Online Sensor Grouping via Multi-Agent Learning Automata: An Ising Model Perspective

Anis Yazidi
Informatics Department, University of Oslo
Oslo, Norway
anis@uio.no

Marco Pinto-Orellana
Integrated Sciences Academy, Rochester Institute of Technology
Rochester, NY, United States
mapcgns@rit.edu

Youcef Djenouri
University of South-Eastern Norway & NORCE Norwegian Research Center
Kongsberg & Oslo, Norway
youcef.djenouri@usn.no

ABSTRACT

We introduce a novel framework for sensor grouping using Learning Automata, drawing inspiration from the Ising model in statistical physics. By representing binary sensor actions as atomic spins, we propose to model the problem as a common Ising-type separation game with interpretable utility functions that promote alignment or anti-alignment based on sensor reliability. In particular, our study presents four reward realizations of this game scenario: ferromagnetic, anti-ferromagnetic, local ferromagnetic, and local anti-ferromagnetic. We rigorously establish that each realization constitutes a potential game aligned with the same separation potential, ensuring that individual sensor incentives contribute to the same overarching separation objective. We also establish that, under standard small-step Learning Automata updates, the induced learning dynamics converge (in the stochastic-approximation sense) to *pure-strategy* Nash equilibria of the corresponding potential game, which are local maximizers of the separation potential. Under the standard two-type reliability model (with $p_i \neq \frac{1}{2}$), these equilibria coincide with the global maximizers and provide perfect separation of fair and unfair sensors (up to a global label flip). Our proposed methods do not require ground truth or local stationarity assumptions about the processes measured by the sensors. By construction, our algorithm uses only current observations and requires minimal state memory in each sensor, making it suitable for near-real-time applications. This work also paves the way for future research in multi-agent coordination leveraging physics principles under noisy feedback conditions, while providing an interpretable framework.

KEYWORDS

Game theory, Ising Model, Sensor Network, Adversarial readings

ACM Reference Format:
Anis Yazidi, Marco Pinto-Orellana, and Youcef Djenouri. 2026. Online Sensor Grouping via Multi-Agent Learning Automata: An Ising Model Perspective. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), Paphos, Cyprus, May 25 – 29, 2026*, IFAAMAS, 9 pages. https://doi.org/10.65109/EBCR6361

Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), C. Amato, L. Dennis, V. Mascardi, J. Thangarajah (eds.), May 25 – 29, 2026, Paphos, Cyprus. © 2026 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). https://doi.org/10.65109/EBCR6361

1 INTRODUCTION

Binary sensors are a particular category of sensors that measure the state of a system with a binary output representing simple states, such as on/off, open/close, or active/inactive [15]. These sensors are often seen as cost-effective solutions with minimal privacy concerns, albeit with a binary level of abstraction over measurable events [5, 21]. For instance, pyroelectric infrared (PIR) sensors are passive sensors that can be used for location tracking and human activity recognition [10], and for independent living support [21], with a versatile installation in environments (e.g., ground or ceiling) [10, 18]. Other categories of binary sensors can be installed on home appliances and furniture to locate specific movements or actions [7]. Data provided by such an array of sensors can be used as biomarkers of dementia, with an F-score of up to 91.60% using ensemble boosted trees [1]. However, depending on the application, some sensors require redundant installation to overlap sensing areas [18] or minimize false positive rates [6]. In some critical contexts, such as real-time elderly assistance, false positives or negatives could lead to an incorrect interpretation of their conditions, resulting in improper care actions [6]. For such critical scenarios, Kodeswaran et al. proposed a system in which daily living activities are monitored with a large number of heterogeneous sensors (n = 14–100) for the main activities, with a failure-detection algorithm to minimize maintenance overhead [9]. In this paper, we focus on the identification of unreliable sensors, i.e., sensors whose measurements have a low probability of corresponding to the inherent process they observe. Sensor fault detection is often based on an expert-provided set of rules [12], pre-trained Markov chain models [14], or anomaly detection in redundant networks. Kodeswaran et al. used the sensor's periodicity to flag abnormalities if it differs considerably from the periodicity of the underlying measurement process [9]. In several scenarios, the ground truth may not be known [12] or obtaining reliable training data is an "obstacle" [21]. In such situations, clustering approaches in networks are often scalable, data-driven alternatives [21], and, by extension, game-theoretic sensor fusion techniques [19, 20], as game theory provides a modeling framework for evolving correlations in asymmetric networks via update strategies [4]. The Ising model is a thermodynamic model for ferromagnetic materials that properly describes macroscopic magnetization using localized, atom-level changes [13, 17]. In this model, the studied atomic property is its spin, which has two possible values (pointing upward or downward). This analogy with the output of binary sensors has inspired game-theoretic learning-automata models where the average magnetization is a utility function and the local energy

is a reward function [11] while providing an interpretable framework [4]. These methods require estimating probabilities using graph properties [11] or known distributions [4].

In this paper, we propose a framework based on principles from the Ising model to address reliability-based sensor grouping challenges in networks without a ground truth reference, where each sensor is modeled as an independent agent. Our framework also avoids assumptions about latent processes, such as local stationarity. This relaxed premise relies on a Learning Automata process where only current samples of sensor agreements are used rather than directly estimating probabilities. Our framework provides four reward realizations with theoretical guarantees to identify reliable versus unreliable sensors: (1) a ferromagnetic realization that rewards global consistency between group labels and measurements, (2) a *signed* anti-ferromagnetic-inspired realization whose microscopic interaction uses ±1 votes within the current group, while still inducing the *same separation incentives* in expectation, (3) a local ferromagnetic realization based on signed votes within the current action group, and (4) a local anti-ferromagnetic realization based on signed votes against the current opposite group. Here, "local" refers to comparisons conditioned on the current action partition rather than to a fixed communication graph. We prove that each realization constitutes a potential game, allowing for a decentralized and scalable method to identify unreliable sensors with reduced computational overhead.

2 DEFINITIONS AND NOTATION

We consider a network of N sensors indexed by $i \in \{1, \dots, N\}$. At each discrete time t, there is an unknown binary *true state* $y(t) \in \{+1, -1\}$ (no ground truth access is assumed). Sensor i produces a binary measurement $x_i(t) \in \{+1, -1\}$ and independently chooses a binary *group label* (action) $s_i(t) \in \{+1, -1\}$ used only for grouping. The sensor actions are abstractions of the spin of an atom in the Ising model [8].

Binary symmetric reliability channel (time-wise i.i.d.). Fix unknown parameters $\{p_i\}_{i=1}^N$ with $p_i \in (0, 1)$. For each sensor i and time t, let $\eta_i(t) \in \{+1, -1\}$ be a noise bit such that

$$\mathbb{P}(\eta_i(t) = +1) = p_i, \qquad \mathbb{P}(\eta_i(t) = -1) = 1 - p_i.$$

Assume $\{\eta_i(t)\}_{i,t}$ are independent across sensors i and time t, and independent of the (possibly nonstationary) true state process $\{y(t)\}$. Measurements are generated as

$$x_i(t) \triangleq y(t)\eta_i(t).$$

Then $p_i = \mathbb{P}(x_i(t) = y(t))$. We classify sensors as **fair** if $p_i > \frac{1}{2}$ and **unfair** if $p_i < \frac{1}{2}$ (classification not used by the algorithm).

Conditional independence and (no) stationarity requirements. Conditional on $y(t)$, the measurements $\{x_i(t)\}_{i=1}^N$ are independent across sensors since $\{\eta_i(t)\}$ are sequentially independent. We do *not* impose stationarity on $\{y(t)\}$: under the binary symmetric channel above, pairwise agreement/disagreement events $\mathbb{I}[x_i(t) = x_j(t)]$ depend only on $(\eta_i(t), \eta_j(t))$ and therefore have time-homogeneous distributions (see Lemma 1).

Profiles. We denote $\boldsymbol{s}(t) = (s_1(t), \dots, s_N(t)) \in \{-1, +1\}^N$ and $\boldsymbol{x}(t) = (x_1(t), \dots, x_N(t)) \in \{-1, +1\}^N$.

Agreement probability and coupling. For $i \neq j$, define the (time-homogeneous) measurement *agreement probability*

$$w_{ij} \triangleq \mathbb{P}\big(x_i(t) = x_j(t)\big).$$

(For completeness we set $w_{ii} \triangleq 1$.) Define the Ising-style coupling

$$J_{ij} \triangleq 2w_{ij} - 1 \in [-1, 1],$$

so that $J_{ij} > 0$ for same-type pairs (both fair or both unfair) and $J_{ij} < 0$ for different-type pairs.

Lemma 1 (Time-homogeneous agreement and factorized coupling). *Under the binary symmetric reliability channel, for any $i \neq j$ and any time t, $w_{ij} = p_i p_j + (1 - p_i)(1 - p_j)$, and $J_{ij} = 2w_{ij} - 1 = (2p_i - 1)(2p_j - 1)$. In particular, w_{ij} and J_{ij} do not depend on the distribution of $y(t)$.*

Proof. For $i \neq j$, $x_i(t) = y(t)\eta_i(t)$ and $x_j(t) = y(t)\eta_j(t)$, hence $x_i(t) = x_j(t)$ iff $\eta_i(t) = \eta_j(t)$. By independence of $\eta_i(t)$ and $\eta_j(t)$,

$$w_{ij} = \mathbb{P}(\eta_i(t) = \eta_j(t)) = p_i p_j + (1 - p_i)(1 - p_j).$$

Then $J_{ij} = 2w_{ij} - 1 = (2p_i - 1)(2p_j - 1)$ by direct expansion. □

Action probabilities. Each sensor i maintains a Bernoulli action distribution $\boldsymbol{P}_i(t) = (p_{i,+1}(t), p_{i,-1}(t))$ with $p_{i,+1}(t) + p_{i,-1}(t) = 1$ and $p_{i,-1}(t) = 1 - p_{i,+1}(t)$.

3 ISING MODEL

The (pairwise) Ising model is a probabilistic model for binary *spins* on a graph [17]. Let $G = (V, E)$ be an undirected graph with $V = \{1, \dots, N\}$, and associate to each node $i \in V$ a spin variable $s_i \in \{-1, +1\}$. A spin configuration is denoted by $\boldsymbol{s} = (s_1, \dots, s_N) \in \{-1, +1\}^N$.

Hamiltonian (energy) and parameters. Given symmetric couplings $\{J_{ij}\}_{(i,j)\in E}$ and (optional) external fields $\{h_i\}_{i\in V}$, the standard Ising Hamiltonian is

$$H(\boldsymbol{s}) \triangleq -\sum_{(i,j)\in E} J_{ij} s_i s_j - \sum_{i\in V} h_i s_i.$$

Under this sign convention, $J_{ij} > 0$ encourages *alignment* ($s_i = s_j$) on edge (i, j) (ferromagnetic interaction), whereas $J_{ij} < 0$ encourages *anti-alignment* ($s_i = -s_j$) (anti-ferromagnetic interaction).

Temperature and Gibbs distribution. For inverse temperature $\beta > 0$, the Gibbs distribution on configurations is

$$\mathbb{P}_\beta(\boldsymbol{s}) \triangleq \frac{1}{Z(\beta)} \exp\big(-\beta H(\boldsymbol{s})\big), \qquad Z(\beta) \triangleq \sum_{\boldsymbol{s}\in\{-1,+1\}^N} \exp\big(-\beta H(\boldsymbol{s})\big),$$

where $Z(\beta)$ is the partition function. In particular, minimizing $H(\boldsymbol{s})$ corresponds to maximizing the score $\sum_{(i,j)\in E} J_{ij} s_i s_j$ when $h_i \equiv 0$.

Spin-flip energy difference. For $h_i \equiv 0$, flipping a single spin s_i to $s_i' = -s_i$ changes the interaction term by

$$H(s_i', \boldsymbol{s}_{-i}) - H(s_i, \boldsymbol{s}_{-i}) = 2s_i \sum_{j:(i,j)\in E} J_{ij} s_j,$$

Mapping to sensor grouping (couplings and separation objective). In our setting, the binary *sensor action* $s_i(t) \in \{-1, +1\}$ is used as a *group label* (spin) for sensor i at time t. The sensor measurement $x_i(t)$ is binary and is *independent* of the chosen group label. Let $w_{ij} = \mathbb{P}(x_i(t) = x_j(t))$ denote the agreement probability between sensors i and j (defined in the previous section), and define the Ising-style coupling

$$J_{ij} \triangleq 2w_{ij} - 1 \in [-1, 1].$$

Under the standard reliability model where $p_i = \mathbb{P}(x_i(t) = y(t))$ and measurements are conditionally independent given the (binary) truth $y(t)$, one has $w_{ij} = p_i p_j + (1 - p_i)(1 - p_j)$ and thus

$$J_{ij} = 2w_{ij} - 1 = (2p_i - 1)(2p_j - 1).$$

Hence $J_{ij} > 0$ for same-type pairs (both fair or both unfair) and $J_{ij} < 0$ for different-type pairs.

We will use the following Ising-type *separation potential*:

$$\Phi_{\text{sep}}(\boldsymbol{s}) \triangleq \sum_{i<j} J_{ij} s_i s_j. \tag{1}$$

which is (up to the conventional sign) the pairwise interaction term of an Ising Hamiltonian on the complete graph. Maximizing Φ_{sep} promotes $s_i = s_j$ when $J_{ij} > 0$ and $s_i = -s_j$ when $J_{ij} < 0$, i.e., it promotes grouping by sensor type.

Assuming $p_i \neq \frac{1}{2}$ for all i and writing $a_i \triangleq 2p_i - 1$, we have $J_{ij} = a_i a_j$ and

$$\Phi_{\text{sep}}(\boldsymbol{s}) = \frac{1}{2}\left(\left(\sum_{i=1}^{N} a_i s_i\right)^2 - \sum_{i=1}^{N} a_i^2\right),$$

so Φ_{sep} is maximized by choosing $s_i = \text{sign}(a_i)$ for all i (and its global sign flip), which yields perfect separation of fair ($p_i > \frac{1}{2}$) and unfair ($p_i < \frac{1}{2}$) sensors.

Proposition 1 (Global maximizers of Φ_{sep}). *Assume $p_i \neq \frac{1}{2}$ for all i and define $a_i \triangleq 2p_i - 1$. Then for all $i < j$ we have $J_{ij} = a_i a_j$ and*

$$\Phi_{\text{sep}}(\boldsymbol{s}) = \frac{1}{2}\left(\left(\sum_{i=1}^{N} a_i s_i\right)^2 - \sum_{i=1}^{N} a_i^2\right).$$

Consequently, Φ_{sep} is maximized by $s_i = \text{sign}(a_i)$ for all i (and by its global sign flip), yielding perfect separation of fair and unfair sensors.

Lemma 2 (No spurious one-flip local maxima). *Under the assumptions of Proposition 1, the only one-flip local maxima of Φ_{sep} on $\{-1, +1\}^N$ are the two global maximizers $\boldsymbol{s} = \text{sign}(\boldsymbol{a})$ and $\boldsymbol{s} = -\text{sign}(\boldsymbol{a})$. Equivalently, if $\Phi_{\text{sep}}(\boldsymbol{s}) \geq \Phi_{\text{sep}}(-s_i, \boldsymbol{s}_{-i})$ for all i, then $\boldsymbol{s} \in \{\text{sign}(\boldsymbol{a}), -\text{sign}(\boldsymbol{a})\}$.*

4 ISING-INSPIRED REWARD DESIGNS FOR SENSOR GROUPING

We now define four reward/utility designs using Ising interactions. At each time t, each sensor i samples a group label $s_i(t)$ and produces a measurement $x_i(t)$. Each sensor then computes a reward $r_i(t) \in [0, 1]$ from *pairwise comparisons* of $(s_i(t), x_i(t))$ with the received $(s_j(t), x_j(t))$ from other sensors (one bit each). No ground truth $y(t)$ is required.

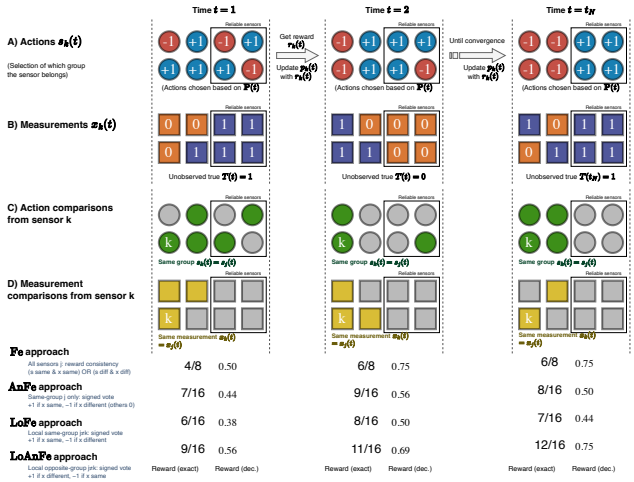

Figure 1: Graphic representation of a game with eight sensors. Each sensor k has (A) an action/group label $s_k(t) \in \{-1, +1\}$ sampled from its action probabilities, and (B) a binary measurement $x_k(t)$ of an unknown true signal. The actions are used only for grouping and do not affect the measurements. Using a reference sensor k, (C) identifies the sensors in the same current action group and (D) the sensors with the same current measurement. These relationships define the four reward realizations Fe, AnFe, LoFe, and LoAnFe. For visual convenience, the figure uses $0/1$ symbols for the binary measurements and hidden truth; in the formal model these correspond to $x_i(t), y(t) \in \{-1, +1\}$, with $T(t)$ denoting $y(t)$. The displayed rewards are exact for $N = 8$: Fe uses $V_k^{\text{Fe}}/8$, while AnFe, LoFe, and LoAnFe use $\frac{1}{2} + V_k/16$; AnFe includes the self term $j = k$, whereas LoFe and LoAnFe exclude self ($j \neq k$).

For each approach, we define a reward $r_i(\boldsymbol{s}, \boldsymbol{x})$ and the corresponding expected utility as the expectation is taken over the measurement conditional on the (fixed) action vector $\boldsymbol{s}$:

$$U_i(\boldsymbol{s}) \triangleq \mathbb{E}_{\boldsymbol{x}}\left[r_i(\boldsymbol{s}, \boldsymbol{x}) | \boldsymbol{s}\right],$$

For brevity, we omit the conditional symbol in the rest of this manuscript.

In contrast to the lattice topology of classical Ising models, our framework imposes no constraints on sensor spatial arrangement or communication topology. The interaction structure is defined by action groupings that emerge dynamically through the Learning Automata process. In LoFe and LoAnFe, *local* therefore refers to comparisons restricted by the current partition rather than to a fixed graph-local neighborhood.

4.1 Ferromagnetic Approach (Fe)

The global ferromagnetic scheme rewards *consistency* between group labels and measurements: agreement inside the same group and disagreement across different groups.

Reward. Define the raw consistency count

$$V_i^{\text{Fe}}(\boldsymbol{s}, \boldsymbol{x}) \triangleq \sum_{j=1}^{N} \Big(\mathbb{I}[s_i = s_j,\ x_i = x_j] + \mathbb{I}[s_i \neq s_j,\ x_i \neq x_j] \Big).$$

where $\mathbb{I}(A)$ is the indicator function (which is 1 when A holds and 0 otherwise).

In this reward, we include the term $j = i$ for simplicity in the notation. It contributes a constant without changing the overall incentives. Then, the normalized reward is

$$r_i^{\text{Fe}}(\boldsymbol{s}, \boldsymbol{x}) \triangleq \frac{1}{N} V_i^{\text{Fe}}(\boldsymbol{s}, \boldsymbol{x}) \in [0, 1].$$

Expected utility. The utility, after taking expectations and using $J_{ij} = 2w_{ij} - 1$ results in:

$$U_i^{\text{Fe}}(\boldsymbol{s}) = \frac{1}{2} + \frac{1}{2N} \sum_{j=1}^{N} J_{ij} s_i s_j.$$

Thus, pairs with $J_{ij} > 0$ increase utility when $s_i = s_j$, while pairs with $J_{ij} < 0$ increase utility when $s_i \neq s_j$, aligning incentives with the separation potential Φ_{sep}.

4.2 Anti-Ferromagnetic Approach (AnFe)

We use a *signed* anti-ferromagnetic-inspired variant whose *microscopic* interaction uses ±1 votes for agreement/disagreement *within the current group*, while still inducing the same separation incentives in expectation.

Reward. Define the signed within-group vote sum

$$V_i^{\text{AnFe}}(\boldsymbol{s}, \boldsymbol{x}) \triangleq \sum_{j=1}^{N} \Big(\mathbb{I}[s_i = s_j,\ x_i = x_j] - \mathbb{I}[s_i = s_j,\ x_i \neq x_j] \Big),$$

where $V_i^{\text{AnFe}} \in [-N, N]$ and normalize/shift to $[0, 1]$:

$$r_i^{\text{AnFe}}(\boldsymbol{s}, \boldsymbol{x}) \triangleq \frac{1}{2} + \frac{1}{2N} V_i^{\text{AnFe}}(\boldsymbol{s}, \boldsymbol{x}) \in [0, 1].$$

Expected utility. A direct computation gives

$$U_i^{\text{AnFe}}(\boldsymbol{s}) = C_i + \frac{1}{4N} \sum_{j=1}^{N} J_{ij} s_i s_j, \qquad C_i \triangleq \frac{1}{2} + \frac{1}{4N} \sum_{j=1}^{N} J_{ij},$$

where C_i is constant with respect to s_i. Consequently, for fixed $\boldsymbol{s}_{-i}$, the best response of sensor i under AnFe is the same as under Fe (both maximize $\sum_j J_{ij} s_i s_j$ up to positive scaling/offset). The full derivation and the induced exact-potential relation are given in Section 6.

4.3 Local Ferromagnetic Approach (LoFe)

The local ferromagnetic scheme compares a sensor only with those currently in the *same* group but keeps *signed* agreement votes. In this approach, we normalize by the total population N to preserve an exact potential-game structure for Φ_{sep}.

Intragroup neighborhood. For a set $\boldsymbol{s}$, define

$$G_i(\boldsymbol{s}) \triangleq \{ j \neq i \mid s_j = s_i \}.$$

Reward. For $j \in G_i(\boldsymbol{s})$, define the signed vote

$$\text{vote}_{ij}^{\text{LoFe}}(\boldsymbol{s}, \boldsymbol{x}) \triangleq \mathbb{I}[x_i = x_j] - \mathbb{I}[x_i \neq x_j] \in \{-1, +1\},$$

and the raw score

$$V_i^{\text{LoFe}}(\boldsymbol{s}, \boldsymbol{x}) \triangleq \sum_{j \in G_i(\boldsymbol{s})} \text{vote}_{ij}^{\text{LoFe}}(\boldsymbol{s}, \boldsymbol{x}).$$

We define the normalized reward

$$r_i^{\text{LoFe}}(\boldsymbol{s}, \boldsymbol{x}) \triangleq \frac{1}{2} + \frac{1}{2N} V_i^{\text{LoFe}}(\boldsymbol{s}, \boldsymbol{x}).$$

If $G_i(\boldsymbol{s}) = \emptyset$, the sum is empty and $r_i^{\text{LoFe}} = 1/2$.

Expected utility. Using $\mathbb{E}[\text{vote}_{ij}^{\text{LoFe}}] = 2w_{ij} - 1 = J_{ij}$, we obtain

$$U_i^{\text{LoFe}}(\boldsymbol{s}) = \frac{1}{2} + \frac{1}{2N} \sum_{j \in G_i(\boldsymbol{s})} J_{ij}.$$

For comparison, replacing N by $|G_i(\boldsymbol{s})|$ results in a group-size-normalized expression closely related to [19]. However, the N-normalization preserves an exact potential-game relation to Φ_{sep}.

4.4 Local Anti-Ferromagnetic Approach (LoAnFe)

The local anti-ferromagnetic scheme compares a sensor only to those currently in the *opposite* group, and rewards disagreements across groups and penalizes agreements.

Opposite-group neighborhood. Define

$$H_i(\boldsymbol{s}) \triangleq \{ j \neq i \mid s_j \neq s_i \}.$$

Reward. For $j \in H_i(\boldsymbol{s})$, define

$$\text{vote}_{ij}^{\text{LoAnFe}}(\boldsymbol{s}, \boldsymbol{x}) \triangleq \mathbb{I}[x_i \neq x_j] - \mathbb{I}[x_i = x_j] \in \{-1, +1\},$$

and

$$V_i^{\text{LoAnFe}}(\boldsymbol{s}, \boldsymbol{x}) \triangleq \sum_{j \in H_i(\boldsymbol{s})} \text{vote}_{ij}^{\text{LoAnFe}}(\boldsymbol{s}, \boldsymbol{x}).$$

The normalized reward is

$$r_i^{\text{LoAnFe}}(\boldsymbol{s}, \boldsymbol{x}) \triangleq \frac{1}{2} + \frac{1}{2N} V_i^{\text{LoAnFe}}(\boldsymbol{s}, \boldsymbol{x}).$$

If $H_i(\boldsymbol{s}) = \emptyset$, we again have $r_i^{\text{LoAnFe}} = 1/2$.

Expected utility. Since $\mathbb{E}[\text{vote}_{ij}^{\text{LoAnFe}}] = 1 - 2w_{ij} = -J_{ij}$ and for $j \in H_i(\boldsymbol{s})$ we have $s_i s_j = -1$, it follows that $-J_{ij} = J_{ij} s_i s_j$. Therefore

$$U_i^{\text{LoAnFe}}(\boldsymbol{s}) = \frac{1}{2} + \frac{1}{2N} \sum_{j \in H_i(\boldsymbol{s})} J_{ij} s_i s_j.$$

Potential-game alignment. Each discussed realization is associated with utilities whose unilateral incentive changes are proportional to the change in the same separation potential $\Phi_{\text{sep}}(\boldsymbol{s}) = \sum_{i<j} J_{ij} s_i s_j$, up to a positive scaling factor. Hence, the four induced games share the same pure-strategy Nash equilibria and are best interpreted as reward realizations of the same separation objective. Detailed derivations, with the exact constants κ for each scheme, are provided in Section 6, see Lemma 4.

5 LEARNING AUTOMATA (LA) ALGORITHM AND CONVERGENCE ANALYSIS

We use a standard Learning Automaton (LA). Each sensor i maintains $\boldsymbol{P}_i(t) = (p_{i,+1}(t), p_{i,-1}(t))$, samples a group label $s_i(t) \sim \boldsymbol{P}_i(t)$, receives a bounded reward $r_i(t) \in [0,1]$ from one of the four schemes in Section 4, and updates its probabilities.

5.1 LA Update Rule

For a step size $\lambda \in (0,1)$, the update is

$$p_{i,s_i(t)}(t+1) = p_{i,s_i(t)}(t) + \lambda\, r_i(t)\big(1 - p_{i,s_i(t)}(t)\big), \tag{2}$$

$$p_{i,k}(t+1) = p_{i,k}(t) - \lambda\, r_i(t)\, p_{i,k}(t), \qquad k \neq s_i(t). \tag{3}$$

Stopping criterion (implementation only). In simulations, we declare convergence when $\max\{p_{i,+1}(t), p_{i,-1}(t)\} \geq 1 - \epsilon$ holds for all i. This is a *termination rule* and is not part of the LA dynamics analyzed below.

5.2 Expected Utilities and Separation Potential

For any fixed action profile $\boldsymbol{s}$, each scheme defines a bounded expected utility $U_i(\boldsymbol{s}) \triangleq \mathbb{E}_{\boldsymbol{x}}\left[r_i(\boldsymbol{s}, \boldsymbol{x})\right] \in [0,1]$. Given a mixed-strategy profile $\boldsymbol{P} = (\boldsymbol{P}_1, \ldots, \boldsymbol{P}_N)$, define the *conditional* expected utility of action $k \in \{-1, +1\}$ for player i:

$$u_{i,k}(\boldsymbol{P}) \triangleq \mathbb{E}\left[U_i(\boldsymbol{s}) \mid \boldsymbol{P},\ s_i = k\right] = \sum_{\boldsymbol{s}_{-i}} U_i(k, \boldsymbol{s}_{-i}) \prod_{j \neq i} p_{j,s_j}. \tag{4}$$

Let $\bar{u}_i(\boldsymbol{P}) \triangleq \sum_k p_{i,k}\, u_{i,k}(\boldsymbol{P})$ denote sensor i's average expected utility under $\boldsymbol{P}$.

We also define the expected separation potential under mixed strategies

$$\Phi(\boldsymbol{P}) \triangleq \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{P}}\left[\Phi_{\text{sep}}(\boldsymbol{s})\right] = \sum_{\boldsymbol{s}} \Phi_{\text{sep}}(\boldsymbol{s}) \prod_j p_{j,s_j}. \tag{5}$$

Lemma 3 (i.i.d. rewards under a frozen policy). *Fix any mixed-strategy profile $\boldsymbol{P}$ and suppose each sensor samples $s_i(t) \sim \boldsymbol{P}_i$ i.i.d. over time (i.e., $\boldsymbol{P}(t) \equiv \boldsymbol{P}$). Under the binary symmetric reliability channel of Section 2, the resulting reward vector $\boldsymbol{r}(t)$ (under any of the four schemes) is i.i.d. over time.*

Proof. Under the four reward definitions, each $r_i(t)$ is a measurable function of $\big(\boldsymbol{s}(t), \boldsymbol{x}(t)\big)$. If $\boldsymbol{P}(t) \equiv \boldsymbol{P}$, then $\boldsymbol{s}(t)$ is i.i.d. over time. By the measurement model, $\boldsymbol{x}(t)$ is generated from i.i.d. noise bits $\{\eta_i(t)\}_i$ that are independent across time; moreover, agreement events $\mathbb{I}[x_i(t) = x_j(t)]$ depend only on $(\eta_i(t), \eta_j(t))$. Hence $\big(\boldsymbol{s}(t), \boldsymbol{x}(t)\big)$ is i.i.d. over time, and so are $\boldsymbol{r}(t)$. □

5.3 Mean ODE Approximation

As $\lambda \to 0$, based on standard stochastic approximation theory, as proved in [2] and [3], we can approximate the discrete updates as continuous-time dynamics.

Theorem 2 (Weak ODE limit of LR–I learning automata). *Consider the linear reward–inaction (LR–I) update with a constant step size $\lambda > 0$. Let $\boldsymbol{P}^{\lambda}(\cdot)$ denote the standard continuous-time interpolation of the discrete iterates $\{\boldsymbol{P}(t)\}_{t\in\mathbb{N}}$ on the λ-rescaled time axis. Under Lemma 3 (i.i.d. rewards when $\boldsymbol{P}$ is held fixed), as $\lambda \to 0$ the interpolated process $\boldsymbol{P}^{\lambda}(\cdot)$ converges weakly to the unique solution of the mean ODE*

$$\dot{p}_{i,k} = p_{i,k}\Big(u_{i,k}(\boldsymbol{P}) - \bar{u}_i(\boldsymbol{P})\Big), \qquad \forall i,\ k \in \{-1,+1\}, \tag{6}$$

with the same initial condition.

Proof. This is the standard ODE method for LR–I learning automata [16]. Lemma 3 provides the key hypothesis used there: if $\boldsymbol{P}(t)$ is held fixed at $\boldsymbol{P}$, then the action–reward process is i.i.d. □

5.4 Potential Along the Mean ODE

For each of the four reward designs, the induced game is an *exact potential game* with potential Φ_{sep} up to a positive scaling constant. Because these constants are positive, all four reward designs induce the same best-response correspondence. Concretely, for any unilateral flip $s_i' = -s_i$ and fixed $\boldsymbol{s}_{-i}$,

$$U_i(s_i', \boldsymbol{s}_{-i}) - U_i(s_i, \boldsymbol{s}_{-i}) = \kappa\Big(\Phi_{\text{sep}}(s_i', \boldsymbol{s}_{-i}) - \Phi_{\text{sep}}(s_i, \boldsymbol{s}_{-i})\Big), \tag{7}$$

with $\kappa = \frac{1}{2N}$ for Fe and $\kappa = \frac{1}{4N}$ for AnFe, LoFe, and LoAnFe (see Section 6.2 for the derivations).

Theorem 3 (Monotonicity of expected separation potential). *Along any solution of the mean ODE (6), the expected separation potential $\Phi(\boldsymbol{P}(t))$ is non-decreasing:*

$$\frac{d}{dt}\Phi(\boldsymbol{P}(t)) \geq 0.$$

More precisely, for the scaling constant κ in (7),

$$\frac{d}{dt}\Phi(\boldsymbol{P}) = \frac{1}{2\kappa}\sum_{i=1}^{N}\sum_{k\in\{-1,+1\}}\sum_{k'\in\{-1,+1\}} p_{i,k}p_{i,k'}\big(u_{i,k}(\boldsymbol{P}) - u_{i,k'}(\boldsymbol{P})\big)^2 \geq 0. \tag{8}$$

Proof. Differentiate (5):

$$\frac{d}{dt}\Phi(\boldsymbol{P}) = \sum_i \sum_k \frac{\partial\Phi(\boldsymbol{P})}{\partial p_{i,k}}\,\dot{p}_{i,k}.$$

From (5),

$$\frac{\partial\Phi(\boldsymbol{P})}{\partial p_{i,k}} = \mathbb{E}\left[\Phi_{\text{sep}}(\boldsymbol{s}) \,\middle|\, \boldsymbol{P},\ s_i = k\right] \triangleq \phi_{i,k}(\boldsymbol{P}).$$

The exact potential property (7) implies that, for each i and any k, k', the conditional expected utilities satisfy

$$u_{i,k}(\boldsymbol{P}) - u_{i,k'}(\boldsymbol{P}) = \kappa\big(\phi_{i,k}(\boldsymbol{P}) - \phi_{i,k'}(\boldsymbol{P})\big),$$

hence $u_{i,k}(\boldsymbol{P}) = \kappa\,\phi_{i,k}(\boldsymbol{P}) + c_i(\boldsymbol{P})$ for a term $c_i(\boldsymbol{P})$ independent of k. Using $\dot{p}_{i,k} = p_{i,k}(u_{i,k} - \bar{u}_i)$ and eliminating c_i gives $\dot{p}_{i,k} = \kappa\, p_{i,k}(\phi_{i,k} - \bar{\phi}_i)$, where $\bar{\phi}_i = \sum_k p_{i,k}\phi_{i,k}$. Substituting into the derivative yields

$$\frac{d}{dt}\Phi(\boldsymbol{P}) = \kappa \sum_i \mathrm{Var}_{k\sim p_i}\big(\phi_{i,k}\big) \geq 0.$$

Rewriting the variance as a pairwise sum gives (8). □

5.5 Convergence to Pure-Strategy Nash Equilibria

Theorem 4 (Stability and strict pure Nash equilibria). *For any of the four schemes, consider the mean ODE* (6)*. Then:*

- *Every pure-strategy profile (corner of the simplex) is a stationary point of the ODE.*
- *Every Nash equilibrium (pure or mixed) is a stationary point of the ODE.*
- *Every stationary point that is not a Nash equilibrium is unstable.*
- *Every strict pure-strategy Nash equilibrium is an asymptotically stable stationary point.*

Proof. These statements follow from [16] for LR–I learning automata. □

5.6 Implication for Sensor Separation

Theorem 4 shows that asymptotically stable stationary points of the mean ODE correspond to *strict* pure-strategy Nash equilibria. Because each of our four reward designs induces an exact potential game with potential Φ_{sep} (Section 4.4 and Section 6), pure-strategy Nash equilibria coincide across the four realizations and with the one-flip local maximizers of Φ_{sep}. Under the reliability model with $p_i \neq \frac{1}{2}$ for all i, Lemma 2 implies that the only one-flip local maximizers are the two perfect separation profiles $\boldsymbol{s} = \text{sign}(\boldsymbol{a})$ and its global sign flip. Hence, every strict pure-strategy Nash equilibrium yields perfect separation of fair and unfair sensors (up to a global label flip), and the LR–I learning automata converges to such equilibria in the standard ODE/weak-convergence sense of Theorem 2.

6 UTILITY DERIVATIONS AND EXACT POTENTIAL PROPERTY

This section collects the derivations of expected utilities and the exact-potential relations used in Sections 4–5.

6.1 Expected utilities for the four schemes

Global ferromagnetic (Fe). Recall

$$r_i^{\text{Fe}}(\boldsymbol{s}, \boldsymbol{x}) = \frac{1}{N} \sum_{j=1}^{N} \Big(\mathbb{I}[s_i = s_j,\, x_i = x_j] + \mathbb{I}[s_i \neq s_j,\, x_i \neq x_j] \Big).$$

For $i \neq j$, let $w_{ij} = \mathbb{P}(x_i = x_j)$ and $J_{ij} = 2w_{ij} - 1$. Then, for fixed $\boldsymbol{s}$,

$$\mathbb{E}[\mathbb{I}(s_i = s_j, x_i = x_j)] = \mathbb{I}(s_i = s_j)\, w_{ij},$$

$$\mathbb{E}[\mathbb{I}(s_i \neq s_j, x_i \neq x_j)] = \mathbb{I}(s_i \neq s_j)\, (1 - w_{ij}).$$

Using $\mathbb{I}(s_i = s_j) = \frac{1+s_i s_j}{2}$ and $\mathbb{I}(s_i \neq s_j) = \frac{1-s_i s_j}{2}$, one obtains

$$U_i^{\text{Fe}}(\boldsymbol{s}) = \frac{1}{2} + \frac{1}{2N} \sum_{j=1}^{N} J_{ij} s_i s_j,$$

where the $j = i$ term contributes only a constant (since $s_i^2 = 1$).

Global signed anti-ferromagnetic (AnFe). Recall

$$r_i^{\text{AnFe}}(\boldsymbol{s}, \boldsymbol{x}) = \frac{1}{2} + \frac{1}{2N} \sum_{j=1}^{N} \Big(\mathbb{I}[s_i = s_j,\, x_i = x_j] - \mathbb{I}[s_i = s_j,\, x_i \neq x_j] \Big).$$

For fixed $\boldsymbol{s}$,

$$\mathbb{E}\big[\mathbb{I}(s_i = s_j, x_i = x_j) - \mathbb{I}(s_i = s_j, x_i \neq x_j)\big] = \mathbb{I}(s_i = s_j)\, J_{ij}.$$

Thus

$$U_i^{\text{AnFe}}(\boldsymbol{s}) = \frac{1}{2} + \frac{1}{4N} \sum_{j=1}^{N} J_{ij} + \frac{1}{4N} \sum_{j=1}^{N} J_{ij} s_i s_j = C_i + \frac{1}{4N} \sum_{j=1}^{N} J_{ij} s_i s_j,$$

with $C_i \triangleq \frac{1}{2} + \frac{1}{4N} \sum_j J_{ij}$ independent of s_i.

Local ferromagnetic (LoFe). Recall $G_i(\boldsymbol{s}) = \{j \neq i : s_j = s_i\}$ and

$$r_i^{\text{LoFe}}(\boldsymbol{s}, \boldsymbol{x}) = \frac{1}{2} + \frac{1}{2N} \sum_{j \in G_i(\boldsymbol{s})} \big(\mathbb{I}[x_i = x_j] - \mathbb{I}[x_i \neq x_j] \big).$$

Since $\mathbb{E}[\mathbb{I}(x_i = x_j) - \mathbb{I}(x_i \neq x_j)] = 2w_{ij} - 1 = J_{ij}$, we get

$$U_i^{\text{LoFe}}(\boldsymbol{s}) = \frac{1}{2} + \frac{1}{2N} \sum_{j \in G_i(\boldsymbol{s})} J_{ij}.$$

Local anti-ferromagnetic (LoAnFe). Recall $H_i(\boldsymbol{s}) = \{j \neq i : s_j \neq s_i\}$ and

$$r_i^{\text{LoAnFe}}(\boldsymbol{s}, \boldsymbol{x}) = \frac{1}{2} + \frac{1}{2N} \sum_{j \in H_i(\boldsymbol{s})} \big(\mathbb{I}[x_i \neq x_j] - \mathbb{I}[x_i = x_j] \big).$$

Here $\mathbb{E}[\mathbb{I}(x_i \neq x_j) - \mathbb{I}(x_i = x_j)] = 1 - 2w_{ij} = -J_{ij}$, and for $j \in H_i(\boldsymbol{s})$ we have $s_i s_j = -1$, hence $-J_{ij} = J_{ij} s_i s_j$. Therefore

$$U_i^{\text{LoAnFe}}(\boldsymbol{s}) = \frac{1}{2} + \frac{1}{2N} \sum_{j \in H_i(\boldsymbol{s})} J_{ij} s_i s_j.$$

Remark (same best responses of the signed realizations). Using the expressions above, the three signed realizations AnFe, LoFe, and LoAnFe differ only by additive terms independent of s_i:

$$U_i^{\text{LoFe}}(\boldsymbol{s}) = \frac{1}{2} + \frac{1}{4N} \sum_{j \neq i} J_{ij} + \frac{1}{4N} \sum_{j \neq i} J_{ij} s_i s_j,$$

$$U_i^{\text{LoAnFe}}(\boldsymbol{s}) = U_i^{\text{LoFe}}(\boldsymbol{s}) - \frac{1}{2N} \sum_{j \neq i} J_{ij}, \qquad U_i^{\text{AnFe}}(\boldsymbol{s}) = U_i^{\text{LoFe}}(\boldsymbol{s}) + \frac{1}{2N}.$$

Consequently, AnFe, LoFe, and LoAnFe induce the same best-response correspondence.

6.2 Exact potential constants

Lemma 4 (Exact potential property and constants). *Let* $\Phi_{\text{sep}}(\boldsymbol{s}) = \sum_{i<j} J_{ij} s_i s_j$ *and let* $s_i' = -s_i$.

For each scheme $X \in \{\text{Fe}, \text{AnFe}, \text{LoFe}, \text{LoAnFe}\}$ *there exists a constant* $\kappa_X > 0$ *such that*

$$U_i^X(s_i', \boldsymbol{s}_{-i}) - U_i^X(s_i, \boldsymbol{s}_{-i}) = \kappa_X \Big(\Phi_{\text{sep}}(s_i', \boldsymbol{s}_{-i}) - \Phi_{\text{sep}}(s_i, \boldsymbol{s}_{-i}) \Big).$$

Concretely, $\kappa_{Fe} = \frac{1}{2N}$ *and* $\kappa_{AnFe} = \kappa_{LoFe} = \kappa_{LoAnFe} = \frac{1}{4N}$.

Proof. For any profile $\boldsymbol{s}$, a single flip satisfies

$$\Phi_{\text{sep}}(s_i', \boldsymbol{s}_{-i}) - \Phi_{\text{sep}}(s_i, \boldsymbol{s}_{-i}) = -2 s_i \sum_{j \neq i} J_{ij} s_j.$$

Global FE: from the utility expression above, $U_i^{\text{FE}}(\boldsymbol{s}) = \frac{1}{2} + \frac{1}{2N}\sum_j J_{ij}s_i s_j$, hence

$$U_i^{\text{FE}}(s_i', \boldsymbol{s}_{-i}) - U_i^{\text{FE}}(s_i, \boldsymbol{s}_{-i}) = \frac{1}{2N}\Delta\Phi_{\text{sep}}.$$

Global ANFE: the only s_i-dependent term is $\frac{1}{4N}\sum_j J_{ij}s_i s_j$, giving

$$U_i^{\text{ANFE}}(s_i', \boldsymbol{s}_{-i}) - U_i^{\text{ANFE}}(s_i, \boldsymbol{s}_{-i}) = \frac{1}{4N}\Delta\Phi_{\text{sep}}.$$

Local LOFE: flipping s_i swaps $G_i(\boldsymbol{s})$ with $H_i(\boldsymbol{s})$, so the difference of utilities is

$$\frac{1}{2N}\Big(\sum_{j\in H_i} J_{ij} - \sum_{j\in G_i} J_{ij}\Big) = \frac{1}{4N}\Delta\Phi_{\text{sep}}.$$

Local LOANFE: similarly,

$$U_i^{\text{LOANFE}}(s_i', \boldsymbol{s}_{-i}) - U_i^{\text{LOANFE}}(s_i, \boldsymbol{s}_{-i}) = \frac{1}{4N}\Delta\Phi_{\text{sep}}.$$

□

7 PROOFS FOR PROPOSITION 1 AND LEMMA 2

PROOF OF PROPOSITION 1. With $a_i \triangleq 2p_i - 1$ and $J_{ij} = a_i a_j$ for $i < j$,

$$\sum_{i<j} a_i a_j s_i s_j = \frac{1}{2}\left(\left(\sum_{i=1}^N a_i s_i\right)^2 - \sum_{i=1}^N a_i^2\right),$$

by expanding the square and sum off-diagonal terms. The second term is constant in $\boldsymbol{s}$, so maximizing Φ_{sep} is equivalent to maximizing $|\sum_i a_i s_i|$, achieved by $s_i = \text{sign}(a_i)$ (and its global sign flip). □

PROOF OF LEMMA 2. Let $L(\boldsymbol{s}) \triangleq \sum_{i=1}^N a_i s_i$. By Proposition 1, $\Phi_{\text{sep}}(\boldsymbol{s}) = \frac{1}{2}(L(\boldsymbol{s})^2 - \sum_i a_i^2)$. If $L(\boldsymbol{s}) = 0$, then flipping any s_i changes L to $L' = -2a_i s_i \neq 0$ and increases L^2, so $\boldsymbol{s}$ cannot be one-flip locally maximal.

Assume $L(\boldsymbol{s}) > 0$. If there exists i with $a_i s_i < 0$, then flipping s_i changes L to $L' = L - 2a_i s_i = L + 2|a_i| > L$, hence strictly increases L^2 and thus Φ_{sep}, contradicting one-flip local maximality. Therefore $a_i s_i > 0$ for all i, i.e., $s_i = \text{sign}(a_i)$ for all i.

If $L(\boldsymbol{s}) < 0$, the same argument applied to $-\boldsymbol{s}$ yields $\boldsymbol{s} = -\text{sign}(\boldsymbol{a})$. □

8 EXPERIMENTAL RESULTS

8.1 Experimental Setup

We consider three simulated scenarios: A) **clearly separable groups**, where sensors have distinctly different reliability levels, making it easier to separate fair and unfair sensors; B) **difficult to separate groups**, where sensors have similar reliability levels, making grouping more challenging; and C) a **large network with moderate difficulty** to assess scalability. For each scenario, we simulate two groups, S_1 and S_2, with a known number of sensors and reliabilities (N_1, p_1, N_2, p_2). Then, we apply the four Ising-inspired reward realizations until convergence as defined in Algorithm 1. After convergence, the two groups obtained are denoted as S_A and S_B with sizes N_A and N_B. The learning rate for all experiments is set to $\lambda = 0.01$.

To evaluate the performance of our approach, we focused on two metrics: the number of iterations until convergence (convergence time) and the error rate. To calculate the latter, we first calculate the accuracy of the sensor identification in groups S_A and S_B: ρ_A and ρ_B, respectively. Then we compute the error rate e as the weighted average of the complement of the accuracy:

$$e = \frac{N_A(1-\rho_A) + N_B(1-\rho_B)}{N_A + N_B} \tag{9}$$

Each scenario is repeated over 100 independent runs, and the mean and 95% CI of the performance metrics are reported. In addition, we included two nonparametric statistical methods to compare convergence times: Kruskal–Wallis (across all groups) and pairwise Bonferroni-corrected Mann–Whitney U.

For comparison with alternative methods, we use LOFE as a reference, since it was previously proposed by Yazidi et al. in [19]. Moreover, for a non-Ising comparison, we added two additional methods: replicator dynamics ([20]) and batch K-means. The first facilitates continuous updates in a manner comparable to our methods, whereas the latter can be assumed to be a reference clustering method. Both methods only use the measurements.

8.2 Results and Discussion

Clearly Separable Groups. Our first simulated scenario includes $N_1 = 10$ fair sensors with reliability $p_1 = 0.9$ and $N_2 = 10$ unfair sensors with reliability $p_2 = 0.1$. Table 1.A shows its convergence time and error rates. All approaches grouped the sensors with zero error rate, indicating effective separation between fair and unfair sensors. Notably, FE achieved convergence in approximately **2119 iterations**, which is about twice as fast as the other methods.

Difficult to Separate Groups. In this scenario, we set $N_1 = 10$, $N_2 = 10$, $p_1 = 0.55$, and $p_2 = 0.45$, making it harder to distinguish between fair and unfair sensors (Table 1.B). The convergence times for all methods are on the order of 10^5 **iterations**, reflecting the increased difficulty of the task. Similarly to the previous scenario, FE converged in approximately **120,000 iterations**, lower than the other approaches.

Moderate Scenario with More Sensors. We consider a larger network with $N_1 = 100$, $N_2 = 100$, $p_1 = 0.7$, and $p_2 = 0.2$ to test scalability (Table 1.C). In this simulation, convergence times of FE

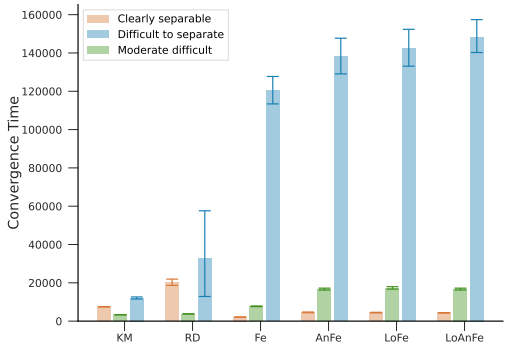

Figure 2: Barplot results of the simulated scenarios. Mean values and bootstrap 95% confidence intervals (CIs) are shown for the six methods: K-means (KM), replicator dynamics (RD), and the FE, ANFE, LOFE, and LOANFE Ising reward realizations.

Table 1: Results of the simulated scenarios: A) clearly separable groups ($N_1 = 10$, $N_2 = 10$, $p_1 = 0.9$, $p_2 = 0.1$), B) difficult-to-separate groups ($N_1 = 10$, $N_2 = 10$, $p_1 = 0.55$, $p_2 = 0.45$), and C) moderate scenario ($N_1 = 100$, $N_2 = 100$, $p_1 = 0.7$, $p_2 = 0.2$)
Mean values and bootstrap 95% confidence intervals (CIs) are shown for the six methods: K-means (KM), replicator dynamics (RD), and the Fe, AnFe, LoFe, and LoAnFe Ising reward realizations.
A Bonferroni-corrected pairwise Mann–Whitney comparison was conducted to analyze the convergence times across the six approaches, with the final column denoting methods exhibiting statistically nonsignificant differences (adjusted p-value > 0.05).

		Convergence Time Mean (CI)	Error Mean (CI)	Similar Convergence Time Method (Adj. P-value)
A) Clearly Separable	KM	7481.07 (7362.90, 7601.26)	0.00 (0.00, 0.00)	-
	RD	20293.15 (18653.05, 21930.61)	0.03 (0.01, 0.04)	-
	Fe	2137.19 (2072.31, 2204.18)	0.00 (0.00, 0.00)	-
	AnFe	4569.23 (4385.61, 4771.07)	0.00 (0.00, 0.00)	LoFe (0.2579),LoAnFe (0.1494)
	LoFe	4439.37 (4277.22, 4616.44)	0.00 (0.00, 0.00)	AnFe (0.2579),LoAnFe (0.7722)
	LoAnFe	4326.61 (4206.56, 4451.35)	0.00 (0.00, 0.00)	AnFe (0.1494),LoFe (0.7722)
B) Difficult to Separate	KM	12065.46 (11551.37, 12614.93)	0.12 (0.10, 0.13)	-
	RD	32754.84 (12868.93, 57609.98)	0.12 (0.11, 0.14)	-
	Fe	120311.35 (113407.53, 127733.69)	0.08 (0.07, 0.09)	-
	AnFe	137965.12 (129083.49, 147699.98)	0.08 (0.07, 0.09)	LoFe (0.4839)
	LoFe	142351.15 (133093.96, 152350.88)	0.09 (0.08, 0.10)	AnFe (0.4839),LoAnFe (0.1883)
	LoAnFe	148276.92 (140238.99, 157416.50)	0.08 (0.07, 0.09)	LoFe (0.1883)
C) Moderate Difficult	KM	3326.94 (3279.47, 3376.43)	0.00 (0.00, 0.00)	-
	RD	3784.08 (3779.00, 3789.26)	0.00 (0.00, 0.00)	-
	Fe	7750.11 (7584.74, 7916.66)	0.00 (0.00, 0.00)	-
	AnFe	16704.50 (16227.64, 17213.16)	0.00 (0.00, 0.00)	LoFe (0.2657),LoAnFe (0.9951)
	LoFe	17328.89 (16711.64, 17979.01)	0.00 (0.00, 0.00)	AnFe (0.2657),LoAnFe (0.2641)
	LoAnFe	16717.08 (16242.79, 17209.88)	0.00 (0.00, 0.00)	AnFe (0.9951),LoFe (0.2641)

and AnFe are around **7750** and **16700** iterations, while LoFe and LoAnFe converge in approximately **17300** and **16700** iterations, respectively. The experiments should be interpreted as a comparison of reward realizations rather than of different equilibrium objectives. Across the three scenarios, Fe consistently converges faster, whereas AnFe, LoFe, and LoAnFe exhibit closely clustered convergence times and similar error profiles. This is consistent with the exact-potential analysis: the three signed realizations share $\kappa = 1/(4N)$, while Fe uses $\kappa = 1/(2N)$. Thus, the main empirical differences lie in microscopic feedback, scaling, and finite-time learning behavior rather than in the equilibrium separation objective. A comprehensive theoretical analysis of the convergence-speed differences between the reward realizations was beyond the scope of the current manuscript.

Furthermore, we need to emphasize that the provided strategies only require each sensor to store its current and immediate past action and a limited memory of the status of each sensor ($O(n)$), while alternative methods require estimating the probabilities [4] ($O(n^2)$). While we compare against a reference approach, a limitation of this study is the number of simulations and comparison methods. Future work involves evaluating our approach in controlled real-life environments to confirm the efficiency of the proposed methods.

9 CONCLUSION

In this work, we developed a novel theoretical framework for distributed sensor/agent coordination inspired by statistical physics principles. The framework comprises four reward realizations of a common separation game: ferromagnetic (Fe), anti-ferromagnetic (AnFe), local ferromagnetic (LoFe), and local anti-ferromagnetic (LoAnFe). Through rigorous mathematical analysis, we demonstrated that these realizations constitute potential games aligned with the same separation potential, where individual sensor objectives align with the same global separation objective.

The main distinction among the four realizations lies in microscopic feedback and finite-time learning behavior. In particular, Fe is a global consistency realization, whereas AnFe, LoFe, and LoAnFe are signed realizations that remain exactly aligned with the same separation incentives and therefore induce the same pure-strategy Nash equilibria. In the local variants, *local* refers to comparisons conditioned on the current action partition rather than to a fixed communication graph.

Our theoretical analysis established convergence guarantees for the LA algorithm in these potential games, demonstrating reliable convergence to Nash equilibria that optimize the common separation potential. This convergence property ensures that the system isolates unreliable sensors and consequently improves measurement accuracy. Using the Ising model, our proposed method allows small-scale interpretable strategies with known effects throughout the system. Experimentally, Fe converged faster, whereas AnFe, LoFe, and LoAnFe exhibited closely clustered behavior, consistent with their shared exact-potential structure. The selection of the proper reward realization therefore depends primarily on microscopic feedback and finite-time behavior rather than on a different equilibrium separation objective. In addition, our findings encourage the development of further statistical physics-inspired reward constructions in game theory.

ACKNOWLEDGMENTS

The work of the first author was supported by Integreat, a Centre of Excellence funded by the Research Council of Norway. The work of the third author was partially funded by the European Commission through the COGNIMAN project under grant agreement number 101058477.

REFERENCES

[1] Farhad Ahamed, Seyed Shahrestani, and Hon Cheung. 2020. Internet of Things and Machine Learning for Healthy Ageing: Identifying the Early Signs of Dementia. *Sensors* 20, 21 (oct 2020), 6031. https://doi.org/10.3390/s20216031

[2] Albert Benveniste, Michel Métivier, and Pierre Priouret. 1990. *Adaptive Algorithms and Stochastic Approximations*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-75894-2

[3] V. S. Borkar and S. P. Meyn. 2000. The O.D.E. Method for Convergence of Stochastic Approximation and Reinforcement Learning. *SIAM Journal on Control and Optimization* 38, 2 (jan 2000), 447–469. https://doi.org/10.1137/s0363012997331639

[4] A.D. Correia, L.L. Leestmaker, H.T.C. Stoof, and J.J. Broere. 2022. Asymmetric games on networks: Towards an Ising-model representation. *Physica A: Statistical Mechanics and its Applications* 593 (may 2022), 126972. https://doi.org/10.1016/j.physa.2022.126972

[5] Dan Ding, Rory A. Cooper, Paul F. Pasquina, and Lavinia Fici-Pasquina. 2011. Sensor technology for smart homes. *Maturitas* 69, 2 (jun 2011), 131–136. https://doi.org/10.1016/j.maturitas.2011.03.016

[6] Nancy E. ElHady and Julien Provost. 2018. A Systematic Survey on Sensor Failure Detection and Fault-Tolerance in Ambient Assisted Living. *Sensors* 18, 7 (jun 2018), 1991. https://doi.org/10.3390/s18071991

[7] Xin Hong, Chris Nugent, Maurice Mulvenna, Sally McClean, Bryan Scotney, and Steven Devlin. 2009. Evidential fusion of sensor data for activity recognition in smart homes. *Pervasive and Mobile Computing* 5, 3 (jun 2009), 236–252. https://doi.org/10.1016/j.pmcj.2008.05.002

[8] Achim Klenke. 2020. *Probability Theory: A Comprehensive Course* (third edition ed.). Springer Nature, Cham, Switzerland. https://doi.org/10.1007/978-3-030-56402-5

[9] Palanivel A. Kodeswaran, Ravi Kokku, Sayandeep Sen, and Mudhakar Srivatsa. 2016. Idea: A System for Efficient Failure Management in Smart IoT Environments. In *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys'16)*. ACM. https://doi.org/10.1145/2906388.2906406

[10] Ranjit Kolkar, V. Geetha, and Sanket Salvi. 2024. *Single Person Occupancy Detection Using PIR Sensors*. Springer Nature Singapore, 549–561. https://doi.org/10.1007/978-981-97-3242-5_37

[11] Andrey Leonidov, Alexey Savvateev, and Andrew G. Semenov. 2024. Ising game on graphs. *Chaos, Solitons & Fractals* 180 (mar 2024), 114540. https://doi.org/10.1016/j.chaos.2024.114540

[12] Beth Logan, Jennifer Healey, Matthai Philipose, Emmanuel Munguia Tapia, and Stephen Intille. 2007. A Long-Term Evaluation of Sensing Modalities for Activity Recognition. In *UbiComp 2007: Ubiquitous Computing*, John Krumm, Gregory D. Abowd, Aruna Seneviratne, and Thomas Strang (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 483–500.

[13] Peter VE McClintock and Aneta Stefanovska. 2021. Introduction to Physics of Biological Oscillators: New Insights into Non-Equilibrium and Non-Autonomous Systems. In *Physics of Biological Oscillators: New Insights into Non-Equilibrium and Non-Autonomous Systems*. Springer.

[14] Ioannis Ch. Paschalidis and Yin Chen. 2010. Statistical anomaly detection with sensor networks. *ACM Transactions on Sensor Networks* 7, 2 (aug 2010), 1–23. https://doi.org/10.1145/1824766.1824773

[15] Jun Qi, Po Yang, Lee Newcombe, Xiyang Peng, Yun Yang, and Zhong Zhao. 2020. An overview of data fusion techniques for Internet of Things enabled physical activity recognition and measure. *Information Fusion* 55 (mar 2020), 269–280. https://doi.org/10.1016/j.inffus.2019.09.002

[16] P Shanti Sastry, Vijay V Phansalkar, and MAL Thathachar. 2002. Decentralized learning of Nash equilibria in multi-person stochastic games with incomplete information. *IEEE Transactions on systems, man, and cybernetics* 24, 5 (2002), 769–777.

[17] A Benjamin Stickler and Ewald Schachinger. 2016. *Basic concepts in computational physics*. Springer.

[18] Bo Yang, Qifan Wei, and Meng Zhang. 2017. Multiple human location in a distributed binary pyroelectric infrared sensor network. *Infrared Physics & Technology* 85 (sep 2017), 216–224. https://doi.org/10.1016/j.infrared.2017.06.007

[19] Anis Yazidi, Hugo L Hammer, Konstantin Samouylov, and Enrique Enrique Herrera-Viedma. 2020. Game-theoretic learning for sensor reliability evaluation without knowledge of the ground truth. *IEEE Transactions on Cybernetics* 51, 12 (2020), 5706–5716.

[20] Anis Yazidi, Marco Antonio Pinto-Orellana, Hugo Hammer, Peyman Mirtaheri, and Enrique Herrera-Viedma. 2020. Solving sensor identification problem without knowledge of the ground truth using replicator dynamics. *IEEE transactions on cybernetics* 52, 1 (2020), 16–24.

[21] Juan Ye, Graeme Stevenson, and Simon Dobson. 2016. Detecting abnormal events on binary sensors in smart home environments. *Pervasive and Mobile Computing* 33 (dec 2016), 32–49. https://doi.org/10.1016/j.pmcj.2016.06.012