IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2020.DOI

Single-cell RNA-sequencing Data Clustering via Locality Preserving Kernel Matrix Alignment

XIAO ZHENG[1,†], JIAJIA CHEN[2,†], CHANG TANG[3], AND SUQIN ZHOU[4,*]
[1]School of Computer, National University of Defense Technology, Changsha 410073, China
[2]Department of Pharmacy, The Affiliated Huai'an Hospital of Xuzhou Medical University, Huai'an 223100, China
[3]School of Computer Science, China University of Geosciences, Wuhan 430074, China
[4]Department of Pharmacy, Lianshui people's Hospital Affiliated to Kangda College, Nanjing Medical University, Huai'an 223300, China

†Xiao Zheng and Jiajia Chen contributed equally as first author to this work.
*Corresponding author: Suqin Zhou (e-mail: sqinzhou@163.com).

This work was supported in part by the National Science Foundation of China under Grant 61701451 and 62076228.

ABSTRACT Single-cell RNA-sequencing (scRNA-seq) data provide opportunities to reveal new insights into many biological problems such as elucidating cell types. An effective approach to elucidate cell types in complex tissues is to partition the cells into several separated subgroups via clustering techniques, where the cells in a specific cluster belong to the same cell type based on gene expression patterns. In this work, we present a novel multiple kernel clustering framework for scRNA-seq data clustering via locality preserving kernel alignment. Specifically, we first generate a series of similarity kernel matrices by using different kernel functions. Then we transfer the clustering task to a multiple kernel k-means problem with the kernels aligned in a local manner, i.e., the similarity of a sample to its k-nearest neighbours are aligned with the ideal similarity matrix. In our method, the clustering process focuses on closer sample pairs that shall stay together, and avoids involving unreliable similarity evaluation for farther sample pairs. In addition, we construct a local Laplacian matrix for each sample to constrain that closer samples should be allocated similar labels. In such a manner, the local structure of the data can be well preserved and utilized to produce better alignment for clustering. An alternate updating algorithm with theoretical analysis is developed to solve the proposed problem. We evaluate the performance of the proposed method on various real scRNA-seq data, and the results show that our method can obtain superior results when compared with other state-of-the-art approaches.

INDEX TERMS scRNA-seq clustering, elucidating cell types, tissues, similarity kernel matrices, locality preservation.

I. INTRODUCTION

Recent literature indicate that single-cell measurements plays an important role in understanding cellular heterogeneity [1]–[5] and cell differentiation [6], [7]. Thanks to the rapid development of Single-cell RNA-sequencing (scRNA-seq) techniques, a tremendous amount of transcriptome datasets have been generated at single-cell resolution [8], [9]. On the one hand, these datasets provide opportunities to reveal new insights into many biological problems, e.g., elucidating cell types, on the other hand, there are also computational challenges due to the amount of data. A straightforward approach to elucidate cell types in complex tissues is to partition the cells into some separated subgroups via clustering techniques [10]–[13], which can be regarded as an unsupervised classification problem [14]–[16]. Many previous clustering techniques can be used for this task, such as principal component analysis (PCA) [17], spectral clustering [18], and k-means [19]. However, different to bulk RNA-seq or gene expression microarrays, there are high level of noise and many missing values in scRNA-seq data due to technical and sampling issues [20]–[22]. In addition, the high variability exists in gene expression levels even between cells of the same type, and this could degenerate the performance of those existing clustering approaches [23]–[27].

In order to address the issues in scRNA-seq data analysis, a various of novel clustering methods have been proposed in recent years. For subtype classification and detection of relationships between the subtypes, some iterative clustering methods have been proposed [28]–[31]. Haghverdi et al. [32] used the diffusion maps to perform dimension reduction of the data, which stresses continuity of cell states along putative developmental pathways. Nonnegative Matrix Factorization (NMF) technique has also been used to decompose the high-dimensional scRNA-seq data into biologically interpretable compositions [33], and the functional cell subgroups and biologically relevant features can be simultaneously obtained with NMF. Several graph theory-based algorithms have also been applied to scRNA-seq data clustering problems. Xu and Su [34] developed a quasi-clique-based clustering algorithm named SNN-Cliq to identify tight groups of highly similar nodes that are likely to belong to the same genuine clusters. In SNN-Cliq, the clusters are identified by using the proposed SNN graph. Spectral clustering, as a typical graph based clustering method, has also been successfully deployed for this task. Park and Zhao [35] first constructed a series of symmetric doubly stochastic similarity matrices by using the Gaussian kernel function with varying parameters, then they learned a target similarity matrix from the previous constructed matrices for spectral clustering. In [36], Wu et al. integrated dimension reduction and clustering of single-cell RNA-sequencing data into a unified framework.

Multiple kernel clustering is also a kind of popular modern clustering method which aims to optimally integrate a group of pre-specified kernels to improve clustering performance. A critical issue in multiple kernel clustering is to learn the kernel combination weights. Margolin [37] made the combination weights adaptively change with respect to samples to better capture their individual characteristics. Du et al. [38] presented a robust multiple kernel k-means algorithm that simultaneously finds the best clustering labels and the optimal combination of multiple kernels by replacing the squared error in k-means with an $l_{2,1}$-norm based one. Lu et al. [39] employed kernel alignment maximization to jointly perform the k-means clustering and multiple kernel learning. Wang et al. [40] presented an analytic framework (SIMLR) via multi-kernel learning which learns a similarity measure from scRNA-seq data in order to perform dimension reduction, clustering and visualization. Compared to other previous methods, SIMLR learns a distance metric that best fits the structure of the data by combining multiple kernels. Standard dimension reduction or clustering algorithms often work under certain statistical assumptions which the diverse statistical characteristics of scRNA-seq data could not easily fit well. Qi et al. [41] proposed to automatically learn similarity information from data and introduced a new clustering method in the form of a multiple kernel combination, which can directly discover groupings in scRNA-seq data. In this paper, we propose a new scRNA-seq data clustering method via locality preserving multiple kernel alignment (referred to as LPKA briefly). Considering that previous kernel alignment methods often rigidly constrain closer and farther sample pairs to be equally aligned to the same ideal similarity, and the intra-cluster variation of samples is inappropriately neglected, we propose to align the kernels in a local manner, i.e., the similarity of a sample to only its k-nearest neighbours are aligned with the ideal similarity matrix. In our method, the clustering process focuses on closer sample pairs that shall stay together, and avoids involving unreliable similarity evaluation for farther sample pairs. In addition, we construct a local Laplacian matrix for each sample to constrain that closer samples should be allocated similar labels. In such a manner, the local structure of the data can be well preserved and utilized to produce better alignment for clustering. Experiments on 9 scRNA-seq datasets are conducted to demonstrate the superiority of our proposed method.

II. MATERIALS AND METHODS

A. DATASETS COLLECTION AND KERNEL MATRICES GENERATION

1) Datasets

In order to evaluate the efficacy of our proposed LPKA, we use some real-world scRNA-seq datasets to test the clustering performance. Similar to [35], we test the performance of LPKA on 9 scRNAseq datasets which represent several types of dynamic processes such as cell cycle, cell differentiation, and response upon external stimulus. For each dataset, the types of cells are known as a priori. The number of cells, number of cell types, number of genes for each dataset are summarized in Table 1.

TABLE 1: Brief information of the 9 used scRNA-seq datasets in our experiments.

Datasets	No. of cells	No. of genes	No. of cell types
Treutlein	80	9352	5
Ting	114	14405	5
Deng	135	12548	7
Ginhoux	251	11834	3
Buettner	182	8989	3
Pollen	249	14805	11
Tasic	1727	5832	49
Zeisel	3005	4412	47
Macosko	6418	12822	39

For the readability and integrity of this paper, we also give the detailed information of the datasets as follows:

- Treutlei [42]. This dataset is composed of single cell RNA-seq expression data for 80 lung epithelial cells at E18.5 together with five putative cell types including AT1, AT2, Clara, BP, and ciliated. Similar to [42], we considered data with selected genes with 959 highest loadings in the first four PCA coefficients. The dataset is downloaded from: https://www.nature.com/articles/nature13173.
- Ting. This dataset contains contains 5 subtypes from Single-cell transcriptomes from MEFs, the NB508 pancreatic cancer cell line, normal WBCs, bulk primary tumors diluted to 10 or 100 pg of RNA, and classical CTC. We downloaded the data from GEO (GSE51372).
- Deng [43]. This dataset consists of transcriptomes for individual cells isolated from mouse embryos at different preimplantation stages. There are 135 cells and

19,703 genes, where cells belong to zygote, early 2-cell-stage, mid 2-cell-stage, late 2-cellstage, 4-cell-stage, 8-cell-stage, and 16-cell-stage. The processed data is downloaded from GEO (GSE45719).

- Ginhoux. This dataset consists of the expression values of 15,752 genes for 251 dendritic cell progenitors in one of following three cellular states: Monocyte and Dendritic cell Progenitors (MDPs), Common Dendritic cell Progenitors (CDPs), and Pre-Dendritic Cells (PreDCs). The dataset contains 59 MDPs, 96 CDPs, and 96 PreDCs. We downloaded the processed data from GEO (GSE60783).
- Buettner [44]. This dataset contains the transcriptional profile of 182 ESCs that has been staged for cell-cycle phase (G1, S, and G2M) based on sorting of the Hoechst 33342-stained cell area of a flow cytometry (FACS) distribution. The cells were sorted for three stages of the cell cycle, and they were validated using gold-standard Hoechst staining. The data have been deposited at ArrayExpress: E-MTAB-2805.
- Pollen. There are 249 single cells from 11 populations using microfluidics, including neural cells and blood cells. The 11 clusters in the dataset were from different sources (CRL-2338, CRL-2339, K562, BJ, HL60, hiPSC, Keratinocyte, Fetal cortex (GW21+3), Fetal cortex (GW21), Fetal cortex (GW16), and NPC) that are expected to show robust differences in gene expression. Data were pre-filtered to exclude genes where more than 90% of cells had zero measurements and include only single cells with greater than 500000 reads (n = 249).
- Tasic [45]. There are 49 transcriptomic cell types in this dataset, including 23 GABAergic, 19 glutamatergic and 7 non-neuronal types. To identify cell types, Tasic et al. [45] applied two parallel and iterative approaches for dimensionality reduction and clustering, iterative principal component analysis (PCA) and iterative weighted gene coexpression network analysis (WGCNA), and validated the cluster membership from each approach using a non-deterministic machine learning method (random forest). We downloaded the processed data from GEO (GSE71585).
- Zeisel [28]. In this dataset, Zeisel et al. [28] used large-scale single-cell RNA sequencing to classify cells in the mouse somatosensory cortex and hippocampal CA1 region. 3005 Cells from the mouse cortex and hippocampus collected. Zeisel et al. (2015) found 47 molecularly distinct subclasses identified by hierarchical biclustering and validated by gene markers.
- Macosko. This dataset contains mouse retina cells with 39 subtypes, it is obtained by droplet-based high-throughput technique. The dataset consists of 44808 cells. The 39 cell types were identified via PCA and density-based clustering, and they were validated by differential gene expression. We filtered out cells with less than 1200 genes (yielding 6418 cells) for clustering analysis. We downloaded the data from GEO (GSE63473).

2) Kernel matrices generation

In our experiments, three kinds of kernels are used to generate kernel matrices, including Gaussian kernel, linear kernel and polynomial kernel.

For Gaussian kernel, it is one of the most widely used kernel function and it can obtain steadily performance whether the data size is large or small. We follow [40] and consider multiple kernel functions to construct kernel matrices as follows:

$$GK_{\theta,k}(i,j) = \exp(-\frac{||x_i - x_j||^2}{2\sigma_{ij}^2}), \tag{1}$$

where $\sigma_{ij} = \frac{\theta(\pi_i+\pi_j)}{2}$, and $\pi_i = \frac{\sum_{k\in KNN(i)} ||x_i-x_j||_2}{k}$. $KNN(i)$ represents a set of sample indices that are the top k nearest neighbors of the sample x_i. As can be seen, parameters θ and k control the width of the neighborhoods. For generality, we also vary θ from $\{1, 1.25, \cdots, 2\}$ and k from $\{10, 12, \cdots, 30\}$. Thus, a total number of 55 Gaussian kernel matrices can be obtained.

For linear kernel, it is suitable for the data samples which are linear separable. In addition, it has no parameter. We generate a linear kernel matrix as follows:

$$LK(i,j) = x_i \cdot x_j. \tag{2}$$

For polynomial kernel, it projects low dimensional feature space to a higher dimensional feature space. It is defined as follows:

$$PK(i,j) = ((x_i \cdot x_j) + 1)^d. \tag{3}$$

In this work, we vary d from $\{0.1, 0.2, \cdots, 1\}$ and obtain 10 polynomial kernel matrices.

Finally, for each dataset, we combine different kinds of kernel matrices to obtain a kernel matrix with size $n \times n \times 66$ for further computation, where n is the number of samples in the dataset.

B. KERNEL K-MEANS CLUSTERING

Given a set of n data samples from k clusters $\{x_i\}_{i=1}^n \subseteq \mathcal{X}$, let $\phi(\cdot) : x \in \mathcal{X} \mapsto \mathcal{H}$ be a feature mapping which maps x from original space onto a reproducing kernel Hilbert space $\mathcal{H}$. Kernel k-means aims to minimize the sum-of-square loss over the cluster assignment matrix $Z \in \{0,1\}^{n\times k}$, and the problem can be solved by minimizing following objective function:

$$\min_{Z\in\{0,1\}^{n\times k}} \sum_{i=1}^{n}\sum_{c=1}^{k} Z_{ic}||\phi(x_i) - \xi_c||_2^2, \quad s.t. \quad \sum_{c=1}^{k} Z_{ic} = 1, \tag{4}$$

where $n_c = \sum_{i=1}^n Z_{ic}$ and $\xi_c = \frac{1}{n_c}\sum_{i=1}^n Z_{ic}\phi(x_i)$ represent the sample number and centroid of the c-th cluster in $\mathcal{H}$.

By some algebra, (4) can be transferred to following matrix form:

$$\min_{Z\in\{0,1\}^{n\times k}} Tr(K) - Tr(L^{\frac{1}{2}} Z^T K Z L^{\frac{1}{2}}), \quad s.t. \quad Z1_k = 1_n, \tag{5}$$

where $Tr(\cdot)$ denotes the trace of a matrix and K is a kernel matrix with $K_{ij} = \phi(x_i)^T\phi(x_i)$, $L = diag([{n_1}^{-1}, {n_2}^{-1}, \cdots, {n_k}^{-1}])$, and $1_l \in \mathbb{R}^l$ is a column vector with all elements 1.

The problem (5) is hard to solve since Z is discrete. Fortunately, this problem can be usually approximated through relaxing Z to take arbitrary real values. By defining $H = ZL^{\frac{1}{2}}$ and letting H take real values, the relaxed version of (5) can be obtained as:

$$\min_{H\in\mathbb{R}^{n\times k}} Tr(K(I_n - HH^T)), \;\; s.t. \;\; H^TH = I_k, \tag{6}$$

where I_k is a $k \times k$ identity matrix. Since $Z^TZ = L^{-1}$, we have $L^{\frac{1}{2}}Z^TZL^{\frac{1}{2}} = I_k$, then it is easy to get the orthogonality constraint on H. Finally, (6) can be solved by taking the k eigenvectors that correspond to the k largest eigenvalues of K.

C. MULTIPLE KERNEL K-MEANS CLUSTERING

With a multiple kernel setting, each sample is represented via a group of feature mappings $\{\phi(\cdot)\}_{p=1}^m$ [46]. In detail, each sample can be represented as $\phi_\xi(x) = [\xi_1\phi_1(x)^T, \xi_2\phi_2(x)^T, \cdots, \xi_m\phi_m(x)^T]^T$, where $\xi = [\xi_1, \xi_2, \cdots, \xi_m]^T$ represents the coefficients of each base kernel that we need to learn. As a consequence, the corresponding kernel function over the above mapping function can be written as:

$$\kappa_\xi(x_i, x_j) = \phi_\xi(x_i)^T\phi_\xi(x_j) = \sum_{p=1}^{m} \xi_p^2\kappa_p(x_i, x_j). \tag{7}$$

By replacing the kernel matrix K in (6) with K_ξ calculated via (7), the multiple kernel k-means clustering problem can be re-rewritten as following form [37]:

$$\min_{H\in\mathbb{R}^{n\times k}, \xi\in R_+^m} Tr(K_\xi(I_n - HH^T)), \; s.t. \; H^TH = I_k, \; \xi^T1_m = 1. \tag{8}$$

D. PROPOSED MULTIPLE KERNEL K-MEANS CLUSTERING VIA LOCALITY PRESERVING MULTIPLE KERNEL ALIGNMENT

Kernel alignment maximization has been widely used to learn kernel parameters in supervised learning. However, it is not suitable to unsupervised learning that the true labels are absent [47]. An effective solution is to update kernel coefficients by maximizing the alignment between the combined kernel K_ξ and HH^T, where H can be regarded as pseudo-labels in the last iteration [48]. In specific, the kernel alignment maximization for multiple kernel clustering can be formulated as following:

$$\max_{H\in\mathbb{R}^{n\times k}, \xi\in R_+^m} \frac{\langle K_\xi, HH^T\rangle}{\sqrt{\langle K_\xi, K_\xi\rangle}} \;\; s.t. \;\; H^TH = I_k, \; \xi^T1_m = 1, \tag{9}$$

where $\langle K_\xi, HH^T\rangle = Tr(K_\xi HH^T)$, $\langle K_\xi, K_\xi\rangle = \tilde{\xi}^TM\tilde{\xi}$ with $\tilde{\xi} = [\xi_1^2, \xi_2^2, \cdots, \xi_m^2]^T$ and M is a positive semi-definite matrix with $M_{pq} = Tr(K_p^TK_q)$ [47]. Directly optimizing (9) is difficult since it is a fourth-order fractional optimization problem. Thus, we need to derive a new and approximated optimization problem.

Following we use Theorem 2 to get a second-order upper bound for the denominator in (9).

Theorem 1: $\xi^TM\xi$ is an upper bound of $\tilde{\xi}^TM\tilde{\xi}$.

Theorem 1: For a pair of semi-definite matrices K_p and K_q, there exists matrices U_p and U_q such that $K_p = U_pU_p^T$ and $K_q = U_qU_q^T$. As a results, we have $M_{pq} = Tr(K_p^TK_qj) = Tr(U_pU_p^TU_qU_q^T) = Tr((U_p^TU_q)(U_p^TU_q)^T) = ||U_p^TU_q||_F^2 \geq 0$, where $||\cdot||_F$ denotes the Frobenius norm of a matrix. We also have $\xi^TM\xi = \sum_{p,q=1}^m M_{pq}\xi_p\xi_q \geq \sum_{p,q=1}^m M_{pq}\xi_p^2\xi_q^2 = \tilde{\xi}^TM\tilde{\xi}$. This completes the proof.

$\xi^TM\xi$ is much easier to handle than $\tilde{\xi}^TM\tilde{\xi}$ since it leads to a well studied quadratic programming. In addition, this term also works as a regularization on the kernel coefficients to prevent ξ_p and ξ_q from being jointly assigned to a large weight if M_{pq} is relatively high.

In addition, we find that minimizing the negative of numerator, i.e., $-Tr(K_\xi HH^T)$, together with $\xi^TM\xi$ simultaneously cannot guarantee that the whole objective is convex w.r.t ξ with fixed H. This would degenerate the quality of solution at each iteration, leading to sub-optimal solution. Here we use the following theorem to give a good substitute of $-Tr(K_\xi HH^T)$ while with convexity.

Theorem 2: $Tr(K_\xi(I_n - HH^T))$ is convex w.r.t ξ with fixed H.

Theorem 2: Since $H^TH = I_k$, we have $HH^TH = H$. By denoting $H = [h_1, h_2, \cdots, h_k]$, we can obtain that $HH^Th_c = h_c$, $\forall 1 \leq c \leq k$. This means HH^T has k eigenvalues with 1 and its rank is no more than k, which implies that it has $n - k$ eigenvalues with 0. As a consequence, $I_n - HH^T$ has $n - k$ and k eigenvalues with 1 and 0. This induces $Tr(K_p(I_n - HH^T)) \geq 0$. With $Tr(K_\xi(I_n - HH^T)) = \sum_{p=1}^m \xi_p^2 Tr(K_p(I_n - HH^T))$, we can conclude that $Tr(K_\xi(I_n - HH^T)$ is convex w.r.t. ξ with fixed H.

According to the above-mentioned observations, the maximization problem described by (9) can be turned to following minimization problem:

$$\begin{aligned} \min_{H\in R^{n\times k}, \xi\in R_+^m} & Tr(K_\xi(I_n - HH^T)) + \frac{\lambda}{2}\xi^TM\xi \\ & s.t. \; H^TH = I_k, \; \xi^T1_m = 1, \end{aligned} \tag{10}$$

where λ is a parameter used to balance the two terms.

As can be seen from (9) and (10), they maximize the kernel alignment between the combined kernel matrices K_ξ and the ideal kernel matrix HH^T globally. In such a way, closer and farther sample pairs will be equally aligned to the same ideal similarity, intra-cluster variation of samples will be neglected. In other words, the discrimination between samples are not fully exploited. In addition, the closer samples are not constrained to share similar label vectors. Therefore, the locality of samples are not well preserved, which is a critical priori in unsupervised learning. In this paper, we propose to locally align the similarity of each sample to its k-nearest neighbours with corresponding ideal kernel matrix rather than enforce the global alignment of all the samples, which is flexible and able to well handle the intra-cluster variations. In addition, for each sample, we construct a local Laplacian matrix to regularize that closer samples being allocated similar labels. Thus the locality of data samples can be well preserved.

If we use $K_\xi^{(i)}$ and $H^{(i)}$ to represent the sub-matrices of K_ξ and H, and their indices are specified by the τ-nearest neighbors of the i-th sample, then we have $K_\xi^{(i)} = {S^{(i)}}^T K_\xi S^{(i)}$ and $H^{(i)} = {S^{(i)}}^T H$, where $S^{(i)} \in \{0,1\}^{n\times\tau}$ is a matrix indicating the τ-nearest neighbors of the i-th sample. For each pair of sample, if they are close to each other, then their label vector should be also similar, this can be formulated as following:

$$\min_{H\in R^{n\times k}} \frac{1}{2}\sum_{i=1}^{n}\sum_{p,q=1}^{\tau} ||h_p^{(i)} - h_p^{(i)}||_2^2 [K_\xi^{(i)}]_{pq}, \qquad (11)$$

where $h_p^{(i)}$ and $h_q^{(i)}$ represent the p-th and q-th row of $H^{(i)}$, respectively. (11) can be easily rewritten as the following trace form:

$$\min_{H\in R^{n\times k}} \sum_{i=1}^{n} Tr({H^{(i)}}^T L_\tau^{(i)} H^{(i)}), \qquad (12)$$

where $L_\tau^{(i)}$ is the local Laplacian matrix of sample i with local similarity matrix $K_\xi^{(i)}$.

Then we rewritten (10) in a locally regularized form and combine it with (12) to induce our final LPKA model:

$$\begin{aligned}\min_{H\in R^{n\times k},\xi\in R_+^m} & \sum_{i=1}^{n} Tr(K_\xi^{(i)}(I_\tau - H^{(i)}{H^{(i)}}^T)) + \frac{\lambda}{2}\xi^T M^{(i)}\xi \\ & + \beta Tr({H^{(i)}}^T L_\tau^{(i)} H^{(i)}) \\ & s.t.\ H^T H = I_k,\ \xi^T 1_m = 1,\end{aligned} \qquad (13)$$

where I_τ and is an identity matrix with size $\tau\times\tau$. By defining $A^{(i)} = S^{(i)}{S^{(i)}}^T$, we obtain the objective function of our proposed LPKA:

$$\begin{aligned}\min_{H\in R^{n\times k},\xi\in R_+^m} & \sum_{i=1}^{n} Tr(K_\xi(A^{(i)} - A^{(i)}HH^T A^{(i)})) + \frac{\lambda}{2}\xi^T M^{(i)}\xi \\ & + \beta Tr(H^T S^{(i)} L_\tau^{(i)} {S^{(i)}}^T H) \\ & s.t.\ H^T H = I_k,\ \xi^T 1_m = 1\end{aligned} \qquad (14)$$

1) Optimization algorithm

Note that (14) is not jointly convex with respect to H and ξ, while it is convex to each variable if the other one is fixed. Thus, we design a two-step algorithm to solve this problem alternately.

Step 1: Updating H with fixed ξ

When ξ is fixed, H can be obtained by solving the following optimization problem:

$$\begin{aligned}\max_{H\in R^{n\times k}} & Tr(H^T \sum_{i=1}^{n}(A^{(i)}K_\xi A^{(i)})H) \\ & - \beta Tr(H^T \sum_{i=1}^{n} S^{(i)} L_\tau^{(i)} {S^{(i)}}^T H) \\ & s.t.\ H^T H = I_k,\ \xi^T 1_m = 1\end{aligned} \qquad (15)$$

If we set $\Omega = \sum_{i=1}^{n} A^{(i)}K_\xi A^{(i)}$ and $\Theta = \sum_{i=1}^{n} S^{(i)}L_\tau^{(i)}{S^{(i)}}^T$, then (15) can be rewritten as:

$$\max_{H\in R^{n\times k}} Tr(H^T(\Omega - \beta\Theta)H),\ \ s.t.\ H^T H = I_k, \qquad (16)$$

which is a standard kernel k-means clustering problem and can be efficiently solved.

Step 2: Updating ξ with fixed H

Given fixed H, the optimization (14) w.r.t ξ is a quadratic programming with linear constraints, which turns to the following problem:

$$\min_{\xi\in R_+^m} \sum_{i=1}^{n} \frac{1}{2}\xi^T(2\Delta + \lambda\Psi)\xi\ \ s.t.\ \ \xi^T 1_m = 1, \qquad (17)$$

where $\Delta = diag([Tr(K_1 V),\cdots,Tr(K_m V)])$, $V = \sum_{i=1}^{n}(A^{(i)} - A^{(i)}HH^T A^{(i)})$, and $\Psi_{pq} = \sum_{i=1}^{n} Tr(K_p A^{(i)} K_q A^{(i)})$.

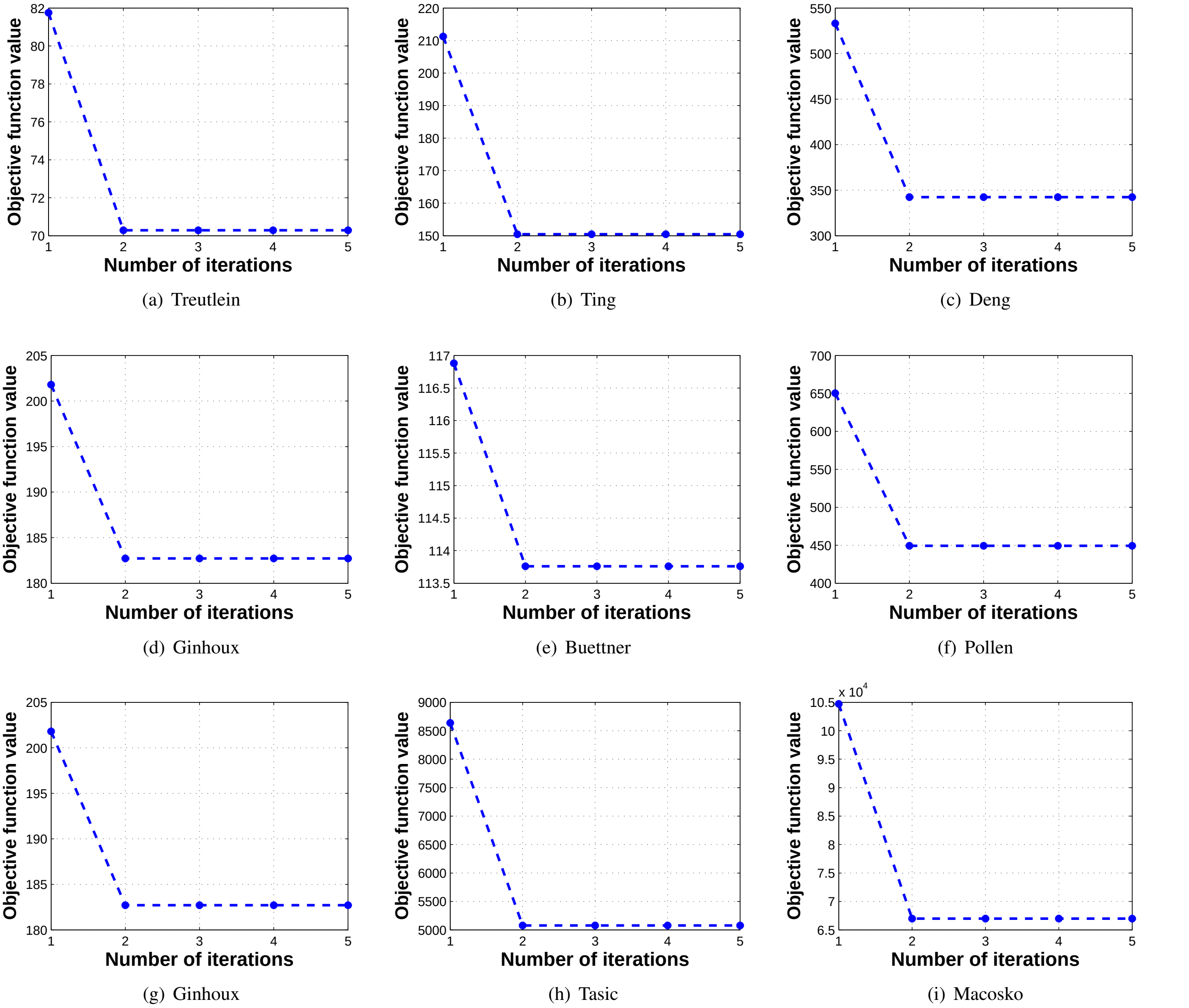
In summary, our algorithm for solving (14) can be outlined in Algorithm 1. obj^t denotes the objective value at the t-th iterations.

Algorithm 1: Optimization Algorithm for the proposed LPKA.

```
Input: Multiple kernel matrices $\{K_p\}_{p=1}^m$, $k$, $\lambda$, $\beta$ and a small positive constant $\varepsilon$.
Initialization: $\xi = 1_m/m$, and $t = 1$. Calculating $S^{(i)}$ for the $i$-th sample ($1 \le i \le n$) by $K_\xi$.
while not converged do
    1. Update $K_\xi$ by $K_\xi^t = \sum_{p=1}^m (\xi_p^t)^2 K_p$;
    2. Update $H^t$ by solving Eq. (15);
    3. Update $\xi^t$ by solving Eq. (17);
    4. t=t+1
    5. Check convergence condition:
    $(obj^{t-1} - obj^t)/obj^t \le \varepsilon$.
end while
Output: $H$ and $\xi$.
```

2) Convergence analysis

In Algorithm 1, the neighborhood of each sample is kept unchanged during the optimization. Specifically, the τ-nearest neighbors of sample i are measured by $K_\xi^{(i)}$. In such a manner, the objective value of Algorithm 1 is guaranteed to be monotonically decreased when updating one variable



(a) Treutlein

(b) Ting

(c) Deng

(d) Ginhoux

(e) Buettner

(f) Pollen

(g) Ginhoux

(h) Tasic

(i) Macosko

FIGURE 1: The objective value of our algorithm at each iteration on each dataset.

with the other fixed at each iteration. Meantime, the whole optimization problem is lower-bounded. As a sequence, the proposed algorithm can be guaranteed to be convergent. In order to empirically study the convergence of Algorithm 1, we show the variation of the objective values of Eq. (14) on different datasets in Figure 1, which demonstrate that our proposed optimization algorithm is very efficient, i.e., the objective value is monotonically decreased and the algorithm quickly converges in less than five iterations.

III. RESULTS AND DISCUSSIONS

A. EXPERIMENT RESULTS

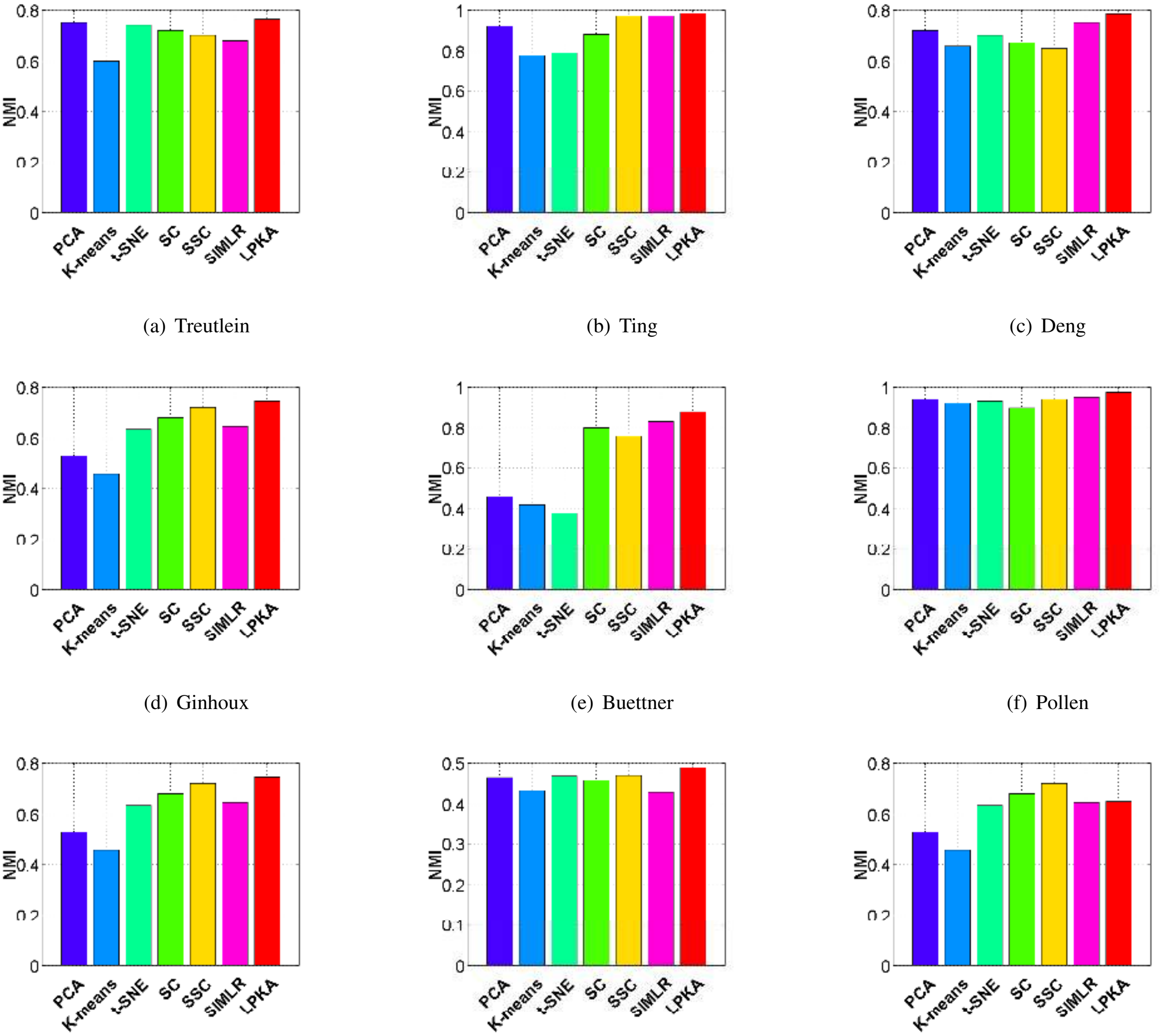
1) Experimental settings

To systematically evaluate the clustering performance of LPKA on the collected datasets, three metrics including Normalized Mutual Information (NMI) [49], Purity [50], and Adjusted Rand Index (ARI) [51] are used in our experiments. NMI and Purity take on values between 0 and 1, but ARI can be negative. The three metrics measure the concordance between the ground truth cell types and the cell types calculated by clustering algorithms, thus higher values indicate better performance for all metrics.

Given two clustering results U and V on a set of N data points with $\mathcal{N}_U$ and $\mathcal{N}_V$ clusters, respectively, the mutual information NMI is defined as

$$NMI(U,V)=\frac{\sum_{p=1}^{\mathcal{N}_U}\sum_{q=1}^{\mathcal{N}_V}|U_p\cap V_q|\log\frac{N|U_p\cap V_q|}{|U_p|\times|V_q|}}{\max\left(-\sum_{p=1}^{\mathcal{N}_U}U_p\log\frac{|U_p|}{N},-\sum_{q=1}^{\mathcal{N}_V}V_q\log\frac{|V_q|}{N}\right)} \tag{18}$$

where the numerator is the mutual information between U and V, and the denominator represents the entropy of the clustering U and V.



(a) Treutlein

(b) Ting

(c) Deng

(d) Ginhoux

(e) Buettner

(f) Pollen

(g) Ginhoux

(h) Tasic

(i) Macosko

FIGURE 2: NMI evaluation of different clustering methods on different datasets.

For Purity, each identified cluster is assigned to the one which is most frequent in the cluster, and then the accuracy of this assignment is computed by counting the number of correctly assigned samples divided by the number N:

$$Purity(U, V) = \frac{\sum_p \max_q |U_p \cap V_q|}{N}. \tag{19}$$

The ARI depends on the following four quantities:

- O_{uv}, the number of objects in a pair that are placed in the same group in U and V;
- O_u, the number of objects in a pair that are placed in the same group in U but in different groups in V;
- O_v, the number of objects in a pair that are placed in the same group in V but in different groups in U;
- O, the number of objects in a pair that are placed in the different group in U and V.

Then, ARI is defined as

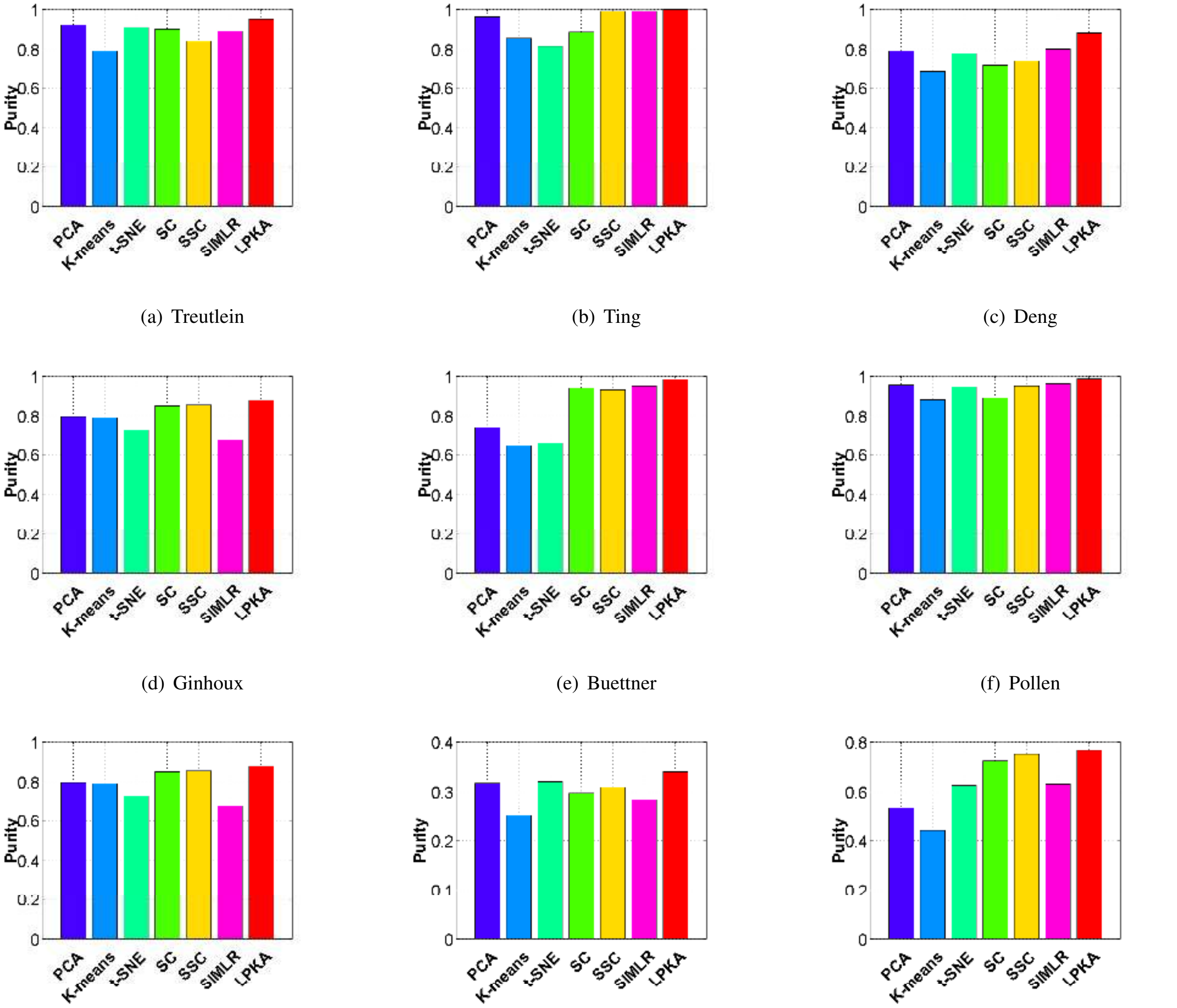
$$ARI(U, V) = \frac{\binom{n}{2}(A) - [(B)(C) + (O_v + O)(O_u + O)]}{\binom{n}{2} - [(B)(C) + (O_v + O)(O_u + O)]}, \tag{20}$$

where $A = O_{uv} + O$, $B = O_{uv} + O_u$ and $C = O_{uv} + O_v$.

As can be seen from Eq. (14), there are three parameters in LPKA including λ, β and τ need to be set. In our experiments, we tune all the parameters by a "grid-search" strategy. Specificaly, λ and β are chosen from $\{0.01, 0.1, 1, 10, 100\}$, and τ is chosen from $\{0.3n, 0.4n, 0.5n\}$. Finally, the best clustering results are used for comparison.

2) Compared with other methods

We compare the proposed LPKA with several existing methods, including PCA, traditional k-means, t-SNE [52], spectral clustering ("SC"), sparse spectral clustering ("SSC") [53] and



(a) Treutlein

(b) Ting

(c) Deng

(d) Ginhoux

(e) Buettner

(f) Pollen

(g) Ginhoux

(h) Tasic

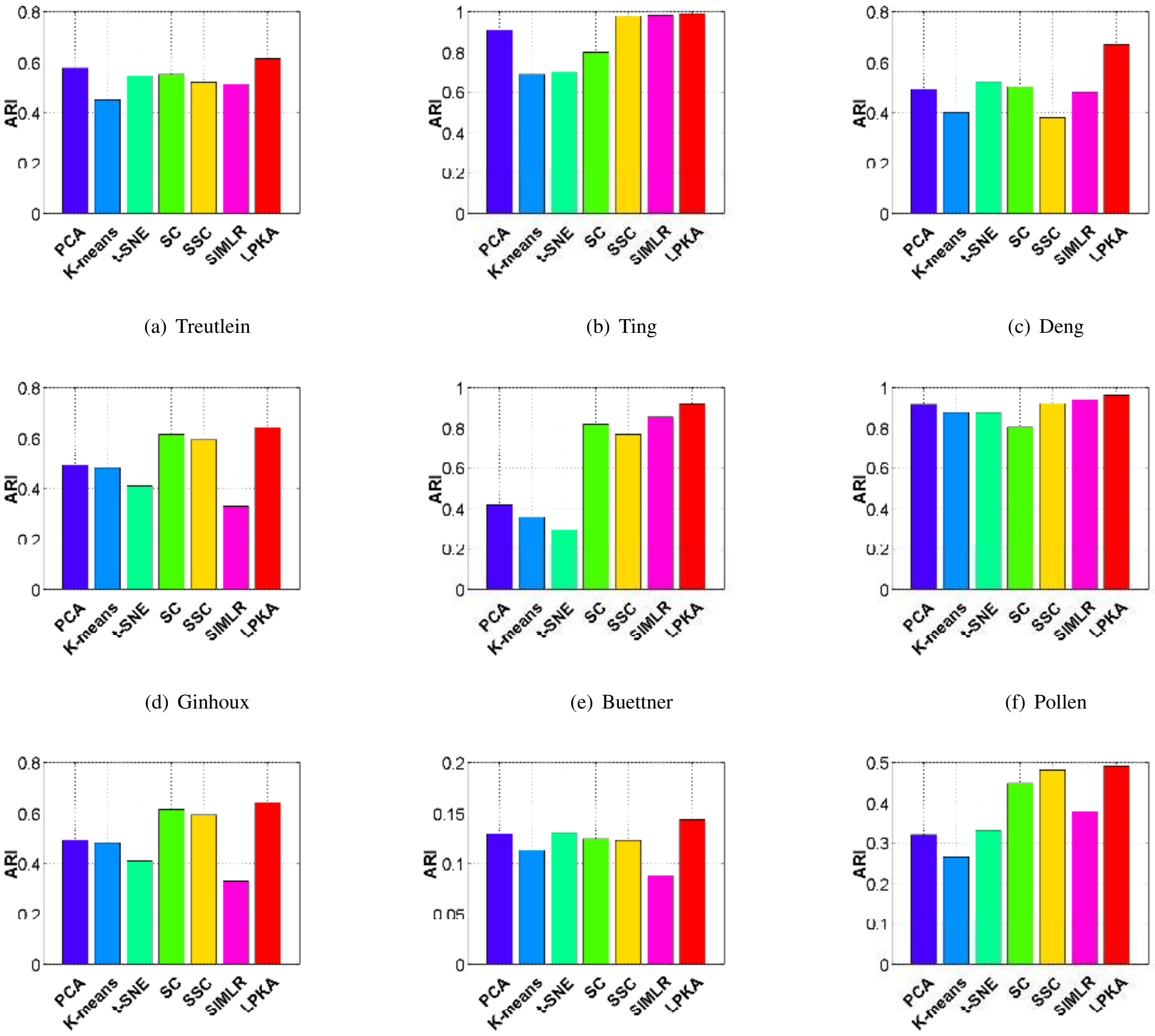
(i) Macosko

FIGURE 3: Purity evaluation of different clustering methods on different datasets.

SIMLR [40]. For all the compared methods, their parameters are turned carefully as suggested in their corresponding papers for fair comparison. We show the NMI, Purity and ARI of different clustering algorithms on different datasets in Figure 2, Figure 3 and Figure 4, respectively. As can be seen, the proposed LPKA has higher performance than all of other methods on the 9 datasets in terms of three different metrics, which demonstrates the superiority of LPKA. Therefore, our proposed LPKA can be used as a reliable pre-processing step to distinguish different cell types, which can reveal new insights into many other biological problems.

3) Parameter sensitivity

In our experiment, we turn the parameters λ, β and τ to obtain the optimal results. To study the sensitivity of the LPKA with regard to the parameters in Eq. (14), we conduct experiments by fixing one of the three parameters and varying the other two ones. Firstly, we fix $\beta = 1$, and vary τ and λ. Figure 5-7 plot the values of the NMI, Purity and ARI, respectively, with fixed β. Secondly, we fix $\lambda = 1$, and vary τ and β. Figure 8-10 plot the values of the NMI, Purity and ARI, respectively, with fixed λ. Thirdly, we fix $\tau = 0.4n$, and vary λ and β. Figure 11-13 plot the values of the NMI, Purity and ARI, respectively, with fixed τ. As can be seen, the clustering results are robust with respect to the varying of λ, while the results are a little sensitive to β and τ to some extend, which demonstrate the significance of preserving the local structure of original data. We can obtain optimal clustering results with different combinations of the three parameters.



(a) Treutlein

(b) Ting

(c) Deng

(d) Ginhoux

(e) Buettner

(f) Pollen

(g) Ginhoux

(h) Tasic

(i) Macosko

FIGURE 4: ARI evaluation of different clustering methods on different datasets.

IV. CONCLUSIONS

In this work, we present a novel multiple kernel clustering framework for scRNA-seq data clustering via locality preserving kernel alignment. A series of similarity kernel matrices are firstly generated by using different kernel functions. Then we transfer the clustering task to a multiple kernel k-means problem with the kernels aligned in a local manner. In order to preserve the local structure of the data for boosting final clustering performance, we construct a local Laplacian matrix for each sample to constrain that closer samples should be allocated similar labels. An alternate updating algorithm with theoretical analysis is developed to solve the proposed problem. Experiments with parameter sensitivity analysis on various real scRNA-seq data are conducted to demonstrate that our method can obtain superior results when compared with other state-of-the-art approaches.

ACKNOWLEDGMENTS

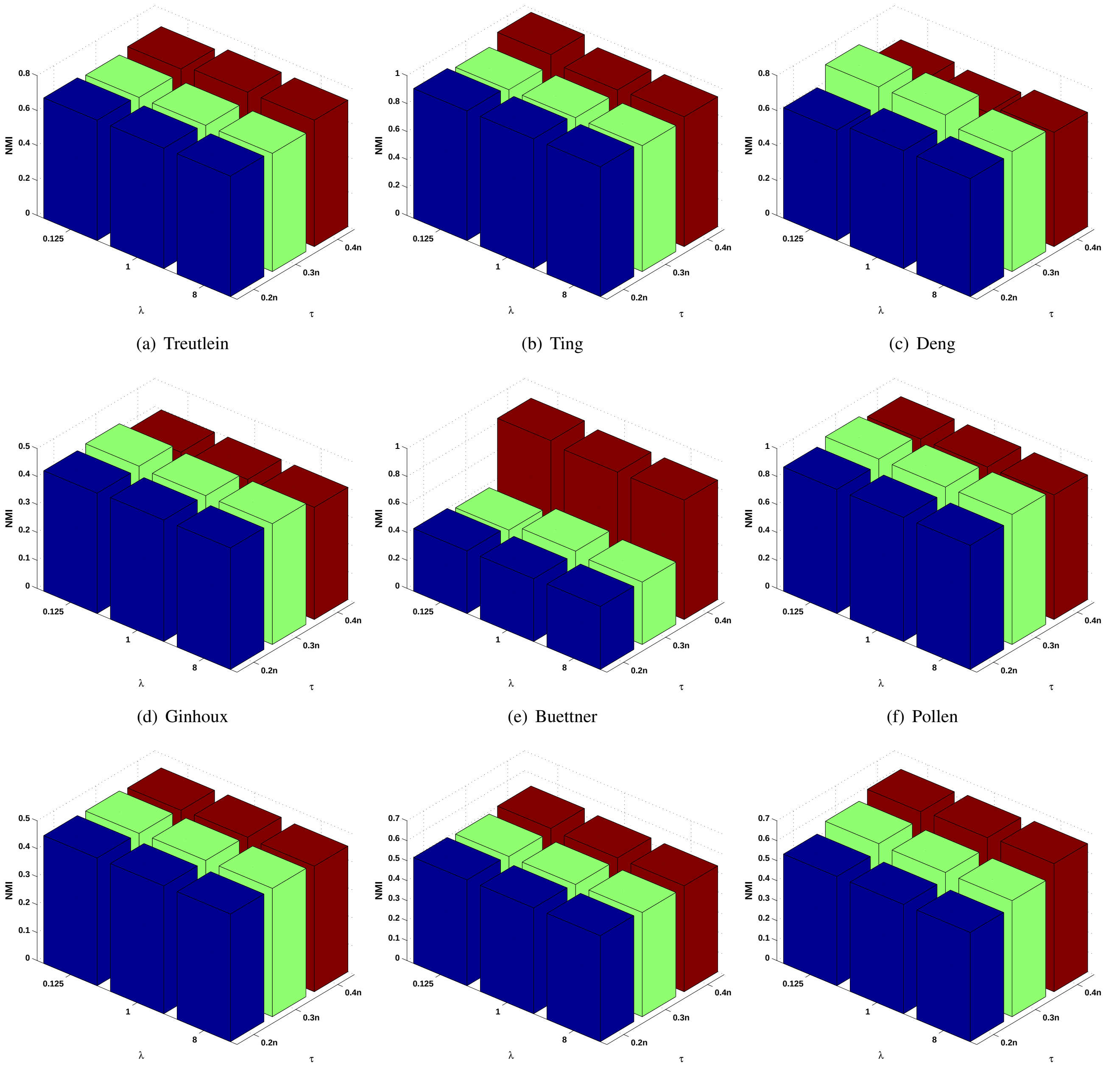
This work was supported in part by the National Natural Science Foundation of China under Grant No. 61701451, in part by the Fundamental Research Funds for the Central Universities, China University of Geosciences (Wuhan) under Grant No.CUG170654.

V. AVAILABILITY OF DATA AND MATERIALS

The implementation code and supporting files are available from http://tangchang.net/codes/LPKACode.zip

REFERENCES

[1] J. M. Raser and E. K. O'Shea, "Control of stochasticity in eukaryotic gene expression," Science, vol. 304, no. 5678, pp. 1811–1814, 2004.



(a) Treutlein

(b) Ting

(c) Deng

(d) Ginhoux

(e) Buettner

(f) Pollen

(g) Tasic

(h) Zeisel

(i) Macosko

FIGURE 5: NMI of our method w.r.t τ and λ on different datasets ($\beta = 1$).

[2] T. Kalisky and S. R. Quake, "Single-cell genomics." Nature Methods, vol. 8, no. 4, pp. 311–314, 2011.

[3] L. Pelkmans, "Cell biology. using cell-to-cell variability–a new era in molecular biology." Science, vol. 336, no. 6080, p. 425, 2012.

[4] Y. Xiao, J. Wu, Z. Lin, and X. Zhao, "A semi-supervised deep learning method based on stacked sparse auto-encoder for cancer prediction using rna-seq data," Computer Methods and Programs in Biomedicine, vol. 166, pp. 99 – 105, 2018.

[5] D. Goksuluk, G. Zararsiz, S. Korkmaz, V. Eldem, G. E. Zararsiz, E. Ozcetin, A. Ozturk, and A. E. Karaagaoglu, "Mlseq: Machine learning interface for rna-sequencing data," Computer Methods and Programs in Biomedicine, vol. 175, pp. 223 – 231, 2019.

[6] Z. Xue, K. Huang, C. Cai, L. Cai, C. Y. Jiang, Y. Feng, Z. Liu, Q. Zeng, L. Cheng, and Y. E. Sun, "Genetic programs in human and mouse early embryos revealed by single-cell rna sequencing." Nature, vol. 500, no. 7464, pp. 593–597, 2013.

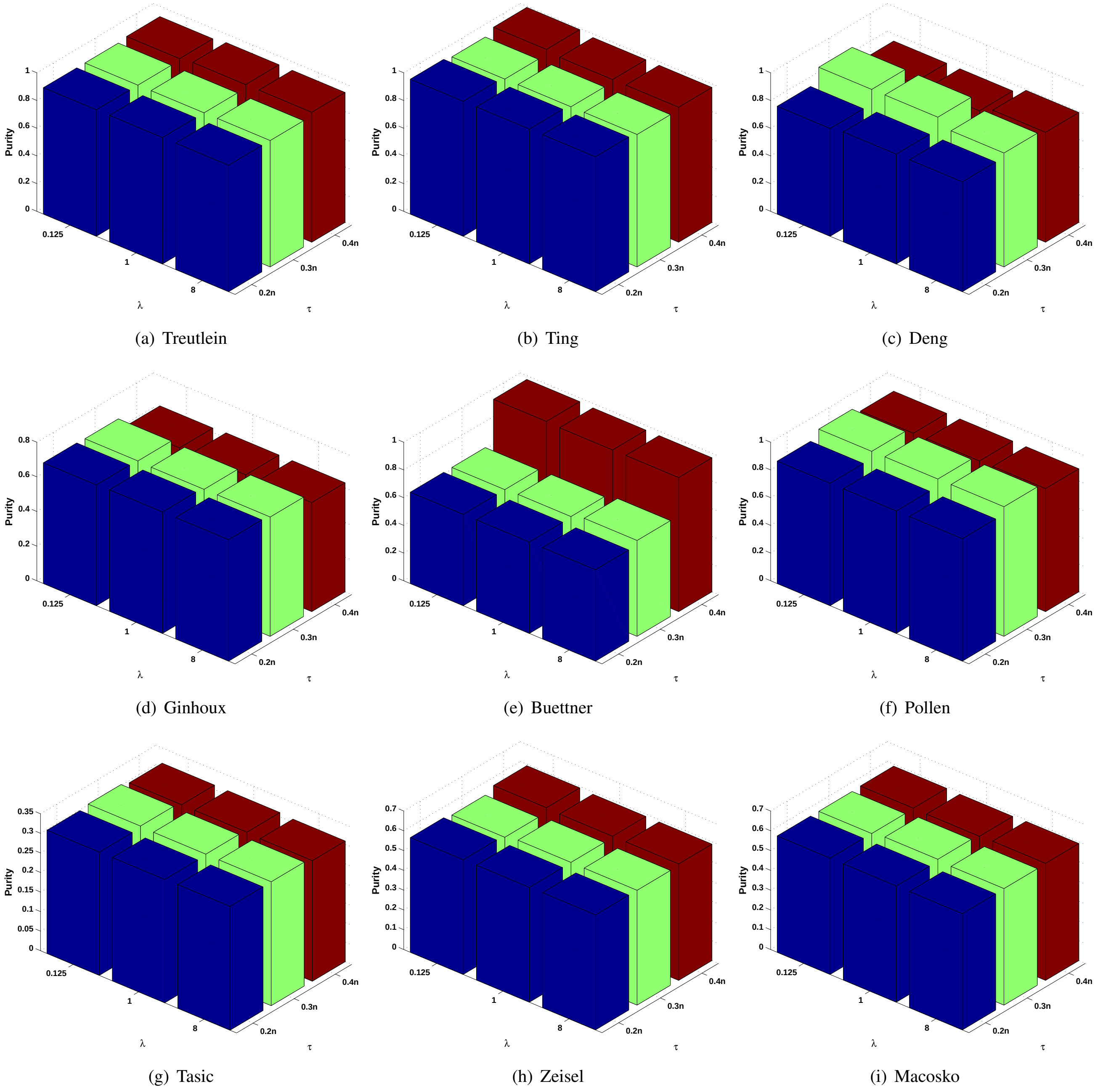
[7] L. Yan, M. Yang, H. Guo, L. Yang, J. Wu, R. Li, P. Liu, Y. Lian, X. Zheng, and J. Yan, "Single-cell rna-seq profiling of human preimplantation embryos and embryonic stem cells," Nature Structural & Molecular Biology, vol. 20, no. 9, pp. 1131–1139, 2013.

[8] A. E. Saliba, A. J. Westermann, S. A. Gorski, and J. Vogel, "Single-cell rna-seq: advances and future challenges," Nucleic Acids Research, vol. 42, no. 14, pp. 8845–60, 2014.

[9] R. Kabir and R. Islam, "Chemical reaction optimization for rna structure prediction," Applied Intelligence, vol. 49, no. 2, pp. 352–375, 2019.

[10] Y. Sun, L. Ouyang, and D.-Q. Dai, "Lrsk: A low-rank self-representation k-means method for clustering single-cell rna-sequencing data," Molecular Omics, 2020.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3036132, IEEE Access



(a) Treutlein (b) Ting (c) Deng

(d) Ginhoux (e) Buettner (f) Pollen

(g) Tasic (h) Zeisel (i) Macosko

FIGURE 6: Purity of our method w.r.t τ and λ on different datasets ($\beta = 1$).

[11] J. Hua, H. Liu, B. Zhang, and S. Jin, “Lak: Lasso and k-means based single-cell rna-seq data clustering analysis,” IEEE Access, vol. 8, pp. 129 679–129 688, 2020.

[12] Y. Wu, Y. Guo, Y. Xiao, and S. Lao, “Aae-sc: A scrna-seq clustering framework based on adversarial autoencoder,” IEEE Access, 2020.

[13] P. Zhou, F. Ye, and L. Du, “Unsupervised robust multiple kernel learning via extracting local and global noises,” IEEE Access, vol. 7, pp. 34 451–34 461, 2019.

[14] C. Tang, X. Zhu, X. Liu, M. Li, P. Wang, C. Zhang, and L. Wang, “Learning a joint affinity graph for multiview subspace clustering,” IEEE Transactions on Multimedia, vol. 21, no. 7, pp. 1724–1736, 2018.

[15] Z. Li, C. Tang, J. Chen, C. Wan, W. Yan, and X. Liu, “Diversity and consistency learning guided spectral embedding for multi-view clustering,” Neurocomputing, vol. 370, pp. 128–139, 2019.

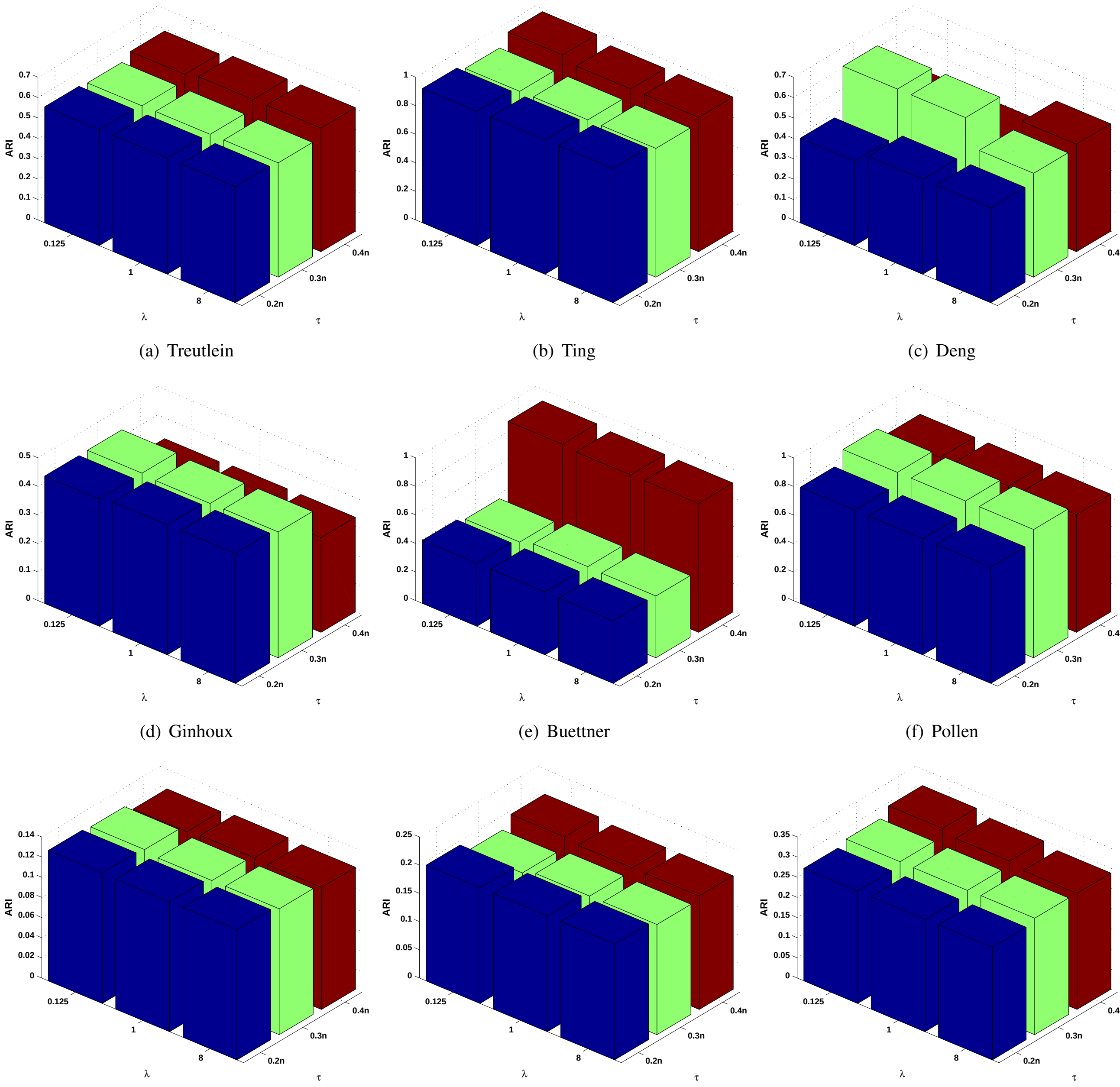
[16] C. Tang, X. Liu, X. Zhu, E. Zhu, Z. Luo, L. Wang, and W. Gao, “Cgd: Multi-view clustering via cross-view graph diffusion.” in AAAI Conference on Artificial Intelligence, 2020, pp. 5924–5931.

[17] R. Vidal, Y. Ma, and S. Sastry, “Generalized principal component analysis (gpca).” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 12, pp. 1945–1959, 2005.

[18] Luxburg and Ulrike, “A tutorial on spectral clustering,” Statistics & Computing, vol. 17, no. 4, pp. 395–416, 2007.

[19] E. W. Forgy, “Cluster analysis of multivariate data : efficiency versus interpretability of classifications,” Biometrics, vol. 21, no. 3, pp. 41–52, 1965.

[20] P. Brennecke, S. Anders, J. K. Kim, A. A. KoÅĆodziejczyk, X. Zhang, V. Proserpio, B. Baying, V. Benes, S. A. Teichmann, and J. C. Marioni, “Accounting for technical noise in single-cell rna-seq experiments.” Nature



(a) Treutlein (b) Ting (c) Deng

(d) Ginhoux (e) Buettner (f) Pollen

(g) Tasic (h) Zeisel (i) Macosko

FIGURE 7: ARI of our method w.r.t τ and λ on different datasets ($\beta = 1$).

Methods, vol. 10, no. 11, pp. 1093–1095, 2013.

[21] D. Grun, L. Kester, and O. A. Van, “Validation of noise models for single-cell transcriptomics,” Nature Methods, vol. 11, no. 6, pp. 637–640, 2014.

[22] R. Bacher and C. Kendziorski, “Design and computational analysis of single-cell rna-sequencing experiments,” Genome Biology, vol. 17, no. 1, p. 63, 2016.

[23] Y. Buganim, D. A. Faddah, A. W. Cheng, E. Itskovich, S. Markoulaki, K. Ganz, S. L. Klemm, O. A. Van, and R. Jaenisch, “Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase,” Cell, vol. 150, no. 6, pp. 1209–1222, 2012.

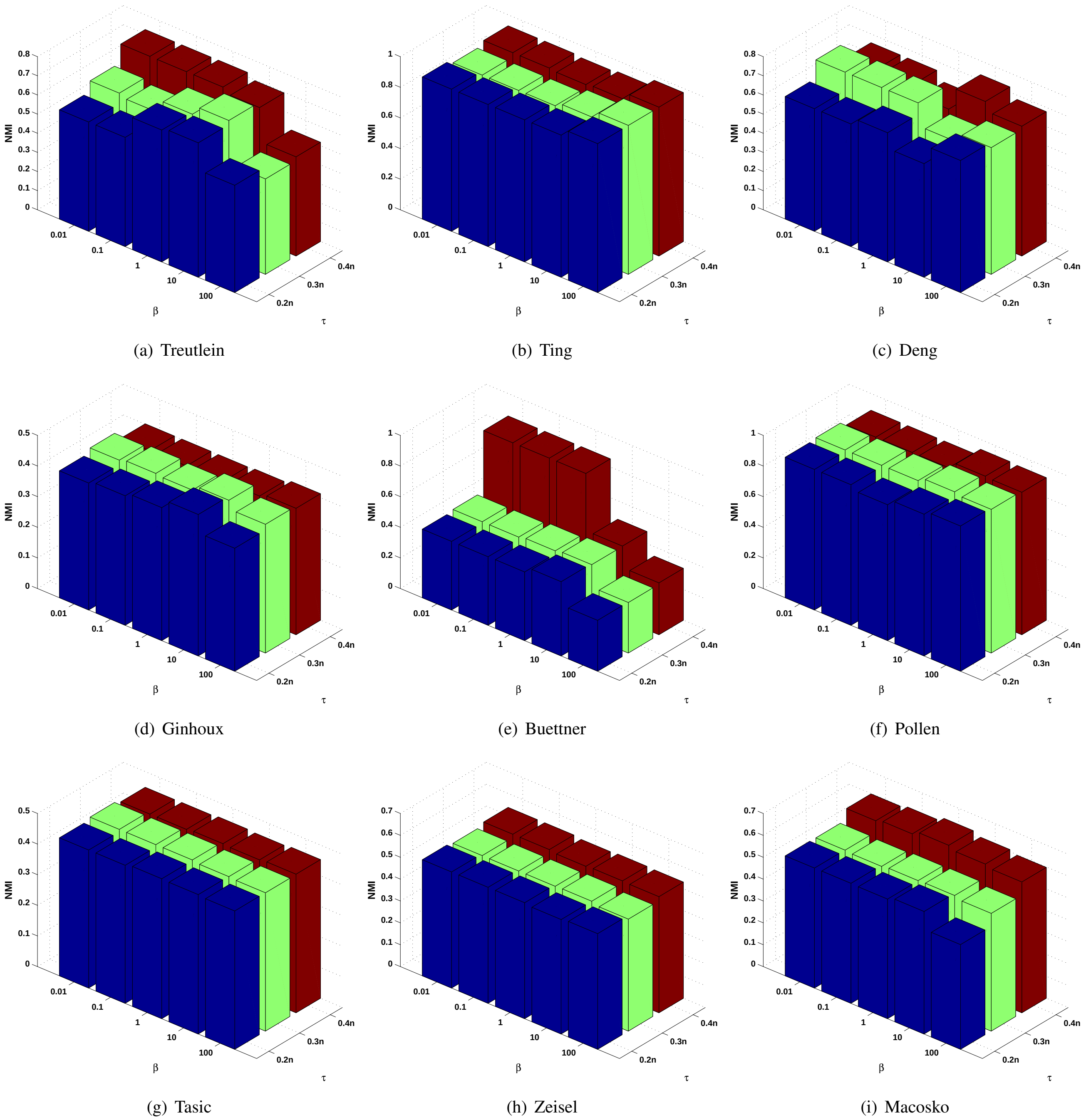
[24] T. Hashimshony, F. Wagner, N. Sher, and I. Yanai, “Cel-seq: Single-cell rna-seq by multiplexed linear amplification,” Cell Reports, vol. 2, no. 3, pp. 666–673, 2012.

[25] V. Y. Kiselev, T. S. Andrews, and M. Hemberg, “Challenges in unsupervised clustering of single-cell rna-seq data,” Nature Reviews Genetics, p. 1, 2019.

[26] H. Peng, X. Zeng, Y. Zhou, D. Zhang, R. Nussinov, and F. Cheng, “A component overlapping attribute clustering (coac) algorithm for single-cell rna sequencing data analysis and potential pathobiological implications,” PLoS computational biology, vol. 15, no. 2, p. e1006772, 2019.

[27] L. Sun, X.-Y. Zhang, Y.-H. Qian, J.-C. Xu, S.-G. Zhang, and Y. Tian, “Joint neighborhood entropy-based gene selection method with fisher score for tumor classification,” Applied Intelligence, vol. 49, no. 4, pp. 1245–1259, 2019.

[28] A. Zeisel, A. B. MuÃśozmanchado, S. Codeluppi, P. LÃűnnerberg, M. G. La, A. JurÃĺus, S. Marques, H. Munguba, L. He, and C. Betsholtz, “Brain structure. cell types in the mouse cortex and hippocampus revealed by single-cell rna-seq.” Science, vol. 347, no. 6226, pp. 1138–42, 2015.



(a) Treutlein

(b) Ting

(c) Deng

(d) Ginhoux

(e) Buettner

(f) Pollen

(g) Tasic

(h) Zeisel

(i) Macosko

FIGURE 8: NMI of our method w.r.t τ and β on different datasets ($\lambda = 1$).

[29] E. Z. Macosko, A. Basu, R. Satija, J. Nemesh, K. Shekhar, M. Goldman, I. Tirosh, A. R. Bialas, N. Kamitaki, and E. M. Martersteck, "Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets." Cell, vol. 161, no. 5, pp. 1202–14, 2015.

[30] B. Tasic, V. Menon, T. N. Nguyen, T. K. Kim, T. Jarsky, Z. Yao, B. Levi, L. T. Gray, S. A. Sorensen, and T. Dolbeare, "Adult mouse cortical cell taxonomy by single cell transcriptomics," Nature Neuroscience, vol. 19, no. 2, pp. 335–346, 2016.

[31] M. Mojarad, S. Nejatian, H. Parvin, and M. Mohammadpoor, "A fuzzy clustering ensemble based on cluster clustering and iterative fusion of base clusters," Applied Intelligence, vol. 49, no. 7, pp. 2567–2581, 2019.

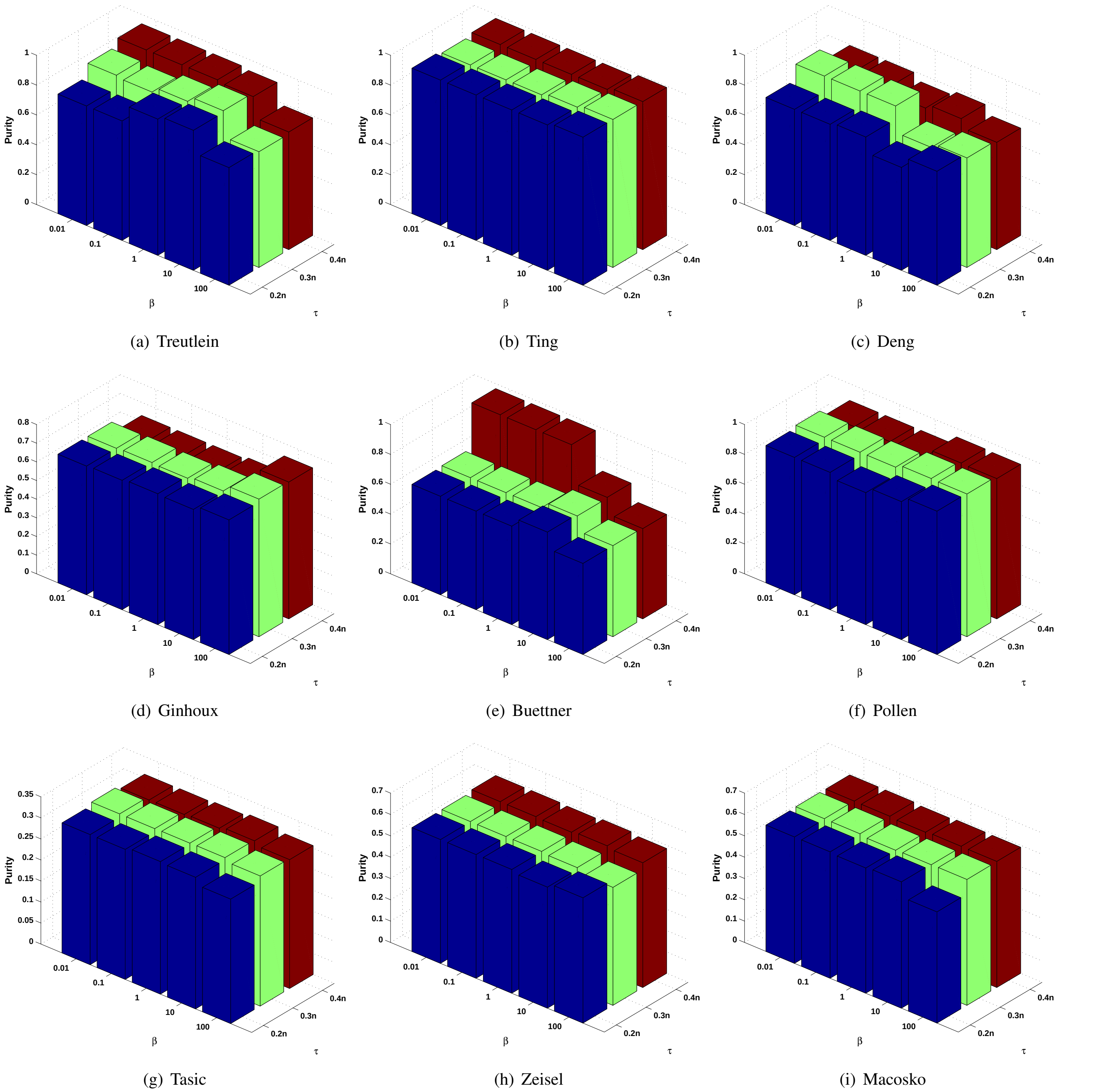
[32] L. Haghverdi, F. Buettner, and F. J. Theis, "Diffusion maps for high-dimensional single-cell analysis of differentiation data," Bioinformatics, vol. 31, no. 18, pp. 2989–2998, 2015.

[33] C. Shao and T. Höfer, "Robust classification of single-cell transcriptome data by nonnegative matrix factorization," Bioinformatics, vol. 33, no. 2, p. 235, 2016.

[34] C. Xu and Z. Su, "Identification of cell types from single-cell transcriptomes using a novel clustering method," Bioinformatics, vol. 31, no. 12, pp. 1974–80, 2015.

[35] S. Park and H. Zhao, "Spectral clustering based on learning similarity matrix." Bioinformatics, 2018.

[36] W. Wu and X. Ma, "Joint learning dimension reduction and clustering of single-cell rna-sequencing data," Bioinformatics, 2020.



(a) Treutlein

(b) Ting

(c) Deng

(d) Ginhoux

(e) Buettner

(f) Pollen

(g) Tasic

(h) Zeisel

(i) Macosko

FIGURE 9: Purity of our method w.r.t τ and β on different datasets ($\lambda = 1$).

[37] A. A. Margolin, "Localized data fusion for kernel k -means clustering with application to cancer biology," in Neural Information Processing Systems, 2014, pp. 1305–1313.

[38] L. Du, P. Zhou, L. Shi, H. Wang, M. Fan, W. Wang, and Y. D. Shen, "Robust multiple kernel k-means using âĎŞ 2;1 -norm," in International Conference on Artificial Intelligence, 2015, pp. 3476–3482.

[39] Y. Lu, L. Wang, J. Lu, J. Yang, and C. Shen, "Multiple kernel clustering based on centered kernel alignment," Pattern Recognition, vol. 47, no. 11, pp. 3656–3664, 2014.

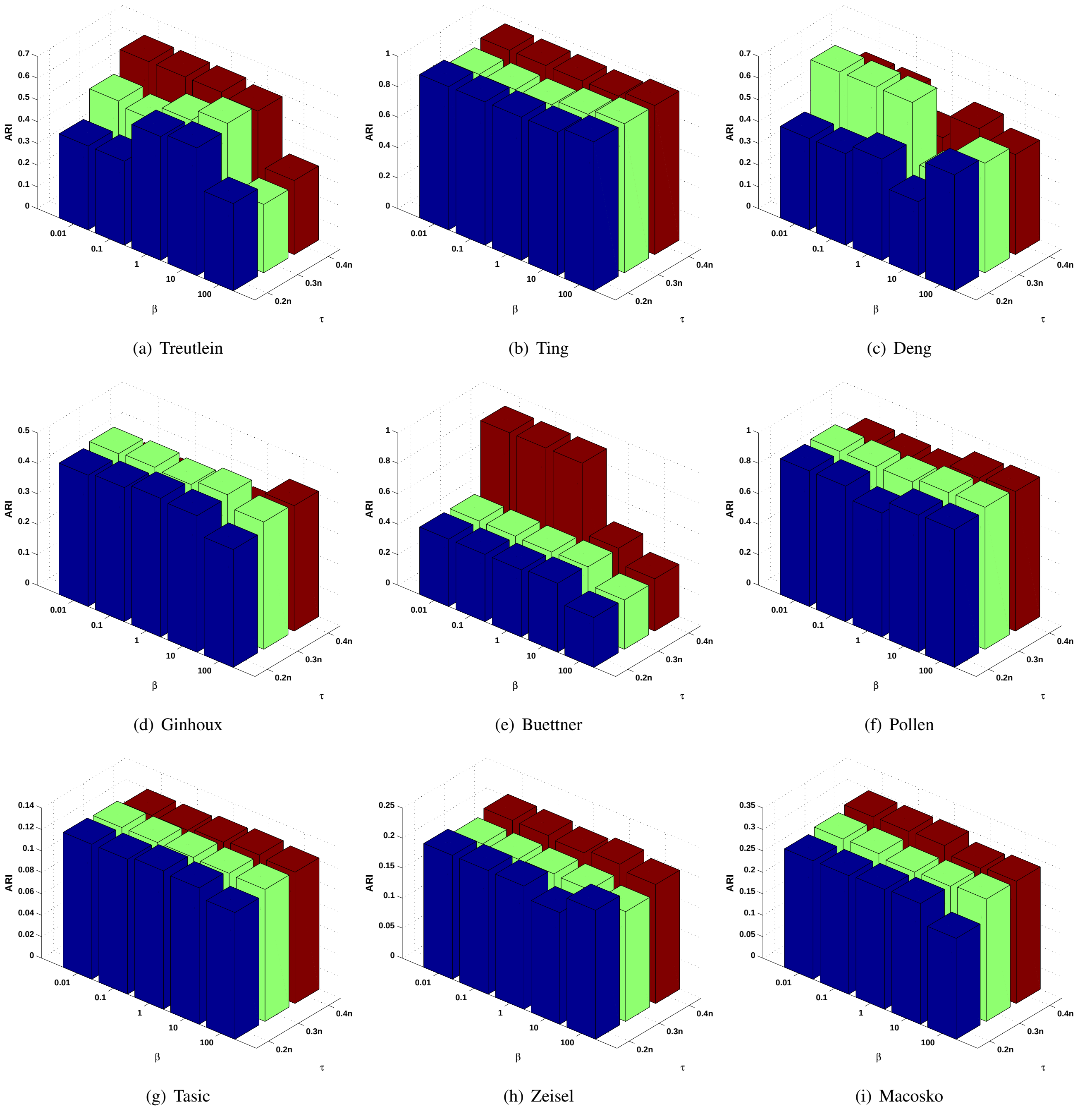
[40] B. Wang, J. Zhu, E. Pierson, D. Ramazzotti, and S. Batzoglou, "Visualization and analysis of single-cell rna-seq data by kernel-based similarity learning," Nature Methods, vol. 14, no. 4, p. 414, 2017.

[41] R. Qi, J. Wu, F. Guo, L. Xu, and Q. Zou, "A spectral clustering with self-weighted multiple kernel learning method for single-cell rna-seq data," Briefings in Bioinformatics, 2020.

[42] B. Treutlein, D. G. Brownfield, A. R. Wu, N. F. Neff, G. L. Mantalas, F. H. Espinoza, T. J. Desai, M. A. Krasnow, and S. R. Quake, "Reconstructing lineage hierarchies of the distal lung epithelium using single-cell rna-seq." Nature, vol. 509, no. 7500, pp. 371–375, 2014.

[43] D. Q, R. D, R. B, and S. R, "Single-cell rna-seq reveals dynamic, random monoallelic gene expression in mammalian cells," Science, vol. 343, no. 6167, pp. 193–196, 2014.

[44] F. Buettner, K. N. Natarajan, F. P. Casale, V. Proserpio, A. Scialdone, F. J. Theis, S. A. Teichmann, J. C. Marioni, and O. Stegle, "Computational



(a) Treutlein (b) Ting (c) Deng

(d) Ginhoux (e) Buettner (f) Pollen

(g) Tasic (h) Zeisel (i) Macosko

FIGURE 10: ARI of our method w.r.t τ and β on different datasets ($\lambda = 1$).

analysis of cell-to-cell heterogeneity in single-cell rna-sequencing data reveals hidden subpopulations of cells," Nature Biotechnology, vol. 33, no. 2, pp. 155–60, 2015.

[45] T. Bosiljka, M. Vilas, N. T. Nghi, K. T. Kyung, J. Tim, Z. Yao, L. Boaz, L. T. Gray, S. A. Sorensen, and D. Tim, "Adult mouse cortical cell taxonomy by single cell transcriptomics:," Nature Neuroscience, vol. 19, no. 2, pp. 335–346, 2016.

[46] B. Zhao, J. T. Kwok, and C. Zhang, "Multiple kernel clustering," SDM, pp. 638–649, 2009.

[47] C. Cortes, M. Mohri, and A. Rostamizadeh, "Algorithms for learning kernels based on centered alignment," Journal of Machine Learning Research, vol. 13, no. 2, pp. 795–828, 2012.

[48] X. Liu, Y. Dou, J. Yin, L. Wang, and E. Zhu, "Multiple kernel k -means clustering with matrix-induced regularization," in AAAI Conference on Artificial Intelligence, 2016, pp. 1888–1894.

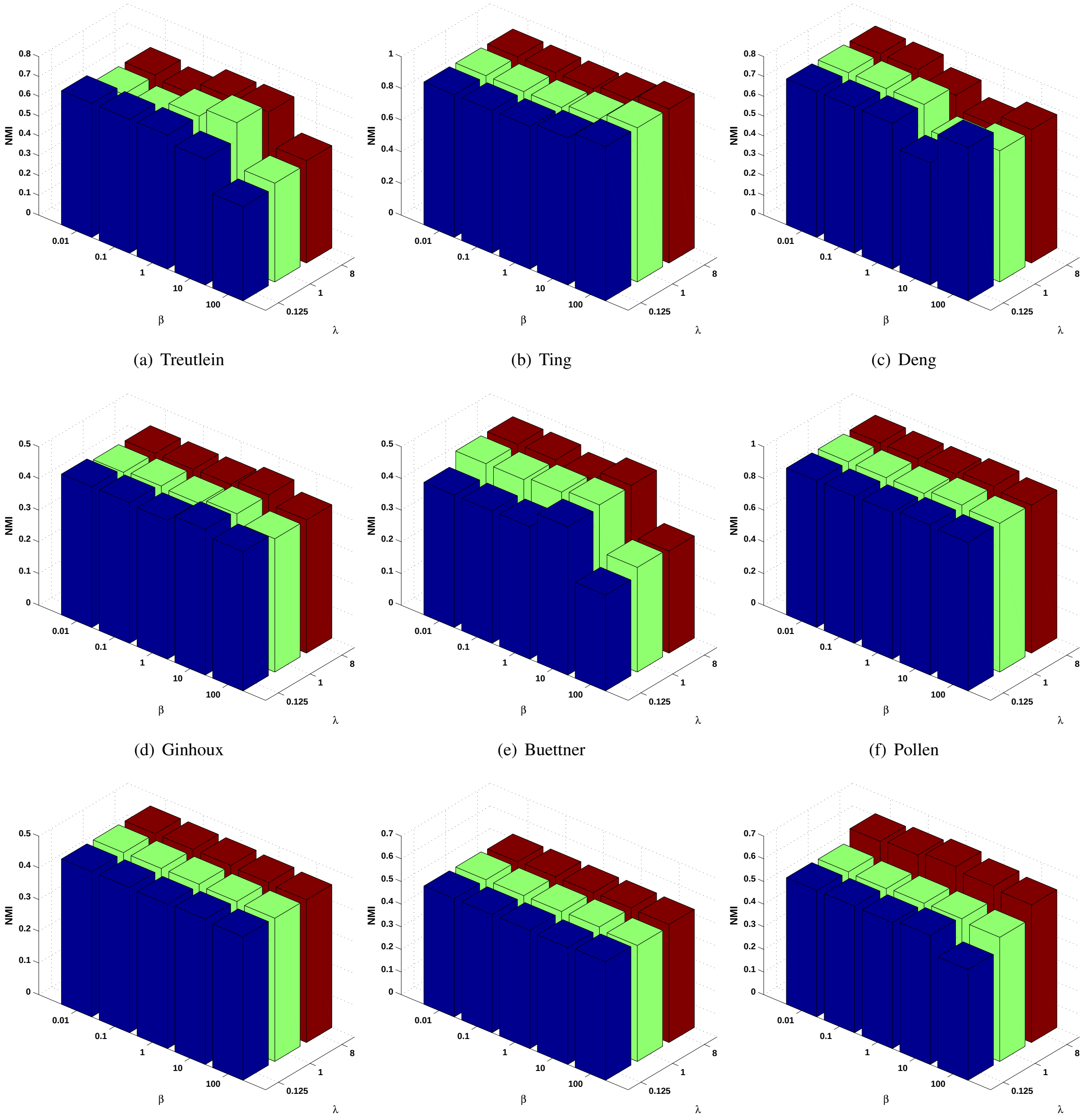
[49] A. Strehl and J. Ghosh, "Cluster ensembles — a knowledge reuse framework for combining multiple partitions," Journal of Machine Learning Research, vol. 3, pp. 583–617, 2003.

[50] S. Wagner and D. Wagner, Comparing clusterings: an overview. Universität Karlsruhe, Fakultät für Informatik Karlsruhe, 2007.

[51] ——, "Comparing clusterings - an overview," Analysis, vol. 4769, pp. 1–19, 2007.

[52] L. van der Maaten and G. Hinton, "Visualizing data using t-sne," Journal of Machine Learning Research, vol. 9, no. 2605, pp. 2579–2605, 2008.

[53] C. Lu, S. Yan, and Z. Lin, "Convex sparse spectral clustering: Single-view to multi-view," IEEE Transactions on Image Processing, vol. 25, no. 6, pp.



(a) Treutlein

(b) Ting

(c) Deng

(d) Ginhoux

(e) Buettner

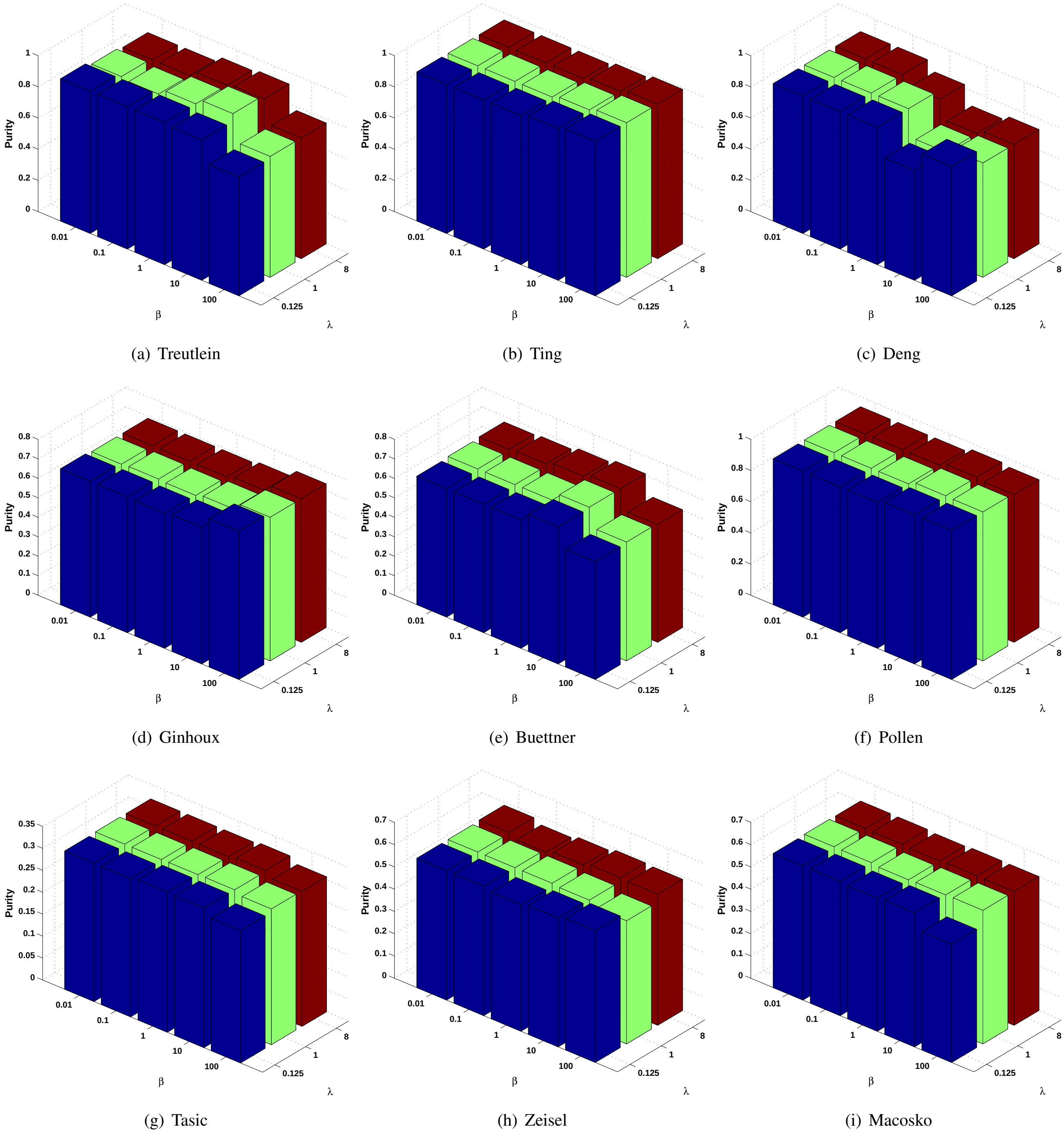
(f) Pollen

(g) Tasic

(h) Zeisel

(i) Macosko

FIGURE 11: NMI of our method w.r.t λ and β on different datasets ($\tau = 0.3$).



(a) Treutlein

(b) Ting

(c) Deng

(d) Ginhoux

(e) Buettner

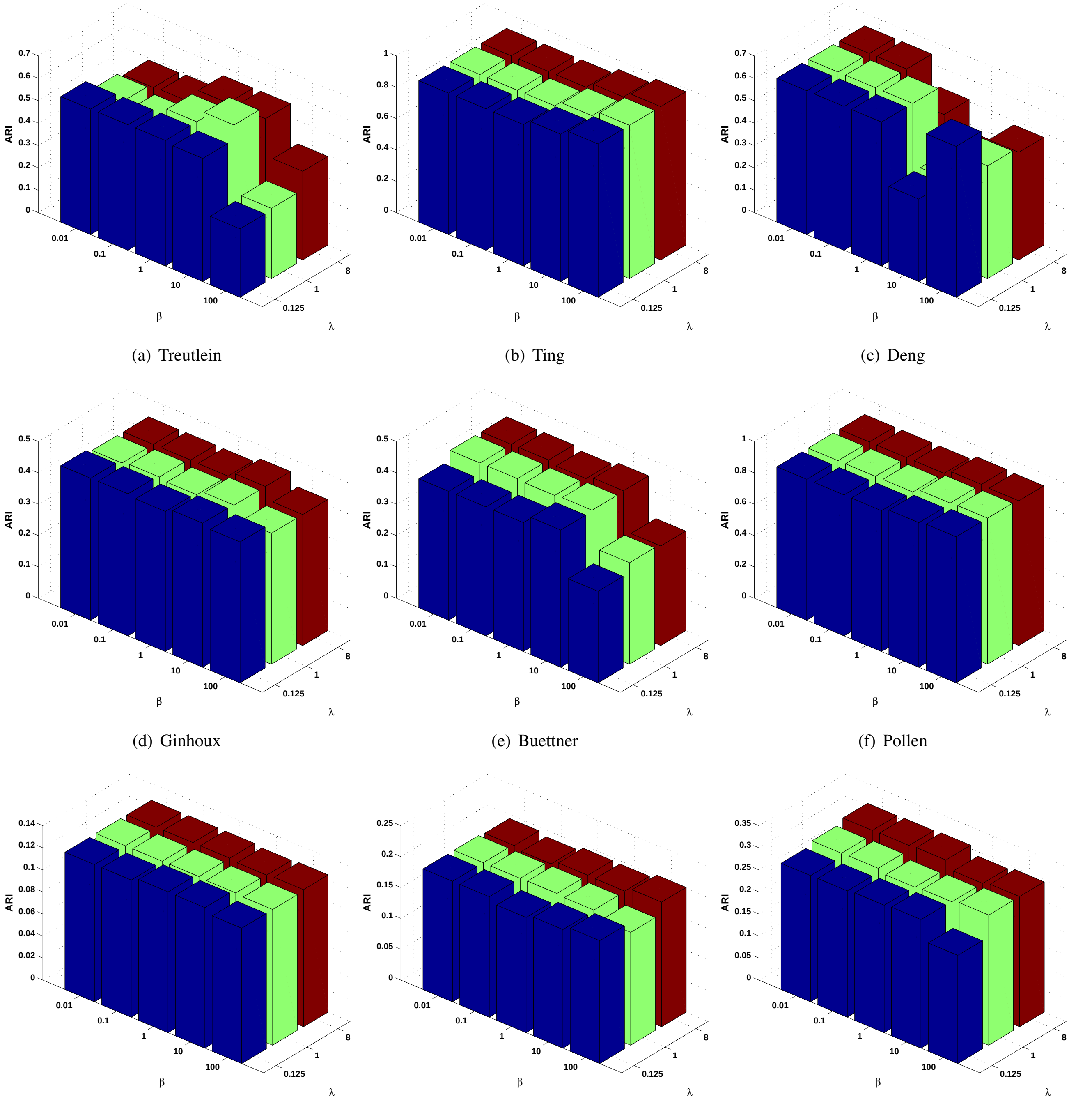
(f) Pollen

(g) Tasic

(h) Zeisel

(i) Macosko

FIGURE 12: Purity of our method w.r.t λ and β on different datasets ($\tau = 0.3$).



(a) Treutlein (b) Ting (c) Deng

(d) Ginhoux (e) Buettner (f) Pollen

(g) Tasic (h) Zeisel (i) Macosko

FIGURE 13: ARI of our method w.r.t λ and β on different datasets ($\tau = 0.3$).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3036132, IEEE Access

2833–2843, 2016.