Q-COTRUST: ECONOMIC AND GOVERNANCE BOUNDARIES OF QUANTUM COORDINATION IN DISTRIBUTED HUMAN–AI DECISIONS

Anonymous authors
Paper under double-blind review

ABSTRACT

Distributed Human–AI decisions increasingly allocate scientific verification, public services, and autonomous actions across agents that hold private information, yet communication, verification, and review consume scarce economic and computational resources. Quantum correlations can expand zero-message coordination, but computational capability has economic significance only if it beats feasible substitutes after implementation and oversight are priced. Existing quantum multi-agent studies and benchmark frameworks rarely identify these margins within one matched evaluation. We introduce Q-COTRUST, a role-separated diagnostic benchmark testing whether quantum-assisted coordination retains *modeled adoption value*—the utility difference between configurations net of specified charges—against feasible classical substitutes, finite quantum resources, and technical controls. Our synthetic two-site scenario routes valuable verification tasks to simulated human or AI workers under private bits, delayed reliability audits, limited messaging, scarce pairs, and capacity-limited review. Q-COTRUST combines an eight-cell structural grid over information, payoff, and communication with a separate eight-regime study of an ideal policy (A), pair infrastructure (I), and technical controls (T), evaluated over 30 paired seeds and an exact labelled-K_3 rendezvous check. In the selected local, joint-payoff, no-message cell, the full regime exceeds an ungated classical backbone by +0.076 [+0.070, +0.083] scenario utility per task; yet the governed classical mean is higher (0.392 versus 0.389), the paired 111–001 interval is unavailable, finite realization changes utility by -0.024 [-0.032, -0.016] relative to ideal A+T at the tested prices, and a timely classical bit exceeds the ideal no-message Bell policy ($2/3$ versus $7/12$). These findings recast quantum advantage as an economic–computational identification problem and yield a staged rule: diagnose the bottleneck, benchmark the strongest timely classical alternative, then price realization and control. Because the evidence is synthetic and two-site, it cannot support procurement or deployment claims; that limit defines a three-level program—artifact validation, price–capacity frontiers, and measured network and participatory field studies—for testing transfer and identifying who bears coordination errors and costs.

1 INTRODUCTION

Consequential decisions are often made by teams whose members hold different pieces of information. Economics formalizes a *team decision problem* as agents sharing an objective while acting on private signals (Radner, 1962); information is valuable only when it improves a decision enough to justify its acquisition and use (Stigler, 1961). Decentralized learning confronts the same structure under partial observability, delayed feedback, and restricted communication (Oliehoek & Amato, 2016; Dafoe et al., 2021). We call it an *information bottleneck* when payoff-relevant signals are dispersed and cannot be pooled accurately, cheaply, and in time. Scientific verification, distributed infrastructure, emergency response, and autonomous systems all face variants of this economic–computational problem.

Quantum correlations offer a distinctive response: separated decision-makers can generate joint actions unavailable to a specified zero-message classical policy class without transmitting their

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.

private signals. This is a claim about *capability*, the attainable correlation law. Adoption concerns *value*: whether that capability still improves a decision after comparison with timely classical communication, finite and noisy resources, and responsible-use controls. Quantum multi-agent work mainly studies policy capability (DeRieux & Saad, 2025; Gardiner et al., 2026; Dahia & Szabo, 2026); network research models pair supply and decay (Wehner et al., 2018; Khatri, 2021); and benchmark research demands matched alternatives and resource accounting (Dehghani et al., 2021; Koch et al., 2026; Aboumrad et al., 2026). These strands rarely identify, within one matched evaluation, the separate margins of an ideal policy, finite realization, and routing and review.

We define Q-COTRUST as a role-separated diagnostic benchmark asking when quantum-assisted coordination retains modeled adoption value after classical substitutes, finite realization, and technical controls are matched. **Q** denotes quantum-assisted correlation and **Co** coordination under dispersed private information; **Trust** is limited to audit-updated reliability, routing, certification, fallback, and capacity-limited review—not interpersonal trust, institutional legitimacy, or social welfare. In our synthetic two-site scenario, each site observes one private bit and routes a task to a simulated human or AI worker; reliability changes, audits arrive late, messages and pairs are scarce, and review is finite. *Modeled adoption value* is the difference in dimensionless scenario utility per task between two named configurations net of specified charges; it is not return on investment or a procurement recommendation.

Three questions organize the contribution.

RQ1. When is additional coordination information useful? A structural grid crosses pooled/local observations, separable/joint payoffs, and one-bit/no-message conditions with explicit classical controls.

RQ2. Which comparison identifies capability, realization, and governed use? Eight configured regimes share an audit-updated classical backbone and distinguish an ideal correlation policy (A), modeled finite pairs (I), and technical routing/review controls (T). They are system coalitions, not invariant one-factor interventions.

RQ3. What evidence is required before a boundary can guide adoption? Paired simulation estimates and exact small-game controls lead to an evidence ladder from mechanism validation to system economics and institutional study.

The structural and configured 2^3 designs answer different questions; no 64-condition Cartesian study was performed.

The comparator changes the conclusion (Figure 1). In the selected local, joint-payoff, no-message cell, the full regime exceeds the ungated audit-updated backbone by +0.076 [+0.070, +0.083] scenario utility per task. Yet its mean is below the governance-matched classical regime (0.389 versus 0.392), and the missing paired interval precludes an incremental quantum-value claim. Finite realization changes utility by -0.024 [-0.032, -0.016] relative to an ideal A+T counterfactual at the tested prices. An exact labelled-K_3 check reveals the same boundary: the ideal no-message Bell policy beats the zero-message classical optimum ($7/12$ versus $5/9$), but one timely classical bit does better ($2/3$).

The transferable contribution is therefore an identification method, not a general quantum-advantage result: *advantage relative to which feasible information, resource, and governance alternative?* It yields a staged rule—diagnose the bottleneck, benchmark the strongest timely classical substitute, then price realization and control. Because the evidence is synthetic and two-site, it cannot establish cost robustness, hardware viability, legitimacy, or deployment value. Those limits define a three-level program: validate the mechanism and artifact, estimate price–capacity frontiers with measured system traces, and test institutional consequences with affected stakeholders.

An anonymous code and evidence artifact is available at https://anonymous.4open.science/r/Q-CoTrust-11CE/.

2 Related Work and the Identification Gap

Information economics and decentralized learning. The economics of information asks when an additional signal changes behavior enough to justify its acquisition cost (Stigler, 1961). Team decision theory sharpens that question for agents with a shared objective but different private observations

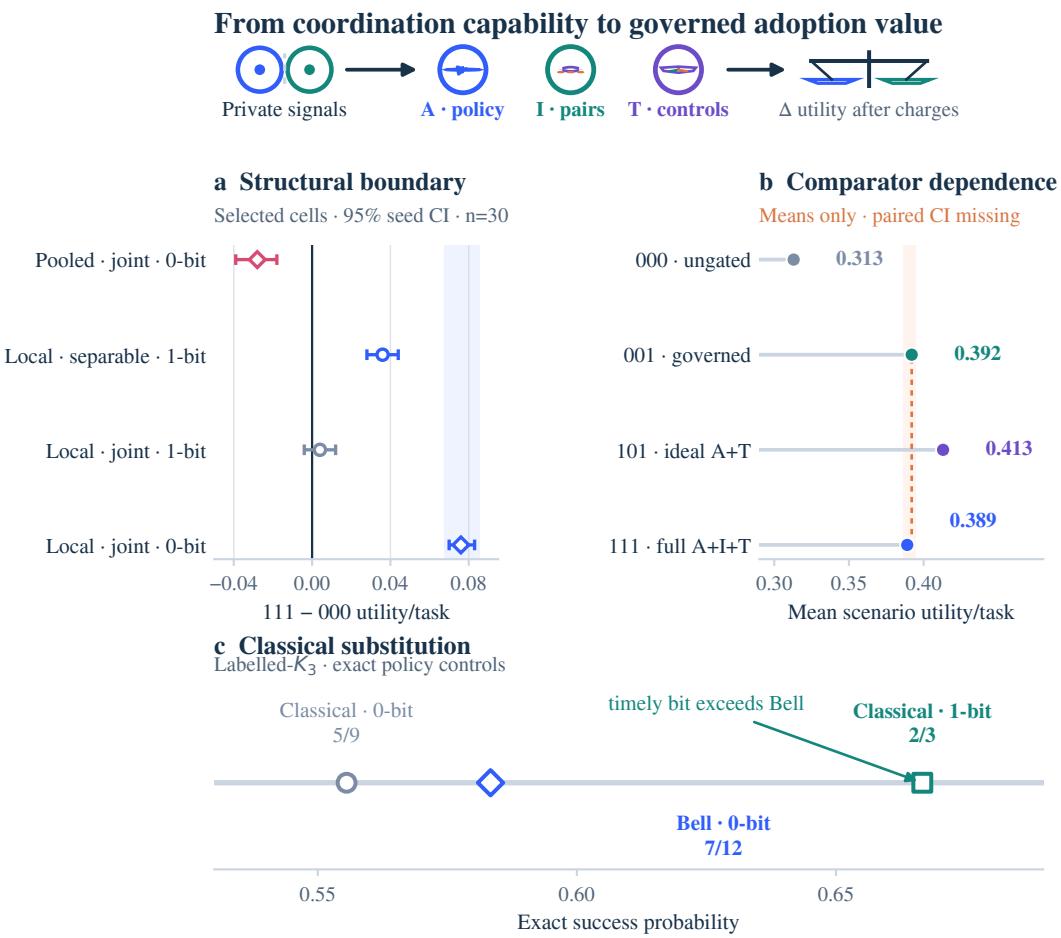


Figure 1: **Q-COTRUST turns coordination capability into a governed adoption-value question.** The conceptual rail uses a Bloch-sphere policy, finite entangled-pair stock, and a branching control shield to separate analytical policy (A), modeled infrastructure (I), and technical controls (T); arrows show reading order, not causality. The result gallery reports (a) four selected 111–000 structural contrasts with 95% seed-bootstrap intervals ($n = 30$), (b) configured-regime means while marking the paired 111–001 interval as unavailable, and (c) exact labelled-K_3 controls showing that a timely classical bit exceeds the ideal no-message Bell policy. Values concern synthetic utility and exact controls, not people, hardware, social welfare, or deployment.

(Radner, 1962). Decentralized partially observable Markov decision processes (Dec-POMDPs) provide a computational formulation: each agent chooses from local information while rewards depend on joint behavior (Oliehoek & Amato, 2016). Centralized-training/decentralized-execution methods learn within that structure (Lowe et al., 2017; Rashid et al., 2018; Yu et al., 2022), and cooperative-AI research emphasizes communication and institutions as part of the coordination problem (Dafoe et al., 2021). Our benchmark links these traditions by treating a quantum correlation as one costly information-and-action technology among classical communication, fallback, and review. It does not estimate a market price, social welfare, or strategic equilibrium.

Quantum coordination and finite resources. Parameterized quantum policies establish a policy class rather than hardware evidence (Jerbi et al., 2021). Quantum multi-agent reinforcement learning has used variational actors, centralized critics, and entangled channels in communications and mobility tasks (Yun et al., 2022; Park et al., 2023; 2024). eQMARL distributes an entangled actor–

critic (DeRieux & Saad, 2025); later work studies sequential entanglement-assisted coordination and invokes nonlocal-game ceilings to distinguish entangled from unentangled policies (Gardiner et al., 2026; Dahia & Szabo, 2026). These results establish that nonclassical correlation can enlarge a specified decentralized policy class. Quantum-network research separately models stochastic pair generation, storage, decay, and routing (Wehner et al., 2018; Khatri, 2021). Neither strand by itself determines whether a finite, noisy, priced implementation remains preferable to a timely message or governed classical route. We therefore treat the idealized eQMARL-style coordinator as an adaptation and optimistic analytical reference, not a reproduction of eQMARL or evidence from a quantum network.

Technical governance and benchmark design. Here *technical governance* means executable controls within the modeled decision path: certification, abstention, fallback, capacity-limited review, and delayed audit updates. It is narrower than institutional legitimacy or public accountability. Trust models aggregate uncertain evidence (Huynh et al., 2006), while human–multi-agent research distinguishes reported confidence from observed performance (Rabby et al., 2024); quantum-inspired trust-aware allocation instead uses classical annealing for assignment (Barroso et al., 2025). We retain only the mechanisms needed to ask how an audit-updated worker selector and a routing/review layer alter a coordination comparison.

Benchmark conclusions depend on task construction, comparator strength, and evaluation access (Dehghani et al., 2021; Raji et al., 2021). AstaBench, an ICLR 2026 datasets-and-benchmarks paper, illustrates a standard interface, controlled execution, broad baselines, and accounting for cost and tool-access confounders (Bragg et al., 2026). Quantum application benchmarks likewise call for matched alternatives, shared instances and metrics, and transparent resource requirements (Koch et al., 2026; Aboumrad et al., 2026). Table 1 makes the cumulative design logic explicit: Q-COTRUST combines a quantum coordination policy, modeled finite realization, technical controls, matched classical alternatives, and a common evaluation interface. GlobalTeam, ClaimVerify, and PublicResponse vary distributions and payoff rules within that interface; they are not three validated application domains. A one-step labelled-K_3 rendezvous protocol supplies an exact, structurally distinct check grounded in prior graph-rendezvous work (Mironowicz, 2023; Viola & Mironowicz, 2024; Tucker et al., 2024).

Table 1: **Q-COTRUST integrates evaluation dimensions developed separately in representative prior work.** Status denotes whether a dimension is central and explicit (full), present but secondary or incomplete (partial), not demonstrated, or outside the work's scope. This is a scoped comparison rather than a systematic review; an unfilled entry does not imply a flaw in the cited work.

Representative work	**Quantum** policy	**Finite** implementation	**Technical** controls	**Matched** alternatives	**Reusable** interface
eQMARL (DeRieux & Saad, 2025) entangled multi-agent learning	●	○	○	◑	◑
Quantum application benchmarks (Koch et al., 2026; Aboumrad et al., 2026) instances, metrics, and resource reporting	○	●	○	●	●
AstaBench (Bragg et al., 2026) controlled scientific-agent evaluation	—	◑	◑	●	●
Q-COTRUST (this study) role-separated adoption-value microbenchmark	●	●	●	●	◑

● fully addressed ◑ partially addressed ○ not demonstrated — outside scope. Q-COTRUST's interface remains partial pending an independently exercised artifact and raw paired record.

3 Diagnostic Benchmark: Information, Resources, and Control

3.1 Distributed decision, information, and value

At task t, sites observe private bits x_t and y_t, while public context contains value $v_t \in \{1, 1.5, 2\}$ and risk flag r_t. Each site selects a synthetic human or AI worker and an intended binary action. Latent

worker reliability $\rho_{i,w,t}$ can enter and recover from compromise; compromised workers over-report it. The common backbone updates Beta–Bernoulli reliability beliefs only from audited outcomes that arrive two steps later. Consequently, the decision may use local observation, public context, pre-t beliefs, available message conditions, and current resource state, but not the task outcome, a future audit, or an unchosen action's score. Appendix Figure 3 makes this temporal information contract explicit.

The *separable* payoff reproduces each site's observation. The *joint* payoff depends on both private signals:

$$a_{1,t} \oplus a_{2,t} = g_d(x_t, y_t), \tag{1}$$

where g_d is an adapter-specific AND or OR rule. Pooled observations or one transmitted bit let a classical policy construct g_d before worker error. Under local observations with no decision-time message, the best deterministic local strategy is enumerated for the adapter's weighted input distribution. This game is not uniform CHSH; CHSH is used only as an implementation check (Appendix D).

The objective is dimensionless scenario utility per task,

$$U_t = \begin{cases} v_t, & \text{success,} \\ -\lambda v_t(1 + 0.5 r_t), & \text{failure,} \end{cases} - c_t^{\text{worker}} - c_t^{\text{message}} - c_t^{A} - c_t^{I} - c_t^{T} - c_t^{\text{audit}} - c_t^{\text{review}}, \tag{2}$$

with $\lambda = 0.7$. "Normalized" in the result files refers to these common scenario units; it does not mean rescaling to $[0, 1]$. Costs are controlled scalars, not measured money, energy, latency, or stakeholder welfare. We report success, cost, correlation invocation or pair use, certification, review, fallback, and belief error where recorded, with denominators tied to scenario tasks.

For two named configurations m and m_0, their comparison-specific *modeled adoption value* is

$$\Delta(m, m_0) = \mathbb{E}[U_t \mid m] - \mathbb{E}[U_t \mid m_0]. \tag{3}$$

The comparator is part of the estimand: changing m_0 changes the question. A positive Δ means only that one configured system has higher expected utility under this synthetic generator and cost vector. It does not imply causal attribution to a component, social welfare, financial return, or institutional approval.

3.2 Configured regimes and common backbone

A—learned analytical policy. A two-party correlation policy learns four measurement angles. The specified joint sampler for modeled visibility ν is

$$E_{xy} = \nu \cos(\alpha_x - \beta_y), \quad p(a_1, a_2 \mid x, y) = \tfrac{1}{4}\left[1 + (-1)^{a_1 \oplus a_2} E_{xy}\right]. \tag{4}$$

It follows that $\Pr[a_1 \oplus a_2 = g_d(x, y)] = \frac{1}{2}[1 + (-1)^{g_d(x,y)} E_{xy}]$. Equation 4 has uniform local marginals, so summing over either output removes dependence on the remote input. When A is on and I is off, the central simulator draws from this law conditional on (x, y). Neither simulated site receives the other's private bit. This *pair-free* condition is an ideal policy-class counterfactual, not a physically implementable decentralized protocol without a correlation resource. The missing executable artifact still prevents source-level verification that its sampler and interfaces implement this specification.

I—modeled infrastructure. At most three pairs arrive stochastically, decay in storage, are consumed on use, and incur a scalar charge. With I alone, the system uses a fixed Bell policy; with A+I, learned angles operate at realized visibility. The model captures inventory, quality, and cost but not density matrices, routing, error correction, hardware latency, or energy.

T—technical controls. Technical governance is the executable control path around a decision, not a claim of institutional legitimacy. Its contextual gate chooses the audit-updated classical path, a certified correlation path, or institutional review. Certification checks pair availability and visibility when I is active; without I, it can attest only the analytical model. Review is eligible only when $r_t = 1$ and $v_t \geq 1.5$ and is capacity limited; ineligible requests fall back to the classical route. All eight regimes retain identical audit-updated worker selection and classical coordination. Configuration 000 therefore retains reliability modeling even though the T bundle is off.

Table 2: **Operational semantics of the eight configured regimes.** Every row retains the audit-updated worker selector and the classical rule allowed by its structural cell. The A/I/T code labels a coalition; it is not an orthogonal primitive-factor design.

Code	Coordination rule	Correlation realization	Additional T routes
000	allowed-information classical rule	none	none
001	allowed-information classical rule	none	gate, fallback, eligible capacity-limited review
010	fixed Bell rule	finite visible stock; decay, consumption, and pair charge	none
011	fixed Bell rule	finite visible stock; decay, consumption, and pair charge	certification, gate, fallback, review
100	learned angles	ideal central joint sampler; no pair stock	none
101	learned angles	ideal central joint sampler; no pair stock	analytical attestation, gate, fallback, review
110	learned angles	finite visible stock; decay, consumption, and pair charge	none
111	learned angles	finite visible stock; decay, consumption, and pair charge	certification, gate, fallback, review

Table 3: **Three evaluation groups.** The structural and configured-regime designs are distinct 2^3 studies; only the latter is run in the selected hard cell. No 64-condition Cartesian study is claimed.

Group	Varied factors	Scope and purpose
Structural grid	pooled/local; separable/joint; one-bit/no-message	all three adapters; locates the information bottleneck (RQ1)
Configured regimes	learned policy A; pair stock I; technical controls T	all adapters in local/joint/no-message; compares named coalitions (RQ2)
Sensitivity and check	unreliable-worker share; visibility; labelled-K_3 protocol	two sweeps plus exact and finite-stock structural checks (RQ3)

The codes in Table 2 describe system coalitions. Their bits are convenient labels, but their operational meanings are conditional: switching I on beside A replaces an ideal sampler with a finite realization, while I without A introduces a fixed Bell rule. We therefore name comparisons by the regimes they connect. Configuration 000, hereafter the *audit-updated classical backbone*, uses the information allowed by the structural cell, delayed Beta–Bernoulli worker updates, and the enumerated classical action rule. It has no T routing or review. Configuration 001 is the governance-matched classical comparator for the full 111 regime. Exact optimality over richer public-context-conditioned or history-dependent classical policies cannot be checked without the executable artifact and remains an author verification item.

3.3 Adapters and extension contract

GlobalTeam uses an AND payoff with cross-site authorization; ClaimVerify and PublicResponse use OR payoffs with different event, risk, and worker-reliability distributions. These narratives explain the generative parameters but do not convert synthetic tasks into organizational evidence. A new adapter supplies observations, latent-state dynamics, payoff, exogenous-event generator, and exact classical controls through the common interface; a new method receives the same allowed inputs and is scored by the unchanged evaluator. Appendix B records intended use, exclusions, data generation, metrics, compute, and extension requirements.

4 EVALUATION PROTOCOL

Training, evaluation, and pairing. Each of 30 study seeds supplies identical exogenous trajectories to compared methods, preventing resource arrivals or task difficulty from being confounded with method choice. Policies train for 120 twelve-task episodes and are evaluated on 80 unseen episodes; all study seeds are retained. MAPPO uses decentralized tabular actors and a centralized critic (Yu et al., 2022); correlation angles use exact expected-reward gradients, and the T gate uses contextual temporal-difference updates. A disjoint three-seed development stage fixed hyperparameters, scenario distributions, and normalized costs before the reported study. A configuration contrast compares complete configured systems and is not interpreted as post-hoc masking of a single fixed policy.

Comparators and structural controls. The comparison set contains an exact reported-reliability rule, tabular MAPPO, an optimistic eQMARL-style adaptation with visibility one and no pair charge, Trust-JMARL with delayed audits, the audit-updated classical backbone (000), the governance-matched classical regime (001), full A+I+T (111), and an unattainable centralized oracle. Pooled information, separable payoff, and timely one-bit communication control the hypothesized information bottleneck. The centralized and idealized references use privileged or cost-free conditions and are interpretive references, not attainable competitors. Methods share task inputs, scoring, exogenous trajectories, and the reported training horizon; physical-resource charges apply only to configurations that consume the modeled stock. The present source does not specify every baseline operation well enough for independent reimplementation; the anonymous artifact must supply the policies, lookup tables, and allowed-input manifests.

Estimands, uncertainty, and robustness. The legacy structural estimand is the seed-paired difference in scenario utility per task between 111 and 000, reported by structural cell and adapter. Means and 95% intervals use 5,000 deterministic bootstrap resamples of paired simulation seeds; tasks within a seed and the three related adapters are not independent replications. These intervals describe variation under the fixed synthetic generator and cost vector, not uncertainty over applications, hardware, or parameter choice. The configured-regime study reports 111–000 and named conditional contrasts. For adoption interpretation, 111–001 is the governance-matched comparison; only rounded cell means are present in the available source, so its paired interval remains unverified. Stress suites vary unreliable workers over $\{0, .1, .2, .3, .4\}$ and mean visibility over $\{.72, .78, .84, .90, .96\}$ with a .76 certification floor. They do not vary the major prices or review capacity. All tested cells and seeds are retained; only the selected hard-cell contrast is treated as primary, and the remaining grids are descriptive.

Exact and external-structure checks. The recorded protocol calls for checks of joint-probability normalization, no-signalling marginals, information isolation, audit timing, deterministic seed keys, resource conservation, review capacity, weighted-game enumeration, and standard CHSH values. Their execution is not assessable without the artifact. The post-design rendezvous check instantiates the published one-step labelled-K_3 game (Tucker et al., 2024): two simulated agents see only their own vertices and must move without waiting or messaging. We enumerate deterministic zero-, one-, and two-bit classical controls and evaluate the published Bell strategy exactly. A 30-seed finite-stock simulation uses 20,000 expectation-scored missions per visibility center. It is a limited structural check, not an independent domain, preregistration, navigation study, or robot experiment.

5 RESULTS: WHERE COORDINATION CREATES—AND LOSES—VALUE

5.1 RQ1: THE BOUNDARY DEPENDS ON INFORMATION, PAYOFF, AND COMMUNICATION

In the selected local, joint-payoff, no-message cell, the full regime obtains 0.389 [0.382, 0.397] scenario utility per task versus 0.313 [0.304, 0.321] for the audit-updated classical backbone, a structural-study paired difference of +0.076 [+0.070, +0.083]. Pairs are consumed on 19.6% [16.9, 22.3] of tasks and review is used on 23.4% [22.3, 24.3]; remaining tasks take or fall back to the classical route. The idealized coordinator reaches 0.303 [0.292, 0.314] without pair charges. The centralized oracle remains far higher (Table 4), showing that better evidence and worker reliability could matter more than correlation expressivity. Learning traces and adapter-level estimates appear in Appendix C.

Table 4: **Selected hard cell in the structural study.** Adapter-averaged seed means with 95% paired simulation-seed bootstrap intervals ($n = 30$). The ideal-correlation reference pays no physical-pair charge; its last-column value is invocation, while the full-regime value is consumed-pair use.

Method	Utility/task ↑	Success ↑	Cost ↓	Corr. invoke/use ↓
Rule-based	0.229 [0.221, 0.237]	0.592 [0.589, 0.594]	0.115 [0.115, 0.115]	0.000 [0.000, 0.000]
Tabular MAPPO	0.115 [0.100, 0.130]	0.549 [0.543, 0.555]	0.115 [0.115, 0.115]	0.000 [0.000, 0.000]
eQMARL-style ideal	0.303 [0.292, 0.314]	0.619 [0.615, 0.624]	0.115 [0.115, 0.115]	1.000 [1.000, 1.000]
Trust-JMARL	0.106 [0.088, 0.125]	0.549 [0.542, 0.556]	0.129 [0.129, 0.130]	0.000 [0.000, 0.000]
Audit-updated classical (000)	0.313 [0.304, 0.321]	0.628 [0.625, 0.630]	0.129 [0.129, 0.130]	0.000 [0.000, 0.000]
Full A+I+T	0.389 [0.382, 0.397]	0.710 [0.706, 0.714]	0.334 [0.326, 0.342]	0.196 [0.168, 0.223]
Central oracle	1.295 [1.292, 1.298]	1.000 [1.000, 1.000]	0.130 [0.130, 0.130]	0.000 [0.000, 0.000]

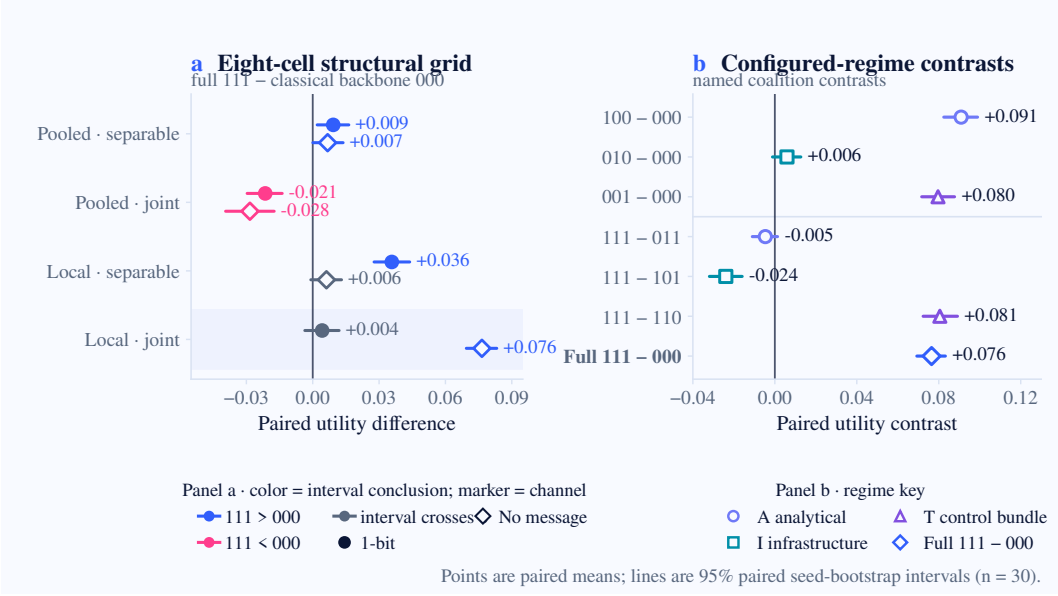


Figure 2: **Structural grid and configured-regime contrasts.** Panel A reports 111 minus 000 in all structural cells: blue paired seed-bootstrap intervals lie above zero, coral below, and slate cross zero; circles denote one-bit communication and diamonds no decision-time message. Panel B reports named regime contrasts; the labels do not imply orthogonal factor effects. Structural and regime estimates come from distinct studies. All non-primary cells are descriptive estimates over 30 paired simulation seeds.

Only the local/joint/no-message cell combines a sizeable 111–000 contrast with nontrivial pair use. Adding one bit changes the contrast to +0.004 [-0.004, +0.012] with 11.5% [9.6, 13.3] pair use; pooling observations yields -0.028 [-0.039, -0.018], and replacing the joint payoff by a separable one yields +0.006 [-0.001, +0.013]. Local/separable/one-bit is positive (+0.036 [+0.028, +0.044]) but invokes pairs on only 0.6% [0.3, 0.9] of tasks. That rate rules out frequent pair-mediated coordination as the explanation, although the available summaries do not isolate routing from fallback. Private information alone is insufficient, and timely classical communication removes the margin under a joint payoff.

5.2 RQ2: CONFIGURED REGIMES ANSWER DIFFERENT COMPARISONS

Relative to 000, the ideal learned-correlation regime 100 differs by +0.091 [+0.083, +0.099], the fixed-pair regime 010 by +0.006 [-0.001, +0.013], and the governed classical regime 001 by +0.080 [+0.072, +0.088]. These are coalition comparisons, not isolated one-role effects. Within

finite-pair regimes, learned rather than fixed angles gives 111–011 $= -0.005[-0.011, +0.001]$. Finite realization relative to the ideal learned-policy-plus-controls counterfactual gives 111–101 $= -0.024[-0.032, -0.016]$. Adding the control bundle to the learned finite-pair regime gives 111–110 $= +0.081[+0.072, +0.089]$, while the whole-package contrast 111–000 is +0.076 [+0.070, +0.084] in the separate regime study. For the governance-matched comparison, the rounded means are 0.389 for 111 and 0.392 for 001. The available result records do not include the seed-paired 111–001 interval, so they do not establish incremental quantum value under matched controls. Appendix E reports all eight regime means.

5.3 RQ3: TESTED ROBUSTNESS IS LIMITED TO RELIABILITY AND VISIBILITY

The 111–000 contrast rises from +0.047 [+0.032, +0.062] with no unreliable workers to +0.133 [+0.106, +0.160] at 40% (Appendix Figure 5). Because routing and review change in the same region, this trend cannot be assigned to quantum coordination. At visibility .72, the contrast is +0.032 [+0.024, +0.041] with 0.0% [0.0, 0.0] pair use. At .96, contrast and pair use are +0.036 [+0.021, +0.050] and 23.4% [18.4, 28.4]. Nearly unchanged value as pair use rises is consistent with the negative 111–101 realization contrast at the chosen prices. These sweeps do not test pair, message, certification, audit, or review prices, so they locate outcomes within one cost configuration rather than a general adoption boundary.

Exact enumeration recovers the no-message classical optimum $5/9 = 0.5556$. The published Bell strategy gives $p_Q(\nu) = 1/3 + \nu/4$ and reaches $7/12 = 0.5833$ ideally, but one-way one-bit and two-way one-bit controls reach 0.6667 and 0.7778. Pre-shared entanglement transmits no location. It beats the zero-message classical value only for $\nu > 0.889$; after the benchmark's .085 charge, the $v = 2, r = 1$ use threshold is $\nu > 0.972$. At visibility centers .96 and 1.00, pair-use/net-gain are 2.8%/+0.0002 and 58.1%/+0.0131. These thresholds depend on the stated value, risk, costs, and noise model; they are not hardware requirements. Appendix J gives the derivation.

6 IMPLICATIONS AND A THREE-LEVEL RESEARCH PROGRAM

Identification before ranking. Q-COTRUST makes adoption value comparator-specific. The structural grid asks whether coordination can help; the regime grid distinguishes whole-package (111–000), finite-versus-ideal (111–101), and governance-matched (111–001) comparisons. Because A/I/T are conditional coalitions, the missing paired 111–001 interval blocks an incremental quantum-value claim.

Information has substitutes and implementation costs. Information economics values a signal by decision improvement net of acquisition cost (Stigler, 1961). Correlation competes with pooling, a timely bit, fallback, and review; it matters only when dispersed signals jointly determine payoff and benefits survive communication and realization costs. Because prices, delays, fidelity, and opportunity costs are unmeasured, the results imply neither return on investment nor welfare. Practitioners should specify timing, benchmark the strongest budget-matched classical substitute, separate policy from realization, and evaluate controls. The Trust label remains technical; legitimacy, distribution, and public value require other evidence and decision rights.

A cumulative research program. **Level 1** validates the mechanism and artifact through raw paired rows, information-flow tests, the 111–001 interval, and clean-room reuse. **Level 2** estimates price–capacity surfaces against equal-budget comparators and network traces. **Level 3** tests multi-site use with measured costs, stakeholder-defined outcomes, consent, appeal, and burden allocation. Appendix H specifies falsifiers and governance gates; missing evidence remains a requirement, not a result.

7 CONCLUSION

Q-COTRUST tests whether quantum coordination retains adoption value after classical substitution, finite realization, and technical control; it is not a general trust score. In this synthetic study, governed classical coordination is not beaten, finite realization is costly, and a timely bit wins the exact K_3 control. Broader use requires artifact, system, and participatory validation.

STATEMENTS EXCLUDED FROM THE MAIN-TEXT PAGE LIMIT

AI USE STATEMENT (REQUIRED)

Large language models and other generative-AI tools assisted literature discovery and organization; research ideation and methodological critique; early code drafting and debugging; drafting, restructuring, and language editing; title, abstract, and keyword development; migration to the ICLR template; and vector-figure design and layout. They did not generate the reported synthetic observations, select the outcomes reported in the paper, produce mathematical proofs, or act as an evaluated policy, evidentiary source, or author.

The authors checked AI-assisted literature suggestions against the cited sources, checked quantitative statements against versioned study outputs, reviewed all AI-assisted prose and graphics, and independently verified the mathematical statements. The authors remain responsible for the hypotheses, methods, interpretation, claims, code, and final submission.

ETHICS STATEMENT (RECOMMENDED)

This study uses synthetic agents, tasks, changing reliability states, scalar costs, modeled entangled-pair stocks, and exact small-game controls. It contains no human participants, personal data, deployed models, physical robots, quantum devices, or network traces. Human and AI workers are simulated roles, not observations about workers or organizations. Modeled adoption value is dimensionless scenario utility, not social welfare, return on investment, or a procurement recommendation. A foreseeable misuse is to lend a technical veneer to institutional trust or quantum procurement. Field translation would require domain-specific consent, security analysis, audit and appeal rights, participatory outcome design, and measurement of who bears coordination errors and costs.

REPRODUCIBILITY STATEMENT (RECOMMENDED)

Sections 3–4 specify the information contract, objective, configured comparisons, and uncertainty procedure; Appendices B–J provide the benchmark card, analytical checks, complete tables, configurations, derivations, and evidence limitations. This source includes numerical macros, tables, and editable vector figures. The separate anonymous code and evidence artifact is available at https://anonymous.4open.science/r/Q-CoTrust-11CE/. Before submission, a clean rerun should verify the reported outputs, paired 111–001 inference, no-signalling interfaces, resource accounting, and adapter reuse against its executable code, raw paired records, environment specification, and tests.

REFERENCES

Willie Aboumrad, Claudio Girotto, Joshua Goings, Luning Zhao, Miguel Angel Lopez-Ruiz, Daiwei Zhu, Ananth Kaushik, Sayonee Ray, Samwel Sekwao, Jason Iaconis, et al. Measuring what matters: A scalable framework for application-level quantum benchmarking, 2026. URL https://arxiv.org/abs/2604.11781. Preprint.

Giselle Roman Barroso, Duy Ho, Luke Miller, Ahmed Alanazi, and Yugyung Lee. Trust-aware task allocation with quantum optimization in adversarial multi-agent systems. In *2025 IEEE International Conference on Data Mining Workshops*, pp. 2716–2725. IEEE, 2025. doi: 10.1109/ICDMW69685.2025.00348. URL https://doi.org/10.1109/ICDMW69685.2025.00348.

Jonathan Bragg, Mike D'Arcy, Nishant Balepur, Dan Bareket, Bhavana Dalvi Mishra, Sergey Feldman, Dany Haddad, Jena Hwang, Peter Jansen, Varsha Kishore, Bodhisattwa Prasad Majumder, Aakanksha Naik, Sigal Rahamimov, Kyle Richardson, Amanpreet Singh, Harshit Surana, Aryeh Tiktinsky, Rosni Vasu, Guy Wiener, Chloe Anastasiades, Stefanus Candra, Jason Dunkelberger, Daniel Emery, Rob Evans, Malachi Hamada, Regan Huff, Rodney Kinney, Matt Latzke, Jaron Lochner, Ruben Lozano-Aguilera, Ngoc-Uyen Nguyen, Smita Rao, Amber Tanaka, Brooke Vlahos, Peter Clark, Doug Downey, Yoav Goldberg, Ashish Sabharwal, and Daniel Weld. AstaBench: Rigorous benchmarking of AI agents with a scientific research suite. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=M7TNf5J26u.

Boris S. Cirel'son. Quantum generalizations of bell's inequality. *Letters in Mathematical Physics*, 4 (2):93–100, 1980. doi: 10.1007/BF00417500. URL https://doi.org/10.1007/BF00417500.

John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt. Proposed experiment to test local hidden-variable theories. *Physical Review Letters*, 23(15):880–884, 1969. doi: 10.1103/PhysRevLett.23.880. URL https://doi.org/10.1103/PhysRevLett.23.880.

Allan Dafoe, Yoram Bachrach, Gillian Hadfield, Eric Horvitz, Kate Larson, and Thore Graepel. Cooperative AI: Machines must learn to find common ground. *Nature*, 593(7857):33–36, 2021. doi: 10.1038/d41586-021-01170-0. URL https://doi.org/10.1038/d41586-021-01170-0. Comment.

Simranjeet Singh Dahia and Claudia Szabo. Quantum advantage in multi agent reinforcement learning, 2026. URL https://arxiv.org/abs/2605.14235. Preprint.

Mostafa Dehghani, Yi Tay, Alexey A. Gritsenko, Zhe Zhao, Neil Houlsby, Fernando Diaz, Donald Metzler, and Oriol Vinyals. The benchmark lottery, 2021. URL https://arxiv.org/abs/2107.07002.

Alexander DeRieux and Walid Saad. eQMARL: Entangled quantum multi-agent reinforcement learning for distributed cooperation over quantum channels. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=cR5GTis5II.

John Gardiner, Orlando Romero, Brendan Tivnan, Nicolò Dal Fabbro, and George J. Pappas. Learning to coordinate via quantum entanglement in multi-agent reinforcement learning, 2026. URL https://arxiv.org/abs/2602.08965. Preprint.

Trung Dong Huynh, Nicholas R. Jennings, and Nigel R. Shadbolt. An integrated trust and reputation model for open multi-agent systems. *Autonomous Agents and Multi-Agent Systems*, 13(2):119–154, 2006. doi: 10.1007/s10458-005-6825-4. URL https://doi.org/10.1007/s10458-005-6825-4.

Sofiene Jerbi, Casper Gyurik, Simon C. Marshall, Hans J. Briegel, and Vedran Dunjko. Parametrized quantum policies for reinforcement learning. In *Advances in Neural Information Processing Systems*, volume 34, pp. 28362–28375, 2021. URL https://proceedings.neurips.cc/paper/2021/hash/eec96a7f788e88184c0e713456026f3f-Abstract.html.

Sumeet Khatri. Policies for elementary links in a quantum network. *Quantum*, 5:537, 2021. doi: 10.22331/q-2021-09-07-537. URL https://doi.org/10.22331/q-2021-09-07-537.

Thorsten Koch, David E. Bernal Neira, Ying Chen, Giorgio Cortiana, Daniel J. Egger, et al. The quantum optimization benchmarking library. *Nature Computational Science*, 6:653–671, 2026. doi: 10.1038/s43588-026-00991-1. URL https://doi.org/10.1038/s43588-026-00991-1.

Ryan Lowe, Yi Wu, Aviv Tamar, Jean Harb, Pieter Abbeel, and Igor Mordatch. Multi-agent actor–critic for mixed cooperative–competitive environments. In *Advances in Neural Information Processing Systems*, volume 30, 2017. URL https://papers.neurips.cc/paper/7217-multi-agent-actor-critic-for-mixed-cooperative-competitive-environments.

Piotr Mironowicz. Entangled rendezvous: A possible application of Bell non-locality for mobile agents on networks. *New Journal of Physics*, 25(1):013023, 2023. doi: 10.1088/1367-2630/acb22d. URL https://doi.org/10.1088/1367-2630/acb22d.

Frans A. Oliehoek and Christopher Amato. *A Concise Introduction to Decentralized POMDPs*. SpringerBriefs in Intelligent Systems. Springer, 2016. doi: 10.1007/978-3-319-28929-8. URL https://doi.org/10.1007/978-3-319-28929-8.

Chanyoung Park, Won Joon Yun, Jae Pyoung Kim, Tiago Koketsu Rodrigues, Soohyun Park, Soyi Jung, and Joongheon Kim. Quantum multiagent actor–critic networks for cooperative mobile access in multi-UAV systems. *IEEE Internet of Things Journal*, 10(22):20033–20048, 2023. doi: 10.1109/JIOT.2023.3282908. URL https://doi.org/10.1109/JIOT.2023.3282908.

Soohyun Park, Jae Pyoung Kim, Chanyoung Park, Soyi Jung, and Joongheon Kim. Quantum multi-agent reinforcement learning for autonomous mobility cooperation. *IEEE Communications Magazine*, 62(6):106–112, 2024. doi: 10.1109/MCOM.020.2300199. URL https://doi.org/10.1109/MCOM.020.2300199.

Md Khurram Monir Rabby, Mubbashar Altaf Khan, Steven Xiaochun Jiang, and Ali Karimoddini. Performance-aware trust modeling within a human–multi-robot collaboration setting. *ACM Transactions on Human-Robot Interaction*, 13(2):1–23, 2024. doi: 10.1145/3660648. URL https://doi.org/10.1145/3660648.

Roy Radner. Team decision problems. *The Annals of Mathematical Statistics*, 33(3):857–881, 1962. doi: 10.1214/aoms/1177704455. URL https://doi.org/10.1214/aoms/1177704455.

Inioluwa Deborah Raji, Emily M. Bender, Amandalynne Paullada, Emily Denton, and Alex Hanna. AI and the everything in the whole wide world benchmark. In *Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks*, volume 1, 2021. URL https://arxiv.org/abs/2111.15366.

Tabish Rashid, Mikayel Samvelyan, Christian Schroeder de Witt, Gregory Farquhar, Jakob Foerster, and Shimon Whiteson. QMIX: Monotonic value function factorisation for deep multi-agent reinforcement learning. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pp. 4295–4304. PMLR, 2018. URL https://proceedings.mlr.press/v80/rashid18a.html.

George J. Stigler. The economics of information. *Journal of Political Economy*, 69(3):213–225, 1961. doi: 10.1086/258464. URL https://doi.org/10.1086/258464.

J. Tucker, P. Strange, P. Mironowicz, and J. Quintanilla. Quantum-assisted rendezvous on graphs: Explicit algorithms and quantum computer simulations. *New Journal of Physics*, 26(9):093038, 2024. doi: 10.1088/1367-2630/ad78f8. URL https://doi.org/10.1088/1367-2630/ad78f8.

Giuseppe Viola and Piotr Mironowicz. Quantum strategies for rendezvous and domination tasks on graphs with mobile agents. *Physical Review A*, 109(4):042201, 2024. doi: 10.1103/PhysRevA.109.042201. URL https://doi.org/10.1103/PhysRevA.109.042201.

Stephanie Wehner, David Elkouss, and Ronald Hanson. Quantum internet: A vision for the road ahead. *Science*, 362(6412):eaam9288, 2018. doi: 10.1126/science.aam9288. URL https://doi.org/10.1126/science.aam9288.

Reinhard F. Werner. Quantum states with einstein–podolsky–rosen correlations admitting a hidden-variable model. *Physical Review A*, 40(8):4277–4281, 1989. doi: 10.1103/PhysRevA.40.4277. URL https://doi.org/10.1103/PhysRevA.40.4277.

Chao Yu, Akash Velu, Eugene Vinitsky, Jiaxuan Gao, Yu Wang, Alexandre Bayen, and Yi Wu. The surprising effectiveness of PPO in cooperative, multi-agent games. In *Advances in Neural Information Processing Systems*, volume 35, pp. 24611–24624, 2022. URL https://proceedings.neurips.cc/paper_files/paper/2022/hash/9c1535a02f0ce079433344e14d910597-Abstract-Datasets_and_Benchmarks.html.

Won Joon Yun, Yunseok Kwak, Jae Pyoung Kim, Hyunhee Cho, Soyi Jung, Jihong Park, and Joongheon Kim. Quantum multi-agent reinforcement learning via variational quantum circuit design. In *2022 IEEE 42nd International Conference on Distributed Computing Systems*, pp. 1332–1335. IEEE, 2022. doi: 10.1109/ICDCS54860.2022.00151. URL https://doi.org/10.1109/ICDCS54860.2022.00151.

A Appendix Roadmap

Appendix B gives the benchmark card and extension contract. Appendix C reports the temporal information contract, learning traces, and evidence horizon moved from the main text. Appendix D defines the probability checks and configured-regime semantics; Appendix E reports every structural cell and regime; and Appendix F records configurations and artifact requirements. Appendix H turns the present evidence limits into a three-level program with testable questions and governance gates. The remaining appendices distinguish hypothetical application translations from evidence, expand responsible-use considerations, and derive the labelled-K_3 rendezvous controls.

B Benchmark Card

Minimum new-method adapter. A conforming implementation exposes initialization from a deterministic seed, a decision function receiving only the fields in Figure 3, an update function triggered after the specified feedback delay, and a manifest declaring centralized or privileged inputs. The evaluator supplies identical exogenous trajectories, applies the same cost ledger and failure coding, and writes one row per seed–condition–method. Extensions must not modify benchmark internals to obtain additional information.

C Supplementary Learning and Evidence Figures

Figure 7 in Appendix J separates evaluated simulation evidence from the robot, hardware, network, and field tests that remain outside the benchmark.

D Quantum probability checks and configured regimes

CHSH is a unit test, not the contribution. For uniformly distributed inputs and the CHSH predicate $a \oplus b = x \wedge y$, enumeration gives a best deterministic local success probability of $3/4$. Ideal Bell measurements give

$$p_{\text{Bell}} = \cos^2(\pi/8) = \frac{2+\sqrt{2}}{4} \approx 0.853553. \tag{5}$$

The release checks these values independently of the learned policy and checks that all sampled outcome probabilities sum to one. These facts and their MARL demonstrations are prior knowledge (Clauser et al., 1969; Cirel'son, 1980; Dahia & Szabo, 2026); they validate code paths only.

The evaluated game is weighted. Each adapter induces a nonuniform joint input distribution through its latent event and two noisy local evidence channels. The best local deterministic policy is therefore enumerated with the exact adapter distribution, rather than assumed to equal 0.75. For OR, local output relabeling uses $x \vee y = (x \wedge y) \oplus x \oplus y$. The analytical policy learns angles under the same weighted distribution.

Table 5: **Q-CoTrust benchmark card.** Items marked pending are pre-submission requirements, not implied permissions or completed artifacts.

Item	Specification
Intended use	Diagnose configured coordination regimes under controlled information, payoff, communication, reliability, resource, and review conditions. The current release is a synthetic microbenchmark specification, not a demonstrated benchmark suite.
Out of scope	Claims about people, deployed AI, organizations, welfare, physical QPUs or networks, robot navigation, security, clinical validity, procurement, or general quantum advantage.
Data and unit	Fully synthetic episodes. The uncertainty unit is a study seed; tasks within a seed and related adapters are not independent replications.
Generation and splits	A disjoint three-seed development stage fixes settings; 30 study seeds each contain 120 training and 80 evaluation episodes of 12 tasks. Exogenous trajectories are paired across methods.
Tasks	GlobalTeam (AND) and ClaimVerify/PublicResponse (OR) vary event, risk, and worker-reliability distributions within one binary two-site interface. They are structural variants, not independent domains.
Metrics	Normalized net utility, success, modeled cost, pair use, certification, abstention, review/escalation, fallback, and reliability-belief error where recorded; every denominator is tasks within the named condition.
Failure reporting	Retain every seed and tested cell; report null and negative contrasts, unavailable/uncertified pair fallback, ineligible review fallback, and resource depletion.
Compute and dependencies	Recorded reference environment: Python 3.12 with NumPy, pandas, SciPy, and Matplotlib on an x86-64 CPU; no GPU, QPU, or robot simulator. Peak memory was not instrumented.
Extension interface	The specified contract asks an adapter to define observations, latent dynamics, an exogenous generator, payoff, and classical controls. A method should receive only permitted decision-time inputs and be scored by an unchanged evaluator; independent use of this contract remains to be demonstrated.
Licensing status	**Pending author action:** the current manuscript repository specifies no license for executable benchmark code or generated data. The anonymous artifact must include explicit code/data licenses before submission.
Artifact status	The Overleaf repository contains manuscript sources, generated tables/macros, and vector figures. A separate artifact is linked at https://anonymous.4open.science/r/Q-CoTrust-11CE/; executable generators, raw seed-level outputs, tests, licenses, and a minimal new-method adapter must be verified there before submission.

Visibility. For a Werner-like scalar model, visibility $\nu \in [0, 1]$ multiplies the ideal correlation. A stored pair's visibility decays by a fixed multiplicative factor each step. Certification rejects an unavailable pair or a pair below the configured threshold. No density-matrix simulator, gate noise, routing stack, error correction, or QPU execution is used.

Joint law and no-signalling requirement. Equation 4 specifies all four joint outcome probabilities, from which the success parity follows. Non-negativity follows from $|E_{xy}| \leq 1$, and summing over either local output gives $1/2$ for the other output, independent of the remote input. The executable release must check these identities for every adapter input, learned and fixed angle set, tested visibility, and configured regime. It must also verify that neither decision interface receives the remote bit, current or future outcomes, or counterfactual scores. These implementation checks are not assessable from the present Overleaf source.

Configured-regime interpretation and information isolation. Configuration 100 samples the learned joint correlation distribution at ideal visibility without consuming a pair and is therefore an analytical counterfactual. The central simulator has both private bits only to draw and score a joint outcome; neither site's policy receives the other site's bit. This condition is not a physically implementable decentralized protocol without a correlation resource. Configuration 010 consumes the

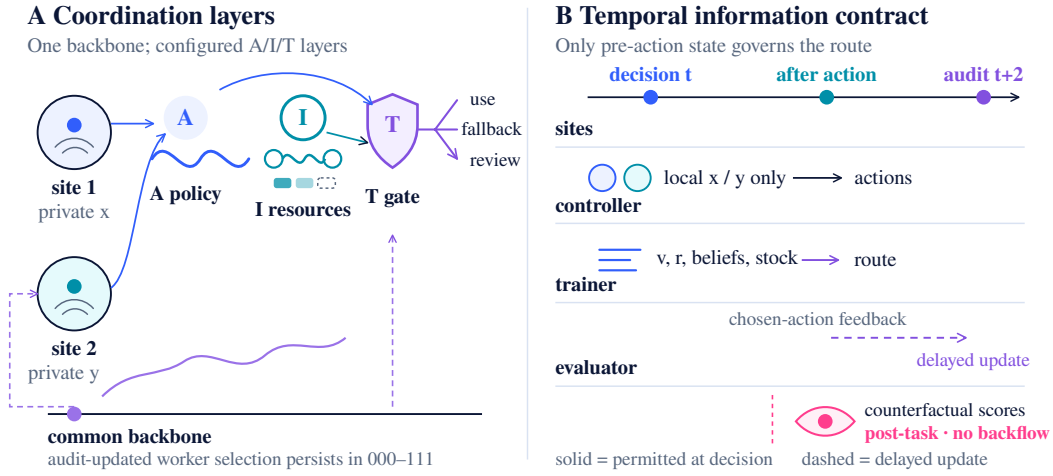


Figure 3: **Coordination layers and temporal information contract.** Panel A separates the analytical correlation policy (A, blue), modeled pair infrastructure (I, teal), and technical routing controls (T, violet) from the audit-updated classical backbone. Panel B shows the decision-time boundary: solid arrows are permitted current inputs; dashed arrows are later outcomes and delayed audits. The evaluator may score counterfactual actions after a task, but that information never enters the current decision. This is a methodological schematic, not measured data or a physical network.

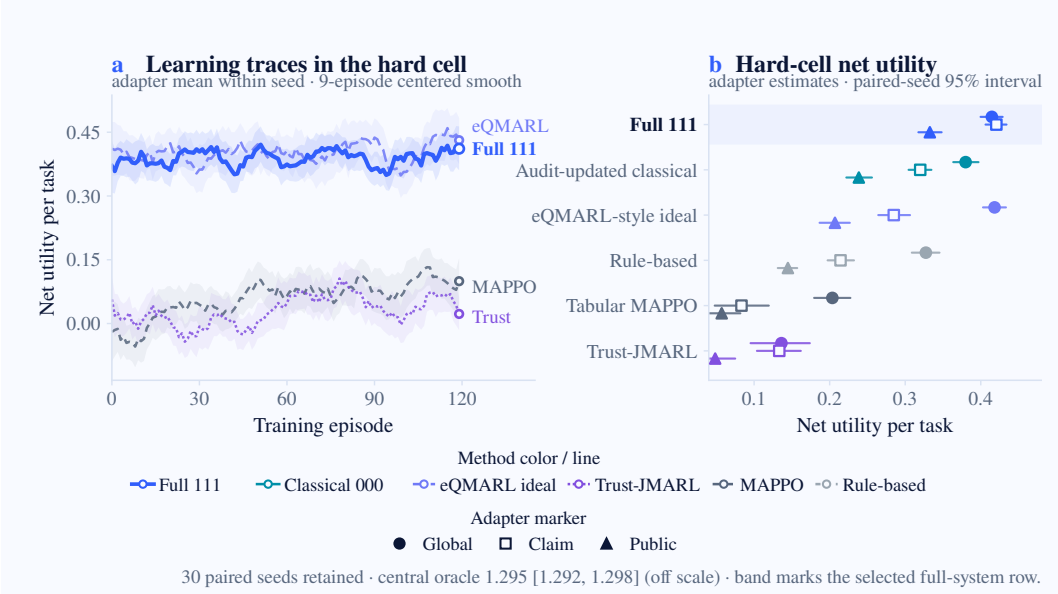


Figure 4: **Learning and hard-cell estimates.** Panel A averages adapters within seed and episode, then shows a centered nine-episode smooth with 95% seed-bootstrap bands. Panel B gives adapter estimates for six on-scale methods; the centralized oracle is off scale. The idealized correlation baseline pays no pair charge.

simulated pair stock with a fixed Bell rule, while 110 couples the learned policy to realized visibility and pair cost. Adding role T introduces the mode gate, certification, abstention, and capacity-limited review; it does not introduce the audit-updated worker selector, which is shared by every setting. The full cube is required because 100 and 110 answer different questions.

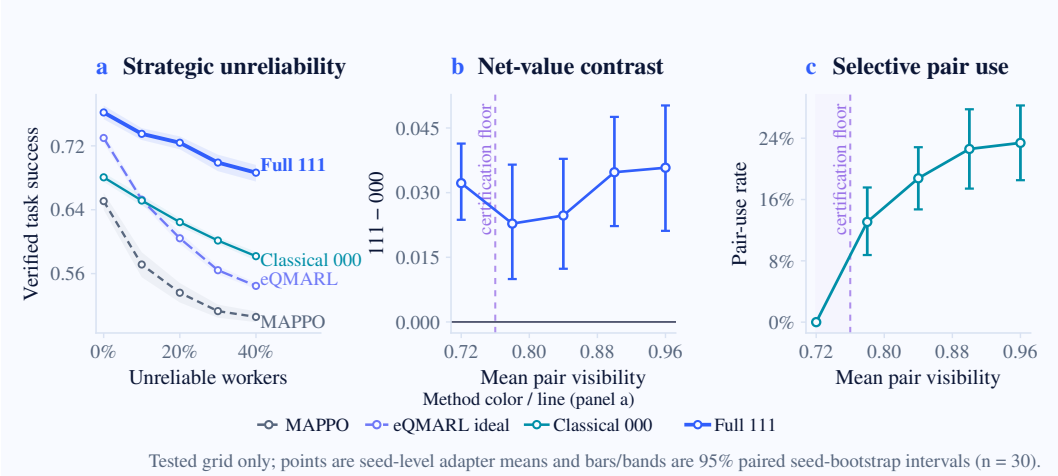


Figure 5: **Sensitivity to unreliable workers and pair visibility.** Points average adapters within paired simulation seeds; intervals are 95% paired seed-bootstrap intervals ($n = 30$). Lines connect only tested values. The violet rule is the certification floor. At low visibility pair use is zero, so the 111–000 margin cannot arise from consumed pairs.

Appendix J independently checks the same distinction in a published rendezvous game: the Bell policy is an analytical result, whereas finite pair availability, visibility, certification, and cost determine whether the simulated system uses it.

E FULL STRUCTURAL AND CONFIGURED-REGIME RESULTS

Table 6: **Complete structural grid.** Paired 111 minus 000 scenario-utility contrasts with 95% simulation-seed bootstrap intervals ($n = 30$); the selected hard cell is primary and the remaining cells are descriptive.

Information	Payoff	Communication	Paired difference
pooled	separable	1-bit	**+0.009 [+0.002, +0.016]**
pooled	separable	no message	+0.007 [-0.000, +0.014]
pooled	joint	1-bit	**-0.021 [-0.029, -0.014]**
pooled	joint	no message	**-0.028 [-0.039, -0.018]**
local	separable	1-bit	**+0.036 [+0.028, +0.044]**
local	separable	no message	+0.006 [-0.001, +0.013]
local	joint	1-bit	+0.004 [-0.004, +0.012]
local	joint	no message	**+0.076 [+0.070, +0.083]**

The tables retain all eight structural cells and all eight configured regimes. They are separate experiments: the structural grid is evaluated across adapters, while the regime grid is evaluated in the selected local/joint/no-message cell. The structural table uses paired seed units and reports adapter-aggregated contrasts; the regime table gives adapter-specific means. The prior artifact contract assigns seed-bootstrap intervals and conditional contrasts to machine-readable 'summary.csv' and 'role_contrasts.csv' files, but those files are absent from this repository. In particular, the rounded means do not provide a paired 111–001 interval. No reported seed or adapter is removed from the displayed tables.

Table 7: **Complete configured-regime grid in the selected hard cell.** Entries are adapter-specific mean scenario utility per task. A/I/T operations follow Table 2; these means alone do not supply a paired 111–001 interval.

Regime (A / I / T)	Global Team	Claim Verify	Public Response	Mean
A I T **(000)**	0.380	0.320	0.239	0.313
A I T (100)	0.466	0.422	0.322	0.404
A I T (010)	0.380	0.330	0.246	0.319
A I T (001)	0.417	0.421	0.339	0.392
A I T (110)	0.372	0.318	0.237	0.309
A I T (101)	0.437	0.441	0.362	0.413
A I T (011)	0.416	0.425	0.341	0.394
A I T **(111)**	0.415	0.421	0.333	0.389

F Reproducibility and configuration

The anonymous artifact is available at https://anonymous.4open.science/r/Q-CoTrust-11CE/README.md. The checks below remain pre-submission verification requirements; linking the artifact does not by itself establish that every reported output reproduces from a clean extraction.

Recorded reference run. The prior experimental record specifies Python 3.12 with NumPy, pandas, SciPy, and Matplotlib. It reports an x86-64 Linux CPU run with no GPU, 1,277 seconds (21.3 minutes) for the main suites, and no instrumented peak-memory value. Exogenous episodes are keyed by study seed, episode, adapter, and stress condition; each method uses a separate deterministic policy stream while receiving the same exogenous trajectory within a paired unit. The submission artifact must record exact package versions, wall time, row counts, and checksum manifests. These runtime values are reproducibility metadata, not hardware-performance evidence.

Core illustrative settings. Training/evaluation/horizon are $120/80/12$. The tabular MAPPO actor, critic, discount, clipping, and epoch settings are $.055/.10/.96/.20/3$. Pair generation probability is .58, memory decay is .985, maximum stock is three, and the default visibility center is .90. Normalized costs are .08 per human, .05 per AI, .14 per message, .018 per analytical quantum call, .055 per consumed pair, .012 per certification, .018 per audit, and .95 per institutional review. Review succeeds with probability .96 on its eligible high-risk tasks; eligibility additionally requires value at least 1.5. These are controlled scenario parameters, not measured deployment costs.

Required commands. The separate anonymous artifact must expose paper-critical operations for unit tests, a fast smoke run, checksum verification, full figure regeneration, and the robot check. 'make robot-pilot' is the recorded command for regenerating the exact protocol table, 30-seed finite-stock outputs, and macros summarized in Figure 6. Each seed evaluates 20,000 missions at each of five visibility centers; only pair generation, visibility, decay, and inventory are sampled, while rendezvous success is integrated exactly. A full experiment rerun should write to a new directory by default. The Overleaf project contains PDF figures and editable SVG masters for portability, but not the executable benchmark.

Traceability and current boundary. The recorded artifact contract uses 'result_index.json' for the structural grid, configured-regime study, and sensitivity tests, and source-data mappings for each quantitative figure. Paper macros and tables must be regenerated by scripts; conceptual figures carry a separate provenance manifest. The prior record states that seventeen tests cover the main benchmark and rendezvous protocol. The executable files and raw seed-level outputs are not present in this Overleaf repository, so those claims, the paired 111–001 contrast, the bootstrap details, and full baseline semantics require verification from a clean extraction of the separate artifact before submission. The same artifact must map outcomes across major cost and capacity ranges before the paper can claim a robust adoption boundary. It must contain no Git history, user path, author identity, external credential, or identifying metadata.

G APPLICATION TRANSLATION MATRIX

Table 8: **How the common interface translates beyond quantum-resource allocation.** The robot row is evaluated only as the stylized pilot in Appendix J; all rows require external validation and none is a deployment.

Domain	Private / dynamic state	Candidate coordination substrate	Required external evidence
Global human–AI organizations	local capability, jurisdiction, workload, audit history	classical messaging; analytical policy; certified service	work and review traces; accountability study
Human-supervised robot teams	private positions, mission deadline, pair quality, radio availability	classical fallback; pre-shared pair under no-message deadline; post-mission audit	physical motion, sensing, radio, QPU/network, safety, and human-oversight study
Federated science / health	protected local data, expertise, evolving site quality	secure aggregation; selective specialist review	governance approval; privacy and clinical/scientific validity
Quantum cloud markets	queue, price, device fidelity, service reliability	remote QPU or pair inventory	measured time, energy, price, and hardware error
Distributed verification	claims, evidence, stake, disputes, strategic participation	classical voting; cryptographic or quantum-safe channel; optional coordinator	public traces plus an externally verified truth subset
Public-interest response	local reports, urgency, source reliability, infrastructure disruption	classical radio; protected link; expert escalation	prospective safety evaluation with affected communities

Public distributed-verification traces could calibrate assertion timing, disputes, participation, rewards, penalties, and review capacity, while counterfactual agents probe rare attacks. This remains an external-validation option rather than evidence in the present paper. Likewise, the K_3 pilot validates a coordination structure, not a robot platform. A protocol's final vote is not automatically objective ground truth, and post-quantum signatures are not physical entanglement.

H THREE-LEVEL FUTURE RESEARCH PROGRAM

Table 9 converts the present evidence limits into cumulative tests. Each level has a distinct unit of evidence and a gate on what may responsibly be inferred. All entries are proposed research, not completed experiments.

Table 9: **Cumulative research program.** Advancement requires the evidence gate in the preceding level; a favorable simulation result does not substitute for artifact, system, or institutional validation.

Level and current boundary	Testable program and evidentiary design	Intellectual contribution, practical value, and gate
1. Mechanism validity Available materials lack the executable benchmark, raw paired rows, tests, lock file, and 111–001 interval.	**Question:** Do the analytical law, information boundaries, resource ledger, and uncertainty procedure reproduce independently? **Design:** release licensed code and data; test probability normalization, no-signalling marginals, audit timing, resource conservation, weighted classical optima, and deterministic seeds; regenerate every table and figure; compute the paired 111–001 contrast; ask a clean-room user to add one method through the public adapter. **Unit:** seed–configuration rows plus exact game states. **Falsifier:** numerical mismatch, prohibited information flow, non-reproducible uncertainty, or evaluator modification required for reuse.	Establishes whether Q-COTRUST is a reusable diagnostic rather than a manuscript-only specification. **Practical value:** auditable mechanism comparison. **Gate:** no reusable-benchmark or incremental quantum-value claim until this level passes.
2. System economics Current evidence fixes major prices and capacities and models visibility rather than measuring a network.	**Question:** Over which price, quality, delay, and capacity regions does a quantum route dominate its strongest feasible classical substitute? **Design:** prespecify response surfaces over pair, message, certification, audit, and review costs and capacities; retrain when policies change; use equal-budget classical comparators; replace scalar pair dynamics with measured latency, loss, fidelity, energy, and availability traces. **Unit:** held-out operating regime and measured system trace. **Falsifier:** no robust region beyond tuning, or dominance by communication, review, or a stronger classical policy.	Identifies the binding scarcity and converts a single cost vector into conditional price–capacity frontiers. **Practical value:** engineering trade-off map, not procurement advice. **Gate:** no hardware, financial-return, or operational-value claim without measured system inputs.
3. Institutional and field translation Current agents, workers, tasks, prices, and robots are simulated; technical controls do not establish legitimacy.	**Question:** Does the boundary survive repeated multi-site operation, strategic behavior, human workload, and domain-specific rights? **Design:** evaluate repeated tasks with motion, sensing, radio, networks, and human supervision where relevant; co-define outcomes with affected stakeholders; measure error incidence, delay, workload, distribution of costs, consent, contestability, and appeal; preregister stopping and safety rules. **Unit:** organization, site, mission, and affected stakeholder group. **Falsifier:** technical gains disappear, shift burdens inequitably, or fail safety, consent, or accountability requirements.	Tests whether a computational boundary becomes institutional value rather than merely moving cost or risk. **Practical value:** evidence for bounded policy deliberation. **Gate:** no deployment, welfare, or public-policy recommendation before domain governance and external validation.

I Responsible-Use Detail

No human participants, personal data, deployed AI outputs, blockchain addresses, or physical quantum devices are used. Consequently, the study does not estimate the behavior of real people, institutions, markets, or emergency systems. Synthetic labels such as "human," "AI," "compromised," and "trust" are functional variables and should not be used to rank populations or automate consequential decisions.

The benchmark could be misused to justify quantum procurement or to give a technical veneer to institutional trust. We mitigate this by pricing modeled resources, retaining classical alterna-

tives, distinguishing configured policy/resource/control regimes, exposing negative regions, limiting escalation, and stating that social legitimacy is outside scope. Any field translation requires domain-specific consent, security analysis, audit rights, redress, and measurement of who bears false-positive, false-negative, delay, and infrastructure costs.

The main-text AI use statement discloses assistance with research ideation and critique, literature discovery and organization, early code drafting, drafting and editing, title and keyword development, formatting, and vector-figure design. AI tools did not generate the reported synthetic observations or serve as evidence. Numerical values in this venue revision were preserved from the study's generated macros and tables; a clean code-and-data rerun remains a separate pre-submission release requirement. The authors remain responsible for the research design, interpretation, claims, and final submission.

J ROBOT-RENDEZVOUS PILOT: PROTOCOL AND EXACT CHECKS

Game and role boundary. We instantiate the one-step labelled-K_3 rendezvous game of Tucker et al. (2024), using their no-wait ($W = 0$), no-edge-meeting ($E = 0$), check-later ($S = 1$) convention. Two agents start independently and uniformly on vertices $x, y \in \{0, 1, 2\}$, including the same vertex, and each observes only its own location. Each must move to one of its two neighbours; success is checked after both moves. There is no decision-time message. The human-supervision layer fixes mission value, risk, admissible resource quality, and fallback before deployment and audits the outcome afterward; it does not supply either robot's missing location during the move.

Exact classical controls. At each vertex, action 0 or 1 selects the lower- or higher-labelled neighbour. We enumerate all $2^3 \times 2^3$ ordered pairs of deterministic local policies, including asymmetric policies, and obtain

$$p_C^* = \frac{5}{9} = 0.5556. \tag{6}$$

We also enumerate every deterministic protocol in which agent A sends one bit before the actions, and every simultaneous protocol with one bit in each direction. Their optima are respectively $2/3 = 0.6667$ and $7/9 = 0.7778$. If both labelled locations are shared, the agents can always select their common remaining neighbour and succeed with probability one. These controls distinguish an advantage under zero messages from a claim of superiority to communication.

Bell strategy and visibility. The agents pre-share $|\Phi^+\rangle$, use local measurement angles $(0, \pi/3, 2\pi/3)$ indexed by their own vertices, measure, and map the output bit to a neighbour. Using a Werner-inspired scalar visibility model (Werner, 1989), local marginals remain uniform and exact enumeration gives

$$p_Q(\nu) = \frac{1}{3} + \frac{\nu}{4}. \tag{7}$$

Thus $p_Q(1) = 7/12 = 0.5833$ and $p_Q(\nu) > p_C^*$ iff $\nu > 8/9 = 0.889$. Neither local measurement reveals the other location; the correlation is useful only because the measurement rule and game were agreed before separation.

Finite-resource simulation. We reuse the main benchmark's pair generation, bounded inventory, memory decay, certification floor, and normalized charges. The pilot uses mission value $v = 2$, risk flag $r = 1$, and failure utility $-0.70v(1 + 0.5r)$, giving a success-to-failure utility span of 4.1. The analytical call, consumed pair, and certification charges total $.018 + .055 + .012 = .085$. A value-aware gate therefore uses a certified quoted pair only when

$$4.1 \left[p_Q(\nu) - \frac{5}{9} \right] > .085, \tag{8}$$

or $\nu > 0.972$. Otherwise it executes the exact classical fallback. Across 30 deterministic seeds and 20,000 missions per visibility center, only the pair process is sampled; action success is scored in expectation. The separate anonymous artifact must include seed-level outputs, the exact protocol table, checksums, and the figure source mapping.

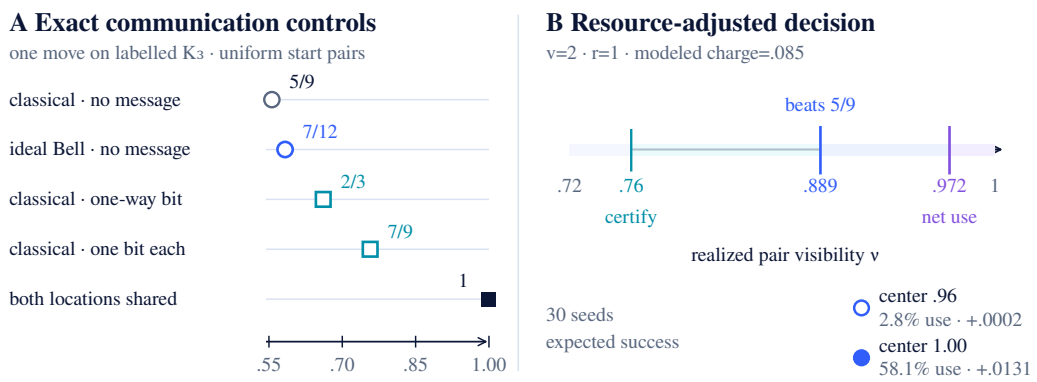


Figure 6: **Limited K_3 rendezvous check.** Panel A gives exact values under the named information policies. Panel B distinguishes the .76 certification floor, $8/9$ no-message advantage threshold, and .972 value-aware threshold, then reports two finite-stock pilot points. Pre-shared correlation sends no location; the check contains no motion, sensing, radio, QPU, or physical robot evidence.

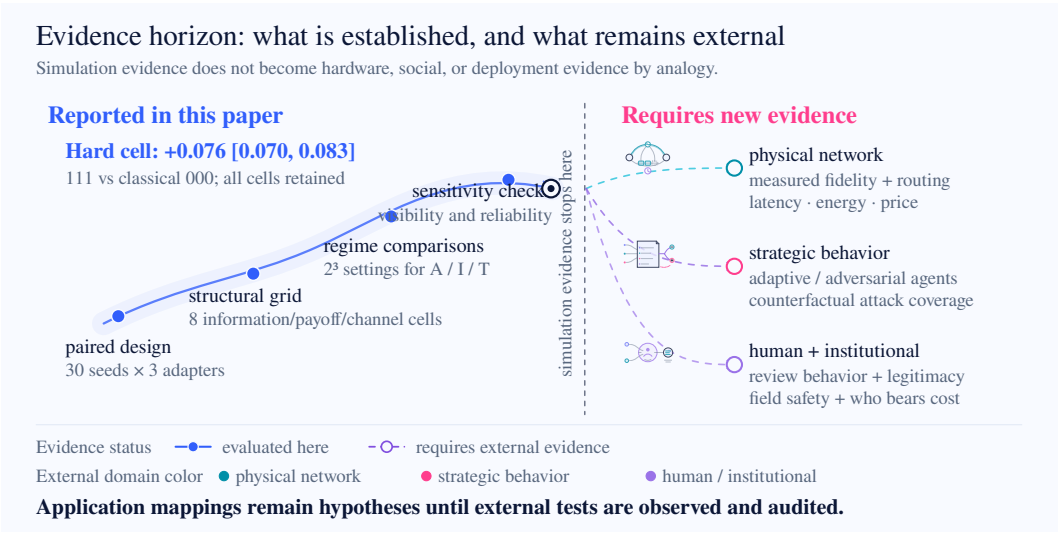


Figure 7: **Evidence horizon after the robot pilot.** The key maps solid/filled to evaluated simulation evidence and dashed/open to unevaluated tests; the vertical rule is the scope boundary. Teal, pink, and violet denote physical-network, strategic-agent, and human/institutional domains. The K_3 pilot adds a simulated external structure but does not cross the boundary into robots, hardware, or field evidence.