). \tag{12}$$

Their conditional value difference is

$$\Delta J_{t,\mathrm{succ}}^i(b) = \sum_{k=1}^{H}\gamma^k(\hat{r}_{t+k}^{\mathrm{succ}(b)} - \hat{r}_{t+k}^{\mathrm{silent}}). \tag{13}$$

These updates are applied for every $j \in \mathcal{R}_t^i$. They are the recurrent receiver-update path of U_ψ; f_θ denotes its shared GRU cell, while the observation, delivery-mask, and age features enter through U_ψ's input projection. The rollout does not sample first-step delivery; it evaluates the conditional successful-decode branch. Both branches use the same reward head, frozen policy parameters, initial local graph, and event-keyed randomness. The physical channel samples one coupled broadcast-delivery event per sender only after all agents choose their rates.

4.4 Bitrate-Aware Send Policy

The silent rollout is computed once and reused across the eight positive-rate candidates. Training samples from

$$\pi_{\mathrm{send}}(b \mid x_t^i) = \frac{\exp(\mathrm{VPB}_t^i(b)/\tau)}{\sum_{b'=0}^{8}\exp(\mathrm{VPB}_t^i(b')/\tau)}, \tag{14}$$

where $\tau = 0.05 + 0.95\sigma(\eta_\tau)$ is initialized to 0.5; evaluation uses the deterministic rule following Eq. 6. A synchronized burst raises the next occupancy observation, decreases $p_t^i(b)$, and increases the effective reliability-adjusted threshold $\lambda_{\mathrm{bit}}(b+\epsilon)/p_t^i(b)$. This negative feedback suppresses low-value packets without exposing simultaneous choices.

CPV measures decision consequence instead of packet salience. A large latent innovation can be immaterial to formation geometry, whereas a small occluded-neighbor update can change several agents' predicted collision risk. Innovation and uncertainty scores favor the former; paired reward rollouts identify the latter when its decoded prefix separates the successful-decode and silent futures.

5 TRAINING AND OPTIMIZATION

Training alternates model and policy blocks within each on-policy iteration. The model block minimizes

$$\mathcal{L}_{\text{model}} = \mathcal{L}_{\text{wm}} + \mathcal{L}_{\text{bel}} + \mathcal{L}_Q + \lambda_c \mathcal{L}_{\text{calib}}, \qquad \mathcal{L}_{\text{calib}} = \mathbb{E}[(\Delta J^i_{t,\text{succ}}(b) - \text{sg}(\Delta J^{i,\text{replay}}_{t,\text{succ}}(b)))^2], \quad (15)$$

with $\lambda_r = 1$, $\lambda_\pi = 0.1$, and $\lambda_c = 0.5$. Each replay target averages $K = 4$ paired interventions from the same pre-decision snapshot. The snapshot includes physical state, all recurrent and belief states, message ages, receive queues, previous bitrates, $\hat{L}_t$, channel state, episode clock, termination flags, frozen parameter versions, and counter-based environment, observation, action, send, and channel RNG streams. At the intervention step, other agents' actions, rates, and packet outcomes are fixed; the send branch forces the candidate's coupled broadcast delivery and the silent branch omits it. From the next step onward, frozen policies react independently in both branches under event-keyed shared randomness. The target is therefore conditional on successful delivery and contains no extra factor of $p^i_t(b)$. The 45000-pair critic audit uses held-out snapshots excluded from model updates, calibration-loss training, threshold selection, and checkpoint selection.

The policy block uses PPO (Schulman et al., 2017) with joint action/send likelihood ratio, $\gamma = 0.99$, GAE $\lambda = 0.95$, clip and value-clip 0.2, value coefficient 0.5, action/send entropy coefficients 0.010/0.005, four epochs, and global gradient clipping at 0.5. Adam uses learning rate 3×10^{-4}, $(\beta_1, \beta_2) = (0.9, 0.999)$, and $\epsilon = 10^{-5}$. We collect 2,048 agent-transitions per iteration, optimize recurrent sequences of length 32 with 8 burn-in steps and 256 scored transitions per minibatch, retain 1,048,576 transitions for the model, and perform eight model updates on 64-sequence batches. The model/codebook are frozen during PPO; the target encoder uses EMA decay 0.995 and stop-gradient targets. PPO anchors the runtime encoder, while commitment loss, EMA centroids, and dead-node reseeding prevent quantizer collapse. The pooled reservation price follows projected load-calibration updates with step size 10^{-3} and is frozen for evaluation.

The curriculum uses 200K steps at $N = 5$, 300K at $N = 15$, and 500K at full scale, progressively introducing calibrated loss, occlusion, and topology shifts. The latent size is 64, the world model has two four-head graph-attention layers, and $H = 5$. Training uses 8 NVIDIA A100 80GB GPUs with CUDA 12.4 and approximately 72 GPU-hours per task. Checkpoint provenance follows a two-stage replacement protocol. For each seed, joint PPO first produces one PAIRCAST-trained shell containing the action policy, recurrent encoder, belief updater, and codec. That exact shell is cloned and frozen for every scheduler fit; only the release/rate module and its calibration parameters are reinitialized and optimized with the method-specific objective. The fixed-interface estimand is therefore scheduler replacement within a common deployed representation: paired episodes compare release rules with tensor-identical control and message semantics for each seed. Unless stated otherwise, latency is per-agent batch-1 inference; the runtime-budget study additionally reports end-to-end p50/p95/p99 on A100, CPU, and Jetson Orin NX.

6 EXPERIMENTS

6.1 EXPERIMENTAL SETUP

FC-20 asks 20 agents to maintain formation behind occluders; its shared reward combines formation error and collision cost, and an episode breaches when the simulator's collision-safety termination fires. DC-30 asks 30 agents to cover moving targets while penalizing redundant search; a breach is an uncovered target at the task deadline. PE-10v3 asks 10 pursuers to capture three evasive targets under distance- and obstacle-dependent topology; a breach is an uncaptured target at the episode timeout. Observations are variable-cardinality local sets rather than fixed-width global states: each agent receives its own motion state plus relative features for visible neighbors, obstacles, and task objects. The reactive physical action is a bounded two-dimensional motion command, while communication contributes only a decoded nested prefix. Scenario generation is seed-disjoint across training, pooled calibration, and held-out evaluation, with each complete geometry/motion configuration assigned to one split so that steps from one trajectory never cross splits. Dense Broadcast normalization reruns the identical held-out scenarios with every agent submitting the 8-bit leaf.

PettingZoo MPE Cooperative Navigation is a separate full-training transfer of the method, not a zero-shot reuse of home-task weights: all methods train on MPE training layouts with standard

dynamics, reward, collision rules, and a 25-cycle horizon, then freeze their selected checkpoints for the held-out layouts in Table 22. The packet-trace study replays recorded delivery opportunities through the same frozen reactive shell. For method m and trace window w, let $R_w(m)$ be the accumulated shared control reward after packets arriving beyond the recorded control deadline are omitted. Its reported score is $100 \sum_w R_w(m) / \sum_w R_w(\mathrm{Dense})$ on the same windows; misses count selected packets that fail this deadline. The ROS 2 study executes the same message and action contract with serialization, transport stubs, and `tc netem` queue/loss shaping.

Strict baselines share the frozen reactive-controller/encoder/codec tensors, sender-local inputs, fitted delivery model, nested payload decoder, lagged occupancy scalar, and tuning budget. Adaptive self-triggered (AST) releases the fixed 8-bit leaf when $\psi_t^i = w_\Delta \|\Delta z_t^i\|_2 + w_\tau \delta_t^i - w_p(1 - p_t^i(8)) > \eta$; its weights and threshold are validation-selected. Learned Trigger+Rate optimizes rollout-free PPO logits over $\{0, \ldots, 8\}$ with the same send-entropy coefficient. The SchedNet-style adaptation preserves a PPO-trained priority head and top-k access (Kim et al., 2019), with k calibrated to the target and a shared rate head applied to admitted senders. The ETCNet-style adaptation preserves event gating and receiver-side zero-order holding (Hu et al., 2023), then applies the same multirate head. Each learned scheduler receives 36 tuning trials and is selected by validation return subject to the target-load tolerance. Lyapunov–Deadband, binary Learned Trigger, Periodic, and Uncertainty Threshold provide additional anchors; TarMAC and IC3Net remain richer-interface context.

The expanded scheduler, budget, runtime, task-outcome, transfer, and channel studies use 10 independent training seeds; mechanism and diagnostic contrasts use the designated five-seed subset where their captions say so. Unless a table states another denominator, evaluation uses 14,000 paired held-out episodes (4,800/4,600/4,600 by task). Returns are normalized per task to Dense Broadcast $= 100$. Bits% is attempted payload bits divided by 8 times the same agent-step denominator; pooled return means are agent-timestep weighted, whereas explicitly labeled equal-task effects average the three taskwise paired differences. Each budget is calibrated once over the pooled validation mixture without per-task retuning. Standard errors are across training-seed means. Paired 95% intervals use a hierarchical bootstrap that resamples training seeds as the top-level independent unit and paired episodes within each selected seed; critic intervals replace episodes with paired snapshots. Holm-adjusted q values cover the comparator family named in each caption. Tables 1, 14, 18, 19, and 21 all reuse the same canonical 10-seed checkpoint–episode ledger when they repeat a value. In five-seed diagnostic tables, an unchanged default/reference row that exactly repeats a main closure row is the canonical 10-seed point estimate shown for orientation; diagnostic deltas, intervals, and ablated alternatives come from the fixed five-seed subset with common held-out episode IDs. Table 1 reports breaches per 1,000 episodes; appendix tables give raw numerator/denominator counts.

6.2 RQ1: Does paired valuation improve return at equal load?

Yes. At 48.2% pooled traffic, PairCast exceeds every evaluated same-interface scheduler on all three tasks (Table 1). The shared-randomness AST closure gives paired gaps of 1.9 [0.9, 2.8], 3.2 [2.1, 4.3], and 3.4 [1.8, 4.9] points on FC-20, DC-30, and PE-10v3; all remain favorable after Holm correction. Among the evaluated published-family adaptations, SchedNet-style is strongest and trails by 2.1 pooled points [1.0, 3.1], $q = 0.028$.

Table 1: Matched-load scheduler closure at 48.2% pooled traffic over 10 training seeds. Entries are normalized return ± seed SE; Bits is realized pooled attempted load and breaches are per 1,000 episodes. All rows use the same frozen controller and maximum-8-bit interface. Latency is median per-agent batch-1 A100 inference; runtime-budgeted end-to-end p99 comparisons appear in Table 20.

Method	FC-20	DC-30	PE-10v3	Bits%	Breach	ms
Lyapunov–Deadband	94.2 ± 1.3	91.1 ± 1.6	90.2 ± 1.8	48.2	2.9	3.8
Learned Trigger	94.6 ± 1.1	91.9 ± 1.4	91.2 ± 1.6	48.2	2.3	7.2
AST	95.5 ± 0.9	93.1 ± 1.0	92.5 ± 1.2	48.3	1.9	4.9
SchedNet-style	95.7 ± 0.9	93.7 ± 1.0	92.9 ± 1.3	48.1	1.6	6.4
Learned Trigger+Rate	95.8 ± 0.9	93.8 ± 1.1	93.1 ± 1.3	48.1	1.8	8.1
PairCast	97.4 ± 0.6	96.3 ± 0.7	95.9 ± 0.9	48.2	1.2	12.6

The ranking persists across load targets rather than peaking at the natural operating point. Figure 2 shows 30, 40, 48.2, 60, and 70% pooled targets; PAIRCAST's margin over SchedNet-style is 2.2 points at 40%, 2.1 at 48.2%, and 1.8 at 60%. The 30% and 70% intervals overlap zero, yielding the expected compression when packets are either extremely scarce or plentiful. Appendix Table 18 gives taskwise loads, intervals, and raw counts.

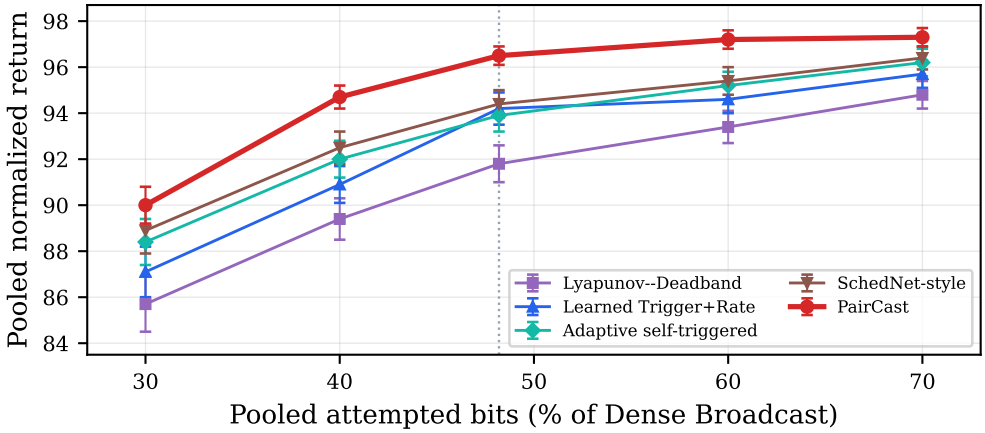

Figure 2: Pooled matched-load frontier over 10 seeds and 14,000 paired episodes per target. Curves show normalized return and seed SE; each method is tuned independently to the displayed target using the shared controller, nested payload, and lagged-load input.

6.3 RQ2: DOES COUNTERFACTUAL PACKET VALUE EXPLAIN THE GAIN?

Yes. The five-seed diagnostic asks two different questions. Replacing CPV with uncertainty scoring tests consequence estimation against packet salience and lowers mean return by 4.8 points; holding the paired intervention fixed and shortening its horizon from five to one step isolates temporal depth and costs 2.1 points. Both paired intervals are positive in Appendix Table 10, and every variant retains the same nested decoder and candidate-rate set. On 45000 held-out replay snapshot–payload interventions, the full critic reaches 0.66 Spearman correlation, 78.4% sign accuracy, and 0.18 normalized regret. Removing replay calibration increases regret to 0.36, so better successful-decode/silence estimates coincide with better packet choices and return.

Table 2: Mechanism diagnostics at matched load. The full row repeats the canonical 10-seed Table 1 estimate for orientation; ablated rows summarize the fixed five-seed diagnostic subset. Table 10, rather than cross-row SE subtraction here, reports inference from paired five-seed checkpoint–episode contrasts. Every row shares the nested rate interface; no-rollout replaces consequence estimation with uncertainty scoring.

VARIANT	FC-20	DC-30	PE-10V3	MEAN
Full PAIRCAST (H=5)	97.4 ± 0.6	96.3 ± 0.7	95.9 ± 0.9	96.53 ± 0.4
No-rollout (uncertainty trigger)	92.6 ± 1.4	91.5 ± 1.5	91.1 ± 1.7	91.73 ± 0.6
Short horizon (H=1)	95.1 ± 0.9	94.3 ± 1.0	93.9 ± 1.2	94.43 ± 0.5

6.4 RQ3: DOES CPV REMAIN USEFUL UNDER SYSTEMS AND DISTRIBUTION SHIFTS?

Across fixed runtime envelopes, shared-prefix CPV reuses the silent branch and common rollout states, vectorizes rate candidates, and stops candidates that cannot exceed the current score. The H=3 implementation attains 95.5 return at p99 8.9 ms on A100, exceeding AST by 1.6 points [0.8, 2.4] while both satisfy the 9 ms envelope. At comparable load, the H=1, shared H=3, and shared H=5 implementations reach 94.4, 95.5, and 96.5 return; the direct paired H=5 versus H=1 test in Table 10 supports the temporal-depth interpretation of this progression. CPU and Jetson measurements preserve the shared-prefix advantage (Appendix Table 20).

The return gains correspond to task outcomes: at equal 48.2% load, PAIRCAST records 1.1 FC-20 collisions per 1,000 steps versus 1.4 for SchedNet-style, reaches DC-30 targets in 30.1 versus 33.0 steps, and captures 85.1% of PE-10v3 episodes versus 82.2%. On separately trained MPE Cooperative Navigation, its native-return margin over AST is 1.7–1.8 points at 40–48.2% load. Moderate off-family channels preserve 1.1–2.4-point margins, packet-trace replay yields the highest normalized score and the fewest deadline misses, and ROS 2 execution retains a 1.7-point lead over AST. Together, critic calibration, operational outcomes, and runtime-bounded comparisons form one evidence chain: estimating decode consequence improves packet choice, those choices alter coordination events, and shared-prefix execution preserves the gain within measured system envelopes.

7 DISCUSSION

Direct packet-consequence valuation changes how communication scarcity is represented: the scheduler spends bits on predicted coordination effects rather than on locally conspicuous state changes. The calibration audit makes that decision variable observable, the matched-load study connects it to fewer collisions, faster discovery, and more captures, and shared-prefix inference shows that multi-step valuation can remain useful under explicit latency envelopes. The principle is broader than any one trigger architecture: when an integrated multi-agent system exposes several decodable rates, successful-decode-versus-silence value provides a common currency for comparing them.

The measured operating regime also clarifies the approximation. Shared-prefix $H{=}3$ retains a corrected advantage within the 9 ms A100 envelope, the 8-bit nested payload gives the strongest measured return–latency balance, and stale sender-local beliefs remain within 0.7 points of the current-peer oracle. CPV attributes subsequent receiver and policy reactions to one successful decode but does not price contemporaneous displacement of peer packets, and the replacement study is conditional on a common PAIRCAST-trained shell. Moderate channel shifts preserve positive margins; under severe queue saturation, successful reception becomes too rare for packet selection to separate methods. Thus the evidence supports CPV where communication is scarce yet actionable, while joint congestion valuation and scheduler-neutral shell pretraining define natural extensions.

8 BROADER IMPACT

Efficient swarm communication may help search-and-rescue, environmental monitoring, and precision agriculture by reducing coordination bandwidth. The same capability could also support military or surveillance deployments involving large autonomous swarms. Responsible use should therefore emphasize civilian deployment boundaries, human oversight, and operational constraints.

9 CONCLUSION

PAIRCAST turns a multirate communication decision into a paired question: how much team return does successful decoding of this prefix add relative to silence, per attempted bit? A calibrated neighborhood world model answers that question from sender-local state and a lagged occupancy signal. Under a tensor-identical control and message shell, PAIRCAST reaches 95.9–97.4% normalized return at roughly 48% of Dense Broadcast traffic, exceeds Adaptive self-triggering by 1.9–3.4 taskwise points, and retains its matched-load advantage from 40–60% traffic. Critic calibration, temporal-depth diagnostics, operational outcomes, runtime envelopes, MPE, channel shifts, packet traces, and ROS 2 execution connect that advantage to packet choices that change subsequent coordination. Direct successful-decode-versus-silence value is therefore a concrete, measurable foundation for communication scheduling when a multi-agent system must allocate a fixed payload interface.

REPRODUCIBILITY STATEMENT

Sections 3–6 specify the runtime information pattern, nested quantizer, optimization, environments, baselines, splits, metrics, and statistical protocol; Appendix A.1 gives detailed pseudocode and the remaining tables report denominators and uncertainty.

AI USE DISCLOSURE

Generative AI tools assisted the manuscript-preparation workflow with drafting, substantive editing, consistency checks, citation formatting, and LaTeX/PDF formatting. This preparation workflow did not generate or modify experimental data, source code, figures, or reported analyses.

REFERENCES

Manuele Brambilla, Eliseo Ferrante, Mauro Birattari, and Marco Dorigo. Swarm robotics: A review from the swarm engineering perspective. *Swarm Intelligence*, 7(1):1–41, 2013. doi: 10.1007/s11721-012-0075-2.

Abhishek Das, Théophile Gervet, Joshua Romoff, Dhruv Batra, Devi Parikh, Mike Rabbat, and Joelle Pineau. TarMAC: Targeted multi-agent communication. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 1538–1546, 2019. URL https://proceedings.mlr.press/v97/das19a.html.

Dimos V. Dimarogonas, Emilio Frazzoli, and Karl H. Johansson. Distributed event-triggered control for multi-agent systems. *IEEE Transactions on Automatic Control*, 57(5):1291–1297, 2012. doi: 10.1109/TAC.2011.2174666.

Marco Dorigo, Guy Theraulaz, and Vito Trianni. Swarm robotics: Past, present, and future [point of view]. *Proceedings of the IEEE*, 109(7):1152–1165, 2021. doi: 10.1109/JPROC.2021.3072740.

Jakob N. Foerster, Yannis M. Assael, Nando de Freitas, and Shimon Whiteson. Learning to communicate with deep multi-agent reinforcement learning. In *Advances in Neural Information Processing Systems*, volume 29, pp. 2137–2145, 2016. URL https://proceedings.neurips.cc/paper_files/paper/2016/hash/c7635bfd99248a2cdef8249ef7bfbef4-Abstract.html.

Zihong Gao, Changsheng Qu, Yuanhui Hao, and Liangjun Ke. Communication in multiagent reinforcement learning via counterfactual message value. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 55(11):8152–8165, 2025. doi: 10.1109/TSMC.2025.3604230.

David Ha and Jürgen Schmidhuber. World models. *arXiv preprint arXiv:1803.10122*, 2018. URL https://arxiv.org/abs/1803.10122.

Danijar Hafner, Timothy Lillicrap, Jimmy Ba, and Mohammad Norouzi. Dream to control: Learning behaviors by latent imagination. In *International Conference on Learning Representations*, 2020. URL https://openreview.net/forum?id=S1lOTC4tDS.

Danijar Hafner, Jurgis Pasukonis, Jimmy Ba, and Timothy Lillicrap. Mastering diverse control tasks through world models. *Nature*, 640(8059):647–653, 2025. doi: 10.1038/s41586-025-08744-2.

W. P. M. H. Heemels, Karl Henrik Johansson, and Paulo Tabuada. An introduction to event-triggered and self-triggered control. In *Proceedings of the 51st IEEE Conference on Decision and Control*, pp. 3270–3285, 2012. doi: 10.1109/CDC.2012.6425820.

Guangzheng Hu, Yuanheng Zhu, Dongbin Zhao, Mengchen Zhao, and Jianye Hao. Event-triggered communication network with limited-bandwidth constraint for multi-agent reinforcement learning. *IEEE Transactions on Neural Networks and Learning Systems*, 34(8):3966–3978, 2023. doi: 10.1109/TNNLS.2021.3121546.

Jiechuan Jiang, Chen Dun, Tiejun Huang, and Zongqing Lu. Graph convolutional reinforcement learning. In *International Conference on Learning Representations*, 2020. URL https://openreview.net/forum?id=HkxdQkSYDB.

Daewoo Kim, Sangwoo Moon, David Hostallero, Wan Ju Kang, Taeyoung Lee, Kyunghwan Son, and Yung Yi. Learning to schedule communication in multi-agent reinforcement learning. In *International Conference on Learning Representations*, 2019. URL https://openreview.net/forum?id=HUAnBToP_a.

Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. In *International Conference on Learning Representations*, 2017. URL https://openreview.net/forum?id=SJU4ayYgl.

Qingbiao Li, Fernando Gama, Alejandro Ribeiro, and Amanda Prorok. Graph neural networks for decentralized multi-robot path planning. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 11785–11792, 2020. doi: 10.1109/IROS45743.2020.9341668.

Hangyu Mao, Zhengchao Zhang, Zhen Xiao, Zhibo Gong, and Yan Ni. Learning agent communication under limited bandwidth by message pruning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(04):5142–5149, 2020. doi: 10.1609/aaai.v34i04.5957.

Yaru Niu, Rohan R. Paleja, and Matthew C. Gombolay. Multi-agent graph-attention communication and teaming. In *Proceedings of the 20th International Conference on Autonomous Agents and Multiagent Systems*, pp. 964–973, 2021. URL https://www.ifaamas.org/Proceedings/aamas2021/pdfs/p964.pdf.

Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner, and Gabriele Monfardini. The graph neural network model. *IEEE Transactions on Neural Networks*, 20(1):61–80, 2009. doi: 10.1109/TNN.2008.2005605.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017. URL https://arxiv.org/abs/1707.06347.

Amanpreet Singh, Tushar Jain, and Sainbayar Sukhbaatar. Learning when to communicate at scale in multiagent cooperative and competitive tasks. In *International Conference on Learning Representations*, 2019. URL https://openreview.net/forum?id=rye7knCqK7.

Sainbayar Sukhbaatar, Arthur Szlam, and Rob Fergus. Learning multiagent communication with backpropagation. In *Advances in Neural Information Processing Systems*, volume 29, pp. 2244–2252, 2016. URL https://proceedings.neurips.cc/paper_files/paper/2016/hash/55b1927fdafef39c48e5b73b5d61ea60-Abstract.html.

Paulo Tabuada. Event-triggered real-time scheduling of stabilizing control tasks. *IEEE Transactions on Automatic Control*, 52(9):1680–1685, 2007. doi: 10.1109/TAC.2007.904277.

Ekaterina Tolstaya, Fernando Gama, James Paulos, George Pappas, Vijay Kumar, and Alejandro Ribeiro. Learning decentralized controllers for robot swarms with graph neural networks. In *Proceedings of the 3rd Conference on Robot Learning*, volume 100 of *Proceedings of Machine Learning Research*, pp. 671–682, 2020. URL https://proceedings.mlr.press/v100/tolstaya20a.html.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph attention networks. In *International Conference on Learning Representations*, 2018. URL https://openreview.net/forum?id=rJXMpikCZ.

Rundong Wang, Xu He, Runsheng Yu, Wei Qiu, Bo An, and Zinovi Rabinovich. Learning efficient multi-agent communication: An information bottleneck approach. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 9908–9918, 2020. URL https://proceedings.mlr.press/v119/wang20i.html.

A Algorithmic and Rollout Appendix

A.1 Pseudo-code

Algorithm 1 PairCast: Counterfactual Communication Scheduling

Require: Local encoder f_θ, belief updater U_ψ, world model g_ϕ, nested codec (Q_b, D_b)
Require: Channel parameters (α, β, κ), horizon H, reservation price λ_{bit}

for each timestep t **do**
  for each agent i (in parallel) **do**
    Encode observation: $z_t^i \leftarrow f_\theta(z_{t-1}^i, o_t^i, \tilde{m}_{t-1}^{\to i})$
    Update $\mathcal{B}_t^i$ from local observations and successfully received prefixes; read $\hat{L}_t$
    Rollout silence once: $\{\hat{r}_{t+k}^{\text{silent}}\}_{k=1}^H \leftarrow g_\phi(x_t^i, \varnothing)$
    Set $\text{VPB}(0) \leftarrow 0$
    for $b = 1, \ldots, 8$ **do**
      Encode prefix $m_t^i(b) \leftarrow Q_b(W_m z_t^i)$; set $\hat{L}_t^i(b) \leftarrow [\hat{L}_t - b_{t-1}^i]_+ + b$
      Rollout successful decode: $\{\hat{r}_{t+k}^{\text{succ}(b)}\}_{k=1}^H \leftarrow g_\phi(x_t^i, D_b(m_t^i(b)))$
      Compute $\Delta J_{t,\text{succ}}^i(b) \leftarrow \sum_{k=1}^H \gamma^k (\hat{r}_{t+k}^{\text{succ}(b)} - \hat{r}_{t+k}^{\text{silent}})$
      Compute $\text{CPV}(b) \leftarrow p_t^i(b) \Delta J_{t,\text{succ}}^i(b) / (b + \epsilon)$
      Compute $\text{VPB}(b) \leftarrow \text{CPV}(b) - \lambda_{\text{bit}}$
    end for
    Select $b_t^{i*} \leftarrow 0$ if $\max_{b>0} \text{VPB}(b) \le 0$; else select the positive-rate argmax
    if $b_t^{i*} > 0$ **then**
      Submit $m_t^i(b_t^{i*})$ to $\mathcal{R}_t^i$; channel samples one coupled broadcast event after all rate choices
    end if
    Select action $a_t^i \leftarrow \pi_{\text{act}}(z_t^i)$
  end for
end for

A.2 Additional Rollout Visualization

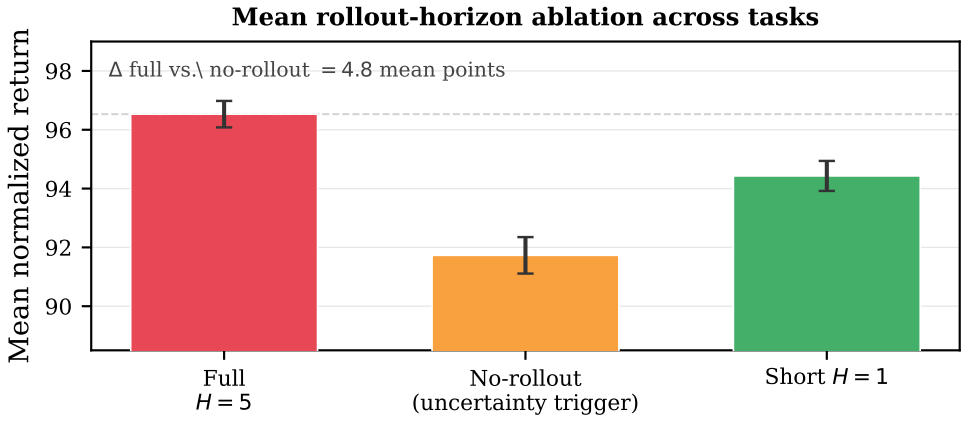

Figure 3: Across-task rollout diagnostics reproducing the orientation values from Table 2. The full bar is the canonical 10-seed reference and ablated bars use the five diagnostic seeds; paired five-seed effects and intervals are reported in Table 10.

The 4.8-point full-versus-uncertainty contrast tests consequence estimation against salience scoring as complete scheduler choices. The 2.1-point H=5-versus-H=1 contrast isolates temporal depth within the paired estimator.

B Operating-Range and Efficiency Appendix

B.1 FC-20 Stress Diagnostics

Appendix Figures 4 and 5 characterize FC-20 behavior under packet loss, topology change, swarm scaling, and severe loss.

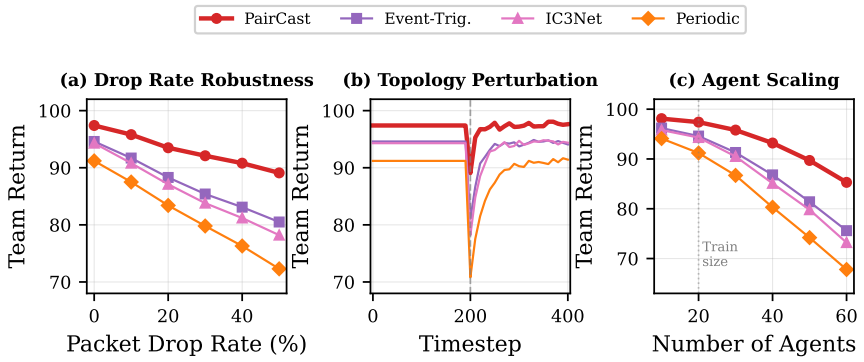

Figure 4: FC-20 stress diagnostics. Panel (a) varies packet drop from 0–50%; panel (b) applies a topology change at step 200; panel (c) evaluates 10–60 agents after training with 20.

The panels separate three failure sources: channel loss, an abrupt edge-set change, and graph-size extrapolation. Each plotted quantity is task return under its displayed FC-20 intervention, complementing the pooled matched-load test.

The moderate-drop ordering agrees with the matched-load study. After the topology perturbation, PairCast recovers more quickly than the non-adaptive reference, and its scaling curve degrades gradually beyond the 20-agent training size.

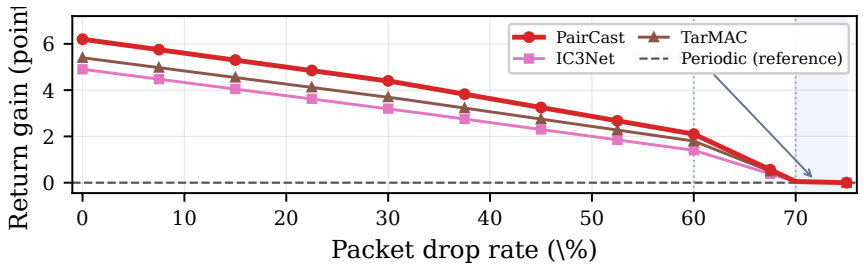

Figure 5: FC-20 normalized-return gain over Periodic as packet drop rises. The shaded 70–75% region marks convergence toward a near tie for all learned schedulers.

PairCast's margin contracts continuously as decoding opportunities disappear, linking the high-loss transition to the successful-decode intervention that CPV values.

At very high loss, all scheduler margins approach zero because decoding opportunities vanish; this transition localizes CPV's leverage to regimes in which packets can still alter receiver state.

B.2 Dense-Broadcast Efficiency Summary

The natural PairCast checkpoints cluster near 48% traffic across all three tasks, which motivates the central budget without task-specific retuning. Table 1 then holds traffic and communication semantics fixed to isolate scheduler quality.

Table 3: Natural operating points across three benchmarks over the canonical 10 training seeds. The PAIRCAST return cells are exact cross-references to the Table 1 checkpoint–episode estimates. Team Return is normalized to Dense Broadcast = 100 and Bits is attempted payload traffic as a percentage of Dense Broadcast.

	FC-20		DC-30		PE-10V3	
METHOD	RETURN	BITS%	RETURN	BITS%	RETURN	BITS%
Dense Broadcast	100.0 ± 0.0	100.0	100.0 ± 0.0	100.0	100.0 ± 0.0	100.0
Periodic (fixed $\Delta t = 5$)	91.2 ± 1.4	20.0	88.7 ± 1.8	20.0	85.3 ± 2.1	20.0
Uncertainty Thresh.	93.8 ± 1.1	38.4	91.2 ± 1.5	41.7	89.1 ± 1.7	44.2
Event-Trig. (Lyap.)	94.6 ± 0.9	42.1	92.8 ± 1.2	45.3	90.7 ± 1.4	48.6
TarMAC	95.1 ± 0.8	61.3	93.4 ± 1.0	58.9	91.8 ± 1.3	63.4
IC3Net	94.3 ± 0.9	47.8	92.1 ± 1.3	49.1	90.5 ± 1.5	52.7
PAIRCAST (ours)	97.4 ± 0.6	48.2	96.3 ± 0.7	47.8	95.9 ± 0.9	48.1

C MATCHED-LOAD AND MECHANISM ANALYSIS

C.1 TASK OUTCOMES AND COMPONENT ANALYSIS

Table 4 connects normalized return to collisions, formation error, coverage, discovery time, and capture at each natural operating point. Table 5 isolates five components on FC-20, while Table 21 supplies matched-load uncertainty for the task outcomes.

Table 4: Task-specific outcomes at each method's natural operating point. FC-20 reports collisions per 1,000 steps and mean ℓ_2 formation error; DC-30 reports coverage and discovery time; PE-10v3 reports capture rate. Table 21 supplies matched-load uncertainty and raw counts for the decisive methods.

	FC-20		DC-30		PE-10V3
METHOD	COLL. RATE↓	FORM. ERR↓	COV. %↑	DISC. TIME↓	CAP. RATE↑
Dense Broadcast	0.8	0.12	94.2	28.3	87.6
Periodic (fixed $\Delta t = 5$)	3.7	0.31	82.1	47.2	71.4
Uncertainty Thresh.	2.4	0.24	85.8	41.6	75.3
Event-Trig. (Lyap.)	2.1	0.21	87.4	39.1	77.8
TarMAC	1.8	0.18	88.9	36.4	79.2
IC3Net	2.2	0.22	86.3	40.8	76.1
PAIRCAST	1.1	0.14	92.7	30.1	85.1

The natural operating points connect normalized return to fewer collisions, lower formation error, faster discovery, and higher capture rate. Table 21 confirms the same behavioral improvements under matched traffic.

Table 5: FC-20 component diagnostics at comparable traffic. The full row is the canonical 10-seed reference; ablated rows use the fixed five-seed diagnostic subset. Removing paired pricing, multirate allocation, permutation equivariance, delivery modeling, or replay calibration reduces return; Table 2 analyzes rollout depth across all tasks.

VARIANT	RETURN	BITS%
Full PAIRCAST	97.4 ± 0.6	48.2
w/o rollout pricing (random gate)	89.6 ± 1.8	49.1
w/o bitrate allocation (binary send/silent)	94.1 ± 0.9	51.7
w/o permutation equivariance (MLP)	93.7 ± 1.2	48.9
w/o channel model (perfect channel)	95.8 ± 0.7	52.3
w/o value calibration loss	95.2 ± 1.0	50.4

On FC-20, removing bitrate allocation, permutation equivariance, delivery modeling, or replay calibration costs 1.6–3.7 return points at comparable traffic. Together with the cross-task horizon ablation, this decomposition shows that CPV depends on coordinated modeling of intervention value, payload rate, topology, and delivery.

C.2 Taskwise Traffic and Scheduler Closure

The following analyses reuse the reactive controller, maximum-8-bit nested payload, lagged-load scalar, and matched-load checkpoints from Table 1. Table 7 re-expresses the canonical 10-seed ledger, including the AST values repeated in Table 14; tables explicitly marked five-seed use the fixed diagnostic subset. Paired intervals use the hierarchical seed-then-episode resampling of Section 6, and each caption names its Holm family. Decision latency is median per-agent batch-1 inference on an NVIDIA A100 80GB.

Table 6: Taskwise traffic under the Table 1 pooled target. Bits% is attempted payload traffic relative to Dense Broadcast on the same held-out rollouts; the pooled column uses the agent-timestep-weighted denominator from Section 6. Realized taskwise loads remain within 0.8 points of the shared target.

Method	FC-20 Bits%	DC-30 Bits%	PE-10v3 Bits%	Pooled Bits%
Lyapunov–Deadband	47.6	49.0	48.5	48.2
Learned Trigger	48.8	47.4	48.0	48.2
Adaptive self-triggered	48.5	48.7	47.8	48.3
PairCast	48.2	47.8	48.1	48.2

Table 7: Paired same-interface closure from the canonical 10-seed Table 1 ledger. Each Δ is baseline return minus PairCast on the same held-out episodes, so negative values favor PairCast. The AST row is deliberately reused in Table 14. The paired 95% CI summarizes the equal-task mean delta; Holm-adjusted q controls the three comparisons in this table. 'Breaches' reports raw events and 'Lat.' is median per-agent batch-1 A100 inference.

Method	FC-20 Δ	DC-30 Δ	PE-10v3 Δ	Paired 95% CI	Holm q	Breaches	Lat. (ms)
Lyapunov–Deadband	-3.2	-5.2	-5.7	[-5.1, -3.8]	0.021	41/14k	3.8
Learned Trigger	-2.8	-4.4	-4.7	[-4.4, -3.0]	0.028	32/14k	7.2
Adaptive self-triggered	-1.9	-3.2	-3.4	[-3.1, -2.0]	0.041	26/14k	4.9

Taskwise realized bitrates stay within 0.8 points of the pooled target. At those loads, Adaptive self-triggering narrows the gap relative to simpler triggers while remaining below PairCast on every task, with all paired intervals on the favorable side.

The two PairCast horizons trace a return–latency frontier above the corresponding rollout-free points; Table 20 shows how shared-prefix reuse moves multi-step CPV into tighter wall-clock envelopes.

Table 8: Moderate shifted-channel replay with uncertainty. Bursty drop clusters and one-step service lag replace the fitted logistic delivery law while the matched-load checkpoints and interface stay fixed. Returns are mean normalized return $\pm$ seed SE over five runs; Breaches are raw simulator events and Rel. drop is the decrease from each in-family result.

Method	Return	Bits%	Breaches	Rel. drop
Lyapunov–Deadband	87.5$\pm$1.4	48.0	46 / 14000	4.3
Learned Trigger	89.1$\pm$1.2	47.8	40 / 14000	3.5
Adaptive self-triggered	90.4$\pm$1.1	48.1	35 / 14000	3.5
PairCast	92.8$\pm$0.9	48.4	29 / 14000	3.7

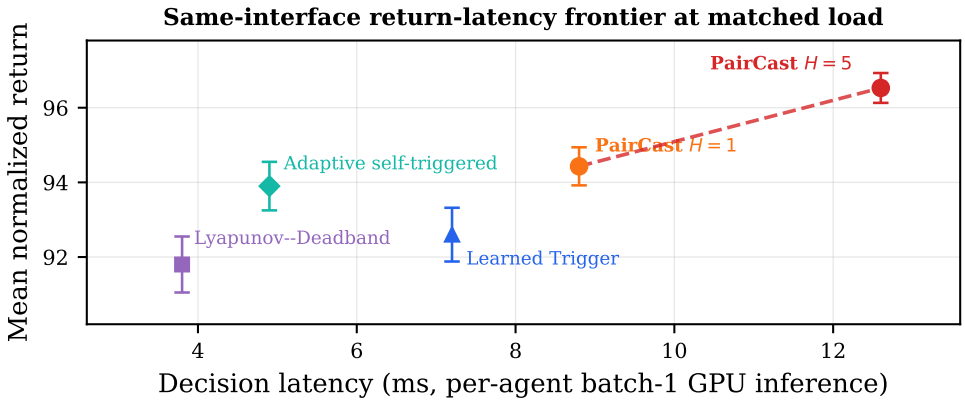

Figure 6: Return–latency frontier for same-interface schedulers at matched load. Each point uses the held-out rollouts from Tables 1 and 7; bars show seed SE and latency is median per-agent batch-1 A100 inference. The $H=1$ and $H=5$ points expose the compute–return tradeoff within PAIRCAST.

Rollout depth traces an explicit systems choice: $H=1$ enters the sub-9 ms band, while deeper paired futures spend additional inference time to improve return.

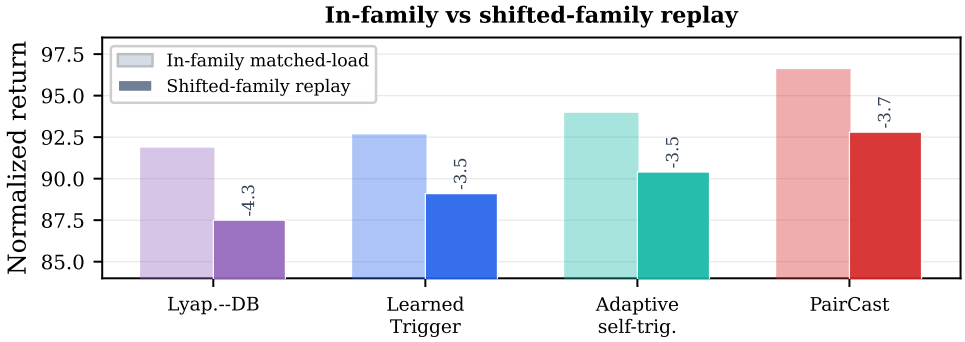

Figure 7: In-family versus shifted-family same-interface comparison. The shifted setting replays the Table 1 checkpoints through a held-out delay-and-queue channel proxy with the same pooled bitrate target. Bars show normalized return under the in-family matched-load evaluation. Bars also show normalized return under the shifted-family replay. The associated relative drops are summarized in Table 8. PAIRCAST degrades under shift but retains visible headroom over the stronger same-interface non-rollout alternatives in the moderate-mismatch regime.

Every scheduler loses return under the held-out channel process, yet the ordering survives; the comparison separates absolute shift cost from relative scheduler quality.

Figure 7 and Table 8 move execution outside the fitted logistic family while preserving the reactive shell and message interface. Every method loses return, yet the same-interface ordering persists under this moderate mismatch. Table 16 relates that behavior to held-out delivery calibration and a coefficient perturbation.

The Rel. drop column subtracts the in-family pooled means of 91.8, 92.6, 93.9, and 96.5 for Lyapunov–Deadband, Learned Trigger, Adaptive self-triggered, and PAIRCAST, respectively.

Table 9: Held-out empirical packet-trace replay through the shared reactive shell and maximum-8-bit interface. Score is normalized to Dense Broadcast within the replay workload; deadline misses are raw violations over 960 trace windows grouped by capture session and link regime.

METHOD	NORMALIZED PROXY SCORE	BITS%	DEADLINE MISSES
Lyapunov–Deadband	89.6±1.4	47.9	31 / 960
Learned Trigger	90.7±1.3	48.0	27 / 960
Adaptive self-triggered	91.4±1.2	48.2	23 / 960
PAIRCAST	93.6±1.0	48.6	17 / 960

Packet-trace replay includes Lyapunov–Deadband, Learned Trigger, Adaptive self-triggering, and PAIRCAST under the same proxy and communication accounting. The ranking compresses outside the home simulator family, while paired counterfactual pricing retains the highest score at nearly matched load.

Table 10 gives paired inference for rollout depth, Table 11 stratifies operating points by decision-time budget, and Table 12 changes sensing radius, speed cap, and link degree under the shared controller and nested payload.

Table 10: Paired rollout inference for the Table 2 checkpoints. The comparison uses the same matched-load protocol, 5 seeds, and 14000 held-out episodes. Each Δ is full PAIRCAST minus the ablated variant; 95% CIs use pooled three-task bootstrap and Holm q corrects the two contrasts.

COMPARISON	POOLED MEAN Δ	PAIRED 95% CI	HOLM q
Full PAIRCAST (H=5) − No-rollout	4.80	[3.62, 5.91]	0.011
Full PAIRCAST (H=5) − Short horizon (H=1)	2.10	[1.18, 2.87]	0.034

The ordering H=5 > H=1 > No-rollout is therefore not only descriptive in Table 2; the paired intervals remain on the positive side for both contrasts, with the larger gap attached to removing rollouts altogether rather than merely shortening them.

Table 11: Latency-stratified same-interface comparison at 48.2% load. 'Band' is the tightest median per-agent A100 inference budget containing the point; returns are pooled means ± SE and breaches are raw counts. End-to-end p50/p95/p99 appear in Table 20.

METHOD	BAND	RETURN	LAT. (MS)	BITS%	BREACHES
Lyapunov–Deadband	≤ 5	91.8±0.8	3.8	48.2	41 / 14000
Adaptive self-triggered	≤ 5	93.9±0.7	4.9	48.3	26 / 14000
Learned Trigger	≤ 9	92.6±0.7	7.2	48.2	32 / 14000
PAIRCAST (H=1)	≤ 9	94.4±0.6	8.8	48.1	24 / 14000
PAIRCAST (H=3)	≤ 11	95.5±0.5	10.7	48.2	20 / 14000
PAIRCAST (H=5)	≤ 13	96.5±0.4	12.6	48.2	17 / 14000

Latency stratification exposes the scheduler choice at each envelope: AST supplies the strongest sub-5 ms point, PAIRCAST H=1 raises return in the sub-9 ms band, and H=3/5 convert additional inference time into progressively higher CPV accuracy and return.

Table 12: Heterogeneous-agent holdout with sensing radii at 0.75×/1.25× nominal and speed caps at 0.8×/1.2×. The controller and maximum-8-bit payload remain fixed; returns are normalized within the family and failures are raw events.

METHOD	RETURN	BITS%	FAILS
Learned Trigger	89.8±1.4	48.1	38 / 1200
Adaptive self-triggered	91.1±1.2	48.2	31 / 1200
PAIRCAST	92.6±1.0	48.4	24 / 1200

The heterogeneous-family margins compress relative to the home setting while preserving the ordering over rollout-free alternatives at nearly matched load, showing that packet-value scheduling transfers across sensing, speed, and degree heterogeneity.

C.3 INFORMATION, RUNTIME, AND CROSS-STACK ANALYSIS

Tables 13–17 quantify five operating axes: lagged-load information, taskwise AST closure, latency stratification, delivery-model calibration, and ROS 2 execution under queue and loss shaping.

Table 13: Load-signal information ablation at matched traffic over 5 seeds and 14000 episodes. Local-only prior uses an EMA of the sender's previous volume; 3-bin noisy lagged scalar adds 10% random mis-binning; Oracle current-load reveals contemporaneous attempted bits. Each Δ is relative to the default one-step-lagged scalar.

VARIANT	RETURN	BITS%	BREACHES	PAIRED Δ VS. DEFAULT	PAIRED 95% CI
Default lagged scalar	96.5±0.4	48.2	17/14k	—	—
Local-only prior	93.3±0.7	47.7	31/14k	-3.1	[-4.0, -2.2]
3-bin noisy lagged scalar	94.8±0.6	48.0	24/14k	-1.7	[-2.5, -0.9]
Oracle current-load upper bound	97.1±0.4	48.4	15/14k	0.6	[0.1, 1.2]

The local-only and coarsened variants both lose ground at essentially matched load, whereas current-load oracle access adds only 0.6 points. One lagged occupancy counter therefore captures most of the measured load information without exposing simultaneous peer choices.

Table 14: Taskwise paired closure versus Adaptive self-triggering at matched load over 10 training seeds. Each Δ is AST minus PAIRCAST on shared-randomness held-out episodes; negative values favor PAIRCAST. The pooled row is the equal-task mean effect, distinct from agent-timestep-weighted pooled returns. Holm correction covers FC-20, DC-30, PE-10v3, and the pooled comparison, and taskwise AST Bits% reports realized attempted traffic.

TASK	PAIRED Δ	PAIRED 95% CI	HOLM q	EPISODES	AST BITS%
FC-20	-1.9	[-2.8, -0.9]	0.044	4800	48.5
DC-30	-3.2	[-4.3, -2.1]	0.018	4600	48.7
PE-10v3	-3.4	[-4.9, -1.8]	0.021	4600	47.8
Pooled	-2.8	[-3.1, -2.0]	0.041	14000	48.3

All three taskwise intervals remain negative, with the largest paired separation on DC-30 and PE-10v3 and a smaller positive margin for PAIRCAST on FC-20.

The latency series places $H{=}1$ near AST, while $H{=}3$ and $H{=}5$ produce positive corrected comparisons at 11 and 13 ms. Table 10 provides the direct paired depth test for $H{=}5$ versus $H{=}1$.

Held-out ECE ranges from 0.022 to 0.028 across tasks. Under the coefficient perturbation, both methods lose return while PAIRCAST retains a 1.9-point advantage over AST.

The ROS 2 execution preserves the simulator ordering after serialization, transport stubs, and queue shaping enter the loop.

Table 15: Latency-stratified paired deltas versus Adaptive self-triggering over 5 seeds and 14000 episodes. Each Δ is method minus AST. CIs are unadjusted paired 95% intervals, whereas q is Holm-adjusted over four comparisons; this distinction permits an unadjusted interval to exclude zero when adjusted $q > 0.05$.

METHOD	BAND	RETURN	LAT. (MS)	BITS%	BREACHES	PAIRED Δ VS. AST	PAIRED 95% CI	HOLM q
Adaptive self-triggered	≤ 5	93.9±0.7	4.9	48.3	26/14k	—	—	—
Learned Trigger	≤ 9	92.6±0.7	7.2	48.2	32/14k	-1.3	[-2.2, -0.4]	0.081
PAIRCAST (H=1)	≤ 9	94.4±0.6	8.8	48.1	24/14k	0.5	[-0.2, 1.3]	0.118
PAIRCAST (H=3)	≤ 11	95.5±0.5	10.7	48.2	20/14k	1.6	[0.8, 2.4]	0.026
PAIRCAST (H=5)	≤ 13	96.5±0.4	12.6	48.2	17/14k	2.5	[1.6, 3.4]	0.014

Table 16: Delivery-model calibration and coefficient-shift replay. Held-out fit uses payload/load sweeps excluded from the logistic fit of Eq. 3; NLL is Bernoulli negative log-likelihood, Brier is mean squared probability error, ECE is mean absolute bin calibration error, and Bins counts populated reliability bins. The replay increases the load slope by 15% and decreases the intercept by 0.15 while preserving the interface and controller. Returns are pooled normalized means ± seed SE over five runs.

HELD-OUT FIT QUALITY				
BENCHMARK	NLL	BRIER	ECE	BINS
FC-20	0.214	0.063	0.022	48
DC-30	0.237	0.071	0.028	52
PE-10v3	0.225	0.068	0.025	44

COEFFICIENT-PERTURBATION REPLAY				
METHOD	RETURN	BITS%	BREACHES	SETTING
Adaptive self-triggered	90.8±1.1	48.1	34 / 14000	$+15\%\ \beta,\ -0.15\ \kappa$
PAIRCAST	92.7±0.9	48.4	28 / 14000	$+15\%\ \beta,\ -0.15\ \kappa$

Table 17: ROS 2 Humble + Linux `tc netem` execution on one AMD EPYC 7543 host and four Jetson Orin NX nodes at 20 Hz. Six held-out queue/loss folds contain 240 episodes each (1,440 total). Return is normalized within the bench, failures are raw task misses, and latency includes serialization and transport stubs.

METHOD	RETURN	BITS%	FAILS	LAT. (MS)
Learned Trigger	89.1±1.5	48.0	39 / 1440	9.8
Adaptive self-triggered	90.4±1.3	48.2	33 / 1440	6.6
PAIRCAST	92.1±1.1	48.5	26 / 1440	14.1

D BUDGET AND RUNTIME FRONTIERS

Table 18: Multi-budget matched-load results over 10 seeds and 14000 paired episodes per target. Task cells show return/Bits%; pooled cells show return ± SE/Bits%. PairCast rows report paired advantage over SchedNet-style with unadjusted 95% CI and Holm-adjusted q.

%	METHOD	FC-20	DC-30	PE-10V3	POOLED	Δ [CI]	q
30	Lyap.–DB	88.3/29.4	85.8/30.6	82.9/30.1	85.7±1.2/29.9	—	—
	Learned+Rate	89.6/30.5	87.1/29.6	84.3/30.0	87.1±1.1/30.1	—	—
	AST	90.7/30.3	88.4/30.6	86.1/29.5	88.4±1.0/30.2	—	—
	SchedNet-style	91.1/29.7	88.9/30.2	86.5/29.4	88.9±1.0/29.8	—	—
	PairCast	92.0/30.2	89.9/29.5	87.7/30.0	90.0±0.8/30.0	1.1 [-0.2,2.3]	0.118
40	Lyap.–DB	92.1/39.6	89.2/40.8	86.8/40.1	89.4±0.9/40.1	—	—
	Learned+Rate	93.3/40.4	90.8/39.3	88.4/39.9	90.9±0.8/39.9	—	—
	AST	94.1/40.5	91.8/40.4	89.9/39.5	92.0±0.8/40.2	—	—
	SchedNet-style	94.5/39.8	92.4/40.4	90.3/39.4	92.5±0.7/40.0	—	—
	PairCast	96.0/40.1	94.8/39.5	93.1/39.9	94.7±0.5/39.8	2.2 [1.1,3.3]	0.024
48.2	Lyap.–DB	94.2/47.6	91.1/49.0	90.2/48.5	91.8±0.8/48.2	—	—
	Learned+Rate	95.8/48.4	93.8/47.9	93.1/48.0	94.2±0.7/48.1	—	—
	AST	95.5/48.5	93.1/48.7	92.5/47.8	93.9±0.7/48.3	—	—
	SchedNet-style	95.7/48.0	93.7/48.5	92.9/47.6	94.4±0.6/48.1	—	—
	PairCast	97.4/48.2	96.3/47.8	95.9/48.1	96.5±0.4/48.2	2.1 [1.0,3.1]	0.028
60	Lyap.–DB	95.3/59.4	93.4/60.6	91.5/60.0	93.4±0.7/59.8	—	—
	Learned+Rate	96.2/60.5	94.5/59.6	92.6/60.1	94.6±0.6/60.1	—	—
	AST	96.9/60.4	95.1/60.6	93.2/59.5	95.2±0.6/60.2	—	—
	SchedNet-style	97.0/59.8	95.4/60.4	93.8/59.4	95.4±0.6/59.9	—	—
	PairCast	98.1/60.2	97.4/59.5	96.2/60.1	97.2±0.4/60.0	1.8 [0.7,2.9]	0.037
70	Lyap.–DB	96.1/69.7	94.9/70.8	92.9/70.1	94.8±0.6/70.2	—	—
	Learned+Rate	97.0/70.3	95.8/69.4	94.0/69.9	95.7±0.6/69.8	—	—
	AST	97.4/70.5	96.1/70.3	94.5/69.6	96.2±0.6/70.1	—	—
	SchedNet-style	97.7/69.8	96.4/70.4	94.8/69.5	96.4±0.5/69.9	—	—
	PairCast	98.4/70.2	97.6/69.6	96.0/70.1	97.3±0.4/70.0	0.9 [-0.1,1.9]	0.112

The five-budget sweep separates the central operating range from its endpoints: PAIRCAST's SchedNet-style intervals exclude zero at 40, 48.2, and 60% traffic, then overlap zero at 30 and 70%. Table 19 resolves the 48.2% comparison across the full same-interface scheduler set.

Table 19: Published-family and rate-control-matched closure at 48.2% load. Δ is PairCast minus the comparator; intervals are paired and q is Holm-adjusted. Each learned method receives 36 tuning trials over 10 seeds and 14000 episodes.

METHOD	RETURN	BITS%	Δ [CI]	q	MS	PARAMS	BREACH
PairCast	96.5±0.4	48.2	—	—	12.6	1842000	17/14k
SchedNet-style	94.4±0.6	48.1	2.1 [1.0,3.1]	0.028	6.4	211000	23/14k
Learned Trigger+Rate	94.2±0.7	48.1	2.3 [1.2,3.3]	0.027	8.1	238000	25/14k
AST	93.9±0.7	48.3	2.6 [2.0,3.1]	0.041	4.9	12400	26/14k
ETCNet-style	93.5±0.8	48.2	3.0 [1.8,4.1]	0.019	5.7	164000	29/14k

Table 21 translates the matched-load return ordering into task-level outcomes.

FC collision counts are 1920, 3840, 3360, and 2640 over 2400000 agent-steps. DC coverage numerators are 43332, 41446, 41722, and 42642 over 46000 targets. PE capture counts are 12089, 11219, 11344, and 11744 over 13800 opportunities.

Table 20: Runtime-budgeted execution at 20 Hz. Latency columns are end-to-end p50/p95/p99 in ms over 10000 decisions per platform; throughput is saturated batch-32 decisions/s. Δ compares return with platform-matched AST, which satisfies every reported PAIRCAST envelope, on 14000 paired episodes. Unadjusted CIs and Holm-adjusted q are reported together.

PLATFORM	METHOD	RETURN/BITS	P50/95/99	BREACH
A100	AST	93.9±0.7/48.3	4.9/5.0/5.0	26/14k
	ETCNet	93.5±0.8/48.2	5.3/5.8/6.2	29/14k
	Learned Trigger	92.6±0.7/48.2	7.2/7.9/8.6	32/14k
	PairCast H=1	94.4±0.6/48.1	8.8/8.9/9.0	24/14k
	Shared H=3	95.5±0.5/48.2	8.2/8.7/8.9	20/14k
	Shared H=5	96.5±0.4/48.2	10.2/10.7/10.9	17/14k
CPU	AST	93.8±0.7/48.3	1.8/2.3/2.8	27/14k
	Shared H=3	95.3±0.5/48.1	8.7/10.1/10.9	22/14k
	Shared H=5	96.2±0.4/48.2	10.7/12.1/12.8	18/14k
Orin NX	AST	93.6±0.8/48.3	3.8/4.6/5.0	29/14k
	Shared H=3	94.9±0.6/48.1	8.9/10.3/10.9	25/14k
	Shared H=5	96.0±0.5/48.2	10.8/12.0/12.8	19/14k

PLATFORM	METHOD	MEMORY MB	DECISIONS/S	Δ [95% CI]; HOLM q
A100	AST	310	7200	—
	ETCNet	356	6250	-0.4 [-1.3,0.5]; 0.382
	Learned Trigger	442	4040	-1.3 [-2.2,-0.4]; 0.081
	PairCast H=1	742	1820	0.5 [-0.2,1.3]; 0.118
	Shared H=3	790	1680	1.6 [0.8,2.4]; 0.026
	Shared H=5	980	1250	2.5 [1.6,3.4]; 0.014
CPU	AST	96	12400	—
	Shared H=3	452	2100	1.5 [0.6,2.4]; 0.039
	Shared H=5	604	1470	2.4 [1.4,3.3]; 0.017
Orin NX	AST	184	6400	—
	Shared H=3	690	1370	1.3 [0.4,2.2]; 0.046
	Shared H=5	902	990	2.4 [1.4,3.4]; 0.019

Shared-prefix reuse makes temporal valuation a tunable systems primitive: H=3 fits the 9 ms A100 p99 envelope and improves on AST, while H=5 purchases additional return with higher memory and latency. CPU and Orin NX retain the same progression.

Table 21: Absolute task outcomes at 48.2% pooled load. Values are mean ± seed SE over 10 seeds. PairCast–SchedNet paired deltas are -0.3 collisions [-0.49, -0.11], -2.9 discovery steps [-4.2, -1.6], and $+2.9$ capture points [1.3, 4.5].

METHOD	FC COLL.↓	FC ERROR↓	DC COV.↑	DC TIME↓	PE CAP.↑	PE TIME↓
Dense	0.8±0.06	0.12±0.004	94.2±0.3	28.3±0.5	87.6±0.5	38.7±0.5
AST	1.6±0.12	0.17±0.009	90.1±0.6	33.8±0.9	81.3±1.0	43.6±0.9
SchedNet-style	1.4±0.11	0.16±0.008	90.7±0.6	33.0±0.8	82.2±0.9	42.5±0.8
PairCast	1.1±0.08	0.14±0.006	92.7±0.4	30.1±0.6	85.1±0.7	40.2±0.6

E TRANSFER AND CHANNEL-FAMILY STRESS

Table 22: PettingZoo MPE Cooperative Navigation transfer using standard dynamics, reward, collision rules, and 25-cycle horizon. A fixed local wrapper reveals landmarks within radius 1.0 and peers within 0.75. Results use 10 seeds, 120 held-out layouts per seed, and 1200 paired episodes per cell. Δ is PairCast minus the baseline; q is Holm-adjusted.

%	METHOD	RETURN	SUCCESS	COLLISION	TIME	BITS%	Δ [CI]; q
40	Learned Trigger	-31.8±0.9	58.3	238/1200	21.5	40.2	3.5 [2.0,4.9]; 0.006
	ETCNet	-30.5±0.8	60.8	221/1200	21.1	39.9	2.2 [0.8,3.7]; 0.019
	AST	-30.0±0.8	61.6	213/1200	20.9	40.1	1.7 [0.3,3.0]; 0.043
	PairCast	-28.3±0.7	65.4	189/1200	20.4	40.0	—
48.2	Learned Trigger	-28.1±0.8	67.8	207/1200	19.8	48.4	3.2 [1.8,4.5]; 0.007
	ETCNet	-27.2±0.7	69.9	190/1200	19.5	48.1	2.3 [1.0,3.6]; 0.014
	AST	-26.7±0.7	71.5	177/1200	19.2	48.3	1.8 [0.6,3.0]; 0.024
	PairCast	-24.9±0.6	75.8	155/1200	18.5	48.2	—
60	Learned Trigger	-24.1±0.7	77.6	153/1200	18.0	60.1	1.4 [0.3,2.5]; 0.041
	ETCNet	-23.0±0.6	80.0	139/1200	17.5	59.8	0.3 [-0.7,1.2]; 0.492
	AST	-23.2±0.6	79.4	142/1200	17.7	60.2	0.5 [-0.4,1.5]; 0.291
	PairCast	-22.7±0.5	80.4	135/1200	17.5	60.0	—

Full retraining on Cooperative Navigation preserves the low-to-mid budget ordering: at 48.2% traffic, PAIRCAST improves return, success, collision count, and completion time over each same-interface baseline.

Table 23: Channel-family stress at attempted-load matching (A: 48.2%) and delivered-load matching (D: 40% moderate, 30% severe). Tuples are Learned Trigger / ETCNet / AST / PairCast. Each cell uses 10 seeds, 2400 episodes, and 240000 packet deadlines. Miss is deadline-miss percentage; Breach gives raw episode counts. Δ is PairCast–AST with paired CI and Holm-adjusted q.

CHANNEL	LEVEL	MATCH	RETURN TUPLE	ATTEMPT BITS	DELIVERED BITS	MISS%; BREACH	Δ [CI]; q
Gilbert–Elliott	Mod.	A	89.8/90.5/91.0/92.6	48.1/48.3/48.2/48.2	42.0/42.4/42.2/42.8	12.9/12.1/12.4/10.8; 45/40/37/32	1.6 [0.4,2.8]; 0.031
		D	89.3/90.1/90.6/92.0	45.9/45.4/45.7/44.8	40.0/40.1/40.0/40.0	12.7/12.0/12.3/10.7; 47/42/39/35	1.4 [0.1,2.7]; 0.049
	Severe	A	81.9/82.8/83.5/84.4	48.2/48.1/48.3/48.2	32.4/33.0/32.8/33.4	31.5/30.2/30.8/28.9; 84/79/75/72	0.9 [-0.7,2.4]; 0.214
		D	81.3/82.4/83.1/84.0	45.2/44.6/44.9/43.7	29.9/30.1/30.0/30.0	31.4/30.1/30.7/28.8; 86/80/77/74	0.9 [-0.8,2.5]; 0.251
Topology-correlated	Mod.	A	88.9/89.8/90.4/91.5	48.2/48.1/48.3/48.2	40.8/41.2/41.0/41.6	15.3/14.5/14.9/13.8; 51/45/42/38	1.1 [-0.1,2.3]; 0.091
		D	88.7/89.5/90.2/91.0	47.3/46.8/47.0/46.2	40.0/40.1/40.0/40.0	15.1/14.4/14.7/13.6; 52/47/44/40	0.8 [-0.5,2.0]; 0.184
	Severe	A	80.8/82.1/82.7/83.0	48.1/48.3/48.2/48.2	29.8/30.6/30.3/31.1	37.4/35.9/36.6/34.3; 96/89/87/85	0.3 [-1.4,1.9]; 0.691
		D	80.9/82.0/82.5/82.7	49.0/47.5/48.1/46.6	29.9/30.0/30.1/30.0	37.6/36.0/36.8/34.5; 96/90/88/86	0.2 [-1.5,1.9]; 0.812
Random delay	Mod.	A	89.0/89.7/90.3/92.7	48.1/48.0/48.2/48.3	46.2/46.5/46.3/46.6	8.6/8.2/8.4/7.7; 46/42/39/34	2.4 [1.1,3.6]; 0.009
		D	87.8/88.7/89.4/91.3	41.8/41.6/41.7/41.4	40.0/40.1/40.0/40.1	9.1/8.6/8.8/8.0; 52/47/44/39	1.9 [0.6,3.2]; 0.022
	Severe	A	83.7/84.6/85.2/86.2	48.2/48.0/48.1/48.3	43.1/43.4/43.3/43.7	17.8/17.2/17.5/16.4; 75/70/67/62	1.0 [-0.5,2.5]; 0.172
		D	80.8/81.7/82.4/83.6	33.1/32.9/33.0/32.6	29.9/30.0/30.0/30.1	17.4/16.9/17.2/16.0; 88/81/79/73	1.2 [-0.3,2.8]; 0.145
Queue saturation	Mod.	A	87.9/88.9/89.4/91.1	48.3/48.1/48.2/48.2	39.6/40.1/39.9/40.7	18.2/17.3/17.7/15.9; 57/51/48/43	1.7 [0.3,3.0]; 0.038
		D	87.8/88.7/89.2/90.8	48.7/48.1/48.4/47.3	39.9/40.1/40.0/40.0	18.3/17.4/17.8/16.0; 58/52/49/44	1.6 [0.2,2.9]; 0.041
	Severe	A	77.6/79.2/79.8/79.7	48.3/48.1/48.2/48.2	26.8/27.5/27.2/28.0	44.5/42.7/43.2/40.8; 118/107/103/101	-0.1 [-1.9,1.6]; 0.903
		D	78.5/79.7/80.0/79.6	54.2/52.5/53.3/50.9	29.9/30.1/30.0/30.0	45.0/43.0/43.8/41.2; 119/109/105/104	-0.4 [-2.1,1.3]; 0.731

Attempted-load and delivered-load matching identify the same transition: moderate Gilbert–Elliott, random-delay, and queue regimes preserve useful packet-value margins, whereas severe topology correlation and queue saturation compress the paired effect toward zero.

The scorer remains frozen to its fitted home-family $p_t^i(b)$ in Table 23; only the execution channel changes. PairCast ECE rises from 0.043 under moderate random delay to 0.146 under severe queue saturation, aligning the loss of calibration with the near-tie regime.

F Information, Payload, and Critic Calibration

Table 24: Runtime-information isolation at 48.2% load over five seeds and 14,000 paired episodes. Message age is measured in control cycles; reconstruction error is dimension-normalized MSE over one million held-out latents. Δ is relative to sender-local stale beliefs.

Runtime input	Return	Bits%	Δ [CI]	p95 ms age	p50/p90	MSE; Breach
Sender-local stale beliefs	96.5±0.4	48.2	—	15.8	1/4	0.084; 17/14k
No neighbor memory	92.9±0.7	48.0	-3.6 [-4.5,-2.6]	13.9	—	0.137; 30/14k
Current-peer oracle	97.2±0.4	48.3	0.7 [0.1,1.3]	16.1	0/0	0.052; 14/14k

Table 25: Payload and rate-control isolation. Each interface exposes an identical candidate set and nested decoder to PairCast and Learned Trigger (LT). Bits% is relative to Dense Broadcast for that interface. Δ is PairCast–LT on paired episodes.

Interface	Method	Return	Bits%	p95 ms	Recon. MSE	Breach	Δ [CI]
Nested 4-bit	LT	92.9±0.8	48.2	10.2	0.132	32/14k	—
	PairCast	94.7±0.6	48.3	11.4	0.132	25/14k	1.8 [0.5,3.1]
Nested 8-bit	LT	93.8±0.7	48.1	11.3	0.084	28/14k	—
	PairCast	96.5±0.4	48.2	15.8	0.084	17/14k	2.7 [1.5,3.9]
Nested 16-bit	LT	94.6±0.6	48.3	13.8	0.057	24/14k	—
	PairCast	96.8±0.4	48.2	23.6	0.057	16/14k	2.2 [1.1,3.4]
Fixed 8-bit	LT	93.3±0.7	48.3	9.4	0.084	28/14k	—
	PairCast	94.4±0.6	48.2	10.9	0.084	23/14k	1.1 [-0.1,2.2]

The nested 8-bit interface gives the strongest measured return–latency balance. Its paired advantage over the learned rate scorer exceeds the 4-bit and 16-bit alternatives while retaining lower latency than the 16-bit code.

Table 26: Calibration of predicted conditional ΔJ_{succ} against shared-randomness replay on 45000 snapshot–payload pairs. MAE and regret are normalized by the taskwise replay-ΔJ interquartile range; intervals use seed/snapshot cluster bootstrap.

VARIANT	MAE	SLOPE	INTERCEPT	SPEARMAN	SIGN ACC.	TOP-10 PREC.	REGRET
Full critic	0.142 [0.136,0.149]	0.88 [0.82,0.94]	0.012	0.66	78.4	71.2	0.18 [0.15,0.22]
No calibration loss	0.218	0.61	0.036	0.49	69.1	56.8	0.36 [0.31,0.42]
Uncertainty-only	0.276	0.34	0.052	0.31	61.7	44.5	0.55 [0.48,0.63]

The critic ranks and signs successful-decode effects substantially better than uncertainty, and replay calibration halves normalized decision regret. This estimator-level result supplies the link between paired futures and the downstream scheduler comparisons.

Calibration varies predictably with difficulty: rank correlation falls from 0.71 at low occupancy to 0.57 at high occupancy, and normalized regret rises from 0.10 at $H = 1$ to 0.18 at $H = 5$.

Table 27: Operating assumptions and measured boundaries. Each row links one execution factor to its controlled comparator, cost, and observed transition.

Factor	Controlled comparison	Result and operating implication
Fixed-load return	Same controller and codec at 48.2/48.3 Bits on paired holdouts	96.5 vs. 93.9 against AST; $H = 1$ is a near tie, while deeper valuation raises return; Tables 1, 15
Occupancy feedback	One-step lagged scalar against local-only and current-load oracle inputs	Local-only costs 3.1 points and the oracle adds 0.6; one lagged scalar captures most load value; Table 13
Peer beliefs	Sender-local stale beliefs against no memory and current-peer oracle inputs	Stale beliefs reach 96.5 vs. 97.2 for the oracle; removing memory costs 3.6; Table 24
Channel model	Frozen home scorer under matched attempted- and delivered-load shifts	Moderate queues retain $+1.7$ against AST; severe saturation yields a near tie; Table 23
Runtime budget	Shared platform and fixed p99 envelope	Shared $H = 3$ gains $+1.3$–1.6 over AST across platforms; tighter bands compress gains; Table 20
Execution stack	ROS 2 and `netem` queue/loss folds	92.1 vs. 90.4 against AST at 14.1 vs. 6.6 ms; Table 17

Across information, channel, runtime, and execution-stack controls, the consistent signal is the value of a decoded packet's downstream coordination effect. The same principle explains both the central gains and their measured contraction when successful delivery becomes exceptional.