Information Homeostasis as a Nash Equilibrium: A Potential Game for Decentralized Multi-Agent Cave Exploration

Avi Soval[1], Keenan Albee[2]

***Abstract*— The coordination of multi-robot teams in communication-constrained subterranean environments is difficult without persistent centralized control or a shared global map. We suggest a decentralized approach inspired by models of biological collective intelligence, and introduce a scalar stress signal - which captures local disagreement in belief states between agents - that can be used to guide the team to an information-rich state. We formulate a multi-agent cave exploration mission as a potential game with time-varying frontier resources and agent utility functions that target information gain, congestion reduction, and belief consensus through this stress coupling. By the convergence properties of standard potential games, we hypothesize that evolution of the system purely via best response dynamics may approach a Nash equilibria that resembles an 'information homeostasis' - where disagreement and uncertainty are uniformly low while coverage is still satisfied. We have currently implemented a base-uncoupled game without the belief consensus that admits a canonical potential game form, and will examine if the perturbed, coupled, information-focused game still possesses any NE convergence guarantees.**

I. INTRODUCTION

Multi-agent robotic exploration of caves, tunnels, and underground infrastructure supports a variety of applications, such as search and rescue, asset coverage and inspection, and even planetary science. From a communications standpoint, teams operating in such environments are often subject to intermittent connectivity, imperfect information exchange, and degradation in shared knowledge over time, in addition to typical limits on range, bandwidth, and latency.

Though groups of robots in these missions can offer significant advantages in performance and redundancy, they must be deployed with a coordination architecture designed with these communication constraints in mind. For example, NASA's upcoming Cooperative Autonomous Distributed Robotic Exploration (CADRE) [1] demonstration will send a team of small rovers to autonomously map a lunar region using a structured leader-follower and election protocol under communication and resource budgets. Similarly, the lessons learned from the DARPA Subterranean challenge [2] document several solutions to this connectivity problem, with many implemented on hardware in realistic environments.

In these extreme, comms-limited or denied settings, assumptions on information parity within the team often cannot be made, and the notion of a shared world model can quickly break down. This compromises even simple exploration-only tasks; missions with stricter requirements on information quality or persistent coverage are even more sensitive. In this work, we study methods that can manage information disparity across a team, while still ensuring that nominal exploration objectives are being satisfied. In this sense, maintaining a degree of belief consensus becomes as or more important than pure exploration itself, and success is naturally framed as a team-wide information condition.

To build a coordination protocol designed for these requirements, we draw from recent work on multi-agent belief-space planning, but also take inspiration from methods applied to the study of biological collective intelligence. In cells, scalar voltage signals propagate via gap junctions, allowing each cell to effectively broadcast its current error from some homeostatic setpoint [3]. High-level target states can be realized as dynamical attractors, and it has been demonstrated that sharing these 'stress' scalars is sufficient for aligning independent actors towards a shared goal, even in the absence of central planning [4].

We seek to adapt this minimal-signaling hypothesis to a multi-agent cooperative exploration setting by defining a scalar stress arising from local belief disagreement and embedding it, alongside information gain and congestion terms, into a multi-agent potential game. In doing so, we investigate if proper utility function and local rule design can lead to *information homeostasis* - a state of persistent belief consensus, and explore whether this goal design can be encoded at a primitive, per-agent information and utility layer, rather than through discrete task and role assignments.

II. RELATED WORK

A. Multi-Agent Potential Games

Congestion and potential games [5], [6] converge to Nash equilibria under pure strategy best response dynamics, at which point no player stands to gain by deviating from their current action. The potential game structure has been well extended to multi-agent systems: [7] uses it to formulate cooperative control problems, [8] looks specifically at networked coverage control (relevant in the context of persistent consensus), and [9] casts multi-agent trajectory planning as a distributed potential game. Critical for our work, [10] introduces a *potentialness* metric that quantifies the dissimilarity between an arbitrary game and a potential game, and predicts the convergence of any learning dynamics. [11] also explores distributed exploration using frontier assignments and behavioral entropy as agent utilities with NE convergence guarantees, but focuses primarily on heterogeneous robots; we differ in that we incorporate agent

---

[1] Department of Aerospace and Mechanical Engineering, University of Southern California, Los Angeles, CA 90089 `soval@usc.edu`

[2] Department of Astronautical Engineering and Department of Aerospace and Mechanical Engineering, University of Southern California, Los Angeles, CA 90089 `kalbee@usc.edu`

belief states as part of the potential, and explicitly identify information homeostasis as the equilibrium objective.

B. Communication-Limited and Belief Space Planning

Coordination architecture and bandwidth strongly constrain multi-robot teams [12]; large-scale systems in the SubT challenge such as NeBula [13] developed methods for belief-space planning under limited communication, while ACHORD [14] directly addressed intermittent connectivity and general planning under uncertainty. Formal planning approaches that incorporate confidence and belief space planning are also being actively developed, with [15] proposing a hybrid belief sharing model for reaching consensus-type objectives. [16] introduces MEF-Explore, a two-tiered communication framework paired with an entropic frontier assignment strategy. Our work is similar in the information exchange policies, but we specifically attempt to formalize the mission as a potential game, which the incentive layer (utilities) are explicitly part of. [17] seeks to manage exploration, consensus, and information acquisition jointly through distributed Gaussian belief propagation despite intermittent connectivity, and is conceptually close to this paper; however, their focus is more on distributed estimation via consensus, whereas we use the stress from lack of consensus to induce a decentralized policy.

III. PROBLEM FORMULATION

We formalize multi-agent exploration as a congestion game over a time-varying set of frontier clusters, and define a global objective for information homeostasis - the team of agents must simultaneously discover the environment and maintain a shared level of confidence in their beliefs. By designing agent utilities that reward exploration and uncertainty reduction, penalize congestion, and encourage belief consensus, we seek to create a potential game whose Nash equilibria we hypothesize to align with a state of optimal information homeostasis. The following III-A and III-B outline the problem setup, III-C formally defines the potential game structure, and III-D presents an interpretation of the predicted solution.

A. The Exploration Game

The environment is structured as a 4-connected binary occupancy grid $G \in \{0,1\}^{H\times W}$ with open set G_{open}. At each discrete timestep t, the joint system state includes:

- **Players:** $\mathcal{N} = \{1, \ldots, n\}$.
- **Agent States:** Position $\mathbf{p}_i \in \mathbb{R}^2$ with neighbor set $\mathcal{N}_i \subseteq \mathcal{N}$, determined by line of sight range
- **Agent Beliefs:** Each agent maintains two sets of beliefs over the grid: a standard navigation map for traversability and frontiers $M_i \in \{\text{unknown}, \text{free}, \text{obs}\}^{H\times W}$, and an *info* (uncertainty) map $I_i(r,c) \in [0,1]^{H\times W}$ that decays over time, where 1.0 denotes full confidence and 0.0 full ignorance
- **Resources:** frontier-cluster set $\mathcal{F}(t)$, computed per-agent but exists as a shared resource
- **Actions:** Possible actions $a_i \in \mathcal{A}_i(t)$ (search, stay, patrol); the joint profile is $\mathbf{a} = (a_1, \ldots, a_n)$.
- **Utilities:** State-dependent utility functions U_i of the form $U_i(a_i; \mathbf{a}_{-i}, B_i) = \alpha\, \mathrm{IG}_i(a_i) - \beta\, C_i(a_i; \mathbf{a}_{-i}) + \gamma\, \Delta\sigma_i(a_i)$ (balancing heuristics for info gain, congestion, and consensus)

B. Objective: Information Homeostasis

We define information homeostasis as the joint satisfaction of two conditions - map uncertainty, and belief consensus - over the entire communication graph $(\mathcal{N}, \mathcal{E})$. We use a belief field B_i (the inverse of I_i), so higher values correspond to more uncertainty and higher rewards for corrective action.

*a) **Residual Uncertainty**:* This upper-bounds the allowable average uncertainty per cell by ϵ_B:

$$\bar{B} \;=\; \frac{1}{n}\sum_{i=1}^{n} B_i \;\rightarrow\; \frac{1}{|G_{\text{open}}|}\sum_{(r,c)\in G_{\text{open}}} \bar{B}(r,c) \;\le\; \epsilon_B.$$

*b) **Belief Consensus**:* This upper-bounds the allowable disagreement in agent beliefs by ϵ_σ, and variation over the whole graph is measured directly:

$$\frac{1}{|\mathcal{E}|}\sum_{(i,j)\in\mathcal{E}} \|B_i - B_j\|_{1,\Omega_{ij}} \;\le\; \epsilon_\sigma$$

With Ω_{ij} as the shared known cells between two agents:

$$\|B_i - B_j\|_{1,\Omega_{ij}} := \frac{1}{|\Omega_{ij}|}\sum_{(r,c)\in\Omega_{ij}} |B_i(r,c) - B_j(r,c)|.$$

Neither condition is observable by any single agent; each knows only its' own B_i and learns of its' neighbors. A global potential function is built from the complete $\bar{B}$ for Nash equilibria analysis, but agents' best response dynamics are determined only by the information available to them.

C. Potential Game

Our choice of actions, resources, and utility functions allows us to formally cast our setup as a potential game. Let $n_f(\mathbf{a})$ denote the number of agents assigned to a frontier cluster f under action profile $\mathbf{a}$. For each resource f, $\pi_f(k)$ specifies the identity of the k-th agent assigned to f. This indexing is fixed, and we construct a global potential as:

$$\Phi(\mathbf{a}) = \sum_{f\in\mathcal{F}}\sum_{k=1}^{n_f(\mathbf{a})} \big(\alpha\, \mathrm{IG}_{\pi_f(k)}(f) - \beta\, k\big) + \gamma\, \Psi(\mathbf{B}),$$

The inner sum is a canonical Rosenthal potential construction, balancing information gain and congestion cost. Our formulation then includes the affine term $\Psi(\mathbf{B})$ to capture the belief layer:

$$\Psi(\mathbf{B}) \;=\; -\frac{1}{|\mathcal{E}|}\sum_{(i,j)\in\mathcal{E}} \|B_i - B_j\|_{1,\Omega_{ij}}$$

Thus, γ is the term that actually introduces coupling into our network. When $\gamma = 0$, the game reduces to a pure congestion game with player-specific rewards IG_i - agents explore in self-interest (still benefiting from each other's knowledge), and do not coordinate actions. In this case, Φ is an exact potential: payoffs map exactly to change in potential and so for any deviation $a_i \to a_i'$,

$$U_i(a_i'; \mathbf{a}_{-i}) - U_i(a_i; \mathbf{a}_{-i}) = \Phi(a_i', \mathbf{a}_{-i}) - \Phi(a_i, \mathbf{a}_{-i}).$$

Consequently, agent behaviors in this uncoupled state should satisfy the core property of any potential game: that any sequence of best-response dynamics strictly increases Φ and converges to a pure-strategy Nash equilibrium, at which point no player stands to gain anything by deviating from their current action profile.

However, when we set $\gamma > 0$, this exact potential structure breaks, since unilateral deviations can alter $\Psi(\mathbf{B})$ in ways not captured solely by the deviating agent's utility. We therefore interpret our setup as an approximate potential game, and seek to explore whether the best response dynamics still correspond to monotonic increase along our constructed global potential. The methodology described in [10] lays out a combinatorial decomposition technique for evaluating the 'potentialness' of a general sum game, which they extend to imperfect-information games like ours. We will examine our perturbed $\gamma > 0$ potential function with this framework, and use it to assess our central hypothesis regarding the existence of an information-homeostatic NE.

D. Nash Equilibrium as Information Homeostasis

As discussed above, no agent can unilaterally improve their utility by changing their action when the system has reached a Nash equilibria $\mathbf{a}^*$. Through our shaping of U_i to reward uncertainty reduction, penalize congestion, and incentivize belief alignment, we hypothesize that these equilibria may resemble an 'information homeostasis' - agents respond to congestion pressures and distribute across separate frontiers ensuring exploration, but intelligently opt to intermittently stay, patrol, or return to previous regions to drive $\bar{B} \to 0$. In this state, the defined homeostatic belief conditions $\epsilon_\sigma, \epsilon_B$ become equilibrium properties of the NE, and variation in them should be balanced by an equivalent change in the per-agent information-gain and congestion penalty terms.

IV. PRELIMINARY AND TARGET RESULTS

We have implemented a 2D cave gridworld with LOS sensing, instantaneous navigation and belief map merging among communication neighbors, canonical frontier clustering, and round-robin best-response assignment for **an uncoupled information-gain and congestion-only set of games** (γ=0). Qualitative runs illustrate the intended behavior: Fig. 1 shows varying belief states between agents for commonly visited cells as they

A. Target Experiments and Evaluation

For nominal configuration settings, we plan to test over an envelope of grid sizes, team sizes, belief thresholds, ratios of sensing to communication radii, in addition to utility weights α, β, γ.

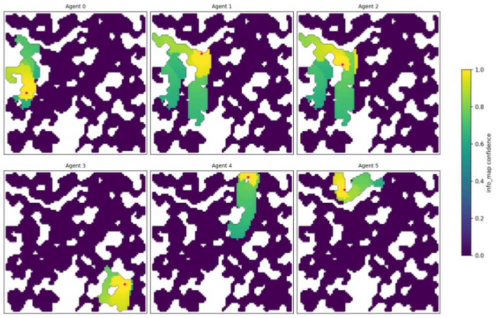

Fig. 1. **Belief maps for each agent. Confidence in each cell decays over time, while cells in sensing radius are always fully known. Agents 1 and 2 have merged maps but possess different beliefs.**

We will evaluate the approach along three technical axes. First, *coverage parity* against standard exploration baselines, to evaluate if consensus-seeking is slowing or starving mapping of the free space. Second, *inter-robot belief alignment* on the communication graph, quantifying how much neighbors disagree on cells they both observe and whether that disagreement decays to a low, stable regime over time. Third, *map informativeness*, using team-level residual uncertainty in the belief field as a proxy that exploration is actually resolving unknown occupancy rather than merely agreeing while leaving large unknown regions. Optionally, we will report perturbation recovery times after communication faults or belief staleness. Metrics and inequalities have been defined for each of these axes, but are not included here.

B. Future Work

The immediate next steps are to implement the stress-coupled game with $\gamma > 0$ using the marginal-$\Psi(\mathbf{B})$ construction, and to log the unilateral potential residual ε_{pot} and variation in potential with agent utilities, to assess if the standard conditions for a NE have been met in addition to our information targets. We will run stress-ablations to test whether the bioelectric minimal-signal hypothesis holds when agents do not share full maps. Longer term, we will relax synchronous the round-robin best response and perfect localization, and incorporate age-of-information–aware belief decay and asynchronous updates to better match realistic exploration conditions.

V. CONCLUSION

We present a preliminary formulation of multi-agent cave exploration as an information-augmented, perturbed potential game, where stress-like signals arising from disagreeing beliefs serve to drive the team to a state of information homeostasis. We have defined a nominal state, action, and utility space with best response dynamics, and used these with local agent beliefs to form a global potential function that can be used to evaluate our predictions about the existence of a Nash equilibria for our system. Ultimately, we hope to show that well-chosen local incentives can lead to global goal alignment within a team of agents, and that planning and coordination can be grounded in beliefs and uncertainty rather than in more direct task and goal assignments.

REFERENCES

[1] S. Nayak, G. Lim, F. Rossi, M. Otte, and J.-P. de la Croix, “Multi-robot exploration for the CADRE mission,” *Autonomous Robots*, vol. 49, no. 17, 2025.

[2] T. H. Chung, V. Orekhov, and A. Maio, “Into the robotic depths: Analysis and insights from the DARPA subterranean challenge,” *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 6, pp. 477–502, 2023.

[3] M. Levin and C. J. Martyniuk, “The bioelectric code: An ancient computational medium for dynamic control of growth and form,” *BioSystems*, vol. 150, pp. 117–133, 2017.

[4] L. Shreesha and M. Levin, “Stress sharing as cognitive glue for collective intelligences: A computational model of stress as a coordinator for morphogenesis,” *Biochemical and Biophysical Research Communications*, 2024.

[5] R. W. Rosenthal, “A class of games possessing pure-strategy Nash equilibria,” *International Journal of Game Theory*, vol. 2, no. 1, pp. 65–67, 1973.

[6] D. Monderer and L. S. Shapley, “Potential games,” *Games and Economic Behavior*, vol. 14, no. 1, pp. 124–143, 1996.

[7] J. R. Marden, G. Arslan, and J. S. Shamma, “Cooperative control and potential games,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 39, no. 6, pp. 1393–1407, 2009.

[8] N. Hayashi, T. Ushio, and T. Kanazawa, “Potential game theoretic approach to power-aware mobile sensor coverage problem,” *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E94-A, no. 3, pp. 929–936, 2011.

[9] Z. Williams, J. Chen, and N. Mehr, “Distributed potential ilqr: Scalable game-theoretic trajectory planning for multi-agent interactions,” 2023.

[10] M. Bichler, D. Legacci, P. Mertikopoulos, M. Oberlechner, and B. Pradelski, “Characterizing the convergence of game dynamics via potentialness,” 2025.

[11] A. Mandal *et al.*, “Behaviorally heterogeneous multi-agent exploration using distributed task allocation,” 2025.

[12] J. Gielis *et al.*, “A critical review of communications in multi-robot systems,” *Current Robotics Reports*, vol. 3, pp. 211–234, 2022.

[13] A.-a. Agha-mohammadi *et al.*, “NeBula: TEAM CoSTAR’s robotic autonomy solution that won phase II of DARPA subterranean challenge,” *Field Robotics*, vol. 4, no. 1, pp. 601–649, 2024.

[14] M. V. Saboia da Silva *et al.*, “ACHORD: Communication-aware multi-robot coordination with intermittent connectivity,” *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 10184–10191, 2022.

[15] B. Franci and F. Fabiani, “Sharing beliefs to learn nash equilibria,” in *2024 European Control Conference (ECC)*, pp. 317–322, 2024.

[16] K. Pongsirijinda, Z. Cao, B. P. L. Lau, R. Liu, C. Yuen, and U.-X. Tan, “Mef-explore: Communication-constrained multi-robot entropy-field-based exploration,” *IEEE Transactions on Automation Science and Engineering*, vol. 22, p. 16062–16078, 2025.

[17] A. Patwardhan and A. J. Davison, “A distributed multi-robot framework for exploration, information acquisition and consensus,” in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 12062–12068, 2024.