

# Supplementary for the Manuscript: “Multi-view dynamic kernelized evidential clustering”

## I. OPTIMIZATION OF MvDKE

In the section, we will show the details of the optimization and derivation process for MvDKE, as a supplement to Sec. III of the manuscript.

### A. Preliminary partition

The objective function of the first step, preliminary partition, in MvDKE is constructed in Eq. (16) of the manuscript. We use  $J_{MvDKE}$  to represent the designed objective function, constrained by Eq. (17) of the manuscript.

The specific updated rules are shown as follows.

- Fix  $\mathbf{V}(t)$  and  $\mathbf{w}(t)$ , update  $\mathbf{U}(t+1)$ ,
- Fix  $\mathbf{U}(t+1)$  and  $\mathbf{V}(t)$ , update  $\mathbf{w}(t+1)$ ,
- Fix  $\mathbf{U}(t+1)$ ,  $\mathbf{w}(t+1)$ , update  $\mathbf{V}(t+1)$ .

**1) Update  $\mathbf{U}$ :** BBAs  $\mathbf{U}$  are updated when  $\mathbf{w}$  and  $\mathbf{V}$  are fixed. We solve the constrained minimization problem with respect to  $\mathbf{U}$  and the Lagrangian functions are given as follows.

$$\begin{aligned} \mathcal{L}(\mathbf{U}, \boldsymbol{\lambda}) = & \sum_{q=1}^Q w_q \sum_{i=1}^N \sum_{j=1}^C (\mu_{ij}^q)^\beta (\mathbf{d}_{ij}^q)^2 + \eta \sum_{q=1}^Q w_q \ln w_q \\ & + \sum_{q=1}^Q w_q \sum_{s=1}^N \sum_{k=1}^N (\mathbf{d}_{sk}^q)^2 \\ & - \sum_{q=1}^Q \sum_{i=1}^N \lambda_i^q \left( \sum_{j=1}^C \mu_{ij}^q - 1 \right). \end{aligned} \quad (1)$$

where  $\boldsymbol{\lambda}$  is the Lagrangian multiplier. By differentiating the Lagrangian with respect to the  $\mu_{ij}^q$  and  $\lambda_i^q$ , we obtain

$$\frac{\partial \mathcal{L}(\mathbf{U}, \boldsymbol{\lambda})}{\partial \mu_{ij}^q} = w_q \cdot \beta \cdot (\mu_{ij}^q)^{\beta-1} \cdot (\mathbf{d}_{ij}^q)^2 - \lambda_i^q = 0, \quad (2)$$

$$\frac{\partial \mathcal{L}(\mathbf{U}, \boldsymbol{\lambda})}{\partial \lambda_i^q} = \sum_{j=1}^C \mu_{ij}^q - 1. \quad (3)$$

Setting Eq. (2) to zero, we obtain

$$\mu_{ij}^q = \left( \frac{\lambda_i^q}{w_q \beta (\mathbf{d}_{ij}^q)^2} \right)^{\frac{1}{\beta-1}} = \left( \frac{\lambda_i^q}{w_q \beta} \right)^{\frac{1}{\beta-1}} \cdot \left( \frac{1}{(\mathbf{d}_{ij}^q)^2} \right)^{\frac{1}{\beta-1}}. \quad (4)$$

Using Eqs. (3)-(4), we have

$$\sum_{j=1}^C \left( \frac{\lambda_i^q}{w_q \beta (\mathbf{d}_{ij}^q)^2} \right)^{\frac{1}{\beta-1}} = 1, \quad (5)$$

and we thereby obtain

$$\left( \frac{\lambda_i^q}{w_q \beta} \right)^{\frac{1}{\beta-1}} = \frac{1}{\sum_{j=1}^C \left[ \frac{1}{(\mathbf{d}_{ij}^q)^2} \right]^{\frac{1}{\beta-1}}}. \quad (6)$$

Returning in Eq. (4), we obtain the optimized solution of  $\mathbf{U}$

$$\mu_{ij}^q = \frac{\left[ (\mathbf{d}_{ij}^q)^2 \right]^{-\frac{1}{\beta-1}}}{\sum_{j=1}^C \left[ (\mathbf{d}_{ij}^q)^2 \right]^{-\frac{1}{\beta-1}}}. \quad (7)$$

**2) Update  $\mathbf{w}$ :** Similarly, we consider that  $\mathbf{U}$  and  $\mathbf{w}$  are fixed and update view weights  $\mathbf{w}$ . We introduce  $Q$  Lagrange multipliers  $\kappa$  and define the Lagrangian functions as follows.

$$\begin{aligned} \mathcal{L}(\mathbf{w}, \kappa) = & \sum_{q=1}^Q w_q \sum_{i=1}^N \sum_{j=1}^C (\mu_{ij}^q)^\beta (\mathbf{d}_{ij}^q)^2 + \eta \sum_{q=1}^Q w_q \ln w_q \\ & + \sum_{q=1}^Q w_q \sum_{s=1}^N \sum_{k=1}^N (\mathbf{d}_{sk}^q)^2 - \kappa \left( \sum_{q=1}^Q w_q - 1 \right). \end{aligned} \quad (8)$$

By differentiating the Lagrangian with respect to  $w_q$  and setting the derivatives to zero, we obtain

$$\begin{aligned} \frac{\partial \mathcal{L}(\mathbf{w}, \kappa)}{\partial w_q} = & \sum_{i=1}^N \sum_{j=1}^C (\mu_{ij}^q)^\beta (\mathbf{d}_{ij}^q)^2 + \sum_{s=1}^N \sum_{k=1}^N (\mathbf{d}_{sk}^q)^2 \\ & + \eta(1 + \ln w_q) - \kappa, \end{aligned} \quad (9)$$

$$\frac{\partial \mathcal{L}(\mathbf{w}, \kappa)}{\partial \kappa} = \sum_{q=1}^Q w_q - 1, \quad (10)$$

where we define

$$\psi \triangleq \sum_{i=1}^N \sum_{j=1}^C (\mu_{ij}^q)^\beta (\mathbf{d}_{ij}^q)^2 + \sum_{s=1}^N \sum_{k=1}^N (\mathbf{d}_{sk}^q)^2. \quad (11)$$

From Eqs. (9) and (11), we thereby obtain

$$w_q = e^{\frac{\kappa-\eta-\psi}{\eta}} = e^{\frac{\kappa-\eta}{\eta}} \cdot e^{-\frac{\psi}{\eta}}. \quad (12)$$

Combining with Eq. (10)

$$\sum_{q=1}^Q e^{\frac{\kappa-\eta}{\eta}} \cdot e^{-\frac{\psi}{\eta}} = 1, \quad (13)$$

we thus obtain

$$e^{\frac{\kappa-\eta}{\eta}} = \frac{1}{\sum_{q=1}^Q e^{-\frac{\psi}{\eta}}}. \quad (14)$$

Returning in Eq. (9), we obtain the optimized  $\mathbf{w}$

$$w_q = \frac{e^{-\frac{\psi}{\eta}}}{\sum_{q=1}^Q e^{-\frac{\psi}{\eta}}}. \quad (15)$$

3) **Update  $\mathbf{V}$ :** Considering  $\mathbf{U}, \mathbf{w}$  are fixed, the minimization of  $\mathbf{V}$  is an unconstrained optimization problem. For  $q$ th view, the partial derivatives of  $J_{MvDKE}$  and  $(\mathcal{d}_{ij}^q)^2$  with respect to the centers are given by

$$\begin{aligned} \frac{\partial J_{MvDKE}}{\partial \phi(\mathbf{v}_j^q)} &= w_q \sum_{i=1}^N (\mu_{ij}^q)^\beta \cdot \frac{\partial (\mathcal{d}_{ij}^q)^2}{\partial \phi(\mathbf{v}_j^q)} \\ &= w_q \sum_{i=1}^N (\mu_{ij}^q)^\beta \cdot 2 \cdot -(\phi(\mathbf{x}_i^q) - \phi(\mathbf{v}_j^q)). \end{aligned} \quad (16)$$

Setting these derivatives to zero, we obtain

$$\phi(\mathbf{v}_j^q) = \frac{\sum_{i=1}^N (\mu_{ij}^q)^\beta \cdot \phi(\mathbf{x}_i^q)}{\sum_{i=1}^N (\mu_{ij}^q)^\beta}. \quad (17)$$

The corresponding inner product can be calculated by

$$\begin{aligned} \phi(\mathbf{v}_s^q)^T \phi(\mathbf{v}_k^q) &= \frac{\sum_{i=1}^N (\mu_{is}^q)^\beta \phi(\mathbf{x}_i^q) \cdot \sum_{i=1}^N (\mu_{ik}^q)^\beta \phi(\mathbf{x}_i^q)}{\sum_{i=1}^N (\mu_{is}^q)^\beta \cdot \sum_{i=1}^N (\mu_{ik}^q)^\beta} \\ &= \frac{\sum_{i=1}^N \sum_{j=1}^N (\mu_{is}^q)^\beta \cdot (\mu_{jk}^q)^\beta \cdot \phi(\mathbf{x}_i^q) \cdot \phi(\mathbf{x}_j^q)}{\sum_{i=1}^N (\mu_{is}^q)^\beta \cdot \sum_{i=1}^N (\mu_{ik}^q)^\beta} \\ &= \frac{\sum_{i=1}^N \sum_{j=1}^N (\mu_{is}^q)^\beta \cdot (\mu_{jk}^q)^\beta \cdot \mathcal{K}(\mathbf{x}_i^q, \mathbf{x}_j^q)}{\sum_{i=1}^N \sum_{j=1}^N (\mu_{is}^q)^\beta \cdot (\mu_{jk}^q)^\beta}, \end{aligned} \quad (18)$$

and

$$\begin{aligned} \phi(\mathbf{x}_i^q)^T \phi(\mathbf{v}_j^q) &= \phi(\mathbf{x}_i^q) \cdot \frac{\sum_{k=1}^N (\mu_{kj}^q)^\beta \phi(\mathbf{x}_k^q)}{\sum_{k=1}^N (\mu_{kj}^q)^\beta} \\ &= \frac{\sum_{k=1}^N (\mu_{kj}^q)^\beta \mathcal{K}(\mathbf{x}_i^q, \mathbf{x}_k^q)}{\sum_{k=1}^N (\mu_{kj}^q)^\beta}. \end{aligned} \quad (19)$$

Therefore, we can obtain the kernelized distance  $\mathcal{d}_{ij}^q$  by taking Eqs. (18)-(19) to

$$(\mathcal{d}_{ij}^q)^2 = \mathcal{K}(\mathbf{x}_i^q, \mathbf{x}_i^q) + \mathcal{K}(\mathbf{v}_j^q, \mathbf{v}_j^q) - 2\mathcal{K}(\mathbf{x}_i^q, \mathbf{v}_j^q). \quad (20)$$

The updated kernelized distance  $\mathcal{d}_{ij}^q$  is used to update  $\mathbf{U}$  in the iteration.

## B. Objects management

No optimization for this step.

### C. Partial reassignment

In this subsection, we first show the optimization process of objective function of partial reassignment, which is represented by Eq. (28) of the manuscript, constrained by Eq. (29) of the manuscript.

The Lagrange multiplier method is used to optimize the variables  $\mathbf{U}$  in Eq. (28). We obtain

$$\begin{aligned} \mathcal{L}(\mathbf{U}, \boldsymbol{\delta}) &= \sum_{q=1}^Q \sum_{i=1}^N \sum_{\{j|A_j \in \mathcal{S}^{\Omega_i}\}} (\mu_{ij}^q)^\beta \cdot \frac{(\mathcal{D}_{ij}^q)^2}{\rho_j^{q*}} \\ &\quad - \sum_{q=1}^Q \sum_{i=1}^N \delta_i^q \left( \sum_{\{j|A_j \in \mathcal{S}^{\Omega_i}\}} \mu_{ij}^q - 1 \right). \end{aligned} \quad (21)$$

Setting the derivatives  $\mathbf{U}$  with respect to  $\delta$  to zero, we have

$$\frac{\partial \mathcal{L}(\mathbf{U}, \boldsymbol{\delta})}{\partial \mu_{ij}^q} = \cdot \beta \cdot (\mu_{ij}^q)^{\beta-1} \cdot \frac{(\mathcal{D}_{ij}^q)^2}{\rho_j^{q*}} - \delta_i^q = 0, \quad (22)$$

we thereby obtain

$$\mu_{ij}^q = \left( \frac{\delta_i^q \cdot \rho_j^{q*}}{\beta (\mathcal{D}_{ij}^q)^2} \right)^{\frac{1}{\beta-1}} = \left( \frac{\delta_i^q}{\beta} \right)^{\frac{1}{\beta-1}} \cdot \left( \frac{\rho_j^{q*}}{(\mathcal{D}_{ij}^q)^2} \right)^{\frac{1}{\beta-1}}, \quad (23)$$

Similarly, by setting the derivatives with respect to  $\delta$  to zero, we have

$$\frac{\partial \mathcal{L}(\mathbf{U}, \boldsymbol{\delta})}{\partial \delta_i^q} = \sum_{\{j|A_j \in \mathcal{S}^{\Omega_i}\}} \mu_{ij}^q - 1. \quad (24)$$

Combining Eqs. (23)-(24), we obtain

$$\left( \frac{\delta_i^q}{\beta} \right)^{\frac{1}{\beta-1}} = \frac{1}{\sum_{\{j|A_j \in \mathcal{S}^{\Omega_i}\}} \left( \frac{\rho_j^{q*}}{(\mathcal{D}_{ij}^q)^2} \right)^{\frac{1}{\beta-1}}}. \quad (25)$$

Returning in Eq. (23), we obtain the optimized  $\mathbf{U}$

$$\mu_{ij}^q = \frac{\left( \frac{\rho_j^{q*}}{(\mathcal{D}_{ij}^q)^2} \right)^{\frac{1}{\beta-1}}}{\sum_{\{j|A_j \in \mathcal{S}^{\Omega_i}\}} \left( \frac{\rho_j^{q*}}{(\mathcal{D}_{ij}^q)^2} \right)^{\frac{1}{\beta-1}}}. \quad (26)$$

Since only one variable  $\mathbf{U}$  in the objective function need to be optimized, we can obtain multi-view BBAs for imprecision objects after only one iteration.

In addition, we give the derivation process of  $\mathcal{D}_{ij}^q$ , represented by Eq. (27) in the manuscript. When  $|A_j| = 1$ , we have  $(\mathcal{D}_{ij}^q)^2 = (\mathcal{d}_{ij}^q)^2$ , which can be calculated by Eq. (20).

When  $|A_j| > 1$ , we have

$$(\mathcal{D}_{ij}^q)^2 = \frac{\sum_{\{k|a_k \in A_j\}} (\mathcal{d}_{ik}^q)^2 + \gamma (\mathcal{d}_{ij}^q)^2}{|A_j| + \gamma}, \quad (27)$$

where the sum of the distance from object  $i$  to singleton clusters  $a_k \in A_j$  can be calculated in Eq. (28). The distance from object  $i$  to meta-cluster  $A_j$  is calculated in Eq. (29). From Eqs. (27)-(29), we thereby obtain Eq. (30).

$$\begin{aligned}
\sum_{\{k|a_k \in A_j\}} (\mathcal{A}_{ik}^q)^2 &= \sum_{\{k|a_k \in A_j\}} \|\phi(\mathbf{x}_i^q) - \phi(\mathbf{v}_k^q)\|^2 \\
&= \sum_{\{k|a_k \in A_j\}} [\mathcal{K}(\mathbf{x}_i^q, \mathbf{x}_i^q) + \mathcal{K}(\mathbf{v}_k^q, \mathbf{v}_k^q) - 2\mathcal{K}(\mathbf{x}_i^q, \mathbf{v}_k^q)] \\
&= |A_j| \mathcal{K}(\mathbf{x}_i^q, \mathbf{x}_i^q) + \sum_{\{k|a_k \in A_j\}} [\mathcal{K}(\mathbf{v}_k^q, \mathbf{v}_k^q) - 2\mathcal{K}(\mathbf{x}_i^q, \mathbf{v}_k^q)]
\end{aligned} \tag{28}$$

$$\begin{aligned}
(\mathcal{D}_{ij}^q)^2 &= \|\phi(\mathbf{x}_i^q) - \phi(\mathbf{v}_j^q)\|^2 \\
&= \left\| \phi(\mathbf{x}_i^q) - \frac{\sum_{\{k|a_k \in A_j\}} \phi(\mathbf{v}_k^q)}{|A_j|} \right\|^2 \\
&= \left[ \phi(\mathbf{x}_i^q) - \frac{1}{|A_j|} \sum_{\{k|a_k \in A_j\}} \phi(\mathbf{v}_k^q) \right]^T \left[ \phi(\mathbf{x}_i^q) - \frac{1}{|A_j|} \sum_{\{k|a_k \in A_j\}} \phi(\mathbf{v}_k^q) \right] \\
&= \frac{1}{|A_j|^2} \left[ \sum_{\{k|a_k \in A_j\}} (\phi(\mathbf{x}_i^q) - \phi(\mathbf{v}_k^q))^T \sum_{\{k|a_k \in A_j\}} (\phi(\mathbf{x}_i^q) - \phi(\mathbf{v}_k^q)) \right] \\
&= \frac{1}{|A_j|^2} \sum_{\{k|a_k \in A_j\}} [\phi(\mathbf{x}_i^q) - \phi(\mathbf{v}_k^q)]^T [\phi(\mathbf{x}_i^q) - \phi(\mathbf{v}_k^q)] + \frac{2}{|A_j|^2} \sum_{\{s,k|\{a_s,a_k\} \subseteq A_j\}} [\phi(\mathbf{x}_i^q) - \phi(\mathbf{v}_s^q)]^T [\phi(\mathbf{x}_i^q) - \phi(\mathbf{v}_k^q)] \\
&= \frac{1}{|A_j|^2} \sum_{\{k|a_k \in A_j\}} [\phi(\mathbf{x}_i^q)^2 + \phi(\mathbf{v}_k^q)^2 - 2\phi(\mathbf{x}_i^q)\phi(\mathbf{v}_k^q)] \\
&\quad + \frac{2}{|A_j|^2} \left[ \frac{|A_j|(|A_j|-1)}{2} \phi(\mathbf{x}_i^q)^2 + \sum_{\{s,k|\{a_s,a_k\} \subseteq A_j\}} \phi(\mathbf{v}_s^q)\phi(\mathbf{v}_k^q) \right] - (|A_j|-1) \sum_{\{k|a_k \in A_j\}} \phi(\mathbf{x}_i^q)\phi(\mathbf{v}_k^q) \\
&= \phi(\mathbf{x}_i^q)^2 + \frac{1}{|A_j|^2} \sum_{\{k|a_k \in A_j\}} \phi(\mathbf{v}_k^q)^2 + \frac{2}{|A_j|^2} \sum_{\{s,k|\{a_s,a_k\} \subseteq A_j\}} \phi(\mathbf{v}_s^q)\phi(\mathbf{v}_k^q) - \frac{1}{|A_j|} \sum_{\{k|a_k \in A_j\}} \phi(\mathbf{x}_i^q)\phi(\mathbf{v}_k^q) \\
&= \mathcal{K}(\mathbf{x}_i^q, \mathbf{x}_i^q) + \frac{1}{|A_j|^2} \sum_{\{k|a_k \in A_j\}} \mathcal{K}(\mathbf{v}_k^q, \mathbf{v}_k^q) + \frac{2}{|A_j|^2} \sum_{\{s,k|\{a_s,a_k\} \subseteq A_j\}} \mathcal{K}(\mathbf{v}_s^q, \mathbf{v}_k^q) - \frac{1}{|A_j|} \sum_{\{k|a_k \in A_j\}} \mathcal{K}(\mathbf{x}_i^q, \mathbf{v}_k^q)
\end{aligned} \tag{29}$$

$$\begin{aligned}
(\mathcal{D}_{ij}^q)^2 &= \left[ (|A_j| + \gamma) \mathcal{K}(\mathbf{x}_i^q, \mathbf{x}_i^q) + \frac{|A_j|^2 + \gamma}{|A_j|^2} \sum_{\{k|a_k \in A_j\}} \mathcal{K}(\mathbf{v}_k^q, \mathbf{v}_k^q) + \frac{2|A_j| - \gamma}{|A_j|} \sum_{\{k|a_k \in A_j\}} \mathcal{K}(\mathbf{x}_i^q, \mathbf{v}_k^q) \right. \\
&\quad \left. + \frac{2\gamma}{|A_j|^2} \sum_{\{s,k|\{a_s,a_k\} \subseteq A_j\}} \mathcal{K}(\mathbf{v}_s^q, \mathbf{v}_k^q) \right] \cdot \frac{1}{|A_j| + \gamma}, \quad |A_j| > 1
\end{aligned} \tag{30}$$

## II. EXPERIMENTS

This section presents complete experimental results and more analysis.

In this section, we first present the visualization results of MvDKE on more toy datasets. And then we give the complete comparison result in five real datasets. We also show the parameter experiment on the more real datasets. In addition, we supplement the kernel selection experiments to present the clustering performance when different kernel functions are applied to MvDKE.

There are 20 clustering methods used in this experiment, including both single-view clustering and multi-view clustering methods, ECM [1], DEC [2], KFCM [3], CDMGC [4], CGD [5], Co-FW-MVFCM [6], CoFKM [7], GMC [8], LMSC [9], LTMSC [10], MCSSC [11], TBGL-MVC [12], MVASM [13], EEOMVC [14], EMGC2F [15], UDBGL [16], FastMICE [17], SMCMB [18], MvWECK [19] and MvDKE.

All the parameters of these methods are set as shown in Table I. For all these methods, we assume that the number of clusters  $C$  is known.

TABLE I  
INFORMATION OF THE METHODS.

Method	view	cluster	Kernel	Parameter Set
ECM	single-view	evidential	-	$\alpha = 2, \beta = 2, \delta = 20$
DEC	single-view	evidential	-	$\beta = 2, \epsilon = 0.3, \delta = 20$
KFCM	single-view	fuzzy	Gaussian	$\beta = 2, \theta = 150$
CDMGC	multi-view	graph-based	-	$\alpha = 1e5, \beta = 1e - 5, knn = 9$
CGD	multi-view	graph-based	-	$\sigma = 0.5$
CoFKM	multi-view	fuzzy	-	$\beta = 1.25$
Co-FW-MVFCM	multi-view	fuzzy	-	$m = 2, \beta = 4$
GMC	multi-view	graph-based	-	$\lambda = 1, pn = 15$
LMSM	multi-view	subspace	-	$\lambda = 0.1$
LTMSC	multi-view	subspace	-	$\lambda = 0.1$
MCSSC	multi-view	graph-based	-	$\lambda = 10, \gamma = 1$
TBGL-MVC	multi-view	graph-based	-	$p = 0.9, \text{anchor rate} = 0.5$
MVASM	multi-view	<i>k</i> -means	-	$\gamma = 0.5, q = 2$
EEOMVC	multi-view	graph-based	-	$k = 15, \lambda = 0.5, \beta = 0.5$ anchor rate = 2%
EMGC2F	multi-view	graph-based	-	-
UDBGL	multi-view	subspace	-	$\alpha = 1e - 3, \beta = 1e - 5$
SMCMB	multi-view	graph-based	-	Number of anchors=100, Number of bipartite graphs=5, $\beta = 1e - 4$
MvWECKM	multi-view	evidential	-	$c = 3, K[t, s] = 1, \beta = 25, \eta = 1$
MvDKE	multi-view	evidential	Gaussian	$\eta = 1, \gamma = 1e - 1, \theta = adpt$

### A. Visualization results on toy dataset

In this subsection, we generate seven toy datasets to demonstrate the effectiveness of MvDKE in characterizing imprecision in multi-view non-spherical data. The information of these generated toy datasets is shown in Table II.

TABLE II  
DETAILS OF TOY DATASETS.

Dataset	N	Q	C
Toy1	2500	2	7
Toy2	1000	2	3
Toy3	5000	3	3
Toy4	200	3	3
Toy5	200	2	3
Toy6	4000	2	4
Toy7	4000	3	4

As a supplement to the manuscript, the visualization of original datasets and clustering results on the Toy3-Toy7 dataset are shown in Figs. 1-4.

### B. Experiments on real dataset

Mnist Handwrite Dataset [20], a widespread benchmark for machine learning, is applied to further show the potential of MvDKE. We show the partial results of hard-to-identify objects in Table III. We can see that non-evidential methods force these hard-to-identify objects into a singleton cluster, most of them are misassigned. In contrast, MvDKE does not assign them to a singleton cluster precisely but narrows down the range of possible singleton clusters, and the correct singleton cluster is always included. Compared with DEC, MvDKE can give a smaller range of possible singleton clusters and has higher clustering accuracy.

### C. Parameter study

1) *Parameters  $\theta$  in Gaussian kernel:* We give additional experiment results for Parameters  $\theta$  to demonstrate the wide applicability of our view-adaptive  $\theta$  calculation. The results on nine UCI datasets are presented in Fig. 5.

2) *Parameters  $\eta$  and  $\gamma$  in MvDKE:* In Subsec. IV-C of the manuscript, we only present the performance on the vehicle dataset for parameter study. Here we provide more results on the other six real datasets in Figs. 6-11.

### REFERENCES

- [1] M.-H. Masson and T. Denoeux, "Ecm: An evidential version of the fuzzy *c*-means algorithm," *Pattern Recognition*, vol. 41, no. 4, pp. 1384–1397, 2008.
- [2] Z.-w. Zhang, Z. Liu, A. Martin, Z.-g. Liu, and K. Zhou, "Dynamic evidential clustering algorithm," *Knowledge-Based Systems*, vol. 213, p. 106643, 2021.
- [3] D.-Q. Zhang and S.-C. Chen, "A novel kernelized fuzzy *c*-means algorithm with application in medical image segmentation," *Artificial Intelligence in Medicine*, vol. 32, no. 1, pp. 37–50, 2004.
- [4] S. Huang, I. W. Tsang, Z. Xu, and J. Lv, "Measuring diversity in graph learning: A unified framework for structured multi-view clustering," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 12, pp. 5869–5883, 2021.
- [5] C. Tang, X. Liu, X. Zhu, E. Zhu, Z. Luo, L. Wang, and W. Gao, "Cgd: Multi-view clustering via cross-view graph diffusion," in *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, 2020, pp. 5924–5931.
- [6] M. Yang and K. P. Sinaga, "Collaborative feature-weighted multi-view fuzzy *c*-means clustering," *Pattern Recognition*, vol. 119, p. 108064, 2021.
- [7] G. Cleuziou, M. Exbrayat, L. Martin, and J.-H. Sublemonier, "Cofkm: A centralized method for multiple-view clustering," in *ICDM 2009, The Ninth IEEE International Conference on Data Mining, Miami, Florida, USA, 6-9 December 2009*. IEEE, 2009, pp. 752–757.

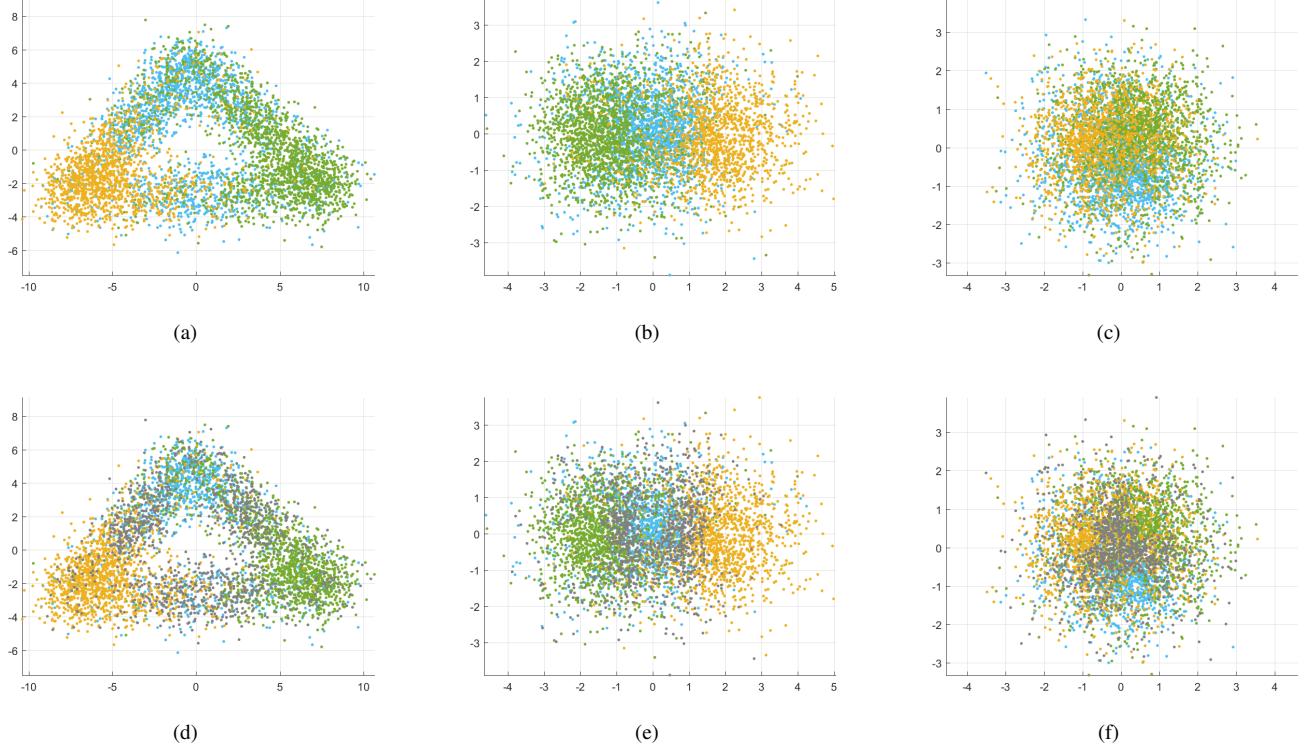


Fig. 1. Results on the Toy3 dataset, (a)-(c) the original 3-view Toy3 dataset, (d)-(f) the results of MvDKE on the Toy3 dataset.

- [8] H. Wang, Y. Yang, and B. Liu, "Gmc: Graph-based multi-view clustering," *IEEE Transactions on Knowledge and Data Engineering*, vol. 32, no. 6, pp. 1116–1129, 2019.
- [9] C. Zhang, H. Fu, Q. Hu, X. Cao, Y. Xie, D. Tao, and D. Xu, "Generalized latent multi-view subspace clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 1, pp. 86–99, 2018.
- [10] C. Zhang, H. Fu, S. Liu, G. Liu, and X. Cao, "Low-rank tensor constrained multiview subspace clustering," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 1582–1590.
- [11] X. Gao, X. Ma, W. Zhang, J. Huang, H. Li, Y. Li, and J. Cui, "Multi-view clustering with self-representation and structural constraint," *IEEE Transactions on Big Data*, vol. 8, no. 4, pp. 882–893, 2021.
- [12] W. Xia, Q. Gao, Q. Wang, X. Gao, C. Ding, and D. Tao, "Tensorized bipartite graph learning for multi-view clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.
- [13] J. Han, J. Xu, F. Nie, and X. Li, "Multi-view  $k$ -means clustering with adaptive sparse memberships and weight allocation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 2, pp. 816–827, 2020.
- [14] J. Wang, C. Tang, Z. Wan, W. Zhang, K. Sun, and A. Y. Zomaya, "Efficient and effective one-step multiview clustering," *IEEE Transactions on Neural Networks and Learning Systems*, 2023.
- [15] D. Wu, J. Lu, F. Nie, R. Wang, and Y. Yuan, "Emgc2f: Efficient multi-view graph clustering with comprehensive fusion," in *IJCAI*, 2022, pp. 3566–3572.
- [16] S.-G. Fang, D. Huang, X.-S. Cai, C.-D. Wang, C. He, and Y. Tang, "Efficient multi-view clustering via unified and discrete bipartite graph learning," *IEEE Transactions on Neural Networks and Learning Systems*, 2023.

TABLE III  
EXAMPLES ON THE MNIST HANDWRITE DATASET.

	0	6	5	2	6	7	7	2	9	9
TrueLabel	0	0	1	2	6	7	7	8	9	9
CDMGC	0	6	7	2	4	1	1	8	1	1
CGD	8	6	6	1	4	7	7	8	7	7
Co-FW-MVFCM	0	9	1	1	4	1	7	8	1	1
CoFKM	6	6	7	2	4	9	9	7	7	1
LMSMC	6	6	1	2	6	1	7	4	1	1
LTMSC	0	6	7	7	4	1	9	8	9	7
MCSSC	6	6	9	1	6	9	9	3	9	1
TBGL-MVC	0	0	1	7	4	1	1	8	1	1
MVASM	4	0	1	7	4	1	7	8	1	1
KFCM	8	6	1	2	2	1	7	3	1	1
DEC	{0,2,6,8}	{0,6}	{1,9}	{1,2,9}	{4,5,6}	{1,7,9}	{7,9}	{2,3,4,8}	{7,9}	{7,9}
MvDKE	{0,6}	0	{1,9}	{2,7}	{4,6}	{1,7}	{7,9}	8	9	{7,9}

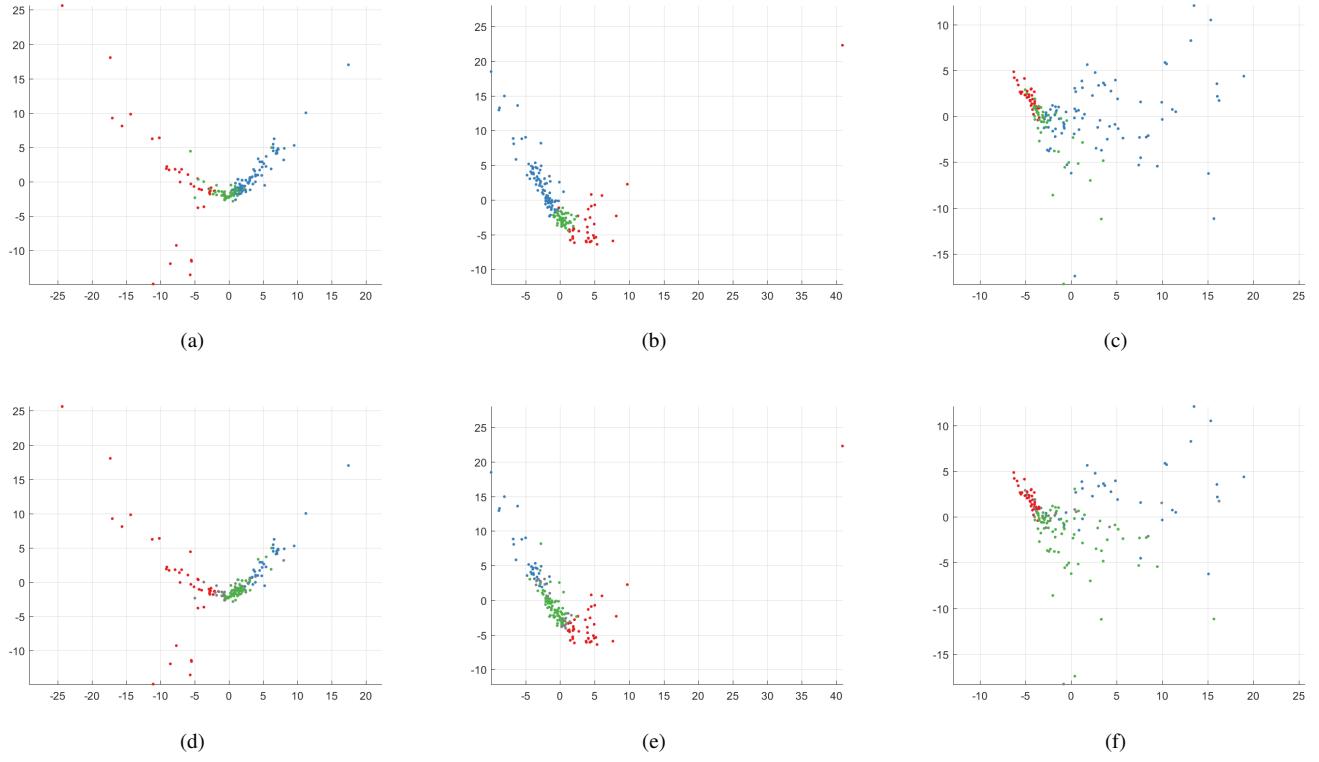


Fig. 2. Results on the Toy4 dataset, (a)-(c) the original 3-view Toy4 dataset, (d)-(f) the results of MvDKE on the Toy4 dataset.

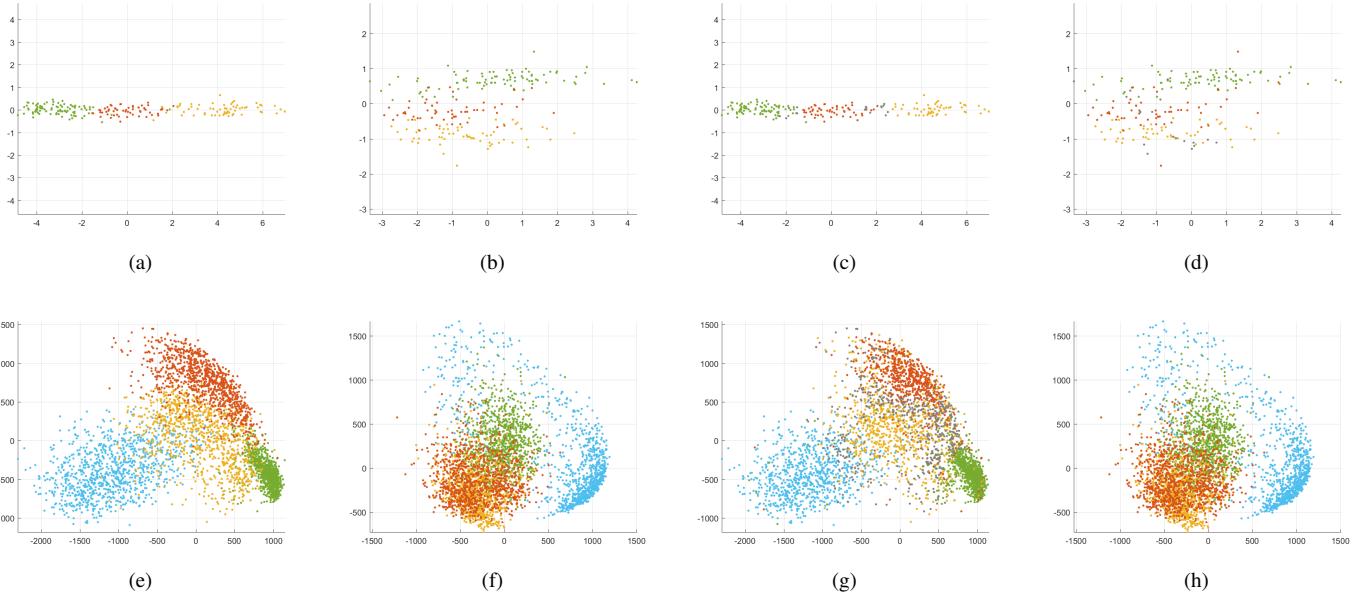


Fig. 3. Results on the Toy5 and Toy6 dataset, (a)-(d) the original 2-view and the results of MvDKE on the Toy5 dataset, (e)-(h) the original 2-view and the results of MvDKE on the Toy6 dataset.

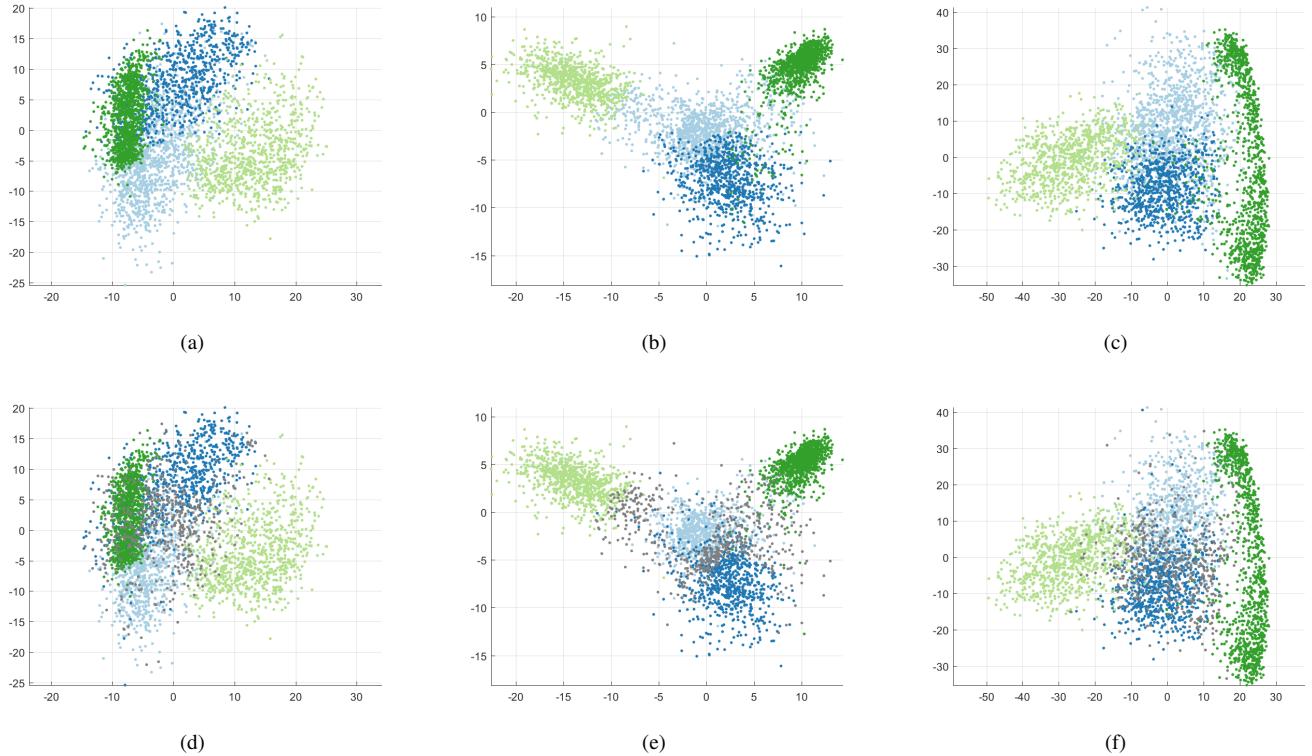


Fig. 4. Results on the Toy7 dataset, (a)-(c) the original 3-view Toy7 dataset, (d)-(f) the results of MvDKE on the Toy7 dataset.

- [17] D. Huang, C.-D. Wang, and J.-H. Lai, “Fast multi-view clustering via ensembles: Towards scalability, superiority, and simplicity,” *IEEE Transactions on Knowledge and Data Engineering*, 2023.
- [18] J. Lao, D. Huang, C.-D. Wang, and J.-H. Lai, “Towards scalable multi-view clustering via joint learning of many bipartite graphs,” *IEEE Transactions on Big Data*, 2023.
- [19] K. Zhou, M. Guo, and M. Jiang, “Evidential weighted multi-view clustering,” in *Belief Functions: Theory and Applications: 6th International Conference, BELIEF 2021, Shanghai, China, October 15–19, 2021, Proceedings 6*. Springer, 2021, pp. 22–32.
- [20] L. Deng, “The mnist database of handwritten digit images for machine learning research [best of the web],” *IEEE Signal Processing Magazine*, vol. 29, no. 6, pp. 141–142, 2012.

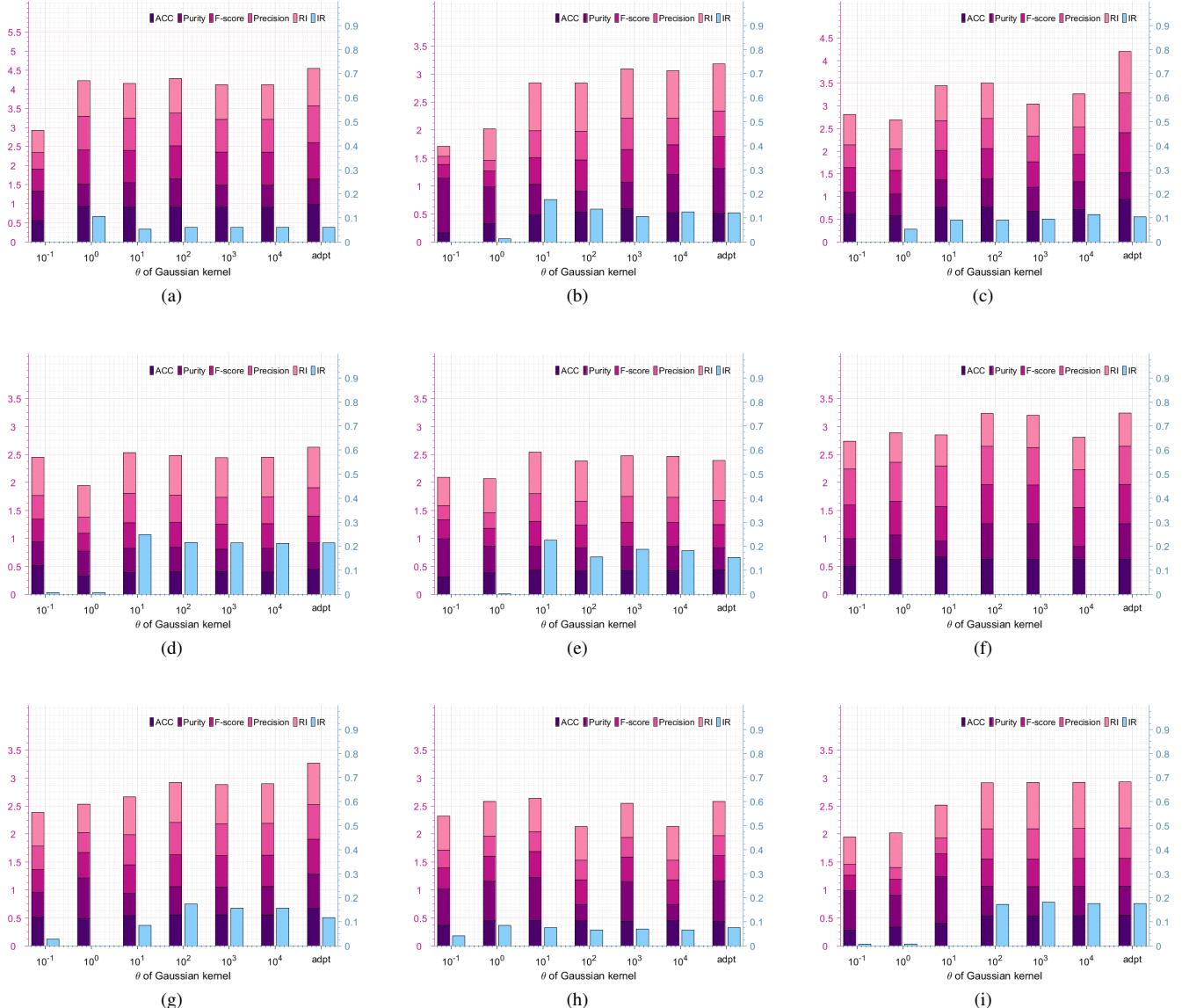


Fig. 5. How Parameter  $\theta$  of gaussian kernel affects the performance of MvDKE on the real-world datasets, (a)Iris, (b)segment, (c)seeds, (d)foresttype, (e)vehicle, (f)spectf, (g)wine, (h)glass, (i)satimage.

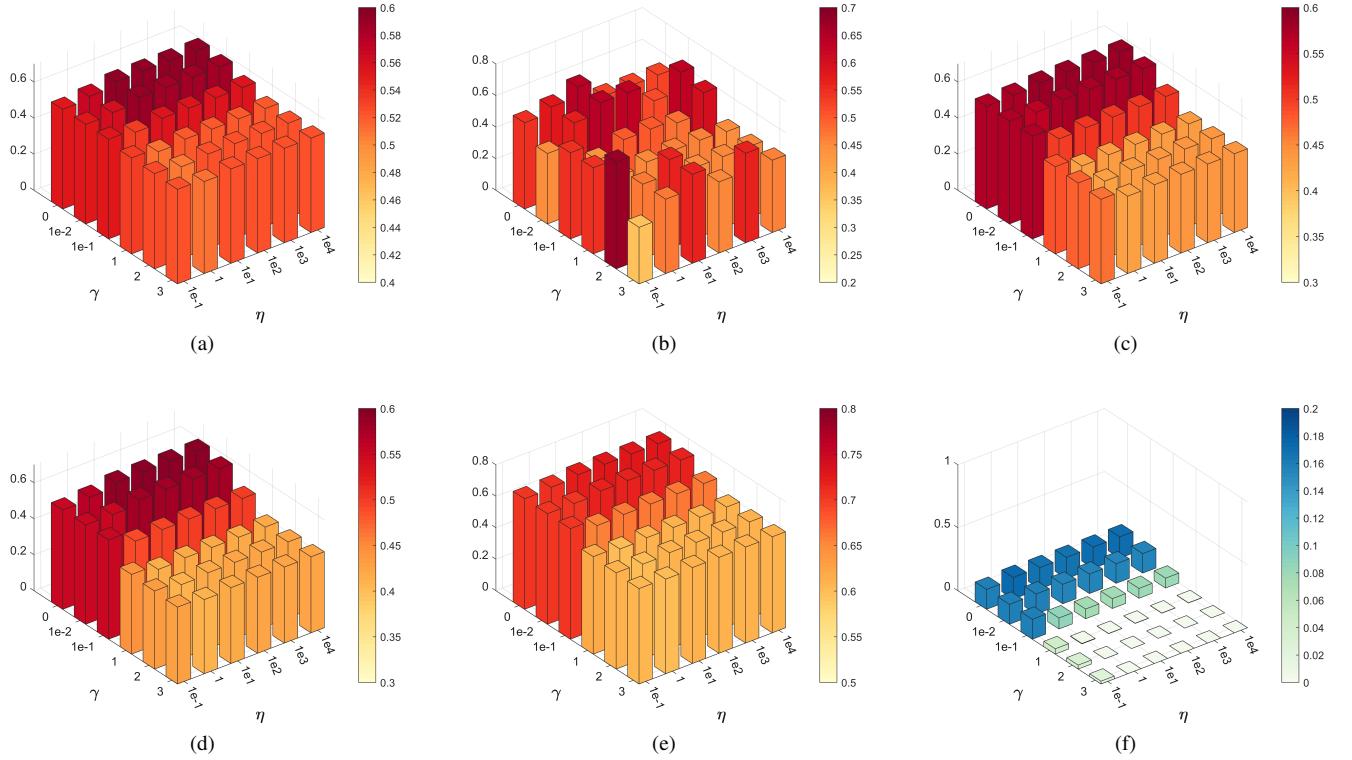


Fig. 6. How Parameter  $\eta$  and  $\gamma$  affect the performance of MvDKE on the abalone dataset, (a)ACC, (b)Purity, (c)F1-score, (d)Precision, (e)RI, (f)IR, where  $x$  and  $y$  axis are  $\eta$  and  $\gamma$ .

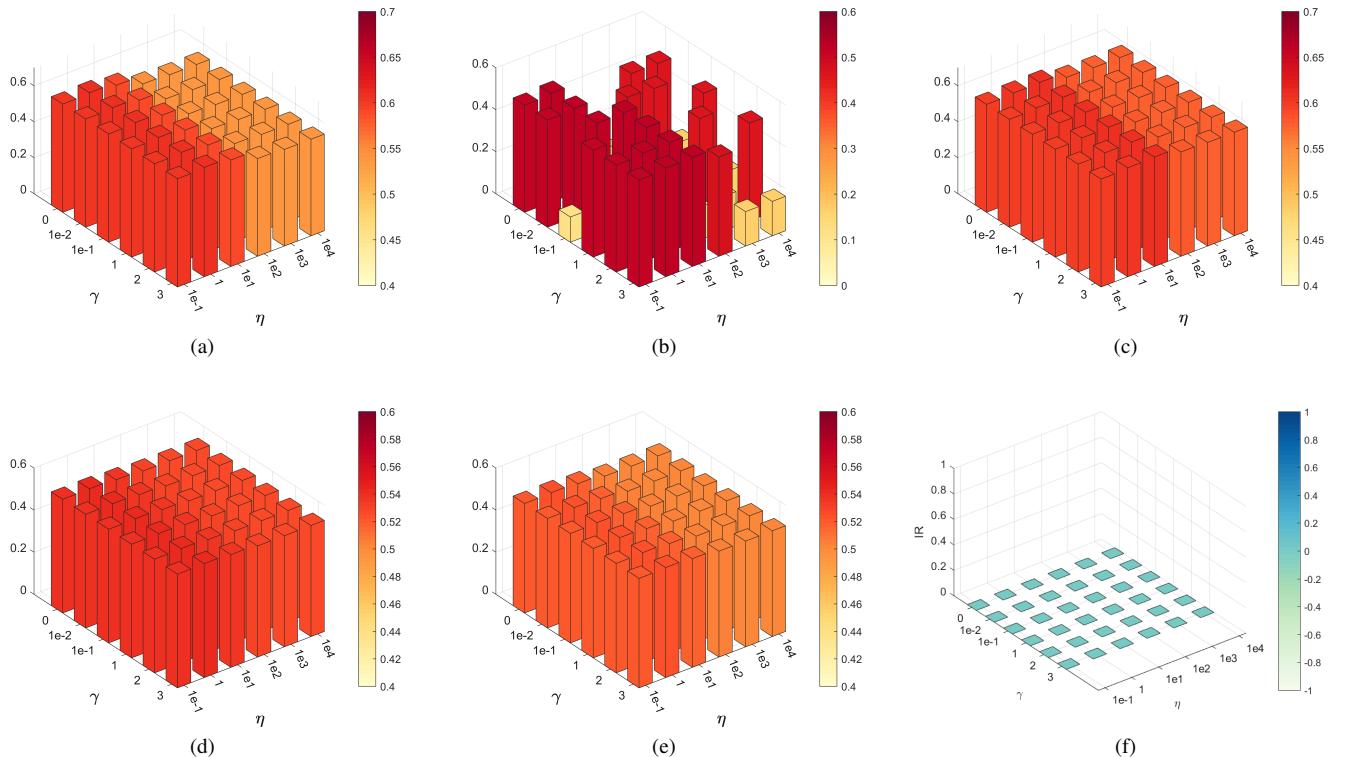


Fig. 7. How Parameter  $\eta$  and  $\gamma$  affect the performance of MvDKE on the german dataset, (a)ACC, (b)Purity, (c)F1-score, (d)Precision, (e)RI, (f)IR, where  $x$  and  $y$  axis are  $\eta$  and  $\gamma$ .

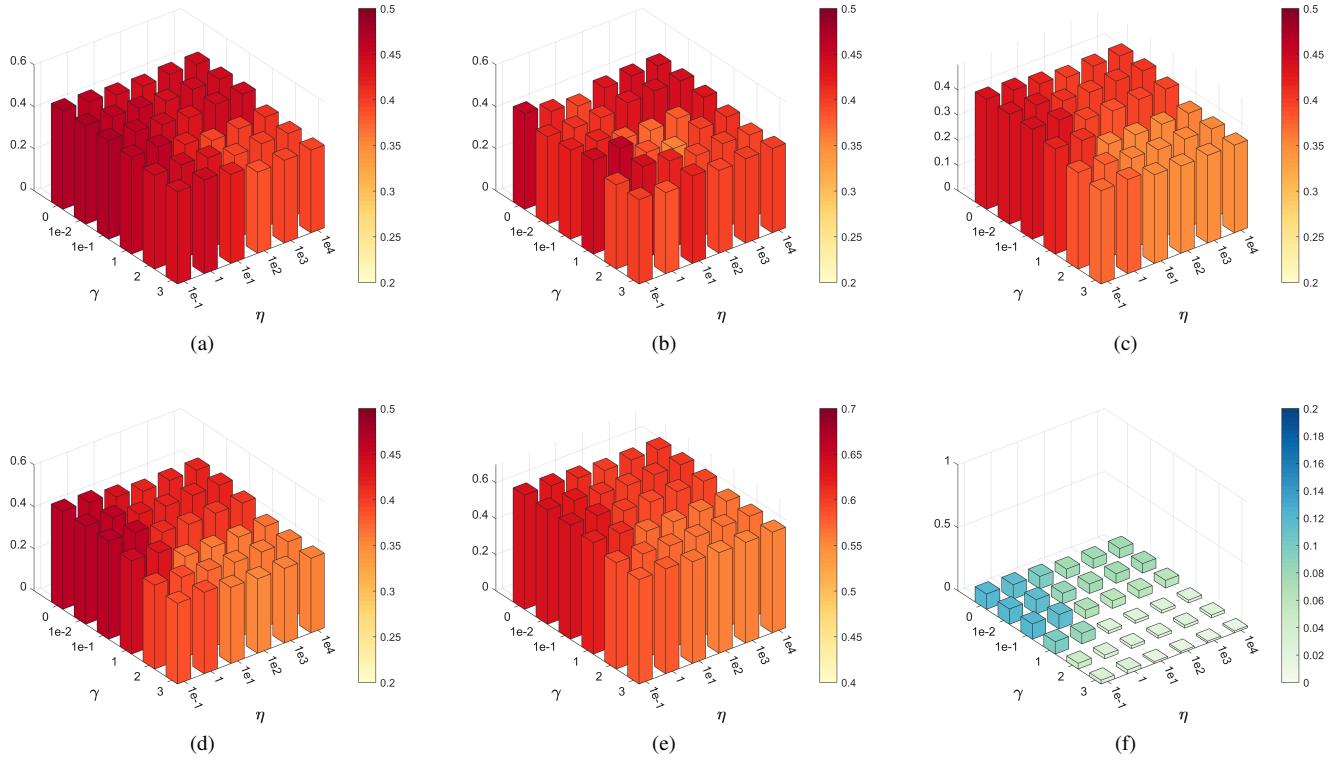


Fig. 8. How Parameter  $\eta$  and  $\gamma$  affect the performance of MvDKE on the hayes dataset, (a)ACC, (b)Purity, (c)F1-score, (d)Precision, (e)RI, (f)IR, where  $x$  and  $y$  axis are  $\eta$  and  $\gamma$ .

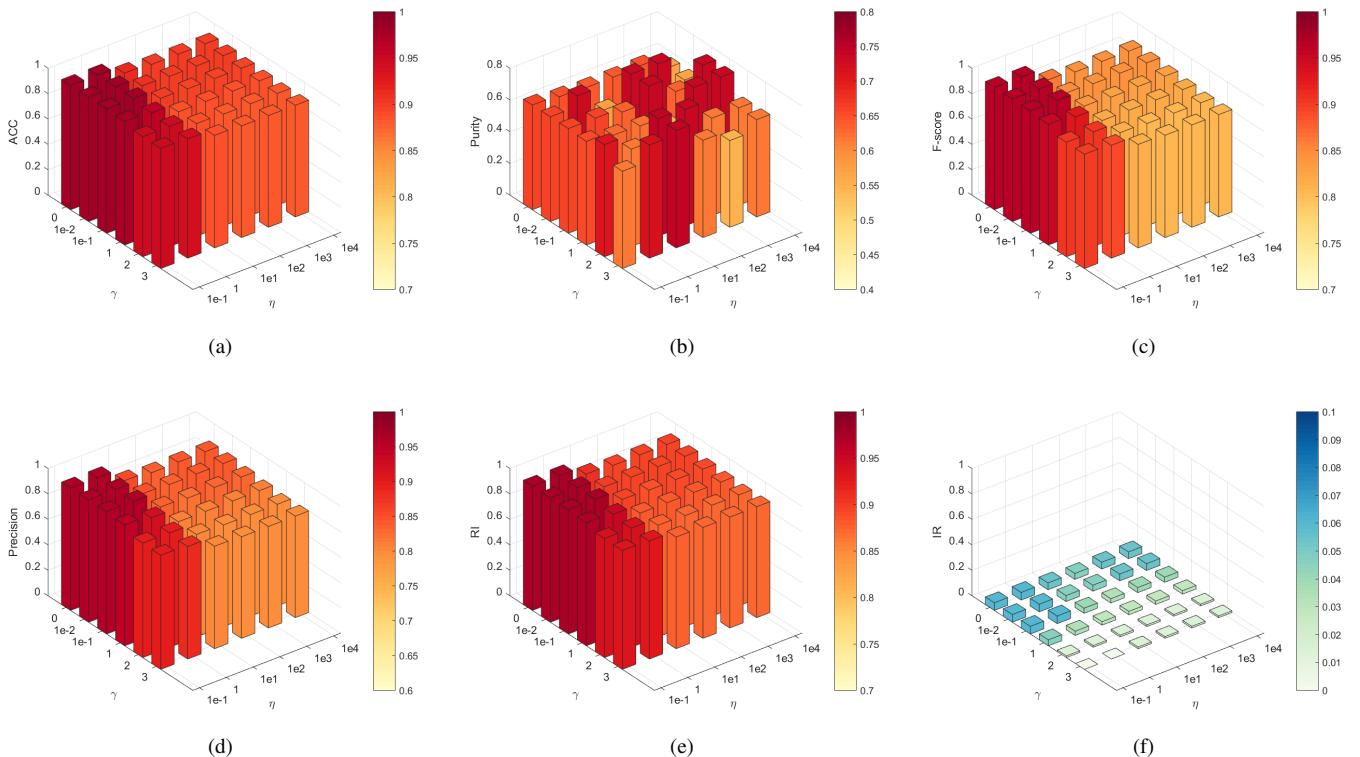


Fig. 9. How Parameter  $\eta$  and  $\gamma$  affect the performance of MvDKE on the iris dataset, (a)ACC, (b)Purity, (c)F1-score, (d)Precision, (e)RI, (f)IR, where  $x$  and  $y$  axis are  $\eta$  and  $\gamma$ .

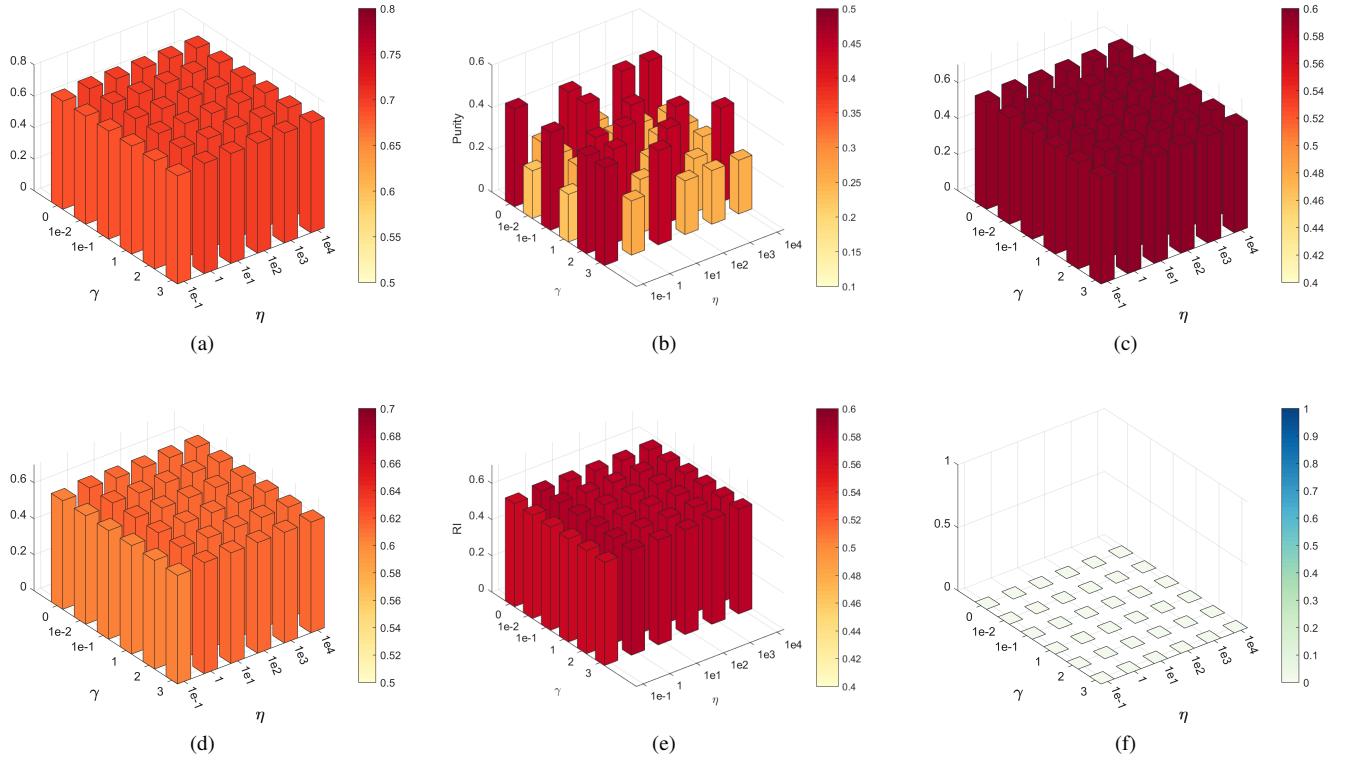


Fig. 10. How Parameter  $\eta$  and  $\gamma$  affect the performance of MvDKE on the ionosphere dataset, (a)ACC, (b)Purity, (c)F1-score, (d)Precision, (e)RI, (f)IR, where  $x$  and  $y$  axis are  $\eta$  and  $\gamma$ .

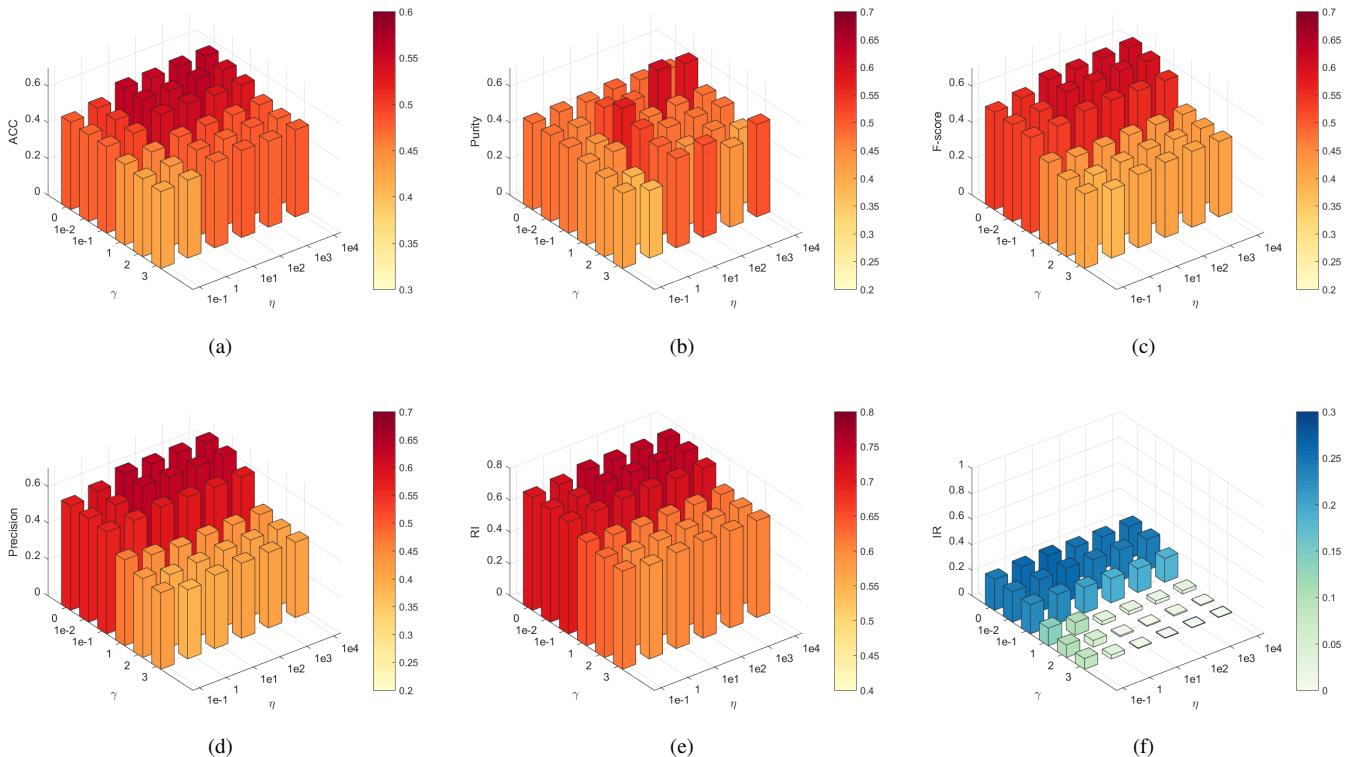


Fig. 11. How Parameter  $\eta$  and  $\gamma$  affect the performance of MvDKE on the waveform dataset, (a)ACC, (b)Purity, (c)F1-score, (d)Precision, (e)RI, (f)IR, where  $x$  and  $y$  axis are  $\eta$  and  $\gamma$ .