THE SHAPE OF COOPERATION ACROSS TASKS

Anonymous authors
Paper under double-blind review

ABSTRACT

When should the same agents communicate through a hub, through groups or directly? We characterize these choices by two fitted summaries of evidence demand, the traffic outside the best hub and the smallest modular detour. Under a faithful transport model that preserves evidence delivery, this signature and the ratio of channel overhead to traffic cost determine the cheapest pattern within the declared family. Identical hub concentration can require different patterns, so concentration alone is insufficient. We derive exact decision boundaries and a confidence bound that charges for learning before authorizing a change. Controlled transport experiments show that matching connectivity to persistent demand can lower total modeled cost after paying to learn it. A direct ablation shows that access to modular patterns improves future cost for concentrated demand but adds no benefit for diffuse demand. More observations improve prediction but can reduce net savings, while replacing future demand reverses the benefit. The resulting characterization explains when fitting a richer communication pattern is useful and when its learning expense or changing demand defeats the gain.

1 INTRODUCTION

Agent teams make decisions not only about answers but also about who should exchange evidence. The same agents might investigate a software failure, reconcile conflicting sources or check a mathematical argument, yet need different communication patterns. Understanding these organization decisions requires more than displaying their graphs. Why should evidence pass through one worker rather than circulate within groups? Which observed need justifies the extra connections? An explanation should identify the alternative, the evidence favoring the choice and the conditions under which that choice should change.

Consider agents investigating a software failure. Findings about separate components may favor exchange within the corresponding investigations. If several components depend on one interface, its investigator may instead be a useful meeting point. However, forwarding through that worker adds expense when most evidence is needed elsewhere. The explanation must therefore connect a pattern to the distribution of evidence requests, not merely call the task coding. It should also predict when the saved connections cease to repay the additional forwarding. This turns an appealing diagram into a comparison that can be checked.

This perspective treats organization as an interpretable decision problem. Following the distinction emphasized by Rudin (2019), we seek a rule whose comparisons are explicit rather than a separate explanation of an opaque selector. Automatic agentic design searches programs (Hu et al., 2024) and communication graphs (Zhuge et al., 2024). G Designer conditions generation on the task (Zhang et al., 2025b), while ARG Designer also selects composition and roles (Li et al., 2025). Our question concerns the reason for choosing one communication pattern over another

Which features of evidence demand explain a team's communication pattern, and when does acting on that explanation improve efficiency?

A topology specifies directed communication permissions among fixed workers. Capabilities, role prompts, tools and decision authority remain unchanged. A communication profile records how often one worker needs another's evidence. Selective messaging is an essential comparison because a fixed graph can already carry different information on different tasks.

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.

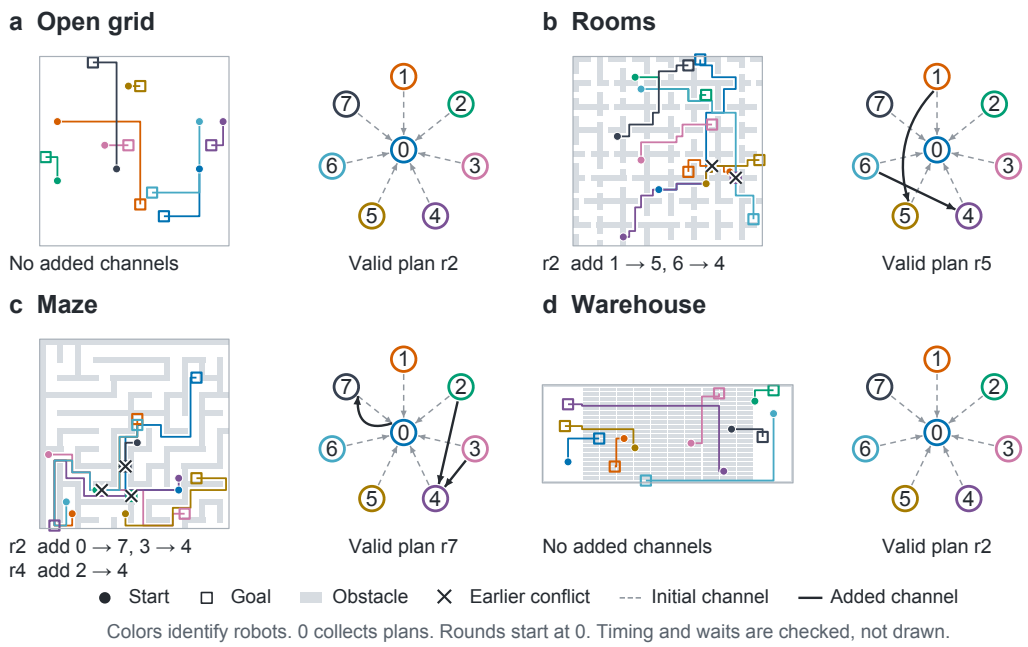

Figure 1: **Route interactions induce different communication networks.** Recorded planning uses eight workers, one conflict rule and prospectively selected seed 48. Gray and solid arrows show initial and added channels. All four tasks succeed; the maze revises twice. Appendix H.5 follows the same workers on four new geometries, including a failed task.

We introduce a task signature that summarizes the forwarding required by a fitted hub and a fitted modular pattern. Together with communication prices, it explains a tradeoff between maintaining channels and relaying evidence. We study this choice for a fixed evidence flow before considering how agents discover and interpret the evidence.

The structural characterization gives an exact comparison of direct exchange, hubs and modules within the candidate family. The explanation uses the same cost differences as the selector, rather than a rationale fitted afterward. Both features matter because equal hub concentration can conceal different opportunities for grouping. The contrasting examples show why one feature cannot resolve this choice, while the gateway construction identifies the actual connections. Thus an explanation of the preferred family and a specification of its endpoints are distinct.

An exact explanation of an observed choice need not predict useful future behavior. Requests are sampled, acquiring them costs resources and the demand may change. The learning bound and paid revision guarantee give sufficient conditions for expected savings after uncertainty, drift and revision expense. They connect the fitted explanation to a testable prediction rather than treating fidelity to the rule as evidence of benefit.

The controlled experiments distinguish three questions. Does the explanation reproduce the implemented decision? Do its connections match the workers who need to exchange evidence? Does the predicted saving survive on future requests after specified selection costs? Endpoint permutation tests the second question; separate future packets and the modular ablation address the third. Gains concentrate where demand persists and fits groups. More observations improve the choice but can reduce net value, and replacement of demand can reverse the gain.

Figure 1 motivates measuring interactions through a separate conflict rule. The signature selector also supplies joint graph proposals to AUTOFORM, but the combined feedback procedure remains unevaluated. Realistic use requires requests that predict useful evidence and measured overhead that makes fewer channels worthwhile.

(i) **An exact structural characterization.** Two fitted demand summaries and a cost ratio determine the cheapest pattern family. Independent gateway choices give an exact factorized implementation.

(ii) **A tested learning tradeoff.** A bound and a controlled experiment identify how observation expense and profile drift limit the benefit of selecting a network.

(iii) **A test of richer connectivity.** A paired modular ablation identifies when fitting groups adds value beyond fitting a hub, retaining its null and adverse regimes.

The result explains an organization choice through its alternative, decisive cost and range of usefulness. Section 6 develops this contrastive interpretation.

2 Related work

Interpretable decision rules Rudin (2019) distinguishes inherently interpretable models from explanations of black boxes. LIME fits a local interpretable surrogate to a classifier (Ribeiro et al., 2016). Here the explanation is the topology selector's own cost comparison, with measured requests and prices as inputs. It explains the selected pattern relative to an alternative, consistent with the contrastive perspective of Miller (2019). This is an explanation of organization, not of workers' internal reasoning.

Evaluating explanations Doshi-Velez & Kim (2017) separate evaluation through formal proxies from studies involving people and applications. We distinguish exact recovery of the selector's choice from prediction of future cost. The permutation and modular controls test the scope and usefulness of the fitted comparisons. These functional tests do not measure human understanding, nor does a compact signature alone establish it.

Automatic organization and interaction GPTSwarm optimizes communication graphs (Zhuge et al., 2024), while ADAS searches executable programs (Hu et al., 2024). G Designer conditions topology on the task (Zhang et al., 2025b); ARG Designer also selects composition and roles (Li et al., 2025). CARD uses environmental signals during training and execution (Wu et al., 2026). ExRole derives roles from trajectories (Liu et al., 2026), while communication pruning removes redundant exchanges (Zhang et al., 2025a). We keep workers and roles fixed and characterize the cost differences underlying a topology choice.

Evidence and value Learned communication (Sukhbaatar et al., 2016) and endogenous network formation (Jackson, 2008) connect information exchange to behavior. Setup and traffic costs are established network design ingredients (Andrews et al., 2016) and supply our transport objective. Complementary set objectives motivate joint proposals (Bai & Bilmes, 2018). Cost and accuracy evaluation (Kapoor et al., 2024) and analyses of agent failures (Cemri et al., 2025) motivate the competitive controls. Established reservation methods (Silver, 2005) and public maps (Stern et al., 2019) supply the retained planning comparison.

3 Problem formulation

We ask whether the same pool of general agents can form useful communication structures across tasks. The target setting fixes agent identities, base policy, roles, tool permissions and context limits. Each agent can perform the same primitive work, but retains a separate history. Differences in their contributions arise through execution, not a supplied specialist decomposition. Changing the agent pool or its capabilities is a separate treatment.

For task ψ, the permitted context x_ψ contains task text, the agent and tool registry, and a total budget B_ψ, but no task graph, reference answer or hidden evaluator label. An agent's history contains this context, its own actions and tool results, and messages actually delivered to it. A shared tool catalog does not imply shared observations or access to another agent's history.

At revision index $k \geq 0$, the same formation policy across tasks chooses directed permissions

$$G_k = \mathcal{F}(\mathcal{H}_k, U_k), \qquad G_k = (V, E_k),$$

where V is the fixed agent set, E_k its directed permissions and U_k exogenous randomization. With V fixed, we identify a graph with its edge set in expressions such as $G \cup T$ and $|G|$ (graph conventions). The organizer history $\mathcal{H}_k$ contains only permitted context, earlier organizations, interaction traces, observable feedback and resource use. Future outcomes and oracle organizations are excluded. Broader organizer trace access must be declared and charged; it does not grant that information to workers. An edge permits delivery rather than guaranteeing that a message is sent. Message selection policies, submission authority and aggregation rules remain fixed when isolating a graph intervention.

Execution produces a canonical decision A, score $Y \in [0, 1]$ and resource use C. The latest committed submission is scored; a missing or malformed submission is null and scores zero. A commitment remains valid after a later execution failure, with termination status and unresolved costs recorded separately. Only declared feedback returns to formation. Execution equivalence requires identical laws of submitted decisions, a stronger condition than equal scores. The shared budget includes formation, revision, worker calls, tools, failed attempts and final aggregation. We report each resource in its native unit and combine costs only under a declared pricing rule.

Freeze the complete policy before evaluation, including its proposal rule, feedback, stopping and any installation gate. It may still adapt within a task. Let $\mu_{\mathcal{F}}(\psi; B_\psi) = \mathbb{E}_{\mathcal{F}}[Y \mid \psi, B_\psi]$ include every execution outcome. For a task distribution $\mathcal{P}$ and declared comparator $\mathcal{F}_0$, the score contrast is

$$\Delta_{\mathcal{P}}(\mathcal{F}, \mathcal{F}_0) = \mathbb{E}_{\psi \sim \mathcal{P}} \left[\mu_{\mathcal{F}}(\psi; B_\psi) - \mu_{\mathcal{F}_0}(\psi; B_\psi) \right]. \tag{1}$$

A frozen organization may reuse saved search resources, while a capable singleton tests collective advantage. Actual costs can be reported separately or combined through the priced policy objective. Note that the target is average policy performance, so individual revisions may help or hurt. Appendix C.8 specifies this study target; Appendix A.3 indexes symbols and proofs.

Section 4 studies a conditional transport problem inside this broader objective. Its profile describes required evidence exchanges at the current information state. It is estimated from observations, not supplied with the task. The completed transport experiment tests this estimation step on constructed requests. Whether naturally observed requests predict final task performance remains a separate empirical question.

4 Task features and topology selection

A task label such as coding or comparison does not determine an efficient network. We characterize the cost of carrying a fixed evidence flow, then ask whether observations predict the cost of carrying future evidence. The distinction separates an exact description of fitted patterns from a test of their predictive value.

4.1 Measurable features of an evidence flow

At a retained checkpoint, fix $N_a = |V| \geq 3$ workers and a horizon of $W \geq 1$ unit evidence packets. The communication profile $p = (p_{ij})_{i \neq j}$ gives the probability that a packet originates at i and is required at j, with $\sum_{i \neq j} p_{ij} = 1$. It describes the task together with the workers' current information, not a supplied dependency graph. Public requests give estimates; their relationship to useful evidence requires validation.

Under the faithful transport assumptions, routing preserves packet contents and destinations. A graph pays declared session overhead $a \geq 0$ per enabled channel and transmission cost $c > 0$ per packet hop. Its expected transport cost is

$$\mathsf{L}_p(G) = a|E(G)| + cW \sum_{i \neq j} p_{ij} d_G(i, j), \tag{2}$$

where d_G is shortest directed distance. All compared graphs deliver every packet with sufficient capacity and at most three hops. This isolates communication cost at preserved evidence delivery, rather than assuming that delivery guarantees a correct answer.

Hub concentration $\gamma_v = \sum_{j \neq v} (p_{vj} + p_{jv})$ measures traffic involving worker v. For a partition $\mathcal{B} = \{B_1, \ldots, B_K\}$ with gateway $g_s \in B_s$, leakage $\epsilon_{\mathcal{B}} = \sum_{s \neq t} \sum_{i \in B_s, j \in B_t} p_{ij}$ measures traffic between groups. The gateway detour $r_{\mathcal{B}}$ counts mean extra hops, with $0 \leq r_{\mathcal{B}} \leq 2\epsilon_{\mathcal{B}}$.

Proposition 4.1 (Task features determine transport comparisons). *Under the transport model, let D be complete direct communication, S_v a bidirectional star centered at v, and $G_{\mathcal{B}}$ complete communication within groups with complete communication between gateways. Write $m_{\mathcal{B}} = \sum_s |B_s|(|B_s|-1) + K(K-1)$. Then*

$$\begin{aligned} \mathsf{L}_p(D) &= aN_a(N_a-1) + cW, \\ \mathsf{L}_p(S_v) &= 2a(N_a-1) + cW(2-\gamma_v), \\ \mathsf{L}_p(G_{\mathcal{B}}) &= am_{\mathcal{B}} + cW(1+r_{\mathcal{B}}). \end{aligned} \tag{3}$$

The cheapest star maximizes γ_v and beats D exactly when $a(N_a-1)(N_a-2) > cW(1-\gamma_v)$. The modular graph beats D exactly when $a[N_a(N_a-1) - m_{\mathcal{B}}] > cWr_{\mathcal{B}}$. For a fixed partition, each optimal gateway independently maximizes its group's external incident traffic. The proof counts channels and paths.

A two dimensional task signature For six workers, let $\mathcal{G}_{\text{module}}$ contain all ninety graphs formed by two groups of three and one gateway per group. Define

$$\Phi(p) = \big(H(p), M(p)\big), \quad H(p) = 1 - \max_v \gamma_v, \quad M(p) = \min_{G \in \mathcal{G}_{\text{module}}} \sum_{i \neq j} p_{ij}\big(d_G(i,j) - 1\big). \tag{4}$$

H measures traffic that the best hub must relay; M measures detour after fitting a modular pattern. Computing M need not evaluate ninety graphs. Let $\mathcal{P}$ be the ten unordered partitions into two triads and use the external incident loads $b_s(v)$. The gateway identity gives

$$M = \min_{\mathcal{B} \in \mathcal{P}} \left[2\epsilon_{\mathcal{B}} - \max_{v \in B_1} b_1(v) - \max_{v \in B_2} b_2(v) \right]. \tag{5}$$

Thus ten partition scores replace the ninety gateway comparisons. Partition fitting remains necessary; Φ is a summary of fitted costs, not a feature available before that work. Put $\lambda = a/(cW)$. Relative to direct delivery, the best hub and module have normalized excess costs

$$\frac{L_{\text{hub}} - \mathsf{L}_p(D)}{cW} = H - 20\lambda, \qquad \frac{L_{\text{module}} - \mathsf{L}_p(D)}{cW} = M - 16\lambda, \tag{6}$$

where $L_{\text{hub}} = \min_v \mathsf{L}_p(S_v)$ and $L_{\text{module}} = \min_{G \in \mathcal{G}_{\text{module}}} \mathsf{L}_p(G)$. Therefore the smallest of 0, $H - 20\lambda$ and $M - 16\lambda$ selects the cheapest family, retaining ties. The attaining hub, partition and gateways specify its connections. The derivation and examples show why both coordinates are needed. Figure 2 displays this identity, then tests predictions on future packets. With free permissions, $\lambda = 0$ and direct delivery is always optimal.

Same concentration, different efficient patterns Consider two exact demand profiles at $a = 4$ and $cW = 256$, before paying to learn them. Uniform demand on the twelve directed pairs within two triads gives $(H, M) = (2/3, 0)$. Modules cost 312 units, below direct delivery at 376. Uniform demand on all thirty pairs has the same $H = 2/3$, but $M = 4/5$. Modules now cost 516.8, so direct delivery wins. The best hub costs $1400/3$ units in both cases. Thus the second coordinate separates localized exchange from diffuse exchange even when hub concentration is identical.

4.2 When an estimated profile can guide revision

Fix a finite candidate family $\mathfrak{G}$ before collecting $m_{\text{obs}} \geq 1$ fresh independent request observations with pair law p^0. Suppose the next horizon has law p^1 with $\text{TV}(p^0, p^1) \leq \nu$. Let $D_{\max}$ bound every candidate distance and $\delta \in (0,1)$. Under the sampling setup, define

$$\eta = cW(D_{\max} - 1)\left[\sqrt{\frac{\log(2|\mathfrak{G}|/\delta)}{2m_{\text{obs}}}} + \nu \right]. \tag{7}$$

Proposition B.2 bounds every estimated cost error by η with probability at least $1 - \delta$. Minimizing estimated cost consequently loses at most 2η against the best candidate under p^1. Its sufficient paid margin compares two uncertain costs. Against direct delivery, however, the reference cost is known exactly and only additional forwarding needs estimation.

For each alternative G, let $r_G = \max_{i \neq j}(d_G(i,j) - 1)$ and let U_G be the simultaneous upper confidence limit for mean normalized detour $(d_G - 1)/r_G$. Proposition B.3 gives

$$\mathsf{L}_{p^1}(D) - \mathsf{L}_{p^1}(G) \geq a\big[N_a(N_a - 1) - |E(G)|\big] - cWr_G \min\{1, U_G + \nu\}. \tag{8}$$

The limit uses the observed detour rather than a worst case error for two costs. If the right side exceeds all learning and switching expenses, the selected change has positive net expected transport gain. This can authorize changes excluded by the original margin. Both bounds require a justified drift bound and a fixed observation window; similar histograms alone do not establish either premise.

A nonempty region for paid revision At $a = 4$ and $cW = 256$ without drift, the original margin cannot accept any change. The refined bound permits a module with zero observed detour to have a positive lower net saving after all learning expense. Appendix F.4 gives the calculations and retrospective comparison, including declined changes.

4.3 Profile selection and feedback guided revision

The transport selector chooses a graph from measured demand. Its cost prediction is distinct from AutoForm's use of task feedback to compare continuations. We first specify the evaluated selection problem, then its interface to the latter procedure.

Exact selection without enumerating gateways The frozen experiment computes all 97 candidate costs from the same observation window. Algorithm 1 implements the same minimization through (5). For each partition, retain the canonical graph among its optimal gateway choices and check the incumbent separately. This preserves the original rule that ties favor the incumbent, then direct communication, then canonical edge order. An empty window retains the incumbent without estimating a profile.

Algorithm 1 Factorized profile selection

Require: Fixed six worker family, incumbent G_0, nonempty observations and prices
1: Count observed endpoint pairs and compute hub incident loads
2: **for** each of the ten triad partitions **do**
3: Compute leakage and external incident loads; maximize each gateway load
4: Retain the canonical minimizing graph and its cost using (5)
5: **end for**
6: Compare retained graphs, six hubs, direct delivery and G_0 with the fixed tie rule
7: Record the selected graph and prediction before continuation; charge all selection work

Whole patterns as AutoForm proposals The adapter converts a target T into a package adding $E(T) \setminus E(G_0)$ and removing $E(G_0) \setminus E(T)$. Workers, roles and checkpoint remain fixed. Appendix C.1 specifies its ranking, package retention and feedback scan. Factorized minimization returns one optimum, not that complete ranking.

Which guarantee belongs to the selected graph? The adapter can leave $\mathfrak{G}$, so its applicable result is the paid search guarantee under accurate paired feedback, not the transport bound. Conversely, factorized selection retains the transport bound's original family size, since its representatives depend on observations. Algorithm 2 and Figure 4 describe the separate feedback procedure.

Connecting requests to task quality The proposed HotpotQA evaluation records requests before filtering, forecasts differences before continuation, and scores answers and supporting facts independently (Yang et al., 2018). Appendix B.6 specifies this unexecuted test of the combined system.

5 Experiments

The controlled study separates decision fidelity, endpoint matching and future cost. Original policy predictions precede future sampling. Appendix F distinguishes prospective controls (Table 2), post hoc ablations and later timing measurements.

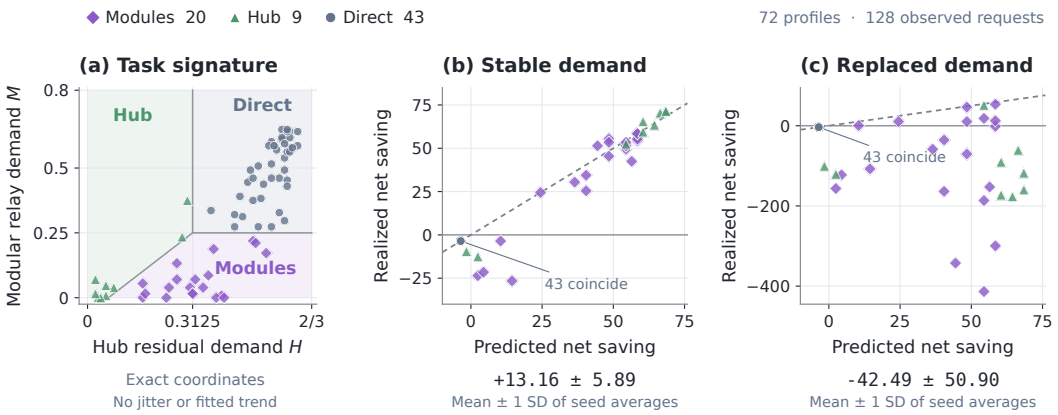

Figure 2: **Task shape predicts the pattern; persistence determines its saving.** At $a = 4, c = 1, W = 256$, 128 requests give each of 72 profiles its signature and cost forecast before continuation. Shading shows the cheapest family in (6). Savings include observation, selection and switching relative to the direct fixed control. Dashed lines show equal predicted and realized savings; vertical scales differ. Summaries give mean $\pm$ one sample SD over 24 seed averages, each averaging three dispersion settings. Coincident profiles are retained without jitter.

Demand and comparisons Six workers share direct communication, six hubs and ninety modular graphs. Independent Dirichlet laws p^0, q on the thirty directed worker pairs have dispersion $\kappa \in \{0.05, 0.5, 5\}$. Observations follow p^0; 256 future packets follow $p^1 = (1 - \omega)p^0 + \omega q$. We cross observation counts $\{32, 128, 512\}$, channel prices $a \in \{0, 1, 4, 16\}$ and replacement weights $\omega \in \{0, 0.25, 0.5, 1\}$, with hop cost $c = 1$. Every graph delivers the same packets. This holds evidence availability fixed while varying the expense of carrying it.

Twenty four development seeds select one fixed graph per price; 24 separate evaluation seeds yield 3456 repeated cells. Methods share future packets, with nested observation windows within each seed and dispersion. Table 2 compares empirical selection with four original controls and the later modular control. The confidence policy receives true simulated drift.

Outcomes and uncertainty Savings subtract fully charged policy cost from fixed graph cost on the same future packets. Profile policies pay $0.02m_{\text{obs}} + 1$ for collection and selection even without a change, plus two for switching. Expected continuation regret excludes those charges. Figures report sample SD; descriptive 95% paired intervals use a t distribution over 24 seed averages, each pooling three dispersions. Parameter cells are not independent replicates, and intervals are not adjusted across settings. Appendix F.2 gives both endpoints and the reconstruction procedure.

5.1 The task signature predicts pattern choice

Figure 2 places observed profiles at $a = 4$ and 128 observations in the regions of (6). Concentrated $\kappa = 0.05$ demand selects seven hubs and seventeen modules; diffuse $\kappa = 5$ demand always selects direct delivery. Both independent reconstruction and factorized selection recover the frozen choices in all 3456 cells. This establishes fidelity to the rule. Only the subsequent packets test whether those choices are useful.

With stable demand and 128 observations, empirical selection reduces total modeled cost by 13.16 units relative to the development fixed graph at $a = 4$, with interval $[10.67, 15.65]$. At $a = 16$, the reduction is 59.26 units with interval $[43.82, 74.69]$. Permuting the same observations instead adds 66.02 and 24.30 units. Notice that relabeling preserves Φ and the set of minimizing families but changes the selected connections. Therefore both pattern choice and matching workers to that pattern matter.

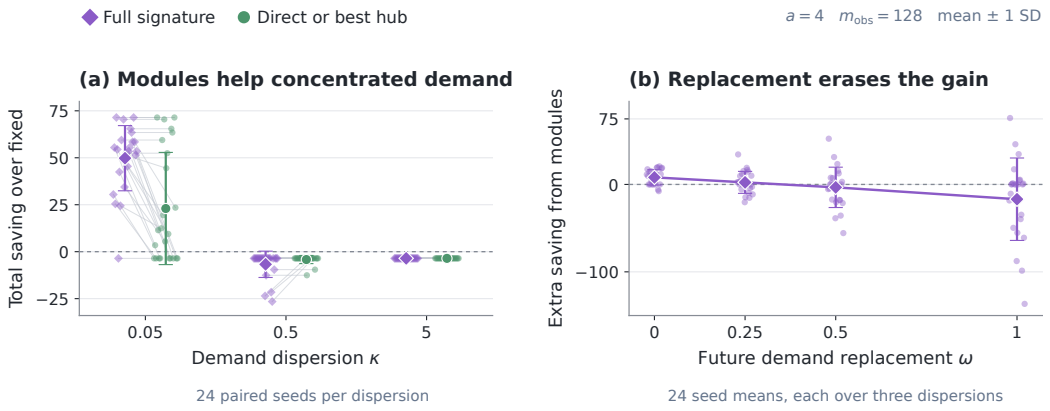

Figure 3: **Modular access helps concentrated demand, not every profile.** Post hoc comparison at $a = 4$ and 128 observations using identical observed pairs and future packets. Left, full selection versus direct delivery or the best hub. Right, the additional saving from modular candidates as demand changes. Small markers retain every seed; horizontal offsets aid visibility and faint connectors pair policies. Large markers and bars show mean and one sample SD over 24 seeds, averaging three dispersion settings within each seed on the right. All specified learning and switching charges are included.

5.2 What modular detour adds beyond concentration

At $a = 4$ and $cW = 256$, modules strictly beat both alternatives when $M < 1/4$ and $H - M > 1/16$. The latter condition requires saved forwarding relative to the best hub to repay four additional sessions. Therefore concentrated demand favors modules when it fits groups without fitting a single hub. The ablation tests whether this fitted distinction predicts lower cost on future packets.

The ablation removes modular candidates while retaining the same observations, future packets and specified selection charges. At $a = 4$, 128 observations and stable demand, access to modules saves another 8.08 units, with paired interval $[4.38, 11.79]$. The effect is concentrated in $\kappa = 0.05$, where the saving is 26.79 units. At $\kappa = 0.5$ it instead adds 2.54 units, with an interval containing zero, while diffuse demand is exactly tied. Figure 3 retains every seed and shows when this added value disappears.

Modular access changes 20 of the 72 profile decisions. Sixteen reduce charged cost, three increase it and one ties; the other 52 choices are unchanged. The gains occur under concentrated demand and the losses at intermediate dispersion. At $a = 16$, however, the additional modular saving is -0.15 units with interval $[-2.31, 2.01]$. Thus a larger gain against the fixed graph need not imply greater value from modules.

Measured selection expense With request counting included, process CPU time is 0.178 ± 0.003 ms for factorization, 0.284 ± 0.005 ms for a cached full scan and 0.081 ± 0.001 ms for the seven candidate control, giving means and sample SD over 24 seed averages. Appendix F.3 separates setup and the remaining transport calibration.

5.3 Observation expense and drift limit the advantage

At $a = 4$ without drift, increasing the fixed observation window from 32 to 128 to 512 reduces expected continuation regret from 5.03 to 1.45 to 0.27 units. However, net savings are 10.85, 13.16 and 6.51 units. The largest window therefore fails to repay its added collection expense.

Replacing future demand reverses the advantage of the stale choice. With 128 observations and $\omega = 1$, the $a = 4$ saving becomes an excess of 42.49 units, with interval $[21.00, 63.99]$. At $a = 16$, the excess is 21.66 with interval $[2.36, 40.95]$. At $a \in \{0, 1\}$, even stable demand does not repay 128 observations. Appendix F.4 separately evaluates the sufficient margin.

Two ways a question combines evidence Illustrative cases

Question	Evidence from documents	Answer
	A Resolve a shared identity	
Which earlier band did the singer who died before Mother Love Bone's Apple play in?	**Mother Love Bone** Andrew Wood fronted the band and died before Apple appeared. **Return to Olympus** Wood previously sang with Malfunkshun.	**Malfunkshun** earlier band
	B Compare separate sources	
Do LostAlone and Guster have equally many members in the cited lineups?	**LostAlone** 3 listed members; **Guster** 3 founding members → 3 = 3	**Yes** equal counts

Table 1: **What useful evidence exchange must accomplish.** Questions and facts are paraphrased from Figure 1 and Table 3 of Yang et al. (2018). The first case links an identity across documents; the second compares two documented lineups. Reference answers follow the source. These are published task illustrations, not recorded agent executions.

5.4 Document reasoning examples

Table 1 illustrates two published HotpotQA questions. One links a singer's identity to his earlier band; the other compares two lineups. Useful communication must connect facts found by different workers. Thus the relevant topology depends on who knows each fact, not only the question type.

The proposed study records requests and connectivity forecasts before continuation, then compares answer and evidence quality against matched rewiring, fixed communication, selective messaging, a manager and a singleton. One inference budget includes discovery costs. Appendix B.6 specifies this unexecuted evaluation. Appendix H retains all six methods and failures from the planning comparisons behind Figure 1.

6 Concluding remarks

The task signature $\Phi(p) = (H(p), M(p))$ and communication prices determine the cheapest permitted family. Its explanation connects evidence demand to the cost of acting on it. The ablation finds value from modules under concentrated, persistent demand but not diffuse demand, making the scope of that explanation experimentally visible.

Explaining why this pattern The cost contrasts explain why a module is preferred. Against direct exchange, it trades fewer maintained channels for additional forwarding. Against a hub, reduced forwarding must repay the extra channels. The paired profiles show how equal hub concentration can favor opposite patterns because grouping changes forwarding. The attaining partition and gateways identify the connections; the signature alone identifies only family costs.

From explanation to prediction Fidelity to the rule does not guarantee future benefit. The learning guarantee gives sufficient conditions, while modular comparisons under replacement remain unresolved. Real agents may discover or combine evidence while forwarding it. The document study must test request usefulness and complete expense together. Human understanding requires separate evaluation.

AI USE STATEMENT

Generative AI assistance, including OpenAI Codex, supported manuscript drafting and editing, checks of mathematical notation and proofs, and the revision and reorganization of sections. The reported numerical evidence comes from executable controlled studies, not from language model responses acting as experimental agents. Verification includes source and artifact hashes, regression tests, execution replay, and reconstruction of reported comparisons as documented in the supplementary material. These checks do not replace author review. The authors retain responsibility for the final text, mathematical claims, code and empirical interpretation.

REPRODUCIBILITY STATEMENT

Appendix B gives the assumptions and proofs for the task profile analysis, while Appendix C specifies AUTOFORM, its conditional guarantees and execution rules. Appendix F documents the controlled transport design, seeds, costs and uncertainty analysis. Appendices G and H describe the supporting constructor and planning comparisons. Reported numerical results were checked by independent reconstruction from recorded executions. The research archive retains the fixed protocols, source code, raw records and reconstruction outputs.

ACKNOWLEDGMENTS

Acknowledgments are withheld for anonymous review.

Ethics statement Automatic organization can reduce computation and expose responsibility, but it can also concentrate authority, amplify erroneous messages, or conceal unsafe delegation. Deployments should log formation and revision, enforce least privilege, test failure containment, preserve human override, and provide an independent shutdown path.

REFERENCES

Hasan Abasi, Nader H. Bshouty, and Hanna Mazzawi. On exact learning monotone DNF from membership queries. *arXiv preprint arXiv:1405.0792*, 2014. URL https://arxiv.org/abs/1405.0792.

Matthew Andrews, Spyridon Antonakopoulos, and Lisa Zhang. Minimum-cost network design with (dis)economies of scale. *SIAM Journal on Computing*, 45(1):49–66, 2016. doi: 10.1137/110825959.

Wenruo Bai and Jeff Bilmes. Greed is still good: Maximizing monotone Submodular+Supermodular (BP) functions. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80, pp. 304–313, 2018. URL https://proceedings.mlr.press/v80/bai18a.html.

Mert Cemri, Melissa Z. Pan, Shuyi Yang, Lakshya A. Agrawal, Bhavya Chopra, Rishabh Tiwari, Kurt Keutzer, Aditya Parameswaran, Dan Klein, Kannan Ramchandran, Matei Zaharia, Joseph E. Gonzalez, and Ion Stoica. Why Do Multi-Agent LLM Systems Fail? *arXiv preprint arXiv:2503.13657*, 2025. URL https://arxiv.org/abs/2503.13657v3. Version 3.

Finale Doshi-Velez and Been Kim. Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608*, 2017.

Shengran Hu, Cong Lu, and Jeff Clune. Automated design of agentic systems. *arXiv preprint arXiv:2408.08435*, 2024. URL https://arxiv.org/abs/2408.08435.

Matthew Jackson. *Social and Economic Networks*. Princeton University Press, 2008.

Sayash Kapoor, Benedikt Stroebl, Zachary S. Siegel, Nitya Nadgir, and Arvind Narayanan. AI Agents That Matter. *arXiv preprint arXiv:2407.01502*, 2024. URL https://arxiv.org/abs/2407.01502.

Shiyuan Li, Yixin Liu, Qingsong Wen, Chengqi Zhang, and Shirui Pan. Assemble your crew: Automatic multi-agent communication topology design via autoregressive graph generation. *arXiv preprint arXiv:2507.18224*, 2025. URL https://arxiv.org/abs/2507.18224v4. Version 4. Accepted at AAAI 2026.

Zhou Liu, Chaoyang Han, Zewei Pan, Zeli Su, and Wentao Zhang. ExRole: From team trajectories to executable roles in multi-agent language models. *arXiv preprint arXiv:2608.11949*, 2026. URL https://arxiv.org/abs/2608.11949.

Tim Miller. Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, 267:1–38, 2019. doi: 10.1016/j.artint.2018.07.007.

Marco Ribeiro, Sameer Singh, and Carlos Guestrin. "why should I trust you?": Explaining the predictions of any classifier. In *Proceedings of NAACL HLT 2016 Demonstrations*, pp. 97–101, 2016. doi: 10.18653/v1/N16-3020.

Cynthia Rudin. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1:206–215, 2019. doi: 10.1038/s42256-019-0048-x.

David Silver. Cooperative pathfinding. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, volume 1, pp. 117–122, 2005. doi: 10.1609/aiide.v1i1.18726.

Roni Stern, Nathan R. Sturtevant, Ariel Felner, Sven Koenig, Hang Ma, Thayne T. Walker, Jiaoyang Li, Dor Atzmon, Liron Cohen, T. K. Satish Kumar, Eli Boyarski, and Roman Barták. Multi-agent pathfinding: Definitions, variants, and benchmarks. In *Proceedings of the International Symposium on Combinatorial Search*, volume 10, pp. 151–158, 2019. doi: 10.1609/socs.v10i1.18510.

Sainbayar Sukhbaatar, Arthur Szlam, and Rob Fergus. Learning multiagent communication with backpropagation. In *Advances in Neural Information Processing Systems*, 2016.

Tongtong Wu, Yanming Li, Ziye Tang, Chen Jiang, Linhao Luo, Guilin Qi, Shirui Pan, and Gholamreza Haffari. CARD: Towards conditional design of multi-agent topological structures. In *International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=JgvJdICc6P.

Zhilin Yang, Peng Qi, Saizheng Zhang, Yoshua Bengio, William Cohen, Ruslan Salakhutdinov, and Christopher D. Manning. HotpotQA: A dataset for diverse, explainable multi-hop question answering. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pp. 2369–2380, 2018. doi: 10.18653/v1/D18-1259. URL https://aclanthology.org/D18-1259/.

Guibin Zhang, Yanwei Yue, Zhixun Li, Sukwon Yun, Guancheng Wan, Kun Wang, Dawei Cheng, Jeffrey Xu Yu, and Tianlong Chen. Cut the crap: An economical communication pipeline for llm based multi-agent systems. In *International Conference on Learning Representations*, 2025a. URL https://openreview.net/forum?id=LkzuPorQ5L.

Guibin Zhang, Yanwei Yue, Xiangguo Sun, Guancheng Wan, Miao Yu, Junfeng Fang, Kun Wang, Tianlong Chen, and Dawei Cheng. G-designer: Architecting multi-agent communication topologies via graph neural networks. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 76678–76692, 2025b. URL https://proceedings.mlr.press/v267/zhang25cu.html.

Mingchen Zhuge, Wenyi Wang, Louis Kirsch, Francesco Faccio, Dmitrii Khizbullin, and Juergen Schmidhuber. GPTSwarm: Language agents as optimizable graphs. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 62743–62767, 2024. URL https://proceedings.mlr.press/v235/zhuge24a.html.

Start with Appendix B for the task measurements and their proofs, and Appendix C for AUTOFORM and its conditional guarantees. Appendix F specifies the controlled profile study. The remaining sections give the notation guide, information limits and supporting comparisons.

A TASK OBJECTIVES AND COMPARISON TARGETS

The task index is ψ and its context x_ψ contains permitted task text, the fixed worker and tool registry, and total budget B_ψ. A graph $G = (V, E(G))$ has fixed finite agent set V and directed edge set $E(G) \subseteq \{(i,j) \in V^2 : i \neq j\}$. We identify G with $E(G)$ in set operations, so $|G| = |E(G)|$ and $G \cup T$ adds an edge package T. A feasible graph class is denoted $\mathcal{S}$. Workers have separate histories. An edge permits future delivery rather than requiring a message. Roles, primitive capabilities, authority, message generation rules and nongraph prompts remain fixed for a topology intervention. Broader fixed organizations o also specify these execution rules.

Execution returns a canonical submitted decision A and score $Y \in [0,1]$. Execution equivalence means equality of the probability laws of A on the same task, not merely equal scores. The endpoint is the latest submission committed through the runtime interface. A missing or malformed submission gives a null decision and zero score, without falling back to an earlier answer. A later runtime failure alone does not erase a commitment. Incorrect answers and citations remain predictions to be scored, not repaired outputs. Termination status and unresolved resource use are retained independently of answer quality. The common evaluator receives no reconstructed answer from uncommitted responses or private histories. Its references and final verdicts are not formation inputs. The selector sees only declared context, permitted observations, earlier organizations and resource use. Organizer access does not grant an observation to workers.

A.1 VALUE OF AN EXECUTION POLICY

Let $C_\psi(\mathcal{F})$ be the total cost incurred by the actual execution of formation policy $\mathcal{F}$. It includes pilot work, successful and rejected proposals, worker and tool calls, attempted communication, final aggregation and any declared switching costs, without double counting. Let $\mathsf{Risk}_\psi(\mathcal{F})$ be its declared risk penalty. These are random execution quantities. For fixed nonnegative risk weight λ_f, define

$$\mathcal{J}_\psi(\mathcal{F}; B_\psi) = \mathbb{E}_{\mathcal{F}}[Y - C_\psi(\mathcal{F}) - \lambda_f \mathsf{Risk}_\psi(\mathcal{F})]\,,$$

subject to the common resource caps. This expectation covers the complete execution, including failed pilots and termination, with histories retained throughout. Prices and penalties are fixed before comparison. When common prices are unavailable, task outcomes and resource dimensions are reported separately.

For a fixed organization o under specified execution rules, use the shorthand

$$\mathsf{J}_\psi(o) = \mathsf{V}_\psi(o) - \mathsf{C}(o) - \lambda_f \mathsf{Risk}_\psi(o), \qquad \mathsf{V}_\psi(o) = \mathbb{E}[Y \mid o]. \tag{9}$$

Here $\mathsf{C}(o)$ and $\mathsf{Risk}_\psi(o)$ are the corresponding expected cost and risk. Proposition D.1 uses this static identity, while comparisons of complete policies use $\mathcal{J}_\psi$.

A.2 ORGANIZATION AND COLLECTIVE GAINS

Fix a positive integer m_{solo} and that many declared singleton controls, indexed by $[m_{\mathrm{solo}}] = \{1, \ldots, m_{\mathrm{solo}}\}$. Let $\mathcal{F}_i^{\mathrm{solo}}$ use one persistent agent with the same permitted information, primitive capabilities and total budget, except for explicitly nontransferable resources. Relative to the strongest declared singleton control, collective gain is

$$\mathsf{CGain}_\psi(\mathcal{F}; B_\psi) = \mathcal{J}_\psi(\mathcal{F}; B_\psi) - \max_{i \in [m_{\mathrm{solo}}]} \mathcal{J}_\psi(\mathcal{F}_i^{\mathrm{solo}}; B_\psi). \tag{10}$$

Let $\mathcal{F}_{\mathrm{freeze}}$ share the initial formation and pilot, then retain the initial graph. It receives the same total budget and may reuse saved revision resources. Organization gain is

$$\mathsf{OGain}_\psi(\mathcal{F}; B_\psi) = \mathcal{J}_\psi(\mathcal{F}; B_\psi) - \mathcal{J}_\psi(\mathcal{F}_{\mathrm{freeze}}; B_\psi). \tag{11}$$

Neither quantity must be positive. Applying a proposal versus a paid sham instead compares connectivity at equal proposal cost and remaining budget. Appendix C specifies that distinction and the treatment of failed prefixes.

A.3 Notation and proof guide

We write $\mathbf{1}\{E\}$ for the indicator of event E, $\mathcal{L}(Z)$ for the law of Z, and $\mathbb{E}$ and $\Pr$ for expectation and probability under the stated law. Graph set operations use the edge set convention. The checkpoint definitions distinguish remaining budget b, search reservation r and deployment cap $B = b - r$. The feedback setup defines Q, n, ξ and its simultaneous accuracy event. Proposition C.1 and its proof use these quantities.

The constructor definitions specify edge scores, package width w and retention limit P. Proposition C.2 uses the strict separation condition; its proof covers both construction and selection. The hidden evidence model instead uses $L, k, b_{\mathrm{qry}}, N$ for channel count, required size, width cap and query limit.

The task profile defines the pair law p and packet horizon W. The pattern definitions give concentration γ_v, leakage $\epsilon_{\mathcal{B}}$ and gateway detour $r_{\mathcal{B}}$. Proposition 4.1 relates them to topology cost through its path calculation. The task signature defines $\Phi(p) = (H(p), M(p))$ and cost ratio $\lambda = a/(cW)$; its decision regions compare family minima. The sampling setup defines the family $\mathfrak{G}$, window m_{obs} and drift bound ν used in Proposition B.2 and its proof. The diagnostic model separately defines history laws and total variation.

The coverage quantities separate missed proposals from feedback error, while the threshold component control tests candidate construction. The search cost definitions specify the reservation for a complete policy. Finally, the task level estimator and matched connectivity control distinguish complete policy value from the effect of selected connections. The feedback definitions separate verifier acceptance from independently evaluated task correctness.

B Task profiles and efficient communication

Setup and traffic costs are standard network design ingredients (Andrews et al., 2016). Here they relate measurable evidence demand to communication patterns for a fixed agent pool.

B.1 Profiles depend on the current information state

Fix a task, resumable checkpoint, worker set V of size $N_a \geq 3$ and all nongraph execution rules. Let $\Omega = \{(i, j) \in V^2 \mid i \neq j\}$. A unit evidence packet has an origin and a required recipient in Ω. Its pair law is $p \in \Delta(\Omega)$, meaning $p_{ij} \geq 0$ and $\sum_{(i,j)\in\Omega} p_{ij} = 1$. The horizon contains a fixed known integer $W \geq 1$ packets. For expected cost it suffices that each packet has marginal pair law p; independence is not needed. Random or unknown volume requires an additional model and is not covered by the estimation result below.

Profiles depend on current information, so two executions of one question can differ. A zero packet horizon pays only session costs and lies outside this normalized model.

Assumption B.1 (Faithful transport with declared channel overhead)**.** *For the fixed task profile, every candidate graph is strongly connected. Workers originate the same packets with the same contents, destinations and pair law under every candidate. An intermediate relay forwards a packet unchanged without exposing it to its own reasoning process. Every destination receives every required packet. Routing uses a shortest directed path and has sufficient capacity, horizon and context allowance to complete delivery. Costs comprise $a \geq 0$ per enabled directed session over the horizon and $c > 0$ per packet per hop, in a common unit. Other work is identical. No compression, multicast saving, congestion, intermediate reasoning or deadline penalty is included.*

Linearity gives (2). Session overhead must represent actual setup or maintenance, with $a = 0$ when absent. Observation, proposal and switching expenses are separate. Changed demand, contents or reasoning after differently timed messages require direct outcome evaluation.

B.2 Three feasible pattern families

The complete graph D contains every pair in Ω. The star S_v contains (i, v) and (v, i) for all $i \neq v$. Fix a partition $\mathcal{B} = \{B_1, \ldots, B_K\}$ of V into nonempty groups, $1 \leq K \leq N_a$, and a gateway $g_s \in B_s$ for each group. The modular graph $G_{\mathcal{B}}$ contains every directed pair within a group and every directed

pair between distinct gateways. Group membership and gateway choice specify connectivity, not new worker roles. Singleton groups and $K = 1, N_a$ are allowed.

The profile features are

$$\gamma_v = \sum_{j \neq v} (p_{vj} + p_{jv}), \tag{12}$$

$$\epsilon_{\mathcal{B}} = \sum_{s \neq t} \sum_{i \in B_s, j \in B_t} p_{ij}, \tag{13}$$

$$r_{\mathcal{B}} = \sum_{s \neq t} \sum_{i \in B_s, j \in B_t} p_{ij} \big(2 - \mathbf{1}\{i = g_s\} - \mathbf{1}\{j = g_t\}\big). \tag{14}$$

Concentration and leakage lie in $[0, 1]$, and $0 \leq r_{\mathcal{B}} \leq 2\epsilon_{\mathcal{B}}$. Detour distinguishes gateway traffic from other cross group traffic, which leakage alone cannot. These features do not measure semantic complementarity.

For worker $v \in B_s$, define its external incident traffic

$$b_s(v) = \sum_{j \notin B_s} (p_{vj} + p_{jv}). \tag{15}$$

Then $r_{\mathcal{B}} = 2\epsilon_{\mathcal{B}} - \sum_s b_s(g_s)$. At a fixed partition, edge count is constant, so each optimal gateway independently maximizes $b_s(v)$ in its group.

Proof of Proposition 4.1. Under Assumption B.1, D uses $N_a(N_a - 1)$ edges and one hop for every packet. The star uses $2(N_a - 1)$ edges. Traffic incident to its hub uses one hop and all other traffic uses two, giving mean distance $2 - \gamma_v$.

Within group edges and between gateway edges are disjoint, so $G_{\mathcal{B}}$ has $m_{\mathcal{B}} = \sum_s |B_s|(|B_s| - 1) + K(K - 1)$ edges. Within group traffic travels directly. A cross group path must pass through the source and destination gateways, so its shortest length is $3 - \mathbf{1}\{i = g_s\} - \mathbf{1}\{j = g_t\}$. Its mean distance is therefore $1 + r_{\mathcal{B}}$. Substitution proves the cost identities and subtraction gives both thresholds. The star edge count is independent of its hub, so maximizing γ_v minimizes cost. Expanding (14) gives the external traffic identity and the independent gateway rule. Finally every delivery between distinct workers needs at least one hop, establishing the lower bound cW when $a = 0$. □

A modular candidate also beats every star exactly when $cW(1 - \max_v \gamma_v - r_{\mathcal{B}}) > a[m_{\mathcal{B}} - 2(N_a - 1)]$. Together with its direct comparison in Proposition 4.1, this determines when it beats both alternatives. Replacing $r_{\mathcal{B}}$ by $2\epsilon_{\mathcal{B}}$ gives sufficient conditions.

B.3 The task signature and its decision regions

For six workers, there are $\binom{6}{3}/2 = 10$ unordered partitions into two groups of three and nine gateway pairs per partition. Thus $\mathcal{G}_{\text{module}}$ in (4) contains ninety graphs. Together with six stars and direct communication, these form the fixed family of 97 candidates. Their edge counts are 14, 10 and 30, respectively. Proposition 4.1 gives

$$L_{\text{hub}} = 10a + cW(1 + H), \qquad L_{\text{module}} = 14a + cW(1 + M), \qquad \mathsf{L}_p(D) = 30a + cW.$$

The finite minima are attained. Subtracting direct cost and dividing by $cW > 0$ proves (6). Equivalently, the set of optimal family labels is

$$\arg\min_{k \in \{\text{direct}, \text{hub}, \text{module}\}} s_k, \qquad (s_{\text{direct}}, s_{\text{hub}}, s_{\text{module}}) = (0, H - 20\lambda, M - 16\lambda).$$

The hub beats direct delivery exactly when $H < 20\lambda$, the module beats direct delivery exactly when $M < 16\lambda$, and the module beats the hub exactly when $H - M > 4\lambda$. Equalities retain ties. At $a = 4, c = 1, W = 256$, these thresholds are 0.3125, 0.25 and 0.0625.

The signature determines family costs, not worker identities. A hub attains $\max_v \gamma_v$, while a module attains the minimum detour over partitions and gateways. Relabeling workers preserves Φ but changes these connections. The implementation resolves cost ties by retaining the incumbent, then preferring direct delivery, then canonical edge order.

Factorization and exact selection For each unordered pair put $q_{ij} = p_{ij} + p_{ji}$. Summing q_{ij} across a partition gives $\epsilon_{\mathcal{B}}$ and its incident loads $b_s(v)$. Minimizing $r_{\mathcal{B}} = 2\epsilon_{\mathcal{B}} - b_1(g_1) - b_2(g_2)$ proves (5). Retain the canonical maximizing gateways for each partition, all six stars and direct delivery. Check the incumbent separately to preserve its priority among ties, then prefer direct delivery and canonical edge order. This preserves the full minimum with exact rational comparisons on integer counts.

Partition enumeration remains necessary. The reduction returns an optimum, not the adapter's ranked list. Because representatives depend on the observations, the learning bound still uses $|\mathfrak{G}| = 97$ and the direct comparison still allocates error over 96 alternatives.

Since $\sum_v \gamma_v = 2$, $0 \le H \le 2/3$. Every modular graph has total detour 24 over the thirty ordered pairs. Symmetry therefore gives mean detour $4/5$ for each pair when averaged over all ninety graphs, so $0 \le M \le 4/5$. These bounds frame the plot; they do not assert that every point in the rectangle is attainable.

B.4 Prediction from observations

Condition on proposal information collected before the estimation sample. Fix a nonempty finite family $\mathfrak{G}$ of strongly connected graphs, an incumbent $G_0 \in \mathfrak{G}$, and an integer $D_{\max} \ge 1$ bounding every pair distance in every candidate. The preceding families permit $D_{\max} = 3$. Fix exact known a, c, W as in Assumption B.1. Collect $m_{\mathrm{obs}} \ge 1$ conditionally independent pair observations Z_t with common law p^0 on Ω. For $Z = (i, j)$ write $d_G(Z) = d_G(i, j)$ and define

$$\widehat{\mathsf{L}}(G) = a|E(G)| + \frac{cW}{m_{\mathrm{obs}}} \sum_{t=1}^{m_{\mathrm{obs}}} d_G(Z_t).$$

All graphs use the same observations. Empirical frequencies estimate the features in (12), (13) and (14). Let the next horizon have a fixed law p^1 under the same conditioning, with

$$\mathrm{TV}(p^0, p^1) = \tfrac{1}{2} \sum_{i \ne j} |p^0_{ij} - p^1_{ij}| \le \nu, \qquad 0 \le \nu \le 1.$$

This is a stipulated distribution change bound, not automatically a bound inferred from observed drift. The guarantees are conditional on the proposal information and these sampling assumptions; integration gives the unconditional statement when they hold almost surely.

Proposition B.2 (Cost prediction and a sufficient revision margin)**.** *Under the estimation setup, fix $\delta \in (0, 1)$ and let η be (7). With probability at least $1 - \delta$,*

$$\max_{G \in \mathfrak{G}} |\widehat{\mathsf{L}}(G) - \mathsf{L}_{p^1}(G)| \le \eta.$$

On this event any empirical minimizer $\widetilde{G} \in \arg\min_{G \in \mathfrak{G}} \widehat{\mathsf{L}}(G)$ satisfies

$$\mathsf{L}_{p^1}(\widetilde{G}) - \min_{G \in \mathfrak{G}} \mathsf{L}_{p^1}(G) \le 2\eta.$$

If $C_{\mathrm{change}} \ge 0$ upper bounds all additional observation, proposal and switching expenses relative to retaining G_0, then

$$\widehat{\mathsf{L}}(G_0) - \widehat{\mathsf{L}}(\widetilde{G}) > C_{\mathrm{change}} + 2\eta \quad \Longrightarrow \quad \mathsf{L}_{p^1}(G_0) - \mathsf{L}_{p^1}(\widetilde{G}) > C_{\mathrm{change}}.$$

Its proof follows.

Proof. For Proposition B.2, first suppose $D_{\max} > 1$. Each $d_G(Z_t)$ lies in $[1, D_{\max}]$. Hoeffding's inequality and a union bound over the fixed family give, simultaneously with probability at least $1 - \delta$,

$$|\widehat{\mathsf{L}}(G) - \mathsf{L}_{p^0}(G)| \le cW(D_{\max} - 1)\sqrt{\frac{\log(2|\mathfrak{G}|/\delta)}{2m_{\mathrm{obs}}}}.$$

For a function with range length $D_{\max} - 1$, changing its law from p^0 to p^1 changes its expectation by at most $(D_{\max} - 1)\mathrm{TV}(p^0, p^1)$. Multiplication by cW and the triangle inequality prove the first claim. For $D_{\max} = 1$, all distances are one and the costs are known exactly, so the claim holds with $\eta = 0$.

Let G^* minimize true cost in the finite family. On the simultaneous event,

$$\mathsf{L}_{p^1}(\widetilde{G}) \le \widehat{\mathsf{L}}(\widetilde{G}) + \eta \le \widehat{\mathsf{L}}(G^*) + \eta \le \mathsf{L}_{p^1}(G^*) + 2\eta.$$

Subtracting the two endpoint errors gives a true saving at least the predicted saving minus 2η, proving the revision margin for any selected candidate in the family. □

The regret bound concerns continuation cost before paying for learning. Only the last comparison charges that work. If probing reduces the continuation volume relative to the incumbent, the common W premise fails. Estimated prices or volume need additional error bounds. Candidate construction using the same sample is not covered unless all possible outputs lie in a fixed larger family whose full size is used. Repeated epochs require a separate simultaneous confidence rule.

B.5 A DIRECT COMPARISON THAT USES OBSERVED DETOUR

Retain the conditional estimation setup with incumbent D. Let $\mathfrak{A} = \mathfrak{G} \setminus \{D\}$ be nonempty and fix $\delta_G > 0$ for $G \in \mathfrak{A}$ before sampling, with $\sum_G \delta_G \le \delta$. Define

$$r_G = \max_{i \ne j}(d_G(i,j) - 1) > 0, \quad X_{G,t} = \frac{d_G(Z_t) - 1}{r_G}, \quad \overline{X}_G = \frac{1}{m_{\mathrm{obs}}} \sum_{t=1}^{m_{\mathrm{obs}}} X_{G,t}.$$

For $x, u \in [0,1]$, let $\mathrm{kl}(x,u) = x\log(x/u) + (1-x)\log((1-x)/(1-u))$, with the continuous endpoint conventions and infinite value for a positive numerator over zero. Set

$$U_G = \sup\left\{ u \in [\overline{X}_G, 1] \;\middle|\; \mathrm{kl}(\overline{X}_G, u) \le \frac{\log(1/\delta_G)}{m_{\mathrm{obs}}} \right\}. \tag{16}$$

The set is nonempty since it contains $\overline{X}_G$. Here $r_G = 1$ for a star and $r_G = 2$ for a module. The study family permits the equal allocation $\delta_G = \delta/96$. The normalized variables need not be Bernoulli.

Proposition B.3 (Paid improvement over direct delivery)**.** *Under the direct comparison setup, with probability at least $1 - \delta$, inequality (8) holds simultaneously for all $G \in \mathfrak{A}$. Let $C_G \ge 0$ upper bound all additional observation, selection and switching expenses relative to retaining D. Any graph chosen from these observations whose lower bound exceeds C_G has*

$$\mathsf{L}_{p^1}(D) - \mathsf{L}_{p^1}(G) - C_G > 0$$

on this event. The proof accounts for selecting the graph on the same observations.

Proof. For Proposition B.3, write $\mu_G^0 = \mathbb{E}_{p^0} X_{G,t}$. For $s < 0$, convexity on $[0,1]$ gives

$$\mathbb{E}e^{sX_{G,t}} \le 1 - \mu_G^0 + \mu_G^0 e^s.$$

Conditional independence, the exponential Markov inequality and minimization over $s < 0$ imply, for $0 \le x < \mu_G^0$, that

$$\Pr\{\overline{X}_G \le x\} \le \exp\{-m_{\mathrm{obs}}\,\mathrm{kl}(x, \mu_G^0)\}.$$

For fixed x, $\mathrm{kl}(x,u)$ increases on $u \ge x$. For fixed μ_G^0, it decreases on $x \le \mu_G^0$. Thus $\mu_G^0 > U_G$ is a lower tail event of probability at most δ_G, with limits covering the endpoints. A union bound yields $\mu_G^0 \le U_G$ for all alternatives with probability at least $1 - \delta$. No independence across candidates is required.

Since X_G takes values in $[0,1]$, total variation gives $\mu_G^1 = \mathbb{E}_{p^1} X_G \le \min\{1, \mu_G^0 + \nu\}$. The exact direct cost and $d_G = 1 + r_G X_G$ now give (8). Simultaneity permits any data dependent choice within the fixed family. Subtracting C_G proves the paid claim. □

The result bounds expected future cost, not every realized packet stream. Declining a change still incurs observation expense. Therefore a sufficient rule for accepted changes need not outperform never collecting observations when averaged across all executions. Its scope is a fixed window and direct incumbent, unlike the general incumbent comparison in Proposition B.2.

B.6 Operational measurements and the prospective test

Record requests before delivery filtering, with origin, intended recipient, timestamp and citations to locally available evidence. Otherwise missing edges can censor demand. A common probe exposes requests; a separate observation window evaluates forecasts after candidate partitions and gateways are fixed. Final answers, hidden support annotations and task graphs remain outside the organizer.

Requested traffic is observable, but useful evidence requires intervention on delivery and independent task scoring. Calibrate session overhead a, hop cost c and collection, selection and switching expense in one common resource unit on excluded development tasks. Hold packet counts and sizes fixed while changing enabled sessions to test whether setup cost is proportional to channel count. Record $a = 0$ for free permissions. If costs instead depend on graph shape or congestion, retain the measured graph costs rather than fitting a positive a to obtain a favorable pattern. Unequal messages require charged equal sized units or a separately specified weighted model. Correlated requests, reasoning at relays, changing contents and deadlines require direct evaluation outside the independent sample guarantee.

Before continuation, record concentration, leakage, detour, candidate costs and predicted savings. Freeze that forecast before assigning connectivity. Compare selected patterns with direct delivery, the best hub and matched rewiring under identical worker histories, tools and total budgets (Appendix C.8). Rewiring must preserve declared edge count and reachability, transferring no donor messages or histories. Score final answers and supporting evidence independently of the feedback used for selection. A capable singleton and manager with the same workers test complete policy competitiveness. The relevant interaction is whether forecasts of greater modular saving predict a larger paired benefit from modules, rather than whether task labels coincide with different graphs.

The controlled study below evaluates sampled demand. The implemented selector and adapter have separate roles; live request capture and priced inference remain unfinished. The realistic study should jointly measure forecast accuracy, complete expense and independently scored answers. Fix task strata before outcomes and retain failures, unchanged graphs and violated transport assumptions in complete policy comparisons.

C Runtime and evaluation contract

This appendix supports Section 4.3. Checkpoint guarantees concern resumable histories; complete policy comparisons also include pilot failures and costs.

Algorithm 2 AutoForm with evidence selected packages

Require: Checkpoint h, graph G_0, cap B, permitted packets, limits w, P, Q, n, ξ

1: Grow connected evidence groups; fix neighborhood $\mathcal{N}$ and both orders before feedback
2: Fix selection mode before evaluation; use completed comparisons at the query limit
3: Reserve $2Qn + 1$ continuations, packet work, feedback and administration
4: Screen using Algorithm 3; retain selected graph $G_\star$, untested candidates and costs
5: On invalid feedback, operational failure or unknown usage, halt dispatch
6: Set deployment graph $G_{\mathrm{dep}} \leftarrow G_\star$
7: Deploy G_{dep} from a fresh copy of h under the reserved cap
8: Retain rejected proposals and costs; unused search reserves stay unavailable

C.1 Whole pattern proposals from a measured profile

The selector ranks direct, star and modular targets from directed request counts over a family fixed before observation, including every available gateway assignment. Ties retain the incumbent, then prefer direct communication and canonical edge order. An empty window produces no packages and retains the incumbent.

Each target $T \neq G_0$ gives one atomic package adding $E(T) \setminus E(G_0)$ and removing $E(G_0) \setminus E(T)$. The adapter checks workers and incumbent. Predicted cost controls retention, not query order. Budget limited mode compares packages before individual edits, each lexicographically; complete scan mode orders all candidates lexicographically.

The wrapper permits one initial scan and at most one acceptance. Complete scan selection is default; using completed gains from a shorter scan requires explicitly choosing budget limited mode. A new epoch rebuilds replacement deltas from its new incumbent.

Regression tests cover replacement, ordering, checkpoint consistency, empty observations and interruption. Request generation, inference and opaque multihop forwarding remain separate requirements. Direct edge filtering does not implement the transport model's relay semantics.

Transport choices remain in $\mathfrak{G}$, whereas individual edits can take the feedback output $\widehat{G}$ into the larger search class $\mathcal{S}$. Proposition B.2 governs the former under its sampling assumptions; Proposition C.1 governs the latter under accurate feedback. Semantic relevance requires separate evaluation.

C.2 Reference search and its conditional guarantee

Algorithm 3 uses the following checkpoint setup and gain estimates.

Algorithm 3 Reference AutoForm search core

Require: Graph G_0, ordered neighbors $\mathcal{N}$, integers $Q \geq 0$, $n \geq 1$, tolerance $\xi > 0$, fixed selection mode

Set $G \leftarrow G_0$ and comparison count $q \leftarrow 0$
loop
  If $\mathcal{N}(G)$ is empty, return G with a complete rejection scan
  for each candidate in fixed order, keeping incumbent G unchanged **do**
    if $q = Q$ **then**
      In budget limited mode, accept the best completed gain above 2ξ, if any
      Return the selected graph without a complete rejection certificate
    end if
    Set $q \leftarrow q + 1$
    Collect n fresh paired gains; on invalid or unavailable feedback return uncertified
  end for
  Let G' maximize the completed mean gains, breaking ties by graph lexicographic order
  if the maximum estimate is at most 2ξ **then**
    Return G with a complete rejection scan
  end if
  Set $G \leftarrow G'$ and repeat, retaining every comparison and charge
end loop

Fix a resumable checkpoint h, its incumbent graph G_0 and a finite nonempty graph class $\mathcal{S}$ containing G_0, using the fixed worker graph convention. Each graph has a well defined continuation from h with the same worker policies, tools and authority. Write $b \geq 0$ for the remaining resource budget, $r \in [0, b]$ for the search reservation and $B = b - r$ for the common continuation cap. A declared deterministic ordered neighborhood $\mathcal{N}(G) \subseteq \mathcal{S} \setminus \{G\}$ can include atomic edge packages. Its dependence on the fixed checkpoint is suppressed. Neither the task nor the continuation objective changes during this search epoch. For $a \in \{b, B\}$, let $J_a(G)$ denote finite expected deployment welfare at cap a, including its declared deployment costs and risk but excluding the common pilot and search costs. Write $\widehat{G}$ for the returned graph and N_{acc} for the number of accepted changes. No result here assumes that natural language feedback estimates this value.

Every candidate comparison uses fresh continuations from separate copies of h. Tool state must also be isolated whenever it is mutable. Evaluation histories remain outside subsequent worker and deployment inputs. Deployment resumes a fresh copy of h under the selected graph rather than continuing a trajectory chosen for its observed outcome. The allowance r must bound all proposal and evaluation work, including failed draws and bookkeeping, and any unused part stays unavailable to deployment. Otherwise the comparison no longer has a common cap B.

For comparison t, let the sigma field $\mathcal{I}_{t-1}$ contain the preceding search information and let the incumbent G_t and candidate G'_t be measurable with respect to it. Fix real bounds $\ell < u$ and an integer sample count $n \geq 1$. Suppose the paired differences $Z_{t,i} \in [\ell, u]$ are conditionally

independent over $i = 1, \dots, n$, with conditional means

$$\mathbb{E}[Z_{t,i} \mid \mathcal{I}_{t-1}] = \Delta_t = J_B(G'_t) - J_B(G_t).$$

The two outcomes within a paired draw need not be independent. They must have the required marginals, and new draws must be fresh. There are at most $Q \geq 1$ comparisons, including an interrupted one, and $\widehat{\Delta}_t = n^{-1} \sum_i Z_{t,i}$ is used only after a complete comparison. For $\delta \in (0, 1)$ and $\xi > 0$, it suffices that

$$n \geq \frac{(u - \ell)^2}{2\xi^2} \log \frac{2Q}{\delta}.$$

Conditional Hoeffding bounds and a union bound over adaptive comparisons then give

$$\Pr(\mathcal{E}_\xi) \geq 1 - \delta, \qquad \mathcal{E}_\xi = \left\{ \left| \widehat{\Delta}_t - \Delta_t \right| \leq \xi \text{ for every completed comparison } t \right\}. \tag{17}$$

For $Q = 0$, the event $\mathcal{E}_\xi$ holds vacuously and no sample bound is needed. If failure can interrupt sampling, this statement requires the same bounded unbiased potential draws before the failure rule; an observed subset is not automatically an unbiased sample. Operational tests cannot establish these distributional assumptions.

Proposition C.1 (Conditional improvement with paid search)**.** *Use the checkpoint setup and Algorithm 3 with an integer comparison limit $Q \geq 0$. Suppose its gain estimates differ from true gains by at most $\xi > 0$. Accept only estimates above 2ξ. After N_{acc} accepted changes, the selected graph $\widehat{G}$ satisfies*

$$J_{b-r}(\widehat{G}) - C_{\mathrm{search}} - J_b(G_0) \geq N_{\mathrm{acc}}\xi - D_r - C_{\mathrm{search}}, \qquad D_r = J_b(G_0) - J_{b-r}(G_0),$$

where C_{search} is actual search expenditure in welfare units. A complete rejection scan at $\widehat{G}$ also bounds every gain in its declared neighborhood $\mathcal{N}(\widehat{G})$ by 3ξ.

The proof follows.

Proof of Proposition C.1. On the simultaneous error event $\mathcal{E}_\xi$ in (17), every accepted move has true gain greater than $2\xi - \xi = \xi$. Summing the gains over the N_{acc} accepted changes gives $J_B(\widehat{G}) - J_B(G_0) \geq N_{\mathrm{acc}}\xi$. Strict increase prevents revisiting a graph, so there are at most $|\mathcal{S}| - 1$ accepted changes. Subtract $J_b(G_0) - J_B(G_0)$ and the actual search expenditure from the telescoping inequality to obtain the stated comparison. Common pilot costs cancel because both policies start from the same checkpoint.

A complete rejection scan holds its incumbent fixed and tests every graph in $\mathcal{N}(\widehat{G})$ using fresh samples. Every estimated gain is at most 2ξ, hence every true gain is at most 3ξ. Selection at the query limit uses the same threshold and only fully completed comparisons, so the improvement proof is unchanged. A truncated scan, failed comparison or unexamined candidate supplies no complete rejection certificate. The guarantee concerns the declared neighborhood, not the global optimum in $\mathcal{S}$. □

Note that $D_r = J_b(G_0) - J_B(G_0)$ is signed because more budget need not improve a fixed policy. Actual search is charged once through C_{search}, in common welfare units, rather than through reservation r. Other units require separate reporting. Releasing reserves, reusing trial histories or changing prompts would change the comparison.

What the proposer and feedback must contribute Fix the checkpoint and caps. For $a \in \{b, B\}$, let $F_a(G) \in [0, 1]$ be expected public verifier acceptance and $R_a(G) \in [0, 1]$ expected final task score from that checkpoint at cap a. These are raw scores rather than priced welfare. The public feedback interface defines an example verifier.

Consider a normal return of the direct budget limited policy, without an operational failure. Let $\mathcal{C}_h$ contain G_0 and every candidate whose comparison with G_0 completed in the first scan. If no comparison completes, $\mathcal{C}_h = \{G_0\}$. Untested candidates do not enter this set, even if the proposer generated them. Define

$$A_h = \max_{G \in \mathcal{S}} R_B(G) - R_B(G_0), \qquad \kappa_h = \max_{G \in \mathcal{S}} R_B(G) - \max_{G \in \mathcal{C}_h} R_B(G).$$

Here A_h is the available improvement and $\kappa_h \geq 0$ is the improvement missed by the compared candidates. Neither quantity is assumed observable by the agents. Suppose the event $\mathcal{E}_\xi$ in (17) holds with $J_B = F_B$, and define

$$\varepsilon_h = \max_{G \in \mathcal{C}_h \cup \{\widehat{G}\}} |F_B(G) - R_B(G)|.$$

Then the existing selection rule implies

$$R_B(\widehat{G}) - R_b(G_0) \geq A_h - \kappa_h - 3\xi - 2\varepsilon_h - D_r^R, \qquad D_r^R = R_b(G_0) - R_B(G_0). \tag{18}$$

To see this, if the first scan rejects, every compared verifier gain is at most 3ξ. If it accepts, including at the query limit, the chosen estimate maximizes the completed estimates and its true verifier value is within 2ξ of their maximum. Later accepted moves only increase verifier value on the same error event. Thus $F_B(\widehat{G}) \geq \max_{G \in \mathcal{C}_h} F_B(G) - 3\xi$. Subtract the two endpoint verifier errors and D_r^R to obtain the display. For $Q = 0$, the returned graph is G_0 and the same inequality holds. A failed comparison does not justify deployment under this argument.

Equation (18) separates missed candidates, feedback error and the cost of reducing deployment resources. Package priority can displace individual comparisons at fixed Q, so joint proposals need not reduce κ_h. On realistic tasks, the package ablation measures final scores rather than treating lexical overlap as recall of unknown dependencies.

The implemented core supports additions, deletions, reversals and packages over a fixed edge universe. Algorithm 2 retains incomplete work and charges while enforcing the common reservation.

C.3 Evidence connected joint proposals

At the checkpoint h, each permitted packet records its owner, a finding, an unresolved question and references to that owner's previously available receipts. Packets do not change worker roles or tools. Let $\mathrm{Tok}(s)$ be the set of lowercase ASCII letter tokens of length at least three in text s, after removal of the fixed stopword list. For packets p from worker i and q from worker j, both citing at least one receipt, define

$$s(p, q) = |\,\mathrm{Tok}(p.\mathrm{finding}) \cap \mathrm{Tok}(q.\mathrm{question})|.$$

For each absent admissible edge $e = (i, j)$, let s_e be the maximum positive packet score $s(p, q)$ and retain a pair attaining it, if one exists. Ties use source and recipient packet identifiers in lexicographic order. The score defines a lexical overlap hypothesis for subsequent feedback evaluation. Let H^+ be the edge set of all retained hypotheses and $M = |H^+|$ its size.

Before evaluation, fix integer width $w \geq 2$ and package limit $P \geq 0$. Start one growth sequence at each edge of H^+. At each step, append the highest scoring unselected edge sharing an endpoint with the current group, breaking ties by directed edge lexicographic order. Ignore direction only for this adjacency test. Stop at w edges or an empty frontier. Emit every prefix of size two through w, remove duplicate edge sets, then rank groups by decreasing summed edge score and lexicographic edge tuple. Retain the first P. No trial score, receipt content or final label enters the ranking. The implementation admits $2 \leq w \leq 15$ and $0 \leq P \leq 16$.

The same packets can propose relays $\{(a, b), (b, c)\}$ or aggregation $\{(a, c), (b, c)\}$ without changing workers.

The reference core retains every admissible single addition, deletion and reversal. Distinct package graphs precede all remaining candidates, with graph lexicographic order within each class. Equal completed scalar gains use global graph lexicographic order, which need not be the comparison order. The incumbent stays fixed throughout a scan. At ordinary query exhaustion, select the best fully compared gain above 2ξ, if any. Record the entire inventory, completed prefix and untested remainder. Partial scans give no neighborhood rejection certificate. Operational failures never authorize selection from an incomplete scan or further deployment.

Packets are bound to the task, prefix, round, actors and graph. Receipt checks establish ownership and availability; packet generation and proposal computation are charged.

Proposition C.2 (Coverage from separated evidence)**.** *The constructor generates at most $M(w-1)$ prefixes. It includes every weakly connected $T \subseteq H^+$ with $2 \leq |T| \leq w$ whose edge scores strictly exceed those outside T incident to its vertices. The condition is vacuous for a whole weak component.*

Suppose T survives truncation and $G_0 \cup T$ is fully compared in the first scan. Under the gain error bound ξ for expected verifier value F_B at cap $B = b - r$, if

$$F_B(G_0 \cup T) - F_B(G_0) > 3\xi,$$

then a normal direct budget limited return satisfies $F_B(\widehat{G}) \geq F_B(G_0 \cup T) - 2\xi$.

The proof follows.

Why the structural condition is sufficient We prove Proposition C.2. A nonempty edge set is weakly connected when its incident vertices are connected after edge directions are ignored. Let $\mathcal{K}(H^+)$ be the weak components containing at least one edge, with m_C directed edges in component C. Before deduplication, the emitted prefix count is exactly

$$\sum_{C \in \mathcal{K}(H^+)} m_C\big(\min\{w, m_C\} - 1\big) \leq M(w-1).$$

Start at any edge of a component. Any proper nonempty edge subset of that component has an unselected adjacent edge. Growth therefore cannot stop before reaching either width w or the complete component. Every root in C emits $\min\{w, m_C\} - 1$ prefixes. Deduplication and truncation cannot increase the count. This bounds candidate generation, not total computation or continuation cost.

For the general claim, use the hypothesis scores $s_e > 0$. For an edge $e = (i, j)$, write $V(e) = \{i, j\}$ and, for a nonempty edge set $T \subseteq H^+$, write $V(T) = \bigcup_{e \in T} V(e)$. Define its external edge boundary by

$$\partial_{H^+} T = \{e \in H^+ \setminus T \mid V(e) \cap V(T) \neq \varnothing\}.$$

This includes extra edges between two vertices of T. The separation premise is

$$s_e > s_f \qquad \text{for all } e \in T \text{ and } f \in \partial_{H^+} T.$$

Take any root in T. While the current prefix is a proper subset of T, weak connectivity supplies an unselected adjacent edge in T. Every eligible external edge belongs to $\partial_{H^+} T$ and has smaller score. Hence growth adds an internal edge. Induction shows that the first $|T|$ edges are exactly T, which is emitted since $2 \leq |T| \leq w$. An empty boundary makes the premise vacuous and recovers the whole component case. Strict separation avoids reliance on the tie order.

Now let the retained edge set be T and suppose its first scan comparison completes. Its true verifier gain above 3ξ yields an estimate above 2ξ. The selected completed estimate is at least as large. Applying the error bound at this candidate and the selected one gives

$$F_B(\widehat{G}) \geq F_B(G_0 \cup T) - 2\xi.$$

Subsequent accepted moves only increase verifier value on the same error event. This proves the claim for a normal direct return. If $|F_B - R_B| \leq \varepsilon$ at G_0, $G_0 \cup T$ and $\widehat{G}$, the package's raw task gain above $3\xi + 2\varepsilon$ suffices and

$$R_B(\widehat{G}) - R_b(G_0) \geq R_B(G_0 \cup T) - R_B(G_0) - 2\xi - 2\varepsilon - D_r^R.$$

Actual search resources remain separate unless a declared price puts them in score units. No installed permission charge is needed for this statement.

A simpler component control Enumerate each whole weak component once and keep those with two through w edges. Rank these packages by the same sum of edge scores and graph tie order, retaining at most P. This implemented control reuses the exact permitted packet matcher and inventory interface. It emits

$$K_w = \sum_{C \in \mathcal{K}(H^+)} \mathbf{1}\{2 \leq m_C \leq w\} \leq \lfloor M/2 \rfloor$$

packages before truncation. Consequently, if $P \geq K_w$, all eligible components survive retention. Rooted growth lacks this retention guarantee because its proper prefixes compete for the same places. The smaller inventory does not itself establish lower total execution cost.

For example, suppose the unknown required set S is one such component and $Y(G) = \mathbf{1}\{S \subseteq G\}$. Starting from empty with exact feedback, $0 < \xi < 1/2$ and no operational failure, the control succeeds when $P, Q \geq K_w$. Its first scan compares S, whose gain one exceeds the acceptance threshold. This needs at most K_w comparisons instead of the sufficient $M(w-1)$ ceiling for all rooted prefixes. Complete permissions also succeed in this uncharged model.

Threshold components isolate the additional hypothesis class Let $\mathcal{T} = \{s_e : e \in H^+\}$ be the distinct positive edge scores. For each $t \in \mathcal{T}$, form $H_t = \{e \in H^+ : s_e \geq t\}$ and enumerate its whole weak components with two through w edges. Deduplicate their directed edge sets, rank by the same score sum and graph tie order, and retain at most P groups. The implemented control uses the same permitted packets and no feedback during construction.

This inventory contains every set covered by Proposition C.2. Indeed, choose $t = \min_{e \in T} s_e$. All edges of T are present, while strict separation excludes every other edge incident to its vertices, including chords. Thus T is exactly a weak component of H_t. Conversely, a component at threshold t has every internal edge score at least t and every external incident score below t, so it satisfies the separation condition. These are precisely the separated connected groups within the width limit.

There are at most M distinct groups before retention. Process thresholds in descending order. A component appearing for the first time must contain a newly admitted edge; otherwise it is unchanged from the preceding threshold. Charge it to one such edge. Components at the same threshold are disjoint and each edge enters once, so the charges are distinct. Deduplication and the width limit only remove groups. This bounds candidate count, not runtime, since enumerating thresholds also incurs local computation.

Rooted growth can add groups outside this class. For scores $s_{a \to b} = 2$ and $s_{b \to c} = s_{a \to d} = 1$ at width two, the root $b \to c$ grows to $T = \{a \to b, b \to c\}$. Threshold two has one edge, while threshold one has an oversized three edge component, so neither emits a package. This identifies additional coverage, not a useful task outcome. Every threshold package occurs in the untruncated rooted inventory by Proposition C.2, but extra rooted candidates can displace it under the same P or Q. Therefore neither retained inventory nor performance is generally ordered. A prospective comparison must test whether the additional groups justify their construction and evaluation costs. Neither new control has been added to the frozen results in Section G.

Where coverage fails Missing matches can omit essential edges, while shared words can merge unrelated exchanges beyond width w. Greedy growth may miss useful subsets or disconnected groups. Strong proper prefixes can displace a useful component even when P equals the component count. Score ranking also need not place it early in the lexicographic query order.

Constructor comparisons hold packets, width, inventory cap, comparison order, feedback and continuation cap fixed. The package ablation removes only joint candidates. Complete policy comparisons instead allow saved resources to fund execution under the same total budget.

C.4 Public feedback and affordable search

A fixed verifier receives the public question, full document collection and committed answer with original sentence citations. It judges citation support, sufficiency for the complete question and contradiction by the context, giving

$$F = \mathbf{1}\{\text{supported and sufficient and not contradicted}\}.$$

Reference answers, annotated support, task categories, graph identity and private trajectories are excluded. Charge comparison judgments to search and the final judgment to deployment evaluation. Independent scoring tests this surrogate's relationship to correctness.

Parse an exact Boolean verdict. Missing or malformed answers score zero without a call; citations remain unrepaired. Invalid verifier output halts search without a sample, rather than supplying a zero. Fix model, decoding, prompt, failure rules and input limits before use; truncation changes the objective.

When proxy gains transfer Use the verifier and final score values $F_B(G)$ and $R_B(G)$ at the same retained checkpoint and cap. Both are raw scores in $[0, 1]$, with costs reported separately. Apply Proposition C.1 to $J_B = F_B$. If, for a constant $\varepsilon \geq 0$,

$$\sup_{G \in \mathcal{S}} |F_B(G) - R_B(G)| \leq \varepsilon,$$

then on its simultaneous estimation event

$$R_B(\widehat{G}) - R_B(G_0) \geq N_{\text{acc}}\xi - 2\varepsilon.$$

Subtract at most ε at each endpoint of the telescoped verifier gain. A complete rejection scan similarly bounds neighboring final score gains by $3\xi + 2\varepsilon$. Average development agreement does not establish the uniform premise for adaptively selected graphs. This comparison holds only at cap B and neither repays search nor establishes advantage over a full budget control.

Development checks verifier judgments against independent final scores and fresh deployment.

The reservation is a feasibility constraint Binary paired gains lie in $[-1, 1]$. Substituting range width $u - \ell = 2$ in the sample bound gives 1235 samples for $Q = 12$, $\xi = 0.1$ and $\delta = 0.05$, hence 29640 search continuations and one deployment, each with at most one verifier judgment. The bound gives a sufficient reservation for the stated confidence guarantee, and each continuation may contain several calls. Smaller sample sizes remain available for empirical policy evaluation without invoking this confidence guarantee.

Write $c_B, c_V, c_P, c_O \geq 0$ for the costs of a worst case continuation, one verifier judgment, the shared pilot and all proposal and bookkeeping work, respectively, in one common resource unit. One searched policy must reserve at least

$$c_P + c_O + (2Qn + 1)(c_B + c_V).$$

The final verifier scores selected deployment; controls need separate allowances. Apply the formula to each compatible resource counter and report wall time separately. Charge the physical pilot once. Budget limited selection can use a partial first scan, whereas a statement about every neighbor requires a complete rejection scan.

C.5 What a worker can observe

Workers retain separate histories containing permitted context, fixed roles, tools, their own actions and results, and admitted messages. Each receives the complete address book. The graph silently filters delivery without permission errors or acknowledgements. Messages enter later rounds; tool results enter subsequent requests. Removing permission cancels pending delivery but never erases received evidence.

The organizer receives only permitted observations. Proposal rationales stay outside worker inputs, and installation changes no roles, capabilities, authority or nongraph prompts. Shared writable state could bypass this boundary, so branches require declared stateless read only tools. Mutable tools would require isolated state and separate validation.

The input wrapper admits only worker cited receipts and checks ownership, not truth or hidden answer leakage. Authentic packet generation and durable provider integration remain unfinished.

C.6 What a graph intervention identifies

The interface in `experiments/revision_interventions.py` forks a resumable pilot with fixed workers and retained histories. Apply installs the proposal, while sham pays the same proposal charge and retains the original graph. Frozen continuation skips that charge and may reuse saved resources under the common total cap. Define $V_{\mathrm{cont}}(h, G, b)$ as expected continuation value at retained state h, installed graph G and remaining budget b, under a declared score or priced welfare objective. Let H be a pilot history whose declared law is supported on resumable states, $G_0(H)$ its incumbent graph, $G_\star$ the selected graph and B' the common postproposal budget. With expectation over the pilot and selection law, apply versus sham targets

$$\Delta_{\mathrm{implement}} = \mathbb{E}[V_{\mathrm{cont}}(H, G_\star, B') - V_{\mathrm{cont}}(H, G_0(H), B')] .$$

Apply versus frozen instead measures whether revision is worth its total cost. The signs need not agree. Rejected and unchanged proposals remain in both accounting and analysis.

Proposition C.3 (Graph effects require changed delivery)**.** *Consider apply and sham with identical task, local histories, pending messages, tool state, incumbent and resource ledger. Fix their actors, roles, authority, dispatch order, horizon, address book and primitive response kernels. Suppose installed edges affect execution only through private message admission; every other transition and charge depends on the preceding nongraph state. Couple primitive randomness through common exogenous tapes and let D denote a delivery difference before termination. Until that difference,*

requests, actions, receipts, reservations and stopping decisions agree. On D^c, final decisions, ledgers and status agree. For scores $Y_{\text{apply}}, Y_{\text{sham}} \in [0,1]$ under the common decision sufficient scoring kernel,

$$|\mathbb{E}[Y_{\text{apply}} - Y_{\text{sham}}]| \leq \Pr(D),$$

The proof follows.

Proof. For Proposition C.3, induct over dispatch and invocation phases, with stopped states absorbing. Equal nongraph states produce equal requests and reservations. Shared primitive tapes give equal responses, tool results and usage. Equal attempted messages yield equal local history updates. Only private filtering can first separate the states. Equal admissions preserve the induction, including the global budget. For a committed decision a, let $\mu(a)$ be its conditional mean score under the common scoring kernel. On D^c the decisions and their μ values agree. Since $0 \leq \mu \leq 1$, taking expectations gives the bound. □

Shared provider seeds do not guarantee this coupling. Independent calls require replications, and changed delivery remains necessary rather than sufficient for a useful effect.

C.7 Admission, failures and resource use

Register task identity and public input before dispatch. Retain failed pilots, unchanged graphs, rejected proposals and incurred charges in the admitted population. Score the latest committed submission under Appendix A; a later operational failure does not erase it. If search fails before deployment, retain only the pilot's commitment, never a trial answer. Missing or malformed commitments score zero, while undispatched controls and pending scores remain incomplete observations.

Before every dispatch, reserve worst case usage. Settle known usage, but retain a conservative charge and halt dispatch when usage is uncertain. Charge packet generation, proposals, comparisons, attempted messages, tools, retries and aggregation. The shared pilot enters each policy's comparison ledger once and the actual study invoice once. Interrupted work is not silently replayed.

C.8 Proposed document reasoning evaluation

The planned HotpotQA study (Yang et al., 2018) tests evidence combination and complete search cost. Excluded development will assess feedback and affordability before fixing policies, models, prompts, membership, prices, caps and analysis. Methods share public documents, tools and verifier access; hidden answers and support remain with the independent evaluator. The HotpotQA study has not run.

Compare joint answer and evidence F1 against selective fixed communication, a frozen graph, a capable singleton, a manager using the same workers and an executable adaptive method. Joint versus individual edits, paid sham and matched rewiring distinguish package and connectivity effects. Report quality against actual inference cost.

Complete policy evaluation Fix the complete policies $\mathcal{F}, \mathcal{F}_0$, task sampling rule, replication counts and total resource caps using development data only. Direct deployment is the proposed mode. Let $N_{\text{task}} \geq 1$ be the integer number of admitted tasks and $R_i \geq 1$ the integer number of prespecified replications for task $i \in \{1, \ldots, N_{\text{task}}\}$ under each policy. Write $Y_{ir}^{\mathcal{F}}$ for the independently evaluated final task score of policy $\mathcal{F}$ on task i in replication r, and similarly for $\mathcal{F}_0$. Define the task differences and their mean by

$$D_i = \frac{1}{R_i} \sum_{r=1}^{R_i} \left(Y_{ir}^{\mathcal{F}} - Y_{ir}^{\mathcal{F}_0} \right), \qquad \widehat{\Delta} = \frac{1}{N_{\text{task}}} \sum_{i=1}^{N_{\text{task}}} D_i. \tag{19}$$

Pairing aligns task identity and replication index, not provider randomness. Independent tasks sampled from $\mathcal{P}$, independent execution groups across tasks and the stipulated policy laws make this estimator unbiased for the population contrast $\Delta_{\mathcal{P}}(\mathcal{F}, \mathcal{F}_0)$ in (1). Replications stay nested within tasks for uncertainty. A fixed benchmark instead targets its own mean expected difference unless its sampling design supports broader inference. No per task significance test or calibrated online verifier is required to estimate this final score contrast.

The selective control chooses recipients from local evidence using the same workers, message policy and public address book, without a supplied relevance rule. Controls may spend saved search resources on reasoning within the common total cap. Include every admitted task under the committed answer rule and retain incomplete pairs as incomplete. Task selection and replacement cannot depend on observed performance.

Equally paid sham retains the original graph after identical search charges at a resumable checkpoint. Fresh deployments isolate connectivity, while complete policy comparisons also include pilot failures and the opportunity cost of search. Final evaluator scores never return to the policy.

A feasible population for matched rewiring Use a declared graph class $\mathcal{S}$ on the fixed actor set V and admissible edge universe. For $G \in \mathcal{S}$, let $\mathrm{Reach}(G)$ contain the ordered pairs $(i, j) \in V^2$ with $i \neq j$ for which a directed path in G leads from i to j. Define

$$\mathcal{M}(G) = \{G' \in \mathcal{S} \setminus \{G\} \mid |G'| = |G|,\ \mathrm{Reach}(G') = \mathrm{Reach}(G)\}.$$

Before continuation, enumerate $\mathcal{M}(G)$ on the fixed actor labels, record eligibility $I(G) = \mathbf{1}\{|\mathcal{M}(G)| > 0\}$ and draw uniformly with a recorded seed. Transfer no messages, prompts or private histories. The matching set may be empty, as for an inward reporting star whose edge count and reachability determine every edge.

Write $Y(G)$ for a fresh continuation's final score from the shared history and cap. The contrast $\mathbb{E}[Y(G) - Y(\widetilde{G}) \mid I(G) = 1]$ averages over pilots, selected graphs, the uniform rewire $\widetilde{G}$ and fresh outcomes. Report eligibility and excluded graphs alongside it. Ineligible tasks remain in complete policy comparisons; with no eligible graph the rewiring effect is unavailable. Matching is fixed before outcomes and preserves reachability, not path lengths or delivery times.

After sealing tasks, compare verifier acceptance with independent scores of the same commitments. Keep these verdicts hidden from the policy and retain every search charge.

D Proofs for task dependent organization

The first result concerns unchanged submitted decisions. The remaining results characterize the observations needed to distinguish useful organizations.

D.1 Execution equivalence at the scored interface

A changed graph need not improve task performance. If submitted decisions stay unchanged, any net gain must come from cost or risk.

Proposition D.1 (Execution equivalence)**.** *Fix a task ψ and two feasible organizations o and o', using the submitted decision A, score Y and execution equivalence and the static welfare quantities in Equation (9). For each $\bar{o} \in \{o, o'\}$, let $Q_{\bar{o}}$ be the law of A. Suppose the evaluator is decision sufficient, meaning that a common scoring kernel $K_\psi(\mathrm{d}y \mid a)$ factors the joint law of (A, Y) as $Q_{\bar{o}}(\mathrm{d}a)K_\psi(\mathrm{d}y \mid a)$ under either organization. If o and o' are execution equivalent, then $\mathsf{V}_\psi(o) = \mathsf{V}_\psi(o')$ and*

$$\mathsf{J}_\psi(o) - \mathsf{J}_\psi(o') = -\big[\mathsf{C}(o) - \mathsf{C}(o')\big] - \lambda_f\big[\mathsf{Risk}_\psi(o) - \mathsf{Risk}_\psi(o')\big].$$

If their declared risks are equal, welfare differs only through resource cost. Under a paired deterministic scorer, pathwise equality of submitted decisions also gives pathwise equality of raw outcomes. The proof follows.

Proof of Proposition D.1. Let Q be the common law of the submitted decision. Decision sufficiency gives

$$\mathsf{V}_\psi(o) = \int\!\!\int y K_\psi(\mathrm{d}y \mid a) Q(\mathrm{d}a) = \mathsf{V}_\psi(o').$$

Subtracting Equation (9) for the two organizations yields the displayed welfare identity. Equal declared risks cancel its last term. In the paired deterministic case, the common scorer maps equal submitted decisions on the same task to equal raw outcomes for every pair. These post execution relations supply neither submitted decisions nor outcomes to the formation policy. □

D.2 Organization identifiability from permitted diagnostics

Take $\mathcal{S} \neq \varnothing$ and $\varepsilon \geq 0$ throughout this subsection.

Here retained state means retained selector state, not worker private histories. The first result uses fixed welfare; its conditional application below permits learning during pilots.

Let Θ index task and agent regimes with welfare $J_\theta \colon \mathcal{S} \to \mathbb{R}$ on finite $\mathcal{S}$, the feasible organization set. An admissible protocol π may randomize pilots with regime independent coins; deterministic protocols use point masses. Let H^π contain every decision input, including text, pilot actions, traces, outcomes, costs, and retained state; let P_θ^π be its law. Define $\mathrm{TV}(P, Q) = \sup_D |P(D) - Q(D)|$, with the supremum over measurable events D. A permitted selector is a common probability kernel from H^π to an output, with any fresh randomness independent of the regime. Its output is either a binary label $\widehat{c} \in \{0, 1\}$ or an organization $\widehat{s} \in \mathcal{S}$. The ε optimal set is

$$\mathcal{A}_\theta(\varepsilon) = \{s \in \mathcal{S} : J_\theta^\star - J_\theta(s) \leq \varepsilon\}, \qquad J_\theta^\star = \max_{s \in \mathcal{S}} J_\theta(s). \tag{20}$$

Theorem D.2 (Quantitative observational barrier)**.** *Under the diagnostic model and good sets in Equation (20), fix $\tau \in [0, 1]$ and regimes θ_0, θ_1 such that $\sup_\pi \mathrm{TV}(P_{\theta_0}^\pi, P_{\theta_1}^\pi) \leq \tau$ over admissible diagnostic protocols. Every permitted binary selector satisfies $e_{\mathrm{cls}} = \frac{1}{2}\sum_{j=0}^{1} \Pr_{\theta_j}\{\widehat{c} \neq j\} \geq (1-\tau)/2$. If $\mathcal{A}_{\theta_0}(\varepsilon) \cap \mathcal{A}_{\theta_1}(\varepsilon) = \varnothing$, every permitted formation rule satisfies $e_{\mathrm{org}} = \frac{1}{2}\sum_{j=0}^{1} \Pr_{\theta_j}\{\widehat{s} \notin \mathcal{A}_{\theta_j}(\varepsilon)\} \geq (1-\tau)/2$. The proof uses contraction of total variation under a common selector.*

When two regimes require disjoint good organizations, reliable formation requires observable evidence that separates them. Task text, pilot interactions, and outcomes can provide that evidence.

Proof of Theorem D.2. Fix any admissible evidence protocol π and a permitted decision rule, either a binary selector or a formation rule. The complete history H^π contains all retained state and coins used by the decision; fresh decision randomness is regime independent. Thus the decision is a common Markov kernel from H^π to its output. Let Q_j be the output law under regime θ_j. Data processing at the actual admissible protocol gives

$$\mathrm{TV}(Q_0, Q_1) \leq \mathrm{TV}\big(P_{\theta_0}^\pi, P_{\theta_1}^\pi\big) \leq \tau.$$

This argument does not require a protocol obtained by fixing pilot coins to remain admissible. For a binary selector, let $C_j = \{j\}$. Then

$$Q_0(C_0) + Q_1(C_1) \leq Q_0(C_0) + Q_0(C_1) + \tau = 1 + \tau.$$

Thus balanced success is at most $(1 + \tau)/2$, so $e_{\mathrm{cls}} \geq (1 - \tau)/2$. For the organization conclusion, write $A_j = \mathcal{A}_{\theta_j}(\varepsilon)$. Because $A_0 \cap A_1 = \varnothing$,

$$\begin{aligned} Q_0(A_0) + Q_1(A_1) &\leq Q_0(A_0) + Q_0(A_1) + \mathrm{TV}(Q_0, Q_1) \\ &\leq 1 + \tau. \end{aligned}$$

Let $e_j = \Pr_{\theta_j}\{\widehat{s} \notin A_j\}$. The left side is $2 - e_0 - e_1$, so $e_{\mathrm{org}} = (e_0 + e_1)/2 \geq (1 - \tau)/2$. The rule was arbitrary, so the result covers adaptive pilot organizations, traces, and outcome feedback. □

Separation under a fixed protocol For a fixed evidence protocol completed before selection, balanced classification accuracy $a_\pi = 1 - e_{\mathrm{cls}}$ satisfies $a_\pi \leq (1 + \tau_\pi)/2$, equivalently $\tau_\pi \geq 2a_\pi - 1$, where $\tau_\pi = \mathrm{TV}(P_{\theta_0}^\pi, P_{\theta_1}^\pi)$ uses the history laws and total variation above.

Assume the diagnostic protocol reaches a resumable checkpoint almost surely under both regimes. Fix a common finite nonempty feasible graph class $\mathcal{S}$. Define finite measurable conditional expected continuation welfare $J_\theta^\pi(h, s)$ by installing s at the pilot's resumable history h, preserving worker states and the same remaining budget. Average hidden worker states conditional on h, without disclosing them. Fix continuation, aggregation, stopping, charging and nongraph policies. Define

$$\mathcal{A}_\theta^\pi(h, \varepsilon) = \left\{ s \in \mathcal{S} : \max_{t \in \mathcal{S}} J_\theta^\pi(h, t) - J_\theta^\pi(h, s) \leq \varepsilon \right\}. \tag{21}$$

Choose measurable versions on a common history space. These values and sets are not supplied to the selector. If the good sets are disjoint for $(P_{\theta_0}^\pi + P_{\theta_1}^\pi)$ almost every h, every common selector

kernel using only h and fresh independent coins has balanced probability at least $(1 - \tau_\pi)/2$ of selecting outside its regime's conditional good set. To prove this, write the joint law of $(H^\pi, \widehat{s})$ as $R_j(\mathrm{d}h, \mathrm{d}s) = P^\pi_{\theta_j}(\mathrm{d}h)K(\mathrm{d}s \mid h)$. Data processing gives $\mathrm{TV}(R_0, R_1) \le \tau_\pi$. The success events $C_j = \{(h, s) : s \in \mathcal{A}^\pi_{\theta_j}(h, \varepsilon)\}$ are disjoint almost surely under both laws. Thus $R_0(C_0) + R_1(C_1) \le R_0(C_0) + R_0(C_1) + \tau_\pi \le 1 + \tau_\pi$. Subtracting balanced success from one proves the claim. History must remain in the joint law because good sets vary with it. If a pilot can terminate without a resumable checkpoint, this statement does not apply to its unrestricted history law. One may instead condition both regime laws on resumability, provided each event has positive probability, and use their conditional total variation throughout. That conditional bound says nothing about pilot failure rates. This necessity result does not bound cumulative regret or guarantee learned utility. Complete policy comparisons must charge pilot and proposal work and retain failed pilots (Appendix C).

E Exploration under complementary evidence

This section gives a controlled requirement model in which locally improving communication cannot initiate useful coordination. Its binary evaluator is a membership query for a monotone conjunction, a single term in the monotone DNF setting studied by Abasi et al. (2014). Complementary set objectives and greedy search are also studied by Bai & Bilmes (2018). We use these established viewpoints to diagnose communication formation.

E.1 A hidden set of jointly required evidence

Fix the permitted initial context $X = x$, integers $L \ge 1$ and $1 \le k \le L$, and an admissible channel set $\mathcal{E}_x$ with $|\mathcal{E}_x| = L$. Each channel carries a distinct evidence item to a fixed decision authority in one round. Agents, roles, item contents, and delivery behavior are fixed. A private validator requires a set $S \subseteq \mathcal{E}_x$ of k evidence items jointly before the authority can commit a prescribed decision $a^\star$. Otherwise the authority abstains with decision $\perp$. We identify a graph with its channel set, following the graph convention. For a queried organization $G \subseteq \mathcal{E}_x$,

$$A_S(G) = \begin{cases} a^\star, & S \subseteq G, \\ \perp, & S \not\subseteq G, \end{cases} \qquad Y_S(G) = \mathbf{1}\{S \subseteq G\}.$$

The validator retains S. The formation policy receives the public dimensions L and k, but neither S nor a task graph. Conditional on $X = x$, assume that S is uniform over the $\binom{L}{k}$ subsets of size k. Thus the prior is specified after all allowed initial task text and registry information, rather than before potentially informative text is observed.

After query G_t, the policy observes $Y_S(G_t)$ and a permitted trace Z_t. Given x, the preceding history h_{t-1}, G_t, and $Y_S(G_t)$, the trace has a common kernel

$$K_t(\mathrm{d}z \mid x, h_{t-1}, G_t, Y_S(G_t))$$

that does not otherwise depend on S. Take trace spaces to be finite or countable. Policy randomization is independent of S conditional on x. Queries may be adaptive, and their traces may include graph dependent resource use, but a failed query reveals no semantic clue about which missing item is required beyond its failure. This last restriction is essential.

Let deployment welfare be $J_S(G) = Y_S(G) - c|G|$, where $c \in [0, 1/k)$ is the deployment price per channel. If an integer number $T \ge 0$ of queries incurs total declared nonnegative cost C_T before final deployment $\widehat{G}$, the net value is

$$\mathcal{V} = Y_S(\widehat{G}) - c|\widehat{G}| - C_T.$$

Query cost includes the charged resources used to obtain feedback. For example, $C_T = \kappa T + \zeta \sum_{t=1}^{T} |G_t|$ charges each query and every channel used in it, with $\kappa, \zeta \ge 0$. Assume C_T is integrable. These prices define the controlled model. In particular, $c = 0$ gives free deployment permissions, a case in which the strict improvement barrier below still holds.

Proposition E.1 (Complementary evidence exploration barrier)**.** *Under the hidden evidence model and cost definitions, the following conclusions hold.*

1. *If $k \geq 2$, a formation procedure initialized at $G = \varnothing$ that tests only single channel additions and accepts only strictly positive marginals of J_S remains empty, regardless of its query budget. Yet $J_S(S) = 1 - ck > 0$.*
2. *Fix integers $N \geq 0$ and $1 \leq k \leq b_{\mathrm{qry}} \leq L$, where b_{qry} is the channel width limit. Every permitted randomized adaptive procedure that uses at most N feedback queries, with $|G_t| \leq b_{\mathrm{qry}}$ and $|\widehat{G}| \leq b_{\mathrm{qry}}$, satisfies*

$$\Pr\{Y_S(\widehat{G}) = 1 \mid X = x\} \leq p_{N,b_{\mathrm{qry}}} := \min\left\{1, (N+1)\frac{\binom{b_{\mathrm{qry}}}{k}}{\binom{L}{k}}\right\}.$$

   Consequently,

$$\mathbb{E}[\mathcal{V} \mid X = x] \leq (1 - ck)p_{N,b_{\mathrm{qry}}} - \mathbb{E}[C_T \mid X = x].$$

3. *For $b_{\mathrm{qry}} = k$, the success bound is attained by querying distinct k item packages, returning the first successful package, and otherwise deploying an unqueried package while one remains.*

If the deployment width instead satisfies $0 \leq b_{\mathrm{qry}} < k$, success is identically zero. The proof follows.

Proof. We prove Proposition E.1. If $k \geq 2$, the empty organization and every singleton channel organization return $\perp$. Each candidate birth therefore has deployment marginal $-c \leq 0$ at the empty state. The marginal is zero when $c = 0$ and negative otherwise. No birth is accepted by a strict improvement rule, so the same state persists after every rejected test. The hidden package S instead enables the committed decision and has value $1 - ck > 0$.

For the exploration bound, condition on x. Presample the policy coins and the coins used by the common trace kernels into a tape U independent of S. For each fixed $U = u$, follow the counterfactual history in which every feedback query returns zero. This determines at most N query sets $G_1^0(u), \ldots, G_N^0(u)$ and a final set $G_{\mathrm{out}}^0(u)$, independently of S. Pad unused queries with empty sets. Trace kernels can be extended arbitrarily on histories of probability zero.

If a real run first receives a positive response at query t, its preceding responses were all zero. Its history and query therefore agree with this counterfactual branch through time t, so $S \subseteq G_t^0(u)$. If no query is positive and final deployment succeeds, then $S \subseteq G_{\mathrm{out}}^0(u)$. Thus every successful requirement set lies in

$$\bigcup_{t=1}^{N} \{S \subseteq G_t^0(u) : |S| = k\} \cup \{S \subseteq G_{\mathrm{out}}^0(u) : |S| = k\}.$$

Every queried or final set has at most b_{qry} channels and therefore contains at most $\binom{b_{\mathrm{qry}}}{k}$ possible requirement sets. The union has cardinality at most $(N+1)\binom{b_{\mathrm{qry}}}{k}$. Dividing by $\binom{L}{k}$, capping at one, and averaging over U proves the probability bound. This argument permits both adaptive proposals and randomized traces.

On successful deployment, $|\widehat{G}| \geq k$. On failure, channel cost is nonnegative. Pointwise,

$$Y_S(\widehat{G}) - c|\widehat{G}| \leq (1 - ck)Y_S(\widehat{G}).$$

Taking expectations, using $1 - ck > 0$, and subtracting the expected total query cost gives the net value bound. No independence between C_T and the final outcome is required.

For $b_{\mathrm{qry}} = k$, one query tests exactly one possible requirement set. Distinct packages and an unqueried final package cover $\min\{N+1, \binom{L}{k}\}$ possibilities. Returning a successful queried package attains the bound under the conditional uniform prior. If $b_{\mathrm{qry}} < k$, the final organization cannot contain S. □

Scope The model fixes workers and roles and places a uniform prior on the required evidence set conditional on permitted context. Its traces provide no additional semantic clues beyond the binary

outcome. If failed work identifies missing evidence, targeted exploration can be easier. At zero channel price, strict ascent still stalls, whereas accepting neutral moves or delivering all evidence can escape. The result therefore motivates joint exploration within this information regime. Wider temporary communication can expose a useful package, with its queries and delivered messages charged to the resource budget. The planning model instead supplies a reservation relevance rule, so selective communication provides the relevant empirical comparison there.

F Learning efficient patterns from sampled demand

The controlled study fixes its sampling, costs and uncertainty analysis before execution. Table 2 separates the questions answered by its controls. The modular ablation and sharper margin calculation subsequently reuse retained records and are explicitly post hoc.

F.1 Demand, candidates and prospective comparisons

For fixed workers $V = \{\mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{d}, \mathtt{e}, \mathtt{f}\}$, the prespecified family contains direct communication, six stars and ninety modules. Modules use two triads and one gateway per triad, giving $\binom{6}{3}/2 = 10$ partitions and nine gateway pairs each.

For each seed and concentration parameter $\kappa \in \{0.05, 0.5, 5\}$, independently draw two laws on the thirty directed worker pairs,

$$p^0, q \sim \text{Dirichlet}(\kappa \mathbf{1}_{30}), \qquad p^1 = (1-\omega)p^0 + \omega q, \qquad \omega \in \{0, 0.25, 0.5, 1\}.$$

Here $\mathbf{1}_{30}$ contains thirty ones. Observations follow p^0, future packets follow p^1 and mixtures share the same sampled laws. Dispersion κ is distinct from hub concentration γ_v in (12); graph types do not generate demand.

Observation counts are $m_{\text{obs}} \in \{32, 128, 512\}$. Each observation is an independent pair from p^0, and the future horizon contains $W = 256$ independently sampled pairs from p^1. Observation and future streams are separate. Within a seed and κ, the observation windows are nested prefixes. Future packets are shared across methods, prices and observation counts at the same mixture weight. Prices are $a \in \{0, 1, 4, 16\}$ per enabled session and $c = 1$ per packet hop. All prices are specified model units.

Development seeds 1000 through 1023 select one fixed graph per price by mean future cost across development profiles and mixtures, with canonical tie breaking. Independent evaluation seeds 2000 through 2023 yield $24 \times 3 \times 3 \times 4 \times 4 = 3456$ cells. Related cells remain in their seed cluster.

Empirical selection minimizes (2); ties prefer the incumbent, direct delivery, then canonical edge order. Shuffling uses a seeded nonidentity permutation of observed worker identities, with future endpoints unchanged. The confidence rule uses Proposition B.2, $\delta = 0.05$ and true simulated drift $\nu = \text{TV}(p^0, p^1)$. It changes only when predicted savings exceed 2η plus all learning and switching charges. All 3456 predictions, graphs and bounds are sealed before future sampling; unchanged choices remain included.

F.2 Costs, endpoints and uncertainty

For future pairs $Z_1^+, \ldots, Z_W^+$, define realized transport cost

$$\ell(G) = a|E(G)| + \sum_{t=1}^{W} d_G(Z_t^+).$$

Direct and fixed policies pay $\ell(G)$. Profile policies also pay $0.02 m_{\text{obs}} + 1$ for collection and selection, plus two units when changing G_0. Observation costs remain after rejection. Costs use shortest paths and full delivery under Assumption B.1, without simulated queues or inference.

The primary cost difference subtracts fixed graph cost from fully charged policy cost; negative values favor selection. Main figures use the opposite sign, saving. Expected regret $\mathsf{L}_{p^1}(G) - \min_{H \in \mathfrak{G}} \mathsf{L}_{p^1}(H)$ uses the future law, while hindsight regret $\ell(G) - \min_{H \in \mathfrak{G}} \ell(H)$ uses realized packets. These oracle diagnostics exclude learning charges and are not deployable controls.

Table 2: **What each comparison establishes.** All controls share empirical selection's future packets. The first four are prospective; the restricted family comparison is post hoc.

Control	Controlled difference and scientific question
Development fixed	Fixed on separate seeds. Does demand fitting repay acquisition and selection expense?
Direct selective	All requested pairs communicate directly. Do fewer channels repay forwarding and learning?
Permuted profile	Pair masses and collection expense stay fixed. Does matching connections to workers matter?
Confidence rule	Same candidates and observations, with a sufficient margin and true drift. Which changes are permitted?
Direct or best hub	Modules removed, with the same observations and stipulated charges. Does grouping improve on direct delivery and hubs?

At each price, observation count and mixture weight, average three κ settings within each seed. Figures show the mean and sample SD of those twenty four averages, using denominator 23. Descriptive 95% intervals use their standard error times $t_{23,0.975}$, separately by setting without simultaneous coverage. Independent reconstruction checks choices, costs, regret and margins from recorded packets using Floyd Warshall rather than the selector's breadth first search. Every planned cell completed; invalid graphs or missing costs would halt execution without replacing seeds.

F.3 Isolating the contribution of modular patterns

For all 3456 cells, minimize estimated cost over direct communication and six hubs with the original incumbent and tie rule. Reconstruct costs on the original future packets. Both selectors pay $0.02m_{\mathrm{obs}}+1$ and two units when changing the incumbent. Equal stipulated selection charges isolate candidate access, not the computational saving from searching seven rather than 97 graphs. No new demand is generated.

At $a = 4$, 128 observations and no drift, the full selector saves 8.08 units relative to this restricted selector, with paired sample SD 8.78 and descriptive 95% interval $[4.38, 11.79]$. For $\kappa = 0.05$, the saving is 26.79 with interval $[16.85, 36.73]$. For $\kappa = 0.5$, it is -2.54 with interval $[-5.46, 0.38]$; $\kappa = 5$ is exactly tied. Twenty of the 72 profiles change graph identity. Sixteen improve cost, three worsen it and one has equal cost despite a different graph. The other 52 retain the same choice.

The effect is conditional on both price and persistence. At $a = 16$ without drift, the additional modular saving is -0.15 with interval $[-2.31, 2.01]$, so the large gain against the development fixed graph does not establish value from M. At $a = 4$, demand replacement weights 0.25, 0.5 and 1 give additional savings of 2.36, -3.33 and -16.90 units; each interval includes zero. Figure 3 displays these unresolved contrasts alongside the stable gain. Complete numerical grids remain in the research archive.

Measured computation cost A fixed local protocol times request counting and selection on every cell, with three repeats in rotated method order. Cached references return only the winning graph and cost; all choices and rational costs agree. On ARM64 macOS with CPython 3.12.14, factorization saves 0.10646 ms against the full scan, with paired interval $[0.10523, 0.10769]$, but adds 0.09626 ms over the restricted scan, with interval $[0.09555, 0.09697]$. SD and intervals use 24 seed averages over cells and repeats on this host, not variation across machines.

Cold setup averages 2.891, 7.337 and 0.508 ms for factorization, full and restricted scans over 24 repeats. Reference setup includes independent distance tables, making the difference implementation specific. Packet outcomes and inference are not rerun. An additional constant selection expense ΔK in transport units reduces the mean modular advantage to $8.08 - \Delta K$. Converting measured CPU to that expense requires calibrated session and hop prices.

F.4 When a sufficient rule can afford to change

At $a = 4$ and zero drift, the original sufficient threshold in Section 4.2 is

$$T(m) = 1024\sqrt{\frac{\log(3880)}{2m}} + 0.02m + 3, \qquad m > 0.$$

Writing $A = 1024\sqrt{\log(3880)/2}$, differentiation gives the minimum at $m = (A/0.04)^{2/3} \approx 1393.82$, where $T(m) \approx 86.63$. This exceeds the largest possible prelearning savings of $20a = 80$ for a hub and $16a = 64$ for a module.

Proposition B.3 instead bounds only additional forwarding. For a module with zero observed detour at 128 observations, $\delta_G = 0.05/96$ gives $U_G = 1 - \exp\{-\log(1920)/128\} = 0.05735$. The lower saving is $64 - 512U_G - 5.56 = 29.08$ after learning and switching expenses. Small observed detour therefore permits an affordable change.

For a retrospective check, apply this bound only to the original empirical choice when its original incumbent is direct. This retains all 2592 eligible cells at $a \in \{0, 1, 4\}$ and excludes the 864 cells at $a = 16$ with a nondirect incumbent. The original observations, windows and privileged drift values are unchanged. Every rejection pays observation and selection expenses. A change pays its switching expense and uses the recorded packet costs for its candidate. This is a new calculation, not a replacement for the frozen confidence arm.

At $a = 4$ without drift, the sharper rule accepts 3, 18 and 21 of 72 profiles at observation counts 32, 128 and 512. Their complete policy savings are 1.47, 11.66 and 5.82 units. At 128 observations the accepted graphs comprise seven hubs and eleven modules. Every accepted candidate has positive expected net gain under the constructed law, and every eligible cell satisfies the simultaneous candidate bound. This checks the bound at each fixed window, not simultaneous coverage across the full experiment. At replacement weight 0.25, one graph is accepted at each of 128 and 512 observations but the complete policy still loses after paying for declines. At weights 0.5 and 1, none is accepted.

G When connected proposals expose useful exchanges

This appendix characterizes the reference constructor and its controlled comparisons.

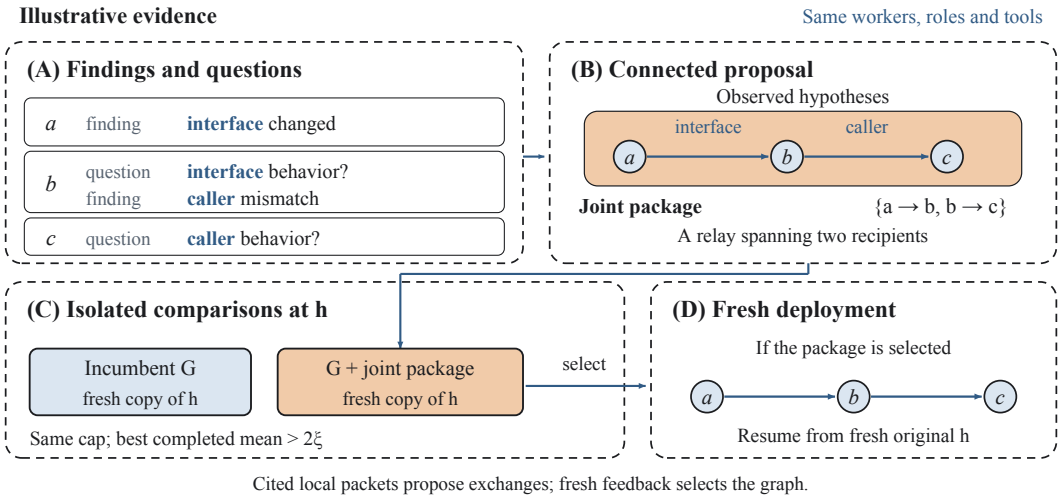

Figure 4: **Evidence suggests exchanges that must be evaluated together.** This schematic illustrates AutoForm's lexical constructor. Cited findings and open questions suggest a relay, whose channels are compared jointly from copies of checkpoint h. The profile adapter instead proposes complete patterns using measured request frequencies. Both use feedback search and fresh deployment, with workers and roles fixed.

A controlled complementary objective Four workers a, b, c, d have twelve admissible directed channels and start without peer permissions. In the required evidence notation, a pair S has one of three shapes, a relay $\{a \to b, b \to c\}$, a common receiver $\{a \to b, c \to b\}$ or disjoint exchanges $\{a \to b, c \to d\}$. The model score is

$$Y_S(G) = \mathbf{1}\{S \subseteq G\}.$$

Extra edges are free. Only the generator and evaluator know the required set. The proposer receives invented packets whose lexical matches encode observed hypotheses.

Evidence quality and comparison limits We cross zero or one missing required match with zero, two, four or eight false matches. Required matches have strength one and false matches strength two, deliberately creating a misleading ranking. Twenty four shared seeds choose actor permutations, the missing edge and nested random subsets of false edges. These 576 constructed cases each use package widths $w \in \{2, 3, 4\}$ and comparison ceilings $Q \in \{1, 2, 4\}$, rather than sampling independent benchmark tasks or natural language errors.

Each constructor uses package limit $P = 4$. The core accepts exact gains above 2ξ, with $\xi = 0.1$ and $n = 1$ paired draw (search notation). Packages precede individual edits, each group ordered lexicographically. Query ceilings are below the initial twelve individual candidates, so every search ends in its first partial scan.

Competitive constructors and fixed controls Controls are common recipient grouping, individual edits and uniform sampling of up to four observed edge subsets of sizes two through w. Uniform sampling permits disconnected groups and uses an unpriced catalogue. Complete permissions succeed without search. Neither permissions nor delivery are charged.

The comparisons below use the prospectively specified width three setting with complete required support. For searched methods, recovery means comparing a graph containing both required edges.

Coverage and comparison order Under clean relay evidence, connected growth recovers all twenty four cases with four comparisons at every tested width, whereas common recipient grouping and individual edits recover none. For a common receiver at width three and four comparisons, recipient grouping recovers all twenty four cases at every false hypothesis level, whereas connected growth declines sharply. Thus the broader constructor does not dominate the targeted one.

Among the 222 covering inventories out of 864 complete support case and width combinations, none loses coverage through truncation. Recovery at one, two and four comparisons is 177, 211 and 222. The remaining 642 inventories fail at construction.

Disconnected requirements and missing evidence At width three, connected growth never recovers disjoint pairs, whereas uniform packages recover every clean case. At width four with two false matches and four comparisons, connected growth recovers twenty two of twenty four pairs. False edges bridge the required exchanges in every recovered case. Their placement can therefore help coverage when width permits and extras are free.

Removing either required hypothesis makes every searched method fail throughout the grid. A positive support constructor cannot propose the unobserved edge, and strict ascent cannot accumulate individually unhelpful additions. Complete permissions still succeed.

Interpretation and reproducibility Two complete executions reproduce every inventory, comparison and summary, with independent reconstruction and full numerical grids retained in the research archive. These results concern initial discovery. The document reasoning study in Appendix C.8 remains prospective.

H PLANNING EVIDENCE AND REPRODUCIBILITY

This appendix gives the deterministic planning comparisons behind Figure 1. The reference AUTOFORM mechanism appears separately in Figure 4.

H.1 TASKS, WORKERS AND ACCOUNTING

The original block uses MovingAI maps `empty-16-16.map`, `room-32-32-4.map`, `maze-32-32-2.map` and `warehouse-10-20-10-2-1.map` (Stern et al., 2019). Four and eight robots with seeds 48 through 71 give 192 tasks and 1152 executions. Starts and goals are sampled without replacement from the largest connected component under public priorities. Four robot starts match the first four of the eight robot assignment. All related assignments and methods remain in the same seed cluster. Development seeds 1004 and 1005 are excluded; every evaluation task is retained.

Robots move to a cardinal neighbor or wait and remain at reached goals. Vertex and opposing edge collisions are forbidden. The horizon is twice the sum of map width and height plus 64. Society workers retain fixed roles, priorities, authority and reservation based A*. Replanning requires changed higher priority reservations and an absent or invalid route. The authority retains its best feasible submission by total travel time, then makespan.

Formation starts with reporting channels to the authority. After a two round pilot and every two rounds thereafter, the organizer reads the latest originating records, including failures, and adds higher to lower priority channels between conflicting successful routes. Channels and histories persist. Global organizer access is charged; workers receive evidence only through their histories and admitted delivery.

The six methods share capabilities and caps. Complete forwarding sends the newest known records; selective fixed sends only changed originating records to lower priority workers and the authority. Both retain complete permissions. Frozen revision keeps the initial organization after its charged pilot, and trace removal hides organizer observations. The centralized planner uses the same A* tool, explores cyclic and reversed priorities within the shared caps and retains its best feasible joint plan.

Caps are 128 shared worker and organizer calls, 128 tool calls and 100000 total expansions, also capped at 100000 per search. Serialized input and output limits are 900000000 and 4000000 bytes. First feasibility counts a committed authority action and its tools, excluding abandoned proposals. An independent schedule checker treats invalid and missing submissions as failures.

H.2 ORIGINAL COMPARISON AND REGISTERED CRITERIA

Figure 5 displays completion against mean total input for all six methods, retaining failures in both coordinates. Its bars show sample standard deviations of the twenty four seed means, not uncertainty over new maps. Peer payload and search work provide complementary comparisons below.

Formation, selective fixed and complete forwarding solve the same 187 tasks. Selective fixed preserves complete forwarding's searches and authority decisions while reducing mean peer payload by 95.4312%. Their ratio is 0.04569 with descriptive interval $[0.04380, 0.04768]$, isolating redundant forwarding.

The registered formation criterion requires a feasibility difference lower bound of at least -2.5 percentage points and a peer ratio upper bound of at most 0.8. The reverse criterion retains the feasibility margin but permits ratio at most one. Formation's ratio is 0.76404 with interval $[0.64476, 0.89386]$. Paired feasibility agrees, but neither payload criterion passes. The interval supports a reduction, not the registered saving of at least 20%.

Intervals use 10000 bootstrap draws with seed 90210 over the twenty four generator seed clusters, keeping maps, team sizes, methods and failures together. Ratios compare resampled arithmetic means. Two zero means give ratio one; positive over zero is unbounded. These descriptive intervals concern the selected maps, without multiplicity adjustment or an equivalence interpretation.

Formation uses 1.27106 MiB mean input against 1.01825 for selective fixed. Their four eight robot maze failures and one warehouse failure also occur under complete forwarding. Formation sends more mean peer payload in both maze settings. Centralized planning adds three maze successes without losses; both revision ablations solve 93 tasks. On 187 joint successes, formation uses 1085.04 more expansions before first feasibility. Centralized priority retries continue afterward, so lower total search need not mean earlier feasibility.

H.3 Transfer without policy changes

Transfer uses `maze-32-32-4.map`, `room-64-64-8.map`, `random-32-32-20.map` and `den312d.map`, selected from catalog metadata and checked as new geometries before freezing. Seeds 72 through 95, both team sizes and six unchanged methods give another 192 tasks and 1152 executions. The blocks have disjoint assignments and are analyzed separately.

The three principal societies again agree on every success. Selective fixed preserves complete forwarding's searches and decisions with 95.7392% less mean peer payload. Formation's ratio to selective fixed is 1.06767 with interval $[0.85972, 1.28769]$. Both payload criteria fail. The reversed mean order establishes neither an advantage nor equivalence. Mean total input is 1.28049 versus 1.00167 MiB.

Formation has less mean peer payload in seven settings but much more in the eight robot maze. Its 17 failures reach the expansion cap, comprising five in that maze, three in the four robot rooms and nine in the eight robot rooms. Centralized planning adds six successes without losing a formed success, a 3.125 percentage point advantage with interval $[1.0417, 5.2083]$. Each ablation solves 77 tasks.

H.4 Variation across settings and individual tasks

Formation and selective fixed submit identical strings on every task in both blocks, including failures. Success, travel time and makespan therefore agree, while formation has larger total input on every pair. Ratios of the means are 1.24828 and 1.27836, giving input penalties of 24.83% and 27.84%.

Formation and centralized planning respectively solve 20 and 23 original eight robot maze tasks, and 15 and 18 transfer eight robot room tasks.

H.5 From individual traces to population comparisons

Four prospectively selected cases use seed 72 and eight robots on every transfer map. The maze and random map succeed without revision. In the game map, three added channels carry route records from robots 0, 3 and 6 into robot 2's next reservation search. Formation succeeds at round 5 while frozen revision fails, but selective communication reaches the same plan at round 4. The room task exhausts 100000 expansions despite five additions. Across these cases, formation sends fewer packets but more bytes than selective fixed. Figure 5 places these examples within the complete populations, retaining all six methods and every failure.

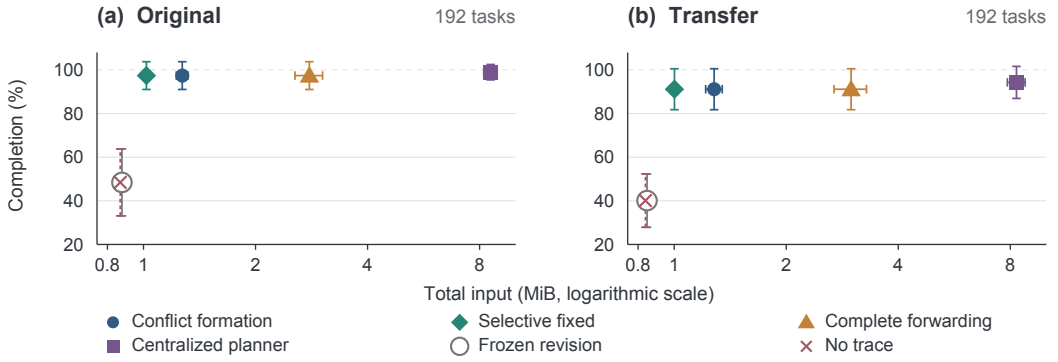

Figure 5: **Completion and input on public planning tasks.** All six methods retain every task and failure in each separate block. Total input includes organizer work. Points show means and bars one sample SD over 24 seed clusters, each containing four maps and both team sizes. The horizontal axis is logarithmic. Section H.1 specifies the tasks and accounting.

Formation reduces peer payload on 156 original and 144 transfer tasks, but increases total input on all 384. All five original and seventeen transfer failures remain included. The largest peer increases occur in the maze settings.