

## A Related Work

**Value Decomposition.** Centralized training with decentralized execution (CTDE) has emerged as a powerful paradigm in MARL [41, 42], where global information can be accessed during centralized training and learned policies are executed with only local information in a decentralized way. Under the CTDE paradigm, value decomposition methods show their strength in expressing the joint value function conditioned on individual value functions. VDN [21] introduces a linear decomposition, representing the joint Q-value as a sum of agent-wise Q-values. However, its additive nature ignores inter-agent interactions, limiting its scalability to complex coordination tasks. QMIX [11] improves representational capacity by employing a nonlinear monotonic mixing network parameterized via hypernetworks, but the imposed monotonicity constraint hinders its flexibility. To overcome this, QTRAN [30] introduces a relaxed transformation-based decomposition to bypass monotonicity, while WQMIX [35] incorporates a weighted projection to enhance approximation quality. QPLEX [34] further refines the decomposition by adopting a duplex dueling architecture that satisfies the Individual-Global-Max (IGM) principle via an advantage-based formulation. Despite their improvements in expressiveness, these methods primarily focus on functional accuracy and provide little insight into the underlying coordination structure. This lack of interpretability becomes particularly problematic in partially observable and interaction-intensive environments, where understanding agent dependencies is crucial for robust credit assignment. To address this, we propose a novel interpretable value decomposition framework that explicitly encodes high-order interactions, offering both performance and transparency.

**Interpretable MARL.** Recent advances in interpretable MARL can be broadly categorized into two paradigms, focusing either on (i) intrinsic interpretability or (ii) post-hoc explanation [43]. Intrinsic interpretability requires the learned model to be self-understandable by nature, which is achieved by using a transparent class of models, whereas post-hoc explanation entails learning a second model to explain an already-trained black-box model. Post-hoc methods provide auxiliary insights without modifying the underlying learning process. For instance, SQDDPG [12] estimates individual agent contributions via Shapley Q-values, while Goto et al. [13] use masked-attention to identify salient observation regions in multi-vehicle coordination tasks. Although informative, these techniques lack robustness guarantees and struggle to recover the relational or temporal structure intrinsic to multi-agent cooperation. In contrast, intrinsically interpretable approaches seek to construct models whose decision logic is understandable by design. Tree-based architectures such as MIXRT [15] and MAVIPER [14] represent agent policies using soft or symbolic decision trees, providing explicit reasoning paths. DTPO [16] advances this line by directly optimizing tree structures via policy gradients, combining transparency with performance. Attention-based models, such as MAAC [44], further enhance interpretability by dynamically identifying inter-agent dependencies, while other methods promote explainability through latent skill inference [45] or constrained policy spaces that encode global objectives [46].

Within the value decomposition framework, central to cooperative MARL, several works also try to understand how agents cooperate via agent-level contributions. VDN [21] factorizes the team reward additively, assuming agent independence. SHAQ [37] adopts Shapley value theory to quantify marginal contributions across coalitions. More recently, NA<sup>2</sup>Q [22] expands the joint value function via a Taylor-like decomposition to capture high-order interactions. However, such expansions scale exponentially with the number of agents, leading to substantial computational and interpretability challenges. These limitations highlight the need for a more principled and scalable formulation that can compactly model high-order agent interactions without sacrificing transparency. To this end, we propose a novel approach that integrates continued fraction networks into the value decomposition framework. By leveraging the recursive structure of continued fractions, our method enables compact and interpretable representations of arbitrary-order interactions while maintaining linear complexity with respect to the number of agents. This formulation provides a powerful alternative to polynomial expansion-based methods, offering both expressive capacity and interpretability in large-scale cooperative MARL settings.

## B Proof

### B.1 Objective Functions for Variational Information Bottlenecks

Considering the Markov chain  $\mathbf{u}^* \leftrightarrow \mathbf{h} \leftrightarrow \mathbf{m}$ , which means the assistive information cannot depend directly on the  $\mathbf{u}^*$ . So we have  $p(\mathbf{m} \mid \mathbf{h}, \mathbf{u}^*) = p(\mathbf{m} \mid \mathbf{h})$ .

As in the IB, the objective can be written as:

$$J_{IB}(\phi) = I(\mathbf{m}, \mathbf{u}^*; \phi) - \beta I(\mathbf{m}, \mathbf{h}; \phi). \quad (15)$$

The  $\beta$  is to realize the trade-off between a succinct representation and inferencing ability.

**Theorem 1** (Lower Bound for  $I(\mathbf{m}, \mathbf{u}^*; \phi)$ ). *Let the representation  $m_i$  be reparameterized as a random variable drawn from a multivariate Gaussian distribution  $m_i \sim \mathcal{N}(f_m(h_i; \phi_m), \mathbf{I})$ , where  $f_m$  is an encoder parameterized by  $\phi_m$ ,  $h_i$  denotes the hidden state of agent  $i$ , and  $\mathbf{I}$  is the identity covariance matrix. Then, the mutual information between the assistive information  $\mathbf{m}$  and the optimal joint action  $\mathbf{u}^*$  is lower-bounded as:*

$$I(\mathbf{m}, \mathbf{u}^*; \phi) \geq \frac{1}{N} \sum_{i=1}^N \mathbb{E}_{\epsilon \sim p(\epsilon)} [-\log q(u_i^* \mid f(h_i, \epsilon))], \quad (16)$$

where  $q(u_i^* \mid m_i)$  is a variational distribution approximating the true posterior  $p(u_i^* \mid m_i)$ , and  $m_i = f(h_i, \epsilon)$  denotes a deterministic function of  $h_i$  and the Gaussian random variable  $\epsilon$ .

*Proof.*

$$\begin{aligned} I(\mathbf{m}, \mathbf{u}^*; \phi) &= \int dm_i du_i^* p(m_i, u_i^*) \log \frac{p(m_i, u_i^*)}{p(m_i) p(u_i^*)} \\ &= \int dm_i du_i^* p(m_i, u_i^*) \log \frac{p(u_i^* \mid m_i)}{p(u_i^*)}, \end{aligned}$$

where  $p(u_i^* \mid m_i)$  is fully defined by our encoder and Markov Chain as follows:

$$\begin{aligned} p(u_i^* \mid m_i) &= \int dh_i p(h_i, u_i^* \mid m_i) \\ &= \int dh_i p(u_i^* \mid h_i) p(h_i \mid m_i) \\ &= \int dh_i \frac{p(u_i^* \mid h_i) p(m_i \mid h_i) p(h_i)}{p(m_i)}. \end{aligned}$$

Since this is intractable in our case, let  $q(u_i^* \mid m_i)$  be a variational approximation to  $p(u_i^* \mid m_i)$ , where this is our decoder which we will take to another neural network with its own set of parameters. Using the fact that Kullback Leibler divergence is always positive, we have

$$\begin{aligned} \text{KL}[p(u_i^* \mid m_i), q(u_i^* \mid m_i)] &\geq 0 \\ \implies \int du_i^* p(u_i^* \mid m_i) \log p(u_i^* \mid m_i) &\geq \int du_i^* p(u_i^* \mid m_i) \log q(u_i^* \mid m_i), \end{aligned}$$

and hence

$$\begin{aligned} I(\mathbf{m}, \mathbf{u}^*; \phi) &\geq \int du_i^* dm_i p(u_i^*, m_i) \log \frac{q(u_i^* \mid m_i)}{p(u_i^*)} \\ &= \int du_i^* dm_i p(u_i^*, m_i) \log q(u_i^* \mid m_i) - \int du_i^* p(u_i^*) \log p(u_i^*) \\ &= \int du_i^* dm_i p(u_i^*, m_i) \log q(u_i^* \mid m_i) + H(u_i^*) \\ &= \int dh_i du_i^* dm_i p(h_i) p(u_i^* \mid h_i) p(m_i \mid h_i) \log q(u_i^* \mid m_i) + H(u_i^*) \\ &= \frac{1}{N} \sum_{i=1}^N \left[ \int dm_i p(m_i \mid \tau_i) \log q(u_i^* \mid m_i) \right] + H(u_i^*). \end{aligned}$$

Notice that the entropy of our labels  $H(u_i^*)$  is independent of our optimization procedure and so can be ignored. And as we can rewrite  $p(m_i | h_i) dm_i = p(\epsilon) d\epsilon$ ,  $m_i = f(h_i, \epsilon)$ . So we have

$$I(\mathbf{m}, \mathbf{u}^*; \phi) \geq \frac{1}{N} \sum_{i=1}^N \mathbb{E}_{\epsilon \sim p(\epsilon)} [-\log q(u_i^* | f(h_i, \epsilon))].$$

□

**Theorem 2** (Upper Bound for  $I(\mathbf{m}, \mathbf{h}; \phi)$ ). *Let  $\tilde{q}(\mathbf{m})$  denote a variational approximation of the marginal distribution  $p(\mathbf{m})$ . Then, the mutual information between the representation  $\mathbf{m}$  and the hidden state  $\mathbf{h}$  admits the following upper bound:*

$$I(\mathbf{m}, \mathbf{h}; \phi) \leq \text{KL}(p(\mathbf{m} | \mathbf{h}_i) \parallel \tilde{q}(\mathbf{m})). \quad (17)$$

*Proof.*

$$\begin{aligned} I(\mathbf{m}, \mathbf{h}; \phi) &= \int dm_i dh_i p(h_i, m_i) \log \frac{p(m_i | h_i)}{p(m_i)} \\ &= \int dm_i dh_i p(h_i, m_i) \log p(m_i | h_i) - \int dm_i p(m_i) \log p(m_i). \end{aligned}$$

Let  $\tilde{q}(m_i)$  be the variational approximation to the marginal distribution  $p(m_i) = \int dh_i p(m_i | h_i) p(h_i)$ . Since  $\text{KL}[p(m_i), \tilde{q}(m_i)] \geq 0 \implies \int dm_i p(m_i) \log p(m_i) \geq \int dm_i p(m_i) \log \tilde{q}(m_i)$ , we have

$$\begin{aligned} I(\mathbf{m}, \mathbf{h}; \phi) &\leq \int dh_i dm_i p(h_i) p(m_i | h_i) \log \frac{p(m_i | h_i)}{\tilde{q}(m_i)} \\ &= \frac{1}{N} \sum_{i=1}^N \left[ p(m_i | h_i) \log \frac{p(m_i | h_i)}{\tilde{q}(m_i)} \right] \\ &= \text{KL}[p(\mathbf{m} | \mathbf{h}_i), \tilde{q}(\mathbf{m})]. \end{aligned}$$

□

Combining *Theorem 1* and *Theorem 2*, we have the objective functions for variational information bottlenecks, which is to minimize

$$\mathcal{L}_{VIB} = \frac{1}{N} \sum_{i=1}^N \mathbb{E}_{\epsilon \sim p(\epsilon)} [-\log q(u_i^* | f(h_i, \epsilon))] + \beta \text{KL}[p(\mathbf{m} | \mathbf{h}_i), \tilde{q}(\mathbf{m})]. \quad (18)$$

## B.2 Correspondence between Continued Fraction Depth and the Order of Agent Interactions

In this section, we establish a one-to-one correspondence between the depth  $d$  of the continued fraction network and the order of agent interactions. This property allows the  $d$ -th order continued fraction to accurately represent the  $d$ -th order approximation of the agent's behavior.

Specifically, a continued fraction network of depth  $d$ ,  $\frac{1}{w_1 Q +} \frac{1}{w_2 Q +} \cdots \frac{1}{w_d Q}$  can be reformulated as  $f(Q) = T_d(Q) + \mathcal{O}_{d+1}(Q)$ , where  $T_n(Q)$  is a degree- $d$  polynomial of  $Q$ , and  $\mathcal{O}_{d+1}(Q)$  denotes terms of order  $d+1$  or higher in  $Q$ .

By setting  $z = \frac{1}{Q}$ , the continued fraction  $\frac{1}{w_1 Q +} \frac{1}{w_2 Q +} \frac{1}{w_3 Q +} \cdots$  can be transformed into

$$\mathbf{K}(z) = \frac{z}{w_1} \frac{z}{w_2} \frac{z}{w_3 + \cdots}, \quad (19)$$

since these approximants are arranged along the "staircase diagonals" of the Padé table.

**Theorem 3.** *For the  $d$ -th order truncation of the continued fraction  $R_k(z) = \frac{A_k(z)}{B_k(z)}$ , the following holds:*

$$p_d = \left\lfloor \frac{d+1}{2} \right\rfloor, \quad q_d = \left\lfloor \frac{d}{2} \right\rfloor, \quad (20)$$

where  $p_d = \deg(A_d)$ ,  $q_d = \deg(B_d)$ .

*Proof.* The  $k$ -th order asymptotic function  $\frac{A_k(z)}{B_k(z)}$  satisfies the recursive relations:

$$\begin{cases} A_k(z) = \mathbf{w}_k A_{k-1}(z) + z A_{k-2}(z) \\ B_k(z) = \mathbf{w}_k B_{k-1}(z) + z B_{k-2}(z) \end{cases},$$

with

$$\begin{bmatrix} A_{-1} & A_0 \\ B_{-1} & B_0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Assume that for all  $k \leq n$ , the following holds:

$$\deg(A_k) = \left\lfloor \frac{k+1}{2} \right\rfloor, \quad \deg(B_k) = \left\lfloor \frac{k}{2} \right\rfloor.$$

**Base Cases:**

$n = 1$ :

$$A_1(z) = \mathbf{w}_1 A_0(z) + z A_{-1}(z) = z \Rightarrow \deg(A_1) = 1, \quad \left\lfloor \frac{1+1}{2} \right\rfloor = 1.$$

$$B_1(z) = \mathbf{w}_1 B_0(z) + z B_{-1}(z) = \mathbf{w}_1 \Rightarrow \deg(B_1) = 0, \quad \left\lfloor \frac{1}{2} \right\rfloor = 0.$$

$n = 2$ :

$$A_2(z) = \mathbf{w}_2 A_1(z) + z A_0(z) = \mathbf{w}_2 z \Rightarrow \deg(A_2) = 1, \quad \left\lfloor \frac{3}{2} \right\rfloor = 1.$$

$$B_2(z) = \mathbf{w}_2 B_1(z) + z B_0(z) = \mathbf{w}_1 \mathbf{w}_2 + z \Rightarrow \deg(B_2) = 1, \quad \left\lfloor \frac{2}{2} \right\rfloor = 1.$$

Hence, the base cases hold.

Then assume the statements hold for  $k = n - 1$  and  $k = n - 2$ , and prove that they also hold for  $k = n$ .

**Degree of the Numerator  $A_n(z)$ :**

From the recurrence:

$$A_n(z) = \mathbf{w}_n A_{n-1}(z) + z A_{n-2}(z),$$

and by the induction hypothesis:

$$\deg(A_{n-1}) = \left\lfloor \frac{n}{2} \right\rfloor, \quad \deg(A_{n-2}) = \left\lfloor \frac{n-1}{2} \right\rfloor.$$

Then:

$$\deg(\mathbf{w}_n A_{n-1}) = \left\lfloor \frac{n}{2} \right\rfloor, \quad \deg(z A_{n-2}) = 1 + \left\lfloor \frac{n-1}{2} \right\rfloor.$$

Case 1:  $n$  even, let  $n = 2m$ :

$$\left\lfloor \frac{n}{2} \right\rfloor = m, \quad \left\lfloor \frac{n-1}{2} \right\rfloor = m - 1.$$

Then

$$\deg(\mathbf{w}_n A_{n-1}) = m, \quad \deg(z A_{n-2}) = 1 + (m - 1) = m \Rightarrow \deg(A_n) = m = \left\lfloor \frac{n+1}{2} \right\rfloor.$$

Case 2:  $n$  odd, let  $n = 2m + 1$ :

$$\left\lfloor \frac{n}{2} \right\rfloor = m, \quad \left\lfloor \frac{n-1}{2} \right\rfloor = m.$$

Then

$$\deg(\mathbf{w}_n A_{n-1}) = m, \quad \deg(z A_{n-2}) = 1 + m = m + 1 \Rightarrow \deg(A_n) = m + 1 = \left\lfloor \frac{n+1}{2} \right\rfloor.$$



**Degree of the Denominator  $B_n(z)$ :**

From the recurrence:

$$B_n(z) = w_n B_{n-1}(z) + z B_{n-2}(z),$$

and using the induction hypothesis:

$$\deg(B_{n-1}) = \left\lfloor \frac{n-1}{2} \right\rfloor, \quad \deg(B_{n-2}) = \left\lfloor \frac{n-2}{2} \right\rfloor.$$

Then:

$$\deg(w_n B_{n-1}) = \left\lfloor \frac{n-1}{2} \right\rfloor, \quad \deg(z B_{n-2}) = 1 + \left\lfloor \frac{n-2}{2} \right\rfloor.$$

Case 1:  $n = 2m$  (even):

$$\left\lfloor \frac{n-1}{2} \right\rfloor = m-1, \quad \left\lfloor \frac{n-2}{2} \right\rfloor = m-1.$$

Then

$$\deg(w_n B_{n-1}) = m-1, \quad \deg(z B_{n-2}) = m \Rightarrow \deg(B_n) = m = \left\lfloor \frac{n}{2} \right\rfloor.$$

Case 2:  $n = 2m+1$  (odd):

$$\left\lfloor \frac{n-1}{2} \right\rfloor = m, \quad \left\lfloor \frac{n-2}{2} \right\rfloor = m-1.$$

Then

$$\deg(w_n B_{n-1}) = m, \quad \deg(z B_{n-2}) = m \Rightarrow \deg(B_n) = m = \left\lfloor \frac{n}{2} \right\rfloor.$$

By mathematical induction, we conclude that:

$$\deg(A_n) = \left\lfloor \frac{n+1}{2} \right\rfloor, \quad \deg(B_n) = \left\lfloor \frac{n}{2} \right\rfloor.$$

Therefore, when the truncation order is  $n = d$ , we have

$$p_d = \left\lfloor \frac{d+1}{2} \right\rfloor, \quad q_d = \left\lfloor \frac{d}{2} \right\rfloor.$$

□

**Theorem 4.** The  $d$ -th order truncation of the continued fraction  $R_d(z) = \frac{A_d(z)}{B_d(z)}$  naturally satisfies the conditions for a Padé approximant, specifically:

$$f(z) - R_d(z) = \mathcal{O}(z^{p_d+q_d+1}), \quad (21)$$

which means that its Taylor expansion coincides with the first  $d$  terms of the original function  $f(z)$ .

*Proof.*

**Definition 1** (Padé Approximant [25, 26]). Let  $C(z) = \sum_{k=0}^{\infty} c_k z^k$  be a formal power series in the variable  $z$ , then the Padé approximant of order  $[L/M]$  is a rational function of the form:

$$R_{L,M}(z) = [A_{L,M}(z)]/[B_{L,M}(z)], \quad (22)$$

where  $A_{L,M}(z)$  and  $B_{L,M}(z)$  are polynomials of degrees at most  $L$  and  $M$ , respectively, chosen such that

$$B_{L,M}(z)C(z) - A_{L,M}(z) = \mathcal{O}(z^{L+M+1}), \quad (23)$$

where notation  $\mathcal{O}(z^k)$  denotes some power series of the form  $\sum_{n=k}^{\infty} \tilde{c}_n z^n$ . This approximation minimizes the difference between the rational function and the power series up to the order  $L+M$ .

Since  $R_d(z) = \frac{A_d(z)}{B_d(z)}$ , we have

$$f(z) - \frac{A_d(z)}{B_d(z)} = \mathcal{O}(z^{p_d+q_d+1}), \quad (24)$$

which implies

$$f(z)B_d(z) - A_d(z) = \mathcal{O}(z^{p_d+q_d+1}) B_d(z) = \mathcal{O}(z^{p_d+q_d+1}). \quad (25)$$

**Lemma 1.** For all  $k \geq 0$ , there exists a polynomial  $S_k(z)$  such that:

$$f(z)B_k(z) - A_k(z) = (-1)^k z^{k+1} S_k(z), \quad (26)$$

and the constant term  $S_k(0) \neq 0$ .

*Proof.* We can prove this lemma by mathematical induction.

**Base Case:**

$k = 0$ :

$$f(z)B_0(z) - A_0(z) = f(z) \cdot 1 - 0 = f(z).$$

By the definition of continued fractions,  $f(z) = \frac{z}{a_1 + \dots}$ , so:

$$f(z) = z \cdot (\text{analytic function}) = zS_0(z), \quad S_0(0) = \frac{1}{a_1} \neq 0.$$

**Inductive Hypothesis ( $k-1$  and  $k-2$  hold):**

$$\begin{aligned} f(z)B_{k-1}(z) - A_{k-1}(z) &= (-1)^{k-1} z^k S_{k-1}(z) \\ f(z)B_{k-2}(z) - A_{k-2}(z) &= (-1)^{k-2} z^{k-1} S_{k-2}(z). \end{aligned}$$

**for  $k$ :**

Substituting the recurrence relations:

$$\begin{aligned} f(z)B_k(z) - A_k(z) &= f(z)(w_k B_{k-1}(z) + zB_{k-2}(z)) - (w_k A_{k-1}(z) + zA_{k-2}(z)) \\ &= w_k (f(z)B_{k-1}(z) - A_{k-1}(z)) + z(f(z)B_{k-2}(z) - A_{k-2}(z)) \\ &= w_k (-1)^{k-1} z^k S_{k-1}(z) + z(-1)^{k-2} z^{k-1} S_{k-2}(z) \\ &= (-1)^k z^k (-w_k S_{k-1}(z) + S_{k-2}(z)) \\ &= (-1)^k z^{k+1} S_k(z), \end{aligned}$$

where  $S_k(z)$  is the polynomial obtained from  $-w_k S_{k-1}(z) + S_{k-2}(z)$ , divided by  $z$ . □

From the lemma, we have:

$$f(z) - \frac{A_d(z)}{B_d(z)} = \frac{f(z)B_d(z) - A_d(z)}{B_d(z)} = \frac{(-1)^d z^{d+1} S_d(z)}{B_d(z)}.$$

Since  $B_d(0) = w_1 w_2 \cdots w_d \neq 0$  (assuming  $w_k \neq 0$ ) and  $S_d(0) \neq 0$ , it follows that:

$$f(z) - R_d(z) = \mathcal{O}(z^{d+1}).$$

According to *Theorem 3*, the degree of the numerator  $A_d(z)$  is  $p_d = \lfloor \frac{d+1}{2} \rfloor$ , and the degree of the denominator  $B_d(z)$  is  $q_d = \lfloor \frac{d}{2} \rfloor$ . When  $d$  is odd, we have  $p_d = q_d = \frac{d}{2}$ ; when  $d$  is even,  $p_d = \frac{d+1}{2}$ ,  $q_d = \frac{d-1}{2}$ . In both cases, it follows that  $p_d + q_d = d$ . Therefore,

$$f(z) - R_d(z) = \mathcal{O}(z^{d+1}) = \mathcal{O}(z^{p_d+q_d+1}),$$

which satisfies the condition of a Padé approximant. □

In conclusion, the depth- $d$  continued fraction network represents the  $d$ -th order truncation of the continued fraction:

$$\frac{1}{w_1 Q +} \frac{1}{w_2 Q +} \cdots \frac{1}{w_d Q},$$

which forms a  $[p_d, q_d]$ -Padé approximant with  $p_d + q_d = d$ . This enables accurate representation of the first  $d$ -th order interactions among agents.

## C Experimental Details

### C.1 Algorithmic Description

---

**Algorithm 1** Continued Fraction Q-Learning

---

```

1: Initialize environment, agent network  $Q_i(\tau_i, u_i; \theta)$ , target network  $Q_i(\tau'_i, u'_i; \hat{\theta})$ , mixing net-
   work  $Q_{tot}$ , and VIB module  $G_\phi(E_{\phi_1}, D_{\phi_2})$ 
2: Initialize replay buffer  $\mathcal{D}$ 
3: repeat
4:   Obtain the initial global state  $s^0$ 
5:   for  $t = 0$  to  $T - 1$  do
6:     For each agent  $i$ , get action-observation history  $\tau_i^t$ 
7:     Calculate individual value function  $Q_i$ 
8:     Get the hidden state  $h_i^t$ 
9:     Select action  $u_i^t$  via value function with probability  $\epsilon$  exploration
10:    Execute joint action  $\mathbf{u}^t$ , receive reward  $r^t$ , next state  $s^{t+1}$ 
11:  end for
12:  Store the episode trajectory in  $\mathcal{D}$ 
13:  Sample a mini-batch  $\mathcal{B}$  of size  $b$  from  $\mathcal{D}$ 
14:  for  $t = 0$  to  $T - 1$  do
15:    Calculate  $\mu, \sigma = E_{\phi_1}(h_i^t)$ 
16:    Generate assistive information  $\mathbf{m}$ 
17:    Get the attention weight  $\alpha_k$  by the intervention function in Eq. 12
18:    Calculate the joint value function  $Q_{tot}$ 
19:  end for
20:  Calculate loss  $\mathcal{L}(\theta) = \mathcal{L}_{Q_{tot}} + \mathcal{L}_{VIB}$  via Eq. 11 and Eq. 14.
21:  Update  $\phi$  and  $\theta$  by minimizing the above loss
22:  Periodically update  $\hat{\theta} \leftarrow \theta$ 
23: until  $Q_i(\tau_i, u_i; \theta)$  converges or maximum steps reached

```

---

### C.2 LBF Description and Hyperparameters Settings

Level-Based Foraging (LBF) [31] is a mixed cooperative-competitive MARL benchmark, where each agent navigates a  $10 \times 10$  grid world. Agents and food items are randomly placed in a 2D grid, and each one is assigned a level. A food item can only be collected when the combined levels of all participating agents equal or exceed its level. The environment induces a spectrum of collaborative behaviors through its level-dependent reward structure: while low-level food items permit independent collection, higher-level resources necessitate coalition formation. Furthermore, we set the penalty reward for movement to  $-0.002$ , and the detailed hyperparameter settings of LBF are shown in Table 1.

**Observation Space.** Each agent observes a  $2 \times 2$  square grid centered on its own position. Within this range, the agent receives a structured array containing the  $(x, y)$  coordinates and levels of all visible food items and other agents. This observation provides both spatial and attribute-level information to support localized decision-making.

**Action Space.** The discrete action space for each agent consists of none, move [direction], and load food. Each agent only moves into one unoccupied grid. If multiple agents attempt to move into the same grid, collisions are resolved by canceling the conflicting moves, leaving the agents in their original positions.

Table 1: The configurations of LBF.

Hyperparameter	Value
Max Episode Length	50
Batch Size	32
Test Interval	10000
Test Episodes	32
Replay Batch Size	5000
Discount Factor	0.99
Start Exploration Rate	1.0
End Exploration Rate	0.05
Anneal Steps	50000
Steps	1M
Target Update Interval	200

**Rewards.** This reward depends on the food’s level, which is distributed among the participating agents in proportion based on their levels. The rewards are normalized to maintain a unit sum across all agents. This design ensures contribution-based fairness in reward distribution while enhancing cooperative efficiency among agents.

### C.3 StarCraft II Description and Hyperparameters Settings

All implementations of algorithms are conducted on StarCraft II unit micro-management tasks (SC2.4.10). We evaluate performance in combat scenarios where enemy units are controlled by the built-in AI with the *difficulty*=7 setting, and each allied unit is controlled by the decentralized agents with reinforcement learning algorithms. During battles, the agents seek to maximize the damage dealt to enemy units while minimizing damage received, requiring the coordination of diverse tactical skills. We assess our method across a variety of challenging scenarios that differ in terms of symmetry, agent composition, and unit count (as shown in Table 3). For clarity, we also outline the core settings of the StarCraft Multi-Agent Challenge (SMAC) [32], including observation, state, action, and reward configurations. The detailed hyperparameter settings of SMAC are shown in Table 2.

Table 2: The configurations of SMAC.

Hyperparameter	Value
Difficulty	7
Batch Size	32
Test Interval	10000
Test Episodes	32
Replay Batch Size	5000
Discount Factor	0.99
Start Exploration Rate	1.0
End Exploration Rate	0.05
Target Update Interval	200
Optimizer	RMSprop
Learning Rate	0.0005

**Observations and States.** At each time step, each agent receives a local observation of units within its field of view. The observation includes the following features for both allied and enemy units: distance, relative X and Y positions, health, shield, and unit type. Note that the agents can only observe the others if they are alive and within their line of sight range, which is set to 9. When a unit (ally or enemy) becomes invisible or is eliminated, its feature vector is reset to all zeros, indicating either death or being outside the field of view. The global state is only available during centralized training, which contains information about all units on the map. Finally, all features, including the global state and the observation of the agent, are normalized by their maximum values.

**Action Space.** Each unit takes an action from the discrete action set: no-op, stop, move [direction], and attack [enemy id]. Agents are allowed to move with a fixed movement amount in four directions: north, south, east, and west, where the unit is allowed to take the attack [enemy id] action only when the enemy is within its shooting range.

**Rewards.** The target goal is to maximize the win rate for each battle scenario. At each time step, the agents receive a shaped reward based on the hit-point damage dealt and enemy units killed, as well as a special bonus for winning the battle. Additionally, agents obtain a 10 positive bonus after killing each enemy and a 200 bonus when killing all enemies, which is consistent with the default reward function of the SMAC.

Table 3: The StarCraft Multi-Agent Challenge benchmark.

Map	Ally Units	Enemy Units	Difficulty	Steps	Anneal Steps	$d$
2s3z	2 Stalkers, 3 Zealots	2 Stalkers, 3 Zealots	Eazy	1.5M	50000	2
2c_vs_64zg	2 Colossus	64 Zerglings	Hard	2M	50000	2
3s_vs_5z	3 Stalkers	5 Zealots	Hard	2M	50000	2
5m_vs_6m	5 Marines	6 Marines	Hard	2M	50000	4
3s5z_vs_3s6z	3 Stalkers, 5 Zealots	3 Stalkers, 6 Zealots	Super Hard	5M	200000	6
6h_vs_8z	6 Hydralisks	8 Zealots	Super Hard	5M	200000	6

#### C.4 SMACv2 Description and Hyperparameters Settings

SMACv2 [33] is an enhanced benchmark for cooperative multi-agent reinforcement learning built on top of StarCraft II. It preserves the original SMAC API while introducing three procedural innovations to increase scenario diversity and challenge contemporary MARL algorithms: randomising start positions, randomising unit types, and changing the unit sight and attack ranges.

**Randomized Start Positions.** Allied and enemy units are spawned either in a "surround" configuration, where enemies encircle the allies, or via a "reflect" scheme that mirrors allied positions across the map center. This ensures that agents cannot overfit to fixed spawn patterns.

**Randomized Unit Types.** Each battle can feature mixed unit compositions rather than uniform rosters. For Terran, Protoss, and Zerg, three unit types are sampled with configurable probabilities through the `team_gen` distribution (as shown in Table 4), promoting adaptable strategies under varied team makeups.

**Unit Sight and Attack Ranges.** Unit vision cones and attack radii are aligned with their true in-game values, increasing realism and preventing agents from exploiting the simplified ranges used in SMAC.

Table 4: The configurations of SMACv2.

Race	Unit	probability
Protoss	Stalker	0.45
	Zealot	0.45
	Colossus	0.1
Terran	Marine	0.45
	Marauder	0.45
	Medivac	0.1
Zerg	Zergling	0.45
	Hydralisk	0.45
	Baneling	0.1

#### C.5 Implementation Details

We compare our method against nine value-based baselines, including VDN [21], QMIX [11], QPLEX [34], Centrally-Weighted QMIX (CW-QMIX) [35], CDS [36], SHAQ [37], GoMARL [38], ReBorn [39], and NA<sup>2</sup>Q [22]. To ensure fairness, we implement all experiments within the PyMARL framework<sup>2</sup>. All hyperparameters of baselines are set identically to our method to compare algorithms fairly. Please refer to PyMARL’s open-source implementation for further training details and fair comparison settings. The depth  $d$  of CFN is determined based on the scale of agents and the complexity of each task.

All scenarios are trained on a system equipped with an NVIDIA RTX 3080TI GPU and an Intel i9-12900k CPU, with training time ranging from 1 to 16 hours per scenario, depending on the task complexity and episode length.

<sup>2</sup>The source code of implementations is from <https://github.com/oxwhirl/pymarl>.

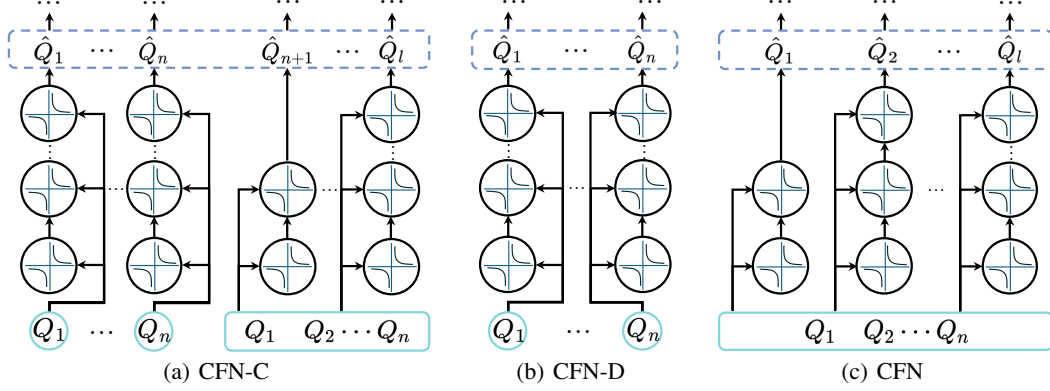


Figure 10: Three variants of the CFN architecture. (a) CFN-C integrates both single-feature ladders, where each ladder processes a single input dimension  $Q_i$ , and full ladders, which take the complete set of individual values  $Q$  as input. (b) CFN-D utilizes only the single-feature ladders. (c) CFN employs only the full ladders with increasing depth.

### C.6 Detailed Description of CFN Structure

As illustrated in *Fig. 10*, the CFN framework includes two structural variants in addition to the main architecture. *Fig. 10(a)* presents CFN-C, a composite architecture inspired by CoFrNet [18], which combines two types of ladders: single-feature ladders, each processing an individual agent utility  $Q_i$ , and full-input ladders, which receive the complete utility vector  $Q$  at every layer. Each ladder yields a partial joint value  $\hat{Q}_k$ , and the aggregation of all ladders constitutes the final joint Q-value.

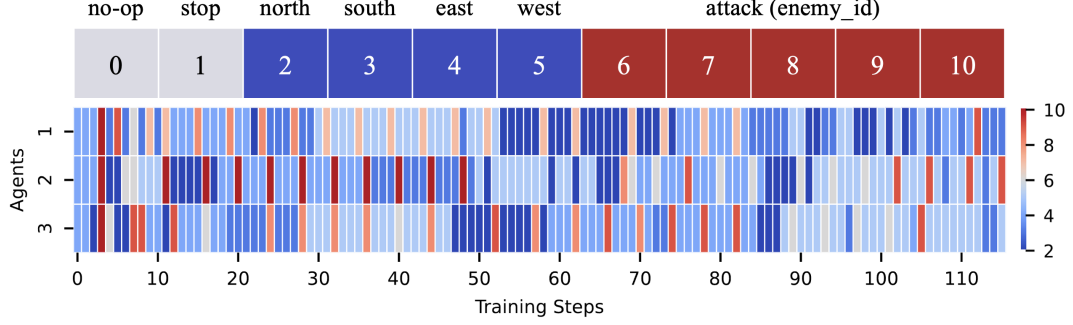
The number of single-feature ladders equals the number of agents, enabling additive modeling of individual effects. In contrast, full-input ladders are deeper and designed to capture complex joint dependencies among agents by recursively combining all inputs, thereby facilitating high-order interaction modeling.

*Fig. 10(b)* and *Fig. 10(c)* depict two simplified variants: 1) CFN-D, which retains only the single-feature ladders, thereby modeling additive effects with strong transparency [47] but lacking the capacity to express inter-agent interactions; 2) CFN, which retains only the full-input ladders, striking a balance between modeling power and computational efficiency.

In our QCoFr algorithm, we adopt the CFN structure with full-input ladders as the default architecture. Compared to CFN-C, this version significantly reduces parameter overhead while preserving the ability to capture arbitrary-order interactions. We further include CFN-D as an ablation baseline to isolate the contribution of high-order modeling: while CFN-D offers interpretability due to its decomposable additive form, its inability to encode dependencies across agents limits its expressiveness in cooperative settings.

Finally, a key advantage of the CFN structure is its linear scalability: the number of parameters grows as  $\mathcal{O}(n)$  with the number of agents, making it particularly suited to large-scale MARL scenarios where modeling expressive joint behavior is critical without incurring prohibitive computational costs.





(a) Actions of each agent.



(b)  $t = 25$



(c)  $t = 75$



(d)  $t = 85$



(e)  $t = 100$

Figure 11: Visualization of evaluation results for QCoFr on 3s\_vs\_5z map. Agents demonstrate a coordinated kite-and-focus-fire strategy: agent 2 initially kites four enemies alone, while agent 1 and agent 3 eliminate another. Agent 3 then draws away two of the remaining enemies, enabling agent 1 and agent 2 to dispatch the others. Finally, all agents regroup to defeat the last enemies.

## D Extended Interpretability Analysis

Fig. 11 illustrates the interpretability of QCoFr on 3s\_vs\_5z scenario. At the beginning of the episode, agent 2 independently kites four enemies, creating a numerical advantage that enables agent 1 and agent 3 to quickly eliminate an isolated opponent. As a result, agent 2 receives the highest individual contribution score (1.337), while the strongest pairwise contribution is observed between agents 1 and 3 due to their effective coordination. As the engagement progresses, agent 3 draws two enemies away, allowing agent 1 and agent 2 to jointly take down the remaining targets. During this phase, the coalition contribution of agents 1 and 2 increases, and agent 3's individual contribution also rises as it delays the enemy. After these enemies are defeated, all three agents regroup to focus fire on the remaining units, resulting in a more balanced distribution of credit across agents. This case study demonstrates that the agents have learned a kite-and-focus-fire strategy. The alignment between observed behaviors and quantitative contribution values confirms the interpretability of QCoFr, which faithfully attributes both individual and coalition-level contributions with high-order interactions in executing complex cooperative tactics.

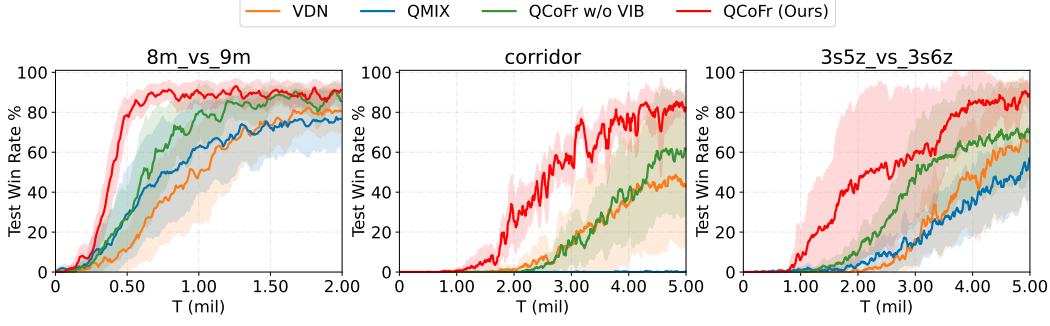


Figure 12: Performance with and without VIB on three extra scenarios of the SMAC benchmark.

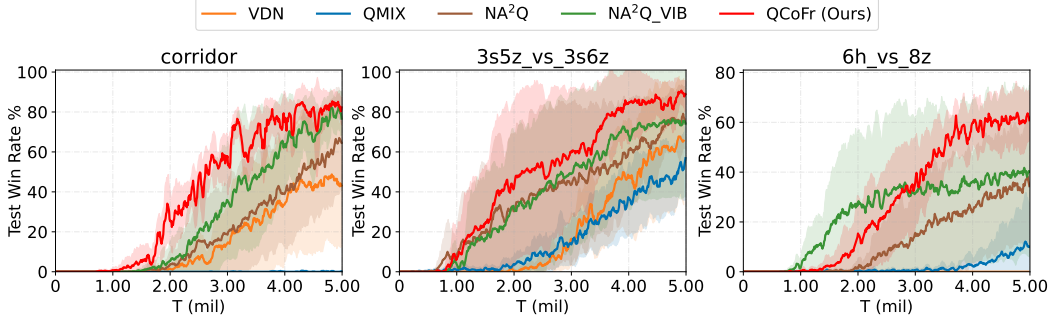


Figure 13: Performance comparison of  $NA^2Q$  with the VIB module and our method.

## E Additional Experiments on SMAC

### E.1 Additional Ablation Experiments

**The Role of the VIB Module.** We ablate the VIB component on three additional SMAC scenarios (Fig. 12), comparing QCoFr with and without VIB under identical settings. With VIB, QCoFr consistently accelerates early learning and achieves higher test win rates, confirming that task-relevant assistive information improves credit assignment and coordination.

**The Role of CFN.** Since  $NA^2Q$  struggles to model higher-order interactions, we equip it with the same VIB module and evaluate on three super-hard SMAC maps, comparing against QCoFr (Fig. 13). This isolates the effect of interaction modeling from that of assistive information. While  $NA^2Q$ +VIB outperforms the original  $NA^2Q$ , a clear gap remains to QCoFr. The results indicate that, on complex tasks, explicitly modeling higher-order dependencies enables more refined cooperative strategies, which highlights the effectiveness of the CFN module beyond what VIB alone can provide.

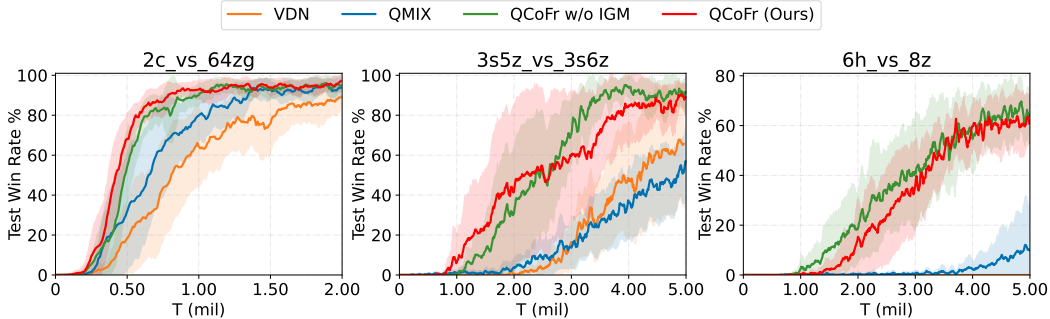


Figure 14: Performance comparison of QCoFr with and without IGM constraint.



## E.2 Discussion on the IGM Constraint

To isolate the effect of the continued-fraction mixing paradigm from non-monotonic joint-action search, we enforce the Individual-Global-Max (IGM) constraint in our framework. Notably, the universal approximation theorem applies to any linear combination of continued fractions and does not require non-negative weights [18], suggesting that the approach can be extended to non-IGM mixers. Integrating CFN with fully IGM-free methods such as DAVE [48] is therefore a natural direction. DAVE emphasizes that most value decomposition methods operate under IGM, which couples the optimal joint action with the optimal individual actions. Relaxing this constraint requires agents to explicitly search for the globally optimal joint action at execution time, often via an auxiliary network. To probe this possibility, we conduct experiments under relaxed IGM assumptions on three SMAC scenarios. As shown in *Fig. 14*, QCoFr achieves comparable or even slightly improved performance without IGM, indicating that our architecture can still recover high-quality joint actions.

## NeurIPS Paper Checklist

The checklist is designed to encourage best practices for responsible machine learning research, addressing issues of reproducibility, transparency, research ethics, and societal impact. Do not remove the checklist: **The papers not including the checklist will be desk rejected.** The checklist should follow the references and follow the (optional) supplemental material. The checklist does NOT count towards the page limit.

Please read the checklist guidelines carefully for information on how to answer these questions. For each question in the checklist:

- You should answer [Yes], [No], or [NA].
- [NA] means either that the question is Not Applicable for that particular paper or the relevant information is Not Available.
- Please provide a short (1-2 sentence) justification right after your answer (even for NA).

**The checklist answers are an integral part of your paper submission.** They are visible to the reviewers, area chairs, senior area chairs, and ethics reviewers. You will be asked to also include it (after eventual revisions) with the final version of your paper, and its final version will be published with the paper.

The reviewers of your paper will be asked to use the checklist as one of the factors in their evaluation. While "[Yes]" is generally preferable to "[No]", it is perfectly acceptable to answer "[No]" provided a proper justification is given (e.g., "error bars are not reported because it would be too computationally expensive" or "we were unable to find the license for the dataset we used"). In general, answering "[No]" or "[NA]" is not grounds for rejection. While the questions are phrased in a binary way, we acknowledge that the true answer is often more nuanced, so please just use your best judgment and write a justification to elaborate. All supporting evidence can appear either in the main paper or the supplemental material, provided in appendix. If you answer [Yes] to a question, in the justification please point to the section(s) where related material for the question can be found.

IMPORTANT, please:

- **Delete this instruction block, but keep the section heading "NeurIPS Paper Checklist",**
- **Keep the checklist subsection headings, questions/answers and guidelines below.**
- **Do not modify the questions and only use the provided macros for your answers.**

### 1. Claims

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

Answer: [Yes]

Justification: The main claims made in the abstract and introduction accurately reflect the paper's contributions and scope. The method can be found in Section C.6 and 4. Experimental results are illustrated in Section 5. The detailed proofs and experimental settings can be found in the Appendix.

Guidelines:

- The answer NA 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 NA 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: We include the limitations in our conclusion. We leave this part as our future work.

Guidelines:

- The answer NA 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: We have provided the full set of assumptions and corresponding complete proofs under each proposed Theorem with detailed derivations included in the Appendix B.

Guidelines:

- The answer NA 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: In addition to the description in the paper, we included more detailed hyperparameters settings in the Appendix.

Guidelines:

- The answer NA 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: [No]

Justification: The model architecture and hyperparameter settings are all included in this work. We believe we have provided enough details. We will make the code available in the near future.

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (<https://nips.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://nips.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, hyper-parameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

Answer: [\[Yes\]](#)

Justification: The details are included in Appendix.

Guidelines:

- The answer NA 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: For every experiment introduced, we run with multiple random seeds and reported both mean and std averaged over these multiple random seeds.

Guidelines:

- The answer NA 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: As mentioned in Appendix, our model runs on an NVIDIA RTX 3080TI GPU and an Intel i9-12900k CPU.

Guidelines:

- The answer NA 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: We are strictly following NeurIPS Code of Ethics.

Guidelines:

- The answer NA 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: [NA]

Justification: Our paper proposes a novel method for MARL, and we do not think there would be any direct social impact of it.

Guidelines:

- The answer NA means that there is no societal impact of the work performed.
- If the authors answer NA 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., pretrained language models, image generators, or scraped datasets)?

Answer: [NA]

Justification: The paper poses no such risks.

Guidelines:

- The answer NA 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: We cited all the creators or original owners of assets mentioned in the paper.

Guidelines:

- The answer NA 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](https://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: [NA]

Justification: The paper does not release new assets.

Guidelines:

- The answer NA 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: [NA]

Justification: The paper does not involve crowdsourcing nor research with human subjects.

Guidelines:

- The answer NA 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: [NA]

Justification: The paper does not involve crowdsourcing nor research with human subjects.

Guidelines:

- The answer NA 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: [NA]

Justification: The core method development in this research does not involve LLMs as any important, original, or non-standard components.

Guidelines:

- The answer NA 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 (<https://neurips.cc/Conferences/2025/LLM>) for what should or should not be described.