A Algorithms

Complexity Analysis. Here, we analyze the computational complexity of one layer in HAMP-I and HAMP-II. Analytically, the time complexity is $\mathcal{O}\left(|\mathcal{V}||\mathcal{E}|c^2 + |\mathcal{V}|c\right)$, where $|\mathcal{V}|$, $|\mathcal{E}|$ and c are the number of nodes, number of hyperedges and number of hidden dimension, respectively. However, the incidence matrix $\mathbf{H}$ is a sparse matrix, so the time complexity is $\mathcal{O}\left((tr(\mathbf{D}_v) + tr(\mathbf{D}_e))c^2 + |\mathcal{V}|c\right)$, where $tr(\mathbf{D}_v)$ is the sum of the degrees of all nodes and $tr(\mathbf{D}_e)$ is the sum of the number of nodes contained in all hyperedges. The detailed process of HAMP-I and HAMP-II are shown in Algorithm 1 and Algorithm 2.

Algorithm 1 The HAMP-I Algorithm for Hypergraph Node Classification.

Input: the incidence matrix $\mathbf{H}$, the node feature $\mathbf{X}$, and the node labels $\mathbf{Y}$.
Output: the model prediction accuracy .
Initialization: the time T, the step size τ and all parameters of model .
while not converged **do**
  Node feature mapping $\mathbf{X} = \text{Linear}_{\text{map}}(\mathbf{X})$;
  Set the initial time $t = 0$, the initial node representation $\mathbf{X}(0) = \mathbf{X}$;
  while $t \leq T$ **do**
    Message passing from $\mathcal{V}$ to $\mathcal{E}$: $\mathbf{X}_{\mathcal{V}\to\mathcal{E}}(t) = \Phi_1(\mathbf{X}(t), \mathbf{H})$;
    Message passing from $\mathcal{E}$ to $\mathcal{V}$: $\mathbf{X}_{\mathcal{E}\to\mathcal{V}}(t) = \Psi\left(\mathbf{X}(t), \Phi_2\left(\mathbf{X}_{\mathcal{V}\to\mathcal{E}}(t), \mathbf{H}\right)\right)$;
    Compute particle dynamics as
$$\mathbf{X}(t+\tau) = \mathbf{X}(t) + \tau\sigma\left(\underbrace{\mathbf{X}_{\mathcal{E}\to\mathcal{V}}(t) - \omega\mathbf{X}(t)}_{\text{Interaction Force}} + \underbrace{\delta f_d(\mathbf{X}(t))}_{\text{Allen-Cahn Force}} + \underbrace{\epsilon\mathbf{B}(t)}_{\text{Noise}} + \beta\mathbf{X}(0)\right);$$
    Updata $t = t + \tau$;
  end while
  Input the node representation into the classifier $\mathbf{X}^{out} = \text{MLP}(\mathbf{X}(T))$;
  Compute the model prediction labels $\widehat{\mathbf{Y}} = \text{Softmax}(\mathbf{X}^{out})$ and compute the loss function ;
  Update all parameter by back propagation using the Adam optimizer ;
end while

Algorithm 2 The HAMP-II Algorithm for Hypergraph Node Classification.

Input: the incidence matrix $\mathbf{H}$, the node feature $\mathbf{X}$, and the node labels $\mathbf{Y}$.
Output: the model prediction accuracy .
Initialization: the time T, the step size τ and all parameters of model .
while not converged **do**
  Node feature mapping $\mathbf{X} = \text{Linear}_{\text{map}}(\mathbf{X})$;
  Set the initial time $t = 0$, the initial node representation $\mathbf{X}(0) = \mathbf{X}$;
  Set the initial velocity $\mathbf{V}(0) = \text{Linear}(\mathbf{X}(0)) - \mathbf{X}(0)$;
  while $t \leq T$ **do**
    Message passing from $\mathcal{V}$ to $\mathcal{E}$: $\mathbf{V}_{\mathcal{V}\to\mathcal{E}}(t) = \Phi_1(\mathbf{V}(t), \mathbf{H})$;
    Message passing from $\mathcal{E}$ to $\mathcal{V}$: $\mathbf{V}_{\mathcal{E}\to\mathcal{V}}(t) = \Psi\left(\mathbf{V}(t), \Phi_2\left(\mathbf{V}_{\mathcal{V}\to\mathcal{E}}(t), \mathbf{H}\right)\right)$;
    Compute the velocity of particle dynamics system as
$$\mathbf{V}(t+\tau) = \mathbf{V}(t) + \tau\sigma\left(\underbrace{\mathbf{V}_{\mathcal{E}\to\mathcal{V}}(t) - \omega\mathbf{V}(t)}_{\text{Interaction Force}} + \underbrace{\delta f_d(\mathbf{V}(t))}_{\text{Allen-Cahn Force}} + \underbrace{\epsilon\mathbf{B}(t)}_{\text{Noise}} + \beta\mathbf{V}(0)\right);$$
    Compute the representation by $\mathbf{X}(t+\tau) = \mathbf{X}(t) + \tau\mathbf{V}(t+\tau)$;
    Updata $t = t + \tau$;
  end while
  Input the node representation into the classifier $\mathbf{X}^{out} = \text{MLP}(\mathbf{X}(T))$;
  Compute the model prediction labels $\widehat{\mathbf{Y}} = \text{Softmax}(\mathbf{X}^{out})$ and compute the loss function ;
  Update all parameter by back propagation using the Adam optimizer ;
end while

Limitations Discussion. Our particle dynamics-based hypergraph message passing framework assumes a static hypergraph topology. While this assumption is valid for social and biological hypergraphs with slowly evolving interactions, it may not hold in highly dynamic scenarios like financial transaction networks, where hyperedge topologies change abruptly. Consequently, effectively modeling temporal hypergraphs with evolving structures remains an open challenge.

B Theoretical Results and Proof

Technically, we suppose there exists $\{f_\beta^e\}$ such that $\mathcal{I} = \{1, \cdots, N\}$ can be divided into two disjoint groups with N_1, N_2 particles respectively: $f_\beta(h_{i,j}^e) \geq 0$, for $\{i, j\} \in \mathcal{I}_1$ or $\mathcal{I}_2$ and $f_\beta(h_{i,j}^e) \leq 0$, otherwise. We designate

$$\{x_i^{(1)}\} := \{\mathbf{x}_i | i \in \mathcal{I}_1\}, \quad \{x_j^{(2)}\} := \{\mathbf{x}_j | j \in \mathcal{I}_2\}. \tag{11}$$

The model is channel-wise, hence we use x instead of $\mathbf{x}$ in the proof.

First, let's define the relevant notation,

The mean value:

$$\bar{x} := \frac{1}{N} \sum_{i=1}^{N} x_i. \tag{12}$$

The deviation values:

$$\hat{x}_i := x_i - \bar{x}. \tag{13}$$

The variance of values within each group:

$$\text{var}(x^{(1)}) = \frac{1}{N_1} \sum (\hat{x}_i^{(1)})^2, \quad \text{var}(x^{(2)}) = \frac{1}{N_2} \sum (\hat{x}_i^{(2)})^2. \tag{14}$$

The second moments:

$$M_2(x^{(1)}) := \sum_{i=1}^{N_1} (x_i^{(1)})^2, \quad M_2(x^{(2)}) := \sum_{i=1}^{N_2} (x_i^{(2)})^2. \tag{15}$$

And others:

$$\widehat{M}_2 := M_2(\hat{x}^{(1)}) + M_2(\hat{x}^{(2)}) = \text{var}(x^{(1)}) + \text{var}(x^{(2)}). \tag{16}$$

$$\psi_i^{e,\pm} := \sum_{j \in e} h_{i,j}^{e,\pm}, \quad \psi_i^{\pm} := \sum_{e \in \mathcal{E}} \sum_{j \in e} h_{i,j}^{e,\pm}. \tag{17}$$

$$k := \max_i \{|\mathcal{E}(i)|\}. \tag{18}$$

$$D_e^- := \max_k \{\psi_k^{e,-}\}, \; D^- := \max_e \{D_e^-\}. \tag{19}$$

$$D_2^- := \max_e \{\|\psi^{e,-}\|_2\}. \tag{20}$$

Technically, we set $N_1 = N_2 := N_0$. This assumption is that N_1 is comparable to N_2, i.e., there exists a positive constant κ satisfying $\frac{1}{\kappa} N_1 \leq N_2 \leq \kappa N_1$.

We can rewrite Eq. 4 as

$$\begin{cases} \dot{u}_i^{(1)} = \dfrac{1}{N_1} \displaystyle\sum_{e \in \mathcal{E}(i)} \sum_{i'=1}^{N_1} h_{i,i'}^{e,+} (x_{i'}^{(1)} - x_i^{(1)}) - \dfrac{1}{N_2} \sum_{e \in \mathcal{E}(i)} \sum_{j=1}^{N_2} h_{i,j}^{e,-} (x_j^{(2)} - x_i^{(1)}) + \delta x_i^{(1)} (1 - (x_i^{(1)})^2) \\ \dot{v}_j^{(2)} = \dfrac{1}{N_2} \displaystyle\sum_{e \in \mathcal{E}(j)} \sum_{j'=1}^{N_2} h_{j,j'}^{e,+} (x_{j'}^{(2)} - x_j^{(2)}) - \dfrac{1}{N_1} \sum_{e \in \mathcal{E}(j)} \sum_{i=1}^{N_1} h_{i,j}^{e,-} (x_i^{(1)} - x_j^{(2)}) + \delta x_j^{(2)} (1 - (x_j^{(2)})^2). \end{cases} \tag{21}$$

Set the matrix A^e as

$$A^e_{i,j} = \begin{cases} h^e_{i,i} \\ -h^e_{i,j}, \end{cases} \tag{22}$$

and and designate $C^A := \min_{e\in\mathcal{E}}\{F(A^e)\}$, where $F(A^e)$ is the *Fiedler number* of A^e.

Lemma B.1 (L_2 estimate for M_2)**.** *There exists a positive constant M_2^∞ such that*

$$\sup_{0\le t<\infty} M_2(t) \le M_2^\infty \le \infty. \tag{23}$$

Proof. Note that $h^{e,\pm}_{i,j} = h^{e,\pm}_{j,i}$, then

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}M_2(x^{(1)}) &= \frac{2}{N_1}\sum_{i=1}^{N_1} x_i^{(1)}\dot{x}_i^{(1)} \\
&= \frac{2}{N_1}\sum_{i=1}^{N_1}\sum_{e\in\mathcal{E}(i)}\sum_{i'\in e} h^{e,+}_{i,i'}(x^{(1)}_{i'} - x^{(1)}_i)x^{(1)}_i - \frac{2}{N_1}\sum_{i=1}^{N_1}\sum_{e\in\mathcal{E}(i)}\sum_{j\in e} h^{e,-}_{i,j}(x^{(2)}_j - x^{(1)}_i)x^{(1)}_i \\
&\quad + \frac{2\delta}{N_1}\sum_{i=1}^{N_1}(x_i^{(1)})^2(1-(x_i^{(1)})^2) \\
&= \frac{2}{N_1}\sum_{e\in\mathcal{E}}\sum_{i,i'\in e} h^{e,+}_{i,i'}(x^{(1)}_{i'} - x^{(1)}_i)x^{(1)}_i - \frac{2}{N_1}\sum_{e\in\mathcal{E}}\sum_{i,j\in e} h^{e,-}_{i,j}(x^{(2)}_j - x^{(1)}_i)x^{(1)}_i \\
&\quad + \frac{2\delta}{N_1}\sum_{i=1}^{N_1}(x_i^{(1)})^2(1-(x_i^{(1)})^2) \\
&= -\frac{1}{N_1}\sum_{e\in\mathcal{E}}\sum_{i,i'\in e} h^{e,+}_{i,i'}(x^{(1)}_{i'} - x^{(1)}_i)^2 - \frac{2}{N_1}\sum_{e\in\mathcal{E}}\sum_{i,j\in e} h^{e,-}_{i,j}(x^{(2)}_j - x^{(1)}_i)x^{(1)}_i \\
&\quad + \frac{2\delta}{N_1}\sum_{i=1}^{N_1}(x_i^{(1)})^2(1-(x_i^{(1)})^2).
\end{aligned} \tag{24}$$

Similarly,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}M_2(x^{(2)}) &= \frac{2}{N_2}\sum_{j=1}^{N_2} x_j^{(2)}\dot{x}_j^{(2)} \\
&= -\frac{1}{N_2}\sum_{e\in\mathcal{E}}\sum_{j,j'\in e} h^{e,+}_{j,j'}(x^{(2)}_{j'} - x^{(2)}_j)^2 - \frac{2}{N_2}\sum_{e\in\mathcal{E}}\sum_{j,i\in e} h^{e,-}_{j,i}(x^{(1)}_i - x^{(2)}_j)x^{(2)}_j \\
&\quad + \frac{2\delta}{N_2}\sum_{j=1}^{N_2}(x_j^{(2)})^2(1-(x_j^{(2)})^2).
\end{aligned} \tag{25}$$

Define the total second moment $M_2 := M_2(x^{(1)}) + M_2(x^{(2)})$. Its time derivative is:

$$\frac{\mathrm{d}}{\mathrm{d}t}M_2 = \frac{\mathrm{d}}{\mathrm{d}t}M_2(x^{(1)}) + \frac{\mathrm{d}}{\mathrm{d}t}M_2(x^{(2)}). \tag{26}$$

By discarding the non-positive squared terms (first two sums), we obtain the inequality:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}M_2 \le &- \frac{2}{N_1}\sum_{e\in\mathcal{E}}\sum_{i,j\in e} h^{e,-}_{i,j}(x^{(2)}_j - x^{(1)}_i)x^{(1)}_i - \frac{2}{N_2}\sum_{e\in\mathcal{E}}\sum_{j,i\in e} h^{e,-}_{j,i}(x^{(1)}_i - x^{(2)}_j)x^{(2)}_j \\
&+ \frac{2\delta}{N_1}\sum_{i=1}^{N_1}(x_i^{(1)})^2(1-(x_i^{(1)})^2) + \frac{2\delta}{N_2}\sum_{i=1}^{N_2}(x_i^{(2)})^2(1-(x_i^{(2)})^2).
\end{aligned} \tag{27}$$

By the Cauchy-Schwarz inequality,

$$
\begin{aligned}
&(M_2^{(1)})^2 = \left(\sum_{i=1}^{N_1}(x_i^{(1)})^2\right)^2 \leq N_1 \sum_{i=1}^{N_1}(x_i^{(1)})^4,\\
&(M_2^{(2)})^2 = \left(\sum_{i=1}^{N_2}(x_i^{(2)})^2\right)^2 \leq N_2 \sum_{i=1}^{N_2}(x_i^{(2)})^4,\\
&(x_i^{(1)} - x_j^{(2)})^2 \leq 2((x_i^{(1)})^2 + (x_j^{(2)})^2).
\end{aligned}
$$

Then we have

$$
\begin{aligned}
&\frac{2\delta}{N_1}\sum_{i=1}^{N_1}(x_i^{(1)})^2\left(1-(x_i^{(1)})^2\right) + \frac{2\delta}{N_2}\sum_{i=1}^{N_2}(x_i^{(2)})^2\left(1-(x_i^{(2)})^2\right)\\
=&\frac{2\delta}{N_1}\sum_{i=1}^{N_1}(x_i^{(1)})^2 - \frac{2\delta}{N_1}\sum_{i=1}^{N_1}(x_i^{(1)})^4 + \frac{2\delta}{N_2}\sum_{i=1}^{N_2}(x_i^{(2)})^2 - \frac{2\delta}{N_2}\sum_{i=1}^{N_2}(x_i^{(2)})^4\\
\leq&\frac{2\delta}{N_1}\sum_{i=1}^{N_1}(x_i^{(1)})^2 - \frac{2\delta}{(N_1)^2}\left(\sum_{i=1}^{N_1}(x_i^{(1)})^2\right)^2 + \frac{2\delta}{N_2}\sum_{i=1}^{N_2}(x_i^{(2)})^2 - \frac{2\delta}{(N_2)^2}\left(\sum_{i=1}^{N_2}(x_i^{(2)})^2\right)^2\\
=&2\delta\left(\frac{M_2^{(1)}}{N_1}+\frac{M_2^{(2)}}{N_2}\right) - 2\delta\left(\left(\frac{M_2^{(1)}}{N_1}\right)^2+\left(\frac{M_2^{(2)}}{N_2}\right)^2\right)\\
\leq&2\delta\left(\frac{M_2^{(1)}}{N_1}+\frac{M_2^{(2)}}{N_2}\right) - \delta\left(\frac{M_2^{(1)}}{N_1}+\frac{M_2^{(2)}}{N_2}\right)^2\\
\leq&\frac{2\delta}{N'}M_2 - \frac{\delta}{N''}(M_2)^2,
\end{aligned}
\tag{28}
$$

where $N' = \min\{N_1, N_2\}$ and $N'' = \max\{N_1, N_2\}$.

So, we have

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}M_2 \leq&\frac{D^-}{N_1}\sum_{e\in\mathcal{E}}\sum_{i,j\in e}\left((x_j^{(2)}-x_i^{(1)})^2+(x_i^{(1)})^2\right) + \frac{D^-}{N_2}\sum_{e\in\mathcal{E}}\sum_{j,i\in e}\left((x_i^{(1)}-x_j^{(2)})^2+(x_j^{(2)})^2\right)\\
&+\frac{2\delta}{N_1}\sum_{i=1}^{N_1}(x_i^{(1)})^2\left(1-(x_i^{(1)})^2\right) + \frac{2\delta}{N_2}\sum_{i=1}^{N_2}(x_i^{(2)})^2\left(1-(x_i^{(2)})^2\right)\\
\leq&\frac{D^-}{N_1}\sum_{e\in\mathcal{E}}\sum_{i,j\in e}\left(3(x_i^{(1)})^2+2(x_j^{(2)})^2\right) + \frac{D^-}{N_2}\sum_{e\in\mathcal{E}}\sum_{i,j\in e}\left(2(x_i^{(1)})^2+3(x_j^{(2)})^2\right) + \frac{2\delta}{N'}M_2 - \frac{\delta}{N''}M_2^2.
\end{aligned}
\tag{29}
$$

These relations yield a Riccati-type differential inequality:

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}M_2 \leq&5D^-\max\{\frac{1}{N_1},\frac{1}{N_2}\}\sum_{e\in\mathcal{E}}\sum_{i,j\in e}\left((x_i^{(1)})^2+(x_j^{(2)})^2\right) + 2\delta M_2 - \delta(M_2)^2\\
\leq&\frac{5D^-k}{N'}M_2 + \frac{2\delta}{N'}M_2 - \frac{\delta}{N''}(M_2)^2\\
\leq&\frac{5D^-k+2\delta}{N'}M_2 - \frac{\delta}{N''}(M_2)^2.
\end{aligned}
\tag{30}
$$

Let y be a solution of the following ODE:

$$
y' = ay - by^2. \tag{31}
$$

Then, by phase line analysis, the solution $y(t)$ to Eq. 31 satisfies

$$
M_2(t) \leq y(t) \leq \max\left\{\frac{a}{b}, M_2(0)\right\} = \max\left\{\frac{(5D^-k+2\delta)N''}{\delta N'}, M_2(0)\right\} =: M_2^\infty. \tag{32}
$$

which yields the desired estimate. ■

Proposition B.2. *For Eq. 4, the distance of the centers of the two clusters is finite.*

Proof of Propostion B.2. By Lemma B.1

$$
\begin{aligned}
\|\bar{x}^{(1)} - \bar{x}^{(2)}\| &= \Big\| \frac{1}{N_1}\sum_{i=1}^{N_1} x_i^{(1)} - \frac{1}{N_2}\sum_{j=1}^{N_2} x_j^{(2)} \Big\| \\
&\le \sqrt{2}\sqrt{\frac{1}{N_1}\sum_{i=1}^{N_1}(x_i^{(1)})^2 + \frac{1}{N_2}\sum_{j=1}^{N_2}(x_j^{(2)})^2} \\
&\le \sqrt{2}\sqrt{M_2^{\infty}},
\end{aligned} \tag{33}
$$

where the first inequality used the Cauchy-Schwarz inequality. ■

Lemma B.3. *Let u, v be the solution to Eq. 21. Then $\|\bar{x}^{(1)} - \bar{x}^{(2)}\|^2$ satisfies*

$$
\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\bar{x}^{(1)} - \bar{x}^{(2)}\|^2 \ge \left(\frac{2c_m}{N_0} - c_1\right)\|\bar{x}^{(1)} - \bar{x}^{(2)}\|^2 - \frac{4(D_2^-)^2}{c_1 N_0} k\widehat{M}_2. \tag{34}
$$

Proof. The time evolution of $\bar{x}^{(1)}$ is given by

$$
\begin{aligned}
\dot{\bar{x}}^{(1)} &= \frac{1}{N_1}\sum_{i=1}^{N_1}\sum_{i' \in \mathcal{E}(i)} h_{i,i'}^{e,+}(x_{i'}^{(1)} - x_i^{(1)}) - \frac{1}{N_1}\sum_{i=1}^{N_1}\sum_{j \in \mathcal{E}(i)} h_{j,i}^{e,-}(x_j^{(2)} - x_i^{(1)}) \\
&= -\frac{1}{N_1}\sum_{e \in \mathcal{E}}\sum_{i,j \in e} h_{i,j}^{e,-}(x_j^{(2)} - x_i^{(1)}) \\
&= -\frac{1}{N_1}\sum_{e \in \mathcal{E}} \psi_j^{e,-} x_j^{(2)} + \frac{1}{N_1}\sum_{e \in \mathcal{E}} \psi_i^{e,-} x_i^{(1)} \\
&= \sum_{e \in \mathcal{E}}\left(-\frac{1}{N_1}\psi_j^{e,-} x_j^{(2)} + \frac{1}{N_1}\psi_i^{e,-} x_i^{(1)}\right),
\end{aligned} \tag{35}
$$

where the first equality uses the relation $\sum_{i=1}^{N_1} \hat{x}_i^{(1)} = 0$.

Then we have

$$
\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\bar{x}^{(1)} - \bar{x}^{(2)}\|^2 && (36)\\
=&(\bar{x}^{(1)} - \bar{x}^{(2)})(\dot{\bar{x}}^{(1)} - \dot{\bar{x}}^{(2)}) && (37)\\
=&(\bar{x}^{(1)} - \bar{x}^{(2)})\left[\sum_{e \in \mathcal{E}}\left(-\frac{1}{N_1}\psi_j^{e,-}x_j^{(2)} + \frac{1}{N_1}\psi_i^{e,-}x_i^{(1)}\right) - \sum_{e \in \mathcal{E}}\left(-\frac{1}{N_2}\psi_i^{e,-}x_i^{(1)} + \frac{1}{N_2}\psi_j^{e,-}x_j^{(2)}\right)\right] && (38)\\
=&(\bar{x}^{(1)} - \bar{x}^{(2)})\sum_{e \in \mathcal{E}}\left(-\frac{1}{N_1}\psi_j^{e,-}x_j^{(2)} - \frac{1}{N_2}\psi_j^{e,-}x_j^{(2)} + \frac{1}{N_1}\psi_i^{e,-}x_i^{(1)} + \frac{1}{N_2}\psi_i^{e,-}x_i^{(1)}\right) && (39)\\
=&(\bar{x}^{(1)} - \bar{x}^{(2)})\sum_{e \in \mathcal{E}}\Big(- \frac{1}{N_1}\psi_j^{e,-}(\bar{x}^{(2)} + \hat{x}_j^{(2)}) - \frac{1}{N_2}\psi_j^{e,-}(\bar{x}^{(2)} + \hat{x}_j^{(2)}) + \frac{1}{N_1}\psi_i^{e,-}(\bar{x}^{(1)} + \hat{x}_i^{(1)}) && (40)\\
&\quad + \frac{1}{N_2}\psi_i^{e,-}(\bar{x}^{(1)} + \hat{x}_j^{(1)})\Big) && (41)\\
=&(\bar{x}^{(1)} - \bar{x}^{(2)})\left\{\left[-\left(\frac{1}{N_1} + \frac{1}{N_2}\right)\sum_{e \in \mathcal{E}}\sum_{j \in e}\psi_j^{e,-}\right]\bar{x}^{(2)} + \left[\left(\frac{1}{N_1} + \frac{1}{N_2}\right)\sum_{e \in \mathcal{E}}\sum_{i \in e}\psi_i^{e,-}\right]\bar{x}^{(1)}\right. && (42)
\end{aligned}
$$

$$\left[-\left(\frac{1}{N_1}+\frac{1}{N_2}\right)\sum_{e\in\mathcal{E}}\sum_{j\in e}\psi_j^{e,-}\right]\hat{x}_j^{(2)}+\left[\left(\frac{1}{N_1}+\frac{1}{N_2}\right)\sum_{e\in\mathcal{E}}\sum_{i\in e}\psi_i^{e,-}\right]\hat{x}_i^{(1)}\Bigg\} \tag{43}$$

$$=\frac{2}{N_0}(\bar{x}^{(1)}-\bar{x}^{(2)})\left[\sum_{e\in\mathcal{E}}\sum_{i\in e}\psi_i^{e,-}\bar{x}^{(1)}-\sum_{e\in\mathcal{E}}\sum_{j\in e}\psi_j^{e,-}\bar{x}^{(2)}\right] \tag{44}$$

$$+\frac{2}{N_0}(\bar{x}^{(1)}-\bar{x}^{(2)})\left[\sum_{e\in\mathcal{E}}\sum_{i\in e}\psi_i^{e,-}\hat{x}_i^{(1)}-\sum_{e\in\mathcal{E}}\sum_{j\in e}\psi_j^{e,-}\hat{x}_j^{(2)}\right]. \tag{45}$$

We denote

$$\mathrm{Pes}(\hat{x}^{(1)},\hat{x}^{(2)}):=\sum_{e\in\mathcal{E}}\sum_{i\in e}\psi_i^{e,-}\hat{x}_i^{(1)}-\sum_{e\in\mathcal{E}}\sum_{j\in e}\psi_j^{e,-}\hat{x}_j^{(2)}.$$

and

$$\mathrm{Pes}(\bar{x}^{(1)},\bar{x}^{(2)}):=\sum_{e\in\mathcal{E}}\sum_{i\in e}\psi_i^{e,-}\bar{x}^{(1)}-\sum_{e\in\mathcal{E}}\sum_{j\in e}\psi_j^{e,-}\bar{x}^{(2)}.$$

Assume there exist constants c_m, c_v, such that

$$\mathrm{Pes}(\bar{x}^{(1)},\bar{x}^{(2)})\geq c_m(\bar{x}^{(1)}-\bar{x}^{(2)}). \tag{46}$$

Then, by Cauchy's inequality, for any c_1, we have

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\bar{x}^{(1)}-\bar{x}^{(2)}\|^2\\
=&\frac{2}{N_0}\mathrm{Pes}(\bar{x}^{(1)},\bar{x}^{(2)})(\bar{x}^{(1)}-\bar{x}^{(2)})+\frac{2}{N_0}\mathrm{Pes}(\hat{x}^{(1)},\hat{x}^{(2)})(\bar{x}^{(1)}-\bar{x}^{(2)})\\
\geq&\left(\frac{2c_m}{N_0}-c_1\right)\|\bar{x}^{(1)}-\bar{x}^{(2)}\|^2+c_1\|\bar{x}^{(1)}-\bar{x}^{(2)}\|^2+\frac{2}{N_0}\mathrm{Pes}(\hat{x}^{(1)},\hat{x}^{(2)})(\bar{x}^{(1)}-\bar{x}^{(2)})\\
\geq&\left(\frac{2c_m}{N_0}-c_1\right)\|\bar{x}^{(1)}-\bar{x}^{(2)}\|^2-\frac{1}{c_1}\frac{2}{N_0}\left(\mathrm{Pes}(\hat{x}^{(1)},\hat{x}^{(2)})\right)^2\\
\geq&\left(\frac{2c_m}{N_0}-c_1\right)\|\bar{x}^{(1)}-\bar{x}^{(2)}\|^2-\frac{4}{c_1N_0}\sum_{e\in\mathcal{E}}\left[\left(\sum_{i\in e}\psi_i^{e,-}\hat{x}_i^{(1)}\right)^2+\left(\sum_{j\in e}\psi_j^{e,-}\hat{x}_j^{(2)}\right)^2\right]\\
\geq&\left(\frac{2c_m}{N_0}-c_1\right)\|\bar{x}^{(1)}-\bar{x}^{(2)}\|^2-\frac{4}{c_1N_0}\sum_{e\in\mathcal{E}}\left[\|\psi^{e,-}\|^2\sum_{i\in e}(\hat{x}_i^{(1)})^2+\|\psi^{e,-}\|^2\sum_{j\in e}(\hat{x}_j^{(2)})^2\right]\\
\geq&\left(\frac{2c_m}{N_0}-c_1\right)\|\bar{x}^{(1)}-\bar{x}^{(2)}\|^2-\frac{4(D_2^-)^2}{c_1N_0}\sum_{e\in\mathcal{E}}\left[\sum_{i\in e}(\hat{x}_i^{(1)})^2+\sum_{j\in e}(\hat{x}_j^{(2)})^2\right]\\
\geq&\left(\frac{2c_m}{N_0}-c_1\right)\|\bar{x}^{(1)}-\bar{x}^{(2)}\|^2-\frac{4(D_2^-)^2}{c_1N_0}k[\widehat{M}_2].
\end{aligned} \tag{47}$$

■

Lemma B.4. *Let u, v be the solution to Eq. 21. Then $\widehat{M}_2$ satisfies*

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\widehat{M}_2\leq C_2\widehat{M}_2+2c_2\|\bar{x}^{(1)}-\bar{x}^{(2)}\|^2, \tag{48}$$

where

$$C_2:=-k\left(C^A-\frac{D^-}{2}+\frac{(D^-)^2}{4c_2}-\frac{\delta}{k}\right), \tag{49}$$

and c_2 is an arbitrary positive constant.

Proof. Subtracting Eq. 35 from Eq. 21 gives $\dot{\hat{x}}_i^{(1)}$. Then we have

$$
\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{N_1}\sum_{i=1}^{N_1}\|\hat{x}_i^{(1)}\|^2\right)\\
=&\frac{1}{N_1}\sum_{i=1}^{N_1}\hat{x}_i^{(1)}\dot{\hat{x}}_i^{(1)}\\
=&\frac{1}{N_1}\sum_{i=1}^{N_1}\hat{x}_i^{(1)}\sum_{e\in\mathcal{E}(i)}\left[\sum_{i'\in e}h_{i,i'}^{e,+}(x_{i'}^{(1)}-x_i^{(1)})-\sum_{j\in e}h_{i,j}^{e,-}(x_j^{(2)}-x_i^{(1)})\right]\\
&\quad+\frac{\delta}{N_1}\sum_{i=1}^{N_1}\hat{x}_i^{(1)}x_i^{(1)}(1-(x_i^{(1)})^2)\\
=&\frac{1}{N_1}\sum_{i=1}^{N_1}\hat{x}_i^{(1)}\sum_{e\in\mathcal{E}(i)}\left[\sum_{i'\in e}h_{i,i'}^{e,+}(\hat{x}_{i'}^{(1)}-\hat{x}_i^{(1)})-\sum_{j\in e}h_{i,j}^{e,-}(x_j^{(2)}-x_i^{(1)})\right]\\
&\quad+\frac{\delta}{N_1}\sum_{i=1}^{N_1}\hat{x}_i^{(1)}x_i^{(1)}(1-(x_i^{(1)})^2)\\
=&\frac{1}{N_1}\sum_{i=1}^{N_1}\hat{x}_i^{(1)}\sum_{e\in\mathcal{E}(i)}\sum_{i'\in e}h_{i,i'}^{e,+}(\hat{x}_{i'}^{(1)}-\hat{x}_i^{(1)})-\frac{1}{N_1}\sum_{i=1}^{N_1}\hat{x}_i^{(1)}\sum_{e\in\mathcal{E}(i)}\sum_{j\in e}h_{i,j}^{e,-}(\hat{x}_j^{(2)}-\hat{x}_i^{(1)})\\
&-\frac{1}{N_1}\sum_{i=1}^{N_1}\hat{x}_i^{(1)}\sum_{e\in\mathcal{E}(i)}\sum_{j\in e}h_{i,j}^{e,-}(\bar{x}^{(2)}-\bar{x}^{(1)})+\frac{\delta}{N_1}\sum_{i=1}^{N_1}\hat{x}_i^{(1)}x_i^{(1)}(1-(x_i^{(1)})^2)\\
=:&\, I_1+I_2+I_3+I_4.
\end{aligned}
\tag{50}
$$

I_1 can be defined by

$$
\begin{aligned}
I_1=&\frac{1}{N_1}\sum_{i=1}^{N_1}\hat{x}_i^{(1)}\sum_{e\in\mathcal{E}(i)}\sum_{i'\in e}h_{i,i'}^{e,+}(\hat{x}_{i'}^{(1)}-\hat{x}_i^{(1)})\\
=&\frac{1}{N_1}\sum_{e\in\mathcal{E}}\sum_{i,i'\in e}h_{i,i'}^{e,+}(\hat{x}_{i'}^{(1)}-\hat{x}_i^{(1)})\hat{x}_i^{(1)}\\
=&\frac{1}{N_1}\sum_{e\in\mathcal{E}}((\hat{x}^{(1)})^e)^\top A^e(\hat{x}^{(1)})^e,
\end{aligned}
\tag{51}
$$

where $(\hat{x}^{(1)})^e:=(\hat{x}_{i_1}^{(1)},\cdots,\hat{x}_{i_{|e|}}^{(1)})^\top$ and A^e is given by Eq. 22 for each e. Thus I_1 is bounded by

$$
\frac{1}{N_1}\sum_{e\in\mathcal{E}}((\hat{x}^{(1)})^e)^\top A^e(\hat{x}^{(1)})^e\le-\frac{1}{N_1}\sum_{e\in\mathcal{E}}C^A\|(\hat{x}^{(1)})^e\|^2\le-\frac{c_r}{N_1}\sum_{i=1}^{N_1}C^A\|\hat{x}_i^{(1)}\|^2,
\tag{52}
$$

where c_r is a constant larger than 1 related to the repetition of $\{\hat{x}_i^{(1)}\}$ in all hyperedges.

I_2 can be controlled by

$$
I_2=-\frac{1}{N_1}\sum_{i=1}^{N_1}\hat{x}_i^{(1)}\sum_{e\in\mathcal{E}(i)}\sum_{j\in e}h_{i,j}^{e,-}(\hat{x}_j^{(2)}-\hat{x}_i^{(1)})
\tag{53}
$$

$$
=-\frac{1}{N_1}\sum_{e\in\mathcal{E}}\sum_{i,j\in e}h_{i,j}^{e,-}(\hat{x}_j^{(2)}-\hat{x}_i^{(1)})\hat{x}_i^{(1)}
\tag{54}
$$

$$
=-\frac{1}{N_1}\sum_{e\in\mathcal{E}}\sum_{i,j\in e}h_{i,j}^{e,-}\hat{x}_j^{(2)}\hat{x}_i^{(1)}+\frac{1}{N_1}\sum_{e\in\mathcal{E}}\sum_{i,j\in e}h_{i,j}^{e,-}\|\hat{x}_i^{(1)}\|^2
\tag{55}
$$

$$\leq\frac{1}{N_1}\sum_{e\in\mathcal{E}}\sum_{i,j\in e}h^{e,-}_{i,j}\frac{1}{2}(\|\hat{x}^{(2)}_j\|^2+\|\hat{x}^{(1)}_i\|^2)+\frac{1}{N_1}\sum_{e\in\mathcal{E}}D^-_e\sum_{i\in e}\|\hat{x}^{(1)}_i\|^2 \tag{56}$$

$$\leq\frac{1}{2N_1}\sum_{e\in\mathcal{E}}\sum_{j\in e}\psi^{e,-}_j\|\hat{x}^{(2)}_j\|^2+\frac{3D^-}{2N_1}\sum_{e\in\mathcal{E}}\sum_{i\in e}\|\hat{x}^{(1)}_j\|^2 \tag{57}$$

$$\leq\frac{D^-}{2N_1}\sum_{e\in\mathcal{E}}\sum_{j\in e}\|\hat{x}^{(2)}_j\|^2+\frac{3D^-}{2N_1}\sum_{e\in\mathcal{E}}\sum_{i\in e}\|\hat{x}^{(1)}_j\|^2 \tag{58}$$

$$\leq\frac{D^-k}{2N_1}\sum_{j=1}^{N_2}\|\hat{x}^{(2)}_j\|^2+\frac{3D^-k}{2N_1}\sum_{i=1}^{N_1}\|\hat{x}^{(1)}_i\|^2. \tag{59}$$

I_3 has the below estimate for any constant $c_2>0$:

$$\begin{aligned}
I_3=&-\frac{1}{N_1}\sum_{i=1}^{N_1}\hat{x}^{(1)}_i\sum_{e\in\mathcal{E}(i)}\sum_{j\in e}h^{e,-}_{i,j}(\bar{x}^{(2)}-\bar{x}^{(1)})\\
\leq& c_2\|\bar{x}^{(1)}-\bar{x}^{(2)}\|^2+\frac{1}{4c_2N_1}\sum_{e\in\mathcal{E}}\sum_{i,j\in e}\|h^{e,-}_{i,j}\|^2\|\hat{x}^{(1)}_i\|^2\\
\leq& c_2\|\bar{x}^{(1)}-\bar{x}^{(2)}\|^2+\frac{(D^-)^2}{4c_2N_1}\sum_{e\in\mathcal{E}}\sum_{i,j\in e}\|\hat{x}^{(1)}_i\|^2\\
\leq& c_2\|\bar{x}^{(1)}-\bar{x}^{(2)}\|^2+\frac{(D^-)^2k}{4c_2N_1}\sum_{e\in\mathcal{E}}\sum_{i=1}^{N_1}\|\hat{x}^{(1)}_i\|^2.
\end{aligned} \tag{60}$$

Define

$$\begin{aligned}
I_4:=&\frac{1}{N_1}\delta\sum_{i=1}^{N_1}\hat{x}^{(1)}_i x^{(1)}_i(1-\|x^{(1)}_i\|^2)\\
=&\frac{\delta}{N_1}\sum_{i=1}^{N_1}\hat{x}^{(1)}_i(\hat{x}^{(1)}_i+\bar{x}^{(1)})(1-(x^{(1)}_i)^2)\\
=&\frac{\delta}{N_1}\sum_{i=1}^{N_1}(\hat{x}^{(1)}_i)^2-\frac{\delta}{N_1}\sum_{i=1}^{N_1}(\hat{x}^{(1)}_i)^2(x^{(1)}_i)^2+\frac{\delta}{N_1}\sum_{i=1}^{N_1}\hat{x}^{(1)}_i\bar{x}^{(1)}-\frac{\delta}{N_1}\sum_{i=1}^{N_1}\hat{x}^{(1)}_i(x^{(1)}_i)^2\bar{x}^{(1)}\\
=&\frac{\delta}{N_1}\sum_{i=1}^{N_1}(\hat{x}^{(1)}_i)^2-\frac{\delta}{N_1}\sum_{i=1}^{N_1}\hat{x}^{(1)}_i\|x^{(1)}_i\|^2x^{(1)}_i.
\end{aligned} \tag{61}$$

Note that

$$\begin{aligned}
\sum_{i=1}^{N_1}\hat{x}^{(1)}_i\|x^{(1)}_i\|^2x^{(1)}_i=&\sum_{i=1}^{N_1}\|x^{(1)}_i\|^2(\|x^{(1)}_i\|^2-x^{(1)}_i\bar{x}^{(1)})\\
\geq&\frac{1}{2}\sum_{i=1}^{N_1}\|x^{(1)}_i\|^2(\|x^{(1)}_i\|^2-\|\bar{x}^{(1)}\|^2)\\
=&\frac{1}{2}\sum_{i=1}^{N_1}\|x^{(1)}_i\|^4-\frac{1}{2}\sum_{i=1}^{N_1}\|x^{(1)}_i\|^2\|\bar{x}^{(1)}\|^2\\
\geq&\frac{1}{2}\sum_{i=1}^{N_1}\|x^{(1)}_i\|^4-\frac{1}{2N_1}\sum_{i=1}^{N_1}(\|x^{(1)}_i\|^2)^2\\
\geq&0.
\end{aligned} \tag{62}$$

Hence,

$$I_4\leq\delta\widehat{M}_2(u). \tag{63}$$

Then

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2N_1}\sum_{i=1}^{N_1}\|\hat{x}_i^{(1)}\|^2\right)\\
&\leq k\left(-\frac{C^A}{N_1}+\frac{3D^-}{2N_1}+\frac{(D^-)^2}{4c_2N_1}\right)\sum_{i=1}^{N_1}\|\hat{x}_i^{(1)}\|^2+k\frac{D^-}{2N_2}\sum_{j=1}^{N_2}\|\hat{x}_j^{(2)}\|^2+c_2\|\bar{x}^{(1)}-\bar{x}^{(2)}\|^2 \qquad (64)\\
&\leq \left[k\left(-C^A+\frac{3D^-}{2}+\frac{(D^-)^2}{4c_2}\right)+\delta\right]\widehat{M}_2(u)+k\frac{D^-}{2}\widehat{M}_2(v)+c_2\|\bar{x}^{(1)}-\bar{x}^{(2)}\|^2.
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2N_2}\sum_{j=1}^{N_2}\|\hat{x}_j^{(2)}\|^2\right)\\
&\leq \left[k\left(-C^A+\frac{3D^-}{2}+\frac{(D^-)^2}{4c_2}\right)+\delta\right]\widehat{M}_2(v)+k\frac{D^-}{2}\widehat{M}_2(u)+c_2\|\bar{x}^{(1)}-\bar{x}^{(2)}\|^2.
\end{aligned} \tag{65}
$$

Summing them together gives

$$
\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}(\widehat{M}_2)\leq -k\left(C^A-\frac{D^-}{2}+\frac{(D^-)^2}{4c_2}-\frac{\delta}{k}\right)(\widehat{M}_2)+2c_2\|\bar{x}^{(1)}-\bar{x}^{(2)}\|^2. \tag{66}
$$

This gives an exponential growth estimate or $\widehat{M}_2$ up to an error term of $\|\bar{x}^{(1)}-\bar{x}^{(2)}\|^2$. ■

Proposition B.5 (L_2 separation of HAMP-I)**.** *For Eq. 4, suppose the above assumptions are satisfied. Define the mean value* $\bar{\mathbf{x}} := \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i$*, and the second moments* $M_2(\mathbf{x}) := \sum_{i=1}^{N}\mathbf{x}_i^2$*. Then for sufficiently large* N_1, N_2*, there exist constants* λ_-, λ_+*, such that if the initial data satisfies*

$$
\lambda(0) := \frac{\widehat{M}_2(0)}{\|\bar{\mathbf{x}}^{(1)}(0)-\bar{\mathbf{x}}^{(2)}(0)\|^2}\leq \lambda_+, \tag{67}
$$

then, there holds that the L_2 *separation*

$$
\lambda(t) := \frac{\widehat{M}_2(t)}{\|\bar{\mathbf{x}}^{(1)}(t)-\bar{\mathbf{x}}^{(2)}(t)\|^2}\leq \lambda_- + (\lambda(0)-\lambda_-)e^{-\mu t}, \tag{68}
$$

with a positive constant μ*, where* $\widehat{M}_2(t) := M_2(\mathbf{x}^{(1)}(t)) + M_2(\mathbf{x}^{(2)}(t))$.

Proof of Proposition 5.2. Lemma B.3 gives an exponential growth estimate of $\|\bar{x}^{(1)}-\bar{x}^{(2)}\|^2$ up to an error term of $\widehat{M}_2$. If N_0 is large enough, the coefficient of the error term is small.

Lemma B.4 gives an exponential growth estimate or $\widehat{M}_2$ up to an error term of $\|\bar{x}^{(1)}-\bar{x}^{(2)}\|^2$. If N_0 is large enough, the coefficient of the error term is small.

Set

$$
A_{11} := \frac{2c_m}{N_0}-c_1,\ A_{12} := \frac{4(D_2^-)^2k}{c_1N_0},\ A_{21}=2c_2,\ A_{22}=k\left(C^A-\frac{D^-}{2}+\frac{(D^-)^2}{4c_2}\right). \tag{69}
$$

If N_0 is large enough, Eq. 69 is satisfied.

$$
(A_{11}+A_{12})^2-4A_{21}A_{12}>0. \tag{70}
$$

Apply Lemma 4.1 in [27] then we obtain the L_2 separation. ■

Remark B.6. We can alternatively prove Theorem 5.2 by the method of [[44] Proposition 2 and 3].

Proposition B.7 (Lower bound of the Dirichlet energy)**.** *If the hypergraph* $\mathcal{H}$ *is a connected one, for Eq. 4 with the conditions of Theorem 5.2, or for Eq. 5 with conditions of Theorem 5.1 in [19], there exists a positive lower bound of the Dirichlet energy.*

Proof of Proposition 5.4. The relative size between $\|\bar{x}^{(1)} - \bar{x}^{(2)}\|^2$ and $\widehat{M_2}$ is an indicator of group separation in the sense of L_2: if $\|\bar{x}^{(1)} - \bar{x}^{(2)}\|^2$ is much larger than $\widehat{M_2}$, then the two groups are well-separated in average sense. Since the hypergraph is connected, there is a positive bound between different clusters, hence the Dirichlet energy does not decay to zero.

Remark B.8. The proof for the second-order system Eq. 5 can be proved in a similar way.

Remark B.9. The separability of Eq. 6 which is the Eq. 4 or Eq. 5 with a stochastic term also holds.

C Experiment Details

C.1 Dataset Details

We selected the most common hypergraph benchmark datasets, and the statistics of nine datasets across different domains are summarized in Tab. 3. The key statistics include the number of nodes, hyperedges, features, classes, average node degree d_v, average hyperedge size $|\mathcal{E}|$, and CE homophily [36]. These variations highlight the diversity in dataset scales and structural patterns, which may influence model performance in hypergraph-related tasks.

Table 3: The summary of data statistics.

Dataset	# nodes	# hyperedges	# features	# classes	avg. d_v	avg. $\|\mathcal{E}\|$	CE homophily
Cora	2708	1579	1433	7	1.767	3.03	0.897
Citeseer	3312	1079	3703	6	1.044	3.200	0.893
Pubmed	19717	7963	500	3	1.756	4.349	0.952
Cora-CA	2708	1072	1433	7	1.693	4.277	0.803
DBLP-CA	41302	22363	1425	6	2.411	4.452	0.869
Congress	1718	83105	100	2	427.237	8.656	0.555
House	1290	340	100	2	9.181	34.730	0.509
Senate	282	315	100	2	19.177	17.168	0.498
Walmart	88860	69906	100	11	5.184	6.589	0.530

C.2 Additional Ablation Studies

Impact of Hidden Dimension on HAMP-I and HAMP-II. We explore the tolerance of our model to different hidden dimensions, as shown in Tab. 4. For simplicity, we only vary the size of hidden dimension, while other parameters remain fixed. Overall, these results demonstrate the robustness of our methods with varying hidden dimensions. It is worth noting that high hidden dimension is key to achieving the best performance of HAMP.

Table 4: Impact of hidden dimension evaluated on the hypergraph datasets.

	HAMP-I				HAMP-II			
	128	256	512	1024	128	256	512	1024
Cora	80.19	80.55	**81.18**	79.85	79.59	79.54	**80.80**	79.60
Citeseer	73.39	72.89	**75.22**	72.40	73.95	74.36	**75.33**	75.19
Pubmed	88.49	88.85	**89.02**	88.82	88.48	88.48	**89.05**	88.78
Senate	63.24	65.35	**69.44**	66.62	63.52	63.10	**70.14**	62.96
House	71.21	70.62	72.66	**72.72**	68.30	69.47	**72.60**	71.21

Impact of Repulsion, Allen-Cahn Force, and Noise on HAMP-I and HAMP-II. We summary ablation studies to investigate the individual and combined effects of repulsion f_β^-, Allen-Cahn force f_d, and noise B_t on both HAMP-I and HAMP-II. Tab. 5 reports the average node classification accuracy with a standard deviation across seven standard hypergraph benchmarks over 10 runs. Models with the repulsion term enabled outperform their counterparts in some dataset, indicating enhanced ability to distinguish node representations in complex hypergraph structures. The synergy between the repulsion and Allen-Cahn terms further boosts performance, confirming that these particle system-inspired mechanisms play complementary roles. Overall, these improvement confirm the validity of the HAMP construction and further highlight the significant advantages of incorporating particle system theory into the hypergraph message passing learning process.

Table 5: Ablation studies on some standard hypergraph benchmarks. The accuracy (%) is reported with a standard deviation from 10 repetitive runs. (Key: f_β^-: repulsion; f_d: Allen-Cahn; B_t: noise.)

Homophilic	f_β^-	f_d	B_t	Cora	Citeseer	Pubmed	Cora-CA
HAMP-I	✖	✖	✖	76.09±1.22	70.53±1.56	87.98±0.38	83.13±1.26
	✔	✖	✖	76.40±1.56	70.85±1.65	88.25±0.50	83.15±1.36
	✖	✔	✖	80.31±1.41	74.83±1.70	88.90±0.45	84.77±1.16
	✖	✖	✔	75.67±1.71	70.59±1.40	87.93±0.52	82.70±1.01
	✔	✔	✖	80.49±1.26	74.96±1.56	88.87±0.40	85.21±1.49
	✖	✔	✔	80.59±1.25	74.67±1.69	88.77±0.44	84.59±1.03
	✔	✔	✔	**81.18±1.30**	**75.22±1.62**	**89.02±0.49**	**85.23±1.15**
HAMP-II	✖	✖	✖	77.42±1.44	71.50±1.49	88.68±0.62	83.60±1.45
	✔	✖	✖	77.08±1.73	72.20±1.14	88.59±0.50	82.95±1.35
	✖	✔	✖	80.13±1.26	74.37±1.59	88.86±0.55	84.37±1.45
	✖	✖	✔	77.18±1.61	71.75±1.68	88.68±0.59	82.91±1.28
	✔	✔	✖	79.70±1.36	73.99±1.75	88.82±0.48	84.30±1.32
	✖	✔	✔	79.50±1.25	74.25±1.28	88.80±0.40	83.31±1.44
	✔	✔	✔	**80.80±1.62**	**75.33±1.61**	**89.05±0.41**	**84.89±1.14**

Table 6: Node Classification on standard hypergraph benchmarks. The accuracy (%) is reported with a standard deviation from 10 repetitive runs. (Key: f_β^-: repulsion; f_d: Allen-Cahn; B_t: noise.)

Heterophilic	f_β^-	f_d	B_t	Congress	Senate	Walmart	House
HAMP-I	✖	✖	✖	93.51±1.08	60.70±8.38	69.64±0.35	70.50±1.45
	✔	✖	✖	93.70±1.02	60.00±8.66	69.80±0.45	70.56±1.96
	✖	✔	✖	94.79±1.14	66.76±5.44	69.77±0.28	71.55±1.53
	✖	✖	✔	93.47±1.13	60.14±7.54	69.86±0.35	69.88±2.48
	✔	✔	✖	94.58±1.25	67.75±8.82	69.76±0.37	71.58±1.87
	✖	✔	✔	94.67±1.02	65.63±2.98	69.73±0.49	71.55±2.58
	✔	✔	✔	**95.09±0.79**	**69.44±6.09**	**69.90±0.38**	**72.72±1.77**
HAMP-II	✖	✖	✖	94.79±0.73	58.73±7.03	69.84±0.25	69.85±1.61
	✔	✖	✖	94.02±1.10	61.55±5.98	69.91±0.30	69.35±1.87
	✖	✔	✖	94.19±1.07	62.82±6.44	69.89±0.31	70.96±2.06
	✖	✖	✔	94.35±1.14	60.14±5.10	69.84±0.37	70.46±2.08
	✔	✔	✖	94.58±0.86	61.97±8.42	69.92±0.33	71.36±1.68
	✖	✔	✔	94.12±0.63	64.51±4.19	69.86±0.34	70.96±2.76
	✔	✔	✔	**95.26±1.34**	**70.14±6.08**	**69.94±0.37**	**72.60±1.23**

C.3 Time-Memory Tradeoff Analysis

We intuitively reveal the differences of different methods with a single-layer network in the time-memory trade-off on Walmart dataset. As shown in Fig. 5, HAMP-I and HAMP-II methods demonstrate a notable trade-off between time efficiency and memory consumption. The experimental results reveal:

- Memory usage: HAMP-I and HAMP-II maintain memory consumption within 11000-12000 MiB, achieving a 20-26% reduction compared to ED-HNN, while being comparable to HDS^{ode}.
- Time efficiency: Although the time consumption for HAMP-I and HAMP-II runtime slightly exceeds ED-HNN (0.1s), it outperforms the baseline HDS^{ode}.

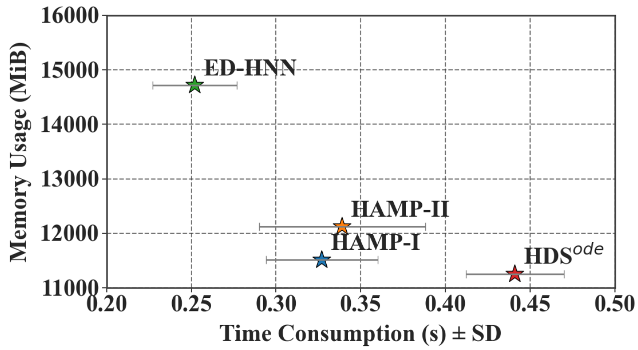

Figure 5: Time-Memory tradeoff analysis of different methods on Walmart dataset. SD denotes the standard deviation of time consumption.

C.4 Hyperparameters

To ensure fairness, we follow the same training recipe as ED-HNN. Specifically, we train the model for 500 epochs using the Adam optimizer with the learning rate of 0.001 and no weight decay during the training phases. And we apply early stopping with a patience of 50. For the stability, we run 10 trials with different seed and report the results of mean and the standard deviation. All experiments are implemented on an NVIDIA RTX 4090 GPU with Pytorch.

We explore the parameter space by grid search, where the search ranges for each critical hyperparameter are delineated below:

- Dropout rate in {0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9};
- Layer number of classifier in {1, 2, 3};
- The hidden dimension of classifier in {128, 256, 512};
- Hidden dimension of model in {128, 256, 512, 1024};
- step size of solver in {0.09, 0.1, 0.15, 0.2, 0.25};
- γ of repulsive force in {0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1, 0.11, 0.12, 0.13, 0.14, 0.15};
- Initial values of learnable parameters δ of damping term in {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15};
- Initial values of learnable parameters ϵ of noise term in {0, 0.1, 0.3};

Tab. 7 and Tab. 8 summarize the best hyperparameters on standard hypergraph benchmarks using HAMP-I and HAMP-II, respectively. For fairness, a linear layer is added to perform feature mapping when conducting HDS^{ode} experiments. The optimal hyperparameters for node classification on standard hypergraph benchmarks is achieved by the HDS^{ode} algorithm, as demonstrated in Tab. 9.

Table 7: The best hyperparameters of Node Classification on standard hypergraph benchmarks using the HAMP-I algorithm.

Dataset	model. hd	cls. hd and # layers	time	step size	δ	γ	dropout	ϵ
Cora	512	128, 1	1	0.1	12	0.05	0.4	0
Citeseer	512	512, 1	0.6	0.1	6	0.05	0.2	0
Pubmed	512	256, 1	0.2	0.1	15	0.1	0.5	0
Cora-CA	512	512, 2	0.4	0.2	4	0.05	0.9	0
DBLP-CA	256	128, 2	1.1	0.1	11	0.12	0.2	0
Congress	128	128, 2	1.4	0.1	1	0.08	0.3	0
House	1024	512, 3	1.05	0.15	3	0.05	0.8	0
Senate	512	256, 2	0.6	0.1	10	0.05	0.7	0
Walmart	256	128, 2	1.75	0.25	0	0.02	0.3	0

Table 8: The best hyperparameters of Node Classification on standard hypergraph benchmarks using the HAMP-II algorithm.

Dataset	model. hd	cls. hd and layers	time	step size	δ	γ	dropout	ϵ
Cora	512	512, 1	1.9	0.1	5	0.12	0.3	0.1
Citeseer	512	512, 1	1.8	0.15	8	0.13	0.6	0
Pubmed	512	256, 1	0.6	0.09	5	0.09	0.3	0
Cora-CA	512	128, 2	0.75	0.25	3	0.01	0.7	0
DBLP-CA	256	128, 2	3.45	0.15	7	0.09	0.3	0
Congress	128	128, 2	6.25	0.25	0	0.01	0.3	0
House	512	256, 2	1.6	0.1	8	0.14	0.8	0
Senate	512	256, 2	3	0.2	13	0.05	0.3	0.3
Walmart	256	128, 2	2.5	0.25	0	0.02	0.3	0

Table 9: The best hyperparameters of Node Classification on standard hypergraph benchmarks using the HDS^{ode} algorithm.

Dataset	model. hd	cls. hd and layers	# layer model.	alpha$_v$	alpha$_e$	step
Cora	512	256, 2	10	0.05	0.9	20
Citeseer	512	512, 1	5	0.05	0.9	20
Pubmed	512	256, 1	12	0.05	0.9	20
Cora-CA	512	512, 2	9	0.05	0.9	20
DBLP-CA	256	256, 2	15	0.05	0.9	20
Congress	256	128, 2	7	0.25	0.9	5
House	512	256, 2	10	0.05	0.9	20
Senate	512	256, 2	9	0.05	0.9	20
Walmart	256	128, 2	6	0.25	0.9	5