# Towards High Semantic Fidelity: Hyperdimensional Symbolic Messages in Multi-Agent Communication

**Shoucheng Song[1,2,3], Chang Yao[1,2], Hao Wu[1,2], Youfang Lin[1,2], Sheng Han[1,2], Kai Lv[1,2]***
[1]School of Computer Science & Technology, Beijing Jiaotong University, Beijing, China
[2]Beijing Key Laboratory of Traffic Data Mining and Embodied Intelligence, Beijing, China
[3]Institute of Artificial Intelligence of China Telecom (TeleAI), Beijing, China
{insis_songsc,yaochang,insis_wh,yflin,shhan,lvkai}@bjtu.edu.cn

## Abstract

Semantic fidelity is essential for multi-agent communication under non-ideal channels, comprising transmission fidelity and conversion fidelity. Yet existing numerical and neural message formats cannot simultaneously achieve both. To address this, we propose a novel message format-**H**yperdimensional **S**ymbolic **M**essages (HSM) and its corresponding communication framework **Herm**. For transmission fidelity, Herm suppresses semantic entanglement via an entity-attribute-value structured encoding pipeline. For conversion fidelity, Herm first leverages a receiver-centric ego-binding mechanism to align semantics. Then, a similarity-aware dual-branch fusion module is proposed to prevent critical semantic dilution. To realize explicit message-semantic conversion, we devise an iterative decoding scheme via an algebraic inverse operation. Experiments exhibit that Herm outperforms baselines across various noisy scenarios while maintaining superior structural readability.

## 1 Introduction

In multi-agent reinforcement learning (MARL) [1, 2, 3, 4, 5, 6], communication serves as the foundation for emerging coordination. In non-ideal channels, the effectiveness of coordination crucially depends on whether the semantics transmitted between agents can be preserved completely and understood accurately. We refer to this critical capability as **Semantic Fidelity**, which features (1) **Transmission Fidelity**: The semantics of messages remain stable and undistorted in noisy or adversarial channels. Even with the perturbed message, the receiver's understanding remains consistent with the sender's original semantics. (2) **Conversion Fidelity**: There exists a bidirectional conversion between agents' native semantics (intentions, trajectories, etc.) and messages. Receivers can recover the native semantics from messages without black-box ambiguity, and transform messages into decision-valid semantics.

As shown in Fig. 1, existing communication methods fall into two paradigms, yet *neither can fully satisfy both fidelity*. **Numerical messages** [7, 8, 9], typically Q-value lists, support basic bidirectional conversion due to their simple structure and limited semantic space. But even mild channel interference can cause catastrophic semantic distortion. **Neural messages** [10, 11, 12, 13, 14, 15, 16] enable improved robustness via tailored losses [17, 18]. However, the black-box encoding entirely fails to support conversion fidelity. Neither agents nor humans can recover the native semantics from neural messages. This inherent dilemma motivates the core question investigated in this paper: ***Can we design a novel message format in MARL that achieves high semantic fidelity?***

Based on the above requirement, Hyperdimensional Computing (HDC) [19, 20] theoretically offers an ideal mathematical framework. The quasi-orthogonality of high-dimensional random vectors

*Corresponding author.

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

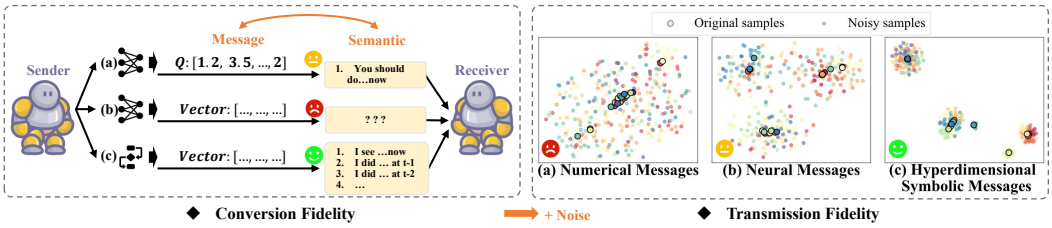



Figure 1: Comparison of Semantic Fidelity of Messages. The left shows the conversion fidelity in terms of the embedding process, message format, and semantic content. The right employs t-SNE to exhibit their transmission fidelity when exposed to noise.

inherently aligns with the noise-robust nature of "transmission fidelity". Meanwhile, its invertible algebraic operations enable "conversion fidelity" by allowing the extraction of original semantics.

However, **in practice**, applying HDC directly to multi-agent communication faces profound challenges that compromise the semantic fidelity. **For transmission fidelity**, standard HDC's flat encoding requires extremely high dimensions when handling MARL's multi-entity perception. With practical low-dimensional settings, HDC compromises the quasi-orthogonality and causes semantic entanglement. **For conversion fidelity**, HDC can only transmit static ego information and fails to provide adaptation for different receivers. Thus, sender-centric encoding leads to semantic misalignment between broadcast messages and receivers' local decision contexts. Moreover, HDC's native equal-weight bundle inevitably dilutes critical decision-valid semantics.

To address the above issues that undermine semantic fidelity, we propose a novel structured algebraic message format: Hyperdimensional Symbolic Messages (HSM) and its corresponding communication framework, named Herm. **To improve transmission fidelity**, Herm devises an entity-attribute-value structured encoding pipeline for the sender. Leveraging HDC's algebraic compositionality, this pipeline strictly aligns the perceptual logic of MARL with HDC operations. This alignment significantly enhances HSMs' semantic expressiveness and suppresses the semantic entanglement. **To ensure conversion fidelity**, Herm enhances the decision-valid semantic conversion for the receiver from two perspectives: semantic misalignment and semantic dilution. First, Herm introduces a receiver-centric ego-binding mechanism. Without any learnable parameters, this mechanism implicitly projects globally broadcasted HSMs into the unique local semantic subspace of receivers, thereby achieving second-order semantics contextualization. Second, Herm designs a similarity-aware dual-branch fusion network to dynamically evaluate the relevance of messages in both perception supplementation and behavior understanding.

Furthermore, we design an iterative decoding scheme for HSMs to realize explicit message-to-semantics conversion. Building upon this, we explore paradigms for integrating HSMs with Large Language Models (LLMs). Extensive experiments across multiple benchmarks (Hallway [11], LBF [21], SMAC [22], SMACv2 [23]) demonstrate Herm's robustness against noise (transmission fidelity) and the readability advantage (conversion fidelity). Our contributions are threefold:

- We pioneer the definition of semantic fidelity in multi-agent communication and the formalization of hyperdimensional symbolic messages.
- We design an entity-attribute-value structured encoding pipeline to eliminate semantic entanglement and guarantee transmission fidelity.
- We devise an ego-binding mechanism and similarity-aware dual-branch fusion to ensure decision-valid conversion fidelity.

## 2 Related Works

**Multi-Agent Robust Communication**. One scheme aims to identify malicious messages and mitigate their adverse effects. Early work employed Gaussian processes to model the message distribution, utilizing posterior probabilities for confidence evaluation [24]. $\mathfrak{R}$-MACRL [25] implemented a two-stage architecture that first detects malicious messages, then performs restoration. AME [26] utilized diverse combinations of ablated messages to mitigate the impact of attacks. ADMAC [27] leveraged

dynamic labeling and binary classifier training to identify malicious messages. Some works trained independent channel adversaries to enhance robustness through adversarial training [28, 29]. They typically involve additional communication stages that increase the training and deployment costs. Furthermore, some studies focus on the robustness of messages via graph information bottleneck [18] or multi-view certification [17]. In contrast, we make a fundamental departure by replacing neural messages with hyperdimensional symbolic messages, achieving robustness (transmission fidelity) and readability (conversion fidelity).

**Hyperdimensional Computing in RL**. Inspired by HDC's applications in federated learning [30] and classification tasks [31], some works have explored its adaptation to the RL domain. HDPG [32] and QHD [33] established the first policy-based and value-based hyperdimensional reinforcement learning paradigms (HRL). Subsequent works have extended HRL to robotic control [34], cyber security [35], and contextual bandit [36]. However, the application of HDC in multi-agent systems remains unexplored. To the best of our knowledge, this work presents the first integration of HDC with multi-agent communication.

## 3 Preliminaries

**Dec-POMDP.** Multi-agent collaboration is typically modeled as a decentralized partially observable Markov decision process (Dec-POMDP), defined by $< \mathcal{N}, \mathcal{S}, \mathcal{A}, \mathcal{O}, \mathcal{R}, \mathcal{P}, \gamma >$. $\mathcal{N}$ denotes multiple agents within the system. $\mathcal{S}$ and $\mathcal{A}$ represent the joint state and action space. $\mathcal{O}$ defines the mapping from joint states to local observations. $\mathcal{R}$ and $\mathcal{P}$ describe the reward and state transition function. $\gamma$ is the discount factor to scale the long-term return. In communication-enabled scenarios, Dec-POMDP can be extended with message space $\mathcal{M}$ to capture inter-agent communication. The objective in Dec-POMDP is to derive an optimal joint policy $\pi^*$ that maximizes the expected discounted return, formulated as $\pi^* = \text{argmax}_\pi \, \mathbb{E}_{\pi,\mathcal{P}} \left[\sum_{t=0}^{\infty} \gamma^t \mathcal{R}(s_t, \boldsymbol{a}_t)\right]$.

**Hyperdimensional Computing.** Hyperdimensional computing (HDC) relies on randomly hypervectors (HVs) from a high-dimensional space. Due to the quasi-orthogonality of HVs, HDC provides a robust approach to information encoding, which inspires our work. Each random HV can represent any single concept, such as an attribute, a numerical value, or a relational structure. Through two fundamental operations—**Bind** ($\otimes$) and **Bundle** ($\oplus$)—HVs can symbolize composite concepts with increasing complexity. Bind computes a composite representation that is dissimilar to each input. Bundle adds multiple items together and retains similarity with all inputs. A simple illustration is shown as: (1) $X =$ "Color" $\otimes$ "Red" $\rightarrow$ "Color = Red", (2) $Y =$ "Shape" $\otimes$ "Circle" $\rightarrow$ "Shape = Circle", (3) X $\oplus$ Y $\rightarrow$ "Red Circle". HVs can then be mapped to specific semantics through unbinding and similarity matching. Based on this, we devise a message-to-semantic transformation to achieve conversion fidelity. Here, we employ the Holographic Reduced Representations (HRR) [19] as the HDC framework. Likewise, our method is compatible with other frameworks (Appendix A.5).

## 4 Methods

The core of our work lies in two fundamental properties for semantic fidelity in communication: transmission fidelity and conversion fidelity, which cannot be traded off in existing numerical and neural message paradigms. In this paper, we design a novel message format, the hyperdimensional symbolic message (HSM), based on training-free hyperdimensional computing (HDC).

While HDC theoretically aligns with the two properties, standard HDC suffers from MARL-specific flaws that directly break the semantic fidelity. Hence, we further propose our method **Herm**, as shown in Fig.2, which makes the corresponding optimization from the view of the sender and receiver.

**On the sender side**, Herm focuses on eliminating the inherent semantic entanglement flaw of standard flat HDC encoding. We design a structured message encoding pipeline (Fig. 2 (b)) to balance semantic expressiveness and anti-interference capability with acceptable overhead. **On the receiver side**, Herm targets the two core barriers in standard HDC — semantic misalignment and semantic dilution. We devise an ego-binding mechanism and similarity-aware dual-branch fusion (Fig. 2 (c)) to ensure decision-valid conversion fidelity during multi-source fusion. Beyond the training, we propose an additional iterative message decoding process for explicit message-to-semantic conversion.

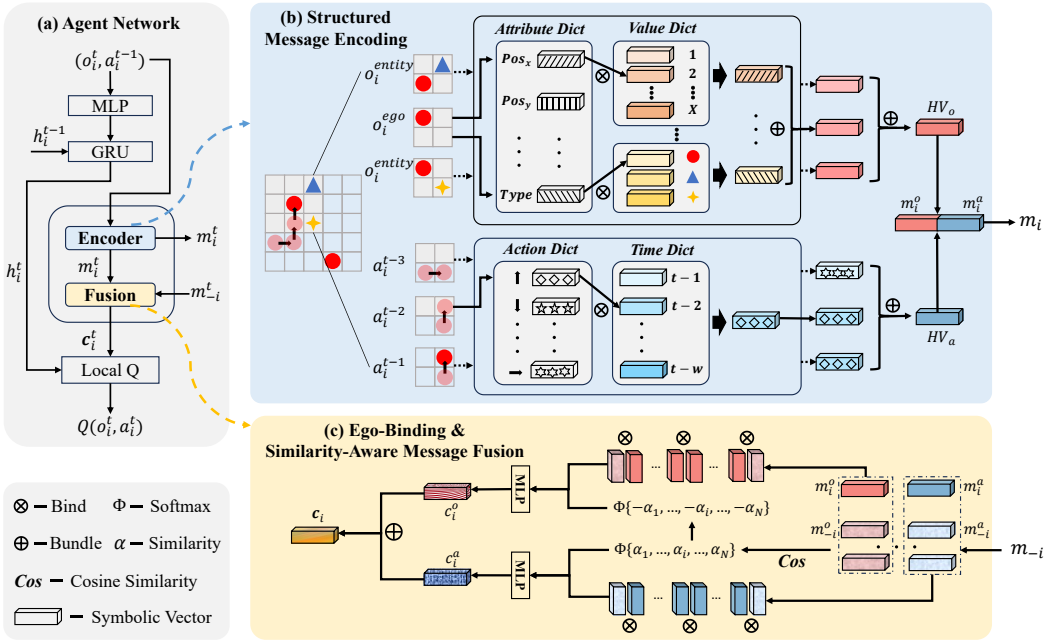



Figure 2: Framework of Herm. (a) The agent network comprises a trajectory-encoding module and a communication module, with their outputs serving as inputs to the local Q-function. (b) On the sender side, the agent encodes its entity-wise observations and historical actions into HSM. (c) On the receiving side, the agent performs ego-binding and fuses messages based on similarity metrics.

### 4.1 Structured Message Encoding for Transmission Fidelity

In the multi-agent system with limited communication bandwidth, standard flat HDC encoding could incur semantic entanglement. Complex multi-entity perception violates HDC's quasi-orthogonality requirement, which is fundamental to transmission fidelity. Hence, it is essential to explore the design of concise and efficient message structures for downstream decision-making.

We posit that communication enhances collaboration through two ways: **(1) supplementing subjective perception**, and **(2) interpreting teammates' objective behaviors**. Given the stability of observations versus the stochastic nature of actions, our raw message $\tilde{m}_i^t$ contains instantaneous environmental perceptions $o_i^t$ alongside temporally-smoothed actions $\{a_i^{t-w}, a_i^{t-w+1}, \cdots, a_i^{t-1}\}$ where $w$ denotes the sliding window. The original $\tilde{m}_i^t$ is then encoded into the symbolic message $m_i^t$ for broadcast through our structured encoding pipeline (Fig.2 (b)). Due to the refinement of $\tilde{m}_i^t$, we employ a modest hyperdimensional space (typically below 256), which incurs acceptable overhead.

#### 4.1.1 Perception Encoding via Observation Transformation.

Agents typically represent their perception of the local environment through entity-wise attribute-value pairs: $o_i = \{\langle id, attr, val\rangle | id \in \mathcal{E}, attr \in \mathcal{A}, val \in \mathcal{V}\}$. $\mathcal{E}$ represents all entities (e.g., teammates, enemies). $\mathcal{A}$ denotes the attribute dictionary (e.g., position, health status) and $\mathcal{V}$ represents the corresponding value dictionary (e.g., X-coordinate, HP). The attribute dictionary is first randomly instantiated in the hyperdimensional space and then processed via Schmidt orthogonalization to maintain strong orthogonality between attribute-aware $HV_{attr}$. For the value dictionary involving continuous quantities, we employ a discretization preprocessing step followed by unified encoding of the discretized intervals. Unlike mutually independent attributes, numerical values exhibit inherent correlations. We utilize cyclic replacement encoding to preserve these relationships. Let $val_{min}$ and $val_{max}$ denote the minimum and maximum values of an attribute. We first sample their hyperdimensional representations: $HV_{min} \sim \mathbb{R}^d, HV_{max} \sim \mathbb{R}^d$. For intermediate value $v \in (val_{min}, val_{max})$, its encoding is formulated as:

$$HV_v = \mathbf{\Theta}(HV_{max}, HV_{min}, \lceil \rho \cdot d \rceil), \tag{1}$$

where $\rho = (v - val_{min})/(val_{max} - val_{min})$ and $\mathbf{\Theta}(\cdot)$ replaces the first $\lceil \rho \cdot d \rceil$ dimensions of $HV_{min}$ with $HV_{max}$'s corresponding dimensions. Our framework then converts agent observations into HSM via four key operations: (i) attribute-value binding ($\otimes$), (ii) attribute bundling ($\oplus$), (iii) identity binding ($\otimes$), and (iv) entity-wise bundling ($\oplus$), defined as:

$$HV_O = \sum_{j}^{\mathcal{E}} \oplus \left[ HV_{id}^{j} \otimes \sum_{attr}^{\mathcal{A}} \oplus (HV_{attr} \otimes HV_v) \right], \tag{2}$$

where $\sum \oplus$ is the multi-bundle operation, and $HV_{id}^{j}$ denotes the identity vector for entity $j$. Although each agent bundles information about all entities, their encoding still differs due to partial observability. Notably, (1) entity-wise observation partitioning is a common practice in MARL research [37, 38, 39, 40] and introduces no manual prior knowledge; (2) while attributes carry semantic meaning for humans, they are treated only as observation dimensions by agents.

#### 4.1.2 Behavior Encoding via Action-Sequence Transformation.

For each action in the sliding window, it can be encoded into $HV_a$ through discretization and sampling in the hyperdimensional space. To model temporal dependencies in action sequences, we define a time dictionary following a construction similar to the numeric dictionary (Equ.1). Unlike perception encoding, actions are self-related and thus do not require identity binding and entity-wise bundling. The transformation can be represented as:

$$HV_A = \sum_{t'=t-w}^{t-1} \oplus (HV_{a_{t'}} \otimes HV_{t'}). \tag{3}$$

### 4.2 Receiver-Centric Fusion for Conversion Fidelity

Standard HDC can only transmit static sender-centric semantic information, but has no control over how this information is interpreted to serve receivers' downstream decisions. This limitation directly breaks the conversion fidelity, inspiring our designs for semantic alignment and preservation.

#### 4.2.1 Ego-Binding for Semantic Alignment

To align the broadcasted messages to the receivers' local context, Herm introduces the Ego-binding mechanism prior to message fusion. Receiver $i$ structurally binds the incoming message with its own encoding $m_i$ (the ego vector) using the circular convolution operator: $m'_{i,j} = m_j \otimes m_i$.

While the binding operation is a native primitive of HDC, existing works only use it for concept composition. Here, we utilize it for semantic contextualization. In the HRR framework, the binding operation functions as a dense multiplicative interaction. By binding $m_j$ and $m_i$, the mechanism inherently constructs a joint state representation at the feature level. This extracts second-order correlations between the sender's and receiver's semantics to support the downstream decision. Moreover, the ego-binding preserves the relative semantic topology between diverse messages, ensuring that $\cos(m_j \otimes m_i, m_k \otimes m_i) \approx \cos(m_j, m_k)$. This property preserves algebraic consistency between semantic contextualization and the subsequent fusion. The verification is in Appendix E.

#### 4.2.2 Similarity-Aware Message Fusion for Semantic Preservation

Received messages inevitably contain redundant information, and indiscriminate fusion will lead to the dilution of decision-valid messages. In contrast to neural vectors whose interrelationships can be captured through attention mechanisms [15], evidence theory [13], or mutual information [9], the correlation between hyperdimensional vectors is predominantly characterized by similarity measures.

We contend that similarity between observations serves as a crucial modulator in multi-agent communication fusion, with downstream effects on decision-making. **For supplementing environment perception**, dissimilar observations carry richer additional information and should be prioritized; **for understanding teammate behavior**, similarity in observations reflects analogous decision contexts that promote understanding and coordination between agents.

Hence, we propose a dual-branch message fusion mechanism that decouples observations from actions, processing each through distinct fusion pathways. Agent $i$ first computes similarity scores as:

$$\boldsymbol{\alpha}_i = \text{softmax}\{\alpha_1, \cdots, \alpha_i, \cdots, \alpha_N\}, \tag{4}$$

where $\alpha_{j\in[1,N]}$ denotes the cosine similarity between its own observational encoding $m_i^o$ and the received ones $m_j^o$. Then, the similarity scores are inverted, resulting in the dissimilarity scores:

$$\dot{\boldsymbol{\alpha}}_i = \text{softmax}\{-\alpha_1, \cdots, -\alpha_i, \cdots, -\alpha_N\}. \tag{5}$$

For the perception branch, we employ dissimilarity scores $\dot{\boldsymbol{\alpha}}$ for fusion, whereas similarity scores $\boldsymbol{\alpha}$ are utilized for the action branch. The fusion is expressed as Equ.6. The fused messages $\mathcal{M}_i^o$ and $\mathcal{M}_i^a$ undergo a mapping transformation into neural vectors $c_i^o$ and $c_i^a$, which are then aggregated to produce the final message representation $\boldsymbol{c}_i$. The overall training architecture adopts the value decomposition framework, where the message representation $\boldsymbol{c}_i$ and the agent's own trajectory encoding $h_i$ are fed into the local Q-network for decision-making. The pseudocode for training is in Appendix H.

$$\mathcal{M}_i^o = \sum_j^N \dot{\boldsymbol{\alpha}}_{\boldsymbol{i,j}} * (m_i^o \otimes m_j^o), \quad \mathcal{M}_i^a = \sum_j^N \boldsymbol{\alpha}_{\boldsymbol{i,j}} * (m_i^a \otimes m_j^a). \tag{6}$$

### 4.3 Iterative Message Decoding via Similarity

For the explicit conversion of HSM, we devise a transformation from messages to semantics. The process makes full use of the following properties of bind ($\otimes$), unbind ($\circledast$), and bundle ($\oplus$): (1) $\delta(HV_p \oplus HV_q, HV_p) \gg 0$, (2) $\delta(HV_p \circledast HV_q, HV_p) \approx 0$, and (3) $\delta(HV_p \otimes HV_q \circledast HV_q, HV_p) \gg 0$, where $\delta$ denotes the similarity function. Here, we specifically refer to the calculation of Holographic Reduced Representation. Proof is illustrated in Appendix B.

Since the perception encoding is obtained through entity-wise bundling in Equ.2, Property 1 implies that $HV_O$ maintains similarity with each individual entity encoding $HV_O^j$. When extracting information about a specific entity from observations, we perform (i) identity unbinding $HV_O \circledast HV_{id}$. With Properties 2 and 3, $HV_O^j$ is successfully recovered while $HV_O^{-j}$ is disrupted. Similarly, since $HV_O^j$ is calculated by attribute-wise bundling, extracting the value of a specific attribute requires performing (ii) attribute unbinding $HV_O^j \circledast HV_{attr}$. Then, we compute the similarity between the resulting $HV_v$ and the value dictionary $\mathcal{V}$ to obtain the final decoded value $v$. The whole process is represented as:

$$v = \underset{v\in\mathcal{V}}{\text{argmax}}\, \delta(HV_v, HV_O \circledast HV_{id} \circledast HV_{attr}). \tag{7}$$

Analogously to Equ.7, we can derive actions at any timestep $t' \in [t-w, t-1]$ from action-sequence encoding $HV_A$ through the unbinding operation of a specific timestep and the similarity matching:

$$a_{t'} = \underset{a\in\mathcal{A}}{\text{argmax}}\, \delta(HV_a, HV_A \circledast HV_{t'}). \tag{8}$$

Notably, the decoded semantics can be converted into natural language, providing a novel insight into the integration of multi-agent communication and large language models (LLMs). Previous work [41] attempts to map agents' messages to the LLM embedding space for implicit collaboration with LLMs, which is constrained by the task complexity and the foundation model. By contrast, we leverage HSM as an explicit bridging medium between agents and LLMs. It establishes a conversion pathway that spans message encoding, symbolic semantics, and natural language. We elaborate on the RL-LLM hybrid decision-making and corresponding experiments in Appendix G.

## 5 Experiments

To validate the proposed Herm, we conduct extensive experiments on standard MARL benchmarks, including SMAC [22], SMACv2 [23], Lbforaging [21], and Hallway [11]. As baselines, we adopt MAIC [9], T2MAC [13], CoDe [14], and TGCNET [15]. For transmission fidelity, we compare CroMAC [17] and several enhanced methods that integrate the AME [26] mechanism as robustness-oriented baselines. Then, we conduct semantic visualization to verify the conversion fidelity.

For fairness, we adopt a full communication topology for all methods across all benchmarks. We clarify that the attribute dictionary only provides a prior for explicit semantic conversion, and agents are completely agnostic to attribute semantics, thus bringing no performance gain.

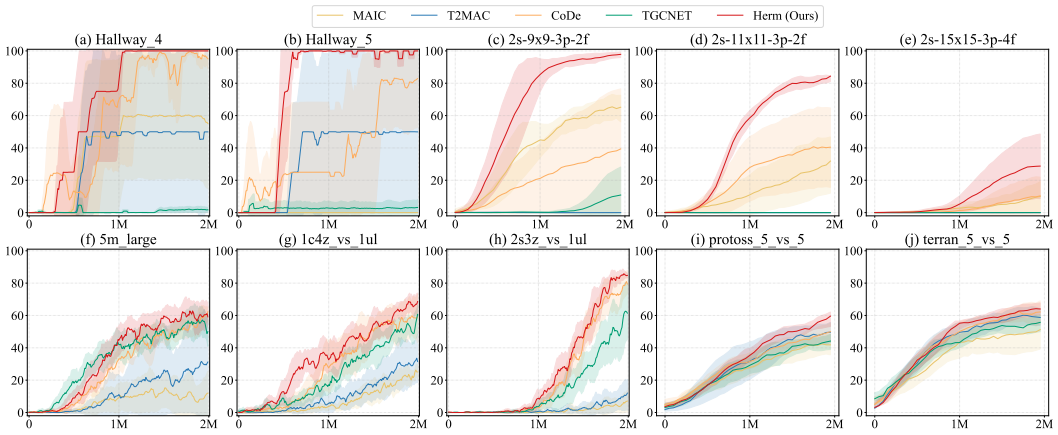



Figure 3: Performance in Noise-Free Environments. Each curve represents the average across four random seeds. (a)-(b) Hallway. (c)-(e) LBF . (f)-(h) SMAC. (i)-(j) SMACv2.

### 5.1 Performance in Noise-Free Environments

We first verify the efficacy of Herm for multi-agent collaboration. In Fig. 3, Herm consistently outperforms all baselines across four benchmarks in noise-free settings. For the Hallway task with a low-dimensional observation space, the orthogonal HSM encoding enables more explicit and efficient transmission of agents' decision contexts, leading to faster convergence and a higher task completion rate. For the partially observable LBF task, Herm achieves remarkable performance gains, especially in large-scale scenarios where baselines almost fail to converge. This is attributed to our similarity-aware dual-branch fusion mechanism, which precisely matches the task's core requirements of unobserved information supplementation and adjacent agent coordination. For communication-intensive SMAC and SMACv2 tasks with high randomness, Herm addresses two core limitations of baselines: 1) numerical messages [9, 13] can only deliver behavioral recommendations, failing to support early-stage joint exploration; 2) neural messages [14, 15] rely on black-box fusion and dilute observation variations caused by environmental stochasticity. In contrast, the binding operation in HSM amplifies discriminative observation features, and our interpretable fusion design provides explicit guidance for both exploration and coordination, thus achieving superior performance.

### 5.2 Performance in Noisy Environments

To verify the transmission fidelity, we conduct the robustness evaluation in two noisy settings.

**Random Noise.** In this setting, each inter-agent channel is subject to independent and identically distributed perturbations. We adopt the noise configuration from MAGI [42] as:

$$\hat{m}_i = m_i + \eta_r \cdot \phi \cdot \epsilon, \tag{9}$$

$\epsilon$ represents Gaussian noise, $\phi$ denotes amplitude computed as the mean of maximum values across all messages, and $\eta_r \in \{0, 0.5, 1, 1.5, 2\}$ is the scaling weight of noise. In Fig.4, Herm maintains superior robustness across varying noise, specifically in the linear (Hallway) or grid-based (LBF) scenarios. We attribute this to the well-preserved orthogonality in symbolic message space, where semantic and similarity relationships remain stable despite noise interference. Although the AME method can somewhat enhance the robustness of baselines, it is not suitable for small-scale scenarios such as Hallway, where each message passing remains essential. While CroMAC attempts to improve message-level robustness, its reliance on neural representations limits its noise resilience.

**Adversarial Noise.** We utilize two types of gradient-based adversarial attacks. (1) Fast Gradient Sign Method (FGSM) [43] is a one-step attack as Equ.10. $\eta_f \in \{0.3, 0.5\}$ is the scaling parameter and $\phi$ denotes the amplitude used in Equ.9. $f$ maximizes the cross-entropy loss between the noisy policy and the noise-free policy.

$$\hat{m}_i = m_i + \eta_f \cdot sign(\nabla_{m_i} f(m_i)) \cdot \phi, \tag{10}$$

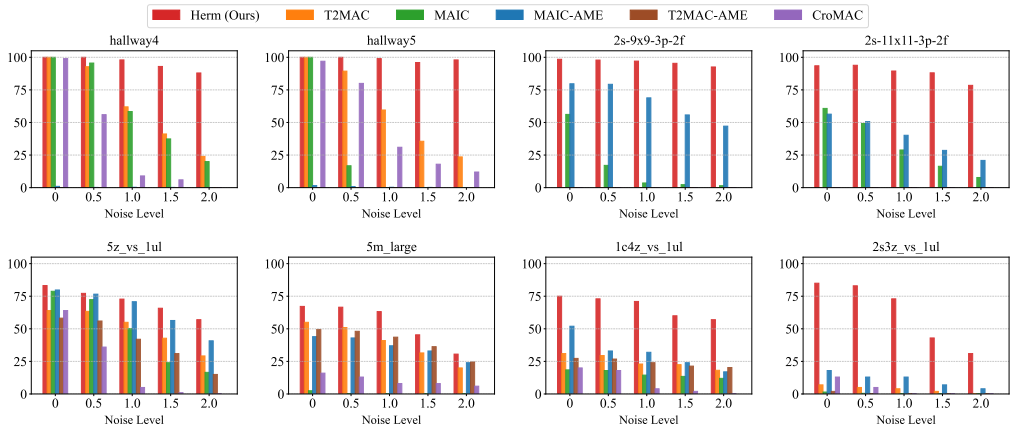



Figure 4: Robustness Evaluation in Randomly Noisy Environments.

Table 1: Robustness Evaluation in Adversarially Noisy Environments.

| Scenario | Method | No-Adv | FGSM (0.1) | FGSM (0.3) | FGSM (0.5) | PGD (5, 0.3) | PGD (5, 0.5) |
|---|---|---|---|---|---|---|---|
| **5m_large** | CroMAC | $16.3 \pm 0.2$ | $10.7 \pm 0.1$ | $3.1 \pm 0.4$ | $0.5 \pm 0.9$ | $9.2 \pm 0.3$ | $1.6 \pm 0.7$ |
| | MAIC-AME | $48.9 \pm 0.8$ | $44.1 \pm 0.6$ | $26.5 \pm 0.5$ | $15.1 \pm 0.2$ | $42 \pm 3.3$ | $\mathbf{36 \pm 2.8}$ |
| | Herm (Ours) | $\mathbf{67.2 \pm 0.5}$ | $\mathbf{58.3 \pm 0.4}$ | $\mathbf{31.8 \pm 0.1}$ | $\mathbf{18.6 \pm 0.3}$ | $\mathbf{43.5 \pm 1.4}$ | $33.9 \pm 3.2$ |
| **5z_vs_1ul** | CroMAC | $64.5 \pm 0.1$ | $10.5 \pm 0.8$ | $0.3 \pm 0.2$ | $0.8 \pm 0.5$ | $31.4 \pm 0.9$ | $24.7 \pm 0.1$ |
| | MAIC-AME | $77.8 \pm 0.3$ | $60.8 \pm 0.7$ | $34.2 \pm 0.6$ | $6.2 \pm 0.1$ | $\mathbf{65.2 \pm 2.4}$ | $37.5 \pm 3.3$ |
| | Herm (Ours) | $\mathbf{79.0 \pm 0.2}$ | $\mathbf{68.5 \pm 0.9}$ | $\mathbf{35.5 \pm 0.3}$ | $\mathbf{25.1 \pm 0.4}$ | $63.5 \pm 4.1$ | $\mathbf{58.5 \pm 3.7}$ |
| **Hallway 4** | CroMAC | $99.3 \pm 0.6$ | $16.2 \pm 0.5$ | $6.0 \pm 0.1$ | $0.0 \pm 0.8$ | $39.5 \pm 0.7$ | $17.3 \pm 0.2$ |
| | MAIC-AME | $5.3 \pm 0.4$ | $0.0 \pm 0.3$ | $0.0 \pm 0.6$ | $0.0 \pm 0.0$ | $0.0 \pm 0.0$ | $0.0 \pm 0.0$ |
| | Herm (Ours) | $\mathbf{100 \pm 0.0}$ | $\mathbf{99.0 \pm 0.2}$ | $\mathbf{95.5 \pm 0.4}$ | $\mathbf{92.7 \pm 0.5}$ | $\mathbf{100.0 \pm 0.0}$ | $\mathbf{97.5 \pm 1.6}$ |
| **Hallway 5** | CroMAC | $98.1 \pm 0.9$ | $96.9 \pm 0.7$ | $24.6 \pm 0.3$ | $13.5 \pm 0.2$ | $70.8 \pm 0.4$ | $44.1 \pm 0.6$ |
| | MAIC-AME | $3.6 \pm 0.5$ | $0.0 \pm 0.0$ | $0.0 \pm 0.9$ | $0.0 \pm 0.0$ | $0.0 \pm 0.0$ | $0.0 \pm 0.0$ |
| | Herm (Ours) | $\mathbf{100.0 \pm 0.0}$ | $\mathbf{98.0 \pm 0.6}$ | $\mathbf{93.5 \pm 0.8}$ | $\mathbf{90.2 \pm 0.7}$ | $\mathbf{100.0 \pm 0.0}$ | $\mathbf{100.0 \pm 0.0}$ |

(2) Projected Gradient Descent (PGD) [44] is a multi-step variant of FGSM, defined as

$$\hat{m}_i^{e+1} = \Pr_{\phi}\left(\hat{m}_i^e + \eta_f \cdot sign(\nabla_{\hat{m}_i^e} f(\hat{m}_i^e)) \cdot \frac{\phi}{5}\right), \tag{11}$$

where $e \in [0, 4]$ means 5-step attacks and $\hat{m}_i^0 = m_i$. Here, we adopt the L2 norm ball constraint $\Pr$.

Results in Tab.1 indicate that Herm almost achieves optimal robustness in two settings. Under PGD attacks, Herm exhibits slight performance fluctuations. This is because the entire communication process only involves a mapping network. This structure enables unobstructed backpropagation of adversarial gradients to original messages, thereby increasing the risk of communication loss.

### 5.3 Message-Semantics Transformation

The explicit conversion fidelity lies in the structural readability, which is guaranteed by the algebraic properties of hyperdimensional computing, and is completely independent of manual semantic labeling. Even without any labels, we can still accurately disassemble the composite message, clarify its information composition, and measure the similarity between different components. Manual labels are only an optional human-readable interface, which maps the disassembled message components to natural language for human inspection, and do not constitute the source of readability itself.

Fig.5 presents the decoding results for the observation components in messages within the LBF. The decoded semantics obtained through attribute-value combinations show near-perfect alignment with manually translated observation texts. Such alignment verifies that hyperdimensional symbolic messages possess excellent conversion fidelity. The conversion for decision-valid semantics is then verified via subsequent ablation experiments. Furthermore, we observe that the accuracy of semantic transformation improves with the expansion of the hyperdimensional space. More semantic transformations and LLM collaboration are presented in Appendix C,D, and G.

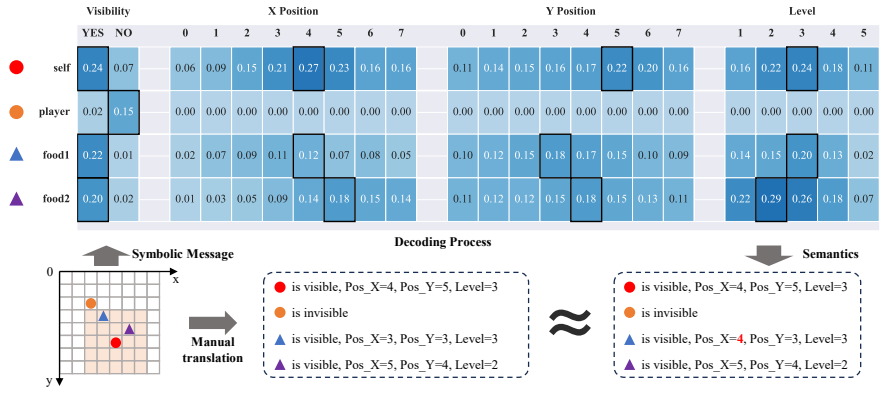
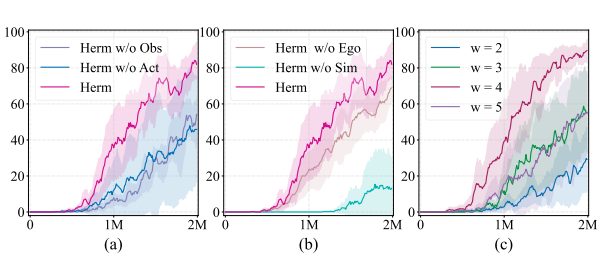
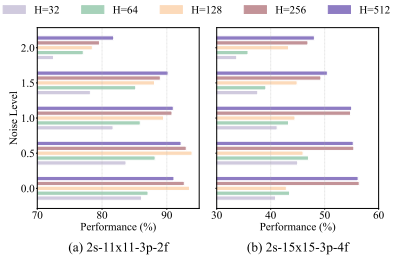



Figure 5: Message-Semantic Transformation in LBF. The upper displays entity-wise similarity measurements, while the lower shows: the raw observations, manually translated observation texts, and the decoded semantics. The black box highlights the value exhibiting maximum similarity.




(a) Analysis of Communication Components in MMM2.




(b) Performance with Varying $\mathcal{H}$ in LBF.

Figure 6: Ablation and Parameter Analysis on the Communication Process.

### 5.4 Ablation over Communication Process

We first analyze components in the sender's message encoding, including 1) observation and 2) action sequence. Fig.6a (a) demonstrates the necessity of both components. Then, we conduct ablations in the receiver's message fusion. In Fig.6a (b), eliminating similarity metrics leads to catastrophic performance degradation. Massive messages result in information redundancy, making it impossible for agents to extract decision-valid semantics. Similarity offers explicit fusion guidance in terms of perception supplementation and behavior coordination. Moreover, without ego-binding, Herm suffers from semantic misalignment, which increases the burden on the downstream network to capture correlations. It further results in reduced sample efficiency and performance.

### 5.5 Parameter Analysis

**Action Sequence Length $w$.** Fig.6a (c) shows significant variations in behavioral intention transmission with different action sequence lengths. When $w = 2$, limited action sequences hinder effective inter-agent behavioral understanding. Performance shows marked improvement with increasing $w$. However, excessively large $w$ can compromise the orthogonality in message encoding, consequently degrading collaborative performance. In our study, we adopt $w = 4$ as the default setting.

**Message Dimension $\mathcal{H}$.** The impact of $\mathcal{H}$ is quantified in Fig.6b. When $\mathcal{H} = 32$, Herm performs poorly in both noiseless and noisy settings. Distinct semantic symbols become entangled in the constrained hyperdimensional space, leading to the ambiguous perception of the environment and teammates. As $\mathcal{H}$ increases, Herm exhibits enhanced robustness. Moreover, the required dimensionality $\mathcal{H}$ varies in different scenarios. For simpler tasks (Fig.6b (a)), $\mathcal{H} = 128$ suffices to handle most conditions, whereas for complex tasks (Fig.6b (b)), $\mathcal{H} = 256$ exhibits significant advantages. In fact, a larger $\mathcal{H}$ delivers better performance in robustness (Fig.6b) and readability (see Appendix C).

Table 2: Spearman Correlation Coefficient $\rho$ under Noise (mean $\pm$ std)

| Task | Encoding | noise=0 | noise=0.5 | noise=1.0 | noise=1.5 | noise=2.0 |
|---|---|---|---|---|---|---|
| $9 \times 9$ | HDC | $\mathbf{0.878 \pm 0.017}$ | $\mathbf{0.872 \pm 0.020}$ | $\mathbf{0.785 \pm 0.034}$ | $\mathbf{0.678 \pm 0.047}$ | $\mathbf{0.615 \pm 0.050}$ |
| $9 \times 9$ | Neural | $0.862 \pm 0.039$ | $0.825 \pm 0.045$ | $0.720 \pm 0.075$ | $0.595 \pm 0.070$ | $0.440 \pm 0.048$ |
| $11 \times 11$ | HDC | $\mathbf{0.882 \pm 0.028}$ | $\mathbf{0.875 \pm 0.020}$ | $\mathbf{0.785 \pm 0.028}$ | $\mathbf{0.705 \pm 0.030}$ | $\mathbf{0.592 \pm 0.038}$ |
| $11 \times 11$ | Neural | $0.858 \pm 0.040$ | $0.815 \pm 0.022$ | $0.695 \pm 0.072$ | $0.510 \pm 0.118$ | $0.385 \pm 0.082$ |

### 5.6 Transmission Fidelity Quantification

We adopt the Spearman correlation coefficient ($\rho$) as the quantitative metric [45]. The Spearman correlation coefficient is a nonparametric statistic that quantifies the strength and direction of monotonic relationships between two variables by evaluating their rank order. We calculate the Spearman correlation between the similarity ranking of raw observations before encoding and the similarity ranking of encoded messages after transmission. The quantitative results are summarized in Table 2. A $\rho$ value closer to 1 indicates that the HSM module better preserves the original semantic similarity relationships. Experimental results show that such similarity relationships can be stably maintained under various noise perturbations, which verifies the strong transmission fidelity of the HSM module.

## 6 Conclusions

This paper pioneers the definition of semantic fidelity in multi-agent communication. Based on this, the hyperdimensional symbolic message and its communication framework are proposed to enhance both transmission fidelity and conversion fidelity. However, this paper still has certain limitations to be addressed. For one thing, the encoding module is decoupled from the policy optimization; future work can explore an end-to-end joint training paradigm from semantic representation to decision-making. For another, the per-dimension message encoding scheme lacks flexibility. We will therefore design an automatic fine-grained attribute selection and dynamic encoding mechanism, which allows agents to decide what to communicate.

## 7 Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No. 62576029).

## References

[1] Enguang Yao, Qidong Liu, Jialu Sun, Kaibo Huang, and Mingliang Xu. Direct: Decentralized intention and latent rule emergence for multi-agent cooperation under intermittent communication. In *Proceedings of the Thirty-Fifth International Joint Conference on Artificial Intelligence, IJCAI*, volume 26, pages 437–445, .

[2] Hao Wu, Shoucheng Song, Chang Yao, Sheng Han, Huaiyu Wan, Youfang Lin, and Kai Lv. Think how your teammates think: Active inference can benefit decentralized execution. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pages 29749–29757, 2026.

[3] Ziyue Chu and Leonardo Stella. Value aggregation with uncertainty in online decentralized marl. In *Forty-third International Conference on Machine Learning*, 2026.

[4] Chang Yao, Youfang Lin, Shoucheng Song, Hao Wu, Yuqing Ma, Sheng Han, and Kai Lv. From general relation patterns to task-specific decision-making in continual multi-agent coordination. In James Kwok, editor, *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence, IJCAI-25*, pages 6821–6829. International Joint Conferences on Artificial Intelligence Organization, 8 2025. doi: 10.24963/ijcai.2025/759. URL `https://doi.org/10.24963/ijcai.2025/759`. Main Track.

[5] Junyong Jiang, Xin Yuan, Longhe Lin, Songze Li, and Lu Dong. Budget-efficient attacks and robustness training for cooperative marl. In *Forty-third International Conference on Machine Learning*, 2026.

[6] Chang Yao, Youfang Lin, Shoucheng Song, Hao Wu, Shengkun Yang, Yuqing Ma, and Kai Lv. Role-level inductive bias for cross-task generalization in multi-agent reinforcement learning. In *Forty-third International Conference on Machine Learning*, .

[7] Sai Qian Zhang, Qi Zhang, and Jieyu Lin. Efficient communication in multi-agent reinforcement learning via variance based control. *Advances in neural information processing systems*, 32, 2019.

[8] Sai Qian Zhang, Qi Zhang, and Jieyu Lin. Succinct and robust multi-agent communication with temporal message control. *Advances in neural information processing systems*, 33:17271–17282, 2020.

[9] Lei Yuan, Jianhao Wang, Fuxiang Zhang, Chenghe Wang, Zongzhang Zhang, Yang Yu, and Chongjie Zhang. Multi-agent incentive communication via decentralized teammate modeling. In *Proceedings of the AAAI conference on artificial intelligence*, volume 36, pages 9466–9474, 2022.

[10] Abhishek Das, Théophile Gervet, Joshua Romoff, Dhruv Batra, Devi Parikh, Mike Rabbat, and Joelle Pineau. Tarmac: Targeted multi-agent communication. In *International Conference on machine learning*, pages 1538–1546. PMLR, 2019.

[11] Tonghan Wang, Jianhao Wang, Chongyi Zheng, and Chongjie Zhang. Learning nearly decomposable value functions via communication minimization. In *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*, 2020.

[12] Xinran Li and Jun Zhang. Context-aware communication for multi-agent reinforcement learning. In *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems*, pages 1156–1164, 2024.

[13] Chuxiong Sun, Zehua Zang, Jiabao Li, Jiangmeng Li, Xiao Xu, Rui Wang, and Changwen Zheng. T2mac: Targeted and trusted multi-agent communication through selective engagement and evidence-driven integration. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 15154–15163, 2024.

[14] Shoucheng Song, Youfang Lin, Sheng Han, Chang Yao, Hao Wu, Shuo Wang, and Kai Lv. Code: Communication delay-tolerant multi-agent collaboration via dual alignment of intent and timeliness. In *AAAI-25, Sponsored by the Association for the Advancement of Artificial Intelligence*, pages 23304–23312, 2025.

[15] Zhuohui Zhang, Bin He, Bin Cheng, and Gang Li. Bridging training and execution via dynamic directed graph-based communication in cooperative multi-agent systems. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 23395–23403, 2025.

[16] Hao Wu, Shoucheng Song, Chang Yao, Haoyu Wang, Huaiyu Wan, Youfang Lin, and Kai Lv. Group perspective matters: Regulating debate relationships can mitigate blind conformity in multi-agent debate, 2026. URL `https://arxiv.org/abs/2608.03648`.

[17] Lei Yuan, Tao Jiang, Lihe Li, Feng Chen, Zongzhang Zhang, and Yang Yu. Robust cooperative multi-agent reinforcement learning via multi-view message certification. *Science China Information Sciences*, 67(4): 142102, 2024.

[18] Shifei Ding, Wei Du, Ling Ding, Lili Guo, and Jian Zhang. Learning efficient and robust multi-agent communication via graph information bottleneck. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 17346–17353, 2024.

[19] T.A. Plate. Holographic reduced representations. *IEEE Transactions on Neural Networks*, 6(3):623–641, 1995.

[20] Pentti Kanerva. Binary spatter-coding of ordered k-tuples. In *Artificial Neural Networks — ICANN 96*, pages 869–873. Springer Berlin Heidelberg, 1996.

[21] Georgios Papoudakis, Filippos Christianos, Lukas Schäfer, and Stefano V. Albrecht. Benchmarking multi-agent deep reinforcement learning algorithms in cooperative tasks. In *Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks (NeurIPS)*, 2021.

[22] Mikayel Samvelyan, Tabish Rashid, Christian Schröder de Witt, Gregory Farquhar, Nantas Nardelli, Tim G. J. Rudner, Chia-Man Hung, Philip H. S. Torr, Jakob N. Foerster, and Shimon Whiteson. The starcraft multi-agent challenge. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*, pages 2186–2188, 2019.

[23] Benjamin Ellis, Jonathan Cook, Skander Moalla, Mikayel Samvelyan, Mingfei Sun, Anuj Mahajan, Jakob N. Foerster, and Shimon Whiteson. Smacv2: An improved benchmark for cooperative multi-agent reinforcement learning. In *Advances in Neural Information Processing Systems 36: Annual Conference on Neural Information Processing Systems*, 2023.

[24] Rupert Mitchell, Jan Blumenkamp, and Amanda Prorok. Gaussian process based message filtering for robust multi-agent cooperation in the presence of adversarial communication. *CoRR*, abs/2012.00508, 2020.

[25] Wanqi Xue, Wei Qiu, Bo An, Zinovi Rabinovich, Svetlana Obraztsova, and Chai Kiat Yeo. Mis-spoke or mis-lead: Achieving robustness in multi-agent communicative reinforcement learning. In *International Conference on Autonomous Agents and Multiagent Systems*, pages 1418–1426, 2022.

[26] Yanchao Sun, Ruijie Zheng, Parisa Hassanzadeh, Yongyuan Liang, Soheil Feizi, Sumitra Ganesh, and Furong Huang. Certifiably robust policy learning against adversarial multi-agent communication. In *The Eleventh International Conference on Learning Representations*, 2022.

[27] Lebin Yu, Yunbo Qiu, Quanming Yao, Yuan Shen, Xudong Zhang, and Jian Wang. Robust communicative multi-agent reinforcement learning with active defense. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 17575–17582, 2024.

[28] Lei Yuan, Feng Chen, Zongzhang Zhang, and Yang Yu. Communication-robust multi-agent learning by adaptable auxiliary multi-agent adversary generation. *Frontiers of Computer Science*, 18(6):186331, 2024.

[29] Xuyan Ma, Yawen Wang, Junjie Wang, Xiaofei Xie, Boyu Wu, Shoubin Li, Fanjiang Xu, and Qing Wang. Robust multi-agent communication based on decentralization-oriented adversarial training. *CoRR*, abs/2504.21278, 2025.

[30] Quanling Zhao, Kai Lee, Jeffrey Liu, Muhammad Huzaifa, Xiaofan Yu, and Tajana Rosing. Fedhd: federated learning with hyperdimensional computing. In *Proceedings of the 28th annual international conference on mobile computing and networking*, pages 791–793, 2022.

[31] Mohsen Imani, Chenyu Huang, Deqian Kong, and Tajana Rosing. Hierarchical hyperdimensional computing for energy efficient classification. In *Proceedings of the 55th Annual Design Automation Conference*, pages 1–6, 2018.

[32] Yang Ni, Mariam Issa, Danny Abraham, Mahdi Imani, Xunzhao Yin, and Mohsen Imani. Hdpg: Hyperdimensional policy-based reinforcement learning for continuous control. In *Proceedings of the 59th ACM/IEEE Design Automation Conference*, pages 1141–1146, 2022.

[33] Yang Ni, Danny Abraham, Mariam Issa, Yeseong Kim, Pietro Mercati, and Mohsen Imani. Efficient off-policy reinforcement learning via brain-inspired computing. In *Proceedings of the Great Lakes Symposium on VLSI 2023*, pages 449–453, 2023.

[34] Hyukjun Kwon, Kangwon Kim, Junyoung Lee, Hyunsei Lee, Jiseung Kim, Jinhyung Kim, Taehyung Kim, Yongnyeon Kim, Yang Ni, Mohsen Imani, et al. Brain-inspired hyperdimensional computing in the wild: Lightweight symbolic learning for sensorimotor controls of wheeled robots. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5176–5182. IEEE, 2024.

[35] Mariam Ali Issa, Hanning Chen, Junyao Wang, and Mohsen Imani. Cyberrl: Brain-inspired reinforcement learning for efficient network intrusion detection. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 44(1):241–250, 2025.

[36] Marco Angioli, Antonello Rosato, Marcello Barbirotta, Rocco Martino, Francesco Menichelli, and Mauro Olivieri. HD-CB: the first exploration of hyperdimensional computing for contextual bandits problems. *CoRR*, abs/2501.16863, 2025.

[37] Fuxiang Zhang, Chengxing Jia, Yi-Chen Li, Lei Yuan, Yang Yu, and Zongzhang Zhang. Discovering generalizable multi-agent coordination skills from multi-task offline data. In *The Eleventh International Conference on Learning Representations*, 2023.

[38] Rongjun Qin, Feng Chen, Tonghan Wang, Lei Yuan, Xiaoran Wu, Yipeng Kang, Zongzhang Zhang, Chongjie Zhang, and Yang Yu. Multi-agent policy transfer via task relationship modeling. *Science China Information Sciences*, 67(8):182101, 2024.

[39] Shunyu Liu, Jie Song, Yihe Zhou, Na Yu, Kaixuan Chen, Zunlei Feng, and Mingli Song. Interaction pattern disentangling for multi-agent reinforcement learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(12):8157–8172, 2024.

[40] Chang Yao, Youfang Lin, Shoucheng Song, Hao Wu, Yuqing Ma, Sheng Han, and Kai Lv. From general relation patterns to task-specific decision-making in continual multi-agent coordination. In *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence*, pages 6821–6829, 2025.

[41] Huao Li, Hossein Nourkhiz Mahjoub, Behdad Chalaki, Vaishnav Tadiparthi, Kwonjoon Lee, Ehsan Moradi Pari, Charles Lewis, and Katia Sycara. Language grounded multi-agent reinforcement learning with human-interpretable communication. *Advances in Neural Information Processing Systems*, 37: 87908–87933, 2024.

[42] Shifei Ding, Wei Du, Ling Ding, Jian Zhang, Lili Guo, and Bo An. Robust multi-agent communication with graph information bottleneck optimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(5):3096–3107, 2024.

[43] Ian J Goodfellow, Jonathon Shlens, and Christian Szegedy. Explaining and harnessing adversarial examples. *arXiv preprint arXiv:1412.6572*, 2014.

[44] Aleksander Madry, Aleksandar Makelov, Ludwig Schmidt, Dimitris Tsipras, and Adrian Vladu. Towards deep learning models resistant to adversarial attacks. *arXiv preprint arXiv:1706.06083*, 2017.

[45] C. Spearman. The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1):72–101, 1904. ISSN 00029556. URL `http://www.jstor.org/stable/1412159`.

[46] George Papadopoulos, Andreas Kontogiannis, Foteini Papadopoulou, Chaido Poulianou, Ioannis Koumentis, and George Vouros. An extended benchmarking of multi-agent reinforcement learning algorithms in complex fully cooperative tasks. In *Proceedings of the 24th International Conference on Autonomous Agents and Multiagent Systems*, page 1613–1622, 2025. ISBN 9798400714269.

[47] Tony A Plate. *Holographic Reduced Representation: Distributed representation for cognitive structures*, volume 150. CSLI Publications Stanford, 2003.

[48] Jan Gosmann and Chris Eliasmith. Vector-derived transformation binding: An improved binding operation for deep symbol-like processing in neural networks. *Neural Computation*, 31(5):849–869, 05 2019. ISSN 0899-7667.

[49] Ryan Lowe, Yi I Wu, Aviv Tamar, Jean Harb, OpenAI Pieter Abbeel, and Igor Mordatch. Multi-agent actor-critic for mixed cooperative-competitive environments. *Advances in neural information processing systems*, 30, 2017.

[50] Chao Yu, Akash Velu, Eugene Vinitsky, Jiaxuan Gao, Yu Wang, Alexandre Bayen, and Yi Wu. The surprising effectiveness of ppo in cooperative multi-agent games. *Advances in neural information processing systems*, 35:24611–24624, 2022.

# A Supplementary Experiments

## A.1 SMAC Configuration

In the experiment, we utilize the communication-intensive tasks proposed in [14]. As illustrated in Fig. 7, the initial positions of units are randomly generated with specific terrain obstacles. Agents are required to explore their surrounding environment to locate allies and enemies. Given that both sides boast comparable military strength and enemies exhibit a tendency to cluster, solo agents are vulnerable to elimination, ultimately resulting in mission failure due to force imbalances. Therefore, agents must wait to form clusters before engaging in combat under conditions of force balance; otherwise, successful mission completion becomes exceedingly difficult. Hence, the imperative of inter-agent communication is further underscored, accompanied by heightened requirements for exploration and coordination. This renders these tasks highly suitable for assessing communication-aware multi-agent reinforcement learning algorithms.

## A.2 Experiment Parameter

In the experiments, both the baselines and our proposed approach are implemented using the PyMARLzoo+ architecture [46], retaining most default parameters (see Tab.3). In the experiment, the CPU used is AMD Ryzen 9 9950X, and the GPU is NVIDIA GeForce RTX 4090.

Table 4 presents the parameters specific to our proposed Herm. Due to the simplicity of the hallway task, we uniformly employ $\mathcal{H} = 64$ and $w = 1$ configurations. For most LBF tasks, we maintain $\mathcal{H} = 64$ and $w = 1$, while increasing to $H = 256$ for more complex 15×15 maps. For SMAC with its greater entity and attribute complexity, we adopt $\mathcal{H} = 256$ with $w = 3$ or 4 to adequately represent the expansive observation space.




Figure 7: Communication-Intensive Tasks in SMAC [14]. Enemies are highlighted in blue circles, while allies are marked in red.

Table 3: Default Parameters

| Parameter | Value |
|---|---|
| Optimizer | Adam |
| Learning Rate | $5 \times 10^{-4}$ |
| Batch Size | 32 |
| Discount Factor | 0.99 |
| GRU Hidden State Dim | 64 |
| Mixing Hidden Dim | 32 |
| Hypernet Embedding Dim | 32 |

To establish a robust communication baseline, we implement the message ablation mechanism from AME, leveraging multi-round voting to mitigate noise. Specifically, during training, we retain $K$ out of $N-1$ message sources for policy learning, where $K$ is set to $N-2$ or $N-3$. During testing, we perform $P=10$ ablation trials for stabilized evaluation.

### A.3 Analysis of Model Size and Efficiency

We compare the model parameter counts and the time required for a single complete interaction across different algorithms in Tab.5. Results demonstrate that Herm benefits from hyperdimensional symbolic messages and eliminates the need for neural networks in the encoding and fusion processes, resulting in a minimal model parameter count. Herm's parameters consist of two components: the first encompasses the GRU and MLP networks for processing agents' local observation, and the second corresponds to the single-layer mapping network after message fusion. In contrast, other baseline methods typically require additional message encoding networks, while adopting attention mechanisms and graph-based aggregation schemes for message fusion, which entails a considerably larger parameter overhead.

In terms of decision-making efficiency, Herm is limited by the programmatic encoding of hyperdimensional computing, exhibiting a slight disadvantage. However, the static encoding process allows for the reuse of previous results, trading space for time—similar to the KV cache mechanism. Results show that storing and reusing the binding outputs of attribute-value pairs can significantly reduce encoding time, bringing it close to baselines.

Table 4: Method-Specific Parameters

| **Symbol** | **Description** | **Value** |
|---|---|---|
| $\mathcal{H}$ | Dimension of the hyperdimensional space | $\{64, 128, 256\}$ |
| $w$ | Length of historical action sequences | $\{1, 3, 4\}$ |
| $Bin$ | Number of discretized partitions | $\{5, 10\}$ |

Table 5: Comparison on Model Size and Efficiency

| | MAIC | T2MAC | CoDe | TGCNET | Herm (Ours) | Herm + Cache |
|---|---|---|---|---|---|---|
| Params (K) | 50.79 | 50.70 | 49.39 | 827.01 | **46.88** | **46.88** |
| Times (s) | 1.481 | 1.528 | **1.363** | 1.656 | 1.993 | 1.392 |

### A.4 Additional Comparison on Robustness

We further supplement additional robustness tests and incorporate the baseline, MAGI [18]. Since MAGI is not open-sourced, some results are missing. Results in Fig.8 demonstrate that our proposed Herm still achieves the optimal robust performance in scenarios with a large number of entities.

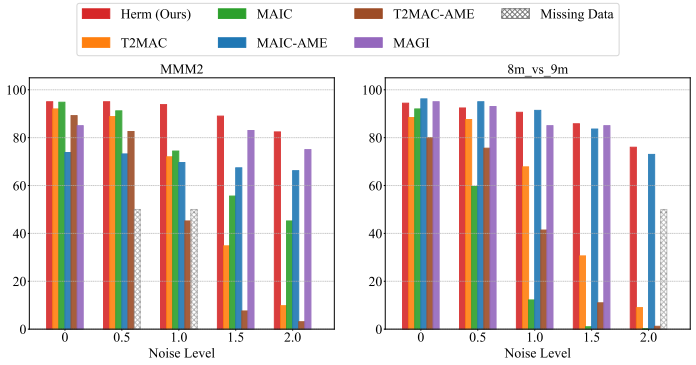



Figure 8: Additional Robustness Comparison.

### A.5 Scalability of Hyperdimensional Computing Framework

Different HDC frameworks differ primarily in their implementations of symbolic operations. We adopt the most conventional HRR framework in our experiments. We provide supplementary performance comparisons of different computational frameworks in Fig.9.

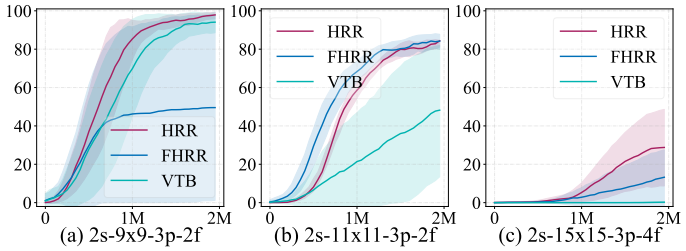



Figure 9: Performance Comparison with Different HDC Framework.

**(1) Holographic Reduced Representation (HRR)** [19] is a technique in hyperdimensional computing that represents compositional structures using circular convolution in distributed representations, allowing for the binding and unbinding of vectors to combine concepts and retrieve them given the combined result. Its key operations involve circular convolution (binding) and circular correlation (unbinding): for vectors $\mathbf{A}$ and $\mathbf{B}$, binding is calculated as:

$$(\mathbf{A} \otimes \mathbf{B})_k = \sum_{i=0}^{D-1} A_i B_{k-i} \bmod D, \tag{12}$$

where $k$ refers to the value of the k-th dimension in the output, and $d$ is the size of the vector space. The unbinding uses a similar process, formulated as:

$$(\mathbf{A} \circledast \mathbf{B})_k = \sum_{i=0}^{D-1} A_i B_{i+k} \bmod D. \tag{13}$$

HRR advantages include 1) representing numerous key-value pairs in a fixed-size vector, 2) efficient associative memory for tasks like natural language processing, and 3) preservation of semantic similarity. However, its commutative binding limits the representation of non-commutative relationships and complex hierarchical structures.

**(2) Fourier Holographic Reduced Representation (FHRR)** [47] extends HRR by performing operations in the frequency domain, with hypervectors of the form $h = (e^{i\theta_1}, \ldots, e^{i\theta_d})$. Bundling and binding are done via element-wise addition and multiplication, respectively. FHRR leverages Fourier transform properties to simplify binding, making it a simple element-wise multiplication in the frequency domain for lower computational complexity $O(D \log D)$, while also supporting sparse vectors well.

**(3) Vector-Derived Transformation Binding (VTB)** [48] proposes a new binding operation that represents relationships as linear transformations in vector space. For vectors $\mathbf{A}$ and $\mathbf{B}$, binding is achieved through a linear transformation matrix W as $\mathbf{A} \otimes \mathbf{B} = \mathbf{W} \cdot (\mathbf{A} + \mathbf{B})$ where $+$ denotes concatenation. It allows flexible vector binding with a learnable transformation matrix $W$, and supports encoding hierarchical structures such as tree-like or knowledge graph relationships. VTB has advantages on high flexibility in representing complex relationships, suitability for neural implementations due to learnable matrices, and minimal influence on vector length beneficial for neural implementation.

### A.6 Extension with Numerical Messages

This section further investigates the integration schemes between hyperdimensional computing (HDC) and the numerical message. Since HDC typically does not involve neural networks, we attempt to realize integration from the perspectives of post-processing. Notably, the design presented in this section is merely intended as an intuitive exploration and verification; more in-depth integration and design will be reserved for future research.

Numerical messages utilized in [7, 8, 9, 13] are usually presented in the form of Q-tables. The sender conveys action suggestions as messages to the receiver, which synthesizes all suggestions to make its own decision. This indicates that a numerical message is formally an array. Therefore, we propose to use hyperdimensional symbolic messages (HSM) to represent this array, thereby enhancing its robustness against noise.

Herein, we propose a simple conversion method to transform an array into a hyperdimensional vector. First, attributes are defined as positive/negative signs and finite digit positions, and the value set covers positive/negative signs and integers 0–9. With this division, the sender adopts the designed hyperdimensional encoding pipeline (see Section 4.1) for conversion and transmission, while the receiver recovers the array from the vector through iterative decoding (see Section 4.3).

Table 6: Performance with HSM-Enhancement

| Task | $\mathcal{H}$ | Noise Level | | | | |
|---|---|---|---|---|---|---|
| | | 0 | 0.5 | 1.0 | 1.5 | 2.0 |
| Hallway-4 | MAIC | $100 \pm 0.0$ | $95.6 \pm 1.7$ | $58.4 \pm 1.9$ | $37.4 \pm 3.3$ | $20.0 \pm 3.9$ |
| | MAIC-E ($\mathcal{H} = 256$) | $98.4 \pm 0.5$ | $96.8 \pm 0.8$ | $94.5 \pm 0.7$ | $64.5 \pm 2.6$ | $44.6 \pm 4.1$ |
| | MAIC-E ($\mathcal{H} = 512$) | $100.0 \pm 0.0$ | $100.0 \pm 0.0$ | $100.0 \pm 0.0$ | $100.0 \pm 0.0$ | $85.5 \pm 1.3$ |
| | MAIC-E ($\mathcal{H} = 1024$) | $100.0 \pm 0.0$ | $100.0 \pm 0.0$ | $100.0 \pm 0.0$ | $100.0 \pm 0.0$ | $100.0 \pm 0.0$ |
| 5z_vs_1ul | MAIC | $78.8 \pm 1.2$ | $72.4 \pm 1.9$ | $50.2 \pm 2.6$ | $24.0 \pm 4.2$ | $16.6 \pm 4.1$ |
| | MAIC-E ($\mathcal{H} = 1024$) | $70.6 \pm 0.8$ | $69.8 \pm 1.3$ | $69.5 \pm 1.6$ | $56.5 \pm 3.5$ | $34.6 \pm 3.7$ |

We select MAIC [9] as the baseline, introduce HSM with different dimensions as the transmission form on the same model, and conduct experimental comparisons. It is shown in Tab.6 that hyperdimensional conversion significantly improves the noise robustness of MAIC in the small-scale task

(hallway), and a positive correlation exists between this improvement and the message dimension $\mathcal{H}$. This is because HSM distributes numerical information across each dimension of the hyper-dimensional space, meaning that other dimensions can mitigate significant noise interference on certain dimensions. By contrast, for more complex tasks (5z_vs_1ul), the Q-table scale increases significantly, posing challenges for accurately converting the value of each digit and requiring a higher dimension. Limited by the conversion accuracy, MAIC-E exhibits a slight performance decline under no-noise conditions, yet it exhibits stronger stability under noisy conditions. This phenomenon can be attributed to the preservation of similarity by HSM under interference, whereby stable mapping to the corresponding digit values is maintained despite numerical alterations in dimensions.

### A.7 Extension with Policy-Based Frameworks

This section further integrates Herm with policy-based methods to verify its scalability. Here, MADDPG [49] and MAPPO [50] are selected as baselines, and experiments are conducted in the MPE environment under continuous-action settings. Experimental results in Fig.10 show that Herm can effectively adapt to the two types of policy-based methods and achieve certain improvements for each. Furthermore, these results also verify that hyperdimensional encoding is equally effective for continuous observation and action spaces. Through discretization, continuous values are converted into discrete semantics, enabling a form of fuzzy control.




(a) MADDPG. (b) MAPPO.

Figure 10: Performance with Policy-based Extension.

### A.8 Ablation Study on Message Format and Fusion Process.

We conduct an ablation study by crossing two message formats (`Neural`, `HSM`) with two receiver fusion strategies (`Conventional`, `Herm`). All model variants adopt identical QMIX backbone, message dimension and network topology. The Neural-Herm variant (`Herm_ne`) replaces HSM with neural encoders while preserving the dual-branch fusion mechanism. The HSM-conventional variant (`Herm_mf`) retains HSM encoding but switches to mean fusion. We further introduce two additional ablated variants: one using neural encoding paired with mean fusion (`Herm_ne_mf`), and another adopting reversed similarity-based fusion (`Herm_is`).

Experimental results show that substituting HSM with neural encoding substantially hurts model performance, leading to over $20\%$ performance degradation on the $11 \times 11$ task under both noise-free and noisy settings. Despite gradient-based optimization, learnable encoders directly operating on raw observations struggle to extract decision-critical information and cannot sustain stable improvement during dynamic updates. In contrast, HSM constructs a smooth manifold mapping from perception-intention to messages, effectively reducing learning stochasticity.

Moreover, both mean fusion and inverse-similarity fusion induce performance drops, where the inverse-similarity variant suffers larger degradation. Mean fusion incurs global semantic dilution, while inverse similarity yields misaligned attention: perceptually similar agents prioritize less discriminative environmental cues, whereas perceptually dissimilar agents overemphasize behavioral signals. These comparisons verify the efficacy of our dual-branch message fusion design.

We conclude that performance gains primarily originate from the HSM message format, while the fusion module acts as an auxiliary downstream component tailored for HSM.

Table 7: Ablation of Herm's Variant under Noise

| Task | Variant | 0.0 | 0.5 | 1.0 | 2.0 |
|---|---|---|---|---|---|
| $9 \times 9$ | Herm | **99.1±0.5** | **97.6±1.2** | **94.8±0.9** | **84.7±1.1** |
| | Herm_ne | $94.8 \pm 0.28$ | $89.6 \pm 3.9$ | $87.4 \pm 3.2$ | $80.5 \pm 4.1$ |
| | Herm_mf | $93.4 \pm 0.4$ | $93.6 \pm 0.2$ | $87.8 \pm 1.3$ | $80.8 \pm 1.6$ |
| | Herm_is | $92.2 \pm 0.7$ | $92.8 \pm 1.1$ | $83.6 \pm 0.8$ | $76.1 \pm 1.2$ |
| | Herm_ne_mf | $90.9 \pm 2.5$ | $87.9 \pm 2.2$ | $83.7 \pm 3.8$ | $78.2 \pm 4.4$ |
| $11 \times 11$ | Herm | **96.1±0.3** | **94.1±0.2** | **90.8±1.1** | **81.8±1.4** |
| | Herm_ne | $74.3 \pm 2.4$ | $70.2 \pm 2.1$ | $66.8 \pm 3.7$ | $60.3 \pm 4.2$ |
| | Herm_mf | $92.9 \pm 0.6$ | $93.1 \pm 0.9$ | $86.8 \pm 0.7$ | $79.8 \pm 1.5$ |
| | Herm_is | $91.7 \pm 0.3$ | $92.3 \pm 0.4$ | $82.7 \pm 1.0$ | $76.7 \pm 1.3$ |
| | Herm_ne_mf | $70.9 \pm 2.2$ | $67.9 \pm 3.3$ | $63.7 \pm 2.9$ | $58.2 \pm 4.0$ |

# B Theoretical Proof for Similarity Proposition

In this section, we provide detailed elaboration and proofs for the three similarity properties involved in the decoding process. Specifically, the following proofs are exclusively applicable to the Holographic Reduced Representation (HRR) framework.

The original description for the three properties is as follows:

1. $\delta(HV_A \oplus HV_B, HV_A) \gg 0$
2. $\delta(HV_A \circledast HV_B, HV_A) \approx 0$
3. $\delta(HV_A \otimes HV_B \circledast HV_B, HV_A) \gg 0$

Since binding and unbinding are mutually inverse operations, Property 2 is equivalently expressed as: $\delta(HV_A \otimes HV_B, HV_A) \approx 0$. Consequently, we prove Property 2 from the perspective of binding operations.

## B.1 Preliminaries and Problem Setup

Let $\mathbb{R}^D$ denote the hyperdimensional space where $D \gg 1$. For simplicity, we denote hyperdimensional vectors by bold lowercase letters (e.g., $\mathbf{a}, \mathbf{b} \in \mathbb{R}^D$) in the following. The similarity between two vectors is measured by the cosine similarity, which simplifies to the dot product for unit-norm vectors. The Hermitian inner product is defined as $\langle \mathbf{u}, \mathbf{v} \rangle \triangleq \mathbf{u}^\top \mathbf{v}$.

Let $\mathcal{F}(\cdot)$ denote the Discrete Fourier Transform (DFT), mapping $\mathbb{R}^D \to \mathbb{C}^D$. By Parseval's theorem, the inner product is preserved (up to a scaling factor) in the frequency domain, where $\overline{\mathcal{F}(\mathbf{v})_k}$ indicates the complex conjugate of the frequency component:

$$\langle \mathbf{u}, \mathbf{v} \rangle = \frac{1}{D} \sum_{k=0}^{D-1} \mathcal{F}(\mathbf{u})_k \overline{\mathcal{F}(\mathbf{v})_k} \tag{14}$$

To derive the statistical bounds, we formalize the properties of atomic hypervectors.

**Assumption B.1** (Hyperdimensional Random Vector)**.** Atomic symbols are represented by random vectors $\mathbf{v} \in \mathbb{R}^D$ with independent and identically distributed (i.i.d.) components sampled from a distribution with zero mean and variance $1/D$.

$$\mathbb{E}[v_i] = 0, \quad \mathrm{Var}[v_i] = \frac{1}{D} \tag{15}$$

This variance scaling ensures that $\mathbb{E}[\|\mathbf{v}\|^2] = 1$. Furthermore, distinct atomic vectors $\mathbf{u}, \mathbf{v}$ are statistically independent, implying quasi-orthogonality in high dimensions: $\mathbb{E}[\langle \mathbf{u}, \mathbf{v} \rangle] = 0$. For norm variance computability, components satisfy the fourth-moment condition $\mathbb{E}[v_i^4] = \frac{3}{D^2}$.

**Assumption B.2** (Fourier Domain Statistics)**.** For large $D$, by the Central Limit Theorem (CLT, satisfied as components are i.i.d. with finite variance), the DFT components of random vectors, denoted as $z_k = \mathcal{F}(\mathbf{v})_k$, converge to a Circularly-Symmetric Complex Gaussian distribution:

$$z_k \sim \mathcal{CN}(0, 1) \tag{16}$$

Key statistical moments required for the following proofs are:

- **Second Moment (Energy):** $\mathbb{E}[|z_k|^2] = 1$.
- **Fourth Moment (Kurtosis):** $\mathbb{E}[|z_k|^4] = 2$.
- **Independence:** For $k \neq j$, $z_k$ and $z_j$ are effectively independent (subject to conjugate symmetry constraints for real inputs, which do not affect asymptotic scaling).

### B.2 Formal Definitions of HRR Operations

We formally define the three core algebraic operations: Bundling ($\oplus$), Binding ($\otimes$), and Unbinding ($\circledast$).

**Definition B.3** (Bundling)**.** Bundling aggregates a set of vectors $\{\mathbf{x}_i\}_{i=1}^n$ into a single composite vector $\mathbf{S}$. To prevent norm growth, we define the normalized bundling as Equ.17. This operation preserves similarity. For any component $\mathbf{x}_k$, the expected similarity $\mathbb{E}[\langle \mathbf{S}, \mathbf{x}_k \rangle] > 0$.

$$\mathbf{S} \triangleq \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{x}_i. \tag{17}$$

**Definition B.4** (Binding via Circular Convolution)**.** Binding associates two vectors $\mathbf{a}$ and $\mathbf{b}$ into a pair $\mathbf{c}$. In HRR, this is realized by circular convolution, denoted as $\mathbf{c} = \mathbf{a} \otimes \mathbf{b}$. There exist two mathematical formulations for binding in the spatial domain and frequency domain as follows, where $\circ$ denotes the element-wise (Hadamard) product. Binding creates a vector that is quasi-orthogonal to its constituents ($\mathbf{c} \perp \mathbf{a}$ and $\mathbf{c} \perp \mathbf{b}$).

- **Spatial Domain:** $c_k = \sum_{j=0}^{D-1} a_j b_{(k-j) \pmod D}$.
- **Frequency Domain:** $\mathcal{F}(\mathbf{c}) = \mathcal{F}(\mathbf{a}) \circ \mathcal{F}(\mathbf{b})$.

**Definition B.5** (Unbinding via Approximate Inverse)**.** Unbinding refers to the operation that extracts a constituent vector from a bound pair. The exact inverse of circular convolution suffers from numerical instability. HRR adopts involution as the approximate inverse, which corresponds to the complex conjugate in the frequency domain. Let $\mathbf{b}^\dagger$ denote the involution of $\mathbf{b}$, defined in the frequency domain as $\mathcal{F}(\mathbf{b}^\dagger)_k = \overline{\mathcal{F}(\mathbf{b})_k}$. The unbinding operation is defined as circular correlation (denoted $\circledast$):

$$\mathbf{a}' \triangleq \mathbf{c} \otimes \mathbf{b}^\dagger \triangleq \mathbf{c} \circledast \mathbf{b} \tag{18}$$

In the frequency domain:

$$\mathcal{F}(\mathbf{a}')_k = \mathcal{F}(\mathbf{c})_k \cdot \overline{\mathcal{F}(\mathbf{b})_k} \tag{19}$$

The unbinding operation is inherently noisy: the circular convolution of $\mathbf{b}$ and its involution $\mathbf{b}^\dagger$ yields $\mathbf{b} \otimes \mathbf{b}^\dagger$, which merely approximates the identity impulse $\boldsymbol{\delta}$; such approximation inevitably induces a residual noise term, and the magnitude of this term is strictly determined by the deviation of $|\mathcal{F}(\mathbf{b})_k|^2$ from unity.

### B.3 Similarity Properties

#### B.3.1 Bundling

Bundling aggregates information while preserving similarity to constituents.

**Theorem B.6** (Bundling Statistics)**.** *For $D \to \infty$ and $n \ll D$, let $\mathbf{S} = \frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbf{a}_i$. For any $k \in \{1, \ldots, n\}$:*

$$\mathbb{E}[\mathbf{S} \cdot \mathbf{a}_k] = \frac{1}{\sqrt{n}}, \quad \textit{Var}[\mathbf{S} \cdot \mathbf{a}_k] \approx \frac{1}{D} \tag{20}$$

*Proof.* Let $\mathbf{S} = \frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbf{a}_i$, decompose the dot product and calculate the expectation:

$$\mathbf{S} \cdot \mathbf{a}_k = \left( \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{a}_i \right) \cdot \mathbf{a}_k = \frac{1}{\sqrt{n}} \left( \|\mathbf{a}_k\|^2 + \sum_{i \neq k} \mathbf{a}_i \cdot \mathbf{a}_k \right) \tag{21}$$

By Assumption B.1, $\mathbb{E}[\|\mathbf{a}_k\|^2] = 1$ and the cross terms have zero mean, i.e., $\mathbb{E}[\mathbf{a}_i \cdot \mathbf{a}_k] = 0$ $(i \neq k)$.

$$\mathbb{E}[\mathbf{S} \cdot \mathbf{a}_k] = \frac{1}{\sqrt{n}} \left( \mathbb{E}[\|\mathbf{a}_k\|^2] + \sum_{i \neq k} \mathbb{E}[\mathbf{a}_i \cdot \mathbf{a}_k] \right) = \frac{1}{\sqrt{n}} \left( 1 + \sum_{i \neq k} 0 \right) = \frac{1}{\sqrt{n}}. \tag{22}$$

Let $\mathbf{a}_k = [a_{k1}, a_{k2}, \ldots, a_{kD}]^T$, the step-by-step calculation is as follows:

Expand by the definition of norm and combine with the variance property $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}^2[X]$:

$$\|\mathbf{a}_k\|^2 = \sum_{i=1}^{D} a_{ki}^2, \tag{23}$$

$$\text{Var}(\|\mathbf{a}_k\|^2) = \text{Var}\left( \sum_{i=1}^{D} a_{ki}^2 \right) = \sum_{i=1}^{D} \left( \mathbb{E}[a_{ki}^4] - \mathbb{E}^2[a_{ki}^2] \right) = \sum_{i=1}^{D} \left( \frac{3}{D^2} - \left( \frac{1}{D} \right)^2 \right) = \frac{2}{D}. \tag{24}$$

Since $\mathbf{a}_i$ and $\mathbf{a}_k$ are mutually independent, thus: $\text{Var}(\mathbf{a}_i \cdot \mathbf{a}_k) = \frac{1}{D}$.

$$\begin{aligned} \text{Var}[\mathbf{S} \cdot \mathbf{a}_k] &= \text{Var}\left[ \frac{1}{\sqrt{n}} \left( \|\mathbf{a}_k\|^2 + \sum_{i \neq k} \mathbf{a}_i \cdot \mathbf{a}_k \right) \right] = \frac{1}{n} \text{Var}(\|\mathbf{a}_k\|^2) + \frac{1}{n} \text{Var}\left( \sum_{i \neq k} \mathbf{a}_i \cdot \mathbf{a}_k \right) \\ &= \frac{\text{Var}(\|\mathbf{a}_k\|^2)}{n} + \frac{n-1}{n} \text{Var}(\mathbf{a}_i \cdot \mathbf{a}_k) = \frac{2}{nD} + \frac{n-1}{nD} = \frac{n+1}{nD} \end{aligned} \tag{25}$$

When $n \ll D$, the term $\frac{2}{nD}$ is negligible, thus: $\text{Var}[\mathbf{S} \cdot \mathbf{a}_k] \approx \frac{1}{D}$. □

### B.3.2 Binding

Binding creates a composite representation orthogonal to its constituents.

**Theorem B.7** (Binding Orthogonality)**.** *For $D \to \infty$, let $\mathbf{c} = \mathbf{a} \otimes \mathbf{b}$. Then $\mathbb{E}[\mathbf{c} \cdot \mathbf{a}] = 0$ and $Var[\mathbf{c} \cdot \mathbf{a}] = 2/D$.*

*Proof.* In the Fourier domain, $\mathcal{F}(\mathbf{c})_k = \mathcal{F}(\mathbf{a})_k \mathcal{F}(\mathbf{b})_k$. By Parseval's theorem:

$$\mathbf{c} \cdot \mathbf{a} = \frac{1}{D} \sum_{k=0}^{D-1} |\mathcal{F}(\mathbf{a})_k|^2 \mathcal{F}(\mathbf{b})_k \tag{26}$$

Since $\mathbf{a}, \mathbf{b}$ are real-valued, $\mathcal{F}(\mathbf{v})_{D-k} = \overline{\mathcal{F}(\mathbf{v})_k}$, ensuring the sum is real. By Assumption B.2, $\mathbb{E}[\mathcal{F}(\mathbf{b})_k] = 0$ and vectors are independent, so $\mathbb{E}[\mathbf{c} \cdot \mathbf{a}] = 0$.

For variance, let $X_k = |\mathcal{F}(\mathbf{a})_k|^2 \mathcal{F}(\mathbf{b})_k$ (real-valued due to conjugate symmetry). $\text{Var}[X_k] = \mathbb{E}[|\mathcal{F}(\mathbf{a})_k|^4] \mathbb{E}[|\mathcal{F}(\mathbf{b})_k|^2] = 2 \cdot 1 = 2$. Summing over $D$ independent components:

$$\text{Var}[\mathbf{c} \cdot \mathbf{a}] = \frac{1}{D^2} \sum_{k=0}^{D-1} \text{Var}[X_k] = \frac{1}{D^2} \cdot 2D = \frac{2}{D}. \tag{27}$$

□

### B.3.3 Unbinding

Unbinding retrieves the original vector from the bound pair using the approximate inverse.

**Theorem B.8** (Unbinding Recovery and SNR)**.** *For $D \to \infty$, let $\mathbf{a}' = (\mathbf{a} \otimes \mathbf{b}) \circledast \mathbf{b}$. The retrieval similarity follows:*

$$\mathbb{E}[\mathbf{a}' \cdot \mathbf{a}] = 1, \quad Var[\mathbf{a}' \cdot \mathbf{a}] = \frac{6}{D}. \tag{28}$$

*Consequently, the Signal-to-Noise Ratio (SNR) scales as $\mathcal{O}(D)$.*

*Proof.* Since the unbinding operation is diagonalized in the Fourier domain, we analyze the statistics of the frequency components. While real-valued vectors impose Hermitian symmetry ($\mathcal{F}(\mathbf{a})_k = \overline{\mathcal{F}(\mathbf{a})_{D-k}}$), we treat the components as independent for asymptotic analysis. This simplifies the derivation without affecting the order of magnitude of the SNR. The unbinding operation in the Fourier domain is $\mathcal{F}(\mathbf{a}')_k = \mathcal{F}(\mathbf{a})_k |\mathcal{F}(\mathbf{b})_k|^2$. Let $Z_k = |\mathcal{F}(\mathbf{a})_k|^2 |\mathcal{F}(\mathbf{b})_k|^2$. The similarity is the sample mean of $Z_k$:

$$\mathbf{a}' \cdot \mathbf{a} = \frac{1}{D} \sum_{k=0}^{D-1} Z_k. \tag{29}$$

**Expectation:** Using the independence of $\mathbf{a}$ and $\mathbf{b}$:

$$\mathbb{E}[Z_k] = \mathbb{E}[|\mathcal{F}(\mathbf{a})_k|^2]\mathbb{E}[|\mathcal{F}(\mathbf{b})_k|^2] = 1 \cdot 1 = 1 \implies \mathbb{E}[\mathbf{a}' \cdot \mathbf{a}] = 1 \tag{30}$$

**Variance:** First, we compute the variance of a single frequency component $Z_k = |\mathcal{F}(\mathbf{a})_k|^2 |\mathcal{F}(\mathbf{b})_k|^2$. Since the elements of $\mathbf{a}, \mathbf{b}$ are i.i.d. Gaussian, the frequency components converge to a complex Gaussian distribution (CLT). Using the fourth-moment property $\mathbb{E}[|z|^4] = 2$ for standard complex Gaussian variables:

$$\text{Var}[Z_k] = \mathbb{E}[Z_k^2] - (\mathbb{E}[Z_k])^2 = \mathbb{E}[|\mathcal{F}(\mathbf{a})_k|^4]\mathbb{E}[|\mathcal{F}(\mathbf{b})_k|^4] - 1^2 = (2 \cdot 2) - 1 = 3. \tag{31}$$

For the frequency components, the variance is formulated as:

$$\text{Var}[\mathbf{a}' \cdot \mathbf{a}] = \text{Var}\left[\frac{1}{D} \sum_{k=0}^{D-1} Z_k\right] = \frac{1}{D^2} \cdot 3D = \frac{3}{D}. \tag{32}$$

For real-valued vectors, the Fourier spectrum satisfies Hermitian symmetry: $\mathcal{F}(\mathbf{x})_k = \overline{\mathcal{F}(\mathbf{x})}_{D-k}$, which implies $Z_k = Z_{D-k}$. Consequently, the components in the sum are pairwise correlated. The variance of the mean is:

$$\text{Var}[\mathbf{a}' \cdot \mathbf{a}] = \text{Var}\left[\frac{1}{D} \sum_{k=0}^{D-1} Z_k\right] = \frac{1}{D^2} \sum_{k=0}^{D-1} \sum_{j=0}^{D-1} \text{Cov}[Z_k, Z_j] \tag{33}$$

Due to the independence of distinct frequencies, $\text{Cov}[Z_k, Z_j] = 0$ for $j \notin \{k, D-k\}$. The non-zero covariance terms are the variance itself ($j = k$) and the symmetry term ($j = D - k$). Since $Z_k = Z_{D-k}$, it results in:

$$\text{Cov}[Z_k, Z_{D-k}] = \text{Var}[Z_k] = 3 \tag{34}$$

Thus, for each $k$, the inner sum contributes $2 \times \text{Var}[Z_k]$:

$$\text{Var}[\mathbf{a}' \cdot \mathbf{a}] \approx \frac{1}{D^2} \sum_{k=0}^{D-1} (\text{Var}[Z_k] + \text{Cov}[Z_k, Z_{D-k}]) = \frac{1}{D^2} \sum_{k=0}^{D-1} (3 + 3) = \frac{6}{D} \tag{35}$$

**SNR:** The SNR is defined as the ratio of the squared signal expectation to the noise variance:

$$\text{SNR} = \frac{(\mathbb{E}[\mathbf{a}' \cdot \mathbf{a}])^2}{\text{Var}[\mathbf{a}' \cdot \mathbf{a}]} = \frac{1}{6/D} = \frac{D}{6} \tag{36}$$

□

*Remark* B.9. A critical feature of HRR is the $\mathcal{O}(D)$ scaling of SNR: as the dimension $D$ increases, SNR grows linearly. It implies the multiplicative noise introduced by the approximate inverse is *effectively suppressed in high-dimensional spaces*. This enables reliable symbol retrieval without exact inverse computation.

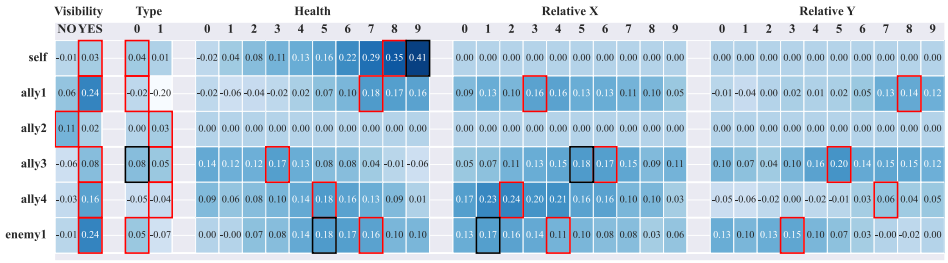
By Chebyshev's inequality, for any $\epsilon > 0$,

$$\mathbb{P}\left(|\mathbf{a}' \cdot \mathbf{a} - 1| \geq \epsilon\right) \leq \frac{\text{Var}[\mathbf{a}' \cdot \mathbf{a}]}{\epsilon^2} = \frac{6}{D\epsilon^2}.$$

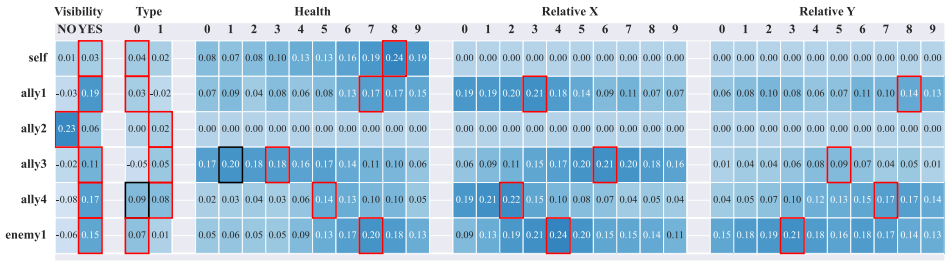
As $D \to \infty$, the right tends to 0, so $\mathbf{a}' \cdot \mathbf{a} \xrightarrow{p} 1$ (converges in probability to 1), confirming the *asymptotic retrieval accuracy* in high dimensions.

## C Conversion Fidelity Analysis on Message Length

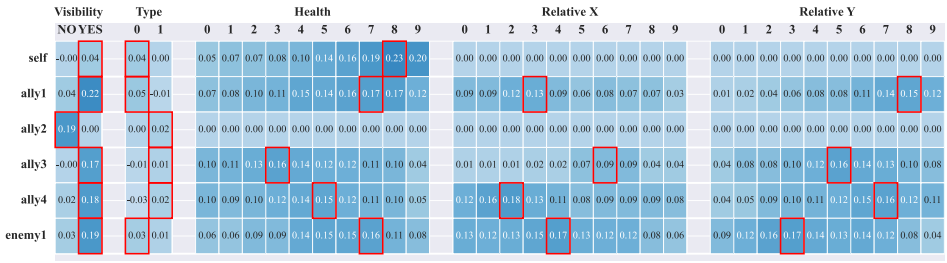
In this section, we investigate the impact of message length on the accuracy of semantic transmission. We dynamically adjust the message length $\mathcal{H}$ in Herm and perform message decoding for identical scenarios, as visualized in Fig.11 (SMAC) and Fig.12 (LBF). The red boxes indicate original semantics, while the black boxes denote erroneously decoded semantics deviating from the originals. The results show that when $\mathcal{H}$ is small, the decoded outputs contain significant errors. Furthermore, similarity metrics exhibit random patterns without discernible trends, contradicting the value dictionary's design principle. We attribute these phenomena to severe degradation of vector orthogonality, which violates the pseudo-inverse requirement for unbinding in hyperdimensional computing. As $\mathcal{H}$ increases, the error rate drops markedly, with a near-perfect decoding achieved at $\mathcal{H}$=512. Meanwhile, similarity scores within the same value dictionary exhibit a distinct peak phenomenon, validating the rationality of the encoding-decoding process.




(a) $\mathcal{H}$ =256




(b) $\mathcal{H}$ =512




(c) $\mathcal{H}$ =1024

Figure 11: Conversion Fidelity in SMAC with Different Message Length.

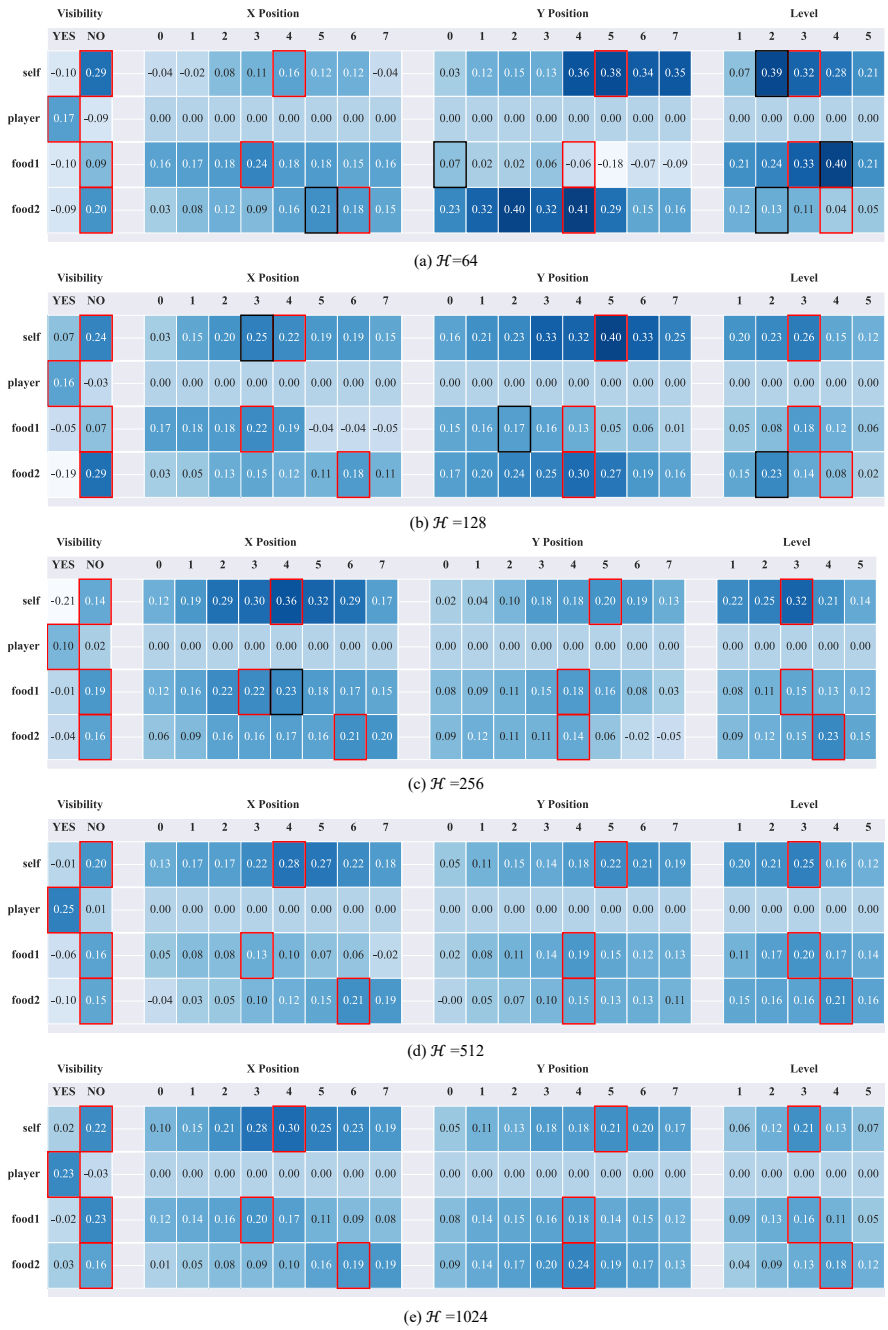



Figure 12: Conversion Fidelity in LBF with Different Message Length.

Fig.13 investigates the semantic discrepancy between decoded hyperdimensional symbolic messages (HSM) and original semantics under noisy settings ($\eta_r$). Here, we perform iterative decoding on one hundred messages and compare the transformed semantics with the original semantics. Results show that message dimension $\mathcal{H}$ regulates the semantic discrepancy: increased decoding accuracy and decreased variance occur with the rise of message dimension. This indicates that the conversion fidelity of HSM is significantly enhanced with the increase in dimension. Furthermore, message dimension also significantly affects decoding robustness—when the dimension exceeds 256, both the mean accuracy and variance remain stable.

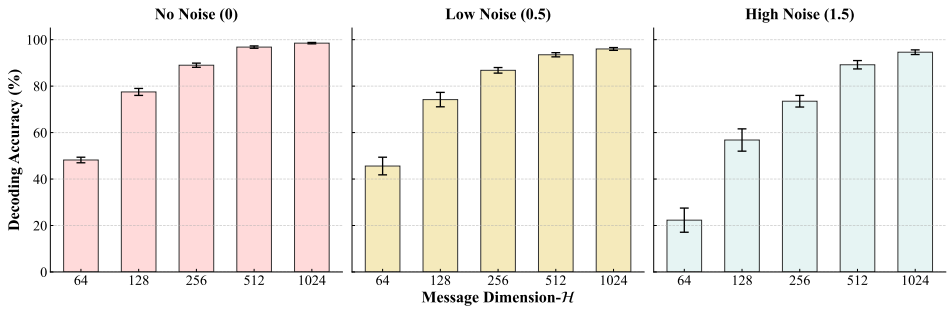
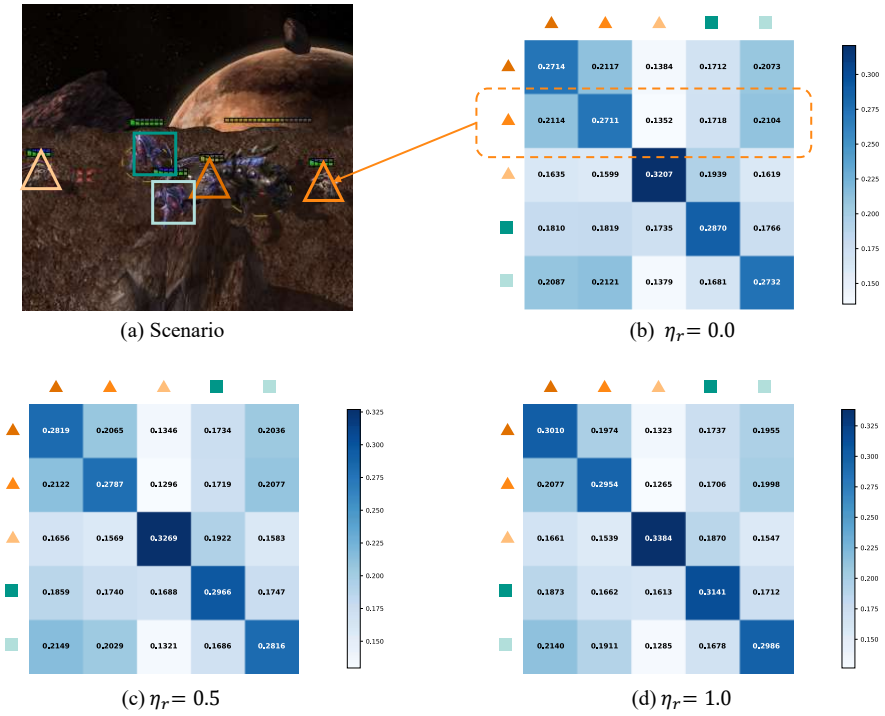



Figure 13: Decoding Accuracy in SMAC with Different Message Length.

## D Message Similarity Analysis under Noise

We conduct a similarity evaluation on the observation component of communication messages in SMAC, with results illustrated in Fig.14 (b). The findings indicate that the observation encoding of the rightmost stalker exhibits the highest similarity to that of the middle stalker (excluding itself). This aligns with their actual scenario characteristics in Fig.14 (a): they belong to the same race and possess similar relative positions with respect to other entities. This phenomenon reveals that hyperdimensional vectors can still effectively preserve semantic similarity even after undergoing multiple bundling and binding. Furthermore, under different noise conditions, the correspondence between semantic similarity and encoding similarity remains intact, as shown in Fig.14 (c-d). This further illustrates the robustness of the proposed hyperdimensional message encoding in this paper.




(a) Scenario (b) $\eta_r = 0.0$

(c) $\eta_r = 0.5$ (d) $\eta_r = 1.0$

Figure 14: Observation Similarity Visualization in SMAC. Triangles denote "stalker" and squares denote "zealot". Subfigures b, c, and d present the observation encoding similarity matrices under different noise conditions.

## E Similarity-Preservation Verification for Ego-Binding

Fig. 15 compares the cosine similarity matrices of agent messages before and after ego-binding. The two matrices show nearly identical global patterns and value distributions, with perfect preservation of mutual similarity and no semantic inversion observed. While ego-binding leads to a rise in the absolute similarity values, the relative similarity relations across agents are preserved. This indicates that semantic alignment is consistent with weight calculation and incurs no conflict.

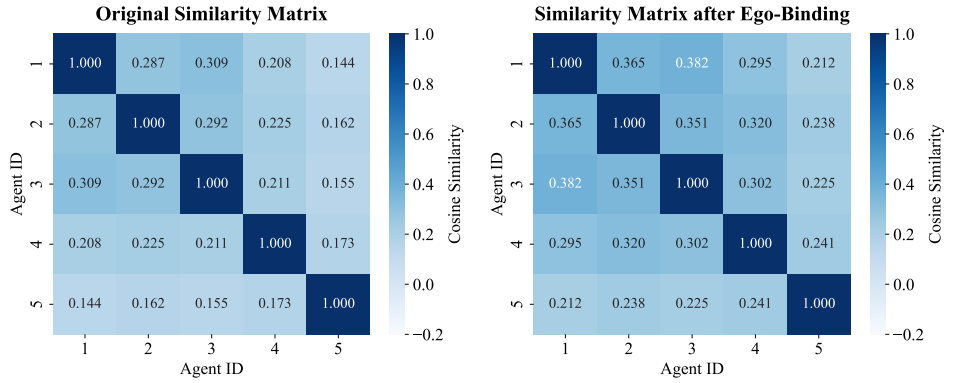



Figure 15: Analysis of the Ego-Binding on Similarity.

## F Hyperdimensional Encoding Process

In this section, we elaborate on the detailed procedures of hyperdimensional message encoding in each experimental environment, including attribute partitioning, value constraints, and computational operations.

### F.1 SMAC / SMACv2

Despite variations in environmental layouts and observational capacities, SMAC and SMACv2 maintain identical fundamental observation and action structures at their core.

#### F.1.1 Attribute Dictionary.

We posit mutual independence among attributes and construct the encoding dictionary through stochastic sampling. The observation is formally defined as an attribute dictionary. Tab 8 introduces the detailed design. Notably, the attribute design here is merely intended for human readability. For agents, it only performs dimension-wise encoding over raw observations.

Table 8: Attribute Properties Specification in SMAC

| Attribute | Explanation | Independence | Discrete | Value Range |
|---|---|---|---|---|
| Agent ID | Unique Identifier | Independent | Yes | $\{1, \cdots, N_{entity}\}$ |
| HP | Current Health Value | Associated | No | $[0, 1]$ |
| Relative X | Normalized X-position Relative to Agent | Associated | No | $(-1, 1)$ |
| Relative Y | Normalized Y-position Relative to Agent | Associated | No | $(-1, 1)$ |
| Unit Type | Race of the Unit | Independent | Yes | $\{1, \cdots, N_{type}\}$ |
| Visibility | Binary Visibility Indicator | Independent | Yes | $\{0, 1\}$ |
| Identity | Indicator Whether the Unit is Self, Ally, or Enemy | Independent | Yes | $\{0, 1, 2\}$ |
| Time | Time Window Index | Associated | No | $\{-1, \cdots, -w\}$ |

The action encoding comprises all executable actions: NOTHING, STAY, MOVE_UP, MOVE_DOWN, MOVE_LEFT, MOVE_RIGHT, ATTACK, HEAL.

### F.1.2 Value Dictionary.

Different attributes correspond to distinct value types, such as independent/associated and discrete/-continuous. We provide detailed specifications for each attribute in Table 8. For independent values, we employ random sampling. For continuous variables, we perform interval discretization followed by cyclic replacement encoding, typically using 10 bins.

### F.1.3 Encoding Process.

The observation encoding follows an entity-wise paradigm, where self features, ally features, and enemy features are separately encoded before being bundled into a composite hyperdimensional vector. Here, $HV_{attr}$ denotes the result after binding attribute-value pairs.

- For self-encoding, we bind the agent's ID, Identity, HP, and Type. The process is formulated as $HV_{id} \otimes HV_{self} \otimes (HV_{hp} \oplus HV_{type})$. A zero vector is returned if the current agent is terminated.
- For ally/enemy encoding, we first perform visibility verification before feature extraction. If invisible, the process is defined as $HV_{id} \otimes HV_{ally/enemy} \otimes HV_{invisible}$. If visible, we bundle its static attributes and perform identity binding, formulated as $HV_{id} \otimes HV_{ally/enemy} \otimes (HV_{visible} \oplus HV_{hp} \oplus HV_x \oplus HV_y \oplus HV_{type})$.

For action sequence encoding, we perform temporal binding of actions within the window and then bundle all features, which is denoted as $(HV_{-1} \otimes HV_{a_{-1}}) \oplus \cdots \oplus (HV_{-w} \otimes HV_{a_{-w}})$

## F.2 Lbforaging (LBF)

Due to the similarity in the encoding process, we only present the attribute dictionary and value dictionary (Table 9) here, while the encoding procedure follows the aforementioned section.

Table 9: Attribute Properties Specification in LBF

| Attribute | Explanation | Independence | Discrete | Value Range |
|---|---|---|---|---|
| Agent ID | Unique Identifier | Independent | Yes | $\{1, \cdots, N_{entity}\}$ |
| Absolute X | Absolute X-position in the Map | Associated | Yes | $\{0, \cdots, X\}$ |
| Absolute Y | Absolute Y-position in the Map | Associated | Yes | $\{0, \cdots, Y\}$ |
| Level | Capture Ability (Predator) / Escape Ability (Prey) | Independent | Yes | $\{1, \cdots, Level\}$ |
| Visibility | Binary Visibility Indicator | Independent | Yes | $\{0, 1\}$ |
| Identity | Indicator Whether the Unit is Self, Predator, or Prey | Independent | Yes | $\{0, 1, 2\}$ |
| Time | Time Window Index | Associated | No | $\{-1, \cdots, -w\}$ |

The action encoding comprises all executable actions: NOTHING, STAY, MOVE_UP, MOVE_DOWN, MOVE_LEFT, MOVE_RIGHT, CAPTURE.

## F.3 Hallway

Due to the similarity in the encoding process, we only present the attribute dictionary and value dictionary (Table 10) here, while the encoding procedure follows the aforementioned section.

Table 10: Attribute Properties Specification in Hallway

| Attribute | Explanation | Independence | Discrete | Value Range |
|---|---|---|---|---|
| Agent ID | Unique Identifier | Independent | Yes | $\{1, \cdots, N_{agent}\}$ |
| Position | Position in its Line | Associated | Yes | $\{0, \cdots, X\}$ |
| Time | Time Window Index | Associated | No | $\{-1, \cdots, -w\}$ |

The action encoding comprises all executable actions: STAY, MOVE_LEFT, MOVE_RIGHT.

# G Collaboration with Large Language Models

Hyperdimensional Symbolic Message (HSM) introduced in our paper constructs a bidirectional conversion pipeline that links (1) message encoding, (2) symbolic semantics, and (3) natural language. In this framework, message encoding enables efficient agent-to-agent communication and decision-making processes. Natural language supports interpretability for both human operators and large language models (LLMs). Symbolic semantics then acts as a core bridging medium that connects these two functionally distinct components. Such explicit conversion fidelity eliminates the algorithm's reliance on foundation model embedding spaces, while ensuring compatibility with various LLMs. Next, we detail the communication and collaborative decision-making processes between agents and LLMs.

## G.1 Message Transformation

```
I'm Agent 2, my type is 1, and my HP is within 0.40-0.50.
I can not see Ally_0.
I can see Ally_1, its type is 0, its HP is within 0.40~0.50, its relative position x to me is within -0.80~-0.60, and its relative position y to me is within -0.80~-0.60.
I can see Ally_3, its type is 0, its HP is within 0.40~0.50, its relative position x to me is within -1.00~-0.80, and its relative position y to me is within -0.60~-0.40.
I can see Ally_4, its type is 1, its HP is within 0.40~0.50, its relative position x to me is within -0.60~-0.40, and its relative position y to me is within -1.00~-0.80.
I can not see Enemy_0.
My previous actions contain ['Move West', 'Move North'].
```

Figure 16: Example of Decoded Message in SMAC

For agents, the messages during communication are actually hyperdimensional vectors. Through the proposed iterative decoding method (Section 4.3), these hyperdimensional vectors are decoded into a series of structured semantic entities, as depicted in Fig.11 and Fig.12. The semantic entities are expressed as attribute-value pairings, which can then be aggregated and mapped to natural language. We present the form of messages following decoding and programmatic transformation in Fig.16, which contains entity-wise information.

## G.2 Decision Making Process

In the team, a single agent is designated as an L-Agent, which is controlled by LLMs; all remaining agents are defined as A-Agents that operate under policies. The joint action of both types of agents is transmitted and executed in the environment.

**A-Agent** $\rightarrow$ **A-Agent**. During each decision time step, every A-Agent encodes its local perceptual data and historical action sequences into hyperdimensional symbolic messages (HSMs) for inter-agent communication and coordination.

**A-Agent** $\rightarrow$ **L-Agent**. Messages from A-Agents are uniformly translated into natural language, aggregated into a structured report, and forwarded to the L-Agent for further processing. The L-Agent then selects the current decision from the set of executable actions, guided by the reasoning process of LLMs.

**L-Agent** $\rightarrow$ **A-Agent**. The L-Agent is also required to convey its local observations and historical action sequences to peer agents. Given that A-Agents are designed only to accept HSMs, the L-Agent utilizes the same hyperdimensional encoding module as A-Agents. The encoded messages are subsequently delivered to A-Agents to facilitate their decision-making processes.

We design the prompt to guide LLM in agent collaboration as Fig.17.

```
Core Role & Positioning:
We are playing StarCraft II micro scenario "2s3z_vs_1ul", trying to control our agents to defeat all of the enemy units.
You are the last member of the multi-agent team, numbered Agent-4, collaborating with 4 algorithm-controlled agents (numbered Agent-0, Agent-1, ..., Agent-3) to complete tasks.
You must make decisions strictly based on messages sent by other agents, with no access to additional global information.

Input Format:
In each step, you will perceive your own observation with natural language and receive structured messages from all agents in the team.
Each message contains the following fields: [Agent-ID] | [Current Observation] | [Previous Actions Sequence].
Furthermore, you will receive your current available action list. Action index is listed as: {
0: No-op only available for dead agents; 1: Stop; 2: Move North; 3: Move South; 4: Move East; 5: Move West; 6 to 6+n_enemy: Attack the enemy with a certain index (index - 6)}

Decision Rules & Collaboration Principles:
1. Team Priority First: Decisions must aim for global task victory, prioritizing coordination with teammates' tactical tendencies (e.g., if teammates focus fire on a specific enemy unit, coordinate the attack; if teammates retreat, provide cover).
2. Observation-Driven Decision Making: Judge strictly based on input message data; Do not assume unobserved information (e.g., enemy units outside vision, status of agents that have not sent messages).
3. Action Effectiveness Constraints: Only select actions from your own list of executable actions; do not perform operations beyond your capabilities.
4. Historical Action Reference: Adjust strategies based on the effect of teammates' previous K steps of actions (e.g., if teammates continuously attack an enemy unit but fail to destroy it, evaluate whether support is needed; if teammates' movement paths show advancement towards the target area, follow up).

Output Format Requirements:
Output only contains SMAC directly executable action commands in the above action list, and the reason why you made the choice.
{
               "action": (action index),
               "reason": (why you do this action)
}

Input & Output Examples: ......

Note:
1. Your answer must follow the given format without any extra output.
2. You must select actions only from the provided list of available actions; actions outside this list are prohibited.
```

Figure 17: Example of LLM Prompt for Multi-Agent Communication

### G.3 Experiments

We evaluate the performance of the algorithm-LLM hybrid decision-making mechanism on the SMAC benchmark, which comprises N-1 algorithm-controlled agents (A-Agents) and one LLM-controlled agent (L-Agent). Given the overhead of real-time interaction, we do not incorporate LLMs into the training of our method, Herm, and only replace one team member with an L-Agent during testing.

We present the LLM's reasoning process during the decision-making phase, as illustrated in Fig.18. It can be observed that the LLM is able to perform reasoning based on the decoded and processed natural language messages and make corresponding decisions.

```
No enemy is visible to me or any ally based on the messages.
I can see Ally_0 to my northwest (relative position x -1.00~-0.80, y
    -0.60~-0.40).
Agent 0's message indicates they see me to their east (relative x
    -0.20~0.00, y -0.60~-0.40), so I am west of them.
To regroup and explore actively (as per principle 5), moving East (action 4)
     will bring me closer to Ally_0 and potentially towards the area where
    other allies might be.
 Because Agent 0 sees Ally_1 to their northwest, this choice aligns with
     team coordination and active participation.
```

Figure 18: Example of LLM Reasoning

Tab.11 presents the hybrid decision-making performance of the algorithm with different LLMs. On the one hand, Herm maintains a certain level of performance when integrated with the three LLMs, which validates the generalizability of the proposed hyperdimensional symbolic messages in hybrid decision-making.

On the other hand, notable performance discrepancies are observed in collaboration with different LLMs, a phenomenon that can be attributed to the intrinsic decision-making competence of LLMs. In terms of the dimensionality of the hyperdimensional space $\mathcal{H}$, performance exhibits an upward trend with increasing dimensionality. This is because a larger hyperdimensional space facilitates accurate semantic transformation. In addition, the upward trend gradually slows down under high-dimensional settings: minor transformation errors are corrected by other message sources during LLM decision-making, leading to a reduced impact on decision outcomes.

Table 11: Performance of Algorithm-LLM Hybrid Decision Making

| $\mathcal{H}$ | Task | Herm | DeepSeek-V3.2 | GLM-4-7 | Doubao-Seed-1.6 |
|---|---|---|---|---|---|
| 256 | 5z_vs_1ul | 79.0±0.2 | 68.4±3.7 | 31.6±3.6 | 53.8±3.5 |
| | 2s3z_vs_1ul | 81.9±0.6 | 64.3±3.3 | 35.6±4.8 | 60.5±3.1 |
| | 1c4z_vs_1ul | 75.1±1.5 | 52.6±3.4 | 28.5±3.3 | 44.6±4.2 |
| 512 | 5z_vs_1ul | 80.3±0.8 | 73.4±2.2 | 34.3±1.5 | 58.0±2.6 |
| | 2s3z_vs_1ul | 80.6±0.5 | 66.8±3.7 | 37.2±3.4 | 59.8±2.8 |
| | 1c4z_vs_1ul | 76.5±0.8 | 55.2±2.8 | 30.5±3.1 | 46.3±3.5 |
| 1024 | 5z_vs_1ul | 79.2±0.5 | 74.1±1.9 | 36.4±1.0 | 59.2±2.4 |
| | 2s3z_vs_1ul | 81.0±1.0 | 68.0±1.2 | 38.8±2.6 | 61.3±2.3 |
| | 1c4z_vs_1ul | 75.8±0.7 | 56.1±2.4 | 30.9±2.8 | 47.1±1.5 |

## H Algorithm Pseudocode

We introduce the decision-making process of a single agent (Alg. 2) and the algorithm training process (Alg. 1) in Herm, respectively. The overall training architecture employs QMIX.

**Algorithm 1:** Training Process in Herm

**Input:** $N$ (number of agents), $O$ (observation space), $A_i$ (action space of agent $i$), $\gamma$ (discount factor), $\alpha$ (learning rate), $\tau$ (soft update coefficient)

**Output:** Policies $\pi_1, \ldots, \pi_N$ for each agent

**Initialization:**

Agent Q-networks $Q_{local}(o_i, a_i; \theta)$ with parameters $\theta$

Mixing network $Mixing(q_1, \ldots, q_N; \beta)$ with $\beta$ (non-negative weights $\beta$)

Target Q-networks $Q_{local}^{\text{target}}$ and target mixing network $Mixing^{\text{target}}$: $\theta^- \leftarrow \theta$, $\beta^- \leftarrow \beta$

Replay buffer $D$ (stores tuples $(s, \boldsymbol{o}, \boldsymbol{a}, r, s', \text{done})$

Exploration strategy $\epsilon$ (e.g., $\epsilon$-greedy)

**for** *each training episode* **do**

  Initialize state $s$ **while** *not terminated* **do**

    Each agent $i$ selects $a_i = \pi_i(o_i)$ via $Q_{\text{local}}(o_i, \cdot; \theta)$ and $\epsilon$ ; `// see Alg. 2`

    Execute $\mathbf{a}$, get reward $r$ and next state $s'$

    Store $(s, \boldsymbol{o}, \boldsymbol{a}, r, s', \text{done})$ in $D$ $s \leftarrow s'$

  **end**

**end**

**for** *each parameter update* **do**

  Sample batch $\{(s_j, \boldsymbol{o}_j, \boldsymbol{a}_j, r_j, s'_j, \text{done}_j)\}_{j=1}^{B}$ from $D$

  **for** *each sample* $j$ **do**

    Each agent $i$ selects action: $a'_i = \arg\max_{a'} Q_{\text{local}}(o'_{i,j}, a'; \theta)$

    Agent target: $y_i^j = Q_{\text{local}}^{\text{target}}(o'_{i,j}, a'_i; \theta^-)$

    Joint target:

$$Y_j = r_j + \gamma(1 - \text{done}_j) \cdot Mixing^{\text{target}}(y_1^j, \ldots, y_N^j; \beta^-)$$

    Current agent Q-values: $q_i^j = Q_{\text{local}}(o_{i,j}, a_{i,j}; \theta)$

    Current joint value: $Q_{\text{total},j} = Mixing(q_1^j, \ldots, q_N^j; \beta)$

  **end**

  Loss: $L = \frac{1}{B} \sum_{j=1}^{B} (Y_j - Q_{\text{total},j})^2$

  Update $\theta$ and $\beta$ via gradient descent (ensure $\beta \geq 0$)

  Soft update targets: $\theta^- \leftarrow \tau\theta + (1 - \tau)\theta^-$ $\beta^- \leftarrow \tau\beta + (1 - \tau)\beta^-$

**end**

**return** Policies $\pi_i(a_i|s) = \arg\max_{a_i} Q_{local}(o_i, a_i; \theta)$

**Algorithm 2:** Decision making of Agent $i$ in Herm

**Initialize:** Local Q Network $Q_{local}$, Observation Projection $\Phi$, Action Projection $\Psi$, Trajectory Encoder $GRU$, Attribute Dictionary $\mathcal{A}$, Value Dictionary $\mathcal{V}$, Action Dictionary $\mathcal{A}ct$, Time Dictionary $\mathcal{T}$, Length of action sequence $\mathcal{W}$

// 1. Message Encoding
// 1.1 Observation Transformation
$HV_O \leftarrow Empty$
**foreach** $\mathbf{e}_j \in Entity$ **do**
  $o_i^j \leftarrow f(o_i)$ ; // Entity-wised Extraction
  $HV_o^j \leftarrow Empty$
  **foreach** $(attr, val) \in o_i^j$ **do**
    $HV_{attr} \leftarrow find(\mathcal{A}, attr)$
    $HV_{val} \leftarrow find(\mathcal{V}, attr)$
    $HV_o^j = HV_o^j \oplus (HV_{attr} \otimes HV_{val})$
  **end**
  $HV_o^j = HV_o^j \otimes HV_{id}^j$
  $HV_O = HV_O \oplus HV_o^j$
**end**
// 1.2 Action Sequence Transformation
$HV_A \leftarrow Empty$
**foreach** $w \in \mathcal{W}$ **do**
  $HV_a^{-w} \leftarrow find(\mathcal{A}ct, a_i^{-w})$
  $HV_t^{-w} \leftarrow find(\mathcal{T}, -w)$
  $HV_A = HV_A \oplus (HV_a^{-w} \otimes HV_t^{-w})$
**end**
// 2. Message Fusion
$\alpha_i = cosine_sim(HV_O, HV_O^{-i})$
$\boldsymbol{\alpha} = softmax(\alpha_i)$
$\dot{\boldsymbol{\alpha}} = softmax(-\alpha_i)$
$c_o^i = \Phi(\sum_j \dot{\boldsymbol{\alpha}}_{i,j} * (HV_O^i \otimes HV_O^{i,j}))$
$c_a^i = \Psi(\sum_j \boldsymbol{\alpha}_{i,j} * (HV_A^i \otimes HV_A^{i,j}))$
$\boldsymbol{c}_i = (c_o^i, c_a^i)$
// 3. Trajectory Encoding
$h_i^t \leftarrow GRU(h_i^{t-1}, o_i^t, a_i^{t-1})$
**return** $Q_{local}(h_i^t, \boldsymbol{c}_i)$

## NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: see Abstract and Introduction.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: see Conclusion.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

Justification: see Appendix.

Guidelines:

- The answer [N/A] means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: see Appendix.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [Yes]

Justification: See Supplementary Material.

Guidelines:

- The answer [N/A] means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [Yes]

   Justification: see Appendix.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes]

   Justification: see Section Experiments.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
   - The assumptions made should be given (e.g., Normally distributed errors).
   - It should be clear whether the error bar is the standard deviation or the standard error of the mean.

- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: see Appendix.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: See Supplementary Material.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

    Justification: See Conclusion.

    Guidelines:

    - The answer [N/A] means that there is no societal impact of the work performed.
    - If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.

- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [N/A]

    Justification: Do not involve.

    Guidelines:

    - The answer [N/A] means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

    Justification: see Appendix.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.

- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [N/A]

    Justification: Do not involve.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: Do not involve.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: Do not involve.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

Answer: [Yes]

Justification: see Appendix H.

Guidelines:

- The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.