Dynamic Tensor Decomposition via Neural Diffusion-Reaction Processes

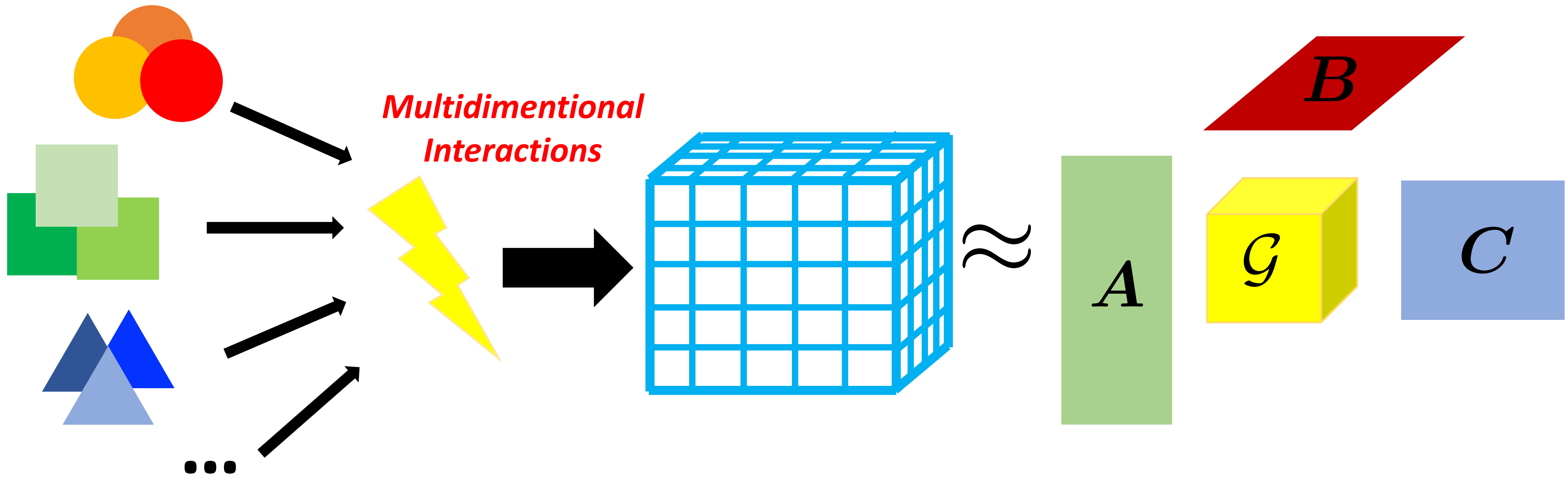
Zheng Wang, Shikai Fang, Shibo Li, Shandian Zhe

Kahlert School of Computing

University of Utah

KAHLERT SCHOOL OF COMPUTING
UNIVERSITY OF UTAH

Quick Overview

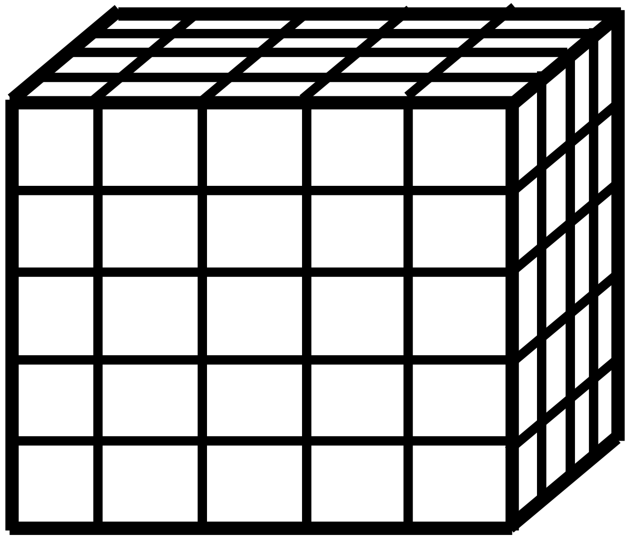


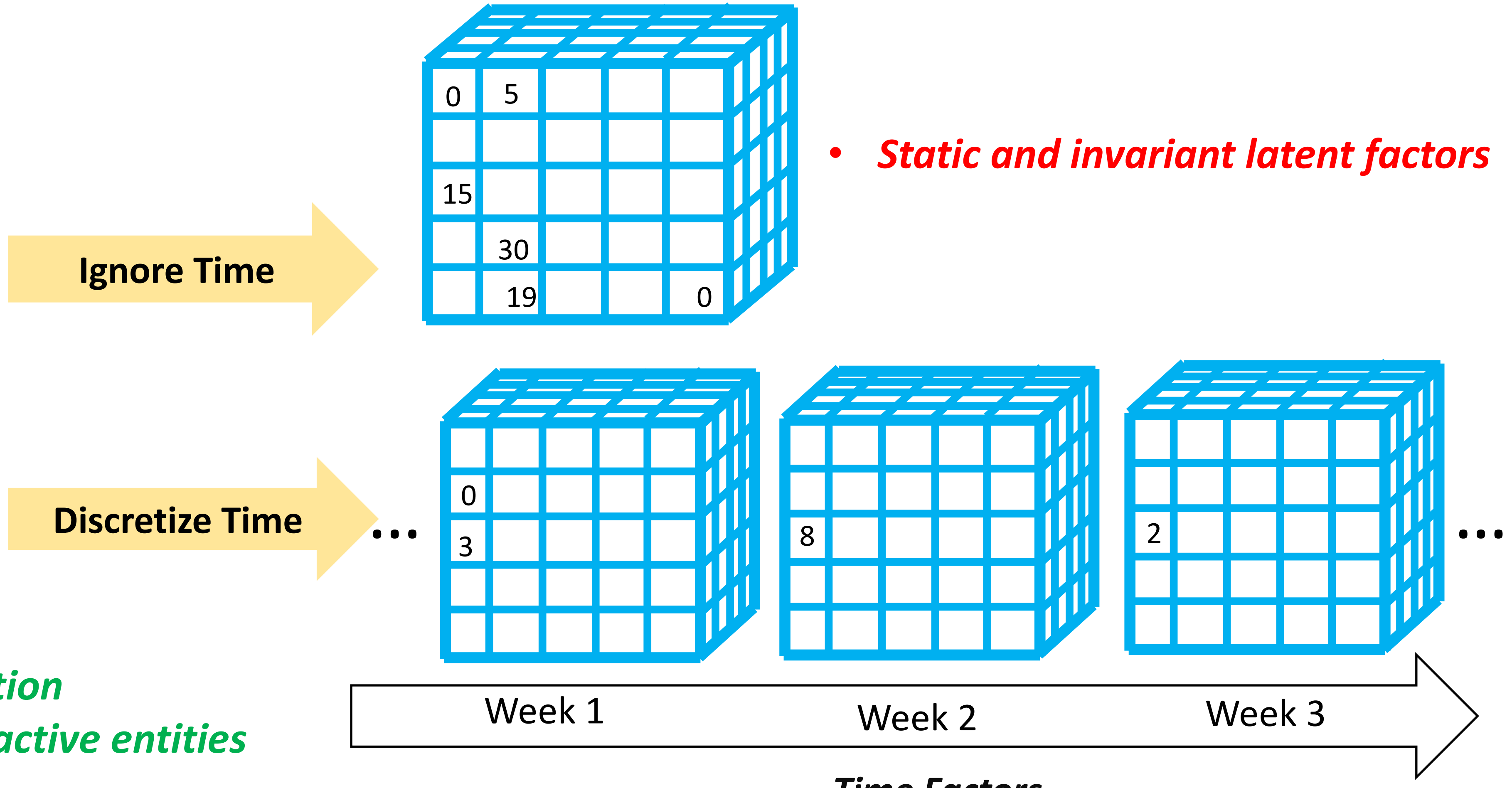
- ➢ Customer, Store, Product
- ➢ Location, Age, Gender
- ➢ Website, Location, Ads Type
- ➢ Latitude, Longitude, Elevation
- ➢ …

- ✓ Online Store Transactions
- ✓ Social Media User Behaviors
- ✓ Advertisement Click Log Data
- ✓ Global Climate Data
- ✓ …

- Tucker decomposition (Tucker, 1966)
- CANDECOMP/PARAFAC (CP) decomposition (Harshman, 1970)
- …

Underexplored Dynamic Tensor Decomposition





- *Static and invariant latent factors*

- *Valuable temporal information*
- *Evolving properties of interactive entities*
- *Structural Information*

Dynamic Tensor Decomposition via Neural Diffusion-Reaction Processes

: Embeddings of entities of each mode

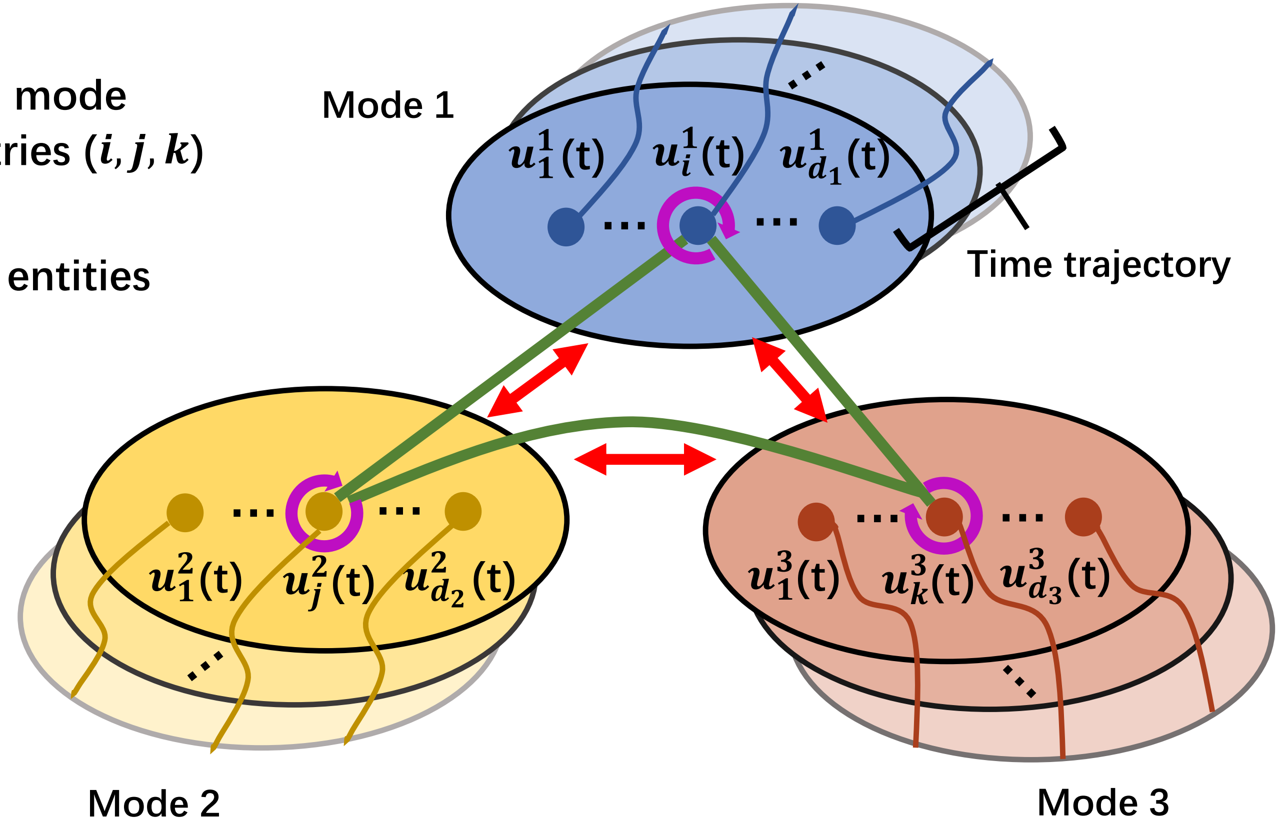
: Edges defined by observed entries (i, j, k)

: Diffusion process along edges

: Reaction process on individual entities

Table of observed data

Mode 1	Mode 2	Mode 3	Time-stamp	Value
i	j	k	t	$y_{ijk}(t)$
…	…	…	…	…



Dynamic Tensor Decomposition via Neural Diffusion-Reaction Processes

$$\frac{\partial \mathcal{U}(t)}{\partial t} = \boxed{(\mathcal{W} - \mathcal{A})\mathcal{U}(t)} + \mathcal{F}(\mathcal{U}, t), \quad \mathcal{U}(0) = \mathcal{U}_0$$

$$\mathcal{W} = \begin{pmatrix} \mathbf{0} & \mathbf{W}^{1,2} & \cdots & \mathbf{W}^{1,K} \\ \mathbf{W}^{2,1} & \mathbf{0} & \cdots & \vdots \\ \vdots & & \ddots & \mathbf{W}^{K-1,K} \\ \mathbf{W}^{K,1} & \cdots & \mathbf{W}^{K,K-1} & \mathbf{0} \end{pmatrix}$$

$$\mathcal{A} = \operatorname{diag}\left(\sum_{s\in\{1\ldots K\}\setminus 1} \mathbf{A}^{1,s}, \ldots, \sum_{s\in\{1\ldots K\}\setminus K} \mathbf{A}^{K,s}\right)$$

- **Diffusion Process on Multi-Partite Graphs**
  - Capture correlations between related entities via diffusion process

$$\frac{\mathrm{d}\mathbf{u}_j^k}{\mathrm{d}t} = \sum_{s\in\{1,\ldots,K\}\setminus k} \sum_{j'=1}^{d_s} [\mathbf{W}^{k,s}]_{j,j'} \left(\mathbf{u}_{j'}^s(t) - \mathbf{u}_j^k(t)\right) = \sum_{s\in\{1,\ldots,K\}\setminus k} \left(\mathbf{w}_j^{k,s}\mathbf{U}^s(t)\right)^\top - a_j^{k,s}\mathbf{u}_j^k,$$

Dynamic Tensor Decomposition via Neural Diffusion-Reaction Processes

$$\frac{\partial \mathcal{U}(t)}{\partial t} = (\mathcal{W} - \mathcal{A})\mathcal{U}(t) + \boxed{\mathcal{F}(\mathcal{U}, t),} \quad \mathcal{U}(0) = \mathcal{U}_0$$

$$\mathcal{W} = \begin{pmatrix} \mathbf{0} & \mathbf{W}^{1,2} & \dots & \mathbf{W}^{1,K} \\ \mathbf{W}^{2,1} & \mathbf{0} & \dots & \vdots \\ \vdots & & \ddots & \mathbf{W}^{K-1,K} \\ \mathbf{W}^{K,1} & \dots & \mathbf{W}^{K,K-1} & \mathbf{0} \end{pmatrix}$$

$$\mathcal{A} = \operatorname{diag}\left(\sum_{s\in\{1\dots K\}\setminus 1} \mathbf{A}^{1,s}, \dots, \sum_{s\in\{1\dots K\}\setminus K} \mathbf{A}^{K,s}\right)$$

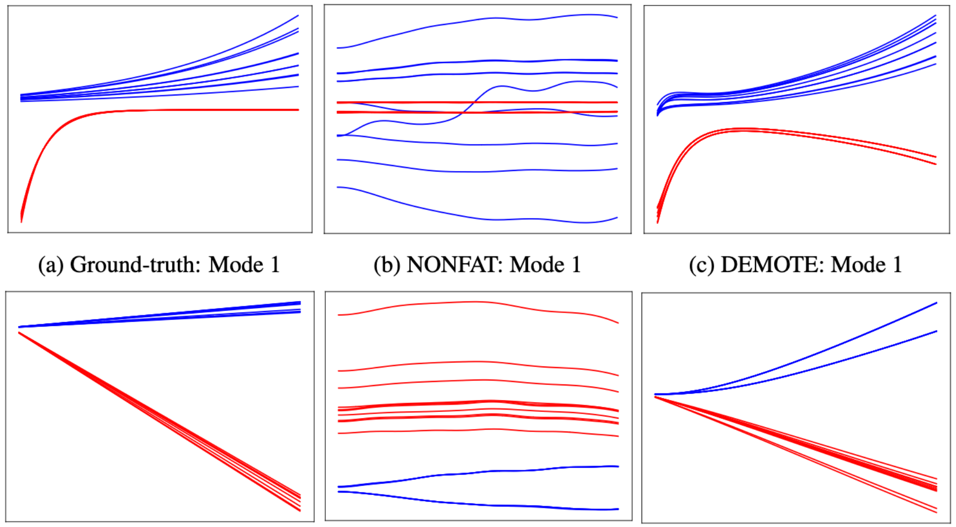
➢ **Diffusion Process on Multi-Partite Graphs**

- Capture correlations between related entities via diffusion process

➢ **Reaction Process of Individual Entities**

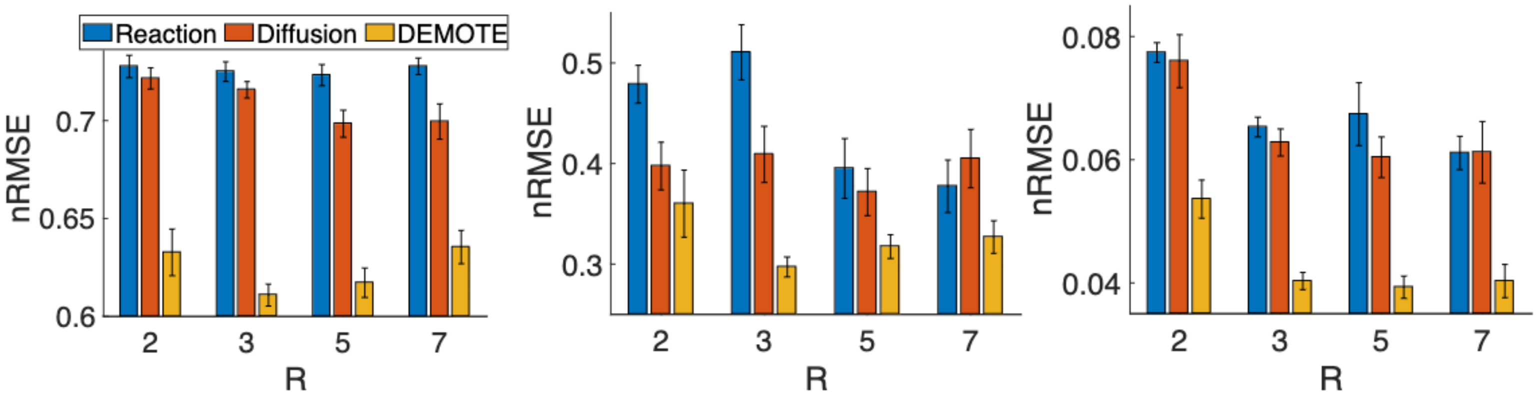
- Formulate entity self-evolvement

To capture underlying dynamics accurately, both diffusion and reaction processes are essential



(a) Ground-truth: Mode 1 (b) NONFAT: Mode 1 (c) DEMOTE: Mode 1

(d) Ground-truth: Mode 2 (e) NONFAT: Mode 2 (f) DEMOTE: Mode 2



(a) *CA Weather* (b) *CA Traffic* (c) *Server Room*

Dynamic Tensor Decomposition via Neural Diffusion-Reaction Processes

$$\frac{\partial \mathcal{U}(t)}{\partial t} = (\mathcal{W} - \mathcal{A})\mathcal{U}(t) + \mathcal{F}(\mathcal{U}, t), \quad \mathcal{U}(0) = \mathcal{U}_0$$

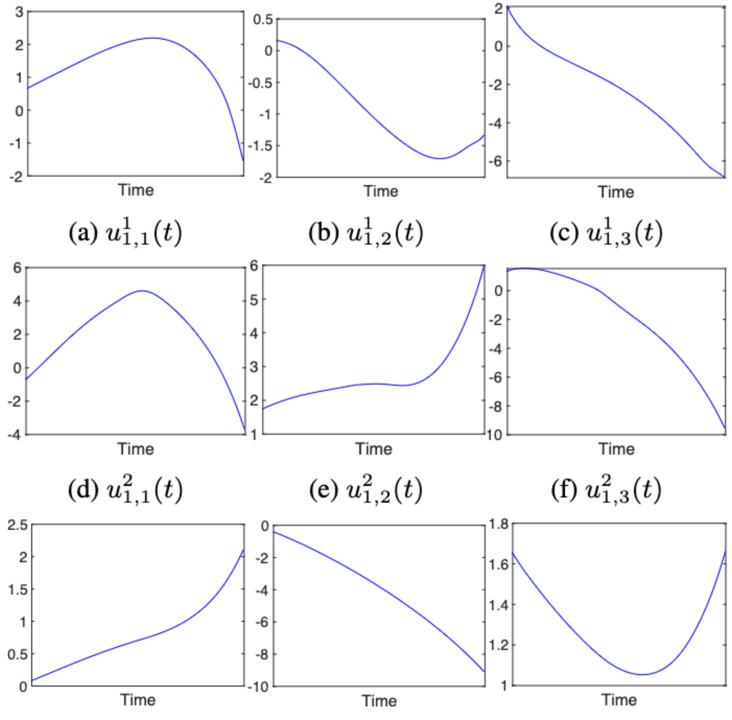
$$\mathcal{W} = \begin{pmatrix} \mathbf{0} & \mathbf{W}^{1,2} & \dots & \mathbf{W}^{1,K} \\ \mathbf{W}^{2,1} & \mathbf{0} & \dots & \vdots \\ \vdots & & \ddots & \mathbf{W}^{K-1,K} \\ \mathbf{W}^{K,1} & \dots & \mathbf{W}^{K,K-1} & \mathbf{0} \end{pmatrix}$$

$$\mathcal{A} = \mathrm{diag}\left(\sum_{s\in\{1\dots K\}\setminus 1} \mathbf{A}^{1,s}, \dots, \sum_{s\in\{1\dots K\}\setminus K} \mathbf{A}^{K,s}\right)$$

- **Diffusion Process on Multi-Partite Graphs**
  - Capture correlations between related entities via diffusion process
- **Reaction Process of Individual Entities**
  - Formulate entity self-evolvement
- **Entry Value Generation**
  - Nonlinear tensor decomposition

$$m_{\boldsymbol{\ell}}(t) = g\left(\mathbf{u}^1_{l_1}(t), \dots, \mathbf{u}^K_{l_K}(t)\right)$$

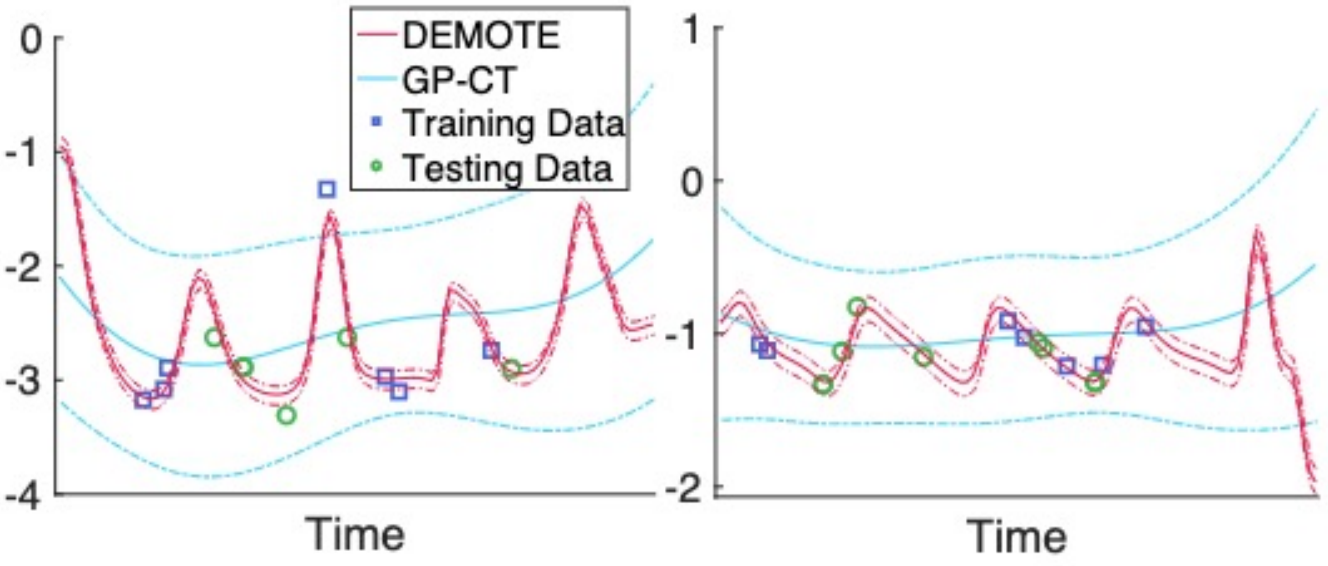
Dynamic Embedding Prediction



(a) $u^1_{1,1}(t)$ (b) $u^1_{1,2}(t)$ (c) $u^1_{1,3}(t)$

(d) $u^2_{1,1}(t)$ (e) $u^2_{1,2}(t)$ (f) $u^2_{1,3}(t)$

(g) $u^3_{1,1}(t)$ (h) $u^3_{1,2}(t)$ (i) $u^3_{1,3}(t)$



(a) (2, 1, 1) (b) (3, 1, 1)

Model Inference

- Maximize the log joint probability

$$\mathcal{L} = \log p(\boldsymbol{\beta}, \{\boldsymbol{\theta}_k\}, \mathbf{y}) = \log(\text{Prior}) - \sum_{n=1}^{N} \log \mathcal{N}\left(y_n | g\left(\mathbf{x}_n\right), \sigma^2 \mathbf{I}\right)$$

- Stratified Mini-Batch Sampling

Enhanced Predictive Performance

CA Weather: 7x6x30x30, 15K

CA Traffic: 7x6x20x20, 30K

Server Room: 34x3x3, 10K

CA Weather	$R=2$	$R=3$	$R=5$	$R=7$
CP-DTLD	0.7440 ± 0.0035	0.7372 ± 0.0040	0.7290 ± 0.0042	0.7270 ± 0.0044
GP-DTLD	0.7417 ± 0.0031	0.7414 ± 0.0036	0.7444 ± 0.0036	0.7449 ± 0.0039
NN-DTLD	0.7228 ± 0.0054	0.7116 ± 0.0033	0.7070 ± 0.0041	0.7065 ± 0.0038
CP-DTND	0.7448 ± 0.0031	0.7360 ± 0.0035	0.7273 ± 0.0037	0.7280 ± 0.0044
GP-DTND	0.7399 ± 0.0034	0.7346 ± 0.0032	0.7448 ± 0.0037	0.7467 ± 0.0031
NN-DTND	0.7113 ± 0.0045	0.6979 ± 0.0126	0.6659 ± 0.0122	0.6543 ± 0.0155
CP-CT	1.0000 ± 0.0096	0.9959 ± 0.0067	1.0010 ± 0.0017	1.0060 ± 0.0034
GP-CT	0.7433 ± 0.0038	0.7354 ± 0.0027	0.7359 ± 0.0034	0.7377 ± 0.0033
NN-CT	0.8697 ± 0.0014	0.8679 ± 0.0022	0.8676 ± 0.0018	0.8695 ± 0.0016
NONFAT	0.7444 ± 0.0042	0.7460 ± 0.0032	0.7645 ± 0.0061	0.7553 ± 0.0029
THIS-ODE	0.7511 ± 0.0052	0.7539 ± 0.0041	0.7614 ± 0.0024	0.7620 ± 0.0032
DEMOTE	$\mathbf{0.6327 \pm 0.0119}$	$\mathbf{0.6109 \pm 0.0056}$	$\mathbf{0.6172 \pm 0.0075}$	$\mathbf{0.6354 \pm 0.0085}$
CA Traffic				
CP-DTLD	0.6498 ± 0.0257	0.6424 ± 0.0266	0.6436 ± 0.0268	0.6405 ± 0.0262
GP-DTLD	0.6309 ± 0.0167	0.6290 ± 0.0185	0.6383 ± 0.0204	0.6496 ± 0.0193
NN-DTLD	0.6528 ± 0.0230	0.6545 ± 0.0244	0.6401 ± 0.0282	0.6136 ± 0.0338
CP-DTND	0.6497 ± 0.0245	0.6456 ± 0.0265	0.6431 ± 0.0263	0.6419 ± 0.0259
GP-DTND	0.6544 ± 0.0213	0.6559 ± 0.0224	0.6604 ± 0.0243	0.6674 ± 0.0214
NN-DTND	0.6578 ± 0.0248	0.6528 ± 0.0256	0.6519 ± 0.0249	0.6482 ± 0.0261
CP-CT	0.9858 ± 0.0120	0.9972 ± 0.0056	0.9816 ± 0.0136	0.9991 ± 0.0120
GP-CT	0.6610 ± 0.0207	0.6668 ± 0.0191	0.6756 ± 0.0190	0.6768 ± 0.0196
NN-CT	0.9804 ± 0.0017	0.9815 ± 0.0015	0.9791 ± 0.0012	0.9802 ± 0.0017
NONFAT	0.4461 ± 0.0247	0.4610 ± 0.0231	0.5031 ± 0.0155	0.6307 ± 0.0847
THIS-ODE	0.6603 ± 0.0230	0.6536 ± 0.0212	0.6838 ± 0.0193	0.6378 ± 0.0142
DEMOTE	$\mathbf{0.3601 \pm 0.0334}$	$\mathbf{0.2972 \pm 0.0099}$	$\mathbf{0.3174 \pm 0.0118}$	$\mathbf{0.3269 \pm 0.0162}$
Server Room				
CP-DTLD	0.4211 ± 0.0029	0.4209 ± 0.0031	0.4208 ± 0.0028	0.4208 ± 0.0028
GP-DTLD	0.0914 ± 0.0020	0.0791 ± 0.0010	0.0739 ± 0.0014	0.0753 ± 0.0013
NN-DTLD	0.4213 ± 0.0032	0.4213 ± 0.0032	0.4212 ± 0.0034	0.4205 ± 0.0030
CP-DTND	0.2835 ± 0.0160	0.1751 ± 0.0020	0.1174 ± 0.0011	0.0829 ± 0.0044
GP-DTND	0.0925 ± 0.0013	0.0784 ± 0.0011	0.0739 ± 0.0009	0.0774 ± 0.0009
NN-DTND	0.4213 ± 0.0032	0.4212 ± 0.0030	0.4211 ± 0.0032	0.4205 ± 0.0030
CP-CT	0.9919 ± 0.0096	0.9951 ± 0.0050	0.9862 ± 0.0109	1.0121 ± 0.0070
GP-CT	0.1385 ± 0.0020	0.1223 ± 0.0016	0.1275 ± 0.0014	0.1365 ± 0.0014
NN-CT	0.1193 ± 0.0030	0.1140 ± 0.0015	0.1113 ± 0.0027	0.1149 ± 0.0028
NONFAT	0.1468 ± 0.0026	0.1407 ± 0.0023	0.1396 ± 0.0022	0.1409 ± 0.0030
THIS-ODE	0.1412 ± 0.0024	0.1312 ± 0.0013	0.1304 ± 0.0016	0.1350 ± 0.0019
DEMOTE	$\mathbf{0.0536 \pm 0.0031}$	$\mathbf{0.0403 \pm 0.0014}$	$\mathbf{0.0393 \pm 0.0018}$	$\mathbf{0.0403 \pm 0.0027}$

Conclusion

The proposed neural diffusion-reaction process model enhanced the predictive performance by learning dynamic embeddings for dynamic tensor decomposition, and the learned embedding trajectories exhibit interesting patterns.

Thanks!