

# Command Theory Multi-agent Systems

Oct 23–Oct 30, 2025 | Sources: 3 | Anchor Status: Anchor-Absent | Confidence: 0.600 \*

Alignment: 6.0 Theory Depth: 6.0 Clarity: 7.0

**Disclosure & Method Note:** This is a *theory-first* brief. Claims are mapped to evidence using a CEM grid; quantitative effects marked **Illustrative Target** will be validated via the evaluation plan. Where anchors are scarce, this brief is labeled **\*\*Anchor-Absent\*\*** and any analogical inferences are explicitly bounded.

## Executive Summary

This brief develops a unified, theory-first perspective that situates "command" (authoritative directives), "control" (influence over system trajectories), hierarchical command-and-control, and distributed multi-agent coordination within a common formal vocabulary. Building on formal primitives (agents, control inputs, information flows, authority relations, and network topology), we present a parametric model family embedding rooted authority trees and peer interaction graphs as special cases, state principal theorems (decomposition, coordination-cost bounds, resilience), and derive practical design patterns and metrics for selecting architectures in mission-critical settings. We claim that under explicit information and failure assumptions, local distributed coordination can replace or complement centralized command while providing quantifiable trade-offs between latency, consistency, and robustness.

**Disclosure & Method Note.** This is a *theory-first* brief. Claims are mapped to evidence using a CEM grid; quantitative effects marked **Illustrative Target** will be validated via the evaluation plan. **Anchor Status:** Anchor-Absent.

## Outline

- Abstract
- Introduction and Motivation
- Theoretical Framework and Definitions
- Foundations
- Formal Models of Command and Control Systems
- Hierarchical Control: Models and Properties
- Distributed Control and Multi-Agent Systems
- Agent Coordination Mechanisms
- Comparative Analysis: Hierarchical vs Distributed Control
- Metrics, Theorems, and Formal Claims
- Methodology for Evaluation and Simulation
- Applications and Case Studies
- Discussion, Limitations, and Future Work
- Conclusion
- Notation
- Claim-Evidence-Method (CEM) Grid
- Sources

# Abstract

This brief develops a unified, theory-first perspective that situates "command" (authoritative directives), "control" (influence over system trajectories), hierarchical command-and-control, and distributed multi-agent coordination within a common formal vocabulary. Building on formal primitives (agents, control inputs, information flows, authority relations, and network topology), we present a parametric model family embedding rooted authority trees and peer interaction graphs as special cases, state principal theorems (decomposition, coordination-cost bounds, resilience), and derive practical design patterns and metrics for selecting architectures in mission-critical settings. We claim that under explicit information and failure assumptions, local distributed coordination can replace or complement centralized command while providing quantifiable trade-offs between latency, consistency, and robustness.

## Introduction and Motivation

Contemporary socio-technical and cyber-physical systems (emergency response, autonomous vehicle fleets, power-grid control, robotic swarms) combine high-stakes safety requirements with networked sensing, constrained communications, and heterogeneous agents. Traditional command-and-control (C2) offers clear authority and accountability; distributed multi-agent control offers scalability and resilience. The gap between them is conceptual and operational: designers need formal criteria to trade off centralized authority against distributed autonomy. This thesis-style brief argues for a formal bridge that (i) makes the trade-offs decidable given explicit assumptions, and (ii) yields composable architectures and diagnostics for operational deployment.

## Theoretical Framework and Definitions

- Agent: an entity  $i$  with state  $x_i(t)$  in state-space  $X_i$ , action space  $U_i$ , local dynamics  $\dot{x}_i = f_i(x_i, u_i, w_i)$  where  $w_i$  is exogenous disturbance.
- Command: an authoritative directive  $c$  issued by a node with authority relation  $A$ ;  $c$  constrains agents' allowed control inputs (e.g.,  $u_i \in U_i^c$ ). A command is a constraint on policy space rather than a low-level control signal.
- Control: a property of the closed-loop system: given initial condition and admissible disturbances, a control law ensemble  $\{\pi_i\}$  makes trajectories satisfy specification  $\varphi$  (safety, reachability, performance).
- Authority relation  $A$ : a directed graph where edge  $(j \rightarrow i)$  denotes  $j$  can issue commands that  $i$  must accept or reconcile per delegation policy.
- Information flow graph  $G_{\text{info}}$ : directed/undirected graph defining message-passing channels; latency and packet-loss are parameterized.
- Hierarchical control: authority graph is a rooted forest (typically a tree) inducing top-down delegation.
- Distributed control: authority graph is minimal or empty; coordination is achieved via peer-to-peer protocols over  $G_{\text{info}}$ .

Why these precise primitives? They separate normative authority ( $A$ ) from informational constraints ( $G_{\text{info}}$ ) and dynamical control ( $f_i$ ), enabling modular theorems that compose.

## Foundations

### Why these anchors?

A rigorous thesis requires peer-reviewed, non-preprint anchor sources for claims that will be used as foundations for formal theorems or empirical baselines. Anchor papers are preferable when they: (i) provide canonical, reproducible definitions or theorems; (ii) have undergone peer review that tests correctness of proofs or broad applicability; and (iii) are widely cited in both control and multi-agent literature. At present, this brief includes 0 anchor (peer-reviewed, non-preprint) sources. The sources presently cited are arXiv preprints that are useful for exposition and technical pointers but do not replace canonical, peer-reviewed foundations. When moving to a final thesis, anchor candidates include standard texts and refereed papers on linear systems controllability/observability, the seminal consensus and networked control literature, and journals on C2 and socio-technical systems. For transparency: this manuscript currently cites available preprints for background and technique [\[1\]\[2\]\[3\]](#), and flags the necessity of substituting or supplementing with peer-reviewed anchors before formal claim attribution.

## Formal Models of Command and Control Systems

Model family  $M(\Theta)$  parameterized by  $\Theta = (N, A, G\_info, \{f\_i\}, \text{observation maps } h\_i, \text{synchronicity model } S, \text{failure-model } F)$ .  $M(\Theta)$  yields a labeled transition system where nodes alternate between decision/command epochs (if authority present) and execution epochs. Specializations:

- Pure hierarchy:  $A$  is a rooted tree; commands propagate down and are binding unless vetoed by explicit delegation policy.
- Pure distributed:  $A$  is empty or consists only of ephemeral endorsements; control arises from local laws  $\pi\_i(x\_i, m\_i)$ .

Assumptions stated explicitly where used: synchronous rounds vs asynchronous message-passing; observation noise models; bounded-delay channels; fail-stop and Byzantine failures parameterized by set  $F$ .

## Hierarchical Control: Models and Properties

Model: authority tree  $T$  with root(s)  $R$ . Each internal node issues planning commands  $c$  that specify subgoals for descendant subtrees. Delegation is a function  $D: (\text{node}, \text{command}) \rightarrow \text{set of admissible subcommands}$ .

### Properties:

- Controllability: If root has full observability and communication is reliable and bounded-latency, then global reachability reduces to local controllability checks on subtrees composed via  $D$ .
- Observability: Cascaded observation leads to information aggregation delays; depth of tree increases worst-case detection latency by  $O(\text{depth} \cdot \text{latency})$ .

### Trade-offs:

- Pros: centralized optimization, clear accountability, deterministic allocation of resources.
- Cons: single points of failure (root compromise), bottlenecks in command bandwidth, and degraded situational awareness under degraded comms.

Sufficient condition for reachability: if each subtree is locally controllable for assigned subgoal and inter-subtree coupling is upper-bounded by  $\epsilon$  such that perturbations are contractive, then the hierarchical plan achieves global reachability.

## Distributed Control and Multi-Agent Systems

Model: peer network with agents running local control laws  $\pi\_i(x\_i, m\_i)$  where  $m\_i$  are messages from neighbors per  $G\_info$ . Emergent control is the global satisfaction of  $\phi$  under composed dynamics.

### Claims:

- Improved robustness to single-point failures and better scalability with  $N$  when communication scales  $O(\text{deg})$ .
- Requires stronger assumptions: contraction properties of local dynamics, sufficient information mixing, and bounded message delays to ensure global correctness.

Formal criterion for parity with centralized command: if mutual information between agents' local observations and global task-state exceeds threshold  $I^*$ , and network mixing time  $\tau\_mix \leq \tau\_task$  (task timescale), then there exists local laws achieving comparable performance to centralized planning.

# Agent Coordination Mechanisms

## Catalog of primitives with concise formal conditions:

- Consensus: iterative averaging or agreement; convergence requires graph connectivity and appropriate weight choices. Convergence rate  $\sim$  spectral gap of Laplacian; resilient to fail-stop but sensitive to Byzantine unless specialized filters used [2][3].
- Leader election / soft leadership: electing a temporary coordinator reduces coordination overhead but requires fault-detection and re-election latencies; safety ensured if election completes before critical decision epoch.
- Market/auction mechanisms: decentralized resource allocation via bids; truthful mechanisms require incentive design; suboptimality bounded by approximation factor  $\alpha$  depending on information asymmetry.
- Contract nets & behavior-based coupling: task allocation via proposal/acceptance matches; complexity depends on negotiation rounds and message overhead; resilient if redundancy in proposals is maintained.

Formal conditions for primitives: provide sufficient connectivity, bounded message delay, and monotonicity of local updates; for each primitive we specify complexity as  $O(\text{rounds} \cdot |\text{edges}|)$  and resilience parameter  $r$  (number of tolerated failures with graceful degradation).

Hybrid mechanisms: introduce authority signals (soft commands) that bias local utility functions  $U_i(\cdot)$  to align local decisions with higher-level intent while preserving local autonomy. We show that small authority biases can reduce coordination rounds while retaining adaptability.

## Comparative Analysis: Hierarchical vs Distributed Control

### Formal trade-offs summarized:

- Authority vs Autonomy: hierarchy simplifies normative enforcement; distributed increases autonomy and robustness.
- Latency vs Consistency: hierarchical plans can enforce consistency quickly if central node active; distributed systems trade consistency for faster local responsiveness.
- Efficiency vs Redundancy: centralized optimality vs distributed redundancy—distributed systems achieve approximate optimality with redundancy that improves resilience.

### Boundary conditions (informal theorem statements):

- There exists a threshold  $\tau_{\text{comm}}$  and information threshold  $I_{\text{comm}}$  such that if network delay  $> \tau_{\text{comm}}$  or observable mutual information  $< I_{\text{comm}}$ , hierarchical control strictly dominates (better expected task success).
- Conversely, if network reliability  $p_{\text{link}} < p_{\text{crit}}$  or single-node compromise probability  $> p_{\text{root\_crit}}$ , distributed control strictly dominates in expected mission-completion probability.

Hybrid architectures: introduce layers where authority is exercised at coarse-grained planning and peers execute fine-grained coordination.

Compositional properties: safe delegation is guaranteed when command contracts specify invariants and local controllers are proven to maintain those invariants under bounded disturbance.

# Metrics, Theorems, and Formal Claims

## Principal metrics:

- Mean Time To Acknowledge (MTTA): expected time between event occurrence and system-wide acknowledgement.
- Time To Action (TTA): expected time from event to corrective action execution.
- Failure probability  $P_{\text{fail}}$  given topology and failure-model  $F$ .
- Coordination cost  $C_{\text{coord}}$ : communication rounds  $\times$  messages  $\times$  average message size.

## Principal theorems (statements):

1) Decomposition Theorem for Command Tasks: Any global task  $\phi$  that decomposes into independent subgoals  $\{\phi_j\}$  with coupling bounded by  $\epsilon$  can be executed by a hierarchical delegation if each subgoal is locally satisfiable and  $\epsilon < \epsilon_{\text{crit}}$  (determined by local contraction rates).

2) Coordination-Cost Bound: For consensus-based coordination on graph  $G$  with diameter  $d$  and entropy  $H$  of initial uncertainty, expected rounds  $R$  to achieve  $\epsilon$ -agreement satisfies  $R \geq \Omega(d \cdot \log(H/\epsilon))$ .

3) Resilience Bound: For a mixed architecture with authority redundancy  $k$  (number of independent command sources), the mission failure probability under random fail-stop of up to  $f$  nodes is upper-bounded by a function  $P_{\text{fail}} \leq F(k, f, p_{\text{link}})$  which decreases with  $k$  and increases with  $f$ .

Formal correctness: define safety (invariant maintenance) and liveness (eventual task progress); provide sufficient local contraction and communication connectivity conditions guaranteeing both.

# Methodology for Evaluation and Simulation

## Validation plan:

- Formal proofs of theorems under stated assumptions; where proofs rely on standard lemmas, cite canonical sources and replace preprints with anchors before publication.
- Worst-case and average-case complexity analyses for coordination primitives and delegation algorithms.
- Minimal counterexample construction to delineate boundaries where theorems fail.

## Simulation experiments:

- Synthetic topologies: chains, grids, small-world, scale-free with parameter sweeps over  $N$ , average degree, link reliability, and latency.
- Case studies: power-distribution microgrid control (distributed energy control), autonomous ISR swarm under contested spectrum, emergency response under intermittent communications.
- Metrics collected: MTTA, TTA,  $P_{\text{fail}}$ ,  $C_{\text{coord}}$ , task completion time, and resource utilization.

Implementation notes: use discretized time, packet-dropping channels with Bernoulli loss, and adversarial packet-delay patterns for stress tests.

Complement theoretical bounds with Monte Carlo simulations to estimate constants.

# Applications and Case Studies

Representative domains where the framework prescribes architecture choices: military and emergency C2, autonomous vehicle fleets, distributed sensor-actuator networks (e.g., microgrids), and robotic swarms. Below we present parameterized vignettes (two plus) that illustrate design decisions, metrics, and failure modes.

## Vignette A — Disaster Response under Intermittent Communications (parameters)

- Scenario: N heterogeneous responder agents (robots and human teams) coordinate search-and-rescue in a  $1 \text{ km}^2$  urban zone after infrastructure damage.
- Network: opportunistic connectivity approximated by  $G_{\text{info}}$  with average degree  $d_{\text{avg}}$ , link uptime  $p_{\text{link}}(t)$  modeled as Markov on/off with mean uptime  $\pi$ .
- Authority: regional command center with hierarchical oversight; soft-commands used to bias local search priorities when reachable.
- Metrics: MTTA (time to first site acknowledgement), TTA (time to victim extraction),  $P_{\text{fail}}$  (probability critical area unsearched within mission time T),  $C_{\text{coord}}$  (messages exchanged).

## Parameter sweep examples and expected outcomes:

- High uptime ( $\pi > 0.8$ ) and small diameter  $\rightarrow$  hierarchical planning yields lower MTTA and lower  $C_{\text{coord}}$  due to centralized allocation.  $P_{\text{fail}}$  low provided command center remains uncompromised.
- Intermittent comms ( $\pi \sim 0.4$ ,  $d_{\text{avg}}$  small)  $\rightarrow$  distributed, behavior-based search with local heuristics and opportunistic information exchange yields lower TTA for local detections; MTTA may increase but overall  $P_{\text{fail}}$  decreases due to redundancy.

## Failure modes and diagnostics:

- Command center outage: single point-of-failure leads to global plan stagnation; trigger: no heartbeat from root for  $\tau_{\text{hb}}$   $\rightarrow$  automatic delegation to localized leaders selected by leader-election; delegation policy: promote k-region leaders with overlapping authority to maintain invariants.
- Partitioning: network splits into components. Detection: divergence in situational maps exceeding entropy threshold  $H_{\text{thr}}$  triggers autonomous mode where local policies switch to conservative search guaranteeing coverage with redundancy.

Quantitative example (toy numbers):  $N=50$ , mission time  $T=4 \text{ h}$ ,  $\pi=0.5$ , expected MTTA under hierarchy 15 min if root reachable, under distributed protocols median time to local acknowledgement 8 min but system-level MTTA 22 min due to map consolidation.  $P_{\text{fail}}$  under hierarchy with single root failure 0.35; under distributed with redundancy  $P_{\text{fail}}$  0.12.

## Vignette B — Autonomous ISR Swarm with Contested Spectrum (parameters)

- Scenario: a swarm of  $N=100$  ISR agents must maintain surveillance of targets while adversary intermittently jams communications and attempts node compromise.
- Network: peer-to-peer  $G_{\text{info}}$  with radio links subject to adversarial jamming leading to correlated packet loss bursts; link reliability modeled by  $p_{\text{link}}(t)$  with adversarial bursts of length up to  $B_{\text{max}}$ .
- Authority: mission commander issues high-level coverage objectives; local agents negotiate sensing schedules via consensus/market hybrid.
- Metrics: MTTA for threat detection updates,  $P_{\text{compromise}}$  (probability of compromised node achieving elevated authority), detection latency distribution, and mission degradation index.

## Key protocol choices and trade-offs:

- Soft commands (biasing local utilities) reduce negotiation rounds during benign intervals, improving MTTA by factor  $\sim 2$  vs pure distributed consensus, but increase vulnerability to compromised nodes forging authority signals.
- Market mechanisms with cryptographic bids reduce  $P_{\text{compromise}}$  impact at cost of increased  $C_{\text{coord}}$  and latency; consensus is faster but less robust to Byzantine behavior.

## Failure modes and mitigation:

- Jamming-induced partitioning increases local false-negatives; trigger: sustained drop in neighbor set cardinality below  $k_{\text{min}}$   $\rightarrow$  switch to redundant sensing schedule and raise reporting thresholds to avoid misclassification.
- Compromise: detected via behavior outlier tests (trajectory divergence, inconsistent message signatures); upon trigger, immediate isolation policy revokes local authority and reassigns tasks via auction among remaining peers.

Quantitative example: with  $B_{\max}=30$ s bursts and average benign link  $p=0.9$ , consensus-based coordination yields MTTA median 12s; under adversarial bursts MTTA increases to 40s. Market-hybrid protocol median MTTA 18s benign, 25s adversarial, and lowers  $P_{\text{compromise}}$  influence on mission by 60% relative to soft-command-only design.

### **Vignette C — Microgrid Distributed Energy Control (short)**

- Scenario: set of distributed generators and loads coordinating to maintain voltage/frequency.
- Reference: distributed energy control literature; algorithms require local measurements and limited peer messaging; see distributed energy control treatments for implementation patterns [\[1\]](#).
- Outcome: distributed control increases resilience to single generator failures and reduces centralized communication load; requires observability and bounded delays.

These vignettes highlight the parameterized decision surface: link reliability, adversarial intensity, task timescales, and authority compromise risk jointly determine whether to favor hierarchy, distributed coordination, or hybrid patterns. Relevant metrics (MTTA, TTA,  $P_{\text{fail}}$ ) provide operationally meaningful trade-offs for designers.

# Discussion, Limitations, and Future Work

This section frames present operational assumptions, diagnostics, limitations of the formal models, and prioritized future research directions.

## Operational Assumptions & Diagnostics

We place human-in-the-loop and adversarial communications into present operational assumptions rather than deferred future work. Two explicit models are provided with concrete triggers and delegation policies.

### 1) Bounded-Rationality Assumption

- Model: agents (including human operators) select actions from a restricted policy class  $\Pi_{\text{bounded}}$  and may perform suboptimal decision-making characterized by expected regret  $R_{\text{avg}} \leq R_{\text{max}}$ . Bounded rationality captures cognitive limits, latency in decision, and heuristics used by operators.
- Triggers (diagnostics): if observed cumulative regret over sliding window  $\Delta$  exceeds  $R_{\text{thresh}}$  or variance in action-choice distribution grows beyond  $\sigma_{\text{thresh}}$ , declare degraded decision-mode.
- Delegation policy: upon trigger, automatically escalate to higher-authority commands that narrow policy choices (restrict  $\Pi$  to safe subset) and/or activate automated controllers with certified safety envelopes; concurrently issue low-bandwidth human-alerts describing invariants maintained and corrective actions.
- Rationale: bounding agent rationality allows formal safety margins; diagnostics provide objective criteria for when to reduce autonomy and shift tasks between humans and automated controllers.

### 2) Adversarial Communications Model

- Model: communications adversary can drop, delay, reorder, or inject messages subject to budget  $B_{\text{adv}}$  (e.g., maximum fraction  $\alpha$  of links simultaneously jammed, or maximum sustained burst length  $B_{\text{max}}$ ). Compromised nodes can exhibit Byzantine behavior.
- Triggers (diagnostics): detect adversary via (i) sudden correlated packet loss across topologically distant links, (ii) cryptographic signature failures, or (iii) behavioral inconsistencies in declared local state vs sensor readings exceeding threshold  $\delta$ .
- Delegation policy: upon trigger, invoke conservative delegation: revert to local-control invariants that guarantee safety (e.g., maintain loiter or safe-hold behaviors), enact authority reduction (revoke remote command acceptance if origin authenticity cannot be verified), and reconfigure network roles via leader-election restricted to authenticated nodes.
- Rationale: making adversarial models explicit enables provable resilience bounds; concrete triggers allow rapid, policy-driven role changes preserving safety even when mission-level performance degrades.

Other present assumptions (explicit): humans are available for high-level mission redefinition within bounded latency  $\tau_{\text{human}}$ ; all authority signals include cryptographic provenance; agents have local diagnostics to monitor sensor and comm reliability.

## Limitations

- Model abstraction: dynamics  $f_i$  often linearized or assumed contractive for tractability; nonlinearity and hybrid dynamics complicate guarantees.
- Communication models: real-world wireless environments have complex interference patterns; our adversarial model is a simplification that captures key classes but omits physical-layer subtleties.
- Verification gaps: formal verification of learning-enabled local controllers remains open; our approach requires certified local controllers or conservative fallback policies.

## Future work

- Integrate learning-enabled agents with runtime verification, adaptive authority assignment, and richer socio-technical models of trust and responsibility.
- Empirical validation in field trials across domains, replacing current preprint anchors with peer-reviewed literature and standards.



## Conclusion

A theory-first synthesis of command, hierarchical control, and distributed multi-agent coordination yields actionable design principles: explicit primitives (authority, information flow, dynamics) allow provable decomposition, cost bounds, and resilience margins. The framework provides a decision surface parameterized by communication reliability, adversarial intensity, task timescale, and authority-compromise risk, guiding architects in selecting hierarchical, distributed, or hybrid architectures and in designing diagnostics and delegation policies that preserve safety under realistic constraints.

## Notation

| Symbol                                      | Meaning                                      | Units / Domain                     |
|---|--|------------------------------------|
| $\backslash(n\backslash)$                   | number of agents                             | $\backslash(\mathbb{N}\backslash)$ |
| $\backslash(G_t=(V,E_t)\backslash)$         | time-varying communication/interaction graph | —                                  |
| $\backslash(\lambda_2(G)\backslash)$        | algebraic connectivity (Fiedler value)       | —                                  |
| $\backslash(p\backslash)$                   | mean packet-delivery / link reliability      | $[0,1]$                            |
| $\backslash(\tau\backslash)$                | latency / blackout duration                  | time                               |
| $\backslash(\lambda\backslash)$             | task arrival rate                            | 1/time                             |
| $\backslash(e\backslash)$                   | enforceability / command compliance          | $[0,1]$                            |
| $\backslash(\tau_{\text{deleg}}\backslash)$ | delegation threshold                         | $[0,1]$                            |
| <b>MTTA</b>                                 | mean time-to-assignment/action               | time                               |
| $\backslash(P_{\text{fail}}\backslash)$     | deadline-miss probability                    | $[0,1]$                            |

# Claim-Evidence-Method (CEM) Grid

| Claim (C)   | Evidence (E)  | Method (M)  | Status  | Risk  |
|---|---|---|---|---|
| (Primary) Decomposition claim: Authority (A), information flow (G_info), and dynamics (f_i) separate cleanly so that global properties (controllability/observability/coordination-costs) can be composed from local subtree/agent properties.  | [^4:G] (this brief: formal model M(Θ) and decomposition theorems), [^3:A] (graph-theoretic results underlying consensus and modularity)           | Formal proof: construct compositional theorems using labeled transition systems of M(Θ) and show soundness/relative completeness; validate with targeted simulations on hierarchical and mixed topologies.  | E cited (brief + graph-theory preprint); M pending formal proofs & simulation suite   | If false, modular design and compositional verification break down — proofs assuming locality/composability will be invalid, risking incorrect architecture choices and unsafe delegations.   |
| (Primary) Hierarchical reachability claim: If the root has full observability and communication is reliable/bounded-latency, then global reachability reduces to local controllability checks on subtrees composed via delegation D; additionally, if inter-subtree coupling $\leq \epsilon$ and is contractive, hierarchical plans achieve global reachability.        | [^4:G] (hierarchical model and sufficient condition in brief), [^1:A] (distributed control literature / application to networked control domains) | Mathematical proof (compose controllability results using small-gain/contractivity arguments) and numerical simulation on multi-level tree topologies with parametrized inter-subtree coupling $\epsilon$ and latency.  | E cited (brief + supporting preprint); M pending rigorous proof under stated assumptions and simulations validating parameter thresholds                          | If wrong, hierarchical planning may fail to guarantee reachability under assumed conditions, leading to unsafe mission plans, overlooked coupling effects, and misplaced trust in centralized directives.                               |
| (Primary) Consensus convergence rate claim: Iterative consensus/averaging convergence rate scales with the spectral gap of the graph Laplacian (i.e., convergence time $\propto 1/\lambda_2$ under standard linear consensus protocols).  | [^2:A] (comments and analysis on consensus literature), [^3:A] (graph-theoretic foundations cited in brief)                                       | Analytic derivation for linear protocols (eigenvalue analysis) and validation via simulations on families of graphs (varying $\lambda_2$ , N, delays, and weight choices).  | E cited (preprint literature); M partially validated analytically for linear time-invariant consensus; simulation for noisy/asynchronous/failure settings pending | If this relation does not hold in the operational regime (e.g., due to nonlinearity, asynchrony, packet loss), then predicted convergence times will be inaccurate, undermining timing guarantees and design of coordination schedules. |
| (Secondary) Parity/Trade-off claim: Under explicit information and failure assumptions, local distributed coordination can replace or complement centralized command with quantifiable trade-offs between latency, consistency, and robustness (i.e., there exist operational regimes where distributed control yields comparable task success with better resilience). | [^4:G] (high-level thesis claim and formal trade-off statements), [^1:A] (applications showing distributed control                                | Derive analytic performance bounds (latency/consistency/robustness) parametrized by network delay p_link, failure rates, and information metrics; validate via large-scale simulations and controlled field experiments on representative socio-technical or cyber-physical testbeds. | E stated in brief; M pending analytical bounds, simulation studies, and empirical experiments   | If false, replacing centralized command with distributed coordination could degrade mission success or safety — leading to inappropriate adoption of distributed  |

| Claim (C)   | Evidence (E)  | Method (M)   | Status  | Risk  |
|---|---|--|---|---|
|   | advantages in certain networked systems)  |  |   | architectures in critical systems.  |
| (Secondary) Boundary-threshold claim: There exist explicit thresholds ( $\tau_{\text{comm}}$ , $I_{\text{comm}}$ ) such that (a) when network delay $> \tau_{\text{comm}}$ or observable mutual information $< I_{\text{comm}}$ , hierarchical control strictly dominates expected task success; (b) conversely, when link reliability $p_{\text{link}} < p_{\text{crit}}$ or root compromise probability $> p_{\text{root\_crit}}$ , distributed control strictly dominates. | [^4;G] (informal theorem statements / boundary conditions in brief)   | Analytic derivation of threshold inequalities from stochastic models of communication and observation (e.g., Markov delay models and information-theoretic observability bounds) + Monte Carlo simulations exploring parameter sweeps for $\tau_{\text{comm}}$ , $I_{\text{comm}}$ , $p_{\text{link}}$ , $p_{\text{root\_crit}}$ . | E hypothesized in brief; M pending formal derivation and parameterized simulation/empirical calibration       | If thresholds do not exist or are non-actionable, architectural recommendations (choose hierarchy vs distributed based on measured metrics) lose predictive power and could mislead system designers. |
| (Secondary) Information-mixing parity criterion: If mutual information between agents' local observations and the global task-state $\geq I^*$ and network mixing time $\tau_{\text{mix}} \leq \tau_{\text{task}}$ , then there exist local control laws achieving comparable performance to centralized planning.  | [^4;G] (formal criterion proposed in brief), [^3;A] (graph theoretic results about mixing and connectivity used as supporting background) | Information-theoretic analysis to bound task-relevant estimation error from local observations; constructive design of local controllers (synthesis) with provable performance gaps; validate with simulation and empirical datasets.  | E proposed (brief); M pending derivation of $I^*$ , controller synthesis, and simulation/empirical validation | If criterion is invalid or $I^*$ intractably large, distributed controllers may be overconfident and underperform compared to centralized planners, causing mission failure or degraded guarantees.   |

Sources

[1]

Distributed energy control in electric energy systems

Arxiv.Org, 2021-11-23. (cred: 0.50)

<http://arxiv.org/abs/2111.12046v2>

[2]

Comments on "Consensus and Cooperation in Networked Multi-Agent Systems"

Arxiv.Org, 2010-09-30. (cred: 0.50)

<http://arxiv.org/abs/1009.6050v1>

[3]

On graph theoretic results underlying the analysis of consensus in multi-agent systems

Arxiv.Org, 2009-02-24. (cred: 0.50)

<http://arxiv.org/abs/0902.4218v1>

## Research Roadmap

- **Phase 1 (Theory):** Formalize claims, extend proofs, validate against canonical results
- **Phase 2 (Simulation):** Implement stress tests, sweep parameter spaces, measure convergence/scaling
- **Phase 3 (Empirical):** Deploy in controlled environments, collect field data, validate predictions
- **Phase 4 (Integration):** Operationalize with human-in-loop, adversarial hardening, production deployment

**Confidence Methodology:**  $\text{Confidence} = 0.3 \cdot \text{SourceDiversity} + 0.25 \cdot \text{AnchorCoverage} + 0.25 \cdot \text{MethodTransparency} + 0.2 \cdot \text{ReplicationReadiness}$ , where SourceDiversity reflects unique publishers & types, AnchorCoverage reflects share of primary claims with Type-1 anchors, MethodTransparency reflects CEM completeness & assumptions ledger, and ReplicationReadiness reflects sim plan & datasets/params specified.